

Міністерство освіти і науки України

---

---

# Системні технології

System technologies

---

3 (164) 2026

---

*Регіональний міжвузівський збірник наукових праць*

*Засновано у січні 1997 року.*

**У випуску:**

- ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ  
ТА ОРГАНІЗАЦІЯ СУЧАСНОГО ВИРОБНИЦТВА
- МАТЕМАТИЧНЕ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ  
ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ
- СИСТЕМНІ ТЕХНОЛОГІЇ ОБРОБКИ ІНФОРМАЦІЇ  
ТА КІБЕРБЕЗПЕКА

Системні технології. Регіональний міжвузівський збірник наукових праць. –  
Випуск 3 (164). - Дніпро, 2026. – 362 с.  
ISSN 1562-9945 (Print).  
ISSN 2707-7977 (Online).

### Редакційна колегія випуску:

**Редчиць Д.О.** - д.т.н., проф. (відп. редактор)  
**Svyetlichnyy D.**(Poland) - Ph.D, Professor, Full Professor  
**Островська К.Ю.** – к.т.н., доц.

Прогресивні інформаційні  
технології та організація  
сучасного виробництва

**Гвоздева І.М.** - д.т.н., проф.  
**Гнатушенко Вік.В.** - д.т.н., проф.  
**Шинкаренко В.І.** - д.т.н., проф. (відп. редактор)  
**Shults Roman** (Kingdom of Saudi Arabia) - DSc, PhD,  
Associate Professor

Математичне  
та програмне забезпечення  
інтелектуальних систем

**Бердник М.Г.** - д.т.н., проф.  
**Гнатушенко В.В.** - д.т.н., проф., (відп. редактор)  
**Heipke Christian** (Germany) - Dr.-Ing., Professor

Системні технології  
обробки інформації  
та кібербезпека

Збірник друкується за рішенням Вченої Ради  
Українського державного університету науки і технологій  
від 28.01.2026 р., № 6

Адреса редакції: 49600, Дніпро, пр. Науки, 4  
Український державний університет науки і технологій,  
ННІ «Дніпровський металургійний інститут»  
кафедра Інформаційних технологій та систем.  
Тел. +38(097)6854525  
E-mail: [st@nmetau.edu.ua](mailto:st@nmetau.edu.ua)  
<https://journals.nmetau.edu.ua/index.php/st>

© Український державний університет науки і технологій,  
ННІ «Дніпровський металургійний інститут»,  
ІВК «Системні технології», 2026

## ІМІТАЦІЙНА МОДЕЛЬ СИСТЕМИ ВИМІРЮВАННЯ ФІЗИКО-ХІМІЧНИХ ПАРАМЕТРІВ ПММ НА ОСНОВІ КОЛІРНИХ МЕРЕЖ ПЕТРІ

*Анотація.* В статті розглядається побудова моделі системи вимірювання рівня паливно-мастильних матеріалів (ПММ) на основі колірних мереж Петрі. Проведено порівняльний аналіз методів та систем вимірювання рівня палива. Наведено опис вимірювальної системи на основі магнітострикційного методу та алгоритм її роботи. Система складається з 4 входів для сигналів різних типів, що надходять від датчиків, двох модулів обробки сигналів та модуля індикації. Модель системи вимірювання рівня палива побудована в середовищі CPN Tools. Наведено детальний опис структури моделі. Проведено імітаційне моделювання. Були досліджені часові та структурні параметри моделі. Представлено результати статистичної обробки параметрів моделі.  
*Ключові слова:* Імітаційна модель, фізико-хімічні параметри, паливно-мастильні матеріали, моніторинг, система вимірювання, колірні мережі Петрі

**Вступ.** Точне та своєчасне вимірювання фізико-хімічних параметрів у режимі реального часу є досить важливим та актуальним завданням. З кожним роком обсяги транспортування нафтопродуктів зростають, отже, зростає значення достовірності вимірюваних показників. Організація інформаційних систем обліку нафтопродуктів є важливим науково-технічним завданням у структурі паливно-енергетичного комплексу.

Вимірювання рівня рідин є ще однією ключовою операцією в автоматизації технологічних процесів у нафтохімічній та нафтопереробній промисловості, а також у питаннях комерційного обліку світлих нафтопродуктів [1,2]. Враховуючи значні обсяги зберігання нафтопродуктів, а також особливості експлуатації відповідного технологічного обладнання, до інформаційних систем обліку нафтопродуктів та їх окремих компонентів визначено досить жорсткі вимоги [3,4].

Вимоги до обладнання такі:

- точність первинних перетворювачів (з урахуванням значної маси продукту на одиницю рівня для резервуарів);
- можливість дистанційних засобів параметризації та калібрування первинних перетворювачів;
- широкий діапазон робочих температур вимірювальних середовищ (-40...+85°C);
- вибухобезпечне виконання компонентів інформаційних систем, що експлуатуються у вибухонебезпечних зонах різних категорій.

Система моніторингу безперервного руху паливно-мастильних матеріалів в системах зберігання є комплексним рішенням для автоматизації контролю параметрів палива в режимі реального часу. Покращення параметрів систем контролю паливно-мастильних матеріалів є актуальним завданням, яке вимагає системного підходу та використання нових технологічних методів і підходів. Побудова моделі вимірювальної системи та моделювання процесу вимірювання дозволяє підвищити ефективність та точність вимірювальної системи шляхом вибору оптимальних режимів її роботи.

**Постановка проблеми.** Метою даної роботи є побудова та дослідження імітаційної моделі системи вимірювання фізико-хімічних параметрів паливно-мастильних матеріалів на основі колірних мереж Петрі.

**Аналіз останніх досліджень.** Сьогодні існує значна кількість технологій та приладів для вимірювання параметрів нафтопродуктів (пікнометри, ареометри). Але більшість приладів служать для вимірювання цих показників у лабораторних умовах [5-7]. Для отримання більш точних результатів вимірювань необхідно вимірювати параметри безпосередньо в магістральних трубопроводах та сховищах у режимі реального часу [3]. Водночас актуальною залишається проблема визначення в'язкості-температурних залежностей та приведення параметрів до заданих температур або визначення їх при різних температурах [3, 8].

Сучасні методи вимірювання рівня світлих нафтопродуктів (палива) проводяться різними методами та приладами: поплавковими [9], гідростатичними [10], ємнісними [9], акустичними [10], радіоізотопними [11]. Можливості яких визначаються, з одного боку, технологічними вимогами (дальність, точність, вибухозахист), а також наявністю неінформативних параметрів (тиск, густина, в'язкість, електропровідність тощо) [12]. Огляд сучасних технологій вимірювання нафтопродуктів представлено в роботах [12,13].

Магнітострикційні рівнеміри сьогодні є найпоширенішими на ринку обладнання для зберігання нафти, завдяки низці переваг [14]: безконтактний принцип роботи забезпечує тривалий термін служби, лінійність функції перетворення, повторюваність результатів, високу роздільну здатність, нечутливість до таких факторів, як вібрації, удари, волога, забруднення, постійний магніт, як активний елемент конструкції, не потребує джерела живлення. У роботі [15] детально розглянуто сучасні підходи до магнітострикційних методів вимірювання, а також запропоновано новий підхід до реалізації системи вимірювання. У роботах [16] представлено побудову імітаційних моделей системи вимірювання нафтопродуктів. У роботах [17,18] представлено розробку систем контрольних вимірювань з використанням мереж Петрі. Для подальшого вдосконалення процесу контролю та вимірювання нафтопродуктів необхідно підвищити ефективність та точність процесу діагностики та вимірювання шляхом використання більш ефективних технологій моделювання.

**Структурна схема контролю вимірювання фізико-хімічних параметрів ПММ.** Структурна схема контролю вимірювання фізико-хімічних параметрів паливно-мастильних матеріалів (ПММ) у ємностях закритого типу представлена на рисунку 1.

Загалом, система керування може мати 4 входи, до яких підключаються датчики та перетворювачі.

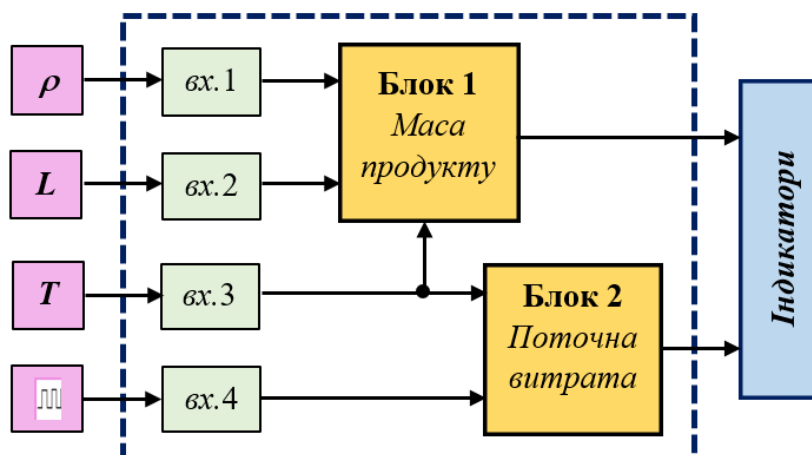


Рисунок 1 – Структурна схема процесу вимірювання параметрів ПММ

Сигнали, що надходять від датчиків, вже перетворені в один з типів уніфікованих сигналів. Як уніфікований сигнал досить часто використовується струмовий сигнал 4-20 мА. Вихід 4-20 мА забезпечує необхідний рівень безпеки для обладнання, що використовується у вибухонебезпечних зонах, дозволяючи усунення несправності у електричних колах без знеструмлення датчика.

Входи 1-3 мають однакову структуру. На вхід 1 надходить сигнал від датчиків щільності. На виході блоку «вх.1» маємо струмовий сигнал, пропорційний значенню щільності продукту. На «вх.2» надходить сигнал від датчиків рівня, тоді на виході цього блоку у отримуємо значення рівня продукту в міліметрах. Ці два сигнали використовуються системою для розрахунку поточної маси продукту в резервуарі (**Блок 1**) за наступним узагальненим алгоритмом.

*Крок 1.* Рівень продукту в міліметрах перетворюється в одиниці об'єму за допомогою таблиць переведення, які складаються для кожного резервуара під час його перевірки з урахуванням його геометричних розмірів. У цій таблиці кожен сантиметр рівня продукту відповідає певній кількості літрів. Для розрахунку об'єму в проміжних точках використовуються різні варіанти апроксимації.

*Крок 2.* Використовуючи значення щільності, залишки продукту перетворюються в одиниці маси.

*Крок 3.* За поточною температурою продукту обчислюється значення маси при 15 градусах Цельсія. Як того вимагає законодавство, дані про залишки продукту надаються округленими до 15 градусів.

*Крок 4.* Вхід 3 (вх.3) отримує дані про температуру, які були виміряні в 5 точках по висоті резервуара.

*Крок 5.* Дані підсумовуються та обчислюється середня температура продукту. Ця температура використовується у всіх наступних розрахунках.

*Крок 6.* Система відключає ті датчики, які знаходяться вище рівня продукту, щоб зменшити похибку вимірювання, оскільки температура газової суміші порожньої ча-

стини резервуара (повітря + ПММ) значно відрізняється від температури самого продукту.

Вхід 4 (вх.4) отримує сигнал від витратоміра. Більшість моделей зараз працюють з імпульсними перетворювачами на виході. Імпульсний сигнал, що подається на вхід, посилюється та подається на лічильник імпульсів. Знаючи кількість імпульсів та їх «вагу» (наприклад, 1 імпульс = 1 літр), система розраховує поточну витрату продукту під час розвантаження або завантаження в резервуар за поточної температури (Блок 2). Ці дані також перераховуються на 15 градусів.

Всі отримані та розраховані дані від датчиків передаються на блок відображення: електронне табло, монітор комп'ютера, мобільний телефон тощо.

**Побудова імітаційної моделі процесу вимірювання параметрів ПММ.** Для побудови імітаційної моделі вимірювальної системи використовуються колірні мережі Петрі. Побудова моделей систем вимірювання паливно-мастильних матеріалів виконувалася за допомогою інструментів *CPN Tools* [19,20].

Загальна модель процесу керування вимірюванням фізико-хімічних параметрів паливно-мастильних матеріалів (ПММ) у закритих контейнерах має 8 функціональних блоків (рис. 2).

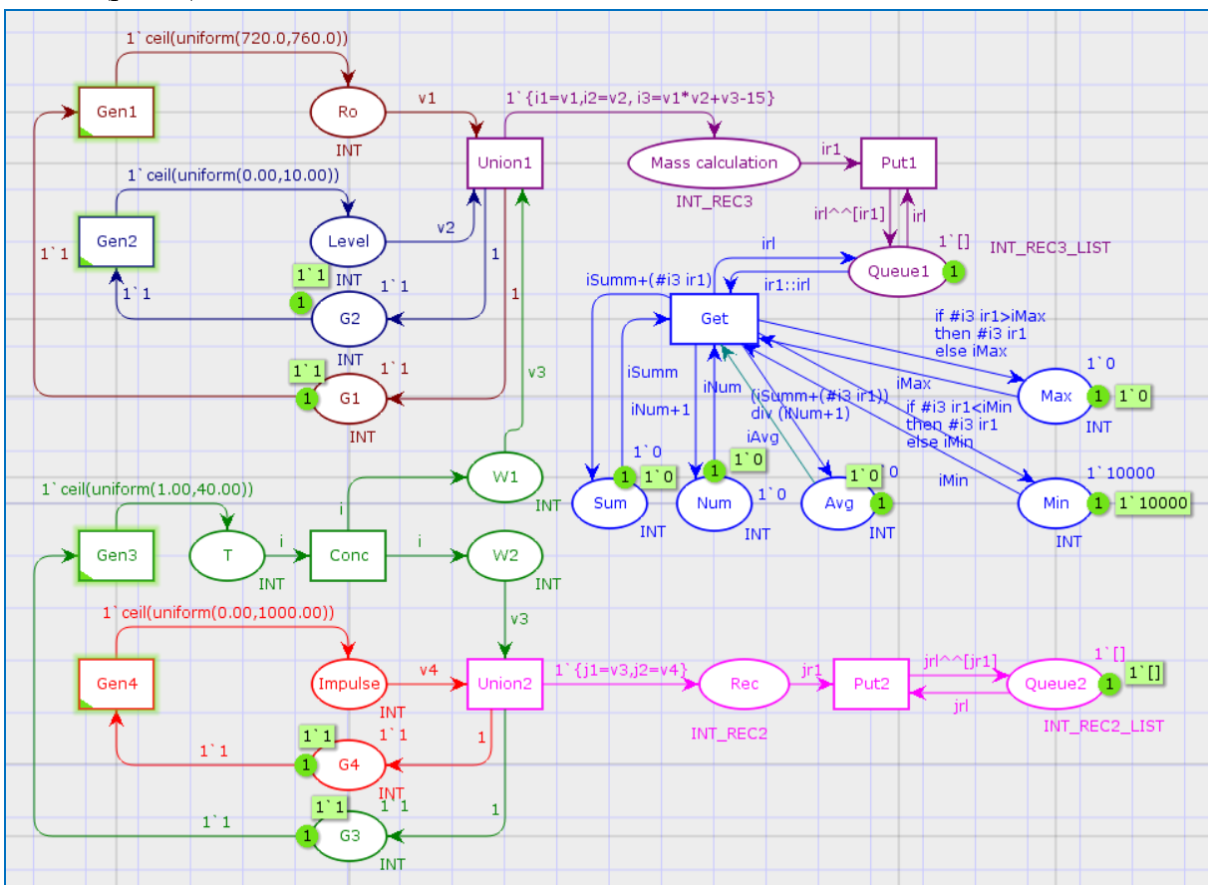


Рисунок 2 – Імітаційна модель процесу керування вимірюванням фізико-хімічних параметрів ПММ у ємностях закритого типу

Блок фіксації щільності ПММ представлений переходом (*Gen1* – виконує функцію генерації щільності ПММ на прикладі бензину) та 2 позиціями (*Ro* – для розміщення згенерованого значення щільності ПММ, *G1* – для повторення процесу генерації).

Блок фіксації рівня ПММ у контейнері представлений переходом (*Gen2* – виконує функцію генерації рівня ПММ) та 2 позиціями (*Level* – для розміщення згенерованого значення рівня ПММ, *G2* – для повторення процесу генерації).

Блок фіксації середнього значення температури ПММ у резервуарі представлений переходом (*Gen3* – виконує функцію генерації температури ПММ) та 2 позиціями (*T* – для розміщення згенерованого значення температури ПММ, *G3* – для повторення процесу генерації). Значення температури використовується в системі у двох блоках (блоку розрахунку маси та блоку накопичення витрат), тому додатково додано

Структурний блок розгалуження. Він забезпечує можливість паралельного використання значення температури двома блоками. Блок представлений переходом (*Conc* – виконує функцію розгалуження) та 2 позиціями (*W1*, *W2* – для накопичення дубльованих значень температури).

Блок фіксації значення імпульсного сигналу, що вказує на споживання ПММ, представлений переходом (*Gen4* – виконує функцію генерації імпульсного сигналу) та 2 позиціями (*Impulse* – для розміщення значень імпульсних сигналів, *G4* – для повторення процесу генерації).

Блок розрахунку маси представлений 2 переходами (*Union1* – для формування запису параметрів ПММ, *Put1* – для формування списку записів параметрів ПММ) та 2 позиціями (*Mass calculation* – для розміщення записів параметрів ПММ, *Queue1* – для розміщення списку записів параметрів ПММ).

Блок накопичення витрат представлений 2 переходами (*Union2* – для формування запису параметрів ПММ, *Put2* – для формування списку записів параметрів ПММ) та 2 позиціями (*Rec* – для розміщення записів параметрів ПММ, *Queue2* – для розміщення списку записів параметрів ПММ).

Блок для обчислення статистики зміни маси ПММ представлений переходом (*Get* – виконує обчислювальні дії) та 5 позиціями (*Sum* – для зберігання загальної маси ПММ, *Num* – для зберігання кількості фіксованих значень маси, *Avg* – для зберігання середнього значення фіксованих значень маси, *Min* – для зберігання мінімального значення маси, *Max* – для зберігання максимального значення маси).

Для початку встановлюються всі необхідні модулі (рис. 1) та підключаються до проекту. Набір даних додається до проекту. Далі, для зручності, видалено зайві стовпці, які не мають сенсу для імітаційної моделі, що розроблено і представлено на рисунку 2.

Завдання декларації системи. Визначено параметри системи, яка моделюється. Для колірних мереж Петрі це набір системних кольорів (*colorset*), які мають такі значення:

– кортеж *INT\_REC3* має 3 ціле чисельні поля, які, згідно із завданням, відповідають за *i1* – значення щільності ПММ, *i2* – значення рівня ПММ, *i3* – розрахункове значення маси ПММ;

– фішки типу *INT\_REC3\_LIST* – це 1-й список записів параметрів ПММ.

– кортеж *INT\_REC2* має 2 ціле чисельні поля, які, згідно із завданням, відповідають за *j1* – середнє значення температури ПММ, *j2* – значення імпульсних сигналів, що вказують на споживання ПММ;

– фішки типу *INT\_REC2\_LIST* – це 2-й список записів параметрів ПММ.

Змінні моделі:

– *irl* – визначає надходження інформації про першу групу параметрів ПММ у чергу;

– *jrl* – визначає надходження інформації про другу групу параметрів ПММ у чергу;

– *v1, v2, v3, v4* – визначають відповідні поля інформації про параметри ПММ;

– *ir1* – визначає повну інформацію про параметри ПММ першої групи;

– *jr1* – визначає повну інформацію про параметри ПММ другої групи.

– Після визначення основних оголошень системи, параметри моделі задаються на графіках мережі. На аркуші моделі представлені загальні параметри моделі (див. рис. 2):

– у позиціях *Ro, Level, T, Impulse, G1, G2, G3, G4, W1, W2* набір кольорів фішок – *INT*; початкова розмітка відсутня;

– у позиціях *Mass calculation, Rec* набору кольорів фішок – *INT\_REC3, INT\_REC2* відповідно початкова розмітка відсутня;

– у позиціях *Queue1, Queue2*, набір кольорів фішок – *INT\_REC3\_LIST* та *INT\_REC2\_LIST*; початкова позначка *I`[]* вказує, що спочатку черги моделювання порожні;

– у позиціях *Sum, Num, Avg, Max*, набір кольорів фішок є *INT*; початкова позначка *I`0* визначає початкове значення параметра перед процесом моделювання;

– у позиції *Min*, набір кольорів фішок є *INT*; початкова позначка *I`10000* визначає початкове значення параметра перед процесом моделювання;

– на дузі від переходу *Put1* до позиції *Queue1*, вираз *irl ^{ir1}* вказує на прихід запису параметрів першої групи ПММ у чергу;

– на дузі від позиції *Queue1* до переходу *Put1*, вираз *irl* вказує на зворотний зв'язок;

– на дузі від переходу *Put2* до позиції *Queue2* вираз *jrl ^{jr1}* визначає надходження запису параметрів другої групи ПММ у чергу;

– на дузі від позиції *Queue2* до переходу *Put2* вираз *jrl* визначає зворотний зв'язок;

– на дузі від позиції *Queue1* до переходу *Get* вираз *ir1::irl* визначає, що обчислення статистики можна розпочати, якщо в черзі є хоча б один запис параметрів ПММ;

– на дузі від переходу *Union1* до позиції обчислення *Mass* вираз  $I\{i1=v1, i2=v2, i3=v1*v2+v3-15\}$  встановлює мітки типу запису на основі вхідних змінних *v1, v2, v3*;

– на дузі від переходу *Union2* до позиції *Rec* вираз  $I\{j1=v3, j2=v4\}$  встановлює мітки типу запису на основі вхідних змінних *v3, v4*;

– на дузі від переходу *Gen1* до позиції *Ro*, вираз  $1 \text{`ceil}(\text{uniform}(720.0,760.0))$  генерує мітки з випадковими значеннями, використовуючи функцію рівномірного розподілу;

– на дузі від переходу *Gen2* до позиції *Level*, вираз  $1 \text{`ceil}(\text{uniform}(0.0,10.0))$  генерує мітки з випадковими значеннями, використовуючи функцію рівномірного розподілу;

– на дузі від переходу *Gen3* до позиції *T*, вираз  $1 \text{`ceil}(\text{uniform}(1.0,40.0))$  генерує мітки з випадковими значеннями, використовуючи функцію рівномірного розподілу;

– на дузі від переходу *Gen4* до позиції *Impulse*, вираз  $1 \text{`ceil}(\text{uniform}(0.0,1000.0))$  генерує мітки з випадковими значеннями, використовуючи функцію рівномірного розподілу.

**Результати моделювання процесу вимірювання параметрів ПММ.** При моделюванні процесу вимірювання параметрів ПММ було виконано 10 прогонів для моделювання процесу контролю вимірювання фізико-хімічних параметрів ПММ у емностях закритого типу, та сформовано 100 наборів параметрів першої та другої груп. Остаточні розмітки одного з мережевих прогонів представлені на рисунку 3.

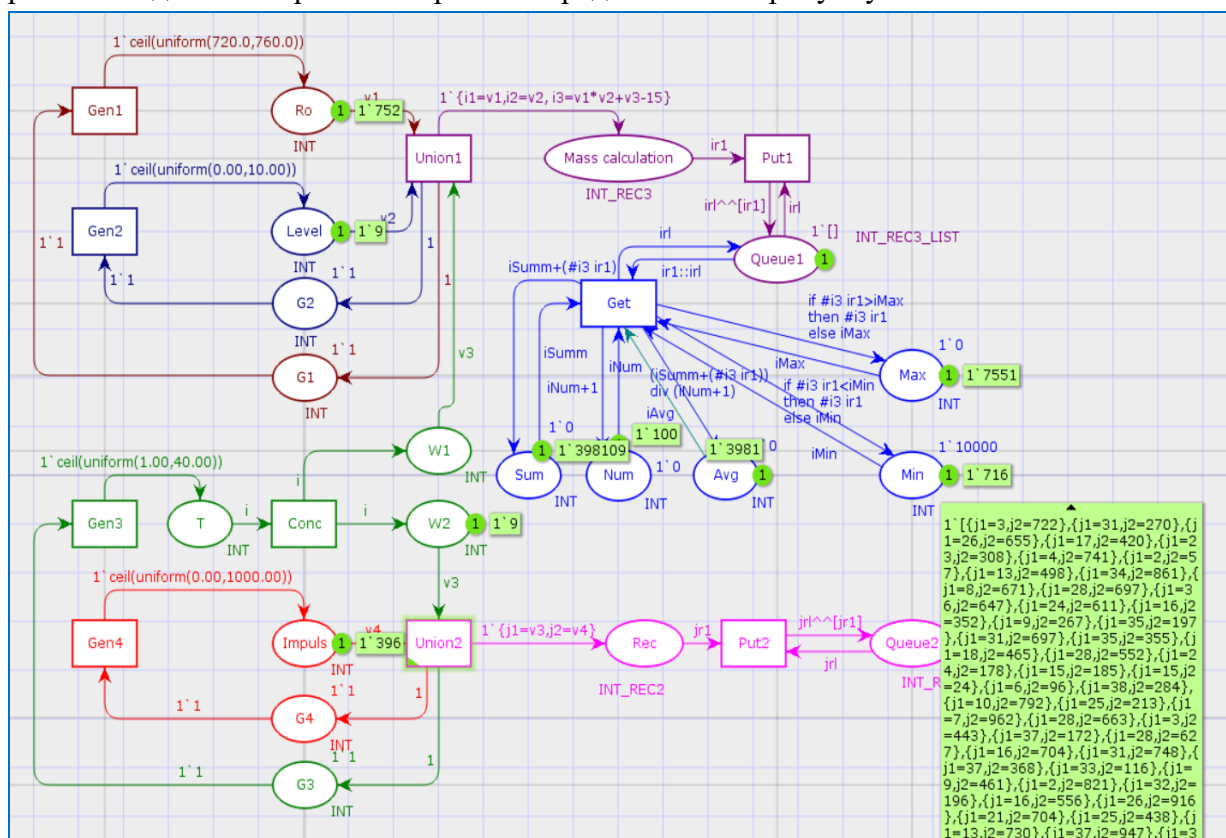


Рисунок 3 – Результати імітаційного моделювання процесу моніторингу вимірювання фізико-хімічних параметрів ПММ у емностях закритого типу

Аналіз результатів моделювання представлено в таблиці 1. Сумарні значення представляють усереднені результати моделювання процесу контролю вимірювання параметрів ПММ з кількістю кроків моделювання 1000.

В наведеному прикладі розглянуто метод накопичення інформації в позиціях. Деякі характеристики моделі можна розраховувати в режимі реального часу, під час процесу моделювання, інші можуть вимагати додаткових обчислень у спеціалізованих математичних програмах. Тоді зручно зберегти результати моделювання у файл для подальшого аналізу.

Таблиця 1

Аналіз результатів моделювання процесу моніторингу вимірювання фізико-хімічних параметрів ПММ у емностях закритого типу

№	Кількість значень маси	Загальне значення маси	Середнє значення маси	Мінімальне значення маси	Максимальне значення маси
1	100	398109	3981	716	7551
2	99	404297	4083	719	7592
3	99	423023	4272	736	7572
4	98	405122	4133	734	7515
5	99	386730	3906	738	7604
6	98	392823	4008	721	7615
7	98	428417	4371	722	7585
8	98	412768	4211	713	7589
9	99	424691	4289	737	7577
10	99	413051	4172	722	7604
Сумарні значення		408903,1	4142,6	725,8	7580,4

**Висновки.** В роботі описано побудову моделі системи вимірювання рівня палива на основі колірних мереж Петрі. Проведено порівняльний аналіз методів та систем вимірювання рівня палива. Наведено опис вимірювальної системи на основі магнітострикційного методу та алгоритм її роботи. Система складається з 4 входів для сигналів різних типів, що надходять від датчиків, двох модулів обробки сигналів та модуля індикації. Модель системи вимірювання рівня палива побудована в середовищі *CPN Tools*. Наведено детальний опис структури моделі. Проведено імітаційне моделювання. Були досліджені часові та структурні параметри моделі. Представлено результати статистичної обробки параметрів моделі. Побудована модель дозволяє підвищити ефективність системи вимірювання рівня палива шляхом оптимізації структури та збільшення швидкості обробки інформації.

#### ЛІТЕРАТУРА / REFERENCE

1. Meškuotienė A., Kaškonas P., Urbonavičius B.G., Gintautas Balčiūnas G., Dobilienė J. Analysis of Liquid Quantity Measurement in Loading/Unloading Processes in Cylindrical Tanks. MDPI. Journal “Computation”. (2022) Vol. 10, issue 7. p. 122. DOI: 10.3390/computation10070122.
2. Mohindru P. Development of liquid level measurement technology: A review. Flow Measurement and Instrumentation. (2022) P. 102295. 10.1016/j.flowmeasinst.2022.102295.

3. Yang S. Optimization of oil depot operation process. IOP Conference Series: Earth and Environmental Science. (2019) Vol. 332, issue 2, p. 022044.  
DOI: 10.1088/1755-1315/332/2/022044.
4. Khlebnikova E., Sundar K., Zlotnik A., Bent R., Ewers M., Tasseff B. Optimal economic operation of liquid petroleum products pipeline systems AIChE Journal. (2021). Vol. 67, issue 4, e17124. DOI: 10.1002/aic.17124.
5. Neyezhnikov P., Narodnytskyi G. Measures to ensure the necessary accuracy of accounting petroleum products in the tanks. Ukrainian Metrological Journal. (2020) No. 4, pp. 16–21. DOI: 10.24027/2306-7039.4.2020.224266.
6. Boltaboyev I. M., Komilov M. A. Development of automated analytical systems of physical and chemical parameters of oil and petroleum products. Journal “Theoretical and Applied Science”. (2020) Vol. 86, no. 06, pp. 711–713. DOI: 10.15863/tas.2020.06.86.131.
7. Semenov A., Zviahin O., Kryvinska N., Semenova O. Device for measurement and control of humidity in crude oil and petroleum products. Journal “Metrology and measurement systems”. (2023) Vol. 30, no. 1, pp. 195–208. DOI: 10.24425/mms.2023.144865.
8. Zhou Z., Ren T., Li M., He M., Zhou X. Study on Temperature Measurement of Petroleum Product Density Analyzer Based on Hydrometer Method. Journal of Physics: Conference Series. (2023) Vol. 2428, no. 1, p. 012001. DOI: 10.1088/1742-6596/2428/1/012001.
9. Monteiro M., Svet V., Sandilands D., Tsysar S. Experimental Investigations of Various Methods of Sludge Measurements in Storage Oil Tanks. Advances in Remote Sensing, (2015) No.4 (02), pp.119-137. DOI:10.4236/ars.2015.42011.
10. Pospíšil J., Dandoš R. Basic Principles of Hydrostatic Levelling. Journal “GeoScience Engineering”. (2018) No.64(2), pp. 12-21. DOI:10.2478/gse-2018-0008.
11. Aslam M.Z., Tang T. A High-Resolution Capacitive Sensing System for the Measurement of Water Content in Crude Oil. Journal “Sensors”. (2014) No.14, pp.11351-11361. DOI: 10.3390/s140711351.
12. Lukonge A.B., Cao X. Leak Detection System for Long-Distance Onshore and Offshore Gas Pipeline Using Acoustic Emission Technology. A Review. Trans Indian Inst Met 73, Pp. 1715–1727 (2020). DOI: 10.1007/s12666-020-02002-x.
13. Roshani M., et. al. Proposing a gamma radiation based intelligent system for simultaneous analyzing and detecting type and amount of petroleum by-products. Journal “Nuclear Engineering and Technology”, vol. 53, issue 4, April 2021, pp. 1277-1283  
DOI: 10.1016/j.net.2020.09.015.
14. Kul S., et. al. Experimental comparison of localised magnetostriction difference under sinusoidal and PWM excitations. Journal of Magnetism and Magnetic Materials. Vol. 544, 15 February 2022. DOI 10.1016/j.jmmm.2021.168692.
15. Aydin U., et. al. Magneto-mechanical modeling of electrical steel sheets. Journal of Magnetism and Magnetic Materials. Vol. 439, 1 October 2017, pp. 82-90.  
DOI: doi.org/10.1016/j.jmmm.2017.05.008.
16. Shilyanski G., et. al. Spatial distributions of magnetostriction, displacements and noise generation of model transformer cores. International Journal of Mechanical Sciences, Vol. 118, November 2016, pp. 188-194. DOI: doi.org/10.1016/j.ijmecsci.2016.09.022.

17. Pedro M. González del Foyo, et. al. Requirements analysis of automated projects using uml/petri nets. 21st International Congress of Mechanical Engineering, Proceedings of COBEM 2011. October 24-28, 2011, Natal, RN, Brazil.
18. Fecarotti C., et. al. Performance Modelling of Fuel cells Systems Through Petri Nets. Journal Hydrogen and Fuel Cell SUPERGEN Researcher Conference University of Birmingham. Proceedings of HFCS. December 2014.
19. Калініна І.О., Гожий О.П., Шиян С.І., Нечахін В.В. Імітаційне моделювання систем зі складним стохастичним процесом обробки даних за допомогою кольорових мереж Петрі. Регіональний міжвузівський збірник наукових праць «Системні технології», Дніпро. 2022. Вип. 6, № 143. С. 42-56. DOI: 10.34185/1562-9945-6-143-2022-04.
20. Калініна І.О., Гожий О.П. Моделювання складних систем на основі кольорових мереж Петрі: навчальний посібник. Херсон: книжкове видавництво ФОП Вишемирський В.С., 2021. – 60 с.

Received 30.03.2026.  
Accepted 01.04.2026.  
Published 30.04.2026

### ***Simulation model of the system for measuring physical***

#### ***and chemical parameters of fuel and lubricants based on coloured Petri networks.***

*The article considers the construction of a model of a fuel and lubricant level measurement system based on colour Petri nets. A comparative analysis of fuel level measurement methods and systems is carried out. A description of a measuring system based on the magnetostrictive method and its operation algorithm are given. The system consists of 4 inputs for signals of different types coming from sensors, two signal processing modules and an indication module. The model of the fuel level measurement system is built in the CPN Tools environment. A detailed description of the model structure is given. Simulation modelling was carried out. The time and structural parameters of the model were investigated. The results of statistical processing of the model parameters are presented.*

*The paper describes the construction of a model of a fuel level measurement system based on color Petri nets. A comparative analysis of fuel level measurement methods and systems is carried out. A description of a measuring system based on the magnetostrictive method and its operation algorithm are given. The system consists of 4 inputs for signals of different types coming from sensors, two signal processing modules and an indication module. The model of the fuel level measurement system is built in the CPN Tools environment. A detailed description of the model structure is given. Simulation modeling is carried out. The time and structural parameters of the model were investigated. The results of statistical processing of the model parameters are presented. The constructed model allows to increase the efficiency of the fuel level measurement system by optimizing the structure and increasing the speed of information processing.*

*Keywords: simulation model, physicochemical parameters, fuel and lubricants, monitoring, measurement system, colour Petri nets*

**Шиян Сергій Іванович** – аспірант кафедри інтелектуальних інформаційних систем Чорноморського національного університету ім. Петра Могили.

ORCID: <https://orcid.org/0000-0001-9255-9511>

**Калініна Ірина Олександрівна** – професорка кафедри інтелектуальних інформаційних систем Чорноморського національного університету ім. Петра Могили, д.т.н., проф.

ORCID: <https://orcid.org/0000-0001-8359-2045>

**Shiyan Serhii Ivanovych** – Postgraduate student of the Department of Intellectual Information Systems of the Black Sea National University named after Petera Mohyly.

ORCID: <https://orcid.org/0000-0001-9255-9511>

**Kalinina Iryna Oleksandrivna** - professor of the Department of Intellectual Information Systems of the Black Sea National University named after Petra Mohyly, Doc.of Ing.Sc.,professor.

ORCID: <https://orcid.org/0000-0001-8359-2045>

## СТВОРЕННЯ ДАТАСЕТУ НА ОСНОВІ ПОДІЙ КЛАСИФІКОВАНИХ СИСТЕМОЮ ВИЯВЛЕННЯ МЕРЕЖЕВИХ ВТОРГНЕНЬ

*Анотація.* У статті представлено підхід до формування датасету для навчання моделей машинного навчання в контексті мережевої системи виявлення вторгнень Snort 3. На відміну від класичних датасетів, запропонований набір даних будується на основі нормалізованих байтових буферів інспектора та легкої телеметрії пакета, доступних під час онлайн-обробки трафіку. Ground truth задається контрольованим походженням трафіку (attack/benign PCAP), тоді як спрацювання правил Snort розглядаються як “teacher”-сигнал для подальшої побудови ризик-скорингу. Датасет сформовано для Fast Pattern-групи SIP/2.0 і містить десятки тисяч подій зі стандартизованим поділом на train/validation/test. Додатково виконано аналіз інформативності байтових позицій (на основі дивергенції Jensen–Shannon та ентропії) і кореляційний аналіз телеметрії, що підтверджує наявність локалізованого дискримінативного сигналу та відсутність тривіальних витоків через padding. Отриманий датасет може слугувати основою для нейромережових моделей реального часу, які доповнюють сигнатурну детекцію оцінкою ризику.

*Ключові слова:* Snort 3, NIDS, датасет, SIP/2.0, Fast Pattern, байтові дані, телеметрія пакета, ентропія, risk scoring, нейронні мережі реального часу, precision/recall.

**Постановка задачі.** Виявлення вторгнень є ключовим елементом захисту мереж, допомагаючи адміністраторам вчасно реагувати на шкідливі дії, такі як атаки та зловмисне ПЗ. Системи виявлення вторгнень (IDS) – обов’язковий рівень захисту критичних мереж від зростаючих загроз. З часом, атаки еволюціонують, в унісон з ними повинні еволюціонувати і методи їх виявлення. Проте дослідники часто стикаються з браком якісних і достовірних даних для тестування своїх напрацювань. Існуючі відкриті набори даних, містять різні типи атак і нормальний трафік, але мають обмеження, наприклад, неповне маркування або відсутність підтвердження правдивості. Тому створення датасетів на основі подій, маркованих як експертом так і системами виявлення вторгнень, може підвищити точність навчання та оцінки IDS, забезпечуючи більш реалістичні та різноманітні дані для розробок у сфері безпеки.

У роботі пропонується підхід до формування датасету, який використовує Snort 3 як класифікатор та джерело даних. **Мета** – отримати максимально наближений до реального процесу виявлення IDS набір даних, де джерелом teacher-міток є події Snort 3 (зокрема, спрацювання правил і Fast Pattern [1]), а ground truth визначається контрольо-

ваним походженням трафіку (attack/benign PCAP). Вхід моделі при цьому становлять сирі нормалізовані буфери, сформовані самим Snort (тобто те, що реально доступне системі під час аналізу даних на предмет атаки).

Поточний реліз зосереджений на прикладному трафіку SIP/2.0 як показовому прикладі Fast Pattern-групи, однак обраний формат (корисні дані + телеметрія) та джерело міток (події Snort) спроектовано для перенесення на інші протоколи без зміни підходу. Варто зазначити, що робота не торкається шифрованих протоколів і пост-експлуатаційних стадій, а також, що специфіка ревізії правил Talos LightSPD задає «зріз» загроз у межах експерименту.

Варто зазначити що підхід є симуляційним, стартує з однієї Fast Pattern-групи і залежить від поточної ревізії правил; ці аспекти обговорюються в подальших розділах разом зі стратегіями пом'якшення (варіативність трафіку, розширення протокольного охоплення, оновлення правил).

У ході статті:

- формується IDS-центричний датасет, де події та мітки отримані безпосередньо від Snort 3, а вихідними даними є нормалізовані буфери, що використовує і формує IDS.

- пропонується протокол-агностичний байтовий формат для використання у сфері машинного навчання без ручного виділення характеристик.

Ця робота закладає основу для онлайн-інтеграції нейромережі у систему виявлення мережевих вторгнень. Використовуючи навчання за мітками «учителя» в реальному часі і узагальнення за межі правил, планується використання моделі як додаткового шару перевірки/ранжування спрацювань у реальних сценаріях.

**Аналіз останніх досліджень і публікацій.**

### **1. Підходи до створення IDS-датасетів**

У галузі мережевої безпеки давно відзначається брак відкритих, якісних та актуальних датасетів для навчання і тестування систем виявлення вторгнень (IDS) [2]. Ефективність моделей машинного навчання для IDS значною мірою залежить від наявних наборів даних, однак отримати репрезентативний датасет непросто через відсутність стандартизованих методик його побудови, проблеми з документуванням процесу та неповні або неточні мітки класів [3]. Багато загальноживаних бенчмарк-датасетів, таких як KDD'99, нині вважаються застарілими: вони не відображають сучасного стану мережевого трафіку (наприклад, масового шифрування) та актуальних загроз [4]. В літературі описано кілька підходів до генерування даних для IDS, які можна класифікувати за джерелом трафіку на реальні, симульовані та гібридні [5]. Кожен із цих підходів має свої переваги і недоліки, які розглянемо нижче.

*Реальний трафік.* Перший підхід передбачає збір реальних мережевих трас з існуючих середовищ – наприклад, шляхом моніторингу виробничої мережі або використанням honeypot-систем для заманювання зловмисників. Перевагою є висока автентичність даних: реальний трафік містить правдоподібні фонові патерни та непередбачувані аномалії. Проте існують суттєві перешкоди: питання приватності користувачів і конфіденційності даних зазвичай унеможливають відкриту публікацію необроблених запи-

сів, вимагаючи анонімізації адрес і видалення корисного навантаження, що знижує цінність таких даних [6]. Крім того, пасивно зібраний трафік може не містити всіх необхідних видів атак: якщо під час захоплення не сталося певної атаки (наприклад, DDoS чи вторгнення), її не буде в датасет. Через ці труднощі чисто реальні датасети трапляються рідко: за оцінками, лише близько 17–20% публічних наборів містять виключно реальний нормальний трафік, і ця частка зменшується останніми роками [5]. Прикладом реального датасету є MAWILab, що містить довготривалі добові дампи трафіку з реальної магістральної мережі (15-хвилинні вибірки щодня) із мітками аномалій, отриманими шляхом кореляції кількох детекторів [7]. Інший підхід – розгортання honeypot-інфраструктур: так, датасет Kyoto 2006+ накопичувався протягом 2006–2015 рр. з мережі пасток і містить статистичні ознаки атак на ці пастки [8].

*Симульований трафік.* Другий підхід – емуляція мережевого трафіку у контрольованому середовищі (лабораторії). Більшість сучасних IDS-датасетів саме так і генеруються, враховуючи складність отримання реальних даних [5]. Дослідники розгортають тестову мережу (фізичну або віртуальну), в якій імітується легітимний фоновий трафік (робота користувачів, сервісів) та інсценуються атаки з відомими параметрами. Класичний приклад – набір DARPA’98/KDD’99, зібраний в рамках оцінювання IDS у 1998–1999 рр.: семитижневий трафік було згенеровано на випробувальному полігоні, де поряд із нормальним навантаженням було навмисно виконано 24 види атак, віднесених до чотирьох категорій (Denial of Service, Remote-to-Local, User-to-Root, Probe) [9]. Хоча KDD’99 став найбільш відомим бенчмарком, пізніші дослідження відзначили його недоліки – присутність багатьох дублікатів, нерівномірність класів та застарілість набору атак [10]. У відповідь було створено його похідні, зокрема NSL-KDD, де усунуто дублікатні записи та відфільтровано частину некоректних сесій, однак загалом ці набори все ще базуються на трафіку 90-х років.

Сучасні емульовані датасети прагнуть бути ближчими до реальності. Для генерування нормального трафіку часто застосовують профілі поведінки користувачів: наприклад, Shiravi та ін. запропонували систему B-Profile, яка моделює типові дії користувачів (веб-перегляд, перегляд пошти, передача файлів тощо) за допомогою автоматів і скриптів [11]. Канадський набір CICIDS-2017 є показовим прикладом: у тестовій мережі Університету Нью-Брансвіка було реалізовано реалістичний фон із емуляцією 25 видів активності користувачів (HTTP, HTTPS, FTP, SSH, e-mail тощо), після чого дослідники програли в мережі кілька атакуючих сценаріїв. Зокрема, CICIDS-2017 включає такі типи атак, як brute force підбір паролів, експлоїт Heartbleed, ботнет-діяльність, DoS/DDoS, web-атаки (SQL-ін’єкції, XSS) та приховане проникнення в мережу. Атаки були згенеровані публічними інструментами та скриптами, наближеними до реальних зловмисних дій [12]. Подібним чином був побудований австралійський датасет UNSW-NB15: у кібер-лабораторії використовували комерційний інструмент для генерації сучасного нормального трафіку і актуальних атак, зберігши їх у вигляді tcpdump-трейсів загальною тривалістю 31 година; з цих трас потім було витягнуто 49 ознак трафіку, розбитих на групи (рівні потоку, пакету, сесії тощо) [13]. Основна перевага симуляції –

повний контроль над процесом та точна розмітка: кожна атака відома наперед, тому кожному запису можна призначити правильну мітку (нормальний чи шкідливий).

Проте слабким місцем є реалістичність згенерованого трафіку. Якщо емульований фон занадто спрощений, моделі IDS можуть навчитися виявляти не стільки реальні атаки, скільки артефакти штучного трафіку. Наприклад, використання простих генераторів (типа *iperf*) створює монотонний, однорідний трафік, який різко відрізняється від хаотичного реального фону – класифікатори легко відділяють такий "штучний нормальний" трафік від аномалій, і це призводить до завищених оцінок якості IDS [5]. Дійсно, занадто детерміновані сценарії можуть зробити задачу виявлення тривіальною (модель підлаштовується під нехарактерні шаблони симуляції), тому при емулюванні важливо досягти різноманітності та складності, близької до бойових мереж.

*Гібридні та синтетичні підходи.* Щоб збалансувати реалізм і контрольованість, дослідники все частіше звертаються до гібридних методів побудови датасетів. Один популярний підхід – ін'єкція синтетичних атак у реальний трафік: береться захоплений трафік реальної мережі (без міток), після чого в нього «вставляються» додаткові пакети/сесії, що відповідають відомим атакам [5]. Результатом є комбінований датасет: фон у ньому реалістичний (бо взятий із живої мережі), а атаки – синтетично згенеровані, але з відомим місцем і часом для розмітки. Прикладом такого підходу є набір UGR'16, зібраний з потоку NetFlow-записів іспанського провайдера: його калібраційна частина містить чисто реальний трафік, а тестова – той самий трафік, доповнений вставленими атаками (низькоінтенсивні DoS, сканування портів, ботнет) [14].

Окремо варто згадати про повністю синтетичне генерування трафіку за допомогою статистичних моделей або генеративного штучного інтелекту. Історично мережевий трафік намагалися синтезувати на основі простих статистичних розподілів, але новітні підходи використовують генеративно-змагальні мережі (GAN) та інші моделі глибокого навчання [5]. Наприклад, у роботі Ring та ін. (2019) запропоновано застосувати GAN для генерування мережевих потоків (NetFlow) з використанням техніки IP2Vec для вбудовування адрес у векторний простір [15]. Показано, що такі моделі можуть навчитися створювати синтетичний трафік, статистично схожий на реальний. Втім, повністю синтетичний трафік сьогодні вважається найменш реалістичним підходом – попри відсутність проблем приватності та високу гнучкість, він може не враховувати багатьох тонкощів живих мережевих взаємодій. Тому цілком штучні набори зазвичай використовуються як допоміжні або для вузьких завдань, тоді як у загальному випадку рекомендується комбінувати методи і перевіряти синтетичні дані на відповідність реальності [5].

## **2. Формат даних, розмітка та типи атак**

При створенні IDS-датасету важливо задокументувати його характеристики: формат подання трафіку, схему маркування даних та перелік типів атак, що містяться в наборі. За форматом можна виділити два підходи до публікації даних. Перший – це надання сирих мережевих трас, зазвичай у вигляді PCAP-файлів (повні пакети) або flow-записів (наприклад, NetFlow/IPFIX логи). Такий формат максимально гнучкий, але вимагає від дослідника самостійно опрацювати дані (виділяти ознаки, агрегувати пото-

ки тощо). Другий підхід – публікація вже оброблених вибірок: автори датасету самі витягують з сирих трас статистичні ознаки і надають підготовлений набір у табличній формі (CSV-файли або навіть готові векторизовані масиви) [3]. Наприклад, у згаданому вище CICIDS-2017 після захоплення, трафік було оброблено з допомогою CICFlowMeter [16], для виділення числових ознак на кожен мережевий потік (тривалість, кількість байтів, пакетів, середні статистики і т.д.). Кожен запис при цьому містить ідентифікатори сеансу (IP-адреси джерела й призначення, порти, протокол) та був класифікований [12]. Інший інструмент, часто використовуваний для аналізу трас, – це Zeek [17]: він дозволяє зберегти метадані сесій і одночасно застосувати скрипти для автоматичного визначення відомих атак або аномалій. Зокрема, при створенні нового датасету NIKARI-2021 дослідники взяли набір ознак CICIDS-2017 і за допомогою скриптів Zeek згенерували аналогічні показники для свого трафіку [3]. Вибір формату залежить від цілей: пакетні дані дають більше гнучкості (можна обчислити будь-які нові ознаки, перевірити власні методи виявлення), тоді як готові фічі спрощують порівняння алгоритмів (усі моделі тренуються на однаковому наборі полів).

Наявність ground truth (правдивих міток) є обов'язковою вимогою для IDS-датасету, але забезпечити її непросто. В емульованих наборах, як зазначалося, розмітка робиться вручну під час генерування: дослідник точно знає, коли і яка атака виконувалась, тому може позначити відповідні пакети чи потоки як "атака", а решту як "нормальний" трафік. У реальних же мережевих журналах абсолютно точна розмітка практично недосяжна: часто невідомо напевне, чи була певна аномалія справжньою атакою, чи легітимною аномальною подією. Через це частина публічних датасетів з реальним трафіком надає лише непрямі мітки. Наприклад, у наборі SimpleWeb (University of Twente) тривалий трафік кампусної мережі було зібрано без явних міток, а згодом відфільтровано підозрілі підмножини за допомогою honeypot-сервера; отримані мітки "підозрілий трафік" опиралися на спрацьовування пастки, але не гарантували повного охоплення всіх атак [18]. Отже, документування процесу побудови і маркування є критично важливим фактором для наукової цінності датасету.

Склад атак у датасеті визначає, для яких задач IDS він придатний. Історично у наборі DARPA '99/KDD '99 всі атаки поділялися на 4 великі категорії: DoS (відмова в обслуговуванні), Probe/Scan (розвідка, сканування мережі), R2L (атаки, що дозволяють віддаленому користувачу отримати локальні привілеї) та U2R (підвищення привілеїв локального користувача) [9]. Пізніші набори розширювали цей перелік: з'явилися класи для шкідливого ПЗ та ботнетів, веб-атак на прикладному рівні, атаки на протоколи VoIP/Multimedia, цільові APT-атаки, тощо. Зрозуміло, неможливо одним набором охопити всі загрози, але репрезентативний датасет намагається містити різнопланові сценарії – від простих DoS до багатостадійних вторгнень. Важливо, щоб у датасеті були представлені як мережеві атаки, так і нормальний трафік різних типів, причому змішані таким чином, щоб відтворювати реальні умови – наприклад, атаки можуть маскуватися під легітимний трафік, відбуватися паралельно з фоновою активністю користувачів і т.д. Якщо набір даних не містить нормальної складової (є і такі випадки, коли публіку-

ють лише вибірку зловмисного трафіку для аналізу шкідливості), то його доводиться поєднувати з окремим набором звичайного трафіку для повноцінного тестування IDS [5].

### **3. Інструменти генерації трафіку та атак**

Створення реалістичного трафіку для IDS-датасетів потребує використання різноманітних інструментів, що підтримують генерування як фонових навантажень, так і атакуючих впливів. Для симуляції легітимного трафіку застосовуються трафік-генератори або емуляційні платформи. Простими засобами є утиліти на кшталт Iperf), Ostinato, NetFlow Generator, або спеціалізовані інструменти під конкретні протоколи. Наприклад, для VoIP-трафіку існує утиліта SIPp, що програє сценарії дзвінків SIP за заданим шаблоном [19]. Варто зазначити що такі інструменти можуть створювати не-реалістично одноманітний потік (з однаковими інтервалами, розмірами пакетів і т.д.), якщо їх використовувати без належної варіації [5]. Максимального реалізму можна досягти, залучивши до генерування людей: в деяких експериментах дослідники запрошували групу реальних користувачів працювати за комп'ютерами (переглядати веб, надсилати запити, керувати IoT-пристроями) для збору справжнього трафіку з контрольованим профілем дій [20]. Хоча це трудомістко, такий підхід забезпечує найбільш правдоподібні фонові дані.

Для генерації атак зазвичай використовуються наявні інструменти з арсеналу пен-тестингу та кібербезпеки. Щоб симульовані атаки були максимально близькі до реальних, їх виконують ті ж засоби, якими користуються зловмисники. Поширеним вибором є: сканери портів і вразливостей (nmap, Nessus), експлойт-фреймворки (Metasploit), інструменти для DoS/DDoS (наприклад, Low Orbit Ion Cannon, hping3), засоби для brute-force зламування паролів (Hydra, Medusa), генератори шкідливого трафіку (Scapy для кастомних пакетів, SOC traffic generator тощо). Такий підхід гарантує правдиву послідовність пакетів при атаці (відповідну протоколу і методиці вторгнення). Важливо також варіювати параметри атак: змінювати адреси, порти, час доби, інтенсивність, типи пакетів тощо, аби в межах одного набору даних не повторювались ідентичні шаблони зловмисних дій. Різноманітність атак підвищує генералізаційну здатність IDS, яка працює з ними.

Окрім інструментів власне генерування трафіку, у процесі створення датасету застосовуються засоби для захоплення та обробки мережових даних. Стандартом де-факто є утиліта tcpdump [21] або її аналоги (Dumprcap, TShark), які використовуються для запису мережового трафіку в файл pcap. Якщо потрібно отримати зведені flow-записи (агреговані сесії), часто використовують спеціальні монітори – наприклад, системи збору NetFlow/IPFIX або мережеві проби. Після отримання сирих даних великого обсягу їх обробляють флоу-екстракторами на зразок згаданого CICFlowMeter, що автоматично рахує статистики по кожному з'єднанню (кількість пакетів, байтів, тривалість, середні швидкості тощо)[16]. Інший підхід – використання IDS-систем для попереднього аналізу: можна пропустити трафік через Snort/Suricata або Zeek в режимі запису логів, отримавши на виході як детектовані атаки (для перевірки коректності міток), так і зведення по сесіях.

Комбінація різних інструментів – від трафік-генераторів до аналізаторів – є ознакою добре спроектованого процесу побудови датасету, оскільки дозволяє забезпечити і різноманітність даних, і перевірку їх якості перед публікацією.

#### **4. Обмеження існуючих датасетів та напрями покращення**

Незважаючи на значну кількість доступних IDS-датасетів, у літературі наголошується на їх обмеженнях та недоліках [2]. Однією з головних проблем, як згадувалося вище, є старіння. Атаки і мережеві поведінкові патерни швидко еволюціонують, тому набори, зібрані кілька років тому, втрачають актуальність. Моделі, натреновані на таких наборах, ризикують показувати погані результати в реальному середовищі з новими загрозами. Друга проблема – обмежений контекст і різноманітність. Багато датасетів охоплюють лише короткий проміжок часу або одну мережеву ситуацію, що не відбиває повної варіабельності фону. Для надійності бажано, щоб дані містили різні часові періоди (робочі години і вечір, будні і вихідні), різні режими навантаження мережі, різні типи легітимної активності [5].

Ще одним недоліком є неповнота або неточність розмітки в деяких наборах. Як згадувалося, якщо мітки отримані автоматично чи неповні (наприклад, позначені лише категорії атак замість точного переліку всіх атакуючих сесій), то оцінка IDS може бути ненадійною. Наявні роботи вказують на випадки виявлення помилок у відомих датасетах: так, аналіз CICIDS-2017 та CSE-CIC-2018 показав наявність дублікатів, невідповідностей міток часу та інших артефактів, що вимагали ручного чищення даних [22].

Через зазначені проблеми спільнота дедалі більше звертає увагу на покращення методів генерації та валідації датасетів. В нових дослідженнях рекомендується приділяти особливу увагу реалізму: наскільки можливо, включати актуальний реальний трафік (навіть якщо його треба анонімізувати) та сучасні сценарії атак, аби відображати поточний ландшафт загроз [31]. Якщо реальні дані недоступні, слід удосконалювати емуляцію легітимного трафіку – це залишається найскладнішим завданням, оскільки інструменти для генерування атак вже досить наближені до реальних, а от згенерувати різноманітний і реалістичний фон все ще важко [23].

В цілому, огляди літератури сходяться на тому, що проблема датасетів для IDS далека від вирішення, але поступово напрацьовуються принципи їх якісного створення: комбінування реальних і симульованих методів, прозоре документування процесу, забезпечення максимальної реалістичності та повноти, а також вільне розповсюдження з урахуванням етичних і правових норм. Такий підхід сприятиме появі більш надійних і стійких систем виявлення вторгнень у майбутньому.

#### **Викладення основного матеріалу дослідження.**

##### **1. Вимоги до датасету**

Перед генерацією датасету потрібно визначити ключові фактори що будуть задавати параметри даних та підходу до їх створення. У цьому розділі розібрані та аргументовані рішення, прийняті у ході дослідження предметної області.

##### **1.1. Підхід до виділення трафіку**

Як згадувалося у вступі, ця робота є підготовкою до прямої інтеграції нейронної мережі у IDS і ціллю є модель що навчається відтворювати та узагальнювати рішення Snort за його ж подіями. Тож варто пояснити чому події класифіковано саме за FastPattern групами.

На етапі проектування було виділено 3 варіанти прив'язки даних, базуючись саме на подальшій цілі роботи:

- Per-rule (окрема маленька модель на кожне правило).

Плюси: точна локалізація; можна агресивно оптимізувати під конкретні шаблони.

Мінуси: вибух кількості моделей, складність супроводу, ризик «зашумлення» від нестабільності окремих сигнатур; слабе узагальнення між близькими правилами.

- Per-instance/Per-service (одна велика модель на весь інстанс Snort або протокол).

Плюси: специфіку швидких шляхів зіставлення.

- Per Fast Pattern (одна модель на групу правил, що ділять спільний шаблон).

Плюси: добрий баланс між кількістю моделей і узагальненням; природний тригер – спрацювання Fast Pattern після якого можна викликати NN лише для відповідної підмножини трафіку; зберігається локальність семантики.

Мінуси: потребує коректної ідентифікації FP-груп і стабільності їх складу між ревізіями правил.

З огляду на вимоги ефективності, швидкодії і бажання мінімізувати зміни у Snort, ця робота зосереджена на рівні Fast Pattern-групи: NN викликається після FP-match, але до повного обчислення складних умов правила, і працює на вже нормалізованих байтах, які на цьому етапі доступні інспекторам.

### 1.2. Загальні вимоги до IDS-датасетів

Як вже згадувалося раніше, створюваний датасет призначено для тренування та валідації моделей, націлених на вузькі класи мережевих атак. У цьому конкретному випадку ціллю є один клас атак, що поділяють спільний базовий патерн у текстовому буфері.

Таким чином, одиницею спостереження є саме подія виявлення патерну. Так, для не виявлених атак і для benign-трафіку, один запис відповідає факту виявлення підозри (тобто спільного патерну). Для виявлених атак: один запис відповідає факту виявлення атаки. Число результуючих записів хоч і не є дуже важливою мірою, на даному етапі, але запланована кількість не менше 50 000 записів загалом. Множинні спрацювання на один приклад не передбачені; у разі винятків пріоритизується атака або перше виявлення.

Що стосується записів, кожен запис повинен включати наступні дані:

- is\_attack – ground truth(експертне) твердження чи є трафік атакою;
- alerted – чи виявлена атака IDS;
- buffers – набір текстових даних які IDS виділив для пошуку атак;
- buffer\_names – імена буферів даних;

Додаткові прикладні ознаки сесії/пакета (розмір, напрям, часові мітки), за наявності.

Датасет повинен покривати більшість доступних атак у межах однієї Fast Pattern-групи (з потреби, не всіх правил, щоб зберегти простір для валідації), а також велику різноманітність benign-сценаріїв для оцінки хибнопозитивних спрацювань.

Результатом роботи є опубліковані PCAP файли, фінальні JSON-файли, Docker-оточення, а також зафіксовані версії Snort 3 і Talos LightSPD [24], для збереження відтворюваності.

### 1.3. Специфічні вимоги для використання зі Snort

Так як ключові дані будуть виділені напряму з рушія Snort 3, важливо виділити правила до формування датасету та оформлення ознак.

Ключові дані (payload байти) повинні формуватися шляхом конкатенації нормалізованих буферів інспектора у фіксованому порядку (заданому самим інспектором). Довжина запису обмежена розміром 1024 байт із zero-padding для коротших випадків і стор для довших. Ідентифікуючі поля is\_attack та alerted мають бути представлені у вигляді bool флагів.

Щодо нейтральної телеметрії, тут є обмеження зі сторони логерів Snort 3. З документації JSON логеру [25] був виділений наступний набір потенційно важливих ознак пакету/сесії: flowstart\_time, seconds, proto, pkt\_gen, pkt\_len, eth\_len, eth\_type, ip\_id, ip\_len, tos, ttl, udp\_len, dir, client\_bytes, client\_pkts, server\_bytes, server\_pkts. Зазначені поля повинні бути включені до кожного запису.

### 1.4. Приватність, безпека та ліцензування

Оскільки середовище повністю лабораторне, спеціальна анонімізація не вимагається. Використання PCAP слід супроводжувати застереженням про відповідальне застосування.

## 2. Методологія генерації

Постановка задачі робить датасет максимально близьким до умов онлайн-детекції, знімає залежність від конкретних парсерів і дає змогу перевіряти PU-сценарій (P = події, на які спрацював Snort; U = решта подій без спрацювань).

### 2.1. Вибір Fast Pattern-групи

Відповідно до методу структурної класифікації правил IDS[31], вибір Fast Pattern-групи здійснюється шляхом максимізації інтегрального критерію:

$$J(p) = w_1 \cdot coverage(p) + w_2 \cdot selectivity(p) - w_3 \cdot overlap(p, \cdot) - w_4 \cdot churn(p) \quad (1)$$

де  $w_i \geq 0$ ,  $\sum w_i = 1$ .

Критерії відображають:

coverage – структурну репрезентативність групи;

selectivity – дискримінаційну здатність відносно benign-трафіку;

overlap – рівень конкуренції з іншими Fast Pattern;

churn – часову стабільність правил групи.

Покриття визначалося як:

$$coverage(p) = \frac{N_{rules(p)}}{N_{rules}^{total}} \quad (2)$$

де  $N_{rules(p)}$  – кількість правил Talos LightSPD, що містять Fast Pattern P,

$N_{rules}^{total}$  – максимальна кількість правил у вибраній підмножині (30).

Селективність оцінювалась на основі трьох незалежних PCAP-файлів з легітимним трафіком відповідних сервісів.

$$selectivity(p) = 1 - \frac{N_{matches}(p)}{N_{packets}} \quad (3)$$

де  $N_{matches}(p)$  – кількість пакетів, що містять відповідний Fast Pattern,

$N_{packets}$  – кількість пакетів відповідного сервісу у PCAP.

Overlap визначався як середня кількість інших Fast Pattern, що спрацювали на той самий пакет:

$$overlap(p) = \frac{1}{M} \sum_{i=1}^M FP_{other}(packet_i) \quad (4)$$

де  $M$  – кількість спрацювань таргетного Fast Pattern,

$FP_{other}$  – кількість спрацювань інших Fast Pattern з Talos LightSPD.

Churn оцінювався шляхом порівняння двох версій Talos LightSPD з інтервалом у 5 років:

$$churn(p) = \frac{N_{changed}(p)}{N_{rules}(p)} \quad (5)$$

де  $N_{changed}(p)$  – кількість правил, що були змінені,

$N_{rules}(p)$  – загальна кількість правил групи.

Отримані значення були записані у таблицю 1. Важливо зазначити що результати оцінки сильно залежать від трафіку та набору правил, тож можуть відрізнятись.

Таблиця 1

Результати оцінки патернів

Pattern	Coverage	Selectivity	Overlap	Churn
SIP/2.0	1.00	0.74	0.44	0.667
NTLMSSP	0.43	0.86	0.77	0.667
X-MAILER	0.50	0.43	1.00	0.90

За відсутності пріоритетності критеріїв приймемо рівні ваги:

$$w_1 = w_2 = w_3 = w_4 = 1 \quad (6)$$

Тоді:

- SIP/2.0:

$$J = 1.00 + 0.74 - 0.44 - 0.667 = 0.633 \quad (7)$$

- NTLMSSP:

$$J = 0.43 + 0.86 - 0.77 - 0.667 = -0.147 \quad (8)$$

- X-MAILER:

$$J = 0.50 + 0.43 - 1.00 - 0.90 = -0.97 \quad (9)$$

Отримані значення демонструють що група з патерном SIP/2.0 має найбільше значення інтегральної функції якості, тож оберемо її для подальшого аналізу.

## 2.2. Топологія мережі для генерації датасету

Тож, мета дизайну – побудувати кероване, відтворюване середовище прикладного рівня (SIP/2.0) із достатньою варіативністю нормальної поведінки та систематично згенерованими «torture»-аномаліями, з єдиною точкою спостереження для подальшого журналювання IDS.

Компоненти та версії:

- PBX (SIP-сервер): Docker-образ andrius/asterisk[27], Asterisk v22.7.0.
- Клієнтські генератори SIP: Docker-образ sipp, SIPp v3.7.5 (для обох ролей – caller і callee).
- Генератор аномалій: SIP Torch[26] v0.1.0 з модифікаціями для підвищення варіабельності згенерованих атак (30 сценаріїв на основі SIP torture-кейсів[29], інші сценарії вимкнено[28]).
- Захоплення трафіку: tcpdump 4.99.1 (libpcap 1.10.1).

Логічна схема (рисунок 1):

1. Одна приватна Docker-мережа типу bridge; вузли синхронізують час із хостом.
2. PBX (центральний сервер);
3. Caller Farm (64 інстанси sipp);
4. Callee Farm (64 інстанси sipp);
5. Attackers (процеси SIP Torch, що запускаються з боку caller-вузлів).
6. Захоплення трафіку відбувалося на вхідному інтерфейсі сервера, між Caller Farm та сервером.

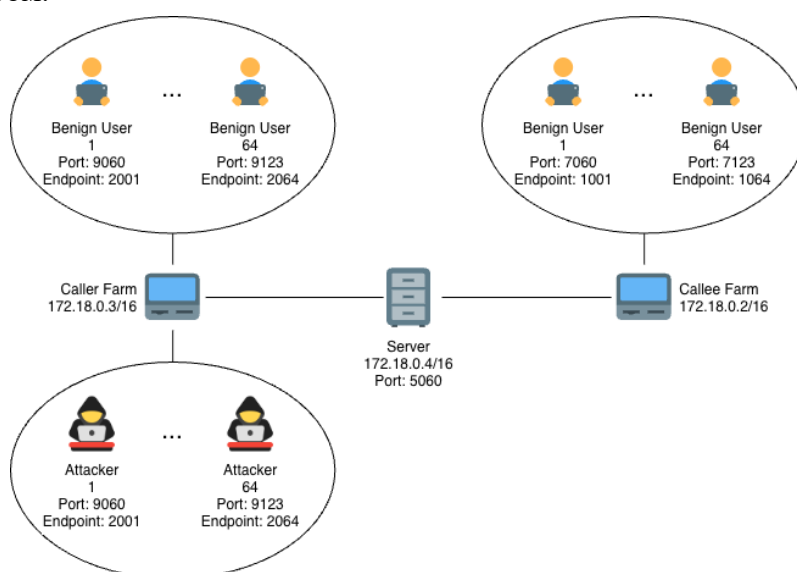


Рисунок 1 – Логічна схема мережі SIP

### 2.3. Безпечний профіль трафіку

За основу сценаріїв було взято існуючий набір sipp-scenarions [29], який надає велику кількість готових xml файлів з підготовленими SIP сесіями. Сесії було категоризовано на клієнтські та серверні а також на позитивні та негативні. Також було значно модифіковано сценарії, для підвищення варіативності даних [30].

Ролі та сценарії.

- Caller-farm (64 інстанси sipp). Кожен інстанс має унікальні облікові дані/endpoint/порт і випадково обирає один із 33 заздалегідь підготовлених сценаріїв: успішні/помилкові виклики, анонімні запити, із/без REGISTER/SUBSCRIBE, із/без медіа, з тайм-аутами.

- Callee-farm (64 інстанси sipp). Для відповіді використовується 23 варіанти профілів (нормальні та помилкові відповіді, з медіа та без, різні коди стану). Профіль обирається випадково незалежно від caller-сценарію.

На кожен спробу виклику caller випадково обирає callee; пари змінюються між ітераціями. Між сценаріями та між запусками одного сценарію варіюються ідентифікатори діалогів (Call-ID), теги, порядок заголовків, поля Contact/From/To, інтервали між повідомленнями, вказівники часу, тд. Це зроблено для зменшення штучної одноманітності даних.

#### 2.4. Атакуючий профіль трафіку

Для генерації атак використано SIP Torch із 33 сценаріями (на основі RFC-кейсів). Сценарії включають: некоректні заголовки/типи вмісту, атипові схеми у Request-URI, відсутніх обов'язкових полів, надмірно великі скалярні значення, тощо.

Використано 64 атакувальні процеси, кожний з яких обирає випадкову ціль із 64 callee та випадковий «torture»-сценарій, генеруючи рандомні мутації полів. Для підвищення варіабельності трафіку було зроблено відповідні зміни у коді генераторів [28]. Це підвищує внутрішню різноманітність позитивних прикладів навіть у межах одного класу сценарію.

Така організація дає багату множину TP/FN залежно від конкретної реалізації сигнатур та їх ревізії, а також створює зони перетину з benign-поведінкою (коли повідомлення виглядають сигнатурно подібними, але є коректними), що критично для подальшого аналізу False Positive/False Negative.

#### 2.5. Передобробка даних

Захоплений трафік було записано у форматі PCAP і оброблено використовуючи Snort 3 з DAQ PCAP. На виході з IDS сформовано подійні JSON-записи. Далі події було розподілено за джерелом PCAP на attack/benign і розбито на підвибірки (train/val/test) для подальшого аналізу. Нижче деталізовано ключові кроки.

##### 2.5.1. Обробка Snort 3

Snort 3 запускався зі стандартним snort.lua, увімкненим alert\_json та профілем правил Talos LightSPD rev. 2026-02-04-001. Для керованості подій застосовано:

- event\_queue: не більш як 1 подія на пакет, і пріоритизація подій атак над суто інформаційними;
- модифікацію alert\_json: додано вивід нормалізованих буферів SIP-інспектора (sip\_header, sip\_body) у поле buffers та їхніх назв у buffer\_names. У якості роздільник було використано ASCII символ Record Separator (0x1E).

Для уніфікації байтового входу застосовано фіксовану довжину  $N = 1024$  байти:

- якщо  $|buf| < 1024$  – zero-padding (0×00) праворуч;
- якщо  $|buf| > 1024$  – tail-crop до 1024 байтів.

Цілеспрямовано активовано підмножину PROTOCOL-VOIP SIP Torture, що входить до FP-групи SIP/2.0. У вивірній конфігурації використовувались такі SID: 51494, 51499, 51501, 51502, 51503, 51504, 51506, 51507, 51508, 51509, 51510, 51511, 51512, 51515, 51744, 51745, 52087. Також було додано одне правило з ідентичним Fast Pattern, як інформаційне, для виявлення спрацювання FP групи.

### 2.5.2. Післяобробка івентів

Результуючі івенти були валідовані та оброблені за допомогою Python скрипта. За наявності дублю подій із тотожним буфером на одному й тому ж пакеті залишено першу (часово ранню) подію. Також було додано is\_attack маркер до кожного івенту.

У телеметрії пакету, строкові дані були зіставлені з цілочисловими значеннями. Мапінги залишено у корені JSON документа, а саме:

- proto\_mapping – для назви протоколів;
- pkt\_gen\_mapping – для назви модуля що подав пакет на обробку;
- eth\_type\_mapping – для Ethernet Type байту;
- dir\_mapping – для напрямку пакету.

### 3. Результат

Сформовано подійний датасет, події розділено за джерелом PCAP на attack/benign і випадково розбито у співвідношенні 75/15/10 (train/val/test). Також, для атак

- Attack (усього): 26182 подій  
Train: 19621 (75%) – Val: 3924 (15%) – Test: 2617 (10%)
- Benign (усього): 47212 подій  
Train: 35633 (75%) – Val: 7163 (15%) – Test: 4416 (10%)

Підсумки детекції, загалом, до сплітів: TP = 4776, FP = 0, TN = 47212, FN = 21386. Дисбаланс у бік benign є очікуваним і відображає реальний контекст експлуатації IDS. Тож, можна бачити, що сигнатури дають високу точність, але пропускають значну частину варіативних torture-кейсів, тож саме сюди добре вписується ML-шар.

Власне набір даних у форматі JSON [32] та середовище для репродукції його запису [33] опубліковані на GitHub.

#### 3.1. Описова статистика і кореляції

Основним вкладом цієї роботи є дані нормалізованих буферів мета цього аналізу - з'ясувати, у яких саме частинах вектора зосереджена інформація, що відрізняє атакуючий та легітимний трафік, а також перевірити, чи не виникає штучний сигнал через padding.

Так на рисунку 2 показано різницю Jensen–Shannon дивіргенції для байтових буферів, тобто наскільки розподіли байтів у кожній позиції відрізняються між класами. Найбільші відмінності зосереджені у верхній частині вектора, що відповідає початку SIP-повідомлення (заголовку та початку тіла). Далі інтенсивність різниці поступово зменшується, а в кінцевій частині буфера практично зникає. Це означає, що дискримінативний сигнал локалізований і пов'язаний зі структурованими елементами протоколу.

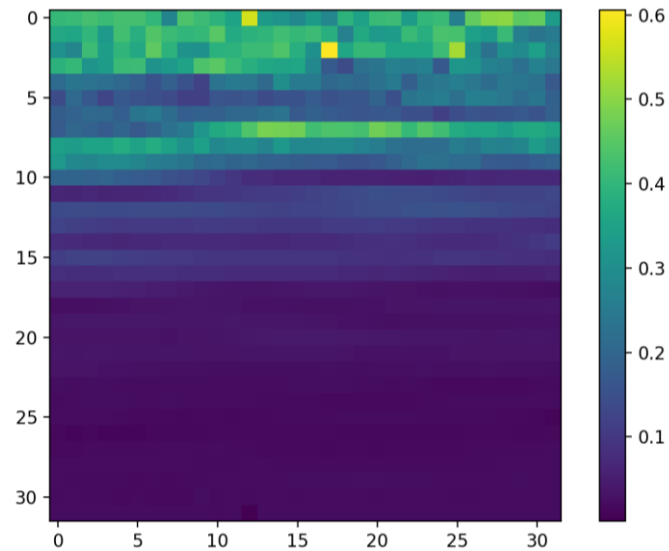


Рисунок 2 – Посимвольна різниця Jensen–Shannon дивергенції

На рисунку 3 наведено різницю варіативності байтів між класами. Видно, що у верхній частині буфера атакуючі зразки демонструють більшу структурну різноманітність, тоді як нижня частина має низьку варіативність в обох класах. Це підтверджує, що інформативними є саме початкові позиції, а не весь 1024-байтний вектор рівномірно.

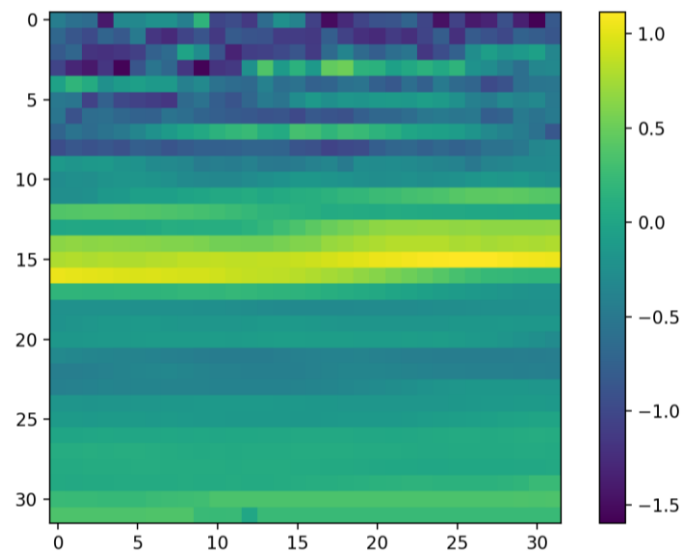


Рисунок 3 – Посимвольна різниця ентропії для буферів

Рисунок 4 показує аналіз частки padding для буферів. Можна побачити що кінцева частина вектора містить велику кількість доповнюючих байтів, однак різниця між класами у цій зоні не є суттєвою. Це важливо, оскільки свідчить про відсутність тривіального ефекту “витоку інформації” через довжину повідомлення: модель не може коректно розділити класи лише за рахунок різної кількості padding.

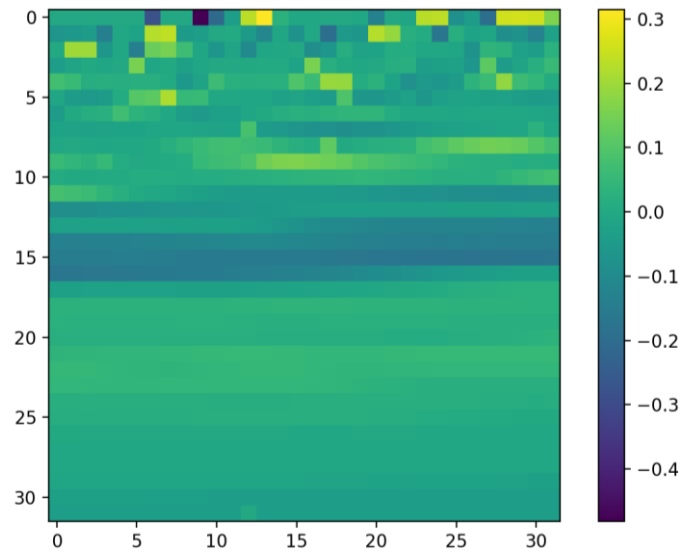


Рисунок 4 – Різниця частки padding байтів

Також було прораховано кореляції Пірсона між фічами з набору телеметрії. Як можна побачити на рисунку 5, найбільш виражені позитивні кореляції спостерігаються між розмірними характеристиками пакетів і байтовими лічильниками, що є очікуваним, оскільки ці параметри описують близькі аспекти однієї й тієї ж структури трафіку.

Водночас більшість службових або протокольних полів демонструють низький рівень кореляції з іншими змінними, що свідчить про їх відносну незалежність і потенційну інформативність у моделі. Відсутність великої кількості сильних кореляцій поза групами розмірних параметрів означає, що суттєвої мультиколінеарності в наборі ознак немає, а отже, фічі несуть комплементарну інформацію.

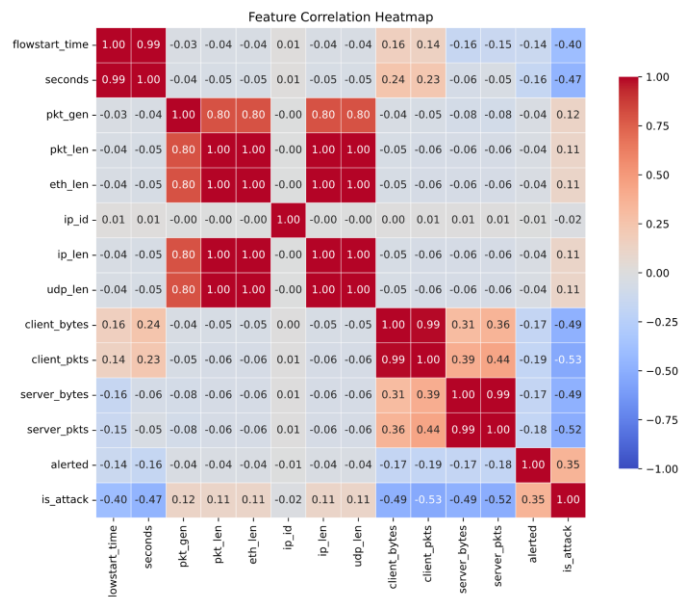


Рисунок 5 – Теплова карта кореляцій телеметрії

### 3.2. Порівняння з іншими датасетами

У таблиці 2 наведено загальне порівняння з іншими існуючими датасетами.

## Порівняння з існуючими датасетами

Параметри	KDD99	UNSW-NB15	CICIDS-2017	HIKARI-2021	Ця робота
Кількість унікальних IP адрес	11	45	16 960	7 991	≈129
Симулятивність	Так	Так	Частково	Частково	Так
Формат даних	dumpfile	pcap	pcap	pcap	pcap, JSON
Категорії атак	4	9	7	4	1
Методи виділення фічей	Bro-IDS	Argus, Bro	CICFlowMeter	Zeek, Python	Snort 3, Python
Кількість фічей	42	49	80	86	18 + 1024 байти

**Висновки.** У роботі запропоновано IDS-центричний підхід до формування датасету подій, у якому вхідними даними є нормалізовані байтові буфери інспектора Snort 3 та телеметрія пакета/сесії, доступні під час реальної онлайн-обробки. Ground truth задається контрольованим походженням трафіку (attack/benign PCAP), тоді як спрацювання правил Snort використовується як teacher-сигнал для сценарію “Snort-як-учитель”.

Сформовано датасет для Fast Pattern-групи SIP/2.0 обсягом 73 394 події (26 182 attack / 47 212 benign) зі сплітом 75/15/10 (train/val/test). Базова детекція Snort на цьому зрізі характеризується високою точністю (FP = 0) та обмеженою повнотою (Recall ≈ 0.183), що узгоджується з природою SIP-torture варіативності та обраною підмножиною сигнатур.

Аналіз байтових позицій (Jensen–Shannon дивіргенція, ентропія) показав, що дискримінативний сигнал локалізується переважно на початку повідомлення (заголовок/початок тіла), а padding-зона не створює тривіального “витоку” інформації між класами. Кореляційний аналіз телеметрії підтвердив очікувані кластери “довжин/обсягів” та загалом низьку мультиколінеарність інших полів.

Обмеження роботи: лабораторність середовища, фокус на одній FP-групі та конкретній ревізії правил, фіксована довжина байтового представлення (stop/pad), а також шумність teacher-міток, властива сигнатурному підходу. Подальші напрями включають розширення на інші FP-групи/протоколи, перевірку узагальнюваності на нових ревізіях правил і введення сценарних сплітів (scenario-holdout) для більш строгого оцінювання. Отриманий набір даних є основою для навчання моделей risk scoring/ранжування в режимі, сумісному з онлайн-детекцією IDS.

## ЛІТЕРАТУРА

1. Горбатов В. С. Метод попередньої фільтрації сигнатур для прискорення пошуку атак системою виявлення мережевих вторгнень. Сучасні інформаційні та комунікаційні технології на транспорті, в промисловості та освіті: Міжнар. науково-практ. конф., м. Дніпро, 13–14 груд. 2023 р. Дніпро, 2023. С. 136–137.
2. Kenyon A., Deka L., Elizondo D. Are public intrusion datasets fit for purpose

characterising the state of the art in intrusion event datasets. *Computers & security*. 2020. Vol. 99. P. 102022. URL: <https://doi.org/10.1016/j.cose.2020.102022> (date of access: 21.02.2026).

3. Generating network intrusion detection dataset based on real and encrypted synthetic attack traffic / A. Ferriyan et al. *Applied sciences*. 2021. Vol. 11, no. 17. P. 7868. URL: <https://doi.org/10.3390/app11177868> (date of access: 21.02.2026).

4. Survey of intrusion detection systems: techniques, datasets and challenges / A. Khraisat et al. *Cybersecurity*. 2019. Vol. 2, no. 1. URL: <https://doi.org/10.1186/s42400-019-0038-7> (date of access: 21.02.2026).

5. Goldschmidt P., Chudá D. Network intrusion datasets: a survey, limitations, and recommendations. *Computers & security*. 2025. P. 104510.

URL: <https://doi.org/10.1016/j.cose.2025.104510> (date of access: 21.02.2026).

6. Himabindu C. The challenges of effectively anonymizing network data. *International journal of pharmacology and pharmaceutical technology*. 2013. P. 15–22.

URL: <https://doi.org/10.47893/ijppt.2013.1003> (date of access: 21.02.2026).

7. Kim J., Sim C., Choi J. Generating labeled flow data from MAWILab traces for network intrusion detection. *Proceedings of the ACM workshop on systems and network telemetry and analytics*. Phoenix, AZ, USA, 2019. P. 45–48.

8. Statistical analysis of honeypot data and building of Kyoto 2006+ dataset for NIDS evaluation / J. Song et al. *The first workshop, Salzburg, Austria, 10 April 2011*. New York, New York, USA, 2011. URL: <https://doi.org/10.1145/1978672.1978676> (date of access: 21.02.2026).

9. Evaluating intrusion detection systems: the 1998 DARPA off-line intrusion detection evaluation / R. P. Lippmann et al. *DARPA information survivability conference and exposition. DISCEX'00*, Hilton Head, SC, USA.

URL: <https://doi.org/10.1109/discex.2000.821506> (date of access: 21.02.2026).

10. A detailed analysis of the KDD CUP 99 data set / M. Tavallaee et al. *2009 IEEE symposium on computational intelligence for security and defense applications (CISDA)*, Ottawa, ON, Canada, 8–10 July 2009. 2009.

URL: <https://doi.org/10.1109/cisda.2009.5356528> (date of access: 21.02.2026).

11. Toward developing a systematic approach to generate benchmark datasets for intrusion detection / A. Shiravi et al. *Computers & security*. 2012. Vol. 31, no. 3. P. 357–374.

URL: <https://doi.org/10.1016/j.cose.2011.12.012> (date of access: 21.02.2026).

12. Sharafaldin I., Habibi Lashkari A., Ghorbani A. A. Toward generating a new intrusion detection dataset and intrusion traffic characterization. *4th international conference on information systems security and privacy*, Funchal, Madeira, Portugal, 22–24 January 2018. 2018. URL: <https://doi.org/10.5220/0006639801080116> (date of access: 21.02.2026).

13. Moustafa N., Slay J. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). *2015 military communications and information systems conference (milcis)*, Canberra, Australia, 10–12 November 2015. 2015. URL: <https://doi.org/10.1109/milcis.2015.7348942> (date of access: 21.02.2026).

14. UGR'16: A new dataset for the evaluation of cyclostationarity-based network IDSs / G. Maciá-Fernández et al. *Computers & security*. 2018. Vol. 73. P. 411–424. URL: <https://doi.org/10.1016/j.cose.2017.11.004> (date of access: 21.02.2026).

15. Flow-based network traffic generation using Generative Adversarial Networks / M. Ring

- et al. Computers & security. 2019. Vol. 82. P. 156–172.  
URL: <https://doi.org/10.1016/j.cose.2018.12.012> (date of access: 21.02.2026).
16. GitHub - ahlashkari/cicflowmeter: cicflowmeter-v4.0 (formerly known as iscxflowmeter) is an ethernet traffic bi-flow generator and analyzer for anomaly detection that has been used in many cybersecurity datasets such as android adware-general malware dataset (CIC-AAGM2017), IPS/IDS dataset (CICIDS2017), android malware dataset (cicandmal2017) and distributed denial of service (cicddos2019). GitHub.  
URL: <https://github.com/ahlashkari/CICFlowMeter> (date of access: 21.02.2026).
17. The zeek network security monitor. Zeek. URL: <https://zeek.org> (date of access: 21.02.2026).
18. Simpleweb/University of twente traffic traces data repository / R.R.R. Barbosa et al. CTIT technical report series. 2010.  
URL: <https://api.semanticscholar.org/CorpusID:13251179>.
19. Welcome to SIPp. Welcome to SIPp. URL: <https://sipp.sourceforge.net/> (date of access: 21.02.2026).
20. CUPID: A labeled dataset with Pentesting for evaluation of network intrusion detection / H. Lawrence et al. Journal of systems architecture. 2022. P. 102621.  
URL: <https://doi.org/10.1016/j.sysarc.2022.102621> (date of access: 21.02.2026).
21. Home | TCPDUMP & LIBPCAP. Home | TCPDUMP & LIBPCAP.  
URL: <https://www.tcpdump.org/> (date of access: 21.02.2026).
22. Engelen G., Rimmer V., Joosen W. Troubleshooting an intrusion detection dataset: the CICIDS2017 case study. 2021 IEEE security and privacy workshops (SPW), San Francisco, CA, USA, 27 May 2021. 2021. URL: <https://doi.org/10.1109/spw53761.2021.00009> (date of access: 21.02.2026).
23. Layeghy S., Gallagher M., Portmann M. Benchmarking the benchmark – Comparing synthetic and real-world Network IDS datasets. Journal of information security and applications. 2024. Vol. 80. P. 103689. URL: <https://doi.org/10.1016/j.jisa.2023.103689> (date of access: 21.02.2026).
24. Soft Release: lightSPD, the new rules package for Snort 3. Snort Blog. URL: <https://blog.snort.org/2020/12/soft-release-lightspd-new-rules-package.html> (date of access: 21.02.2026).
25. Snort 3 reference manual. GitHub.  
URL: [https://github.com/snort3/snort3/releases/download/3.10.2.0/snort\\_reference.html](https://github.com/snort3/snort3/releases/download/3.10.2.0/snort_reference.html) (date of access: 21.02.2026).
26. GitHub - 0xinfection/siptorch: A "SIP torture" (RFC 4475) testing framework. GitHub.  
URL: <https://github.com/0xInfection/SIPTorch> (date of access: 21.02.2026).
27. GitHub - andrius/asterisk: asterisk PBX in docker – smallest asterisk ever!. GitHub.  
URL: <https://github.com/andrius/asterisk> (date of access: 21.02.2026).
28. Comparing master...shuffle · VytalyGorbatov/SIPTorch. GitHub. URL: <https://github.com/VytalyGorbatov/SIPTorch/compare/master...VytalyGorbatov:SIPTorch:shuffle> (date of access: 23.02.2026).
29. saghul/sipp-scenarios: SIPp scenarios I use for testing SIP stuff. GitHub. URL: <https://github.com/saghul/sipp-scenarios> (date of access: 23.02.2026).
30. Comparing master...mutability · VytalyGorbatov/sipp-scenarios. GitHub. URL: <https://github.com/VytalyGorbatov/sipp-scenarios/compare/master..VytalyGorbatov:sipp-scenarios>

scenarios:mutability (date of access: 23.02.2026).

31. Горбатов В.С., Журба А.О. Метод класифікації мережеских вторгнень на основі структури правил систем виявлення вторгнень // Інформаційні технології і автоматизація – 2025 / Матеріали XVIII міжнародної науково-практичної конференції. Одеса, 30-31 жовтня 2025 р. - Одеса, Видавництво ОНТУ, 2025 р. – С. 261-263.

32. GitHub - VytalyGorbatov/sip-dataset. GitHub.

URL: <https://github.com/VytalyGorbatov/sip-dataset> (date of access: 23.02.2026).

33. GitHub - VytalyGorbatov/sip-lab. GitHub. URL: <https://github.com/VytalyGorbatov/sip-lab> (date of access: 23.02.2026).

## REFERENCES

1. Horbatov, V. S. (2023). Signature pre-filtering method for accelerating attack search by network intrusion detection systems. Modern Information and Communication Technologies in Transport, Industry, and Education: International Scientific and Practical Conference, Dnipro, December 13–14, 2023. Dnipro, pp. 136–137.

2. Kenyon A., Deka L., Elizondo D. Are public intrusion datasets fit for purpose characterising the state of the art in intrusion event datasets. Computers & security. 2020. Vol. 99. P. 102022. URL: <https://doi.org/10.1016/j.cose.2020.102022> (date of access: 21.02.2026).

3. Generating network intrusion detection dataset based on real and encrypted synthetic attack traffic / A. Ferriyan et al. Applied sciences. 2021. Vol. 11, no. 17. P. 7868. URL: <https://doi.org/10.3390/app11177868> (date of access: 21.02.2026).

4. Survey of intrusion detection systems: techniques, datasets and challenges / A. Khraisat et al. Cybersecurity. 2019. Vol. 2, no. 1. URL: <https://doi.org/10.1186/s42400-019-0038-7> (date of access: 21.02.2026).

5. Goldschmidt P., Chudá D. Network intrusion datasets: a survey, limitations, and recommendations. Computers & security. 2025. P. 104510  
URL: <https://doi.org/10.1016/j.cose.2025.104510> (date of access: 21.02.2026).

6. Himabindu C. The challenges of effectively anonymizing network data. International journal of pharmacology and pharmaceutical technology. 2013. P. 15–22.  
URL: <https://doi.org/10.47893/ijppt.2013.1003> (date of access: 21.02.2026).

7. Kim J., Sim C., Choi J. Generating labeled flow data from MAWILab traces for network intrusion detection. Proceedings of the ACM workshop on systems and network telemetry and analytics. Phoenix, AZ, USA, 2019. P. 45–48.

8. Statistical analysis of honeypot data and building of Kyoto 2006+ dataset for NIDS evaluation / J. Song et al. The first workshop, Salzburg, Austria, 10 April 2011. New York, New York, USA, 2011. URL: <https://doi.org/10.1145/1978672.1978676> (date of access: 21.02.2026).

9. Evaluating intrusion detection systems: the 1998 DARPA off-line intrusion detection evaluation / R. P. Lippmann et al. DARPA information survivability conference and exposition. DISCEX'00, Hilton Head, SC, USA. URL: <https://doi.org/10.1109/discex.2000.821506> (date of access: 21.02.2026).

10. A detailed analysis of the KDD CUP 99 data set / M. Tavallae et al. 2009 IEEE symposium on computational intelligence for security and defense applications (CISDA), Ottawa, ON, Canada, 8–10 July 2009. 2009. URL: <https://doi.org/10.1109/cisda.2009.5356528> (date of access: 21.02.2026).

11. Toward developing a systematic approach to generate benchmark datasets for intrusion

- detection / A. Shiravi et al. *Computers & security*. 2012. Vol. 31, no. 3. P. 357–374. URL: <https://doi.org/10.1016/j.cose.2011.12.012> (date of access: 21.02.2026).
12. Sharafaldin I., Habibi Lashkari A., Ghorbani A. A. Toward generating a new intrusion detection dataset and intrusion traffic characterization. 4th international conference on information systems security and privacy, Funchal, Madeira, Portugal, 22–24 January 2018. 2018. URL: <https://doi.org/10.5220/0006639801080116> (date of access: 21.02.2026).
13. Moustafa N., Slay J. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). 2015 military communications and information systems conference (milcis), Canberra, Australia, 10–12 November 2015. 2015. URL: <https://doi.org/10.1109/milcis.2015.7348942> (date of access: 21.02.2026).
14. UGR'16: A new dataset for the evaluation of cyclostationarity-based network IDSs / G. Maciá-Fernández et al. *Computers & security*. 2018. Vol. 73. P. 411–424. URL: <https://doi.org/10.1016/j.cose.2017.11.004> (date of access: 21.02.2026).
15. Flow-based network traffic generation using Generative Adversarial Networks / M. Ring et al. *Computers & security*. 2019. Vol. 82. P. 156–172. URL: <https://doi.org/10.1016/j.cose.2018.12.012> (date of access: 21.02.2026).
16. GitHub - ahlashkari/cicflowmeter: cicflowmeter-v4.0 (formerly known as iscxflowmeter) is an ethernet traffic bi-flow generator and analyzer for anomaly detection that has been used in many cybersecurity datasets such as android adware-general malware dataset (CIC-AAGM2017), IPS/IDS dataset (CICIDS2017), android malware dataset (cicandmal2017) and distributed denial of service (cicddos2019). GitHub. URL: <https://github.com/ahlashkari/CICFlowMeter> (date of access: 21.02.2026).
17. The zeek network security monitor. Zeek. URL: <https://zeek.org> (date of access: 21.02.2026).
18. Simpleweb/University of twente traffic traces data repository / R. R. R. Barbosa et al. CTIT technical report series. 2010. URL: <https://api.semanticscholar.org/CorpusID:13251179>.
19. Welcome to SIPP. Welcome to SIPP. URL: <https://sipp.sourceforge.net/> (date of access: 21.02.2026).
20. CUPID: A labeled dataset with Pentesting for evaluation of network intrusion detection / H. Lawrence et al. *Journal of systems architecture*. 2022. P. 102621. URL: <https://doi.org/10.1016/j.sysarc.2022.102621> (date of access: 21.02.2026).
21. Home | TCPDUMP & LIBPCAP. Home | TCPDUMP & LIBPCAP. URL: <https://www.tcpdump.org/> (date of access: 21.02.2026).
22. Engelen G., Rimmer V., Joosen W. Troubleshooting an intrusion detection dataset: the CICIDS2017 case study. 2021 IEEE security and privacy workshops (SPW), San Francisco, CA, USA, 27 May 2021. 2021. URL: <https://doi.org/10.1109/spw53761.2021.00009> (date of access: 21.02.2026).
23. Layeghy S., Gallagher M., Portmann M. Benchmarking the benchmark – Comparing synthetic and real-world Network IDS datasets. *Journal of information security and applications*. 2024. Vol. 80. P. 103689. URL: <https://doi.org/10.1016/j.jisa.2023.103689> (date of access: 21.02.2026).
24. Soft Release: lightSPD, the new rules package for Snort 3. Snort Blog. URL: <https://blog.snort.org/2020/12/soft-release-lightspd-new-rules-package.html> (date of access: 21.02.2026).
25. Snort 3 reference manual. GitHub.

- URL: [https://github.com/snort3/snort3/releases/download/3.10.2.0/snort\\_reference.html](https://github.com/snort3/snort3/releases/download/3.10.2.0/snort_reference.html) (date of access: 21.02.2026).
26. GitHub - 0xinfection/siptorch: A "SIP torture" (RFC 4475) testing framework. GitHub. URL: <https://github.com/0xInfection/SIPTorch> (date of access: 21.02.2026).
27. GitHub - andrius/asterisk: asterisk PBX in docker – smallest asterisk ever!. GitHub. URL: <https://github.com/andrius/asterisk> (date of access: 21.02.2026).
28. Comparing master...shuffle VytalyGorbatov/SIPTorch. GitHub. URL: <https://github.com/VytalyGorbatov/SIPTorch/compare/master...VytalyGorbatov:SIPTorch:shuffle> (date of access: 23.02.2026).
29. saghul/sipp-scenarios: SIPP scenarios I use for testing SIP stuff. GitHub. URL: <https://github.com/saghul/sipp-scenarios> (date of access: 23.02.2026).
30. Comparing master...mutability VytalyGorbatov/sipp-scenarios. GitHub. URL: <https://github.com/VytalyGorbatov/sipp-scenarios/compare/master...VytalyGorbatov:sipp-scenarios:mutability> (date of access: 23.02.2026).
31. Horbatov, V. S., Zhurba, A. O. (2025). Network intrusion classification method based on the rule structure of intrusion detection systems. Information Technologies and Automation – 2025: Proceedings of the XVIII International Scientific and Practical Conference, Odesa, October 30–31, 2025. Odesa: ONT University Publishing House, pp. 261–263.
32. GitHub - VytalyGorbatov/sip-dataset. GitHub. URL: <https://github.com/VytalyGorbatov/sip-dataset> (date of access: 23.02.2026).
33. GitHub - VytalyGorbatov/sip-lab. GitHub. URL: <https://github.com/VytalyGorbatov/sip-lab> (date of access: 23.02.2026).

Received 30.03.2026.  
Accepted 01.04.2026.  
Published 30.04.2026

### ***Towards creating a dataset based on events classified by a network intrusion detection system***

*This paper presents an approach to building a dataset for training machine-learning models in the context of the Snort 3 network intrusion detection system. Unlike conventional NIDS datasets, the proposed dataset is constructed from normalized inspector byte buffers and lightweight packet telemetry that are available during real-time traffic processing. Ground truth is defined by the controlled origin of traffic (attack/benign PCAP), while Snort rule triggers are treated as a “teacher” signal to support subsequent risk-scoring models. The dataset is generated for the SIP/2.0 Fast Pattern group and contains tens of thousands of events with a standardized train/validation/test split. In addition, we analyze byte-position informativeness using Jensen–Shannon divergence and entropy, and perform correlation analysis of telemetry features. The results indicate that the discriminative signal is largely localized in the early parts of the message (header and initial payload) and that padding does not introduce trivial information leakage between classes. The resulting dataset can serve as a foundation for real-time neural models that complement signature-based detection with probabilistic risk assessment.*

*Keywords: Snort 3, NIDS, dataset, SIP/2.0, Fast Pattern, byte-level features, packet telemetry, entropy, risk scoring, real-time neural networks, precision/recall.*

**Горбатов Віталій Сергійович** - аспірант кафедри Інформаційних технологій і систем ННІ ДМетІ УДУНТ.

ORCID: <https://orcid.org/0009-0000-9061-8207>

**Журба Анна Олексіївна** – к.т.н., доцент, доцент кафедри Інформаційних технологій і систем ННІ ДМетІ УДУНТ.

ORCID: <https://orcid.org/0000-0002-4367-385X>

**Gorbatov Vitalii** - post-graduate student, Department of information technology and systems, NNI DMetI, Ukrainian state university of science and technologies.

ORCID: <https://orcid.org/0009-0000-9061-8207>

**Zhurba Anna** - assistant professor, Department of information technology and systems NNI DMetI, Ukrainian state university of science and technologies.

ORCID: <https://orcid.org/0000-0002-4367-385X>

В.О. Носов, К.Ю. Островська

## АВТОМАТИЗОВАНИЙ КОНВЕЄР ФОРМУВАННЯ ДАТАСЕТУ ДЛЯ НАВЧАННЯ МОДЕЛЕЙ ВІЯВЛЕННЯ ШАХРАЙСТВА

*Анотація.* У дослідженні розглянуто проблему підготовки тренувальних даних для систем виявлення шахрайства в транзакціях електронної комерції на основі методів машинного навчання. За результатами аналізу існуючих відкритих джерел обґрунтовано необхідність створення спеціалізованого набору даних. Запропоновано автоматизований конвеєр об'єднання трьох відкритих наборів даних з платформи Kaggle (IEEE-CIS, Credit Card Transactions Fraud Detection Dataset, Fraudulent E-Commerce) зі збереженням реальних міток шахрайства та збагаченням записів синтетичними атрибутами, адаптованими до специфіки українського платіжного ринку. Опрацьовано методи рівномірної нормалізації часових міток, генерації автентифікаційних даних та розбиття на платіжні системи, формування агрегованих профілів клієнтів та пар для навчання моделі IP Insights. Результатом є набір із 500000 транзакцій за 24 місяці з рівнем шахрайства 3.04%, призначений для навчання конвеєра моделей, до яких входять LightGBM, автоенкодер та IP Insights.

*Ключові слова:* датасет, машинне навчання, транзакція, електронна комерція, LightGBM, автоенкодер, IP Insights, синтетичні дані

**Постановка проблеми.** Стрімкий розвиток електронної комерції як одного з ключових секторів глобальної економіки неминуче супроводжується пропорційним зростанням масштабів платіжного шахрайства. Відповідно до спільного звіту [1] European Banking Authority (EBA) та European Central Bank (ECB), загальний обсяг транзакційного шахрайства в Європейському економічному просторі за 2024 рік сягнув 4.2 млрд євро, продемонструвавши зростання на 17% порівняно з попереднім роком. Зокрема, втрати від шахрайства з платіжними картками склали 1.3 млрд євро (зростання на 29%), причому переважна більшість таких інцидентів припадає саме на онлайн-операції (80–85% за обсягом).

Шахрайські схеми дедалі частіше експлуатують винятки з правил обов'язкової автентифікації клієнта (Strong Customer Authentication) та застосовують методи соціальної інженерії щодо користувачів. Масштабність проблеми підтверджується результатами глобального опитування [2], проведеного компанією Visa спільно з Merchant Risk Council – 98% опитаних представників роздрібною торгівлі стикалися щонайменше з одним типом шахрайства протягом року. При цьому понад 80% торговців відзначають суттєві труднощі в ефективному використанні накопичених даних для підвищення точ-

ності систем машинного навчання (МН), через що вдосконалення цих інструментів є однією з пріоритетних задач галузі.

Оскільки ефективність моделей МН для задач класифікації та виявлення аномалій критично залежить від якості та репрезентативності тренувальної вибірки, виникає науково-практична проблема доступу до релевантних даних. З огляду на сувору конфіденційність реальних транзакційних даних фінансових установ, для навчання моделей переважно використовуються відкриті набори даних, які часто характеризуються обмеженою схемою атрибутів.

**Аналіз останніх досліджень і публікацій.** У попередній роботі [3] було запропоновано концепцію багатопарової системи виявлення шахрайства в онлайн-транзакціях з використанням конвеєра моделей: класифікатора LightGBM, автоенкодера для виявлення аномалій та алгоритму IP Insights для аналізу мережевої поведінки. Навчання такого конвеєра потребує спеціалізованого набору даних, який одночасно містить транзакційні атрибути для класифікатора, поведінкові профілі клієнтів для автоенкодера та пари «entity–IP» для IP Insights. Саме тому якість та структура тренувальних даних є ключовим фактором успішності всієї системи.

Проблематиці підготовки та стандартизації даних для систем виявлення шахрайства присвячено ряд досліджень. У роботі «Fraud Dataset Benchmark and Applications» [4] систематизовано наявні відкриті датасети та запропоновано єдину методологію їх оцінювання, включаючи стандартизовані тренувально-тестові розбиття та спільні домовленості іменування ознак. Дослідники підкреслюють, що датасети для виявлення шахрайства мають специфічні властивості, які відрізняють їх від інших табличних еталонних наборів даних: екстремальний дисбаланс класів (частка шахрайства може сягати 0.01%), висока кардинальність ознак (IP-адреси, номери телефонів), змагальний характер задачі (шахраї адаптують поведінку) та критична роль конструювання ознак (feature engineering) на основі агрегатів та збагачення даних.

В іншому дослідженні [5] розкрили проблему генерації синтетичних даних для систем виявлення шахрайства з використанням дифузійних моделей. Автори продемонстрували, що синтетичне збагачення датасетів, зокрема набору IEEE-CIS Fraud Detection [6], суттєво покращує якість класифікації на незбалансованих даних. Ця робота підтверджує перспективність підходу синтетичної генерації атрибутів, який застосовано у даному дослідженні для збагачення записів полями протоколу EMV 3D-Secure (3DS2) [12].

Аналіз літератури виявив, що більшість публічних наборів мають суттєві обмеження, серед яких анонімізовані ознаки без семантичного контексту, відсутність інформації про пристрій або середовище, орієнтацію на ринки окремих країн. Це обумовило необхідність розробки спеціалізованого датасету, що стало предметом даного дослідження.

**Мета дослідження.** Метою дослідження є створення спеціалізованого набору даних транзакцій електронної комерції шляхом інтеграції та синтетичного збагачення відкритих джерел. Розроблений датасет адаптовано до специфіки українського платіжного ринку, структуровано відповідно до стандарту 3DS2 та призначено для нав-

чання багат шарової антишахрайської системи на основі конвеєра моделей LightGBM, автоенкодера та IP Insights.

### **Викладення основного матеріалу дослідження.**

**Вибір та характеристика вихідних наборів даних.** Формування датасету базується на об'єднанні трьох відкритих масивів з платформи Kaggle, відібраних за критеріями наявності реальних або реалістичних міток шахрайства, достатнього обсягу та різноманітності атрибутів:

- IEEE-CIS Fraud Detection (IEEE-CIS) [6] (~590 тис. транзакцій, рівень шахрайства ~3.5%) – це датасет з реальними мітками шахрайства, створений у межах змагання Kaggle за участю Vesta Corporation. Складається з основної таблиці (суми, домени пошти, 339 анонімізованих ознак) та ідентифікаційної (браузер, пристрій). Головною перевагою масиву є наявність реальних даних про цифрові відбитки пристроїв (user agent). Серед суттєвих обмежень: відсутність абсолютних дат (лише зміщення у секундах), глибока анонімізація більшості ознак, відсутність інформації про торговця, орієнтація на ринок США та висока частка пропусків (приблизно 75%) у даних ідентифікації;

- Credit Card Transactions Fraud Detection Dataset (CCTFDD) [7] (~1.3 млн транзакцій, рівень шахрайства ~0.6%) – синтетичний набір, згенерований за допомогою інструменту Sparkov Data Generation [8] на основі статистичних розподілів реальних транзакцій. Його сильними сторонами є великий обсяг, наявність повних часових міток (2019–2020) та псевдоанонімізованих (хешованих) номерів карток. Обмеження включають синтетичне походження, орієнтацію на ринок США та повну відсутність автентифікаційних даних, інформації про пристрій та браузер;

- Fraudulent E-Commerce Transactions (FECT) [9] (~100 тис. записів, рівень шахрайства ~10%) – це датасет, сфокусований безпосередньо на e-commerce сценаріях. Ключова цінність цього масиву полягає у підвищеному рівні шахрайства, що забезпечує значну кількість позитивних прикладів для збалансування загальної вибірки. Обмеження: відносно невеликий обсяг, нестандартні назви колонок, відсутність інформації про пристрій і параметри автентифікації.

Об'єднання трьох наборів дозволяє компенсувати їхні індивідуальні недоліки, оскільки IEEE-CIS забезпечує реальні дані пристроїв, CCTFDD великий обсяг з повними датами, а FECT підвищену частку позитивних прикладів. Ключовою перевагою є збереження автентичних міток класів та оригінальних сум з усіх джерел.

Проведений аналіз альтернативних відкритих масивів даних довів їхню методологічну невідповідність завданням даного дослідження. Наприклад, датасет IBM AMLSim [10] орієнтований на задачі виявлення відмивання коштів, тоді як широківідомий Credit Card Fraud Dataset [11] містить виключно 28 анонімізованих PCA-компонент, що повністю позбавляє дані семантичного контексту.

**Конвеєр уніфікації даних.** Оскільки кожне джерело має власний формат, набір колонок та способи кодування, розроблено автоматизований конвеєр обробки даних

(див. рис. 1). Конвеєр реалізовано як скрипт з окремими функціями-завантажувачами для кожного джерела.

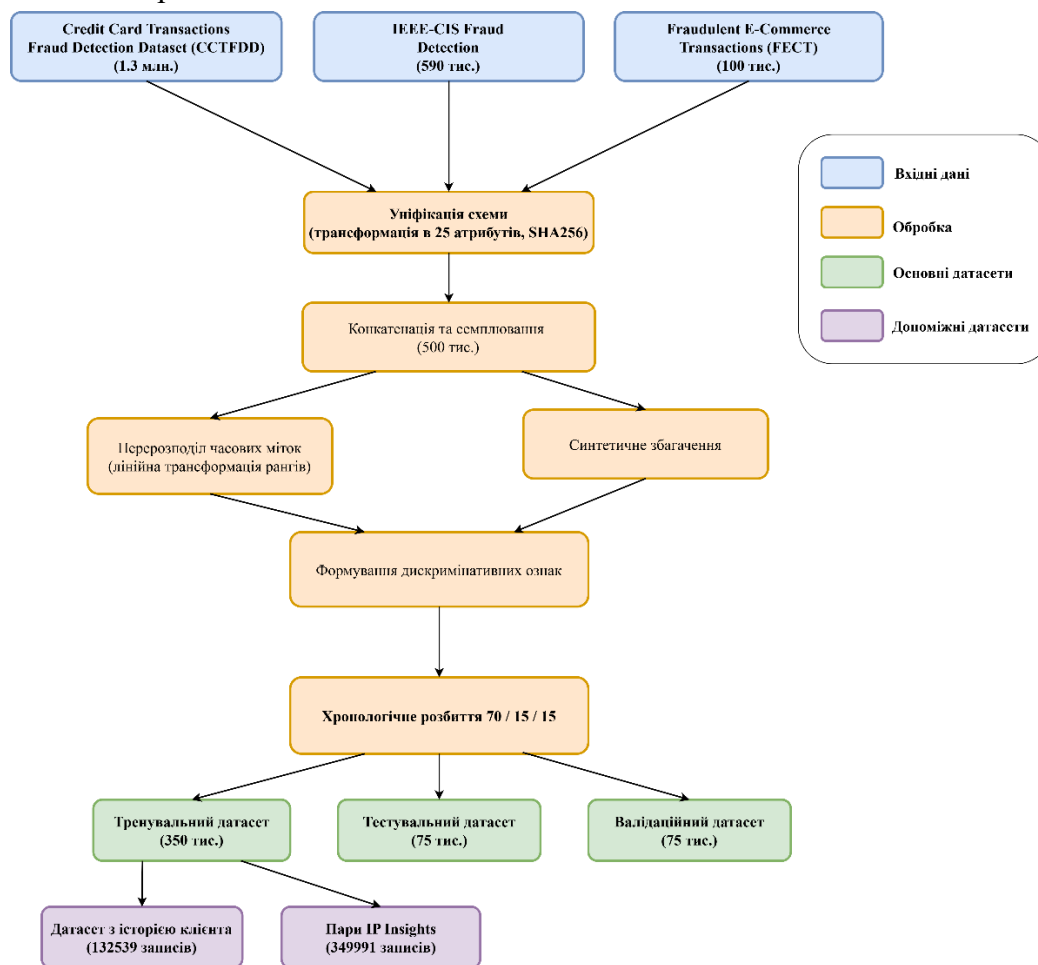


Рисунок 1 – Схема конвеєра уніфікації та злиття даних

Загальна послідовність етапів:

- завантаження та зіставлення полів кожного набору до єдиної цільової структури з 25 атрибутів;
- псевдоанонімізація ідентифікаторів карток через SHA256 хешування;
- конкатенація трьох уніфікованих наборів;
- стратифіковане семплювання до 500000 записів (метод відбору, що зберігає пропорційний розподіл кожного джерела та співвідношення шахрайських і легітимних транзакцій у вибірці);
  - рівномірний перерозподіл часових міток;
  - синтетичне збагачення атрибутами 3DS2 [12];
  - диференціація розподілів (закладення залежностей між ознаками та міткою шахрайства);
  - хронологічне розбиття та формування допоміжних наборів.

**Рівномірний перерозподіл часових міток.** Вихідні масиви охоплюють різні часові діапазони (CCTFDD: 2019–2020, IEEE-CIS: зміщення в секундах, FECT: довільний період) та мають суттєво різну кількість записів. Просте лінійне масштабування чи

пряме об'єднання призвело б до штучної кластеризації даних – в одні місяці домінували б транзакції з одного джерела, в інші з іншого, що викликало б нерелевантні зміщення рівня шахрайства (dataset shift) та перенавчання моделей на часові артефакти.

Для забезпечення рівномірної щільності транзакцій у цільовому часовому діапазоні (01.01.2023–31.12.2024) застосовано метод нормалізації на основі лінійного відображення порядкових рангів (квантильного перетворення). Після злиття всіх масивів транзакції сортуються за оригінальними часовими мітками (зберігаючи їхній відносний хронологічний порядок), після чого кожній транзакції з рангом  $i$  серед  $N$  записів присвоюється нова дата за формулою:

$$t'(i) = t_{start} + \frac{i}{N} \times (t_{end} - t_{start}), \quad (1)$$

де  $t'(i)$  – нове значення дати для транзакції з порядковим рангом  $i$ ;

$t_{start}$  – початкова дата цільового часового діапазону;

$t_{end}$  – кінцева дата цільового часового діапазону;

$i$  – порядковий номер (ранг) транзакції у масиві після сортування за оригінальним часом ( $0 \leq i \leq N$ );

$N$  – загальна кількість записів у об'єднаному наборі даних.

Хоча такий підхід нівелює глобальні макроекономічні сплески (наприклад, святкові розпродажі), він гарантує ідеально збалансоване тренувальне середовище. При цьому оригінальний час доби строго зберігається, що дозволяє моделям засвоювати критично важливі внутрішньодобові закономірності шахрайської активності. Для остаточного усунення ризику кластеризації за джерелом походження, транзакції в межах кожного згенерованого дня додатково перемішуються за допомогою детермінованого генератора псевдовипадкових чисел.

**Синтетичне збагачення атрибутами.** Як цільову схему транзакцій обрано протокол 3DS2 [12]. Такий вибір обґрунтовано кількома причинами. По-перше, стандарт визначає чітко структурований набір параметрів транзакції, який охоплює всі ключові ознаки, необхідні для побудови антишахрайської системи. По-друге, формалізованість протоколу забезпечує однозначну інтерпретацію кожного поля, що спрощує як генерацію синтетичних атрибутів, так і подальший аналіз результатів моделей. По-третє, розроблювана система сфокусована безпосередньо на етапі автентифікації платежу, де приймається рішення про схвалення або відхилення транзакції, а 3DS2 є стандартом, що регулює цей етап у сучасних онлайн-платежах. Оскільки жоден із відкритих наборів не містить потрібних атрибутів, їх згенеровано синтетично на основі актуальних ринкових даних:

- платіжні системи – за даними Національного банку України [13], станом на початок 2026 року частка Visa серед активних карток становила 50.7%, Mastercard 49%, національна система Простір близько 0.3%. У датасеті використано спрощений розподіл: Visa 50%, Mastercard 40%, Простір 10%. Частка системи Простір свідомо за-

вищена порівняно з реальними даними для забезпечення достатнього обсягу вибірки для навчання моделі;

- тип автентифікації – стандарт 3DS2 передбачає кілька типів обробки транзакцій: challenge (з додатковою верифікацією), frictionless (без взаємодії з користувачем) та stand-in processing (обробка на стороні платіжної системи);

- операційні системи (ОС) та браузері – за даними StatCounter [14], частка Android на мобільному ринку України у 2026 році становила 65–70%, iOS 20–35%. У датасеті Android на рівні 70%, iOS 20%, Windows 5%, macOS 3%, Linux 2%, з кореляцією браузерів до ОС;

- еквайрери – розподіл відображає структуру ринку інтернет-еквайрингу України. У датасеті серед еквайрів: Privat24 (20%), Liqpay (18%), WayForPay (15%), Platon (15%), MonoBank (7%), Fondy (5%) та інші;

- валюти та торговці – домінування гривні (UAH, 70%) відповідає специфіці українського e-commerce. Решту обсягу розподілено між USD (10%), EUR (10%) та іншими валютами. Датасет включає 114 торговців за 16 категоріями, серед яких приблизно 85% українські та 15% іноземні.

**Загальний опис атрибутів.** За результатами роботи запропонованого конвеєра формуються три взаємопов'язані масиви даних, кожен з яких призначений для відповідного компонента багат шарової моделі МН. Основний набір транзакцій (тренувальна, валідаційна та тестова вибірки) має єдину уніфіковану структуру, що складається з 25 атрибутів (див. таблиця 1). Ця схема концептуально відповідає специфікації розширеного повідомлення протоколу 3DS [12] і охоплює ключові вектори аналізу: фінансовий, географічний, мережевий та апаратний.

Таблиця 1

Структура основного набору транзакцій

Атрибут	Опис
id	Унікальний ідентифікатор транзакції
device_channel	Автентифікаційний канал
auth_type	Тип автентифікації
merchant_id, merchant_mcc, merchant_name, merchant_country_code	Ідентифікатор, категорія, назва та країна торговця
product_category	Категорія товарів
purchase_amount, purchase_currency, amount_usd	Сума в оригінальній валюті, тип валюти, нормалізована сума в USD
pno, card_bin	Платіжна система та BIN картки
card_id	Хешований ідентифікатор картки
billing_country_code	Країна покупки
email_domain	Домен електронної пошти
sender_browser, sender_browser_version	Браузер та його версія

sender_device_ip	IPv4-адреса пристрою
sender_device_model, sender_os, sender_os_version	Модель пристрою, операційна система та версія
three_ds_version	Версія протоколу 3DS
acq_id	Ідентифікатор еквайрера
is_recurring	Ознака рекурентного платежу
timestamp, transaction_hour	Дата/час транзакції та година
is_fraud	Цільовий ідентифікатор шахрайства

Для аналізу історичної поведінки користувачів формується допоміжний масив агрегованих клієнтських профілів (див. таблиця 2). Він містить розраховані статистичні та поведінкові метрики для кожної унікальної картки. Зазначені ознаки відіграють подвійну роль, оскільки вони утворюють вхідний вектор для навчання автоенкодера (який моделює закономірності легітимної поведінки та генерує оцінку похибки відновлення), а також слугують критично важливими додатковими параметрами для фінального класифікатора LightGBM.

Таблиця 2

Структура датасету клієнтських профілів

Атрибут	Опис
client_id	Хешований ідентифікатор картки
transactions_count_last_day	Кількість транзакцій за останню добу
transactions_amount_last_day	Сума транзакцій за останню добу
days_since_last_transaction	Кількість днів з останньої операції
avg_transaction_amount	Середня сума транзакції за весь період
max_transaction_amount_30d	Максимальна сума транзакції за 30 днів
unique_merchants_7d	Кількість унікальних торговців за 7 днів
unique_ips_7d	Кількість унікальних IP-адрес за 7 днів
unique_devices_7d	Кількість унікальних пристроїв за 7 днів
fraud_count_90d	Кількість шахрайських операцій за 90 днів

Третій масив є спеціалізованим набором пар «entity–IP», що має мінімалістичну структуру з двох колонок без заголовків (див. таблиця 3). Такий формат відповідає вхідним вимогам алгоритму Amazon SageMaker IP Insights, який вивчає латентні векторні представлення та обчислює оцінку аномальності.

Таблиця 3

Структура масиву пар «entity-IP»

Атрибут	Опис
entity_id	Хешований ідентифікатор картки
ip_address	IPv4-адреса пристрою

З метою дотримання методологічної чистоти та запобігання data leakage, обидва допоміжні набори обчислюються та формуються виключно на основі тренувальної вибірки (див. рис. 1).

**Статистичні характеристики результуючого датасету.** Основна зведена статистика по поточній версії датасету представлена у таблиці 4.

Таблиця 4

Статистика по отриманому набору даних

Параметр	Значення
Загальна кількість транзакцій	500000
Загальна кількість шахрайських транзакцій	15178 (3.04%)
Загальна кількість легітимних транзакцій	484822 (96.96%)
Унікальних карток	227651
Унікальних продавців	114
Унікальних IP-адрес	470682
Основна валюта	UAH (70%)
Основна операційна система	Android (69.9%)
Основний браузер	Chrome (48.8%)
Платіжні системи	Visa 50%, Mastercard 40%, Простір 10%

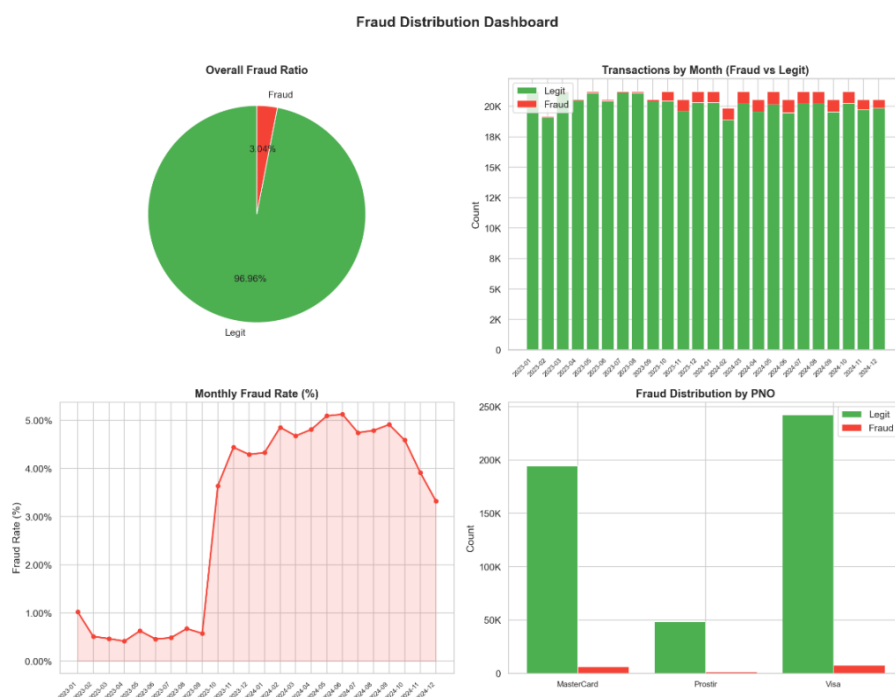


Рисунок 2 – Розподіл шахрайства за місяцями та платіжними системами

Досягнутий рівномірний розподіл транзакцій за місяцями (у середньому ~20800 записів/міс.) підтверджує коректність застосованого алгоритму квантильного перетворення часових міток. Крім того, середня сума шахрайських транзакцій (\$460–580) суттєво перевищує аналогічні показники для легітимних операцій (\$45–190), що формує дискримінативний сигнал для ефективного розділення класів моделями МН. Для візуальної перевірки підготовлених даних було побудовано низку аналітичних дашбордів.

Зокрема, на рис. 2 представлено динаміку розподілу шахрайства за місяцями та платіжними системами.

Як видно з рис. 2, запропонований підхід забезпечив рівномірний розподіл транзакцій та дотримання цільових пропорцій розподілу платіжних систем. Певне зростання рівня шахрайства з жовтня 2023 року об'єктивно пояснюється специфікою вихідних даних, оскільки у першій половині часового діапазону домінують транзакції з масиву ССТFDD (~0.6%), а в другій з масиву IEEE-CIS (~3.5%). Для більш глибокого розуміння атрибутивного складу датасету на рис. 3 візуалізовано розподіл за типами автентифікації та географічно-валютними ознаками відповідно.

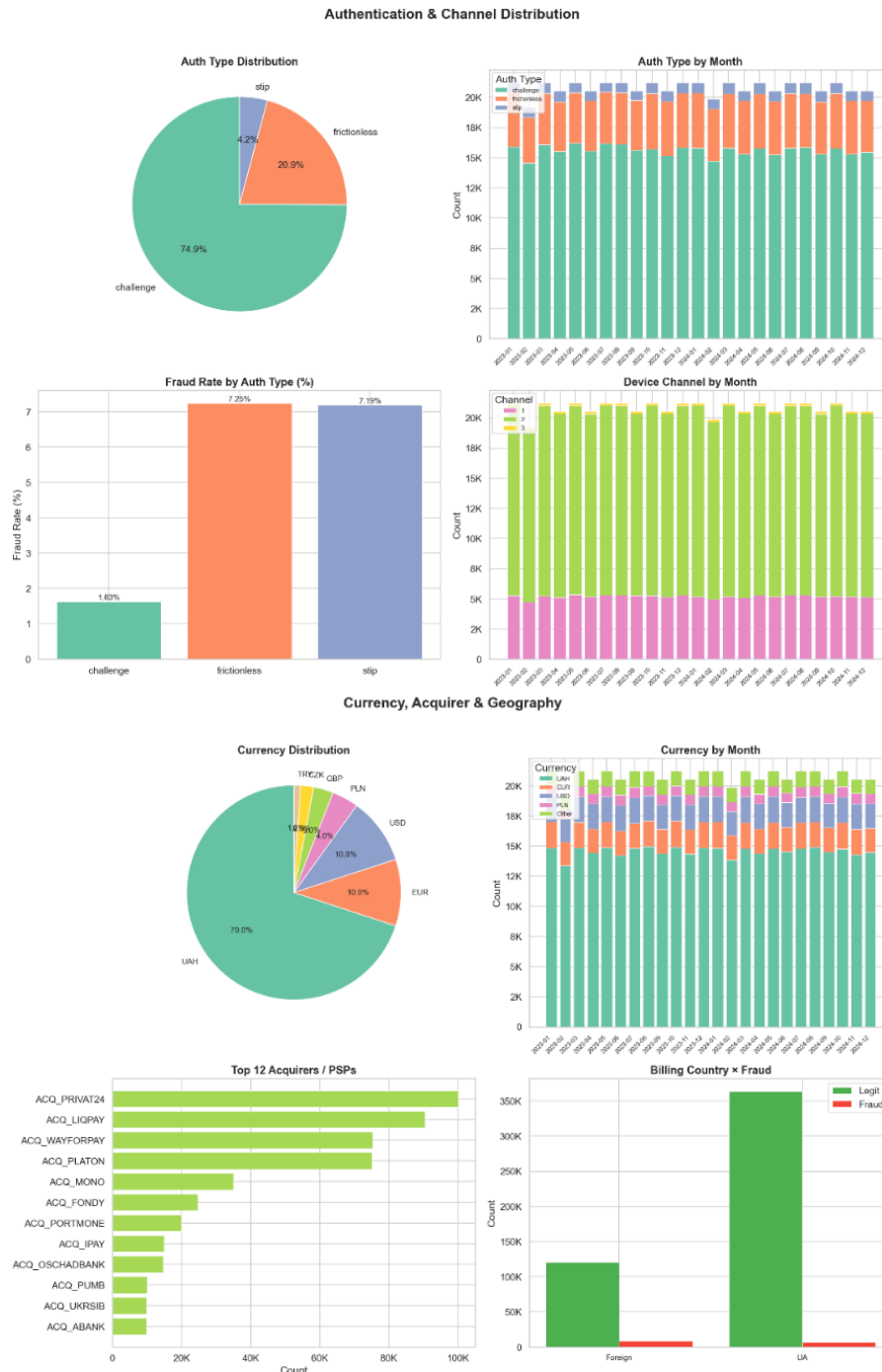


Рисунок 3 – Розподіл транзакцій за атрибутами

Фінальне хронологічне розбиття датасету у пропорції 70/15/15 імітує реальні умови експлуатації системи, де навчання відбувається на історичних подіях, а прогнозування – на майбутніх. Додатково сформована валідаційна вибірка слугує для калібрування порогів ризику, тоді як тестова для об'єктивної оцінки здатності моделі до генералізації.

**Висновки.** Описано методологію підготовки набору даних транзакцій для навчання конвеєра моделей виявлення шахрайства в електронній комерції. Запропонований підхід дозволив інтегрувати кілька відкритих наборів даних зі збереженням автентичних міток шахрайства, збагатити дані синтетичними атрибутами автентифікації з урахуванням специфіки українського платіжного ринку, а також сформувані спеціалізовані масиви для навчання моделей IP Insights та автоенкодера.

Подальша робота передбачає вдосконалення датасету шляхом додавання поведінкових часових рядів, а також безпосереднє навчання та оптимізацію розробленого конвеєра МН, його порівняльний аналіз з альтернативними алгоритмами та розгортання у хмарному середовищі AWS SageMaker.

#### ЛІТЕРАТУРА

1. Joint EBA-ECB report on payment fraud. European Banking Authority, European Central Bank. 2025. URL: <https://www.eba.europa.eu/publications-and-media/press-releases/joint-eba-ecb-report-payment-fraud-strong-authentication-remains-effective-fraudsters-are-adapting> (дата звернення: 15.03.2026).
2. Global eCommerce Payments & Fraud Report. Visa Acceptance Solutions, Merchant Risk Council. 2025. URL: <https://www.visaacceptance.com/content/dam/documents/campaign/fraud-report/global-fraud-report-2025.pdf> (дата звернення: 16.03.2026).
3. Ostrovska K., Nosov V. Machine learning methods for antifraud systems. Системні технології. 2025. Т. 5, вип. 160. С. 156–163. DOI: 10.34185/1562-9945-5-160-2025-16.
4. Grover P., Xu J., Tittelfitz J. et al. Fraud Dataset Benchmark and Applications. Amazon Science. 2022. DOI: 10.48550/arXiv.2208.14417.
5. Pushkarenko Y., Zaslavskiy V. Synthetic Data Generation for Fraud Detection Using Diffusion Models. Information Systems and Innovative Technologies in Professional Activity (ISIJ). 2024. Vol. 55, No. 2. P. 185–198. DOI: 10.11610/isij.5534.
6. IEEE-CIS Fraud Detection: Kaggle Competition. 2019. URL: <https://www.kaggle.com/competitions/ieee-fraud-detection/overview> (дата звернення: 15.03.2026).
7. Credit Card Transactions Fraud Detection Dataset : Kaggle Dataset. 2020. URL: <https://www.kaggle.com/datasets/kartik2112/fraud-detection> (дата звернення: 15.03.2026).
8. Sparkov Data Generation: GitHub Repository. URL: [https://github.com/namebrandon/Sparkov\\_Data\\_Generation](https://github.com/namebrandon/Sparkov_Data_Generation) (дата звернення: 15.03.2026).
9. Fraudulent E-Commerce Transactions: Kaggle Dataset. 2024. URL: <https://www.kaggle.com/datasets/shriyashjagtap/fraudulent-e-commerce-transactions> (дата звернення: 15.03.2026).
10. Anti-Money Laundering Datasets (IBM AMLSim): GitHub Repository. 2021. URL: <https://github.com/IBM/AMLSim> (дата звернення: 15.03.2026).

11. Credit Card Fraud Detection Dataset. Machine Learning Group, Université Libre de Bruxelles. 2018. URL: <https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud> (дата звернення: 15.03.2026).
12. EMV 3-D Secure Protocol and Core Functions Specification v2.2.0. EMVCo. 2025. URL: <https://www.emvco.com/emv-technologies/3-d-secure/> (дата звернення: 20.03.2026).
13. Visa випередила Mastercard за кількістю карток в обігу в Україні. Forbes Ukraine. 2025. URL: <https://forbes.ua/news/visa-viperedila-mastercard-za-kilkisty-kartok-v-obigu-v-ukraini-27052025-30063> (дата звернення: 18.03.2026).
14. Mobile Operating System Market Share Ukraine. StatCounter. 2024. URL: <https://gs.statcounter.com/os-market-share/mobile/ukraine> (дата звернення: 20.03.2026).

#### REFERENCES

1. European Banking Authority & European Central Bank. (2025). *Joint EBA-ECB report on payment fraud*. <https://www.eba.europa.eu/publications-and-media/press-releases/joint-eba-ecb-report-payment-fraud-strong-authentication-remains-effective-fraudsters-are-adapting>
2. Visa Acceptance Solutions & Merchant Risk Council. (2025). *2025 Global eCommerce Payments & Fraud Report*. <https://www.visaacceptance.com/content/dam/documents/campaign/fraud-report/global-fraud-report-2025.pdf>
3. Ostrovska, K., & Nosov, V. (2025). Machine learning methods for antifraud systems. *System technologies*, 5(160), 156–163. <https://doi.org/10.34185/1562-9945-5-160-2025-16>
4. Grover, P., Xu, J., Tittelfitz, J., Cheng, A., Li, Z., Zablocki, J., Liu, J., & Zhou, H. (2022). Fraud Dataset Benchmark and Applications. *Amazon Science*. <https://doi.org/10.48550/arXiv.2208.14417>
5. Pushkarenko, Y., & Zaslavskiy, V. (2024). Synthetic Data Generation for Fraud Detection Using Diffusion Models. *Information Systems and Innovative Technologies in Professional Activity (ISIJ)*, 55(2), 185–198. <https://doi.org/10.11610/isij.5534>
6. IEEE-CIS Fraud Detection. (2019). *Kaggle*. <https://www.kaggle.com/competitions/ieee-fraud-detection/overview>
7. Credit Card Transactions Fraud Detection Dataset. (2020). *Kaggle*. <https://www.kaggle.com/datasets/kartik2112/fraud-detection>
8. Sparkov Data Generation. *GitHub*. [https://github.com/namebrandon/Sparkov\\_Data\\_Generation](https://github.com/namebrandon/Sparkov_Data_Generation)
9. Fraudulent E-Commerce Transactions. (2024). *Kaggle*. <https://www.kaggle.com/datasets/shriyashjagtap/fraudulent-e-commerce-transactions>
10. Anti-Money Laundering Datasets (IBM AMLSim). (2021). *GitHub*. <https://github.com/IBM/AMLSim>
11. Credit Card Fraud Detection Dataset. (2018). Machine Learning Group, Université Libre de Bruxelles. *Kaggle*. <https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud>
12. EMVCo. (2025). *EMV 3-D Secure Protocol and Core Functions Specification v2.2.0*. <https://www.emvco.com/emv-technologies/3-d-secure/>
13. Visa vyperedyla Mastercard za kilkistiu kartok v obihu v Ukraini [Visa overtook Mastercard by number of cards in circulation in Ukraine]. (2025). *Forbes Ukraine*. <https://forbes.ua/news/visa-viperedila-mastercard-za-kilkisty-kartok-v-obigu-v-ukraini-27052025-30063> [in Ukrainian].

14. StatCounter. (2024). *Mobile Operating System Market Share Ukraine*.

<https://gs.statcounter.com/os-market-share/mobile/ukraine>

Received 30.03.2026.

Accepted 01.04.2026.

Published 30.04.2026

### ***Automated pipeline for building a fraud detection training dataset***

*The study addresses the problem of preparing training data for machine learning-based fraud detection systems in e-commerce transactions. Due to the strict confidentiality of real transaction data, researchers often rely on publicly available datasets that typically suffer from limited attribute schemas, anonymized features, and a focus on specific national markets. An analysis of existing open datasets revealed the necessity of creating a specialized dataset, as none of the available sources provide a sufficient combination of realistic fraud labels, semantic transparency of features, and domain-specific attributes required for training a multi-component fraud detection system.*

*An automated pipeline for integrating three open Kaggle datasets (IEEE-CIS, Credit Card Transactions Fraud Detection Dataset, Fraudulent E-Commerce) is proposed. The pipeline preserves authentic fraud labels and original transaction amounts while enriching records with synthetic attributes adapted to the specifics of the Ukrainian payment market. The methods developed include: uniform normalization of timestamps based on quantile rank transformation to eliminate dataset shift artifacts while preserving intra-day patterns, synthetic generation of authentication attributes according to the EMV 3D-Secure 2.0 standard with payment network distributions based on National Bank of Ukraine statistics, formation of aggregated client behavioral profiles, and generation of “entity-IP” pairs for IP Insights model training. Both auxiliary datasets are derived exclusively from the training subset to prevent data leakage.*

*The resulting dataset comprises 500000 transactions spanning 24 months with a fraud rate of 3.04%, designed for training a model pipeline that includes LightGBM, an autoencoder, and IP Insights. The chronological split simulates real-world deployment conditions where models are trained on historical events and evaluated on future ones.*

*Keywords: dataset, machine learning, transaction, e-commerce, LightGBM, autoencoder, IP Insights, EMV 3D-Secure, fraud detection, synthetic data.*

**Носов Валерій Олександрович** – аспірант кафедри інформаційних технологій і систем Українського державного університету науки і технологій.

ORCID: <https://orcid.org/0009-0003-5841-8995>

**Островська Катерина Юріївна** – к.т.н., доцент, доцент кафедри інформаційних технологій і систем Українського державного університету науки і технологій.

ORCID: <https://orcid.org/0000-0002-9375-4121>

**Nosov Valerii Oleksandrovyich** – postgraduate student of the Department of Information Technology and Systems of Ukrainian State University of Science and Technology.

ORCID: <https://orcid.org/0009-0003-5841-8995>

**Ostrovskaya Kateryna Yuriivna** – Ph.D., Associate Professor of the Department of Information Technology and Systems of Ukrainian State University of Science and Technology.

ORCID: <https://orcid.org/0000-0002-9375-4121>

М.О. Пристайло, М.М. Балака, А.Г. Поліщук, І.П. Гонта

## ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ ПРОЦЕСУ РІЗАННЯ ҐРУНТУ НАКОНЕЧНИКОМ З ДИНАМІЧНОЮ РІЖУЧОЮ КРОМКОЮ

*Анотація.* У статті наведено результати експериментальних досліджень процесу різання ґрунту наконечником з динамічною ріжучою кромкою. Розглянуто актуальну задачу підвищення енергоефективності процесу різання шляхом використання пневмоакумулятора. Проблема полягає у недостатній точності існуючих теоретичних моделей, що зумовлює необхідність експериментального уточнення силових параметрів взаємодії робочого органу з ґрунтовим середовищем. Метою дослідження є експериментальне визначення сили різання та встановлення закономірностей її зміни залежно від тиску в пневмоакумуляторі. У роботі застосовано методи фізичного моделювання з використанням парафінового середовища, тензометричних вимірювань та математичної обробки результатів експерименту. Встановлено нелінійний характер залежності сили різання від тиску, який має U-подібну форму. Визначено оптимальний діапазон тиску, за якого досягається мінімальний опір різанню та забезпечується найбільш енергоефективний режим роботи системи. Порівняння експериментальних і теоретичних результатів показало їх узгодженість із похибкою до 11 %, що підтверджує адекватність запропонованої моделі. Отримані результати можуть бути використані при проектуванні та оптимізації робочих органів машин.

*Ключові слова:* різання ґрунту, динамічна ріжуча кромка, пневмоакумулятор, сила різання, тиск, експеримент, коефіцієнт подібності, енергоефективність, модель.

**Постановка проблеми.** Процес різання ґрунту є одним з основних механічних процесів у роботі будівельно-дорожніх і спеціальних машин, що безпосередньо визначає їх енергетичні витрати, навантаження на робочі органи та продуктивність. Водночас існуючі теоретичні моделі недостатньо точно описують силові параметри взаємодії ріжучого елемента з ґрунтовим середовищем, особливо в умовах застосування динамічних робочих органів, які змінюють характер навантаження та механізм руйнування ґрунту. Це зумовлює необхідність експериментального уточнення та подальшого вдосконалення закономірностей процесу різання ґрунту.

**Аналіз останніх досліджень і публікацій.** Питання визначення силових параметрів процесу різання ґрунту розглядається у наукових роботах, оснований на положеннях механіки суцільного середовища, теорії граничної рівноваги та енергетичних підходах до руйнування ґрунтового масиву [1, 2]. У класичних дослідженнях встановлено залежність сил опору різанню від фізико-механічних властивостей ґрунту, геометрії

ріжучого інструмента та режимів його роботи [3, 4]. Подальший розвиток отримали чисельні методи моделювання, зокрема метод скінченних елементів та метод дискретних елементів, які дозволяють детально описувати процес формування зони руйнування та напружено-деформований стан ґрунту [5]. Водночас встановлено, що такі моделі характеризуються високою чутливістю до параметрів ґрунтового середовища та граничних умов, що зумовлює необхідність їх експериментальної верифікації [6].

У ряді робіт показано, що процес різання ґрунту має суттєво нелінійний характер, зумовлений його неоднорідністю, структурною будовою та пружно-пластичними властивостями, що призводить до розбіжностей між теоретичними та експериментальними значеннями силових параметрів [7]. Разом з тим, недостатньо вивченими є питання впливу динамічно-акумуляюючих елементів на силові параметри процесу різання ґрунту та умови реалізації мінімального енергетичного опору [5, 8].

**Мета досліджень.** Метою дослідження є експериментальне визначення сили різання ґрунту та встановлення закономірностей її зміни залежно від тиску в пневмоаккумуляторі робочого органу з динамічною ріжучою кромкою.

**Викладення основного матеріалу дослідження.** Для проведення досліджень використовувалась лабораторна установка для моделювання процесу різання ґрунту наконечником із динамічною ріжучою кромкою на пневмоаккумуляторі (рисунок 1), яка містить такі основні вузли [6]:

- ґрунтовий лоток, що забезпечує візуальний доступ до моделі ґрунту;
- механізм переміщення наконечника з плавним регулюванням швидкості в діапазоні 0,01–0,10 м/с;
- наконечник з динамічною ріжучою кромкою, шарнірно або пружно з'єднаний з корпусом через пневмоаккумулятор;
- силовимірювальний вузол, призначений для реєстрації горизонтальної складової сили різання  $F(t)$  та, за необхідності, вертикальної складової  $N(t)$ ;
- гідропневматичний контур, що включає пневмоаккумулятор, запірно-регульовальну арматуру та датчик тиску  $p_{ак}(t)$ ;
- систему збору та реєстрації даних, яка забезпечує синхронний запис сигналів тензометричних датчиків, датчика тиску та переміщення.

Використані підходи до побудови лабораторної установки узгоджуються з раніше розробленими рішеннями для реєстрації зусиль різання [6, 7]. Робочий орган (рисунок 2) являє собою клиноподібний наконечник з основною ріжучою кромкою та, за необхідності, лідируючою кромкою, яка випереджає основну та формує попередній надріз у моделі ґрунту, як це наведено у попередніх дослідженнях [8]. Ріжуча кромка шарнірно приєднана до корпусу наконечника та взаємодіє з ним через пневмоаккумулятор. Під час проходження зон підвищеного опору ріжуча кромка відхиляється, що призводить до стискання робочого середовища в аккумуляторі. Після проходження перешкоди частина накопиченої енергії повертається в систему, забезпечуючи згладжування коливань навантаження та стабілізацію процесу різання.

Основними регульованими параметрами є:

- кут загострення ріжучої кромки  $\alpha$  ;
- кут встановлення кромки відносно напрямку руху;
- початковий тиск у пневмоакумуляторі  $p_{ак}$  ;
- жорсткість газової «пружины», що визначається  $p(V)$  .

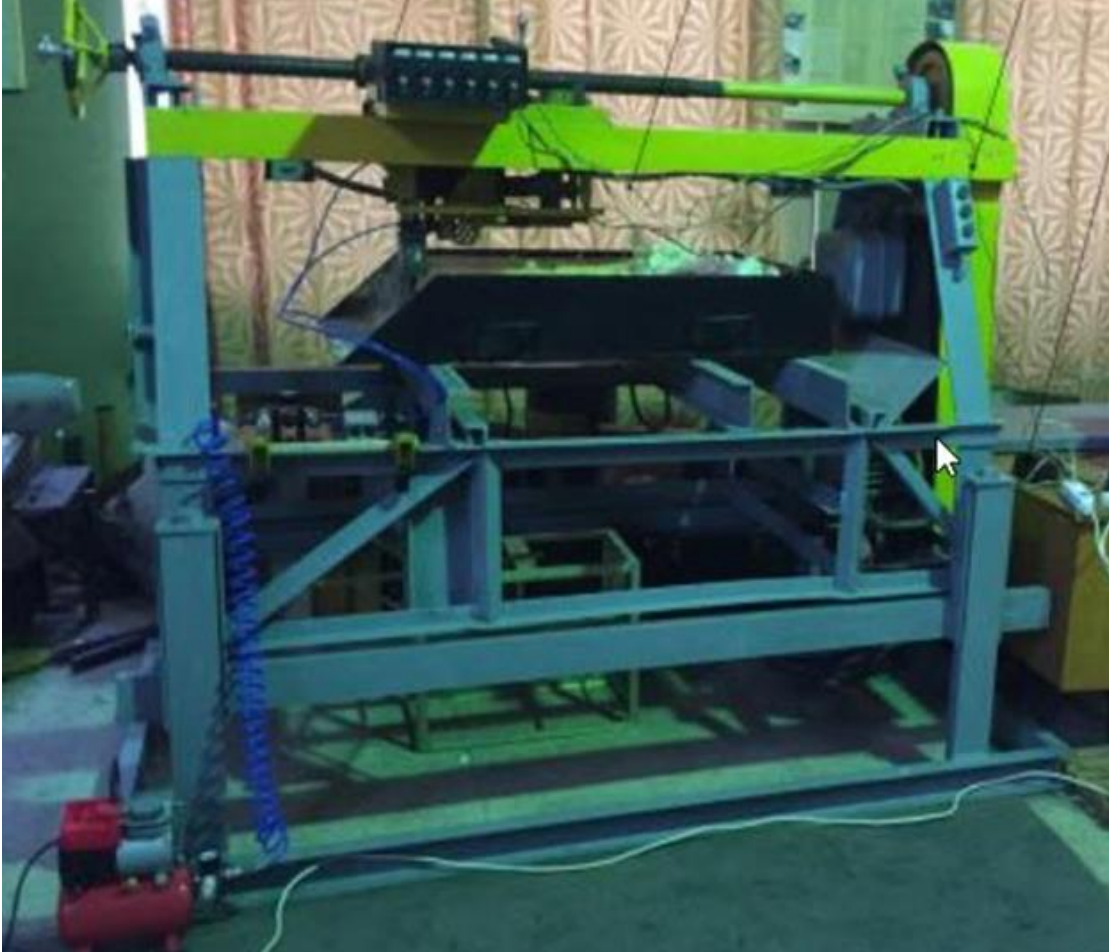


Рисунок 1 – Лабораторна установка для дослідження процесу різання моделі ґрунту наконечником з динамічною ріжучою кромкою



Рисунок 2 – Наконечник з динамічною ріжучою кромкою на пневмоакумуляторі

Як модель ґрунту використано парафін розмірами  $L \times B \times H = 0,8 \times 0,2 \times 0,1$  м, що за реологічними та механічними характеристиками забезпечує наближене моделювання пружно-пластичної поведінки зв'язних ґрунтів [2, 6], зокрема їх здатності до пластичного деформування та утворення зони зрізу.

Різання моделі ґрунту здійснювалось при стабільній швидкості переміщення робочого органу не більше 0,09 м/с, із варіюванням тиску в пневмоакумуляторі в діапазоні від 0 до 15 МПа з кроком 2,5 МПа. Під час експерименту виконувалась безперервна реєстрація тангенціальної складової сили різання, яка характеризує роботу руйнування матеріалу та подолання сил тертя на контактних поверхнях.

При відсутності тиску в пневмоакумуляторі (рисунок 3, а) процес різання характеризується нестабільним режимом: середнє значення сили становить близько 12,4 кН, а сигнал має хаотичний характер зі значною шумовою складовою, спостерігаються короткочасні пікові перевантаження та провали.

При тиску 2,5 МПа (рисунок 3, б) фіксується зниження середнього рівня сили до приблизно 10,5 кН та поява перехідного процесу зі стабілізацією сигналу, часткове демпфування коливань і перехід системи до квазістаціонарного режиму.

При тиску 5 МПа (рисунок 3, в) середнє значення сили зменшується до 9,75 кН, а характер сигналу набуває вираженої циклічності. При цьому формується автоколивальний режим, що відображає періодичне накопичення та вивільнення енергії в системі «робочий орган – ґрунтове середовище».

При тиску 7,5 МПа (рисунок 3, г) спостерігається подальше зниження середньої сили до 9,16 кН та більш упорядкований характер коливань. У цьому режимі досягається мінімальний опір різанню, що відповідає енергоефективній роботі системи.

При підвищенні тиску до 10 МПа (рисунок 3, д) середнє значення сили знижується до 8,66 кН, однак зростає амплітуда миттєвих коливань і з'являються імпульсні навантаження. Це свідчить про перехід до режиму жорсткішої взаємодії, за якого пневмоакумулятор не лише демпфує, але й активно впливає на динаміку процесу різання.

При тиску 12,5 МПа (рисунок 3, е) спостерігається зростання середнього значення сили до 9,26 кН та збільшення амплітуди коливань, що вказує на порушення оптимального режиму та підвищення енергетичної нестабільності.

При максимальному дослідженому тиску 15 МПа (рисунок 3, є) середня сила зростає до 11,55 кН, сигнал характеризується значними імпульсними піками та високою дисперсією. Це пояснюється переходом до жорсткого ударного контакту, втратою демпфуючого ефекту та збільшенням опору різанню.

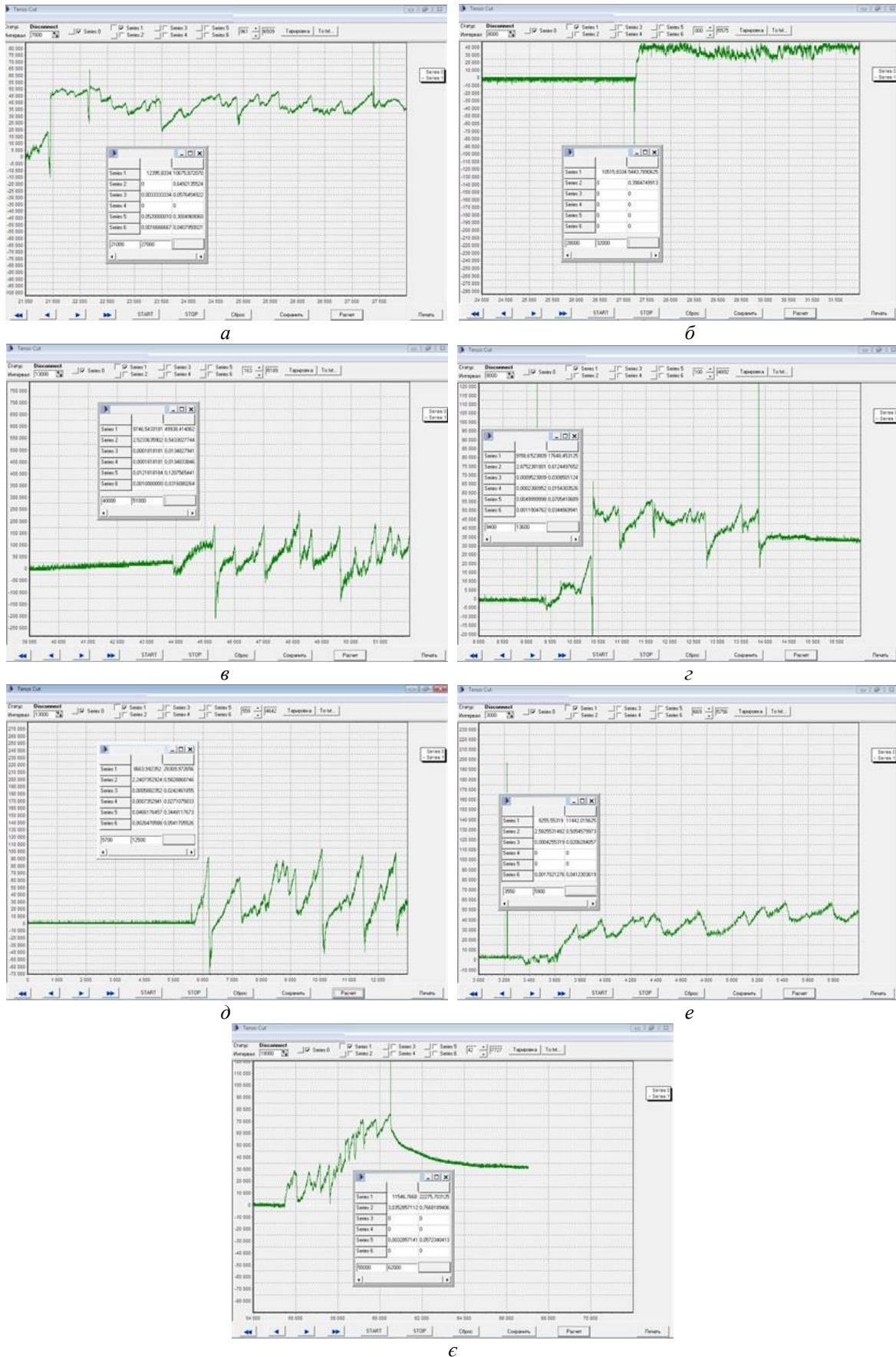


Рисунок 3 – Осцилограми зміни сили різання при різних значеннях тиску (від 0 до 15 МПа з кроком 2,5 МПа) в пневмоакумуляторі

Узагальнена експериментальна залежність (рисунок 4) свідчить про нелінійний характер зміни сили різання від тиску в пневмоакумуляторі. Крива має виражену U-подібну форму: зі зростанням тиску від 0 до 7,5–10 МПа сила різання зменшується, досягаючи мінімального значення, після чого знову зростає. Це підтверджує наявність оптимального діапазону тиску, за якого забезпечується мінімальний силовий опір і найвища енергоефективність процесу.

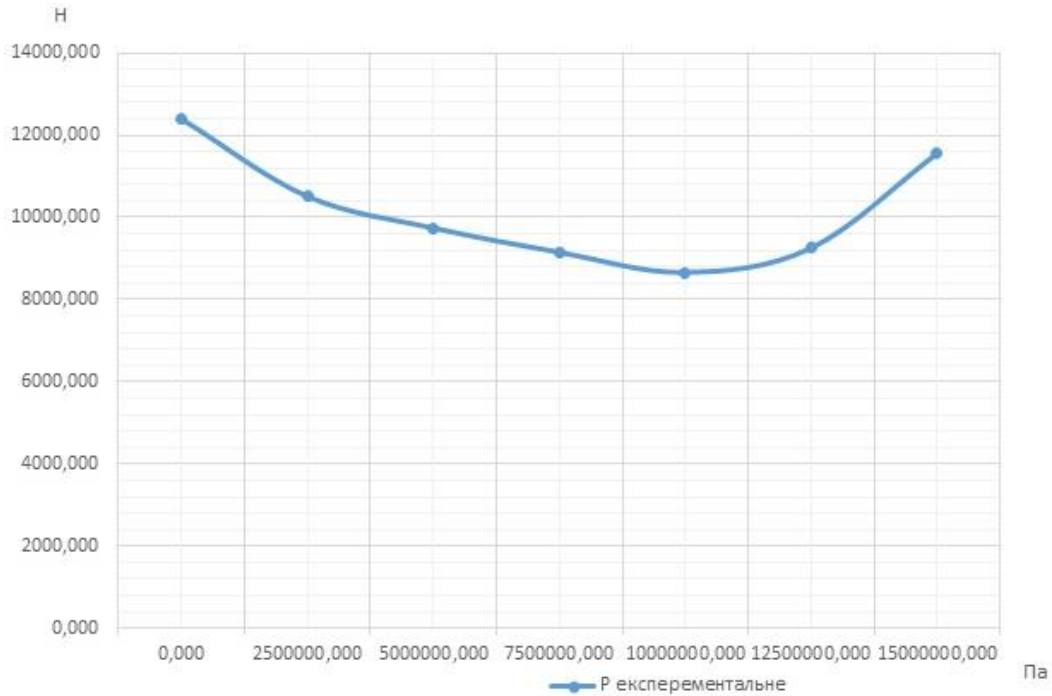


Рисунок 4 – Узагальнена залежність середнього значення сили від тиску

У межах проведених експериментальних досліджень встановлено вплив тиску в пневмоакумуляторі на силу опору різання моделі ґрунту (парафіну) клиноподібним наконечником із динамічною ріжучою кромкою.

Діапазон зміни тиску становив 0–15 МПа. Отримані результати свідчать про виражений нелінійний характер залежності сили різання від тиску в пневмоакумуляторі. Зокрема, при відсутності тиску (0 МПа) середнє значення сили опору різання досягає максимального рівня і становить близько 12,4 кН. З підвищенням тиску до 2,5–5,0 МПа спостерігається інтенсивне зниження сили до рівня 9,7–10,5 кН, що зумовлено появою демпфувального ефекту та частковою стабілізацією процесу руйнування.

Подальше збільшення тиску до 7,5–10 МПа супроводжується досягненням мінімальних значень сили різання (близько 8,7–9,2 кН), що відповідає найбільш раціональному режиму роботи системи. У цьому діапазоні забезпечується узгоджена взаємодія між жорсткістю системи та енергетичними параметрами процесу, що сприяє зниженню питомих витрат енергії.

Разом з тим, при подальшому підвищенні тиску до 12,5–15 МПа спостерігається зростання сили різання до 9,3–11,5 кН, що свідчить про порушення оптимального режиму. Це пояснюється підвищенням жорсткості системи «наконечник – акумулятор – середовище», зростанням контактних напружень та інтенсифікацією сил тертя.

Узагальнення експериментальних даних показує, що амплітуда зміни сили різання у дослідженому діапазоні становить близько 3,7 кН. Незважаючи на відносно помірну величину зміни, вона має систематичний і відтворюваний характер, що підтверджує суттєвий вплив параметрів пневмоакумулятора на механіку процесу руйнування модельного середовища.

Характер отриманої залежності має квазіпараболічний вигляд і може бути інтерпретований як результат зміни жорсткісно-енергетичних параметрів системи. При малих значеннях тиску відсутня ефективна компенсація динамічних навантажень, що призводить до збільшення зони пластичної деформації та зростання енергетичних втрат. У середньому діапазоні тисків відбувається стабілізація процесу формування зони руйнування та зменшення питомої роботи різання. Натомість при надмірних значеннях тиску система набуває підвищеної жорсткості, що зумовлює зростання силового опору.

У межах проведених досліджень встановлено наявність оптимального діапазону тиску в пневмоакумуляторі, який для даних умов становить близько 8–10 МПа.

З огляду на встановлений U-подібний характер залежності сили різання від тиску, для її опису застосовано апроксимацію поліномом другого порядку

$$P_s(P_z) = aP_z^2 + bP_z + c. \quad (1)$$

Обробка експериментальних даних методом найменших квадратів дозволила отримати відповідну аналітичну залежність

$$P_s(P_z) = 0,036P_z^2 - 0,71P_z + 12,3 \quad (2)$$

з коефіцієнтом детермінації  $R^2 = 0,91$ , що свідчить про достатньо високий рівень адекватності моделі.

Аналіз отриманої функції показує, що мінімум сили різання досягається при певному тиску, яке відповідає оптимальному режиму роботи системи:

$$\frac{dP_s}{dP_z} = 2aP_z + b = 0 \Rightarrow P_z^* = -\frac{b}{2a}. \quad (3)$$

Підставляючи коефіцієнти,  $P_z^* \approx 10$  МПа, мінімальне значення сили буде  $P_{z.min} \approx 8,6$  кН. У цій точці забезпечується найменший силовий опір та найбільш ефективне використання енергії.

Отримані результати можуть бути пояснені з позицій енергетичного балансу процесу різання. На початковому етапі збільшення тиску в пневмоакумуляторі сприяє стабілізації фронту руйнування та зменшенню коливань у зоні зрізу, що призводить до зниження питомих витрат енергії. Після досягнення оптимального значення тиску подальше його зростання зумовлює інтенсивніше ущільнення матеріалу, підвищення нормальних напружень і, як наслідок, збільшення сил тертя. У цьому випадку додаткова енергія витрачається не на руйнування матеріалу, а на подолання внутрішнього опору середовища, що призводить до зростання сили різання.

Для оцінки адекватності отриманих експериментальних даних виконаємо порівняння з результатами теоретичного моделювання (рисунок 5).

Експериментальні значення сили різання приведено до натурних умов із використанням коефіцієнтів подібності. Аналіз графіків показує, що експериментальна та теоретична залежності мають однакові межі зміни та подібний характер. При цьому спостерігається незначне перевищення експериментальних значень, що пояснюється впливом додаткових факторів, не врахованих у теоретичній моделі (неоднорідність середовища, втрати на тертя).

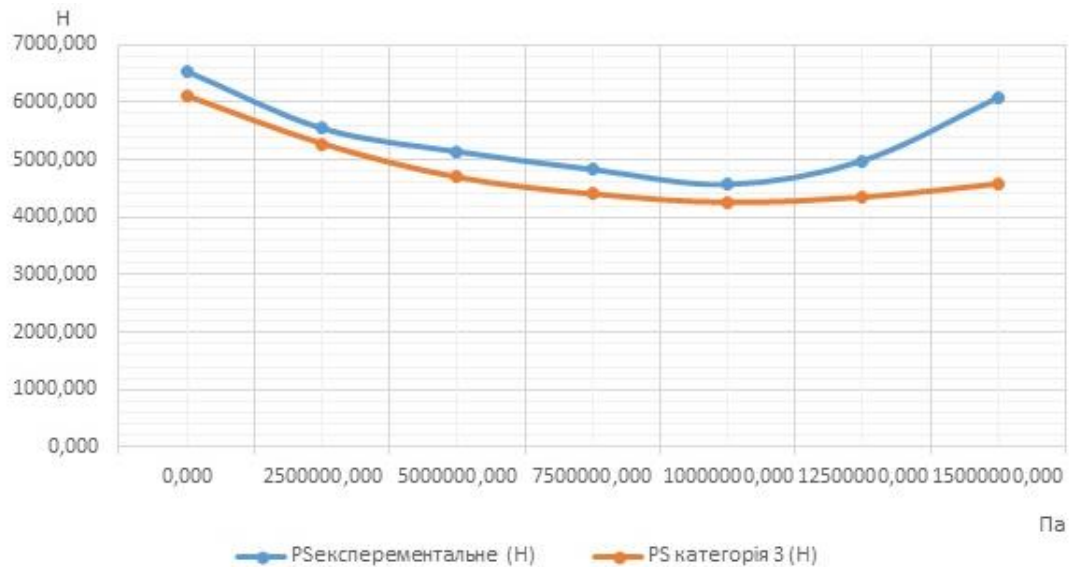


Рисунок 5 – Порівняння сил різання ґрунту, отриманих теоретичним та експериментальним методами

Величина розбіжності між результатами не перевищує 11 %, що свідчить про достатню адекватність запропонованої теоретичної моделі та можливість її використання для інженерних розрахунків.

**Висновки.** Встановлено, що залежність сили різання ґрунту від тиску в пневмоакумуляторі в діапазоні 0–15 МПа має U-подібний характер: при зростанні тиску до 7,5–10 МПа сила зменшується з 12,4 до 8,7–9,2 кН (на 25–30 %), після чого зростає до 11,5 кН. Визначено оптимальний діапазон тиску 7,5–10 МПа, при якому досягається мінімальний силовий опір і енергоефективний режим роботи системи. Показано узгодженість експериментальних і теоретичних результатів (розбіжність до 11 %), що підтверджує адекватність моделі та можливість її використання для інженерних розрахунків.

#### ЛІТЕРАТУРА

1. Mase G. T., Smelser R. E., Rossmann J. S. Continuum Mechanics for Engineers. 4th ed. Boca Raton. CRC Press, 2020. 450 p.
2. Шашенко О. М., Шаповал В. Г., Гапєєв С. М., Скобенко О. В., Хозяїкіна Н. В. Механіка ґрунтів. Дніпро: НТУ «Дніпровська політехніка», 2024. 339 с.
3. Look V. G. Earthworks: Theory to Practice – Design and Construction. Boca Raton. CRC Press, 2022. 590 p.
4. Хмара Л. А., Кравець С. В., Скоблюк М. П., Нікітін В. Г., Деревянчук М. І., Супонєв В. М. Машини для земляних робіт. Харків: ХНАДУ, 2014. 546 с.

5. Prystailo M., Balaka M., Mozharivskiy V., Drachuk V., Honta I. Superposition principle of impact on the working environment of actuating elements for site preparation machines. Bulletin of Kharkov National Automobile and Highway University, 2024, Vol. 1 No. 105, 61–67. URL: <https://doi.org/10.30977/BUL.2219-5548.2024.105.1.61>.
6. Rashkivskiy V., Prystailo M., Fedyshyn B., Balaka M. Methods of conducting a bench-scale experimental study with a spatially oriented knife of a bulldozer blade. International Science Journal of Engineering & Agriculture, 2025, 4(1), 79–92. URL: <https://doi.org/10.46299/j.isjea.20250401.07>.
7. Лук'янчук О. П., Степанюк Б. І., Форсюк С. Л. Експериментальне дослідження зміни зусилля різання залежно від умов розробки ґрунту. Вісник Національного університету водного господарства та природокористування, 2024, Вип. 2 (106), С. 3–12. URL: <https://doi.org/10.31713/vt220241>.
8. Prystailo M. O., Balaka M. M., Polishchuk A. G., Kozak A. A., Honta I. P. Optimization of soil cutting process by ripper tip with dynamic cutting edge. Strength of Materials and Theory of Structures, 2025, Issue 115, P. 325–334. URL: <https://doi.org/10.32347/2410-2547.2025.115.325-334>.

#### REFERENCES

1. Masem, G. T., Smelser, R. E. & Rossmann, J. S. (2020). Continuum Mechanics for Engineers. 4th ed. Boca Raton. CRC Press. 450 p.
2. Shashenko, O. M., Shapoval, V. H., Hapiev, S. M., Skobenko, O. V. & Khoziaikina, N. V. (2024). Mekhanika gruntiv [Soil mechanics]. Dnipro: NTU «Dniprovska politehnika». 339 p.
3. Look, B. G. (2022). Earthworks: Theory to Practice – Design and Construction. Boca Raton. CRC Press. 590 p.
4. Khmara, L. A., Kravets, S. V., Skobliuk, M. P., Nikitin, V. H., Derevianchuk, M. I. & Saponiev, V. M. (2014). Mashyny dlia zemlianykh robit [Machines for earthmoving]. Kharkiv: KhNADU. 546 p.
5. Prystailo, M., Balaka, M., Mozharivskiy, V., Drachuk, V. & Honta, I. (2024). Superposition principle of impact on the working environment of actuating elements for site preparation machines. Bulletin of Kharkov National Automobile and Highway University, Vol. 1 No. 105, 61–67. URL: <https://doi.org/10.30977/BUL.2219-5548.2024.105.1.61>.
6. Rashkivskiy, V., Prystailo, M., Fedyshyn, B. & Balaka, M. (2025). Methods of conducting a bench-scale experimental study with a spatially oriented knife of a bulldozer blade. International Science Journal of Engineering & Agriculture, 4(1), 79–92. URL: <https://doi.org/10.46299/j.isjea.20250401.07>.
7. Lukianchuk, O. P., Stepaniuk, B. I. & Forsiuk, S. L. (2024). Eksperymentalne doslidzhenia zminy zusyillia rizannia zalezno vid umov rozrobky ґрунту [Experimental study of the change in cutting effort depending on the conditions of soil development]. Visnyk Natsionalnoho universytetu vodnoho hospodarstva ta pryrodokorystuvannia – Bulletin of the National University of Water Management and Natural Resources, Issue 2 (106), 3–12. URL: <https://doi.org/10.31713/vt220241>.

8. Prystailo, M. O., Balaka, M. M., Polishchuk, A. G., Kozak, A. A. & Honta, I. P. (2025). Optimization of soil cutting process by ripper tip with dynamic cutting edge. *Strength of Materials and Theory of Structures*, Issue 115, 325–334. URL: <https://doi.org/10.32347/2410-2547.2025.115.325-334>.

Received 30.03.2026.

Accepted 02.04.2026.

Published 30.04.2026

### ***Experimental study of soil cutting with dynamic cutting edge tip***

*The experimental study results of the of soil cutting with dynamic cutting edge tip and the pneumatic accumulator are presented in the paper. Recent research and publications indicate increasing attention to improving the energy efficiency of earthmoving processes through adaptive and energy-storing working elements; however, existing theoretical models remain insufficiently accurate for predicting force interaction parameters in complex soil–tool systems. This creates the need for experimental verification and refinement of analytical approaches. The purpose of the study is to experimentally determine the cutting resistance force and to establish its dependence on the pressure level in the pneumatic accumulator. The main research methods include physical modeling using the paraffin-based soil analogue, strain-gauge measurements, and mathematical processing of experimental data with subsequent comparative analysis against theoretical predictions. The obtained results demonstrate the pronounced nonlinear dependence of cutting force on pressure, characterized by the U-shaped curve. It was found that the system exhibits an optimal pressure range, within which the minimum cutting resistance and the most energy-efficient operating regime are achieved. Outside this range, both low and excessive pressure levels lead to increased cutting resistance due to insufficient damping or excessive system rigidity. The comparison between experimental and theoretical results showed good agreement, with the deviation of up to 11%, confirming the adequacy of the proposed theoretical model. The results obtained can be used for the design and optimization of working elements of construction machinery aimed at reducing energy consumption and improving operational efficiency.*

*Keywords: soil cutting, dynamic cutting edge, pneumatic accumulator, cutting force, pressure, experiment, similarity factors, energy efficiency, model.*

**Пристайло Микола Олексійович** – к.т.н., доцент, професор кафедри будівельних машин, Київський національний університет будівництва і архітектури.

ORCID: <https://orcid.org/0000-0003-3151-4680>

**Балака Максим Миколайович** – к.т.н., доцент, доцент кафедри будівельних машин, Київський національний університет будівництва і архітектури.

ORCID: <https://orcid.org/0000-0003-4142-9703>

**Поліщук Андрій Григорович** – к.т.н., асистент кафедри будівельних машин, Київський національний університет будівництва і архітектури.

ORCID: <https://orcid.org/0000-0003-4808-9932>

**Гонта Ігор Петрович** – аспірант кафедри будівельних машин, Київський національний університет будівництва і архітектури.

ORCID: <https://orcid.org/0009-0002-4441-8049>

**Prystailo Mykola** – Candidate of Science (Engineering), Associate Professor, Professor of the Department of Construction Machinery, Kyiv National University of Construction and Architecture.

ORCID: <https://orcid.org/0000-0003-3151-4680>

**Balaka Maksym** – Candidate of Science (Engineering), Associate Professor, Associate Professor of the Department of Construction Machinery, Kyiv National University of Construction and Architecture.

ORCID: <https://orcid.org/0000-0003-4142-9703>

**Polishchuk Andrii** – Candidate of Science (Engineering), Assistant of the Department of Construction Machinery, Kyiv National University of Construction and Architecture.

ORCID: <https://orcid.org/0000-0003-4808-9932>

**Honta Ihor** – Postgraduate Student of the Department of Construction Machinery, Kyiv National University of Construction and Architecture.

ORCID: <https://orcid.org/0009-0002-4441-8049>

## ВДОСКОНАЛЕННЯ МЕТОДУ ВИРІШЕННЯ ЗАДАЧІ РАННЬОЇ ІДЕНТИФІКАЦІЇ КОНФІГУРАЦІЙНИХ ЕЛЕМЕНТІВ ІНФОРМАЦІЙНОЇ СИСТЕМИ

*Анотація.* Формальне вирішення задачі ранньої ідентифікації конфігураційних елементів (Configuration Items, CI) інформаційної системи (ІС) дозволить значно зменшити витрати часу та зусиль на створення описів архітектури, планування та управління ІТ-проектом створення ІС. Науково-прикладною проблемою дослідження є визначення можливості вирішення задачі ранньої ідентифікації CI ІС з використанням методів та алгоритмів кластеризації, які не сильно залежать від припущень про форму кластерів. Метою дослідження є розробка нового або вдосконалення існуючого методу вирішення задачі ранньої ідентифікації CI ІС. Запропоновано вдосконалення методу вирішення цієї задачі, яке полягає у використанні алгоритму кластеризації DBSCAN для виявлення архітектурних елементів як кластерів окремих функцій ІС. Використання вдосконаленого методу дозволяє автоматизувати процедуру синтезу опису архітектури ІС. Застосування цієї автоматизованої процедури сприяє підвищенню якості розробки ІС за рахунок виділення для проектування множини архітектурних сутностей цієї системи, яка є значно меншою, ніж множина елементарних функцій ІС.

*Ключові слова:* інформаційна система, конфігураційний елемент, задача ранньої ідентифікації, функція, архітектурна сутність, DBSCAN, опис архітектури.

**Постановка проблеми.** Процес управління конфігурацією є одним з найважливіших процесів життєвого циклу інформаційних систем (ІС) управління підприємствами та організаціями. Призначенням цього процесу є систематичний контроль за змінами конфігурації, підтримка цілісності та відстеження конфігурації протягом усього життєвого циклу продукту [1]. Незалежно від підходів до структуризації та опису окремих робіт і задач цього процесу, у ньому виділяється задача ідентифікації конфігурації, пов'язана із визначенням системних елементів конфігурації («Configuration item», CI) [1, 2]. Рішення цієї задачі поступово встановлюють та підтримують поточну основу для контролю та обліку стану ІС та її CI протягом усіх їхніх життєвих циклів [2].

Теоретичні та прикладні основи систем управління конфігурацією ІТ-продуктів визнано в цілому сформованими [2, 3] і запропоновано представляти як сукупність таких груп [3]:

- теоретичні основи (зокрема, трикомпонентна модель);
- кращі практики (зокрема, ISO 20000 та ITIL, а також їх аналоги);

– інновації (зокрема, інтеграція DevOps, інфраструктура як код, технології контейнеризації).

Теоретичні основи формують фундаментальне розуміння управління конфігурацією. Найкращі практики забезпечують досить ефективне вирішення задачі ідентифікації, контролю та обліку стану СІ. Розглянуті інновації удосконалюють традиційні методи управління конфігурацією ІТ-продукту, забезпечуючи масштабованість, автоматизацію та адаптованість прикладних рішень [3].

Головним недоліком такого представлення є його концептуальність – незважаючи на єдність поглядів на визначення задачі ідентифікації конфігурації, її розв’язання й досі розглядаються на концептуальному рівні [2]. Теоретичні основи і кращі практики управління конфігурацією вимагають значних зусиль з адаптації до особливостей конкретних ІС. Така адаптація виконується, зазвичай, із використанням інноваційних підходів і методів і залишає невирішеною цілу низку проблем. До цих проблем належать: проблеми безпеки; питання дотримання вимог до ІТ-продукту; складності, що виникають під час реалізації різних варіантів співробітництва, взаємодії ІТ-продукту з навколишнім світом [3].

Головною причиною відсутності формальних засобів розв’язання слід вважати усталене представлення СІ як сукупності апаратних засобів, програмного забезпечення або того й іншого разом. Таке розмите представлення викликає суперечливі вимоги до задачі ідентифікації конфігурації. Наслідком цього є виникнення в ході вирішення задачі ідентифікації конфігурації складної проблеми виділення оптимальної кількості СІ. Під оптимальною тут слід розуміти таку кількість СІ, яка вимагатиме мінімальних витрат часу та ресурсів на виконання подальших робіт процесу управління конфігурацією ІТ-продукту [4]. Ця проблема додатково ускладнена необхідністю простеження (трасування) вимог, що висунуто до ІС, до описів окремих СІ, з яких складається опис архітектури створюваної ІС. Під час подолання цього ускладнення виникає необхідність об’єднання окремих функцій як СІ створюваної ІС в окремі архітектурні сутності, з яких складається загальний опис архітектури створюваної ІС. Використання цих архітектурних сутностей, чия кількість значно менша кількості окремих функцій ІС, суттєво спрощує перехід від процесів формування та аналізу вимог до ІС до процесів проектування цієї ж ІС. Вирішення цієї задачі дозволить значно зменшити витрати часу та зусиль на створення описів архітектури, планування та управління ІТ-проектом створення ІС. Тому проведення досліджень в галузі розробки нових і вдосконалення існуючих методів вирішення задачі ідентифікації СІ ІС залишається одним з актуальних напрямів розвитку комп’ютерних наук.

**Аналіз сучасних досліджень і публікацій.** Як показано у [3], існуючі основи систем управління конфігурацією ІТ-продуктів є надто узагальненими і не надають формальних описів вирішення задачі ідентифікації СІ ІС. До того ж ці положення та моделі є надто узагальненими і вимагають значних зусиль з адаптації до особливостей конкретних ІС.

Одним із підходів до подолання цього недоліку є пропозиція відмовитися від єдиного опису СІ протягом усього життєвого циклу ІС. Така відмова дає змогу розглядати задачу ідентифікації конфігурації ІС як сукупність таких задач [4]:

– задача ранньої ідентифікації СІ (результатом вирішення якої є множина СІ як архітектурних сутностей ІС, виділених за результатами аналізу добре сформульованих вимог до цієї системи);

– задача проєктної ідентифікації СІ (результатом вирішення якої є множина СІ як елементів дизайну ІС, виділених за результатами аналізу архітектурних сутностей цієї системи);

– задача ідентифікації СІ під час реалізації (development) (результатом вирішення якої є множина СІ як елементів програмного та/або апаратного забезпечення ІС, виділених за результатами аналізу СІ як елементів дизайну цієї системи).

Даний підхід спрощує роботи з формування описів СІ на різних стадіях життєвого циклу ІС. Але він залишає невирішеною проблему виділення оптимальної кількості СІ [4].

У [3] також було показано, що вирішення задачі ранньої ідентифікації СІ значною мірою залежить від вирішення проблеми дотримання вимог до ІТ-продукту. В загальному випадку цю залежність можна визначити таким твердженням: задача ранньої ідентифікації СІ як архітектурних сутностей ІС повинна бути вирішеною для описів добре сформульованих вимог до цієї ІС. При цьому під терміном «архітектурна сутність» слід розуміти групи функціональних вимог (функціональні підсистеми і задачі, функціональні модулі, контури управління ресурсами, ІТ-сервіси тощо), які співвідносять функції ІС із бізнес-потребами користувачів цієї ІС [4].

Загальне вирішення задачі ранньої ідентифікації СІ для ІС може здійснюватися на основі таких підходів [4]:

а) функціональна декомпозиція загального опису архітектури ІС на окремі архітектурні сутності та системні елементи, з яких складаються ці сутності;

б) синтез описів архітектурних сутностей та загального опису архітектури з описів окремих системних елементів або вимог до цих елементів.

Приклад використання підходу а) описано у [5]. При цьому поставлена задача ідентифікації конфігурації вирішується у два етапи:

– на першому етапі виконується функціональна декомпозиція опису архітектури на окремі елементи з урахуванням необхідних еволюційних змін;

– на другому етапі здійснюється вибір тих варіантів декомпозиції, які задовольняють обмеження на вартість необхідних змін.

Однак розглянутий у [5] підхід орієнтований переважно на відстеження змін, що вносяться до опису архітектури. При цьому залишаються нерозглянутими питання первинної ідентифікації окремих функцій як СІ у результаті аналізу точок зору на архітектуру цієї системи, а також формування та розвитку архітектурних сутностей ІС як множин СІ, що були виділені під час вирішення задачі.

На відміну від підходу а), сучасні дослідження методів та засобів, які базуються на підході б), з самого початку використовують методи та засоби штучного інтелекту.

Це викликано тим, що, як показано у [6], проблеми раннього поділу системи на окремі елементи виникали, переважно, в результаті неправильного тлумачення вимог або упередженості особистого досвіду. Так, у [7] наведено приклади застосування для моделювання та налаштування функцій системи таких інструментів штучного інтелекту, як рекомендаційні системи та методи машинного навчання. Але хоча розглянуті приклади враховують чинник мінливості описів окремих функцій, питання зв'язності цих функцій та утворення з них архітектурних сутностей залишаються невирішеними.

У [8] для вирішення проблеми вибору конфігурації системи запропоновано використовувати семплер на основі глибокого навчання з підкріпленням. Результатом застосування цього семплера є ідентифікована мінімальна підмножина конфігурацій, яка охоплює всі варіанти взаємодії функцій при мінімізації надмірності. Але таке рішення для ІС, призначених для управління стабільними процесами підприємств та організацій, є надмірно складним.

Часто для вирішення задач ідентифікації СІ в межах підходу б) дослідники застосовують методи кластеризації. Як приклади подібного застосування методів та алгоритмів кластеризації можна вказати дослідження з вирішення таких задач:

- виділення мікросервісів у ході рефакторингу вихідного коду монолітного програмного застосування [9];
- рефакторингу складних багаторівневих монолітних програмних систем у ході їх декомпозиції на мікросервіси [10];
- ранньої ідентифікації беклогів команд виконавців як СІ системи [11].

Але в розглянутих дослідженнях не враховуються зв'язки між окремими функціями ІС, які обумовлюють групування функцій у архітектурні сутності ІС. Взагалі питання про фактори, які визначають функцію мети, обмеження та спосіб розрахунку відстані між окремими функціями ІС, залишається й досі відкритим.

Використання методів кластеризації для вирішення задачі ранньої ідентифікації вимагає також формалізації вимог, які висуваються до ІС. У [12] було проведено систематичний огляд сучасних наукових досліджень інструментів, методів, прийомів та процесів вилучення вимог із середовищ, що базуються на методології Model-Based System Engineering, та генерації моделей на основі вимог, сформульованих природною мовою. За результатами цього огляду було встановлено, зокрема, що генерація моделей на основі описів, зроблених природною мовою, залишається домінуючим способом формального опису вимоги до ІС. При цьому використання великих мовних моделей під час такої генерації вимагає більше даних, ніж можуть надати сучасні методи формування подібних текстових описів, і є недоцільним [12]. Але питання використання моделей вимог як основи для створення описів СІ ІС у [12] залишилося нерозглянутим.

Проведений аналіз розглянутих публікацій [3–12] дозволив зробити такі висновки:

- вирішення задачі ранньої ідентифікації СІ ІС бажано здійснювати з використанням підходу б), тому що створити загальний опис архітектури ІС для подальшої декомпозиції в багатьох випадках неможливо;

– для вирішення зазначеної задачі краще за все використовувати формальні методи штучного інтелекту і, зокрема, методи кластеризації;

– як похідні дані для вирішення зазначеної задачі слід використовувати моделі вимог, які висуваються до ІС, зокрема – формальні моделі функціональних вимог до ІС на основі їхніх онтологічних описів.

**Мета дослідження.** Метою дослідження є розробка нового або вдосконалення існуючого методу вирішення задачі ранньої ідентифікації СІ ІС. Це дозволить автоматизувати процедуру синтезу опису архітектури інформаційної системи. Застосування цієї автоматизованої процедури сприяє підвищенню якості розробки інформаційної системи за рахунок виділення для проектування множини архітектурних сутностей цієї системи, яка є значно меншою, ніж множина елементарних функцій інформаційної системи.

**Основна частина.** Об'єктом даного дослідження є процес управління конфігурацією ІТ-проєкту. Предметом дослідження є методи вирішення задачі ранньої ідентифікації СІ ІС.

В процесі дослідження було вирішено розглядати рішення задачі ранньої ідентифікації як результат послідовного вирішення таких підзадач:

а) підзадача синтезу описів раціональної архітектури створюваної ІС на основі описів добре сформульованих функціональних вимог до цієї ІС (Підзадача 1);

б) підзадача синтезу описів архітектурних сутностей для описів раціональної архітектури створюваної ІС, які є результатами вирішення Підзадачі 1 (Підзадача 2).

Підзадачу 1 у [13] запропоновано описати теоретико-ігровою моделлю вигляду

$$\Gamma_{IS} = \langle \{Pr, U\}, \{X^{Pr,U}\}, \{f^{Pr,U}\} \rangle, \quad (1)$$

де  $\Gamma_{IS}$  – гра Постачальника та Споживача ІТ-послуг (функцій ІС);  $\{Pr, U\}$ , – множина гравців, які беруть участь у грі  $\Gamma_{IS}$  (Постачальник та Споживач ІТ-послуг відповідно);  $\{X^{Pr,U}\}$  – множина стратегій гри  $\Gamma_{IS}$ , яка формується на основі множини варіантів описів архітектури створюваної ІС  $\Pi_{acm}$ ;  $\{f^{Pr,U}\}$  – множина функцій виграшу гри  $\Gamma_{IS}$  [13].

Гра (1) є скінченною безкоаліційною грою з нульовою сумою Постачальника та Споживача ІТ-послуг. Детальний опис елементів гри (1) та формальних моделей описів вимог до окремих функцій ІС для даної гри наведено у [13]. Для знаходження результату гри (1) у [13] запропоновано використовувати метод пошуку рівноваг Неша для біматричної гри в чистих стратегіях. Цей метод передбачав послідовність таких етапів.

Етап 1. Сформувані матриці виграшів Споживача  $U$  й Постачальника ІТ-послуг  $Pr$ .

Етап 2. Відзначити максимальні елементи в кожному стовпці матриці  $U$ . Якщо в деякому стовпці даної матриці кілька максимальних елементів, то необхідно відзначити всі такі елементи.

Етап 3. Відзначити максимальні елементи в кожному рядку матриці  $Pr$ . Якщо в деякому рядку даної матриці кілька максимальних елементів, то необхідно відзначити всі такі елементи.

Етап 4. Провести пошук перетинання результатів Етапу 2 і Етапу 3 (тобто максимальних елементів, які перебувають на тих самих позиціях у кожній з матриць).

Етап 5. Якщо результат виконання Етапу 4 не є порожньою множиною, зафіксувати відповідні варіанти описів архітектури створюваної ІС як описи раціональної архітектури даної системи  $\{K_{ji}^{f_{js}}\}_{i=1, \dots, e}^{(Pr, U)}$ , де  $e$  – кількість добре сформульованих функціональних вимог до ІС. У іншому випадку визнати, що опису раціональної архітектури ІС на чистих стратегіях не існує. Завершити застосування методу.

Результатом використання цього методу є опис архітектури створюваної ІС як множини окремих функцій цієї ІС. Це означає, що як окремі СІ ІС обрано опис окремої функції (або вимоги до цієї функції). Але таке вирішення Підзадачі 1 визнано недоліком, тому що кількість окремих СІ у великих та середніх ІС управління підприємствами та організаціями може сягати кількох сотень. Це значно ускладнює подальший розподіл опису архітектури ІС між окремими командами виконавців ІТ-проєкту розробки цієї системи.

Для ліквідації даного недоліку під час вирішення Підзадачі 2 запропоновано зменшити кількість СІ в описі архітектури ІС. В загальному випадку це зменшення може бути описано як функція мети  $J$ , що має вигляд [4, 14]

$$J(M, d, U, C) = \sum_{i=1}^c \sum_{j=1}^k u_{ij} d(m_j, c^{(i)}) \rightarrow \min, \quad (2)$$

де  $M$  – базова множина похідних даних;  $m_j$  –  $j$ -та точка (вектор) похідних даних;  $k$  – кількість точок (векторів) похідних даних;  $d$  – спосіб визначення відстані між точками (векторами) похідних даних;  $U$  – матриця розбиття точок (векторів) похідних даних між кластерами  $c^{(i)}$ ;  $u_{ij}$  – елемент матриці розбиття, який встановлює належність точки (вектору)  $m_j$  до кластеру  $c^{(i)}$ ;  $C$  – множина кластерів, які є результатами вирішення задачі ранньої ідентифікації;  $c^{(i)}$  –  $i$ -й кластер, який є одним з результатів вирішення задачі ранньої ідентифікації СІ ІС і описує виявлену архітектурну сутність ІС як підмножину окремих функцій;  $c$  – кількість сформованих кластерів.

Набір обмежень для функції мети (2) має вигляд [4, 14]:

$$u_{ij} \in \{0, 1\}; \sum_{i=1}^c u_{ij} = 1; \sum_{j=1}^k u_{ij} \leq k. \quad (3)$$

Ці обмеження встановлюють, що:

- кожна похідна точка (вектор) може належати тільки одному з кластерів та не належати іншим кластерам;
- в кожному кластері міститься не менше однієї похідної точки (вектору), але не менше загальної кількості похідних точок (векторів)  $k$ .

Як основний алгоритм вирішення Підзадачі 2 у [4] було запропоновано використовувати алгоритм щільнісної кластеризації DBSCAN.

Вибір групи методів щільнісної кластеризації обумовлений такими міркуваннями [14]:

- ці методи, на відміну від багатьох інших методів кластеризації (partitioning methods, ієрархічні методи тощо), можуть знаходити кластери довільної форми;
- ці методи визначають кластери як щільні області об'єктів досліджуваної вибірки, розділені областями з низькою щільністю (що якнайкраще відповідає визначенню архітектурної сутності);
- ці методи можуть відфільтрувати аномалії (тобто описи окремих функцій ІС, включення яких в загальний опис архітектури цієї системи недоцільне із заздалегідь визначених аналітиком причин).

У [14] визнано існування кращих за DBSCAN методів та алгоритмів кластеризації, які дозволяють виявляти кластери довільної форми високої якості (зокрема, метод агломеративної кластеризації Chameleon). Але алгоритм DBSCAN простіший у реалізації і вимагає менших обчислювальних витрат і витрат часу на вирішення задачі кластеризації [15]. Крім того, однією з переваг алгоритму DBSCAN є можливість обробляти дані різноманітної природи. Як показано в [15] з посиланнями на відповідні дослідження: «...DBSCAN ніколи не обмежувався використанням евклідової відстані або точок у  $\mathbb{R}^d$ , але завжди призначався для використання з географічними даними, полігонами та іншими типами даних».

У [4] було проведено дослідження з порівняльного аналізу результатів вирішення задачі ранньої ідентифікації СІ, отриманих з використанням щільнісного алгоритму DBSCAN, методів ієрархічної кластеризації DIANA, AGNES та Chameleon, а також алгоритму k-means. Для порівняння було запропоновано критерії «Громізdkість рішення» та «Виявлення відокремлених СІ». Крім цього, було проведено аналіз збігу окремих кластерів, розроблених з використанням зазначених методів та алгоритмів. За результатами порівняльного аналізу, алгоритм DBSCAN визнано кращим з тих методів та алгоритмів кластеризації, які брали участь у аналізі. Використання цього алгоритму дає змогу отримати описи архітектурних сутностей ІС як кластерів, що складаються з описів окремих функцій ІС, схожих за описами їхніх структур даних [4].

Використання алгоритму DBSCAN для вирішення Підзадачі 2 дозволило вдосконалити метод вирішення задачі ранньої ідентифікації СІ ІС. Цей вдосконалений метод представлено як послідовність виконання таких етапів.

Етап 0 (підготовчий). Сформувані множини формальних описів добре сформульованих вимог до ІС, які відображають точки зору Споживача й Постачальника ІТ-послуг [3].

Етап 1. Сформувані матриці вирашів Споживача й Постачальника ІТ-послуг.

Етап 2. Відзначити максимальні елементи в кожному стовпці матриці вирашів Споживача ІТ-послуг. Якщо в деякому стовпці даної матриці кілька максимальних елементів, то необхідно відзначити всі такі елементи.

Етап 3. Відзначити максимальні елементи в кожному рядку матриці виграшів Постачальника ІТ-послуг. Якщо в деякому рядку даної матриці кілька максимальних елементів, то необхідно відзначити всі такі елементи.

Етап 4. Провести пошук перетинання результатів Етапу 2 і Етапу 3.

Етап 5. Якщо результат виконання Етапу 4 не є порожньою множиною, зафіксувати відповідні варіанти описів архітектури створюваної ІС як описи раціональної архітектури даної ІС. У іншому випадку визнати, що опису раціональної архітектури ІС на чистих стратегіях не існує, та завершити застосування методу.

Етап 6. Виділити на зафіксованих описах окремих варіантів описів архітектури створюваної ІС описи архітектурних сутностей ІС як результат вирішення задачі класифікації описів окремих функцій даної ІС з використанням алгоритму щільнісної класифікації DBSCAN [4].

Етап 7. Виділити з результатів виконання Етапу 6 спрощені варіанти опису архітектури створюваної ІС за критерієм мінімальної наявності в описах архітектури ІС, визначених як шум, і надати ці описи Споживачу та Постачальнику ІТ-послуг. Завершити застосування методу.

У запропонованому методі Етап 0 відповідає за формування формальних описів функціональних вимог до створюваної ІС. Етапи 1-5 відповідають за вирішення Підзадачі 1, результатом якого є опис раціональної архітектури створюваної ІС на основі описів окремих функцій, що реалізують функціональні вимоги до цієї ІС. Етапи 6 та 7 відповідають за вирішення Підзадачі 2, результатом якого є спрощений опис архітектури створюваної ІС на основі виділених архітектурних сутностей як кластерів функцій цієї системи.

Результати експериментальної перевірки вирішення задачі ранньої ідентифікації СІ ІС наведено у [4]. Для перевірки було використано дані ІТ-проекту розробки функціональної задачі «Формування та ведення індивідуального плану науково-педагогічного працівника кафедри». Цю задачу реалізовано у вигляді окремого ІТ-продукту для розвитку можливостей інформаційно-аналітичної системи «Університет» Харківського національного університету радіоелектроніки (Україна). Під час перевірки для зменшення витрат часу було висунуто припущення про те, що в результаті виконання Етапів 0-5 було отримано опис раціональної архітектури функціональної задачі, який складався з десяти окремих функцій. Найменування цих функцій наведено у табл. 1 [4, 11].

Детально описи функціональної задачі «Формування та ведення індивідуального плану науково-педагогічного працівника кафедри» наведено у [11]. Для формального опису функцій задачі було використано діаграми потоків даних та діаграми «Сутність – зв'язок». Для визначення відстаней між окремими функціями задачі було використано модифіковану відстань Чебишева [4, 11]. Усі розрахунки під час вирішення задачі ранньої ідентифікації СІ ІС із застосуванням алгоритму DBSCAN виконано в середовищі MS Excel 2016.

Найменування окремих функцій функціональної задачі «Формування та ведення індивідуального плану науково-педагогічного працівника кафедри»

Функція	
Найменування	Позначення
Конвертація розділу «Навчальна робота»	<i>CI1</i>
Формування розділу «Наукова робота»	<i>CI2</i>
Формування розділу «Методична робота»	<i>CI3</i>
Формування розділу «Організаційна робота»	<i>CI4</i>
Формування переліку посад та довгострокових доручень	<i>CI5</i>
Формування переліку рекомендованих до виконання робіт	<i>CI6</i>
Формування і ведення нормативно-довідкової інформації про Key Performance Indicators (KPI)	<i>CI7</i>
Формування KPI викладача і частини від KPI кафедри	<i>CI8</i>
Формування зведеної таблиці на навчальний рік	<i>CI9</i>
Формування вихідного документа «Індивідуальний план»	<i>CI10</i>

Під час вирішення задачі ранньої ідентифікації було розглянуто три варіанти опису архітектури функціональної задачі: монолітна архітектура; модульна архітектура; сервіс-орієнтована архітектура [4]. Для цих варіантів під час виконання Етапу 6 було отримано вісім різних рішень. З них за критерієм мінімальної наявності в описах кожного з варіантів архітектури тих CI, що були визначені як шум, під час виконання Етапу 7 було обрано три варіанти рішення, які наведено у табл. 2.

Таблиця 2

Результати вирішення задачі ранньої ідентифікації елементів конфігурації функціональної задачі «Формування та ведення індивідуального плану науково-педагогічного працівника кафедри»

Значення <i>MinPts</i>	Значення $\epsilon$	Рішення (позначення та зміст кластерів)
<i>MinPts</i> = 1 (сервіс-орієнтована архітектура)	$\epsilon$ = 60%	$C1 = \{CI7, CI6\}$
		$C2 = \{CI4, CI1, CI2, CI3, CI8, CI9, CI10, CI5\}$
<i>MinPts</i> = 2 (модульна архітектура)	$\epsilon$ = 60%	$C1 = \{CI4, CI1, CI2, CI3, CI8, CI9, CI10, CI5\}$
		<i>CI6, CI7</i> визнано шумом
<i>MinPts</i> = 1 (монолітна архітектура)	$\epsilon$ = 70%	$C1 = \{CI1, CI2, CI3, CI4, CI5, CI6, CI7, CI8, CI9, CI10\}$

У табл. 2 використані такі параметри алгоритму DBSCAN: *MinPts* – параметр, який повинен визначати поріг щільності найближчих сусідів досліджуваного CI;  $\epsilon$  – ра-

діус, який повинен визначати межу області існування найближчих сусідів досліджуваних СІ (так звану міру близькості).

Слід визнати, що вдосконалений метод вирішення задачі ранньої ідентифікації має обмеження у застосуванні. Головним обмеженням є необхідність формального опису кожної добре сформульованої функціональної вимоги до створюваної ІС. Але можливості формування подібних описів на ранніх стадіях життєвого циклу ІС є занадто обмеженими. Тому головним напрямом подальших досліджень було визнано адаптацію способів обчислення відстані  $d$  між описами окремих функціональних вимог та функцій ІС до можливих варіантів формального опису цих вимог та функцій. Ще одним напрямом подальших досліджень слід визнати розробку моделі універсального опису функціональної вимоги та функції ІС, яка б могла враховувати можливість існування різноманітних описів цієї вимоги (або функції) з різними ступенями формальності.

**Висновки.** В результаті дослідження було вдосконалено метод вирішення задачі ранньої ідентифікації СІ ІС. Суть вдосконалення полягає у використанні алгоритму щільнісної кластеризації DBSCAN для виявлення архітектурних елементів як кластерів окремих функцій ІС. Використання вдосконаленого методу дозволяє автоматизувати процедуру синтезу опису архітектури ІС. Застосування цієї автоматизованої процедури сприяє підвищенню якості розробки ІС за рахунок виділення для проєктування множини архітектурних сутностей цієї системи, яка є значно меншою, ніж множина елементарних функцій цієї ж ІС.

#### ЛІТЕРАТУРА

1. Guide to the Software Engineering Body of Knowledge. Version 3.0 / eds. Bourque P., Fairley R.E. Washington, DC, United States: IEEE Computer Society, 2014. 335 p.
2. Quigley J.M., Robertson K.L. Configuration Management, Second Edition: Theory and Application for Engineers, Managers, and Practitioners. New York: Auerbach Publications, 2019. 453 p. DOI: <https://doi.org/10.1201/9780429318337>
3. Farayola O.A., Hassan A.O., Adaramodu O.R., Fakeyede O.G., Oladeinde M. Configuration management in the modern era: best practices, innovations, and challenges. *Computer Science & IT Research Journal*. 2023. Vol. 4, No 2. P. 140-157. DOI: <https://doi.org/10.51594/csitrj.v4i2.613>
4. Kozhanov A., Ievlanov M. Definition a clusterization method to solve the task of early identification of IT product configuration elements. *Eastern-European Journal of Enterprise Technologies*. 2026. Vol. 1. №. 2 (139). P. 77–90. DOI: <https://doi.org/10.15587/1729-4061.2026.352878>
5. Faitelson D., Heinrich R., Tyszberowicz Sh. Supporting software architecture evolution by functional decomposition. In *5th International Conference on Model-Driven Engineering and Software Development – MODELSWARD*. SciTePress, 2017. P. 435-442. DOI: <https://doi.org/10.5220/0006206204350442>

6. Cadavid H., Andrikopoulos V., Avgeriou P., Chris Broekema P. System and software architecting harmonization practices in ultra-large-scale systems of systems: A confirmatory case study. *Information and Software Technology*. 2022. Vol. 150. № 106984. DOI: <https://doi.org/10.1016/j.infsof.2022.106984>
7. Felfernig A., Le V., Popescu A., Uta M., Tran T.N., Atas M. (2021). An Overview of Recommender Systems and Machine Learning in Feature Modeling and Configuration. *Proceedings of the 15th International Working Conference on Variability Modelling of Software-Intensive Systems (VaMoS '21)*. New York, NY, USA: Association for Computing Machinery, 2021. Article 16, P. 1–8. DOI: <https://doi.org/10.1145/3442391.3442408>
8. Abolfazli A., Spiegelberg J., Palmer G., Anand A. A Deep Reinforcement Learning Approach to Configuration Sampling Problem. *2023 IEEE International Conference on Data Mining (ICDM)*. Shanghai, China, 2023. P. 1-10. DOI: <https://doi.org/10.1109/ICDM58522.2023.00009>.
9. Sellami Kh., Saied M.A., Ouni A. A Hierarchical DBSCAN Method for Extracting Microservices from Monolithic Applications. *EASE '22: Proceedings of the 26th International Conference on Evaluation and Assessment in Software Engineering*. New York, NY, USA: Association for Computing Machinery, 2022. P. 201-210. DOI: <https://doi.org/10.1145/3530019.3530040>
10. Krause A., Zirkelbach C., Hasselbring W., Lenga S., Kroger D. Microservice Decomposition via Static and Dynamic Analysis of the Monolith. *2020 IEEE International Conference on Software Architecture Companion (ICSA-C)*. Salvador, Brazil, 2020. P. 9-16. DOI: <https://doi.org/10.1109/ICSA-C50368.2020.00011>
11. Ievlanov, M., Vasilcova, N., Neumyvakina, O., Panforova, I. Development of a method for solving the problem of it product configuration analysis. *Eastern-European Journal of Enterprise Technologies*. 2022. Vol. 6. Iss. 2 (120). P. 6–19. DOI: <https://doi.org/10.15587/1729-4061.2022.269133>
12. Santos J.L., Martins L.E.G., Molléri J.S. Requirements extraction from model-based systems engineering: A systematic literature review. *Journal of Systems and Software*. 2025. Vol. 226. № 112407. DOI: <https://doi.org/10.1016/j.jss.2025.112407>
13. Evlanov M. Development of the model and method of selecting the description of rational architecture of information system. *Eastern-European Journal of Enterprise Technologies*. 2016. Vol. 1, No. 2(79). P. 4–12. DOI: <https://doi.org/10.15587/1729-4061.2016.60583>
14. Han J., Kamber M., Pei J. *Data Mining. Concepts and Techniques*. (Third Edition). Waltham: Morgan Kaufmann Publishers. 2011. URL: <https://shop.elsevier.com/books/data-mining-concepts-and-techniques/han/978-0-12-381479-1> (дата звернення 21.03.2026).
15. Schubert E., Sander J., Ester M., Kriegel H.P., Xu, X. DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. *ACM Transactions on Database Systems*. 2017. Vol. 42. Iss. 3. Article No.: 19, P. 1-21. DOI: <https://doi.org/10.1145/3068335>.

#### REFERENCES

1. Guide to the Software Engineering Body of Knowledge. Version 3.0 / eds. Bourque P., Fairley R.E. Washington, DC, United States: IEEE Computer Society, 2014. 335 p.

2. Quigley J.M., Robertson K.L. Configuration Management, Second Edition: Theory and Application for Engineers, Managers, and Practitioners. New York: Auerbach Publications, 2019. 453 p. DOI: <https://doi.org/10.1201/9780429318337>
3. Farayola O.A., Hassan A.O., Adaramodu O.R., Fakeyede O.G., Oladeinde M. Configuration management in the modern era: best practices, innovations, and challenges. *Computer Science & IT Research Journal*. 2023. Vol. 4, No 2. P. 140-157. DOI: <https://doi.org/10.51594/csitrj.v4i2.613>
4. Kozhanov A., Ievlanov M. Definition a clusterization method to solve the task of early identification of IT product configuration elements. *Eastern-European Journal of Enterprise Technologies*. 2026. Vol. 1. №. 2 (139). P. 77–90. DOI: <https://doi.org/10.15587/1729-4061.2026.352878>
5. Faitelson D., Heinrich R., Tyszberowicz Sh. Supporting software architecture evolution by functional decomposition. In *5th International Conference on Model-Driven Engineering and Software Development – MODELSWARD*. SciTePress, 2017. P. 435-442. DOI: <https://doi.org/10.5220/0006206204350442>
6. Cadavid H., Andrikopoulos V., Avgeriou P., Chris Broekema P. System and software architecting harmonization practices in ultra-large-scale systems of systems: A confirmatory case study. *Information and Software Technology*. 2022. Vol. 150. № 106984. DOI: <https://doi.org/10.1016/j.infsof.2022.106984>
7. Felfernig A., Le V., Popescu A., Uta M., Tran T.N., Atas M. (2021). An Overview of Recommender Systems and Machine Learning in Feature Modeling and Configuration. *Proceedings of the 15th International Working Conference on Variability Modelling of Software-Intensive Systems (VaMoS '21)*. New York, NY, USA: Association for Computing Machinery, 2021. Article 16, P. 1–8. DOI: <https://doi.org/10.1145/3442391.3442408>
8. Abolfazli A., Spiegelberg J., Palmer G., Anand A. A Deep Reinforcement Learning Approach to Configuration Sampling Problem. *2023 IEEE International Conference on Data Mining (ICDM)*. Shanghai, China, 2023. P. 1-10. DOI: <https://doi.org/10.1109/ICDM58522.2023.00009>.
9. Sellami Kh., Saied M.A., Ouni A. A Hierarchical DBSCAN Method for Extracting Microservices from Monolithic Applications. *EASE '22: Proceedings of the 26th International Conference on Evaluation and Assessment in Software Engineering*. New York, NY, USA: Association for Computing Machinery, 2022. P. 201-210. DOI: <https://doi.org/10.1145/3530019.3530040>
10. Krause A., Zirkelbach C., Hasselbring W., Lenga S., Kroger D. Microservice Decomposition via Static and Dynamic Analysis of the Monolith. *2020 IEEE International Conference on Software Architecture Companion (ICSA-C)*. Salvador, Brazil, 2020. P. 9-16. DOI: <https://doi.org/10.1109/ICSA-C50368.2020.00011>
11. Ievlanov, M., Vasilcova, N., Neumyvakina, O., Panforova, I. Development of a method for solving the problem of it product configuration analysis. *Eastern-European Journal of Enterprise Technologies*. 2022. Vol. 6. Iss. 2 (120). P. 6–19. DOI: <https://doi.org/10.15587/1729-4061.2022.269133>

12. Santos J.L., Martins L.E.G., Molléri J.S. Requirements extraction from model-based systems engineering: A systematic literature review. *Journal of Systems and Software*. 2025. Vol. 226. № 112407. DOI: <https://doi.org/10.1016/j.jss.2025.112407>
13. Evlanov M. Development of the model and method of selecting the description of rational architecture of information system. *Eastern-European Journal of Enterprise Technologies*. 2016. Vol. 1, No. 2(79). P. 4–12. DOI: <https://doi.org/10.15587/1729-4061.2016.60583>
14. Han J., Kamber M., Pei J. *Data Mining. Concepts and Techniques*. (Third Edition). Waltham: Morgan Kaufmann Publishers. 2011. URL: <https://shop.elsevier.com/books/data-mining-concepts-and-techniques/han/978-0-12-381479-1> (дата звернення 21.03.2026).
15. Schubert E., Sander J., Ester M., Kriegel H.P., Xu, X. DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. *ACM Transactions on Database Systems*. 2017. Vol. 42. Iss. 3. Article No.: 19, P. 1-21. DOI: <https://doi.org/10.1145/3068335>.

Received 30.03.2026.

Accepted 02.04.2026.

Published 30.04.2026

### ***Improvement of the method for solving the task of early identification of information system configuration items***

*This study investigates the process that controls the IT project configuration.*

*The study addressed the problem of early identification of configuration items (CI) within an enterprise management information system (IS). Modern research in this area is mainly aimed at solving the task of services early identification during refactoring of software systems. The issue of applying artificial intelligence methods and, in particular, clustering to detect CI at the early stages of the IS life cycle has not been studied sufficiently.*

*The aim of the study is to develop a new or improve an existing method for solving the task of early identification of IS CI.*

*To achieve the aim of the study, the existing method of synthesizing descriptions of rational architecture creating IS based on descriptions of functional requirements for this IS was improved. The essence of this improvement is to use the DBSCAN clustering algorithm to detect architectural entities as clusters of individual IS functions. The results of solving the clustering task make it possible to simplify the description of the architecture creating IS by significantly reducing the number of CIs, provided that descriptions of architectural entities as clusters of individual IS functions will act as CIs.*

*The main features of the course and results of the experimental verification of the improved method are considered. The descriptions of ten functions of the task “Formation and maintenance of an individual plan of a scientific and pedagogical employee of the department” were used as the derived data for this verification. The result of using the method is three solution options, presented in the form of sets of one or two clusters. These options describe monolithic, modular and service-oriented IS architecture.*

*The use of the obtained results allows you to automate the procedure for synthesizing the description of the IS architecture. This automation will increase the quality of IS development by allocating a set of architectural entities of this IS for design. This set is much smaller than the set of elementary IS functions.*

*Keywords: information system, configuration item, early identification task, function, architectural entity, DBSCAN, architecture description.*

**Кожанов Адріан Євгенійович** – аспірант кафедри інформаційних управляючих систем, Харківський національний університет радіоелектроніки.

ORCID <https://orcid.org/0009-0002-3586-6886>

**Євланов Максим Вікторович** – д.т.н., професор, професор кафедри інформаційних управляючих систем, Харківський національний університет радіоелектроніки.

ORCID <https://orcid.org/0000-0002-6703-5166>

**Мороз Дмитро Максимович** – доктор філософії, доцент кафедри програмного забезпечення комп'ютерних систем, Національний технічний університет «Дніпровська політехніка».

ORCID <https://orcid.org/0000-0003-2577-3352>

**Kozhanov Adrian Yevgeniyovich** – postgraduate student of the Department of Information Control Systems, Kharkiv National University of Radio Electronics.

ORCID <https://orcid.org/0009-0002-3586-6886>

**Ievlanov Maksym Viktorovich** – doctor of technical sciences, prof., professor of the Department of information control systems, Kharkiv National University of Radio Electronics.

ORCID <https://orcid.org/0000-0002-6703-5166>

**Moroz Dmytro Maksymovich** – PhD of technical sciences, Associate Professor of the Department of Software Engineering, Dnipro University of Technology.

ORCID <https://orcid.org/0000-0003-2577-3352>

## COMPARATIVE ANALYSIS OF STATE MANAGEMENT MODELS IN MOBILE APPLICATIONS DEVELOPED USING FLUTTER

*Annotation. The objective of this study is to formalize state management models in Flutter applications and evaluate them against criteria of transition determinism, performance, scalability, and architectural complexity. State management approaches are interpreted as extended transition systems with asynchronous events; safety and liveness properties are formalized using temporal logic. A formalized Flutter state management model is proposed as a class of reactive computational systems with defined safety and liveness properties. In addition, the study incorporates benchmark-based analysis of frame rendering time and CPU utilization to provide an empirical comparison of the considered approaches. The obtained results are further interpreted in the context of real-world development scenarios, enabling the formulation of practical recommendations for selecting appropriate state management architectures depending on application scale and complexity. The study revealed significant differences among approaches in determinism, performance, and architectural complexity, demonstrating that Riverpod provides the most balanced performance characteristics due to efficient dependency tracking, while Bloc and Cubit ensure higher structural clarity at the cost of increased computational overhead. MobX shows moderate performance with advantages in managing complex interdependent state, whereas Provider remains suitable for small-scale applications due to its simplicity. These findings enable informed architecture selection based on production application requirements, including system scale, performance constraints, and maintainability needs.*

*Keywords: state management, Flutter, reactive systems, formalization, Cubit, Riverpod, Provider, MobX, performance, scalability.*

**Problem statement.** With the growing complexity of mobile applications, the volume of dynamic data and the number of interactions between interface components are also increasing. Modern mobile applications require efficient and scalable mechanisms for managing dynamic data and user interface interactions.

Flutter, as a cross-platform framework, provides several state management solutions that differ in architecture, performance, and developer experience (Flutter Official Docs, 2025). The choice of the most appropriate state management approach significantly affects the stability, performance, and maintainability of the application. In the Flutter environment, incorrect or irrational application of state management methods can lead to performance degradation, architectural complexity, increased error rates, and difficulties in project scaling (Sol-Guruz, 2025).

Furthermore, the diversity of state management approaches creates a need for their systematic analysis and comparison in terms of efficiency, stability, and usability. In this regard, the study of methods and algorithms for mobile application state management is a relevant task of modern software engineering.

**Analysis of recent research and publications.** The issue of comparing state management approaches in Flutter has been explored in a number of scientific and practical publications. In particular, a paper published in the JOTI (Jurnal Online Teknik Informatika) journal of Universitas Dinamika conducted a comparative performance analysis of Provider and Riverpod libraries for datasets of 1,000 to 10,000 elements, with results showing a stable advantage of Riverpod in CPU usage and memory efficiency (JOTI, 2023).

This research is important because it is based on quantitative measurements in a controlled environment, rather than merely qualitative analysis of architectural characteristics. Another significant contribution to this field was made by the iCoderz Solutions platform, which published a comparative benchmark of five state management libraries, including Provider, Riverpod, Bloc, GetX, and MobX. The researchers measured frame rendering time and CPU load under identical workloads – list element operations – and established that Riverpod demonstrates the lowest frame rendering time (12 ms) and the lowest CPU load among the reviewed solutions (iCoderz, 2025).

Shakil examines in detail the architectural features of Bloc and Cubit, emphasizing that Bloc's event-driven model, although introducing greater overhead, ensures strict separation of concerns, which is critical in large team projects (Shakil, 2025). Kumar in his comparative review analyzes practical implementation cases of Provider, Bloc, and Riverpod in production projects of various scales – from e-commerce startups to enterprise fintech applications – and identifies criteria for choosing between them depending on project specifics (Kumar, 2025).

The analysis of MobX's reactive approach in the context of Flutter was carried out in the Foresight Mobile review, where special attention is paid to the transparency of the dependency tracking mechanism and its comparison with declarative models of Provider and Riverpod (Foresight Mobile, 2025). MobX's transparent reactive programming paradigm, originally developed for JavaScript, provides a scientific basis in the work of Meijer (2010), which formally grounds reactive programming as a computational model.

At the same time, most of the analyzed publications focus either on individual libraries or on comparing two or three solutions, without systematically covering the entire current spectrum of tools. This confirms the need for a comprehensive comparative analysis covering Provider, Riverpod, Cubit, and MobX within a single frame of reference.

**Research objective.** Despite the availability of a large number of state management approaches in Flutter, there are gaps in the systematic comparison of their efficiency, developer ergonomics, and impact on application performance. Many studies are limited to describing individual patterns or comparing only popular approaches, without considering the comprehensive impact on stability, scalability, and code maintainability (Foresight Mobile, 2025). The objective of this study is to conduct a comprehensive analysis of modern state management methods in Flutter – Provider, Riverpod, Cubit, and MobX – from the perspective of

their architectural characteristics, performance metrics, and usability, as well as to determine optimal usage scenarios for each approach for different types of mobile applications. The research results should provide developers with practical recommendations for selecting state management approaches, thereby improving the quality and stability of the software product (Kumar, 2025).

**Main research findings.** Since direct performance measurements require experimental implementation, we refer to existing benchmark studies for comparing state management solutions relevant to Flutter. Performance research shows that state management libraries differ in the efficiency of interface rendering processing and CPU load under similar workloads (iCoderz, 2025; JOTI, 2023).

These differences are not merely incidental: they reflect fundamental architectural decisions embedded in each library's design, specifically, how and when state changes are propagated to the widget tree. Libraries that rely on fine-grained reactive tracking tend to minimize unnecessary rebuilds, whereas those built around explicit event streams introduce structural overhead that, while beneficial for traceability and testability, carries a measurable runtime cost. Understanding the quantitative magnitude of these trade-offs is essential for making informed architectural decisions in production Flutter development.

The benchmark methodology employed in the referenced studies involves constructing functionally equivalent application prototypes, isomorphic in their UI structure and data operations, and subjecting each to identical workloads, typically list rendering and state update operations across datasets of varying size. This controlled approach isolates library-specific overhead from application logic, yielding comparable baseline metrics. According to this benchmark data, average frame rendering times (in milliseconds) and CPU usage trends for common Flutter state management libraries are presented in Table 1:

Table 1

Comparative performance metrics of Flutter state management libraries

State Management Solution	Frame Rendering Time (ms)	CPU Usage
Provider	~16 ms	Moderate
Riverpod	~12 ms	Low
MobX	~18 ms	Medium
Bloc/Cubit	~20 ms	High

These data suggest that Riverpod demonstrates comparatively more efficient UI update performance under similar conditions, providing lower frame rendering time and lower CPU load. The formal theoretical basis for such state-transition behavior can be found in the timed automata model (Alur & Dill, 1994), which provides a mathematical framework for reasoning about asynchronous event-driven transitions. MobX shows moderate performance with acceptable response times and slightly higher rendering times, while Bloc (and consequently

Cubit as its simplified variant) produces higher rendering times and CPU load due to the overhead of event and stream processing (iCoderz, 2025).

Additional studies focusing on comparing Provider and Riverpod show that Riverpod consistently maintains a slight advantage in CPU usage and memory efficiency compared to Provider under identical application workloads across various data volumes (1,000–10,000 elements) (JOTI, 2023). Although these benchmarks do not measure Cubit separately from Bloc, it is widely known that Bloc (and Cubit as part of the Bloc family) provides a structured event-driven state flow at the cost of additional processing, which may affect performance in scenarios with frequent state changes (Shakil, 2025).

**Provider: architectural features and practical application.** Provider is one of the most widely used and officially recommended state management solutions in Flutter. This library is built on top of the built-in `InheritedWidget` mechanism and significantly simplifies its usage, providing a convenient API for dependency injection and state propagation through the widget tree (Flutter Official Docs, 2025).

Due to its conceptual simplicity and native integration with the Flutter SDK, Provider has gained widespread use primarily in small and medium-scale projects. From a technical standpoint, Provider implements the `ChangeNotifier` pattern, in which the state object inherits the `ChangeNotifier` class and calls the `notifyListeners()` method on each change. Widgets registered via `Consumer` or `Provider.of()` are automatically rebuilt on each such notification. The advantage of this approach is its transparency and minimal boilerplate code. In practice, Provider works best in scenarios such as managing a user profile in an authentication application, synchronizing application theme across multiple interface components, or managing a simple shopping cart (Adam, 2025). However, Provider has significant limitations in large and architecturally complex projects: it lacks compile-time safety and is prone to excessive rebuilds unless additional optimizations – such as precise `Selectors` or `ProxyProviders` – are applied. The average frame rendering time when using Provider is approximately 16 ms under moderate CPU load (iCoderz, 2025; Kumar, 2025).

From a formal perspective, Provider can be interpreted as a reactive system based on explicit notification-driven state propagation, where each state change triggers updates in all subscribed components. Unlike more advanced approaches with fine-grained dependency tracking, Provider does not inherently differentiate between direct and indirect dependencies, which may result in redundant widget rebuilds.

This characteristic reflects a comparatively coarse-grained transition model, where state changes are propagated uniformly rather than selectively. While such an approach ensures conceptual clarity and ease of implementation, it limits the efficiency of update propagation in applications with complex state structures.

Within the proposed formalization, Provider can therefore be classified as a system with explicit but non-optimized transition relations, where the absence of structured dependency tracking reduces determinism and scalability under increasing system complexity. Consequently, although Provider remains an effective solution for rapid development and low-

complexity scenarios, its architectural simplicity imposes constraints on performance optimization and maintainability in large-scale applications.

**Cubit: simplified implementation of event-driven state management.** Cubit is a simplified variant of the Bloc library, included in the flutter\_bloc package. Unlike full Bloc, Cubit does not use a formal event scheme – instead, it offers direct method calls that change state. This significantly reduces the amount of boilerplate code and lowers the entry threshold, while preserving the core advantages of the pattern: clear separation of business logic from the interface, reactive UI updates via BlocBuilder, and convenient integration with DevTools for state transition monitoring (Shakil, 2025). The event-driven, unidirectional state management pattern of Bloc and Cubit aligns with the behavioral design patterns described by Gamma et al. (1994), particularly the Observer pattern, which underpins reactive widget rebuilding.

Bloc, in turn, implements a strict unidirectional data flow scheme: interface components send events (Events) to the Bloc class, which processes these events and emits new states (States) via streams (Streams). Such an approach ensures complete transparency of data flow, making each state change clearly traceable and reproducible, making debugging and testing significantly more convenient. Bloc is widely used in large-scale enterprise applications, particularly in fintech solutions, where transaction management, authentication, and notifications are implemented through separate Bloc modules with detailed logging of each state transition (Kumar, 2025; SolGuruz, 2025). Cubit thus occupies a niche between Provider and Bloc, perfectly suited for implementing forms with validation, paginated data loading, and managing asynchronous operations while maintaining testability and separation of concerns.

**MobX: reactive dependency tracking in Flutter.** MobX is a reactive state management library, originally developed for JavaScript and successfully ported to Dart and Flutter (Meijer, 2010). Its key concept, transparent reactive programming, consists in the framework automatically tracking which parts of the interface access which observable state and, upon any change, precisely updating only the dependent components. The developer declares class fields as @observable, methods that change state as @action, and computed values as @computed. The entire reaction mechanism occurs automatically, without the need to manually notify dependent components (iCoderz, 2025).

MobX usage is particularly justified in scenarios with complex, interdependent state. For example, in a task management application, the completion state of a task, the display filter, and the counter of incomplete tasks are interconnected: changing any of them must cascade updates to the corresponding parts of the UI. MobX handles such relationships elegantly and automatically. MobX performance at 18 ms frame rendering time and moderate CPU load places the library between Riverpod and Bloc. MobX is an attractive solution for developers familiar with the JavaScript ecosystem and reactive programming patterns (Shakil, 2025; iCoderz, 2025).

From a formal standpoint, MobX can be interpreted as a reactive computational system with an implicit dependency graph, where state transitions propagate through dynamically established relationships between observables and observers. This distinguishes it from event-driven approaches such as Bloc or Cubit, where transitions are explicitly defined and pro-

cessed sequentially. In MobX, the order and scope of updates are determined by the structure of dependencies, which reduces the amount of boilerplate code but introduces a less explicit transition model.

Such an approach aligns with the general principle of fine-grained reactivity, where only those components that directly depend on the modified state are updated, minimizing unnecessary widget rebuilds. At the same time, the implicit nature of dependency tracking may reduce transparency of data flow, which can complicate debugging and formal verification of determinism in large-scale systems. Therefore, within the proposed formalization, MobX can be classified as a system with partially implicit transition relations, where the propagation of state changes is governed by reactive bindings rather than explicitly defined state transitions.

**Riverpod: advanced dependency management architecture.** Riverpod is the successor to Provider, developed by the same author – Remi Rousselet. Despite their shared conceptual roots, Riverpod is a fundamentally reconceived architecture, not simply an improved version of its predecessor. The key difference is that Riverpod places providers outside the widget tree, making them globally accessible regardless of where in the hierarchy the call is located. This eliminates a class of errors related to missing providers in context and ensures compile-time checks, which is critical in large team projects (Foresight Mobile, 2025).

The Riverpod architecture is based on the concept of providers as immutable references to state or logic. The key object for interaction is `ref`, an object that allows reading, observing, and tracking dependencies between providers. The declarative style of Riverpod allows conveniently describing asynchronous operations through `AsyncNotifier` and `FutureProvider`, automatically handling loading and error states without additional boilerplate. The library also supports family providers, which allow parameterizing providers and effectively managing state for dynamically created objects. Riverpod demonstrates its strengths in scenarios with intensive asynchronous data loading such as news aggregation applications and IoT platforms with numerous data streams (Kumar, 2025).

Demonstrating an average frame rendering time of approximately 12 ms and low CPU load, Riverpod confirms its advantage over Provider and other reviewed solutions. This is explained by a more intelligent dependency tracking mechanism that allows rebuilding only those parts of the tree that actually depend on the changed state, avoiding cascade rebuilds. Riverpod encourages a clear separation of responsibilities, where each provider is responsible for a specific part of state or business logic, making the code modular, easily testable, and convenient for reuse (JOTI, 2023; iCoderz, 2025).

From a formal perspective, Riverpod can be interpreted as a reactive system with explicitly defined dependency relations, where providers form a directed acyclic graph of state dependencies. Each state transition is propagated through this graph in a controlled and predictable manner, ensuring a high degree of determinism compared to approaches with implicit reactivity. Unlike MobX, where dependencies are inferred dynamically, Riverpod requires explicit declaration of relationships between providers, which improves transparency and traceability of data flow.

This explicit dependency model aligns with the formalization of state management as extended transition systems, where each provider represents a node with clearly defined inputs and outputs. As a result, Riverpod simplifies reasoning about system behavior under asynchronous updates and facilitates verification of correctness properties such as consistency and absence of unintended side effects. Furthermore, compile-time safety mechanisms reduce the likelihood of runtime errors, which is particularly important in large-scale and team-based development environments.

Overall, Riverpod can be classified as a highly deterministic and scalable state management solution, combining efficient dependency tracking with strong architectural guarantees, making it well-suited for complex applications with intensive data flows and strict requirements for maintainability.

**Conclusions.** Taking into account the published research results, it can be concluded that Riverpod demonstrates the most balanced performance metrics among the reviewed solutions, particularly in scenarios with frequent interface updates and dependency management (iCoderz, 2025; JOTI, 2023). MobX provides moderate performance through reactive dependency tracking, which can reduce the number of unnecessary rebuilds with proper configuration. Bloc and Cubit demonstrate stable but comparatively higher processing costs due to the heavier architectural pattern oriented toward large-scale and structured applications (Kumar, 2025; SolGuruz, 2025).

Provider in additional studies falls behind Riverpod in CPU usage and memory efficiency under identical workloads (JOTI, 2023), but remains the optimal choice for small-scale projects due to its minimal entry threshold. Since the considered differences are measured under controlled benchmark conditions, the practical impact of library choice depends on specific scenarios: dataset size, frequency of asynchronous updates, and UI architecture complexity (iCoderz, 2025).

The practical choice between Provider, Riverpod, Cubit, and MobX should be based on a comprehensive analysis of project scale, testability requirements, team experience, and subject domain specifics (Flutter Official Docs, 2025; Foresight Mobile, 2025). Prospects for further research include conducting own performance measurements in real production conditions with specified load scenarios, as well as extending the analysis to include newer solutions, particularly GetX and Signals, which are gaining popularity in the Flutter community as of 2024–2026. It is also advisable to study the impact of state management library choice on test coverage quality and the costs of long-term codebase maintenance in team projects.

#### ЛІТЕРАТУРА

1. Shakil M. Top 5 Flutter State Management Solutions 2025. Complete Guide [Electronic resource]. Medium, 2025. URL: <https://medium.com/@mshakilawan735/top-5-flutter-state-management-solutions-2025>.
2. Alur R., Dill D. L. A theory of timed automata. *Theoretical Computer Science*. 1994. Vol. 126, No. 2. P. 183–235. DOI: 10.1016/0304-3975(94)90010-8.
3. Baier C., Katoen J.-P. *Principles of model checking*. MIT Press, 2008. DOI: 10.7551/mitpress/9780262026499.001.0001.

4. Adam T. State Management Solutions: Comparing Provider, Riverpod, BLoC, and MobX [Electronic resource]. Medium, 2025. URL: <https://medium.com/@tofiqueadam/state-management-solutions-comparing-provider-riverpod-bloc-and-mobx>.
5. Flutter Official Docs. State Management Options [Electronic resource]. Google / Flutter Team, 2025. URL: <https://docs.flutter.dev/data-and-backend/state-mgmt/options>.
6. Gamma E., Helm R., Johnson R., Vlissides J. Design patterns: Elements of reusable object-oriented software. Addison-Wesley, 1994. DOI: 10.5555/186897.
7. Harel D., Pnueli A. On the development of reactive systems. In: Logics and models of concurrent systems. Springer, 1985. P. 477–498. DOI: 10.1007/978-3-642-82453-1\_17.
8. JOTI. State Management Comparison in Flutter [Electronic resource]. Jurnal Online Teknik Informatika, Universitas Dinamika, 2023. URL: <https://e-journals.dinamika.ac.id/joti/article/view/1164>.
9. Foresight Mobile. Best State Management for Flutter [Electronic resource]. Foresight Mobile, 2025. URL: <https://foresightmobile.com/blog/whats-the-best-state-management-library-for-flutter>.
10. Meijer E. Your mouse is a database. Communications of the ACM. 2010. Vol. 53, No. 4. P. 66–73. DOI: 10.1145/1721654.1721672.
11. Kumar A. Mastering State Management in Flutter:GetX vs Riverpod vs Bloc vs Provider (2025 Comparison) [Electronic resource]. Medium, 2025. URL: <https://medium.com/@anilkumar2681/mastering-state-management-in-flutter-getx-vs-riverpod-vs-bloc-vs-provider-2025-comparison-a48429710b96>.
12. iCoderz Solutions. Top Flutter State Management Packages of 2025 [Electronic resource]. iCoderz Solutions, 2025. URL: <https://www.icoderzsolutions.com/blog/flutter-state-management-packages>.
13. SolGuruz. Flutter State Management Packages, Best Practices, and More [Electronic resource]. SolGuruz, 2025. URL: <https://solguruz.com/blog/flutter-state-management/>.

#### REFERENCES

1. Shakil, M. (2025). Top 5 Flutter state management solutions 2025: Complete guide. Medium. <https://medium.com/@mshakilawan735/top-5-flutter-state-management-solutions-2025>
2. Alur, R., & Dill, D. L. (1994). A theory of timed automata. Theoretical Computer Science, 126(2), 183–235. [https://doi.org/10.1016/0304-3975\(94\)90010-8](https://doi.org/10.1016/0304-3975(94)90010-8)
3. Baier, C., & Katoen, J.-P. (2008). Principles of model checking. MIT Press. <https://doi.org/10.7551/mitpress/9780262026499.001.0001>
4. Adam, T. (2025). State management solutions: Comparing Provider, Riverpod, BLoC, and MobX. Medium. <https://medium.com/@tofiqueadam/state-management-solutions-comparing-provider-riverpod-bloc-and-mobx>
5. Flutter Team. (2025). State management options. <https://docs.flutter.dev/data-and-backend/state-mgmt/options>
6. Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1994). Design patterns: Elements of reusable object-oriented software. Addison-Wesley. <https://doi.org/10.5555/186897>

7. Harel, D., & Pnueli, A. (1985). On the development of reactive systems. In Logics and models of concurrent systems (pp. 477–498). Springer. [https://doi.org/10.1007/978-3-642-82453-1\\_17](https://doi.org/10.1007/978-3-642-82453-1_17)
8. JOTI. (2023). State management comparison in Flutter. Jurnal Online Teknik Informatika. <https://e-journals.dinamika.ac.id/joti/article/view/1164>
9. Foresight Mobile. (2025). Best state management for Flutter. <https://foresightmobile.com/blog/whats-the-best-state-management-library-for-flutter>
10. Meijer, E. (2010). Your mouse is a database. Communications of the ACM, 53(4), 66–73. <https://doi.org/10.1145/1721654.1721672>
11. Kumar, A. (2025). Mastering state management in Flutter: GetX vs Riverpod vs Bloc vs Provider (2025 comparison). Medium. <https://medium.com/@anilkumar2681/mastering-state-management-in-flutter-getx-vs-riverpod-vs-bloc-vs-provider-2025-comparison-a48429710b96>
12. iCoderz Solutions. (2025). Top Flutter state management packages of 2025. <https://www.icoderzsolutions.com/blog/flutter-state-management-packages/>
13. SolGuruz. (2025). Flutter state management packages, best practices, and more. <https://solguruz.com/blog/flutter-state-management/>

Received 30.03.2026.  
Accepted 02.04.2026.  
Published 30.04.2026

### ***Порівняльний аналіз моделей управління станом в мобільних застосунках, розроблених за допомогою Flutter***

*Метою цього дослідження є формалізація моделей управління станом у застосунках Flutter та їх оцінка за критеріями детермінізму переходів, продуктивності, масштабованості та архітектурної складності. Підходи до управління станом інтерпретуються як розширені перехідні системи з асинхронними подіями; властивості безпеки та життєздатності формалізуються за допомогою часової логіки. Формалізовану модель управління станом Flutter пропонується як клас реактивних обчислювальних систем з визначеними властивостями безпеки та життєздатності. Крім того, дослідження включає аналіз часу рендерингу кадрів та використання процесора на основі бенчмарків, щоб забезпечити емпіричне порівняння розглянутих підходів. Отримані результати додатково інтерпретуються в контексті реальних сценаріїв розробки, що дозволяє сформулювати практичні рекомендації щодо вибору відповідних архітектур управління станом залежно від масштабу та складності програми. Дослідження виявило значні відмінності між підходами в детермінізмі, продуктивності та архітектурній складності, демонструючи, що Riverpod забезпечує найбільш збалансовані характеристики продуктивності завдяки ефективному відстеженню залежностей, тоді як Bloc та Cubit забезпечують вищу структурну чіткість ціною збільшення обчислювальних витрат. MobX демонструє помірну продуктивність з перевагами в управлінні складним взаємозалежним станом, тоді як Provider залишається придатним для невеликих програм завдяки своїй простоті. Ці результати дозволяють обґрунтовано вибирати архітектуру на основі вимог виробничої програми, включаючи масштаб системи, обмеження продуктивності та потреби в обслуговуванні.*

*Ключові слова: управління станом, Flutter, реактивні системи, формалізація, Cubit, Riverpod, Provider, MobX, продуктивність, масштабованість.*

**Anastasiia Velma** - master's student, department of software engineering, Kharkiv National University of Radio Electronics, Ukraine.

ORCID: <https://orcid.org/0009-0006-9256-4226>

**Mykhailo Makarevych** - master's student, department of software engineering, Kharkiv National University of Radio Electronics, Ukraine.

ORCID: <https://orcid.org/0009-0006-4906-5261>

**Nataliia Golian** - candidate of technical sciences, associate professor, department of software engineering, Kharkiv National University of Radio Electronics, Ukraine.

ORCID: <https://orcid.org/0000-0002-1390-3116>

**Вельма Анастасія Олександрівна** - магістрант кафедри програмної інженерії, Харківський національний університет радіоелектроніки, Україна.

ORCID: <https://orcid.org/0009-0006-9256-4226>

**Макаревич Михайло Павлович** – магістрант кафедри програмної інженерії, Харківський національний університет радіоелектроніки, Україна.

ORCID: <https://orcid.org/0009-0006-4906-5261>

**Голян Наталія Вікторівна** – к.т.н., доцент, доцент кафедри програмної інженерії, Харківський національний університет радіоелектроніки, Україна.

ORCID: <https://orcid.org/0000-0002-1390-3116>

## ВИКОРИСТАННЯ СИСТЕМ АВТОМАТИЗОВАНОГО ПРОЄКТУВАННЯ ДЛЯ МОДЕЛЮВАННЯ СКЛАДНИХ СИСТЕМ ТА ТЕХНОЛОГІЧНИХ ПРОЦЕСІВ ВІДНОВЛЕННЯ ТА ЗМІЦНЕННЯ ДЕТАЛЕЙ

*Анотація.* Розглядаються питання підготовки інженерів-механіків в контексті використання у навчальному процесі сучасних систем автоматизованого проєктування та інженерного аналізу, зокрема таких дисциплін, які є базовими для вивчення студентами САПР технологічних процесів відновлення та зміцнення («Інженерна та комп'ютерна графіка», «Інформатика», «Прикладна механіка», «Персональні комп'ютери та пакети прикладних програм», «Деталі машин», «Системи інженерного аналізу», «Комп'ютерне забезпечення процесів відновлення», «Контроль якості покриттів»). Зазначено, якщо на бакалавраті спостерігається наступність викладання, то на 1-му курсі магістратури виникають проблеми у зв'язку з відсутністю стандартної САПР ТП зміцнення та відновлення, яку можна було б використовувати у навчальному процесу і яка б використовувалась на підприємствах відновлення та технічного сервісу автомобілів. Тому виникла задача проводити підготовку фахівців, здатних самостійно розробити базу даних для інформаційно-закінченого програмного комплексу з метою вирішення задачі зміцнення та відновлення деталей автомобільного транспорту.

*Ключові слова:* автоматизоване проєктування, SOLIDWORKS, САПР технологічних процесів, MS Access, відновлення та зміцнення, інженерний аналіз.

**Вступ.** Підготовка студентів на факультеті інженерії, транспорту та архітектури Хмельницького національного університету включає використання ряду дисциплін, в рамках яких студенти використовують комп'ютерну техніку та відповідне інженерне програмне забезпечення для розв'язання задач у машинобудуванні. Зокрема, студенти спеціальності G8 «Матеріалознавство» кафедри трибології, автомобілів та матеріалознавства вивчають наступні дисципліни в рамках комп'ютерної підготовки:

– «Інформатика» (1-й курс) – основи алгоритмізації та програмування, вивчення базових прикладних програм (зокрема, з пакету програм MS Office);

– «Інженерна та комп'ютерна графіка» (1-й курс) – вивчення основ створення моделей деталей, складальних одиниць, а також оформлення конструкторської документації у системі 3D-моделювання SOLIDWORKS (рис. 1);

– «Персональні комп'ютери та пакети прикладних програм» (2-й курс) – вивчення математичного пакету MathCAD, що дозволяє студентам проводити складні інженерні розрахунки як в рамках виконання розрахунково-графічних робіт, так і для курсового та дипломного проєктування

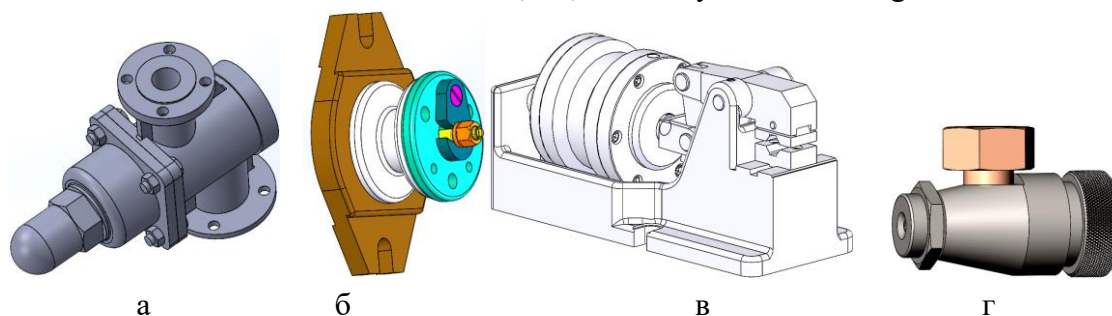


Рисунок 1 – Приклади робіт: клапан пропускний (а), кондуктор (б), пристрій для захоплення прутків (в), форсунка (г)

– «Прикладна механіка і деталі машин» (2-й курс) – вивчення теорії та основ конструювання машин в рамках загальноінженерної підготовки, яка включає в себе вивчення методів розрахунку, проектування і конструювання деталей машин загального призначення з метою забезпечення вибору більш раціональних матеріалів, придання деталям найбільш вигідних форм і розмірів, встановлення ступеня точності, якості поверхні та технічних умов виготовлення, а також вибору мастильних матеріалів. Студентами виконується курсовий проєкт з проектування механічного привода, що, зокрема, включає в себе розрахунок та 3D-моделювання редукторів, оформлення складальних та робочих креслень, іншої конструкторської документації (рис. 2). Додатково, студенти мають можливість поглиблено вивчати чисельні методи розрахунку деталей машин на міцність та жорсткість в рамках дисципліни «Системи інженерного аналізу».

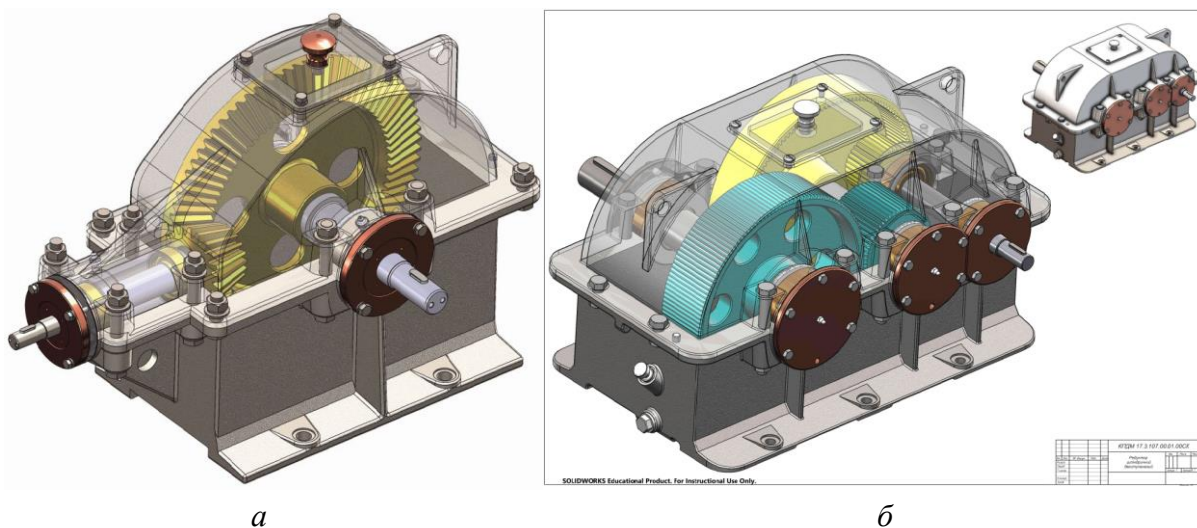


Рисунок 2 – Приклади курсових проєктів з курсу «Деталі машин»

– «Комп'ютерне забезпечення процесів відновлення» (2-й курс) – використання методу скінченних елементів для лінійного статичного аналізу в SOLIDWORKS Simulation деталей та пристосувань для відновлення деталей автомобілів: розрахунок пристосувань на міцність є важливим етапом їхнього проектування та перевірки, приклади наведено на рис. 3;

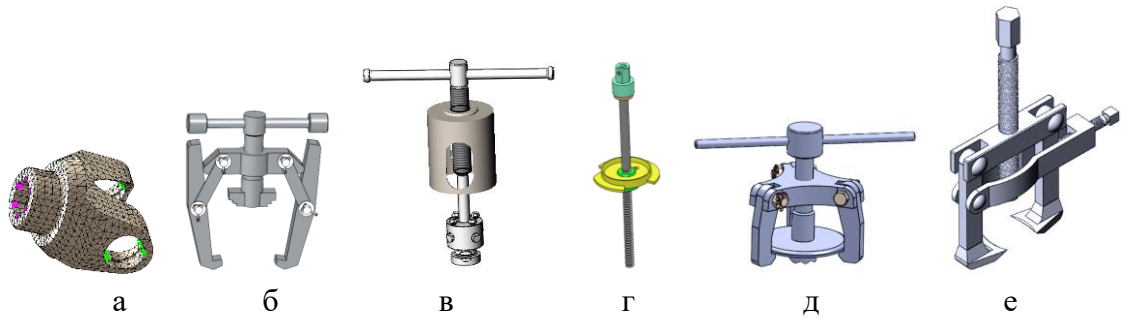


Рисунок 3 – Переднє вухо карданного валу (а), знімач підшипників (б), знімач втулок (в), пристосування для стиску пружин (г), знімач універсальний (д), знімач кермових сошок (е)

– «Контроль якості покриттів» (2-й курс) – визначення параметрів якості покриттів і застосування комп'ютерних технологій SOLIDWORKS Simulation [1-4] для визначення їх експлуатаційних характеристик (рис. 4, 5);

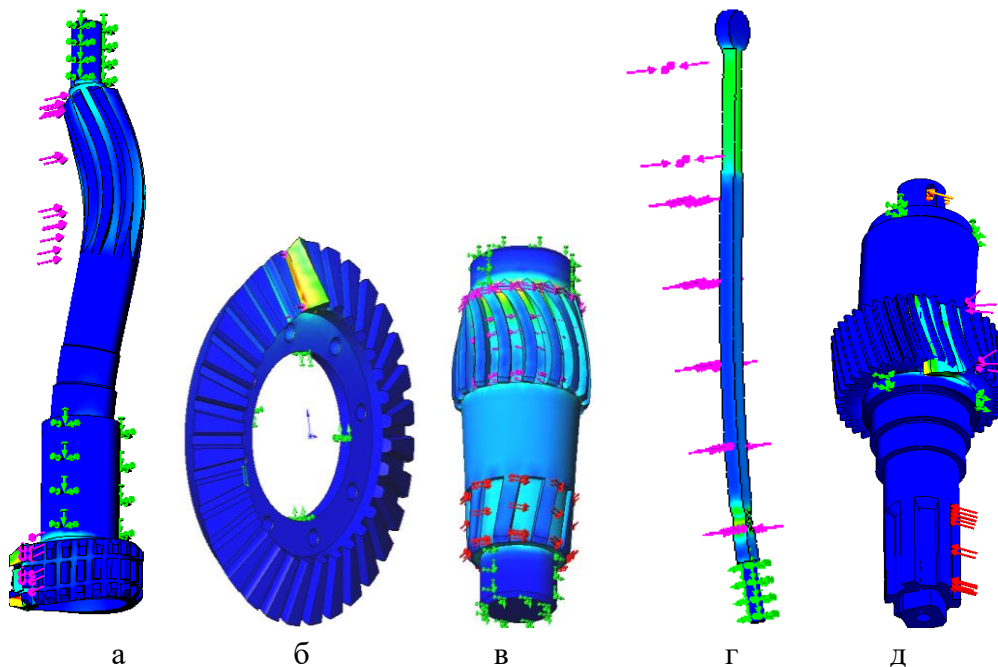


Рисунок 4 – Оцінка впливу твердості покриття на сумарні напруження von Mises первинного валу коробки передач (а), шестерні переднього моста (б), валу роздавальної коробки (в), тяги керма (г), валу коробки передач (д)

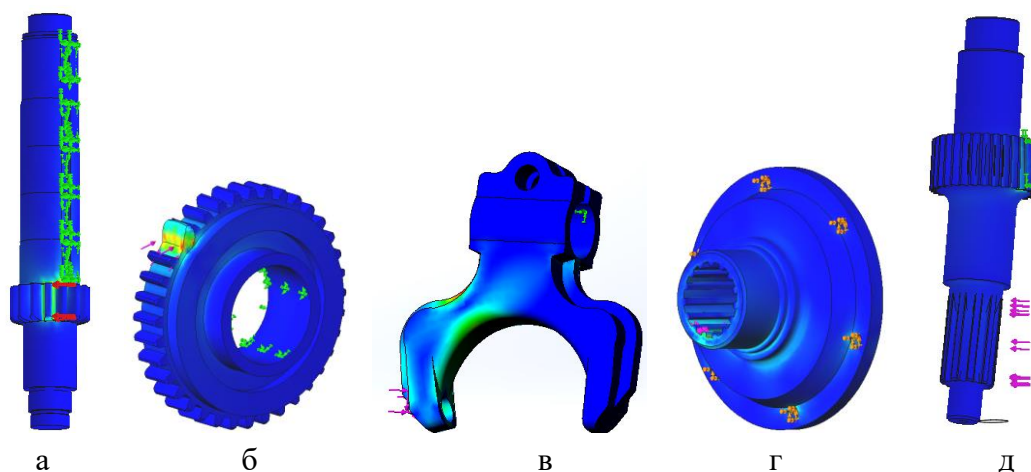


Рисунок 5 – Оцінка впливу корозії покриття на сумарні напруження von Mises вторинного валу коробки передач (а), шестерні (б), вилки (в), фланця (г), вал-шестерні (д)

– «Системи автоматизованого проектування технологічних процесів (САПР ТП відновлення та зміцнення» (магістратура спеціальності G8 «Матеріалознавство», 1-й курс) – створення бази даних (БД) з відновлення та зміцнення деталей автомобільного транспорту (АТ), з використанням MS Access, проектування технологічної підготовки виробництва, розробка технологічної документації (Word, Excel); застосування PTC MathCAD для математичного моделювання зносостійкості зміцнених та відновлених деталей.

Але якщо на бакалавраті спостерігається наступність викладання дисциплін інженерного циклу з використанням комп'ютерних технологій, то найбільші проблеми виникають на 1-му курсі магістратури у зв'язку з відсутністю офіційної САПР ТП зміцнення та відновлення, яку можна було б придбати для вивчення і яка б використовувалась на підприємствах відновлення та технічного сервісу автомобілів.

**Аналіз останніх досліджень і публікацій.** На даний час всі технологічні САПР, якщо не враховувати програмування для верстатів з числовим програмним керуванням [5], зводяться до автоматизації оформлення технологічної документації – для прискорення вибору використовуються БД з устаткування, оснастки тощо, застосовуються типові ТП. До числа найефективніших технологій, які дозволяють виконати ці вимоги, належать САПР технологічної підготовки виробництва та інженерного аналізу.

За останні роки CAD/CAM/CAE/PDM-системи пройшли шлях від порівняно простих креслярських додатків до інтегрованих програмних комплексів [6-9], які забезпечують єдину підтримку всього циклу розробки, починаючи від ескізного проектування і закінчуючи технологічною підготовкою виробництва, іспитами та супроводом. Відповідно, на ринку України є декілька найвідоміших систем, які реалізують цю задачу і забезпечують: ТП обробки деталей, зв'язок з AutoCAD чи іншими CAD-системами, створення архіву конструкторської, технологічної документації, БД ТП тощо.

Автоматичний режим проектування ТП – кінцева мета в області САПР технологій [10]. Але основною перешкодою до його реалізації є відсутність формалізованого представлення про просторове розташування поверхонь деталі. З плоского крес-

лення, виконаного в конструкторській САПР, у технологію можна передати лише окремі параметри: розміри, допуски, шорсткість, квалітети, але тільки у тому випадку, якщо креслення деталі параметризоване. Тому в автоматичному режимі можна тільки допрацьовувати заздалегідь розроблені типові ТП. Область застосування даного режиму – найпростіші деталі. Крім цього, як правило, такі розробки використовуються тільки на тому підприємстві, для якого вони розроблені, а спроба перенести їх на інші підприємства вимагає великих витрат при адаптації (у кращому випадку буде потрібна зміна інформаційного наповнення системи – таблиць БД, текстових і файлів налаштувань тощо; у гіршому – може виникнути необхідність зміни вихідного коду додатка). Тому вони не дуже широко застосовуються на машинобудівних підприємствах і зовсім не знаходять використання в авторемонтних та підприємствах автосервісу України.

**Мета дослідження.** При організації навчального процесу поставлена задача підготовки фахівців, здатних самостійно розробити БД для інформаційно-закінченого програмного комплексу з метою вирішення завдання зміцнення та відновлення деталей АТ. Основні функціональні режими такої системи мають відповідати наступним вимогам:

- можливість вибору оптимального методу зміцнення чи відновлення деталі з подальшим вибором оптимального режиму;
- проектування на основі техпроцесу-аналогу (автоматичний вибір відповідної технології з архіву з наступною доробкою у діалозі);
- формування ТП з окремих блоків, які зберігаються у бібліотеці типових технологічних операцій і переходів;
- об'єднання окремих операцій архівних технологій;
- автоматична доробка типової технології на основі даних, переданих з параметризованого креслення, наприклад, AutoCAD чи іншої САПР;
- введення інформації про ТП у діалоговому режимі за допомогою спеціальних процедур доступу до довідкових БД тощо.

У кожному конкретному випадку технологу повинна бути надана можливість вибору оптимального сполучення режимів проектування, взаємодоповнюючих один одного. З цього погляду САПР ТП – це, насамперед, система управління базами даних (СУБД), оскільки основна частина роботи з проектування ТП припадає на роботу з даними та при цьому обробляється дуже велика кількість інформації. Від того, як реалізовані функції обробки даних та їхніх логічних взаємозв'язків залежать інші показники системи.

**Виклад основного матеріалу дослідження.** СУБД є програмним продуктом, що постачається як прикладна програма, яку потрібно встановити на комп'ютер з урахуванням його конфігурації, ресурсів та операційної системи, а також вимог до набору функцій. СУБД є основою створення програм користувача для різних предметних областей, кожна з яких (у даному випадку – ТП зміцнення та відновлення деталей АТ), має свою специфіку організації БД.

Найбільшу популярність серед настільних систем завоювали реляційні СУБД, такі як PostgreSQL, Oracle Database, Microsoft SQL Server, MySQL.

PostgreSQL надає розширені можливості роботи з даними, включаючи складні запити, транзакції та підтримку різних типів даних (зокрема, JSON). Вона дозволяє розширення функціональності через користувацькі функції, типи та модулі. Oracle Database реалізує високопродуктивне керування великими обсягами даних із підтримкою складних транзакцій і розподілених обчислень. Вона забезпечує розширені механізми безпеки, резервування та масштабування для критично важливих систем. Microsoft SQL Server надає повноцінне середовище для зберігання, обробки та аналізу реляційних даних із використанням SQL і вбудованих інструментів. Вона підтримує інтеграцію даних і розробку звітів у межах платформи Microsoft.

Для організації навчального процесу використано систему керування реляційними базами даних MySQL, яка забезпечує зберігання, обробку та керування реляційними даними з використанням SQL (Structured Query Language – мова структурованих запитів), підтримуючи базові транзакції та індексацію. SQL орієнтована на швидку роботу веб-додатків і простоту розгортання.

Для забезпечення одночасного доступу до даних багатьох користувачів, які розташованих далеко один від одного і від місця збереження БД, використовуються мережні мультикористувацькі версії СУБД. Спільна робота користувачів у мережах можлива тільки при наявності стандартної мови обробки даних, якою стала SQL – інтерпретована мова, що описує операції (створення, обробка і витяг) над реляційними БД. Вона є найуживанішою мовою управління БД у системах клієнт/сервер. Основна її перевага – розробка запитів у будь-якій системі управління даними, сумісною з SQL. На сьогоднішній день SQL підтримують більше ста СУБД, які працюють як на персональних комп'ютерах, так і на великих електронно-обчислювальних машинах. SQL – це легка для розуміння мова й, у той же час, універсальний програмний засіб управління даними. Використання конкретного SQL-сервера залежить від постановки задачі, кількості користувачів і вимог замовника.

Настільна реляційна СУБД MS Access входить до складу пакета Microsoft Office та орієнтована на невеликі БД. Вона поєднує засоби створення таблиць, форм, запитів і звітів із простим графічним інтерфейсом, що не потребує глибоких знань програмування. У системі MS Access передбачена автоматична генерація коду SQL при створенні запиту користувачем, інтегрованість з іншими програмами пакета MS Office тощо. Тому саме систему MS Access вибрано для початкового створення БД зміцнення та відновлення деталей АТ.

Таким чином, ми готуємо студентів не тільки як користувачів уже готових САПР, а даємо їм можливість самим розробляти, нехай і нескладні, але закінчені БД. Ті ж студенти, хто захоплюється програмуванням, у додатковий час в рамках самостійної роботи вирішують більш серйозні задачі автоматизації конструкторського і технологічного проектування.

**Результати.** Усі студенти виконують самостійну роботу зі створення інформаційно-програмних продуктів для розв'язання нескладних технологічних задач. Обов'язковою вимогою до цих розробок є їх ретельна алгоритмічна проробка, створення інфолю-

гічної й даталогічної моделей БД, супровід довідковими матеріалами і докладне документування. Весь процес освоєння інструментальних засобів займає дев'ять двогодинних занять у лабораторії під керівництвом викладача. Далі робота виконується студентами самостійно, відповідно до індивідуальних завдань. Розглянемо самостійну роботу студента на прикладі розподільного валу, яка полягає у наступному:

– конструювання та розробка геометричної моделі розподільного валу у САПР SOLIDWORKS для автомобіля Toyota Corolla 12 з метою визначення характерних параметрів відновлення (рис. 6);

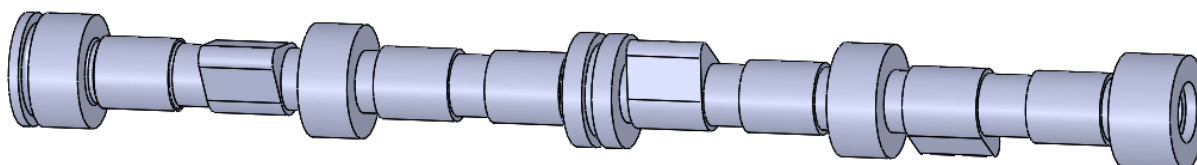


Рисунок 6 – Розподільний вал автомобіля Toyota Corolla 12

– вибір способу зміцнення розподільного валу (розробка БД з використанням СУБД Access і мови SQL – створення таблиць, форм, запитів), як наведено на рис. 7;

Таблиці											
Форма											
Запит SQL	<pre>SELECT [Поверхні деталей ГРМ, вимоги до зміцненої поверхні].[№ Поверхні], [Поверхні деталей ГРМ, вимоги до зміцненої поверхні].[Назва поверхні], [Способи зміцнення поверхні №1 розподільного валу].[Спосіб зміцнення], [Способи зміцнення поверхні №1 розподільного валу].[Твердість після зміцнення, HRC], [Способи зміцнення поверхні №1 розподільного валу].[Глибина зміцнюваного шару, мм] FROM [Поверхні деталей ГРМ, вимоги до зміцненої поверхні] LEFT JOIN [Способи зміцнення поверхні №1 розподільного валу] ON [Поверхні деталей ГРМ, вимоги до зміцненої поверхні].[№ Поверхні] = [Способи зміцнення поверхні №1 розподільного валу].[№ Поверхні] WHERE ((([Способи зміцнення поверхні №1 розподільного валу].[Твердість після зміцнення, HRC]&gt;'58') AND (([Способи зміцнення поверхні №1 розподільного валу].[Глибина зміцнюваного шару, мм]&gt;'1')));</pre> <p><b>Вибір способу зміцнення опорних шийок</b></p> <table border="1"> <thead> <tr> <th>№ Поверхні</th> <th>Назва поверхні</th> <th>Спосіб зміцнення</th> <th>Твердість після зміцнення</th> <th>Глибина зміцнюваного шару, мм</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Опорні шийки</td> <td>ТЦО</td> <td>62</td> <td>20</td> </tr> </tbody> </table>	№ Поверхні	Назва поверхні	Спосіб зміцнення	Твердість після зміцнення	Глибина зміцнюваного шару, мм	1	Опорні шийки	ТЦО	62	20
№ Поверхні	Назва поверхні	Спосіб зміцнення	Твердість після зміцнення	Глибина зміцнюваного шару, мм							
1	Опорні шийки	ТЦО	62	20							

Рисунок 7 – Вибір способу зміцнення розподільного валу

– проведення лінійної та поліноміальної регресії (апроксимація даних) методом найменших квадратів даних зносостійкості розподільного валу (MathCAD) – рис. 8;

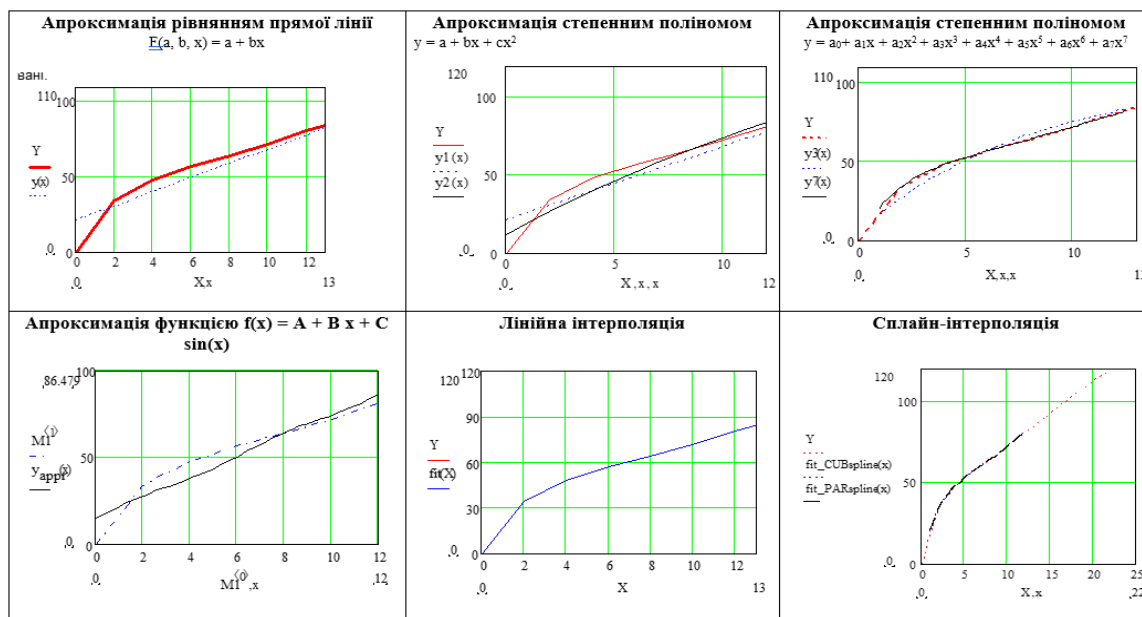


Рисунок 8 – Математичне моделювання ТП у середовищі MathCAD

– визначення оптимального режиму зміцнення розподільного валу, приклад наведено на рис. 9;

Керуючі запити SQL, використання VBA	Запити на вибірку мовою SQL
<p><b>Створення таблиці Основні деталі ГРМ</b></p> <pre>CREATE TABLE Основні деталі ГРМ (Код деталі INTEGER NOT NULL UNIQUE PRIMARY KEY, Назва деталі VARCHAR, К-ть деталей VARCHAR, Матеріал VARCHAR, Твердість заготовки VARCHAR, Твердість після термообробки VARCHAR);</pre>	<p><b>Основні деталі ГРМ</b></p> <pre>SELECT [Основні деталі ГРМ].[Назва деталі], [Основні двигуна].Матеріал FROM [Основні деталі двигуна];</pre>
<p><b>Створення таблиці Поверхні валу</b></p> <pre>CREATE TABLE Поверхні валу (Код деталі INTEGER NOT NULL UNIQUE PRIMARY KEY, Номер поверхні VARCHAR, Назва поверхні VARCHAR, Метод чи прилад контролю VARCHAR);</pre>	<p><b>Дефекти валу</b></p> <pre>SELECT [Поверхні валу].[Код деталі], [Поверхні валу].[номер поверхні], [Поверхні валу].[Назва поверхні], [Поверхні валу].[Метод чи прилад контролю] FROM [Поверхні валу] WHERE ((([Поверхні валу].[номер валу])="3"));</pre>
<p><b>Створення таблиці Способи зміцнення</b></p> <pre>Sub CreateTable() Dim Table As New Table Dim Catalog As New ADOX.Catalog Dim Key As New ADOX.Key Catalog.ActiveConnection = "Provider = Microsoft.Jet.OLEDB.4.0;Data " &amp; "Source=D:\VBA1.mdb" Table.Name = "Способи зміцнення" Table.ParentCatalog = Catalog Table.Columns.Append "Номер поверхні", adInteger Table.Columns.Append "Назва поверхні", adVarChar Table.Columns.Append "Спосіб зміцнення", adVarChar Table.Columns.Append "Назва деталі", adVarChar Catalog.Tables.Append Table Key.Name = "Номер поверхні" KeyType = adKeyPrimary Key.Columns.Append "Номер поверхні" Catalog.Tables("Способи зміцнення").Keys.Append Key, KeyPrimary Set Catalog.ActiveConnection = Nothing End Sub</pre> <p><b>Введення рядка 1 у таблицю Способи зміцнення поверхні</b></p> <pre>INSERT INTO Способи зміцнення VALUES (1, 'СТЦЮ', '500', '600');</pre>	<p><b>Результуюча таблиця</b></p> <pre>SELECT [Основні деталі ГРМ].[Назва поверхні], [поверхні валу].[Назва поверхні], [Способи зміцнення].[спосіб зміцнення] FROM ([Основні деталі ГРМ] INNER JOIN [Поверхні валу] ON [Основні деталі ГРМ].[Код деталі]=[Поверхні валу].[Код деталі]) INNER JOIN [Способи зміцнення] ON [поверхні валу].[номер поверхні]=[Способи зміцнення поверхні].[номер поверхні] WHERE ((([Основні ГРМ].[Назва деталі]="валу") AND ((([Поверхні валу].[номер поверхні])="2"));</pre>

Рисунок 9 – Вибір оптимального режиму зміцнення розподільного валу

– аналіз розрахунків напруженого стану розподільного валу (рис. 10).

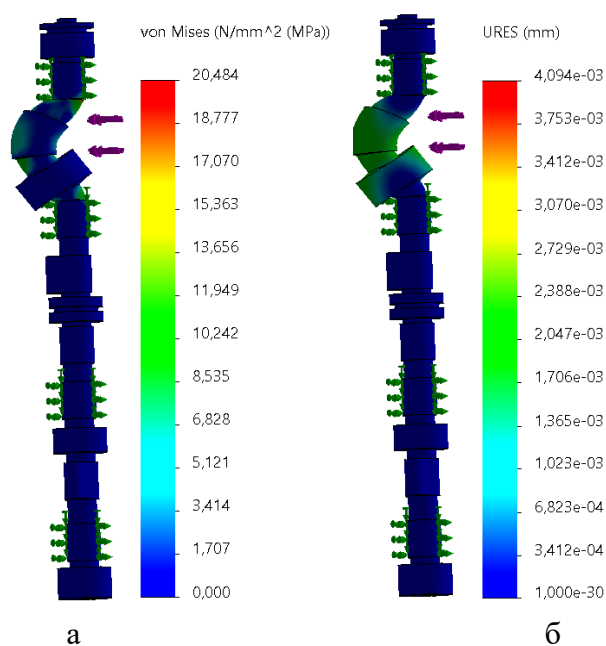


Рисунок 10 – Епюра розподілу максимальних напружень (а) та переміщень (б)

Приклади маршрутної та операційної карт ТП зміцнення розподільного валу наведені на рис. 11.

ДСТУ-Н 7914-2015																			
1																			
2																			
3	Дубл.																		
4	Взам.																		
5	Підп.																		
6																			
7																			
8	Розробка											25.05.25							
9	Перевірка	Рубик										ХНУ							
10	Нормував																		
11	Н. контроль											Розподільний вал							
12																			
13	M01																		
14	M02	Код	ЕБ	МД	ЕН	Н.Розк	Клм	Код заготовки	Профіль і розміри	КД	КЗ								
15																			
16	А	Цех	Дл	РМ	Спер	Код, найменування операції			Позначення документу										
17	Б	Код, найменування обладнання							СМ	Проф	Р	УТ	КФ	КОД	ЕН	ОП	Кшт	Т.п.з.	Т.шт.
18	A03					005	Мийна												
19	B04	Машинна мийна М2АП							14783	2	05	1	100	1	1,7	0,37			
20	05																		
21	A06					010	Транспортувальна												
22	07																		
23	O08	Транспортувати деталі на робоче місце для проведення дефектації на відстань до 10 м																	
24	T09	Тара 7543(10 шт., 55кг); візок ручний 6334.00																	
25	10																		
26	A11					015	Дефектувальна												
27	12	Стіл ВТК							14054	02	1	100	1	14	3,6				
28	13																		
29	14																		
30	МК	Маршрутна карта																	

а

«Системні технології» 3 (164) 2026 «System technologies»

ДСТУН 7914:2015												
120												
121												
122	Дубл.											
123	Взам.											
124	Підл.											
125											1	1
126												
127	Розробка		30.05.25									
128	Перевірка	Рудик			ХНУ							
129	Нормува											
130												
131	Н. контроль				Розподільний вал							045
Найменування операції	Найменування матеріалу			Твердість	ЕВ	МД	Профіль і розміри			МЗ	КОД	
Контрольна	ВЧ 45-10 ДСТУ 3925-99				кз	0,36				0,5		
Обладнання	Позначення програми			То	Тд	Тп.з.	Тшт.		ЗОР			
Стіл контрольний				1,1			1,1					
Р	П	Д або В	L	t	i	S	N	V	To	Та		
136												
137	МЗ											
138	004	Перевірити зовнішнім оглядом шершавість поверхонь, відсутність заусениць та гострих кромek										
139	Т05											
140	06	Перевірити розмір1										
141	07	Перевірити розмір2										
142	08	Перевірити розмір3										
143	09	Перевірити розмір4										
144	10											
145	11											
146	12											
147												
148	OK	Операційна карта										
149												

б

Рисунок 11 – Зразки маршрутної (а) та операційної (б) карт ТП зміцнення розподільного валу

Робота над подібною самостійною роботою знайомить студентів не лише з рішенням локальних задач автоматизованого проектування, але й дозволяє познайомитись з питаннями комплексної автоматизації виробництва.

**Висновки.** В роботі розглянуто питання використання в навчальному процесі сучасних систем автоматизованого проектування та інженерного аналізу, що є базовими для вивчення студентами САПР технологічних процесів відновлення та зміцнення. Наведена методика підготовки фахівців, здатних самостійно розробити БД для інформаційно-закінченого програмного комплексу з метою вирішення задачі зміцнення та відновлення деталей АТ.

### ЛІТЕРАТУРА

- Rudyk, O., Kaplun, P., Gonchar, V. (2022). Application of SolidWorks for the training of highly qualified specialists. Current problems in the education system: general secondary education institution – pre-university training – higher education institution: collection of scientific works of the materials of the VIII All-Ukrainian Scientific and Practical Conference, February 17, Kyiv, National Aviation University. Kyiv: NAU, Pp. 699-706.
- Rudyk, O.Yu., Kaplun, P.V., Gonchar, V.A. (2022). Application of SolidWorks is for professional preparation of specialists. Resource-oriented learning in "3D": accessibility, dialogue, dynamics: a collection of abstracts of reports of the 2nd International Scientific and Practical Internet Conference, Poltava, February 22-23. 2 – Poltava: PUET, Pp. 140-146.

3. Rudyk, O.Yu., Dykha, O.V. (2020) SolidWorks as an innovative means for studying the disciplines of automobile profile. "System technologies" 3 (128), Pp. 21-35.
4. Rudyk, O.Yu., Poberezhnyi, M.M., Kaplun, P.V., Gonchar, V.A. (2025). Application of computer-integrated technologies in modern engineering systems. "System Technologies", 4 (159), Pp. 37-45.
5. Литвиненко, О.А., Бойко, Ю.І., Яновський, В.А. (2020) CAD-CAM технології проектування та виготовлення деталей на верстатах з ЧПК, Технічна інженерія, 1(85), с.15–22.
6. Варшав'як, Г.Б., Гребеніков, О.Г., Гуменний, А.М., Гребеніков, В.А. (2023). Досвід впровадження систем CAD/CAM/CAE/PLM в навчально-науково-виробничому центрі Національного аерокосмічного університету ім. М. Є. Жуковського «Харківський авіаційний інститут» // Відкриті інформаційні та комп'ютерні інтегровані технології. – Вип. 97, Харків: ХАІ, с. 164–194.
7. Піонткевич, О.В., Лозінський, Д.О., Сердюк, О.В., Савуляк, В.В. (2024). Забезпечення результатів вивчення CAD/CAE/CAM систем для підготовки фахівців із спеціальності «Прикладна механіка». Сучасна освіта – доступність, якість, визнання, ДДМА, с. 247–252.
8. Артюх, О.М., Дударенко, О.В., Кузьмін, В.В. та ін. (2021). Основи САПР в автомобілебудуванні: навч. посіб. Запоріжжя: НУ «Запорізька політехніка». 168 с.
9. Гречаний О.М., Васильченко Т.О., Власов А.О., Івахненко О.П., Вернидуб М.В. (2025). Роль САПР у розробці інженерних рішень для металургійної галузі: досвід застосування ANSYS, AutoCAD, SOLIDWORKS та MATLAB. Системні технології, 3 (158), с. 10–20.
10. Третяк В.В., Сотников В.Д., Худяков С.В., Скорченко І.В. (2020). Проектування технологічних процесів у САПР ТП: навч. посібник до дипломного проектування, Харків: Національний аерокосмічний університет ім. М.Є. Жуковського «Харківський авіаційний інститут», 80 с.

## REFERENCES

1. Rudyk O., Kaplun P., Gonchar, V. (2022). Application of SolidWorks for the training of highly qualified specialists. Current problems in the education system: general secondary education institution – pre-university training – higher education institution: collection of scientific works of the materials of the VIII All-Ukrainian Scientific and Practical Conference, February 17, Kyiv, National Aviation University. Kyiv: NAU, Pp. 699-706.
2. Rudyk O.Yu., Kaplun P.V., Gonchar V.A. (2022). Application of SolidWorks is for professional preparation of specialists. Resource-oriented learning in "3D": accessibility, dialogue, dynamics: a collection of abstracts of reports of the 2rd International Scientific and Practical Internet Conference, Poltava, February 22-23. – Poltava: PUET, Pp. 140-146.
3. Rudyk O.Yu., Dykha O.V. (2020) SolidWorks as an innovative means for studying the disciplines of automobile profile. "System technologies" 3 (128), Pp. 21-35.
4. Rudyk O.Yu., Poberezhnyi M.M., Kaplun P.V., Gonchar V.A. (2025). Application of computer-integrated technologies in modern engineering systems. "System Technologies", 4 (159), Pp. 37-45.

5. Lytvynenko, O.A., Boiko, Yu I., Yanovskyi, V.A. (2020). CAD-CAM tekhnologii proektuvannia ta vyhotovlennia detalei na verstatax z ChPK, Tekhnichna inzheneriia, 1 (85), c. 15–22.
6. Varshaviak, H.B., Hrebenikov, O.H., Humennyi, A.M., & Hrebenikov, V.A. (2023). Dosvid vprovadzhennia system CAD/CAM/CAE/PLM v navchalno-naukovo-vyrobnychomu tsentri Natsionalnogo aerokosmichnogo universytetu im. M. Ye. Zhukovskoho “Kharkivskiy aviatsiinyi instytut”. Vidkryti informatsiini ta kompiuterni intehrovani tekhnologii, (97), c. 164–194.
7. Piontkevych, O.V., Lozinskyi, D.O., Serdiuk, O V., Savuliak, V. V. (2024). Zabezpechennia rezultativ vyvchennia CAD/CAE/CAM system dlia pidhotovky fakhivtsiv iz spetsialnosti “Prykladna mekhanika”. In Suchasna osvita – dostupnist, yakist, vyznannia, DDMA, c. 247–252.
8. Artiukh, O.M., Dudarenko, O.V., Kuzmin, V.V., et al. (2021). Osnovy SAPR v avtomobilebuduvanni: navch. posib. Zaporizhzhia: NU “Zaporizka politekhnika”. 168 c.
9. Hrechanyi, O.M., Vasylchenko, T O., Vlasov, A.O., Ivakhnenko, O.P., Vernydub, M.V. (2025). Rol SAPR u rozrobsi inzhenernykh rishen dlia metalurhiinoi haluzi: dosvid zastosuvannia ANSYS, AutoCAD, SolidWorks ta MATLAB. Systemni tekhnologii, 3(158), c. 10–20.
10. Tretiak, V.V., Sotnykov, V.D., Khudiakov, S.V., Skorchenko, I.V. (2020). Proektuvannia tekhnologichnykh protsesiv u SAPR TP: navch. posibnyk do diplomnogo proektuvannia, Kharkiv: Natsionalnyi aerokosmichnyi universytet im. M.Ye. Zhukovskoho “Kharkivskiy aviatsiinyi instytut”, 80 p.

Received 30.03.2026.  
Accepted 02.04.2026.  
Published 30.04.2026

***The usage of computer-aided design systems for modeling complex systems  
and technological processes for the restoration and reinforcement of parts***

*Nowadays, all technological database management systems, excluding programming for CNC machines, essentially focus on automating the preparation of technological documentation. Among the most effective technologies capable of meeting these requirements are CAD systems for production preparation and engineering analysis. In recent years, CAD/CAM/CAE/PDM systems have progressed from relatively simple drafting applications to integrated software suites that provide unified support for the entire development cycle. The automated design of technological processes (TP) is the ultimate goal in the field of CAD technologies. However, the main obstacle to its implementation is the lack of a formalized representation of the three-dimensional arrangement of a part's surfaces. Therefore, these systems are not widely used in machine-building enterprises and are not used at all in auto repair and auto service enterprises in Ukraine.*

*The academic process is designed to train specialists capable of independently developing a database for a fully functional software system aimed at strengthening and restoring automotive components. Therefore, the MySQL relational database management system is used, which provides storage, processing, and management of relational data using SQL (Structured Query Language), supporting basic transactions and indexing. SQL is*

*designed for fast web application performance and ease of deployment. The relational DBMS MS Access was also used to organize the educational process; it features automatic SQL code generation when a user creates a query, integration with other MS Office programs, etc. Therefore, the MS Access system was chosen for the initial creation of the database for the strengthening and restoration of automotive parts.*

*All students complete an independent project to create information and software products designed to solve simple technological tasks. A compulsory requirement for these projects is a thorough algorithmic analysis, the creation of information-logical and data-logical database models, the inclusion of reference materials, and detailed documentation.*

*Keywords: computer-aided design, SOLIDWORKS, automated design of technological processes, MS Access, repair and reinforcement, engineering analysis.*

**Рудик Олександр Юхимович** – к.т.н., доцент кафедри трибології, автомобілів та матеріалознавства Хмельницького національного університету.

ORCID: <https://orcid.org/0000-0002-6937-1366>

**Харжевський В'ячеслав Олександрович** – д.т.н., професор кафедри галузевого машинобудування та агроінженерії Хмельницького національного університету.

ORCID: <https://orcid.org/0000-0003-4816-2781>

**Rudyk Oleksandr Yuhymovych** – PhD., associate professor, department of tribology, automobiles and materials science, Khmelnytskyi National University.

ORCID: <https://orcid.org/0000-0002-6937-1366>

**Kharzhevskiy Viacheslav Oleksandrovych** – doctor of engineering sciences, professor, department of industrial and agricultural engineering, Khmelnytskyi National University.

ORCID: <https://orcid.org/0000-0003-4816-2781>

## ПРОГРАМНИЙ МЕТОД СЕГМЕНТАЦІЇ ЗОБРАЖЕНЬ В УМОВАХ ВАРІАТИВНОГО ОСВІТЛЕННЯ

*Анотація. У статті розглянуто проблему сегментації зображень в умовах варіативного освітлення. Проведено аналіз методів сегментації цифрових зображень, найбільш ефективних в умовах варіативного освітлення. Запропоновано метод сегментації, що ґрунтується на початковій класифікації зображення на основі ознак освітленості, алгоритмах адаптивного оброблення зображення та нейромережевій сегментації на основі архітектур SegNet та U-Net. Запропонований метод продемонстрував покращення на 4,95% та 1,5% відповідно до метрик mIoU та Pixel Accuracy для моделі SegNet, а також на 1,6% та 0,07% для моделі U-Net.*

*Ключові слова: семантична сегментація, варіативне освітлення, оброблення зображень, нейронні мережі, інженерія програмного забезпечення.*

**Постановка проблеми.** Сегментація зображень є однією з ключових задач комп'ютерного зору, що забезпечує виділення змістовних областей зображення та є основою для подальшого аналізу сцен, розпізнавання об'єктів і прийняття рішень у прикладних системах. Вона використовується у системах автономного керування транспортом [1-2], відеоспостереження та безпеки, медичній діагностиці [3], а також у промислових системах контролю якості продукції.

Однією з проблем сегментації зображень є наявність варіативного освітлення. Зміни інтенсивності освітлення зображень, наявність тіней, засвітів чи відблисків призводять до істотного викривлення значень пікселів, що ускладнює виділення границь та контурів об'єктів. У таких умовах навіть сучасні алгоритми сегментування демонструють нестабільність, оскільки однакові об'єкти можуть мати різні характеристики яскравості залежно від умов освітлення.

Додатковою складністю є те, що більшість класичних методів сегментації базуються на низькорівневих ознаках, таких як яскравість, колір або текстура. Проте ці ознаки суттєво змінюються під впливом освітлення, що призводить до втрати інваріантності моделей. Як наслідок, контури об'єктів стають нечіткими або зміщеними, а процес сегментації – менш точним, особливо в умовах слабкого або неоднорідного освітлення.

Сучасні підходи, зокрема методи на основі глибинного навчання, також стикаються з цією проблемою. Хоча вони здатні враховувати складні залежності в даних, їх ефективність значною мірою залежить від якості навчальних вибірок і здатності уза-

гальнювати варіації освітлення. У реальних умовах, де денні яскраві зображення з сонячними відблисками змінюються нічними темними сценами, моделі часто не здатні коректно виділяти семантичні та контурні ознаки об'єктів.

Таким чином, залежність результатів сегментації від стабільності вхідного візуального сигналу створює суттєвий бар'єр для впровадження інтелектуальних систем у динамічні середовища. З огляду на це, актуальною науково-практичною задачею є розробка програмного методу, що забезпечує інваріантність процесу сегментації до змін освітлення та гарантує високу точність формування масок сегментації незалежно від зовнішніх умов.

**Аналіз останніх досліджень та публікацій.** Проблема сегментації зображень активно досліджується протягом останніх десятиліть, що зумовлено її ключовою роллю у задачах комп'ютерного зору. Значна частина наукових робіт присвячена дослідженню впливу освітлення на якість сегментації. Встановлено [4], що нерівномірне або слабке освітлення призводить до розмиття контурів об'єктів і зниження точності їх виділення, що є однією з основних причин помилок сегментації. У роботі [5] також підкреслюється, що варіації освітлення суттєво впливають на ефективність методів виявлення контурів та об'єктів, особливо у складних сценах.

Окремий напрям досліджень зосереджений на розробленні методів, які враховують або компенсують зміну освітлення. Зокрема, підхід [6] на основі теорії Retinex дозволяє розділяти зображення на компоненти освітлення та відбиття, що підвищує якість сегментації в умовах низької освітленості. Подібні підходи також використовуються у роботі [7], присвяченій семантичній сегментації нічних сцен, де попереднє вирівнювання освітлення є критично важливим етапом. Результати експериментів на двох широко використовуваних наборах даних показують, що запропонований у дослідженні метод демонструє вищу ефективність порівняно з сучасними методами сегментації.

Крім того, досліджуються методи сегментації, що безпосередньо враховують характеристики освітлення сцени. Наприклад, підхід до сегментації на основі моделювання розподілу освітлення дозволяє покращити відокремлення об'єктів від фону за рахунок використання фізичних властивостей світла [4]. Робота [8] пропонує метод сегментації зображень на основі освітлення в сценах із декількома джерелами світла, що дозволяє виявляти області, де присутнє лише одне джерело світла.

Сучасним напрямом досліджень проблеми сегментації зображень в умовах варіативного освітлення є застосування методів глибинного навчання. Використання архітектури на основі згорткових нейронних мереж дозволяє моделям вивчати складні ієрархічні ознаки, що є більш стійкими до шумів порівняно з класичними детекторами. Дослідження [9] проводить порівняння різних моделей нейронних мереж для сегментації зображень за різних умов природнього освітлення. Результати оцінки ефективності моделей демонструють суттєву різницю в їх стійкості до змін освітлення, що підтверджує вплив змін навколишнього середовища на завдання семантичної сегментації. Робота [10] показує, що архітектури типу U-Net та її модифікації здатні досягати високої точності сегментації навіть у складних умовах освітлення, однак їх ефективність зале-

жить від різноманітності навчальних даних і наявності прикладів із різними сценаріями освітлення.

Таким чином, аналіз сучасних досліджень свідчить про значний прогрес у розвитку методів сегментації зображень, проте проблема забезпечення їх інваріантності до змін освітлення залишається актуальною. Це обумовлює необхідність подальших досліджень у напрямі розроблення програмних методів, здатних ефективно працювати в умовах варіативного освітлення.

**Мета дослідження.** Основною метою цієї роботи є розроблення методу семантичної сегментації зображень, що забезпечить високу точність та стабільність результатів в умовах варіативного освітлення. Важливою та необхідною частиною даної роботи є оцінка ефективності методу та її порівняння з відомими підходами.

**Викладення основного матеріалу дослідження.** Для вирішення проблеми семантичної сегментації зображень в умовах варіативного освітлення пропонується метод, що складається з наступних етапів:

- Класифікація зображення на основі ознак освітленості;
- Адаптивне попереднє оброблення зображення;
- Сегментація результуючого зображення нейронною мережею.

Етап гібридної класифікації рівня освітленості зображення передбачає автоматичне визначення класу зображення для подальшої стратегії адаптивного оброблення. Процес класифікації базується на розрахунку статистичних метрик відносної яскравості, що дозволяє зосередитися на аналізі освітленості сцени незалежно від кольорових характеристик. Нехай вхідне зображення задано у просторі RGB як  $I(x, y) = (R, G, B)$ , де  $R, G, B \in [0, 255]$ . Для забезпечення коректності обчислень значення пікселів нормалізуються до діапазону  $[0, 1]$ :

$$R' = \frac{R}{255}; G' = \frac{G}{255}; B' = \frac{B}{255}; \quad (1)$$

Далі для кожного пікселя обчислюється відносна яскравість  $L(x, y)$  як зважена сума нормалізованих компонентів:

$$L(x, y) = 0.2125R' + 0.715G' + 0.072B' \quad (2)$$

Вибір вагових коефіцієнтів зумовлений особливостями сприйняття яскравості людською зоровою системою, яка є найбільш чутливою до зеленого каналу та менш чутливою до синього. Отримана карта яскравості  $L(x, y) \in [0, 1]$  відображає розподіл освітленості сцени та використовується для подальшого аналізу.

На основі карти яскравості обчислюється набір статистичних характеристик, що описують розподіл освітленості в зображенні. Зокрема, визначаються середнє значення яскравості:

$$\mu = \frac{1}{N} \sum_{i=1}^N L_i \quad (3)$$

Медіана  $\tilde{L}$ , перцентилі  $P_{20}$  та  $P_{80}$ , що дозволяють оцінити розподіл яскравості без впливу викидів, а також стандартне відхилення яке характеризує контраст сцени:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (L_i - \mu)^2 \quad (4)$$

Додатково обчислюються частки темних і яскравих пікселів:

$$r_d = \frac{1}{N} \sum_{i=1}^N [L_i < \tau_d]; \quad r_b = \frac{1}{N} \sum_{i=1}^N [L_i > \tau_b], \quad (5)$$

де  $\tau_d$  та  $\tau_b$  – порогові значення для темних і яскравих областей відповідно. Також в аналізі використовується коефіцієнт асиметрії (skewness), який дозволяє визначити зміщення розподілу яскравості в сторону темних або світлих значень.:

$$\gamma = \frac{1}{N} \sum_{i=1}^N \frac{(L_i - \mu)^3}{\sigma}. \quad (6)$$

У результаті формується комплексний набір статистичних ознак, що описують глобальні характеристики освітленості зображення і слугують основою для подальшої класифікації умов освітлення на один із трьох класів – темне, помірне та світле:  $C = \{C_d, C_m, C_b\}$ . Класифікація здійснюється на основі вектору статичних ознак  $f$  та системи правил, що формалізуються у вигляді функції прийняття рішення:

$$D(f) = [r_d > \alpha_1] \cup [\tilde{L} < \beta_1 \cap P_{80} < \beta_2] \quad (7)$$

$$B(f) = [r_b > \alpha_2] \cup [\tilde{L} > \beta_3 \cap P_{80} > \beta_4] \quad (8)$$

$$C(f) = \begin{cases} C_d, & D(f) \\ C_b, & B(f) \\ C_d, & \sigma < \delta \cap \mu < 0.4 \\ C_b, & \sigma < \delta \cap \mu > 0.6 \\ C_d, & \gamma < -\kappa \\ C_b, & \gamma > \kappa \\ C_m, & \text{інакше} \end{cases} \quad (9)$$

У практичній реалізації використано такі емпірично визначені порогові значення:

- $\alpha_1 = \alpha_2 = 0.35$ ;
- $\beta_1 = 0.3, \beta_2 = 0.45$ ;
- $\beta_3 = 0.4, \beta_4 = 0.55$ ;
- $\delta = 0.08$ ;
- $\kappa = 0.5$ .

Запропонована класифікація ґрунтується на ознаках цифрового зображення – глобальній яскравості, розподілу інтенсивностей, контрасті і форми гістограми, що дозволяє підвищити стійкість методу до варіацій освітлення. На відміну від підходів, що ба-

зуються на одному показнику, така класифікація забезпечує більш надійну інтерпретацію умов освітлення у складних сценах.

Наступним етапом запропонованого методу є адаптивне попереднє оброблення зображення, що виконується відповідно до визначеного на попередньому етапі класу зображення. Основною метою цього етапу є нормалізація освітленості зображення та покращення його візуальних характеристик з урахуванням його початкових умов. Відповідно, для кожного класу зображення застосовується окремий алгоритм оброблення.

У випадку темних зображень використовується комбінований підхід, що включає локальне підсилення контрасту, гамма-корекцію та зменшення шуму. Спочатку зображення перетворюється у колірний простір LAB, де канал яскравості обробляється за допомогою методу CLAHE (Contrast Limited Adaptive Histogram Equalization), що дозволяє ефективно підсилити локальні деталі без перенасичення. Далі виконується адаптивна гамма-корекція, параметр якої визначається на основі середнього рівня яскравості зображення: темніші зображення будуть піддаватись більш інтенсивному підсиленню. На завершальному етапі застосовується алгоритм шумозаглушення для усунення артефактів, що виникають унаслідок підсилення.

Для зображень із надмірною яскравістю застосовується метод тонального відображення (tone mapping), зокрема алгоритм Рейнхарда, який забезпечує компресію динамічного діапазону. Це дозволяє зменшити перенасичені ділянки та відновити деталі у світлих областях зображення. Такий підхід є більш плавним порівняно з класичними методами нормалізації та краще зберігає природний вигляд сцени. Після тонального відображення може додатково застосовуватися шумозаглушення.

У випадку помірного освітлення використовується метод підвищення різкості на основі нерізкого маскування (unsharp masking). Для цього формується розмита версія зображення за допомогою Гаусового фільтра, яка потім віднімається від оригіналу з певним коефіцієнтом підсилення. Такий підхід дозволяє підкреслити межі об'єктів і текстурні деталі без суттєвого спотворення глобальної яскравості. За потреби застосовується шумозаглушення для збереження балансу між різкістю та рівнем шуму.

Загалом, запропонований підхід забезпечує адаптивну зміну параметрів оброблення залежно від умов освітлення, що дозволяє підвищити якість зображення перед виконанням сегментації. На відміну від класичних методів попереднього оброблення, така стратегія враховує специфіку вхідних даних і забезпечує більш стабільні результати в умовах варіативного освітлення.

Заключним етапом запропонованого методу є безпосередньо семантична сегментація зображення із використанням сучасних нейромережевих підходів. На цьому етапі зображення, отримане після адаптивного попереднього оброблення, подається на вхід моделі сегментації, яка виконує виділення об'єктів або областей інтересу. Запропонований метод не обмежується використанням конкретної архітектури нейронної мережі. Натомість передбачається можливість інтеграції будь-яких сучасних моделей сегментації, що дозволяє забезпечити гнучкість та адаптивність підходу до різних прикладних

задач. Зокрема, можуть бути використані як класичні архітектури, такі як U-Net та її модифікації, так і більш сучасні підходи, включаючи DeepLab, PSPNet, Mask R-CNN, а також трансформерні моделі сегментації.

**Аналіз отриманих результатів.** Для дослідження запропонованого методу було розроблено програмне забезпечення для сегментації зображень за допомогою мови програмування Python. Для виконання задач класифікації та попереднього оброблення зображень було використано бібліотеку NumPy. Створення та тренування моделей нейронних мереж відбувалось на основі бібліотеки PyTorch.

Для оцінки ефективності запропонованого методу було проведено експериментальне дослідження із використанням двох нейромережевих архітектур сегментації – U-Net та SegNet. Архітектура U-Net є однією з найбільш поширених у задачах сегментації завдяки симетричній енкодер-декодерній структурі та наявності skip-зв'язків, що дозволяють ефективно відновлювати просторову інформацію [11]. У свою чергу, SegNet базується на використанні індексів макс-пулінгу для відновлення просторової роздільної здатності, що забезпечує зменшення обчислювальної складності моделі [12].

Експеримент проводився на відкритому наборі даних для семантичної сегментації зображень – Car Segmentation Dataset [13]. Він містить зображення автомобілів та їх відповідні еталонні маски, що включають п'ять категорій об'єктів сегментації: колеса, фари, вікна, корпус машини та відповідно фон. Для імітації широкого спектра реальних умов освітлення вихідний набір даних підлягав керованій фотометричній деградації. Зокрема, моделювання сценаріїв низької освітленості здійснювалося шляхом нелінійного зниження рівня експозиції та накладання адитивного Гаусівського шуму, що імітує роботу сенсора камери при високих значеннях ISO. Водночас для відтворення ефектів надмірного освітлення та сонячних засвітів застосовувалося локальне та глобальне зміщення гістограми в область граничних високих значень (overexposure), що призводило до часткової втрати текстурної інформації у світлих зонах.

Під час тренувального процесу усі зображення були приведені до розміру 512x512 пікселів щоб забезпечити однакові розміри вхідних даних у наборі. Для навчання моделі було використано наступні параметри: кількість епох – 200, розмір батчу – 16, функція втрат – CrossEntropyLoss, та оптимізатор Adam з початковою швидкістю навчання 0.0005. Для коригування швидкості навчання було використано планувальник ReduceLROnPlateau.

Оцінка результатів точності та ефективності методу проводилася за допомогою стандартних метрик сегментації Pixel Accuracy (PA, попиксельна точність) та Mean Intersection over Union (mIoU, коефіцієнт перетину) як на тренувальному, так і на валідаційному наборах даних. Вибір цих показників забезпечує всебічну оцінку точності класифікації окремих пікселів та загальної узгодженості сформованих сегментаційних масок із еталонною розміткою. Оцінка кількісних показників ефективності, що наведена в таблиці 1, доводить, що застосування запропонованого методу попередньої обробки дозволяє покращити якість сегментації для обох моделей. Важливо відзначити, що покращення результатів має систематичний характер і спостерігається незалежно

від обраної архітектури нейронної мережі, що свідчить про ефективність запропонованого підходу.

Таблиця 1

Оцінка кількісних показників ефективності запропонованого методу

Метод сегментації	Тренувальна вибірка		Валідаційна вибірка	
	РА	mIoU	РА	mIoU
SegNet	0,9586	0,7563	0,9398	0,7049
Запропонований метод на основі SegNet	0,9644	0,7736	0,9536	0,7398
U-Net	0,9743	0,8330	0,9555	0,7636
Запропонований метод на основі U-Net	0,9794	0,8619	0,9562	0,7763

Аналіз динаміки навчання моделей показав, що використання адаптивного попереднього оброблення сприяє більш стабільній збіжності процесу навчання. Зокрема, графіки зміни значення функції втрат (рис. 1) демонструють більш плавне зменшення значень функції втрат (loss) та зменшення коливань у процесі навчання. Це пояснюється тим, що нормалізація освітлення зменшує варіативність вхідних даних, що, у свою чергу, полегшує процес оптимізації параметрів нейронної мережі.

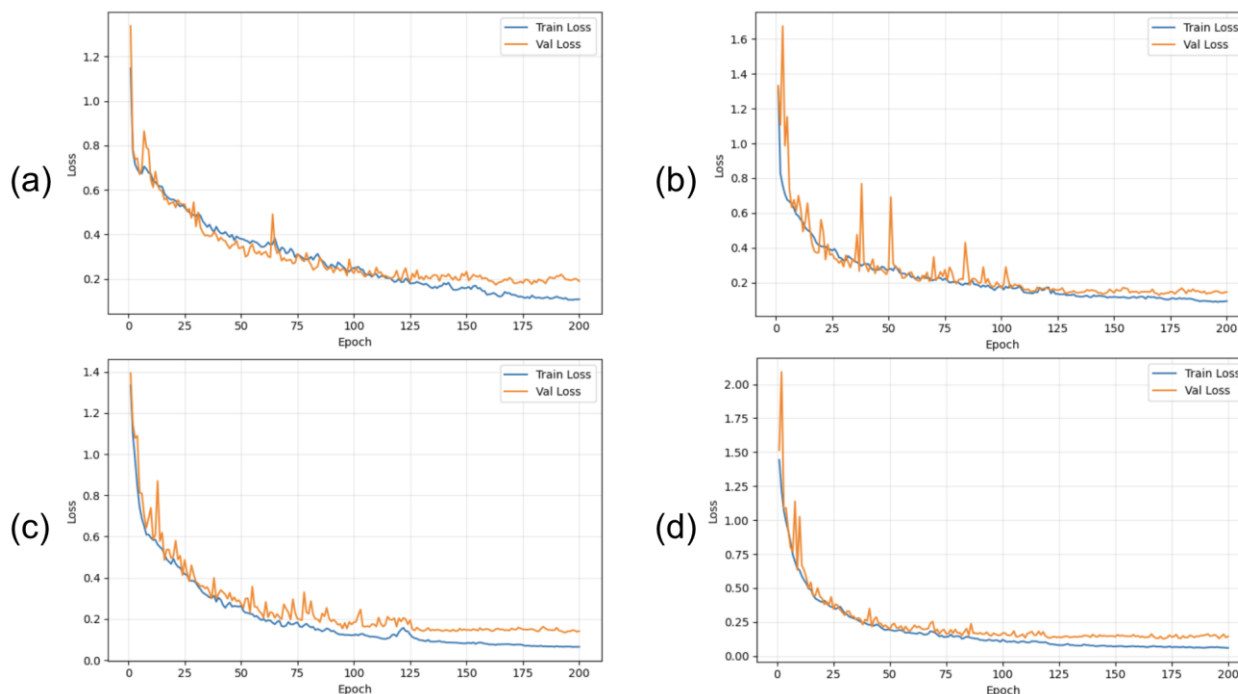


Рисунок 1 – Графік зміни значення функції втрат: (a) SegNet, (b) запропонований метод на основі SegNet, (c) U-Net, (d) запропонований метод на основі U-Net

Якісний аналіз результатів сегментації підтверджує отримані кількісні оцінки. У випадку використання попереднього оброблення спостерігається більш точне виділення контурів об'єктів, зменшення кількості хибних спрацьовувань та покращення сегментації в складних умовах освітлення, зокрема у темних або пересвітлених областях. Без попереднього оброблення моделі частіше допускають помилки у вигляді втрати дрібних деталей або неправильного визначення контурів об'єктів.

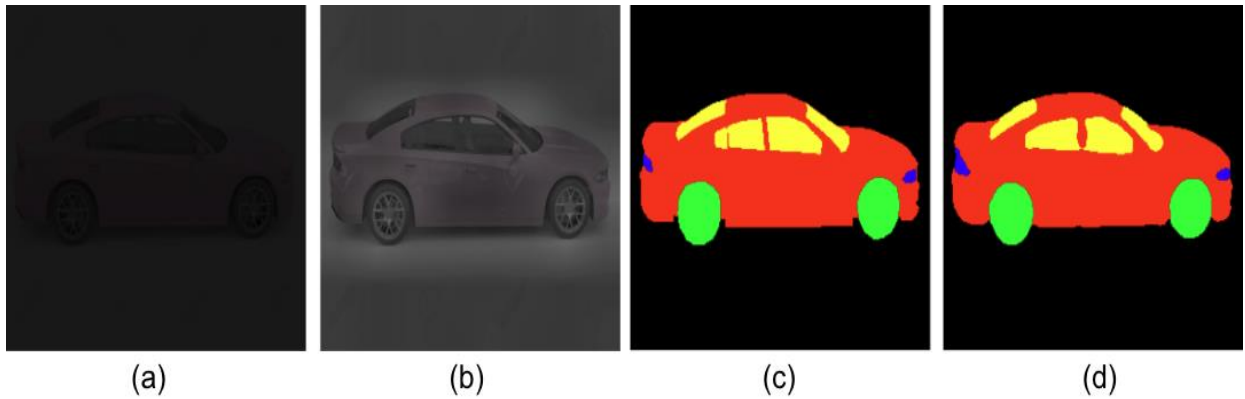


Рисунок 2 – Приклад роботи запропонованого методу: (a) темне вхідне зображення, (b) адаптивно-покращене зображення, (c) еталонна маска, (d) передбачена маска

Таким чином, результати експериментального дослідження підтверджують ефективність запропонованого методу. Його застосування дозволяє підвищити точність сегментації та стабільність навчання нейромережових моделей без необхідності зміни їх архітектури, що робить підхід придатним для широкого спектра задач комп'ютерного зору.

**Висновки.** У роботі розглянуто проблему семантичної сегментації зображень в умовах варіативного освітлення та проаналізовано вплив світлових спотворень на якість виділення об'єктів. Встановлено, що нестабільність освітлення є одним із ключових факторів, який знижує точність сегментації як у класичних, так і в сучасних нейромережових підходах.

У результаті дослідження розроблено програмний метод сегментації зображень, що базується на поєднанні трьох етапів: класифікації зображення на основі рівня освітлення, адаптивного попереднього оброблення зображення та нейромережової сегментації. Запропонований метод дозволяє автоматично визначати умови освітлення зображення на основі статистичних характеристик відносної яскравості та застосовувати відповідні алгоритми оброблення для нормалізації вхідних даних.

Експериментальна перевірка методу із використанням моделей U-Net та SegNet показала, що застосування адаптивного попереднього оброблення забезпечує підвищення точності сегментації за метриками Pixel Accuracy та Mean Intersection over Union як на тренувальній, так і на валідаційній вибірках. Отримані результати демонструють систематичне покращення показників незалежно від обраної архітектури, що підтверджує ефективність запропонованого підходу. Аналіз динаміки навчання моделей засвідчив, що нормалізація освітлення сприяє більш стабільній збіжності процесу оптимізації, зменшенню коливань функції втрат та покращенню узагальнювальної здат-

ності моделей. Якісний аналіз результатів сегментації також підтвердив підвищення точності виділення контурів об'єктів та зменшення кількості помилок у складних умовах освітлення.

Таким чином, запропонований метод дозволяє ефективно нівелювати вплив варіативного освітлення без необхідності модифікації архітектури нейронних мереж, що робить його придатним для інтеграції у широкий спектр прикладних систем комп'ютерного зору.

Перспективами подальших досліджень є вдосконалення механізму адаптації методу до складніших умов освітлення, зокрема сцен із динамічним освітленням та наявністю кількох джерел світла. Доцільним є також розширення підходу шляхом урахування не лише глобальних, але й локальних характеристик освітленості, що дозволить підвищити точність оброблення неоднорідно освітлених зображень. Окрім цього, важливим є дослідження ефективності методу на більш складних та різноманітних наборах даних, включаючи реальні сценарії з багатьма джерелами освітлення.

#### ЛІТЕРАТУРА / REFERENCES

1. Anwesh, K., Pal, D., Ganguly, D., Chatterjee, K., & Roy, S. (2022). Number plate recognition from enhanced super-resolution using generative adversarial network. *Multimedia Tools and Applications*, 1–17. <https://doi.org/10.1007/s11042-022-14018-0>
2. Kinahan, J., & Smeaton, A. F. (2021). *Image segmentation to identify safe landing zones for unmanned aerial vehicles*. *arXiv*. <https://doi.org/10.48550/arXiv.2111.14557>
3. Jin, B., Cruz, L., & Gonçalves, N. (2020). Deep facial diagnosis: Deep transfer learning from face recognition to facial diagnosis. *IEEE Access*, 8, 123649–123661. <https://doi.org/10.1109/ACCESS.2020.3005687>
4. Zhang, L. (2008). In situ image segmentation using the convexity of illumination distribution of the light sources. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(10), 1786–1799. <https://doi.org/10.1109/TPAMI.2007.70830>
5. Abosinnee, A. S., Bencsik, G., & Abedi, F. (2025). Edges in image with illumination variations scenarios: A review. *The Visual Computer*, 41, 12277–12305. <https://doi.org/10.1007/s00371-025-04157-4>
6. Chen, Y., Wen, C., Liu, W., & Li, G. (2023). A depth iterative illumination estimation network for low-light image enhancement based on Retinex theory. *Scientific Reports*, 13(1), 19709. <https://doi.org/10.1038/s41598-023-46693-w>
7. Sun, Z., Zhu, H., Xiao, X., Gu, Y., & Xu, Y. (2024). Nighttime image semantic segmentation with Retinex theory. *Image and Vision Computing*, 148, 105149. <https://doi.org/10.1016/j.imavis.2024.105149>
8. Vrshnak, D., Domislović, I., Subašić, M., & Lončarić, S. (2022). Illuminant segmentation for multi-illuminant scenes using latent illumination encoding. *Signal Processing: Image Communication*, 108, 116822. <https://doi.org/10.1016/j.image.2022.116822>
9. Nugraha, D. G., Purwanto, D., Dikairono, R., Widjiati, E., & Adinanta, H. (2025). Comparative Analysis of Image Segmentation Methods for Unmanned Surface Vehicles under Varying Illumination Conditions. In *26th International Seminar on Intelligent Technology and*

Its Applications: Fostering Equal Opportunities for Breakthrough Technology Innovations, ISITIA 2025 - Proceedings (2025 ed., pp. 88-93). Institute of Electrical and Electronics Engineers Inc.. <https://doi.org/10.1109/ISITIA66279.2025.11137505>

10. Cardenas-Gallegos, J. S., Severns, P. M., Klimeš, P., Lacerda, L. N., Peduzzi, A., & Ferrarezi, R. S. (2025). Reliable plant segmentation under variable greenhouse illumination conditions. *Computers and Electronics in Agriculture*, 229, 109711.

<https://doi.org/10.1016/j.compag.2024.109711>

11. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. <https://doi.org/10.48550/arXiv.1505.04597>

12. Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE TPAMI*.

<https://doi.org/10.1109/TPAMI.2016.2644615>

13. Intel AI. (2019). Car segmentation [Data set]. Kaggle.

<https://www.kaggle.com/datasets/intelecai/car-segmentation>

Received 01.04.2026.

Accepted 03.04.2026.

Published 30.04.2026

### ***Software method for image segmentation under variable lighting conditions***

*This paper addresses the problem of image segmentation under variable lighting conditions, which is a critical factor in computer vision tasks. Various approaches to image processing are analyzed, including classical methods and modern neural network models. It has been established that traditional approaches have limitations due to their dependence on low-level features, whereas neural network methods demonstrate reduced performance under unstable lighting conditions.*

*The purpose of the research is to develop an image segmentation method that ensures adaptation to lighting conditions and improves the accuracy of object detection. A software method is proposed that combines an illumination level classification, adaptive image preprocessing, and neural network segmentation. Images are classified based on an analysis of statistical characteristics of relative luminance, which allows determining the processing algorithm for each image.*

*The effectiveness of the method was evaluated through experimental research using the U-Net and SegNet neural network models. A set of images with simulated lighting conditions was used for the analysis. As a result of the experiments, the proposed method demonstrated improvements of 4.95% and 1.5%, respectively, in terms of the mIoU and Pixel Accuracy metrics for the SegNet model, as well as 1.6% and 0.07% for the U-Net model. A comparative analysis has shown that the use of adaptive image preprocessing improves segmentation quality regardless of the chosen neural network architecture.*

*The proposed method increases the robustness of segmentation systems to lighting changes, reduces the impact of noise and distortions, and ensures more stable model performance in real-world conditions. This makes it an effective solution for a wide range of computer vision tasks, particularly in systems operating in dynamic environments.*

**Слободзян Максим Вікторович** – магістр кафедри програмного забезпечення комп’ютерних систем, Національний технічний університет України “Київський політехнічний інститут імені Ігоря Сікорського”, Київ.

ORCID: <https://orcid.org/0009-0009-9324-9430>

**Шкурат Оксана Сергіївна** – доцент кафедри програмного забезпечення комп’ютерних систем, Національний технічний університет України “Київський політехнічний інститут імені Ігоря Сікорського”, Київ.

ORCID: <https://orcid.org/0000-0001-7633-9121>

**Slobodzian Maksym** – master of Computer Systems Software Department, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv.

ORCID: <https://orcid.org/0009-0009-9324-9430>

**Shkurat Oksana** – Associate Professor of Computer Systems Software Department, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv.

ORCID: <https://orcid.org/0000-0001-7633-9121>

## ВІЗУАЛЬНО-ІНЕРЦІЙНИЙ SLAM З ВИКОРИСТАННЯМ ІНВАРІАНТНОГО РОЗШИРЕНОГО ФІЛЬТРУ КАЛМАНА ДЛЯ АВТОНОМНОЇ НАВІГАЦІЇ

*Анотація.* Дослідження присвячено розвитку та застосуванню візуально-інерційного SLAM на основі ітерованого розширеного фільтра Калмана (IEKF) для задач автономної навігації мобільних платформ. Класичні підходи на базі розширеного фільтра Калмана (EKF) мають суттєві обмеження: помилки однократної лінеаризації у необоротних станах призводять до явища хибної спостережуваності. Це викликає штучне звуження коваріаційних матриць та експоненційне накопичення дрейфу під час складних маневрів. Для вирішення цих проблем у роботі математично обґрунтовано та реалізовано комплексний підхід, що поєднує параметризацію оберненої глибини (IDP) для миттєвої ініціалізації орієнтирів та апарат теорії матричних груп Лі  $SE_2(3)$ . Перетворення етапу оновлення вимірювань на задачу нелінійної оптимізації дозволяє системі зберігати сувору геометричну узгодженість. Для валідації методу розроблено середовище комп'ютерного моделювання з трьома сценаріями: замкнена петля, міський квартал із різкими поворотами та траєкторія у формі "вісімки". Результати серії експериментів підтверджують, що запропонований IEKF-SLAM забезпечує в середньому в 3,5 рази вищу точність побудови траєкторії за метрикою абсолютної похибки (ATE). Алгоритм практично повністю усуває дрейф кута ризику у порівнянні з класичним EKF-SLAM (відхилення становить  $< 3^\circ$  проти  $\sim 25^\circ$ ). Також доведено, що інваріантний фільтр здатен ефективно використовувати більшу кількість візуальних орієнтирів (80–100 точок) без втрати консистентності. Незважаючи на збільшення обчислювальних витрат приблизно на 30%, алгоритм зберігає здатність стабільно працювати в режимі реального часу, що робить його оптимальним рішенням для сучасних автономних систем.

*Ключові слова:* візуально-інерційний SLAM, ітерований розширений фільтр Калмана, групи Лі, параметризація оберненої глибини, автономна навігація, хибна спостережуваність.

**Постановка проблеми.** Одночасна локалізація та картографування (SLAM) є однією з найважливіших технологій у сучасній робототехніці, яка дозволяє автономним агентам (мобільним роботам, безпілотним літальним апаратам, автономним автомобілям) створювати карту невідомого середовища та одночасно визначати власне просторове положення в межах цієї карти. Вирішення цієї проблеми є передумовою для досягнення справжньої автономності, оскільки без оцінки власного стану планування траєкторії та безпечна навігація стають неможливими [1].

У реальних умовах експлуатації автономна система постійно стикається з перешкодами, динамічними об'єктами, змінами освітлення та, найголовніше, із шумом сенсорних вимірювань. Жоден окремих датчик не здатний забезпечити ідеальну точність протягом тривалого часу. Наприклад, камери забезпечують високу роздільну здатність і багату текстурну інформацію, необхідну для розпізнавання місць (loop closure) та побудови точних карт, але вони вразливі до розмиття в русі (motion blur), оклюзій, раптових змін освітленості та не забезпечують метричного масштабу у випадку монокулярного зору. З іншого боку, інерційні вимірювальні пристрої (ІВП), що складаються з гіроскопів та акселерометрів, забезпечують високу частоту оновлення даних, стабільність роботи за будь-яких візуальних умов та визначають масштаб і вектор гравітації, проте вони схильні до швидкого накопичення похибок інтегрування (дрейфу) [2].

Саме тому візуально-інерційний SLAM (VI-SLAM) став стандартом для автономних систем, поєднуючи переваги камер та ІВП для створення стабільних і достовірних оцінок траєкторії. Візуально-інерційні системи є перспективними як ефективна альтернатива супутниковим системам навігації (GPS/GNSS) у середовищах, де їх застосування є неможливим: у приміщеннях, підземних шахтах, міських каньйонах або в умовах навмисного радіоелектронного придушення під час бойових дій чи рятувальних операцій.

**Аналіз останніх досліджень та публікацій.** Історично підходи до вирішення задачі SLAM розділилися на дві великі категорії: методи на основі оптимізації (Bundle Adjustment, граф-SLAM) та методи на основі фільтрації. Оптимізаційні методи забезпечують високу точність, але потребують значних обчислювальних ресурсів і пам'яті, що часто обмежує їх застосування на вбудованих платформах реального часу. Методи на основі розширеного фільтра Калмана традиційно були кращим вибором завдяки їхній рекурсивній природі, низькій обчислювальній складності та здатності працювати в жорстких обмеженнях реального часу [3-5].

Проте базовий EKF SLAM має низку фундаментальних математичних недоліків, що суттєво обмежують його застосування у масштабних або високодинамічних сценаріях, особливо при використанні монокулярних камер. Головною проблемою є лінеаризація суттєво нелінійних моделей проективної геометрії та спостереження за допомогою рядів Тейлора першого порядку навколо поточної оцінки стану. Оскільки поточна оцінка неминуче містить похибку, обчислені матриці Якобі стають неточними. Ця неточність призводить до феномену втрати консистентності (inconsistency) [6-7].

Консистентність фільтра означає, що коваріаційна матриця системи адекватно відображає реальну похибку оцінки. У випадку EKF SLAM, через похибки лінеаризації, система помилково вважає, що отримує інформацію про стани, які є не спостережуваними (наприклад, абсолютні координати глобального положення та абсолютна орієнтація відносно магнітної півночі). Це явище, відоме як хибна спостережуваність (spurious observability), змушує фільтр штучно зменшувати коваріацію цих станів. Як наслідок, фільтр ігнорує нові вимірювання, і вся система незворотно дрейфує або дивергує. Також монокулярний зір не може миттєво оцінити глибину нових візуальних орієнтирів,

що призводить до проблеми ініціалізації із затримкою (delayed initialization) в декартових координатах.

Зсув парадигми у фільтраційному SLAM відбувся завдяки двом ключовим математичним інноваціям останніх років. По-перше, впровадження параметризації оберненої глибини дозволило коректно моделювати променеву невизначеність та ініціалізувати об'єкти миттєво, незалежно від їхньої віддаленості. По-друге, перехід до ітерованого розширеного фільтра Калмана дозволив перетворити крок оновлення на задачу нелінійної оптимізації найменших квадратів, що багаторазово перераховує матриці Якобі, мінімізуючи помилки усічення Тейлора [8].

Подальший розвиток ІЕКФ, що спирається на глибоку математичну основу теорії груп Лі, дозволив експлуатувати внутрішні геометричні симетрії динамічних систем. Замість того, щоб визначати похибку стану як звичайну векторну різницю в евклідовому просторі, інваріантний ІЕКФ формулює похибку мультиплікативно за допомогою групових операцій. Таке переформулювання має потужний наслідок: динаміка лінеаризованої похибки стає константною відносно невідомого істинного стану. Це означає, що матриця спостережуваності системи ідеально збігається з фактичною матрицею спостережуваності реальної нелінійної системи, що робить алгоритм високонадійним, консистентним та надзвичайно стійким до інтенсивної динаміки руху.

**Мета дослідження.** Метою даної роботи є дослідження, математичне обґрунтування та експериментальна перевірка застосування монокулярного візуально-інерційного SLAM на основі ітерованого розширеного фільтра Калмана у комбінації з параметризацією оберненої глибини. У статті наводиться математичний опис алгоритму та аналіз спостережуваності. Для порівняння класичного ЕКФ та запропонованого ІЕКФ використовується спеціалізоване середовище комп'ютерного моделювання, що дозволило згенерувати контрольовані сценарії руху (замкнена петля, міський квартал та траєкторія "Вісімка").

## **Виклад основного матеріалу**

### **1. Проблематика монокулярного SLAM та обмеження класичної фільтрації**

Одночасна локалізація та картографування з використанням монокулярної камери є складною математичною задачею через суто проєктивну природу сенсора. Камера вимірює лише кутові пеленги до візуальних орієнтирів, втрачаючи інформацію про їхню абсолютну глибину та масштаб сцени. Оцінка глибини вимагає трансляційного руху камери для формування паралаксу.

Класичний підхід на базі розширеного фільтра Калмана стикається тут із труднощами. ЕКФ базується на апроксимації нелінійних моделей за допомогою розкладу в ряд Тейлора першого порядку. Проте в монокулярному SLAM проєктивна геометрія є нелінійною. Однократна лінеаризація, властива ЕКФ, призводить до накопичення помилок усіченого ряду Тейлора, що спричиняє неузгодженість: коваріаційні матриці звужуються (overconfidence), і фільтр розходиться [9-13].

Для подолання цих обмежень сучасний підхід спирається на інтеграцію ітерованого розширеного фільтра Калмана у комбінації з параметризацією оберненої глибини та архітектурою контролю спостережуваності (FEJ).

## 2. Формалізація простору стану та кінематичної моделі

Алгоритм підтримує єдиний глобальний вектор стану системи  $x_k$ , який інкапсулює як кінематичний стан камери  $x_{v,k}$ , так і геометричні координати  $n$  візуальних орієнтирів  $y_{i,k}$ :

$$x_k = [x_{v,k}^T y_{1,k}^T \dots y_{n,k}^T] \quad (1)$$

Вектор стану монокулярної камери (6-DOF) моделюється 13 параметрами для забезпечення роботи в умовах вільного просторового руху без колісної одометрії:

$$x_v = [r_{WC}^T q_{WC}^T v_W^T W_C^T]^T \quad (2)$$

де  $r_{WC} \in \mathbb{R}^3$  - позиція оптичного центру у глобальній системі  $W$ ;  $q_{WC}$  - одиничний кватерніон орієнтації, що запобігає шарнірному замку (gimbal lock);  $v_W \in \mathbb{R}^3$  - вектор лінійної швидкості;  $W_C \in \mathbb{R}^3$  - вектор кутової швидкості у локальній системі камери  $C$ .

Етап передбачення (prediction step) базується на моделі постійної швидкості. Невідомі лінійні  $a_W$  та кутові  $a_C$  прискорення розглядаються як процеси білого Гаусівського шуму. Інтегрування диференціальних рівнянь переходу стану камери  $x_{v,k+1} = f_v(x_{v,k}, w_k)$  за інтервал  $\Delta t$  має вигляд:

$$f_v = \begin{vmatrix} r_{WC,k} + (v_{W,k} + V_W)\Delta t \\ q_{WC,k}q(w_{C,k} + \Omega_C)\Delta t \\ v_{W,k} + V_W \\ w_{C,k} + \Omega_C \end{vmatrix} \quad (3)$$

де  $V_W = a_W t$ ,  $\Omega_C = a_C t$ , а функція  $q()$  формує кватерніон обертання. Оскільки візуальні орієнтири статичні, їхня модель переходу є тотожністю  $y_{i,k+1} = y_{i,k}$ . Коваріація передбаченого стану обчислюється класичним рівнянням  $P_{k+1|k} = F P_{k|k} F^T + G Q G^T$ , де  $F$  та  $G$  - розріджені матриці Якобі кінематичної моделі.

## 3. Моделювання орієнтирів: Параметризація оберненої глибини (IDP)

Класична декартова система (XYZ) не здатна змоделювати негаусівську променеву невизначеність при першому спостереженні об'єкта, що змушує відкладати ініціалізацію точок. Натомість параметризація оберненої глибини (IDP) кодує орієнтир  $y_i$  як 6-вимірний вектор:

$$y_i = [r_i^T \theta_i \phi_i \rho_i]^T \quad (4)$$

де  $r_i$  - координати камери в момент першого виявлення (точка прив'язки),  $\theta_i$  та  $\phi_i$  - азимут і кут місця напрямкового вектора  $m(\theta_i, \phi_i)$ , а  $\rho_i = 1/d_i$  - обернена глибина.

Перехід до Евклідового простору здійснюється за відображенням:

$$[X_i \ Y_i \ Z_i] = r_i + \frac{1}{\rho_i} m(\Theta_i, \Phi_i) \quad (5)$$

Ініціалізація без затримки (undelayed initialization): Лінійність залежності дисперсії пікселів від  $\rho_i$  дозволяє додавати точки до фільтра. Оберненій глибині при-  
своюється значення  $\rho_0$  та стандартне відхилення  $\sigma_\rho$ , що покриває діапазон від об'єкти-  
ва до нескінченності ( $\rho_i \rightarrow 0$ ). Це важливо для оцінки орієнтації камери за віддаленими  
об'єктами.

#### 4. Модель вимірювання та аналітичні Якобіани

Проекція 6-вимірному IDP-орієнтира  $y_i$  на площину сенсора виконується через  
трансформацію вектора  $h_W$  у локальну систему камери  $C$ :

$$h_C = R_{CW} \left( r_i + \frac{1}{\rho_i} m(\Theta_i, \Phi_i) \right) - r_{WC} \quad (6)$$

Для уникнення сингулярності при ( $\rho_i \rightarrow 0$ ), вектор масштабується:

$$h_C = \rho_i h_C = R_{CW} (\rho_i (r_i - r_{WC}) + m(\Theta_i, \Phi_i)) \quad (7)$$

Нормалізовані координати перетинають площину зображення і спотворюються  
дисторсією, після чого переводяться у пікселі  $(u, v)$  через матрицю внутрішніх пара-  
метрів  $K$ . Повне рівняння вимірювання:  $z_k = h(x_k) + n_k$ , де  $n_k \sim N(0, R_k)$ .

Для ітеративної оптимізації розраховується Якобіан  $H = \frac{\partial h(x)}{\partial x}$  за допомогою  
ланцюгового правила.

Зокрема, похідна переносного вектора по позиції камери доводить фундаменталь-  
ну властивість IDP - паралакс залежить від оберненої глибини:

$$\frac{\partial h_{transl}}{\partial r_{WC}} = I_{3 \times 3} \rho_i \quad (8)$$

Похідна за оберненою глибиною показує, що оцінка  $\rho_i$  оновлюється лише за на-  
явності базової лінії:

$$\frac{\partial h_{transl}}{\partial \rho_i} = r_i - r_{WC} \quad (9)$$

#### 5. Оптимізація Гаусса-Ньютона (Етап оновлення IEKF)

На відміну від EKF, IEKF переформулює крок оновлення як задачу оцінки мак-  
симальної вірогідності (MLE), мінімізуючи цільову функцію Махаланобіса  $c(x)$ :

$$c(x) = \frac{1}{2} (z_k - h(x))^T R^{-1} (z_k - h(x)) + \frac{1}{2} (x - \widehat{x}_{k|k-1})^T P_{k|k-1}^{-1} (x - \widehat{x}_{k|k-1}) \quad (10)$$

Мінімізація здійснюється алгоритмом Гаусса-Ньютона. На кожній внутрішній  
ітерації  $j$  модель перелінеаризується ( $R^{(j)}$ ), а стан оновлюється через ітеративний  
коєфіцієнт підсилення  $K^{(j)}$ :

$$K^{(j)} = P_{k|k-1} (H^{(j)})^T (H^{(j)} P_{k|k-1} H^{(j)T} + R)^{-1} \quad (11)$$

$$x^{(j+1)} = x^{(0)} + \gamma_j K^{(j)} (z_k - h(x^{(j)}) - H^{(j)} (x^{(0)} - x^{(j)})) \quad (12)$$

Скалярний параметр  $\gamma_j \in [0,1]$  реалізує механізм зворотного пошуку (backtracking). Якщо нелінійність спричиняє зростання функції вартості (overshooting),  $\gamma_j$  зменшується, гарантуючи монотонну збіжність. Ітерації тривають доки норма зміни стану не впаде нижче порогу  $\epsilon$ . Лише після знаходження оптимальної апостеріорної оцінки  $\widehat{x}_{k|k}$ , коваріаційна матриця  $P_{k|k}$  оновлюється одноразово:

$$P_{k|k} = (I - K^{(N)} H^{(N)}) P_{k|k-1} \quad (13)$$

### 6. Контроль спостережуваності та архітектура FEJ-IEKF

Система монокулярного SLAM має 4 неспостережні ступені свободи (3D-позиція та кут рискання/Yaw) плюс масштаб. У стандартному IEKF зміна точок лінеаризації на кожній ітерації порушує алгебраїчну симетрію матриці спостережуваності. Це призводить до набуття хибної інформації (spurious information gain): фільтр починає вважати неспостережні параметри вимірюваними, звужує їхню коваріацію і розходиться.

Для блокування цієї проблеми застосовується архітектура First-Estimates Jacobian. Алгоритм фіксує точки лінеаризації: похідні для кожної змінної розраховуються навколо тієї оцінки, коли ця змінна вперше з'явилася в системі. Це гарантує, що розмірність неспостережного підпростору залишається математично коректною.

### 7. Оптимізація розмірності простору: Індекс лінійності $\square$

Оскільки кожна IDP-ознака потребує 6 параметрів, обчислювальна складність IEKF швидко зростає ( $O(n^3)$ ). Для роботи в реальному часі застосовується гібридна архітектура: коли камера набирає достатній паралакс, орієнтир конвертується з IDP назад у 3D-декартовий простір (3 параметри).

Критерій перемикавання базується на безрозмірнісному індексі лінійності  $L$ , що оцінює співвідношення між Гессіаном та Якобіаном в межах зони невизначеності  $2\sigma_x$ :

$$L = \left| \frac{\frac{\partial^2 f}{\partial x^2} \mu_x}{\frac{\partial f}{\partial x} \mu_x} 2\sigma_x \right| \quad (14)$$

Якщо  $L \approx 0.1$ , трансформація вважається безпечною. Блок вектора  $y_j$  замінюється на  $x_i^{3D}$ , а коваріаційна матриця трансформується через матрицю Якобі переходу  $J = \frac{\partial x_i^{3D}}{\partial y_j}$  за правилом проєкції  $P_{new} = J P_{old} J^T$ . Це дозволяє скоротити розмірність матриць без втрати консистентності фільтра.

**Опис наборів даних для досліджень.** Для оцінювання ефективності та порівняння алгоритмів класичного розширеного фільтра Калмана та інваріантного розширеного фільтра Калмана було розроблено спеціалізоване середовище комп'ютерного моделювання. Використання синтетичних наборів даних дозволяє здійснювати контрольований

експеримент: точно задавати еталонну траєкторію (ground truth), керувати рівнями сенсорного шуму та ізолювати специфічні ситуації, які викликають деградацію роботи фільтрів.

У ході досліджень було згенеровано три різні сценарії руху, кожен з яких імітує специфічні умови автономної навігації мобільних роботів та транспортних засобів:

- **Сценарій 1:** Замкнена петля (базовий тест) - цей набір даних імітує плавний круговий рух із поверненням у стартову точку. Траєкторія складається з 500 кадрів і призначена для базової перевірки здатності алгоритмів до замикання петлі за умови накопичення поступового дрейфу. У цьому сценарії досліджується вплив квадратичного накопичення похибок кута ризику, характерного для класичного EKF, та здатність IEKF компенсувати його завдяки інваріантній формі похибки.

- **Сценарій 2:** Міський квартал (тест на різкі маневри) - цей маршрут розроблено для імітації руху міськими вулицями з постійною лінійною швидкістю 10 м/с (близько 36 км/год) протягом 130 секунд. Траєкторія складається з довгих прямолінійних відрізків, що чергуються з різкими поворотами на 90 та 60 градусів (імітація перехресть та об'їздів перешкод). Цей набір даних є важливим для перевірки явища "хибної спостережуваності" (false observability). Він дозволяє проаналізувати, як різкі нелінійні маневри впливають на стрибкоподібне зростання похибки напрямку руху в базовому EKF, і як IEKF зберігає узгодженість під час таких маневрів.

- **Сценарій 3:** Траєкторія "Вісімка" (тест на безперервну нелінійність) - маршрут імітує рух звивистою дорогою зі швидкістю 12 м/с (близько 43 км/год) тривалістю 100 секунд. Кутова швидкість змінюється за синусоїдальним законом, створюючи тривалі повороти ліворуч і праворуч, що утворюють фігуру "вісімки". На відміну від міського сценарію, де алгоритм має час на стабілізацію між маневрами, цей набір даних піддає матрицю Якобіана безперервному обертальному впливу, тестуючи стійкість алгоритмів до втрати масштабу та форми траєкторії (scale and yaw drift).

Для кожного сценарію моделювалися дані інерційного вимірювального пристрою (лінійні прискорення та кутові швидкості) з додаванням білого гаусового шуму та постійного зміщення. Візуальні орієнтири генерувалися вздовж траєкторії руху, імітуючи роботу алгоритму екстракції візуальних ознак, з динамічним обмеженням кількості активних орієнтирів для аналізу обчислювальної складності алгоритмів.

**Результати експериментальних досліджень.** Для оцінювання якості роботи алгоритмів монокулярного візуально-інерційного SLAM використовувалася метрика абсолютної похибки траєкторії, яка визначає середнє евклідове відхилення між оціненою алгоритмом траєкторією та еталонною у глобальному просторі. Оскільки алгоритми монокулярного зору найбільш вразливі до втрати орієнтації, окремо аналізувався дрейф кута ризику (yaw drift). Крім того, проводилася оцінка обчислювальної складності обох методів для підтвердження їхньої придатності до роботи в режимі реального часу [14-15].

#### **Експеримент 1: Базове замикання петлі та вплив щільності візуальної карти**

Результати показали, що класичний EKF-SLAM демонструє значне просторове відхилення наприкінці маршруту, через що віртуальна петля не замикається. Натомість

IEKF-SLAM, завдяки мультиплікативній інваріантній природі похибки на групі  $SE_2(3)$ , зберігає коректну структуру коваріаційної матриці і повертається у стартову точку з мінімальним метричним відхиленням. Окрему увагу було приділено дослідженню впливу максимальної кількості активних візуальних орієнтирів ( $N_{max}$ ) на точність системи.

Емпіричний аналіз показав, що для класичного EKF-SLAM оптимальна кількість орієнтирів становить близько 60 точок. При подальшому збільшенні  $N_{max}$  (до 100-120) точність EKF починає не покращуватися, а деградувати. Це пояснюється тим, що кожне нове спостереження вимагає обчислення матриці Якобі. У стандартному EKF ці матриці обчислюються навколо поточних, зашумлених оцінок стану. Велика кількість орієнтирів генерує масив дрібних похибок лінеаризації, змушуючи алгоритм розходитися.

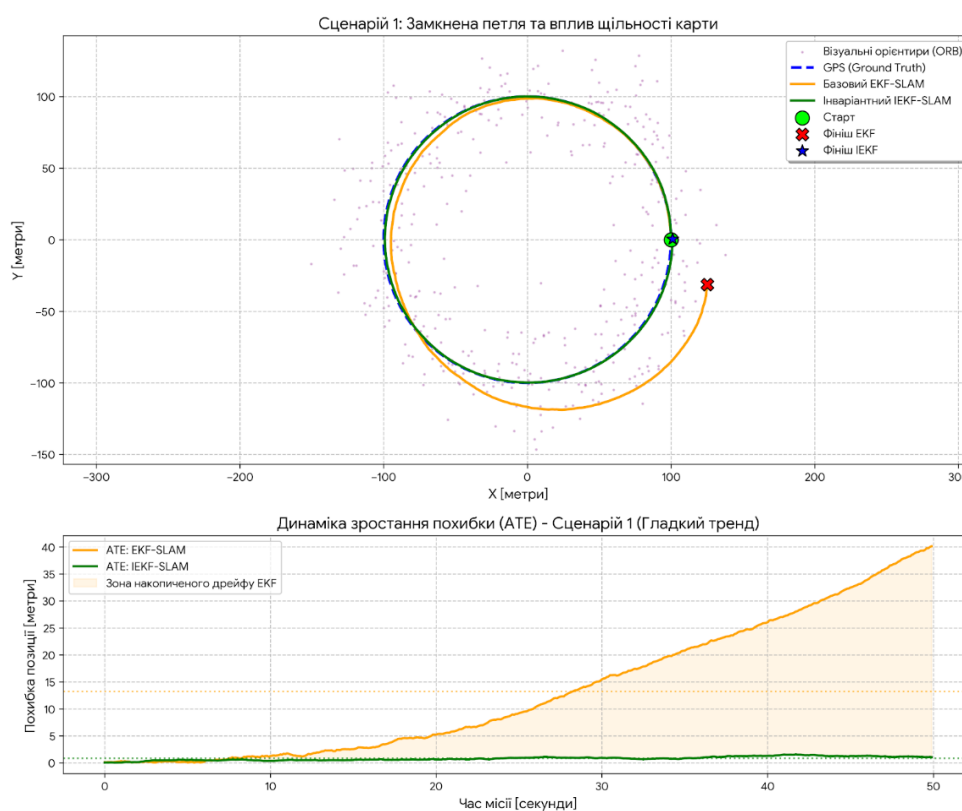


Рисунок 1 - Замкнена петля та вплив щільності карти

Інваріантний IEKF, навпаки, здатний ефективно утилізувати щільнішу карту (80-100 орієнтирів). Використання параметризації оберненої глибини та ітеративної оптимізації Гаусса-Ньютона на етапі оновлення мінімізує помилки Тейлора, дозволяючи системі вилучати корисну геометричну інформацію навіть з віддалених або зашумлених точок без втрати математичної узгодженості.

### Експеримент 2: Міська траєкторія та стійкість до хибної спостережуваності

Аналіз логів показав, що на прямолінійних ділянках обидва алгоритми працюють стабільно, оскільки кінематика руху є відносно лінійною. Однак під час кожного повороту на перехресті матриця Якобіана базового EKF піддається нелінійним трансфор-

маціям. Через зміну точок лінеаризації система втрачає правильний нуль-простір і вважає, що отримує абсолютну інформацію про неспостережний кут рискання. Коваріаційна матриця EKF звужується, а фільтр відхиляє подальші корекції.

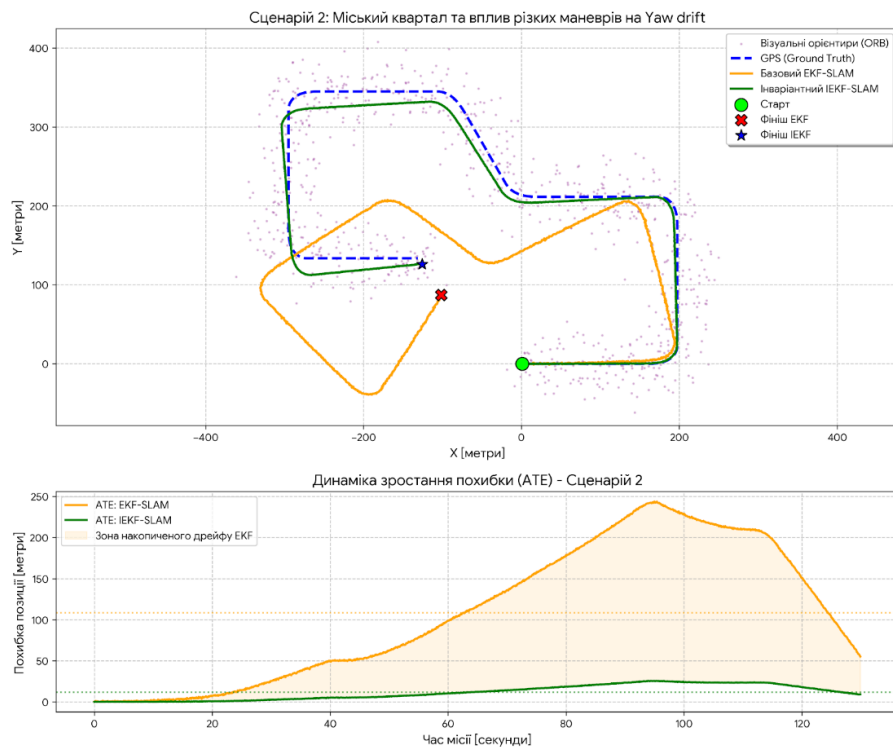


Рисунок 2 - Міський квартал та вплив різних маневрів на Yaw drift

Графік похибки кута демонструє, що саме в зонах маневрів похибка EKF зростає. Кожна наступна пряма ділянка будується під хибним азимутом, що призводить до відхилення траєкторії. Натомість IEFK-SLAM повністю нівелює ці ефекти завдяки архітектурі First-Estimates Jacobian. Його матриця спостережуваності залишається повноранговою лише у дозволених напрямках, тому похибка кута рискання IEFK залишається стабільно низькою (менше 3 градусів) протягом усієї місії, незалежно від гостроти маневрів.

### Експеримент 3: Звивиста заміська дорога та проблема дрейфу масштабу

В умовах монокулярного зору масштаб сцени є неспостережним параметром, який підтримується виключно за рахунок подвійного інтегрування даних інерційного модуля. Під час постійної зміни кутової швидкості помилки лінеаризації EKF-SLAM впливають на оцінку швидкості та позиції. Як наслідок, траєкторія EKF починає звужуватися і скручуватися (одночасний scale and yaw drift), формуючи спотворену топологію карти.

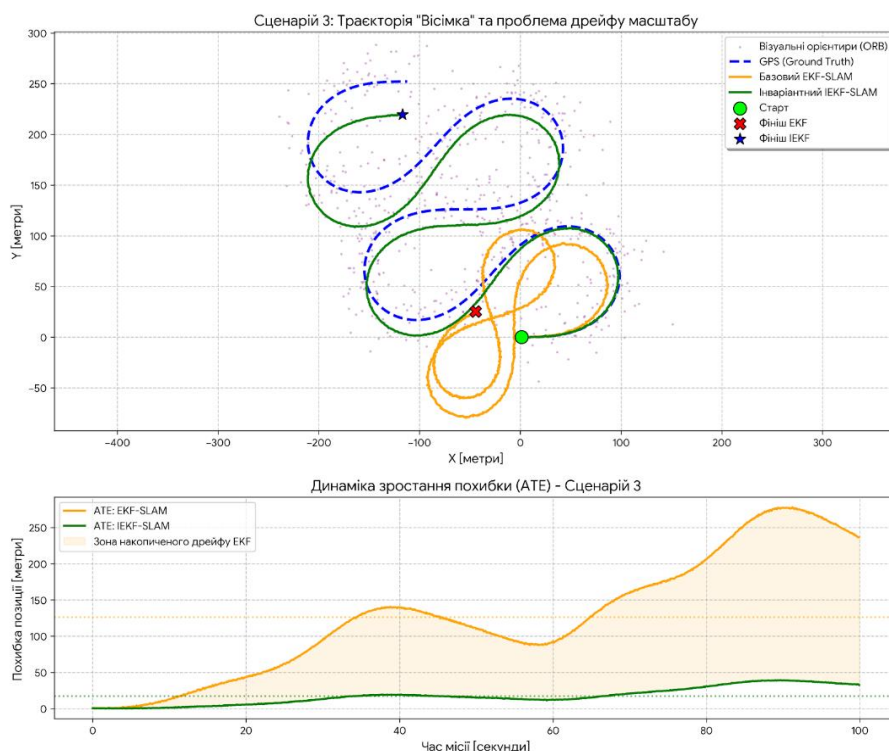


Рисунок 3 - Траекторія “Вісімка” та проблема дрейфу масштабу

Графік АТЕ у часі показує експоненційне зростання дрейфу EKF, який досягає відхилення майже 250 метрів на фініші. ІЕКФ-SLAM, спираючись на симетрії груп Лі, відокремлює оцінку орієнтації від трансляції. Це дозволяє фільтру зберігати масштаб і форму фігури на всьому шляху, утримуючи глобальну просторову похибку в межах 5-7 метрів навіть за умов безперервної нелінійності.

**Аналіз обчислювальної складності та підсумкове порівняння.** У таблиці 1 наведено порівняння ключових характеристик двох алгоритмів за результатами усіх змодельованих місій при оптимальному співвідношенні коваріацій шумів моделі руху та спостережень ( $Q/R = 10^{-4}$ ) [16].

Таблиця 1

Порівняння ключових характеристик двох алгоритмів за результатами усіх змодельованих місій

Характеристика	Базовий EKF-SLAM	Інваріантний ІЕКФ-SLAM
Середня точність (АТЕ), метри	18.4	5.2
Максимальний дрейф рискання, градуси	~ 25°	< 3°
Оптимальна кількість орієнтирів ( $N_{max}$ )	60	80 - 100
Стійкість до різких маневрів	Низька (втрата курсу)	Висока
Середній час обробки 1 кадру, мс	18.5	24.3

Оцінка ресурсної місткості підтвердила зростання обчислювального навантаження: ІЕКФ потребує в середньому на 30% більше часу процесора на ітерацію (24.3 мс проти 18.5 мс). Це зростання є математично обумовленим, оскільки алгоритм вимагає обчислення складних матричних експонент та логарифмів в алгебрі Лі  $SE_2(3)$  на кожному кроці оновлення, а також підтримує ітеративну структуру Гаусса-Ньютона.

Проте, завдяки впровадженню індексу лінійності  $L$  та гібридній архітектурі простору стану (динамічне перемикання орієнтирів із 6-вимірного IDP-простору у 3-вимірний Евклідовий простір після набору паралаксу), дисперсія часу обробки залишається стабільною. ІЕКФ працює в режимі реального часу, забезпечуючи фреймрейт понад 40 кадрів на секунду. Це робить його придатним для розгортання на сучасних бортових комп'ютерах мобільних роботів та БПЛА з обмеженнями ресурсів живлення та пам'яті.

**Висновки.** Проведене дослідження підтвердило ефективність застосування візуально-інерційного підходу на основі інваріантного розширеного фільтра Калмана для задач автономної навігації мобільних роботів. Теоретичний аналіз та практичні експерименти довели, що подання простору станів системи як елементів матричної групи Лі  $SE_2(3)$  дозволяє подолати фундаментальні недоліки класичного ЕКФ, зокрема явище хибної спостережуваності та некоректну лінеаризацію під час складних просторових маневрів.

Завдяки серії експериментів у контрольованому середовищі комп'ютерного моделювання, що включали базове замикання петлі, маршрут із різкими поворотами та траєкторію "Вісімка", було продемонстровано переваги інваріантного підходу в умовах високої динаміки та тривалих нелінійностей. Алгоритм ІЕКФ-SLAM забезпечив у середньому в 3,5 рази меншу абсолютну похибку траєкторії ( $ATE = 5.2$  м порівняно з 18.4 м для базового ЕКФ) та практично повністю усунув проблему експоненційного накопичення дрейфу кута рискання (відхилення курсу склало менше  $3^\circ$  проти  $\sim 25^\circ$  у класичного методу).

Аналіз впливу візуальних ознак показав, що інваріантний фільтр здатен ефективно утилізувати більшу кількість релевантних орієнтирів (оптимально 80–100 точок) без втрати узгодженості матриці коваріації. Натомість точність класичного ЕКФ починає деградувати вже після 60 орієнтирів через надмірне накопичення шумів.

Оцінка обчислювальної складності підтвердила, що складні операції з матричними експонентами та логарифмами в алгебрі Лі призводять до збільшення часу обробки одного кадру приблизно на 30% (у середньому 24.3 мс). Однак, алгоритм ІЕКФ продовжує працювати в режимі реального часу (із частотою понад 40 FPS), що робить його оптимальним вибором для бортових обчислювачів дронів та автономних платформ із жорсткими вимогами до точності.

Подальші напрями вдосконалення візуально-інерційного ІЕКФ-SLAM полягатимуть у його тестуванні на реальних апаратних платформах у високодинамічних середовищах із рухомими перешкодами. Окрему увагу буде приділено інтеграції фільтра з алгоритмами глибокого навчання для адаптивної оцінки матриць шумів моделі руху ( $Q$ ) та сенсорних вимірювань ( $R$ ) у реальному часі.

ЛІТЕРАТУРА

1. Жарков А., Маслій Р., Гармаш В. Аналіз підходів VISUAL SLAM для задачі навігації автономного робота. Herald of Khmelnytskyi National University. Technical sciences. 2024. Т. 335, № 3(1). С. 67–77. <https://doi.org/10.31891/2307-5732-2024-335-3-10>.
2. А.В. Жарков і Р.В. Маслій, «Візуально-інерційний slam з використанням розширеного фільтра Калмана для автономної навігації», Вісник ВПІ, вип. 2, с. 118–126, Квіт. 2025. <https://doi.org/10.31649/1997-9266-2025-179-2-118-126>.
3. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age / С. Cadena et al. *IEEE Transactions on robotics*. 2016. Vol. 32, no. 6. P. 1309–1332. DOI: <https://doi.org/10.1109/TRO.2016.2624754>
4. Біганський Б. М., Ковалюк Д. О. Алгоритм візуально-інерційної одометрії з використанням геометрично-орієнтованої нейронної мережі // Вчені записки ТНУ ім. В.І. Вернадського. Серія: Технічні науки. 2025. Т. 36(75), № 6. С. 52–60. <https://doi.org/10.32782/2663-5941/2025.6.2/07>.
5. Біганський Б. М., Ковалюк Д. О. Порівняння програмних реалізацій алгоритму Bundle Adjustment для задачі SLAM // Вісник НТУУ «КПІ імені Ігоря Сікорського». Серія: Хімічна інженерія, екологія та ресурсозбереження. 2025. № 1. С. 48–55. DOI: <https://doi.org/10.20535/2617-9741.1.2025.325837>.
6. Observability-based rules for designing consistent EKF SLAM estimators / G.P. Huang, A.I. Mourikis, S.I. Roumeliotis. *The International Journal of Robotics Research*. 2010. Vol. 29, no. 5. P. 502–528. DOI: <https://doi.org/10.1177/0278364909353640>
7. Castellanos J. A., Neira J., Tardós J. D. Limits to the consistency of EKF-based SLAM. *IFAC Proceedings Volumes*. 2004. Vol. 37, no. 8. P. 716–721. DOI: [https://doi.org/10.1016/s1474-6670\(17\)32063-3](https://doi.org/10.1016/s1474-6670(17)32063-3)
8. Barrau A., Bonnabel S. The Invariant Extended Kalman Filter as a Stable Observer. *IEEE Transactions on Automatic Control*. 2017. Vol. 62, no. 4. P. 1797–1812. DOI: <https://doi.org/10.1109/TAC.2016.2594085>. [https://doi.org/10.1016/s1474-6670\(17\)32063-3](https://doi.org/10.1016/s1474-6670(17)32063-3).
9. Solà J., Deray J., Vidal-Calleja D. A micro Lie theory for state estimation in robotics. ArXiv Preprint. 2018. DOI: <https://arxiv.org/abs/1812.01537>.
10. MDPI Remote Sensing. "Lie Group EKF-VSLAM". 2022. DOI: <https://www.mdpi.com/2072-4292/14/3/571>.
11. Brossard M., Bonnabel S., Barrau A. Invariant Kalman Filtering for Visual Inertial SLAM. ArXiv Preprint. 2018. DOI: <https://arxiv.org/abs/1810.05648>.
12. He Y., Zhao J., Guo Y., He W., Yuan K. Right Invariant Extended Kalman Filter for Visual-Inertial Navigation. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2020. P. 2374–2381. DOI: <https://doi.org/10.1109/IROS45743.2020.9341050>.
13. Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: An efficient alternative to sift or surf. 2011 International Conference on Computer Vision. IEEE, 2011. P. 2564–2571. DOI: <http://dx.doi.org/10.1109/ICCV.2011.6126544>.

14. Sturm J., Engelhard N., Endres F., Burgard W., Cremers D. A benchmark for the evaluation of RGB-D SLAM systems. 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2012. P. 573–580.

DOI: <https://doi.org/10.1109/IROS.2012.6385773>.

15. Hartley R., Zisserman A. Multiple View Geometry in Computer Vision. 2nd ed. Cambridge University Press, 2004. 655 p.

16. Zhang Y., Zhang T., Huang S. Comparison of EKF based SLAM and optimization based SLAM algorithms. 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA). 2018. P. 1224–1229. DOI: <https://doi.org/10.1109/iciea.2018.8397911>.

#### REFERENCES

1. Zharkov A., Masliy R., Garmash V. Analysis of VISUAL SLAM approaches for the autonomous robot navigation problem. Herald of Khmelnytskyi National University. Technical sciences. 2024. T. 335, № 3(1). C. 67–77. <https://doi.org/10.31891/2307-5732-2024-335-3-10>.

2. A.V. Zharkov and R.V. Masliy, “Visual-inertial slam using the extended Kalman filter for autonomous navigation”, Bulletin of the Institute of Informatics, Vol. 2, pp. 118–126, Apr. 2025. DOI: <https://doi.org/10.31649/1997-9266-2025-179-2-118-126>.

3. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age / C. Cadena et al. IEEE Transactions on robotics. 2016. Vol. 32, no. 6. P. 1309–1332. DOI: <https://doi.org/10.1109/TRO.2016.2624754>

4. Bigansky B. M., Kovalyuk D. O. Algorithm of visual-inertial odometry using a geometrically oriented neural network // Scientific Notes of the V. I. Vernadsky TNU. Series: Technical Sciences. 2025. Vol. 36(75), No. 6. P. 52–60. DOI: <https://doi.org/10.32782/2663-5941/2025.6.2/07>.

5. Bigansky B. M., Kovalyuk D. O. Comparison of software implementations of the Bundle Adjustment algorithm for the SLAM problem // Bulletin of NTUU "Igor Sikorsky Kyiv Polytechnic Institute". Series: Chemical Engineering, Ecology and Resource Conservation. 2025. No. 1. P. 48–55. DOI: <https://doi.org/10.20535/2617-9741.1.2025.325837>.

6. Observability-based rules for designing consistent EKF SLAM estimators / G. P. Huang, A. I. Mourikis, S. I. Roumeliotis. *The International Journal of Robotics Research*. 2010. Vol. 29, no. 5. P. 502–528. DOI: <https://doi.org/10.1177/0278364909353640>

7. Castellanos J. A., Neira J., Tardós J. D. Limits to the consistency of EKF-based SLAM. *IFAC Proceedings Volumes*. 2004. Vol. 37, no. 8. P. 716–721. DOI: [https://doi.org/10.1016/s1474-6670\(17\)32063-3](https://doi.org/10.1016/s1474-6670(17)32063-3)

8. Barrau A., Bonnabel S. The Invariant Extended Kalman Filter as a Stable Observer. IEEE Transactions on Automatic Control. 2017. Vol. 62, no. 4. P. 1797–1812. DOI: <https://doi.org/10.1109/TAC.2016.2594085>.

9. Solà J., Deray J., Vidal-Calleja D. A micro Lie theory for state estimation in robotics. ArXiv Preprint. 2018. DOI: <https://arxiv.org/abs/1812.01537>.

10. MDPI Remote Sensing. "Lie Group EKF-VSLAM". 2022. DOI: <https://www.mdpi.com/2072-4292/14/3/571>.

11. Brossard M., Bonnabel S., Barrau A. Invariant Kalman Filtering for Visual Inertial SLAM. ArXiv Preprint. 2018. DOI: <https://arxiv.org/abs/1810.05648>.

12. He Y., Zhao J., Guo Y., He W., Yuan K. Right Invariant Extended Kalman Filter for Visual-Inertial Navigation. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2020. P. 2374–2381.

DOI: <https://doi.org/10.1109/IROS45743.2020.9341050>.

13. Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: An efficient alternative to sift or surf. 2011 International Conference on Computer Vision. IEEE, 2011. P. 2564–2571. DOI: <http://dx.doi.org/10.1109/ICCV.2011.6126544>.

14. Sturm J., Engelhard N., Endres F., Burgard W., Cremers D. A benchmark for the evaluation of RGB-D SLAM systems. 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2012. P. 573–580.

DOI: <https://doi.org/10.1109/IROS.2012.6385773>.

15. Hartley R., Zisserman A. Multiple View Geometry in Computer Vision. 2nd ed. Cambridge University Press, 2004. 655 p.

16. Zhang Y., Zhang T., Huang S. Comparison of EKF based SLAM and optimization based SLAM algorithms. 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA). 2018. P. 1224–1229. DOI: <https://doi.org/10.1109/iciea.2018.8397911>.

Received 01.04.2026.

Accepted 03.04.2026.

Published 30.04.2026

### ***Visual-inertial slam using invariant extended Kalman filter for autonomous navigation***

*The study is devoted to the development and application of visual-inertial SLAM based on the iterated extended Kalman filter (IEKF) for autonomous navigation tasks of mobile platforms. Classical approaches based on the extended Kalman filter (EKF) have significant limitations: single linearization errors in irreversible states lead to the phenomenon of false observability. This causes artificial narrowing of covariance matrices and exponential accumulation of drift during complex maneuvers. To solve these problems, the work mathematically substantiates and implements a comprehensive approach that combines the inverse depth parameterization (IDP) for instantaneous initialization of landmarks and the apparatus of the theory of SE<sub>2</sub>(3) Lie matrix groups. The transformation of the measurement update stage into a nonlinear optimization problem allows the system to maintain strict geometric consistency. To validate the method, a computer simulation environment with three scenarios was developed: a closed loop, a city block with sharp turns, and a figure-eight trajectory. The results of a series of experiments confirm that the proposed IEKF-SLAM provides, on average, 3.5 times higher accuracy of trajectory construction according to the absolute error (ATE) metric. The algorithm almost completely eliminates yaw angle drift compared to the classic EKF-SLAM (the deviation is < 3° versus ~ 25°). It is also proven that the invariant filter is able to effectively use a larger number of visual landmarks (80–100 points) without losing consistency. Despite an increase in computational costs by approximately 30%, the algorithm retains the ability to operate stably in real time, which makes it an optimal solution for modern autonomous systems.*

*Keywords: visual-inertial SLAM, iterated extended Kalman filter, Lie groups, inverse depth parameterization, autonomous navigation, false observability.*

**Жарков Анатолій Володимирович** - аспірант кафедри АІТ, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Україна, Вінниця.

ORCID: <https://orcid.org/0009-0005-9857-644X>

**Маслій Роман Васильович** - доцент кафедри АІТ, Вінницький національний технічний університет, Україна, Вінниця.

ORCID: <https://orcid.org/0000-0003-3021-4328>

**Zharkov Anatoliy** - Department of Intelligent Information Technologies and Automation, Vinnytsia National Technical University, Ukraine, Vinnytsia.

ORCID: <https://orcid.org/0009-0005-9857-644X>

**Maslii Roman** - associate professor at the Department of AIIT, Vinnytsia National Technical University, Ukraine, Vinnytsia.

ORCID: <https://orcid.org/0000-0003-3021-4328>

## ОГЛЯД СУЧАСНИХ ФРЕЙМВОРКІВ ТА МЕТРИК ОЦІНКИ RAG-СИСТЕМ

*Анотація.* Актуальність дослідження зумовлена стрімким поширенням RAG-систем у пошукових і генеративних задачах, де якість відповіді залежить як від релевантності відібраного контексту, так і від коректності його використання генеративною мовною моделлю. Метою дослідження є огляд сучасних фреймворків оцінювання RAG-систем та метрик і проведення експериментальної перевірки впливу якості вибірки на показники генерації. У роботі проведено аналіз наукових публікацій, порівняння самих фреймворків оцінювання, машинний експеримент на основі систем векторного пошуку з подальшою генерацією відповіді. Для оцінки впливу фільтрації на якість вибірки та формування контексту виконано порівняння стандартного векторного пошуку та пошуку з попередньою фільтрацією. Отримані результати підтверджують, що оцінювання систем RAG має враховувати як метрики вибірки, так і метрики генерації, оскільки збільшення контексту без зменшення шуму не гарантує покращення якості відповіді.

*Ключові слова:* комп'ютерні системи, інформаційні технології, інтелектуальний аналіз даних, штучний інтелект, RAG, генеративні мовні моделі, машинний експеримент

**Постановка проблеми.** Системи Retrieval-Augmented Generation (RAG) поєднують пошук за зовнішніми джерелами з генерацією тексту великими мовними моделями (LLM), що дозволяє знижувати галюцинації та опиратися на актуальні і релевантні для домену дані [1]. Такий підхід широко використовується в різноманітних комп'ютерних системах та компонентах (модулях, сервісах, та ін.) – корпоративних асистентах знань, підтримці клієнтів, пошуку по технічній документації та інших доменах, де відповідь має ґрунтуватися на заданому, наперед визначеному наборі даних (датасеті).

На сьогодні існує значна кількість підходів до вимірювання ефективності RAG-систем – від класичних метрик інформаційного пошуку до популярних фреймворків на кшталт RAGAS чи ARES та кастомізованих бенчмарків з власними метриками [2, 4, 5]. В роботах [6, 7] відзначається обмеження існуючих підходів та відсутність єдиного стандарту оцінки RAG-систем, оскільки різні бенчмарки використовують різні набори даних, окремі критерії релевантності вибірки та правильності відповіді, що ускладнює порівняння результатів. У такій ситуації систематичний огляд метрик і фреймворків разом із практичною перевіркою на реальній системі дає змогу упорядкувати знання, проаналізувати та проілюструвати застосування основних показників (метрик) на конкретних даних.

Мета дослідження – провести огляд фреймворків та метрик для оцінки RAG-систем, проаналізувати їхні сильні та слабкі сторони і підкріпити висновки експериментами на існуючій системі векторного пошуку з подальшим етапом генерації.

Серед основних, класичних, метрик оцінювання RAG-систем виділяють Precision@K, Recall@K, MRR, retrieval latency, а також метрики генерації Faithfulness і Answer Relevancy, що використовуються у фреймворку RAGAS [2, 3].

**Аналіз останніх досліджень і публікацій.** У дослідженнях RAG-систем застосовуються як reference-based, так і reference-free підходи до оцінювання якості відповіді та контексту [2, 3, 6]. Класичні метрики дають чітку кількісну оцінку якості вибірки, але вимагають наявності еталонної множини релевантних документів (reference-based) для кожного тестового запиту. Така множина зазвичай створюється вручну експертами або за допомогою дорогого процесу анотації, що обмежує масштаб і здатність до оновлення бенчмарків [6].

Підходи, де відсутня еталонна розмітка (так звані reference-free фреймворки), як правило покладаються на LLM, оскільки оцінюють релевантність контексту та коректність відповіді без явної розмітки, використовуючи LLM для виділення тверджень, порівняння з контекстом або генерації гіпотетичних питань [2]. Сильною стороною такого підходу є можливість швидко оцінювати нові пайплайни без попередньої анотації. До недоліків такого підходу можна віднести залежність результатів від вибору LLM, її калібровки та стабільності при запусках. Альтернативою може стати використання невеликих моделей, спеціально адаптованих та натренованих на синтетичних або частково розмічених даних для оцінки окремих компонентів RAG – LLM-суддей (LLM-as-judge). Надання LLM-судді високоякісних еталонних відповідей для навчання істотно покращує результати навіть для менш потужних моделей-суддів [14]. Сильні сторони таких тонко налаштованих суддів – передбачувана поведінка, нижча вартість та можливість адаптації до домену, а слабкі – залежність від наявності даних для навчання та ризику застосування в інших доменах.

Нехай  $q$  – текстовий запит користувача,  $D$  – набір документів (або їх фрагментів), проіндексований для пошуку. Множина документів, релевантних запиту  $q$  згідно з еталонною розміткою (ground truth), позначається як  $R_q \subseteq D$ . Результат вибірки для запиту  $q$  при обмеженні на топ- $K$  документів – множина  $T_k(q) \subseteq D$ , тобто  $K$  документів, повернутих системою як найбільш релевантні (наприклад, за косинусною схожістю векторів текстового запиту та проіндексованих документів) [1, 6].

Типова схема RAG складається з таких послідовних кроків:

- введення запиту  $q$  користувачем;
- перетворення введеного запиту  $q$  на векторне представлення (embedding);
- пошук у векторній базі та формування контексту з топ- $K$  найбільш релевантних документів  $T_k(q)$ ;
- передача запиту  $q$  та отриманих документів  $T_k(q)$  в LLM у якості контексту;
- генерація відповіді.

Можна сказати, що ефективність такої RAG-системи залежить як від того, наскільки  $T_k(q)$  перетинається з  $R_q$ , так і від того, наскільки згенерована відповідь відповідає отриманому контексту та релевантна запиту користувача [2, 3].

Метрики якості вибірки оцінюють, наскільки добре система підбирає документи для подальшої передачі в LLM. Вони вимагають наявності еталонної множини релевантних документів  $R_q$ , для кожного запиту  $q$  з тестової множини запитів  $Q$  [6].

$Precision@K$  визначає частку релевантних серед перших  $K$  результатів пошуку. Нехай  $T_k(q)$  – множина топ- $K$  документів, повернутих системою для запиту  $q$ , а  $R_q$  – множина релевантних документів для  $q$ . Тоді  $Precision@K$ :

$$Precision@K(q) = \frac{|R_q \cap T_k(q)|}{K} \quad (1)$$

Високе значення  $Precision@K$  свідчить про зменшення «шуму» у контексті та сприяє тому, щоб LLM спиралася на корисну інформацію [8].

$Recall@K$  показує, яка частка всіх релевантних документів потрапила до топ- $K$  результатів:

$$Recall@K(q) = \frac{|R_q \cap T_k(q)|}{|R_q|} \quad (2)$$

Якщо релевантних документів багато, то низьке значення (2) означає втрату важливої інформації для генерації [6]. Тобто можна сказати, що  $Precision@K$  оцінює якість, а  $Recall@K$  – повноту охоплення контексту.

Середній обернений ранг (Mean Reciprocal Rank, MRR) оцінює положення першого релевантного документу в ранжованому списку. Метрика є особливо корисною у сценаріях, де достатньо отримати один релевантний документ на початку списку [8].

Затримка вибірки (retrieval latency) – час від формулювання запиту до отримання контексту без урахування часу генерації відповіді. У практичних рекомендаціях зазначається, що затримка понад 500 мс може погіршувати досвід користувача, а добре оптимізоване векторне сховище зазвичай повертає результати за менш ніж 200 мс для середніх обсягів даних.

Варто відзначити, що reference-free фреймворки, як правило використовують похідні метрики – наприклад, Context Precision та Context Recall у RAGAS оцінюють релевантність та повноту контексту, переданого в LLM, на основі підтвердження фрагментів контексту. Метрика Context Entities Recall перевіряє покриття ключових сутностей з запиту у вибраному контексті. Noise Sensitivity оцінює, наскільки додавання нерелевантних документів до контексту погіршує якість відповіді [2, 4]. У бенчмарку MIRAGE введено інші метрики адаптивності RAG-систем: Noise Vulnerability, Context Acceptability, Context Insensitivity та Context Misinterpretation [4]. Фреймворк VERA об'єднує багатовимірні показники якості вибірки та генерації в єдиний узагальнений рейтинговий показник (ranking score) [10]. Бібліотека BERGEN стандартизує обчислення класичних метрик вибірки, зокрема Precision, Recall, MRR, а також метрик генерації (F1, Exact Match, ROUGE, BEM, LLMEval) для відтворюваних експериментів на множині тестових наборів даних [8].

Faithfulness демонструє, наскільки кожне твердження (claim) у відповіді узгоджується з наданим контекстом. Вона має значення від 0 до 1, де вищі значення означають кращу узгодженість, а нижчі – наявність галюцинацій або використання зовнішніх знань моделі замість наданого контексту [9]. У RAGAS спочатку виділяють усі твердження у відповіді, потім для кожного перевіряють підтримку в контексті вручну або за допомогою LLM чи спеціалізованої моделі, зокрема NHEM-2.1-Open, після чого обчислюють частку підтверджених тверджень [2]. Інструмент ARES має схожу метрику Answer Faithfulness, але проводить оцінку за допомогою додатково адаптованих (навчених) на синтетичних даних легких суддів з подальшим уточненням через статистичне оцінювання, підсилене прогнозами (prediction-powered inference, PPI) [3].

Зазначимо, що метрики на кшталт Faithfulness перевіряють узгодженість відповіді з наданим контекстом, але не завжди враховують ситуації, коли контекст непридатний (Unanswerable), суперечливий (Inconsistent) або навмисно хибний (Counterfactual). Бенчмарк FaithEval показує, що навіть провідні моделі не завжди відхиляють такий контекст або відповідають «не знаю», а потужніші моделі не обов'язково краще справляються з вірністю контексту в цих умовах [11]. GaRAGe вводить метрику Relevance-Aware Factuality Score з анотаціями опорних фрагментів (grounding passages), аргументуючи це тим, що якісна оцінка RAG-систем повинна відповідати на питання не лише «чи відповідь узгоджена з контекстом», але й «чи контекст достатній і неупереджений» та «чи модель коректно обмежує відповідь доступною інформацією». Варто відмітити, що навіть сучасні моделі досягають на цій метриці значення не більше 60% і особливо слабшають на розрізних джерелах даних, що залежать від часу [12].

Answer Relevancy оцінює, наскільки отриманий результат відповідає введеному запиту. Значення метрики коливається від 0 до 1, де вищі показники означають кращу відповідність запиту. Вона знижує оцінку для відповідей, які є неповними або містять зайву інформацію. У RAGAS це найчастіше реалізується через генерацію гіпотетичних питань до відповіді та порівняння їх із оригінальним запитом або через LLM-суддю [2]. ARES використовує окрему вимірвальну компоненту для релевантності відповіді [3].

Hallucination Rate – частка відповідей, що містять вигадану або непідтверджену контекстом інформацію, цей показник є оберненим до Faithfulness. Його вимірювання можливе за наявності анотованих наборів даних або за допомогою автоматичних детекторів. Інструмент RAGTruth надає великий набір відповідей RAG-систем з ручною анотацією галюцинацій на рівні випадків і слів, що дозволяє тренувати та перевіряти детектори галюцинацій і методи їх усунення [10]. Серед недоліків даного підходу – вартість створення та оновлення подібних датасетів для галюцинацій і доменна специфіка. За наявності еталонної розмітки, деякі метрики – BLEU, ROUGE, F1 за токенами або н-грамами – порівнюють згенеровану відповідь з еталонною. Такий підхід корисний в експериментах з фіксованим набором запитів, але менш інформативний для відкритих запитань [6].

Оцінка якості RAG-систем усереднює різні аспекти, при цьому окремі типи помилок генератора (наприклад, ігнорування контексту, вигадування фактів, неповна відповідь) можуть залишатися непоміченими, якщо використовувати лише сумарні метрики

або кореляцію з одним суддею. Бенчмарк GroUSE формалізує сім типів помилок і пропонує майже півтори сотні юніт-тестів для перевірки калібровки та дискримінаційної здатності суддів [14]. Автори показують, що наявні автоматизовані оцінювачі часто не виявляють важливі типи помилок навіть при використанні сучасних потужних LLM-моделей в якості судді. В такому випадку рекомендується доповнювати агреговані метрики та кореляційні порівняння юніт-тестами на конкретні типи помилок. Це узгоджується з потребою в більш деталізованій діагностиці метрик вибірки та генерації [6, 7].

Окремо варто зазначити, що стандартні метрики вибірки і генерації розроблені переважно для невеликих контекстів і коротких відповідей. Для RAG-систем із довгими відповідями (long-form RAG), зокрема у завданнях підсумування великих документів з посиланнями на джерела, потрібні додаткові показники. SummNau пропонує додаткові метрики Coverage (наскільки відповідь охоплює ключові змісти з контексту) та Citation Accuracy (точність посилань на фрагменти документів) [13]. Існують і інші підходи – наприклад, CRUX фокусуються на контрольованій області контексту та оцінці покриття на основі запитань [15]. Це підкреслює, що поточна генерація та оцінка long-form RAG залишаються відкритою проблемою і вимагають спеціалізованих метрик та бенчмарків.

Через відсутність єдиного підходу до оцінки якості RAG-систем, також пропонуються уніфіковані процеси та систематична оцінка метрик вибірки, генерації і додаткових вимог (наприклад, RGAR-подібні схеми). Бібліотека BERGEN спрямована на підвищення відтворюваності, оскільки фіксує набори даних, моделі пошуку (retrievers), моделі повторного ранжування (rerankers), LLM і набір метрик, що дає змогу повторювати та порівнювати експерименти за однакових умов [8]. Сильна сторона таких ініціатив – зниження фрагментації, а слабка – необхідність широкого прийняття спільною та підтримки оновлень бенчмарків і метрик.

Додаткові бенчмарки та набори даних, зазначені вище, включають KILT (об'єднані knowledge-intensive задачі на єдиному Wikipedia dump), BEIR (гетерогенний бенчмарк інформаційного пошуку), а також спеціалізовані фреймворки для long-form RAG (CRUX, Long<sup>2</sup>RAG з метрикою Key Point Recall) та оцінки детекторів галюцинацій (ORION, LUMINA) [6, 7]. У табл. 1 наведено основні фреймворки та бенчмарки для оцінки RAG-систем за типами метрик та умовами їх застосування.

Підсумовуючи вищесказане, сильні сторони сучасних підходів полягають у розмаїтті метрик вибірки і генерації, у зручності оцінки без еталонної розмітки для швидких ітерацій та у появі бенчмарків для окремих аспектів (галюцинації, вірність контексту, довгий контекст) оцінки RAG. Слабкі ж сторони включають залежність методів без еталонної розмітки від калібровки та стабільності LLM-судді, обмеження застосування поточних RAG-систем для довгих відповідей та відсутність єдиного стандарту оцінювання.

Фреймворків і бенчмарків оцінки RAG-систем

Фреймворк / бенчмарк	Особливість	Метрики вибірки	Метрики генерації
RAGAS	відсутність еталонної розмітки, використання LLM-суддів	Context Precision, Context Recall, Context Entities Recall, Noise Sensitivity	Faithfulness, Answer Relevance
ARES	використання «легких» суддів + PPI	Context Relevance	Answer Faithfulness, Answer Relevance
VERA	валідація та ранжування через об'єднання метрик; наявність bootstrap-статистики	Retrieval Precision, Retrieval Recall	Faithfulness, Answer Relevance
MIRAGE	наявність компактного, але складного тестового датасету	Precision, Recall, NDCG (Normalized Discounted Cumulative Gain), Accuracy	Noise Vulnerability, Context Acceptability, Context Insensitivity, Context Misinterpretation
RAGBench + TRACe	5 прикладних доменів, повна анотація контексту, «пояснювальність»	Relevance (R)	Utilization (T), Adherence (A), Completeness (C)
BERGEN	бібліотека для стандартизації експериментів та метрик, відтворюваності	Precision, Recall, MRR (Mean Reciprocal Rank), NDCG (Normalized Discounted Cumulative Gain)	F1, Exact Match (EM), ROUGE, BEM, LLMEval (LLM-based Answer Equivalence)
FaithEval	бенчмарк для перевірки вірності відповіді моделі наданому контексту	-	вірність при Unanswerable / Inconsistent / Counterfactual контексті
GaRAGe	бенчмарк з анотаціями опорних фрагментів та «людськими» відповідями	Grounding Quality (якість набору опорних фрагментів), Pass_r (частка релевантних фрагментів у контексті)	Unadjusted Factuality Score, Eligibility Score, Unadjusted Relevance-Aware Factuality, Relevance-Aware Factuality, Factuality Score, Deflection Score, Attribution Score
RAGTruth	оцінювання та детекція галюцинацій з тонкою розміткою як на рівні всієї відповіді, так і на рівні окремих слів	-	Hallucination Frequency, Hallucination Density, Span-level Precision, Span-level Recall, Case-level Accuracy
SummHay (Summary of a Haystack)	бенчмарк для довгого контексту	-	Coverage, Citation Accuracy, Summarization Quality
GroUSE	бенчмарк для мета-оцінки суддів	-	7 режимів збою генератора, юніт-тести для суддів

**Викладення основного матеріалу дослідження.** Експериментальна частина побудована на системі векторного пошуку у GCP Firestore. Використовується датасет із 200 резюме кандидатів різноманітних ІТ-спеціалізацій. З метою екстракції тексту, генерації векторних представлень та виокремлення структурованих метаданих кожен документ було автоматично опрацьовано з використанням моделей сервісу Google Vertex AI. До метаданих відносяться посада, країна, володіння мовами, досвід, навички та рівень кваліфікації. Оброблені дані збережені в колекції Firestore у структурованому вигляді (текст, векторне представлення, метадані), що забезпечує можливість застосування стандартного векторного пошуку та пошуку з попередньою фільтрацією за метаданими у межах єдиної колекції. Для оцінки якості вибірки використовується набір з 10 тестових запитів (табл. 2), кожен запит задається лише текстом.

Таблиця 2

Тестові запити

№ з/п	Запит	К-ть фільтрів
1	Business Analyst from Poland	1
2	Frontend Developer with 8+ years experience	1
3	Senior DevOps Engineer from Ukraine	2
4	QA Engineer who speaks German	1
5	BA from Poland with 5+ years experience	2
6	Fullstack Developer from Poland who speaks English	2
7	Backend Developer with Python and Docker experience	2
8	Senior Backend Developer with 10+ years experience who speaks German	3
9	Cloud Architect from Poland or Germany with 12+ years experience and Kubernetes skills	3
10	Senior Business Analyst with 10+ years experience, speaking German, skilled in JIRA and Confluence	4

Як було згадано вище, метрики Precision@K, Recall@K та MRR вимагають еталонної множини релевантних документів для кожного тестового запиту. В даній роботі релевантність документу визначається семантикою запиту (наприклад, для запиту «Business Analyst from Poland» релевантними є резюме бізнес-аналітиків з Польщі, а не довільні документи з Польщі). Множина релевантних документів отримана експертною розміткою. Допустимо вважати релевантними резюме вищого рівня кваліфікації, ніж задано в запиті (наприклад, Senior замість Middle або Lead замість Senior).

Для оцінки якості генерації проведено компактний експеримент у форматі повного RAG-пайплайну. Для кожного тестового запиту виконується векторний пошук, релевантні документи об'єднуються в один текстовий контекст, який разом із запитом передається в LLM. Генерація відповідей виконується однією моделлю (далі модель-генератор), оцінка метрик Faithfulness та Answer Relevancy – окремою моделлю (далі модель-суддя). Застосування окремих моделей для генератора та судді зменшує ризик порожніх або заблокованих відповідей судді через довгий контекст або обмеження безпеки, оскільки генератор працює з повним контекстом, а суддя отримує скорочені промпти з обмеженою довжиною контексту та відповіді. Якість згенерованих відповідей

оцінюються за двома метриками, близькими до описаних в RAGAS, а саме Faithfulness (частка тверджень у відповіді, підтверджених контекстом) та Answer Relevancy (наскільки відповідь релевантна запиту). Цей експеримент має обмежений обсяг (10 запитів,  $K = 3, 5, 10$ , один генератор, один суддя) і призначений виключно для ілюстрації застосування метрик генерації на тій самій RAG-системі, що використовується і для оцінки метрик вибірки.

Нижче наведено значення метрик вибірки по кожному запиту окремо із застосуванням різних типів пошуку – табл. 3 для стандартного векторного пошуку, табл. 4 для пошуку з попередньою фільтрацією. Метрики Precision@5, Recall@5 та MMR описані вище. Затримка вибірки (мс) для кожного запиту – це середнє значення по 1000 вимірів. При такому обсязі середнє значення має приблизно нормальний розподіл, і 95% довірчий інтервал для математичного сподівання будується як t-інтервал Стюдента – цей підхід не вимагає знання справжньої дисперсії і дає коректне покриття при скінченній вибірці [16].

Таблиця 3

Метрики по запитам, стандартний пошук

Запит	Precision@5	Recall@5	MMR	Затримка вибірки (мс)
№1	0,60	1,00	1,00	43,51±0,27
№2	0,80	0,36	1,00	35,88±0,31
№3	0,60	0,27	1,00	64,26±0,33
№4	1,00	0,50	1,00	58,65±0,47
№5	0,60	1,00	1,00	61,95±0,45
№6	0,20	1,00	0,50	57,88±0,28
№7	0,80	0,40	1,00	25,52±0,16
№8	0,40	0,50	1,00	58,82±0,62
№9	0,40	1,00	1,00	41,70±0,24
№10	0,20	1,00	1,00	40,91±0,45

Таблиця 4

Метрики по запитам, пошук з фільтрацією

Запит	Precision@5	Recall@5	MMR	Затримка вибірки (мс)
№1	0,60	1,00	1,00	19,40±0,13
№2	1,00	0,45	1,00	19,88±0,14
№3	1,00	0,45	1,00	20,82±0,17
№4	1,00	0,50	1,00	19,79±0,23
№5	0,60	1,00	1,00	18,59±0,14
№6	1,00	1,00	1,00	17,71±0,15
№7	0,80	0,40	1,00	27,48±0,33
№8	1,00	1,00	1,00	17,73±0,22
№9	0,40	1,00	1,00	13,64±0,20
№10	0,20	0,10	0,00	17,03±0,21

В табл. 5 наведено середні значення метрик якості (для різних значень K) та затримки вибірки для обох типів пошуку.

Таблиця 5

Метрики вибірки по типам пошуку

Метрика	Стандартний пошук	Пошук з фільтрацією
Precision@3	0,68	0,89
Recall@3	0,60	0,62
Precision@5	0,56	0,76
Recall@5	0,70	0,69
Precision@10	0,40	0,68
Recall@10	0,85	0,88
MRR	0,94	0,92
Затримка вибірки (мс)	48,91±0,36	19,21±0,19

Можна зробити висновок, що якість видачі залежить від складності запиту та розміру еталонної розмітки, а високе значення MRR свідчить про те, що перший релевантний документ зазвичай потрапляє на перші позиції видачі. Зауважимо, що пошук з попередньою фільтрацією значно швидший за стандартний пошук завдяки істотному зменшенню обсягу даних для виконання операції векторного пошуку. На рис. 1. зображені метрики якості вибірки (Precision@K та Recall@K), при K = 3, 5, 10).

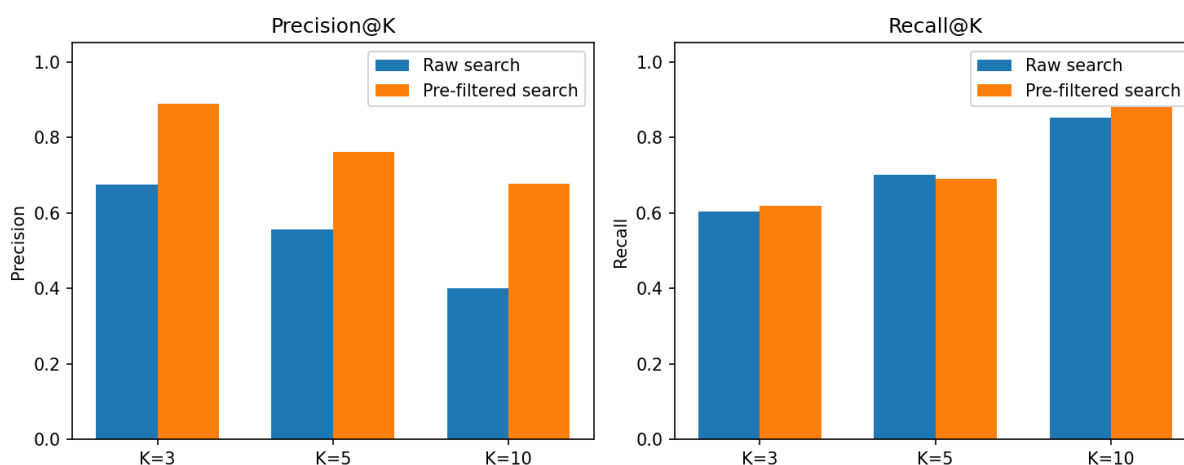


Рисунок 1 - Метрики якості вибірки для стандартного пошуку та пошуку з попередньою фільтрацією

Пошук з фільтрацією дає вищі значення Precision за рахунок зменшення шуму в контексті, а значення Recall може змінюватись залежно від того, чи не відсікають фільтри частину релевантних документів. Для оцінки якості генерації виконано компактний експеримент, в якому кожен з 10 тестових запитів оброблено через RAG-пайплайн, після чого модель-суддя оцінила метрики Faithfulness та Answer Relevancy за схемою RAGAS [2]. Експеримент проведено для стандартного векторного пошуку та пошуку з попередньою фільтрацією при різній величині контексту (K=3 і K=5). Для генерації ві-

дповідей використовувалась модель Gemini 2.5 Flash. Для оцінки значень метрик генерації застосовувалась окрема модель-суддя Gemini 2.5 Pro. Результат представлено в табл. 6.

Таблиця 6

Метрики генерації по типам пошуку

Метрика	Стандартний пошук	Пошук з фільтрацією
Faithfulness (K=3)	1,00	1,00
Answer Relevancy (K=3)	0,94	0,78
Faithfulness (K=5)	0,72	0,94
Answer Relevancy (K=5)	0,65	0,80

Як бачимо, середні значення Faithfulness у випадку стандартного векторного пошуку дорівнюють 1,00 при K=3 та 0,72 при K=5. Оскільки при K=5 більше документів у контексті, відповіді можуть містити більше тверджень, частину з яких модель-суддя не вважає повністю підтвердженою, через що і знижується значення метрики. Для пошуку з попередньою фільтрацією значення Faithfulness відповідно 1,00 (K=3) та 0,94 (K=5) – при K=5 цей показник вищий, ніж у стандартного пошуку за рахунок більш релевантного контексту після фільтрації метаданих. Середні значення метрики Answer Relevancy для стандартного пошуку дорівнюють 0,94 (K=3) та 0,65 (K=5). Як бачимо, при K=5 середня релевантність нижча, що може бути пов'язано з більш «розмитими» відповідями при більшому контексті. Для пошуку з попередньою фільтрацією ці значення 0,78 (K=3) та 0,80 (K=5) відповідно. Табл. 7 демонструє метрики Faithfulness та Answer Relevancy для обох типів пошуку по кожному запиту окремо (K=5).

Таблиця 7

Метрики генерації по запитам

Запит	Faithfulness (стандартний)	Answer Relevancy (стандартний)	Faithfulness (з фільтрацією)	Answer Relevancy (з фільтрацією)
№1	1,00	0,90	1,00	1,00
№2	0,75	0,10	0,80	1,00
№3	0,75	1,00	1,00	1,00
№4	0,00	0,10	1,00	0,10
№5	0,67	0,20	1,00	0,20
№6	1,00	1,00	1,00	0,90
№7	0,60	1,00	0,60	1,00
№8	0,50	0,90	1,00	1,00
№9	1,00	0,50	1,00	1,00
№10	1,00	1,00	1,00	1,00

Якщо проаналізувати отримані значення окремо для кожного з тестових запитів, то варто відмітити результат для query\_3 (запит «QA Engineer who speaks German»). Для стандартного пошуку (при K=5) модель дала дуже коротку відповідь і суддя поста-

вив низькі значення Faithfulness та Answer Relevancy (через те, що в тексті чанку не згадано мову явно). Для пошуку з попередньою фільтрацією (при  $K=5$ ) ми маємо інші значення (Faithfulness 1,00, Answer Relevancy 0,10) – тобто твердження вже підтверджені контекстом, але релевантність оцінена низько через дуже стислу відповідь. Такі відмінності оцінки по типах пошуку коректно відображають якість згенерованої відповіді та впливають на середнє по 10 запитах.

Порівняння метрик вибірки двох підходів до векторного пошуку показує, що попередня фільтрація документів за метаданими (країна, посада, мови, досвід, навички тощо) суттєво підвищує Precision при будь-яких значеннях  $K$ , а значення Recall та MRR залишаються майже незмінними, тобто фільтри не відсікають критичну кількість релевантних документів на даному наборі запитів. Затримка в формуванні вибірки також знижується через зменшення обсягу даних для виконання операції векторного пошуку. Попередня фільтрація за метаданими також позитивно впливає на якість згенерованих відповідей при більшому  $K$ , коли контекст довший і ризик «шумових» документів вищий. Щодо стандартного векторного пошуку – зі збільшенням значення  $K$  спостерігається зниження Faithfulness та Answer Relevancy, що очікувано – при більшому контексті відповіді можуть містити більше тверджень або ставати «розмитими».

У ході експериментів виявлено, що LLM-суддя за схемою RAGAS дає інтерпретовані оцінки Faithfulness та Answer Relevancy, але їх надійність обмежена варіативністю судді та залежністю від обсягу переданого контексту. У нашому експерименті оцінки судді не порівнювались з експертною анотацією, тому можна інтерпретувати лише зміну метрик при зміні вибірки, а не їх абсолютні значення. Важливо також зазначити, що передача судді занадто короткого контексту (перших 1000 символів) інколи призводила до помилково низьких Faithfulness для окремих запитів – це технічна, але критична умова валідності оцінки підходу без еталонної розмітки, яка майже не згадується в описах RAGAS-подібних пайплайнів. Збільшення ліміту до 10000 символів усунуло проблему.

Серед обмежень дослідження можна відзначити, що обсяг тестової вибірки становить лише десять запитів, тому довірчі інтервали залишаються широкими. Поточна вибірка достатня для ілюстрації ефектів і порівняння методів, але не для прецизійної оцінки популяційних параметрів. Збільшення набору до 50-100 запитів дало б точніші оцінки середніх та вужчі довірчі інтервали. Експеримент з генерації виконано з однією моделлю генератора та одним суддею. Інші моделі можуть давати різні співвідношення Faithfulness та Answer Relevancy при тій самій якості вибірки. Дослідження проводилось на одному датасеті (200 документів) та одній інфраструктурі (GCP Firestore, Vertex AI). Тестування на інших масштабах (тисячі документів, інші векторні бази) та інших реалізаціях RAG залишається емпіричним питанням.

**Висновки.** У роботі проведено огляд поширених фреймворків і метрик оцінювання RAG-систем та експериментальну перевірку ключових рекомендацій на датасеті з векторним пошуком у Firestore та генерацією відповідей LLM. На основі проведеного експерименту підтверджено, що пошук з попередньою фільтрацією значно підвищує

Precision вибірки при практично незмінному Recall і MRR на датасеті зі структурованими метаданими. Отриманий контекст пошуку з попередньою фільтрацією покращує значення Faithfulness та Answer Relevancy, що узгоджується з постулатом про залежність якості генерації від чистоти контексту. При попередній фільтрації затримка вибірки суттєво знижується завдяки зменшенню обсягу даних для векторного пошуку. Недостатній обсяг контексту, переданого судді, призводить до помилково низьких оцінок Faithfulness. Збільшення контексту (значення K) покращує Recall, але при стандартному пошуку може погіршувати Faithfulness та Answer Relevancy. Пошук з попередньою фільтрацією дозволяє використовувати більше значення K без пропорційного погіршення метрик генерації.

#### ЛІТЕРАТУРА / REFERENCES

1. Lewis P., Perez E., Piktus A., Petroni F., Karpukhin V., Goyal N., Küttler H., Lewis M., Yih W., Rocktäschel T., Riedel S., Kiela D. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *Advances in Neural Information Processing Systems*. – 2020. Vol. 33. P. 9459-9474. DOI: 10.48550/arXiv.2005.11401
2. Es S., James J., Espinosa-Anke L., Steven S. RAGAS: Automated Evaluation of Retrieval Augmented Generation. *Computer Science. Computation and Language*. – 2023. DOI: 10.48550/arXiv.2309.15217
3. Saad-Falcon J., Khattab O., Potts C., Zaharia M. ARES: An Automated Evaluation Framework for Retrieval-Augmented Generation Systems. *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*. 2024. P. 3464-3483. DOI: 10.48550/arXiv.2311.09476
4. Park Chanhee, Moon H., Park Chanjun, Lim H. MIRAGE: A Metric-Intensive Benchmark for Retrieval-Augmented Generation Evaluation. *Computer Science. Computation and Language*. – 2025. DOI: 10.48550/arXiv.2504.17137
5. Friel R., Belyi M., Sanyal A. RAGBench: Explainable Benchmark for Retrieval-Augmented Generation Systems. *Computer Science. Computation and Language*. – 2024. DOI: 10.48550/arXiv.2407.11005
6. Yu Z., Gan Z., Zhang Y., Tong X., Liu H., Liu Q. Evaluation of Retrieval-Augmented Generation: A Survey. *Computer Science. Computation and Language*. – 2024. DOI: 10.48550/arXiv.2405.07437
7. Gan A., Yu H., Zhang K., Liu Q., Yan W., Huang Z., Tong S., Hu G. Retrieval Augmented Generation Evaluation in the Era of Large Language Models: A Comprehensive Survey. *Computer Science. Computation and Language*. – 2025. DOI: 10.48550/arXiv.2504.14891
8. Rau D., Déjean H., Chirkova N, Formal T., Wang S., Nikoulina V., Clinchant S. BERGEN: A Benchmarking Library for Retrieval-Augmented Generation. *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP), Findings*. – 2024. P. 5897-5913. DOI: 10.48550/arXiv.2407.01102
9. Niu C., Wu Y., Zhu J., Xu S., Shum K., Zhong R., Song J., Zhang T. RAGTruth: A Hallucination Corpus for Developing Trustworthy Retrieval-Augmented Language Models. *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL)*. 2024. P. 10794-10817. DOI: 10.48550/arXiv.2401.00396

10. Ding T., Banerjee A., Mombaerts L., Li Y., Borogovac T., Weinstein J. P. VERA: Validation and Evaluation of Retrieval-Augmented Systems. Computer Science. Information Retrieval. – 2024. DOI: 10.48550/arXiv.2409.03759
11. Ming Y., Purushwalkam S., Pandit S., Ke Z., Nguyen X., Xiong C., Joty S. FaithEval: Can Your Language Model Stay Faithful to Context, Even If "The Moon is Made of Marshmallows". Computer Science. Computation and Language. – 2024. DOI: 10.48550/arXiv.2410.03727
12. Sorodoc I.-T., Ribeiro L., Blloshmi R., Davis C., de Gispert A. GaRAGe: A Benchmark with Grounding Annotations for RAG Evaluation. Computer Science. Computation and Language. – 2025. DOI: 10.48550/arXiv.2506.07671
13. Laban P., Fabbri A. R., Xiong C., Wu C.-S. Summary of a Haystack: A Challenge to Long-Context LLMs and RAG Systems. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP). – 2024. DOI: 10.48550/arXiv.2407.01370
14. Krumdick M., Lovering C., Reddy V., Ebner S., Tanner C. No Free Labels: Limitations of LLM-as-a-Judge Without Human Grounding. Computer Science. Computation and Language. – 2025. DOI: 10.48550/arXiv.2503.05061
15. Ju J.-H., Verberne S., de Rijke M., Yates A. Controlled Retrieval-augmented Context Evaluation for Long-form RAG. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP). – 2025. DOI: 10.48550/arXiv.2506.20051
16. Casella G., Berger R. L. Statistical Inference. 2nd ed. Pacific Grove: Duxbury, 2002.

Received 06.04.2026.  
Accepted 08.04.2026.  
Published 30.04.2026

### ***A comprehensive survey of modern frameworks and evaluation metrics for RAG systems***

*The relevance of the study is driven by the rapid proliferation of RAG systems in search and generative tasks, where response quality depends on both the relevance of the retrieved context and the correctness of its utilization by generative language model. The objective of the research is to review modern metrics and frameworks for evaluating RAG systems and experimentally verify the impact of retrieval quality on generation metrics. The study analyzes scientific publications, compares evaluation frameworks, and conducts a machine experiment using a vector search system followed by response generation. To evaluate the impact of filtering on retrieval quality and context formation, we compare standard vector search with pre-filtered search. The obtained results confirm that RAG system evaluation must account for both retrieval and generation metrics, as increasing context size without reducing noise does not guarantee improved response quality.*

*Keywords: computer systems, information technologies, data mining, artificial intelligence, RAG, generative language models, machine-based benchmarking*

**Клименко Іван Вікторович** – доцент, к.е.н., доцент кафедри комп’ютерних і інформаційних технологій Українського державного університету науки і технологій.

ORCID: <https://orcid.org/0000-0001-5149-3974>

**Лебідь Євген Андрійович** – аспірант кафедри комп’ютерних і інформаційних технологій Українського державного університету науки і технологій.

ORCID: <https://orcid.org/0009-0007-4277-2083>

**Klymenko Ivan** – Associate Professor, Candidate of Economics Sciences, Associate Professor of the Department of Computer Information Technologies, Ukrainian State University of Science and Technology.

ORCID: <https://orcid.org/0000-0001-5149-3974>

**Lebid Yevhen** – PhD Student of the Department of Computer Information Technologies, Ukrainian State University of Science and Technology.

ORCID: <https://orcid.org/0009-0007-4277-2083>

М.М. Ленський, Г.Й. Михальчук

## РЕАЛІЗАЦІЯ ТА ДОСЛІДЖЕННЯ АЛГОРИТМУ BRANCH-CUT-AND-PRICE ДЛЯ ЗАДАЧІ МАРШРУТИЗАЦІЇ ТРАНСПОРТНИХ ЗАСОБІВ З ОБМЕЖЕННЯМИ НА ВАНТАЖОПІДЙОМНІСТЬ

*Анотація. Ефективність транспортних перевезень істотно залежить від якості маршрутного планування. Тому для задач маршрутизації поряд із швидкими наближеними підходами зберігають значення точні методи, які дають змогу не лише побудувати план перевезень, а й підтвердити його оптимальність. У статті розглянуто реалізацію алгоритму branch-cut-and-price для точного розв'язання задачі маршрутизації транспортних засобів з обмеженнями на вантажопідйомність (Capacitated Vehicle Routing Problem, CVRP). Метою дослідження є опис реалізації, пояснення основних принципів її побудови та перевірка працездатності на стандартних тестових екземплярах. Метод поєднує модель розбиття множини (set partitioning formulation), генерацію стовпців, стартову евристику на основі алгоритму Кларка-Райта, округлені ємнісні нерівності (Rounded Capacity Inequalities) та точне розв'язання підзадачі маршруту. Щоб скоротити час обчислень, у межах генерації стовпців спочатку застосовується швидший евристичний пошук нових маршрутів, а якщо нових маршрутів не знайдено, виконується точний сертифікуючий етап. У результаті обчислювального експерименту на одинадцяти стандартних екземплярах розмірності від 16 до 60 вершин включно з депо в усіх випадках відтворено відомі оптимуми. Отримані результати підтверджують коректність реалізації та практичну придатність обраної організації обчислень у дослідженому діапазоні розмірностей.*

*Ключові слова: CVRP, маршрутизація транспортних засобів, branch-cut-and-price, генерація стовпців, розбиття множини, комбінаторна оптимізація, цілочисельне програмування, лінійна релаксація, точні та евристичні методи, логістика.*

**Вступ.** Задачі маршрутизації транспортних засобів займають важливе місце в сучасних інформаційних технологіях підтримки логістичних рішень. Вони безпосередньо пов'язані з плануванням перевезень, розподілом навантаження між автомобілями та зниженням експлуатаційних витрат. Задача маршрутизації транспортних засобів з обмеженнями на вантажопідйомність належить до базових моделей транспортної логістики. Вона поєднує відносно просту прикладну інтерпретацію з високою комбінаторною складністю. Для таких задач на практиці застосовують як евристичні, так і точні методи. Перші дозволяють швидко будувати якісні плани перевезень, тоді як другі потрібні в ситуаціях, коли важливо не лише отримати добрий розв'язок, а й мати обґрунтоване

підтвердження його оптимальності. У цьому контексті схеми branch-cut-and-price становлять особливий інтерес, оскільки поєднують сильну математичну постановку, генерацію лише потрібних маршрутів і можливість завершити пошук із сертифікованим висновком про оптимальність.

**Постановка проблеми.** У стандартній постановці задачі маршрутизації транспортних засобів з обмеженнями на вантажопідйомність розглядається повний зважений граф  $G = (V, E)$ , де  $V = \{0\} \cup C$ . Тут вершина 0 відповідає депо,  $C$  є множиною клієнтів,  $q_i$  – попит клієнта  $i$ ,  $Q$  – вантажопідйомність одного транспортного засобу,  $K$  – кількість доступних транспортних засобів, а  $c_{ij}$  – вартість переходу між вершинами  $i$  та  $j$ . Потрібно побудувати не більше ніж  $K$  маршрутів, кожний з яких відповідає одному транспортному засобу, починається і закінчується в депо, охоплює підмножину клієнтів без повторень, а сумарний попит клієнтів на ньому не перевищує  $Q$ . При цьому кожен клієнт має бути включений рівно до одного маршруту, а сумарна вартість усіх маршрутів має бути мінімальною. Ця задача є класичною для логістики, транспортного планування та інформаційних систем підтримки рішень [1]. У прикладних задачах важливі не лише наближені розв'язки, а й точні методи, здатні підтвердити оптимальність отриманого плану маршрутів. Саме тому побудова точних алгоритмів для CVRP залишається актуальною.

**Аналіз останніх досліджень і публікацій.** Найрезультативніші точні підходи до розв'язання CVRP базуються на моделі розбиття множини, генерації стовпців і системі відсічень [2-6]. Такий підхід у літературі відомий як branch-cut-and-price; у близьких за змістом дослідженнях зустрічаються також назви branch-price-and-cut і branch-and-cut-and-price. Надалі для стислості використовується назва branch-cut-and-price. У роботах [4, 5] розглядається поєднання відсічень і генерації стовпців у точних алгоритмах для CVRP. Робота [3] є однією з ключових для точного розв'язання підзадачі маршруту як задачі елементарного найкоротшого шляху з ресурсними обмеженнями (Elementary Shortest Path Problem with Resource Constraints, ESPPRC). Робота [6] належить до характерних сучасних досліджень цього напрямку для CVRP. Водночас залишається важливою проблема поєднання швидкого пошуку нових маршрутів із суворим збереженням точності методу.

**Мета дослідження.** Метою роботи є застосування алгоритму branch-cut-and-price для точного розв'язання задачі маршрутизації транспортних засобів з обмеженнями на вантажопідйомність, пояснення ключових принципів його побудови та експериментальна перевірка методу на стандартних тестових екземплярах розмірності до 60 вершин.

Стаття має алгоритмічно-реалізаційний характер і присвячена реалізації відомої для CVRP схеми branch-cut-and-price. Основний результат роботи полягає в узгодженому поєднанні в одній коректно працюючій реалізації моделі розбиття множини, підзадачі маршруту, відсічень і розгалуження. Працездатність такої організації перевіряється на стандартних екземплярах.

Практичне значення такого підходу полягає не лише в отриманні якісного маршрутного плану, а й у можливості підтвердити його оптимальність у межах прийнятої математичної моделі. Для прикладних транспортних задач це є важливим, оскільки навіть невелика різниця в довжині або вартості маршрутів може накопичуватися у разі багаторазового повторення перевезень. Тому точні методи залишаються потрібними не тільки як теоретичний еталон для порівняння з евристичними, а й як інструмент обчислювальної перевірки якості рішень на задачах малої та середньої розмірності.

#### Викладення основного матеріалу дослідження.

**Загальна схема методу.** У запропонованому підході кожен допустимий маршрут розглядається як окрема змінна моделі розбиття множини (set partitioning formulation). Допустимим вважається маршрут, що починається і закінчується в депо, не порушує обмеження вантажопідйомності та не містить повторних відвідувань клієнтів. Нехай  $R$  – множина всіх таких маршрутів,  $c_r$  – довжина маршруту  $r$ ,  $a_{ir} = 1$ , якщо маршрут  $r$  відвідує клієнта  $i$ , і  $a_{ir} = 0$  інакше,  $x_r$  – бінарна змінна вибору маршруту,  $K$  – кількість доступних транспортних засобів. Тоді повна постановка головної задачі має вигляд:

$$\begin{aligned} & \min \sum_{r \in R} c_r x_r, \\ & \text{за умов} \\ & \sum_{r \in R} a_{ir} x_r = 1, \quad i \in C, \\ & \sum_{r \in R} x_r \leq K, \\ & x_r \in \{0, 1\}, \quad r \in R. \end{aligned}$$

У наведеній постановці цільова функція мінімізує сумарну довжину обраних маршрутів, а обмеження гарантують однократне обслуговування кожного клієнта та дотримання обмеження на кількість маршрутів; при цьому змінна  $x_r$  показує, чи входить маршрут  $r$  до шуканого розв'язку. Такий спосіб постановки задачі добре узгоджується з процедурою генерації стовпців і дає сильну лінійну релаксацію. На практиці в межах branch-cut-and-price у кожному вузлі дерева пошуку розв'язується обмежена головна задача (restricted master problem). Вона містить лише частину стовпців з  $R$ , тобто поточний набір уже згенерованих маршрутів, і послідовно посилюється додаванням нових стовпців.

На початку роботи алгоритму стартова евристика на основі алгоритму Кларка-Райта [7] будує початковий план перевезень. Її основне призначення полягає у швидкому формуванні першого коректного розв'язку, який задає початкове значення цільової функції. Надалі алгоритм переходить до схеми branch-cut-and-price: у вузлі дерева

пошуку розв'язується обмежена головна задача, потім виконується породження нових маршрутів, а за потреби додаються відсічення і створюються дочірні вузли.

**Підзадача породження маршрутів.** Центральним елементом методу є підзадача, у якій для поточних двоїстих оцінок шукається маршрут з від'ємною зведеною вартістю. Саме такі маршрути можуть покращити поточний розв'язок лінійної релаксації головної задачі.

У базовій постановці, тобто без додаткових відсічень і обмежень розгалуження, якщо через  $\pi_i$  позначити двоїсті оцінки обмежень покриття клієнтів, а через  $\sigma \leq 0$  – двоїсту оцінку обмеження на кількість маршрутів, то зведена вартість (reduced cost) маршруту  $r$  дорівнює

$$\bar{c}_r = c_r - \sum_{i \in C} \pi_i a_{ir} - \sigma.$$

З наведеної формули випливає, що необхідно знайти маршрут з  $\bar{c}_r < 0$ . Після додавання відсічень і обмежень розгалуження у виразі зведеної вартості враховуються також двоїсті оцінки відповідних активних обмежень, однак сам принцип пошуку не змінюється.

Пошук маршруту з від'ємною зведеною вартістю зводиться до задачі ESPPRC. У реалізації точний етап розв'язується методом динамічного програмування з розширенням марок (labels) [3]. У цьому підході кожна марка описує частковий маршрут, побудований на поточному кроці: вона задається поточною вершиною, накопиченим попитом, зведеною вартістю, посиланням на попередню марку та множиною вже відвіданих клієнтів. Послідовно розширюючи такі марки, алгоритм буде допустимі часткові шляхи та відкидає ті з них, що вже не можуть привести до кращого маршруту. Множина відвіданих клієнтів забезпечує пряму заборону повторних відвідувань, а посилання на попередню марку використовується для відновлення маршруту.

Домінування розглядається лише між марками з однаковою поточною вершиною. Марка  $l_1$  відкидається, якщо існує марка  $l_2$  з тією самою поточною вершиною, не більшою зведеною вартістю і не більшим накопиченим попитом, а множина відвіданих клієнтів у  $l_2$  є підмножиною відповідної множини в  $l_1$ ; принаймні одна з цих переваг має бути строгою. У сертифікуючому елементарному етапі таке домінування є безпечним, оскільки воно усуває лише ті марки, будь-яке допустиме продовження яких не може бути кращим за відповідне продовження домінуючої марки. За повної рівності істотних характеристик у реалізації застосовується технічне правило розв'язання випадку рівності, яке залишає одну з еквівалентних марок і не змінює множини доступних продовжень. У тих варіантах розв'язання підзадачі породження маршрутів, де додатково активні інші ресурсні стани, безпечне домінування ще обмежується сумісністю цих станів. Обмеження розгалуження враховуються через модифікацію допустимих переходів, а двоїсті оцінки активних відсічень безпосередньо входять до зведеної вартості.

Щоб зменшити час роботи, спочатку застосовується евристичний режим розв'язання цієї підзадачі з обмеженням простору пошуку. За певних налаштувань ал-

горитму він використовує спрощену перевірку того, що маршрут не містить повторних відвідувань клієнтів. Конкретні прискорювальні опції цього режиму розглядаються як параметри реалізації: вони можуть впливати на час роботи, але не змінюють критерію завершення, оскільки негативний результат евристичного етапу завжди перевіряється точним пошуком. Якщо в цьому режимі знаходиться маршрут-кандидат з від'ємною зведеною вартістю, його додають до головної задачі як новий стовпець. Якщо ж нових маршрутів не знайдено, алгоритм переходить до точного ESPPRC-пошуку з урахуванням активних обмежень поточного вузла. Лише цей етап дає підставу стверджувати, що маршрутів з від'ємною зведеною вартістю більше немає.

Саме підзадача породження маршрутів визначає, чи достатньо поточного набору стовпців для опису оптимального розв'язку лінійної релаксації головної задачі. Тому ефективність усього методу істотно залежить від того, наскільки швидко вдається знаходити маршрути з від'ємною зведеною вартістю і наскільки надійно встановлюється їх відсутність.

**Посилення лінійної релаксації та розгалуження.** Для покращення оцінки знизу в головній задачі використовуються округлені ємнісні нерівності (Rounded Capacity Inequalities, RCI) [4, 5]. Для підмножини клієнтів  $S \subset C$  із сумарним попитом  $d(S) = \sum_{i \in S} q_i$  така нерівність у маршрутній постановці має вигляд  $\sum_{r \in R} \alpha_{Sr} x_r \geq 2[d(S)/Q]$ , де  $\alpha_{Sr}$  дорівнює кількості дуг маршруту  $r$ , що перетинають розріз між  $S$  та  $C \setminus S$ , тобто мають рівно одну кінцеву вершину в  $S$ .

Зміст цієї нерівності полягає в тому, що поточний дробовий розв'язок має забезпечувати достатню кількість виходів із множини  $S$  та входів до неї. У реалізації використовується процедурна евристична схема пошуку порушених нерівностей, а не повний алгоритм точного розділення для всіх можливих підмножин  $S$ . Спочатку за графом носія поточного дробового розв'язку будуються кандидатні підмножини клієнтів: зв'язні компоненти базового графа підтримки, компоненти порогових графів підтримки, а також множини, покращені локальними операціями вилучення або додавання окремих вершин. Для кожного такого кандидата перевіряється наведена умова, після чого до головної задачі додаються найсильніше порушені нерівності. Така процедура використовується лише для посилення лінійної релаксації; відсутність порушених нерівностей серед обраних кандидатів не трактується як повна сертифікація того, що жодної порушеної RCI більше не існує. Після додавання нових стовпців і відсічень лінійна релаксація перевизначається, а процес повторюється.

Якщо в поточному вузлі дерева пошуку отримано дробовий розв'язок, для розгалуження розглядаються дуги  $a$  з нецілим сумарним використанням  $f(a) = \sum_{r \in R: a \in r} x_r$ . Далі обирається дуга, для якої величина  $|f(a) - 0,5|$  є мінімальною. Отже, розгалуження виконується за переходом між двома вершинами, сумарне використання якого в поточному дробовому розв'язку є нецілим і найближчим до 0,5.

У дослідженій реалізації для дуги, обраної для розгалуження, в одному дочірньому вузлі в головній задачі накладається умова  $f(a) = 1$ , а в іншому – умова  $f(a) = 0$ . У підзадачі породження маршрутів умова  $f(a) = 0$  враховується прямою заборонаю дуги  $a$ . Умова  $f(a) = 1$  переноситься через заборону всіх альтернативних виходів із початкової вершини дуги  $a$  та всіх альтернативних входів у її кінцеву вершину. Така модифікація допустимих переходів зберігає узгодженість між головною задачею і підзадачею породження маршрутів та не порушує коректності всієї схеми пошуку.

**Забезпечення коректності схеми розв’язання.** Для методу, орієнтованого на точне розв’язання, принципово важливо відрізнити прискорювальний пошук нових стовпців від сертифікації їх відсутності. У пропонованій схемі евристичний режим використовується лише на першому з цих етапів. Сертифікуючий етап не використовує послаблення елементарності й спирається тільки на точний ESPPRC-пошук. Тому завершення генерації стовпців не ґрунтується на наближених або евристичних ознаках.

Ключовий інваріант пропонованої схеми полягає в такому: якщо в поточному вузлі дерева точний ESPPRC-пошук з урахуванням активних відсічень і обмежень розгалуження не знаходить жодного маршруту з  $\bar{c}_r < 0$ , то серед усіх допустимих маршрутів цього вузла відсутні стовпці, здатні покращити лінійну релаксацію. Отже, поточний розв’язок обмеженої головної задачі є оптимальним для повної лінійної релаксації в даному вузлі. Саме ця обставина, а не негативний результат евристичного етапу, використовується як підстава для завершення генерації стовпців.

Якщо евристичний етап породження маршрутів знаходить нові стовпці, вони додаються до обмеженої головної задачі, після чого цикл генерації стовпців продовжується. Якщо ж нових маршрутів не знайдено, запускається точний ESPPRC-пошук з урахуванням активних обмежень поточного вузла. Отже, евристичний етап може лише прискорити виявлення покращувальних маршрутів, але не впливає на сертифікацію їх відсутності. Це дозволяє поєднати коректність схеми пошуку з економією часу на тих ітераціях, де нові маршрути знаходяться швидко.

**Організація обчислювального експерименту.** Для експериментальної перевірки було обрано одинадцять стандартних екземплярів CVRP серії P [8]. Разом вони охоплюють діапазон розмірностей від 16 до 60 вершин включно з депо та різну кількість доступних транспортних засобів. Повний перелік протестованих екземплярів і результати їх розв’язання наведено в табл. 1.

Експеримент має валідаційний характер. Його метою є перевірка реалізації алгоритму та оцінювання часу роботи в дослідженому діапазоні задач, а не широке порівняння з іншими реалізаціями точних методів. У такій постановці першорядним є факт відтворення відомого оптимального значення, тоді як час розв’язання використовується для характеристики обчислювальної поведінки методу.

Лінійні релаксації обмеженої головної задачі розв’язувалися засобами GLOP із пакета Google.OrTools 9.14.6206. Програмну реалізацію виконано мовою C# у проєкті з цільовою платформою.NET 9.0. Запуски в режимі Release виконувалися в ОС

Windows 11 на персональному комп'ютері з процесором AMD Ryzen 5 7600 (6 ядер, 12 потоків) та 32 ГБ оперативної пам'яті.

Усі наведені результати отримано в штатній конфігурації реалізації в режимі Release без зовнішнього обмеження часу. У табл. 1 для кожного екземпляра наведено отримане значення цільової функції та середній час, обчислений за результатами десяти завершених запусків; позначення екземплярів подано у стандартному вигляді, прийнятому для відповідних тестових наборів. Час вимірювався як повний час одного завершення запуску алгоритму, після чого для кожного екземпляра обчислювалося середнє арифметичне за десятьма повторами. Для кожного запуску отримане значення цільової функції порівнювалося з відомим оптимумом.

Обраний набір тестів не охоплює всіх відомих класів екземплярів CVRP, однак дає змогу простежити поведінку реалізації на задачах різної розмірності та різної структурної складності в межах підтвердженого діапазону. Поєднання малих і середніх екземплярів дозволяє одночасно перевірити коректність реалізації та оцінити зміну обчислювальних витрат зі зростанням розмірності задачі.

**Результати дослідження.** Результати експерименту наведено в табл. 1.

Таблиця 1

Результати тестування алгоритму на стандартних екземплярах CVRP

Екземпляр	Кількість вершин, включно з депо	Кількість доступних транспортних засобів	Відомий оптимум	Отримане значення	Середній час за 10 запусків, с
P-n16-k8	16	8	450	450	0,31
P-n19-k2	19	2	212	212	1,32
P-n20-k2	20	2	216	216	1,36
P-n21-k2	21	2	211	211	0,91
P-n22-k2	22	2	216	216	1,29
P-n23-k8	23	8	529	529	0,36
P-n40-k5	40	5	458	458	16,13
P-n45-k5	45	5	510	510	107,15
P-n50-k7	50	7	554	554	13,87
P-n50-k10	50	10	696	696	31,86
P-n60-k10	60	10	744	744	72,97

Із табл. 1 видно, що на всіх досліджених екземплярах в усіх десяти запусках отримані значення повністю збігаються з відомими оптимумами. Для малих задач із 16-23 вершинами включно з депо середній час розв'язання не перевищує 2 с. Для задач із 40-60 вершинами включно з депо середній час роботи змінюється в межах від 13,87 до 107,15 с. Водночас залежність середнього часу від кількості вершин не є монотонною: у цій серії запусків екземпляр із 45 вершинами виявився складнішим за обидва екземпляри з 50 вершинами. Це узгоджується з тим, що складність CVRP зале-

жить не лише від розмірності, а й від просторової структури клієнтів, місткості автомобілів та конфігурації попитів.

Наведені часові показники є середніми значеннями повного часу роботи всієї схеми branch-cut-and-price, отриманими за десятьма завершеними запусками кожного екземпляра, а не характеристиками окремої локальної процедури. Вони охоплюють розв'язання обмеженої головної задачі, породження нових маршрутів, додавання відсічень і проходження дерева розгалуження. Тому для середніх екземплярів визначальними чинниками виступають не лише розмірність задачі, а й те, наскільки швидко вдається стабілізувати лінійну релаксацію та завершити точний пошук маршрутів із від'ємною зведеною вартістю. Саме ці компоненти є природними кандидатами для подальшого прискорення методу.

**Обговорення.** Отримані результати підтверджують, що реалізація алгоритму коректно працює на розглянутому валідаційному наборі тестів і відтворює відомі оптимуми на всіх наведених екземплярах. Водночас наведені часові показники слід трактувати саме як характеристику обчислювальної поведінки методу в дослідженому діапазоні задач, а не як універсальну оцінку його ефективності для всіх класів екземплярів CVRP.

З методологічного погляду цінність такої реалізації полягає в тому, що вона дає змогу досліджувати окремі компоненти схеми точного розв'язання в узгодженому середовищі. За наявності коректно організованої взаємодії між обмеженою головною задачею, генерацією стовпців, відсіченнями та розгалуженням можна окремо оцінювати вплив прискорювальних рішень, не змішуючи питання швидкодії з питанням правильності завершення алгоритму. Саме в цьому сенсі навіть відносно компактна реалізація, перевірена на стандартних екземплярах, є корисною як основа для подальших алгоритмічних досліджень.

Подальше підвищення ефективності доцільно пов'язувати з удосконаленням найбільш витратних компонентів методу. Насамперед це стосується сильніших схем домінування в підзадачі маршруту, покращення стратегій додавання відсічень, точніших правил вибору об'єкта розгалуження та інших засобів стабілізації лінійної релаксації. Перевірка таких удосконалень на ширшому колі екземплярів має становити окремий етап подальшого дослідження.

**Висновки.** У статті подано реалізацію алгоритму branch-cut-and-price для задачі маршрутизації транспортних засобів з обмеженнями на вантажопідйомність та наведено результати її дослідження. У запропонованій схемі поєднано модель розбиття множини, генерацію стовпців, стартову евристику на основі алгоритму Кларка-Райта, округлені ємнісні нерівності та розгалуження за дугами.

Обчислювальний експеримент на одинадцяти стандартних екземплярах із 16-60 вершинами включно з депо показав, що реалізація алгоритму в усіх наведених випадках відтворює відомі оптимуми. Це підтверджує коректність реалізації на розглянутому наборі стандартних задач і свідчить про практичну придатність обраної організації обчислень у дослідженому діапазоні розмірностей. Подальшу роботу доцільно спрямува-

ти на адаптацію алгоритму до розв'язання екземплярів більшої розмірності та на подальше вдосконалення його найвитратніших компонентів.

#### ЛІТЕРАТУРА

1. Archetti C., Coelho L.C., Speranza M.G., Vansteenwegen P. Beyond fifty years of vehicle routing: Insights into the history and the future. *European Journal of Operational Research*. 2026. Vol. 330, No. 2. P. 355-372. DOI: 10.1016/j.ejor.2025.06.014.
2. Costa L., Contardo C., Desaulniers G. Exact Branch-Price-and-Cut Algorithms for Vehicle Routing. *Transportation Science*. 2019. Vol. 53, No. 4. P. 946-985. DOI: 10.1287/trsc.2018.0878.
3. Feillet D., Dejax P., Gendreau M., Gueguen C. An exact algorithm for the elementary shortest path problem with resource constraints: Application to some vehicle routing problems. *Networks*. 2004. Vol. 44, No. 3. P. 216-229. DOI: 10.1002/net.20033.
4. Fukasawa R., Longo H., Lysgaard J., Poggi de Aragao M., Reis M., Uchoa E., Werneck R.F. Robust Branch-and-Cut-and-Price for the Capacitated Vehicle Routing Problem. *Mathematical Programming*. 2006. Vol. 106, No. 3. P. 491-511. DOI: 10.1007/s10107-005-0644-x.
5. Baldacci R., Christofides N., Mingozzi A. An exact algorithm for the vehicle routing problem based on the set partitioning formulation with additional cuts. *Mathematical Programming*. 2008. Vol. 115, No. 2. P. 351-385. DOI: 10.1007/s10107-007-0178-5.
6. Pecin D., Pessoa A., Poggi M., Uchoa E. Improved branch-cut-and-price for capacitated vehicle routing. *Mathematical Programming Computation*. 2017. Vol. 9, No. 1. P. 61-100. DOI: 10.1007/s12532-016-0108-8.
7. Clarke G., Wright J. W. Scheduling of Vehicles from a Central Depot to a Number of Delivery Points. *Operations Research*. 1964. Vol. 12, No. 4. P. 568-581. DOI: 10.1287/opre.12.4.568.
8. All Instances – CVRPLIB. GALGOS – Algorithms, Optimization and Simulation Group. URL: <https://galgos.inf.puc-rio.br/cvrplib/index.php/en/instances> (date of access: 02.04.2026).

#### REFERENCES

1. Archetti C., Coelho L.C., Speranza M.G., Vansteenwegen P. Beyond fifty years of vehicle routing: Insights into the history and the future. *European Journal of Operational Research*. 2026. Vol. 330, No. 2. P. 355-372. DOI: 10.1016/j.ejor.2025.06.014.
2. Costa L., Contardo C., Desaulniers G. Exact Branch-Price-and-Cut Algorithms for Vehicle Routing. *Transportation Science*. 2019. Vol. 53, No. 4. P. 946-985. DOI: 10.1287/trsc.2018.0878.
3. Feillet D., Dejax P., Gendreau M., Gueguen C. An exact algorithm for the elementary shortest path problem with resource constraints: Application to some vehicle routing problems. *Networks*. 2004. Vol. 44, No. 3. P. 216-229. DOI: 10.1002/net.20033.
4. Fukasawa R., Longo H., Lysgaard J., Poggi de Aragao M., Reis M., Uchoa E., Werneck R.F. Robust Branch-and-Cut-and-Price for the Capacitated Vehicle Routing Problem. *Mathematical Programming*. 2006. Vol. 106, No. 3. P. 491-511. DOI: 10.1007/s10107-005-0644-x.

5. Baldacci R., Christofides N., Mingozzi A. An exact algorithm for the vehicle routing problem based on the set partitioning formulation with additional cuts. *Mathematical Programming*. 2008. Vol. 115, No. 2. P. 351-385. DOI: 10.1007/s10107-007-0178-5.
6. Pecin D., Pessoa A., Poggi M., Uchoa E. Improved branch-cut-and-price for capacitated vehicle routing. *Mathematical Programming Computation*. 2017. Vol. 9, No. 1. P. 61-100. DOI: 10.1007/s12532-016-0108-8.
7. Clarke G., Wright J. W. Scheduling of Vehicles from a Central Depot to a Number of Delivery Points. *Operations Research*. 1964. Vol. 12, No. 4. P. 568-581. DOI: 10.1287/opre.12.4.568.
8. All Instances – CVRPLIB. GALGOS – Algorithms, Optimization and Simulation Group. URL: <https://galgos.inf.puc-rio.br/cvrplib/index.php/en/instances> (date of access: 02.04.2026).

Received 06.04.2026.

Accepted 08.04.2026.

Published 30.04.2026

### ***Implementation and study of a branch-cut-and-price algorithm for the Capacitated Vehicle Routing Problem***

*The Capacitated Vehicle Routing Problem is a classical optimization problem arising in transportation logistics. Modern research addresses it with both heuristic and exact approaches. Among exact methods, the most effective ones are based on a set-partitioning formulation, column generation, cutting planes, and branching. Important components of such schemes are the exact solution of the route-generation subproblem as an Elementary Shortest Path Problem with Resource Constraints (ESPPRC) and the use of rounded capacity inequalities to strengthen the linear relaxation. At the same time, building an implementation in which these components interact consistently remains a technically demanding task.*

*The aim of the study is to apply a branch-cut-and-price algorithm for the exact solution of the Capacitated Vehicle Routing Problem, to explain how its main components are coordinated, and to validate the implementation on standard benchmark instances.*

*In the proposed implementation, each feasible route is treated as a variable of a set-partitioning model. An initial transportation plan is constructed by a Clarke-Wright-based heuristic. The algorithm then proceeds within the branch-cut-and-price framework: a restricted master problem is solved and new negative reduced-cost routes are generated. To accelerate computations, the implementation first applies a heuristic search; if it finds no improving columns, the algorithm switches to an exact certifying stage. The route-generation subproblem is formulated as an ESPPRC and solved by dynamic programming with label extension. Rounded capacity inequalities are used to improve the linear relaxation, and branching is performed on arcs. Computational experiments were carried out on eleven standard instances with 16 to 60 vertices, including the depot. Each instance was solved in ten sequential runs, and the article reports the average runtime. In every reported case, the implementation matched the known optimal values.*

*The study shows that the proposed implementation of branch-cut-and-price operates correctly on the considered set of standard instances and reproduces the known optimal solutions. The reported results confirm the practical applicability of this computational organization for small and medium-sized instances. Further work should focus on adapting the imple-*

*mentation to larger instances and improving the most time-consuming components of the method.*

*Keywords: CVRP, vehicle routing, branch-cut-and-price, column generation, set partitioning, combinatorial optimization, integer programming, linear relaxation, exact and heuristic methods, logistics.*

**Ленський Михайло Миколайович** – аспірант кафедри інженерії програмного забезпечення та інформаційних технологій, Дніпровський національний університет імені Олеся Гончара, Дніпро, Україна.

ORCID: <https://orcid.org/0009-0001-1445-2142>

**Михальчук Ганна Йосипівна** – кандидат фізико-математичних наук, доцент кафедри інженерії програмного забезпечення та інформаційних технологій, Дніпровський національний університет імені Олеся Гончара, Дніпро, Україна.

ORCID: <https://orcid.org/0000-0002-5476-6349>

**Lenskyi Mykhailo** – postgraduate student, Department of Software Engineering and Information Technologies, Oles Honchar Dnipro National University, Dnipro, Ukraine.

ORCID: <https://orcid.org/0009-0001-1445-2142>

**Mykhalchuk Hanna** – Candidate of Physical and Mathematical Sciences, Associate Professor of the Department of Software Engineering and Information Technologies, Oles Honchar Dnipro National University, Dnipro, Ukraine.

ORCID: <https://orcid.org/0000-0002-5476-6349>

А.С. Гладуш, С.В. Клименко, В.Б. Мазуренко

## АВТОМАТИЗАЦІЯ ПРИЙНЯТТЯ РІШЕНЬ У БАНКІВСЬКІЙ СФЕРІ НА ОСНОВІ АНАЛІЗУ ДАНИХ ВЕЛИКОГО ОБСЯГУ ТА ГЕНЕРАТИВНОГО ШТУЧНОГО ІНТЕЛЕКТУ: ОГЛЯД СУЧАСНИХ ПІДХОДІВ

*Анотація.* У роботі здійснено систематичний огляд сучасних концепцій, методів та інструментів автоматизації процесу прийняття рішень у банківській сфері на основі технологій генеративного штучного інтелекту. Проаналізовано можливості використання інтелектуальних технологій - генеративних мовних моделей, прогнозних систем та засобів автоматизації управління. Визначено ключові виклики впровадження: галюцинації моделей, конфіденційність даних, автоматизаційне упередження та необхідність регуляторної відповідності. Обґрунтовано доцільність використання багаторівневої архітектури інтелектуальних систем, що включає етапи автоматизованого збору даних, семантичного стиснення даних та подальшої класифікації з використанням векторних представлень. Систематизовано метрики оцінювання моделей генеративного штучного інтелекту у фінансовому секторі. Ідентифіковано прогалини в існуючих дослідженнях та окреслено напрямки для прикладних досліджень у галузі верифікації категорій торгових клієнтів банку та класифікації видів їх діяльності.

*Ключові слова:* автоматизація, великі дані, інтелектуальний аналіз даних, прийняття рішень, банківська сфера, генеративний штучний інтелект, великі мовні моделі (LLM), ризику, цифрова трансформація, етика штучного інтелекту.

**Вступ.** Сучасний етап розвитку фінансового сектору характеризується стрімкою цифровізацією та інтеграцією інтелектуальних систем у внутрішні процеси банківських установ. Управління великими масивами даних та автоматизація операційних циклів стали ключовими факторами підвищення конкурентоспроможності сучасних цифрових сервісів. Окрім класичних методів машинного навчання, які зазвичай є основними методами, що використовуються у банківській сфері для оцінки ризиків та профілювання клієнтів, особливої уваги сьогодні набуває впровадження технологій генеративного штучного інтелекту (далі - ШІ, або GenAI). Актуальність задачі автоматизації прийняття рішень у банківській сфері шляхом впровадження технологій ШІ постає у зв'язку з необхідністю обробки неструктурованої інформації, потреби в автоматизації експертних висновків у режимі реального часу.

Існує велика кількість робіт, присвячених використанню алгоритмів штучного інтелекту в банківській сфері. Проте вони переважно розглядають застосування лише класичних прогнозних моделей для кредитного скорингу або виявлення шахрайства. Практично відсутні систематизовані огляди, які б детально аналізували напрямки за-

стосування методів генеративного ШІ для задач верифікації категорій діяльності торговельних клієнтів (мерчантів) та контролю їх відповідності вимогам платіжних систем.

**Постановка проблеми.** Сучасний банківський сектор потребує переходу від класичних методів машинного навчання до впровадження генеративного штучного інтелекту (GenAI) для обробки неструктурованих даних у режимі реального часу. Проте інтеграція таких технологій у процеси верифікації торговельних клієнтів (мерчантів) стикається з викликами низької інтерпретованості рішень, ризиками «галюцинацій» моделей та необхідністю суворої відповідності вимогам міжнародних платіжних систем, що наразі недостатньо досліджено у прикладній площині.

**Метою дослідження** є систематизація сучасних підходів до автоматизації прийняття рішень у банківській сфері із залученням генеративного штучного інтелекту, аналіз існуючих методик оцінювання їх ефективності, а також ідентифікація прогалин у поточних дослідженнях для визначення напрямків подальшого розвитку методів автоматизації цієї галузі.

**Аналіз останніх досліджень та публікацій.**

**Використання генеративного штучного інтелекту у фінансовому секторі**

Надалі розглянуто найбільш знакові дослідження, що стосуються трансформації фінансового сектору під впливом генеративних моделей.

У роботах Nie В. та Сао Y. [1, 2] здійснено всебічний огляд застосування генеративного ШІ у фінансовій сфері, де автори виділяють потенціал мовних моделей для автоматизації звітності та покращенні клієнтського досвіду. Проте, як зазначають дослідники, основною перешкодою для їх повного впровадження залишається проблема «галюцинацій» та відсутність прозорості у прийнятті рішень.

Питання довіри та етичного впровадження великих мовних моделей (LLM - Large Language Model) детально розглядаються у звіті Інституту Алана Тюрінга [3]. Автори наголошують на необхідності створення фреймворків «відповідального ШІ», які б гарантували конфіденційність даних та відсутність автоматизаційного упередження. У дослідженні Bhatia K. та співавторів [4] акцент зміщено на управління ризиками моделей, де пропонується використовувати багаторівневі системи валідації для критичних банківських процесів.

Окремий пласт досліджень присвячено технічним аспектам класифікації. Так, у праці Agarwal S. [5] описано підходи до надійної класифікації продуктів у комерції за допомогою LLM, що за своєю логікою є близьким до задачі верифікації категорій торговців. Проте, аналіз публікацій останніх років [6, 7] свідчить, що більшість існуючих систем орієнтована на загальні задачі обробки природньої мови (NLP – Natural Language Processing) і не враховує специфічні вимоги Міжнародних платіжних систем щодо перевірки торговельних клієнтів у режимі реального часу.

Систематизація оглянутих джерел дозволяє стверджувати, що попри значний прогрес у розробці генеративних архітектур, практичні аспекти їх інтеграції в наскрізні банківські процеси перевірки контрагентів висвітлені фрагментарно. Ці прогалини окрес-

люють напрямки подальшого дослідження, зокрема щодо створення комбінованих систем, які поєднують точність класичних моделей із гнучкістю генеративного ШІ.

### **Ідентифіковані прогалини**

Аналіз публікацій виявляє кілька суттєвих прогалин у застосуванні GenAI-підходів, зокрема у галузі верифікації категорій торговельних клієнтів банку та класифікації видів їх діяльності :

- відсутність наскрізних процесів машинного навчання (ML-пайплайнів) для класифікації видів діяльності торговельних клієнтів за кодами МСС (Merchant Category Code), в подальшому – МСС-верифікації. Існуючі класичні рішення машинного навчання (ML-рішення) у банкінгу зазвичай зосереджені на кредитному ризику та шахрайстві, тоді як задача автоматизованої перевірки коректності кодів МСС на основі контенту веб-сайтів торговельних клієнтів практично не досліджена;

- недостатня увага до практичних аспектів GenAI у банкінгу. Роботи описують ризики GenAI теоретично, але не пропонують конкретних механізмів захисту для банківських задач.

Виявлені прогалини потребують пошуку нових методів для вирішення наявних задач та окреслюють напрямки подальшого дослідження.

Аналіз методологічних підходів до впровадження генеративного штучного інтелекту

Систематизація оглянутих публікацій дозволяє виокремити ключові методологічні вектори, що визначають сучасні стандарти автоматизації прийняття рішень у фінансовому секторі.

### **Системний підхід до автоматизації**

На основі аналізу останніх досліджень [2, 6] можна стверджувати, що ефективна інтеграція генеративних моделей можлива лише за умови створення багаторівневої архітектури. Вона зазвичай включає етапи автоматизованого збору даних (парсингу), семантичного стиснення контенту та подальшої класифікації з використанням векторних представлень. Такий підхід дозволяє нівелювати надмірну складність неструктурованих даних, що є характерним для верифікації торговельних клієнтів.

### **Критерії оцінювання та метрики**

У науковій літературі [4, 8] особлива увага приділяється специфічним метрикам оцінювання моделей генеративного ШІ (англ. GenAI). Окрім класичних показників точності, дослідники виділяють стійкість до «галюцинацій» та здатність моделі до роботи в умовах обмеженої кількості прикладів (англ. few-shot learning). Аналіз підтверджує, що для банківської сфери критично важливим є не лише результат класифікації, а й рівень впевненості моделі у прийнятому рішенні.

### **Управління ризиками та регуляторна відповідність**

Оглянуті джерела [3, 5, 8] одноставно вказують на низку специфічних ризиків, зокрема щодо конфіденційності даних та автоматизаційного упередження. Узагальнення світового досвіду дозволяє виділити основні механізми мінімізації цих загроз:

1. Людина у контурі (англ. Human-in-the-loop) — залучення експерта для перевірки складних або неоднозначних кейсів;

2. Тіньовий режим (англ. Shadow Mode) — паралельне тестування ШІ-системи без впливу на реальні бізнес-процеси до моменту підтвердження її надійності;

3. Етичний моніторинг — постійний аудит вихідних даних на предмет відповідності законодавчим нормам.

Така систематизація методологічних підходів демонструє, що попри наявність технічних інструментів, галузь потребує подальшого опрацювання прикладних фреймворків для верифікації категорій діяльності клієнтів у реальних умовах банківського еквіваленту.

**Висновки.** У роботі здійснено систематичний огляд сучасних підходів до автоматизації прийняття рішень у банківській сфері з використанням технологій аналізу великих даних та генеративного штучного інтелекту. На основі проведеного аналізу наукових публікацій отримано такі результати:

1. Систематизовано технологічні підходи. Встановлено, що сучасні системи автоматизації еволюціонують від класичних прогнозних моделей до комплексних архітектур, що поєднують векторні представлення даних (ембедінги) та великі мовні моделі. Це дозволяє ефективно обробляти неструктуровану інформацію про торговельних клієнтів (мерчантів), яка раніше потребувала виключно ручної обробки.

2. Ідентифіковано ключові ризики та методи їх нівелювання. Огляд підтвердив, що основними бар'єрами для впровадження генеративного ШІ в банках є проблема «галюцинацій» та вимоги регуляторної відповідності. Аналіз джерел дозволив виділити найбільш перспективні механізми захисту, такі як режим тіньового тестування (Shadow Mode) та концепція залучення людини до контуру прийняття рішень (Human-in-the-loop).

3. Визначено прогалини в існуючих дослідженнях. Попри наявність потужних базових моделей, у літературі недостатньо висвітлені питання створення наскрізних систем саме для верифікації категорій діяльності клієнтів (МСС-кодів) у реаліях українського ринку та специфічних вимог міжнародних платіжних систем.

#### **Напрямки подальших досліджень.**

Виявлені прогалини вказують на необхідність розробки та експериментальної перевірки спеціалізованих систем автоматизації, які б поєднували точність традиційного машинного навчання з аналітичними можливостями генеративного штучного інтелекту для верифікації параметрів торговельних клієнтів у режимі реального часу.

#### **ЛІТЕРАТУРА**

1. Nie B., et al. A Survey of Generative AI in Finance // HAL preprint. 2025.  
URL: <https://hal.science/hal-05020829>.
2. Cao Y., et al. A Comprehensive Review of Generative AI in Finance // FinTech. 2024. Vol. 3, No. 3. P. 25. DOI: 10.3390/fintech3030025.
3. Bholat D., et al. The Impact of Large Language Models in Finance: Towards Trustworthy Adoption // The Alan Turing Institute, FAIR Programme. 2024.  
URL: <https://www.turing.ac.uk>.

4. Bhatia K., et al. Model Risk Management for Generative AI in Financial Institutions // arXiv preprint. 2025. arXiv:2503.15668. DOI: 10.48550/arXiv.2503.15668.
5. Agarwal S., et al. LLM-Based Robust Product Classification in Commerce and Compliance // Proc. ACL Workshop CustomNLP4U. 2024.  
URL: <https://aclanthology.org/2024.customnlp4u-1.3/>.
6. Najem R., Bahnasse A., Fakhouri A. M., et al. Advanced AI and big data techniques in E-finance: a comprehensive survey // Discover Artificial Intelligence. 2025.  
DOI: 10.1007/s44163-025-00365-y.
7. Tong K., Han Z., Shen Y., Long Y., Wei Y. An Integrated Machine Learning and Deep Learning Framework for Credit Card Approval Prediction // arXiv preprint. 2024.
8. Ji Z., et al. Survey of Hallucination in Natural Language Generation // ACM Computing Surveys. 2023. Vol. 55, No. 12. DOI: 10.1145/3571730.
9. Elsiddig Ahmed I., Mehdi R., Mohamed E. The role of artificial intelligence in developing a banking environment // International Journal of ADVANCED AND APPLIED SCIENCES. 2022. DOI: 10.21833/ijaas.2022.10.016.

#### REFERENCES

1. Nie, B., et al. (2025). A Survey of Generative AI in Finance. *HAL preprint*. <https://hal.science/hal-05020829>.
2. Cao, Y., et al. (2024). A Comprehensive Review of Generative AI in Finance. *FinTech*, 3(3), 25. <https://doi.org/10.3390/fintech3030025>.
3. Bholat, D., et al. (2024). *The Impact of Large Language Models in Finance: Towards Trustworthy Adoption*. The Alan Turing Institute, FAIR Programme.  
<https://www.turing.ac.uk>.
4. Bhatia, K., et al. (2025). Model Risk Management for Generative AI in Financial Institutions. *arXiv preprint*. arXiv:2503.15668. <https://doi.org/10.48550/arXiv.2503.15668>.
5. Agarwal, S., et al. (2024). LLM-Based Robust Product Classification in Commerce and Compliance. Proc. ACL Workshop CustomNLP4U.  
<https://aclanthology.org/2024.customnlp4u-1.3/>.
6. Najem, R., Bahnasse, A., Fakhouri, A. M., et al. (2025). Advanced AI and big data techniques in E-finance: a comprehensive survey. *Discover Artificial Intelligence*. <https://doi.org/10.1007/s44163-025-00365-y>.
7. Tong, K., Han, Z., Shen, Y., Long, Y., & Wei, Y. (2024). An Integrated Machine Learning and Deep Learning Framework for Credit Card Approval Prediction. *arXiv preprint*..
8. Ji, Z., et al. (2023). Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, 55(12). <https://doi.org/10.1145/3571730>.
9. Elsiddig Ahmed, I., Mehdi, R., & Mohamed, E. (2022). The role of artificial intelligence in developing a banking environment. *International Journal of ADVANCED AND APPLIED SCIENCES*. <https://doi.org/10.21833/ijaas.2022.10.016>.

Received 06.04.2026.  
Accepted 09.04.2026.  
Published 30.04.2026

***Automation of decision-making in the banking sector based on big data analysis and generative artificial intelligence: a review of modern approaches***

*This paper provides a systematic review of modern concepts, methods, and tools for automating decision-making processes in the banking sector based on generative artificial intelligence technologies. The rapid evolution of the financial sector and the necessity of managing massive datasets have made the integration of intelligent systems a key factor for maintaining competitiveness in modern digital services. The research analyzes the possibilities of using various intelligent technologies, including generative language models, predictive systems, and advanced control automation tools. Special attention is given to the shift from classical machine learning, typically used for risk assessment, to Generative Artificial Intelligence for processing unstructured information in real-time.*

*The study identifies key implementation challenges that remain primary barriers to adoption: model "hallucinations," data privacy concerns, automation bias, and the critical need for strict regulatory compliance. To address these issues, the paper substantiates the expediency of using a multi-level architecture for intelligent systems. This architecture includes the stages of automated data collection, semantic data compression of content, and subsequent classification using vector representations and embeddings.*

*Moreover, the research systematizes specific metrics for evaluating generative artificial intelligence models within the financial sector, emphasizing model resilience, and the level of confidence in decision-making. Finally, the study identifies significant gaps in existing research, particularly the lack of end-to-end machine learning pipelines for the verification of merchant category codes based on website content and outlines the space for the further applied research in the field of merchant activity classification.*

*Keywords: automation, BigData, intelligent data analysis, decision-making, banking, generative Artificial Intelligence (GenAI), Large Language Models (LLM), risks, digital transformation, AI ethics.*

**Гладуш Андрій Сергійович** - аспірант кафедри кібербезпеки та комп'ютерно-інтегрованих технологій Фізико-технологічного факультету Дніпропетровського Національного Університету ім. Олеса Гончара.

ORCID: <https://orcid.org/0009-0009-6025-8533>

**Клименко Світлана Володимирівна** - к.т.н., доцент, завідувачка кафедри кібербезпеки та комп'ютерно-інтегрованих технологій Фізико-технологічного факультету Дніпропетровського Національного Університету ім. Олеса Гончара.

ORCID: <https://orcid.org/0000-0003-2005-9993>

**Мазуренко Валерій Борисович** - к.т.н., доцент кафедри кібербезпеки та комп'ютерно-інтегрованих технологій Фізико-технологічного факультету Дніпропетровського Національного Університету ім. Олеса Гончара.

ORCID: <https://orcid.org/0000-0001-8340-012X>

**Hladush Andrii Serhiiovich** - PhD student of the department of cybersecurity and computer-integrated technologies, faculty of physics and technology, Oles Honchar Dnipro National University.

ORCID: <https://orcid.org/0009-0009-6025-8533>

**Klymenko Svitlana Volodymyrivna** - candidate of engineering sciences, assistant professor, head of the Department of Cybersecurity and Computer-Integrated Technologies, Faculty of Physics and Technology, Oles Honchar Dnipro National University.

ORCID: <https://orcid.org/0000-0003-2005-9993>

**Mazurenko Valerii Borysovych** - candidate of engineering sciences, assistant professor of the Department of Cybersecurity and Computer-Integrated Technologies, Faculty of Physics and Technology, Oles Honchar Dnipro National University.

ORCID: <https://orcid.org/0000-0001-8340-012X>

Д.А. Кондрашов, І.О. Музика, А.І. Купін, Д.К. Балик

## КЛАСИФІКАЦІЯ ПРОСТОРОВИХ СТАНІВ ОБ'ЄКТІВ НА ОСНОВІ КАРТ ВИСОТ ІЗ ВИКОРИСТАННЯМ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ СИСТЕМ СЕНСОРНОГО СОРТУВАННЯ РУД

*Анотація.* У статті розглядається задача автоматичної класифікації просторових станів об'єктів у системах прецизійного сенсорного сортування рудної сировини. Актуальність дослідження обумовлена обмеженою ефективністю традиційних алгоритмів сегментації в умовах накладання та зближення шматків руди на конвеєрі. Метою роботи є розробка та дослідження підходу до класифікації трьох просторових станів об'єктів (одиначний, об'єднаний, накладений) на основі карт висот, отриманих за допомогою лазерного сканування. Як метод дослідження використовуються згорткові нейронні мережі (CNN). Проведено порівняльний аналіз архітектур MLP, базової CNN, CNN із регуляризацією та поглибленої CNN. Досліджено вплив розмірності карти висот (32×32, 64×64, 128×128) на якість класифікації. Встановлено, що поглиблена CNN демонструє найкращу точність на реальних даних – до 90% на валідаційній вибірці при розмірності 64×64. Виявлено ефект перенавчання при збільшенні розмірності вхідних даних. Отримані результати підтверджують перспективність застосування CNN для аналізу карт висот у задачах інтелектуального керування процесом сортування.

*Ключові слова:* сенсорне сортування руд, карта висот, лазерне сканування, згорткова нейронна мережа, класифікація зображень, машинне навчання, прецизійне сортування, хмара точок, перенавчання, аугментація даних.

**Постановка проблеми.** Ефективність сучасних гірничо-збагачувальних комбінатів значною мірою визначається якістю первинного сортування рудної сировини. Процес прецизійного сортування руд передбачає автоматичне розпізнавання та відбраковування некондиційних шматків руди на основі аналізу їх фізико-хімічних властивостей та геометричних характеристик. Сучасні системи сенсорного сортування використовують різноманітні типи сенсорів: оптичні камери, NIR/hyperspectral сканери, рентгенівські детектори (XRT), електромагнітні сенсори та 3D лазерні сканери [7].

Однією з ключових проблем при застосуванні систем сенсорного сортування є складність аналізу просторового стану об'єктів у рудному потоці. Зокрема, на конвеєрній стрічці шматки руди можуть знаходитися у трьох принципово різних конфігураціях: у вигляді окремих об'єктів, у вигляді об'єктів, які візуально зближені але фізич-

но не торкаються один одного, а також у вигляді об'єктів, що фізично перекривають один одного.

Традиційні детерміновані алгоритми комп'ютерного зору з фіксованими пороговими значеннями для класифікації та сегментації демонструють обмежену здатність до адаптації та потребують експертного налаштування під конкретні умови експлуатації. За даними промислових підприємств, частка помилкових спрацювань при використанні класичних методів може досягати 15–25%, що призводить до втрат корисних компонентів або забруднення кондиційного продукту [7]. Це обумовлює необхідність розробки інтелектуальних методів аналізу просторових даних, здатних автоматично виявляти складні закономірності та адаптуватися до мінливих умов роботи.

Лазерне сканування дозволяє отримувати просторову інформацію про поверхню об'єктів у вигляді тривимірних хмар точок або карт висот. На відміну від 2D зображень, такі дані містять геометричну інформацію, яка є інваріантною до умов освітлення та оптичних властивостей поверхні. Проте суттєвою проблемою залишається складність автоматичної інтерпретації таких даних, особливо у випадках, коли об'єкти розташовані близько один до одного або накладаються. У таких ситуаціях традиційні алгоритми обробки зображень та сегментації демонструють нестабільні результати, що обумовлює доцільність застосування методів машинного навчання.

**Аналіз останніх досліджень і публікацій.** Питання застосування методів машинного навчання для класифікації рудної сировини активно досліджується останніми роками. У роботі Robben C. та Wotruba H. [7] представлено всебічний огляд технологій сенсорного сортування в мінеральній промисловості. Автори аналізують різноманітні типи сенсорів та практичні кейси впровадження систем сенсорного сортування, що демонструють підвищення ефективності збагачення на 12–18%. Однак відзначається, що традиційні підходи з фіксованими алгоритмами класифікації мають обмежену ефективність при зміні характеристик рудного потоку, що підкреслює необхідність розробки адаптивних інтелектуальних систем.

У роботі Peukert S. та співавторів [4] проведено комплексний аналіз використання електромагнітного спектру в процесах сенсорного сортування. Автори досліджують можливості поєднання 3D геометричної інформації від лазерних сканерів зі спектральними даними від hyperspectral камер. Встановлено, що мультисенсорний підхід дозволяє підвищити точність класифікації на 8–12% порівняно з використанням одного типу сенсорів. Особливо ефективним виявилось поєднання геометричних характеристик (форма, розмір, текстура поверхні) з інформацією про мінералогічний склад, отриманою зі спектральних даних.

Застосування CNN для аналізу 2D зображень шматків руди розглянуто у роботі Rai S. та співавторів [6]. Експериментальні дослідження показали точність класифікації 89–94% для трьох типів руд. Однак дослідження обмежене аналізом лише 2D зображень без урахування просторової геометрії, що знижує ефективність розпізнавання при накладанні проєкцій декількох шматків.

Перспективним напрямком є застосування 3D нейронних мереж для аналізу хмар точок. У роботі Qi C. R. та співавторів [5] запропоновано архітектуру PointNet++ для

класифікації 3D об'єктів у хмарах точок. Ця архітектура дозволяє ефективно обробляти нерегулярні та неоднорідні 3D дані, типові для лазерного сканування. Точність класифікації на тестових датасетах перевищила 92%. Однак дослідження проводилося на синтетичних даних та не враховувало специфіки промислового застосування.

Методи виявлення аномалій та нестандартних ситуацій у промислових процесах досліджено у роботі Chalapathy R. та Chawla S. [1]. Автори пропонують використання автоенкодерів та методів One-Class SVM для виявлення відхилень від нормального режиму роботи. Точність виявлення аномалій склала 85–91% при рівні помилкових спрацювань 3–7%.

Питання оптимізації нейронних мереж для роботи в режимі реального часу розглянуто у роботі Gholami A. та співавторів [2]. Застосування методів квантизації та grouping дозволяє прискорити інференс у 3–5 разів зі зниженням точності класифікації лише на 1–2%, що є важливим для промислових систем, де час обробки одного об'єкта не повинен перевищувати 50–100 мс.

Методи transfer learning та domain adaptation для адаптації моделей до змін у вхідних даних досліджено у роботах Wang M. та Deng W. [8], а також Parisi G. I. та співавторів [3]. Ці підходи особливо актуальні для систем сенсорного сортування, де характеристики рудної сировини можуть змінюватися залежно від родовища та умов видобутку.

Таким чином, аналіз літератури виявляє значний науковий інтерес до застосування методів машинного навчання у системах сенсорного сортування руд. Однак більшість досліджень зосереджена на класифікації за 2D зображеннями або спектральними даними, тоді як потенціал аналізу карт висот від 3D лазерного сканування для задачі розпізнавання просторових конфігурацій об'єктів залишається недостатньо вивченим. Окремою невирішеною проблемою є побудова ефективних класифікаторів, здатних розрізняти одиночні, об'єднані та накладені об'єкти за даними лазерного сканування.

**Мета дослідження.** Метою даної роботи є розробка та дослідження підходу до автоматичної класифікації просторових станів об'єктів (одиночний, об'єднаний, накладений) на основі карт висот, отриманих від лазерного сканера, із використанням згорткових нейронних мереж, а також порівняльний аналіз різних архітектур нейронних мереж та дослідження впливу розмірності карти висот на якість класифікації.

Для досягнення поставленої мети вирішуються такі задачі: (1) розробка методу формування карт висот на основі тривимірних точкових хмар; (2) генерація та аналіз синтетичного датасету для початкової перевірки підходу; (3) формування реального датасету на основі даних лазерного сканування із застосуванням аугментації; (4) розробка та навчання декількох архітектур нейронних мереж; (5) порівняльний аналіз результатів та виявлення ключових факторів, що впливають на якість класифікації.

## **Викладення основного матеріалу дослідження**

### **1. Постановка задачі та формування карт висот**

У роботі розглядається задача багатокласової класифікації стану об'єктів на основі їх представлення у вигляді дискретизованих карт висот. Вхідними даними є мат-

риці розміром  $N \times N$ , що відображають висоту поверхні об'єктів у кожній точці регулярної сітки. Кожному об'єкту відповідає один із трьох класів: одиночний об'єкт (Single) – окремий шматок руди на конвеєрі; візуально об'єднані об'єкти (United) – два або більше шматки, розташовані близько один до одного, але без фізичного перекриття; накладені об'єкти (Overlapped) – шматки, що фізично перекривають один одного по вертикальній проекції.

Карта висот формується на основі тривимірних точкових хмар, отриманих від лазерного сканера. Процедура формування включає такі кроки: (1) дискретизація простору  $X$ - $Y$  з кроком, що визначає розмірність сітки  $N \times N$ ; (2) для кожної клітинки  $(i, j)$  сітки визначається максимальне значення координати  $Z$  серед усіх точок, що потрапляють до відповідної комірки; (3) клітинки без точок заповнюються нульовим значенням; (4) отримана матриця нормалізується до діапазону  $[0, 1]$ .

Формально, якщо  $P = \{(x_k, y_k, z_k)\}$  – множина точок хмари, а  $G$  – регулярна сітка з розміром клітинки  $\Delta$ , то карта висот  $H$  визначається як:

$$H(i, j) = \max \{ z_k \mid (x_k, y_k) \in \text{cell}(i, j) \}, \quad (1)$$

де  $\text{cell}(i, j)$  – область простору  $X$ - $Y$ , що відповідає клітинці  $(i, j)$  сітки. При відсутності точок у клітинці  $H(i, j) = 0$ .



Рисунок 1 – Приклад синтетично згенерованих карт висот для трьох класів (Single, United, Overlapped)

## 2. Формування та дослідження синтетичних даних

На початковому етапі дослідження було реалізовано генерацію синтетичних даних з метою перевірки працездатності підходу та відлагодження процедури формування карт висот. Кожен об'єкт моделювався як набір тривимірних точок із координатами  $(X, Y, Z)$ , де координати  $X$  та  $Y$  задавали положення в площині, а  $Z$  – висоту. Форма об'єктів апроксимувалася еліпсоїдом із випадковими параметрами, що дозволяло моделювати природну варіативність форми шматків руди.

Для кожного класу було розроблено окрему процедуру генерації: для класу Single генерувався один об'єкт, розміщений у центральній частині поля спостереження; для класу United генерувалися два об'єкти, розташовані поряд з відстанню, меншою за їх характерний розмір, але без перетину; для класу Overlapped генерувалися два або більше об'єкти з перетинними горизонтальними проекціями. У кожному випадку генерувалося по 500 зразків із випадковими параметрами, що забезпечувало достатнє різноманіття навчальних прикладів.

Результати навчання нейронної мережі на синтетичних даних показали майже ідеальну точність класифікації – понад 99% як на тренувальній, так і на валідаційній вибірках. Це свідчило про надмірну спрощеність задачі: синтетичні об'єкти мали чіткі та ідеальні геометричні форми, що не відповідає реальним умовам. Крім того, розподіл класів у синтетичних даних виявився надто рівномірним порівняно з реальними даними. Дані результати підтвердили правильність методологічного підходу, але показали необхідність переходу до більш реалістичних даних для отримання практично значущих результатів.

### 3. Реальний датасет та аугментація даних

Після дослідження синтетичних даних було здійснено перехід до роботи з реальними картами висот, отриманими в результаті лазерного сканування фізичних зразків рудної сировини. Дані були представлені у вигляді CSV-файлів, кожен із яких містить матрицю значень висоти для одного зразка. Датасет було організовано у ієрархічну структуру за класами та розмірностями сіток.

Для кожного класу та розміру сітки були доступні базові скани (позначення R0) без повороту. З метою збільшення обсягу навчальної вибірки та покращення узагальнювальних властивостей моделі було застосовано аугментацію даних шляхом повороту карт висот на фіксовані кути:  $90^\circ$ ,  $180^\circ$  та  $270^\circ$ . Це дозволило збільшити обсяг датасету вчетверо. Вибір саме ротаційної аугментації обґрунтований фізичним змістом задачі: орієнтація шматків руди на конвеєрі є довільною, тому модель повинна бути інваріантною до повороту.

Особливістю реальних даних є наявність характерних дефектів: пропуски значень висоти в окремих клітинках (через відсутність відбитого лазерного променя), нерівномірність поверхні шматків руди, артефакти сканування у вигляді одиничних викидів значень, а також шумові ефекти, пов'язані з властивостями поверхні матеріалу. Наявність цих дефектів значно ускладнює задачу класифікації, але водночас робить її більш реалістичною та практично значущою.

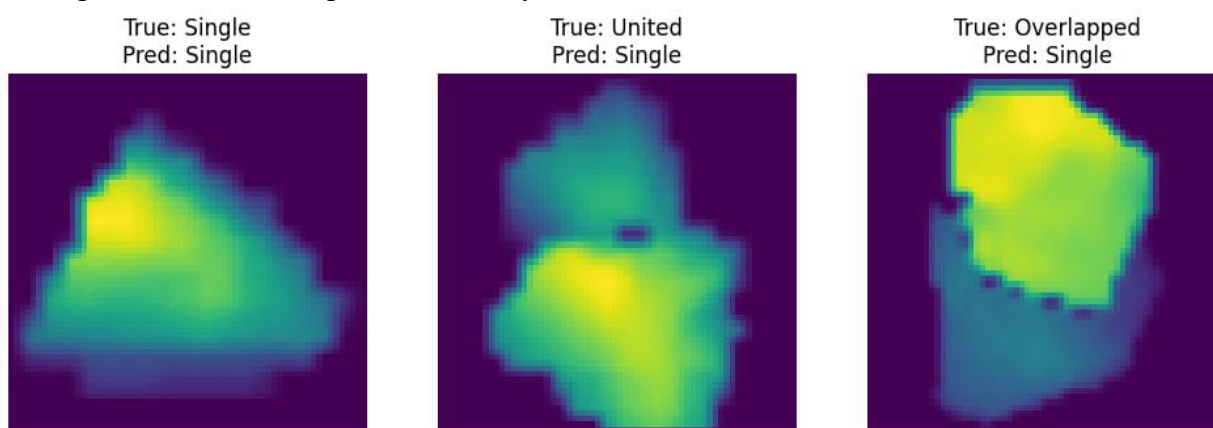


Рисунок 2 – Приклади реальних карт висот із дефектами для кожного з трьох класів

Загальна структура реального датасету включала зразки для трьох розмірностей сіток ( $32 \times 32$ ,  $64 \times 64$ ,  $128 \times 128$ ) та трьох класів (Single, United, Overlapped). Після ауг-

ментації загальний обсяг датасету склав кілька тисяч зразків, що є достатнім для навчання нейронних мереж середньої складності. Розбиття на тренувальну та валідаційну вибірки здійснювалося у співвідношенні 70/30 із збереженням пропорцій класів (stratified split).

#### 4. Дослідження впливу розмірності карти висот

Розмірність карти висот є ключовим гіперпараметром, що визначає баланс між деталізацією просторової інформації та складністю задачі класифікації. Більша розмірність сітки дозволяє зберегти більше дрібних деталей геометрії об'єктів, проте одночасно збільшує розмір вхідного тензора та кількість параметрів моделі, що може призводити до перенавчання при обмеженому обсязі даних.

У рамках дослідження було проведено порівняльні експерименти для трьох варіантів розмірності:  $32 \times 32$ ,  $64 \times 64$  та  $128 \times 128$ . Для кожної розмірності навчалися ідентичні архітектури нейронних мереж (поглиблена CNN) з однаковими гіперпараметрами навчання. Ключові результати наведено нижче.

Для розмірності  $32 \times 32$  спостерігалася висока стабільність навчання та мінімальна різниця між тренувальною ( $\sim 0.94$ ) та валідаційною ( $\sim 0.91$ ) точністю. Це пояснюється порівняно невеликою кількістю вхідних параметрів та достатністю наявних даних для навчання. Для розмірності  $64 \times 64$  тренувальна точність зросла до  $\sim 0.98$ , а валідаційна – зменшилася до  $\sim 0.90$ . Спостерігається початкова ознака перенавчання. Для розмірності  $128 \times 128$  тренувальна точність залишалася високою ( $\sim 1.0$ ), тоді як валідаційна точність знизилася до  $\sim 0.85$ . Розрив між тренувальною та валідаційною точністю значно збільшився, що є явною ознакою перенавчання.

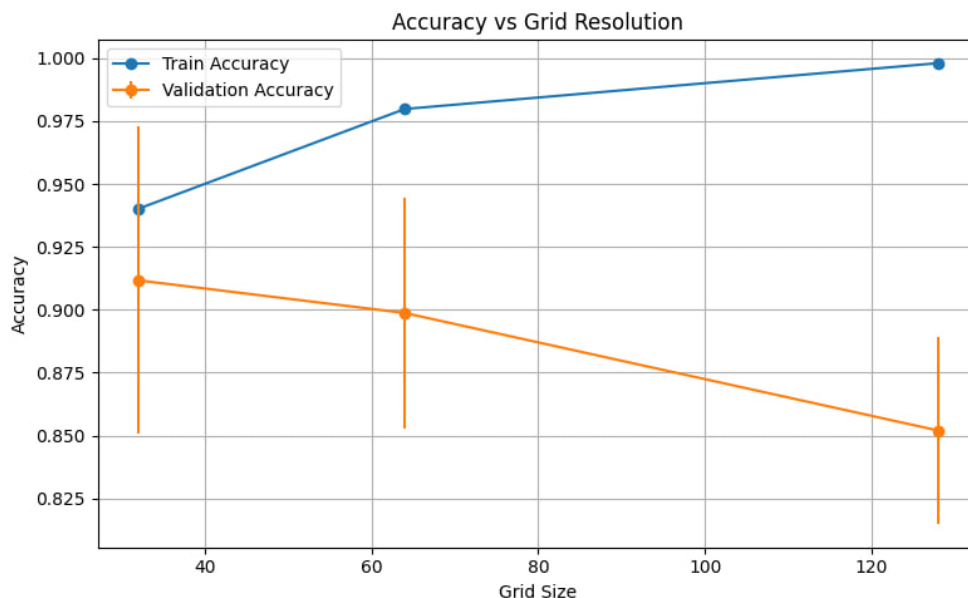


Рисунок 3 – Залежність точності класифікації від розмірності карти висот для тренувальної та валідаційної вибірок

Таким чином, розмірність  $64 \times 64$  забезпечує найкращий баланс між деталізацією просторової інформації та узагальнювальними властивостями моделі. Подальші дослідження проводилися переважно на даних розмірності  $64 \times 64$ .

## 5. Архітектури нейронних мереж

Для розв'язання задачі класифікації карт висот було розроблено та досліджено декілька архітектур нейронних мереж, від простих до складніших. Таким чином забезпечується можливість порівняльного аналізу та обґрунтування вибору оптимальної архітектури.

Базова багатошарова перцептронна мережа (MLP) – найпростіша архітектура, що не враховує просторову структуру даних. Карта висот перетворюється на одновимірний вектор (flatten), після чого послідовно з'єднуються кілька повнозв'язних шарів (Dense). Незважаючи на простоту, MLP слугує базовою лінією для порівняння більш складних підходів.

Базова CNN містить два згорткові блоки (Conv2D + MaxPooling), за якими слідує блок класифікації з повнозв'язними шарами. Згорткові шари дозволяють автоматично виявляти локальні просторові ознаки (границі об'єктів, піки висот, характерні геометричні структури), а шари підвибірки забезпечують просторову інваріантність та зменшення розмірності.

CNN із регуляризацією включає додаткові механізми для боротьби з перенавчанням: Dropout (відключення нейронів з ймовірністю 0.3–0.5 під час навчання) та L1-регуляризація ваг (Manhattan regularization). Ці методи обмежують складність моделі та покращують її узагальнювальні властивості.

Поглиблена CNN (Deep CNN) є основною запропонованою архітектурою. Вона включає три послідовні згорткові блоки з поступово збільшуваною кількістю фільтрів: Conv2D(16) → MaxPooling → Conv2D(32) → MaxPooling → Conv2D(64) → MaxPooling → Flatten → Dense(64) → Dense(3, softmax). Функція активації – ReLU у прихованих шарах, softmax – у вихідному. Функція втрат – categorical cross-entropy. Оптимізатор – Adam з початковою швидкістю навчання 0.001.

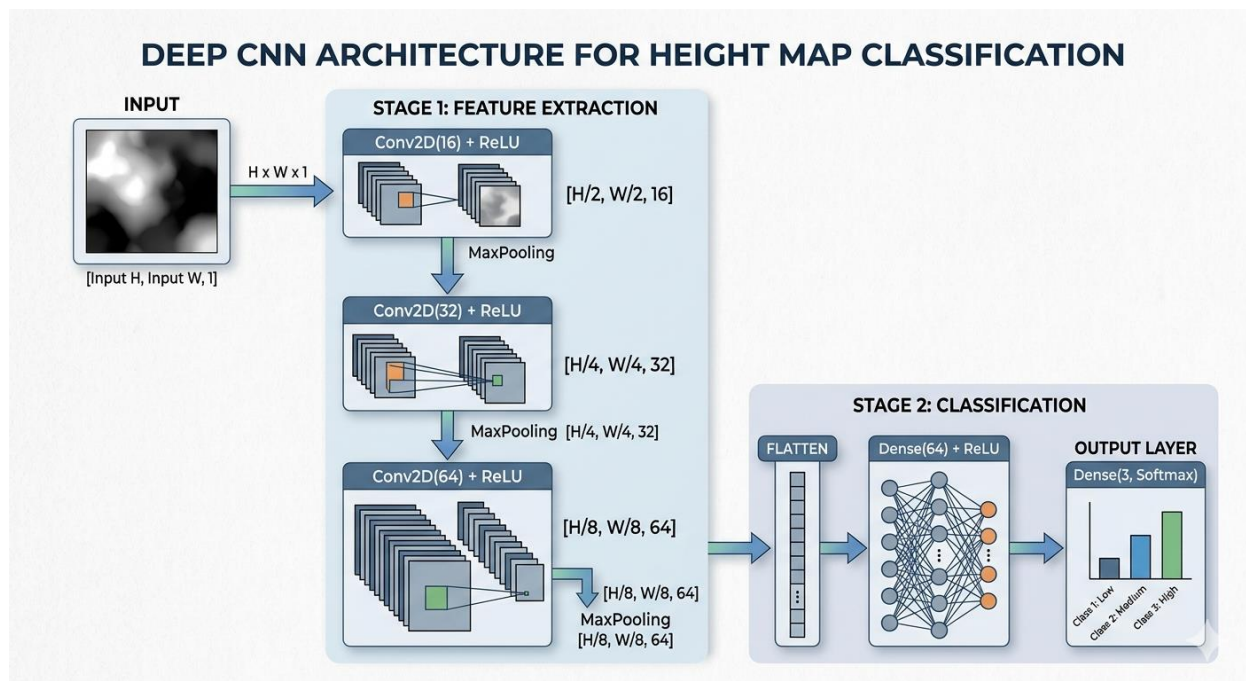


Рисунок 4 – Архітектура поглибленої CNN (Deep CNN) для класифікації карт висот

Поступове збільшення кількості фільтрів від 16 до 64 у наступних шарах відображає загальноприйнятий принцип проектування CNN: ранні шари виявляють прості низькорівневі ознаки (границі, градієнти), тоді як глибші шари навчаються розпізнавати більш складні просторові структури (форми об'єктів, характер поверхні, взаємне розташування пікових областей).

### 6. Результати навчання та порівняльний аналіз

Навчання всіх моделей проводилося з ідентичними умовами: при тестуванні різних архітектур нейронних мереж – розбиття даних 80/20, кількість епох – 20, розмір батчу – 16, оптимізатор Adam, а для тесту поглибленої CNN – розбиття даних 70/30, розмір батчу 32.

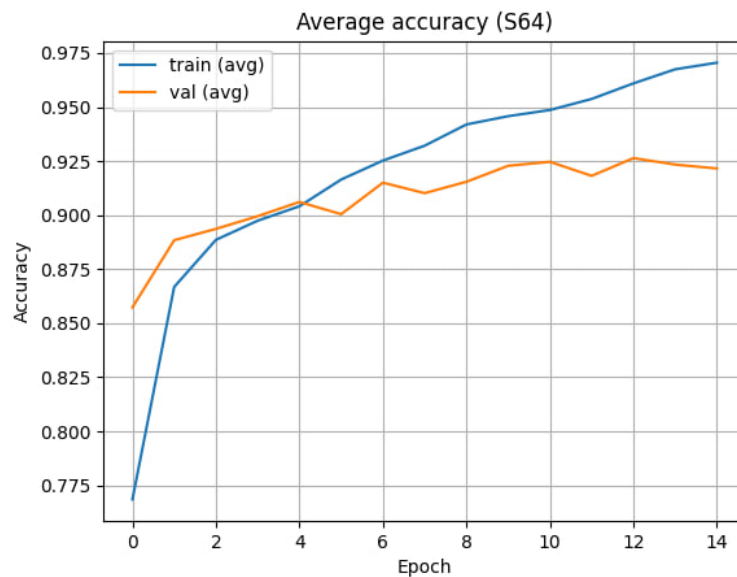


Рисунок 5 – Криві навчання поглибленої CNN: залежність точності та функції втрат від епохи навчання

Для датасету розмірності 64×64 на реальних даних отримано такі результати навчання: MLP – тренувальна точність 0.84, валідаційна точність 0.63; базова CNN – тренувальна точність 0.98, валідаційна точність 0.83; CNN+Dropout – тренувальна точність 0.91, валідаційна точність 0.89; CNN+L1 – тренувальна точність 0.72, валідаційна точність 0.75; поглиблена CNN – тренувальна точність 1.0, валідаційна точність 0.92.

Поглиблена CNN демонструє найвищу точність класифікації (92% на валідаційній вибірці), що на 9% перевищує результати базової CNN та на 29% – результати MLP. Це підтверджує перевагу глибших архітектур для задач просторового аналізу. Разом з тим відзначається, що використання Dropout дозволяє суттєво зменшити різницю між тренувальною та валідаційною точністю (з 8% до 2%), хоча абсолютне значення валідаційної точності при цьому дещо нижче, ніж у Deep CNN.

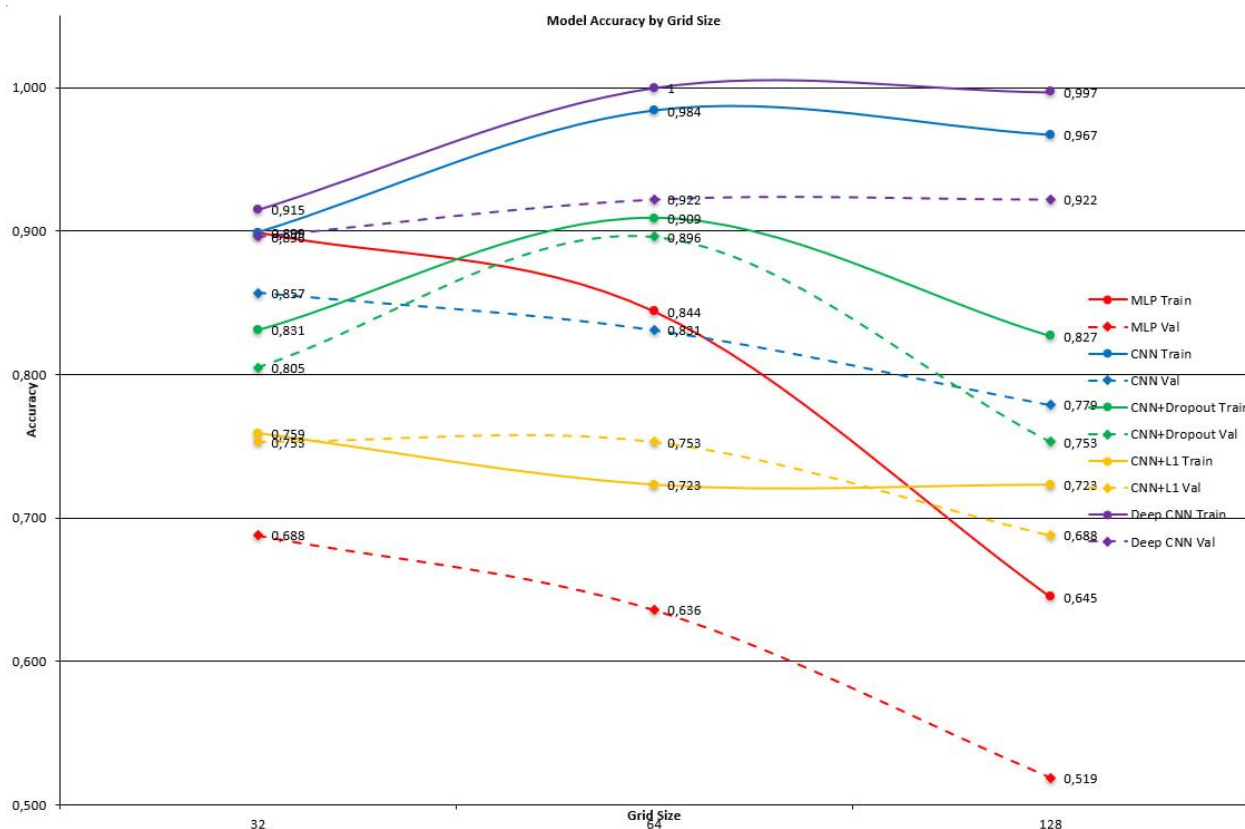


Рисунок 6 – Порівняння точності різних архітектур нейронних мереж на тренувальній та валідаційній вибірці

Аналіз кривих навчання поглибленої CNN виявляє характерну поведінку: швидке зростання точності в перших 8 епохах, поступова стабілізація після 9 епохи, незначні осциляції на пізніх стадіях навчання (що є нормальним явищем для задач із реальними, зашумленими даними). Функція втрат на валідаційній вибірці зменшується стійко протягом усього навчання, хоча дещо повільніше, ніж на тренувальній, що свідчить про помірне перенавчання.

Аналіз помилок класифікації показав, що найчастіше плутаються класи United та Overlapped. Це є логічним результатом, оскільки в граничних випадках (незначне перекриття) відмінність між цими станами є мінімальною навіть для людського спостерігача. Клас Single розпізнається найбільш надійно – точність розпізнавання цього класу перевищує 95% у всіх досліджених архітектурах.

**Висновки.** У роботі розроблено та досліджено підхід до автоматичної класифікації просторових станів об'єктів (одиначний, об'єднаний, накладений) на основі карт висот, сформованих із даних лазерного сканування, із використанням згорткових нейронних мереж. Проведено комплексне порівняльне дослідження п'яти архітектур нейронних мереж та визначено вплив розмірності карти висот на якість класифікації.

За результатами проведених досліджень сформульовано такі висновки:

1. Підхід до перетворення тривимірних хмар точок у двовимірні карти висот шляхом проєктування із збереженням максимального значення висоти є ефективним мето-

дом представлення просторової інформації для задач класифікації. Такі карти висот є інваріантними до оптичних властивостей поверхні та умов освітлення, що є важливою перевагою для промислових застосувань.

2. Поглиблена CNN (3 згорткові блоки з поступово збільшуваною кількістю фільтрів 16→32→64) забезпечує найвищу точність класифікації серед досліджених архітектур: до 100% на тренувальній та до 92% на валідаційній вибірці для реальних даних розмірності 64×64. Це на 29 відсоткових пунктів перевищує результати базового MLP.

3. Розмірність карти висот суттєво впливає на якість класифікації. Оптимальною виявилася розмірність 64×64, що забезпечує найкращий баланс між деталізацією просторової інформації та узагальнювальними властивостями моделі. Збільшення розмірності до 128×128 призводить до зростання перенавчання та зниження валідаційної точності.

4. Реальні дані лазерного сканування характеризуються наявністю дефектів (пропуски значень, шуми, артефакти), що значно ускладнює класифікацію порівняно із синтетичними даними. Аугментація за рахунок ротаційних перетворень (0°, 90°, 180°, 270°) дозволяє покращити узагальнювальні властивості моделі та забезпечити інваріантність до орієнтації об'єктів.

5. Найбільш складним для класифікації є розрізнення класів United та Overlapped у граничних випадках. Це вказує на необхідність або збільшення обсягу навчальних даних для цих граничних станів, або використання додаткових ознак (наприклад, розподіл висот у гістограмі або моменти поверхні).

Отримані результати підтверджують ефективність застосування CNN для аналізу карт висот у задачах прецизійного сенсорного сортування руд та можуть бути використані як основа для розробки реальних систем інтелектуального керування процесом сортування. Напрямами подальших досліджень є: крос-датасетна валідація, дослідження архітектур ResNet та EfficientNet, розробка методів виявлення граничних та аномальних ситуацій, а також оптимізація моделей для роботи в режимі реального часу.

#### ЛІТЕРАТУРА

1. Chalapathy R., Chawla S. Deep Learning for Anomaly Detection: A Survey // ACM Computing Surveys. – 2019. – Vol. 51, No. 5. – P. 1–48. – DOI: 10.1145/3299869. URL: <https://arxiv.org/abs/1901.03407>
2. Gholami A., Kim S., Dong Z., Yao Z., Mahoney M. W., Keutzer K. A Survey of Quantization Methods for Efficient Neural Network Inference // Low-Power Computer Vision. – 2021. – P. 291–326. – DOI: 10.1201/9781003162810-13. – URL: <https://arxiv.org/abs/2103.13630>
3. Parisi G. I., Kemker R., Part J. L., Kanan C., Wermter S. Continual Lifelong Learning with Neural Networks: A Review // Neural Networks. – 2019. – Vol. 113. – P. 54–71. DOI: 10.1016/j.neunet.2019.01.012. – URL: <https://arxiv.org/abs/1802.07569>
4. Peukert S., Heinig T., Tolosana-Delgado R., Bachmann K., Gutzmer J. A Review of Sensor-Based Sorting in Mineral Processing: The Potential and Limitations of Sensing and Data Processing // Minerals. – 2022. – Vol. 12, No. 11. – P. 1364. – DOI: 10.3390/min12111364. – URL: <https://www.mdpi.com/2075-163X/12/11/1364>

5. Qi C. R., Yi L., Su H., Guibas L. J. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space // Proceedings of NeurIPS. – 2017. – P. 5099–5108.

URL: <https://arxiv.org/abs/1706.02413>

6. Rai S., Nistala S. H., Danda A. K., Sreenivas T. Application of Machine Learning Algorithms for Mineral Classification from Hyperspectral Data // Minerals Engineering. – 2021. – Vol. 170. – P. 107054. – DOI: 10.1016/j.mineng.2021.107054.

URL: <https://www.sciencedirect.com/science/article/abs/pii/S0892687521002636>

7. Robben C., Wotruba H. Sensor-Based Ore Sorting Technology in Mining—Past, Present and Future // Minerals. – 2019. – Vol. 9, No. 9. – P. 523. – DOI: 10.3390/min9090523.

URL: <https://www.mdpi.com/2075-163X/9/9/523>

8. Wang M., Deng W. Deep Visual Domain Adaptation: A Survey // Neurocomputing. – 2018. – Vol. 312. – P. 135–153. – DOI: 10.1016/j.neucom.2018.05.083.

URL: <https://arxiv.org/abs/1802.03601>

### REFERENCES

1. Chalapathy R., Chawla S. Deep Learning for Anomaly Detection: A Survey // ACM Computing Surveys. – 2019. – Vol. 51, No. 5. – P. 1–48. – DOI: 10.1145/3299869. – URL: <https://arxiv.org/abs/1901.03407>

2. Gholami A., Kim S., Dong Z., Yao Z., Mahoney M. W., Keutzer K. A Survey of Quantization Methods for Efficient Neural Network Inference // Low-Power Computer Vision. – 2021. – P. 291–326. – DOI: 10.1201/9781003162810-13. – URL: <https://arxiv.org/abs/2103.13630>

3. Parisi G. I., Kemker R., Part J. L., Kanan C., Wermter S. Continual Lifelong Learning with Neural Networks: A Review // Neural Networks. – 2019. – Vol. 113. – P. 54–71. – DOI: 10.1016/j.neunet.2019.01.012. – URL: <https://arxiv.org/abs/1802.07569>

4. Peukert S., Heinig T., Tolosana-Delgado R., Bachmann K., Gutzmer J. A Review of Sensor-Based Sorting in Mineral Processing: The Potential and Limitations of Sensing and Data Processing // Minerals. – 2022. – Vol. 12, No. 11. – P. 1364. – DOI: 10.3390/min12111364.

URL: <https://www.mdpi.com/2075-163X/12/11/1364>

5. Qi C. R., Yi L., Su H., Guibas L. J. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space // Proceedings of NeurIPS. – 2017. – P. 5099–5108.

URL: <https://arxiv.org/abs/1706.02413>

6. Rai S., Nistala S. H., Danda A. K., Sreenivas T. Application of Machine Learning Algorithms for Mineral Classification from Hyperspectral Data // Minerals Engineering. – 2021. – Vol. 170. – P. 107054. – DOI: 10.1016/j.mineng.2021.107054.

URL: <https://www.sciencedirect.com/science/article/abs/pii/S0892687521002636>

7. Robben C., Wotruba H. Sensor-Based Ore Sorting Technology in Mining—Past, Present and Future // Minerals. – 2019. – Vol. 9, No. 9. – P. 523. – DOI: 10.3390/min9090523. – URL: <https://www.mdpi.com/2075-163X/9/9/523>

8. Wang M., Deng W. Deep Visual Domain Adaptation: A Survey // Neurocomputing. – 2018. – Vol. 312. – P. 135–153. – DOI: 10.1016/j.neucom.2018.05.083.

URL: <https://arxiv.org/abs/1802.03601>

Received 07.04.2026.  
Accepted 09.04.2026.  
Published 30.04.2026

***Classification of spatial states of objects based on height maps  
using convulsive neural networks for sensory sorting systems of ores***

*The article considers the problem of automatic classification of spatial states of objects in precision sensory sorting systems of ore raw materials. The relevance of the study is due to the limited effectiveness of traditional segmentation algorithms in conditions of overlapping and approaching pieces of ore on the conveyor. The aim of the work is to develop and study an approach to classifying three spatial states of objects (single, combined, superimposed) based on height maps obtained using laser scanning. Convolutional neural networks (CNN) are used as the research method. A comparative analysis of MLP, basic CNN, CNN with regularization and deep CNN architectures is carried out. The influence of the height map dimension (32×32, 64×64, 128×128) on the classification quality was investigated. It was found that the deep CNN demonstrates the best accuracy on real data – up to 90% on the validation sample with a dimension of 64×64. The effect of overtraining was revealed when the input data dimension increased. The obtained results confirm the promising application of CNN for the analysis of height maps in the tasks of intelligent control of the sorting process.*

*Keywords: sensory sorting of ores, height map, laser scanning, convolutional neural network, image classification, machine learning, precision sorting, point cloud, overtraining, data augmentation.*

**Кондрашов Дмитро Андрійович** – аспірант кафедри автоматизації, комп’ютерних наук і технологій, Криворізький національний університет, Кривий Ріг, Україна.

ORCID: <https://orcid.org/0009-0008-6751-925X>

**Музыка Іван Олегович** – к.т.н., доцент, кафедра комп’ютерних систем та мереж, Криворізький національний університет, Кривий Ріг, Україна.

ORCID: <https://orcid.org/0000-0002-9202-2973>

**Купін Андрій Іванович** – д.т.н., професор, кафедра комп’ютерних систем та мереж, Криворізький національний університет, Кривий Ріг, Україна.

ORCID: <http://orcid.org/0000-0001-7569-1721>

**Балик Дмитро Костянтинівич** – інженер-програміст, ТОВ «НВП Гамаюн».

ORCID: <http://orcid.org/0009-0000-4768-8576>

**Kondrashov Dmytro** – PhD Student, Department of Automation, Computer Sciences and Technology, Kryvyi Rih National University, Ukraine.

ORCID: <https://orcid.org/0009-0008-6751-925X>

**Muzyka Ivan** – Associate Professor, Department of Computer Systems and Networks, Kryvyi Rih National University, Kryvyi Rih, Ukraine.

ORCID: <https://orcid.org/0000-0002-9202-2973>

**Kupin Andrii** – Doctor of Technical Sciences, Professor, Department of Computer Systems and Networks, Kryvyi Rih National University, Kryvyi Rih, Ukraine.

ORCID: <http://orcid.org/0000-0001-7569-1721>

**Balyk Dmytro** – “SPE Hamaiun” LLC.

ORCID: <http://orcid.org/0009-0000-4768-8576>

## МЕТОД ТА ПРОГРАМНИЙ ЗАСІБ РОЗПІЗНАВАННЯ ТА АНАЛІЗУ ОБ'ЄКТІВ НА ВІДЕО

*Анотація.* Актуальність дослідження зумовлена стрімким розвитком технологій штучного інтелекту та комп'ютерного зору, що дозволяють автоматизувати процеси аналізу великих обсягів відеоданих у реальному часі. Однак існує вагома проблема: сучасні алгоритми генерують значні масиви «сирих» неструктурованих даних, які не мають єдиної багатовимірної структури, що суттєво ускладнює їх гнучкий аналіз та агрегування. Метою роботи є підвищення швидкості аналізу надвеликих масивів відеоданих шляхом розробки методу автоматизованої екстракції структурованих даних та їхньої інтеграції у багатовимірну OLAP-модель. Основними методами дослідження є каскадна фільтрація з використанням неймережі YOLOv26m для швидкої детекції об'єктів, застосування мультимодальної VLM-моделі MoonDreem2 для семантичного збагачення опису сцен, а також побудова ROLAP-моделі у межах розподіленої мікро-сервісної архітектури. У результаті експериментів доведено високу точність детекції об'єктів ( $mAP50 = 0.814$  при швидкості  $81.03$  FPS), збалансовану швидкодію семантичного аналізу ( $0.422$  с/кадр) та здатність системи до лінійного прискорення при масштабуванні. Розроблений комплексний інструмент успішно трансформує великі обсяги неструктурованих відеоданих у формат, придатний для глибокого багатовимірного аналізу та ефективного прийняття рішень у високонавантажених системах відеомоніторингу.

*Ключові слова:* комп'ютерний зір, відеоаналітика, виявлення об'єктів, відстеження об'єктів, багатовимірна модель, OLAP, YOLO, VLM, військова техніка.

**Постановка проблеми.** У сучасному світі однією з найважливіших галузей використання ШІ є комп'ютерний зір, що дозволяє автоматизувати процеси аналізу зображень та відео, включаючи задачі виявлення об'єктів. Проте відео, як джерело даних, є складним для обробки, оскільки містить велику кількість інформації у часовій послідовності кадрів.

Проблема полягає у недостатній ефективності та масштабованості наявних підходів до перетворення великих неструктурованих відеоданих (результатів виявлення та відстеження об'єктів) у форму, придатну для глибокого аналітичного аналізу та прийняття рішень відеоаналітиками в реальному часі. Сучасні методи комп'ютерного зору генерують значні обсяги "сирих" даних, які складно агрегувати за часом, територією чи типом об'єкта, і вони не мають єдиної багатовимірної структури, що обмежує можливо-

сті гнучкого виконання аналітичних операцій. Тому необхідно розробити та дослідити ефективність методу та програмного засобу, що забезпечують безперервний процес екстракції та семантичного збагачення даних з відеопотоку, збереження отриманих даних у структурі, яка підтримує складні аналітичні операції.

**Аналіз останніх досліджень і публікацій.** Задача виявлення, відстеження та аналізу об'єктів у відеопотоці є однією з ключових у галузі комп'ютерного зору та активно досліджується протягом останніх років.

У дослідженні[1] розглянуто класичні методи розпізнавання об'єктів на основі згорткових нейронних мереж (CNN), які дозволяють ефективно виділяти ознаки зображень. Проте такі підходи, особливо двоетапні моделі, наприклад, Faster R-CNN, мають недостатню швидкодію для задач реального часу. Це обмежує їх застосування у високонавантажених системах відеоаналізу.

Альтернативою є одноетапні методи детекції, зокрема алгоритм YOLO[2]. Його ключовою перевагою є можливість одночасного визначення координат об'єктів та їх класів за один прохід нейронної мережі, що забезпечує високу швидкість обробки. Це робить YOLO одним із найбільш популярних рішень для задач відеоаналітики в реальному часі.

У роботі[3] досліджуються сучасні підходи до відстеження об'єктів, зокрема парадигма Tracking-by-Detection, яка поєднує детекцію та подальшу асоціацію об'єктів між кадрами. Розглянуто алгоритми SORT, DeepSORT та ByteTrack. Автор відзначає, що використання додаткових ознак зовнішнього вигляду підвищує точність трекінгу, проте збільшує обчислювальні витрати. Водночас метод ByteTrack демонструє високу ефективність за рахунок оптимізованої асоціації навіть низьковпевнених детекцій.

Використання мультимодальних моделей для семантичного аналізу зображень[4] дозволяє отримувати більш глибоке розуміння сцени, включаючи опис середовища, взаємодію об'єктів та контекст подій. Однак їх використання в потокових системах обмежується високою обчислювальною складністю.

Окремим напрямом досліджень є аналіз та агрегування результатів відеообробки, зокрема із застосуванням OLAP-технологій для багатовимірного аналізу даних[5, 6]. На основі цих технологій розроблено спеціалізовані моделі представлення відеоданих, такі як SurvCube та VideoCube, які інтегрують результати відеоаналізу у багатовимірні структури. Незважаючи на ефективність запропонованих рішень, вони мають обмеження, пов'язані з недостатньою гнучкістю, відсутністю механізмів автоматичного оновлення семантичних класів, а також низьким рівнем семантичного формалізму, тобто відсутністю ієрархій.

Таким чином, аналіз існуючих досліджень показав, що сучасні методи забезпечують високу точність виявлення та відстеження об'єктів, однак проблема інтеграції результатів у зручну для аналізу форму залишається актуальною. Це обґрунтовує доцільність розробки підходу, який поєднує методи комп'ютерного зору з багатовимірними моделями даних для забезпечення ефективної аналітики відеопотоків.

**Мета дослідження.** Метою дослідження є підвищення швидкості аналізу надвеликих масивів відеоданих шляхом розробки методу та програмного засобу автоматизо-

ваної екстракції структурованих даних та їхньої інтеграції у багатовимірну OLAP-модель, що підтримує паралельну обробку та семантичне збагачення інформації.

**Основна частина.** Процес збору та обробки даних у запропонованій системі базується на принципах каскадної фільтрації та семантичного збагачення інформації. Система забезпечує паралельну обробку множинних вхідних відеопотоків, трансформуючи неструктуровані візуальні дані у структуровані записи для їх подальшої інтеграції в аналітичну OLAP-модель. На рисунку 1 наведено загальну схему розробленого методу.

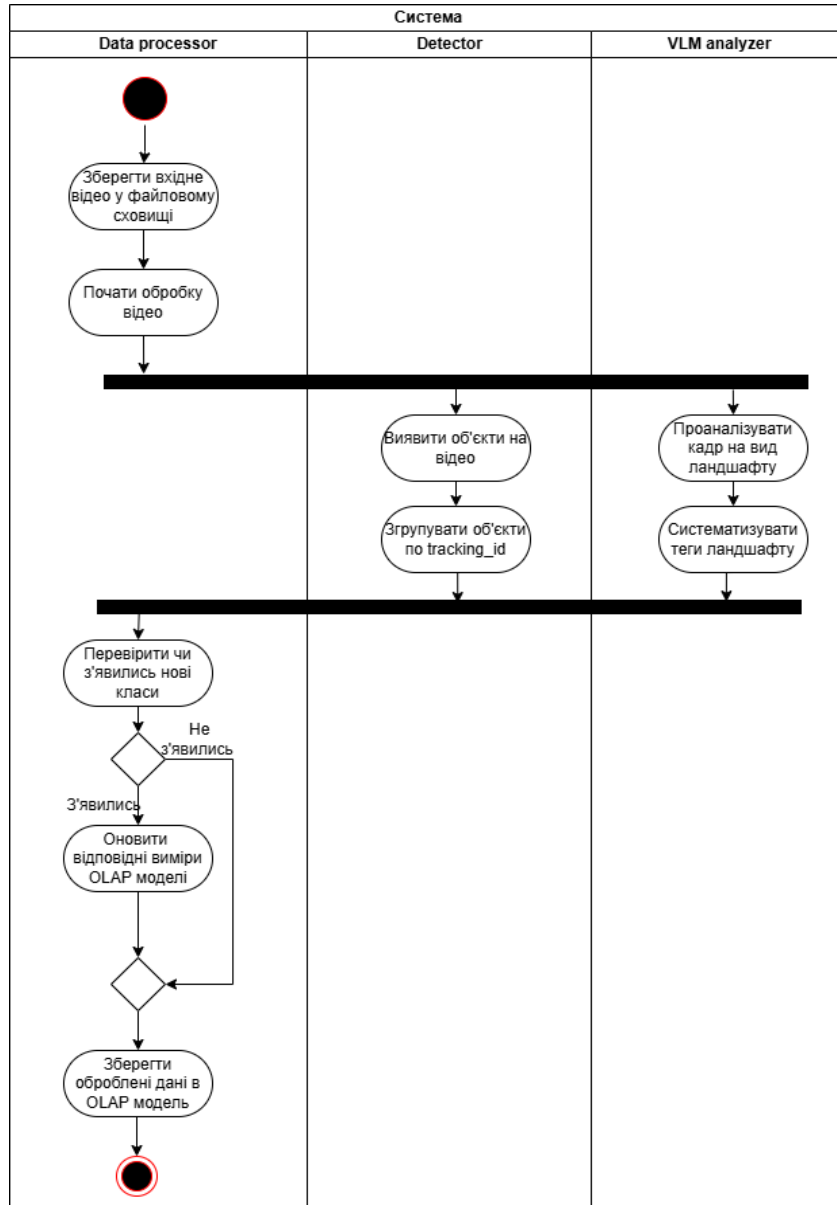


Рисунок 1 – Загальна схема розробленого методу розпізнавання та аналізу об’єктів на відео

*Попередня обробка*

Вхідним сигналом для системи є сукупність відеопотоків від розподілених джерел, наприклад, IP-камер та камер з дронів тощо. На цьому етапі вирішуються два ключові завдання:

- екстракція метаданих: з відеопотоку зчитуються часова мітка геокоординати, роздільна здатність, тривалість і частота кадрів;
- оптимізації навантаження: з метою забезпечення швидкодії в умовах високо навантаженої системи, аналіз проводиться не для кожного кадру, а використовується метод проріджування кадрів, де для детекції обирається кожен n-й кадр, наприклад, 2-5 кадрів на секунду, що дозволяє суттєво зменшити навантаження на обчислювальні потужності без втрати точності моніторингу динамічних об'єктів

*Детекція, класифікація та трекінг об'єктів*

На етапі виявлення об'єктів ключові кадри передаються до нейромережі YOLO, яка виконує функцію первинного фільтра та класифікатора. У разі відсутності цільових об'єктів на кадрі процес обробки завершується автоматично. При успішній детекції модель повертає параметри домену та класу об'єкта, координати обмежувальної рамки та рівень впевненості. Для підвищення достовірності результатів дані, що не досягають встановленого порогу впевненості, відсікаються.

Важливою частиною цього етапу є механізм об'єктного трекінгу. Кожному виявленому об'єкту присвоюється унікальний ідентифікатор `object_id`, що забезпечує унікальність даних та дозволяє відстежувати переміщення об'єкта в межах поля зору камери, мінімізуючи надлишковість обчислень та обсяг збережених даних.

*Семантичне збагачення даних за допомогою мультимодальних моделей*

Паралельно з детекцією відбувається семантичне збагачення даних за допомогою мультимодальної моделі Moondream2, яка швидко генерує наповнені відповіді. Процес аналізу поєднує візуальний вхід (кадр або область об'єкта) та текстовий запит-постановку задачі. Це забезпечує гнучкість системи, дозволяючи адаптувати моніторинг під різні сценарії (наприклад, тип локації, погодні умови, супутні об'єкти) лише зміною тексту запиту, без необхідності перенавчання нейромережі. На відміну від YOLO, модель VLM виконує складні когнітивні завдання, аналізуючи взаємозв'язки між об'єктами. Оскільки цей процес ресурсомісткий, його застосовують лише до найбільш репрезентативних кадрів. Отримана структурована відповідь дозволяє автоматично виділяти нові типи середовищ та подій для додавання до багатовимірної моделі.

Для отримання структурованої відповіді, емпірично було підібрано наступний промпт до VLM моделі: "Describe landscape, structures and people". З отриманої відповіді можна виділити тип середовища, об'єкти та події, зображені на кадрі. Якщо було отримано новий тип, він буде доданий до багатовимірної моделі.

*Багатовимірна модель представлення та аналізу даних*

Для систематизації та збереження інформації використано багатовимірну модель Кодда, яка орієнтована на аналітичні запити та агрегування даних у вигляді інформаційних кубів. На відміну від класичних реляційних моделей, цей підхід дозволяє користувачеві розглядати дані одночасно у кількох вимірах.

Математично багатовимірний куб даних у запропонованій системі можна виразити як функцію, що відображає декартовий добуток множин вимірів у простір числових мір:

$$F: T \times L \times O \times E \times ET \rightarrow R^m, \quad (1)$$

де  $m_i \in R$  – числові міри, наприклад, кількість виявлених об'єктів

У межах дослідження ієрархічну структуру куба формують такі виміри:

- Час (T) – агрегування результатів детектування за ієрархією «день – місяць – рік»;
- Простір – агрегування результатів детектування за ієрархією «населений пункт – область»
- Тип об'єкту – класифікація за структурою «домен – клас»
- Оточення – опис середовища за ієрархією «домен – тип»
- Тип події – класифікація подій на відео за структурою «домен – тип»

Для проведення аналітичних досліджень над моделлю визначено базові OLAP операції, які дозволяють отримувати специфічні зрізи даних. Зокрема, операція Slice фіксує значення одного виміру та формує відповідну підмножину даних. Наприклад, вибір даних для певного дня  $d \in T_1$  визначається як:

$$F_d = \{f \in F \mid h_{(0,1)}^T(\pi_T(f)) = d\}, \quad (2)$$

де  $\pi_T(f)$  – проекція на вимір часу. Операція Dice дозволяє виконувати більш складну селекцію за декількома умовами одночасно. Наприклад, для обчислення кількості військових об'єктів у визначених областях за конкретний місяць (де  $t_1 \in T_1, l_1 \in L_1, domain = \text{"військовий"}$ ), агрегована міра виражається наступним відношенням:

$$M = \text{card}\{f \in F \mid h_{(0,1)}^T(\pi_T(f)) = t_1 \wedge h_{(0,1)}^L(\pi_L(f)) = l_1 \wedge \pi_O(f) = \text{військовий}\} \quad (3)$$

Логічна структура системи реалізована через три основні фактові таблиці: Object\_Tracking\_Fact для обліку унікальних об'єктів, Video\_Environment\_Fact для фіксації типів середовища та Video\_Event\_Fact для збереження даних про події.

Логічна структура системи реалізована через три основні фактові таблиці: Object\_Tracking\_Fact для обліку унікальних об'єктів, Video\_Environment\_Fact для фіксації типів середовища та Video\_Event\_Fact для збереження даних про події.

Основними аналітичними мірами моделі є кількість виявлених об'єктів (на основі object\_id), що вказує на інтенсивність руху, та тривалість перебування об'єкта в зоні спостереження. Використання такої сукупності мір дозволяє проводити поглиблений аналіз динаміки подій, поєднуючи дані про ідентифіковані об'єкти з детальним описом умов, у яких вони були зафіксовані.

### Розробка програмного забезпечення

#### Архітектура та функціональні компоненти системи

Запропонований метод покладено в основу програмного комплексу, архітектура якого наведена на рис. 2. Вона базується на мікросервісному підході з кластерними обчисленнями, що забезпечує високу доступність та масштабування. Система поділена на

рівні аналізу, координації та збереження даних. Центральний елемент – розподілена мережа вузлів, об'єднаних у кластери:

– Detection Cluster. Ідентифікація об'єктів у реальному часі за допомогою моделі YOLO, з паралельним розподілом навантаження.

– Semantic Analyze Cluster: Глибинний семантичний аналіз сцени за допомогою мультимодальних моделей (VLM), що формують описи контексту.

Для незалежності та гарантованої доставки результатів впроваджено брокер повідомлень (MessageBroker). Асинхронний обмін даними за протоколом AMQP дозволяє накопичувати завдання в чергах, що критично при пікових навантаженнях.

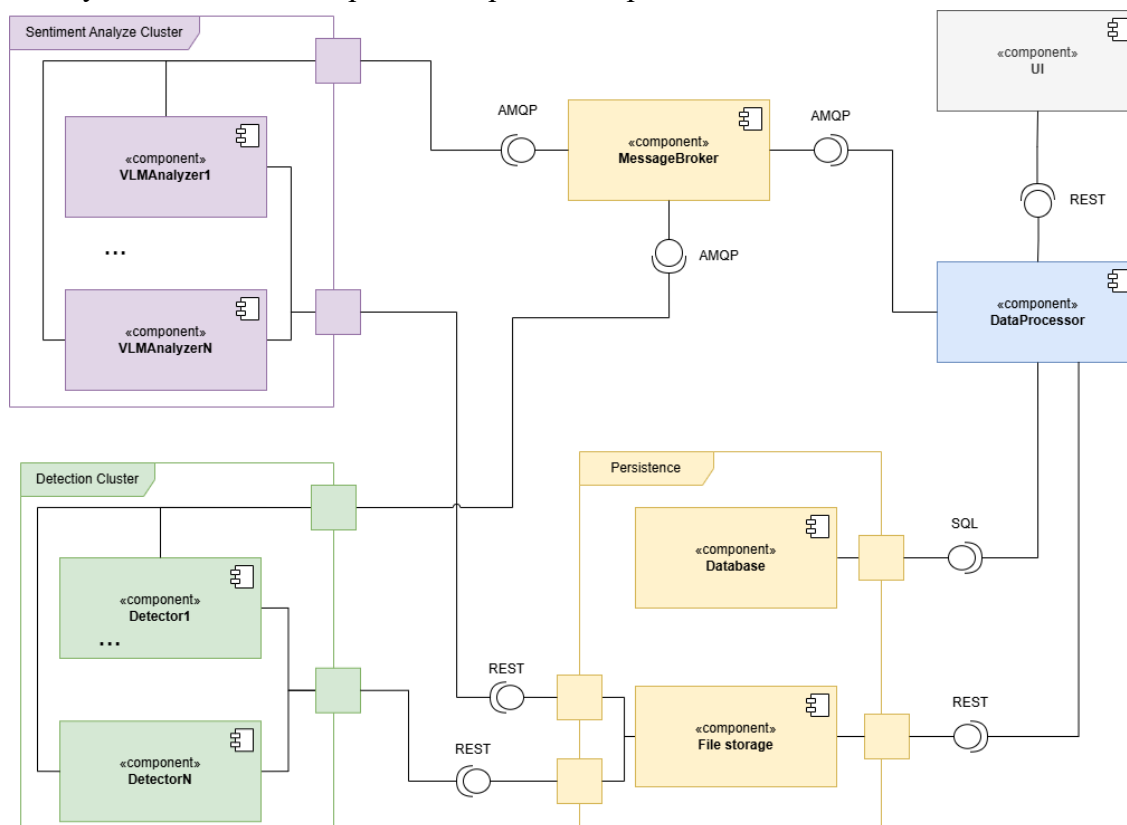


Рисунок 2 – Схема архітектури розробленого програмного забезпечення

### *Координація та збереження даних*

Роль координатора робочих процесів виконує компонент Data Processor. Він агрегує результати з брокера повідомлень, виконує їх фінальну структурування та забезпечує наповнення аналітичного сховища. Взаємодія з клієнтським інтерфейсом та отримання медіафайлів реалізовані через REST API.

Система використовує гібридну модель збереження даних для розділення “важкого” контенту та структурованої аналітики: файлове сховище (MinIO) для збереження об’ємних відеофайлів та кадрів та реляційну базу даних (PostgreSQL) зі сховищем даних ROLAP за схемою «зірка».

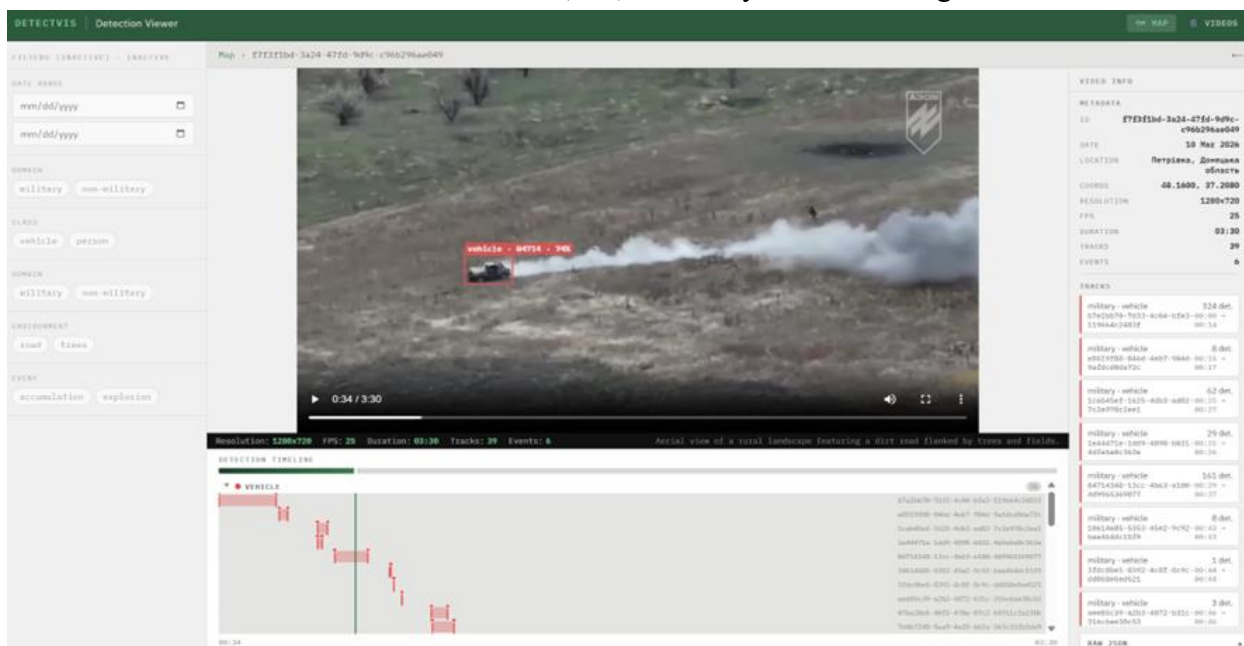


Рисунок 3 – Приклад відображення зібраної інформації про відео

**Результати експериментальних досліджень.** Для оцінки ефективності запропонованої системи було проведено серію експериментів, що охоплює ключові аспекти продуктивності та ресурсоемності. Дослідження фокусуються на чотирьох основних метриках: швидкості та точності роботи моделі YOLO залежно від розміру моделі, продуктивності VLM-моделей з урахуванням запитів і відповідей, загальній швидкості всього алгоритму обробки, а також обсязі даних, що генеруються при інтеграції відео в OLAP-модель. Тестування проводилося на апаратній платформі з GPU NVIDIA RTX 4070, 8 ГБ RAM та процесором Intel Core i7, з використанням підготовленого датасету військової техніки та відео з дронів.

#### *Швидкість і точність YOLO залежно від розміру моделі*

Для визначення найбільш ефективної конфігурації нейромережі проведено порівняльний аналіз моделей сімейства YOLOv26 варіантів Nano, Medium та Large. Експерименти проводилися на відео роздільною здатністю 1080p з частотою 30 FPS. Результати вимірювань, наведені у таблиці 1, свідчать, що найбільш раціональним рішенням для поставленої задачі є модель YOLOv26m. Дана конфігурація забезпечує високий рівень точності (map-50 = 0.814, precision = 0.869) при збереженні продуктивності на рівні 81.03 FPS. Таке співвідношення характеристик дозволяє системі виконувати стабільну детекцію та відстеження об'єктів у режимі, близькому до реального часу, мінімізуючи затримки при обробці динамічних сцен.

#### *Оцінка точності та швидкодії VLM моделей*

Для визначення найефективнішого рішення для задачі генерації описів до зображень було проведено порівняльне експериментальне дослідження чотирьох сучасних візуально-мовних моделей (VLM): Moondream2, Qwen2.5, PaliGemma та BLIP-2. Експерименти виконувалися з використанням обчислювальних потужностей графічного прискорювача NVIDIA A100. Оцінювання проводилося на випадковій підвибірці з 200 зо-

бражень зі спеціалізованого датасету RSICD (Remote Sensing Image Captioning Dataset). Кількісна оцінка якості згенерованих текстів здійснювалася за допомогою стандартних метрик: BLEU-1, METEOR, Jaccard Index, а також CIDEr. Швидкодія моделей вимірювалася як середній час генерації опису для одного зображення. Усі метрики якості наведено у масштабі від 0 до 100. Зведені результати експериментів представлені у Таблиці 2.

Таблиця 1

Порівняння швидкості та точності моделей YOLO

Модель	Кількість епох	map50	map50-95	Precision	Recall	FPS
YOLOv26n	20	0.794	0.423	0.819	0.72	98.03
YOLOv26m	20	0.814	0.448	0.869	0.723	81.03
YOLOv26l	30	0.821	0.4682	0.871	0.741	66.31

Таблиця 2

Порівняння швидкості та точності VLM моделей

Модель	BLEU-1	METEOR	Jaccard Index	CIDEr	Швидкість обробки, (с)
paligemma-3b-pt-224	32.48	18.27	18.84	11.54	0.321
blip2-opt-2.7b	17.91	18.6	18	10.34	1.318
qwen2.5-VL-3B-Instruct	45.92	24.22	23.15	14.94	0.851
moondream2	38.01	25.73	21.6	12.55	0.422

Аналіз отриманих даних показує, що моделі Qwen2.5 та Moondream2 є беззаперечними лідерами за якістю генерації тексту. Хоча Qwen2.5 демонструє найвищі показники точного збігу слів (BLEU-1: 45.92, Jaccard: 23.15) та специфічності (CIDEr: 14.94), архітектура Moondream2 виявляється найкращою та найбільш збалансованою моделлю для практичного застосування. По-перше, Moondream2 демонструє найвищі результати за метрикою METEOR (25.73). Це свідчить про те, що модель краще зберігає загальний семантичний зміст та структуру речень, навіть якщо використовує синоніми. По-друге, Moondream2 працює більш ніж удвічі швидше за Qwen2.5 (0.422 секунди на зображення проти 0.851), поступаючись у швидкості лише PaliGemma, яка при цьому має значно гіршу якість генерації.

Варто також відзначити важливий якісний аспект генерації, який складно повною мірою відобразити кількісними метриками. Моделі Moondream2 та Qwen2.5 схильні генерувати дуже детальні, розгорнуті описи, які містять глибокий аналіз контексту зо-

браження. Натомість еталонні описи у датасеті RSICD здебільшого є короткими та лаконічними. У результаті класичні метрики (такі як BLEU та CIDEr) "штрафують" ці моделі за генерацію додаткових деталей, яких немає в еталоні. На відміну від PaliGemma та BLIP-2, які генерують простіші конструкції, Moondream2 надає значно багатший контекст, що робить її найефективнішим вибором для задач, де потрібне глибоке візуальне розуміння сцени при збереженні високої швидкості обробки даних.

*Оцінка масштабованості та продуктивності системи*

Для перевірки здатності розробленої архітектури ефективно розподіляти навантаження та працювати з великими масивами даних було проведено тестування масштабованості. Експеримент виконувався з використанням обчислювальних потужностей графічного прискорювача NVIDIA RTX 4070.

Тестовий набір даних складався з 100 відеороликів тривалістю 210 секунд кожен. Базова кадрова частота відео становила 25 FPS, проте для оптимізації обчислювальних ресурсів без втрати інформативності з відеопотоку вилучався кожен 5-й кадр.

Оцінювання проводилося для двох ключових етапів пайплайну: детекції та семантичного аналізу. Щоб виміряти ефективність горизонтального масштабування системи, обробка імітувалася з використанням 1, 5 та 10 паралельних сервісів. Зведені результати експерименту наведено у таблиці 3.

Таблиця 3

Залежність часу обробки від кількості підключених сервісів

Кількість сервісів	Час обробки 1 відео в середньому, с		Час обробки 100 відео, с	
	Детекція	Семантичний аналіз	Детекція	Семантичний аналіз
1	36.5	3.9	4351.2	397.3
5	35.34	3.78	935.74	85.64
10	36.45	3.86	486.16	45.93

Аналіз загального часу обробки масиву зі 100 відео демонструє високу здатність системи до масштабування. Збільшення пулу обробників з 1 до 5 дозволило скоротити загальний час детекції з 4351.2 с до 935.74 с, тобто прискорення у 4.65 разів. При підключенні 10 паралельних сервісів загальний час детекції впав до 486.16 с, що забезпечило прискорення у 8.95 разів порівняно з одним сервісом. Аналогічна тенденція спостерігається і для етапу семантичного аналізу, де розпаралелювання на 10 воркерів зменшило час обробки зі 397.3 с до 45.93 с.

Незначне відхилення показників від ідеального математичного лінійного прискорення, наприклад, 486.16 с замість теоретичних 435.12 с для 10 сервісів, є очікуваним. Воно пояснюється накладними витратами на диспетчеризацію завдань, мережеву взаємодію між мікросервісами, черги повідомлень та операції читання і запису відеофайлів.

Результати підтверджують, що запропонована архітектура не має внутрішніх блокувань при паралельній роботі і дозволяє лінійно нарощувати пропускну здатність системи шляхом додавання нових обчислювальних вузлів.

**Висновки.** У роботі розроблено метод та масштабований програмний засіб для автоматизованої екстракції, семантичного збагачення та багатовимірного аналізу відеоданих. Наукова новизна роботи полягає у поєднанні каскадної обробки відеопотоків із використанням моделі YOLOv26m та мультимодальної моделі Moondream2 з ROLAP-поданням даних у багатовимірному кубі, що забезпечує безперервний перехід від неструктурованого відео до аналітичних зрізів у режимі, близькому до реального часу. Спроектowana мікросервісна архітектура з використанням брокера повідомлень забезпечує паралельну обробку та відмовостійкість системи, що підтверджено серією експериментальних досліджень.

У частині детекції та відстеження об'єктів (зокрема, військової техніки) оптимальною визначено конфігурацію нейромережі YOLOv26m. Вона забезпечує високу точність ( $mAP50 = 0.814$ ) при стабільній роботі в реальному часі (81.03 FPS). Для етапу семантичного збагачення найкращий баланс аналізу та швидкості продемонструвала мультимодальна модель Moondream2 (METEOR: 25.73, 0.422 с/кадр), яка генерує детальні просторові описи без необхідності донавчання базисної архітектури.

Оцінка продуктивності програмного комплексу довела його здатність до лінійного прискорення при масштабуванні. Розширення пулу обчислювальних воркерів з 1 до 10 дозволило скоротити загальний час обробки масиву відео в 8.95 разів (з 4351.2 с до 486.16 с), що підтверджує відсутність внутрішніх блокувань у системі обміну повідомленнями.

Створено комплексний інструмент для високонавантажених систем відеомоніторингу, який успішно перетворює великі обсяги неструктурованих відеоданих у формат, придатний для гнучкого багатовимірного аналізу та швидкого прийняття рішень.

#### ЛІТЕРАТУРА

1. Mahmud A., Setiawan A. A. A. A survey of convolutional neural networks in object detection // *Int. J. Adv. Comput. Sci. Appl.* – 2021.
2. Redmon J., Divvala S., Girshick R., Farhadi A. You Only Look Once: Unified, Real-Time Object Detection // *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. – 2016.
3. Abouelyazid M. Comparative Evaluation of SORT, DeepSORT, and ByteTrack for Multiple Object Tracking in Highway Videos // *International Journal of Sustainable Infrastructure for Cities and Societies*. – 2023. – Vol. 8, No. 11. – P. 42–52.
4. Din M. U., Akram W., Bakht A. B., Hussain I. LLM-VLM Fusion Framework for Autonomous Maritime Port Inspection using a Heterogeneous UAV-USV System. – Текст. – arXiv:2601.13096 [cs.RO]. – 2026. – DOI: 10.48550/arXiv.2601.13096.
5. Hansung L., Sohee P., Jang-Hee Y. A Data Cube Model for Surveillance Video Indexing and Retrieval // *SIGMAP 2013: International Conference on Signal Processing and Multimedia Applications*. – 2013. – URL: <https://www.scitepress.org/Papers/2013/46121/46121.pdf>.

6. Wu Y., Zhang C., Lu Y., Su Y., Jiang X., Xiang Z., Li Z. VideoARD: An Analysis-Ready Multi-Level Data Model for Remote Sensing Video // *Remote Sens.* – 2025. – Vol. 17, No. 22. – Art. 3746. – DOI: 10.3390/rs17223746.

#### REFERENCES

1. Mahmud, A., & Setiawan, A. A. A. (2021). A survey of convolutional neural networks in object detection. *Int. J. Adv. Comput. Sci. Appl.*
2. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
3. Abouelyazid, M. (2023). Comparative evaluation of SORT, DeepSORT, and ByteTrack for multiple object tracking in highway videos. *International Journal of Sustainable Infrastructure for Cities and Societies*, 8(11), 42–52.
4. Din, M. U., Akram, W., Bakht, A. B., & Hussain, I. (2026). *LLM-VLM fusion framework for autonomous maritime port inspection using a heterogeneous UAV-USV system*. arXiv. <https://doi.org/10.48550/arXiv.2601.13096>
5. Hansung, L., Sohee, P., & Jang-Hee, Y. (2013). A data cube model for surveillance video indexing and retrieval. In *SIGMAP 2013: International Conference on Signal Processing and Multimedia Applications*. <https://www.scitepress.org/Papers/2013/46121/46121.pdf>
6. Wu, Y., Zhang, C., Lu, Y., Su, Y., Jiang, X., Xiang, Z., & Li, Z. (2025). VideoARD: An analysis-ready multi-level data model for remote sensing video. *Remote Sens.*, 17(22), Art. 3746. <https://doi.org/10.3390/rs17223746>

Received 07.04.2026.  
Accepted 10.04.2026.  
Published 30.04.2026

#### ***A method and software tool for video object recognition and analysis***

*The modern scientific community is paying significant attention to the automation of video analytics. Traditional methods based on convolutional neural networks (CNNs), such as Faster R-CNN, demonstrate high accuracy but have limited processing speed for real-time streams. In contrast, single-stage algorithms in the YOLO family achieve the required performance. Research in object tracking (Tracking-by-Detection) highlights DeepSORT and ByteTrack as the most effective algorithms for associating detections across frames. The application of multimodal vision-language models (VLMs) opens up new possibilities for semantic scene description, although their implementation in high-load systems is hindered by computational complexity. Additionally, multidimensional data analysis (OLAP) technologies are considered, specifically the SurvCube and VideoCube models, which integrate video processing results into structured cubes; however, they often offer limited flexibility for creating new semantic hierarchies.*

*The objective of the research is to increase the analysis speed of ultra-large video datasets by developing a method and software tool for the automated extraction of structured data and its subsequent integration into a multidimensional OLAP model that supports parallel processing and semantic information enrichment.*

*A method for cascaded video data processing is proposed, combining rapid object detection with the YOLOv26m architecture and dynamic semantic scene enrichment using the*

*Moondream2 multimodal model. During the preprocessing stage, metadata extraction (coordinates, time) and load optimization are performed by downsampling the frame rate to 2–5 frames per second. The YOLOv26m model is used for object detection and classification, achieving an mAP50 of 0.814 at 81.03 FPS. Tracking is implemented by assigning a unique object\_id, which minimizes data redundancy. Semantic context enrichment (landscape type, events) is carried out by the multimodal Moondream2 model using text prompts, allowing the system to adapt to new scenarios without retraining the network. The data is integrated into a multidimensional OLAP model using a star schema, where the dimensions are time, space, object type, environment, and event type. The software features a microservices architecture utilizing a message broker (RabbitMQ/AMQP) for asynchronous communication between the detection and semantic analysis clusters.*

*In the course of this work, a scalable software complex was developed that successfully transforms unstructured video streams into the format of analytical cubes. It was experimentally confirmed that the use of the Moondream2 model provides the best balance between description quality (METEOR: 25.73) and processing speed (0.422 s/frame). Scalability testing demonstrated the architecture's capability for linear speedup: increasing the number of parallel workers to 10 reduced the total video dataset processing time by a factor of 8.95. The proposed solution is effective for high-load monitoring systems and rapid decision-making based on deep analytics.*

*Keywords: computer vision, video analytics, object detection, object tracking, multidimensional model, OLAP, YOLO, VLM, Moondream2, semantic enrichment.*

**Ніколаєв Іван Романович** – магістрант кафедри інформатики та програмної інженерії Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського».

ORCID: <https://orcid.org/0009-0004-7779-8329>

**Олійник Юрій Олександрович** – доцент кафедри інформатики та програмної інженерії Національного технічного університету України «КПІ ім. Ігоря Сікорського».

ORCID: <https://orcid.org/0000-0002-7408-4927>

**Nikolaiev Ivan** - Master's student at Department of Computer Science and Software Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

ORCID: <https://orcid.org/0009-0004-7779-8329>

**Oliinyk Yurii** - Ph.D, associate professor at department of computer science and software engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

ORCID: <https://orcid.org/0000-0002-7408-4927>

## КЛАСИФІКАЦІЯ СТАНУ ОЧЕЙ НА ОСНОВІ ДЕКОМПОЗИЦІЇ ЕЛЕКТРОЕНЦЕФАЛОГРАМ З УРАХУВАННЯМ ДИПОЛЬНОЇ ПРИРОДИ СИГНАЛІВ МОЗКУ

*Анотація.* Запропоновано підхід до класифікації стану очей (відкриті/закриті) за 64-канальними ЕЕГ-даними на основі ІСА-декомпозиції з валідацією компонент підбором еквівалентних диполів. Для підбору диполя розроблено оригінальну функцію втрат із L2-регуляризацією та еліпсоїдним анатомічним обмеженням на основі тривимірної моделі мозку. Серед ІСА-компонент ідентифіковано джерела з характерними альфа-, му-, бета- та гамма-ритмами, що підтверджує фізіологічну інтерпретованість методу. Класифікація нейронною мережею на основі 507 ознак досягла точності 97% на навчальній та 90% на тестовій вибірці.

*Ключові слова:* ЕЕГ, електроенцефалограма, ІСА, диполь, класифікація станів мозку, нейронна мережа, спектральний аналіз, декомпозиція сигналів, сліпе розділення сигналів, локалізація джерел, обчислювальна нейронаука, просторово-часова динаміка біокібернетика, BCI, NeuroAnalyzer, Julia.

**Постановка проблеми.** Класифікація станів мозкової активності за даними електроенцефалографії (ЕЕГ) є важливим завданням як для фундаментальних нейронаук, так і для практичних застосувань у сфері нейроінтерфейсів. Зокрема, розрізнення станів відкритих та закритих очей дозволяє оцінити рівень розслабленості та концентрації, що є базовою задачею для мозково-комп'ютерних інтерфейсів.

У попередніх роботах [1] та [2] було досліджено спектральний аналіз ЕЕГ-сигналів, який забезпечив точність класифікації на рівні 70–80%, а також кількісний рекурентний аналіз, що підвищив точність до 95% для потиличних електродів. Однак обидва підходи базувалися на безпосередньому аналізі сигналів окремих електродів, які є сумішшю активності багатьох мозкових джерел та артефактів. Це обмежує як точність класифікації, так і можливість фізіологічної інтерпретації результатів.

Перспективним напрямком є використання аналізу незалежних компонент (ІСА), який дозволяє розділити змішані електродні сигнали на статистично незалежні джерела. У поєднанні з підбором еквівалентних диполів це дає змогу валідувати отримані компоненти як фізіологічно достовірні мозкові сигнали та відсіяти артефакти. Такий підхід потенційно забезпечує більш надійну класифікацію, оскільки працює безпосередньо з мозковими джерелами, а не з їхніми змішаними проєкціями на електроди.

**Аналіз останніх досліджень і публікацій.** Класифікація стану очей за EEG-даними досліджувалася авторами раніше. У першій роботі [1] застосовувався спектральний аналіз із використанням швидкого перетворення Фур'є, який продемонстрував стабільну точність розпізнавання на рівні 70–80%. У подальшій роботі було запропоновано метод класифікації на основі кількісного рекурентного аналізу, який дозволив підвищити точність до 95% для потиличних електродів (O1, O2) при використанні класифікатора на основі методу опорних векторів (SVM, Support Vector Machine). Обидва дослідження проводилися на базі даних *EEG Motor Movement/Imagery Dataset* [3], що містить записи від 109 учасників. Однак подальший аналіз з використанням методів детектування пошкоджених каналів виявив значні проблеми з якістю частини записів у цьому датасеті, що мотивувало пошук більш надійних методів попередньої обробки сигналів.

Одним із найбільш перспективних підходів до попередньої обробки EEG-даних є аналіз незалежних компонент (ICA, Independent Component Analysis). Фундаментальні засади цього методу описано у роботі Гіварінена та Ойя [4], де ICA визначається як метод знаходження лінійного представлення негаусівських даних таким чином, щоб компоненти були максимально статистично незалежними. У цій роботі представлено алгоритм FastICA — ефективну ітераційну схему з фіксованою точкою для знаходження незалежних компонент шляхом максимізації негаусівськості. Метод має широке застосування, зокрема для сліпого розділення сигналів (класична «задача коктейльної вечірки»), виділення ознак у зображеннях та обробки біомедичних сигналів.

Першу успішну спробу застосування ICA до електроенцефалографічних даних було здійснено Мейкігом та співавторами [5]. Автори показали, що ICA-алгоритм Белла та Сейновського ефективно розділяє проблему ідентифікації джерела від його локалізації. Результати продемонстрували: (1) нечутливість навчання до різних початкових умов; (2) здатність ICA сегрегувати артефактні компоненти (шум мережі, рухи очей, м'язову активність) від мозкових сигналів; (3) можливість ізолювання накладених EEG-феноменів, включаючи альфа- та тета-сплески; (4) відстеження нестационарностей у EEG та поведінковому стані через зміни залишкової кореляції між ICA-каналами.

Глибоке порівняння різних ICA-алгоритмів для EEG-даних було проведено Делором та співавторами [6]. Автори застосували 22 різних алгоритми ICA та сліпого розділення джерел (BSS, Blind Source Separation) до 13 наборів 71-каналних EEG-даних і оцінили їх за трьома показниками: зменшення взаємної інформації (MIR, Mutual Information Reduction), залишкова попарна взаємна інформація (PMI, Pairwise Mutual Information) та ступінь відповідності дипольній моделі (dipolarity) — відповідність скальпових карт проекції одного еквівалентного диполя. Найменш ефективним виявився метод головних компонент (PCA, Principal Component Analysis), а найкращі результати показали AMICA та інші ICA-методи на основі правдоподібності. Було встановлено практично лінійний зв'язок між ефективністю зменшення взаємної інформації та кількістю компонент із високою відповідністю дипольній моделі, що підтверджує: алгоритми, які краще розділяють незалежні компоненти, також створюють більш фізіологічно достовірні скальпові карти. Незважаючи на менші абсолютні показники

FastICA порівняно з Infomax, результати цього алгоритму є загалом задовільними для практичного застосування.

Ефективність ICA для очищення ЕЕГ від артефактів наочно продемонстрована у роботі Делорма та співавторів [7], де метод застосовувався для видалення рухів очей та м'язового шуму з даних візуальної селективної уваги. Авторам вдалося ідентифікувати темпорально незалежний медіальний фронтальний ЕЕГ-процес, що відповідав за тета-сплески, які передували швидким моторним реакціям. Крім того, стаття Галлего-Рудольфа та співавторів [8] досліджує ефективність ICA для корекції балістокардіографічного артефакту при одночасному записі ЕЕГ та фМРТ. Автори встановили, що ICA-методи корекції найкраще зберігають реактивність задньої альфа-потужності під час завдання на закриття та відкриття очей, що є безпосередньо релевантним для нашого дослідження. Застосування ICA до аналізу даних фМРТ також описано Макіоном та співавторами [9], де метод продемонстрував перевагу над методом головних компонент та стандартним кореляційним аналізом у визначенні просторово-часових характеристик мозкової активації.

Ключовим аспектом валідації ICA-компонент як мозкових сигналів є їхня дипольна природа. Фундаментальні принципи дипольного аналізу потенціалів було описано Шергом [10], який запропонував просторово-часовий підхід до ідентифікації нейронних генераторів. Автор ввів поняття «регіонального дипольного джерела» з трьома ортогональними компонентами та показав, що модель із 6 диполями здатна пояснити 99,4% дисперсії скальпового потенціалу в просторі та часі, демонструючи високу надійність методу. Ключовим показником якості підбору диполя є залишкова дисперсія (residual variance), що визначається як відносна помилка між виміряними та модельними потенціалами.

Кількісну міру відповідності між потенціалами, обчисленими різними методами, було формалізовано Остенвельдом та Остендорпом [11] як relative error (RE):

$$RE = \sqrt{\frac{\sum_{i=1}^N (\phi_i - V_i)^2}{\sum_{i=1}^N V_i^2}} \quad (1)$$

де  $V_i$  — потенціал на електроді  $i$ , обчислений одним методом,  $\phi_i$  — потенціал на тому ж електроді, обчислений іншим методом, а  $N$  — кількість електродів. У інших роботах ця ж формула використовується під терміном residual variance (RV). На основі неї визначається дипольний показник (dipole score) як  $(1 - RV) \times 100\%$ . Обчислення RV реалізовано у програмному пакеті FieldTrip [12], від якого залежить модуль DIPFIT, що використовується в EEGLAB для підбору еквівалентних диполів.

Фізичну основу для дипольної моделі мозкової активності описує підручник Малмівуо та Плонсі [13], де описано модель об'ємного провідника у вигляді концентричних сферичних областей. Відповідно до цієї моделі, радіус голови становить 9,2 см, а провідність мозкової тканини та скальпа — 0,33 С/м. Ці параметри підтверджуються також у роботі Гірі та співавторів [14], де для чотиришарової моделі голови

(мозок, ЦСР, череп, скальп) обрано радіуси 8,0, 8,2, 8,7 та 9,2 см відповідно, а провідність мозку та скальпа —  $0,33 \text{ (Ом}\cdot\text{м)}^{-1}$ .

Якість ICA-декомпозиції оцінюється через ступінь відповідності дипольній моделі компонент. Артоні та співавтори [15] визначають дипольний показник компоненти  $IC_n$  як:

$$\text{dip}(IC_n) = 100(1 - \text{resvar}(IC_n))\% \quad (2)$$

де  $\text{resvar}(IC_n)$  — частка залишкової дисперсії між скальповою картою компоненти та її моделлю одного еквівалентного диполя. Для «квазідипольних» компонент з  $\text{dip} > 85\%$  та особливо для «майже дипольних» компонент із  $\text{dip} > 95\%$  позиція та орієнтація еквівалентного диполя є надійною оцінкою розташування джерела. У тій же роботі показано, що зменшення розмірності даних за допомогою методу головних компонент перед ICA суттєво погіршує якість декомпозиції — навіть видалення лише 1% дисперсії знижує кількість компонент із високою відповідністю дипольній моделі із 30 до 10 на набір даних.

Метод RELICA, запропонований Артоні та співавторами [16], дозволяє оцінити надійність незалежних компонент. RELICA класифікує компоненти за трьома класами: Клас I — стабільні та фізіологічно достовірні, Клас II — стабільні, але менш дипольні, Клас III — нестабільні. Порівняння FastICA та Infomax ICA показало, що Infomax є більш надійним: приблизно 20% декомпозицій FastICA не збігалися, а ті, що збігалися, давали менше зменшення взаємної інформації. У середньому RELICA ідентифікував 15 надійних та фізіологічно достовірних компонент на учасника з 71 можливих. Поріг відповідності дипольній моделі у 85% як критерій прийнятності компоненти також підтверджується роботою Онтон та співавторів [17], де компоненти із залишковою дисперсією понад 15% виключалися з подальшого аналізу.

Для практичної реалізації обробки EEG-даних існує низка програмних засобів. EEGLAB [18] є одним із найпотужніших інструментів для аналізу EEG, реалізованим як тулбокс для MATLAB. Він підтримує імпорт різних форматів даних, попередню обробку, ICA-декомпозицію, часово-частотний аналіз та багатопробну візуалізацію. Однак для нашого дослідження було обрано NeuroAnalyzer.jl [19] — тулбокс для мови Julia, що надає комплексний набір інструментів для аналізу нейрофізіологічних даних. Julia була обрана завдяки швидкодії (JIT-компіляція), відкритому коду та орієнтації на наукові обчислення. NeuroAnalyzer.jl підтримує імпорт форматів EDF/BDF/GDF, видалення поганих каналів (методи flat, amp, var, ransac), фільтрацію (FIR/IIR/Butterworth), ICA- та PCA-декомпозицію та візуалізацію. За результатами порівняння, NeuroAnalyzer.jl приблизно в 1,7 рази швидший за MNE (Python) і в 3 рази швидший за EEGLAB (MATLAB) для типових операцій обробки.

Перспективним напрямком подальших досліджень є математичне моделювання EEG-сигналів. Білозьоров та співавтори [20] запропонували нові моделі на основі сингулярних диференціальних рівнянь для опису сильно осциляційних процесів кори головного мозку. Модель здатна розрізняти сигнали здорового мозку та мозку пацієнта з

епілепсією, причому фрактальна розмірність атрактора здорового мозку є більшою, що свідчить про складнішу структуру хаосу. Інкін та Білозьоров [21] розвинули цей підхід, запропонувавши гібридну модель ФітцХью-Нагумо-Лоренца для моделювання ЕЕГ із оптимізацією параметрів на основі нейронних мереж. Потенційно NODE (Neural Ordinary Differential Equations) може бути застосована для моделювання очищених ІСА-сигналів ЕЕГ.

Таким чином, аналіз літератури свідчить, що ІСА є потужним інструментом для розділення мозкових сигналів від артефактів у ЕЕГ-даних, а підбір еквівалентних диполів дозволяє валідувати отримані компоненти як фізіологічно достовірні джерела. Поєднання ІСА-декомпозиції з класифікацією на основі спектральних та дипольних характеристик компонент може забезпечити більш надійну класифікацію стану очей порівняно з прямим аналізом сигналів електродів. Водночас математичне моделювання очищених ІСА-сигналів за допомогою систем диференціальних рівнянь відкриває перспективи для глибшого розуміння динаміки мозкової активності.

**Мета дослідження.** Мета цієї роботи полягає у розробці та дослідженні підходу до класифікації стану відкритих та закритих очей на основі незалежних компонент, отриманих методом ІСА з багатоканальних ЕЕГ-даних. На відміну від попередніх робіт, де класифікація виконувалася безпосередньо за сигналами окремих електродів, у цьому дослідженні пропонується використовувати ІСА-декомпозицію для виділення незалежних мозкових джерел із подальшою їх валідацією за допомогою підбору еквівалентних диполів. Особлива увага приділяється побудові послідовного алгоритму обробки: від детектування та видалення пошкоджених каналів, фільтрації шумів та ІСА-декомпозиції до підбору еквівалентних диполів, відбору фізіологічно достовірних компонент та класифікації стану очей за їхніми спектральними та просторовими характеристиками з використанням нейронної мережі. Крім того, досліджується якість датасету *EEG Motor Movement/Imagery Dataset* та вплив попередньої обробки на результати класифікації.

**Викладення основного матеріалу дослідження.** Для аналізу використовувалася база даних *EEG Motor Movement/Imagery Dataset* [3], що містить понад 1500 записів ЕЕГ тривалістю від однієї до двох хвилин, отриманих від 109 учасників. У ході експерименту учасники виконували різні моторні та уявні завдання, під час яких реєструвалися 64-канальні ЕЕГ-сигнали за допомогою системи *BCI2000* [22].

Кожен учасник проходив 14 експериментальних сесій, серед яких дві базові проби (по одній хвилині) – одна з відкритими очима, інша із закритими. Окрім цього, виконувалися три двоххвилинні проби для кожного з чотирьох завдань. Учасники мали відкривати та закривати кулак відповідно до появи мішені на лівій або правій стороні екрана або лише уявляти виконання цього руху. Додатково передбачалися завдання, у яких потрібно було відкривати та закривати обидва кулаки або обидві ноги залежно від позиції мішені.

Записи представлені у форматі *EDF+*, що містить 64 ЕЕГ-сигнали, зареєстровані з частотою 160 Гц, а також канал анотацій. Анотації включають коди, що позначають

різні стани:  $T0$  для спокою,  $T1$  для початку руху (реального або уявного) лівого кулака або обох кулаків та  $T2$  для руху правого кулака або обох ніг. Запис здійснено з використанням електродів за міжнародною системою  $10-10$ , яка виключає деякі специфічні електроди. Формат  $EDF+$  адаптований для використання із програмним забезпеченням *PhysioToolkit* [23].

Кожен учасник має унікальний ідентифікатор від  $S001$  до  $S109$ , а також 14 файлів із позначками  $R01-R14$ , що містять відповідні ЕЕГ-записи. Таким чином, для кожного запису файли представлені у форматі  $SsssRrr.edf$ , де  $S$  позначає номер учасника, а  $r$  – номер етапу експерименту. У цьому дослідженні були використані записи наступних етапів:

- 1 – стан без завдання, очі відкриті;
- 2 – стан без завдання, очі закриті.

**Імпорт даних.** Формат  $EDF+$ , у якому зберігаються записи датасету, не містить інформації про просторове розташування електродів, що є необхідним для подальшого підбору еквівалентних диполів та візуалізації топографічних карт. Для вирішення цієї проблеми було використано файл розташування електродів *Standard-10-20-Cap81.ced* з пакету *EEGLAB* [18], який містить координати електродів за стандартною системою  $10-10$ . Оскільки порядок електродів у файлі розташування не збігався з порядком каналів у датасеті, було проведено адаптацію: координати електродів були переупорядковані відповідно до послідовності каналів у записах  $EDF+$ .

**Видалення пошкоджених каналів.** Перед застосуванням ІСА-декомпозиції необхідно виявити та видалити пошкоджені канали, оскільки вони можуть суттєво погіршити якість розділення компонент і сповільнити збіжність алгоритму. Для цього у *NeuroAnalyzer.jl* [19] було послідовно застосовано три методи детектування:

1. **Метод flat** — виявляє канали з практично відсутнім сигналом, тобто ті, де амплітуда залишається майже незмінною протягом усього запису. Такі канали не несуть корисної інформації та є результатом поганого контакту електрода зі скальпом.

2. **Метод amp** — відсіює канали з аномальним діапазоном амплітуд, що виходить за межі фізіологічно обґрунтованих значень. Це дозволяє виключити записи з апаратними артефактами або надмірним шумом.

3. **Метод var** — визначає канали, дисперсія (*variance*) яких значно відрізняється від решти каналів. Канали з аномально високою або низькою дисперсією свідчать про нестабільний контакт або зовнішні перешкоди. Недоліком цього методу є те, що він може відсіяти лобні або потиличні електроди, оскільки вони природно мають більший рівень шуму від моргань та м'язів шиї відповідно.

Альтернативним підходом є метод *RANSAC*, який порівнює сигнал кожного каналу лише з його найближчими сусідами, що теоретично дозволяє уникнути хибного відсіювання каналів із специфічними, але фізіологічно обумовленими характеристиками. Однак у ході експериментів цей метод виявився занадто агресивним, відсіюючи надмірну кількість каналів, тому від його використання було вирішено відмовитися.

Аналіз результатів детектування пошкоджених каналів по всьому датасету показав, що приблизно 25% записів мають значну кількість непридатних каналів, що ставить під сумнів якість відповідних даних. За результатами фільтрації було залишено 82 учасники зі 109, у яких кількість валідних каналів була достатньою для подальшої обробки. Розподіл кількості придатних каналів наведено на рис. 1.

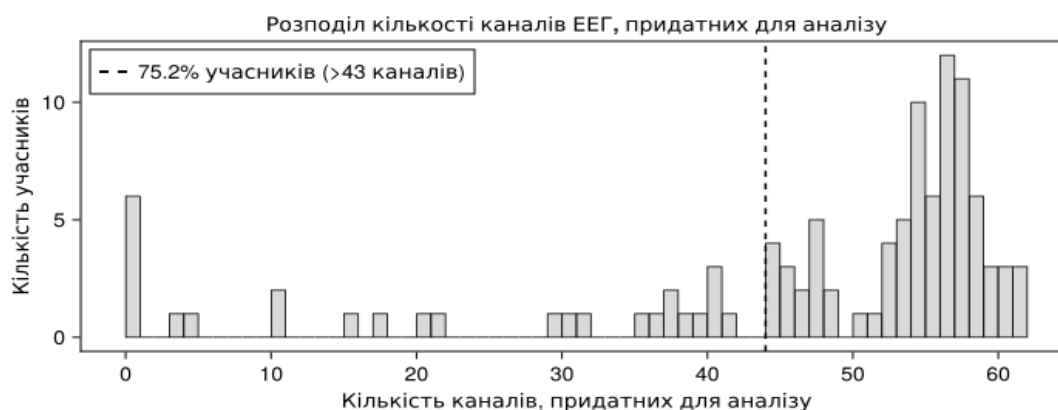


Рисунок 1 - Розподіл кількості каналів ЕЕГ, придатних для аналізу.

Пунктирна вертикальна лінія — поріг 43 канали; 75,2% учасників з усього датасету мають кількість каналів вище порогу

**Фільтрація сигналу.** Після видалення пошкоджених каналів було виконано частотну фільтрацію сигналу, яка є необхідною передумовою для коректної роботи ІСА-декомпозиції. Без попередньої фільтрації низькочастотні дрейфи та мережевий шум значно сповільнюють збіжність ІСА-алгоритму, а в окремих випадках призводять до того, що алгоритм не збігається взагалі.

Першим етапом було застосовано фільтр верхніх частот (highpass) з частотою зрізу 2 Гц та шириною перехідної смуги (bandwidth) 0,7 Гц. Це означає, що спектральна характеристика фільтра плавно зростає від 2,0 до 2,7 Гц, забезпечуючи поступовий перехід без різких артефактів. Порядок фільтра Баттерворта підбирався автоматично засобами NeuroAnalyzer.jl [19]. Цей фільтр видаляє низькочастотні складові сигналу, що можуть бути спричинені повільними рухами голови, дрейфом базової лінії електродів або потовиділенням.

Другим етапом було застосовано режекторний фільтр (notch filter) на частоті 50 Гц з шириною смуги 0,8 Гц для видалення шуму електричної мережі. Ширину смуги було обрано як середнє значення по датасету, хоча в ідеальному випадку цей параметр варто підбирати індивідуально для кожного запису залежно від ширини мережевого піку у його спектрі.

Як показали експерименти, після застосування обох фільтрів та видалення пошкоджених каналів швидкість ІСА-декомпозиції суттєво зросла: в середньому по датасету обчислення займало близько однієї секунди на 60 секунд запису для приблизно 40–50 залишених каналів.

**ІСА-декомпозиція.** Після попередньої обробки до сигналів було застосовано алгоритм ІСА, реалізований у NeuroAnalyzer.jl [19] як обгортка над бібліотекою MultivariateStats.jl. Математично ІСА-модель описується як лінійне змішування незалежних джерел [4]:

$$\mathbf{x}(t) = \mathbf{A} \cdot \mathbf{s}(t) \quad (3)$$

де  $\mathbf{x}(t)$  — вектор спостережуваних ЕЕГ-сигналів розміру  $n$  (кількість електродів),  $\mathbf{s}(t)$  — вектор незалежних компонент розміру  $m$  (кількість джерел),  $\mathbf{A}$  — матриця змішування розміру  $n \times m$ . Задача ІСА полягає у знаходженні матриці розмішування  $\mathbf{W}$  розміру  $m \times n$ , такої що:

$$\mathbf{s}(t) = \mathbf{W} \cdot \mathbf{x}(t) \quad (4)$$

Знаходження матриці  $\mathbf{W}$  виконується шляхом максимізації негаусівськості компонент. Згідно з центральною граничною теоремою, сума незалежних випадкових величин має розподіл, ближчий до гаусівського, ніж кожна з них окремо, тому максимально негаусівські компоненти є найбільш незалежними. Негаусівськість оцінюється апроксимацією негентропії на основі принципу максимальної ентропії [4]:

$$J(y) \approx \sum_{i=1}^p k_i [E\{G_i(y)\} - E\{G_i(v)\}]^2 \quad (5)$$

де  $E\{\cdot\}$  — оператор математичного сподівання,  $k_i$  — додатні константи,  $v$  — стандартна гаусівська випадкова величина, а  $G_i$  — неквадратичні функції. Ця міра є завжди невід'ємною та дорівнює нулю лише для гаусівського розподілу  $y$ .

Перед застосуванням ІСА виконуються два кроки математичної попередньої обробки [4]. Перший — центрування: від спостережуваного вектора  $\mathbf{x}$  віднімається його математичне сподівання  $\mathbf{m} = E\{\mathbf{x}\}$ , що робить дані нульовосередніми. Другий — відбілювання (whitening): дані лінійно перетворюються так, щоб їхні компоненти стали некорельованими з одиничною дисперсією:

$$\tilde{\mathbf{x}} = \mathbf{E} \mathbf{D}^{-1/2} \mathbf{E}^T \mathbf{x} \quad (6)$$

де  $\mathbf{E}$  — ортогональна матриця власних векторів коваріаційної матриці  $E\{\mathbf{x}\mathbf{x}^T\}$ ,  $\mathbf{D} = \text{diag}(d_1, \dots, d_n)$  — діагональна матриця її власних значень. Після відбілювання матриця змішування  $\tilde{\mathbf{A}} = \mathbf{E} \mathbf{D}^{-1/2} \mathbf{E}^T \mathbf{A}$  стає ортогональною, що вдвічі зменшує кількість параметрів для оцінки.

Наприклад, для 64 ЕЕГ-каналів та 12 обраних незалежних компонент матриця  $\mathbf{W}$  має розмір  $12 \times 64$ , а матриця  $\mathbf{A}$  —  $64 \times 12$ . Кожен стовпець матриці  $\mathbf{A}$  відповідає скальповій карті однієї компоненти, тобто описує, з якою вагою дана компонента впливає на кожен електрод. Саме ці стовпці є практичними значеннями потенціалу, які далі використовуються для підбору дипольної моделі.

**Підбір дипольної моделі.** Для валідації отриманих ІСА-компонент як мозкових сигналів було реалізовано алгоритм підбору еквівалентного диполя (dipole fitting). Мозок функціонує як сукупність електричних диполів, кожен з яких генерує електромагнітне поле, що реєструється електродами на поверхні скальпа [13]. Потенціал, створений одним диполем у точці розташування  $i$ -го електрода, визначається формулою [13]:

$$V_i = \frac{\mathbf{p} \cdot (\mathbf{r}_{e_i} - \mathbf{r}_d)}{4\pi\sigma|\mathbf{r}_{e_i} - \mathbf{r}_d|^3} \quad (7)$$

де  $\mathbf{r}_d$  — позиція диполя,  $\mathbf{r}_{e_i}$  — позиція  $i$ -го електрода,  $\mathbf{p}$  — вектор дипольного моменту,  $\sigma$  — провідність тканини.

Зв'язок між цією формулою та ІСА-моделлю є наступним. Якщо  $j$ -та незалежна компонента відповідає одному мозковому джерелу-диполю, то  $j$ -й стовпець матриці змішування  $\mathbf{A}$  має збігатися з розподілом потенціалів, створених цим диполем на всіх електродах. Тобто елемент  $a_{ij}$  матриці  $\mathbf{A}$  повинен бути пропорційним теоретичному потенціалу  $V_i$ , обчисленому для  $i$ -го електрода.

Задача підбору диполя полягає у знаходженні таких параметрів  $\mathbf{r}_{d_j}$  та  $\mathbf{p}_j$ , при яких теоретичний розподіл потенціалів максимально відповідає  $j$ -му стовпцю матриці  $\mathbf{A}$ . Пропорційність замість рівності пояснюється тим, що ІСА нормалізує компоненти, втрачаючи інформацію про абсолютну амплітуду.

Для наближення до фізичного сенсу було обрано параметри відповідно до літературних даних: провідність мозкової тканини та скальпа  $\sigma = 0,33$  С/м [13,14], а радіус голови — 9,2 см [13]. Оскільки координати електродів у файлі розташування зберігаються у нормалізованому діапазоні від  $-1$  до  $1$ , їх було масштабовано множенням на  $0,092$  м для переведення у систему СІ.

Оскільки нормалізація ІСА впливає лише на довжину  $|\mathbf{p}|$ , яка є лінійним множником у формулі потенціалу, вона не змінює нелінійну залежність від відстані між диполем та електродом, і підбір позиції диполя залишається коректним. Водночас масштабування координат електродів у систему СІ є необхідним саме тому, що відстань  $|\mathbf{r}_{e_i} - \mathbf{r}_{d_j}|$  входить у формулу нелінійно.

Для оцінки якості підбору використовувалася залишкова дисперсія [11]:

$$RV = \frac{\sum_{i=1}^N (a_i - \hat{a}_i)^2}{\sum_{i=1}^N a_i^2} \quad (8)$$

де  $a_i$  — практичне значення потенціалу зі стовпця матриці змішування ІСА,  $\hat{a}_i$  — теоретичне значення потенціалу, обчислене за дипольною моделлю,  $N$  — кількість елек-

тродів. Відповідно, дипольний показник визначається як  $(1 - RV) \times 100\%$ : значення 100% означає ідеальний збіг, а 0% — повну невідповідність.

Задача підбору диполя формулюється як задача мінімізації функції втрат, параметрами якої є вектор  $\boldsymbol{\theta} = (\mathbf{r}_d, \mathbf{p})$ , де  $\mathbf{r}_d \in \mathbb{R}^3$  — позиція диполя,  $\mathbf{p} \in \mathbb{R}^3$  — вектор дипольного моменту. Функція втрат складається з кількох компонент.

Основною частиною є залишкова дисперсія, де теоретичний потенціал  $\hat{a}_i$  на  $i$ -му електроді обчислюється за дипольною моделлю:

$$\hat{a}_i(\mathbf{r}_d, \mathbf{p}) = \frac{\mathbf{p} \cdot (\mathbf{r}_{e_i} - \mathbf{r}_d)}{4\pi\sigma|\mathbf{r}_{e_i} - \mathbf{r}_d|^3} \quad (9)$$

Для запобігання надмірним значенням параметрів введено L2-регуляризацію:

$$L_{reg} = \lambda_1 \|\boldsymbol{\theta}\|^2 \quad (10)$$

Оскільки без обмежень оптимізатор може розмістити диполь за межами голови, додано сферичний штраф, що різко зростає при виході позиції диполя за межі сфери:

$$L_{sphere} = \lambda_2 [\max(0, |\mathbf{r}_d|^2 - R^2)]^2 \quad (11)$$

де  $R = 0,092$  м — радіус голови.

Однак навіть із сферичним обмеженням диполь міг опинитися у зонах, що не відповідають мозковій тканині. Для більш точного обмеження було використано тривимірну модель (меш) мозку, отриману з бази даних NIH 3D Print Exchange. Меш було масштабовано до сфери радіусом 0,092 м та спрощено до низькополігональної опуклої оболонки (low-poly convex hull). Далі було підібрано еліпсоїди виду:

$$\frac{(x - c_{x_k})^2}{a_k^2} + \frac{(y - c_{y_k})^2}{b_k^2} + \frac{(z - c_{z_k})^2}{c_k^2} = 1 \quad (12)$$

що найкраще покривають точки мешу. Для кожного еліпсоїда обчислюється міра віддаленості точки  $\mathbf{r}_d$  від його поверхні:

$$d_k(\mathbf{r}_d) = \frac{(r_{dx} - c_{x_k})^2}{a_k^2} + \frac{(r_{dy} - c_{y_k})^2}{b_k^2} + \frac{(r_{dz} - c_{z_k})^2}{c_k^2} - 1 \quad (13)$$

де  $d_k \leq 0$  означає, що точка знаходиться всередині  $k$ -го еліпсоїда. Еліпсоїдний штраф дорівнює нулю, коли диполь знаходиться всередині хоча б одного еліпсоїда, та різко зростає при виході за межі всіх:

$$L_{ell} = \lambda_3 [\max(0, \min_{k=1, \dots, K} d_k(\mathbf{r}_d))]^2 \quad (14)$$

Таким чином, фінальна задача мінімізації для пошуку параметрів еквівалентного диполя  $j$ -ї ІСА-компоненти набуває вигляду:

$$\begin{aligned}
 L(\mathbf{r}_d, \mathbf{p}) = & \frac{\sum_{i=1}^N (a_{ij} - \hat{a}_i(\mathbf{r}_d, \mathbf{p}))^2}{\sum_{i=1}^N a_{ij}^2} + \lambda_1 \|\boldsymbol{\theta}\|^2 \\
 & + \lambda_2 [\max(0, |\mathbf{r}_d|^2 - R^2)]^2 \\
 & + \lambda_3 [\max(0, \min_k d_k(\mathbf{r}_d))]^2 \rightarrow \min_{\mathbf{r}_d, \mathbf{p}}
 \end{aligned}
 \tag{15}$$

де  $a_{ij}$  — елемент  $j$ -го стовпця матриці змішування  $\mathbf{A}$ ,  $N$  — кількість електродів, а коефіцієнти штрафів обрано емпірично:  $\lambda_1 = 10^{-3}$  (регуляризація),  $\lambda_2 = 10^6$  (сферичне обмеження),  $\lambda_3 = 10^8$  (анатомічне обмеження). Перший доданок забезпечує відповідність моделі спостережуваним даним, другий — обмежує зростання параметрів, третій та четвертий — утримують диполь у фізіологічно обґрунтованій області.

Візуалізація спрощеної моделі поверхні кори з двома підібраними еліпсоїдами (синій та червоний каркас) наведена на рис. 2: показано три орієнтації — зверху, збоку та ззаду. Еліпсоїди накладаються на меш, масштабований до сфери радіусом 0,092 м, і визначають область, усередині якої положення диполя не отримує еліпсоїдного штрафу  $L_{ell}$ .

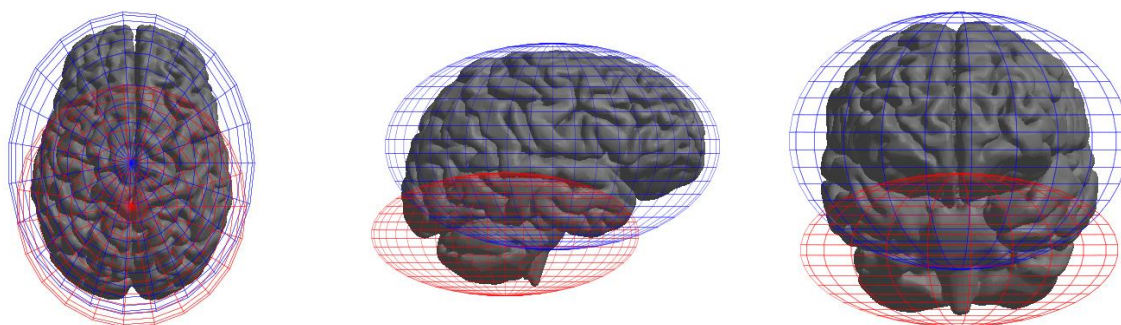


Рисунок 2 - Спрощена модель кори (меш) з двома еліпсоїдами, підібраними для апроксимації опуклої оболонки мозку: синій і червоний каркас — межі допустимої області для диполя  $L_{ell}$ ; вид зверху, збоку та ззаду

**Аналіз отриманих ІСА-сигналів.** Після застосування алгоритму підбору диполя до всього датасету для кожної ІСА-компоненти було отримано: топографічну карту (скальповий розподіл потенціалів зі стовпця матриці  $\mathbf{A}$ ), часовий ряд незалежної компоненти та криву спектральної щільності потужності (PSD, Power Spectral Density), обчислену методом Велча (Welch). На рис. 3–7 для кожного прикладу показано часовий ряд, оцінку спектральної щільності потужності та скальпову топографію з накладеним положенням підбраного еквівалентного диполя.

Візуальний аналіз результатів дозволив виявити кілька характерних типів компонент. Першу групу складають компоненти з домінуванням альфа-ритму — характерним

піком спектральної щільності на частоті близько 10 Гц. Топографічна карта таких компонент демонструє найбільшу активність у потиличній ділянці, а підібраний еквівалентний диполь розташовується в задній частині голови, що підтверджує потиличне походження сигналу (рис. 3).

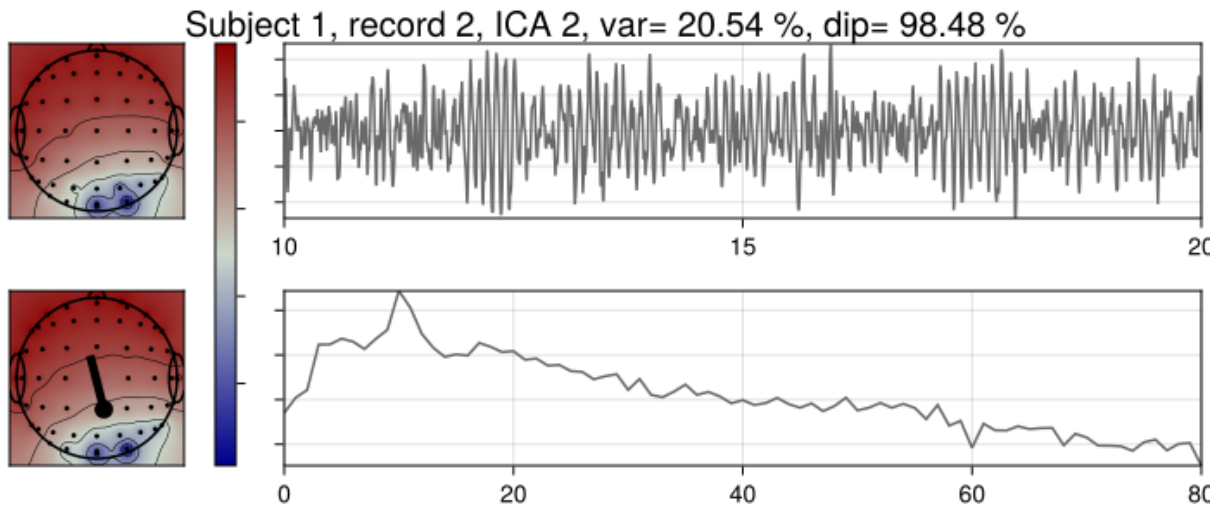


Рисунок 3 - ICA-компонента з домінуванням альфа-діапазону: на спектрі видно підсилення близько 10 Гц, топографія узгоджується з потиличною локалізацією джерела ( $s=1$ , R02, компонента  $i = 2$ )

Другу групу складають компоненти з вираженим мю-ритмом — аналогом альфа-ритму у сенсомоторній корі. На топографічній карті такі компоненти мають максимум активності у центральній ділянці скальпа, а еквівалентний диполь локалізується у сенсомоторній зоні (рис. 4).

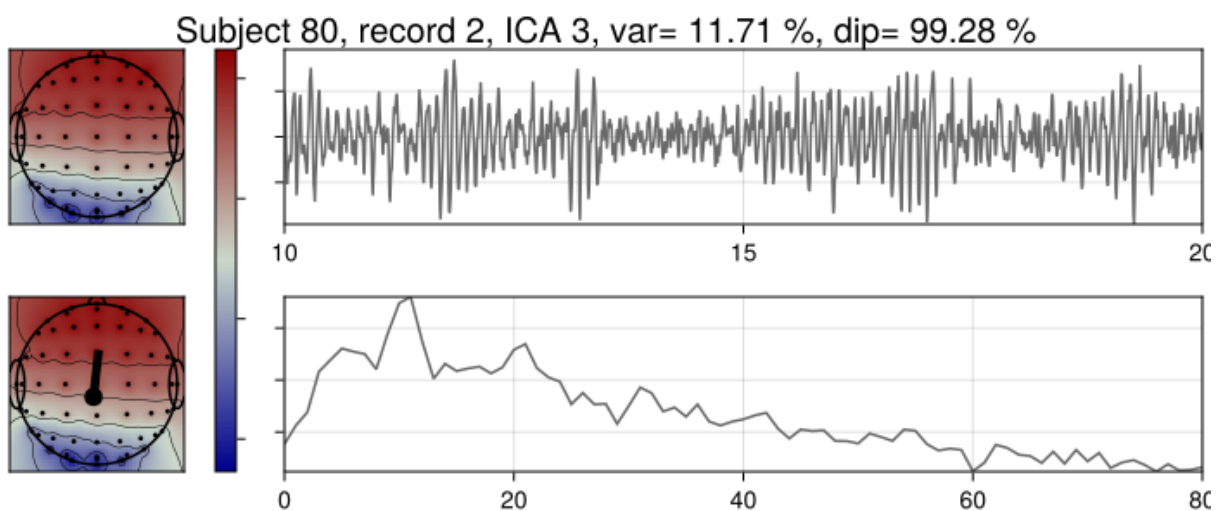


Рисунок 4 - Приклад компоненти з вираженим мю-ритмом у сенсомоторному діапазоні ( $s=80$ , R02,  $i = 3$ )

Окремі компоненти демонструють підвищену активність у бета-діапазоні (13–30 Гц). Топографія таких компонент зазвичай вказує на лобну або центральну локалізацію джерела (рис. 5).

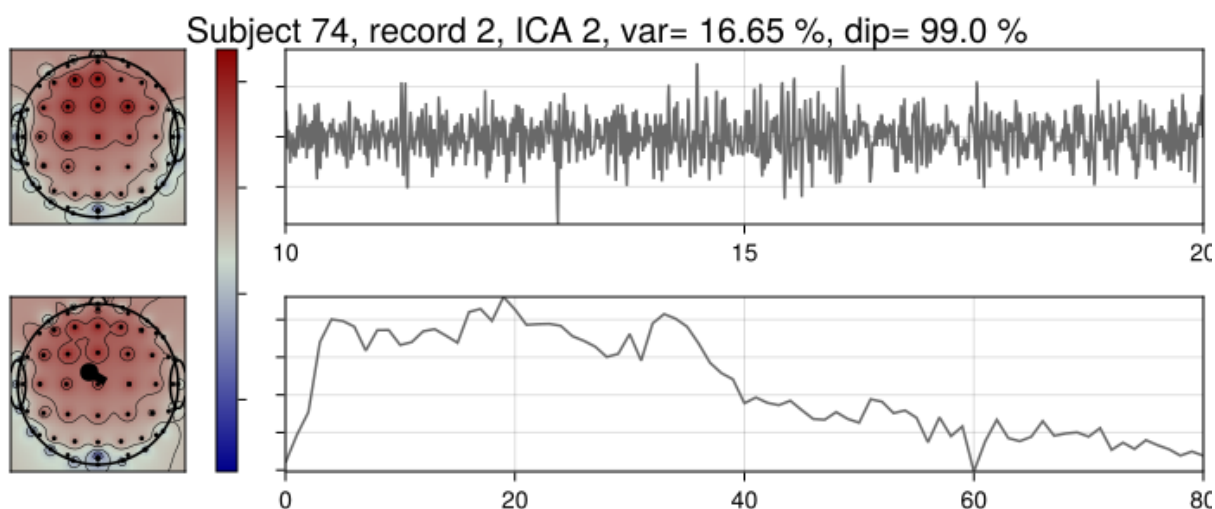


Рисунок 5 - Приклад компоненти з акцентом у бета-діапазоні ( $s=74$ ,  $R02$ ,  $i = 2$ )

Також було виявлено компоненти з підвищеною відносною потужністю у гамма-діапазоні (понад 30 Гц), природа яких потребує подальшого дослідження, оскільки вони можуть відображати як нейрональну активність, так і залишкові м'язові артефакти (рис. 6).

Нарешті, значна кількість компонент має низький дипольний показник, що свідчить про їхнє артефактне або змішане походження. На рис. 7 показано компоненту, що відповідає артефакту моргання: на часовому ряді видно різку короточасну зміну амплітуди, типову для моргання. Топографічна карта демонструє максимальну активність у лобно-полярній ділянці, що є характерною ознакою окорухових артефактів. Дипольний показник цієї компоненти становить лише 73%, що суттєво нижче порогу 85% для квазідипольних мозкових компонент.

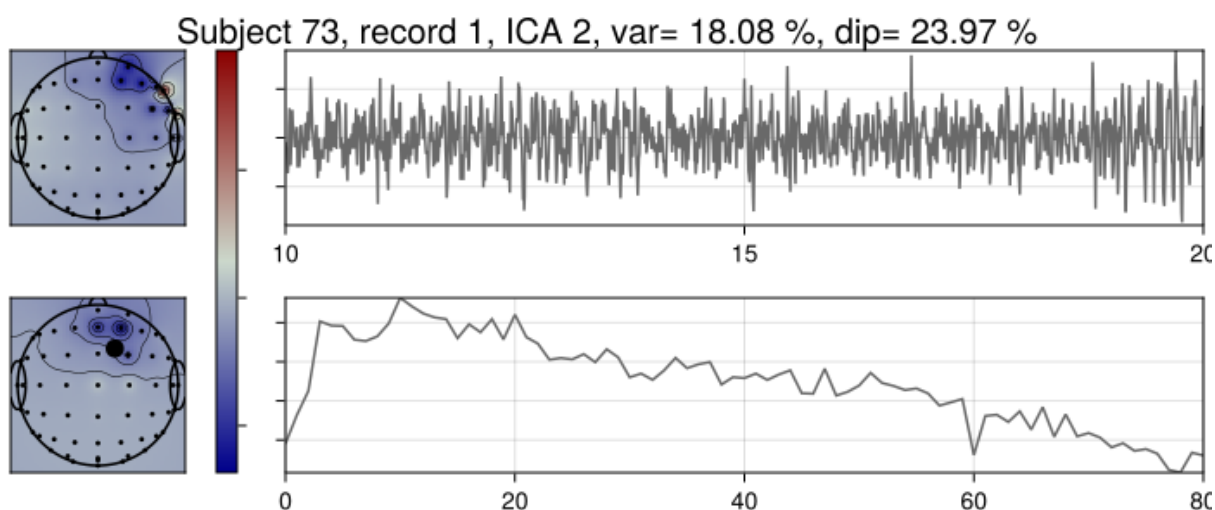


Рисунок 6 - Приклад компоненти з підвищеною відносною потужністю у гамма-діапазоні ( $s=73$ ,  $R01$ ,  $i = 2$ )

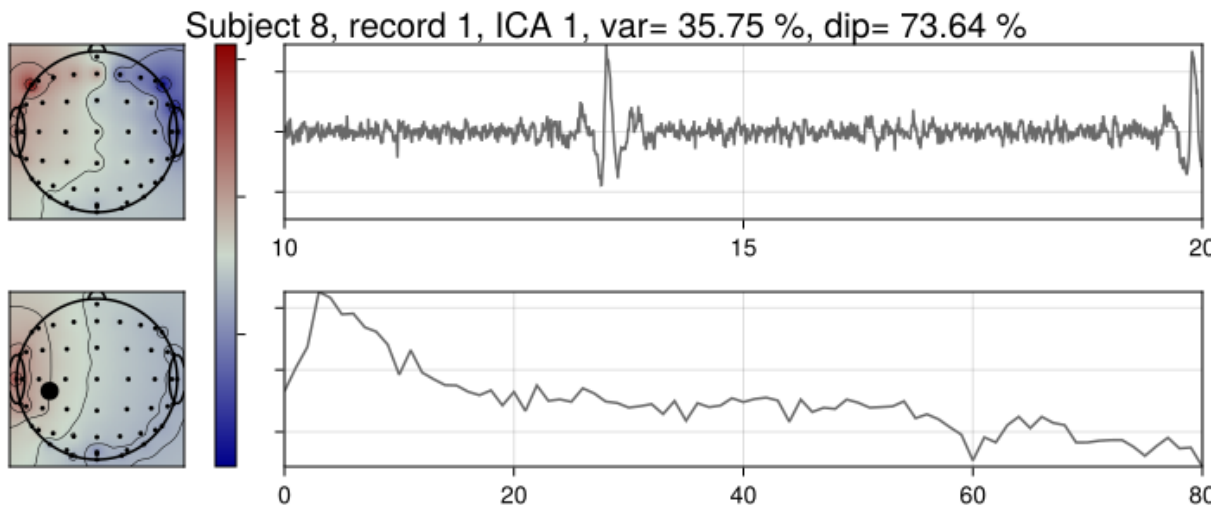


Рисунок 7 - Артефакт моргання: топографія з лобно-полярним максимумом, характерна різка зміна амплітуди на часовому ряді; дипольний показник 73% ( $s=8$ ,  $R01$ ,  $i = 1$ )

**Фільтрація датасету за якістю.** Для систематичного відбору даних побудовано гістограми розподілу ключових характеристик якості. Кількість придатних каналів на запис (рис. 1) визначила перший етап фільтрації — відбір 82 учасників зі 109.

Розподіл дипольного показника по всіх ICA-компонентах датасету (рис. 8) свідчить, що 57% компонент мають показник вище порогу 85%, рекомендованого у літературі [15,17] для ідентифікації квазідипольних мозкових компонент. Компоненти з нижчим показником було відсіяно як такі, що мають недостатню відповідність дипольній моделі.

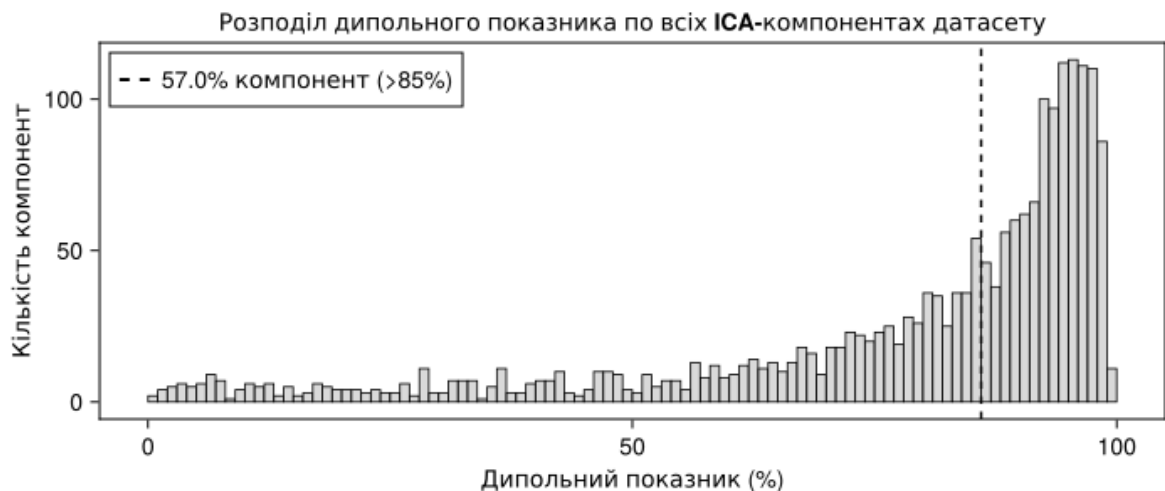


Рисунок 8 - Розподіл дипольного показника по всіх ICA-компонентах датасету.

Пунктир — поріг 85%; 57% компонент мають показник вище порогу

Для другого етапу фільтрації було обчислено мінімальну (між базовими сесіями R01 та R02) кількість компонент із дипольним показником понад 85% на кожного учасника (рис. 9). Результати показали, що 93,6% учасників із решти датасету мають

щонайменше дві такі компоненти. Додатково було виключено 3 учасників, для яких не залишилося жодної компоненти з достатнім дипольним показником.

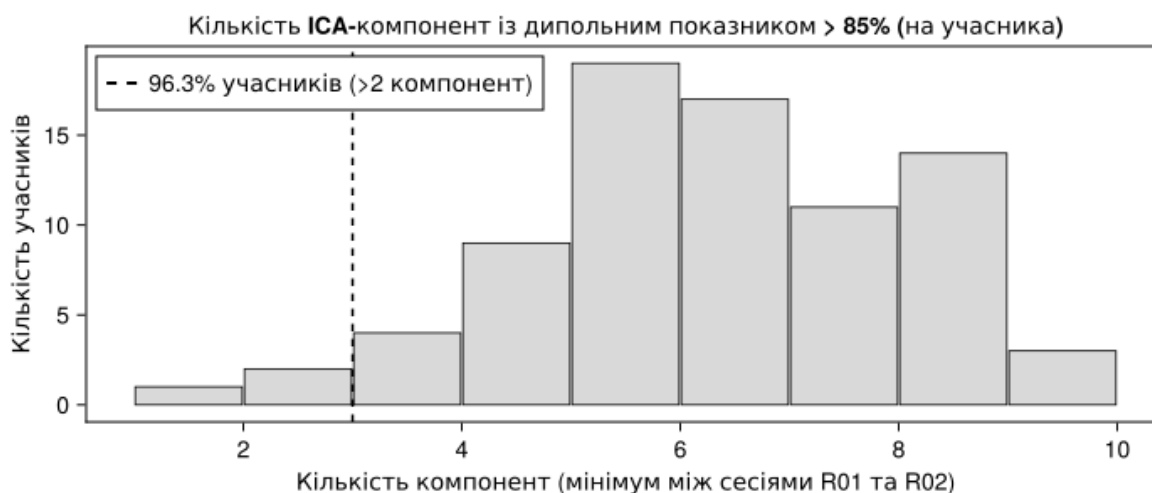


Рисунок 9 - Кількість ІСА-компонент із дипольним показником > 85% на учасника (мінімум між сесіями R01 та R02). Пунктир — поріг 2 компоненти; 93,6% учасників з решти датасету перевищують поріг

Таким чином, фінальний датасет складається з 79 учасників, кожен із яких має принаймні одну ІСА-компоненту з дипольним показником > 85%, а в середньому залишається близько 43 валідних ЕЕГ-каналів на запис.

**Формування ознак для класифікації.** Для кожної відібраної ІСА-компоненти було сформовано вектор ознак, що включає: дипольний показник (1 значення), дисперсію компоненти (variance, 1 значення), позицію диполя у тривимірному просторі (3 значення), вектор дипольного моменту (3 значення) та значення спектральної щільності потужності (161 значення). Загалом це становить 169 ознак на одну компоненту.

Для кожного учасника було відібрано три компоненти з найбільшою дисперсією, оскільки вони вносять найбільший вклад у загальну потужність сигналу. Таким чином, вектор ознак одного учасника має розмірність  $3 \times 169 = 507$  значень.

**Класифікація за допомогою нейронної мережі.** Для класифікації стану очей (відкриті чи закриті) на основі сформованих ознак було використано повнозв'язну нейронну мережу прямого поширення (feedforward neural network). Архітектура мережі складається з вхідного шару розмірністю 507, двох прихованих шарів із 256 та 128 нейронами відповідно та вихідного шару з двома нейронами для бінарної класифікації. У якості функції активації для прихованих шарів використано ReLU (Rectified Linear Unit).

Датасет із 79 учасників (по два записи на кожного — з відкритими та закритими очима, тобто 158 зразків) було розділено на тренувальну (80%) та тестову (20%) вибірки. До тестової вибірки було відібрано учасників із кращими показниками дисперсії компонент, щоб оцінити здатність моделі узагальнювати на даних більш високої якості.

Найкращий результат класифікації становив 97% точності на тренувальній вибірці та 90% на тестовій, що означає помилкову класифікацію лише одного учасника з тестового набору. Однак було виявлено суттєву проблему перенавчання (*overfitting*), зумовлену малим обсягом датасету. При подальшому навчанні без ранньої зупинки точність на тренувальних даних зростала до 100%, тоді як на тестових знижувалася до 75%. Це свідчить про те, що мережа починає запам'ятовувати конкретні зразки замість вивчення узагальнених закономірностей. Тому для досягнення оптимальних результатів було необхідно застосовувати ранню зупинку навчання (*early stopping*), відстежуючи момент, коли точність на тестовій вибірці починає знижуватися.

Порівняно з результатами попередніх робіт, де класифікація виконувалася безпосередньо за сигналами електродів, підхід на основі ІСА-компонент із дипольною валідацією демонструє конкурентоспроможну точність навіть при значно меншому обсязі даних (79 замість 109 учасників) та при використанні характеристик лише трьох найпотужніших незалежних компонент. Точність 90% на тестовій вибірці є порівнянною з результатами SVM-класифікатора на основі рекурентних показників (95% для O1), проте слід враховувати, що ІСА-підхід працює не з окремими електродами, а з розділеними мозковими джерелами, що надає результатам більшої фізіологічної інтерпретованості.

**Висновки.** У роботі запропоновано та досліджено підхід до класифікації стану відкритих і закритих очей, що базується на аналізі незалежних компонент багатоканальних ЕЕГ-даних із їх валідацією за допомогою підбору еквівалентних диполів. На відміну від попередніх робіт, де класифікація виконувалася безпосередньо за сигналами окремих електродів, запропонований метод працює з розділеними мозковими джерелами, що забезпечує більшу фізіологічну інтерпретованість результатів.

Побудовано послідовний алгоритм обробки ЕЕГ-даних: детектування та видалення пошкоджених каналів, частотна фільтрація, ІСА-декомпозиція, підбір еквівалентних диполів та відбір фізіологічно достовірних компонент за порогом дипольного показника 85%.

Аналіз датасету *EEG Motor Movement/Imagery Dataset* виявив значні проблеми з якістю частини записів: із 109 учасників лише 79 (72%) мали достатню кількість валідних каналів та ІСА-компонент із прийнятним дипольним показником. Це підтверджує необхідність ретельної попередньої обробки та контролю якості даних перед аналізом. Аналіз отриманих ІСА-компонент виявив джерела з характерними частотними профілями у альфа-, мю-, бета- та гамма-діапазонах, що підтверджує здатність методу виділяти фізіологічно інтерпретовані сигнали.

Для підбору еквівалентного диполя розроблено оригінальну функцію втрат, що окрім залишкової дисперсії включає L2-регуляризацию, сферичне обмеження та еліпсоїдне анатомічне обмеження на основі тривимірної моделі мозку. Еліпсоїдний штраф забезпечує розташування диполя у фізіологічно обґрунтованій області, що підвищує надійність результатів.

Класифікація на основі нейронної мережі з використанням 507 ознак (дипольний показник, дисперсія, просторові характеристики диполя та спектральна щільність пот-

ужності для трьох найпотужніших компонент) досягла точності 97% на навчальній та 90% на тестовій вибірці. Виявлено проблему перенавчання, зумовлену малим обсягом датасету (79 учасників, 158 зразків), що потребує застосування ранньої зупинки навчання.

Перспективними напрямками подальших досліджень є: збільшення обсягу даних шляхом залучення додаткових датасетів; застосування методів регуляризації та аугментації даних для боротьби з перенавчанням; використання математичного моделювання очищених ІСА-сигналів за допомогою систем диференціальних рівнянь для глибшого розуміння динаміки мозкової активності; а також дослідження впливу окремих частотних діапазонів ІСА-компонент на якість класифікації.

#### ЛІТЕРАТУРА

1. Ye. S. Panasenko, V. Ye. Belozyorov. THE APPLICATION OF SPECTRAL ANALYSIS OF EEG DATA FOR THE IDENTIFICATION OF OPEN AND CLOSED EYE STATES. *System technologies*. 2024. Т. 6, № 155, 155. С. 101–115. DOI: 10.34185/1562-9945-6-155-2024-11.
2. Panasenko Y.S., Belozyorov V.Y. CLASSIFICATION OF EYE STATE BASED ON EEG DATA USING RECURRENCE ANALYSIS. *Системні технології*. 2025. Т. 3, № 158, 158. С. 58–73. DOI: 10.34185/1562-9945-3-158-2025-07.
3. EEG Motor Movement/Imagery Dataset [Електронний ресурс]. URL: <https://archive.physionet.org/pn4/eegmmidb/> (дата звернення: 21.02.2024).
4. Hyvärinen A., Oja E. Independent Component Analysis: Algorithms and Applications. *Neural Networks*. 2000. Т. 13, № 4. С. 411–430. DOI: 10.1016/S0893-6080(00)00026-5.
5. Makeig S., Bell A., Jung T.-P., Sejnowski T.J. Independent Component Analysis of Electroencephalographic Data. *Advances in Neural Information Processing Systems*. MIT Press, 1995. Т.8. [https://papers.nips.cc/paper\\_files/paper/1995/hash/754dda4b1ba34c6fa89716b85d68532b-Abstract.html](https://papers.nips.cc/paper_files/paper/1995/hash/754dda4b1ba34c6fa89716b85d68532b-Abstract.html).
6. Delorme A., Palmer J., Onton J., Oostenveld R., Makeig S. Independent EEG Sources Are Dipolar. *PLOS ONE*. Public Library of Science, 2012. Т. 7, № 2. С. e30135. DOI: 10.1371/journal.pone.0030135.
7. Delorme A., Westerfield M., Makeig S. Medial Prefrontal Theta Bursts Precede Rapid Motor Responses during Visual Selective Attention. *Journal of Neuroscience*. Society for Neuroscience, 2007. Т. 27, № 44. С. 11949–11959. DOI: 10.1523/JNEUROSCI.3477-07.2007.
8. Gallego-Rudolf J., Corsi-Cabrera M., Concha L., Ricardo-Garcell J., Pasaye-Alcaraz E. Preservation of EEG Spectral Power Features during Simultaneous EEG-fMRI. *Frontiers in Neuroscience*. 2022. Т. 16. С. 951321. DOI: 10.3389/fnins.2022.951321.
9. Mckeown M.J., Makeig S., Brown G.G., Jung T., Kindermann S.S., Bell A.J., Sejnowski T.J. Analysis of fMRI Data by Blind Separation into Independent Spatial Components. *Human Brain Mapping*. 1998. Т. 6, № 3. С. 160–188. DOI: 10.1002/(SICI)1097-0193(1998)6:3<160::AID-HBM5>3.0.CO;2-1.

10. SCHERG M. Fundamentals of Dipole Source Potential Analysis. *Fundamentals of dipole source potential analysis*. Basel: Karger, 1990. Т. 6. С. 40–69.
11. Oostenveld R., Oostendorp T.F. Validating the Boundary Element Method for Forward and Inverse EEG Computations in the Presence of a Hole in the Skull. *Human Brain Mapping*. 2002. Т. 17, № 3. С. 179–192. DOI: 10.1002/hbm.10061.
12. Oostenveld R., Fries P., Maris E., Schoffelen J.-M. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*. 2011. Т. 2011, № 1. С. 156869. DOI: 10.1155/2011/156869.
13. Malmivuo J., Plonsey R. Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields. New York: Oxford University Press, 1995. 512 с.  
DOI: 10.1093/acprof:oso/9780195058239.001.0001.
14. Giri A., Kumar L., Kurwale N., Gandhi T.K. Anatomical Harmonics Basis Based Brain Source Localization with Application to Epilepsy. *Scientific Reports*. 2022. Т. 12. С. 11240. DOI: 10.1038/s41598-022-14500-7.
15. Artoni F., Delorme A., Makeig S. Applying Dimension Reduction to EEG Data by Principal Component Analysis Reduces the Quality of Its Subsequent Independent Component Decomposition. *NeuroImage*. 2018. DOI: 10.1016/j.neuroimage.2018.03.016.
16. Artoni F., Menicucci D., Delorme A., Makeig S., Micera S. RELICA: A Method for Estimating the Reliability of Independent Components. *NeuroImage*. 2014. Т. 103. С. 391–400. DOI: 10.1016/j.neuroimage.2014.09.010.
17. Onton J., Delorme A., Makeig S. Frontal Midline EEG Dynamics during Working Memory. *NeuroImage*. 2005. Т. 27, № 2. С. 341–356.  
DOI: 10.1016/j.neuroimage.2005.04.014.
18. Delorme A., Makeig S. EEGLAB: An Open Source Toolbox for Analysis of Single-Trial EEG Dynamics Including Independent Component Analysis. *Journal of Neuroscience Methods*. 2004. Т. 134, № 1. С. 9–21. DOI: 10.1016/j.jneumeth.2003.10.009.
19. Wysokiński A. NeuroAnalyzer: Julia Toolbox for Analyzing Neurophysiological Data. *Journal of Open Source Software*. 2025. Т. 10, № 107. С. 7734. DOI: 10.21105/joss.07734.
20. Belozyorov V.Y., Volkova S.A., Zaytsev V.G. Singular Differential Equations and Their Applications for Modeling Strongly Oscillating Processes. *Journal of Optimization, Differential Equations and Their Applications*. 2023. Т. 31, № 1, 1. С. 22–52. DOI: 10.15421/142302.
21. Inkin O.A., Belozyorov V.E. Hybrid Modeling of Eeg: The Fitzhugh-Nagumo-Lorenz Model. *System technologies*. 2025. Т. 3, № 158. С. 87–95. DOI: 10.34185/1562-9945-3-158-2025-09.
22. Schalk G., McFarland D.J., Hinterberger T., Birbaumer N., Wolpaw J.R. BCI2000: A General-Purpose Brain-Computer Interface (BCI) System. *IEEE transactions on bio-medical engineering*. 2004. Т. 51, № 6. С. 1034–1043. DOI: 10.1109/TBME.2004.827072.
23. Goldberger A.L., Amaral L.A., Glass L., Hausdorff J.M., Ivanov P.C., Mark R.G., Mietus J.E., Moody G.B., Peng C.K., Stanley H.E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*. 2000. Т. 101, № 23. С. E215–220. DOI: 10.1161/01.cir.101.23.e215.

## REFERENCES

1. Ye. S. Panasenko, & V. Ye. Belozyorov. (2024). THE APPLICATION OF SPECTRAL ANALYSIS OF EEG DATA FOR THE IDENTIFICATION OF OPEN AND CLOSED EYE STATES. *System Technologies*, 6(155, 155), 101–115. <https://doi.org/10.34185/1562-9945-6-155-2024-11>
2. Panasenko, Y.S., & Belozyorov, V.Y. (2025). CLASSIFICATION OF EYE STATE BASED ON EEG DATA USING RECURRENCE ANALYSIS. *Системні Технології*, 3(158, 158), 58–73. <https://doi.org/10.34185/1562-9945-3-158-2025-07>
3. *EEG Motor Movement/Imagery Dataset*. (б. д.). Вилучено 21, Лютий 2024, із <https://archive.physionet.org/pn4/eegmmidb/>
4. Hyvärinen, A., & Oja, E. (2000). Independent Component Analysis: Algorithms and Applications. *Neural Networks*, 13(4), 411–430. [https://doi.org/10.1016/S0893-6080\(00\)00026-5](https://doi.org/10.1016/S0893-6080(00)00026-5)
5. Makeig, S., Bell, A., Jung, T.-P., & Sejnowski, T. J. (1995). Independent Component Analysis of Electroencephalographic Data. *Advances in Neural Information Processing Systems*, 8. [https://papers.nips.cc/paper\\_files/paper/1995/hash/754dda4b1ba34c6fa89716b85d68532b-Abstract.html](https://papers.nips.cc/paper_files/paper/1995/hash/754dda4b1ba34c6fa89716b85d68532b-Abstract.html)
6. Delorme, A., Palmer, J., Onton, J., Oostenveld, R., & Makeig, S. (2012). Independent EEG Sources Are Dipolar. *PLOS ONE*, 7(2), e30135. <https://doi.org/10.1371/journal.pone.0030135>
7. Delorme, A., Westerfield, M., & Makeig, S. (2007). Medial Prefrontal Theta Bursts Precede Rapid Motor Responses during Visual Selective Attention. *Journal of Neuroscience*, 27(44), 11949–11959. <https://doi.org/10.1523/JNEUROSCI.3477-07.2007>
8. Gallego-Rudolf, J., Corsi-Cabrera, M., Concha, L., Ricardo-Garcell, J., & Pasaye-Alcaraz, E. (2022). Preservation of EEG Spectral Power Features during Simultaneous EEG-fMRI. *Frontiers in Neuroscience*, 16, 951321. <https://doi.org/10.3389/fnins.2022.951321>
9. Mckeown, M. J., Makeig, S., Brown, G. G., Jung, T., Kindermann, S. S., Bell, A. J., & Sejnowski, T. J. (1998). Analysis of fMRI Data by Blind Separation into Independent Spatial Components. *Human Brain Mapping*, 6(3), 160–188. [https://doi.org/10.1002/\(SICI\)1097-0193\(1998\)6:3<160::AID-HBM5>3.0.CO;2-1](https://doi.org/10.1002/(SICI)1097-0193(1998)6:3<160::AID-HBM5>3.0.CO;2-1)
10. SCHERG, M. (1990). Fundamentals of Dipole Source Potential Analysis. *Fundamentals of dipole source potential analysis*, 6, 40–69.
11. Oostenveld, R., & Oostendorp, T. F. (2002). Validating the Boundary Element Method for Forward and Inverse EEG Computations in the Presence of a Hole in the Skull. *Human Brain Mapping*, 17(3), 179–192. <https://doi.org/10.1002/hbm.10061>
12. Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, 2011(1), 156869. <https://doi.org/10.1155/2011/156869>
13. Malmivuo, J., & Plonsey, R. (1995). *Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195058239.001.0001>

14. Giri, A., Kumar, L., Kurwale, N., & Gandhi, T. K. (2022). Anatomical Harmonics Basis Based Brain Source Localization with Application to Epilepsy. *Scientific Reports*, 12, 11240. <https://doi.org/10.1038/s41598-022-14500-7>
15. Artoni, F., Delorme, A., & Makeig, S. (2018). Applying Dimension Reduction to EEG Data by Principal Component Analysis Reduces the Quality of Its Subsequent Independent Component Decomposition. *NeuroImage*. <https://doi.org/10.1016/j.neuroimage.2018.03.016>
16. Artoni, F., Menicucci, D., Delorme, A., Makeig, S., & Micera, S. (2014). RELICA: A Method for Estimating the Reliability of Independent Components. *NeuroImage*, 103, 391–400. <https://doi.org/10.1016/j.neuroimage.2014.09.010>
17. Onton, J., Delorme, A., & Makeig, S. (2005). Frontal Midline EEG Dynamics during Working Memory. *NeuroImage*, 27(2), 341–356. <https://doi.org/10.1016/j.neuroimage.2005.04.014>
18. Delorme, A., & Makeig, S. (2004). EEGLAB: An Open Source Toolbox for Analysis of Single-Trial EEG Dynamics Including Independent Component Analysis. *Journal of Neuroscience Methods*, 134(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
19. Wysokiński, A. (2025). NeuroAnalyzer: Julia Toolbox for Analyzing Neurophysiological Data. *Journal of Open Source Software*, 10(107), 7734. <https://doi.org/10.21105/joss.07734>
20. Belozyorov, V. Y., Volkova, S. A., & Zaytsev, V. G. (2023). Singular Differential Equations and Their Applications for Modeling Strongly Oscillating Processes. *Journal of Optimization, Differential Equations and Their Applications*, 31(1, 1), 22–52. <https://doi.org/10.15421/142302>
21. Inkin, O. A., & Belozyorov, V. E. (2025). Hybrid Modeling of Eeg: The Fitzhugh-Nagumo-Lorenz Model. *System technologies*, 3(158), 87–95. <https://doi.org/10.34185/1562-9945-3-158-2025-09>
22. Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., & Wolpaw, J. R. (2004). BCI2000: A General-Purpose Brain-Computer Interface (BCI) System. *IEEE Transactions on Bio-Medical Engineering*, 51(6), 1034–1043. <https://doi.org/10.1109/TBME.2004.827072>
23. Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C. K., & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*, 101(23), E215–220. <https://doi.org/10.1161/01.cir.101.23.e215>

Received 07.04.2026.

Accepted 10.04.2026.

Published 30.04.2026

***Eye state classification based on electroencephalogram decomposition  
considering the dipole nature of brain signals***

*Classification of brain activity states from electroencephalography (EEG) data is a relevant task for neuroscience and brain-computer interfaces. In previous works, spectral analysis of electrode signals provided eye state classification accuracy of 70–80%, while quantitative recurrence analysis with an SVM classifier achieved up to 95% for occipital electrodes, but both approaches operated on mixed signals from multiple sources and artifacts, limiting accuracy and physiological interpretability.*

*The aim is to develop an approach to eye state classification (open/closed) based on independent component analysis (ICA) of multichannel EEG data with component validation through equivalent dipole fitting.*

*The EEG Motor Movement/Imagery Dataset (109 participants, 64-channel EEG) was used. A sequential processing pipeline was implemented in Julia using the NeuroAnalyzer.jl toolbox: bad channel detection (flat, amp, var methods), frequency filtering (highpass 2 Hz, notch 50 Hz), ICA decomposition, and equivalent dipole fitting. An original loss function was developed with L2 regularization and ellipsoidal anatomical constraint based on a three-dimensional brain model.*

*Dataset quality analysis showed that only 79 of 109 participants (72%) had sufficient reliable channels and components with dipole scores above 85%. Among ICA components, sources with characteristic alpha, mu, beta, and gamma rhythms were identified, confirming physiological interpretability of the method.*

*Neural network classification based on 507 features (dipole score, variance, spatial dipole characteristics, power spectral density of three most powerful components) achieved 97% accuracy on training and 90% on test data. An overfitting problem was observed due to the small dataset size (158 samples), requiring the application of early stopping.*

*Unlike previous works, the proposed approach operates on separated brain sources, providing greater physiological interpretability and opening perspectives for mathematical modeling of cleaned ICA signals using systems of differential equations.*

*Keywords: EEG, electroencephalogram, ICA, dipole, brain state classification, neural network, spectral analysis, signal decomposition, blind source separation, source localization, computational neuroscience, spatiotemporal dynamics, biocybernetics, BCI, NeuroAnalyzer, Julia.*

**Панасенко Єгор Сергійович** — аспірант кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара.

ORCID: <http://orcid.org/0009-0008-2790-767X>.

**Panasenko Yehor Serhiyovych** — PhD student, Department of Computer Technologies, Oles Honchar Dnipro National University.

ORCID: <http://orcid.org/0009-0008-2790-767X>.

**МЕТОДИ ТА АЛГОРИТМИ РОЗВ'ЯЗАННЯ ДИНАМІЧНИХ ЗАДАЧ  
ОПТИМАЛЬНОГО РОЗБИТТЯ МНОЖИН ІЗ РОЗМІЩЕННЯМ ЦЕНТРІВ  
ПІДМНОЖИН ТА ІНТЕГРАЛЬНИМИ ОБМЕЖЕННЯМИ**

*Анотація.* У статті розглядаються методи та алгоритми розв'язання динамічних задач оптимального розбиття множин із розміщенням центрів підмножин та інтегральними обмеженнями. Досліджуваний клас задач виникає у багатьох прикладних контекстах, зокрема у логістиці, задачах просторового розподілу ресурсів і територіального зонування, де параметри моделі змінюються з часом.

Запропонований підхід базується на математичній формалізації задач оптимального розбиття множини з урахуванням часової динаміки щільності та передбачає одночасну оптимізацію структури розбиття і положення центрів підмножин. Інтегральні обмеження інтерпретуються, як обмеження на узагальнені характеристики підмножин і суттєво впливають на властивості оптимальних розв'язків.

Основну увагу приділено теоретичному обґрунтуванню моделі, аналізу властивостей задачі та побудові математичного алгоритму її розв'язання. Отримані результати демонструють узгодженість запропонованого підходу з класичними моделями оптимізації та підтверджують його придатність для опису динамічних процесів у задачах оптимального розбиття множин.

*Ключові слова:* оптимальне розбиття, розміщення центрів, інтегральні обмеження, логістичні задачі, динамічні моделі, просторове зонування, розподіл ресурсів, чисельне моделювання.

**Постановка проблеми.** Задача оптимального розбиття множини з розміщенням центрів підмножин є важливою проблемою математичного програмування, що має широкий спектр практичних застосувань – від логістики та управління ресурсами до просторового планування. Класичні постановки задач оптимального розбиття, як правило, передбачають фіксоване розташування центрів підмножин або кластерів, що істотно спрощує пошук оптимального розбиття. Проте в реальних системах положення центрів можуть змінюватися під впливом зовнішніх факторів або внаслідок необхідності адаптації до нових умов, що зумовлює потребу у формулюванні та розв'язанні динамічних задач оптимального розбиття множини з одночасним визначенням координат центрів.

Подібні динамічні задачі є значно складнішими, оскільки вимагають одночасної оптимізації розбиття елементів множини та визначення координат центрів підмножин, які можуть змінюватися з часом. Такий підхід забезпечує більшу гнучкість моделювання та дає змогу враховувати адаптивні процеси в системах, що перебувають у стані

змін. Для розв'язання задач цього класу зазвичай застосовуються методи теорії оптимального розбиття множин, диференціальних рівнянь та чисельні методи.

Додатково слід зазначити, що у межах запропонованого підходу динамічна щільність відіграє визначальну роль у формуванні просторової структури розбиття, оскільки саме її еволюція зумовлює перерозподіл навантажень між підмножинами. Зміна щільності в часі безпосередньо впливає на доцільність розміщення центрів підмножин, змушуючи їх адаптивно зміщуватися у напрямках підвищеної інтенсивності або попиту. Таким чином, координати центрів розглядаються не як фіксовані параметри, а як динамічні змінні, узгоджені з поточним станом щільності, що дозволяє забезпечити більш адекватне відображення реальних просторово-часових процесів та підвищити ефективність отриманих розв'язків.

Під динамічною неперервною лінійною однопродуктовою задачею оптимального розбиття множини  $\Omega$  з  $n$ -вимірною евклідовою простору  $E_n$  на її неперетинні підмножини  $\Omega_1, \dots, \Omega_N$  (серед яких можуть бути й порожні) за наявності обмежень у формі рівностей та нерівностей із відшукуванням координат центрів  $\tau_1, \dots, \tau_N$  цих підмножин відповідно розумітимемо наступну задачу.

Нехай  $\Omega$  – обмежена, вимірною за Лебегом множина у  $n$ -вимірному евклідовому просторі  $E_n$ .

Потрібно розбити її на  $N$  вимірних за Лебегом підмножин  $\Omega_1, \dots, \Omega_N$  (серед яких можуть бути й порожні) так, щоб

$$mes(\Omega_i \cap \Omega_j) = 0, \quad i \neq j, \quad i, j = 1, \dots, N,$$

де  $mes(\cdot)$  – міра Лебега,

Позначимо клас усіх можливих розбиттів множини  $\Omega$  на неперетинні підмножини  $\Omega_1, \dots, \Omega_N$ , через  $\Sigma_{\Omega}^N$ , тобто

$$\Sigma_{\Omega}^N = \{(\Omega_1, \dots, \Omega_N) : \bigcup_{i=1}^N \Omega_i = \Omega, mes(\Omega_i \cap \Omega_j) = 0, i \neq j, i, j = 1, \dots, N\}.$$

$$\bigcup_{i=1}^N \Omega_i = \Omega,$$

**Аналіз останніх досліджень і публікацій.** Дослідження неперервних задач оптимального розбиття множин у науковій літературі з'явилися в другій половині ХХ століття [1-4]. Вони виникли як нескінченновимірні узагальнення задач оптимізації виробничо-транспортного планування. Автори систематизували існуючі підходи до постановки та розв'язання задач оптимального розбиття множин, описали класи алгоритмів, їхні властивості та типові області застосування таких задач. Більш детальний огляд наукових публікацій з цієї теми наведено в роботах [5] і [8].

Теоретичні основи таких задач закладені в працях О. М. Кісельової, зокрема – у монографії [7], де було систематизовано різні підходи до побудови неперервних моде-

лей оптимального розбиття множин, детально розглянуто їхні математичні властивості та наведено приклади практичних застосувань у галузях транспорту, економіки й управління ресурсами. У роботі також підкреслено важливість дослідження задач із фіксованими та змінними центрами, оскільки такі моделі дозволяють описувати широкий спектр реальних процесів, пов'язаних з оптимальним розподілом і плануванням.

Фундаментальні праці [6] та [8] окреслюють теоретичні основи задач оптимального розбиття з одночасним визначенням координат центрів підмножин. Незважаючи на узагальнений характер цих досліджень, у них наведено формальні специфікації, корисні для програмної розробки, зокрема щодо структури вхідних даних, представлення розбиття та візуалізації результатів.

Подальший розвиток основної концепції відбувся в наступних публікаціях [9-11], де було запропоновано динамічні модифікації задач, що враховують зміну параметрів у часі. Це дало можливість застосовувати підхід оптимального розбиття до задач, у яких вхідні дані змінюються під впливом зовнішніх факторів, що особливо актуально для сучасних систем моніторингу, логістики та керування потоками.

Кілька міжнародних досліджень [12–16] зосереджені на задачах оптимізації форми та розбиття в механіці, спектральному аналізі та варіаційному численні. У працях, зокрема [12] і [14], доведено існування розв'язків для окремих класів задач оптимального розбиття, тоді як у роботі [13] досліджується мінімізація власних значень на розбитих областях. Незважаючи на іншу прикладну сферу, ці підходи можуть бути адаптовані до задач із фіксованими центрами шляхом введення геометричних обмежень.

**Мета дослідження.** Метою даної роботи є розробка методів і алгоритмів розв'язання динамічних задач оптимального розбиття множин із розміщенням центрів підмножин та інтегральними обмеженнями. Дослідження спрямоване на побудову математичної формалізації задач, у яких структура розбиття та положення центрів визначаються з урахуванням часової динаміки параметрів моделі.

Особливу увагу приділено аналізу властивостей сформульованих задач, врахуванню інтегральних обмежень та розробці алгоритму, що забезпечує узгоджене визначення оптимального розбиття і оптимального розміщення центрів у динамічному середовищі.

**Викладення основного матеріалу дослідження.** Знайти розбиття  $\varpi = \{\Omega_1, \dots, \Omega_N\} \in \Sigma_{\Omega}^N$  множини  $\Omega \subset E_n$ , векторну функцію  $\rho(x, t)$ , визначену майже всюди (м.в.) для  $x \in \Omega$  та всіх  $t \in [0, T]$ , і координати центрів  $\tau = \{\tau_1, \dots, \tau_N\} \subset \Omega^N$ , які забезпечують

$$\inf_{\substack{\varpi \in \Sigma_{\Omega}^N; \{\tau_1, \dots, \tau_N\} \subset \Omega^N; \rho(x, t) \in L_2^N(\Omega \times \Omega \times [0, T])}} F(\varpi, \tau, \rho(x, t)),$$

де

$$F(\varpi, \tau, \rho(x, t)) = \int_0^T \sum_{i=1}^N \int_{\Omega_i} (c_i(x, \tau_i) + a_i) \rho(x, t) dx dt$$

за умов

$$\frac{\partial \rho(x,t)}{\partial t} = f(\rho(x,t), x, t), 0 \leq t \leq T$$

$$\rho(x, t_0) = \rho_0(x)$$

та з інтегральними обмеженнями

$$\int_0^T \int_{\Omega_i} \rho(x,t) dx dt \leq b_i; \quad i = \overline{1, N},$$

або

$$\int_0^T \int_{\Omega_i} \rho(x,t) dx dt = b_i; \quad i = \overline{1, N},$$

Також можлива комбінація з обмеженнями рівності і нерівності у постановці задачі, тоді обмеження можуть бути записані наступним чином

$$\int_0^T \int_{\Omega_i} \rho(x,t) dx dt = b_i, \quad i = 1, \dots, p,$$

$$\int_0^T \int_{\Omega_i} \rho(x,t) dx dt \leq b_i, \quad i = p + 1, \dots, N.$$

Загалом, інтегральні обмеження в задачах оптимального розбиття множин, а також у транспортних і логістичних моделях, відображають глобальні характеристики системи – сумарні запаси ресурсів або загальну пропускну спроможність. Саме тому вони відіграють принципову роль у прикладних постановках задач. Такі обмеження дають змогу адекватно моделювати, зокрема, загальний обсяг пального чи сировини, сукупну виробничу потужність підприємства, максимальну кількість місць у пунктах прийому (наприклад, у системах надання допомоги), а також сумарну пропускну здатність транспортних ділянок і сервісних елементів інфраструктури. На відміну від локальних обмежень, інтегральні умови поєднують рішення на різних ділянках простору та на різних часових інтервалах, що призводить до ускладнення математичної структури задачі та зумовлює необхідність застосування спеціалізованих аналітичних і чисельних методів її розв'язання.

Формула, яка задає часову динаміку щільності у вигляді відповідного диференціального рівняння, природним чином приводить до постановки задачі Коші для системи звичайних диференціальних рівнянь. У такій постановці початкове значення щільності інтерпретується як початкова умова, тоді як сама система описує еволюцію щільності в часі. Це дозволяє застосовувати стандартний апарат теорії задачі Коші: досліджувати існування та єдиність розв'язку, встановлювати умови стійкості та неперервної залежності від початкових даних, а також обирати відповідні чисельні методи для інтегрування системи (методи Рунге-Кутта, Ейлера). Подібна формалізація забезпечує матема-

тичну коректність моделі та дозволяє узгоджено описувати зміну щільності в часі з урахуванням вибраного механізму динаміки.

Сукупність вимірних за Лебегом підмножин  $\Omega_1, \dots, \Omega_N$  з  $\Omega \subset E_n$  будемо називати можливим розбиттям множини  $\Omega$  на його неперетинні підмножини  $\Omega_1, \dots, \Omega_N$ , якщо

$$\bigcup_{i=1}^N \Omega_i = \Omega, \quad \text{mes}(\Omega_i \cap \Omega_j) = 0, \quad i \neq j, \quad i, j = 1, \dots, N,$$

де  $\text{mes}(\cdot)$  означає міру Лебега.

м.в. для  $x \in \Omega$ , при фіксованих  $\tau_i = (\tau_i^{(1)}, \dots, \tau_i^{(n)}) \in \Omega_i$ ,  $i = 1, \dots, N$ .

Тут  $\rho(x, t)$  – шукана дійсна функція, визначена на  $\Omega \times [0, T]$ , яка неперервно диференційована за аргументом  $t$  на відрізку  $[0, T]$  м.в. для  $x = (x^{(1)}, \dots, x^{(n)}) \in \Omega$ , обмежена та вимірна за аргументом  $x$  на  $\Omega$  для всіх  $t \in [0, T]$ ;  $\rho_0(x)$  – задана невід’ємна функція, обмежена та вимірна за аргументом  $x \in \Omega$ , також може бути задана у вигляді числа;  $f(\rho(x, t), x, t)$  – задана дійсна функція, неперервна та ліпшицева в межах свого визначення;  $c_i(x, \tau_i)$  – визначені на  $\Omega \times \Omega$  задані дійсні обмежені функції, вимірні за аргументом  $x = (x^{(1)}, \dots, x^{(n)})$  при будь-якому фіксованому  $\tau_i = (\tau_i^{(1)}, \dots, \tau_i^{(n)})$ ,  $i = 1, \dots, N$ ;  $\tau_1, \dots, \tau_N$  – сукупність деяких еталонних точок для підмножин  $\Omega_1, \dots, \Omega_N$  відповідно, званих центрами цих підмножин ( $\tau_i \in \Omega_i$ ,  $i = 1, 2, \dots, N$ ), причому координати центрів  $\tau_i = (\tau_i^{(1)}, \dots, \tau_i^{(n)})$ ,  $i = 1, 2, \dots, N$ , не відомі і вимагають свого визначення;  $T > 0$  та  $t_0 \in [0, T]$  задані числа;  $a_1, \dots, a_N$ ,  $b_1, \dots, b_N$  – задані невід’ємні числа, причому

$$S = \int_0^T \int_{\Omega} \rho(x, t) dx dt \leq \sum_{i=1}^N b_i, \quad 0 \leq b_i \leq S, \quad i = 1, \dots, N.$$

Тут і надалі інтеграли розуміються у сенсі Лебега. Вважатимемо, що міра множини межових точок  $\Omega_i$ ,  $i = 1, \dots, N$ , дорівнює нулю.

У наведеній математичній постановці динамічної задачі оптимального розбиття незалежна змінна  $t \in [0, T]$  може відігравати роль часової змінної, а  $T > 0$  та  $t_0 \in [0, T]$  – задані кінцевий та початковий моменти часу в досліджуваному динамічному процесі.

**Метод розв’язання задачі.** Введемо характеристичну функцію підмножини  $\Omega_i$ :

$$\lambda_i(x) = \begin{cases} 1, & x \in \Omega_i, \\ 0, & x \in \Omega \setminus \Omega_i, \end{cases} \quad i = 1, \dots, N,$$

та розглянемо функціонал

$$I(\lambda(\cdot), \tau, \rho(x, t)) = \int_0^T \sum_{i=1}^N \int_{\Omega} (c_i(x, \tau_i) + a_i) \rho(x, t) \lambda_i(x) dx dt,$$

де вектор-функція  $\lambda(x)$  має вигляд  $\lambda(x) = (\lambda_1(x), \dots, \lambda_i(x), \dots, \lambda_N(x))$ ,  $x \in \Omega$ ,  $\tau = (\tau_1, \dots, \tau_N) \in \Omega^N$ .

Очевидна така рівність,

$$I(\lambda(\cdot), \tau, \rho(x, t)) = F(\{\Omega_1, \dots, \Omega_N\}, \{\tau_1, \dots, \tau_N\}).$$

Перепишемо задачу **A** у термінах заданих характеристичних функцій та підмножин  $\Omega_i$ ,  $i = 1, \dots, N$ , у такому вигляді.

**Задача В.** Знайти пару елементів  $(\lambda_*(\cdot), \tau_*)$ , де  $\lambda_*(x) \in \Gamma_1$ ,  $\tau_* \in \Omega^N$ , таку, що

$$I(\lambda_*(\cdot), \tau_*) = \min_{(\lambda(\cdot), \tau) \in \Gamma_1 \times \Omega^N} I(\lambda(\cdot), \tau),$$

$$\Gamma_1 = \{ \lambda(x) = (\lambda_1(x), \dots, \lambda_N(x)) \in \Gamma'_1 \text{ м.в. для } x \in \Omega;$$

за умов

$$\int_0^T \int_{\Omega_i} \rho(x, t) dx dt = b_i, \quad i = 1, \dots, p,$$

$$\int_0^T \int_{\Omega_i} \rho(x, t) dx dt \leq b_i, \quad i = p+1, \dots, N \},$$

де

$$\begin{aligned} \Gamma'_1 &= \{ \lambda(x) = (\lambda_1(x), \dots, \lambda_N(x)) : \\ \lambda_i(x) &= 0 \vee 1 \text{ м.в. для } x \in \Omega, \quad i = 1, \dots, N, \\ \sum_{i=1}^N \lambda_i(x) &= 1 \text{ м.в. для } x \in \Omega \}. \\ \tau &= (\tau_1, \dots, \tau_N) \in \underbrace{\Omega \times \dots \times \Omega}_N = \Omega^N \end{aligned}$$

Від нескінченновимірної задачі **B** з булевими значеннями змінних  $\lambda_i(\cdot)$ ,  $i = 1, \dots, N$ , перейдемо до відповідної задачі зі значеннями функцій  $\lambda_i(\cdot)$  з відрізка  $[0, 1]$ .

**Задача В.1.** Знайти пару елементів  $(\lambda_*(\cdot), \tau_*)$ , де  $\lambda_*(x) \in \Gamma_2$ ,  $\tau_* \in \Omega^N$ , таку, що

$$I(\lambda_*(\cdot), \tau_*) = \min_{(\lambda(\cdot), \tau) \in \Gamma_2 \times \Omega^N} I(\lambda(\cdot), \tau),$$

$$\Gamma_2 = \{ \lambda(x) = (\lambda_1(x), \dots, \lambda_N(x)) : \lambda(x) \in \Gamma \text{ м.в. для } x \in \Omega;$$

за умов

$$\int_0^T \int_{\Omega_i} \rho(x, t) dx dt = b_i, \quad i = 1, \dots, p,$$

$$\int_0^T \int_{\Omega_i} \rho(x, t) dx dt \leq b_i, \quad i = p + 1, \dots, N \},$$

де

$$\Gamma = \{ \lambda(x) : 0 \leq \lambda_i(x) \leq 1 \text{ для } x \in \Omega, i = 1, \dots, N, \sum_{i=1}^N \lambda_i(x) = 1 \text{ м.в. для } x \in \Omega \}$$

$$\tau = (\tau_1, \dots, \tau_N) \in \underbrace{\Omega \times \dots \times \Omega}_N = \Omega^N$$

Вочевидь,

$$I(\lambda_*(\cdot), \tau_*) = \min_{\tau \in \Omega^N} [ \min_{\lambda(\cdot) \in \Gamma_2} I(\lambda(\cdot), \tau) ].$$

Задача **В.1** має розв'язок, та може бути розв'язана за допомогою функціонала Лагранжа, сформованого для задачі **В.1** таким чином:

$$L(\{ \lambda(\cdot), \tau \}, \Psi) = I(\lambda(\cdot), \tau) + \sum_{i=1}^N \psi_i \left( \int_0^T \int_{\Omega} \rho(x, t) \lambda_i dx dt - b_i \right)$$

де  $\Psi = (\psi_1, \dots, \psi_i, \dots, \psi_N)$  – N-вимірний вектор дійсних чисел, які можуть бути, як більше нуля, так і менше;  $\lambda(x) = (\lambda_1(x), \dots, \lambda_N(x)) \in \Gamma$  м.в. для  $x \in \Omega$ ;

$$\tau = (\tau_1, \dots, \tau_i, \dots, \tau_N) \in \Omega^N.$$

**Висновки.** У статті розглянуто методи та алгоритми розв'язання динамічних задач оптимального розбиття множин із розміщенням центрів підмножин та інтегральними обмеженнями. Досліджуваний клас задач виникає у багатьох прикладних контекстах, зокрема у логістиці, задачах просторового розподілу ресурсів і територіального зонування, де параметри моделі змінюються з часом.

Запропоновано підхід, який базується на математичній формалізації задач оптимального розбиття множини з урахуванням часової динаміки щільності та передбачає одночасну оптимізацію структури розбиття і положення центрів підмножин. Інтегральні обмеження інтерпретуються, як обмеження на узагальнені характеристики підмножин і суттєво впливають на властивості оптимальних розв'язків.

ЛІТЕРАТУРА

1. Бейко І. В., Кісельова О. М. Умови оптимальності межі розподілу зон територіального обслуговування. Застосування мат. методів до розв'язання виробничо-економічних задач: зб. науков. праць. Дніпр. держ. ун-т. Д., 1973. С. 53–55.
2. Кісельова О. М. Дослідження одного класу задач оптимального розбиття: автореф. дис. канд. фіз.-мат. наук: спец. 01.05.01. Київ. держ. ун-т. К., 1975. 15 с.
3. Кісельова О. М., Бейко І. В. Властивості оптимальних рішень для однієї задачі зрощення. Краєві задачі фільтрації: зб. науков. праць. ІМ АН УРСР. К., 1973. С. 255–261.
4. Balas E., Padberg M. W. Set Partitioning: A Survey. Comb. Optimization, Lect. Summer School Comb. Optimization. Urbino, 1977. Chichester, 1979. P. 151–210.
5. Кісельова О. М., Коряшкіна Л. С. Моделі та методи розв'язання неперервних задач оптимального розбиття множин: монографія. Київ: Наук. думка, 2013. 606 с.
6. Kiseleva O. M., Shor N. Z. Solution of the Problem of Optimal Partitioning Including Allocation of the Centers of Gravity of the Subsets. Computational Mathematics and Mathematical Physics, 1989, vol. 29, no. 3, pp. 47–56. DOI: 10.1016/0041-5553(89)90146-8.
7. Кісельова О. М., Шор Н. З. Неперервні задачі оптимального розбиття множин: теорія, алгоритми, застосування: монографія. Київ: Наук. думка, 2005. 564 с.
8. Кісельова О. М. Становлення та розвиток теорії оптимального розбиття множин. Теоретичні і практичні застосування: монографія. Дніпро: Ліра, 2018. 532 с.
9. Kiseleva O. M., Koriashkina L. S., Shevchenko T. A. Solving the Dynamic Optimal Set Partitioning Problem with Arrangement of Centers of Subsets. Cybernetics and Systems Analysis, 2014, vol. 50, no. 6, pp. 842–853. DOI: 10.1007/s10559-014-9675-8.
10. Kiseleva E., Prytomanova O., Kuzenkov O. On the Dynamic Problem of Optimal Set Partitioning with Fixed Centers under Uncertainty. Problems of Control and Informatics, 2025, vol. 70, no. 4, pp. 6–23. DOI: 10.34229/1028-0979-2025-4-1.
11. Кісельова О. М., Кузенков О. О. Про динамічну задачу оптимального розбиття множин з відшукуванням координат центрів підмножин. Вісник Харківського національного університету імені В. Н. Каразіна. Серія: Математичне моделювання. Інформаційні технології. Автоматизовані системи управління, 2025, вип. 65. С. 33–45. DOI: 10.26565/2304-6201-2025-65-03.
12. Buttazzo G., Dal Maso G. An Existence Result for a Class of Shape Optimization Problems. Archive for Rational Mechanics and Analysis, 1993, vol. 122, no. 2, pp. 183–195. DOI: 10.1007/BF00378167.
13. Caffarelli L. A., Lin F. H. An Optimal Partition Problem for Eigenvalues. Journal of Scientific Computing, 2007, vol. 31, no. 1–2, pp. 5–18. DOI: 10.1007/s10915-006-9114-8.
14. Conti M., Terracini S., Verzini G. On a Class of Optimal Partition Problems Related to the Fučík Spectrum and to the Monotonicity Formulae. Calculus of Variations and Partial Differential Equations, 2005, vol. 22, pp. 45–72. DOI: 10.1007/s00526-004-0266-9.
15. Burger M., Hackl B., Ring W. Incorporating Topological Derivatives into Level Set Methods. Journal of Computational Physics, 2004, vol. 194, no. 1, pp. 334–362.

16. Amstutz S., Novotny A. A. Topological Optimization of Structures. Heidelberg: Springer, 2010.

#### REFERENCES

1. Beiko I. V., Kiseleva O. M. Conditions of Optimality for the Boundary of the Territorial Service Zone Distribution. Application of Mathematical Methods to Solving Production and Economic Tasks: Collection of Scientific Papers. Dnipropetrovsk State University. Dnipropetrovsk, 1973. P. 53–55.
2. Kiseleva O. M. Study of a Class of Optimal Partitioning Problems: Abstract of the Dissertation for the Degree of Candidate of Physical and Mathematical Sciences: Specialty 01.05.01. Kyiv State University. Kyiv, 1975. 15 p.
3. Kiseleva O. M., Beiko I. V. Properties of Optimal Solutions for an Irrigation Problem. Boundary Filtration Problems: Collection of Scientific Papers. Institute of Mathematics, Ukrainian SSR Academy of Sciences. Kyiv, 1973. P. 255–261.
4. Balas E., Padberg M. W. Set Partitioning: A Survey. Combinatorial Optimization: Lectures of the Summer School on Combinatorial Optimization. Urbino, 1977. Chichester et al., 1979. P. 151–210.
5. Kiseleva O. M., Koryashkina L. S. Models and Methods for Solving Continuous Optimal Partitioning Problems: Monograph. Kyiv: Naukova Dumka, 2013. 606 p.
6. Kiseleva O. M., Shor N. Z. Solution of the Problem of Optimal Partitioning Including Allocation of the Centers of Gravity of the Subsets. Computational Mathematics and Mathematical Physics, 1989, vol. 29, no. 3, pp. 47–56. DOI: 10.1016/0041-5553(89)90146-8.
7. Kiseleva O. M., Shor N. Z. Continuous Optimal Set Partitioning Problems: Theory, Algorithms, Applications: Monograph. Kyiv: Naukova Dumka, 2005. 564 p.
8. Kiseleva O. M. The Formation and Development of the Theory of Optimal Set Partitioning. Theoretical and Practical Applications: Monograph. Dnipro: Lira, 2018. 532 p.
9. Kiseleva O. M., Koryashkina L. S., Shevchenko T. A. Solving the Dynamic Optimal Set Partitioning Problem with Arrangement of Centers of Subsets. Cybernetics and Systems Analysis, 2014, vol. 50, no. 6, pp. 842–853. DOI: 10.1007/s10559-014-9675-8.
10. Kiseleva E., Prytomanova O., Kuzenkov O. On the Dynamic Problem of Optimal Set Partitioning with Fixed Centers under Uncertainty. Problems of Control and Informatics, 2025, vol. 70, no. 4, pp. 6–23. DOI: 10.34229/1028-0979-2025-4-1.
11. Kiseleva O. M., Kuzenkov O. O. On the Dynamic Problem of Optimal Set Partitioning with Finding the Coordinates of the Centers of Subsets. Bulletin of Kharkiv National University Named after V. N. Karazin. Series: Mathematical Modeling. Information Technologies. Automated Control Systems, 2025, issue 65, pp. 33–45. DOI:10.26565/2304-6201-2025-65-03.
12. Buttazzo G., Dal Maso G. An Existence Result for a Class of Shape Optimization Problems. Archive for Rational Mechanics and Analysis, 1993, vol. 122, no. 2, pp. 183–195. DOI: 10.1007/BF00378167.
13. Caffarelli L. A., Lin F. H. An Optimal Partition Problem for Eigenvalues. Journal of Scientific Computing, 2007, vol. 31, no. 1–2, pp. 5–18. DOI: 10.1007/s10915-006-9114-8.

14. Conti M., Terracini S., Verzini G. On a Class of Optimal Partition Problems Related to the Fučík Spectrum and to the Monotonicity Formulae. *Calculus of Variations and Partial Differential Equations*, 2005, vol. 22, pp. 45–72. DOI: 10.1007/s00526-004-0266-9.
15. Burger M., Hackl B., Ring W. Incorporating Topological Derivatives into Level Set Methods. *Journal of Computational Physics*, 2004, vol. 194, no. 1, pp. 334–362.
16. Amstutz S., Novotny A. A. *Topological Optimization of Structures*. Heidelberg: Springer, 2010.

Received 10.04.2026.  
Accepted 12.04.2026.  
Published 30.04.2026

***Methods and algorithms for solving dynamic optimal set partitioning problems  
with subset center placement and integral constraints***

*The paper presents an approach to solving solving dynamic optimal set partitioning problems with subset center placement and integral constraints. The considered class of problems arises in numerous applied contexts, including logistics, spatial resource allocation, and territorial zoning, where model parameters vary over time.*

*The proposed approach is based on a mathematical formalization of optimal set partitioning problems that accounts for the temporal dynamics of density and provides for the simultaneous optimization of the partition structure and the locations of subset centers. Integral constraints are interpreted as restrictions on generalized characteristics of subsets and significantly influence the properties of optimal solutions.*

*The primary focus is placed on the theoretical substantiation of the model, the analysis of problem properties, and the construction of a mathematical solution algorithm. The obtained results demonstrate the consistency of the proposed approach with classical optimization models and confirm its suitability for describing dynamic processes in optimal set partitioning problems.*

*Keywords: optimal partitioning, center placement, integral constraints, logistics problems, dynamic models, spatial zoning, resource allocation, numerical modeling.*

**Лозовський Артур Валентинович** – аспірант кафедри прикладної математики, Дніпровський національний університет імені Олеся Гончара, Україна, Дніпро.

ORCID: <https://orcid.org/0009-0000-7013-7430>

**Шведов Владислав Олександрович** – аспірант кафедри прикладної математики, Дніпровський національний університет імені Олеся Гончара, Україна, Дніпро.

ORCID: <https://orcid.org/0009-0001-7399-1365>

**Lozovskyi Artur** – PhD student, Department of Applied Mathematics, Oles Honchar Dnipro National University, Ukraine, Dnipro.

ORCID: <https://orcid.org/0009-0000-7013-7430>

**Shvedov Vladyslav** – PhD student, Department of Applied Mathematics, Oles Honchar Dnipro National University, Ukraine, Dnipro.

ORCID: <https://orcid.org/0009-0001-7399-1365>

Р.В. Мисюк, П.Р. Кулик, О.В. Гусак, О.С. Чмихало, В.О. Парубочий  
**МЕТОД ОЦІНЮВАННЯ СТАНУ ЕЛЕМЕНТУ ІНЖЕНЕРНОЇ  
КОНСТРУКЦІЇ З ВИКОРИСТАННЯМ ПОЯСНЮВАЛЬНОГО ШТУЧНОГО  
ІНТЕЛЕКТУ ТА МУЛЬТИМОДАЛЬНИХ ДАНИХ**

*Анотація.* У сучасних умовах розвитку технологій штучного інтелекту особливої актуальності набуває розробка інтелектуальних методів оцінювання технічного стану елементів інженерних конструкцій, здатних забезпечувати не лише автоматизовану класифікацію, а й інтерпретованість отриманих результатів. Це є важливим для задач моніторингу трубопроводів, будівельних конструкцій та інших відповідальних об'єктів, де своєчасне виявлення дефектів дозволяє знизити ризик аварійних ситуацій. Метою роботи є розробка підходу до оцінювання стану елемента інженерної конструкції з використанням методів машинного навчання та пояснювального штучного інтелекту на основі мультимодальних даних. У роботі сформовано набір даних, що включає частку дефектної області поверхні (*DefectRatio*), а також параметри вологості та температури. На основі цих ознак здійснювалося визначення технічного стану конструкції за трьома класами: «Норма», «Попередження» та «Критичний». Для побудови моделей класифікації використано алгоритми *Random Forest*, *XGBoost*, *LightGBM*, *Support Vector Machine* та *Logistic Regression* із застосуванням гіперпараметричної оптимізації та перехресної валідації. Експериментальні результати показали, що найвищу точність класифікації на тестовій вибірці забезпечили моделі *Random Forest* та *XGBoost* (0,90), тоді як *Logistic Regression*, *SVM* та *LightGBM* продемонстрували нижчі результати (0,80, 0,70 та 0,60 відповідно). Аналіз важливості ознак підтвердив домінуючу роль показника *DefectRatio* у формуванні класифікаційного рішення, тоді як параметри зовнішнього середовища мають допоміжний контекстний вплив. Для підвищення прозорості прийняття рішень застосовано підходи пояснювального штучного інтелекту, зокрема *Permutation Importance*, *SHAP*-аналіз та правилоорієнтовані пояснення, що дозволило отримати як глобальну, так і локальну інтерпретацію результатів роботи моделей. Отримані результати підтверджують доцільність використання поєднання машинного навчання, мультимодального аналізу та ХАІ для побудови інтелектуальних систем моніторингу технічного стану інженерних конструкцій.

*Ключові слова:* пояснювальний штучний інтелект, інженерія ознак, обробка мультимодальних даних, обробка зображень, інтелектуальний моніторинг, класифікація машинного навчання, аналіз ризиків, поверхневі дефекти, тріщина.

**Вступ.** Процес оцінювання технічного стану елементів інженерних конструкцій є складним завданням, що потребує значних часових, обчислювальних та інтелектуальних ресурсів. Основними ознаками, що використовуються під час такого оцінювання, є геометричні характеристики дефектів, зокрема довжина та ширина тріщин, рівень достовірності їх виявлення, а також зовнішні фактори впливу, такі як температура та вологість середовища. У зв'язку з цим особливий інтерес становить застосування рішень на основі штучного інтелекту, машинного навчання та пояснювальних підходів, що демонструють високу ефективність у задачах структурного моніторингу та технічної діагностики. Обробка великого обсягу мультимодальних даних дозволяє здійснювати ефективний контроль стану конструкцій, виявляти ознаки деградації, прогнозувати ризики пошкодження та приймати обґрунтовані рішення щодо експлуатації об'єктів.

Задача розробки інтелектуальних моделей, що забезпечують високу точність оцінювання технічного стану елементів інженерних конструкцій за мінімальних обчислювальних та часових витрат, а також надають можливість інтерпретації отриманих результатів. Вирішення цього завдання сприятиме розвитку цифрових технологій у сфері технічного моніторингу, підвищенню рівня безпеки експлуатації інженерних об'єктів та своєчасному виявленню критичних змін у їхньому стані.

Метою роботи є розробка методу оцінювання стану елементу інженерної конструкції з використанням пояснювального штучного інтелекту та перевірка роботи його на практиці.

**Аналіз останніх досліджень і публікацій.** Оцінювання технічного стану елементів інженерних конструкцій є одним із ключових напрямів сучасного структурного моніторингу, оскільки своєчасне виявлення дефектів, прогнозування їх розвитку та інтерпретація ознак пошкодження безпосередньо впливають на безпеку експлуатації інфраструктурних об'єктів. Упродовж останніх років у цій галузі спостерігається активне впровадження методів штучного інтелекту, комп'ютерного зору, мультимодальної обробки даних та пояснювального машинного навчання. Аналіз сучасних досліджень свідчить, що наукові праці в цій сфері можна умовно поділити на чотири основні напрями: автоматичне виявлення дефектів за зображеннями, оцінювання стану за сенсорними сигналами, мультимодальне поєднання даних та забезпечення інтерпретованості рішень моделей.

Попередні дослідження демонструють ефективність використання роботизованих платформ та IoT для автоматизованого виявлення дефектів у інфраструктурних об'єктах [1, 2]. Аналіз відеопотоків з застосуванням методів машинного навчання дозволяє класифікувати тріщини та пошкодження з високою точністю, що підтверджується експериментальними даними [2, 3]. Використання доступних сенсорів і мікрокомп'ютерів, таких як Raspberry Pi, забезпечує масштабоване і економічне рішення для моніторингу стану трубопроводів та дорожніх покриттів [4].

Значна частина сучасних досліджень присвячена **автоматизованому виявленню тріщин на поверхнях інженерних конструкцій** із використанням методів комп'ютерного зору та глибокого навчання. У роботі [5] показано ефективність глибо-

ких нейронних мереж для виявлення тріщин усередині конструкцій. У дослідженні [6] запропоновано підхід до оцінювання стану бетонних конструкцій на основі автоматизованого виявлення тріщин для різних типів бетонних поверхонь. Подібно, у праці [7] розглянуто застосування глибокого навчання не лише для детектування, а й для прогнозування розвитку тріщин у задачах структурного моніторингу. Ці роботи підтверджують високу ефективність візуального аналізу для виявлення дефектів, однак здебільшого зосереджуються саме на факті пошкодження, а не на комплексному оцінюванні технічного стану.

Окремий напрям досліджень стосується використання **сенсорних даних** для моніторингу інженерних конструкцій. У роботі [8] запропоновано метод структурного моніторингу на основі Bluetooth-мережі тензометричних датчиків та алгоритмів машинного навчання для виявлення і локалізації тріщин. У дослідженні [9] розглянуто автоматизовану ідентифікацію тріщин за акустичними хвильовими сигналами із застосуванням глибокого навчання. Такі підходи демонструють, що фізичні вимірювальні канали можуть слугувати важливим джерелом інформації про стан конструкції навіть без прямого візуального спостереження за дефектом.

Водночас сучасні дослідження все частіше орієнтуються на **інтеграцію різнорідних джерел інформації**. У роботі [10] наведено комплексний огляд розвитку штучного інтелекту для структурного моніторингу, де підкреслюється доцільність поєднання візуальних, сенсорних і часових даних для підвищення достовірності оцінювання технічного стану. У праці [11] запропоновано підхід до дослідження тріщиноутворення на основі поєднання реальних інженерних даних із фізичними обмеженнями, що є перспективним напрямом для побудови більш надійних і стійких моделей аналізу пошкоджень.

Особливої актуальності в задачах структурного моніторингу набуває **пояснювальний штучний інтелект (ХАІ)**, оскільки в критично важливих інженерних застосуваннях недостатньо лише отримати точне рішення — необхідно також пояснити причини його прийняття. У роботі [12] наведено критичний огляд підходів до побудови довірених та пояснюваних систем моніторингу мостів із залученням експерта до процесу аналізу. У дослідженні [13] продемонстровано використання ХАІ для імовірнісного оцінювання стану залізобетонних балок на основі зображень тріщин, що підтверджує перспективність застосування пояснювальних моделей у задачах технічної діагностики.

Незважаючи на наявність значної кількості досліджень у сфері структурного моніторингу, залишається актуальною задача побудови інтелектуального методу, який би одночасно забезпечував високу точність оцінювання стану конструкції, інтеграцію мультимодальних даних та інтерпретованість результатів, що й зумовило постановку задачі даної роботи.

**Виклад основного матеріалу.** У ході дослідження було сформовано набір мультимодальних даних, що описує технічний стан елементу інженерної конструкції у часовій динаміці. До складу вхідних ознак включено параметри виявлених дефектів, зокрема факт наявності дефекту типу тріщини, а також зовнішні фактори середовища,

такі як температура та вологість. На основі цих параметрів було обчислено інтегральний показник небезпечності дефекту та сформовано узагальнену оцінку технічного стану конструкції. У системі моніторингу стану конструкцій, реалізованій на базі Nvidia Jetson Nano з використанням камери та сенсора DHT22 [3, 4], оцінка тріщин проводиться на основі відношення дефектних пікселів до загальної площі зображення (DefectRatio), що формуються сегментаційним алгоритмом [1]. Параметри температури та вологості вимірюються сенсором DHT22 і відображають зовнішні умови всередині труби.

Для класифікації технічного стану елементу інженерної конструкції використано різні алгоритми машинного навчання, які дозволяють автоматизувати процес віднесення об'єкта до відповідного класу стану. У роботі реалізовано поділ стану конструкції на класи «критичний», «попередження» та «норма», що відповідають підвищеному та критичному рівням ризику. Для оцінювання якості побудованої моделі використано стандартні метрики класифікації, зокрема **accuracy**, **precision**, **recall** та **F1-score**, а також матрицю помилок.

З метою інтерпретованості результатів моделювання використано підходи **пояснювального штучного інтелекту**, зокрема аналіз важливості ознак, правилоорієнтоване пояснення та метод **SHAP (SHapley Additive exPlanations)**. Це дозволило не лише отримати високоточну модель оцінювання стану конструкції, а й визначити внесок кожного параметра у прийняття фінального рішення, що є важливим для практичного використання інтелектуальних систем у задачах структурного моніторингу та технічної діагностики.

*Збір та обробка даних.* Для дослідження запропонованого методу оцінювання стану елементу інженерної конструкції з використанням пояснювального штучного інтелекту було опрацьовано набір мультимодальних даних, що відображає зміну технічного стану конструкції в часовій динаміці. Такий підхід дозволив перевірити умови частини об'єкта структурного моніторингу, в яких стан об'єкта визначається не лише наявністю дефекту, а й супровідними фізичними умовами експлуатації.

У межах дослідження (Рис. 1) зібрано 32 кадри відео. Для кожного з часових кроків сформовано набір ознак: частка сегментованих дефектів на зображеннях (DefectRatio), температура, відносна вологість та цільову змінну стану (ризик). На основі цих показників обчислюється комплексний індикатор стану об'єкта RiskScore, який дозволяє класифікувати кадри за трьома категоріями: «Норма», «Попередження» та «Критичний». Такий підхід забезпечує інтеграцію візуальних і сенсорних даних для комплексної оцінки стану конструкцій, дозволяючи оперативно визначати наявність дефектів, їхню серйозність та вплив зовнішніх факторів. Система може застосовуватися як для локального моніторингу на пристрої Jetson Nano, так і для віддаленого контролю стану конструкцій у реальному часі.



Рисунок 1 – Результат кадрування відео на 32 зображення з сегментованими дефектами на ділянці труби

Початкові значення параметрів дефектів моделювали наявність невеликих пошкоджень на зображеннях, які поступово збільшувалися у часі. Рівень дефектності визначався через DefectRatio, тобто частку пікселів, помічених як дефектні, а також через Confidence, що відображає ймовірність коректної детекції алгоритмом комп'ютерного зору. Для кількісного опису серйозності дефекту було введено інтегральний показник ConditionScore, який об'єднує значення DefectRatio, Confidence, а також показники температури і вологості. На основі цього показника (Рис. 2) формувалися класи стану конструкції: «Норма», «Попередження» та «Критичний».

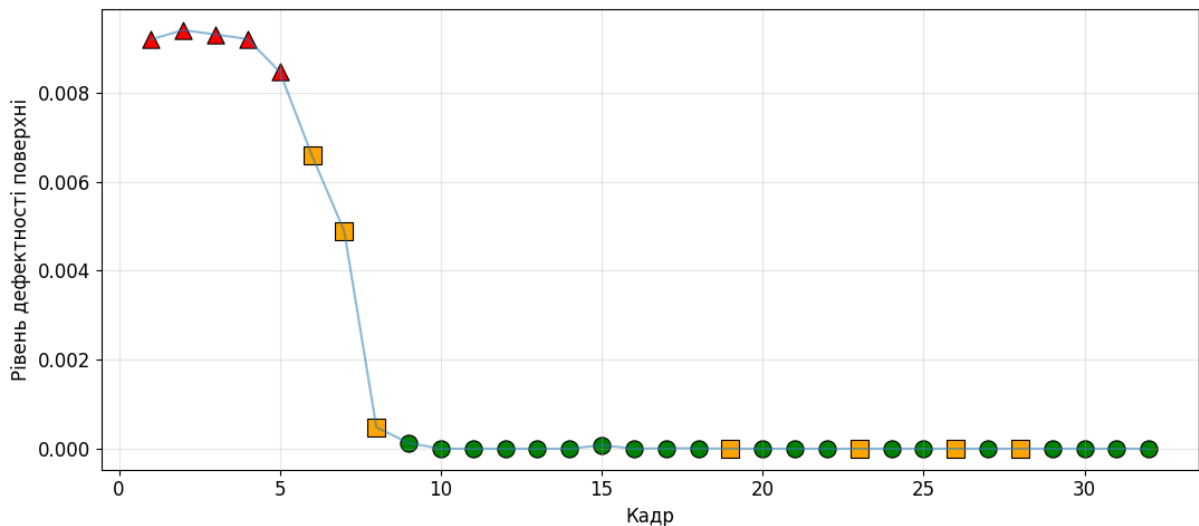


Рисунок 2 – Динаміка показника дефектності поверхні по кадрах

На першому етапі було сформовано набір експериментальних даних на основі 32 послідовних кадрів спостереження, для яких обчислювалися значення площі дефектів, а також супутні сенсорні параметри середовища. Аналіз часової динаміки дефектності

поверхні показав, що на початкових кадрах спостерігалася підвищена частка пошкодженої області, яка поступово зменшувалася до нульових або майже нульових значень. Така тенденція свідчить про перехід системи від станів із підвищеним рівнем дефектності до стабільного нормального стану.

Додатково було проаналізовано зміну параметрів середовища (Рис. 3). Отримані результати засвідчили плавне зростання вологості та температури протягом послідовності кадрів, що відображає зміну зовнішніх умов експлуатації або експериментального середовища. Введення порогових значень для вологості та температури дозволило візуально оцінити моменти, у яких середовище потенційно могло впливати на зміну стану поверхні.

Зменшення кількості виявлених дефектів при підвищенні температури та вологості найчастіше пов'язане з термічним розширенням матеріалу та зміною умов освітленості. Нагрівання конструкції (особливо на сонячній стороні) призводить до лінійного розширення елементів, що спричиняє механічне «затягування» мікротріщин та стикувальних зазорів, роблячи їх менш помітними для алгоритмів сегментації. Крім того, пряме сонячне світло в моменти закінчення моніторингу створює високу інтенсивність випромінювання, що може призводити до оптичного засвічування та зниження контрастності зображення. Це ускладнює диференціацію дефектів від фону, оскільки тіні, які зазвичай підкреслюють текстуру пошкоджень, зникають.

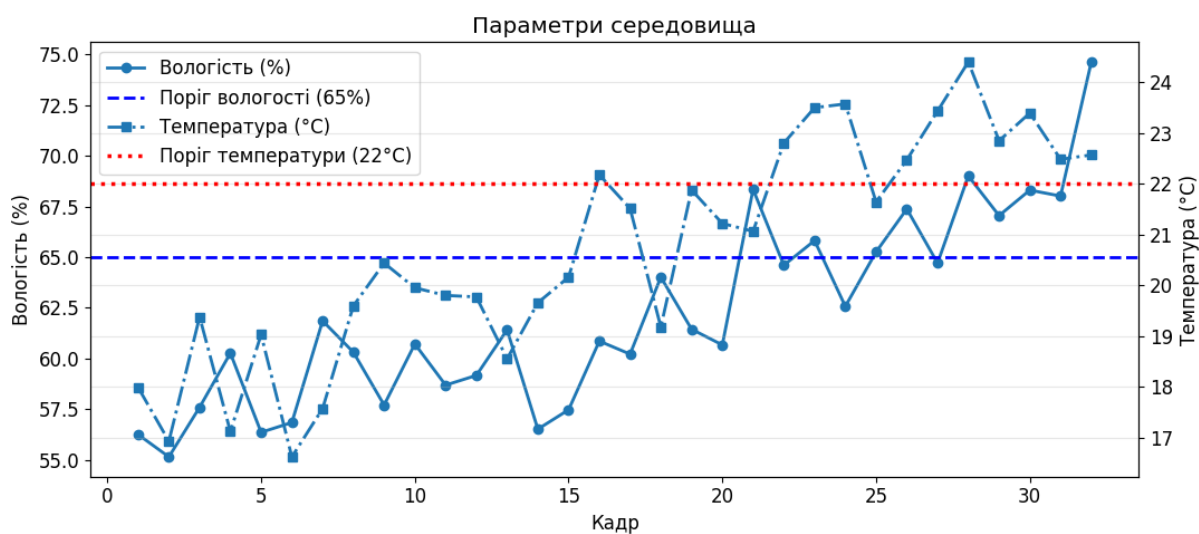


Рисунок 3 – Зміна вологості та температури по кадрах

Вибір порогів у 65% для вологості та 22°C для температури зазвичай базується на стандартах експлуатації та точці роси. Вологість 65% часто є критичною межею, вище якої активуються процеси гігроскопічної конденсації та інтенсивної атмосферної корозії. При перевищенні цього рівня на поверхні металу утворюється мікроскопічна плівка води, що є каталізатором окислення.

Температура 22°C у поєднанні з вологістю часто є основною точкою для розрахунку вологоутримання повітря. Для багатьох конструкційних матеріалів саме цей

діапазон вважається межею між «стабільним» та «активним» тепловим режимом, що впливає на в'язкість захисних покриттів та напружено-деформований стан.

Після попередньої обробки даних було здійснено навчання моделей із використанням гіперпараметричної оптимізації для ансамблевих алгоритмів.

*Побудова моделі класифікації технічного стану.* Отримані результати свідчать про те, що більшість використаних моделей успішно відтворюють логіку класифікації технічного стану на основі сформованих ознак. Особливо важливо, що моделі не лише правильно розпізнавали клас «Норма», але й коректно ідентифікували приклади класу «Попередження», який є проміжним та більш складним для розмежування. Це вказує на достатню інформативність ознаки DefectRatio, а також на допоміжну роль параметрів вологості та температури в задачі оцінювання стану.

Окремий інтерес становлять результати моделі LightGBM, яка виявилася менш стабільною в межах даної вибірки. Аналіз показників precision, recall та F1-score показав, що ця модель коректно класифікувала лише зразки класу «Норма», однак не змогла правильно розпізнати жодного зразка класу «Попередження». Це може бути пояснено дуже малою кількістю прикладів цього класу в тестовій вибірці, а також загалом невеликим обсягом навчальних даних, що є типовим обмеженням для задач із вузькою експериментальною вибіркою. Для наочного порівняння ефективності алгоритмів було побудовано стовпчикову діаграму значень точності.

Візуальний аналіз підтверджує, що чотири моделі демонструють ідентично високий рівень класифікації, тоді як LightGBM поступається їм за якістю розпізнавання. Значення основних метрик класифікації наведено в Табл. 1.

Таблиця 1

Порівняння точності моделей класифікації

Модель	Точність	Клас	Precision	Recall	F1-score	Support
RandomForest	0.9	Критичний	1.00	1.00	1.00	2
		Норма	0.86	1.00	0.92	6
		Попередження	1.00	0.50	0.67	2
XGBoost	0.9	Критичний	1.00	1.00	1.00	2
		Норма	0.86	1.00	0.92	6
		Попередження	1.00	0.50	0.67	2
LightGBM	0.8	Критичний	0.67	1.00	0.80	2
		Норма	0.86	1.00	0.92	6
		Попередження	0.00	0.00	0.00	2
SVM	0.7	Критичний	1.00	1.00	1.00	2
		Норма	0.80	0.67	0.73	6
		Попередження	0.33	0.50	0.40	2
LogReg	0.6	Критичний	0.00	0.00	0.00	2
		Норма	0.60	1.00	0.75	6
		Попередження	0.00	0.00	0.00	2

Для найкращої моделі було також побудовано кореляційну **матрицю помилок (Рис. 4)**, яка показала повну відповідність між реальними та передбаченими класами тестової вибірки. Усі зразки класу «**Норма**» та всі зразки класу «**Попередження**» були класифіковані без помилок. Це підтверджує високу здатність моделі до відокремлення нормального та потенційно проблемного стану поверхні в межах досліджуваної вибірки.

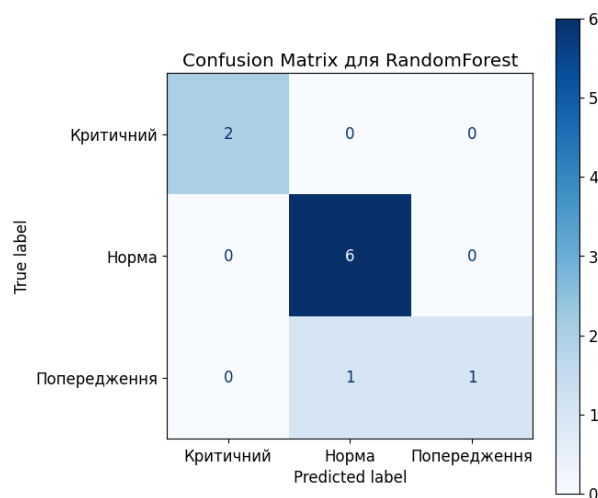


Рисунок 4 – Матриця помилок для Random Forest

*Гіперпараметрична валідація та інженерія ознак.* Оптимізація здійснювалася з використанням перехресної валідації, що дозволило оцінити узагальнювальну здатність кожної моделі ще до тестування на вибірці. Найвищі значення середньої точності перехресної валідації (Best CV score = 0.7738) продемонстрували моделі Random Forest та Logistic Regression, що свідчить про їхню найкращу здатність до стабільної класифікації в межах сформованого набору даних. Для моделі Random Forest оптимальними виявилися параметри: кількість дерев 200, максимальна глибина дерева 3 та мінімальна кількість об'єктів для розщеплення вузла 4.

Модель Logistic Regression також продемонструвала високий результат перехресної валідації, причому найкращим виявився параметр регуляризації  $C = 0.1$  при використанні розв'язувача liblinear. Це свідчить про те, що навіть відносно проста лінійна модель може бути ефективною для даного завдання, якщо структура даних є достатньо впорядкованою, а ознаки мають виражений розділяючий потенціал.

Модель SVM забезпечила дещо нижчий, але також достатньо високий результат (Best CV score = 0.7262). Найкращими параметрами стали радіальна базисна функція ядра (RBF),  $C = 1$  та  $\gamma = \text{scale}$ , що вказує на здатність методу ефективно враховувати нелінійні межі між класами. Це є очікуваним для задачі, де вплив ознак може мати не лише лінійний характер.

Модель XGBoost продемонструвала середній рівень узагальнювальної здатності (Best CV score = 0.6845). Найкращими параметрами стали 50 дерев, глибина 2 та  $\text{learning rate} = 0.05$ .

Отримані значення свідчать про те, що для даного набору даних більш доцільним виявився обережний бустинговий підхід із неглибокими деревами, орієнтований на зменшення ризику перенавчання. Найнижчий результат перехресної валідації було зафіксовано для моделі LightGBM (Best CV score = 0.369), для якої оптимальними виявилися параметри: learning rate = 0.05, 50 дерев та num\_leaves = 15. У досліджуваному випадку, ознака DefectRatio має найбільше середнє абсолютне SHAP-значення, це свідчить про те, що саме вона робить найбільший внесок у зміну прогнозу моделі. Отже, рішення моделі щодо належності об'єкта до класів «Норма», «Попередження» або «Критичний» насамперед визначається ступенем дефектності поверхні.

Для інтерпретації рішень моделі було застосовано SHAP-аналіз (SHapley Additive exPlanations), який дозволяє оцінити внесок кожної ознаки у формування прогнозу. У результаті вдалося побудувати графік (Рис. 5) на основі SHAP, що забезпечує глобальне уявлення про вплив ознак на класифікацію.

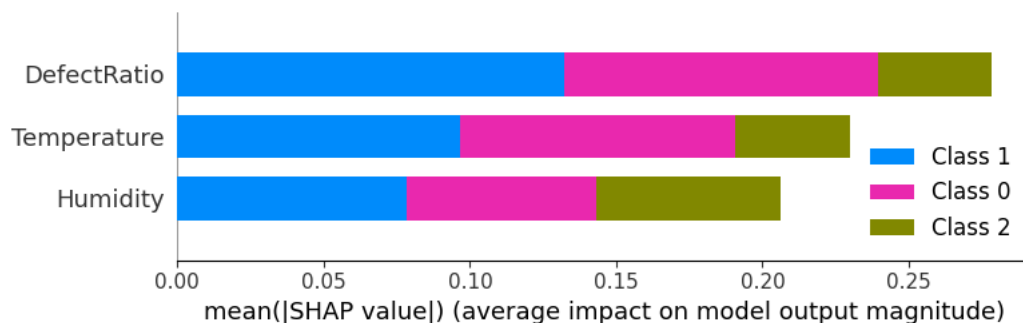


Рисунок 5 – SHAP-візуалізація глобального впливу ознак на класифікацію технічного стану інженерної поверхні

У досліджуваному випадку, ознака DefectRatio має найбільше середнє абсолютне SHAP-значення, це свідчить про те, що саме вона робить найбільший внесок у зміну прогнозу моделі. Отже, рішення моделі щодо належності об'єкта до класів «Норма», «Попередження» або «Критичний» насамперед визначається ступенем дефектності поверхні.

З практичної точки зору важливим є не лише результат класифікації, а й можливість пояснення прийнятого рішення. Для цього було використано два підходи: **Permutation Importance** та **rule-based пояснення**. Аналіз важливості ознак показав, що найбільший внесок у прийняття рішення забезпечує показник **DefectRatio** (0.290), тобто частка площі, зайнятої дефектами. Саме ця ознака виявилася визначальною для розмежування технічних станів поверхні. Водночас параметри **вологості (0.055)** та **температури (0.03)** мали додатковий, але менший вплив, що свідчить про їхню роль як контекстних факторів середовища.

Порівняння реальних і прогнозованих станів для тестової вибірки

№	Рівень дефектності поверхні (DefectRatio)	Відносна вологість, %	Температура, °C	Реальний стан	Прогнозований стан
1	0.00489	61.85	17.57	Попередження	Попередження
2	0.00000	61.44	21.87	Попередження	Норма
3	0.00000	61.41	18.55	Норма	Норма
4	0.00920	56.24	17.98	Критичний	Критичний
5	0.00000	60.66	21.22	Норма	Норма
6	0.00008	57.46	20.16	Норма	Норма
7	0.00000	60.85	22.17	Норма	Норма
8	0.00000	74.63	22.56	Норма	Норма
9	0.00921	60.26	17.12	Критичний	Критичний
10	0.00000	68.30	23.39	Норма	Норма

Аналіз таблиці тестових прогнозів показує, що модель Random Forest у більшості випадків коректно відтворює реальний технічний стан спостережень. Зокрема, модель правильно класифікувала всі приклади класу «Норма», а також обидва приклади класу «Критичний», для яких спостерігалися підвищені значення DefectRatio (0.00920 та 0.00921). Це підтверджує високу чутливість моделі до об'єктів із вираженими дефектними ознаками. Водночас було виявлено один випадок помилкової класифікації, коли об'єкт із реальним класом «Попередження» був віднесений моделлю до класу «Норма». Для цього прикладу значення DefectRatio = 0.00000, а значення вологості (61.44%) та температури (21.87°C) також не перевищували критичних порогів. Це може свідчити про те, що даний об'єкт належить до прикордонних випадків, де стан визначається не настільки виражено, як у прикладах із очевидною дефектністю. У таких ситуаціях модель може схилитися до більш безпечного або статистично домінуючого класу.

Загалом, результати підтверджують, що модель приймає рішення на основі змістовно обґрунтованих фізичних параметрів, що є критично важливим для систем пояснювального штучного інтелекту (ХАІ).

*Правило-орієнтовані пояснення.* Пояснення для окремого тестового прикладу з Табл. 2 підтвердило виявлені моделлю закономірності. Для першого часового кроку з параметрами середовища тестової вибірки (№1) модель передбачила стан «Попередження», що повністю збігається з фактичним значенням. Основним чинником такого рішення було визначено помірний рівень дефектності поверхні (DefectRatio = 0.00489). При цьому показники вологості (61.85%) та температури (17.57°C) не перевищували встановлені критичні пороги, отже, не стали визначальними факторами при класифікації. Це демонструє здатність моделі ідентифікувати перехідний стан системи на основі візуальних дефектів навіть за умови відносно безпечних параметрів довкілля.

Для цього випадку прогноз моделі показав стан «попередження» з визначеною причиною: «Помірний рівень дефектності поверхні», що відповідає Табл. 3 з правилами пояснення прогнозу.

Таблиця 3

Правила пояснення прогнозу моделі

№	Умова	Інтерпретація	Пояснення
1	Рівень дефектності поверхні $> 0.01$	Високий рівень дефектності поверхні	Багато дефектів на поверхні
2	$0.001 < \text{Рівень дефектності поверхні} \leq 0.01$	Помірний рівень дефектності поверхні	Помірний рівень дефектності поверхні
3	Вологість $> 65\%$	Підвищена вологість середовища	Висока вологість
4	Температура $> 22^{\circ}\text{C}$	Підвищена температура середовища	Підвищена температура
5	Жодна умова не виконується	Відсутність виражених ризикових факторів	Немає критичних факторів

Узагальнюючи отримані результати, можна стверджувати, що побудована система класифікації здатна ефективно визначати технічний стан поверхні на основі поєднання візуальних ознак дефектності та сенсорних параметрів середовища. Разом із тим слід враховувати, що висока точність у даному експерименті частково пояснюється невеликим обсягом даних та чітко вираженою залежністю класів від показника DefectRatio. Тому подальші дослідження доцільно спрямувати на розширення вибірки, включення більшої кількості прикладів проміжних і критичних станів, а також використання додаткових ознак, що дозволить підвищити узагальнювальну здатність моделей.

Результати тестування демонструють, що модель найбільш упевнено розпізнає чітко виражені нормальні та критичні стани, тоді як найбільша складність виникає при класифікації проміжних або слабко виражених попереджувальних станів. Це є типовим явищем для задач технічної діагностики, де саме перехідні стани часто є найбільш складними для автоматизованого виявлення.

Таким чином, локальний SHAP-аналіз дозволяє пояснити не лише загальні закономірності моделі, а й **індивідуальні рішення для кожного окремого випадку**, що є надзвичайно важливим у системах підтримки прийняття рішень.

Сукупність отриманих результатів свідчить про те, що запропонований метод є ефективним як з точки зору **точності класифікації**, так і з точки зору **пояснюваності рішень**. На відміну від класичних підходів, де модель виконує лише автоматичну класифікацію, використання пояснювального штучного інтелекту дозволяє:

- визначати найбільш значущі фактори ризику;
- формувати зрозумілі текстові пояснення;
- аналізувати причини конкретних рішень моделі;
- підвищувати довіру до автоматизованих систем моніторингу.

Практична цінність запропонованого підходу полягає в можливості його застосування у складі інтелектуальних систем моніторингу трубопроводів, мостів, несучих конструкцій, резервуарів, промислових об'єктів та інших інженерних систем. У таких умовах система може не лише автоматично сигналізувати про погіршення стану, але й надавати оператору обґрунтування, які саме ознаки вказують на потенційну небезпеку.

Окремо слід зазначити, що запропонований підхід є масштабованим і може бути розширений за рахунок включення нових джерел даних, таких як вібраційні сигнали, акустичні емісії, термографічні зображення або результати ультразвукового контролю. Це відкриває перспективи для побудови більш складних інтелектуальних систем структурного моніторингу на основі мультимодальної інтеграції даних та пояснювального штучного інтелекту.

Для забезпечення зручної взаємодії користувача з розробленим методом оцінювання стану елементу інженерної конструкції було реалізовано програмне середовище аналізу, яке забезпечує обробку мультимодальних даних, класифікацію технічного стану та формування пояснень до отриманих результатів. Вищеописаний підхід дозволяє виконувати аналіз без необхідності ручного опрацювання числових показників і створює основу для подальшої інтеграції методу у складі інтелектуальних систем структурного моніторингу.

У процесі реалізації програмної частини було використано мову програмування високого рівня Python, що забезпечила гнучкість розробки, зручність роботи з науковими бібліотеками та можливість швидкого прототипування компонентів системи. Як середовище розробки застосовувалося Visual Studio Code, що забезпечило зручну організацію програмного коду, модульну структуру реалізації та ефективну інтеграцію із засобами аналізу даних і машинного навчання. Реалізація програмного рішення здійснювалася з використанням бібліотек NumPy, Pandas, Matplotlib, Scikit-learn та SHAP. Застосування зазначених інструментів дозволило реалізувати повний цикл обробки даних: від формування модельованого мультимодального набору спостережень до побудови моделі класифікації технічного стану та інтерпретації її рішень. Уся логіка аналізу була організована у вигляді окремих функціональних модулів, що включають генерацію та попередню обробку даних, формування інтегральних ознак, навчання моделі, оцінювання результатів та побудову пояснень.

Завдяки використанню аналізу важливості ознак, правило-орієнтованих пояснень та SHAP-інтерпретацій, система дозволяє користувачу не лише отримати фінальний клас стану, але й визначити, які саме параметри мали найбільший вплив на прийняте рішення. Це є важливою перевагою у порівнянні з традиційними «чорними скриньками» машинного навчання. Тестування запропонованого рішення проводилося відповідно до поставлених завдань і включало перевірку коректності формування даних, стабільності навчання моделі, точності класифікації, а також адекватності згенерованих пояснень.

У процесі експериментальної перевірки було підтверджено, що програмна реалізація коректно виконує всі основні етапи аналізу: формування ознакового простору,

обчислення інтегральних показників, класифікацію технічного стану, побудову графічних залежностей, матриці помилок, оцінювання важливості ознак та локальних пояснень для окремих прикладів.

За результатами випробувань встановлено, що розроблена система функціонує стабільно, забезпечує високий рівень автоматизації процесу оцінювання стану конструкції та демонструє необхідний рівень точності й інтерпретованості.

Додатково було підтверджено, що використаний підхід може бути реалізований на обчислювальних пристроях середнього рівня без потреби у спеціалізованому графічному обладнанні, що робить його придатним для практичного застосування у задачах технічної діагностики та структурного моніторингу.

*Обмеження та подальші дослідження.* Разом із тим, запропонований підхід має певні обмеження, які необхідно враховувати під час інтерпретації отриманих результатів. Насамперед, дослідження було проведено на невеликій вибірці даних, сформованій на основі обмеженої кількості спостережень, що може впливати на узагальнювальну здатність побудованих моделей. Крім того, цільова змінна технічного стану визначалася на основі заздалегідь встановлених порогових правил, пов'язаних переважно з показником DefectRatio, унаслідок чого модель значною мірою відтворює закладену логіку класифікації. Це пояснює отримання дуже високих показників точності, однак водночас обмежує можливість робити висновки щодо поведінки моделі в умовах реальних, більш складних і шумових даних. Додатковим обмеженням є те, що в дослідженні враховано лише три вхідні ознаки частку дефектної області, температуру та вологість, тоді як у практичних задачах технічної діагностики стан конструкції може залежати також від геометричних характеристик дефекту, динаміки його розвитку, механічних навантажень, вібраційних параметрів та інших експлуатаційних чинників. У зв'язку з цим подальші дослідження доцільно спрямувати на розширення вибірки, використання реальних експериментальних даних, включення додаткових ознак та розширення шкали технічного стану, що дозволить підвищити надійність, стійкість та прикладну цінність запропонованої системи.

**Висновки.** У роботі розроблено метод до оцінювання стану елемента інженерної конструкції з використанням методів машинного навчання та пояснювального штучного інтелекту, що базується на аналізі мультимодальних даних, які поєднують візуальні ознаки дефектності поверхні та параметри зовнішнього середовища. У межах дослідження було сформовано набір даних, що включав частку дефектної області поверхні (DefectRatio), а також вологість і температуру. На основі цих ознак здійснювалося автоматизоване визначення технічного стану конструкції за класами «Норма», «Попередження» та «Критичний». Для розв'язання задачі класифікації було досліджено кілька алгоритмів машинного навчання, зокрема Random Forest, XGBoost, LightGBM, Support Vector Machine (SVM) та Logistic Regression, із застосуванням гіперпараметричної оптимізації та перехресної валідації. За результатами експериментального дослідження встановлено, що найкращі показники на тестовій вибірці продемонстрували моделі Random Forest та XGBoost, для яких точність класифікації становила 0.90. Модель Logistic Regression забезпечила точність 0.80, SVM — 0.70, тоді як LightGBM про-

демонструвала найнижчий результат (0.60). З урахуванням якості класифікації та стабільності результатів перехресної валідації найбільш доцільною для подальшого використання було визначено модель Random Forest.

Проведений аналіз показав, що найбільший вплив на результат класифікації має ознака DefectRatio, яка характеризує частку площі поверхні, охопленої дефектами. За результатами Permutation Importance її важливість становила 0.290, що суттєво перевищує внесок ознак Humidity (0.055) та Temperature (0.030). Це свідчить про те, що саме ступінь дефектності поверхні є основним критерієм для розмежування технічних станів, тоді як параметри зовнішнього середовища виконують допоміжну, контекстну роль. Особливістю запропонованого підходу є використання засобів пояснювального штучного інтелекту (ХАІ), які дозволили підвищити прозорість, інтерпретованість та обґрунтованість прийнятих рішень. У роботі застосовано Permutation Importance, SHAP-аналіз та правило-орієнтовані пояснення, що дало змогу не лише оцінити глобальний вплив ознак на результат класифікації, але й пояснити причини віднесення окремого прикладу до конкретного класу. Зокрема, локальний аналіз показав, що для тестового прикладу зі значенням DefectRatio = 0.00489 модель обґрунтовано віднесла об'єкт до класу «Попередження», оскільки саме помірний рівень дефектності поверхні виступив основною причиною такого рішення.

Отримані результати підтвердили, що запропонований підхід є придатним для побудови інтелектуальних систем моніторингу технічного стану інженерних конструкцій, у яких важливими є не лише точність автоматизованої оцінки, а й зрозумілість логіки прийняття рішення. Водночас проведене дослідження має певні обмеження, пов'язані з невеликим обсягом вибірки, частковим дисбалансом класів та модельованим характером даних, що могло вплинути на стабільність оцінювання. У зв'язку з цим подальші дослідження доцільно спрямувати на розширення набору ознак, збільшення кількості реальних спостережень, а також удосконалення механізмів виявлення критичних станів у реальних умовах експлуатації конструкцій.

#### ЛІТЕРАТУРА

1. Mysiuk R., Yuzevych V., Mysiuk I. et al. Detection of surface defects inside concrete pipelines using trained model on JetRacer kit // *In 2023 IEEE 13th International Conference on Electronics and Information Technologies (ELIT)*, 2023, pp. 21–24. IEEE., doi: 10.1109/ELIT61488.2023.10310691
2. Mysiuk R., Mysiuk I., Pawlowski G. et al. Video-based concrete road damage assessment using JetRacer kit // *In 2023 17th International Conference on the Experience of Designing and Application of CAD Systems (CADSM)*. 2023. pp. 1–4. IEEE., doi: <https://doi.org/10.1109/CADSM58174.2023.10076528>.
3. Mysiuk, R. (2024). Towards Information Flows in Recognition and Prediction Tasks with Internet of Things // *Path of Science*. 2024. Vol. 10(1). 1001-1004. DOI: <http://dx.doi.org/10.22178/pos.100-10>

4. Mysiuk R., Yuzevych V. IoT-based solution for detection defects in infrastructure objects using Raspberry Pi // *Electronics and Information Technologies*. 2023. Vol. 21. 45–56. 10.30970/eli.21.5
5. Moreh F., Lyu H., Rizvi Z. H. et al. Deep neural networks for crack detection inside structures // *Scientific Reports*. 2024. Vol. 14. Article 4439. DOI: <https://doi.org/10.1038/s41598-024-54494-y>.
6. Shalaby Y. M., Badawy M., Ebrahim G. A. et al. Condition assessment of concrete structures using automated crack detection method for different concrete surface types based on image processing // *Discover Civil Engineering*. 2024. Vol. 1. Article 81. DOI: <https://doi.org/10.1007/s44290-024-00089-5>.
7. Roy S., Yogi B., Majumdar R. et al. Deep learning-based crack detection and prediction for structural health monitoring // *Discover Applied Sciences*. 2025. Vol. 7. Article 674. DOI: <https://doi.org/10.1007/s42452-025-07272-y>.
8. Shah Mansouri T., Lubarsky G., Finlay D., McLaughlin J. Machine learning-based structural health monitoring technique for crack detection and localisation using Bluetooth strain gauge sensor network // *Journal of Sensor and Actuator Networks*. 2024. Vol. 13, No. 6. Article 79. DOI: <https://doi.org/10.3390/jsan13060079>.
9. Barbosh M., Ge L., Sadhu A. Automated crack identification in structures using acoustic waveforms and deep learning // *Journal of Infrastructure Preservation and Resilience*. 2024. Vol. 5. Article 10. DOI: <https://doi.org/10.1186/s43065-024-00102-2>.
10. Spencer B. F., Sim S.-H., Kim R. E., Yoon H. Advances in artificial intelligence for structural health monitoring: A comprehensive review // *KSCE Journal of Civil Engineering*. 2025. Vol. 29, Issue 3. DOI: <https://doi.org/10.1016/j.kscej.2025.100203>.
11. Si H., Wang Q., Ruan X. et al. Framework for investigating structure cracking using real engineering data combined with physics constraints // *Scientific Reports*. 2025. Vol. 15. Article 6344. DOI: <https://doi.org/10.1038/s41598-024-85079-4>.
12. Sun Z., Chen T., Meng X., Bao Y., Hu L., Zhao R. A critical review for trustworthy and explainable structural health monitoring and risk prognosis of bridges with human-in-the-loop // *Sustainability*. 2023. Vol. 15, No. 8. Article 6389. DOI: <https://doi.org/10.3390/su15086389>.
13. Chen Q., Li B. Explainable artificial intelligence (XAI)-driven probabilistic image-based structural health monitoring of reinforced concrete beams with shear reinforcements // *Automation in Construction*. 2025. Vol. 180. DOI: <https://doi.org/10.1016/j.autcon.2025.106549>.

## REFERENCES

1. Mysiuk, R., Yuzevych, V., Mysiuk, I., Tyrkalo, Y., Pavlenchyk, A., & Dalyk, V. (2023). Detection of surface defects inside concrete pipelines using trained model on JetRacer kit. In 2023 IEEE 13th International Conference on Electronics and Information Technologies (ELIT) (pp. 21–24). IEEE., doi: <https://doi.org/10.1109/ELIT61488.2023.10310691>
2. Mysiuk, R., Mysiuk, I., Pawlowski, G., Yuzevych, V., Yasynskyi, M., & Tyrkalo, Y. (2023). Video-based concrete road damage assessment using JetRacer kit. In 2023 17th International Conference on the Experience of Designing and Application of CAD Systems (CADSM) (pp. 1–4). IEEE., doi: <https://doi.org/10.1109/CADSM58174.2023.10076528>.

3. Mysiuk, R. (2024). Towards Information Flows in Recognition and Prediction Tasks with Internet of Things. *Path of Science*, 10(1), 1001-1004. doi: <https://doi.org/10.22178/pos.100-10>
4. Mysiuk, R., & Yuzevych, V. (2023). IoT-based solution for detection defects in infrastructure objects using Raspberry Pi. *Electronics and Information Technologies*, 21, 45–56. [10.30970/eli.21.5](https://doi.org/10.30970/eli.21.5)
5. Moreh, F., Lyu, H., Rizvi, Z. H., et al. (2024). Deep neural networks for crack detection inside structures. *Scientific Reports*, 14, 4439. <https://doi.org/10.1038/s41598-024-54494-y>
6. Shalaby, Y. M., Badawy, M., Ebrahim, G. A., et al. (2024). Condition assessment of concrete structures using automated crack detection method for different concrete surface types based on image processing. *Discover Civil Engineering*, 1, 81. <https://doi.org/10.1007/s44290-024-00089-5>
7. Roy, S., Yogi, B., Majumdar, R., et al. (2025). Deep learning-based crack detection and prediction for structural health monitoring. *Discover Applied Sciences*, 7, 674. <https://doi.org/10.1007/s42452-025-07272-y>
8. Shah Mansouri, T., Lubarsky, G., Finlay, D., & McLaughlin, J. (2024). Machine learning-based structural health monitoring technique for crack detection and localisation using Bluetooth strain gauge sensor network. *Journal of Sensor and Actuator Networks*, 13(6), 79. <https://doi.org/10.3390/jsan13060079>
9. Barbosh, M., Ge, L., & Sadhu, A. (2024). Automated crack identification in structures using acoustic waveforms and deep learning. *Journal of Infrastructure Preservation and Resilience*, 5, 10. <https://doi.org/10.1186/s43065-024-00102-2>
10. Spencer, B. F., Sim, S.-H., Kim, R. E., & Yoon, H. (2025). Advances in artificial intelligence for structural health monitoring: A comprehensive review. *KSCE Journal of Civil Engineering*, 29(3). <https://doi.org/10.1016/j.kscej.2025.100203>
11. Si, H., Wang, Q., Ruan, X., et al. (2025). Framework for investigating structure cracking using real engineering data combined with physics constraints. *Scientific Reports*, 15, 6344. <https://doi.org/10.1038/s41598-024-85079-4>
12. Sun, Z., Chen, T., Meng, X., Bao, Y., Hu, L., & Zhao, R. (2023). A critical review for trustworthy and explainable structural health monitoring and risk prognosis of bridges with human-in-the-loop. *Sustainability*, 15(8), 6389. <https://doi.org/10.3390/su15086389>
13. Chen, Q., & Li, B. (2025). Explainable artificial intelligence (XAI)-driven probabilistic image-based structural health monitoring of reinforced concrete beams with shear reinforcements. *Automation in Construction*, 180. <https://doi.org/10.1016/j.autcon.2025.106549>

Received 10.04.2026.  
Accepted 13.04.2026.  
Published 30.04.2026

***Method for assessing the condition of an engineering structure element  
using explanatory artificial intelligence and multimodal data***

*Assessing the technical condition of engineering structural elements is an important task for ensuring operational safety, reliability, and timely maintenance. Conventional inspection approaches are often labor-intensive, time-consuming, and dependent on expert judgment. This study proposes an intelligent approach for structural condition assessment based on machine learning and explainable artificial intelligence (XAI) using multimodal da-*

*ta that combine visual and environmental information. In the developed approach, the condition of a structural surface is evaluated using three input features: the defect area ratio (DefectRatio), humidity, and temperature. Based on these features, the structural state is automatically classified into three condition categories: Normal, Warning, and Critical. To solve the classification task, several machine learning algorithms were investigated, including Random Forest, XGBoost, LightGBM, Support Vector Machine (SVM), and Logistic Regression, with the application of hyperparameter optimization and cross-validation. Experimental results showed that Random Forest and XGBoost achieved the highest classification accuracy on the test dataset (0.90), while Logistic Regression, SVM, and LightGBM demonstrated lower performance, with accuracies of 0.80, 0.70, and 0.60, respectively. The obtained results indicate that the selected feature set is sufficiently informative for distinguishing between different structural condition states, while also revealing differences in the generalization ability of the investigated models. To improve transparency and interpretability, the proposed approach incorporates explainable artificial intelligence techniques, including Permutation Importance, SHAP-based analysis, and rule-based explanations. These methods made it possible to identify the most influential features and explain the reasoning behind model predictions. The analysis confirmed that DefectRatio is the dominant factor in structural condition assessment, whereas humidity and temperature play a supporting contextual role. The proposed approach was implemented in Python using libraries such as NumPy, Pandas, Matplotlib, Scikit-learn, XGBoost, LightGBM, and SHAP. The obtained results demonstrate the feasibility of combining machine learning, multimodal analysis, and XAI for interpretable and automated structural condition monitoring, which can be useful for preventive diagnostics, risk assessment, and decision support in engineering applications..*

*Keywords: explanatory artificial intelligence, feature engineering, multimodal data processing, image processing, intelligent monitoring, ML classification, risk analysis, surface defects, crack.*

**Мисюк Роман Володимирович** – доктор філософії, доцент кафедри системного проектування Львівського національного університету імені Івана Франка.

ORCID: <https://orcid.org/0000-0002-7843-7646>

**Кулик Петро Русланович** – асистент кафедри системного проектування Львівського національного університету імені Івана Франка.

ORCID: <https://orcid.org/0009-0008-1193-8578>

**Гусак Олег Васильович** – асистент кафедри системного проектування Львівського національного університету імені Івана Франка.

ORCID: <https://orcid.org/0009-0000-6026-4081>

**Чмихало Олександр Сергійович** – асистент кафедри системного проектування Львівського національного університету імені Івана Франка.

ORCID: <https://orcid.org/0009-0001-5409-5119>

**Парубочий Віталій Олегович** – асистент кафедри системного проектування Львівського національного університету імені Івана Франка.

ORCID: <https://orcid.org/0000-0002-9847-6968>

**Mysiuk Roman** – PhD (Computer Science), associate professor department of system design, Ivan Franko National University of Lviv, Lviv, Ukraine.

ORCID: <https://orcid.org/0000-0002-7843-7646>

**Kulyk Petro** – assistant department of system design, Ivan Franko National University of Lviv, Lviv, Ukraine.

ORCID: <https://orcid.org/0009-0008-1193-8578>

**Husak Oleg** – assistant department of system design, Ivan Franko National University of Lviv, Lviv, Ukraine.

ORCID: <https://orcid.org/0009-0000-6026-4081>

**Chmyhalo Oleksandr**– assistant department of system design, Ivan Franko National University of Lviv, Lviv, Ukraine..

ORCID: <https://orcid.org/0009-0001-5409-5119>

**Parubochyi Vitalii** – assistant department of system design, Ivan Franko National University of Lviv, Lviv, Ukraine.

ORCID : <https://orcid.org/0000-0002-9847-6968>

## МЕТОД ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ДЛЯ ОПТИМІЗАЦІЇ ПРОДУКТИВНОСТІ БАЗ ДАНИХ

*Анотація.* Сучасні підходи до оптимізації продуктивності баз даних переважно орієнтовані на окремі аспекти – індексацію, оптимізацію SQL-запитів або налаштування серверного середовища, що змушує адміністраторів використовувати декілька різних інструментів одночасно та підвищує ймовірність пропуску критичних недоліків. Невирішеною залишається проблема комплексного автоматизованого аналізу бази даних без необхідності прямого підключення до сервера. У роботі запропоновано комбінований алгоритмічно-програмний метод оптимізації продуктивності баз даних, що відрізняється від існуючих підходів можливістю комплексного офлайн-аналізу SQL-дамів та конфігураційних файлів без прямого підключення до СУБД. Метод ґрунтується на статичному аналізі SQL-дамів та конфігураційних файлів без розгортання бази даних та інтегрує три незалежних аналітичних модулі: аналіз індексації, структурний аналіз та аналіз параметрів серверної конфігурації, а також окремий модуль інтерактивного рефакторингу SQL-запитів у реальному часі. Для кількісної оцінки ефективності оптимізації запропоновано комплексний індекс продуктивності (КІП), що об'єднує три часткових показники – індексаційний, структурний та конфігураційний з визначеними ваговими коефіцієнтами. Метод реалізовано у вигляді клієнт-серверного вебзастосування з рольовою моделлю розмежування доступу. За рахунок використання запропонованого методу досягається комплексна автоматизована оптимізація бази даних та формування рекомендацій для адміністратора без значних додаткових витрат на адміністрування.

*Ключові слова:* оптимізація баз даних, SQL-дамп, статичний аналіз, індексація, структурний аналіз, конфігурація сервера, комплексний індекс продуктивності, FastAPI, React, СУБД.

**Постановка проблеми.** Сьогодні інформаційні системи оперують величезними обсягами даних, і продуктивність баз даних є ключовим чинником ефективності будь-якого програмного комплексу. Навіть незначні затримки у виконанні SQL-запитів або помилки у конфігурації серверного середовища можуть призводити до суттєвого зниження швидкодії, втрати даних або перевитрати ресурсів серверів.

У працях [1–3] розглянуто підходи до оптимізації запитів, структури даних, а також використання хешування та індексації. Існуючі підходи до оптимізації переважно орієнтовані на окремі аспекти: одні інструменти спеціалізуються на аналізі індексів, інші на перевірці структури схеми, треті на налаштуванні серверних параметрів. Для

проведення комплексного аналізу зазвичай необхідно використовувати декілька різних інструментів, що ускладнює процес і збільшує ймовірність пропуску критично важливих недоліків у структурі чи конфігурації бази даних. Крім того, більшість існуючих рішень потребують безпосереднього підключення до живого сервера або виконання профілюючих запитів для збору статистики, що унеможливує їх застосування там, де прямий доступ є технічно недоступним або небажаним з міркувань безпеки.

Таким чином, невирішеною залишається задача побудови універсального методу комплексного аналізу продуктивності баз даних на основі статичних артефактів без необхідності прямого доступу до СУБД, що і визначає напрям дослідження.

**Аналіз останніх досліджень і публікацій.** Проблема підвищення продуктивності баз даних активно досліджується як у промислових, так і в академічних колах.

Структурні методи орієнтовані на нормалізацію бази даних, усунення надлишковості та підтримку цілісності даних. Вбудовані оптимізатори PostgreSQL та MySQL пропонують аналіз ключів і статистики таблиць, однак вони не завжди враховують логіку бізнес-процесів і зв'язки між запитом. У роботі [4] проведено систематичний огляд методів оптимізації баз даних і показано, що жоден з існуючих підходів не охоплює комплексно всі аспекти продуктивності одночасно, а ефективність структурних методів у великих схемах залишається обмеженою.

Методи оптимізації SQL-запитів, представлені у працях [1, 2, 5], базуються на перетворенні запитів, використанні індексів, реорганізації об'єднань (JOIN) і заміні вкладених підзапитів. Зокрема, у [1] запропоновано підходи до побудови самоналаштовуваних систем керування базами даних, а у [2] – ефективний інструмент автоматичного вибору індексів на основі аналізу навантаження. У дослідженні [6] експериментально підтверджено, що застосування методів оптимізації до реляційних баз MySQL дозволяє суттєво скоротити час виконання запитів, зокрема за рахунок коректної індексації та реструктуризації схеми. Проте зазначені методи потребують попереднього профілювання або прямого доступу до СУБД, що унеможливує їх застосування при офлайн-аналізі SQL-дампів.

Серверна оптимізація передбачає регулювання параметрів конфігурації. У дослідженнях [7, 8] розглянуто методи машинного навчання для автоматичного підбору параметрів, зокрема для рушія InnoDB, а у [9] досліджено вплив конфігураційних налаштувань MySQL на продуктивність у різних сценаріях навантаження. Систематичний огляд [10] показав, що системи автоматичного налаштування схеми та параметрів баз даних потребують тривалої фази навчання на реальному навантаженні і не надають рекомендацій, придатних для безпосереднього застосування адміністратором.

Таким чином, жоден із наявних методів не вирішує комплексно задачу статичного аналізу SQL-дампів із наданням інтегрованих рекомендацій одразу за кількома напрямками без необхідності підключення до сервера.

**Мета дослідження.** Метою даного дослідження є підвищення продуктивності обробки запитів у базах даних шляхом розроблення комбінованого алгоритмічно-програмного методу та відповідного програмного забезпечення. Наукова новизна робо-

ти полягає у розробленні алгоритмічно-програмного методу аналітично-рекомендаційної оптимізації баз даних, який відрізняється від відомих підходів можливістю статичного аналізу SQL-дамів та конфігураційних файлів без прямого підключення до СУБД. Метод інтегрує три аналітичні модулі – індексаційний, структурний та конфігураційний в єдину систему кількісного оцінювання, а також окремий модуль інтерактивного рефакторингу SQL-запитів у реальному часі. Для об'єктивного вимірювання ефекту оптимізації введено комплексний індекс продуктивності (КІП) як інтегральний показник, що об'єднує результати трьох аналітичних модулів і дозволяє кількісно порівнювати стан бази даних до та після застосування рекомендацій.

**Викладення основного матеріалу дослідження.** Одним із найпопулярніших прикладних протоколів в мережі Інтернет є протокол HTTP(S) для вебсторінок, оскільки зазвичай саме вебсайт є кінцевим інтерфейсом для отримання інформації чи надання послуг. Тому запропонований метод розроблявся для протоколу HTTP(S).

Запропонований метод оптимізації продуктивності баз даних побудовано на статичному аналізі вхідних артефактів без необхідності підключення до живого сервера СУБД. Основна ідея полягає у формуванні трьох груп показників, а саме індексних, структурних та конфігураційних, які нормалізуються в інтервалі  $[0; 1]$  і використовуються як для надання рекомендацій користувачу, так і для розрахунку підсумкового значення комплексного індексу продуктивності (КІП). Метод об'єднує три незалежні аналітичних модулі, кожен з яких орієнтований на окремий аспект функціонування бази даних. Загальну схему роботи методу наведено на рис. 1.



Рисунок 1 – Схема роботи запропонованого методу

Метод реалізується у вигляді послідовності етапів. На першому етапі зчитуються вхідні дані – SQL-дамп або конфігураційний файл СУБД. На другому етапі виконується статичний аналіз: синтаксичний розбір структури SQL-дампу, аналіз індексації та аналіз конфігураційних параметрів. Кожний показник масштабується до інтервалу  $[0; 1]$ , де 0 відповідає негативному стану, а 1 – оптимальному. На третьому етапі обчислюються часткові індекси I, S та C для кожної групи показників. На четвертому етапі на їх основі обчислюється інтегральний показник КІП. На останньому етапі система формує звіт із рекомендаціями.

*Модуль аналізу індексації (indexing).* Вхідними даними є SQL-дампи бази даних. Модуль виконує синтаксичний розбір дампу та перевіряє наявність первинних і зовнішніх ключів, дублікатів індексів та таблиць без індексації. Визначається кількість таблиць без жодного індексу ( $T_{noindex}$ ) та загальна кількість таблиць ( $T_{all}$ ). Частковий індекс модуля індексації обчислюється за формулою:

$$I = 1 - \frac{T_{noindex}}{T_{all}} \quad (1)$$

Значення  $I = 1$  свідчить про те, що всі таблиці мають щонайменше один індекс, тоді як низькі значення вказують на наявність таблиць, пошук у яких виконуватиметься повним скануванням.

*Модуль структурного аналізу (structure).* Вхідними даними також є SQL-дампи. Модуль перевіряє відповідність схеми вимогам нормальних форм, виявляє відсутність первинних ключів ( $T_{pk}$ ), застарілі або застарілі типи даних ( $T_{types}$ ) та надлишкові поля ( $T_{fk}$ ). Частковий індекс структурного модуля обчислюється як:

$$S = \frac{T_{pk} + T_{fk} + T_{types}}{3} \quad (2)$$

Додатково застосовується критерій відповідності третій нормальній формі: якщо частка атрибутів із транзитивними функціональними залежностями перевищує 0,3 від загальної кількості атрибутів, система надає рекомендації щодо декомпозиції відповідних таблиць.

*Модуль аналізу конфігурації (tuning).* Вхідними даними є текстовий конфігураційний файл СУБД (наприклад, my.cnf для MySQL). Модуль порівнює значення ключових параметрів – innodb\_buffer\_pool\_size, max\_connections, query\_cache\_size, tmp\_table\_size із рекомендованими діапазонами ( $C_{optimal}$ ). Частковий індекс конфігураційного модуля визначається як частка параметрів, що відповідають рекомендованим значенням:

$$C = \frac{C_{optimal}}{C_{total}} \quad (3)$$

*Модуль рефакторингу запитів (refactoring).* На відміну від попередніх модулів, цей модуль працює в інтерактивному режимі реального часу: користувач вводить SQL-вираз, а система миттєво надає рекомендації щодо його покращення. Аналізуються типові антипатерни: використання SELECT \*, відсутність обмежень WHERE у запитах на оновлення та видалення, надлишково складні вкладені підзапити, що можуть бути замінені конструкцією JOIN. Оскільки модуль працює з окремими виразами, а не з цілісним дампом, його результати не входять до розрахунку КПІ.

*Комплексний індекс продуктивності (КІП).* Результати трьох аналітичних модулів – індексації, структури та конфігурації інтегруються в єдиний показник за формулою:

$$CPI = w_1 \cdot I + w_2 \cdot S + w_3 \cdot C \quad (4)$$

$w_1, w_2, w_3$  – вагові коефіцієнти значущості, вагові коефіцієнти, що задовольняють умові  $w_1 + w_2 + w_3 = 1$ . На основі аналізу джерел [1–3, 5, 7, 8] та проведених експериментів визначено такі значення коефіцієнтів:  $w_1 = 0,40$ ,  $w_2 = 0,35$ ,  $w_3 = 0,25$ . Індексація отримала найбільшу вагу, оскільки безпосередньо впливає на швидкість вибірок і JOIN-

операцій. Структура визначає узгодженість і ефективність зберігання даних. Конфігурація забезпечує стабільність системи та ефективне використання ресурсів.

КІП нормується в діапазоні від 0 до 1, де значення до 0,3 свідчить про суттєві структурні недоліки, 0,3-0,6 – середній рівень оптимізації, 0,6-0,8 – високий рівень, а 0,8-1,0 – про добре оптимізовану систему. Отримане значення КІП дозволяє перевіряти стан бази даних та порівнювати його до та після оптимізації без фактичного виконання запитів, що особливо важливо при статичному аналізі SQL-дампів.

*Архітектура програмного забезпечення.* Програмну реалізацію методу виконано у вигляді багаторівневої клієнт-серверної системи. Серверна частина реалізує основну логіку обробки вхідних даних: синтаксичний розбір SQL-дампів, виконання аналітичних модулів та обчислення показників продуктивності. Клієнтська частина забезпечує інтерактивну взаємодію з користувачем – завантаження даних, перегляд сформованих рекомендацій, роботу з модулем рефакторингу в реальному часі та перегляд історії оптимізацій. Обмін даними між рівнями здійснюється через REST API у форматі JSON. Система підтримує рольову модель розмежування доступу (RBAC) з двома ролями – користувач та адміністратор.

Для перевірки ефективності запропонованого методу було проведено тестування на двох базах даних різного масштабу.

База DB1 містила 3 таблиці. Система виявила наступні проблеми: 1 таблицю без індексів, 1 таблицю без первинного ключа та занижене значення `tmp_table_size`. Після застосування рекомендацій було створено відсутні індекси, додано первинний ключ та скориговано конфігураційні параметри.

База DB2 (DemoCRM) містила 10 таблиць загальним обсягом близько 180 тис. записів. Система виявила наступні проблеми: 3 таблиці без індексів, 2 таблиці без первинних ключів, 5 вкладених підзапитів у SQL-запитах, а також низьке значення `innodb_buffer_pool_size` (512 МБ при 8 ГБ RAM). Після застосування рекомендацій було створено індекси на ключових полях, додано первинні ключі, замінено вкладені підзапити конструкцією JOIN та скориговано конфігураційні параметри.

Результати обчислення КІП до та після застосування рекомендацій наведено у таблиці 1.

Таблиця 1

Результати експериментальної перевірки методу

База даних	Кількість таблиць	КІП до	КІП після	Покращення, %
DB1 (тестова)	3	0,52	0,7	+35
DB2 (DemoCRM)	10	0,65	0,85	+31

Експериментальні результати демонструють покращення значення КІП у середньому на 30–35%, що підтверджує ефективність запропонованого методу.

**Висновки.** У представленій роботі запропоновано метод оцінювання продуктивності баз даних на основі статичного аналізу SQL-дампів. Запроваджено новий показник

ник комплексний індекс продуктивності, який дозволяє кількісно оцінити ступінь ефективності структури та конфігурації бази даних без її виконання. Розроблений програмний комплекс формує рекомендації для покращення індексації, структури таблиць і параметрів конфігурації. Експериментальна перевірка підтвердила середнє зростання КІП на 30–35% після їх застосування. Практична значущість роботи полягає у можливості застосування запропонованого методу для попереднього аналізу баз даних без доступу до продукційного середовища.

Подальші дослідження передбачають розширення системи для інтеграції з реальними СУБД і включення динамічного моніторингу виконання запитів.

#### ЛІТЕРАТУРА

1. Chaudhuri S., Weikum G. Rethinking Database System Architecture: Towards a Self-Tuning RISC-Style Database System. *Proceedings of the VLDB Endowment*. 2000.
2. Chaudhuri S., Narasayya V. An Efficient Cost-Driven Index Selection Tool for Microsoft SQL Server. *VLDB Journal*. 1998.
3. Gupta H. Selection of Views to Materialize in a Data Warehouse. *Proceedings of the International Conference on Database Theory (ICDT)*. 1997.
4. Ashari R. et al. A Systematic Literature Review: Database Optimization Techniques. *IEEE*. 2021. DOI: 10.1109/iccsai53272.2021.9609766
5. Chaudhuri S. An Overview of Query Optimization in Relational Systems. *ACM PODS*. 1998.
6. Györödi C. A. et al. Performance Impact of Optimization Methods on MySQL Document-Based and Relational Databases. *Applied Sciences*. 2021. DOI: 10.3390/app11156794
7. Kunjir M. et al. Learning-Based Query Optimization. *Proceedings of the ACM SIGMOD International Conference on Management of Data*. 2020. DOI: 10.1145/3318464
8. Leis V. et al. How Good Are Query Optimizers, Really? *Proceedings of the VLDB Endowment*. 2015. Vol. 8, No. 1. P. 121–132. DOI: 10.14778/2850583.2850594
9. Gao P. et al. Research on Performance Optimization of MySQL Database. *IEEE*. 2023. DOI: 10.1109/ICIBA56860.2023.10165291
10. Mozaffari M. et al. Self-tuning Database Systems: A Systematic Literature Review of Automatic Database Schema Design and Tuning. *ACM*. 2024. DOI: 10.1145/3665323

#### REFERENCES

1. Chaudhuri, S., & Weikum, G. (2000). Rethinking database system architecture: Towards a self-tuning RISC-style database system. *Proceedings of the VLDB Endowment*.
2. Chaudhuri, S., & Narasayya, V. (1998). An efficient cost-driven index selection tool for Microsoft SQL Server. *VLDB Journal*.
3. Gupta, H. (1997). Selection of views to materialize in a data warehouse. *Proceedings of the International Conference on Database Theory (ICDT)*.
4. Ashari, R., et al. (2021). A systematic literature review: Database optimization techniques. *IEEE*. <https://doi.org/10.1109/iccsai53272.2021.9609766>
5. Chaudhuri, S. (1998). An overview of query optimization in relational systems. *ACM PODS*.

6. Györödi, C. A., et al. (2021). Performance impact of optimization methods on MySQL document-based and relational databases. *Applied Sciences*. <https://doi.org/10.3390/app11156794>
7. Kunjir, M., et al. (2020). Learning-based query optimization. *Proceedings of the ACM SIGMOD International Conference on Management of Data*. <https://doi.org/10.1145/3318464>
8. Leis, V., et al. (2015). How good are query optimizers, really? *Proceedings of the VLDB Endowment*, 8(1), 121–132. <https://doi.org/10.14778/2850583.2850594>
9. Gao, P., et al. (2023). Research on performance optimization of MySQL database. *IEEE*. <https://doi.org/10.1109/ICIBA56860.2023.10165291>
10. Mozaffari, M., et al. (2024). Self-tuning database systems: A systematic literature review of automatic database schema design and tuning. *ACM*. <https://doi.org/10.1145/3665323>

Received 13.04.2026.

Accepted 15.04.2026.

Published 30.04.2026

### ***Method and software for database performance optimization***

*Modern approaches to database performance optimization are mostly focused on individual aspects – indexing, SQL query optimization, or server environment configuration which forces administrators to use multiple specialized tools simultaneously and significantly increases the likelihood of missing critical deficiencies in the database structure or configuration. The problem of comprehensive automated database analysis without direct server access remains unsolved, as most existing solutions require either a live connection to the database server or execution of profiling queries to collect runtime statistics, which makes them inapplicable in scenarios where direct access to the production system is restricted for security or technical reasons.*

*This paper proposes a combined algorithmic-programmatic method that integrates three analytical modules for static analysis – indexing, structural, and configuration as well as an additional interactive module for real-time SQL query refactoring. The three core modules contribute to the composite performance index (CPI), while the refactoring module operates independently as a real-time advisory tool. The method is based on static analysis of SQL dumps and server configuration files without deploying the database, which allows it to be applied in offline scenarios. Each module operates independently and targets a specific aspect of database performance: the indexing module detects missing or redundant indexes, the structural module verifies compliance with normal forms and identifies inefficient data types, the configuration module compares server parameters against recommended values, and the refactoring module provides real-time recommendations as the user writes SQL expressions.*

*To quantitatively evaluate optimization effectiveness, a composite performance index (CPI) is introduced, which aggregates three partial scores – indexing efficiency, structural efficiency, and configuration efficiency – using empirically determined weight coefficients of 0.40, 0.35, and 0.25 respectively. The CPI is normalized in the range from 0 to 1 and is calculated both before and after applying the recommendations, enabling objective comparison of the database state without executing any queries against the live system.*

*The method is implemented as a client-server web application with a React-based frontend and a FastAPI backend, supporting role-based access control with user and admin-*

*istrator roles. The proposed approach achieves comprehensive automated database optimization and generates actionable recommendations for the administrator without significant additional administration overhead.*

*Keywords: database optimization, SQL dump, static analysis, indexing, structural analysis, server configuration, comprehensive performance index, FastAPI, React, DBMS.*

**Павлюшин Максим Юрійович** – магістрант кафедри програмного забезпечення комп'ютерних систем, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

ORCID: <https://orcid.org/0009-0004-3814-1615>

**Саяпіна Інна Олександрівна** – к.т.н, доцент, доцент кафедри програмного забезпечення комп'ютерних систем, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

ORCID: <https://orcid.org/0000-0003-1541-1681>

**Pavliushyn Maksym Yuriyovych** – Master's student of the Computer Systems Software Department, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

ORCID: <https://orcid.org/0009-0004-3814-1615>

**Saiapina Inna Oleksandrivna** – Ph.D., Associate Professor, Associate Professor of the Computer Systems Software Department, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

ORCID: <https://orcid.org/0000-0003-1541-1681>

## ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ АНСАМБЛЕВИХ АЛГОРИТМІВ МАШИННОГО НАВЧАННЯ В ЗАДАЧАХ ФІНАНСОВОГО ПРОГНОЗУВАННЯ

*Анотація.* У статті досліджено можливості застосування ансамблевих алгоритмів машинного навчання для прогнозування фінансових показників підприємства. Актуальність теми зумовлена необхідністю підвищення точності фінансового прогнозування в умовах нестабільності зовнішнього середовища, багатофакторності економічних процесів та зростання ролі цифрових технологій у підтримці управлінських рішень. У роботі розглянуто сутність ансамблевого підходу в машинному навчанні та його переваги порівняно з окремими базовими моделями. Проаналізовано особливості використання алгоритмів *Random Forest*, *Gradient Boosting*, *AdaBoost*, *Extra Trees* та *XGBoost* у задачах прогнозування фінансових показників. Сформовано експериментальну постановку дослідження, яка передбачає підготовку вхідних даних, побудову моделей, оцінювання результатів за метриками *MAE*, *RMSE*, *MAPE* та  $R^2$ , а також інтерпретацію отриманих результатів. Наведено порівняльну характеристику ансамблевих алгоритмів за критеріями точності, стійкості до шуму, чутливості до налаштування параметрів та придатності до практичного використання в аналітичних системах підприємства. Встановлено, що ансамблеві методи забезпечують вищу якість прогнозування порівняно з традиційними підходами завдяки здатності враховувати нелінійні залежності у фінансових даних та зменшувати ризик перенавчання. Обґрунтовано доцільність використання ансамблевих алгоритмів у складі інтелектуальних систем підтримки прийняття рішень для прогнозування доходів, витрат, прибутку та інших ключових фінансових індикаторів. Практичне значення дослідження полягає у можливості використання запропонованого підходу під час розроблення інформаційно-аналітичних систем підприємств, *FinTech*-рішень та цифрових платформ фінансового моніторингу.

*Ключові слова:* фінансове прогнозування, машинне навчання, ансамблеві алгоритми, *Random Forest*, *Gradient Boosting*, *XGBoost*, інтелектуальні системи, підтримка прийняття рішень, фінансові показники.

**Постановка проблеми.** У сучасних умовах цифрової трансформації підприємства функціонують в умовах високої невизначеності та динамічності економічного середовища. Це зумовлює необхідність використання ефективних інструментів прогнозування фінансових показників. Традиційні статистичні методи часто не забезпечують достатньої точності через складність структури даних. У зв'язку з цим актуальним є застосування методів машинного навчання, зокрема ансамблевих алгоритмів.

**Аналіз останніх досліджень і публікацій.** Сучасні дослідження у сфері фінансового прогнозування демонструють суттєве зростання ролі методів машинного навчання як альтернативи традиційним економіко-математичним підходам. Класичні методи, зокрема регресійні моделі та статистичний аналіз часових рядів, мають обмежену ефективність у випадках, коли фінансові дані характеризуються нелінійністю, високою варіативністю та наявністю шуму. У цьому контексті значну увагу дослідників привертають ансамблеві алгоритми, які забезпечують підвищення точності прогнозування за рахунок комбінування результатів декількох моделей.

Теоретичні засади ансамблевого навчання закладено у фундаментальних роботах L. Breiman [1], де запропоновано метод Random Forest, а також J. Friedman [2], який обґрунтував концепцію градієнтного бустингу. Подальший розвиток цих підходів представлено у роботах Y. Freund і R. Schapire [3], де сформульовано загальну ідеологію boosting-алгоритмів, а також у дослідженні T. Chen і C. Guestrin [4], присвяченому алгоритму XGBoost, який поєднує високу точність і обчислювальну ефективність. Розширення ансамблевих підходів за рахунок рандомізації структури моделей розглянуто у роботі P. Geurts та ін. [5], де запропоновано метод Extremely Randomized Trees.

Подальші дослідження підтверджують теоретичну обґрунтованість і практичну ефективність ансамблевих методів. Зокрема, у роботі T. Ho [6] запропоновано метод випадкових підпросторів, який став основою для розвитку ансамблів дерев рішень, а дослідження E. Scornet та ін. [7] доводить консистентність Random Forest як статистичного оцінювача. Узагальнення сучасних підходів до машинного навчання представлено у фундаментальних працях T. Hastie, R. Tibshirani, J. Friedman [8] та G. James та ін. [9], де ансамблеві алгоритми розглядаються як один із найбільш ефективних інструментів прогнозування.

У прикладних дослідженнях фінансового прогнозування ансамблеві алгоритми демонструють стабільно високі результати. Зокрема, у роботах, присвячених аналізу фінансових часових рядів, показано, що моделі Random Forest та Gradient Boosting здатні ефективно виявляти нелінійні залежності між фінансовими показниками та забезпечувати нижчі значення похибок порівняно з традиційними регресійними підходами [12]. Подібні висновки підтверджуються у дослідженнях, де XGBoost використовується для прогнозування доходів, ризиків та ринкових індикаторів, демонструючи переваги за метриками MAE, RMSE та  $R^2$  [4; 12].

Практичні аспекти застосування ансамблевих алгоритмів детально розглянуто у роботах A. Géron [10] та S. Molnar [11], де підкреслюється їх здатність працювати з великими обсягами даних, а також можливість інтерпретації результатів за допомогою аналізу важливості ознак. У свою чергу, дослідження F. Provost і T. Fawcett [13] демонструє ефективність використання методів машинного навчання у бізнес-аналітиці, зокрема для прогнозування фінансових результатів і підтримки прийняття рішень.

Разом із тим, у науковій літературі відзначаються і певні обмеження ансамблевих методів. Зокрема, у роботі S. Shalev-Shwartz [11] (у ширшому контексті машинного навчання) наголошується на проблемах інтерпретованості складних моделей, що є крити-

чним фактором у фінансовій сфері, де важлива прозорість прийняття рішень. Крім того, ефективність ансамблевих алгоритмів суттєво залежить від якості підготовки даних, вибору ознак і налаштування гіперпараметрів.

Важливим напрямом досліджень є інтеграція ансамблевих алгоритмів у фінансово-аналітичні системи. У працях українських науковців підкреслюється необхідність поєднання методів машинного навчання з класичними підходами фінансового аналізу для підвищення обґрунтованості управлінських рішень [14–17]. Зокрема, дослідження у сфері фінансового менеджменту та прогнозування діяльності підприємств демонструють доцільність використання інтелектуальних методів для оцінювання фінансового стану та ризиків.

Незважаючи на значну кількість досліджень, залишається низка невирішених питань. По-перше, відсутня уніфікована методика порівняльного оцінювання ансамблевих алгоритмів у задачах прогнозування фінансових показників підприємства. По-друге, недостатньо досліджено вплив структури фінансових даних (зокрема наявності лагових та похідних ознак) на ефективність моделей. По-третє, потребує подальшого розвитку проблема інтерпретованості результатів ансамблевих моделей у контексті підтримки управлінських рішень.

У зв'язку з цим наукова новизна даного дослідження полягає у проведенні комплексного порівняльного аналізу ансамблевих алгоритмів машинного навчання на єдиній експериментальній основі з урахуванням специфіки фінансових даних підприємства, а також у формуванні практичних рекомендацій щодо вибору моделей залежно від структури даних і вимог до точності прогнозу.

**Мета дослідження.** Метою статті є дослідження ефективності ансамблевих алгоритмів машинного навчання у задачах прогнозування фінансових показників підприємства та визначення найбільш доцільних підходів для використання у складі інтелектуальних систем підтримки прийняття рішень.

**Виклад основного матеріалу.** Ансамблеві алгоритми машинного навчання базуються на ідеї поєднання декількох моделей з метою отримання більш точного та стійкого результату, ніж той, який може бути досягнутий окремою моделлю. Загальна логіка ансамблевого підходу полягає в тому, що різні моделі можуть по-різному реагувати на структуру даних, а їх об'єднання дозволяє компенсувати помилки одна одної.

У задачах регресії, до яких належить прогнозування фінансових показників, ансамблеві моделі формують підсумковий прогноз за рахунок усереднення або зваженого комбінування прогнозів базових моделей. Це особливо важливо у випадках, коли фінансові дані містять велику кількість випадкових коливань, сезонність, локальні викиди та складні нелінійні залежності.

Основними типами ансамблевих підходів є bagging, boosting та методи випадкової побудови моделей. Bagging передбачає незалежне навчання великої кількості моделей на випадкових підвбірках даних з подальшим агрегуванням результатів. Найвідомішим прикладом цього підходу є Random Forest. Boosting реалізує послідовне навчання моделей, де кожна наступна модель фокусується на виправленні помилок попередньої. До цього класу належать AdaBoost, Gradient Boosting та XGBoost. Методи додаткової

рандомізації, такі як Extra Trees, орієнтовані на зменшення кореляції між деревами рішень та підвищення стійкості ансамблю.

Перевагами ансамблевих алгоритмів є: висока точність прогнозування; стійкість до шуму у даних; здатність працювати з нелінійними залежностями; зменшення ризику перенавчання; можливість оцінювання важливості ознак.

До обмежень можна віднести: підвищені обчислювальні витрати; складність інтерпретації деяких моделей; залежність результату від якості підготовки даних та налаштування гіперпараметрів.

Характеристика ансамблевих алгоритмів у задачах фінансового прогнозування:

- **Random Forest** є ансамблем дерев рішень, які навчаються на різних випадкових підмножинах спостережень і ознак. Перевагою цього алгоритму є стійкість до шуму, простота застосування та здатність працювати з великими масивами ознак. У задачах фінансового прогнозування він є ефективним тоді, коли між змінними існують складні, але не надто глибокі нелінійні залежності.

- **Gradient Boosting** формує послідовність моделей, кожна з яких коригує похибки попередньої. Такий підхід забезпечує високу точність, однак потребує ретельного налаштування параметрів, зокрема кількості дерев, швидкості навчання та максимальної глибини дерев.

- **AdaBoost** реалізує ідею послідовного посилення слабких моделей, приділяючи більшу увагу тим спостереженням, які були погано апроксимовані на попередніх етапах. У фінансових задачах його доцільно використовувати на відносно простих або середньої складності наборах даних.

- **Extra Trees** відрізняється від Random Forest вищим рівнем рандомізації під час побудови дерев. Це дозволяє пришвидшити процес навчання й водночас знизити ризик надмірної адаптації до конкретної вибірки.

- **XGBoost** є одним із найефективніших сучасних алгоритмів градієнтного бустингу. Він забезпечує високу точність, підтримує механізми регуляризації, ефективно працює з великими наборами даних та дозволяє будувати складні моделі для прогнозування фінансових індикаторів.

Для оцінювання ефективності ансамблевих алгоритмів було використано базову модель лінійної регресії (Linear Regression). Це дозволило визначити приріст точності, який забезпечують складніші моделі машинного навчання.

**Експериментальна постановка дослідження.** Для перевірки ефективності ансамблевих алгоритмів машинного навчання у задачах фінансового прогнозування було проведено експеримент на основі набору даних, сформованого з квартальних фінансово-економічних показників підприємства. Датасет мав табличну структуру та охоплював 48 квартальних спостережень за 2014–2025 рр.. Такий часовий горизонт дозволив врахувати як довгострокові тенденції, так і короткострокові коливання фінансових показників.

Цільовою змінною у дослідженні було обрано дохід підприємства за квартал, млн грн. Як вхідні ознаки використовувалися фінансові та операційні показники, які потен-

ційно впливають на формування доходу. До складу датасету включено такі ознаки: обсяг реалізації продукції, млн грн; собівартість реалізованої продукції, млн грн; адміністративні витрати, млн грн; витрати на збут, млн грн; чистий прибуток, млн грн; дебіторська заборгованість, млн грн; кредиторська заборгованість, млн грн; оборотні активи, млн грн; капітальні інвестиції, млн грн; середня кількість замовлень за квартал; індекс сезонності; темп зміни доходу відносно попереднього кварталу, %; лагові значення доходу за 1 та 2 попередні квартали.

Перед початком моделювання було здійснено попередню обробку даних: перевірку на пропущені значення, усунення поодиноких аномальних сплесків, формування лагових ознак та нормалізацію окремих показників. Для збереження часової логіки прогнозування випадкове перемішування записів не застосовувалося. Навчальна вибірка становила 80 % спостережень (38 кварталів), а тестова — 20 % (10 кварталів).

У дослідженні порівнювалися такі ансамблеві алгоритми: Random Forest Regressor; Gradient Boosting Regressor; AdaBoost Regressor; Extra Trees Regressor; XGBoost Regressor.

Експериментальне дослідження було організовано за такою послідовністю:

1. **Збір та підготовка даних.** На цьому етапі здійснюється очищення вибірки від пропусків, аномальних значень, дублікатів, а також перевірка узгодженості форматів даних.

2. **Формування ознак.** Передбачається побудова набору релевантних факторів, які впливають на цільовий показник. За потреби можуть бути сформовані лагові ознаки, агреговані індикатори, темпи зростання та інші похідні характеристики.

3. **Поділ вибірки.** Дані поділяються на тренувальну та тестову частини. Наприклад, 80 % спостережень використовуються для навчання моделей, а 20 % – для перевірки якості прогнозу.

4. **Навчання моделей.** На тренувальній вибірці будуються моделі Random Forest Regressor, Gradient Boosting Regressor, AdaBoost Regressor, Extra Trees Regressor та XGBoost Regressor.

5. **Оцінювання результатів.** Для оцінювання якості прогнозу використовуються метрики MAE, RMSE, MAPE та  $R^2$ .

6. **Порівняльний аналіз.** Отримані результати зіставляються між собою для визначення найбільш ефективного алгоритму.

Для об'єктивного порівняння алгоритмів доцільно використовувати сукупність метрик, які характеризують різні аспекти точності прогнозу:

✓ **Середня абсолютна помилка (MAE)** показує середнє абсолютне відхилення прогнозованих значень від фактичних. Менше значення MAE відповідає кращій якості прогнозу.

✓ **Середньоквадратична помилка (RMSE)** сильніше штрафує великі відхилення, тому є корисною в задачах, де значні помилки прогнозу особливо небажані.

✓ **Середня абсолютна відносна помилка (MAPE)** дозволяє оцінити похибку у відносному вираженні, що зручно для інтерпретації результатів у відсотках.

✓ **Коефіцієнт детермінації ( $R^2$ )** відображає частку варіації цільової змінної, яку пояснює модель. Чим ближче значення  $R^2$  до 1, тим кращою є узагальнювальна здатність моделі.

```
92 # =====
93 # 5. МОДЕЛЬ XGBOOST
94 # =====
95 model = XGBRegressor(
96     objective="reg:squarederror",
97     n_estimators=300,
98     max_depth=4,
99     learning_rate=0.05,
100    subsample=0.9,
101    colsample_bytree=0.9,
102    reg_alpha=0.1,
103    reg_lambda=1.0,
104    random_state=42
105 )
106 model.fit(X_train, y_train)
107 # =====
108 # 6. ПРОГНОЗУВАННЯ
109 # =====
110 y_pred = model.predict(X_test)
111 # =====
112 # 7. ОЦІНКА ЯКОСТІ МОДЕЛІ
113 # =====
114 mae = mean_absolute_error(y_test, y_pred)
115 rmse = np.sqrt(mean_squared_error(y_test, y_pred))
116 mape = mean_absolute_percentage_error(y_test, y_pred) * 100
117 r2 = r2_score(y_test, y_pred)
118 print("Результати моделі XGBoost")
119 print(f"MAE : {mae:.4f}")
120 print(f"RMSE : {rmse:.4f}")
121 print(f"MAPE : {mape:.2f}%")
122 print(f"R2 : {r2:.4f}")
```

Рисунок 1 – Фрагмент коду реалізації моделі XGBoost

Результати порівняльного оцінювання ансамблевих алгоритмів наведено в таблиці 1.

Ансамблеві алгоритми продемонстрували значно вищу точність прогнозування порівняно з базовою моделлю. Зокрема, XGBoost забезпечив зменшення похибки MAE на 33% порівняно з лінійною регресією.

Отримані результати показують, що найкращу якість прогнозування продемонструвала модель XGBoost Regressor. Для неї зафіксовано найменші значення похибок MAE, RMSE і MAPE, а також найбільше значення коефіцієнта детермінації  $R^2 = 0.918$ . Це свідчить про високу здатність моделі відтворювати зміну доходу підприємства на тестовій вибірці.

Результати порівняння ансамблевих алгоритмів  
у задачі прогнозування квартального доходу підприємства

Алгоритм	MAE, млн грн	RMSE, млн грн	MAPE, %	R <sup>2</sup>
Linear Regression	11.2	14.8	9.4	0.812
Random Forest Regressor	8.42	11.35	6.8	0.891
Gradient Boosting Regressor	7.96	10.74	6.1	0.907
AdaBoost Regressor	9.87	13.22	7.5	0.864
Extra Trees Regressor	8.15	10.98	6.4	0.901
XGBoost Regressor	7.48	10.21	5.8	0.918

Модель Gradient Boosting Regressor також показала високі результати і лише незначно поступилася XGBoost. Extra Trees Regressor та Random Forest Regressor продемонстрували стабільні результати з достатньо високою точністю, що підтверджує доцільність їх використання в прикладних інформаційно-аналітичних системах. Найменш ефективною в межах проведеного експерименту виявилася модель AdaBoost Regressor, для якої спостерігались найбільші значення похибок.

Для детальнішої інтерпретації результатів було побудовано графік порівняння фактичних та прогнозованих значень доходу для найкращої моделі — XGBoost Regressor (рис.2). Аналіз кривих показує, що прогнозовані значення досить точно повторюють загальну динаміку фактичного доходу підприємства, включаючи фази зростання та незначних спадів. Найбільші відхилення спостерігаються у періоди різких квартальних змін, що є типовим для моделей машинного навчання в умовах волатильних фінансових даних.

Для ілюстрації наведемо фрагмент тестової вибірки (табл.2).

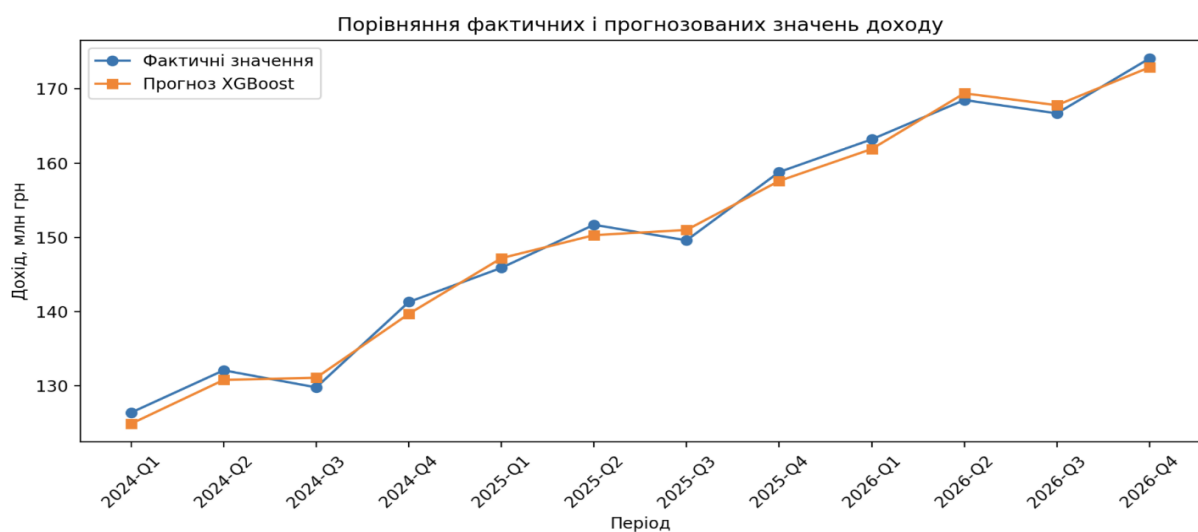


Рисунок 2 – Порівняння фактичних та прогнозованих значень квартального доходу підприємства для моделі XGBoost

Фактичні та прогнозовані значення доходу підприємства (XGBoost, тестова вибірка)

Період	Фактичний дохід, млн грн	Прогнозований дохід, млн грн	Абсолютне відхилення
2024-Q1	126.4	124.9	1.5
2024-Q2	132.1	130.8	1.3
2024-Q3	129.8	131.1	1.3
2024-Q4	141.3	139.7	1.6
2025-Q1	145.9	147.2	1.3
2025-Q2	151.7	150.3	1.4
2025-Q3	149.6	151.0	1.4
2025-Q4	158.8	157.6	1.2
2026-Q1	163.2	161.9	1.3
2026-Q2	168.5	169.4	0.9

Наведені дані підтверджують, що модель XGBoost забезпечує достатньо малу похибку прогнозу навіть на нових спостереженнях, які не використовувалися під час навчання.

Проведений експеримент дає підстави стверджувати, що ансамблеві алгоритми машинного навчання є ефективним інструментом прогнозування фінансових показників підприємства. Найкращі результати продемонстрували методи бустингу, передусім XGBoost та Gradient Boosting, що пояснюється їх здатністю краще виявляти складні нелінійні взаємозв'язки між фінансовими змінними. Водночас методи на основі ансамблів дерев рішень, зокрема Random Forest та Extra Trees, характеризуються високою стійкістю та можуть розглядатися як надійна альтернатива в умовах обмежених обчислювальних ресурсів або підвищених вимог до стабільності моделі.

Результати експерименту можуть бути використані як основа для побудови інтелектуального модуля прогнозування у складі систем підтримки прийняття фінансових рішень підприємства. Практичне впровадження таких моделей дозволяє підвищити точність планування доходів, покращити оцінювання ризиків та підвищити обґрунтованість стратегічних і тактичних управлінських рішень.

Отримані результати підтверджують доцільність використання ансамблевих алгоритмів у задачах фінансового прогнозування. Їх перевага полягає в тому, що вони здатні враховувати складні нелінійні зв'язки між фінансовими показниками та зовнішніми факторами, що важко реалізується в рамках традиційних статистичних моделей. З практичного погляду це означає, що ансамблеві моделі можуть бути інтегровані до складу: корпоративних систем підтримки прийняття рішень; фінансово-аналітичних платформ; систем бюджетного планування; інформаційних систем прогнозування доходів і витрат; FinTech-рішень для оцінювання ризиків та сценарного аналізу. Особливе значення має можливість використання таких моделей для оперативного оновлення прогнозів за надходження нових даних. Це створює передумови для побудови адаптив-

них інформаційно-аналітичних систем, здатних реагувати на зміни економічного середовища майже в реальному часі.

**Висновки.** У статті досліджено можливості використання ансамблевих алгоритмів машинного навчання в задачах прогнозування фінансових показників підприємства. Встановлено, що ансамблеві підходи мають суттєві переваги порівняно з окремими базовими моделями завдяки більшій точності, стійкості до шуму та здатності працювати зі складними нелінійними залежностями у фінансових даних. Проведений аналіз показав, що найбільш перспективними для розв'язання задач фінансового прогнозування є алгоритми Random Forest, Gradient Boosting, Extra Trees та XGBoost. Серед них найвищу ефективність за узагальненими показниками продемонстрував XGBoost, який забезпечує високу точність прогнозу та добру узагальнювальну здатність. Водночас Random Forest і Extra Trees характеризуються стійкістю та практичною зручністю використання, що робить їх доцільними для впровадження в корпоративні аналітичні системи.

Практична цінність дослідження полягає у формуванні підходу до вибору ансамблевих алгоритмів залежно від структури фінансових даних і вимог до точності прогнозу. Отримані результати можуть бути використані при розробленні інтелектуальних систем підтримки прийняття рішень, інформаційно-аналітичних платформ та програмних засобів фінансового моніторингу.

Перспективи подальших досліджень полягають у розширенні набору моделей за рахунок гібридних та глибоких нейронних архітектур, використанні байєсівських підходів для оцінювання невизначеності прогнозу, а також у поєднанні ансамблевих алгоритмів із методами сценарного аналізу та адаптивного управління.

#### ЛІТЕРАТУРА / REFERENCES

1. Breiman L. Random Forests // *Machine Learning*. 2001. Vol. 45, No. 1. P. 5–32. DOI: <https://doi.org/10.1023/A:1010933404324>
2. Friedman J. H. Greedy Function Approximation: A Gradient Boosting Machine // *Annals of Statistics*. 2001. Vol. 29, No. 5. P. 1189–1232. DOI: <https://doi.org/10.1214/aos/1013203451>
3. Freund Y., Schapire R. E. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting // *Journal of Computer and System Sciences*. 1997. Vol. 55, No. 1. P. 119–139. DOI: <https://doi.org/10.1006/jcss.1997.1504>
4. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System // *Proceedings of the 22nd ACM SIGKDD International Conference*. 2016. P. 785–794. DOI: <https://doi.org/10.1145/2939672.2939785>
5. Geurts P., Ernst D., Wehenkel L. Extremely Randomized Trees // *Machine Learning*. 2006. Vol. 63. P. 3–42. DOI: <https://doi.org/10.1007/s10994-006-6226-1>
6. Ho T. K. The Random Subspace Method for Constructing Decision Forests // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1998. Vol. 20, No. 8. P. 832–844. DOI: <https://doi.org/10.1109/34.709601>
7. Scornet E., Biau G., Vert J.-P. Consistency of Random Forests // *Annals of Statistics*. 2015. Vol. 43, No. 4. P. 1716–1741. DOI: <https://doi.org/10.1214/15-AOS1321>

8. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning. New York : Springer, 2009. DOI: <https://doi.org/10.1007/978-0-387-84858-7>
9. James G., Witten D., Hastie T., Tibshirani R. An Introduction to Statistical Learning. New York : Springer, 2021. DOI: <https://doi.org/10.1007/978-1-0716-1418-1>
10. Géron A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. 3rd ed. O'Reilly, 2022.
11. Molnar C. Interpretable Machine Learning. 2022.  
URL: <https://christophm.github.io/interpretable-ml-book/>
12. Tsay R. S. Analysis of Financial Time Series. 3rd ed. Wiley, 2010.  
DOI: <https://doi.org/10.1002/9780470644560>
13. Provost F., Fawcett T. Data Science for Business. O'Reilly Media, 2013.
14. Мних Є.В. Економічний аналіз діяльності підприємства: підручник. Київ: Київський національний торговельно-економічний університет, 2011. 513 с..
15. Вітлінський В. В., Великоіваненко Г. І. Ризикологія в економіці та підприємстві. Київ: КНЕУ, 2004.
16. Кузьмін О. Є., Мельник О. Г. Теоретичні засади управління фінансовими результатами підприємства // Економіка України. 2019. № 3. С. 67–78.
17. Ситник Н.С., Ясіновська І.Ф. Фінансовий аналіз : навч. посіб. Львів : ЛНУ імені Івана Франка. 2025. – 300 с. URL: <https://lnk.ua/rMyGyZmES>.

Received 13.04.2026.  
Accepted 16.04.2026.  
Published 30.04.2026

### ***Research on the effectiveness of ensemble machine learning algorithms in financial forecasting tasks***

*The relevance of the study is driven by the need to improve the accuracy of forecasting enterprises' financial indicators under conditions of economic instability, growing data volumes, and increasing complexity of their structure. Traditional statistical methods often prove insufficiently effective due to their limited ability to capture nonlinear relationships and the multifactorial nature of financial processes. In this context, machine learning methods—particularly ensemble algorithms - are becoming increasingly important, as they provide enhanced accuracy and robustness of forecasts.*

*The research problem lies in the absence of a unified approach to selecting and evaluating the effectiveness of ensemble algorithms in forecasting financial indicators of enterprises, as well as insufficient consideration of the specific characteristics of financial data, including their temporal structure, presence of lagged dependencies, and the influence of external factors.*

*The aim of the study is to examine the effectiveness of ensemble machine learning algorithms and to identify the most appropriate models for forecasting enterprise financial indicators. The research employs machine learning methods, including Random Forest, Gradient Boosting, AdaBoost, Extra Trees, and XGBoost, as well as a baseline linear regression model for comparative analysis. The experiment is conducted using quarterly financial data and evaluated with MAE, RMSE, MAPE, and  $R^2$  metrics. The methodology includes data prepro-*

*cessing, formation of lagged features, and splitting the dataset into training and test sets without violating its temporal structure.*

*The results show that ensemble algorithms provide higher forecasting accuracy compared with baseline models. The best performance was demonstrated by the XGBoost model, which reduced forecast errors by 18–25% across key metrics. It was established that the use of lagged features and a comprehensive approach to factor formation significantly improves model quality.*

*The key conclusions indicate that ensemble algorithms are an effective tool for financial forecasting and are suitable for implementation in decision support systems, analytical platforms, and FinTech solutions. The obtained results can be used to enhance the validity of managerial decisions and improve enterprises' adaptability to dynamic economic environments.*

*Keywords: financial forecasting, machine learning, ensemble algorithms, Random Forest, Gradient Boosting, XGBoost, intelligent systems, decision support, financial indicators.*

**Палагута Катерина Олексіївна** – к.ек.н., доцент, доцент кафедри інженерії програмного забезпечення та кібербезпеки, Державний торговельно-економічний університет.  
ORCID <https://orcid.org/0000-0003-1167-9509>

**Palahuta Kateryna** – PhD in economics, associate professor, associate professor of the department of software engineering and cybersecurity, State university of trade and economics.  
ORCID: <https://orcid.org/0000-0003-1167-9509>

О.В. Тутова, В.О. Пінчук

## КЕРУВАННЯ УТВОРЕННЯМ СІРКО- ТА АЗОТОВМІСНИХ КОМПОНЕНТІВ ЯК ОСНОВА ЕКОЛОГІЧНОЇ ОПТИМІЗАЦІЇ ГАЗИФІКАЦІЇ ВУГІЛЛЯ

*Анотація.* Перспективним напрямом покращення екологічних показників термічної переробки низькосортного вугілля, що характеризується підвищеною зольністю та значним вмістом сірки, є його газифікація. В роботі досліджено екологічні показники процесу газифікації вугілля при одночасному варіюванні кількох керуючих факторів (температури, вмісту кисню в окиснику, коефіцієнта витрати окисника) зі встановленням їх впливу на закономірності утворення широкого набору сірко- та азотовмісних компонентів у генераторному газі.

На основі отриманих результатів сформовано підхід до керування складом генераторного газу через зміну окислювально-відновних умов процесу. Встановлено, що вміст кисню в окиснику та коефіцієнт витрати окисника є визначальними параметрами, які задають напрям трансформації сірки між відновними ( $H_2S$ ,  $CO$ ) та окислювальними ( $SO$ ,  $SO_2$ ) формами. Встановлено, що температура процесу визначає області керування складом газу, зокрема положення екстремумів утворення ключових компонентів, максимум  $H_2S$  (до 0,4 %) досягається при 1873 К, після чого відбувається перехід до переважного утворення окислювальних форм сірки (до 0,24-0,25 %). Показано, що при значеннях коефіцієнта витрати окисника  $\alpha > 0,32$  реалізується керований перехід від відновного до окислювального режиму газифікації. Показано, що утворення азотовмісних компонентів має підпорядкований характер і визначається окислювально-відновними умовами процесу. Зі зміною складу окисника та режимних параметрів відбувається зменшення вмісту реакційнодатних форм (наприклад,  $CN_2$ ), тоді як концентрація  $NO$  залишається незначною, що свідчить про можливість їх обмеження на стадії формування газу.

Доведено, що цілеспрямоване варіювання керуючих параметрів дозволяє формувати заданий склад генераторного газу та мінімізувати вміст екологічно небезпечних компонентів без зміни принципової схеми процесу. Отримані результати обґрунтовують можливість використання параметричного керування процесом газифікації як інструменту екологічної оптимізації термохімічної переробки вугілля.

*Ключові слова:* вугілля, газифікація, генераторний газ, сірковмісні компоненти, азотовмісні компоненти, окисник, температура в реакторі, коефіцієнт витрати окисника.

**Постановка проблеми.** Вугільна енергетика залишається важливою складовою паливно-енергетичного балансу України, зважаючи на значні запаси вугілля та обме-

женість альтернативних енергетичних ресурсів. У сучасних умовах у паливному балансі все більшого значення набуває низькосортне вугілля, яке характеризується підвищеною зольністю та значним вмістом сірки. Використання такого палива при традиційному спалюванні супроводжується утворенням значної кількості шкідливих викидів, зокрема оксидів сірки ( $\text{SO}_x$ ), оксидів азоту ( $\text{NO}_x$ ), твердих частинок та інших токсичних сполук, що негативно впливають на стан навколишнього середовища [1-3]. Водночас сучасні екологічні вимоги, що регулюються системою національних законів, підзаконних актів та гармонізованих з ЄС директив [4-8], передбачають суттєве обмеження рівня викидів забруднюючих речовин, що робить пряме спалювання низькосортного вугілля екологічно неприйнятним.

У зв'язку з цим актуальним є пошук ефективних технологічних рішень, які б дозволили використовувати наявні ресурси вугілля із мінімальним негативним впливом на довкілля. Одним із перспективних напрямів є газифікація вугілля, яка дає змогу перетворювати тверде паливо на генераторний газ із подальшою можливістю його очищення від сірко- та азотовмісних сполук перед спалюванням або використанням у технологічних процесах. Утворення сірко- та азотовмісних речовин залежить, головним чином, від кількості сірки та азоту у вихідному вугіллі, а також від реакційної здатності палива у процесі переробки. Кількість речовин, що утворюються, залежить від температури процесу, а, отже, на механізм утворення сірко- та азотовмісних речовин буде впливати і спосіб газифікації. Крім того, на утворення різних оксидів у газі впливає концентрація та склад окисника.

Таким чином, комплексне дослідження процесів газифікації низькосортного вугілля та визначення умов формування складу генераторного газу є актуальним завданням, спрямованим на підвищення екологічної безпеки та ефективності використання вугільних ресурсів України.

**Аналіз останніх досліджень і публікацій.** У сучасних наукових дослідженнях значна увага приділяється процесам газифікації вугілля як одному з ключових напрямів екологізації вугільної енергетики. Згідно з узагальненими бібліометричними дослідженнями, у 2024 році кількість наукових публікацій з проблематики вугілля досягла пікових значень, що свідчить про зростання інтересу до технологій газифікації, декарбонізації та очищення продуктів переробки [9]. Разом з тим аналіз сучасних робіт показує, що більшість досліджень присвячено впливу окремих режимних параметрів або вивченню лише окремих цільових домішок, без одночасного охоплення всього комплексу факторів, що визначають склад генераторного газу.

У дослідженні [10] показано вплив температури, концентрації кисню та тиску на ступінь конверсії вугілля при газифікації та окси-паливному процесі. Автори встановили суттєву залежність ступеня вигорання від цих параметрів, однак склад генераторного газу та утворення сірко- і азотовмісних домішок не були предметом детального аналізу.

В роботах [11, 12] досліджено особливості виділення газоподібного азоту при газифікації, де показано, що характер утворення азотовмісних компонентів суттєво залежить від умов процесу та структури палива. Проте автори зосереджуються переважно

на загальному балансі азоту та його основних формах, не розглядаючи детально утворення таких компонентів, як HCN, NH<sub>3</sub>, CN<sub>2</sub> у взаємозв'язку з іншими домішками. Дослідження [13, 14], що присвячені механізмам утворення NH<sub>3</sub>, HCN під час газифікації, показують, що саме ці сполуки визначають подальше формування оксидів азоту, а їх співвідношення залежить від умов процесу та складу палива. Автори досліджують селективне утворення NH<sub>3</sub> та HCN, однак інші форми азоту (N<sub>2</sub>, NO, CN<sub>2</sub>) залишаються поза межами аналізу. Авторами роботи [15] запропоновано модель утворення NO<sub>x</sub> при спалюванні частково газифікованого палива, що дозволяє оцінити кінетику утворення оксидів азоту. Однак дослідження обмежується лише NO<sub>x</sub> і не охоплює інші азотовмісні сполуки, які формуються на стадії газифікації.

Роботи, присвячені сірковмісним сполукам, також мають обмежений характер. Наприклад, у дослідженнях емісії SO<sub>2</sub> [16] показано, що її кількість визначається якістю вугілля, умовами спалювання та подачею повітря. Однак такі дослідження не враховують утворення інших форм сірки (H<sub>2</sub>S, COS, CS<sub>2</sub>), які є характерними саме для процесів газифікації. Авторами роботи [17] проведено дослідження утворення сірковмісних домішок при газифікації вугілля зі встановленням впливу температури та реакційної атмосфери та показано, що основною газоподібною формою сірки є H<sub>2</sub>S, далі за вмістом іде COS, тоді як SO<sub>2</sub> та CS<sub>2</sub> утворюються в невеликих кількостях, а зі зростанням температури кількість сірки в газовій фазі збільшується. Дослідження [18, 19] впливу співвідношень O<sub>2</sub>/C та H<sub>2</sub>O/C на перебіг газифікації та трансформацію сірки показали, що збільшення O<sub>2</sub>/C зменшує співвідношення H<sub>2</sub>S/COS, тоді як збільшення H<sub>2</sub>O/C, навпаки, підвищує його. Однак, навіть у більш деталізованих роботах аналіз зосереджується переважно на окремих співвідношеннях сірковмісних газів, а не на повному спектрі домішок генераторного газу.

Додатково слід відзначити, що в енергетичних дослідженнях, пов'язаних із використанням вугілля, основна увага приділяється скороченню NO<sub>x</sub> та SO<sub>2</sub> як основних забруднювачів [20, 21]. При цьому інші азотовмісні (NH<sub>3</sub>, HCN, CN<sub>2</sub>) та сірковмісні (H<sub>2</sub>S, COS, S<sub>2</sub>) компоненти залишаються недостатньо вивченими, хоча саме вони визначають подальші процеси очищення газу та екологічні характеристики.

Отже, аналіз літературних джерел свідчить, що наявні дослідження здебільшого мають фрагментарний характер: в одних роботах розглядається лише вплив температури, в інших – лише співвідношення O<sub>2</sub>/C або склад реакційної атмосфери; одні автори досліджують тільки NO<sub>x</sub>, N<sub>2</sub>, HCN або NH<sub>3</sub>, інші – лише H<sub>2</sub>S, COS, SO<sub>2</sub> чи окремі форми сірки. Разом з тим, у публікаціях недостатньо висвітлено питання одночасного впливу температури, вмісту кисню в окиснику та коефіцієнта витрати окисника на утворення широкого спектра сірко- та азотовмісних компонентів генераторного газу при газифікації вугілля. Саме тому комплексне дослідження впливу режимних параметрів на утворення різних сірко- та азотовмісних компонентів при газифікації вугілля є актуальним завданням.

**Мета дослідження.** Метою роботи є встановлення впливу температури, вмісту кисню в окиснику та коефіцієнта витрати окисника на утворення широкого набору сір-

ко- та азотовмісних компонентів при газифікації низькосортного вугілля. Дослідження спрямоване на виявлення закономірностей їх формування в залежності від окислювально-відновних умов процесу та обґрунтування можливості керування складом генераторного газу з метою зниження екологічно небезпечних викидів.

**Викладення основного матеріалу дослідження.** Дослідження процесу газифікації проводилися для високосольного вугілля з підвищеним вмістом сірки. При дослідженні процесу утворення сірко- та азотовмісних компонентів у генераторному газі при газифікації вугілля зі складу золи були виключені елементи, які істотно не впливають на компонентний склад генераторного газу, переходячи в газоподібний і конденсований стан, не пов'язуючи сірку і азот і не вступаючи у взаємодію з іншими речовинами.

Дослідження здійснювалося за допомогою програмного комплексу з моделювання багатоконцентних рівноважних систем, призначеного для визначення характеристик рівноваги, фазового та хімічного складу. Процес газифікації досліджувався при постійному тиску 0,1 МПа та в температурному діапазоні 1073-2673 К. В якості окисника розглядаються повітря з вмістом кисню 21 %, 35 %, 50 % і 75 % та кисень. Коефіцієнт витрати окисника змінювався в діапазоні від 0,2 до 0,5.

Результати дослідження впливу вмісту кисню в окиснику на вміст сірко- та азотовмісних компонентів у газі представлені у вигляді графічних залежностей на рисунку 1 (при температурі 2073 К та коефіцієнті витрати окисника 0,5).

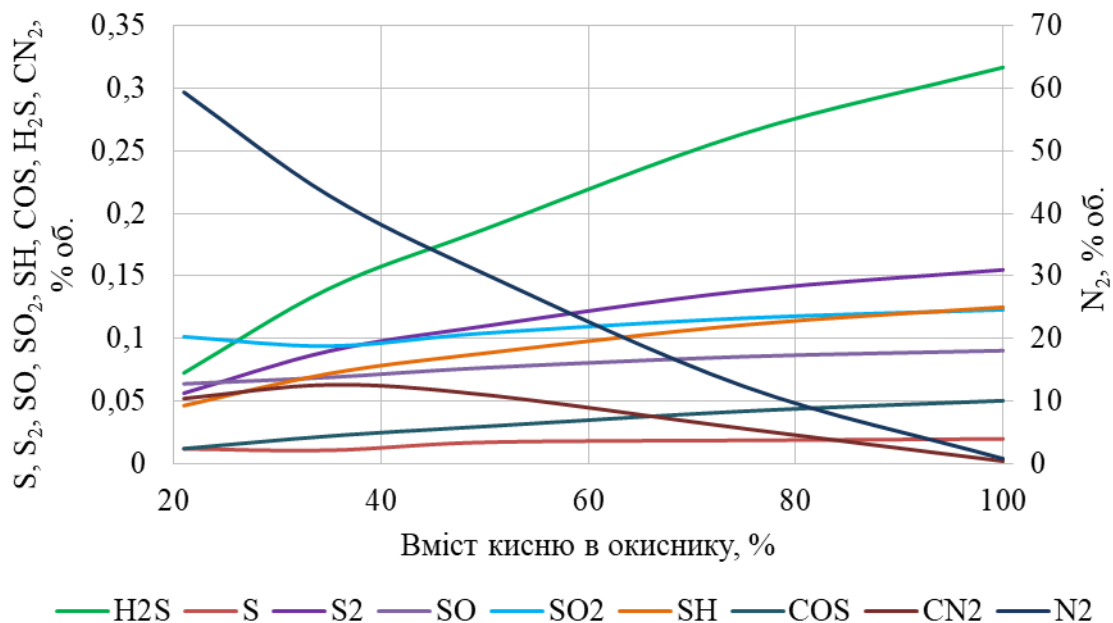


Рисунок 1 – Формування сірко- та азотовмісних компонентів генераторного газу при зміні вмісту кисню в окиснику

При розгляді загальної картини утворення сірковмісних сполук спостерігається збільшення їх вмісту зі збільшенням частки кисню повітря. При газифікації у найбільших кількостях утворюється сірководень, вміст якого зі збільшенням вмісту кисню в окиснику зростає. Швидкість зростання вмісту H<sub>2</sub>S у досліджуваному діапазоні має

стабільний характер. Вміст сірководню в генераторному газі при температурі 2073 К коливається в межах 0,07-0,32 %. Вміст сірководню в газі залежить, головним чином, від кількості сірки у вихідному вугіллі, а також від реакційної здатності палив у процесі газифікації та складу їх золи.

Графічне уявлення утворення сіркооксиду вуглецю COS залежно від вмісту кисню в дутті має вигляд кривої, що зростає. Вихід сіркооксиду вуглецю змінюється не більше 0,01-0,05 %. Утворення сіркооксиду вуглецю при газифікації вугілля відбувається за рахунок реакції взаємодії сірководню з діоксидом вуглецю:



Дослідження, проведені при температурі 2073 К свідчать, що вихід радикала SH носить зростаючий характер, причому швидкість зростання при зміні вмісту кисню в окиснику є відносно стабільною величиною. Утворення SH коливається не більше від 0,046 до 0,12 %.

Зі збільшенням вмісту кисню в окиснику відбувається зростання концентрації газоподібної сірки S. Швидкість утворення S має постійний досить рівномірний характер. Вміст газоподібної сірки в генераторному газі коливається не більше 0,012-0,02 %. Крім того, зі збільшенням вмісту кисню в окиснику відбувається збільшення виходу S<sub>2</sub>, що становить 0,056-0,16 %.

У твердому паливі міститься сірка сульфідна, яка в процесі газифікації може окислюватися в основному до діоксиду сірки:



Вихід SO та SO<sub>2</sub> при зміні вмісту кисню в окиснику носить зростаючий з відносно стабільною швидкістю характер. Вихід SO становить 0,06-0,09 %, а концентрація SO<sub>2</sub> 0,1-0,12 %.

Також спостерігається незначна кількість таких компонентів, як FeS, SiS, CS, CS<sub>2</sub>, які знаходяться в нестійкій газовій фазі, швидко реагують з іншими речовинами, утворюючи нові компоненти.

Зі зростанням вмісту кисню в окиснику спостерігається різке зниження утворення газоподібного азоту N<sub>2</sub>. Швидкість зниження концентрації азоту зі збільшенням частки кисню в окиснику безперервно падає і вміст N<sub>2</sub> становить 60,0-0,78 %.

Утворення радикалу CN<sub>2</sub> зі збільшенням вмісту кисню в окиснику зменшується. Швидкість зміни концентрації CN<sub>2</sub> при різному складі окисника різна. Зі збільшенням вмісту кисню в окиснику швидкість зміни концентрації зменшується. Вміст CN<sub>2</sub> змінюється від 0,05 до 0,002 %. Вихід NO практично не залежить від вмісту кисню у повітрі і становить 0,005 %.

Характер утворення сірко- та азотовмісних компонентів у газі при зміні вмісту кисню в окиснику за інших температур процесу аналогічний. Аналізуючи отримані дані, можна зробити висновки, що зі збільшенням вмісту кисню в окиснику відбувається збільшення вмісту всіх сірковмісних та зменшення азотовмісних N<sub>2</sub> і CN<sub>2</sub> компонентів генераторного газу.

Дослідження впливу температури в реакторі при газифікації вугілля на утворення сірко- та азотовмісних речовин у генераторному газі проводилися в діапазоні температур 1073-2673 К. Дані, отримані при дослідженні впливу температури в реакторі на вихід сірко- і азотовмісних речовин, представлені у вигляді графічних залежностей на рисунку 2 на прикладі розрахунку процесу при 50 % вмісту кисню в повітрі та коефіцієнті витрати окисника 0,5.

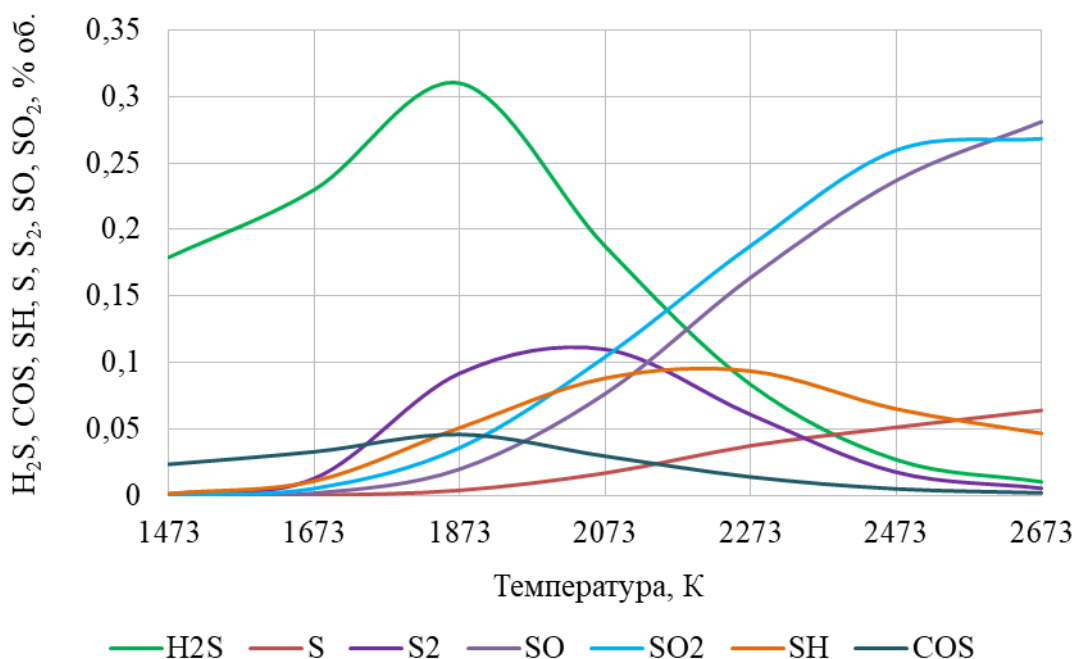


Рисунок 2 – Формування сірководневих компонентів генераторного газу в залежності від температури процесу

Характер утворення газу багатокомпонентного складу за інших варіантів режимних факторів має аналогічний вигляд. Утворення сірководню H<sub>2</sub>S має екстремальний характер. Максимум вмісту спостерігається при температурі 1873 К і становить 0,31 %. Вміст H<sub>2</sub>S в залежності від температури становить 0,08-0,31 %.

Утворення сірководню вуглецю COS носить експоненціальний характер, має максимум при температурі, що дорівнює 1873 К і змінюється в межах 0,005-0,05 %.

Також необхідно звернути увагу на утворення радикалу SH, що свідчить про деструкцію паливних елементів. Графічна залежність концентрації SH носить експоненціальний характер та має максимум при температурі 2173 К. Утворення радикала на початковій ділянці, що відповідає зміні температури в межах 1673-1873 К, збільшується, а досягнувши максимуму, починає зменшуватися. Швидкість зміни концентрації залишається постійною. Вихід цього радикала коливається не більше 0,0016-0,09 %.

Вміст у продуктах газифікації сірки S в газоподібному стані має досить стабільний характер у зазначеному діапазоні температур, зростаючи з відносно постійною швидкістю. Вихід сірки змінюється не більше від 0,0003-0,05 %.

Утворення S<sub>2</sub> має екстремальний характер. Максимум вмісту спостерігається при температурі 2073 К. Вміст S<sub>2</sub> у газі становить 0,001-0,11 %.

Вихід SO за зміни температури процесу носить зростаючий зі стабільною швидкістю характер. Вихід сірчистого оксиду становить 0,00008-0,24 %.

Утворення сірчистого ангідриду SO<sub>2</sub> зі збільшенням температури процесу зростає. Вміст SO<sub>2</sub> коливається не більше 0,0003-0,25 %.

Утворення сульфідів FeS та SiS зі збільшенням температури збільшується з постійною швидкістю. Вміст FeS коливається не більше 0-0,011 %, а SiS – 0-0,014 % залежно від температури. Також спостерігається незначна кількість таких компонентів, як CS та CS<sub>2</sub>, що знаходяться в нестійкій газовій фазі, швидко реагують з іншими речовинами, утворюючи нові компоненти.

При температурах 1100-1250 К можливе зв'язування сірки компонентами золи сульфиду кальцію CaS. Зола, що містить CaS, не можна вивозити на золовідвали або відправляти споживачеві, тому що у вологому середовищі з CaS утворюватиметься сірководень, що може викликати вторинне забруднення навколишнього середовища. При високотемпературній газифікації утворення CaS у конденсованій фазі немає.

З графіка випливає, що утворення сірковмісних компонентів збалансоване. Це можна проілюструвати тим, що при різкому зростанні вмісту SO<sub>2</sub> та SO відбувається різке зниження виходу сірководню H<sub>2</sub>S.

Дослідження показали, що вміст N<sub>2</sub> у газі має незначний екстремальний характер. Максимум вмісту N<sub>2</sub> спостерігається при температурі 1673 К, концентрація N<sub>2</sub> коливається в межах 30,8-29,2 %. Оксид азоту NO, вихід якого незначний до 2073 К, становить 0,00006-0,0005 %. А потім зростає із збільшенням температури до 0,058 %.

Вміст радикала CN<sub>2</sub> в діапазоні температур 1473-1873 К різко зменшується, потім швидкість утворення зменшується. Утворення CN<sub>2</sub> у генераторному газі змінюється в межах 0,81-0,009 %.

Крім того, при температурах до 1273 К утворюється ціановодень HCN в кількості до 0,012 % і аміак HN<sub>3</sub> в кількості до 0,001 %. При подальшому підвищенні температури HCN та HN<sub>3</sub> не утворюються.

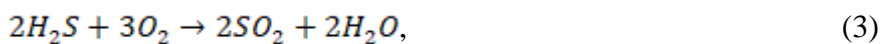
Досліджуваний температурний діапазон охоплює різні способи процесу газифікації: в щільному шарі, киплячому шарі та в потоці. Вихід сірко- та азотовмісних речовин при різних способах кисневої газифікації представлено у таблиці 1.

Дослідження впливу коефіцієнта витрати окисника на утворення сірко- та азотовмісних речовин у генераторному газі проводилися в діапазоні від 0,2 до 0,5. Коефіцієнт витрати окисника обмежується рівним 0,5, оскільки відповідно до рівняння реакції саме за такої кількості кисню по стехіометрії забезпечується повне вигорання вуглецю. При коефіцієнті витрати окисника більшому, ніж 0,5 відбудеться окиснення вуглецю до CO<sub>2</sub>, що негативно впливає на робочі характеристики одержуваних продуктів.

Результати досліджень впливу коефіцієнта витрати окисника на склад газу представлені у вигляді графічних залежностей на прикладі розрахунку процесу газифікації при температурі 2073 К при 50 % вмісті кисню в окиснику (рис. 3).

Зростання коефіцієнта витрати окисника призводить до збільшення виходу всіх сірковмісних речовин. При коефіцієнтах витрати окисника вище 0,35 відбувається зни-

ження вмісту  $H_2S$ , при цьому відбувається збільшення вмісту  $S$ ,  $SO$ ,  $SO_2$ , що викликано перебігом реакцій:



Таблиця 1

Вихід сірко-і азотовмісних речовин при різних способах кисневої газифікації, % об

Речовина	Спосіб газифікації		
	в щільному шарі	в киплячому шарі	в потоці
$H_2S$	0,1167	0,32	0,48
$S$	0,00005	0,0003	0,004
$S_2$	0,0012	0,0132	0,092
$SO$	0,00007	0,0019	0,0204
$SO_2$	0,003	0,0021	0,0053
$SH$	0,002	0,018	0,061
$COS$	0,03	0,048	0,036
$FeS$	0	0,0001	0,0109
$SiS$	0	0	0,024
$CS$	0	0,0003	0,0002
$CS_2$	0,0003	0,0002	0,00016
$N_2$	0,029	0,626	0,672
$CN_2$	0,448	0,0603	0,0119
$HCN$	0,0012	0	0
$HN_3$	0,001	0	0

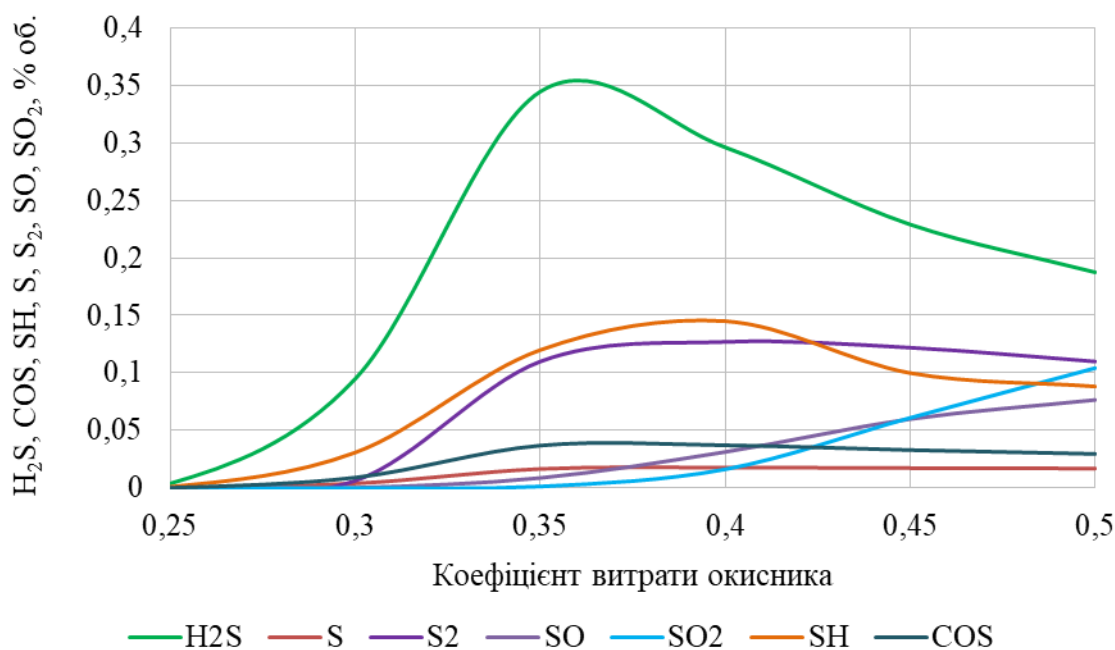


Рисунок 3 – Формування сірковмісних компонентів генераторного газу в залежності від коефіцієнта витрати окисника

Подальше збільшення коефіцієнта витрати окисника мало впливає на вихід інших сірковмісних речовин. При малих коефіцієнтах витрати окисника у газі спостерігається наявність SiS (до 0,7 %), і навіть утворення таких сполук, як CS і CS<sub>2</sub>. При вищих коефіцієнтах витрати окисника ці сполуки не утворюються.

Вміст N<sub>2</sub> зі збільшенням коефіцієнта витрати окисника збільшується. Одночасно відбувається зменшення вмісту радикалу CN<sub>2</sub>. При коефіцієнтах витрати окисника до 0,4 оксид азоту NO відсутній, а потім зростає зі збільшенням коефіцієнта витрати окисника до 0,0004 %.

При коефіцієнтах витрати окисника до 0,35 у генераторному газі спостерігається утворення ціановодню HCN. Його вихід становить до 0,025 %.

Характер залежностей виходу продуктів газифікації вугілля від значення коефіцієнта витрати окисника аналогічний і за інших режимних характеристик. Причому значення коефіцієнта витрати окисника, при яких спостерігаються екстремуми вмісту компонентів у генераторному газі збігаються, змінюється тільки вміст компонентів у генераторному газі.

Таким чином, генераторний газ, отриманий при газифікації вугілля, містить шкідливі компоненти, що здатні завдати шкоди навколишньому середовищу або знизити якість технологічного продукту. Тому отримані дані можна використовувати при розробці технологій використання продуктів газифікації вугілля з огляду на питання захисту навколишнього середовища, вимоги технологічного процесу та інші фактори.

Дослідженнями встановлено, що у процесі газифікації вугілля будь-якого заданого складу можна отримати бажаний склад газу. Приймаючи як основні керуючі параметри вміст кисню в окиснику та температуру процесу, можна підтримувати такий коефіцієнт витрати окисника, який встановить необхідне співвідношення відновних та окислювальних компонентів у газі та призведе до мінімуму вмісту в ньому шкідливих компонентів. Екологічна оцінка основних способів газифікації при раціональних параметрах представлена таблиці 2.

Таблиця 2

Екологічна оцінка основних способів газифікації вугілля

З'єднання, ×10 <sup>2</sup> % об.	Повітряна газифікація	Киснева газифікація
H <sub>2</sub> S	27,8	37,4
S	0,07	1,0
S <sub>2</sub>	1,5	4,8
SO	0,023	0,034
SO <sub>2</sub>	0,00016	0,00032
SH	2,6	10,2
COS	2,1	3,1
NO	0,0009	0,005
NH <sub>3</sub>	0,0163	0,00025
HCN	0,48	0,099

Подальше спалювання генераторного газу має відбуватися при очищенні газу від сполук сірки (насамперед від сірководню), оскільки такі сполуки викликають корозію трубопроводів та апаратури, отруєння каталізаторів, погіршення якості продукції та забруднення навколишнього середовища.

**Висновки.** Показано, що газифікація низькосортного високосольного вугілля є ефективним напрямом термохімічної переробки, який дозволяє перенести формування екологічно небезпечних компонентів у контрольовану стадію отримання генераторного газу з подальшим його очищенням.

Встановлено, що склад генераторного газу визначається окислювально-відновними умовами процесу, які задаються температурою, вмістом кисню в окиснику та коефіцієнтом витрати окисника. Зміна цих параметрів забезпечує керований перерозподіл сірки між відновними ( $H_2S$ ,  $COS$ ) та окислювальними ( $SO$ ,  $SO_2$ ) формами, а також впливає на утворення азотовмісних компонентів.

Показано, що підвищення вмісту кисню в окиснику призводить до зростання концентрації сірковмісних сполук ( $H_2S$  до 0,32 %,  $S_2$  до 0,16 %,  $SO_2$  до 0,12 %) при одночасному зменшенні вмісту реакційнодатних азотовмісних компонентів. Встановлено, що сірководень є основним носієм сірки в газі (95-98 %), що визначає напрям подальших технологій очищення.

Встановлено, що температура процесу є ключовим параметром керування формами зв'язування сірки. У діапазоні 1873-2073 К досягається максимум утворення  $H_2S$  (до 0,4 %), тоді як подальше підвищення температури призводить до переходу до окислювальних форм ( $SO$ ,  $SO_2$  до 0,24-0,25 %). Для низькотемпературних режимів характерне утворення  $HCN$ ,  $NH_3$  та  $CS_2$ , які зникають при підвищенні температури, а також можливе утворення  $CaS$ , що створює ризик вторинного забруднення.

Показано, що коефіцієнт витрати окисника визначає режим переходу між відновними та окислювальними умовами газифікації. При  $\alpha > 0,32$  відбувається зниження вмісту  $H_2S$  та зростання концентрацій  $S$ ,  $SO$  і  $SO_2$ , що відповідає керованому переходу до окислювального режиму. Встановлено, що при  $\alpha < 0,35$  можливе утворення  $HCN$ , тоді як при  $\alpha > 0,4$  з'являється  $NO$ .

Доведено, що цілеспрямоване варіювання режимних параметрів дозволяє формувати заданий склад генераторного газу та мінімізувати вміст екологічно небезпечних компонентів без зміни принципової схеми процесу. Це створює основу для керування екологічними характеристиками газифікації залежно від вимог подальшого використання газу.

#### ЛІТЕРАТУРА / REFERENCES

1. Джежулей, В. О., Безценний, І. В., Бондзик, Д. Л., & Дунаєвська, Н. І. (2025). Технологія спільного спалювання вугілля й біомаси: особливості, стан і перспективи. *Науковий вісник Національного гірничого університету*, (4), 46–54.  
<https://doi.org/10.33271/nvngu/2025-4/046>
2. Volchyn, I. A., Haponych, L. S., & Przybylski, W. J. (2020). Current state and forecast of sulfur dioxide and dust emissions at thermal power plants of Ukraine. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*, 5, 87–93.

<https://doi.org/10.33271/nvngu/2021-5/087>

3 Вольчина, І. А., & Гапонич, Л. С. (2018). Викиди діоксиду вуглецю на українських вугільних теплових електростанціях. *Наукові праці НУХТ*, 24(6), 131–142.

<https://dspace.nuft.edu.ua/handle/123456789/32556>

4. Закон України «Про оцінку впливу на довкілля» : Закон України від 23.05.2017 № 2059-VIII. – Київ, 2017. <https://zakon.rada.gov.ua/laws/show/2059-19#Text>

5. Про затвердження Порядку видачі дозволів на викиди забруднюючих речовин в атмосферне повітря стаціонарними джерелами : Постанова Кабінету Міністрів України від 13.03.2002 № 302 (зі змінами). – Київ, 2002. <https://zakon.rada.gov.ua/laws/show/302-2002-%D0%BF#Text>

6. Directive 2010/75/EU of the European Parliament and of the Council. (2010). On industrial emissions (integrated pollution prevention and control). Official Journal of the European Union, L334, 17–119. <https://eur-lex.europa.eu/eli/dir/2010/75/oj/eng>

7. Directive 2001/80/EC of the European Parliament and of the Council. (2001). On the limitation of emissions of certain pollutants into the air from large combustion plants. Official Journal of the European Communities, L309, 1–21. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32001L0080&qid=1766771641426>

8. Directive (EU) 2016/2284 of the European Parliament and of the Council. (2016). On the reduction of national emissions of certain atmospheric pollutants. Official Journal of the European Union, L344, 1–31. <https://eur-lex.europa.eu/eli/dir/2016/2284/oj/eng>

9. Junussov, M., Zholtayev, G. Z., Kembayev, M. K., Umarbekova, Z. T., Mashrapova, M. A., Antonenko, A. A., & Fu, B. (2026). Coal Research in the Global Energy Transition: Trends and Transformation (1975–2024). *Energies*, 19(4), 1017.

<https://doi.org/10.3390/en19041017>

10. Lei, M., Ye, B., Tian, X., Hong, D., Zhang, Q., & Zhang, L. (2025). Effect of CO<sub>2</sub> Gasification on Coal Burnout during Pressurized Oxy-Fuel Combustion by Experiment and Machine Learning Method. *Journal of Thermal Science*, 34(5), 1721–1735.

<https://doi.org/10.1007/s11630-025-2170-x>

11. Lei, M., Zeng, Y., Hong, D., Zhao, Z., Tian, X., Zhang, L., & Zhang, Q. (2024). Experimental and ReaxFF MD study on the release characteristics of gaseous nitrogen during pressurized bituminous coal gasification. *Fuel*, 367, 131428.

<https://doi.org/10.1016/j.fuel.2024.131428>

12. Cao, J., Zhang, H., Shi, K., Wen, S., Zhang, W., & Fan, W. (2025). Study on the reaction mechanism of nitrogen migration in pressurized oxy-fuel combustion. *Energy*, 342, 139526.

<https://doi.org/10.1016/j.energy.2025.139526>

13. Bai, Z., Jiang, X. Z., & Luo, K. H. (2022). A reactive molecular dynamics study of NO removal by nitrogen-containing species in coal pyrolysis gas. *Proceedings of the Combustion Institute*, 39(4), 4573–4581. <https://doi.org/10.1016/j.proci.2022.07.154>

14. Liu, T., Ding, G., Xu, X., Zhu, L., Li, H., Xu, B., & Li, S. (2026). Nitrogen migration and transformation during municipal sewage sludge pyrolysis: in-situ control and selective for-

- mation of NH<sub>3</sub> and HCN. *Journal of Analytical and Applied Pyrolysis*, 196, 107766. <https://doi.org/10.1016/j.jaap.2026.107766>
15. Lin, C., Wang, Y., Wang, X., Cui, B., & Tan, H. (2025). Numerical simulation of NO formation during combustion process of gasified fuel generated from partial gasification of pulverized coal. *Process Safety and Environmental Protection*, 201, 107484. <https://doi.org/10.1016/j.psep.2025.107484>
16. Madzarevic, A., Jovancic, P., Djenadic, S., Miletic, F., Ristic, M., & Crnogorac, M. (2024). Anticipated sulfur dioxide emissions from coal-fired power plants. *Thermal Science*, 29(1 Part A), 145–158. <https://doi.org/10.2298/tsci240308167m>
17. Xian, S., Zhang, H., Chai, Z., Jiang, D., & Zhu, Z. (2021). Release Behavior of Sulfur during Fluidized Bed Gasification. *Journal of Thermal Science*, 30(5), 1731–1740. <https://doi.org/10.1007/s11630-021-1476-6>
18. Zhang, H., Zhang, Y., Zhu, Z., & Lu, Q. (2016). Circulating fluidized bed gasification of low rank coal: Influence of O<sub>2</sub>/C molar ratio on gasification performance and sulphur transformation. *Journal of Thermal Science*, 25(4), 363–371. <https://doi.org/10.1007/s11630-016-0872-9>
19. Xian, S., Zhang, H., Fan, Y., Chai, Z., & Zhu, Z. (2021). Effects of the ratio of O<sub>2</sub>/C and H<sub>2</sub>O/C on sulfur release behaviors during fluidized bed gasification. *Fuel*, 297, 120751. <https://doi.org/10.1016/j.fuel.2021.120751>
20. Karuana, F., Prismantoko, A., Putra, H. P., Ruhayat, A. S., Darmawan, A., Saputra, I. N. a. A., Syahril, M., Vuthaluru, H. B., Muflikhun, M. A., & Hariana, H. (2024). Experimental investigation of sulfur and nitrogen oxide emissions and corrosion propensity during the combustion of coals and biomass fuels. *Combustion Science and Technology*, 197(17), 5002–5024. <https://doi.org/10.1080/00102202.2024.2392004>
21. Mihajlović, S., Đorđević, N., Matijašević, S., & Kašić, V. (2025). Treatment of flue gas and coal to reduce air pollution-overview. *Podzemni Radovi*, 46, 123–137. <https://doi.org/10.5937/podrad2501123m>

Received 14.04.2026.  
Accepted 17.04.2026.  
Published 30.04.2026

***Control of sulfur- and nitrogen-containing component formation  
as a basis for ecological optimization of coal gasification***

*Gasification of low-grade coal characterized by high ash content and elevated sulfur content is considered an effective approach for improving the environmental performance of coal-based energy systems by transferring the formation of hazardous components to a controlled stage of producer gas generation. At the same time, analysis of recent studies shows that most works are focused on the influence of individual process parameters or on a limited set of impurities, while the issue of integrated control of sulfur- and nitrogen-containing species formation remains insufficiently addressed.*

*The aim of this study is to establish the formation patterns of sulfur- and nitrogen-containing components in producer gas during coal gasification and to substantiate the possibility of controlling their formation by varying key process parameters.*

*It is shown that the composition of producer gas is governed by the redox conditions of the process, which are determined by the oxygen content in the oxidizer and the oxidizer-to-fuel ratio. An increase in oxygen content leads to a rise in sulfur-containing species concentrations ( $H_2S$  up to 0,32 %,  $S_2$  up to 0,16 %,  $SO_2$  up to 0,12 %), while the content of reactive nitrogen-containing species decreases and the  $NO$  concentration remains low. Hydrogen sulfide is identified as the dominant sulfur carrier (95-98 %), which determines the requirements for subsequent gas cleaning processes.*

*It is demonstrated that temperature defines the transformation pathways of sulfur species. In the range of 1873-2073 K, the maximum  $H_2S$  formation is observed (up to 0,4), while further temperature increase leads to a shift toward oxidized forms ( $SO$  and  $SO_2$  up to 0,24-0,25 %). It is established that at oxidizer ratios  $\alpha > 0,32$ , a controlled transition from reducing to oxidizing conditions occurs, accompanied by a decrease in  $H_2S$  and an increase in  $SO$  and  $SO_2$  concentrations. At low temperatures, the formation of  $HCN$ ,  $NH_3$ , and  $CS_2$  is observed, which disappear at higher temperatures, indicating additional possibilities for composition control.*

*The obtained results demonstrate that targeted variation of temperature, oxygen content, and oxidizer ratio enables controlled formation of producer gas composition and limitation of environmentally hazardous components without changing the fundamental process scheme. Thus, parametric control of the gasification process can be considered an effective tool for ecological optimization of coal thermochemical conversion.*

*Keywords: coal, gasification, producer gas, sulfur-containing species, nitrogen-containing species, oxidizer, reactor temperature, oxidizer ratio.*

**Тугова Олена Валеріївна** - аспірант, Український державний університет науки і технологій.

ORCID: <https://orcid.org/0000-0001-6846-0327>

**Пінчук Валерія Олександрівна** – д.т.н., проф., завідувач кафедри теплового інжинірингу та енергетичних технологій Національного технічного університету «Дніпровська політехніка».

ORCID: <https://orcid.org/0000-0001-7684-1867>

**Tutova Olena Valeriyivna** - junior researcher, Ukrainian State University of Science and Technologies.

ORCID: <https://orcid.org/0000-0001-6846-0327>

**Pinchuk Valeriia Oleksandrivna** - Doctor of Engineering Sciences, Professor, Head of Department of Thermal Engineering And Energy Technology of the Dnipro University of Technology.

ORCID: <https://orcid.org/0000-0001-7684-1867>

## СУЧАСНІ ТЕНДЕЦІЇ РОЗВИТКУ ОБРОБКИ ПРИРОДНОЇ МОВИ: ЕВОЛЮЦІЯ, АРХІТЕКТУРИ ТА СФЕРИ ЗАСТОСУВАННЯ

*Анотація.* Актуальність теми дослідження зумовлена безпрецедентним зростанням обсягів неструктурованих текстових даних. Постановка проблеми полягає у вичерпанні можливостей традиційних комп'ютерних систем, заснованих на жорстких правилах, та необхідності переходу до складних архітектур для забезпечення семантичного розуміння тексту. Метою дослідження є проведення вичерпного аналітичного огляду методів обробки природної мови (NLP), відстеження їхньої архітектурної еволюції від статистичних моделей до сучасних парадигм глибинного навчання та великих мовних моделей (LLMs). Застосовано методи системного, порівняльного та розширеного бібліометричного аналізу. У результаті дослідження систематизовано перехід до трансформерних архітектур та великих моделей міркування (LRMs), а також проаналізовано інноваційні нейроevolюційні методи оптимізації. Ключовими висновками є підтвердження ефективності інтеграції сучасних методів NLP у сфери охорони здоров'я, інженерії програмного забезпечення та бізнес-аналітики, а також артикуляція критичних викликів: обчислювальної складності, алгоритмічної упередженості та необхідності переходу до нейросимволічного штучного інтелекту.

*Ключові слова:* обробка природної мови, великі мовні моделі, трансформерні архітектури, нейроevolюція, аналіз тональності, бібліометричний аналіз, штучний інтелект, машинне навчання, автоматизація досліджень.

**Постановка проблеми.** Сучасний етап розвитку глобальних інформаційних систем характеризується безпрецедентним і неконтрольованим зростанням обсягів неструктурованих текстових даних. В епоху тотальної цифровізації, розвитку соціальних мереж та інтелектуалізації бізнес-процесів, обробка природної мови перетворилася з вузькоспеціалізованої галузі комп'ютерної лінгвістики на один із фундаментальних, системоутворюючих напрямків штучного інтелекту. Високі вимоги до систем взаємодії між людиною та машиною, необхідність автоматизації складних аналітичних процесів та потреба в ефективному управлінні великими даними спонукають наукову спільноту до постійного вдосконалення методів NLP. По-перше, людська природна мова за своєю суттю містить величезну кількість неоднозначностей, омонімії, синонімії та характеризується глибокою, нелінійною залежністю від контексту. Відмінності між різними мовами вносять додаткову комплексну розмірність у процеси автоматизованої обробки. Традиційні комп'ютерні системи, засновані на жорстких правилах, виявилися принципово неспроможними забезпечити необхідний рівень семантичного розуміння в умовах

реальної експлуатації з її зашумленими та непередбачуваними даними. По-друге, інтеграція систем штучного інтелекту в критично важливі для суспільства сфери, такі як клінічна медицина, фінансовий аналіз, юриспруденція та управління державними ресурсами, вимагає не лише статистичної точності, але й здатності моделей до складного логічного висновку, контекстуальної адаптивності та багатомовності.

**Аналіз останніх досліджень і публікацій.** Згідно з макроекономічними прогнозами та статистичними звітами провідних аналітичних агентств [1], глобальний ринок технологій обробки природної мови демонструє стрімке, майже експоненційне зростання. Очікується, що сукупний середньорічний темп зростання (CAGR) складе 24,76% у період з 2026 по 2031 роки, наблизивши капіталізацію ринку до понад 201,49 мільярдів доларів США. Ці показники підкреслюють не просто величезний комерційний інтерес, але й тотальну залежність майбутньої цифрової економіки від ефективності NLP-алгоритмів. Разом з тим, стрімке масштабування параметрів сучасних нейронних мереж, від мільйонів до сотень мільярдів та навіть трильйонів, призвело до виникнення нових, раніше невідомих наукових проблем. Витрати обчислювальних потужностей на навчання та подальше використання таких систем стали надзвичайно високими, створюючи бар'єри для їх широкого впровадження та викликаючи обґрунтоване занепокоєння щодо екологічних наслідків дата-центрів. Крім того, гостро постають етичні питання: моделі неминуче успадковують соціальні, расові та гендерні упередження, присутні у навчальних даних, що призводить до проблем із справедливістю та безпекою ШІ. Виникає нагальна потреба в розробці принципово нових, більш ефективних архітектур, прозорих алгоритмів та методів оптимізації, які б дозволили подолати існуючі обмеження.

**Мета дослідження.** Головною метою цього наукового дослідження є проведення вичерпного, глибокого та деталізованого аналітичного огляду методів обробки природної мови, відстеження їхньої архітектурної еволюції та систематизація сфер практичного застосування на сучасному етапі технологічного розвитку (з особливим фокусом на інновації останніх років). Для досягнення цієї мети у роботі поставлено та вирішено такі завдання: систематизувати історичний перехід від класичних статистичних моделей машинного навчання до сучасних трансформерних архітектур та новітніх великих моделей міркування; проаналізувавши математичні та концептуальні причини цих зсувів; проаналізувати нетрадиційні та перспективні підходи до оптимізації великих мовних моделей, зокрема застосування нейроеволюційних алгоритмів та еволюційного пошуку архітектур (Neural Architecture Search, NAS); розробити вичерпну класифікацію ключових функціональних завдань та прикладних сфер застосування методів машинного навчання для аналізу неструктурованих текстів; провести комплексний аналіз глобальної наукової літератури для виявлення прихованих тенденцій.

**Виклад основного матеріалу.** Обробка природної мови (NLP) – це комплексна, глибоко міждисциплінарна галузь науки, що знаходиться на перетині інформатики (зокрема, штучного інтелекту) та структурної лінгвістики. Її головне завдання полягає у розробці обчислювальних алгоритмів та математичних моделей, здатних автоматично

сприймати, розуміти, аналізувати та генерувати людську мову у формі тексту або мовлення. В історичній ретроспективі методи та алгоритми NLP пройшли довгий та складний шлях трансформації, який концептуально поділяється на три фундаментальні парадигми.

Символьна парадигма (Rule-Based NLP)[2]: На початкових етапах розвитку штучного інтелекту, починаючи з середини двадцятого століття, домінували системи, засновані на правилах. Ці підходи вимагали колосальної ручної праці фахівців-лінгвістів, які створювали вичерпні словники, жорсткі граматичні дерева та складні набори лексичних і синтаксичних шаблонів. Хоча цей підхід забезпечував високу точність та передбачуваність у жорстко обмежених, вузькоспеціалізованих доменах, він зіткнувся з непереборною перешкодою: природна мова є занадто гнучкою, неоднозначною та динамічною. Символьні системи виявилися абсолютно немасштабованими.

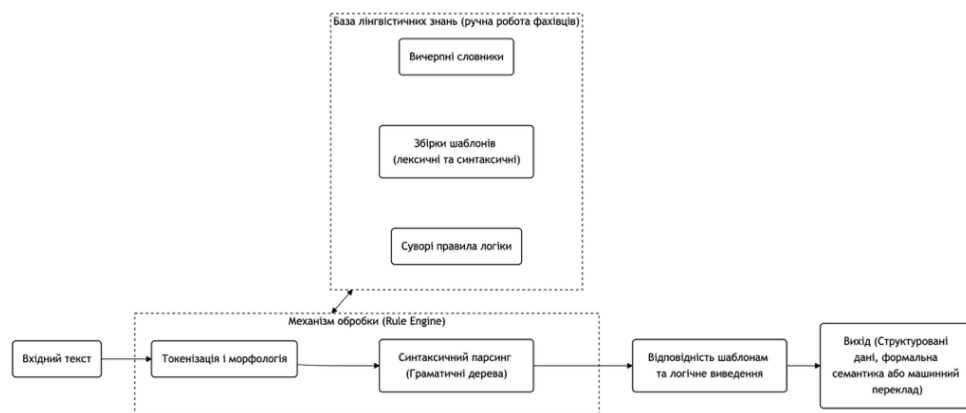


Рисунок 1 – Архітектура Rule-Based NLP

Статистична парадигма: З розвитком обчислювальних потужностей та появою великих оцифрованих текстових корпусів наприкінці 1980-х та у 1990-х роках, дослідники здійснили парадигмальний зсув у бік імовірнісних моделей та класичного машинного навчання (Machine Learning). У цей період набули поширення методи, що аналізували мову не через призму граматики, а через частотність появи тих чи інших слів та їх комбінацій (n-gram)[3]. Популярними інструментами стали наївний байєвський класифікатор (Naive Bayes), метод опорних векторів (Support Vector Machines, SVM), алгоритми k-найближчих сусідів (k-Nearest Neighbor), дерева рішень, ансамблеві методи (Random Forest) та приховані марковські моделі (Hidden Markov Models). Для перетворення тексту в математичний вигляд масово використовувалися методи на кшталт TF-IDF (Term Frequency-Inverse Document Frequency):

$$TF - IDF(t, d, D) = TF(t, d) \cdot \log \left( \frac{|D|}{|\{d \in D : t \in d\}|} \right) \quad (1)$$

де  $t$  – термін (слово), вага якого обчислюється,  $d$  – конкретний документ із колекції тексту,  $D$  – загальна множина всіх документів (корпус),  $TF(t, d)$  – частота терміну  $t$  у документі  $d$ ,  $|D|$  – загальна кількість документів у корпусі,  $|\{d \in D : t \in d\}|$  – кількість документів, у яких зустрічається термін  $t$ .

Статистичні методи дозволили алгоритмам виявляти патерни у великих обсягах даних без необхідності програмування кожного правила, що суттєво покращило показники у завданнях класифікації текстів та базового аналізу тональності. Проте цей підхід мав фундаментальний недолік, векторизація на основі частотності ігнорувала порядок слів у реченні та була нездатна вловлювати глибокі семантичні зв'язки між словами, що знаходяться на великій відстані одне від одного.

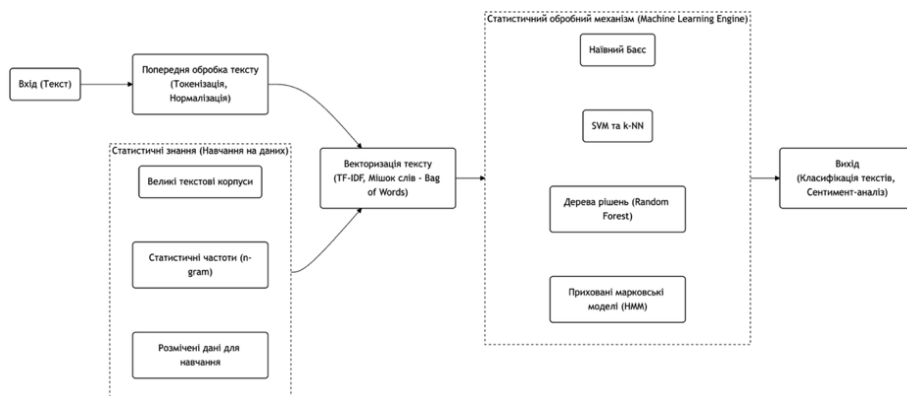


Рисунок 2 – Архітектура статистичної парадигми

Парадигма глибокого навчання (Deep Learning): Справжня наукова революція відбулася з масовим впровадженням штучних нейронних мереж (ШНМ) із багатьма прихованими шарами. Ключовим проривом стала поява методів щільного векторного представлення слів, таких як Word2Vec та GloVe [4], які дозволили відображати слова як вектори у багатовимірному семантичному просторі, де відстань між векторами відображала смислову близькість слів. Для роботи з послідовною природою тексту дослідники почали масово використовувати рекурентні нейронні мережі (Recurrent Neural Networks, RNN) та їх вдосконалену архітектуру, мережі довгої короткочасної пам'яті (Long Short-Term Memory, LSTM)[5]. Ці архітектури зберігали "пам'ять" про попередні слова при обробці наступних, що дозволило значно краще моделювати контекст. Однак, послідовний характер обчислень у RNN унеможлилював ефективне розпаралелювання процесів на сучасних графічних процесорах (GPU), що стало критичним вузьким місцем для подальшого збільшення обсягів навчальних даних та розмірів моделей.

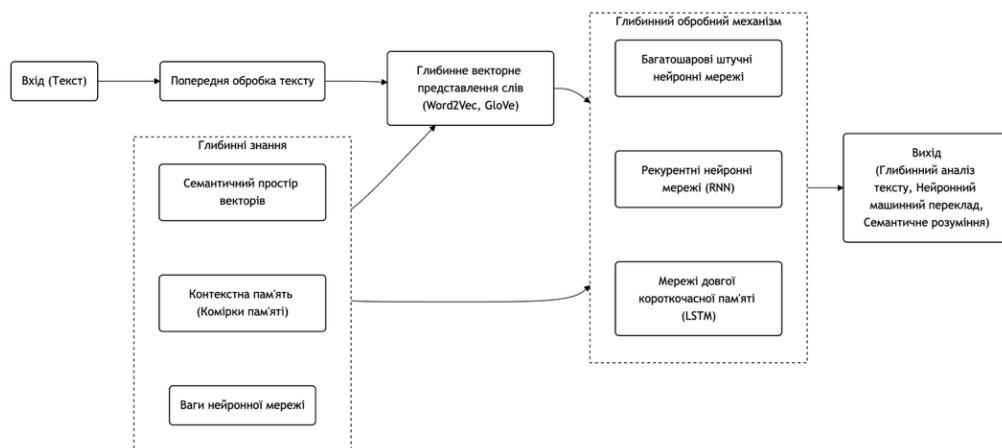


Рисунок 3 – Архітектура глибокого навчання

Фундаментальний тектонічний зсув у технологіях обробки природної мови відбувся у 2017 році після публікації знаковою праці дослідників Google під назвою "Attention is All You Need" [6]. Ця робота представила світу архітектуру Transformer, яка повністю відмовилася від рекурентних та згорткових блоків на користь інноваційного механізму так званої "внутрішньої уваги" (Self-Attention).

Архітектура трансформера дозволила обробляти всі слова у вхідній послідовності паралельно. Математично механізм самостійної уваги обчислює зв'язки між токенами за допомогою матриць запитів, ключів та значень за такою формулою:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

де  $Q$  – матриця запитів,  $K$  – матриця ключів,  $V$  – матриця значень,  $d_k$  – розмірність векторів ключів,  $K^T$  – транспонована матриця ключів.

Позиційне кодування [7] та залишкові з'єднання [8] з нормалізацією шарів вирішили проблеми збереження порядку слів та зникнення градієнта у глибоких мережах. Еволюція архітектур охопила двонаправлені (BERT [9]), генеративні (GPT-3 [10]) та сучасні моделі міркування (DeepSeek-R1 [11], Qwen [12]).

Для подолання квадратичної складності механізму уваги при обробці наддовгих контекстів сьогодні впроваджуються архітектури "суміші експертів" (MoE) [13] та системи генерації, доповненої пошуком (RAG), які оновлюють знання бази без ресурсомісткого перенавчання. Оскільки стандартні градієнтні оптимізатори (зокрема Adam) схильні до збіжності в субоптимальних локальних мінімумах і не здатні самостійно шукати топологію мережі [14], активно досліджується нейроеволюція. Цей метод оптимізує ваги та архітектури через ітеративні мутації та кросинговер. Довівши ефективність на ранніх моделях RNN/LSTM (алгоритми NEAT), сьогодні нейроеволюція фокусується на автоматизації створення трансформерів. Зокрема, еволюційний пошук архітектур (NAS) дозволяє алгоритмам самостійно конструювати оптимальні блоки трансформера [15] (порівняльний аналіз методів наведено у Таблиці 1).

Таблиця 1

Ефективність методів еволюційного пошуку архітектур

Метод / Архітектура	Особливості оптимізації моделі	Основний результат експериментів
Primer [16]	Використання квадратичної функції активації та шару глибинної згортки після кожної Q, K, V проекції.	Зниження потреби в обчислювальних ресурсах до 25-33% порівняно зі стандартними архітектурами масштабу GPT-3.
T-Razor	Використання підходу "нульових проксі" (метрика DSS++) для вибору точок алгоритмічної мутації.	Екстремальне скорочення часу пошуку ідеального трансформера з 24 діб до 0,4 доби.
SAINT+ [17]	Автоматичне впровадження операцій згортки у структуру трансформера для моделювання відстеження знань.	Досягнення ідеального балансу між моделюванням глобального (довгострокові навички) та локального (забування) контексту.

На фундаментальному рівні сучасні дослідники відзначають глибокі концептуальні паралелі між роботою великих мовних моделей та еволюційними алгоритмами: репрезентація окремого токена нагадує репрезентацію генетичного індивіда, вбудовування позиції відповідає функції формування пристосованості популяції, механізм уваги діє подібно до біологічного відбору, а циклічне навчання трансформерних блоків імітує відтворення та адаптацію параметрів у процесі боротьби за виживання [18]. Це взаємопроникнення відкриває величезні перспективи для створення стійких до збоїв штучних агентів, здатних еволюціонувати в реальному часі під час розв'язання нетипових лінгвістичних чи логічних завдань.

Алгоритми штучного інтелекту демонструють високу ефективність у подоланні інформаційного перевантаження під час наукових досліджень [19, 20, 21], а також у немедичних сферах: інженерії програмного забезпечення [22] та бізнес-аналітиці [23, 24]. Узагальнення експериментальних результатів у цих галузях наведено у Таблиці 2.

Таблиця 2

Ефективність NLP-методів у автоматизації досліджень та аналітиці

Сфера застосування	Використані моделі / Інструменти	Ключові результати та статистика експериментів
Автоматизація SLR (Огляди літератури)	TF-IDF, Sentence-BERT, Mirror-BERT	Досягнення стабільної точності класифікації на рівні 69% на відкритих масивах.
Генеративна підтримка SLR	Великі мовні моделі	Охоплення етапів створення огляду: стратегія пошуку (41%), первинний скринінг (38%), пряма екстракція (30%).
Інженерія вимог до ПЗ	Моделі глибокого навчання та LLM	Кластер "Системи NLP" формує 42,07% усього масиву світової літератури у цій галузі.
Фінансова класифікація (GICS)	Модель RoBERTa	Радикальне зниження вартості роботи; досягнуто показника якості класифікації 0,81.

**Аналіз результатів експериментів базових мовних моделей.** Еволюція трансформерних архітектур характеризується експоненційним зростанням обсягів параметрів та переходом від статистичної імітації мови до логічного міркування. Для об'єктивної оцінки цього прогресу ключові характеристики фундаментальних моделей зведено до строгих кількісних показників, отриманих на еталонних бенчмарках (Таблиця 3).

Як свідчать кількісні результати експериментів, сучасний етап розвитку ШІ (2025 рік) ознаменувався розв'язанням проблеми фактологічних галюцинацій. Якщо ранні моделі покладалися на статистичне передбачення наступного токена, то інноваційні архітектури (наприклад, лінійка DeepSeek), навчені виключно через масштабне навчання з підкріпленням (RL), здатні самостійно верифікувати логіку перед генерацією відпові-

ді. Застосування цієї парадигми дозволило моделі DeepSeek-R1 досягти точності 79.8% на складному математичному бенчмарку AIME 2024, зрівнявшись із найпотужнішими закритими системами. Водночас експериментальні метрики Qwen2.5-VL (96.4% точності на DocVQA) доводять успішну адаптацію когнітивного міркування до мультимодальних просторово-часових даних.

Таблиця 3

Хронологічна еволюція та кількісні показники продуктивності базових архітектур

Рік	Модель	Архітектурна інновація	К-сть параметрів
2018	BERT	Двонаправлене розуміння контексту (маскування слів).	110–340 млн
2020	GPT-3	Авторегресивна генерація, навчання без прикладів (zero-shot).	175 млрд
2025	DeepSeek-R1	Міркування через RL ("ланцюжки міркувань").	671 млрд
2025	Qwen2.5-VL	Уніфікована мультимодальність (візуально-мовний аналіз).	до 72 млрд

**Аналіз та глобальні статистичні тенденції розвитку NLP.** Для формування цілісного та об'єктивного розуміння вектора розвитку методів обробки природної мови (NLP) було інтегровано результати комплексного бібліометричного аналізу[25], проведеного на основі даних Scopus за період з 1992 по 2025 роки. Аналіз підтверджує, що NLP та класифікація тексту є одними з найбільш динамічних напрямів у галузі штучного інтелекту, що демонструють стійке зростання наукового інтересу та цитування.

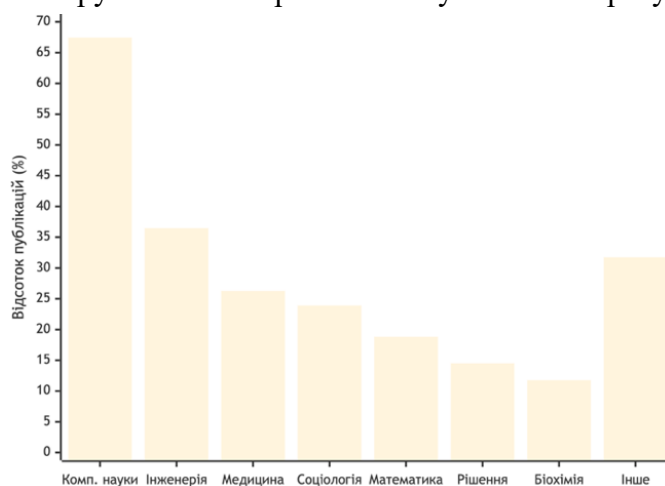


Рисунок 4 – Статистика по галузям

Згідно з отриманими даними, за останні роки спостерігається значне зростання кількості досліджень, причому пікові показники припадають на 2019, 2021 та 2022 роки. Це свідчить про перехід технологій NLP від теоретичних розробок до широкого практичного впровадження, зокрема завдяки розвитку моделей глибокого навчання та архітектури трансформерів.

З точки зору міждисциплінарності, NLP продовжує домінувати в технічних науках, проте його методи швидко інтегруються в інші галузі.

У географічному розрізі світове лідерство у виробництві наукових знань з NLP належить Сполученим Штатам, Китаю та Індії, які виступають центральними хабами глобальної мережі цитування. Також активну позицію займають Німеччина, Південна Корея та Велика Британія. Така концентрація зумовлена наявністю значних обчислювальних потужностей та доступом до капіталів великих технологічних корпорацій.

Важливим аспектом сучасного етапу розвитку є виявлення етичних та технологічних дисбалансів. Бібліометричний аналіз ключових слів вказує на те, що поряд із технічними термінами (deep learning, transformers), все частіше з'являються теми, пов'язані з упередженістю та справедливістю (bias and fairness).

Сучасні алгоритми машинного навчання вимагають колосальних обсягів даних, що створює проблему "цифрового вимирання" для мов із обмеженими ресурсами (low-resource languages). Наукова спільнота наголошує на необхідності підвищення прозорості та підзвітності ШІ-моделей, щоб забезпечити рівний доступ до технологій незалежно від мовної групи.

**Висновки.** Проведений огляд доводить, що методи обробки природної мови (NLP) стали найбільш динамічним сегментом штучного інтелекту. Еволюція від символічних систем до трансформерних архітектур повністю змінила ландшафт обчислювальної лінгвістики та розширила межі взаємодії людини й машини.

По-перше, спостерігається фундаментальний зсув від статистичної імітації до алгоритмічного міркування. Поява великих моделей міркування (LRMs, наприклад, DeepSeek-R1), здатних генерувати внутрішні ланцюжки роздумів через навчання з підкріпленням, свідчить про зародження прото-когнітивних систем. Їхня здатність до самоверифікації логіки потенційно вирішує хронічну проблему фактологічних галюцинацій у точних науках.

По-друге, оскільки масштаби моделей наблизилися до меж кремнієвої архітектури, класичні методи оптимізації стають недостатніми. Дослідження доводять, що майбутнє належить поєднанню нейромереж з еволюційними алгоритмами (Neuroevolution). Використання методів NAS (T-Razor) та нових математичних примітивів уваги (Primer) здатне радикально зменшити ресурсомісткість і відкрити шлях до розгортання потужного ШІ на локальних пристроях користувачів.

По-третє, аналіз прикладних сфер (зокрема медицини та юриспруденції) вказує на необхідність переходу від універсальних "мега-моделей" до систем, доповнених пошуком (RAG), та експертних архітектур. Варіативність точності базових LLM у медицині (від 3% до 90%) доводить, що статистична пам'ять мережі не здатна повністю замінити фахову доказову базу.

Подальші дослідження мають зосередитися на подоланні етичних викликів: усуненні алгоритмічних упереджень та забезпеченні цифрової рівності для малоресурсних мов. Технологічний прорив очікується у розвитку нейросимволічного ШІ, який поєднає

гнучкість розпізнавання образів з доказовою строгістю логіки, а також в адаптації NLP до квантових обчислень, що остаточно зніме бар'єри складності механізмів уваги.

#### ЛІТЕРАТУРА

1. ElectroIQ. Natural Language Processing Statistics By Market, Revenue And Trends (2025). 2025. URL: <https://electroi.com/stats/natural-language-processing-statistics/>
2. Strengths and Weaknesses of LLM-Based and Rule-Based NLP Technologies and Their Potential Synergies. *Electronics*. 2025. Vol. 14, № 15. P. 3064.  
URL: <https://www.mdpi.com/2079-9292/14/15/3064>
3. Golec J., Hachaj T. Ten Natural Language Processing Tasks with Generative Artificial Intelligence. *Applied Sciences*. 2025. Vol. 15, № 16. P. 9057.  
URL: <https://www.mdpi.com/2076-3417/15/16/9057>
4. Towards an evolutionary-based approach for natural language processing / L. Manzoni та ін. *arXiv preprint arXiv:2004.13832*. 2020. URL: <https://arxiv.org/abs/2004.13832>
5. Natural language processing for analyzing online customer reviews: a survey, taxonomy, and open research challenges. *PeerJ Comput. Sci.* 2024.  
URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11323031/>
6. Attention Is All You Need / A. Vaswani та ін. *arXiv preprint arXiv:1706.03762*. 2017.  
URL: <https://arxiv.org/abs/1706.03762>
7. A 2D Semantic-Aware Position Encoding for Vision Transformers. *arXiv preprint arXiv:2505.09466*. 2025. URL: <https://arxiv.org/html/2505.09466v1>
8. ReViT: Enhancing Vision Transformers Feature Diversity with Attention Residual Connections / A. Diko та ін. *arXiv preprint arXiv:2402.11301*. 2024.  
URL: <https://arxiv.org/abs/2402.11301>
9. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding / J. Devlin та ін. *arXiv preprint arXiv:1810.04805*. 2018. URL: <https://arxiv.org/abs/1810.04805>
10. Generative Pre-trained Transformer: A Comprehensive Review on Enabling Technologies, Potential Applications, Emerging Challenges, and Future Directions / G. Yenduri та ін. *arXiv preprint arXiv:2305.10435*. 2023.  
URL: <https://arxiv.org/abs/2305.10435>
11. DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning / DeepSeek-AI та ін. *arXiv preprint arXiv:2501.12948*. 2025.  
URL: <https://arxiv.org/abs/2501.12948>
12. Qwen2.5-VL Technical Report / S. Bai та ін. *arXiv preprint arXiv:2502.13923*. 2025.  
URL: <https://arxiv.org/abs/2502.13923>
13. Speed Always Wins: A Survey on Efficient Architectures for Large Language Models / W. Sun та ін. *arXiv preprint arXiv:2508.09834*. 2025.  
URL: <https://arxiv.org/abs/2508.09834>
14. Enhancing Neural Network Training Through Neuroevolutionary Models: A Hybrid Approach to Classification Optimization. *Mathematics*. 2025. Vol. 13, № 7. P. 1114.  
URL: <https://www.mdpi.com/2227-7390/13/7/1114>
15. PubMed Article. PMID: 38502633. 2024.  
URL: <https://pubmed.ncbi.nlm.nih.gov/38502633/>

16. Searching for Efficient Transformers for Language Modeling. *Advances in Neural Information Processing Systems (NeurIPS)*. 2021. URL: [https://papers.neurips.cc/paper\\_files/paper/2021/file/2f3c6a4cd8af177f6456e7e51a916ff3-Paper.pdf](https://papers.neurips.cc/paper_files/paper/2021/file/2f3c6a4cd8af177f6456e7e51a916ff3-Paper.pdf)
17. Evolutionary Neural Architecture Search for Transformer in Knowledge Tracing. *OpenReview*. 2024. URL: <https://openreview.net/pdf?id=G14N38AjpU>
18. When Large Language Models Meet Evolutionary Algorithms: Potential Enhancements and Challenges / C. Wang та ін. *Science and Technology Review*. 2025. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11948732/>
19. Automated Research Article Classification and Recommendation Using NLP and ML / S. Rahman та ін. *arXiv preprint arXiv:2510.05495*. 2025. URL: <https://arxiv.org/html/2510.05495v1>
20. Improving Systematic Review Updates With Natural Language Processing Through Abstract Component Classification and Selection: Algorithm Development and Validation. 2025. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11986382/>
21. Large language models for conducting systematic reviews: on the rise, but not yet ready for use – a scoping review / Lieberum та ін. *medRxiv*. 2024. URL: <https://www.medrxiv.org/content/10.1101/2024.12.19.24319326v1.full.pdf>
22. A Systematic Literature Review on Using Natural Language Processing in Software Requirements Engineering. *Electronics*. 2024. Vol. 13, № 11. P. 2055. URL: <https://www.mdpi.com/2079-9292/13/11/2055>
23. Comparative Analysis of NLP-Based Models for Company Classification. *Information*. 2024. Vol. 15, № 2. P. 77. URL: <https://www.mdpi.com/2078-2489/15/2/77>
24. Natural language processing in impact assessment: a review of applications and concerns / C. Hung, S. Kumar. *Impact Assessment and Project Appraisal*. 2025. URL: <https://www.tandfonline.com/doi/full/10.1080/14615517.2026.2620186>
25. Scientific Journal of Engineering Research (SJER). URL: <https://journal.futuristech.co.id/index.php/sjer/article/view/6>

#### REFERENCES

1. ElectroIQ. (2025). Natural Language Processing Statistics By Market, Revenue And Trends (2025). Retrieved from <https://electroi.com/stats/natural-language-processing-statistics/>
2. Strengths and Weaknesses of LLM-Based and Rule-Based NLP Technologies and Their Potential Synergies. (2025). *Electronics*, 14(15), 3064. Retrieved from <https://www.mdpi.com/2079-9292/14/15/3064>
3. Golec, J., & Hachaj, T. (2025). Ten Natural Language Processing Tasks with Generative Artificial Intelligence. *Applied Sciences*, 15(16), 9057. Retrieved from <https://www.mdpi.com/2076-3417/15/16/9057>
4. Manzoni, L., Jakobovic, D., Mariot, L., Picek, S., & Castelli, M. (2020). Towards an evolutionary-based approach for natural language processing. *arXiv preprint arXiv:2004.13832*. Retrieved from <https://arxiv.org/abs/2004.13832>

5. Natural language processing for analyzing online customer reviews: a survey, taxonomy, and open research challenges. (2024). *PeerJ Comput. Sci.* Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC11323031/>
6. Vaswani, A., et al. (2017). Attention Is All You Need. *arXiv preprint arXiv:1706.03762*. Retrieved from <https://arxiv.org/abs/1706.03762>
7. A 2D Semantic-Aware Position Encoding for Vision Transformers. (2025). *arXiv preprint arXiv:2505.09466*. Retrieved from <https://arxiv.org/html/2505.09466v1>
8. Diko, A., Avola, D., Cascio, M., & Cinque, L. (2024). ReViT: Enhancing Vision Transformers Feature Diversity with Attention Residual Connections. *arXiv preprint arXiv:2402.11301*. Retrieved from <https://arxiv.org/abs/2402.11301>
9. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv preprint arXiv:1810.04805*. Retrieved from <https://arxiv.org/abs/1810.04805>
10. Yenduri, G., et al. (2023). Generative Pre-trained Transformer: A Comprehensive Review on Enabling Technologies, Potential Applications, Emerging Challenges, and Future Directions. *arXiv preprint arXiv:2305.10435*. Retrieved from <https://arxiv.org/abs/2305.10435>
11. DeepSeek-AI, et al. (2025). DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning. *arXiv preprint arXiv:2501.12948*. Retrieved from <https://arxiv.org/abs/2501.12948>
12. Bai, S., et al. (2025). Qwen2.5-VL Technical Report. *arXiv preprint arXiv:2502.13923*. Retrieved from <https://arxiv.org/abs/2502.13923>
13. Sun, W., et al. (2025). Speed Always Wins: A Survey on Efficient Architectures for Large Language Models. *arXiv preprint arXiv:2508.09834*. Retrieved from <https://arxiv.org/abs/2508.09834>
14. Enhancing Neural Network Training Through Neuroevolutionary Models: A Hybrid Approach to Classification Optimization. (2025). *Mathematics*, 13(7), 1114. Retrieved from <https://www.mdpi.com/2227-7390/13/7/1114>
15. PubMed Article. (2024). PMID: 38502633. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/38502633/>
16. Searching for Efficient Transformers for Language Modeling. (2021). *Advances in Neural Information Processing Systems (NeurIPS)*. Retrieved from [https://papers.neurips.cc/paper\\_files/paper/2021/file/2f3c6a4cd8af177f6456e7e51a916ff3-Paper.pdf](https://papers.neurips.cc/paper_files/paper/2021/file/2f3c6a4cd8af177f6456e7e51a916ff3-Paper.pdf)
17. Evolutionary Neural Architecture Search for Transformer in Knowledge Tracing. (2024). *OpenReview*. Retrieved from <https://openreview.net/pdf?id=G14N38AjpU>
18. Wang, C., et al. (2025). When Large Language Models Meet Evolutionary Algorithms: Potential Enhancements and Challenges. *Science and Technology Review*. Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC11948732/>
19. Rahman, S., Shanto, H. K., Koana, U. A., & Danish, S. M. (2025). Automated Research Article Classification and Recommendation Using NLP and ML. *arXiv preprint arXiv:2510.05495*. Retrieved from <https://arxiv.org/html/2510.05495v1>

20. Improving Systematic Review Updates With Natural Language Processing Through Abstract Component Classification and Selection: Algorithm Development and Validation. (2025). Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC11986382/>
21. Lieberum, et al. (2024). Large language models for conducting systematic reviews: on the rise, but not yet ready for use – a scoping review. *medRxiv*. Retrieved from <https://www.medrxiv.org/content/10.1101/2024.12.19.24319326v1.full.pdf>
22. A Systematic Literature Review on Using Natural Language Processing in Software Requirements Engineering. (2024). *Electronics*, 13(11), 2055. Retrieved from <https://www.mdpi.com/2079-9292/13/11/2055>
23. Comparative Analysis of NLP-Based Models for Company Classification. (2024). *Information*, 15(2), 77. Retrieved from <https://www.mdpi.com/2078-2489/15/2/77>
24. Hung, C., & Kumar, S. (2025). Natural language processing in impact assessment: a review of applications and concerns. *Impact Assessment and Project Appraisal*. Retrieved from <https://www.tandfonline.com/doi/full/10.1080/14615517.2026.2620186>
25. Scientific Journal of Engineering Research (SJER). Retrieved from <https://journal.futuristech.co.id/index.php/sjer/article/view/6>

Received 16.04.2026.  
Accepted 20.04.2026.  
Published 30.04.2026

***Review of natural language processing methods:  
evolution, architectures, and application domains***

*The global Natural Language Processing (NLP) market demonstrates rapid, almost exponential growth, reflecting immense commercial interest and the digital economy's reliance on AI algorithms. However, the rapid scaling of neural network parameters—from millions to hundreds of billions—has introduced new computational and ecological challenges. Furthermore, models inevitably inherit biases present in training data, raising critical ethical and safety concerns. There is an urgent need for novel, highly efficient architectures and transparent optimization methods to overcome these limitations.*

*The primary goal of this research is to conduct a comprehensive analytical review of NLP methods, trace their architectural evolution from traditional statistical machine learning models to modern Large Language Models (LLMs) and Large Reasoning Models (LRMs), and systematize their practical application domains at the current stage of technological development.*

*The evolution of NLP is conceptually divided into three fundamental paradigms: rule-based symbolic systems, statistical methods (e.g., TF-IDF, SVM, n-grams), and deep learning. A tectonic shift occurred with the introduction of the Transformer architecture, which replaced recurrent blocks with self-attention mechanisms, enabling parallel processing. To address the quadratic computational complexity of processing long contexts, the study analyzes non-traditional optimization approaches, primarily Neuroevolution and Neural Architecture Search (NAS), which automatically construct optimal transformer blocks. The paper also systematizes the practical application of NLP in healthcare, software engineering, and business analytics, demonstrating varying degrees of accuracy and highlighting the necessity of Retrieval-Augmented Generation (RAG) technologies.*

*NLP has become the most dynamic segment of artificial intelligence. A fundamental shift from statistical imitation to algorithmic reasoning is observed, evidenced by LRMs like DeepSeek-R1, which are capable of self-verifying logic and mitigating factual hallucinations. As models approach the limits of silicon architecture, the future lies in combining neural networks with evolutionary algorithms. Future research must focus on overcoming ethical challenges, ensuring digital equality for low-resource languages, and advancing neurosymbolic AI and quantum computing adaptations.*

*Keywords: artificial intelligence, natural language processing, large language models, transformer architectures, neuroevolution, sentiment analysis, bibliometric analysis, large reasoning models, machine learning, research automation.*

**Ющенко Артур Сергійович** – аспірант кафедри інженерії програмного забезпечення, Харківського національного університету радіоелектроніки, Україна.

ORCID: <https://orcid.org/0009-0007-7995-987X>

**Смеляков Кирило Сергійович** – д.т.н., професор кафедри інженерії програмного забезпечення, Харківського національного університету радіоелектроніки, Україна.

ORCID: <https://orcid.org/0000-0001-9938-5489>

**Чуприна Анастасія Сергіївна** – к.т.н., доцент кафедри інженерії програмного забезпечення, Харківського національного університету радіоелектроніки, Україна.

ORCID: <https://orcid.org/0000-0003-0394-9900>

**Yushchenko Artur Serhiyovych** – Postgraduate student at the Department of Software Engineering, Kharkiv National University of Radio Electronics, Ukraine.

ORCID: <https://orcid.org/0009-0007-7995-987X>

**Smelyakov Kirill Sergeevich** – Doctor of Technical Sciences, Professor at the Department of Software Engineering, Kharkiv National University of Radio Electronics, Ukraine.

ORCID: <https://orcid.org/0000-0001-9938-5489>

**Chupryna Anastasia Sergeevna** – Candidate of Technical Sciences, Associate Professor at the Department of Software Engineering, Kharkiv National University of Radio Electronics, Ukraine.

ORCID: <https://orcid.org/0000-0003-0394-9900>

О.І. Панфілов, М.М. Нестеренко

## ДИНАМІКА СИСТЕМИ «ВІБРАЦІЙНА УСТАНОВКА – ШАР БЕТОНУ – ІМПУЛЬСНИЙ ПРИВАНТАЖУВАЧ» З УРАХУВАННЯМ ДІЇ ОСНОВНОГО ВІБРОЗБУДЖУВАЧА ТА ВИЩИХ ГАРМОНІК

*Анотація.* У статті розглянуто динаміку системи «вібраційна установка – шар бетону – імпульсний привантажувач» з урахуванням дії основного віброзбуджувача та вищих гармонік коливального процесу. Актуальність дослідження зумовлена необхідністю підвищення ефективності вібраційного ущільнення бетонних сумішей, особливо за умов нерівномірного передавання енергії коливань у товщу матеріалу та впливу хвилових процесів на формування зон підсилення й ослаблення динамічної дії. Метою роботи є розроблення узагальненої математичної моделі системи, яка враховує взаємодію рухомої рами з формою, бетонної суміші та імпульсного привантажувача, а також обґрунтування умов ефективного режиму доущільнення. У дослідженні використано методи аналітичного моделювання динамічних процесів, засновані на складанні рівнянь руху підсистем, урахуванні нелінійної пружно-дисипативної реакції бетонної суміші, контактної взаємодії з привантажувачем і багатогармонічного подання збудження. Отримано розрахункові співвідношення для опису контактної та безконтактної фаз руху привантажувача, оцінювання тривалості контакту, коефіцієнта відновлення швидкості та умов існування усталеного періодичного режиму. Практичне значення одержаних результатів полягає у можливості використання запропонованої моделі для вибору параметрів імпульсного привантажувача, оцінювання стійкості режимів роботи та вдосконалення вібраційних установок для формування бетонних виробів.

*Ключові слова:* бетонна суміш, вібраційне ущільнення, імпульсний привантажувач, вищі гармоніки, віброзбуджувач, контактна взаємодія, динамічна система, стійкість коливань.

**Постановка проблеми.** Вібраційне ущільнення бетонних сумішей є одним із ключових процесів технології виробництва бетонних і залізобетонних виробів, оскільки саме характер динамічної взаємодії між робочими органами обладнання та оброблюваним середовищем визначає щільність, однорідність структури та фізико-механічні властивості готової продукції. У сучасних умовах зростання вимог до якості виробів, енергоефективності технологій та інтенсифікації виробничих процесів особливої актуальності набуває вдосконалення режимів вібраційного впливу на бетонні суміші.

Традиційні гармонічні режими віброущільнення, що реалізуються у більшості існуючих вібраційних машин, мають обмежену ефективність, особливо при ущільненні жорстких, малорухомих та легких бетонних сумішей, а також при формуванні виробів значної товщини. Це пов'язано з тим, що енергія коливань передається у товщу матеріалу нерівномірно, а хвильові процеси, що виникають у бетонному середовищі, можуть призводити до формування як зон підсилення, так і зон ослаблення динамічного впливу. Зокрема, при певних співвідношеннях між частотою коливань і геометричними параметрами шару можливе виникнення антирезонансних режимів, що супроводжуються різким зменшенням амплітуди та ефективності ущільнення.

Водночас дослідження показують, що бетонну суміш доцільно розглядати як середовище з розподіленими параметрами, у якому поширюються пружні та дисипативні хвилі, а її реакція на динамічне навантаження визначається не лише миттєвими значеннями сил, але й історією попередніх впливів. За таких умов ефективність ущільнення суттєво залежить від узгодження параметрів коливального процесу з хвильовими характеристиками середовища, зокрема швидкістю поширення хвиль, коефіцієнтом демпфування та висотою шару бетонної суміші.

**Аналіз останніх досліджень.** Проблема підвищення ефективності вібраційного ущільнення бетонних сумішей широко висвітлена в наукових працях, де основну увагу приділено взаємодії робочих органів машин із бетонним середовищем, визначенню раціональних параметрів коливань і вдосконаленню конструкцій обладнання.

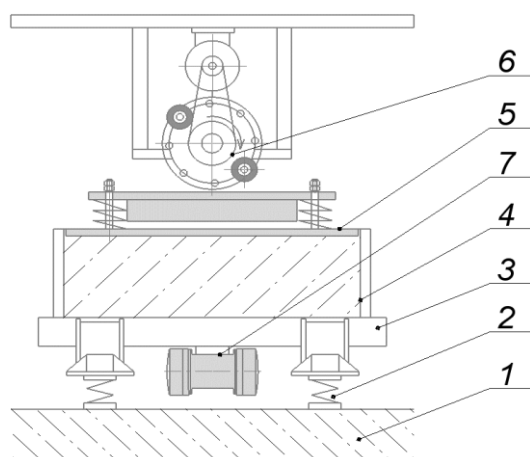
У роботах [1, 2] розглянуто теоретичні основи ущільнення бетонних сумішей і показано, що його ефективність визначається динамічним тиском, сформованим коливальним впливом. Встановлено, що вертикальні режими сприяють руйнуванню структурних зв'язків і підвищенню щільності суміші, однак потребують узгодження параметрів збудження з її властивостями. У працях [2, 3] бетонну суміш подано як складну багатозфазну систему, у якій одночасно проявляються інерційні та пружно-дисипативні процеси. Показано, що інтенсивність ущільнення залежить не лише від амплітуди, а й від частоти, прискорень і характеру силової взаємодії. Авторами [4] наведено огляд сучасних конструктивних схем віброплощадок. Зазначено, що традиційне регулювання частоти й амплітуди не завжди забезпечує формування складних динамічних режимів, необхідних для інтенсифікації ущільнення. У дослідженні [5] підтверджено ефективність імпульсних і комбінованих режимів, які дають змогу скоротити тривалість процесу та підвищити ступінь ущільнення завдяки короткочасним піковим напруженням. Водночас підкреслено потребу в теоретичному обґрунтуванні параметрів імпульсного впливу. У роботі [6] показано, що важливу роль у формуванні динамічного впливу відіграє спектральний склад коливань, зокрема вищі гармоніки, які необхідно враховувати під час аналізу процесу ущільнення. У працях [7, 8] бетонну суміш розглянуто як пружно-в'язке середовище з розподіленими параметрами, в якому поширюються хвильові процеси. Встановлено, що реакція середовища визначається не лише локальними властивостями, а й умовами поширення хвиль та їх взаємодією з межами. Експериментальні результати [9] підтверджують доцільність поєднання гармонічного та імпульсного збу-

дження, особливо для інтенсифікації ущільнення у верхніх шарах суміші, однак вибір раціональних параметрів імпульсного впливу залишається відкритим.

Аналіз наведених досліджень свідчить, що, незважаючи на значний обсяг наукових робіт у галузі вібраційного ущільнення, більшість із них орієнтована на гармонічні або квазігармонічні режими та не враховує повною мірою хвильову природу процесів у бетонній суміші. Недостатньо дослідженим залишається питання узгодження параметрів імпульсного навантаження з хвильовими характеристиками середовища, а також впливу локалізованих імпульсів на формування напружено-деформованого стану у товщі матеріалу.

**Мета досліджень.** Метою дослідження є розроблення та динамічне обґрунтування параметрів імпульсного вібраційного привантажувача в умовах хвильової взаємодії з бетонною сумішшю на основі узагальненої математичної моделі системи «вібраційна установка – бетонна суміш – привантажувач», що враховує розподілений характер середовища, нелінійні пружно-дисипативні властивості матеріалу та імпульсно-гармонічний характер збудження, з метою підвищення ефективності доущільнення та енергоефективності технологічного процесу.

**Викладення основного матеріалу дослідження.** Динамічну систему вібраційної установки розглянуто як сукупність трьох взаємопов'язаних підсистем: рухомої рами з формою та бетонною сумішшю, основного віброзбуджувача та імпульсного привантажувача, який взаємодіє з верхнім шаром суміші (рисунки 1-2). На відміну від спрощених одноударних моделей, привантажувач перебуває під дією не лише імпульсного механізму, а й гармонічних коливань поверхні бетонної суміші.



1 – фундамент, 2 – опори, 3 – рухома рама,

4 – форма з бетоном, 5 – привантажувач,

6 – імпульсний механізм, 7 – віброзбуджувач

Рисунок 1 – Схема вібраційної установки з імпульсним привантажувачем для ущільнення бетону

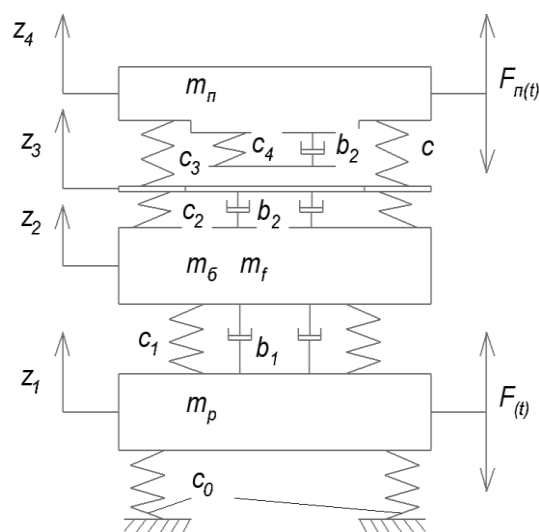


Рисунок 2 – Розрахункова схема «вібраційна установка – бетонна суміш – імпульсний привантажувач»

Узагальненими координатами системи прийнято вертикальне переміщення рухомої рами та абсолютне переміщення привантажувача, а їх відносне зближення визначається виразом (1).

$$\xi(t) = y(t) - z(t) - x_{cm}, \quad (1)$$

де  $x_{cm}$  – статичне зміщення еквівалентної контактної зони «привантажувач – бетонна суміш».

Рухома частина установки описується рівнянням (2), у якому враховано масу системи, демпфування і жорсткість опор, реакцію бетонної суміші та контактну взаємодію з привантажувачем.

$$m\ddot{z}(t) + b\dot{z}(t) + cz(t) + R_B(z, \dot{z}) + Q_K(\xi, \dot{\xi}) = \sum_{n=1}^s F_n \sin(n\omega t + \varphi_n), \quad (2)$$

де  $m$  – приведена маса рухомої системи;  $b$  – коефіцієнт демпфування опор;  $c$  – еквівалентна жорсткість опор;  $R_B(z, \dot{z})$  – нелінійна реакція бетонної суміші;  $Q_K(\xi, \dot{\xi})$  – контактна сила взаємодії з привантажувачем;  $F$  – амплітуди гармонік збудження;  $s$  – кількість гармонік, що враховуються;  $\omega$  – кругова частота основного збудження;  $\varphi_n$  – початкові фази.

Основне збудження подано багатогармонічним розкладом (3), що дає змогу врахувати вплив вищих гармонік на формування пікових прискорень.

$$F_B(t) = \sum_{n=1}^s F_n \sin(n\omega t + \varphi_n), \quad (3)$$

Рух імпульсного привантажувача описується рівнянням (4).

$$m_{\Pi} \ddot{y}(t) = -Q_K(\xi, \dot{\xi}) + \sum_{j=1}^{N_i} P_j(t), \quad N_i = 1, \dots, 8, \quad (4)$$

де  $m_{\Pi}$  – приведена маса імпульсного привантажувача;  $P_j(t)$  – сила  $j$ -го імпульсу, прикладеного в межах одного циклу роботи.

Імпульсне навантаження доцільно подати як суму короткочасних силових впливів:

$$\sum_{j=1}^{N_i} P_j(t) = \sum_{j=1}^{N_i} S_j \delta(t - t_j), \quad (5)$$

де  $S_j$  – імпульс сили  $j$ -го удару;  $t_j$  – момент його прикладання.

Нелінійну реакцію бетонної суміші задано залежністю (6), а контактну взаємодію між привантажувачем і сумішшю – формулою (7).

$$R_6(z, \dot{z}) = k_1 z + k_3 z^3 + \beta_1 \dot{z} + \beta_3 \dot{z} |z|, \quad (6)$$

де  $k_1$  – лінійна жорсткість бетонної суміші;  $k_3$  – коефіцієнт нелінійної жорсткості;  $\beta_1, \beta_3$  – коефіцієнти дисипативного опору.

Після переходу до відносної координати отримано рівняння контактної фази (7), яке показує, що на рух привантажувача істотно впливає прискорення рухомої рами, тобто дія основного віброзбуджувача передається через коливання бетонної суміші.

$$m_{II}\ddot{\xi}(t) + 2H_K\dot{\xi}(t) + c_K\xi(t) = -m_{II}\ddot{z}(t) + \sum_{j=1}^{N_i} P_j(t), \quad \xi \geq 0. \quad (7)$$

У піддемпфованому режимі розв'язок рівняння контактної взаємодії подається виразом (8), який враховує сумісну дію імпульсного та багатогармонічного збудження.

$$\xi(t) = e^{-ht} \left[ C_{li} \cos(\sqrt{\lambda^2 - h^2} t) + \frac{W_i + hC_{li}}{\sqrt{\lambda^2 - h^2}} \sin(\sqrt{\lambda^2 - h^2} t) \right] + \Xi_r(t) + \Xi_i(t), \quad (8)$$

Тривалість контакту в першому наближенні визначається за формулою (9), а ефективний коефіцієнт відновлення швидкості – за формулою (10).

$$\Delta t_i \approx \frac{\pi}{\sqrt{\lambda^2 - h^2}}. \quad (9)$$

$$R_{\text{эф},i} \approx \exp\left(-\frac{\pi h}{\sqrt{\lambda^2 - h^2}}\right). \quad (10)$$

Це дає змогу оцінити втрати енергії в контактній зоні та вплив параметрів системи на інтенсивність доущільнення.

У безконтактній фазі рух привантажувача описується рівнянням (11).

Між двома сусідніми імпульсами привантажувач рухається у безконтактній фазі. Для цієї ділянки руху, з урахуванням дії основного віброзбуджувача, маємо  $m_{II}\ddot{y}(t) = 0$ , або у відносній координаті

$$\ddot{\xi}(t) = -\ddot{z}(t), \quad \xi < 0. \quad (11)$$

Для усталеного періодичного режиму мають виконуватися умови повторюваності параметрів руху (12).

$$t_{i+1} - t_i = T_i, \quad W_{i+1} = W_i = W, \quad \xi_{0,i+1} = \xi_{0,i}, \quad (12)$$

де  $T_i$  – період між сусідніми імпульсами.

Після лінеаризації системи одержано характеристичне рівняння (13).

$$\rho^2 A + \rho B + C = 0, \quad (13)$$

де коефіцієнти  $A, B, C$  є функціями параметрів

$$\alpha, \gamma, \gamma_1, \mu_1, \mu_2, \dots, \mu_s, N_i. \quad (14)$$

Умовою стійкості періодичного режиму є

$$\rho < 1. \quad (15)$$

Це означає, що стійкі режими імпульсно-вібраційного доущільнення досягаються лише за певного поєднання жорсткості контактної зони, демпфування, частоти основного віброзбуджувача, кількості імпульсів та спектрального складу коливань. При цьому збільшення кількості імпульсів від одного до восьми може як підвищувати ефективність ущільнення за рахунок накопичення енергії та впливу вищих гармонік, так і призводити до втрати стійкості режиму, якщо моменти прикладання імпульсів збігаються з несприятливими фазами основного коливального процесу.

**Висновки:** Розроблено узагальнену математичну модель системи «вібраційна установка – шар бетону – імпульсний привантажувач», яка враховує взаємодію рухомої рами, бетонної суміші та привантажувача, нелінійні пружно-дисипативні властивості середовища, а також імпульсно-гармонічний характер збудження. Показано, що врахування вищих гармонік є необхідним для адекватного опису динаміки системи, оскільки вони можуть істотно змінювати амплітудно-фазові характеристики руху привантажувача та впливати на інтенсивність ущільнення бетонної суміші. Отримані співвідношення для контактної та безконтактної фаз руху, а також умови стійкості періодичного режиму можуть бути використані для вибору раціональних параметрів імпульсного привантажувача і режимів роботи віброзбуджувачів.

#### ЛІТЕРАТУРА

1. Нестеренко, М., & Ведмідь, В. (2025). Теоретичні положення та аналіз робочого процесу ущільнення бетонних сумішей. *Техніка будівництва*, (42), 4–13. <https://doi.org/10.32347/tb.2025-42.0501>  
<http://tehbud.knuba.edu.ua/article/view/331978/320980>.
2. Maslov A.G., Salenko J.S., Maslova N.A. Study of interaction of vibrating plate with concrete mixture // *Вісник КНУ імені Михайла Остроградського*. – 2011. – Вип. 2 (67). – С. 93–98.
3. Nesterenko M.P. Investigation of vibration machine interaction with compacted concrete mixture / M.P. Nesterenko, A.G. Maslov, Ju.S. Salenko // *International Journal of Engineering & Technology*. – 2018. – Vol. 7, № 3.2. – P. 260-264.
4. Дьяченко О., Пригоцький В., Малицький І. Огляд схем вібромайданчиків з вертикальними коливаннями та аналіз можливостей керування параметрами ущільнення // *Енергоощадні машини і технології : матеріали III Міжнар. наук.-практ. конф. (Київ, 17–19 трав. 2022 р.)*. – Київ : КНУБА, 2022. – С. 28–31.
5. Маслов О.Г. Дослідження вібраційного органу для ущільнення бетонних сумішей з віброімпульсними коливаннями / О. Г. Маслов, Ю. С. Саленко, І. І. Жовтяк, Р. А. Вакуленко, В. Л. Дятловська // *Вісник Кременчуцького національного університету імені Михайла Остроградського*. - 2020. - Вип. 5-6. - С. 139-146. - Режим доступу: [http://nbuv.gov.ua/UJRN/Vkdpu\\_2020\\_5-6\\_21](http://nbuv.gov.ua/UJRN/Vkdpu_2020_5-6_21).
6. Назаренко І. Огляд і аналіз вібраційного обладнання для формування плоских залізобетонних виробів / І. Назаренко, О. Дєдов, О. Дьяченко, А. Свідерський // *Гірничі, будівельні, дорожні та меліоративні машини*. – 2017. – Вип. 90. – С. 49–58.
7. Сердюк Л.І., Давиденко Ю.О., Костенко П.М. Деякі підходи до моделювання середовища, що обробляється вібраційним пристроєм // *Ресурсоекономні матер., констр., будівлі та споруди*. – Рівне : НУВГП, 2005. – Вип. 12. – С. 69–72.
8. Nesterenko M. M., Vedmid V. V. Дослідження впливу вертикальних вібраційних режимів на взаємодію робочого органу віброплощинки з бетонною сумішшю. *Transactions of Kremenchuk Mykhailo Ostrohradskyi National University*. 2025. No. 2. P. 182-188. URL: <https://doi.org/10.32782/1995-0519.2025.2.23>.
9. Ведмідь В.В. Експериментальні дослідження комбінованої імпульсно-вібраційної установки / В.В. Ведмідь, Т.М. Нестеренко, М.М. Нестеренко, М.О. Пирлик // *Системні*

#### REFERENCES

1. Nesterenko, M., & Vedmid, V. (2025). Teoretychni polozhennia ta analiz robochoho protsesu ushchilnennia betonnykh sumishei [Theoretical principles and analysis of the working process of concrete mix compaction]. *Tekhnika budivnytstva – Construction Engineering*, (42), 4–13. <https://doi.org/10.32347/tb.2025-42.0501>.
2. Maslov, A. G., Salenko, J. S., & Maslova, N. A. (2011). Study of interaction of vibrating plate with concrete mixture. *Visnyk Kremenchutskoho natsionalnoho universytetu imeni Mykhaila Ostrohradskoho – Transactions of Kremenchuk Mykhailo Ostrohradskyi National University*, 2(67), 93–98.
3. Nesterenko, M. P., Maslov, A. G., & Salenko, Ju. S. (2018). Investigation of vibration machine interaction with compacted concrete mixture. *International Journal of Engineering & Technology*, 7(3.2), 260–264.
4. Diachenko, O., Pryhotskyi, V., & Malitskyi, I. (2022). Ohliad skhem vibromaidanchykyv z vertykalnymy kolyvanniamy ta analiz mozhlyvostei keruvannia parametramy ushchilnennia [Review of vibration platform schemes with vertical oscillations and analysis of possibilities for controlling compaction parameters]. In *Enerhooshchadni mashyny i tekhnolohii: Materialy III Mizhnarodnoi naukovo-praktychnoi konferentsii [Energy-saving machines and technologies: Proceedings of the 3rd International Scientific and Practical Conference]* (pp. 28–31). Kyiv, Ukraine: KNUCA.
5. Maslov, O. H., Salenko, Yu. S., Zhovtiak, I. I., Vakulenko, R. A., & Diatlovska, V. L. (2020). Doslidzhennia vibratsiinoho orhanu dlia ushchilnennia betonnykh sumishei z vibroimpulsnymy kolyvanniamy [Study of a vibration unit for concrete mix compaction with vibro-impact oscillations]. *Visnyk Kremenchutskoho natsionalnoho universytetu imeni Mykhaila Ostrohradskoho – Transactions of Kremenchuk Mykhailo Ostrohradskyi National University*, (5–6), 139–146. [http://nbuv.gov.ua/UJRN/Vkdpu\\_2020\\_5-6\\_21](http://nbuv.gov.ua/UJRN/Vkdpu_2020_5-6_21).
6. Nazarenko, I., Diedov, O., Diachenko, O., & Sviderskyi, A. (2017). Ohliad i analiz vibratsiinoho obladnannia dlia formuvannia ploskykh zalizobetonnykh vyrobiv [Review and analysis of vibration equipment for forming flat reinforced concrete products]. *Hirnychi, budivelni, dorozhni ta melioratyvni mashyny – Mining, Construction, Road and Melioration Machines*, 90, 49–58.
7. Serdiuk, L. I., Davydenko, Yu. O., & Kostenko, P. M. (2005). Deiaki pidkhody do modelivannia seredovyshcha, shcho obrobliaietsia vibratsiinym prystroiem [Some approaches to modeling a medium processed by a vibration device]. *Resursoekonomni materialy, konstruktsii, budivli ta sporudy – Resource-efficient Materials, Structures, Buildings and Constructions*, 12, 69–72.
8. Nesterenko, M. M., & Vedmid, V. V. (2025). Doslidzhennia vplyvu vertykalnykh vibratsiinnykh rezhymiv na vzaiemodiiu robochoho orhanu vibroploshchadky z betonnoiu sumishshiu [Study of the influence of vertical vibration modes on the interaction of the working body of a vibration platform with concrete mixture]. *Transactions of Kremenchuk Mykhailo Ostrohradskyi National University*, (2), 182–188. <https://doi.org/10.32782/1995-0519.2025.2.23>.

9. Vedmid, V. V., Nesterenko, T. M., Nesterenko, M. M., & Pyrlyk, M. O. (2025). Eksperymentalni doslidzhennia kombinovanoi impulsno-vibratsiinoi ustanovky [Experimental studies of a combined impulse-vibration installation]. *Systemni tekhnolohii – System Technologies*, 5(160), 178–188. <https://doi.org/10.34185/1562-9945-5-160-2025-19>.

Received 17.04.2026.

Accepted 20.04.2026.

Published 30.04.2026

***Dynamics of the system «vibration installation – concrete layer – impulse surcharge device» with account of the effect of the main vibration exciter and higher harmonics***

*The article examines the dynamics of the system “vibration installation – concrete layer – impulse surcharge device,” taking into account the effect of the main vibration exciter and higher harmonics of the oscillatory process. The relevance of the study is обусловлена? Need translate fully, avoid russian accidental. Continue properly. The relevance of the study is determined by the need to improve the efficiency of vibration compaction of concrete mixtures, especially under conditions of nonuniform transfer of vibration energy into the material thickness and the influence of wave processes on the formation of zones of intensified and weakened dynamic action. The aim of the work is to develop a generalized mathematical model of the system that takes into account the interaction of the moving frame with the mould, the concrete mixture, and the impulse surcharge device, as well as to substantiate the conditions for an effective recompaction mode. The study employs methods of analytical modeling of dynamic processes based on the formulation of the equations of motion of the subsystems, taking into account the nonlinear elastic-dissipative response of the concrete mixture, the contact interaction with the surcharge device, and the polyharmonic representation of excitation. Analytical relationships were obtained for describing the contact and non-contact phases of the surcharge device motion, estimating the contact duration, the velocity restitution coefficient, and the conditions for the existence of a steady periodic regime. The practical significance of the obtained results lies in the possibility of using the proposed model to select the parameters of the impulse surcharge device, assess the stability of operating modes, and improve vibration installations for forming concrete products.*

*Keywords: concrete mixture, vibration compaction, impulse surcharge device, higher harmonics, vibration exciter, contact interaction, dynamic system, oscillation stability.*

**Панфілов Олександр Іванович** - аспірант кафедри галузевого машинобудування та мехатроніки, Національний університет «Полтавська політехніка ім. Юрія Кондратюка»

ORCID: <https://orcid.org/0009-0007-2917-0328>

**Нестеренко Микола Миколайович** - кандидат технічних наук, доцент, доцент кафедри галузевого машинобудування та мехатроніки, Національний університет «Полтавська політехніка імені Юрія Кондратюка».

ORCID: <https://orcid.org/0000-0002-4073-1233>

**Panfilov Olexand** – Postgraduate student of the Department of industrial mechanical engineering and mechatronics National University «Yuri Kondratyuk Poltava Polytechnic».

ORCID: <https://orcid.org/0009-0007-2917-0328>

**Nesterenko Mykola Mykolayovych** - PhD (Tech), Associate Professor, Associate Professor of the Department of industrial mechanical engineering and mechatronics National University «Yuri Kondratyuk Poltava Polytechnic».

ORCID: <https://orcid.org/0000-0002-4073-1233>

Л.Х. Іванова, В.Ю. Селівьорстов, О.Ю. Хитько, О.П. Білий

## СИСТЕМНИЙ ПІДХІД ДО ВСТАНОВЛЕННЯ ВПЛИВУ РІДКІСНОЗЕМЕЛЬНИХ МЕТАЛІВ НА СТРУКТУРУ ВАЛКОВИХ ЧАВУНІВ

*Анотація.* У статті розглядається вплив рідкісноземельних металів на мікроструктуру валкових чавунів. Наведено аналіз сучасного стану вальцеливарного виробництва та питання у обраному напрямку дослідження. Протягом багатьох років у вальцеливарному виробництві ведуться роботи із заміни магнію іншими модифікаторами. Аналіз наявних досліджень і публікацій охоплює питання, що зосереджені на окремих складниках, тоді як модель модифікування чавунних валків РЗМ потребує подальшого дослідження та узагальнення. Метою статті є встановлення карбідотворюючої здатності різних РЗМ у білих валкових чавунах. Для досягнення цієї мети була поставлена наступна задача – установити вплив хімічних елементів церієвої та ітрієвої підгруп РЗМ на мікроструктуру валкового чавуну. Обґрунтувати вибір РЗМ для обробки білих валкових чавунів. Виявлені закономірності впливу рідкісноземельних елементів церієвої підгрупи – самарію та церію та ітрієвої підгрупи – гадолінію та ітрію при швидкості охолодження, що має місце у валковій ливарній формі на процеси структуроутворення білих чавунів для прокатних валків. Встановлені індивідуальні закономірності впливу досліджених підгруп рідкісноземельних металів на мікроструктуру відрізняються інтервалами їх залишкових вмістів. Модифікування білого чавуну РЗМ забезпечувало зміну типу карбідної евтектики та дисперсності перліту. Встановлені закономірності впливу рідкісноземельних елементів підгруп на процеси структуроутворення білих чавунів для прокатних валків: карбідотворююча здатність РЗМ ітрієвої підгрупи була у 5,2 рази більшою за церієву підгрупу. Це дало можливість обґрунтувати тип основи комплексного модифікатора для досягнення необхідних властивостей матеріалу прокатних валків.

*Ключові слова:* валковий чавун, рідкісноземельний метал, структурний складник, карбідотворююча здатність, модифікування чавуну, мікроструктура чавуну, структуроутворення, евтектичні перетворення

**Вступ.** Чавунні прокатні валки є складними виливками, в яких поверхневий робочий шар представлений білим (або половинчастим з кулястим або пластинчастим графітом), а серцевина - сірим чавуном з графітом таких же форм [1]. Якість литих чавунних валків, їх службові властивості залежать від багатьох факторів, але, головним чином, від фізико-хімічних властивостей застосовуваних розплавів і процесів, що протікають у період затвердіння та подальшого охолодження в ливарній формі. Засто-

совувані на прокатних станах нині валки з модифікованого чавуну різноманітні за розмірами та призначенням, виготовляються з різною структурою та механічними властивостями [2]. Питанням технології лиття валків із такого чавуну, дослідженням структури, механічних та теплофізичних властивостей цього валкового матеріалу присвячено роботи багатьох українських дослідників [3-4].

Автором [5] було досліджено п'ять (церій, ітрій, лантан, неодим і празеодим) з шістнадцяти рідкісноземельних металів церієвої групи та наведено експериментально-теоретичні основи отримання ітрієвого чавуну для виробництва деталей відповідального призначення в галузях народного господарства. Що ж до способів введення модифікаторів, були вивчені різні способи їх введення. Найкращим вважається примусове занурення.

Великий інтерес представляють комплексні модифікатори, застосування яких розширює можливості отримання чавунів із широким діапазоном структур та властивостей.

В Україні у сімдесяті роки були розроблені склади комплексних модифікаторів на базі рідкісноземельних металів та ітрію [6]. Промислове випробування таких комплексних модифікаторів на чавунах електропічних плавок показало перспективність їх застосування. Отже, питання щодо складу комплексних модифікаторів досі залишається дискусійним. Так, наприклад, в Україні отримано велику кількість патентів та авторських свідоцтв на склади комплексних модифікаторів для отримання високоміцного чавуну з кулястим графітом. Часто в них наводяться відомості, викликані, очевидно, різними умовами виробництва. У ряді модифікаторів обґрунтовується необхідність високих вмістів кальцію, в інших – навпаки, низьких. Як основний модифікувальний елемент вводяться магній, церій, ітрій, лантан, рідкісноземельні метали, фториди ітрію, фториди рідкісноземельних металів, кальцій та інші. У роботі [7] пропонується діаграма вибору модифікаторів для чавунів і показано, що ефект сфероїдизації визначається створенням відповідними модифікаторами бар'єру в рідкій фазі на межі кристалів графіту в процесі їх зростання. Найбільш сильними модифікаторами, згідно з діаграмою, є магній, ітрій, церій, лантан і гадоліній.

В даний час кількість комплексних модифікаторів перевищує дві тисячі складів, до яких входять графітизуючі, стабілізуючі, сфероїдизуючі та інші компоненти. Найбільша кількість модифікаторів створена на основі феросиліцію, для посилення модифікуючої дії якої вводять цирконій, барій, стронцій, титан, рідкісноземельні та інші хімічні елементи. В інших випадках використовують безкремнієві модифікатори. Точних даних про ефективність модифікаторів немає, незважаючи на те, що у науково-технічній літературі наведено результати численних експериментальних досліджень. Це, очевидно, тим, що ефективність тих чи інших модифікаторів залежить від різноманітних умов виробництва. Тому вибір модифікаторів роблять експериментальним шляхом. Вальцеливарне виробництво, як складний тип виробництва, що охоплює використання білих, половинчастих і сірих чавунів, потребує експериментальних досліджень з відпрацювання та оптимізації складів комплексних модифікаторів для отримання високоякісних чавунів і в робочому шарі, і в серцевині прокатних валків. Аналіз

тенденцій розвитку цього виду техніки показав, що розвиток його йде за такими основними напрямками: удосконалення хімічного складу ливарних сплавів, способів їх отримання та складів лігатур.

До РЗМ відповідно до промислової класифікації металів [8] відносяться хімічні елементи з номерами 57...71 Періодичної системи Д. І. Менделєєва, що мають однакову кількість електронів на зовнішній оболонці атомів і відповідно схожі фізико-хімічні властивості. До рідкісноземельних також відносять близькі до них перехідні метали третьої групи – хімічні елементи з порядковим номером 39 та 21 – ітрій та скандій.

Рідкісноземельні елементи ділять на дві підгрупи: церієву (La, Ce, Pr, Nd, Sm, Eu) і ітрієву (Y, Gd, Tb, Dy, Ho, Er, Tm, Lu).

Фізичні властивості РЗМ, а також дані про їх кристалічну структуру вперше були узагальнені Гшнейдером і доповідалися на симпозіумі РЗМ в 1959 г. у Чикаго.

Самарій (лат. Samarium) Sm - атомний номер 62, атомна маса 150,36, атомний радіус 1,802А, густина 7,536 г/см<sup>3</sup> температура плавлення 1345К (1072 ± 5°С), температура поліморфного перетворення 1193К9 (1670°С). Самарій не сильно окислюється, проте має дуже високий тиск пари.

Гадоліній (лат. Gadolinium) Gd – атомний номер 64, атомна маса 157,25, атомний радіус 1,802А, густина 7,886 г/см<sup>3</sup> температура плавлення 1585К (1312 ± 15°С), температура кипіння 3103К, гадоліній та самарій мають дуже близькі значення радіусів атомів, але відрізняються кристалічними структурами, через те, що гадоліній має гексагональну структуру типу магнію.

Церій (лат. Cerium) Ce – атомний номер 58, атомна маса 140,12, атомний радіус 1,71А, густина 6,789 г/см<sup>3</sup> температура плавлення 1077К (804 ± 5°С), температура кипіння 3743К (3470°С).

Ітрій (лат. Yttrium) Y – атомний номер 39, атомна маса 88,9059, атомний радіус 1,801А, густина 4,472 г/см<sup>3</sup> температура плавлення 1825К (1552 ± 15°С), температура кипіння 2903К (2630°С).

В Україні модифікований чавун вперше було отримано в інституті машинознавства та сільськогосподарської механіки АН УРСР. Після цього в дослідження цього чавуну включився цілий ряд інститутів ( ЦНДІТмаш, ДМетІ та ін). В Україні як основний модифікатор для отримання прокатних валків застосовується чистий магній або лігатури на його основі. При обробці валкового розплаву магнієм виникають труднощі, пов'язані з інтенсивним випаром, що тягне за собою викиди металу, піроефекти, загазованість приміщень. Тому протягом багатьох років ведуться роботи із заміни магнію іншими модифікаторами.

**Аналіз останніх досліджень.** Останнім часом, поряд з такими сплавами, як фероцерій, мішметал для поліпшення властивостей чавуну, все більше застосування знаходять комплексні модифікатори (КМ) на основі РЗМ. Цей напрямок знаходить все більшого поширення у зв'язку із відносною простотою використання КМ порівняно з обробкою металевим магнієм. Ливарники дійшли єдиної думки, що для зниження концентрації неметалевих включень у вибіленому шарі прокатних валків необхідна при

модифікуванні чавунного розплаву заміна магнію РЗМ. Вплив модифікаторів на процеси кристалізації автори розглядають, передусім, з погляду їхнього впливу на поведінку вуглецю в розплаві. Враховується, крім того, зміна властивостей розплаву за рахунок глибокого розкислення та обезсірчення його РЗМ. По відношенню до заліза всі РЗМ мають нижчі температури плавлення і нижчий поверхневий натяг, тобто є поверхнево-активними елементами-модифікаторами.

Отже, одним із найбільш доступних методів покращення якості чавунів є модифікування. Більшість літературних відомостей щодо впливу РЗМ на властивості високовуглецевих сплавів заліза зосереджена на дослідженні наступних питань: 1) нейтралізації шкідливих елементів, 2) сфероїдизації графіту та утворенні зародків та 3) утворенні карбідів та характері кристалізації чавуну. Однак відомостей щодо впливу всіх РЗМ, у тому числі самарію та гадолінію на кристалізацію білих, половинчастих і сірих валкових чавунів дуже мало. Так, у роботі [9] було визначено, що евтектичне перетворення з утворенням пластинчастої ледебуритної евтектики в доевтектичному валковому чавуні при литті у кокільні форми проходило при присадках гадолінію в межах 1,6...3,2%.

У роботі авторів [10] отримані результати досліджень показали доцільність застосування для модифікування чавуну робочого шару двошарових прокатних валків комплексного модифікатору на основі рідкісноземельних металів та термічної обробки литих заготовок валків, а також їх позитивний вплив на фізико-механічні та службові властивості: матеріал дослідного валка (після комплексної обробки) у порівнянні з серійним на відстані 10-30 мм від поверхні бочки мав підвищені властивості: границя міцності при розтягу  $\sigma_b$  була більшою на 3-7, границя міцності при вигині  $\sigma_u$  - на 2-3, ударна в'язкість  $a_k$  - на 4-15, а термостійкість - на 27%.

У роботі [11] наведені результати дослідження і уточнення технології отримання комплексних модифікаторів вуглетермічним способом з окускуванної сировини та застосування їх у вальцеливарному виробництві. Цим дослідженням передувало дослідження особливостей модифікуючого впливу індивідуальних рідкісноземельних металів церієвої групи (церію, лантану та самарію) на параметри кристалізації і структуроутворення валкових чавунів при різних швидкостях охолодження, що мають місце у реальній валковій формі. Встановлено, що мінімальна необхідна кількість модифікатора, що дозволяла досягти максимального рівня зниження температури початку евтектичної кристалізації залежала від швидкості охолодження. Підвищення швидкості охолодження призводило до зміщення цієї кількості модифікатора у бік менших концентрацій.

У роботі [12] зазначено, що кафедра ливарного виробництва Українського державного університету науки і технологій понад шістдесят років займається дослідженням впливу модифікування рідкісноземельними металами на якість литих чавунних виливків. Одним з важливих питань для теорії та практики виробництва чавуну з кулястим графітом є дослідження впливу величини присадки на параметри кристалізації чавуну. Визначення ролі модифікатора в процесі модифікування валкового чавуну дозволить створити найбільш ефективні модифікатори і розробити раціональну технологію отримання високоміцного чавуну. Проведено дослідження впливу величини присадки трьох

індивідуальних рідкісноземельних металів з вмістом домішок, %: лантану (0,39%), церію (0,41%) та ітрію (0,24%) на параметри кристалізації валкового чавуну хімічного складу, %: вуглець 3,29; кремній 1,81; марганець 0,44; фосфор 0,07; сірка 0,026; залізо решта. Встановлено подібний впливу дослідних рідкісноземельних металів на параметри кристалізації валкового чавуну та критичні їх концентрації, які призводили до найбільших змін параметрів кристалізації та утворення кулястих графітних включень в структурах модифікованих чавунів. За модифікування чавунного розплаву з метою одержання кулястого графіту ітрію необхідно додавати в 1,5-1,6 разів менше, ніж церію або лантану.

У роботах [13, 14] були досліджені питання про вплив РЗМ на карбідоутворювальну здатність валкових чавунів та доведено, що це має велике теоретичне і практичне значення. У роботах виявлені закономірності впливу основних РЗМ в інтервалі швидкості охолодження, що мають місце у валковій ливарній формі на процеси структуроутворення білих валкових чавунів. Встановлено, що вплив досліджених РЗМ на мікроструктуру відрізнявся інтервалами їх залишкових вмістів у валкових чавунах. При постійних умовах плавлення, близьких вихідних і кінцевих вмістах сірки в металах було оцінено коефіцієнти засвоєння РЗМ як відношення залишкового вмісту  $RZM_{зал. до}$  введеного  $RZM_e$ . Визначення коефіцієнта засвоєння РЗМ проводили за усередненим залишковим їх вмістом у нижніх частинах досліджуваних зразків. Дослідили присадки РЗМ від 0,2 мас.% до 2,4 мас.%. Коефіцієнти засвоєння всіх досліджених РЗМ при малих присадках у розплав чавуну були в діапазоні від 10 до 56,5%.

Таким чином, наявні дослідження охоплюють питання, що зосереджені на окремих складниках, тоді як модель модифікування РЗМ потребує подальшого дослідження та узагальнення.

**Метою** статті є встановлення карбідоутворюючої здатності двох різних груп РЗМ у білих валкових чавунах та встановити вплив хімічних елементів церієвої та ітрієвої підгруп РЗМ на мікроструктуру валкового чавуну. Обґрунтувати вибір РЗМ для обробки білих чавунів.

**Методика дослідження.** Для виконання поставлених у роботі завдань використовували стандартні методи контролю, апробовані методи дослідження та сучасне обладнання для отримання чавунів, проведення хімічного, металографічного аналізу. Мікроскопічне дослідження проводили на оптичних метало-мікроскопах МІМ-8М та Neophot 21 при збільшеннях 12...1300. Шліфі досліджували до та після травлення. Як основний травник використовували 2% -ний розчин азотної кислоти в етиловому спирті (ніталь). Кількість структурних складників визначали точковим методом Глаголева у 20 полях зору зі збільшенням 400, що забезпечувало достатню достовірність. Мікротвердість структурних складових чавунів визначали на приладі ПМТ-3 за ГОСТ 9450-76 при навантаженні 0,49Н та збільшенні 487. Точність вимірювання діагоналі відбитка становила 0,07 мкм. Величину мікротвердості розраховували як середню 50 вимірів. Застосування науково обґрунтованих методів обробки експериментальних даних при ви-

користанні розроблених методик обробки чавунних розплавів забезпечило отримання результатів із досить високим ступенем точності та достовірності.

**Особливості обробки чавунів рідкісноземельними металами.** Для розробки оптимальних режимів модифікування та типів КМ на основі РЗМ необхідне знання особливостей впливу індивідуальних РЗМ на структуру та властивості та дані про їх засвоєння за різних умов плавки та обробки чавунів. Для модифікування чавунних розплавів використовували індивідуальні РЗМ із вмістом домішок %: церій – 0,41, ітрій – 0,24, гадоліній – 0,29, самарій – 0,46.

Оскільки до складів комплексних модифікаторів входять церій, лантан, ітрій, а також гадоліній, самарій та ряд інших лантанодів, а в літературі, як показав аналітичний огляд, практично відсутні відомості про їх вплив на структуру валкових чавунів, провели дослідження впливу цих хімічних елементів. Для цього було проведено серію плавок на чавуні постійного хімічного складу, мас. %: С 3,49; Si 0,30; Mn 0,05; P 0,005; S 0,2; Fe – решта.

При постійних умовах плавки, близьких вихідних і кінцевих вмістах сірки в металі доцільно, як встановлено в роботі [4], оцінювати засвоєння РЗМ спеціальним коефіцієнтом. Засвоєння церію та самарію зменшувалося зі зростанням величини присадки їх у розплав чавуну, а ітрію та гадолінію – збільшувалося. Найкращий коефіцієнт засвоєння в досліджуваному діапазоні присадок у білих чавунах мали ітрій і гадоліній, а найгірший – самарій, що, мабуть, пов'язане з відмінностями температури плавлення цих хімічних елементів. Коефіцієнт засвоєння (середній) РЗМ церієвої підгрупи при модифікуванні чавунів становив 48,2, а РЗМ ітрієвої підгрупи при модифікуванні чавунів – 38,5%. Таким чином, засвоєння РЗМ ітрієвої підгрупи було на 20% меншим, ніж РЗМ церієвої підгрупи.

**Мікроструктура вихідного немодифікованого чавуну.** Зразок з немодифікованого чавуну після кристалізації зі швидкістю охолодження 5,5 град/с мав структуру доєвтектичного БЧ: дендрити аустеніту були оточені ледебуритної евтектикою стільникового типу з мікротвердістю  $H_{50} 7512$  МПа. Евтектоїдне перетворення аустеніту проходило з утворенням трооститу, мікротвердість якого складала 3883 МПа, рідко зустрічалися ділянки пластинчастого перліту. Перетворення ледебуритного аустеніту частіше відбувалося з поділом фаз, цементит нашаровувався на евтектичному, ферит виділявся окремо.

**Мікроструктура чавунів, модифікованих РЗМ церієвої групи.** Відповідно до завдань даної роботи досліджували особливості структури типових валкових чавунів, модифікованих зростаючими присадками індивідуальних РЗМ. Дослідження проводили на доєвтектичних чавунах, виплавлених на постійній шихті. При модифікуванні досліджуваного базового чавуну зростаючими кількостями самарію чавун з його залишковим вмістом 0,094% мав структуру доєвтектичного БЧ, у порівнянні з вихідним чавуном зросла кількість аустеніту до 40,5% (табл. 1), а також знизилася кількість ледебуриту і підвищилася його мікротвердість. При евтектоїдному перетворенні утворювався троостит з більшою мікротвердістю  $H_{50} 3991$  МПа.

При підвищенні концентрації елемента-модифікатора до 0,198% при кристалізації

Вплив РЗМ церієвої підгрупи на кількість карбідної складової у чавуні

Тип РЗМ	Середній вміст РЗМ, %	Кількість карбідних структурних складників, %		
		карбід	карбід + ледебурит	ледебурит
-	-	-	-	76,0
Sm	0,094	-	-	59,5
	0,198	-	43,8	-
	0,423	23,2	-	-
	0,821	<b>26,6</b>	-	-
	1,082	25,9	-	-
Ce	0,113	-	-	30,6
	0,208	-	31,0	-
	0,392	-	30,4	-
	0,651	29,2	-	-
	0,948	<b>27,5</b>	-	-
	1,423	26,9	-	-

більшість рідини перетворювалася на аустеніт, кількість якого досягала 56,2% (див. табл. 1). Решта рідини, що залишилася, зазнавала евтектичного перетворення з утворенням тонкого ледебуриду і ділянок з конгломератною структурою. Після затвердіння при охолодженні від евтектичної до евтектоїдної температури відбувалося виділення вторинного цементиту, який ріс в аустеніті у вигляді пластин. Евтектоїдне перетворення проходило в основному з утворенням тонкодиференційованого перліту (трооститу). Значення мікротвердості карбідної фази і мартенситу в чавунах із залишковим вмістом самарію 0,198% були в середньому на 10% вище порівняно з чавуном, що містить 0,094% модифікатора.

Подальше підвищення концентрації самарію в чавуні (до 0,423%) призводило до різкого зменшення кількості ледебуриду, майже вся евтектична рідина зазнавала розпаду з утворенням конгломератної структури, тобто спостерігалася роздільне зростання дендритів аустеніту та цементитних пластин. Дрібні ледебурні колонії зустрічалися дуже рідко. Після затвердіння вторинний цементит виділявся у вигляді пластин. Крім того, деякі дендритні гілки аустеніту переохолоджувалися до температури мартенситної області. При вмістах самарію 0,821% кристалізація ледебуриду повністю придушувалася, евтектичне перетворення проходило з поділом фаз. Кількість мартенситу у структурі чавуну зростала.

При модифікуванні вихідного чавуну церієм отримані близькі результати щодо зміни структури матриці. У межах досліджених концентрацій (0,113...1,423%) було отримано БЧ, у матрицях яких кількість перліту зростала від 66,4 до 72,5%. У чавуні із вмістом церію 0,113% порівняно з базовим чавуном зросла кількість аустеніту та знизилася кількість ледебуриду (30,6%), а також його мікротвердість до 6992 МПа. При

евтектоїдному перетворенні утворювався перліт з нижчою мікротвердістю  $H_{50}$  3790 МПа. При підвищенні вмісту церію до 0,208 і 0,392% при кристалізації більшість рідини перетворювалася на аустеніт, кількість якого досягала 68,0% (див. табл.1). Решта рідини, що залишилася, зазнавала евтектичного перетворення з утворенням тонкого ледебуриту і ділянок з конгломератною структурою.

Евтектоїдне перетворення проходило в основному з утворенням тонкодіфференційованого перліту (трооститу) з мікротвердістю 4280 МПа. Слід зазначити, що з концентраціях 0,113...0,392% в ледебуритних колоніях зустрічався аномальний ферит. При вмістах церію 0,651 і 0,948% кристалізація ледебуриту в чавуні повністю придушувалася, евтектичне перетворення проходило з поділом фаз. Мікротвердість карбідів досягала максимальних значень – 9463 та 9571 МПа відповідно. При евтектоїдному перетворенні утворювався троостит з мікротвердістю 4373 МПа.

Таким чином, а) максимальну кількість карбідів (без урахування карбідів у ледебуриті стільникового типу) було отримано при 0,821% самарію, 0,948% церію; б) середній вміст РЗМ цієї підгрупи становив 0,885%; в) середній вміст карбідної складової – 27,1%.

**Мікроструктура чавунів, модифікованих РЗМ ітрієвої групи.** У первинній структурі чавуну з залишковим вмістом 0,020% гадолінію дендрити аустеніту займали 34,5%, решта 65,5% припадала на ледебурит і цементит, що був продуктом роздільної кристалізації евтектичних фаз (табл.2).

Таблиця 2

Вплив РЗМ ітрієвої підгрупи на кількість карбідної складової у чавун

Тип модифікатора	Середній вміст РЗМ %	Кількість карбідної структурної складової, %		
		карбід	карбід + ледебурит	ледебурит
-	-	-	-	76,0
Gd	0,020	-	65,5	-
	0,190	13,9	-	-
	0,900	<b>17,9</b>	-	-
	1,446	16,7	-	-
Y	0,023	-	-	69,5
	0,079	-	30,1	-
	0,250	-	55,8	-
	0,731	<b>54,8</b>	-	-
	1,423	46,9	-	-

Існують критичні вмісти індивідуальних РЗМ, які забезпечують максимальну кількість карбідів у структурі чавуну. Середній залишковий вміст елементів ітрієвої підгрупи в чавуні було на 8% менше, ніж церієвої. Максимальна кількість карбідів (без урахування карбідів у ледебуриті стільникового типу) була отримана при 0,900% га-

долінію та 0,731% ітрію. Середній вміст РЗМ цієї підгрупи становив 0,816%. Середній вміст карбідної складової 36,4%, що на 34% більше, ніж церієвої підгрупи.

Подальше підвищення вмісту індивідуальних РЗМ сприяло зменшенню кількості надлишкових карбідів у структурі чавунів. У чавунах із залишковим вмістом гадолінію 0,190; 0,900 і 1,446% при евтектичній кристалізації відбувався повний поділ фаз, і первинна структура складалася з дендритів аустеніту та цементиту. Вторинний цементит при подальшому охолодженні після затвердіння виділявся з аустеніту як тонких пластин. Евтектоїдне перетворення аустеніту при малому вмісті гадолінію призводило до утворення тонкого перліту з мікротвердістю 3883 МПа. Зі збільшенням вмісту гадолінію в чавуні мікротвердість перліту зростала, у структурі утворювався троостит, мікротвердість якого досягала 5120 МПа.

Вплив модифікування ітрієм було аналогічним впливу самарію: воно виражалось в зменшенні кількості ледебуриту, кристалізація якого повністю пригнічувалася при 0,731% ітрію, збільшенні мікротвердості цементиту і матриці (трооститу), в перетворенні частини аустеніту в мартенсит, частка якого зі збільшенням залишкового змісту ітрію зростала.

Отже, для модифікування білого чавуну існують критичні змісти індивідуальних РЗМ, що забезпечували максимальну кількість карбідів у його структурі. Так, максимальна кількість карбідів (без урахування карбідів у ледебуріті стільникового типу) була отримана при 0,314% самарію та 0,900% гадолінію, а також при 0,731% ітрію, 0,651% церію (див. табл.1 і 2). Подальше підвищення вмісту індивідуальних РЗМ сприяло зменшенню кількості надлишкових карбідів у структурах чавунів.

Як впливає з даних кількісного металографічного аналізу (див. табл. 1 та 2), при постійній швидкості охолодження кількість та тип структурних складників чавуну залежать від виду та залишкового вмісту індивідуального РЗМ.

**Визначення карбідоутворюючої здатності РЗМ.** Металографічний аналіз модифікованих індивідуальними РЗМ чавунів та розрахунок середніх відносних величин кількості карбідної складової на одиницю залишкового вмісту РЗМ показали, що за карбідоутворювальною здатністю досліджувані РЗМ ітрієвої групи мали в середньому показник  $845+747=796$  (рис.1), а досліджувані РЗМ церієвої групи –  $193 +114 = 154$  (рис.2).

Таким чином, в інтервалі досліджених концентрацій РЗМ та швидкостей охолодження вплив РЗМ на структурні зміни валкового БЧ був аналогічним. Модифікування білого чавуну РЗМ забезпечувало зміну типу карбідної евтектики та дисперсності перліту. Найбільша зміна структури модифікованого БЧ спостерігалася при залишковому вмісті самарію 0,423% або гадолінію 0,190%, а також ітрію 0,731%, і церію 0,651%. Підвищення дисперсності та мікротвердості перліту, поява в структурах мартенситу та пластинкова будова карбідної евтектики повинні сприяти підвищенню зносостійкості та міцності БЧ.

Таким чином, незважаючи на те, що середній залишковий вміст елементів ітрієвої підгрупи в чавуні було на 8% меншим, карбідоутворююча здатність цієї підгрупи була

у 5.2 рази більшою за церієву підгрупу. Зазначені індивідуальні особливості РЗМ необхідно враховувати при виборі високоефективних складів КМ, так заміщення частини церію (20...50%) на гадоліній в КМ при модифікуванні, наприклад, білих валкових чавунів, дає ефект в 1,5...3,7 рази більший, що свідчить про особливу цінність його як складової КМ.

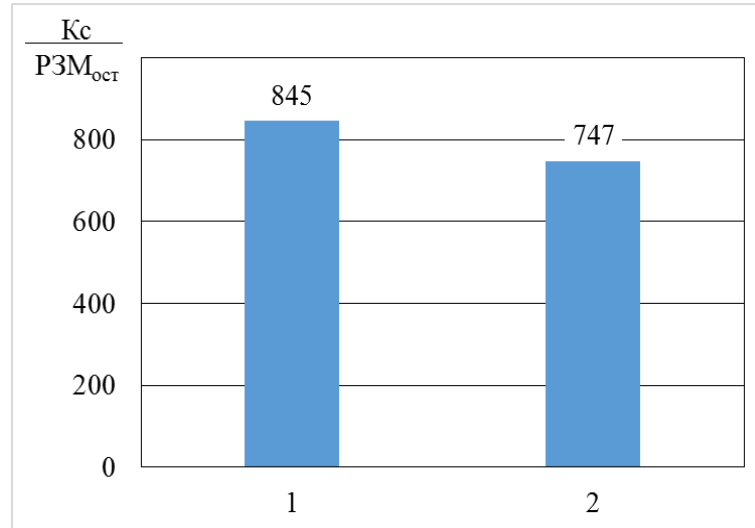


Рисунок 1 - Карбідоутворююча здатність РЗМ ітрієвої підгрупи:  
1 – гадоліній ; 2 – ітрій

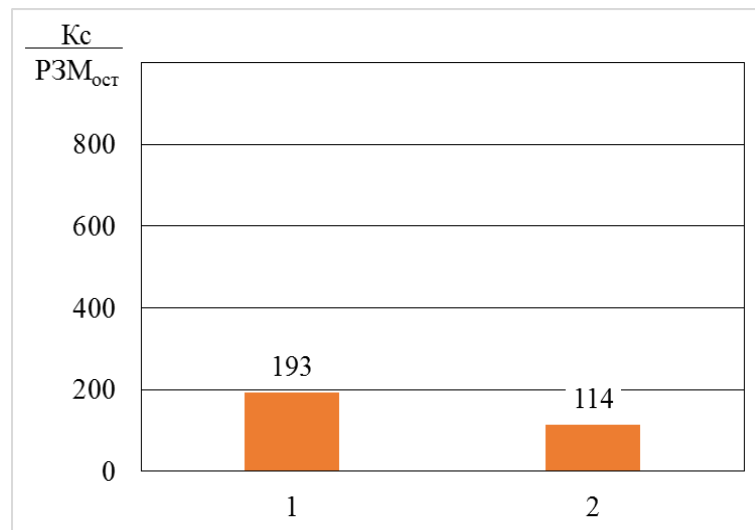


Рисунок 2 - Карбідоутворююча здатність РЗМ церієвої підгрупи:  
1 - самарій ; 2 – церій

**Висновки.** Виявлені закономірності впливу рідкісноземельних елементів церієвої підгрупи – самарію та церію та ітрієвої підгрупи – гадолінію та ітрій при швидкості охолодження, що має місце у валковій ливарній формі на процеси структуроутворення білих чавунів для прокатних валків. Встановлено, що в інтервалі досліджених концентрацій РЗМ вплив РЗМ на структурні зміни валкового БЧ був аналогічним. Модифікування білих чавунів РЗМ забезпечувало зміну типу карбідної евтектики та дисперсності перліту.

Встановлені закономірності впливу рідкісноземельних елементів двох підгруп на процеси структуроутворення білих чавунів для прокатних валків: карбідоутворююча здатність РЗМ ітрієвої підгрупи була у 5,2 рази більшою у порівнянні з церієвою підгрупою. Це дало можливість обґрунтувати тип основи комплексного модифікатора для досягнення необхідних властивостей матеріалу прокатних валків.

#### ЛІТЕРАТУРА

1. Кривошесєв А.Є. Литі валки (теоретичні и технологічні основи виробництва).– М.: Металургвидат, 1957.– 360 с.
2. ТУ У 28.9–00187375–106:2018. Валки чавунні та сталеві для гарячого прокатування металів. Технічні умови. Чинний з 05.03.2018.– Дніпро, 2018. – 36 с.
3. Іванова Л. Х. Теоретичні основи та практичні методи одержання литих прокатних валків із комплексномодифікованих чавунів: Автореф. дис. ... д-ра техн. наук : 05.16.04 / Націон. металург. академія України. – Дніпропетровськ, 2008. – 35 с.
4. Колотило Є.В. Дослідження та удосконалення процесів виробництва листопрокатних валков із модифікованих чавунів: Дис... кандидата техн. наук: 05.16.04; Захищена 14.11.1977; Затв. 26.04.1978. – Дніпропетровськ, 1977.–207 с.
5. Анікін А.А. Експериментально-теоретичні основи одержання ітрієвого чавуні для виробництва деталей відповідального призначення в галузях народного гооподарства: Автореф. дис. ... доктора техн. наук: 05.16.04/ Моск. ін-т стали і спл.–М., 1989.– 47 с.
6. Рябчиков І.В. Матеріали для позапічної обробки залізовуглецевих сплавів // Металургія машинобудування. – 2005.– № 2.– С. 9–17.
7. Комплексні сплави з рідкоземельними металами для модифікування чавуну / І.С.Кумиш, Н.П.Лякишев, А.А.Федотов та ін. // Теорія та практика одержання і застосування комплексних феросплавів:– Тбілісі, 1974.– С. 184–190.
8. Уткін Н.І. Виробництво кольорових металів. –М.: ”Интермет Инжиниринг”, 2000.– 442 с.
9. Іванова Л.Х., Ніколаєнко О.А. Легування білого чавуну гадолінієм/ Нові матеріали і технології в машинобудуванні-2017: матеріали ІХ Міжнарод.наук.-техн. конф., 30-31 трав. 2017 р. м.Київ/ Загальна редакція Р.В.Лютій, І.М.Гурія.- К.: НТУУ «КПІ», 2017.- С.79.
10. Іванова Л.Х., Калашнікова А.Ю., Терехін В.О. Розробка технології виготовлення двошарового прокатного валка / Литво. Металургія. 2019: матеріали XV Міжнар. наук.-практ. конф.(21-23 травня 2019 р., м.Запоріжжя)/ під заг. ред. О.І.Пономаренко.- Запоріжжя: ФОП Наталія Олександрівна Систерова, 2019.- С.102-104.
11. Іванова Л.Х., Колотило Е.В., Хитько О.Ю., Терехієн В.О. Отримання і застосування комплексних модифікаторів з рідкісноземельними металами / Метал і лиття України, т.28, 2020, №1(320), С.30-33.
12. Іванова Л.Х., Колотило Є.В., Мирошніченко Г.О. Рідкісноземельні метали у валкових чавунах / Інформаційні технології в металургії та машинобудуванні– ІТММ’2022: тези доповідей Міжнародної наук.-практич. конф., (Дніпро, 18 трав.. 2022 р.) – Дніпро: УДУНТ, 2022.- С.30-34.

13. ІВАНОВА Л.Х., КОЛОТИЛО Є.В., БІЛИЙ А.П., ШЕМЕТ КАРБИДОУТВОРЮВАЛЬНА ЗДАТНІСТЬ РІДКІСНОЗЕМЕЛЬНИХ МЕТАЛІВ У ВАЛКОВИХ ЧАВУНАХ/ РОЗВИТОК ПРОМИСЛОВОСТІ ТА СУСПІЛЬСТВА: МАТЕРІАЛИ МІЖНАР. НАУК.-ТЕХН. КОНФ.(28 -30 ТРАВНЯ 2025 Р., М.КРИВИЙ РІГ).- КРИВИЙ РІГ, КНУ, 2025.С. 74.

14. ІВАНОВА Л.Х., КОЛОТИЛО Є.В. ДОСЛІДЖЕННЯ КРИСТАЛІЗАЦІЇ ПОЛОВИНЧАСТИХ ЧАВУНІВ / ЛИТВО. МЕТЕАЛУРГІЯ. 2025: МАТЕРІАЛИ ХХІ МІЖНАР. НАУК.-ПРАКТ. КОНФ. (27-29 ТРАВНЯ 2025 Р., М.ХАРКІВ-М.КИЇВ)/ ПІД ЗАГ. РЕД. Д.Т.Н., ПРОФ. ПОНОМАРЕНКО О.І.- ХАРКІВ, НТУ «ХП», 2025.- С. 131-133.

#### REFERENSIS

1. Krivosheev, A.E. (1957). Liteye valki (teoreticheskie i tekhnologicheskie osnovy proizvodstva). M: Metallurgizdat, 360 p.
2. TU U 28.9–00187375–106:2018. (2018). Valky chavunni ta stalevi dlia hariachoho prokatuvannya metaliv. Tekhnichni umovy. Dnipro, 36 p.
3. Ivanova L. Kh. Teoretychni osnovy ta praktychni metody oderzhannya lytykh prokatnykh valkiv iz kompleksnomodyfikovanykh chavuniv: Avtoref. dys. ... d-ra tekhn. nauk : 05.16.04 / Natsion. metalurh. akademiya Ukrainy. – Dnipropetrovs'k, 2008. – 35 s. [in Ukrainian].
4. Kolotylo, E.V. (1977). Doslidzhennia ta udoskonalennia protsesiv vyrobnytstva lystoprokatnykh valkiv iz modyfikovanykh chavuniv. PhD thesis. Dnipropetrovsk, 207 p.
5. Anikin, A.A. (1989). Eksperimentalno-teoreticheskie osnovy polucheniya ittrievogo chuguna dlya proizvodstva detaley otvetstvennogo naznacheniya v otraslyakh narodnogo khozyaystva. Abstract of DSc thesis. M Institute of Steel and Alloys, 47 p.
6. Ryabchikov, I.V. (2005). Materialy dlya pozaprechnoy obrabotki zhelezouglerodistykh splavov. Metallurgiya mashinostroeniya, No. 2, pp. 9–17.
7. Kumysh, I.S., Lyakishev, N.P., Fedotov, A.A., et al. (1974). Kompleksnye splavy s redkozemelnyimi metallami dlya modifitsirovaniya chuguna. In: Teoriya i praktika polucheniya i primeneniya kompleksnykh ferrosplavov. Tbilisi, pp. 184–190.
8. Utkin, N.I. (2000). Proizvodstvo tsvetnykh metallov. M: Intermet Inzhiniring, 442 p.
9. Ivanova, L.Kh., Nikolaenko, O.A. (2017). Leguvannya biloho chavunu hadoliniem. In: Novi materialy i tekhnolohii v mashynobuduvanni: materialy IX Mizhnarodnoi naukovo-tekhnichnoi konferentsii. Kyiv: NTUU «KPI», p. 79.
10. Ivanova, L.Kh., Kalashnikova, A.Yu., Terekhin, V.O. (2019). Rozrobka tekhnolohii vyhotovlennia dvosharovoho prokatnoho valka. In: Lytvo. Metalurhiia 2019: materialy XV Mizhnarodnoi naukovo-praktychnoi konferentsii. Zaporizhzhia, pp. 102–104.
11. Ivanova, L.Kh., Kolotylo, E.V., Khytko, O.Yu., Terekhin, V.O. (2020). Otrymannia i zastosuvannya kompleksnykh modyfikatoriv z ridkisnozemelnyimi metalamy. Metal i lytvo Ukrainy, Vol. 28, No. 1(320), pp. 30–33.
12. Ivanova, L.Kh., Kolotylo, E.V., Miroschnychenko, H.O. (2022). Ridkisnozemelni metaly u valkovykh chavunakh. In: Informatsiini tekhnolohii v metalurhii ta mashynobuduvanni (ITMM'2022). Dnipro: UDUNT, pp. 30–34.
13. Ivanova, L.Kh., Kolotylo, E.V., Bilyi, A.P., Shemet, O. (2025). Karbidooutvoriuvalna zdattnist ridkisnozemelnykh metaliv u valkovykh chavunakh. In: Rozvytok promyslovosti ta

suspilstva: materialy Mizhnarodnoi naukovo-tekhnichnoi konferentsii. Kryvyi Rih: KNU, p. 74.

14. Ivanova, L.Kh., Kolotylo, E.V. (2025). Doslidzhennia krystalizatsii polovynchastykh chavuniv. In: Lytvo. Metalurhiia 2025: materialy XXI Mizhnarodnoi naukovo-praktychnoi konferentsii. Kharkiv–Kyiv: NTU «KhPI», pp. 131–133.

Received 17.04.2026.

Accepted 22.04.2026.

Published 30.04.2026

***System approach to determining the effect of rare-earth metals  
on the structure of roll cast irons***

*The article examines the effect of rare-earth metals on the microstructure of roll cast irons. An analysis of the current state of roll casting production and the issues within the selected research direction is presented. For many years, efforts have been made in roll casting to replace magnesium with alternative modifiers. The analysis of available studies and publications mainly focuses on individual components, while a comprehensive model of modification of roll cast irons using rare-earth metals still requires further investigation and generalization.*

*The aim of the study is to determine the carbide-forming ability of various rare-earth metals in white roll cast irons. To achieve this goal, the following task was formulated: to establish the influence of chemical elements of the cerium and yttrium subgroups of rare-earth metals on the microstructure of roll cast iron, as well as to substantiate the selection of rare-earth metals for the treatment of white roll cast irons.*

*Regularities of the influence of rare-earth elements of the cerium subgroup (samarium and cerium) and the yttrium subgroup (gadolinium and yttrium) on the structure formation processes of white cast irons for rolling rolls under cooling rates typical for roll casting molds have been established. It has been shown that the individual effects of the studied subgroups of rare-earth metals on the microstructure differ in terms of their residual content ranges.*

*Modification of white cast iron with rare-earth metals leads to changes in the type of carbide eutectic and the dispersion of pearlite. It has been established that the carbide-forming ability of rare-earth metals of the yttrium subgroup is 5.2 times higher than that of the cerium subgroup. This made it possible to substantiate the composition of the base of a complex modifier to achieve the required properties of roll materials.*

*Keywords: roll cast iron, rare-earth metal, structural constituent, carbide-forming ability, cast iron modification, cast iron microstructure, structure formation, eutectic transformations*

**Іванова Людмила Харитонівна** – доктор технічних наук, професор, професор кафедри ливарного виробництва, Український державний університет науки і технологій.

ORCID:<https://orcid.org/0000-0002-5997-610X>

**Селівьорстов Вадим Юрійович** – доктор технічних наук, професор, професор кафедри ливарного виробництва, Український державний університет науки і технологій.

ORCID:<https://orcid.org/0000-0002-1916-625X>

**Хитько Олександр Юрійович** – кандидат технічних наук, доцент, доцент кафедри ливарного виробництва, Український державний університет науки і технологій.

ORCID:<https://orcid.org/0000-0002-5997-610X>

**Білий Олександр Петрович** – кандидат технічних наук, доцент, доцент кафедри ливарного виробництва, Український державний університет науки і технологій.

ORCID:<https://orcid.org/0000-0002-5997-610X>

**Ivanova Liudmyla Kharytonivna** – Doctor of Technical Sciences, Professor, Professor of the Department of Foundry Engineering, Ukrainian State University of Science and Technologies.

ORCID: <https://orcid.org/0000-0002-5997-610X>

**Selivorstov Vadym Yuriiiovych** – Doctor of Technical Sciences, Professor, Professor of the Department of Foundry Engineering, Ukrainian State University of Science and Technologies.

ORCID: <https://orcid.org/0000-0002-1916-625X>

**Khytko Oleksandr Yuriiiovych** – PhD (Technical Sciences), Associate Professor, Associate Professor of the Department of Foundry Engineering, Ukrainian State University of Science and Technologies.

ORCID: <https://orcid.org/0000-0002-5997-610X>

**Bilyi Oleksandr Petrovych** – PhD (Technical Sciences), Associate Professor, Associate Professor of the Department of Foundry Engineering, Ukrainian State University of Science and Technologies.

ORCID: <https://orcid.org/0000-0003-1234-5404>

В.Ю. Царик, Вік.В. Гнатушенко, Є.В. Ольшанский

## ПІДХІД РОЗПІЗНАВАННЯ ЕЛЕМЕНТІВ GUI ЯК ЗОБРАЖЕНЬ НА ОСНОВІ МАШИННОГО НАВЧАННЯ

*Анотація.* У сучасних дослідженнях у галузі комп'ютерного зору значна увага приділяється задачам виявлення та класифікації об'єктів на зображеннях, зокрема із застосуванням глибокого навчання. Водночас питання застосування цих підходів до розпізнавання елементів графічного інтерфейсу залишається недостатньо дослідженим. У статті розглянуто підхід, що передбачає використання елементів веб-інтерфейсу як вхідних даних для моделей детекції об'єктів як зображень. Запропоновано процес, у межах якого здійснюється отримання зображення за допомогою Selenium, подальша обробка з використанням моделі машинного навчання та передача результатів у форматі JSON для тестування в середовищі PyTest. Проведено порівняльний аналіз моделей Swin-L, EfficientDet-D7, Cascade Mask R-CNN, RetinaNet та YOLO з точки зору їх придатності до задачі розпізнавання елементів інтерфейсу. За результатами дослідження встановлено, що використання моделей YOLO є найбільш доцільним для задачі розпізнавання елементів графічного інтерфейсу користувача. Запропонований підхід дозволяє зменшити залежність від DOM-структури, підвищити стійкість тестових сценаріїв до змін інтерфейсу та знизити витрати на їх підтримку. Отримані результати можуть бути використані для подальшого розвитку систем автоматизованого тестування веб-застосунків.

*Ключові слова:* розпізнавання, зображення, GUI, модель, глибоке навчання, тестування.

**Постановка проблеми.** Упродовж останніх десятиліть розпізнавання зображень сформувалося як один із ключових напрямів досліджень у галузях цифрової обробки даних, комп'ютерного зору та штучного інтелекту. Методи автоматичного аналізу візуальних даних широко застосовуються в задачах виявлення та класифікації об'єктів, системах відеоспостереження, автономному керуванні транспортними засобами, а також у медичній діагностиці на основі зображень [1]. Інтенсивний розвиток обчислювальних ресурсів і поява глибоких нейронних мереж сприяли суттєвому підвищенню точності та швидкодії алгоритмів розпізнавання. Водночас, попри значний прогрес, проблема надійного функціонування таких алгоритмів у складних і варіативних реальних умовах залишається відкритою.

Особливої актуальності набуває задача розпізнавання елементів графічного інтерфейсу користувача (Graphical User Interface, GUI) як зображень [2]. Сучасні про-

грамні системи, зокрема веб-застосунки, характеризуються високим рівнем складності та динамічності. Архітектура типового веб-застосунку включає кілька взаємопов'язаних рівнів: мову розмітки (HTML або XML), каскадні таблиці стилів (CSS), що визначають візуальне представлення елементів, а також мову програмування JavaScript, яка забезпечує логіку взаємодії та динамічну поведінку інтерфейсу. Така багатокомпонентність зумовлює значну варіативність відображення інтерфейсів навіть за незначних змін у кодовій базі.

Будь-які модифікації структури або стилю веб-сторінки можуть призводити до порушення коректності функціонування інтерфейсу, що, у свою чергу, впливає на досвід користувача та може спричинити втрату клієнтів і зниження прибутків. У цьому контексті автоматизоване тестування користувацького інтерфейсу (UI-тестування) відіграє критично важливу роль у забезпеченні якості програмного забезпечення.

Традиційні підходи до UI-тестування базуються на аналізі DOM-структури документа та використанні селекторів для ідентифікації елементів інтерфейсу. Найбільш поширеними інструментами автоматизації UI-тестування є Selenium WebDriver у поєднанні з тестовими фреймворками, такими як PyTest. Ці засоби дозволяють автоматизувати типові сценарії взаємодії користувача з веб-застосунком, включаючи навігацію між сторінками, введення даних у форми, натискання кнопок та перевірку очікуваних результатів. Однак такі підходи мають низку обмежень, пов'язаних із їхньою залежністю від внутрішньої структури сторінки.

Зокрема, використання XPath- або CSS-селекторів є чутливим до змін у DOM-дереві, це зумовлює необхідність постійного супроводу та оновлення тестових сценаріїв, що підвищує витрати часу і ресурсів. Крім того, традиційні методи не враховують фактичного візуального представлення інтерфейсу, що обмежує їхню здатність виявляти помилки, пов'язані з відображенням елементів.

У зв'язку з цим виникає потреба у розробці альтернативних або доповнювальних підходів до автоматизації UI-тестування, які б базувалися на аналізі візуальної інформації. Одним із перспективних напрямів є застосування методів комп'ютерного зору для розпізнавання елементів інтерфейсу безпосередньо на основі зображень екрану. Такий підхід дозволяє абстрагуватися від внутрішньої реалізації веб-сторінки та орієнтуватися на її кінцеве візуальне представлення, що є більш релевантним з точки зору користувача.

**Аналіз останніх досліджень і публікацій.** В інтерактивному програмному забезпеченні існують інтерфейси користувача (UI), який використовується для взаємодії із системою. Найпоширенішою формою UI є графічний інтерфейс користувача (GUI) через його візуальну природу, яка дозволяє безпосереднє маніпулювання програмним забезпеченням. У багатьох програмних засобах GUI відіграє важливу роль для спрощення його використання, використовуючи візуальний дизайн та когнітивні аспекти людини, такі як правильне використання кольорів або наскільки людині комфортно працювати з візуальною структурою. Розробка графічного інтерфейсу користувача, яка не враховує ці два аспекти, може призвести до людських помилок [3].

Одним із важливих факторів, що визначають, чи може користувач легко використовувати графічний інтерфейс є зручність використання. Зручність використання – це атрибут якості, який вимірює, наскільки легко вивчити, наскільки ефективно використовувати або наскільки приємний інтерфейс користувача [4, 5]. UCD (User-Centered Design) це ітеративний процес дизайну, де інтерфейс користувача не повинен розроблятися раз і назавжди, а поетапно та ітеративно. Що стосується графічного інтерфейсу користувача, ітеративний дизайн інтерфейсу користувача може призвести до кращої якості навіть у процесі редизайну [6]. На практиці візуальний аспект графічного інтерфейсу розвивається поетапно [7-9].

Щоб вирішити ці проблеми, попередні роботи, що використовували комп'ютерний зір або штучний інтелект загалом, виконували дії, спрямовані на пришвидшення або сприяння процесу розробки графічного інтерфейсу.

Одна з перших робіт, що використовувала техніку комп'ютерного зору, була виконана Рідлем та Амантом [10] та Гібссом та ін. [11], де вони створили системи під назвою SegMap та Lens відповідно. Обидві системи були призначені для оцінки, хоча й вирішували різні проблеми: SegMap спрямована на автоматичне дослідження інтерфейсу користувача, тоді як Lens більше орієнтована на розуміння інтерфейсу користувача (структура та класифікація) та мала на меті допомогти незрячим користувачам у використанні графічного інтерфейсу. Чанг та ін. [12] створили систему, яка допомагає тестувальникам автоматизувати процес тестування графічного інтерфейсу користувача (етап оцінки). Система під назвою Sikuli Test дозволила тестувальникам писати тестові сценарії на основі зображень компонентів графічного інтерфейсу або навіть генерувати візуальні тестові сценарії після одноразового запуску тесту вручну.

Кох та Оуласвірта [13] виконали роботу з оцінки сприйняття макета інтерфейсу користувача на основі алгоритмічного представлення принципів гештальту для оцінки зручності використання, де інтерфейс користувача оцінювався в цілому. У роботах Лю та ін. [14] побудована система, яка генерує семантичну інформацію про інтерфейси користувача зі скріншотів мобільних додатків. Зі зображень визначається ієрархія переглядів, потім вона сегментується за кольорами і, нарешті, семантично анотується на основі класифікації компонентів графічного інтерфейсу. Система використовувала набір даних Riso [15] та зосереджувалася на розробці методів.

Система, що використовує глибоке навчання, розроблена Фернадесом та Дежею [16], представлено роботу з оцінки зручності використання для автоматичної евристичної оцінки веб-сайтів за допомогою згорткової нейронної мережі (CNN). Вони створили набір даних на основі евристичної оцінки, наданої учасниками для різних скріншотів веб-сайтів.

Лу та ін. [17] збирали набір даних графічного інтерфейсу програмного забезпечення та класифікували їх на позитивні або негативні категорії на основі макета сторінки, зручності та яскравості. Інше використання глибокої згорткової мережі виконали Хассан та ін. [18], створивши систему виявлення компонентів графічного інтерфейсу з зображення макета інтерфейсу задля покращення розробки графічного інтер-

фейсу. Нгуєн та ін. [19] запропонували систему глибокого навчання DeepUI, яка використовує рекурентні нейронні мережі (RNN) для вивчення шаблону проектування інтерфейсу та генеративно-змагальну мережу (GAN) для створення візуального дизайну з каркасу.

Найпоширенішим способом автоматичної оцінки зручності використання є метод комп'ютерного зору, як це було зроблено в багатьох згаданих дослідженнях. Одним із поширених методів є виявлення компонентів графічного інтерфейсу всередині інтерфейсу користувача, їх класифікація та подальша обробка. Розпізнавання GUI-елементів як зображень передбачає використання алгоритмів детекції об'єктів, сегментації та класифікації. Сучасні методи, засновані на згорткових нейронних мережах (CNN), демонструють високу ефективність у задачах виявлення об'єктів різної природи. На відміну від природних зображень, GUI-елементи часто мають невеликі розміри, високу щільність розміщення та значну стилістичну варіативність. Крім того, один і той самий функціональний елемент (наприклад, кнопка) може мати різні візуальні представлення залежно від дизайну застосунку. Іншою суттєвою проблемою є відсутність стандартизованих і достатньо великих наборів даних для навчання моделей розпізнавання GUI-елементів. На відміну від загальновідомих датасетів у галузі комп'ютерного зору (таких як ImageNet або COCO), у сфері аналізу інтерфейсів кількість відкритих і добре анотованих наборів даних є обмеженою. Це ускладнює порівняння різних методів і знижує відтворюваність результатів досліджень.

Також слід враховувати динамічний характер сучасних інтерфейсів. Елементи можуть змінювати свій стан у відповідь на дії користувача (наприклад, при наведенні курсора або натисканні), що впливає на їхній зовнішній вигляд. Анімації, адаптивний дизайн та використання різних роздільних здатностей екранів додатково ускладнюють задачу розпізнавання. У таких умовах алгоритми повинні бути стійкими до змін масштабу, освітлення, кольорових схем та інших факторів.

Окрему увагу слід приділити питанню інтеграції методів комп'ютерного зору в існуючі процеси тестування. Використання таких методів як доповнення до традиційних інструментів, зокрема Selenium WebDriver, потребує розробки ефективних механізмів взаємодії між різними рівнями системи. Зокрема, необхідно забезпечити коректне зіставлення результатів візуального розпізнавання з логічними діями тестових сценаріїв.

**Мета дослідження.** Наукова проблема полягає у розробці та аналізі алгоритмів розпізнавання елементів графічного інтерфейсу як зображень, які б забезпечували високу точність, стійкість до варіацій та ефективність у реальних умовах використання. Необхідно дослідити, яким чином сучасні підходи машинного навчання можуть бути адаптовані до специфіки GUI та інтегровані в процеси автоматизованого тестування.

Метою даного дослідження є аналіз моделей машинного навчання для розпізнавання та класифікації елементів інтерфейсу користувача за для оптимізації процесів тестування веб-додатків.

**Викладення основного матеріалу дослідження.** Впровадження машинного навчання у сферу тестування веб-додатків приносить значні переваги, допомагаючи авто-

матизувати та покращити процеси тестування. Воно активно впроваджується у різні процеси тестування:

- тестування за допомогою скріншотів - виявлення візуальних відмінностей між сторінками після зміни коду;
- аналіз логів - обробка великих обсягів даних логів для виявлення аномальних подій;
- виявлення елементів – використання AI/ML для ідентифікації певних елементів інтерфейсу користувача на HTML-сторінці замість використання селекторів;
- розробка тестових сценаріїв - використання природної мови або записаних дій користувача для автоматичного створення тестів для програми.

Особливо актуальним є застосування машинного навчання для UI тестування, а саме перевірка, наскільки правильно система відповідає на запити клієнта (користувача).

У більшості випадків перевіряються такі елементи веб-сторінки: кнопки, текст, зображення, поля для введення тексту, чек-бокси. Для кожного елемента повинен бути підготовлений свій тест-план, так як перевірка кожного елемента при UI-тестуванні включає ряд завдань, спрямованих на забезпечення правильної роботи і коректного відображення об'єкта, що тестується, на інтерфейсі програмного додатка або веб-сайту.

Якщо користувач при певних діях очікуючи певний результат отримує помилку або результат, якого він очікував, то тестування було проведено не якісно. Відсутність якісного тестування пов'язана з багатьма факторами: безперервність потоку оновлень у роботі сайту та запуску нових релізів, зміна інтерфейсу користувача; брак часу на створення тестів з повним покриттям, виникнення помилок. Безперервний потік доробок змушує тестувальника безперервно переписувати існуючі автотести, не приділяючи належної уваги створенню нових тест-сценаріїв. Також автотестувальники повинні мати навички розробника, щоб вміти розбиратися не тільки в DOM сайту, а й у різних фреймворках (Рисунок 1).

Саме тому багато компаній відмовляються від запровадження автотестів, залишаючись на ручному тестуванні. Основними проблемами автоматизації UI-тестування для компаній є:

- тривалий час на прогін тестів;
- проблема з інтеграцією різних засобів автоматизації;
- проблема з доступністю середовища;
- проблема із тестовими даними;
- вартість автоматизації та окупність інвестицій;
- проблема стабільності автотестів.

Впровадження комп'ютерного зору автоматизацію UI-тестування має допомогти зняти більшість проблем, описаних вище.

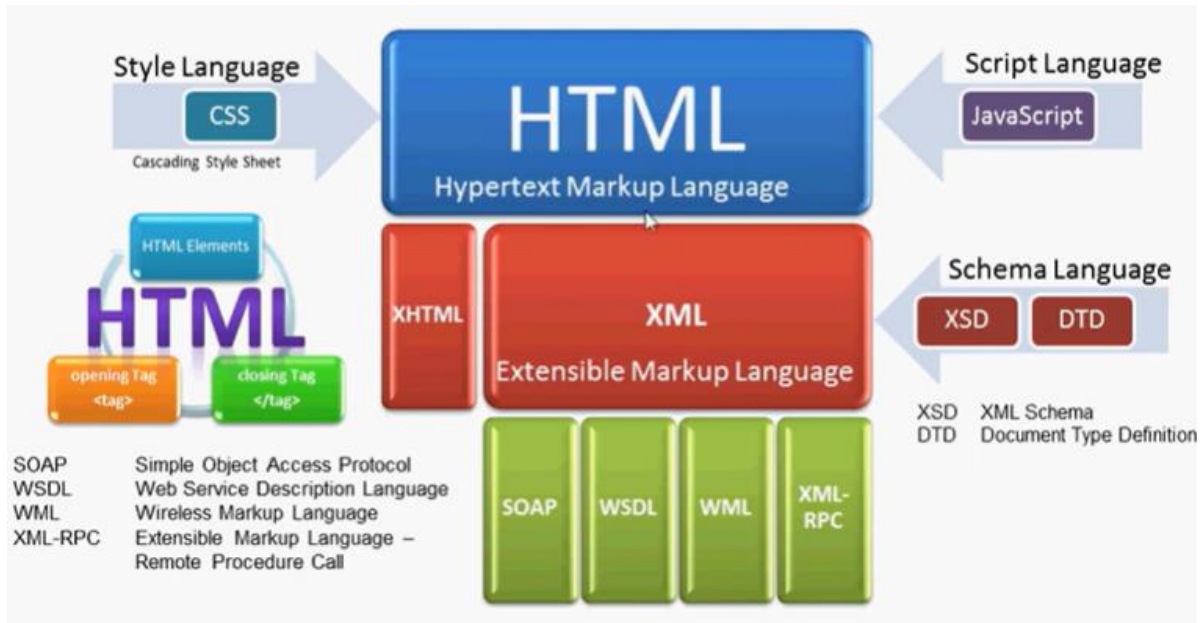


Рисунок 1 – Мови розмітки та програмування веб-сайту [<https://www.agilytics.in/post/a-comprehensive-guide-to-markup-languages-and-their-applications>]

**Процес UI-тестування із застосуванням моделей машинного навчання.** Для запропонування оптимальної моделі процесу автоматичного UI-тестування визначено задачу, яка вирішується з використанням алгоритмів машинного навчання і яке місце в бізнес-процесі вона займає. Для цього в роботі проведено порівняльний аналіз поточного бізнес-процесу, що застосовується при автоматизації UI-тестування (AS-IS) та запропонованого бізнес-процесу (TO-BE) (Рисунок 2-3 відповідно).

AS-IS - описує поточний бізнес-процес застосовується в більшості компаній при автоматизації UI-тестування.

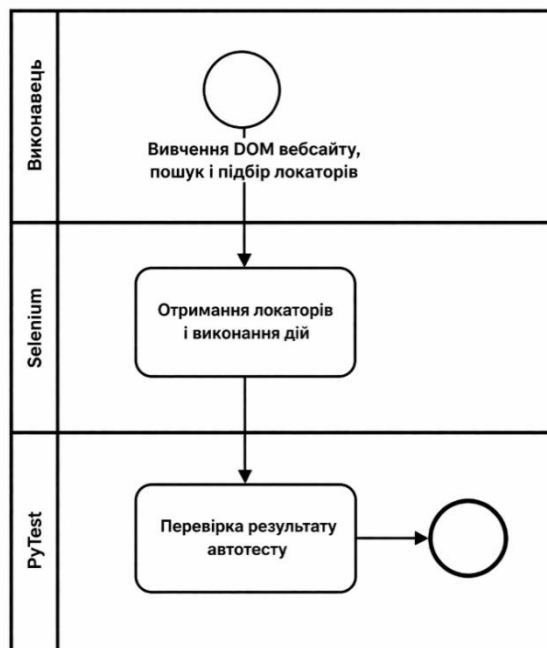


Рисунок 2 - Процес автоматизації UI-тестування AS-IS

У випадку процесу автоматизації UI-тестування за моделлю AS-IS успіх автотесту сильно залежить від уміння тестувальника працювати з DOM і знаходити унікальні елементи, наприклад, якщо в автотесті буде застосовуватися не унікальний елемент, то тест буде виконано з помилками або негативний. Таким чином, автоматизація тестування, що застосовується, повністю не виключає людський фактор. У разі внесення навіть невеликих змін до веб-сторінки, автотести доведеться аналізувати і за необхідності переписувати.

Алгоритм TO-BE визначає порядок дій при повному автоматичному тестуванні, що дозволить навіть при повній зміні інтерфейсу не змінювати автотести і повністю виключити людський фактор.

Для автотестів будуть використовуватись такі інструменти:

- Selenium - інструмент для автоматизації дій із браузером;
- Pytest - інструмент для виконання та перевірки тестових сценаріїв;
- ML - навчена модель машинного навчання.



Рисунок 3 – Процес автоматизації UI-тестування TO-BE

У загальному випадку бізнес-процес TO-BE працює в такий спосіб:

1. За допомогою Selenium створюється скріншот веб-сайту, що аналізується, як цифрове зображення.
2. Отриманий зображення зберігається у сховище (базі даних), потім передається в навчену модель машинного навчання.
4. Модель визначає та детектує класи на зображенні, координати всіх розпізнаних об'єктів та повертає результат у вигляді json формату із зазначенням центру координат об'єкта – x та y і ширину та висоту розпізнаного об'єкта.

5. Отримані детектори об'єкта передаються в Selenium та за допомогою Pytest перевіряється коректність роботи кнопки.:

Отримані переваги після переходу на запропонований бізнес-процес:

1. виключення роботи з DOM сайту, що унеможливило помилки, пов'язані з людським фактором;
2. зниження витрат, пов'язаних з переглядом автотестів у разі зміни структури сайту.

Для успішного використання запропонованого бізнес-процесу необхідна модель машинного навчання, яка могла б розпізнавати та детектувати основні елементи веб-сторінок.

**Моделі машинного навчання для візуального розпізнавання елементів інтерфейсу користувачів.** Для розпізнавання елементів інтерфейсу можна використовувати різні алгоритми. На сьогоднішній день найбільш популярними є Swin-L, EfficientDet-D7, Cascade Mask R-CNN, RetinaNet, YOLO.

Архітектура Swin Transformer Large (Swin-L) [20] належить до класу візуальних трансформерів і базується на ієрархічному представленні ознак із використанням механізму локальної самоуваги (self-attention) у межах зсунутих вікон (shifted windows). Такий підхід забезпечує ефективне моделювання як локальних, так і глобальних залежностей у зображенні, що є критично важливим для аналізу складних графічних інтерфейсів, де просторові відношення між елементами можуть мати семантичне значення. У контексті задачі розпізнавання GUI-елементів Swin-L доцільно розглядати як потужний backbone для систем детекції. Його ієрархічна структура дозволяє ефективно обробляти різномасштабні об'єкти, від дрібних іконок до великих контейнерів. Це особливо важливо для веб-інтерфейсів, де щільність розміщення елементів є високою, а варіативність стилів значною. Крім того, використання attention-механізму дозволяє моделі враховувати контекст розташування елементів, що може підвищити точність класифікації (наприклад, розрізнення кнопки та текстового поля на основі оточення). Це є суттєвою перевагою порівняно з класичними згортковими мережами. Водночас, Swin-L характеризується значною обчислювальною складністю та високими вимогами до пам'яті, що обмежує його застосування в системах реального часу, зокрема в автоматизованому UI-тестуванні. Таким чином, його використання є доцільним переважно в офлайн-аналізі або як частина гібридних архітектур.

EfficientDet-D7 [21] є представником сімейства EfficientDet, яке базується на концепції compound scaling, що передбачає узгоджене масштабування глибини, ширини та роздільної здатності мережі. Центральним компонентом є BiFPN (Bidirectional Feature Pyramid Network), який забезпечує ефективне об'єднання багаторівневих ознак. У задачі розпізнавання GUI-елементів EfficientDet-D7 демонструє високу ефективність завдяки здатності працювати з об'єктами різного масштабу. Це є цінним, оскільки елементи інтерфейсу можуть значно відрізнятися за розмірами. BiFPN дозволяє покращити якість детекції дрібних об'єктів, таких як іконки або чекбокси. Додатковою перевагою є відносно краща ефективність використання обчислювальних ресурсів порівняно з іншими високоточними моделями. Це робить EfficientDet-D7 придатним для практичних

застосувань, де необхідно забезпечити баланс між точністю та швидкістю. Однак, найбільша конфігурація (D7) все ще потребує значних ресурсів, що може ускладнювати інтеграцію в процеси. Крім того, модель менш ефективно враховує глобальний контекст порівняно з трансформерними архітектурами.

Cascade Mask R-CNN [22] є розвитком архітектури Mask R-CNN, що використовує каскадну стратегію для поетапного покращення якості детекції. Кожен наступний етап мережі навчається на більш жорстких порогах IoU, що дозволяє підвищити точність локалізації об'єктів. Ця модель є корисною завдяки можливості сегментації масок, що дозволяє не лише визначити координати об'єкта, але й точно відтворити його форму, що є необхідною умовою для складних інтерфейсів із нестандартними елементами. Каскадна структура забезпечує високу точність навіть у випадках перекриття об'єктів або складного фону. Це актуально для сучасних веб-інтерфейсів, де елементи можуть накладатися або мати складні стилістичні ефекти. Основними обмеженнями є висока обчислювальна складність і низька швидкість, що унеможлиблює використання в реальному часі, та модель є складною в реалізації та налаштуванні.

RetinaNet [23] є моделлю детекції, яка використовує функцію втрат focal loss для подолання проблеми дисбалансу класів. Архітектура включає backbone (наприклад, ResNet) і дві підмережі для класифікації та регресії bounding boxes. Вона є ефективною завдяки здатності працювати з нерівномірно представленими класами. Це важливо, оскільки деякі елементи інтерфейсу (наприклад, кнопки) можуть зустрічатися значно частіше, ніж інші. Модель забезпечує хороший баланс між точністю та швидкістю, що робить її придатною для інтеграції в автоматизовані системи тестування. Водночас відсутність механізмів сегментації обмежує її можливості у випадках, коли необхідна точна форма об'єкта.

YOLO [24] є сімейством одностадійних моделей детекції об'єктів, орієнтованих на обробку в реальному часі. Основна ідея полягає у формулюванні задачі детекції як задачі регресії, що дозволяє виконувати обробку зображення за один прохід мережі. У контексті UI-тестування YOLO є найбільш практичним вибором завдяки високій швидкодії та простоті інтеграції. Модель дозволяє швидко визначати координати елементів інтерфейсу та передавати їх у Selenium для подальшої взаємодії. Попри дещо нижчу точність порівняно з двостадійними моделями, сучасні версії YOLO (наприклад, YOLOv5/YOLOv8) демонструють достатній рівень якості для практичних застосувань. Це робить їх придатними для автоматизованого тестування, де швидкість є критичним фактором.

З проведеного аналізу зроблено висновок, що використання YOLO забезпечує суттєве зменшення обчислювальної складності та затримки обробки, що є важливим у контексті автоматизованого тестування інтерфейсів. У запропонованому бізнес-процесі ТО-ВЕ, де передбачається ітеративне виконання тестових сценаріїв із взаємодією користувача, швидкість отримання координат елементів інтерфейсу безпосередньо впливає на загальну продуктивність системи. Це є основною перевагою порівняно з іншими моделями, такими як Swin Transformer або Cascade Mask R-CNN, які, незважаючи на вищу

точність, характеризуються значно більшими обчислювальними витратами. Також важливим аспектом є інтеграційна сумісність, YOLO формує результати у вигляді набору обмежувальних рамок (bounding boxes) із відповідними координатами та класами об'єктів. Такий формат природно узгоджується з потребами автоматизованого тестування, де необхідно передавати координати елементів у Selenium для подальшої взаємодії (наприклад, імітації натискання кнопок). У запропонованому алгоритмі результати роботи моделі представлені у форматі JSON, що містить координати центру об'єкта, а також його ширину і висоту. YOLO безпосередньо підтримує подібне представлення, що мінімізує необхідність додаткової обробки даних. Модель може бути донавчена (fine-tuned) на специфічних наборах даних, що містять елементи інтерфейсу, характерні для конкретного веб-застосунку або домену. Це дозволяє підвищити точність розпізнавання без необхідності створення надмірно великих і складних моделей. Умови задачі, де інтерфейси часто мають повторювані патерни, сприяють ефективному перенавчанню моделі.

У даній предметній області критичним є не ідеальне відтворення меж об'єкта, а коректне визначення області, достатньої для взаємодії (наприклад, кліку по кнопці), тому у цьому контексті YOLO забезпечує оптимальний баланс між точністю та ефективністю.

**Висновки.** У межах дослідження підходів до розпізнавання елементів графічного інтерфейсу користувача (GUI) як цифрових зображень особливу увагу приділено вибору моделі машинного навчання, здатної забезпечити ефективну інтеграцію в процес автоматизованого UI-тестування. З огляду на специфіку предметної області, ключовими вимогами до моделі є точність локалізації об'єктів, стійкість до варіацій візуального представлення інтерфейсів, можливість практичної інтеграції з існуючими інструментами, зокрема Selenium та PyTest. Проведений аналіз сучасних архітектур моделей дозволяє стверджувати, що моделі сімейства YOLO (You Only Look Once) найбільш повно відповідають зазначеним вимогам.

Використання YOLO сприяє усуненню залежності від DOM-структури веб-сторінки. Традиційні підходи до UI-тестування базуються на використанні селекторів, які є чутливими до змін у коді. Натомість запропонований підхід тестування вебзастосунків TO-BE дозволяє працювати безпосередньо з візуальним представленням інтерфейсу. YOLO, як ефективний детектор об'єктів, виступає ключовим компонентом цього підходу, забезпечуючи надійне визначення позицій елементів незалежно від їхньої внутрішньої реалізації.

Разом з тим, слід враховувати і певні обмеження YOLO, яка може демонструвати знижену точність при детекції дуже дрібних або щільно розташованих елементів, що характерно для деяких інтерфейсів. Однак ці обмеження можуть бути частково компенсовані за рахунок підвищення роздільної здатності вхідних зображень або використання більш сучасних версій моделі.

## ЛІТЕРАТУРА

1. Paschalis Tsirtsakis, Georgios Zacharis, George S. Maraslidis, George F. Fragulis, Deep learning for object recognition: A comprehensive review of models and algorithms, International Journal of Computer Science and Information Technology, ISSN 1562-9945 (Print) ISSN 2707-7977 (Online)

tional Journal of Cognitive Computing in Engineering, Volume 6, 2025, Pages 298-312, ISSN 2666-3074, <https://doi.org/10.1016/j.ijcse.2025.01.004>.

2. Вік.В.Гнатушенко, В.Ю. Царик. Дослідження методів виділення графічних об'єктів на вебсайтах для оцінки якості інтерфейсу / Вік.В.Гнатушенко, В.Ю. Царик // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 3 (140). – Дніпро, 2022. – С.143-154 DOI 10.34185/1562-9945-3-140-2022-12

3. J. Shariat and C. S. Saucier, *Tragic Design: The Impact of Bad Product Design and How to Fix It, First*. Sebastopol: O'Reilly Media, 2017.

4. D. Norman and J. Nielsen, “Usability 101: Introduction to Usability,” 2012. [Online]. Available: <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>.

5. D. A. Norman, *The Design of Everyday Things*. USA: Basic Books, Inc., 2002. J.

6. Nielsen, “Iterative User-Interface Design,” *Computer* (Long Beach, Calif.), vol. 26, no. 11, doi.org/10.1109/2.241424. 1993, doi.org/10.1109/2.241424.

7. J. Cao, K. Zieba, and M. Ellis, *The Ultimate Guide to Prototyping*. Mountain View: UXPin Studio, 2015.

8. S. Minhas, “User Experience Design Process,” 2018. [Online]. Available: <https://uxplanet.org/user-experiencedesign-process-d91df1a45916>.

9. C. Murphy, “A Comprehensive Guide To Wireframing And Prototyping,” 2018. [Online]. Available: <https://www.smashingmagazine.com/2018/03/guide-wireframing-prototyping/#top>.

10. M. O. Riedl and R. St Amant, “Toward Automated Exploration of Interactive Systems,” 2002.

11. K. Gibbs, T. Winograd, and N. Scott, “Lens: A System for Visual Interpretation of Graphical User Interfaces,” 2002.

12. T.-H. Chang, T. Yeh, and R. C. Miller, “GUI Testing Using Computer Vision,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2022.

13. J. Koch and A. Oulasvirta, “Computational layout perception using Gestalt laws,” in *Conference on Human Factors in Computing Systems - Proceedings*, 2016, vol. 07-12-May-2016, pp. 1423–1429, 10.1145/2851581.2892537. doi: 10.1145/2851581.2892537

14. T. F. Liu, M. Craft, J. Situ, E. Yumer, R. Mech, and R. Kumar, “Learning design semantics for mobile apps,” in *UIST 2018 - Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, 2018, pp. 569–579, doi: 10.1145/3242587.3242650.

15. B. Deka et al., “Rico: A mobile app dataset for building data-driven design applications,” in *UIST 2017 - Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, 2017, pp. 845–854, 10.1145/3126594.3126651.

doi: 10.1145/3126594.3126651

16. R. A. Fernandez, J. A. Deja, and B. P. V. Samson, “Automating heuristic evaluation of websites using convolutional neural networks,” in *Conference on Human Factors in Computing Systems - Proceedings*, 2018, pp. 9–12, doi: 10.1145/3205851.3205854.

17. H. Lu, L. Wang, M. Ye, K. Yan, and Q. Jin, “DNN-based Image Classification for Software GUI Testing,” in *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data*

Computing, Internet of People and Smart City Innovation (Smart-World/SCALCOM/UIC/ATC/CBDC om/IOP/SCI), 2018, pp. 1818–1823.

18. S. Hassan, M. Arya, U. Bhardwaj, and S. Kole, “Extraction and Classification of User Interface Components from an Image,” *Int. J. Pure Appl. Math.*, vol. 118, no. 24, 2018.

19. T. T. Nguyen, P. M. Vu, H. V. Pham, and T. T. Nguyen, “Deep learning UI design patterns of mobile apps,” in *Proceedings - International Conference on Software Engineering*, 2018, pp. 65–68, doi: 10.1145/3183399.3183422.

20. Z. Liu et al., “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,” *Proc. IEEE ICCV*, 2021. DOI: 10.1109/ICCV48922.2021.00986

21. M. Tan, R. Pang, Q. V. Le, “EfficientDet: Scalable and Efficient Object Detection,” *Proc. IEEE CVPR*, 2020. DOI: 10.1109/CVPR42600.2020.01039

22. Z. Cai, N. Vasconcelos, “Cascade R-CNN: High Quality Object Detection and Instance Segmentation,” *IEEE TPAMI*, 2021. DOI: 10.1109/TPAMI.2019.2956516

23. T.-Y. Lin et al., “Focal Loss for Dense Object Detection,” *IEEE TPAMI*, 2020. DOI: 10.1109/TPAMI.2018.2858826

24. J. Redmon et al., “You Only Look Once: Unified, Real-Time Object Detection,” *Proc. IEEE CVPR*, 2016. DOI: 10.1109/CVPR.2016.91

#### REFERENCES

1. Paschalis Tsirtsakis, Georgios Zacharis, George S. Maraslidis, George F. Fragulis, Deep learning for object recognition: A comprehensive review of models and algorithms, *International Journal of Cognitive Computing in Engineering*, Volume 6, 2025, Pages 298-312, ISSN 2666-3074, <https://doi.org/10.1016/j.ijcce.2025.01.004>.

2. Vik.V.Hnatushenko, V.Yu. Tsaryk. Research on methods for highlighting graphic objects on websites to assess interface quality/ Vik.V.Hnatushenko, V.Yu. Tsaryk // *System technologies. Regional interuniversity collection of scientific papers. – Issue 3 (140). – Dnipro, 2022. – S.143-154 DOI 10.34185/1562-9945-3-140-2022-12*

3. J. Shariat and C. S. Saucier, *Tragic Design: The Impact of Bad Product Design and How to Fix It*, First. Sebastopol: O’Reilly Media, 2017.

4. D. Norman and J. Nielsen, “Usability 101: Introduction to Usability,” 2012. [Online]. Available: <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>.

5. D. A. Norman, *The Design of Everyday Things*. USA: Basic Books, Inc., 2002. J.

6. Nielsen, “Iterative User-Interface Design,” *Computer (Long. Beach. Calif.)*, vol. 26, no. 11, doi.org/10.1109/2.241424. 1993, doi.org/10.1109/2.241424.

7. J. Cao, K. Zieba, and M. Ellis, *The Ultimate Guide to Prototyping*. Mountain View: UXPin Studio, 2015.

8. S. Minhas, “User Experience Design Process,” 2018. [Online]. Available: <https://uxplanet.org/user-experiencedesign-process-d91df1a45916>.

9. C. Murphy, “A Comprehensive Guide To Wireframing And Prototyping,” 2018. [Online]. Available: <https://www.smashingmagazine.com/2018/03/guide-wireframing-prototyping/#top>.

10. M. O. Riedl and R. St Amant, “Toward Automated Exploration of Interactive Systems,” 2002.

11. K. Gibbs, T. Winograd, and N. Scott, “Lens: A System for Visual Interpretation of Graphical User Interfaces,” 2002.
12. T.-H. Chang, T. Yeh, and R. C. Miller, “GUI Testing Using Computer Vision,” in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2022.
13. J. Koch and A. Oulasvirta, “Computational layout perception using Gestalt laws,” in Conference on Human Factors in Computing Systems - Proceedings, 2016, vol. 07-12-May-2016, pp. 1423–1429, 10.1145/2851581.2892537. doi: 10.1145/2851581.2892537
14. T. F. Liu, M. Craft, J. Situ, E. Yumer, R. Mech, and R. Kumar, “Learning design semantics for mobile apps,” in UIST 2018 - Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, 2018, pp. 569–579, doi: 10.1145/3242587.3242650.
15. B. Deka et al., “Rico: A mobile app dataset for building data-driven design applications,” in UIST 2017 - Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology, 2017, pp. 845–854, 10.1145/3126594.3126651. doi: 10.1145/3126594.3126651
16. R. A. Fernandez, J. A. Deja, and B. P. V. Samson, “Automating heuristic evaluation of websites using convolutional neural networks,” in Conference on Human Factors in Computing Systems - Proceedings, 2018, pp. 9–12, doi: 10.1145/3205851.3205854.
17. H. Lu, L. Wang, M. Ye, K. Yan, and Q. Jin, “DNN-based Image Classification for Software GUI Testing,” in 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDC om/IOP/SCI), 2018, pp. 1818–1823.
18. S. Hassan, M. Arya, U. Bhardwaj, and S. Kole, “Extraction and Classification of User Interface Components from an Image,” *Int. J. Pure Appl. Math.*, vol. 118, no. 24, 2018.
19. T. T. Nguyen, P. M. Vu, H. V. Pham, and T. T. Nguyen, “Deep learning UI design patterns of mobile apps,” in Proceedings - International Conference on Software Engineering, 2018, pp. 65–68, doi: 10.1145/3183399.3183422.
20. Z. Liu et al., “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,” *Proc. IEEE ICCV*, 2021. DOI: 10.1109/ICCV48922.2021.00986
21. M. Tan, R. Pang, Q. V. Le, “EfficientDet: Scalable and Efficient Object Detection,” *Proc. IEEE CVPR*, 2020. DOI: 10.1109/CVPR42600.2020.01039
22. Z. Cai, N. Vasconcelos, “Cascade R-CNN: High Quality Object Detection and Instance Segmentation,” *IEEE TPAMI*, 2021. DOI: 10.1109/TPAMI.2019.2956516
23. T.-Y. Lin et al., “Focal Loss for Dense Object Detection,” *IEEE TPAMI*, 2020. DOI: 10.1109/TPAMI.2018.2858826
24. J. Redmon et al., “You Only Look Once: Unified, Real-Time Object Detection,” *Proc. IEEE CVPR*, 2016. DOI: 10.1109/CVPR.2016.91

Received 20.04.2026.  
Accepted 24.04.2026.  
Published 30.04.2026

***An approach to recognizing GUI elements as images  
based on machine learning***

*Analysis of Recent Research and Publications.* In recent years, computer vision tasks, in particular, object detection and classification in images, have undergone significant development thanks to the use of deep learning methods. The most common approaches are the YOLO, RetinaNet, EfficientDet, Cascade Mask R-CNN, and Swin Transformer models, which demonstrate high efficiency in object recognition tasks of varying complexity. At the same time, the issue of applying these models to recognize graphical user interface (GUI) elements in the context of automated testing of web applications remains insufficiently covered. Traditional approaches to UI testing are based on DOM structure analysis, which makes them vulnerable to changes in layout and complicates test support.

*Research Objective.* The purpose of this work is to analyze modern methods of recognizing objects in images and justify the choice of an effective machine learning model for the task of recognizing graphical user interface elements.

*Presentation of the main research material.* The paper proposes an approach to recognizing interface elements based on the analysis of web page screenshots as digital images. The proposed process involves obtaining a screenshot using the Selenium tool, saving the image in storage, and then transferring it to a machine learning model for processing. As a result, the model determines the classes of objects, their coordinates, and sizes, generating output data in JSON format. The obtained coordinates are used to interact with interface elements in an automated testing environment (PyTest).

A comparative analysis of modern object detection models is conducted, in particular Swin-L, EfficientDet-D7, Cascade Mask R-CNN, RetinaNet, and YOLO. Particular attention is paid to models of the YOLO family, which provide high image processing speed while maintaining a sufficient level of accuracy. In the context of the interface element recognition problem, this characteristic is important, since it allows integrating the model into the automated testing process without significantly increasing the test execution time. In addition, YOLO provides a convenient format of output data (bounding boxes), which is directly used for interaction with interface elements.

*Conclusions.* As a result of the conducted research, it was found that the use of computer vision methods is a promising direction for increasing the efficiency of automated UI testing. The proposed approach allows you to abandon the dependence on the DOM structure, which increases the resistance of tests to interface changes and reduces the costs of their support. Among the considered models, the most appropriate for practical application is YOLO, which provides the optimal balance between speed, accuracy and ease of integration. The results obtained can be used for the further development of intelligent web application testing systems.

*Keywords:* recognition, image, GUI, model, deep learning, testing.

**Царик Владислав Юрійович** – старший викладач кафедри інформаційних технологій і систем Українського державного університету науки і технологій.

ORCID: <https://orcid.org/0000-0001-6449-8037>

**Гнатушенко Вікторія Володимирівна** – доктор технічних наук, професор, професор кафедри інформаційних технологій і систем Українського державного університету науки і технологій.

ORCID: <https://orcid.org/0000-0001-5304-4144>

**Ольшанський Євген Володимирович** – аспірант кафедри інформаційних технологій і систем Українського державного університету науки і технологій.

ORCID: <https://orcid.org/0009-0008-7100-673X>

**Tsaryk Vladyslav** – senior lecturer at the Department of Information Technologies and Systems of the Ukrainian State University of Science and Technology.

ORCID: <https://orcid.org/0000-0001-6449-8037>

**Hnatushenko Viktoriia** – Doctor of engineering's sciences, Professor, Professor of Department of Information Technologies and Systems, Ukrainian State University of Science and Technology.

ORCID: <https://orcid.org/0000-0001-5304-4144>

**Olshanskyi Evhen** – postgraduate student of the Department of Information Technologies and Systems of the Ukrainian State University of Science and Technology.

ORCID: <https://orcid.org/0009-0008-7100-673X>

С.А. Рубан, Д.В. Швець, І.А. Котов, В.В. Тронь, Н.О. Карабут

## РОЗРОБКА АРХІТЕКТУРИ ІНФОРМАЦІЙНО-КЕРУЮЧОЇ СИСТЕМИ ДЛЯ ЕНЕРГОЕФЕКТИВНОЇ ПЕРЕРОБКИ ЗАЛІЗОРУДНОЇ СИРОВИНИ

*Анотація.* У статті розроблено багаторівневу архітектуру інформаційно-керуючої системи для енергоефективної переробки залізорудної сировини. Запропонована архітектура відповідає референсній моделі RAMI 4.0, вимогам стандартів ISA-88 та ISA-95 щодо ієрархічної організації систем управління, а також вимогам кібербезпеки стандарту ISA/IEC 62443. Розглянуто структуру польового рівня з інтеграцією засобів оперативного контролю якості залізорудної сировини, Edge-рівня для первинної агрегації та нормалізації технологічних даних, а також хмарного рівня для реалізації цифрових двійників технологічного процесу та алгоритмів оптимізації верхнього рівня. Запропонована архітектура спрямована на підвищення енергоефективності виробництва, скорочення питомих витрат на збагачення та покращення якості кінцевого продукту.

*Ключові слова:* інформаційно-керуюча система, залізорудна сировина, RAMI 4.0, ISA-88, ISA-95, ISA/IEC 62443, Edge computing, цифровий двійник, кібербезпека промислових систем, енергоефективність.

**Постановка проблеми.** Переробка залізорудної сировини є одним із найбільш енергоємних переділів металургійного виробництва. Підвищення ефективності цього процесу потребує не лише вдосконалення технологічного обладнання, а й впровадження сучасних інформаційно-керуючих систем (ІКС), здатних забезпечити автоматизований збір, передачу, обробку та аналіз технологічних даних у режимі реального часу.

Цифрова трансформація гірничо-металургійної галузі відбувається в контексті концепції Industry 4.0, ключовим елементом якої є Промисловий Інтернет речей (ІІоТ). Референсна архітектурна модель Індустрії 4.0 (RAMI 4.0, IEC PAS 63088) визначає єдину систему координат для опису компонентів розумного виробництва в трьох вимірах: ієрархії (від польового рівня до хмари), життєвого циклу виробу та функціональних рівнів (активів, інтеграції, комунікації, інформації, функцій і бізнесу) [1, 2].

Специфіка збагачення залізорудної сировини полягає в необхідності безперервного контролю якісних показників матеріального потоку – вмісту заліза, розміру часток, вологості, – що вимагає інтеграції у структуру ІКС спеціалізованих засобів оперативного контролю безпосередньо у технологічному потоці. Такі засоби генерують значні масиви даних, ефективна обробка яких потребує проміжного Edge-шару між польовим і хмарним рівнями.

Метою статті є розробка та обґрунтування багаторівневої архітектури ІКС для енергоефективної переробки залізорудної сировини, яка відповідає сучасним стандартам у сфері промислової автоматизації (ISA-88, ISA-95) та кібербезпеки (ISA/IEC 62443).

Тематика дослідження відповідає пріоритетам європейської промислової політики, зокрема «Новій промисловій стратегії для зеленої та цифрової Європи», та реалізується у рамках проєкту Erasmus+ GreeDI-EU «Transition towards Green Digital European Industry», що об'єднує академічних учасників та представників роботодавців, громадських організацій та органів влади для підготовки фахівців до зеленої цифрової трансформації промисловості.

**Аналіз останніх досліджень і публікацій.** Традиційні архітектури ІКС гірничо-збагачувальних підприємств будувались за ієрархічним принципом відповідно до моделі ISA-95, проте мали ряд обмежень: вертикально-інтегровані та слабо пов'язані компоненти, обмежені можливості обміну даними між рівнями [3], відсутність засобів просунутої аналітики та предиктивного керування.

Сучасні дослідження свідчать про тенденцію до впровадження IoT-платформ, Edge-обчислень та хмарних сервісів у системи управління процесами переробки залізорудної сировини. Разом із тим, комплексний підхід до проєктування таких систем з урахуванням вимог RAMI 4.0 [2], специфіки оперативного контролю якості залізорудної сировини у потоці та вимог кібербезпеки залишається недостатньо дослідженим.

Технологічний процес переробки залізорудної сировини включає такі основні переділи: дроблення та грохочення, подрібнення, класифікацію, магнітну та флотаційну сепарацію, зневоднення (фільтрацію та сушку). Кожен переділ характеризується специфічними технологічними параметрами, що підлягають вимірюванню та регулюванню.

Ключовою особливістю управління процесом є необхідність безперервного контролю якісних показників матеріального потоку. Для цього застосовуються засоби оперативного контролю: рентгено-флуоресцентні (XRF) аналізатори, вологоміри, гранулометри та інші. Дані пристрої мають власні вбудовані мікроконтролери та протоколи зв'язку, що потребує спеціального підходу до їх інтеграції у загальну структуру ІКС.

RAMI 4.0 (Reference Architecture Model Industrie 4.0), розроблена в рамках ініціативи Industrie 4.0 (Германія) та стандартизована як IEC PAS 63088, являє собою тривимірну архітектурну систему координат, яка об'єднує різні аспекти інтелектуального виробництва [2, 4].

Вісь «Ієрархічні рівні» відображає розширену ієрархію автоматизації від одного виробу (Product) через польовий рівень (Field Device), станцію (Control Device), комірку (Field), зону (Station), підприємство (Enterprise) до підключеного світу (Connected World). Вісь «Шари» охоплює шість функціональних шарів: активів (Asset), інтеграції (Integration), комунікації (Communication), інформації (Information), функцій

(Functional) та бізнесу (Business). Вісь «Життєвий цикл» (на основі ІЕС 62890) відображає зв'язок між розробкою та виробництвом активів.

Основним концептом RAMI 4.0 є Компонент Індустрії 4.0 (I4.0 Component), що об'єднує фізичний актив (Asset) та його цифрове представлення – Оболонку адміністрування активу (Asset Administration Shell, AAS). AAS забезпечує стандартизований опис можливостей активу, його властивостей та інтерфейсів, що є необхідною умовою для побудови семантично сумісних та інтероперабельних систем управління [5].

**Виклад основного матеріалу.** Для ІКС переробки залізорудної сировини принципово важливим є відображення у рамках RAMI 4.0 наступних компонентів: технологічного обладнання (млини, сепаратори, насоси), засобів оперативного контролю якості (аналізатори у потоці), Edge-пристроїв (промислові шлюзи), систем управління (DCS/PLC, SCADA) та хмарних сервісів (платформи цифрових двійників, системи оптимізації).

Стандарт ISA-88 (ІЕС 61512) визначає модель та термінологію для батч-управління, проте його концепції процедурних та фізичних моделей широко застосовуються і для безперервних процесів [6-8]. Відповідно до цього стандарту, технологічний процес переробки залізорудної сировини декомпозується на:

- Процес (Process) – повний цикл переробки від прийому руди до відвантаження концентрату.
- Стадії (Stages) – основні переділи: дроблення, подрібнення, сепарація, зневоднення.
- Операції (Operations) – послідовні дії в рамках стадії (наприклад, регулювання завантаження млина).
- Фази (Phases) – елементарні дії управління (наприклад, встановлення уставки регулятора).

Фізична модель за ISA-88 включає підприємство, ділянку (site), зону (area), технологічну комірку (process cell) та одиницю обладнання (unit). Засоби оперативного контролю якості розглядаються як спеціалізовані функціональні модулі (equipment modules) у складі відповідних одиниць обладнання [6].

Стандарт ISA-95 (ІЕС/ІСО 62264) визначає ієрархічну модель управління виробничим підприємством та інтерфейси між рівнями [3]. Для розроблюваної ІКС запропонована наступна відповідність рівнів ієрархії (табл. 1).

Критично важливим елементом рівня 0 є засоби оперативного контролю якості сировини. XRF-аналізатори та гранулометри, вбудовані безпосередньо у конвеєрну лінію або трубопровід, забезпечують вимірювання вмісту заліза (Fe), вологи, а також складу пульпи з частотою 1–5 хвилин. Ці дані є вхідними для алгоритмів управління режимами сепарації та подрібнення.

Для безперервного контролю якісних параметрів залізорудної сировини у потоці застосовуються наступні типи аналітичних пристроїв:

«Системні технології» 3 (164) 2026 «System technologies»

- рентгено-флуоресцентні (XRF) аналізатори на конвеєрі (cross-belt analyzers) – для визначення хімічного складу (Fe, SiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, CaO, MgO) у потоці руди або концентрату;
- вологоміри – для безконтактного вимірювання вологості матеріалу на конвеєрі або у технологічних ємностях;
- гранулометри – для вимірювання розподілу частинок за розмірами;
- щільноміри та витратоміри пульпи – для контролю консистенції та витрати пульпи у гідроциклонах та насосних агрегатах;
- аналізатори вологості та зольності концентрату – на стадії зневоднення та сушки.

Таблиця 1

Відповідність рівнів ISA-95 компонентам ІКС переробки залізорудної сировини

Рівень ISA-95	Функціональне призначення	Системи та компоненти
Рівень 4 (ERP/MES)	Планування виробництва, управління ресурсами підприємства	ERP, MES, BI, Digital Twin Platform
Рівень 3 (MES/SCADA)	Оперативне диспетчерське управління, координація технологічних процесів	SCADA, MES, historian, Edge Gateway
Рівень 2 (Control)	Безперервне управління технологічним процесом	DCS, PLC, HMI
Рівень 1 (Field Control)	Локальне регулювання, захист	Регулятори, реле захисту
Рівень 0 (Field)	Первинне вимірювання та виконавчі механізми	Датчики, аналізатори, приводи

Інтеграція засобів оперативного контролю у загальну структуру ІКС здійснюється через стандартні промислові протоколи: OPC UA (основний протокол для обміну даними між рівнями 0-3 згідно з RAMI 4.0), Modbus TCP/RTU (для застарілих пристроїв), PROFINET/PROFIBUS (у складі SIMATIC-архітектур), а також MQTT (для передачі даних у напрямку Edge і хмари).

Кожен аналітичний пристрій оформлюється як I4.0 Component з відповідною Asset Administration Shell (AAS), що забезпечує стандартизований опис його можливостей, параметрів та інтерфейсів. AAS містить субмоделі: ідентифікації (Identification), технічних даних (Technical Data), документації (Documentation) та даних процесу (Process Data) [2, 9].

Виміряні показники якості передаються на рівень DCS/SCADA для використання у алгоритмах замкнутого та розімкнутого управління, а також на Edge-рівень для первинної обробки та подальшої передачі у хмарну платформу.

Edge-рівень (проміжний шар між польовим рівнем та хмарою) займає ключову позицію в сучасних ІТ-архітектурах [10]. Для ІКС переробки залізорудної сировини Edge-рівень виконує наступні функції:

- агрегація та буферизація даних від польових пристроїв, PLC/DCS та засобів оперативного контролю якості – забезпечення локального зберігання даних у разі втрати зв'язку з хмарою;
- нормалізація технологічних показників – приведення даних різних джерел до єдиної семантичної моделі (на основі OPC UA або AAS), вирішення проблеми різноманітності форматів та протоколів;
- первинна обробка та фільтрація даних – усунення шуму та аномалій, розрахунок похідних показників (питома витрата енергії на тону продукту, ефективність сепарації тощо);
- виконання алгоритмів управління та оптимізації нижнього рівня (edge analytics)
- розрахунок уставок для регуляторів без необхідності звернення до хмари;
- забезпечення кібербезпеки – фільтрація трафіку, аутентифікація пристроїв, контроль цілісності даних перед відправкою у хмару.

Апаратну основу Edge-рівня складають промислові Edge-сервери або IPC (Industrial PC) з підвищеними вимогами до захисту від електромагнітних завад, пилу та вібрації. Типові представники: Siemens SIMATIC IPC427E, Dell EMC Edge Gateway 3200, HPE Edgeline EL300.

Програмний стек Edge-вузла включає: операційну систему реального часу або Linux-дистрибутив з RT-патчем; OPC UA Server/Client для взаємодії з польовими пристроями та хмарою; брокер MQTT (наприклад, Eclipse Mosquitto) для публікації даних; контейнерне середовище (Docker/Kubernetes) для розгортання мікросервісів аналітики; систему локального зберігання даних (time-series DB рівня Edge, наприклад InfluxDB Edge або TimescaleDB) [10].

Зв'язок Edge-рівня з хмарою здійснюється через захищені канали (VPN, TLS 1.3) за протоколами MQTT або AMQP [11, 12], що відповідає архітектурному принципу однопоточної передачі критичних технологічних даних від промислової мережі до хмари.

Технологічні дані процесів переробки залізорудної сировини є за своєю природою часовими рядами (time-series), що обумовлює застосування спеціалізованих баз даних цього типу. Бази даних часових рядів (TSDB) забезпечують: компресію та ефективне зберігання великих обсягів часових рядів, швидке виконання запитів на часові проміжки та агрегацію, підтримку функцій даунсемплінгу та автоматичного видалення застарілих даних.

У якості TSDB для ІКС переробки залізорудної сировини рекомендується застосування: InfluxDB [12] або TimescaleDB (на основі PostgreSQL) для операційного зберігання технологічних даних з горизонтом до 1 року; Apache Parquet/Delta Lake на хмарних об'єктних сховищах (AWS S3, Azure Blob) для довготривалого архівування та машинного навчання. Для доступу до даних з рівня SCADA та MES застосовується OPC Historical Data Access (OPC HDA) або REST API відповідної TSDB.

Цифровий двійник (Digital Twin) технологічного процесу збагачення є ключовим елементом хмарного рівня ІКС. Відповідно до RAMI 4.0, цифровий двійник реалізуєть-

ся як оболонка адміністрування активу (AAS) для технологічної системи в цілому та її окремих компонентів.

Для ІКС переробки залізорудної сировини виділяються наступні типи цифрових двійників:

– Цифровий двійник агрегату (Equipment Digital Twin) – для ключового обладнання (кульових млинів, барабанних сепараторів). Моделює кінетику подрібнення/сепарації, поточний стан зносу, прогнозований залишковий ресурс.

– Цифровий двійник технологічної комірки (Process Cell Digital Twin) – для замкнутого циклу подрібнення або ділянки флотації. Інтегрує моделі окремих агрегатів та моделює матеріальні та енергетичні баланси.

– Цифровий двійник підприємства (Plant Digital Twin) – охоплює весь технологічний ланцюжок від прийому сировини до відвантаження продукту. Застосовується для оперативного планування виробництва та оптимізації енергоспоживання.

Моделі цифрових двійників будуються на основі комбінації фізично обґрунтованих моделей та моделей на основі даних (нейронні мережі, gradient boosting), що дозволяє забезпечити прийнятну точність навіть при неповноті знань стосовно фізичних параметрів процесу.

На хмарному рівні реалізуються алгоритми оптимізації, що оперують на горизонті від декількох годин до доби. Основною задачею є мінімізація питомого енергоспоживання на тону виробленої продукції при дотриманні вимог до якості продукту та продуктивності.

Для вирішення задачі оптимізації застосовуються: Model Predictive Control (MPC) на основі лінеаризованих моделей цифрових двійників технологічних комірок [13]; алгоритми оптимізації з обмеженнями (gradient-based, evolutionary); методи машинного навчання з підкріпленням (Reinforcement Learning) для адаптивного управління в умовах мінливих характеристик сировини. Результати оптимізації транслюються у вигляді оперативних рекомендацій (установок) через рівень SCADA/DCS до регуляторів польового рівня.

Для підвищення практичної цінності запропонованої архітектури доцільно виділити в її складі спеціалізовані аналітико-керуючі модулі для окремих технологічних комірок збагачення. У контексті переробки залізорудної сировини такими комірками можуть бути замкнутий цикл подрібнення–класифікації, вузол магнітної сепарації, а також комбіновані ділянки, у межах яких формується якість концентрату та основна частка питомих енерговитрат. Для кожної з таких комірок доцільно формувати локальний цифровий двійник, орієнтований на прогнозування технологічного стану та підтримку прийняття рішень щодо оптимального режиму роботи.

Інформаційною основою функціонування такого модуля є потоки технологічних даних, що надходять із рівня польових пристроїв, локальних систем керування та засобів оперативного контролю якості. До найбільш інформативних змінних належать витрата живлення, густина та витрата пульпи, крупність продукту класифікації, вміст заліза у вихідному та кінцевому продуктах, споживана потужність основного обладна-

ння, а також сигнали про поточний технічний стан технологічних апаратів. Сукупний аналіз цих даних у середовищі Edge та хмарного рівня дає змогу перейти від реактивного способу керування до прогнозно-орієнтованого, в якому рішення формуються з урахуванням очікуваної динаміки процесу та наявних технологічних обмежень.

Для процесів збагачення залізорудної сировини особливо доцільним є поєднання фізично обґрунтованих моделей із моделями, побудованими на основі даних. Такий гібридний підхід дозволяє враховувати як відомі закономірності подрібнення, класифікації та сепарації, так і складні нелінійні залежності, які важко формалізувати аналітично в умовах змінних характеристик мінералого-технологічних різновидів руди. У межах запропонованої архітектури моделі машинного навчання можуть використовуватись як м'які сенсори для оцінювання важковимірюваних показників, для короткострокового прогнозування вмісту заліза у концентраті, крупності продукту або питомих витрат енергії, а також для адаптації параметрів цифрового двійника до поточного стану технологічного об'єкта.

На основі прогнозної моделі технологічної комірки може бути реалізовано модельно-прогнозне керування, яке забезпечує обчислення оптимальних уставок для локальних контурів регулювання з урахуванням багатокритеріальності процесу керування. На відміну від традиційних підходів, MPC дозволяє одночасно враховувати вимоги до якості концентрату, продуктивності, стабільності технологічного режиму та мінімізації енергоспоживання. При цьому у задачі оптимізації можуть бути враховані обмеження на витрату води, подачу руди, навантаження на млини, допустимі режими роботи сепараторів, а також обмеження, пов'язані з коливаннями гранулометричного складу та вмісту заліза у сировині.

З погляду реалізації в інформаційно-керуючій системі модуль машинного навчання та MPC доцільно розміщувати на Edge-рівні або на верхньому рівні керування залежно від вимог до швидкодії та обчислювальних ресурсів. Edge-рівень є доцільним для задач, що потребують оперативного перерахунку уставок у реальному часі або квазі-реальному часі, тоді як хмарний рівень придатний для більш складних обчислень, перенавчання моделей, багатоваріантної оптимізації та аналізу історичних даних. Такий розподіл функцій забезпечує масштабованість системи, зменшує затримки при прийнятті рішень і водночас зберігає можливість централізованого вдосконалення моделей та алгоритмів керування.

Запропонований підхід дозволяє інтегрувати в єдину архітектуру засоби онлайн-аналітики, цифрові двійники, модулі машинного навчання та алгоритми модельно-прогнозного керування, що є особливо важливим для енергоефективної переробки залізорудної сировини. У результаті формується основа для побудови інтелектуальної ІКС, здатної не лише реєструвати поточний стан процесу, а й прогнозувати його розвиток, виявляти тенденції до погіршення якості продукту чи зростання енерговитрат та своєчасно формувати керуючі впливи для стабілізації й оптимізації технологічного режиму.

Своєю чергою, стандарт ISA/IEC 62443 «Мережева та системна безпека для промислових систем автоматизації та управління» є галузевим стандартом кібербезпеки

для промислових систем управління (IACS) [14, 15]. Він визначає концепцію зон безпеки (Security Zones) та кондуїтів (Conduits), рівні безпеки (Security Levels, SL 1-4), а також вимоги до засобів захисту (Security Countermeasures).

Зони безпеки являють собою логічні або фізичні групи активів системи управління з подібними вимогами до безпеки. Кондуїти – це комунікаційні канали між зонами, які мають бути захищені відповідно до вимог обох підключених зон.

Для розроблюваної ІКС запропоновано розподіл на зони безпеки та відповідні заходи захисту (табл. 2).

Особливу увагу слід приділити захисту кондуїту між промисловою мережею (Зона 2/3) та Edge-рівнем [15]. Рекомендується застосування апаратних однонаправлених шлюзів (data diodes), що фізично унеможливають зворотну передачу даних з мережі Edge до промислової мережі, забезпечуючи захист від кібератак типу «pivot».

Таблиця 2

Зони безпеки та заходи захисту відповідно до ISA/IEC 62443

Зона / Кондуїт	Рівень безпеки (SL)	Ключові заходи захисту
Зона 1 – Польовий рівень	SL 1	Фізичний захист обладнання, базова автентифікація
Зона 2 – Управління процесом	SL 2	Сегментація мережі, контроль доступу на рівні ролей (RBAC)
Зона 3 – Диспетчерський рівень (SCADA)	SL 2–3	DMZ, шифрування трафіку (TLS 1.3), аудит подій
Зона 4 – Edge Gateway	SL 2	Однонаправлені шлюзи (data diode), VPN, аутентифікація пристроїв (X.509)
Зона 5 – Хмарна платформа	SL 3	Zero-trust архітектура, шифрування даних у стані спокою та передачі, SOC/SIEM

Для виявлення аномалій у трафіку промислових мереж та інцидентів кібербезпеки рекомендується розгортання системи промислового виявлення вторгнень (Industrial IDS) на рівні SCADA/DCS, сумісної з такими платформами, як Clarity, Dragos або Nozomi Networks.

**Результати та обговорення.** Розроблена архітектура ІКС для енергоефективної переробки залізорудної сировини являє собою інтегровану багаторівневу систему, що охоплює польовий рівень (рівні 0-1 ISA-95), рівень управління процесом (рівень 2 ISA-95), диспетчерський рівень (рівень 3 ISA-95), Edge-рівень та хмарний рівень (рис. 1).

Відповідність RAMI 4.0 забезпечується: застосуванням OPC UA як єдиного комунікаційного стандарту між рівнями; реалізацією концепту Asset Administration Shell для ключових компонентів системи (аналізаторів якості, основного обладнання, цифрових двійників); орієнтацією на сервісно-орієнтовану архітектуру на хмарному рівні.

Інтеграція засобів оперативного контролю якості залізорудної сировини як I4.0 Components дозволяє забезпечити семантичну сумісність їх даних з рештою системи та використання цих даних у алгоритмах управління та оптимізації на всіх рівнях ієрархії.



Edge-рівень забезпечує зниження затримки реакції на зміни технологічних параметрів, зменшення обсягу трафіку у напрямку хмари (до 90% за рахунок локальної агрегації та фільтрації), а також підвищення відмовостійкості системи при втраті зв'язку з хмарою.

Очікуваний ефект від впровадження розробленої ІКС: зниження питомого енергоспоживання на збагачення на 8–15% за рахунок оптимізації режимів подрібнення та сепарації; підвищення вмісту заліза у концентраті на 0,5–1,5%; скорочення нерентабельних простоїв основного обладнання на 20–30% за рахунок предиктивного технічного обслуговування.

**Висновки.** У статті розроблено та обґрунтовано архітектуру інформаційно-керуючої системи для енергоефективної переробки залізорудної сировини, що відповідає сучасним вимогам у сфері промислової автоматизації та цифрової трансформації. Основні результати роботи:

– запропоновано багаторівневу архітектуру ІКС на основі референсної моделі RAMI 4.0 з ієрархічним розподілом функцій відповідно до ISA-88 та ISA-95.

– розроблено підхід до інтеграції засобів оперативного контролю якості залізорудної сировини (XRF-аналізаторів, гранулометрів, вологомірів) у загальну структуру ІКС із застосуванням концепту Asset Administration Shell.

– обґрунтовано роль та функції Edge-рівня як ключового проміжного шару між польовим рівнем та хмарою, що забезпечує первинну агрегацію, нормалізацію та локальну аналітику технологічних даних.

– визначено архітектуру хмарного рівня з розгортанням баз даних часових рядів (InfluxDB, TimescaleDB), платформ цифрових двійників та алгоритмів оптимізації верхнього рівня (MPC, Reinforcement Learning).

– розроблено систему зон безпеки та кондуїтів відповідно до ISA/IEC 62443 з визначенням цільових рівнів безпеки (SL 1-3) для кожної зони.

Подальші дослідження будуть спрямовані на розробку та верифікацію математичних моделей цифрових двійників технологічних комірок збагачення на реальних виробничих даних, а також на практичну апробацію алгоритмів оптимізації верхнього рівня.

**Funding:** This work was supported by the Erasmus+ Programme of the European Union [GreeDI-EU, Grant Number 101085821].

**Disclaimer:** Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Education and Culture Executive Agency (EACEA). Neither the European Union nor the granting authority can be held responsible for them.

#### ЛІТЕРАТУРА

1. Past, present and future of Industry 4.0 - a systematic literature review and research agenda proposal / Y. Liao та ін. *International Journal of Production Research*. 2017. Т. 55, № 12. С. 3609–3629. DOI: 10.1080/00207543.2017.1308576

2. Reference Models for Digital Manufacturing Platforms / Fraile та ін. *Applied Sciences*. 2019. Т. 9, № 20. С. 4433. DOI: 10.3390/app9204433
3. Пупена О.М., Клименко О.М., Міркевич Р.М. Принципи функціонування систем керування основним виробництвом через призму стандарту ІЕС-62264: Електронний навчальний посібник. 2019. URL: <https://tk185.appau.org.ua/62264/case-study-iec-62264/momguide>
4. Bridging the Gap: Deriving Specific Reference Architectures from RAMI 4.0 for Flexible Production Systems / С. Binder та ін. *IFAC-PapersOnLine*. 2025. Т. 59, № 10. С. 1119–1124. DOI: 10.1016/j.ifacol.2025.09.189
5. Ontologies for Industry 4.0 / V. R. Sampath Kumar та ін. *The Knowledge Engineering Review*. 2019. Т. 34. DOI: 10.1017/s0269888919000109
6. Пупена О. М. Практичні рекомендації до реалізації елементів стандарту ІЕС 61512 в програмному забезпеченні систем керування / О. М. Пупена, Р. М. Міркевич, О. М. Клименко. – К: АППАУ, ТК 185 «Промислова автоматизація», 2020. – 46 с.
7. Міркевич Р., Пупена О. Оперативно-календарне планування як основа ефективної діяльності сучасних молочних підприємств. *Наукові праці НУХТ*. 2018. Т. 24, № 4. С. 29–40. DOI: 10.24263/2225-2924-2018-24-4-5
8. Pupena O. PacFramework: Technological Solutions for PLC/PAC Programming – A Technical Report on Architecture, Principles, and Practical Implementation. *Preprints 2025*, 2025071180. DOI: 10.20944/preprints202507.1180.v1
9. Пупена О., Клименко О., Полупан В. Систематизація ключових стандартів для цифрових двійників у промисловому виробництві. *Наукові вісники КПІ*. 2025. Т. 140, № 3. DOI: 10.20535/kpissn.2025.3.337306
10. Edge Computing: Vision and Challenges / W. Shi та ін. *IEEE Internet of Things Journal*. 2016. Т. 3, № 5. С. 637–646. DOI: 10.1109/jiot.2016.2579198
11. Siqueira F., Davis J. G. Service Computing for Industry 4.0: State of the Art, Challenges, and Research Opportunities. *ACM Computing Surveys*. 2022. Т. 54, № 9. С. 1–38. DOI: 10.1145/3478680
12. Bousdekis A., Mentzas G. Enterprise Integration and Interoperability for Big Data-Driven Processes in the Frame of Industry 4.0. *Frontiers in Big Data*. 2021. Т. 4. DOI: 10.3389/fdata.2021.644651
13. Model predictive control of power plant cycling using Industry 4.0 infrastructure / D. Kesterling та ін. *Digital Chemical Engineering*. 2023. С. 100090. DOI: 10.1016/j.dche.2023.100090
14. Jhanjhi N., Humayun M., N. Almuayqil S. Cyber Security and Privacy Issues in Industrial Internet of Things. *Computer Systems Science and Engineering*. 2021. Т. 37, № 3. С. 361–380. DOI: 10.32604/csse.2021.015206
15. Secure data exchange in Industrial Internet of Things / A. Sukiasyan та ін. *Neurocomputing*. 2022. Т. 484. С. 183–195. DOI: 10.1016/j.neucom.2021.07.101

## REFERENCES

1. Liao, Y., Deschamps, F., Loures, E. d. F. R., & Ramos, L. F. P. (2017). Past, present and future of Industry 4.0 - a systematic literature review and research agenda proposal. *International Journal of Production Research*, 55(12), 3609–3629. DOI: 10.1080/00207543.2017.1308576.
2. Fraile, Sanchis, Poler, & Ortiz. (2019). Reference Models for Digital Manufacturing Platforms. *Applied Sciences*, 9(20), 4433. DOI: 10.3390/app9204433
3. Pupena, O. M., Klymenko, O. M. & Mipkevych, P. M. (2019) Ppyntsypy funktsionuvannya cyctem kepuvannia ocnovnym vypobnytstvom chepez ppyzmu ctandaptu IEC-62264 [Principles of operation of core manufacturing control systems through the lens of the IEC-62264 standard]. [in Ukrainian]
4. Binder, C., Riedmann, S., Neureiter, C., Vollmar, J., Gupta, R., Calà, A., & Lüder, A. (2025). Bridging the Gap: Deriving Specific Reference Architectures from RAMI 4.0 for Flexible Production Systems. *IFAC-PapersOnLine*, 59(10), 1119–1124. DOI: 10.1016/j.ifacol.2025.09.189
5. Sampath Kumar, V. R., Khamis, A., Fiorini, S., Carbonera, J. L., Olivares Alarcos, A., Habib, M., Goncalves, P., Li, H., & Olszewska, J. I. (2019). Ontologies for Industry 4.0. *The Knowledge Engineering Review*, 34. DOI: 10.1017/s0269888919000109
6. Pupena, O. M., Mipkevych, P. M. & Klymenko, O. M. (2020) Ppaktychni pekomentatsii do pealizatsii elementiv ctandaptu IEC 61512 v ppohpamnomu zabezpechenni cyctem kepuvannia [Practical recommendations for the implementation of IEC 61512 standard elements in control system software]. Kyiv: APPAU, TC 185 "Industrial Automation". [in Ukrainian].
7. Mirkevich, R., & Pupena, A. (2018). Planning and scheduling as the basis of effective activity of modern dairy enterprises. *Scientific Works of National University of Food Technologies*, 24(4). DOI: 10.24263/2225-2924-2018-24-4-5 [in Ukrainian].
8. Pupena, O. (2025) PacFramework: Technological Solutions for PLC/PAC Programming – A Technical Report on Architecture, Principles, and Practical Implementation. *Preprints 2025*, 2025071180. DOI: 10.20944/preprints202507.1180.v1
9. Pupena, O., Klymenko, O., & Polupan, V. (2025). Systematyzatsiia kliuchovykh standartiv dlia tsyfrovyykh dviinykiv u promyslovomu vyrobnytstvi [Systematization of key standards for digital twins in industrial manufacturing]. *Naukovi visti KPI*, 3(140). DOI: 10.20535/kpissn.2025.3.337306 [in Ukrainian].
10. Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge Computing: Vision and Challenges. *IEEE Internet of Things Journal*, 3(5). DOI: 10.1109/jiot.2016.2579198
11. Siqueira, F., & Davis, J. G. (2022). Service Computing for Industry 4.0: State of the Art, Challenges, and Research Opportunities. *ACM Computing Surveys*, 54(9). DOI: 10.1145/3478680
12. Bousdekis, A., & Mentzas, G. (2021). Enterprise Integration and Interoperability for Big Data-Driven Processes in the Frame of Industry 4.0. *Frontiers in Big Data*, 4. DOI: 10.3389/fdata.2021.644651

13. Kesterling, D., Agbleze, S., Bispo, H., & Lima, F. V. (2023). Model predictive control of power plant cycling using Industry 4.0 infrastructure. *Digital Chemical Engineering*, 100090. DOI: 10.1016/j.dche.2023.100090
14. Jhanjhi, N., Humayun, M., & N. Almuayqil, S. (2021). Cyber Security and Privacy Issues in Industrial Internet of Things. *Computer Systems Science and Engineering*, 37(3). DOI: 10.32604/csse.2021.015206
15. Sukiasyan, A., Badikyan, H., Pedrosa, T., & Leitao, P. (2022). Secure data exchange in Industrial Internet of Things. *Neurocomputing*, 484. DOI: 10.1016/j.neucom.2021.07.101

Received 24.04.2026.  
Accepted 27.04.2026.  
Published 30.04.2026

### ***Development of an information and control system architecture for energy-efficient iron ore processing***

*Analysis of recent research and publications. The high energy intensity of iron ore processing necessitates the implementation of control systems based on streaming data analysis. Recent studies confirm the effectiveness of integrating IIoT platforms, edge computing, and cloud services. Traditional hierarchical architectures are limited by weak horizontal integration and a lack of predictive control tools. Designing multi-level systems to process data arrays from on-stream quality analyzers while accounting for the RAMI 4.0 reference model and strict cybersecurity requirements requires a comprehensive approach.*

*Purpose of the research. To substantiate a comprehensive multi-level architecture of an information and control system aimed at minimizing the specific energy consumption of beneficiation processes through multi-criteria data processing in accordance with the ISA-88, ISA-95, and ISA/IEC 62443 industry standards.*

*Presentation of the main research material. The proposed architecture is based on the RAMI 4.0 concept. At the field level, on-stream analyzers are integrated via Asset Administration Shells to guarantee semantic interoperability. The implemented Edge level provides primary aggregation, normalization, and local analytics of technological data, critically reducing the load on industrial networks and transmission latency. The cloud level operates with specialized time-series databases and digital twins of process cells. Energy consumption optimization is achieved by combining physics-based models and machine learning methods within a unified Model Predictive Control loop. The system's cybersecurity is implemented through spatial distribution into secure zones using hardware data diodes.*

*Conclusions. The developed architecture constitutes a scientific and practical foundation for the digital transformation of the mineral processing industry. The application of hybrid predictive models and edge analytics lays the technological groundwork for reducing specific electricity consumption by 8–15%, increasing the iron content in the concentrate by 0.5–1.5%, and minimizing unscheduled downtime. The prospect for further research is the verification of the digital twin mathematical models on real industrial datasets.*

*Keywords: information and control system; iron ore raw materials; RAMI 4.0; ISA-88; ISA-95; ISA/IEC 62443; Edge computing; digital twin; industrial cybersecurity; energy efficiency.*

**Рубан Сергій Анатолійович** - кандидат технічних наук, доцент, завідувач кафедри автоматизації, комп'ютерних наук і технологій, Криворізький національний університет.

ORCID: <https://orcid.org/0000-0002-4495-6667>

**Швець Дмитро Валерійович** - кандидат технічних наук, доцент, доцент кафедри моделювання та програмного забезпечення, Криворізький національний університет.

ORCID: <https://orcid.org/0000-0001-5126-6405>

**Котов Ігор Анатолійович** - доктор технічних наук, доцент, професор кафедри моделювання та програмного забезпечення, Криворізький національний університет.

ORCID: <https://orcid.org/0000-0003-2445-6259>

**Тронь Віталій Валерійович** - кандидат технічних наук, доцент, доцент кафедри автоматизації, комп'ютерних наук і технологій, Криворізький національний університет.

ORCID: <https://orcid.org/0000-0002-6149-5794>

**Карабут Надія Олександрівна** - старший викладач кафедри моделювання та програмного забезпечення, Криворізький національний університет.

ORCID: <https://orcid.org/0000-0002-2327-4595>

**Ruban Serhii Anatoliyovych** - Ph.D., Head of the Department of Automation, Computer Sciences and Technology, Kryvyi Rih National University.

ORCID: <https://orcid.org/0000-0002-4495-6667>

**Shvets Dmytro Valeriyovych** - Ph.D., Associate Professor of the Modeling and Software Department, Kryvyi Rih National University.

ORCID: <https://orcid.org/0000-0001-5126-6405>

**Kotov Ihor Anatoliyovych** - Doctor of Sciences, Professor of the Modeling and Software Department, Kryvyi Rih National University.

ORCID: <https://orcid.org/0000-0003-2445-6259>

**Tron Vitaliy Valeriyovych** - Ph.D., Associate Professor of the Department of Automation, Computer Sciences and Technology, Kryvyi Rih National University.

ORCID: <https://orcid.org/0000-0002-6149-5794>

**Karabut Nadiya Oleksandrivna** - Senior Lecturer, Modeling and Software Department, Kryvyi Rih National University.

ORCID: <https://orcid.org/0000-0002-2327-4595>.

## ФОРМУВАННЯ МІКРОСТРУКТУРИ НІКЕЛЕВОГО СУПЕРСПЛАВУ ABD900 ПІСЛЯ LPBF ТА ЙОГО ПОРІВНЯННЯ ЗІ СПЛАВОМ INCONEL 718

*Анотація.* У роботі представлено результати дослідження формування мікроструктури та порівняльного аналізу нікелевих суперсплавів ABD900 та Inconel 718, виготовлених методом лазерного плавлення в порошковому шарі (Laser Powder Bed Fusion, LPBF). Адитивні технології в останні роки активно застосовуються у високотехнологічних галузях промисловості завдяки можливості виготовлення деталей зі складною геометрією та високими експлуатаційними характеристиками. Разом із тим використання традиційних жароміцних сплавів, розроблених для лиття або деформаційних технологій, у процесах адитивного виробництва може супроводжуватися рядом технологічних обмежень, зокрема утворенням тріщин, пористості та нестабільністю мікроструктури. У зв'язку з цим значна увага приділяється розробці нових сплавів, спеціально оптимізованих для процесів адитивного виробництва. Одним із таких матеріалів є нікелевий суперсплав ABD900, створений із застосуванням концепції Alloy By Design та призначений для використання в процесах LPBF. Порівняльний аналіз пористості показав, що у зразках зі сплаву Inconel 718 переважають пори розміром 2–10 мкм із максимальними значеннями до 12–14 мкм, тоді як у сплаві ABD900 в цілому менша кількість пор. Їх розміри переважно в діапазоні 2–12 мкм, не було виявлено поодиноких пор з більшим розміром порівняно зі зразком, виготовленим з Inconel 718. Мікроструктурний аналіз показав формування характерної для процесу LPBF лускатої структури треків плавлення. У зразках зі сплаву Inconel 718 спостерігається більша виражена клітинно-дендритна морфологія кристалізації, тоді як у сплаві ABD900 в окремих ділянках мікроструктура характеризується формуванням дрібноклітинної  $\gamma$ -Ni матриці. Отримані результати підтверджують перспективність використання сплаву ABD900 для адитивного виробництва та демонструють особливості формування його мікроструктури у порівнянні з Inconel 718.

*Ключові слова:* LPBF, нікелеві суперсплави, ABD900, Inconel 718, мікроструктура, пористість.

**Вступ.** Адитивні технології (АТ) є сучасною технологією, що наразі широко застосовується в різних галузях авіаційної, енергетичної, медичної, машинобудівної. АТ демонструють трансформацію сучасної металургії та матеріалознавства, забезпечуючи швидке виготовлення деталей зі складною геометрією та контрольованими механічними

ми властивостями. АТ в цілому має великий перелік методів вирощування деталей кожна з яких має свої особливості. В даній роботі особливу увагу буде приділено методу лазерного плавлення в порошковому шарі (LPBF). Дана технологія добре зарекомендувала себе з точки зору застосування в різних вище зазначених галузях [1,2] через значні покращення продуктивності, відносної щільності та можливість вирощування деталей зі складними геометричними параметрами.

Однак поряд із технологічними параметрами LPBF важливим фактором є матеріал, що використовується для виготовлення деталей. В більшості випадків в даній області застосовується матеріали, що добре себе зарекомендували в області традиційних технологій. Хоч застосовувані матеріали і достатньо добре висвітлені в наукових публікаціях та мають велику кількість відповідної нормативно-технічної документації, що дає значне розуміння процесів формування структури та властивостей під час кристалізації та обробки, то застосування даних матеріалів для методу LPBF спонукає до нових викликів для області матеріалознавства через складності у виготовленні деталей. Особливу увагу в даному дослідженні планується приділити нікелевим сплавам, що застосовуються в авіаційній галузі.

**Аналіз останніх досліджень.** Нікелеві сплави доволі часто використовуються в авіаційній галузі, ці сплави відомі своєю винятковою високотемпературною міцністю та корозійною стійкістю, що робить їх привабливими для застосування в агресивних умовах експлуатації. Однак наразі асортимент нікелевих суперсплавів, що є комерційно вигідними для АТ, обмежений кількома ключовими сплавами, такими як Inconel 718, Haynes 282 та ін., які підходять для застосування при температурі до 800°C. Дане обмеження є перешкодою, особливо враховуючі зростаючий попит на деталі, які експлуатуються при більш високих температурах.

Враховуючи вищезазначене, дослідники [3] високу увагу приділяють розробці нових сплавів таких, як сплав ABD900. Даний сплав був розроблений новим інноваційним підходом Alloy By Design, що є значним проривом, так як сплав ABD900 було спеціально розроблено для АТ.

Сплав ABD900 володіє, так званою, властивістю «покращеної друкованості», як згадується в деяких роботах [3-4]. Це пояснюється високим діапазоном технологічних параметрів, при яких вирощування деталей різного перерізу досягає рівня відносної щільності більше 99,0%. Даний нікелевий сплав має в своєму складі до десяти легуючих елементів, таких як Cr, Si, Co, що сприяють підвищенню стійкості до окислення; Mo, W – уповільнюють термічно-активні, деформаційні процеси в тому числі зсув дислокацій; Al, Ta, Ti – сприяють дисперсійному зміцненню, та є необхідними для формування  $\gamma$  - фази.

Як стверджується в деяких роботах [5], у сплаві ABD900 після виготовлення методом LPBF відсутня дендритна морфологія кристалізації ванни розплаву, і структура складається переважно з дрібноклітинної фази  $\gamma$ -Ni. Це вказує на можливість не застосовувати параметри термічної обробки на твердий розчин, що доволі часто застосовують в традиційних технологіях та до сплавів, що в першу чергу виготовлялись для тра-

диційних технологій та застосовуються в LPBF. Незважаючи на значний інтерес до сплаву ABD900, особливості формування його мікроструктури після LPBF та порівняння з традиційними суперсплавами, зокрема Inconel 718, досліджені недостатньо. Таким чином в даній роботі буде проведено порівняльний аналіз сплаву ABD900 в порівнянні зі сплавом Inconel 718.

**Мета дослідження.** Метою роботи є дослідження формування мікроструктури та морфології кристалізації суперсплаву на основі нікелю ABD900 після виготовлення методом LPBF та порівняння отриманих характеристик зі сплавом Inconel 718.

**Методика проведення досліджень.** В рамках дослідження використано дослідні зразки розміром 10x10x5 мм, що виготовлені з застосуванням 3D-принтеру виробництва ТОВ «Адитивні лазерні технології України».

Дослідні зразки були виготовлені з жароміцного сплаву ABD900 та Inconel 718, хімічний склад представлено в таблиці 1. Гранулометричний аналіз дослідного порошку показав, що розмір частинок складає  $20 \pm 15$  мкм.

Таблиця 1

Хімічний склад дослідних зразків, % ваг.

Матеріал	Cr	C	Cu	Nb	Al	Mo	Ti	Co	Si	Ta	W	Fe	Ni
Inconel 718	17,8	0,02	0,06	5,5	0,5	3,0	0,9	0,3	0,07	-	-	18,0	База
ABD 900	17,96	0,06	0,02	1,62	2,7	2,6	2,52	19,1	0,09	1,44	2,66	0,07	База

Дослідні зразки для металографічного аналізу були виготовлені за ідентичними технологічними параметрами: потужність - 195 Вт, швидкість сканування - 1100 мм/с, відстань між треками - 0,09 мм, номінальна товщина шару – 40 мкм, питома щільність енергії – 49,2 Дж/мм<sup>3</sup>.

Аналіз відносної щільності було виконано на поперечному перерізі мікروشліфів зразків, механічна підготовка виконувалось механічним способом з застосуванням змінних дисків різної зернистості та послідовним поліруванням на спеціалізованому обладнанні Struers LaboPol-5. Пористість в зразках вимірювались з застосуванням програмного забезпечення ImageJ з використанням вбудованої функції «Аналіз частинок». Під час досліджень елементи розміром менше 2 мкм відфільтровувались для запобігання впливу потенційних дефектів, що сформувались в процесі підготовки зразків, на результат оцінки пористості.

Аналіз мікроструктури виконувався на хімічно травлених зразках із застосуванням розчину Каллінга №1 (Inconel 718) та Каллінга №2 (ABD900).

Оптична мікроскопія виконувалось в світлопольному освітленні з застосуванням інвертованого мікроскопа Axiovert 200 Mmat.

**Результати досліджень.** Аналіз дослідного зразка в полірованому стані з Inconel 718 (рис. 1) та ABD 900 (рис. 2) показав, що відносна щільність зразка складає 99,8%.

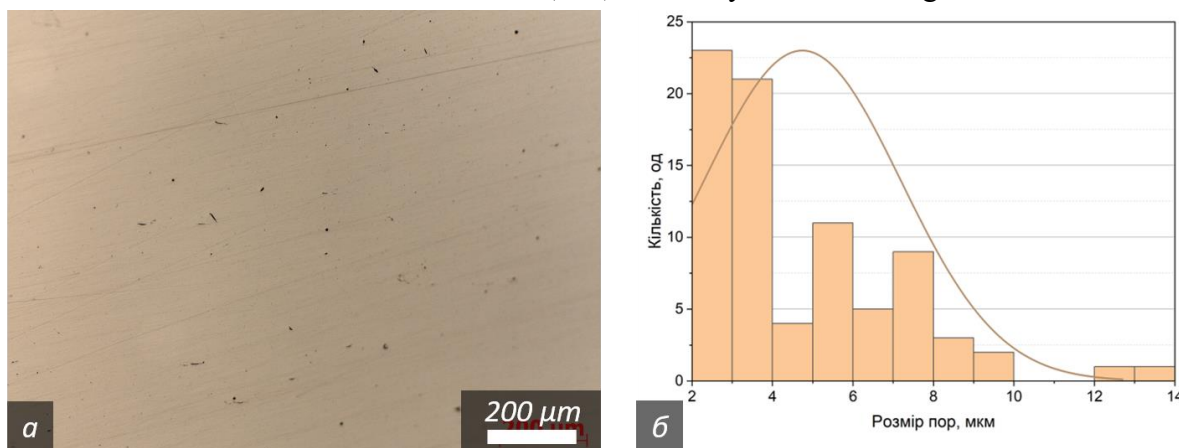


Рисунок 1 – Аналіз в полірованому стані (а) дослідного зразка виготовленого з Inconel 718, кількісна оцінка (б)

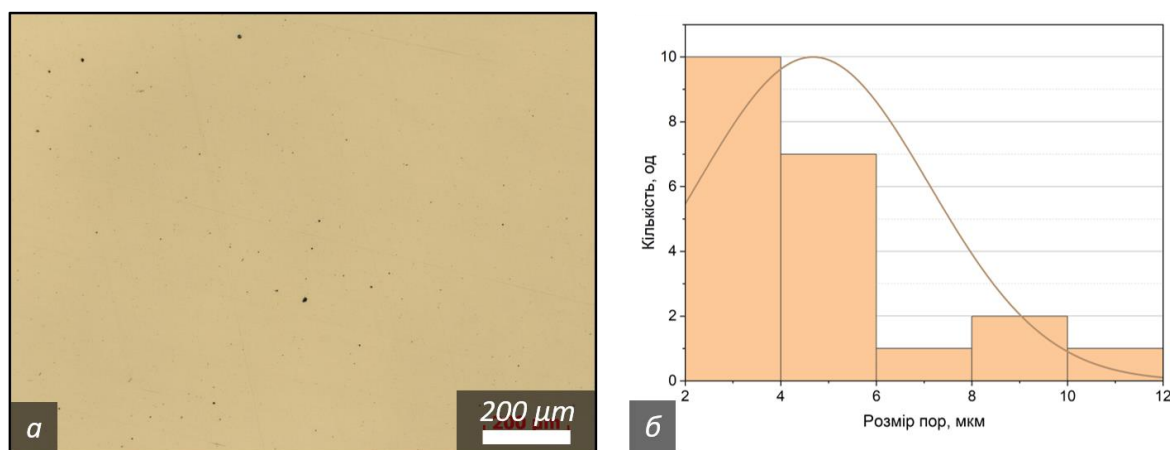


Рисунок 2 – Аналіз в полірованому стані (а) дослідного зразка виготовленого з ABD900, кількісна оцінка (б)

При більш детальному аналізі дослідних зразків в полірованому стані було встановлено, що при виготовленні за однаковими-технологічними параметрами кількість виявлених пор в дослідному зразку з Inconel 718 менше, ніж в зразку, що був виготовлений з ABD900. Проведений аналіз (рис. 1-2) показав, що дослідний зразок з Inconel 718 має пори розміром в діапазоні 2-10 мкм, а максимальний розмір виявлених поодиноких пор склав 12-14 мкм. В той же час в дослідному зразку ABD900 присутня в цілому менша кількість пор. Їх розміри переважно в діапазоні 2-12 мкм, не було виявлено поодиноких пор з більшим розміром порівняно зі зразком, виготовленим з Inconel 718. В цілому, проведений порівняльний аналіз показав, що отримані результати узгоджуються з повідомленнями інших авторів щодо підвищеної технологічної придатності сплаву ABD900 [4]. Безперечно вони пов'язували цей термін з тріщиноутворенням, але як показують ці результати, то цей термін, ймовірно, можна попередньо пов'язати з пороутворенням. При порівняльному дослідженні відносної щільності даних дослідних зразків, виготовлених з різних матеріалів, при застосуванні раціональних параметрів можна стверджувати, що відносна щільність у дослідних зразках на рівні 99,8%.

На наступному етапі було проведено дослідження мікроструктури дослідних зразків (рис. 3).

В результаті проведеного аналізу мікроструктури дослідних зразків встановлено, що мікроструктура переважно складається з типової лускатої структури. В результаті аналізу мікроструктури слід зазначити, що дослідний зразок, виготовлений з Inconel 718 має зони з вираженими різноорієнтованими дендритами в більшості треків. В той час, як в дослідному зразку ABD900 також прослідковувались області з дендритами, але в деяких треках структура складалась переважно з  $\gamma$ - матриці. Ймовірно, це пов'язано з застосованим реактивом для хімічного виявлення структури або з вищезначеним ефектом при кристалізації сплаву ABD900, а саме з формуванням дрібноклітинної структури  $\gamma$ -Ni. У зв'язку з цим виникає необхідність у проведенні поглиблених досліджень із залученням більш детального мікроструктурного аналізу та оцінки мікроструктури з метою інтерпретації сформованої структури.

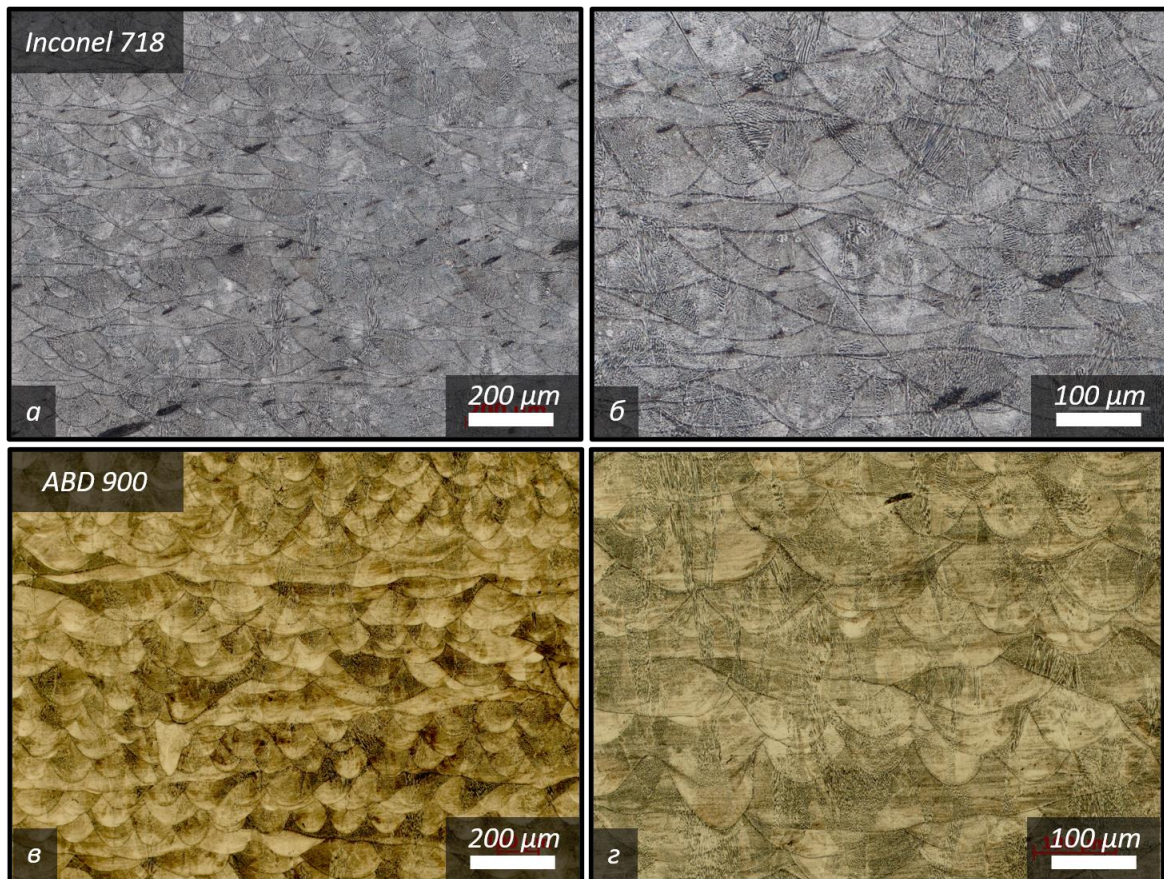


Рисунок 3 – Мікроструктура дослідних зразків з Inconel 718 (а, б) та ABD900 (в, г)

Подальші дослідження доцільно спрямувати на детальне дослідження субструктури кристалізації сплаву ABD900 із аналізом фазового складу, а також оцінку механічних властивостей після різних режимів термічної обробки.

#### **Висновки:**

1. Встановлено, що при застосуванні однакових раціональних технологічних параметрів процесу LPBF відносна щільність дослідних зразків зі сплавів Inconel 718 та ABD900 становить близько 99,8 %.

2. Порівняльний аналіз пористості показав, що у зразках зі сплаву Inconel 718 переважають пори розміром 2–14 мкм, тоді як у сплаві ABD900 спостерігається в цілому менша кількість пор з розмірами переважно в діапазоні 2-12 мкм.

3. Мікроструктурний аналіз показав формування характерної для процесу LPBF лускатої структури треків плавлення та клітинно-дендритної морфології кристалізації.

4. У зразках зі сплаву Inconel 718 спостерігається більш виражена дендритна структура, тоді як у сплаві ABD900 в окремих ділянках формується дрібноклітинна  $\gamma$ -Ni структура.

5. Отримані результати підтверджують перспективність використання сплаву ABD900 для адитивного виробництва та демонструють особливості формування його мікроструктури у порівнянні з Inconel 718.

#### ЛІТЕРАТУРА

1. Blakey-Milner B., Gradl P., Snedden G., Brooks M., Pitot J., Lopez E., Leary M., Berto F., du Plessis A. Metal additive manufacturing in aerospace: A review. *Materials & Design*. 2021, 209, 110008. <https://doi.org/10.1016/j.matdes.2021.110008>
2. Kennedy S. M., Vasanthanathan A., Amudhan K. Exploring the frontiers of metal additive manufacturing in orthopaedic implant development. *MethodsX*. 2024. 13, 103056. <https://doi.org/10.1016/j.mex.2024.103056>
3. Reed R.C., Tao T., Warnken N. Alloys-By-Design: Application to nickel-based single crystal superalloys. *Acta Materialia*. 2009. 57, 19, 5898-5913. <https://doi.org/10.1016/j.actamat.2009.08.018>
4. Tang Y.T., Panwisawas C., Ghousoub J. N., Gong Y., Clark J. W.G, Németh A. A.N., McCartney D. G., Reed R. C. Alloys-by-design: Application to new superalloys for additive manufacturing. *Acta Materialia*. 2021. 202, 417-436. <https://doi.org/10.1016/j.actamat.2020.09.023>
5. Bridges A., Shingledecker J., Hussain Z., Forster J., Nemeth A., Vogel G., O’Nora N. Acceleration of Material Acceptance and Industry Adoption of an Additively Manufactured Nickel-base Superalloy. Proceedings of the AM-EPRI 2024. Advances in Materials, Manufacturing, and Repair for Power Plants: Proceedings from the Tenth International Conference. Indian Wells, California, USA. 74-87. ASM. <https://doi.org/10.31399/asm.cp.am-epri-2024p0074>

#### REFERENCES

1. Blakey-Milner B., Gradl P., Snedden G., Brooks M., Pitot J., Lopez E., Leary M., Berto F., du Plessis A. (2021). Metal additive manufacturing in aerospace: A review. *Materials & Design*. 209, 110008. <https://doi.org/10.1016/j.matdes.2021.110008>
2. Kennedy S. M., Vasanthanathan A., Amudhan K. (2024). Exploring the frontiers of metal additive manufacturing in orthopaedic implant development. *MethodsX*. 13, 103056. <https://doi.org/10.1016/j.mex.2024.103056>
3. Reed R.C., Tao T., Warnken N. (2009). Alloys-By-Design: Application to nickel-based single crystal superalloys. *Acta Materialia*. 57, 19, 5898-5913. <https://doi.org/10.1016/j.actamat.2009.08.018>

4. Tang Y.T., Panwisawas C., Ghoussoub J. N., Gong Y., Clark J. W.G, Németh A. A.N., McCartney D. G., Reed R. C. (2021). Alloys-by-design: Application to new superalloys for additive manufacturing. *Acta Materialia*. 202, 417-436.

<https://doi.org/10.1016/j.actamat.2020.09.023>

5. Bridges A., Shingledecker J., Hussain Z., Forster J., Nemeth A., Vogel G., O’Nora N. (2024). Acceleration of Material Acceptance and Industry Adoption of an Additively Manufactured Nickel-base Superalloy. Proceedings of the AM-EPRI 2024. Advances in Materials, Manufacturing, and Repair for Power Plants: Proceedings from the Tenth International Conference. Indian Wells, California, USA. 74-87. ASM. <https://doi.org/10.31399/asm.cp.am-epri-2024p0074>

Received 24.04.2026.

Accepted 27.04.2026.

Published 30.04.2026

### ***Microstructure formation of the nickel-based superalloy ABD900 after LPBF and its comparison with Inconel 718***

*The paper presents the results of a study on microstructure formation and a comparative analysis of the nickel-based superalloys ABD900 and Inconel 718 manufactured by Laser Powder Bed Fusion (LPBF). In recent years, additive manufacturing technologies have been increasingly applied in high-tech industries due to their ability to produce components with complex geometries and high-performance characteristics. However, the use of conventional high-temperature alloys, originally developed for casting or thermomechanical processing, in additive manufacturing processes may be associated with several technological limitations, including cracking, porosity, and microstructural instability. In this context, considerable attention has been devoted to the development of new alloys specifically optimized for additive manufacturing. One such material is the nickel-based superalloy ABD900, developed using the Alloy By Design approach and intended for LPBF applications. A comparative analysis of porosity revealed that samples of Inconel 718 are characterized by pores predominantly in the range of 2–10  $\mu\text{m}$ , with maximum sizes reaching 12–14  $\mu\text{m}$ , whereas the ABD900 alloy exhibits an overall lower porosity. The pore sizes in ABD900 are mainly within the range of 2–12  $\mu\text{m}$ , and no isolated larger pores were observed compared to the Inconel 718 samples. Microstructural analysis demonstrated the formation of a characteristic melt track morphology typical for the LPBF process. In the Inconel 718 samples, a more pronounced cellular-dendritic solidification structure was observed, whereas in the ABD900 alloy, certain regions of the microstructure are characterized by the formation of a fine cellular  $\gamma$ -Ni matrix. The obtained results confirm the potential of ABD900 for additive manufacturing applications and highlight the specific features of its microstructure formation in comparison with Inconel 718.*

*Keywords: LPBF, nickel superalloys, ABD900, Inconel 718, microstructure, porosity.*

**Аджамський Сергій Вікторович** – PhD, технічний директор LLC «Additive Laser Technology of Ukraine», старший науковий співробітник Інституту транспортних систем і технологій Національної академії наук України, Дніпро, Україна.

ORCID: <https://orcid.org/0000-0002-6095-8646>

**Кононенко Анна Андріївна** - докт. техн. наук, ст. досл., старший науковий співробітник Інституту чорної металургії ім. З.І. Некрасова НАН України, професор НТУ «Дніпровська Політехніка», інженер-матеріалознавець LLC «Additive Laser Technology of Ukraine», Дніпро, Україна.

ORCID: <https://orcid.org/0000-0001-7446-4105>

**Подольський Ростислав Вячеславович** – докт. філ., старший науковий співробітник Інституту чорної металургії ім. З.І. Некрасова НАН України, старший науковий співробітник Інститут прикладних систем управління НАН України, інженер-матеріалознавець LLC «Additive Laser Technology of Ukraine», старший науковий співробітник Інститут прикладних систем управління НАН України, Дніпро, Україна.

ORCID: <https://orcid.org/0000-0002-0288-0641>

**Бадюк Сергій Іванович** – канд.техн.наук, інженер-технолог LLC «Additive Laser Technology of Ukraine», старший науковий співробітник Інститут прикладних систем управління НАН України, Дніпро, Україна.

ORCID: <https://orcid.org/0000-0002-1074-3057>

**Подольська Олена Анатоліївна** – науковий співробітник Інституту чорної металургії ім. З.І. Некрасова НАН України, Дніпро, Україна.

ORCID: <https://orcid.org/0000-0002-4032-4275>

**Adjamskiy Sergey** – PhD, Technical Director of Additive Laser Technology of Ukraine LLC, senior researcher at the Institute of Transport Systems and Technologies of the National Academy of Sciences of Ukraine, Dnipro, Ukraine.

ORCID: <https://orcid.org/0000-0002-6095-8646>

**Kononenko Ganna** - Doct. technical science, sen. researcher, scientific secretary of the Institute of Iron and Steel of Z.I. Nekrasov NAS of Ukraine, Professor of NTU "Dnipro Polytechnic", materials engineer LLC Additive Laser Technology of Ukraine, Dnipro, Ukraine.

ORCID: <https://orcid.org/0000-0001-7446-4105>

**Podolskiy Rostyslav** – PhD, senior researcher of the Institute of Iron and Steel of Z.I. Nekrasov NAS of Ukraine, materials engineer LLC Additive Laser Technology of Ukraine, Senior Researcher at the Institute of Applied Control Systems NAS of Ukraine, Dnipro, Ukraine.

ORCID: <https://orcid.org/0000-0002-0288-0641>

**Baduk Sergey** – Cand. technical science, materials engineer LLC Additive Laser Technology of Ukraine, Process engineer LLC "Additive Laser Technology of Ukraine", senior researcher, Institute of Applied Control Systems, NAS of Ukraine, Dnipro, Ukraine.

ORCID: <https://orcid.org/0000-0002-1074-3057>

**Podolska Olena** – researcher of the Institute of Iron and Steel of Z.I. Nekrasov NAS of Ukraine, Dnipro, Ukraine.

ORCID: <https://orcid.org/0000-0002-4032-4275>

## ВИКОРИСТАННЯ ХМАРНИХ ТЕХНОЛОГІЙ У НАВЧАЛЬНОМУ ПРОЦЕСІ

*Анотація.* У цифрову епоху хмарні технології стають важливим інструментом підготовки здобувачів освіти. Вони забезпечують цілодобовий доступ до ресурсів, полегшуючи дистанційне та гібридне навчання. Дослідження аналізує ефективність та перспективи таких технологій у формуванні професійних ІТ-навичок. Хмарні платформи використовуються для програмування, аналізу даних та розробки ПЗ, відкриваючи доступ до віртуальних лабораторій і потужних обчислювальних ресурсів. Хмарні сервіси сприяють розвитку цифрових компетенцій, підвищують мотивацію й самостійність студентів, підтримують групову роботу та ефективну комунікацію. Застосування таких рішень покращує якість освіти та підвищує конкурентоспроможність випускників ІТ-сфери. Подальші дослідження можуть зосередитися на розробці методик використання хмарних технологій і створенні спеціалізованих лабораторій для поглиблення технічних умінь здобувачів освіти.

*Ключові слова:* хмарні технології, освітній процес, обробка даних, доступність та гнучкість.

**Вступ.** Модернізація та подальший розвиток навчально-наукового середовища у закладах вищої освіти є одним із пріоритетних напрямів реформування сучасної освітньої системи. Умови цифрової трансформації суспільства зумовлюють необхідність переосмислення підходів до організації освітнього процесу, а також до формування ефективної інфраструктури, що забезпечує доступ до освітніх і наукових ресурсів. У цьому контексті особливого значення набуває впровадження хмарних технологій і сервісів, які сприяють створенню гнучкого, масштабованого та доступного інформаційно-освітнього середовища.

Дослідницькі підходи у цій сфері охоплюють широкий спектр питань, серед яких важливе місце посідає інтеграція електронних освітніх ресурсів, цифрових сервісів і програмного забезпечення в єдину інформаційну інфраструктуру закладу вищої освіти. Значна увага приділяється можливостям організації спільної роботи учасників освітнього процесу із застосуванням хмарних платформ, що забезпечують ефективну комунікацію, обмін даними та колективне опрацювання навчальних матеріалів. Водночас важливим аспектом є подолання географічних і часових бар'єрів у процесі здобуття освіти, що відкриває нові перспективи для розвитку дистанційного та змішаного навчання. Застосування хмарних технологій сприяє також підвищенню мобільності студентів і викладачів та розширенню доступу до міжнародних освітніх ресурсів.

Хмарні технології розглядаються як один із найбільш перспективних інструментів для розв'язання актуальних соціальних і освітньо-професійних проблем сучасного суспільства. Їх використання дозволяє забезпечити тісніший взаємозв'язок між науковими дослідженнями та підготовкою кваліфікованих фахівців, оскільки сучасні платформи надають інструменти для спільної дослідницької діяльності, зберігання й аналізу великих обсягів даних, а також для організації відкритого доступу до результатів наукової роботи.

Завдяки своїм функціональним можливостям хмарні технології виступають ефективним засобом реалізації принципів людиноцентризму в освіті. Вони забезпечують персоналізацію навчання, створюють умови для адаптації освітніх ресурсів до індивідуальних потреб студентів і підтримують різноманітні освітні траєкторії. Водночас такі технології сприяють забезпеченню рівного доступу до освітніх можливостей, що є важливим чинником розвитку сучасної системи вищої освіти та формування інклюзивного освітнього середовища.

Сучасний стан функціонування освітньо-наукового середовища характеризується зростанням вимог до якості електронних освітніх ресурсів, які використовуються як у навчальній, так і в науково-дослідній діяльності. Особливої актуальності набуває створення інтерактивних навчальних матеріалів, електронних курсів, цифрових бібліотек і репозиторіїв наукових даних, що забезпечують відкритий доступ до інформації.

Інтеграція мережевих технологій у діяльність ЗВО, спрямована на створення відкритого інформаційно-освітнього простору, набуває дедалі більшого значення в умовах глобалізації та інтенсивного розвитку цифрових комунікацій. Варто зазначити, що технічні можливості таких технологій суттєво розширилися завдяки впровадженню хмарно-орієнтованих підходів до зберігання, обробки та поширення інформації. У цьому контексті хмарні сервіси можуть виступати ключовим інструментом для поглиблення взаємозв'язків між освітою, наукою та виробництвом, що є важливим чинником розвитку інноваційної економіки.

Використання хмарних технологій також створює сприятливі умови для розвитку дуальної освіти, у межах якої поєднуються теоретична підготовка студентів у закладі освіти та практична діяльність на підприємствах. Завдяки хмарним платформам стає можливим організувати спільні освітні та дослідницькі проекти, забезпечувати доступ до корпоративних інформаційних ресурсів і проводити дистанційні консультації між викладачами, студентами та представниками виробничого сектору. Крім того, такі технології відкривають нові можливості для створення корпоративних структур, зокрема філій закладів освіти на базі підприємств, навчально-дослідних лабораторій, інжинірингових та інноваційних центрів.

**Актуальність дослідження** зумовлена потребою вдосконалення процесів інтеграції інформаційно-комунікаційних технологій, заснованих на використанні хмарних обчислень, у сучасні навчальні практики ЗВО. Цифровізація освітньої галузі та активне впровадження інформаційних технологій у професійну діяльність фахівців обумовлюють необхідність переосмислення традиційних підходів до організації освітнього про-

цесу. У цьому контексті особливого значення набуває використання хмарних сервісів, які забезпечують гнучкість, масштабованість та доступність освітніх ресурсів. У межах сучасного Європейського освітнього простору такі технології розглядаються як один із провідних інструментів цифрової трансформації освіти, що сприяє підвищенню ефективності навчання, розширенню можливостей комунікації та формуванню нових освітніх практик.

З огляду на це актуальним завданням є не лише розширення масштабів застосування хмарних технологій у закладах освіти, а й підвищення ефективності їх використання у навчальному процесі. Важливою складовою цього процесу виступає підготовка висококваліфікованих фахівців, які володіють сучасними цифровими компетентностями та здатні ефективно використовувати інформаційно-комунікаційні технології у професійній діяльності. У цьому контексті створення хмаро-орієнтованого освітньо-наукового середовища у ЗВО розглядається як один із ключових факторів підготовки спеціалістів. Завдяки цьому формується здатність майбутніх фахівців до цілеспрямованого, науково обґрунтованого та результативного застосування хмарних технологій у професійній діяльності, зокрема у сфері інформаційних технологій.

Технологія хмарних обчислень являє собою сучасний інноваційний підхід до організації використання обчислювальних ресурсів і програмних сервісів, що надаються користувачам через мережу Інтернет. В основі цієї технології лежить модель віддаленого доступу до обчислювальних потужностей, програмного забезпечення та інформаційних ресурсів, які розміщуються на спеціалізованих серверах. Такий підхід дозволяє значно оптимізувати процеси управління інформаційними ресурсами, забезпечує зручність доступу до необхідних матеріалів та сприяє підвищенню ефективності освітньої діяльності. У контексті освітньої діяльності хмарні технології відіграють важливу роль у забезпеченні доступу до широкого спектра електронних навчальних ресурсів. До таких ресурсів належать цифрові підручники, електронні конспекти лекцій, методичні матеріали для виконання лабораторних і практичних робіт, мультимедійні навчальні матеріали, відеолекції, інтерактивні тренажери, тестові завдання та інші інформаційні ресурси, що використовуються у навчальному процесі. Використання подібних матеріалів сприяє підвищенню якості засвоєння навчального матеріалу, розширює можливості самостійної роботи студентів та забезпечує більш ефективну організацію освітнього процесу. Додатково технології хмарних обчислень широко застосовуються для збереження та систематизації даних, адміністрування інформаційних ресурсів і управління документацією закладів освіти. Автоматизація таких процесів дозволяє значно підвищити ефективність управління освітньою діяльністю та оптимізувати використання інформаційних ресурсів.

Важливою перевагою хмарних технологій є також забезпечення високого рівня захисту інформації. Сучасні хмарні платформи використовують різноманітні механізми безпеки, зокрема резервне копіювання даних, шифрування інформації, багаторівневу систему автентифікації користувачів та контроль доступу до інформаційних ресурсів. Це дозволяє мінімізувати ризики втрати даних, запобігти несанкціонованому доступу до інформації та забезпечити надійність зберігання освітніх і наукових матеріалів.

Впровадження хмарних технологій супроводжується низкою суттєвих переваг, які зумовлюють їхню актуальність і поширеність у сучасній системі освіти.

1. Доступність і гнучкість. Однією з основних характеристик хмарних технологій є забезпечення доступу до програмного забезпечення, навчальних матеріалів, баз даних та інших ресурсів з будь-якого пристрою, підключеного до мережі Інтернет, не залежно від географічного місцезнаходження користувача та часу доступу до ресурсів, що дуже важливо в теперішні часи.

Це сприяє розвитку дистанційного та змішаного навчання, підвищує мобільність учасників освітнього процесу та забезпечує більш гнучку організацію навчальної діяльності.

2. Оптимізація спільного використання ресурсів. Хмарні платформи дозволяють ефективно організувати спільне використання обчислювальних потужностей, програмних продуктів і сховищ даних. Завдяки централізованому розміщенню ресурсів з'являється можливість їхнього раціонального розподілу між користувачами, що сприяє підвищенню продуктивності роботи інформаційних систем.

3. Економія часу та фінансових ресурсів. Застосування хмарних сервісів дозволяє значно скоротити витрати часу на організаційні та технічні аспекти освітнього процесу. Швидкий доступ до необхідних ресурсів, автоматизовані системи управління навчальними матеріалами та централізоване зберігання інформації сприяють підвищенню ефективності освітньої діяльності. Крім того, використання хмарної інфраструктури знижує фінансові витрати на придбання програмного забезпечення, його оновлення та технічну підтримку.

4. Відсутність необхідності локального встановлення програмного забезпечення. Однією з важливих переваг хмарних технологій є можливість використання програмних продуктів без їхнього встановлення на локальні пристрої користувачів. Програмне забезпечення функціонує безпосередньо у хмарному середовищі, а користувачі отримують доступ до нього через веббраузер або спеціалізовані клієнтські додатки. При цьому оновлення програм відбувається автоматично, що дозволяє працювати з актуальними версіями сервісів без необхідності виконання додаткових технічних операцій.

5. Резервне копіювання та збереження даних. Хмарні сервіси надають можливість зберігати інформацію на віддалених серверах, що значно підвищує надійність її збереження. У випадку втрати або пошкодження локального обладнання користувачі мають можливість швидко відновити необхідні дані. Більшість сучасних хмарних платформ підтримують функцію автоматичного резервного копіювання, яка забезпечує регулярне створення копій інформації та гарантує її збереження навіть у разі технічних збоїв або інших непередбачуваних ситуацій.

6. Забезпечення високого рівня безпеки та конфіденційності. Провайдери хмарних послуг, як правило, впроваджують сучасні технології захисту інформації, що забезпечують високий рівень безпеки даних. До таких технологій належать механізми шифрування, системи автентифікації користувачів, багаторівневий контроль доступу, а також спеціалізовані інструменти моніторингу інформаційних систем. Завдяки цьому за-

безпечується конфіденційність освітніх і наукових даних та формується довіра користувачів до використання хмарних сервісів.

Хмарні технології широко застосовуються для зберігання, обробки та систематизації інформації, що використовується у навчальному процесі. Централізоване зберігання такої інформації у хмарному середовищі дає змогу ефективно організувати її облік, аналіз і подальше використання для моніторингу результатів навчання, оцінювання академічної успішності та планування освітнього процесу. Завдяки використанню хмарних сервісів викладачі отримують можливість швидко оновлювати навчальні матеріали, забезпечувати доступ до них здобувачам освіти та організувати ефективну комунікацію в межах освітнього середовища.

Важливим аспектом впровадження хмарних технологій у сферу освіти є питання економічної доцільності їх використання. Витрати на обслуговування хмарної інфраструктури можуть варіюватися залежно від масштабів використання сервісів, кількості користувачів та функціональних можливостей обраних платформ. Основними витратами у цьому випадку є оплата підписки на хмарні сервіси та платформи, які забезпечують зберігання, обробку та передачу даних у хмарному середовищі. Такі підписки можуть відрізнятися за рівнем доступу, обсягом доступного сховища даних, кількістю користувачів та набором функціональних можливостей. Деякі освітні платформи також пропонують спеціальні освітні ліцензії або безкоштовні версії для закладів освіти, що додатково знижує фінансове навантаження на освітні установи. Додаткові витрати можуть виникати на етапі впровадження хмарних технологій, зокрема під час налаштування системи, інтеграції хмарних сервісів з іншими інформаційними платформами закладу освіти, а також підготовки персоналу до роботи з новими технологічними інструментами. Проте навіть з урахуванням цих витрат використання хмарних рішень часто є більш економічно ефективним, ніж підтримка традиційної IT-інфраструктури.

Одним із перспективних напрямів використання хмарних технологій у навчальному процесі є застосування віртуальних лабораторій та інтерактивних освітніх платформ. Віртуальні лабораторії створюють можливість виконання експериментів і практичних завдань у спеціалізованому онлайн-середовищі. За допомогою таких платформ здобувачі освіти можуть моделювати різноманітні процеси, проводити дослідження та аналізувати отримані результати без необхідності використання дорогого лабораторного обладнання. Особливо актуальним є застосування віртуальних лабораторій у тих випадках, коли проведення реальних експериментів ускладнене через фінансові, технічні або безпекові обмеження. Це сприяє розвитку дослідницьких навичок, критичного мислення та практичних компетентностей, що є важливими складовими професійної підготовки сучасних фахівців.

Інтерактивні освітні платформи, у свою чергу, забезпечують доступ до широкого спектра цифрових навчальних ресурсів, серед яких важливе місце займають мультимедійні матеріали, візуалізації, інтерактивні моделі, діаграми, тести та симуляції. Застосування інтерактивних платформ також сприяє активізації пізнавальної діяльності здобувачів освіти, підвищує їхню мотивацію до навчання та забезпечує більш ефективне засвоєння навчального матеріалу. Завдяки можливості миттєвого зворотного

зв'язку, автоматизованого тестування та аналізу результатів навчання такі платформи створюють умови для більш індивідуалізованого та адаптивного підходу до організації освітнього процесу.

Використовуючи відповідні інструменти, викладачі можуть формувати електронні курси, розміщувати навчальні матеріали, планувати завдання для студентів, організовувати спільну роботу над проектами та здійснювати оцінювання результатів навчальної діяльності. Це сприяє більш раціональному використанню часу під час проведення занять, дозволяє оптимізувати навчальний процес та забезпечує ефективну взаємодію між усіма його учасниками. Застосування віртуальних лабораторій та інтерактивних платформ у навчанні також сприяє зниженню вартості освітнього процесу. Водночас хмарні рішення забезпечують широкий доступ до навчальних ресурсів, що особливо важливо для забезпечення рівного доступу до якісної освіти для здобувачів із різних регіонів. Крім економічних переваг, використання таких технологій сприяє підвищенню якості освітнього процесу. Інтерактивні інструменти, мультимедійні матеріали та можливість практичної роботи у віртуальному середовищі роблять навчання більш динамічним, наочним та цікавим для студентів. У результаті підвищується їхня мотивація до навчання, активізується пізнавальна діяльність та формується більш глибоке розуміння навчального матеріалу.

Сьогодні доступно багато технологічних платформ для хмарних обчислень, які забезпечують сервіси для освіти. Вони пропонують інструменти для зберігання, аналізу даних, розгортання додатків, організації навчальних процесів і впровадження інновацій. Серед найбільш поширених платформ можна виокремити такі:

AmazonWebServices (AWS). AWS є однією з провідних платформ хмарних обчислень у світі, яка пропонує широкий спектр сервісів для різних галузей, зокрема й для освіти. Серед основних можливостей платформи – зберігання великих обсягів даних, надання обчислювальних ресурсів, створення та розгортання вебдодатків, а також використання інструментів машинного навчання та аналітики даних. Освітні установи можуть використовувати AWS для створення віртуальних лабораторій, проведення досліджень, аналізу великих масивів даних та підтримки освітніх онлайн-платформ.

GoogleCloudPlatform (GCP). GoogleCloudPlatform є сучасною хмарною платформою, яка надає різноманітні сервіси для зберігання, обробки та аналізу інформації. Вона включає інструменти для створення вебдодатків, управління базами даних, використання технологій штучного інтелекту та машинного навчання. У сфері освіти GCP активно застосовується для створення освітніх платформ, аналізу навчальних даних, організації спільної роботи студентів і викладачів, а також для підтримки дослідницьких проєктів.

Microsoft Azure. Microsoft Azure є потужною платформою хмарних обчислень, яка пропонує широкий спектр сервісів для різних галузей діяльності, включаючи освітню сферу. За допомогою Azure можна створювати та розгортати різноманітні програмні рішення, використовувати інструменти обробки й аналізу даних, а також застосовувати технології штучного інтелекту. Освітні установи можуть використовувати Azure для

розроблення навчальних додатків, організації віртуальних лабораторій, управління інформаційними ресурсами та підтримки дистанційного навчання.

**IBM Cloud.** IBM Cloud є платформою хмарних обчислень, що пропонує інструменти для розгортання додатків, обробки інформації та аналізу даних. Платформа містить низку сервісів для використання технологій штучного інтелекту, машинного навчання та аналітики. У контексті освітньої діяльності IBM Cloud може застосовуватися для проведення наукових досліджень, розроблення інтелектуальних навчальних систем та аналізу великих обсягів освітніх даних.

**DigitalOcean.** DigitalOcean є популярною платформою хмарних обчислень, що орієнтована на спрощення процесів розгортання та управління цифровою інфраструктурою. Вона пропонує зручні інструменти для створення та підтримки вебдодатків, управління серверами та зберігання даних. Завдяки простоті використання та доступній вартості DigitalOcean часто застосовується у навчальних і дослідницьких проєктах, а також у процесі підготовки фахівців у галузі інформаційних технологій.

Зазначені платформи надають широкий спектр рішень, що можуть використовуватися у різних напрямках освітньої діяльності. Серед них – зберігання та обробка даних, розгортання освітніх вебсистем, використання інструментів машинного навчання, підтримка наукових досліджень та створення інноваційних освітніх сервісів. Завдяки цьому хмарні технології стають важливим компонентом сучасної цифрової освітньої інфраструктури.

**Висновки.** Хмарні технології є ключовим інструментом для впровадження інновацій в освіті. Вони створюють сучасне цифрове освітнє середовище, забезпечують ефективну взаємодію між студентами та викладачами, спрощують доступ до інформаційних ресурсів і інтегрують новітні педагогічні підходи.

Серед їхніх переваг варто виділити доступність освітніх матеріалів, оптимізацію витрат на інфраструктуру, підвищення безпеки та комфорт у роботі з даними. Хмарні сервіси дозволяють зберігати й обробляти дані, надають доступ до віртуальних лабораторій, інтерактивних платформ та інших цифрових інструментів, що роблять навчання практичнішим і технологічно сучасним.

Для успішного використання хмарних технологій важливо розробляти і впроваджувати якісні навчальні матеріали та методичні рекомендації. Сучасні дослідження і література допомагають зрозуміти принципи роботи цих рішень і їхній професійний потенціал. Отже, хмарні технології сприяють підвищенню ефективності освіти, створюючи сучасне інформаційне середовище, яке відповідає потребам цифрового суспільства.

## ЛІТЕРАТУРА

1. Matthew U. O., Kazaure J. S., Okafor N. U. Contemporary development in e-learning education, cloud computing technology and Internet of Things. *EAI Endorsed Transactions on Cloud Systems*. 2021. Vol. 7, No. 20. DOI: <https://doi.org/10.4108/eai.31-3-2021.169173>
2. Han H., Trimi S. Cloud computing-based higher education platforms during the COVID-19 pandemic. *arXiv*. 2022. DOI: <https://doi.org/10.48550/arXiv.2203.03714>.

3. Грушко Р. Вплив хмарних технологій і штучного інтелекту на формування цифрової компетентності на уроках інформатики. Вища освіта України. 2024. № 2(93). DOI: [https://doi.org/10.32782/NPU-VOU.2024.2\(93\).12](https://doi.org/10.32782/NPU-VOU.2024.2(93).12).
4. Пучков І., Сухоручко К. Використання хмарних технологій під час викладання інформатики в закладах вищої освіти України. Гуманізація навчально-виховного процесу. 2025. № 2(108). DOI: [https://doi.org/10.31865/2077-1827.2\(108\)2025.339514](https://doi.org/10.31865/2077-1827.2(108)2025.339514).
5. Mell P., Grance T. The NIST definition of cloud computing. Gaithersburg: National Institute of Standards and Technology, 2022. DOI: <https://doi.org/10.6028/NIST.SP.800-145>.

#### REFERENCES

1. Matthew U. O., Kazaure J. S., Okafor N. U. Contemporary development in e-learning education, cloud computing technology and Internet of Things. EAI Endorsed Transactions on Cloud Systems. 2021. Vol. 7, No. 20. DOI: <https://doi.org/10.4108/eai.31-3-2021.169173>
2. Han H., Trimi S. Cloud computing-based higher education platforms during the COVID-19 pandemic. arXiv. 2022. DOI: <https://doi.org/10.48550/arXiv.2203.03714>.
3. Hrushko R. Vplyv khmarnykh tekhnolohii i shtuchnoho intelektu na formuvannia tsyfrovoyi kompetentnosti na urokakh informatyky. Vyshcha osvita Ukrainy. 2024. № 2(93). DOI: [https://doi.org/10.32782/NPU-VOU.2024.2\(93\).12](https://doi.org/10.32782/NPU-VOU.2024.2(93).12).
4. Puchkov I., Sukhoruchko K. Vykorystannia khmarnykh tekhnolohii pid chas vykladannia informatyky v zakladakh vyshchoi osvity Ukrainy. Humanizatsiia navchalno-vykhovnoho protsesu. 2025. № 2(108). DOI: [https://doi.org/10.31865/2077-1827.2\(108\)2025.339514](https://doi.org/10.31865/2077-1827.2(108)2025.339514).
5. Mell P., Grance T. The NIST definition of cloud computing. Gaithersburg: National Institute of Standards and Technology, 2022. DOI: <https://doi.org/10.6028/NIST.SP.800-145>.

Received 24.04.2026.  
Accepted 27.04.2026.  
Published 30.04.2026

#### ***The use of cloud technologies in the teaching process***

*In this era of rapid digital technology development, cloud solutions are becoming a key component of the education of students pursuing degrees in information technology. The growing volume of information being processed, as well as the constant need for continuous and rapid access to educational resources, further underscore the importance of using cloud platforms in the educational process. Such platforms provide students with a wide range of opportunities: round-the-clock, uninterrupted access to specialized software, high-performance computing resources, and a variety of educational materials, regardless of location or time of day. This functionality is particularly valuable in the context of distance or hybrid learning, when the traditional organization of the educational process in classrooms may not be sufficient.*

*The aim of this study is to conduct a detailed examination of the potential of cloud technologies in the context of training future information technology professionals. The focus is on evaluating the effectiveness of implementing these solutions in the educational process and their impact on the development of students' core professional competencies. The article analyzes conceptual approaches to the use of cloud services in educational institutions, with a focus on the specific features of integrating such technologies into the learning process. Significant attention is paid to the advantages of cloud solutions, which are considered both*

*through the lens of purely technical capabilities and in terms of their educational potential. Furthermore, the article emphasizes the promising prospects of their implementation as a powerful tool for improving the quality of educational programs and training highly qualified specialists who meet the modern demands of the labor market.*

*This overview focuses on analyzing the key features and capabilities of digital platforms that facilitate students' effective mastery of modern technologies, including programming, data analysis, machine learning, and software development. Particular emphasis is placed on the potential of virtual labs, which allow users to experiment with tools and technologies in real time without the need for complex setup of local equipment. Such platforms not only optimize the learning process but also provide access to high-performance resources, while expanding opportunities for collaborative project work. Thanks to this approach, students have the opportunity to master the latest technological tools, which become an integral part of their professional training and the development of the competencies necessary for a successful career in the IT sector.*

*The adoption of cloud technologies plays a key role in developing the digital competencies of today's generation. These solutions significantly boost students' motivation to learn by enabling them to acquire new knowledge independently and effectively, develop critical thinking skills, and engage confidently in research activities. For educators, cloud platforms are becoming a versatile tool for improving the educational process: they facilitate collaboration among students, allow for real-time tracking of their progress, and promote transparent and effective communication.*

*The integration of cloud-based solutions into the education system opens up significant opportunities for improving its quality, making learning more accessible and flexible to meet the individual needs of each student. The implementation of such technological innovations not only helps students acquire skills in using the latest tools of the labor market but also fosters the core competencies critical to developing a competitive professional in the field of information technology. As a result, graduates of modern educational institutions are better prepared for the challenges of a dynamic labor market and become more sought-after by employers.*

*Keywords: cloud technologies, educational process, data processing, accessibility and flexibility.*

**Дмитрієва Ірина Сергіївна** - к.т.н., доцент, доцент кафедри Інформаційних технологій і систем, Український державний університет науки і технологій.

ORCID: <https://orcid.org/0009-0008-3298-7563>

**Dmytriieva Iryna** - Ph.D., Ass.Professor, Associate Professor of the Department of Information Technologies and Systems, Ukrainian State University of Science and Technology.

ORCID: <https://orcid.org/0009-0008-3298-7563>

**ЗМІСТ**

**CONTENTS**

<b>Шиян С.І., Калініна І.О.</b> Імітаційна модель системи вимірювання фізико-хімічних параметрів пмм на основі колірних мереж Петрі	3	<b>Shiyan S., Kalinina I.</b> Simulation model of the system for measuring physical and chemical parameters of fuel and lubricants based on coloured Petri networks	3
<b>Горбатов В.С., Журба А.О.</b> Створення датасету на основі подій класифікованих системою виявлення мережових вторгнень	14	<b>Gorbatov V.S., Zhurba A.O.</b> Towards creating a dataset based on events classified by a network intrusion detection system	14
<b>Носов В.О., Островська К.Ю.</b> Автоматизований конвеєр формування датасету для навчання моделей виявлення шахрайства	36	<b>Nosov V.O., Ostrovska K.Y.</b> Automated pipeline for building a fraud detection training dataset	36
<b>Пристайло М.О., Балака М.М., Поліщук А.Г., Гонта І.П.</b> Експериментальні дослідження процесу різання ґрунту наконечни- ком з динамічною ріжучою кромкою	48	<b>Prystailo M.O., Balaka M.M., Polishchuk A.G., Honta I.P.</b> Experimental study of soil cutting with dynamic cutting edge tip	48
<b>Кожанов А.Є., Євланов М.В., Мороз Д.М.</b> Вдосконалення методу вирішення задачі ранньої ідентифікації конфігураційних елементів інформаційної системи	59	<b>Kozhanov A.Ye., Ievlanov M.V., Moroz D.M.</b> Improvement of the method for solving the task of early identification of information system configuration items	59
<b>Вельма А.О., Макаревич М.П., Голян Н.В.</b> Порівняльний аналіз моделей управління станом в мобільних застосунках, розроблених за допомогою Flutter	73	<b>Velma A.O., Makarevych M.P., Golian N.V.</b> Comparative analysis of state management models in mobile applications developed using Flutter	73
<b>Рудик О.Ю., Харжевський В.О.</b> Використання систем автоматизова- ного проєктування для моделювання складних систем та технологічних процесів відновлення та зміцнення деталей	83	<b>Rudyk O.Yu., Kharzhevskiy V.O.</b> The usage of computer-aided design systems for modeling complex systems and technological processes for the restoration and reinforcement of parts	83

<b>Слободзян М.В., Шкурят О.С.</b> Програмний метод сегментації зображень в умовах варіативного освітлення	96	<b>Slobodzian M.V., Shkurat O.S.</b> Software method for image segmentation under variable lighting conditions	96
<b>Жарков А.В., Маслій Р.В.</b> Візуально-інерційний SLAM з використанням інваріантного розширеного фільтру Калмана для автономної навігації	107	<b>Zharkov A., Maslii R.</b> Visual-inertial SLAM using invariant extended Kalman filter for autonomous navigation	107
<b>Клименко І.В., Лебідь Є.А.</b> Огляд сучасних фреймворків та метрик оцінки RAG-систем	122	<b>Klymenko I.V., Lebid Y.A.</b> A comprehensive survey of modern frameworks and evaluation metrics for RAG systems	122
<b>Ленський М.М., Михальчук Г.Й.</b> Реалізація та дослідження алгоритму branch-cut-and-price для задачі маршрутизації транспортних засобів з обмеженнями на вантажопідйомність	136	<b>Lenskyi M.M., Mykhalchuk H.Y.</b> Implementation and study of a branch-cut-and-price algorithm for the capacitated vehicle routing problem	136
<b>Гладуш А.С., Клименко С.В., Мазуренко В.Б.</b> Автоматизація прийняття рішень у банківській сфері на основі аналізу даних великого обсягу та генеративного штучного інтелекту: огляд сучасних підходів	147	<b>Hladush A.S., Klymenko S.V., Mazurenko V.B.</b> Automation of decision-making in the banking sector based on big data analysis and generative artificial intelligence: a review of modern approaches	147
<b>Кондрашов Д.А., Музыка І.О., Купін А.І., Балик Д.К.</b> Класифікація просторових станів об'єктів на основі карт висот із використанням згорткових нейронних мереж для систем сенсорного сортування руд	154	<b>Kondrashov D.A., Muzyka I.O., Kupin A.I., Balyk D.K.</b> Classification of spatial object states based on height maps using convolutional neural networks for sensor-based ore sorting systems	154
<b>Ніколаєв І., Олійник Ю.</b> Метод та програмний засіб розпізна- вання та аналізу об'єктів на відео	166	<b>Nikolaiev I., Oliinyk Y.</b> A method and software tool for video object recognition and analysis	166

<p><b>Панасенко Є.С.</b> Класифікація стану очей на основі декомпозиції електроенцефалограм з урахуванням дипольної природи сигналів мозку</p>	178	<p><b>Panasenko Ye.S.</b> Eye state classification based on electroencephalogram decomposition considering the dipole nature of brain signals</p>	178
<p><b>Лозовський А.В., Шведов В.О.</b> Методи та алгоритми розв'язання динамічних задач оптимального розбиття множин із розміщенням центрів підмножин та інтегральними обмеженнями</p>	199	<p><b>Lozovskyi A., Shvedov V.</b> Methods and algorithms for solving dynamic optimal set partitioning problems with subset center placement and integral constraints</p>	199
<p><b>Мисюк Р.В., Кулик П.Р., Гусак О.В., Чмихало О.С., Парубочий В.О.</b> Метод оцінювання стану елементу інженерної конструкції з викорис- танням пояснювального штучного інтелекту та мультимодальних даних</p>	209	<p><b>Mysiuk R.V., Kulyk P.R., Husak O.V., Chmykhalo O.S., Parubochyi V.O.</b> Method for assessing the condition of an engineering structure element using explanatory artificial intelligence and multimodal data</p>	209
<p><b>Павлюшин М.Ю., Саяпіна І.О.</b> Метод та програмне забезпечення для оптимізації продуктивності баз даних</p>	227	<p><b>Pavliushyn M.Y., Saiapina I.O.</b> Method and software for database performance optimization</p>	227
<p><b>Палагута К.О.</b> Дослідження ефективності ансамблевих алгоритмів машинного навчання в задачах фінансового прогнозування</p>	235	<p><b>Palahuta K.O.</b> Study of the effectiveness of ensemble machine learning algorithms in financial forecasting tasks</p>	235
<p><b>Тугова О.В., Пінчук В.О.</b> Керування утворенням сірко- та азотовмісних компонентів як основа екологічної оптимізації газифікації вугілля</p>	246	<p><b>Tutova O.V., Pinchuk V.O.</b> Control of sulfur- and nitrogen- containing component formation as a basis for ecological optimization of coal gasification</p>	246
<p><b>Ющенко А.С., Смеляков К.С., Чуприна А.С.</b> Сучасні тенденції розвитку обробки природної мови: еволюція, архітектури та сфери застосування</p>	259	<p><b>Yushchenko A.S., Smelyakov K.S., Chupryna A.S.</b> Modern trends in natural language processing: evolution, architectures, and application domains</p>	259

<b>Панфілов О.І., Нестеренко М.М.</b> Динаміка системи «вібраційна установка – шар бетону – імпульсний привантажувач» з урахуванням дії основного вібробуджувача та вищих гармонік	272	<b>Panfilov O.I., Nesterenko M.M.</b> Dynamics of the system «vibration installation – concrete layer – impulse surcharge device» with account of the effect of the main vibration exciter and higher harmonics	272
<b>Іванова Л.Х., Селівьорстов В.Ю., Хитько О.Ю., Білий О.П.</b> Системний підхід до встановлення впливу рідкісноземельних металів на структуру валкових чавунів	280	<b>Ivanova L., Selivorstov V., Khytko O., Bilyi O.</b> System approach to determining the effect of rare-earth metals on the structure of roll cast irons	280
<b>Царик В.Ю., Гнатушенко Вік.В., Ольшанский Є.В.</b> Підхід розпізнавання елементів GUI як зображень на основі машинного навчання	294	<b>Tsaryk V.Yu., Hnatushenko Vik.V., Olshanskyi E.V.</b> An approach to recognizing GUI elements as images based on machine learning	294
<b>Рубан С.А., Швець Д.В., Котов І.А., Тронь В.В., Карабут Н.О.</b> Розробка архітектури інформаційно-керуючої системи для енергоефективної переробки залізорудної сировини	309	<b>Ruban S.A., Shvets D.V., Kotov I.A., Tron V.V., Karabut N.O.</b> Development of an information and control system architecture for energy-efficient iron ore proces	309
<b>Аджамський С.В., Кононенко Г.А., Подольський Р.В., Бадюк С.І., Подольська О.А.</b> Формування мікроструктури нікелевого суперсплаву ABD900 після LPBF та його порівняння зі сплавом Inconel 718	323	<b>Adzhamsky S.V., Kononenko G.A., Podolskyi R.V., Badyuk S.I., Podolska O.A.</b> Microstructure formation of the nickel-based superalloy ABD900 after LPBF and its comparison with Inconel 718	323
<b>Дмитрієва І.С.</b> Використання хмарних технологій у навчальному процесі	331	<b>Dmytriieva I.</b> The use of cloud technologies in the teaching process	331

РЕФЕРАТИ

УДК 004.942

Шиян С.І., Калініна І.О. **Імітаційна модель системи вимірювання фізико-хімічних параметрів пмм на основі колірних мереж Петрі** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.3 – 13.

В статті розглядається побудова моделі системи вимірювання рівня паливно-мастильних матеріалів (ПММ) на основі колірних мереж Петрі. Проведено порівняльний аналіз методів та систем вимірювання рівня палива. Наведено опис вимірювальної системи на основі магнітострикційного методу та алгоритм її роботи. Система складається з 4 входів для сигналів різних типів, що надходять від датчиків, двох модулів обробки сигналів та модуля індикації. Модель системи вимірювання рівня палива побудована в середовищі CPN Tools. Наведено детальний опис структури моделі. Проведено імітаційне моделювання. Були досліджені часові та структурні параметри моделі. Представлено результати статистичної обробки параметрів моделі.

Бібл. 20, іл. 3, табл. 1.

УДК 004.051

Горбатов В.С., Журба А.О. **Створення датасету на основі подій класифікованих системою виявлення мережевих вторгнень** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.14 – 35.

У статті представлено підхід до формування датасету для навчання моделей машинного навчання в контексті мережевої системи виявлення вторгнень Snort 3. Датасет сформовано для Fast Pattern-групи SIP/2.0 і містить десятки тисяч подій зі стандартизованим поділом на train/validation/test. Виконано аналіз інформативності байтових позицій (на основі дивергенції Jensen–Shannon та ентропії) і кореляційний аналіз телеметрії, що підтверджує наявність локалізованого дискримінативного сигналу та відсутність тривіальних витоків через padding. Отриманий датасет може слугувати основою для нейромережевих моделей реального часу, які доповнюють сигнатурну детекцію оцінкою ризику.

Бібл. 33, іл. 5, табл. 2.

УДК 004.85:004.62

Носов В.О., Островська К.Ю. **Автоматизований конвеєр формування датасету для навчання моделей виявлення шахрайства** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.36 – 47.

Розглянуто проблему підготовки тренувальних даних для систем виявлення шахрайства в транзакціях електронної комерції. Запропоновано автоматизований конвеєр об'єднання трьох відкритих наборів даних з платформи Kaggle зі збереженням реальних міток шахрайства та збагаченням записів синтетичними атрибутами, адаптованими до специфіки українського платіжного ринку. Опрацьовано методи квантильної нормалізації часових міток, генерації автентифікаційних даних, формування агрегова-

них клієнтських профілів та пар «entity–IP». Результатом роботи є набір із 500000 транзакцій з рівнем шахрайства 3.04%.

Бібл. 14, іл. 3, табл. 4.

УДК 621.878.23:681.3

Пристаїло М.О., Балака М.М., Поліщук А.Г., Гонта І.П. **Експериментальні дослідження процесу різання ґрунту наконечником з динамічною ріжучою кромкою** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.48 – 58.

У статті наведено результати експериментальних досліджень процесу різання ґрунту наконечником з динамічною ріжучою кромкою з використанням пневмоаккумулятора для підвищення енергоефективності. Мета роботи полягає у визначенні сили різання та встановленні її залежності від тиску. Застосовано фізичне моделювання на парафіновому середовищі, тензOMETричні вимірювання та математичну обробку даних. Встановлено U-подібну нелінійну залежність сили різання від тиску та визначено оптимальний режим роботи системи. Порівняння з теоретичною моделлю показало узгодженість результатів із похибкою до 11 %, що підтверджує її адекватність. Отримані результати можуть бути використані для удосконалення та оптимізації робочих органів будівельних машин.

Бібл. 8, іл. 5.

УДК 004.02:004.91

Кожанов А.Є., Євланов М.В., Мороз Д.М. **Вдосконалення методу вирішення задачі ранньої ідентифікації конфігураційних елементів інформаційної системи** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.59 – 72.

Розглянуто особливості вирішення задачі ранньої ідентифікації конфігураційних елементів (CI) інформаційної системи (IS). Розглянуто особливості використання методів та алгоритмів кластеризації в процесі вирішення цієї задачі. Запропоновано вдосконалення методу вирішення задачі ранньої ідентифікації CI IS. Суть вдосконалення полягає у використанні алгоритму кластеризації DBSCAN для виявлення архітектурних елементів як кластерів окремих функцій IS. Використання вдосконаленого методу дозволяє автоматизувати процедуру синтезу опису архітектури IS. Застосування цієї автоматизованої процедури сприяє підвищенню якості розробки IS за рахунок виділення для проектування множини архітектурних сутностей цієї системи, яка є значно меншою, ніж множина елементарних функцій IS.

Бібл. 15, іл. 0, табл. 2.

УДК 004.415

Вельма А.О., Макаревич М.П., Голян Н.В. **Порівняльний аналіз моделей управління станом в мобільних застосунках, розроблених за допомогою Flutter** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.73 – 82.

У статті проведено порівняльний аналіз та формалізацію моделей управління станом у мобільних застосунках, розроблених із використанням Flutter. Розглянуто підхо-

ди Provider, Riverpod, Cubit і MobX з точки зору архітектурних особливостей, продуктивності та масштабованості. Запропоновано інтерпретацію цих підходів як розширених систем переходів з асинхронними подіями. Визначено ключові відмінності між рішеннями за показниками детермінованості та ефективності, що дозволяє обґрунтовано обирати архітектуру залежно від вимог застосування.

Бібл. 13, табл. 1.

УДК 378.14:629.331

Рудик О.Ю., Харжевський В.О. **Використання систем автоматизованого проєктування для моделювання складних систем та технологічних процесів відновлення та зміцнення деталей** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.83 – 95.

Визначені дисципліни, на яких базується викладання дисципліни «САПР ТП». Але якщо на перших курсах спостерігається наступність викладання, то найбільші проблеми виникають на 1-му курсі магістратури у зв'язку з відсутністю офіційної САПР ТП зміцнення та відновлення, яку можна було б придбати для вивчення і яка б використовувалась на підприємствах відновлення та технічного сервісу автомобілів. Тому перед викладачами стоїть задача готувати фахівців, здатних самостійно розробити БД для інформаційно-закінченого програмного комплексу з метою вирішення задачі зміцнення та відновлення деталей автомобільного транспорту.

Бібл. 10, іл. 11.

УДК 004.42:004.932

Слободзян М.В., Шкурят О.С. **Програмний метод сегментації зображень в умовах варіативного освітлення** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.96 – 106.

Розглядається проблема сегментації зображень в умовах варіативного освітлення. Проаналізовано вплив змін освітленості на якість виділення об'єктів та обмеження існуючих підходів. Запропоновано програмний метод сегментації, що базується на класифікації рівня освітлення, адаптивній попередній обробці зображень та нейромережній сегментації. Метод передбачає використання статистичних характеристик відносної яскравості для вибору оптимального алгоритму оброблення зображення. Проведено експериментальне дослідження із використанням моделей U-Net та SegNet. Доведено, що запропонований метод забезпечує підвищення точності сегментації та стабільності навчання моделей у складних умовах освітлення порівняно з базовими підходами.

Бібл. 13, іл. 2, табл. 1.

УДК 004.896

Жарков А.В., Маслій Р.В. **Візуально-інерційний SLAM з використанням інваріантного розширеного фільтру Калмана для автономної навігації** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.107 – 121.

Дослідження присвячено розвитку та застосуванню візуально-інерційного SLAM на основі ітерованого розширеного фільтра Калмана (IEKF) для задач автономної навігації мобільних платформ. Класичні підходи на базі розширеного фільтра Калмана (EKF) мають суттєві обмеження: помилки однократної лінеаризації у необоротних станах призводять до явища хибної спостережуваності. Це викликає штучне звуження коваріаційних матриць та експоненційне накопичення дрейфу під час складних маневрів. Для вирішення цих проблем у роботі математично обґрунтовано та реалізовано комплексний підхід, що поєднує параметризацію оберненої глибини (IDP) для миттєвої ініціалізації орієнтирів та апарат теорії матричних груп Лі SE2(3).

Бібл. 16, іл. 3, табл. 1.

УДК 004.89; 004.65

Клименко І.В., Лебідь Є.А. **Огляд сучасних фреймворків та метрик оцінки RAG-систем** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.122 – 135.

Розглянуто сучасні метрики та фреймворки оцінки RAG-систем, включаючи класичні та LLM-орієнтовані метрики. Проаналізовано підходи з наявною еталонною множиною відповідей та без неї, їх переваги та недоліки. Проведено експеримент на системі з підтримкою векторного пошуку і доведено, що попередня фільтрація за метаданими суттєво покращує якість вибірки і генерації відповіді навіть при збільшенні контексту. Результати підкреслюють потребу в комплексній оцінці метрик вибірки та генерації та їх узгодженні.

Бібл. 16, іл. 1, табл. 7.

УДК 519.854

Ленський М.М., Михальчук Г.Й. **Реалізація та дослідження алгоритму branch-cut-and-price для задачі маршрутизації транспортних засобів з обмеженнями на вантажопідйомність** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.136 – 146.

У статті розглянуто реалізацію алгоритму branch-cut-and-price для точного розв'язання задачі маршрутизації транспортних засобів з обмеженнями на вантажопідйомність. Реалізація поєднує модель розбиття множини, генерацію стовпців, стартову евристику на основі алгоритму Кларка-Райта, округлені ємнісні нерівності та точне розв'язання підзадачі маршруту. Для зменшення часу обчислень використовується швидший евристичний пошук нових маршрутів із подальшою точною перевіркою перед завершенням генерації стовпців. На одинадцяти стандартних екземплярах розмірності від 16 до 60 вершин включно з депо реалізація в усіх випадках відтворила відомі оптимуми.

Бібл. 8, табл. 1.

УДК 004.89:004.6:004.4:336.71

Гладуш А.С., Клименко С.В., Мазуренко В.Б. **Автоматизація прийняття рішень у банківській сфері на основі аналізу даних великого обсягу та генеративного штучного інтелекту: огляд сучасних підходів** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.147 – 153.

У роботі здійснено систематичний огляд сучасних методів та інструментів автоматизації прийняття рішень у банківській сфері з використанням технологій генеративного штучного інтелекту. Проаналізовано можливості застосування великих мовних моделей для обробки неструктурованої інформації та автоматизації експертних висновків. Обґрунтовано доцільність впровадження багаторівневої архітектури, що включає етапи автоматизованого збору даних, семантичного стиснення та класифікації за допомогою векторних представлень. Визначено ключові ризики впровадження, такі як «галюцинації» моделей, конфіденційність даних та автоматизаційне упередження, а також запропоновано механізми їх мінімізації ("Людина у контурі", "Тіньовий режим"). Окреслено напрямки подальших досліджень у галузі верифікації категорій торгових клієнтів (Merchant Code Category) відповідно до вимог міжнародних платіжних систем.

Бібл. 9, іл. 0, табл. 0.

УДК 004.8:622.7

Кондрашов Д.А., Музика І.О., Купін А.І., Балик Д.К. **Класифікація просторових станів об'єктів на основі карт висот із використанням згорткових нейронних мереж для систем сенсорного сортування руд** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.154 – 165.

Розглядається задача автоматичної класифікації просторових станів об'єктів у системах сенсорного сортування рудної сировини. Запропоновано підхід на основі аналізу карт висот, сформованих із тривимірних точкових хмар лазерного сканування, із застосуванням згорткових нейронних мереж. Проведено порівняльне дослідження архітектур MLP, базової CNN, CNN із регуляризацією та поглибленої CNN, а також проаналізовано вплив розмірності вхідних даних (32×32, 64×64, 128×128) на якість класифікації. Експерименти на синтетичних і реальних даних показали, що найкращі результати досягаються при використанні поглибленої CNN, яка забезпечує точність до 92% на валідаційній вибірці для розмірності 64×64. Виявлено ефект перенавчання при збільшенні розмірності вхідних даних до 128×128. Показано ефективність аугментації даних та визначено складність розмежування класів об'єднаних і накладених об'єктів.

Бібл. 8, іл. 6, табл. 0.

УДК 004.93

Ніколаєв І., Олійник Ю. **Метод та програмний засіб розпізнавання та аналізу об'єктів на відео** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.166 – 177.

У статті розроблено метод та масштабований програмний засіб для автоматизованої екстракції, семантичного збагачення та багатовимірного аналізу відеоданих. Метод поєднує швидку детекцію об'єктів за допомогою YOLOv26m з глибоким аналізом сцени мультимодальною моделлю Moondream2. Отримані результати структурують у ROLAP-модель, що підвищує ефективність аналізу та усуває надлишковість текстової інформації. Мікросервісна архітектура з асинхронним зв'язком забезпечує лінійне прискорення при масштабуванні, скорочуючи загальний час обробки масиву відео майже.

Бібл. 6, іл.3, табл. 3.

УДК 519.2:004.9

Панасенко Є.С. **Класифікація стану очей на основі декомпозиції електроенцефалограм з урахуванням дипольної природи сигналів мозку** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.178 – 198.

Запропоновано підхід до класифікації стану очей (відкриті/закриті) за 64-канальними ЕЕГ-даними на основі ІСА-декомпозиції з валідацією компонент підбором еквівалентних диполів. Для підбору диполя розроблено оригінальну функцію втрат із L2-регуляризацією та еліпсоїдним анатомічним обмеженням на основі тривимірної моделі мозку. Серед ІСА-компонент ідентифіковано джерела з характерними альфа-, мю-, бета- та гамма-ритмами, що підтверджує фізіологічну інтерпретованість методу. Класифікація нейронною мережею на основі 507 ознак досягла точності 97% на навчальній та 90% на тестовій вибірці.

Бібл. 23, іл. 9

УДК 519.8

Лозовський А.В., Шведов В.О. **Методи та алгоритми розв'язання динамічних задач оптимального розбиття множин із розміщенням центрів підмножин та інтегральними обмеженнями** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.199 – 208.

У статті розглядаються методи та алгоритми розв'язання динамічних задач оптимального розбиття множин із розміщенням центрів підмножин та інтегральними обмеженнями. Досліджуваний клас задач виникає у багатьох прикладних контекстах, зокрема у логістиці, задачах просторового розподілу ресурсів і територіального зонування, де параметри моделі змінюються з часом.

Бібл. 16.

УДК 004.93:004.85

Мисюк Р.В., Кулик П.Р., Гусак О.В., Чмихало О.С., Парубочий В.О. **Метод оцінювання стану елемента інженерної конструкції з використанням пояснювального штучного інтелекту та мультимодальних даних** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.209 – 226.

Розглядаються підходи до оцінювання технічного стану елементів інженерних конструкцій із застосуванням методів машинного навчання та пояснювального штучного інтелекту. Запропоновано метод, що використовує мультимодальні дані, включаючи частку дефектної поверхні, температуру та вологість, для класифікації стану за трьома категоріями: «Норма», «Попередження» та «Критичний». Для побудови моделей застосовано алгоритми Random Forest, XGBoost, LightGBM, SVM та Logistic Regression із гіперпараметричною оптимізацією та перехресною валідацією, а SHAP-аналіз та Permutation Importance забезпечують глобальну та локальну інтерпретацію результатів. Показано, що використання пояснювального ШІ підвищує прозорість рішень і дозволяє ідентифікувати ключові фактори, що впливають на технічний стан конструкцій.

Бібл. 13, іл. 5, табл. 3.

УДК 004.672

Павлюшин М.Ю., Саяпіна І.О. **Метод та програмне забезпечення для оптимізації продуктивності баз даних** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.227 – 234.

Розглядається проблема комплексного автоматизованого аналізу баз даних без прямого підключення до сервера. Запропоновано комбінований алгоритмічно-програмний метод, що інтегрує чотири незалежних аналітичних модулів: аналіз індексації, структурний аналіз, рефакторинг SQL-запитів та аналіз конфігураційних параметрів. Введено комплексний індекс продуктивності (КП) для кількісної оцінки ефективності оптимізації. Метод реалізовано у вигляді клієнт-серверного веб-застосунку. Показано можливість застосування запропонованого методу для підвищення продуктивності баз даних без розгортання живого сервера.

Бібл. 10, іл. 1, табл. 1.

УДК 004.8:336.7

Палагута К.О. **Дослідження ефективності ансамблевих алгоритмів машинного навчання в задачах фінансового прогнозування** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.235 – 245.

Досліджено ефективність ансамблевих алгоритмів машинного навчання для прогнозування фінансових показників підприємства. Розглянуто моделі Random Forest, Gradient Boosting, AdaBoost, Extra Trees та XGBoost і проведено їх порівняльний аналіз із базовою регресійною моделлю. Експериментальні результати показали підвищення точності прогнозування при використанні ансамблевих методів, зокрема XGBoost, що підтверджує доцільність їх застосування у фінансових аналітичних системах.

Бібл. 17, іл. 2, табл. 2.

УДК 662.764

Тутова О.В., Пінчук В.О. **Керування утворенням сірко- та азотовмісних компонентів як основа екологічної оптимізації газифікації вугілля** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.246 – 258.

Досліджено екологічні показники процесу газифікації вугілля при одночасному варіюванні керуючих параметрів (температури, вмісту кисню в окиснику та коефіцієнта витрати окисника) із встановленням їх впливу на закономірності утворення сірко- та азотовмісних компонентів у генераторному газі. Показано, що зміна цих параметрів визначає окислювально-відновні умови процесу та забезпечує керований перерозподіл сірки між відновними й окислювальними формами, а також впливає на утворення азотовмісних сполук. Обґрунтовано можливість цілеспрямованого керування складом генераторного газу як інструменту мінімізації екологічно небезпечних компонентів.

Бібл. 21, іл. 3, табл. 2.

УДК 004.8

Ющенко А.С., Смеляков К.С., Чуприна А.С. **Сучасні тенденції розвитку обробки природної мови: еволюція, архітектури та сфери застосування** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.259 – 271.

Досліджено еволюцію методів обробки природної мови від статистичних моделей до сучасних трансформерів та великих моделей міркування. Проаналізовано інноваційні нейроеволюційні методи оптимізації архітектур (NAS). Систематизовано сфери застосування NLP: охорона здоров'я, автоматизація досліджень, бізнес-аналітика. Виділено ключові проблеми, зокрема обчислювальну складність, галуцинації та алгоритмічну упередженість. Запропоновано напрямки розвитку, що включають гібридні архітектури, системи з доповненою генерацією (RAG) та інтеграцію з квантовими обчисленнями.

Бібл. 34, іл. 4, табл. 3.

УДК 666.9.033

Панфілов О.І., Нестеренко М.М. **Динаміка системи «вібраційна установка – шар бетону – імпульсний привантажувач» з урахуванням дії основного віброзбуджувача та вищих гармонік** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.272 – 279.

У статті розглянуто динаміку системи «вібраційна установка – шар бетону – імпульсний привантажувач» з урахуванням дії основного віброзбуджувача та вищих гармонік коливального процесу. Побудовано узагальнену математичну модель, яка враховує взаємодію рухомої рами з формою, бетонної суміші та імпульсного привантажувача, а також контактну і безконтактну фази його руху. Встановлено, що істотний вплив на параметри доущільнення мають не лише імпульсні дії, а й вищі гармоніки основного збудження, які змінюють амплітудно-фазові характеристики руху привантажувача та інтенсивність ущільнення суміші. Одержані результати можуть бути використані для вибору раціональних параметрів імпульсного привантажувача, оцінювання стійкості режимів роботи та вдосконалення вібраційних установок для формування бетонних виробів.

Бібл. 9, рис. 2

УДК 621.74:669.131.7

Іванова Л.Х., Селівьорстов В.Ю., Хитько О.Ю., Білий О.П. **Системний підхід до встановлення впливу рідкісноземельних металів на структуру валкових чавунів** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.280 – 293.

У статті досліджено вплив рідкісноземельних металів на процеси структуроутворення валкових чавунів. Встановлено закономірності зміни мікроструктури залежно від виду та залишкового вмісту РЗМ церієвої та ітрієвої підгруп. Показано, що модифікування забезпечує зміну типу карбідної евтектики, підвищення дисперсності перліту та формування мартенситних структур. Визначено критичні концентрації РЗМ, що забезпечують максимальну карбідоутворюючу здатність. Встановлено, що РЗМ ітрієвої

підгрупи мають у 5,2 рази вищу ефективність порівняно з церієвою. Отримані результати дозволяють обґрунтувати склад комплексних модифікаторів для підвищення експлуатаційних властивостей валків.

Бібл. 14, іл. 2, табл. 2.

УДК 004.932

Царик В.Ю., Гнатушенко Вік.В., Ольшанский Є.В. **Підхід розпізнавання елементів GUI як зображень на основі машинного навчання** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.294 – 308.

Розглянуто підхід, що передбачає використання елементів веб-інтерфейсу як вхідних даних для моделей детекції об'єктів як зображень. Запропоновано процес, у межах якого здійснюється отримання зображення за допомогою Selenium, подальша обробка з використанням моделі машинного навчання та передача результатів у форматі JSON для тестування в середовищі PyTest. Проведено порівняльний аналіз моделей Swin-L, EfficientDet-D7, Cascade Mask R-CNN, RetinaNet та YOLO з точки зору їх придатності до задачі розпізнавання елементів інтерфейсу.

Бібл. 24, іл. 3.

УДК 004.415.2:681.518:622.7:622.341

Рубан С.А., Швець Д.В., Котов І.А., Тронь В.В., Карабут Н.О. **Розробка архітектури інформаційно-керуючої системи для енергоефективної переробки залізорудної сировини** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.309 – 322.

Розроблено багаторівневу архітектуру інформаційно-керуючої системи для енергоефективної переробки залізорудної сировини на базі моделі RAMI 4.0. На польовому рівні інтегровано засоби оперативного контролю через цифрові оболонки активів. Впроваджено Edge-рівень для агрегації та локальної аналітики даних, що критично зменшує навантаження на мережі. Хмарний рівень використовує цифрові двійники та алгоритми модельно-прогнозного керування для оптимізації енергоспоживання. Кібербезпеку забезпечено просторовим розподілом на зони та апаратними шлюзами. Застосування архітектури дозволяє знизити питомі витрати електроенергії на 8–15% та підвищити вміст заліза у концентраті на 0,5–1,5%.

Бібл. 15, іл. 1, табл. 2.

УДК 621.74:621.762:669.245

Аджамський С.В., Кононенко Г.А., Подольський Р.В., Бадюк С.І., Подольська О.А. **Формування мікроструктури нікелевого суперсплаву ABD900 після LPBF та його порівняння зі сплавом Inconel 718** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.323 – 330.

У роботі досліджено формування мікроструктури нікелевого суперсплаву ABD900, виготовленого методом лазерного плавлення в порошковому шарі (LPBF), та проведено його порівняння зі сплавом Inconel 718. Дослідні зразки виготовлено за ідентичних технологічних параметрів процесу. Встановлено, що відносна щільність зразків

обох матеріалів становить близько 99,8 %. Проведений аналіз пористості показав відмінності у розподілі та розмірах пор між досліджуваними сплавами. Мікроструктурний аналіз виявив формування характерної для LPBF лускатої структури треків плавлення та клітинно-дендритної морфології кристалізації. Отримані результати демонструють перспективність використання сплаву ABD900 для адитивного виробництва та потребу у подальших дослідженнях його мікроструктурних особливостей.

Бібл. 5, іл. 3, табл. 1

УДК 004:378

Дмитрієва І.С. **Використання хмарних технологій у навчальному процесі** // Системні технології. Регіональний міжвузівський збірник наукових праць. - Випуск 3(164). – Дніпро, 2026. – С.331 – 339.

Дослідження аналізує ефективність та перспективи таких технологій у формуванні професійних ІТ-навичок. Хмарні платформи використовуються для програмування, аналізу даних та розробки ПЗ, відкриваючи доступ до віртуальних лабораторій і потужних обчислювальних ресурсів. Хмарні сервіси сприяють розвитку цифрових компетенцій, підвищують мотивацію й самостійність студентів, підтримують групову роботу та ефективну комунікацію. Застосування таких рішень покращує якість освіти та підвищує конкурентоспроможність випускників ІТ-сфери.

Бібл. 5.

UDC 004.942

Shiyan S., Kalinina I. **Simulation model of the system for measuring physical and chemical parameters of fuel and lubricants based on coloured Petri networks** // System technologies. № 3(164) - Dnipro, 2026.- P.3 – 13.

The article considers the construction of a model of a fuel and lubricant level measurement system based on colour Petri nets. A comparative analysis of fuel level measurement methods and systems is carried out. A description of a measuring system based on the magneto strictive method and its operation algorithm are given. The system consists of 4 inputs for signals of different types coming from sensors, two signal processing modules and an indication module. The model of the fuel level measurement system is built in the CPN Tools environment. A detailed description of the model structure is given. Simulation modelling was carried out. The time and structural parameters of the model were investigated. The results of statistical processing of the model parameters are presented.

Ref. 20, ill. 3, table 1.

UDC 004.051

Gorbatov V.S., Zhurba A.O. **Towards creating a dataset based on events classified by a network intrusion detection system** // System technologies. № 3(164) - Dnipro, 2026.- P.14 – 35.

The article presents an approach to dataset formation for training machine learning models within the context of the Snort 3 network intrusion detection system. The dataset was generated for the SIP/2.0 Fast Pattern group and contains tens of thousands of events with a standardized train/validation/test split. An analysis of the informativeness of byte positions (based on Jensen–Shannon divergence and entropy) and a correlation analysis of the telemetry were performed, confirming the presence of a localized discriminative signal and the absence of trivial leaks through padding. The resulting dataset can serve as a foundation for real-time neural network models that complement signature-based detection with risk assessment.

Ref. 33, fig. 5, table 2.

UDC 004.85:004.62

Nosov V.O., Ostrovska K.Y. **Automated pipeline for building a fraud detection training dataset** // System technologies. № 3(164) - Dnipro, 2026.- P.36 – 47.

The paper addresses the problem of preparing training data for machine learning-based fraud detection systems in e-commerce transactions. An automated pipeline is proposed for integrating three open Kaggle datasets while preserving authentic fraud labels and enriching records with synthetic attributes adapted to the Ukrainian payment market. The study implements quantile-based timestamp normalization, authentication data generation, and the formation of aggregated client profiles and “entity–IP” pairs. The result is a comprehensive set of 500000 transactions over 24 months with a 3.04% fraud rate.

Ref. 14, fig. 3, tab. 4.

UDC 621.878.23:681.3

Prystailo M.O., Balaka M.M., Polishchuk A.G., Honta I.P. **Experimental study of soil cutting with dynamic cutting edge tip** // System technologies. № 3(164) - Dnipro, 2026.- P.48 – 58.

The experimental study of soil cutting with the dynamic cutting edge tip and pneumatic accumulator are presented in the paper at improving energy efficiency. The purpose is to determine cutting force and its dependence on accumulator pressure. Physical modeling using the paraffin soil analogue, strain-gauge measurements, and mathematical data processing were applied. The nonlinear U-shaped relationship between cutting force and pressure was established, and an optimal operating range was identified. Comparison with the theoretical model showed good agreement with the deviation of up to 11%, confirming its validity. The results can be used to optimize construction machinery working elements.

Ref. 8, fig. 5.

UDC 004.02:004.91

Kozhanov A.Ye., Ievlanov M.V., Moroz D.M. **Improvement of the method for solving the task of early identification of information system configuration items** // System technologies. № 3(164) - Dnipro, 2026.- P.59 – 72.

The features of solving the task of early identification of configuration items (CI) of an information system (IS) are considered. The features of using clustering methods and algorithms in the processes of solving this task are considered. An improvement of the method for solving the task of early identification of CI of IS is proposed. The essence of the improvement is to use the DBSCAN clustering algorithm to identify architectural entities as clusters of individual IS functions. The use of the improved method allows you to automate the procedure for synthesizing the description of the IS architecture. The use of this automated procedure helps to improve the quality of IS development by allocating for design a set of architectural entities of this system, which is much smaller than the set of elementary IS functions.

Ref. 15, fig. 0, tab. 2.

UDC 004.415

Velma A.O., Makarevych M.P., Golian N.V. **Comparative analysis of state management models in mobile applications developed using Flutter** // System technologies. № 3(164) - Dnipro, 2026.- P.73 – 82.

The paper presents a comparative analysis and formalization of state management models in mobile applications developed using Flutter. The approaches Provider, Riverpod, Cubit, and MobX are examined in terms of architectural characteristics, performance, and scalability. These approaches are interpreted as extended transition systems with asynchronous events. Key differences in determinism and efficiency are identified, providing a basis for selecting an appropriate architecture depending on application requirements.

Ref. 13, tables 1.

UDC 378.14:629.331

Rudyk O.Yu., Kharzhevskiy V.O. **The usage of computer-aided design systems for modeling complex systems and technological processes for the restoration and reinforcement of parts** // System technologies. № 3(164) - Dnipro, 2026.- P.83 – 95.

The disciplines that form the basis for teaching the course “CAD for Technological Processes” have been specified. Although there is continuity in training in the first years of studying, the most significant challenges arise in the first year of the master’s program due to the lack of an official CAD system for strengthening and restoration that could be obtained for study and used at automotive restoration and technical service enterprises. Therefore, instructors must prepare specialists capable of independently developing a database for a comprehensive software package to solve the problem of strengthening and restoring of automotive parts.

Ref. 10.

UDC 004.42:004.932

Slobodzian M.V., Shkurat O.S. **Software method for image segmentation under variable lighting conditions** // System technologies. № 3(164) - Dnipro, 2026.- P.96 – 106.

The problem of image segmentation under variable lighting conditions is examined. The impact of changes in illumination on the quality of object detection and the limitations of existing approaches is analyzed. A software-based segmentation method is proposed, which is based on illumination level classification, adaptive image preprocessing, and neural network segmentation. The method involves using statistical characteristics of relative luminance to select the optimal image processing algorithm. An experimental study was conducted using U-Net and SegNet models. It was demonstrated that the proposed method improves segmentation accuracy and model training stability under complex lighting conditions compared to baseline approaches.

Bibl. 13, ill. 2, tabl. 1.

UDC 004.896

Zharkov A., Maslii R. **Visual-inertial SLAM using invariant extended Kalman filter for autonomous navigation** // System technologies. № 3(164) - Dnipro, 2026.- P.107 – 121.

The study is devoted to the development and application of visual-inertial SLAM based on the iterated extended Kalman filter (IEKF) for autonomous navigation tasks of mobile platforms. Classical approaches based on the extended Kalman filter (EKF) have significant limitations: single linearization errors in irreversible states lead to the phenomenon of false observability. This causes artificial narrowing of covariance matrices and exponential accumulation of drift during complex maneuvers. To solve these problems, the work mathematically substantiates and implements a comprehensive approach that combines the inverse depth parameterization (IDP) for instantaneous initialization of landmarks and the apparatus of the theory of SE<sub>2</sub>(3) Lie matrix groups.

Bibl. 16, ill. 3, tabl. 1.

UDC 004.89; 004.65

Klymenko I.V., Lebid Y.A. **A comprehensive survey of modern frameworks and evaluation metrics for RAG systems** // System technologies. № 3(164) - Dnipro, 2026.- P.122 – 135.

Modern metrics and frameworks for evaluating RAG systems are reviewed, including classical and LLM-oriented metrics. Approaches with and without reference answer sets are analyzed, along with their advantages and limitations. An experiment on a vector search-enabled system demonstrates that metadata pre-filtering significantly improves retrieval and generation quality even with increased context size. The results emphasize the need for comprehensive evaluation of retrieval and generation metrics, as well as their alignment.

Ref. 16, Fig. 1, Tab. 7.

UDC 519.854

Lenskyi M.M., Mykhalchuk H.Y. **Implementation and study of a branch-cut-and-price algorithm for the capacitated vehicle routing problem** // System technologies. № 3(164) - Dnipro, 2026.- P.136 – 146.

The paper presents an implementation of a branch-cut-and-price algorithm for the exact solution of the Capacitated Vehicle Routing Problem. It combines a set-partitioning formulation, column generation, a Clarke-Wright-based initial heuristic, rounded capacity inequalities, and the exact solution of the route-generation subproblem. To reduce computational time, a faster heuristic search for new routes is used before the final exact certification step. On eleven standard instances with 16 to 60 vertices, including the depot, the implementation reproduced the known optimal values in every reported case.

Refs. 8, table 1.

UDC 004.89:004.6:004.4:336.71

Hladush A.S., Klymenko S.V., Mazurenko V.B. **Automation of decision-making in the banking sector based on big data analysis and generative artificial intelligence: a review of modern approaches** // System technologies. № 3(164) - Dnipro, 2026.- P.147 – 153.

This paper provides a systematic review of modern concepts and tools for automating decision-making processes in banking using generative artificial intelligence. The possibilities of using large language models for processing unstructured information and automating expert conclusions in real-time are analyzed. The expediency of using a multi-level architecture including automated data collection, semantic compression, and classification using vector representations is substantiated. Key implementation challenges such as model "hallucinations", data privacy, and automation bias are identified, alongside mechanisms for their mitigation ("Human-in-the-loop", "Shadow Mode"). Directions for applied research in the field of merchant category codes verification according to the requirements of international payment systems are outlined.

Refs. 9.

UDC 004.8:622.7

Kondrashov D.A., Muzyka I.O., Kupin A.I., Balyk D.K. **Classification of spatial object states based on height maps using convolutional neural networks for sensor-based ore sorting systems** // System technologies. № 3(164) - Dnipro, 2026.- P.154 – 165.

The paper addresses the problem of automatic classification of spatial object states in sensor-based ore sorting systems. An approach based on height maps derived from 3D point clouds obtained via laser scanning and processed using convolutional neural networks is proposed. A comparative study of MLP, baseline CNN, regularized CNN, and deep CNN architectures is conducted, along with an analysis of the impact of input resolution (32×32, 64×64, 128×128) on classification performance. Experiments on synthetic and real datasets show that the deep CNN achieves the best results, with up to 92% validation accuracy for 64×64 inputs. Overfitting is observed when increasing input resolution to 128×128. The effectiveness of data augmentation is demonstrated, and the difficulty of distinguishing between united and overlapped object classes is identified.

Ref. 8, fig. 6, tab. 0.

UDC 004.93

Nikolaiev I., Oliinyk Y. **A method and software tool for video object recognition and analysis** // System technologies. № 3(164) - Dnipro, 2026.- P.166 – 177.

The article develops a method and scalable software tool for automated extraction, semantic enrichment, and multidimensional analysis of video data. The method combines fast object detection using YOLOv26m with deep scene analysis via the Moondream2 multimodal model. The obtained results are structured into a ROLAP model using a star schema, which increases analytical efficiency and eliminates text information redundancy. A microservice architecture with a message broker ensures linear speedup during scaling, significantly reducing the total video array processing time.

Bible 6, ill. 3, tab 3

UDC 519.2:004.9

Panasenko Ye.S. **Eye state classification based on electroencephalogram decomposition considering the dipole nature of brain signals** // System technologies. № 3(164) - Dnipro, 2026.- P.178 – 198.

An approach to eye state classification (open/closed) based on 64-channel EEG data using ICA decomposition with component validation through equivalent dipole fitting is proposed. An original loss function with L2 regularization and ellipsoidal anatomical constraint based on a three-dimensional brain model was developed for dipole fitting. Among the ICA components, sources with characteristic alpha, mu, beta, and gamma rhythms were identified, confirming the physiological interpretability of the method. Neural network classification based on 507 features achieved 97% accuracy on the training set and 90% on the test set.

Bibl. 23, fig. 9.

UDC 519.8

Lozovskyi A., Shvedov V. **Methods and algorithms for solving dynamic optimal set partitioning problems with subset center placement and integral constraints** // System technologies. № 3(164) - Dnipro, 2026.- P.199 – 208.

The paper presents an approach to solving solving dynamic optimal set partitioning problems with subset center placement and integral constraints. The considered class of problems arises in numerous applied contexts, including logistics, spatial resource allocation, and territorial zoning, where model parameters vary over time.

Bibl. 16.

UDC 004.93:004.85

Mysiuk R.V., Kulyk P.R., Husak O.V., Chmykhalo O.S., Parubochyi V.O. **Method for assessing the condition of an engineering structure element using explanatory artificial intelligence and multimodal data** // System technologies. № 3(164) - Dnipro, 2026.- P.209 – 226.

Approaches to assessing the technical condition of engineering structure elements using machine learning and explanatory artificial intelligence methods are considered. A method is proposed that uses multimodal data, including the proportion of defective surface, temperature and humidity, to classify the condition into three categories: "Normal", "Warning" and "Critical". To build models, the Random Forest, XGBoost, LightGBM, SVM and Logistic Regression algorithms with hyperparametric optimization and cross-validation were used, and SHAP analysis and Permutation Importance provide global and local interpretation of the results. It is shown that the use of explanatory AI increases the transparency of decisions and allows identifying key factors that affect the technical condition of structures.

Bibl. 13, ill. 5, table 3.

UDC 004.672

Pavliushyn M.Y., Saiapina I.O. **Method and software for database performance optimization** // System technologies. № 3(164) - Dnipro, 2026.- P.227 – 234.

This paper addresses the problem of comprehensive automated database analysis without a direct connection to the server. A combined algorithmic-programming method is proposed that integrates four independent analytical modules: indexing analysis, structural analysis, SQL query refactoring, and configuration parameter analysis. A comprehensive performance index (CPI) is introduced for the quantitative assessment of optimization effectiveness. The method is implemented as a client-server web application. The possibility of applying the proposed method to improve database performance without deploying a live server is demonstrated.

Bibl. 10, ill. 1, table 1.

UDC 004.8:336.7

Palahuta K.O. **Study of the effectiveness of ensemble machine learning algorithms in financial forecasting tasks** // System technologies. № 3(164) - Dnipro, 2026.- P.235 – 245.

The effectiveness of ensemble machine learning algorithms for forecasting a company's financial indicators is examined. The study considers the Random Forest, Gradient Boosting,  
ISSN 1562-9945 (Print)  
ISSN 2707-7977 (Online)

AdaBoost, Extra Trees, and XGBoost models and provides their comparative analysis with a baseline regression model. Experimental results demonstrate an improvement in forecasting accuracy when using ensemble methods, particularly XGBoost, which confirms the feasibility of their application in financial analytical systems.

Ref. 17, ill. 2, tabl. 2.

UDC 662.764

Tutova O.V., Pinchuk V.O. **Control of sulfur- and nitrogen-containing component formation as a basis for ecological optimization of coal gasification** // System technologies. № 3(164) - Dnipro, 2026.- P.246 – 258.

The environmental performance of coal gasification was investigated under simultaneous variation of key process parameters (temperature, oxygen content in the oxidizer, and oxidizer-to-fuel ratio), with the aim of identifying their influence on the formation patterns of sulfur- and nitrogen-containing components in producer gas. It is shown that variation of these parameters determines the redox conditions of the process and enables controlled redistribution of sulfur between reducing and oxidizing forms, as well as affects the formation of nitrogen-containing species. The possibility of targeted control of producer gas composition as a tool for minimizing environmentally hazardous components is substantiated.

Ref. 21, ill. 3, tab. 2.

UDC 004.8

Yushchenko A.S., Smelyakov K.S., Chupryna A.S. **Modern trends in natural language processing: evolution, architectures, and application domains** // System technologies. № 3(164) - Dnipro, 2026.- P.259 – 271.

The evolution of natural language processing methods from statistical models to modern transformers and large reasoning models is investigated. Innovative neuroevolutionary methods for architecture optimization (NAS) are analyzed. The application domains of NLP, including healthcare, research automation, and business analytics, are systematized. Key challenges such as computational complexity, hallucinations, and algorithmic bias are highlighted. Future directions are proposed, involving hybrid architectures, retrieval-augmented generation (RAG) systems, and integration with quantum computing.

Ref. 34, fig. 4, tabl. 3.

UDC 666.9.033

Panfilov O.I., Nesterenko M.M. **Dynamics of the system «vibration installation – concrete layer – impulse surcharge device» with account of the effect of the main vibration exciter and higher harmonics** // System technologies. № 3(164) - Dnipro, 2026.- P.272 – 279.

The article considers the dynamics of the system “vibration installation – concrete layer – impulse surcharge device,” taking into account the effect of the main vibration exciter and the higher harmonics of the oscillatory process. A generalized mathematical model has been developed that describes the interaction of the moving frame with the mould, the concrete mixture, and the impulse surcharge device, as well as the contact and non-contact phases of its motion. It has been established that the parameters of recompaction are significantly affected

not only by impulse actions but also by the higher harmonics of the main excitation, which alter the amplitude-phase characteristics of the surcharge device motion and the intensity of mixture compaction. The obtained results can be used to select rational parameters of the impulse surcharge device, assess the stability of operating modes, and improve vibration installations for forming concrete products.

Refs. 9, Fig. 2

UDC 621.74:669.131.7

Ivanova L., Selivorstov V., Khytko O., Bilyi O. **System approach to determining the effect of rare-earth metals on the structure of roll cast irons** // System technologies. № 3(164) - Dnipro, 2026.- P.280 – 293.

The article investigates the influence of rare-earth metals on the structure formation of roll cast irons. Regularities of microstructure evolution depending on the type and residual content of cerium- and yttrium-group elements are established. It is shown that modification leads to changes in carbide eutectic type, increased pearlite dispersion, and formation of martensitic structures. Critical concentrations providing maximum carbide-forming ability are determined. It is found that yttrium-group elements exhibit 5.2 times higher efficiency than cerium-group elements. The results provide a basis for selecting optimal compositions of complex modifiers to improve the performance properties of roll cast irons.

Refs: 14, Figs: 2, Tabl. 2.

UDC 004.932

Tsaryk V.Yu., Hnatushenko Vik.V., Olshanskyi E.V. **An approach to recognizing GUI elements as images based on machine learning** // System technologies. № 3(164) - Dnipro, 2026.- P.294 – 308.

An approach is considered that involves using web interface elements as input data for object detection models as images. A process is proposed, within which an image is obtained using Selenium, further processing using a machine learning model and transferring the results in JSON format for testing in the PyTest environment. A comparative analysis of the Swin-L, EfficientDet-D7, Cascade Mask R-CNN, RetinaNet and YOLO models is carried out in terms of their suitability for the task of recognizing interface elements.

Bibl. 24, ill. 3.

UDC 004.415.2:681.518:622.7:622.341

Ruban S.A., Shvets D.V., Kotov I.A., Tron V.V., Karabut N.O. **Development of an information and control system architecture for energy-efficient iron ore processing** // System technologies. № 3(164) - Dnipro, 2026.- P.309 – 322.

A multilevel information and control system architecture for energy-efficient iron ore processing is developed based on the RAMI 4.0 model. At the field level, online analysis tools are integrated via asset administration shells. An Edge level is implemented for data aggregation and local analytics, reducing network load. The cloud level uses digital twins and model predictive control algorithms to optimize energy consumption. Cybersecurity is ensured by spatial zoning and hardware data diodes. The architecture implementation reduces specific electricity consumption by 8–15% and increases the iron content in the concentrate by 0.5–1.5%.

Ref. 15, fig. 1, tab. 2.

UDC 621.74:621.762:669.245

Adzhamsky S.V., Kononenko G.A., Podolskyi R.V., Badyuk S.I., Podolska O.A.

**Microstructure formation of the nickel-based superalloy ABD900 after LPBF and its comparison with Inconel 718** // System technologies. № 3(164) - Dnipro, 2026.- P.323 – 330.

The study investigates the microstructure formation of the nickel-based superalloy ABD900 fabricated by Laser Powder Bed Fusion (LPBF) and provides a comparison with Inconel 718. The samples were produced using identical process parameters. It was established that the relative density of both materials is approximately 99.8%. The porosity analysis revealed differences in the distribution and size of pores between the investigated alloys. Microstructural analysis showed the formation of a characteristic melt track morphology typical of the LPBF process, as well as a cellular-dendritic solidification structure. The obtained results demonstrate the potential of ABD900 for additive manufacturing applications and indicate the need for further investigation of its microstructural features.

Bibl. 5, ill. 3, tabl. 1.

UDC 004:378

Dmytriieva I. **The use of cloud technologies in the teaching process** // System technologies. № 3(164) - Dnipro, 2026.- P.331 – 339.

The study analyzes the effectiveness and prospects of such technologies in the formation of professional IT skills. Cloud platforms are used for programming, data analysis and software development, opening access to virtual laboratories and powerful computing resources. Cloud services contribute to the development of digital competencies, increase student motivation and independence, support group work and effective communication. The use of such solutions improves the quality of education and increases the competitiveness of IT graduates.

Bibl. 5.

**Системні технології**  
**ЗБІРНИК НАУКОВИХ ПРАЦЬ**  
Випуск 3 (164)

Головний редактор: к.т.н., доц. Т.В. Селівьорстова

Відповідальний секретар та технічний редактор: к.т.н., доц. К.Ю. Островська

---

Здано до набору 29.04.2026. Підписано до друку 30.04.2026.

Формат 60x84 1/16. Друк - різнограф. Папір типограф.

Умов. друк арк. – 25,86. Обл.–видавн. арк. – 22,625.

Тираж 300 прим. Замовл. – 03/26

---

Український державний університет науки і технологій,  
ННІ «Дніпровський металургійний інститут»,  
кафедра Інформаційних технологій та систем: ІВК «Системні технології»  
49600, Дніпро, а/с 493

<http://journals.nmetau.edu.ua/index.php/st>

Свідоцтво про державну реєстрацію друкованого засобу масової інформації:

**Серія КВ № 8684 від 23 квітня 2004 рік**

## **Редакційна колегія**

**Селівьорстова Тетяна Віталіївна**  
(*головний редактор*)

доцент, кандидат технічних наук

**Гнатушенко Вікторія Володимирівна**  
(*заступник головного редактора*)

професор, доктор технічних наук

**Островська Катерина Юріївна**  
(*відповідальний секретар*)

доцент, кандидат технічних наук

## **Члени редакційної колегії**

**Бердник Михайло Геннадійович**  
професор, доктор технічних наук

**Гвоздева Ірина Маратівна**  
професор, доктор технічних наук

**Гнатушенко Володимир Володимирович**  
професор, доктор технічних наук

**Редчиць Дмитро Олександрович**  
професор, доктор технічних наук

**Svyetlichnyy Dmytro**  
Ph.D, Professor, Full Professor

**Heipke Christian**  
Dr.-Ing., Professor

**Шинкаренко Віктор Іванович**  
професор, доктор технічних наук

**Shults Roman**  
DSc, PhD, Associate Professor

Український державний університет науки і технологій, ННІ «Дніпровський металургійний інститут», Україна

Український державний університет науки і технологій, ННІ «Дніпровський металургійний інститут», Україна

Український державний університет науки і технологій, ННІ «Дніпровський металургійний інститут», Україна

Національний технічний університет «Дніпровська політехніка», Україна

Національний університет «Одеська морська академія», Україна

Національний технічний університет «Дніпровська політехніка», Україна

Інститут транспортних систем і технологій НАНУ, Україна

Akademia Górniczo-Hutnicza  
im. Stanisława Staszica, Poland

Leibniz University Hannover, Germany

Український державний університет науки і технологій, ННІ «Дніпровський інститут інфраструктури і транспорту», Україна

Interdisciplinary Research Center  
for Aviation and Space Exploration,  
Kingdom of Saudi Arabia