DOI https://doi.org/10.32689/maup.it.2026.1.8
УДК 004.93

КАСКАДНИЙ МЕТОД УДОСКОНАЛЕННЯ СИСТЕМ ДЕТЕКЦІЇ ОБ'ЄКТІВ

Л. Г. Юдіна, Ю. В. Дегтяр

A CASCADE METHOD FOR IMPROVING OBJECT DETECTION SYSTEMS

Lyudmila Yudina, Yuri Degtyar

Анотація

У статті досліджується проблема нестабільності функціонування нейромережевих моделей детекції об'єктів у системах комп'ютерного зору безпілотних літальних апаратів (БпЛА) в умовах динамічної зміни зовнішнього середовища. Показано, що традиційні універсальні моделі, зокрема архітектури сімейства YOLO, демонструють суттєве зниження показників точності (Precision) та повноти (Recall) при зміні погодних умов, освітлення та просторового масштабу об'єктів.

Метою дослідження є підвищення ефективності та стійкості систем детекції об'єктів у реальному часі шляхом розробки адаптивного каскадного методу обробки відеопотоку. Запропоновано підхід Cascade YOLO, що базується на послідовному використанні спеціалізованих нейромережевих моделей із механізмом адаптивного перемикання залежно від рівня впевненості детекції. На відміну від класичних ансамблевих методів, у запропонованому підході повторна обробка кадру не виконується, а наступна модель застосовується до нового кадру, що дозволяє уникнути накопичення затримок.

Методологія дослідження включає побудову математичної моделі каскадного алгоритму, програмну реалізацію та експериментальну перевірку на відеоданих із варіативними умовами спостереження. Проведене порівняння з базовою моделлю Single-YOLO показало підвищення Precision до 0.867 та Recall до 0.824 при незначному зниженні швидкодії (до 5–10 %).

Отримані результати підтверджують ефективність запропонованого підходу для підвищення надійності систем технічного зору БпЛА. Практична цінність полягає у можливості впровадження методу в бортові системи обробки даних для забезпечення стабільної роботи в умовах невизначеності.

Ключові слова: *комп'ютерний зір, безпілотні літальні апарати, YOLO, каскадні алгоритми, детекція об'єктів, складні погодні умови.*

Abstract

The paper addresses the problem of instability of neural network-based object detection systems deployed on unmanned aerial vehicles (UAVs) under dynamically changing environmental conditions. It is shown that conventional universal models, including the YOLO family architectures, demonstrate a significant degradation of key performance metrics such as Precision and Recall when operating under varying weather conditions, illumination changes, and different object scales.

The aim of the study is to improve the efficiency and robustness of real-time object detection systems by developing an adaptive cascade-based processing method. The proposed Cascade YOLO approach is based on sequential utilization of specialized neural network models combined with an adaptive switching mechanism depending on the confidence level of the current detection. Unlike classical ensemble approaches, the proposed method avoids reprocessing the same frame and instead applies the next model to a new incoming frame, which prevents latency accumulation and preserves real-time performance.

The research methodology includes the development of a mathematical model of the cascade algorithm, its software implementation, and experimental validation on heterogeneous video datasets with varying observation conditions. Comparative analysis with a baseline Single-YOLO model demonstrated an increase in Precision up to 0.867 and Recall up t 0.824, with only a minor decrease in processing speed (within 5–10 %).

The obtained results confirm the effectiveness of the proposed approach in enhancing the reliability of UAV-based computer vision systems. The practical significance of the research lies in the possibility of integrating the method into onboard data processing systems to ensure stable performance under uncertainty and dynamic environmental conditions.

Key words: *computer vision, unmanned aerial vehicles, YOLO, cascade algorithms, object detection, complex weather conditions.*

1. Вступ. Сучасні безпілотні літальні апарати (БпЛА) відіграють ключову роль у задачах моніторингу, розвідки та інспекції інфраструктури. Ефективність виконання цих завдань безпосередньо залежить від роботи бортових систем комп'ютерного зору, основою яких є згорткові нейронні мережі (CNN), зокрема архітектури сімейства YOLO. Вони стали стандартом де-факто завдяки оптимальному балансу між швидкістю та точністю на еталонних наборах даних. Проте функціонування БпЛА відбувається у неструктурованому, високодинамічному середовищі, яке характеризується значною варіативністю зовнішніх факторів. Критичною проблемою залишається нестабільність роботи нейромереж при зміні умов спостереження. Стандартні моделі, навчені на збалансованих датасетах, демонструють

© Юдіна Л. Г., Дегтяр Ю. В., 2026
Стаття поширюється на умовах ліцензії відкритого доступу CC BY 4.0

різке падіння показників точності (Precision) та повноти (Recall) при виникненні атмосферних перешкод (дощ, туман, сніг), складного освітлення (відблиски, сутінки) або вібрацій камери.

Додатковим викликом є варіативність просторового масштабу об'єктів. Під час польоту висота та відстань до цілі постійно змінюються: об'єкт може займати як значну частину кадру, так і лише декілька пікселів. Універсальні моделі часто виявляються нездатними однаково ефективно детектувати об'єкти на різних дистанціях: налаштування, оптимальні для «ближнього бою», дають велику кількість хибних спрацювань на дальніх дистанціях, і навпаки. Існуючі методи вирішення цієї проблеми мають суттєві обмеження: використання однієї «важкої» універсальної моделі вимагає значних обчислювальних ресурсів, що є критичним для вбудованих систем БпЛА з обмеженим енергоспоживанням; застосування класичних ансамблів моделей, де кожен кадр паралельно обробляється кількома мережами, призводить до неприпустимого зниження частоти кадрів (FPS) та збільшення затримки (latency), що унеможливлює керування апаратом у реальному часі.

Таким чином, виникає нагальна потреба у розробці адаптивних алгоритмів, які здатні динамічно підлаштовуватися під поточні умови сцени (погода, масштаб) без втрати швидкодії. Актуальним науково-прикладним завданням є створення методу, який дозволив би використовувати переваги спеціалізованих нейромереж, уникаючи при цьому затримок, властивих традиційним підходам до агрегації моделей.

У роботі вирішено актуальне науково-прикладне завдання підвищення надійності та стабільності систем комп'ютерного зору безпілотних літальних апаратів, що функціонують у динамічному середовищі. Результати досліджень підтверджують, що розроблений каскадний метод є перспективним напрямком удосконалення систем автономної навігації та технічного зору, поєднуючи високу адаптивність до зовнішніх умов із вимогами до швидкодії вбудованих систем.

Сучасний розвиток систем технічного зору для безпілотних апаратів базується на глибоких згорткових нейронних мережах. Фундаментальні праці в цій галузі [1;8-12;17;18;] визначають ключову роль машинного навчання у задачах розпізнавання образів. Серед існуючих архітектур особливе місце посідають одностадійні детектори, зокрема сімейство YOLO (You Only Look Once), розвиток якого від версії YOLOv3 [10] до YOLOv4 [2] та сучасних ітерацій YOLOv8 [7] і YOLOv10 [17] дозволив досягти балансу між точністю та швидкістю інференсу. На відміну від двостадійних методів, таких як Faster R-CNN [14], або трансформерних архітектур на кшталт DETR [3], алгоритми YOLO оптимізовані для роботи в реальному часі, що є критичним для БпЛА [16].

Однак, як зазначається у дослідженнях стійкості нейромереж [15], універсальні моделі демонструють падіння ефективності при зміні домену даних (наприклад, перехід від ясної погоди до туману або зміна масштабу цілі). Для вирішення цієї проблеми було запропоновано каскадний метод адаптивного перемикання моделей.

Тому, мета дослідження полягає у підвищенні ефективності та стійкості систем детекції об'єктів на безпілотних літальних апаратах у складних умовах спостереження (атмосферні перешкоди, змінна дальність) шляхом розробки та застосування каскадного методу нейромережевої обробки зображень [5; 6].

2. Матеріали і методи. Методологія базується на гіпотезі, що спеціалізовані нейронні мережі демонструють вищу ефективність у вузьких сценаріях (конкретні погодні умови, різні дистанції до об'єктів), ніж одна універсальна модель. Дослідження передбачає розробку та тестування каскадного алгоритму, який аналізує рівень впевненості поточної моделі і, у разі його падіння нижче порогового значення, ініціює обробку наступного кадру відеопотоку іншою, більш релевантною моделлю. Такий підхід дозволяє уникнути затримок на повторну обробку одного зображення та підтримувати частоту кадрів на рівні пропускної здатності апаратної платформи.

Математична модель та алгоритм роботи. Суть запропонованого підходу полягає у відмові від використання однієї універсальної моделі на користь впорядкованого набору (каскаду) спеціалізованих моделей $M = \{M1, M2, ..., Mn\}$, де кожна модель Mi оптимізована під конкретні умови середовища (ясна погода, опади, низька освітленість) або просторовий масштаб об'єктів.

Для кількісної оцінки ефективності детекції кожного кадру використовуються стандартні метрики точності (Precision) та повноти (Recall). Вони визначаються через елементи матриці помилок за наступними формулами:

$$Precision = \frac{TP}{TP + FP},$$

$$Recall = \frac{TP}{TP + FN},$$

де: *TP* (True Positive) – кількість коректно виявлених об'єктів; *FP* (False Positive) – кількість хибних спрацювань (детектор знайшов об'єкт там, де його немає); *FN* (False Negative) – кількість пропущених об'єктів.

ISSN 2786-5460 (Print), ISSN 2786-5479 (Online)

Як інтегральна метрика якості використовується mAP (mean Average Precision), що розраховується як середнє значення точності (AP) за всіма N класами об'єктів:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i,$$

де $\sum$ – сума значень середньої точності для кожного класу і від 1 до N.

На відміну від класичних каскадів, де один кадр послідовно обробляється всіма моделями (що призводить до падіння FPS кратно кількості моделей), розроблений алгоритм використовує стратегію «Look-Ahead Switching».

Нехай F_t поточний кадр відеопотоку в момент часу t. Алгоритм аналізує максимальну впевненість детекції (confidence score) $C(F_t)$ оточної активної моделі M_{curr}.

Логіка перемикання описується системою умов:

$$M_{active}(t+1) = \begin{cases} M_{curr} \text{ якщо } & (C)F_t \geqslant \sigma \\ M_{next} \text{ якщо } & (C)F_t \leqslant \sigma \end{cases}$$

де σ – експериментально встановлений поріг впевненості (наприклад, σ = 0.5), M_{next} – наступна модель у каскаді, натренована на складніші умови. Важливою особливістю є те, що при перемиканні модель M_{next} застосовується вже до нового кадру F_{t+1}, а не до поточного F_t.

Схематичне зображення роботи запропонованого алгоритму наведено на рисунку 1.

Як видно з рисунка 1, система складається з трьох рівнів:

1. $M1$ (Basic/Far): Модель для ясної погоди та загальних планів.
2. $M2$ (Cloudy/Mid): Модель для умов середньої складності.
3. $M3$ (Hard/Near): Модель для складних метеоумов (дощ, туман) або специфічних дистанцій.

Експериментальні дослідження.

Для верифікації методу було проведено порівняльний аналіз базового підходу (Single-YOLOv11) та розробленого каскаду (Cascade YOLO). Тестування проводилося на апаратній платформі Intel Core i7 з використанням бібліотек OpenCV та PyTorch [5; 6].

Результати експериментів, отримані на тестовій вибірці змішаного типу (відео з різними погодними умовами), зведені у таблицю 1.

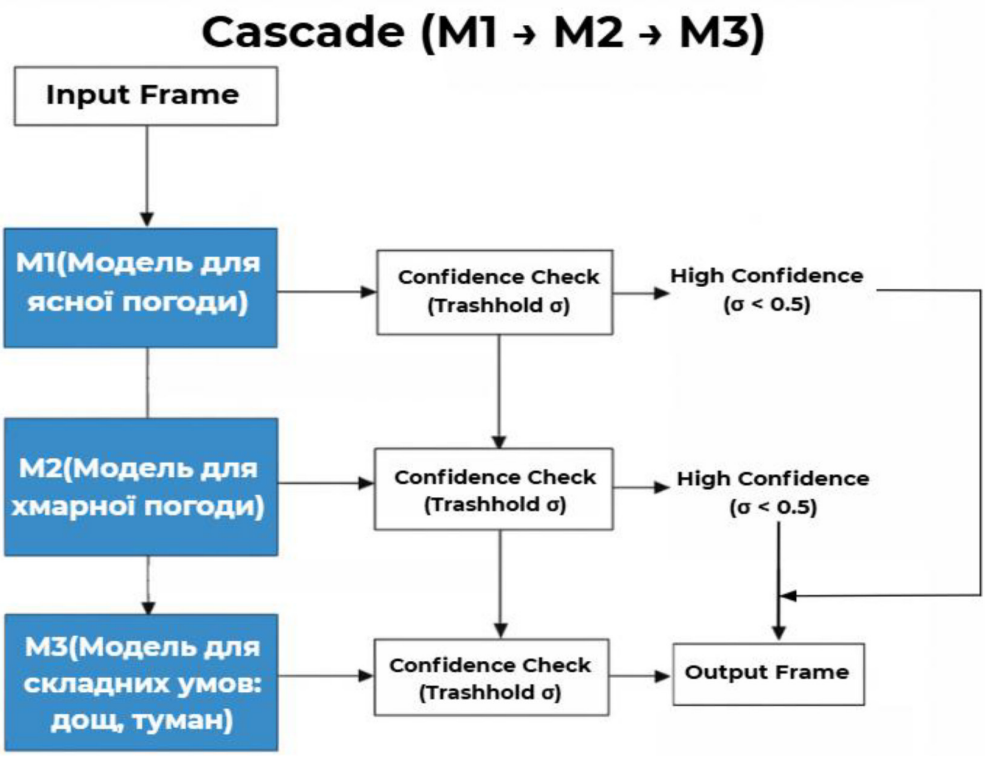

Рис. 1. Блок-схема адаптивного каскадного алгоритму

Таблиця 1

Порівняння ефективності алгоритмів детекції

Алгоритм	Precision	Recall	mAP	Середній FPS	Час обробки кадру (latency)
Single-YOLO (Baseline)	0.734	0.707	0.718	19	50 ms
Cascade ($M1 \rightarrow M2 \rightarrow M3$)	0.867	0.824	0.839	17-18	55 ms

3. Результати і обговорення. Аналізуючи дані таблиці 1, можна констатувати, що каскадний метод забезпечив приріст точності (P) на 18.0 % та повноти (R) на 16.5 %. Показник mAP зріс до 0.839, що свідчить про значно вищу надійність системи у складних сценах.

Особливу увагу слід звернути на показник FPS. У класичних підходах послідовне застосування трьох моделей знизило б швидкість обробки приблизно в 3 рази (до ~6 FPS).

Однак, завдяки застосуванню стратегії обробки нового кадру, середня частота кадрів знизилася лише на 5–10 % (до 17–18 FPS). Це падіння зумовлене накладними витратами на програмне перемикання контексту та завантаження ваг моделей у оперативну пам'ять, але воно не є критичним для задач навігації та трекінгу цілей.

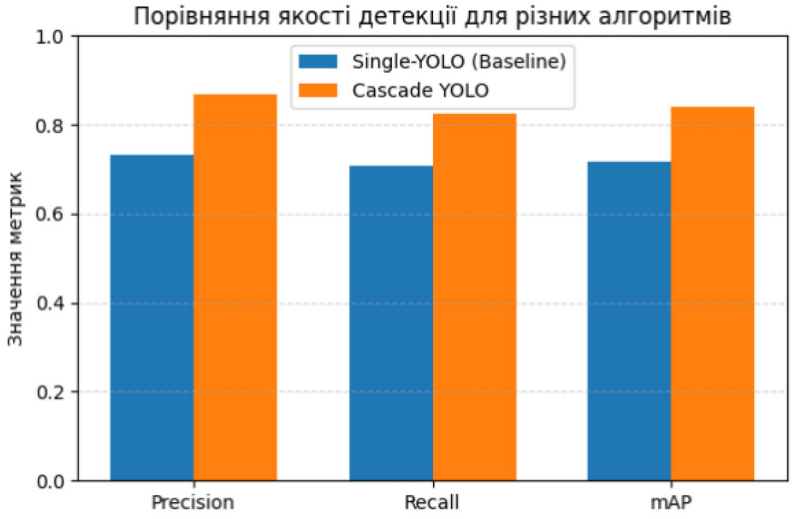

Рис. 2. Порівняння значень Precision, Recall та mAP

Для наочної інтерпретації отриманих результатів побудовано графік зміни основних метрик якості для базового алгоритму та запропонованого каскадного підходу.

Як видно з рис. 2, запропонований каскадний підхід демонструє суттєве підвищення якості детекції порівняно з одиничною моделлю. Значення Precision зростає з 0.734 до 0.867 (приблизно на 18.1 %), Recall – з 0.707 до 0.824 (на 16.5 %), а інтегральний показник mAP – з 0.718 до 0.839 (на 16.9 %). Це свідчить про зменшення кількості як хибних спрацювань, так і пропусків об'єктів у складних сценах.

Таким чином, результати підтверджують переваги спеціалізації моделей, про що згадується в роботах з оптимізації нейромереж [4; 18], але в ході дослідження це досягнуто без втрати реального часу виконання.

У дослідженні вирішено актуальне науково-прикладне завдання підвищення надійності та стабільності систем комп'ютерного зору безпілотних літальних апаратів, що функціонують у динамічному середовищі. На основі проведених досліджень та експериментальних даних можна зробити наступні висновки:

Підтверджено ефективність спеціалізації. Доведено, що використання набору вузькоспеціалізованих моделей YOLO, натренованих окремо для різних погодних умов (ясна погода, дощ, туман) та різних масштабів об'єктів (ближня/дальня дистанція), забезпечує вищі показники точності (Precision) та повноти (Recall) порівняно з використанням однієї універсальної моделі, яка часто втрачає ефективність при зміні контексту сцени.

Оптимізовано швидкодію каскаду. Розроблено та протестовано унікальну стратегію адаптивного перемикання, яка при низькій впевненості поточної моделі ініціює обробку наступною нейромережею нового актуального кадру відеопотоку, а не повторну обробку попереднього. Це дозволило усунути головний недолік класичних каскадних схем – накопичення затримок (latency).

ISSN 2786-5460 (Print), ISSN 2786-5479 (Online)

Експериментально встановлено, що зниження середнього показника FPS при такій схемі не перевищує 5–10 % порівняно з одиночною моделлю, що є критично важливим для керування БпЛА в реальному часі.

Збалансовано точність та ресурсоємність. Запропонований метод дозволяє динамічно адаптувати обчислювальне навантаження: у простих умовах система працює з максимальною продуктивністю, використовуючи базову модель, а у складних (атмосферні перешкоди або зміна відстані до цілі) – автоматично задіює більш релевантні моделі для підтримки високої якості детекції.

Практична цінність. Результати досліджень можуть бути імплементовані в бортові системи обробки інформації розвідувальних та ударних дронів, забезпечуючи стабільне захоплення та супровід цілей незалежно від погодних умов та маневрування апарата за висотою/відстанню.

Таким чином, розроблений каскадний метод є перспективним напрямком удосконалення систем автономної навігації та технічного зору, поєднуючи високу адаптивність до зовнішніх умов із вимогами до швидкодії вбудованих систем.

4. Висновки. Вперше запропоновано метод адаптивної каскадної детекції об'єктів для БпЛА, який, на відміну від класичних ансамблевих підходів, здійснює динамічне перемикання між нейромережами, спеціалізованими не лише за типами об'єктів, а й за умовами спостереження (атмосферні явища, освітлення) та просторовим масштабом (відстань до цілі). Це дозволяє забезпечити високу точність розпізнавання у широкому діапазоні висот польоту та погодних умов.

Удосконалено алгоритм управління обчислювальним потоком у системах комп'ютерного зору реального часу: розроблено стратегію перемикання моделей, яка при недостатній впевненості детекції ініціює обробку не поточного, а наступного актуального кадру відеопотоку. Таке рішення дозволило усунути затримки, властиві традиційним каскадним методам, і зберегти швидкодію системи на рівні, близькому до одиночної моделі (втрати FPS не перевищують 5–10 %).

Набуло подальшого розвитку дослідження стійкості архітектур YOLO до візуальних перешкод, що дозволило експериментально визначити оптимальні порогові значення впевненості (confidence threshold) для автоматичного перемикання між моделями «ближнього/дальнього» плану та «ясних/складних» погодних умов.

Конфлікт інтересів. Автори декларують, що не мають конфлікту інтересів стосовно даного дослідження, в тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

Фінансування. Дослідження проводилося без фінансової підтримки.

Доступність даних. Рукопис має пов'язані дані у сховищі даних.

Використання даних штучного інтелекту. Автори підтверджують, що не використовували технології штучного інтелекту при створенні представленої роботи.

Внесок авторів. Людмила Юдіна: формування наукової концепції, математичне моделювання процесів, обробка даних; Юрій Дегтяр: підготовка ілюстративного матеріалу, реалізація експериментальної частини, перевірка працездатності методу.

References:

1. Bochkovskiy, A., Wang, C.-Y., Liao, H.-Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934,

2. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., Zagoruyko, S. (2020). End-to-End Object Detection with Transformers (DETR). ECCV,

3. Ge, Z., Liu, S., Wang, F., Li, Z., Sun, J. (2021). YOLOX: Exceeding YOLO Series in 2021. arXiv:2107.08430,

4. Кашкевич, С. О., Нечипорук, О. П., Апенько, Н. В., Брановицька, І. В. (2025). Метод оптимізації маршрутів на основі поведінки змій в системах обміну даними БпЛА. *Інформаційні технології та суспільство*. № 4 (19). С. 78–82. https://doi.org/10.32689/maup.it.2025.4.13

5. IBM. Computer Vision. https://www.ibm.com/think/topics/computer-vision

6. Jocher, G. et al. (2023). YOLOv8: Ultralytics Next-Generation Object Detector. GitHub release,

7. Khan, A. A., Laghari, A. A., Awan, S. A. Machine Learning in Computer Vision: A Review.

8. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S. (2016). SSD: Single Shot MultiBox Detector. ECCV

9. Redmon, J., Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv:1804.02767,

10. Ren, S., He, K., Girshick, R., Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. NeurIPS,

11. Szeliski, R. (2021). Computer Vision: Algorithms and Applications, 2nd Edition. Final draft, September 30, © 2022 Springer.

12. Svitlana Kashkevich (Ed.) (2025). Decision support systems: mathematical support. Kharkiv : TECHNOLOGY CENTER PC. https://doi.org/10.15587/978-617-8360-13-9.

13. Svitlana Kashkevich, Illia Dmytriiev, Inna Shevchenko, Oleksandr Lytvynenko, Lyubov Shabanova-Kushnarenko, Nataliia Apenko. (2024). Scientific-method apparatus for improving the efficiency of information processing using artificial

intelligence. Information and control systems: modelling and optimizations: collective monograph. Kharkiv : TECHNOLOGY CENTER PC, pp. 137–167. https://doi.org/10.15587/978-617-8360-04-7.ch5.

14. Tan, M., Pang, R., Le, Q.V. (2020). EfficientDet: Scalable and Efficient Object Detection. CVPR,

15. Ultralytics. YOLO Documentation. https://docs.ultralytics.com

16. Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., Ding, G. (2024). YOLOv10: Real-time end-to-end object detection. arXiv preprint arXiv:2405.14458,

17. Wang, C.Y., Bochkovskiy, A., Liao, H.Y.M. (2021). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv:2207.02696,

18. Xu, Y., et al. (2023). DAMO-YOLO: A Report on Real-Time Object Detection Design. arXiv:2303.04784

Відомості про авторів

Англ.	Укр.
Yudina Lyudmila Postgraduate Student Department of Computer Information Technologies, Kyiv Aviation Institute 1 Lyubomyr Huzar Avenue, Kyiv, 03058 ORCID: 0000-0001-7496-8418	Юдіна Людмила Геннадіївна аспірант Кафедра комп'ютерних інформаційних технологій, Київський авіаційний інститут просп. Любомира Гузара, 1, Київ, 03058 ORCID: 0000-0001-7496-8418
Yuri Degtyar Senior Lecturer in the Department of Intelligent Cybernetic Systems, Kyiv Aviation Institute 1 Lyubomyr Huzar Avenue, Kyiv, 03058 6433061@stud.kai.edu.ua yurii.dehtiar@npp.kai.edu.ua ORCID: 0000-0001-5615-2474	Дегтяр Юрій Віталійович старший викладач кафедри інтелектуальних кібернетичних систем, Київський авіаційний інститут просп. Любомира Гузара, 1, Київ, 03058 6433061@stud.kai.edu.ua yurii.dehtiar@npp.kai.edu.ua ORCID: 0000-0001-5615-2474

Дата надходження статті: 20.03.2026
Дата надходження виправленої версії статті: 07.04.2026
Дата прийняття статті: 17.04.2026
Дата публікації статті: 01.06.2026