

ПОЛТАВСЬКИЙ ДЕРЖАВНИЙ АГРАРНИЙ УНІВЕРСИТЕТ
Навчально-науковий інститут економіки, управління, права та
інформаційних технологій
Кафедра інформаційних систем та технологій

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття ступеня вищої освіти магістр

на тему: **«Повітряний моніторинг сільськогосподарських угідь
на основі Object Detection»**

Виконала: здобувачка вищої освіти
за освітньою програмою
Інформаційні управляючі системи та
технології
спеціальності 126 Інформаційні
системи та технології
ступеня вищої освіти магістр
групи 126ІСТ_мз_2023[1](л. н.)
Кулінченко Ірина Русланівна
Керівник: Слюсар Вадим Іванович
Рецензент: Муравльов Володимир
В'ячеславович

Полтава – 2024 року

ВСТУП

Актуальність теми кваліфікаційної роботи підтверджується необхідністю реалізації концепції Agriculture 4.0, яка передбачає моніторинг сільськогосподарських угідь. Підвищити якість його проведення можливо за рахунок залучення рішень на основі інтеграції дронів і технологій комп'ютерного зору та штучного інтелекту для виконання завдань виявлення об'єктів (Object Detection). Однак, питання розробки рекомендацій щодо використання архітектури YOLO для нейронної мережі Object Detection потребують додаткових досліджень. Все це свідчить про актуальність теми роботи.

Зв'язок роботи з науковими програмами, темами. Робота відповідає дослідженням в межах науково-дослідної ініціативної тематики «Організаційно-методологічні аспекти впровадження інформаційно-комунікаційних систем і технологій в управлінні діяльністю сучасних організацій та підприємств за умов переходу до цифрової економіки» (ДРН 0123U105060, 2023-2028 рр.), що реалізується на кафедрі інформаційних систем та технологій, тематиці досліджень навчально-дослідної лабораторії інтелектуальних систем, комп'ютерних мереж та інтернет речей кафедри інформаційних систем та технологій Полтавського державного аграрного університету.

Метою кваліфікаційної роботи є підвищення ефективності моніторингу сільськогосподарських угідь за рахунок використання технології Object Detection.

Завданнями кваліфікаційної роботи є:

- аналіз варіантів використання дронів під час моніторингу сільськогосподарських угідь;
- аналіз особливостей існуючих підходів щодо виконання процедури Object Detection;
- аналіз архітектури нейронних мереж типу Yolo;

- визначення точності запропонованих моделей нейронних мереж;
- формування рекомендацій щодо використання синтезованих архітектур нейронних мереж.

Об'єктом дослідження є процес виявлення об'єктів за допомогою штучних нейронних мереж.

Предметом дослідження є точність нейронних мереж Object Detection.

Методами дослідження є: аналіз, синтез нейронних мереж Object Detection, вбудовані в фреймворк Keras математичні бібліотеки.

Інформаційна база кваліфікаційної роботи сформована з ресурсів, що містять інформацію про методи сегментації зображень, нейронні мережі та їх компоненти, що використовуються для виконання Object Detection, а також інструментарій для розробки та дослідження згорткових нейронних мереж.

Елементи наукової новизни роботи полягають в розробці архітектури нейронної мережі для реалізації функцій Object Detection, порівняльній оцінці точності нейронних мереж на основі модифікованих архітектур Yolo.

Практична значущість роботи полягає в розробці рекомендацій щодо використання запропонованої архітектури нейронної мережі в інтересах аграрного сектору, які можуть бути використані для подальших досліджень за даною тематикою та при реалізації концепції Agriculture 4.0.

Апробація результатів відбувалася в рамках VII Міжнародної студентської наукової конференції «Наука сьогодення: від досліджень до стратегічних рішень» (листопад 2024 р., м. Кривий Ріг) та VII Міжнародної студентської наукової конференції «Розвиток суспільства та науки в умовах цифрової трансформації» (листопад 2024 р., м. Тернопіль).

Структура кваліфікаційної роботи логічно пов'язана з завданнями досліджень і містить вступ, три розділи основної частини, висновки, список використаних джерел, додатки. Загальний обсяг пояснювальної записки кваліфікаційної роботи складає 79 сторінок формату А4. Вона містить 34 рисунки і 2 таблиці.

РОЗДІЛ 1

АНАЛІЗ ВЛАСТИВОСТЕЙ ПОВІТРЯНОГО МОНІТОРИНГУ СІЛЬСЬКОГОСПОДАРСЬКИХ УГІДЬ

1.1 Особливості реалізації повітряного моніторингу сільськогосподарських угідь

Як відомо [1], повітряний моніторинг у сільському господарстві – це процес збирання, аналізу та інтерпретації даних про стан сільськогосподарських угідь за допомогою засобів повітряного спостереження. Це може бути здійснено за допомогою супутників, літаків дронів (Unmanned Aerial Vehicle, UAV), або інших платформ, оснащених спеціалізованими сенсорами та камерами. До появи UAV моніторинг проводився традиційним методом об'їзду та обходу своїх полів. Ці методи були вкрай витратними, трудомісткими і часто необ'єктивними. Все це призводило до того, що агроном не міг об'єктивно визначити проблемну ділянку. Надалі це позначалося на неправильному рішенні за нормою мінеральних добрив, пестицидів і насіння.

Основною метою цього процесу є отримання детальної інформації про стан рослинності, ґрунту та кліматичних умов, що дозволяє приймати рішення про догляд за культурами, оптимізації ресурсів і підвищення врожайності, а саме:

- оцінювати стан культур на великих площах у короткі терміни;
- виявляти проблемні ділянки (посуху, ерозію ґрунту, шкідників) задовго до того, як вони стануть візуально помітними;
- оптимізувати використання добрив, води та засобів захисту рослин, зменшуючи їхнє нераціональне використання та ін.

Застосування повітряного моніторингу стало одним із ключових інструментів у рамках концепції точного землеробства. Сьогодні, коли аграрний сектор стикається з викликами, такими як зміна клімату, деградація

ґрунтів, обмеженість водних ресурсів і зростаюче населення, інноваційні технології стають критично важливими.

Основні компоненти повітряного моніторингу можна умовно поділити на три категорії:

- програмне забезпечення – системи для збору, обробки та візуалізації даних (наприклад, GIS-системи або платформи аналізу зображень);
- камери (включають мультиспектральні, гіперспектральні камери, тепловізори) та сенсори (в тому числі радарні або лідарні сенсори);
- мережі передачі даних, що забезпечують швидке отримання та обробку інформації у реальному часі.

Ефективний аналіз даних, зібраних під час повітряного моніторингу, можливий завдяки використанню сучасного програмного забезпечення (ПЗ). Ці програми дозволяють перетворювати необроблені дані в корисну інформацію, яка використовується для прийняття рішень у сільському господарстві. До прикладів такого ПЗ можна віднести Pix4D, DroneDeploy, ArcGIS.

Pix4D (рис. 1.1) – це спеціалізоване ПЗ для обробки аерофотознімків і створення ортофотопланів, теплових карт і 3D-моделей територій [2]. Дозволяє аналізувати площі, об'єми, рельєф та висоти культур при вимірюванні параметрів полів. Для оцінки здоров'я рослин ведеться розрахунок індексів вегетації на основі NDVI. При цьому підтримується багато моделей UAV та сенсорів.

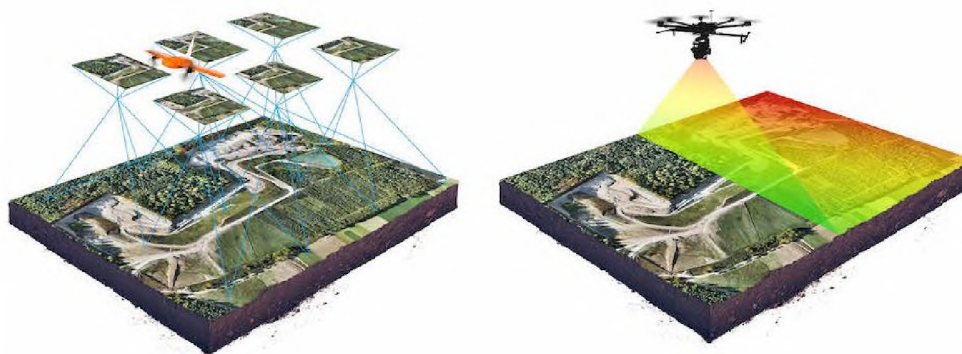


Рисунок 1.1 – ПЗ Pix4D

В цілому, це ПЗ виконує локальну обробку даних без підключення до Інтернет та має інтуїтивний інтерфейс і можливість налаштування звітів під потреби користувача. В якості недоліків можна вказати високу вартість ліцензії, та необхідність потужного обладнання для обробки великих обсягів даних.

DroneDeploy [3] (рис. 1.2) – хмарна платформа для збору, обробки й аналізу даних із дронів. Вона забезпечує завантаження аерофотознімків у платформу для автоматизованого створення карт, також розраховує індекси вегетації (NDVI, VARI, GNDVI тощо) та надає можливість отримати попередні результати прямо в полі. ПЗ інтегрується з іншими інструментами, наприклад, John Deere Operations Center.

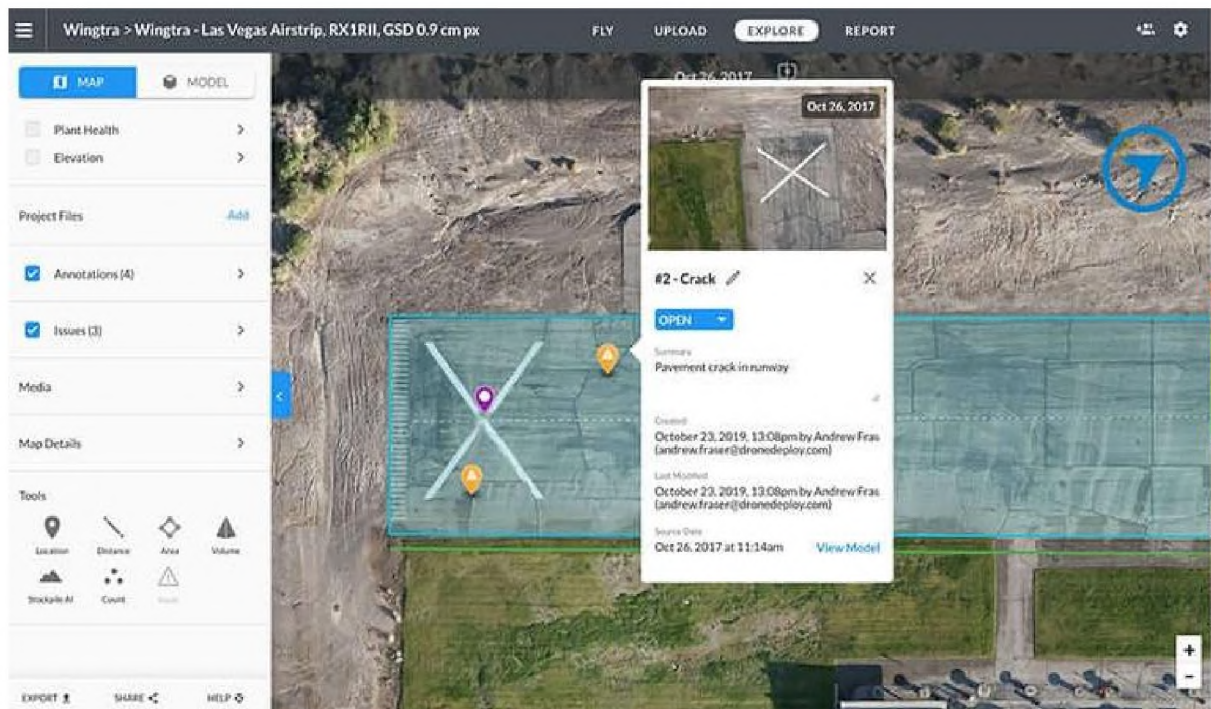


Рисунок 1.2 – ПЗ DroneDeploy

До переваг варто віднести можливість спільного доступу до проєктів між командами, простоту у використанні навіть для новачків, нема потреби у потужному локальному обладнанні. Недоліками є залежність від підключення до Інтернет, ліцензія оплачується за підпискою, що може бути витратним для малих господарств.

ArcGIS [4] (рис. 1.3) – універсальна платформа для аналізу геопросторових даних, яка широко використовується в різних галузях, включно із сільським господарством. Вона має аналіз рельєфу, гідрологічні моделі, прогнозування зміни врожайності, а також можливість об'єднувати дані з дронів, супутників і наземних датчиків та підтримує Python для створення скриптів автоматизації процесів ПЗ створює тематичні карти, геостатистичний аналіз та прогнозування.

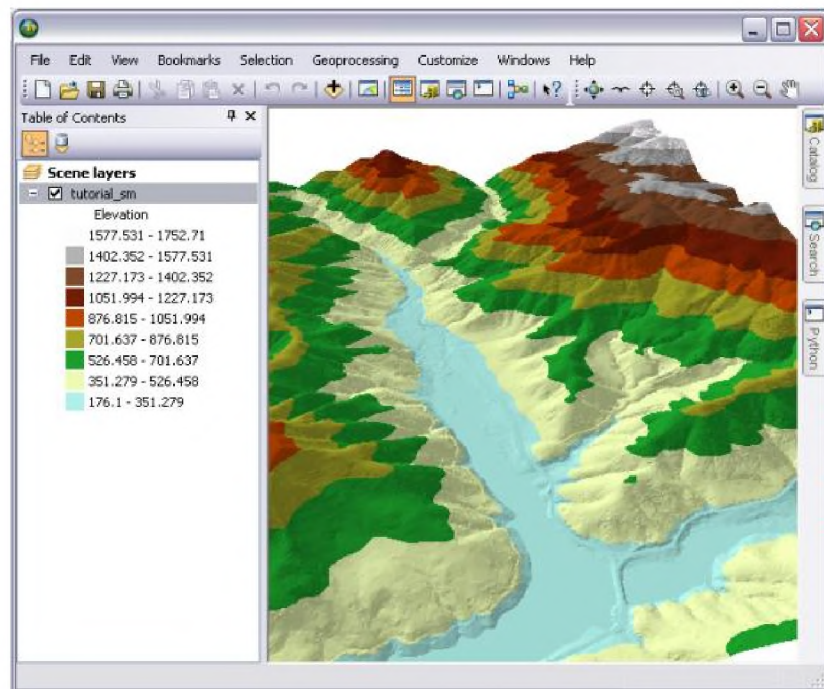


Рисунок 1.3 – ПЗ ArcGIS

ПЗ виступає як потужний інструмент для великих і складних проєктів, що підтримує великі масиви даних і їх детальний аналіз. Має високий рівень адаптивності до потреб користувача. Недоліками є висока вартість ліцензії та крута крива навчання для освоєння всіх функцій.

Таким чином, вибір ПЗ залежить від потреб користувача, наприклад Для малих господарств – підійде DroneDeploy через простоту та оперативність; Для великих проєктів – ArcGIS завдяки гнучким інструментам аналізу; Для спеціалізованих завдань – Pix4D стане ідеальним рішенням завдяки високій точності обробки.

Макроекономічні виклики сучасності змушують переходити сільськогосподарських землекористувачів на цифрову модель ведення бізнесу – точне землеробство. Одним із основних елементів точного землеробства є мультиспектральний моніторинг сільськогосподарських угідь, створення карток неоднорідностей поля з підбором вегетаційних індексів, а також створення карток-приписів для диференційованого внесення мінеральних добрив та пестицидів. За допомогою мультиспектрального зондування полів (рис. 1.4) можна отримувати інформацію про стан культур, ступінь зволоженості певних ділянок поля, загрозу шкідників, бур'янів, а також прогнозувати врожайність та приймати адекватні та своєчасні рішення.

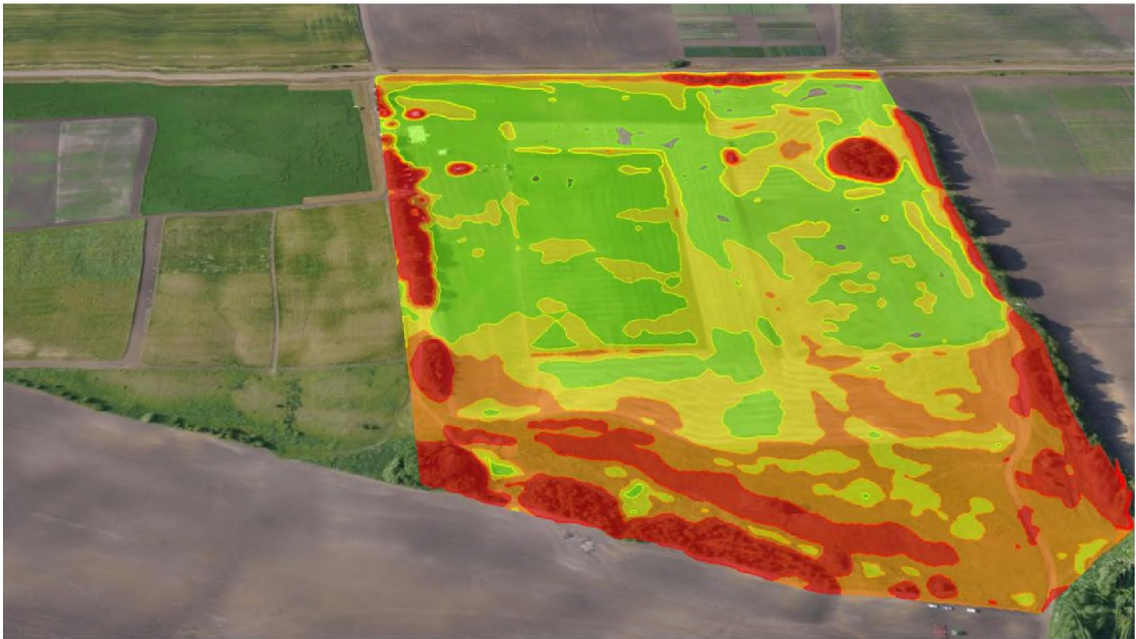


Рисунок 1.4 – Мультиспектральний моніторинг сільськогосподарських угідь

Щоб виявити найбільш родючі ділянки, області з підвищеним рівнем вмісту азоту, відстежувати розвиток посівів, обчислити вегетаційні індекси та індекси листової поверхні необхідно, щоб UAV підтримував встановлення різних видів камер. Обробка даних мультиспектральної зйомки дозволяє отримати карти вегетаційних індексів (NDVI, NDRE, VARI, LCI та ін.), якими можна оцінити продуктивність сільськогосподарських угідь і спрогнозувати врожайність.

На даний час, існують три основні методи повітряного моніторингу: Дистанційне зондування із супутників, літаки та пілотовані платформи, а також UAV (дрони). Перший метод забезпечує глобальне покриття та дозволяє проводити регулярний моніторинг великих територій. Дані з супутників можуть мати низьку роздільну здатність, але вони зручні для аналізу загальної картини. Цей метод більш технологічний, але має низьку недоліків. За рахунок великої висоти супутників деталізація знімків невелика. Кадри високого розрізнення коштують дорого, тому найчастіше аграріям доводиться купувати цифрову картку не лише своєї ділянки, а й сусідніх. Над певним полем супутник може пролітати раз на 5 днів, що не дає можливості оперативно дізнатися про стан своїх угідь у потрібний момент часу. Імовірність хмарності зменшує шанси на отримання якісних даних ще в кілька разів. Другий метод використовуються для середніх за масштабом територій. Він менш поширені через високу вартість. Третій метод підходить для локального моніторингу. Аерофотозйомка з дрона завжди чіткіша і детальніша, ніж космічний знімок. Як правило, робоча висота дронів від 100 до 500 метрів. Ортофотоплани полів, отримані за допомогою дрона, виходять детальними та з високою роздільною здатністю. Також зйомка з коптерів можлива навіть за умов низької хмарності. А завдяки сучасним системам передачі відеосигналу оператор може отримувати знімки та відео з дронавої камери в режимі реального часу. Карти розпоряджень, отримані за рахунок мультиспектральної зйомки з дрону, допомагають аграріям заощадити до 25 % за рахунок диференційованого внесення азотних добрив та пестицидів. Для прикладу, варто розглянути одне з безпілотних рішень для мультиспектрального зондування полів від компанії DJI – P4 Multispectral (рис. 1.5) [5]. Це високоточна модель дрону, оснащена мультиспектральною системою обробки зображень. Він має 5 мультиспектральних камер та одну камеру в діапазоні RGB. Можливість відзняти 100 га за 20 хв. на висоті польоту 500 м заощадить ще й час (табл. 1.1). Отримані з його допомогою дані дозволяють фермерам та агрономам оперативно отримувати

максимально точні дані про стан експлуатованих угідь та сільськогосподарських культур, що вирощуються.



Рисунок 1.5 – P4 Multispectral

Таблиця 1.1 – Продуктивність DJI P4 Multispectral на площі 100 га при перекритті фотографій (%): поздовжнє – 70, поперечне – 80

Висота (метр)	Час (хв.)	Кількість акумуляторів (шт.)
500	20	2
450	22	2
400	24	2
350	26	2
300	28	2
250	35	2
200	43	3
150	57	4

На основі проведених досліджень можна визначити основні цілі впровадження технологій повітряного моніторингу в сільському господарстві:

- визначення рівня здоров'я рослин, вмісту хлорофілу та індексів продуктивності (наприклад, NDVI [6]);
- допомагає оптимізувати полив і знижувати витрати на воду;
- раннє визначення інфікованих ділянок дозволяє вчасно застосувати заходи для захисту посівів;

- фіксація змін рельєфу, утворення ярів чи засоленості ґрунту;
- збір даних для аналізу динаміки росту культур та прогнозування врожаю;
- аналіз впливу внесення добрив або пестицидів на стан рослинності.

В цілому, повітряний моніторинг сприяє переходу до сталого сільського господарства. Висока деталізація даних дозволяє виявляти проблеми на ранніх стадіях. Своєчасне реагування на проблеми сприяє збільшенню продуктивності. Швидке охоплення великих територій відбувається без значних витрат часу та ресурсів. Завдяки сучасним технологіям, таким як дрони та супутникові системи, фермери отримують доступ до детальної інформації про стан полів, що дозволяє приймати обґрунтовані рішення та впроваджувати принципи точного землеробства. Підвищення врожайності та точності управління та зменшення витрат на ресурси (вода, добрива, пестициди) можливе за рахунок використання штучного інтелекту (Artificial Intelligence, AI) та машинного навчання (Machine Learning, ML) для точного прогнозування. Надалі доцільно розглянути варіанти застосування UAV в інтересах проведення повітряного моніторингу сільськогосподарських угідь.

1.2 Аналіз напрямів використання дронів

На основі проведених досліджень можна виділити кілька прикладних аспектів залучення дронів до повітряного моніторингу.

Виявлення та класифікація рослин. На великих сільськогосподарських полях людям дуже важко виявити такі рослини, як трава та бур'яни. Якщо справжні посіви та бур'яни не будуть виявлені та класифіковані, це може вплинути на прибуток, а також на навколишнє середовище через надмірне використання добрив, пестицидів та інсектицидів. Багато рослин є сировиною для медичної промисловості, і ідентифікація цих рослин є виснажливою роботою для людини. Багато рідкісних рослин потребують

збереження та ідентифікації. Тобто можна використовувати UAV для збору інформації, а використання ML повинно бути дуже простим, економічним і менш трудомістким. Пояснення класифікації рослин за допомогою UAV наведено на рис. 1.6. UAV роблять знімки, потім ці зображення проходять через алгоритми класифікації рослин.

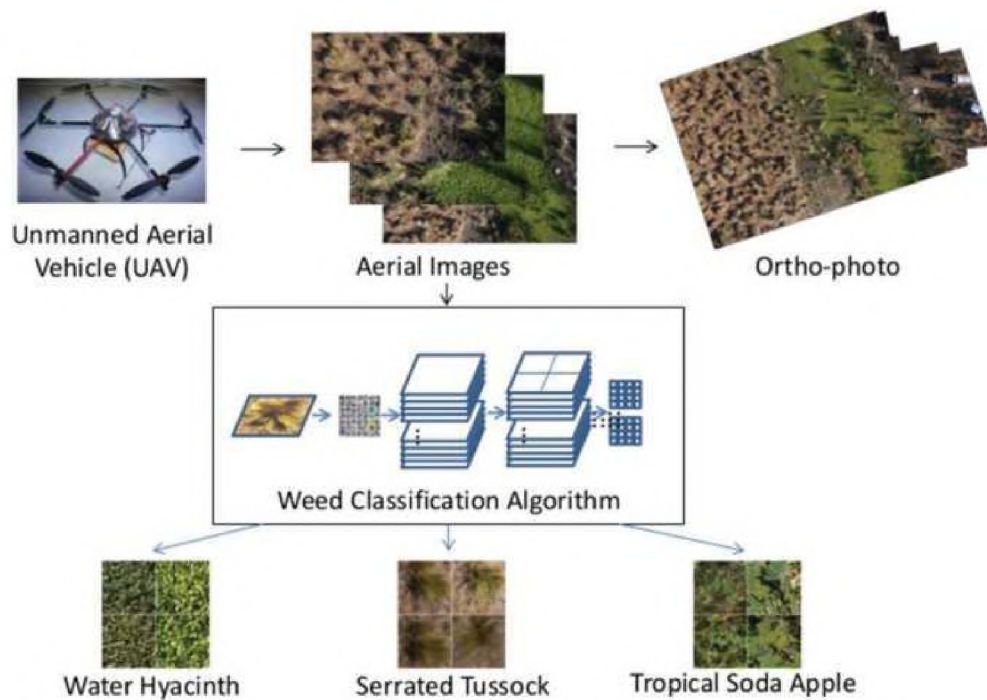


Рисунок 1.6 – Виявлення рослин за допомогою зображень UAV

Отриманні з UAV зображення використовуються для вивчення рослинності для ідентифікації трави на луках. В основі лежить платформа з подвійною камерою на основі UAV [7]. Використовуючи класифікацію к-середніх, проводиться оцінка кількості квіток ріпаку. В [8] запропонований алгоритм на основі методів навчання для ідентифікації бур'янів на полях за допомогою зображень UAV. Автори стверджують, що досягли 90 % точності. В [9] наведена інтегрована конвеєрна методологія з використанням дерев рішень із посиленням градієнта для картографування трав і рослинності в посушливих землях із точністю близько 96 %. У [10] запропонована архітектура CNN для класифікації рослин із послідовностей зображень з точністю 97,47 % для класифікації 16 видів рослин. У [11] CNN

продемонструвала середню точність 86,2 % при класифікації 22 види рослин, включаючи бур'яни та зернові культури.

Оцінка здоров'я рослин. Оскільки технологія продовжує вдосконалюватися, вона також потребує впровадження систем автоматизації на сільськогосподарських полях. Щоб розвивати високотехнологічні сільськогосподарські системи, потрібно зосередитися на здоров'ї рослин, де можна виділити кілька напрямів:

- моніторинг здоров'я посівів;
- виявлення хвороб і пошкоджень культур;
- сортування та аналіз форми плодів сільськогосподарської продукції.

Моніторинг стану рослин. Робота [12] представляє використання ML для розробки методології, яка аналізує цифрові зображення цитрусових рослин, зібраних з UAV. Наземні датчики, включаючи вимірювач водного потенціалу, хлорофілометр і спектрорадіометр, використовуються для моніторингу стану ферми. Тут за допомогою цього методу застосували техніку точного землеробства. Зібрані зображення та основні правдиві дані потім використовуються як навчальні дані для ML. Це дає численні результати щодо використання комерційних датчиків RGB для надання фермерам важливої інформації про стан полів. Головним мотивом їх досліджень є зменшення внесення добрив і підвищення врожайності. Автори досягли 82 % продуктивності моделі.

Виявлення хвороб і пошкоджень культур. В якості прикладу слід розглянути роботу [13], в якій описано експеримент з посівами пшениці на двох полях одне було заражене гелмінтоспоріозом (HLB), а інший – нормальним. HLB-інфекція призвела до жовтизни листя на початковій стадії. UAV (Phantom 4) з бортовою камерою RGB використовувався для зйомки зображень 4000×3000 пікселів з висоти 80 метрів нейронна мережа CNN досягла точності 91,43 %, а похибка становила 0,83 %, перевершуючи інші за рівнем точності та стабільності. В [14] висвячені результати експериментів з октокоптером DJI S1000 і 5-діапазонною мультиспектральною камерою.

Систему випробували на озимій пшениці, зараженій інокулятом жовтої іржі. UAV робив знімки на відстані від 16 до 24 метрів із роздільною здатністю від 1 до 1,5 см/піксель. Розроблена система досягла середньої точності 89,2 %, запам'ятовування 89,4 % і точності 89,3 %. В [15] збирали гіперспектральні дані за допомогою DJI Matrice 600Pro Hexacopter і гіперспектральної камери. Вони використовували ARI і TCARI 1, які точно виявляли рослини з раковою інфекцією. В [16] описаний експеримент з двома UAV, оснащеними датчиками RGB (12 МП). Вони зробили 40 зображень розміром 4000×3000 пікселів. Використовуючи GoogLeNet, вони виявили фузаріоз з точністю понад 90 %. Згідно [17], запроваджена система виявлення хвороб рослин за допомогою системи обробки зображень використовує алгоритм RELIEF-F для виявлення різних захворювань. На рис. 1.7 показано кроки для виявлення хвороб рослин і класифікації хвороб.

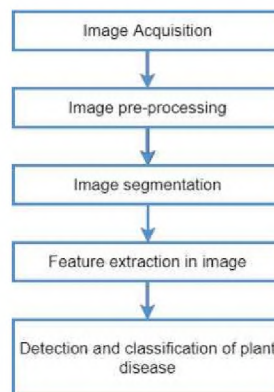


Рисунок 1.7 – Основні етапи виявлення та класифікації хвороб рослин [17]

У [18] наведена модель CNN для виявлення та діагностики захворювань рослин, використовуючи зображення листя здорових і хворих рослин, використовуючи методології глибокого навчання. На основі відкритої бази даних, яка містить 87848 зображень 25 різних рослин, наборів із 58 різних класів комбінацій (рослин, хвороб). При цьому нейронна мережа досягла 99,53 % успіху. В [19] наведена техніка для ідентифікації та класифікації грибкових захворювань. Вивчались чотири різні сільськогосподарські/садівничі культури: овочеві культури, фруктові

культури, зернові культури та комерційні культури. Пропонована робота спрямована на розробку методології класифікації та ідентифікації симптомів грибкових захворювань садових і сільськогосподарських культур.

Розумна боротьба зі шкідниками та травами. Шкідники – головна проблема сільському господарстві. Ефективне використання пестицидів необхідне як доквілля, так сільськогосподарських культур. Надмірне використання пестицидів також є неекономічним. Перевага розпилення пестицидів за допомогою UAV полягає у прискоренні процесу та запобіганні надмірному поглинанню пестицидів ґрунтом. Однак через швидкість і напрям вітру велика кількість пестицидів може переноситися на прилеглі поля. З підвищеною точністю можна обмежити втрати пестицидів під час розпилення зверху. Наприклад, в [20] запропонована архітектура, яка використовує UAV з системою поєднаних розпилювачів, які можуть спілкуватися з бездротовою мережею датчиків. Ця система має на меті надсилати відгуки про погодні умови та рівень обприскування на полі цільової культури. На основі інформації UAV відповідним чином застосовує оптимізований маршрут. В [21] автори провели порівняльне дослідження серед архітектур глибокого навчання VGG-19, ResNet-50, Inception-v3, VGG-16 і Xception із 5000 зображеннями шкідників сої та виявили, що архітектура ResNet-50 досягла точності до 93,82 %. У [22] запропонований мобільний додаток під назвою BioLeaf, який може вимірювати пошкодження листя сої. Додаток було протестовано на реальних зображеннях, зібраних із соєвих полів, і інструмент можна використовувати для культур різної ширини та вузького листя. В [23] проведено дослідження раннього моніторингу злакових бур'янів на полях кукурудзи з використанням зображень UAV та алгоритмів OBIA (рис. 1.8).

Аналіз поля та оцінка врожайності сільськогосподарських культур є важливим завданням в управлінні та маркетингу товарів. Точне прогнозування врожайності допомагає фермерам покращити якість їхнього врожаю. Це також допомагає зменшити експлуатаційні витрати, приймаючи

кращі рішення щодо інтенсивності збору врожаю та необхідної праці. Загалом оцінки врожайності здійснюються з використанням попередніх даних, а працівники вручну підраховують плоди у вибраних місцях поля. Існує два типи комп'ютерних методів оцінки продуктивності: регіональні методи та методи підрахунку.



Рисунок 1.8 – Приклад отриманого UAV зображення поля соняшника з аналізом зображення, яке відокремлює бур'яни від соняшників

В [24] розглянута автоматична система для оцінки врожайності яблук, у якій зображення яблуневих садів фіксуються вночі для зменшення непередбачуваного природного освітлення вдень. В [25] використано кольорові зображення, щоб підрахувати кількість яблук у садах при природному освітленні, модель може виявити яблука з точністю понад 85 %. В [26] наведена методика оцінки врожайності насіння ріпаку за допомогою даних дистанційного зондування методом спектрального аналізу суміші. Оцінка врожайності є важливою задачею в сільському господарстві; точніша оцінка приносить більше прибутку. У майбутньому UAV повинні автоматично й точніше оцінювати ресурси та врожай у спосіб, який буде менш трудомістким і більш прибутковим. Можна виділити основні домени: виявлення та класифікація рослин, оцінка здоров'я рослин, інтелектуальна боротьба зі шкідниками та травами, а також аналіз поля та оцінка врожайності.

1.3 Object Detection у завданнях повітряного моніторингу

Дослідивши різноманітність додатків, з якими AI може допомогти у сільськогосподарській галузі, стає зрозуміло, що інтеграція комп'ютерного зору (Computer Vision, CV) на сільське господарство дає як економічні, і екологічні переваги, змінюючи традиційні методи з допомогою інноваційних технологій [27]. CV забезпечує стабільні, надійні результати, знижуючи кількість людських помилок у таких завданнях, як виявлення фруктів, виявлення шкідників та оцінка стану врожаю. Автоматизовані системи знижують потребу в ручній праці, що призводить до економії коштів при виконанні таких трудомістких завдань, як збирання врожаю та боротьба зі шкідниками. CV здатне обробляти великі обсяги даних, CV підтримує масштабні сільськогосподарські операції, забезпечуючи розуміння величезних полів за мінімального втручання людини. Сприяючи цілеспрямованому ефективному розподілу ресурсів, CV підтримує стійкі методи господарювання, що відповідають екологічним цілям.

Виявлення об'єктів (Object Detection) має фундаментальне значення для повітряного моніторингу, забезпечуючи точність, необхідну для виявлення, класифікації та аналізу об'єктів у різних умовах ферми де важливі швидкість та точність [28]. Використовуючи Object Detection для повітряного моніторингу, аграрії отримують потужний інструмент для прийняття обґрунтованих рішень, що сприяє сталому розвитку сільського господарства. Відповідно, Object Detection – це технологія CV, яка дозволяє автоматично виявляти та класифікувати об'єкти на зображеннях або відео, визначаючи їх місцезнаходження та тип. Це здійснюється за допомогою алгоритмів ML та AI, які аналізують візуальні дані для ідентифікації об'єктів, таких як транспортні засоби, люди, тварини, рослини тощо.

Таким чином, Object Detection для повітряного моніторингу сільськогосподарських угідь можна використовувати за кількома напрямки. За допомогою дронів або супутникових зображень можна виявляти хвороби,

шкідників або дефіцит поживних речовин у рослинах. Object Detection дозволяє автоматично ідентифікувати проблемні ділянки для швидкого реагування. Автоматичне розпізнавання небажаних рослин або комах допомагає оптимізувати використання добрив, зменшуючи витрати та екологічний вплив. Аналізуючи густоту та стан посівів, можна прогнозувати потенційну врожайність та планувати збір урожаю. Виявлення зон з недостатнім або надмірним зволоженням допомагає оптимізувати системи зрошення, зберігаючи воду та підвищуючи ефективність вирощування. На пасовищах об'єктне розпізнавання може використовуватися для моніторингу худоби, виявлення втеч або потенційних загроз від хижаків. Виявлення ознак ерозії або інших ґрунтових проблем дозволяє вжити заходів для збереження родючості земель. Створення детальних карт угідь з точним розташуванням різних культур та об'єктів інфраструктури.

Object Detection можна реалізувати за допомогою алгоритму You Only Look Once (YOLO). YOLO – це одна з популярних та ефективних мереж для Object Detection та класифікації на зображеннях (відео) у реальному часі. На відміну від інших методів, які розбивають зображення на частини або використовують кілька мереж, YOLO обробляє все зображення за один раз. Це забезпечує високу швидкість обробки. YOLO передбачає координати обмежувальних рамок (Bounding Boxes) та ймовірності класів об'єктів безпосередньо з пікселів зображення. YOLO можна навчити розпізнавати ознаки хвороб або наявності шкідників на рослинах, аналізуючи зображення з дронів. Модель може розрізняти різні типи рослин, допомагаючи виявляти бур'яни серед культурних насаджень. На полях та пасовищах YOLO може використовуватися для відстеження руху худоби або виявлення диких тварин, розпізнавання ознак ерозії, сухості або інших проблем з ґрунтом.

На даний час, актуальною є версія YOLO11. За допомогою таких завдань, як сегментація екземплярів, YOLO11 може розрізняти плоди, листя, шкідників та багато іншого, що робить його універсальним інструментом для

різних застосувань – від посадки рослин до післязбиральних процесів. Під час навчання на наборах даних, специфічних для сільського господарства, YOLO11 може розпізнавати тонкі деталі, такі як індикатори шкідників або різні стадії росту фруктів. Наприклад, передові моделі виявлення об'єктів, включаючи YOLO11, можна навчити відрізняти плоди від листя та гілок навіть у густому листі (рис. 1.9). Це може зробити оцінку врожайності більш точною та надійною, дозволяючи краще управляти ресурсами та прогнозувати прибуток.



Рисунок 1.9 – YOLO для виявлення та оцінки врожайності зелених яблук.

YOLO11 може підтримувати точний збір урожаю, визначаючи та сегментуючи стиглі фрукти в режимі реального часу. Для таких цінних культур, як яблука, ця технологія дозволяє звести до мінімуму використання ручної праці та зменшити пошкодження, пов'язані з обробкою, що зрештою підвищує якість зібраного врожаю. Щоб почати роботу з YOLO треба зібрати та анотувати зображення, які відповідають завданню (наприклад, фотографії полів з різними типами рослин або проблемами). Потім треба використовувати фреймворки для навчання моделі YOLO на власних даних. Після навчання модель можна інтегрувати в систему моніторингу, яка аналізуватиме нові зображення та надаватиме результати в реальному часі.

Таким чином, YOLO є потужним інструментом для реалізації Object Detection в сфері сільського господарства. Ця архітектура дозволяє швидко та ефективно аналізувати великі обсяги візуальних даних, допомагаючи аграріям приймати обґрунтовані рішення та підвищувати ефективність господарювання. Вона підходить для задач, де потрібна висока швидкість, наприклад, обробка відео в реальному часі. Сучасні версії YOLO (наприклад, YOLOv8, YOLO11) забезпечують високу точність Object Detection. Її можна навчити розпізнавати практично будь-які об'єкти, за умови наявності відповідних даних для навчання.

Висновки до розділу 1

Повітряний моніторинг сільськогосподарських угідь відіграє важливу роль у впровадженні концепції точного землеробства. Застосування дронів, супутників та інших платформ, оснащених сенсорами, дозволяє аграріям отримувати детальну та об'єктивну інформацію про стан рослинності, ґрунту та кліматичних умов. Це значно полегшує прийняття рішень щодо догляду за культурами, оптимізації використання ресурсів та підвищення врожайності. Технології повітряного моніторингу дозволяють оцінювати стан культур на великих площах у короткі терміни, виявляти проблемні ділянки на ранніх етапах. Висока деталізація даних сприяє своєчасному реагуванню на проблеми, що дозволяє зменшити витрати та зберегти навколишнє середовище. Застосування сучасних програмних засобів робить процес моніторингу ефективним і гнучким, залежно від потреб користувача.

Впровадження дронів у сільське господарство відкриває широкі можливості для підвищення ефективності ведення аграрного бізнесу. Зокрема, UAV успішно використовуються для моніторингу сільськогосподарських угідь, ідентифікації рослин, оцінки їх здоров'я, а

також для боротьби зі шкідниками та травами. Одним із ключових напрямів застосування дронів є виявлення та класифікація рослин, що допомагає оптимізувати використання пестицидів і зменшити негативний вплив на навколишнє середовище. За допомогою UAV можна вчасно виявити бур'яни, рідкісні або корисні рослини, що сприяє підвищенню якості управління ресурсами.

Моніторинг здоров'я рослин із використанням машинного навчання та обробки зображень дозволяє оперативно виявляти хвороби та пошкодження культур, підвищуючи точність застосування захисних заходів. Це, в свою чергу, зменшує надмірне використання хімічних засобів, підвищуючи рентабельність виробництва та мінімізуючи вплив на екологію. Також дрони відіграють важливу роль у боротьбі зі шкідниками та травами, дозволяючи ефективніше застосовувати пестициди, зменшуючи втрати при їх розпиленні та оптимізуючи використання. Аналіз поля та оцінка врожайності є ще одним важливим напрямом застосування дронів, що сприяє підвищенню ефективності управління ресурсами та прогнозуванню врожайності.

Інтеграція технологій CV у сільське господарство забезпечує суттєві економічні та екологічні переваги, трансформуючи традиційні методи управління аграрними процесами. Застосування технологій Object Detection, зокрема на базі алгоритмів YOLO, дозволяє автоматично ідентифікувати та класифікувати об'єкти на зображеннях, що сприяє значному підвищенню ефективності господарювання. Object Detection за допомогою дронів та супутникових зображень дозволяє вчасно виявляти проблеми, такі як шкідники, хвороби рослин або дефіцит поживних речовин. Це сприяє оптимізації використання добрив. Технології CV також допомагають аналізувати стан посівів, прогнозувати врожайність та планувати збір урожаю, що дозволяє аграріям приймати обґрунтовані рішення.

Застосування сучасних версій YOLO, таких як YOLO11, забезпечує високу точність і швидкість Object Detection у реальному часі. Це особливо важливо для таких завдань, як автоматизований збір урожаю та моніторинг

стану угідь. Використання дронів для аналізу густоти посівів, моніторингу худоби, виявлення ознак ерозії та інших проблем дозволяє ефективно розподіляти ресурси, зберігати воду та підтримувати екологічно стійкі методи господарювання.

РОЗДІЛ 2

СИНТЕЗ АРХІТЕКТУРИ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ РЕАЛІЗАЦІЇ ФУНКЦІЙ ОБ'ЄКТ DETECTION

2.1 Визначення компонентів архітектури YOLO

CV, галузь, яка швидко розвивається, дозволяє машинам інтерпретувати та розуміти візуальні дані [29]. Вирішальним аспектом цього домену є виявлення об'єктів [30], яке передбачає точну ідентифікацію та локалізацію об'єктів у зображеннях або відеопотоках [31]. Останні роки стали свідками значного прогресу в алгоритмічних підходах до вирішення цієї проблеми [32].

Ключовий прорив у виявленні об'єктів стався з впровадженням алгоритму YOLO Редмоном та ін. у 2015 р. [33]. Цей інноваційний підхід, як впливає з назви, обробляє все зображення за один прохід, щоб виявити об'єкти та їх розташування. Методологія YOLO відрізняється від традиційних двоетапних процесів виявлення, розглядаючи виявлення об'єктів як проблему регресії [33]. У ньому використовується єдина CNN для одночасного прогнозування обмежувальних рамок і ймовірностей класів по всьому зображенню [34], оптимізуючи конвеєр виявлення порівняно з складними традиційними методами.

YOLO11 є останньою ітерацією в серії YOLO та являє собою значний крок вперед у технології Object Detection у реальному часі. У цій новій версії суттєво вдосконалено як архітектуру, так і методику навчання, розсуваючи межі точності, швидкості та ефективності. Інноваційний дизайн YOLO11 включає в себе вдосконалені методи вилучення функцій, що дозволяє охоплювати більш тонкі деталі, зберігаючи мінімальну кількість параметрів. Це призводить до підвищення точності в різноманітних завдань CV, від виявлення об'єктів до класифікації. Крім того, YOLO11 досягає значного приросту швидкості обробки, суттєво підвищуючи продуктивність у

реальному часі. Табл. 2.1 ілюструє розвиток моделей YOLO від їх створення до останніх версій. Кожна ітерація принесла значні покращення в можливостях виявлення об'єктів, обчислювальній ефективності та універсальності в обробці різноманітних завдань CV.

Таблиця 2.1 – Еволюція моделей YOLO

Реліз	Рік	Завдання	Компоненти
YOLO [33]	2015	Object Detection, Basic Classification	Single-stage object detector
YOLOv2 [35]	2016	Object Detection, Improved Classification	Multi-scale training, dimension clustering
YOLOv3 [36]	2018	Object Detection, Multi-scale Detection	SPP block, Darknet-53 backbone
YOLOv4 [37]	2020	Object Detection, Basic Object Tracking	Mish activation, CSPDarknet-53 backbone
YOLOv5 [38]	2020	Object Detection, Basic Instance Segmentation (via custom modifications)	Anchor-free detection, SWISH activation, PANet
YOLOv6 [39]	2022	Object Detection, Instance Segmentation	Self-attention, anchor-free OD
YOLOv7 [40]	2022	Object Detection, Object Tracking, Instance Segmentation	Transformers, E-ELAN reparameterisation
YOLOv8 [41]	2023	Object Detection, Instance Segmentation, Panoptic Segmentation, Keypoint Estimation	GANs, anchor-free detection
YOLOv9 [42]	2024	Object Detection, Instance Segmentation	PGI and GELAN
YOLOv10 [43]	2024	Object Detection	Consistent dual assignments for NMS-free training

YOLO11 [44] вирізняється своєю розширеною адаптивністю, підтримкою розширеного діапазону завдань CV, крім традиційного виявлення об'єктів. Примітними серед них є оцінка положення та сегментація екземплярів, що розширює застосовність моделі в різних областях. Конструкція YOLO11 зосереджена на балансі потужності та практичності, спрямованому на вирішення конкретних завдань у різних галузях з підвищеною точністю та ефективністю. Ця остання модель демонструє постійну еволюцію технології. Універсальність і покращення продуктивності позиціонують YOLO11 як значний прогрес у галузі, потенційно відкриваючи

нові шляхи для реального впровадження в різних секторах. За своєю суттю архітектура YOLO складається з трьох основних компонентів. По-перше, магістраль (Backbone) служить основним екстрактором ознак, використовуючи CNN для перетворення необроблених даних зображення в багатомасштабні карти функцій. По-друге, компонент шиї (Neck) діє як проміжний етап обробки, використовуючи спеціалізовані шари для агрегування та покращення репрезентації ознак у різних масштабах. По-третє, голова (Head) – компонент функціонує як механізм прогнозування, генеруючи остаточні результати для локалізації та класифікації об'єктів на основі вдосконалених карт ознак. Спираючись на цю архітектуру, YOLO11 розширює та вдосконалює основу, закладену в YOLOv8, запроваджуючи архітектурні інновації та оптимізуючи параметри для досягнення чудової продуктивності Object Detection (рис. 2.1).

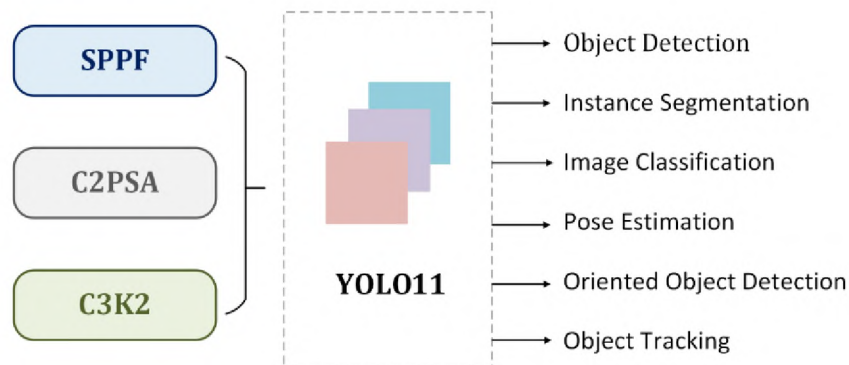


Рисунок 2.1 – Ключові архітектурні модулі в YOLO11

Магістраль є ключовим компонентом архітектури YOLO, відповідальним за вилучення функцій із вхідного зображення в різних масштабах. Цей процес включає укладання згорткових шарів і спеціалізованих блоків для створення карт об'єктів з різною роздільною здатністю. YOLO11 зберігає структуру, подібну до своїх попередників, використовуючи початкові згорткові шари для зменшення дискретизації зображення. Ці шари формують основу процесу виділення ознак, поступово зменшуючи просторові розміри та збільшуючи кількість каналів. Значним

покращенням у YOLO11 є введення блоку C3k2, який замінює блок C2f, що використовувався в попередніх версіях [45]. Блок C3k2 є більш ефективною з обчислювальної точки зору реалізацією вузького місця Cross Stage Partial (CSP). Він використовує дві менші згортки замість однієї великої згортки, як видно в YOLOv8 [41]. «k2» у C3k2 вказує на менший розмір ядра, що сприяє швидшій обробці при збереженні продуктивності. YOLO11 зберігає блок Spatial Pyramid Pooling Fast (SPPF) з попередніх версій, але вводить новий блок Cross Stage Partial with Spatial Attention (C2PSA) після нього [45]. Блок C2PSA є помітним доповненням, яке покращує просторову увагу на картах функцій.

Цей механізм просторової уваги дозволяє моделі ефективніше фокусуватися на важливих областях зображення. Завдяки просторовому об'єднанню функцій блок C2PSA дозволяє YOLO11 зосередитися на конкретних областях інтересу, потенційно покращуючи точність виявлення об'єктів різного розміру та положення. Neck поєднує риси в різних масштабах і передає їх Head для прогнозування. Цей процес, зазвичай, передбачає підвищення дискретизації та конкатенацію карт функцій з різних рівнів, що дозволяє моделі ефективно отримувати багатомасштабну інформацію.

Розглянемо магістраль. Блок C3k2 розроблений як швидший і ефективніший, підвищуючи загальну продуктивність процесу агрегування функцій. Після підвищення дискретизації та конкатенації шия в YOLO11 включає цей покращений блок, що призводить до покращеної швидкості та продуктивності [45]. Помітним доповненням до YOLO11 є посилення уваги до просторової уваги через модуль C2PSA. Цей механізм уваги дозволяє моделі зосередитися на ключових регіонах зображення, потенційно сприяючи точнішому виявленню, особливо менших або частково закритих об'єктів. Включення C2PSA відрізняє YOLO11 від його попередника YOLOv8, якому бракує цього специфічного механізму уваги [45].

Розглянемо голову YOLO11, яка відповідає за генерацію остаточних прогнозів щодо виявлення та класифікації об'єктів. Вона обробляє карти

функцій, передані з шиї, зрештою виводячи обмежувальні рамки та мітки класів для об'єктів на зображенні. Head використовує кілька блоків C3k2 для ефективної обробки та вдосконалення карт функцій. Блоки C3k2 розміщені в кількох шляхах у голові, функціонуючи для обробки багатомасштабних елементів на різних глибинах. Блок C3k2 проявляє гнучкість в залежності від значення параметра C3k:

- коли $c3k = \text{False}$, модуль C3k2 поводить себе подібно до блоку C2f, використовуючи стандартну структуру вузького місця;

- коли $c3k = \text{True}$, структура вузького місця замінюється модулем C3, який дозволяє глибше та складніше витягувати ознаки.

C3k2 має такі основні характеристики:

- використання двох менших згорток зменшує накладні витрати на обчислення порівняно з однією великою згорткою, що призводить до швидшого вилучення ознак;

- C3k2 є більш компактною версією вузького місця CSP, що робить архітектуру більш ефективною з точки зору кількості параметрів, які можна навчити.

Іншим помітним доповненням є блок C3k, який забезпечує підвищену гнучкість, дозволяючи налаштовувати розміри ядра. Адаптивність C3k особливо корисна для отримання більш детальних характеристик із зображень, що сприяє підвищенню точності виявлення.

Head YOLO11 включає кілька шарів Convolution-BatchNorm-Silu (CBS) [46] після блоків C3k2. Блоки CBS служать основоположними компонентами як для вилучення ознак, так і для процесу виявлення, забезпечуючи передачу уточнених карт функцій до наступних рівнів для обмежувальної рамки та прогнозування класифікації. Ці шари додатково вдосконалюють карти функцій за допомогою властивостей:

- вилучення відповідних функцій для точного виявлення об'єктів;

- стабілізація та нормалізація потоку даних за допомогою пакетної нормалізації;

– використання функції активації Sigmoid Linear Unit (SiLU) для нелінійності, що покращує продуктивність моделі.

Тепер розглянемо кінцеві згорткові шари та рівень виявлення. Кожна гілка виявлення закінчується набором шарів Conv2D, які зменшують функції до необхідної кількості виходів для координат обмежувальної рамки та передбачень класу. Останній рівень виявлення об'єднує ці передбачення, які включають:

- координати обмежувальної рамки для локалізації об'єктів на зображенні;
- оцінки об'єктності, які вказують на наявність об'єктів;
- бали класу для визначення класу виявленого об'єкта.

YOLO11 підтримує широкий спектр завдань CV, демонструючи свою універсальність і потужність у різних програмах. Вона відмінно справляється з ідентифікацією та локалізацією об'єктів у зображеннях або відеокадрах, надаючи обмежувальні рамки для кожного виявленого елемента [47]. Ця здатність знаходить застосування в системах відеоспостереження, автономних транспортних засобах і роздрібній аналітиці, де точна ідентифікація об'єктів має вирішальне значення [48]. Виходячи за рамки простого виявлення, YOLO11 може ідентифікувати та відокремлювати окремі об'єкти на зображенні аж до піксельного рівня [47]. Ця сегментація особливо цінна в медичній візуалізації для точного окреслення органу чи пухлини, а також у виробництві для детального виявлення дефектів [48]. YOLO11 здатний класифікувати цілі зображення за заздалегідь визначеними категоріями, що робить його оптимальним для програм, таких як категоризація продуктів на платформах електронної комерції або моніторинг дикої природи в екологічних дослідженнях [48]. Модель може виявляти певні ключові точки в зображеннях або відеокадрах, щоб відстежувати рухи або пози. Ця можливість є корисною для додатків відстеження фізичної активності, аналізу спортивних результатів і різноманітних програм охорони здоров'я, які вимагають оцінки руху [48]. YOLO11 представляє можливість виявляти

об'єкти за кутом орієнтації, що дозволяє точніше локалізувати об'єкти, що обертаються. Ця функція особливо цінна в аналізі аерофотознімків, робототехніці та завданнях автоматизації складів, де об'єктна орієнтація має вирішальне значення [48]. YOLO11 ідентифікує та відстежує шлях об'єктів у послідовності зображень або відеокадрів [48]. Ця можливість відстеження в реальному часі необхідна для таких додатків, як моніторинг руху, спортивний аналіз і системи безпеки.

YOLO11 представляє діапазон розмірів моделей, від нано до надзвичайно великих, що задовольняють різноманітні потреби застосування. Ця масштабованість дозволяє розгортати в різних сценаріях, починаючи від крайових пристроїв з обмеженими ресурсами до високопродуктивних обчислювальних середовищ. Варіант Nano, зокрема, демонструє вражаючі покращення швидкості та ефективності порівняно зі своїм попередником. Модель містить нові архітектурні елементи, які покращують її можливості вилучення функцій і обробки. Включення нових елементів, таких як блок C3k2, SPPF і C2PSA, сприяє більш ефективному вилученню та обробці функцій. Ці вдосконалення дозволяють моделі краще аналізувати та інтерпретувати складну візуальну інформацію, потенційно сприяючи підвищенню точності виявлення в різних сценаріях. Універсальність YOLO11 виходить за рамки виявлення об'єктів, охоплюючи такі завдання, як сегментація екземплярів, класифікація зображень, оцінка пози та орієнтоване виявлення об'єктів. Цей багатогранний підхід позиціонує YOLO11 як комплексне рішення для різноманітних завдань CV. Ключовим прогресом у YOLO11 є інтеграція складних механізмів просторової уваги, зокрема компонента c2PSA. Ця функція дозволяє моделі ефективніше зосереджуватися на критичних областях зображення, підвищуючи її здатність виявляти й аналізувати об'єкти. Покращена здатність концентрувати увагу особливо корисна для ідентифікації складних або частково закритих об'єктів, вирішуючи загальну проблему в задачах виявлення об'єктів. Це удосконалення просторового сприйняття сприяє загальному покращенню

продуктивності YOLO11, особливо в складних візуальних середовищах. Порівняльний аналіз показує чудову продуктивність YOLO11, особливо в його менших варіантах. Модель nano, незважаючи на незначне збільшення параметрів, демонструє підвищену швидкість виведення та кількість кадрів в секунду (FPS) порівняно з попередником. Це покращення свідчить про те, що YOLO11 досягає сприятливого балансу між обчислювальною ефективністю та точністю виявлення.

Прогрес у YOLO11 має значні наслідки для різних галузей. Його підвищена ефективність і багатозадачність роблять його особливо придатним для застосування в автономних транспортних засобах, системах спостереження та промисловій автоматизації. Здатність моделі добре працювати в різних масштабах також відкриває нові можливості для розгортання в середовищах з обмеженими ресурсами без шкоди для продуктивності.

2.2 Виявлення невеликих об'єктів при повітряному моніторингу

В останні роки розвиток дронів був швидким і значним. У цивільному секторі потенційне застосування дронів поширюється на численні сфери, включаючи транспорт, сільське господарство, логістику, безпеку, міське будівництво, попередження про катастрофи тощо. Більшість аерофотознімків містять значну кількість невеликих об'єктів. Набір даних Common Objects in Context (COCO) [49] – це набір даних, який зазвичай використовується для виявлення об'єктів, у якому визначення розміру об'єкта полягає в тому, що об'єкти з розмірами менше 32×32 пікселів класифікуються як малі об'єкти, об'єкти з розмірами між 32×32 пікселів і 96×96 пікселів позначаються як середні об'єкти, а об'єкти з розмірами більше 96×96 пікселів позначаються як великі об'єкти. Як ілюстрація, набір даних VisDrone [50] демонструє майже дві третини малих об'єктів, як показано на рис. 2.2.

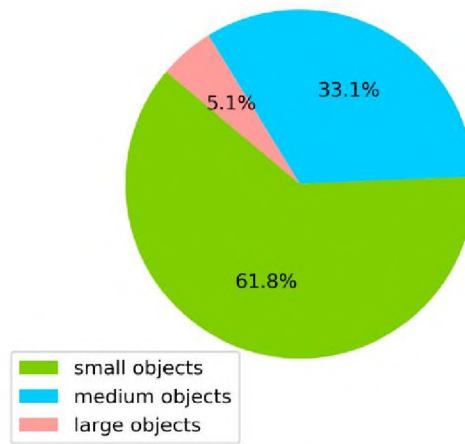


Рисунок 2.2 – Співвідношення великих, середніх і малих об’єктів у наборі даних VisDrone

Отже, у вищезазначених областях швидка та точна ідентифікація малих об’єктів може помітно підвищити ефективність роботи. Тому велике значення мають алгоритми виявлення малих об’єктів на аерофотознімках. В свою чергу, зображення, що отримані за допомогою UAV, мають відмінні характеристики, такі як різні масштаби цілі, складний фон, сильна оклюзія, маленькі цілі та щільний розподіл. У результаті виявлення об’єктів на аерофотознімках стикається з проблемами, такими як труднощі з вилученням інформації про невеликий об’єкт і погана інтеграція просторових і семантичних даних. Крім того, існуючі алгоритми Object Detection мають велику кількість параметрів, що створює проблему для розгортання на UAV з обмеженими апаратними ресурсами. Крім того, алгоритм Object Detection для аерофотознімків з часом буде розгорнуто на термінальних пристроях дронів, які переважно обмежені апаратним забезпеченням і не мають надійного графічного процесора (GPU). Деякі існуючі складні моделі глибокого навчання, хоча і мають відносно високу точність виявлення в теорії, не можуть працювати в режимі реального часу на дронах через надмірні параметри та велику кількість обчислень, що обмежує їхню цінність у практичних застосуваннях UAV. Отже, алгоритм повинен враховувати власну складність, щоб переконатися, що він не займає надмірно обмежену доступну обчислювальну потужність. Крім того, він повинен досягти ефективного

балансу між точністю та швидкістю. Як наслідок, в роботі [51] запропонований варіант YOLO, структура якої розроблена для вирішення проблеми розпізнавання щільних малих об'єктів на зображеннях (Efficient Small Object Detection YOLO, ESOD-YOLO) – рис. 2.3. Модуль RepNIBMS пропонується як заміна модулю C2f мережі вилучення функцій.

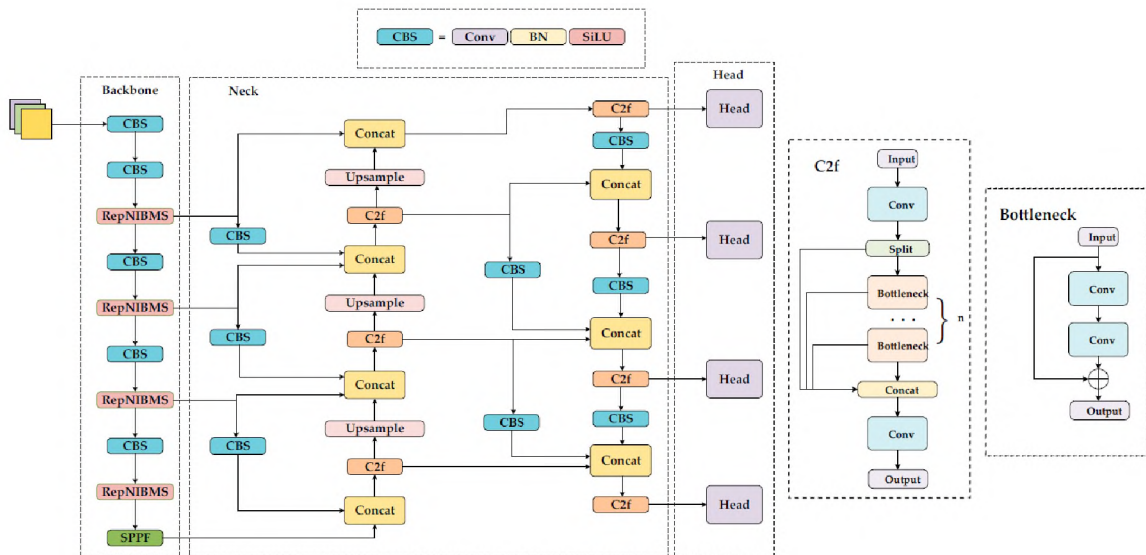
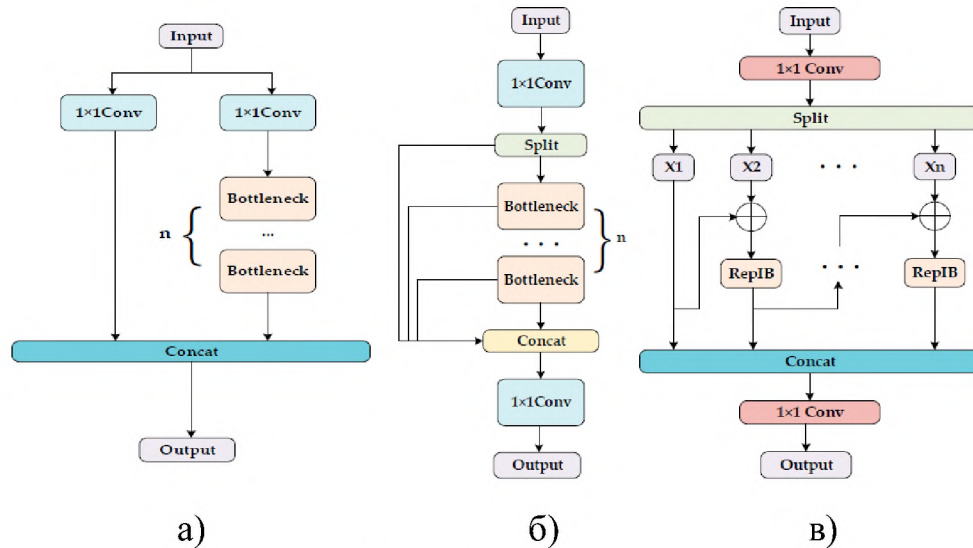


Рисунок 2.3 – Структура ESOD-YOLO

Його мета полягає в тому, щоб підвищити здатність моделі обробляти об'єкти різного масштабу та об'єднувати характеристики різних рівнів в ієрархічний спосіб. Частина Neck конструкції було модифіковано для включення багатомасштабної структури злиття функцій на кількох шарах. Ця інтеграція семантичної та просторової інформації з карт об'єктів різних масштабів на поточному, верхньому та нижньому рівнях має на меті покращити здатність виражати ознаки моделі після об'єднання. Масштаб цілі на аерофотознімку неоднорідний. Тобто, розміри різних об'єктів демонструють значну варіабельність. Крім того, масштаб об'єкта на зображенні може змінюватися залежно від кута зйомки та відстані, навіть якщо сам об'єкт залишається незмінним. Тому очевидно, що обробка багатомасштабних ознак відіграє невід'ємну роль у навчанні та виявленні моделей. Спосіб, у який обробляються багатомасштабні характеристики, як з

точки зору простоти, так і ефективності, має глибокий вплив на загальну ефективність і точність алгоритму. На рис. 2.4.а показано модуль вилучення функцій CSP, який використовується YOLOv5.



Умовні позначки:

- а) – CSP;
- б) – C2f;
- в) – RepNIBMS.

Рисунок 2.4 – Структури блоків YOLO

Основна мета дизайну цієї структури полягає в тому, щоб мінімізувати обчислювальне навантаження шляхом розділення карт функцій на різних етапах і вирішення проблеми надлишкової інформації про градієнти. Отже, це підвищує як точність, так і швидкість моделі. Однак ця структура має обмежену гнучкість і не оптимально використовує інформацію про градієнтний потік. Таким чином, YOLO пропонує модуль C2f, як показано на рис. 2.4.б. Модуль спрощує процес об'єднання функцій і зменшує кількість зайвих обчислень. Для поділу та злиття каналів використовується більш ефективний метод, який зменшує обчислювальну складність і покращує загальну ефективність моделі. Крім того, структура оптимізує шлях градієнтного потоку, щоб забезпечити потік градієнтів у глибокій мережі, тим

самим підвищуючи стабільність навчання. Хоча цей модуль із серії YOLO демонструє похвальну продуктивність у загальних завданнях виявлення об'єктів, він демонструє певні обмеження при застосуванні до конкретних контекстів, таких як фотознімки, зроблені дронами. Зокрема, проблеми, пов'язані з виявленням малих об'єктів на фотознімках, складною фоновою обробкою та обмеженнями обчислювальних ресурсів вимагають більш спеціалізованих рішень. Однак більшість існуючих алгоритмів YOLO просто виконують злиття ознак у структурі ший, щоб отримати вичерпну семантичну інформацію, не враховуючи при цьому значення мереж виділення ознак у побудові багатомасштабних представлень ознак. Таким чином, необхідно вдосконалити модуль, щоб гарантувати його здатність вивчати більш детальну інформацію. Паралельно розроблено модуль RepNIBMS (рис. 2.4.в) шляхом злиття концепцій ієрархічної мережі агрегації DLA [28] і Res2Net [29], що забезпечує агрегацію ієрархічних залишків. Ця структура має здатність збільшити кількість масштабів, присутніх у вихідних функціях, полегшити вивчення більш виразних багатомасштабних функцій і підтримувати швидку швидкість обчислень. Зокрема, модуль RepNIBMS спочатку розбиває вхідні дані на кілька підмножин, кожна з яких обробляється модулем RepIB незалежно (рис. 2.5).

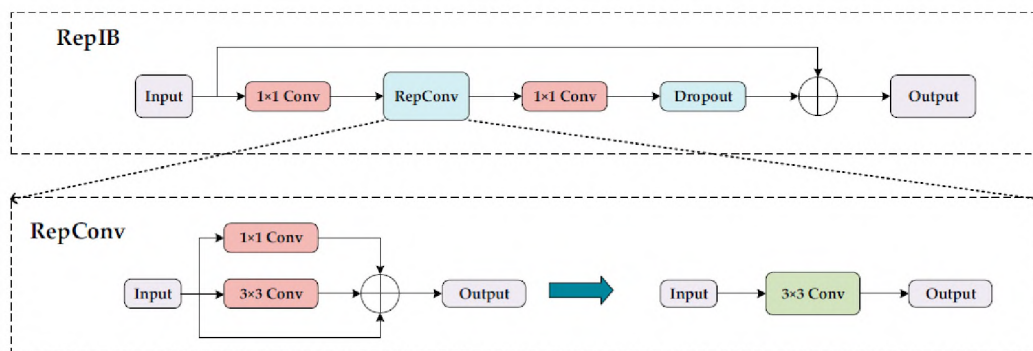


Рисунок 2.5 – Структура модуля RepIB

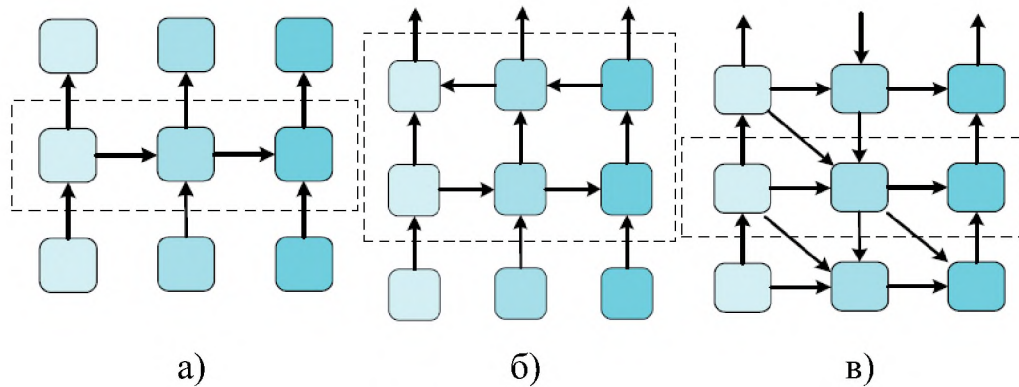
Модуль RepIB дає змогу агрегувати залишки на різних рівнях, таким чином дозволяючи моделі отримати більш багате представлення функцій і

досягти глибокої агрегації ознак разом з ефективною обробкою. Ця методологія ефективно фіксує інформацію про особливості в різних масштабах, таким чином покращуючи вивчення більш виразних багатомасштабних функцій. Одночасно RepNIBMS забезпечує високу швидкість обчислень, що робить його придатним для програм реального часу. Модуль репараметризованого інвертованого блоку (RepIB) використовується відповідно до вищезгаданої формули для полегшення міжшарового з'єднання, таким чином забезпечуючи більш плавний потік градієнтів у глибокій мережі.

Нарешті, усі ієрархічні елементи з'єднані, а обмін інформацією здійснюється через згортку 1×1 , що покращує здатність моделі виявляти об'єкти в складних сценах. Згодом мережа зменшується на згортковий рівень і проходить через рівень вилучення нейрони відкидаються, що призводить до іншої структури мережі на кожному етапі навчання. Це запобігає переобладнанню мережею навчальних даних, тим самим покращуючи здатність до узагальнення та надійність моделі. Коефіцієнт відсіву встановлено на 0,1.

Модуль WFPN. Інформація про особливості кількох масштабів витягується в мережі вилучення ознак. Згодом вилучені функції потребують подальшої обробки та об'єднання. Низькорівнева карта функцій генерується на початковому рівні мережі та містить значну кількість просторових деталей та інформації про межі, а також більшість вихідної інформації, присутньої у вхідному зображенні. Карта функцій високого рівня генерується на глибокому рівні мережі. Він містить більше семантичної інформації з меншою кількістю просторових деталей і здатний отримувати глобальну семантичну інформацію. Комбінація цих карт ознак різних масштабів і рівнів призводить до генерації більш тонких і розрізняючих ознак. У серії алгоритмів YOLO використовуються структури пірамідної мережі (FPN) [52] і мережі агрегації шляхів (PAN) [53] – рис. 2.6. Ці структури дозволяють мережі використовувати функції різного масштабу, тим самим полегшуючи отримання контекстної інформації та підвищуючи здатність алгоритму

виявляти як малі, так і великі об'єкти. Однак шлях потоку контекстної інформації в цій структурі є фіксованим і недостатньо гнучким. Отже, пропонується міжрівнева багатомасштабна структура злиття ознак, WFPN.



Умовні позначки:

а) – FPN;

б) – PAN;

в) – WFPN.

Рисунок 2.6 – Структуру блоків YOLO

У структурі WFPN інформація про особливості поширюється аналогічним чином. Замість того, щоб текти в одному напрямку, він пливе через кілька масштабів одночасно; кожна карта функцій функціонує як хвиля, яка отримує інформацію про вертикальний «стрибок» від попереднього шару, одночасно «зіткнувшись» і «зливаючись» з горизонтально з'єднаними однолітками на тому самому рівні. Ця структура демонструє чудову продуктивність в обробці багатомасштабних функцій. Включення додаткових функцій масштабу та більш складних стратегій синтезу дозволяє точніше обробляти об'єкти в більш широкому діапазоні масштабів, таким чином підвищуючи здатність моделі виявляти об'єкти різного розміру.

У будові частини Neck має бути повністю поєднана ознакова інформація різних масштабів. Однак механізм синтезу Generalized-FPN (GFPN) [54] призведе до збільшення кількості параметрів і обчислень. Збільшення кількості об'єднань і обробки складних функцій неминуче

приведе до збільшення споживання ресурсів. Відповідно, конфігурація WFPN одночасно встановлює горизонтальні зв'язки між картами функцій на кожному рівні та вертикальними зв'язками з попереднього рівня, а також приймає вхідні дані з попереднього рівня горизонтальних зв'язків (рис. 2.7).

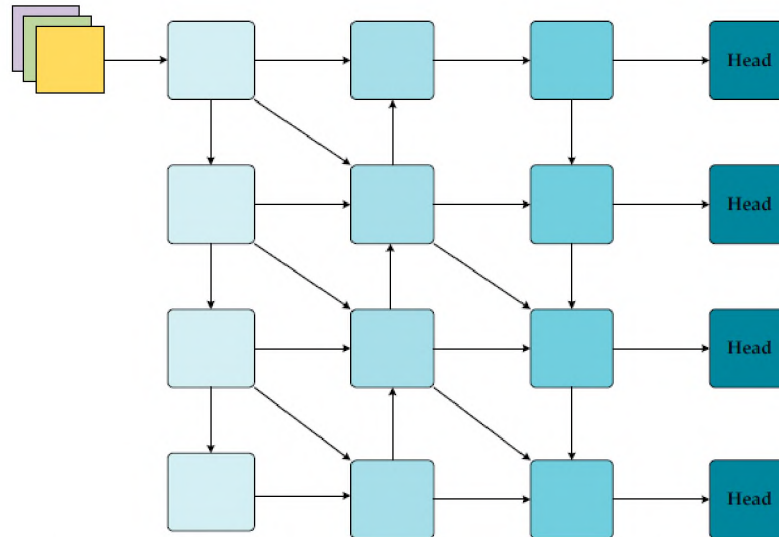


Рисунок 2.7 – Використання структури WFPN у моделі мережі

Ця структура не обмежена фіксованим шляхом потоку контекстної інформації, а впровадження багатошляхових з'єднань дозволяє вільно встановлювати зв'язки між картами об'єктів різного масштабу. Кожна карта функцій може отримати користь від інформації в різних масштабах. Карти ознак низького рівня (що містять детальну інформацію) і карти ознак високого рівня (що містять семантичну інформацію) доповнюють одна одну, тим самим покращуючи загальну здатність вираження ознак. З'єднання моделі прості та ефективні, що забезпечує багаторівневе поєднання функцій.

Алгоритм YOLOv8 використовує функцію втрат, що складається з двох окремих компонентів: втрати класифікації та втрати регресії. Класифікаційні втрати використовують функцію втрат BCE (двійкова крос-ентропія). Як наслідок усунення втрати об'єкта та достовірності об'єкта у вихідних даних за допомогою YOLOv8, оцінки достовірності для кожної категорії виводяться безпосередньо, при цьому максимальне значення береться та

використовується як достовірність блоку прив'язки. Втрата регресії визначається як втрата DFL [55] у поєднанні з втратою CIOU. Функція втрати DFL призначена для оптимізації ймовірності двох позицій, які найбільше збігаються з міткою. Це дозволяє мережі швидко зосередитися на цільовій позиції та розподілі сусідніх регіонів. CIOU враховує три геометричні чинники: мінімізацію нормалізованої відстані центроїда, площу перекриття та узгодженість співвідношення сторін.

Таким чином, модуль WFPN підвищує точність виявлення моделі, зберігаючи відносно низьке збільшення параметрів і обчислювального навантаження. Вбудована компактна головка виявлення об'єктів помітно підвищує точність можливостей виявлення моделі. Однак це також вимагає введення додаткових карт функцій для виявлення, що, як наслідок, призводить до більш складного об'єднання та обробки ознак. Ці кроки призводять до помітного збільшення кількості параметрів і кількості обчислень, необхідних для моделі. Модуль RepNIBMS підвищує точність моделі, одночасно зменшуючи кількість параметрів і обчислень, пов'язаних із моделлю. В цілому, використання 3-фокальної функції підвищує точність моделі.

2.3 Модифікація YOLO11 для виявлення об'єктів з великими відмінностями в розмірі

Проведені дослідження свідчать, що моделі YOLO11 у практичних застосуваннях все ж таки мають певні недоліки. Наприклад, модель має великий об'єм, що ускладнює розгортання на платформах з обмеженими ресурсами (інтеграція не квантизованих версій на мобільних пристроях або одноплатних комп'ютерах та ін.). Це завдання є особливо важливим для розробки ефективної та практичної системи виявлення об'єктів з великими відмінностями в розмірі. Незважаючи на те, що YOLO11 добре працює у

певних сферах, таких як візуальне розпізнавання продукції, все ще потрібні подальша оптимізація та налаштування. Щоб вирішити ці проблеми, модель YOLO11 може бути вдосконалена в наступних трьох напрямках (рис. 2.8) [56].

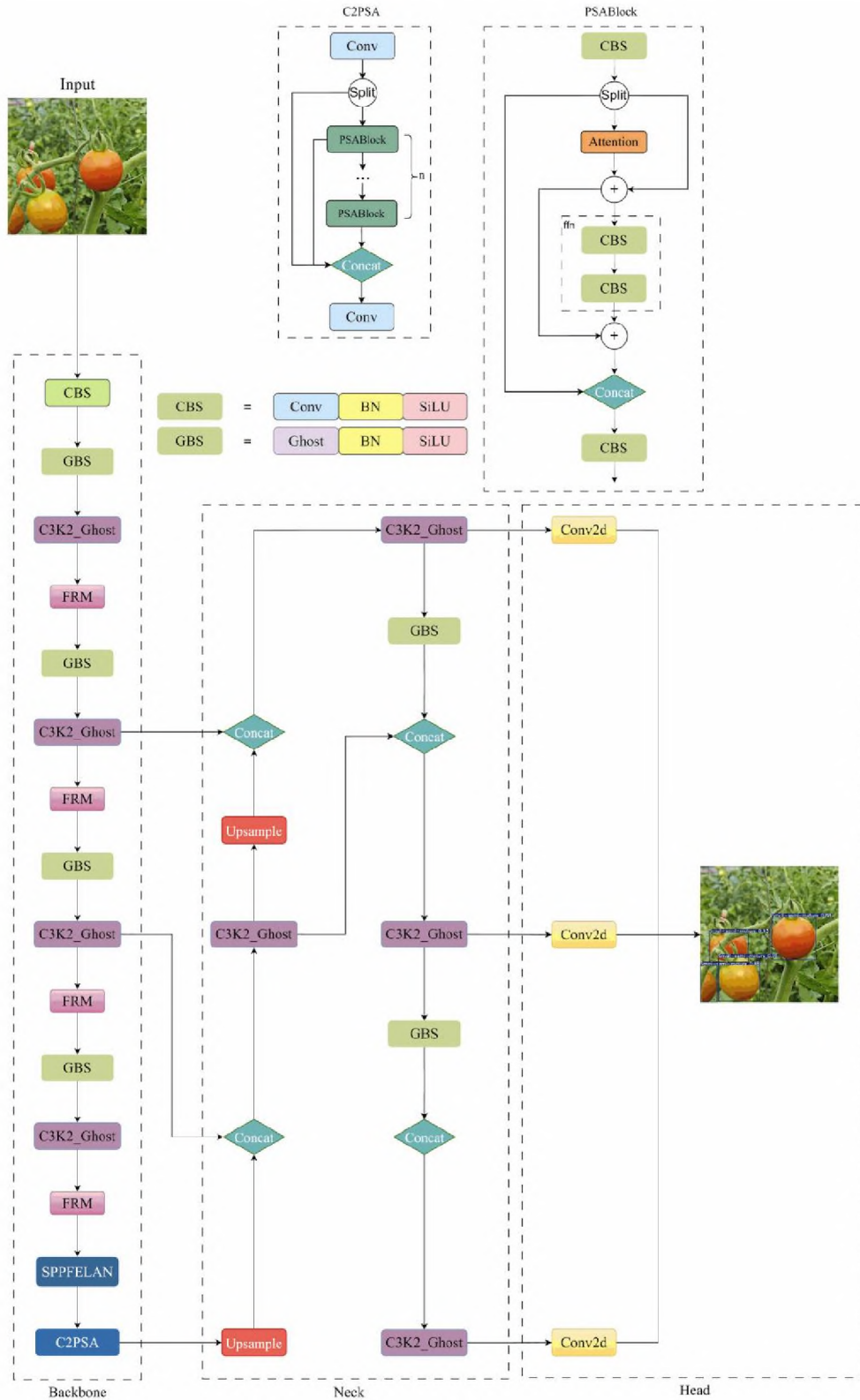


Рисунок 2.8 – Структурна схема моделі GFS-YOLO11 [56]

Модуль C3k2_Ghost запропоновано замінити модуль C3k2 у вихідній мережі, щоб зменшити обчислювальну складність і споживання пам'яті моделлю, а також покращити міркування швидкість моделі. Зважаючи на те, що полегшена конструкція може втратити частину інформації про особливості об'єкту, вбудовують FRM після кожного модуля виділення ознак (C3k2) в магістральну мережу, щоб покращити здатність вилучення ознак звичайних розмірів і малих в складних польових середовищах. Модуль SPPF запропоновано замінити на SPPFELAN у вихідній мережі, що додатково покращує здатність виявлення різних розмірів.

Модуль C3K2_Ghost (рис. 2.9). У оптимізації моделей глибокої нейронної мережі, полегшення є критичною метою, особливо під час розгортання модів на ресурсах. Завдяки введенню структури GhostBottleneck [57], яка замінює традиційну структуру вузьких місць, витрати на обчислення моделі та використання пам'яті значно зменшуються без значного зниження продуктивності моделі.

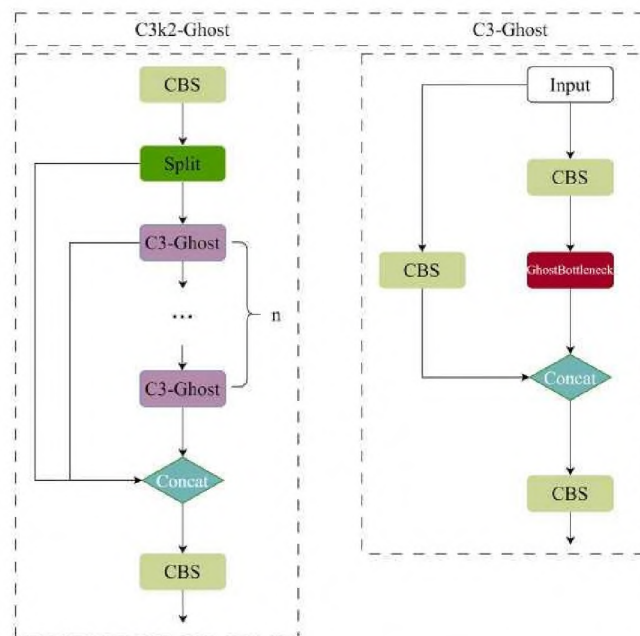


Рисунок 2.9 – Структура мережі модуля C3K2_Ghost

Модуль C3K2 – це модуль вилучення функцій останньої моделі YOLO11. Базуючись на структурі CSPNet [58], він обробляє вхідний графік

ознак, розділяючи його на дві частини та використовуючи модуль вузького місця для багатомасштабного вилучення ознак. Однак, традиційні структури вузьких місць, зазвичай, містять кілька шарів згортки і вимагають багато обчислень. Введення структури GhostBottleneck ефективно вирішує цю проблему. Порівняно з традиційною структурою вузьких місць, GhostBottleneck пропонує більш ефективний спосіб вилучення функцій, який використовує операції GhostConv як своє ядро. Ghost-Conv генерує частину карти функцій за допомогою меншої кількості ядер згортки, а потім розширює її за допомогою недорогих операцій лінійного перетворення для створення більш різноманітних представлень функцій, таким чином значно зменшуючи обчислювальні зусилля, зберігаючи продуктивність. Нарешті, дві частини карти об'єктів з'єднуються разом, щоб отримати кінцеву вихідну карту об'єктів. Завдяки цій філософії проектування GhostBottleneck підтримує або навіть покращує продуктивність моделі, значно зменшуючи кількість параметрів і обчислень. Замінивши структуру вузького місця в модулі C3k2 на структуру GhostBottleneck, можна значно зменшити обчислювальні витрати на етапі виділення ознак моделі, таким чином досягнувши легкої моделі. Безпосередні переваги цього вдосконалення включають вищу швидкість виводу та менші розміри моделі. Більша швидкість виводу дозволяє моделі краще задовольняти потреби додатків у реальному часі.

Модуль FRM (рис. 2.10). Незважаючи на те, що легкий метод може ефективно зменшити обчислювальні витрати, він також може спричинити зниження чутливості моделі до деяких особливостей, таких як чутливість текстури поверхні об'єкту, градієнт кольору та інші тонкі особливості [59]. Щоб вирішити цю проблему можна застосувати структуру під назвою модуль уточнення функцій (FRM), вбудовану після модуля C3K2 магістральної мережі YOLO11. Цей модуль має на меті покращити здатність фіксувати представлення ознак шляхом інтеграції локальних особливостей і інформації про глобальний контекст, щоб покращити продуктивність моделі у визначенні

об'єктів. Основна ідея FRM полягає в зосередженні як на локальних деталях, так і на інформації про глобальний контекст. Для локальних особливостей FRM використовує ефективну згортку з роздільною глибиною [60] для їх виділення.

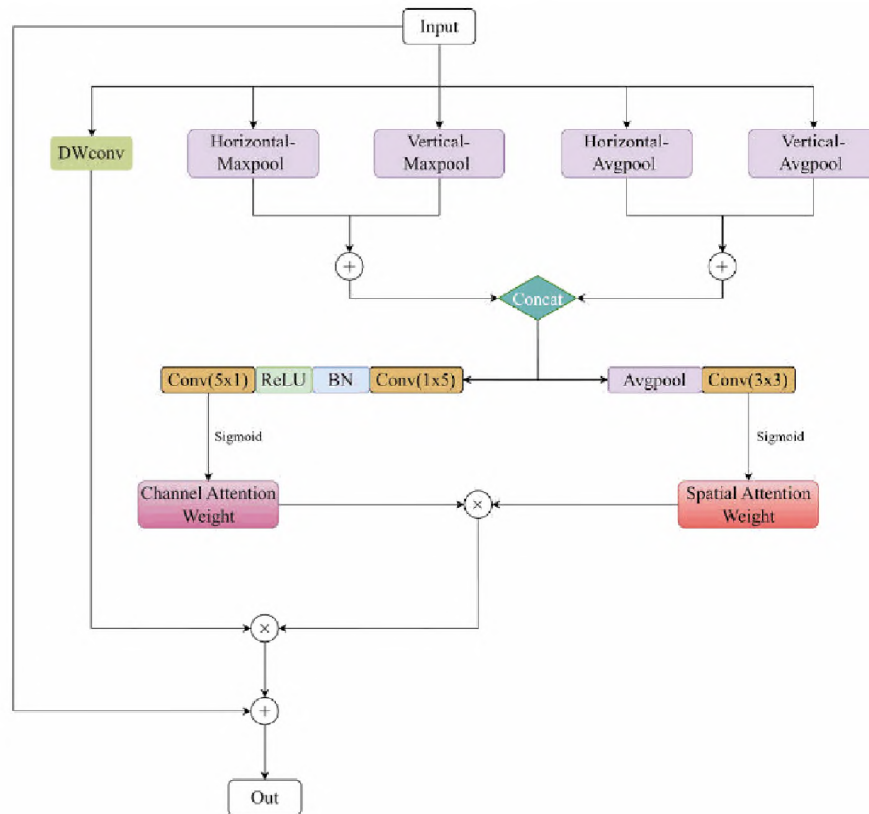


Рисунок 2.10 – Структурна схема моделі FRM

Цей метод значно знижує витрати на обчислення, зберігаючи при цьому чутливість до локальної просторової інформації, що сприяє фіксації тонких особливостей, таких як зміни текстури та кольору на поверхні об'єктів. Отримання глобальної контекстної інформації залежить від операцій об'єднання. FRM використовує середнє об'єднання та максимальне об'єднання в горизонтальному та вертикальному напрямках для ефективного захоплення інформації про функції в різних масштабах.

Середнє об'єднання може відображати загальні статистичні характеристики карт об'єктів, а максимальне об'єднання висвітлює найважливіші місцеві особливості. Поєднання результатів горизонтального та

вертикального об'єднання може більш вичерпно описати просторовий розподіл ознак, що допомагає зрозуміти загальну форму та розподіл кольору об'єктів. Щоб ще більше підвищити здатність вираження ознак, FRM вводить механізми уваги каналу та просторової уваги. Механізм звернення уваги на канали зосереджений на перевірці важливих каналів функцій. Дизайн групової згортки може ефективно зменшити обчислювальну складність, зберігаючи здатність моделювати зв'язок між каналами. Механізм просторової уваги фокусується на важливості різних позицій на карті ознак. По-перше, FRM виконує глобальне середнє об'єднання об'єднаних багатомасштабних ознак, щоб отримати глобальний вектор ознак, який представляє цілі характеристики зображення. Потім за допомогою операції згортки з розміром ядра 3 глобальний вектор ознак перетворюється для генерації ваги просторової уваги. Нарешті, сигмоїдна функція активації використовується для нормалізації значення діаграми уваги між 0 і 1, що вказує на важливість кожного місця. Діаграма уваги каналу та діаграма уваги простору множаться на елементи, щоб отримати остаточну маску уваги. Маска враховує важливість функцій як у вимірі каналу, так і в просторовому вимірі, щоб точніше спрямовувати мережу, щоб зосередитися на важливих областях і функціях, таких як зміни кольору або особливості текстури на поверхні помідорів. Застосування маски уваги до виділених локальних особливостей може ефективно покращити ключові характеристики та придушити фоновий шум, таким чином покращуючи розрізняльність представлення ознак. Нарешті, FRM приймає структуру залишкового з'єднання. Залишкове підключення додає вхід модуля безпосередньо до виходу, ефективно усуває проблему зникнення градієнта, сприяє навчанню мережі та зберігає оригінальну інформацію про функції. Ця конструкція дозволяє краще інтегрувати FRM в магістральну мережу YOLO11 і покращити загальну ефективність виявлення.

Модуль SPPFELAN. У задачах виявлення цілей ефективно виділення та об'єднання багатомасштабних ознак є важливим для розпізнавання об'єктів

різних розмірів і форм [61]. SPP і його ефективна версія SPPF використовують ядра Pool фіксованого розміру для перевірки вхідної карти функцій і видалення багатомасштабних функцій, що ефективно покращує здатність моделі виявляти об'єкти різних розмірів. Однак, SPP і SPPF використовують прості операції конкатенації для інтеграції функцій, отриманих з різних рівнів об'єднання, які можуть не повністю використовувати додаткову інформацію між багатомасштабними функціями, що обмежує продуктивність моделі. Efficient Layer Aggregation Networks (ELAN) [62] демонструє унікальні переваги в об'єднанні функцій. За допомогою паралельних кількох гілок вони витягують різні рівні інформації про функції та виконують ефективне агрегування, що покращує продуктивність, зберігаючи низькі витрати на обчислення. Для того, щоб поєднати переваги SPPF та ELAN та покращити здатність моделі виявляти об'єкти різного розміру, у [56] запропоновано новий модуль під назвою SPPFELAN (рис. 2.11).

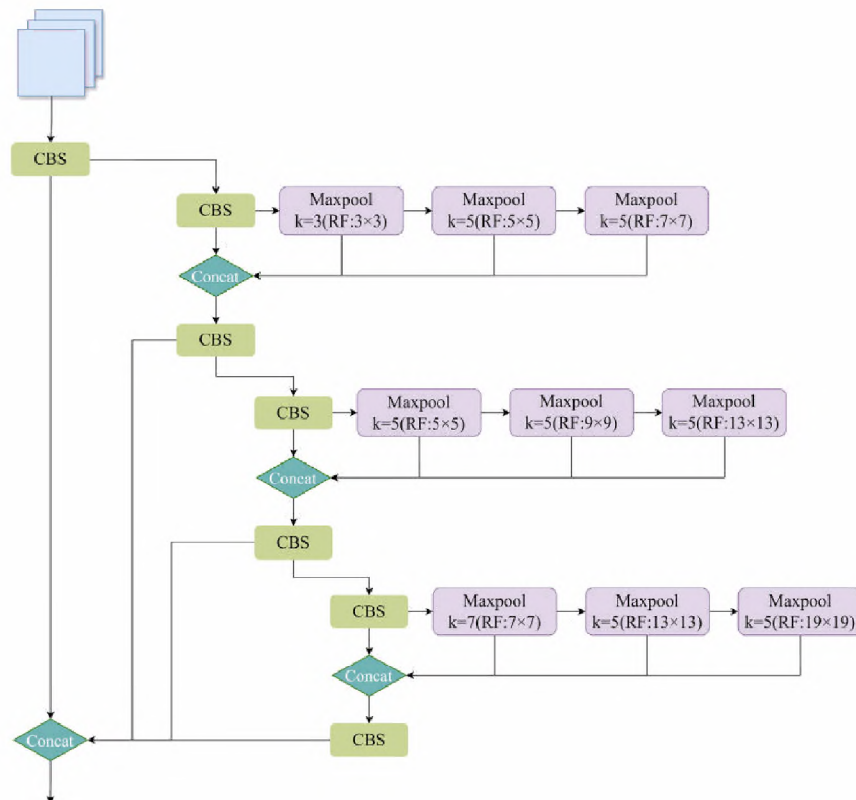


Рисунок 2.11 – Схема структури моделі SPPFELAN

Модуль запозичив ідею ELAN паралельного створення кількох гілок для вилучення багатомасштабних функцій і вміло використовує модуль SPPF як конкретну реалізацію гілок. Кожен модуль SPPF має ядра об'єднання різних розмірів (3, 5, 7), які можуть отримувати різні рівні інформації про функції: ядра об'єднання великого розміру можуть фіксувати глобальні характеристики, такі як загальна форма та розподіл кольору помідора, а малі – ядра об'єднання розмірів можуть зосереджуватися на детальних характеристиках, таких як текстура поверхні та локальні дефекти. Нарешті, модуль SPPFELAN з'єднує та об'єднує вихідні дані всіх гілок, що є простим і ефективним і може інтегрувати інформацію про ознаки з різних рівнів для формування більш повного вираження ознак, додатково покращуючи здатність моделі виявляти зрілість помідорів різних типів, розміри.

Висновки до розділу 2

Сучасний розвиток алгоритмів виявлення об'єктів, зокрема серії YOLO, демонструє значний прогрес у галузі CV. Остання ітерація (YOLO11) впроваджує інновації, які роблять її одним із найефективніших інструментів для вирішення завдань у реальному часі. YOLO11 продемонструвала свою універсальність, підтримуючи широкий спектр завдань, включаючи сегментацію екземплярів, оцінку пози, орієнтоване виявлення об'єктів і об'єктне відстеження. Ця багатозадачність відкриває нові можливості для використання моделі в таких сферах, як автономні транспортні засоби, відеоспостереження, медична візуалізація та промислова автоматизація.

Архітектурні вдосконалення, такі як блоки C3k2 та C2PSA, значно підвищують точність і швидкість роботи моделі, зокрема в умовах складних і високовимогливих середовищ. Інтеграція складних механізмів просторової уваги та вдосконалених методів вилучення функцій дозволяє YOLO11 зберігати високу продуктивність навіть за обмежених обчислювальних

ресурсів, що особливо актуально для крайових пристроїв. Крім того, модель демонструє відмінний баланс між ефективністю та точністю, що робить її придатною для широкого спектра застосувань, починаючи від спрощених пристроїв до високопродуктивних систем.

Задача виявлення малих об'єктів при повітряному моніторингу є актуальною в умовах зростаючого використання дронів у різних галузях, включаючи сільське господарство, безпеку, логістику та попередження про катастрофи. Проблеми, пов'язані зі складним фоном, різноманітними масштабами об'єктів та обмеженими апаратними ресурсами, потребують розробки нових підходів у структурі алгоритмів.

Розглянута модифікація YOLO дозволяє підвищити ефективність у вирішенні завдань розпізнавання щільних малих об'єктів. Інтеграція модуля RepNIBMS, заснованого на ієрархічній агрегації ознак, забезпечила покращене вилучення багатомасштабних ознак, одночасно оптимізуючи обчислювальні ресурси. Модифікована структура WFPN, що включає багатошляхове злиття ознак, підвищує точність виявлення об'єктів різних масштабів за рахунок більш гнучкого синтезу просторової та семантичної інформації. Поєднання нових механізмів обробки ознак, таких як RepIB і WFPN, дозволить досягти значного покращення в точності моделі без критичного збільшення кількості параметрів. Крім того, вдосконалена функція втрат, яка комбінує BCE, CIOU і DFL, сприяє більш точній локалізації та класифікації об'єктів у складних сценах. Розроблений підхід демонструє високий потенціал для впровадження в реальних додатках, зокрема у системах відеоспостереження, сільськогосподарських моніторингових системах та автономних дронах.

Проте подальші дослідження можуть бути спрямовані на оптимізацію архітектури для ще більш ефективного використання обмежених обчислювальних ресурсів у реальному часі.

Ще одна модифікація YOLO11, що досліджена в роботі забезпечує оптимізацію алгоритмів виявлення об'єктів із великими відмінностями у

розмірі. Основні вдосконалення спрямовані на зменшення обчислювальної складності, покращення продуктивності та адаптацію до обмежених апаратних ресурсів. Основні зміни стосуються введення модулів C3K2_Ghost, FRM і SPPFELAN. Використання структури GhostBottleneck у модулі C3K2_Ghost дозволяє значно скоротити витрати на обчислення та зменшити обсяг пам'яті, необхідної для роботи моделі, зберігаючи при цьому її продуктивність. Це забезпечує підвищену швидкість роботи та зменшення розміру моделі, що є важливим для реального часу. Модуль FRM покращує чутливість моделі до тонких особливостей, таких як текстура та градієнт кольору, шляхом інтеграції локальних і глобальних ознак. Використання механізмів уваги сприяє точнішій фокусуванні на важливих областях об'єкта, що підвищує якість розпізнавання. Модуль SPPFELAN об'єднує переваги ELAN і SPPF для ефективною агрегації багатомасштабних функцій. Цей підхід забезпечує моделі можливість однаково ефективно обробляти як великі, так і малі об'єкти.

Таким чином, розглянуті варіанти оптимізації архітектури YOLO11 може не лише підвищити точність і швидкість виявлення, але й робить моделі більш придатними для розгортання на пристроях із обмеженими обчислювальними ресурсами.

РОЗДІЛ 3

РЕКОМЕНДАЦІЇ ЩОДО ВИКОРИСТАННЯ ЗАПРОПОНОВАНОЇ АРХІТЕКТУРИ НЕЙРОННОЇ МЕРЕЖІ

3.1 Формування датасету

Згідно п. 2.1, для реалізації завдання Object Detection в рамках повітряного моніторингу вибрана архітектура нейронної мережі YOLO. Щоб перевірити адекватність роботи нейронної мережі на її основі необхідно використовувати датасет. В якості такого вибраний відкритий датасет з ресурсу Roboflow [63], що містить дані фотозйомки з дрону полів сорго [64] (рис. 3.1). Сорго (лат. *Sorghum bicolor*) [65] (рис. 3.2) – це рід рослин сімейства злакові, що включає однорічні та багаторічні трави.



Рисунок 3.1 – Ресурс Roboflow

Сорго є однією з найважливіших сільськогосподарських культур, що вирощується для різних цілей, таких як виробництво продовольства, корми для тварин та технічних продуктів. вирощують у регіонах зі спекотним кліматом, особливо в посушливих та напівзасушливих зонах.

Завдяки своїй посухостійкості ця рослина стала важливою сільськогосподарською культурою у південних регіонах Європи з обмеженими водними ресурсами, де є проблеми із посушливим кліматом. Сорго стає альтернативою кукурудзі. У Європі основними країнами-виробниками є Франція, Італія, Іспанія та Україна. Використовується сорго як кормова культура та для виробництва біоенергії.



Рисунок 3.2 – Рослинна культура сорго [66]

Згідно з [66], найбільші посіви були у Запорізькій, Херсонській, Миколаївській, Одеській областях. Незважаючи на те, що сорго зараз є нішевою культурою, фахівці звертають увагу на її перспективність, особливо з огляду на природно-кліматичні умови. Це одна з найкращих культур, яка за нинішніх умов дає стабільні результати. Коефіцієнт використання води виготовлення одиниці продукції вдвічі менше, ніж в кукурудзи. Сорго також наполовину менше, ніж кукурудза, висота рослин досягає 120-130 см (потрібно менше вологи та поживних речовин).

Використання open-source датасетів з RoboFlow надає значні переваги, такі як економія часу, доступ до якісних даних, зручні інструменти обробки та прозорість ліцензування. Це робить платформу гарним вибором як для

новачків, так і для професіоналів у галузі CV. RoboFlow пропонує множину попередньо оброблених датасетів для різних завдань (Object Detection, класифікація, сегментація), що заощаджує час на збирання та підготовку даних. Багато датсетів надаються безкоштовно, що знижує витрати на їх створення або купівлю. Можна знайти датасети, які складно зібрати самостійно (наприклад, рідкісні види об'єктів або складні завдання сегментації). Більшість датасетів розмічена вручну або за допомогою напівавтоматизованих підходів високої якості. RoboFlow пропонує дані у зручних форматах (YOLO, COCO, PASCAL VOC та ін.), які легко інтегруються з популярними фреймворками (YOLO, TensorFlow, PyTorch, Detectron2 та ін.). Дані часто вже оброблені (нормалізація, кадрування), що прискорює процес підготовки до навчання. При цьому, RoboFlow дозволяє виконувати аугментацію даних (масштабування, обертання, додавання шуму), змінювати розмір зображень, конвертувати датасети у потрібний формат. RoboFlow дозволяє комбінувати свої дані з відкритими датасетами, створювати власні версії датасетів та ділитися ними. Є можливість налаштувати додаткові варіанти даних для покращення якості навчання. Ці інструменти роблять датасети більш універсальними для різних моделей та завдань. Платформа підтримується активною спільнотою, яка ділиться своїми датасетами, ідеями та кодом. RoboFlow надає детальні інструкції щодо використання датасетів та інструментів, що спрощує процес впровадження. З готовими датасетами можна відразу розпочати навчання моделей, замість витратити тижні або місяці на збір даних. Датасети RoboFlow часто використовуються у дослідженнях, що дозволяє легко порівнювати продуктивність своєї моделі з іншими. Більшість open-source датасетів на RoboFlow мають чітко визначені ліцензії (наприклад, MIT, CC-BY), що спрощує їх використання у проектах. Користувач може бути впевненим у легальності використання даних, що особливо важливо для комерційних програм. При цьому датасет повинен бути представлений у форматі даних YOLO. Він простий, мінімалістичний та призначений для високоефективної

обробки даних. Його популярність обумовлена ефективністю роботи та зручністю інтеграції з бібліотеками, такими як PyTorch та TensorFlow. Кожне зображення в датасеті має окремий текстовий файл розмітки (.txt) Ім'я текстового файлу збігається з ім'ям зображення (наприклад, image1.jpg → image1.txt). У кожному текстовому файлі рядки описують окремі об'єкти на зображенні. При цьому, формат рядка має вигляд:

```
<class_id> <x_center> <y_center> <width> <height>
```

В даному випадку, `class_id` – ID класу об'єкта (ціле число починається з 0); `x_center` – координата X центру об'єкта (нормалізована щодо ширини зображення: від 0 до 1); `y_center` – координата Y центру об'єкта (нормалізована щодо висоти зображення: від 0 до 1); `width` – ширина об'єкта (нормалізована щодо ширини зображення: від 0 до 1); `height` – висота об'єкта (нормалізована щодо висоти зображення: від 0 до 1). Усі значення координат (`x_center`, `y_center`, `width`, `height`) нормалізуються в діапазоні від 0 до 1. Це робить формат незалежним від розмірів зображення. Зображення та файли розмітки зберігаються в окремих папках (рис. 3.3). Для навчання моделей YOLO використовується файл `data.yaml` з описом структури датасету, наприклад:

```
train: dataset/images/train
val: dataset/images/val
nc: 3 # Кількість класів
names: ['cat', 'dog', 'car'] # Імена класів
```

Варто мати на увазі, якщо дані розмічуються вручну, потрібне чітке дотримання формату. Формат не підходить для складніших завдань (наприклад, сегментації об'єктів).

Таким чином, використовуваний датасет описується кількома параметрами:

- загальний обсяг даних – 864 зображень;
- навчальна вибірка – 807 зображень (93%);
- валідаційна вибірка – 25 зображень (3%);
- тестова вибірка – 32 зображення (4%).

Усі зображення були масштабовані до роздільної здатності 416×416 пікселів з розтягуванням пропорцій.

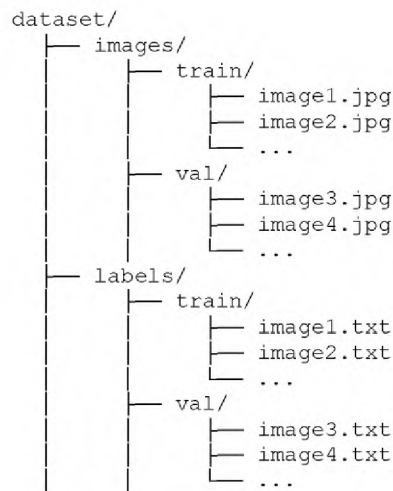


Рисунок 3.3 – Директорії датасета у форматі YOLO

Кожен приклад у тренувальній вибірці генерує 3 варіації, використовуючи наступні налаштування:

- насиченість – в діапазоні від -30 до +30 %;
- яскравість – змінюється від -30 до +30 %;
- експозиція – варіюється в межах від -12 до +12 %;
- розмиття – додається до 1 пікселя;
- шум – торкається до 2 % пікселів.
- виріз – на зображення накладаються 3 прямокутні області розмірами 10 % від загального розміру кожної.

Цей підхід дозволяє значно збільшити різноманітність даних для навчання моделі, покращуючи її стійкість до реальних умов.

3.2 Вибір метрик для оцінки точності нейронних мереж

Щоб оцінити точність нейронної мережі, що використовується для завдання Object Detection необхідно використовувати певні метрики. Розглянемо їх більше детально. Вони дозволяють оцінити, наскільки добре

модель справляється із завданням на кожному етапі навчання. Функції втрат (Losses). Втрати задають напрям для коригування параметрів моделі (ваги та зміщення) під час зворотного поширення помилки (Backpropagation). Це дозволяє моделі покращувати свої прогнози на наступних ітераціях. Високі втрати вказують на те, що модель не навчилася належним чином. Якщо втрати не зменшуються зі збільшенням кількості епох, це може свідчити про проблеми, такі як неправильний підбір гіперпараметрів або недостатньо даних. Зниження втрат на тренувальній вибірці показує, що модель вчиться краще передбачати дані, на яких вона навчається. Зниження втрат на валідаційній вибірці підтверджує, що модель узагальнює свої знання і здатна правильно передбачати дані, які вона не бачила раніше. Якщо втрати на тренувальних даних продовжують знижуватися, але на валідаційних даних починають зростати, це може свідчити про перенавчання (Overfitting), коли модель починає запам'ятовувати дані замість того, щоб вчитися узагальнювати. Таким чином, метрики втрат (Losses) вимірюють розбіжність між прогнозами моделі та реальними (цільовими) значеннями. Вони використовуються як цільові функції, які модель намагається мінімізувати під час навчання. Чим нижче значення втрат, тим ближчі передбачення моделі до істини. Моделі YOLO використовують кілька спеціалізованих функцій втрат, кожна з яких відповідає за різні аспекти навчання. Box Loss (train/box_loss, val/box_loss) – відповідає за точність передбачення координат рамок (bounding boxes); порівнює передбачені рамки з реальними, оцінюючи їх перекриття (IoU). Мінімізація цієї втрати допомагає моделі точніше локалізувати об'єкти. Classification Loss (train/cls_loss, val/cls_loss) відповідає за правильне визначення класу об'єкта та вимірює розбіжність між передбачуваними ймовірностями для кожного класу та справжнім класом. Мінімізація цієї втрати покращує здатність моделі класифікувати об'єкти. Distillation for Localization (DFL) Loss (train/df_loss, val/df_loss). DFL Loss – втрати, що пов'язані з точністю розподілу ймовірностей навколо координат. Вона допомагає покращити деталізацію локалізації об'єктів. Наведемо

прикладі використання Loss на практиці. Якщо втрати не зменшуються після кількох епох, це сигналізує про необхідність перегляду гіперпараметрів (розмір партії, швидкість навчання тощо) або даних (додати більше даних або покращити їхню якість). Якщо втрати на тренувальній вибірці падають, а на валідаційній ростуть, це говорить про перенавчання. Якщо обидві метрики втрат падають, навчання йде успішно. Наприклад, якщо `box_loss` низький, а `cls_loss` залишається високим, це може означати, що модель добре локалізує об'єкти, але їх слабо класифікує. Якщо втрати (Loss) послідовно зменшуються як у тренувальній, і на валідаційній вибірці, – це добрий індикатор того, що модель успішно навчається. Стабілізація Loss – після певної кількості епох графіки втрат стабілізуються, що говорить про досягнення моделі оптимуму (модель «навчилася»).

Метрики втрат. Втрати на тренувальних даних, пов'язані з регресією координат боксів (кордонів об'єктів) – `train/box_loss`. Зниження цього показника вказує на покращення точності визначення координат об'єктів. Втрати, пов'язані з класифікацією об'єктів у тренувальних даних – `train/cls_loss`. Ця метрика відображає, наскільки добре модель розрізняє класи об'єктів. Втрати на валідаційних даних, що пов'язані з регресією координат боксів, описуються `train/df_l_loss`, яка використовується для уточнення локалізації об'єктів. Аналогічні метрики застосовуються для валідаційної та тесових вибірок датасету. Метрики втрат є критично важливим інструментом для управління процесом навчання; розуміння продуктивності моделі; виявлення проблем (недонавчання або перенавчання) та оптимізації гіперпараметрів і архітектури моделі. Без них процес навчання моделі був би хаотичним та некерованим, тому що не було б явного показника, який показує, наскільки добре модель виконує своє завдання.

Метрики якості використовуються для оцінки ефективності та точності моделей ML та AI, тренування та налаштування гіперпараметрів до фінального розгортання. `Metrics/precision` – частка правильно передбачених об'єктів серед усіх передбачень моделі (`Precision` – точність). Висока точність

означає, що модель робить менше хибних позитивних передбачень. Metrics/recall – частка правильно знайдених об'єктів серед реальних об'єктів (Recall – повнота). Висока повнота означає, що модель пропускає менше реальних об'єктів.

Metrics/mAP50 – середня точність (Mean Average Precision) на порозі IoU = 50 %. Відображає, наскільки точно модель розпізнає об'єкти за умови, що прогнози перекриваються з реальними об'єктами хоча б 50 %.

Metrics/mAP50-95 – середня точність при різних порогах IoU (від 50 до 95 % з кроком 5 %). Суворіша метрика, що враховує різні рівні збігу між передбаченими та реальними боксами.

3.3 Оцінка точності нейронних мереж на основі модифікованих архітектур YOLO

Для підтвердження практичної реалізації архітектури YOLO для виконання завдань Object Detection в роботі запропонований програмний код на мові Python (Додаток А). Він демонструє використання бібліотеки Ultralytics YOLO для навчання та оцінки моделі на наборі даних користувача. Код добре підходить для базової візуалізації YOLO-анотацій та передбачає наступну функціональність. Встановлюються бібліотеки та відповідні залежності. Файл датасету завантажується за допомогою посилання Google Drive, що формується через gdown. Завантажений ZIP-файл розпаковується в каталог /content/dataset. З файлу data.yaml витягуються відомості про дані. Передбачається, що файл YAML містить структуру з описом шляхів або метаданих. Функція draw_boxes працює з YOLO-форматом розмітки, де координати центру, ширина та висота представлені у відносних одиницях. Рисується Bounding Boxes на зображенні зеленими прямокутниками (0, 255, 0). Кожне зображення виводиться з використанням OpenCV та Matplotlib (рис. 3.4). Перші 5 зображень з папки train/images відкриваються, розмітка з

відповідних файлів train/labels зчитується та застосовується до зображень. Встановлюється бібліотека Ultralytics, яка підтримує роботу з різними версіями YOLO (в даній роботі розглядаються лише n та s). Імпортується клас YOLO для взаємодії з моделлю. Після ініціалізації моделі проводиться донавчання моделі.

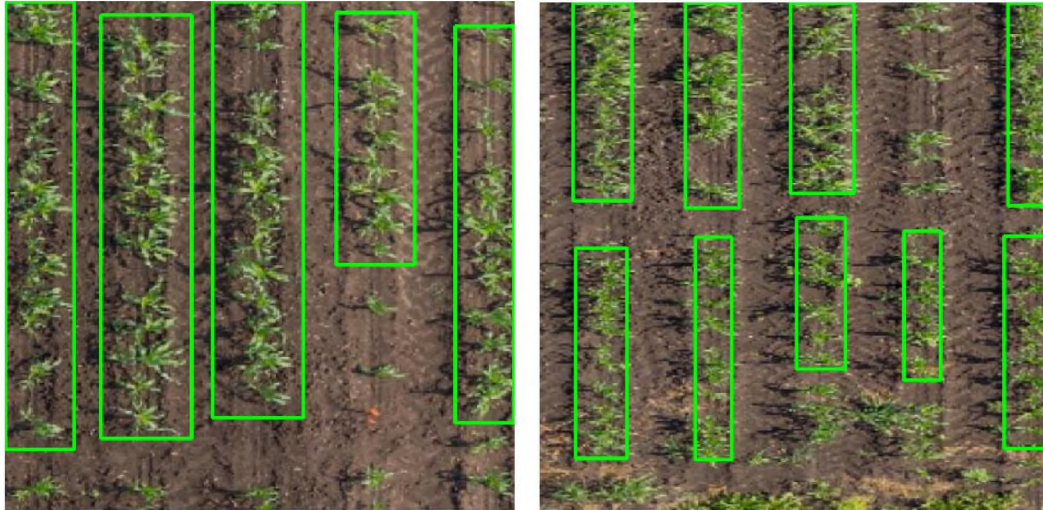


Рисунок 3.4 – Приклад зображення з Bounding Boxes

Функція `os.listdir` отримує список усіх файлів у папці з результатами. Папка `results_dir` (зберігаються результати навчання) створюється автоматично. YOLO версії Nano (n) оптимізована для роботи на пристроях з обмеженими обчислювальними ресурсами, наприклад, вбудовані системи, мобільні пристрої, мікроконтролери та ін. (рис. 3.5).

	from	n	params	module	arguments	
0		-1	1	464	ultralytics.nn.modules.conv.Conv	[3, 16, 3, 2]
1		-1	1	4672	ultralytics.nn.modules.conv.Conv	[16, 32, 3, 2]
2		-1	1	7360	ultralytics.nn.modules.block.C2f	[32, 32, 1, True]
3		-1	1	18560	ultralytics.nn.modules.conv.Conv	[32, 64, 3, 2]
4		-1	2	49664	ultralytics.nn.modules.block.C2f	[64, 64, 2, True]
5		-1	1	73984	ultralytics.nn.modules.conv.Conv	[64, 128, 3, 2]
6		-1	2	197632	ultralytics.nn.modules.block.C2f	[128, 128, 2, True]
7		-1	1	295424	ultralytics.nn.modules.conv.Conv	[128, 256, 3, 2]
8		-1	1	460288	ultralytics.nn.modules.block.C2f	[256, 256, 1, True]
9		-1	1	164608	ultralytics.nn.modules.block.SPPF	[256, 256, 5]
10		-1	1	0	torch.nn.modules.upsampling.Upsample	[None, 2, 'nearest']
11		[-1, 6]	1	0	ultralytics.nn.modules.conv.Concat	[1]
12		-1	1	148224	ultralytics.nn.modules.block.C2f	[384, 128, 1]
13		-1	1	0	torch.nn.modules.upsampling.Upsample	[None, 2, 'nearest']
14		[-1, 4]	1	0	ultralytics.nn.modules.conv.Concat	[1]
15		-1	1	37248	ultralytics.nn.modules.block.C2f	[192, 64, 1]
16		-1	1	36992	ultralytics.nn.modules.conv.Conv	[64, 64, 3, 2]
17		[-1, 12]	1	0	ultralytics.nn.modules.conv.Concat	[1]
18		-1	1	123648	ultralytics.nn.modules.block.C2f	[192, 128, 1]
19		-1	1	147712	ultralytics.nn.modules.conv.Conv	[128, 128, 3, 2]
20		[-1, 9]	1	0	ultralytics.nn.modules.conv.Concat	[1]
21		-1	1	493056	ultralytics.nn.modules.block.C2f	[384, 256, 1]
22		[15, 18, 21]	1	751507	ultralytics.nn.modules.head.Detect	[1, [64, 128, 256]]

Model summary: 225 layers, 3,011,043 parameters, 3,011,027 gradients, 8.2 GFLOPs

Рисунок 3.5 – Архітектура YOLO версії Nano

Мережа добре підходить для додатків відеоспостереження, дрон-навігації або класифікації об'єктів. Вона має розмір складає 6 МБ а кількість параметрів: $\approx 3,2$ млн. Швидкість роботи та GPU $\approx 150 \dots 200$ FPS.

Друга версія – Small (s) підходить для завдань з складнішими вимогами до якості розпізнавання, наприклад, промислові програми, автономні транспортні засоби, моніторинг з високою якістю (рис. 3.6). Вона має розмір складає 11 МБ а кількість параметрів: $\approx 11,2$ млн. Швидкість роботи та GPU $\approx 100 \dots 150$ FPS.

	from	n	params	module	arguments
0	-1	1	928	ultralytics.nn.modules.conv.Conv	[3, 32, 3, 2]
1	-1	1	18560	ultralytics.nn.modules.conv.Conv	[32, 64, 3, 2]
2	-1	1	29056	ultralytics.nn.modules.block.C2f	[64, 64, 1, True]
3	-1	1	73984	ultralytics.nn.modules.conv.Conv	[64, 128, 3, 2]
4	-1	2	197632	ultralytics.nn.modules.block.C2f	[128, 128, 2, True]
5	-1	1	295424	ultralytics.nn.modules.conv.Conv	[128, 256, 3, 2]
6	-1	2	788480	ultralytics.nn.modules.block.C2f	[256, 256, 2, True]
7	-1	1	1180672	ultralytics.nn.modules.conv.Conv	[256, 512, 3, 2]
8	-1	1	1838080	ultralytics.nn.modules.block.C2f	[512, 512, 1, True]
9	-1	1	656896	ultralytics.nn.modules.block.SPPF	[512, 512, 5]
10	-1	1	0	torch.nn.modules.upsampling.Upsample	[None, 2, 'nearest']
11	[-1, 6]	1	0	ultralytics.nn.modules.conv.Concat	[1]
12	-1	1	591360	ultralytics.nn.modules.block.C2f	[768, 256, 1]
13	-1	1	0	torch.nn.modules.upsampling.Upsample	[None, 2, 'nearest']
14	[-1, 4]	1	0	ultralytics.nn.modules.conv.Concat	[1]
15	-1	1	148224	ultralytics.nn.modules.block.C2f	[384, 128, 1]
16	-1	1	147712	ultralytics.nn.modules.conv.Conv	[128, 128, 3, 2]
17	[-1, 12]	1	0	ultralytics.nn.modules.conv.Concat	[1]
18	-1	1	493056	ultralytics.nn.modules.block.C2f	[384, 256, 1]
19	-1	1	590336	ultralytics.nn.modules.conv.Conv	[256, 256, 3, 2]
20	[-1, 9]	1	0	ultralytics.nn.modules.conv.Concat	[1]
21	-1	1	1969152	ultralytics.nn.modules.block.C2f	[768, 512, 1]
22	[15, 18, 21]	1	2116435	ultralytics.nn.modules.head.Detect	[1, [128, 256, 512]]

Model summary: 225 layers, 11,135,987 parameters, 11,135,971 gradients, 28.6 GFLOPs

Рисунок 3.6 – Архітектура YOLO версії Small

Матриці помилок (Confusion Matrix) для цих двох версій наведені на рис. 3.7 та 3.8. У випадку YOLO, яка є мережею для виявлення об'єктів, важливо досягти балансу між точністю та повнотою. Точність дозволяє уникнути зайвих хибних тривог, тоді як висока повнота дозволяє не пропускати важливі об'єкти. Матриця помилок допомагає аналізувати ці аспекти та відстежувати, які саме помилки частіше трапляються – хибні спрацьовування або пропуски об'єктів.

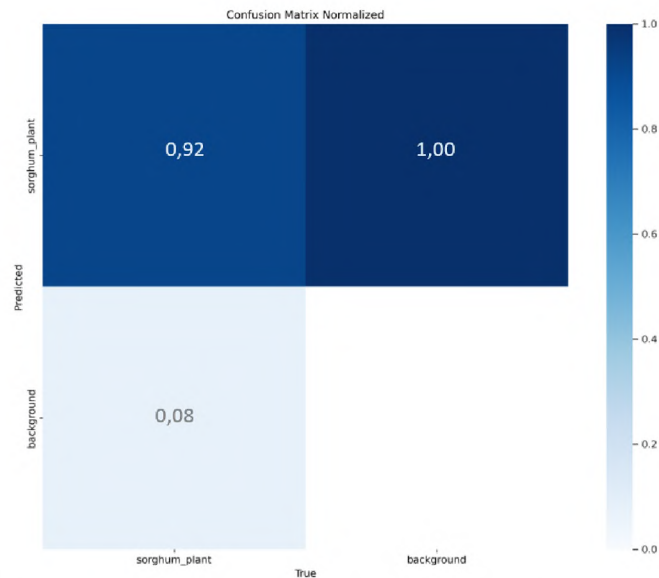


Рисунок 3.7 – Нормалізована матриця помилок для версії Nano

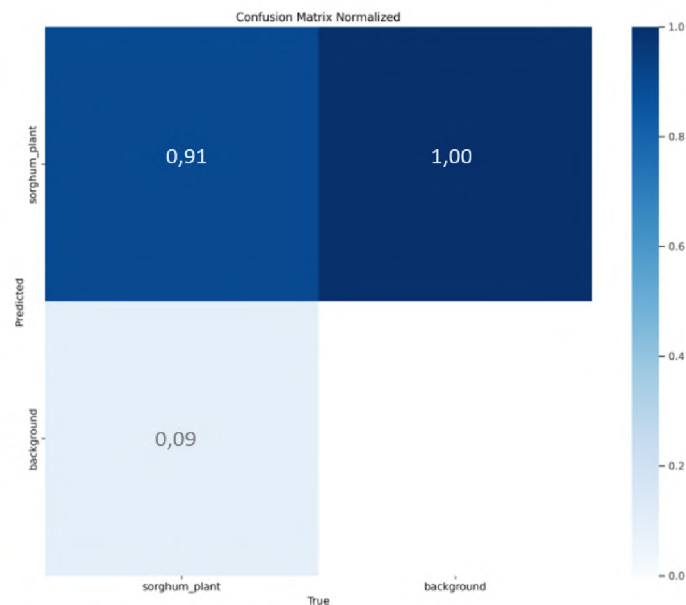


Рисунок 3.8 – Нормалізована матриця помилок для версії Small

Зовнішній вигляд отриманих файлів з результатами наведено на рис. 3.9 та 3.10. Проведемо порівняння за метриками. YOLO Small значно важче за кількістю параметрів та обчислювальної складності (11125971 параметрів проти 3005843 у Nano).

Це означає, що Small вимагає більше обчислювальних ресурсів, але при цьому може досягти кращих результатів. Small показала трохи краще значення Precision (0,864 проти 0,842), що вказує на те, що модель більш точна при класифікації знайдених об'єктів. Однак Recall у Nano виявилось

трохи вищим (0,852 проти 0,838), що говорить про те, що Nano може знаходити більше об'єктів, хоча з меншою точністю.

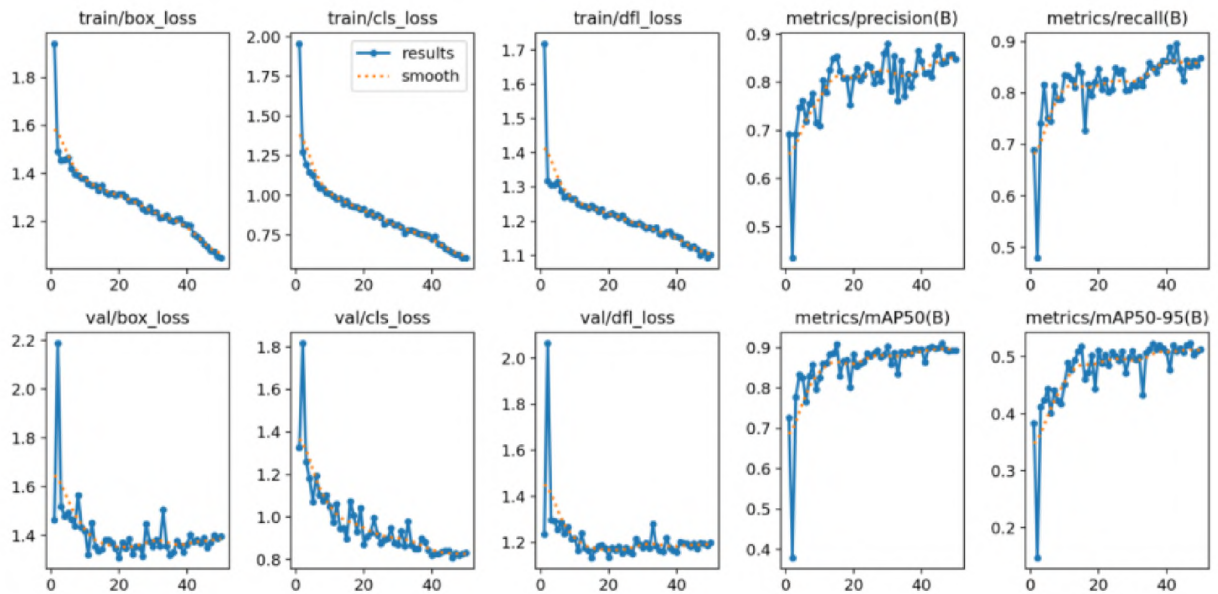


Рисунок 3.9 – Графіки результатів для версії Nano

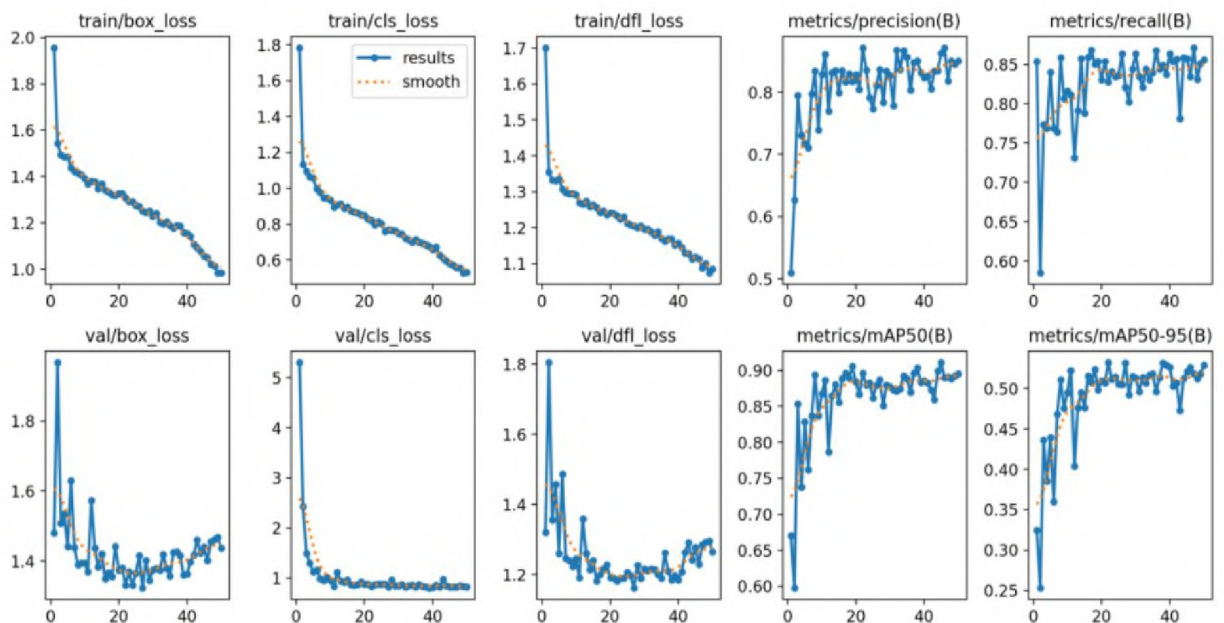


Рисунок 3.10 – Графіки результатів для версії Small

Метрика $mAP@50$ у Small краще (0,904 проти 0,896), що підтверджує її вищу точність при детекції об'єктів, а $mAP@50-95$ також трохи вище у Small (0,532 проти 0,523), що вказує на її здатність більш точно прогнозувати

об'єкти при різних рівнях IoU. Далі виконаємо порівняння за графіками. Втрати – на графіках обох моделей спостерігається подібне поведінка у зниженні втрат за часом (box, cls і dfl losses). Однак у Small більш згладжені та стабільні валідаційні втрати, що вказує на кращу здатність моделі узагальнювати дані та уникати перенавчання. Коливання у Nano трохи вищі, особливо на етапі валідації. Графіки Precision та Recall для обох моделей стабільні після перших 10 епох. Однак Small виглядає більш стабільною у плані Precision, а її значення залишаються вищими протягом усього процесу навчання. Для Small mAP@50 та mAP@50-95 трохи краще – значення більш стабільні, тоді як Nano помітні невеликі коливання.

Приклад практичного застосування YOLO для Object Detection поля з культурою сорго наведено на рис. 3.11.

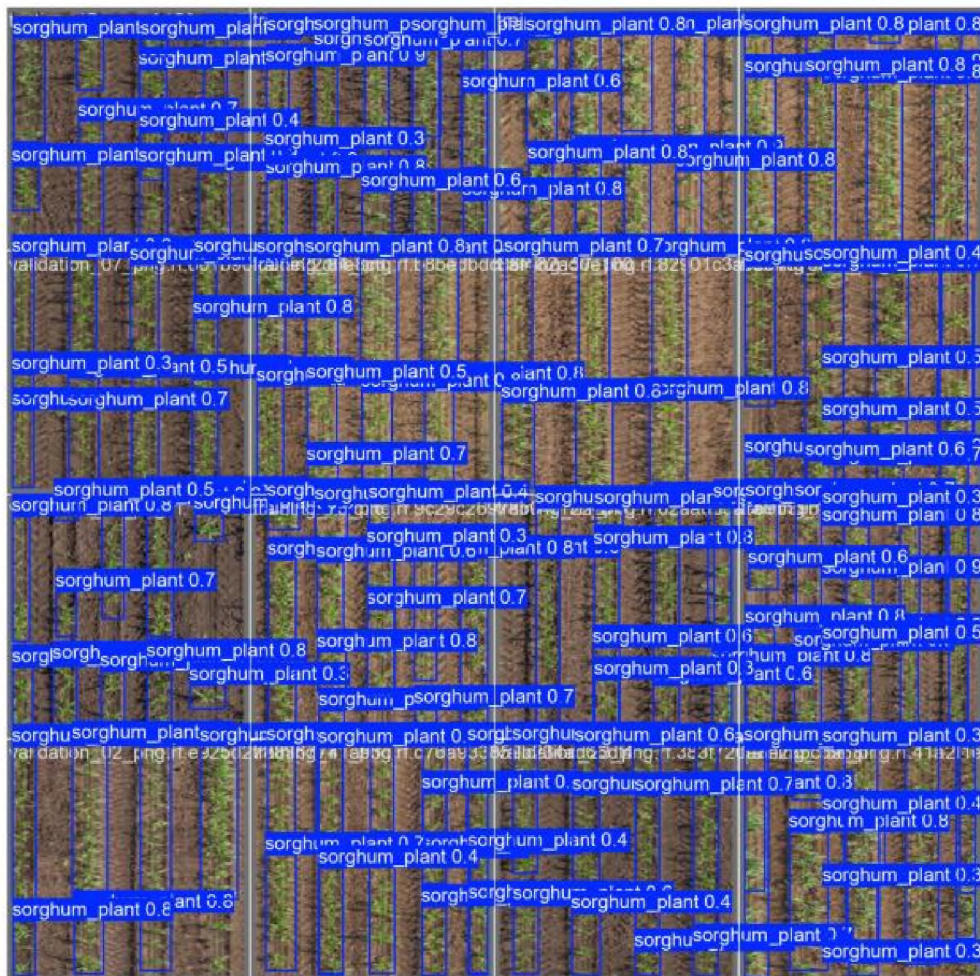


Рисунок 3.11 – Object Detection поля з сорго за допомогою YOLO

В цілому, Small є більш потужною моделлю, демонструючи кращу якість передбачень з точки зору Precision і mAP. Nano більш легка і швидка, що робить її хорошим вибором, якщо важлива швидкість виводу та використання меншого об'єму ресурсів. Таким чином, якщо пріоритетом є точність, стабільність та ресурси не є обмеженням, Small буде оптимальним вибором. Якщо важлива швидкість і легші моделі, Nano може бути кращим.

На основі проведених досліджень можна зробити висновок, що Для підвищення продуктивності розглянутої моделі глибокого навчання за рахунок кількох положень. По-перше, збільшення об'єму даних. Коливання на валідаційних втратах можуть бути викликані недостатнім об'ємом даних. По-друге, підбір гіперпараметрів. Збільшення або зменшення batch size може покращити навчання. Коливання можуть бути викликані занадто високим значенням learning rate. Замість однієї валідаційної вибірки можна використати крос-валідацію, щоб краще оцінити, наскільки модель узагальнює на різних вибірках даних. Щоб зменшити ймовірність перенавчання проводять регуляризацію. Підвищення коефіцієнта регуляризації для wag може допомогти покращити узагальнюючу здатність моделі. Неправильна розмітка в тренувальних та валідаційних даних може призвести до нестабільності під час навчання. Тобто треба перевірит якість розмітки. Нарешті, за умов узгодження з наявними обчислювальними потужностями можна на більш потужні версії YOLO.

3.4 Техніко-економічне обґрунтування прийнятих рішень

Запропонована в роботі модель Object Detection використовує сучасну архітектуру YOLO. Її донавчання проводиться за допомогою open source датасета агрокультури сорго. Для інтеграції нейронних мереж YOLO необхідно узгоджувати наявні обчислювальні ресурси з вимогами щодо швидкості Object Detection. Мінімальний рівень продуктивності такого

компромiсного рiшення може спиратись на версії архiтектур Nano або Small. Для цього можна використовувати рiшення Edge AI. На даний час, багато виробникiв пропонують AI-акселератори в малих форм-факторах, наприклад, рис. 3.12.

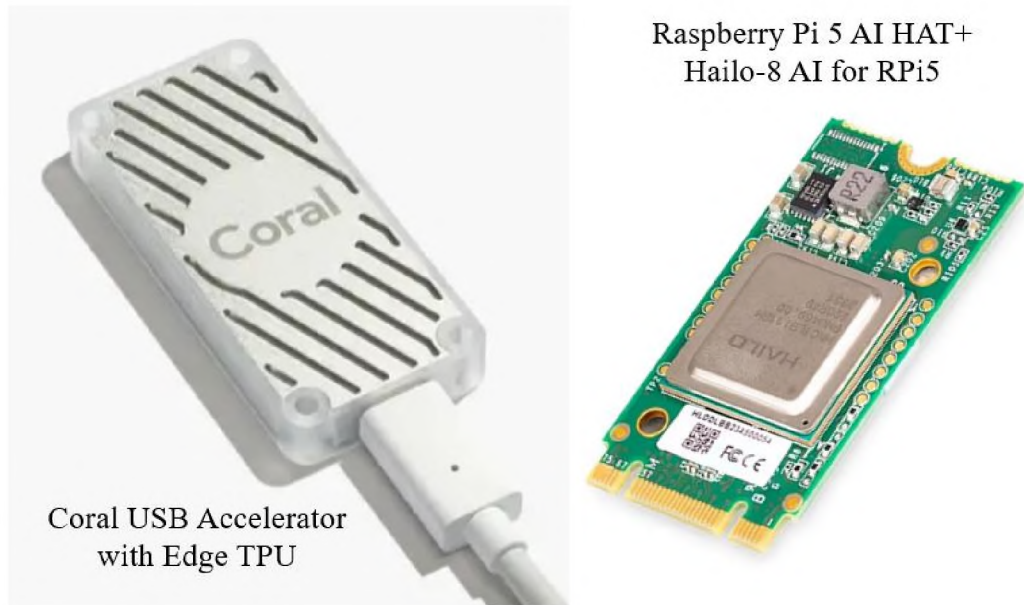


Рисунок 3.12 – AI-акселератори [67, 68]

Крiм цього, сучаснi платформи польотних контролерiв UAV спецiально адаптованi пiд Edge AI (рис. 3.13) [69], а також готовi та запрограмованi комплекти з компаньйон-комп'ютером (рис. 3.14) [70].



Рисунок 3.13 – Плата польотного контролера ARK Jetson PAV Carrier пiд компаньйон-комп'ютер Jetson Orin



Рисунок 3.14 – Польотний контролер Pixhawk 6X Pro Jetson Orin NX 16GB

В якості прикладу розглянемо апаратну складову на базі Raspberry Pi 5 AI NAT+ (Hailo-8 AI). Його вартість складає 4200 грн. Сам мікрокомп'ютер Raspberry Pi 5 (8ГБ) коштує 4700 грн [71]. В якості оптичного сенсору можна застосувати Raspberry Pi Camera Module 3 (12 Мп), вартість якого складає 1660 грн [72]. Тобто витрати на основні компоненти апаратного забезпечення складають 10560 грн. Враховуючі наявність довідкового матеріалу по нейронних мережах YOLO, необхідний робочий час на розробку коду та донавчання мережі складає 8 год при використанні хмарного сервісу Colaboratory від к. Google [73]. Витрати на заробітну плату програміста Python з відповідними навичками та досвідом роботи з проектами AI складає ≈ 48000 грн. Як наслідок, сумарні витрати очікуються на рівні:

$$V = V_1 + V_2 \quad (3.1)$$

де V_1 – вартість апаратного сегменту;

V_2 – витрати на заробітну плату.

Таким чином, орієнтовні витрати на створення макету прототипу для реалізації запропонованої моделі виявлення об'єктів складають

$V = 12560$ грн. При цьому розглядається ситуація наявності готової бази анатованих зображень. Приведення датасету до висунутих вимог може відбуватись за рахунок розширення та збагачення на основі синтетичних даних.

Висновки до розділу 3

На основі проведеного дослідження датасету для завдання об'єктного детектування в межах повітряного моніторингу можна зробити наступні висновки. Використання відкритого датасету з платформи RoboFlow надає значні переваги, зокрема доступ до якісних даних, зручність у форматуванні та готовність до інтеграції з популярними фреймворками, такими як YOLO. Застосування датасету, що містить 864 зображення, дозволило створити структуровані вибірки для навчання, валідації та тестування, забезпечуючи належну різноманітність і стійкість моделі до реальних умов завдяки методам аугментації.

Посухостійкість культури сорго, яка є основним об'єктом дослідження, та її важливість для регіонів із засушливим кліматом додатково підкреслюють актуальність аналізу. Використання RoboFlow як джерела даних знижує витрати та спрощує процес підготовки, дозволяючи зосередитися на розробці моделей та аналізі результатів. Формат YOLO довів свою ефективність у створенні компактних і гнучких рішень для завдань CV. Описаний підхід із застосуванням аугментацій покращує загальні характеристики моделей, роблячи їх адаптивними до практичних умов.

На основі аналізу метрик, використовуваних для оцінки точності нейронних мереж у завданнях Object Detection, можна зробити такі висновки. Метрики втрат (Losses) відіграють ключову роль у процесі навчання моделей, забезпечуючи зворотний зв'язок щодо коригування параметрів і

виявлення потенційних проблем, таких як перенавчання чи недостатнє навчання. Висока ефективність моделей залежить від оптимізації функцій втрат, таких як Box Loss, Classification Loss та DFL Loss, які дозволяють досягати точності у локалізації та класифікації об'єктів.

Оцінка якості роботи моделей за допомогою метрик, таких як Precision, Recall, mAP50 і mAP50-95, забезпечує всебічну перевірку здатності моделі до точного розпізнавання об'єктів і мінімізації хибних передбачень. Високі показники Precision і Recall свідчать про збалансованість моделі, а використання строгих метрик, таких як mAP50-95, дозволяє оцінити її стійкість у різних сценаріях.

Таким чином, правильний вибір і інтерпретація метрик втрат і якості є невід'ємною частиною процесу розробки, налаштування та оцінки нейронних мереж, що гарантує високу продуктивність і надійність у реальних умовах застосування.

На основі аналізу двох модифікованих архітектур YOLO, Small та Nano, можна зробити висновок, що вибір оптимальної моделі залежить від конкретних вимог завдання. YOLO Small демонструє вищі показники точності (Precision) та середньої точності (mAP@50 та mAP@50-95), а також більш стабільну поведінку під час навчання, що робить її придатною для задач, де важлива висока якість розпізнавання та узагальнення даних. Проте ця архітектура вимагає більше обчислювальних ресурсів.

YOLO Nano, навпаки, є легкою та швидкою моделлю, яка оптимально підходить для використання на пристроях з обмеженими обчислювальними можливостями, таких як вбудовані системи або дрони. Її перевагою є вища швидкість роботи та менші вимоги до ресурсів, хоча точність дещо нижча порівняно з Small.

Результати дослідження також підкреслюють важливість таких факторів, як збільшення обсягу даних, покращення якості розмітки, налаштування гіперпараметрів (наприклад, batch size та learning rate), і застосування регуляризації для підвищення стабільності навчання. При

належному використанні цих підходів обидві архітектури можуть бути успішно адаптовані для вирішення конкретних завдань, враховуючи обмеження та пріоритети системи.

У роботі виконано техніко-економічне обґрунтування запропонованих рішень для реалізації моделі виявлення об'єктів на базі архітектури YOLO. Зокрема, було оцінено витрати на апаратне забезпечення, що включає використання Raspberry Pi 5 AI HAT+ (Hailo-8 AI), Raspberry Pi 5 (8ГБ) та оптичного сенсору Raspberry Pi Camera Module 3. Загальна вартість цих компонентів становить 10560 грн. Розробка коду, донавчання моделі та використання хмарного сервісу Colaboratory обраховано з урахуванням середньої заробітної плати фахівця Python для AI-проектів, що призводить до загальних витрат у 12 560 грн.

Значним чинником економічної ефективності є використання готової бази анатованих зображень, що скорочує витрати на створення датасету. У разі необхідності модифікації бази можливо використовувати синтетичні дані для її розширення та збагачення.

Таким чином, запропоноване рішення є фінансово доцільним та забезпечує ефективність реалізації системи Object Detection, враховуючи доступність сучасних апаратних платформ Edge AI, таких як Raspberry Pi 5, та готових інструментів для роботи з нейронними мережами YOLO.

ВИСНОВКИ

Таким чином, повітряний моніторинг сприяє сталому розвитку сільського господарства, підвищенню врожайності та забезпеченню раціонального використання ресурсів, що актуально у контексті сучасних макроекономічних викликів.

Загалом, впровадження дронів у сільське господарство дозволяє зменшити витрати, підвищити продуктивність і забезпечити стійке використання ресурсів, що забезпечує продовольчу безпеку та підвищує рентабельність агросектору.

Загалом, технології CV та Object Detection забезпечують новий рівень автоматизації та точності в аграрному секторі, що сприяє сталому розвитку сільського господарства, підвищенню врожайності та зменшенню використання ресурсів. Це дозволяє аграріям мінімізувати ручну працю, покращити управління полями та оптимізувати витрати, що є важливим кроком до впровадження високотехнологічного та ефективного сільського господарства.

Проведене дослідження продемонструвало ефективність використання відкритого датасету з платформи RoboFlow для задачі об'єктного детектування в межах повітряного моніторингу. Використання якісних даних, зручність у форматуванні та готовність до інтеграції з популярними фреймворками, такими як YOLO, забезпечують високу продуктивність моделі та її адаптацію до реальних умов. Посухостійкість культури сорго, яка є основним об'єктом дослідження, підкреслює важливість використання таких інструментів для забезпечення сталого розвитку аграрних регіонів із засушливим кліматом.

Ефективність нейронних мереж оцінювалася за допомогою метрик втрат та якості, таких як Precision, Recall, mAP50 і mAP50-95. Високі значення Precision та Recall підтверджують здатність моделі до точного розпізнавання об'єктів, а метрика mAP50-95 дозволяє всебічно оцінити

стійкість моделі у різних сценаріях. Порівняння модифікованих архітектур YOLO, зокрема Small та Nano, показало, що вибір між ними залежить від вимог конкретного завдання: YOLO Small демонструє вищу точність, тоді як YOLO Nano є оптимальною для пристроїв з обмеженими обчислювальними можливостями.

Окрім технічних аспектів, дослідження підкреслює економічну доцільність запропонованих рішень, що базуються на використанні апаратних платформ Raspberry Pi 5 AI HAT+ та готових анатованих датасетів. Використання таких рішень забезпечує ефективність впровадження системи об'єктного детектування в умовах реального часу та знижує загальні витрати на реалізацію проекту. Отже, інтеграція сучасних нейронних мереж у процес повітряного моніторингу сприяє не лише підвищенню продуктивності аграрних процесів, а й збереженню ресурсів, зменшенню витрат та забезпеченню стійкого розвитку сільського господарства.

Таким чином, результатами роботи є архітектура нейронної мережі для реалізації функцій Object Detection, порівняльна оцінка точності нейронних мереж на основі модифікованих архітектур Yolo; рекомендації щодо використання запропонованої архітектури нейронної мережі в інтересах аграрного сектору. Вони можуть бути використані для подальших досліджень за даною тематикою та при реалізації концепції Agriculture 4.0.