

АДАПТАЦІЯ СУЧАСНИХ МОДЕЛЕЙ КОМП'ЮТЕРНОГО ЗОРУ (YOLOv8-nano) ДЛЯ РОБОТИ В РЕАЛЬНОМУ ЧАСІ В СИСТЕМАХ ВІДЕОАНАЛІТИКИ ДОРОЖНЬОГО РУХУ

Вінницького національного технічного університету

Анотація

У роботі розглянуто підхід до оптимізації моделі комп'ютерного зору YOLOv8-nano для задач відеоаналітики дорожнього руху в режимі реального часу. Запропоновано використання полегшених архітектурних компонентів, зокрема Ghost-модулів та механізму уваги CBAM, що дозволяє зменшити обчислювальну складність моделі при збереженні прийнятної рівня точності детекції. Проведено експериментальне дослідження впливу архітектурних модифікацій на точність виявлення об'єктів, кількість параметрів та обчислювальну складність моделі. Отримані результати підтверджують доцільність використання комбінованої оптимізації архітектури для систем відеоаналітики дорожнього руху, що працюють на edge-платформах.

Ключові слова: комп'ютерний зір, YOLOv8, відеоаналітика, детекція об'єктів, інтелектуальні транспортні системи, edge-обчислення.

Abstract

The paper considers an approach to optimizing the YOLOv8-nano computer vision model for real-time traffic video analytics tasks. The use of lightweight architectural components, in particular Ghost modules and the CBAM attention mechanism, is proposed, which allows reducing the computational complexity of the model while maintaining an acceptable level of detection accuracy. An experimental study of the impact of architectural modifications on the accuracy of object detection, the number of parameters, and the computational complexity of the model was conducted. The results obtained confirm the feasibility of using combined architecture optimization for traffic video analytics systems operating on edge platforms.

Keywords: computer vision, YOLOv8, video analytics, object detection, intelligent transportation systems, edge computing.

Вступ

Сьогодні розвиток інтелектуальних транспортних систем суттєво змінює підходи до організації дорожнього руху та забезпечення безпеки на автошляхах. Одним із ключових напрямів цього розвитку є використання технологій комп'ютерного зору для автоматизованого аналізу відеопотоків з камер дорожнього спостереження. Практичні аспекти впровадження таких систем у реальних умовах розглянуто у роботах [1–4]. Водночас практичне впровадження таких систем ускладнюється необхідністю обробки даних у режимі реального часу та обмеженими обчислювальними ресурсами на місцях (edge-пристроїв).

Результати дослідження

Сучасні моделі глибокого навчання, зокрема одноетапні детектори об'єктів сімейства YOLO[5], демонструють високу точність виявлення транспортних засобів, однак їх застосування в реальних умовах міської інфраструктури часто супроводжується проблемами швидкодії, особливо за наявності щільного трафіку, часткових оклюзій та змін освітлення. Це зумовлює необхідність адаптації архітектури нейронної мережі з урахуванням специфіки дорожніх сцен [6]. Питання оптимізації моделей детекції для дорожніх сцен детально досліджено в роботі [1].

У сучасних дослідженнях значна увага приділяється оптимізації архітектур комп'ютерного зору з метою досягнення балансу між точністю детекції та швидкістю обробки[6]. Перспективним підходом є використання полегшених архітектур та механізмів уваги, що дозволяють зменшити кількість параметрів моделі без істотної втрати якості. Зокрема, застосування Ghost-модулів дозволяє генерувати додаткові ознаки з меншими обчислювальними витратами [7], а механізми уваги типу CBAM підвищують інформативність просторових і каналних ознак [8].

У межах даної роботи розглядається підхід до адаптації моделі YOLOv8-nano для задач дорожнього моніторингу шляхом заміни стандартних блоків на ефективніші архітектурні рішення у поєднанні з механізмами уваги [1]. Запропонований підхід дозволяє забезпечити стабільну роботу системи в режимі реального часу та підвищити її придатність до використання на edge-платформах.

У проведених експериментах досліджувався вплив окремих архітектурних модифікацій на точність та обчислювальну складність моделі. Зокрема, блоки C3Ghost використовувалися як заміна стандартних C2f-модулів з метою зменшення кількості параметрів за рахунок генерації «дешевих» ознак на основі лінійних перетворень. Модулі GhostConv дозволяють скоротити кількість згорткових операцій, зменшуючи обчислювальні витрати без суттєвої втрати інформативності ознак. Механізм уваги CBAM (Convolutional Block Attention Module) інтегрується для покращення виділення релевантних просторових і каналних ознак, що особливо важливо для дорожніх сцен із частковими оклюзіями та складним фоном.

Таким чином, експерименти спрямовані на дослідження компромісу між скороченням кількості параметрів, зменшенням GFLOPS та збереженням прийнятного рівня точності детекції, відповідні результати експериментів з вилученням модулів представлені у таблиці 1.

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

Результати експериментів показують, що заміна стандартних блоків на C3Ghost дозволяє зменшити кількість параметрів моделі з 3.0 до 2.3 млн та знизити обчислювальну складність із 8.1 до 7.3 GFLOPS при незначному зниженні показника mAP. Додаткове використання GhostConv ще більше скорочує параметри до 2.2 млн, проте супроводжується незначним зменшенням точності. Інтеграція механізму уваги CBAM частково компенсує втрати точності, підвищуючи mAP@(0.50–0.95) до 0.525 при збереженні зниженого рівня обчислювальної складності. Отримані результати узгоджуються з висновками досліджень щодо ефективності полегшених YOLO-моделей для задач інтелектуальних транспортних систем [9], що підтверджує доцільність використання комбінованої архітектурної оптимізації для забезпечення балансу між точністю та швидкістю в задачах реального часу.

Висновки

Отже, адаптація сучасних моделей комп'ютерного зору до умов реальної експлуатації є необхідною передумовою створення ефективних систем відеоаналітики дорожнього руху. Проведена оптимізація архітектури YOLOv8-nano шляхом використання Ghost-модулів та механізмів уваги дозволяє зменшити обчислювальне навантаження без критичної втрати точності. Це створює передумови для впровадження систем автоматизованого моніторингу на edge-платформах із обмеженими ресурсами та забезпечує можливість масштабування рішень на рівні міської інфраструктури.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Романець В., Бісікало О. Виявлення об'єктів дорожнього руху з камер відеоспостереження // Herald of Khmelnytskyi National University. Technical Sciences. 2025. Т. 355, № 4. С. 491–497. <https://doi.org/10.31891/2307-5732-2025-355-70>.
2. Романець В., Маслій Р. В. Виявлення аномалій у відеопотоці трафіку транспорту засобами комп'ютерного зору та глибинного навчання // Вісник ВПІ. 2025. Вип. 5. С. 146–155.
3. Романець В., Бісікало О. Метод багатооб'єктного трекінгу з обробкою оклюзій для систем відеоспостереження // Measuring and Computing Devices in Technological Processes. 2025. Т. 84. № 4. С. 440–445. <https://doi.org/10.31891/2219-9365-2025-84-53>.
4. Романець В., Бісікало О. Information technology in customer service: use of computer vision and decision support systems // SWorldJournal. 2023. Т. 1, № 21-01. С. 65–72. <https://doi.org/10.30888/2663-5712.2023-21-01-046>.
5. Redmon J., Divvala S., Girshick R., Farhadi A. You only look once: Unified, real-time object detection // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2016. – P. 779–788.
6. Lin T.-Y., Dollár P., Girshick R., He K., Hariharan B., Belongie S. Feature pyramid networks for object detection // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2017. – Vol. 30. – P. 2117–2125. – DOI: <https://doi.org/10.1109/CVPR.2017.106>.
7. Han K., Wang Y., Tian Q., Guo J., Xu C., Xu C. GhostNet: More features from cheap operations // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2020. – Vol. 33. – P. 1580–1589. – DOI: <https://doi.org/10.1109/CVPR42600.2020.00165>.
8. Woo S., Park J., Lee J. Y., Kweon I. S. CBAM: Convolutional block attention module // European Conference on Computer Vision (ECCV). – 2018. – Vol. 15. – P. 3–19. – DOI: https://doi.org/10.1007/978-3-030-01234-2_1.
9. Bakirci M. Real-time vehicle detection using YOLOv8-Nano for intelligent transportation systems // Traitement du Signal. – 2024. – Vol. 41, № 4. – P. 1727

Романець Владислав Олегович – аспірант групи 126-22а, факультет інтелектуальних інформаційних технологій та автоматизації, ВНТУ, Вінниця, e-mail: romanets.vlad@gmail.com

Бісікало Олег Володимирович – д-р техн. наук, професор, завідувач кафедри автоматизації та інтелектуальних інформаційних технологій, ВНТУ, м. Вінниця, e-mail: obisikalo@vntu.edu.ua

Маслій Роман Васильович – канд. техн. наук, доцент, доцент кафедри автоматизації та інтелектуальних інформаційних технологій, ВНТУ, м. Вінниця, e-mail: maslij.r.v@vntu.edu.ua

Romanets Vladyslav O. – post-graduate student, group 126-22a, faculty of intellectual information technologies and automation, VNTU, Vinnytsia, e-mail: romanets.vlad@gmail.com

Bisikalo Oleg V. – Doctor of Technical Sciences, Professor of the Department of Automation and Intelligent Information Technologies, VNTU, Vinnytsia, e-mail: obisikalo@vntu.edu.ua

Maslii Roman V. – Candidate of Technical Sciences, associate professor at the Department of Automation and Intelligent Information Technologies, VNTU, Vinnytsia, e-mail: maslij.r.v@vntu.edu.ua