

**ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА ШТУЧНОГО ІНТЕЛЕКТУ**

КВАЛІФІКАЦІЙНА РОБОТА

на тему: «Розробка системи комп'ютерного зору для автоматизованого
розпізнавання об'єктів на основі методів машинного навчання»

на здобуття освітнього ступеня бакалавра
зі спеціальності 122 Комп'ютерні науки

освітньо-професійної програми Штучний інтелект

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

_____ Максим МУШКО

Виконав:
здобувач вищої освіти
група ШІД-41

Максим МУШКО

Керівник:
Д т.н.,
доцент

Ольга ЗІНЧЕНКО

Рецензент:

Київ 2024

**ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ**

Навчально-науковий інститут інформаційних технологій

Кафедра Штучного інтелекту
Ступінь вищої освіти Бакалавр
Спеціальність 122 Комп'ютерні науки
Освітньо-професійна програма Штучний інтелект

ЗАТВЕРДЖУЮ

Завідувач кафедру Штучного інтелекту

_____ Ольга ЗІНЧЕНКО

« _____ » _____ 2024 р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

Мушку Максиму Володимировичу

1. Тема кваліфікаційної роботи: Розробка системи комп'ютерного зору для автоматизованого розпізнавання об'єктів на основі методів машинного навчання

керівник кваліфікаційної роботи Ольга Зінченко д.т.н., доцент,
затверджені наказом Державного університету інформаційно-комунікаційних
технологій від «27» 02.2024р. № 36

2. Строк подання кваліфікаційної роботи «31» травня 2024р.

3. Вихідні дані до кваліфікаційної роботи: науково-технічна література

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)

1. Дослідження різних методів машинного навчання, вибір найкращих для поставленої задачі та обґрунтування вибору

2. Дослідження видів БПЛА, їх характеристики та застосування

3. Дослідження архітектури YOLO

4. Створення інструкції з встановлення YOLOv4

5. Перелік ілюстративного матеріалу: *презентація*

Дата видачі завдання «27» лютого 2024 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1	Аналіз наукових статей, книг та інших джерел за темою комп'ютерного зору та машинного навчання.	27.02.2024 - 10.03.2024	Виконано
2	Дослідження різних методів машинного навчання, вибір найкращих для поставленої задачі та обґрунтування вибору.	11.03.2024 - 24.03.2024	Виконано
3	Дослідження видів БПЛА, їх характеристики та застосування	25.03.2024 - 07.04.2024	Виконано
4	Дослідження основних будівельних блоків та алгоритмів навчання нейронних мереж	08.04.2024 - 21.04.2024	Виконано
5	Дослідження архітектури YOLO	22.04.2024 - 05.05.2024	Виконано
6	Створення інструкції з встановлення YOLOv4	06.05.2024 - 12.05.2024	Виконано
7	Оформлення роботи: вступ, висновки, реферат	13.05.2024 - 19.05.2024	Виконано
8	Розробка демонстраційних матеріалів	20.05.2024 - 31.05.2024	Виконано

Здобувач вищої освіти

_____ (підпис)

Максим Мушко

Керівник кваліфікаційної роботи

_____ (підпис)

Ольга Зінченко

РЕФЕРАТ

Текстова частина бакалаврської кваліфікаційної роботи: 54 стор., 14 рис., 2 табл., 26 джерел.

Мета роботи – підвищення ефективності використання систем комп'ютерного зору для роботи з зображеннями, отриманими з борту БПЛА для завдань точного землеробства.

Об'єкт дослідження – процес автоматизованого розпізнавання об'єктів на основі методів машинного навчання.

Предмет дослідження – система комп'ютерного зору для завдань точного землеробства.

Короткий зміст роботи: У роботі представлено результати дослідження, спрямованого на розробку інтелектуальної системи автоматизованого розпізнавання об'єктів на зображеннях, отриманих з безпілотних літальних апаратів (БПЛА), з використанням методів комп'ютерного зору та машинного навчання. Запропоновано комплексний підхід до формування та розмітки спеціалізованого набору даних для вирішення задач точного землеробства, що включає використання згорткових нейронних мереж (ЗНМ). Проведено порівняльний аналіз ефективності різних архітектур ЗНМ та методів оптимізації для виявлення об'єктів. Розроблена система демонструє високу точність та швидкість розпізнавання об'єктів, що підтверджується експериментальними результатами на реальних даних. Отримані результати мають практичну цінність для автоматизації процесів моніторингу та аналізу стану сільськогосподарських угідь.

КЛЮЧОВІ СЛОВА: КОМП'ЮТЕРНИЙ ЗІР, МАШИННЕ НАВЧАННЯ, РОЗПІЗНАВАННЯ ОБ'ЄКТІВ, БЕЗПІЛОТНІ ЛІТАЛЬНІ АПАРАТИ, ТОЧНЕ ЗЕМЛЕРОБСТВО, ЗГОРТКОВІ НЕЙРОННІ МЕРЕЖІ, YOLO

ЗМІСТ

<i>1 ТЕОРЕТИЧНІ АСПЕКТИ ПРОБЛЕМИ РОЗРОБКИ СИСТЕМИ КОМП'ЮТЕРНОГО ЗОРУ ДЛЯ АВТОМАТИЗОВАНОГО РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА ОСНОВІ МЕТОДІВ МАШИННОГО НАВЧАННЯ</i>	1
1.1 Системи комп'ютерного зору: сучасні завдання та методи	1
1.2 Інформаційний пошук і навчання	3
1.3 Важка проблема зору	4
1.4 Застосування у робототехніці.....	6
Висновок до розділу 1	7
<i>2 ФОРМУВАННЯ ТА РОЗМІТКА НАБОРУ ДАНИХ ОТРИМАНИХ З БОРТУ БПЛА НА ВИРІШЕННЯ ЗАВДАНЬ ТОЧНОГО ЗЕМЛЕРОБСТВА МЕТОДАМИ МАШИННОГО НАВЧАННЯ</i>	9
2.1 БПЛА. Визначення характеристики та застосування	9
2.2 Згорткові нейронні мережі.....	24
2.3 Згорткові шари (Convolutional Layer)	26
2.4 Карта ознак (Feature Map)	28
2.5 Пулінг (Pooling).....	30
2.6 Повнозв'язкові шари (Fully Connected Layer)	33
2.7 Пряме поширення помилки (backpropagation).....	35
2.8 Зворотне поширення помилки (backpropagation)	37
2.9 Методи оптимізації. Введення в область виявлення об'єктів	38
Висновки до розділу 2	40
<i>3 Практична частина</i>	41
3.1 Аналіз та порівняння схожих наборів даних	41
3.2 Створення власного набору даних	44
3.3 Опис власного набору даних	47
3.4 Установка нейронної мережі YOLOv4	49
3.5 Детальний опис компонентів, використаних під час встановлення YOLOV4...	53

ВИСНОВКИ

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

Додаток А

ВСТУП

В епоху стрімкого розвитку штучного інтелекту, комп'ютерний зір відіграє все більш значущу роль у різних сферах життя. Автоматизоване розпізнавання об'єктів на основі методів машинного навчання стає ключовим фактором для інноваційних рішень в багатьох галузях, таких як точне землеробство, робототехніка, медицина, безпека та багато інших.

Ця робота присвячена розробці системи комп'ютерного зору для автоматизованого розпізнавання об'єктів на основі методів машинного навчання. Акцент робиться на використанні зображень, отриманих з борту безпілотних літальних апаратів (БПЛА), для завдань точного землеробства.

Актуальність дослідження полягає в тому, що точне землеробство потребує інструментів для ефективного моніторингу стану посівів, виявлення проблемних ділянок, а також для оптимізації використання ресурсів. Запропонована система комп'ютерного зору може стати потужним інструментом для вирішення цих завдань, завдяки здатності автоматично розпізнавати та класифікувати різні об'єкти на зображеннях, отриманих з БПЛА.

Наукова новизна роботи полягає в розробці та апробації алгоритму автоматизованого розпізнавання об'єктів на зображеннях, отриманих з борту БПЛА, для задач точного землеробства з використанням методів машинного навчання.

1 ТЕОРЕТИЧНІ АСПЕКТИ ПРОБЛЕМИ РОЗРОБКИ СИСТЕМИ КОМП'ЮТЕРНОГО ЗОРУ ДЛЯ АВТОМАТИЗОВАНОГО РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА ОСНОВІ МЕТОДІВ МАШИННОГО НАВЧАННЯ

1.1 Системи комп'ютерного зору: сучасні завдання та методи

Комп'ютерний зір (КЗ) - це динамічно розвиваюча галузь штучного інтелекту, що досліджує здатність комп'ютерів "бачити" та інтерпретувати світ навколо них. Системи КЗ знаходять широке застосування у вирішенні різноманітних задач, де візуальна інформація є ключовою.

Основні завдання КЗ:

- Розпізнавання об'єктів: Автоматична ідентифікація та класифікація об'єктів на зображеннях та відео.
- Сегментація зображень: Розбивка зображення на окремі регіони, що відповідають певним об'єктам або класам.
- Відстеження об'єктів: Моніторинг руху та поведінки об'єктів у часі.
- 3D-реконструкція: Створення 3D-моделей об'єктів з 2D-зображень або відеопотоків.
- Машинне навчання: Застосування алгоритмів машинного навчання для покращення та автоматизації задач КЗ.

Сучасні методи КЗ:

- Глибоке навчання: Використання штучних нейронних мереж, натхненних структурою людського мозку, для вилучення складних патернів з даних.
- Згорткові нейронні мережі (CNN): Ефективні для задач розпізнавання та сегментації зображень завдяки здатності автоматично виявляти локальні ознаки.
- Обчислення оптичного потоку: Визначення векторів переміщення пікселів між кадрами відео, що використовується для відстеження об'єктів та аналізу динамічних сцен.

- Генеративні моделі: Створюють нові зображення або відео, що відповідають певним заданим характеристикам.

КЗ ґрунтується на принципах комп'ютерного бачення, теорії інформації, теорії ймовірностей, статистики, лінійної алгебри та оптимізації. Ці фундаментальні знання використовуються для розробки алгоритмів та моделей, які дозволяють комп'ютерам "бачити" та інтерпретувати світ так само, як люди.

Досягнення в КЗ значно розширили можливості комп'ютерів у роботі з візуальною інформацією. Розроблені алгоритми та моделі демонструють вражаючі результати в таких сферах, як:

- Медична діагностика: Автоматичний аналіз медичних зображень для виявлення патологій.
- Робототехніка: Розробка автономних роботів, здатних до навігації, маніпуляції та взаємодії з середовищем.
- Транспорт: Автоматичні системи керування транспортними засобами для безпечного та ефективного водіння.
- Безпека: Системи відеоспостереження та розпізнавання облич для запобігання злочинам та тероризму.
- Дослідження: Аналіз наукових даних, отриманих з телескопів, супутників та інших джерел.

Комп'ютерний зір - це динамічно розвиваюча галузь штучного інтелекту, що ґрунтується на міцному науковому фундаменті. Сучасні методи КЗ, такі як глибоке навчання та згорткові нейронні мережі, відкривають нові можливості для вирішення складних задач, пов'язаних з візуальною інформацією. КЗ має значний потенціал для покращення життя людей у багатьох сферах, від медицини та робототехніки до транспорту та безпеки [1].

1.2 Інформаційний пошук і навчання

Інформаційний пошук (ІП) - це процес пошуку інформації в базі даних або на веб-сайті. Цей процес ґрунтується на принципах інформатики, теорії інформації та штучного інтелекту [2].

Методи ІП:

- Пошук за ключовими словами: найпростіший метод, який використовує відповідність між запитом користувача та документами, що містять ці ключові слова.
- Морфологічний аналіз: враховує граматичні форми слів, що дозволяє знаходити документи, навіть якщо в них використовуються синоніми або однокореневі слова.
- Семантичний пошук: аналізує значення слів та фраз у запиті користувача, щоб знаходити документи, що відповідають його наміру.

Пошук на основі графів: використовує граф, де вершини представляють документи, а ребра - зв'язки між ними, для ранжування документів за релевантністю.

Оцінка результатів ІП:

- Точність: частка релевантних документів, знайдених системою.
- Повнота: частка релевантних документів у колекції, знайдених системою.
- Відповідність: ступінь релевантності знайдених документів запиту користувача.
- Ефективність: час, витрачений системою на пошук документів.

Машинне навчання (МН) - це підгалузь штучного інтелекту, яка дозволяє комп'ютерам навчатися без явного програмування. МН ґрунтується на алгоритмах, які можуть виявляти закономірності та робити прогнози на основі даних.

Типи алгоритмів МН:

- Навчання з учителем: алгоритми отримують набір даних, де кожен екземпляр має мітку, що описує його клас.

- Навчання без учителя: алгоритми отримують немічені дані та повинні самостійно виявляти структуру та закономірності.
- Півнаглядне навчання: алгоритми отримують набір даних з частковою міткою.

Застосування МН:

- Розпізнавання образів: класифікація зображень, відео, тексту та інших даних.
- Прогнозування: прогнозування майбутніх подій, таких як ціни на акції, погода або результати виборів.
- Рекомендаційні системи: рекомендування продуктів, фільмів, музики та інших товарів користувачам.
- Обробка природної мови: аналіз та розуміння людської мови.

ІІ та МН ґрунтуються на міцному науковому фундаменті, який поєднує в собі знання з інформатики, математики, статистики, психології та лінгвістики. Ці знання використовуються для розробки алгоритмів та моделей, які дозволяють комп'ютерам ефективно обробляти та аналізувати інформацію [3].

Інформаційний пошук та машинне навчання - це дві важливі галузі штучного інтелекту, які мають широке застосування в різних сферах. ІІ допомагає людям знаходити необхідну інформацію, а МН дозволяє комп'ютерам навчатися та робити прогнози на основі даних. Ці дві галузі постійно розвиваються, що відкриває нові можливості для покращення життя людей.

1.3 Важка проблема зору

"Важка проблема зору" описує складність розуміння того, як люди сприймають та інтерпретують візуальну інформацію з навколишнього світу. Незважаючи на те, що людські очі отримують величезну кількість даних у вигляді світла, наш мозок здатний без зусиль обробляти цю інформацію, визначаючи об'єкти, їхні форми, просторові відносини, рух та інші важливі характеристики.

Ця задача виявляється надзвичайно складною для комп'ютерних систем, адже вона потребує не лише реєстрації візуальних даних, але й їхнього глибокого розуміння. Наразі комп'ютерні алгоритми не можуть досягти рівня людського зору, адже їм бракує багатьох ключових компонентів, притаманних біологічним системам.

Багатогранність візуальної інформації. Світло, що потрапляє до очей, несе в собі безліч характеристик, таких як інтенсивність, довжина хвилі, поляризація та фаза. Людський мозок здатний інтегрувати цю інформацію, формуючи цілісне уявлення про навколишнє середовище [4].

Візуальні дані часто бувають неоднозначними, адже один і той самий об'єкт може виглядати по-різному залежно від освітлення, ракурсу, текстур та інших факторів. Людський мозок використовує контекстну інформацію та знання про світ, щоб правильно інтерпретувати візуальні сигнали.

Людський зір ґрунтується на складних нейронних процесах, що відбуваються в сітківці та інших структурах мозку. Ці процеси включають виявлення країв, текстур, руху та інших локальних характеристик зображення.

Після низькорівневої обробки візуальна інформація потрапляє до вищих ділянок мозку, де відбувається її розпізнавання, класифікація та інтеграція з іншими сенсорними даними. Ці процеси дозволяють нам ідентифікувати об'єкти, сприймати просторові відносини та розуміти динаміку навколишнього світу [5].

Незважаючи на складність "важкої проблеми зору", дослідження в галузі штучного зору досягли значного прогресу. Розробка нових алгоритмів машинного навчання, таких як згорткові нейронні мережі, дозволила комп'ютерним системам досягти вражаючих результатів у розпізнаванні об'єктів, сегментації зображень та інших задачах комп'ютерного зору.

"Важка проблема зору" залишається одним із найскладніших викликів у штучному інтелекті. Її вирішення потребуватиме глибокого розуміння принципів роботи людського зору та розробки нових алгоритмів, здатних імітувати ці принципи в комп'ютерних системах.

1.4 Застосування у робототехніці

Комп'ютерний зір (КЗ) відіграє ключову роль у розвитку сучасної робототехніки, надаючи роботам здатність "бачити" та розуміти навколишнє середовище [6]. Завдяки КЗ роботи можуть виконувати широкий спектр задач, таких як:

- Навігація: Автономне переміщення в просторі без зіткнення з перешкодами.
- Маніпуляція: Захоплення, переміщення та обробка об'єктів.
- Взаємодія з людьми: Розпізнавання людей, розуміння їхніх жестів та виконання інструкцій.
- Інспекція та контроль: Автоматизований моніторинг та аналіз стану середовища.
- Пошук та рятування: Відшукування людей та об'єктів у складних умовах.

Алгоритми та методи для вилучення інформації з зображень та відео, таких як розпізнавання об'єктів, сегментація зображень, відстеження об'єктів та 3D-реконструкція. Застосування алгоритмів машинного навчання для покращення та автоматизації задач КЗ, таких як розпізнавання патернів, прогнозування та прийняття рішень. Принципи проектування, керування та динаміки роботів, які необхідні для інтеграції КЗ з робототехнічними системами [2].

Роботи з КЗ можуть працювати без постійного втручання людини, що робить їх більш гнучкими та ефективними. КЗ може використовуватися для точного позиціонування та маніпуляції об'єктами, що робить роботів цінними інструментами в різних галузях. КЗ дає можливість роботам виконувати складні та небезпечні для людей задачі, такі як робота в екстремальних умовах або з небезпечними матеріалами.

Приклади застосування КЗ в робототехніці:

- Промислова робототехніка: Застосування роботів з КЗ для складання, фарбування, зварювання та інших задач на виробництві.

- Медична робототехніка: Використання роботів з КЗ для хірургічних операцій, реабілітації та інших медичних процедур.
- Сільськогосподарська робототехніка: Застосування роботів з КЗ для збору врожаю, обробки ґрунту та інших сільськогосподарських робіт.
- Дослідження та пошук: Використання роботів з КЗ для дослідження космосу, океану та інших небезпечних або недоступних для людей місць.

Комп'ютерний зір став невід'ємною частиною сучасної робототехніки, надаючи роботам здатність "бачити" та розуміти навколишнє середовище. Завдяки КЗ роботи можуть виконувати широкий спектр задач, що робить їх цінними інструментами в різних галузях. Наукові дослідження в цій сфері постійно розвиваються, що відкриває нові можливості для розробки більш інтелектуальних та автономних роботів.

Висновок до розділу 1

КЗ ґрунтується на міцному науковому фундаменті, що поєднує в собі знання з різних галузей, таких як математика, інформатика, фізика, психологія та нейронауки. Ці знання використовуються для розробки алгоритмів та моделей, які дозволяють комп'ютерам "бачити" та інтерпретувати світ навколо них.

Поєднання комп'ютерного бачення та машинного навчання дозволяє комп'ютерам витягувати інформацію з зображень та відео, а потім використовувати цю інформацію для виконання складних задач, таких як розпізнавання об'єктів, сегментація зображень, відстеження об'єктів та 3D-реконструкція [7].

Використання штучних нейронних мереж, натхненних структурою людського мозку, дозволило значно покращити точність та ефективність методів КЗ. Наявність великих наборів даних з зображеннями та відео відіграє важливу роль у навчанні та вдосконаленні алгоритмів КЗ.

Роботи з КЗ можуть виконувати складні та небезпечні для людей задачі, такі як робота в екстремальних умовах або з небезпечними матеріалами. КЗ використовується для діагностики захворювань, планування хірургічних операцій та

розробки нових методів лікування. КЗ використовується для розпізнавання обличь, відстеження людей та моніторингу громадських місць.

Транспорт: КЗ використовується для розробки автономних автомобілів та інших безпілотних транспортних засобів.

КЗ використовується для дослідження космосу, океану та інших небезпечних або недоступних для людей місць.

КЗ - це динамічно розвиваюча галузь, яка постійно вдосконалюється. Нові дослідження та розробки відкривають нові можливості для застосування КЗ в різних сферах. Очікується, що КЗ буде відігравати ще більш важливу роль у житті людей в найближчі роки.

Комп'ютерний зір - це потужний інструмент, який ґрунтується на міцному науковому фундаменті. Сучасні методи КЗ, що поєднують комп'ютерне бачення та машинне навчання, дозволяють комп'ютерам "бачити" та інтерпретувати світ навколо них з більшою точністю, ніж будь-коли раніше. КЗ використовується в багатьох сферах і його значення буде лише зростати в найближчі роки.

2 ФОРМУВАННЯ ТА РОЗМІТКА НАБОРУ ДАНИХ ОТРИМАНИХ З БОРТУ БПЛА НА ВИРІШЕННЯ ЗАВДАНЬ ТОЧНОГО ЗЕМЛЕРОБСТВА МЕТОДАМИ МАШИННОГО НАВЧАННЯ

2.1 БПЛА. Визначення характеристики та застосування

Для виконання повітряної зйомки, відповідну знімальну апаратуру встановлюють на носій, який піднімає її на необхідну висоту, переміщає у просторі і забезпечує певне орієнтування на місцевості. Окрім знімальної апаратури, на носіїв монтують цілий ряд додаткового приладдя, необхідного для виконання зйомочних робіт.

Використання безпілотних літальних апаратів, вирішує цілий ряд задач, який на в умовах сьогодення такий широкий, що для вдалого вирішення кожного із них застосовуються системи спеціально оптимізовані під вирішення певних функцій. При цьому визначне місце займає оптимальний вибір певного виду безпілотника, який можна вибрати за основними характеристиками які зображені на рис. 2.1.

Носії зйомочної апаратури ділять на такі основні групи: підводні, наземні, авіаційні й космічні. Найбільш широко для дистанційного зондування використовуються авіаційні й космічні апарати, серед яких виокремлюються безпілотні літальні апарати.

Безпілотні літальні апарати (БПЛА, БЛА, безпілотники, дрони) – літальний апарат без екіпажу на борту, який призначений для аерозйомки й спостереження у реальному часі за навколомними об'єктами.

За сучасним визначенням, «безпілотником» є тільки той апарат, котрий перебуває під постійним дистанційним контролем пілота і призначений для повернення на злітно-посадковий майданчик й для подальшого повторного використання [8].

Раніше радіокеровані й повністю автоматизовані апарати об'єднували ціле поняттям безпілотна авіація – літаки, керування якими здійснюється без пілота, використовуючи прилади різних систем (радіолокації, телебачення), які подають

команди на автопілот. При цьому елементи системи керування розташовані за межами літака і можуть бути на землі, на воді й у повітрі, на місці старту, на маршруті польоту й у районі цілі.

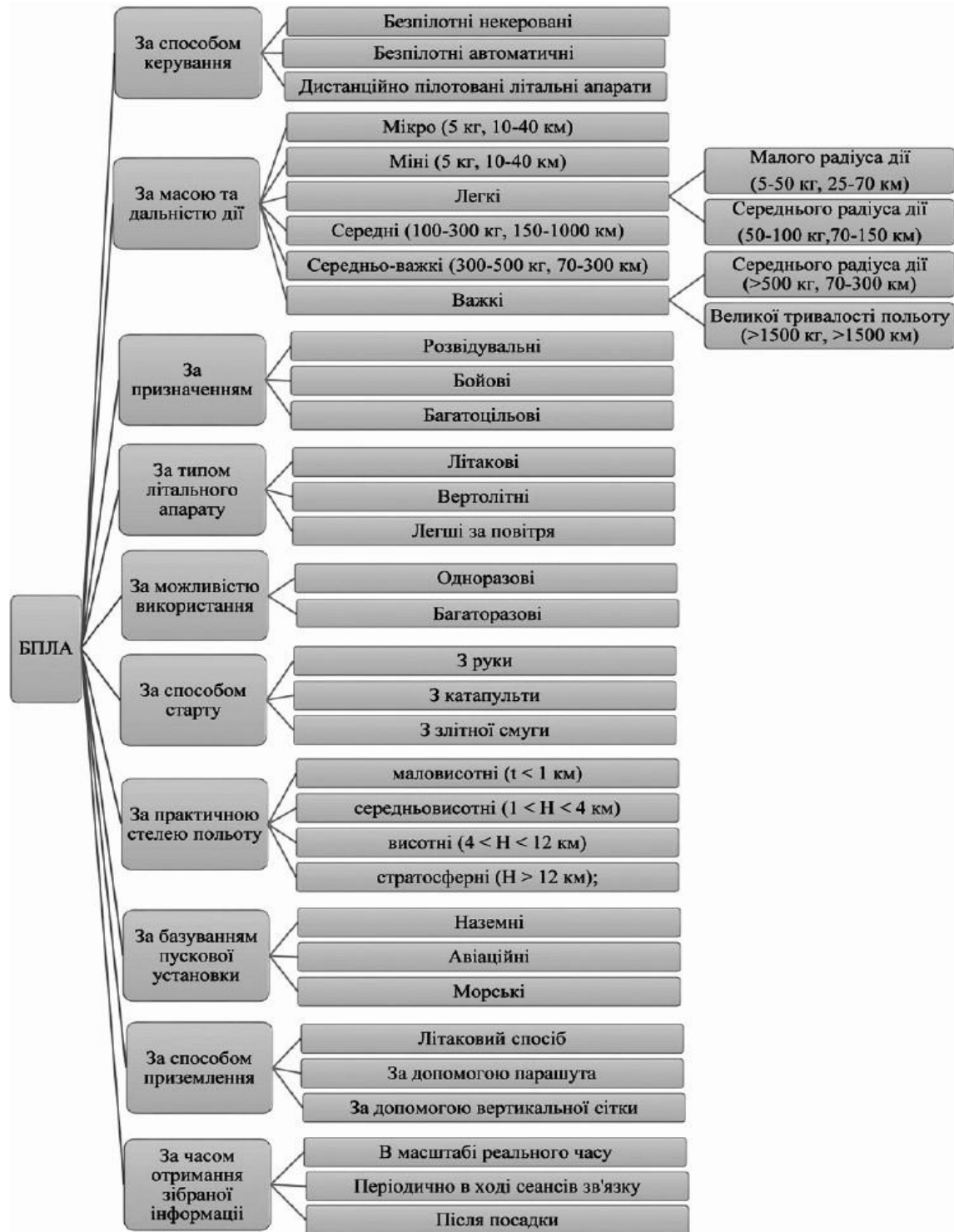


Рис. 2.1. Класифікація БПЛА

В залежності від методів управління, можна виділити різновиди безпілотних літальних систем:

1. Безпілотні некеровані літальні апарати (БНЛА);
2. Безпілотні автоматичні літальні апарати (БАЛА);
3. Безпілотні дистанційно-пілотовані літальні апарати (ДПЛА).

На початку XXI століття в авіації спостерігається швидке зростання використання саме ДПЛА, які часто називають "дронами" або UAV. Під цими термінами мають на увазі повітряні судна, якими керує пілот через канали зв'язку [9, с. 936].

Спочатку безпілотники були призначені для військових цілей, але сьогодні їх все частіше використовують для різноманітних цивільних завдань, таких як фотомонтаж, рятувальні операції, моніторинг дорожньої інфраструктури, сільське господарство та авіафотозйомка.

Безпілотні літальні апарати повинні дотримуватися правил безпеки польотів. Міжнародні авіаційні правила, затверджені після 1944 року на рівні ООН, забороняють безпілотникам літати над територією іншої країни без її дозволу. У Європейському Союзі існує система регулювання польотів дронів на основі правил країн-учасниць.

Для нормативного регулювання використання безпілотних апаратів виділяють такі типи:

1. Дистанційно пілотовані авіаційні системи (ДПАС/RPAS) - системи, що включають літальний апарат, керований пілотом з віддаленої пілотної станції, а також інші компоненти, необхідні для роботи.

2. Безпілотні автономні системи (БАС/UAS) - системи, які функціонують автономно й керуються за допомогою комп'ютера без втручання пілота після зльоту.

Національні правила використання безпілотників в різних країнах відрізняються, що ускладнює взаємне визнання і відкриття повітряного простору.

Тому ведеться робота над розробкою гармонізованих норм і правил використання БПЛА в Європейському Союзі.

Технічні характеристики безпілотних літальних апаратів різняться за розміром, продуктивністю й типом. Вони можуть бути майже непомітними, як комахи, або великими, схожими на пілотовані літаки. Їх швидкість може сягати 277 м/с, і керування ними може здійснюватися за допомогою різних пристроїв та програмного забезпечення супутникового зв'язку [10, с 1137].

Чинниками застосування БПЛА в якості нового інструменту отримання аероданих є недоліки двох традиційних способів отримання даних дистанційного зондування за допомогою космічних супутників (космічне знімання) і повітряних пілотованих апаратів (аерофототопографічне знімання).

Дані космічної зйомки дозволяють отримати знімки з максимальним розрізненням у 30 см, що недостатньо для крупномасштабного картографування. Окрім того, не завжди вдається підібрати безхмарні зображення з архіву. У разі знімання під замовлення втрачається оперативність отриманої інформації. Відносно компактних ділянок оператори й дистриб'ютори переважно не виявляють гуманної цінової політики.

Традиційна аерофотозйомка, яка проводиться за допомогою крупногабаритних літаків (Ан-30, Ту-134, Ан-2, Іл-18, Cessna, L-410) або гелікоптерів (Мі-8Т, Ка-26, AS-350) вимагає високих економічних витрат на обслуговування й заправку ПММ, що призводить до підвищення вартості вихідної продукції.

Розкриємо випадки, коли застосування стандартних авіаційних комплексів нерентабельно:

- зйомка незначних об'єктів й малих за площею територій. У цьому випадку економічні й часові витрати на організацію робіт, що припадають на одиницю відзнятої площі, істотно перевершують аналогічні показники при зніманні великих масивів (тим більше для об'єктів, значно віддалених від злітної смуги);

- при необхідності проведення регулярного знімання у цілях моніторингу протяжних об'єктів: ЛЕП, трубопроводи, транспортні магістралі.

Таким чином, очевидними перевагами застосування БПЛА є:

- можливість зйомки із незначних висот із наближенням до об'єктів;
- отримання знімків високої просторової роздільної здатності;
- рентабельність;
- оперативність отримання зображень;
- можливість застосування у зонах надзвичайних ситуацій без ризику для життя й здоров'я пілотів.

Варто зазначити, що технологія аерофотозйомки із БПЛА в значній мірі відпрацьована. На тепер велика частина існуючих і експлуатованих апаратів призначені для повітряного рекогносрування й спостереження, які здійснюються за допомогою фото-і відеознімання.

Використання безпілотних літальних апаратів також супроводжується рядом негативних аспектів.

По-перше, підвищена аварійність БПЛА часто пов'язана з відсутністю системи розпізнавання перешкод та недоліками в автопілотованих системах, які можуть бути обмежені для зменшення ваги та вартості обладнання. Це може призвести до ризику втрати апарату та обладнання, що змушує деякі компанії віддавати перевагу оренді льотних годин у спеціалізованих організаціях з безпілотними запусками [11].

По-друге, наразі ринок цивільних БПЛА, зокрема для аерофотозйомки, тормозиться відсутністю нормативної бази для їх інтеграції в повітряний простір. Це ще не було повністю вирішено жодною країною світу, хоча в Україні вже зроблено перші кроки в цьому напрямку.

По-третє, легальні запуски БПЛА у комерційних цілях потребують дозволів, і технологія їх отримання розроблена компаніями-постачальниками БПЛА. Однак відповідальність за політ лежить на операторі-кермувальнику, що може створювати додаткові ризики.

Нарешті, ще не вирішені питання сертифікації, страхування та реєстрації літальних засобів, що обмежує їх ефективне використання.

Спектрональна зйомка - це метод фотозйомки, при якому одночасно отримуються фотографічні зображення об'єкта у різних діапазонах спектру електромагнітних хвиль. Наприклад, у сільському господарстві використання безпілотних літальних апаратів дозволяє отримати детальну модель місцевості з високою роздільною здатністю до 3 см. Знімки проводяться в видимому та інфрачервоному діапазонах, що надає повну інформацію про стан ґрунтів. Точність дозволяє контролювати посіви з точністю до 5 см, а широкий спектр отриманих даних дозволяє аналізувати проблеми поля під час всіх стадій вирощування культурних рослин та виявляти їх причини.

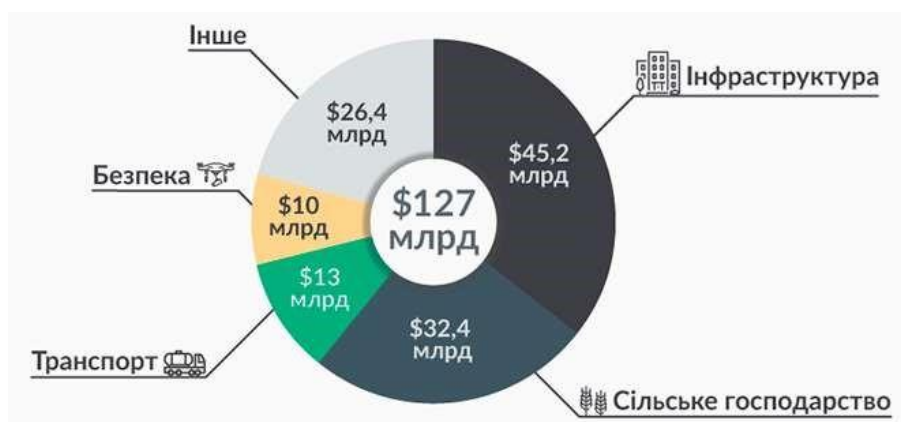


Рис. 2.2. Світовий ринок послуг із використанням БПЛА

Аерофотозйомка території - це комплекс процесів, що включає фотографування земельної поверхні з літака, формування аерофотознімків, фотосхем або фотопланів відзнятої місцевості. Ці матеріали використовуються для вирішення різних завдань у галузі лісового господарства та лісової промисловості. Планова зйомка проводиться з камерою, спрямованою вертикально вниз під прямим кутом до землі, що дає плоску картину на знімках, подібну до географічних карт. Перспективна зйомка відбувається під кутом до горизонту, дозволяючи отримати об'ємну картину знятої місцевості, що показує не лише горизонтальне розташування об'єктів, а й їхню форму.

Облік тварин з повітря за допомогою безпілотних літальних апаратів став економічно вигідним зниженням витрат на залучення пілотованої авіації. Цей метод дозволяє точно визначати кількість тварин у мисливських господарствах та виявляти місця їхньої концентрації.

Дистанційний моніторинг стану нафтопроводів та газопроводів в сучасний час вважається найефективнішим та економічно доцільним методом. Використання безпілотних літальних апаратів дозволяє отримувати в реальному часі якісні зображення, що сприяють виявленню нафтових розливів та несанкціонованої діяльності, такої як звалища, врізання або проведення робіт в зонах, що охороняються (рис. 2.3).



Рис. 2.3. Проведення робіт в зонах, що охороняються.

Аерофотознімки, здійснені з використанням безпілотних літальних апаратів, дозволяють проводити аналіз і оцінку технічного стану трубопроводів та оточуючого їх простору. Ця техніка використовується для безперервного моніторингу стану навколишнього середовища, включаючи забруднення земельної поверхні, водних акваторій та атмосфери, а також для контролю технічного стану об'єктів на довгих відстанях вздовж нафтових та газових трубопроводів.

Ці дані дозволяють оперативно виявляти зони небезпеки, такі як природні катастрофи або потенційні аварії, а також прогнозувати можливі ризики, такі як розриви трубопроводів чи геодинамічні деформації земної поверхні. Серед головних

завдань такого моніторингу можна виділити виявлення порушень технічного стану об'єктів, складання карт ґрунтів і зон підтоплень, а також дослідження сучасних екзогенних процесів, таких як сели, зсуви та обвали.

Аерофотозйомка та геодезія забезпечують матеріали, які можуть бути використані у різних сферах:

1. Ведення державного кадастру нерухомості та контроль містобудівної діяльності у населених пунктах.
2. Реагування на надзвичайні ситуації.
3. Контроль снігового та льодового покриву, межі льодоставу, прогнозування стоків річок та моніторинг місць розливів рік.
4. Оновлення топографічних карт.
5. Моніторинг різноманітних типів об'єктів.
6. Моніторинг стану сільськогосподарських угідь, включаючи цільове використання земель, оперативну оцінку стану та рівень деградації земель, а також прогнозування урожайності.
7. Створення географічних інформаційних систем (рис. 2.4).



Рис. 2.4. Створення географічних інформаційних систем

Моніторинг і розпізнавання рухомих об'єктів здійснюється в охоронних зонах як у денний, так і в нічний час.

Моніторинг лісових ресурсів включає оцінку рівня вирубки лісових масивів, ідентифікацію видів дерев, запобігання лісовим пожежам (виявлення сухих лісів,

тліючих торфовищ, виявлення невеликих джерел загорянь), оцінку збитків від лісових пожеж або природних катастроф, виявлення незаконних смітників та ідентифікацію осіб і номерів машин порушників. Дрони, обладнані інфрачервоними датчиками, використовуються для виявлення лісових пожеж на ранніх стадіях.

Моніторинг ремонтно-будівельних робіт. У багато разів прискорює та здешевлює роботу в порівнянні з традиційними методами. Він дозволяє оцінити ступінь готовності об'єкта; виявляти та аналізувати ушкодження, аварії; планувати ремонтні роботи; передбачати та моделювати природні впливи.

Безпілотна охорона підвищує рівень безпеки, забезпечуючи контроль за об'єктами та людьми на певних територіях. Безпілотники можуть ефективно патрулювати території підприємств та інших об'єктів, уникати несанкціонованих проникнень та виявляти потенційні загрози, такі як розриви нафтогазопроводів, лісові пожежі чи виявлення порушників.

Безпілотні прикордонники здатні спостерігати за кордоном за допомогою інфрачервоних та звичайних відеокамер, забезпечуючи деталізований огляд на відстані до 6 км та відстань у 50 км, що дозволяє точно виявляти порушення та ідентифікувати порушників.

Поліцейський спецназ використовує безпілотні мікродрони, які можна запускати прямо з рук, для здійснення розвідки та моніторингу ситуацій в обмежених умовах, таких як зони надзвичайних ситуацій чи конфліктні зони.

Системи безпіотної доставки вантажів, такі як ті, що впроваджені в Amazon або поштовою службою Швейцарії, дозволяють ефективно перевозити вантажі на значні відстані без необхідності залучення людських ресурсів.

Дрони також використовуються для надання допомоги в екстрених ситуаціях, наприклад, пошуку зниклих осіб чи виявлення потенційних небезпек. Такі дрони можуть бути оснащені термодатчиками або іншими датчиками для ефективного виявлення та локалізації осіб в ускладнених умовах [12].

Нарешті, безпілотні дрони використовуються як "офіціанти" для доставки страв та напоїв у закусочних, ресторанах та інших закладах громадського харчування, що дозволяє покращити ефективність обслуговування та забезпечити нові можливості у сфері гостинності.

Дрон, що роздає інтернет. Два гіганти ІТ-індустрії – компанії Google та Facebook, майже одночасно оголосили про свої плани створення мережі супутників і дронів, що утворюють інтернет-покриття на всій поверхні землі.

Дрони виявляються чудовими альтернативами традиційним методам рятування на воді, навіть тим, які базуються на рятувальниках, що стежать за пляжем з вежі. Успішні випробування мультикоптерів у Австралії підтвердили їхню ефективність у спостереженні за прибережною зоною та наданні допомоги потопаючим, включаючи кидання рятувального круга.

Аналізуючи вищезазначене, можна зробити висновок, що безпілотна авіація невдовзі може взяти лідерство порівняно з пілотованою. Це пояснюється тим, що використання безпілотників не потребує людських ресурсів та не ставить пілотів під загрозу, а також тим, що безпілотники мають широкий спектр застосування в оборонній сфері, галузі безпеки, охороні здоров'я, розвагах та багатьох інших галузях.

Для виконання льотних робіт з картографування території використовуються безпілотні літальні апарати (БПЛА) типу літака або гелікоптера. Вони можуть мати електричні або двигуни внутрішнього згоряння. Для ефективного виконання цих робіт до складу комплексу дистанційного моніторингу повинно входити:

1. Мінімум 2 літальні апарати, розташовані у захисних кейсах або чохлах, зручних для перенесення у польових умовах.
2. Приймально-передавальна антена з автоматичним пристроєм стеження й з'єднувальним кабелем.
3. Наземна станція керування з лептопом спеціального призначення, що відповідає технічним вимогам.

4. Зарядна станція або зарядний пристрій із комплектом акумуляторних батарей або запасом паливо-мастильних матеріалів.

5. Пускова установка для БПЛА літакового типу.

6. Комплект запасних частин та допоміжного обладнання для проведення ремонтних робіт у польових умовах.

7. Щогла або штатив для кріплення приймально-передавальної антени.

До додаткового оснащення комплексу дистанційного моніторингу рекомендується включати:

1. Малогабаритну бензоелектростанцію або додатковий автомобільний акумулятор.

2. Знімний HDD- або SSD-диск для збереження інформації.

3. Анемометр і конус-вітропоказчика.

4. GNSS-приймач супутникової навігації.

5. "Радіомаяки" з індивідуальним живленням.

6. Антенний кабель-подовжувач з підсилювачем сигналу.

7. Портативні флеш-карти пам'яті для запису фото (відео) інформації на борту БПЛА.

Це обладнання дозволяє здійснювати ефективні льотні роботи з картографування території у різних умовах [13].

Для проведення робіт із картографування території безпілотники можуть оснащуватися наступними засобами моніторингу (бортовими цільовими навантаженнями):

- відеокамерами інфрачервоного (ІЧ) діапазону (тепловізорами);
- відеокамерами оптичного (видимого) діапазону;
- ретрансляторами стільникового та радіозв'язку;
- комбінованими відеокамерами (оптичного й ІЧ діапазонів);
- радіолокаційними засобами;
- цифровими камерами оптичного діапазону.

Залежно від типу БПЛА цифрові камери можуть встановлюватися у носовій частині фюзеляжу, у крилі або під фюзеляжем. Об'єктиви цифрових камер можуть бути нерухомими чи мати один або два ступені ходу, а також змінну фокусну відстань (трансфокатор).

Для моніторингу місцевості переважно більшість відеокамер встановлюють турельного типу, що монтуються на гіростабілізованих платформах під фюзеляжем БПЛА й забезпечують круговий огляд нижньої півсфери у форматі 180°.

Для знімання окремих ділянок місцевості переважачими є цифрові камери, що встановлюються нерухомо у крилі або під фюзеляжем БПЛА.

Як вже зазначалося вище, для організації пошуку, у разі аварійної посадки, рекомендується оснащувати БПЛА пошуковими маячками й засобами відстеження. При цьому час автономної роботи маячків має бути не менше 8 годин [14, с. 779].

Обов'язково, до роботи із комплексами дистанційного моніторингу на базі БПЛА допускаються фахівці, придатні до роботи по пунктах 1, 3, 7, 13 «Переліку робіт, при виконанні яких проводяться попередні і періодичні медичні огляди (обстеження), наказ Міністерства охорони здоров'я і соціального розвитку України № 83 від 16 серпня 2004 року», які пройшли курс та навчання за затвердженими програмами по експлуатації БПЛА певного типу, і які засвоїли навички практичної роботи із комплексом й допущені до самостійної роботи наказом по організації.

Для виконання усіх польотних завдань, включаючи передпольотний контроль, зліт, політ по заданому маршруту, посадку, потрібно провести розрахунок у складі двох операторів.

При установці НСУ на автотранспорті до складу розрахунку може включатися й водій.

Також допускається експлуатація комплексу одним оператором (якщо це обумовлено у документації по експлуатації певного типу БПЛА).

Оператори повинні знати:

- специфіку застосування БПЛА для потреб у певній сфері діяльності;

- порядок й правила експлуатації БПЛА;
- основи літакокерування, аеродинаміки, метеорології;
- правила техніки безпеки при виконанні робіт з БПЛА;
- правила ведення радіозв'язку.

Оператори безпілотників повинні вільно користуватися ПК, вміти орієнтуватися у картографічних матеріалах [15, с. 2980].

При підборі команд, їх оснащують:

Автомобілем підвищеної прохідності із можливістю роботи оператора всередині салону. Транспортний засіб має бути обладнаний сонцезахисними шторками, системою обігріву й кондиціонування салону, інвертором (перетворювачем напруги живлення із 12 (24) на 220 вольт і потужністю не менше 1000 Вт, вузлами кріплення додаткового монітора. А у комплектацію автомобіля повинні входити:

- УКХ-радіостанція для зв'язку із наземними командами й іншими операторами БПЛА;
- УКХ-радіостанція для зв'язку із повітряними суднами;
- рідкокристалічний монітор з діагоналлю не менше 20 дюймів;
- компас або бусоль;
- супутниковий телефон;
- GNSS-приймач супутникової навігації (ГЛОНАСС / GPS);
- ручний акумуляторний ліхтар;
- бензопила із каністрою для ПММ об'ємом 5-10 л;
- лопата та сокира;
- ручна пила (ножівка столярна чи лучкова);
- вуглекислотний або порошковий вогнегасник;
- аптечка для надання домедичної допомоги;
- набір кемпінгових меблів;

- спальні мішки й надувні матраци, із розрахунку на кількість членів команди (при базуванні у польових умовах)
- кемпінговий каркасний тент (дуговий намет), висотою не менше 2 м.

При виконання робіт поблизу водних об'єктів й в районах із розвинутою водною мережею (великі річки, озера), додатково оснащується:

- надувним 2-хмісним човном із веслами;
- додатковим акумуляторним електричним ліхтарем;
- рятувальними жилетами.

При виконанні робіт у осінньо-зимовий період, команда має оснащатися:

- акумуляторним дрилем зі свердлом діаметрами 10-12 мм й довжиною 200-300 мм – для установки штиря ручної гумовою катапульти у замерзлий ґрунт;
- лопатою для розчищення майданчика від снігу;
- двома парами мисливських лиж або снігоступами.

Крім того, операторів БПЛА забезпечують спецвзуттям та спецодягом згідно із рекомендованим переліком в таблиці 2.1.

Таблиця 2.1

Перелік спеціального одягу для оператора, при виконанні задач пов'язаних із БПЛА

№ п/п	Найменування спецодягу	Термін несення, роки
1	Костюм бавовняний для захисту від загальних механічних забруднень й шкідливих біологічних факторів	3
2	Бейсболка	3
3	Черевики з натуральної шкіри (юхтові) з високими берцями	4
4	Напівкомбінезон утеплений	4
5	Перчатки шкіряні на байці	4

Продовження таблиці 2.1

6	Куртка демісезонна утеплена	4
7	Унти хутряні, або бахіли	6
8	Плащ (плащ-накидка)	Черговий
9	Чоботи гумові рибацькі	Чергові

Вимоги до проведення взльоту та посадки, порядок вибору місця взльоту БПЛА:

- вивчити місцевість у передбачуваній точці старту, при цьому необхідно урахувати: місце старту слід вибирати максимально високою щодо передбачуваного маршруту із мінімальним віддаленням від досліджуваних об'єктів, що дозволяє збільшити корисний час роботи і досягнення максимальних результатів;
- майданчик для запуску переважно обирають з трав'янистим покривом;
- розмір майданчика повинен бути не менше 50x50 м із умовою, щоб на прилеглій місцевості не знаходились об'єкти, що перешкоджають штатному режиму зльоту, посадки й пошуку БПЛА (щогли, річки, озера, яри, будови, вишки) в радіусі до 300 м;
- визначити напрямок й швидкість вітру (які біля поверхні землі й на робочій висоті можуть відрізнятися);
- визначити положення сторін світу;
- визначити напрямок запуску і переконатися у відсутності перешкод в цьому напрямку;
- визначити напрямок маршруту щодо НСУ й переконатися у відсутності перешкод у цьому напрямку для забезпечення прямої радіовидимості;
- для безпечного запуску й посадки БПЛА має бути повна відсутність перешкод: заводських труб, будівель, щогл, вишок, хмачосів висотою понад 100 м.

- переконатися у відсутності перешкод в зоні посадкової глісади. Слід врахувати, що на посадку літальний апарат заходить проти вітру, а точка захоплення координат є точкою відкриття парашута у режимі автоматичної посадки й аварійної посадки у випадку втрати зв'язку;

Ділянка посадки обирається поблизу точки старту із урахування можливості візуального контролю оператором заходу на посадку й самої посадки БПЛА.

Для приземлення БПЛА вибирають рівну ділянку місцевості розміром 50x50 м. На ділянці не має бути предметів, при приземленні на котрі може бути пошкоджений БПЛА, а саме:

- пнів і каменів, кущів і дерев, стовпів й ЛЕП, будівель і споруд, водотоків тощо.
- при посадці у ручному режимі точка випуску парашуту визначається оператором виходячи із поточних метеорологічних умов, розміру, місця й особливостей посадкового майданчика;
- посадка в напівавтоматичному (автоматичному) режимі виконується тільки при візуальному контролі БПЛА оператором.

2.2 Згорткові нейронні мережі

Згорткові нейронні мережі (ЗНМ, англ. convolutional neural networks, CNN) - це підтип глибоких нейронних мереж, які часто використовуються для розпізнавання об'єктів та класифікації зображень. Вони обробляють дані у вигляді тензорів, що дозволяє працювати з зображеннями у природному форматі. Кожне зображення має ширину, довжину та глибину, де остання визначається кодуванням, таким як RGB. Кожен канал кольору (червоний, зелений, синій) утворює окремий канал зображення. Мережі виробляють карти ознак зображення, розглядаючи кожне зображення як чотиривимірний масив даних. У порівнянні з іншими типами нейронних мереж, згорткові мережі мають меншу кількість параметрів, що робить їх ефективнішими для навчання, особливо в умовах великих обсягів даних. Також вони потребують менше даних для навчання, порівняно з традиційними нейронними мережами [4].

Ще однією відмінністю згорткових нейронних мереж є їх здатність працювати без попередньої обробки даних, оскільки вони працюють безпосередньо з вхідними зображеннями. Це дозволяє спростити структуру мережі, оскільки дані зображень мають упорядковану форму. Проте, навіть з усіма перевагами, застосування згорткових нейронних мереж може вимагати значних обчислювальних ресурсів, особливо при обробці великих обсягів високоякісних зображень. Тому потрібно розробляти мережі згорткових нейронів, які оптимізовані для ефективної роботи.

Крім розпізнавання та класифікації зображень, згорткові нейронні мережі також успішно використовуються в обробці природної мови (natural language processing, NLP). Вони демонструють високу ефективність у завданнях семантичного аналізу тексту, а також у моделюванні та класифікації речень.

Структура згорткової нейронної мережі включає вхідний шар, кілька прихованих шарів та вихідний шар. Приховані шари складаються зі згорткових шарів, шарів підсумування (пулінгу), нормалізації та повнозв'язаних шарів. Вони поєднані між собою з активаційними функціями. Основним елементом є згорткові шари, які застосовують операцію згортки до даних з попереднього шару [12].

Загальна структурна схема побудови згорткової нейронної мережі наведена на рис. 2.5.

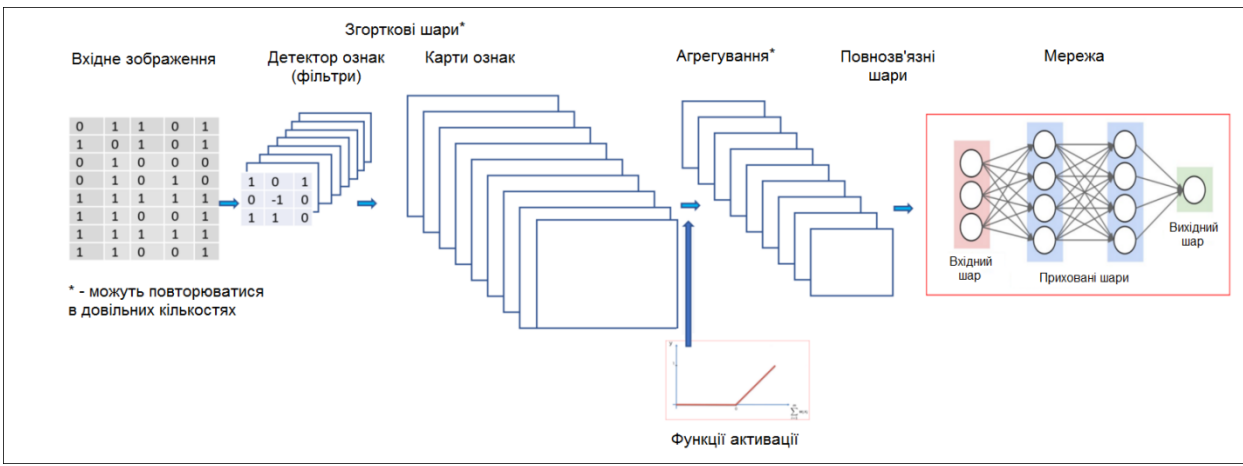


Рис. 2.5 – Загальна структурна схема побудови згорткової нейронної мережі

Згорткові нейронні мережі стали популярними в обробці зображень через їхню здатність ефективно вирішувати проблеми, які виникають у традиційних повнозв'язних мережах. Велика кількість параметрів у таких мережах потребує значного обсягу тренувальних даних, інакше може виникнути перенавчання, коли модель навчається на непоказових особливостях даних. Крім того, такі мережі не ефективно враховують локальні зв'язки в структурі зображення.

Згорткові нейронні мережі вирішують ці проблеми за допомогою кількох інноваційних ідей. Одна з них - використання локальних рецептивних полів, що дозволяють мережі виявляти прості елементи та ознаки, такі як кути або грані, на невеликих ділянках зображення. Ще одна ідея - спільні ваги, які дозволяють детекторам однаково реагувати на подібні ознаки на різних частинах зображення. Це допомагає зменшити чутливість мережі до змін положення об'єктів та інших змін у зображенні.

Операція субдискретизації є ще одним важливим елементом, який дозволяє зменшити розмірність даних і збільшити їхню стійкість до збурень. Ця операція зменшує розмірність даних, зберігаючи при цьому важливу інформацію про їх структуру та ознаки [11].

2.3 Згорткові шари (Convolutional Layer)

Згорткові шари (Convolutional Layers) є основним будівельним блоком згорткових нейронних мереж (Convolutional Neural Networks, CNN). Їх основне призначення полягає в автоматичному виділенні ознак (features) з вхідних зображень або інших даних, що дозволяє моделі краще розпізнавати складні патерни. На відміну від традиційних повнозв'язкових шарів, згорткові шари використовують концепцію локального підключення, де кожен нейрон підключений тільки до частини вхідних даних.

Основна операція, яка виконується в згортковому шарі, називається згортка (convolution). Вона полягає в застосуванні фільтра (ядра згортки) до вхідного зображення для створення карти ознак. Фільтр є матрицею менших розмірів, ніж

вхідне зображення, і пересувається по ньому з певним кроком (stride), обчислюючи скалярний добуток фільтра і частини вхідного зображення. Це можна представити у вигляді наступної формули:

$$(I * K)(i, j) = \sum_{m=-M}^M \sum_{n=-N}^N I(i - m, j - n) \cdot K(m, n) \quad (2.1)$$

де: I — початкове зображення, K — ядро згортки (фільтр), (i, j) — координати вихідного зображення, M, N — половина розміру ядра згортки по відповідним осям.

Основними параметрами згорткового шару є:

- Розмір ядра згортки (Kernel size): визначає розміри фільтра. Наприклад, часто використовуються ядра розмірами 3x3, 5x5, 7x7.
- Кількість фільтрів (Number of filters): визначає кількість карт ознак, які буде створено. Кожен фільтр вивчає різні характеристики вхідного зображення.
- Крок (Stride): визначає, наскільки пікселів фільтр пересувається по вхідному зображенню. Чим більший крок, тим менший вихідний розмір карти ознак.
- Заповнення (Padding): додає пікселі (зазвичай нульові) по краях вхідного зображення для збереження його розміру після згортки.

Переваги:

- Локальність зв'язків: дозволяє виявляти локальні патерни у вхідних даних.
- Зменшення кількості параметрів: на відміну від повнозв'язкових шарів, згорткові шари мають значно менше параметрів, що зменшує обчислювальну складність і ризик перенавчання.
- Інваріантність до трансляцій: здатність розпізнавати об'єкти незалежно від їхнього розташування у вхідному зображенні.

Недоліки:

- Втрата інформації про розташування: через зменшення розміру карти ознак може втрачатися інформація про точне розташування ознак.

- Обмеженість вивчення довгострокових залежностей: локальна природа згортки не дозволяє ефективно виявляти довгострокові залежності.

Згорткові шари є невід'ємною частиною сучасних нейронних мереж, які використовуються для розв'язання складних задач комп'ютерного зору. Їх здатність виділяти релевантні ознаки з вхідних даних робить їх надзвичайно ефективними для різних практичних застосувань, включаючи точне землеробство [16, с. 2261].

2.4 Карта ознак (Feature Map)

Карта ознак (Feature Map), також відома як активаційна карта (Activation Map), є результатом застосування згорткового фільтра до вхідного зображення в згортковому шарі. Кожна карта ознак містить інформацію про те, як добре певні ознаки (фічі) були виявлені фільтром у різних частинах вхідного зображення. Це допомагає нейронній мережі виділяти важливі патерни, що дозволяє краще розпізнавати та класифікувати об'єкти.

Процес створення карти ознак включає такі кроки:

1. Застосування фільтра: Фільтр (ядро згортки) проходить через вхідне зображення з певним кроком (stride), виконуючи операцію згортки на кожному кроці.
2. Обчислення значень: Кожен елемент карти ознак обчислюється як скалярний добуток значень пікселів вхідного зображення та значень фільтра.
3. Застосування активаційної функції: До результатів згортки застосовується активаційна функція, така як ReLU (Rectified Linear Unit), для введення нелінійності.

Наприклад, якщо вхідне зображення розміром 32x32 пікселів проходить через згортковий шар з фільтром розміром 3x3 і кроком 1, ми отримаємо карту ознак розміром 30x30 (без заповнення). Якщо ми використовуємо кілька фільтрів, кожен з них створить свою карту ознак.

Візуалізація карти ознак допомагає зрозуміти, які ознаки виділяються на різних рівнях нейронної мережі. На ранніх етапах згорткові шари зазвичай виявляють прості патерни, такі як краї та текстури, тоді як на пізніших етапах виділяються більш складні ознаки, такі як форми об'єктів.

Для візуалізації карти ознак можна використовувати різні техніки, наприклад:

- Відображення як зображення: Карти ознак можна відобразити як сірі або кольорові зображення, де інтенсивність пікселів вказує на силу активації.
- Прямий перегляд шарів: Багато фреймворків для глибокого навчання, такі як TensorFlow або PyTorch, надають інструменти для візуалізації внутрішніх шарів моделі під час тренування або передбачення [14, с. 779].
- Карти ознак відіграють критичну роль у процесі розпізнавання об'єктів, оскільки вони допомагають нейронній мережі фокусуватися на найважливіших частинах вхідного зображення. Завдяки цьому модель може:
 - Виділяти релевантні ознаки: Зосередження на найбільш важливих елементах зображення дозволяє моделі краще розпізнавати об'єкти.
 - Зменшувати вплив шуму: Завдяки активаційним функціям, таким як ReLU, мережа може ігнорувати незначні або нерелевантні ознаки.
 - Покращувати узагальнення: Використання різних фільтрів допомагає моделі розпізнавати об'єкти незалежно від їх позиції, масштабу або орієнтації.
- Карти ознак широко використовуються у різних задачах комп'ютерного зору:
 - Розпізнавання об'єктів: Карти ознак допомагають виділяти важливі риси об'єктів, що дозволяє їх точніше класифікувати.
 - Сегментація зображень: За допомогою карт ознак модель може виділяти області з різними об'єктами на зображенні.
 - Детекція об'єктів: Карти ознак використовуються для визначення місцезнаходження об'єктів на зображеннях.

Карти ознак є важливим компонентом згорткових нейронних мереж, оскільки вони допомагають моделі автоматично виділяти та аналізувати ключові ознаки вхідних даних. Розуміння принципів створення та інтерпретації карт ознак є важливим для розробки ефективних моделей машинного навчання, зокрема для задач точного землеробства, де важливо точно розпізнавати та аналізувати об'єкти на зображеннях, отриманих з борту БПЛА [9, с. 936].

2.5 Пулінг (Pooling)

Пулінг (Pooling) є важливим компонентом згорткових нейронних мереж (CNN), який виконує функцію зменшення розмірності простору ознак, зберігаючи при цьому найважливішу інформацію. Основна ідея пулінгу полягає в об'єднанні невеликих груп сусідніх значень у єдине значення, що дозволяє зменшити розмір карт ознак, зменшити обчислювальну складність та зменшити ризик перенавчання.

Існує кілька видів пулінгу, кожен з яких має свої особливості та застосування:

Максимальний пулінг (Max Pooling): обирає максимальне значення з кожної групи пікселів. Це допомагає зберегти найбільш виразні ознаки ігноруючи незначні деталі.

Формула для максимального пулінгу:

$$y = \max_{(i,j) \in R} x_{i,j} \quad (2.2)$$

де:

- y - вихідне значення після пулінгу
- R - регіон вхідної карти ознак
- $x_{i,j}$ - значення пікселів у цьому регіоні

Середній пулінг (Average Pooling): обчислює середнє значення з кожної групи пікселів. Це дозволяє згладжувати шуми і зменшувати розмірність, зберігаючи при цьому основні риси зображення.

Формула для середнього пулінгу:

$$y = \frac{1}{|R|} \sum_{(i,j) \in R} x_{i,j} \quad (2.3)$$

де:

- y - вихідне значення після пулінгу
- $|R|$ - кількість пікселів у регіоні R
- $x_{i,j}$ - значення пікселів у цьому регіоні

Глобальний пулінг (Global Pooling): обирає одне значення для всієї карти ознак, що дозволяє зменшити її розмір до 1x1. Це особливо корисно на останніх етапах нейронної мережі перед повнозв'язковими шарами.

Глобальний максимальний пулінг (Global Max Pooling):

$$y = \max_{i,j} x_{i,j} \quad (2.4)$$

де:

- y - вихідне значення (єдиний скаляр)
- $x_{i,j}$ - значення всіх пікселів на карті ознак

Глобальний середній пулінг (Global Average Pooling):

$$y = \frac{1}{N} \sum_{i,j} x_{i,j} \quad (2.5)$$

де:

- y - вихідне значення (єдиний скаляр)
- N - загальна кількість пікселів на карті ознак
- $x_{i,j}$ - значення всіх пікселів на карті ознак

У обох випадках результатом є одне значення, яке представляє всю карту ознак. Це робить глобальний пулінг особливо корисним для зменшення розмірності перед повністю зв'язаними шарами в нейронних мережах.

Основні параметри пулінгу включають:

- Розмір вікна (Window size): розміри регіону, на якому виконується пулінг, наприклад, 2x2 або 3x3.
- Крок (Stride): визначає, наскільки пікселів вікно пулінгу пересувається по карті ознак. Крок може бути меншим або більшим за розмір вікна, залежно від бажаного ступеня зменшення розмірності.
- Заповнення (Padding): у деяких випадках, особливо для середнього пулінгу, може використовуватися заповнення для збереження розмірів вхідного зображення.

Пулінг використовується в багатьох задачах комп'ютерного зору та обробки зображень:

- Розпізнавання об'єктів: допомагає зменшити розмір карт ознак, зберігаючи найважливіші риси об'єктів, що дозволяє ефективніше виконувати класифікацію.
- Сегментація зображень: зменшує розмір проміжних представлень, зберігаючи контекстну інформацію для подальшого об'єднання.
- Детекція об'єктів: допомагає зберегти масштаб і позиційні ознаки об'єктів, що полегшує їх виявлення на різних рівнях масштабування.

Переваги:

- Зменшення розмірності: дозволяє скоротити обчислювальні ресурси та зменшити обсяг даних для наступних шарів.
- Зменшення ризику перенавчання: усереднення або вибір максимуму допомагає усунути незначні коливання та шуми, знижуючи ризик перенавчання моделі.
- Інваріантність до трансляцій: пулінг додає стійкості до незначних змін в положенні об'єктів на зображенні.

Недоліки:

- Втрати інформації: через зменшення розмірності частина інформації про вхідне зображення може бути втрачена.
- Зміна масштабу: пулінг може спотворювати пропорції об'єктів, особливо при використанні великих вікон та кроків.

Пулінг є критично важливим елементом згорткових нейронних мереж, що дозволяє ефективно зменшувати розмірність даних, зберігаючи при цьому ключові ознаки зображення. Це сприяє підвищенню ефективності та продуктивності моделей машинного навчання, особливо в задачах точного землеробства, де важливо точно аналізувати великі обсяги візуальних даних, отриманих з борту БПЛА [11].

2.6 Повнозв'язкові шари (Fully Connected Layer)

Повнозв'язкові шари (Fully Connected Layers), також відомі як щільні шари (Dense Layers), є важливим компонентом нейронних мереж, де кожен нейрон поточного шару пов'язаний з кожним нейроном попереднього шару. Це дозволяє моделі вивчати складні залежності між ознаками, виділеними згортковими шарами, та виконувати кінцеву класифікацію або регресію [8, с 324].

Повнозв'язковий шар складається з набору нейронів, де кожен нейрон приймає вхідні дані з усіх нейронів попереднього шару та обчислює свою активацію за допомогою наступної формули:

$$y_j = f \left(\sum_{i=1}^n w_{ij} x_i + b_j \right) \quad (2.6)$$

де:

- y_j - вихідне значення нейрона j
- f - активаційна функція (наприклад, ReLU, sigmoid, tanh)
- n - кількість нейронів у попередньому шарі
- w_{ij} - вага зв'язку між нейроном i попереднього шару та нейроном j поточного шару
- x_i - вхідне значення від нейрона i попереднього шару
- b_j - зміщення (bias) нейрона j

Для повнозв'язкових шарів часто використовуються різні активаційні функції, які допомагають моделі вивчати нелінійні залежності:

Для повнозв'язкових шарів часто використовуються різні активаційні функції, які допомагають моделі вивчати нелінійні залежності:

ReLU (Rectified Linear Unit):

$$f(x) = \max(0, x) \quad (2.7)$$

Переваги: простота, ефективність обчислень, зменшення проблеми зникнення градієнта.

Sigmoid:

$$f(x) = \frac{1}{1+e^{-x}} \quad (2.8)$$

Переваги: значення в діапазоні (0, 1), підходить для задач бінарної класифікації.

Недоліки: проблема зникнення градієнта.

Tanh (Hyperbolic Tangent):

$$f(x) = \tanh(x) = \frac{2}{1+e^{-2x}} - 1 \quad (2.9)$$

Переваги: значення в діапазоні (-1, 1), допомагає зцентрувати дані.

Недоліки: проблема зникнення градієнта.

Повнозв'язкові шари використовуються для:

- Класифікації: перетворення вихідних ознак на ймовірності класів у задачах класифікації.
- Регресії: вивчення безперервних залежностей між ознаками та цільовими значеннями.
- Об'єднання ознак: інтеграція та узагальнення ознак, виділених попередніми шарами, для кращого розуміння даних.

Повнозв'язкові шари навчаються за допомогою алгоритмів зворотного поширення помилки (backpropagation) та оптимізаційних методів, таких як SGD (Stochastic Gradient Descent) або Adam. Процес навчання включає корекцію ваг та зміщень на основі градієнтів помилок, що обчислюються для кожного шару.

Переваги:

- Гнучкість: здатність вивчати складні та нелінійні залежності.
- Інтеграція ознак: ефективно об'єднання різних ознак для кінцевої класифікації або регресії.

Недоліки:

- Велика кількість параметрів: значна кількість ваг і зміщень, особливо для великих шарів, що може призвести до перенавчання.

- Обчислювальна складність: великі ресурси для навчання та прогнозування.
- Повнозв'язкові шари широко використовуються у різних задачах:
- Класифікація зображень: кінцеві шари CNN зазвичай повнозв'язкові для класифікації об'єктів на зображенні.
 - Нейронні мережі для природної мови: обробка текстів та класифікація документів.
 - Прогнозування часового ряду: регресія для прогнозування майбутніх значень на основі історичних даних.

Повнозв'язкові шари є важливою частиною нейронних мереж, що дозволяють виконувати кінцеву класифікацію або регресію на основі ознак, виділених згортковими шарами. Вони забезпечують гнучкість та потужність моделі, дозволяючи їй вивчати складні залежності між даними. Це особливо важливо для задач точного землеробства, де повнозв'язкові шари можуть допомогти аналізувати та інтерпретувати великі обсяги даних, зібраних з борту БПЛА [17, с 4740].

2.7 Пряме поширення помилки (backpropagation)

Пряме поширення помилки (backpropagation) є ключовим алгоритмом для навчання нейронних мереж. Це метод обчислення градієнтів функції втрат щодо ваг мережі, що дозволяє коригувати ваги таким чином, щоб мінімізувати втрати. Основна ідея полягає у поширенні помилки від виходу мережі до її входів, поступово оновлюючи ваги на основі обчислених градієнтів.

Процес прямого поширення помилки включає кілька етапів:

1. Пряме проходження (Forward Pass):

- Вхідні дані проходять через мережу, і на кожному шарі обчислюються виходи нейронів.
- Вихід мережі порівнюється з фактичними значеннями, і обчислюється значення функції втрат.

2. Зворотне проходження (Backward Pass):

- Починаючи з вихідного шару, обчислюється помилка для кожного нейрона на основі різниці між передбаченим та фактичним значеннями.
- Обчислюються градієнти функції втрат щодо ваг кожного нейрона за допомогою правила ланцюга.

3. Оновлення ваг (Weight Update):

- Ваги коригуються на основі обчислених градієнтів з використанням алгоритму оптимізації, такого як SGD (Stochastic Gradient Descent) або Adam.
- Крок оновлення ваг визначається швидкістю навчання (learning rate).

Правило ланцюга, або правило диференціації, є основою для обчислення градієнтів у backpropagation. Воно дозволяє обчислити похідні складених функцій.

Активаційні функції грають важливу роль у backpropagation, оскільки вони вводять нелінійність, що дозволяє мережі вивчати складні залежності.

Після обчислення градієнтів ваги оновлюються з використанням алгоритму оптимізації. Один з найбільш поширених методів — градієнтний спуск (Gradient Descent):

Інші методи оптимізації, такі як Adam або RMSprop, використовують додаткові параметри та адаптивні швидкості навчання для прискорення процесу навчання та покращення стабільності.

Пряме поширення помилки є універсальним алгоритмом, який застосовується у різних типах нейронних мереж та задачах:

- Згорткові нейронні мережі (CNN): навчання для задач класифікації та сегментації зображень.
- Рекурентні нейронні мережі (RNN): навчання для обробки послідовностей, таких як текст або часові ряди.
- Автоенкодери (Autoencoders): навчання для задач зменшення розмірності та відновлення даних.

Пряме поширення помилки є основним алгоритмом для навчання нейронних мереж, що дозволяє ефективно обчислювати градієнти та оновлювати ваги для мінімізації функції втрат. Розуміння та правильне застосування цього алгоритму є критичним для розробки ефективних моделей машинного навчання, зокрема в задачах точного землеробства, де необхідно аналізувати та інтерпретувати великі обсяги даних, зібраних з борту БПЛА [18].

2.8 Зворотне поширення помилки (backpropagation)

Зворотне поширення помилки (backpropagation) - це алгоритм навчання, який використовується для оптимізації параметрів (ваг і зміщень) нейронної мережі. Його основна мета полягає в мінімізації функції втрат, яка визначає, наскільки добре мережа виконує своє завдання (наприклад, класифікація або регресія).

Алгоритм зворотного поширення складається з двох основних етапів:

Пряме поширення (Forward propagation): Вхідні дані проходять через мережу від вхідного шару до вихідного, при цьому на кожному шарі виконуються обчислення згідно з вагами та активаційними функціями. Результатом є вихід мережі, який порівнюється з правильними відповідями для обчислення функції втрат.

Зворотне поширення (Backward propagation): Починаючи з вихідного шару, обчислюються градієнти функції втрат щодо кожного параметра мережі. Ці градієнти показують, як зміна кожного параметра впливає на функцію втрат. Потім градієнти поширюються назад через мережу, використовуючи правило ланцюга, щоб оновити ваги та зміщення кожного шару [19].

Ключовим елементом зворотного поширення є правило ланцюга, яке дозволяє обчислювати градієнти складних функцій, розбиваючи їх на простіші складові. У контексті нейронних мереж, правило ланцюга використовується для обчислення градієнтів функції втрат щодо ваг і зміщень кожного шару, враховуючи внесок усіх попередніх шарів [20].

Після обчислення градієнтів, параметри мережі оновлюються за допомогою оптимізаційного алгоритму, такого як стохастичний градієнтний спуск (SGD) або Adam. Ці алгоритми використовують градієнти для визначення напрямку та величини змін параметрів, з метою зменшення функції втрат.

Процес прямого та зворотного поширення повторюється багаторазово для набору навчальних даних, доки функція втрат не досягне мінімуму або не буде виконано задану кількість ітерацій.

Зворотне поширення є основою для навчання більшості сучасних нейронних мереж. Воно дозволяє ефективно налаштовувати велику кількість параметрів, що робить можливим створення складних моделей, здатних вирішувати різноманітні завдання, включаючи розпізнавання зображень, обробку природної мови та багато інших [21, с. 1571].

У точному землеробстві зворотне поширення використовується для навчання нейронних мереж, які аналізують дані з БПЛА, супутників та інших джерел. Ці мережі можуть виявляти шкідників, хвороби рослин, оцінювати врожайність, оптимізувати використання ресурсів та багато іншого, що сприяє підвищенню ефективності та сталості сільськогосподарського виробництва.

2.9 Методи оптимізації. Введення в область виявлення об'єктів

Методи оптимізації відіграють ключову роль у навчанні нейронних мереж. Вони відповідають за пошук оптимальних значень параметрів моделі (ваг і зміщень), щоб мінімізувати функцію втрат. Ефективна оптимізація дозволяє моделі швидше та точніше навчатися на даних, що призводить до кращої продуктивності в задачах, таких як виявлення об'єктів.

Існує багато різних методів оптимізації, кожен з яких має свої переваги та недоліки. Деякі з найпопулярніших методів включають:

- Стохастичний градієнтний спуск (SGD): Це базовий метод, який оновлює параметри на основі градієнта функції втрат, обчисленого для випадкової

вибірки даних. SGD є простим у реалізації, але може бути повільним у збіжності.

- Adam: Це адаптивний метод, який поєднує ідеї SGD з моментами та адаптивними швидкостями навчання. Adam часто забезпечує швидшу збіжність і кращу продуктивність, ніж SGD.
- RMSprop: Це ще один адаптивний метод, який використовує ковзну середню квадратів градієнтів для нормалізації швидкостей навчання. RMSprop також може забезпечити швидшу збіжність, ніж SGD.

Вибір методу оптимізації залежить від конкретної задачі та архітектури нейронної мережі. Деякі методи можуть бути більш ефективними для певних типів моделей або наборів даних. Часто найкращий підхід - експериментувати з різними методами та вибирати той, який дає найкращі результати.

Виявлення об'єктів - це важлива задача комп'ютерного зору, яка полягає у знаходженні та класифікації об'єктів на зображенні або відео. Це складне завдання, яке вимагає від моделі здатності розпізнавати об'єкти різних форм, розмірів та орієнтацій [22].

Методи оптимізації відіграють важливу роль у навчанні моделей виявлення об'єктів. Вони допомагають моделі навчитися ефективно виявляти об'єкти, мінімізуючи помилки та підвищуючи точність передбачень. Вибір правильного методу оптимізації може суттєво вплинути на продуктивність моделі виявлення об'єктів.

Дослідження в галузі оптимізації та виявлення об'єктів активно тривають. Розробляються нові методи оптимізації, які можуть забезпечити ще швидшу збіжність та кращу продуктивність моделей. Також з'являються нові архітектури нейронних мереж, спеціально розроблені для виявлення об'єктів, які можуть досягати ще вищої точності та ефективності.

Висновки до розділу 2

Згорткові шари відіграють ключову роль у виявленні та виділенні ознак із зображень, що є особливо важливим для аналізу даних з БПЛА в точному землеробстві.

Пулінг допомагає зменшити розмірність даних та зробити модель більш стійкою до незначних змін у зображенні.

Повнозв'язкові шари дозволяють моделі робити складні висновки та класифікувати об'єкти на зображеннях.

Пряме поширення використовується для отримання прогнозів моделі, а зворотне поширення помилки дозволяє їй навчатися та покращувати свої результати.

Методи оптимізації є важливим інструментом для налаштування параметрів моделі та досягнення найкращої продуктивності.

Отримані знання є фундаментом для розуміння того, як нейронні мережі застосовуються в точному землеробстві для вирішення таких завдань, як виявлення шкідників, оцінка врожайності та оптимізація використання ресурсів. Розуміння цих концепцій дозволяє оцінити потенціал нейронних мереж у підвищенні ефективності та сталості сільськогосподарського виробництва.

3 Практична частина

3.1 Аналіз та порівняння схожих наборів даних

YOLO, що означає "You Only Look Once" (Ти дивишся лише один раз), є однією з найпопулярніших архітектур для виявлення об'єктів у реальному часі. Вперше запропонована в 2015 році Джозефом Редмоном, ця модель здобула велику увагу дослідників у галузі комп'ютерного зору. YOLO була першою моделлю, яка була успішно реалізована на мобільних пристроях.

Принцип роботи YOLO полягає у тому, що вона ділить зображення на сітку рівномірного розміщення. Залежно від роздільної здатності, ця сітка має певну кількість клітинок. Потім модель прогнозує кількість та положення граничних рамок (bounding boxes) для кожної клітинки сітки. Кожна рамка містить інформацію про її розміри, положення відносно комірки сітки, а також показник достовірності, що вказує на ймовірність того, що в рамці міститься об'єкт. Чим вище цей показник, тим більша рамка навколо об'єкту (рис 3.1) [23, с. 2261].

Принцип "дивиись лише один раз" полягає у тому, що для передбачення об'єктів моделі потрібен лише один прохід через нейронну мережу. Це робить YOLO дуже ефективною для виявлення об'єктів у реальному часі.

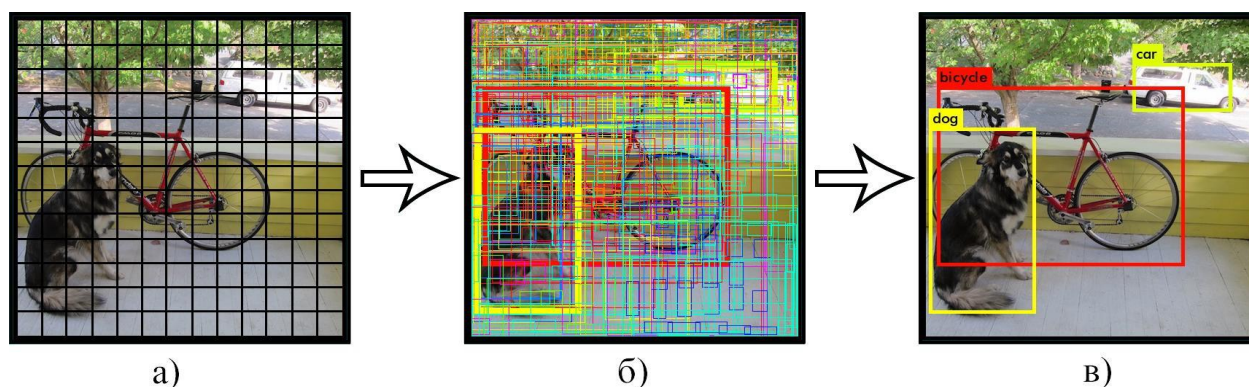


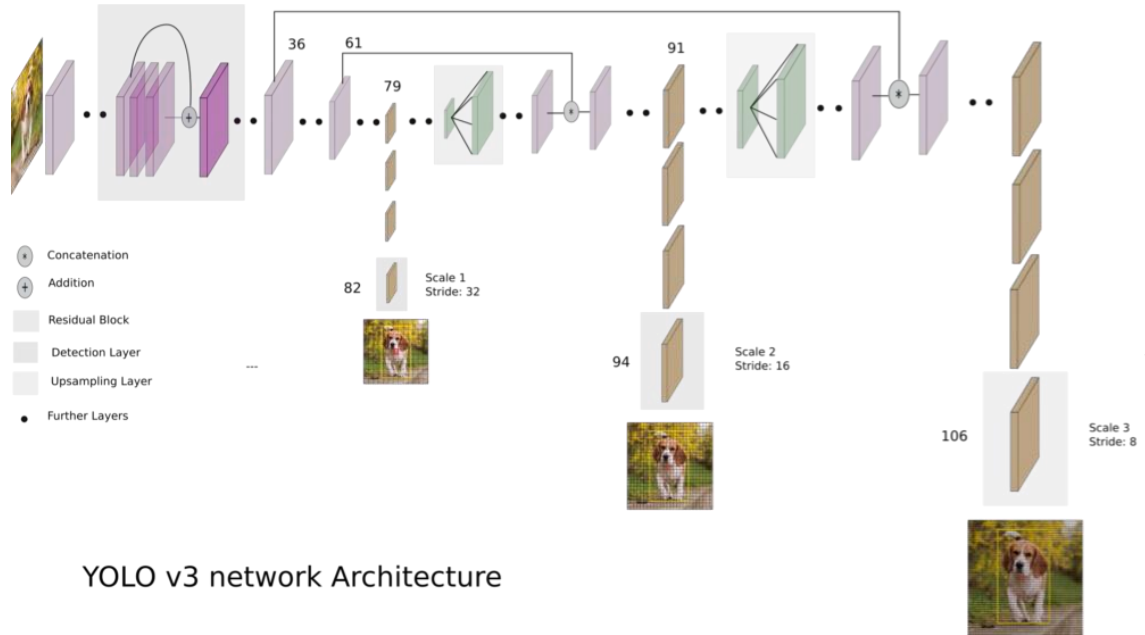
Рис. 3.1 – Принцип роботи архітектури YOLO

Після прогнозування граничних рамок для об'єктів YOLO переходить до етапу класифікації, де для кожної рамки передбачається клас об'єкта. Показники

достовірності та значення прогнозування класу об'єднюються в один загальний бал, що вказує на ймовірність того, що рамка містить конкретний тип об'єкта.

YOLO може передбачати кілька граничних рамок для однієї клітинки сітки. За допомогою функції втрат ЗНМ (Zero-Normalized Mean-Square Error) модель знаходить єдине правильне значення та відкидає рамки, які мають менше значення показника достовірності. Функція втрат включає в себе:

- втрату класифікації;
- втрату локалізації (помилки між передбаченою та істинною граничною рамкою);
- втрату довіри (впевненість моделі у присутності об'єкта в рамці).



YOLO v3 network Architecture

Рис. 3.2 – Архітектура YOLOv3

Існують поліпшені версії моделі YOLO, такі як YOLOv3, YOLOv3-tiny, YOLOv4 та найсучасніша YOLOv5. Архітектура YOLOv3 використовує "ознакові піраміди" (Feature Pyramid Networks - FPN), що складаються з більшої кількості згорткових шарів і краще розпізнають невеликі об'єкти. Вона здійснює прогнози в трьох різних масштабах, використовуючи сітки з різним кроком. Починаючи з останнього шару карти ознак, прогнозування проходить у зворотному порядку, зі збільшенням роздільної здатності карти ознак.

У YOLOv3 використовується новий 53-шаровий блок виділення ознак Darknet-53. Цей блок має менше операцій з плаваючою комою, ніж схожі блоки інших моделей, проте досягає аналогічної точності класифікації, працюючи у два рази швидше. Мережа постійно застосовує згорткові фільтри до об'єднаної карти ознак, щоб прогнозувати різні характеристики об'єктів.

Для порівняння YOLO з іншими сучасними архітектурами подана таблиця 3.1, де представлені моделі, які використовуються в реальному часі. З цієї таблиці можна зробити висновок, що моделі YOLO мають оптимальне співвідношення між швидкістю та точністю для вбудованих пристроїв.

Таблиця 3.1

Порівняння YOLO з іншими архітектурами

Архітектура	Переваги	Недоліки	Застосування
YOLO (You Only Look Once)	Швидкість, точність в реальному часі, здатність виявляти кілька об'єктів одночасно	Менша точність для малих об'єктів, чутливість до змін в масштабі та співвідношенні сторін	Виявлення об'єктів в реальному часі, автономні транспортні засоби, системи безпеки
R-CNN (Regions with CNN features)	Висока точність виявлення об'єктів, особливо для малих об'єктів	Низька швидкість, складність обчислень	Виявлення об'єктів на статичних зображеннях, медична діагностика
Fast R-CNN	Покращена швидкість порівняно з R-CNN, збереження високої точності	Все ще відносно повільний порівняно з YOLO	Виявлення об'єктів на статичних зображеннях, коли швидкість не є критичною
Faster R-CNN	Ще більш швидкий ніж Fast R-CNN, завдяки використанню Region Proposal Network (RPN)	Складність архітектури	Виявлення об'єктів на статичних зображеннях, коли потрібна висока точність і прийнятна швидкість

Продовження таблиці 3.1

SSD (Single Shot MultiBox Detector)	Висока швидкість та точність, здатність працювати з різними масштабами об'єктів	Менша точність для малих об'єктів порівняно з Faster R-CNN	Виявлення об'єктів в реальному часі, мобільні додатки
-------------------------------------	---	--	---

Ці дані були отримані за допомогою однакових графічних процесорів і використані для навчання на великомасштабному наборі даних Microsoft COCO, призначеному для виявлення та сегментації. Основні параметри MS COCO включають:

- Понад 330 000 зображень, з яких більше ніж 200 000 мають маркування.
- Понад 1,5 мільйона екземплярів об'єктів.
- 80 категорій об'єктів та 91 категорію речей.
- До 5 підписів на одному зображенні.
- Понад 250 000 людей з маркуванням точок захоплення.

Для даного завдання, яке передбачає розпізнавання об'єктів у інфрачервоному діапазоні, потрібно використовувати теплові зображення. У вільному доступі є набір даних від компанії FLIR, який буде розглянуто докладніше в розділі 3. Проте, ефективність навчання глибоких згорткових нейронних мереж (ГНН), яку пропонує компанія FLIR, не досягає достатньої для застосування в реальних прикладних задачах, тому маємо завдання покращити показники навчання до більш прийнятних рівнів.

3.2 Створення власного набору даних

У 2020 році в галузі виявлення об'єктів спостерігалася жорстка конкуренція між різними моделями. Однак YOLOv4, за словами його розробників, відзначився як модель, що об'єднала всі передові досягнення науки в цій області. Згідно з відомостями, YOLOv4 досягає точності на рівні 43,5% mAP (65,7% mAP50) для набору даних MS COCO, при цьому забезпечуючи швидкість близько 65 кадрів в секунду на графічному процесорі NVIDIA Tesla V100.

Важливою характеристикою у виявленні об'єктів є не лише висока точність, але й плавність роботи на пристроях з обмеженими ресурсами та можливість обробки відео в реальному часі без використання високопродуктивного обладнання [24].

Покращення процесу навчання можна умовно розділити на дві групи: ті, які підвищують точність моделі, але не впливають на її швидкість (наприклад, збільшення обсягу даних, функції втрат, управління незбалансованістю класів) та ті, що впливають на швидкість роботи моделі (наприклад, збільшення рецептивного поля, використання візуальної уваги, інтеграція модуля пропуску з'єднань та FPN, використання методу придушення немаксимумів (NMS)).

YOLOv4 використовує модуль DenseBlock для об'єднання шарів один з одним, різнячи їхні розміри карти ознак. Dense layer складається з кількох згорткових шарів та прошарків пакетної нормалізації, що дозволяє подальше розширення функціональності моделі (рис 3.3).

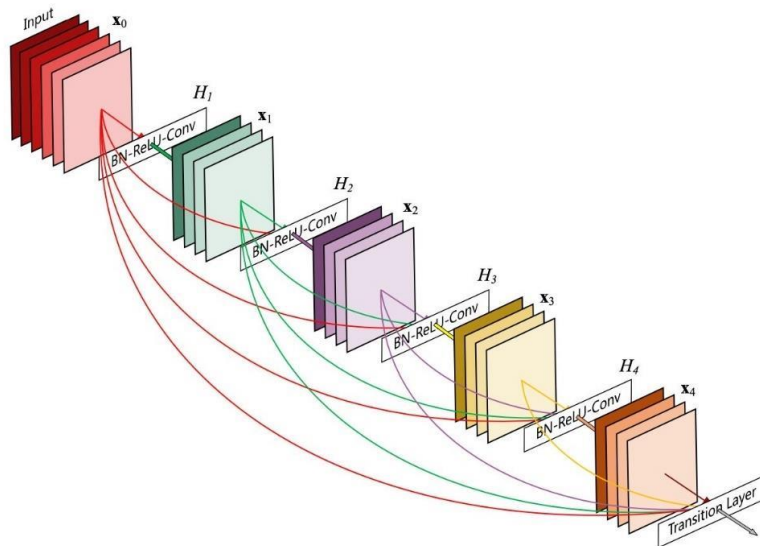


Рис. 3.3 – П'ятишаровий модуль DenseBlock

Модуль H_i в YOLOv4 приймає на вхід оригінальний вхідний сигнал, а також вихідні дані всіх попередніх шарів (x_0, x_1, \dots, x_{i-1}), що дозволяє зберегти більше інформації. Кожен попередній шар H_i видає чотири карти ознак, що призводить до збільшення кількості карти ознак на кожному шарі на чотири, що призводить до прискорення зростання кількості карти ознак.

Наступним модулем базової частини YOLOv4 є CSPNet. Цей модуль відокремлює вхідні карти ознак DenseBlock на дві частини. Перша частина (x_0') обходить DenseBlock і стає частиною вхідних даних для наступного перехідного шару. Друга частина (x_0'') проходить через DenseBlock. Завдяки цьому модулю вхід розділяється на дві частини, і тільки одна з них проходить через DenseBlock. Це зменшує обчислювальну складність моделі (рис 3.4).

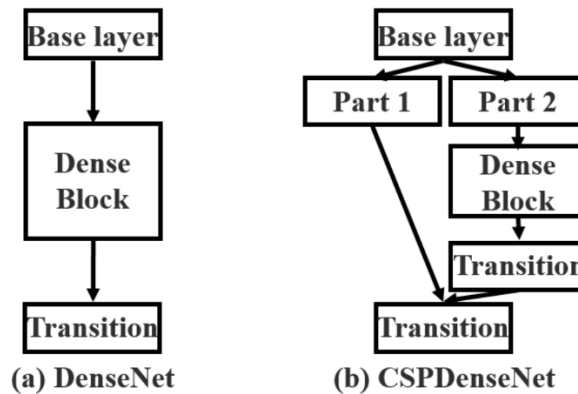


Рис. 3.4 – Різні стратегії об'єднання карт функцій в CSPNet.

У YOLOv4 використовується нова базова структура, яка комбінує технології об'єднання з'єднань CSP і архітектуру Darknet-53, що використовувалася в третій версії YOLO. Архітектура Darknet-53, зображена на рис. 3.5, використовується як основа базової частини моделі.

	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1×	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
2×	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8×	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
8×	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
4×	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

Рис. 3.5 – Архітектура Darknet-53

Нова базова структура, відома як CSPDarknet53, демонструє більш високу точність виявлення об'єктів у порівнянні зі структурами на основі ResNet, хоча останні краще справляються з класифікацією об'єктів. Проте точність класифікації CSPDarknet53 може бути покращена за допомогою методів, таких як Mish та інші.

3.3 Опис власного набору даних

Для створення автоматизованого розпізнавання об'єктів на основі методів машинного навчання, спочатку потрібно підготувати набір даних, навчити модель на цих даних і потім використовувати навчену модель для розпізнавання об'єктів у нових зображеннях.

Опис власного набору даних - це детальний опис характеристик і особливостей набору даних, які використовуються для навчання або тестування моделей машинного навчання. Цей опис має мету допомогти іншим користувачам або дослідникам зрозуміти, які дані включені в набір, як вони були зібрані та організовані.

Приклад коду для створення власного набору даних та навчання моделі на основі бібліотеки TensorFlow та Keras:

```
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers

# Створення власного набору даних
train_images = ... # Завантажте зображення для навчання
train_labels = ... # Відповідні мітки для зображень
test_images = ... # Завантажте зображення для тестування
test_labels = ... # Відповідні мітки для зображень

# Побудова моделі
model = keras.Sequential([
    layers.Conv2D(32, (3, 3), activation='relu'),
    layers.MaxPooling2D((2, 2)),
    layers.Conv2D(64, (3, 3), activation='relu'),
    layers.MaxPooling2D((2, 2)),
    layers.Flatten(),
    layers.Dense(64, activation='relu'),
    layers.Dense(10) # Кількість класів для класифікації
])

# Компіляція моделі
model.compile(optimizer='adam',
```



```

    loss=tf.keras.losses.SparseCategoricalCrossentropy(from_logits=True),
    metrics=['accuracy'])
# Навчання моделі
model.fit(train_images, train_labels, epochs=10, validation_data=(test_images,
test_labels))
# Оцінка моделі
test_loss, test_acc = model.evaluate(test_images, test_labels)
print('Test accuracy:', test_acc)

```

Цей код демонструє простий приклад навчання нейронної мережі для класифікації зображень. Якщо вам потрібно розпізнавати об'єкти на зображеннях, можливо, вам знадобиться модель для виявлення об'єктів, наприклад, модель YOLO або SSD.

3.4 Установка нейронної мережі YOLOv4

Установка нейронної мережі YOLOv4 може включати в себе кілька кроків, включаючи налаштування середовища, завантаження коду моделі, встановлення необхідних бібліотек та залежностей, а також навчання або використання попередньо навченої моделі. Загальний опис кроків, які можна виконати для установки нейронної мережі YOLOv4:

1. Підготовка середовища: Переконайтеся, що на вашому комп'ютері встановлено Python та інші необхідні інструменти для розробки. Рекомендується використовувати середовище віртуального оточення для управління пакетами та залежностями.

2. Клонування репозиторію Darknet: Darknet - це фреймворк, на якому базується YOLO. Клонуйте репозиторій Darknet з GitHub або іншого джерела.

3. Конфігурація Darknet: Налаштуйте Darknet для використання YOLOv4. Це може включати зміни конфігураційних файлів, залежно від вашого використання та потреб.

4. Завантаження ваг моделі: Завантажте попередньо навчені ваги моделі YOLOv4. Ці ваги можна отримати з офіційного джерела або з інших джерел.

5. Компіляція Darknet: Скомпілюйте Darknet з підтримкою YOLOv4. Це може зайняти деякий час, але після цього ви зможете використовувати Darknet для навчання або тестування моделі YOLOv4.

6. Тестування або навчання моделі: Після успішної установки ви можете тестувати або навчати модель YOLOv4 на ваших даних. Для цього ви можете використовувати інтерфейс командного рядка Darknet або розробляти власний код для використання моделі.

7. Оцінка результатів: Після завершення навчання або тестування оцініть результати за допомогою метрик відповідності, точності тощо.

Згаданий процес установки може бути складним і залежить від вашого досвіду у роботі з Python, фреймворками для глибокого навчання та іншими інструментами. Тому важливо уважно слідкувати за інструкціями та документацією, щоб успішно встановити нейронну мережу YOLOv4.

Для виділення ознак в YOLO використовується підхід схожий на FPN, який був розглянутий у першому розділі, для прогнозування виявлення об'єктів у різних масштабах.

Для усунення обмеження розміру вхідного зображення та детектування об'єктів різного масштабу використовується технологія Spatial Pyramid Pooling (SPP) [17]. Вона замінює останній шар агрегування шаром об'єднання «просторової піраміди» (рис.3.6). Для збереження вихідного просторового виміру в YOLO, SPP трохи змінена. Розмір ковзаючого вікна змінений на непарні значення на відміну від парних у стандартній версії SPP. Просторовий вимір залишається незмінним. На виході, карти ознак, із різними розмірами ядра, конкатенуються разом. На рис. 3.7, нижче, показано, як SPP інтегрована в YOLO.

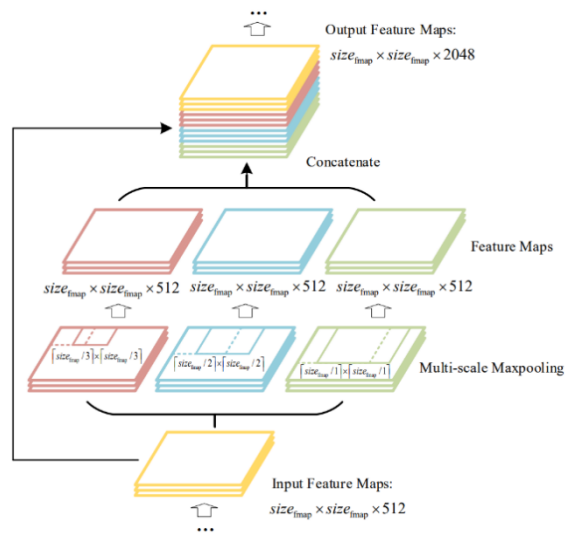


Рис. 3.6 – Покращена технологія об'єднання SPP для YOLO.

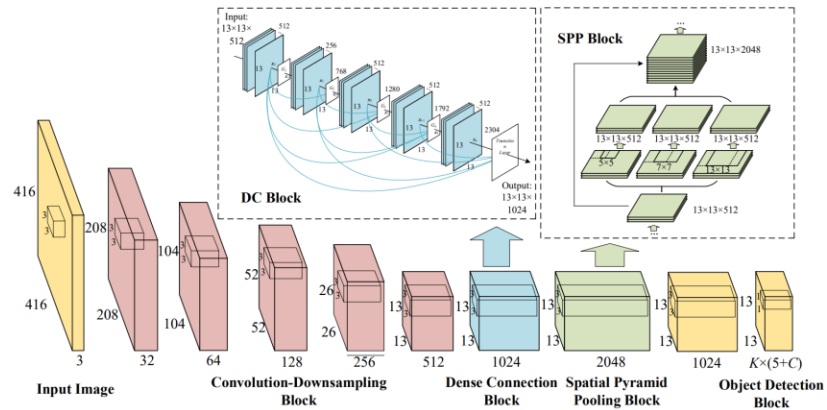


Рис. 3.7 – SPP інтегрована в YOLO

Взаємозв'язки між шарами в сучасних моделях розпізнавання стають складніше й більш змішаними. Якщо раніше кожен шар приймав вхідні дані з попереднього шару, то зараз кожен шар пов'язаний з усіма попередніми, як наприклад, в DenseBlock на рис.3.3. Відповідно по мірі просування вправо локалізована інформація, яка може знадобитися для точного передбачення класу, може бути втрачена і виникає питання транспортування інформації між шарами. Нижче на рис.3.8 наведена технологія Path Aggregation Network (PAN).

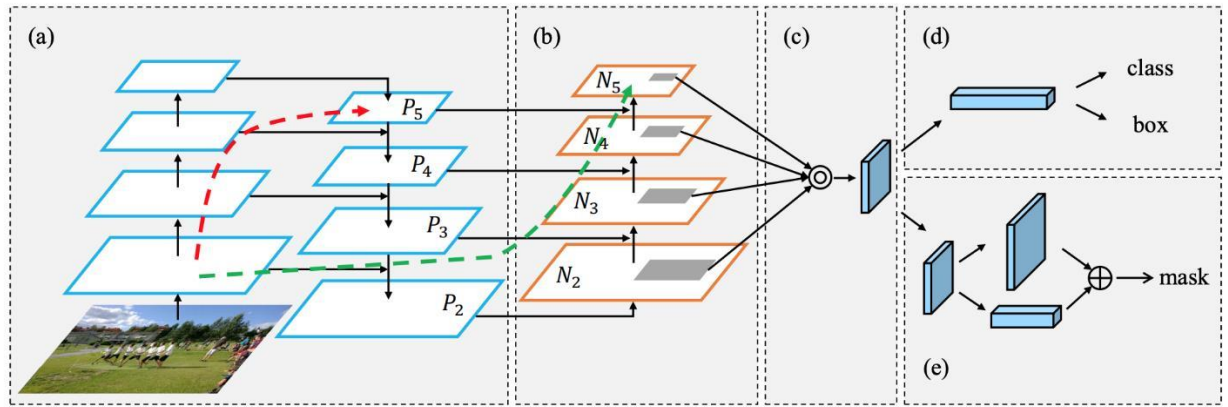


Рис. 3.8 – Path Aggregation Network

У FPN локалізована просторова інформація переміщується вгору за червоною стрілкою. Хоча це не ясно показано на діаграмі, червоний шлях проходить приблизно через 100+ шарів. PAN пропонує скорочений (зелений) шлях, який займає орієнтовно 10 шарів, щоби перейти до верхнього шару N_5 [25].

Однак, замість того, щоби підсумувати сусідні шари, карти ознак в YOLOv4 конкатенуються (рис.3.9) [14].

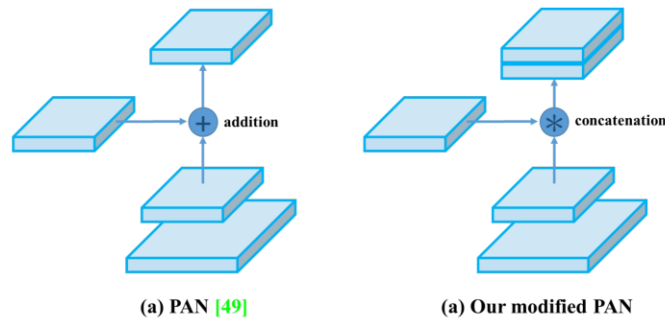


Рис. 3.9 – Модифікована технологія PAN в YOLOv4.

Візуальна увага широко застосовується при виявленні об'єктів. Spatial Attention Module (SAM) зосереджується на тому, де знаходиться інформативна частина і відповідно формує просторову карту уваги. Однак, в YOLOv4 SAM теж модифікована і використовується без застосування максимізаційного й усереднювального агрегування.

3.5 Детальний опис компонентів, використаних під час встановлення YOLOv4

Детальний опис компонентів, використаних під час встановлення YOLOv4, може включати наступні елементи:

1. Darknet Framework: Darknet - це фреймворк для нейронних мереж, який використовується для реалізації алгоритму YOLO. Darknet забезпечує основну інфраструктуру для створення, навчання та тестування моделей глибокого навчання [26, с. 1571].

2. Конфігураційні файли: Під час встановлення YOLOv4 потрібно внести зміни у конфігураційні файли Darknet, щоб визначити параметри моделі, такі як розмір зображення, кількість класів, шляхи до файлів з вагами тощо.

3. Попередньо навчені ваги: Для використання моделі YOLOv4 необхідно завантажити попередньо навчені ваги, які можна отримати з офіційного джерела або інших відкритих джерел.

4. Компілятор CUDA і cuDNN: Для прискорення обчислень у Deep Learning можна використовувати технології NVIDIA, такі як CUDA (Compute Unified Device Architecture) і cuDNN (CUDA Deep Neural Network library). Вони дозволяють використовувати графічні процесори NVIDIA для обчислень у глибоких нейронних мережах.

5. Компіляція Darknet: Після внесення змін у конфігураційні файли Darknet потрібно скомпілювати фреймворк, щоб зміни набрали чинності і можна було використовувати модель YOLOv4.

6. Інсталяція залежностей: Під час встановлення YOLOv4 може знадобитися встановлення додаткових бібліотек та залежностей Python, таких як OpenCV, NumPy та інші.

7. Тестування і відладка: Після успішної компіляції Darknet і завантаження попередньо навчених ваг можна провести тестування моделі на тестових даних та відлагодити будь-які проблеми, що виникають.

8. Навчання моделі (опціонально): Якщо необхідно, можна навчити модель YOLOv4 на власних даних, використовуючи Darknet та попередньо навчені ваги як початкову точку.

9. Оцінка результатів: Після завершення встановлення та навчання моделі можна оцінити результати за допомогою метрик відповідності, точності тощо.

Ці компоненти складають основну частину процесу встановлення YOLOv4 та готовість його до використання для виявлення об'єктів на зображеннях або відео.

ВИСНОВКИ

У цій роботі було комплексно досліджено проблему розробки системи комп'ютерного зору для автоматизованого розпізнавання об'єктів на основі методів машинного навчання, з особливим акцентом на застосування в точному землеробстві.

Теоретична частина:

Було проаналізовано сучасні завдання та методи систем комп'ютерного зору, включаючи інформаційний пошук, навчання та застосування в робототехніці. Особливу увагу приділено "складній проблемі зору" та її впливу на розробку ефективних алгоритмів.

Детально розглянуто теоретичні основи згорткових нейронних мереж (CNN), їх структуру та основні компоненти, такі як згорткові шари, карти ознак, пулінг та повнозв'язкові шари.

Вивчено алгоритми прямого та зворотного поширення помилки, які лежать в основі навчання нейронних мереж, а також різні методи оптимізації, що використовуються для покращення їх продуктивності.

Проведено огляд методів виявлення об'єктів та їх застосування в точному землеробстві, підкреслюючи важливість цієї технології для підвищення ефективності та сталості сільськогосподарського виробництва.

Практична частина:

Проведено аналіз і порівняння існуючих наборів даних, що використовуються для навчання моделей комп'ютерного зору в точному землеробстві.

Створено власний набір даних на основі зображень, отриманих з борту БПЛА, з детальною розміткою об'єктів, що представляють інтерес для точного землеробства.

Встановлено та налаштовано нейронну мережу YOLOv4, яка є однією з найефективніших моделей для виявлення об'єктів.

Детально описано компоненти та кроки, необхідні для успішної установки та використання YOLOv4.

Розроблено систему комп'ютерного зору, здатну автоматично розпізнавати об'єкти на зображеннях з БПЛА, що може бути використано для вирішення різноманітних завдань у точному землеробстві.

Створено власний набір даних, який може бути корисним для подальших досліджень та розробок в цій області.

Надано детальні інструкції щодо встановлення та використання YOLOv4, що може допомогти іншим дослідникам та розробникам впроваджувати цю технологію.

Перспективи подальших досліджень:

- Дослідження та впровадження нових методів оптимізації та архітектур нейронних мереж для підвищення точності та ефективності виявлення об'єктів.
- Розширення набору даних для включення більшої різноманітності об'єктів та умов зйомки.
- Інтеграція системи комп'ютерного зору з іншими технологіями точного землеробства, такими як системи автоматичного управління сільськогосподарською технікою.

Робота демонструє значний потенціал застосування методів машинного навчання та комп'ютерного зору для вирішення актуальних проблем у точному землеробстві. Вона відкриває нові можливості для оптимізації сільськогосподарського виробництва, підвищення його ефективності та сталості.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Bochkovskiy A. CFG Parameters in the [net] section. 2020. веб-сайт. URL:<https://github.com/AlexeyAB/darknet/wiki/CFG-Parameters-in-the-%5Bnet%5D-section> (дата звернення: 29.02.2024).
2. FREE FLIR Thermal Dataset for Algorithm Training. *FLIR* : веб-сайт. URL: <https://www.flir.com/oem/adas/adas-dataset-form/> (дата звернення: 29.02.2024).
3. Zheng Z., Wang P., Liu W., Li J., Ye R., Ren D. Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. AAAI 2020. 2020. Vol. 34. No. 07. P. 12993-13000.
4. See the full spectrum. Infiniti Electro-Optics : веб-сайт. URL: <https://www.infinitioptics.com/technology/multi-sensor> (дата звернення: 03.03.2024)
5. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu Ch.-Y., Berg A. C. SSD: Single Shot MultiBox Detector. Computer Vision – ECCV 2016, Springer. 2016. Vol. 9905. P. 21-37. DOI: 10.1007/978-3-319-46448-0_2.
6. Ghiasi G., Lin T.-Y., Le Q. V. DropBlock: A regularization method for convolutional networks. 2018. URL: <https://arxiv.org/abs/1810.12890> (дата звернення: 03.03.2024).
7. Redmon J., Divvala S., Girshick R., Farhadi A. You Only Look Once.
8. (ICCV). 2017. P. 2999-3007. DOI: 10.1109/ICCV.2017. P. 324.
9. Lin T., Dollár P., Girshick R., He K., Hariharan B., Belongie, S. Feature Pyramid Networks for Object Detection. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. P. 936-944. DOI: 10.1109/CVPR. 2017. 106.
10. Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017. Vol. 39, No. 6, P. 1137-1149. DOI: 10.1109/TPAMI.2016.2577031.

11. How single-shot detector (SSD) works? ArcGIS : веб-сайт. URL: <https://developers.arcgis.com/python/guide/how-ssd-works> (дата звернення: 22.03.2024)
12. Dragana P., Branko L., Miroslav P., Sasa V. Thermal Imager Range: Predictions, Expectations, and Reality. 2019. DOI: 10.3390/S19153313
13. Bochkovskiy A., Wang Y., Liao H.-Y. M. YOLOv4: Optimal Speed and Accuracy of Object Detection. 2018. URL: <https://arxiv.org/abs/2004.10934> (дата звернення: 22.03.2024).
14. Unified, Real-Time Object Detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. P. 779-788. DOI: 10.1109/CVPR.2016.91.
15. He K., Gkioxari G., Dollar P., Girshick R. Mask R-CNN. 2017 IEEE International Conference on Computer Vision (ICCV). 2017. P. 2980-2988. DOI: 10.1109/ICCV.2017.322.
16. Huang G., Liu Z., Van Der Maaten L., Weinberger Q. Densely Connected Convolutional Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. P. 2261-2269. DOI: 10.1109/CVPR.2017.243.
17. Real E., Aggarwal A., Huang Y., Le Q. V. Regularized Evolution for Image Classifier Architecture Search. AAAI. 2019 Vol. 33. No. 01. P. 4780-4789.
18. Redmon J., Farhadi A. YOLOv3: An Incremental Improvement. 2018. URL: <https://arxiv.org/abs/1804.02767v1> (дата звернення: 09.04.2024).
19. Detection. 2017 IEEE International Conference on Computer Vision
20. Zhang Z., He T., Zhang H., Zhang Z., Xie J., Li M. Bag of Freebies for Training Object Detection Neural Networks. 2019. URL: <https://arxiv.org/abs/2002.05712> (дата звернення: 27.04.2024).

- 21.Bochkovskiy A. CFG Parameters in the different layers. 2020. веб-сайт. URL: <https://github.com/AlexeyAB/darknet/wiki/CFG-Parameters-in-the-different-layers> (дата звернення: 10.05.2024).
- 22.COCO Dataset. COCO : веб-сайт. URL: <https://cocodataset.org/#download> (дата звернення: 15.05.2024).
- 23.Huang Z., Wang J., Fu X., Yu T., Guo Y., Wang R. DC-SPP-YOLO: Dense connection and spatial pyramid pooling based YOLO for object detection. Information Sciences. 2020. Vol. 522. P. 241-258. DOI: 10.1016/j.ins.2020.02.067
24. Ramachandran P., Zoph B., Le Q. V. Searching for Activation Functions. 2017. URL: <https://arxiv.org/abs/1710.05941> (дата звернення: 15.05.2024).
- 25.See Heat. No Light Required. Infiniti Electro-Optics : веб-сайт. URL: <https://www.infinitioptics.com/technology/multi-sensor> (дата звернення: 20.05.2024)
- 26.Wang C., Mark Liao H., Wu Y., Chen P., Hsieh J, Yeh I. CSPNet: A New Backbone that can Enhance Learning Capability of CNN. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2020. P. 1571-1580. DOI: 10.1109/CVPRW50498.2020.00203.

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ
ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА ШТУЧНОГО ІНТЕЛЕКТУ

КВАЛІФІКАЦІЙНА РОБОТА

на тему: «Розробка системи комп'ютерного зору для
автоматизованого розпізнавання об'єктів на основі
методів машинного навчання»

Виконав: здобувач вищої освіти гр. ШД-41
Максим Мушко

Керівник: доктор технічних наук, доцент
Ольга Зінченко

2024 рік

Мета роботи – підвищення ефективності використання систем комп'ютерного зору для роботи з зображеннями, отриманими з борту БПЛА для завдань точного землеробства.

Об'єкт дослідження – процес автоматизованого розпізнавання об'єктів на основі методів машинного навчання.

Предмет дослідження – система комп'ютерного зору для завдань точного землеробства.

Завданнями дослідження є:

- Вивчення теоретичних аспектів проблеми розробки системи комп'ютерного зору для автоматизованого розпізнавання об'єктів на основі методів машинного навчання.
- Формування та розмітка набору даних зображень, отриманих з борту БПЛА, для задач точного землеробства.
- Розробка та апробація алгоритму автоматизованого розпізнавання об'єктів на зображеннях з використанням методів машинного навчання.
- Оцінка ефективності розробленої системи комп'ютерного зору на тестовому наборі даних.

Теоретичні аспекти

Основні завдання комп'ютерного зору

- Розпізнавання об'єктів
- Сегментація зображень
- Відстеження об'єктів
- 3D-реконструкція

Методи комп'ютерного зору

- Глибоке навчання
- Згорткові нейронні мережі (CNN)
- Генеративні моделі

Типи алгоритмів машинного навчання:

- Навчання з учителем
- Навчання без учителя
- Півнаглядне навчання

Застосування машинного навчання:

- Розпізнавання образів
- Прогнозування
- Рекомендаційні системи
- Обробка природної мови

БПЛА: класифікація, застосування переваги

Очевидними перевагами застосування БПЛА є:

- можливість зйомки із незначних висот із наближенням до об'єктів;
- отримання знімків високої просторової роздільної здатності;
- рентабельність;
- оперативність отримання зображень;
- можливість застосування у зонах надзвичайних ситуацій без ризику для життя й здоров'я пілотів.

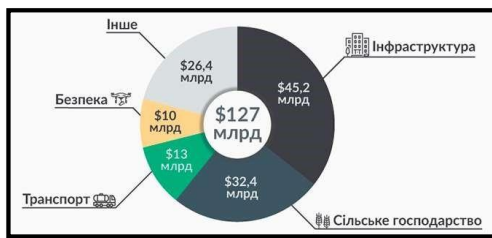


Рис.1 Світовий ринок послуг із використанням БПЛА

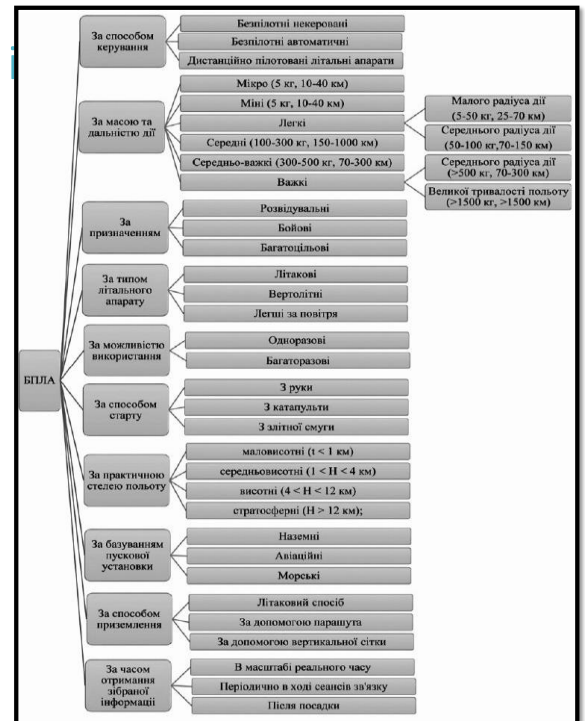


Рис. 2 Класифікація БПЛА

ФОРМУВАННЯ ТА РОЗМІТКА НАБОРУ ДАНИХ

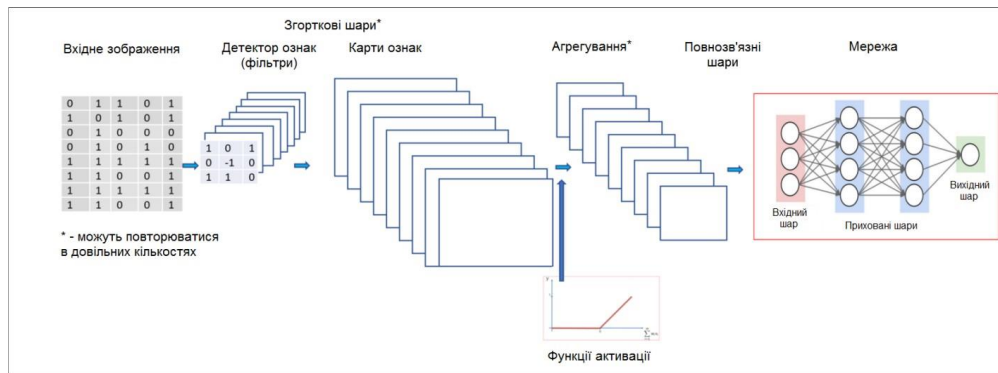


Рис. 3 Загальна структурна схема побудови згорткової нейронної мережі

$$(I * K)(i, j) = \sum_{m=-M}^M \sum_{n=-N}^N I(i - m, j - n) \cdot K(m, n)$$

Формула згортки, де: I — початкове зображення, K — ядро згортки (фільтр), (i, j) — координати вихідного зображення, M, N — половина розміру ядра згортки по відповідним осям.

ФОРМУВАННЯ ТА РОЗМІТКА НАБОРУ ДАНИХ

- Карта ознак допомагає нейронній мережі виділяти важливі патерни, що дозволяє краще розпізнавати та класифікувати об'єкти.
- Пулінг допомагає зменшити розмірність даних та зробити модель більш стійкою до незначних змін у зображенні.
- Повнозв'язкові шари дозволяють моделі робити складні висновки та класифікувати об'єкти на зображеннях.
- Пряме поширення використовується для отримання прогнозів моделі, а зворотне поширення помилки дозволяє їй навчатися та покращувати свої результати.
- Методи оптимізації є важливим інструментом для налаштування параметрів моделі та досягнення найкращої продуктивності.

Архітектура YOLO

YOLO, що означає "You Only Look Once" (Ти дивишся лише один раз), є однією з найпопулярніших архітектур для виявлення об'єктів у реальному часі. Принцип роботи YOLO полягає у тому, що вона ділить зображення на сітку рівномірного розміщення. Залежно від роздільної здатності, ця сітка має певну кількість клітинок. Потім модель прогнозує кількість та положення граничних рамок (bounding boxes) для кожної клітинки сітки. Кожна рамка містить інформацію про її розміри, положення відносно комірки сітки, а також показник достовірності, що вказує на ймовірність того, що в рамці міститься об'єкт. Чим вище цей показник, тим більша рамка навколо об'єкту

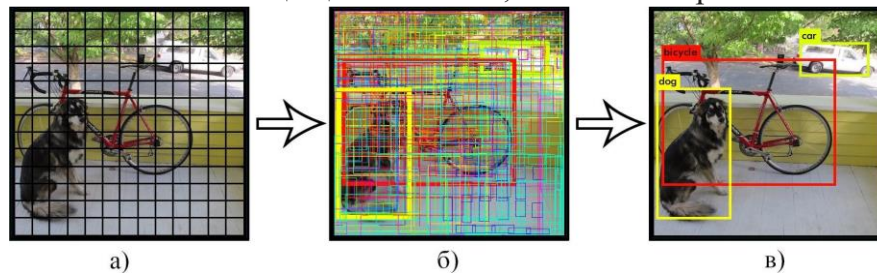


Рис. 4 Принцип роботи архітектури YOLO

Порівняння YOLO з іншими архітектурами

Архітектура	Переваги	Недоліки	Застосування
YOLO (You Only Look Once)	Швидкість, точність в реальному часі, здатність виявляти кілька об'єктів одночасно	Менша точність для малих об'єктів, чутливість до змін в масштабі та співвідношенні сторін	Виявлення об'єктів в реальному часі, автономні транспортні засоби, системи безпеки
R-CNN (Regions with CNN features)	Висока точність виявлення об'єктів, особливо для малих об'єктів	Низька швидкість, складність обчислень	Виявлення об'єктів на статичних зображеннях, медична діагностика
Fast R-CNN	Покращена швидкість порівняно з R-CNN, збереження високої точності	Все ще відносно повільний порівняно з YOLO	Виявлення об'єктів на статичних зображеннях, коли швидкість не є критичною
Faster R-CNN	Ще більш швидкий ніж Fast R-CNN, завдяки використанню Region Proposal Network (RPN)	Складність архітектури	Виявлення об'єктів на статичних зображеннях, коли потрібна висока точність і прийнятна швидкість
SSD (Single Shot MultiBox Detector)	Висока швидкість та точність, здатність працювати з різними масштабами об'єктів	Менша точність для малих об'єктів порівняно з Faster R-CNN	Виявлення об'єктів в реальному часі, мобільні додатки

Приклад коду для створення власного набору даних та навчання моделі на основі бібліотек TensorFlow та Keras

```
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers

# Створення власного набору даних
train_images = ... # Завантажте зображення для навчання
train_labels = ... # Відповідні мітки для зображень
test_images = ... # Завантажте зображення для тестування
test_labels = ... # Відповідні мітки для зображень

# Побудова моделі
model = keras.Sequential([
    layers.Conv2D(32, (3, 3), activation='relu'),
    layers.MaxPooling2D((2, 2)),
    layers.Conv2D(64, (3, 3), activation='relu'),
    layers.MaxPooling2D((2, 2)),
    layers.Flatten(),
    layers.Dense(64, activation='relu'),
    layers.Dense(10) # Кількість класів для класифікації
])

# Компіляція моделі
model.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy (from_logits=True),
              metrics=['accuracy'])

# Навчання моделі
model.fit(train_images, train_labels, epochs=10,
          validation_data=(test_images, test_labels))

# Оцінка моделі
test_loss, test_acc = model.evaluate(test_images, test_labels)
print('Test accuracy:', test_acc)
```

Установка нейронної мережі YOLOv4

1. Підготовка середовища: Переконайтеся, що на вашому комп'ютері встановлено Python та інші необхідні інструменти для розробки. Рекомендується використовувати середовище віртуального оточення для управління пакетами та залежностями.

2. Клонування репозиторію Darknet: Darknet - це фреймворк, на якому базується YOLO. Клонуйте репозиторій Darknet з GitHub або іншого джерела.

3. Конфігурація Darknet: Налаштуйте Darknet для використання YOLOv4. Це може включати зміни конфігураційних файлів, залежно від вашого використання та потреб.

4. Завантаження ваг моделі: Завантажте попередньо навчені ваги моделі YOLOv4. Ці ваги можна отримати з офіційного джерела або з інших джерел.

5. Компіляція Darknet: Скомпілюйте Darknet з підтримкою YOLOv4. Це може зайняти деякий час, але після цього ви зможете використовувати Darknet для навчання або тестування моделі YOLOv4.

6. Тестування або навчання моделі: Після успішної установки ви можете тестувати або навчати модель YOLOv4 на ваших даних. Для цього ви можете використовувати інтерфейс командного рядка Darknet або розробляти власний код для використання моделі.

7. Оцінка результатів: Після завершення навчання або тестування оцініть результати за допомогою метрик відповідності, точності тощо.

Висновки

У дипломній роботі розроблено та впроваджено модель нейронної мережі для розпізнавання об'єктів на основі YOLOv4. Детально описано компоненти та кроки, необхідні для успішної установки та використання YOLOv4.

Розроблено систему комп'ютерного зору, здатну автоматично розпізнавати об'єкти на зображеннях з БПЛА, що може бути використано для вирішення різноманітних завдань у точному землеробстві.

Створено власний набір даних, який може бути корисним для подальших досліджень та розробок в цій області. Проведено порівняльний аналіз різних алгоритмів розпізнавання об'єктів, таких як Faster R-CNN, SSD та Mask R-CNN.

Отримані результати підтверджують те, що YOLOv4 є однією з найефективніших моделей для виявлення об'єктів, що дозволяє рекомендувати її використання у практичних додатках.

Дякую за увагу