

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**  
**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ**  
**«ДНІПРОВСЬКА ПОЛІТЕХНІКА»**

Кваліфікаційна наукова  
праця на правах рукопису

**КАЗИМИРЕНКО ОЛЕКСІЙ ВОЛОДИМИРОВИЧ**

УДК 004.932.72`1:519.688

**ДИСЕРТАЦІЯ**

**НЕЙРОМЕРЕЖЕВЕ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ ТРАНСПОРТУ НА  
АЕРОКОСМІЧНИХ ЗОБРАЖЕННЯХ**

126 – Інформаційні системи та технології

12 – Інформаційні технології

Подається на здобуття ступеня **доктора філософії**

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело



О.В. Казимиренко

Науковий керівник: **ГНАТУШЕНКО Володимир Володимирович**,  
доктор технічних наук, професор

Дніпро – 2026

## АНОТАЦІЯ

**Казмиренко О.В. Нейромережеве розпізнавання об'єктів транспорту на аерокосмічних зображеннях. – Рукопис**

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 126 – Інформаційні системи та технології. – Національний технічний університет «Дніпровська політехніка», Дніпро, 2026.

В дисертаційному дослідженні розв'язана важлива науково-прикладна задача розробки інформаційної технології розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях з використанням методів глибинного навчання.

Дисертація складається зі вступу, чотирьох розділів, висновків, списку використаних джерел та додатків. Повний обсяг дисертації – 168 сторінок; список використаних джерел зі 142 найменування, 3 додатки. Робота проілюстрована 46 рисунками та містить 10 таблиць.

У вступі обґрунтовано актуальність теми, сформульовані мета і завдання дослідження, визначені методи роботи, наведена загальна характеристика дослідження та його структури, визначені наукова новизна, практичне значення отриманих результатів і особистий внесок автора, оцінена достовірність результатів, а також наведені відомості про публікації та апробацію результатів дослідження.

У першому розділі проаналізовано сучасний стан розвитку методів аерокосмічного спостереження та обробки зображень високої просторової роздільної здатності. Розглянуто технічні аспекти формування аерокосмічних знімків, їх характеристики та спектральні особливості, а також методи попередньої обробки даних. Проаналізовано існуючі підходи до класифікації та розпізнавання об'єктів на аерокосмічних зображеннях, зокрема транспортних засобів, а також сучасні методи машинного навчання і глибинного навчання, включно зі згортковими нейронними мережами. Визначено проблеми і виклики автоматизованого розпізнавання об'єктів

транспорту на аерокосмічних знімках і сформульовано завдання розробки інформаційної технології їх просторово-орієнтованого розпізнавання.

У другому розділі запропоновано методи та інформаційні технології семантичної сегментації та просторово-орієнтованого розпізнавання транспортних засобів на аерокосмічних знімках високої просторової роздільної здатності. Розроблено нейромережевий підхід на основі архітектури DeepLab із глибинною залишковою мережею ResNet, що забезпечує багатомасштабне вилучення ознак і застосування функції втрат на основі коефіцієнта Dice, дозволяючи точно виділяти транспортні засоби навіть у складних багатокласових сценах та за часткового перекриття. Запропоновано методи підвищення точності розпізнавання транспортних засобів на аерокосмічних знімках: для виявлення повітряного транспорту застосовано інтеграцію контурних ознак із оригінальними зображеннями, а для виявлення автомобілів – модифіковану архітектуру YOLOv11 з орієнтованими обмежувальними рамками. Розроблений метод просторово-орієнтованого розпізнавання транспортних засобів довільного розташування забезпечує узгоджене багатомасштабне ознакове подання та контекстуальне узгодження ознак із семантичним поданням сцени, що сприяє підвищенню точності локалізації в умовах складного фону та щільного розташування об'єктів.

У третьому розділі розроблено інформаційну технологію нейромережевого розпізнавання транспортних засобів довільного розташування на аерокосмічних знімках високої просторової роздільної здатності. Технологія базується на багатокомпонентній архітектурі, що поєднує глибоке вилучення ознак через Backbone Network, просторову стабілізацію за допомогою Spatial Transformer Network та розпізнавання транспортних засобів із використанням орієнтованих обмежувальних рамок і багатозадачної функції втрат. Представлено розробку програмного додатку для реалізації запропонованих методів, описано архітектуру системи, функціональні модулі, алгоритми обробки даних та інтеграцію результатів

нейромережевого розпізнавання з інтерфейсом користувача. Продемонстровано функціональні можливості додатку, включно з автоматичним та ручним режимами налаштування параметрів розпізнавання, формування аналітичної звітності та верифікації результатів. Експериментальні дослідження підтвердили ефективність розробленого програмного продукту для автоматизованого аналізу аерокосмічних знімків.

У четвертому розділі проведено експериментальне дослідження інформаційної технології нейромережевого розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях. Сформовано спеціалізований набір даних із знімків камери SONY DSC-WX220 із тайлінгом 200×200 пікселів, що дозволило мінімізувати вплив складного фону та оптимізувати навчання моделі. Здійснено тестування розроблених моделей із використанням спеціалізованих навчальних наборів даних, проведено порівняння продуктивності запропонованого підходу з класичними та сучасними нейромережевими методами. Експериментальні дослідження забезпечують високу точність виявлення, стійкість до змін масштабу та орієнтації транспортних засобів, а також ефективну локалізацію об'єктів навіть у складних умовах аерокосмічної зйомки та багатокласових сцен.

Впровадження розроблених у межах дисертаційного дослідження інформаційних технологій та методів нейромережевого розпізнавання транспортних засобів довільного просторового розташування на аерокосмічних зображеннях, а також створеного на їх основі програмного забезпечення у виробничу діяльність підприємств забезпечило скорочення тривалості обробки аерокосмічних даних, зниження ресурсомісткості процесів аналізу та підвищення точності автоматизованого розпізнавання об'єктів.

**Ключові слова:** інформаційна технологія, машинне навчання, нейронна мережа, аерокосмічні зображення, розпізнавання, транспортні засоби, сегментація.

## SUMMARY

### **Kazymyrenko O.V. Neural network recognition of transport objects in aerospace images. – Manuscript.**

Thesis for a scientific degree of Doctor of Philosophy in the specialty 126 «Information systems and technologies». – Dnipro University of Technology, Dnipro, 2026.

In the dissertation research, an important scientific and applied problem is solved, namely the development of an information technology for spatially oriented recognition of vehicles in aerospace imagery based on deep learning methods.

The dissertation consists of an introduction, four chapters, conclusions, a list of references, and appendices. The total volume of the dissertation is 168 pages; the list of references includes 142 sources, and there are 3 appendices. The work is illustrated with 46 figures and contains 10 tables.

The introduction substantiates the relevance of the research topic, formulates the aim and objectives of the study, defines the research methods, and presents a general description of the research and its structure. The scientific novelty, practical significance of the obtained results, and the author's personal contribution are identified. The reliability of the results is assessed, and information on publications and approbation of the research results is provided.

The first chapter analyzes the current state of development of aerospace observation methods and high spatial resolution image processing. Technical aspects of aerospace image formation, their characteristics and spectral features, as well as data preprocessing methods, are considered. Existing approaches to object classification and recognition in aerospace imagery, in particular vehicles, are analyzed, along with modern machine learning and deep learning methods, including convolutional neural networks. The problems and challenges of automated vehicle object recognition in aerospace imagery are identified, and the

tasks of developing an information technology for their spatially oriented recognition are formulated.

In the second chapter, methods and information technologies for semantic segmentation and spatially oriented recognition of vehicles in high spatial resolution aerospace imagery are proposed. A neural network–based approach built upon the DeepLab architecture with a deep residual ResNet backbone is developed, enabling multi-scale feature extraction and the use of a Dice coefficient–based loss function. This approach allows accurate delineation of vehicles even in complex multi-class scenes and under partial occlusion. Methods for improving the accuracy of vehicle recognition in aerospace imagery are proposed: for aerial transport detection, the integration of contour-based features with the original images is applied, while for road vehicle detection, a modified YOLOv11 architecture with oriented bounding boxes is employed. The developed method for spatially oriented recognition of vehicles with arbitrary orientations provides consistent multi-scale feature representation and contextual alignment of features with the semantic scene representation, thereby improving localization accuracy under complex background conditions and dense object distributions.

In the third chapter, an information technology for neural network–based recognition of arbitrarily oriented vehicles in high spatial resolution aerospace imagery is developed. The technology is based on a multicomponent architecture that combines deep feature extraction via a backbone network, spatial stabilization using a Spatial Transformer Network, and vehicle recognition employing oriented bounding boxes and a multi-task loss function. The development of a software application implementing the proposed methods is presented, including a description of the system architecture, functional modules, data processing algorithms, and integration of neural network recognition results with the user interface. The functional capabilities of the application are demonstrated, including automatic and manual modes for configuring recognition parameters, generation of analytical reports, and verification of results.

Experimental studies confirmed the effectiveness of the developed software product for automated analysis of aerospace imagery.

In the fourth chapter, an experimental study of the information technology for neural network–based recognition of arbitrarily oriented vehicles in aerospace imagery is conducted. A specialized dataset is formed using images acquired by a SONY DSC-WX220 camera, with tiling into  $200 \times 200$  pixel patches, which made it possible to minimize the influence of complex backgrounds and to optimize model training. The developed models are tested using specialized training datasets, and the performance of the proposed approach is compared with classical and state-of-the-art neural network–based methods. Experimental results demonstrate that the proposed methods provide high detection accuracy, robustness to variations in vehicle scale and orientation, as well as effective object localization even under challenging aerospace imaging conditions and in multi-class scenes.

The implementation of the information technologies and methods for neural network–based recognition of vehicles with arbitrary spatial orientations developed within the scope of the dissertation research, as well as the software created on their basis, in the operational activities of enterprises has resulted in reduced aerospace data processing time, decreased resource intensity of analysis processes, and improved accuracy of automated object recognition.

**Keywords:** information technology, machine learning, neural network, aerospace imagery, recognition, vehicles, segmentation.

## СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

### *Публікації у виданнях, включених до переліку наукових фахових видань*

#### *України:*

1. Каштан В.Ю., **Казимиренко О.В.**, Гнатушенко В.В. Порівняльний аналіз ефективності методів розпізнавання автомобілів на аерокосмічних знімках. Control Systems and Computers. 2024. 4. С. 50-60. <https://doi.org/10.15407/csc.2024.04.050>.

### *Публікації в наукових виданнях, включених до наукометричних баз*

#### *Scopus та/або Web of Sc.:*

2. Radionov Y.D., Kashtan V.Yu., Hnatushenko V.V., **Kazymyrenko O.V.** Aircraft detection with deep neural networks and contour-based methods. The scientific journal «Radio Electronics, Computer Science, Control». – Zaporizhzhia, 2024. №1 4(71). – P.121-129. DOI [10.15588/1607-3274-2024-4-12](https://doi.org/10.15588/1607-3274-2024-4-12). [WoS]
3. Kashtan V. Yu., Hnatushenko V. V., Udovyk I. M., **Kazymyrenko O. V.**, Radionov Y. D. A neural network approach to semantic segmentation of vehicles in very high resolution images. The scientific journal «Radio Electronics, Computer Science, Control». – Zaporizhzhia, 2025. № 3(74). – P.77-85. DOI <https://doi.org/10.15588/1607-3274-2025-3-8> [WoS]
4. Kashtan V. Yu., **Kazymyrenko O.V.**, Hnatushenko V.V. Neural network method for invariant recognition of vehicles in aerospace images. Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu, 2026, № 1. P.130-137. DOI <https://doi.org/10.33271/nvngu/2026-1/130> [Scopus]

### *Публікації у матеріалах наукових конференцій:*

5. Hnatushenko V.V., Kashtan V. Yu., **Kazymyrenko O.V.** Information technology for detecting cars on aerial imaging using a modified YOLO-OBb architecture. MoDaST 2025: Modern Data Science Technologies Doctoral Consortium, June, 15, 2025, Lviv, Ukraine, Pp. 293-304. <https://ceur-ws.org/Vol-4005/paper20.pdf>. [Scopus]

6. **Казмиренко О.В.**, Гнатушенко В.В. Нейромережеве розпізнавання об'єктів транспорту за даними аерокосмічної зйомки. Матеріали VIII міжнародної науково-технічної конференції «Комп'ютерне моделювання та оптимізація складних систем». 1-3 листопада 2023 року, м. Дніпро, Україна. С.114.
7. **Казмиренко О.В.** Нейромережева модель YOLOv6 для виявлення автомобілів на аерокосмічному знімку. Автоматизація та біомедичні і комп'ютерні технології: тези доповідей Всеукраїнської науково-технічної інтернет-конференції. (Дніпро, 26 березня 2025 р.) / ДВНЗ «ПДТУ».– Дніпро: ПДТУ, 2025, С.156-158.
8. **Казмиренко О.В.** Огляд нейромережевих методів виявлення транспортних засобів. «Наукова весна» 2025: матеріали XV Міжнародної науково-технічної конференції аспірантів та молодих вчених, Дніпро, 26–28 березня 2025 року / Національний технічний університет «Дніпровська політехніка» – Дніпро : НТУ «ДП», 2025, С.187 -188.

*Авторське свідоцтво:*

9. Радіонов Є. Д., Каштан В. Ю., Гнатушенко В. В., **Казмиренко О. В.** Свідоцтво про реєстрацію авторського права на твір № 136518 від 23.05.2025 за наукову статтю «Aircraft detection with deep neural networks and contour-based methods». Свідоцтво про реєстрацію авторського права на твір № 136518 від 23.05.2025.

## ЗМІСТ

<b>ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ</b> .....	13
<b>ВСТУП</b> .....	14
<b>РОЗДІЛ 1 ОГЛЯД СТАНУ ПРОБЛЕМИ НЕЙРОМЕРЕЖЕВОГО РОЗПІЗНАВАННЯ ОБ'ЄКТІВ ТРАНСПОРТУ НА АЕРОКОСМІЧНИХ ЗОБРАЖЕННЯХ</b> .....	24
1.1. Технічні аспекти формування аерокосмічних зображень .....	24
1.2. Обробка аерокосмічних зображень.....	29
1.3. Класифікація та розпізнавання аерокосмічних зображень.....	32
1.3.1. Класифікація аерокосмічних зображень.....	32
1.3.2. Розпізнавання аерокосмічних зображень .....	34
1.4. Характеристики аерокосмічних зображень.....	36
1.5. Методи виявлення та розпізнавання транспортних засобів на аерокосмічних зображеннях .....	38
1.5.1. Згорткові нейронні мережі .....	44
1.5.2. Глибоке навчання .....	47
1.6. Проблеми та виклики розпізнавання об'єктів транспорту на аерокосмічних зображеннях .....	49
1.7. Постановка задачі та теоретичний апарат дослідження .....	52
1.8. Висновки до першого розділу.....	54
<b>РОЗДІЛ 2 НЕЙРОМЕРЕЖЕВІ МЕТОДИ РОЗПІЗНАВАННЯ ТРАНСПОРТНИХ ОБ'ЄКТІВ НА АЕРОКОСМІЧНИХ ЗОБРАЖЕННЯХ</b> .....	56
2.1. Нейромережевий підхід до семантичної сегментації транспортних засобів на аерокосмічних зображеннях надвисокого просторового розрізнення.....	56
2.2. Нейромережеве розпізнавання повітряних транспортних засобів на аерокосмічних зображеннях з урахуванням просторово-геометричних ознак.....	62

2.3. Нейромережеве розпізнавання автомобілів на аерокосмічних зображеннях з використанням модифікованої архітектури YOLO-OBV .....	66
2.4. Метод просторово-орієнтованого розпізнавання транспортних засобів довільного розташування .....	73
2.5. Висновки до другого розділу .....	78

### **РОЗДІЛ 3 РОЗРОБКА ТА ПРАКТИЧНА РЕАЛІЗАЦІЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ НЕЙРОМЕРЕЖЕВОГО РОЗПІЗНАВАННЯ ТРАНСПОРТНИХ ЗАСОБІВ ДОВІЛЬНОГО РОЗТАШУВАННЯ НА АЕРОКОСМІЧНИХ ЗОБРАЖЕННЯХ .....**

3.1. Інформаційна технологія нейромережевого розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях .....	80
3.2. Розробка програмного додатку розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях.....	87
3.3. Функціональні можливості програмного додатку.....	91
3.4. Висновки до третього розділу .....	96

### **РОЗДІЛ 4 ЕКСПЕРИМЕНТАЛЬНА ПЕРЕВІРКА ЕФЕКТИВНОСТІ ЗАПРОПОНОВАНИХ НЕЙРОМЕРЕЖЕВИХ МЕТОДІВ І ТЕХНОЛОГІЙ РОЗПІЗНАВАННЯ ТРАНСПОРТНИХ ОБ'ЄКТІВ НА АЕРОКОСМІЧНИХ ЗОБРАЖЕННЯХ .....**

4.1. Методика експериментальних досліджень.....	98
4.1.1. Набори даних .....	98
4.1.2. Формування власного набору даних .....	103
4.1.3. Кількісні метрики .....	108
4.2. Експериментальне дослідження ефективності методів розпізнавання автомобілів на аерокосмічних зображеннях .....	111
4.3. Експериментальна оцінка нейромережевого підходу до семантичної сегментації транспортних засобів на зображеннях надвисокого просторового розрізнення .....	118

4.4. Дослідження ефективності нейромережевого розпізнавання повітряних транспортних засобів з урахуванням просторово-геометричних ознак.....	124
4.5. Експериментальне дослідження нейромережевого розпізнавання автомобілів на аерокосмічних зображеннях з використанням архітектури YOLO-OBV .....	130
4.6. Оцінка ефективності нейромережевої технології розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях.....	137
4.7. Висновки до четвертого розділу.....	146
<b>ВИСНОВКИ .....</b>	<b>149</b>
<b>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ .....</b>	<b>151</b>
<b>ДОДАТОК А СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ АВТОРА ЗА ТЕМОЮ ДИСЕРТАЦІЇ .....</b>	<b>164</b>
<b>ДОДАТОК Б ДОКУМЕНТИ ЩОДО ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ ДОСЛІДЖЕНЬ .....</b>	<b>166</b>
<b>ДОДАТОК В АВТОРСЬКЕ СВІДОЦТВО .....</b>	<b>168</b>

## ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

- CNN Convolutional Neural Network (згорткова нейронна мережа);
- DNN Deep Neural Network (глибинна нейронна мережа);
- DOTA Dataset for Object Detection in Aerial Images;
- DL Deep Learning (глибоке навчання);
- FP False Positive (хибнопозитивне спрацювання);
- FN False Negative (хибнонегативне спрацювання);
- TP True Positive (істинно позитивне спрацювання);
- FPS Frames Per Second (кількість кадрів за секунду);
- GUI Graphical User Interface (графічний інтерфейс користувача);
- IoU Intersection over Union (міра перекриття областей);
- mAP mean Average Precision (середня точність за класами);
- RGB Red, Green, Blue (адитивна колірна модель);
- ROI Region of Interest (область інтересу);
- UAV Unmanned Aerial Vehicle (безпілотний літальний апарат);
- YOLO You Only Look Once (однокроковий алгоритм детектування об'єктів);
- БПЛА Безпілотний літальний апарат;
- ГЛОНАСС Глобальна навігаційна супутникова система (Global Navigation Satellite System);
- ДЗЗ Дистанційне зондування Землі;
- ЦСОЗ Цифрова система обробки зображень;
- БПЛА безпілотний літальний апарат;
- ІТ інформаційна технологія;
- ПЗ програмне забезпечення;
- НМ нейронна мережа;
- БД база даних;
- ПЗ програмне забезпечення;
- ТЗ транспортні засоби.

## ВСТУП

**Актуальність теми.** Розвиток аерокосмічних технологій характеризується стрімким зростанням обсягів зображувальної інформації, що отримується із супутників, літальних апаратів та безпілотних авіаційних систем. Використання сенсорів високої та надвисокої просторової розрізненості забезпечує детальну фіксацію об'єктів земної поверхні та створює передумови для вирішення широкого спектра прикладних задач, зокрема у сфері моніторингу транспортної інфраструктури, управління дорожнім рухом, безпеки, оборони, реагування на надзвичайні ситуації та ліквідації наслідків стихійних лих.

В умовах надзвичайних ситуацій, таких як землетруси, повені, зсуви або техногенні катастрофи, особливо критичним є оперативне отримання достовірної інформації про стан територій та наявність транспортних засобів. Як показують численні дослідження, перші 72 години після виникнення надзвичайної ситуації є вирішальними для збереження людських життів, що зумовлює необхідність швидкого та ефективного проведення пошуково-рятувальних операцій. Аерокосмічні зображення дозволяють забезпечити ситуаційну обізнаність у режимі, близькому до реального часу, що суттєво перевищує можливості традиційних наземних методів оцінювання, які є трудомісткими, обмеженими за доступністю та часовими витратами.

Окремим важливим напрямом застосування аерокосмічних зображень є автоматизований моніторинг дорожнього руху та транспортних потоків. Зростання інтенсивності трафіку та розвиток автономних транспортних систем зумовлюють підвищений попит на точні та масштабовані методи виявлення й аналізу транспортних засобів. На відміну від стаціонарних камер або наземних сенсорів, аерокосмічні зображення забезпечують охоплення великих територій одним сенсором і дозволяють одночасно фіксувати велику кількість об'єктів та їх просторову взаємодію.

Разом із тим автоматизоване розпізнавання транспортних засобів на аерокосмічних зображеннях є складною науково-прикладною задачею. Це зумовлено значною варіативністю умов зйомки, довільною орієнтацією об'єктів, різноманітністю масштабів, наявністю тіней, складного фону, оклюзій та спектральної неоднорідності. Особливої складності набуває аналіз різночасових зображень високої та надвисокої просторової розрізненості, для яких характерні істотні відмінності фотометричних і просторових характеристик навіть для одних і тих самих об'єктів.

Аналіз наукових публікацій, присвячених автоматизованому виявленню та розпізнаванню транспортних засобів на аерокосмічних зображеннях, свідчить про поступовий перехід від класичних підходів до методів глибинного навчання. У ранніх роботах значна увага приділялася використанню ознак ручного дешифрування у поєднанні з методами машинного навчання, що обмежувало ефективність розпізнавання в умовах складного фону, змін освітленості та високої щільності об'єктів.

Дослідження зарубіжних авторів, зокрема Liao W., Chen X., Yang J., Roth S., Goesele M., Yang M., та ін., спрямовані на застосування згорткових нейронних мереж (НМ) для виявлення транспортних засобів на зображеннях високої просторової розрізненості. Запропоновані підходи продемонстрували суттєве підвищення точності порівняно з традиційними методами, однак у більшості випадків орієнтувалися на прямокутні обмежувальні рамки та не враховували довільну орієнтацію об'єктів, що є характерною для аерокосмічних зображеннях із надірною та похилою зйомкою.

Таким чином, наведений аналіз сучасного стану досліджень підтверджує актуальність вирішення важливої науково-прикладної задачі розробки інформаційної технології розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях з використанням методів глибинного навчання. Реалізація такої технології дозволяє підвищити точність та надійність автоматизованого аналізу, забезпечити

швидкодію обробки великих обсягів даних і розширити можливості практичного використання аерокосмічних зображень у задачах використання аерокосмічних даних у задачах моніторингу транспортної інфраструктури, безпеки та територіального управління.

### **Зв'язок роботи з науковими програмами, планами, темами.**

Обраний напрямок досліджень пов'язаний із виконанням автором науково-дослідних робіт кафедри інформаційних технологій та комп'ютерної інженерії Національного технічного університету «Дніпровська політехніка» «Моделі й інформаційні технології обробки та аналізу даних в складних комп'ютерних системах і мережах» (реєстраційний номер 0121U114523, 2022-2024 рр.), «Методи та інформаційні технології інтелектуального аналізу неструктурованих даних в розподілених комп'ютерних системах» (реєстраційний номер 0125U000076).

**Мета і задачі дослідження.** Метою дисертаційного дослідження є підвищення ефективності інтелектуального розпізнавання транспортних засобів на аерокосмічних зображеннях високої роздільної здатності, в умовах змін масштабу, просторової орієнтації об'єктів та варіативних умов зйомки.

Для досягнення поставленої мети дослідження необхідно вирішити наступні задачі:

1. Провести аналіз сучасного стану методів виявлення та розпізнавання транспортних засобів на аерокосмічних зображеннях, зокрема класичних алгоритмів обробки зображень та нейромережевих підходів, і визначити їх переваги та обмеження з точки зору точності, інваріантності до просторових трансформацій та обчислювальної складності.

2. Дослідити ефективність класичних методів виділення ознак (Canny, Sobel, Laplacian, Prewitt, Scharr) та сучасних нейромережевих моделей виявлення об'єктів при розпізнаванні транспортних засобів на

аерокосмічних і аерофотознімках з використанням кількісних метрик Precision, Recall, F1 Score та *mAP*.

3. Розробити метод просторово-орієнтованого розпізнавання транспортних засобів, що базується на використанні глибоких нейронних мереж із механізмами автоматичної компенсації просторових трансформацій та багатомасштабного аналізу ознак.

4. Запропонувати методи попередньої обробки та аугментації аерокосмічних зображень, спрямовані на підвищення стійкості розпізнавання до змін масштабу, кута зйомки, освітлення та контрастності сцен.

5. Дослідити можливості застосування сучасних архітектур глибокого навчання для задач виявлення та сегментації транспортних засобів, зокрема моделей сімейства YOLO з підтримкою орієнтованих рамок (ОВВ) та сегментаційних архітектур типу DeepLab із глибокими магістральними мережами.

6. Розробити інформаційну технологію обробки та аналізу аерокосмічних зображень, яка інтегрує запропоновані методи розпізнавання транспортних засобів, модулі попередньої обробки, сегментації, виявлення та оцінки якості результатів.

7. Реалізувати програмний додаток, що забезпечує практичне застосування розробленої інформаційної технології, включаючи графічний інтерфейс користувача, засоби завантаження та обробки зображень, візуалізації результатів розпізнавання та аналізу показників ефективності.

8. Провести експериментальну перевірку розробленого методу та інформаційної технології на відкритих і спеціалізованих наборах даних аерокосмічних та БПЛА-знімків, а також здійснити порівняльний аналіз з існуючими одноетапними та двоетапними алгоритмами виявлення об'єктів.

9. Оцінити обчислювальну ефективність і можливість застосування розробленого програмного додатку в режимі, наближеному до

реального часу, що є важливим для задач моніторингу транспортної інфраструктури та безпеки.

**Об'єкт дослідження** – процес автоматизованого нейромережевого розпізнавання транспортних засобів на аерокосмічних зображеннях високої просторової розрізненості.

**Предмет дослідження** – методи та засоби інформаційної технології розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях високої просторової розрізненості з використанням глибинних нейронних мереж.

**Методи дослідження.** Теоретичну та методологічну основу роботи складають методи попередньої обробки та аналізу аерокосмічних зображень, методи комп'ютерного зору і машинного навчання, принципи побудови глибоких нейронних мереж, зокрема моделей сімейства YOLO та сегментаційних архітектур типу DeepLab, (Backbone Network + STN + Rotated RoI Align), методи виділення ознак (Canny, Sobel, Laplacian, Prewitt, Scharr) та механізми багатомасштабного аналізу.

**Наукова новизна одержаних результатів** полягає в тому, що автором:

**Вперше:**

– запропоновано комплексний багатокomпонентний підхід до нейромережевого розпізнавання транспортних засобів на аерокосмічних зображеннях високої просторової роздільної здатності, що ґрунтується на інтеграції просторової нормалізації ознак, виділення областей інтересу, багатомасштабного представлення ознак та семантичної сегментації в єдиній нейромережевій архітектурі для виявлення та класифікації об'єктів транспорту довільної орієнтації та масштабу;

– розроблено інформаційну технологію нейромережевого розпізнавання транспортних засобів на аерокосмічних зображеннях високої просторової роздільної здатності на основі запропонованої архітектури

глибоких нейронних мереж. Запропонована технологія забезпечує розпізнавання транспортних засобів довільного розташування на різночасових аерокосмічних зображеннях та оптимізує час навчання моделей і обробки даних.

### **Вдосконалено:**

- архітектуру нейромережевого детектора (YOLOv11) об'єктів із підтримкою орієнтованих обмежувальних рамок, що базується на багатомасштабному представленні просторових ознак та адаптивному механізмі їх агрегації й уточнення, що забезпечує підвищення точності виявлення та локалізації транспортних засобів у складних сценах на аерокосмічних зображеннях високої просторової розрізненості;

- методи автоматизованого виявлення транспортних засобів на аерокосмічних зображеннях за рахунок поєднання класичних алгоритмів виділення ознак із сучасними нейромережевими моделями, що забезпечує стабільність роботи системи за умов змінної освітленості, контрастності та спектральної неоднорідності сцен (Precision – 99,5%, Recall – 96,8%, F1 Score – 98,1%);

- підходи до підготовки та навчання нейромережових моделей на комбінованих наборах аерокосмічних даних (DOTA та спеціалізовані аерофотозображення) із застосуванням методів аугментації та функції втрат, що дозволяє зменшити вплив дисбалансу класів та підвищити якість виявлення об'єктів різних розмірів (*mAP* – 96,8%, *OBV IoU* – 98,5%);

- методику формування анотованих наборів даних із використанням орієнтованих обмежувальних рамок (OBV), що мінімізує вплив фону та забезпечує коректне навчання моделей для підвищення точності виявлення транспортних засобів. На спеціалізованому наборі даних модель досягла точності = 100%, FP = 0, Recall = 95,5% (виявлено 107 із 112 ТЗ).

### **Набули подальшого розвитку:**

- метод семантичної сегментації аерокосмічних зображень на основі глибоких згорткових нейронних мереж та функції втрат типу Dice, адаптовані до задач виділення транспортних засобів;

- комплексний підхід до автоматизованого розпізнавання транспортних засобів на аерокосмічних зображеннях високої просторової роздільної здатності, що інтегрує багатомасштабне представлення ознак, сегментацію контекстуальної інформації та орієнтовані механізми виявлення об'єктів, підвищуючи точність розпізнавання транспортних засобів у складних сценах.

**Практичне значення одержаних результатів** полягає у створенні інформаційної технології нейромережевого розпізнавання транспортних засобів на аерокосмічних зображеннях високої просторової роздільної здатності, яка забезпечує інтеграцію багатомасштабного аналізу ознак, орієнтоване виявлення об'єктів, семантичну сегментацію контекстуальної інформації та зручний графічний інтерфейс для візуалізації результатів та оцінки ефективності розпізнавання з мінімальними витратами часу і обчислювальних ресурсів.

Розроблена технологія включає такі основні компоненти:

- розроблений метод просторово-орієнтованого розпізнавання транспортних засобів довільного розташування із використанням багатокomпонентної нейромережевої архітектури YOLOv11-OBV та DeepLab з ResNet-101, що забезпечує високу точність локалізації та класифікації об'єктів різного масштабу і орієнтації у складних аерокосмічних сценах;

- розроблену автоматизовану інформаційну технологію обробки та аналізу аерокосмічних зображень, яка включає попередню обробку, сегментацію, орієнтоване виявлення транспортних засобів та оцінку результатів за ключовими метриками (Precision, Recall, F1 Score, *mAP*, OBV IoU), що дозволяє ефективно застосовувати технологію для моніторингу транспортної інфраструктури.

**Впровадження одержаних результатів.** Практичне значення результатів підтверджується впровадженнями результатів дисертаційної роботи у практику робіт «Дніпрокосмос» філії Національного центру управління та випробувань космічних засобів Державного космічного агентства України та впровадженням в освітній процес Національного технічного університету «Дніпровська політехніка».

**Особистий внесок здобувача.** Усі теоретичні та практичні результати, що складають наукову новизну дисертаційної роботи отримані особисто. Результати дисертаційної роботи опубліковані в у працях [1-9] (Додаток А).

У наукових публікаціях, створених у співавторстві, здобувачу належить: проведення порівняльного аналізу ефективності методів виявлення автомобілів на аерокосмічних зображеннях, який охоплює класичні градієнтні та контурні оператори [1]; розробка програмного забезпечення та проведення тестування методу розпізнавання повітряних транспортних засобів на оптичних космічних зображеннях високої роздільної здатності з використанням архітектури YOLOv8 [2]; метод семантичної сегментації аерокосмічних зображень на основі глибоких згорткових нейронних мереж із використанням функцій втрат типу Dice, адаптовані до задач виділення транспортних засобів [3]; комплексний багатокomпонентний підхід до нейромережевого розпізнавання транспортних засобів на аерокосмічних зображеннях високої просторової роздільної здатності [4]; вдосконалення архітектури нейромережевого детектора (YOLOv11) об'єктів із підтримкою орієнтованих обмежувальних рамок [5].

Особистий внесок інших співавторів у спільних публікаціях: у статті [1] – Каштан В. належить розробка методології порівняльного аналізу дослідження; Гнатушенко В. – здійснював математичне редагування; у статті [2] – Радіонову Є. (Radionov Y.) належить, вибір архітектури YOLOv8n для тестування, навчання моделі та оптимізація параметрів

розпізнавання; Каштан В. – програмна реалізація алгоритмів контурного виділення (Canny, Sobel, Laplacian та ін.); Гнатушенку В. (Hnatushenko V.) – підготовка аерокосмічних знімків для експериментальних досліджень та алгоритм попередньої обробки даних; у статті [3] Каштан В. (Kashtan V.) належить обґрунтування заміни Xception на ResNet-101 в архітектурі DeepLab v3+; Гнатушенку В. (Hnatushenko V.) – проведення тестування моделі на валідаційних вибірках; Удовик І. (Udovuk I.) – літературний огляд сучасних методів сегментації транспортних засобів на аерокосмічних знімках; Радіонову Є. (Radionov Y.) – проведення двоетапного навчання на цільових класах; у статті [4] Каштан В. (Kashtan V.) запропоновано використання інваріантних компонентів модулів Spatial Transformer Network та Rotated RoI Align в рамках нейромережевої інформаційної технології; Гнатушенку В. (Hnatushenko V.) – обґрунтування теоретичної бази для забезпечення стійкості моделі до повороту, масштабу та зміщення об'єктів; у статті [5] Каштан В. (Kashtan V.) запропоновано використання компонентів Freq-SpaFEM (частотно-просторове вилучення ознак) та DAGLF (динамічне групування ознак) для виявлення об'єктів складної геометрії в умовах щільної забудови; Гнатушенку В. (Hnatushenko V.) – статистичне порівняння метрик mAP@50, Precision та Recall.

**Апробація результатів дисертації.** Результати дисертаційної роботи доповідались і обговорювались на наукових семінарах кафедри інформаційних технологій та комп'ютерної інженерії Національного технічного університету «Дніпровська політехніка»; VIII міжнародній науково-технічній конференції «Комп'ютерне моделювання та оптимізація складних систем», 2023 р. (Дніпро, Україна); Всеукраїнській науково-технічній інтернет-конференції, 2025 р. (Дніпро, Україна); Міжнародній науково-технічній конференції аспірантів та молодих вчених, 2025 р. (Дніпро, Україна), Міжнародній конференції «Modern Data Science Technologies Doctoral Consortium», 2025 р. (Львів, Україна).

**Публікації.** Основні результати дисертаційної роботи опубліковано у 9 наукових працях. Чотири статті опубліковано у наукових виданнях, включених до переліку фахових видань України (всі індексуються у НМБД Index Copernicus), три з них – категорії А (Scopus, Web of Sc.), чотири наукових праці опубліковано у збірниках наукових праць та матеріалах міжнародних конференцій, одну з яких проіндексовано у НМБ Scopus; отримано одне свідоцтво про реєстрацію авторського права на твір.

**Структура і обсяг роботи.** Дисертаційна робота складається зі вступу, чотирьох розділів, висновків, списку використаних джерел та додатків. Загальний обсяг дисертації становить 167 сторінок, містить 150 сторінку основної частини, включає 46 рисунків, 10 таблиць, 142 літературних джерел.

# РОЗДІЛ 1

## ОГЛЯД СТАНУ ПРОБЛЕМИ НЕЙРОМЕРЕЖЕВОГО РОЗПІЗНАВАННЯ ОБ'ЄКТІВ ТРАНСПОРТУ НА АЕРОКОСМІЧНИХ ЗОБРАЖЕННЯХ

### 1.1. Технічні аспекти формування аерокосмічних зображень

У межах даної дисертаційної роботи під аерокосмічними зображеннями розуміємо двовимірні цифрові зображення земної поверхні, отримані в результаті дистанційної реєстрації власного або відбитого електромагнітного випромінювання за допомогою бортових технічних засобів, розміщених на аерокосмічних платформах, зокрема супутниках та безпілотних літальних апаратах. Такі зображення призначені для виявлення, якісного та кількісного аналізу об'єктів, явищ і процесів шляхом дешифрування, вимірювання та картографування. Формування аерокосмічних зображень забезпечує реєстрацію просторово узгоджених даних на великих площах земної поверхні, отриманих за однакових умов спостереження [1], що є важливим для коректного порівняльного аналізу та автоматизованого розпізнавання об'єктів.

Однією з характеристик аерокосмічних зображень є просторова розрізненість, яка визначається як мінімальний лінійний розмір об'єкта на земній поверхні, що може бути достовірно відображений на знімку. До зображень дуже низької просторової розрізненості належать дані, для яких характерне узагальнене відображення поверхні Землі з деталізацією в десятки кілометрів; на сучасному етапі такі зображення використовуються обмежено та переважно представлені радіометричними спостереженнями. Зображення низької просторової розрізненості, що забезпечують відображення об'єктів розмірами в декілька кілометрів, широко застосовуються в метеорологічних дослідженнях і формуються телевізійними та сканерними системами метеорологічних і ресурсних

супутників. Зображення середньої просторової розрізненості, які дозволяють відображати об'єкти розміром у сотні метрів, отримують за допомогою сканерної апаратури середньої розрізненості та теплових інфрачервоних сенсорів. Особливе значення для задач аналізу окремих об'єктів мають зображення високої та дуже високої просторової розрізненості, для яких характерна деталізація від десятків метрів до одиниць метрів і менше. Такі зображення формуються картографічними супутниками, орбітальними станціями та засобами аерозйомки з літальних апаратів. Саме цей клас аерокосмічних зображень є найбільш придатним для задач автоматизованого виявлення та розпізнавання транспортних засобів.

Серед джерел отримання зображень високої та дуже високої просторової розрізненості особливе місце посідають безпілотні літальні апарати (БПЛА), які дозволяють здійснювати зйомку з малих висот, оперативність отримання даних та високу деталізацію сцен спостереження. Управління БПЛА та передача отриманих зображень реалізується за допомогою двох основних каналів (рис. 1.1) [2, 3] : прямий канал (рис.1.1 а) – безпосередній обмін даними між БПЛА та цифровою системою обробки зображень (ЦСОЗ) у межах радіовидимості [4]; супутниковий канал (рис.1.1б) – передача даних через ретранслятор (супутник), що дозволяє керувати апаратом на глобальних відстанях.

Окрему групу аерокосмічних даних становлять зображення, отримані з використанням безпілотних літальних апаратів. Формування таких зображень пов'язане із застосуванням супутникових навігаційних систем [1], які забезпечують просторову прив'язку знімків, визначення координат носія знімальної апаратури та параметрів його орієнтації у просторі. Просторове положення платформи зйомки описується трьома основними координатами: географічною широтою, географічною довготою та висотою над середнім рівнем моря. Географічна широта характеризує положення об'єкта у напрямку північ–південь і змінюється в межах від  $-90^\circ$  до  $+90^\circ$ ,

тоді як географічна довгота визначає положення у напрямку захід–схід відносно нульового меридіана і змінюється в діапазоні від  $-180^\circ$  до  $+180^\circ$ . Висота задає відстань платформи зйомки від середнього рівня моря та, як правило, вимірюється в метрах [5].



а)



б)

Рисунок 1.1 – Принцип управління БПЛА з використанням з використанням цифрової системи обробки зображень (ЦСОЗ): а) прямий

канал передачі даних між БПЛА та ЦСОЗ; б) канал передачі даних між БПЛА та ЦСОЗ через супутник

Крім просторового положення, для коректного формування аерокосмічних зображень необхідно враховувати орієнтацію літального апарата у просторі. Вона задається поворотами апарата навколо трьох взаємно перпендикулярних осей: навколо поздовжньої осі, навколо поперечної осі, навколо вертикальної осі. Сукупність цих осей утворює ортонормовану систему координат, що дозволяє однозначно описати кутове положення носія знімальної апаратури.

Для опису орієнтації часто застосовується опорна система координат типу «північ–схід–вниз». У цій системі початок координат розташовується в центрі літального апарата, поздовжня вісь, спрямована на північ, поперечна вісь – на схід, а вертикальна вісь є перпендикулярною до них і спрямована вниз, у бік центра Землі. Така система координат є правобічною та зручною для навігаційних і фотограмметричних задач.

Власна система координат літального апарата визначається його миттєвою орієнтацією у просторі: поздовжня вісь спрямована вздовж напрямку руху, поперечна вісь – праворуч від нього, а вертикальна вісь спрямована вниз відносно корпусу апарата. Для переходу від системи координат NED до бортової системи координат використовуються три кути повороту: кут повздовжньої осі  $\phi$ , кут поперечної осі  $\theta$  та кут вертикальної осі  $\psi$ .

Процес перетворення координат реалізується шляхом послідовного обертання навігаційної системи координат навколо відповідних осей: спочатку навколо вертикальної осі, далі – навколо поперечної осі і, на завершальному етапі, навколо повздовжньої осі. У результаті формується матриця повороту розмірності  $R 3 \times 3$ , яка використовується для узгодження координат точок місцевості, параметрів зйомки та геометрії аерокосмічного зображення.

Для опису геометричних властивостей об'єктів та їх перетворень у просторі традиційно використовується  $n$ -вимірний евклідів простір  $\mathbb{R}^n$ , у якому будь-яка точка задається вектором [6]:

$$\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n. \quad (1.1)$$

Оскільки реальні сцени зйомки є тривимірними, особливе практичне значення має евклідів простір  $\mathbb{R}^3$ . Водночас у задачах проєкції, зокрема при моделюванні формування зображень камерою, доцільним є використання проєктивного простору, який спрощує аналітичний запис геометричних перетворень.

Проєктивний простір  $\mathbb{P}^n$  визначається як [1, 6]:

$$\mathbb{P}^n := \mathbb{R}^{n+1} \setminus \{0\}, \quad (1.2)$$

тобто як множина ненульових векторів розмірності  $n+1$ . Він розширює евклідів простір шляхом введення додаткової проєктивної координати  $w$ . Точка:

$$\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_n, \tilde{w})^T \in \mathbb{P}^n, \quad (1.3)$$

відповідає точці евклідового простору:

$$\mathbf{x} = \left( \frac{\tilde{x}_1}{\tilde{w}}, \dots, \frac{\tilde{x}_n}{\tilde{w}} \right)^T \in \mathbb{R}^n. \quad (1.4)$$

Координати  $\mathbf{x}$  називають декартовими, тоді як  $\tilde{\mathbf{x}}$  – однорідними. Важливою властивістю проєктивного простору є те, що всі вектори описують одну й ту саму точку евклідового простору. Завдяки цьому при переході до однорідних координат проєктивну координату зазвичай фіксують як  $w=1$ , якщо не зазначено інакше. Надалі в роботі вектори з тильдою використовуються для позначення однорідних координат, а без тильди – для відповідних векторів евклідового простору.

Використання однорідного подання точок у проєктивному просторі дає змогу уніфіковано описувати геометричні перетворення в тривимірному просторі, що є особливо важливим у задачах фотограмметрії та формування аерокосмічних зображень. У такому поданні будь-яке просторове перетворення задається у вигляді матриці, яка одночасно враховує як

обертання, так і лінійні зсуви, а застосування перетворення до точки реалізується шляхом матричного множення.

Процес формування аерокосмічного зображення описується камерною моделлю, яка встановлює відповідність між координатами точок сцени у світовій системі координат та їх проєкціями на двовимірну площину сенсора. Камерна модель включає зовнішні параметри, що визначають положення та орієнтацію камери, а також внутрішні параметри, які характеризують оптичні властивості знімальної апаратури. Сукупність цих параметрів забезпечує однозначний опис перспективної проєкції об'єктів місцевості на площину зображення [7 – 9].

У системах аерокосмічного моніторингу з використанням безпілотних літальних апаратів найбільш поширеними є монокулярні оптичні камери, що формують зображення у видимому спектральному діапазоні. Такі сенсори характеризуються компактністю, відносною простотою калібрування та можливістю отримання зображень високої просторової розрізненості, що робить їх придатними для детального аналізу земної поверхні. Альтернативні рішення, зокрема багатокамерні та глибинні системи, хоча й забезпечують додаткову просторову інформацію, мають суттєві обмеження щодо масо-габаритних характеристик, енергоспоживання та складності інтеграції на легкі аерокосмічні платформи.

З огляду на зазначені особливості, монокулярні RGB-камери залишаються одним із найбільш доцільних джерел вхідних даних для задач автоматизованого аналізу аерокосмічних зображень, зокрема у задачах виявлення та локалізації транспортних засобів.

## **1.2. Обробка аерокосмічних зображень**

Реалізація процедур обробки аерокосмічних зображень забезпечує підготовку даних до подальшого аналізу, усунення шумів і атмосферних

спотворень, корекцію геометричних та радіометричних параметрів, а також покращення контрастності й деталізації об'єктів. Така підготовка є необхідною перед застосуванням алгоритмів класифікації та розпізнавання, оскільки якість вихідних даних безпосередньо визначає точність і надійність автоматичного виділення об'єктів та їх категоризації. Обробка аерокосмічних зображень включає реалізацію різних процедур класифікації, як на основі попереднього навчання моделей, так і без використання навчальних вибірок. Комп'ютерна обробка цифрових аерокосмічних зображень відкриває нові можливості, такі як підготовка зображень для моніторингу, покращення їх якості (наприклад, усунення атмосферних впливів), синтез кольорових зображень, автоматизована обробка та отримання кількісних даних (координати, відстані, площі тощо). Результати такої обробки є основою для створення картографічних продуктів, які можуть зберігатися в цифровому вигляді або друкуватися на паперових носіях [10, 11].

Цифрове зображення складається з сегментів – пікселів, які формують сітку рядків і стовпців. Кожен піксель має координати та характеризується значенням яскравості, що визначається у умовних одиницях. Значення яскравості пов'язане зі здатністю об'єктів поверхні Землі відбивати сонячне випромінювання, що впливає на точність подальшої обробки. Основні результати обробки цифрових зображень включають [10]:

- визначення характеристик яскравості об'єктів;
- роботу з багатоканальними зображеннями;
- синтез кольорових зображень;
- автоматизовану обробку за допомогою спеціалізованого програмного забезпечення;
- проведення вимірювань за зображеннями (координати, площі, відстані тощо).

Алгоритм обробки аерокосмічних зображень, показаний на рис. 1.2, ілюструє послідовність основних етапів роботи з цифровими даними:

- попередня підготовка зображень – усунення шумів та атмосферних впливів, калібрування яскравості;
- геоприв'язка – відображення зображень у заданій системі координат для точного співставлення з іншими картографічними даними;
- класифікація об'єктів – автоматичне визначення типів поверхні або об'єктів за спектральними та текстурними характеристиками;
- постобробка та аналіз – об'єднання результатів класифікації, усунення помилок, генерація кількісних даних;
- формування картографічних продуктів – створення карт, цифрових моделей та інших аналітичних матеріалів для використання в моніторингу та управлінні.



Рисунок 1.2 – Алгоритм обробки аерокосмічних зображень

### **1.3. Класифікація та розпізнавання аерокосмічних зображень**

Класифікація аерокосмічних зображень є одним із базових етапів обробки даних дистанційного зондування Землі та полягає у формуванні впорядкованої системи, що забезпечує віднесення значень пікселів зображення до інформативно значущих класів, які відповідають певним об'єктам або явищам земної поверхні.

У контексті розпізнавання транспортних об'єктів класифікація забезпечує виділення автомобілів, залізничних складів, суден, літальних апаратів та елементів транспортної інфраструктури на аерокосмічних зображеннях.

#### **1.3.1. Класифікація аерокосмічних зображень**

Класифікацію доцільно поділяти на три основні категорії залежно від ступеня участі оператора та рівня автоматизації [10]:

- автоматичні методи;
- ручні методи;
- гібридні методи.

Ручна класифікація базується на візуальному аналізі зображень фахівцем-інтерпретатором з використанням ознак форми, текстур, тону, контексту та просторового розташування об'єктів. Такий підхід характеризується високою точністю для локальних задач, однак є трудомістким, суб'єктивним і малопридатним для обробки великих масивів даних.

Автоматична класифікація передбачає використання алгоритмів цифрової обробки зображень, методів машинного навчання та нейромережових моделей для віднесення пікселів або сегментів зображення до заданих класів без безпосереднього втручання оператора. Саме автоматичні методи на сьогодні є найбільш перспективними та широко

застосовуваними у задачах просторової класифікації, зокрема при обробці високороздільних аерокосмічних зображень, що містять складні об'єкти транспортної інфраструктури (рис. 1.3).

Гібридні методи поєднують автоматичні алгоритми з елементами експертного аналізу та корекції результатів, що дозволяє підвищити достовірність класифікації у складних умовах, наприклад за наявності шумів, тіней, перекриттів об'єктів або різномірних спектральних характеристик.

Класифікація аерокосмічних зображень належить до ключових етапів обробки просторових даних, оскільки забезпечує перетворення первинної растрової інформації на впорядковані тематичні набори, придатні для подальшої аналітичної інтерпретації. Отримані результати використовуються при формуванні тематичних карт, виконанні просторового аналізу та інформаційному забезпеченні процесів планування, моніторингу й управління транспортними системами.



Рисунок 1.3 – Процес класифікації аерокосмічних даних

Стандартний підхід до класифікації аерокосмічних зображень передбачає послідовність взаємопов'язаних етапів, які забезпечують

інтеграцію даних дистанційного зондування з результатами польових досліджень та аналітичної обробки.

### **1.3.2. Розпізнавання аерокосмічних зображень**

Розпізнавання об'єктів полягає у виявленні конкретного об'єкта на зображенні, зазвичай за його контурами, формою або іншими характерними ознаками, без визначення більш широкого класу або категорії. Розпізнавання часто використовується як попередній етап для подальшої класифікації, коли знайдені об'єкти підлягають віднесенню до конкретного класу (наприклад, тип автомобіля чи модель судна). Таким чином, розпізнавання можна розглядати як підзадачу класифікації, орієнтовану на локалізацію об'єкта, тоді як класифікація визначає його приналежність до категорії або класу.

Серед методів розпізнавання виділяють кілька основних підходів, що відрізняються принципом роботи, точністю та вимогами до обчислювальних ресурсів. Метод Віоли–Джонса призначений для розпізнавання об'єктів на зображенні і базується на принципі скануючого вікна [12]. Основною ідеєю методу є інтегральне представлення зображення, що дозволяє швидко обчислювати сумарну яскравість прямокутників на зображенні та виділяти ознаки Хаара для виявлення об'єкта [13]. Для прийняття рішення ознаки передаються на вхід класифікатора, який навчається з використанням бустинг-методів, що дозволяє створювати складніші класифікатори на основі композиції простих. Алгоритм методу складається з етапів: вибірка навчальних даних і обчислення примітивів Хаара, визначення порогів для класифікації, відбір кращих примітивів і порогів для кожного примітива. Метод Віоли–Джонса є популярним, проте його застосування у системах реального часу та на обмежених ресурсах ускладнене через високу обчислювальну складність та потребу у великому обсязі навчальних даних. Існують численні модифікації цього методу, зокрема поєднання з алгоритмами виділення шкіри для

виявлення обличчя, адаптивне скануюче вікно, використання операторів Лапласа та гаусового фільтра для підвищення точності розпізнавання складних об'єктів [14]. Контурний аналіз є групою методів, що використовуються для виділення, опису, збереження та пошуку об'єктів за їхніми контурами. Контурна інформація часто достатня для ідентифікації об'єкта, що зменшує обчислювальну складність. Серед найбільш поширених методів виділяють:

- метод активних контурів, де контур ініціалізується як лінія і деформується до границі об'єкта, максимізуючи «енергію» моделі;
- детектор границь Кенні, який включає розмивання зображення, визначення градієнтів, пошук локальних максимумів, подвійне порогове відсікання та формування підсумкових контурів [15];
- метод прослідковування контурів, при якому скануюча точка рухається по зображенню для знаходження меж між об'єктом і фоном.

Методи пошуку шаблону призначені для виділення областей на зображенні, що максимально збігаються із заданим зразком. Заданий шаблон переміщується по зображенню, оцінюється ступінь відповідності та виділяється область із найбільшим збігом.

Методи на основі ключових точок спрямовані на виділення характерних точок об'єкта і оцінку їх взаємного розташування. Процес включає три етапи: детектор точок, формування дескрипторів ознак та порівняння точок (матчінг). Деякі підходи поєднують виділення триангульованих структур на основі дескрипторів для зменшення обсягу перебору при зіставленні аерофотознімків.

Згорткові нейронні мережі (CNN) є сучасним і ефективним інструментом розпізнавання об'єктів. Вони забезпечують стійкість до різних видів спотворень зображення, таких як повороти, зміни масштабу, ракурсу або шумів. Основна відмінність CNN від повнозв'язних мереж полягає у локальній пов'язаності нейронів та структурі шарів на «простий» і «складний» підшари: перший виявляє характерні ознаки об'єкта, другий

узагальнює їх для класифікації. Для навчання використовуються методи з учителем, зокрема зворотне поширення помилки, а також без нагляду, наприклад patch-based training. CNN дозволяють не лише визначати наявність об'єктів, а й локалізувати пікселі, які їм належать, за допомогою архітектур типу VGG, ResNet, SegNet та їх модифікацій, включно з Bayesian SegNet, а також адаптованих AlexNet та GoogLeNet для задач сегментації.

#### **1.4. Характеристики аерокосмічних зображень**

Розуміння спектральних і просторових характеристик зображень безпосередньо впливає на вибір методів виявлення та розпізнавання об'єктів, оскільки ефективність алгоритмів залежить від роздільної здатності, контрастності та спектральної інформації. Зважаючи на це, розробка методів виявлення та розпізнавання транспортних засобів і інших об'єктів земної поверхні потребує адаптації алгоритмів до конкретних характеристик даних.

У процесі роботи з аерокосмічними зображеннями одним із чинників є спектральний діапазон зйомки, оскільки саме він зумовлює характер відображення об'єктів земної поверхні та можливості подальшої обробки отриманих даних. Спектральні характеристики зйомки визначають особливості відображення фізичних і спектральних властивостей об'єктів земної поверхні на аерокосмічних зображеннях, а також впливають на технологію формування цифрових даних.

Зазначені особливості визначають візуальні, радіометричні та геометричні властивості аерокосмічних зображень, які, у свою чергу, формують основу для їх класифікації з урахуванням можливостей автоматизованої та нейромережевої обробки. Таким чином, спектральний діапазон зйомки виступає базовим рівнем класифікації аерокосмічних зображень, оскільки відображає відбивні та випромінювальні характеристики об'єктів, зафіксованих сенсорами.

З цієї точки зору виділяють три основні типи аерокосмічних зображень [6, 10]:

- зображення у видимому, ближньому та середньому інфрачервоному діапазонах, які формуються за рахунок відбитого сонячного випромінювання;
- зображення у тепловому інфрачервоному діапазоні, що відображають власне теплове випромінювання об'єктів;
- зображення у радіочастотному діапазоні, отримані з використанням активних радіолокаційних систем.

Кожен із зазначених типів аерокосмічних зображень характеризується специфічними інформаційними властивостями та обмеженнями, що визначають доцільність їх використання для різних прикладних задач, зокрема для виявлення, класифікації та розпізнавання транспортних об'єктів.

У загальному розумінні класифікація розглядається як процедура віднесення об'єкта до одного з наперед визначених класів на основі наявності у нього певних ознак. Об'єкти вважаються такими, що належать до одного класу, якщо вони характеризуються подібними спектральними, просторовими, текстурними або контекстними характеристиками. У результаті класифікації множина об'єктів поділяється на окремі класи, що відображають їх семантичну належність.

Процес класифікації аерокосмічних зображень у загальному випадку включає кілька послідовних етапів. На першому етапі здійснюється ідентифікація основних розрізнявальних ознак, які відображають відмінності між об'єктами різних класів. Далі формується просторове представлення цих ознак, що враховує як спектральну інформацію, так і геометричні характеристики об'єктів.

Наступним етапом є визначення або обчислення значень відповідних ознак для досліджуваних об'єктів, на основі яких здійснюється їх класифікація. Завершальним етапом є прийняття рішення щодо належності

кожного об'єкта до одного з класів із використанням певного вирішального правила або моделі, зокрема методів машинного навчання та нейромережових підходів.

Зазначена послідовність етапів створює методологічну основу для розроблення та застосування сучасних алгоритмів автоматизованої класифікації аерокосмічних зображень, що є критично важливим для задач нейромережового розпізнавання об'єктів транспорту в умовах великої різноманітності сенсорних даних і складних сцен зйомки.

### **1.5. Методи виявлення та розпізнавання транспортних засобів на аерокосмічних зображеннях**

Враховуючи зазначені фізичні, спектральні та просторові особливості аерокосмічних знімків, з'являється потреба у спеціалізованих методах автоматизованого виявлення та розпізнавання об'єктів, зокрема транспортних засобів. Використання таких методів забезпечує підвищення точності і надійності аналізу, що є критично важливим для задач моніторингу інфраструктури, управління транспортними потоками та підтримки рішень у містобудівних і екологічних дослідженнях.

Виявлення і розпізнавання об'єктів на аерокосмічних зображеннях, отриманих за допомогою дистанційного зондування, є важливим та складним завданням [16]. Воно використовується для ідентифікації штучних об'єктів (дороги, будівлі, транспортні засоби, кораблі) [17] та природних елементів (дерева, ліси) і застосовується для оновлення географічних інформаційних систем, моніторингу навколишнього середовища, картографування зон забруднення, точного землеробства та міського планування.

Штучні об'єкти зазвичай мають чіткі контури, а природні – розмиті та нечіткі, що ускладнює процес їх виділення. Додаткові труднощі виникають через зміни кута зйомки, часткову оклюзію, складний фон, освітленість [18]

та тінь. Низька просторова роздільна здатність попередніх знімків обмежує точність виявлення об'єктів, що стимулює розвиток методів вилучення характеристик областей для підвищення ефективності алгоритмів розпізнавання.

Протягом останніх десятиліть було розроблено широкий спектр методів для виявлення транспортних засобів на аерокосмічних зображеннях та реалізується двома основними підходами [19 – 22]:

- традиційні методи, які базуються на заздалегідь визначених ознаках (handcrafted features). До них відносяться методи, що використовують форму, контури, текстуру або спектральні характеристики транспортних засобів для їх детекції;

- методи на основі глибокого навчання, які дозволяють автоматично виділяти ознаки транспортних засобів із зображень та забезпечують вищу точність і стійкість до змін масштабу, освітлення або орієнтації об'єктів.

На рис. 1.4 представлено узагальнену класифікацію методів розпізнавання об'єктів на зображеннях з високим просторовим розрізненням [23]. До класичних підходів належать оператори Canny, Sobel, Laplacian, Prewitt та Scharr [23], принцип дії яких ґрунтується на аналізі градієнтів яскравості для виділення контурів об'єктів. Виділення контурної інформації є важливим етапом попередньої обробки, оскільки коректне визначення меж об'єктів безпосередньо впливає на точність їх подальшої ідентифікації. Зазначені методи демонструють задовільну ефективність за умов однорідного освітлення та низького рівня шуму. Водночас у випадках наявності інтенсивних шумів, тіньових зон або різких перепадів яскравості їх результативність суттєво погіршується, що ускладнює процес розпізнавання та може призводити до зниження точності результатів.

На відміну від класичних алгоритмів, сучасні нейромережеві підходи, зокрема моделі сімейства YOLO (You Only Look Once) [24], забезпечують високі показники швидкодії та точності, оскільки в межах одного

обчислювального циклу виконують як локалізацію, так і класифікацію об'єктів на зображенні.

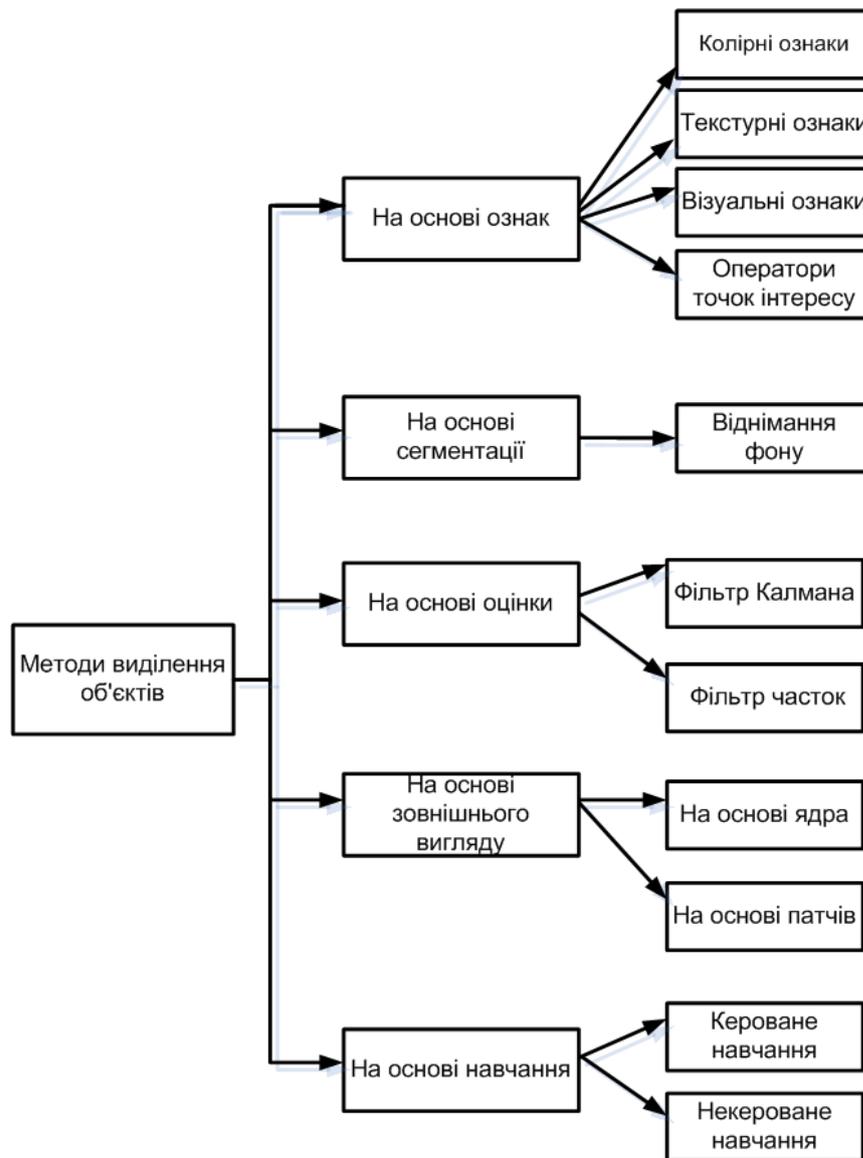


Рисунок 1.4 – Систематизація підходів до ідентифікації об'єктів на зображеннях високої роздільної здатності [25]

Методи, засновані на використанні ознак, передбачають виділення об'єктів за кольоровими, текстурними та візуальними характеристиками, а також із застосуванням операторів точок інтересу. Такий підхід дозволяє здійснювати ідентифікацію об'єктів на основі їхніх індивідуальних властивостей. На початковому етапі формуються ключові ознаки, зокрема текстурні, градієнтні та кольорові характеристики, які мають бути достатньо інформативними та унікальними для надійного представлення об'єктів у

просторі ознак. Комбінування декількох типів ознак сприяє підвищенню стійкості методів до змін умов освітлення, масштабування, обертання та часткових перекриттів об'єктів.

Колір є однією з базових ознак, що широко застосовується в задачах виявлення та розпізнавання об'єктів. У сучасних методах колірна інформація часто подається у вигляді кольорових гістограм, які використовуються для моделювання об'єкта. Перевагами такого підходу є відносно невеликі обчислювальні витрати, а також інваріантність до змін масштабу, обертання та часткової оклюзії. У роботі [26] запропоновано метод, що поєднує фонові моделі з градуйованими характеристиками назв кольорів у рамках алгоритму MeanShift. Застосування цього підходу дозволяє підвищити ефективність розпізнавання за рахунок адаптації до змін освітлення та кольорових властивостей об'єкта. Інтеграція градуйованих кольорових ознак сприяє формуванню більш стійких і точних моделей, що позитивно впливає на якість розпізнавання в реальному часі.

Локальні особливості розглядаються як характерні шаблони зображення, які дозволяють відрізнити об'єкт від фону. До таких особливостей можуть належати ключові точки, фрагменти контурів або локальні області зображення. У разі використання координат, отриманих на етапі виділення об'єктів, для подальшого аналізу ці елементи інтерпретуються як точки інтересу. Протягом останніх років було запропоновано низку модифікацій алгоритму SIFT. Початково цей метод орієнтувався на обробку монохромних зображень, однак згодом його було адаптовано до кольорових даних шляхом упровадження кольорового локального інваріантного дескриптора ознак (CSIFT). Використання CSIFT забезпечує вищу стійкість до змін кольору та фотометричних умов порівняно з класичним SIFT. У дослідженні [27] наведено порівняльний аналіз масштабно-інваріантного перетворення ознак (SIFT) і прискореного робастного перетворення ознак (SURF), а також запропоновано

використання геометричних перетворень для виявлення ключових точок у межах описуваних областей.

Текстурні ознаки визначаються варіаціями інтенсивності поверхні та кількісно описують такі властивості, як регулярність і гладкість, що є важливими для ідентифікації текстурних об'єктів, зокрема аеропортів [28], будівель, міських територій і транспортних засобів. Однією з найбільш досліджених характеристик текстури є вейвлети Габора, основною перевагою яких є інваріантність до змін освітлення, обертання, масштабу та зсуву. Це робить фільтри Габора ефективним інструментом для задач розпізнавання й відстеження об'єктів. Іншим поширеним текстурним дескриптором є локальні бінарні патерни (LBP), які широко використовуються в задачах розпізнавання облич та аналізу руху. Оператор LBP, інваріантний до обертання, характеризує локальну текстуру та не залежить від градацій сірого, що забезпечує його ефективність у різних системах комп'ютерного зору.

До методів, заснованих на використанні ознак, також належать градієнтні оператори Canny, Sobel, Laplacian, Prewitt і Scharr, які забезпечують виявлення контурів об'єктів. Контурна інформація є важливою для подальшого аналізу та ідентифікації, оскільки використання градієнтів яскравості дозволяє точно визначати форму й межі об'єктів. Такі методи особливо ефективні у випадках, коли об'єкти мають чітко виражені краї.

Сегментація зображень є одним із ключових етапів візуального аналізу. Процес сегментації переднього плану спрямований на відокремлення об'єктів інтересу від фонові сцени. Як правило, до об'єктів переднього плану належать елементи з високим контрастом або значною варіативністю відносно фону, що зумовлює необхідність їх чіткого виділення для подальшого аналізу. У роботі [29] запропоновано метод об'єктного аналізу зображень дистанційного зондування, який поєднує алгоритми обробки зображень із функціональними можливостями

геоінформаційних систем для інтеграції спектральної та контекстної інформації. У дослідженні [29] наведено огляд підходів до сегментації, де на прикладах практичних застосувань показано, що використання сегментаційних та об'єктно-орієнтованих методів забезпечує вищу точність аналізу й класифікації порівняно з традиційними піксельними підходами. На відміну від піксельної обробки, об'єкти зображень можуть характеризуватися розширеним набором атрибутів, зокрема формою, контекстом і просторовими взаємозв'язками між регіонами.

Методи, що базуються на оцінюванні або фільтрації, формулюють задачу розпізнавання об'єктів як задачу оцінювання стану, де об'єкт описується вектором стану. Байєсівські підходи забезпечують універсальну математичну основу для розв'язання задач оцінювання динамічного стану. Функція щільності ймовірності вектора стану обчислюється з урахуванням усієї наявної інформації на момент оцінювання, а апостеріорний розподіл відображає повну інформацію про стан об'єкта. Для цього використовуються модель стану, яка описує еволюцію системи, та модель спостереження, що встановлює зв'язок між вимірюваннями і станом об'єкта. Процес фільтрації включає етапи прогнозування та корекції, на яких відповідно оцінюється майбутній стан і уточнюється поточна оцінка на основі спостережень.

Типовими представниками Байєсівських методів є фільтр Калмана та фільтр частинок, які широко застосовуються для оцінювання траєкторій руху об'єктів на зображеннях і дозволяють ефективно інтегрувати дані з різних джерел з метою підвищення точності оцінювання.

Слід зазначити, що об'єкти реального світу характеризуються складною структурою та можуть змінювати свою форму з часом. Для надійного розпізнавання й відстеження таких об'єктів алгоритми повинні враховувати варіації їх зовнішнього вигляду та характеристик. Попри значну кількість наукових досліджень, проблема побудови адекватних моделей, здатних описувати динамічні зміни ознак об'єктів, залишається

актуальною, що ускладнює створення високонадійних систем комп'ютерного зору [25].

### 1.5.1. Згорткові нейронні мережі

Згорткові нейронні мережі (CNN) – це спеціалізований тип нейронних мереж, призначених для обробки багатовимірних даних із відомою сітчастою топологією, наприклад, зображень чи відеокадрів. Типові CNN складаються з трьох основних типів шарів: згорткових шарів, шарів підвибірки (pooling) та повнозв'язних шарів [30]. Щоб зменшити кількість параметрів і, відповідно, складність мережі, згорткові шари використовують дві основні ідеї: замість з'єднання кожного нейрона з усіма нейронами попереднього шару, нейрони з'єднуються лише з нейронами в межах локальної області. Таким чином, матриця ваг стає розрідженою. Крім того, ваги спільні для всіх нейронів шару, що робить їх незалежними від позиції нейрона (рис.1.5).

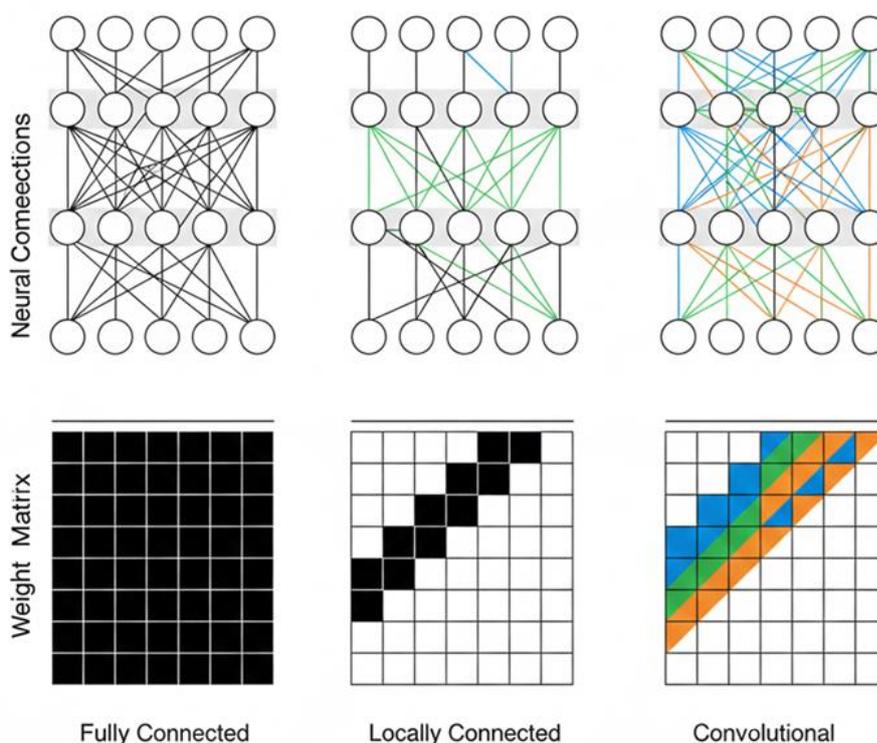


Рисунок 1.5 – Перехід від повнозв'язного шару до 1D згорткового шару та відповідної матриці ваг

Згорткові шари складаються з набору навчених фільтрів (каналів), оскільки одиночні фільтри не здатні захопити достатньо інформації з попереднього шару. Кожен фільтр характеризується розміром ядра, що визначає локальне поле зору (receptive field). Вага не спільна між фільтрами, тому кожен фільтр навчається окремо. Результатом роботи шару є карта ознак, що містить вектор ознак у кожній позиції, розмірність якого дорівнює кількості фільтрів  $C$  у шарі. Для багатоканальних входів  $D$ , кожен фільтр складається з  $D$  ядер розміром  $k \times k$ , які посуваються по відповідному каналу і результати підсумовуються (рис.1.6).

Для наближення складних функцій прийняття рішень у CNN застосовують нелінійності за допомогою активаційних функцій. Найчастіше використовують ReLU:

$$\phi(x) = \max(0, x). \quad (1.5)$$

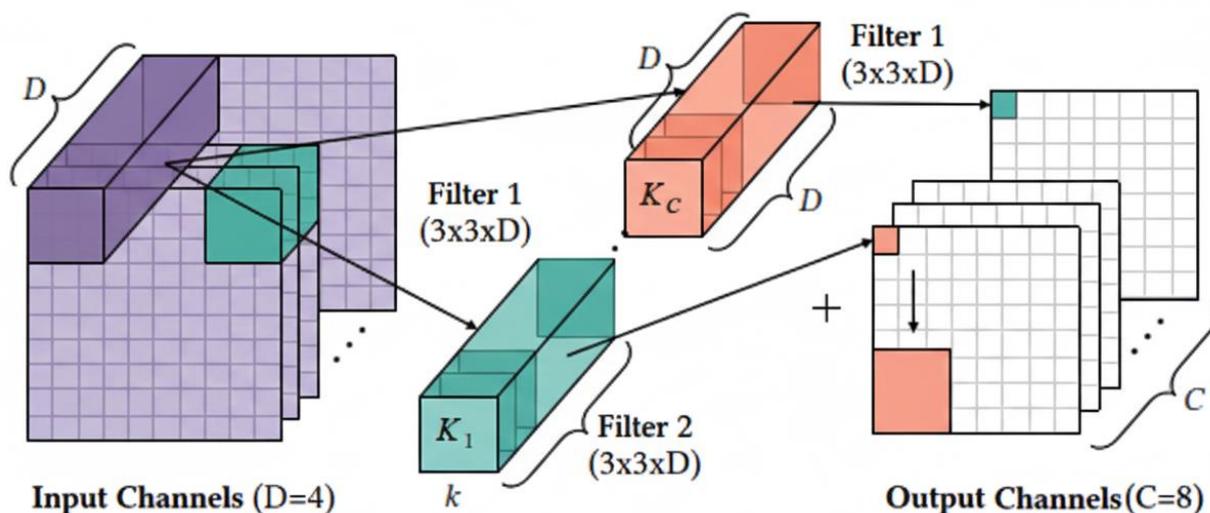


Рисунок 1.6 – Схематичне зображення згортки [31]

Функція ReLU замінює всі негативні значення карти ознак на нуль (рис. 1.7) [32].

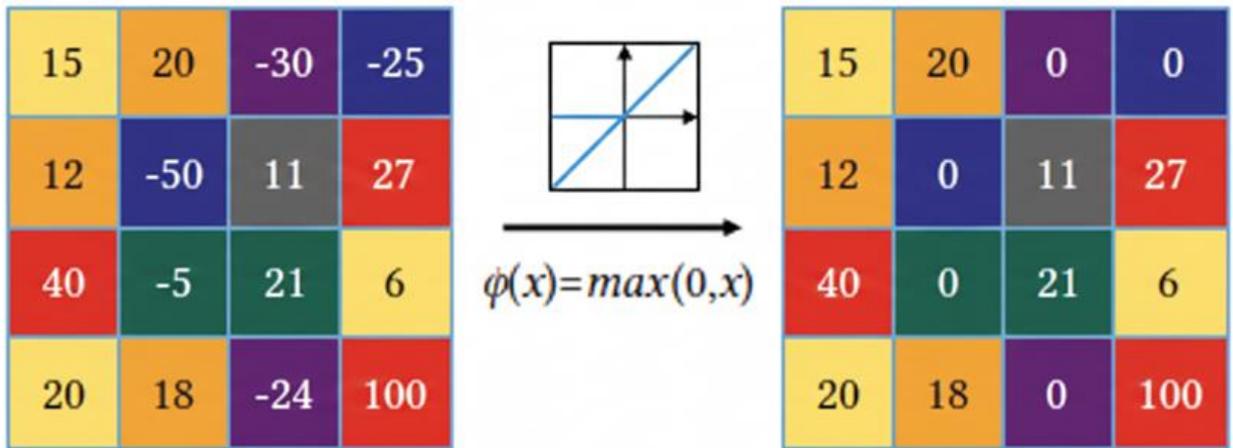


Рисунок 1.7 – Функція ReLU [32]

Шари підвибірки періодично вставляються після згорткових шарів для зменшення просторового розміру карти ознак та кількості параметрів, що підвищує ефективність обчислень і стійкість до невеликих зсувів вхідних даних. Зазвичай застосовують max pooling з ядром  $2 \times 2$  або  $3 \times 3$  і кроком 2 (рис. 1.8) [33].

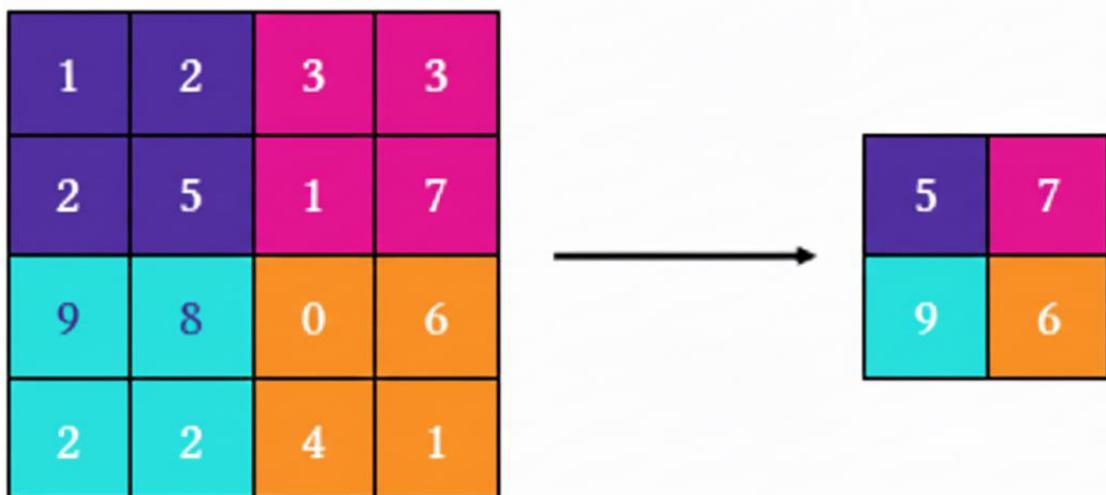


Рисунок 1.8 – Max pooling  $2 \times 2$ , stride=2

Повнозв'язні шари на завершальному етапі мережі поєднують високорівневі ознаки для задач класифікації, регресії або інших комп'ютерно-зорючих завдань.

До появи глибинного навчання виявлення транспортних засобів у одиночних аерокосмічних зображеннях здійснювалося шляхом виділення ручних ознак (handcrafted features) і подальшої класифікації або каскаду класифікаторів. Автори в роботі [34] поєднали геометричні характеристики форми, градації сірого та моменти  $xу$ . В роботі [35] автори застосували SIFT для виділення ключових точок із класифікацією за SVM. Використання ознак HOG з SVM [36] для зображень UAV. Автори в роботі [37] використали HOG та карти диспаратності. В [38] застосували ICF та AdaBoost у soft-cascade структурі. Додатково використовувалися 3D-інформація та стерео-сумісні алгоритми для підвищення точності [39].

Альтернативні підходи базуються на явних моделях транспортних засобів, які порівнюють форму автомобіля або групують ознаки для створення структур, подібних до моделі.

### **1.5.2. Глибоке навчання**

У останні роки традиційні методи виявлення транспортних засобів на аерокосмічних зображеннях значною мірою були замінені методами на основі глибинного навчання через обмежену представницьку здатність ручних ознак. Перші підходи на основі глибинного навчання застосовували згорткові нейронні мережі (CNN) як екстрактор ознак і частково як класифікатор у алгоритмах sliding window.

Автори [40] дослідили невелику CNN, яка складалася з трьох згорткових шарів з наступними max pooling та одним повнозв'язним шаром для екстракції ознак і класифікації у sliding window. Запропонована CNN продемонструвала кращі результати порівняно з традиційними методами на основі ручних ознак, такими як HOG+SVM та LBP+SVM, на супутникових зображеннях Google Earth. У наступній роботі [41] архітектуру було модифіковано шляхом поділу виходів останнього згорткового шару на блоки з різними receptive fields, що дозволило виділяти ознаки різних масштабів і підвищило точність виявлення.

Щоб зменшити обчислювальні витрати, ряд авторів застосовував методи генерації кандидатних регіонів (region proposal), які створюють менший набір кандидатних областей для класифікації. В роботі [42] автори застосували Binary Normed Gradients (BING) для генерації кандидатних регіонів замість exhaustive sliding window, що суттєво зменшило час обчислення без втрати точності. В роботі [43] використано R-CNN з методом Selective Search для побудови кандидатних регіонів. Для кожного регіону обчислювались CNN-ознаки через AlexNet, після чого проводилась класифікація за допомогою SVM. Авторами [44] покращено результати, запропонувавши нову функцію мети для тренування AlexNet, яка забезпечує інваріантність до обертання об'єктів у топ-в'ю через регуляризаційний термін. В роботі [45] проведено виявлення автомобілів у UAV-зображеннях у три етапи: генерація кандидатів через mean-shift – CNN (VGG16) – класифікація SVM. В роботі [46] використано ансамбль двох CNN (AlexNet та GoogleNet) для класифікації кандидатних регіонів і запропонували алгоритм unsupervised bounding box regression для підвищення точності локалізації. Автори [47] поєднали BING та CNN зі spatial pyramid pooling (SPPNet) і двоетапною каскадною класифікацією SVM, що дозволяє отримувати фіксовану довжину вектора ознак без деформації та обрізки вхідного регіону.

Для підвищення точності виявлення у аерокосмічних зображеннях були запропоновані інші підходи – використали дві послідовні CNN: перша генерує vehicle-like регіони, друга виконує класифікацію [48]. Автори [49] запропонували двошарову схему: семантична сегментація для виділення кандидатів, а потім класифікація типів транспортних засобів та фону. Обчислення CNN-ознак для кожного регіону окремо є обчислювально дорогим, тому були розроблені підходи, що обчислюють ознаки для всього зображення одночасно. При цьому не потребувалося адаптацій для специфіки аерокосмічних зображень завдяки низькій відстані між центрами пікселів на землі (GSD).

Для покращення продуктивності виявлення також застосовували [50] мультишарове поєднання ознак для RPN і класифікації у multi-scale Faster R-CNN. Сучасні дослідження також концентруються на орієнтованих bounding boxes [51 – 53] для точного обліку положення транспортних засобів на зображеннях та підвищення точності локалізації.

Алгоритм YOLO [54] належить до найбільш відомих підходів до виявлення об'єктів, що базуються на методах керованого навчання. Він характеризується відносно простою архітектурою та високими показниками точності. Обробка зображення в YOLO здійснюється за один обчислювальний прохід, що забезпечує можливість виявлення об'єктів у режимі реального часу, на відміну від традиційних методів, які потребують багатоетапної обробки [55]. Модифікація YOLOv6n [56] відрізняється спрощеною архітектурою порівняно з попередніми версіями за рахунок застосування мобільних блоків і полегшених шарів, що дозволяє зменшити обчислювальні витрати без суттєвого погіршення якості розпізнавання. Це зумовлює доцільність використання даної моделі в умовах обмежених апаратних ресурсів, зокрема на мобільних пристроях та у вбудованих системах.

## **1.6. Проблеми та виклики розпізнавання об'єктів транспорту на аерокосмічних зображеннях**

Розглянуті у попередніх підрозділах технічні аспекти формування аерокосмічних зображень, підходи до їх обробки, класифікації та розпізнавання, а також сучасні методи виявлення транспортних засобів, зокрема на основі згорткових нейронних мереж і глибокого навчання, демонструють значний потенціал застосування аерокосмічних даних для моніторингу земної поверхні. Водночас ефективність таких методів істотно залежить від умов зйомки, характеристик сенсорів, просторової роздільності зображень і властивостей самих об'єктів спостереження. Це

зумовлює наявність низки проблем і викликів, які необхідно враховувати при розробці та впровадженні алгоритмів розпізнавання об'єктів транспорту на аерокосмічних зображеннях.

Розпізнавання транспортних засобів на аерокосмічних зображеннях, отриманих із сенсорних платформ, таких як літаки, безпілотні літальні апарати або супутники, є складним завданням, що зумовлено значною відстанню між сенсором і поверхнею, умовами зйомки та різноманітністю сценаріїв через зміну часу доби та регіональні особливості. До основних викликів, що виникають при обробці аерокосмічних зображень, належать фактори, пов'язані з отриманням зображень, варіації об'єктів і сцени, а також практичні вимоги до алгоритмів виявлення.

Серед проблем, пов'язаних із отриманням зображень, можна виділити шум, який виникає як випадкове відхилення від реальних значень інтенсивності пікселів і зумовлений статистичними квантовими флуктуаціями, фізичними властивостями сенсора та обмеженням точності цифрового представлення інтенсивності пікселів. Розмиття зображень може бути спричинене не лише рухом об'єктів або камери, а й недостатньою експозицією при слабкому освітленні, що призводить до зниження різкості та контрастності. Освітлення суттєво впливає на якість зображення, наприклад, низьке освітлення потребує тривалішої експозиції, що збільшує шум та спричиняє розмиття, тоді як надмірне освітлення може спричинити насичення сенсора та втрату деталей. Низька просторова роздільність, яка виникає через великі відстані між камерою і землею, призводить до того, що транспортні засоби на зображенні займають лише кілька пікселів, що ускладнює їхню локалізацію та класифікацію, а варіації роздільності можуть спричинити помилки між класами, наприклад між «автомобілем» та «вантажівкою».

Варіації об'єктів і сцени також суттєво ускладнюють задачу. Висока внутрішньокласова різноманітність транспортних засобів за кольором, формою та розміром ускладнює навчання стійкого представлення ознак для

розрізнення транспортних засобів та фону, при цьому орієнтація об'єктів у перспективі зйомки додатково збільшує цю різноманітність. Низька міжкласова відмінність, що проявляється в подібності розмірів і форм різних категорій об'єктів, таких як автомобіль і вантажівка, може призводити до помилкової класифікації. Складний фон, зокрема в міських та промислових районах, підвищує ймовірність помилкових спрацьовувань через наявність об'єктів, подібних за розміром і формою до транспортних засобів. Часткове перекриття об'єктів, наприклад деревами або дорожніми знаками, змінює видимі ознаки і ускладнює класифікацію, а наявність тіней, що виникають при частковому або повному перекритті прямого світла, призводить до деформації форми об'єктів та зниження контрастності у тіньових областях.

Крім того, реальні застосування накладають додаткові вимоги до алгоритмів виявлення транспортних засобів. Алгоритми повинні бути універсальними, щоб забезпечити стійкість до варіацій даних, адже навчальні набори обмежені, а моделі повинні коректно працювати на нових, раніше невідомих зображеннях. Важливою є також можливість обробки у реальному часі, яка передбачає завершення обробки одного кадру до надходження наступного, що особливо критично при високій частоті кадрів. Нарешті, обмеження апаратного забезпечення, пов'язані з обмеженою вантажопідйомністю та ресурсами обчислювальної платформи на борту літальних апаратів, впливають на швидкість обробки та складність моделей глибинного навчання, оскільки ресурси GPU та оперативної пам'яті обмежені порівняно із серверними установками.

Таким чином, розпізнавання транспортних об'єктів на аерокосмічних зображеннях є комплексним завданням, що включає численні технічні, фізичні та обчислювальні виклики, які вимагають спеціалізованих методів обробки та адаптації алгоритмів для забезпечення високої точності, надійності та швидкості розпізнавання. Це дозволяє нам сформулювати постановку задачі.

## 1.7. Постановка задачі та теоретичний апарат дослідження

Незважаючи на розвиток сучасних методів обробки аерокосмічних зображень, завдання автоматизованого розпізнавання транспортних засобів на зображеннях високої просторової роздільної здатності залишається складним і потребує вдосконалення. Основними труднощами є зміни масштабу, орієнтації об'єктів, складні умови зйомки, висока щільність об'єктів та варіації освітлення і контрасту сцени. Існуючі класичні методи обробки зображень виявляють низьку точність і мають обмежену інваріантність до просторових трансформацій. При цьому сучасні нейромережеві підходи, що базуються на згорткових нейронних мережах і архітектурах типу YOLO та DeepLab, демонструють значно вищу ефективність, проте потребують адаптації до орієнтованого та багатомасштабного розпізнавання об'єктів у складних аерокосмічних сценах.

Аналіз наукових робіт показав, що більшість існуючих рішень фокусуються на виявленні об'єктів окремого типу або у локальних наборах даних, не забезпечують інваріантності до просторових трансформацій та не інтегрують багатокomпонентний підхід, який включає сегментацію контекстуальної інформації (дороги, смуги руху) для підвищення точності локалізації. Таким чином, виникає потреба у створенні інформаційної технології нейромережевого розпізнавання транспортних засобів довільного розташування, здатної ефективно працювати з аерокосмічними знімками високої просторової роздільної здатності та забезпечувати практичне застосування для моніторингу транспортної інфраструктури.

Виходячи з цього, метою дисертаційного дослідження є підвищення точності, інваріантності та автоматизації розпізнавання транспортних засобів на аерокосмічних зображеннях високої просторової роздільної здатності шляхом розробки багатокomпонентної нейромережевої інформаційної технології.

Для досягнення визначеної мети передбачаються такі основні завдання:

- провести аналіз сучасного стану методів виявлення та розпізнавання транспортних засобів на аерокосмічних зображеннях, зокрема класичних алгоритмів обробки зображень та нейромережових підходів, і визначити їх переваги та обмеження з точки зору точності, інваріантності до просторових трансформацій та обчислювальної складності;

- дослідити ефективність класичних методів виділення ознак (Canny, Sobel, Laplacian, Prewitt, Scharr) та сучасних нейромережових моделей виявлення об'єктів при розпізнаванні транспортних засобів на аерокосмічних зображеннях з використанням кількісних метрик Precision, Recall, F1 Score та mAP;

- розробити метод просторово-орієнтованого розпізнавання транспортних засобів, що базується на використанні глибоких нейронних мереж із механізмами автоматичної компенсації просторових трансформацій та багатомасштабного аналізу ознак;

- запропонувати методи попередньої обробки та аугментації аерокосмічних зображень, спрямовані на підвищення стійкості розпізнавання до змін масштабу, кута зйомки, освітлення та контрастності сцен;

- дослідити можливості застосування сучасних архітектур глибокого навчання для задач виявлення та сегментації транспортних засобів, зокрема моделей сімейства YOLO з підтримкою орієнтованих рамок (ОВВ) та сегментаційних архітектур типу DeepLab із глибокими магістральними мережами;

- розробити інформаційну технологію обробки та аналізу аерокосмічних зображень, яка інтегрує запропоновані методи розпізнавання транспортних засобів, модулі попередньої обробки, сегментації, виявлення та оцінки якості результатів;

- реалізувати програмний додаток, що забезпечує практичне застосування розробленої інформаційної технології, включаючи графічний інтерфейс користувача, засоби завантаження та обробки зображень, візуалізації результатів розпізнавання та аналізу показників ефективності;
- провести експериментальну перевірку розробленого методу та інформаційної технології на відкритих і спеціалізованих наборах даних аерокосмічних та БПЛА-знімків, а також здійснити порівняльний аналіз з існуючими алгоритмами виявлення об'єктів;
- оцінити обчислювальну ефективність і можливість застосування розробленого програмного додатку в режимі, наближеному до реального часу, що є важливим для задач моніторингу транспортної інфраструктури та безпеки.

## **1.8. Висновки до першого розділу**

1. Проаналізовано технічні аспекти формування аерокосмічних зображень, зокрема просторову та оптичну параметризацію, використання проєктивної геометрії для переходу від тривімірних об'єктів до двовимірних зображень, а також особливості застосування монокулярних RGB-камер на безпілотних літальних апаратах. Обґрунтовано доцільність використання цифрових систем обробки зображень у складі систем управління БПЛА для підвищення надійності, автономності та точності збору аерокосмічних даних, що є важливим для автоматизованого розпізнавання транспортних засобів на знімках.

2. У розділі проаналізовано основні етапи обробки аерокосмічних зображень, включаючи попередню підготовку, геоприв'язку, класифікацію та постобробку результатів. Розглянуто ручні, автоматичні та гібридні методи класифікації, а також підходи до розпізнавання об'єктів на основі контурного аналізу, шаблонів, ключових точок і згорткових нейронних мереж. Аналіз літературних джерел показав, що класичні методи

розпізнавання мають обмежену стійкість до змін масштабу, ракурсу та умов зйомки, що знижує їх ефективність при обробці складних аерокосмічних сцен.

3. Проведено порівняльний аналіз ефективності методів виявлення та розпізнавання транспортних засобів на аерокосмічних зображеннях. Аналіз класичних методів та підходів глибокого навчання на основі згорткових нейронних мереж, засвідчив перспективність CNN для автоматизованої обробки аерокосмічних зображень. На основі проведеного огляду сформульовано мету та основні завдання подальшого дослідження, спрямованого на розробку ефективних методів виявлення та нейромережевого розпізнавання транспортних засобів на аерокосмічних зображеннях.

Основні результати даного розділу опубліковано в роботі автора [25].

## РОЗДІЛ 2

### НЕЙРОМЕРЕЖЕВІ МЕТОДИ РОЗПІЗНАВАННЯ ТРАНСПОРТНИХ ОБ'ЄКТІВ НА АЕРОКОСМІЧНИХ ЗОБРАЖЕННЯХ

#### **2.1. Нейромережевий підхід до семантичної сегментації транспортних засобів на аерокосмічних зображеннях надвисокого просторового розрізнення**

Семантична сегментація є одним із підходів для автоматизованого аналізу аерокосмічних зображень, зокрема отриманих з БПЛА [57, 58]. Цей метод комп'ютерного зору полягає в класифікації кожного пікселя зображення відповідно до його класу [59]. Семантична сегментація дозволяє з високою точністю виділяти транспортні засоби на складних міських та дорожніх сценах [60].

Актуальність обумовлена потребою у розробках нових методів розпізнавання транспортних засобів для інтелектуальних транспортних систем, що включають в себе моніторинг дорожнього руху, управління логістичними процесами, а також покращення безпеки на дорогах. Використання зображень з БПЛА у поєднанні з архітектурою глибокого навчання дозволить підвищити точність і швидкість автоматизованого виявлення транспортних засобів в режимі реального часу.

Існуючі підходи до семантичної сегментації транспортних засобів можна поділити на традиційні методи та методи, засновані на глибокому навчанні. Традиційні методи сегментації транспортних засобів передбачають ручне виділення ознак, а потім використання методів машинного навчання, таких як SVM, AdaBoost та інші, для класифікації [61]. Ці методи мали істотні обмеження, оскільки потребували тривалої попередньої обробки для виділення ознак і налаштування порогових значень. Це також ускладнювало застосування їх до складних сцен, таких як аерофотозображення з малорозмірними об'єктами, оскільки вони здатні

виділяти лише поверхневі об'єкти і не справляються з більш складними випадками.

З появою глибокого навчання, зокрема завдяки впровадженню згорткових нейронних мереж та повністю згорткових мереж (FCN), ситуація змінилася, і методи семантичної сегментації значно удосконалилися. Авторами в роботі [62] запропоновано загальну мультимодальну систему глибокого навчання, яка використовує п'ять типів мереж злиття для інтеграції ознак гіперспектрального зображення (HSI), зображення LiDAR та зображення SAR для покращення ефективності сегментації зображень. Серія моделей Deeplab [63] заснована на збільшенні згорткових шарів, що вирішує проблему зниження роздільної здатності, що виникає на етапі максимальному об'єднанні шарів.

Серед підходів до виявлення об'єктів за допомогою глибокого навчання можна виділити дві основні категорії: двоетапні та одноетапні алгоритми. Двоетапні алгоритми, такі як Fast R-CNN [64], спочатку виділяють області інтересу, а потім виконується локалізація і класифікація об'єктів. Наприклад, метод, запропонований в роботі [65], показав задовільні результати для виявлення літаючих об'єктів з використанням Faster R-CNN і VGG-16, досягнувши середньої точності 66% (mAP). Однак ці методи мають значну обчислювальну складність і можуть бути менш ефективними при виявленні малих об'єктів. В роботі [66], використовують паралельні мережі RPN (Region Proposal Networks) для покращення виявлення щільних областей на аерофотозображеннях. Запропонований авторами [67] метод, заснований на CNN, використовує Xception для класифікації та U-Net з ResNet18 як кодер для точної сегментації кораблів на оптичних супутникових зображеннях, досягаючи точності понад 84%. Але його застосування до сегментації транспортних засобів на аерокосмічних зображеннях надвисокого розрізнення має низку обмежень: відмінності в характеристиках об'єктів, різні просторові особливості зображень, відсутність спеціалізованого донавчання та обмеження у

виділенні дрібних структурних об'єктів при використанні U-Net. У роботі [68] було запропоновано метод виявлення транспортних засобів на аерофотозображеннях, який використовує згорткову нейронну мережу з подвійними фокусними втратами (MFL CNN). Автори підкреслюють складність завдання виявлення транспортних засобів, зокрема через їх малий розмір та складний фон зображень. У роботі продемонстровано переваги запропонованого підходу порівняно з базовими моделями, що підтверджується результатами на наборах даних EAGLE та XWHEEL.

Однак складність моделі і двоетапний процес виявлення не відповідають вимогам до реального часу. Водночас, одноетапні алгоритми, зокрема YOLO [67–71], демонструють значні переваги в швидкості і точності порівняно з двоетапними методами, але також мають певні обмеження, зокрема у вирішенні проблеми хибних спрацювань і складних фонових умов.

Алгоритми на основі глибокого навчання значно покращили точність і ефективність виявлення об'єктів, включаючи транспортні засоби. Вони можуть автоматично навчатися на великих наборах даних і не залежать від ручного виділення ознак. Проте залишаються невирішеними проблеми: велика кількість помилкових спрацювань при виявленні об'єктів виникає через те, що деякі нетранспортні об'єкти мають схожий зовнішній вигляд з транспортними засобами; існуючі детектори транспортних засобів на основі CNN завжди мають два виходи: координати обмежувальної рамки та ймовірність того, що об'єкт, який знаходиться в межах цієї рамки, є транспортним засобом.

Нейромережевий підхід семантичної сегментації транспортних засобів з використанням зображень надвисокого просторового розрізнення представлено на рис.2.1, базується на архітектурі DeepLab + ResNet з використанням багаторівневого вилучення ознак.

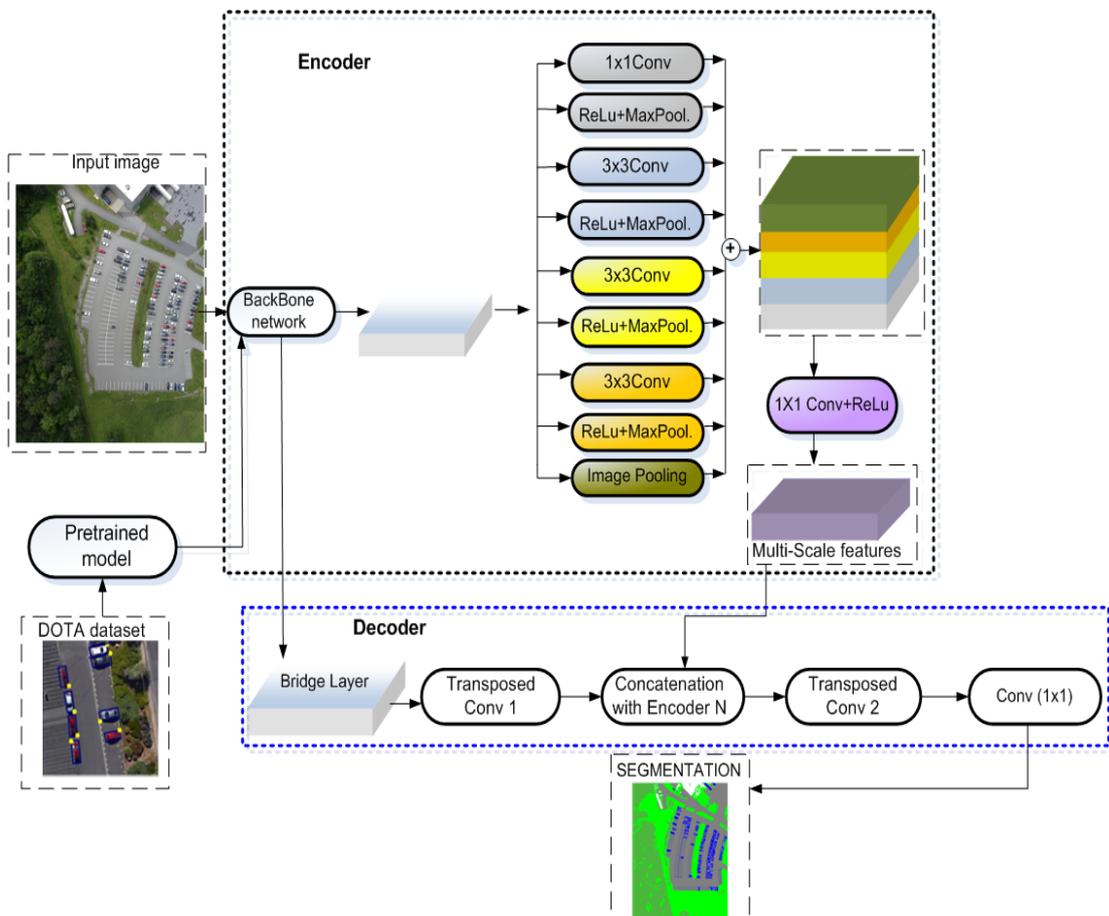


Рисунок 2.1 – Структурна схема неймережевого підходу семантичної сегментації транспортних засобів

На першому етапі реалізації методу виконується завантаження аерокосмічного зображення. Далі, вхідне зображення обробляється основною мережею (Backbone network), яка була ініціалізована вагами, отриманими під час навчання на наборі даних DOTA (Dataset for Object Detection in Aerial Images) [72], а потім донавчена на власних спеціалізованих наборах даних для семантичної сегментації із класами: автомобілі (індивідуальні транспортні засоби, паркінги, дороги); зелена зона (рослинність, газони, парки); будівлі (житлова та промислова забудова); дороги (основні магістралі, другорядні вулиці, перехрестя). На цьому етапі виконано попередню обробку [73]: нормалізацію, масштабування та розмітку об'єктів, щоб забезпечити коректне навчання неймережі:

$$I_{\text{norm}} = \frac{I - \mu(I)}{\sigma(I)}, \quad (2.1)$$

де  $\mu$  і  $\sigma$  представляють середнє значення і стандартне відхилення зображення  $I$ .

Глибинні нейронні мережі, що складаються з великої кількості послідовно з'єднаних шарів, схильні до проблеми зникаючого градієнта. Ця проблема виникає в процесі зворотного поширення помилки, коли градієнти, що використовуються для оновлення ваг мережі, зменшуються експоненціально з глибиною мережі, наближаючись до нуля. Внаслідок цього, шари, розташовані ближче до входу мережі, практично не навчаються, що обмежує здатність мережі вивчати складні залежності. Для вирішення цієї проблеми в запропонованій методиці застосовуються в Backbone network використовуємо мережу ResNet-101 замість Xception як магістральну мережу, що дозволяє зберегти градієнт при збільшенні глибини мережі та ефективно вилучати ознаки на різних масштабах. Це досягається шляхом додавання вхідних даних до виходу одного або кількох шарів, що дозволяє градієнтам поширюватися на попередні шари. Остаточний навчальний блок (A residual building block) може бути визначений за формулою (2.2) [73]:

$$y = F(x, \{W_i\}) + x, \quad (2.2)$$

де  $y$  – вихід блоку,  $F(x, \{W_i\})$  – функція відображення, що представляє собою послідовність шарів з параметрами  $\{W_i\}$ ,  $x$  – вхідні дані.

Кодер складається з послідовності згорткових шарів  $1 \times 1$  згортки для зменшення розмірності ознак без втрати інформації та  $3 \times 3$  згортки з функцією активації ReLU, доповнені операціями MaxPooling для ієрархічного агрегування просторово-контекстної інформації. Для забезпечення багатомасштабної обробки застосовується механізм об'єднання багаторівневих ознак (Multi-Scale Features), який включає шари глобального згорткового згладжування (Image Pooling) та подальшу трансформацію через  $1 \times 1$  згортку з активацією ReLU. Це дозволяє нейромережі одночасно аналізувати як локальні, так і глобальні контексти, що покращує точність сегментації об'єктів різного розміру, зокрема

транспортних засобів. Однак, при багатокласовій сегментації виникає проблема дисбалансу класів. Це коли кількість зразків одного класу значно перевищує кількість зразків інших класів, що може призвести до погіршення розпізнавання менш представлених класів. У найгірших випадках модель може повністю ігнорувати класи з низькою представленістю, якщо кількість їхніх навчальних зразків є недостатньою. Для вирішення цієї проблеми в нашому методі використовуються кастомізована функція втрат на основі коефіцієнта Dice:

$$D(p, q) = \frac{2 \sum_i^N p_i q_i + \epsilon}{\sum_i^N p_i^2 + \sum_i^N q_i^2 + \epsilon}, \quad (2.3)$$

де  $\epsilon$  – невелике додатне число, що додається, щоб уникнути ділення на нуль у разі відсутності перетину між передбаченим та реальним сегментами;  $p$  – передбачена сегментація;  $q$  – реальна розмітка.

Для формулювання функції втрат використано її доповнення:

$$L_{Dice} = 1 - D(p, q). \quad (2.4)$$

Декодер відновлює просторову роздільну здатність сегментаційного зображення шляхом послідовного використання транспонованих згорткових операцій (Transposed Convolution), що дозволяють поступово відновлювати структуру об'єктів. Крім того, застосовується механізм конкатенації (skip connections) з відповідними шарами енкодера для збереження високорівневої інформації та покращення деталізації сегментації. Фінальний згортковий шар ( $1 \times 1$  Convolution) виводить зображення до потрібної кількості каналів для кожного класу сегментації.

Запропонована архітектура DeepLab + ResNet забезпечує ефективне вилучення багатомасштабних ознак, що сприяє підвищенню точності сегментації завдяки збереженню просторової та семантичної інформації [74].

## **2.2. Нейромережеве розпізнавання повітряних транспортних засобів на аерокосмічних зображеннях з урахуванням просторово-геометричних ознак**

Розглянуті у попередньому підрозділі методи семантичної сегментації показали ефективність нейромережових підходів для виділення транспортних засобів на зображеннях високого просторового розрізнення. Проте виділення та класифікація повітряних транспортних засобів, зокрема літаків, на аерокосмічних зображеннях залишаються складним завданням через їхній малий розмір на зображенні, мінливий фон, нечіткі краї об'єктів та коливання умов освітлення. Це створює потребу у додаткових підходах, що враховують просторово-геометричні ознаки об'єктів, а також спеціальні методи попередньої обробки для підвищення точності моделей глибокого навчання. Наступний підрозділ присвячений саме таким методам та підходам до нейромережевого розпізнавання повітряних транспортних засобів на аерокосмічних зображеннях.

Техніки глибокого навчання, зокрема згорткові нейронні мережі, за останні роки продемонстрували високу ефективність у завданнях розпізнавання та класифікації об'єктів [75 – 80].

Для завдань виявлення невеликих об'єктів на високороздільних зображеннях дистанційного зондування автори в роботі [81] запропонували структуру з деконволюційним модулем для уточнення карт ознак та відновлення просторової інформації, втраченої під час об'єднання шарів. Механізм уваги «стиснення та збудження» спрямований на виділення інформативних ознак, що є важливим для точного виявлення невеликих літаків на складному фоні. RefineContourNet – багатошляхова CNN на основі ResNet, спеціально розроблена для уточнення меж об'єктів літаків після їх первинного виявлення моделями глибокого навчання [82]. В іншому дослідженні [83] було запропоновано алгоритмічну техніку сегментації літаків на основі нечіткої логіки, яка досягає точності 92,5%.

Вдосконалені моделі YOLO також демонструють значні покращення. Наприклад, алгоритм YOLO-extract [84] забезпечує швидшу збіжність, зменшує обсяг обчислень та знижує кількість параметрів на 10,526 млн, одночасно збільшуючи *mAP* на 8,1% та прискорюючи швидкість виявлення літаків у 3 рази порівняно з YOLOv5. Модель YOLO-class [85] є покращеною версією YOLO-extract із модифікацією архітектури мережі, що підвищує точність виявлення з 0,608 до 0,704 та FPS з 36,16 до 39,598. Для обробки даних SAR архітектура YOLO може бути адаптована та вдосконалена [86]. Додавання модуля Attention-Efficient Layer Aggregation Network-Head дозволяє пріоритизувати ключові ознаки, що збільшує *mAP50* на 2,1% та *mAP50-95* на 1,9%.

Таким чином, сучасні дослідження зазвичай зосереджені на одному з двох підходів: виявлення об'єктів або контурів. Хоча ці методи ефективні окремо, їхня інтеграція залишається недостатньо вивченою, що відкриває перспективи для подальших розробок. Тому, запропоновано новий підхід до виявлення повітряного транспорту на аерокосмічних зображеннях складається з чотирьох етапів, як показано на рис. 2.2.

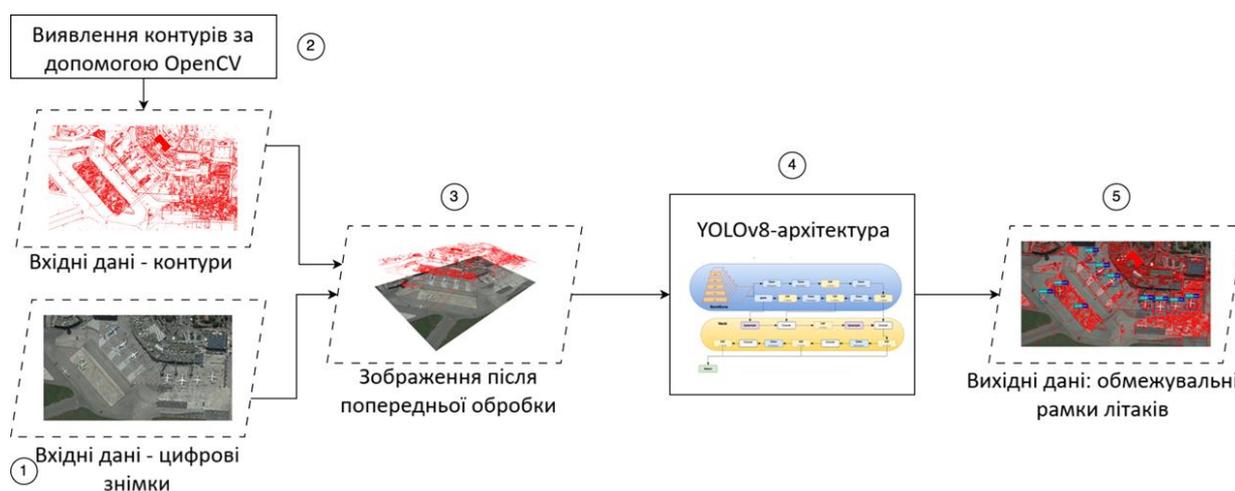


Рисунок 2.2 – Структурна схема запропонованого підходу на основі YOLOv8m-архітектури

Приклад алгоритму Сузукі показаний на рис. 2.3, в нашому випадку *ob* означає зовнішню межу, *hb* позначає межу отвору, а батьківська межа вказує, що зовнішній шар виступає батьківським для внутрішнього шару.

Обведені пікселі в (a–г) відповідають описам меж для кожного посилення праворуч.

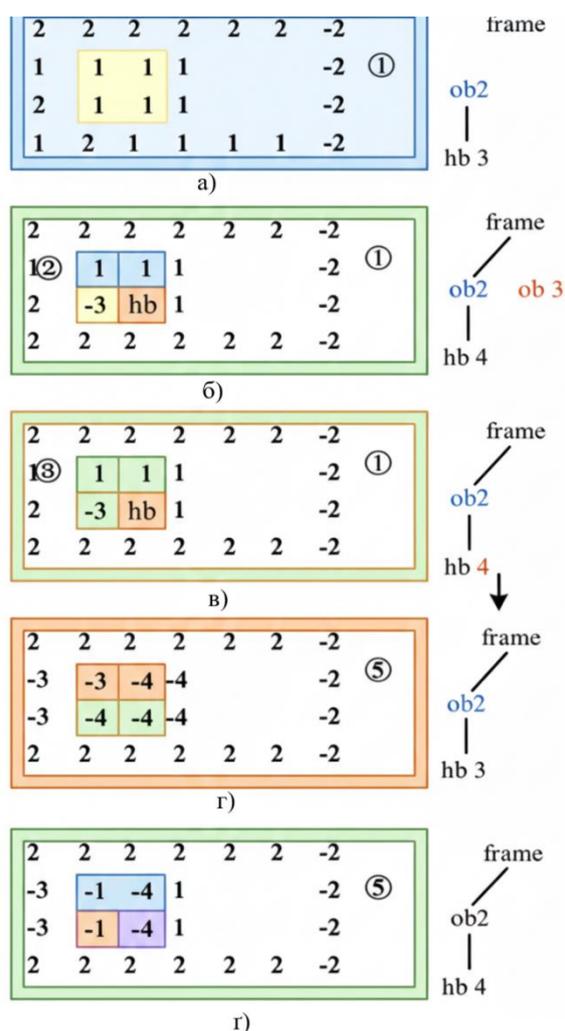


Рисунок 2.3 – Приклад діаграми алгоритму відстеження меж [88]

Третій етап – попередня обробка даних шляхом поєднання контурів з оригінальним зображенням. На четвертому етапі запропоновано навчити модель виявлення об'єктів YOLOv8m. Архітектура моделі показана на рис. 2.4.

Вибір моделі обумовлений сприятливим балансом між складністю моделі (295 шарів і 25,9 мільйона параметрів) і заявленою продуктивністю.

Ця характеристика дозволила збалансувати точність виявлення об'єктів і обчислювальну ефективність.

Архітектура YOLOv8 дотримується принципу модульного дизайну і може бути розділена на два основних компоненти: основу та голову.

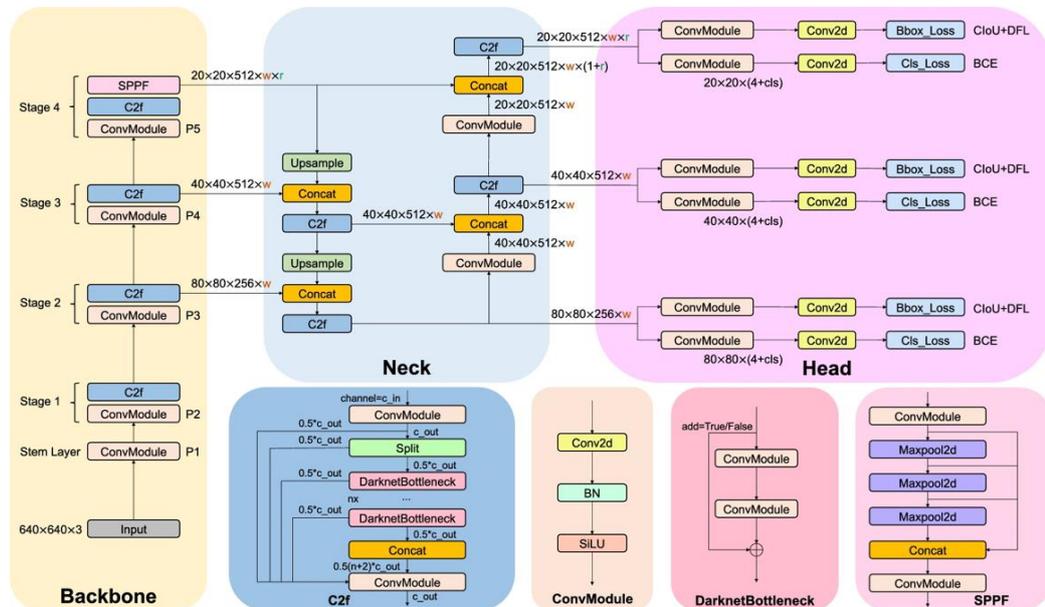


Рисунок 2.4 – Структурна схема YOLOv8-архітектури [89]

Мережа служить основою для вилучення ознак. Вона складається з 53 конволюційних шарів, посилені міжетапними частковими з'єднаннями. YOLOv8 пропонує гнучкість завдяки використанню різноманітних варіантів основ, включаючи CSPDarknet53 та EfficientDet. Цей вибір дозволяє досягти компромісу між здатністю витягувати інформативні ознаки та пов'язаною з цим обчислювальною складністю.

Головна частина генерує прогнози на основі ознак, витягнутих базовою мережею та архітектурою шиї. Вона прогнозує координати обмежувальної рамки, оцінки об'єктності та ймовірності класу для кожної анкерної рамки, пов'язаної з коміркою сітки. Архітектура використовує анкерні рамки для ефективного прогнозування об'єктів різних форм і розмірів.

Перший етап полягає у завантаженні набору даних HRPlanesv2 [87], другої ітерації набору даних HRPlanes. Він містить 2120 зображень Google Earth з дуже високою роздільною здатністю, на яких позначено загалом 14335 літаків. Набір даних розділено на три частини: 70% для навчання, 20% для валідації та 10% для тестування.

Другий етап – це вилучення контурів за допомогою виявлення контурів OpenCV. Техніка, заснована на алгоритмі відстеження меж [86], для ідентифікації та аналізу пов'язаних областей у зображенні, які мають подібні характеристики інтенсивності або кольору. Вона аналізує бінарне зображення для ідентифікації з'єднаних компонентів, де всі пікселі мають однакову інтенсивність (у нашому випадку білий колір) і є сусідніми один з одним. Ці з'єднані компоненти представляють межі об'єкта або контури. Ці області часто відповідають чітким формам або об'єктам, присутнім на зображенні.

Чотири межі зображення називаються його рамкою. Зображення з шириною  $w$  і висотою  $h$  може бути представлене як матриця порядку  $h \times w$ , що складається з окремих пікселів. Рядки 1 і  $h$  та стовпці 1 і  $w$  утворюють рамку зображення. Піксель із сірим значенням нуль визначається як нульовий піксель, а піксель із сірим значенням один називається одиничним пікселем. У цьому алгоритмі рамки бінаризованого зображення розглядаються як нульові пікселі. Якщо рамка вхідного зображення містить будь-який одиничний піксель, вони перетворюються на нульові пікселі [89].

### **2.3. Нейромережеве розпізнавання автомобілів на аерокосмічних зображеннях з використанням модифікованої архітектури YOLO-OBV**

Після розгляду методів нейромережевого розпізнавання повітряних транспортних засобів стає очевидним, що аналогічні виклики виникають і при виявленні автомобілів на аерофотозображеннях. Незважаючи на відмінності у розмірі та морфології об'єктів, автомобілі також

характеризуються малими розмірами на зображеннях, щільним розташуванням у транспортних вузлах, схожістю колірно-текстурних ознак з оточенням та різноманітністю форм і орієнтацій. Ці фактори значно ускладнюють автоматизовану ідентифікацію та класифікацію транспортних засобів, що обумовлює потребу в спеціалізованих нейромережових підходах, здатних забезпечити високу точність локалізації та класифікації в умовах складного фону та варіативності масштабу.

Значний вплив на якість аналізу даних мають атмосферні явища, такі як хмарність, туман, тіньові ділянки та неоднорідне освітлення, які можуть спотворювати візуальні характеристики об'єктів і знижувати достовірність автоматизованої ідентифікації [91, 92]. Практичний інтерес серед об'єктів аерофотознімків становлять саме автомобілі, оскільки їхнє виявлення має важливе значення для систем інтелектуального моніторингу, моделювання транспортних потоків, урбаністичного планування та підвищення ефективності безпекових технологій.

Одним із найбільш вивчених напрямів є багатомасштабне виявлення, що має особливе значення для автомобілів, розміри яких значною мірою залежать від висоти зйомки, фокусної відстані камери та параметрів зображення. Для подолання масштабної варіативності застосовуються архітектури з багаторівневою побудовою ознак. Так, мережа пірамід ознак (Feature Pyramid Network, FPN) [93] реалізує ієрархічну структуру побудови ознак із горизонтальними з'єднаннями, дозволяючи формувати карти ознак на різних рівнях ієрархії та покращуючи точність виявлення дрібних об'єктів. Разом із тим, базовий підхід FPN має обмеження, пов'язані з контекстним узгодженням ознак та апсемплінгом, що може призводити до втрати просторових деталей [94]. Для подолання цих обмежень запропоновано Adaptive Feature Pyramid Network, яка більш точно інтегрує просторову інформацію з різних рівнів ієрархії ознак.

Для покращення багатомасштабного виявлення автомобілів вдосконалюють також одноетапні детектори. У роботі [95] представлено

Feature Fusion SSD (FSSD) – модифікацію SSD, що забезпечує більш ефективне об'єднання ознак з різних рівнів та масштабів, формуючи єдине представлення для детекційних блоків і підвищуючи точність виявлення дрібних об'єктів без значного збільшення обчислювальних витрат. Принципово новий підхід RepPoints [96] описує об'єкт як набір навчальних точок, що одночасно слугують для локалізації та семантичного опису, відмовляючись від anchor boxes. Для підвищення точності локалізації запропоновано динамічне призначення міток на основі прогнозованих IoU [97]. Метод SCRDet++ [98] використовує шумозаглушення на рівні екземпляра та коефіцієнт постійного IoU для врахування орієнтації об'єктів. Для інкрементного навчання застосовується Constraint-based Fine-tuning (CFA) [99], що підвищує узгодженість знань між базовими та новими класами. Методи глибокого навчання [100], зокрема CNN, значно підвищили ефективність виявлення об'єктів [101, 102]. Двохетапні архітектури R-FCN, Faster та MS-Faster R-CNN [103, 104] демонструють високу точність завдяки попередній генерації регіонів-пропозицій. Одноетапні моделі, зокрема YOLO [67, 105–107], забезпечують високу швидкість обробки, що важливо для реального часу, проте можуть поступатися точністю при дрібних або скупчених об'єктах.

Незважаючи на розвиток CNN, залишаються обмеження, пов'язані з горизонтальними обмежувальними рамками (ОВВ), що не завжди дозволяють точно описати орієнтовані об'єкти на аерозображеннях. Використання орієнтованих обмежувальних рамок (ОВВ) [97, 70] дозволяє враховувати орієнтацію та точніше розмежовувати щільно розташовані екземпляри, формуючи задачу виявлення орієнтованих об'єктів (oriented object detection) та вимагаючи модифікації детекторів і функцій втрат [108]. Додатково обмеження на розмір вхідного зображення моделей [109, 110] потребують масштабування або фрагментації знімків із великими розмірами, характерними для ITCVD або DOTA.

В роботі запропоновано інформаційну технологію, представлену на рисунку 2.5, що функціонально структурована на п'ять ключових етапів. Кожен із цих етапів є складовою загальної послідовності обробки даних, процесу побудови, валідації моделі, що забезпечує систематичний підхід до вирішення поставленої задачі виявлення автомобілів на аерофотозображеннях.

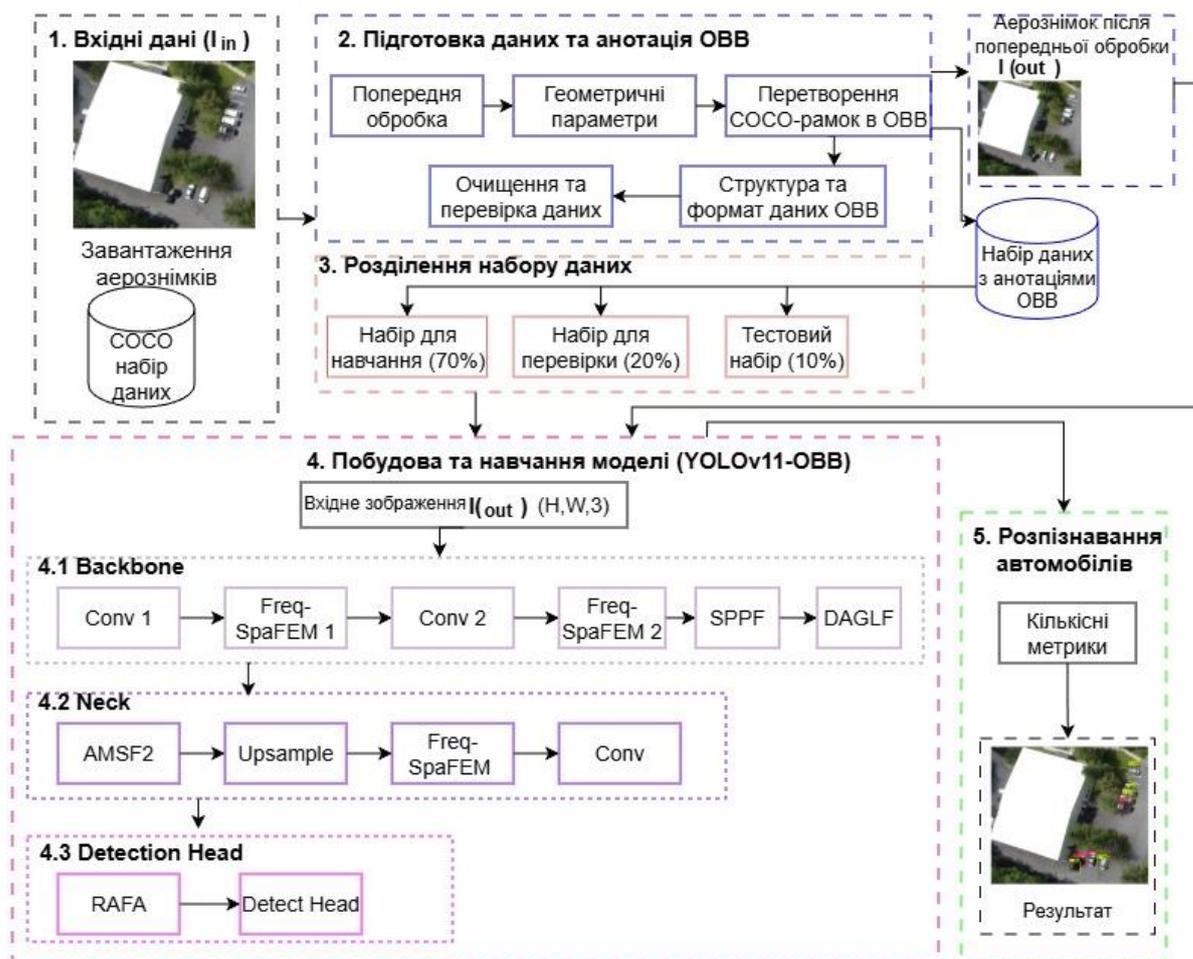


Рисунок 2.5 – Структурна схема інформаційної технології розпізнавання автомобілів

На першому етапі реалізації технології виконується завантаження вхідного аерофотознімка, а також використовується COCO-датасет [112] як базовий набір даних із розміткою транспортних засобів. На вхід інформаційної технології подається аерофотознімок

$$I \in \mathbb{R}^{H \times W \times 3}, \quad (2.5)$$

де  $H$  – висота зображення,  $W$  – ширина зображення, 3 – кількість каналів (RGB).

На другому етапі відбувається підготовка даних та ОБВ-анотація. Спочатку відбувається попередня обробка завантажених аерофотознімків [111]: нормалізація, зміна розміру (Resizing), адаптивне вирівнювання гістограми з обмеженням контрастності (CLAHE), та корекція кольору. Для забезпечення стабільності навчання нейронної мережі первинні данні проходять етап нормалізації:

$$I'_{(x,y,c)} = \frac{I_{(x,y,c)} - \mu_c}{\sigma_c}, \quad (2.6)$$

де  $I_{(x,y,c)}$  – інтенсивність пікселя у координаті  $(x,y)$  у каналі  $c$ ,  $\mu_c$ ,  $\sigma_c$  – середнє значення та стандартне відхилення піксельних значень по каналу,  $I'$  – нормалізоване зображення.

З метою уніфікації вхідних даних зображення масштабуються до фіксованого розміру (640×640) пікселів. Щоб уникнути спотворення співвідношення сторін, використовується padding (доповнення), яке зберігає початкову геометрію сцени. Далі використовується метод адаптивного вирівнювання гістограми з обмеженням контрастності (CLAHE) для підвищення локального контрасту зображення, що є особливо важливим для аерофотознімків, де об'єкти можуть мати низький контраст з фоном через умови освітлення або тіні. На відміну від глобального вирівнювання гістограми, CLAHE обробляє зображення по невеликих, непересічних областях (тайлах), а потім використовує білінійну інтерполяцію для згладжування меж між тайлами. Для кожного тайла, нехай  $p(k)$  – це нормалізована гістограма інтенсивності пікселів, де  $k$  – рівень інтенсивності. Трансформація інтенсивності  $I_{in}$  до  $I_{out}$  здійснюється за функцією розподілу ймовірностей (CDF)  $C(k)$ :

$$I_{out} = C(I_{in}) = \sum_{j=0}^{I_{in}} p(j). \quad (2.7)$$

Після попередньої обробки виконується вилучення геометричних параметрів. Для кожного об'єкта розраховується орієнтована обмежувальна рамка:

$$B_{OVB} = (x_c, y_c, w, h, \theta), \quad (2.8)$$

де  $(x_c, y_c)$  – координати центру рамки;  $w, h$  – ширина і висота рамки,  $\theta$  – кут обертання рамки відносно осі  $x$ .

Це дає змогу перейти від НВВ до ОВВ за допомогою трансформаційної матриці:

$$\begin{bmatrix} x'_i \\ y'_i \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_i - x_c \\ y_i - y_c \end{bmatrix} + \begin{bmatrix} x_c \\ y_c \end{bmatrix}. \quad (2.9)$$

Потім виконується фільтрація об'єктів за площею та перекриттям, очищення некоректних міток, збереження у форматі, сумісному з обробкою орієнтованих рамок. Нарешті, відбувається структурування та форматування всіх ОВВ-даних, що дозволяє створити єдиний ОВВ-анотований набір даних, готовий для навчання моделі. Результатами цього етапу є попередньо оброблений аерофотознімок та структурований ОВВ-анотований набір даних.

На третьому етапі відбувається розподіл даних. З метою забезпечення об'єктивного оцінювання ефективності моделі та запобігання упередженості результатів, орієнтовано-анотований (ОВВ) набір даних, що включає 2758 аерофотознімків, було логічно розподілено на три взаємозалежні підмножини. Навчальна вибірка охоплює 70% загального обсягу даних, що становить 1931 зображення, і використовується безпосередньо для оптимізації ваг нейронної мережі під час навчання. Валідаційна підмножина містить 552 зображення (20%) і призначена для контролю проміжних результатів, а також для налаштування гіперпараметрів моделі з метою зниження ризику перенавчання. Тестовий набір, сформований з решти 10%

(275 зображень), використовується виключно на завершальному етапі для незалежного оцінювання генералізаційної здатності моделі, що дозволяє об'єктивно визначити її ефективність на раніше невідомих даних.

Четвертим етапом є побудова моделі YOLOv11-OBV. Цей етап є основним для реалізації запропонованої інформаційної технології. Модель приймає на вхід зображення Iout. Архітектура моделі умовно поділяється на три функціональні компоненти: бекбоун, Neck та детекційну голову. Бекбоун виконує вилучення ієрархічних ознак із зображення після другого етапу, використовуючи послідовність згорткових шарів (Conv1, Conv2), інтегрованих із модулями частотно-просторового вилучення ознак Freq-SpaFEM1 та Freq-SpaFEM2 (Frequency-Spatial Feature Extraction Module). Freq-SpaFEM забезпечує одночасне вилучення локальних просторових та глобальних частотних характеристик, що є важливим при обробці даних зі складними текстурами й дрібними об'єктами, характерними для аерофотознімків. Завершують бекбоун шари SPPF (Spatial Pyramid Pooling Fast) [113] та DAGLF (Dual Attention Global Local Fusion). SPPF дозволяє агрегувати контекстні ознаки різних масштабів, а DAGLF дозволяє одночасно фокусуватися на локальних деталях і глобальному контексті сцени, покращуючи семантичну виразність ознак і зменшуючи вплив фонових завад.

Компонент Neck виступає проміжною ланкою між блоком вилучення ознак (Backbone) та детекційною головою, забезпечуючи злиття ознак з різних рівнів глибини мережі. Він включає модуль AMSF2 (Adaptive Multi-Scale Feature Fusion), що дає змогу покращити узагальнення для об'єктів різних розмірів. Шари Upsample (для узгодження з просторовими розмірами вищих рівнів мережі), Freq-SpaFEM (для посилення просторово-частотної узгодженості карт ознак) та Conv (для подальшого уточнення ознак перед їх передачею до детекційного блоку).

Блок Head є завершальним компонентом архітектури, який відповідає за безпосереднє формування прогнозів щодо виявлення та класифікації

об'єктів. Він містить модуль RAFA (Region-Aware Feature Aggregation), що дозволяє моделі краще відрізнити об'єкти від фону та враховувати контекст навколо кожного об'єкта, та Detect Head, що генерує остаточні прогнози ОБВ, включаючи координати, розміри, кути нахилу та формує вихідні орієнтовані обмежувальні рамки.

Завершальний етап інформаційної технології включає практичне застосування навченої моделі для виявлення автомобілів та оцінку її продуктивності. Проводиться комплексний аналіз результатів за допомогою стандартних метрик оцінки [114] якості виявлення об'єктів (Precision, Recall, mAP, ОБВ IoU) для підтвердження ефективності запропонованої технології. Крім того, здійснюється візуалізація результатів виявлення, що дозволяє наочно продемонструвати здатність моделі точно локалізувати та класифікувати автомобілі на аерофотозображеннях [115].

#### **2.4. Метод просторово-орієнтованого розпізнавання транспортних засобів довільного розташування**

Попередньо розглянуті методи нейромережевого виявлення транспортних засобів на аерофотозображеннях, включно з багатомасштабними та адаптованими одноетапними детекторами, забезпечують високу точність локалізації та класифікації автомобілів. Проте їхня ефективність обмежена у випадках, коли об'єкти мають довільну орієнтацію, нестандартні пропорції або складну геометрію. Традиційні обмежувальні рамки (ОБВ), навіть у модифікованих архітектурах YOLO-ОБВ, частково враховують орієнтацію, але не завжди дозволяють повністю інтегрувати просторово-геометричну інформацію об'єктів у процесі розпізнавання.

Формування аерокосмічних зображень визначається складною сукупністю геометричних факторів, зокрема просторовим положенням та орієнтацією носія знімальної апаратури, параметрами камери, а також

перспективними спотвореннями, що виникають у процесі проєкції тривимірної сцени на двовимірну площину зображення. У результаті об'єкти на аерокосмічних зображеннях, зокрема транспортні засоби, можуть мати довільну орієнтацію, змінні масштаби та нестандартні геометричні форми, що істотно ускладнює їх подальше автоматичне виявлення та локалізацію.

Зважаючи на ці особливості, виникає необхідність переходу від класичних підходів до методів, здатних явно враховувати просторову орієнтацію об'єктів у сцені. Це визначає актуальність розробки спеціалізованих методів просторово-орієнтованого розпізнавання транспортних засобів довільного розташування, які є предметом наступного підрозділу.

Описані в підрозділі 1.2 технічні аспекти формування аерокосмічних зображень обґрунтовують необхідність переходу від класичних підходів до просторово-орієнтованого опису об'єктів до методів, здатних явно враховувати просторову орієнтацію цільових об'єктів у сцені. Це, у свою чергу, зумовлює актуальність розробки спеціалізованих методів просторово-орієнтованого розпізнавання транспортних засобів довільного розташування.

Точне виявлення транспортних засобів на аерокосмічних зображеннях є складним завданням з огляду на малі геометричні розміри об'єктів, обмежену кількість виразних локальних ознак, а також складний і структурно неоднорідний наземний фон [116]. Додаткову складність становить довільна просторова орієнтація транспортних засобів, зумовлена як умовами зйомки, так і реальним плануванням транспортної інфраструктури, що було детально розглянуто у підрозділі 1.2. Попередні методи, що враховували геометричні спотворення через додавання відповідних варіацій у тренувальні набори, часто призводили до значних обчислювальних витрат та потенційного зниження продуктивності мережі.

Методи глибокого навчання досягли значних успіхів у завданнях обробки зображень та комп'ютерного зору, зокрема у відновленні зображень, розпізнаванні об'єктів та класифікації. Однак у разі неякісних зображень, геометричних спотворень або атмосферних артефактів точність моделей на основі глибокого навчання може істотно знижуватися. Наприклад, при розпізнаванні об'єктів на зображеннях із камер дальнього бачення або аерокосмічних платформ спотворення структури об'єктів через турбулентність повітря чи перспективні ефекти можуть призводити до помилкової класифікації [117].

Одним з підходів для вирішення цієї проблеми є додавання штучно модифікованих зображень до тренувального набору (наприклад, обертання, масштабування, додавання шуму) для покращення навчання моделі та підвищення її здатності коректно виявляти об'єкти у різноманітних умовах, однак це значно збільшує обчислювальні витрати та може погіршити продуктивність за рахунок додаткової варіації у даних. Альтернативним підходом є використання моделей для опису геометричних трансформацій зображень, проте підбір точної моделі для різних типів трансформацій є складним і трудомістким завданням. Тому актуальним є застосування нейромережевих підходів для розпізнавання транспортних засобів, здатних автоматично виконувати просторову нормалізацію ознак та формувати стійкі внутрішні подання сцени без необхідності явного задання моделей геометричних спотворень.

Існуючі підходи до виявлення транспортних засобів на аерокосмічних зображеннях можна розділити на традиційні методи та методи, засновані на глибокому навчанні. Традиційні алгоритми передбачають ручне витягування ознак [118], після чого виконують класифікацію за допомогою методів машинного навчання, таких як SVM або AdaBoost. Такий підхід є трудомістким і обмеженим, оскільки дозволяє використовувати лише поверхневі ознаки, що значно знижує ефективність при обробці сцен із малими цілями та складним фоном, характерними для аерофотозйомки.

Таким чином, аналіз сучасних методів показує, що існує компроміс між точністю та швидкістю виявлення, а також існують труднощі при роботі з малими об'єктами, складним фоном, варіативними умовами зйомки та обмеженими наборами даних. Це обумовлює необхідність розробки методів, які поєднують просторову нормалізацію ознак та багатомасштабне формування інформативних ознак.

Запропонований метод просторово-орієнтованого розпізнавання транспортних засобів на аерокосмічних зображеннях базується на багатокомпонентній нейромережевій архітектурі, що поєднує просторове узгодження ознак, багатомасштабне вилучення інформативних характеристик, семантичну сегментацію дорожньої інфраструктури та орієнтоване виявлення транспортних засобів.

Архітектура методу включає модуль вилучення ознак на основі глибокої згорткової мережі, який формує ієрархічне подання просторових та семантичних характеристик сцени [119]. Для забезпечення інваріантності до геометричних перетворень використовується механізм просторової нормалізації ознак, що дозволяє компенсувати вплив поворотів та масштабних змін об'єктів.

Нехай  $RGB \in \mathbb{R}^{H \times W \times D}$  позначає вхідний тензор ознак для довільного згорткового шару, де вхідне аерокосмічне RGB-зображення містить транспортні засоби довільної орієнтації та масштабу.

Нехай  $T(\cdot)$  – оператор просторового перетворення, що включає масштабування, обертання та афінні деформації:

$$T(I) = AI, \quad (2.10)$$

де  $A \in \mathbb{R}^{2 \times 3}$  – матриця афінного перетворення.

Нехай  $F(I) \in \mathbb{R}^{h \times w \times d}$  – багатовимірний тензор ознак, вилучений із вхідного зображення за допомогою згорткового відображення:

$$F(I) = C(I), \quad (2.11)$$

де  $C(\cdot)$  – композиція згорткових і нелінійних операторів.

Для компенсації впливу геометричних перетворень вводиться оператор просторової нормалізації:

$$\hat{F} = S(F, \theta), \quad (2.12)$$

де  $\theta$  – параметри просторового перетворення, оцінені з ознак, а  $S(\cdot)$ — оператор приведення локального просторового подання до канонічної форми.

Тоді для будь-якого допустимого афінного перетворення  $T$  виконується:

$$S(C(T(I)), \theta_T) \approx S(C(I), \theta), \quad (2.13)$$

що забезпечує просторове подання до змін орієнтації та масштабу об'єктів.

Масштабна стійкість внутрішнього подання реалізується за рахунок формування ознак на різних рівнях просторової роздільності. Нехай:

$$F_l \in \mathbb{R}^{h_l \times w_l \times d_l}, \quad (2.14)$$

де  $l=1, \dots, L$ .

Формується узагальнене багатомасштабне подання:

$$F^* = \bigoplus_{l=1}^L \psi_l(F_l), \quad (2.15)$$

де  $\psi_l(\cdot)$  – оператор приведення ознак до спільного простору, а  $\bigoplus$  – операція агрегування.

Таким чином, для об'єктів різного масштабу виконується:

$$F^*(I) \approx F^*(S_\lambda(I)), \quad (2.16)$$

де  $S_\lambda$  – оператор масштабування з коефіцієнтом  $\lambda$ .

Для коректного опису транспортних засобів, що мають довільний кут повороту, результат локалізації представляється у вигляді орієнтованої рамки:

$$O=(x, y, w, h, \alpha), \quad (2.17)$$

де  $x, y$  – координати центра,  $w, h$  – розміри,  $\alpha$  – кут орієнтації.

Для будь-якого повороту зображення на кут  $\Delta\alpha$  виконується:

$$O(T\Delta\alpha(I))=(x, y, w, h, \alpha+\Delta\alpha), \quad (2.18)$$

що забезпечує коректне врахування орієнтації об'єктів у процесі формування ознак та сталість класифікаційного рішення до її змін.

Для зменшення впливу складного фону використовується контекстуальне подання сцени у вигляді семантичної карти:

$$M = \mathcal{G}(I), \quad (2.19)$$

де  $\mathcal{G}(\cdot)$  – оператор семантичного узагальнення.

Ознакове подання, що не залежить від масштабних перетворень, задається співвідношенням:

$$F^{ctx} = F^* \odot M, \quad (2.20)$$

де  $\odot$  – операція узгодження просторових ознак з контекстом.

Це дозволяє зберігати стабільність виявлення транспортних засобів навіть за наявності локальних геометричних спотворень фону.

Навчання методу здійснюється шляхом мінімізації узагальненої функції втрат:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{det} + \lambda_2 \mathcal{L}_{orient} + \lambda_3 \mathcal{L}_{inv}, \quad (2.21)$$

де  $\mathcal{L}_{det}$  – втрата локалізації та класифікації,  $\mathcal{L}_{orient}$  – втрата оцінювання орієнтації,  $\mathcal{L}_{inv}$  – втрата просторової узгодженості ознак.

## 2.5. Висновки до другого розділу

1. Запропоновано нейромережевий підхід до семантичної сегментації транспортних засобів на зображеннях надвисокого просторового розрізнення, який базується на використанні архітектури DeepLab у поєднанні з глибинною залишковою мережею ResNet. Застосування механізмів багатомасштабного вилучення ознак та функції втрат на основі коефіцієнта Dice забезпечує високу точність сегментації транспортних засобів, зокрема в умовах багатокласової сегментації. Такий підхід ефективно вирішує проблему дисбалансу класів і дозволяє коректно виділяти об'єкти навіть за наявності складного фону та часткових перекриттів.

2. Розроблено метод підвищення точності розпізнавання літаків на супутникових зображеннях високої просторової розрізненості, який ґрунтується на поєднанні попередньої контурної обробки та глибинного навчання. Метод включає формування навчальних даних на основі набору HRPlanesv2, виділення контурів об'єктів за допомогою алгоритму послідовного обходу меж та інтеграцію отриманих контурних ознак з оригінальними зображеннями. Подальше навчання моделі YOLOv8m на різних варіантах наборів даних дозволило оцінити вплив додаткових просторово-геометричних характеристик на ефективність виявлення літаків і підтвердило доцільність використання комбінованого представлення вхідних даних.

3. Розроблено інформаційну технологію виявлення автомобілів на аерофотозображеннях, що базується на модифікованій архітектурі YOLOv11 з підтримкою орієнтованих обмежувальних рамок (Oriented Bounding Boxes). Запропонований підхід дозволяє коректно локалізувати транспортні засоби довільної орієнтації, що є важливим для аерофотознімків.

4. Запропоновано метод просторово-орієнтованого розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях, який забезпечує стійкість результатів до змін масштабу та орієнтації. Метод ґрунтується на формуванні просторово нормалізованого багатомасштабного ознакового подання та узгодженому описі параметрів орієнтації транспортних засобів у вигляді орієнтованих обмежувальних рамок. Для підвищення точності розпізнавання в умовах складного фону введено механізм контекстуального узгодження ознакових представлень із семантичним поданням сцени, що дозволяє зменшити кількість хибних спрацювань і підвищити точність локалізації близько розташованих транспортних засобів за варіативних умов аерокосмічної зйомки.

Основні результати даного розділу опубліковані в роботах автора [57, 71, 75, 91, 116, 120].

## РОЗДІЛ 3

# РОЗРОБКА ТА ПРАКТИЧНА РЕАЛІЗАЦІЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ НЕЙРОМЕРЕЖЕВОВОГО РОЗПІЗНАВАННЯ ТРАНСПОРТНИХ ЗАСОБІВ ДОВІЛЬНОГО РОЗТАШУВАННЯ НА АЕРОКОСМІЧНИХ ЗОБРАЖЕННЯХ

### 3.1. Інформаційна технологія нейромережевого розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях

Незважаючи на ефективність запропонованої у розділі 2 інформаційної технології виявлення автомобілів на аерофотозображеннях на основі модифікованої архітектури YOLOv11-OBV, проведений аналіз результатів експериментальних досліджень дозволив виявити низку обмежень, які зумовлюють доцільність подальшого розвитку підходів до розпізнавання транспортних засобів довільного розташування. По-перше, архітектура YOLOv11-OBV орієнтована переважно на локальне виявлення окремих об'єктів і не враховує просторову щільність транспортних засобів у сцені, що є характерним для сцен із щільним розміщенням транспортних засобів, зокрема на парковках, у транспортних вузлах та на міських територіях. У таких умовах спостерігається часткове перекриття орієнтованих обмежувальних рамок, зменшення відстаней між об'єктами та підвищений вплив фону, що ускладнює коректне розділення близько розташованих транспортних засобів. По-друге, попередня технологія на основі архітектури YOLOv11-OBV не використовує апріорну семантичну інформацію про структуру сцени, зокрема про дорожню інфраструктуру та зони ймовірного розміщення транспортних засобів (дороги, парковочні місця). Розпізнавання виконується без явного урахування контексту, що обмежує здатність моделі відрізнити транспортні засоби від об'єктів фону зі схожими спектральними та текстурними характеристиками, особливо в

умовах складного освітлення або тіней. По-третє, хоча використання орієнтованих обмежувальних рамок підвищує точність локалізації об'єктів довільної орієнтації, інваріантність до геометричних перетворень досягається опосередковано, без явного механізму просторового узгодження ознак. Це знижує стійкість моделі до значних змін ракурсу зйомки, перспективних спотворень та локальних деформацій, притаманних низьковисотним аерокосмічним знімкам. Подібні проблеми спостерігалися і в інших дослідженнях [120–130], де аналізували ефективність різних підходів глибокого навчання для розпізнавання транспортних засобів на аерофотозображеннях, підкреслюючи обмеження одноетапних та двоетапних моделей у складних сценах і при обробці щільно розташованих об'єктів.

Вище зазначені обмеження обумовили необхідність розробки вдосконаленої інформаційної технології виявлення транспортних засобів довільного розташування, яка поєднує (рис.3.1):

- просторове узгодження ознакових представлень із використанням нейронної мережі просторових перетворень (Spatial Transformer Network, STN);
- багатомасштабне вилучення та інтеграцію ознак, спрямовані на підвищення точності локалізації транспортних засобів малого розміру та коректне виявлення близько розташованих транспортних засобів;
- семантичну сегментацію елементів дорожньої інфраструктури з метою формування контекстно-орієнтованих карт ознак;
- орієнтоване виявлення транспортних засобів у рамках єдиної інтегрованої нейромережевої моделі, що забезпечує спільне навчання взаємопов'язаних завдань аналізу сцени.

Підготовка набору даних здійснювалася на основі аерокосмічних знімків високої роздільної здатності, отриманих камерою SONY DSC-WX220 у режимі низьковисотного моніторингу. Вихідні зображення високого просторового розрізнення були просторово фрагментовані на

тайли фіксованого розміру  $200 \times 200$  пікселів, що забезпечило уніфікацію вхідних даних та можливість ефективної пакетної обробки нейронною мережею. Обраний розмір тайлів дозволив одночасно зберегти достатній локальний контекст для ідентифікації малих транспортних засобів та зменшити обчислювальне навантаження під час навчання моделі. На основі сформованих тайлів було створено підмножини для навчання (70%), валідації(15%) та тестування(15%), що забезпечило репрезентативність даних та коректність подальшої оцінки якості моделі.

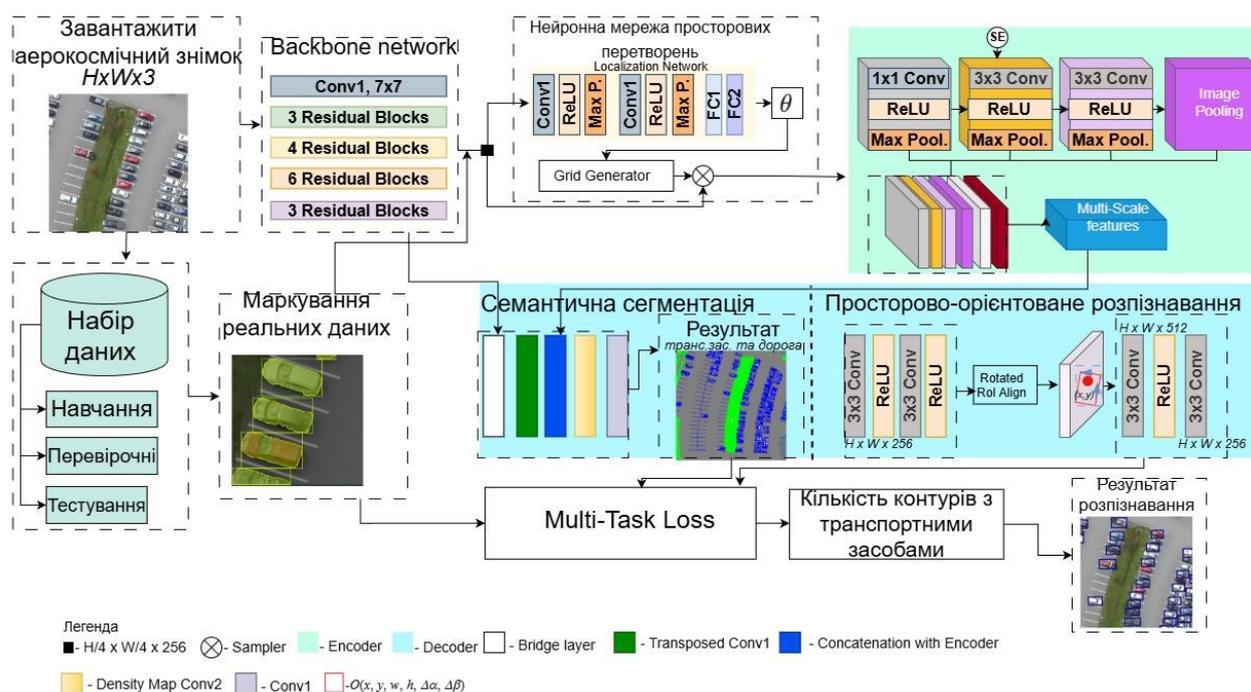


Рисунок 3.1 – Структурна схема технології нейромережевого інваріантного розпізнавання транспортних засобів

Анотація транспортних засобів виконувалася для кожного тайла окремо та передбачала точне виділення об'єктів із фіксацією їхнього положення та орієнтації. Замість традиційних axis-aligned прямокутників використовувалися орієнтовані обмежувальні прямокутники (Oriented Bounding Boxes, OBB), що є більш придатними для аерокосмічних зображень, де транспортні засоби можуть бути повернуті під довільним кутом. Кожна мітка містила координати центру, ширину, висоту та кут обертання, що дозволяло мінімізувати зайвий фон та підвищити точність

навчання моделі на малих об'єктах. Анотації зберігалися у текстових файлах, де кожний рядок описував один об'єкт у форматі, сумісному з компонентом орієнтованого детектування в структурі Multi-Task Loss. Усі мітки були організовані у відповідних директоріях train/labels та valid/labels, що забезпечило чітку структуру набору даних та можливість його прямого використання під час навчання моделі.

На вхід моделі подається аерокосмічне RGB-зображення розмірності  $H \times W \times 3$  для вилучення ієрархічних ознак за допомогою Базової мережі (Backbone Network). Ця мережа побудована на принципах глибокої залишкової архітектури ResNet. Обробка починається з початкового згорткового шару (Conv1) з розміром ядра  $7 \times 7$ , який виконує виділення низькорівневих ознак, таких як контури та текстури. Далі слідує послідовність Резидуальних Блоків (Residual Blocks) (глибиною 3, 4, 6 та 3 блоки відповідно), які використовують структуру залишкових з'єднань (skip connections). Ці блоки формують ієрархічні багаторівневі представлення, які здатні описувати як початкові контури, так і складніші семантичні структури. У процесі проходження через ці блоки формуються проміжні тензори ознак, які слугують основою для подальших паралельних обчислень. Кожен згортковий шар у Backbone Network виконує операцію згортки відповідно до принципів глибокого навчання.

Нехай  $RGB \in \mathbb{R}^{H \times W \times D}$  позначає вхідний тензор ознак для довільного згорткового шару. Набір навчених згорткових ядер (фільтрів) позначається як  $F = [g_1, g_2, \dots, g_c]$ . Вихідний канал ознак  $V^c$  для  $c$ -го фільтра  $g_c \in \mathbb{R}^{K \times K \times D}$  обчислюється як згортка [131]:

$$V^c = g_c * I = \sum_{s=1}^D g_c^s * Y^s. \quad (3.1)$$

$Y_s$  – це  $s$ -й вхідний канал тензора RGB,  $g_c \in \mathbb{R}^{K \times K}$  це 2D просторове ядро, що є частиною  $c$ -го фільтра. Після операції згортки для стабілізації навчання та прискорення збіжності застосовується шар Batch Normalization (BN). Для посилення репрезентативної здатності виконується лінійна

трансформація масштабування ( $\gamma$ ) та зміщення ( $\beta$ ), після чого застосовується нелінійна активаційна функція  $\sigma(\text{ReLU})$  [132]:

$$V_{out}^c = \sigma(\gamma \cdot \text{BN}(V^c) + \beta). \quad (3.2)$$

Для забезпечення інваріантності використовується модуль нейронна мережа просторових перетворень (Spatial Transformer Network, STN). Локалізаційна мережа STN складається із згорткових шарів та шарів повного з'єднання (Fully Connected), які оцінюють параметри афінної трансформації. На основі цих параметрів модуль Grid Generator формує сітку вибірки, а Sampler виконує просторове перетворення ознак. STN дозволяє приводити локальні області до стандартної орієнтації, покращуючи всю подальшу обробку. Після нормалізації ознак за допомогою STN формується багатомасштабне подання, побудоване на основі комбінації згорткових шарів  $1 \times 1$  та  $3 \times 3$ , активації ReLU, операцій Max Pooling та глобального агрегування (Image Pooling). Узгоджене представлення різних просторових масштабів забезпечує високу чутливість детектора до малих транспортних засобів, особливо в умовах складного фону.

На наступному етапі декодер отримує ознаки, сформовані Backbone Network та STN, та використовує їх для паралельного виконання семантичної сегментації елементів дорожньої інфраструктури і виявлення транспортних засобів із забезпеченням інваріантності до їх орієнтації у сцені. Модуль семантичної сегментації, який реалізований у вигляді U-подібної архітектури. Ця структура ефективно працює для задач піксельної класифікації, оскільки вона забезпечує збереження як глобальних семантичних ознак (високорівневі ознаки), так і точних просторових деталей (низькорівневі ознаки). Архітектура складається з трьох послідовних компонентів: кодера (Encoder), який використовує згорткові шари для поступового зменшення просторової роздільної здатності та вилучення високосемантичних ознак; шару з'єднання енкодера та декодера (Bridge Layer), що містить найбільш стиснене семантичне представлення

сцени; та декодера (Decoder). Декодер відновлює повну просторову роздільну здатність та виконує класифікацію пікселів, використовуючи шари транспонованих згорток (Transposed Convolutions) для поетапного збільшення роздільної здатності карти ознак. Ключовим елементом U-подібної структури є з'єднання із проміжними рівнями ознак (skip connections). На етапі декодування відбувається об'єднання каналів (concatenation) інформації, отриманої від декодера. Це дозволяє ефективно передавати деталізовану просторову інформацію від початкових шарів до кінцевих шарів, запобігаючи втраті точних границь об'єктів. Кінцевим результатом роботи модуля є контекстуальна карта. Ця карта являє собою класифікацію кожного пікселя зображення, що містить семантичні відомості про ключові елементи сцени: дороги, транспортні засоби та загальний фон. Генерація контекстуальної карти забезпечує семантичну структурованість сцени та передає ці контекстуальні відомості до модуля виявлення. Це підвищує точність локалізацію невеликих об'єктів (малих автомобілів), оскільки модуль детектування отримує апріорну інформацію щодо найбільш імовірних регіонів для присутності цільових об'єктів.

Етап забезпечення інваріантності до орієнтації транспортних засобів у сцені та їхнього точного виявлення реалізується у спеціалізованому модулі Oriented Detection. Його основне функціональне призначення – забезпечити точну просторову локалізацію та визначення кута обертання цільових об'єктів, незалежно від їхньої довільної орієнтації на аерокосмічному знімку. Модуль Oriented Detection отримує на вхід багатомасштабні ознаки від Backbone Network (зокрема, після їхньої трансформації модулем STN) та контекстуальну карту від модуля семантичної сегментації. Об'єднані карти ознак піддаються додатковій обробці через послідовність згорткових блоків (3x3 із функцією активації ReLU) для посилення їхніх семантичних та просторових характеристик. Операція Rotated RoI Align виконує просторове вирівнювання ознак усередині кутових (орієнтованих) областей інтересу, усуваючи залежність від кута обертання об'єкта.

Механізм орієнтованого вирівнювання ознак формує стандартизоване та інваріантне представлення ознак, що забезпечує підвищену точність подальших завдань класифікації та регресії. Після вирівнювання ознак здійснюється регресія параметрів орієнтованого обмежувального прямокутника ( $O$ ) та класифікація ("транспортний засіб"/"не транспортний засіб"). Кожен орієнтований обмежувальний прямокутник ( $O$ ) описується параметрами ( $x, y, w, h, \Delta\alpha, \Delta\beta$ ), які формують вершини об'єкта відповідно до Midpoint Offset Representation (репрезентації зсуву відносно середин сторін) [133]:

$$\begin{aligned} v_1 &= \left(x, y - \frac{h}{2}\right) + (\Delta\alpha, 0) \\ v_2 &= \left(x + \frac{w}{2}, y\right) + (0, \Delta\beta) \\ v_3 &= \left(x - \frac{w}{2}, y\right) + (0, -\Delta\beta) \end{aligned} \quad (3.3)$$

Результат оптимізації моделі формується шляхом мінімізації багатозадачної функції втрат ( $\mathcal{L}$ ), яка об'єднує результати всіх паралельних завдань. Це забезпечує одночасну оптимізацію інваріантності, сегментації та розпізнавання транспортних засобів:

$$\mathcal{L} = \alpha \mathcal{L}_{Detection} + \beta \mathcal{L}_{Segmentation} + \gamma \mathcal{L}_{STN}, \quad (3.4)$$

де  $\mathcal{L}_{Detection}$  включає класифікаційну та регресійну втрати для ОБВ,  $\mathcal{L}_{Segmentation}$  втрату для піксельної класифікації,  $\mathcal{L}_{STN}$  - втрату, пов'язану з якістю просторової трансформації. Мінімізація  $\mathcal{L}$  гарантує, що модель є робушною як у виявленні, так і в семантичному розумінні сцени.

Результатом є розпізнавання транспортних засобів на аерокосмічному знімку, виділених орієнтованими рамками. Розрахунок загальної кількості транспортних засобів (Total Vehicles Counts) відбувається після порогової фільтрації передбачень та застосування алгоритму не-максимального пригнічення (Non-Maximum Suppression), загальна кількість ТЗ еквівалентна кількості елементів у фінальному, ненадлишковому наборі орієнтованих обмежувальних рамок. Інтеграція просторово адаптивних механізмів,

контекстуальної інформації та багатомасштабного аналізу ознак забезпечує високу точність виявлення транспортних засобів у складних умовах [134].

### **3.2. Розробка програмного додатку розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях**

На основі розробленої неймережевої інформаційної технології неймережевого розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях було створено програмний додаток, що забезпечує практичну реалізацію запропонованого підходу та дозволяє виконувати аналіз аерокосмічних знімків у реальних умовах. Розроблена інформаційна технологія реалізована у вигляді багаторівневої програмної системи, архітектура якої побудована за принципом функціонального розділення на рівні представлення, даних та обробки. Такий підхід забезпечує масштабованість, модульність та можливість адаптації технології до різних типів аерокосмічних даних і умов зйомки.

Архітектура програмного додатку представлено на рис.3.2 включає чотири логічні рівні:

- рівень представлення;
- рівень даних;
- рівень обробки даних;
- рівень програмно-технологічної реалізації.

Рівень представлення відповідає за взаємодію користувача з програмним додатком та візуалізацію результатів обробки аерокосмічних зображень. На цьому рівні реалізовано клієнтський інтерфейс і модуль візуалізації, які забезпечують завантаження вхідних даних, ініціалізацію процесів обробки, перегляд результатів сегментації та розпізнавання транспортних засобів, а також формування звітної інформації. Відокремлення рівня представлення дозволяє ізолювати логіку взаємодії з користувачем від алгоритмічної частини системи.

Рівень даних призначений для зберігання та управління наборами аерокосмічних зображень і відповідних анотованих даних транспортних засобів. Він включає хмарне сховище, яке забезпечує централізований доступ до навчальних і тестових вибірок, а також підтримує операції

оновлення, розширення та повторного використання наборів даних. Використання хмарного підходу спрощує масштабування системи та підвищує відтворюваність експериментальних досліджень.

Рівень обробки даних є ключовим функціональним рівнем інформаційної технології та охоплює етапи попередньої обробки аерокосмічних зображень і реалізацію методів машинного навчання. На цьому рівні здійснюється нормалізація вхідних даних, усунення шумів, приведення зображень до уніфікованого формату, а також виконуються процедури семантичної сегментації дорожньої інфраструктури й розпізнавання транспортних засобів довільного розташування. Логіка цього рівня реалізує запропонований нейромережевий метод та забезпечує його стійкість до змін масштабу, орієнтації та геометричних спотворень об'єктів.



Рисунок 3.2 – Структурна схема архітектури програмного додатку

Рівень програмно-технологічної реалізації визначає сукупність програмних і апаратних засобів, на основі яких реалізовано інформаційну технологію. Він включає бібліотеки комп'ютерного зору, фреймворки глибокого навчання, засоби побудови графічного інтерфейсу та апаратні

ресурси прискорення обчислень. Даний рівень не виконує безпосередньої обробки даних, але забезпечує ефективну реалізацію алгоритмів, їх інтеграцію та практичну придатність системи для аналізу аерокосмічних зображень великого обсягу.

Графічний інтерфейс програмного додатку орієнтований на поетапне виконання процесу аналізу та візуалізацію отриманих результатів у зрозумілій формі (рис.3.3).

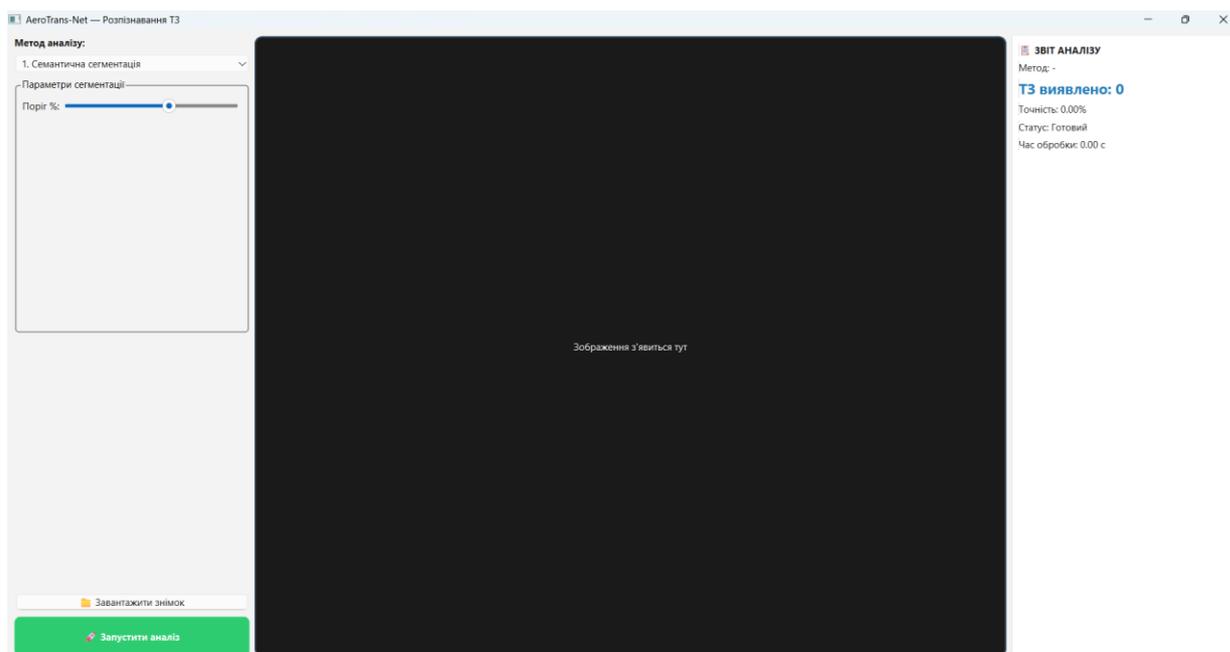


Рисунок 3.3 – Графічний інтерфейс

Головне вікно додатку поділено на функціональні області відповідно до логіки роботи інформаційної технології. У лівій частині інтерфейсу розташовано панель керування параметрами аналізу, яка дозволяє користувачеві обрати метод обробки аерокосмічних зображень (семантична сегментація або розпізнавання транспортних засобів довільного розташування), а також задати ключові параметри алгоритмів, зокрема поріг імовірності виявлення об'єктів та мінімальну допустиму площу цільових об'єктів. Такий підхід забезпечує гнучке налаштування процесу аналізу залежно від характеристик вхідних даних та умов зйомки.

Центральна частина інтерфейсу призначена для відображення вхідних аерокосмічних зображень і результатів їх обробки. У цій області здійснюється візуалізація сегментованих областей, обмежувальних рамок транспортних засобів, а також результатів розпізнавання довільного

розташування з урахуванням просторової орієнтації об'єктів. Візуальне представлення результатів дозволяє оперативно оцінити коректність локалізації та якість роботи нейромережових моделей.

У правій частині головного вікна розміщено інформаційний блок формування звіту аналізу, який відображає узагальнені числові характеристики результатів обробки. До таких характеристик належать кількість виявлених транспортних засобів, значення показників точності (зокрема *mAP*), а також час обробки зображення. Наявність цього блоку забезпечує можливість порівняльної оцінки ефективності різних методів та параметрів аналізу без необхідності звернення до зовнішніх засобів обробки результатів.

З метою формалізації структури програмного додатку та сценаріїв його використання в роботі застосовано UML-моделювання, яке дозволяє наочно представити функціональні можливості системи та взаємодію її компонентів.

Діаграма варіантів використання відображає основні сценарії взаємодії користувача з програмним додатком (рис.3.4). Основним актором системи є користувач, який ініціює процеси завантаження аерокосмічних зображень, вибору методу аналізу, налаштування параметрів обробки та запуску розпізнавання транспортних засобів довільного розташування. Окремими варіантами використання є перегляд результатів обробки, аналіз кількісних показників та формування звітної інформації.

Діаграма класів відображає внутрішню структуру програмного додатку та взаємозв'язки між його основними компонентами (рис.3.5). У моделі виділено класи, що відповідають за графічний інтерфейс користувача, модулі обробки аерокосмічних зображень, нейромережові моделі сегментації та розпізнавання довільного розташування, а також класи для візуалізації результатів і формування звітних даних. Взаємодія між класами реалізує логіку послідовного виконання етапів аналізу, що відповідає архітектурі інформаційної технології, представленої на структурній схемі.

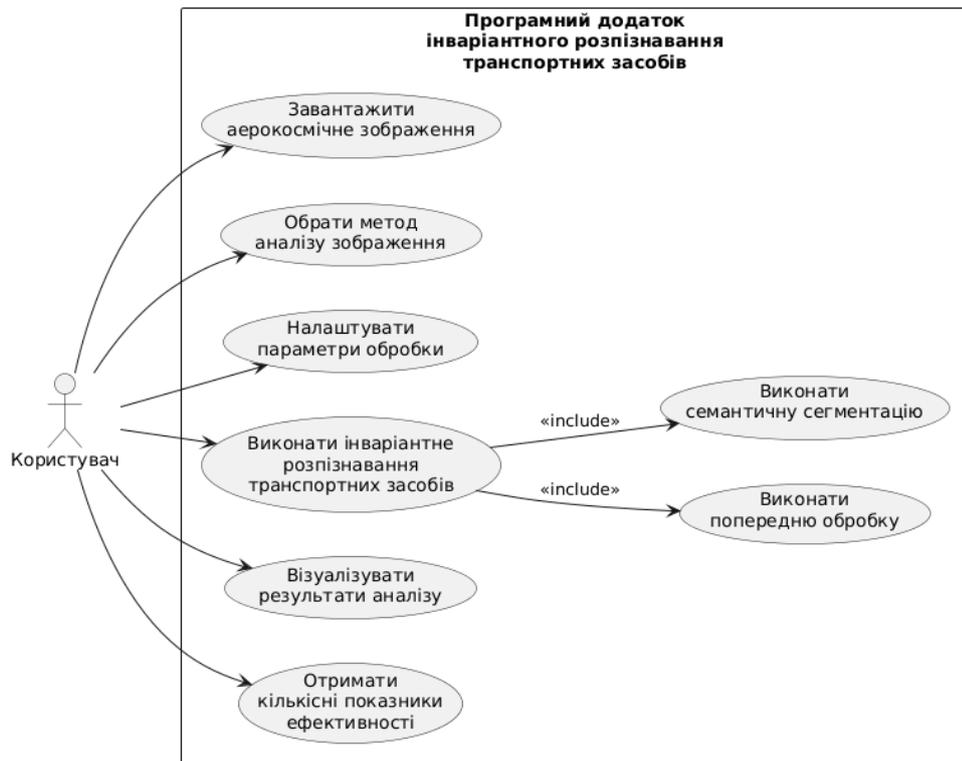


Рисунок 3.4 – UML-діаграма варіантів використання програмного додатку розпізнавання транспортних засобів довільного розташування

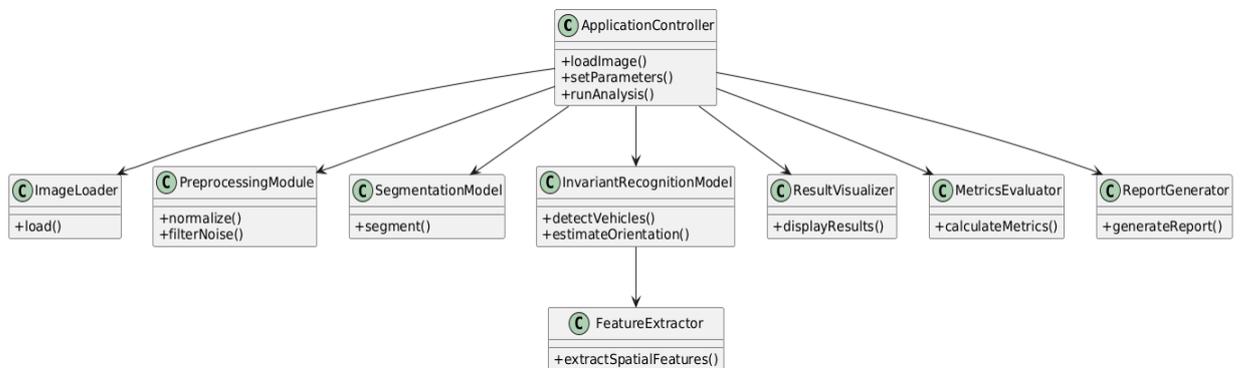


Рисунок 3.5 – UML-діаграма класів програмного додатку для розпізнавання транспортних засобів довільного розташування

### 3.3. Функціональні можливості програмного додатку

У межах дослідження розроблений програмний додаток забезпечує інтерактивну взаємодію користувача з інформаційною технологією розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях. Процес використання додатку структуровано у

вигляді послідовних етапів, що відповідають логіці обробки даних та отримання аналітичних результатів (рис.3.6).

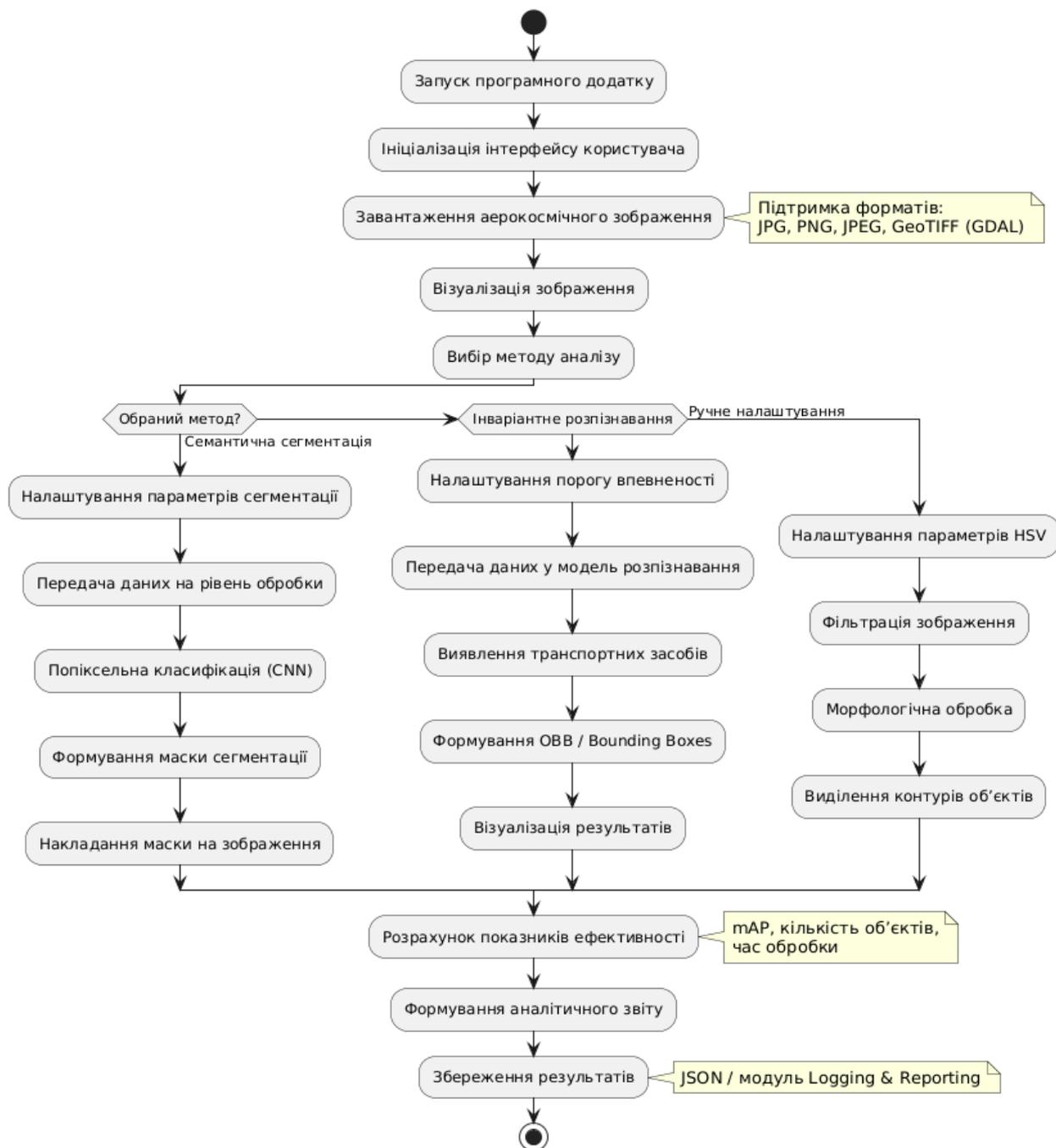


Рисунок 3.6 – Логічна схема роботи програмного додатку розпізнавання ТЗ довільного розташування

На першому етапі здійснюється ініціалізація та завантаження вхідних даних. Користувач обирає аерокосмічне зображення з локальної файлової системи. Програмний додаток підтримує імпорт растрових зображень високої роздільної здатності у поширених форматах (.jpg, .png, .jpeg), а

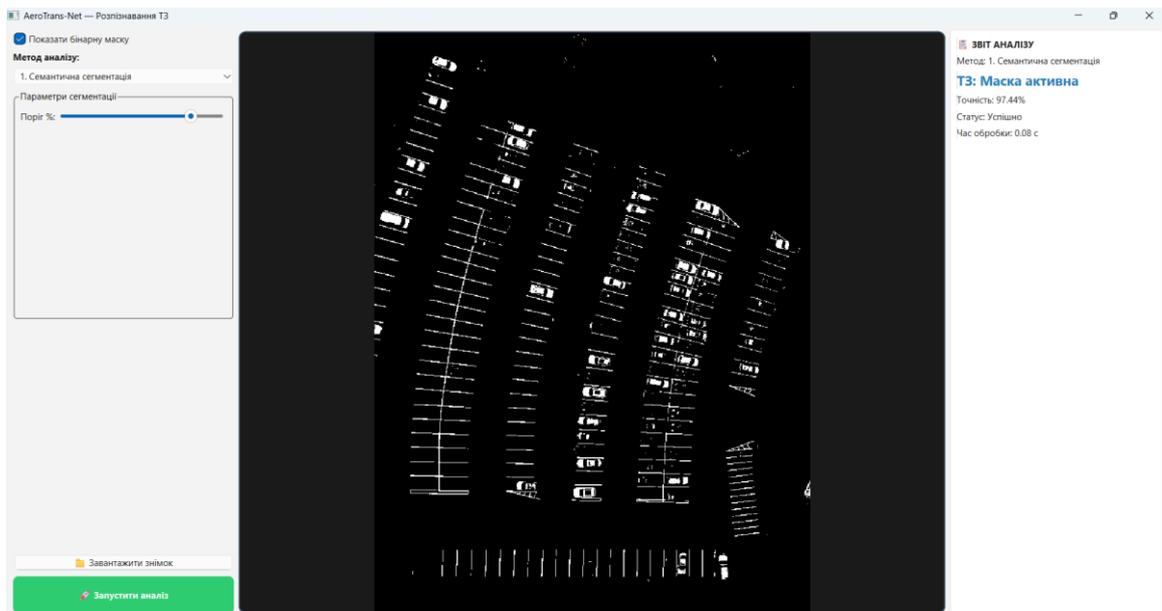
також професійних геопросторових даних формату GeoTIFF [126], що забезпечується інтеграцією бібліотеки GDAL. Після завантаження зображення відображається у центральному вікні візуалізації, що підтверджує коректну ініціалізацію внутрішніх модулів обробки даних.

На другому етапі користувач виконує налаштування параметрів аналізу та обирає відповідний метод обробки. Через панель параметрів здійснюється вибір між режимом семантичної сегментації, який призначений для виділення просторових контурів об'єктів, та режимом розпізнавання транспортних засобів довільного розташування, орієнтованого на ідентифікацію окремих класів об'єктів незалежно від їх орієнтації, масштабу та умов зйомки. Додатково користувач має змогу регулювати порогові значення ймовірності розпізнавання, що дозволяє мінімізувати кількість хибнопозитивних результатів та адаптувати систему до специфіки конкретної сцени.

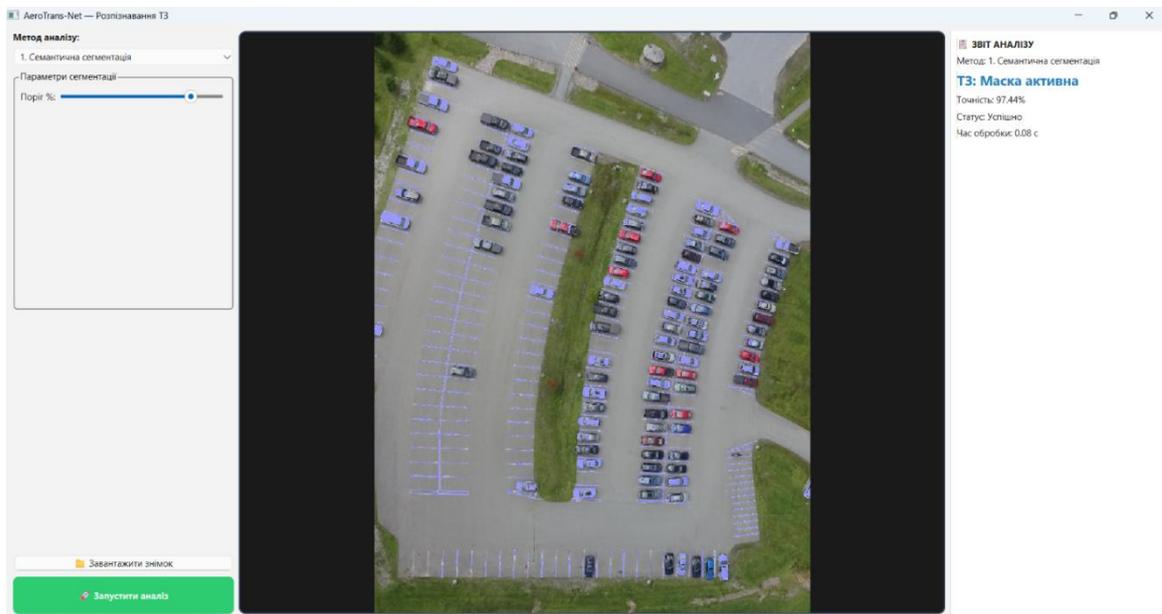
Третій етап передбачає запуск процесу аналізу та формування результатів. Після ініціації обчислювального процесу дані передаються на рівень обробки, де здійснюються обчислювальні ресурси графічного процесора з використанням паралельних обчислень.

У режимі семантичної сегментації (рис.3.7) виконується попіксельна класифікація з формуванням бінарних або багатоканальних масок. Отримані результати накладаються на вихідне зображення у вигляді напівпрозорої кольорової маски, що дозволяє візуально оцінити просторове розташування та межі виявлених транспортних засобів.

Завершальним етапом є формування аналітичної звітності та верифікація результатів. У правій панелі додатку автоматично відображаються кількісні показники ефективності, зокрема значення середньої точності розпізнавання (mAP), кількість виявлених об'єктів та час обробки одного зображення. Отримані значення свідчать про високу швидкодію та придатність запропонованої інформаційної технології до використання в режимі, близькому до реального часу. Результати аналізу зберігаються у структурованому вигляді та можуть передаватися до модуля звітності для подальшого накопичення та аналізу.



а)



б)

Рисунок 3.7 – Результат семантичної сегментації: а) бінарна; б) кольорова

Окремим сценарієм використання є режим розпізнавання транспортних засобів довільного розташування, у якому система виконує розпізнавання транспортного засобу з формуванням обмежувальних рамок. Користувач має можливість змінювати порогові значення впевненості, що дозволяє адаптувати результати до умов високої щільності об'єктів. На тестових сценах з паркувальними майданчиками система демонструє високі показники точності розпізнавання при прийнятному часі обробки навіть для складних сцен (рис.3.8).

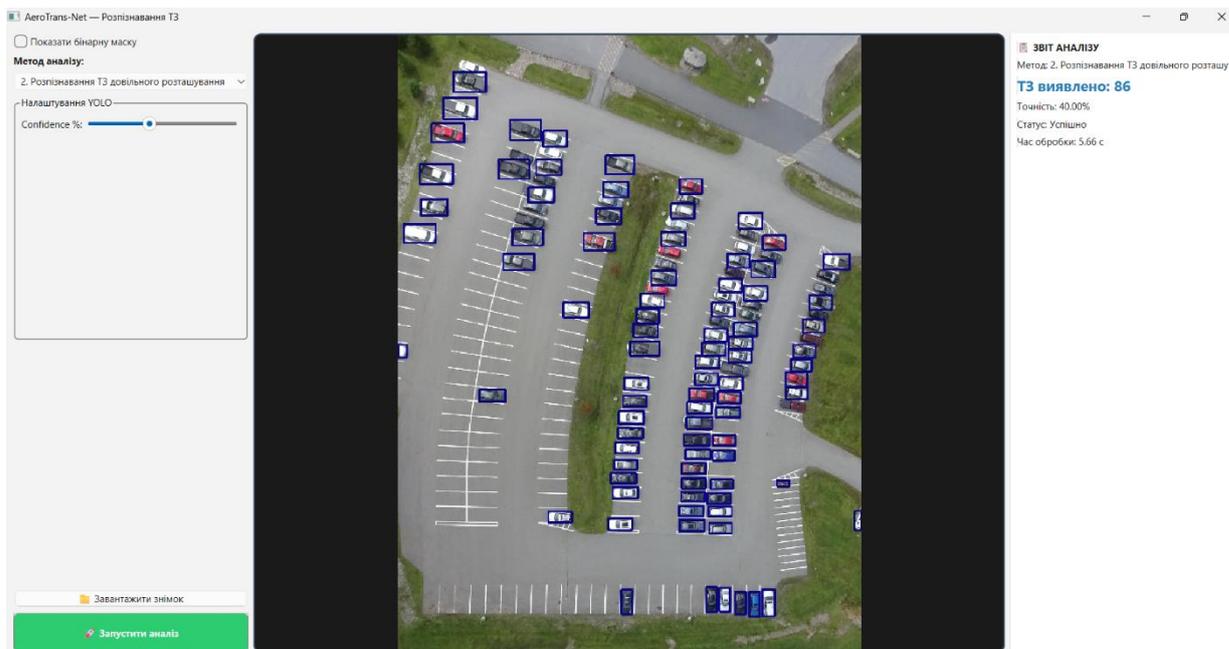


Рисунок 3.8 – Результат розпізнавання ТЗ довільного розташування

У межах програмного додатку реалізовано режим ручного налаштування розпізнавання об'єктів, який використовується як допоміжний інструмент для уточнення результатів у складних умовах зйомки або за наявності специфічних спектральних характеристик транспортних засобів. При активації методу «Ручне налаштування розпізнавання об'єктів» ліва панель управління змінює склад доступних інструментів, надаючи користувачеві доступ до блоку «Налаштувати параметри» (рис.3.9).

Ручний режим базується на обробці зображень у колірному просторі HSV (Hue, Saturation, Value), що дозволяє ефективно відокремлювати об'єкти від фону за їх спектральними та фотометричними ознаками. Зокрема, параметр Hue (тон) використовується для вибору цільового колірною діапазону, що дає змогу акцентувати увагу на транспортних засобах із певними колірними характеристиками. Параметр Saturation (насиченість) застосовується для відсікання ахроматичних компонентів підстильної поверхні та зменшення впливу шумів, пов'язаних із дорожнім покриттям або тінями. Параметр Value (яскравість) забезпечує адаптацію алгоритму до умов освітленості знімка, включаючи ділянки з інтенсивними відблисками або затіненими зонами.

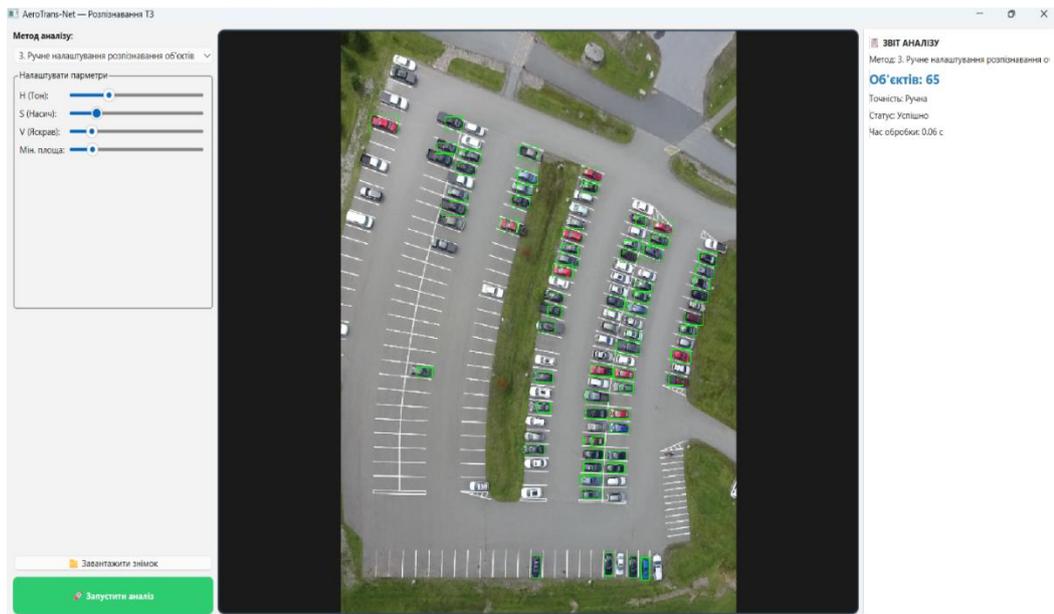


Рисунок 3.9 – Результат ручного налаштування параметрів для розпізнавання ТЗ

Додатково реалізовано параметр мінімальної площі об'єкта, який виконує функцію морфологічного фільтра. Його використання дозволяє ігнорувати дрібні елементи ландшафту або артефакти зображення, площа яких у пікселях не перевищує заданий поріг, що сприяє зниженню кількості хибнопозитивних спрацьовувань.

Взаємодія користувача з системою у ручному режимі реалізується за ітераційним сценарієм. Після завантаження вихідного аерокосмічного знімка та вибору відповідного методу аналізу користувач здійснює інтерактивне коригування параметрів HSV. Зміна значень параметрів супроводжується миттєвим візуальним відгуком у центральному вікні візуалізації, де система динамічно підсвічує виявлені контури транспортних засобів. Після досягнення задовільного результату виконується фіналізація аналізу з автоматичним розрахунком метрик та формуванням звітних даних.

### 3.4. Висновки до третього розділу

1. Запропоновано інформаційну технологію нейромережевого розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях високої просторової роздільної здатності. Технологія базується на багатокомпонентній архітектурі, яка поєднує

глибоке вилучення ознак через Backbone Network (ResNet), просторову стабілізацію за допомогою Spatial Transformer Network (STN) та розпізнавання ТЗ з використанням орієнтованих обмежувальних рамок (ОВВ) і багатозадачної функції втрат (Multi-Task Loss).

2. Розроблено програмний додаток розпізнавання транспортних засобів довільного розташування із багаторівневою архітектурою (Presentation, Logic, Data, Implementation layers). Програмна реалізація виконана мовою Python із використанням PySide6 для графічного інтерфейсу та фреймворку PyTorch для неймережевих обчислень. Реалізовано два основні режими роботи додатку: семантичну сегментацію та розпізнавання транспортних засобів довільного розташування. Режим семантичної сегментації забезпечує попіксельну класифікацію з формуванням бінарних або багатоканальних масок, що накладаються на вихідне зображення для візуальної оцінки розташування та меж об'єктів. Режим розпізнавання ТЗ довільного розташування дозволяє визначати положення транспортних засобів у довільній орієнтації та масштабі, з можливістю адаптації порогів ймовірності розпізнавання для мінімізації хибнопозитивних результатів навіть у сценах з високою щільністю об'єктів.

3. Впроваджено сервіс-орієнтований підхід через модуль Logging & Transfer, що забезпечує автоматичне формування звітів та передачу результатів у форматі JSON до хмарного сховища та інтегрування технології у середовище віддаленого моніторингу транспортної інфраструктури. Доведено ефективність такої архітектури для асинхронної обробки потоків даних із високою швидкістю та надійністю результатів.

Результати досліджень даного розділу опубліковано в роботах автора [120, 135, 140].

## РОЗДІЛ 4

### ЕКСПЕРИМЕНТАЛЬНА ПЕРЕВІРКА ЕФЕКТИВНОСТІ ЗАПРОПОНОВАНИХ НЕЙРОМЕРЕЖЕВИХ МЕТОДІВ І ТЕХНОЛОГІЙ РОЗПІЗНАВАННЯ ТРАНСПОРТНИХ ОБ’ЄКТІВ НА АЕРОКОСМІЧНИХ ЗОБРАЖЕННЯХ

#### 4.1. Методика експериментальних досліджень

##### 4.1.1. Набори даних

Для проведення експериментальних досліджень ефективності нейромережових методів розпізнавання транспортних засобів на аерокосмічних зображеннях використано відкриті (публічно доступних) набори даних, що широко застосовуються у задачах розпізнавання об’єктів та семантичного аналізу зображень. Огляд основних наборів даних для виявлення транспортних засобів на аерофотозображеннях та супутникових зображеннях, а також їх ключових характеристик, наведено в таблиці 4.1. При цьому кількість екземплярів обмежено лише транспортними засобами, оскільки інші категорії об’єктів (будівлі, мости, порти тощо) не входять до предметної області даного дослідження.

Як базовий набір даних для дослідження запропонованого конвеєра виявлення транспортних засобів обрано DLR 3K Munich Vehicle Aerial Image Dataset (DLR 3K) було отримано з висоти приблизно 1000 м над поверхнею землі над територією міста Мюнхен (Німеччина). Зйомка охоплює переважно міські та житлові райони, що характеризуються складною просторовою структурою, високою щільністю забудови та різноманітними фоновими об’єктами, типовими для урбанізованих територій. Набір даних DLR 3K містить 20 аерофотознімків (рис.4.1) з просторовою роздільною здатністю  $5616 \times 3744$  пікселів та середнім значенням просторового кроку зйомки (GSD) близько 13 см. Еталонні розмітки (ground truth, GT) у вигляді орієнтованих обмежувальних рамок

для різних типів транспортних засобів, зокрема car, truck та trailer, надані лише для 10 зображень, що обмежує обсяг доступних анотованих даних.

Таблиця 4.1 – Набори даних для виявлення транспортних засобів на аерофотозображеннях

Набір даних	Тип анотації	Кількість зображень	Ширина зображення (пікселів)	Кількість об'єктів	GSD (см)
TAS	Прямокутник (горизонтальний)	30	792	1,319	-
NWPU	Прямокутник (горизонтальний)	800	1	477	8-200
UCAS-AOD	Прямокутник (довільно орієнтований)	910	1	2,819	-
DLR 3K	Прямокутник (довільно орієнтований)	20	5,616	14,235	13
VEDAI	Прямокутник (довільно орієнтований)	1,268	1,024	2,95	12.5
COWC	Позначка центру (точка)	53	2k-19k	32,716	15
DOTA	Прямокутник (довільно орієнтований)	2,806	800-13k	~180k	10-100
ITCVD	Прямокутник (горизонтальний)	173	5,616	29,088	10
xView	Прямокутник (горизонтальний)	1,127	700-4k	~250k	30

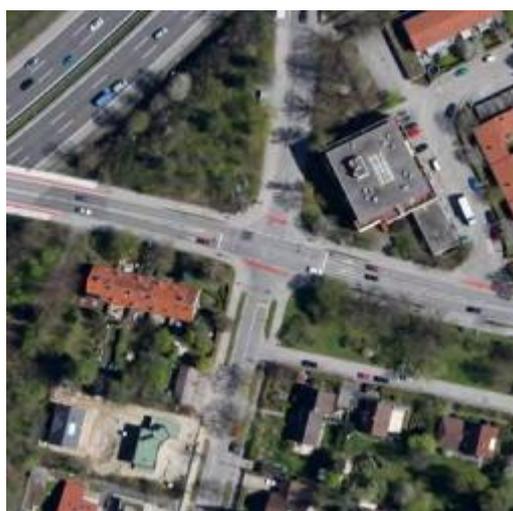


Рисунок 4.1 – Приклади набору даних аерофотознімків транспортних засобів DLR 3K [133]

Набір даних Cars Overhead with Context (COWC), який містить ще більшу кількість транспортних засобів, у межах даної роботи не використовувався як основний, оскільки анотації в ньому подані лише у вигляді координат центрів об'єктів без повноцінних обмежувальних рамок. Це обмежує можливості застосування сучасних методів детекції, які потребують точного опису геометричних параметрів об'єктів.

З метою оцінювання узагальнювальної здатності запропонованих нейромережових підходів додаткові експерименти проведено з використанням набору даних Vehicle Detection in Aerial Imagery (VEDAI) [133], який є загальноприйнятим еталонним набором для задач виявлення транспортних засобів на аерофотозображеннях. Набір даних VEDAI містить супутникові зображення, отримані агентством Utah Automated Geographic Reference Center (AGRC) під час аерозйомки території штату Юта (США) навесні 2012 року. Початкові зображення представлені у вигляді чотириканальних даних (RGB та ближній інфрачервоний діапазон), однак у межах даного дослідження для обробки використовувалися лише RGB-канали, що відповідає умовам більшості практичних застосувань нейромережових методів розпізнавання.



Рисунок 4.2 – Приклади набору даних аерофотознімків транспортних засобів VEDAI [133]

Як показано на рис. 4.2, зображення VEDAI характеризуються значною варіативністю фону, охоплюючи сільськогосподарські, сільські та міські території. Така різноманітність сцен створює додаткові труднощі для автоматичного виявлення транспортних засобів, зокрема через складний фон, змінну текстуру поверхні та різні рівні щільності об'єктів. Усього набір даних VEDAI містить 1268 зображень розміром  $1024 \times 1024$  пікселів із середнім значенням GSD 12.5 см. Крім того, доступна зменшена версія цього набору із GSD 25 см, отримана шляхом масштабування вихідних зображень. Відповідно, у роботі розрізняються:

- LCI (Large-size Color Images) — зображення з GSD 12.5 см;
- SCI (Small-size Color Images) — зображення з GSD 25 см.

Еталонні розмітки (GT) у наборі VEDAI надані у вигляді орієнтованих обмежувальних рамок для дев'яти типів транспортних засобів, при цьому класи car, pickup та van об'єднані в узагальнену категорію малих наземних транспортних засобів (small land vehicles).

Зазначені набори даних суттєво відрізняються між собою за просторовою роздільною здатністю (GSD), розмірами об'єктів, типами транспортних засобів, складністю фону та кількістю об'єктів на одному зображенні. Це дозволяє комплексно оцінити стабільність та адаптивність запропонованих методів до різних умов зйомки та сцен.

Експерименти на наборах даних DOTA, ITCVD та xView виконано переважно у якісному режимі з метою аналізу здатності моделей до перенесення знань між різними доменами даних. Кількісне оцінювання для цих наборів даних не проводилося через їх відносно пізню доступність, а також часткову неоднорідність і обмежену якість анотацій, що могло б призвести до викривлення результатів порівняльного аналізу.

Набір даних DOTA (Dataset for Object Detection in Aerial Images) [73] складається з 2806 зображень, отриманих з використанням різних сенсорів і платформ, зокрема сервісу Google Earth, супутника JL-1, а також супутника GF-2 Китайського центру супутникових даних та їх застосування (China

Centre for Resources Satellite Data and Application). Така різноманітність джерел зйомки забезпечує високу варіативність характеристик зображень і дозволяє оцінити здатність методів розпізнавання до узагальнення за умов зміни сенсорів та параметрів зйомки. Розміри зображень у наборі DOTA істотно варіюються – від  $800 \times 800$  до  $4000 \times 4000$  пікселів, при цьому значення просторової роздільної здатності на місцевості (GSD) задається індивідуально для кожного зображення і знаходиться в діапазоні від 10 до 100 см. Як показано на рис. 4.3, набір охоплює різноманітні типи сцен, включаючи міські, промислові, портові та сільськогосподарські території, а також характеризується широким діапазоном масштабів і довільними орієнтаціями об'єктів, що є типовим для аерокосмічних знімків надвисокої роздільної здатності.

Еталонні розмітки (GT) у наборі DOTA надані для 1869 зображень та охоплюють 15 категорій об'єктів. Анотації представлені у вигляді орієнтованих (oriented) та осьово-орієнтованих (axis-aligned) обмежувальних рамок, що робить цей набір особливо цінним для досліджень, спрямованих на розпізнавання об'єктів із довільною орієнтацією, зокрема транспортних засобів. Разом із тим, слід зазначити, що якість анотацій у наборі DOTA є неоднорідною. Зокрема, для категорій *small vehicle* та *large vehicle* спостерігається значна кількість пропущених розміток, що істотно ускладнює проведення коректної кількісної оцінки результатів детекції. У зв'язку з цим у межах даного дослідження експерименти з використанням набору DOTA мають переважно якісний характер та спрямовані на аналіз переносимості та узагальнювальної здатності запропонованих нейромережевих методів у складних сценах із багатомасштабними та орієнтованими об'єктами.

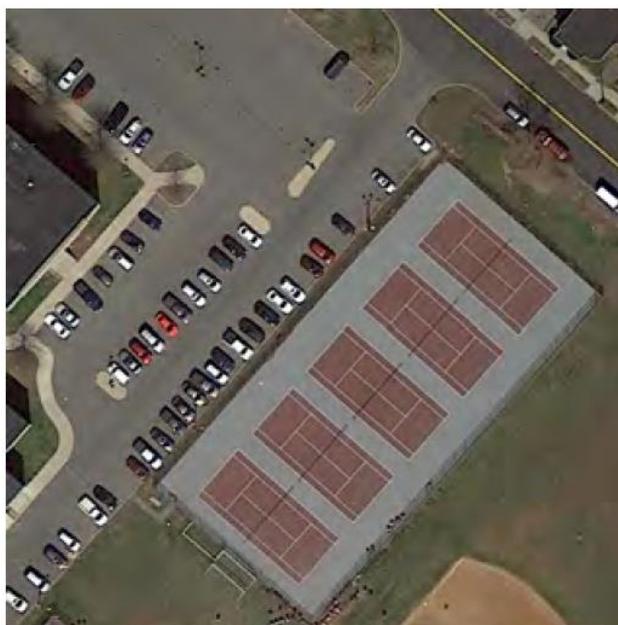


Рисунок 4.3 – Приклади набору даних аерофотознімків транспортних засобів DOTA [70]

Таким чином, набір даних DOTA використовується не стільки для формальної метричної оцінки, скільки для верифікації працездатності розроблених моделей в умовах реальних аерокосмічних зйомок, що характеризуються високою складністю фону, значною кількістю об'єктів та істотною різницею в їх просторово-геометричних параметрах.

Для дослідження впливу семантичної сегментації на якість виявлення транспортних засобів використано набір даних ISPRS 2D Semantic Labeling Challenge Potsdam, який є одним із найбільш поширених еталонів для задач семантичного маркування аерофотознімків. Зазначений набір даних дозволяє оцінити ефективність інтеграції контекстної інформації (дороги, забудова, рослинність) у процес детекції транспортних об'єктів.

#### **4.1.2. Формування власного набору даних**

Проведений у п.4.1.1 аналіз експериментальних результатів на відкритих наборах даних DLR 3K, VEDAI та DOTA показав, що, попри їхнє широке використання у задачах розпізнавання транспортних засобів на

аерокосмічних зображеннях, зазначені датасети мають низку обмежень. Зокрема, вони характеризуються обмеженою кількістю анотованих зображень для окремих класів, значною варіативністю просторової роздільної здатності, неоднорідними умовами зйомки та, у ряді випадків, недостатньою якістю розмітки, що особливо критично для малорозмірних транспортних об'єктів. Крім того, більшість відкритих наборів даних орієнтовані на загальні сценарії міської забудови та не повною мірою відображають специфіку низьковисотного аерофотознімання, характерного для локального моніторингу транспортної інфраструктури.

Отримані результати експериментів на існуючих наборах даних засвідчили, що узагальнювальна здатність нейромережових моделей істотно знижується при їх застосуванні до зображень, отриманих в інших умовах зйомки, з іншими параметрами сенсорів та геометрією спостереження (рис.4.4). У зв'язку з цим виникла об'єктивна необхідність формування власного спеціалізованого набору даних, який би відповідав реальним умовам застосування розробленого нейромережового методу та забезпечував коректну оцінку його ефективності.



а)

б)

Рисунок 4.4 – Розпізнавання транспортних засобів на наборах даних: а) DOTA; б) VEDAI [136]

Підготовка власного набору даних здійснювалася на основі аерокосмічних знімків високої роздільної здатності, отриманих камерою SONY DSC-WX220 у режимі низьковисотного моніторингу. Використання даного типу сенсора дозволило сформувати зображення з високим рівнем деталізації транспортних об'єктів, що є критично важливим для задач їх точного виявлення та локалізації. Сформований власний набір даних (рис.4.5) включає 1122 зображення (тайли), на яких розміщено 1351 анотацію об'єктів класу «vehicles». Таким чином, середня щільність становить близько 2,9 транспортних об'єкта на один фрагмент зображення, що відповідає реальним умовам аерофотознімання парковок і дорожніх ділянок із помірною концентрацією транспортних засобів.

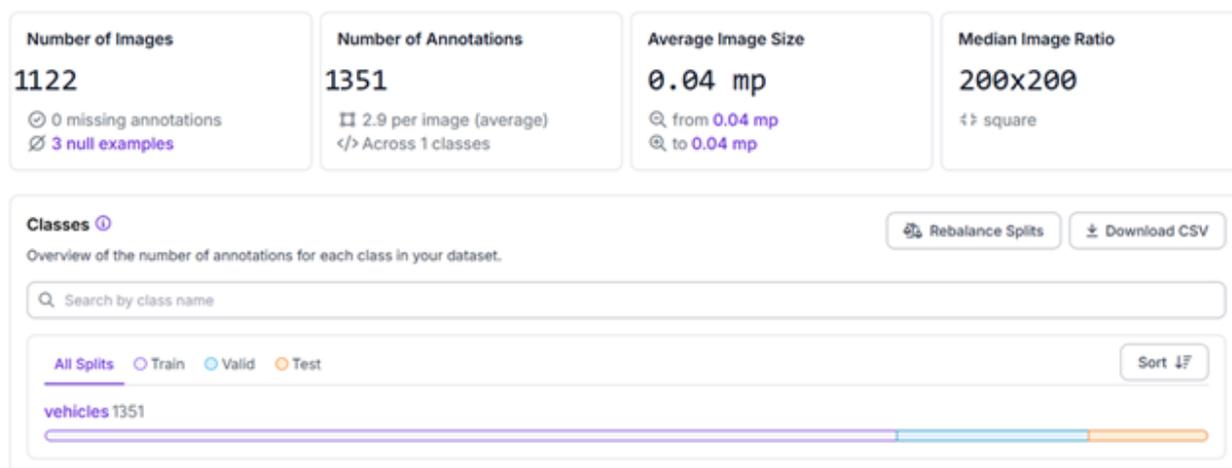


Рисунок 4.5 – Формування власного набору даних

Усі вхідні зображення мають фіксований розмір  $200 \times 200$  пікселів, що відповідає площі 0,04 Мп, та квадратне співвідношення сторін (square). Така уніфікація виключає геометричні спотворення об'єктів під час масштабування та забезпечує стабільність процесу навчання нейронної мережі незалежно від архітектурних особливостей моделі. Обраний розмір тайлів є компромісним та дозволяє, з одного боку, зберегти достатній локальний контекст для ідентифікації малих транспортних засобів, а з

іншого — зменшити обчислювальне навантаження та стабілізувати процес навчання моделі. На основі сформованого набору тайлів дані були розподілені на підмножини для навчання (70 %), валідації (15 %) та тестування (15 %), що забезпечило репрезентативність вибірки та коректність подальшої кількісної оцінки якості нейромережевого розпізнавання.

Анотація транспортних засобів виконувалася для кожного тайла окремо та передбачала точне виділення об'єктів із фіксацією їхнього просторового положення та орієнтації. На відміну від традиційних осьово-орієнтованих обмежувальних прямокутників (*axis-aligned bounding boxes*), у даному дослідженні використовувалися орієнтовані обмежувальні прямокутники (*Oriented Bounding Boxes*, *ОВВ*), що є більш адекватними для аерокосмічних зображень, де транспортні засоби можуть бути розташовані під довільними кутами відносно осей зображення. Кожна мітка містила координати центру об'єкта, його ширину, висоту та кут обертання, що дозволяло мінімізувати включення фонового оточення та підвищити точність навчання моделі, особливо для малорозмірних і щільно розташованих об'єктів.

Анотації зберігалися у текстовому форматі, де кожний рядок відповідав окремому транспортному засобу та був сумісний з компонентом орієнтованого детектування у структурі *Multi-Task Loss*. Усі мітки були організовані у відповідних директоріях *train/labels* та *valid/labels*, що забезпечило чітку структуру набору даних і можливість його безпосереднього використання під час навчання та валідації нейромережевої моделі.

Результати аналізу кількості анотованих об'єктів на зображеннях, представлені у вигляді гістограми, свідчать про збалансований характер сформованого набору даних. Найбільшу частку становлять тайли, що містять 2–3 транспортні засоби (43 зображення), що є типовим для сцен аерокосмічного моніторингу парковок та прилеглих дорожніх ділянок.

Значна кількість зображень з одиничними об'єктами (33 тайли) та з помірно підвищеною щільністю (4–5 об'єктів, 33 тайли) забезпечує варіативність просторових конфігурацій і дозволяє моделі навчатися як на простих, так і на складніших сценах.

Наявність меншої кількості зображень із високою щільністю об'єктів (6–9 транспортних засобів) відображає реальні умови зйомки та створює додаткові складні приклади, що є важливими для перевірки стійкості моделі до часткових перекриттів і близького розташування об'єктів. Окремо варто відзначити присутність фонових зображень без анотацій (3 тайли), які сприяють зниженню кількості хибнопозитивних спрацювань та покращують узагальнювальну здатність моделі.

На прикладах анотованих зображень продемонстровано практичну реалізацію орієнтованих обмежувальних рамок (Oriented Bounding Boxes, ОВВ). Візуальний аналіз результатів підтверджує високу щільність прилягання рамок до контурів транспортних засобів незалежно від їхньої орієнтації та напрямку руху. Це є особливо важливим для аерокосмічних зображень, де об'єкти можуть бути розміщені під довільними кутами відносно осей зображення.

На прикладах анотованих зображень (рис. 4.6) продемонстровано практичну реалізацію орієнтованих обмежувальних рамок. Візуальний аналіз підтверджує високу точність прилягання рамок до контурів транспортних засобів незалежно від їх орієнтації та напрямку руху, що є критично важливим для задач розпізнавання транспортних об'єктів на аерокосмічних зображеннях.



Рисунок 4.6 – Фрагмент створеного власного набору даних

### 4.1.3. Кількісні метрики

Для оцінювання ефективності нейромережових методів розпізнавання та сегментації транспортних об'єктів на аерокосмічних зображеннях у роботі застосовано комплекс кількісних метрик, що дозволяють всебічно охарактеризувати якість виявлення, локалізації та піксельної класифікації об'єктів. Вибір метрик зумовлений специфікою задач розпізнавання та семантичної сегментації, а також вимогами до об'єктивного порівняння різних нейромережових архітектур.

Для аналізу ефективності методів розпізнавання автомобілів, зокрема моделей класу YOLO, використано стандартні метрики оцінювання якості виявлення об'єктів: *Precision*, *Recall*, *F1 Score* та середню точність (*mAP*). Зазначені показники є базовими у задачах розпізнавання об'єктів і дозволяють оцінити баланс між точністю і повнотою розпізнавання.

Метрика *Precision* характеризує частку правильно виявлених об'єктів серед усіх об'єктів, які були класифіковані моделлю як позитивні, і визначається формулою (4.1) [130, 137]:

$$Precision = TP / (TP + FP), \quad (4.1)$$

де *TP* (true positive) – кількість правильно розпізнаних об'єктів, *FP* (false positive) – кількість хибно розпізнаних об'єктів.

Метрика *Recall* відображає здатність моделі виявляти всі наявні об'єкти заданого класу та обчислюється за формулою (4.2) [130, 137]:

$$Recall = TP / (TP + FN), \quad (4.2)$$

де *FN* (false negative) – кількість пропущених об'єктів.

Для узагальненої оцінки якості детекції використано *F1 Score*, який є гармонійним середнім між *Precision* та *Recall* і дозволяє оцінити збалансованість моделі за формулою (4.3) [130]:

$$F_1 Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}. \quad (4.3)$$

Середня точність *mAP* (mean Average Precision) застосовується як інтегральний показник якості розпізнавання об'єктів, що враховує зміну *Precision–Recall* при різних порогах впевненості та використовується для порівняння моделей у задачах багатокласового розпізнавання об'єктів.

Для оцінювання якості семантичної сегментації транспортних засобів на аерокосмічних зображеннях застосовано піксельні метрики, які дозволяють оцінити відповідність результатів сегментації еталонним розміткам. Однією з базових метрик є точність пікселів (Pixel Accuracy, *PA*),

що визначає частку правильно класифікованих пікселів відносно загальної кількості пікселів у зображенні. Розрахунок  $PA$  здійснюється за формулою (4.4) [137]:

$$PA = \frac{TP}{TP + FP + FN + TN}, \quad (4.4)$$

де  $TP$  – кількість правильно класифікованих позитивних пікселів,  $FP$  – кількість хибних позитивних пікселів,  $FN$  – кількість хибних негативних пікселів,  $TN$  – кількість правильно класифікованих негативних пікселів.

Точність пікселів дозволяє оцінити, наскільки добре модель справляється з класифікацією кожного окремого пікселя, що є важливим при сегментації на рівні детального зображення, наприклад, для точного виділення транспортних засобів на супутникових зображеннях.

Для узагальненого оцінювання якості сегментації по всіх класах використано середню точність (Mean Accuracy,  $MA$ ), яка відображає середнє значення точності класифікації для кожної категорії об'єктів. Розрахунок  $MA$  здійснюється за формулою (4.5) [137]:

$$MA = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FP_i + FN_i}, \quad (4.5)$$

де  $N$  – кількість класів у зображенні (наприклад, кількість категорій об'єктів, що сегментуються),  $TP_i$  – кількість правильно класифікованих пікселів класу  $i$ ,  $FP_i$  – кількість хибних позитивних пікселів для класу  $i$ ,  $FN_i$  – кількість хибних негативних пікселів для класу  $i$ .

Середня точність дозволяє отримати узагальнений показник ефективності сегментації по всіх класах і забезпечує уявлення про те, наскільки добре модель працює з різними типами об'єктів на зображеннях.

Ключовим показником якості семантичної сегментації у роботі є середнє перетинання по об'єднанню (mean Intersection over Union,  $mIoU$ ), яке широко використовується у задачах комп'ютерного зору. Метрика  $mIoU$  дозволяє оцінити ступінь перекриття між передбаченими сегментами та

еталонними розмітками з урахуванням усіх класів і визначається за формулою (4.6) [137]:

$$mIoU = \frac{(X_{ii} / T_i + \sum_{j=1}^N (X_{ji} - X_{ii}))}{N}, \quad (4.6)$$

де  $N$  являє собою кількість категорій пікселів зображення;  $T_i$  – загальна кількість пікселів класу  $i$ ;  $X_{ii}$  – загальна кількість пікселів з фактичним типом  $i$  та типом  $i$  передбачення;  $X_{ji}$  – загальна кількість пікселів з фактичним типом  $i$  та типом передбачення  $j$ .

Метрика  $mIoU$  є одним з найкращих індексів для комплексної оцінки результатів сегментації, оскільки дозволяє врахувати точність як для позитивних, так і для негативних пікселів, а також забезпечує збалансоване оцінювання на основі всіх класів. Оскільки цей показник враховує як перетинання, так і об'єднання між сегментованими класами, він дає більш об'єктивну оцінку якості моделі, що особливо важливо для задач, де є кілька класів об'єктів (наприклад, транспортні засоби, дороги та інші елементи на зображеннях).

#### **4.2. Експериментальне дослідження ефективності методів розпізнавання автомобілів на аерокосмічних зображеннях**

Після визначення характеристик наборів даних та кількісних метрик оцінки ефективності у п.4.1, наступним кроком стало безпосереднє проведення експериментальних досліджень для порівняльної оцінки запропонованих та класичних методів розпізнавання транспортних засобів на аерокосмічних зображеннях. Це дозволило порівняти ефективність алгоритмів виявлення об'єктів із різними підходами до обробки зображень та визначити оптимальні методи для подальшого впровадження.

У ході експериментальних досліджень було використано аерокосмічне зображення, отримане з відкритого ресурсу OpenAerialMap

[138] (рис. 4.7а). На знімку представлено фрагмент території з чітко візуалізованими автомобілями, які обрано як об'єкти для подальшого аналізу та виявлення. Для оцінювання ефективності класичних методів виділення контурів у задачі розпізнавання автомобілів застосовано алгоритми Канні (рис. 4.7б), Собеля (рис. 4.7в), Лапласіана (рис. 4.7г), Прюїтта (рис. 4.7г) та Шарра (рис. 4.7д). Реалізація зазначених методів здійснювалася з використанням мови програмування Python [25].

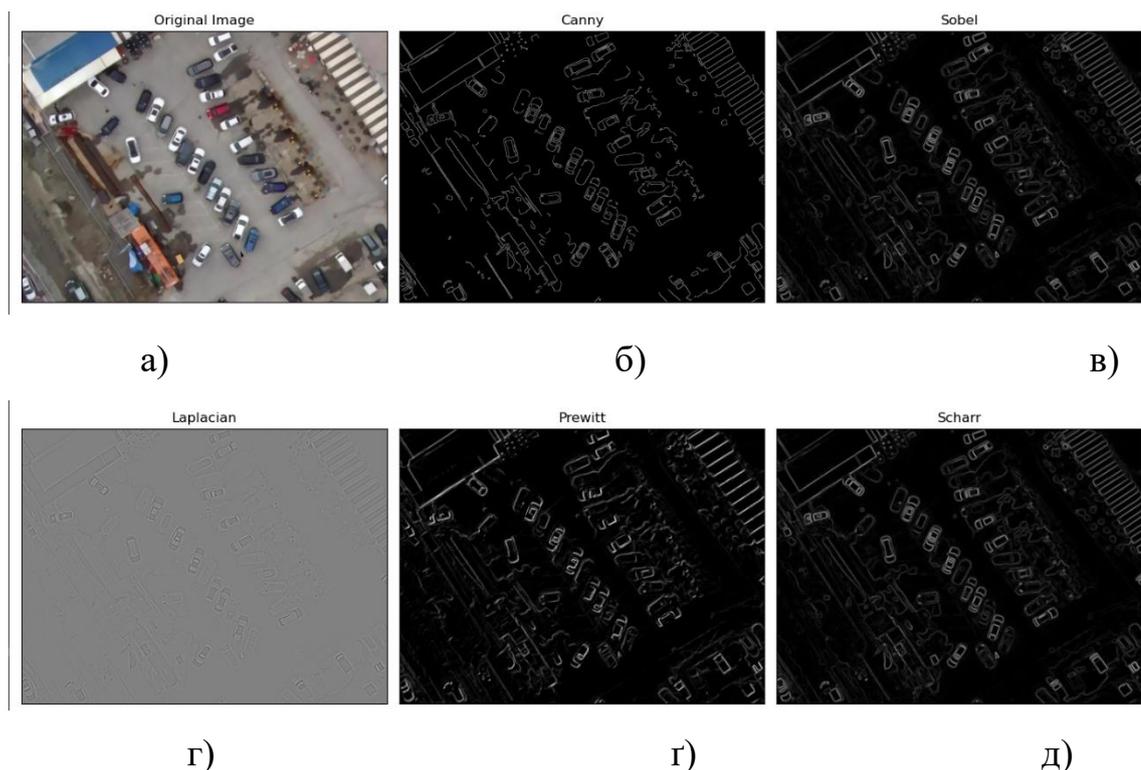


Рисунок 4.7 – Порівняльний аналіз результатів виявлення об'єктів на фрагменті аерокосмічного знімку: а) первинний знімок; б) метод Канні; в) метод Собеля; г) метод Лапласіана; г) метод Прюїтта; д) метод Шарра [25]

На рис. 4.8 та 4.9 наведено результати накладання виділених контурів на вихідне зображення, що дозволяє виконати візуальну оцінку якості розпізнавання автомобілів для кожного з розглянутих підходів. На рис. 4.10 представлено результат роботи нейромережевої моделі, розробленої для автоматичного виявлення автомобілів на аерокосмічних знімках. Модель здійснює локалізацію об'єктів за допомогою обмежувальних прямокутників, кожен з яких супроводжується значенням рівня довіри, що

відображає ймовірність коректного виявлення. Нейронна мережа була попередньо навчена на спеціалізованому наборі даних для ідентифікації автомобілів на зображеннях, отриманих з безпілотних літальних апаратів. Використовувана модель реалізована на основі архітектури Roboflow 2.0 Object Detection (Fast) та навчена на датасеті COCO (Common Objects in Context) [139] із застосуванням архітектури YOLOv6n [140].

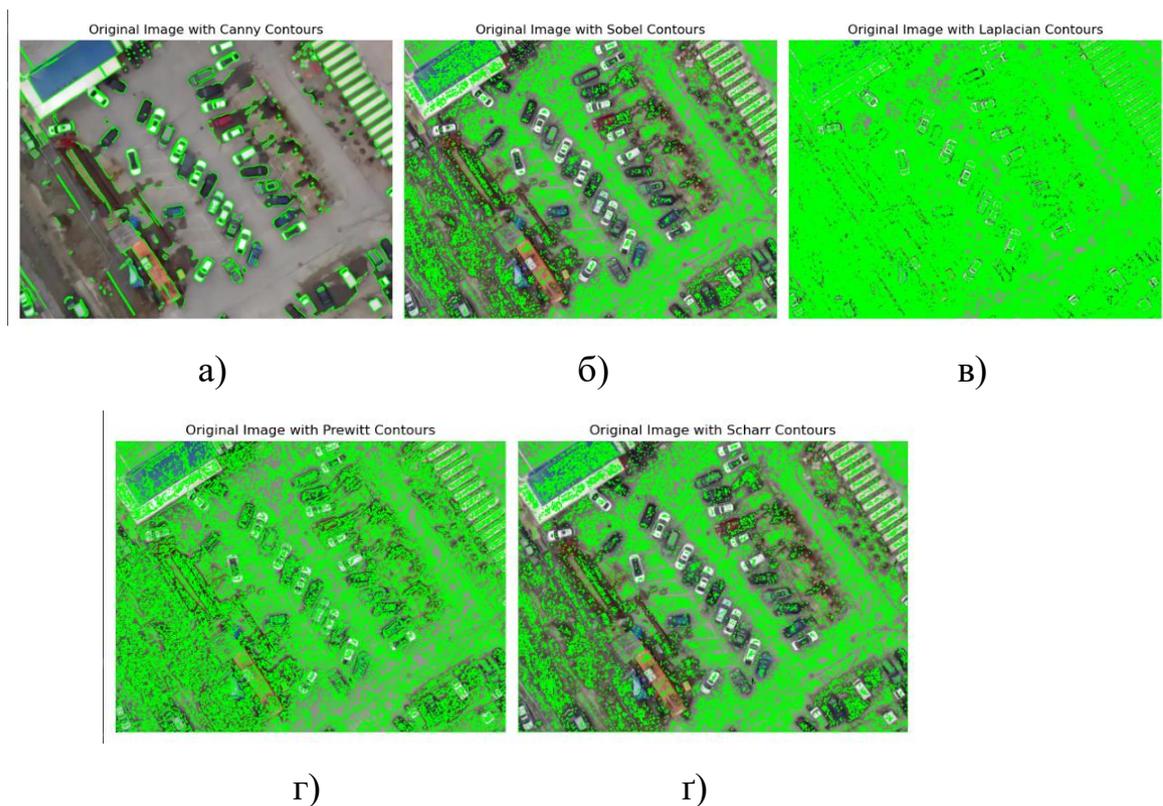


Рисунок 4.8 – Результат розпізнавання автомобілів на аерозображенні: а) метод Канні; б) метод Собеля; в) метод Лапласіана; г) метод Прюїтта; г) метод Шарра [25]

Візуальний аналіз результатів, отриманих за допомогою методу Канні (рис. 4.7б та рис. 4.8), свідчить про його здатність формувати чітко виражені контури на зображенні. Водночас алгоритм виділяє не лише автомобілі, а й інші елементи сцени, зокрема дорожню розмітку, бордюри та дрібні деталі. Використання подвійного порогового значення дозволяє контролювати рівень шумів, однак метод не є селективним щодо конкретного класу об'єктів. За високих значень порогів (рис. 4.9б) частина автомобілів не

виділяється, що зумовлено орієнтацією алгоритму на найбільш контрастні границі. У таких умовах об'єкти з менш вираженими межами, а також автомобілі, розташовані в затінених або складних за освітленням ділянках, можуть залишатися невиявленими. Натомість зменшення порогових значень (рис. 4.9а) призводить до появи значної кількості сторонніх контурів, зокрема текстур поверхні та тіней, що ускладнює подальшу сегментацію та аналіз.

Метод Собеля (рис. 4.7в та рис. 4.8б) дозволяє ідентифікувати автомобілі, проте сформовані контури є менш чіткими та частково розмитими порівняно з результатами Канні. Спостерігається підвищений рівень шуму, особливо в області дрібних деталей. Алгоритм ефективно реагує на горизонтальні та вертикальні границі, однак його застосування є обмеженим у задачах, де вимагається висока точність окреслення контурів. Згладжування країв зменшує кількість дрібних шумових елементів, проте одночасно знижує деталізацію.

Результати, отримані за допомогою методу Лапласіана (рис. 4.7г та рис. 4.8в), характеризуються ще більшою розмитістю контурів і підвищеною чутливістю до шуму, що негативно впливає на точність виявлення автомобілів, особливо на складних або текстурованих фонах. Зазначений метод доцільніше застосовувати для загального виділення контурів, а не для ідентифікації конкретних об'єктів.

Метод Прюїтта (рис. 4.7г та рис. 4.8г) демонструє результати, близькі до Собеля, однак із вищим рівнем шуму. Контури автомобілів формуються недостатньо чітко, що ускладнює їх відокремлення від фонових об'єктів. Через високу чутливість до завад даний метод є малоефективним для задач точного розпізнавання автомобілів.

Алгоритм Шарра (рис. 4.7д та рис. 4.8г) забезпечує дещо детальніше виділення контурів порівняно з Собеля і Прюїтта, однак також супроводжується значним рівнем шуму. Незважаючи на покращену якість

контурів, за показниками чіткості та точності Шарра поступається методу Канні [25].

Як показано на рис. 4.10, неймережева модель YOLOv6n забезпечує коректне виявлення автомобілів на аерокосмічному знімку, демонструючи високий рівень точності та стабільність локалізації, що підтверджує її ефективність у задачах автоматизованого розпізнавання.



Рисунок 4.9 – Результат розпізнавання автомобілів на аерозображенні методом Канні: а) порогові значення 80/150; б) порогові значення 200/250 [25]

На рис. 4.11 наведено результати порівняльного аналізу значень метрик Precision, Recall, F1 Score та mAP для досліджуваних методів розпізнавання автомобілів на аерокосмічних зображеннях. Для оцінювання ефективності було обрано метрики Precision, Recall і F1 Score, а також середню точність (mAP) [46] для неймережевої моделі YOLOv6n, оскільки вони комплексно відображають як точність виявлення, так і повноту ідентифікації об'єктів.

Для методу Канні отримано низькі значення всіх метрик (Precision – 0,038; Recall – 0,064; F1 Score – 0,048), що свідчить про його обмежену придатність до кількісного розпізнавання автомобілів, незважаючи на візуально чіткі контури. Алгоритм Собеля характеризується дуже низьким

Precision (0,031) при високому Recall (0,968), що вказує на значну кількість хибнопозитивних результатів і підтверджується низьким значенням F1 Score (0,060). Подібна ситуація спостерігається і для методу Шарра [25].

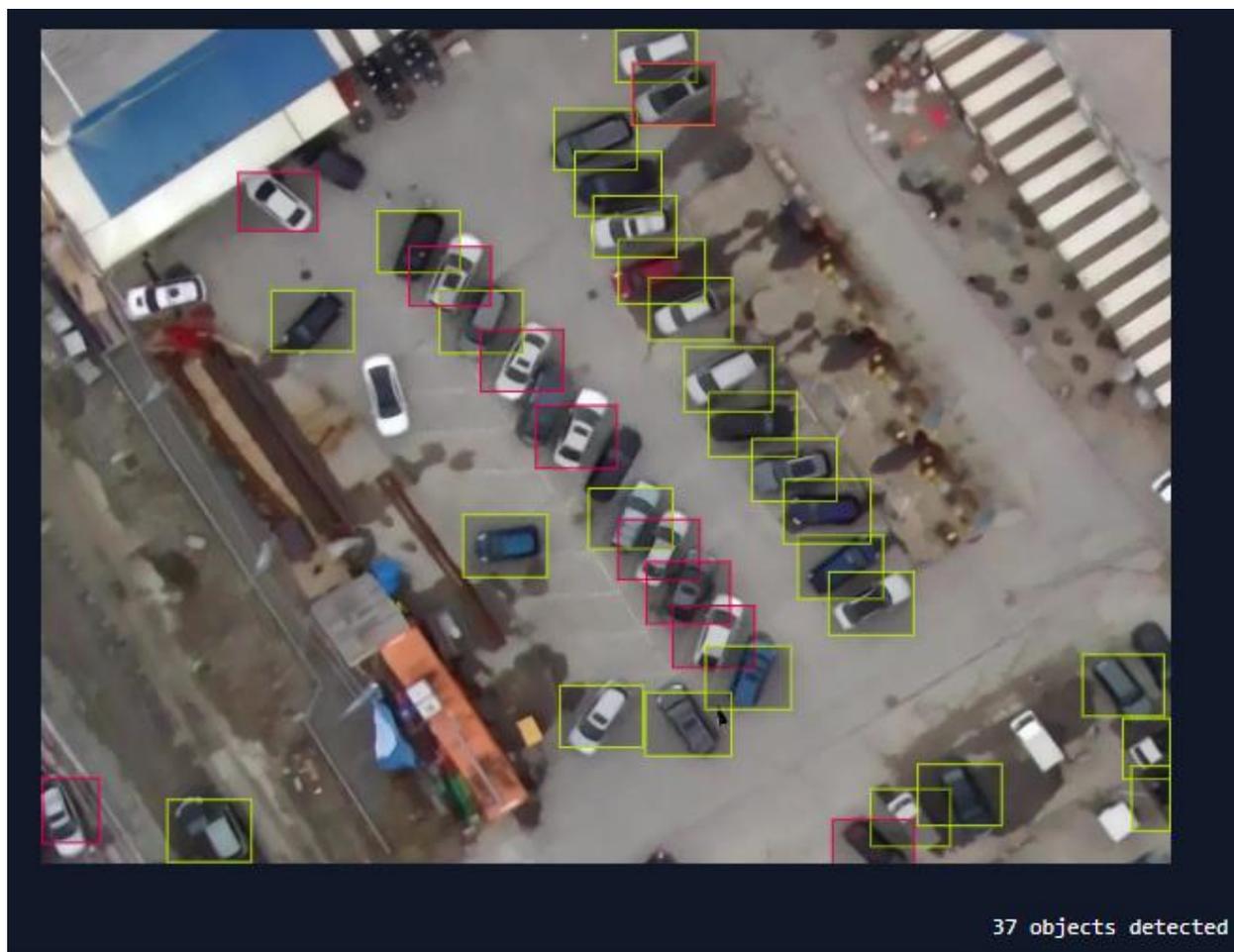


Рисунок 4.10 – Розпізнавання автомобілів на аерокосмічному знімку за допомогою неймережі YOLOv6 [25]

На рис. 4.11 наведено результати порівняння метрик Precision, Recall, F1 Score та mAP для кожного з досліджуваних методів розпізнавання автомобілів на аерокосмічних зображеннях.

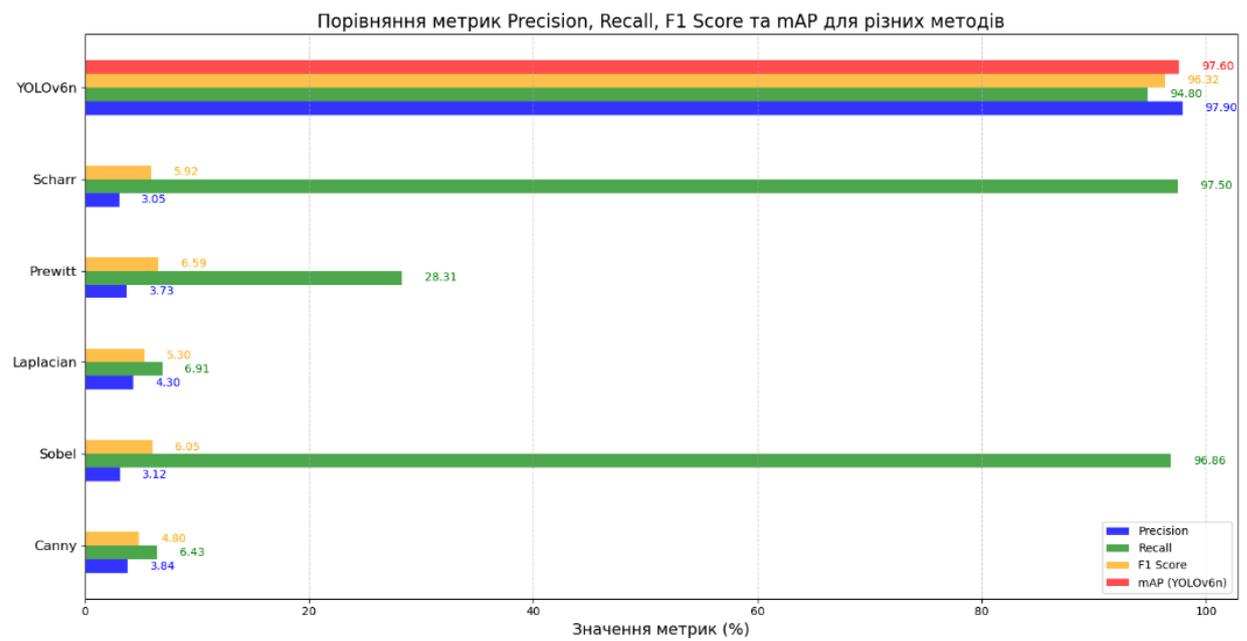


Рисунок 4.11 – Результати метрик Precision, Recall, F1 Score та mAP [25]

Метод Лапласіана демонструє дещо кращі значення Precision (0,043) та F1 Score (0,053), однак низький Recall (0,069) свідчить про нездатність алгоритму виявляти більшість об'єктів. Для Прюїтта спостерігається відносно більш збалансований результат (Precision – 0,037; Recall – 0,283; F1 Score – 0,066), проте рівень шуму залишається значним.

На відміну від класичних методів, нейромережева модель YOLOv6n демонструє високі значення всіх показників: Precision – 97,9%, Recall – 94,8%, F1 Score – 96,32%, mAP – 97,6%. Отримані результати свідчать про високу точність і надійність моделі в задачах розпізнавання автомобілів, із мінімальним рівнем хибнопозитивних та хибнонегативних спрацювань.

Узагальнюючи результати дослідження, можна стверджувати, що традиційні методи виділення контурів мають суттєві обмеження у задачах розпізнавання автомобілів на аерокосмічних знімках, тоді як використання нейромережеских моделей, зокрема YOLOv6n, забезпечує значно вищу ефективність та практичну придатність для автоматизованого аналізу зображень [25].

### **4.3. Експериментальна оцінка нейромережевого підходу до семантичної сегментації транспортних засобів на зображеннях надвисокого просторового розрізнення**

Після оцінки ефективності класичних методів виділення контурів для розпізнавання автомобілів, розглянутих в п.1.5 виявлено їх обмеження при обробці складних сцен із високою щільністю транспортних засобів, частковим перекриттям об'єктів, а також різноманітними умовами освітлення та тінями. Для подолання цих обмежень і підвищення точності локалізації та ідентифікації транспортних засобів доцільним є застосування нейромережевих методів семантичної сегментації, які здатні автоматично виділяти складні ознаки об'єктів, враховувати контекст сцени та забезпечувати інваріантність до геометричних перетворень.

Для проведення експериментів було використано набір даних, що складається з зображень, отриманих з безпілотних літальних апаратів на висоті 15 см, та відповідних еталонних масок сегментації. Еталонні маски були створені вручну експертами, що забезпечило високу якість анотацій. Набір даних було розділено на три частини: навчальну, валідаційну та тестову. Навчальна вибірка використовувалася для навчання моделі, валідаційна – для налаштування гіперпараметрів та моніторингу процесу навчання, а тестова – для оцінки узагальнюючої здатності моделі. Навчання проводилося до досягнення стабілізації значення функції втрат на валідаційній вибірці. Для покращення узагальнюючої здатності моделі було застосовано аугментацію даних, включаючи обертання, віддзеркалення та масштабування зображень.

Для оцінки якості сегментації було використано три метрики: точність пікселів ( $PA$ ), середня точність ( $MA$ ) та середнє перетин по об'єднанню ( $mIoU$ ). У експериментальних дослідженнях використовуються вищеописані метрики для порівняння ефективності різних моделей сегментації, а також для оцінки результатів, отриманих за допомогою запропонованої

архітектури DeepLab + ResNet. Зокрема, в контексті досліджень на зображеннях надвисокої просторової роздільної здатності, оцінка за допомогою RA, MA та mIoU дає змогу детально проаналізувати якість сегментації транспортних засобів.

Результати експериментального дослідження демонструють перевагу запропонованого нейромережевого підходу над існуючими підходами сегментації за ключовими метриками оцінки якості. Аналіз показників Loss, Accuracy, RA та mIoU підтверджує ефективність навчання та високу точність сегментації.

Отримані результати, представлені в таблиці 4.2, дозволяють оцінити ефективність різних методів сегментації на основі метрик True Positive (TP), True Negative (TN), False Positive (FP) та False Negative (FN). Запропонований метод демонструє найвищі значення TP (8600) та TN (9700), що свідчить про його здатність точно визначати цільові об'єкти та правильно класифікувати фон. Крім того, запропонований метод має найменші значення FP (900) та FN (600), що вказує на знижену кількість хибно позитивних і хибно негативних класифікацій. Це підтверджує його високу точність у виявленні цільових об'єктів при мінімізації помилок сегментації. У порівнянні з іншими моделями, такими як DeepLab v3, U-Net та SegNet, запропонований метод показує кращий баланс між правильними та хибними класифікаціями, що робить його ефективним для семантичної сегментації транспортних засобів.

Таблиця 4.2 – Кількість правильно та неправильно класифікованих пікселів для різних моделей

Модель	TP	TN	FP	FN
DeepLab v3	8200	9500	1200	850
U-Net	8100	9400	1300	900
SegNet	7000	9200	1600	1200
FCN8s	6800	9100	1700	1400
ENet	6600	8900	2000	1600
Запропонована	8600	9700	900	600

Метрика Loss (функція втрат) відображає розбіжність між прогнозованими та фактичними значеннями, тобто, чим нижче значення Loss, тим краще модель відповідає навчальним даним. Згідно з таблицею 4.3, запропонований метод демонструє найнижчі значення Loss як на етапі навчання (0.3), так і на етапі валідації (0.4). Це свідчить про те, що запропонована архітектура нейронної мережі ефективно мінімізує похибку під час навчання та добре узагальнює отримані знання на нових, невідомих даних (валідаційна вибірка). Низька різниця між значеннями Loss на навчальній та валідаційній вибірках вказує на стабільність процесу навчання та відсутність перенавчання.

Таблиця 4.3 – Значення функції втрат (Loss) на навчальній та перевіірочній вибірках

Модель	Епохи	Loss (навчання)	Loss (перевіірочної)
DeepLab v3	100	0.7	0.8
U-Net	100	0.8	0.9
SegNet	100	0.4	0.5
FCN8s	100	0.5	0.6
ENet	100	0.6	0.7
Запропонована	100	0.3	0.4

Метрика Ассурасу (точність) вимірює відсоток правильно класифікованих пікселів і є важливим показником ефективності моделі в задачах сегментації. Валідаційна точність відображає здатність моделі узагальнювати набуті знання на нових, що є важливим для оцінки її узагальнюючої здатності та стійкості до перенавчання. Згідно з Таблицею 4.4, запропонований метод демонструє найвищу валідаційну точність (90%), що свідчить про його здатність ефективно класифікувати пікселі на нових зображеннях. Він також досягає високої точності на тренувальній вибірці (95%), що вказує на хорошу збіжність моделі. DeepLab v3 (88%) та U-Net (85%) показують дещо нижчі результати валідаційної точності, але все ще демонструють відносно ефективне узагальнення. Моделі SegNet (80%),

FCN8s (65%) та ENet (70%) мають значно нижчі значення точності перевірочних даних, що свідчить про обмежену здатність до узагальнення та вищу похибку на перевірочних даних.

Таблиця 4.4 – Порівняльна оцінка точності на навчальній та перевірочній вибірках для різних моделей

Модель	Епохи	Accuracy (навчальної)	Accuracy (перевірочної)
DeepLab v3	100	0.91	0.88
U-Net	100	0.9	0.85
SegNet	100	0.85	0.8
FCN8s	100	0.8	0.65
ENet	100	0.75	0.7
Запропонована	100	0.95	0.9

Як показано в Таблиці 4.5, запропонований метод досягає найвищого значення Pixel Accuracy (95.1%), що перевищує результати всіх інших розглянутих моделей, включаючи PSANet (94.8%), DANet (94.6%) та OCNNet (92.1%). Висока точність PA свідчить про здатність моделі ефективно ідентифікувати об'єкти на зображенні, мінімізуючи як помилки класифікації фонового шуму.

Таблиця 4.5 – Значення метрики PA

Модель	PA (%)
DeepLab v3	91.8
U-Net	90.1
SegNet	81.2
FCN8s	86.4
ENet	74.8
Proposed method	95.1

Метрика mIoU є одним із ключових показників для оцінки якості сегментації зображень (табл.4.6). Вона визначає ступінь відповідності між передбаченими та реальними сегментами об'єктів, обчислюючи відношення площі їхнього перетину до площі їхнього об'єднання. Високе

значення mIoU свідчить про здатність моделі точно ідентифікувати межі об'єктів, зменшуючи кількість неправильно класифікованих пікселів для забезпечення високої точності сегментації. Запропонований метод досяг 82.3%, що значно перевищує показники DeepLab v3 (74.0%), ENet (70.8%) та U-Net (73.3%). При цьому SegNet та FCN8s мають однакове значення (56.7%), що вказує на їхню обмежену здатність до точного розділення об'єктів.

Результати експериментів показують, що запропонований метод демонструє високу ефективність у сегментації транспортних засобів на зображеннях з БПЛА, досягаючи найкращих результатів за метриками РА та mIoU, а також низького значення Loss. Це свідчить про його здатність точно класифікувати пікселі та сегментувати об'єкти.

Високі значення РА та mIoU, досягнуті запропонованим методом, можуть бути пояснені використанням багатомасштабних ознак та транспонованих згорток, що дозволяє ефективно виявляти та сегментувати об'єкти різних розмірів та форм. Низьке значення Loss свідчить про ефективне навчання моделі.

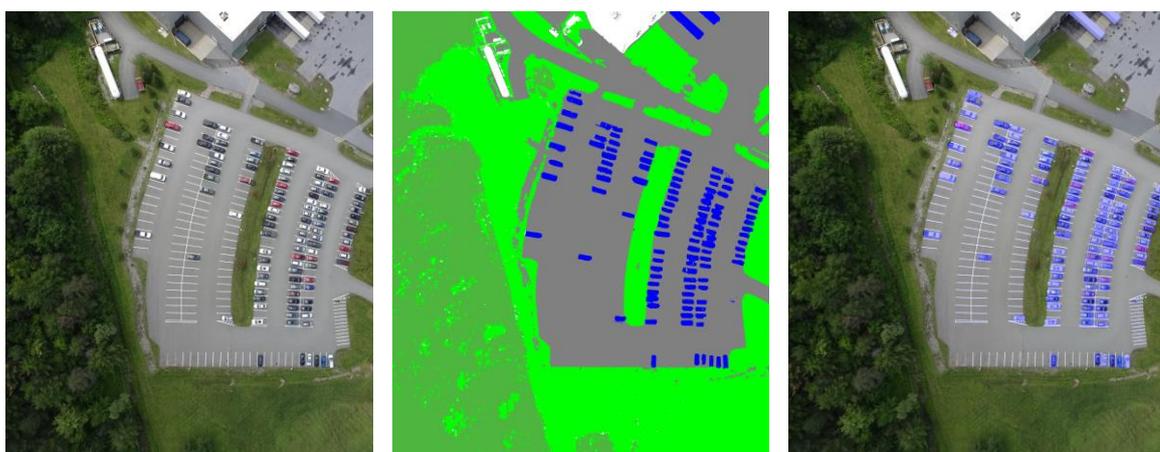
Таблиця 4.6 – Значення метрики mIoU

Модель	mIoU (%)
DeepLab v3	74.0
U-Net	73.3
SegNet	56.7
FCN8s	56.7
ENet	70.8
Запропонована	82.3

Візуальний аналіз отриманих результатів підтверджує ефективність запропонованого методу у задачі семантичної сегментації транспортних засобів. Як показано на рисунку 4.12, запропонований метод забезпечує чітке та точне виявлення транспортних засобів, що підтверджується якістю бінарної маски та кольорового сегментованого зображення.

Зокрема, транспортні засоби коректно виділені на бінарній масці без значних пропусків чи хибних позитивних сегментацій. Крім того, сегментоване зображення демонструє високу точність розділення класів об'єктів, таких як дороги, зелені насадження та будівлі. Важливим фактором є те, що запропонований метод ефективно розрізняє об'єкти зі схожими спектральними характеристиками, що часто є проблемою для традиційних підходів.

Порівняно з існуючими моделями, запропонований метод демонструє краще збереження контурів об'єктів та мінімізацію шуму у сегментації. Це особливо важливо для застосувань, що вимагають високої деталізації результатів, зокрема для моніторингу транспортного потоку, аналізу паркувальних зон та міського планування.



Legend: ■ - trees; ■ - grass; ■ - roads; ■ - vehicles;  - unclassified.

а)

б)

в)

Рисунок 4.12 – Зображення з БПЛА: а) оригінальний набір даних; б) результат застосування запропонованого методу на основі нейронної мережі; в) результат сегментації класу «транспортні засоби» на оригінальному зображенні з використанням запропонованого методу

Результати експериментів показують, що запропонований метод демонструє високу ефективність у сегментації транспортних засобів на зображеннях з БПЛА, досягаючи найкращих результатів за метриками  $PA$  та  $mIoU$ , а також низького значення  $Loss$ . Це свідчить про його здатність точно класифікувати пікселі та сегментувати об'єкти.

Високі значення  $PA$  та  $mIoU$ , досягнуті запропонованим методом, можуть бути пояснені використанням багатомасштабних ознак та транспонованих згорток, що дозволяє ефективно виявляти та сегментувати об'єкти різних розмірів та форм. Низьке значення  $Loss$  свідчить про ефективне навчання моделі.

#### **4.4. Дослідження ефективності нейромережевого розпізнавання повітряних транспортних засобів з урахуванням просторово-геометричних ознак**

Після проведення експериментальної оцінки нейромережевого підходу до семантичної сегментації автомобілів на надвисокороздільних зображеннях доцільним є розширення досліджень на інший клас об'єктів – повітряні транспортні засоби. Вони характеризуються специфічними просторово-геометричними ознаками, зокрема великою варіативністю розмірів, форм та орієнтації на зображенні. Для перевірки ефективності запропонованого нейромережевого підходу було використано набір даних  $HRPlanesv2$ , що містить високоякісні аерокосмічні зображення літаків та вертольотів із відповідними анотаціями.

У рамках досліджень виконано порівняння різних методів виявлення контурів:

1. Виявлення країв за методом Канні. Багатоступеневий алгоритм, призначений для надійного вилучення країв на зображеннях. Він забезпечує баланс між високою швидкістю виявлення країв, точним визначенням їхнього розташування та мінімальною кількістю помилкових

спрацьовувань. Алгоритм використовує низку кроків, включаючи зменшення шуму за допомогою гауссового фільтрування, обчислення градієнтів зображення, придушення не максимальних значень для витончення країв та гістерезисне порогове значення для надійного вибору країв.

2. Алгоритм відстеження меж для виявлення контурів.

3. Лапласіан. Математичний оператор, що використовується в обробці зображень для виявлення країв та підвищення чіткості зображення. Він обчислює другу похідну інтенсивності зображення, виділяючи області з швидкими змінами інтенсивності, які часто відповідають краям. Позитивні значення Лапласіана позначають області, де інтенсивність світлішає до центру (опуклі області), а негативні значення позначають області, де інтенсивність темнішає (увігнуті області). Нульове значення позначає плоскі області або краї, де напрямок зміни інтенсивності змінюється.

4. Фільтр Собела. Основний інструмент, що використовується для виявлення країв у обробці зображень. Він апроксимує градієнт зображення, представляючи напрямок і величину зміни інтенсивності в зображенні. Фільтр використовує два невеликих ядра, одне призначене для виявлення горизонтальних країв, а інше – для вертикальних, щоб оцінити ці градієнти. Фільтр Собела відіграє вирішальну роль у виявленні країв і вилученні ознак і служить будівельним блоком для алгоритму виявлення країв Канні.

5. Послідовне застосування операторів Канні та Лапласа.

Оскільки YOLO вимірює точність моделі з точки зору її здатності точно ідентифікувати та локалізувати об'єкти, що цікавлять, в зображенні, було використано *mAP*. *mAP* – це комбінація значень точності та відтворення, обчислених за декількома порогами достовірності, які також називаються порогами перетину над об'єднанням. Зміна порогу IoU призведе до різних значень істинних позитивних та хибних позитивних результатів.

У таблиці 4.7 наведено показники точності, досягнуті моделлю виявлення об'єктів на наборах даних, які пройшли різні методи попередньої обробки. Представлені показники дають уявлення про ефективність різних методів попередньої обробки в покращенні здатності моделі виявляти об'єкти в наборах даних.

Таблиця 4.7 – Показники точності виявлення літальних апаратів [71]

Pre-processing type	Precision	Recall	mAP50	mAP50-95
None	0.990522	<b>0.981679</b>	0.991079	0.831246
Канні	0.902949	0.782908	0.859798	0.592251
Border-following	<b>0.996403</b>	0.975793	<b>0.994272</b>	<b>0.863925</b>
Лапласіана	0.951500	0.885149	0.958168	0.713322
Собеля	0.965968	0.958904	0.982615	0.809901
Канні + Лапласіана	0.948966	0.827006	0.897425	0.654177

Значення показника Recall і другий за величиною показник mAP50-95. Це свідчить про те, що необроблені зображення вже містять достатню інформацію для того, щоб модель могла ідентифікувати значну частину присутніх літаків. Про це свідчить її чудова продуктивність за показниками Precision, mAP50 і mAP50-95. Результати свідчать про те, що витягнуті контури значно покращують здатність моделі точно виявляти і відрізняти літаки від елементів фону. Отже, ця техніка була обрана для запропонованого методу.

Техніки попередньої обробки оператором Лапласа та фільтром Собеля продемонстрували непогану продуктивність, але значно нижчу, ніж оригінальний набір даних та виявлення контурів, що слідує за межами. Однак ці методи призвели до значної втрати інформації порівняно з оригінальними зображеннями. Це свідчить про те, що, хоча вони можуть

підкреслювати потенційні краї, вони можуть вносити артефакти або усунути важливі деталі, необхідні для точного виявлення об'єктів. Алгоритм виявлення країв Канні показав найнижчу ефективність за всіма показниками. Це можна пояснити наявністю шуму або невідповідностей у супутникових зображеннях з високою роздільною здатністю. Комбінований підхід Канні-Лапласіана дав середні результати, що знаходяться між ефективністю попередньої обробки за допомогою одного Канні та одного Лапласіана. Хоча він перевершив Канні, він залишився менш ефективним, ніж попередня обробка за допомогою Лапласіана. Це свідчить про те, що комбінований підхід, можливо, не дозволив ефективно використати переваги обох технік, потенційно вводячи надлишкову або суперечливу інформацію для моделі. Ці висновки підкреслюють важливість вибору відповідних технік попередньої обробки для конкретних наборів даних зображень і завдань. Хоча виявлення країв може покращити вилучення ознак, вибір методу повинен забезпечувати баланс між окресленням країв і збереженням інформації для досягнення оптимальної продуктивності в моделях глибокого навчання для розпізнавання об'єктів.

Матриця помилок, представлена на рис. 4.13, ілюструє продуктивність моделей класифікації, відображаючи кількість правильно та неправильно класифікованих об'єктів у кожній категорії. Результати, показані на рис. 4.13, демонструють порівняння матриць плутанини для двох підходів: використання оригінального набору даних (рис. 4.13а) та модифікованого набору даних із використанням запропонованої технології (рис. 4.13б). В оригінальному наборі даних кількість правильно класифікованих об'єктів у класі «літак» становила 2521, з 71 помилковим позитивним результатом та 34 помилковими негативними результатами. Після впровадження запропонованого методу кількість помилкових позитивних результатів знизилася до 31, що свідчить про підвищення точності. Однак кількість помилкових негативних результатів збільшилася до 45, що означає певне зниження повноти. Таким чином, запропонована технологія підвищує

точність класифікації об'єктів класу «літак», але це досягається за рахунок незначного зниження індексу повноти.

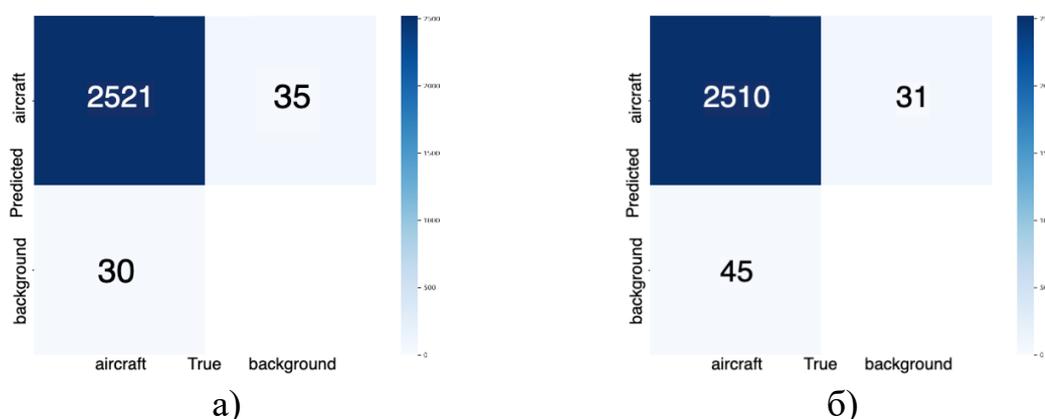
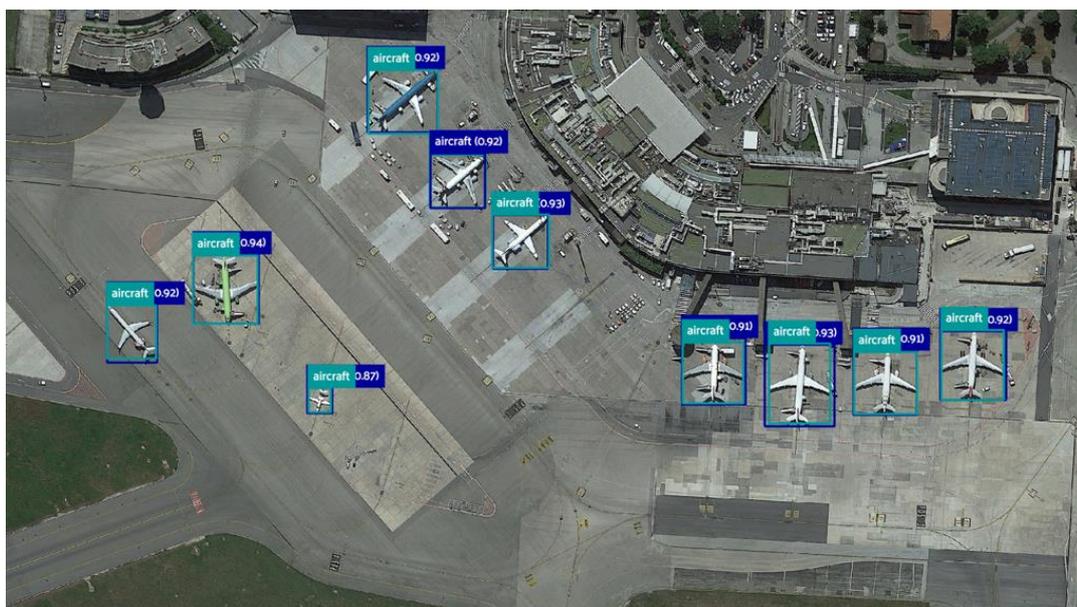


Рисунок 4.13 – Матриця помилок для: а) оригінального набору даних; б) набору даних із застосуванням запропонованою технологією

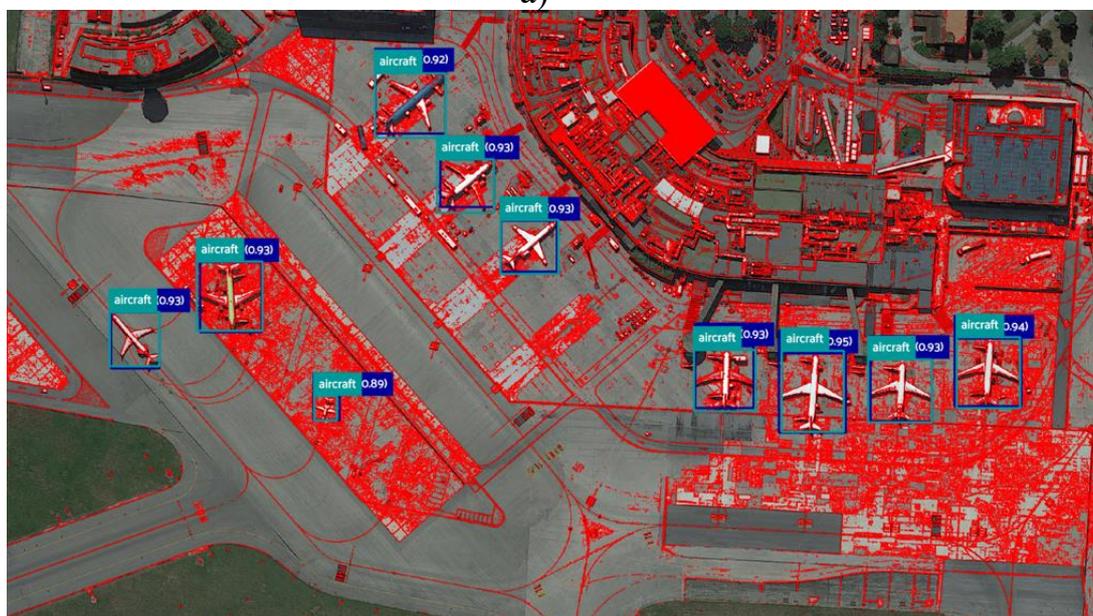
На рис. 4.14 показано результати виявлення для вихідного набору даних без попередньої обробки (рис.4.14а) та набору даних після застосування методу виявлення контурів, що слідують за межами (рис.4.14б). На обох зображеннях блакитні обмежувальні рамки позначають анотації істинних значень, а сині обмежувальні рамки показують виявлення, зроблені відповідною моделлю. Дослідження різних методів виявлення країв дало цінні результати. Найефективнішим виявився алгоритм відстеження меж, реалізований в OpenCV.

Цікаво, що комбінований підхід Канні-Лапласіана виявився перспективним, що свідчить про можливість подальшої оптимізації в майбутніх дослідженнях з метою вивчення потенційної синергії між різними методами. Хоча запропонований метод демонструє очевидний прогрес у виявленні літаків за допомогою супутникових знімків високої роздільної здатності, існують певні обмеження, які слід враховувати. Оцінка проводилася з використанням єдиного набору даних (HRPlanesV2). Крім того, дослідження альтернативних моделей глибокого навчання з архітектурами, спеціально розробленими для завдань розпізнавання

зображень з високою роздільною здатністю, може потенційно привести до подальших поліпшень. Крім того, дослідження ще більш вдосконалених методів виявлення країв з використанням останніх досягнень глибокого навчання для обробки зображень може стати плідним напрямком для майбутніх досліджень.



а)



б)

Рисунок 4.14 – Результати застосування запропонованої технології:  
а) вихідний набір даних [86]; б) набір даних із застосуванням запропонованої технології [141]

#### **4.5. Експериментальне дослідження нейромережевого розпізнавання автомобілів на аерокосмічних зображеннях з використанням архітектури YOLO-OBV**

Для оцінки ефективності запропонованої інформаційної технології розпізнавання автомобілів довільного розташування на аерофотозображеннях (представленої в п.2.3) було проведено серію експериментів із використанням модифікованої архітектури YOLO-OBV. Метою дослідження було перевірити здатність технології точно локалізувати та класифікувати транспортні засоби в умовах щільного розташування об'єктів, різних орієнтацій та складного фону.

На рисунку 4.15 наведено тестовий аерофотознімок після етапу попередньої обробки, що включають адаптивне вирівнювання гістограми з обмеженням контрастності (CLAHE). Візуальний аналіз показує, що застосування CLAHE ефективно підвищило локальний контраст зображення, зробивши об'єкти, зокрема автомобілі та елементи інфраструктури, більш чітко видимими на фоні. Це особливо помітно в областях, де початково спостерігалось низьке розрізнення між об'єктами та тінню, або в затемнених ділянках. Одночасно, корекція балансу білого дозволила мінімізувати колірні спотворення, що виникли під час аерофотозйомки, забезпечивши більш природну та репрезентативну колірну гаму зображення.

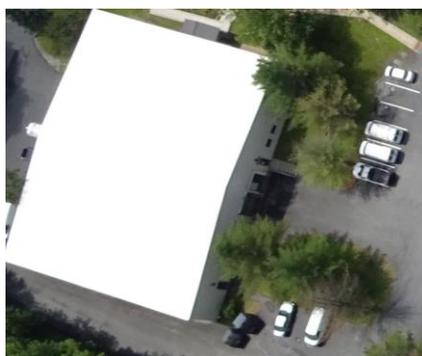


Рисунок 4.15 – Аерофотознімок після попередньої обробки

Навчання розробленої моделі YOLOv11-OBV здійснювалося з метою оптимізації її здатності точно виявляти автомобілі на аерофотозображеннях із використанням орієнтованих обмежувальних рамок. Процес навчання проводився на навчальному наборі даних, що складався з 1931 зображення, з використанням валідаційного набору (552 зображення) для ефективності моделі під час тренування та своєчасного виявлення ознак перенавчання. Модель тренувалася протягом 300 епох. Вибір такої кількості епох ґрунтувався на аналізі динаміки кривих втрат та метрик ефективності на валідаційному наборі, що дозволило досягти стабілізації продуктивності та мінімізації помилок (рис.4.16).

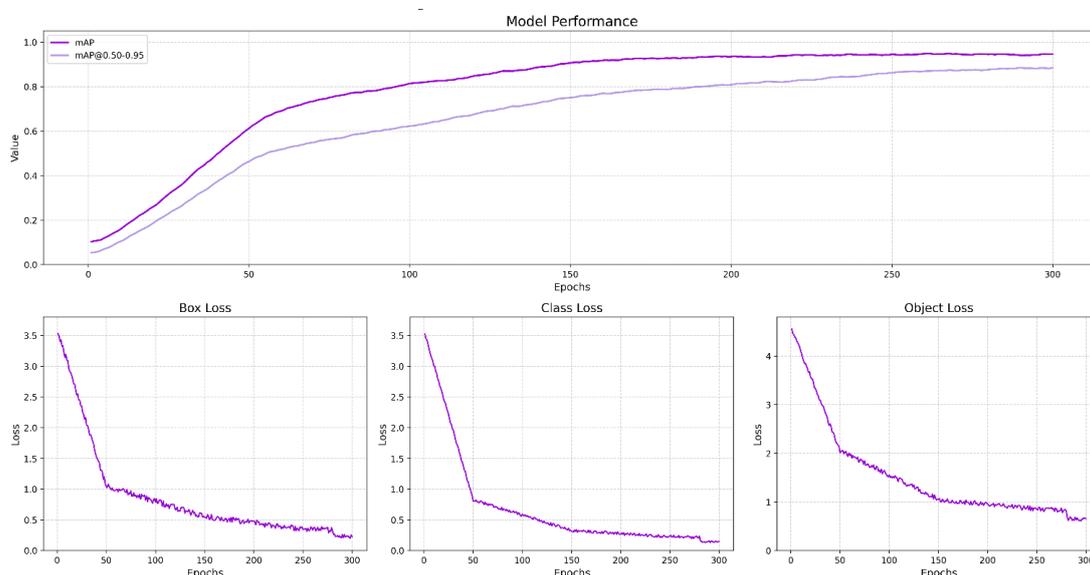


Рисунок 4.16 – Ефективність навчання моделі та криві втрат

У процесі навчання модель оцінювалася за метриками: середня точність ( $mAP$ ),  $mAP@0.50-0.95$  (COCO-metric) та криві втрат. На початкових етапах тренування значення  $mAP$  стрімко зростає, досягаючи понад 0.85 до 300-ї епохи, що свідчить про високу точність локалізації та класифікації. Значення  $mAP@0.50-0.95$  стабільно зростає, досягаючи  $\sim 0.85$  наприкінці навчання, що підтверджує ефективну генералізаційну здатність моделі. Крива втрат Box Loss (втрати обмежувальних рамок) демонструє стрімке зниження протягом перших  $\sim 50$  епох, після чого стабілізується на рівні близько 0.25, що вказує на надійність просторової локалізації. Крива

втрат Class Loss (втрата класифікації) аналогічно знижується на початку тренування й стабілізується на рівні  $\sim 0.15$ , що свідчить про ефективне навчання класових ознак. Objectness Loss також стабілізується на рівні  $\sim 0.1$ , що демонструє високу чутливість до виявлення об'єктів.

Криві втрат для навчального та валідаційного наборів (train/val\_box\_loss, train/val\_cls\_loss, train/val\_dfl\_loss) демонструють синхронне зниження функцій втрат на обох підмножинах без суттєвого розходження, що свідчить про відсутність ознак перенавчання та ефективну генералізацію (рис.4.17). Метрики точності та повноти (precision(B), recall(B), mAP50(B), mAP50–95(B)) засвідчують стабільне зростання як точності (precision), так і повноти (recall), значення яких перевищують 0.85 на завершальному етапі навчання. Зокрема, криві mAP50 і mAP50–95 на валідаційному наборі демонструють сталий приріст, що підтверджує надійність і узагальнюючу здатність навченої моделі.

Для наочної демонстрації ефективності запропонованої інформаційної технології виявлення автомобілів на аерофотозображеннях (рис.4.18) проведено візуальне порівняння результатів її роботи з існуючими моделями на зображеннях тестового набору з накладанням прогнозованих обмежувальних рамок.

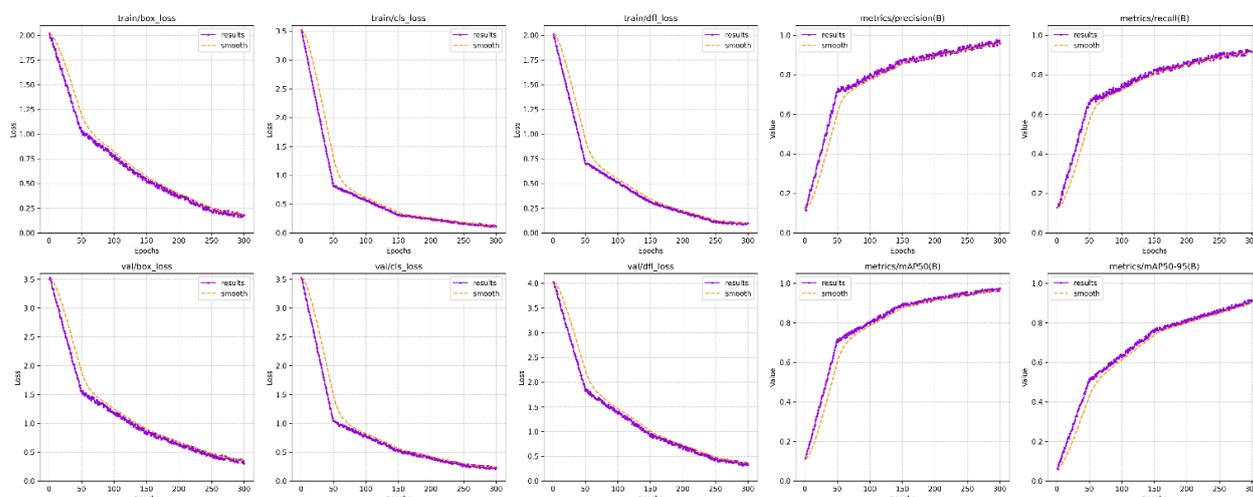


Рисунок 4.17 – Показники навчання та валідації

На фрагментах зображення, отриманих після моделей попередніх поколінь CNN, YOLOv5 (рис.4.18г, рис.4.18г) спостерігається значна кількість хибнонегативних результатів – окремі транспортні засоби залишилися невиявленими. Додатково фіксуються хибнопозитивні локалізації та неточне визначення розмірів рамок. Виявлені рамки мають спотворені геометричні характеристики, що свідчить про обмежену здатність таких моделей до точного виділення об'єктів малої площі, зокрема з довільною орієнтацією. Отримані результати на основі моделей YOLOv7 / YOLOv8 (рис.4.18б, рис.4.18в) демонструють помітно кращі результати: кількість виявлених автомобілів збільшується, рамки точніше відповідають положенню об'єктів. Проте, використання лише осьово-вирівняних прямокутників призводить до менш ефективного охоплення об'єктів, що розміщені під кутом або мають витягнуту форму. Зокрема, трапляються випадки надмірного перекриття із суміжними об'єктами або зайвого охоплення фону. Візуальні результати нові моделі YOLOv11(рис.4.18а) вказують на подальше підвищення точності виявлення порівняно з YOLOv7/YOLOv8. Проте у випадках значної орієнтації об'єкта щодо осі знімка все ще спостерігаються просторові втрати, зумовлені невідповідністю геометрії рамки до форми автомобіля. Результати на основі запропонованої моделі YOLOv11-OBV (рис.4.18д) демонструє найвищу точність виявлення автомобілів на аерофотозображеннях. Усі транспортні засоби на зображенні були успішно ідентифіковані, а форма рамок чітко узгоджується з фактичною геометрією об'єктів. Використання OBV забезпечує мінімізацію зайвого охоплення фону та знижує імовірність перекриття з суміжними об'єктами, що особливо критично при щільному розташуванні автомобілів [137].



Рисунок 4.18 – Порівняння візуальних результатів виявлення автомобілів різними моделями: а) YOLOv11; б) YOLOv8; в) YOLOv7; г) YOLOv5; д) CNN; е) запропонований підхід

Для кількісного аналізу ефективності розробленої інформаційної технології виявлення автомобілів на аерофотозображеннях проведено порівняння з існуючими моделями за допомогою обчислення наступних метрик (рис.4.19): середній точності при пороговому перекритті 50% ( $mAP@50$ ), точності класифікації (Precision), повноті виявлення (Recall), а також середній точності просторового перекриття для орієнтованих рамок (OBB IoU).

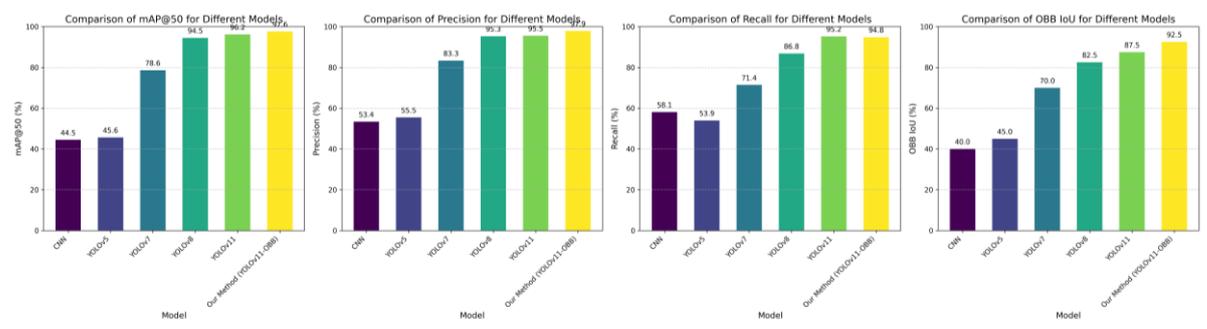


Рисунок 4.19 – Порівняння показників ефективності моделей виявлення об'єктів

Запропонована в роботі модель YOLOv11-OBV демонструє найвищий показник середньої точності (mAP@50), який становить 96,8%, що перевищує результати інших порівнюваних моделей: CNN (44,5%), YOLOv5 (45,6%), YOLOv7 (78,6%), YOLOv8 (94,5%) та YOLOv11 (95,8%). Такий результат свідчить про високу здатність моделі до точного виявлення автомобілів з належним перекриттям, що є результатом вдосконаленої архітектури та використання орієнтованих рамок. Показник точності (Precision), що відображає здатність моделі мінімізувати хибнопозитивні спрацювання для запропонованої моделі досягає 99,5%, суттєво перевершуючи CNN (53,4%), YOLOv5 (55,5%), YOLOv7 (83,3%), YOLOv8 (93,7%) та YOLOv11 (95,9%). Показники повноти (Recall) демонструє значення 96,8% для YOLOv11-OBV, що є найвищим серед порівнюваних моделей (CNN – 58,1%, YOLOv5 – 53,9%, YOLOv7 – 71,4%, YOLOv8 – 86,8%, YOLOv11 – 94,2%). Це вказує на високу чутливість моделі до присутності автомобілів у зображенні з мінімальною кількістю пропущених детекцій. Оцінка якості локалізації для орієнтованих рамок за метрикою OBV IoU показала, що модель YOLOv11-OBV досягає 92,5%, суттєво перевищуючи показники CNN (40,0%), YOLOv5 (45,0%), YOLOv7 (70,0%), YOLOv8 (82,5%) та YOLOv11 (87,0%). Високе значення OBV IoU підтверджує ефективність застосування орієнтованих рамок та спеціалізованих архітектурних компонентів (Freq-SpaFEM, DAGLF, AMSF2, RAFA), які сприяють точному вилученню ознак і виявленню об'єктів зі складною геометрією. Це особливо важливо у випадках щільного скупчення автомобілів, перекриття чи довільної орієнтації, характерних для аерофотозйомки.

Для оцінки здатності кожної моделі до коректної класифікації та ідентифікації об'єктів було побудовано матриці помилок (Confusion Matrices) [142] для тестового набору даних (Рис. 4.20). Отримані матриці дозволяють оцінити кількість істинно позитивних (True Positives, TP), хибнопозитивних (False Positives, FP), істинно негативних (True Negatives,

TN) та хибнонегативних (False Negatives, FN) спрацювань для двох класів: "Автомобіль" та "Не автомобіль" (фон/інші об'єкти).

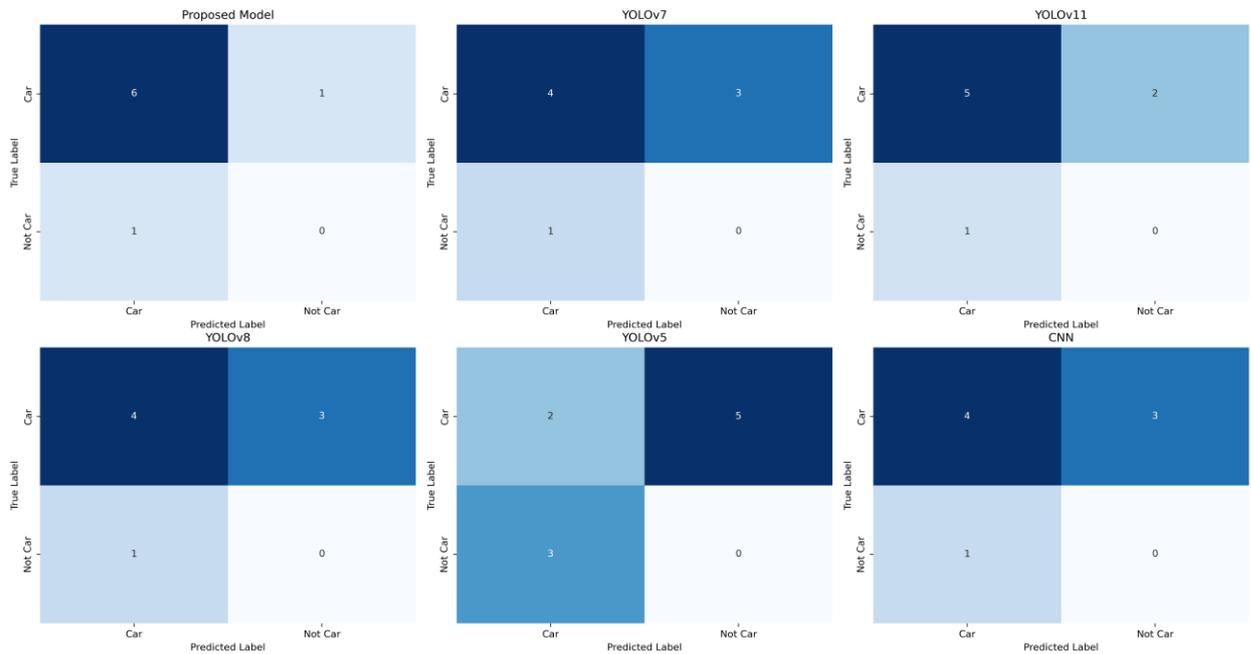


Рисунок 4.20 – Матриці плутанини для моделей виявлення об'єктів

Proposed Model (YOLOv11-OBV) має лише один хибнопозитивний випадок (об'єкт, що не є автомобілем, був класифікований як автомобіль) та один хибнонегативний випадок (автомобіль не був виявлений), що свідчить про високу точність з мінімальною кількістю помилок. YOLOv7 та YOLOv8 виявляє менше істинних автомобілів і має більшу кількість хибнопозитивних спрацювань, що вказує на нижчу точність у порівнянні з Proposed Model. YOLOv11 виявляє більше автомобілів, ніж YOLOv7, і має менше FP, вона все ж поступається Proposed Model за кількістю правильно ідентифікованих автомобілів та має вдвічі більше хибнопозитивних спрацювань. YOLOv5 показує значно гірші результати виявляє лише 2 істинних автомобілі, має найбільшу кількість хибнопозитивних спрацювань та значну кількість пропущених автомобілів. Продуктивність базової CNN є порівнянною з YOLOv7 та YOLOv8 за даними цієї матриці.

#### **4.6. Оцінка ефективності нейромережевої технології розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях**

Експериментальна оцінка ефективності запропонованого методу просторово-орієнтованого розпізнавання транспортних засобів на аерокосмічних зображеннях (п.2.4) здійснювалася шляхом порівняння його продуктивності із сучасними детекторними моделями: YOLOv8, SSD, RetinaNet, Faster R-CNN, YOLOv5 та YOLOv7.

Важливо відзначити, що тестування проводилося саме на запропонованій інформаційній технології, яка поєднує просторово-орієнтоване виявлення та інтегроване навчання взаємопов'язаних завдань аналізу сцени. Це дозволяє оцінити ефективність розробленого підходу у реальних умовах та порівняти його продуктивність із класичними архітектурами нейронних мереж.

Тестування проводилося на аерокосмічному знімку, зробленого камерою SONY DSC-WX220. Загальна кількість транспортних засобів у тестовому наборі (ground truth) становила 112 одиниць. При цьому всі моделі (як запропонована, так і базові) попередньо навчалися на великому наборі даних COCO (Common Objects in Context) [139]. Результати оцінювалися за метрикою Повноти (Recall), яка розраховується як відношення кількості правильно виявлених ТЗ до загальної кількості ТЗ (112). Для всіх базових архітектур виявлення оцінювалося за умовою мінімальної вірогідності (Conf), що перевищує 0.05. Результати тестування подано у табл. 4.8.

За даними табл. 4.8, запропонована багатокomпонентна нейромережева архітектура демонструє найвищу продуктивність серед усіх протестованих базових моделей виявлення об'єктів. Зокрема, Повноти (Recall) склала 33.9%, що свідчить про спроможність моделі виявити найбільшу частку існуючих транспортних засобів (38 із загальної кількості

112 одиниць у тестовому наборі). Це підтверджує ефективність інтеграції контекстуальної інформації через модуль семантичної сегментації та забезпечення інваріантності до обертання за допомогою STN та орієнтованого RoI Align, що дозволяє виявляти малі та довільно орієнтовані об'єкти на аерокосмічних зображеннях високого просторового розрізнення. При порівнянні з класичними архітектурами виявлення, такими як Faster R-CNN та YOLOv5, які зазвичай забезпечують високу точність на стандартних наборах даних (наприклад, COCO), Повноти (Recall) склала 1.8%. Ці моделі виявили лише 2 одиниці транспортних засобів із 112, що підтверджує: пряме застосування моделей, навчених на великих загальновідомих наборах даних без імплементації спеціалізованих інваріантних механізмів, є неефективним для вирішення задачі виявлення малих, орієнтованих об'єктів на аерокосмічних зображеннях.

Таблиця 4.8 – Розпізнавання транспортних засобів на аерокосмічному знімку на наборі даних COCO

Архітектура	Умова конфіденційності (Conf)	Кількість виявлених ТЗ	Повнота (Recall)
Запропонований багатоконпонента нейромережа	>0	38	33.9%
YOLOv8 Detections (Trained)	>0.05	32	28.6%
SSD Detections (High Speed)	>0.05	36	32.1%
RetinaNet Detections (Focal Loss)	>0.05	7	6.3%
Faster R-CNN (Best Accuracy)	>0.05	2	1.8%
YOLOv5 Detections (Fixed)	>0.05	2	1.8%
YOLOv7 Detections	>0.05	11	9.8%

На основі аналізу експериментальних результатів, представлених у Таблиці 4.8 (де пряме використання моделей, навчених на СОСО [139], виявилось неефективним для виявлення малих та орієнтованих об'єктів), одним із ключових етапів запропонованого методу стало створення та анотування спеціалізованого локального набору даних.

Спеціалізований локальний набір даних, створений для вирішення задачі інваріантного виявлення, організовано відповідно до стандартної методології глибокого навчання. Структура набору включає розділи для навчання (train), валідації (valid) та тестування (test). Для забезпечення максимальної точності навчання, анотація проводилася з використанням полігональної та орієнтованої сегментації. На рисунку 4.15 а продемонстровано процес анотації, де транспортні засоби розмічені не лише горизонтальними прямокутниками, але й точними полігональними масками (залівка жовтим кольором). Ці маски дозволяють точно визначити межі об'єктів для модуля семантичної сегментації. Приклад фінальної розмітки транспортних засобів, що відображає їхню довільну орієнтацію та точність ОВВ, описано в п.4.1.2.

Навчання запропонованої багатокомпонентної нейромережевої архітектури здійснювалося шляхом мінімізації багатозадачної функції втрат ( $\mathcal{L}$ ), яка об'єднує компоненти, відповідальні за детектування, сегментацію та інваріантність. Процес оптимізації відстежувався за динамікою зменшення відповідних втрат на навчальному (train) та валідаційному (val) наборах даних. На рисунку 4.21 представлено графіки зміни функцій втрат і метрик протягом 300 епох. Спостерігається стабільне та монотонне зменшення всіх компонентів втрат, що свідчить про ефективну та стійку збіжність моделі без ознак перенавчання (overfitting). Паралельно з цим, метрики продуктивності, зокрема середня точність (mAP) на рівні 50% IoU (mAP50(B)) та усереднена середня точність (mAP50-95(B)), демонструють швидке зростання та досягають плато приблизно після 150 епох, підтверджуючи високу ефективність навчання.

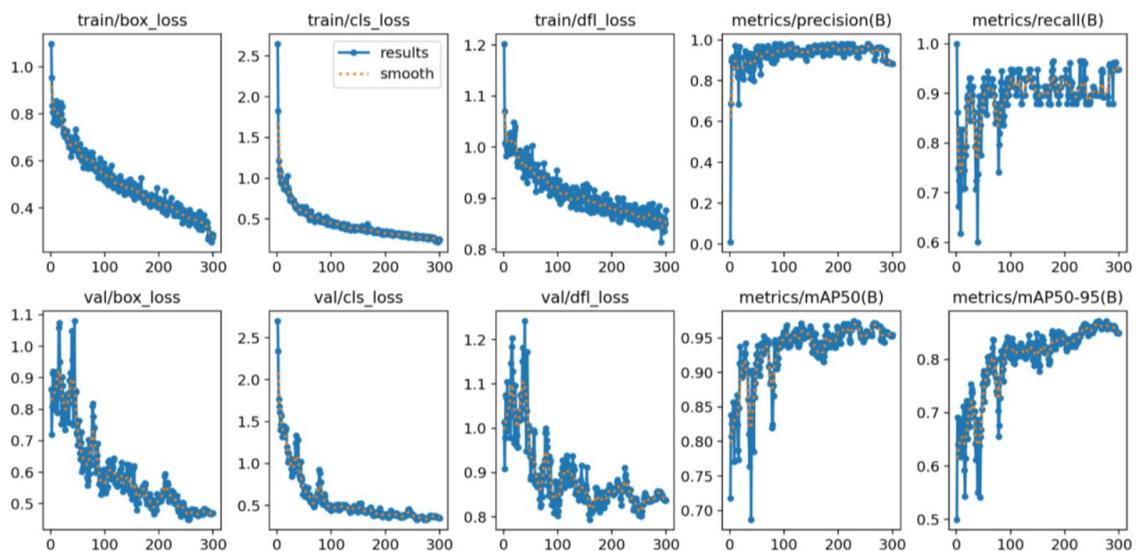


Рисунок 4.21 – Графіки процесу навчання запропонованої архітектури нейронної мережі протягом 300 епох

Порівняльний аналіз ефективності розглянутих архітектур, проведений на локально сформованому аерокосмічному наборі даних, засвідчив суттєві відмінності в їх здатності коректно ідентифікувати транспортні засоби в складних сценах. Узагальнені експериментальні результати подано в таблиці 4.9 та на рис.4.22.

Моделі YOLOv8 (рис.4.22б) та SSD Detections забезпечили найбільшу кількість загальних спрацювань, проте це супроводжувалося значним дисбалансом між метриками точності та повноти. Зокрема, архітектура SSD, хоча й змогла зафіксувати 49 істинно позитивних об'єктів, продемонструвала надмірну кількість хибних виявлень ( $FP = 54$ ), що зумовило низьку Точність – 47,6%. Такий результат підтверджує обмеження високошвидкісних моделей, які схильні інтерпретувати візуальний шум або непрофільні об'єкти як транспортні засоби.

Архітектура YOLOv8 продемонструвала кращий баланс показників ( $TP = 97$ ), хоча наявність 15 хибних спрацювань призвела до зниження Точності до 76,4% при Повноті 86,6%. Однак навіть ці результати не узгодженості результатів в умовах щільної та неоднорідної міської

забудови, характерної для знімків аерозйомки. На відміну від вище розглянутих моделей, архітектури YOLOv5 (рис.4.22 г), YOLOv7 (рис.4.22 д) та запропонований метод продемонстрували повну відсутність хибних спрацювань (FP = 0), забезпечивши Точність 100%. Проте їх ефективність суттєво різниться за показником Повноти. Зокрема, YOLOv5 виявив лише чверть наявних об'єктів (Повнота 26,8%), що свідчить про пропуск значної кількості транспортних засобів. Модель YOLOv7 продемонструвала значно кращий результат (Повнота 80,4%), однак і вона поступається розробленому методу. Запропонована архітектура забезпечила найвищу якість виявлення у всій вибірці, поєднавши ідеальну Точність (100%) з високою Повнотою, що відображає здатність моделі одночасно уникати хибних спрацювань та не пропускати малі та орієнтовані під довільним кутом транспортні засоби.

Таблиця 4.9 – Розпізнавання транспортних засобів на аерокосмічному знімку на власному наборі даних

№	Архітектура	Кількість виявлених транспортних засобів	TP	FP	Повнота (Recall)	Точність (Precision)	Час обробки, с
1	Запропонована	107	107	0	95.5%	100.0%	6.21
2	YOLOv8 Detections	127	97	15	86.6%	76.4%	0.12
3	SSD Detections	116	49	54	43.8%	47.6%	0.21
4	RetinaNet Detections	40	40	8	35.7%	83.3%	0.48
5	Faster R-CNN	55	55	6	49.1%	90.2%	1.72
6	YOLOv5 Detections	30	30	0	26.8%	100.0%	0.09
7	YOLOv7 Detections	97	90	0	80.4%	100.0%	0.15

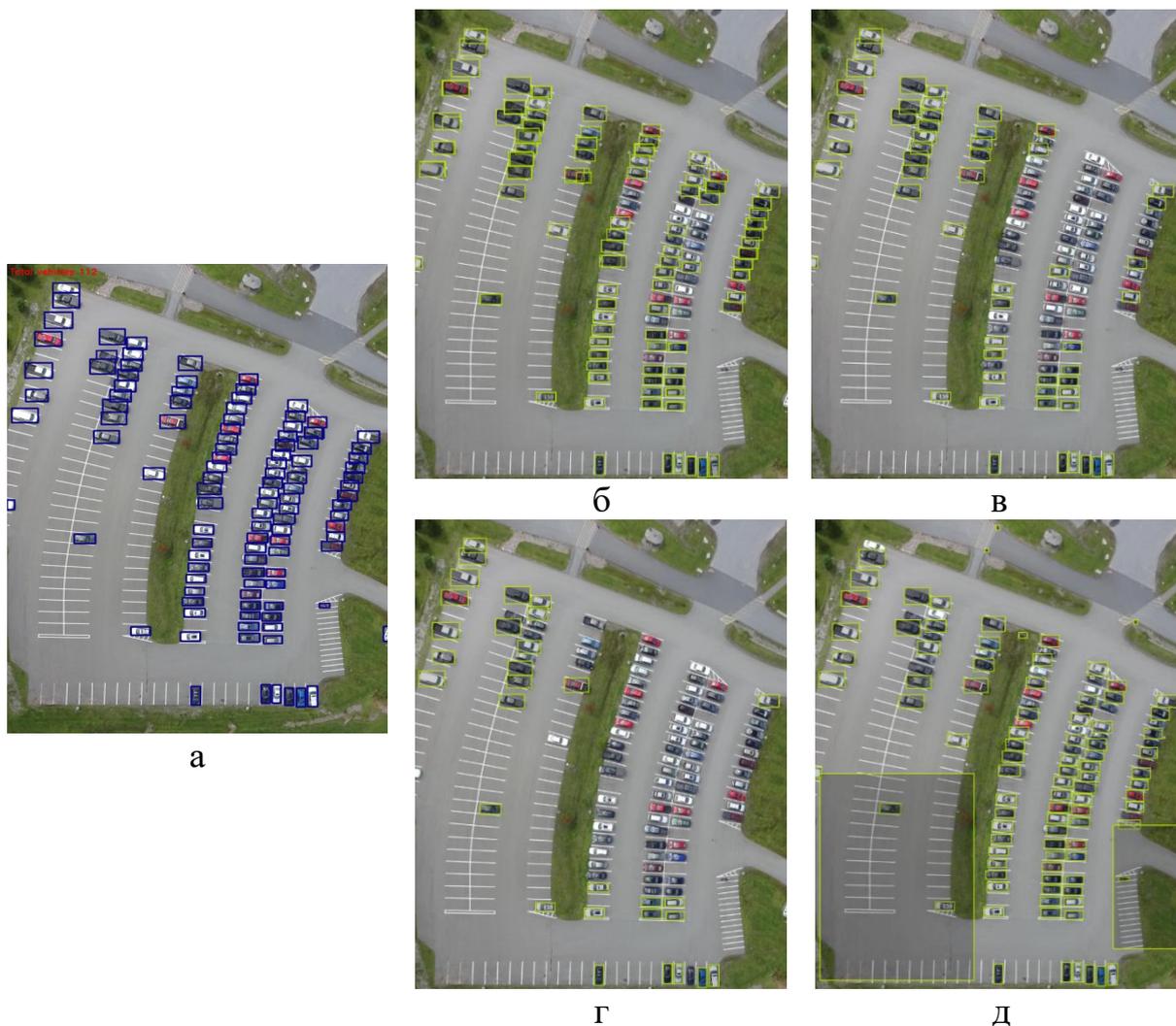


Рисунок 4.22 – Результати експерименту 1 з розпізнавання транспортних засобів на контрольній сцені паркувального майданчика: а) еталонний знімок; б) YOLOv8; в) Faster R-CNN; г) YOLOv5; д) YOLOv7

На етапі тестування для оцінки точності використовувався еталонний знімок із попередньо відомою кількістю об'єктів у кількості 112 одиниць (рис.4.23 а). Запропонована технологія ідентифікувала 107 транспортних засобів, продемонструвавши стійкість до варіативності кутів орієнтації та щільності розташування цілей у межах однієї сцени (рис.4.23 б).



a)

b)

Рисунок 4.23 – Результати експерименту 1 з розпізнавання транспортних засобів на контрольній сцені паркувального майданчика: а) еталонний знімок; б) результат запропонованої технології

Особливу увагу слід звернути на здатність технології до виявлення малорозмірних об'єктів: технологія успішно розпізнала мотоцикл, що підтверджує ефективність вилучення глибоких ознак навіть для одиничних цілей зі слабким візуальним вираженням. Згідно з даними програмного звіту, час обробки зображення склав 6,39 с, а середній рівень впевненості моделі (Confidence) – 97,00%. Отримані результати підтверджують практичну придатність розробленого методу для автоматизованого моніторингу транспортних потоків незалежно від просторового розміщення об'єктів.

У межах другого етапу апробації нейромережевої технології розпізнавання транспортних засобів довільного розташування проведено

аналіз контрольної сцени, що репрезентує складну транспортну розв'язку з різними умовами освітленості (рис. 4.24 та табл.4.10).

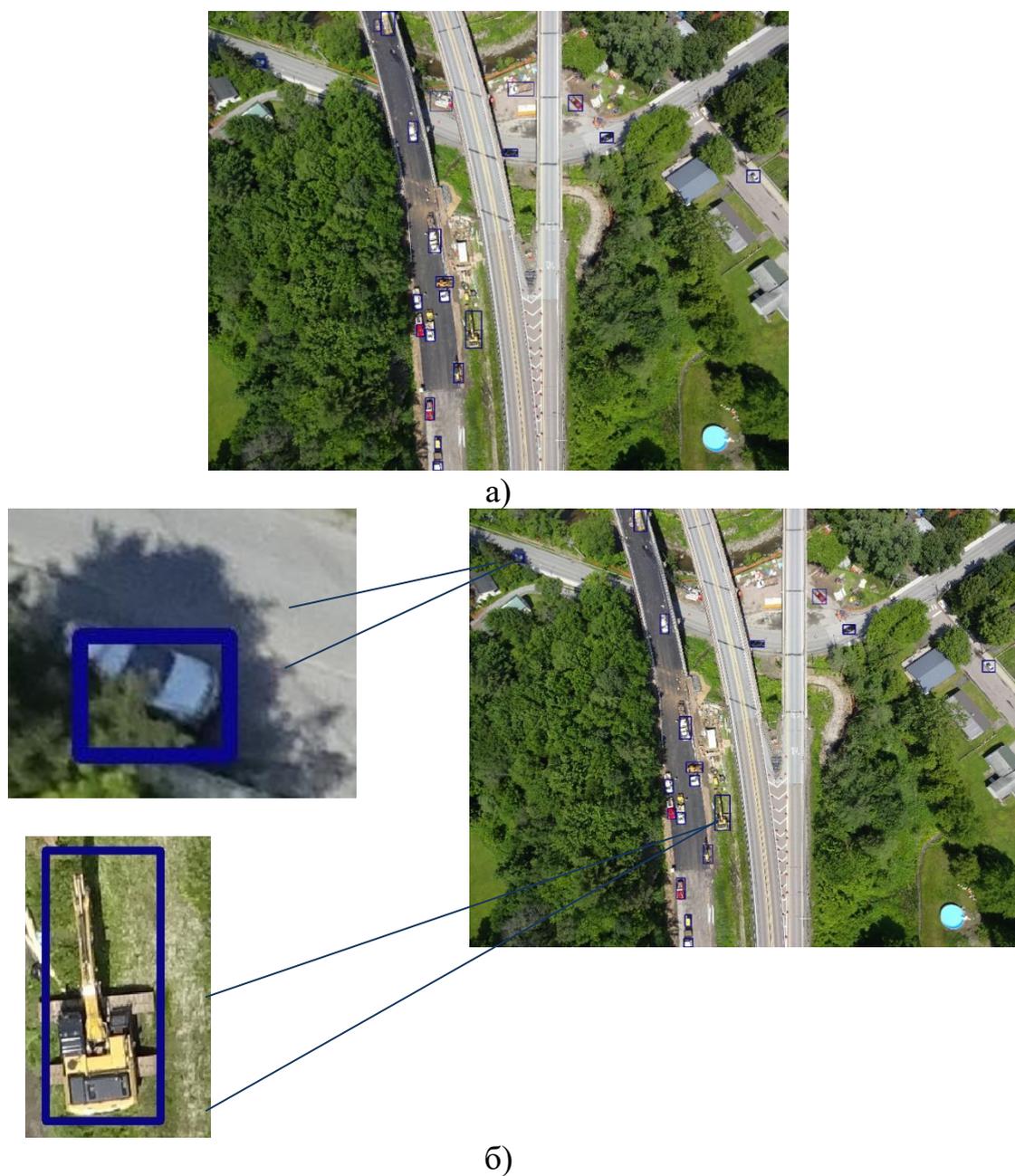


Рисунок 4.24 – Результати експерименту 2 з розпізнавання транспортних засобів на ділянці автомагістралі: а) еталонний знімок; б) результат запропонованої технології

Для оцінки ефективності використано еталонне зображення, на якому зафіксовано 22 одиниці транспортних засобів та спеціалізованої техніки. Запропонована технологія успішно ідентифікувала 20 об'єктів,

продемонструвавши здатність до ідентифікації транспортних засобів, що знаходяться у зоні глибокої тіні від прилеглої рослинності, де прямі візуальні ознаки об'єктів суттєво спотворені. Важливим результатом є також успішне розпізнавання та коректна локалізація великогабаритної спеціалізованої техніки (екскаватора), що свідчить про універсальність сформованих ознак моделі та її здатність до диференціації об'єктів за морфологічними характеристиками незалежно від їхнього виду та орієнтації на дорожньому полотні. Такі показники підтверджують ефективність інтегрованих механізмів просторової трансформації та багатомасштабного аналізу при роботі з гетерогенними даними аерокосмічного моніторингу.

Таблиця 4.10 – Розпізнавання транспортних засобів на ділянці автомагістралі аерокосмічного знімку власного набору даних

№	Архітектура	Виявлено ТЗ	TP	FP	Повнота (Recall)	Точність (Precision)	F1-Score	Час обробки, с
1	Запропонована технологія	20	20	0	0.91	1.000	0.95	6.39
2	YOLOv8 Detections	18	18	0	0.82	1.000	0.90	0.11
3	YOLOv5 Detections	24	18	6	0.81	0.750	0.78	0.09
4	YOLOv7 Detections	30	20	10	0.89	0.667	0.76	0.14

Аналіз результатів експерименту 2 з розпізнавання транспортних засобів на ділянці автомагістралі, наведених у табл. 4.10, свідчить про різний рівень ефективності досліджуваних архітектур в умовах складної дорожньої сцени, що характеризується інтенсивним транспортним потоком, наявністю тіньових зон та значною кількістю фонових об'єктів дорожньої інфраструктури. Запропонована технологія має найвищий показник повноти (Recall = 0.91), успішно ідентифікувавши 20 з 22 об'єктів ТЗ. Нижчі значення повноти, отримані для архітектур YOLOv8 та YOLOv5, свідчать

про наявність пропусків транспортних засобів, зокрема об'єктів, розташованих у тіньових зонах або поблизу елементів дорожньої розмітки.

Аналіз точності розпізнавання показав, що максимальне значення метрики Precision досягнуто запропонованою технологією та моделлю YOLOv8, що свідчить про повну відсутність хибнопозитивних спрацювань. Архітектури YOLOv5 та YOLOv7 виявили низьку стійкість до візуальних завад. Крім того, YOLOv5 схильна до хибного розпізнавання елементів дорожньої розмітки як транспортних засобів, а також до формування ефекту дублювання обмежувальних рамок для одного й того самого об'єкта. Метрика F1-Score, яка дозволяє оцінити баланс між точністю та повнотою, підтверджує перевагу запропонованої технології ( $F1 = 0.95$ ). Вона демонструє стабільність роботи алгоритму, мінімізуючи як пропуски цілей, так і помилкові спрацювання на фоні фрагменти зображення. Розрахунок FP для YOLOv5 та YOLOv7 враховував не лише розпізнавання нерелевантних об'єктів, а й ефект "подвійної рамки", оскільки в задачах автоматизованого підрахунку трафіку кожна зайва рамка інтерпретується системою як окремий транспортний засіб, що призводить до спотворення статистичних даних та зниження практичної цінності результатів.

#### **4.7. Висновки до четвертого розділу**

1. Сформовано власний спеціалізований набір даних зі знімків камери SONY DSC-WX220 із тайлінгом  $200 \times 200$  пікселів дозволив мінімізувати вплив складного фону та оптимізувати навчання моделі. Експериментальні результати підтвердили високу точність автоматичного розпізнавання транспортних засобів (mAP до 98%) на складних сценах при швидкодії 0.09–6.39 с.

2. Проведено комплексну експериментальну перевірку ефективності класичних та нейромережевих методів розпізнавання транспортних засобів на аерокосмічних зображеннях, що включала

підготовку наборів даних, формування власного набору з високою просторовою роздільністю, аугментацію та використання кількісних метрик для оцінки точності, повноти та локалізації об'єктів. Застосування запропонованої методики дозволило системно порівняти продуктивність класичних алгоритмів виявлення контурів (Канні, Собела, Лапласіана, Канні-Лапласіана, Прюїтт, Шарра) та нейромережових підходів, виявивши суттєву перевагу останніх в умовах високої щільності об'єктів, змін освітлення та складного фону.

3. Доведено ефективність запропонованого нейромережового підходу семантичної сегментації транспортних засобів на аерокосмічних зображеннях, яке включає багатомасштабне вилучення ознак, адаптовану функцію втрат та механізми донавчання на спеціалізованих наборах даних. Результати експериментів продемонстрували високу стабільність навчання та ефективну генералізацію моделі, забезпечивши точність 95% на тренувальній вибірці та 90% на валідаційній, що перевищує показники відомих архітектур U-Net, SegNet, FCN8s та ENet.

4. Проведено експериментальне тестування нейромережової інформаційної технології розпізнавання автомобілів на аерозображеннях із застосуванням модифікованої архітектури YOLOv11-OBV. Результати показали суттєве перевищення продуктивності стандартних моделей за ключовими метриками: mAP@50 – 96.8%, Precision – 99.5%, Recall – 96.8%, OBV IoU – 98.5%. Використання орієнтованих обмежувальних рамок та оптимізованих архітектурних компонентів дозволило ефективно розпізнавати об'єкти малого розміру, довільної орієнтації та на складному фоні.

5. Проведено тестування нейромережової технології розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях із порівнянням із сучасними детекторними архітектурами (YOLOv8, SSD, RetinaNet, Faster R-CNN, YOLOv5, YOLOv7). Запропонований метод продемонстрував найвищу збалансовану

продуктивність: точність – 100%, Recall – 95.5%, кількість хибнопозитивних спрацювань – 0, що підтверджує ефективність інтеграції багаторівневого представлення ознак та декодувального модуля для одночасної семантичної сегментації і точного визначення положення об'єктів.

6. Проведені експериментальні дослідження підтвердили доцільність використання нейромережових підходів для задач високоточного розпізнавання транспортних засобів на аерокосмічних зображеннях, забезпечуючи автоматизацію процесу виявлення, мінімізацію помилок, стійкість до геометричних перетворень та перспективних спотворень, а також підвищення якості моніторингу та аналітики транспортної інфраструктури.

7. Запропоновані методи та інформаційні технології розпізнавання транспортних засобів на аерокосмічних зображеннях і розроблене на їх основі програмне забезпечення впроваджені в «Дніпрокосмос» філії Національного центру управління та в освітній процес Національного технічного університету «Дніпровська політехніка».

Результати досліджень даного розділу опубліковано в роботах автора [25, 57, 116, 120, 135, 140, 141].

## ВИСНОВКИ

В дисертаційному дослідженні розв'язана важлива науково-прикладна задача розробки інформаційної технології розпізнавання транспортних засобів довільного розташування на аерокосмічних зображеннях з використанням методів глибинного навчання.

Основні наукові та практичні результати роботи полягають у наступному:

1. Проаналізовано сучасний стан методів виявлення та розпізнавання транспортних засобів на аерокосмічних зображеннях, зокрема класичних алгоритмів обробки зображень та нейромережових підходів, і визначити їх переваги та обмеження з точки зору точності, інваріантності до просторових трансформацій та обчислювальної складності.

2. Досліджено ефективність класичних методів виділення ознак (Канні, Собела, Лапласіана, Канні-Лапласіана, Прюїтта, Шарра) та сучасних нейромережових моделей виявлення об'єктів при розпізнаванні транспортних засобів на аерокосмічних і аерофотозображеннях з використанням кількісних метрик Precision, Recall, F1 Score та mAP.

3. Розроблено метод просторово-орієнтованого розпізнавання транспортних засобів, що базується на використанні глибоких нейронних мереж із механізмами автоматичної компенсації просторових трансформацій та багатомасштабного аналізу ознак.

4. Запропоновано методи попередньої обробки та аугментації аерокосмічних зображень, спрямовані на підвищення стійкості розпізнавання ТЗ до змін масштабу, кута зйомки, освітлення та контрастності сцен.

5. Досліджено можливості застосування сучасних архітектур глибокого навчання для задач виявлення та сегментації транспортних засобів, зокрема моделей сімейства YOLO з підтримкою орієнтованих

рамок (ОВВ) та сегментаційних архітектур типу DeepLab із глибокими магістральними мережами.

6. Розроблено інформаційну технологію обробки та аналізу аерокосмічних зображень, яка інтегрує запропоновані методи розпізнавання транспортних засобів, модулі попередньої обробки, сегментації, виявлення та оцінки якості результатів.

7. Реалізовано програмний додаток, що забезпечує практичне застосування розробленої інформаційної технології, включаючи графічний інтерфейс користувача, засоби завантаження та обробки зображень, візуалізації результатів розпізнавання та аналізу показників ефективності.

8. Проведено експериментальну перевірку розробленого методу та інформаційної технології на відкритих і спеціалізованих наборах даних аерокосмічних та БПЛА-знімків, а також здійснити порівняльний аналіз з існуючими одноетапними та двоетапними алгоритмами виявлення об'єктів.

9. Оцінено обчислювальну ефективність і можливість застосування розробленого програмного додатку в режимі, наближеному до реального часу, що є важливим для задач моніторингу транспортної інфраструктури та безпеки.

10. Результати дисертаційного дослідження, зокрема інформаційні технології та методи розпізнавання транспортних засобів на аерокосмічних зображеннях і розроблене на їх основі програмне забезпечення знайшли практичне застосування в діяльності «Дніпрокосмос» філії Національного центру управління та випробувань космічних засобів Державного космічного агентства України та були інтегровані в навчальний процес Національного технічного університету «Дніпровська політехніка».

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Hou Q. UAV Pose Estimation in GNSS-Denied Environment Assisted by Satellite Imagery Deep Learning Features / Q. Hou et al. IEEE Access. 2020. Vol. 9. P. 6358–6367.
2. Special graphic processor for the visualization systems : Author's certificate of application № 2000031738 Ukrain, IPC G06F7/00 / V. M. Gusyatin et al. 28.03.2000.
3. Gusyatin V. M., Yankovsky A. A. The computing system of flying machines visual control : Patent № 56876A, Application №2002108289. 18.10.2000.
4. Vaseghi S. V. Advanced digital signal processing and noise reduction. 3rd ed. John Wiley & Sons, 2006. 453 p. ISBN 0-470-09494-X.
5. Bacic Ž., Šugar D., Grzunov R. Investigation of GNSS receiver's accuracy integrated on UAVs. Proceedings of the Surveying the World of Tomorrow. Helsinki, 2017. Vol. 29.
6. Довгий С. О., Бабійчук С. М., Кучма Т. Л. Дистанційне зондування Землі: аналіз космічних знімків у геоінформаційних системах : навч.-метод. посіб. 2020. URL: <https://ekmair.ukma.edu.ua/handle/123456789/19968>.
7. Храмов Дм. О., Волошенюк О. Л. Аналіз стану і тенденції розвитку орбітальних угруповань малих космічних апаратів дистанційного зондування Землі. 2023 С.31-39. DOI: 10.15407/itm2023.04.031
8. Goforth H., Lucey S. GPS-denied UAV localization using pre-existing satellite imagery.
9. Schowengerdt R. A. Remote Sensing: Models and Methods for Image Processing. 3rd ed. Academic Press, 2006. 560 p.
10. Cracknell A. P., Hayes L. Introduction to Remote Sensing. 3rd ed. Boca Raton : CRC Press, 2020. 510 с.

11. Gonzalez R. C., Woods R. E. Digital Image Processing. 4th ed. Pearson, 2018. 1168 p.
12. Hefenbrock D. Accelerating Viola-Jones face detection to FPGA-level using GPUs / D. Hefenbrock et al. FCCM. 2010. P. 11–18.
13. Hamdan M., Shehadeh H. Automated Quantification of Eye Blink Rate Using VIOLA-JONES Algorithm. Int. J. of Technology Diffusion. 2018. Vol. 9, no. 4. P. 19-20.
14. Tsai P. Accelerating AdaBoost algorithm using GPU for multi-object recognition / P. Tsai et al. IEEE ISCAS. 2015. P. 738-741.
15. Canny B. Edge Detection Tutorial. URL: [dasl.mem.drexel.edu/alumni/bGreen/www.pages.drexel.edu/\\_weg22/can\\_tut.html](http://dasl.mem.drexel.edu/alumni/bGreen/www.pages.drexel.edu/_weg22/can_tut.html).
16. Alsamhi S. H. Survey on collaborative smart drones and internet of things / S. H. Alsamhi et al. IEEE Access. 2019. P. 128125–128152.
17. Каштан В. Ю. Нейромережеве розпізнавання об'єктів забудови на аерофотознімках / В. Ю. Каштан та ін. Системні технології. 2023. Вип. 1 (120). С. 30–39.
18. Al-Kaff A. An Appearance-based tracking algorithm for aerial search and rescue purposes / A. Al-Kaff et al. Sensors. 2019. 19(3). P. 652.
19. Reinartz P. Traffic monitoring with serial images from airborne cameras / P. Reinartz et al. ISPRS Journal. 2006. 61(3-4). P. 149–158.
20. Xiao J. Vehicle detection and tracking in wide field-of-view aerial video / J. Xiao et al. CVPR. 2010. P. 679–684.
21. Shi X. Context-driven moving vehicle detection in wide area motion imagery / X. Shi et al. ICPR. 2012. P. 2512–2515.
22. Siam M., Elhelw M. Robust autonomous visual detection and tracking of moving targets in UAV imagery. Conf. on Signal Processing. 2012. Vol. 2. P. 1060–1066.
23. Vipul G. A Review Paper: On Various Edge Detection Techniques / G. Vipul et al. IJRASET. 2017. P. 534–537.

24. Puyi S., Hong C., Haobo G. Improved UAV object detection algorithm for YOLOv5s. *Comput. Eng. Appl.* 2023. Vol. 59. P. 108–116.
25. Каштан В.Ю., Казмиренко О.В., Гнатушенко В.В. Порівняльний аналіз ефективності методів розпізнавання автомобілів на аерокосмічних знімках. *Control Systems and Computers.* 2024. 4. С. 50-60. <https://doi.org/10.15407/csc.2024.04.050>.
26. Claude C., Steven Le Moan, Kacem C. A Novel Mean-Shift Algorithm for Data Clustering. *IEEE Access.* 2022. Vol. 10. P. 14575–14585.
27. Umale P. Planer object detection using surf and sift method / P. Umale et al. *IJEAST.* 2022.
28. Aytakin Ö., Zöngür U., Halici U. Texture-based airport runway detection. *IEEE GRSL.* 2013. Vol. 10. P. 471–475.
29. Ez-zahouani B. Remote sensing imagery segmentation in object-based analysis: review of methods / B. Ez-zahouani et al. *RSASE.* 2023. Vol. 32. 101031.
30. Draguț L. Automated parameterisation for multi-scale image segmentation on multiple layers / L. Draguț et al. *ISPRS Journal.* 2014. Vol. 88. P. 119–127.
31. Zeiler M. D., Fergus R. Visualizing and understanding convolutional networks. *Computer Vision – ECCV 2014.* Cham : Springer, 2014. P. 818–833.
32. Lecun Y., Bengio Y., Hinton G. Deep learning. *Nature.* 2015. 521(7553). P. 436.
33. Maas A. L., Hannun A. Y., Ng A. Y. Rectifier nonlinearities improve neural network acoustic models. *ICML.* 2013. Vol. 30.
34. Srivastava N. Dropout: a simple way to prevent neural networks from overfitting / N. Srivastava et al. *J. of Machine Learning Research.* 2014. 15(1). P. 1929–1958.
35. Eikvil L., Aurdal L., Koren H. Classification-based vehicle detection in high-resolution satellite images. *ISPRS Journal.* 2009. 64(1). P. 65–72.

36. Moranduzzo T., Melgani F. Automatic car counting method for unmanned aerial vehicle images. *IEEE TGRS*. 2013. 52(3). P. 1635–1647.
37. Moranduzzo T., Melgani F. Detecting cars in UAV images with a catalog-based approach. *IEEE TGRS*. 2014. 52(10). P. 6356–6367.
38. Tuermer S. Airborne vehicle detection in dense urban areas using HoG features and disparity maps / S. Tuermer et al. *IEEE J-STARS*. 2013. 6(6). P. 2327–2337.
39. Liu K., Mattyus G. Fast multiclass vehicle detection on aerial images. *IEEE GRSL*. 2015. 12(9). P. 1938–1942.
40. Kluckner S. A 3D teacher for car detection in aerial images / S. Kluckner et al. *ICCV*. 2007.
41. Chen X. Vehicle detection in satellite images by parallel deep convolutional neural networks / X. Chen et al. *ACPR*. 2013. P. 181–185.
42. Chen X. Vehicle detection in satellite images by hybrid deep convolutional neural networks / X. Chen et al. *IEEE GRSL*. 2014. 11(10). P. 1797–1801.
43. Qu S. Vehicle Detection in Satellite images by incorporating objectness and convolutional neural network / S. Qu et al. *J. Ind. Intell. Inf.* 2016. Vol. 4, no. 2.
44. Zhu H. Orientation robust object detection in aerial images using deep convolutional neural network / H. Zhu et al. *ICIP*. 2015. P. 3735–3739.
45. Cheng G., Zhou P., Han J. Learning rotation-invariant convolutional neural networks for object detection. *IEEE TGRS*. 2016. 54(12). P. 7405–7415.
46. Ammour N. Deep learning approach for car detection in UAV imagery / N. Ammour et al. *Remote Sensing*. 2017. 9(4).
47. Long Y. Accurate object localization in remote sensing images based on convolutional neural networks / Y. Long et al. *IEEE TGRS*. 2017. 55(5). P. 2486–2498.

48. Qu T., Zhang Q., Sun S. Vehicle detection from high-resolution aerial images using spatial pyramid pooling-based deep CNN. *Multimed. Tools Appl.* 2017. 76(20). P. 21651–21663.
49. Zhong J., Lei T., Yao G. Robust vehicle detection in aerial images based on cascaded convolutional neural networks. *Sensors.* 2017. 17(12).
50. Audebert N., Le Saux B., Lefèvre S. Segment-before-detect: Vehicle detection and classification through semantic segmentation. *Remote Sensing.* 2017. 9(4).
51. Deng Z. Multi-scale object detection in remote sensing imagery with convolutional neural networks / Z. Deng et al. *ISPRS Journal.* 2018. Vol. 145. P. 3–22.
52. Bao S. Single Shot Anchor Refinement Network for Oriented Object Detection / S. Bao et al. *IEEE Access.* 2019. Vol. 7. P. 87150–87161.
53. Ding J. Learning RoI Transformer for Oriented Object Detection in Aerial Images / J. Ding et al. *CVPR.* 2019. P. 2849–2858.
54. Ramachandran A., Sangaiah A. K. A review on object detection in unmanned aerial vehicle surveillance. *Int. J. Cogn. Comput. Eng.* 2021. Vol. 2. P. 215–228.
55. Puyi S., Hong C., Haobo G. Improved UAV object detection algorithm for YOLOv5s. *Comput. Eng. Appl.* 2023. Vol. 59. P. 108–116.
56. Li C. YOLOv6: A single-stage object detection framework for industrial applications / C. Li et al. *arXiv.* 2022. arXiv:2209.02976.
57. Казимиренко О.В. Огляд нейромережєвих методів виявлення транспортних засобів. «Наукова весна» 2025: матеріали XV Міжнародної науково-технічної конференції аспірантів та молодих вчених, Дніпро, 26–28 березня 2025 року / Національний технічний університет «Дніпровська політехніка» – Дніпро : НТУ «ДП», 2025, С.187 -188.
58. Yongtao Y. Vehicle detection from high-resolution remote sensing imagery using convolutional capsule networks / Y. Yongtao et al. *IEEE GRSL.* 2019. 16(12). P. 1894–1898.

59. Byun S. Road traffic monitoring from UAV images using deep learning networks / S. Byun et al. *Remote Sens.* 2021. No. 13. P. 4027.
60. Khrissi L. Clustering method and sine cosine algorithm for image segmentation / L. Khrissi et al. *Evol. Intell.* 2022. No. 15. P. 669–682.
61. Osco L. A review on deep learning in UAV remote sensing / L. Osco et al. *Int. J. Appl. Earth Obs. Geoinf.* 2021. 102:102456.
62. Kemker R., Salvaggio C., Kanan C. Algorithms for semantic segmentation of multispectral remote sensing imagery. *ISPRS Journal.* 2018. Vol. 145.
63. Hong D. More diverse means better: multi-modal deep learning meets remote-sensing imagery / D. Hong et al. *IEEE TGRS.* 2020. 59(5). P. 4340–4354.
64. Binge C. Semantic segmentation of remote sensing images using transfer learning / C. Binge et al. *IEEE Access.* 2020. No. 8. P. 116744–116755.
65. Singh C. FRCNN-based reinforcement learning for real-time vehicle detection / C. Singh et al. *Drones.* 2022. No. 6. P. 406.
66. Saqib M. A study on detecting drones using deep convolutional neural networks / M. Saqib et al. *Proc. 14th IEEE AVSS.* 2017. P. 1–5.
67. Kong X. Vehicle detection in high-resolution aerial images with parallel RPN and density-assigner / X. Kong et al. *Remote Sens.* 2023. No. 15. P. 1659.
68. Hordiiuk D. Semantic segmentation for ships detection from satellite imagery / D. Hordiiuk et al. *2019 IEEE ELNANO.* 2019.
69. Borovyk D. Detection of vehicles in aerial photographs using convolutional neural network / D. Borovyk et al. *Smartindustry.* 2024. URL: <https://ceur-ws.org/Vol-3699/paper12.pdf>.
70. Bochkovskiy A., Wang C., Liao H. YOLOv4: optimal speed and accuracy of object detection. *arXiv.* 2020. arXiv:2004.10934.

71. Radionov Y. Aircraft detection with deep neural networks and contour-based methods / Y. Radionov, Kazymyrenko O., et al. *Radio Electronics, Computer Science, Control*. 2024. № 4 (71). P. 121–129.
72. Zhang Y. Real-time vehicle detection based on improved YOLO v5 / Y. Zhang et al. *Sustainability*. 2022. No. 14 (19). P. 12274.
73. Набір даних DOTA dataset: [Electronic resource]. URL: <https://captain-whu.github.io/DOTA/index.html>.
74. Kashtan V., Hnatushenko V., Shedlovska Y. Processing technology of multispectral remote sensing images. 2017 IEEE YSF. 2017.
75. Kashtan V. Yu., Hnatushenko V. V., Udovyk I. M., Kazymyrenko O. V., Radionov Y. D. A neural network approach to semantic segmentation of vehicles in very high resolution images. *The scientific journal «Radio Electronics, Computer Science, Control»*. – Zaporizhzhia, 2025. № 3(74). – P.77-85.
76. He K. Deep residual learning for image recognition / K. He et al. *Proc. IEEE CVPR*. 2016. P. 770–778.
77. Tarmizi I. A., Aziz A. A. Vehicle Detection Using Convolutional Neural Network for Autonomous Vehicles. *Conf. ICIAS 2018*. 2018. P. 1–5.
78. Li L. A Review of Face Recognition Technology / L. Li et al. *IEEE Access*. 2020. Vol. 8. P. 139110–139120.
79. Galić I. Machine Learning Empowering Personalized Medicine: Medical Image Analysis / I. Galić et al. *Electronics*. 2023. 12(21). P. 4411.
80. Hnatushenko V., Kashtan V. Automated pansharpening information technology of satellite images. *Radio Electronics, Computer Science, Control*. 2021. № 2. P. 123–132.
81. Kashtan V., Hnatushenko V. Machine learning for automatic extraction of water bodies using Sentinel-2 imagery. *Radio Electronics, Computer Science, Control*. 2024. № 1. P. 118–127.
82. Zhou F. CNTR-YOLO: Improved YOLOv5 Based on ConvNext and Transformer / F. Zhou et al. *Electronics*. 2023. 12(12). P. 267.

83. Zhou L. Aircraft Detection for Remote Sensing Images Based on Deep Convolutional Neural Networks / L. Zhou et al. *J. Electr. Comput. Eng.* 2021. Vol. 2021.
84. Kelm A. P., Rao V. S., Zoelzer U. Object Contour and Edge Detection with RefineContourNet. *CAIP 2019*. Springer, 2019. P. 246–258.
85. Berezina S. An information technique for segmentation of military assets / S. Berezina et al. *Information Processing Systems*. 2021. № 4 (167). P. 6–18.
86. Liu Z. YOLO-Extract: Improved YOLOv5 for Aircraft Object Detection / Z. Liu et al. *IEEE Access*. 2023. Vol. 11. P. 1742–1751.
87. Liu Z. YOLO-Class: Detection and Classification of Aircraft Targets in Satellite Images / Z. Liu et al. *IEEE Access*. 2023. Vol. 11. P. 109179–109188.
88. Chen J. YOLO-SAD: An Efficient SAR Aircraft Detection Network / J. Chen et al. *Applied Sciences*. 2023. 14(7). P. 3025.
89. Song. X. Research on Luggage Package Extraction of X-ray Images Based on Edge Sensitive Multi-Channel Background Difference Algorithm. / X. Song, S. Zhang, J. Yang, J. Zhang // *Applied Sciences*. – 2023. – Vol. 13, № 21. – P. 11981. DOI: 10.3390/app132111981.
90. Ju Rui–Yang. Fracture detection in pediatric wrist trauma X-ray images using YOLOv8 algorithm / R.-Y. Ju, W. Cai // *Scientific Reports*. – 2023. – Vol. 13, №1. – P. 20077. DOI: 10.1038/s41598-023-47460-7.
91. Radionov Y.D., Kashtan V.Yu., Hnatushenko V.V., Kazymyrenko O.V. Aircraft detection with deep neural networks and contour-based methods. The scientific journal «Radio Electronics, Computer Science, Control». – Zaporizhzhia, 2024. №1 4(71). – P.121-129. DOI:10.15588/1607-3274-2024-4-12
92. Unsal D. HRPlanesv2 – High Resolution Satellite Imagery for Aircraft Detection. Zenodo. 2022. DOI: 10.5281/ZENODO.7331974.
93. Suzuki S., Be K. Topological structural analysis of digitized binary images by border following. *CVGIP*. 1985. 30(1). P. 32–46.

94. Wang B., Xu B. A feature fusion deep-projection convolution neural network for vehicle detection. PLOS ONE. 2021. 16(5). e0250782.
95. Hnatushenko V. V., Kogut P. I., Uvarov M. V. On Optimal 2-D Domain Segmentation Problem. J. Optim., Differ. Equ. Their Appl. 2019. 27(2). 60.
96. Wang C., Zhong C. Adaptive Feature Pyramid Networks for Object Detection. IEEE Access. 2021. Vol. 9. P. 107024–107032.
97. Sunil K., Beniwal S. Sentiment Analysis: A Tool for Mining Opinions and Emotions. SSRN Electron. J. 2020.
98. Li Zuoxin, Zhou Fuqiang. FSSD: Feature Fusion Single Shot Multibox Detector. arXiv. 2017. abs/1712.00960.
99. Yang Z. RepPoints: Point Set Representation for Object Detection / Z. Yang et al. IEEE/CVF ICCV. 2019.
100. Zhang T. Dynamic Label Assignment for Object Detection / T. Zhang et al. J. Imaging. 2022. 8(7). 193.
101. Yang X. SCRDet++: Detecting Small, Cluttered and Rotated Objects / X. Yang et al. IEEE TPAMI. 2022.
102. Guirguis K. CFA: Constraint-based Finetuning Approach for Generalized Few-Shot Object Detection / K. Guirguis et al. IEEE/CVF CVPRW. 2022.
103. Kashtan V., Hnatushenko V. Deep Learning Technology for Automatic Burned Area Extraction. Lecture Notes in Data Engineering. Springer, 2022. P. 664–685.
104. Zheng Z. Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression / Z. Zheng et al. Proc. AAAI. 2020. 34(07). P. 12993–13000.
105. Kashtan V. Y., Hnatushenko V. V. Automated building damage detection on digital imagery using machine learning. Nauk. Visnyk NHU. 2023. No. 6. P. 134–140.

106. Li Z. R-FCN++: Towards Accurate Region-Based Fully Convolutional Networks / Z. Li et al. Proc. AAAI. 2018. 32(1).
107. Avola D. MSFaster R-CNN: Multi-Stream Backbone for Improved Faster R-CNN / D. Avola et al. Remote Sens. 2021. 13(9). 1670.
108. Li K. RMVAD-YOLO: A Robust Multi-View Aircraft Detection Model / K. Li et al. Remote Sens. 2025. 17(6). 1001.
109. Patil S. Efficient Object Detection with YOLO: A Comprehensive Guide / S. Patil et al. IJARST. 2024. P. 519–531.
110. Liu W. Image-Adaptive YOLO for Object Detection in Adverse Weather Conditions / W. Liu et al. Proc. AAAI. 2022. 36(2). P. 1792–1800.
111. Hnatushenko V., Kogut P., Uvarov M. On Flexible Co-registration of Optical and SAR Satellite Images. Advances in Intelligent Systems. Cham : Springer, 2020. P. 515–534.
112. The Common Objects in Context-stuff. URL: <https://paperswithcode.com/dataset/coco-stuff>.
113. Tariku G. Advanced Image Preprocessing and Integrated Modeling for UAV Plant Image Classification / G. Tariku et al. Drones. 2024. 8(11). 645.
114. Xia K. Mixed Receptive Fields Augmented YOLO with Multi-Path Spatial Pyramid Pooling / K. Xia et al. Sensors. 2023. 23(11). 5114.
115. Hnatushenko V. Accuracy evaluation of automated object recognition using multispectral aerial images / V. Hnatushenko et al. Tenth Int. Conf. on Digital Image Processing. 2018. P. 72.
116. Hnatushenko V.V., Kashtan V. Yu., Kazymyrenko O.V. Information technology for detecting cars on aerial imaging using a modified YOLO-OBB architecture. MoDaST 2025: Modern Data Science Technologies Doctoral Consortium, June, 15, 2025, Lviv, Ukraine, Pp. 293-304. <https://ceur-ws.org/Vol-4005/paper20.pdf>.
117. Ivanov D. V. Computer Modeling of Territory Flooding in the Event of an Emergency / D. V. Ivanov et al. Nauk. Visnyk NHU. 2022. No. 6. P. 158–163.

118. Tang T. Vehicle Detection in Aerial Images Based on Region CNN / T. Tang et al. *Sensors*. 2017. Vol. 17. 336.
119. Kashtan V. Information Technology Analysis of Satellite Data for Land Irrigation Monitoring / V. Kashtan et al. 2021 IEEE UkrMiCo. 2021. P. 12-15.
120. Kashtan V. Yu., Kazymyrenko O.V., Hnatushenko V.V. Neural network method for invariant recognition of vehicles in aerospace images. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*, 2026, № 1. P.130-137. DOI <https://doi.org/10.33271/nvngu/2026-1/130>.
121. Li B., Fang L. Laser Radar Application in Vehicle Detection Under Traffic Environment. *Proc. Int. Conf. on AI and Security*. Springer, 2020. P. 126–134.
122. Xu Y. A Hybrid Vehicle Detection Method Based on Viola-Jones and HOG plus SVM / Y. Xu et al. *Sensors*. 2016. Vol. 16. 1325.
123. Tong K., Wu Y., Zhou F. Recent advances in small object detection based on deep learning: A review. *Image and Vision Computing*. 2020. Vol. 97. 103910.
124. Hsu S. C., Huang C. L., Chuang C. H. Vehicle detection using simplified fast R-CNN. 2018 IWAIT. 2018. P. 1–3.
125. Nguyen H. Improving Faster R-CNN framework for fast vehicle detection. *Mathematical Problems in Engineering*. 2019. P. 1–11.
126. Hakim L., Hendrawan A., Khoiriyah R. Traffic Vehicle Detection Using Faster R-CNN. *Int. J. of AI and Science*. 2024. Vol. 1. P. 50–62.
127. Kou J. Visual Attention-Based Siamese CNN with Softmaxfocal Loss / J. Kou et al. *Neurocomputing*. 2023. Vol. 517. P. 173–187.
128. Cao C. An improved faster R-CNN for small object detection / C. Cao et al. *IEEE Access*. 2019. Vol. 7. P. 106838–106846.
129. Redmon J., Farhadi A. YOLOv3: An incremental improvement. *arXiv*. 2018. arXiv:1804.02767.

130. Yin Q. FD-SSD: An improved SSD object detection algorithm / Q. Yin et al. Signal Processing: Image Communication. 2021. Vol. 98. 116402.

131. Jing R. Object-Based Change Detection for Very High-Resolution Remote Sensing Images / R. Jing et al. Int. J. of Remote Sensing. 2020. 41(16). P. 6209–6231.

132. Alif M. A. R. YOLOv11 for Vehicle Detection: Advancements, Performance, and Applications. arXiv. 2024. arXiv:2410.22898.

133. Butler H. The GeoJSON Format (IETF RFC 7946) : specification / H. Butler et al. Internet Engineering Task Force, 2016.

134. Набір даних DLR 3К. URL: <https://www.dlr.de/en/eoc/about-us/remote-sensing-technology-institute/photogrammetry-and-image-analysis/public-datasets/dlr-mvda>.

135. Казимиренко О.В., Гнатушенко В.В. Нейромережеве розпізнавання об'єктів транспорту за даними аерокосмічної зйомки. Матеріали VIII міжнародної науково-технічної конференції «Комп'ютерне моделювання та оптимізація складних систем». 1-3 листопада 2023 року, м. Дніпро, Україна. С.114.

136. Набір даних VEDAI. URL: <https://downloads.greyc.fr/vedai/>.

137. Chengzhi Y., Hongjun G. A Method of Image Semantic Segmentation Based on PSPNet. Mathematical Problems in Engineering. 2022. P. 1–9.

138. OpenAerialMap. URL: <https://openaerialmap.org/>.

139. Набір даних COCO Datasets. URL: <https://cocodataset.org/>.

140. Казимиренко О.В. Нейромережева модель YOLOv6 для виявлення автомобілів на аерокосмічному знімку. Автоматизація та біомедичні і комп'ютерні технології: тези доповідей Всеукраїнської науково-технічної інтернет-конференції. (Дніпро, 26 березня 2025 р.) / ДВНЗ «ПДТУ». – Дніпро: ПДТУ, 2025, С.156-158.

141. Радіонов Є. Д., Каштан В. Ю., Гнатушенко В. В., Казимиренко О. В. Свідоцтво про реєстрацію авторського права на твір № 136518 від

23.05.2025 за наукову статтю «Aircraft detection with deep neural networks and contour-based methods». Свідоцтво про реєстрацію авторського права на твір № 136518 від 23.05.2025.

142. Vijayakumar A., Vairavasundaram S. YOLO-based Object Detection Models: A Review and its Applications. *Multimed. Tools Appl.* 2024.

**ДОДАТОК А**  
**СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ АВТОРА ЗА ТЕМОЮ**  
**ДИСЕРТАЦІЇ**

*Публікації у виданнях, включених до переліку наукових фахових видань  
України:*

1. Каштан В.Ю., **Казимиренко О.В.**, Гнатушенко В.В. Порівняльний аналіз ефективності методів розпізнавання автомобілів на аерокосмічних знімках. Control Systems and Computers. 2024. 4. С. 50-60.  
<https://doi.org/10.15407/csc.2024.04.050>.

*Публікації в наукових виданнях, включених до наукометричних баз  
Scopus та/або Web of Sc.:*

2. Radionov Y.D., Kashtan V.Yu., Hnatushenko V.V., **Kazymyrenko O.V.** Aircraft detection with deep neural networks and contour-based methods. The scientific journal «Radio Electronics, Computer Science, Control». – Zaporizhzhia, 2024. №1 4(71). – P.121-129. DOI [10.15588/1607-3274-2024-4-12](https://doi.org/10.15588/1607-3274-2024-4-12). [WoS]
3. Kashtan V. Yu., Hnatushenko V. V., Udovyk I. M., **Kazymyrenko O. V.**, Radionov Y. D. A neural network approach to semantic segmentation of vehicles in very high resolution images. The scientific journal «Radio Electronics, Computer Science, Control». – Zaporizhzhia, 2025. № 3(74). – P.77-85. DOI <https://doi.org/10.15588/1607-3274-2025-3-8> [WoS]
4. Kashtan V. Yu., **Kazymyrenko O.V.**, Hnatushenko V.V. Neural network method for invariant recognition of vehicles in aerospace images. Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu, 2026, № 1. . P.130-137. DOI <https://doi.org/10.33271/nvngu/2026-1/130> [Scopus]

*Публікації у матеріалах наукових конференцій:*

5. Hnatushenko V.V., Kashtan V. Yu., **Kazymyrenko O.V.** Information technology for detecting cars on aerial imaging using a modified YOLO-OBb architecture. MoDaST 2025: Modern Data Science Technologies Doctoral

Consortium, June, 15, 2025, Lviv, Ukraine, Pp. 293-304. <https://ceur-ws.org/Vol-4005/paper20.pdf>. [Scopus]

6. **Казимиренко О.В.**, Гнатушенко В.В. Нейромережеве розпізнавання об'єктів транспорту за даними аерокосмічної зйомки. Матеріали VIII міжнародної науково-технічної конференції «Комп'ютерне моделювання та оптимізація складних систем». 1-3 листопада 2023 року, м. Дніпро, Україна. С.114.
7. **Казимиренко О.В.** Нейромережева модель YOLOv6 для виявлення автомобілів на аерокосмічному знімку. Автоматизація та біомедичні і комп'ютерні технології: тези доповідей Всеукраїнської науково-технічної інтернет-конференції. (Дніпро, 26 березня 2025 р.) / ДВНЗ «ПДТУ».– Дніпро: ПДТУ, 2025, С.156-158.
8. **Казимиренко О.В.** Огляд нейромережевих методів виявлення транспортних засобів. «Наукова весна» 2025: матеріали XV Міжнародної науково-технічної конференції аспірантів та молодих вчених, Дніпро, 26–28 березня 2025 року / Національний технічний університет «Дніпровська політехніка» – Дніпро : НТУ «ДП», 2025, С.187 -188.

*Авторське свідоцтво:*

9. Радіонов Є. Д., Каштан В. Ю., Гнатушенко В. В., **Казимиренко О. В.** Свідоцтво про реєстрацію авторського права на твір № 136518 від 23.05.2025 за наукову статтю «Aircraft detection with deep neural networks and contour-based methods». Свідоцтво про реєстрацію авторського права на твір № 136518 від 23.05.2025.

**ДОДАТОК Б**  
**ДОКУМЕНТИ ЩОДО ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ**  
**ДОСЛІДЖЕНЬ**



**Державне космічне агентство України**  
**«ДНІПРОКОСМОС» ФІЛІЯ НАЦІОНАЛЬНОГО ЦЕНТРУ**  
**УПРАВЛІННЯ ТА ВИПРОБУВАНЬ КОСМІЧНИХ ЗАСОБІВ**

вул. Криворізька, 1, м. Дніпро, 49047, телефон: +380636825630, E-mail: dk@dfdk.dp.ua  
код ЄДРПОУ 38294544

№ \_\_\_\_\_  
На № \_\_\_\_\_ від 15.12.2025р.

**А К Т**

про впровадження наукових результатів дисертаційної роботи Казимиренка Олексія Володимировича на тему «Нейромережеве розпізнавання об'єктів транспорту на аерокосмічних зображеннях», поданої на здобуття наукового ступеня доктора філософії за спеціальністю 126 – Інформаційні системи та технології

Результати досліджень, виконаних у дисертації О.В. Казимиренка, використані у практиці робіт «Дніпрокосмос» філії Національного центру управління та випробувань космічних засобів при розробці та впровадженні методів і технологій нейромережевого виявлення та розпізнавання об'єктів транспорту за даними дистанційного зондування Землі високої просторової роздільної здатності для вирішення прикладних тематичних задач.

У межах дисертаційного дослідження запропоновано новий комплексний багатокомпонентний підхід до нейромережевого розпізнавання транспортних засобів на аерокосмічних знімках високої просторової роздільної здатності, що ґрунтується на інтеграції механізмів просторової нормалізації ознак, виділення областей інтересу, багато масштабного представлення ознак та семантичної сегментації в єдиній нейромережевій архітектурі для розпізнавання об'єктів транспорту. Автором дисертації модифіковано архітектуру YOLOv11 з підтримкою орієнтованих обмежувальних рамок, у якій інтегровано спеціалізовані модулі Freq-SpaFEM, SPPF, DAGLF, AMSF2 та RAFA, що забезпечують ефективне вилучення, уточнення та агрегацію ознак. Реалізація запропонованих архітектурних рішень дозволила підвищити точність і надійність нейромережевого розпізнавання об'єктів транспорту в складних сценах аерокосмічних зображень високої просторової роздільної здатності.

Отримані результати мають практичну цінність та використовуються при виконанні прикладних робіт з оброблення аерокосмічних зображень, а також можуть бути рекомендовані для впровадження в системах дистанційного моніторингу транспортної інфраструктури.

Начальник філії



Віталій КАВАЦ

ЗАТВЕРДЖУЮ



Проректор з науково-педагогічної та  
навчально-виховної роботи

Ігор НІКІТЕНКО

«25» срудня 2025 р.

### АКТ

**впровадження результатів дисертаційного дослідження  
Казмиренка Олексія Володимировича на тему  
«Нейромережеве розпізнавання об'єктів транспорту на аерокосмічних  
зображеннях» в навчальний процес**

Наукові та науково-практичні результати дисертаційного дослідження на тему «Нейромережеве розпізнавання об'єктів транспорту на аерокосмічних зображеннях», що виконувалися протягом 2022-2025 рр., впроваджено в 2025/2026 навчальному році у навчальному процесі на факультеті інформаційних технологій на кафедрі інформаційних технологій та комп'ютерної інженерії при розробці дисциплін: «Обробка зображень та комп'ютерний зір» (проф. Гнатушенко В.В.), «Цифрова обробка просторових даних» (доц. Каштан В.Ю.) для бакалаврів спеціальностей 123 Комп'ютерна інженерія та 126 Інформаційні системи та технології; «Системи штучного інтелекту» (доц. Соколова Н.О.) для бакалаврів спеціальності 126 Інформаційні системи та технології; «Наукові проблеми штучного інтелекту» (проф. Гнатушенко В.В.); для докторів філософії спеціальності 126 Інформаційні системи та технології.

Декан факультету  
інформаційних технологій

Григор Григор Григор Григор Григор  
Грина УДОВИК  
«21» чрудня 2025 р.

Завідувач кафедри  
інформаційних технологій та  
комп'ютерної інженерії

Гнатушенко  
Володимир ГНАТУШЕНКО  
«25» срудня 2025 р.

ДОДАТОК В  
АВТОРСЬКЕ СВДОЦТВО

УКРАЇНА



**СВДОЦТВО**

про реєстрацію авторського права на твір

№ 136518

**Наукова стаття «Aircraft detection with deep neural networks and contour-based methods»**

(вид, назва твору)

**Автор (співавтори) Радіонов Євген Дмитрович, Каштан Віта Юрївна, Гнатушенко Володимир Володимирович, Казимиренко Олексій Володимирович**

(прізвище, ім'я, по батькові (за наявності), псевдонім (за наявності))

**Твір оприлюднено: Оpubлікування: Y. D. Radionov, V. Yu. Kashtan, V. V. Hnatushenko, O. V. Kazymyrenko. Aircraft detection with deep neural networks and contour-based methods // Radio Electronics, Computer Science. - 2024. - № 4. Режим доступу: <https://doi.org/10.15588/1607-3274-2024-4-12>.**

(відомості про факт і дату оприлюднення твору (за наявності))

**Авторські майнові права належать спільно Радіонов Євген Дмитрович, вул. Волонтерів, 14, кв. 25, м. Павлоград, Дніпропетровська обл., 51413; Каштан Віта Юрївна, б-р Слави, 2 А, кв. 88, м. Дніпро, 49130; Гнатушенко Володимир Володимирович, пров. С. Крушельницької, 4/24, м. Дніпро, 49125; Казимиренко Олексій Володимирович, вул. Велика Діївська, 4, корп., 3, кв. 89, м. Дніпро, 49125**

(прізвище, ім'я, по батькові (за наявності) фізичної особи / найменування юридичної особи, адреса)

Дата реєстрації 23 травня 2025 р.

**Директор Державної організації «Український національний офіс інтелектуальної власності та інновацій»**

  
**Олена ОРЛЮК**



УКРНОІВІ