

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерної інженерії та управління
(повна назва)

Кафедра Електронних обчислюваних машин
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

Рівень вищої освіти Другий (магістерський)

Модель інтелектуальної обробки зображень

(тема)

Виконав:

студент II курсу, групи СПМ-22-2
Яковлєва Н.Ю.
(прізвище, ініціали)

Спеціальність 123 Комп'ютерна інженерія та управління
(код і повна назва спеціальності)

Тип програми Освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма Системне програмування
(повна назва освітньої програми)

Керівник: ст.викл. Єршоміна Н.С.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. Кафедри ЕОМ

(підпис)

Коваленко А. А.

(прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет Комп'ютерної інженерії та управління

Кафедра Електронних обчислюваних машин

Рівень вищої освіти Другий (магістерський)

Спеціальність Комп'ютерна інженерія
(код і повна назва)

Тип програми Освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма Системне програмування
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

“ _____ ” _____ 20__ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

студенту Яковлевої Наталії Юріївни
(прізвище, ім'я, по батькові)

1. Тема роботи Модель інтелектуальної обробки зображень

затверджена наказом по університету від “ 6 ” листопада 2023р. № 1299 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 15 січня 2024 р.

3. Вхідні дані до роботи Andriod портативний девайс, камера, аналогове зображення

4. Перелік питань, що потрібно опрацювати у роботі _____

1) Аналіз проблеми та огляд існуючих рішень _____

2) Вибір технології розробки та інструментальних засобів _____

3) Програмно-апаратний комплекс _____

4) Програмна реалізація та тестування проєкту _____

5) Висновок _____

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 64 с., 23 рис., 1 дод., 13 джерел.

ДОПОВНЕНА РЕАЛЬНІСТЬ, ВІРТУАЛЬНА РЕАЛЬНІСТЬ, ІНТЕЛКТУАЛЬНА ОБРОБКА ЗОБРАЖЕНЬ, ОБРОБКА ЗОБРАЖЕНЬ, ЗОБРАЖЕННЯ, ЗОБРАЖЕННЯ У РЕАЛЬНОМУ ЧАСІ, UNITY3D, КРОСС-ПЛАТФОРМНИЙ ДОДАТОК, МОБІЛЬНИЙ ДОДАТОК, VUFORIA.

Метою кваліфікаційної роботи є розробка моделі інтелектуальної обробки зображень. У ході роботи було досліджено існуючі методи розв'язання поставленої задачі, способи оцінки їх якості, а також існуючі проблеми та обмеження. Реалізація розробленої моделі проводилась на базі графічного рушія Unity 3d, з можливістю кросс-платформного використання.

Розроблена модель інтелектуальної обробки зображень представляє собою віртуальний музей з використанням технологій доповненої реальності на базі мобільного девайсу з операційною системою Android.

ABSTRACT

Master`s thesis 64 pages, 23 figures, 1 appendices, 13 sources.

AUGMENTED REALITY, VIRTUAL REALITY, IMAGE PROCESSING, IMAGE PROCESSING, IMAGES, REAL-TIME IMAGES, UNITY3D, CROSS-PLATFORM APPLICATION, MOBILE APPLICATION, VUFORIA.

The major goal of this thesis is to develop a model for intelligent image processing. During the work, existing methods for solving the stated problem, ways of assessing their quality, as well as existing problems and limitations were investigated. The implementation of the developed model was carried out using the Unity 3D graphics engine, with the possibility of cross-platform usage.

The developed model of intelligent image processing represents a virtual museum using augmented reality technologies on a mobile device with the Android operating system.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ	8
ВСТУП	9
1 ДОСЛІДЖЕННЯ МЕТОДІВ ІНТЕЛЕКТУАЛЬНОЇ ОБРОБКИ ЗОБРАЖЕНЬ ТА ГАЛУЗЕЙ ЇХ ЗАСТОСУВАННЯ.....	10
1.1 Загальні відомості про інтелектуальну обробку зображень	10
1.1.1 Зображення та їх отримання для інтелектуальної обробки.....	12
1.2 Види та способи інтелектуальної обробки зображень	13
1.2.1 Сегментація зображень.....	15
1.2.2 Розпізнавання об'єктів	19
1.2.3 Відстеження руху	20
1.2.4 Розпізнавання обличчя	21
1.2.5 Фільтрація та покращення зображень.....	24
1.3 Галузі застосування інтелектуальної обробки зображень	28
1.3.1 Медицина	28
1.3.2 Автомобільна промисловість.....	30
1.3.3 Безпека.....	30
1.3.4 Маркетинг	31
1.3.5 Мистецтво	32
1.4 Мета та постановка завдання	33
2 ПРОГРАМНО-АПАРАТНИЙ КОМПЛЕКС	34
2.1 Вибір платформи для реалізації проекту	34
2.2 Вибір технологій інтелектуальної обробки зображень та інструментів.....	36
2.2.1 AR технологія як метод інтелектуальної обробки зображень...	36
2.2.2 Unity 3d як середовище розробки програмних додатків.....	42

2.3 Вибір мови програмування	44
3 ПРОГРАМНА РЕАЛІЗАЦІЯ ПРОЕКТУ	46
3.1 Етапи реалізації проєкту.....	46
3.2 Налаштування середовища Unity 3d	47
3.3 Підключення плагінів	48
3.4 Створення та налаштування макета-додатка обробки зображень у реальному часі	50
ВИСНОВКИ.....	55
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	56
ДОДАТОК А.....	58

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ
І ТЕРМІНІВ

AR – Augmented reality (Доповнена реальність)

VR – Virtual reality (Віртуальна реальність)

ІОЗ – Інтелектуальна обробка зображень

ВСТУП

Інтелектуальна обробка зображень є швидко розвиваючою галуззю обчислювальної науки, яка знаходить застосування в різних аспектах буденного життя. Завдяки поєднанню технологій обробки зображень та іншим розширеним методам, вона перетворює звичайні зображення у джерело цінної інформації та забезпечує здатність розпізнавати об'єкти, аналізувати образи та приймати рішення на основі зображень.

Спричиненими стрімкими зростаннями в обчислювальних ресурсах та розвитком алгоритмів і штучних нейронних мереж, інтелектуальна обробка зображень стала важливим інструментом у різних галузях, включаючи медицину, автомобільну промисловість, робототехніку, безпеку, мультимедіа та інше.

У цьому дослідженні було розглянуто основні методи та технології інтелектуальної обробки зображень, їх застосування та перспективи розвитку. Було проаналізовано, як інтелектуальна обробка зображень допомагає вирішувати складні завдання та вносить вагомий внесок у вдосконалення сучасних технологій.

Ця тема надає можливість зрозуміти, як обробка зображень впливає на наше оточення і способи, якими ми взаємодіємо з ним, відкриває нові можливості та розширює горизонти для подальших досліджень та застосувань.

1 ДОСЛІДЖЕННЯ МЕТОДІВ ІНТЕЛЕКТУАЛЬНОЇ ОБРОБКИ ЗОБРАЖЕНЬ ТА ГАЛУЗЕЙ ЇХ ЗАСТОСУВАННЯ

1.1 Загальні відомості про інтелектуальну обробку зображень

Сьогодні використання комп'ютерів як ефективного інструменту полегшило контроль інженерних систем, збір даних, аналіз даних, обробку даних і навіть прийняття рішень на різних етапах процесу. Обробка зображень, нині обробка цифрових зображень, (Intelligent Image Processing) – це розділ інформатики, який займається обробкою цифрових сигналів, які представляють зображення, зроблені цифровою камерою або скановані сканером [1]. Обробка зображень тепер називається обробкою цифрових зображень, яка потребує знань комп'ютера та обробляє цифровий сигнал, отриманий камерою чи сканером. Обробка зображень є одним із основних компонентів інтелектуальних систем, де приймаються рішення. Обробка цифрових зображень застосовується комп'ютерними системами. Різноманітні застосування обробки зображень у різних галузях техніки, промисловості, міст, медицини та науки зробили її дуже активною темою серед галузей досліджень.

У конкретному розумінні обробка зображень – це будь-який тип обробки сигналу, який є вхідним сигналом для зображення, такого як фотографія або сцена з фільму. Вихідні дані процесора зображень можуть бути відображенням або набором спеціальних символів або пов'язаних із зображенням змінних.

Більшість методів обробки зображень передбачає роботу із зображенням як із двовимірним сигналом. До зображень застосовуються стандартні методи обробки сигналів. Обробка зображень часто стосується цифрової обробки зображень, але існує також оптична та аналогова обробка зображень.

Переваги цифрової обробки зображень:

- покращена якість зображення. Алгоритми цифрової обробки зображень можуть покращити візуальну якість зображень, роблячи їх чіткішими та більш

інформативними;

- автоматизовані завдання на основі зображень: обробка цифрових зображень може автоматизувати багато завдань на основі зображень, таких як розпізнавання об'єктів, виявлення шаблонів і вимірювання;

- підвищена ефективність. Алгоритми обробки цифрових зображень можуть обробляти зображення набагато швидше, ніж люди, що дає змогу аналізувати великі обсяги даних за короткий проміжок часу;

- підвищена точність. Алгоритми обробки цифрових зображень можуть надавати більш точні результати, ніж людські, особливо для завдань, які вимагають точних вимірювань або кількісного аналізу.

Недоліки цифрової обробки зображень:

- високі обчислювальні витрати: деякі алгоритми обробки цифрових зображень потребують великих обчислень і значні обчислювальні ресурси;

- обмежена інтерпретація: деякі алгоритми обробки цифрових зображень можуть давати результати, які важко інтерпретувати людині, особливо для складних алгоритмів;

- залежність від якості вхідних даних: Якість вихідних даних алгоритмів обробки цифрових зображень сильно залежить від якості вхідних зображень. Вхідні зображення низької якості можуть призвести до низької якості результату;

- обмеження алгоритмів. Алгоритми обробки цифрових зображень мають обмеження, такі як труднощі розпізнавання об'єктів у засвітлених або погано освітлених сценах або нездатність розпізнавати об'єкти зі значними деформаціями чи оклюзіями;

- залежність від якісних навчальних даних: продуктивність багатьох алгоритмів обробки цифрових зображень залежить від якості навчальних даних, які використовуються для розробки алгоритмів.

1.1.1 Зображення та їх отримання для інтелектуальної обробки

Зображення визначається як двовимірна функція $F(x,y)$, де x і y – просторові координати, а амплітуда F у будь-якій парі координат (x,y) називається інтенсивністю цього зображення у цій точці. Коли значення x , y та амплітуди F скінченні, ми називаємо це цифровим зображенням.

Іншими словами, зображення може бути визначено двовимірним масивом, спеціально організованим у рядки та стовпці.

Цифрове зображення складається зі скінченної кількості елементів, кожен з яких має певне значення в певному місці. Ці елементи називаються елементами зображення та пікселями. Піксель найбільш широко використовується для позначення елементів цифрове зображення.

Функція $f(x, y)$ характеризується двома компонентами: кількістю джерела освітлення, що падає на сцену, що переглядається і кількістю освітлення, відбитого об'єктами в сцені. Відповідно, вони називаються компонентами освітлення та відбиття та позначаються $i(x, y)$ та $r(x, y)$ відповідно. Дві функції поєднуються як добуток, утворюючи $f(x, y)$ (1.1):

$$F(x, y) = i(x, y)r(x, y) \quad (1.1)$$

де $0 \leq r(x, y) \leq 1$ та $0 \leq i(x, y) < \infty$.

Таким чином, коефіцієнт відбиття обмежений 0 (повне поглинання) і 1 (повний коефіцієнт відбиття). Характер $i(x, y)$ (визначається джерелом освітлення, а $r(x, y)$ – характеристиками об'єктів зображення. Ці вирази застосовні також до зображень, сформованих через передачу світла через середовище, таке як рентген грудної клітки. У цьому випадку ми матимемо справу з пропускнуою здатністю замість функції відбиття, але обмеження будуть такими ж, як у рівнянні $r(x,y)$, а сформована функція зображення буде змодельована як добуток у рівнянні $F(x, y)$.

Цифрові зображення бувають декількох видів:

- бінарне зображення – як випливає з назви, бінарне зображення містить лише два піксельні елементи, тобто 0 і 1, де 0 означає чорний колір, а 1 – білий. Це зображення також відоме як монохромне;

- чорно-біле зображення – зображення, яке складається лише з чорного та білого кольорів, називається чорно-білим зображенням;

- 8-бітний формат кольорів – це найвідоміший формат зображення. Він містить 256 різних відтінків кольорів і широко відомий як зображення в градаціях сірого. У цьому форматі 0 означає чорний, 255 – білий, а 127 – сірий;

- 16-бітний формат кольору – це формат кольорового зображення. Він містить 65 536 різних кольорів. Він також відомий як формат високого кольору. У цьому форматі розподіл кольорів не такий самий, як зображення у відтінках сірого.



Рисунок 1.1 – Приклад цифрового зображення на якому можна побачити інструмент переводу реальності у цифрове зображення

1.2 Види та способи інтелектуальної обробки зображень

Геометричні коригування, такі як зміна розміру, обертання та коригування

кольору, є основними операціями обробки зображень. Існують також інші методи обробки, такі, як змішування зображень із яскравістю, чіткістю чи кольоровим простором і об'єднання двох або більше зображень, які стискають зображення, наприклад, зменшення розміру областей зображення, що підвищує якість файлу, зменшення шуму та підвищення контрастності. Застосування Secure Socket Layer у мережевій та веб-безпеці досліджено Dastres і Soori [2] для посилення заходів безпеки в мережі даних. Дастрес і Сури [2] вивчають вплив «розривної діри» на різні процесори та операційні системи, щоб підвищити безпеку виробництва ЦП шляхом запобігання захопленню даних на комп'ютерах або смартфонах злоумисниками.

Обробка зображень має дві основні гілки: покращення зображень і машинний зір. Удосконалення зображень включає такі методи, як використання фільтра розмиття та підвищення контрастності для покращення візуальної якості зображень і забезпечення їх правильного відображення в цільовому середовищі (наприклад, комп'ютерний монітор), тоді як машинне зір має справу з методами. З їх допомогою можна зрозуміти значення та зміст зображень для використання в таких галузях, як робототехніка та вісь зображення [2]. Етапи обробки зображень показано на рисунку 1.1.

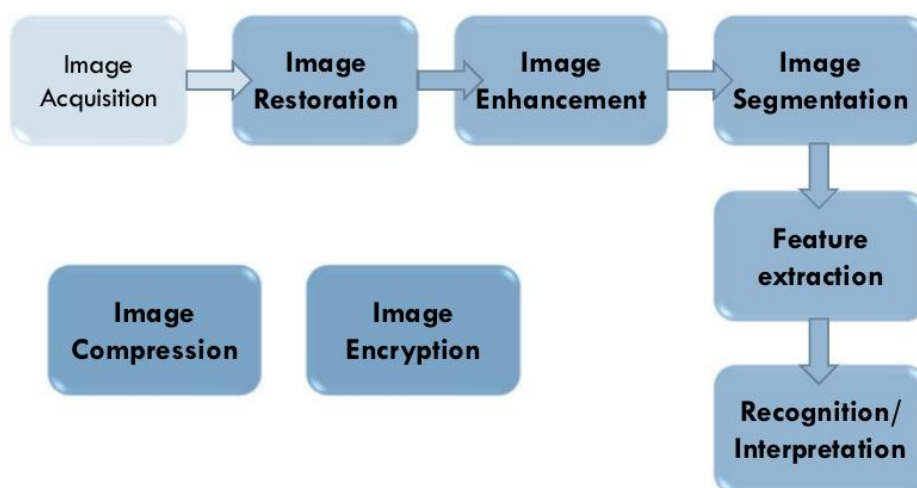


Рисунок 1.2 – Етапи обробки зображень

Основні етапи обробки цифрових зображень складаються з:

- 1) отримання зображення – це процес захоплення зображення за допомогою цифрової камери чи сканера або імпорт існуючого зображення на комп'ютер;
- 2) покращення зображення передбачає собою покращення візуальної якості зображення, наприклад підвищення контрастності, зменшення шуму та видалення артефактів;
- 3) відновлення зображення представляє собою усунення погіршення якості зображення, наприклад розмиття, шум та деформація;
- 4) сегментація зображення: передбачає поділ зображення на області або сегменти, кожен з яких відповідає певному об'єкту чи функції зображення;
- 5) репрезентація та опис зображення – це етап представлення зображення таким чином, щоб його можна було проаналізувати та обробити комп'ютером, а також опис характеристик зображення в стислому та змістовному вигляді;
- 6) аналіз зображення – використання алгоритмів і математичних моделей для отримання інформації із зображення, наприклад розпізнавання об'єктів, виявлення шаблонів і кількісного визначення характеристик;
- 7) синтез і стиснення зображень: цей етап передбачає створення нових зображень або стиснення існуючих зображень для зменшення вимог до зберігання та передачі.

1.2.1 Сегментація зображень

Сегментація зображення – це метод поділу цифрового зображення на підгрупи, які називаються сегментами зображення, що зменшує складність зображення та забезпечує подальшу обробку чи аналіз кожного сегмента зображення. Технічно сегментація – це призначення міток пікселям для ідентифікації об'єктів, людей або інших важливих елементів зображення.

Є два поширених підходи до сегментації зображення:

- підхід подібності: цей підхід передбачає виявлення подібності між пікселями зображення для формування сегмента на основі заданого порогового значення. Алгоритми машинного навчання, такі як кластеризація, базуються на такому підході до сегментування зображення;

- розривний підхід: цей підхід ґрунтується на непостійності значень інтенсивності пікселів зображення. Методи визначення ліній, точок і країв використовують цей тип підходу для отримання проміжних результатів сегментації, які пізніше можна обробити для отримання остаточного сегментованого зображення.

Існує п'ять поширених методів сегментації зображення:

- сегментація на основі порогів. Порогова сегментація зображення є простою формою сегментації зображення. Це спосіб створення бінарного або багатоколірного зображення на основі встановлення порогового значення інтенсивності пікселів вихідного зображення. У цьому процесі порогового значення розглядається гістограма інтенсивності всіх пікселів зображення. Потім встановлюється поріг для розділення зображення на частини. Наприклад, якщо розглядати пікселі зображення в діапазоні від 0 до 255, то можна встановити поріг 60. Таким чином, усі пікселі зі значеннями, меншими або рівними 60, матимуть значення 0 (чорний), а всі пікселі зі значенням більше 60 буде надано значення 255 (білий). Якщо розглядати зображення з фоном і об'єктом, можливо розділити зображення на області на основі інтенсивності об'єкта та фону. Але цей поріг має бути ідеально встановлений, щоб сегментувати зображення на об'єкт і фон;

- сегментація зображення на основі країв. Сегментація на основі країв базується на краях, знайдених на зображенні за допомогою різних операторів виявлення країв. Ці краї позначають місця розривів зображення в рівнях сірого, кольорі, текстурі тощо. Коли відбувається перехід від однієї області до іншої, рівень сірого може змінюватися. Отже, якщо з'являється можливість знайти цей розрив, також стає можливим знайти це ребро. Доступні різні оператори

визначення країв, але отримане зображення є проміжним результатом сегментації, і його не слід плутати з остаточним сегментованим зображенням. Наступним кроком виконується подальша обробка зображення, щоб його сегментувати. Додаткові кроки включають об'єднання отриманих сегментів країв в один сегмент, щоб зменшити кількість сегментів, а не фрагментів малих меж, які можуть перешкоджати процесу заповнення області. Це робиться для отримання безшовної межі об'єкта. Метою сегментації країв є отримання проміжного результату сегментації, до якого можливо застосувати сегментацію на основі регіону або будь-який інший тип сегментації, щоб отримати остаточне сегментоване зображення;

- сегментація зображення на основі регіонів: регіони можна класифікувати як пов'язані групи пікселів, що мають схожі характеристики. Подібність пікселів може стосуватися інтенсивності або кольору. У цьому типі сегментації існує набір заздалегідь визначених правил, яких слід дотримуватися, щоб класифікувати пікселі в схожі регіони. Методи сегментації на основі областей є кращими за методи сегментації на основі країв для зашумлених зображень;

- сегментація зображень на основі кластеризації: кластеризація - це тип неконтрольованого алгоритму машинного навчання. Його часто використовують для сегментації зображень. Одним з найпоширеніших алгоритмів кластеризації, що використовується для сегментації, є кластеризація за методом К-середніх. Цей тип кластеризації можна використовувати для створення сегментів кольорових зображень. Щоб це було зрозуміліше, уявімо двовимірний набір даних. Спочатку випадковим чином ініціалізуються центри в наборі даних (вибрані користувачем). Потім обчислюються відстані між усіма точками і всіма кластерами, і точки відносяться до кластера з найменшою відстанню. Потім центри всіх кластерів перераховуються, і середнє значення кластерів приймається за центр. Точки даних приписуються до цих кластерів. Цей процес продовжується до тих пір, поки алгоритм не знайде оптимальний розв'язок. Зазвичай кількість ітерацій, необхідних алгоритму для досягнення розв'язку,

дуже мала;

- сегментація на основі штучної нейронної мережі.

Сегментація зображення зазвичай використовується для виявлення об'єктів. Замість обробки всього зображення поширеною практикою є спочатку використання алгоритму сегментації зображення для пошуку об'єктів, що цікавлять зображення. Тоді детектор об'єктів може працювати з обмежувальною рамкою, уже визначеною алгоритмом сегментації. Це перешкоджає детектору обробляти все зображення, покращуючи точність і скорочуючи час висновку.

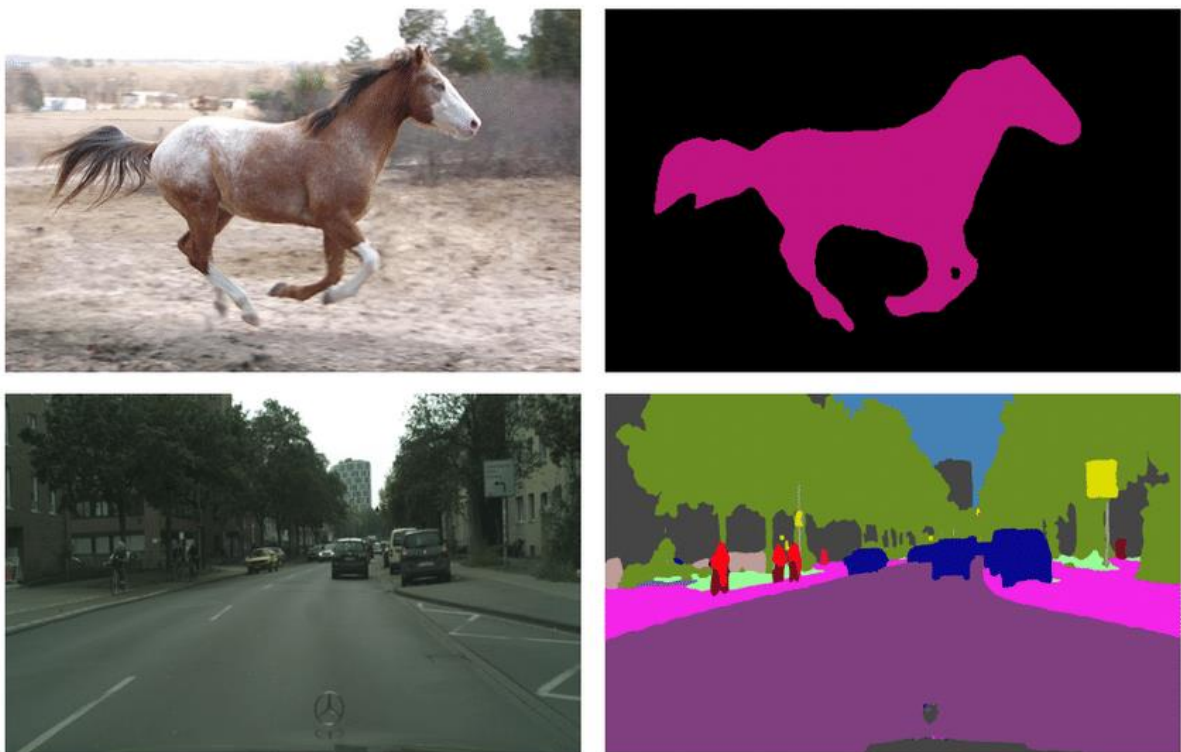


Рисунок 1.3 – Приклади сегментації зображень

Сегментація зображення є ключовим будівельним блоком технологій і алгоритмів комп'ютерного зору. Він використовується для багатьох практичних застосувань, включаючи аналіз медичних зображень, комп'ютерне бачення для автономних транспортних засобів, розпізнавання та виявлення облич, відеоспостереження та аналіз супутникових зображень.

Також цей метод може бути використаний для виділення обличь людей на

фотографіях, визначення об'єктів на зображеннях з багатьма об'єктами.

1.2.2 Розпізнавання об'єктів

Розпізнавання об'єктів – це технологія візуального штучного інтелекту для ідентифікації об'єктів на фото чи відеозображенні. Ця технологія є результатом роботи над глибоким навчанням і алгоритмами машинного навчання. Коли люди, які використовують цю технологію, переглядають фото або відео, вони можуть швидко ідентифікувати людей, об'єкти та багато іншого на зображенні. Метою роботи над технологією розпізнавання об'єктів часто є надання можливості штучному інтелекту вивчати речі людської природи та об'єктів, якими люди користуються. Розпізнавання об'єктів полягає в розпізнаванні, ідентифікації та позиціонуванні об'єктів на зображеннях на фото чи відео з певним ступенем впевненості.

Під час розпізнавання об'єктів є чотири основні завдання:

- класифікація: можна вирішити класифікацію об'єктів за допомогою згорткових нейронних мереж (CNN). Існує багато алгоритмів класифікації зображень на основі згорткових нейронних мереж. Більшість цих алгоритмів мають основу, яка використовує інфраструктуру CNN. Прикладами таких є Resnet, LeNet-5, AlexNet тощо. Класифікація об'єктів використовує тип моделі класифікатора ознак під час навчання моделі. Під час класифікації об'єктів алгоритм запам'ятовує, що є лише один об'єкт, ігноруючи всі інші класи;
- теггінг: теги об'єктів можуть розпізнавати кілька об'єктів для даного зображення. На відміну від класифікації об'єктів, цей процес намагається повернути всі найкращі класи, що відповідають зображенню під час додавання тегів;
- виявлення: виявлення об'єктів – це технологія обробки зображень, яка допомагає виявляти екземпляри об'єктів певного класу в цифрових зображеннях і відео;
- сегментація. Сегментація об'єкта, як було розглянуто у минулому

пункті, це процес поділу об'єкта на менші об'єкти фіксованого розміру для покращення зберігання та використання ресурсів для великих об'єктів;

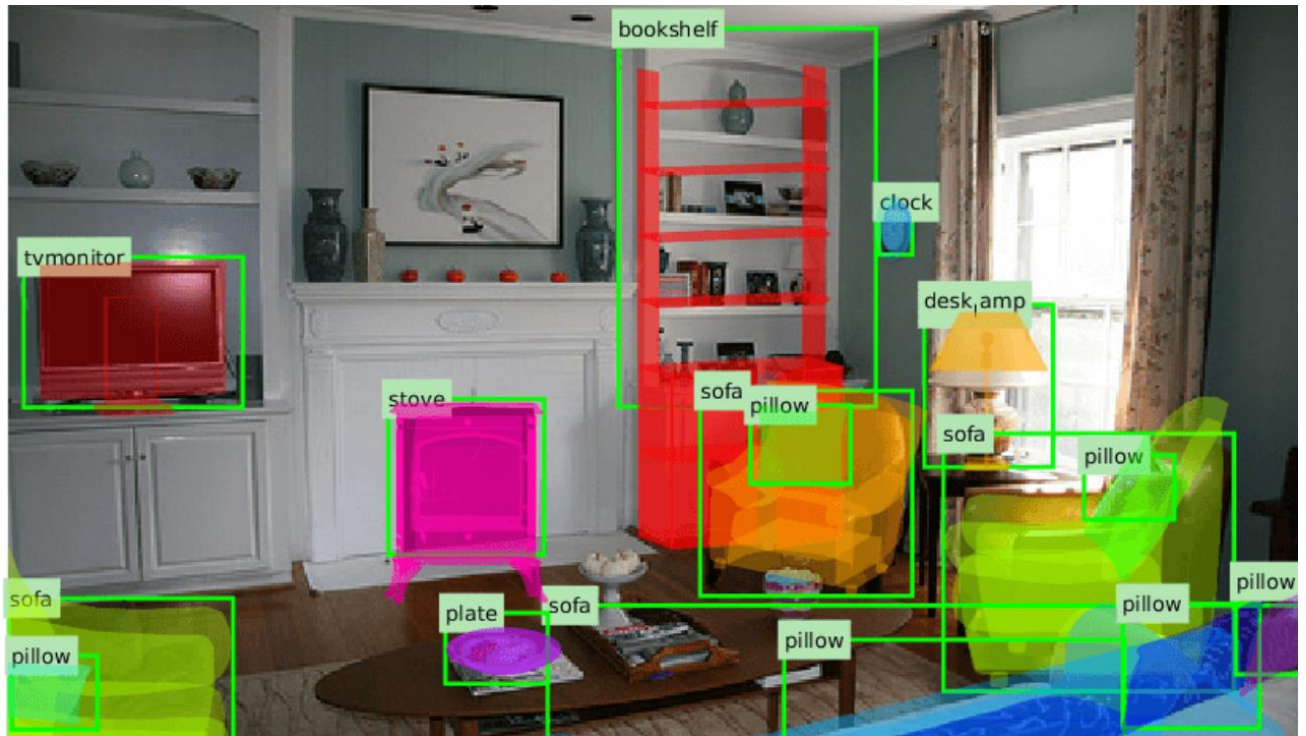


Рисунок 1.4 – Розпізнавання об'єктів кімнати

Технологія розпізнавання об'єктів використовується в багатьох сферах, таких як безпека, людські ресурси компаній, зв'язки з громадськістю та реклама, банківські послуги, охорона здоров'я та роботизовані системи зору. Однією зі сфер застосування цієї технології є технологія автономних транспортних засобів. Технологія розпізнавання об'єктів, яка використовується в автономних транспортних засобах, дозволяє ідентифікувати дорожні знаки, відмінності між пішоходом і нерухомим об'єктом і, коротко кажучи, ідентифікувати об'єкти на дорозі.

1.2.3 Відстеження руху

Відстеження руху полягає в визначенні руху об'єктів на послідовних кадрах відео. Це може бути використано для відстеження об'єктів у системах

відеоспостереження або для створення ефектів відеоредакторів.

Відстеження руху використовується в системах безпеки для виявлення підозрілих дій, а також в ігровій індустрії для створення інтерактивних ігор.

1.2.4 Розпізнавання обличчя

Цей метод полягає в ідентифікації та аналізі обличчя на зображеннях або відео. Він використовує нейронні мережі та алгоритми для визначення основних ознак обличчя людей.

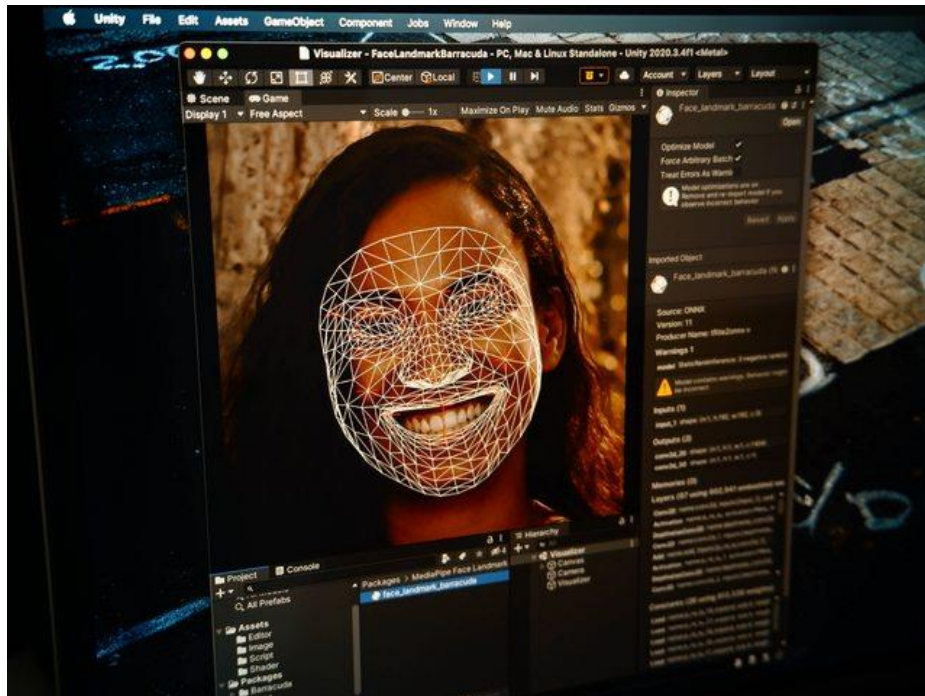


Рисунок 1.5 – Розпізнавання обличчя людини

Перші спроби ідентифікації особи за допомогою порівняння частини обличчя на фотографії було відзвітовано на британському суді в 1871 році [6]. Визнання обличчя є однією з найважливіших технік правоохоронних органів у випадках, коли є відеоматеріал або фотографії на місці злочину. Юристи проводять ручний тест на визнання обличчя для зіставлення з обличчям підозрюваного. Автоматизовані технології визнання обличчя підвищили

ефективність роботи працівників суду та спростили процес порівняння [7].

Сьогодні визнання обличчя, пов'язане з техніками штучного інтелекту, дозволяє ідентифікувати особу за її обличчям або перевірити, що вона твердить бути. Технологія визнання обличчя може аналізувати риси обличчя та інші біометричні дані, такі як очі, і порівнювати їх з фотографіями або відео. Звинувачення у широкому розповсюдженні спостереження створює багато обурення серед його супротивників, які бояться порушень конфіденційності даних і прав людини. Для своїх захисників визнання обличчя надає точну, швидку і безпечну аутентифікацію для захисту від різних форм обману. Згідно з звітом аналітичної компанії Mordor-intelligence [8], світовий ринок визнання обличчя був оцінений на рівні 4,4 мільярда доларів у 2019 році і очікується, що він перевищить 10,9 мільярда у 2025 році. Ця технологія вже стала популярною в деяких країнах, таких як Китай.

Завдяки технологіям штучного інтелекту в сфері розпізнання обличчя відбулися значні досягнення. У ранні часи інтереси досліджень були спрямовані, переважно, на визнання обличчя в контрольованих умовах, де прості класичні підходи надавали відмінну ефективність. Сьогодні дослідження фокусуються на неконтрольованих умовах, де технологія глибокого навчання [9] стає все більш популярною, оскільки вона має високу стійкість до численних змін, які можуть впливати на процес визнання.

Розпізнавання обличчя широко використовується в системах розблокування смартфонів за допомогою обличчя, в системах безпеки та в соціальних мережах для позначення друзів на фотографіях.

Розпізнавання обличчя є одним із кількох біометричних методів, який має високу точність і, на відміну від інших методів перевірки, користувач може легко увійти в процес перевірки зі своїм власним обличчям. Розпізнавання обличчя за допомогою комп'ютера є однією з найпривабливіших областей біометричних досліджень, яка включає різні наукові галузі, такі як машинне бачення, обчислювальний інтелект, розпізнавання образів і психологія. Тому використання програмного забезпечення, оснащеного системою розпізнавання

обличчя людини, може значно підвищити надійність систем контролю дорожнього руху, особливо на кордонах і в пунктах в'їзду в країну. Крім того, розпізнавання облич широко використовується в робототехніці для порятунку людського життя в небезпечних умовах.

Кожна система розпізнавання облич складається з чотирьох основних частин як структура та процес системи розпізнавання облич: зображення та виявлення, попередня обробка, вилучення ознак, реалізація та прийняття рішень.

Удосконалена система реорганізації обличчя, заснована на трьох рівнях згорткової нейронної мережі (CNN), розроблена Altameem для підвищення потужності системи виявлення обличчя в програмах обробки зображень у системах охорони здоров'я.

У процесі розпізнавання обличчя застосовуються три рівні, такі як класифікація текстури, кореляція та виявлення візуалізації обличчя за допомогою збереженої інформації. Вхідними даними є зображення обличчя, тобто ключові характеристики витягуються та аналізуються вчасно, уникаючи помилкового виявлення та етапів обчислення. Нейронна мережа, яка класифікує характеристики та розпізнає їх, це CNN. Крім того, алгоритм розроблено в програмах виявлення dase на основі вилучення властивостей з кількома роздільними властивостями. Застосовуючи цей алгоритм, швидкість обробки буде значно збільшена, фактично, за допомогою цього алгоритму він виділить більш ефективні функції та зменшить навантаження на програму.

Фаза впровадження та прийняття рішення є останнім кроком алгоритму виявлення обличчя. Система порівнює всі характеристики бази даних із зображеннями бази даних і на основі цього порівняння визначається потрібний клас зображення (людина). Типи класифікаторів, які використовуються для розпізнавання облич, включають класифікатори для опорної векторної машини найближчого сусіда та штучні нейронні мережі (Рисунок 1.12).

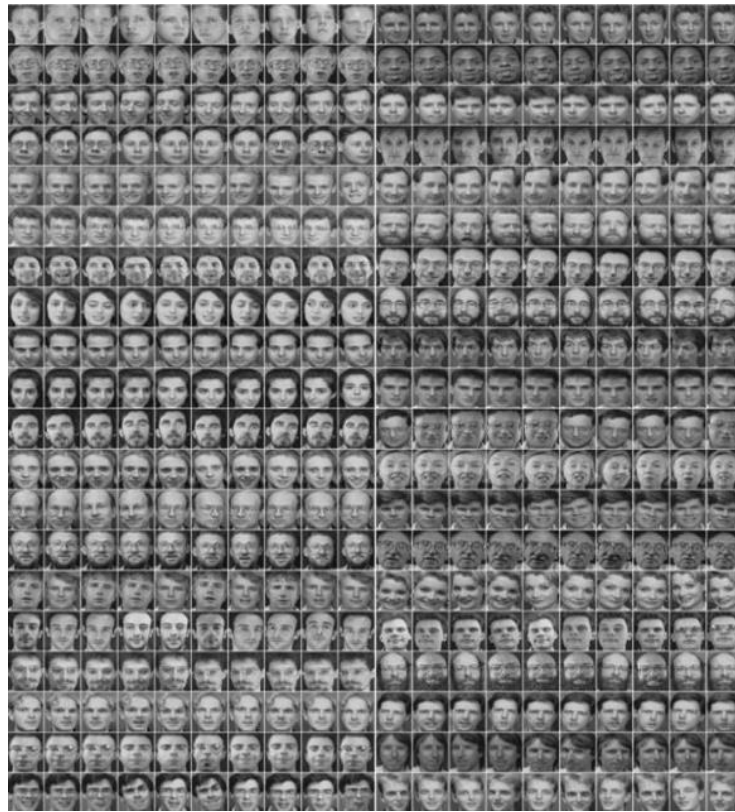


Рисунок 1.6 – База даних обличчь

1.2.5 Фільтрація та покращення зображень

Ці методи інтелектуальної обробки зображень використовуються в різних галузях, включаючи комп'ютерний зір, медицину, автоматизовану обробку зображень, розважальну індустрію та багато інших. Вони сприяють покращенню роботи систем, які взаємодіють з візуальною інформацією та роблять її більш доступною та корисною для людей.

Вони включають в себе застосування різних фільтрів та алгоритмів для покращення якості зображень, видалення шуму, збільшення різкості та змінення кольорової гами.

Фільтрацію зображень можна розділити на дві основні категорії. Просторова фільтрація – це традиційний метод фільтрації зображень. Вона використовується безпосередньо на пікселях зображення. Фільтри в області частот використовуються для видалення високих та низьких частот і для

згладжування.

Нелінійні фільтри використовуються для виявлення країв. Ці техніки фільтрації ефективніші, ніж лінійні фільтри. При лінійній фільтрації деталі та краї зображення мають тенденцію розмитися. Гауссів фільтр, Лапласів фільтр та фільтр середнього значення в околі можуть бути ідентифіковані як приклади лінійних фільтрів.

Ці методи використовуються для видалення шуму, збільшення різкості, підвищення контрасту, зміни кольорової гами та інших покращень. Нижче розглянемо деякі основні аспекти фільтрації та покращення зображень:

Згладжування (smoothing): згладжування зображень використовується для видалення шуму та зменшення деталей. Основна ідея полягає в обчисленні середнього значення пікселів в околі кожного пікселя та заміні його на це середнє значення. Згладжування використовується в медичній обробці зображень для зменшення шуму на рентгенівських знімках.



Рисунок 1.7 – Згладжування зображення

Підвищення різкості (sharpening)[6]: підвищення різкості включає в себе виділення граней та покращення контрасту між сусідніми пікселями. Для цього

використовуються спеціальні фільтри, такі як фільтр Шарра або фільтр Робертса. Підвищення різкості застосовується в фотографії та відео редакторах для покращення якості зображень.



Рисунок 1.8 – Посилення різкості

Збільшення яскравості та контрасту (brightness and contrast enhancement): цей метод збільшує яскравість та контраст зображення, зробивши світлі області світлішими і темні області темнішими. Збільшення яскравості та контрасту використовується в фотографії для покращення зображень та підвищення їх виразності.

Фільтрація кольору (color filtering): фільтрація кольору дозволяє змінювати колірну гаму зображення, наприклад, перетворюючи кольорові зображення в чорно-білі або надавати їм специфічний кольоровий відтінок. Фільтрація кольору використовується в фотофільтрах та фотофіксаторах для створення різних ефектів.

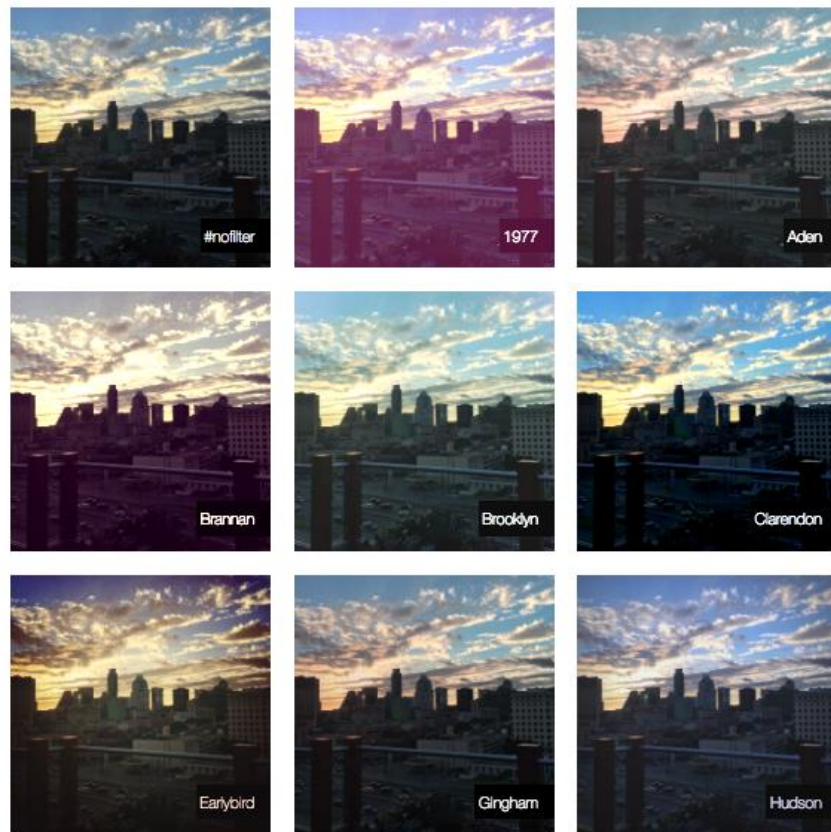


Рисунок 1.9 – Приклад застосування фільтрації кольору

Видалення шуму (noise reduction): шум на зображенні може бути видалено за допомогою різних алгоритмів, таких як фільтр медіани, фільтр гаусса або алгоритми машинного навчання. Видалення шуму важливо в фотографії, особливо при роботі з високочутливими камерами або при зйомці в умовах недостатнього освітлення.

Покращення роздільності (resolution enhancement): цей метод використовується для збільшення розмірів зображення, зберігаючи якість. Він використовує алгоритми інтерполяції для додавання додаткових деталей у зображення. Покращення роздільності використовується в системах спостереження, які мають обмежену розмірність зображення, а також для відновлення старих низькорозмірних фотографій[6].

Принцип роботи цих методів може варіюватися, але в багатьох випадках вони використовують технології машинного навчання, зокрема нейронні мережі.

Наприклад, для суперрезольюції можуть використовуватися глибокі згорткові нейронні мережі, що навчаються визначати шаблони та структуру зображення для збільшення роздільної здатності. У випадку усунення шуму, нейронні мережі можуть навчатися розпізнавати шум та видаляти його з зображення.

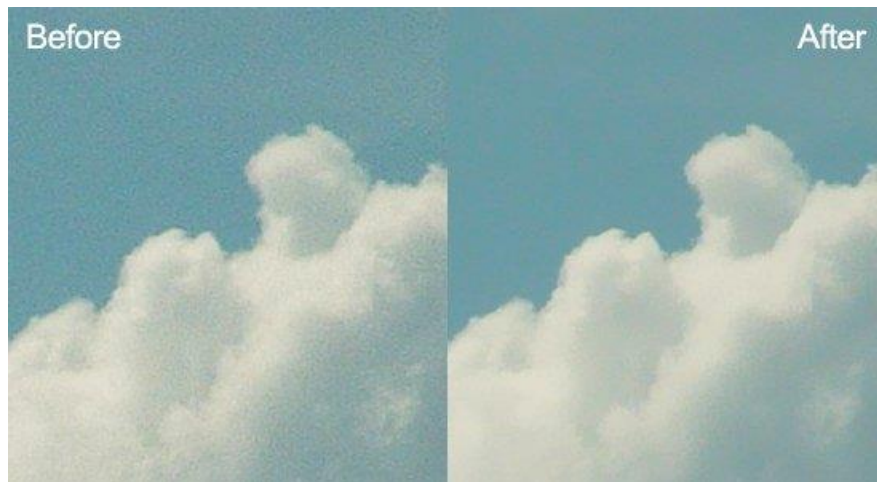


Рисунок 1.10 – Приклад застосування видалення шуму зображення

1.3 Галузі застосування інтелектуальної обробки зображень

Інтелектуальна обробка зображень має безліч застосувань у різних галузях, дозволяє комп'ютерам аналізувати, розуміти та взаємодіяти з візуальною інформацією. Завдяки швидкому розвитку технологій та збільшенню доступу до великих обсягів даних, інтелектуальна обробка зображень дозволяє нам аналізувати, розпізнавати та визначати широкий спектр об'єктів і явищ. Розглянемо роль інтелектуальної обробки зображень у таких галузях, як медицина, автомобільна промисловість, безпека, маркетинг та мистецтво.

1.3.1 Медицина

Інтелектуальна обробка зображень має значний вплив на медичну галузь. Вона використовується для аналізу медичних зображень, таких як рентгени, КТ-сканування, магнітно-резонансна томографія (МРТ) та ультразвукові

зображення. Роль інтелектуальної обробки зображень в цій галузі полягає в наступному:

1) діагностика та розпізнавання хвороб: За допомогою алгоритмів обробки зображень, медичні фахівці можуть вчасно виявляти та діагностувати хвороби, такі як рак, захворювання серця та захворювання нирок;

2) слідкування за пацієнтами: Обробка зображень дозволяє стежити за динамікою змін в стані пацієнтів, що є важливим у хірургії та реабілітації;

3) автоматизація діагностики: Інтелектуальні системи можуть автоматизувати процеси аналізу медичних зображень, що допомагає пришвидшити діагностику та зменшити ризик помилок.

Завдяки обробці зображень, можливо здійснювати віддалені консультації, надаючи пацієнтам доступ до медичної допомоги навіть не знаходячись у цей момент в одному приміщенні.

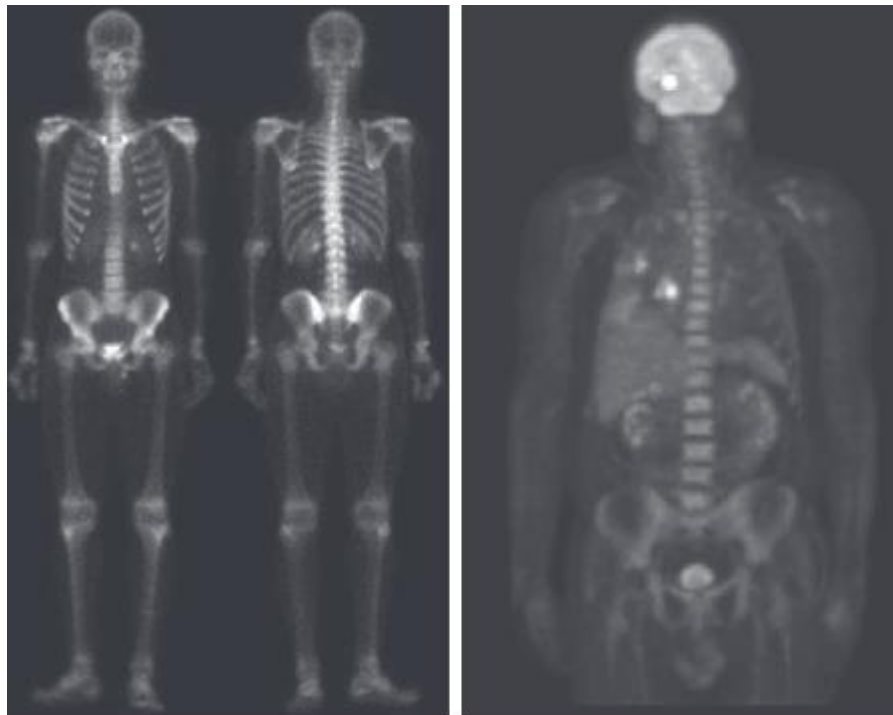


Рисунок 1.11 – Ренген-зображення тіла людини дуже популярний спосіб діагностики стану здоров'я

1.3.2 Автомобільна промисловість

У сучасних автомобілях інтелектуальна обробка зображень грає важливу роль в безпеці та зручності водіїв. Основні аспекти використання цієї технології в автомобільній промисловості включають:

1) автономні автомобілі: Системи інтелектуальної обробки зображень встановлені в автономних автомобілях для розпізнавання дорожніх знаків, інших транспортних засобів та пішоходів, що допомагає у прийнятті рішень та уникненні аварій;

2) паркування: Завдяки камерам та сенсорам, інтелектуальна обробка зображень дозволяє автомобілям паркуватися автоматично, надійно уникати перешкод та забезпечує безпеку при паркуванні;

3) системи слідкування за водієм: Вбудовані камери в автомобілях можуть визначати ознаки втоми або розсіяності водіїв, сприяючи збереженню життя на дорозі.

1.3.3 Безпека

Системи інтелектуальної обробки зображень грають важливу роль у галузі безпеки. Їх використовують для розпізнавання та відстеження осіб, виявлення незвичайної діяльності та контролю доступу. Основні аспекти включають:

1) системи відеоспостереження: Вони використовуються для виявлення злочинів, нагляду за об'єктами та об'єктами інфраструктури, ідентифікації потенційно небезпечних ситуацій;

2) розпізнавання облич: Технології розпізнавання облич використовуються для автоматичного визначення осіб на відео та фото, що може бути корисним у справах безпеки та розслідуванні кримінальних справ;

3) контроль доступу: Інтелектуальна обробка зображень застосовується для визначення допуску до об'єктів, будівель та інших областей, забезпечуючи безпеку приміщень.

1.3.4 Маркетинг

У маркетингу інтелектуальна обробка зображень використовується для покращення рекламних кампаній, аналізу споживчого ринку та розробки нових продуктів. Основні аспекти включають:

1) розпізнавання товарів: Завдяки інтелектуальній обробці зображень можливо розпізнавати товари на зображеннях, що спрощує процес онлайн-покупок та пошуку інформації про товари;

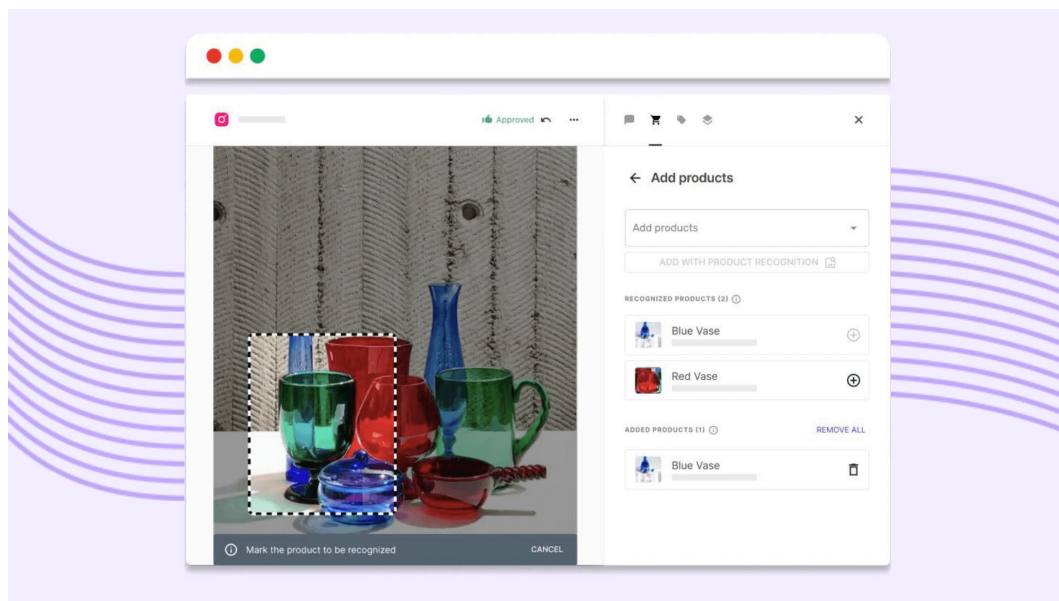


Рисунок 1.12 – Розпізнавання продукту на фотографії

2) аналіз соціальних мереж: Маркетологи використовують інтелектуальну обробку зображень для аналізу зображень та відео, які споживачі публікують у соціальних мережах, щоб зрозуміти їхні вподобання та настрої;

3) персоналізація реклами: Інтелектуальна обробка зображень допомагає створювати персоналізовані рекламні матеріали, які залучають увагу споживачів та підвищують ефективність рекламних кампаній.

1.3.5 Мистецтво

В мистецтві інтелектуальна обробка зображень розширює можливості художників та створює нові форми виразності [5]. Основні аспекти включають:

1) творення мистецьких творів: художники використовують інтелектуальну обробку зображень для створення унікальних мистецьких творів, які поєднують традиційні мистецькі техніки з сучасними технологіями;

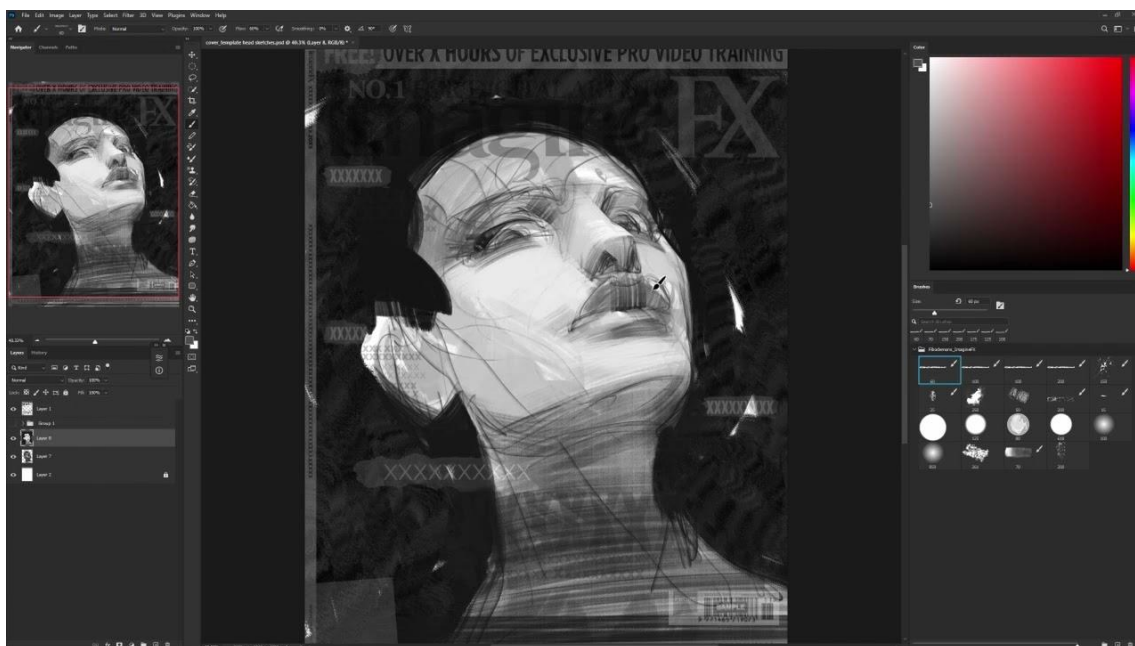


Рисунок 1.13 – Малювання за допомогою сучасної програми редагування зображень

2) візуальні ефекти у кіно та іграх: У сучасних фільмах та відеоіграх інтелектуальна обробка зображень використовується для створення реалістичних візуальних ефектів, таких як CGI (комп'ютерна графіка) та анімація;

3) культурна спадщина та віртуальні музеї: Інтелектуальна обробка зображень дозволяє відтворювати та відновлювати культурну спадщину, а також створювати віртуальні музеї для дослідження та освіти.

Інтелектуальна обробка зображень має безмежний потенціал у багатьох галузях і продовжує трансформувати наше сприйняття та використання зображень. Вона стала невід'ємною частиною сучасного світу та продовжує

розвиватися, надаючи нові можливості та вирішуючи важливі завдання в різних сферах нашого життя.

1.4 Мета та постанова завдання

Метою кваліфікаційної роботи є дослідження методів інтелектуальної обробки зображень та розробка моделі на базі обробки зображень з реального світу за допомогою AR технологій, на прикладі створення додатку розпізнавання книжкових обкладинок.

Виходячи з поставленої мети, у ході кваліфікаційної роботи повинні бути опрацьовані наступні завдання проектування:

- аналіз та вибір технології та методу інтелектуальної обробки зображень;
- вибір платформи для реалізації моделі;
- аналіз, вибір мов програмування та середовищ розробки додатку;
- створення прототипу додатку з розпізнаванням та відтворенням зображення з реального світу;
- доповнення додатку функціями розпізнавання, відстеження та доповнення об'єктів реального світу.

2 ПРОГРАМНО-АПАРАТНИЙ КОМПЛЕКС

Розробка моделі інтелектуальної обробки зображень полягає у розробці додатку, який би використовував Augmented reality технології для редагування оцифрованого зображення об'єкта реального світу. Перевагою такого додатку буде обробка зображення у реальному часі через камеру девайса, ідентифікація конкретного об'єкта та його доповнення. Даний додаток буде показувати обробку зображення у реальному часі на прикладі умов інтерактивної бібліотеки. Ціль моделі буде полягати у розпізнанні та обробці заздалегідь підготовленого зображення обкладинки книги, або автора, після обробки якого до зображення буде додаватись детальна інформація про об'єкт. Підчас створення моделі інтелектуальної обробки зображень було також розглянуто переваги та блокуючі фактори використання AR технологій як інструмент обробки зображень.

2.1 Вибір платформи реалізації проєкту

Оскільки однією з цілей цього проєкту є розробка портативного та доступного кожному додатка, у ході пошуку платформи для реалізації проєкту було розглянуто різні операційні системи мобільних девайсів, планшетів, портативних ноутбуків. Першим був відкинутий варіант використання Microsoft Windows, так як головною частиною девайсу для використання додатка є наявність камери, яка не завжди є в комп'ютерах, а процент мобільних девайсів, які працюють на ОС Windows значно менший ніж під управлінням Android або iOS.

Смартфони та планшети суттєво змінили наше життя сьогодні. Найпопулярнішими операційними системами для смарт-пристроїв є iOS від Apple та Android від Google.

Операційна система Android базується на ядрі Linux. Вона є відкритою системою, що дозволяє розробникам вносити зміни до коду та створювати

програмне забезпечення для різних пристроїв, таких як смартфони, планшети, телевізори і автомобільні системи.

Однією з основних переваг Android є його широка база користувачів. Завдяки великій кількості користувачів ця платформа стає привабливою для розробників, оскільки їх програми можуть отримати доступ до великої аудиторії.

Android також підтримує різноманітні пристрої від різних виробників, що дозволяє розробникам створювати програми для різних умов і характеристик.

Google Play Store, як офіційний магазин додатків для Android, забезпечує зручний спосіб розповсюдження і публікації додатків, що також створює сприятливі умови для розробників.

Щодо мов програмування, Android підтримує різні, такі як Java, Kotlin і C++, надаючи розробникам вибір у тому, якою мовою їм зручніше працювати.

Для розробників важливою є також можливість налаштувань та персоналізації, яку Android надає для інтерфейсу та функціоналу пристроїв.

У свою чергу, для розробників програмного забезпечення, Android пропонує великий ринок, доступ до внутрішніх функцій системи та апаратного забезпечення, широкий інструментарій розробки.

Суттєвою особливістю iOS є її замкненість. Вона використовується тільки на пристроях Apple, що дозволяє компанії забезпечити високий рівень оптимізації і контролю за якістю продукту, а також зберегти унікальність своїх девайсів на ринку портативної техніки. Ця система входить в широкую екосистему продуктів Apple, таких як Mac, Apple Watch і Apple TV, що сприяє легкій інтеграції різних пристроїв і сервісів.

iOS славиться своїм простим інтуїтивним інтерфейсом, який полегшує використання пристроїв. Крім того, велика увага приділяється системі безпеки, яка включає в себе шифрування даних, ідентифікацію за допомогою Touch ID або Face ID, а також строгий контроль за додатками в App Store.

Для розробників програмного забезпечення iOS пропонує стандартизоване середовище роботи. Розробники можуть працювати в умовах, де їм необхідно тестувати і оптимізувати додатки лише для обмеженої кількості пристроїв Apple.

Apple встановлює строгі стандарти для додатків, доступних в App Store.

Також розробник ОС надає інструменти для розробки додатків, такі як Xcode, які допомагають у створенні, тестуванні і оптимізації додатків для iOS.

У ході порівняння цих двох систем, для реалізації проєкту було вирішено обрати Android, у зв'язку з відкритою та прозорою системою файлів, доступністю у використанні, та легкістю тестування, а також великим списком підтримуваних мов програмування. Але у подальшому розвитку додатку, планується додати iOS до списку підтримуваних операційних систем додатку.

2.2 Вибір технологій інтелектуальної обробки зображень та інструментів

Для розробки моделі інтелектуальної обробки зображень була обрана технологія Доповненої реальності (Augmented Reality) за її новизну та потенціал в поліпшенні життя звичайних людей, наприклад, бібліотекарів. Робота з AR функціями та програмування додатку потребує використання графічного рушія, в якому і буде проводитись уся розробка макету. У ролі такого рушія була обрана середа розробки Unity 3d, яка підтримує більшість мов програмування, має вражаючий функціонал та велику кількість плагінів, у томи числі AR. Також Unity дозволить родити крос платформні додатки для мобільних девайсів.

2.2.1 AR технологія як метод інтелектуальної обробки зображень

Доповнена реальність (AR) визначається як прямий або опосередкований перегляд у реальному часі фізичного оточення реального світу, покращений/розширений шляхом додавання віртуальної комп'ютерної інформації [12]. AR є інтерактивною, записаною у форматі 3D і поєднує реальні та віртуальні об'єкти. Реально-віртуальний континуум Мілгрема був визначений Полом Мілгремом і Фумію Кішіно як континуум, що охоплює реальний світ і віртуальне середовище і включає в себе доповнену реальність і віртуальну

реальність (VR).

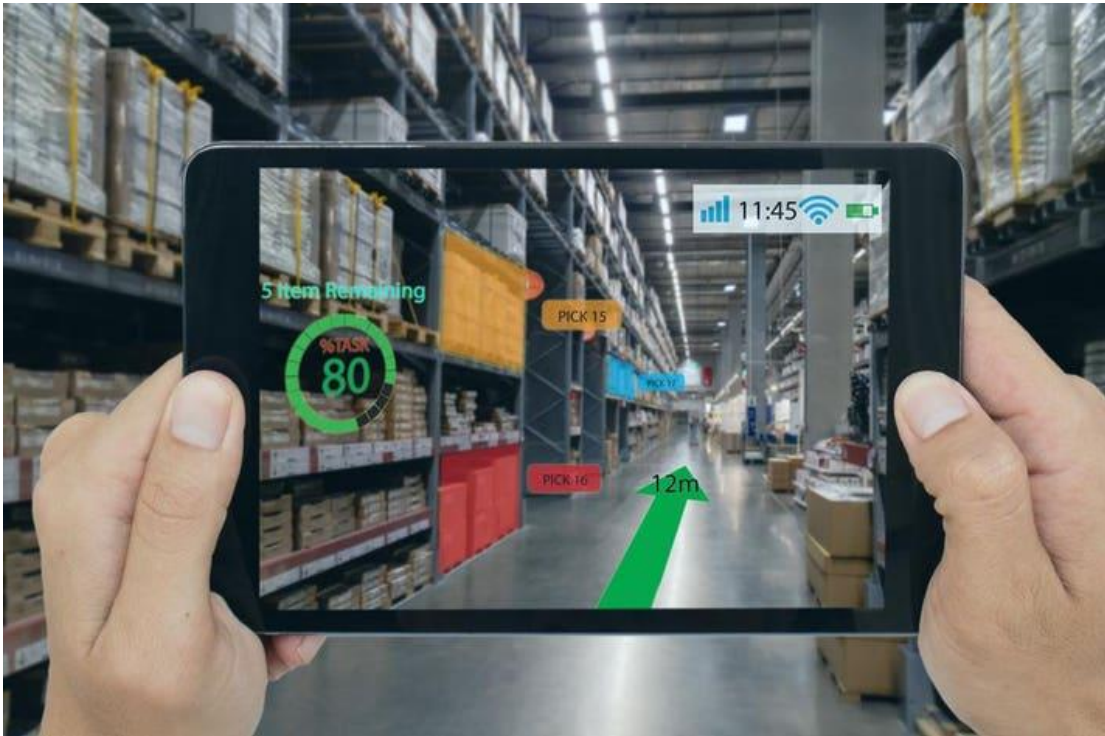


Рисунок 2.1 – Приклад використання AR

Доповнена реальність має на меті полегшити життя користувачів, надаючи віртуальну інформацію не лише про їхнє безпосереднє оточення, але й про всі опосередковані види реального світу, такі як потокове відео в реальному часі; доповнена реальність покращує сприйняття та взаємодію користувачів з реальним світом. Віртуальна реальність – це новий тип доповненої реальності, покликаний покращити досвід користувача в реальному світі. Технологія віртуальної реальності (VR), або віртуальне середовище в термінах Мілгрема, занурює користувача в повністю штучний світ, без можливості бачити реальний світ, тоді як технологія доповненої реальності покращує реальність, накладаючи віртуальні об'єкти і підказки на реальний світ в режимі реального часу. Доповнена реальність не обмежується певним типом технології відображення, наприклад, дисплеєм, що монтується на голові, або зором; AR має потенціал для застосування до всіх органів чуття і може покращити нюх, дотик і слух; AR

також може бути використана для доповнення або заміни втрачених органів чуття користувача за допомогою сенсорного заміщення. Наприклад, зір сліпого або користувача зі слабким зором може бути посилений за допомогою аудіосигналів, а слух глухого користувача – за допомогою візуальних сигналів.

Azuma та ін. [4] також розглядають додатки доповненої реальності, які, окрім додавання віртуальних об'єктів, повинні видаляти реальні об'єкти з навколишнього середовища. Фактично, видалення об'єкта з реального світу відповідає накладенню на нього віртуальної інформації, яка відповідає фону, щоб створити у користувача враження, що об'єкта там немає. Віртуальні об'єкти, додані в реальне середовище, показують користувачеві інформацію, яка не може бути безпосередньо сприйнята органами чуття. Наприклад, цифрова інформація може відображатися через гарнітуру, щоб допомогти працівникам орієнтуватися в проводах в літаку. Іноді ця інформація слугує лише для розважальних цілей, як у Wikitude та інших мобільних додатках доповненої реальності. Існує багато інших типів додатків доповненої реальності, включаючи медичну візуалізацію, розваги, рекламу, технічне обслуговування і ремонт, анотації та планування маршрутів роботів.

Комп'ютерний зір візуалізує тривимірні віртуальні об'єкти з тієї ж перспективи, що й зображення реальної сцени, зняті камерами спостереження. Для запису зображень у доповненій реальності використовуються різні методи комп'ютерного зору, здебільшого пов'язані з відеоспостереженням.

Ці методи зазвичай складаються з двох етапів: відстеження та реконструкції/розпізнавання. На першому етапі на зображенні з камери виявляються орієнтири, оптичні зображення або об'єкти, що становлять інтерес.

На етапі відстеження для інтерпретації зображення з камери можуть використовуватися функції виявлення особливостей, виявлення країв або інші методи обробки зображень. У комп'ютерному зорі більшість існуючих методів відстеження можна розділити на два класи: на основі об'єктів і на основі моделей [7]. Об'єктно-орієнтовані методи базуються на виявленні взаємозв'язку між об'єктами на 2D-зображеннях та їхніми координатами в 3D-картині світу [6].

Методи на основі моделей використовують функціональні моделі об'єкта, що відстежується, такі як CAD-моделі або шаблони 2D-об'єктів на основі характерних ознак [7]. Після встановлення зв'язку між 2D-зображенням і 3D-картиною світу 3D-координати об'єкта проєктуються на координати спостережуваного 2D-зображення, і відповідне положення камери може бути отримане шляхом мінімізації відстані до 2D-об'єкта. Обмеження для оцінки положення камери зазвичай задаються за допомогою точкових характеристик. На етапі реконструкції/розпізнавання реальна система координат реконструюється з використанням даних, отриманих на першому етапі.

За наявності відкаліброваної камери та моделі перспективної проєкції, якщо точка має координати $(x, y, z)^T$ у системі координат камери, її проєкція на площину зображення $(x/z, y/z, 1)^T$

У точкових обмеженнях ми маємо дві головні системи координат (рисунок 1.1), світову систему координат W і систему координат двовимірного зображення. Нехай $p_i(x_i, y_i, z_i)^T$, $i = 1, \dots, n$, з $n \geq 3$, бути набором тривимірних неколінеарних опорних точок у світовій системі координат і $(q_i^x; q_i^y; q_i^z)^T$ – відповідні координати простору камери, p_i і q_i пов'язані таким перетворенням:

$$q_i = R p_i + T \quad (2.1)$$

$$\text{де } T = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \text{ і } R = \begin{pmatrix} r_1^T \\ r_2^T \\ r_3^T \end{pmatrix} \quad (2.2)$$

є матрицею обертання та вектором трансляції відповідно.

Помилка простору зображення дає зв'язок між опорними точками 3D, їхніми відповідними точками виділеного 2D зображення та параметрами пози камери та відповідає обмеженням точки [6].

Деякі методи припускають наявність опорних маркерів у середовищі чи об'єкті з відомою 3D-геометрією та використовують ці дані. Інші мають

попередньо розраховану 3D-структуру сцени, наприклад пристрій AR-View Хуанга та ін. [12]; однак пристрій має бути нерухомим, а його положення відомо. Якщо вся сцена не відома заздалегідь, техніка одночасної локалізації та відображення (SLAM) використовується для відображення опорних маркерів або відносних положень 3D-моделей. У випадку, коли неможливо зробити припущення щодо 3D-геометрії сцени, використовується метод структури з руху (SfM). Метод SfM можна розділити на дві частини: відстеження точки характеру та оцінка параметрів камери.

Методи відстеження в доповненій реальності здебільшого залежать від типу середовища, у якому буде працювати пристрій доповненої реальності, а також від типу системи доповненої реальності. Середовище може бути в приміщенні, на відкритому повітрі або в поєднанні обох. Таким же чином система може бути мобільною або статичною (з фіксованим положенням). Наприклад, якщо пристрій AR є пристроєм із фіксованим положенням для зовнішнього реального середовища, таким як пристрій AR-View Хуанга та ін. [12], розробники можуть використовувати механічне відстеження, оскільки всі рухи, які слід відстежувати, будуть механічний, оскільки відомо положення пристрою. Цей тип середовища та системи полегшує відстеження середовища для збільшення оточення. З іншого боку, якщо пристрій AR мобільний і розроблений для зовнішнього середовища, відстеження стає набагато складнішим, а різні методи мають певні переваги та недоліки. Наприклад, Nilsson et al. [7] створили систему виявлення пішоходів для уникнення автомобільних зіткнень за допомогою AR. Їх система мобільна і вулична. Для камери, що рухається в невідомому середовищі, проблема комп'ютерного зору полягає в тому, щоб реконструювати як рух камери, так і структуру сцени за допомогою зображення та додаткових послідовностей даних сенсора. У цьому випадку, оскільки неможливо зробити припущення щодо тривимірної геометрії сцени, метод SfM використовується для реконструкції сцени.

Розробники також мають вибір використовувати існуючі бібліотеки AR, такі як ARToolKit. ARToolKit, розроблений у 1999 році Хірокадзу Като з

Інституту науки і технологій Нарі та випущений лабораторією НІТ Університету Вашингтона, є бібліотекою відстеження комп'ютерного зору, яка дозволяє користувачеві створювати програми доповненої реальності [7]. Він використовує можливості відстеження відео, щоб обчислити в режимі реального часу реальне положення та орієнтацію камери відносно фізичних маркерів. Коли стане відомо реальне положення камери, віртуальну камеру можна розмістити в тому самому точному положенні та намалювати 3D-модель комп'ютерної графіки для накладання маркерів. Розширеною версією ARToolKit є ARToolKitPlus, яка додала багато функцій до ARToolKit, зокрема API на основі класів; однак він більше не розробляється і вже має наступника: Studierstube Tracker. Концепції Studierstube Tracker дуже схожі на ARToolKitPlus; однак його кодова база зовсім інша, і він не є відкритим кодом, тому недоступний для завантаження. Він підтримує мобільний телефон із Studierstube ES, а також комп'ютери, що робить його вимоги до пам'яті дуже низькими (100 КБ або 5–10% від ARToolKitPlus) і дуже швидкою обробкою (приблизно вдвічі швидше, ніж ARToolKitPlus на мобільних телефонах, і приблизно 1 мс на кадр на ПК) [9]. Studierstube Tracker дуже модульний; розробники можуть розширити його, створивши для нього нові функції. Коли вперше представляли Studierstube [9], розробники мали на увазі користувальницький інтерфейс, який «використовує доповнену реальність для спільної роботи, щоб поєднати кілька вимірів користувацького інтерфейсу: кілька користувачів, контексти та локалі, а також програми, 3D-вікна, хости, платформи відображення та операційні системи». Більше інформації про Studierstube можна знайти на [9].

Хоча візуальне відстеження тепер має можливість розпізнавати та відстежувати багато речей, воно здебільшого покладається на інші методи, такі як GPS і акселерометри. Наприклад, комп'ютеру виявити та розпізнати автомобіль дуже важко. Поверхня більшості автомобілів одночасно блискуча і гладка, і більшість характерних точок походять від відображень і, отже, не мають значення для оцінки пози та навіть іноді розпізнавання [9]. Кілька стабільних елементів, які можна сподіватися розпізнати, наприклад, кути вікон або колеса,

надзвичайно важко зіставити через відображення та прозорі частини. Хоча цей приклад є дещо екстремальним, він показує труднощі та проблеми, з якими стикається комп'ютерний зір із більшістю об'єктів неправильної форми, такими як їжа, квіти та більшість предметів мистецтва.

Художники десятиліттями використовували технології розширеної реальності (XR), такі як AR і віртуальна реальність (VR), але нещодавно AR зацікавився різними галузями. VR представляв першу хвилю золотої лихоманки XR. Незважаючи на величезний приплив інвестицій і значні технологічні досягнення, це залишається дорогим і трудомістким у створенні, і для цього потрібні спеціальні гарнітури. Навпаки, більшість сучасних смартфонів і планшетів уже мають вбудовані можливості AR, які вдосконалюються з кожним днем. Багато спостерігачів очікують, що поворотний момент – у вигляді доступних, легких, повнофункціональних окулярів AR – настане десь цього десятиліття. Ходять чутки, що Apple представить свою надсекретну модель у 2022 році, але хто знає, чи стане цей пристрій святим Граалем, яким не змогли стати Google Glass у 2013 році чи Magic Leap у 2019. Тим часом ціле покоління стало звикнути до роботи зі смартфоном AR, від Pokémon GO до створених користувачами фільтрів Snapchat і Instagram. AR тепер є доступною версією старої науково-фантастичної мрії про змішану реальність, головним чином тому, що накладання віртуальних зображень на реальний світ (на відміну від повного занурення VR у віртуальне) має свою особливу привабливість для хоста концептуальних та естетичних причин.

2.2.2 Unity 3d як середовище розробки програмних додатків

Unity 3D – це інтегроване середовище розробки для створення крос-платформних ігор і додатків. Воно надає розробникам потужні інструменти для створення і візуалізації 2D та 3D графіки, а також для розробки фізики, штучного інтелекту, аудіо та інших складових ігрового середовища. Unity 3D дозволяє створювати додатки для різних платформ, таких як Windows, macOS, Linux, iOS,

Android, Xbox, PlayStation, VR-пристрої та інші. Ось декілька ключових аспектів Unity 3D як серед розробки крос-платформних додатків:

Мови програмування: Unity 3D підтримує декілька мов програмування, таких як C#, JavaScript (але варто відзначити, що підтримка JavaScript була припинена після версії Unity 2018.2), Boo. Саме C# є найбільш популярним вибором серед розробників.

Графіка і анімація: Unity 3D надає потужні інструменти для створення відмінної 2D та 3D графіки, а також анімації. Ви можете створювати 2D та 3D моделі, використовуючи власні або імпортовані ресурси.

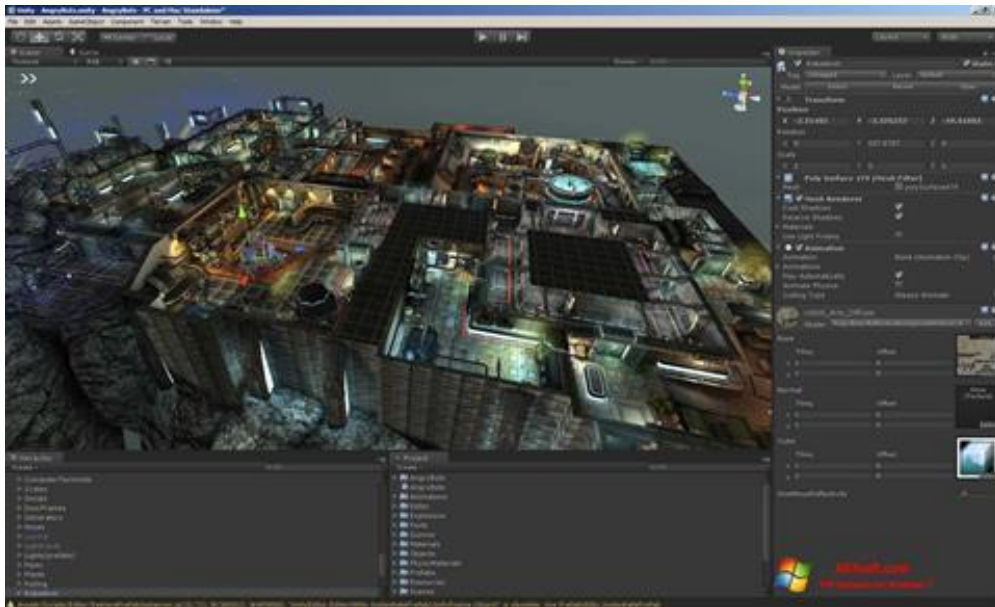


Рисунок 2.2 – Серед розробки Unity

Фізика: Unity має вбудований фізичний рушій, що дозволяє створювати реалістичну поведінку об'єктів у грі. Також в Unity можливо налаштовувати фізичні параметри для об'єктів та створювати реалістичну динаміку гри.

Крос-платформна розробка: Unity підтримує розробку додатків для різних платформ. Можливо створювати гру один раз і публікувати її на різних платформах, забезпечуючи крос-платформну сумісність.

Розвиваюча спільнота: Unity має велику та активну спільноту розробників,

що допомагає розв'язувати проблеми та надає безліч ресурсів, включаючи форуми, документацію та онлайн-курси.

Віртуальна реальність та доповнена реальність: Unity 3D підтримує розробку додатків для віртуальної реальності (VR) та доповненої реальності (AR), що робить його ідеальним вибором для створення додатків для таких пристроїв, як Oculus Rift, HTC Vive, HoloLens та інші.

Unity 3D – це потужне середовище розробки, яке дозволяє створювати крос-платформні додатки для ігор, віртуальної реальності, доповненої реальності та інших застосувань. Завдяки широкому спектру можливостей та підтримці різних платформ, воно є популярним вибором серед розробників у галузі ігрової та розширеної реальності.

2.3 Вибір мови програмування

Мовою програмування додатку була обрана C# у зв'язку з її функціоналом та повною сумісністю з Unity3D.

C# (C-Sharp) – це об'єктно-орієнтована мова програмування, розроблена Microsoft, яка є частиною .NET Framework. Вона використовується для розробки різноманітних додатків, включаючи ігри, веб-додатки та служби.

В контексті Unity 3D, C# використовується для написання скриптів, які визначають поведінку об'єктів у грі. Вона забезпечує можливості контролю за різними аспектами гри, такими як введення користувача, реалізація штучного інтелекту, інтеракція з гравцем та багато іншого.

Однією з переваг використання C# в Unity є його легкість вивчення та зрозумілий синтаксис. Це робить її досить доступною для новачків та забезпечує читабельність коду.

Мова підтримує об'єктно-орієнтовану парадигму, дозволяючи розробникам структурувати код за допомогою класів та об'єктів. Це полегшує розробку та обслуговування кодової бази.

Інтеграція C# з API Unity робить його ідеальним для роботи з ігровим

двигуном. Розробники можуть взаємодіяти з різноманітними функціями та можливостями Unity, отримуючи доступ до різних функціональних можливостей для створення складних ігрових механік.

Загалом, використання C# в Unity 3D дозволяє розробникам створювати інтерактивні 3D-додатки з потужними можливостями, інтуїтивним синтаксисом та зручним API.

3 ПРОГРАМНА РЕАЛІЗАЦІЯ ПРОЕКТУ

3.1 Етапи реалізації проєкту

Розробка інтелектуальної моделі обробки зображень з використанням Unity 3D та AR є досить складною та часомісткою задачею. Враховуючи це, було створено загальний план дій, якого було дотримано під час створення моделі:

- 1) встановлення робочого середовища;
- 2) встановлення Unity 3D та створення нового проєкта;
- 3) встановлення AR бібліотеки;
- 4) використання AR бібліотеку, таку як AR Foundation, Vuforia або ARKit/ARCore, щоб додати можливості розпізнавання та відслідковування маркерів чи зображень;
- 5) створення 3D моделей;
- 6) створення 3D об'єктів, які необхідно додати до реального зображення. Також буде використано інструменти моделювання в Unity або імпорт готових моделей;
- 7) створення реалізації розпізнавання зображень;
- 8) використання AR бібліотеки для розпізнавання маркерів або зображень через камеру смартфона або іншого пристрою;
- 9) додавання цієї логіки до Unity проєкту, яка буде визначати, коли реальне зображення (маркер) знаходиться в полі зору камери;
- 10) розміщення 3D об'єктів:
 - i) після розпізнавання реального зображення додавання 3D об'єктів до сцени Unity. У даному кроці можливе використання позиційної інформації, отриману від AR бібліотеки, для правильного розміщення об'єктів;
- 11) відображення додаткових об'єктів:
 - i) налаштування анімації та поведінки додаткових об'єктів. Наприклад, вони можуть з'являтися, коли маркер розпізнається і зникають, коли маркер

виходить з полі зору камери;

12) Тестування та налагодження:

i) Тестування проєкта з різними маркерами та об'єктами. Переконайтесь, що об'єкти правильно розміщуються та реагують на рухи камери;

13) Оптимізація та публікація:

i) Оптимізування проєкта для публікації на платформі (наприклад, iOS, Android). Розповсюдження додатка, який використовує цю модель обробки зображень через камеру.

3.2 Налаштування середовища Unity 3d

У зв'язку з тим, що Unity 3d залежна від обраної версії (існує дуже багато версій, закріплених за роком розробки версії рушія), необхідно обрати таку версію Unity, яка б підтримувала розробку AR додатків. Такою була обрана версія 2022.1.5f1. Також, ця версія має можливість підключення Android, iOS, Windows підтримку збірки білдів, що також важливо для розробки додатку, та його тестування.

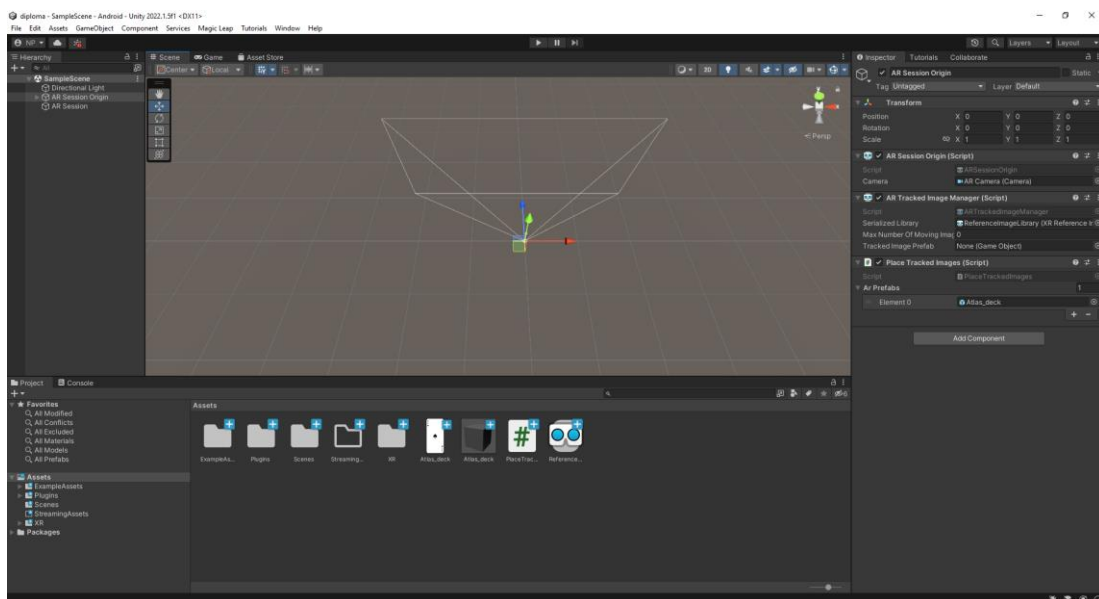


Рисунок 3.1 – Середовище розробки Unity

Також, важливо правильно виставити Player's settings, які відповідають за

коректе функціонування програми та збірку білдів. У цих налаштуваннях також можна виставити плагін для роботи з AR функціями.

Для коректної роботи усіх компонентів Unity 3d необхідно переконатись, що усі потрібні модулі підключені та оновлені до останньої версії. Для цього треба скористатись Package Manager з опції Window головної панелі інструментів.

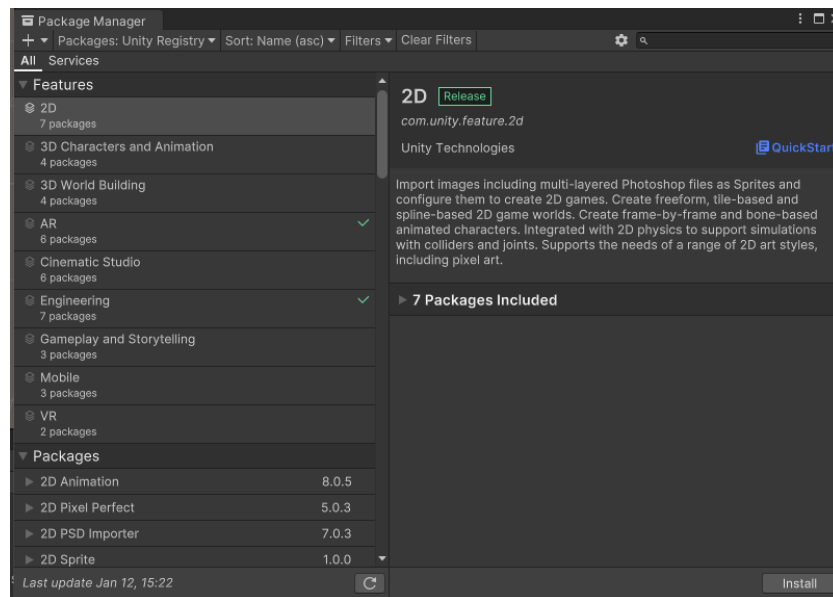


Рисунок 3.2 – Середовище розробки Unity

Таким чином, робоче середовище Unity можна вважати налаштованим та готовим для розробки моделі іфнтелектуальної обробки зображень за допомогою Доповненої реальності.

3.3 Підключення плагінів

Розглянемо плагіни, які доступні в Unity, для розробки додатку з доповненою реальністю, та оберемо той, який найкраще підійде для охоплення цілей проекту.

Один із таких інструментів – Vuforia. Це розширена платформа для

створення AR-додатків, яка включає в себе функціонал розпізнавання зображень та відстеження об'єктів. Інтегрована з Unity, вона дозволяє створювати додатки для різних платформ, включаючи Android і iOS.

Ще один інструмент – AR Foundation, розроблений Unity. Цей фреймворк служить абстракцією для взаємодії з AR-технологіями на різних платформах. AR Foundation дозволяє розробляти додатки, які сумісні як з ARKit (iOS), так і з ARCore (Android), використовуючи спільний код.

Також варто згадати про Wikitude, платформу для розробки AR-додатків, яка надає Unity SDK і підтримує функції, такі як розпізнавання зображень, геолокація та розпізнавання об'єктів.

Кожен з цих інструментів має свої особливості та можливості, спрямовані на полегшення розробки AR-додатків в Unity та забезпечення високого рівня взаємодії з доповненою реальністю на різних платформах, але в рамках проекту було прийнято рішення розробляти додаток використовуючи AR Foundation, так як цей плагін є нативним та розробленим від розробників Unity для самої Unity та має багатий арсенал функцій, які допоможуть виконати задані цілі проекту та навіть лишають за собою потенціал для майбутніх розширень додатку.

Для підключення потрібного плагіну необхідно перейти до налаштувань проекту та клікнути на вкладку «Плагіни». Оскільки модель обробки зображень за допомогою доповненої реальності буде цілитись на запуск на портативних девайсах, необхідно обрати ARCore for Android.

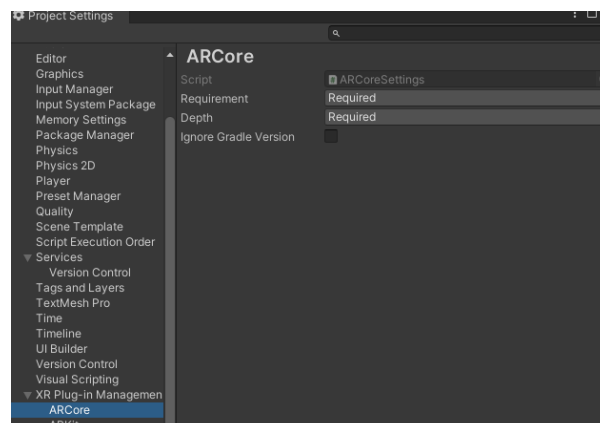


Рисунок 3.3 – Налаштування плагіну доповненої реальності ARCore

3.4 Створення та налаштування макета-додатка обробки зображень у реальному часі

Першим кроком у створенні макету інтелектуальної обробки зображень був вибір тематики самого додатку. Сама обробка зображення полягає в використанні технології доповненої реальності, а середа застосування була обрана як інтерактивна бібліотека. Суть проєкта полягає у спрощенні роботи з книгами у бібліотеці, їх пошуку, маркування та знаходження потрібної книги.

Основною метою проєкта поставимо ідею розпізнавання мобільним девайсом обкладинки книги та доповнення її зображення, отриманого з реального світу, додатковою інформацією. Такою, як фотографія автора, рік видання, основна інформація про книгу.

По-перше, після обрання потрібної версії середі розробки, необхідно створити проєкт з пресетом під розробку AR додатків, для автоматичного доналаштування проєкту та розміщення необхідних файлів у папці проєкту системою розробки.

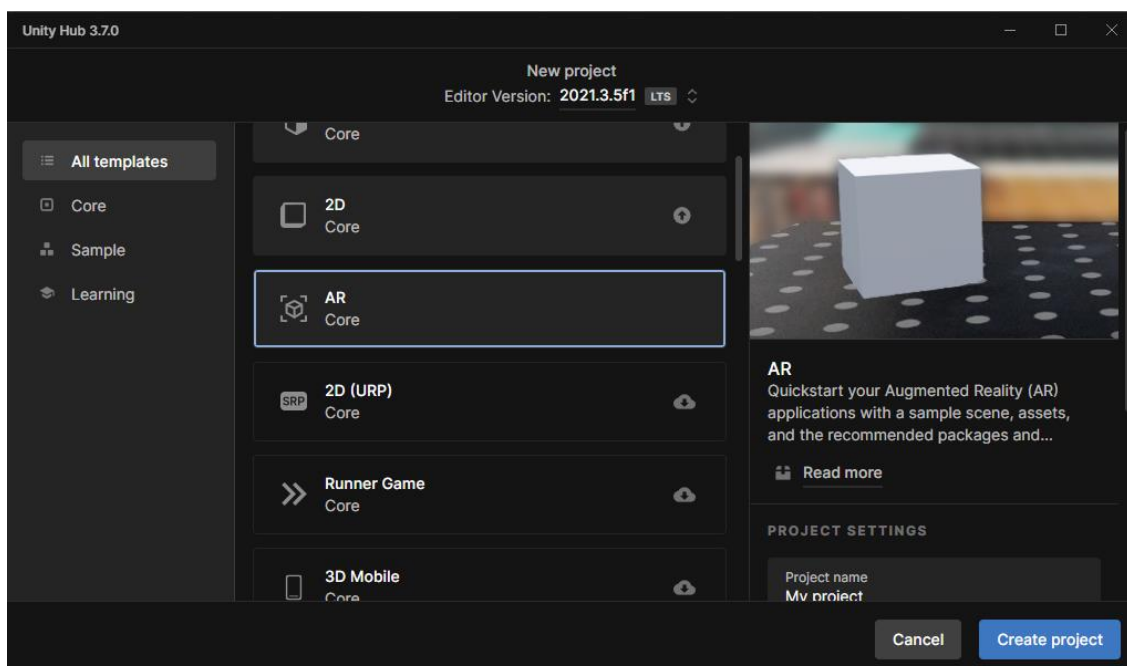


Рисунок 3.4 – Створення проєкту

Після створення проєкту, відкривається середовище розробки, де вже з'явилися три об'єкти сцени проєкту: Directional Light, AR Session Origin, AR Session.

Перший об'єкт відповідає за постановку світла на інтерактивному екрані сцени, AR Session Origin відповідає за виконання скриптів та модулів, що відповідають за імплементацію функцій роботи з доповненою реальністю. Саме сюди у рамках проєкту будуть додаватися модулі відслідковування об'єктів та їх доповнення графікою, а також буде розроблений скрипт керування поведінкою додатка у тих чи інших триггерних ситуаціях.

У нижній частині екрану розташований File Explorer і папка Assets де будуть зберігатися усі необхідні файли проєкту.

Наступним кроком буде розробка модуля для трекінгу об'єкта реального світу. Для цього скористаємось AR Tracked Image Manager. AR Tracked Image Manager потребує зображення об'єкта, який буде відстежуватися за допомогою камери, тому необхідно додати

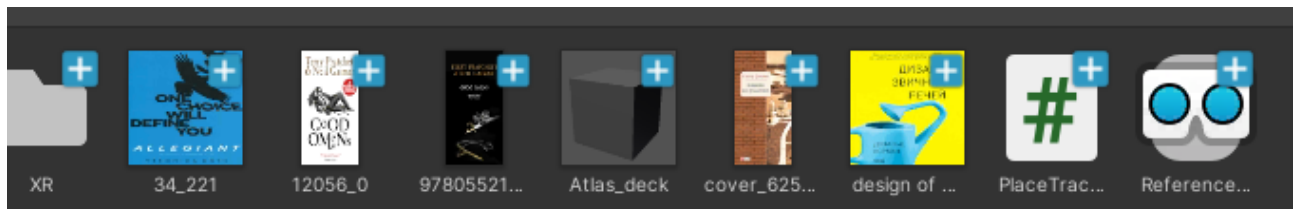


Рисунок 3.5 – Приклад доданих ассетів обкладинок

Після додавання необхідних ассетів, необхідно скласти з них бібліотеку референсів, якими буде оперувати машинний зір камери та, якщо таке зображення відтвориться у реальному світі через камеру, то це спричинить розгортання триггерного скрипту. Бібліотека це теж ассет, і додається через головне меню, опція Компоненти.

За створенням бібліотеки референсних зображень, потрібно підключити її як

опцію в компоненті AR Tracked Image Manager. Далі створюється скрипт, який відповідає за те, щоб при появленні одного з цих зображень, воно доповнювалось префабом з додатковою інформацією та моделями.

Скрипт складається з декількох функцій, основна з яких PlaceTrackedImages, де відбуваються дії пошуку збігів зображення реальності та зображень з бібліотеки референсів, і, якщо таке зображення буде знайдено, програма захоплює назву зображення та шукає префаб з такою назвою, якщо префаб знайдений – він накладається на об'єкт з реального світу.

Лістинг 3.1 – Функції пошуку відстежуваного зображення та накладання префабу

```
// Loop through all new tracked images that have been detected
foreach (var trackedImage in eventArgs.added) {
// Get the name of the reference image
var imageName = trackedImage.referenceImage.name;
// Now loop over the array of prefabs
    foreach (var curPrefab in ArPrefabs) {
        // Check whether this prefab matches the tracked image name,
        and that
            // the prefab hasn't already been created
        if (string.Compare(curPrefab.name, imageName,
StringComparison.OrdinalIgnoreCase) == 0
            && !_instantiatedPrefabs.ContainsKey(imageName)) {
// Instantiate the prefab, parenting it to the ARTrackedImage
var newPrefab = Instantiate(curPrefab, trackedImage.transform);
// Add the created prefab to our array
_instantiatedPrefabs[imageName] = newPrefab;
}
}
}
```

Написаний скрипт необхідно також додати як новий компонент у AR Session Origin, після чого, можна розробляти префаби для кожного зображення в бібліотеці референсів. Назва префабу обов'язково повинна співпадати з назвою відстежуваного зображення, до якого цей префаб повинет додаватись.

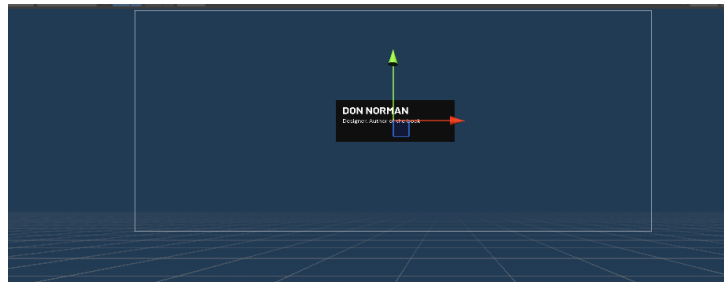


Рисунок 3.6 – Створення префабу з додатковими зображеннями

Після того, як усі префаби будуть створені, їх потрібно закріпити за скриптом, додавши їх як контент для змінних.

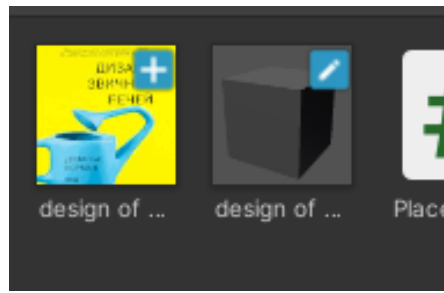


Рисунок 3.7 – Зображення і відповідний йому префаб

Для того, щоб протестувати додаток, можливо скористатись функцією збірки білдів, які потім можна запустити на портативному девайсі. У результаті збірки білда утвориться файл з розширенням ark, який втсановлюється на смартфонах, як звичайний Android-додаток. Важливо також надати усі потрібні доступи додатку при першому відкритті.

Результатом розробки додатку є ark з автоматично ввімкненою камерою девайсу, через яку можна вести пошук об'єктів-книжок, при наведенні на один з таких об'єктів, вихідна інформація з камери доповнюється додатковими зображеннями, накладеними за допомогою доповненої реальності.



Рисунок 3.8 – Результат роботи додатку

ВИСНОВКИ

Під час виконання кваліфікаційної роботи були досліджені методи інтелектуальної обробки зображень та розроблено додаток інтерактивної бібліотеки з використанням AR-технологій.

На даній ітерації макету інтелектуальної обробки зображень є достатній функціонал для аналізування, розпізнавання та доповнення об'єктів реального світу за допомогою камери мобільного Android-девайсу та AR технологій.

В планах на подальші версії доповнення бази даних з інформацією про книги, доповнення функціоналу (додавання архівних номерів на книги, можливість додавати інформацію та нові книги за допомогою інтерфейсу додатка).

Інтелектуальна обробка зображень є потужним інструментом, який змінює багато аспектів нашого життя. Від медицини до транспорту та маркетингу, вона допомагає автоматизувати процеси аналізу та прийняття рішень. ІОЗ продовжує розвиватися і розширювати свої можливості, вносячи значний внесок у різні галузі і покращуючи якість нашого життя.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Augmented Reality Art Creates New Opportunities—and New Contexts—for Works [Електронний ресурс] // 2022 – Режим доступу до ресурсу: <https://redshift.autodesk.com/articles/augmented-reality-art>.
2. What Is the Metaverse, and Will It Be Worth the Wait? [Електронний ресурс]. – 2022. – Режим доступу до ресурсу: https://www.washingtonpost.com/business/what-is-the-metaverse-and-will-it-be-worth-the-wait/2022/10/27/68cd52bc-5666-11ed-ac8b-08bbfab1c5a5_story.html.
3. Bajura M. Dynamic Registration Correction in Video-Based Augmented Reality Systems / M. Bajura, U. Neumann., 5. – (IEEE Computer Graphics and Applications). – (15).
4. T. Azuma et al., “A survey of augmented reality,” *Presence*, vol. 6, no. 4, pp. 355–385, 1997.
5. R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, “Recent advances in augmented reality,” *Computer Graphics and Applications*, IEEE, vol. 21, no. 6, pp. 34–47, 2001.
6. I. E. Sutherland and C. A. Mead, “Microelectronics and computer science,” *Scientific American*, vol. 237, pp. 210–228, 1977.
7. T. P. Caudell and D. W. Mizell, “Augmented reality: An application of heads-up display technology to manual manufacturing processes,” in *System Sciences*, 1992. *Proceedings of the Twenty-Fifth Hawaii International Conference on*, vol. 2. IEEE, 1992, pp. 659–669.
8. A. L. Janin, D. W. Mizell, and T. P. Caudell, “Calibration of head-mounted displays for augmented reality applications,” in *Virtual Reality Annual International Symposium*, 1993., 1993 IEEE. IEEE, 1993, pp. 246–255.
9. R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, “Recent advances in augmented reality,” *Computer Graphics and Applications*, IEEE, vol. 21, no. 6, pp. 34–47, 2001.

10. M. de S'á and E. Churchill, "Mobile augmented reality: exploring design and prototyping techniques," in Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services. ACM, 2012, pp. 221–230.
11. O. Bimber, R. Raskar, and M. Inami, Spatial augmented reality. AK Peters Wellesley, 2005.
12. F. Zhou, H. B.-L. Duh, and M. Billinghurst, "Trends in augmented reality tracking, interaction and display: A review of ten years of ismar," in Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality. IEEE Computer Society, 2008, pp. 193–202.
13. A. Shatte, J. Holdsworth, and I. Lee, "Mobile augmented reality based context-aware library management system," Expert Systems with Applications, vol. 41, no. 5, pp. 2174–2185, 2014.