

Міністерство освіти і науки України
Криворізький національний університет
Факультет інформаційних технологій
Кафедра автоматизації, комп'ютерних наук і технологій

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття ступеня вищої освіти – бакалавр
за освітньо-професійною програмою
«Автоматизація та комп'ютерно-інтегровані технології»

зі спеціальності
174 – Автоматизація, комп'ютерно-інтегровані технології та
робототехніка

тема роботи:

***«Дослідження інтелектуальних алгоритмів детекції та
класифікації безпілотних літальних апаратів у складі систем
спостереження»***

Виконав студент гр. АКІТР-23ск _____ Концевич Д. В.

Керівник _____ Тронь В. В.

Нормоконтроль _____ Маринич І. А.

Завідувач кафедри _____ Рубан С. А.

Кривий Ріг – 2026

КРИВОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет: інформаційних технологій

Кафедра: автоматизації, комп'ютерних наук і технологій

Ступінь вищої освіти: Бакалавр

Спеціальність: 174 – Автоматизація, комп'ютерно-інтегровані технології та
робототехніка

ЗАТВЕРДЖУЮ

Зав. кафедрою: к.т.н. Рубан С.А.

« 26 » січня 2026 р.

ЗАВДАННЯ

на кваліфікаційну роботу бакалавра

студентові групи АКІТР-23ск Концевичу Данилу Володимировичу

1. Тема кваліфікаційної роботи: «Дослідження інтелектуальних алгоритмів
детекції та класифікації безпілотних літальних апаратів у складі систем
спостереження»

затверджено наказом по університету № 172с від 30.03.2026 р.

2. Термін здачі кваліфікаційної роботи: 10.06.2026 р.

3. Склад кваліфікаційної роботи: Пояснювальна записка, презентація у
Microsoft PowerPoint в електронному та друкованому вигляді

4. Консультанти кваліфікаційної роботи:

Розділ 1-2

доц. Тронь В. В.

Нормоконтроль

доц. Маринич І. А.

5. Календарний план:

№	Етапи роботи	Термін виконання
1	<i>Вступ</i>	<i>01.03.26</i>
2	<i>Розділ 1</i>	<i>05.04.26</i>
3	<i>Розділ 2</i>	<i>01.05.26</i>
4	<i>Висновки</i>	<i>25.05.26</i>
5	<i>Оформлення кваліфікаційної роботи</i>	<i>28.05.26</i>
6	<i>Підготовка презентації та графічного матеріалу</i>	<i>20.05.26</i>
7	<i>Підготовка доповіді до захисту</i>	<i>05.06.26</i>

6. Дата видачі завдання: 28.01.2026р.

Керівник _____ / Тронь В. В./

7. Запевнення: Я, Концевич Данило Володимирович, запевняю, що ця кваліфікаційна робота виконана самостійно, не містить академічного плагіату, фабрикації, фальсифікації. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело.

Із чинним Положенням про академічну доброчесність Криворізького національного університету ознайомлений.

Чітко усвідомлюю, що в разі виявлення у кваліфікаційній роботі умисних порушень робота не допускається до захисту або оцінюється незадовільно.

Студент _____ / Концевич Д. В./

АНОТАЦІЯ

Концевич Д. В. Дослідження інтелектуальних алгоритмів детекції та класифікації безпілотних літальних апаратів у складі систем спостереження: кваліфікаційна робота бакалавра : 174 – Автоматизація, комп'ютерно-інтегровані технології та робототехніка. Кривий Ріг. Криворізький національний університет, 2026. 63 с.

Метою роботи є дослідження інтелектуальних алгоритмів на основі сучасних методів штучного інтелекту й машинного навчання для детекції й класифікації БПЛА основних типів.

У першому розділі розглянуто сучасні підходи до класифікації та детекції зображень для виявлення БПЛА. Визначено, що ефективність моделей залежить від якості підготовки даних, аугментації, балансування класів та правильного вибору архітектури з урахуванням співвідношення точності й швидкодії. Обґрунтовано доцільність використання моделей YOLO для систем реального часу та перспективність трансформерних архітектур DETR для складних задач детекції.

У другому розділі. У другому розділі проведено порівняльне дослідження моделей YOLO та RtDETR для автоматичного виявлення й класифікації основних типів БПЛА. Встановлено, що YOLOv12s та YOLOv11s показали найкращі результати серед моделей YOLO, особливо для квадрокоптерів, тоді як RtDETR забезпечила високу точність детекції БПЛА літакового типу. Отримані результати підтверджують практичну цінність роботи для систем моніторингу, логування та аналізу польотів БПЛА.

БПЛА, ДЕТЕКЦІЯ ОБ'ЄКТІВ, КЛАСИФІКАЦІЯ ОБ'ЄКТІВ, КВАДРОКОПТЕРИ, КРИТЕРІЇ ОЦІНКИ МОДЕЛЕЙ, МАШИНИЙ ЗІР, МАШИНЕ НАВЧАННЯ, GOOGLE COLAB, VS CODE, PYCHARM.

					КНУ КРБ.174.26.02.00.ПЗ						
Змн.	Арк.	№ докум.	Підпис	Дата							
Розроб.		Концевич Д.В.			АНОТАЦІЯ			Лім.	Арк.	Аркушів	
Перевір.		Трогь В. В.								3	1
Н. Контр.		Маринич І.А.									
Затверд.		Маринич І.А.									
								АКНТ АКІТР-23ск			

ЗМІСТ

ВСТУП.....	6
РОЗДІЛ 1 АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ КЛАСИФІКАЦІЇ ТА ДЕТЕКЦІЇ ОБ'ЄКТІВ НА ЗОБРАЖЕННЯХ.....	10
1.1 Класифікація та детекція об'єктів.....	10
1.2 Огляд та аналіз питань класифікації зображень.....	11
1.3 Огляд і аналіз питань детекції зображень.....	15
1.4 Ключові фактори.....	21
1.5 Якість даних.....	23
Висновки за розділом.....	24
РОЗДІЛ 2 ВИБІР Й НАВЧАННЯ МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ, ІНТЕРПРЕТАЦІЯ РЕЗУЛЬТАТІВ.....	26
2.1 Опис моделі YOLOv10.....	26
2.2 Опис моделі YOLOv11.....	28
2.3 Опис моделі YOLOv12.....	29
2.4 Опис моделі RT-DETR.....	30
2.5 Збір даних та аналіз зображень.....	33
2.6 Огляд результатів навчання моделі YOLOv10.....	38
2.7 Огляд результатів навчання моделі YOLOv11.....	41
2.8 Огляд результатів навчання моделі YOLOv12.....	44
2.9 Огляд результатів навчання моделі RT-DETR.....	46
2.10 Інтерпретація отриманих результатів.....	49
2.11 Застосування навченої моделі RT-DETR на реальних даних.....	53
Висновки до розділу.....	55

					КНУ КРБ.174.26.02.00.ПЗ						
Змн.	Арк.	№ докум.	Підпис	Дата							
Розроб.		Концевич Д.В.			ЗМІСТ			Літ.	Арк.	Аркушів	
Перевір.		Трогь В. В.								4	2
								АКНТ АКІТР-23ск			
Н. Контр.		Маринич І.А.									
Затверд.		Маринич І.А.									

ВИСНОВКИ.....57

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ.....59

ВСТУП

Сучасний розвиток методів штучного інтелекту (ШІ) та машинного навчання (ML) відкриває нові можливості у сфері моніторингу та аналізу повітряного простору. Вдосконалення алгоритмів дозволяє створювати високоточні системи автоматичного виявлення та класифікації безпілотних літальних апаратів (БПЛА), включаючи мультикоптери (наприклад, квадрокоптери) та апарати типу літаків. Актуальність цього завдання суттєво зростає через зростання популярності використання БПЛА як у комерційних, так і в несанкціонованих цілях, що становить значну загрозу безпеці аеродромів, а також промислових об'єктів.

Актуальність розробки моделі машинного зору для виявлення та класифікації БПЛА зумовлена такими практичними проблемами:

1 Під час виробництва у виробника, під час випробувань на носій і приймальні, виникають такі проблеми:

- під час зльоту та посадки параметри тангажу та крену можуть різко змінюватися, наприклад, через різку зміну максимально допустимого значення бокового вітру, внаслідок чого БПЛА може різко зміститися вбік;

- під час польоту, через зниження повітряних потоків, можлива різка зміна висоти або положення БПЛА;

- у «сирому» або розробницькому (тестовому) програмному забезпеченні та математиці спостерігається аномальна поведінка БПЛА, наприклад, можлива зміна траєкторії або «коливання».

2 Під час роботи прототипів критичні системи можуть виходити з ладу, що також може призвести до зміни, наприклад, траєкторії планування.

					КНУ КРБ.174.26.02.00.ПЗ						
Змн.	Арк.	№ докум.	Підпис	Дата							
Розроб.		Концевич Д.В.			ВСТУП			Лім.	Арк.	Аркушів	
Перевір.		Трогь В. В.								6	4
								АКНТ АКІТР-23ск			
Н. Контр.		Маринич І.А.									
Затверд.		Маринич І.А.									

Відмова бортового обладнання або аномальна поведінка БПЛА під час польоту потенційно становлять загрозу безпеці персоналу та цивільних, оскільки зовнішні пілоти та оператори розташовані поблизу стартової точки БПЛА і присутні вздовж злітно-посадкової смуги протягом усього польоту, проводячи зовнішнє спостереження.

Наразі фото- та відеодокументація польоту, наприклад під час випробувань, здійснюється виключно в ручному режимі, коли оператор, незалежно від погодних умов чи часових інтервалів (ранок, день, вечір), записує випробування.

Автоматична система виявлення БПЛА та призначення їх до певного типу вирішила б кілька проблем одночасно:

- утримувати об'єкт дослідження в певних межах, водночас усуваючи проблему відволікання уваги, як у випадку людського фактора;
- автоматизація реєстрації польотних ступенів (прискорення, зліт, політ, посадка);
- стандартизацію випробувань і вилучення візуальних ознак у польоті, а також порівняння вимог до технічних характеристик із фактичними характеристиками;
- можливість детального аналізу подій. Наприклад, під час тестів із незадовільною оцінкою польоту система дозволяє записувати тип продукту, час події, зафіксувати аномальну подію, а також покадровий огляд і аналіз стабілізованого відео, каліброваного вздовж злітно-посадкової смуги у точно визначеному положенні. Це усуває вплив людських помилок, таких як тремтіння камери, зум або нахил;
- у разі втрати зв'язку з БПЛА або його ненормальної поведінки виникає потенційно небезпечна ситуація, особливо якщо БПЛА залишає злітно-посадкову смугу. Завдяки зовнішньому автоматизованому спостереженню оператори можуть залишатися в безпечному місці та отримувати більше інформації про характер польоту дрона.

					КНУ КРБ.174.26.02.00.ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		7

Наукові дослідження в галузі автоматичного виявлення та класифікації БПЛА зосереджені переважно на завданнях загального моніторингу повітряного простору. Однак спеціалізовані підходи, спрямовані на ідентифікацію та розрізнення конкретних типів БПЛА (наприклад, квадрокоптери та авіаційні апарати), що працюють поблизу аеродромів, залишаються недостатньо розвиненими. Це створює потребу у глибшому та більш сфокусованому вивченні проблеми, розробці методів, адаптованих до реальних сценаріїв роботи з різними джерелами даних.

Актуальність роботи пов'язана з необхідністю розробки експлуатаційних та високоточних рішень для виявлення та класифікації БПЛА, які можна інтегрувати в існуючу систему керування об'єктами. Цей напрямок має велике значення на тлі зростання виробництва БПЛА та, як наслідок, їх розробки до випуску продукції.

Предметом роботи є дослідження й вибір найкращих метрик для розв'язання задач детекції і класифікації основних типів БПЛА.

Об'єктом дослідження є основні типи БПЛА (квадрокоптери й БПЛА з фіксованим крилом) та їх візуальні параметри, що використовуються у системах автоматичного виявлення та класифікації.

Метою роботи є дослідження інтелектуальних алгоритмів на основі сучасних методів штучного інтелекту й машинного навчання для детекції й класифікації БПЛА основних типів.

Для досягнення поставленої мети необхідно розв'язати наступні завдання:

Для досягнення поставленої мети необхідно розв'язати такі завдання:

- Провести аналіз науково-технічної літератури щодо сучасних методів класифікації об'єктів на зображеннях.
- Дослідити існуючі підходи до детекції об'єктів на зображеннях із використанням методів штучного інтелекту та машинного навчання.
- Визначити ключові фактори, що впливають на точність виявлення та класифікації БПЛА різних типів.

					КНУ КРБ.174.26.02.00.ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		8

- Сформувати та підготувати набір зображень БПЛА для навчання й тестування моделей машинного навчання.
- Обґрунтувати вибір сучасних моделей машинного навчання для задачі виявлення та класифікації БПЛА, зокрема YOLOv10, YOLOv11, YOLOv12 та RT-DETR.
- Провести навчання обраних моделей на підготовленому наборі даних.
- Виконати порівняльний аналіз результатів навчання моделей за основними метриками якості детекції та класифікації.
- Інтерпретувати отримані результати та визначити найбільш ефективну модель для автоматичного виявлення БПЛА.

Практична цінність роботи полягає у можливості використання розробленого детектора та класифікатора БПЛА для автоматичного виявлення і розпізнавання безпілотних літальних апаратів на зображеннях. Отримані результати можуть бути застосовані для підвищення ефективності випробувань серійних зразків льотної техніки, скорочення часу аналізу візуальних даних, зменшення впливу людського фактора та підвищення безпеки осіб, відповідальних за політ в умовах, наближених до реальних.

					КНУ КРБ.174.26.02.00.ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		9

РОЗДІЛ 1

АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ КЛАСИФІКАЦІЇ ТА ДЕТЕКЦІЇ ОБ'ЄКТІВ НА ЗОБРАЖЕННЯХ

1.1 Класифікація та детекція об'єктів

Перед розглядом сучасних методів детекції та класифікації безпілотних літальних апаратів доцільно проаналізувати науково-технічну літературу, присвячену загальним підходам до обробки зображень, розпізнавання об'єктів та використання алгоритмів штучного інтелекту в задачах комп'ютерного зору. Це дозволяє визначити основні принципи побудови інтелектуальних систем аналізу зображень, їх переваги, обмеження та сфери практичного застосування.

Особливу увагу в межах дослідження необхідно приділити двом взаємопов'язаним напрямкам: класифікації зображень та детекції об'єктів на зображеннях. Класифікація зображень передбачає визначення належності зображення або його фрагмента до певного класу, тоді як детекція спрямована не лише на розпізнавання об'єкта, а й на встановлення його місця розташування на зображенні. Саме поєднання цих підходів є важливим для створення ефективної системи автоматичного виявлення та класифікації БПЛА.

У зв'язку з цим у цьому розділі розглядаються науково-технічні джерела, присвячені сучасним методам класифікації зображень, а також підходам до детекції об'єктів із використанням алгоритмів машинного навчання та глибоких нейронних мереж. Такий аналіз є основою для подальшого вибору моделей, підготовки даних, навчання алгоритмів і оцінювання їх ефективності в задачі автоматичного розпізнавання безпілотних літальних апаратів.

					КНУ КРБ.174.26.02.01.ПЗ						
Змн.	Арк.	№ докум.	Підпис	Дата	РОЗДІЛ 1				Літ.	Арк.	Аркушів
Розроб.		Концевич Д.В.									
Перевір.		Трогь В. В.								10	17
									АКНТ АКІТР-23ск		
Н. Контр.		Маринич І.А.									
Затверд.		Маринич І.А.									

1.2 Огляд та аналіз питань класифікації зображень

В статті [1] представлено новий підхід до масштабування згорткових нейронних мереж (CNN) під назвою Compound Scaling, який дозволяє одночасно збільшувати глибину, ширину та роздільну здатність вхідних зображень з використанням єдиного коефіцієнта масштабування. Цей метод був застосований до архітектури EfficientNet, що призвело до створення серії моделей (від B0 до B7), які демонструють state-of-the-art точність на наборах даних ImageNet, CIFAR-10 та CIFAR-100 при меншій кількості параметрів та обчислювальних витратах. Наприклад, EfficientNet-B7 досяг точності 84.4% Top-1 та 97.1% Top-5 на ImageNet, перевершивши такі моделі, як ResNet, DenseNet та NASNet як за точністю, так і за ефективністю. Основною перевагою EfficientNet є збалансоване масштабування, що забезпечує оптимальне співвідношення між точністю та обчислювальною складністю.

В статті [2] представлена нова архітектура CrossViT, яка покращує Vision Transformer (ViT) за рахунок використання перехресної уваги (крос-уваги) між різними масштабами зображень. CrossViT обробляє зображення на двох масштабах (великому та дрібному) і застосовує механізм крос-уваги для ефективного обміну інформацією між ними. Модель була протестована на наборах даних ImageNet, CIFAR-10/100 та Oxford Flowers-102, показавши state-of-the-art результати: 88.2% Top-1 accuracy на ImageNet, 98.1% на CIFAR-10 та 91.3% на CIFAR-100, а також покращення точності на 2-3% на Oxford Flowers-102 порівняно з ViT. CrossViT демонструє найкраще співвідношення точності та обчислювальних витрат, підтверджуючи ефективність мультимасштабної обробки в архітектурах на основі Transformer.

В статті [3] представлена нова архітектура Swin Transformer, яка об'єднує ієрархічну структуру, характерну для CNN, та механізми уваги, що використовуються в Transformer. Основна інновація полягає в застосуванні зсувних вікон (shifted windows) для ефективного обчислення self-attention, що знижує обчислювальну складність і робить модель більш масштабованою. Swin

					КНУ КРБ.174.26.02.01.ПЗ	Арк.
						11
Змн.	Арк.	№ докум.	Підпис	Дата		

Transformer був протестований на наборах даних ImageNet, COCO та ADE20K, показавши state-of-the-art результати: 87.3% Top-1 accuracy на ImageNet, 58.7 box AP та 51.1 mask AP на COCO, а також 53.5 mIoU на ADE20K. Модель продемонструвала лінійну обчислювальну складність, на відміну від квадратичної у стандартних Transformer, що підтверджує її ефективність та потенціал для задач класифікації, виявлення об'єктів та семантичної сегментації.

В статті [4] представлено всебічний огляд глибоких згорткових нейронних мереж (CNN) для задач класифікації зображень. Автори розглянули еволюцію архітектур CNN, починаючи з LeNet і закінчуючи сучасними моделями, такими як ResNet, DenseNet та EfficientNet, а також ключові компоненти CNN, включаючи згорткові шари, пулінг, функції активації та методи регуляризації. У статті проаналізовано застосування CNN у задачах класифікації, виявлення об'єктів та сегментації, а також проведено порівняння різних архітектур за точністю, обчислювальною складністю та ефективністю. Набори даних, такі як ImageNet, CIFAR-10/100 та MNIST, використовувалися для оцінки моделей. Результати показали, що сучасні CNN, такі як ResNet та EfficientNet, досягли state-of-the-art точності (до 84.4% Top-1 accuracy на ImageNet), при цьому EfficientNet виділяється найкращим співвідношенням точності та обчислювальних витрат. Методи регуляризації, такі як Dropout та Batch Normalization, значно покращують узагальнюваність моделей, а трансферне навчання на великих наборах даних підвищує точність на невеликих наборах. На завершення зазначається, що CNN залишаються основним інструментом для класифікації зображень, але майбутні дослідження мають бути спрямовані на покращення інтерпретованості та ефективності на невеликих наборах даних.

В статті [5] запропоновано метод класифікації зображень зловмисних безпілотних літальних апаратів (БПЛА) з використанням згорткових нейронних мереж (CNN). Автори використовували попередньо навчені моделі, такі як ResNet та VGG, адаптувавши їх для задачі класифікації на основі візуальних

					КНУ КРБ.174.26.02.01.ПЗ	Арк.
						12
Змн.	Арк.	№ докум.	Підпис	Дата		

характеристик БПЛА. Модель була протестована на власному наборі даних, що включає зображення зловмисних та доброякісних БПЛА, а також на додаткових даних з відкритих джерел. Результати показали високу точність (96.8% accuracy) і швидкість обробки, що дозволяє використовувати метод у реальному часі. CNN перевершила традиційні методи машинного навчання, такі як SVM та Random Forest, за точністю та швидкістю, а також продемонструвала стійкість до зміни умов освітлення та кутів зйомки. Цей підхід може бути ефективно застосований для підвищення безпеки та запобігання загрозам, пов'язаним зі зловмисними БПЛА.

В статті [6] представлена нова архітектура MLP-Mixer, яка повністю заснована на багатошарових перцептронах (MLP) і не використовує згорткові шари (CNN) або механізми уваги (Transformer). Архітектура складається з двох типів MLP-шарів: token-mixing MLP для обміну інформацією між патчами зображення та channel-mixing MLP для обміну інформацією між каналами. MLP-Mixer була протестована на наборах даних ImageNet, CIFAR-10/100 та JFT-300M, показавши конкурентоспроможні результати: 87.94% Top-1 accuracy на ImageNet при навчанні на JFT-300M, 94.7% на CIFAR-10 та 81.5% на CIFAR-100. Архітектура продемонструвала, що MLP-мережі можуть досягати високої точності без складних механізмів, таких як CNN або Transformer, що робить їх простішими та універсальнішими. MLP-Mixer відкриває нові напрямки для досліджень, показуючи, що state-of-the-art результати в задачах комп'ютерного зору можуть бути досягнуті з мінімальними індуктивними припущеннями.

В статті [7] проведено огляд застосування глибокого навчання (DL) в задачах дистанційного зондування з використанням безпілотних літальних апаратів (БПЛА). Розглянуто методи DL, такі як CNN, RNN та GAN, для вирішення задач класифікації, виявлення об'єктів, сегментації та 3D-реконструкції, а також особливості даних, включаючи високу роздільну здатність, часові ряди та мультиспектральні дані. CNN показали високу точність (до 95%) у класифікації землекористування, методи на основі Faster R-CNN та YOLO досягли до 80% mAP у виявленні об'єктів, а U-Net та його

					КНУ КРБ.174.26.02.01.ПЗ	Арк.
						13
Змн.	Арк.	№ докум.	Підпис	Дата		

модифікації продемонстрували найкращі результати в сегментації. Глибоке навчання також покращило точність 3D-реконструкції. Незважаючи на значні успіхи, залишаються виклики, такі як обробка часових рядів та мультиспектральних даних, що відкриває напрямки для майбутніх досліджень.

В статті [8] запропоновано гібридний підхід, що поєднує глибоке навчання (CNN) та алгоритм оптимізації Pelican Optimization Algorithm (POA) для класифікації зображень, отриманих з безпілотних літальних апаратів (БПЛА), у контексті M2M (machine-to-machine) комунікацій. Цей підхід досяг 92.5% асигуру на наборі даних UAVDT, перевершивши традиційні методи, і покращив швидкість обробки на 15-20% завдяки оптимізації з використанням POA. Модель успішно інтегрована в систему M2M комунікацій, забезпечуючи класифікацію в реальному часі. Гібридний підхід показав найкращі результати порівняно з CNN та іншими методами оптимізації, демонструючи високу ефективність для задач класифікації зображень з БПЛА та їх застосування в реальних умовах.

В статті [9] запропоновано покращену систему виявлення вторгнень (IDS) для багатокласової класифікації атак у мережах безпілотних літальних апаратів (БПЛА). Використовуючи глибоке навчання (CNN та RNN) та методи оптимізації, система досягла 98.2% асигуру на наборі даних CICIDS2017, перевершивши традиційні методи, і покращила швидкість обробки на 10-15%. Система успішно класифікувала різні типи атак, такі як DDoS, MITM та spoofing, з високою точністю. Покращена IDS показала найкращі результати порівняно з існуючими методами, що робить її ефективним інструментом для підвищення безпеки та надійності мереж БПЛА.

Стаття [10] представляє покращену версію моделі YOLOv7, названу ATS-YOLOv7, для виявлення об'єктів на аерофотознімках з БПЛА в реальному часі. Основні покращення включають багатомасштабне виявлення для роботи з об'єктами різних розмірів, оптимізацію архітектури для підвищення швидкості та точності, а також адаптацію до даних БПЛА. Модель показала вищу точність (mAP) і швидкість (FPS) порівняно з базовою YOLOv7 та іншими методами,

					КНУ КРБ.174.26.02.01.ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		14

такими як Faster R-CNN та SSD, особливо в задачах виявлення дрібних об'єктів. ATS-YOLOv7 підходить для застосування в моніторингу дорожнього руху, сільському господарстві та пошуково-рятувальних операціях, забезпечуючи ефективний аналіз даних з БПЛА в реальному часі.

1.3 Огляд і аналіз питань детекції зображень

Детекція об'єктів на зображеннях - це ключова задача комп'ютерного зору, яка полягає у виявленні та локалізації об'єктів на зображенні. Зазвичай детекція об'єктів відбувається за допомогою обмежувальних рамок (bounding boxes), проте у складніших випадках задача може включати також сегментацію об'єктів (піксельну маску). Сучасні методи на основі глибокого навчання використовують згорткові нейронні мережі (CNN), які автоматично вилучають ознаки та навчаються на великих обсягах даних. Ці методи значно перевершують традиційні за точністю та стійкістю.

Двоетапні методи (Two-Stage Detectors):

R-CNN (Region-based CNN) - спочатку виділяються регіони-кандидати, а потім класифікуються.

Fast R-CNN - покращення R-CNN, де згорткові обчислення використовуються один раз для всього зображення.

Faster R-CNN - вводить Region Proposal Network (RPN) для генерації регіонів-кандидатів, що прискорює процес.

Mask R-CNN - розширення Faster R-CNN для одночасної детекції та сегментації об'єктів.

Одноетапні методи (Single-Stage Detectors):

YOLO (You Only Look Once) - ділить зображення на сітку та передбачає bounding boxes і класи за один прохід.

SSD (Single Shot MultiBox Detector) - використовує кілька шарів CNN для детекції об'єктів різних масштабів.

					КНУ КРБ.174.26.02.01.ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		15

RetinaNet - вводить фокальну функцію втрат (focal loss) для боротьби з дисбалансом класів.

Трансформерні методи для детекції:

DETR (DEtection Transformer) - використовує архітектуру трансформерів для детекції об'єктів, усуваючи необхідність в anchor boxes та NMS (Non-Maximum Suppression).

Основні показники оцінки детекції об'єктів

Для оцінки якості детекції об'єктів використовуються такі метрики:

- Precision (Точність):

$$Precision = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Positives\ (FP)} \quad (1.1)$$

- Recall (Повнота):

$$Recall = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Negatives\ (FN)} \quad (1.2)$$

- F1-Score:

$$F1 - Score = 2x \frac{Precision * Recall}{Precision + Recall} \quad (1.3)$$

- IoU (Intersection over Union) — міра перетину між передбаченим і істинним обмежувальним прямокутником:

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \quad (1.4)$$

- mAP (mean Average Precision) - середня точність (AP) за всіма класами, усереднена за різними порогоми IoU. Це основний показник для оцінки детекторів.

- AR (Average Recall) - середня повнота за всіма класами.

У роботі [11] представляє YOLOv7 - нову модель для виявлення об'єктів, яка перевершує всі відомі аналоги за швидкістю (від 5 до 160 FPS) та точністю (56.8% AP) на GPU V100. YOLOv7-E6 (56 FPS, 55.9% AP) значно випереджає

					КНУ КРБ.174.26.02.01.ПЗ	Арк.
						16
Змн.	Арк.	№ докум.	Підпис	Дата		

трансформерні та згорткові моделі, такі як SWIN-L Cascade-Mask R-CNN та ConvNeXt-XL, а також інші версії YOLO (YOLOR, YOLOX, YOLOv5). Модель навчалася виключно на MS COCO без попередніх ваг, а її архітектура включає інновації, такі як метод trainable bag-of-freebies, що вирішує проблеми динамічного призначення міток.

Робота [12] представляє YOLOX - нову версію детектора об'єктів, яка перевершує попередні моделі YOLO та інші сучасні детектори за балансом швидкості та точності. YOLOX використовує такі інновації, як anchor-free підхід, decoupled head та просунуті методи призначення міток, що дозволяє їй досягти найкращих результатів на наборі даних COCO. Наприклад, YOLOX-L показує 50.0% AP, перевершуючи YOLOv3 на 3.0% AP. Модель також посіла 1-е місце на Streaming Perception Challenge (CVPR 2021), демонструючи ефективність у задачах з урахуванням затримок. YOLOX пропонує покращений баланс швидкості та точності для різних розмірів моделей, що робить її придатною як у промисловості, так і в дослідженнях.

Стаття [13] представляє Anchor-free Oriented Proposal Generator (AOPG) - новий метод для детекції орієнтованих об'єктів, який повністю виключає використання горизонтальних bounding box, що покращує точність виявлення. AOPG демонструє state-of-the-art результати на наборах даних DOTA (80.66% mAP з ResNet50-FPN) та HRSC2016 (96.22% mAP за метрикою VOC12). Метод також включає Coarse Localization Module (CLM) та його покращену версію CLM+, які генерують точніші орієнтовані пропозиції (proposals), підвищуючи recall та mAP. AOPG ефективний у складних сценах із щільно розташованими об'єктами, такими як мости та великі транспортні засоби. У статті також представлено новий набір даних DIOR-R для задач орієнтованої детекції. Результати підтверджують, що AOPG значно перевершує існуючі методи, встановлюючи новий стандарт у цій галузі.

Стаття [14] пропонує новий підхід до покращення продуктивності детекторів об'єктів за рахунок уніфікації «голів» (heads) з використанням механізмів уваги. Автори вводять модуль Dynamic Head, який об'єднує три типу

					КНУ КРБ.174.26.02.01.ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		17

уваги: масштабну (scale-aware), просторову (spatial-aware) та задачно-залежну (task-aware), що дозволяє моделі краще враховувати контекст та взаємозв'язки між об'єктами. Експерименти на наборі даних COCO демонструють, що Dynamic Head значно покращує точність виявлення об'єктів (до 3.6 AP порівняно з базовою моделлю) і може бути легко інтегрований у різні архітектури детекторів, такі як Faster R-CNN, RetinaNet та FCOS. Підхід встановлює новий state-of-the-art на COCO, досягаючи 54.0 AP, і показує високу ефективність як за точністю, так і за обчислювальними витратами, що робить його універсальним і потужним інструментом для задач виявлення об'єктів.

Стаття [15] пропонує інноваційний підхід для детекції дрібних об'єктів на зображеннях високої роздільної здатності, вирішуючи проблему високої обчислювальної складності традиційних методів. Основна ідея полягає у використанні каскадних розріджених запитів (Cascaded Sparse Query, CSQ), які фокусуються лише на релевантних областях зображення, ігноруючи неінформативні частини. Метод складається з двох етапів: coarse detection для швидкого сканування зображення з низькою роздільною здатністю та fine detection для уточнення деталей у вибраних областях. QueryDet демонструє значне покращення точності (на 5-10% вище mAP на наборі даних VisDrone) і швидкості (у 2-3 рази швидше) порівняно з методами на кшталт Faster R-CNN та YOLO, що робить його придатним для задач реального часу, таких як обробка аерофотознімків або супутникових даних. Однак метод вимагає складної реалізації та залежить від якості початкових пропозицій. Загалом, QueryDet пропонує ефективне рішення для детекції дрібних об'єктів, поєднуючи високу точність та обчислювальну ефективність.

Стаття [16] присвячена розробці високоефективної архітектури Vision Transformer (ViT), яка поєднує високу точність і швидкість, порівнянну з легковагомими CNN, такими як MobileNet. Основна мета — зробити ViT придатними для мобільних пристроїв, де важливі низька затримка та продуктивність. EfficientFormer розділений на два блоки: 4D-блок (згорткові шари для обробки просторових даних) та 3D-блок (Multi-Head Self-Attention для

					КНУ КРБ.174.26.02.01.ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		18

глобальних залежностей). Ключові інновації включають використання згорткового stem (зменшує затримку на 48%) та оптимізацію архітектури для мобільних пристроїв. На наборі даних ImageNet-1K EfficientFormer-L1 досягає 80.2% точності із затримкою 1.6 мс, а EfficientFormer-L3 - 82.4% точності із затримкою 3.0 мс на iPhone 12. Метод перевершує MobileNet та інші легковагові CNN за точністю та швидкістю, зберігаючи баланс між продуктивністю та затримкою. EfficientFormer являє собою значний крок вперед, пропонуючи універсальне рішення для мобільних додатків з високими вимогами до швидкості та точності.

Стаття [17] представляє покращену версію моделі YOLO, оптимізовану для детекції об'єктів на аерофотознімках з безпілотних літальних апаратів (БПЛА). Основна мета - створити легковагову та ефективну модель для виявлення дрібних об'єктів на складних фонах, таких як міські пейзажі. YOLO-LE включає ключові модифікації: покращений backbone з модулями C2f та LDown для вилучення дрібнозернистих ознак, neck з модулем AMFF (Adaptive Multi-scale Feature Fusion) для інтеграції ознак різних масштабів та head з LEF-Head, що використовує локальний self-attention для підвищення точності. Модель демонструє значне покращення точності (mAP збільшився з 36.4% до 39.9%), особливо для дрібних об'єктів (наприклад, точність для велосипедистів зросла з 13.9% до 21.5%). При цьому YOLO-LE зберігає низьку обчислювальну складність (4.0 млн параметрів та 8.5 млрд FLOPS), що робить її придатною для пристроїв з обмеженими ресурсами. Візуалізація результатів підтверджує, що YOLO-LE краще справляється з виявленням дрібних об'єктів на складних фонах порівняно з YOLOv8n та іншими моделями. Загалом, YOLO-LE [28] пропонує ефективне рішення для задач детекції на аерофотознімках, поєднуючи високу точність та низьку обчислювальну складність.

Стаття [18] представляє новий еталонний набір даних для задач детекції та відстеження об'єктів на відео, знятих з безпілотних літальних апаратів (БПЛА). Набір даних включає 100 відео (близько 80 000 кадрів) з анотаціями bounding box'ів та атрибутами, такими як погодні умови, висота польоту та кут

					КНУ КРБ.174.26.02.01.ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		19

огляду камери. Він призначений для задач детекції (DET), множинного (MOT) та одиночного відстеження об'єктів (SOT). Сучасні методи показують гірші результати на цьому наборі даних через складність сценаріїв, таких як дрібні об'єкти, складні фони та різкі рухи. Основні напрямки для майбутніх досліджень включають оптимізацію алгоритмів для роботи в реальному часі, врахування контексту сцени, використання інформації про рух та покращення виявлення дрібних об'єктів. Набір даних служить важливою платформою для оцінки та розробки стійкіших методів, здатних справлятися з викликами реального світу.

Стаття [19] представляє огляд сучасних методів глибокого навчання для задач детекції та відстеження об'єктів на даних з БПЛА. Основну увагу приділено наборам даних, таким як VisDrone, та методам, включаючи YOLO, Faster R-CNN, SSD, DeepSORT та SiamRPN, які адаптовані для роботи з дрібними об'єктами та складними сценаріями. Однак точність сучасних методів залишається низькою (наприклад, 65.2% AP для VisDrone-VID). Ключові проблеми включають необхідність оптимізації для роботи в реальному часі, покращення детекції дрібних об'єктів, використання мультисенсорних даних та інтеграцію детекції з відстеженням. Стаття підкреслює, що майбутні дослідження мають бути спрямовані на створення ефективніших і універсальніших методів для задач з БПЛА.

Стаття [20] пропонує спрощений підхід до виявлення об'єктів у виробничій сфері, орієнтований на малі та середні підприємства (SMEs). Основна мета - надати низькозатратне рішення для виявлення дефектів на поверхні об'єктів з використанням низькоякісних зображень та відкритих методів комп'ютерного зору. Автори представляють новий набір даних на основі зображень пластикових цеглин з дефектами, що враховує дисбаланс класів. Для оцінки використовується модель YOLOv5, яка демонструє високу точність на навчальних даних (mAP@0.5: 0.995, mAP@0.5:0.95: 0.858), але показує перенавчання на тестових даних (mAP@0.5: 0.668, mAP@0.5:0.95: 0.507). Найкращі результати на тестових даних досягаються для збалансованого

					КНУ КРБ.174.26.02.01.ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		20

набору (mAP@0.5: 0.774, mAP@0.5:0.95: 0.594). Підхід підтримує принципи FAIR, що робить його доступним для SMEs. У майбутньому планується розширення набору даних та розробка універсальніших методів для застосування в різних виробничих задачах. Загалом, стаття демонструє, що спрощені методи можуть бути ефективними навіть при обмежених ресурсах, що особливо важливо для малих і середніх підприємств.

В статті [21] досліджуються сучасні методи, метрики та набори даних для задачі виявлення об'єктів. Основні результати включають: 1) опис ключових метрик, таких як Intersection over Union (IoU), Average Precision (AP) та Mean Average Precision (mAP), що використовуються для оцінки продуктивності моделей; 2) аналіз популярних наборів даних, включаючи PASCAL VOC та DOTA; 3) огляд еволюції алгоритмів від традиційних методів (Viola-Jones, HOG) до сучасних моделей на основі глибокого навчання (R-CNN, YOLO) з акцентом на їхню точність і швидкість. Виділено поточні проблеми, такі як виявлення дрібних та множинних об'єктів, і запропоновано напрямки для майбутніх досліджень, включаючи багатомасштабні стратегії та механізми уваги.

В статті [22] порівнюються версії YOLOv8 (nano, small, medium, large, extra-large) для виявлення засобів індивідуального захисту (PPE) на наборі даних CHV. YOLOv8l та YOLOv8x показали найвищу точність (mAP), але вимагають більше часу для навчання, YOLOv8n та YOLOv8s є швидшими, але менш точними, а YOLOv8m пропонує баланс між точністю та швидкістю. Моделі краще виявляють «людину» та «жилет», ніж «каска» (особливо «синю»). Дослідження рекомендує вибирати модель залежно від вимог до точності, швидкості та ресурсів.

1.4 Ключові фактори

В основному, БПЛА поділяються на 2 основні типи: квадрокоптери та апарати літакового типу.

					КНУ КРБ.174.26.02.01.ПЗ	Арк.
						21
Змн.	Арк.	№ докум.	Підпис	Дата		

Квадрокоптери є найбільш поширеним видом безпілотних літальних апаратів. Основний принцип їх функціонування базується на використанні кількох гвинтів, установлених на стабілізаційних рамах, які створюють підйомну силу та забезпечують стабільність польоту. Завдяки незалежному управлінню швидкістю обертання кожного гвинта квадрокоптери мають високу маневреність, здатність до різкої зміни траєкторії та можливість тривалого зависання в одній точці повітряного простору.

Безпілотні апарати літакового типу відрізняються наявністю статичних крил. Підйомну силу в цьому випадку створюється за рахунок різниці тиску повітряних потоків над і під поверхнею крила. БПЛА цього типу, як правило, характеризуються більшою швидкістю, тривалістю польоту та передбачуваною прямолінійною траєкторією руху з відносно постійною швидкістю.

Конструктивні та аеродинамічні особливості кожної з зазначених категорій обумовлюють відмінності в їх виявленні, класифікації та прогнозуванні поведінки в повітряному просторі. Так, висока маневреність квадрокоптерів та їхній малий розмір відносно другого класу БПЛА значно ускладнює задачу автоматичної детекції та відстеження, тоді як детектування БПЛА літакового типу виявляється вищим за рахунок більших розмірів та чітких контурів фюзеляжу.

Ці відмінності є критично важливими факторами при виборі методів машинного навчання та стратегії підготовки навчальних даних. Для ефективної роботи систем детекції та класифікації необхідно враховувати як візуальні особливості кожного типу апаратів, так і їхні поведінкові характеристики в динаміці.

У процесі формування навчальних вибірок необхідно забезпечувати репрезентативність даних, що відображають реальні умови експлуатації, включаючи різноманітність типів апаратів, а також складність фонових сцен.

Таким чином, розуміння особливостей функціонування та поведінки різних типів БПЛА є фундаментальним для побудови ефективних систем автоматичного виявлення, класифікації та супроводження безпілотних об'єктів.

					КНУ КРБ.174.26.02.01.ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		22

1.5 Якість даних

Якість навчальних даних є одним із ключових факторів, що визначають ефективність побудови моделей машинного навчання для задач детекції та класифікації БПЛА. При формуванні навчальних вибірок необхідно враховувати такі аспекти:

Роздільна здатність зображень. Висока просторова роздільна здатність зображень суттєво підвищує здатність моделі до виявлення та розрізнення дрібних об'єктів, таких як квадрокоптери, особливо при їх знаходженні на великих висотах. Однак на практиці зображення та відеоматеріали супроводжуються наявністю шуму, розмитості, викликані умовами зйомки, компресією або особливостями обладнання. Наявність шуму погіршує якість ознак, що може негативно позначитися на точності розпізнавання.

Різноманітність даних. Для забезпечення високої узагальнюючої здатності моделі необхідно включати до навчальних наборів зображення, зняті в різних умовах: при різній освітленості (денний та нічний час, сутінки), при різних погодних умовах (дощ, туман, сніг), з різних ракурсів та під різними кутами спостереження. Така різноманітність дозволяє моделі адаптуватися до широкого спектра реальних сценаріїв експлуатації, підвищуючи її стійкість до варіацій зовнішнього середовища та знижуючи ймовірність помилок у нестандартних умовах.

Баланс даних. Незбалансованість класів у навчальному наборі, наприклад, наявність значно більшої кількості зображень одного типу БПЛА порівняно з іншим, призводить до зміщення моделі в бік домінуючого класу та погіршення якості класифікації рідкісних класів. Для мінімізації цього ефекту застосовуються різні підходи, наприклад, такі як аугментація даних - штучне збільшення кількості зображень за рахунок варіативності дій над зображеннями, до прикладу: обертання, віддзеркалення, затемнення або засвітлення тощо.

					КНУ КРБ.174.26.02.01.ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		23

Налаштування гіперпараметрів. Правильне налаштування гіперпараметрів моделі, включаючи швидкість навчання (learning rate), розмір пакета даних (batch size), кількість епох навчання та параметри оптимізаторів, відіграє важливу роль у досягненні високої точності класифікації. Ефективне налаштування гіперпараметрів дозволяє моделі краще узагальнювати знання, мінімізуючи помилку на валідаційних та тестових даних.

Використання предобучених моделей (трансферне навчання). При обмеженому обсязі даних доцільно використовувати методи трансферного навчання, які полягають у донавчанні моделей, попередньо навчених на великих загальнодоступних датасетах (наприклад, ImageNet, COCO). Це дозволяє суттєво прискорити процес навчання, підвищити початковий рівень точності та скоротити вимоги до обсягу навчальних даних.

Якість розмітки даних. Не менш важливим аспектом є якість анотацій навчальних зображень. Помилки в розмітці (наприклад, неточні контури об'єктів, пропущені екземпляри) можуть суттєво знизити якість моделі. Тому критично важливо забезпечити високу точність анотування та проводити регулярну перевірку якості розмітки на всіх етапах підготовки датасету.

Таким чином, комплексний підхід до забезпечення якості даних, що враховує роздільну здатність зображень, їх різноманітність, баланс класів, оптимізацію гіперпараметрів, використання трансферного навчання та високу якість розмітки, є необхідною умовою для ефективного навчання моделей детекції та класифікації безпілотних літальних апаратів.

Висновки за розділом:

У розділі проведено всебічний огляд сучасних підходів до класифікації та детекції зображень, а також їх застосування до задач виявлення та класифікації БПЛА, можна зробити кілька ключових висновків:

по-перше, успішне вирішення таких задач вимагає глибокого розуміння особливостей даних, їхньої специфіки, розмірів об'єктів, мінливості кутів огляду, фонових перешкод та можливих артефактів. Це підкреслює важливість

					КНУ КРБ.174.26.02.01.ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		24

етапу предобробки даних, що включає аугментацію (наприклад, повороти, зміну яскравості, додавання шуму) та баланс класів, щоб модель могла узагальнюватися для різних сценаріїв;

по-друге, вибір архітектури нейронної мережі повинен враховувати компроміс між точністю та швидкістю, особливо якщо система призначена для роботи в реальному часі. Сучасні архітектури, такі як YOLO, демонструють високу ефективність у задачах детекції, але їхня продуктивність може варіюватися залежно від складності даних. При огляді науково-технічної літератури видно явний акцент на застосування моделей сімейства YOLO, які забезпечують високу швидкість обробки при збереженні прийнятної точності, що є критичним для систем реального часу. Також цікавим є використання трансформерів DETR для зображень, які, незважаючи на високі обчислювальні витрати, демонструють чудові результати на складних даних. Однак їх застосування вимагає ретельного налаштування та великих обчислювальних ресурсів.

Для оцінки якості моделей класифікації та детекції БПЛА необхідно використовувати метрики, які відображають як точність виявлення, так і здатність моделі розрізняти об'єкти на фоні шуму. Для класифікації ключовими метриками є accuracy, precision, recall та F1-score, особливо якщо класи незбалансовані. Для задач детекції найбільш важливими метриками є mAP та IoU, які дозволяють оцінити точність локалізації об'єктів та їхньої класифікації. Визначено ключові фактори при класифікації та детекції даних та сформульовані вимоги до якості даних необхідних при навчанні.

РОЗДІЛ 2

ВИБІР Й НАВЧАННЯ МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ, ІНТЕРПРЕТАЦІЯ РЕЗУЛЬТАТІВ

2.1 Опис моделі YOLOv10

Модель YOLOv10 є сучасною одноетапною архітектурою для виявлення об'єктів, яка інтегрує найбільш ефективні методики з попередніх версій сімейства YOLO, а також включає ряд суттєвих удосконалень, спрямованих на оновлення точності розпізнавання та обчислювальної ефективності.

Архітектура моделі YOLOv10, створена на базі пакета Ultralytics, надає покращений підхід до виявлення об'єктів у реальному часі, усуваючи недоліки архітектур минулих моделей, а також усуває немаксимальне подавлення NMS (подавлення максимумів - техніка пост-обробки, за допомогою якої відфільтровуються моделлю після генерації кількох рамок, тобто залишається одна рамка, яка оптимально підходить для об'єкта). Архітектура лінійки моделей YOLOv10 представлена на рисунку 2.1:

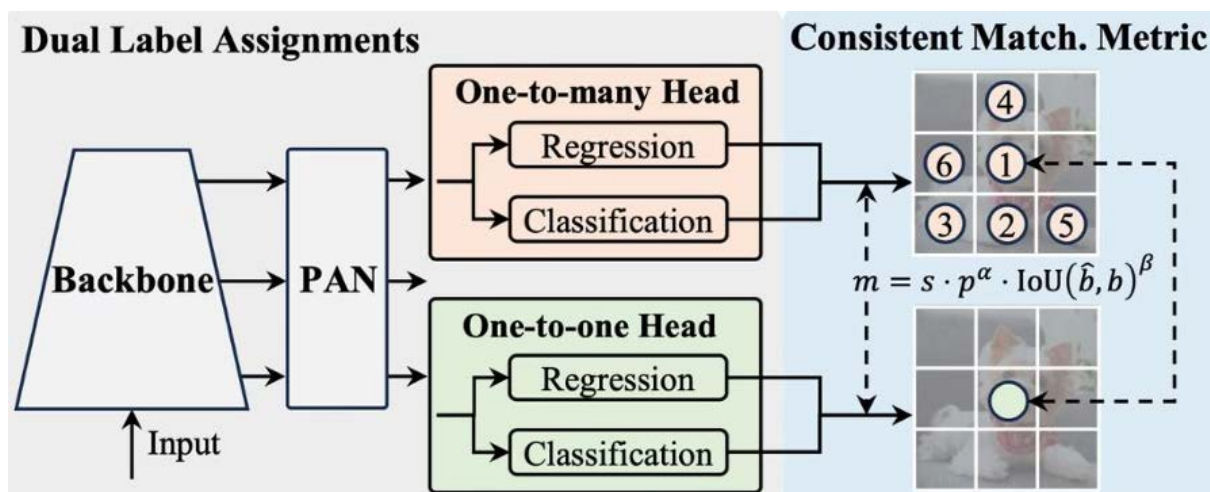


Рисунок 2.1 – Архітектура моделі YOLOv10

					<i>КНУ КРБ.174.26.02.02.ПЗ</i>		
Змн.	Арк.	№ докум.	Підпис	Дата			
Розроб.		Концевич Д.В.			<i>РОЗДІЛ 2</i>		
Перевір.		Трогь В. В.					
Н. Контр.		Маринич І.А.					
Затверд.		Маринич І.А.					
					Літ.	Арк.	Аркушів
						26	31
					<i>АКНТ АКІТР-23ск</i>		

– основна частина, що відповідає за вилучення ознак об'єктів, у YOLOv10 використовує покращену версію CSPNet (Cross Stage Partial Network) для покращення потоку градієнтів та зменшення надмірності обчислень;

– Neck призначена для об'єднання ознак різних масштабів і передачі їх у head. Вона включає в себе шари PAN (Path Aggregation Network) для ефективного об'єднання багатомасштабних ознак;

– модуль «Один до багатьох» генерує кілька прогнозів для кожного об'єкта під час навчання, забезпечуючи насичені сигнали керування та підвищуючи точність навчання;

– модуль «Один до одного» генерує один найкращий прогноз для кожного об'єкта під час логічного висновку (інференсу), усуваючи необхідність в NMS, тим самим скорочуючи час очікування та підвищуючи ефективність.

Лінійка моделей YOLOv10 (демонструє перевагу над сучасними детекторами як за швидкістю, так і за якістю передбачень) представлена у кількох масштабах, що дозволяє вибрати оптимальний варіант під різні умови експлуатації:

– YOLOv10n (Nano) - ультракомпактна версія для пристроїв з обмеженими ресурсами;

– YOLOv10s (Small) - компактна модель із балансом між швидкістю та точністю (працює в 1,8 раза швидше за модель RT-DETR при аналогічному рівні точності на датасеті COCO);

– YOLOv10m (Medium) - універсальна версія для широкого кола задач;

– YOLOv10b (Balanced) - більш широка модель, спрямована на підвищення точності при помірних ресурсних витратах (має на 25% менше параметрів і на 46% меншу затримку, ніж YOLOv9-C, при збереженні ідентичної точності);

– YOLOv10l (Large) - збільшена версія для високоточних застосунків;

					КНУ КРБ.151.25.03.02.ПЗ	Арк.
						27
Змн.	Арк.	№ докум.	Підпис	Дата		

– YOLOv10x (Extra-large) - максимально продуктивна та точна модель (разом з попередньою перевершують відповідні версії YOLOv8 на 0,3–0,5 AP при 1,8–2,3-кратному зниженні кількості параметрів).

2.2 Опис моделі YOLOv11

YOLOv11 вносить значні покращення в архітектуру та методи навчання, що робить його універсальним вибором для вирішення широкого спектра задач комп'ютерного зору. Покращення в архітектурі backbone та neck, що забезпечує більш ефективне вилучення ознак та підвищує точність детекції об'єктів в умовах складних сцен;

Інтегрує інноваційні архітектурні рішення та оптимізовані конвеєри навчання, що сприяє збільшенню швидкості інференсу при збереженні балансу між точністю (accuracy) та продуктивністю (throughput);

Досягає вищого значення середньої точності (mean Average Precision, mAP) на тестовому наборі даних COCO, використовуючи при цьому на 22% менше параметрів порівняно з YOLOv8m;

Підтримує розгортання в гетерогенних середовищах, включаючи периферійні пристрої (edge devices), хмарні платформи (cloud computing) та системи з прискоренням на графічних процесорах NVIDIA, що забезпечує гнучкість у різних сценаріях застосування.

Лінійка моделей YOLOv11:

– YOLOv11n (Nano): це найменший і найшвидший варіант, що містить 2,6 мільйона параметрів. Він призначений для мобільних пристроїв, пристроїв IoT та інших середовищ з обмеженими ресурсами, де важлива швидкість.

– YOLOv11s (Small): цей варіант пропонує хороший баланс між швидкістю та точністю, містить 9,4 мільйона параметрів. Він підходить для виявлення об'єктів у реальному часі на пристроях із помірними обчислювальними ресурсами. PC-YOLOv11s - це покращений метод виявлення невеликих об'єктів на основі моделі YOLOv11. PC-YOLOv11s

					<i>КНУ КРБ.174.26.02.02.ПЗ</i>	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		28

скоригував ієрархічну структуру мережі виявлення та додав шар P2 для виявлення невеликих об'єктів.

– YOLOv11m (Medium): цей варіант пропонує вищу точність, ніж варіанти small та nano, містить 20,1 мільйона параметрів. Він підходить для задач, що потребують вищої точності виявлення та доступних обчислювальних ресурсів.

– YOLOv11l (Large): цей варіант призначений для задач, що потребують високої точності, містить 25,3 мільйона параметрів. Його можна використовувати в середовищах, де обчислювальні ресурси не є проблемою.

– YOLOv11x (Extra Large): це найбільший і найточніший варіант, що містить 56,9 мільйона параметрів. Він призначений для складних задач виявлення об'єктів, де потрібна максимальна точність.

2.3 Опис моделі YOLOv12

YOLOv12 - це сучасна модель детекції об'єктів, заснована на механізмах уваги (attention), що відрізняє її від традиційних CNN-архітектур, використовуваних у попередніх версіях YOLO. Незважаючи на складність, модель зберігає високу швидкість роботи, що робить її придатною для задач реального часу.

Механізм уваги до областей замість стандартного self-attention: застосовується поділ карт ознак на області, що знижує обчислювальні витрати;

R-ELAN (Residual Efficient Layer Aggregation Network) - покращений модуль агрегації ознак, оптимізований для великомасштабних моделей;

Оптимізований механізм уваги включає FlashAttention для прискорення роботи, видалення позиційного кодування та налаштування коефіцієнтів MLP.

Підтримка багатьох задач: детекція, сегментація, класифікація, оцінка пози та OBB (орієнтована обмежувальна рамка / oriented bounding box).

YOLOv12 демонструє перевагу над попередніми версіями (YOLOv10, YOLOv11) та конкурентами (RT-DETR), показуючи приріст точності, а саме:

					КНУ КРБ.174.26.02.02.ПЗ	Арк.
						29
Змн.	Арк.	№ докум.	Підпис	Дата		

YOLOv12n покращує mAP на +2,1% проти YOLOv10n та +1,2% проти YOLOv11n (COCO val2017), а YOLOv12s перевершує RT-DETR на 1,5% за точністю з 42% перевагою у швидкості. Модель зберігає оптимальний баланс між швидкістю та деталізацією, залишаючись гнучкою для розгортання на різних пристроях - від мобільних до хмарних серверів. YOLOv12, як і інші моделі YOLO, підтримує кілька варіантів ваг, оптимізованих під різні задачі:

- YOLOv12n (Nano) - найлегша версія для пристроїв з обмеженими ресурсами.
- YOLOv12s (Small) - баланс швидкості та точності.
- YOLOv12m (Medium) - підвищена точність для більш складних сцен.
- YOLOv12l (Large) - максимальна деталізація, але потребує більше обчислювальних ресурсів.
- YOLOv12x (Extra Large) - найточніша версія для спеціалізованих задач.

2.4 Опис моделі RT-DETR

RT-DETR (Real-Time Detection Transformer) - модель детекції об'єктів на основі трансформерної архітектури, розроблена Baidu для задач, що потребують обробки в реальному часі. На відміну від класичного DETR (Detection Transformer), дана модифікація демонструє підвищену обчислювальну ефективність при збереженні точності детекції за рахунок оптимізованої схеми обробки ознак та модифікованого механізму уваги. Архітектурна схема моделі RT-DETR представлена на рисунку 2.2.

RT-DETR використовує згорткову нейромережу як backbone, зазвичай це ResNet або CSP. Така архітектура дозволяє вилучати ознаки з різних рівнів (наприклад, S3, S4, S5), забезпечуючи багатомасштабне представлення вхідного зображення. Завдяки цьому RT-DETR виявляється легшим і швидшим порівняно з моделями, які повністю засновані на трансформерах.

					КНУ КРБ.174.26.02.02.ПЗ	Арк.
						30
Змн.	Арк.	№ докум.	Підпис	Дата		

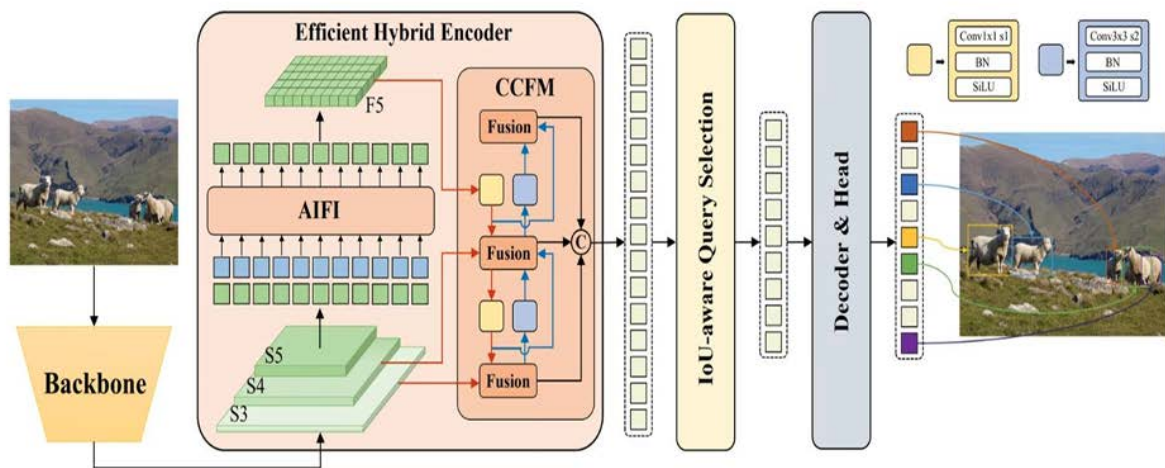


Рисунок 2.2 – Архітектурна схема моделі RT-DETR

Для обробки вилучених ознак застосовується гібридний кодер, який працює у два етапи. Спочатку модуль AIFI аналізує ознаки всередині кожного масштабу, виявляючи контекстні взаємозв'язки. Потім модуль CCFM об'єднує інформацію між різними рівнями, що особливо корисно для коректного розпізнавання як дрібних, так і великих об'єктів. Такий поділ обробки робить кодер менш ресурсомістким порівняно з класичним DETR, знижуючи обчислювальну складність.

На етапі вибору запитів використовується метод, орієнтований на IoU. З усієї карти ознак вибираються найбільш релевантні області - ті, в яких модель очікує високе перетинання з об'єктами (IoU). Ці області стають запитами (object queries), які далі обробляються декодером. Такий підхід покращує початкову точку виявлення об'єктів і сприяє більш точним передбаченням.

Декодер побудований на трансформерній архітектурі та ітеративно уточнює запити. На кожному етапі він покращує як локалізацію об'єктів, так і їх класифікацію. Важливою особливістю є можливість адаптивного вибору кількості шарів у декодері, що дозволяє прискорити модель без необхідності повторного навчання.

На відміну від багатьох інших моделей, RT-DETR не застосовує постобробку у вигляді подавлення перетинних передбачень (NMS). Замість

цього вона напряму видає фінальні передбачення об'єктів та їхні межі, робивши процес більш прямолінійним і швидким.

Модель RT-DETR характеризується рядом архітектурних і функціональних покращень, спрямованих на підвищення ефективності детекції об'єктів при знижених обчислювальних витратах. Одним із ключових компонентів є гібридний кодер, що поєднує внутрішньомасштабну взаємодію ознак (Across-Instance Feature Interaction, AIFI) та міжмасштабне злиття (Cross-Scale Feature Fusion Module, CCFM). Така організація обробки дозволяє ефективно використовувати багаторівневу інформацію, забезпечуючи багатий контекст при мінімальному навантаженні на обчислювальні ресурси.

Механізм вибору запитів, заснований на припущеній величині IoU (Intersection over Union), забезпечує якіснішу ініціалізацію object queries, що сприяє підвищенню точності подальшої детекції. Даний підхід дозволяє декодеру спочатку фокусуватися на найбільш ймовірних областях присутності об'єктів, тим самим покращуючи результативність моделі.

На окрему увагу заслуговує механізм адаптації швидкості виведення. Архітектура декодера допускає динамічну зміну числа шарів без необхідності додаткового навчання. Це надає можливість гнучко налаштовувати модель під конкретні апаратні умови або вимоги до продуктивності, зберігаючи при цьому високу точність.

RT-DETR реалізує повністю end-to-end підхід без використання класичних механізмів, таких як подавлення немаксимумів (Non-Maximum Suppression, NMS) та якоря (anchors). Відмова від цих елементів сприяє спрощенню архітектури та підвищує її узагальнюючу здатність, особливо при перенесенні на нові домени або умови.

У порівнянні з попередніми архітектурами, RT-DETR демонструє кращу масштабованість на прискорених обчислювальних платформах, таких як CUDA з TensorRT, забезпечуючи оптимальний баланс між точністю та швидкістю роботи. Архітектура моделі відрізняється високою адаптивністю,

					КНУ КРБ.174.26.02.02.ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		32

що робить її придатною до широкого спектра задач без необхідності перенавчання.

2.5 Збір даних та аналіз зображень

Набір даних обох типів БПЛА був зібраний з відкритих джерел, включаючи спеціалізовані датасети, відеоматеріали реальних випробувань. Дані містять зображення БПЛА в різних режимах роботи, а саме: у польоті на різних висотах і під різними кутами, на землі, у складних міських умовах забудови, при зльоті та при посадці.

Особливістю набору даних є його різноманітність, обумовлена колірною різноманітністю, геопросторовим положенням, габаритами та умовами освітленості. Приклади зібраних зображень представлені нижче на рисунку 2.3.



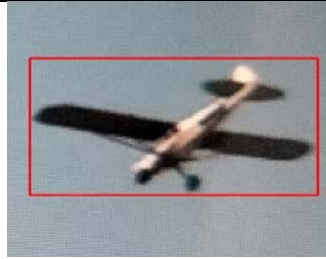
Рисунок 2.3 – Зібраний датасет

Зображення анотовані з використанням платформи Roboflow. Формат анотацій YOLO представлений у таблиці 2.1 і має такий вигляд:

Таблиця 2.1 – Формат анотацій та приклади зображень

Тип БПЛА	X_center	Y_center	Width	Height
0 або 1	Центральна точка прямокутника по осі X	Центральна точка прямокутника по осі Y	Ширина прямокутника	Висота прямокутника

Приклад зображень та формат анотації до нього:



1	0.50338942307	0.48855769237	0.87103365386	0.53026442307
---	---------------	---------------	---------------	---------------



0	0,454427 0,407118	0,454427 0,407118	0,138021 0,144676	0,138021 0,144676
---	----------------------	----------------------	----------------------	----------------------

Враховуючи, що загальний обсяг класів БПЛА типу 1 складається з 1487 зображень, а типу 0 складається з 1528 зображень, можна знехтувати несуттєвою різницею і зробити припущення, що класи є збалансованими. Співвідношення класів зображень представлено на рисунках 2.4 і 2.5 нижче:

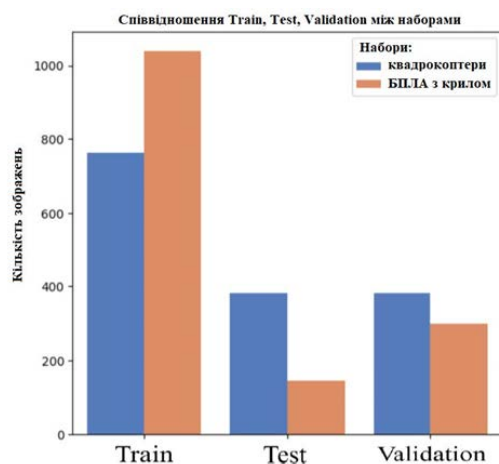


Рисунок 2.4 – Співвідношення наборів Train, Test, Validation

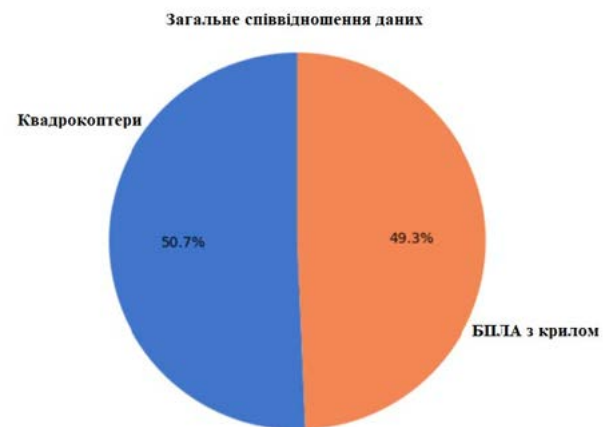


Рисунок 2.5 – Загальне співвідношення даних

У результаті сформовано загальний набір зображень. Набір включає понад 3000 зображень, рівномірно розподілених між двома класами: квадрокоптери (клас 1) та БПЛА літакового типу з фіксованим крилом (клас 0). Таким чином забезпечується баланс класів, що є важливим фактором для

коректного навчання моделей машинного навчання та запобігання зміщенню в бік одного з класів.

Для аналізу коректності обмежувальних рамок було написано код на Python, який зчитує розмітку в анотаціях і наносить її на зображення, назва якого збігається з назвою файлу анотації.

Аналіз набору показав, що існує приблизно 5–7% зображень, у яких обмежувальні рамки не збігаються з фактичним розташуванням об'єкта на зображенні. На рисунках 2.6 – 2.7 наведено приклади таких зображень.



Рисунок 2.6 – Невідповідність Bounding Box з фактичним розташуванням БПЛА



Рисунок 2.7 – Невідповідність Bounding Box з фактичним розташуванням БПЛА

У результаті було реалізовано повторну розмітку зображень з використанням платформи Roboflow та збереження анотацій у правильному форматі YOLO в .txt файлі.

Аналіз характеристик зображень показує наявність значних відмінностей у візуальних ознаках об'єктів двох класів. Наприклад, квадрокоптери зазвичай мають малі габарити і, відповідно, малі розміри обмежувальних рамок на зображеннях. Вони можуть перебувати як на тлі неба, так і в складному міському середовищі, що створює додаткові труднощі у вигляді шуму для їх детектування: наявність безлічі об'єктів та текстур фону призводить до зниження відношення сигнал/шум. На відміну від них, БПЛА

					КНУ КРБ.174.26.02.02.ПЗ	Арк.
						35
Змн.	Арк.	№ докум.	Підпис	Дата		

літакового типу мають більші розміри і займають істотно більшу частину зображення, що полегшує їх візуальне виявлення.

Кількісна характеристика зображень за їхніми розмірами представлена на рисунку 2.8.

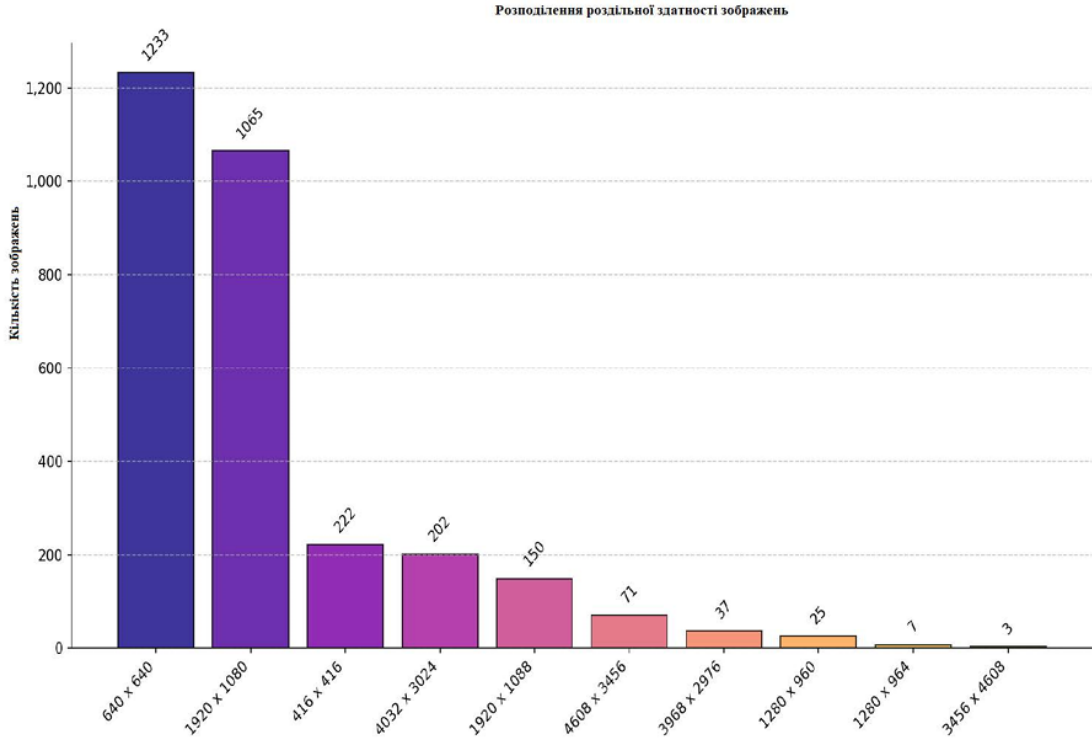


Рисунок 2.8 – Розподіл роздільної здатності зображень у датасеті

Такий розкид розмірів призводить до необхідності стандартизації зображень перед подачею на навчання моделей. Додатково варто зазначити, що дані умови зйомки передбачають значну різноманітність ракурсів, освітлення та якості зображень, що робить задачу більш наближеною до реальних умов експлуатації детектуючих систем.

Таким чином, сформований набір даних має такі характеристики:

- баланс між класами, що мінімізує зміщення моделі;
- наявність об'єктів різної видимості та складності виявлення;
- різноманітність сцен та умов зйомки під різними кутами;
- широкий діапазон роздільних здатностей зображень.

Для забезпечення коректного навчання моделей (YOLOv10, YOLOv11, YOLOv12 та RT-DETR) буде організовано таку структуру датасету:

					КНУ КРБ.174.26.02.02.ПЗ	Арк.
						36
Змн.	Арк.	№ докум.	Підпис	Дата		

```

D:\dataset\
├── train\
│   ├── images\ # Тренувальні зображення
│   └── labels\ # Анотації у форматі YOLO (.txt)
└── val\
    ├── images\ # Валідаційні зображення
    └── labels\ # Анотації для валідації (.txt)

```

Крім цього, існує конфігураційний файл `data.yaml`, який містить інформацію про шляхи до наборів, кількість класів та їхні назви, і має такий вигляд (Рисунок 2.9):

```

train: D:/dataset/train # Шлях до тренувальних зображень
val: D:/dataset/val     # Шлях до валідаційних зображень
# test: D:/dataset/test

nc: 2 # Кількість класів
names: ['quadrocopter', 'fixed-wing'] # Назви класів

```

Рисунок 2.9 – Вміст файлу `data.yaml`

Приведення зображень до єдиного розміру: Оскільки вихідні дані мають різну роздільну здатність, перед подачею на вхід моделям зображення необхідно привести до єдиного стандарту. У межах цієї роботи прийнято рішення масштабувати зображення до роздільної здатності 800×800 пікселів. Такий розмір є розумним компромісом між збереженням деталей дрібних об'єктів (особливо квадрокоптерів) та оптимізацією обчислювальних ресурсів.

При масштабуванні важливо використовувати методи, що зберігають пропорції `bounding boxes`, такі як:

- масштабування із збереженням аспектного співвідношення (пропорцій), доповнене паддингом (наприклад, із заповненням кольором);
- лінійна інтерполяція для зміни розмірів зображень.

Аугментація даних: Для підвищення узагальнюючої здатності моделей

					КНУ КРБ.174.26.02.02.ПЗ	Арк.
						37
Змн.	Арк.	№ докум.	Підпис	Дата		

та покращення їхньої стійкості до змін умов зйомки рекомендується застосовувати методи аугментації. Особливо це актуально для класів з невеликими об'єктами (квадрокоптерів), де можливі проблеми з перенавчанням на певні шаблони. Пропоновані методи аугментації:

- Геометричні трансформації: повороти, відображення, зміни масштабу;
- Фотометричні трансформації: зміна яскравості, контрастності, насиченості та відтінку;
- Додавання шуму: гаусівський шум, розмиття, імітація погіршення якості зображення;
- Random Crop (випадкове кадрування): акуратно, щоб не втратити об'єкти;
- Mosaic-аугментація (особливо актуальна для YOLO): об'єднання кількох зображень в одне для імітації складних сцен.

Адаптація розмітки: При зміні розмірів зображень та застосуванні аугментацій необхідно відповідним чином коригувати координати обмежувальних рамок (bounding boxes), щоб зберегти коректність анотацій.

Розподіл на навчальну та валідаційну вибірки: Набори зображень відповідатимуть пропорціям: 75% для навчання та 25% для валідації.

2.6 Огляд результатів навчання моделі YOLOv10

У межах цього дослідження було обрано дві модифікації моделі - YOLOv10n та YOLOv10s. Обидві моделі навчалися на зображеннях розміром 800×800 пікселів протягом 100 епох. Метрики навчання моделей YOLOv10 представлені на рисунках 2.10 – 2.13.

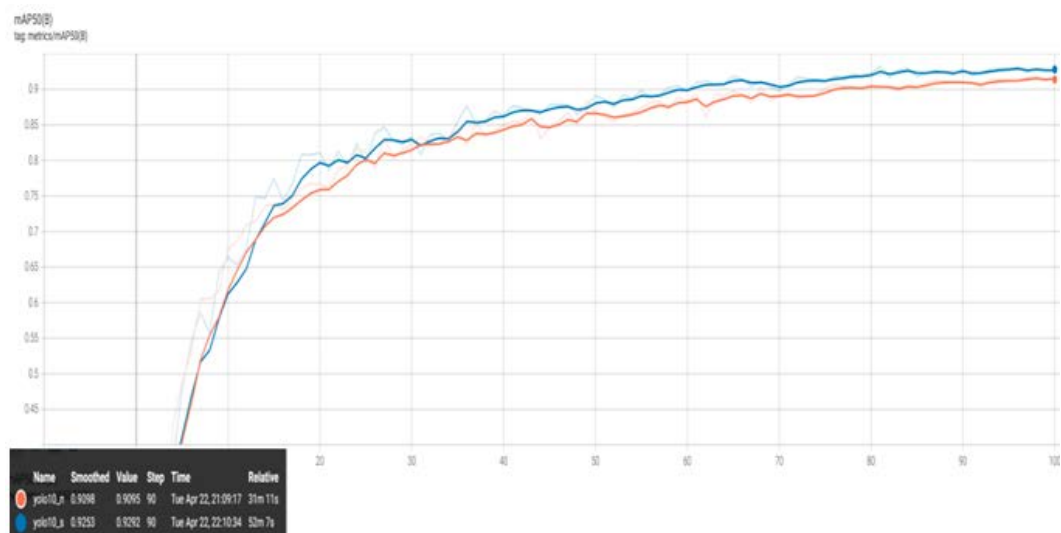


Рисунок 2.10 – Метрика середньої точності mAP@0.5

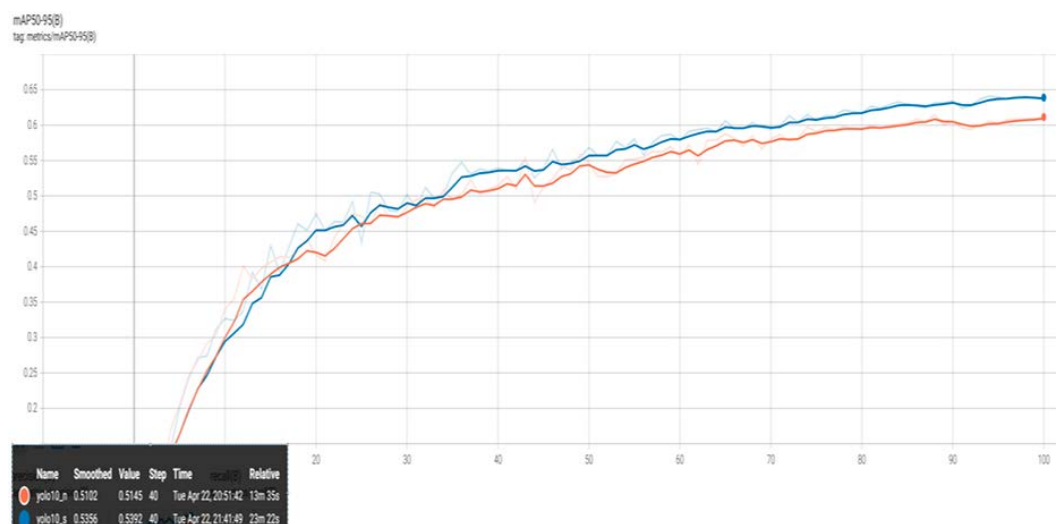


Рисунок 2.11 – Метрика середньої точності mAP@0.5:0.95



Рисунок 2.12 – Метрика точності (precision)

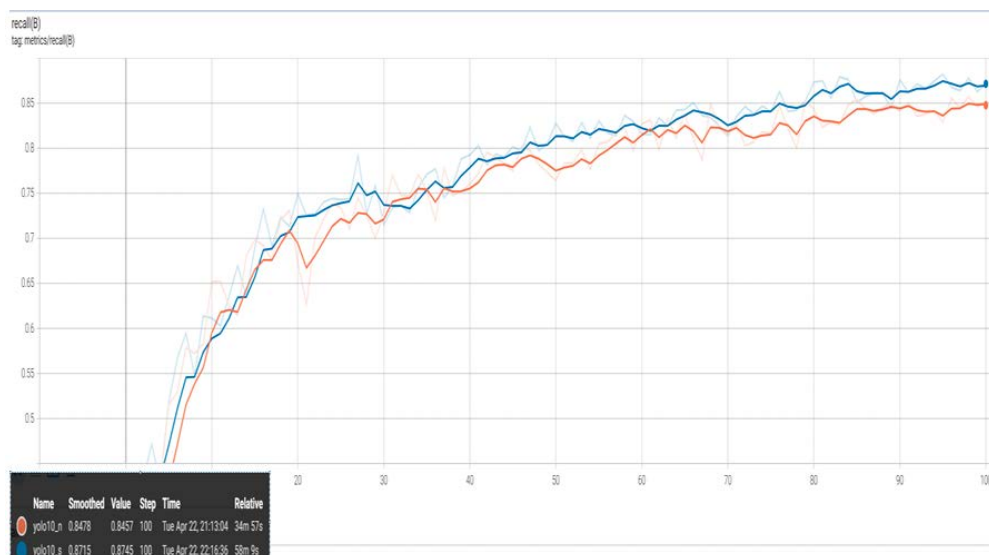


Рисунок 2.13 – Метрика повноти (recall)

Матриці помилок для моделей YOLOv10n та YOLOv10s наведені в таблицях 2.2 і 2.3 відповідно.

Таблиця 2.2 – Матриця помилок для моделі YOLOv10n

Реальний \ Передбачений	quadrocopter	fixed-wing	background
quadrocopter	0.85	-	0.27
fixed-wing	-	0.93	0.73
background	0.15	0.07	-

Таблиця 2.3 – Матриця помилок для моделі YOLOv10s

Реальний \ Передбачений	quadrocopter	fixed-wing	background
quadrocopter	0.87	-	0.39
fixed-wing	-	0.93	0.61
background	0.13	0.07	-

Клас «квадрокоптер» (quadr.): Для моделі YOLOv10n правильна класифікація досягається у 85% випадків, при цьому 27% квадрокоптерів помилково визначаються як фон. Модель YOLOv10s показала невелике покращення точності до 87%, проте частка помилкової класифікації квадрокоптерів як фон збільшилася до 39%.

Клас «БПЛА літакового типу» (fix.wings): Обидва варіанти моделі (YOLOv10n та YOLOv10s) забезпечують однаково високу точність класифікації БПЛА літакового типу — 93%. При цьому YOLOv10s показала зменшення кількості помилок класифікації як фон: 61% порівняно з 73% у YOLOv10n.

Фон (backgr.): У випадку фонових областей частина об'єктів помилково визначається як об'єкти: YOLOv10n помилково класифікує фон як квадрокоптер у 15% випадків і як літак у 7% випадків. YOLOv10s знижує помилку на фоні до 13% для квадрокоптерів, зберігаючи 7% помилкової класифікації як літак.

2.7 Огляд результатів навчання моделі YOLOv11

У межах цього дослідження було обрано дві модифікації моделі - YOLOv11n та YOLOv11s. Обидві моделі навчалися на зображеннях розміром 800×800 пікселів протягом 100 епох. Метрики навчання моделей YOLOv11 представлені на зображеннях 2.14 – 2.17.

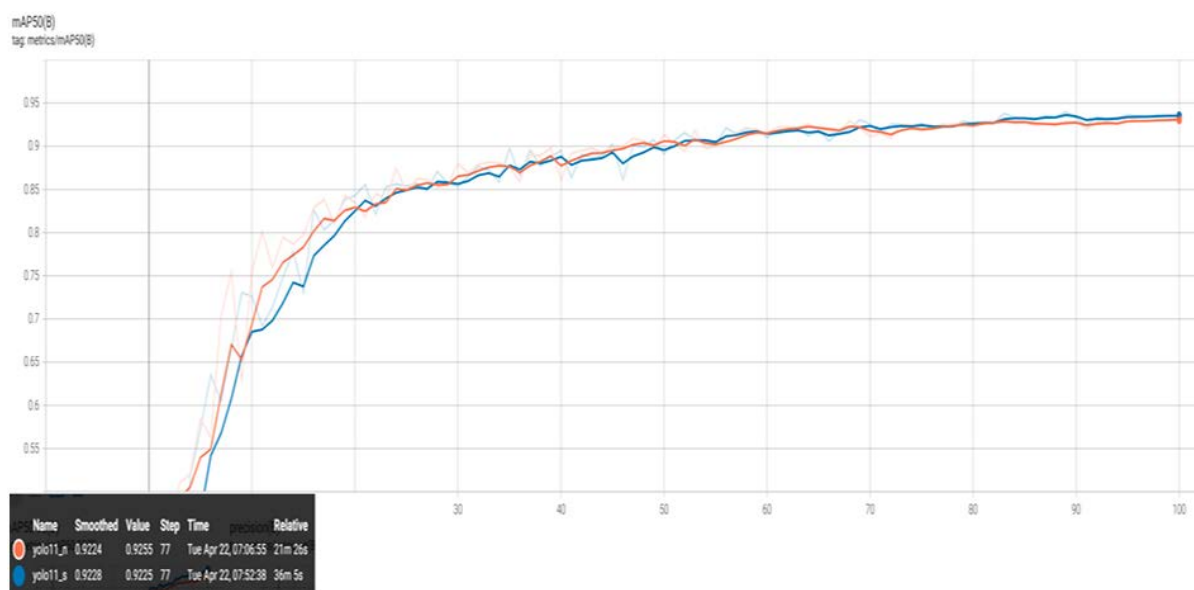


Рисунок 2.14 – Метрика середньої точності mAP@0.5

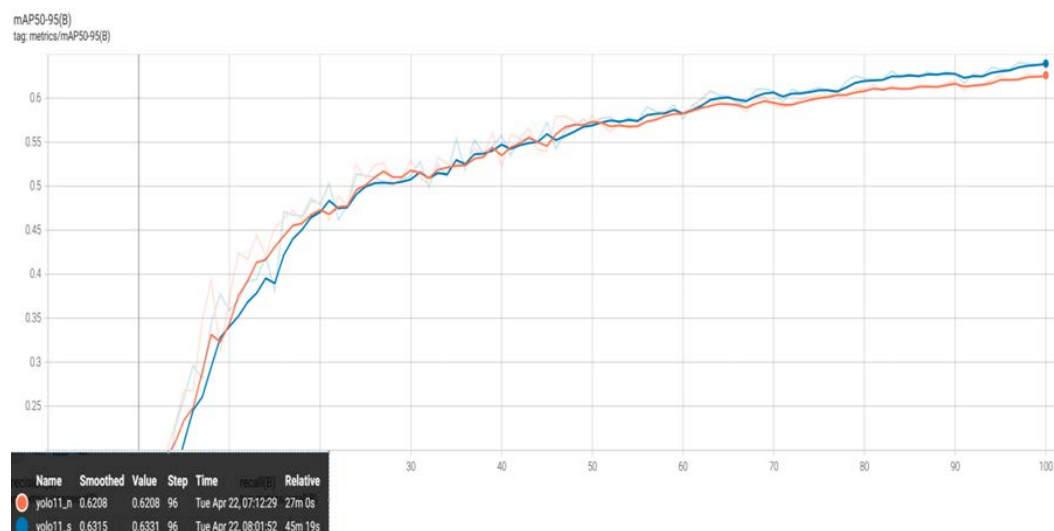


Рисунок 2.15 – Метрика середньої точності mAP@0.5:0.95

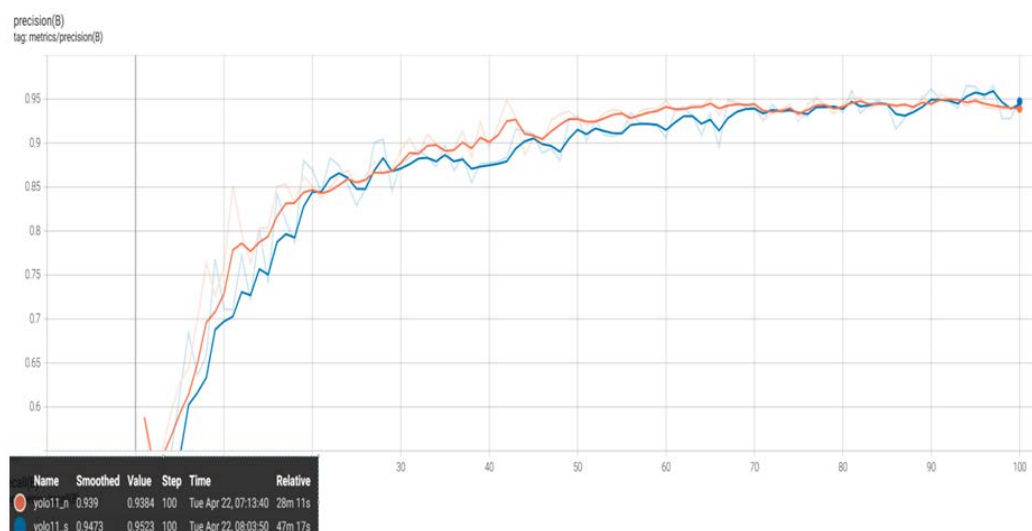


Рисунок 2.16 – Метрика точності (precision)

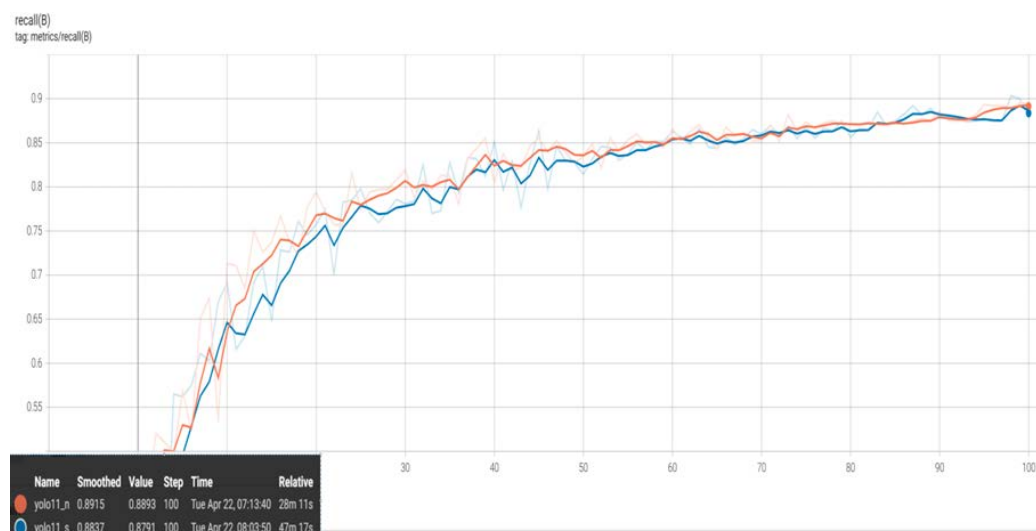


Рисунок 2.17 – Метрика повноти (recall)

Матриці помилок для моделей YOLOv11n та YOLOv11s наведені в таблицях 2.4 і 2.5 відповідно.

Таблиця 2.4 – Матриця помилок для моделі YOLOv11n

Реальний \ Передбачений	quadrocopter	fixed-wing	background
quadrocopter	0.86	0.01	0.20
fixed-wing	-	0.96	0.80
background	0.14	0.04	-

Таблиця 2.5 – Матриця помилок для моделі YOLOv11s

Реальний \ Передбачений	quadrocopter	fixed-wing	background
quadrocopter	0.88	-	0.36
fixed-wing	-	0.96	0.64
background	0.12	0.04	-

Клас «квадрокоптер» (quadr.): Модель YOLOv11n правильно класифікує квадрокоптери у 86% випадків, а YOLOv11s - у 88%. У YOLOv11n з'явилася невелика помилка перекласифікації квадрокоптерів як літаків (1%), якої не було в попередніх моделях. Помилкова класифікація квадрокоптерів як фон зменшилася порівняно з YOLOv10n (20% проти 27%), но для YOLOv11s залишилася на рівні 36% (аналогічно YOLOv10s).

Клас «БПЛА літакового типу» (fix.wings): Обидві моделі YOLOv11 демонструють підвищення точності розпізнавання БПЛА до 96% порівняно з 93% у YOLOv10n та YOLOv10s. Помилкова класифікація літаків як фон також зменшилася: для YOLOv11n - 80% (гірше, ніж 73% у YOLOv10n, але помітно краще, ніж у YOLOv10s - 61%), для YOLOv11s - 64% (краще, ніж 61% у YOLOv10s).

Фон (backgr.): Фон помилково класифікується як об'єкти: у YOLOv11n 14% фону приймаються за квадрокоптери та 4% за літаки; в YOLOv11s ці показники трохи покращуються: 12% та 4% відповідно.

2.8 Огляд результатів навчання моделі YOLOv12

У межах цього дослідження було обрано дві модифікації моделі - YOLOv12n та YOLOv12s. Обидві моделі навчалися на зображеннях розміром 800×800 пікселів протягом 100 епох. Метрики навчання моделей YOLOv12 представлені на рисунках 2.18 – 2.21.

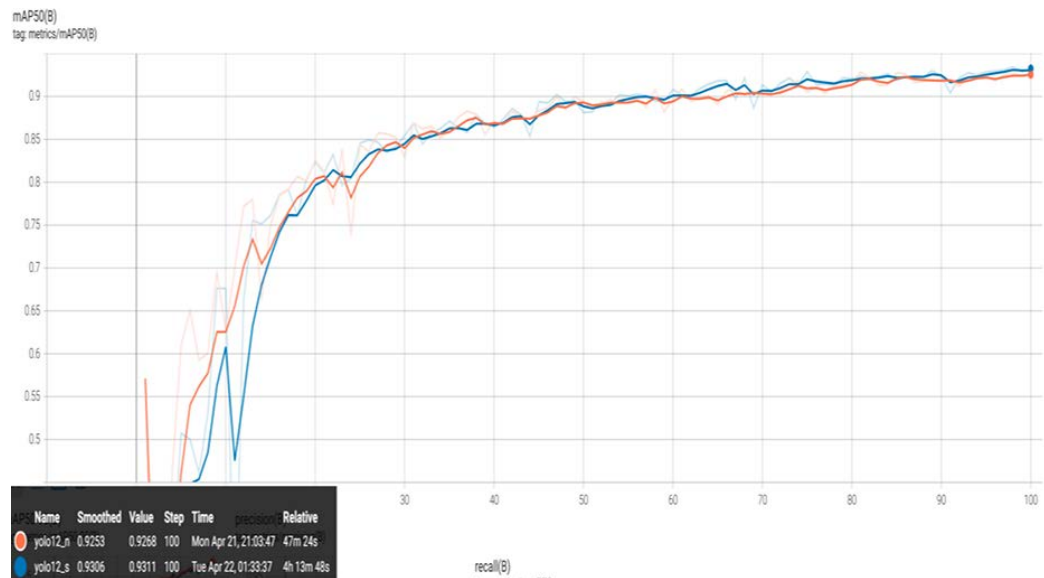


Рисунок 2.18 – Метрика середньої точності mAP@0.5

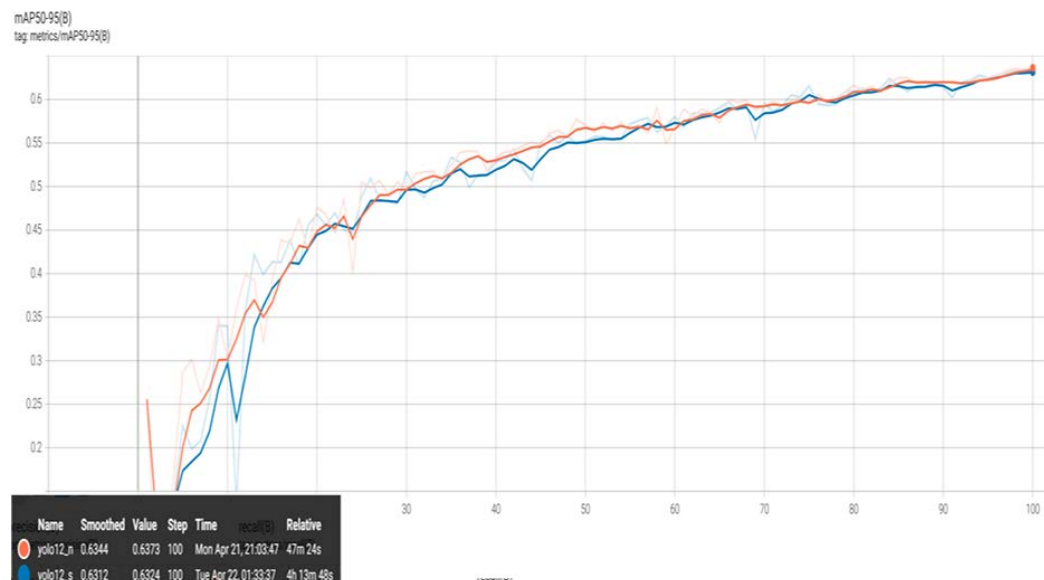


Рисунок 2.19 – Метрика середньої точності mAP@0.5:0.95

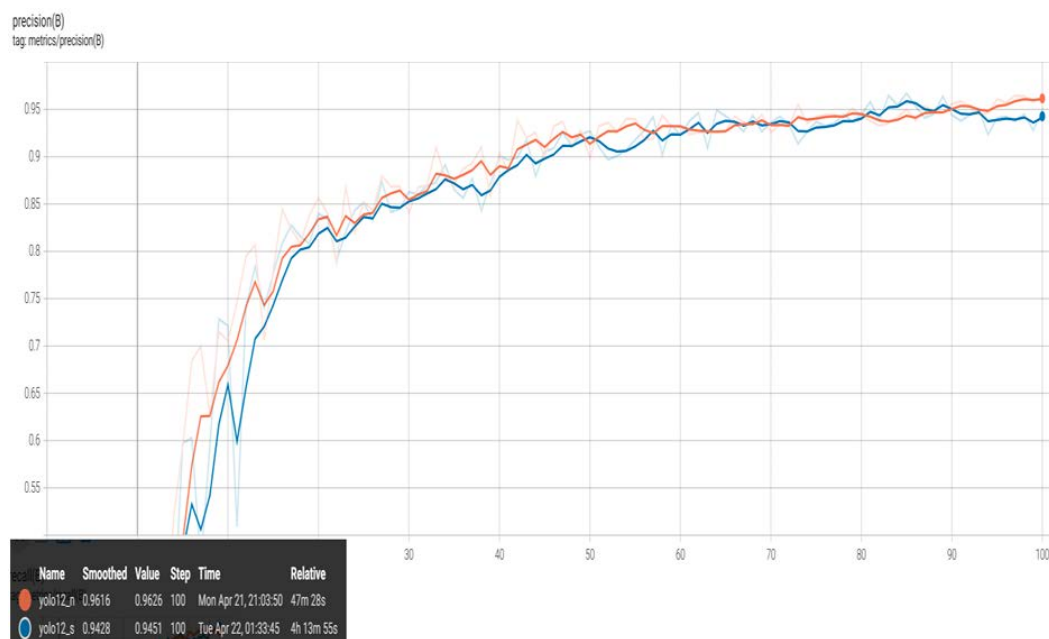


Рисунок 2.20 – Метрика точності (precision)

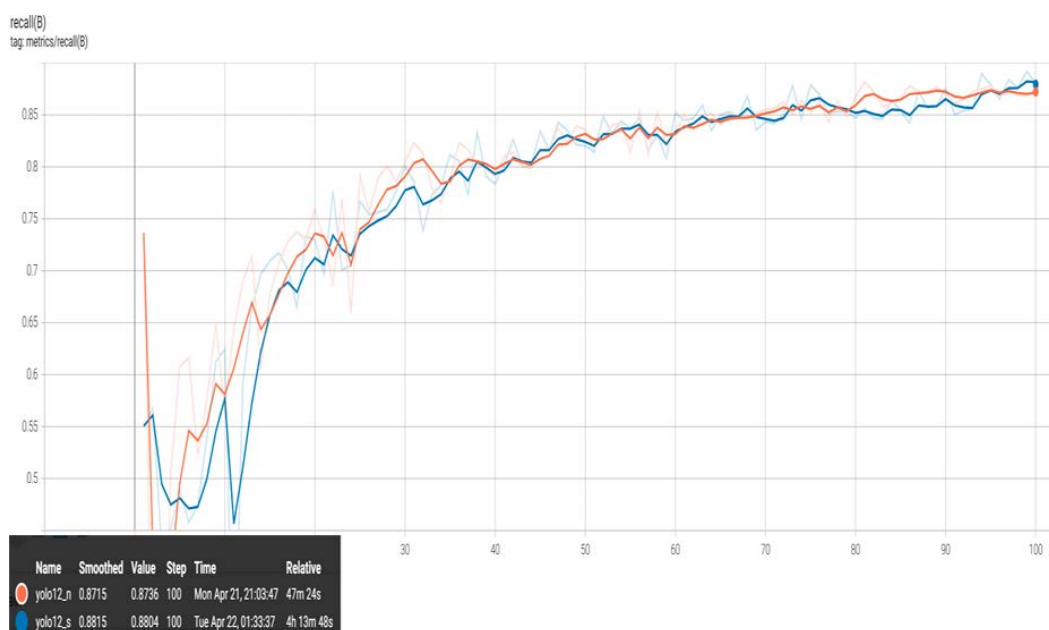


Рисунок 2.21 – Метрика повноти (recall)

Матриці помилок для моделей YOLOv12n та YOLOv12s наведені в таблицях 2.6 і 2.7 відповідно.

Клас «квадрокоптер» (quadr.): Модель YOLOv12n показує правильну класифікацію квадрокоптерів у 85% випадків, при цьому 25% квадрокоптерів помилково класифікуються як фон. Модель YOLOv12s досягає 88% точності з помилкою класифікації квадрокоптерів як фон у 27%, що краще, ніж у

YOLOv11s (36%).

Таблиця 2.6 – Матриця помилок для моделі YOLOv12n

Реальний \ Передбачений	quadrocopter	fixed-wing	background
quadrocopter	0.85	-	0.25
fixed-wing	-	0.96	0.75
background	0.15	0.04	-

Таблиця 7 – Матриця помилок для моделі YOLOv12s

Реальний \ Передбачений	quadrocopter	fixed-wing	background
quadrocopter	0.88	-	0.27
fixed-wing	-	0.95	0.73
background	0.12	0.05	-

Клас «БПЛА літакового типу» (fix.wings): Точність класифікації літаків зберігається на високому рівні: 96% для YOLOv12n та 95% для YOLOv12s. Помилкова класифікація літаків як фон становить 75% для YOLOv12n та 73% для YOLOv12s - трохи гірше, ніж у моделей YOLOv11.

Фон (backgr.): Помилкова класифікація фонових даних у YOLOv12 аналогічна результатам YOLOv11: 15% (YOLOv12n) та 12% (YOLOv12s) фону помилково визначаються як квадрокоптери, 4–5% - як літаки.

2.9 Огляд результатів навчання моделі RT-DETR

Обидві моделі навчалися на зображеннях розміром 800×800 пікселів протягом 100 епох. Метрики навчання моделі RT-DETR представлені на рисунках 2.22 – 2.25.

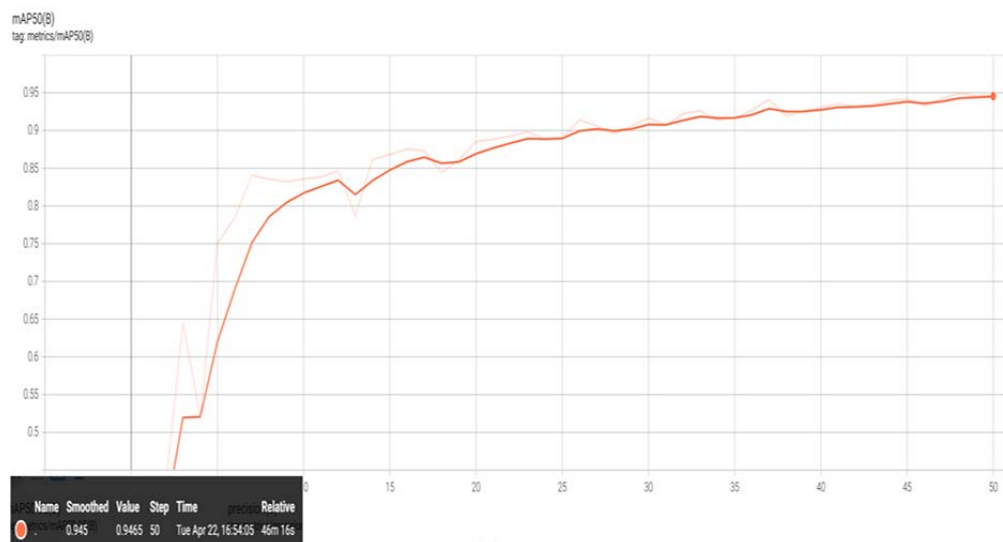


Рисунок 2.22 – Метрика середньої точності mAP@0.5

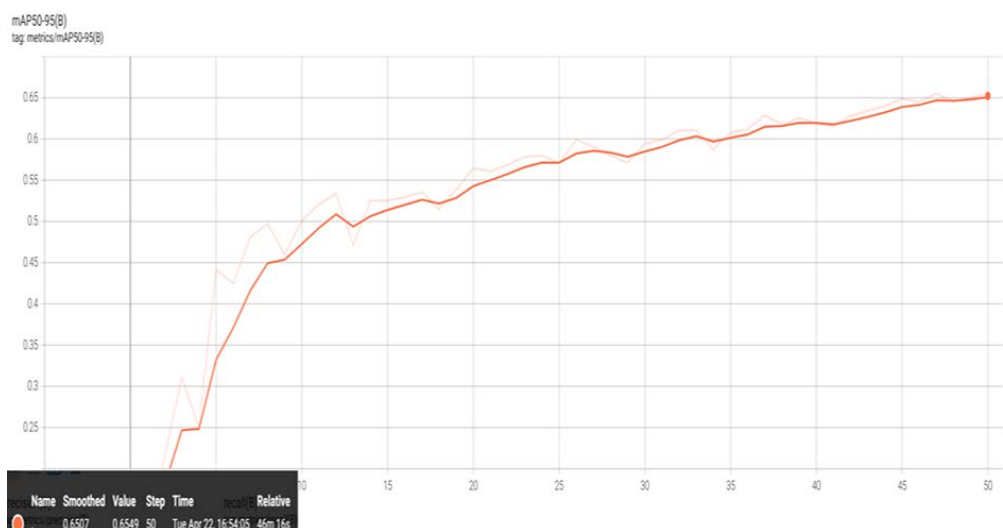


Рисунок 2.23 – Метрика середньої точності mAP@0.5:0.95

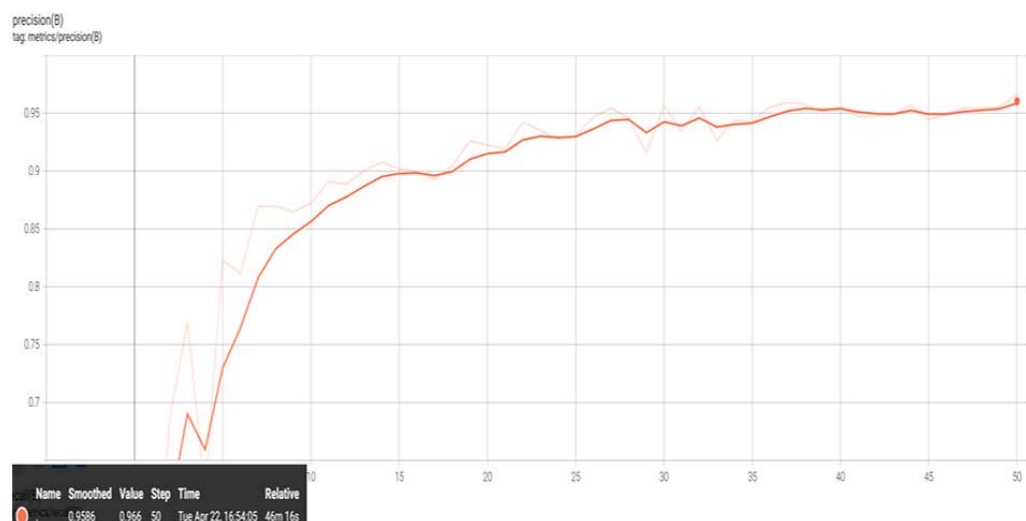


Рисунок 2.24 – Метрика точності (precision)

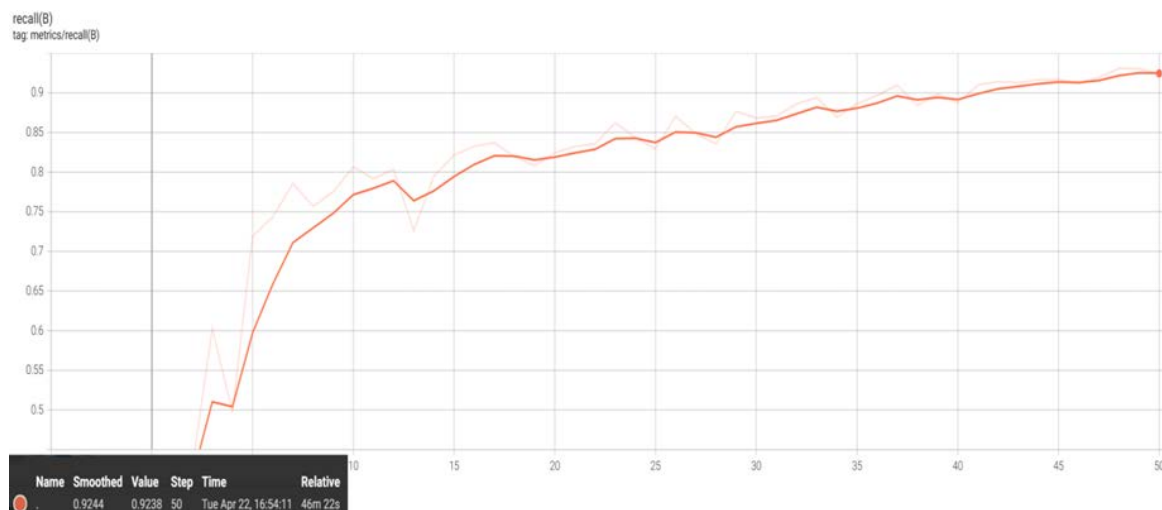


Рисунок 2.25 – Метрика повноти (recall)

Матриця помилок для моделі RT-DETR на 50 епохах наведена в таблиці 2.8.

Таблиця 2.8 – Матриця помилок для моделі RT-DETR

Реальний \ Передбачений	quadrocopter	fixed-wing	background
quadrocopter	0.94	-	0.46
fixed-wing	-	0.95	0.54
background	0.06	0.05	-

Точність класифікації квадрокоптерів у RT-DETR значно вища, ніж у моделей YOLOv12. Помилка «квадрокоптер» як «фон» у RT-DETR є вищою (46%), що свідчить про більшу складність для моделі відрізнити квадрокоптери від фону. Точність класифікації літаків порівнянна з YOLOv12, но помилка «літак» як «фон» є нижчою у RT-DETR (54% проти 73–75%). Помилка на фоні у RT-DETR значно нижча (6–5%) порівняно з YOLOv12 (12–15% для квадрокоптерів).

Порівняння моделей YOLOv10, YOLOv11, YOLOv12 та RT-DETR за точністю наведено в таблиці 2.9.

Таблиця 2.9 – Порівняння метрик моделей YOLO та RT-DETR за точністю

Модель	TP quadr.	TP fix.wings	FN quadr.	FN fix.wings
YOLO10n	0.85	0.93	0.27	0.73
YOLO10s	0.87	0.93	0.39	0.61
YOLO11n	0.86	0.96	0.20	0.80
YOLO11s	0.88	0.96	0.36	0.64
YOLO12n	0.85	0.96	0.25	0.75
YOLO12s	0.88	0.95	0.27	0.73
RT-DETR	0.94	0.95	0.46	0.54

2.10 Інтерпретація отриманих результатів

Статистична значущість

Припустимо, що точність (ассурасу) - це ймовірність успіху в біноміальному експерименті, тобто частка правильно класифікованих об'єктів серед загальної кількості прикладів:

$$\hat{P} = \frac{TP + TN}{N} \quad (2.1)$$

де $\hat{P}$ - оцінка точності, яка задається у відсотках (%); N - загальна кількість зображень у валідації (603 зображення).

Розрахунок 95% інтервалу:

$$CI_{95} = \hat{P} \pm 1.96 * \sqrt{\frac{\hat{P}(1 - \hat{P})}{n}} \quad (2.2)$$

де 1.96 - коефіцієнт для 95% довірчого інтервалу. Для задач детекції об'єктів статистична значущість для моделей оцінюється за основною метрикою mAP@0.5:0.95. 95% довірчі інтервали для моделей наведені в таблиці 2.10.

Таблиця 2.10 – 95% Довірчі інтервали за метрикою $mAP@0.5:0.95$

Модель	$mAP@0.5:0.95$	95%-й довірчий інтервал
yolo10n	0.5145	[0.4743 - 0.5556]
yolo10s	0.5392	[0.4992 - 0.5788]
yolo11n	0.6208	[0.5804 - 0.6600]
yolo11s	0.6331	[0.5937 - 0.6716]
yolo12n	0.6373	[0.5987 - 0.6750]
yolo12s	0.6324	[0.5937 - 0.6700]
RT-DETR	0.6507	[0.6119 - 0.6882]

Графік з 95% довірчими інтервалами для метрики $mAP@0.5:0.95$ за кожною з моделей наведено на рисунку 2.26.

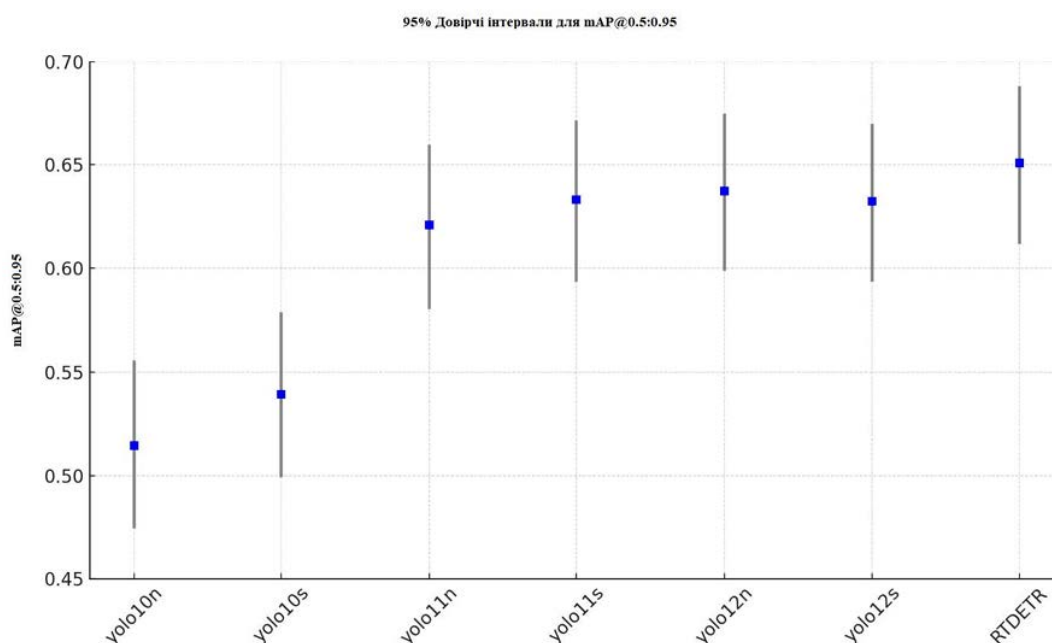


Рисунок 2.26 – Графік метрики $mAP@0.5:0.95$ для кожної моделі

Графік візуалізації 95% довірчих інтервалів для кожної моделі показує:

- в моделях YOLO10n та YOLO10s інтервали нижчі за інші;
- в моделях YOLO11/12 та RT-DETR інтервали перекриваються, що свідчить про відсутність статистично значущих відмінностей між ними.

Таким чином, на рівні значущості моделі YOLO версії 11 і 12 та модель RT-DETR працюють приблизно однаково, проте різниця між ними та YOLO

версії 10 є статистично значущою.

Для оцінки статистичної значущості для задачі класифікації типів БПЛА використовується основна метрика Ассигасу. Тоді отримуємо інтервали точності для класу «квадрокоптери», які наведені в таблиці 2.11 нижче:

Таблиця 2.11 – 95% Довірчі інтервали за точністю для класу квадрокоптери

Модель	Точність для класу квадрокоптери	95% Довірчий інтервал
YOLOv10n	85.00%	[82.15%, 87.85%]
YOLOv10s	87.00%	[82.32%, 89.68%]
YOLOv11n	86.00%	[83.23%, 88.77%]
YOLOv11s	88.00%	[85.41%, 90.59%]
YOLOv12n	85.00%	[82.15%, 87.85%]
YOLOv12s	88.00%	[85.41%, 90.59%]
RT-DETR	94.00%	[92.10%, 95.90%]

Наочна демонстрація розподілу 95% довірчих інтервалів за метрикою Ассигасу для класу «Квадрокоптери» наведена на рисунку 2.27.

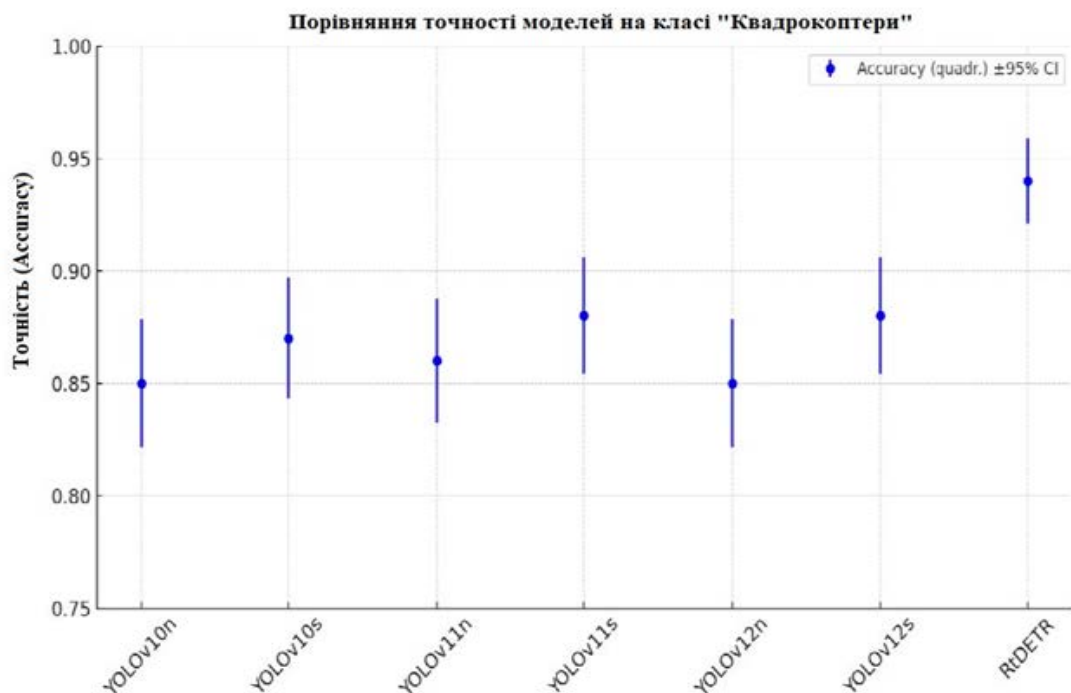


Рисунок 2.27 – Графік порівняння точності моделей на класі «квадрокоптери» з 95% довірчими інтервалами

Видно, що модель RT-DETR має значно вищу точність, ніж YOLO-моделі. Решта моделей відрізняються між собою незначно (їхні інтервали перекриваються). І також отримуємо інтервали точності для класу «БПЛА з фіксованим крилом» у таблиці 2.12.

Таблиця 2.12 – 95% Довірчі інтервали за точністю для класу БПЛА з фіксованим крилом

Модель	Точність для класу	95% Довірчий інтервал
YOLOv10n	93.00%	[90.96%, 95.04%]
YOLOv10s	93.00%	[90.96%, 95.04%]
YOLOv11n	96.00%	[94.44%, 97.56%]
YOLOv11s	96.00%	[94.44%, 97.56%]
YOLOv12n	96.00%	[94.44%, 97.56%]
YOLOv12s	95.00%	[93.26%, 96.74%]
RT-DETR	95.00%	[93.26%, 96.74%]

Наочна демонстрація розподілу 95% довірчих інтервалів за метрикою Ассигасу для класу «БПЛА літакового типу» наведена на рисунку 2.28.

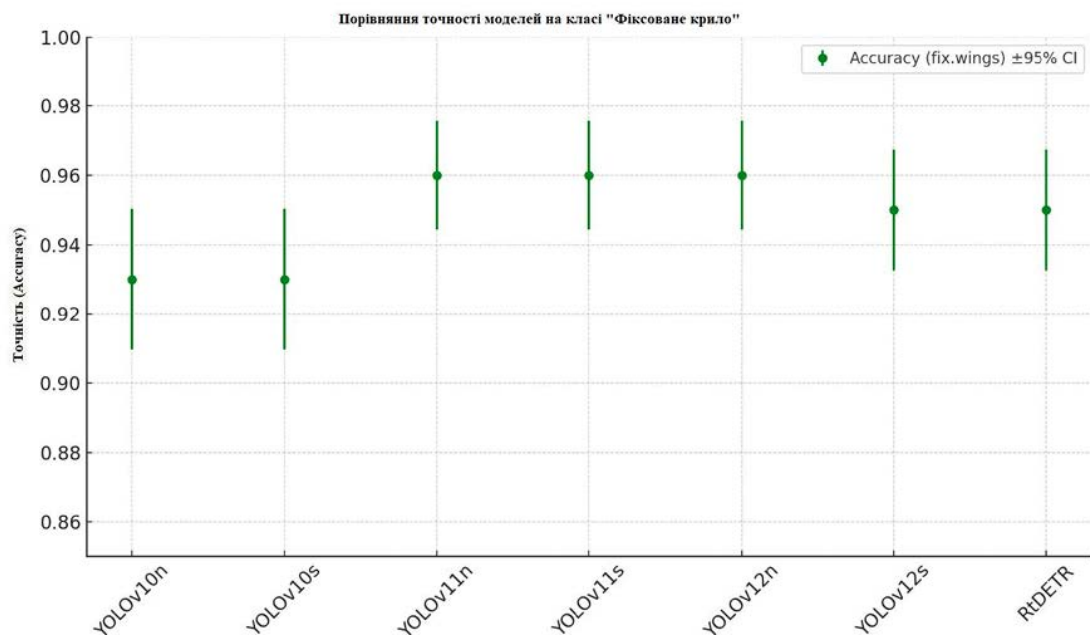


Рисунок 2.28 – Графік точності за класом «БПЛА з фіксованим крилом»

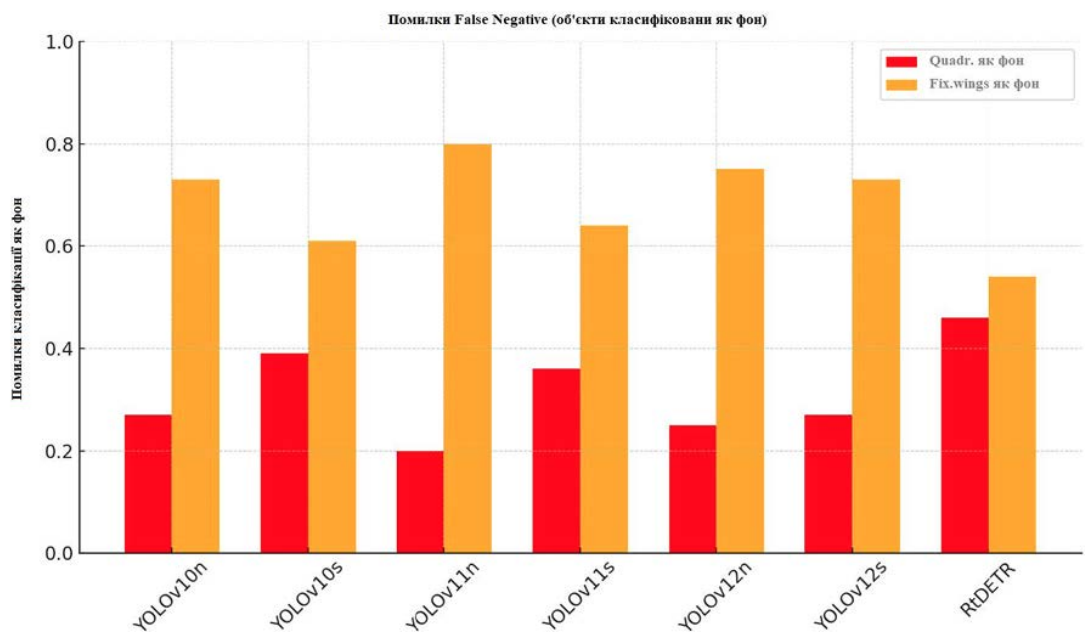


Рисунок 2.29 – Порівняльний графік помилок False Negative

RT-DETR має найбільшу помилку щодо квадрокоптерів як фон (46%), незважаючи на високу точність. У моделей YOLOv11n та YOLOv12n помилка фону у fix.wings становить 80% та 75%, що свідчить про часті пропуски цих об'єктів. Найменшу помилку за «fix.wings як фон» показує RT-DETR (54%).

2.11 Застосування навченої моделі RT-DETR на реальних даних

Для оцінки придатності моделі в реальних умовах було проведено інференс на кількох відеофайлах, які не входили до навчальної вибірки. Відеодані відображають реальні експлуатаційні характеристики (роздільна здатність відеокамери, світло, фон, спрямованість уздовж ЗПС) у прийнятно-здавальних випробуваннях БПЛА літакового типу на аеродромі.

Результати перевірки показують ефективність застосування детектора БПЛА в автоматизованому режимі. БПЛА добре розпізнаються моделлю, після чого система може автоматизовано логувати відеофайли в систему збору даних на різних етапах польоту із зазначенням належності до типу БПЛА.



Рисунок 2.30 – Розгін та зліт БПЛА (класифікація: fixed-wing 0.88)



Рисунок 2.31 – Захід на посадку БПЛА (класифікація: fixed-wing 0.82)

Однак існує невелика проблема в коректності визначення класу моделі, а саме: визначення БПЛА літакового типу як квадрокоптер на значній відстані від пункту спостереження.



Рисунок 2.32 – Помилкове визначення класу БПЛА літакового типу як клас квадрокоптер (quadrocopter 0.71)

Ця проблема викликана фізичними можливостями камери, а також відсутністю автоматичного зуму, оскільки цю роботу виконує оператор у ручному режимі. Тобто всі навчені моделі сприймали клас 1 (квадрокоптери) як об'єкти маленького масштабу (іноді навіть менше 20 пікселів), тоді як БПЛА літакового типу сприймалися моделлю як об'єкти, які займають значну частину зображення (мінімум 10–15% від усього розміру і більше).

Проте на стартовій позиції, при заході на посадку та в польоті БПЛА у видимій близькості від пункту спостереження правильна детекція БПЛА відбувається у понад 90% випадків.

Висновки до розділу:

У другому розділі, на основі результатів розгляду та аналізу науково-технічної літератури у першому розділі, було проведено порівняльне дослідження сучасних моделей детекції об'єктів для виявлення та класифікації основних типів БПЛА, зокрема моделей сімейства YOLO та RtDETR.. За результатами роботи сформовано результати навчання у вигляді отриманих метрик.

					КНУ КРБ.174.26.02.02.ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		55

Встановлено, що моделі YOLO демонструють високу точність і швидкодію, причому найкращі результати серед них показали YOLOv12s та YOLOv11s, особливо для класу квадрокоптерів. Водночас модель RtDETR забезпечила високі показники точності для БПЛА літакового типу, що підтверджує доцільність її використання у задачах, де пріоритетом є точність детекції. Отримані результати мають практичне значення, оскільки можуть бути використані для вдосконалення наземних пунктів управління БПЛА, систем моніторингу польотів, прийому телеметрії та аналізу польотної інформації, а також для подальшої адаптації легких моделей до мобільних пристроїв і компактних планшетів.

Таким чином, у межах цієї роботи вирішується головна задача - автоматичне виявлення безпілотних літальних апаратів (БПЛА) та визначення їхнього типу в потенційно небезпечних зонах: при зльоті, посадці, а також у межах видимості. Система здійснює детектування, класифікацію типу БПЛА та веде логування результатів у систему збору даних.

					КНУ КРБ.174.26.02.02.ПЗ	Арк.
						56
Змн.	Арк.	№ докум.	Підпис	Дата		

ВИСНОВКИ

У ході даного дослідження було проведено багатогранну оцінку ефективності різних моделей детекції об'єктів для задачі класифікації та виявлення основних класів БПЛА. Було досліджено різні версії моделей YOLO: YOLOv10n, YOLOv10s, YOLOv11n, YOLOv11s, YOLOv12n і YOLOv12s - у порівнянні з методами на основі Transformer, такими як RtDETR.

За результатами досліджень було отримано такі результати:

– Усі моделі YOLO продемонстрували високу точність детекції, однак відмінності в їхній ефективності були помітними залежно від структури моделі - сильно або слабо спрощеної версії;

– YOLOv12s показала найкращі результати серед моделей YOLO з найвищими значеннями Precision і Recall для обох класів, тоді як моделі YOLOv10n і YOLOv10s продемонстрували нижчі результати щодо точності детекції та локалізації;

– Модель RtDETR, незважаючи на використання складнішої архітектури, показала порівняно хороші результати, але не перевершила за точністю усі версії YOLO, особливо щодо швидкості роботи та детекції на зображеннях з об'єктами меншого розміру;

– Застосування методів статистичної оцінки показало, що різниця в результатах моделей YOLO і RtDETR за метрикою Precision є статистично значущою. Це підтверджується відмінностями в довірчих інтервалах між моделями, що свідчить про суттєву перевагу моделі RtDETR над моделями YOLO у задачі детекції БПЛА літакового типу.

– Зокрема, моделі YOLOv12s та YOLOv11s продемонстрували мінімальні значення False Negative (FN) та False Positive (FP) для класу квадрокоптерів, що вказує на їхню високу здатність до точної локалізації та

					<i>КНУ КРБ.174.26.02.00.ПЗ</i>		
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>	<i>ВИСНОВКИі</i>		
<i>Розроб.</i>		<i>Концевич Д.В.</i>					
<i>Перевір.</i>		<i>Трогь В. В.</i>					
<i>Н. Контр.</i>		<i>Маринич І.А.</i>					
<i>Затверд.</i>		<i>Маринич І.А.</i>			<i>АКНТ АКІТР-23ск</i>		
					<i>Літ.</i>	<i>Арк.</i>	<i>Аркушів</i>
						57	2

класифікації.

Висновки та рекомендації:

– Беручи до уваги, що більша частина продукції безпілотних літальних апаратів серійного виробництва складається з БПЛА з фіксованим крилом, для задач детекції, які потребують високої точності та швидкодії, рекомендується використовувати модель RtDETR, оскільки вона забезпечує найкращі показники точності.

– Подальші дослідження можуть бути спрямовані на виявлення аномалій зовнішнього корпусу апаратів під час зльоту та посадки, коли нерідко трапляються випадки пошкодження фюзеляжу під високими навантаженнями, або на виявлення аномальної поведінки самого БПЛА під час польоту.

Практичне значення:

– Результати цього дослідження можуть бути інтегровані у наявні наземні пункти управління БПЛА, де здійснюється керування, моніторинг польотів, прийом телеметрії, логування інформації та аналіз польотів.

– Більш легкі моделі, такі як YOLOv10n і YOLOv11n, можуть бути адаптовані для мобільних додатків та компактних планшетів.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks
URL: <https://arxiv.org/abs/1905.11946> (дата звернення 15.04.2026)
2. CrossViT: Cross-Attention Multi-Scale Vision Transformer for Image Classification URL: https://www.researchgate.net/publication/350484319_CrossViT_Cross-Attention_Multi-Scale_Vision_Transformer_for_Image_Classification (дата звернення 15.04.2026)
3. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows
URL: https://www.researchgate.net/publication/350398051_Swin_Transformer_Hierarchical_Vision_Transformer_using_Shifted_Windows (дата звернення 15.04.2026)
4. Deep Learning for Image Classification: A Comprehensive Review URL: https://www.researchgate.net/publication/317496930_Deep_Convolutional_Neural_Networks_for_Image_Classification_A_Comprehensive_Review (дата звернення 15.04.2026)
5. Data Augmentation for Improving Image Classification: A Survey URL: https://www.researchgate.net/publication/325920702_Data_augmentation_for_improving_deep_learning_in_image_classification_problem (дата звернення 15.04.2026)
6. ResNeSt: Split-Attention Networks URL: https://www.researchgate.net/publication/340805846_ResNeSt_Split-Attention_Networks (дата звернення 15.04.2026)
7. CNN Based Image Classification of Malicious UAVs
URL: https://www.researchgate.net/publication/366608244_CNN_Based_Image_Classification_of_Malicious_UAVs (дата звернення 15.04.2026)

					КНУ КРБ.174.26.00.00.ПЗ				
Змн.	Арк.	№ докум.	Підпис	Дата					
Розроб.		Концевич Д.В.			СПИСОК ЛІТЕРАТУРИ	Літ.		Арк.	Аркушів
Перевір.		Трогь В. В.						59	4
						АКНТ АКІТР-23ск			
Н. Контр.		Маринич І.А.							
Затвердив		Маринич І.А.							

8. MLP-Mixer: An all-MLP Architecture for Vision URL: https://www.researchgate.net/publication/351342431_MLP-Mixer_An_all-MLP_Architecture_for_Vision (дата звернення 15.04.2026)

9. A Comprehensive Survey of Unmanned Aerial Vehicles Detection and Classification Using Machine Learning Approach: Challenges, Solutions, and Future Directions URL: https://www.researchgate.net/publication/378662950_A_Comprehensive_Survey_of_Unmanned_Aerial_Vehicles_Detection_and_Classification_Using_Machine_Learning_Approach_Challenges_Solutions_and_Future_Directions (дата звернення 15.04.2026)

10. Bag of Tricks for Image Classification with Convolutional Neural Networks URL: https://www.researchgate.net/publication/338506275_Bag_of_Tricks_for_Image_Classification_with_Convolutional_Neural_Networks (дата звернення 15.04.2026)

11. Knowledge Distillation for Image Classification: A Survey URL: https://www.researchgate.net/publication/382551168_Knowledge_Distillation_in_Image_Classification_The_Impact_of_Datasets (дата звернення 15.04.2026)

12. A Review on Deep Learning in UAV Remote Sensing URL: https://www.researchgate.net/publication/353482369_A_Review_on_Deep_Learning_in_UAV_Remote_Sensing (дата звернення 25.05.2025)

13. Emerging Properties in Self-Supervised Vision Transformers URL: https://www.researchgate.net/publication/351221840_Emerging_Properties_in_Self-Supervised_Vision_Transformers (дата звернення 15.04.2026)

14. Hybrid deep learning with pelican optimization algorithm for M2M communication on UAV image classification URL: <https://ijeecs.iaescore.com/index.php/IJECS/article/view/36663> (дата звернення 25.05.2025)

15. Enhanced Intrusion Detection System for Multiclass Classification in UAV Networks URL: https://www.researchgate.net/publication/381484926_Enhanced_Intrusion_Detection_System_for_Multiclass_Classification_in_UAV_Networks (дата звернення 15.04.2026)

16. UAV image crop classification based on deep learning with spatial and

					<i>КНУ КРБ.174.26.02.00.ПЗ</i>	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		60

spectral features URL: https://www.researchgate.net/publication/352277623_UAV_image_crop_classification_based_on_deep_learning_with_spatial_and_spectral_features (дата звернення 15.04.2026)

17. Tree Species Classification from UAV Canopy Images with Deep Learning Models URL: https://www.researchgate.net/publication/384960921_Tree_Species_Classification_from_UAV_Canopy_Images_with_Deep_Learning_Models(дата звернення 15.04.2026)

18. ATS-YOLOv7: A Real-Time Multi-Scale Object Detection Method for UAV Aerial Images Based on Improved YOLOv7 URL: https://www.researchgate.net/publication/376221750_ATS-YOLOv7_A_Real-Time_Multi-Scale_Object_Detection_Method_for_UAV_Aerial_Images_Based_on_Improved_YOLOv7 (дата звернення 15.04.2026)

19. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors URL: https://www.researchgate.net/publication/361807900_YOLOv7_Trainable_bag-of-freebies_sets_new_state-of-the-art_for_real-time_object_detectors (дата звернення 15.04.2026)

20. Dynamic Head: Unifying Object Detection Heads with Attentions URL: https://www.researchgate.net/publication/352424393_Dynamic_Head_Unifying_Object_Detection_Heads_with_Attentions (дата звернення 15.04.2026)

21. QueryDet: Cascaded Sparse Query for Accelerating High-Resolution Small Object Detection URL https://www.researchgate.net/publication/350104746_QueryDet_Cascaded_Sparse_Query_for_Accelerating_High-Resolution_Small_Object_Detection (дата звернення 15.04.2026)

22. YOLO-LE: A lightweight and efficient UAV aerial image target detection model URL: https://www.researchgate.net/publication/384012383_YOLO-LE_A_lightweight_and_efficient_UAV_aerial_image_target_detection_model (дата звернення 15.04.2026)

23. Deep Learning for UAV-based Object Detection and Tracking: A Survey URL: https://www.researchgate.net/publication/355582689_Deep_Learning_for_UAV- (дата звернення 15.04.2026)

					<i>КНУ КРБ.174.26.02.00.ПЗ</i>	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		61

24. YOLO-Based UAV Technology: A Review of the Research and Its Applications URL: https://www.researchgate.net/publication/369216598_YOLO-Based_UAV_Technology_A_Review_of_the_Research_and_Its_Applications
(дата звернення 15.04.2026)

25. Real-Time Drone Detection Using Deep Learning Approach URL: https://www.researchgate.net/publication/328241210_Real-Time_Drone_Detection_Using_Deep_Learning_Approach_Third_International_Conference_MLI_COM_2018_Hangzhou_China_July_6-8_2018_Proceedings(дата звернення 15.04.2026)

26. Simplified Object Detection for Manufacturing: Introducing a Low-Resolution Dataset URL: https://www.researchgate.net/publication/388277621_Simplified_Object_Detection_for_Manufacturing_Introducing_a_Low-Resolution_Dataset(дата звернення 15.04.2026)

27. Моркун Н. В., Маринич І. А. Методичні вказівки до виконання кваліфікаційної роботи бакалавру для студентів спеціальності 151 “Автоматизація та комп’ютерно-інтегровані технології”. Кривий Ріг : Видавничий центр КНУ, 2019. 50 с.

28. ДСТУ 3008:2015. Інформація та документація. Звіти у сфері науки і техніки. Структура і правила оформлення. Київ, ДП «УкрННЦ», 2015. 26с. (Інформація та документація).

29. ДСТУ 8302:2015. Бібліографічне посилання. Загальні вимоги та правила складання Київ, ДП «УкрННЦ», 2016. 16 с. (Інформація та документація).

30. ДСТУ 3582:2013. Бібліографічний опис. Скорочення слів і словосполучень в українській мові. Загальні вимоги та правила. Київ, ДП «УкрННЦ», 2013. 23 с. (Інформація та документація)

31. ДСТУ 3651.0-97 Метрологія. Одиниці фізичних величин. Основні одиниці фізичних величин Міжнародної системи одиниць. Основні положення, назви та позначення Київ, Держстандарт України, 1998. 27 с. (Інформація та документація)..

					КНУ КРБ.174.26.02.00.ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		62

