

## НЕЙРОМЕРЕЖЕВЕ РОЗПІЗНАВАННЯ ВІЙСЬКОВИХ ТРАНСПОРТНИХ ЗАСОБІВ НА АЕРОКОСМІЧНИХ ЗОБРАЖЕННЯХ З ВИКОРИСТАННЯМ МОДИФІКОВАНОЇ АРХІТЕКТУРИ YOLOV8 ТА ЧАСТОТНО-ПРОСТОРОВИХ ОЗНАК

У роботі представлено модифіковану архітектуру YOLOv8 для автоматизованого розпізнавання військових транспортних засобів на аерокосмічних зображеннях високої роздільної здатності. Запропонована архітектура інтегрує модуль частотно-просторового підсилення ознак (Freq-SpaFEM), модифіковану двонаправлену піраміду ознак (BiFPN), а також орієнтовані обмежувальні рамки (OBB). Процес навчання нейронної моделі здійснено протягом 320 epoch, що забезпечило стабільну збіжність та високу здатність до узагальнення на валідаційних даних. Експериментальні результати засвідчили ефективність запропонованого методу: досягнуто значення mAP@50 на рівні 98,6 %, точності (Precision) – 98,7 % та повноти (Recall) – 95,7 %. Проведений порівняльний аналіз модифікованої архітектури YOLOv8 із базовими моделями, зокрема YOLOv5, YOLOv9, YOLOv11 та стандартною YOLOv8, підтвердив підвищення точності розпізнавання військових транспортних засобів на аерокосмічних зображеннях, зменшення кількості хибнопозитивних і хибнонегативних результатів. Наукова новизна полягає у розробленні нейромережевого методу розпізнавання військових транспортних засобів на аерокосмічних зображеннях, який базується на поєднанні комбінованого частотно-просторового механізму ознак із модифікованою багаторівневою структурою BiFPN, механізмами уваги та застосуванням орієнтованих обмежувальних рамок у межах єдиної глибокої нейромережевої моделі YOLOv8. Аналіз матриць помилок підтвердив високу здатність моделі до розрізнення об'єктів і фонового середовища, що проявляється у високому рівні істинно позитивних результатів та мінімальному рівні хибнопозитивних спрацювань. Це свідчить про ефективність використання запропонованих модулів для підвищення інформативності ознак і точності локалізації. Для практичної реалізації розроблено програмний застосунок із використанням мови програмування Python та бібліотек глибокого навчання та інструментів комп'ютерного зору, що забезпечує можливість візуалізації та аналізу результатів у режимі реального часу.

**Ключові слова:** нейронні мережі, глибоке навчання, YOLO, аерокосмічні зображення, військові транспортні засоби, розпізнавання об'єктів.

## NEURAL NETWORK RECOGNITION OF MILITARY VEHICLES IN AERIAL IMAGES USING A MODIFIED YOLOV8 ARCHITECTURE AND FREQUENCY-SPATIAL FEATURES

This paper presents a modified YOLOv8-based object detection architecture for automated recognition of military transport vehicles in high-resolution aerospace imagery. The proposed approach aims to improve detection accuracy under conditions of complex background, low object contrast, and arbitrary spatial orientation. The architecture integrates a frequency-spatial feature enhancement module (Freq-SpaFEM), a modified bidirectional feature pyramid network (BiFPN), and oriented bounding boxes with rotation parameter estimation. The training process was conducted over 320 epochs, ensuring stable convergence and high generalization performance on validation data. Experimental results demonstrate that the proposed model achieves superior detection performance, reaching mAP@50 of 98.6 %, Precision of 98.7 %, and Recall of 95.7 %. Comparative analysis with baseline models, including YOLOv5, YOLOv9, YOLOv11, and the standard YOLOv8, confirms significant improvements in detection accuracy, reductions in false positives and false negatives, and enhanced robustness to challenging scene conditions. The scientific novelty of the study lies in the development of an integrated approach that combines frequency-domain and spatial feature analysis within a unified deep learning framework, along with the implementation of rotation-aware feature alignment and oriented bounding-box parameterization for aerospace object detection. It enables more accurate representation of object geometry and improves detection performance for partially occluded and arbitrarily oriented targets. Analysis of error matrices confirmed the high ability of the model to distinguish objects and the background environment, which is manifested in a high level of true positive results and a minimum level of false positives. This indicates the effectiveness of using the proposed modules to increase the informativeness of features and localization accuracy. A software application based on deep

*learning libraries and computer vision tools was developed to validate the proposed method in practice, providing real-time visualization and analysis capabilities. The obtained results confirm the effectiveness of the proposed method for aerospace image analysis and highlight its potential for practical deployment in monitoring, surveillance, and geospatial intelligence systems.*

**Keywords:** *neural networks, deep learning, YOLO, aerial images, military vehicles, object recognition.*

### Постановка проблеми

Розвиток військових технологій та поява нових типів загроз [1] зумовлюють підвищені вимоги до оперативності, точності та достовірності обробки розвідувальної інформації. У цьому контексті методи глибокого навчання набули важливого значення в системах національної безпеки та оборони [2; 3], забезпечуючи ефективне вирішення завдань виявлення та ідентифікації потенційних загроз [4].

Особливу роль у сучасних розвідувальних системах відіграють зображення набору даних IMINT [5], що базуються на використанні даних з оптичних, інфрачервоних, радіолокаційних та інших сенсорів, розміщених на наземних, повітряних та космічних платформах. Сьогодні спостерігається тісна інтеграція IMINT із розвідкою з відкритих джерел (OSINT) [6]. Значна частина таких даних формується у вигляді аерокосмічних знімків високого просторового розрізнення, отриманих із супутникових систем та безпілотних літальних апаратів, що забезпечують детальне відображення об'єктів і процесів на земній поверхні. Але автоматизоване розпізнавання військової техніки на аерокосмічних зображеннях супроводжується рядом суттєвих науково-технічних труднощів. Зокрема, складність задачі зумовлена наявністю малорозмірних об'єктів на фоні високодеталізованих і неоднорідних текстур підстильної поверхні; значною варіативністю просторової роздільної здатності вхідних даних; впливом атмосферних явищ, сенсорних шумів та геометричних і радіометричних спотворень; довільною просторовою орієнтацією об'єктів; а також застосуванням засобів маскування та камуфляжу, що знижують контрастність та виразність цільових ознак.

Окремого науково-прикладного значення набуває проблема формування навчальних вибірок для нейромережових моделей. Побудова репрезентативних наборів даних потребує значних часових і обчислювальних ресурсів, що зумовлено необхідністю високоточної анотації об'єктів, зокрема із застосуванням орієнтованих обмежувальних рамок, які забезпечують адекватне відображення геометричних характеристик об'єктів довільної орієнтації. Додатковою складністю є обмежена доступність якісних, збалансованих та достатньо різноманітних даних у військовій сфері, що обумовлює зниження здатності моделей до узагальнення та їх стійкості до варіацій вхідних даних. У задачах військового призначення системи виявлення та розпізнавання об'єктів повинні забезпечувати високу ефективність функціонування в умовах апріорної невизначеності та обмеженості інформації. Зокрема, необхідним є забезпечення ідентифікації об'єктів невеликих розмірів, урахування варіативності просторової роздільної здатності зображень, а також інваріантності до впливу атмосферних спотворень. Сучасні підходи до розв'язання задач розпізнавання об'єктів, зокрема моделі на основі трансформерних архітектур (наприклад, DETR) [7], а також гібридні методи злиття ознак, такі як Pyramid Converge-and-Assign Fusion [8], демонструють високий потенціал щодо підвищення точності розпізнавання об'єктів в складних сценах. Але їх практичне впровадження в інформаційно-аналітичні системи оборонного призначення суттєво обмежується значними обчислювальними витратами, складністю архітектурної реалізації та труднощами обробки даних в режимі реального часу.

Таким чином, актуальною науково-прикладною проблемою є розробка ефективних методів нейромережового розпізнавання військової техніки на аерокосмічних зображеннях, які забезпечують високу точність розпізнавання в умовах складного фону та довільної орієнтації об'єктів.

### Аналіз останніх досліджень і публікацій

Сучасні підходи до автоматизованого розпізнавання об'єктів на зображеннях високого просторового розрізнення можна класифікувати за типом використовуваних архітектур: регіонно-орієнтовані згорткові моделі, одноступеневі детектори, трансформерні підходи та гібридні архітектури. Ранні роботи використовували дискримінативно-структурні моделі, де об'єкт ідентифікувався як сукупність характерних геометричних та просторових ознак. Застосування класичних підходів машинного навчання, таких як метод опорних векторів із використанням методів злиття ознак [9], дозволило розпізнавати військові об'єкти, однак обмеженням таких підходів стала їх низька робастність до варіацій фону та нелінійна складність обробки знімків високої розмірності.

Регіонно-орієнтовані згорткові нейронні мережі (R-CNN, Fast R-CNN, Faster R-CNN) [10] є двоступеневими схемами розпізнавання об'єктів на аерокосмічних знімках. Дані підходи забезпечили суттєве підвищення точності виявлення об'єктів за рахунок попередньої генерації регіональних пропозицій (Region Proposals) з подальшою класифікацією. Такий підхід є ефективним у випадках, коли об'єкти мають чітко виражені ознаки та їх можна відокремити від фону. Проте двотапність процесу зумовлює надмірну обчислювальну інерційність. Велика латентність (inference time) фактично унеможливує використання таких архітектур у системах повітряної розвідки, що працюють у режимі реального часу.

Альтернативою стали одноступеневі детектори (YOLO, SSD), що реалізують концепцію регресії обмежувальних рамок безпосередньо з ознак цілого аерокосмічного зображення. Авторами в роботах [11; 12] продемонстровано ефективність модифікацій YOLO для задач розпізнавання військової техніки. Впровадження механізмів уваги (Attention Mechanisms), деформованих згорток (Deformable Convolutions) та стратегій агресивної аугментації дозволило адаптувати ці моделі до умов обмежених навчальних вибірок. Проте стандартні версії цих мереж часто втрачають ефективність при екстремально малих розмірах цілей на складному ландшафті.

Перспективним напрямом є використання трансформерних моделей, які забезпечують моделювання довготривалих просторових залежностей. Зокрема, модель DETR [13] реалізує наскрізне розпізнавання об'єктів із урахуванням глобального контексту. Проте такі підходи характеризуються значною обчислювальною складністю, повільною збіжністю та обмеженою ефективністю при розпізнаванні об'єктів невеликих розмірів на зображеннях високої роздільної здатності, що є типовим для аерокосмічних даних.

З метою поєднання переваг CNN та трансформерів активно розвиваються гібридні архітектури та методи багатомасштабного злиття ознак. У роботі [14] запропоновано методи, орієнтовані на підвищення точності розпізнавання об'єктів невеликих розмірів на основі використання багатомасштабних та інваріантних ознак, що частково вирішують проблеми розмиття та низької контрастності.

Таким чином, аналіз останніх досліджень свідчить про наявність невирішених проблем, пов'язаних із забезпеченням високої точності розпізнавання військових об'єктів на аерокосмічних зображеннях за умов обмежених обчислювальних ресурсів, складного фону та варіативності просторових характеристик даних. Це обумовлює доцільність розроблення нових підходів, орієнтованих на поєднання ефективних механізмів вилучення локальних і глобальних ознак із урахуванням специфіки аерокосмічних зображень.

### Мета дослідження

Метою роботи є розробка та дослідження модифікованої нейромережевої архітектури YOLOv8, що інтегрує блоки частотно-просторового підсилення ознак (Freq-SpaFEM) у поєднанні з двонаправленою пірамідою ознак (BiFPN), механізмами уваги та орієнтованих обмежувальних рамок (ОВВ) для підвищення точності автоматизованого розпізнавання об'єктів

військових транспортних засобів на аерокосмічних зображеннях в умовах складного природного та антропогенного ландшафту.

### Виклад основного матеріалу дослідження

Запропонований в даній роботі метод автоматизованого розпізнавання військових транспортних засобів на аерокосмічних зображеннях з використанням модифікованої архітектури YOLOv8 представлено на рис. 1.

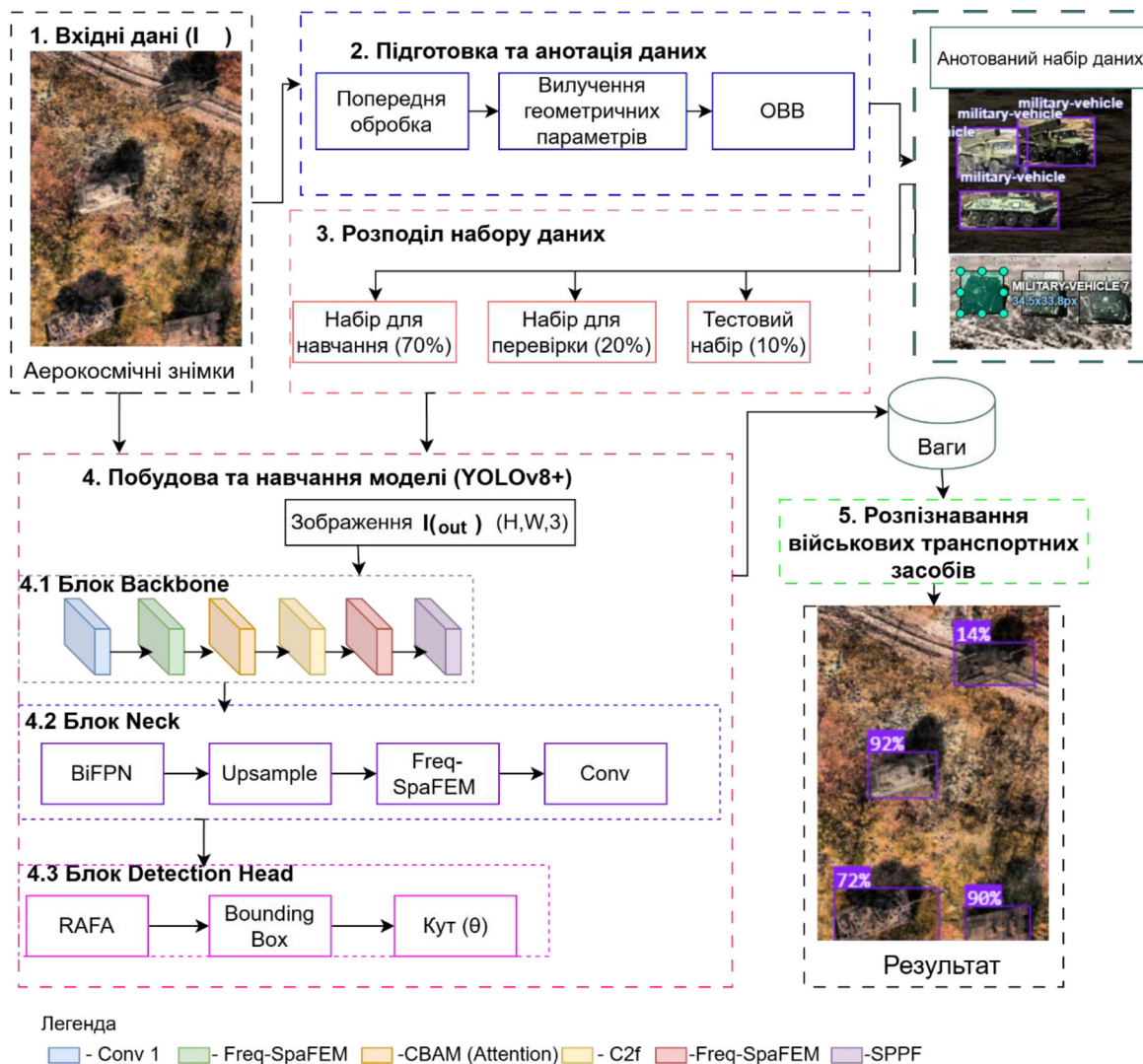


Рис. 1. Структурна схема методу неймережевого розпізнавання військових транспортних засобів на аерокосмічних зображеннях

На першому етапі здійснюється формування вхідних даних шляхом завантаження аерокосмічних знімків високого просторового розрізнення. Джерелами даних є супутникові зображення та знімки, отримані з безпілотних літальних апаратів, що дозволяє забезпечити різноманітність сцен спостереження та умов зйомки.

Другий етап передбачає попередню обробку та анотацію даних. Набір даних, використаний у роботі, сформовано на основі відкритих джерел, зокрема офіційних веб-ресурсів, спеціалізованих платформ із відкритим доступом та матеріалів аерозйомки. Сформований набір містить різноманітну військову техніку, включаючи засоби наземних сил (танки, бойові машини піхоти, бронетранспортери) та повітряних сил (гелікоптери і літаки різного призначення), представлені в різних ракурсах, масштабах і умовах спостереження. Кожне зображення було

анотовано із використанням спеціалізованого інструмента розмітки CVAT (рис. 2). На цьому етапі для кожного об'єкта класу «військові транспортні засоби» проводиться вилучення геометричних параметрів та формування орієнтованих обмежувальних рамок (ОВВ) для коректного навчання моделі в умовах довільної орієнтації об'єктів.. Загальний обсяг сформованого набору становить 2 424 зображення.



Рис. 2. Фрагмент анотації даних

На третьому етапі здійснюється розподіл набору даних на підмножини для навчання, перевірки та тестування у співвідношенні 70/20/10 відповідно. У результаті навчальний набір містить 1701 зображення, перевірки – 480, тестовий – 243 зображення. Такий розподіл забезпечує коректну оцінку узагальнюючої здатності моделі розпізнавати об'єкти на аерокосмічних знімках.

Четвертий етап полягає у побудові та навчанні модифікованої нейромережевої архітектури на основі YOLOv8. Вибір цієї архітектури обумовлений її здатністю забезпечувати баланс між точністю та швидкістю, що є важливим для задач аналізу аерокосмічних зображень у режимі, наближеному до реального часу. На вхід моделі подаються зображення фіксованого формату  $I_{out} \in \mathbb{R}^{H \times W \times 3}$ , отримані після попередньої обробки. Архітектура моделі складається з трьох основних функціональних блоків. У першому блоці (4.1 Backbone) виконується вилучення ознак із використанням каскадної структури згорткових перетворень із інтеграцією модулів Freq-SpaFEM, які забезпечують підсилення інформативних ознак шляхом поєднання аналізу частотних характеристик та просторових структур зображення [15]:

$$F_{enhanced} = f^{-1}(\Phi(f(F_{spatial}))), \quad (1)$$

де  $f$  – пряме перетворення Фур'є,  $f^{-1}$  – обернене перетворення Фур'є,  $\Phi$  – частотна фільтрація.

Додатково використано механізми уваги типу СВМ, що дозволяють адаптивно виділяти найбільш значущі каналні та просторові компоненти ознак. Другий блок (4.2 Neck) спрямований на багатомасштабне агрегування ознак. У ньому реалізовано модифіковану архітектуру ViFPN, яка забезпечує двонаправлене об'єднання ознак різних рівнів просторової роздільної здатності [16]:

$$P_i^{td} = \frac{w_1 \cdot P_i + w_2 \cdot Up(P_{i+1})}{w_1 + w_2 + \epsilon}, \quad P_i^{bu} = \frac{w_3 \cdot P_i + w_4 \cdot Down(P_{i-1})}{w_3 + w_4 + \epsilon}, \quad (2)$$

де  $w_i$  – вагові коефіцієнти, що навчаються,  $Up(\cdot)$  та  $Down(\cdot)$  – операції масштабування.

Такий підхід дозволяє ефективно поєднувати як високорівневі семантичні, так і низькорівневі просторові характеристики зображення. Додатково після операції підвищення просторової роздільної здатності (Upsample) інтегровано модуль Freq-SpaFEM, що забезпечує уточнення частотно-просторових ознак та сприяє покращенню розпізнавання об'єктів невеликого розміру.

Третій блок (4.3 Detection Head) відповідає за формування результатів розпізнавання об'єктів. У ньому реалізовано механізм незалежного визначення просторових параметрів об'єктів на аерокосмічних знімках та їхніх семантичних ознак, що сприяє підвищенню точності локалізації. Ключовою особливістю цього блоку є інтеграція модуля RAFA (Rotation-Aware Feature Alignment) для адаптивного аналізу ознак, що дозволяє мережі враховувати геометричну специфіку військових транспортних засобів. Додатково в архітектурі реалізовано розширену параметризацію обмежувальних рамок, яка передбачає оцінювання орієнтації об'єктів шляхом визначення кута повороту  $\theta$ . На відміну від класичних підходів із осьово-орієнтованими рамками, це дозволяє враховувати довільне просторове розташування об'єктів на аерокосмічних знімках. Застосування орієнтованих обмежувальних рамок забезпечує більш точне узгодження з геометрією об'єктів.

Навчання моделі здійснюється на основі анотованого набору даних із використанням ітеративної оптимізації параметрів шляхом мінімізації функції втрат [17]:

$$L = \lambda_{box} L_{box} + \lambda_{cls} L_{cls} + \lambda_{obj} L_{obj}, \quad (3)$$

де  $L_{box}$  – функція втрат, що характеризує похибку визначення просторових параметрів об'єкта, зокрема координат та розмірів обмежувальної рамки відносно еталонної анотації,  $L_{cls}$  – функція втрат, що характеризує похибку визначення семантичної належності об'єкта до відповідного класу,  $L_{obj}$  – функція втрат, що характеризує похибку визначення наявності об'єкта в заданій області зображення.

У процесі навчання відбувається узгодження передбачених параметрів об'єктів із еталонними анотаціями, що включають координати обмежувальних рамок, належність до класу та орієнтаційні характеристики. У результаті формується набір навчених вагових коефіцієнтів моделі, які зберігають інформацію про виявлені закономірності у даних і надалі використовуються для виконання задачі детекції.

П'ятий етап полягає у застосуванні навченої моделі для автоматизованого виявлення військових транспортних засобів на аерокосмічних зображеннях. На цьому етапі на вхід моделі подаються нові зображення, які не використовувалися під час навчання. Модель, використовуючи отримані вагові коефіцієнти, виконує прогнозування просторового положення об'єктів та їх належності до цільового класу. Результати детекції формуються у вигляді обмежувальних рамок (bounding boxes), кожна з яких містить інформацію про: координати положення об'єкта, оцінку ймовірності (достовірності) виявлення об'єкту та мітки класу (class label).

Процес навчання модифікованої нейромережевої архітектури YOLOv8 здійснювався протягом 320 епох (рис. 3), що дозволило забезпечити стабільну збіжність моделі та високу точність розпізнавання військових транспортних засобів.

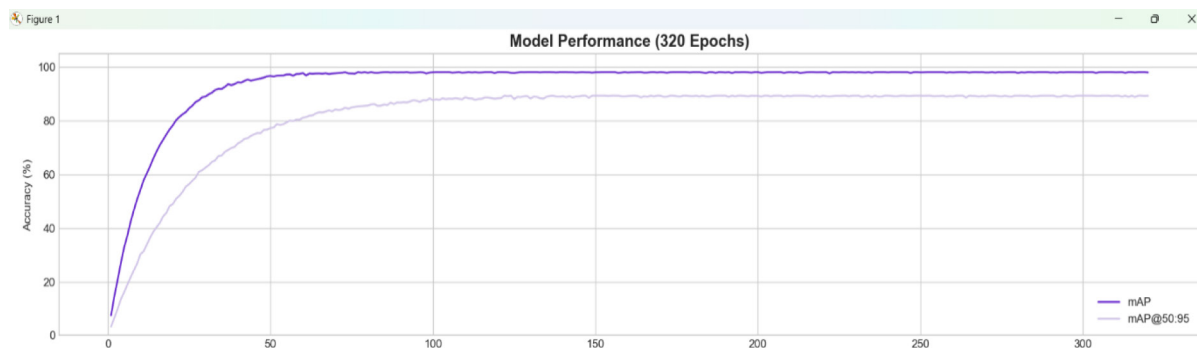


Рис. 3. Результат процесу навчання модифікованої нейромережевої архітектури YOLOv8

Оптимізація параметрів мережі виконувалася ітеративно із застосуванням функції втрат, що поєднує складові локалізації, класифікації та об'єктності. Оцінювання ефективності

навчання проводилося на основі метрик середньої точності (mAP), а також аналізу динаміки функцій втрат, результати яких наведено на рис. 4.

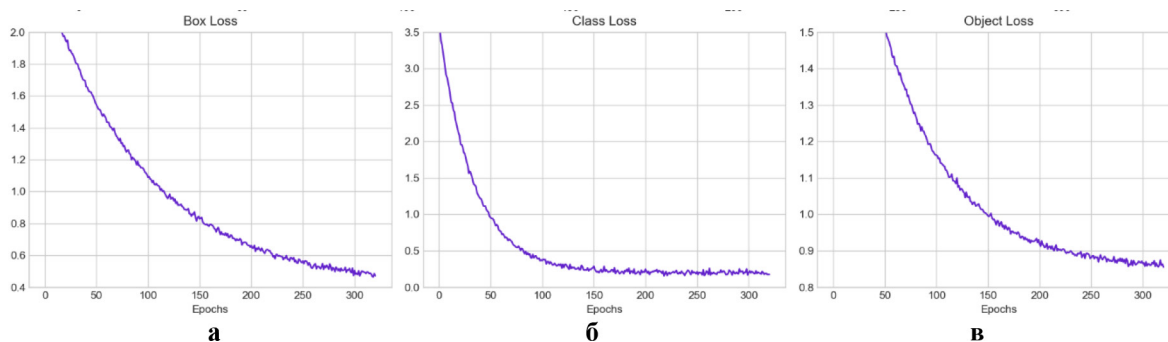


Рис. 4. Графіки функції втрат: а – орієнтованих обмежувальних рамок; б – розпізнавання класів об'єктів; в – впевненості у розпізнаванні об'єкта

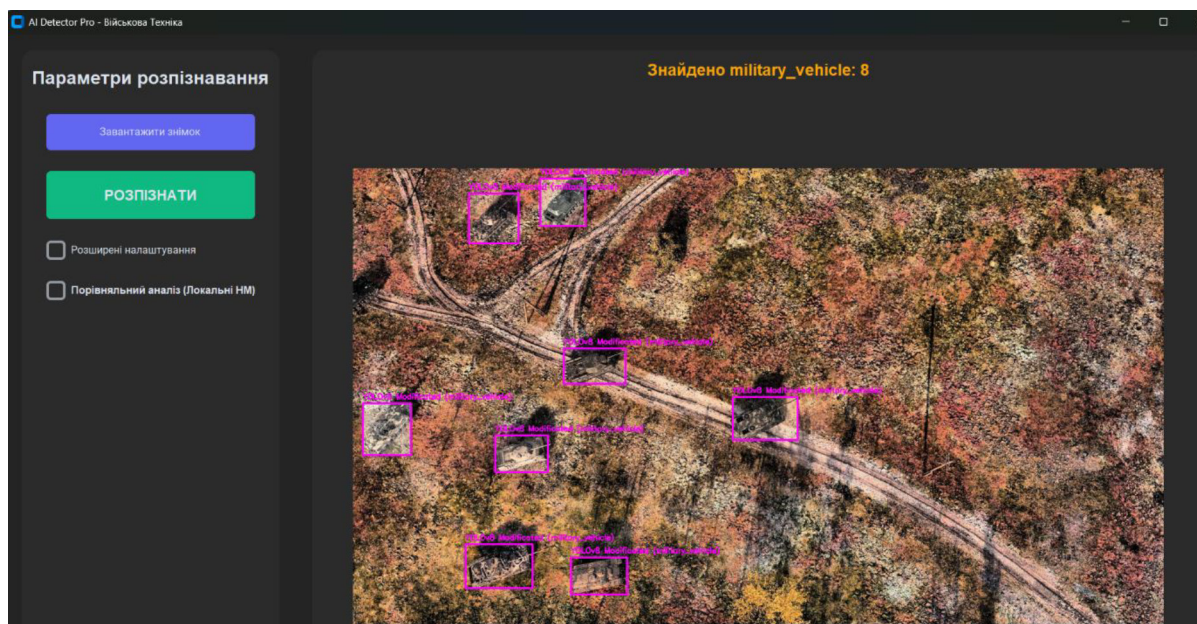
Аналіз метрик точності показав, що запропонована модифікована архітектура демонструє високий рівень узагальнювальної здатності розпізнавати військові транспортні засоби. Зокрема, значення метрики mAP@50 досягло 98.2 %, при цьому основна фаза зростання спостерігалася до 80-ї епохи. Це свідчить про досягнення оптимального балансу між навчанням та перевіркою моделі на валідаційній вибірці. Метрика mAP@50:95 досягла рівня 89.4 %, причому вже на 92-й епосі значення становило 0.71 з подальшим поступовим зростанням до завершення навчання. Отримані результати підтверджують високу точність локалізації об'єктів, зокрема завдяки використанню орієнтованих обмежувальних рамок, що дозволяють враховувати більш точну геометрію об'єктів у сцені аерокосмічного зображення.

Додатковий аналіз функцій втрат (рис. 4) дозволив оцінити ефективність окремих компонентів архітектури. Функція втрат орієнтованих обмежувальних рамок  $\mathcal{L}_{box}$  (рис. 4, а), що характеризує похибку визначення просторових параметрів об'єктів, продемонструвала стабільне зниження з початкового рівня 1.9 до 0.48. Це свідчить про високу прецизійність локалізації та підтверджує ефективність використання модуля вирівнювання ознак із урахуванням обертання, який забезпечує коректне визначення координат і кута орієнтації  $\theta$ .

Функція втрат розпізнавання класів об'єктів  $\mathcal{L}_{cls}$  (рис. 4, б), що характеризує похибку віднесення об'єкта до відповідного класу, демонструє найбільш інтенсивне зниження серед усіх складових. Її значення зменшилося до рівня 0.25, що свідчить про високу здатність моделі до розрізнення класів об'єктів. Такий результат досягається завдяки поєднанню модулів Freq-SpaFEM та ViFPN, які забезпечують формування інформативних ознак і їх ефективне багатомасштабне агрегування.

Функція втрат впевненості у розпізнаванні об'єкта  $\mathcal{L}_{obj}$  (рис. 4, в), що характеризує похибку визначення наявності об'єкта в заданій області зображення, також продемонструвала стабільну збіжність моделі, досягнувши рівня 0.85. Це вказує на сформовану здатність моделі впевнено відокремлювати об'єкти від фонового середовища навіть у складних умовах, що, у свою чергу, забезпечує зниження кількості хибнопозитивних спрацювань.

Для практичної апробації розробленого методу неймережевого розпізнавання військових транспортних засобів було створено спеціалізований програмний додаток. Програмна реалізація виконана мовою програмування Python із використанням бібліотек глибокого навчання та інструментів комп'ютерного зору. Розроблений додаток оснащено графічним інтерфейсом користувача, який забезпечує можливість завантаження аерокосмічних знімків, ініціації процесу обробки та візуалізації результатів розпізнавання у режимі, наближеному до реального часу (рис. 5).



**Рис. 5.** Розроблений програмний додаток для розпізнавання військових транспортних засобів на основі модифікованої нейромережевої архітектури YOLOv8

Процес експериментального дослідження продемонстровано на прикладі аналізу аерокосмічного зображення. На рис. 5 представлено завантажений знімок земної поверхні з високим рівнем семантичної складності фону, на якому присутні об'єкти військових транспортних засобів. Візуальний аналіз вихідного зображення підтверджує складність задачі розпізнавання, що обумовлено низьким контрастом цільових об'єктів, а також наявністю природних об'єктів, зокрема рослинного покриву та ґрунтових доріг, які ускладнюють процес виділення релевантних ознак. Результати ідентифікації, отримані за допомогою модифікованої моделі на основі YOLOv8, представлені на рис. 5. Експериментально встановлено, що інтеграція модулів частотно-просторового підсилення ознак (Freq-SpaFEM) та механізмів орієнтованих обмежувальних рамок (ОВР) забезпечує ефективне розпізнавання об'єктів класу «military\_vehicle». Зокрема, модель коректно ідентифікувала 8 одиниць військових транспортних засобів. Аналіз отриманих результатів свідчить, що запропонований підхід забезпечує високу точність локалізації навіть у випадках часткового перекриття об'єктів або їх довільної орієнтації відносно площини знімка.

Порівняльний аналіз результатів розпізнавання військових транспортних засобів на аерокосмічному зображенні з використанням різних архітектур класу YOLO представлено на рис. 6. Візуально продемонстровано переваги запропонованої модифікованої моделі на основі YOLOv8 порівняно з іншими архітектурами, зокрема YOLOv5, YOLOv9 та YOLOv11.

Експериментальні результати свідчать, що модифікована архітектура YOLOv8+ забезпечує найбільш повне виявлення об'єктів класу «military\_vehicle», ідентифікуючи 8 одиниць техніки, тоді як альтернативні моделі демонструють нижчі показники: YOLOv11 – 6 об'єктів, YOLOv8\_s – 9 об'єктів, YOLOv9 – 1 об'єкт та YOLOv5 – 1 об'єкт. Така різниця у результатах пояснюється як покращеною здатністю моделі до вилучення інформативних ознак, так і ефективністю використаних механізмів багатомасштабної обробки. Візуальний аналіз отриманих результатів показує, що базові моделі YOLO часто «пропускають» об'єкти з низьким контрастом або ті, що частково перекриті елементами фону (рослинністю, дорожньою інфраструктурою). Натомість запропонована модель демонструє вищу чутливість до таких складних випадків, забезпечуючи стабільну детекцію навіть за умов складної текстурної неоднорідності сцени. Це досягається завдяки інтеграції модуля частотно-просторового підсилення ознак (Freq-SpaFEM),

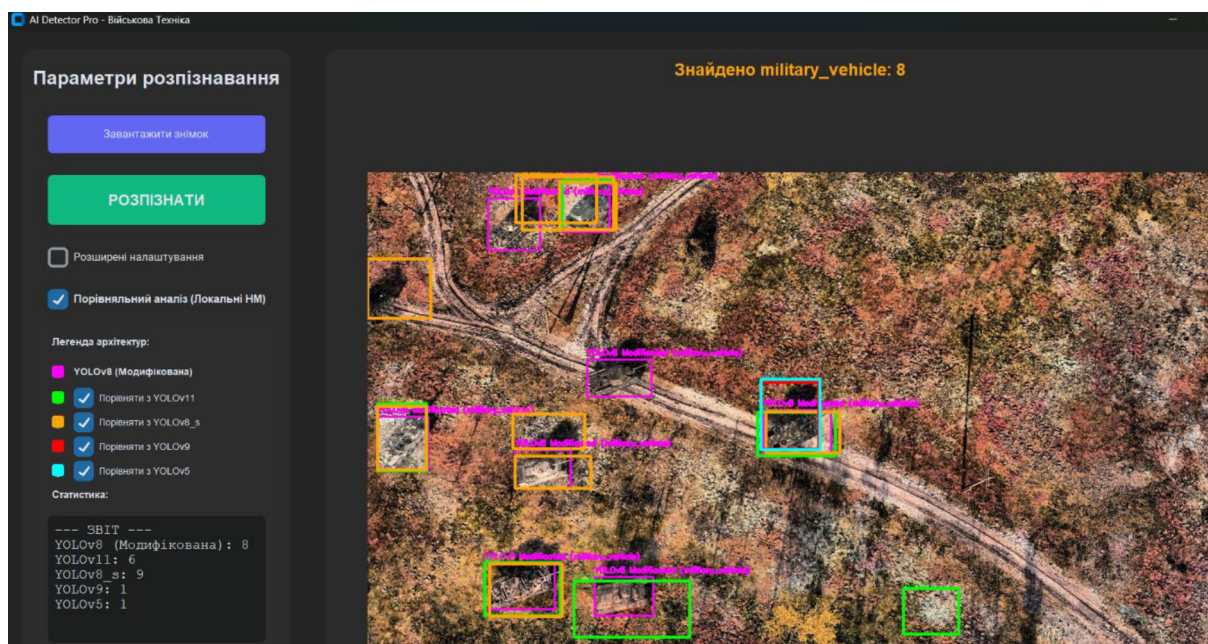


Рис. 6. Результат розпізнавання військових транспортних засобів архітекурами YOLO

який покращує роздільну здатність ознак у спектральних та просторових ознаках. Додатковою перевагою є використання орієнтованих обмежувальних рамок (ОВВ), що дозволяє більш точно узгоджувати форму розпізнаних об’єктів із їхнім реальним положенням у сцені аерокосмічного зображення. У той час як стандартні моделі, що використовують осьово-орієнтовані рамки, демонструють зниження точності при довільній орієнтації об’єктів, запропонований метод забезпечує коректне урахування просторово-геометричних характеристик об’єктів.

Кількісна оцінка ефективності запропонованого методу нейромережевого розпізнавання військових транспортних засобів проводилася шляхом порівняння з популярними архітектурами класу YOLO, зокрема YOLOv8 (базова версія), YOLOv5, YOLOv9 та YOLOv11. Для оцінювання якості розпізнавання використано стандартні метрики: середню точність (mAP@50), точність (Precision), повноту (Recall), а також фактичну кількість виявлених об’єктів. Результати представлено в табл. 1.

Таблиця 1

**Кілткісні метрики точності розпізнавання військових транспортних засобів для нейромережевих архітектур**

Архітектура	mAP@50	Precision	Recall	Виявлено об’єктів
YOLOv8 (Модифікована)	98.6 %	98.7 %	95.7 %	8
YOLOv8n (Base)	28.0 %	39.5 %	33.1 %	9
YOLO11n (Standard)	27.5 %	50.5 %	27.8 %	6
YOLOv9 (Base)	12.5 %	35.0 %	15.0 %	1
YOLOv5 (Base)	8.4 %	28.0 %	10.0 %	1

Отримані результати свідчать про перевагу запропонованої модифікованої архітектури YOLOv8 над іншими архітектурами. Зокрема, значення mAP@50 для запропонованої моделі становить 98.6 %, тоді як для YOLOv8n – 28.0 %, YOLO11n – 27.5 %, YOLOv9 – 12.5 % та YOLOv5 – 8.4 %. Аналогічна тенденція спостерігається і для метрик точності та повноти: YOLOv8 (Модифікована) досягає Precision на рівні 98.7 % та Recall – 95.7 %, що значно перевищує відповідні показники інших моделей. При цьому кількість коректно розпізнаних об’єктів становить 8, що відповідає фактичній кількості цілей на зображенні, тоді як альтернативні

моделі демонструють або пропуски об'єктів (низький Recall), або надлишкові детекції (завищена кількість виявлених об'єктів). Зокрема, базова модель YOLOv8n виявила 9 об'єктів, що перевищує реальну кількість і свідчить про наявність хибнопозитивних спрацювань, при цьому її Precision (39.5 %) та Recall (33.1 %) залишаються на низькому рівні. Модель YOLO11n продемонструвала дещо вищу точність (50. %), проте низький Recall (27.8 %) вказує на значну кількість пропущених об'єктів. Найгірші результати показали YOLOv9 та YOLOv5, для яких характерні як низька точність, так і низька здатність до розпізнавання об'єктів.

Для більш детального аналізу якості класифікації та здатності моделей розрізняти об'єкти на фоні підстильної поверхні було побудовано матриці помилок (рис. 7).

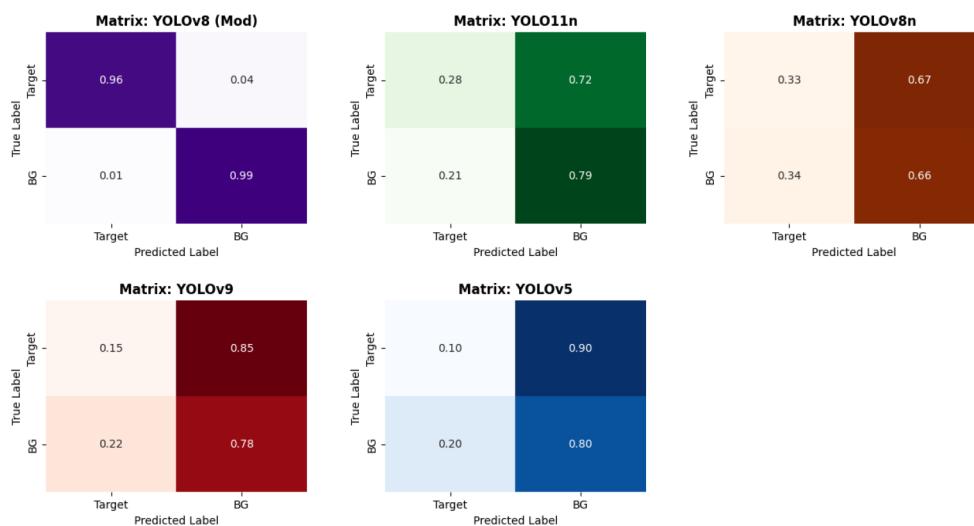


Рис. 7. Матриці помилок для розпізнавання військових транспортних засобів архітектурами YOLO

Результати аналізу підтверджують високу ефективність запропонованого в роботі методу розпізнавання військових транспортних засобів на аерокосмічних зображеннях. Зокрема, для YOLOv8 (Модифікована) значення істинно позитивних спрацювань (True Positive) становить 0.96, що свідчить про майже повну коректність розпізнавання об'єктів військових транспортних засобів. Крім того, рівень хибнопозитивних спрацювань (False Positive) є мінімальним і дорівнює 0.01, що вказує на ефективне відокремлення об'єктів від фону.

### Висновки

У результаті проведених досліджень в роботі розроблено та досліджено модифіковану неймережеву архітектуру YOLOv8 для задачі автоматизованого розпізнавання військових транспортних засобів на аерокосмічних зображеннях. Запропонований метод базується на інтеграції частотно-просторового підсилення ознак, модифікованого механізму багатомасштабного агрегування ознак та використанні орієнтованих обмежувальних рамок.

Під час проведення експериментальних досліджень встановлено, що збільшення тривалості навчання до 320 епох забезпечує досягнення стабільної збіжності моделі та високої узгодженості результатів на валідаційних даних. Модифікована архітектура YOLOv8 продемонструвала високі значення основних метрик якості, зокрема mAP@50 – 98.6 %, Precision – 98.7 % та Recall – 95.7 %, що перевищують показники базових з архітектур класу YOLO.

Порівняльний аналіз із сучасними архітектурами YOLOv5, YOLOv9, YOLOv11 та базовою версією YOLOv8 показав, що запропонована модифікована архітектура YOLOv8 забезпечує значно вищу точність розпізнавання військових транспортних засобів, зменшення кількості хибнопозитивних і хибнонегативних спрацювань, а також більш повне виявлення об'єктів у складних умовах аерокосмічних сцен.

Аналіз матриць помилок підтвердив високу здатність моделі до розрізнення об'єктів і фонового середовища, що проявляється у високому рівні істинно позитивних результатів (0.96) та мінімальному рівні хибнопозитивних спрацювань (0.01). Це свідчить про ефективність використання запропонованих модулів для підвищення інформативності ознак і точності локалізації.

Практична реалізація у вигляді програмного додатка із використанням бібліотек глибокого навчання та інструментів комп'ютерного зору підтвердила можливість застосування розробленого методу в реальних умовах. Забезпечено автоматизоване завантаження зображень, обробку та візуалізацію результатів детекції у режимі, наближеному до реального часу.

Дослідження статті виконані в рамках наукового проєкту «Інтелектуальні технології аналізу просторово-часових змін на аерокосмічних знімках для прийняття рішень в умовах збройної агресії» (номер держреєстрації № 0126U000995), що фінансується Міністерством освіти і науки України за рахунок державного бюджету.

### Список використаної літератури

1. Robinson K., McKenna B., Rooney D. The relationship of risk to rules, values, virtues, and moral complexity: What we can learn from the moral struggles of military leaders. *J. Bus. Ethics*. 2022. Vol. 179. P. 749–766. <https://doi.org/10.1007/s10551-021-04842-4>
2. Peng H., Zhang Y., Yang S., Song B. Battlefield image situational awareness application based on deep learning. *IEEE Intell. Syst.* 2019. Vol. 35. P. 36–43. <https://doi.org/10.1109/MIS.2019.2942701>
3. Mei C., Fan Z., Zhu Q., Yang P., Hou Z., Jin H. A Novel scene matching navigation system for UAVs based on vision/inertial fusion. *IEEE Sens. J.* 2023. Vol. 23. P. 6192–6203. <https://doi.org/10.1109/JSEN.2023.3242045>
4. Zhou L., Leng S., Liu Q., Wang Q. Intelligent UAV swarm cooperation for multiple targets tracking. *IEEE Internet Things J.* 2021. Vol. 9. P. 743–754. <https://doi.org/10.1109/JIOT.2021.3106339>
5. Meng Q., Zhang L., Sun Z., Meng F., Wang L., Sun Y. Characterizing spatial and temporal trends of surface urban heat island effect in an urban main built-up area: A 12-year case study in Beijing, China. *Remote Sens. Environ.* 2018. Vol. 204. P. 826–837. <https://doi.org/10.1016/j.rse.2017.10.027>
6. Scuro V. Open-Source Intelligence (OSINT) for Researchers and Practitioners. 2025. [https://doi.org/10.1007/978-3-032-02014-7\\_2](https://doi.org/10.1007/978-3-032-02014-7_2)
7. Carion N., Massa F., Synnaeve G. et al. End-to-end object detection with transformers. In *Proceedings of the European Conference on Computer Vision (ECCV)*. Glasgow, UK, 23–28 August 2020. Springer, Cham. P. 213–229. [https://doi.org/10.1007/978-3-030-58452-8\\_13](https://doi.org/10.1007/978-3-030-58452-8_13)
8. Wang W. et al. Pyramid vision transformer: A versatile backbone for dense prediction without convolutions. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. Virtual, 11–17 October 2021. P. 568–578. <https://doi.org/10.1109/ICCV48922.2021.00061>
9. Sumari A. D. W. et al. Automatic target recognition and identification for military ground-to-air observation tasks using support vector machine and information fusion. In *Proceedings of the 2022 International Conference on ICT for Smart Society (ICISS)*. Bandung, Indonesia, 10–11 August 2022. P. 1–8. <https://doi.org/10.1109/ICISS55894.2022.9915201>
10. Su H. et al. Object detection and instance segmentation in remote sensing imagery based on precise mask R-CNN. In *Proceedings of the IGARSS 2019 – 2019 IEEE International Geoscience and Remote Sensing Symposium*. Yokohama, Japan, 28 July–2 August 2019. P. 1454–1457. <https://doi.org/10.1109/IGARSS.2019.8898160>
11. Guo Z., Mu X., Chang C. et al. An Enhanced Framework for Small Object Detection with MiddleOrder Interaction and Adaptive Cross-Scale Aggregation. *Engineering Applications of Artificial Intelligence*. 2025. Vol. 130. <https://doi.org/10.1016/j.engappai.2024.107736>

12. Zhao X., Xia Y., Zhang W., Zheng C., Zhang Z. YOLO-ViT-Based Method for Unmanned Aerial Vehicle Infrared Vehicle Target Detection. *Remote Sens.* 2023. Vol. 15. 3778. <https://doi.org/10.3390/rs15153778>
13. Alrayes F. S., Ahmad N., Alshuhail A. et al. Convolutional Transform Learning-Based Fusion for Scale-Invariant Long-Term Target Detection in UAVs. *Scientific Reports.* 2025. Vol. 15. <https://doi.org/10.1038/s41598-024-51123-x>
14. Chutke S., Kumar V. V., Saravanakumar M. AI-Enhanced Image Processing for Target Detection and Threat Recognition in Defence Applications. *In Proc. IEEE Conf. on Smart Technologies and Virtual Environments.* 2025.
15. Wang J., Wu Y., Liu B. et al. Fourier aids CNN and transformer for semantic segmentation of remote sensing images. *Multimedia Systems.* 2025. Vol. 31. 451. <https://doi.org/10.1007/s00530-025-02029-4>
16. Tan M., Pang R., Le Q. EfficientDet: Scalable and Efficient Object Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).* 2020. P. 10778-10787. <https://doi.org/10.1109/CVPR42600.2020.01079>
17. Affah V., Erniwati S. YOLOv8 for Object Detection: A Comprehensive Review of Advances, Techniques, and Applications. *International Journal of Advanced Computing and Informatics.* 2025. Vol. 2. P. 53–61. <https://doi.org/10.71129/ijaci.v2i1.pp53-61>

### References

1. Robinson, K., McKenna, B., & Rooney, D. (2022). The relationship of risk to rules, values, virtues, and moral complexity: What we can learn from the moral struggles of military leaders. *Journal of Business Ethics*, 179, 749–766. <https://doi.org/10.1007/s10551-021-04842-4>
2. Peng, H., Zhang, Y., Yang, S., & Song, B. (2019). Battlefield image situational awareness application based on deep learning. *IEEE Intelligent Systems*, 35, 36–43. <https://doi.org/10.1109/MIS.2019.2942701>
3. Mei, C., Fan, Z., Zhu, Q., Yang, P., Hou, Z., & Jin, H. (2023). A Novel scene matching navigation system for UAVs based on vision/inertial fusion. *IEEE Sensors Journal*, 23, 6192–6203. <https://doi.org/10.1109/JSEN.2023.3242045>
4. Zhou, L., Leng, S., Liu, Q., & Wang, Q. (2021). Intelligent UAV swarm cooperation for multiple targets tracking. *IEEE Internet of Things Journal*, 9, 743–754. <https://doi.org/10.1109/JIOT.2021.3106339>
5. Meng, Q., Zhang, L., Sun, Z., Meng, F., Wang, L., & Sun, Y. (2018). Characterizing spatial and temporal trends of surface urban heat island effect in an urban main built-up area: A 12-year case study in Beijing, China. *Remote Sensing of Environment*, 204, 826–837. <https://doi.org/10.1016/j.rse.2017.10.027>
6. Scuro, V. (2025). *Open-Source Intelligence (OSINT) for Researchers and Practitioners.* Springer. [https://doi.org/10.1007/978-3-032-02014-7\\_2](https://doi.org/10.1007/978-3-032-02014-7_2)
7. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). End-to-end object detection with transformers. *In Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 213–229). Springer, Cham. [https://doi.org/10.1007/978-3-030-58452-8\\_13](https://doi.org/10.1007/978-3-030-58452-8_13)
8. Wang, W., Xie, E., Li, X., Fan, D. P., Song, K., Liang, D.,... & Shao, L. (2021). Pyramid vision transformer: A versatile backbone for dense prediction without convolutions. *In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (pp. 568–578). <https://doi.org/10.1109/ICCV48922.2021.00061>
9. Sumari, A. D. W., Pranata, A. S., Mashudi, I. A., Syamsiana, I. N., & Sereati, C. O. (2022). Automatic target recognition and identification for military ground-to-air observation tasks using support vector machine and information fusion. *In Proceedings of the 2022 International Conference on ICT for Smart Society (ICISS)* (pp. 1–8). <https://doi.org/10.1109/ICISS55894.2022.9915201>

10. Su, H., Wei, S., Yan, M., Wang, C., Shi, J., & Zhang, X. (2019). Object detection and instance segmentation in remote sensing imagery based on precise mask R-CNN. *In Proceedings of the IGARSS 2019 – 2019 IEEE International Geoscience and Remote Sensing Symposium* (pp. 1454–1457). <https://doi.org/10.1109/IGARSS.2019.8898160>
11. Guo, Z., Mu, X., Chang, C., et al. (2025). An Enhanced Framework for Small Object Detection with MiddleOrder Interaction and Adaptive Cross-Scale Aggregation. *Engineering Applications of Artificial Intelligence*, 130. <https://doi.org/10.1016/j.engappai.2024.107736>
12. Zhao, X., Xia, Y., Zhang, W., Zheng, C., & Zhang, Z. (2023). YOLO-ViT-Based Method for Unmanned Aerial Vehicle Infrared Vehicle Target Detection. *Remote Sensing*, 15(15), 3778. <https://doi.org/10.3390/rs15153778>
13. Alrayes, F. S., Ahmad, N., Alshuhail, A., et al. (2025). Convolutional Transform Learning-Based Fusion for Scale-Invariant Long-Term Target Detection in UAVs. *Scientific Reports*, 15. <https://doi.org/10.1038/s41598-024-51123-x>
14. Chutke, S., Kumar, V. V., & Saravanakumar, M. (2025). AI-Enhanced Image Processing for Target Detection and Threat Recognition in Defence Applications. *In Proceedings of the IEEE Conference on Smart Technologies and Virtual Environments*.
15. Wang, J., Wu, Y., Liu, B., et al. (2025). Fourier aids CNN and transformer for semantic segmentation of remote sensing images. *Multimedia Systems*, 31, 451. <https://doi.org/10.1007/s00530-025-02029-4>
16. Tan, M., Pang, R., & Le, Q. (2020). EfficientDet: Scalable and Efficient Object Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 10778–10787. <https://doi.org/10.1109/CVPR42600.2020.01079>
17. Afifah, V., & Erniwati, S. (2025). YOLOv8 for Object Detection: A Comprehensive Review of Advances, Techniques, and Applications. *International Journal of Advanced Computing and Informatics*, 2, 53-61. <https://doi.org/10.71129/ijaci.v2i1.pp53-61>

Каштан Віта Юріївна – к.т.н., доцент, доцент кафедри інформаційних технологій та комп'ютерної інженерії Національного технічного університету «Дніпровська політехніка». E-mail: [kashtan.v.yu@nmu.one](mailto:kashtan.v.yu@nmu.one), ORCID: 0000-0002-0395-5895.

Гнатушенко Володимир Володимирович – д.т.н., професор, завідувач кафедри інформаційних технологій та комп'ютерної інженерії Національного технічного університету «Дніпровська політехніка». E-mail: [vvgnat@ukr.net](mailto:vvgnat@ukr.net), ORCID: 0000-0003-3140-3788.

Kashtan Vita Yuriiivna – Candidate of Technical Sciences, Associate Professor, Associate Professor at the Information Technology and Computer Engineering Department of the Dnipro University of Technology. E-mail: [kashtan.v.yu@nmu.one](mailto:kashtan.v.yu@nmu.one), ORCID: 0000-0002-0395-5895.

Hnatushenko Volodymyr Volodymyrovych – Doctor of Technical Sciences, Professor, Head of the Information Technology and Computer Engineering Department of the Dnipro University of Technology. E-mail: [vvgnat@ukr.net](mailto:vvgnat@ukr.net), ORCID: 0000-0003-3140-3788.

Дата першого надходження статті до видання: 10.04.2026

Дата прийняття статті до друку після рецензування: 11.05.2026

Дата публікації (оприлюднення) статті: 01.07.2026



Стаття поширюється на умовах ліцензії відкритого доступу (CC BY 4.0)