

ПРОБЛЕМИ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ВБУДОВАНИХ СИСТЕМ КЕРУВАННЯ З МОДУЛЕМ КОМП'ЮТЕРНОГО ЗОРУ

Анотація. У статті досліджується актуальна проблема підвищення ефективності вбудованих систем автоматичного керування (САК), в яких робота віртуальних датчиків виконується за алгоритмами комп'ютерного зору на базі глибоких нейромереж. Виконаний аналіз сучасних підходів до компресії нейромережевих моделей (квантування, проріджування) у контексті парадигм Edge AI та TinyML. За результатом огляду робіт інших авторів висвітлена недостатня узгодженість параметрів САК: оптимізація моделей виключно за метриками машинного навчання, без урахування динаміки об'єкта, що знижує запаси стійкості системи через зміну статистики помилок та внесення стохастичного шуму. Сформульовані завдання дослідження, що поєднує цифрові двійники, емуляцію та апаратну реалізацію (Hardware-in-the-loop) поставленої задачі пошуку оптимального компромісу між затримкою інференсу та точністю оцінки стану об'єкта для підвищення ефективності автономних систем.

Ключові слова: комп'ютерний зір, вбудовані системи, автоматичне керування, TinyML, компресія нейромереж, запас стійкості, цифрові двійники, апаратно-орієнтована оптимізація.

Вступ та обґрунтування актуальності дослідження. Розвиток автономної робототехніки, безпілотних літальних апаратів (БПЛА) вимагає розширення сенсорних можливостей кіберфізичних систем. Тому в сучасних системах керування все частіше відбувається заміна коштовних стандартних фізичних датчиків на модулі комп'ютерного зору (КЗ), які забезпечують візуальний зворотній зв'язок керованого об'єкта з системою автоматичного керування (САК), тобто за наявності вказаного елемента САК механізми здатні орієнтуватись в просторі, розпізнавати складні об'єкти, будувати адаптивні маршрути і таке інше. Крім того, скорочення кількості фізичних датчиків підвищує автономність САК без суттєвої втрати точності та гнучкості роботи керованого об'єкта в цілому [1].

Проте інтеграція сучасних алгоритмів КЗ (на базі згорткових нейромереж чи трансформерів) у периферійні пристрої (Edge devices) ускладнюється жорсткими вимогами до апаратного забезпечення [1]. У вбудованих системах керування проблема застосування мікроконтролерів та одноплатних комп'ютерів полягає в обмеженнях основних характеристик, до яких відносять лімітовану пам'ять, обчислювальну потужність і ресурси енергоспоживання. Таким чином, для впровадження в периферійні при-

строї "важких" нейромереж, необхідно адаптувати та стискати (компресувати) їх математичні моделі.

Водночас треба враховувати відомий факт, що агресивна компресія має критичний недолік - вона знижує точність розпізнавання. Крім того, у замкненому контурі керування, який реалізується в периферійних пристроях, навіть незначна похибка моделі спричиняє серйозні системні наслідки: хибні рішення регулятора, виникнення автоколивань, втрату цілі або повну дестабілізацію керованого об'єкта. Отже, стиснення нейромережевої моделі повинно відбуватись за визначеними умовами для конкретного керованого об'єкта, тобто традиційної "програмної" оптимізації в таких задачах недостатньо.

Таким чином актуальність роботи полягає в проблемі ефективного поєднання обчислювальної спроможності нейромережі та коректної роботи САК автономних систем. Досягнення цього балансу дозволить підвищити загальну (інтегральну) надійність та ефективність сучасних автономних пристроїв.

Сучасні методи підвищення обчислювальної ефективності (Edge AI та TinyML). Як зазначено в огляді С. El Zeinaty та співавторів [1], перенесення алгоритмів комп'ютерного зору з хмарних серверів безпосередньо на бортові пристрої реалізується завдяки підходам Edge Computing та TinyML. При цьому для забезпечення стабільної частоти кадрів (FPS) на ресурсно-обмежених апаратних платформах розробники застосовують комплекс технік просторово-часової оптимізації моделей:

- **Квантування (Quantization):** Перетворення значень ваг та активацій нейромережі з точного формату з рухомою комою (FP32) у цілочисельні формати меншої рядності (INT16, INT8, INT4). Це зменшує споживання пам'яті у 4-8 разів і дозволяє процесорам використовувати енергоефективні інструкції для цілочисельних обчислень.

- **Проріджування (Pruning):** Виявлення та видалення найменш важливих синаптичних зв'язків у нейромережі, а іноді - цілих каналів чи фільтрів. Наприклад, В. Jiang та співавтори [2] пропонують об'єднати цей підхід із квантуванням (стратегія Single-shot pruning and quantization), що прискорює інференс у 6-8 разів.

- **Дистиляція знань (Knowledge Distillation) та пошук архітектур (NAS):** Навчання компактних "моделей-учнів" за допомогою великих "моделей-вчителів", а також автоматичний пошук найкращої структури мережі для конкретного мікропроцесора.

Ще один відомий підхід підвищення обчислювальної ефективності має в основі адаптивне розподілення обчислень. Зокрема, С. McKee [6] пропонує архітектуру edge-cloud для систем інспекції: рутинні задачі обробляє периферійний пристрій, а складні випадки передаються у хмару. Однак для систем реального часу (як-от стабілізація БПЛА) втрата зв'язку з хмарою є неприпустимою, тому модель має гарантовано та повністю працювати на бортовому комп'ютері.

Попри всі переваги цих методів, Park J. та співавтори [3] зазначають: будь-яка компресія неминуче вимагає компромісу між точністю (Accuracy) та затримкою (Latency). Крім того, М. de Prado [4] наголошує, що квантовані моделі стають менш

стійкими (робастними) до зовнішніх збурень, а це створює пряму загрозу безпеці системи керування.

Узгодження між обчислювальною ефективністю та якістю керування. Однією з проблем сучасних досліджень є те, що фахівці з Data Science оптимізують нейромережі переважно ізольовано від об'єкта керування. Вони оцінюють якість детектора статичними метриками (mAP, F1-score). Проте інтеграція такого модуля в контур зворотного зв'язку реальної фізичної системи докорінно змінює характер його роботи.

- Відеосенсор як динамічна ланка із запізненням. Загальновідомо, що в класичній теорії автоматичного керування (ТАК) часову затримку прийнято розглядати як ланку чистого запізнювання, що описується передавальною функцією наступного вигляду:

$$W_{\tau}(s) = e^{-sT} \quad (1)$$

де T - час затримки.

Така ланка створює фазове запізнення, яке лінійно зростає зі збільшенням частоти, що закономірно призводить до зниження запасу стійкості системи по фазі. Згідно з теорією автоматичного керування, перевищення так званої граничної затримки (Delay Margin) робить систему нестійкою попри будь-яку якість вимірювань. Наочним прикладом є використання «важкої» моделі (для розпізнавання та обчислення координат керуємого об'єкту) з часом інференсу 150 мс для керування об'єктом із механічною сталою часу 50 мс - це гарантовано спричинить втрату керованості та руйнування системи, незважаючи на абсолютно безпомилкову роботу комп'ютерного зору.

- Вплив компресії на статистику помилок та стохастичний шум. Щоб подолати проблему затримки, розробники компресують моделі. Це різко зменшує час інференсу T (наприклад, до 15 мс) і відновлює фазовий запас стійкості.

Однак стиснення спотворює статистику помилок моделі. Зокрема, зниження розрядності ваг генерує шум квантування. На практиці він виглядає як високочастотний "джиттер" - хаотичне тремтіння координат обмежувальної рамки. Математично сигнал $y(t)$ на виході такого віртуального датчика описується рівнянням:

$$y(t) = f_{cnn} (x(t - \tau(t))) + w(t), \quad (2)$$

де f_{cnn} (function of Convolutional Neural Network) - це функція перетворення згорткової нейромережі (модуля комп'ютерного зору), яка відображає реальний просторовий стан об'єкта у вихідні координати обмежувальної рамки; $x(t)$ - реальний стан керованого об'єкта, $\tau(t)$ - змінна затримка інференсу, а $w(t)$ - стохастичний шум вимірювань, спричинений компресією; t - час.

Наприклад, у контурі керування цей шум діє вкрай деструктивно, особливо на диференціальну складову ПІД (пропорційно-інтегрально-диференціальний) - регулятора. У відповідь на хибні коливання координат регулятор починає хаотично "смикати" виконавчі механізми. Це різко погіршує якість перехідного процесу та прискорює механічний знос обладнання.

Стан досліджень в українському та світовому науковому просторі. Дослідження в галузі інтеграції комп'ютерного зору в системи керування в Україні набули критичної актуальності національного масштабу, що прямо пов'язано з нагальними потребами у розробці автономних БПЛА, систем автосупроводу цілей та наземних роботизованих платформ.

Аналіз вітчизняного наукового простору свідчить про значні успіхи в окремих напрямках цієї галузі. Наприклад, у роботі [7] успішно вирішуються завдання розпізнавання та моніторингу об'єктів за допомогою машинного навчання, проте оптимізація розглядається переважно в контексті офлайн-аналізу.

Дослідження І. Худякова та ін. [8] охоплюють моделювання інформаційних систем для дистанційного моніторингу транспортних засобів, де акцент робиться на архітектурі передачі даних, а не на динаміці керування в реальному часі на Edge-пристроях.

Крім того, вітчизняні науковці активно досліджують методи зменшення надлишковості самих візуальних даних. Наприклад, у роботі фахівців ВНТУ О. Кавки та ін. [10] детально проаналізовано алгоритми стиснення зображень із втратами (на основі дискретного косинусного перетворення та вейвлетів). Проте вплив неминучих артефактів такого стиснення на точність нейромережевого інференсу, коли зображення надходить безпосередньо в контур керування, залишається недослідженим.

Аналітичні підходи до обробки інформації в складних технічних системах (зокрема аерокосмічного профілю) активно розробляються фахівцями Національного аерокосмічного університету ім. М.Є. Жуковського «ХАІ» [9], однак специфіка нелінійних похибок глибоких нейромереж часто залишається поза увагою.

Таким чином, огляд праць українських науковців щодо задач інтеграції комп'ютерного зору в системи керування показав, що спостерігається певна неузгодженість підходів, характерна як для вітчизняної, так і для світової наукової думки, що полягає в наступному:

- Фахівці в галузі інформаційних технологій зосереджуються на архітектурі мереж та їх апаратному розгортанні, фокусуючись переважно на максимізації показників швидкодії (FPS) та зменшенні розміру моделі. Питання про те, як саме шум стисненої YOLO-мережі вплине на поведінку ПІД-регулятора, зазвичай ігнорується.

- Фахівці в галузі автоматичного керування розробляють робастні системи (LQR, MPC), найчастіше приймаючи відеодатчик як класичну ланку із детермінованим запізненням та білим гауссівським шумом, що може не відповідати реальній стохастичній природі нейромережевого інференсу.

Визначена ізольованість науковців різних галузей показала нагальну потребу у проведенні міждисциплінарних (Co-design) досліджень, здатних інтегрувати параметри компресії нейромереж безпосередньо в рівняння динаміки керованих електромеханічних об'єктів.

Недосліджені прогалини. Проведений аналіз наукових джерел виявив низку невирішених проблем у процесі інтеграції нейромереж на Edge-пристроях, які гальмують ефективне впровадження комп'ютерного зору у контури зворотного зв'язку:

- **Брак міждисциплінарних (Co-design) метрик оптимізації.** Досі немає комплексної функції втрат, яка б під час стиснення мережі враховувала динаміку фізичного об'єкта та мінімізувала інтегральну квадратичну похибку (ISE) системи керування.

- **Запит на методологію аналізу стійкості систем із модулем комп'ютерного зору.** Класичні методи теорії керування не адаптовані до специфіки впровадження TinyML, оскільки вони не дозволяють комплексно оцінити спільний дестабілізуючий вплив змінної затримки інференсу $\tau(t)$ та випадкового шуму квантування $w(t)$ на контур зворотного зв'язку.

- **Дефіцит експериментальних даних щодо стійкості.** У літературі бракує кількісних оцінок того, як послідовне проріджування мережі на фізичних стендах змінює запаси стійкості системи керування по фазі та амплітуді.

- **Відсутність інженерних методик для розробників.** Існує нагальна потреба в математично обґрунтованих критеріях. Інженерам потрібен інструмент, щоб чітко визначати гранично допустимий ступінь стиснення мережі, за якого об'єкт не втратить керованість.

Завдання комплексного експериментального дослідження. Для вирішення вищезазначених проблем з урахуванням необхідності компромісу між затримкою і точністю у процесі інтеграції нейромереж на Edge-пристроях сформульовані постановки поетапних завдань комплексного експериментального дослідження. Спираючись на сучасні практики розробки кіберфізичних систем, пропонована багаторівнева методологія передбачає послідовний перехід від створення ідеалізованих цифрових двійників до напівнатурного (Hardware-in-the-loop, HIL) та натурального моделювання. Цей процес складається з наступних шести завдань (Рис. 1):

Завдання 1: Створення цифрового двійника (фізична та візуальна модель).

- Розробка в САПР (CAD) точної 3D-моделі об'єкта (наприклад, зворотного маятника чи мобільної платформи).
- Імпортування 3D-моделі об'єкта у середовища фото та відео реалістичної симуляції (Unity або NVIDIA Isaac Sim) та підключенням фізичного рушія (наприклад MuJoCo для середовища Unity).

Мета завдання - забезпечити адекватну динаміку твердих тіл і високоякісний рендер сцени.

Завдання 2: Синтез ідеального керування.

- Побудова класичної системи керування (ПД, LQR) із використанням "ідеальних" віртуальних датчиків симулятора.

Мета завдання - отримати значення еталонних параметрів (Ground Truth) перехідних процесів оскільки "ідеальні" віртуальні датчики миттєво і без шумів видають точну позицію, швидкість та прискорення.

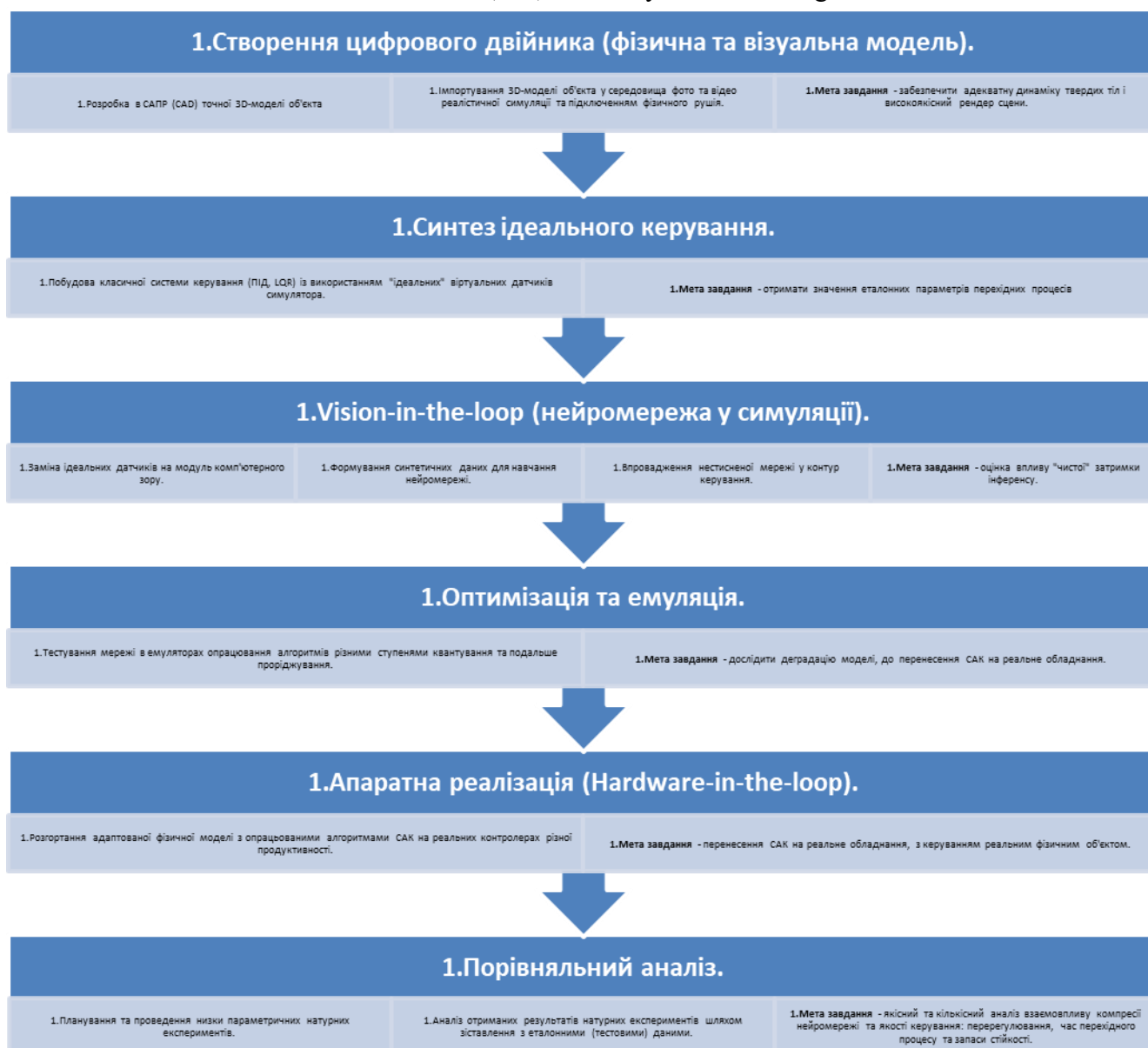


Рисунок 1 - Блок-схема пропонованої багаторівневої методології - комплексного експериментального дослідження

Завдання 3: Vision-in-the-loop (нейромережа у симуляції).

- Заміна ідеальних датчиків на модуль комп'ютерного зору.
- Формування синтетичних даних для навчання нейромережі.
- Впровадження нестисненої мережі (FP32) у контур керування.

Мета завдання - оцінка впливу "чистої" затримки інференсу.

Завдання 4: Оптимізація та емуляція (TinyML).

- Тестування мережі з використанням емуляторів (наприклад QEMU або Renode): опрацювання алгоритмів різними ступенями квантування та подальше проріджування.

Мета завдання - дослідити деградацію моделі, до перенесення САК на реальне обладнання.

Завдання 5: Апаратна реалізація (Hardware-in-the-loop).

- Розгортання (впровадження) адаптованої фізичної моделі з опрацьованими алгоритмами САК на реальних контролерах різної продуктивності: мікроконтролерах (STM32), одноплатниках (Raspberry Pi) або тензорних обчислювачах (NVIDIA Jetson).

Мета завдання - перенесення САК на реальне обладнання, з керуванням реальним фізичним об'єктом.

Завдання 6: Порівняльний аналіз.

- Планування та проведення низки параметричних натурних експериментів.

- Аналіз отриманих результатів натурних експериментів шляхом зіставлення з еталонними (тестовими) даними (із Завдання 2).

Мета завдання - якісний та кількісний аналіз взаємовпливу компресії нейромережі та якості керування: перерегулювання, час перехідного процесу та запаси стійкості.

Висновки. Проведений аналіз дозволяє припустити, що підвищення ефективності вбудованих систем керування з модулем комп'ютерного зору є комплексною задачею, що потребує міждисциплінарного підходу. Традиційні методи ізольованої оптимізації нейромереж наразі стикаються з певними обмеженнями, оскільки підвищення швидкості інференсу задля розгортання моделей на вбудованих системах, без урахування динамічних характеристик об'єкта, може призводити до небажаних ефектів у контурі керування.

Окреслений у роботі запит на новий підхід вказує на потенціал зближення підходів Data Science та теорії автоматичного керування. Це дозволить більш глибоко дослідити взаємозв'язок між параметрами апаратної компресії (TinyML) та показниками якості керування, такими як перерегулювання та час перехідного процесу.

Сформульовані постановки послідовних завдань комплексного експериментального дослідження, що охоплюють етапи від моделювання в Unity/Isaac Sim до HiL-тестування на платформах NVIDIA Jetson та STM32, та становлять основу для подальшого розвитку визначеної проблеми підвищення ефективності вбудованих систем керування з модулем комп'ютерного зору.

У перспективі реалізація послідовних завдань комплексного експериментального дослідження спрямована на розвиток математичної моделі контуру керування САК із врахуванням впливу змінної затримки та шуму квантування для подальшого формулювання обґрунтованих рекомендацій з підвищення сукупної ефективності сучасних автономних систем.

ЛІТЕРАТУРА

1. Designing Object Detection Models for TinyML: Foundations, Comparative Analysis, Challenges, and Emerging Solutions / C. El Zeinaty, W. Hamidouche, G. Herrou, D. Menard. *ACM Computing Surveys*. 2024. Vol. 56. P. 1–46. DOI: <https://doi.org/10.1145/3744339>.
2. Jiang B., Chen J., Liu Y. Single-shot pruning and quantization for hardware-friendly neural network acceleration. *Engineering Applications of Artificial Intelligence*. 2023. Vol. 126, Part B. 106816. DOI: <https://doi.org/10.1016/j.engappai.2023.106816>.

3. Park J., Kim P., Ko D. Real-time open-vocabulary perception for mobile robots on edge devices: a systematic analysis of the accuracy-latency trade-off. *Frontiers in Robotics and AI*. 2025. Vol. 12. 1693988. DOI: <https://doi.org/10.3389/frobt.2025.1693988>.
4. Robustifying the Deployment of tinyML Models for Autonomous Mini-Vehicles / M. de Prado, M. Rusci, A. Capotondi et al. *Sensors*. 2021. Vol. 21, no. 4. 1339. DOI: <https://doi.org/10.3390/s21041339>.
5. Alaklabi S., Alharbi S. DRL-TinyEdge: Energy- and Latency-Aware Deep Reinforcement Learning for Adaptive TinyML at the 6G Edge. *Future Internet*. 2026. Vol. 18, no. 1. 31. DOI: <https://doi.org/10.3390/fi18010031>.
6. McKee C. Design, Embedded Implementation, and Performance Optimization of a Real-Time AI-Driven Vision Inspection System for Automated Industrial Quality Control: Technical Report. 2025.
7. Розпізнавання та моніторинг водних об'єктів на оптичних супутникових зображеннях із використанням машинного навчання / В. Каштан, В. Гнатушенко, І. Удовик, О. Шевцова. *Information Technology: Computer Science, Software Engineering and Cyber Security*. 2023. Вип. 3. С. 32–42. DOI: <https://doi.org/10.32782/IT/2023-3-4>.
8. Особливості моделювання та побудови інформаційної системи дистанційного моніторингу технічного стану транспортних засобів / І. В. Худяков, І. В. Грицук, В. В. Черненко та ін. *Вісник машинобудування та транспорту*. 2021. № 2 (14). С. 140–148. DOI: <https://doi.org/10.31649/2413-4503-2021-14-2-140-148>.
9. Берднікова А. Л., Манжос Ю. С. Інформаційна технологія моделювання складних систем. *Системи обробки інформації*. 2012. Вип. 3 (101), т. 2. С. 2–7.
10. Аналіз алгоритмів стиснення зображень із втратами / О. О. Кавка, В. П. Майда-нюк, О.Н. Романюк, Є.К. Завальнюк. *Інформаційні технології та комп'ютерна інженерія*. 2023. № 3. С. 59–64. DOI: <https://doi.org/10.31649/1999-9941-2023-58-3-59-64>.

REFERENCES

1. El Zeinaty, C., Hamidouche, W., Herrou, G., & Menard, D. (2024). Designing object detection models for TinyML: Foundations, comparative analysis, challenges, and emerging solutions. *ACM Computing Surveys*, 56(8), 1–46. <https://doi.org/10.1145/3744339>
2. Jiang, B., Chen, J., & Liu, Y. (2023). Single-shot pruning and quantization for hardware-friendly neural network acceleration. *Engineering Applications of Artificial Intelligence*, 126(Part B), Article 106816. <https://www.google.com/search?q=https://doi.org/10.1016/j.engappai.2023.106816>
3. Park, J., Kim, P., & Ko, D. (2025). Real-time open-vocabulary perception for mobile robots on edge devices: A systematic analysis of the accuracy-latency trade-off. *Frontiers in Robotics and AI*, 12, Article 1693988. <https://www.google.com/search?q=https://doi.org/10.3389/frobt.2025.1693988>
4. de Prado, M., Rusci, M., Capotondi, A., Donze, R., Benini, L., & Pazos, N. (2021). Robustifying the deployment of tinyML models for autonomous mini-vehicles. *Sensors*, 21(4), Article 1339. <https://www.google.com/search?q=https://doi.org/10.3390/s21041339>

5. Alaklabi, S., & Alharbi, S. (2026). DRL-TinyEdge: Energy- and latency-aware deep reinforcement learning for adaptive TinyML at the 6G edge. *Future Internet*, 18(1), Article 31. <https://www.google.com/search?q=https://doi.org/10.3390/fi18010031>
6. McKee, C. (2025). *Design, embedded implementation, and performance optimization of a real-time AI-driven vision inspection system for automated industrial quality control* [Technical report].
7. Kashtan, V., Hnatushenko, V., Udovik, I., & Shevtsova, O. (2023). Rozpiznavannia ta monitorynh vodnykh ob'ektiv na optychnykh suputnykovykh zobrazhenniakh iz vykorystanniam mashynnoho navchannia [Recognition and monitoring of water objects on optical satellite images using machine learning]. *Information Technology: Computer Science, Software Engineering and Cyber Security*, 3, 32–42. <https://doi.org/10.32782/IT/2023-3-4>
8. Khudiakov, I. V., Gritsuk, I. V., Chernenko, V. V., Gritsuk, Y. V., Pohorletskyi, D. S., Makarova, T. V., & Manzhelei, V. S. (2021). Osoblyvosti modeliuvannia ta pobudovy informatsiinoi systemy dystantsiinoho monitorynhu tekhnichnoho stanu transportnykh zasobiv [Features of modeling and construction of the information system of remote monitoring of the technical condition of vehicles]. *Visnyk mashynobuduvannia ta transportu* [*Herald of Mechanical Engineering and Transport*], 14(2), 140–148. <https://doi.org/10.31649/2413-4503-2021-14-2-140-148>
9. Berdnikova, A. L., & Manzhos, Y. S. (2012). Informatsiina tekhnolohiia modeliuvannia skladnykh system [Information technology for modeling of complex systems]. *Systemy obrobky informatsii* [*Information Processing Systems*], 101(3), 2–7.
10. Kavka, O. O., Maidaniuk, V. P., Romanyuk, O. N., & Zavalniuk, Y. K. (2023). Analiz alhorytmiv stysnennia zobrazhen iz vtratamy [Analysis of the lossy image compression algorithms]. *Informatsiini tekhnolohii ta kompiuterna inzheneriia* [*Information Technologies and Computer Engineering*], 58(3), 59–64. <https://doi.org/10.31649/1999-9941-2023-58-3-59-64>

Received 17.04.2026.
Accepted 22.04.2026.
Published 30.04.2026

Issues of improving the efficiency of embedded control systems with a computer vision module

The development of autonomous robotics, UAVs, and Industry 4.0 requires expanding the sensory capabilities of cyber-physical systems through the implementation of computer vision (CV) modules on edge devices [1, 2]. The transfer of CV algorithms is implemented via Edge Computing and TinyML paradigms, utilizing spatio-temporal optimization techniques such as quantization (transition from FP32 to INT8/INT4) and synaptic pruning [3, 4].

Global studies indicate that any compression requires a trade-off between accuracy and latency, and also makes models less robust to external disturbances [5]. Domestic scientists have significant achievements in the fields of machine learning, monitoring, and image compression; however, these issues are primarily considered in the context of offline analysis or data transmission [6, 7].

In the scientific community, there is an inconsistency of approaches: Data Science specialists evaluate models in isolation using static metrics and ignore the impact of compressed network noise on the control system [8]. Meanwhile, automatic control specialists perceive the video sensor as a classic link with deterministic delay, which may not correspond to the stochastic nature of inference [9].

Finding the optimal trade-off between the computational efficiency of the neural network (inference latency) and the performance quality of the automatic control system itself to achieve maximum reliability and efficiency of autonomous devices.

Integrating modern computer vision algorithms into resource-limited edge devices leads to a fundamental contradiction.

In classical automatic control theory, a video sensor acts as a dynamic link of pure delay, which creates a phase lag and linearly reduces the system's phase stability margin [10]. Exceeding the delay limit inevitably leads to a loss of controllability of the object. To avoid this, developers use aggressive compression to reduce inference time, which restores the phase margin but distorts error statistics [11]. Specifically, bit-depth reduction generates quantization noise (high-frequency jitter), which acts extremely destructively on the differential component of the regulator and causes chaotic oscillations of the actuators [12].

Since classic optimization metrics are not suitable for such cases, the authors propose a comprehensive multi-level methodology (Co-design) that allows integrating compression parameters into the dynamic's equations. It encompasses six sequential tasks: creating a digital twin using the MuJoCo physics engine in Unity or NVIDIA Isaac Sim environments [13, 14]; synthesis of ideal control to obtain reference data; Vision-in-the-loop simulation with an uncompressed network; optimization and emulation in QEMU/Renode for TinyML testing [15]; hardware implementation of Hardware-in-the-loop on Nvidia Jetson/Raspberry Pi/STM32 controllers; and a final comparative analysis of the impact of compression on overshoot and stability margins.

The conducted analysis suggests that increasing the efficiency of embedded ACS with computer vision is a complex interdisciplinary task. Optimizing neural networks solely by machine learning metrics, without considering the physical object's dynamics, is not effective and can destabilize the control loop.

The proposed approach, which combines Data Science and control theory methods (from simulation to HIL testing), creates a solid foundation for researching the relationship between hardware compression parameters and control quality indicators. In the future, its implementation will help develop more accurate mathematical descriptions of the control loop, considering variable delay and quantization noise, as well as formulate clear engineering recommendations for developers of autonomous systems.

Keywords: computer vision, embedded systems, automatic control, TinyML, neural network compression, stability margin, hardware-oriented optimization.

Гусєв Артем Юрійович - аспірант кафедри КБКІТ, Дніпровський національний університет імені Олеся Гончара.

ORCID: <http://orcid.org/0009-0002-3846-6204>

Клим Вікторія Юрїївна - канд. наук, доц., Дніпровський національний університет імені Олеся Гончара.

ORCID: <http://orcid.org/0000-0002-5887-1955>

Husiev Artem - PhD student at the Department of Cybersecurity and Computer-Integrated Technologies, Oles Honchar Dnipro National University.

ORCID: <http://orcid.org/0009-0002-3846-6204>

Klym Viktoriia - cand. sc., assoc. prof., Oles Honchar Dnipro National University.

ORCID: <http://orcid.org/0000-0002-5887-1955>