

Міністерство освіти і науки України
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛЯНСЬКА АКАДЕМІЯ»
Кафедра інформатики факультету інформатики



**Знаходження безпілотних повітряних об'єктів методами глибинного
навчання**

**Текстова частина до дипломної роботи
за спеціальністю „Комп'ютерні науки” - 122**

Керівник курсової роботи
доцент Ющенко Ю.О.

“ ____ ” _____ 2023 р.
(підпис)

Виконав студент
Федюченко М.І.
“ ____ ” _____ 2023 р.

ЗМІСТ

ВСТУП	7
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ	10
1.1 АРХІТЕКТУРА UAV	10
1.2 МЕТОДИ ВИЯВЛЕННЯ ТА СУПРОВОДУ ДРОНІВ	14
1.2.1 ВИЯВЛЕННЯ ЗА ДОПОМОГОЮ РАДАРІВ	14
1.2.2 ВИЯВЛЕННЯ ЗА ТЕРМАЛЬНИХ КАМЕР	16
1.2.3 ВИЯВЛЕННЯ ЗА ДОПОМОГОЮ РАДІОЧАСТОТ	17
1.2.4 АУДІО-ВИЯВЛЕННЯ	18
1.2.5 ЕЛЕКТРООПТИЧНЕ ВИЯВЛЕННЯ	20
1.2.5.1 ОПТИЧНЕ ВИЯВЛЕННЯ	22
1.3 ГЛИБИННЕ НАВЧАННЯ ДЛЯ ВИЗНАЧЕННЯ ОБ'ЄКТІВ	22
1.3.1 ДВОХЕТАПНЕ ВИЗНАЧЕННЯ ОБ'ЄКТУ	23
1.3.2 ОДНОЕТАПНЕ ВИЗНАЧЕННЯ ОБ'ЄКТУ	28
1.4 YOLO	29
1.4.1 YOLOV5	32
1.5 ПРОБЛЕМА ВИЗНАЧЕННЯ МАЛИХ ОБ'ЄКТІВ	36
1.5.1 ПЛИТКОВЕ РОЗБИТТЯ	36
2 МЕТОДОЛОГІЯ	38
2.1 ЗБІР НАЧАЛЬНИХ ДАНИХ	38
2.1.1 ДЖЕРЕЛА НАБОРІВ ДАНИХ	39
2.2 ЗАСТОСУВАННЯ ПЛИТКОВОГО РОЗБИТТЯ	41
2.3 ОБРАНІ МЕТРИКИ	43
2.3.1 ТОЧНІСТЬ ТА ПОВНОТА	43
2.3.2 РЕЗУЛЬТАТИВНИЙ ЧАС	44
2.3.3 ТЕСТ ФРІДМЕНА ТА ВІЛКОКСОНА ДЛЯ СТАТИСТИЧНОГО АНАЛІЗУ	44
2.4 ПРОВЕДЕНІ ЕКСПЕРИМЕНТИ	46
2.4.1 ОЦІНКА YOLO МОДЕЛЕЙ	46

	3
2.4.2 ЕВОЛЮЦІЯ КОНФІГУРАЦІЙНИХ ПАРАМЕТРІВ	47
2.4.3 ЗАСТОСУВАННЯ ПЛИТУВАННЯ НА НАБОІ ДАНИХ З СИМУЛЯЦІЇ	48
2.5 СИСТЕМНА ДОКУМЕНТАЦІЯ	48
2.5.1 КОНФІГУРАЦІЙНІ ЗНАЧЕННЯ.....	48
2.5.2 СЕРЕДОВИЩЕ РОЗРОБКИ	50
3 АНАЛІЗ РЕЗУЛЬТАТІВ.....	51
3.1 ВИБІР НАВЧАЛЬНОЇ МОДЕЛІ	51
3.1.1 НАВЧАННЯ МОДЕЛЕЙ YOLOV5	51
3.1.2 НАВЧАННЯ МОДЕЛЕЙ YOLOV8	53
3.1.3 ПОРІВНЯННЯ РЕЗУЛЬТАТІВ НАВЧЕНИХ МОДЕЛЕЙ	54
3.2 ВПЛИВ НА РЕЗУЛЬТАТИ ЕВОЛЮЦІЄЮ ПАРАМЕТРІВ	55
3.3 ВПЛИВ НА РЕЗУЛЬТАТИ ЕВОЛЮЦІЄЮ ПАРАМЕТРІВ	56
ВИСНОВКИ	59
СПИСОК ЛІТЕРАТУРИ	61
ДОДАТОК А (ОБОВ'ЯЗКОВИЙ) НАЙТИПОВІШІ ТИПИ ДРОНІВ	64

Міністерство освіти і науки України
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛЯНСЬКА АКАДЕМІЯ»
Кафедра інформатики факультету інформатики

ЗАТВЕРДЖУЮ
Зав.кафедри інформатики,
доцент., к. ф.-м. н. Гороховський С.С.

(підпис)
« ____ » _____ 2022 р.

ІНДИВІДУАЛЬНЕ ЗАВДАННЯ
на дипломну роботу

студенту Федюченку Михайлу Ігоровичу факультету інформатики 2-го курсу
ТЕМА Знаходження безпілотних повітряних об'єктів методами глибинного
навчання

Зміст ТЧ до курсової роботи:

1. Індивідуальне завдання
2. Календарний план
3. Анотація
4. Вступ
5. Аналіз предметної області
6. Огляд методології
7. Аналіз отриманих результатів
8. Висновки
9. Список використаної джерел

Дата видачі „ ____ ” _____ 2022 р. Керівник _____
(підпис)

Завдання отримав _____
(підпис)

Календарний план виконання роботи

№ п/п	Назва етапу курсової роботи	Термін виконання етапу	Примітка
1.	Отримання теми курсової роботи	29.09.2022	
2.	Огляд літератури за темою роботи	18.10.2022	
3.	Виконання аналізу предметної області	21.11.2022	
4.	Огляд літератури одно та двоетапних детекторів	14.01.2023	
5.	Написання першого розділу	23.01.2023	
6.	Пошук наборів даних для навчання моделей	20.02.2023	
7.	Написання другого розділу	30.03.2023	
8.	Виконання експериментів	04.04.2023	
9.	Написання останнього розділу	01.05.2023	
10.	Корегування роботи згідно із зауваженнями керівника	03.06.2023	

Студент Федюченко М.І.

Керівник Ющенко Ю.О.

“ _____ ”

Анотація

Дипломна робота присвячена виявленню безпілотних повітряних апаратів (дронів) з використанням методів глибокого навчання. В процесі дослідження було проаналізовано вже існуючі методи та системи виявлення безпілотників, та на основі цих даних обрано потрібну методологію. Було виконано збір релевантних наборів даних та проведено тренування моделей YOLOv5 та YOLOv8 на відповідних наборах. Було проаналізовано ефективність використання еволюційних алгоритмів для конфігураційних параметрів моделей. Після тренування були зібрані тестові дані, що включали зображення з реальними сценаріями, включаючи різні умови освітлення, шум та зміну масштабу об'єктів. У результаті порівняння встановлено, що використання техніки плиточного розбиття покращує детекцію малих об'єктів, забезпечуючи більш точні результати, але тільки у випадку високої роздільної якості зображень. Виявлено, що обидві архітектури YOLOv5 і YOLOv8 демонструють високу ефективність у виявленні дронів, проте YOLOv8 продемонстрував дещо кращі результати у порівнянні з YOLOv5. Актуальність роботи полягає в тому, що отримані результати можна застосувати для виявлення безпілотних повітряних апаратів, що є корисними для подальшого розвитку систем безпілотного контролю та безпеки повітряного простору.

Вступ

Важливим елементом розвитку людства є створення рішень, що покращують та полегшують людське існування. Не зважаючи на вирішення поставлених проблем та створення нових можливостей це в свою чергу може нести нові загрози та виклики з котрим людство ще ніколи не стикалось. Наприклад, зростаюче використання безпілотних літальних апаратів у різних сферах життя відкриває нові можливості, але також породжує проблеми безпеки та приватності.

Зі зниженням ціни на комплектуючі, та обчислювальні потужності використання дронів виходить за рамки простого хобі або розваги і стає реальним викликом для безпеки інфраструктури та громадського порядку. Дрони можуть використовуватися для здійснення незаконних наглядів, незаконного збирання інформації, транспортування незаконних речей або навіть для атак на об'єкти та особи. Яскравим прикладом може бути використання різноманітних безпілотних літальних апаратів під час Російсько-Української війни. Зважаючи на дешевизну самих апаратів, вони активно використовуються у якості розвідників, для планування спеціальних операцій, виявлення живої сили та техніки противника, корегування вогню, а також враження цілей у випадку дронів, що переносять вибуховий заряд.

Особливу увагу слід приділяти використанню дронів у заборонених або недоступних зонах, таких як військові об'єкти, повітряні простори аеропортів або приватна територія. Це може створювати серйозні загрози для національної безпеки, інфраструктури та особистої безпеки громадян.

Одним із викликів полягає у розробці ефективних систем виявлення та відстеження дронів, що порушують заборонений периметр, оскільки традиційні технології, що використовуються в системах відстеження повітряних цілей, такі як радар або радіочастотний детектор, не є достатньо точними у випадках, коли дрон перебуває в місці з перешкодами сигналу або коли отриманий сигнал блокується.

Крім того, радіолокаційне обладнання, яке використовується для виявлення дронів, зазвичай коштує доволі багато. Синхронізація цього обладнання між собою можуть бути складними завданнями, особливо при покритті великих територій. Це обмежує масштаби і ефективність систем, оснащених традиційними засобами виявлення та супроводження повітряних цілей. У зв'язку з цим, розробка нових технологічних рішень, які дозволяють ефективно виявляти, відстежувати та супроводжувати дрони, стає важливою. Підхід, заснований на використанні штучного інтелекту та глибокого навчання, може виявитися більш доступним та економічно ефективним. Використання таких технологій дозволяє досягти високої точності без необхідності в дорогому та складному обладнанні. Крім того, системи, засновані на штучному інтелекті, можуть бути легко масштабовані для покриття великих територій, що робить їх більш ефективними в контролі та відстеженні дронів. Технології штучного інтелекту, зокрема глибоке навчання, може бути використане для автоматичного виявлення та класифікації дронів з високою точністю.

Предметом дослідження дипломною роботи є розробка та вивчення ефективних методів та моделей для автоматичного виявлення дронів з використанням технологій глибокого навчання. Дослідження спрямоване на розробку алгоритмів та систем, які здатні точно та швидко виявляти присутність дронів на зображеннях або відео з різних джерел. У межах дослідження буде проведений аналіз сучасних моделей глибокого навчання, які використовуються для виявлення об'єктів, і визначені найбільш ефективні моделі для виявлення дронів. Також будуть вивчені різні підходи до навчання моделей з використанням різних архітектур нейронних мереж, методів попереднього навчання та оптимізації параметрів.

Основна мета дослідження полягає в розробці та валідації моделей для виявлення дронів, які будуть відповідати високим стандартам точності, швидкодії та стійкості до різних умов зйомки. А також створення системи, котра зможе з

високою ймовірністю ідентифікувати безпілотний літальний апарат в заданому секторі використовуючи недороге та просте у підтримці обладнання. Що є дуже актуальним в умовах Російсько-Української війни, коли є постійна небезпека авіаційних атак із застосуванням дронів. Було виконано збір навчальних даних шляхом завантаження загальнодоступних наборів даних з різними типами дронів, а також створення власних даних. Після огляду було вибрано кілька передових моделей, які слугують кандидатами для порівняльного аналізу. Подальше дослідження включає внесення модифікацій до моделі та набору даних з метою вдосконалення ефективності виявлення. Для тестування різних стратегій навчання виконувалося тренування моделей з різних початкових умов та наборів даних.

Текстова частина курсової роботи складається з трьох розділів

У першому розділі розглядаються методи виявлення та супроводження повітряних об'єктів, проаналізовано переваги та недоліки кожного методу.

У другому розділі розглянуто популярні моделі генеративного визначення повітряних об'єктів. У розділі наведено детальний огляд набору даних, використаних метрик для оцінки результатів, а також опис структури моделі та обладнання, яке було використано під час проведення експериментів. Також розглянуто основні аспекти середовища, в якому здійснювалися експерименти.

Третій розділ висвітлює результати проведених експериментів з прикладами проміжних результатів з урахуванням різних середовищ. Також виконується аналіз результатів дослідження для кожного методу та різних налаштувань.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

Для виявлення безпілотних літальних апаратів потрібно перш за все зрозуміти, що це саме таке, які є різновиди, як вони функціонують, що можливо є спільного у дронів. Важливо дізнатися які є переваги та недоліки функціонування таких пристроїв.

1.1 Архітектура UAV

Дрони, як правило, класифікуються на чотири категорії: багатороторні, крилофіксовані, планерні, однороторні-гелікоптерні та вертикально-злітно-посадкові системи (ВЗПС). У сфері безпілотної авіації, серед розглянутих типів дронів значну популярність набули мультироторні дрони, які характеризуються використанням множини роторів для генерації підйомної сили та керування польотом. Основними представниками мультироторних дронів є квадрокоптери, які мають чотири ротори, розташовані на протилежних кутах платформи. Цей тип дронів відрізняється своєю простою конструкцією та високою маневреністю.

Поміж інших типів мультироторних дронів можна виділити шестикоптери, які мають шість роторів, розташованих на шести кутах платформи. Шестикоптери забезпечують більшу стійкість і підйомну силу, що дозволяє їм перевозити більші вантажі та працювати у менш сприятливих умовах. Також належить відзначити октокоптери, які мають вісім роторів, розташованих на восьми кутах платформи. Вони забезпечують ще більшу стійкість, маневреність та підйомну силу, що дозволяє їм працювати в умовах високої обтічної швидкості.

Крилофіксовані дрони представляють собою конструкцію без прямокутного хвостового крила та фюзеляжа, що має велике крило, яке охоплює практично всю довжину апарата. Цей тип дрона характеризується відсутністю вертикального

хвостового стабілізатора, а також доповнюється системою контролю положення і керування, яка забезпечує стабільність та маневреність в повітрі. Найпопулярніші різновиди: Delta Wing (Трикутне крило). Такий тип дрона має трикутну форму крила з вузьким носом і широкою базою. Наведена у приклад конфігурація забезпечує велику підйомну силу та маневреність під час польоту. Delta Wing дрони часто використовуються для забезпечення швидкого та стійкого політного характеру. Існує багато різновидів, найпопулярніші з них: Flying Wing (Крило без фюзеляжа), що характеризується відсутністю окремого фюзеляжу. Велике крило безпосередньо об'єднує площину керування, системи автономного польоту та навігації. Крилофіксовані дрони типу Flying Wing забезпечують високу ефективність, довгий час польоту та велику площину для розміщення сенсорів і легкого вантажу. Wing Body має плавний перехід між крилом і фюзеляжем, створюючи змішану форму. Це дозволяє забезпечити кращу аеродинаміку, знизити опір повітря і покращити енергоефективність. Такі дрони мають велику вантажопідйомність і можуть бути використані для довготривалого нагляду або транспортування великих вантажів.

Серед літакоподібних дронів, варто визначити A-tail, V-tail та standart plane дрони, які відрізняються своєю конструкцією планера. A-tail дрони володіють конфігурацією хвостової частини, яка нагадує літеру "А". Це означає, що вони мають вертикальний стабілізатор та два горизонтальних поверхи, розташовані у вигляді літери "А". Така конфігурація забезпечує дрону стабільність польоту та керуваність у повітрі. V-tail дрони, натомість, мають хвостову частину у формі літери "V". Це означає, що вони поєднують вертикальний стабілізатор та горизонтальний поверх в одній конструкції. Така конфігурація сприяє зниженню опору повітря та покращенню аеродинамічних характеристик дрона. Standart plane дрони є типовими планерами з фіксованим крилом та горизонтальним та вертикальним стабілізаторами. Їх конструкція близька до традиційних літаків. Вони володіють доброю стійкістю польоту та здатністю до плавного керування. Цей тип

дронів часто використовується для виконання довгих місій, аерофотозйомки та наукових досліджень.

VTOLs дрони характеризуються особливою конструкцією та функціональністю. Вони можуть мати різні конфігурації та аеродинамічні рішення, включаючи квадрокоптерну, гібридну та інші. Квадрокоптерні VTOLs дрони мають чотири вертикально розташовані ротори, що забезпечують їм здатність вертикального зльоту та посадки, а також маневреність у повітрі. Гібридні VTOLs дрони поєднують фіксоване крило для польоту на великій швидкості та вертикальний ротор для вертикального зльоту та посадки.

Приклади описаних конструкцій можна побачити у додатку А.

Спільним для літальних апаратів є інтегровані підсистеми, що відповідають за запуск, політ, функціонування, обробку, передачу та отримання команд з віддалених або наземних станцій. Можна визначити основні елементи архітектури БПЛА:

- 1) блок живлення - розроблений з метою забезпечення тривалого часу роботи БПЛА без необхідності його заряджання;
- 2) модуль зв'язку – отримання команд та надсилання отриманої інформації через радіочастоти. БЛА оснащені високоякісним бездротовим комунікаційним обладнанням, включаючи 5G¹, WiF²i, Bluetooth³ та ідентифікацію за радіочастотою (RFID⁴), для забезпечення зв'язку з наземною станцією керування або інтернетом;
- 3) основний обчислювальний пристрій - обробляє команди на основі зібраних даних з інших підсистем або компонентів. Обробляє отриману інформацію,

¹ 5G є п'ятою генерацією мобільних телекомунікаційних мереж і представляє собою новий стандарт бездротового зв'язку.

² Wi-Fi є технологією бездротового мережевого з'єднання

³ Bluetooth є бездротовим комунікаційним протоколом, що дозволяє обмінювати дані між електронними пристроями на короткій відстані. Він широко використовується для підключення телефонів, планшетів, навушників, гучномовців, клавіатур, мишей, принтерів та багатьох інших пристроїв.

⁴ RFID (Radio Frequency Identification) є технологією безконтактного ідентифікаційного зчитування, яка використовує радіочастотний сигнал для обміну даними між RFID-системою та об'єктами, які мають RFID-теги або мітки.

передає її назад, керує польотними операціями та забезпечує зв'язок з наземним контролем;

- 1) датчикова плата – відповідає за функціонування GPS⁵, датчиків, гіроскопів, акселерометрів, антен, приймачів та інших різноманітних датчиків потрібних для функціонування дрону чи виконання поставленої задачі;
- 2) Наземна станція керування (НСК): Ця базова станція в основному використовується для моніторингу та керування БЛА під час його роботи. Польотні операції постійно контролюються і можуть бути керовані для зміни місії.

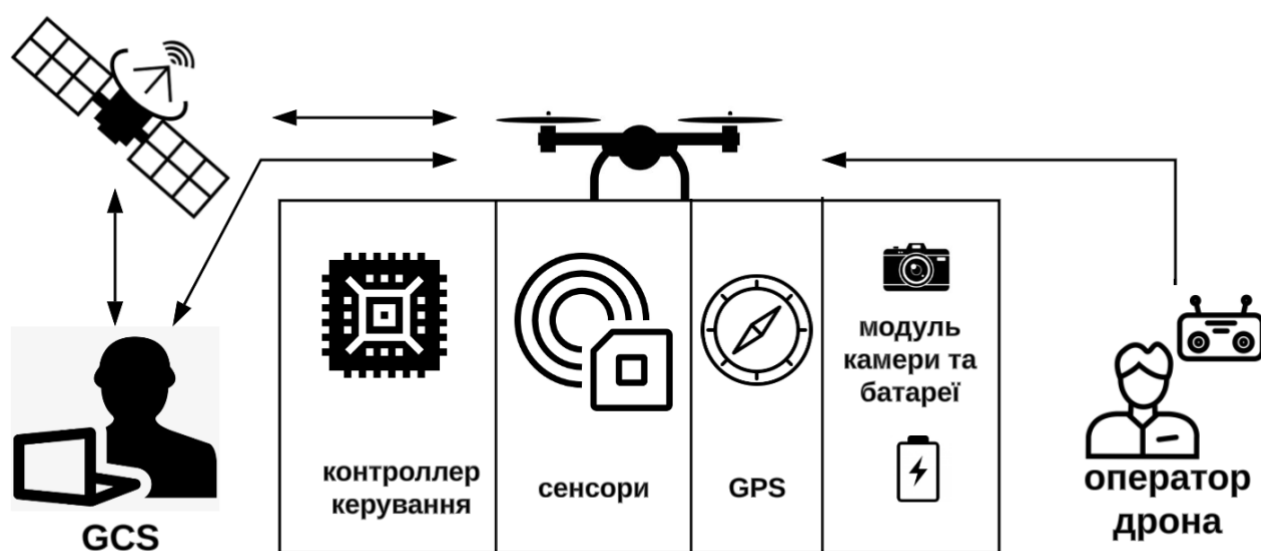


Рисунок 1.1 - високорівнева архітектура UAV

⁵ GPS (Global Positioning System) є глобальною навігаційною системою, яка дозволяє визначати точне місцезнаходження на Землі за допомогою супутникового сигналу. Система GPS складається з супутників, наземних станцій контролю і приймачів GPS.

1.2 Методи виявлення та супроводу дронів

Виявлення дронів є складним завданням, оскільки у повітрі присутні багато об'єктів, таких як птахи, хмари або комерційні літаки. Для виявлення дронів використовуються різноманітні методи:

- 1) Виявлення за допомогою радару: Системи радару використовують радіохвилі для виявлення та відстеження об'єктів у повітряному просторі;
- 2) Акустичне виявлення: дрони виробляють характерні звуки, особливо ті, що оснащені роторам;
- 3) Візуальне виявлення: Методи візуального виявлення використовують камери та алгоритми комп'ютерного зору для ідентифікації дронів за їх зовнішнім виглядом та характеристиками руху;
- 4) Виявлення за допомогою радіочастот (RF);
- 5) Виявлення за допомогою інфрачервоного (IR) випромінювання;
- 6) Комбінований підхід;

Варто зазначити, що кожен метод виявлення має свої переваги, обмеження та застосовність, залежно від таких факторів, як умови навколишнього середовища, характеристики дронів та конкретні вимоги до виявлення.

1.2.1 Виявлення за допомогою радарів

Обробка радарних радіосигналів є одним з класичних підходів до виявлення повітряних суден та безпілотників, оскільки він може використовуватись у всіх погодних умовах з цілодобовою роботою на відміну від методів акустичного та візуального виявлення. Цей підхід отримання сигналу характеризується для виявлення ехо, доплерівської сигнатури або радіолокаційної площі перехрещення (RCS) для виявлення та відстеження цілі. Класичні методи обробки радіосигналів

мають обмеження у точному відрізненні малих безпілотників від птахів через їх меншу RCS. Існують методи на основі штучного інтелекту, що виконують пост-обробку ознаки з радіосигналів для часткового вирішення цієї проблеми.

У радарному виявленні використовується радіоенергія для виявлення цілі та визначення її положення. Зазвичай радарна система виявлення має три основні компоненти: радар РЧ⁶, збір даних та обробка сигналу. У радарі РЧ електромагнітна енергія випромінюється в простір та зустрічає тіло БПЛА, що перебуває у зоні спостереження. Відбита хвиля БПЛА повертається та отримується системою, вимірюється та обробляється в реальному часі (збір даних та обробка сигналу). Таким чином, успішно визначається положення БПЛА, а його шлях руху відстежується системою.

Метод відстеження	Специфікація	Ідентифікація	Класифікація	Локалізація та супроводження	Продуктивність
FMCW/CW радар	Принцип ефекту доплера	+	-	-	Залежно від потужності обладнання
CW радар	Частота опрацювання: 35 ГГц	+	-	+	Точність 85%
5G міліметровий радар	Початкова частота становить 25 ГГц, що	+	-	+	Виявлення на відстані 300 м зі швидкістю 157,9 об/с і

⁶ Радар РЧ (радіолокаційний радіовимірювальний апарат радіочастотного діапазону) – це пристрій який використовує радіохвилі в радіочастотному діапазоні для вимірювання відстаней, напрямку, швидкості та інших характеристик об'єктів у просторі.

	міститься в діапазоні 5G				на 850,2 м зі швидкістю 88 об/с
--	-----------------------------	--	--	--	---------------------------------------

Таблиця 1- Підсумок найпопулярніших методів відстеження дронів з використанням радарів

1.2.2 Виявлення за термальних камер

Термальне виявлення є ще одним методом, який використовується для виявлення безпілотних літальних апаратів, особливо для дронів з фіксованим крилом з рушійними двигунами, що генерують гарячі вихлопні гази. За допомогою виявлення теплового підпису термальні сенсори можуть потенційно виявляти дрони на відстані до 110 метрів. Проте цей метод має свої обмеження.

Пластикові квадрокоптери з електричними двигунами, випромінюють менше тепла і їх складніше виявити за допомогою термальних сенсорів. Це призводить до помилково позитивних результатів, коли птахи або інші теплові об'єкти помилково ідентифікуються як дрони через їх вищу теплову підпису в порівнянні з пластиковими дронами. Тому термальне виявлення само по собі може бути ненадійним і частіше використовується як додатковий компонент у поєднанні з іншими методами виявлення.

Важливо зазначити, що вартість впровадження термального виявлення може бути високою, оскільки потрібні спеціалізовані термальні сенсори та обладнання. Крім того, швидкість виявлення за допомогою термальних сенсорів самотійно зазвичай невисока, а дальність обмежена відносно невеликими відстанями. Тому комплексна система виявлення дронів зазвичай поєднує кілька технік, включаючи

термальне виявлення, радар, акустичні сенсори та візуальну ідентифікацію, для досягнення вищої точності та надійності.

1.2.3 Виявлення за допомогою радіочастот

Методи виявлення на основі радіочастот включають моніторинг радіочастотного спектру для ідентифікації сигналів, що випромінюються дронами. Ці сигнали можуть включати комунікаційні сигнали між дроном та його контролером, а також телеметричні дані, що надсилаються дроном. Різні дослідження повідомляють про перспективні результати, використовуючи різні алгоритми та техніки для виявлення БПЛА на основі радіосигналів. Проте, наявність шуму впливає на точність та дальність виявлення. У Таблиці 2 наведено узагальнення оглядових робіт та перелік особливостей та точності методології виявлення БПЛА на основі радіосигналів.

У переглянутих джерелах зазначено, що метод пасивного виявлення за допомогою радіосигналів має переваги низької вартості, відсутності необхідності в ліцензіях, великого діапазону дальності та можливості раннього попередження.

Метод відстеження	Ідентифікація	Класифікація	Локалізація та супроводження	Продуктивність
2.4 ГГц	+	+	-	Середня оцінка 97%
20 МГц	+		+	Не визначено
1-6 ГГц	+	+	+	Середня оцінка 99%

ISM Bands	+	-	+	Не визначено
-----------	---	---	---	--------------

Таблиця 2- Підсумок найпопулярніших методів відстеження дронів шляхом використання радіочастот

1.2.4 Аудіо-виявлення

Акустичні сенсори, такі як мікрофонні масиви, захоплюють звук, що створюється роторами та пропелерами дрона, і потім порівнюють отримані характеристики з базами акустичних сигнатур для виявлення та класифікації дронів та безпілотних літальних апаратів (БПЛА) за допомогою традиційних та штучно-інтелектуальних архітектур. Для цього використовуються методи аналізу звуку, такі як мел-частотні кепстральні коефіцієнти і короткочасне перетворення Фур'є. Більшість мікрофонів можуть сприймати звук з відстані 5-10 метрів. Звукові хвилі потім обробляються і фільтруються, й аналізується цільова частота. Більшість дронів оснащені гучними моторами, які генерують шиплячий звук високої частоти близько 40 кГц. Цей звук є характерним для більшості дронів. За допомогою цифрової обробки сигналів можна ідентифікувати конкретну частоту і визначити наявність дрона неподалік. MFCC⁷ представляють собою набір характеристик, які відображають сприйняття людиною звуків і застосовуються для аудіо класифікації з використанням методів машинного навчання. STFT⁸ є проміжною

⁷ MFCC (Mel Frequency Cepstral Coefficients) є методом аналізу та представлення звукових сигналів в обробці мови та розпізнаванні мови. Він базується на принципах спектрального аналізу звуку та моделювання людського слуху

⁸ STFT (Short-Time Fourier Transform) є методом аналізу часово-частотного представлення сигналу. Він використовується для виявлення змін спектральної складової сигналу в залежності від часу.

характеристикою, яка містить більше інформації та шуму, ніж MFCC, що дає йому перевагу. Глибинні моделі навчання здатні легко обробляти STFT і працювати з більш складними даними. Застосування акустичних сенсорів у поєднанні з цими алгоритмами дозволяє ефективно виявляти та класифікувати дрони та БПЛА на основі їх акустичних сигнатур. На відміну від радарних та радіочастотних підходів, акустичне рішення не вимагає прямої видимості. Однак у цього рішення є виклики, такі як обмежений зони прийому, потреба у великій базі акустичних сигнатур та вразливість до шумів та завад, особливо в міських районах, а також при тихій роботі безпілотної. Детекція пілота дрона також може бути неможливою за допомогою акустичних сенсорів.

Метод відстеження	Ідентифікація	Класифікація	Локалізація та супроводження	Продуктивність
BiLSTM	+	+	-	Середня оцінка 94.02%
Concurrent neural networks	+	-	+	Середня оцінка 96.3%
TDoa, SRP-PHAT	+	-	+	SRP-PHAT перевершує TDoA
MFCC, STFT,CNN,SVM	+	-	+	Не визначено
SRP-PHAT	+	-	+	Шум дуже впливає на виявлення

Таблиця 3- Підсумок найпопулярніших методів відстеження дронів шляхом аудіопрослуховування

1.2.5 Електрооптичне виявлення

Система електрооптичного виявлення передає, виявляє та аналізує випромінювання в оптичному спектрі, включаючи видиме світло, інфрачервоне та ультрафіолетове випромінювання. Вона може працювати з великими дальностями зображення та має надійні результати при різних рівнях освітлення. Компоненти, пов'язані з такими системами, включають оптику, лазер, детектори, камеру, обчислювальний блок та інше. Такі системи застосовуються для виявлення, визначення напрямку та локалізації БПЛА незалежно від погодних умов.

У деяких електрооптичних рішеннях виявлені дані, передані на етап аналізу, включаючи розширену обробку, машинне зір або машинне навчання, недостатньо точні для ефективного відстеження малих БПЛА.

Метод електро-оптичного виявлення потребує врахування наступних факторів: розмір і рух у тривимірному просторі, швидкість виявлених повітряних об'єктів, максимальна відстань виявлених об'єктів від положення камери, опис оптичної лінзи та роздільна здатність лінійних об'єктів на зображенні. Оскільки ці фактори безпосередньо пов'язані з методами обробки зображень, виявлення відстані та результати виявлення можуть бути негативно піддані впливу, якщо один з вищезазначених факторів містить неправильну або неточну інформацію. Виявлення рухомих об'єктів на максимальній відстані від камери в реальному часі є основною метою цього методу. Виявлення за допомогою електрично-оптичного походу є важливим інструментом, але воно має певний набір обмежень. Камери можуть зафіксувати дрон на відстані приблизно 110 метрів з задовільною роздільною здатністю, що робить такий підхід економічно вигідним варіантом. Проте можуть виникати проблеми з класифікацією літальних об'єктів і птахів.

Метод відстеження	Ідентифікація	Класифікація	Локалізація та супроводження	Продуктивність
Датчик 3D LADAR, 3D віднімання фону, V-RBNN	+	-	+	Дальність виявлення до двох кілометрів
Комбінація EO/IR з акустичними сигналами	+	-	+	Обмеження прямої видимості
DVS камера, скронева фільтрація, триангуляція	+	-	+	Точна дальність виявлення 30 м
Віднімання фону, CNN	+	-	-	Постійно змінні фонові залежності

Таблиця 3- Підсумок найпопулярніших методів відстеження дронів шляхом аудіопрослуховування

Ефективність рішень, основою яких є зорове спостереження, погіршується у випадку відсутності прямої видимості, правильного куту зйомки, а також якщо низька якість об'єктива, у туманних, темних та запиленних середовищах та при високій температурі фону. Вищезазначені обмеження можна, до певної міри, усунути за допомогою ІЧ-камери, тобто виявлення на основі теплового випромінювання елементів дрона. Однак, це значно збільшує вартість системи та обмежує дальність виявлення та ефективність в різних середовищах через чутливість датчиків, що вимірюють теплову різницю між дроном та фоном.

1.2.5.1 Оптичне виявлення

Залежно від методу отримання ознак для класифікації та визначення об'єктів, можна поділити електрооптичний метод на два підходи – традиційний ручний підхід та підхід з використання глибинного навчання.

На основі ручного вилучення ознак та алгоритмів машинного навчання зазвичай працюють традиційні підходи виявлення об'єктів. Ці підходи передбачають початкове знаходження потрібної області за допомогою таких методів, як віднімання фону. Адже у цих програмах відео потоки зазвичай складаються з квазістатичних камер, де камери або рухаються дуже повільно, або залишаються нерухомими, тоді як цілі у відео потоках є рухомими. Дескриптори ознак, такі як гістограма орієнтованих градієнтів (HOG), дескриптори Фур'є і прискорені надійні функції (SURF), витягують у свою чергу з ідентифікованих областей інтересу ознаки та пересилають до класифікаторів для розпізнавання та класифікації. Хоча ці методи за більшістю досліджень швидші, вони, як правило, мають нижчий рівень розпізнавання порівняно з підходами глибинного навчання.

Підхід з використанням глибинного навчання зазвичай отримує ознаки за допомогою згорткових шарів, що дають можливість отримати як високорівневі семантичні ознаки із зображень, так і ознаки із необроблених пікселів. У результаті глибинне навчання може генерувати кращі ієрархічні ознаки.

1.3 Глибинне навчання для визначення об'єктів

У площині комп'ютереного зору задачу знаходження об'єкта можна розділити та три підзадачі: класифікацію, виявлення та сегментацію. Тобто важливо визначити наявність і місцезнаходження дрона на кадрах, побудувавши навколо нього обмежувальну рамку, що відповідає саме задачі виявлення. Виявлення

об'єктів у глибинному навчанні зазвичай здійснюється за допомогою двох типів моделей: двохетапних детекторів, або ж одноетапних.

1.3.1 Двохетапне визначення об'єкту

У контексті виявлення об'єктів, двохетапні детектори складаються з першого етапу, що призначений для генерації набору областей, які можуть містити цільові об'єкти, та другого етапу, котрий зосереджується на класифікації запропонованих областей та створенні обмежувальних рамок. Зазвичай методи виявлення об'єктів виконують три головні кроки:

- 1) Створення декількох пропозицій регіону, що є кандидатами в котрих можуть міститися об'єкти. Для цього використовують такі алгоритми як EdgeBoxes та Selective Search. Після виконання алгоритмів створюються тисячі регіонів.
- 2) З кожного запропонованого регіону витягується вектор ознак фіксованої довжини за допомогою різних дескрипторів зображення, наприклад, гістограми орієнтованих градієнтів (HOG). Далі даний вектор має правильно описати об'єкт, навіть у випадку його збільшення чи переміщення.
- 3) Потім вектор ознак використовується для призначення кожної пропозиції регіону або фоновому класу, або одному з класів об'єктів. Важливо зазначити, що зі збільшенням кількості класів складність побудови моделі, яка може диференціювати всі ці об'єкти, зростає. Однією з популярних моделей, що використовуються для класифікації пропозицій регіону, є опорна векторна машина (SVM).

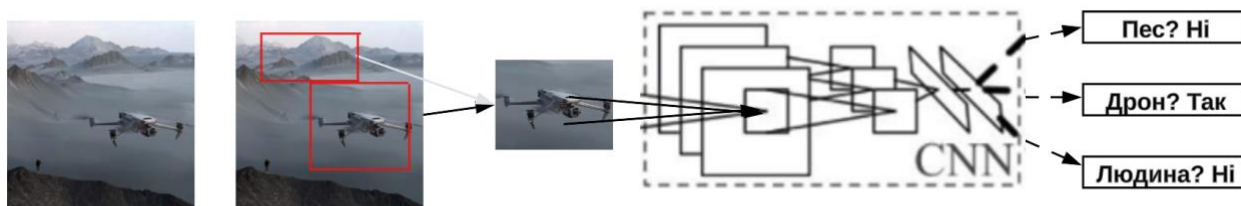


Рисунок 2 - Діаграма з прикладом структури мережі R-CNN

. Ця модель має й недоліки:

- Вона кешує витягнуті ознаки з попередньо навченого CNN на диску, щоб пізніше навчити SVM. Для цього потрібні сотні гігабайт вільної пам'яті.
- R-CNN використовує алгоритм вибіркового пошуку для генерації пропозицій регіонів, що займає багато часу і його не можна налаштувати відповідно до проблеми виявлення.
- Це багатоступенева модель, де кожен етап є незалежним компонентом.
- Не можна у реальному часі працювати, адже кожний запропонований регіон незалежно надсилається на CNN.

Для вирішення перелічених обмежень створили розширену й швидшу Fast R-CNN, що була розроблена Річардом Гіршаком. На перевагу від R-CNN, нова мережа має один етап та додатковий рівень ROI Pooling, що витягує вектори ознак однакової довжини з усіх запропонованих регіонів та дозволяє поділитися обчисленнями згорткового шару одразу з усіма кандидатами регіонів. У цьому також виливає непотреба використання кешування для витягнутих ознак. Тобто на вхід йде зображення як вхідні дані, а повертається обмежувальний прямокутник виявленого об'єкту та ймовірності класу.

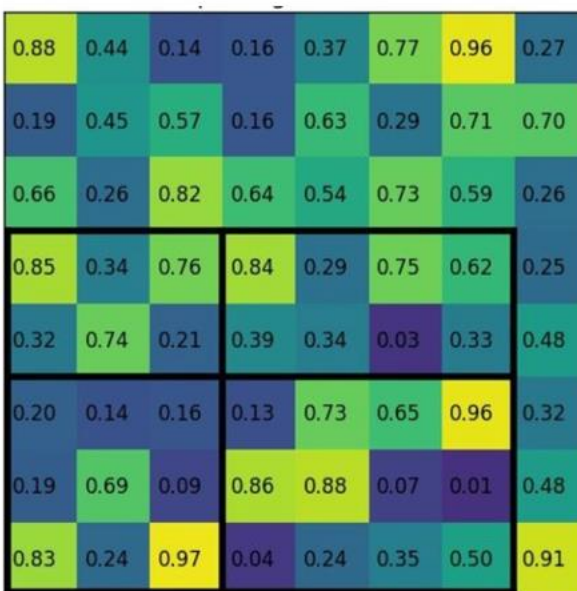


Рисунок 3 - Мапа ознак отримана з CNN шару

На рисунку 3 можна побачити карту ознак з останнього згорткового шару, котра передається на шар ROI Pooling, щоб отримати вектор ознак фіксованої довжини з кожного запропонованого регіону. Рівень ROI Pooling розбиває кожну пропозицію регіону на сітку комірок. Операція максимуму застосовується до кожної комірки сітки для повернення єдиного значення. Усі значення з усіх клітинок представляють вектор ознак. Далі даний вектор передається до деяких шарів FC. Останній вихід з цього шару розділений на 2 гілки:

- 1) Рівень Softmax для прогнозування балів класу;
- 2) Рівень FC для прогнозування обмежувальних рамок виявлених об'єктів.

Незважаючи на переваги моделі Fast R-CNN, існує критичний недолік - це використання вибіркового пошуку для створення кандидатів регіону. Метод вибіркового пошуку не можна налаштувати для конкретного завдання знаходження об'єкта. Таким чином, він може бути неточним для виявлення всіх цільових об'єктів у наборі даних.

Для заміни алгоритму вибіркового пошуку створили Faster R-CNN, що використовує мережу пропозиції регіонів (RPN), яка є повністю згортковою

мережею. RPN генерує пропозиції з різними масштабами та співвідношеннями сторін. Замість використання кількох екземплярів зображення, але в різних масштабах, чи фільтрів із різними розмірами – використовується концепція опорних блоків. Опорний блок – це довідкова коробка певного масштабу та співвідношення сторін. Тобто може існувати декілька масштабів та пропорцій для одного регіону завдяки кільком інформаційним коробкам.

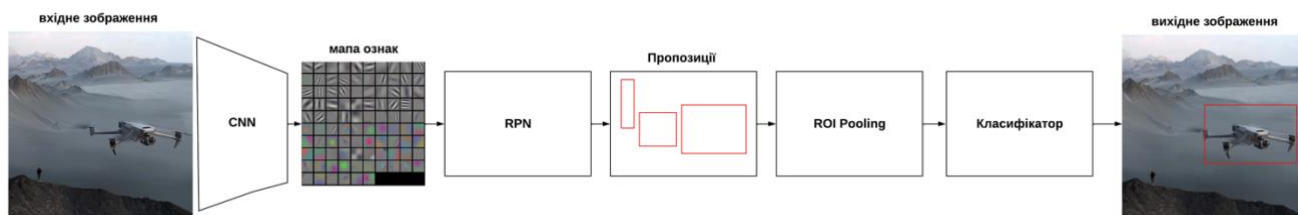


Рисунок 4 - Діаграма із зображенням шарів нейронної мережі Faster R-CNN

Створюється відображення для кожного регіону на відповідну коробку, що дозволяє виявити об'єкти в різних масштабах і з різним співвідношенням сторін. Перевагою RPN над алгоритмом вибіркового пошуку є можливість створення кандидатів регіонів мережею, котру можна навчити та налаштувати відповідно до поставленого завдання. Завдяки спільному використанню однакових згорткових рівнів RPN і Fast R-CNN можна об'єднати в одну мережу. Таким чином, навчання проводиться тільки один раз.

Можна визначати основні два модулі Faster R-CNN:

- 1) RPN: для створення пропозицій регіонів.
- 2) Fast R-CNN: для виявлення об'єктів у запропонованих регіонах.

З RPN рівня пропозиції регіонів фільтруються на основі їх оцінки об'єктивності. Для кожної пропозиції області $n \times n$ витягується вектор ознак. Потім цей вектор передається на 2 однорідні повністю зв'язані шари FC:

- 1) Перший рівень FC називається cls і є двійковим класифікатором, що генерує оцінку об'єктності для кожної пропозиції області (тобто, чи містить область об'єкт чи є частиною фону).
- 2) Другий шар FC - повертає 4-D вектор, що визначає обмежувальну рамку регіону.

Для тренування RPN кожен інформаційний блок отримує позитивну або негативну оцінку об'єктності на основі відношення між площею перетину з опорною коробкою та опорної коробки до площі об'єднання двох коробок. Це називається Intersection-over-Union (IoU) й значення може коливатися від 0,0 до 1,0. Якщо немає перетину, IoU дорівнює 0.0 . Коли 2 коробки наближаються одна до одної, IoU збільшується, поки не досягне 1.0, що буде означати, що коробки ідентичні.

Існує чотири головні умови, щоб визначити, яку оцінку присвоїти інформаційному блоку:

- 1) Інформаційний блок, чиє значення IoU вище 0,7 з будь-яким базовим полем істинності, отримує мітку позитивної об'єктивності.
- 2) Якщо немає інформаційної коробки зі значенням IoU, вищим за 0,7, тоді призначити позитивну мітку інформаційній коробці із найвищим значенням IoU за допомогою ground-truth box.
- 3) Негативний бал об'єктивності призначається непозитивній інформаційній коробці, коли значення IoU для всіх коробок ground-truth менше 0,3. Негативний бал об'єктивності означає, що інформаційна коробка класифікується як фон.
- 4) Інформаційні коробки, що не позитивні, але й не негативні, не корисні для навчання.

Це можна об'єднати у формулу:

$$Objectness_{score}(IoU) = \begin{cases} Positive \rightarrow IoU > 0.7 \\ Positive \rightarrow 0.5 < IoU \leq 0.7 \\ Negative \rightarrow IoU < 0.3 \\ Not\ Negative/Positive \rightarrow 0.3 \leq IoU \leq 0.5 \end{cases}$$

Одним із недоліків Faster R-CNN є те, що RPN навчається, тільки коли всі інформаційні блоки в пакеті розміром 256 витягуються з одного зображення. Оскільки всі зразки з одного зображення можуть бути корельованими процес мережі може зайняти багато часу для досягнення збіжності.

1.3.2 Одноетапне визначення об'єкту

На відміну від двохетапних детекторів, одноетапні детектори не використовують етап вибору областей, а безпосередньо передбачають обмежувальні рамки та клас об'єкта одночасно. Зберігаючи високу точність виявлення, одноетапні детектори здатні працювати у режимі реального часу.

Одноетапні детектори використовують архітектури глибокого навчання, такі як Single Shot MultiBox Detector (SSD), You Only Look Once (YOLO) і RetinaNet, щоб отримувати важливі ознаки та робити точні прогнози. Ці архітектури, як правило, навчаються на великомасштабних наборах даних із анотованими мітками обмежувальних рамок. Це дозволяє їм вивчати складні шаблони та представлення об'єктів. Хоча одноетапні детектори пропонують помітні переваги у швидкості, вони можуть бути не настільки точними, як двоетапні. Однак останні досягнення в дизайні архітектури, методології виділення функцій і стратегії навчання значно скоротили цей розрив у продуктивності. Наприклад такі покращення, що зосередженні на підвищенні точності одноетапних детекторів шляхом включення таких методів, як мережі пірамід функцій, уточнення прив'язки та контекстне моделювання. Таким чином, одноетапні методи виявлення об'єктів забезпечують

ефективне рішення в режимі реального часу для локалізації та ідентифікації об'єктів. Завдяки інтеграції передових архітектури глибокого навчання та методів навчання можна віднайти баланс між швидкістю та точністю.

1.4 YOLO

YOLO (You Only Look Once) є швидким і ефективним алгоритмом для виявлення об'єктів у зображеннях. Він відрізняється від інших методів виявлення об'єктів тим, що виконує виявлення і класифікацію об'єктів одночасно, а не шляхом сканування зображення кілька разів для кожного об'єкта. Цей алгоритм базується на глибинних нейронних мережах і використовує концепцію сіток для поділу вхідного зображення на рівні частини. Кожна сітка відповідає за виявлення об'єктів у певному регіоні зображення. Крім того, для кожної сітки алгоритм передбачає фіксовану кількість рамок (bounding boxes) і відповідних ймовірностей належності до певних класів об'єктів.

У процесі роботи алгоритм здійснює передачу ознак через глибинну нейронну мережу, що дозволяє отримати вектор ознак для кожного пікселя зображення. Далі ці вектори ознак обробляються шаром, що забезпечує виявлення об'єктів, визначення їх координат і розмірів, а також відповідних ймовірностей класифікації. Кінцевим результатом є вектори ознак, що містять інформацію про об'єкти, їх класи і локацію.

Одна з головних переваг YOLO полягає у його швидкості, оскільки він виконує виявлення об'єктів в реальному часі. Алгоритм працює швидко завдяки використанню однієї глибокої нейронної мережі для виявлення об'єктів, у порівнянні з іншими методами, які вимагають багаторазового сканування зображення. YOLO також відомий своєю здатністю до обробки зображень різної роздільної здатності без втрати точності. Алгоритм може працювати з

зображеннями високої роздільності, що робить його придатним для застосування в вимогливих задачах.

Еволюція алгоритмів YOLO почалася з YOLOv1, що використовує підхід на основі сітки, розбиваючи зображення на $S \times S$ комірок однакового розміру. Кожна окрема комірка визначеної сітки відповідає за виявлення об'єктів, якщо центр об'єкта знаходиться в її межах. У кожній комірці прогнозується заздалегідь визначена кількість граничних регіонів, які супроводжуються відповідними довірчими оцінками. Ці межі складаються з п'яти атрибутів: x , y , w , h та довірчої ймовірності. Варто зауважити, що x та y - це координати центру обмежувальної рамки, а w та h - її ширина та висота, відповідно. Для визначення найточнішої межі об'єкта в комірці сітки YOLOv1 використовує вже згадану метрику IOU. Крім того, YOLOv1 використовує техніку, відому як не-максимальне придушення, щоб усунути надлишкові обмежувальні рамки. У випадку, якщо IOU перевищує або дорівнює 0,5, не-максимальне придушення вибірково видаляє надлишкові граничні рамки з низькими оцінками передбачення. Для обчислення функції втрат YOLO використовує суму квадратичних помилок.

У YOLOv2 було запроваджено пакетну нормалізацію в поєднанні з шарами згортки, щоб підвищити точність і зменшити проблеми з надмірним перенавчанням. У YOLOv3 основу виділення ознак було модернізовано з Darknet19, яка мала труднощі з виявленням малих об'єктів, до Darknet53 для усунення цього обмеження. Це вдосконалення включало введення залишкових блоків, пропускових з'єднань і методів підвищеної вибірки, що значно підвищило точність алгоритму. YOLOv4 у свою чергу модифікував основу екстрактора ознак, використовуючи CSPDarknet53, що призвело до значного покращення швидкості та точності визначення. YOLOv5 представляє останню ітерацію і є полегшеним варіантом попередніх алгоритмів YOLO, використовуючи фреймворк PyTorch замість Darknet. Тобто зазначається, що голова і тип нейронної мережі однакові для всіх алгоритмів, тоді як скелет (backbone), шия (neck) і функція втрат відрізняються.

В YOLOv3 архітектура Darknet53 слугує основою для вилучення ознак в алгоритмі YOLO. Кістяк, який складається зі згорткових шарів, відіграє вирішальну роль у вилученні важливих ознак із вхідного зображення. Крім того, YOLOv3 включає мережу піраміди ознак (feature pyramid network, FPN) як компонент "шиї". Шия відповідає за вилучення карт ознак з різних етапів, використовуючи комбінацію висхідних і низхідних шляхів. Потім ці карти особливостей передаються до "голови" мережі, яка складається з шару YOLO.

Голова мережі YOLOv3, яка є одноступеневим детектором, виконує остаточне передбачення, генеруючи вектор, що містить координати обмежувальних рамок (ширина, висота), мітку класу та ймовірність класу. Вхідне зображення спочатку обробляється Darknet53 для виділення релевантних ознак. Згодом ці ознаки подаються в мережу піраміди ознак для виконання злиття ознак і покращення репрезентативності. Нарешті, шар YOLO використовує об'єднані ознаки для отримання бажаних результатів, таких як виявлення та класифікація об'єктів.

У YOLOv4, модифікованій версії YOLOv3, основу для вилучення ознак замінено на Cross Stage Partial Network (CSPNet) в Darknet, що призвело до появи нового дистрибутора ознак під назвою CSPDarknet53. Згорткова архітектура CSPDarknet53 базується на модифікованій версії DenseNet. Ця архітектура передає копію мапи ознак з базового шару на наступний через щільний блок. Використання DenseNet має кілька переваг, включно з вирішенням проблеми зникнення градієнта, полегшення зворотного поширення, усунення обчислювальних вузьких місць та покращення процесу навчання.

"Шия" YOLOv4 складається з шару просторового об'єднання пірамід (SPP) і мережі агрегації шляхів (PANet). Шар SPP і агрегація шляхів PANet використовуються для агрегації ознак, покращення рецептивного поля і відбору важливих ознак з кістяка. "Головою" мережі залишається шар YOLO, який відповідає за виконання остаточних прогнозів.

Як і в YOLOv3, зображення спочатку обробляється CSPDarknet53 для виділення релевантних ознак. Потім ознаки проходять через мережу агрегації шляхів (Path Aggregation Network, PANet) для злиття. Нарешті, шар YOLO генерує бажані результати, такі як виявлення та класифікація об'єктів. YOLOv4 включає в себе використання методів "bag of freebies" і "bag of specials" для підвищення продуктивності алгоритму. Пакет "bag of freebies" відповідає за інтеграцію різних поліпшень, таких як повна втрата IOU (CIoU), регуляризація блоків падіння і різні методи аугментації. З іншого боку, пакет спеціальних можливостей представляє специфічні покращення, такі як функція активації Mish, Diou-NMS (Non-Maximum Suppression) та модифікації мереж агрегації шляхів. Вищезазначені доповнення та модифікації в YOLOv4 спрямовані на підвищення продуктивності та точності алгоритму, що робить його більш ефективним у задачах виявлення об'єктів.

1.4.1 YOLOv5

У порівнянні з іншими алгоритмами YOLOv5 застосовує фреймворк PyTorch замість Darknet. Оскільки CSPDarknet53 ефективно вирішує проблему повторюваності градієнтної інформації у великих "скелетах". Скелет" інтегрує градієнтні зміни в мапи особливостей, що призводить до зменшення часу для отримання результату, підвищення точності та зменшення розміру моделі за рахунок зменшення кількості параметрів.

Для компоненти "шиї" YOLOv5 використовує мережу агрегації шляхів (PANet), яка покращує потік інформації. PANet використовує нову архітектуру Feature Pyramid Network (FPN), яка включає декілька висхідних і низхідних шарів. Ця архітектура покращує розповсюдження низькорівневих ознак в моделі, тим самим підвищуючи точність локалізації. PANet особливо покращує локалізацію в нижніх шарах, що призводить до покращення локалізації об'єктів. Подібно до

YOLOv4 і YOLOv3, "голова" в YOLOv5 генерує три різні набори мап особливостей, що дозволяє здійснювати різномасштабне прогнозування і ефективно обробляти об'єкти різного розміру. Такий підхід підвищує точність прогнозування для об'єктів від малого до великого розміру. Зображення спочатку обробляється CSPDarknet53 для виділення особливостей, після чого відбувається злиття особливостей у компоненті PANet. Шар YOLO генерує бажані результати, такі як виявлення та класифікація об'єктів.

Рисунок 5 надає огляд архітектури алгоритму YOLOv5l. Слід зазначити, що YOLOv5 вводить шар "Фокус", який є еволюцією структури YOLOv3. Він замінює перші три шари YOLOv3 і об'єднує їх в один шар в YOLOv5. Крім того, архітектура включає різні компоненти, такі як шари згортки (Conv), модуль з каскадними вузькими місцями (C3), просторове пірамідалльне об'єднання (SPP) для гнучкого об'єднання, апсемплінг (Upsample) для збільшення масштабу злиття попередніх шарів, конкатенація (Concat) для нарізки попередніх шарів і модулі виявлення (Conv2d) в голові мережі.

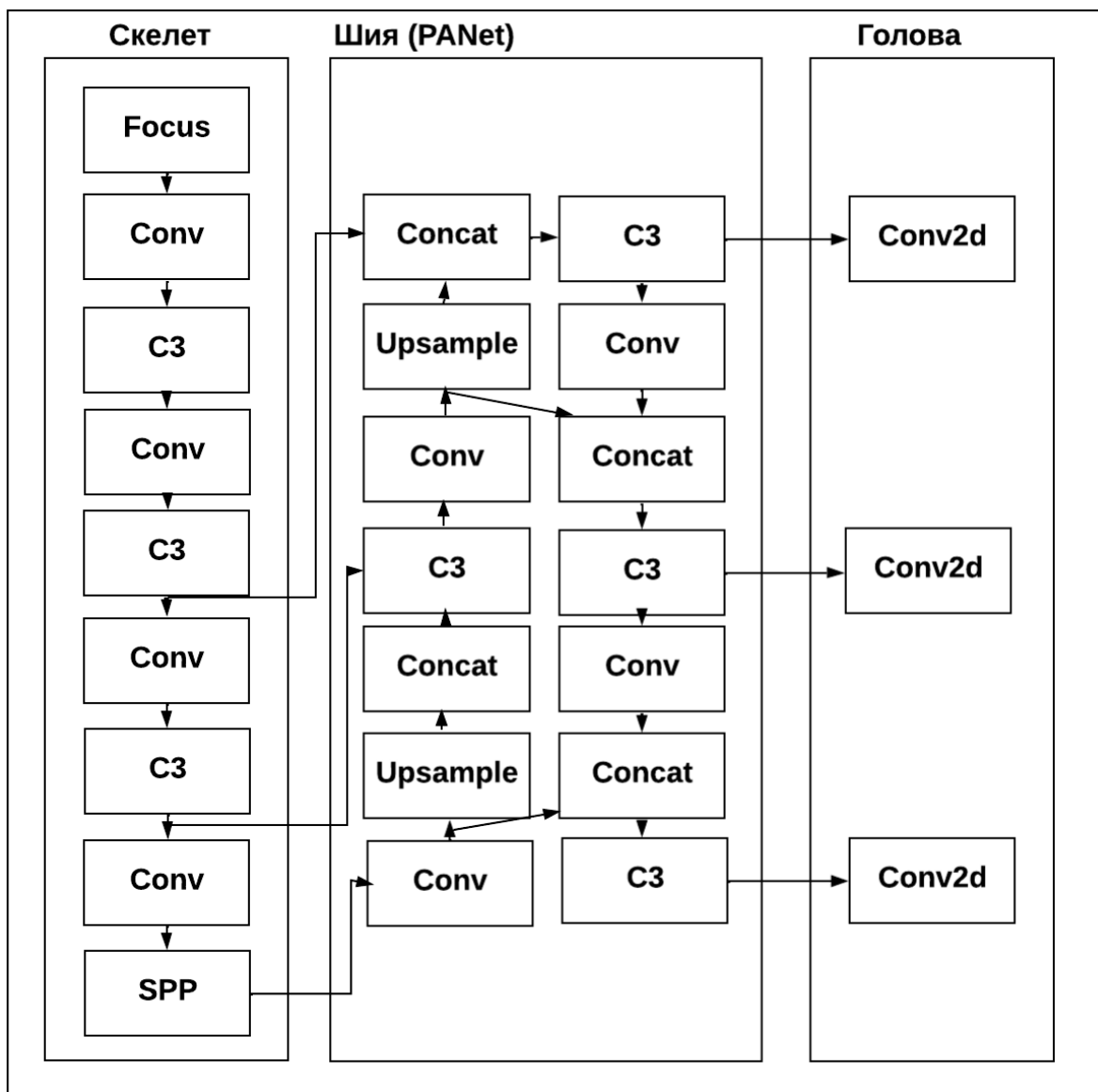


Рисунок 5- Архітектурна схема моделі YOLOv5

Ці архітектурні елементи та модифікації в YOLOv5 сприяють покращенню продуктивності моделі, що дозволяє ефективно і точно вирішувати задачі виявлення об'єктів.

Основні архітектурні відмінності між YOLOv3, YOLOv4 та YOLOv5:

- 1) YOLOv3 використовує кістяк Darknet53 для вилучення ознак. Darknet53 складається зі загорткових шарів, які відповідають за виділення релевантних ознак із вхідного зображення.
- 2) У YOLOv4 архітектура використовує CSPdarknet53 як основу. CSPdarknet53 - це покращена версія Darknet53, яка вирішує проблему повторюваності градієнтної інформації. Вона інтегрує зміни градієнта в карти особливостей, що призводить до покращення швидкості виводу, точності та зменшення розміру моделі.

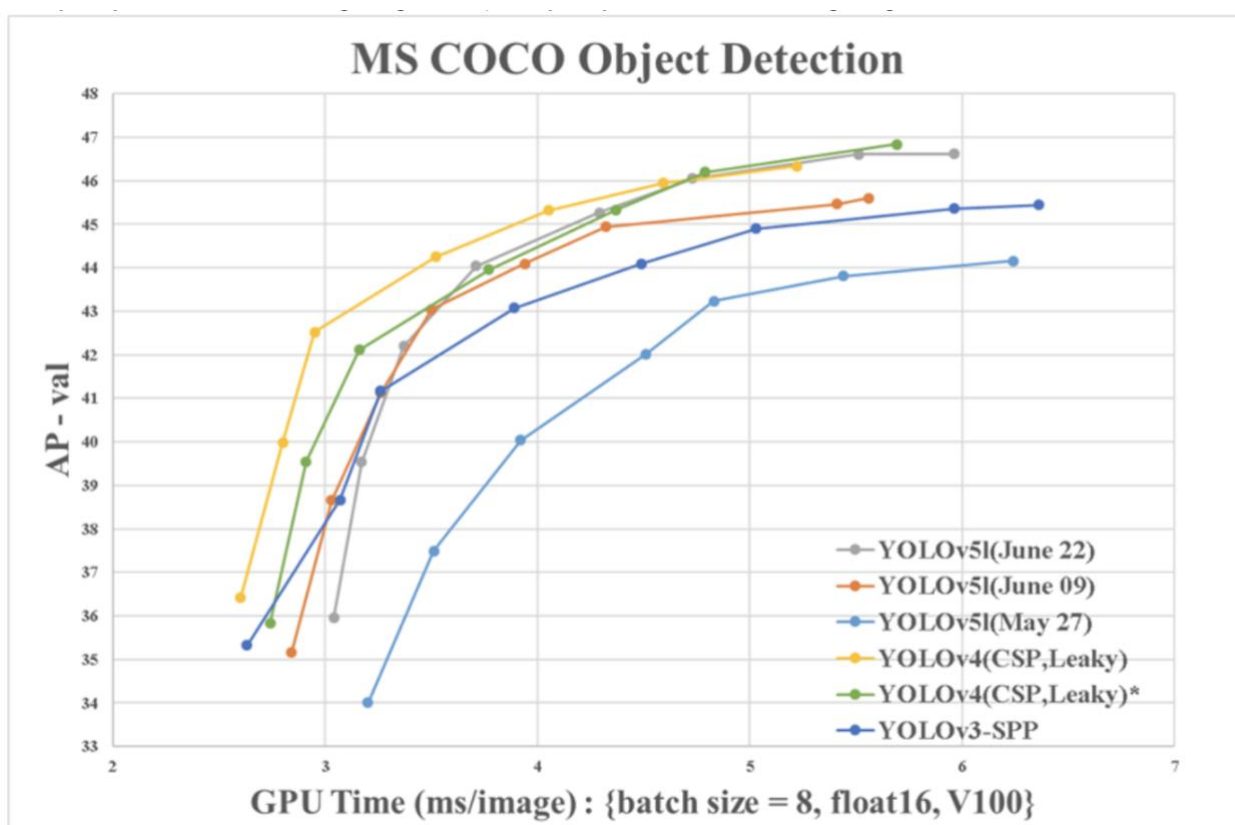


Рисунок 6 - Порівняння швидкості YOLO моделей на наборі даних COCO

- 3) YOLOv5 представляє структуру Focus, яка є новим доповненням до архітектури. Шар Focus замінює перші три шари алгоритму YOLOv3. Переваги використання шару Focus включають зменшення вимог до пам'яті CUDA, зменшення кількості шарів та підвищення ефективності прямого та зворотного поширення.

1.5 Проблема визначення малих об'єктів

Однією з визначених підпроблем є визначення об'єктів, що можуть мати малий розмір на вхідному зображенні. Традиційно малий об'єкти визначають як такий, чий регіон займає менше ніж 10% площі зображення, або мають розмір менший-рівний 32x32 пікселі. Виявлення дрібних об'єктів на зображеннях з низькою роздільною здатністю може бути складним завданням через відсутність характерних форм і текстур. У 2022 році було створено багато методів, спрямованих на покращення виявлення малих об'єктів.

1.5.1 Плиткове розбиття

Мережі виявлення об'єктів, такі як YOLOv3, YOLOv4 і YOLOv5, мають певні розміри вхідних зображень. Наприклад, YOLOv3 і YOLOv4 зазвичай приймають вхідні зображення розміром 416×416 пікселів, тоді як YOLOv5 очікує вхідні зображення розміром 640×640 пікселів. Якщо вхідне зображення перевищує вказаний розмір, його потрібно змінити, щоб воно відповідало вимогам мережі. Тобто для зображення розміром 2560×2560 пікселів, яке містить невеликий об'єкт, наприклад, дрон розміром 16×16 пікселів. Коли це зображення передається на вхід YOLOv5, його розмір буде змінено до 640×640 пікселів, щоб відповідати розміру вхідних даних мережі. Отже, дрон, який спочатку мав розмір 16×16 пікселів, зменшиться до 4×4 пікселів завдяки процесу зміни розміру.

Крім того, ознаки малих об'єктів можуть стати надзвичайно малими або навіть зникнути взагалі після проходження таких операцій, як згортка шарів та максимальне об'єднання шарів. Це явище створює проблему для точного прогнозування, оскільки зменшені або зниклі ознаки можуть перешкоджати

здатності мережі ефективно виявляти та розпізнавати малі об'єкти. Одним з підходів до вирішення проблеми зміни розміру вхідних зображень є метод розбиття на плитки. Цей метод передбачає поділ зображення на менші ділянки або плитки, які потім передаються мережі окремо. Таким чином, відпадає потреба у зміні розміру вхідного зображення, а деталізація об'єктів може бути збережена, оскільки роздільна здатність об'єктів залишається незмінною.

Однак такий підхід має певні компроміси. При збереженні деталізації об'єктів значно збільшується час виведення, оскільки кожен ділянку зображення потрібно окремо подавати в мережу. Крім того, потрібен додатковий час для обрізання зображення на фрагменти та об'єднання прогнозованих обмежувальних рамок, отриманих з окремих фрагментів. Незважаючи на потенційне збільшення часу виведення, метод плиткового розбиття дозволяє зберегти цілісність детальних ознак і може бути корисним у сценаріях, де збереження дрібнозернистої інформації має вирішальне значення для точного виявлення об'єктів.

2 Методологія

У даному розділі описано проведені експерименти та їх особливості. Методологію збору наборів навчальних даних та їх застосування для вирішення поставленої задачі.

2.1 Збір початкових даних

Збір даних є одним з важливих етапів в підготовці навчального набору для тренування моделі YOLO. Цей етап передбачає збирання великої кількості зображень, на яких відображені об'єкти, що підлягають детекції алгоритмом.

Для роботи моделі потрібно визначення об'єкти, які необхідно визначити, та їх розмітка або позначення на зображеннях, що будуть подаватися на вхід. Цей процес може бути виконаний вручну людиною або за допомогою автоматизованих інструментів для розмітки даних. Під час збору даних було враховано різноманітні аспекти, такі як різноманітність об'єктів, освітлення, перспектива, масштаб, а також можливі зміни в умовах зйомки, адже була потреба у створенні репрезентативного набору даних, котрий містив широкий спектр сценаріїв, щоб забезпечити роботу алгоритму в різних умовах.

Після збору дані були підготовлені для використання в навчанні YOLO моделей. Було виконано процес попередньої обробки, такий як масштабування зображень, нормалізацію пікселів, аугментацію даних та створення файлів конфігурації, які містять інформацію про розміри зображень, шляхи до зображень та анотації.

2.1.1 Джерела наборів даних

Оскільки в Україні через військовий стан використання дронів обмежене, було вирішено отримати дані з публічно доступних джерел. Для знаходження та подальшого керування датасетами було вирішено скористатися Roboflow. Roboflow - це хмарна платформа для комп'ютерного зору, яка надає інструменти та інфраструктуру для спрощення процесу керування, попередньої обробки та розширення наборів даних зображень для проектів машинного навчання. Вона пропонує широкий спектр функцій та можливостей, щоб допомогти дослідникам та розробникам оптимізувати свої робочі процеси з комп'ютерним зором.

Було вирішено скористатися нею через нижче зазначені переваги:

- **Управління наборами даних:** Roboflow дозволяє користувачам організовувати та зберігати свої набори даних зображень в централізованому місці. Вона надає засоби для версіонування даних, співпраці та зручного доступу до набору даних для навчання та оцінки моделей.
- **Попередня обробка даних:** платформа Roboflow спрощує процес попередньої обробки зображень, надаючи різні опції трансформації. Вона підтримує зміну розміру, обрізку, поворот та застосування фільтрів до зображень, що дозволяє користувачам попередньо обробляти набори даних відповідно до їх конкретних вимог.
- **Розширення даних:** Roboflow пропонує різні техніки розширення даних для збільшення різноманітності набору даних. Вона включає такі опції як випадкові відображення, обертання, зміна яскравості та інші. Розширення даних допомагає покращити узагальненість та стійкість моделі.
- **Інтеграція та експорт:** Roboflow легко інтегрується з популярними фреймворками та бібліотеками машинного навчання. Вона підтримує експорт

наборів даних у різних форматах, сумісних з популярними фреймворками глибокого навчання, такими як TensorFlow та PyTorch.

- Моніторинг продуктивності: Roboflow надає метрики продуктивності та візуалізації для оцінки роботи навчених моделей. Це допомагає користувачам аналізувати точність, повноту та інші метрики моделі для оцінки її ефективності.

Усі набори даних, що були використані в експериментах можна розділити на кілька категорій:

- 1) Зображення зняті на камеру без спеціального обладнання
- 2) Зображення створені шляхом симуляції

Набір зображень знятих на камеру без спеціального обладнання був створений шляхом об'єднання кількох датасетів знайдених на платформі Roboflow. Датасети складаються з фотографій квадрокоптерів на різній відстані від місця зйомки.

Зображення створені шляхом симуляції були створені власноруч у AirSim. AirSim - це відкритий крос-платформний симулятор, розроблений компанією Microsoft для досліджень в галузі автономних транспортних засобів. Він є потужним інструментом для тренування та тестування алгоритмів, пов'язаних з сприйняттям, керуванням та плануванням в сфері автономних систем. AirSim переважно призначений для моделювання поведінки безпілотних повітряних апаратів (БПЛА), але також підтримує наземні транспортні засоби. Однією з важливих особливостей AirSim є підтримка датчиків, які часто використовуються в автономних системах, таких як камери, LIDAR та GPS. Ці віртуальні датчики генерують синтетичні дані, які наближаються до реальних даних з датчиків. Це дозволило зібрати велику кількість визначених даних для тренування алгоритмів сприйняття без потреби у фізичних датчиках або зборі реальних даних. Крім того, AirSim надає багатий набір інструментів для генерації сценаріїв, що дозволяє

створювати складні сценарії з динамічними об'єктами, погодними умовами та дорожнім рухом. Що допомогло робити зображення на фоні дерев та різноманітних об'єктів. Було створено зображення з використанням DJI Mavic 2 Pro, DJI Phantom Inspire I та DJI Phantom 4 Pro. Створений набір зображень був використаний для порівняння ефективності моделі, Важливо зазначити що усі зображення в наборах даних мали розмір 640 x 640.

2.2 Застосування плиткового розбиття

Оскільки виникає проблема визначення малих об'єктів було вирішено застосувати розбиття на плитки, щоб зберегти зображення дронів і їх навчальну роль. Зображення з симуляції мали розмір HD та Full HD, тобто 1280 x 720 та 1920 x 1080 відповідно. Щоб не втратити зображення малих дронів у даній роботі застосовано плиткування з накладанням щоб визначений об'єкт не було поділено. За ширину накладання було взято найбільшу довжину границі серед об'єктів з набору даних. Оскільки як для YoLov5 та YoLov8 моделей вхідні зображення мають мати розмір 640x640 – усі зображення з набору даних було поділено на 8 плиток з різною довжиною накладання. Було застосовано формулу () для обрахунку довжини накладання, де L_n – це довжина накладання плиток, n – кількість плиток горизонтально, чи вертикально, L_p – довжина плитки вертикально, чи горизонтально, L_3 – справжня довжина оригінального зображення.

$$L_o = \frac{nL_{\Pi} - L_3}{n - 1} \quad (2.1)$$

Після розбиття зображення на плитки, необхідно також було адаптувати мітки до нових координат. Усі мітки оригінальних зображень використовують формат YOLO. Для кожного зображення є “.txt” файл з таким самим ім'ям, що містить анотації об'єктів, включаючи клас об'єкту, координати x-y, ширину і висоту. Останні чотири значення x, y, w, h обчислюються за наступними рівняннями:

$$x = \frac{x_{\min} + x_{\max}}{2} \frac{1}{w_{image}}$$

$$y = \frac{y_{\min} + y_{\max}}{2} \frac{1}{h_{image}}$$

$$w = \frac{x_{\max} - x_{\min}}{w_{image}}$$

$$h = \frac{y_{\max} - y_{\min}}{h_{image}}$$

(2.2)

Система координат використовує піксельні індекси, де піксель служить базовою одиницею. Початкова точка знаходиться в верхньому лівому куті зображення, а індекси збільшуються зліва направо і зверху вниз. Клас об'єкту відображається як ціле число, тоді як решта значень (x, y, w, h) виражені у вигляді чисел з рухомою комою від 0 до 1. Ці значення обчислюються згідно з наведеними рівняннями. За координати верхнього лівого кута границі відповідають $x_{\min}$ та $y_{\min}$, а $x_{\max}$ та $y_{\max}$ - це координати нижнього правого кута границі. Крім того, w_{image} вказує ширину зображення, а h_{image} - його висоту.

2.3 Обрані метрики

Для знаходження та вибору найкращих моделей з поміж інших було вирішено скористатися тестом Фрідмена та Неменного. Для оцінки продуктивності моделей Використовувалися такі метрики як точність, відклик та результативний час.

2.3.1 Точність та повнота

Для оцінки продуктивності моделей визначення об'єктів часто використовують концепт перетину об'єднання. IoU Це значення обчислюється як співвідношення перетину до об'єднання двох обмежувальних рамок: обмежувальних рамок істинного положення та передбачених обмежувальних рамок. Обмежувальні рамки істинного положення вручну анотовані, а передбачені обмежувальні рамки є результатами роботи моделі. За значенням IoU можна оцінити, наскільки точно передбачена рамка відповідає рамці істинного положення. Для проведених експериментів було встановлено поріг величиною 0.65 для визначення, чи вважати виявлений об'єкт дійсним. Залежно від значення IoU застосовуються такі класифікації:

- Якщо $\text{IoU} \geq 0.65$, виявлений об'єкт вважається правильним позитивним (TP).
- Якщо $\text{IoU} \leq 0.65$, виявлений об'єкт вважається неправильним позитивним (FP).
- Випадок, коли модель не виявляє істинне положення об'єкта на зображенні, класифікується як неправильний негативний (FN).
- Крім того, на зображеннях можуть бути частини без передбачених або істинних об'єктів, які класифікуються як без інформаційні області (TN).

Для обрахунку точності та повторюваності використовується значення правильно позитивної оцінки, неправильно позитивної та неправильно негативної, у випадку обрахунку відклику.

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN}$$

(2.3)

Точність (Precision) вказує, чи правильно відбувається виявлення об'єктів, а показник повторюваності (Recall) вказує, чи вдалося виявити всі об'єкти.

2.3.2 Результативний час

Останньою оцінкою ефективності моделі є результативний час. Оскільки запланована система має працювати в реальному часі необхідно щоб модель досягала швидкості більше ніж 10 оброблених кадрів за секунду, тобто результативний час для обробки одного зображення має бути не більше 100 мс. Оскільки початкове зображення має бути розбите на 8 плиток для ефективного виявлення результативний час має бути не більше 13 мс для обробки однієї плитки.

2.3.3 Тест Фрідмена та Вілкоксона для статистичного аналізу

У статистичному висновку головною метою є перевірка гіпотез. Цей процес полягає у відборі вибірки з популяції та оцінці, наскільки ця вибірка може репрезентувати популяцію. У рамках гіпотезного тестування формулюються

нульова гіпотеза (H_0) і альтернативна гіпотеза (H_1) на основі вивчення проблеми порівняння класифікаторів. Обидві гіпотези мають бути стислими твердженнями, що мають дві версії істини. Ці версії повинні бути взаємовиключними та вичерпними, що означає, що вони охоплюють всі можливості. У контексті порівняння моделей:

- H_0 : Класифікатори рівні, тобто середні значення метрик не відрізняються.
- H_1 : Класифікатори різні, середнє значення метрик для однієї чи декількох моделей відрізняються.

Для цього було взято таблицю з середніми значеннями для кожної метрики по кожній моделі. Далі було створено таблицю рангів по значенням метрик. Далі було знайдено ксі-квадрат за формулою:

$$X_F^2 = \frac{12N}{k(k+1)} \left[\sum_{j=1}^k R_j^2 - \frac{k(k+1)^2}{4} \right] \quad (2.4)$$

Де, $N = 3-4$ (кількість рядків), k (кількість моделей). R – середнє значення рангу. Маючи значення ксі-квадрату можна обрахувати FF :

$$F_F = \frac{(N-1)x_F^2}{N(k-1) - x_F^2} \quad (2.5)$$

У подальшому значення FF ми порівнюємо з критичним значенням $F_{k-1, (k-1)(n-1)}$. Якщо FF більше за критичне значення H_0 спростовано і можна зробити висновок, що між класифікаторами існує велика різниця.

Для порівняння пар моделей було використано ранговий тест Вілкоксона.

У контексті порівняння моделей YOLO, тест Вілкоксона є непараметричним статистичним тестом, який дозволяє визначити наявність статистично значущої різниці між двома незалежними вибірками. Для застосування тесту Вілкоксона маємо дві вибірки результатів моделей YOLO. Позначимо їх як вибірку А та вибірку В. Використовуючи ці вибірки, можна обчислити ранги для кожного значення в кожній вибірці. Далі, для кожного спареного значення (значення з вибірки А та відповідне значення з вибірки В) обчислюємо різницю рангів. Відносні ранги можуть бути використані, якщо вибірки мають різний розмір або мають однакові значення. Після цього обчислюється сума рангів різниць, яка слугує статистикою тесту. Значення цієї статистики в порівнянні з таблицею критичних значень дозволяє зробити висновок про наявність статистично значущої різниці між моделями YOLO. Тест Вілкоксона має нульову гіпотезу, якщо розподіли вибірок А та В є однаковими. Альтернативна гіпотеза стверджує, що розподіли вибірок відрізняються. За допомогою визначення статистики та порівняння її з критичним значенням, можна прийняти рішення щодо відхилення або прийняття нульової гіпотези і у свою чергу обрати кращу модель.

2.4 Проведені експерименти

2.4.1 Оцінка YOLO моделей

Було проведено кілька експериментів, що допомогли обрати найкращу YOLO модель ґрунтуючись на поставленій задачі визначення малих повітряних об'єктів у реальному часі. Розділили експерименти на три умовних частини:

- 1) Порівняння модифікованих моделей версії YOLOv5, а саме N, S, M та L, що використовуються для позначення різних варіантів або розмірів моделі. Ці літери представляють різні масштаби та конфігурації архітектури моделі

YOLOv5. Тренування та обрання найкращої на основі отриманих результатів та проведених тестів.

- 2) Порівняння найкращої модифікованої YOLOv5 моделі зі стандартною конфігурацією.
- 3) Порівняння модифікованих моделей версії YOLOv8, а саме N, S, M. У випадку моделі YOLOv8 було вирішено не розглядати модель конфігурації L через довший результативний час, що не підходить для поставленої задачі обробки в реальному часі. Тренування та обрання найкращої моделі на основі отриманих результатів та проведених тестів.

Усі проведені експерименти з порівняння моделей були виконані на наборі даних зі справжніми фотографіями розміром 4008 фотографій з різноманітними квадрокоптерами. Набір поділено пропорційно 70% навчальні дані, 20% валідаційні та 10% тестові. Кожна модель була тренована на графічному процесорі на трьохста епохах по 16 зображень в пакеті.

2.4.2 Еволюція конфігураційних параметрів

Для найкращої моделі було вирішено провести еволюцію гіперпараметрів шляхом застосування генетичного алгоритму. Оскільки YOLOv8 було створене на основі YOLOv5 вони мають 30 гіперпараметрів. Під час ініціалізації еволюції було використано стандартні гіперпараметри, що використовувалися під час попередніх навчань. Функція здоров'я представлена як зважена комбінація метрик: $mAP@0.5$, $mAP@0.5:0.95$, а також метрик точності та повноти. Еволюція виконувалася 600 поколінь. Усі гіперпараметри кожного покоління завантажуються разом з оцінкою ефективності цього покоління. Далі 5 найкращих поколінь додаються в список. З кожного з 5 поколінь віднімається значення оцінки ефективності, яке найменше в

цьому поколінні. Поколінню з найнижчою оцінкою ефективності присвоюється значення оцінки 10^{-6} . 5 поколінь тепер сортуються випадковим порядком з ваговими коефіцієнтами, які визначаються зменшеною оцінкою ефективності з кроком 3. Обирається найкраще покоління зі списку. Далі виконується мутація без кросоверу. Після чого виконується нова оцінка функцією здоров'я та навчання з новим набором параметрів.

2.4.3 Застосування плиткування на наборі даних з симуляції

Для розуміння ефективності методу плиткового розкриття, було порівняно ефективність моделей, що були навчені на початковому наборі даних з симуляції та такому ж наборі даних розбитих на плитки.

2.5 Системна документація

2.5.1 Конфігураційні значення

Усі YOLO моделі під час порівняння мали однаковий набір гіперпараметрів з Таблиці 2:

Гіперпараметр	Значення	Коментарі
lr0	0.01	початкова швидкість навчання
lrf	0.01	кінцева швидкість навчання OneCycleLR ($lr0 \times lrf$)
momentum	0.937	моментум SGD / beta1 Adam
weight decay	0.0005	зважене зниження оптимізатора
warmup_epochs	3.0	епохи підготовки
warmup_momentum	0.8	початковий моментум підготовки

warmup_bias_lr	0.1	початковий коефіцієнт зміщення підготовки
box	0.05	коефіцієнт втрати для області
cls	0.5	коефіцієнт втрати для класу
cls_pw	1.0	позитивна вага cls BCELoss
obj	1.0	коефіцієнт втрати для об'єкту
obj_pw	1.0	позитивна вага obj BCELoss
iou_t	0.20	поріг навчання IoU
anchor_t	4.0	поріг множника якорів
anchors	3	якорі на вихідному шарі
fl_gamma	0.0	гамма фокусної втрати
hsv_h	0.015	аугментація відтінку зображення HSV (частка)
hsv_s	0.7	аугментація насиченості зображення HSV (частка)
hsv_v	0.4	аугментація значення яскравості зображення HSV
degrees	0.0	обертання зображення (+/- градуси)
translate	0.1	переклад зображення (+/- частка)
scale	0.5	масштабування зображення (+/- множник)
shear	0.0	зсув зображення (+/- градуси)
perspective	0.0	перспектива зображення (+/- частка), діапазон 0-0.001
flipud	0.0	відображення зображення вгору-вниз
fliplr	0.5	відображення зображення ліворуч-праворуч
mosaic	1.0	мозаїка зображень (ймовірність)
mixup	0.0	змішування зображень (ймовірність)
copy_paste	0.0	копіювання та вставка сегментів (ймовірність)

Таблиця 4 - Список початкових гіперпараметрів для усіх моделей

2.5.2 Середовище розробки

Створення моделей їх тестування, основні інструменти аналізу, що були використані наведені у таблиці .

НАЗВА	ВЕРСІЯ
GOOGLE COLAB	11.8
TORCH	2.0.1
MATPLOTLIB	3.7.1
NUMPY	1.23.4
PANDAS	2.0.2
SCIPY	1.10.1
ULTRALYTICS	8.0.114
PYTHON	3.10.11
CUDA	11.8

Таблиця 5 – Інструменти та середовища розробки

Специфікація обладнання на котрому виконувалася розробка та виконувалися експерименти подані у таблиці 6.

КОМПОНЕНТ	СПЕЦИФІКАЦІЯ
CPU	Intel(R) Xeon(R) CPU @ 2.20GHz
GPU	NVIDIA Tesla T4 GPU 16GB GDDR6
MEMORY	32 GB 256-bit 137 GB/s

Таблиця 6 – Специфікація обладнання

3 Аналіз результатів

3.1 Вибір навчальної моделі

3.1.1 Навчання моделей YOLOv5

Метрики за результатами навчання моделей YOLOv5 продемонстровані у таблиці 3.1. Цікаво, що одні з найкращих результатів точності та повноти має модель YOLOv5s. Найшвидшою моделлю є YOLOv5n, хоча різниця у результативному часі є незначною якщо порівнювати з YOLOv5s.

<i>Модель</i>	<i>Точність</i>	<i>Повнота</i>	<i>Результативний час (ms)</i>
<i>YOLOv5n</i>	92.8 ± 0.5	90.0 ± 0.4	2.1 ± 0.3
<i>YOLOv5s</i>	94.4 ± 0.3	93.3 ± 0.3	2.7 ± 0.4
<i>YOLOv5m</i>	94.0 ± 0.3	90.3 ± 0.2	3.8 ± 0.5
<i>YOLOv5l</i>	92.6 ± 0.2	90.8 ± 0.3	6.1 ± 0.3
<i>Тест Фрідмена p-value</i>	0.014	0.013	0.007

Таблиця 3.1 – Середні значення метрик та p-value за тестом Фрідмена для YOLOv5

Для усіх метрик використовувався рівень значущості $\alpha = 0.05$, тому можна помітити, що усі метрики мають статистично значні відмінності. Для знаходження

відмінною моделі застосовувався тест Неманного. Результати можна побачити нижче на малюнку.

	n	s	l	m
n	1.000000	0.9	0.593413	0.900000
s	0.900000	1.0	0.900000	0.900000
l	0.593413	0.9	1.000000	0.747133
m	0.900000	0.9	0.747133	1.000000

Рисунок 3.1 - результати тесту Неманного для метрики точності

	n	s	l	m
n	1.000000	0.670273	0.900000	0.900000
s	0.670273	1.000000	0.823993	0.516551
l	0.900000	0.823993	1.000000	0.900000
m	0.900000	0.516551	0.900000	1.000000

Рисунок 3.2 - результати тесту Неманного для метрики повноти

	n	s	l	m
n	1.000000	0.065622	0.823993	0.900000
s	0.065622	1.000000	0.354855	0.065622
l	0.823993	0.354855	1.000000	0.823993
m	0.900000	0.065622	0.823993	1.000000

Рисунок 3.3 - результати тесту Неманного для метрики результативної швидкості

Проаналізувавши таблиці можна помітити, що YOLOv5n та YOLOv5s мають значно різні середні значення результативної точності, повноти та швидкості. Цікаво, що повнота була доволі відмінною між YOLOv5m та YOLOvS. Результативна

швидкість набагато краща у YOLOv5s у порівнянні з YOLOv5L. Зважаючи на отримані результати було вирішено обрати серед моделей YOLOv5s.

3.1.2 Навчання моделей YOLOv8

Метрики по результатам навчання моделей YOLOv5 продемонстровані у таблиці 3.2. Можна помітити, що YOLOv8s має значні переваги у точності та повноті. За швидкістю на першому місці перебуває YOLOv8n.

Модель	Точність	Повнота	Результативний час (ms)
YOLOv8n	97.9 ± 0.2	69.8 ± 0.2	3.2 ± 0.3
YOLOv8s	98.3 ± 0.3	70.6 ± 0.4	5.4 ± 0.4
YOLOv8m	96.6 ± 0.2	63.5 ± 0.4	11.2 ± 0.5
Тест Фрідмена p-value	0.049	0.049	0.049

Таблиця 3.2 – Середні значення метрик та p-value за тестом Фрідмена для YOLOv8

	n	s	m
n	1.000000	0.900000	0.326531
s	0.900000	1.000000	0.559666
m	0.326531	0.559666	1.000000

Рисунок 3.4 - результати тесту Неминного для метрики точності YOLOv8

	n	s	m
n	1.00000	0.44041	0.90000
s	0.44041	1.00000	0.44041
m	0.90000	0.44041	1.00000

Рисунок 3.5 - результати тесту Неминного для метрики повноти YOLOv8

	n	s	m
n	1.000000	0.793203	0.440410
s	0.793203	1.000000	0.157722
m	0.440410	0.157722	1.000000

Рисунок 3.6 - результати тесту Неминного для метрики результативної швидкості YOLOv8

Значення p тесту Фрідмана демонструють, що метрики мають статистичні відмінності. Зважаючи, що для тесту Неменного метрики результативного часу рівень значущості дорівнює 0.5, то модель YOLOv8n виконує обрахунки набагато швидше від YOLOv8m. Тому було вирішено обрати модель YOLOv8s.

3.1.3 Порівняння результатів навчених моделей

Усі моделі YOLO були ефективними у визначенні дронів на зображеннях, особливо з урахуванням того, що майже всі дрони в наборі даних були надзвичайно маленькими. Важливо розуміти, що набір даних, використаний у цьому

експерименті, включав лише клас дронів. Також зображення у навчальному датасеті були зняті у сприятливому середовищі, в умовах хорошого освітлення. Оцінюючи моделі, можна помітити, що YOLOv8 значно перевершує клас YOLOv5 моделей за точністю, однак за повнотою та результативною швидкістю кращим вибором буде YOLOv5. Проте, навіть не зважаючи на це результативна швидкість більшої моделі YOLOv8 дозволяє використовувати модель у реальному часі для знаходження безпілотних літальних апаратів на фото чи потоковому відео.

3.2 Вплив на результати еволюцією параметрів

На наборі даних зі справжніми фотографіями було виконано навчання з використанням еволюції параметрів для двох найкращих YOLO моделей з кожної архітектури. У наведеній нижче таблиці можна помітити, що середні значення кожної метрики зросло. У подальшому було вирішено використовувати модель YOLOv8s з отриманим набором конфігурацій.

<i>Модель</i>	<i>Точність</i>	<i>Повнота</i>	<i>Результативний час (ms)</i>	<i>Точність після GA</i>	<i>Повнота після GA</i>	<i>Результативний час (ms) після GA</i>
YOLOv5s	94.4	93.3	2.7	95.2	94.8	3.8
YOLOv8s	98.3	70.6	5.4	98.8	81.3	6.1

Таблиця 3.3 – Середні значення метрик при стандартному наборі конфігурацій та після виконання генетичного алгоритму

Використання генетичного алгоритму (GA) для пошуку найкращих гіперпараметрів – це ефективний метод, але його недоліком є потреба у потужних обчислювальних потужностях та величезній кількості процесо-годин для

отримання якісних результатів. Виконання еволюційного алгоритму для кожної з YOLO моделей виконувалося більше трьох днів, адже було виконано понад 300 еволюцій по 300-ста епохам.

3.3 Вплив на результати еволюцією параметрів

Для набору даних із зображеннями симуляцій було застосовано плиткування. Початково в наборі даних існувало 1326 зображень, кількість котрих зросла у 9 разів після плиткування, адже було вирішено поділити зображення 3:3. У новоствореному наборі стало 10608 зображень з котрих 9540 були позначені мітками, а усі інші стали фоном. Виконуючи порівняння міток на графіках нижче, можна помітити що велика кількість дронів була посередині зображення, але після виконання плиткування дрони стали більш однорідно розподіленими на зображеннях. Також можна помітити, як масштаб границь дронів став більшим.

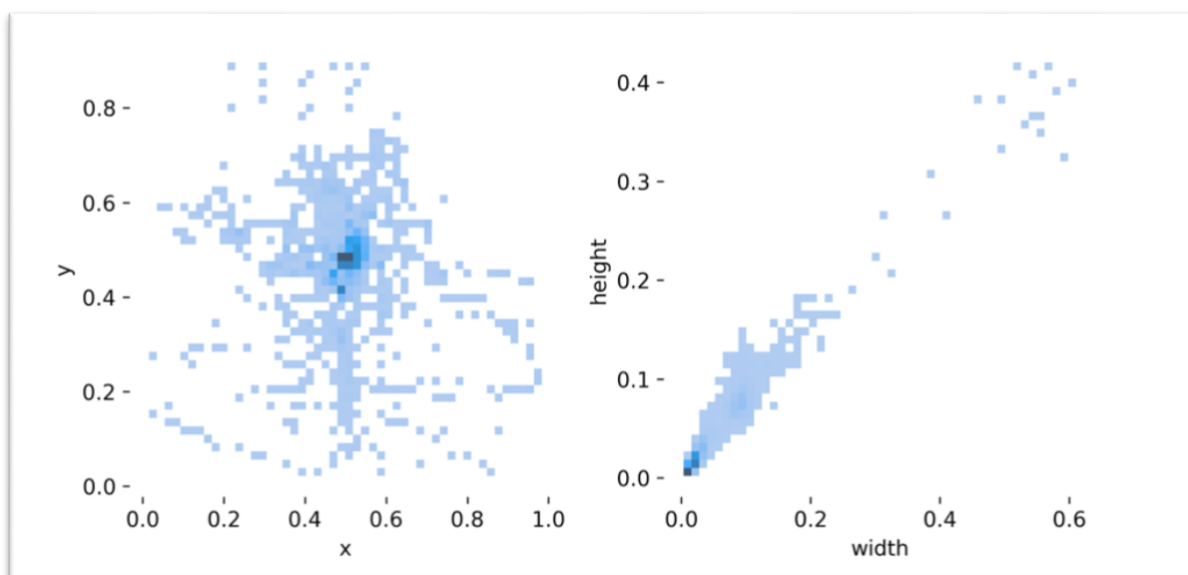


Рисунок 3.7 – розподіл розташування границь із дронами на зображеннях та розподіл розміру зображень відносно усього зображення для початкових зображень

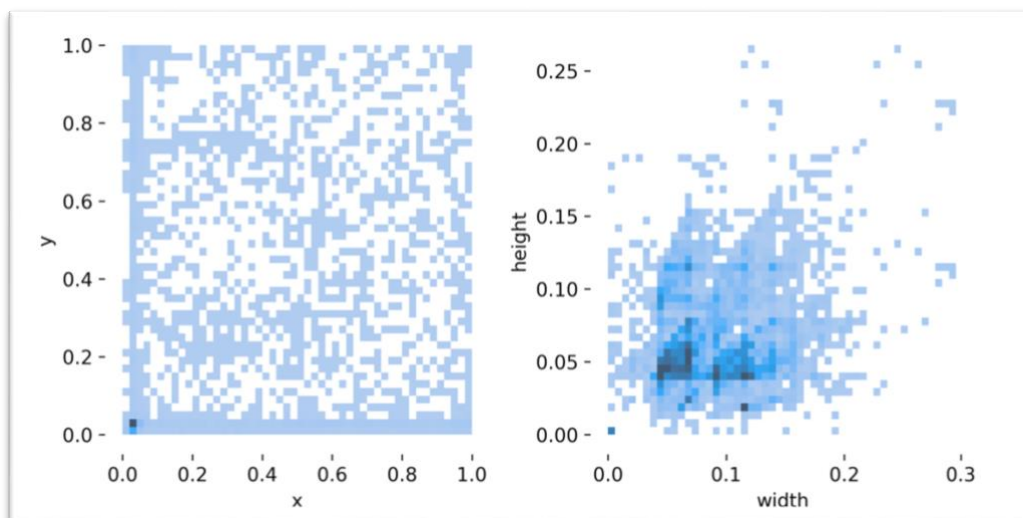


Рисунок 3.8 - розподіл границь із дронами на зображеннях та розподіл розміру зображень відносно усього зображення для поділених на плитки даних

Однак після оцінки середніх значень метрик можна зазначити, що через низьку роздільну здатність фотографій симуляцій точність знизилася в навченій моделі.

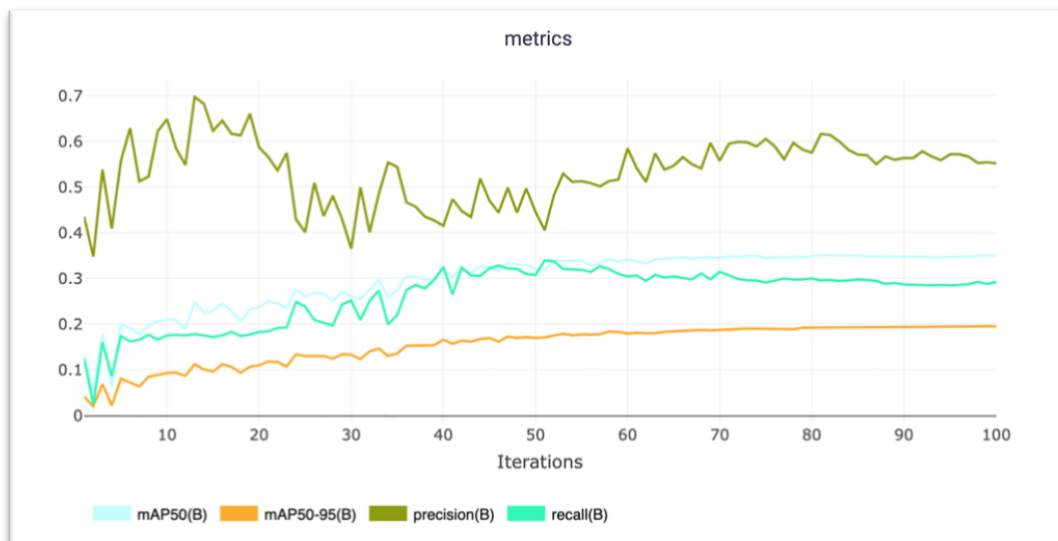


Рисунок 3.9 – метрики під час навчання на поділеному наборі даних з симуляції SimUAV

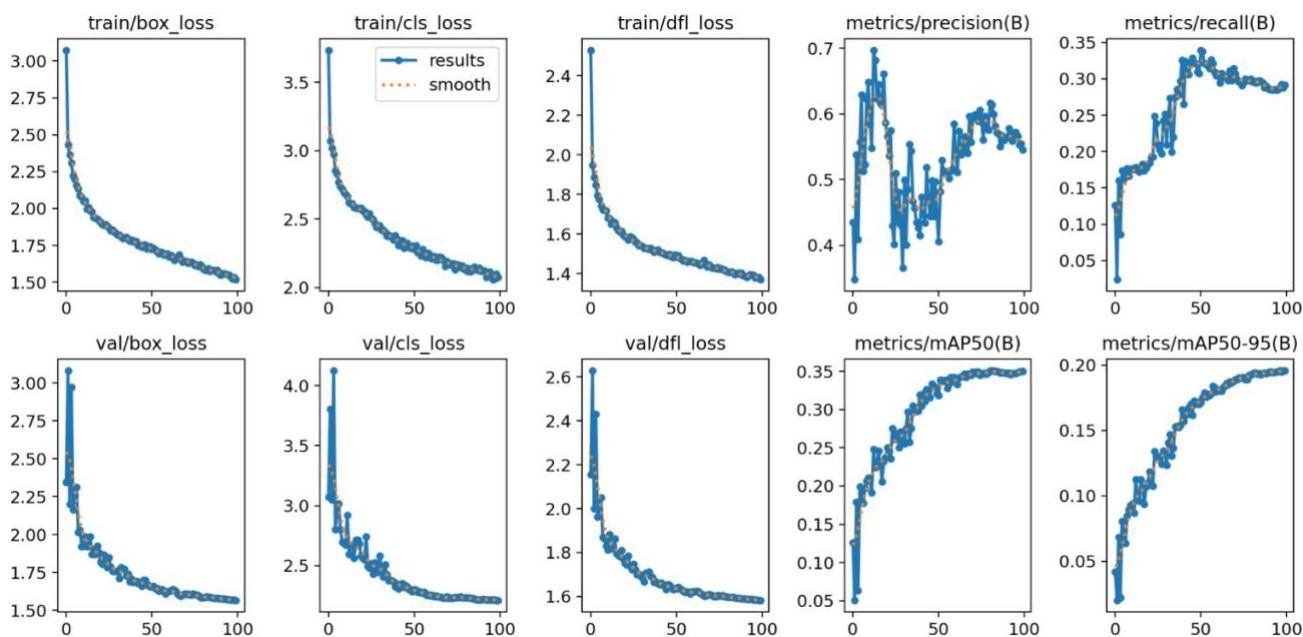


Рисунок 3.9.1 - метрики під час навчання на поділеному наборі даних з симуляції SimUAV

Дані метрики показують потребу у використанні камери з високою роздільною здатністю для фіксування малих безпілотних апаратів. Адже застосувавши плиткування на зображеннях низької якості ми збільшили позитивно негативні знаходження. Одним з аспектів, на який слід звернути увагу, є те, що в більшості випадків фоном для дрона був небосхил. Вилучення інших фонових зображень не дозволяло моделі вивчити, які ознаки не є безпілотникам, що призводило до зниження точності.

Висновки

Проведені дослідження підтверджують високий потенціал глибинного навчання та використання YOLO моделей для знаходження та класифікації безпілотних повітряних апаратів в режимі реального часу, що може мати значний вплив на безпеку та ефективність використання дронів. За результатами експериментів чудові результати показала YOLOv5s та YOLOv8s моделі. У подальшому на їх основі можна побудувати якісну систему виявлення дронів. Для вирішення проблеми визначення маленьких об'єктів на зображенні було застосовано плиточне розбиття. За результатами цього експерименту було додано вимогу до обладнання, що буде відповідати за збір інформації. Адже хоч плиточне розбиття є чудовим рішенням для покращення точності виявлення малих об'єктів, воно негативно впливає на модель у випадку низького розширення зображень. Одним із важливих кроків для покращення існуючої моделі є знаходження найкращих конфігурацій моделі шляхом використання еволюційного алгоритму, але варто зважати на складність та часову складність цього методу. Основними результатами роботи є розробка ефективного алгоритму для виявлення та ідентифікації об'єктів на зображеннях з використанням нейронних мереж, що забезпечує високу точність і швидкість обробки даних.

Зважаючи на отримані результати та проаналізовані джерела можна покращити визначення безпілотних повітряних апаратів додавши акустичний метод виявлення до оптичного з використанням глибинного навчання. У подальшому можна розробити такий пристрій, бо ціна початкового пристрою для детекції не має критично змінитися, адже у найкращій реалізації пристрій фіксації повинен тільки передавати дані, а обчислення будуть виконані на стороні хмари. Важливо зазначити, що навчання моделей були виконані на порівняно невеликих даних і без великої кількості фонових зображень та додаткових класів, що може негативно відобразитися на детекції у випадку потрапляння пташки в поле видимості чи

забрудненості об'єктива. Також варто дослідити можливості передбачення поведінки дронів, їх траєкторій та дій на основі аналізу відеоданих з використанням глибинного навчання. Це може допомогти в розробці більш складних моделей, які здатні передбачати дії дронів та вжити відповідних заходів безпеки.

Список літератури

- 1) Kannadaguli, P. "Система виявлення людей на основі YOLO v4 з використанням аеротермального зображення для безпілотних повітряних апаратів у сфері нагляду." У Протоколи Міжнародної конференції з наукового допоміжного прийняття рішень та застосування (DASA), Сакхір, Бахрейн, 2020, с. 1213–1219.
- 2) Flórez, J., Ortega, J., Betancourt, A., García, A., Bedoya, M., & Botero, J.S. "Огляд алгоритмів, методів та технік виявлення БПЛА та БЛА з використанням аудіо, радіочастоти та відео." *TecnoLógicas* 2020, 23, с. 262–278.
- 3) Bouzayene, I., Mabrouk, K., Gharsallah, A., & Kholodnyak, D. "Сканувальний радар з використанням рівномірного прямокутного масиву для виявлення безпілотних апаратів з низьким КСП." У Протоколи 2019 р. 19-го Середземноморського міжнародного мікрохвильового симпозіуму (MMS), Гамамет, Туніс, 31 жовтня – 2 листопада 2019, с. 1–4.
- 4) R. Jiang, Y. Zhou, та Y. Peng, "Огляд виявлення вторгнення дронів на основі глибинного навчання," у збірнику 2021 року IEEE 4-ої конференції з управління інформацією, комунікаціями, електронікою та автоматизованим контролем (IMSEC), том 4, 2021. doi: 10.1109/IMSEC51613.2021.9482092, с. 1032–1039. [Сторінки 6 і 10.]
- 5) Narayanan, S., & Kuo. Розпізнавання оточуючих звуків за допомогою аудіофункцій час-частоти. Обробка аудіо-, мовленнєвої та мовної інформації, транзакції IEEE.
- 6) Довідка про структури апаратних засобів. Dronecode [Електронний ресурс]. Режим доступу: https://dev.px4.io/master/en/airframes/airframe_reference.html
- 7) Girshick, R. "Швидка R-CNN." У Протоколи Міжнародної конференції з комп'ютерного зору IEEE, 2015, с. 1-7.

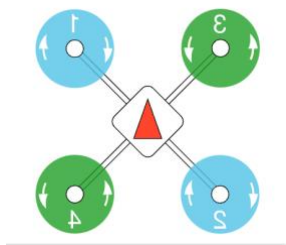
- 8) Guishan Xiang, "Відстеження об'єктів, що швидко рухаються в реальному часі з активною камерою", Обробка зображень та сигналів, 2009, CISP '09, 2-й Міжнародний конгрес.
- 9) Girshick, R., et al. "Багатошарові ієрархії рис для точного виявлення об'єктів та семантичного сегментації." У Протоколи конференції IEEE з комп'ютерного зору та розпізнавання образів, 2014, с. 1-4.
- 10) Siewert, S.B., Andalibi, M., Bruder, S., & Rizor, S. "Мережа оптичного та інфрачервоного електрооптичного сенсора з плавним пошуком для виявлення, відстеження та ідентифікації малих БПЛА." У Протоколи Форуму AIAA Scitech, Сан-Дієго, Каліфорнія, США, 7–11 січня 2019, с. 2264.
- 11) Документація з розмітки зображень Roboflow [Електронний ресурс]. Режим доступу: <https://docs.roboflow.com/image-tagging>
- 12) Документація з Airsim [Електронний ресурс]. Режим доступу: <https://microsoft.github.io/AirSim/>
- 13) Hosang, J., Benenson, R., Dollar, P., & Schiele, B. "Що є ефективними пропозиціями для виявлення?" ' arXiv:1502.05082, 2015, с. 6-9.
- 14) Cui, Y. "Використання глибокого машинного навчання для об'єктно-орієнтованої ідентифікації та виявлення руху на відеоспостереженні з метою забезпечення безпеки." У Протоколи конференції "Симпозіум з міжнародного контролю": Будівництво майбутніх засобів міжнародного контролю, Відень, Австрія, 2018, с. 4-13.
- 15) Документація Python API TensorFlow Core v2.9.1 [Електронний ресурс]. Режим доступу: https://www.tensorflow.org/api_docs/python/tf
- 16) G.J. et. al., "ultralytics/yolov5: v6.0 - Моделі YOLOv5n 'Nano', інтеграція Roboflow, експорт TensorFlow, підтримка OpenCV DNN," Жовтень, 2021. [Електронний ресурс]. Режим доступу: <https://doi.org/10.5281/zenodo.5563715>, с.9.

- 17) Ünel, F., Özkalaycı, B.O., & Çiğla, C. "Сила плитки для виявлення малих об'єктів." У Протоколи конференції IEEE/CVF з комп'ютерного зору та розпізнавання образів (CVPRW), 2019, с. 582–591.
- 18) Isaac-Medina, B.K., Poyser, M., & Organisciak, D. "Візуальне виявлення та відстеження безпілотних повітряних апаратів за допомогою глибоких нейронних мереж: показники ефективності." arXiv 2021, arXiv:2103.13933.

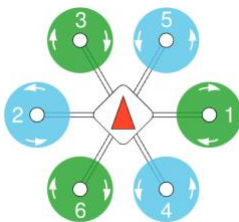
Додаток А

(обов'язковий)

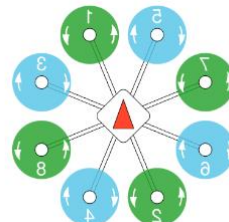
Найпопулярніші типи дронів за архітектурою



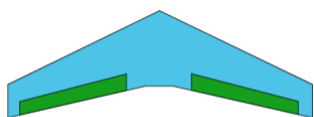
7 Квадрокоптер



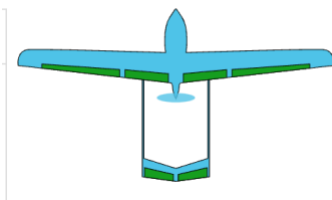
8 Гексакоптер



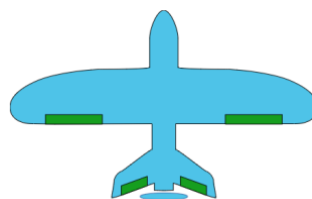
9 Октокоптер



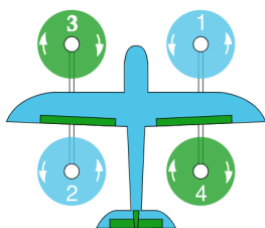
10 Фіксоване крило



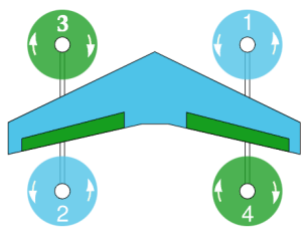
11 А-подібний хвіст



12 V-подібний хвіст



13 Стандартний VTOL



14 VTOL квадро-тейлсіттер



15 Коаксiальний гелiкоптер