

К.Ю. Островська, А.Ю. Зимогляд, А.О. Козакова

МОДЕЛЮВАННЯ НЕЙРОМЕРЕЖЕВОГО АЛГОРИТМУ ДЛЯ АНАЛІЗУ ДЕФЕКТІВ ВІДЕОПОТОКУ

Анотація. В роботі зроблено огляд на існуючі наукові роботи з аналізу відео на виявлення дефектів та технологій цифрового мовлення. Розглядається створення датасета, розподілу даних по вибіркам. Описується навчання нейромережових моделей та отримання результатів навчання. Реалізації методу використання навченої нейронної мережової моделі.

Ключові слова: нейромережа, згорткова нейронна мережа, дата-сет, дефект, відеофайл, відеотрансляція, відеомовлення.

Вступ. Сучасні джерела інформації витіснили телебачення з графіка повсякденного життя. Зараз замість аналогового телебачення перевагу використовують цифрове. Протоколи передачі аудіо та відео в цифрових пакетах продовжують активно використовуватися не тільки в цифровому телебаченні, але і в інших сферах.

Метою роботи є моделювання нейромережового алгоритму для аналізу дефектів відеопотоку.

Опис вимог для вирішення поставленого завдання. На основі проведеного аналізу для виконання поставленого завдання необхідно виконати такі пункти.

1. Зробити набір зображень, що містять кадри з дефектами трансляції.
2. Розмітити отриманий набір зображень типи помилок, зображені на кадрі.
3. Вибрати існуючу нейромережну модель та навчити її на створеному датасеті.
4. Написати програму, яка використовуватиме навчену модель для виявлення дефектів на відео.

Для реалізації поточних вимог також необхідно використовувати додаткове ПЗ:

1. Для відтворення потокового відеотрансляції буде використано VLC media player. Цей медіаплеєр має можливість відтворення файлів потокового мовлення, програє більшість кодеків, має відкритий вихідний код. Для навчання нейронної мережі необхідно створити датасет із зображеннями дефектів відеотрансляції.

2. Для створення зображень із дефектами потрібно отримати файл відеотрансляції із пріоритетом помилок пакетів третього рівня. Для внесення в пакет помилок буде використано програмне забезпечення TSDuck.

TSDuck це open source ПЗ для маніпуляцій з TS-потоками. Як вход вміє працювати з IP, http, hls, dvb-тюнерами, dektec dvb-asi демодулятором, є внутрішній генератор TS-потоків та читання з файлів. Як вихід може бути запис у файл, IP, hls, dektec dvb-asi

та HiDes модулятори, плеєри (mplayer, vlc, xine) та drop. Між входом і виходом можна включити різні процесори трафіку.

3. Щоб отримувати технічну інформацію про файл, що містить відеотрансляцію, буде використано програму Mediainfo.

MediaInfo – це невелика безкоштовна утиліта, яка дозволяє отримувати технічну інформацію про медіафайли.

4. Для вирізання кадрів із відео використовується програма Kdenlive.

Kdenlive – це абревіатура від KDE Non-Linear Video Editor. Це доступний редактор з відкритим вихідним кодом, сумісний із пристроями Linux, Mac та Windows. Він надає стандартні для галузі параметри редагування та широкий спектр аудіо-, відео- та композиційних ефектів.

Набір використаних даних. Підбір відео для створення набору даних. Для навчання нейромережевої моделі необхідно мати датасет із зображеннями дефектів трансляції. Вирішили створити свій датасет даних використовуючи вільні джерела. Завантажено відео з вільних джерел, які транслюються цифровим телебаченням. Було завантажено відеотрансляцію цифрових телепередач наступних форматів: мультфільми, спорт, новини, наука, детективи, твори художньої культури, передачі про тварин, реклама, ситкоми та комедійні шоу. На кожен формат було завантажено близько 10 відеозаписів. У сумі було завантажено 100 відео файлів. Усі відеозаписи були конвертовані у формат потокового мовлення ".TS" з використанням програми VLCMediaPlayer.

Для створення первинного набору даних використовується програмне забезпечення TSDuck.

Експериментальним шляхом було виявлено, що для отримання максимальної кількості дефектів відеопотоку, необхідно збільшити максимальну кількість бітових помилок на пакет (max_pkt_err).

Щоб перейти до розмітки датасета, необхідно з файлів відеотрансляції вирізати кадри, що містять дефекти зображення.

Програма для відеомонтажу Kdenlive має інструментарій для обробки файлів відеотрансляції «TS». За допомогою цього програмного забезпечення було вирізано 1000 кадрів з роздільною здатністю високої чіткості (1280x720) у форматі «.png».

Зображення з низькою якістю видалено.

Розмітка первинних даних. Для розмітки зображень було використано програму Label Studio.

Для навчання нейромережевої моделі вирізаний набір кадрів з дефектами було розміщено у програмі Label Studio. Для кожного типу дефекту було визначено клас. Спосіб розмітки сегментів – Bounding boxes.

Основне завдання нейромережевої моделі – виявлення лише кадру з дефектом.

Кількість розмічених областей кожного класу. У роботі нейромережі було використано самостійно підготовлений набір даних з урахуванням файлів відеотрансляцій з дефектами кадрів.

Датасет складається з 1016 зображень: 816 зображень із дефектами та 200 зображень із звичайними кадрами відеотрансляції, які не містять дефектів (Рисунок 1).

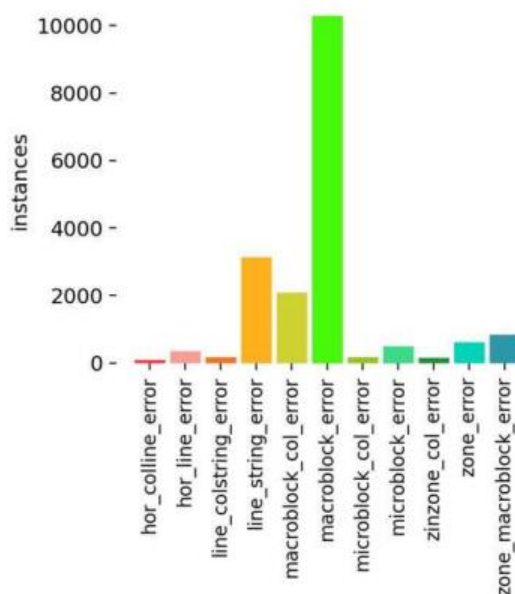


Рисунок 1 - Кількість розмічених областей кожного класу

Збір тренувальної, тестової, валідаційної вибірки. У даному випадку, щоб неймережа правильно навчилася виявляти помилки відеотрансляції, необхідно розподілити зображення на тренувальну, тестову та валідаційну вибірку так, щоб кількість зображень з певним класом приблизно задовольняла співвідношенню: 70x20x10. Після розподілу даних, вийшло: 686 зображень у тренувальній вибірці, 225 у тестовій, 105 у валідаційній (Рисунок 2).

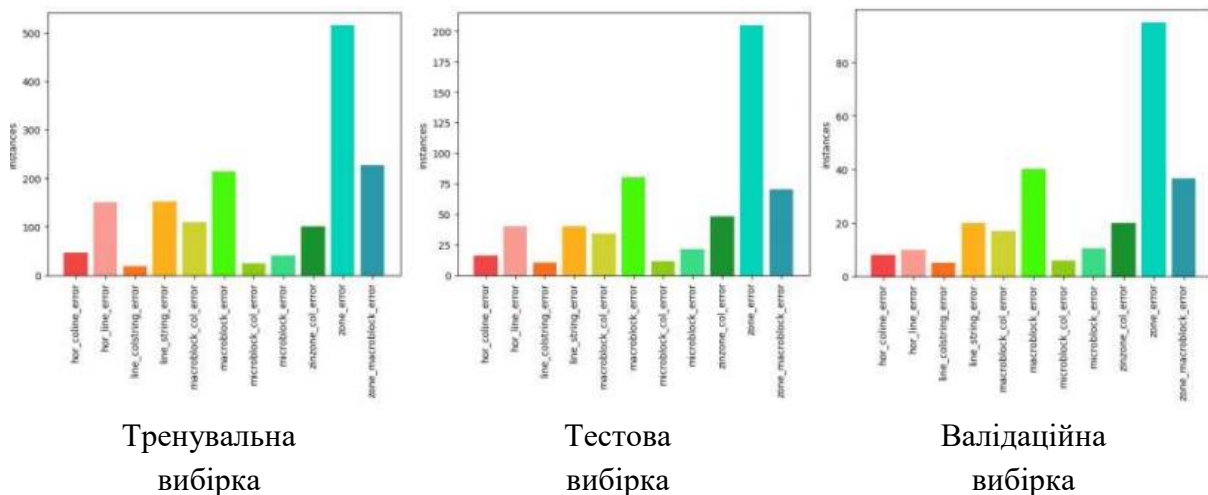


Рисунок 2 - Збір тренувальної, тестової, валідаційної вибірки

Реалізація неймережної моделі. Для навчання було обрано модель виявлення об'єктів, YOLO.

YOLO – це система (мережа) виявлення об'єктів. Моделі YOLO підходять для завдань: виявлення об'єктів, сегментації екземплярів, класифікації, орієнтованого виявлення. Найбільшою перевагою YOLO над іншими архітектурами є швидкість.

Для подальшого навчання були обрані найшвидша та найточніша модель версій YOLOv5 та YOLOv8, щоб дізнатися яка з моделей продемонструє найкращий результат. У таблиці 1 представлено кілька вибраних моделей YOLO та його показники.

Таблиця 1

Показники моделей YOLO

№	Найменування моделі YOLO	Максимальний розмір зображення	Швидкість обробки однієї фотографії (мс)	mAPval (середня точність виявлення об'єктів)
1	yolov5x6u.pt	1280	1,06	56,8
2	yolov5n6u.pt	1280	3,81	42,1
3	YOLOv8x.pt	640	3,53	53,9
4	YOLOv8n.pt	640	0,99	37,3

Щоб зрозуміти які необхідні параметри, необхідно було провести перше оцінне навчання. Була обрана модель з найбільшою швидкістю навчання – YOLOv8n.pt, з кількістю епох – 50, та роздільною здатністю зображень – 400x400. Результати навчання представлені на рисунках 3 і 4

З рисунків 3 і 4 можна дійти невтішного висновку, що модель не відрізняє помилки відеотрансляції від звичайних об'єктів, що знаходяться на кадрі, але знаходить зону кадру, де знаходяться помилки.

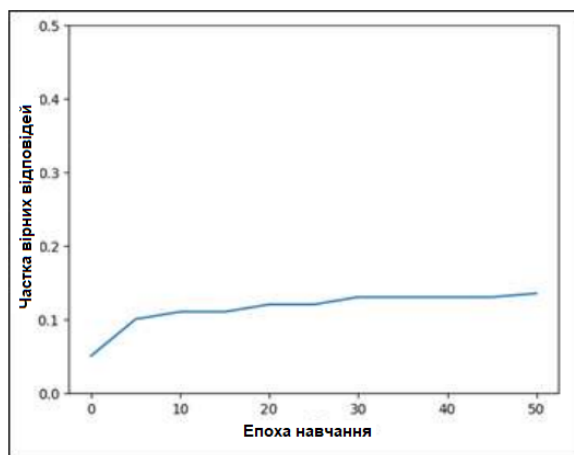


Рисунок 3– Результат навчання моделі

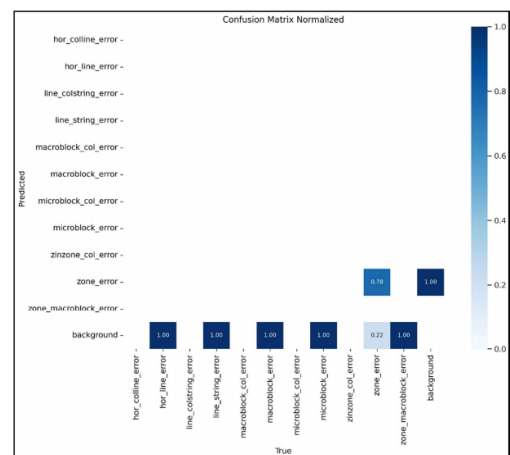


Рисунок 4 – Матриця помилок першого навчання

Для повторного тестового навчання змінимо показник роздільної здатності картини до максимуму (640). Кількість епох навчання залишається незмінною, для більш точних моделей кількість епох буде зменшена, у зв'язку з обмеженням на обчислювальні ресурси. Модель навчання не змінюється.

На рисунках 5 та 6 показані результати повторного тестового навчання моделі.

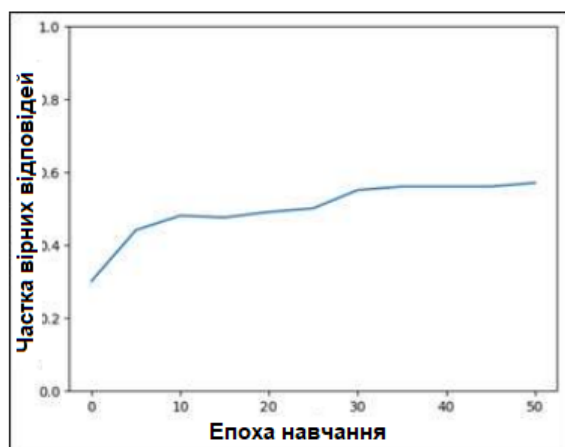


Рисунок 5 – Результат другої спроби навчання моделі

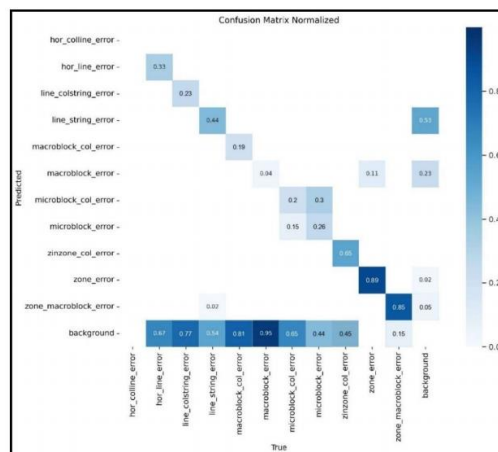


Рисунок 6 – Матриця помилок другої спроби навчання

Навчання моделей YOLO та результати навчання. На основі перших тестових спроб навчання можна зробити висновок, що роздільна здатність зображень – вкрай важливий параметр. Якщо якість зображення надто низька, то нейромережа не зможе відрізнити дрібні помилки зображення від об'єктів розташованих на зображенні, таким чином, необхідно встановити максимально можливий дозвіл для навчання кожної моделі. Також для отримання найкращих результатів змінювалися гіперпараметри моделей. Під час навчання моделей було помічено, що на підсумковий результат впливає значення гіперпараметра "batch" (кількість зображень, що обробляються за один прохід).

Нейросетевые моделі YOLO захищені від ефекту перенавчання – це досягається автоматичним збереженням параметрів, у яких нейронна мережу має мінімальне значення помилки.

У таблиці 2 наведено фінальні значення змінних навчання кожної моделі.

Таблиця 2

Значення параметрів навчання моделей

№	Модель YOLO	Максимальний розмір зображення	Швидкість обробки однієї фотографії (мс)	mAPval (середня точність виявлення об'єктів)	Кількість зображень, що обробляються за один прохід
1	yolov5x6u.pt	1280	1,06	56,8	8
2	yolov5n6u.pt	1280	3,81	42,1	4
3	YOLOv8x.pt	640	3,53	53,9	16
4	YOLOv8n.pt	640	0,99	37,3	12

Найкращі результати навчених моделей навчання наведені на рисунках 7 – 10.

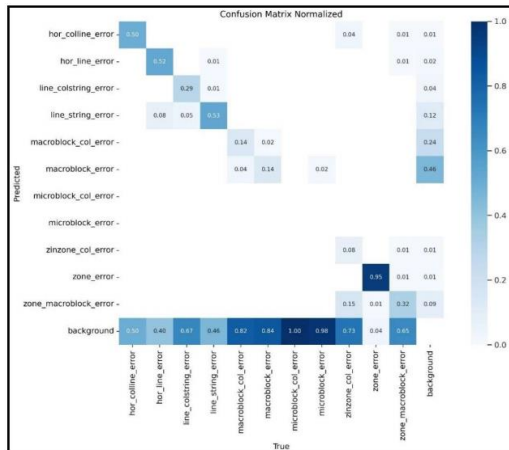


Рисунок 7 – Найкращий результат матриці помилок моделі yolov5x6u.pt

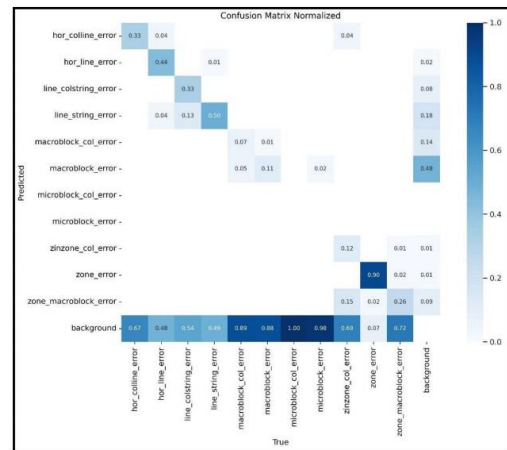


Рисунок 8 – Найкращий результат матриці помилок моделі yolov5n6u.pt

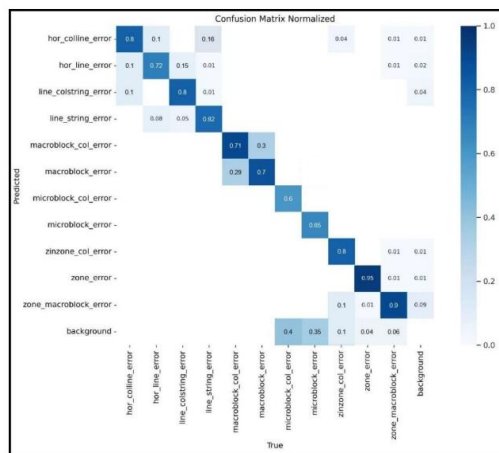


Рисунок 9 – Найкращий результат матриці помилок моделі YOLOv8x.pt

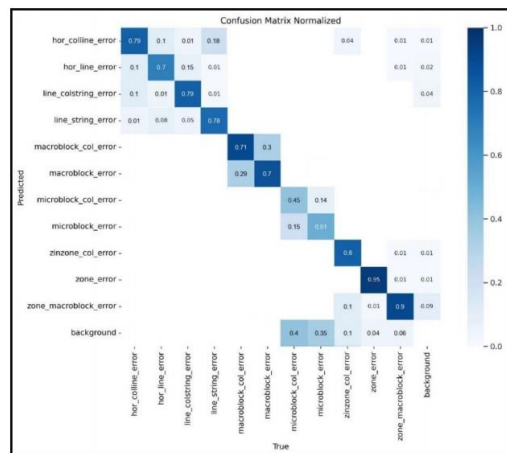


Рисунок 10 – Найкращий результат матриці помилки моделі YOLOv8n.pt

Висновок. З отриманих результатів можна дійти неутішного висновку, що найбільшій точності у виявленні дефектів має навчена модель YOLOv8x.pt.

Після отримання найкращих результатів навчання моделей було виявлено, що найбільший дозвіл впливає якість виявлення помилок лише до певного значення.

На даній вибірці навчання достатньою роздільною здатністю для детектування зображень є 540x540 пікселів.

Обчислювальні ресурси сервісу Google Colaboratory обмежені для безкоштовного використання, тому для моделей YOLOv8x.pt та yolov5x6u.pt було скорочено кількість епох навчання. Навіть із меншою кількістю епох навчання модель YOLOv8x.pt змогла показати найкращий результат.

Для подальшого використання буде використовуватись навчена нейромережева модель YOLOv8x.pt.

ЛІТЕРАТУРА / REFERENCES

1. Mehedi H., Tasneem R., Kiok A., Oksam C. Artifacts Detection and Error Block Analysis від Broadcasted Videos. // arXiv preprint, 2018 – 14 с.

2. Glavota I., Kaprocki Z., Vranjes M., Herceg M. No-reference realtime video transmission artifact detection for video signals. // Journal of RealTime Image Processing, 2020. - С. 799-821.
3. Morals D., Silva A.F., Mylene C.Q., A Correlation-Based NoReference Packet-Loss. // Conference Paper, 2016. - 6 с.
4. Mylene C.Q. Відмінні contributions of blocky, blurry, noisy, i ringing synthetic artifacts до overall annoyance. // Journal of Electronic Imaging, 2012. - 23 с.
5. Charles, P. Digital video and HDTV/P Charles. - San Francisco: Morgan Kaufmann, 2001. - 700 с.

Received 03.04.2025.
Accepted 07.04.2025.

Simulation of a neural network algorithm for the analysis of video flow defects

The paper provides an overview of existing scientific works on video analysis for defect detection and digital broadcasting technologies. The creation of a data set, data distribution by samples are considered. The training of neural network models and obtaining training results are described. Implementation of the method of using a trained neural network model.

From the results obtained, we can draw an unfortunate conclusion that the trained model YOLOv8x.pt has the highest accuracy in detecting defects.

After obtaining the best results of model training, it was found that the highest resolution affects the quality of error detection only up to a certain value.

On this training sample, a sufficient resolution for image detection is 540x540 pixels.

The computing resources of the Google Colaboratory service are limited for free use, so the number of training epochs was reduced for the YOLOv8x.pt and yolov5x6u.pt models.

Even with fewer training epochs, the YOLOv8x.pt model was able to show the best result. For further use, the trained neural network model YOLOv8x.pt will be used.

Keywords: neural network, convolutional neural network, data set, defect, video file, video broadcast, video broadcast.

Островська Катерина Юріївна - доцент, кандидат технічних наук, доцент кафедри Інформаційних технологій і систем ННІ «ДМетІ» УДУНТ.

Зимогляд Андрій Юрійович - доцент, кандидат технічних наук, доцент кафедри Інформаційних технологій і систем ННІ «ДМетІ» УДУНТ.

Козакова Анастасія Олександрівна - магістр кафедри Інформаційних технологій і систем ННІ «ДМетІ» УДУНТ.

Ostrovska Kateryna – Associate Professor, Candidate of Technical Sciences, Associate Professor of the Department of Information Technologies and Systems of the Scientific Research Institute "DMetI" UDUNT.

Zymogliad Andriy – Associate Professor, Candidate of Technical Sciences, Associate Professor of the Department of Information Technologies and Systems of the Scientific Research Institute "DMetI" UDUNT.

Kozakova Anastasia - Master of the Department of Information Technologies and Systems of the Scientific Research Institute "DMetI" UDUNT.