

2022p.

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ РАДІОЕЛЕКТРОНІКИ

Факультет Автоматики і комп'ютеризованих технологій
Кафедра Комп'ютерно-інтегрованих технологій, автоматизації та мехатроніки
Рівень вищої освіти другий (магістерський)
Спеціальність 151 Автоматизація та комп'ютерно-інтегровані технології
Тип програми Освітньо-професійна
Освітня програма Комп'ютерно-інтегровані технологічні процеси та виробництва
(шифр і назва)

ЗАТВЕРДЖУЮ:
Зав. кафедри КІТАМ _____
(підпис)
« _____ » _____ 2022 р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

студентові Брюховецькому Олександрю Андрійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Розроблення системи розпізнавання об'єктів на основі комп'ютерного зору безпілотних літальних апаратів
Затверджена наказом по університету від 7.11.22 № 1464 Ст
2. Термін подання студентом роботи до екзаменаційної комісії _____
3. Вихідні дані до роботи: системи розпізнавання об'єктів, мова програмування Python, бібліотека TensorFlow
- _____
- _____
- _____
- _____
4. Перелік питань, що потрібно опрацювати в роботі 4.1 Вступ
4.2 Огляд предметної області
4.3 Огляд систем розпізнавання об'єктів
4.4 Розроблення системи розпізнавання об'єктів
4.5 Висновки

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп’ютерних ілюстрацій Слайди у форматі Power Point 12 шт.

6. Консультанти розділів роботи

Найменування розділу	Консультант (посада, прізвище, ім’я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Отримання завдання на до атестаційної роботи	01.09.2022	виконано
2	Аналіз завдання, підбір літератури	01.09.2022-1.10.2022	виконано
3	Огляд технологічних засобів	01.09.2022-15.10.2022	виконано
4	Огляд існуючих методів	15.10.2022-30.10.2022	виконано
5	Програмна реалізація	30.10.2022-1.11.2022	виконано
6	Проведення дослідження обраних моделей	1.11.2022-11.11.2022	виконано
7	Проведення експертного оцінювання	11.11.2022-29.11.2022	виконано
8	Оформлення пояснювальної записки	29.11.2022-19.12.2022	виконано
9	Перевірка на плагіат	19.12.2022	виконано
11	Рецензування	20.12.2022	виконано
12	Підготовка презентації та доповіді	21.12.2022	виконано
13	Попередній захист атестаційної роботи	22.12.2022	виконано

Дата видачі завдання 01.09.22

Студент _____
(підпис)

Брюховецький О.А.
(прізвище, ініціали)

Керівник роботи _____
(підпис)

проф. Сезонова І.К.
(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка: 55 с., 5 табл., 14 рис., 1 дод., 31 джерел.

СИСТЕМА КОМП'ЮТЕРНОГО ЗОРУ, ШТУЧНИЙ ІНТЕЛЕКТ, НЕЙРОННА МЕРЕЖА, НАВЧАННЯ З ВЧИТЕЛЕМ, БЕЗПЛОТНИЙ ЛІТАЛЬНИЙ АПАРАТ, DATASET IMAGENET, БІБЛІОТЕКА TENSORFLOW.

Об'єкт дослідження – системи розпізнавання визначеного об'єкта за допомогою системи комп'ютерного зору безпілотного літального апарату (БПЛА), просторові моделі геометричних об'єктів, методи їх ідентифікації та розпізнавання.

Предмет дослідження – геометричні характеристики визначеного об'єкта щодо автоматизації розпізнавання та програмні модулі для супроводження його розпізнавання системами комп'ютерного зору БПЛА.

Мета кваліфікаційної роботи – автоматизація системи розпізнавання об'єктів системою комп'ютерного зору БПЛА. ідентифікація параметрів конструкцій прес-форм м'яких роботизованих маніпуляторів.

Методи дослідження – аналіз та синтез теоретичних і практичних знань, експериментальний за допомогою моделювання ситуації щодо розпізнавання предметів, програмний.

У кваліфікаційній роботі досліджено геометричні характеристики визначеного об'єкта щодо автоматизації розпізнавання та програмні модулі для супроводження його розпізнавання системами комп'ютерного зору БПЛА.

Для цього проаналізовано існуючі системи розпізнавання об'єктів на основі комп'ютерного зору.

На базі проведеного дослідження розроблено модель та програму для БПЛА щодо розпізнавання визначеного об'єкту.

Для забезпечення безпечних умов роботи у лабораторії з виготовлення прес-форми для м'яких роботизованих захватів проведені необхідні розрахунки з усунення основних шкідливий виробничий факторів в робочій зоні.

Результати кваліфікаційної роботи апробовані у 3 фахових статтях та 2 міжнародних конференціях.

ABSTRACT

Explanatory note: 55 pp., 5 tables, 14 figures, 1 appendix, 31 sources.

COMPUTER VISION SYSTEM, ARTIFICIAL INTELLIGENCE, NEURAL NETWORK, LEARNING WITH A TEACHER, UAV, DATASET IMAGENET, TENSORFLOW LIBRARY.

The object of research is the recognition systems of a specified object using the computer vision system of an unmanned aerial vehicle (UAV), spatial models of geometric objects, and methods of identification and recognition.

The subject of the study is the geometric characteristics of the specified object regarding automation of recognition and software modules to support its recognition by UAV computer vision systems.

The purpose of the qualification work is the automation of the object recognition system by the UAV computer vision system. identification of structural parameters of soft robotic manipulator presses forms.

Research methods – analysis and synthesis of theoretical and practical knowledge, experimental with the help of modeling the situation regarding object recognition, software.

In the qualification work, the geometric characteristics of the specified object about the automation of recognition and software modules to support its recognition by UAV computer vision systems were investigated.

For this purpose, existing object recognition systems based on computer vision were analyzed.

Based on the conducted research, a model and a program for the recognition of the specified object were developed for the UAV.

To ensure safe working conditions in the laboratory for the production of molds for soft robotic grippers, the necessary calculations were made to eliminate the main harmful production factors in the work area.

The results of the qualification work were tested in 3 professional articles and 2 international conferences.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів 7

Вступ 8

1 Огляд предметної області..... 9

 1.1 Термінологія..... 9

 1.2 Історія БПЛА..... 10

 1.3 Класифікація БПЛА..... 14

 1.4 Датчики на борту БПЛА..... 15

 1.5 Використання штучного інтелекту в сучасних БПЛА..... 17

 1.6 Способи обробки даних БПЛА..... 20

 1.7 Особливості глибокого навчання для БПЛА 24

2 Огляд систем розпізнавання об’єктів..... 25

 2.1 Вхідні дані..... 25

 2.2 Історія розвитку комп’ютерного зору..... 25

3 Розроблення системи розпізнавання об’єктів..... 33

 3.1 Структура згорткових нейронних мереж..... 33

 3.2 Середовище навчання нейронної мережі..... 34

 3.2.1 TensorFlow..... 34

 3.2.2 Keras..... 35

 3.2.3 TensorFlow Lite..... 35

 3.2.4 Tensorflow detection API..... 36

 3.2.5 Google Colaboratory 36

 3.3 Виявлення об’єкту. 37

 3.3.1 Оцінка достовірності..... 39

 3.3.2 Розташування..... 40

 3.4 Навчання мережі..... 40

 3.5 Донавчання мережі..... 41

 3.6 Встановлення моделі на Android смартфон..... 42

 3.7 Тестування моделі розпізнавання об’єктів..... 43

4 Охорона праці..... 47

Висновки..... 49

Перелік джерел посилання..... 50

Додаток А Демонстраційний матеріал 54

ПЕРЕЛІК СКОРОЧЕНЬ І ТЕРМІНІВ

БПЛА – безпілотний літальний апарат;

ДДЗ – датчик динамічного зору;

ЛА – літальний апарат;

ПЗ – програмне забезпечення;

ПК – персональний комп'ютер;

ШІ – штучний інтелект;

API – Application Programming Interface;

CAA – Civil Aviation Authority;

CCD – charge-coupled device;

CNN – Convolutional neural network;

CV – computer vision;

DoD – Department of Defense;

GNSS – Global Navigation Satellite System;

ILSVRC – ImageNet Large Scale Visual Recognition Challenge;

IMU – inertial measurement unit.

IoU – Intersection Over Union;

LiDAR – Light Detection and Ranging;

SLAM – simultaneous localization and mapping;

UAV – unmanned aerial vehicle.

ВСТУП

Безпілотні літальні апарати (БПЛА), які також широко відомі як дрони – це літальні апарати (ЛА), які літають без пілота на борту. Використання БПЛА надає численні переваги в порівнянні з керованими ЛА, це виготовлення ЛА різних розмірів, зменшення часу навчання, а також вартості та розмірів літака, що має величезний вплив на діапазон можливих застосувань, а найголовніше це збереження життя пілотів. Цифри, що стоять за галуззю БПЛА, вражають: за оцінками Value Market Research, ринок БПЛА з вертикальним зльотом і посадкою досягне приблизно 10 163 мільйонів доларів США до 2024 року. Інший звіт PwC передбачає, що у 2030 році в небі Великобританії працюватиме 76 000 дронів, що залучатиме загалом 628 000 робочих місць [1]. Український ринок БПЛА також зростає, з 2016 року виручка розробників «Фурії» та «Лелеки» зросла з 4,25 млн грн до 110,3 млн грн в 2022-му [2].

Головне завдання розвитку БПЛА сьогодні – це досягнення повністю автономних оперативних можливостей для польоту та розуміння реальних і складних сценаріїв.

Таким чином метою кваліфікаційної роботи є розробка системи розпізнавання визначеного об'єкта за допомогою системи комп'ютерного зору безпілотного літального апарату.

Об'єкт дослідження – просторові моделі геометричних об'єктів, методи їх ідентифікації та автоматизація система розпізнавання об'єктів БПЛА.

Предмет дослідження – геометричні характеристики визначеного об'єкта щодо автоматизації розпізнавання та програмні модулі для супроводження його розпізнавання системами комп'ютерного зору БПЛА.

Методи дослідження – структурний аналіз та проектування автоматизованих систем, проектування і програмування, методи і засоби розробки програм.

Для досягнення поставленої мети необхідно вирішити такі завдання:

- проаналізувати існуючі системи розпізнавання об'єктів на основі комп'ютерного зору;
- розробити модель та програму для БПЛА щодо розпізнавання визначеного об'єкту;
- оформити пояснювальну записку згідно з рекомендаціями, керуючись навчальним посібником з дипломного проєкту [3], ДСТУ 3008 – 15 [4] та методичними вказівками [5].

1 ОГЛЯД ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Термінологія предметної області

Хоча в літературі найчастіше цитується термін БПЛА, назви та аббревіатури літальних апаратів без людей на борту включають безпілотну літальну систему (UAS), віддалено пілотований літальний апарат (RPA) або транспортний засіб (RPV), віддалено керований літальний апарат (ROA).), а також безлюдні літальні апарати. Технічно між цими термінами є відмінності [6], але також є випадок семантичної плутанини.

Наприклад, останній БПЛА зазвичай відноситься до пристроїв, що використовуються у військових цілях, тоді як перший відноситься до пристроїв, які використовуються новачками та професіоналами в цивільних і неквапливих додатках. В офіційних повідомленнях Міністерства оборони США (DoD) та Управління цивільної авіації (CAA) Сполученого Королівства термін UAS використовується для опису системи, що включає безпілотний літальний апарат (UA), наземну систему управління (GCS) і канал передачі даних [6]. За визначенням Коломіні та Моліни припустили, що БПЛА – це система систем: набір взаємодоповнюючих технологій, об'єднаних для виконання конкретного завдання, і як такі існують різноманітні системи. Крім того, в літературі наводяться додаткові відмінності та класифікації БПЛА на основі їхніх характеристик (розмір, вага, витривалість, аеродинаміка, дальність польоту та висота польоту), а також способів їх застосування [6].

Виходячи з пошукових запитів, загалом термін «дрон» використовується для опису будь-якого БПЛА (або системи), який оснащено системою автономного польоту. Загалом, на протязі 10 років (від 2005 до 2015), інтерес до цих пристроїв зростає, і використання терміну «БПЛА» є найпопулярнішим для їх позначення. Крім того, протягом останнього десятиліття спостерігалось постійне зростання використання термінів «дрон» і «БПЛА», причому кожен термін рідко з'являвся в перші роки, що розглядаються. Разом «дрон» і «БПЛА» складають понад 20% результатів пошуку Google Scholar і приблизно 10% результатів пошуку SCOPUS у 2015 році [7]. Відносно обмежене використання термінів «БПЛА» і «дрон» в літературі протягом років не відображають їх поширеність у повсякденних засобах масової інформації, а FAA інституціоналізує термін «UAS».

1.2 Історія БПЛА

Подібно до багатьох об'єктів, які використовуються навіть у цивільному житті, БПЛА є результатом розробок, проведених під час військових конфліктів [8]. Ці типи літальних апаратів з'явилися в Першу світову війну. Вони представляли собою звичайні літаки зі встановленим автопілотом та використовувались як летючі бомби для терору міст які за теорією Джуліо Дзе повинні були підірвати моральний дух населення [9].

Після початку Другої світової війни попит військових на безпілотні машини-мішені почав стрімко зростати. Ось чому радіолітак (рис. 1.1) під кодовою назвою OQ-2 був створений у США Реджинальдом Денні. Це був дерев'яний моноплан, який приводився в рух екіпажем. БПЛА міг приземлитися за допомогою парашута, якщо з якихось причин був втрачений канал передачі даних, і ця функція доводить, що це відносно розвинена модель. Крім того, він міг злетіти зі своїх шасі на звичайній злітно-посадковій смугі. На відміну від попередніх БПЛА, завдяки цим функціям можна було уникнути пошкоджень.



Рисунок 1.1 – OQ-2 в музеї

Після цього компанія Реджинальда Денні випустила кілька OQ-моделей, найдосконалішою з цієї серії була OQ-14, яка залишалася на озброєнні навіть після Другої світової війни.

Свого розвитку отримали і літаючі бомби. З 1931 року в Німеччині розроблялись перші в світі крилаті ракети V-1, Vergeltungswaffe-1 («Зброя відплати-1»), або «Фау-1».

Для керування польотом німецькі інженери використовували звичайну систему автопілота, що складається з 3-ступеневого гіроскопа, барометричного висотоміра та обладнання для вимірювання відстані, причому працював він на повітрі. Після досягнення заданої дистанції двигун вимикався і почав пікірувати V-1, діяв як бомба, випущена з бомбардувальника [9] [10].

З початком холодної війни та завдяки глобальній ядерній загрозі, стратегічна розвідка стала важливою. Символом цієї епохи став літак-розвідник U-2, який був розроблений на замовлення ЦРУ на початку 1950-х років. Вважалося на висоті понад 20 км його виявлення та перехоплення було малоімовірним. Однак 1 травня 1960 року радянська ППО збила U-2 над Радянським Союзом. Так народилося нове застосування БПЛА – розвідувальні дрони.

У 1960 році ВПС США доручили Ryan Aeronautical переробити Ryan Model 147 РТА на розвідувальні безпілотники. Цей новий БПЛА отримав назву «Lightning Bug» під кодовою назвою BQM-34 Firebee. Цей ракетний комплекс міг бути запущений за допомогою катапульти як із землі, так і з корабля. Однак найчастіше це запускали з літаків DC-130, які могли нести 4 безпілотники одночасно. Спочатку, після завершення своєї місії, Firebee приземлювався з парашутом, пізніше спеціально модифікований гелікоптер СН-3 «зловив» його, а потім поставив на землю в описаному місці. Вперше цей новий безпілотний літальний апарат виявився корисним активом у війні у В'єтнамі [11].

Наступний крок в розвитку безпілотників став зменшення розмірів ЛА та використання сучасних на той час розробок в області мікроелектроніки та кібернетики. Подальший розвиток військових безпілотників складався в сегментації ЛА, розробці під конкретні цілі.

У 1980 році США розпочали програму БПЛА з метою створення недорогого безпілотника для швидких операцій. Розробка RQ-2 «Pioneer» стала результатом спільного проекту США та Ізраїлю в 1986 році. Це був літак-розвідник середнього розміру, який міг працювати за будь-якої погоди, в день та ніч. Був керований пілотом, а також міг виконувати не складні операції такі як автономний обліт заданої території (рисунок 1.2) [9] [10] [12].



Рисунок 1.2 – RQ-2 Pioneer

У цей період розробники дронів були зосереджені на альтернативних послугах електроживлення для дронів, де основною послугою була сонячна енергія. Це призвело до кількох цікавих дронів на сонячних батареях під назвою AERO-VironmentHAL SOL. Безпілотний літальний апарат високої амплітуди (AAL SOL) на сонячних батареях був спонсорований ЦРУ в 1980-х роках як перший літальний апарат на сонячних батареях.

Міні- та мікроверсії БПЛА були представлені в 1990 році, тоді як дрон Predator був представлений в 2000 році.

Величезне розширення інновацій у галузі дронів було очевидним за останні 10 років і в комерційному інтересі. Наприклад, на початку 2010-х дрони використовувалися не тільки як військова техніка, а й як засоби доставки. У 2015 році було видано близько 1000 комерційних дозволів на дрони, і попит на дрони продовжує експоненціально зростати. 2010 рік став так званим «золотим віком дронів». За рік їх кількість зросла втричі. В результаті злиття радіокерованих літальних апаратів і технологій смартфонів, оснащення дронів камерами стало новим звичним явищем у комерційній фотографії та відеозйомці [8].

Сучасні БПЛА можуть дистанційно керуватися пілотом або літати автономно. У першому сценарії пілот (оператор) знаходиться на значній відстані від апарату. Найпростіша така станція може складатись складається з пульта дистанційного

керування з додатковим екраном, найчастіше роль екрану може виконувати звичайний телефон або планшет. В останньому випадку станція може складатись з цілого комплексу керування польотом по заздалегідь розробленому плану польоту та динамічна система автоматизації. Головне завдання розвитку БПЛА сьогодні – це досягнення повністю автономних оперативних можливостей для польоту та розуміння реальних і складних сценаріїв.

В даний час удосконалення програмного та апаратного забезпечення у вбудованих системах і мініатюризація призвели до створення недорогих датчиків та інерційних вимірювальних пристроїв (IMU), які можуть отримувати корисну інформацію на борту, таку як кутова швидкість і орієнтація. Датчики стабілізації призначені для переведення БПЛА у стан рівноваги за рахунок вимірювання кутових координат та вирівнювання положення камери для відеозйомки. За це відповідає сукупність інерційних сенсорів: трьохосьовий акселерометр, трьохосьовий гіроскоп та трьохосьовий магнітометр. Така комбінація дає змогу зменшити габарити та споживання енергії, зберігаючи високу точність. Вимірювання проводяться у діапазоні $2/\pm 4/\pm 8/\pm 16$ g для акселерометра, $\pm 125/\pm 245/\pm 500/\pm 1000/\pm 2000$ °/с для гіроскопа та $2/\pm 4/\pm 8/\pm 16^\circ$ для магнітометра. Було також запропоновано багато підходів і конфігурацій, щоб отримати суттєві знання про навколишнє середовище з даних, отриманих споживчими RGB-камерами, датчиками глибини, LiDAR і камерами на основі подій. Також були представлені комплексні та повні набори датчиків, які об'єднують кілька даних.

Типовими операціями БПЛА є спостереження, відстеження людей, планування шляху, уникнення перешкод. Для цього оператор виводить БПЛА на точку виконання завдання, потім здійснюється моніторинг об'єкта за допомогою наявного цільового обладнання, після чого отримана інформація або транслюється в режимі реального часу, або зберігається на ЛА. Подальша її обробка може здійснюватися онлайн, але це вимагає певного рівня якості зв'язку який буде захищений від втрати інформації буде стабільним та захищеним від перешкод і не менш важливо, потребує участі оператора в керуванні. Коли попередньо перелічені критерії не вдається задовольнити набагато вигідніше робити обчислення на самому ЛА. Тому у сучасних системах управління мобільними роботами важливу роль відіграють технології комп'ютерного зору та машинного навчання. Завдяки якісному стрибку, що відбувся в останні роки, у розвитку цих технологій, вони знаходять все ширше застосування при вирішенні завдань навігації та управління рухом мобільними робототехнічними комплексами.

Отримання інформації для обробки та подальшого оцінювання стану об'єктів

здійснюється за допомогою спеціалізованих БПЛА. Особливістю таких безпілотних апаратів є те, що вони повинні відповідати певним ознакам, які дозволяють виконувати моніторинг стану об'єктів з подальшим їх оцінюванням. До них відносять тактичні (дальність польоту – 80 км); мало або середньорозмірні (від 200 кг до 2000 кг); малої (до 6 год), середньої (від 6 год до 12 год) чи великої тривалості польоту (більше 12 год); маловисокі (до 1 км); з літаковою або гелікоптерною аеродинамічною схемою; наземного чи морського базування; військові або цивільні; горизонтального чи вертикального підйому / посадки; багаторазові та з широким діапазоном радіуса дії.

1.3 Класифікація БПЛА

На сьогоднішній день існують різні типи БПЛА, які можна класифікувати за різними критеріями [13], такими як підйомна система, розмір, максимальна злітна маса, вантажопідйомність, дальність польоту, швидкість польоту, висота польоту, витривалість, умови експлуатації або рівень автономності тощо.

Дуже розширеним критерієм для встановлення класифікації є підйомна система, яка розрізняє:

- легші за повітря;
- літаки з нерухомим крилом;
- гвинтокрили або мультикоптери;
- машущіки;
- гібридні конфігурації.

Найбільш поширеними є конфігурації з фіксованим крилом і гвинтокрилом: БПЛА з фіксованим крилом завдяки своїй використовуються на великій дальності та висоті. Таким ЛА потрібна спеціально підготовлене місце або систему запуску та посадки. Гвинтокрили мають переваги у вертикальному зльоті та посадці, зависанні на місці та польотах на низькій швидкості, що робить їх ідеальними для навіть найбільшої кількості застосувань. Серед гвинтокрилих найпоширенішими типами є:

- вертолітний;
- мультироторний.

Вертольоти мають більшу корисну вантажопідйомність і витривалість з недовіком, що вони механічно складніші, складніші в керуванні та небезпечні (оскільки вони мають більші гвинти), ніж багатомоторні. З усіх вищезазначених причин мультикоптери є найбільш широко використовуваним типом БПЛА в цивільних і військових цілях, особливо на висотах польоту менше 120 м.

1.4 Датчики на борту БПЛА

БПЛА оснащені бортовими датчиками для отримання інформації про навколишнє середовище або про сам БПЛА. Їх спосіб використання залежить від багатьох факторів, таких як середовище в якому вони будуть застосовані, галузь використання, розміри та вантажопідйомність БПЛА, собівартість, рівень захищеності; рівень автономності БПЛА, тощо. Кожен датчик має різні робочі характеристики, тому він має свої переваги та недоліки. З усіх відомих технологій найбільш популярними або перспективними є:

- камери RGB – це пасивні датчики, які фіксують інформацію про інтенсивність візуального спектру світла за допомогою трьох різних каналів, червоного, зеленого та синього (RGB). Основними проблемами під час встановлення на борту БПЛА є швидкість пересування транспортного засобу або різкі дії, які можуть спричинити розмиття та шуми на зображенні. З цією метою зазвичай надають перевагу КОМП-камерам, особливістю яких є те, що вся область зображення сканується одночасно на відмінну CCD-камери, в яких сканування відбувається стрічково (вертикально або горизонтально), як для аналога з рухомими затворами [14].

Камери у градаціях сірого забезпечують лише один канал із інформацією про середню інтенсивність візуального спектру світла спостережуваної сцени. Відеокамера призначена для моніторингу стану об'єктів, тому повинна мати високу роздільну здатність не менше 12 Мп для фото- та відеозйомки, можливість запису 4K відео, високий динамічний діапазон, частоту кадрів більше 30 к/с та можливість запису в умовах поганої освітленості;

- камери подій або датчик динамічного зору (ДДЗ), добре підходять для аналізу руху в реальному часі, демонструючи значно вищу часову роздільну здатність і чутливість до світла та низьку затримку. Останнім часом він використовується на борту БПЛА [15]. Їх вихід, у будь-якому випадку, складається не з класичного зображення інтенсивності, а з послідовності асинхронних подій, тому необхідність подальших досліджень і розробки нових конвеєрів обробки зображень є дуже актуальною темою [16];

- тепловізори або інфрачервоні камери – це пасивні датчики, які вловлюють інфрачервоне випромінювання, що випускається всіма об'єктами з температурою вище абсолютного нуля. Якщо області їх застосування спочатку обмежувалися спостереженням і нічним баченням для військових цілей, то нещодавнє падіння ціни відкрило широке поле застосування [17].

Тепловізори можуть усунути проблеми недостачі освітлення звичайних камер у відтінках сірого та RGB, забезпечуючи точне рішення в усіх сферах, де необхідно виявляти та відстежувати життя, і БПЛА не є винятком [18];

– 3D-камери можуть знімати сцену, кольорове зображення, а також 3D-інформацію. Наразі три домінуючі технології для створення таких камер – це стереобачення, Time-of-flight (камера ToF) і Structured Light камери. Наприклад, комерційні стереокамери такі як камера Stereolab ZED, були використані для проблеми виявлення об’єктів БПЛА [19]. У будь-якому випадку, вони здебільшого засновані на тріангуляції, тому наявність текстури на зображенні необхідна для отримання інформації про глибину.

Завдяки останнім технологічним досягненням, відносно невеликі та недорогі комерційні сенсори глибини Microsoft Kinect [11], ASUS Xtion Pro Live використовуються для виявлення об’єктів встановлених на БПЛА. У будь-якому випадку, вищезгадані датчики засновані на структурованому світлі (рис. 1.3).

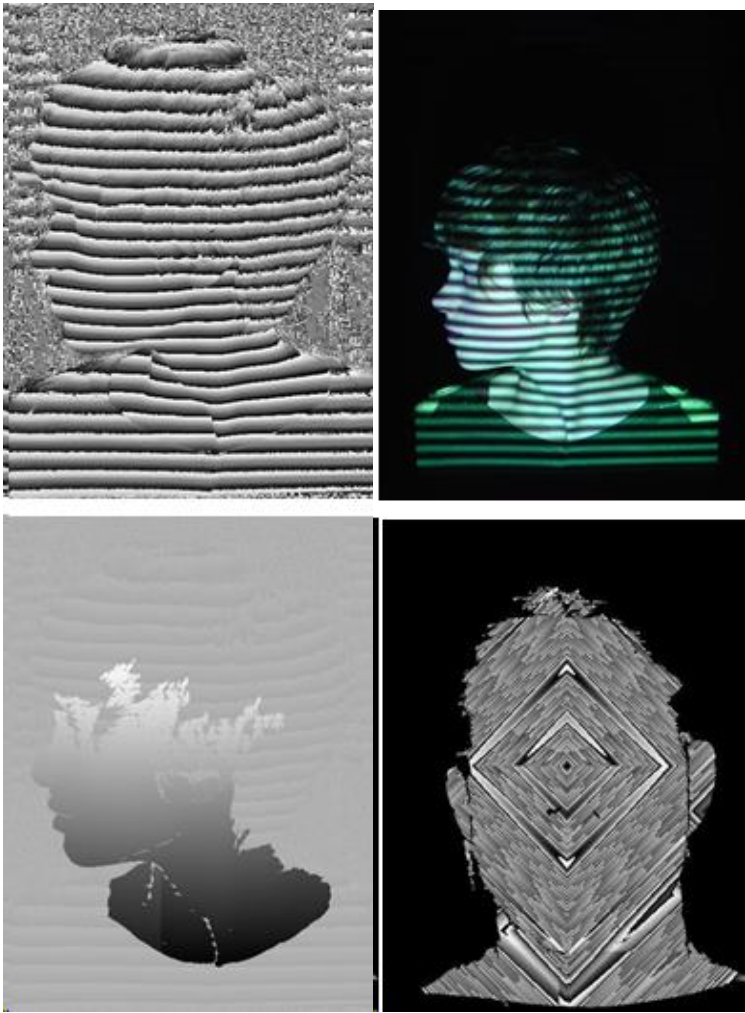


Рисунок 1.3 – Схема роботи структурованого світла

Оскільки «смуги» на різних об’єктах будуть зміщуватися пропорційно відстані, вимірюючи це зміщення ми можемо розрахувати відстань до об’єктів. Із за технічних

особливостей цих камер вони мають недоліки у вигляді «тіней» від об'єктів та чутливі до інших випромінювачів ІЧ світла, тому більше підходять для застосування всередині приміщень або контрольованих середовищ.

Камери ToF засновані на вимірюванні round-trip затримки світла, за який випромінюване інфрачервоне світло проходить до камери, відбиваючись від поверхонь об'єктів (рис. 1.4).

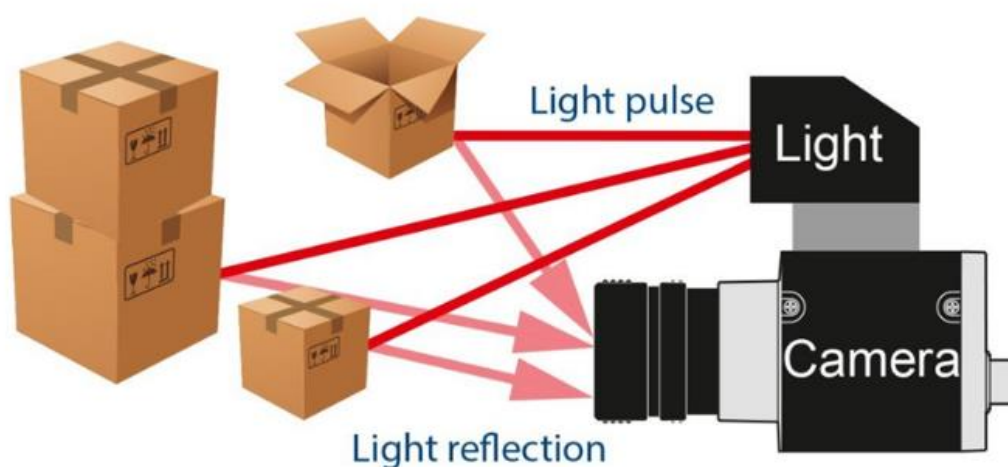


Рисунок 1.4 – Схема роботи камери ToF

Завдяки розвитку сучасних процесорів за 1 такт процесора с з частотою 3 ГГц світло встигає переміститись на 10 сантиметрів. Вони не мають «тіней» типових для структурованого світла, але вразливі до навколишнього світла та мають малу швидкість зйомки. Крім того, вони пропонують нижчу роздільну здатність, ніж датчики RGB;

– LiDAR – це активні датчики, які вимірюють відстані шляхом попадання на ціль лазерного світла та вимірювання відбиття за допомогою перетворювача. Раніше LiDAR були занадто великими, важкими та дорогими для використання на борту БПЛА, однак останні досягнення в твердотільних технологіях, наприклад LiDAR Ouster, значно зменшили їхній розмір, вага та вартість перетворюють їх на все більш популярний вибір бортових БПЛА.

1.5 Використання штучного інтелекту в сучасних БПЛА

Ще кілька років тому польотами БПЛА в основному керував дистанційно пілот: зліт, посадка, а також виконання даного польоту були під контролем людини-оператора, хоча прості траєкторії, такі як маршрут по точкам, контролювалися автопілотом. За останні 5–10 років ми спостерігаємо зростаючу кількість комерційних платформ, здатних надійно злітати, приземлятися та виконувати польоти в автономному режимі

відповідно до більш складних різних. Ці рішення виконуються відповідно до заздалегідь запланованих польотів (оскільки автономне рішення не приймається) і покладаються на навігацію на основі GNSS. Додаткові датчики (такі як камери та датчики відстані) були додані на нових комерційних платформах для виявлення великих перешкод і уникнення зіткнень, запобігаючи їхньому польоту надто близько до перешкод (або принаймні зупиняючи їх перед зіткненням), але лише дуже небагато останніх моделей мають можливість виявлення та обходити перешкоди наприклад, такі як останні лінійки DJI Mavic Air.

Автономна навігація в невідомому середовищі вимагає постійного просторово-часового сприйняття елементів навколишнього середовища, розуміння ситуації на місці події за допомогою даних та інформації, прийняття рішень щодо наступних етапів і здатності приймати швидкі рішення на основі знань попередніх елементів з мінімальними витратами, або відсутність втручання людини-оператора. Усе це було відкритою темою для досліджень у робототехніці протягом останніх двох десятиліть, про що свідчить величезна кількість опублікованих наукових робіт, більшість із яких присвячені внутрішньому середовищу (De Croon and De Wagter, 2018). Підходи до створення карт за допомогою мобільних платформ є або пасивними, коли людина сприймає оточення, щоб просто побудувати карту, або активними, де додатково планується, що траєкторія платформи рухається крізь середовище (Stachniss, 2009). Ці останні підходи часто можна розділити на три різні ітераційні кроки: (I) зондування, (II) локалізація та відображення та (III) планування шляху. Для автономної навігації дрона зазвичай використовуються різні датчики: моно- або стереокамери, а також лазери використовуються для картографування навколишнього середовища. Інерційні блоки та GNSS можна об'єднати, щоб локалізувати платформу, тоді як нещодавно були додані ультразвукові та ІЧ-камери для виявлення перешкод поблизу БПЛА. Незважаючи на розробку мініатюрних активних датчиків, сучасною тенденцією в літературі є використання RGB-камер, які, окрім їх зменшеної ваги, вартості та споживання енергії, можуть ефективно оброблятися за допомогою бортових пристроїв, що забезпечують більш повну та ефективну роботу гнучка інформація в режимі реального часу.

Автономне планування шляху стосується послідовності рішень, які приймає автономний транспортний засіб, щоб створити набір маршрутних точок без зіткнень і досягти пункту призначення. Для кожного автономного польоту потрібна загальна мета, яка полягає в тому, щоб (I) досягти цільової точки або (II) дослідити невідоме середовище з метою максимізації його інформаційного покриття. Тоді шлях обмежується різними елементами, такими як повнота збору даних, час польоту,

споживання енергії, найкоротший шлях до польоту або межі маневреності використовуваної платформи.

Якщо метою підходу є досягнення пункту призначення, у літературі можна знайти як офлайн, так і онлайн алгоритми. У автономних підходах траєкторії плануються заздалегідь, враховуючи вже приблизні геометричні знання навколишнього середовища, що дозволяє спланувати безпечний політ. Для цього завдання можна використовувати різні стратегії: у більшості випадків тривимірне середовище дискретизується в наборах вузлів, а потім політ оптимізується відповідно до таких алгоритмів, як Rapid-exploring Random. Дерева (RRT), RRT-зірка (RRT*), A-зірка (A*), ймовірнісні дорожні карти (PRM) або оптимізація роїв частинок (PSO), спрямовані на генерацію відсутності зіткнень і найкоротші шляхи. Відповідно до реалізації, ці алгоритми можуть працювати виключно в автономному режимі або можуть оновлювати план, враховуючи нову інформацію, зібрану під час польоту: у цьому випадку наступний хід платформи ітеративно перераховується для досягнення кінцевої мети. Ці онлайн-методи розглядають динамічне та частково невідоме середовище, де дослідження нової частини простору поступово створює навколишнє 3D оточення на карті. Перешкоди зазвичай можна розділити на статичні та динамічні, і в багатьох підходах злиття двох різних типів шляху (глобального та локального) поєднується. Глобальна траєкторія польоту визначає найкращу можливу траєкторію відповідно до попередньої інформації про траєкторію, тоді як місцеві планувальники перераховують траєкторію, щоб уникнути можливих динамічних перешкод. Взаємодія між цими двома планувальниками може змінюватися залежно від реалізації, яка змінює спосіб уникнення перешкод дроном.

Відсутність цільового призначення дозволяє досліджувати абсолютно недосліджені середовища: загальна ідея полягає в тому, щоб максимізувати приріст інформації, що досягається при кожному русі дрона, визначаючи наступну точку зору на межі між відомим і недослідженим простором. Ці підходи, що базуються на кордоні, зазвичай призначені для створення 3D реконструкції, виявлення об'єктів або класифікації захоплених регіонів. Алгоритми, призначені для 3D реконструкції, часто називають наступним найкращим видом і оптимізують функцію з урахуванням таких параметрів, як 3D-невизначеність 3D-реконструйованих точок або повнота згенерованої хмари точок. Інші методи спрямовані на максимізацію охоплення певної території відповідно до функції корисності, яка враховує певний прибуток інформації, або збільшення знань про територію відповідно до певного обґрунтування, такого як наявність об'єкта інтересів по сусідству. Більшість ранніх реалізацій автономної навігації було реалізовано для дослідження замкнутих (у приміщенні) невідомих

середовищ, тоді як тенденція до зростання автономних підходів в даний час розробляється для відкритих просторів і більшої різноманітності завдань.

Розробка стратегій автономного водіння в поєднанні зі сплеском глибокого навчання передала кілька рішень (наприклад, стратегії уникнення зіткнень), які вплинули на розробку алгоритмів дронів. Традиційні алгоритми SLAM часто підтримуються мережами глибокого навчання, а 3D-реконструкція поєднується із семантичною сегментацією, щоб зменшити вплив динамічних об'єктів у сцени або поєднання геометричної та семантичної інформації, що взаємно покращує розуміння сцени тривимірного середовища. Створення швидких карт глибини завдяки оцінкам глибини одиночних і стереозображень, а також вдосконалення мініатюрних пристроїв GPU встановлені на дронах сприяли широкому застосуванню цих алгоритмів для виявлення перешкод і масштабування сцени, спираючись лише на зображення. З іншого боку, уникнення зіткнень може бути вбудовано в наскрізні рішення, керовані даними, де згорткові нейронні мережі (CNN) виводять такі команди, як швидкість і зміна напрямку від попередньо запланованого шляху, замість виведення простих карт глибини або щоб уникнути перешкод. Мережі, навчені спеціальними наборами даних, також показали, що слідують шляху, рейках або дорогах, динамічно уникаючи перешкод. Автономну навігацію за певними траєкторіями також можна навчити за допомогою навчання з підкріпленням, коли БПЛА отримує винагороду за дії, вжиті в навколишньому середовищі, які сприяють досягненню його цілей.

Більшість підходів було розроблено для внутрішнього (замкнутого) середовища або для дотримання структурних особливостей, таких як дороги чи річки, однак деякі повністю автономні рішення для навігації в захаращених середовищах також доступні. Різні фреймворки, такі як навчання з підкріпленням із змодельованих або реальних наборів даних або мережі, що вивчають часові залежності (такі як рекурентні нейронні мережі) для послідовностей зображень ще більше сприяють застосуванню цих методів у нових і більш складних програмах. Незважаючи на великий інтерес до всіх цих рішень, більшість із них все ще є «дослідницькими» рішеннями, тоді як для їх впровадження на ринок знадобиться проведення подальших суворих випробувань для оцінки та сертифікації їх безпеки в умовах експлуатації без нагляду [20].

1.6 Способи обробки даних БПЛА

І картографування в реальному часі, і автономна навігація передбачають здатність швидко обробляти отримані зображення та надійно передавати необхідну інформацію.

У зв'язку з цим, бортові обчислювальні та комунікаційні технології сприяють створенню інтелектуальних малих картографічних систем БПЛА. Комунікації між мобільними платформами або між платформами та інфраструктурою через Wi-Fi, радіо, а також периферійні обчислення служать для обміну завданнями та інформацією про обмін даними. Швидка (майже в реальному часі) обробка може бути досягнута за допомогою бортових спеціальних пристроїв або потокової передачі даних на віддалені комп'ютери.

У галузі аерофотограмметрії деякі перші приклади 3D-картографування в режимі реального часу були реалізовані на платформах БПЛА з метою забезпечення орієнтації зображень у реальному часі та створення грубих ортофотографій. Ці методи в основному адаптували робототехнічні підходи (такі як алгоритми SLAM) до фотограмметричних потреб з метою надання геометрично узгоджених карт, враховуючи швидкість обробки як пріоритет більше, ніж геометричну точність наданих результатів. Основне вузьке місце/проблема полягає в обмежених можливостях обробки, що обмежує кількість завдань, які можна виконувати на борту. Це часто пом'якшується шляхом використання наближених орієнтацій зображення, прийнятих у робототехніці, або спрощених проміжних продуктів (таких як використання розріджених хмар точок або грубих стратегій ущільнення хмар точок) у створенні ортофото.

Сплеск глибокого навчання нещодавно був викликаний розробкою ефективних графічних процесорів, здатних обробляти складні CNN за скорочений час. Окрім використання графічного процесора, встановленого на потужних робочих станціях, були розроблені мініатюрні рішення для обробки даних на межі (тобто поблизу датчика), щоб запобігти проблемам через обмежене або відсутність з'єднання в полі. Зобов'язання багатьох провідних компаній (таких як NVIDIA, Google і Intel) розробити відносно дешеві (кілька сотень євро) пристрої, які легко програмувати за допомогою звичайних бібліотек (наприклад, PyTorch, TensorFlow), сприяє їх широкому застосуванню на БПЛА, що розгортаються в полі. [21]. Розробка алгоритмів глибокого навчання для уникнення зіткнень і щільних реконструкцій ще більше розширює потенціал цих плат для традиційних цілей картографування [22]. Швидкість обробки даних цих пристроїв не порівнянна з системними платами високого класу для настільних комп'ютерів (майже на порядок повільніше), але кілька додатків із відносно неглибокими мережами можна виконувати в режимі реального часу [23]. Ці пристрої мають відносно низьке енергоспоживання порівняно з їх повнорозмірними версіями, і вони все ще можуть бути вбудовані в багато БПЛА [24]. Приклади цих методів наразі можна знайти в моніторингу доріг [25], управлінні надзвичайними ситуаціями,

виявленні транспортних засобів [26] і відстеження кількох об'єктів, серед інших. Зусилля щодо мініатюризації цих пристроїв призвели до розробки автономних платформ вагою кілька грамів, призначених для дослідницьких цілей. Ці рішення пронизують ринок із появою комерційних дронів, здатних відстежувати об'єкти, виконувати 3D-реконструкції та семантичне розуміння для рекреаційних цілей.

Хоча мініатюрні рішення зростають, бортові обчислення все ще мають певні обмеження, пов'язані зі швидкістю обробки, енергоспоживанням і рівнем налаштування дрона. Якщо невеликі затримки прийнятні, потокова передача даних на віддалений ПК для обробки все ще може бути варіантом. Кілька прикладів цих підходів було розроблено з використанням локальних мереж (таких як Wi-Fi) або радіозв'язків і хмарних механізмів, щоб забезпечити потік зображень, відео та геолокалізацію дрона. Традиційне з'єднання GSM (наприклад, 4G) також можна використовувати для відносно обмежених обсягів даних. Розробка комплекту розробки програмного забезпечення (SDK) для багатьох комерційних платформ також дозволила реалізувати ці рішення для різних програм. Поява зв'язку 5G може ще більше розширити ці типології архітектури в найближчому майбутньому. Подальші периферійні обчислення та рішення IoT (Інтернет речей) дозволять виконувати обчислення на джерелі даних або поблизу нього.

Спеціалізовані цілодобові рішення, спрямовані на спостереження та моніторинг великих промислових підприємств з використанням БПЛА, наразі пропонуються все більшою кількістю компаній у світі, які поєднують БПЛА зі спеціальними док-станціями для підзарядки/заміни батарей та обміну даними. У багатьох випадках БПЛА вважаються літаючими датчиками, вже вбудованими у великі мережі Інтернету речей: отримані зображення стають частиною більшої он-лайн обробки різнорідних даних у реальному часі. У цих рішеннях розробка гібридних стратегій із вбудованими обчисленнями та віддаленою обробкою в хмарі здається найефективнішим рішенням на даний момент. Майбутній розвиток завдяки швидшому зв'язку та меншим затримкам (тобто 5G), ймовірно, ще більше революціонізує спосіб обробки даних, підштовхнувши розробку рішень у режимі реального часу та нових послуг із використанням БПЛА.

1.7 Особливості глибокого навчання для БПЛА

Незважаючи на великий інтерес в останні роки, кількість опублікованих робіт, які використовують методи глибокого навчання для зображень БПЛА для семантичного аналізу сцени, все ще відносно обмежена. БПЛА пропонують високий потенціал і

додаткові проблеми для вилучення семантичної інформації завдяки дуже високій просторовій роздільній здатності та можливості збору 3D-даних. Документи з використанням глибокого навчання в основному зосереджувалися на (I) класифікації зображень, (II) семантичній сегментації (або класифікації по пікселях) і (III) виявленні об'єктів. Приклади класифікації зображень представлені для класифікації окремих видів дерев за допомогою RGB-зображень високої роздільної здатності, або для виявлення структурних пошкоджень будівель після землетрусів. Алгоритми сегментації використовуються для обробки даних БПЛА в широкому діапазоні застосувань, таких як картографування бур'янів над рисовими полями, семантична сегментація видів рослин із зображень БПЛА з високою роздільною здатністю, моніторинг гірничодобувної діяльності та визначення видимих кадастрових меж. Застосування в навколишньому середовищі, такі як виявлення ссавців в африканській савані, характеристика купи наносів, хвороби культур також можна знайти в літературі.

Використовувані мережі відрізняються за глибиною, кількістю параметрів, складністю, пам'яттю та обчислювальною вартістю. Більшість цих робіт починаються з популярних мереж комп'ютерного зору, які часто дають задовільні результати завдяки тонкому налаштуванню або невеликим модифікаціям обробки даних БПЛА. Мережі, подібні до Resnet, часто застосовуються в класифікації, тоді як мережі кодувальника-декодера [26], розширені згортки або їх комбінації, такі як ефективність DeepLab, часто використовуються в семантичній сегментації. Сучасні мережі Yolo [27] зазвичай використовуються як відправна точка для виявлення об'єктів. Це можна пояснити тим, що дані БПЛА мають вищу роздільну здатність порівняно з повітряними та супутниковими знімками. У зв'язку з цим традиційні методи машинного навчання, засновані на створенні функцій вручну, легко стагнуть в аналізі зображень найнижчої роздільної здатності субметрової та субдециметрової роздільної здатності, де просторово-контекстуальні характеристики несуть більшу частину дискримінаційної інформації. Глибоке навчання показало здатність підвищувати точність шляхом вивчення високорівневих глибших просторових особливостей. Це часто призводило до того, що отримані дані БПЛА можна порівняти зі звичайними програмами комп'ютерного зору, не звертаючи уваги на додаткову варіативність використовуваних датчиків, типологій отримання даних і додатків, що представляють інтерес.

Успішні готові моделі комп'ютерного бачення не завжди є найкращим рішенням для поставленого завдання, оскільки знання домену часто можуть допомогти розробити або адаптувати мережі до конкретних характеристик даних БПЛА та семантичної проблеми. Мережі комп'ютерного бачення розроблені для зображень RGB, але їх

потрібно відповідним чином модифікувати, щоб повністю використовувати спектральну інформацію, отриману мультиспектральними та гіперспектральними датчиками. Це може вимагати більш ніж простого збільшення кількості фільтрів на першому мережевому рівні [28], які часто не враховуються. У порівнянні з бортовими та супутниковими даними, просторова роздільна здатність БПЛА впливає на очікувану просторову автокореляцію та, отже, на необхідне сприйнятливє поле мережі. Тому очікується, що зображення з вищою роздільною здатністю потребують мереж із більшим полем зору фільтрів. Ще одна інформація, яка суттєво впливає на завдання аналізу (і проектування мереж глибокого навчання), – це геометрія отримання зображень БПЛА. На відміну від великих наборів даних комп'ютерного зору (які часто використовуються для попереднього навчання), де один і той самий об'єкт (наприклад, людина) може перебувати на різних відстанях від камери, на зображеннях БПЛА відстань до об'єкта інтересу, як правило, фіксована, хоча похилий і найнижчий вигляд мають велике значення для зовнішнього вигляду об'єктів. Це часто може зменшити варіабельність зовнішнього вигляду об'єктів і може надати важливу попередню інформацію про середній розмір об'єкта (у пікселях). У цих випадках можна віддати перевагу більш легким мережам, пристосованим до характеристик даних і поточної проблеми, щоб полегшити навчання, зменшити витрати на обчислення та покращити здатність до узагальнення.

Значною проблемою, яка часто стоїть перед використанням глибокого навчання для БПЛА, є відсутність вільно доступних наборів даних для навчання. Деякі нещодавно запроваджені тести намагаються пом'якшити цю проблему, але збір польових даних або оцифрування вручну є єдиним рішенням для створення репрезентативного навчального набору. Відсутність спеціалізованих наборів даних також обмежує використання багатозадачних методів навчання, незважаючи на їх високу продуктивність, досягнуту завдяки взаємозв'язку між завданнями.

У більшості отриманих даних дані БПЛА надають додаткову можливість отримувати 3D-інформацію у формі цифрових моделей або хмар точок. Ця додаткова інформація сприяє характеристиці просторових особливостей цільових об'єктів у 3D-середовищі. Багатообіцяючі результати були показані в [29], де 2.5D топографічні та 3D геометричні об'єкти інтегровані в класичний робочий процес машинного навчання для підвищення точності складних проблем класифікації. Ця інтеграція 2D і 3D інформації може призвести до підвищення продуктивності в багатьох інших програмах у майбутньому.

2 ОГЛЯД СИСТЕМ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ

2.1 Постановка завдання

Згідно технічному завданню (ТЗ) необхідно розробити геометричні характеристики визначеного об'єкта щодо автоматизації розпізнавання та програмні модулі для супроводження його розпізнавання системами комп'ютерного зору БПЛА.

Найбільш перспективним напрямом використання нейронних мереж у задачах розпізнавання об'єктів є можливість використання невеликої комп'ютерної системи з функцією машинного зору у складі безпілотного літального апарату.

2.2 Історія розвитку комп'ютерного зору

Багато в чому історія комп'ютерного зору – це історія штучного інтелекту. Обидві дисципліни імітують біологічні процеси, засновані на розумінні того, як працює мозок, і кожна з них просунулася вперед завдяки появі штучних нейронних мереж, досконаліших обчислювальних ресурсів та великих даних.

Багато чого з того, що ми знаємо сьогодні про зорове сприйняття, прийшло до нас завдяки нейрофізіологічним дослідженням, проведеним на кішках у 1950-х та 1960-х роках. Вивчаючи, як нейрони реагують на різні стимули, вчені помітили, що людський зір ієрархічний. Нейрони виявляють прості ознаки, такі як краї, потім реагують на більш складні ознаки, такі як форми, та потім, зрештою, реагують більш складні візуальні уявлення.

Озброївшись цими знаннями, вчені-комп'ютерники зосередилися на відтворенні неврологічних структур людини у цифровій формі. Подібно до своїх біологічних аналогів, системи комп'ютерного зору використовують ієрархічний підхід до сприйняття та аналізу візуальних стимулів.

Область штучного інтелекту була заснована на літньому семінарі, проведеному на території Дартмутського коледжу в 1956, коли вчені об'єднали кілька розрізнених областей для уточнення та розвитку ідей про мислячі машини. "Це було перше свідчення інституційної тенденції до самовпевнених прогнозів, яка від початку вразила область штучного інтелекту", – пояснює розробник машинного навчання в Motion Metrics, що спеціалізується на глибоких нейронних мережах.

Комп'ютерний зір всерйоз зародився у 1960-х роках в університетах, які

розглядали цей проєкт як сходинку до штучного інтелекту. Перші дослідники були надзвичайно оптимістичні щодо майбутнього цих суміжних областей і просували ШІ як технологію, яка може змінити світ. Дехто передбачав, що протягом одного покоління буде створено машину, таку ж інтелектуальну, як людина. Галас приніс дослідникам мільйони доларів державного та приватного фінансування. По всьому світу з'явилися дослідні центри. Однак міжнародні зусилля по розробці ШІ були зведені нанівець через те, що вони не виправдали надій, що покладалися на них.

Завдяки своєму величезному оптимізму дослідники покладали на громадськість надмірно високі надії, не розуміючи при цьому складності завдання, яке вони поставили перед собою. Коли обіцяні результати не виправдали надій, область зазнала різкої критики і зіткнулася з серйозними фінансовими труднощами.

Перші обчислювальні ресурси не могли технічно відповідати складності проблем, які ставили перед собою вчені, і навіть найвражаючі проєкти вирішували лише тривіальні завдання. Більш того, більшість дослідників працювали в ізольованих групах і не мали наукової підтримки для значного розвитку цієї галузі.

У 1966 році американський комп'ютерний вчений і співзасновник лабораторії штучного інтелекту Массачусетського технологічного інституту Марвін Мінський отримав літній грант, щоб найняти першокурсника Джеральда Сассмана, який мав провести літо, з'єднавши камеру з комп'ютером і змусивши комп'ютер описати те, що він бачить. "Нема потреби говорити, що Сассман не вклався вчасно". Зір виявився однією з найскладніших проблем, що розчаровують, у штучному інтелекті протягом наступних чотирьох десятиліть". Як одного разу зауважив експерт з машинного зору Бертольд Хорн: "Сассман вирішив більше ніколи не працювати в галузі зору".

До середини 1970-х років уряди та корпорації втратили віру у ШІ. Фінансування закінчилося, і період, що пішов за цим, став відомий як "Зима штучного інтелекту". Хоча у 1980-х та 1990-х роках спостерігалось невелике відродження, штучний інтелект в основному був віднесений до галузі наукової фантастики, і серйозні вчені-комп'ютерники уникали цього терміна.

З появою Інтернету комп'ютерники отримали доступ до більшої кількості даних, ніж будь-коли раніше. Обчислювальне обладнання продовжувало удосконалюватися у міру зниження вартості. Рудиментарні нейронні мережі та алгоритми, розроблені у 1980-90-х роках, удосконалювалися. Зараз, коли штучному інтелекту вже понад півстоліття, в 2012 році в галузі штучного інтелекту нарешті стався прорив у вирішенні задачі візуального розпізнавання ImageNet Large Scale Visual Recognition Challenge (ILSVRC).

ILSVRC – це щорічне змагання з класифікації зображень, в якому дослідницькі групи оцінюють власні алгоритми на заданому наборі даних, а потім змагаються у досягненні вищої точності при вирішенні кількох задач візуального рапізнання. У 2010-2011 роках коефіцієнт помилок переможців ILSVRC коливався близько 26%. Потім, у 2012 році, команда з Університету Торонто представила глибоку нейронну мережу “AlexNet”, яка змінила гру для проєктів у галузі штучного інтелекту та комп’ютерного зору (рис. 2.1) [10].

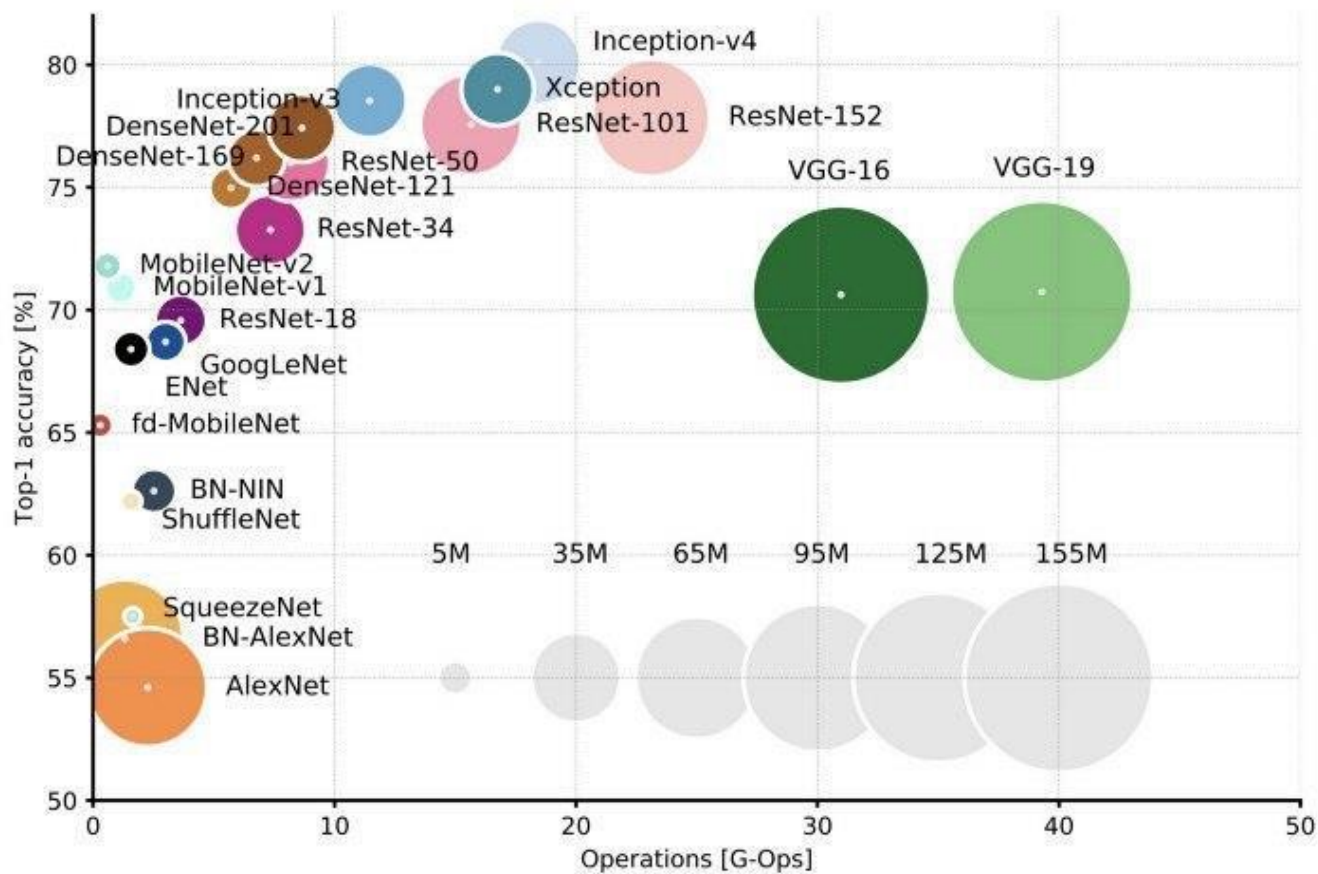


Рисунок 2.1 – Графік точності існуючих мереж відносно кількості операцій

Це була загорткова нейронна мережа (CNN) для класифікації зображень. AlexNet складається з 5 загорткових шарів і 3 повністю зв'язаних шарів.

Глибокі нейронні мережі зробили революцію у сфері штучного інтелекту. AlexNet досягла рівня помилок у 16,4%, і в наступні роки рівень помилок у ILSRVC знизився до кількох відсотків; тепер глибокі нейронні мережі є золотим стандартом для розпізнавання зображень. Ці здобутки проклали шлях для проникнення штучного інтелекту в Кремнієву долину.

Штучний інтелект вже органічно увійшов до багатьох аспектів нашого повсякденного життя. В останні роки штучний інтелект досяг величезного успіху в багатьох областях досліджень. Ігрові системи, такі як AlphaGo, використовують підкріплююче навчання для навчання новим стратегіям. Слухові апарати

використовують алгоритми глибокого навчання для фільтрації навколишнього шуму. Ці технології використовуються навіть у системах обробки та перекладу природної мови, розпізнавання об'єктів та зіставлення образів, які ми сприймаємо як належне у Google, Amazon, iTunes та подібних сервісах.

Ця тенденція не виявляє ознак уповільнення існує безліч дрібних завдань, які ми можемо автоматизувати, щоб звільнити наш час. Хоча ми досягли неймовірних успіхів у галузі штучного інтелекту, ми все ще маємо бути реалістами щодо його застосування для комп'ютерного зору – мине ще багато часу, перш ніж комп'ютери зможуть інтерпретувати зображення так само добре, як це робить людина.

У короткостроковій перспективі ймовірніше, що штучний інтелект буде використовуватися для доповнення та розширення людських можливостей.

Сектор охорони здоров'я є очевидним прикладом позитивної взаємодії людини та машини. Дещо історичної перспективи дозволяє припустити, що такі інновації, як протезування кори головного мозку, стануть мейнстрімом. Ці футуристичні пристрої стануть такими ж повсюдними і звичайними, як технічні кросівки для бігу або окуляри за рецептом.

Інтерпретація цифрових зображень і відео стане в нагоді в багатьох областях. Наприклад.

Медична діагностика. Класифікація зображень та розпізнавання образів широко використовуються для розробки програмних систем, які допомагають лікарям діагностувати небезпечні захворювання, такі як рак легенів. Група дослідників навчила систему штучного інтелекту аналізувати комп'ютерні томограми онкологічних пацієнтів. Алгоритм показав 95% точності, тоді як людина – лише 65%.

Керування виробництвом. Важливо виявляти дефекти на виробництві з максимальною точністю, але це непросто, оскільки часто потрібний контроль мікромасштабу. Наприклад, коли необхідно перевірити нарізку різьблення сотень тисяч гвинтів. Система комп'ютерного зору використовує дані з камер у режимі реального часу та застосовує алгоритми ML для аналізу потоків даних. Таким чином можна легко знайти недоброякісні товари.

Роздрібна торгівля. Amazon стала першою компанією, що відкрила магазин, що працює без касирів та касових машин. Amazon Go має сотні камер комп'ютерного зору. Ці пристрої відстежують, які товари покупці кладуть у візки. Камери також можуть відстежувати, чи покупець повертає товар на полицю і чи прибирає він його з віртуального кошика. Плата з покупців стягується через додаток Amazon Go, що позбавляє їх необхідності стояти в черзі. Камери також запобігають крадіжкам у

магазинах та відсутності товару.

Системи безпеки. Розпізнавання осіб використовується на підприємствах, школах, громадському транспорті та, скрізь, де важлива безпека. У школах США технологія розпізнавання обличчя використовується для зниження потенційної загрози, виявлення сексуальних, фізичних та інших злочинців. Таке програмне забезпечення також може розпізнавати зброю для запобігання актам насильства у школах. Тим часом деякі авіакомпанії використовують розпізнавання осіб для ідентифікації та реєстрації пасажирів, заощаджуючи час та знижуючи вартість перевірки квитків.

Охорона тварин. Екологам вигідно використовувати комп'ютерний зір для отримання даних про дику природу, включаючи відстеження переміщень рідкісних видів, моделей їхньої поведінки тощо, не турбуючи тварин. CV підвищує ефективність та точність аналізу зображень для наукових відкриттів.

Самокеровані автомобілі. За допомогою датчиків та камер автомобілі навчилися розпізнавати бампери, дерева, стовпи та припарковані автомобілі навколо себе. Комп'ютерний зір дозволяє їм вільно переміщатися у навколишньому середовищі без контролю з боку людини.

2.3 Нейронні мережі для мобільних пристроїв

В Нейронні мережі, спеціально розроблені для мобільних пристроїв, являються наступним поколінням еволюції ШІ. Нарешті, завдяки подальшому розвитку виконання логічних висновків стало доступним на пристроях із низьким енергоспоживанням і пам'яттю. Представлено методи подальшого покращення продуктивності під час виконання, зберігаючи точність.

SqueezeNet була випущена в 2016 році. Представляє новий будівельний блок під назвою пожежний модуль, який складається з 1×1 згортковий шар, щоб зменшити кількість каналів з мінімальною кількістю параметрів з наступним поєднанням 1×1 і 3×3 звивини, що утворюють розширений блок, який знову збільшує розмір глибини. Крім того, цей модуль поєднується з проектним рішенням відкладеної дискретизації, за допомогою якої роздільна здатність карт функцій зменшується пізніше в мережі для підвищення точності.

SqueezeNext являє собою модернізацію вищезгаданої мережі, але вносить ряд змін. По-перше, використовується більш агресивне зменшення каналу шляхом включення двоступеневого модуля стискання. Це значно зменшує загальну кількість параметрів, що використовуються для згорток 3×3 . По-друге, він використовує

роздільні згортки 3×3 для подальшого зменшення розміру моделі та видаляє додаткову гілку 1×1 після модуля стискання. По-третє, мережа використовує поелементне з'єднання пропуску додавання, подібне до архітектури ResNet.

ShuffleNet використовує дві нові операції, поточкову групову згортку та перетасування каналів, щоб зменшити витрати на обчислення, зберігаючи при цьому точність.

MobileNet з'явилася у квітні 2017 року. Її побудовано на концепції роздільних по глибині згорток, вперше представлених і популяризованих мережею Xception. Цей обчислювальний модуль складається з одного першого рівня ядер, який працює на кожному вхідному каналі незалежно, глибинної згортки (рис. 2.2),

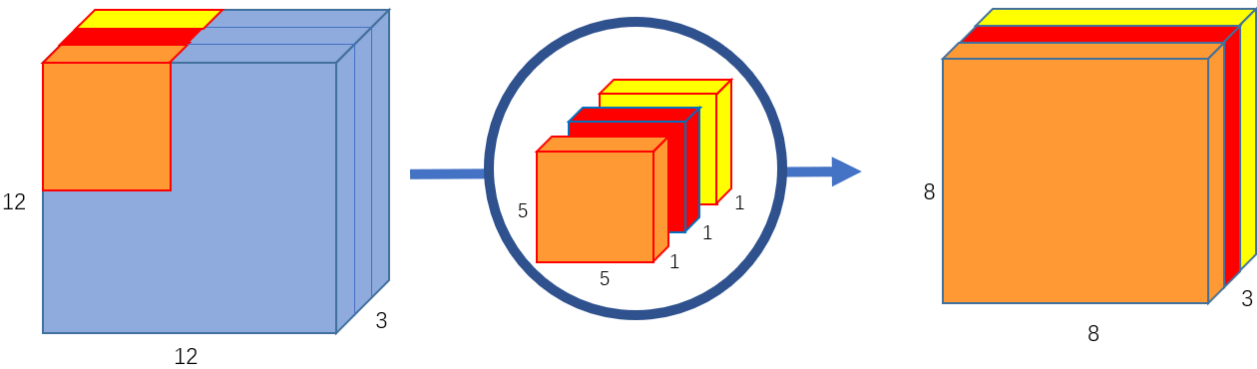


Рисунок 2.2 – Операція згортки у глибину

за якою слідує другий, точковий (рис. 2.3), який використовує ядро 1×1 : ядро, яке повторює кожну точку. Це ядро має глибину стільки ж каналів, скільки має вхідне зображення.

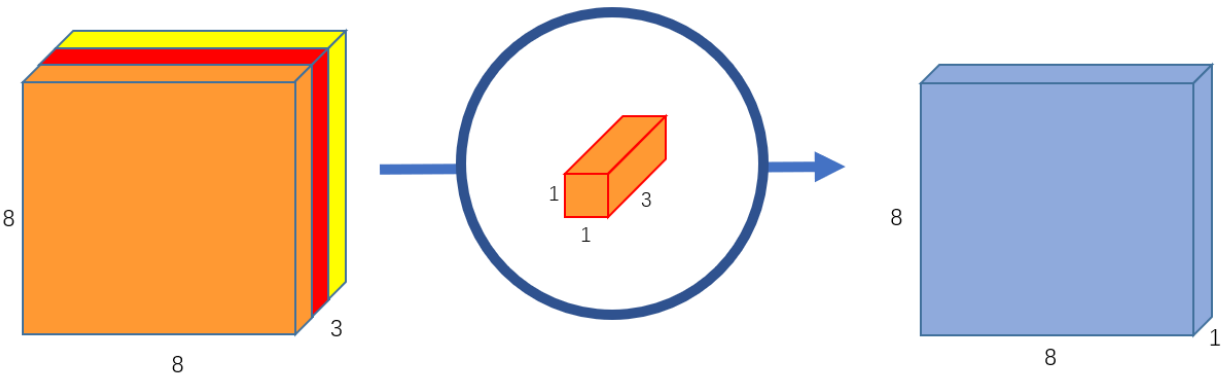


Рисунок 2.3 – Операція точкової згортки

Крім того, у MobileNet є гіперпараметр під назвою множник ширини для контролю

кількості вихідних каналів у кінці кожного блоку та загальної кількості параметрів [30] [31].

MobileNetV2 з'явилась через рік. V2 було додано лінійні шари вузьких місця та інвертовані залишки. Залишкові блоки з'єднують початок і кінець згорткового блоку зв'язком пропуску. Додавши ці два стани, мережа має можливість отримати доступ до попередніх активацій, які не були змінені в згортковому блоці. Цей підхід виявився необхідним для побудови мереж великої глибини. вихідний залишковий блок дотримується підходу широкий->вузький->широкий або вузький->широкий->вузький (рис. 2.4) щодо кількості каналів. Автори описують цю ідею як перевернутий залишковий блок, оскільки між вузькими частинами мережі існують з'єднання пропуску, що є протилежним тому, як працює початкове залишкове з'єднання. З перевернутими залишковими блоками ми стискаємо шари, де з'єднані пропуски з'єднання. Це погіршує продуктивність мережі. Автори представили ідею лінійного вузького місця, де остання згортка залишкового блоку має лінійний вихід перед тим, як її додати до початкових активацій.

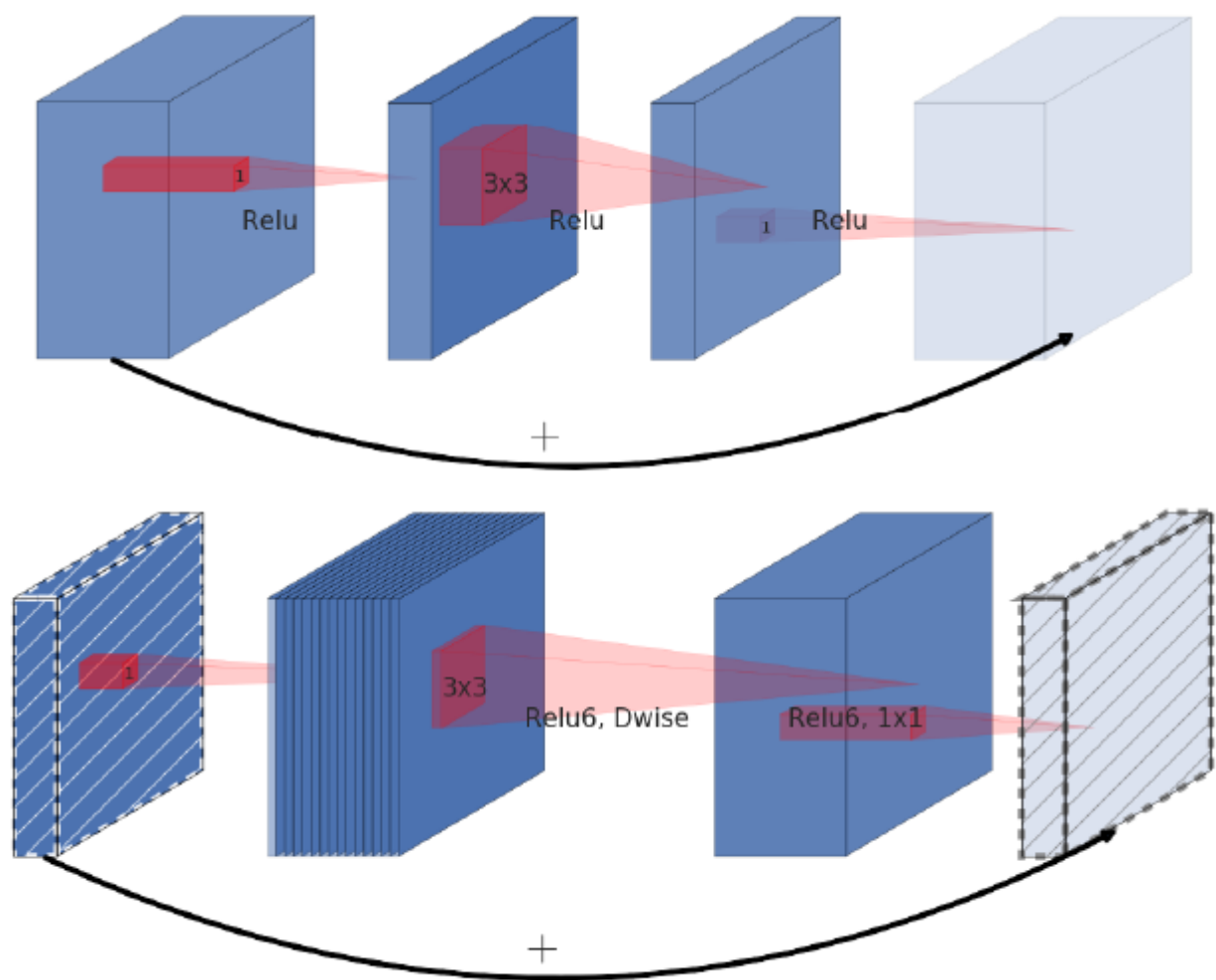


Рисунок 2.4 – Еволюція роздільних блоків згортки

Пошук нейронної архітектури (NAS) – це нова плідна галузь, метою якої є

використання алгоритмів пошуку, як правило, методів підкріпленого навчання або еволюційної оптимізації, щоб знайти комбінацію модулів, яка забезпечує оптимальний компроміс між затримкою та точністю. Наприклад, метрику оптимізації можна вказати у формі багатоцільової функції Парето, як показано в MnasNet. FBNet, ChamNet і MobilenetV3 є іншими прикладами послідовних внесків у пошук архітектури мобільних мереж. Читач може звернутися до для ретельного аналізу цієї техніки.

EfficientNet покращує можливість ручного налаштування складності моделі, дозволяючи користувачам вибрати бажаний компроміс між ефективністю та точністю. Слід зазначити, що автори пропонують комплексний гіперпараметр масштабування, який налаштовує глибину мережі, тобто кількість шарів або блоків, ширину, тобто кількість каналів, і розмір вхідного зображення, що впливає на роздільну здатність карт внутрішніх функцій, все це одразу оптимальним чином. Цей метод масштабування впливає зі спостереження, що розміри мережі незалежно не впливають на компроміс між затримкою та точністю. Нарешті, ця нова концепція масштабування застосовується до базової мережі, знайденої за допомогою багатоцільової оптимізації та простору пошуку, визначеного в MnasNet.

3 РОЗРОБЛЕННЯ СИСТЕМИ РОЗПІЗНАВАННЯ ОБ’ЄКТІВ

3.1 Структура згорткових нейронних мереж

Згорткові нейронні мережі (CNN) є найдосконалішими штучними нейронними мережами [22], які використовуються для високоточного виявлення об’єктів у цьому десятилітті.

Принцип роботи CNN дуже схожий на те як працює людське око (рис. 3.1). В загальних рисах згорткові мережі складаються з двох основних частин: екстракторів ознак і класифікатора. Екстрактори ознак складаються із згорткових шарів у якому згортковий фільтр проходить уздовж вхідної матриці і шарів максимального об’єднання, зменшення матриці (або матриць), створеної попереднім згортковим шаром, до меншої матриці. Об’єднання зазвичай передбачає взяття максимального або середнього значення в об’єднаній області. Операція об’єднання, як і операція згортки, ділить матрицю на фрагменти, а потім згортає цю операцію кроками . Наприклад, припустимо, що операція об’єднання ділить згортку на фрагменти 2x2 із кроком 1x1.

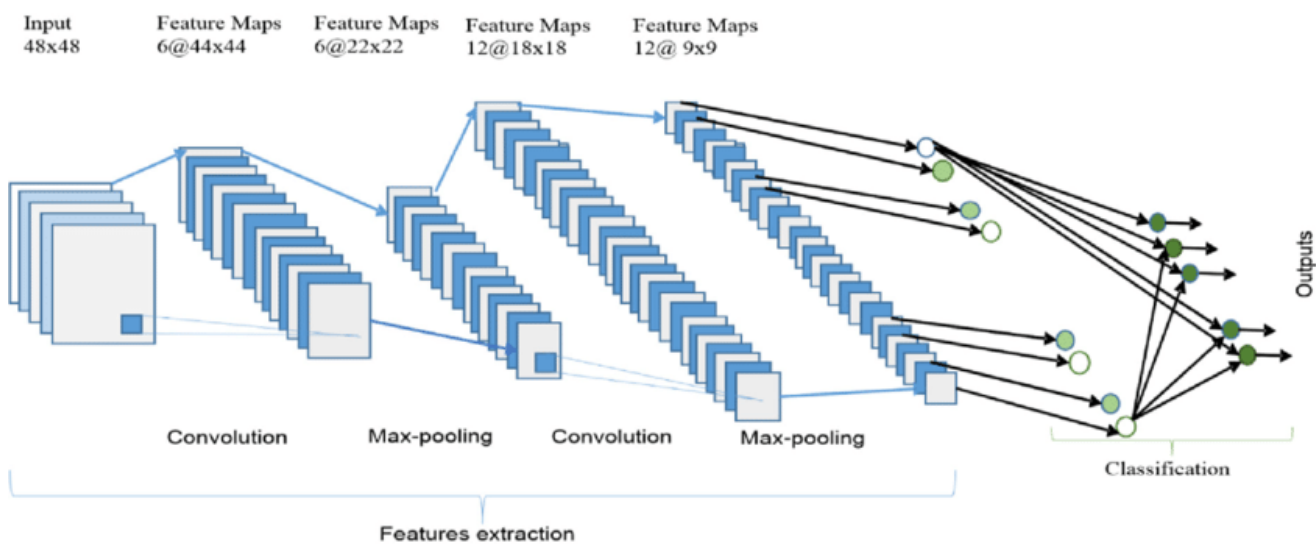


Рисунок 3.1 – Загальна архітектура CNN

Як показано на рисунку 3.2, відбуваються чотири операції об’єднання. Вихідні дані останнього рівня CNN використовуються як вхідні дані для повністю підключеної мережі, яка є класифікаційним рівнем.

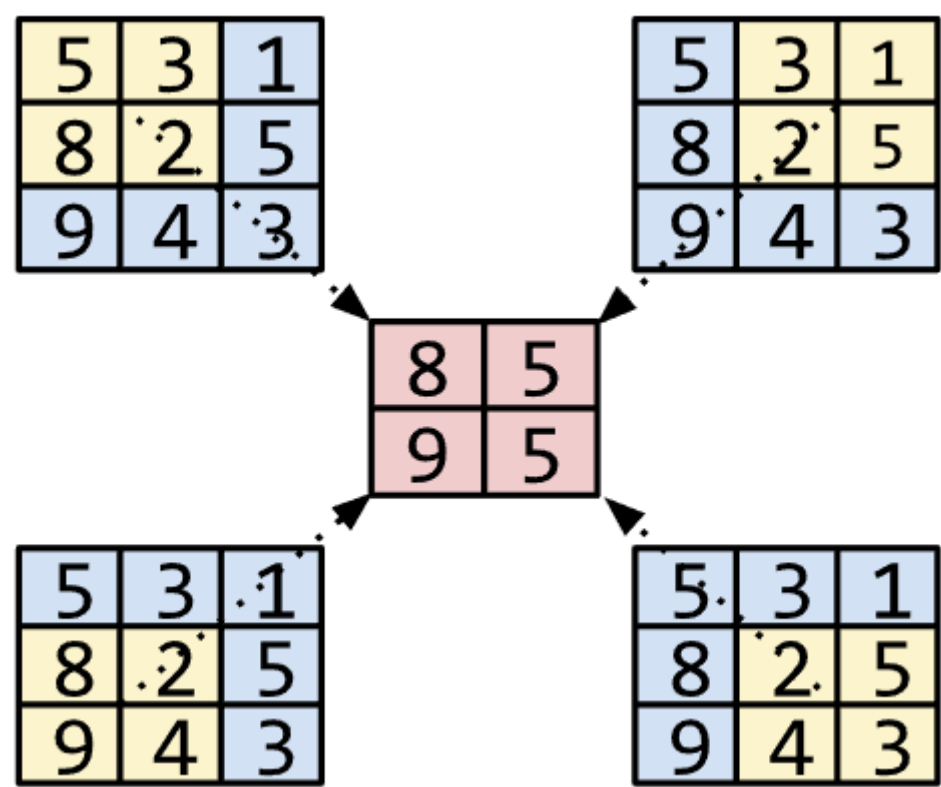


Рисунок 3.2 – Процес об’єднання

Об’єднання допомагає забезпечити трансляційну інваріантність у вхідній матриці, тобто здатність алгоритму успішно класифікувати зображення, навіть коли положення об’єктів на зображенні змінюється. Наприклад, алгоритм все одно може ідентифікувати собаку, незалежно від того, чи знаходиться вона в центрі кадру чи в лівому кінці кадру.

3.2 Середовище навчання нейронної мережі

Щоб впорядкувати процес навчання, було використано операційну систему Windows 10 x64, хоча цей процес можна виконувати в кількох середовищах, середовище від Microsoft було вибрано через сумісність між деякими програмами та фреймворками, які ми збираємося описати у наступних розділах.

3.2.1 TensorFlow

Фреймворк Tensorflow розроблено командою Google Brain Team і поєднує кілька методів обчислювальної алгебри з оптимізацією обчислення нейронної мережі моделі, таким чином полегшуючи обчислення різних функцій, необхідних моделям, виконання яких за інших обставин вимагатиме багато часу. Він підтримує таке:

- визначення, оптимізація та ефективне числення математичних операцій з багатовимірними масивами (введення концепції «тензора»);
- підтримка більшості методів, які використовуються в навчанні нейронних

мереж, наприклад, регресійних функцій, шарових моделей, оптимізаторів, відсіву тощо;

– операції обчислення виконуються автономно для користувача, тому Tensorflow аналізуватиме те, що обчислюється на центральному процесорі, а що обчислюватиметься за допомогою графічного процесора.

Усі ці функції є ключовими функціями, завдяки яким Tensorflow стає винятковою бібліотекою в процесі розробки нейронної мережі, зосередженою на важливих частинах створення, модифікації та навчання обчислювальних моделей.

Іншим ключовим фактором фреймворку є безкоштовний і легкий доступ до величезної кількості протестованих моделей, які підтримуються як командою Tensorflow, так і спільнотою. Моделі доступні в репозиторії Github

Це основа інших фреймворків, які забезпечують високий рівень абстракції своїх методів і додають деякі додаткові функції, які можуть бути дійсно корисними залежно від програми. TensorFlow працює з багатовимірними масивами або тензорами, представленими у вигляді об'єктів `tf.Tensor`. Тензори – це багатовимірні масиви з єдиним типом (називається `dtype`).

Його популярність пояснюється простотою роботи з декількома раніше розробленими моделями нейронних мереж, що забезпечує різні методи навчання залежно від наших вимог на додаток до величезної кількості функцій і методів, до яких ми можемо отримати доступ для зміни або створення власної моделі.

3.2.2 Keras

Говорячи про Keras, це високорівневий API (Прикладний програмний інтерфейс), який може використовувати функції TensorFlow (також інші бібліотеки ML, такі, як Theano). Keras був розроблений із зручністю та модульністю як керівні принципи. З практичної точки зору Keras дозволяє реалізувати безліч потужних, але найчастіше складних функцій TensorFlow максимально просто, до того ж він налаштований для роботи з Python без серйозних змін або налаштувань.

3.2.3 TensorFlow Lite

TensorFlow Lite – це набір інструментів, який забезпечує машинне навчання на пристрої, допомагаючи розробникам запускати свої моделі на мобільних, вбудованих та периферійних пристроях.

Ключова особливість:

– оптимізовано для машинного навчання на пристрої за рахунок усунення 5 ключових обмежень: затримка (немає можливості звернення до сервера);

конфіденційність (особисті дані не залишають пристрій), можливість підключення (підключення до Інтернету не потрібно), розмір (зменшена модель та двійковий розмір) та енергоспоживання (ефективний висновок та відсутність мережових підключень);

– підтримка декількох платформ, включаючи пристрої Android та iOS, вбудований Linux та мікроконтролери;

– підтримка різних мов, включаючи Java, Swift, Objective-C, C++ та Python;

– висока продуктивність з апаратним прискоренням та оптимізацією моделі;

– наскрізні приклади для загальних завдань машинного навчання, таких як класифікація зображень, виявлення об'єктів, оцінка пози, відповіді питання, класифікація тексту тощо на кількох платформах.

3.2.4 Tensorflow detection API

Щоб розробити програму, було використано Tensorflow detection API.

Причиною такого рішення є тип необхідної програми. Це найкращий варіант завдяки тому, що він працює над Tensorflow і надає кілька прикладів і попередньо створених моделей, зручних для проектів виявлення об'єктів.

3.2.5 Google Colaboratory

Google Colaboratory, або також відомий як Google Colab являється хмарним сервісом від Google Research. Це IDE, яка дозволяє будь-якому користувачеві писати вихідний код у своєму редакторі та запускати його з браузера. Зокрема, з 2021 року він підтримує мову програмування Python 3.6.9 та орієнтований на завдання машинного навчання, аналіз даних, навчальні проекти тощо. Також з головних плюсів Colab є створення моделей TensorFlow та підтримка TensorBoard з «коробки».

Це середовище засноване на Jupyter Notebook, та працює абсолютно безкоштовно за допомогою облікового запису Gmail, і він не вимагає налаштування, а також не потрібно завантажувати чи встановлювати Jupyter. Очевидно, що як щось безкоштовне, Google Colaboratory не має необмежених ресурсів і не гарантує їх, але вони змінюються залежно від використання, яке надається системі.

Важливо зазначити, що коли ви отримуєте доступ до Colab за допомогою свого облікового запису, ви отримуєте віртуальну машину, на якій ви можете запускати свій код, ізольований від інших користувачів і ресурсів. Тому ви можете відновити віртуальну машину до початкового стану, якщо у вас виникнуть проблеми. Це також означає, що якщо ви виконаєте деякий код у своїй віртуальній машині та закриваєте браузер, машини будуть видалені після певного періоду бездіяльності, щоб звільнити

ресурси. Однак ви матимете свої блокноти в GDrive, якщо ви їх зберегли, або ви зможете завантажити їх локально (формат Jupyter з відкритим кодом .ipynb).

3.3 Виявлення об'єкту

За наявності зображення або відеопотоку модель виявлення об'єктів може визначити, які з відомого набору об'єктів можуть бути присутніми, та надати інформацію про їх позиції на зображенні.

Наприклад, на рисунку 3.3 показано, як було розпізнано два об'єкти та анотовано їх положення.

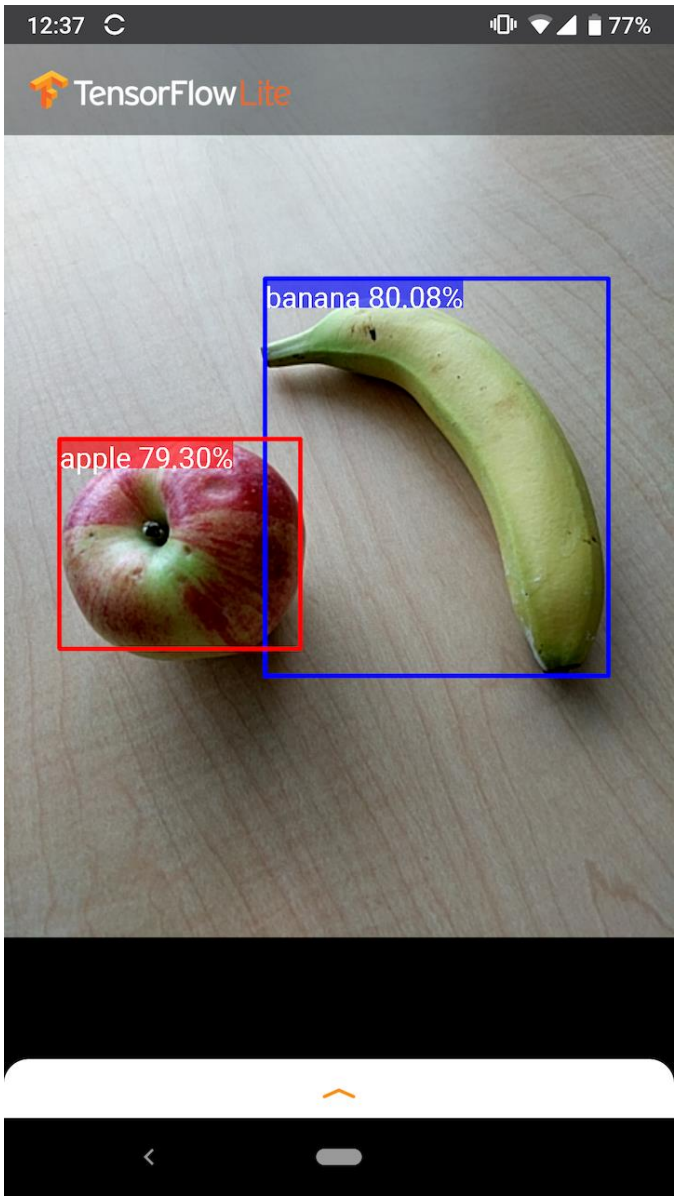


Рисунок 3.3 – Розпізнавання об'єктів

Модель виявлення об'єктів навчена виявляти присутність та місце розташування кількох класів об'єктів. Наприклад, модель можна навчати із зображеннями, що містять різні шматочки фруктів, разом із міткою, що вказує клас фруктів, які вони

представляють (наприклад, яблуко, банан або полуниця), та даними, що вказують, де кожен об'єкт з'являється у зображенні. зображення.

Коли зображення згодом надається моделі, вона виводить список виявлених об'єктів, розташування рамки, що містить кожний об'єкт, і оцінку, що вказує на впевненість у тому, що виявлення було правильним.

Вхідний підпис

Модель приймає зображення як вхідні дані.

Припустимо, що очікуване зображення має розмір 300×300 пікселів із трьома каналами (червоний, синій та зелений) на піксель. Це має бути передано до моделі у вигляді згладженого буфера з 270 000 байтових значень (300x300x3). Якщо модель квантована, кожне значення має бути одним байтом, що становить від 0 до 255.

Модель виводить чотири масиви, що відображаються на індексах 0–4. Масиви 0, 1 і 2 описують N виявлених об'єктів, причому кожному об'єкту відповідає один елемент у кожному масиві (таблиця 3.1).

Таблиця 3.1 – Розпізнавання об'єктів

Індекс	Ім'я	Опис
0	Розташування	Багатовимірний масив [N][4] значень із плаваючою комою від 0 до 1, внутрішні масиви представляють обмежувальні рамки у формі [верхній, лівий, нижній, правий]
1	Класи	Масив з N цілих чисел (виведених як значення з плаваючою комою), кожне з яких вказує на індекс мітки класу з файлу міток
2	Результати	Масив із N значень із плаваючою комою від 0 до 1, що представляє ймовірність того, що клас було виявлено
3	Число виявлень	Ціле значення N

Наприклад, уявіть, що модель навчена виявляти яблука, банани та полуницю. При наданні зображення буде виведено задану кількість результатів виявлення у цьому прикладі наведено нижче (таблиця 3.2).

Таблиця 3.2 – Приклад класифікації

Клас	Оцінка	Розташування
Яблуко	0,92	[18, 21, 57, 63]
Банан	0,88	[100, 30, 180, 150]
Полуниця	0,87	[7, 82, 89, 163]
Банан	0,23	[42, 66, 57, 83]
Яблуко	0,11	[6, 42, 31, 58]

3.3.1 Оцінка достовірності

Щоб інтерпретувати ці результати, ми можемо подивитися на оцінку та розташування кожного виявленого об'єкта. Оцінка є числом від 0 до 1 і також має таку назву як Перетин через об'єднання (IoU), яке вказує на впевненість у тому, що об'єкт дійсно виявлений. Чим ближче число до 1, тим впевненіша модель.

Цей поріг відсікання можна встановити власноруч, та залежить від того що необхідно: хибні спрацьовування (неправильно ідентифіковані об'єкти або області зображення, які помилково ідентифікуються як об'єкти, хоча насправді такими не є) або хибнонегативні (справжні об'єкти, які пропустили, тому що їхня впевненість була низькою).

Наприклад, на рисунку 3.4 груша (яка не є об'єктом, для виявлення якого модель була навчена) була помилково ідентифікована як "людина". Це приклад хибнопозитивного результату, який можна було б ігнорувати, вибравши відповідну відсічку. У цьому випадку граничне значення 0,6 (або 60%) зручно виключає хибнопозитивний результат.

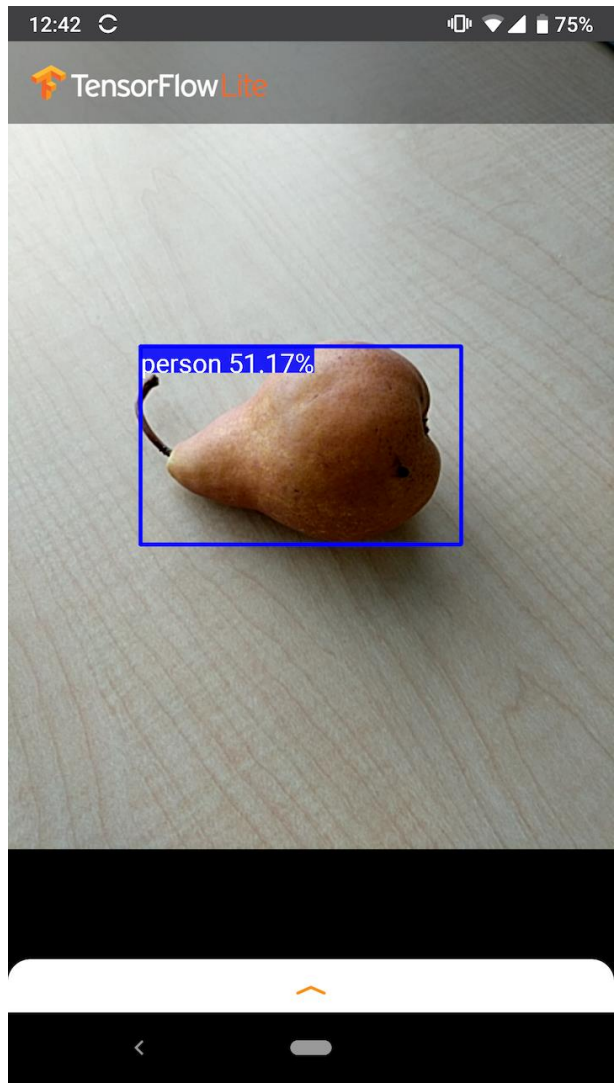


Рисунок 3.4 – Приклад неправильно ідентифікованого об'єкту

3.3.2 Розташування

Для кожного виявленого об'єкта модель поверне масив з чотирьох чисел, що представляють прямокутник, що обмежує, навколишній його положення. Для наданої початкової моделі номери розташовані послідовно так:

- зверху;
- зліва;
- знизу;
- справа.

Верхнє значення представляє відстань верхнього краю прямокутника від верху зображення в пікселях. Ліве значення представляє відстань лівого краю від лівого краю вхідного зображення. Інші значення представляють нижній і правий краї подібним чином.

3.4 Навчання мережі

Для того щоб мережа розпізнавала об'єкти її потрібно навчити, потрібно зібрати

свої дані та оформити їх так, щоб нейромережа могла навчатися на них. Це включає збір зображень та їх маркування. Підготовка даних – це мистецтво саме собою, пов'язане з вирішенням таких проблем, як пропущені значення, пошкоджені дані, дані в неправильному форматі, неправильні мітки тощо. Зазвичай чим більший набір даних (Dataset) тим більш точніша буде майбутня мережа, тому найчастіше використовують вже готові набори даних для розпізнавання об'єктів: COCO, ImageNet, Open Images та ін.

Після підготовки або вибору набору даних потрібно вибрати модель нейронної мережі яка буде використовуватись в роботі. З огляду в попередніх частинах роботи було вибрано нейронні мережі MobileNet та EfficientDet Lite (таблиця 3.3) як найбільш відповідні до поставлених умов.

Таблиця 3.3 – Характеристики нейронних мереж

Архітектура моделі	Розмір, Мб	Час кадру, ms	mAP, %
EfficientDet-Lite0	4.4	37	25.69
EfficientDet-Lite1	5.8	49	30.55
EfficientDet-Lite2	7.2	69	33.97
MobileNetV1	6.7	24	20.2

Процес навчання залежить від обладнання на якому воно проходить та може займати великий проміжок часу, тому для прискорення цього процесу використовують GPU який істотно зменшує час навчання. Також швидкість навчання залежить від кількості епох навчання, чим більше епох пройдено тим більш ефективною стає мережа, але збільшується шанс перенавчити мережу (розглянуто в 3.3.1).

3.5 Донавчання мережі

Вибрані мережі були навчені на наборі даних COCO 2017 тому не дуже підходять для необхідних цілей. Щоб не перенавчати мережу заново та зберегти відповідну точність TensorFlow має бібліотеку яка спрощує адаптацію нових даних використовується бібліотека Model Maker. Ця бібліотека використовує трансферне навчання, щоб зменшити кількість необхідних навчальних даних та скоротити час навчання всього за кілька рядків коду.

Для цього потрібно виконати декілька рядків в Colab:

- встановити необхідні пакети;

```

!sudo apt -y install libportaudio2
!pip install -q --use-deprecated=legacy-resolver tf-lite-model-maker
!pip install -q pycocotools
!pip install -q opencv-python-headless==4.1.2.30
!pip uninstall -y tensorflow && pip install -q tensorflow==2.8.0
    – імпортувати необхідні пакети;
import numpy as np
import os
from tf_lite_model_maker.config import QuantizationConfig
from tf_lite_model_maker.config import ExportFormat
from tf_lite_model_maker import model_spec
from tf_lite_model_maker import object_detector
import tensorflow as tf
assert tf.__version__.startswith('2')
tf.get_logger().setLevel('ERROR')
from absl import logging
logging.set_verbosity(logging.ERROR)
    – підготувати набір даних, який надається у форматі CSV;
    – вибрати архітектуру моделі виявлення об'єктів;
spec = model_spec.get('efficientdet_lite1')
    – завантажити датасет;
train_data, validation_data, test_data = object_detector.DataLoader.from_csv('gs://cloud-ml-
data/img/openimage/csv/drone_ml_use.csv')
    – прооНавчання TensorFlow моделі з навчальними даними;
model = object_detector.create(train_data, model_spec=spec, batch_size=8,
train_whole_model=True, validation_data=validation_data)
    – оцінка моделі з тестовими даними;
model.evaluate(test_data)
    – експорт моделі як TensorFlow Lite.

```

3.6 Встановлення моделі на Android смартфон

Для розпізнавання об'єктів за допомогою нейронної мережі був вибраний смартфон на базі ОС Android як тестова платформа. Головними критеріями вибору були доступність, простота встановлення.

Для встановлення було на ПК встановлено Android Studio версії Dolphin. В Android Studio було імпортовано та запущено проєкт, вибрано файл build.gradle, змінено код, додано файл моделі .tflite до каталогу src/main/assets та виконано такі кроки.

1. Встановлено змінну з назвою моделі:

```
val modelName = "mobilenetv1.tflite"
```

2. Встановлено параметри моделі, такі як поріг передбачення та розмір набору результатів:

```
val optionsBuilder =  
    ObjectDetector.ObjectDetectorOptions.builder()  
        .setScoreThreshold(threshold)  
        .setMaxResults(maxResults)
```

3. Увімкнено прискорення графічного процесора за допомогою параметрів і дозвольте коду коректно завершуватися збоєм, якщо прискорення не підтримується на пристрої:

```
try {  
    optionsBuilder.useGpu()  
} catch(e: Exception) {  
    objectDetectorListener.onError("GPU is not supported on this device")  
}
```

4. Встановлено такі налаштування цього об'єкта для створення об'єкта TensorFlow Lite ObjectDetector, що містить модель:

```
objectDetector =  
    ObjectDetector.createFromFileAndOptions(  
        context, modelName, optionsBuilder.build())
```

Після цих кроків почато синхронізацію з Gradle. На телефоні в налаштуваннях було включено режим розробника та підключено телефон до ПК. Після цих дій телефон відображається в Android Studio. Після успішної синхронізації з Gradle та телефоном можна натискати на зелену кнопку Run. Після завершення зборки Android Studio відобразить повідомлення BUILD SUCCESSFUL на панелі стану Build Output .

3.7 Тестування моделі розпізнавання об'єктів

В встановленому додатку на смартфоні було встановлено 4 моделі розпізнавання об'єктів:

- MobileNetV1;

- EfficientDet-Lite0;
- EfficientDet-Lite1;
- EfficientDet-Lite2.

Для тесту був вибраний смартфон Google Pixel 2 з такими найголовнішими характеристиками (таблиця 3.4).

Таблиця 3.4 – Характеристики Pixel 2

Процесор	Qualcomm Snapdragon 835
Графічний процесор	Adreno 540
Оперативна пам'ять	4 ГБ

Результати тестувань продуктивності моделей нейронних мереж показані в таблиці 3.5.

Таблиця 3.5 – Результати тестувань

Назва моделі	mAP, %	Час кадру на GPU, ms	Час кадру на CPU, ms
EfficientDet 0	33.6	146–157	98–108
EfficientDet 1	54	230–280	167–180
EfficientDet 2	67	420–436	242–260
MobileNet V1	48	90–100	45–50

За цими результатами, моделей нейронних мереж які працювали на GPU в середньому в 1,66 разім мають більший час кадру тобто обробка на CPU в 1,66 разів швидша.

В процесі тестування моделей нейронних мереж були виявлені такі проблеми:

Проблема малих розмірів. На рисунку 3.5 вхідні розміни зображення. При польоті на середніх та великих висотах, для розпізнавання об’єктів (автомобілів)

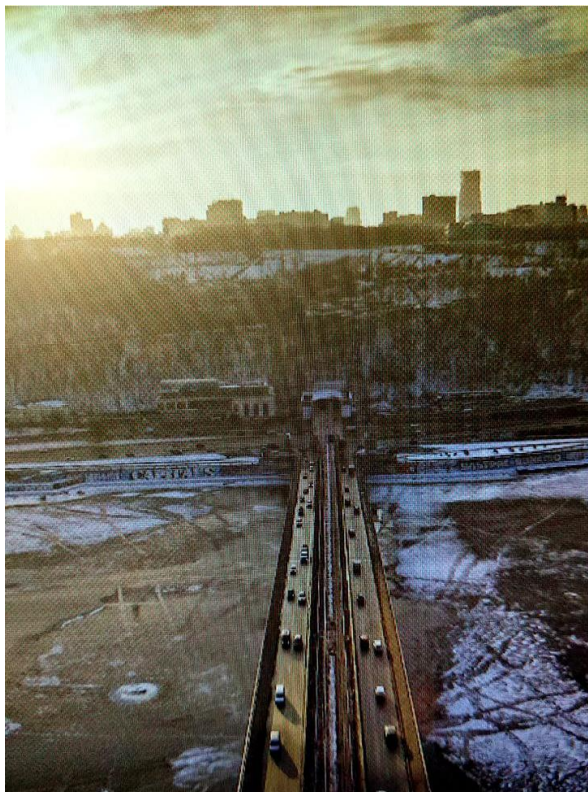


Рисунок 3.5 – Вхідні дані

потрібно дуже сильно збільшувати розмір вхідної картинки (рис. 3.6).



Рисунок 3.6 – Вихідний результат

Для вирішення даної проблеми можна використати один з способів збільшення зображення:

- оптичне, використання оптичного зуму. Даний спосіб потребує встановлення

камери з оптикою яка зможе достатньо збільшувати зображення. Такого типу камери встановлюються на ЛА великого розміру які виконують польоти на великих висотах;

– цифрове, використання цифрового зуму. Даний спосіб потребує використання камери великої розділової здатності та додаткових обчислювальних потужностей. Такого типу камери можна встановлювати на дрони невеликих розмірів та які виконують польоти на середніх висотах.

Проблема нагріву. При роботі програми телефон нагрівається, що приводить до зменшення частоти процесору та збільшення часу кадру в 2 рази. Дана проблема може стати критичною при більш сурових способах використання. Для вирішення даної проблеми в майбутніх сценаріях використання БПЛА потрібно забезпечити обчислювальний блок більш потужнішою системою охолодження чим встановлена на телефоні.

4 ОХОРОНА ПРАЦІ

Охорона праці при роботі з комп'ютером.

Сучасний розвиток технічного та технологічного стану виробництва передбачає постійну автоматизацію та оптимізацію виробничих процесів. Сьогодні, напевно, важко уявити компанію, господарська діяльність в якій здійснювалась би без використання комп'ютерної техніки. Через масовий характер робіт, що виконуються працівниками за допомогою комп'ютера, законодавством України чітко врегульовано норми та вимоги до використання комп'ютерної техніки на підприємстві, безпосередньо й охорона праці при роботі з комп'ютером.

Вимоги до приміщення.

Приміщення, в яких планується установка та подальша робота з комп'ютером, повинні відповідати проектній документації будинку, погодженій з уповноваженими державними органами. Крім того, роботодавець повинен враховувати санітарні нормативи освітлення, вимоги до параметрів мікроклімату (температура, відносна вологість), ступеня і сили вібрації, звукового шуму і вогнестійкості приміщення, а також характеристики електромагнітного, ультрафіолетового та інфрачервоного полів. Конкретні показники зазначених санітарних норм див. в Державних санітарних правилах і нормах роботи з візуальними дисплейними терміналами електронно-обчислювальних машин ДСанПН 3.3.2.007–98, затверджених Постановою Головного державного санітарного лікаря України №7 від 10 грудня 1998 року. Правила поширюються на умови й організацію праці при роботі з візуальними дисплейними терміналами (ВДТ) усіх типів вітчизняного та зарубіжного виробництва на основі електронно-променевих трубок (ЕПТ), що використовуються в електронно-обчислювальних машинах (ЕОМ) колективного використання та персональних ЕОМ (ПЕОМ). Так, наприклад, роботодавцю заборонено установлювати комп'ютери в приміщеннях, розташованих у підвалах будинків. Для уникнення можливих аварій та замикань, поряд з приміщеннями, де вестиметься робота з комп'ютером (над чи під ними), також не дозволяється проведення робіт, що потребують здійснення надмірно вологих технологічних процесів. Відповідне приміщення повинно бути укомплектоване системами центрального або індивідуального опалення, кондиціонування чи вентиляції повітря. Але при установці зазначених систем, необхідно переконатись, що батареї опалення, водопровідні труби, вентиляційні кабелі тощо, надійно сховані під захисними щитками, які перешкоджатимуть можливому потраплянню робітника під напругу.

У кожній кімнаті, де обладнуватимуться робочі місця співробітників, що

працюватимуть на комп'ютері, повинні бути наявні елементи природного та штучного освітлення. При цьому, на вікнах слід встановити легко регульовані жалюзі чи штори, які дозволять працівникам коригувати рівень освітлення в приміщенні. Бажано розмістити комп'ютери в кімнаті таким чином, щоб світло потрапляло на екрани моніторів з півдня чи північного сходу. З метою досягнення максимального рівня безпеки і охорони праці при роботі з комп'ютером, виробничі приміщення необхідно обладнати аптечками першої медичної допомоги, системами автоматичної пожежної сигналізації і вогнегасниками. В приміщенні, в якому разом працюють 5 або більше комп'ютерів, на видимому місці встановлюється службовий вимикач, який у разі потреби дозволить повністю відключити електричне живлення кімнати.

Вимоги до особистого робочого місця працівника

Роботодавець, який використовує найману працю робітників, повинен забезпечити відповідність їхніх робочих місць комфортним та безпечним умовам. Розмір одного робочого місця має становити не менше 6 квадратних метрів. При необхідності, суміжні робочі місця співробітників, що працюють з комп'ютером, слід розділити перегородками висотою до 2 метрів. При визначенні достатнього розміру приміщення і робочого місця на одну особу необхідно додатково враховувати шафи, сейфи, тумби або інші предмети меблів чи обладнання, які знаходяться в кімнаті. На столі працівника можливо розмістити допоміжні для роботи пристрої (принтери, колонки, сканери), а також місця для зберігання документів, за умови, що це не обмежуватиме видимість екрану і не заважатиме працівнику. У разі надмірного шуму чи вібрації технічного обладнання, роботодавець повинен забезпечити працівників антивібраційними килимками. Робочий стілець співробітника має бути підйомно-поворотним, легко регульованим за висотою та забезпечувати належну підтримку та зручне положення спини і хребта особи. Щодня необхідно проводити вологе прибирання приміщення, та очищати робоче місце та безпосередньо монітор комп'ютера від запиленості. На підприємстві забороняється: проводити ремонт та технічне обслуговування комп'ютера за робочим місцем працівника; самочинно ремонтувати або намагатись здійснити технічне налагодження комп'ютера без залучення компетентних спеціалістів; складувати на робочому місці зайві документи, деталі та предмети, що не потрібні для роботи; використовувати монітори з нечітким зображенням та монітори, у яких наявні поламки екрану; працювати з матричним принтером без антивібраційного покриття та зі знятою кришкою. Допускати до роботи осіб, які не пройшли затверджений на підприємстві курс охорони праці для роботи з комп'ютером, не дозволяється.

ВИСНОВКИ

За прогнозами, до 2026 року світовий ринок БПЛА досягне 34,5 млрд доларів, щорічне зростання складе 32%. У міру того, як БПЛА дешевшають і стають доступнішими для споживачів, кількість знімків з БПЛА також стрімко зростає. Рішення проблеми вбачається в штучному інтелекті та комп'ютерному зорі для БПЛА. Завдяки технологічним досягненням останніх років, моделі комп'ютерного зору можуть зрівнятися або навіть перевершити людські здібності у завданнях класифікації зображень та виявлення об'єктів – лише за частки секунди.

БПЛА використовуються в різних галузях, включаючи:

- природні ресурси для відстеження популяцій диких тварин, моніторингу зусиль з лісовідновлення, відстеження худоби;
- охорона, ведення розвідки та спостереження;
- екстрені служби для пошуку тих, хто вижив, або доставки медикаментів у віддалені райони.

Замість того, щоб відправляти зібрані дані в хмару, штучний інтелект виконує обробку безпосередньо на так званих вбудованих пристроях, розташованих близько до джерела збору.

Результатом виконаного завдання буде розробка системи штучного інтелекту БПЛА, за допомогою якої можна ідентифікувати будь-який об'єкт, який ми його навчимо ідентифікувати, будь то тварини, люди або звичайні предмети. Це дешевше, безпечніше та точніше, ніж традиційні методи аналізу зображень, отриманих з камер спостереження за допомогою людської праці. Ці переваги переважають витрати на розробку та впровадження.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. PricewaterhouseCoopers. Skies without Limits–Drones-Taking the UK's Economy to New Heights [Електронний ресурс] / PricewaterhouseCoopers // PwC. – 2018. – Режим доступу до ресурсу: <https://www.pwc.co.uk/intelligent-digital/drones/Drones-impact-on-the-UK-economy-FINAL.pdf>.
2. Красномовець П. Народжені літати. З 2014 року в Україні виросла ціла галузь безпілотників. Чи допоможе війна їй стати глобальним гравцем [Електронний ресурс] / Павло Красномовець. – 2022. – Режим доступу до ресурсу: <https://forbes.ua/innovations/narodzheni-litati-z-2014-roku-v-kraini-virosla-tsila-galuz-virobnitstva-bpla-chi-dopomozhe-viyna-ukrainskomu-military-tech-stati-globalnim-gravtsem-06092022-8150>.
3. Невлюдов, І.Ш. дипломне проєктування для студентів усіх форм навчання спеціальностей 151 «автоматизація та комп'ютерно-інтегровані технології» [Текст] / І.Ш. Невлюдов, А. О. Андрусевич, О. В. Токарева, Г. В. Пономарьова – Київ-58, пр. Космонавта Комарова, 1, 2016 – 320с.
4. ДСТУ 3008-15. документація. звіти у сфері науки і техніки. структура і правила оформлення [Текст] – Введ. 2015-06-22. – К. Держстандарт України, 2017 – 29 с.
5. Методичні вказівки до підготовки атестаційної роботи бакалавра для студентів усіх форм навчання спеціальності 151 Автоматизація та комп'ютерно-інтегровані технології / Упоряд.: І.Ш. Невлюдов, О.В. Токарева, Г.В. Понома-рьова – Харків: ХНУРЕ, – 2018. –34с.4.
6. Colomina I. Unmanned aerial systems for photogrammetry and remote sensing: A review [Електронний ресурс] / I. Colomina, P. Molina // ISPRS Journal of Photogrammetry and Remote Sensing. – 2014. – Режим доступу до ресурсу: <https://www.sciencedirect.com/science/article/pii/S0924271614000501>.
7. The Rise of UAVs [Електронний ресурс] / A.Cummings, A. McKee, K. Kulkarni, N. Markandey // Photogrammetric Engineering and Remote Sensing. – 2017. – Режим доступу до ресурсу: https://www.researchgate.net/publication/315946402_The_rise_of_UAVs.
8. PROSPECTS OF THE DEVELOPMENT OF UNMANNED AERIAL VEHICLES (UAVs) [Електронний ресурс] / [R. Shokirov, N. Abdujabarov, J. Takhirov та ін.] // Kadambay Saytov's Lab. – 2020. – Режим доступу до ресурсу:

https://www.researchgate.net/publication/345306836_PROSPECTS_OF_THE_DEVELOPMENT_OF_UNMANNED_AERIAL_VEHICLES_UAVs.

9. Dr. Palik M. Brief history of UAV development [Електронний ресурс] / M. Dr. Palik, N. Máté // Repüléstudományi Közlemények. – 2019. – Режим доступа до ресурсу: https://www.researchgate.net/publication/333584348_Brief_history_of_UAV_development.

10. Sárhidai G. Robotrepülőgépek / Gyula Sárhidai. – Budapest: Zrínyi Katonai Könyv- és Lapkiadó, 1986. – 63 с.

11. A gesture based kinect for quadrotor control [Електронний ресурс] / A.Mashood, H. Noura, I. Jawhar, N. Mohamed // ICTRC. – 2015. – Режим доступа до ресурсу: <https://ieeexplore.ieee.org/abstract/document/7156481/authors#authors>.

12. Krizhevsky A. ImageNet Classification with Deep Convolutional Neural Networks [Електронний ресурс] / A. Krizhevsky, I. Sutskever, G. E. Hinton // NIPS. – 2012. – Режим доступа до ресурсу: <https://papers.nips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>.

13. EVOLUTION OF THE UNMANNED AERIAL VEHICLES UNTIL PRESENT [Електронний ресурс] / [C. Cuerno-Rejado, L. García-Hernández, A. o Sánchez-Carmona та ін.] // DYNA. – 2016. – Режим доступа до ресурсу: https://orbilu.uni.lu/bitstream/10993/45787/1/7781_ENGLISH.pdf.

14. Baca T. Autonomous landing on a moving car with unmanned aerial vehicle [Електронний ресурс] / T. Baca, P. Stepan, M. Saska // ECMR. – 2017. – Режим доступа до ресурсу: <https://ieeexplore.ieee.org/abstract/document/8098700>.

15. Event-Based Moving Object Detection and Tracking [Електронний ресурс] / A.Mitrokhin, C. Fermüller, C. Parameshwara, Y. Aloimonos // IROS. – 2018. – Режим доступа до ресурсу: <https://ieeexplore.ieee.org/abstract/document/8593805>.

16. Event-Based Feature Extraction Using Adaptive Selection Thresholds [Електронний ресурс] / [S. Afshar, N. Ralph, Y. Xu та ін.] // Sensors 2020. – 2020. – Режим доступа до ресурсу: <https://www.mdpi.com/1424-8220/20/6/1600/htm>.

17. Gade R., Moeslund T. B. Thermal cameras and applications: a survey. URL: <https://doi.org/10.1007/s00138-013-0570-5> (date of access: 01.10.2022).

18. Portmann J., Lynen S., Chli M., Siegwart R., People detection and tracking from aerial thermal views. URL: <https://ieeexplore.ieee.org/document/6907094> (date of access: 01.10.2022).

19. Ramon Soria P. Detection, Location and Grasping Objects Using a Stereo Sensor on UAV in Outdoor Environments [Електронний ресурс] / P. Ramon Soria, B. C. Arrue, A.

Ollero // Sensors 2017. – 2017. – Режим доступу до ресурсу: <https://www.mdpi.com/1424-8220/17/1/103>.

20. UAV in the advent of the twenties: Where we stand and what is next [Електронний ресурс] / [F. Nex, C. Armenakis, M. Cramer та ін.] // ISPRS Journal of Photogrammetry and Remote Sensing. – 2022. – Режим доступу до ресурсу: https://www.researchgate.net/publication/357884334_UAV_in_the_advent_of_the_twenties_Where_we_stand_and_what_is_next.

21. Enabling High-Performance Onboard Computing with Virtualization for Unmanned Aerial Systems [Електронний ресурс] / [B. Wang, J. Xie, S. Li та ін.] // International Conference on Unmanned Aircraft Systems (ICUAS). – 2018. – Режим доступу до ресурсу: <https://ieeexplore.ieee.org/document/8453368>.

22. CodeSLAM - Learning a Compact, Optimisable Representation for Dense Visual SLAM [Електронний ресурс] / [M. Bloesch, J. Czarowski, R. Clark та ін.] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2018. – Режим доступу до ресурсу: <https://ieeexplore.ieee.org/document/8578369>.

23. DroNet: Learning to Fly by Driving [Електронний ресурс] / A. Loquercio, A. I. Maqueda, C. R. del-Blanco, D. Scaramuzza // IEEE. – 2018. – Режим доступу до ресурсу: <https://ieeexplore.ieee.org/document/8264734>.

24. Comparing Energy Efficiency of CPU, GPU and FPGA Implementations for Vision Kernels [Електронний ресурс] / [M. Qasaimeh, K. Denolf, J. Lo та ін.] // IEEE. – 2019. – Режим доступу до ресурсу: <https://ieeexplore.ieee.org/document/8782524>.

25. Real-Time Dense Stereo Embedded in a UAV for Road Inspection [Електронний ресурс] / [R. Fan, J. Jiao, J. Pan та ін.] // IEEE. – 2019. – Режим доступу до ресурсу: <https://ieeexplore.ieee.org/document/9025600>.

26. Balamuralidhar N. MultEYE: Monitoring System for Real-Time Vehicle Detection, Tracking and Speed Estimation from UAV Imagery on Edge-Computing Platforms [Електронний ресурс] / N. Balamuralidhar, S. Tilon, F. Nex // Remote Sens. – 2021. – Режим доступу до ресурсу: <https://doi.org/10.3390/rs13040573>.

27. UAV-YOLO: Small Object Detection on Unmanned Aerial Vehicle Perspective [Електронний ресурс] / [M. Liu, X. Wang, A. Zhou та ін.] // Sensors. – 2020. – Режим доступу до ресурсу: <https://doi.org/10.3390/s20082238>.

28. Deep Learning for Hyperspectral Image Classification: An Overview [Електронний ресурс] / [S. Li, W. Song, L. Fang та ін.] // IEEE. – 2019. – Режим доступу до ресурсу: <https://ieeexplore.ieee.org/document/8697135>.

29. A deep learning approach to DTM extraction from imagery using rule-based training labels [Електронний ресурс] / C. M. Gevaert, C. Persello, F. Nex, G. Vosselman // ISPRS Journal of Photogrammetry and Remote Sensing. – 2018. – Режим доступу до ресурсу: <https://doi.org/10.1016/j.isprsjprs.2018.06.001>.
30. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications [Електронний ресурс] / [A. G. Howard, M. Zhu, B. Chen та ін.] // arXiv. – 2017. – Режим доступу до ресурсу: <https://doi.org/10.48550/arXiv.1704.04861>.
31. Miličević Z. M. From the early days of 962 unmanned aerial vehicles (UAVs) to their integration into wireless networks [Електронний ресурс] / Z. M. Miličević, Z. S. Bojković // Military Technical Courier. – 2021. – Режим доступу до ресурсу: https://www.researchgate.net/publication/366012651_Path_Planning_with_Time_Windows_for_Multiple_UAVs_Based_on_Gray_Wolf_Algorithm.