

<https://doi.org/10.31891/2307-5732-2026-363-8>

УДК 004.93:629.735

ДОБРОДИЙ РОМАН

Національний технічний університет України «Київський Політехнічний Інститут імені Ігоря Сікорського»
<https://orcid.org/0009-0000-3755-809X>
e-mail: dobru2001@gmail.com

БОНДАРЕНКО ВІКТОР

Національний технічний університет України «Київський Політехнічний Інститут імені Ігоря Сікорського»
<https://orcid.org/0000-0003-1663-4799>
e-mail: vicnbondarenko@gmail.com

БОНДАРЕНКО НАТАЛІЯ

Національний технічний університет України «Київський Політехнічний Інститут імені Ігоря Сікорського»
<https://orcid.org/0000-0002-9237-8187>
e-mail: nalbondarenko@gmail.com

АПАРАТНО-ПРОГРАМНИЙ КОМПЛЕКС ОБРОБКИ ВІДЕОПОТОКІВ

Стаття присвячена вирішенню актуальної науково-прикладної задачі створення апаратно-програмного комплексу обробки відеопотоків в режимі реального часу для автономних систем спостереження на основі перенесення обчислень безпосередньо в систему із застосуванням концепції Edge AI. Запропоноване рішення поєднує гетерогенну платформу системи на кристалі з оптимізованим виявленням об'єктів на основі нейронної мережі, досягаючи високої продуктивності та низької затримки за жорстких обмежень енергоспоживання. Результати експериментальних досліджень підтверджують досягнення продуктивності понад 50 кадрів за секунду при сумарній затримці обробки менш ніж 40 мс та споживаній потужності до 12 Вт.

Ключові слова: автономна система спостереження, відеопотік, комп'ютерний зір, нейронна мережа

DOBRODIY ROMAN, BONDARENKO VIKTOR, BONDARENKO NATALIYA,
National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute»

HARDWARE AND SOFTWARE VIDEO PROCESSING COMPLEX

The article is devoted to solving the current scientific and applied problem of creating a hardware-software complex for processing video streams in autonomous observation systems, focused on real-time operation in conditions of limited computing and energy resources. It is shown that traditional approaches to transmitting a "raw" video stream to control stations do not provide the necessary delay and resistance to radio-electronic interference. The feasibility of transferring calculations directly to the autonomous observation system using the Edge AI concept is substantiated. An analysis of modern hardware architectures for building machine vision systems is carried out, in particular microcontrollers, FPGAs and heterogeneous systems on a crystal with AI accelerators. A device heterogeneous architecture based on the NVIDIA Jetson Orin NX module and a video sensor with a MIPI CSI 2 interface is proposed, which provides minimal image capture delay. The algorithmic support for object detection based on the YOLOv8 neural network and methods for its optimization using the TensorRT library with INT8 quantization are described. Using the YOLOv8 model ensured high detection accuracy. The integration of optimization algorithms allowed us to significantly improve system performance compared to standard use of PyTorch. The results of experimental studies are presented, confirming the achievement of performance of more than 50 frames per second with a total processing delay (latency) of less than 40 ms and power consumption of up to 12 W. Low latency is a critical factor for using the complex in the control loop of high-speed observation systems. The results obtained indicate the possibility of practical application of the hardware and software complex as part of modern autonomous observation systems.

Keywords: autonomous observation system, video stream, computer vision, neural network

Стаття надійшла до редакції / Received 12.01.2026
Прийнята до друку / Accepted 11.02.2026
Опубліковано / Published 26.03.2026



This is an Open Access article distributed under the terms of the [Creative Commons CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)

© Добродій Роман, Бондаренко Віктор, Бондаренко Наталія

Вступ

Стрімкий розвиток автономних систем спостереження зумовив зростання вимог до обробки інформації, зокрема відеоданих. Сучасні автономні системи спостереження застосовуються для розвідки, моніторингу, навігації та виконання автономних місій у складних умовах, де використання віддаленої обробки відео є обмеженим через затримки каналів зв'язку та вплив засобів створення радіоелектронних завад. У зв'язку з цим актуальною є задача проектування компактних і енергоефективних систем технічного зору, здатних виконувати аналіз відеопотоків безпосередньо в автономній системі.

Метою роботи є розробка ефективного та доступного апаратно-програмного комплексу для обробки відеопотоків з урахуванням вимог реального часу, обмежень енергоспоживання та точності детекції.

Аналіз сучасних рішень

У сучасних системах технічного зору для автономних систем спостереження застосовуються різні апаратні та програмні підходи, вибір яких визначається вимогами до швидкодії, енергоспоживання та гнучкості алгоритмів. Найпростішим рішенням є використання мікроконтролерів (MCU), які широко застосовуються у сучасних системах спостереження. Однак обчислювальні ресурси MCU є недостатніми для обробки відеопотоків високої роздільної здатності та реалізації сучасних алгоритмів комп'ютерного зору [1].

Більш продуктивним підходом є застосування програмованих логічних інтегральних схем (ПЛІС). Завдяки апаратному паралелізму, ПЛІС забезпечують мінімальні затримки обробки відеоданих та високу детермінованість роботи. Разом з тим, розробка та модифікація алгоритмів на ПЛІС є складною та трудомісткою, а реалізація гнучких нейронних мереж потребує значних ресурсів [2].

Найбільш перспективним напрямом є використання гетерогенних систем на кристалі (SoC), які

поєднують центральний процесор, графічний процесор та спеціалізовані AI-акселератори. Такі платформи дають змогу реалізувати складні алгоритми глибокого навчання з використанням стандартних фреймворків, забезпечуючи компроміс між продуктивністю, енергоефективністю та гнучкістю розробки [3, 4].

Для наочного порівняння сучасних підходів до обробки відеопотоків в автономних системах спостереження доцільно виділити ключові параметри, що визначають придатність тієї чи іншої архітектури: продуктивність обробки відео, затримка, енергоспоживання та складність розробки. Мікроконтрольні рішення забезпечують мінімальне енергоспоживання, однак їх продуктивність обмежується одиницями кадрів за секунду при обробці відео високої роздільної здатності. ПЛІС дають змогу досягати затримок на рівні одиниць мілісекунд, проте потребують складного проектування та не підтримують швидку модифікацію алгоритмів.

Гетерогенні SoC-платформи, зокрема сімейство NVIDIA Jetson, забезпечують продуктивність на рівні десятків трильйонів операцій за секунду при споживаній потужності до 15 Вт, що робить їх оптимальним вибором для автономних систем спостереження. Саме гетерогенну SoC-платформу покладено в основу запропонованого у статті рішення.

Архітектура запропонованого комплексу

Структурна схема розробленого апаратно-програмного комплексу наведена на рис. 1. Вона відображає логіку взаємодії основних функціональних блоків комплексу та їх інтеграцію у загальну систему керування автономною літаючою системою спостереження, обраною як приклад реалізації. Запропонований апаратно-програмний комплекс побудовано на базі модуля NVIDIA Jetson Orin NX з 16 ГБ оперативної пам'яті, який виконує функцію центрального обчислювального вузла. До складу комплексу також входять відеосенсор, підсистема живлення, інтерфейси зв'язку та програмні модулі обробки даних. Для захоплення відео використано камеру з сенсором Sony IMX477 та інтерфейсом MIPI CSI-2, що забезпечує прямий доступ до апаратного блока обробки зображень (ISP) та мінімізує затримки передачі даних. Обробка первинного відеопотоку (демозаїка, корекція кольору, масштабування) виконується апаратними засобами SoC без залучення центрального процесора. Інтеграція з автопілотом здійснюється через інтерфейс UART із використанням протоколу MAVLink, що дає змогу передавати координати виявлених об'єктів, службову телеметрію та керуючі команди в реальному часі. Підсистема живлення реалізована на базі імпульсного DC–DC перетворювача, який забезпечує стабільну роботу пристрою при коливаннях бортової напруги. Запропонована структура забезпечує потокову обробку відеоданих та мінімальну сумарну затримку.

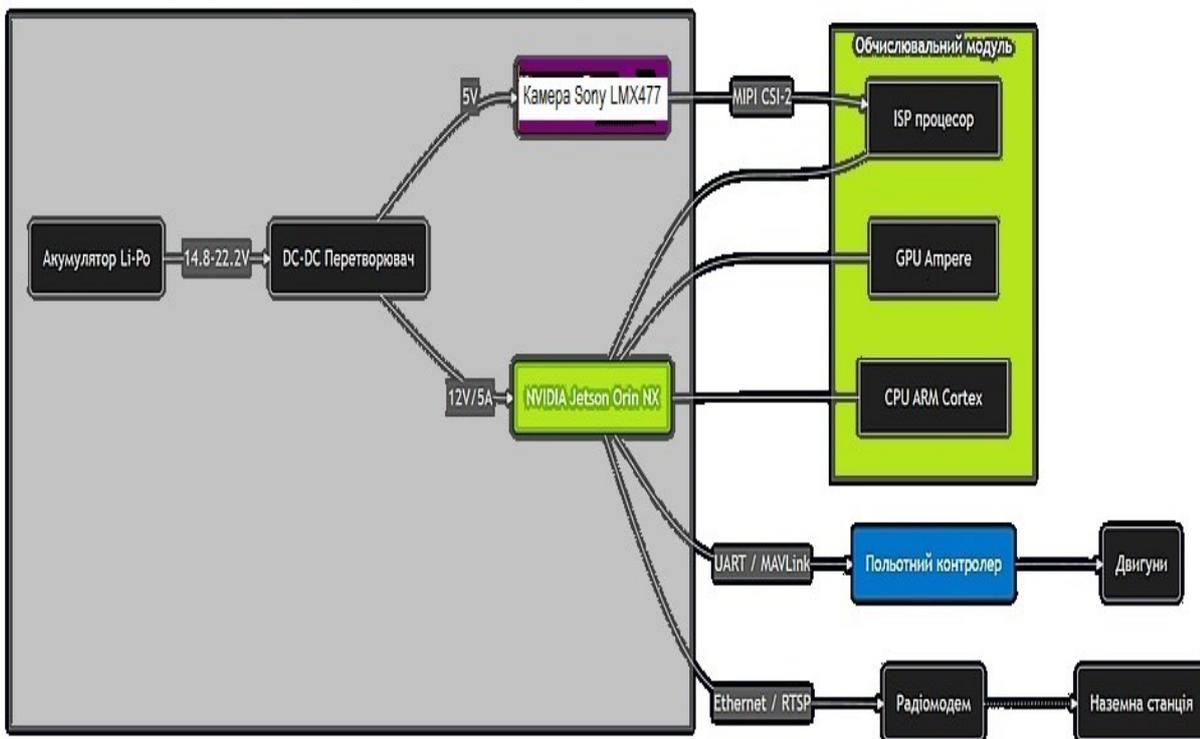


Рис. 1. Структурна схема апаратно-програмного комплексу обробки відеопотоків
Алгоритмічне та програмне забезпечення

Алгоритмічне забезпечення розробленого комплексу орієнтоване на виконання задач детекції об'єктів у режимі реального часу. В якості базового алгоритму використано згорткову нейронну мережу YOLOv8, яка характеризується високим співвідношенням точності та швидкодії. Архітектура YOLOv8 дозволяє виконувати одночасну локалізацію та класифікацію об'єктів за один прохід мережі, що суттєво зменшує обчислювальні витрати.

Для забезпечення роботи в умовах обмежених ресурсів виконано низку оптимізацій. Зокрема, модель перетворено у формат TensorRT, що дало змогу використати апаратні можливості GPU та AI-акселераторів.

Додатково застосовано квантування з плаваючою точністю FP32 до цілочисельного формату INT8, що зменшило обсяг обчислень та використання пам'яті без суттєвої втрати точності детекції.

Обробка відеопотоку реалізована у вигляді конвеєра, який включає захоплення кадру, попередню обробку, інференс нейронної мережі та формування результатів. Такий підхід забезпечує паралельне виконання операцій та мінімізує прості обчислювальних ресурсів. Програмний стек реалізовано на базі ОС Ubuntu 20.04 та SDK NVIDIA JetPack, що забезпечує сумісність із бібліотеками CUDA, cuDNN та TensorRT. Блок-схема алгоритму обробки відеопотоку наведена на рис. 2.

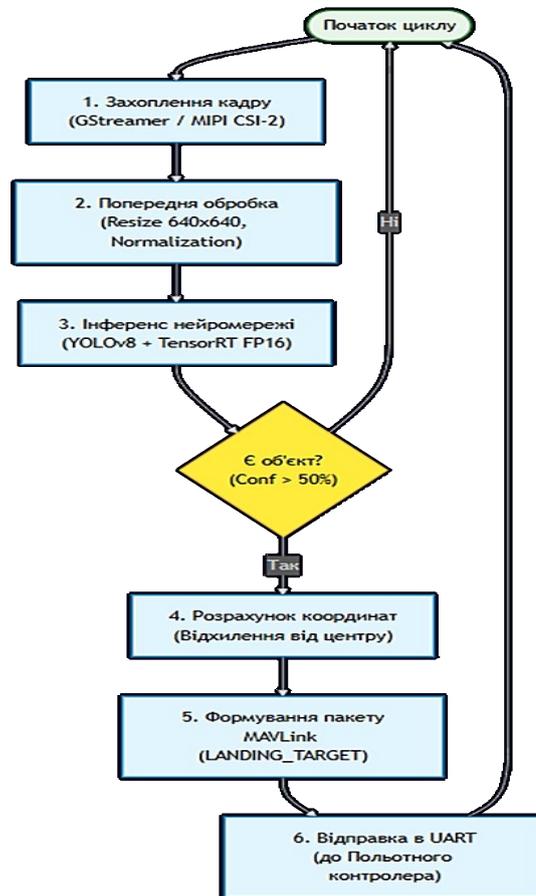


Рис. 2. Блок-схема алгоритму обробки відеопотоку

Експериментальні дослідження

Експериментальні дослідження проведені з метою оцінки обчислювальної продуктивності, часових затримок та енергоспоживання розробленого апаратно-програмного комплексу. Випробування здійснювалися на спеціалізованому лабораторному стенді із застосуванням відеопотоку роздільної здатності 1920×1080 пікселів.

Основним показником ефективності комплексу є частота обробки кадрів (FPS). У процесі тестування встановлено, що після оптимізації нейронної мережі YOLOv8 із застосуванням бібліотеки TensorRT та квантуванням INT8 середня продуктивність становила 54 кадри за секунду. Це значення перевищує мінімальні вимоги до систем реального часу та забезпечує коректну роботу автономної системи спостереження у динамічних умовах [5].

Окрему увагу приділено аналізу сумарної затримки обробки відеоданих. До її складу входять затримка захоплення зображення, час інференсу нейронної мережі та затримка кодування відеопотоку. Експериментально визначено, що загальна затримка не перевищує 40 мс, що є прийнятним для використання у контурах автоматичного керування автономними системами спостереження.

Енергетичні характеристики комплексу досліджувалися шляхом вимірювання споживаної потужності у номінальному режимі роботи. Отримані результати показали, що споживана потужність не перевищує 12 Вт, що дає змогу застосовувати пристрій в автономній літаючій системі спостереження без суттєвого зменшення тривалості польоту. Таким чином, експериментальні дослідження підтвердили відповідність розробленого комплексу основним технічним вимогам.

Висновки

Апаратно-програмний комплекс обробки відеопотоків для автономних систем спостереження поєднує високу продуктивність, низьку затримку та енергоефективність. Експериментальні випробування підтвердили працездатність комплексу в режимі реального часу. Середня продуктивність склала 54 кадри за секунду, а сумарна затримка обробки відеоданих не перевищила 40 мс. Споживана потужність у номінальному режимі не перевищила 12 Вт, що відповідає вимогам до автономних систем спостереження. Отримані результати підтверджують доцільність застосування гетерогенних SoC-платформ для реалізації комплексів технічного зору в сучасних автономних системах спостереження.

Запропоноване рішення дасть змогу, з незначними зусиллями, перетворити звичайну автономну систему спостереження на інтелектуальну систему на сучасній апаратній платформі, стійку до засобів створення радіоелектронних завад.

Література

1. Sullivan, G. J., Ohm, J.-R., Han, W.-J., Wiegand, T. Overview of the High Efficiency Video Coding (HEVC) Standard. *IEEE Transactions on Circuits and Systems for Video Technology*, – 2012 Torralba, Antonio. Foundations of Computer Vision (Adaptive Computation and Machine Learning series)/ Antonio Torralba, Phillip Isola, William T. Freeman. – Cambridge, Massachusetts, USA: The MIT Press, – 2024. [Електронний ресурс]. – Режим доступу: <https://visionbook.mit.edu> – (Дата звернення 20.12.2025 р.).
2. Sze, V., Chen, Y.-H., Yang, T.-J., Emer, J. *Efficient Processing of Deep Neural Networks*. San Rafael, CA: Morgan & Claypool Publishers, – 2020. – 295p.
3. NVIDIA Corporation. *NVIDIA Jetson Orin NX Technical Specifications*. [Електронний ресурс]. – Режим доступу: <https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-orin> – (Дата звернення 20.12.2025 р.).
4. Ultralytics. *YOLOv8 Documentation*. [Електронний ресурс]. – Режим доступу: <https://docs.ultralytics.com> – (Дата звернення: 20.12.2025 р.).
5. Sullivan, G. J., Ohm, J.-R., Han, W.-J., Wiegand, T. Overview of the High Efficiency Video Coding (HEVC) Standard. *IEEE Transactions on Circuits and Systems for Video Technology*, – 2012. – Vol. 22, – No. 12, P. 1649–1668. – DOI: 10.1109/TCSVT.2012.2221191

References

1. Torralba, Antonio. Foundations of Computer Vision (Adaptive Computation and Machine Learning series)/ Antonio Torralba, Phillip Isola, William T. Freeman. – Cambridge, Massachusetts, USA: The MIT Press, – 2024. [Electronic resource]. – Rezhym dostupu: <https://visionbook.mit.edu> – (Data zvernennia 20.12.2025 r.).
2. Sze, V., Chen, Y.-H., Yang, T.-J., Emer, J. *Efficient Processing of Deep Neural Networks*. San Rafael, CA: Morgan & Claypool Publishers. – 2020. – 295 p.
3. NVIDIA Corporation. *NVIDIA Jetson Orin NX Technical Specifications*. [Electronic resource]. – Rezhym dostupu: <https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-orin/> – (Data zvernennia 20.12.2025 r.).
4. Ultralytics. *YOLOv8 Documentation*. [Electronic resource]. – Rezhym dostupu: <https://docs.ultralytics.com> – (Data zvernennia 20.12.2025 r.).
5. Sullivan, G. J., Ohm, J.-R., Han, W.-J., Wiegand, T. Overview of the High Efficiency Video Coding (HEVC) Standard. *IEEE Transactions on Circuits and Systems for Video Technology*, – 2012. – Vol. 22, – No.12, P. 1649–1668. – DOI: 10.1109/TCSVT.2012.2221191