

РОМАНЕЦЬ ВЛАДИСЛАВ

Вінницький національний технічний університет
<https://orcid.org/0009-0009-6075-0210>
e-mail: romanets.vlad@gmail.com

БІСІКАЛО ОЛЕГ

Вінницький національний технічний університет
<https://orcid.org/0000-0002-7607-1943>
e-mail: obisikalo@vntu.edu.ua

ВИЯВЛЕННЯ ОБ'ЄКТІВ ДОРОЖНЬОГО РУХУ З КАМЕР ВІДЕОСПОСТЕРЕЖЕННЯ

У статті досліджено задачу виявлення об'єктів дорожнього руху на основі даних із камер відеоспостереження. Проведено аналіз методів виявлення об'єктів: від класичних підходів, що використовують вручну створені ознаки (SIFT, HOG, LBP) і класифікатори SVM, до сучасних методів на основі одноетапних згорткових нейронних мереж YOLO та трансформерів.

Метою дослідження є оптимізація архітектури нейронної мережі YOLOv8 для розпізнавання об'єктів дорожнього руху, зокрема транспортних засобів, із камер відеоспостереження. Особливу увагу приділено забезпеченню високої якості виявлення в складних умовах, таких як нічний час або низька якість зображень.

Для оцінки ефективності запропонованого підходу створено набір даних із 1400 зображень, отриманих із відкритих джерел. Зображення охоплюють різні часи доби, несприятливі погодні умови та низьку якість, що відображає реальні сценарії роботи камер відеоспостереження.

Для оптимізації архітектури YOLOv8-nano запропоновано замінити окремі блоки C2f модулем C3Ghost у поєднанні з блоком уваги CBAM. Це забезпечило баланс між якістю виявлення та обчислювальною складністю моделі. У результаті оптимізації кількість параметрів зменшено на 26%, а обчислювальна складність (GFLOPS) — на 10%. При цьому якість виявлення залишилася порівнянною з базовою моделлю: за показником mAP@0.50:0.95 знизилася незначно — з 0.531 до 0.525.

Результати дослідження підтверджують перспективність використання модифікованої архітектури YOLOv8-nano для задач відеоспостереження в інтелектуальних транспортних системах.

Ключові слова: виявлення об'єктів, відеоспостереження, трафік, yolov8, C3Ghost, cbam

ROMANETS VLADYSLAV, OLEG BISIKALO

Vinnitsia National Technical University

DETECTION OF TRAFFIC OBJECTS FROM VIDEO SURVEILLANCE CAMERAS

The article investigates the problem of detecting traffic objects based on data from video surveillance cameras. The analysis of object detection methods is carried out: from classical approaches using manually created features (SIFT, HOG, LBP) and SVM classifiers, to modern methods based on single-stage convolutional neural networks YOLO and transformers.

The aim of the study is to optimize the architecture of the YOLOv8 neural network for recognizing traffic objects, specifically vehicles, from video surveillance cameras. The vehicle class includes cars, buses, and trucks. Particular attention is paid to ensuring high detection quality in difficult conditions, such as night time, adverse weather conditions like rain, fog, or snow, and low image quality. To achieve this, various modifications to the network architecture were explored to balance detection accuracy and computational efficiency, making the model suitable for real-time applications.

To evaluate the effectiveness of the proposed approach, a dataset of 1400 images obtained from open sources was created. The images cover different times of day, including dawn, dusk, and nighttime, as well as adverse weather conditions and low-quality images, reflecting real-world scenarios encountered by video surveillance cameras. The dataset was annotated to identify the vehicle class, encompassing cars, buses, and trucks, ensuring robust evaluation under challenging conditions.

To optimize the YOLOv8-nano architecture, it was proposed to replace individual C2f blocks with the C3Ghost module in combination with the CBAM attention block. This modification provided a balance between detection quality and computational complexity. As a result of the optimization, the number of parameters was reduced by 26%, and the computational complexity (GFLOPS) was reduced by 10%. At the same time, the detection quality remained comparable to the base model: the average accuracy (mAP@0.50:0.95) decreased slightly from 0.531 to 0.525.

The results of the study confirm the promising use of the modified YOLOv8-nano architecture for video surveillance tasks in intelligent transport systems. The proposed approach can be applied in traffic monitoring systems, autonomous vehicle navigation, and other domains requiring fast and accurate video data processing in real time. Future work may focus on further refining the model to handle additional vehicle classes, such as motorcycles and bicycles, to enhance its applicability in diverse scenarios. Additionally, integrating the model into embedded systems for widespread deployment and exploring optimizations for specific camera types or extreme conditions could further improve performance.

Keywords: object detection, video surveillance, traffic, yolov8, C3Ghost, cbam

Стаття надійшла до редакції / Received 27.05.2025

Прийнята до друку / Accepted 26.06.2025

Вступ

Виявлення об'єктів дорожнього руху за допомогою камер спостереження, а також відстеження й ідентифікація інцидентів є ключовими технологіями в сучасних інтелектуальних транспортних системах (ІТС). Із зростанням кількості транспортних засобів і підвищенням потреби в безпеці на дорогах автоматизовані системи аналізу відеопотоків стають невід'ємною складовою управління дорожнім рухом.

Ефективність ІТС значною мірою залежить від якості даних, зібраних сенсорами, та продуктивності алгоритмів їхньої автоматичної обробки. Такі системи забезпечують розпізнавання транспортних засобів, пішоходів, дорожніх знаків та інших об'єктів у реальному часі, що дозволяє оперативно реагувати на потенційні загрози, оптимізувати рух транспорту й зменшувати кількість аварій [1].

Камери спостереження за дорожнім рухом є одними з найпоширеніших пристроїв у системах ІТС. Їхня популярність пояснюється економічною вигодою, простотою встановлення, а також здатністю надавати деталізовані візуальні дані з широкою зоною охоплення. Завдяки стрімкому розвитку комп'ютерного зору сучасні ІТС можуть ефективно обробляти відеодані з цих камер, значно підвищуючи якість управління дорожнім рухом.

Аналіз останніх публікацій

Ранні підходи до виявлення об'єктів дорожнього руху базувалися на класичних методах комп'ютерного зору. Одним із популярних інструментів був алгоритм Viola-Jones, який використовував Хаар-подібні ознаки, каскадну архітектуру та Adaboost класифікатор для швидкого виявлення об'єктів. Хоча початково цей алгоритм був розрахований на виявлення облич, автори робіт [2, 3] адаптували його виявлення автомобілів у камерах трафіку. Цей метод вирізняється високою швидкістю обробки, але має обмеження в точності, особливо в умовах змінного освітлення чи складного фону. Модифікація цього підходу, що дещо компенсувала втрату точності завдяки попередній обробці зображень, розглядалася в роботі [4].

Інша група методів для вилучення ознак використовувала такі дескриптори зображень як SIFT, HOG чи LBP для аналізу яскравості, кольору, текстур та контурів, з подальшою класифікацією за допомогою класифікатора SVM, бустінгу, байєсівських мереж, чи нейронних мереж. Зокрема в роботі [5] автори модифікували дескриптор SIFT для класифікації транспортних засобів у відео трафіка, В роботі [6] запропоновано метод виявлення перешкод на трасах при використанні дескриптора HOG та класифікатора SVM. Недоліком таких підходів є необхідність ретельного налаштування параметрів та чутливість до шумів.

Ще одна група методів базувалися на аналізі руху через оптичний потік або сегментацію фону. Наприклад, алгоритми типу гусівських сумішів (Gaussian Mixture Model, GMM) використовувалися для відокремлення рухомих об'єктів (автомобілів, пішоходів) від статичного тла. Зокрема, автори в роботі [7] використали GMM для виявлення транспортних засобів та фільтр Калмана для їх відстеження. Хоча ці методи були економічними з точки зору обчислювальних ресурсів, їхня ефективність знижувалася в умовах щільного трафіку чи при наявності перешкод, таких як тіні чи погодні явища.

З появою глибокого навчання методи виявлення об'єктів досягли значного прогресу. Згорткові нейронні мережі (CNN) стали основою сучасних систем розпізнавання об'єктів. Однією з перших моделей, призначених для виявлення та класифікації об'єктів дорожнього руху з камер відеоспостереження в реальному часі, стала нейронна мережа DetectNet, розроблена компанією NVIDIA. Ця мережа є адаптацією архітектури GoogLeNet для задачі детекції об'єктів [8]. У роботах [9, 10] автори застосували DetectNet для виявлення автомобілів, використовуючи середовище NVIDIA DIGITS для моделювання.

Важливим етапом у розвитку методів виявлення об'єктів стала архітектура R-CNN (Region-based CNN), яка поєднувала генерацію регіональних пропозицій із класифікацією на основі згорткових нейронних мереж, реалізуючи двоетапний підхід. Подальші вдосконалення, такі як Fast R-CNN і Faster R-CNN, значно підвищили швидкість і точність, зробивши ці моделі придатними для використання в реальному часі. У роботі [11] автори застосували Faster R-CNN для виявлення та класифікації транспортних засобів у відеопотоці дорожнього руху.

Одночасно з двоетапними розвивалися одноетапні методи виявлення об'єктів, такі як YOLO (You Only Look Once) і SSD (Single Shot MultiBox Detector). YOLO, зокрема, здобув популярність у задачах ІТС завдяки своїй здатності обробляти відеопотоки в реальному часі з високою точністю. У роботі [12] автори запропонували підхід до підрахунку кількості транспортних засобів у відеопотоці дорожнього руху на основі детектора SSD. У дослідженні [13] для відстеження транспортних засобів у відео автори застосували алгоритм DeepSort, використавши PP-YOLO як детектор. У роботі [14] для відстеження транспортних засобів і пішоходів було використано алгоритм SORT у поєднанні з YOLOv5 як детектором об'єктів. У дослідженні [15] автори здійснювали виявлення нещасних випадків на дорогах у відеопотоці трафіку, застосовуючи модель YOLOv8.

Трансформери, які спочатку застосовувалися в обробці природної мови, також почали інтегруватися в комп'ютерний зір. Моделі на основі Vision Transformer (ViT) і їхні похідні, такі як DETR (DEtection TRansformer), показали високу ефективність у задачах виявлення об'єктів із складними взаємозв'язками, наприклад, при аналізі щільних дорожніх сцен із численними учасниками руху. Зокрема у роботі [16] автори запропонували використовувати Swin трансформер разом з YOLOv5 та модуля уваги SimAM для покращення виявлення транспортних засобів в умові щільних дорожніх сцен.

Традиційні методи, хоча й поступаються за точністю, досі використовуються в системах із обмеженими обчислювальними ресурсами, наприклад, у вбудованих пристроях. Натомість методи глибокого навчання домінують у високопродуктивних ІТС, де доступні потужні GPU чи TPU. Одними із ключових задач залишається адаптація моделей до різних умов – нічного часу, поганої погоди чи низької якості зображень а також адаптація моделей глибоких нейронних мереж до застосування в системах із обмеженими обчислювальними ресурсами.

Формулювання цілей статті

Мета дослідження полягає в оптимізації архітектури нейронної мережі YOLOv8 для розпізнавання об'єктів дорожнього руху з камер відеоспостереження, зокрема транспортних засобів, при збереженні високого рівня виявлення об'єктів, зокрема в нічний час чи за умов низької якості зображень.

Для досягнення зазначеної мети визначено такі основні завдання дослідження: провести аналіз сучасних методів виявлення об'єктів дорожнього руху, модифікувати архітектуру моделі YOLOv8, здійснити навчання та тестування моделі. Результати дослідження можна буде використовувати у камерах відеоспостереження ІТС.

Основний виклад матеріалу

За основу для удосконалення у роботі використана архітектура моделі YOLOv8. Вона є сучасною одноетапною моделлю детекції об'єктів, що поєднує ефективність та високу точність завдяки використанню вдосконаленої модульної архітектури. Основою мережі є C2f-блоки, які забезпечують компактність та глибину представлення ознак. Neck-частина моделі базується на комбінації FPN і PAN, що дозволяє ефективно агрегувати ознаки з різних рівнів. У частині head застосовується anchor-free підхід, який спрощує навчання та підвищує узагальнювальну здатність моделі. Крім того, YOLOv8 використовує вдосконалену функцію втрат CIOU та підтримує масштабування архітектури під різні ресурси, що робить її придатною для широкого спектру застосувань у реальному часі.

Для зменшення обчислювальної складності та кількості параметрів моделі YOLOv8-nano при збереженні продуктивності запропоновано використати модуль C3Ghost, що є модифікацією блоку C3, використаного в YOLOv5. C3Ghost базується на Ghost-згортках [17], які зменшують надлишковість ознакових карт. Замість повної згортки створюється обмежена кількість первинних ознакових карт стандартною згорткою, які потім трансформуються у більше "ghost"-ознак за допомогою економних глибинних згорток (3×3 або 5×5). Це дозволяє генерувати більше ознак із меншими витратами ресурсів. У C3Ghost стандартні Bottleneck-блоки замінено на GhostBottleneck.

Для покращення розпізнавання об'єктів дорожнього руху в складних умовах (ніч, дощ, низька якість зображень) до C3Ghost пропонується інтегрувати модуль уваги CBAM [18]. Модуль CBAM призначений для посилення як каналної, так і просторової уваги. Канальна увага дозволяє моделі визначати найбільш інформативні канали та надавати їм більшу вагу, тоді як просторова увага фокусується на ключових областях зображення, ігноруючи менш релевантні. Зокрема, канальна увага може допомогти краще розпізнавати важливі ознаки транспортних засобів, такі як форма чи текстура. Просторова увага, в свою чергу, може зменшити вплив фону (наприклад, дерев або неба), концентруючись на самих автомобілях, що буде сприяти підвищенню якості виявлення. Пропонується додати модуль CBAM у C3Ghost після конкатенації перед останньою згорткою. Це дозволить фільтрувати надлишкову інформацію після об'єднання ознакових карт та компенсувати втрати від економних Ghost-згорток, підсилюючи релевантні ознаки. Графічне представлення запропонованого модуля C3Ghost-CBAM наведено на рисунку 1.

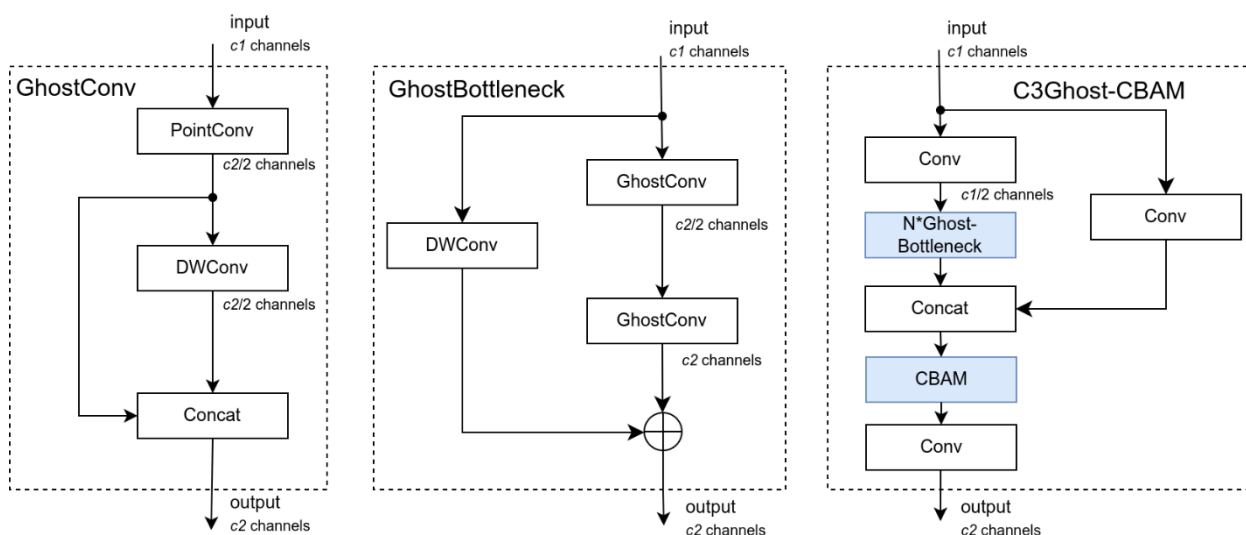


Рис. 1 – C3Ghost-CBAM модуль

Пропонується модифікувати базову архітектуру YOLOv8 шляхом додавання блоків C3Ghost-CBAM замість C2f у Neck частину а також полегшеної згортки GhostConv. Крім того для спрощення архітектури пропонується замінити один модуль C2f у Backbone частину перед модулем SPPF.

Запропонована модифікація YOLOv8 представлена на рисунку 2.

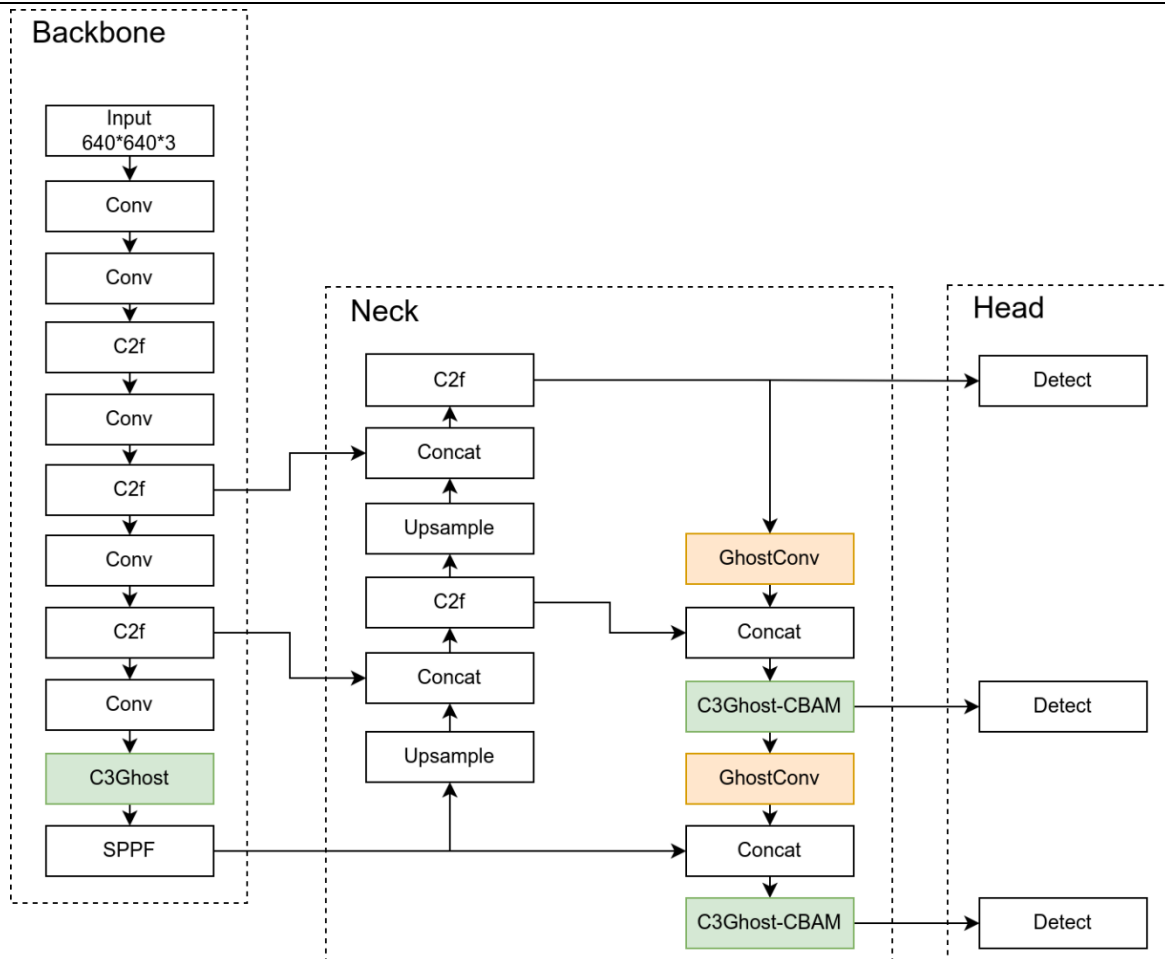


Рис. 2 – Запропонована архітектура нейронної мережі

Дослідження ефективності запропонованого методу

Для оцінки ефективності запропонованого методу було створено набір зображень з камер відеоспостереження трафіку, які отримані з відкритих джерел. Датасет складає 1400 зображень, які були поділені на тренувальні, валідаційні та тестові у співвідношенні - 7:2:1. Велика частина зображень характеризуються складними умовами, зокрема, отриманими в різний час доби, за умов поганої погоди чи характеризуються низькою якістю. Типові приклади зображень отриманого датасету наведені на рисунку 3.



Рис. 3 – Приклади зображень створеного датасету

Для експериментів була використана така апаратна конфігурація – графічна карта NVIDIA GeForce RTX 1080 з 8 ГБ, фреймворк глибинного навчання PyTorch версії 2.5.1, Python версії 3.11, CUDA версії 11.0 та операційна система Ubuntu 20.04.

Максимальна кількість епох для тренування набору даних становить 100, з першими трьома епохами для розігріву. Для регулювання швидкості навчання використовується оптимізатор AdamW з початковою швидкістю навчання 0.02. Через наявність об'єктів малого розміру у наборі даних, а також для забезпечення балансу між якістю та швидкістю виявлення у пристроях з обмеженими ресурсами розмір кожного зображення пропонується нормалізувати до 640×640 пікселів.

Для оцінки ефективності моделі виявлення транспортних засобів в цій роботі використовувалася метрика mean Average Precision (mAP). Обчислювальну складність моделі оцінювали за такими показниками, як кількість параметрів і обсяг операцій з плаваючою комою (FLOPs). Оскільки в дослідженні розглядався лише один клас (транспортний засіб), значення mAP еквівалентне значенню Average Precision (AP) для цього класу.

Метрика AP визначається як площа під кривою Precision-Recall і розраховується наступним чином:

$$AP = \int_0^1 \text{Precision}(\text{Recall})d(\text{Recall}) \quad (1)$$

Precision відображає частку правильних детекцій серед усіх позитивних передбачень моделі, а Recall – частку правильних детекцій серед усіх реальних об'єктів класу “транспортний засіб”. Рівняння для обчислення Precision та Recall представлені на у формулах (2) та (3) відповідно:

$$\text{Precision} = \frac{TP}{TP+FP}, \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN}, \quad (3)$$

де TP позначає кількість правильно виявлених позитивних зразків, FN позначає кількість неправильно виявлених негативних зразків, а FP вказує на кількість неправильно виявлених позитивних зразків.

У дослідженні використовувалися два показники mAP: mAP@0.5, що відповідає значенню mAP при порозі IoU (Intersection over Union), рівному 0.5, та mAP@0.5:0.95, яке є середнім значенням mAP, розрахованим для порогів IoU від 0.5 до 0.95 з кроком 0.05.

На основі створеного набору зображень, були проведені експерименти з виявлення транспортних засобів у зображеннях з камер відеоспостереження трафіку, для дослідження впливу кожного доданого модуля на загальну продуктивність моделі. За базову взята оригінальна модель YOLOv8n, модель 1 включала модуль C3Ghost замість модуля C2f як у backbone так і у neck відповідно до рисунка 2. У моделі 2 окрім C3Ghost модулів з моделі 1 замість двох Conv блоків у neck були використані GhostConv блоки, у моделі 3 були використані GhostConv у neck та доданий CBAM у C3Ghost модулі лише у neck. Зміни показників оцінювання для чотирьох моделей були кількісно досліджені та порівнянні. Результати оцінювання нових моделей на створеному наборі даних наведені в таблиці 1, де найкращі результати для кожного показника оцінювання були позначені жирним.

Таблиця 1

Результати експериментів з вилученням модулів

Модель	C3Ghost	GhostConv	CBAM	mAP@(0.50-0.95)	mAP@0.50	Кількість параметрів, млн	GFLOPS
YOLOv8-nano	-	-	-	0.531	0.817	3.0	8.1
1	+	-	-	0.518	0.809	2.3	7.3
2	+	+	-	0.515	0.805	2.2	7.2
3	+	+	+	0.525	0.810	2.3	7.3

Згідно з таблицею 1, базова модель YOLOv8-nano показала найвищу якість виявлення: mAP@(0.50:0.95) — 0.531, mAP@0.50 — 0.817, при 3.0 млн параметрів і 8.1 GFLOPS. Модель 2 досягла найменшої обчислювальної складності: 2.2 млн параметрів, 7.2 GFLOPS, але з нижчим mAP@(0.50:0.95) — 0.515 і mAP@0.50 — 0.805. Модель 3 забезпечила оптимальний баланс: mAP@(0.50:0.95) — 0.525, mAP@0.50 — 0.810, 2.3 млн параметрів і 7.3 GFLOPS.

Результати роботи моделі 3 на тестових зображеннях наведені на рисунку 4.

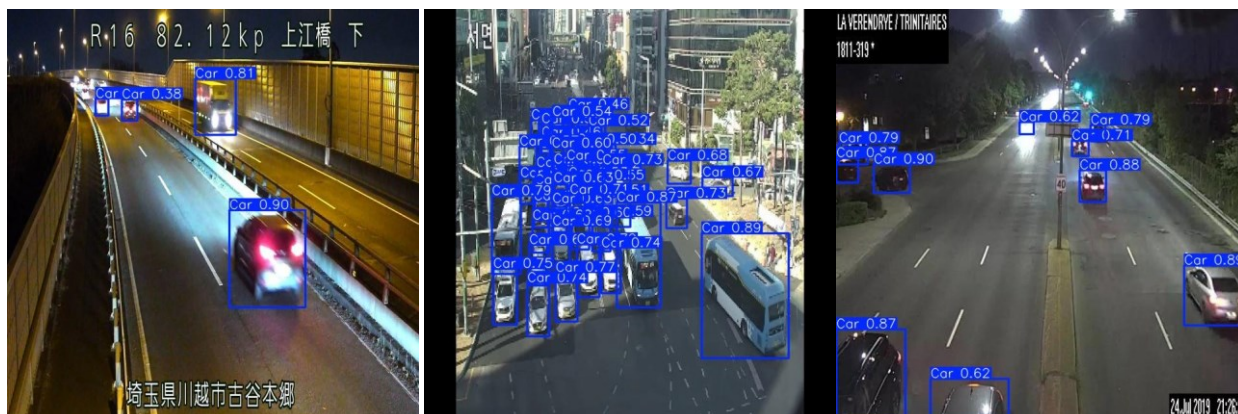


Рис. 4. Приклади детекції об'єктів дорожнього руху

Висновки

Аналіз літературних джерел показав, що ключовими задачами при виявленні об'єктів дорожнього руху в інтелектуальних транспортних системах є адаптація моделей до складних умов – нічного часу, несприятливої погоди чи низької якості зображень, а також оптимізація глибоких нейронних мереж для використання в системах з обмеженими обчислювальними ресурсами. У зв'язку з цим фокус дослідження було спрямовано на оптимізацію моделей виявлення об'єктів із фокусом на складні умови отримання зображень.

Застосування модуля C3Ghost замість C2f в оригінальній архітектурі YOLOv8 дозволило зменшити кількість параметрів на 26% і обчислювальну складність на 10%. Для підвищення здатності моделі фокусуватися на об'єктах дорожнього руху та ігнорувати нерелевантний фон запропоновано інтегрувати модуль уваги CBAM у блок C3Ghost. Це компенсувало потенційні втрати інформації, спричинені використанням C3Ghost. У результаті якість виявлення запропонованої моделі залишилася порівнянною з базовою: середня точність (mAP@0.50:0.95) знизилася незначно – з 0.531 до 0.525.

Розроблена модифікована архітектура YOLOv8-nano є перспективною для використання в задачах відеоспостереження в інтелектуальних транспортних системах.

Література

1. Ghahremannezhad, H., Shi, H., & Liu, C. (2023). Object detection in traffic videos: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 24(7), 6780–6799.
2. Choudhury, S., Chattopadhyay, S. P., & Hazra, T. K. (2017). Vehicle detection and counting using haar feature-based classifier. In *2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON)* (pp. 106–109). IEEE.
3. Ramadhani, M. I., Minarno, A. E., & Cahyono, E. B. (2018). Vehicle classification using haar cascade classifier method in traffic surveillance system. *Kinetik*, 3(1), 57–64.
4. Kvyetnyy, R., Maslii, R., Harmash, V., Bogach, I., Kotyra, A., Grądz, Ż., Zhanpeisova, A., & Askarova, N. (2017). Object detection in images with low light condition. In *Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2017* (Vol. 10445, 104450W). SPIE. <https://doi.org/10.1117/12.2281001>
5. Ma, X., & Grimson, W. E. L. (2005). Edge-based rich representation for vehicle classification. In *Tenth IEEE International Conference on Computer Vision (ICCV'05)* (Vol. 2, pp. 1185–1192). IEEE.
6. Lan, J., Jiang, Y., Fan, G., Yu, D., & Zhang, Q. (2016). Real-time automatic obstacle detection method for traffic surveillance in urban traffic. *Journal of Signal Processing Systems*, 82(3), 357–371.
7. Yusliana, B. R., Sari, A. I., & Prayogi, A. A. (2016). Vehicle detection and tracking using Gaussian mixture model and Kalman filter. In *2016 International Conference on Computational Intelligence and Cybernetics*. IEEE.
8. NVIDIA. *NVIDIA DIGITS* [Electronic resource]. <https://developer.nvidia.com/digits>
9. Коцюбинський, В., Кветний, Р., Кириленко, О., Бісікало, О., & Маслій, Р. (2018). Розв'язання задачі підрахунку об'єктів при складних умовах отримання зображень. У *Інформаційні технології та комп'ютерне моделювання: матеріали міжнар. наук.-практ. конф.* (с. 44–47). Івано-Франківськ.
10. Кветний, Р. Н., Маслій, Р. В., & Кириленко, О. М. (2020). Виявлення та класифікація об'єктів дорожнього руху при використанні середовища DIGITS. *Оптико-електронні інформаційно-енергетичні технології*, 39(1), 14–20.
11. Faruque, M. O., Ghahremannezhad, H., & Liu, C. (2019). Vehicle classification in video using deep learning. In *Machine Learning and Data Mining in Pattern Recognition (MLDM)* (pp. 117–131). ibai publishing.
12. Dai, Z., Song, H., Wang, X., Fang, Y., Yun, X., Zhang, Z., & Li, H. (2019). Video-based vehicle counting framework. *IEEE Access*, 7, 64460–64470.
13. Lu, J., Xia, M., Gao, X., Yang, X., Tao, T., Meng, H., & Ding, E. (2021). Robust and online vehicle counting at crowded intersections. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4002–4008).
14. Rezaei, M., Azarmi, M., & Mir, F. M. P. (2021). Traffic-net: 3D traffic monitoring using a single camera. *arXiv preprint. arXiv:2109.09165*
15. Minh, Q. N. H., Dinh, N. N., Ho, L. V., & Huu, C. P. (2025). Real-time traffic accident detection using YOLOv8. *Transportation Research Procedia*, 85, 68–75.
16. Zhang, Y., & Sun, Z. (2024). An advanced YOLOv5s approach for vehicle detection integrating Swin Transformer and SimAM in dense traffic surveillance. *Journal of Industrial Intelligence*, 2(1), 31–41.
17. Han, K., Wang, Y., Tian, Q., Guo, J., Xu, C., & Xu, C. (2020). GhostNet: More features from cheap operations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1580–1589).
18. Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 3–19). Munich, Germany.

References

1. Ghahremannezhad, H., Shi, H., & Liu, C. (2023). Object detection in traffic videos: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 24(7), 6780–6799.
2. Choudhury, S., Chattopadhyay, S. P., & Hazra, T. K. (2017). Vehicle detection and counting using haar feature-based classifier. In *2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON)* (pp. 106–109). IEEE.
3. Ramadhani, M. I., Minarno, A. E., & Cahyono, E. B. (2018). Vehicle classification using haar cascade classifier method in traffic surveillance system. *Kinetik*, 3(1), 57–64.
4. Kvyetnyy, R., Maslii, R., Harmash, V., Bogach, I., Kotyra, A., Grądz, Ż., Zhanpeisova, A., & Askarova, N. (2017). Object detection in images with low light condition. In *Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2017* (Vol. 10445, 104450W). SPIE. <https://doi.org/10.1117/12.2281001>
5. Ma, X., & Grimson, W. E. L. (2005). Edge-based rich representation for vehicle classification. In *Tenth IEEE International Conference on Computer Vision (ICCV'05)* (Vol. 2, pp. 1185–1192). IEEE.
6. Lan, J., Jiang, Y., Fan, G., Yu, D., & Zhang, Q. (2016). Real-time automatic obstacle detection method for traffic surveillance in urban traffic. *Journal of Signal Processing Systems*, 82(3), 357–371.
7. Yusliana, B. R., Sari, A. I., & Prayogi, A. A. (2016). Vehicle detection and tracking using Gaussian mixture model and Kalman filter. In *2016 International Conference on Computational Intelligence and Cybernetics*. IEEE.
8. NVIDIA. NVIDIA DIGITS [Electronic resource]. <https://developer.nvidia.com/digits>
9. Kotsiubynskyi, V., Kvietyi, R., Kyrylenko, O., Bisikalo, O., & Maslii, R. (2018). Rozviazannia zadachi pidrakhunku ob'ektiv pry skladnykh umovakh otrymannia zobrazhen. U Informatsiini tekhnolohii ta kompiuterni modeliuvannia: materialy mizhnar. nauk.-prakt. konf. (s. 44–47). Ivano-Frankivsk.
10. Kvietyi, R. N., Maslii, R. V., & Kyrylenko, O. M. (2020). Vyiavleniia ta klasyfikatsiia ob'ektiv dorozhnoho rukhu pry vykorystanni seredovyscha DIGITS. *Optyko-elektronni informatsiino-enerhetychni tekhnolohii*, 39(1), 14–20.
11. Faruque, M. O., Ghahremannezhad, H., & Liu, C. (2019). Vehicle classification in video using deep learning. In *Machine Learning and Data Mining in Pattern Recognition (MLDM)* (pp. 117–131). ibai publishing.
12. Dai, Z., Song, H., Wang, X., Fang, Y., Yun, X., Zhang, Z., & Li, H. (2019). Video-based vehicle counting framework. *IEEE Access*, 7, 64460–64470.
13. Lu, J., Xia, M., Gao, X., Yang, X., Tao, T., Meng, H., & Ding, E. (2021). Robust and online vehicle counting at crowded intersections. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4002–4008).
14. Rezaei, M., Azarmi, M., & Mir, F. M. P. (2021). Traffic-net: 3D traffic monitoring using a single camera. *arXiv preprint. arXiv:2109.09165*
15. Minh, Q. N. H., Dinh, N. N., Ho, L. V., & Huu, C. P. (2025). Real-time traffic accident detection using YOLOv8. *Transportation Research Procedia*, 85, 68–75.
16. Zhang, Y., & Sun, Z. (2024). An advanced YOLOv5s approach for vehicle detection integrating Swin Transformer and SimAM in dense traffic surveillance. *Journal of Industrial Intelligence*, 2(1), 31–41.
17. Han, K., Wang, Y., Tian, Q., Guo, J., Xu, C., & Xu, C. (2020). GhostNet: More features from cheap operations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1580–1589).
18. Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 3–19). Munich, Germany.