

Хмельницький національний університет  
Факультет інформаційних технологій  
Кафедра комп'ютерної інженерії та інформаційних систем

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

Галузь знань 12 – Інформаційні технології

Спеціальність 123 – Комп'ютерна інженерія

на тему:

«Метод та засоби ідентифікації об'єктів у тривимірних хмарах технологіями комп'ютерного зору та машинного навчання»

КвРКІ. 170172.21.01.15 ПЗ

Виконав: студент 2 курсу, група КІ2М-21-1

Керівник к.т.н., доцент  
Науковий ступінь, вчене звання

До захисту допускаю:  
Зав. кафедри КІС, д.т.н, проф.  
Науковий ступінь, вчене звання

26 05 2023 р.

  
Підпис

В.О Корольков  
Ініціали, прізвище

  
Підпис

К.М Березька  
Ініціали, прізвище

  
Підпис

Т.О. Говорущенко  
Ініціали, прізвище

Хмельницький, 2023

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Кафедра КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ ТА ІНФОРМАЦІЙНИХ СИСТЕМ

Освітній рівень МАГІСТР

Галузь знань 12 ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

Спеціальність 123 КОМП'ЮТЕРНА ІНЖЕНЕРІЯ

Освітня програма ОСВІТНЬО-НАУКОВА ПРОГРАМА «КОМП'ЮТЕРНА ІНЖЕНЕРІЯ  
ТА ПРОГРАМУВАННЯ»

ЗАТВЕРДЖУЮ

Зав. кафедри Т.О.Говорущенко

“ 01 ” 09 2022 р.

**ЗАВДАННЯ**

**НА КВАЛІФІКАЦІЙНУ РОБОТУ МАГІСТРА**

Королькову Владиславу Олександровичу

Прізвище, ім'я, по батькові студента

1. Тема проекту (роботи) Метод та засоби ідентифікації об'єктів у тривимірних хмарах технологіями комп'ютерного зору та машинного навчання

Керівник проекту (роботи) Березька К.М., к.т.н., доцент

Прізвище, ім'я, по батькові, науковий ступінь, вчене звання

Затверджена наказом ректора університету від 09.01.2023 р. № 1

2. Строк подання студентом проекту (роботи) на кафедру 19.05.2023 р.

3. Вихідні дані до проекту (роботи) Завдання на дипломне проектування

4. Зміст пояснювальної записки (перелік питань, які потрібно розробити) \_\_\_\_\_

Аналіз предметної області та постановка задачі





— Архітектура дистрибутивної системи

— Штучні імунні системи та нейронні мережі

— Реалізація апаратно-програмного комплексу для виконання навігації і розпізнавання людських образів та результати

5. Перелік графічного матеріалу (із зазначенням обов'язкових креслень) \_\_\_\_\_

6. Консультанти розділів кваліфікаційної роботи магістра

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Нормоконтроль	Лисенко С.М, професор кафедри КПС		
Антиплагіат	Нічепорук А.О, доцент кафедри КПС		

7. Дата видачі завдання « 06 » 09 2022р.

**КАЛЕНДАРНИЙ ПЛАН**

№з/п	Назва етапів (розділів) кваліфікаційної роботи магістра	Термін виконання етапів проекту (роботи)	Примітка
1	Вибір напрямку дослідження та узгодження тематики КвРМ з керівником	05.09.2022	виконано
2	Ознайомлення з предметною областю; формулювання мети та задач дослідження; визначення об'єкта та предмета дослідження	05.10.2022	виконано
3	Робота над розділом 1 – аналіз відомих моделей, методів за темою; постановка задачі	05.11.2022	виконано
4	Робота над розділом 2 – розробка моделей для вирішення поставленої задачі	05.12.2022	виконано
5	Робота над науковою статтею	05.01.2023	виконано
6	Робота над розділом 3 – розробка методів для вирішення поставленої задачі	15.02.2022	виконано
7	Робота над розділом 4 – проектування та розробка ПЗ для вирішення поставленої задачі, експериментальна частина	05.04.2023	виконано
8	Оформлення пояснювальної записки згідно вимог	15.04.2023	виконано
9	Попередній захист ДРМ	18.04.2023	виконано
10	Захист ДРМ на засіданні ЕК	19.05.2023	

Студент

  
Підпис

В.О Корольков  
Ініціали, прізвище

Керівник роботи

  
Підпис

К.М Березька  
Ініціали, прізвище

## РЕФЕРАТ

Тема дипломної роботи: Метод та засоби ідентифікації об'єктів у тривимірних хмарах технологіями комп'ютерного зору та машинного навчання

Автор роботи: В.О Корольков

Керівник роботи: К.М Березька

Пояснювальна записка: 70с, 40 рис, 3 дод, 83 джерел

ПЕРЕЛІК КЛЮЧОВИХ СЛІВ: ідентифікація об'єктів, тривимірні хмари, доповнена реальність, комп'ютерний зір, віртуальна реальність, маркери, машинне навчання.

Об'єктом дослідження є процес ідентифікації об'єктів за допомогою машинного навчання та комп'ютерного зору.

Предметом дослідження підсистема ідентифікації об'єктів у тривимірних хмарах

Метою дипломної роботи є створення системи, що буде виконувати розпізнавання людських обличь, а саме слідкувати за ними у горизонтальному положенні і передавати інформацію на ПК.

Наукова новизна. В роботі використано технології комп'ютерного зору для задач детекції та розпізнавання людських обличь і розроблено метод ідентифікації об'єктів в тривимірному просторі.

Практична новизна. Бібліотека OpenCV забезпечує обробку зображень в режимі реального часу. Також в роботі запропоновано програмно-апаратний зв'язок проаналізованої вхідної інформації з механічною частиною самохідного пристрою, що дозволяє йому самостійно змінювати своє положення, оскільки обробка зображення ведеться на комп'ютері, а управління сервоприводами здійснюється за допомогою контролера

Для розв'язання поставлених задач використовувалися методи:

1. Рухатись у горизонтальній поверхні для слідкування;
2. Збирати і опрацьовувати зображення з веб-камери;
3. Надвати можливість оператору керувати даною схемою, яка в свою чергу має збирати і передавати інформацію на зовнішній носій.

## ЗМІСТ

<b>СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ .....</b>	<b>7</b>
<b>ВСТУП.....</b>	<b>8</b>
<b>1 АНАЛІЗ ВІДОМИХ МЕТОДІВ ТА РІШЕНЬ.....</b>	<b>11</b>
1.1 Аналіз проблемної галузі. ....	11
1.2 Аналіз та класифікація машинного зору. ....	12
1.3 Стан сучасної системи машинного зору.....	15
1.3.1 Визначення поняття, змішана, доповнена та віртуальна реальність, віртуальність континуум .....	24
1.3.2 Супутня робота точок хмари 3D-екземплярів.....	27
1.4 Висновки до розділу 1 .....	28
<b>2. ТЕХНОЛОГІЯ ТА ЗАСОБИ ДОПОВНЕНОЇ РЕАЛЬНОСТІ .....</b>	<b>28</b>
2.1 Технологія доповненої реальності та використання її у повсякденному житті .....	28
2.1.1 Доповнена реальність - Маркерна.....	33
2.1.2 Доповнена реальність - Безмаркерна .....	35
2.1.3 Доповнена реальність - Просторова .....	37
2.2 Інструменти доповненої реальності.....	38
2.3 Висновки до розділу 2 .....	42
<b>3. АПАРАТНА ЧАСТИНА .....</b>	<b>43</b>
3.1 Контролер Arduino Leonardo.....	43
3.2. Камера пристрою .....	48
3.3.Схема підключення пристрою .....	50
3.4 Висновки до розділу 3 .....	52
<b>4. ДЖЕРЕЛА ТА ЕКСПЕРИМЕНТИ РОЗРОБКИ.....</b>	<b>53</b>
4.1 Використання бібліотек машинного зору задля ідентифікації динамічних об'єктів .....	53
4.2 Розробка прототипу системи машинного зору освітнього призначення. ....	55
4.3 Використання Microsoft Azure для розпізнавання динамічних облич .....	64

4.4 Тестування розробленого прототипу .....	74
4.5.Висновок до розділу 4 .....	75
<b>ВИСНОВОК</b> .....	76
<b>ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ</b> .....	78
<b>ДОДАТОК А</b> Клієнтська частина прототипу.....	87
<b>ДОДАТОК Б</b> Копія наукової публікації.....	90
<b>ДОДАТОК В</b> Презентація магістерської роботи .....	

## СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

БПР – блок прийняття рішень

РО – розпізнавання обличь

МН – машинне навчання

КЗ – Комп'ютерний зір

АС – академічне середовище

ЕС – експертна система

БД – база даних

ЕВ – експертний висновок

ММ – математична модель

VR – віртуальна реальність

AR – доповнена реальність

RCNN - Region-Based Convolutional Neural Network

LiDAR - Light Detection and Ranging

TOID - TOpographic IDentifier

AIA – асоціація автоматизованих зображень

ModR – модульована реальність

MfR – Модифікована реальність

DR – Зменшена реальність

## ВСТУП

Система розпізнавання обличчя – це процес ідентифікації або підтвердження особи людини за її обличчям. Він фіксує, аналізує та порівнює шаблони на основі деталей обличчя людини. Процес виявлення обличчя є важливим кроком у виявленні місцезнаходження людей на зображеннях і відео, процес захоплення обличчя перетворює аналогову інформацію у цифрову, на основі рис обличчя людини, процес зіставлення обличчя перевіряє чи належать два обличчя одній людині.

Для вивчення сучасного алгоритму виявлення 3D-об'єктів, основанийого на глибокому вивченні, в моїй роботі я розберу алгоритм знаходження 3D-об'єктів, який використовує багатосаровий перцептрон для вилучення точкових ознак. В моїй роботі запропоновано метод, заснований на точковому RCNN. Алгоритм в три етапи знаходження 3D-об'єктів підвищує точність алгоритму за рахунок з'єднання інформації про зображення. Алгоритм добре поєднує інформацію про зображення трьох етапів, що покращує використання інформації всього алгоритму. Зрівняно з традиційним алгоритмом знаходження 3D-цілей структура алгоритму в моїй роботі більш компактна, що дає змогу ефективніше використовувати інформацію.

Мета і завдання роботи. В моїй роботі досліджені основні етапи створення пристрою, що має розпізнавати обличчя і фокусуватись на них за допомогою червоної рамки. Про кольори рамок та як розуміти чи зфокусувався пристрій, червона рамка – фокус виявлено, жовта рамка, низький відсоток фокусування, зелена рамка – фокус ідеальний.

За допомогою веб камери, встановленої на поворотному механізмі, яка керується операційною системою Windows із встановленим ПЗ OpenCV буде створюватись відео фіксація та ідентифікація об'єктів. В апарат ідентифікації, який підключений до комп'ютера через USB передаються координати, після того, як програма виявить в полі зору веб камери об'єкт, в нашому випадку це обличчя, то обчислюється його центр.

Інструментом дослідження роботи виступає контролер Arduino IDE, яка написана на мові Java. Для завантаження і написання програм на сумісні з Arduino плати використовується якраз ця програма, з середовищем програмування Microsoft Visual C++ і бібліотека для роботи з обробкою зображення та веб пристроями, тобто Open CV.

Наукова новизна. В роботі була використана технологія комп'ютерного зору для задач детекції та розпізнавання людських обличчя, за допомогою пристроїв зчитування інформації з об'єктів завдяки, бібліотеці OpenCV, яка забезпечує обробку зображень в режимі реального часу. Програмно-апаратний зв'язок проаналізованої вхідної інформації, запропонований в моїй роботі, що дозволить йому самостійно змінювати своє положення, обробка зображення ведеться на комп'ютері, а управління за допомогою контролера передається на сервоприводи.

Дана робота створена для виготовлення пристрою, що буде використовуватись в будь-якій сфері життя, починаючи від звичайних смартфонів, де потрібно розпізнавання обличчя аби розблокувати телефон, закінчуючи системою захисту у будинках та приватних територіях для ідентифікації людини за відбитком пальця або райдужної оболонки ока.

По темі дипломної роботи була опублікована стаття у фаховому науковому виданні «ЗБІРНИК НАУКОВИХ ПРАЦЬ за матеріалами XIV» Всеукраїнської науково-практичної конференції «Актуальні проблеми комп'ютерних наук АПКН-2022».

# 1 АНАЛІЗ ВІДОМИХ МЕТОДІВ ТА РІШЕНЬ ДЛЯ ІДЕНТИФІКАЦІЇ ОБ'ЄКТІВ

## 1.1 Аналіз проблемної галузі

Аналіз та виявлення зображень являються одним з основних напрямків досліджень в комп'ютерній області. Виявлення цілей грає важливу роль в традиційних дослідженнях. Наприклад відома бібліотека аналізу зображень з відкритим початковим кодом OpenCV вже підтримує такі алгоритми, як відслідковування обличчя та відстежування виявлення транспортних засобів. Технологія відслідковування об'єктів широко використовується в військовій сфері для виявлення цілей, до прикладу виявлення цілей для виявлення куль, кораблів або радіолокаційних цілей. В цивільній сфері, тісно пов'язаній з життям людей, до прикладу в медицині, використовуються виявлення об'єктів для автоматичної діагностики діабетичних уражень очного дна. Однак традиційні методи потребують ручного проектування функцій для виявлення об'єктів, що займає багато часу та сил. В епоху розвитку глибокого навчання дуже важливо застосовувати глибоке навчання для виявлення об'єктів.

Технологія виявлення об'єктів так само грає важливу роль в області автономного водіння. Автономне водіння завжди було предметом досліджень в області машинобудування. Сучасні системи автономного водіння можна розділити на два типи:

1. Безпілотне.
2. ADAS – удосконалена система допомоги водію.

Ця форма зорієнтована на повністю автономне водіння автомобілю для економії витрат на оплату праці водія, а інша направлена, зниження стресу водія під час водіння та підвищена безпека транспортного засобу.

Обидва використовують різні датчики, встановлені на транспортному засобі, для збору даних та об'єднання картографічних даних для системних розрахунків, аби реалізувати планування маршруту руху та контролювати

транспортний засіб для досягнення завданого положення. Прориви в області штучного інтелекту такі

як машинний зір та глибоке вивчення, мають велике значення для розвитку автономного водіння.

Якщо брати безпілотні автомобілі то в них частіше за все використовується LiDAR та декілька камер, встановлених в різних місцях автомобілю, для збору даних про сприйняття, а потім аналізу та локації зібраних візуальних даних для визначення місцезнаходження таких об'єктів, як смуги руху, транспортні засоби та пішоходи. Останнім часом, з швидким розвитком технологій глибокого вивчення машинного зору та штучного інтелекту, можливості комп'ютерів по аналізу даних зображення значно покращились зрівняно з традиційними методами, тому багато компаній почали розробляти чіпи, спеціально призначені для штучного інтелекту. В інтелектуальних системах водіння основні сенсорні рішення ввімкнуть LiDAR, камери та радар міліметрового діапазону. Перевага LiDAR полягає в 3D-модельованні, широкому діапазоні виявлення та високої точності фокусування.

Через це важливо вивчити швидке виявлення цілей датчиками LiDAR на основі вихідних даних точок хмари. Запропонований в роботі удосконалений метод знаходження цілей точок хмари с прискоренням розрідженої згортки має теоретичне та практичне значення, а також визначену дослідницьку та прикладку цінність. Проведені модельні експерименти, та доведено, що можна досягти високої точності виявлення цілей, що досягає 79.51%. Точність моделі на 0.66% вище базової, що демонструє ефективність класифікатора точок хмари третього рівня та класифікатора зображень.

## 1.2 Аналіз та класифікація машинного зору

Машинний зір, допомагає надійно та послідовно вирішувати складні промислові завдання, за даними асоціації автоматизованих зображень, машинне бачення охоплює всі промислові та не промислові програми, за

допоомгою унікальних ідентифікаторів UUID та глобальний унікальний ідентифікатор GUID, для стандартної унікальної ідентифікації інформації та програм в комп'ютерному середовищі, цифровий ідентифікатор об'єкта DOI. Перші системи унікальної ідентифікації геопросторових об'єктів були розроблені та впроваджені в різних країнах для цілей ведення кадастрів на земельні ділянки та нерухомого майна. Зазвичай ці системи ґрунтуються на кодуванні об'єктів адміністративно-територіального устрою країни, номерів земельних ділянок в межах кадастрових кварталів.

Для унікальної ідентифікації будівель застосовуються різні системи, але найчастіше це системи, ідентифікатор будівель в яких похідним від кадастрового номера ділянки є код їхнього поштового адресу.

Комп'ютерний зір має підкатегорію, яка називається машинним зором. Обидва терміни взаємозамінні. Робота його системи вимагає використання комп'ютера та спеціального програмного забезпечення, але для процесу комп'ютерного зору не потрібна машина. Комп'ютерне бачення може не тільки сканувати цифрові веб-фотографії чи відео, але й аналізувати «зображення» з детекторів руху, інфрачервоних датчиків та інших джерел. Завдяки комп'ютерному зору всі типи машин, керованих комп'ютером, тепер можуть працювати більш інтелектуально та безпечно. Комп'ютерний зір дозволяє роботам працювати краще та різноманітніше, ніж будь-коли раніше, від масивного фабричного та сільськогосподарського обладнання до крихітних дронів, які можуть розпізнавати людей та слідкувати за ними за ними автономно.

Переваги його для цілей перевірки давно визнані у важкій промисловості. Камери та комп'ютери можуть записувати та обробляти зображення значно точніше та швидше, ніж люди. Не може бути помилок у делікатному виробництві виробничих ліній, таких як генерування компонентів для кардіостимуляторів.

Люди-інспектори надто небезпечні для таких ретельних перевірок, і легко зрозуміти, чому, якщо врахувати людські обмеження та можливості комп'ютерного ока та мозку:

1. Щоб просто подивитися на фотографії, надіслані в Snapchat за останню годину, людині знадобиться десять років.

2. Багато сучасних виробничих підприємств не могли б конкурувати, якби вони не включали в свою діяльність комп'ютерну перевірку машин. Виробництво, пакування та доставка їжі є одними з найпоширеніших видів використання. Мною запропонована система на рисунку 1.1 класифікації унікальної ідентифікації об'єктів у геопросторових даних.

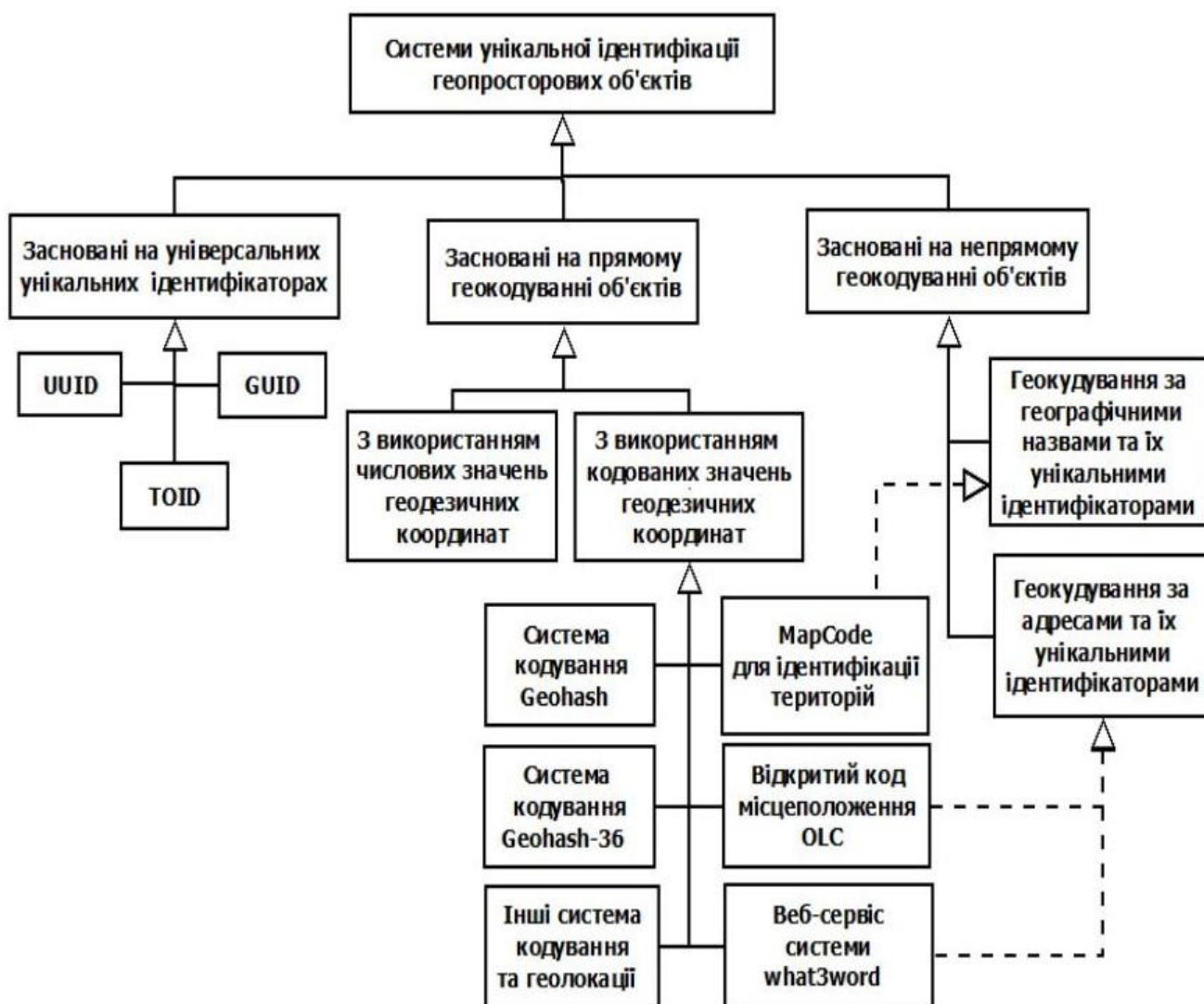


Рисунок 1.1 – Система унікальної ідентифікації геопросторових об'єктів

Унікальні ідентифікатори об'єктів на основі прямих методів геокодування мають семантичне навантаження на систему, оскільки містять безпосередньо в закодованому виді інформацію про місцезнаходження об'єкта на місцевості, а як

правило належать до обов'язкових атрибутів типів самих об'єктів, які можуть мати географічну назву або адресу. Кожному значенню OSAPR відповідає TOID, це забезпечує логічний зв'язок геопросторових об'єктів OS MasterMap з масивом. База даних реєстрій назв та адрес досить широко використовуються в пошукових сервісах геопорталів.

### 1.3 Стан сучасної системи машинного зору

Система машинного зору насамперед дозволяє комп'ютеру розпізнавати та оцінювати зображення. Це схоже на технологію розпізнавання голосу, але замість нього використовуються зображення.

Система машинного зору зазвичай складається з цифрових камер і апаратного та програмного забезпечення для обробки зображень. Камера на передньому кінці знімає зображення з навколишнього середовища або з сфокусованого об'єкта, а потім надсилає їх до системи обробки. Залежно від конструкції або потреби MVS, захоплені зображення зберігаються або обробляються відповідним чином.

Сучасні камери можуть отримувати візуальну інформацію з високою точністю, але фотографувати об'єкти – це не те саме, що бачити ці об'єкти. Комп'ютерний зір дозволяє комп'ютерам отримувати цінну інформацію із зображень і відео (рисунок 1.2).

Зараз ця технологія використовується в поєднанні з машинним навчанням і штучним інтелектом для розпізнавання облич, керування дорожнім рухом, автономної навігації автомобіля, обробки природної мови тощо. Перегляд комп'ютерів — це наступний крок у розвитку штучного інтелекту, який може дати потужний поштовх вашому бізнесу. Сьогодні ми збираємося обговорити, як вона вже змінила індустрії та яку користь ви можете отримати від цієї технології.

Оскільки ціни на камери впали, обчислювальна потужність зросла, а алгоритми вдосконалилися, машинний зір допоміг робототехніці вийти зі статично запрограмованого стану, який Том Хаммел описав.

Завдяки останнім досягненням у глибокому навчанні, які дозволяють роботам фактично аналізувати те, що вони бачать, роботи здатні виконувати завдання, які колись здавалися неможливими або занадто дорогими, наприклад, вибирати певні предмети зі смітника, сказав Хаммел.

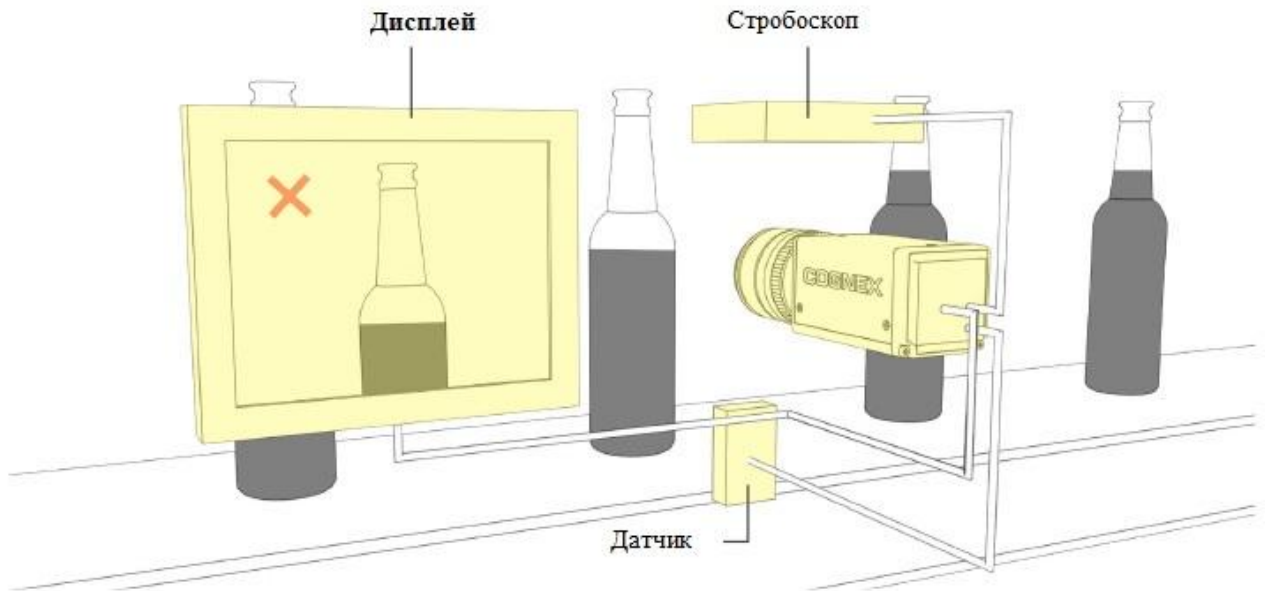


Рисунок 1.2 – Прилад для вимірювання повноти пляшки.

За даними Асоціації автоматизованих зображень (AIA), машинне бачення охоплює всі промислові та непромислові програми, у яких комбінація апаратного та програмного забезпечення забезпечує оперативне керування пристроями під час виконання їхніх функцій на основі захоплення та обробки зображень. Хоча промислове комп'ютерне бачення використовує багато тих самих алгоритмів і підходів, що й академічні/освітні та урядові/військові програми комп'ютерного бачення, обмеження відрізняються.

Як приклад розглянемо систему перевірки рівня заповнення на пивоварному заводі (рисунок 1.3). Кожна пляшка пива проходить через датчик перевірки, який запускає систему зору, яка блимає стробоскопом і робить знімок пляшки. Отримавши зображення та зберігши його в пам'яті, програмне

забезпечення Vision обробляє або аналізує його та видає відповідь «пройшов-не пройшов» на основі рівня заповнення пляшки. Якщо система виявляє неправильно наповнену пляшку (несправність), вона сигналізує перенаправленню відхилити пляшку. Оператор може переглядати забраковані пляшки та статистику поточного процесу на дисплеї.

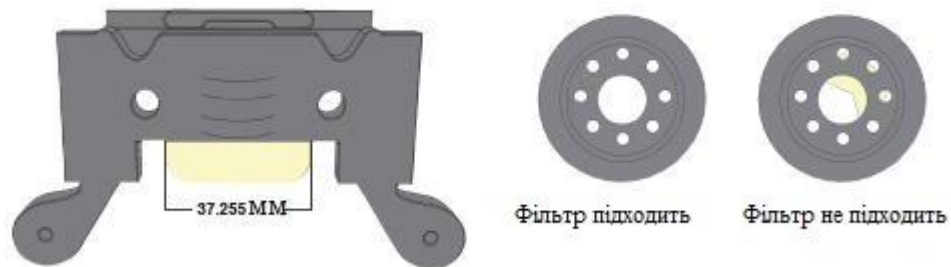


Рисунок 1.3 – Вимірювання рівня олії по фільтру.

Системи машинного бачення також можуть виконувати об'єктивні вимірювання, такі як визначення зазору свічки запалювання або надання інформації про місцезнаходження, яка спрямовує робота для вирівнювання деталей у виробничому процесі. На малюнку 1.6 показано приклади того, як системи машинного зору можуть використовуватися для проходження чи несправності масляних фільтрів (праворуч) і вимірювання ширини центральної пластини на кронштейні (ліворуч).

Після обробки зображень та інших даних ця інформація повертається до робота або інших машин, які працюють поруч із ним. Звідти машини можуть приймати відповідні рішення, чи то зупинка на пішохідному переході, чи вибір потрібного товару для замовлення, що підвищує ефективність і безпеку. Використовуючи штучний інтелект і алгоритми машинного навчання, а також камери та інші засоби зондування, такі як LiDAR, радар і тепловізійні камери, які вимірюють і збирають місцезнаходження, швидкість та інші важливі дані про проїжджу частину, електричні безпілотні автомобілі Zoox здатні бачити навколишній світ і взаємодіяти з ним. Автомобілі, які випробовуються в районі затоки Сан-Франциско, Лас-Вегасі та Сіетлі, можуть навіть передбачити, що

можуть або не можуть робити автомобілі, вантажівки, велосипедисти та пішоходи поблизу, наприклад, якщо людина виходить на проїжджу частину.



Рисунок 1.6-Класифікація методів машинного навчання

Існує чотири основні категорії алгоритмів машинного навчання: Контрольоване, не контрольоване, напівконтрольоване, та навчання з підкріпленням. Не зважаючи на те, що класифікація та регресія належать до категорії навчання під наглядом, вони не є одним і тим же.

Задача прогнозування, це класифікація, коли цільова змінна являється дискретною. Додаток, це ідентифікація основного настрою фрагменту тексту.

Задача прогнозування представляє собою регресію, коли цільова змінна безперервна. До прикладу може бути прогнозування заробітної плати людини з урахуванням його освіти, попереднього місця роботи, географічного положення та рівня досвіду роботи.

В машинному навчанні також є чотири основні задачі класифікації, такі як, бінарна, багатокласова, багато рівнева та не збалансована класифікація.

Бінарна класифікація являє собою класифікацію вхідних даних по двом взаємовиключним категоріям.

Багатокласова класифікація має принаймні дві взаємовиключні мітки класів, ціль котрих, передбачити якому класу належить даний вхідний приклад.

Багаторівнева класифікація з декількома мітками передбачає 0 або більше класів для кожного прикладу.

Використання звичайних прогностичних моделей, таких як дерева рішень, логістична регресія тощо, не може бути ефективним при роботі з незбалансованим набором даних, оскільки вони можуть бути упередженими в бік прогнозування класу з найбільшою кількістю спостережень і розглядати ті з меншими числами як шум.

Також хотілось би розглянути дерево рішень, це контрольований алгоритм навчання, який ідеально підходить для проблем класифікації, оскільки він здатний упорядковувати класи на точному рівні. Він працює як блок-схема, розділяючи точки даних на дві подібні категорії одночасно від «стовбура дерева» до «гілок» і «листя», де категорії стають більш кінцево подібними. Це створює категорії в категоріях, що дозволяє органічну класифікацію з обмеженим наглядом людини.

З іншого боку, ліниві учні або ті, хто навчається на основі екземплярів, не створюють жодної моделі відразу з навчальних даних, і саме звідси походить ледачий аспект. Вони просто запам'ятовують навчальні дані, і кожного разу, коли виникає потреба зробити прогноз, вони шукають найближчого сусіда з усіх навчальних даних, що робить їх дуже повільними під час прогнозування.

Давайте розглянемо наступний сценарій класифікації з 3 класів, де навчальні дані містять: 60% вантажівок, 25% літаків і 15% човнів. В одному аспекті втілення способу оцінки пози об'єкта з використанням машинного бачення може бути підсумовано таким, що включає ідентифікацію області об'єкта зображення, на якому виконується оцінка пози на основі відповідності між щонайменше частиною представлення об'єкта в області об'єкта зображення і, щонайменше, відповідна одна з безлічі еталонних двовимірних моделей об'єкта, причому область об'єкта є частиною зображення, яка містить представлення,

щонайменше, частини об'єкта; і визначення тривимірного положення об'єкта на основі щонайменше однієї з безлічі еталонних тривимірних моделей об'єкта та тривимірного представлення області об'єкта під час виконання, де зв'язок «точка-точка» між трьома еталонними -вимірні моделі об'єкта та тривимірне представлення області об'єкта під час виконання не обов'язково є раніше відомими.

У додатковому аспекті варіант здійснення системи для виконання тривимірної оцінки пози може бути підсумований як такий, що включає щонайменше один датчик; хоча б один процесор.

Спосіб оцінки пози об'єкта з використанням машинного бачення, що включає ідентифікація об'єктної області зображення, на якій виконується оцінка пози на основі відповідності між щонайменше частиною представлення об'єкта на зображенні та щонайменше відповідним одним із множини опорних двовимірних моделей об'єкта область об'єкта є частиною зображення, яка містить представлення щонайменше частини об'єкта; і визначення тривимірного положення об'єкта на основі щонайменше однієї з безлічі еталонних тривимірних моделей об'єкта та тривимірного представлення області об'єкта під час виконання, де зв'язок «точка-точка» між еталонними тривимірними моделями розмірні моделі об'єкта та тривимірне представлення області об'єкта під час виконання не обов'язково є раніше відомими.

Енергозчитуваний комп'ютером носій, який зберігає інструкції для того, щоб змусити комп'ютер виконати оцінку пози об'єкта за допомогою машинного бачення, шляхом ідентифікація об'єктної області зображення на основі відповідності між принаймні частиною представлення об'єкта в об'єктній області зображення та принаймні відповідною однією з безлічі опорних двовимірних моделей об'єкта, об'єкта область, яка є частиною зображення, яка містить представлення щонайменше частини об'єкта; і визначення тривимірного положення об'єкта на основі щонайменше однієї з безлічі еталонних тривимірних моделей об'єкта та тривимірного представлення області об'єкта під час виконання, де зв'язок «точка-точка» між еталонними тривимірними моделями розмірні

моделі об'єкта та тривимірне представлення області об'єкта під час виконання не обов'язково є раніше відомими. Система для виконання тривимірної оцінки пози, яка містить:

1. Хоча б один датчик.
2. Хоча б один процесор і щонайменше одна виконувана інструкція процесора, що зберігає пам'ять, яка змушує щонайменше один процесор сегментувати зображення.

Він проводить аналіз даних знову і знову, доки не розпізнає відмінності й остаточно не дійде до розпізнавання зображення. До прикладу, аби комп'ютер навчився розпізнавати автомобільні шини, йому потрібно надати дуже велику кількість зображень шин і предметів, пов'язаних із шинами, щоб дізнатися про відмінності та розпізнати шину, особливо без дефектів використовує, для перегляду, пошуку та отримання зображень із великих сховищ даних на основі вмісту зображень, а не пов'язаних із ними тегів метаданих. Це завдання може включати автоматичне анотування зображення, яке замінює ручне тегування зображення. Незважаючи на те, що стає легше отримати ресурси для розробки програм комп'ютерного зору, важливе запитання, на яке потрібно відповісти на ранній стадії: що саме будуть робити ці програми? Розуміння та визначення конкретних завдань комп'ютерного зору може зосередити та перевірити проекти та програми та полегшити початок роботи. Застосування технологій машинного зору в автоматизації та промислових лініях добре відомо. Його системи дозволяють системі мінімізувати час, за який люди беруть участь у кількох завданнях. Це може статися під час перевірки або виробництва. Правильне застосування його систем у кінцевій установці підвищує продуктивність і покращує коректність результатів роботи шляхом виявлення помилок до отримання клієнта. Оскільки він може бути з'єднаний з іншими системами, такими як конвеєри, його можна використовувати в потенційно небезпечних або чистих середовищах, де людина може бути забруднена або травмована. Системи бачення підвищують якість продукції, зменшуючи людські помилки та

забезпечуючи перевірку якості всіх товарів, що проходять через лінію (рисунок 1.7).

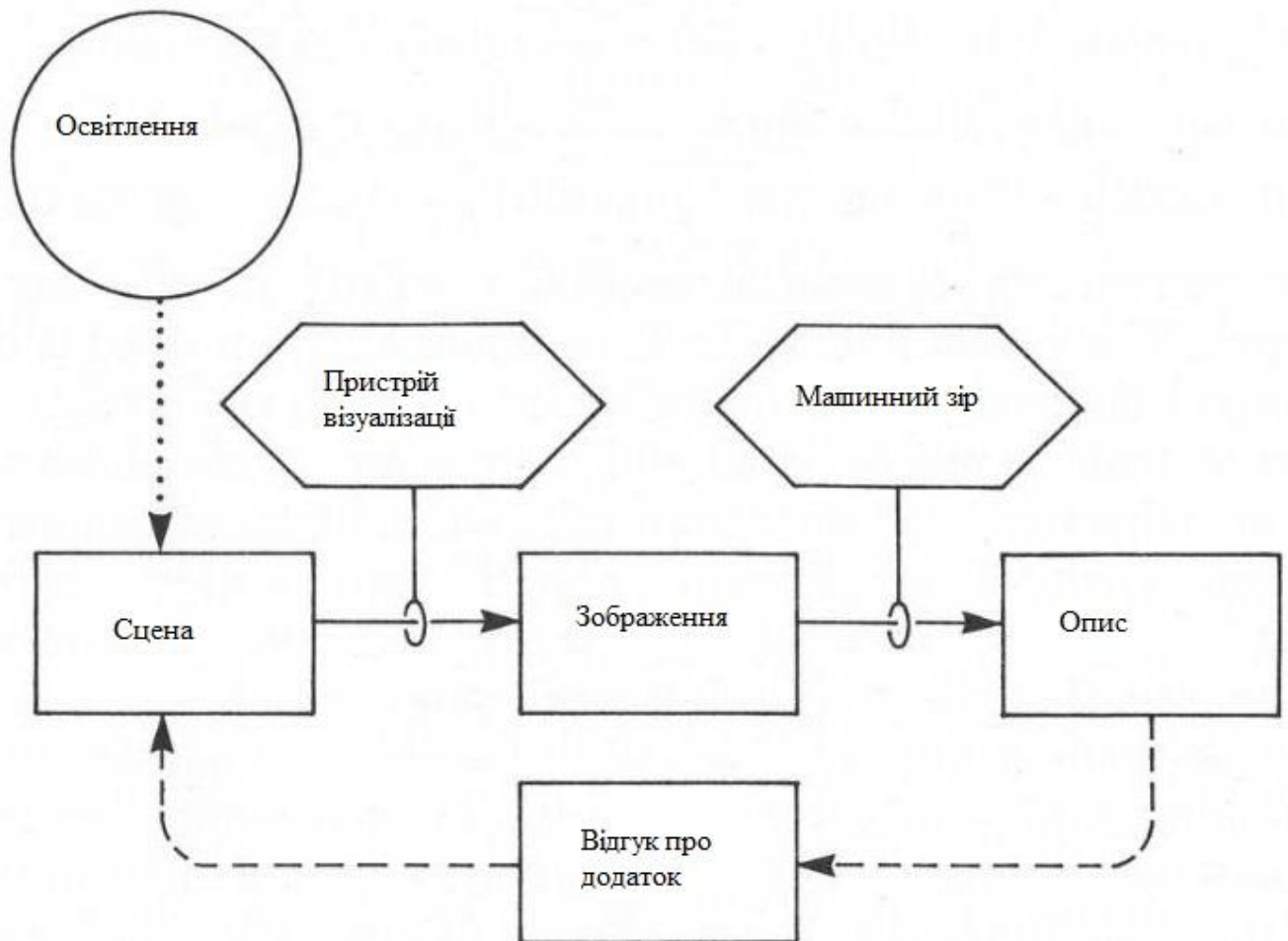


Рисунок 1.7 – Схема принципу роботи машинного зору

Комп'ютерний зір використовується в різних галузях промисловості, від енергетики та комунальних послуг до виробництва та автомобілебудування – і ринок продовжує зростати. Очікується, що до 2022 року він досягне 48,6 млрд доларів США. Комп'ютерний зір працює майже так само, як і людський, за винятком того, що люди мають перевагу. Людський зір має перевагу тривалості контексту, щоб навчитися розрізняти об'єкти, як далеко вони знаходяться, чи рухаються вони та чи є щось не так на зображенні. Комп'ютерний зір дозволяє комп'ютерам мислити тільки якщо ІІ дає на це дозвіл, також вони можуть спостерігати, бачити та розуміти об'єкти.

Оскільки система, навчена перевіряти продукти або спостерігати за виробничим активом, може аналізувати тисячі продуктів або процесів за хвилину, помічаючи непомітні дефекти чи проблеми, вона може швидко перевершити людські можливості. Подібно до того, як людина створює зображення на відстані, CNN спочатку розрізняє жорсткі грані та прості форми, а потім заповнює інформацію, коли відбувається ітерація прогнозів програми. CNN використовується для розуміння окремих зображень. Подібним чином використовується повторювальна нейронна мережа, аби допомогти комп'ютерам зрозуміти, як зображення в серії кадрів пов'язані одне з одним.

1.3.1 Визначення поняття, доповнена, змішана та віртуальна реальність, віртуальність континуум.

Безперервну шкалу в діапазоні між повністю віртуальним, досконало реальною та реальністю представляє віртуальність континуум. Континуум реальності, це віртуальність, яка включає в себе всі можливі композиції та варіації реальних та віртуальних об'єктів.

Це було описано в вигляді концепції нових медіа та інформатики, але насправді, також цей опис можливо розглянути в вигляді питання з антропології. Пол Мілграм вперше представив цю концепцію, в якій область між двомісними краями де змішане та реальне та віртуальне, називається змішаною реальністю, у свою чергу, він як є з повної реальності, де віртуальне поповнення реального, так і з повної віртуальності, де реальне поповнення віртуального. На рисунку 1.8 можна побачити розширений континуум до двовимірної площини віртуальності і медіальності.

1. V – включає в себе реальність, доповнену реальністю.
2. R – позначає незмінну реальність.

Також, таксономія включає в себе модифікацію віртуальності або реальності, або також будь-яку їх комбінацію, змішана та доповнена реальність використовуються інколи як синоніми, правіше і вище у нас віртуальні світи, які

реагують на змінену сильно версію реальності. Так, як саме ця концепція у багатьох аспектах пов'язана з тим, яким чином люди продовжують змінювати спосіб спілкування, спосіб формування ідентичності і спосіб їх взаємодії зі світом і в середині нього, тому його більш точно можна описати як предмет антропології. Зміни до ставлення до технологій та засобів масової інформації, збільшили їх доступ до зміни і розширення способу їх використання, таких як до прикладу:

1. SMS – один до одного.
2. Електронна пошта – один до багатьох.
3. Чати – багатьох до багатьох.

Відмінності між реальністю і віртуальністю стали не чітко видними, так як люди включаються в значній мірі у віртуальність в свої особисті щоденні реальності, використовуючи предмети, використання яких зробило колись чіткі відмінності, між реальністю і віртуальністю, онлайн та оффлайн застарілими, оскільки люди включаються в значній мірі покладаються в свої особисті щоденні реальності.

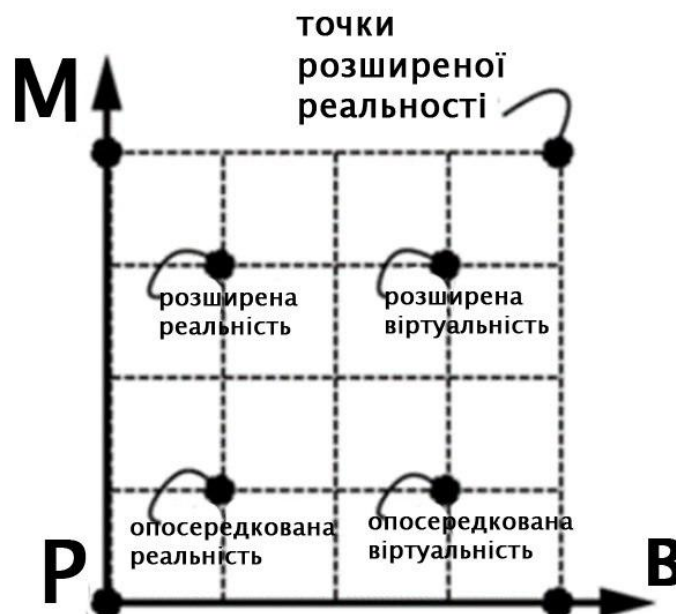


Рисунок 1.8 – Опосередкована реальність, в якій позується чотири точки на осях

Термін «опосередкована реальність» був введений піонером носимих комп'ютерів Стівом Манном, він стосується здатності додавати, віднімати інформацію або іншим чином маніпулювати своїм сприйняттям реальності за допомогою комп'ютерних пристроїв, таких як переносні (наприклад, гарнітури з ефектом занурення) або портативні пристроїв (наприклад, смартфонів). Посередницька реальність — це надмножина віртуальної, доповненої та змішаної реальності, а також доповненої віртуальності (AV), зменшеної реальності (DR), модульованої реальності (ModR) і модифікованої реальності (MfR).

Посередницькі реалії зазвичай передбачають перехід між «опосередкуванням» сприйняття реальності користувачем, додаванням, видаленням, розширенням, покращенням або зміною їх сприйняття її, як і в багатьох технологіях, є позитивні та негативні сторони. Комп'ютерно-опосередкована реальність використовується для покращення візуального сприйняття як допоміжний засіб для людей із вадами зору, як це стало очевидним в останні роки, опосередкована реальність, як-от віртуальна, доповнена та змішана реальності, продовжує відігравати важливу роль у багатьох галузях, таких як ігри, медицина, інженерія тощо. Існує багато аспектів структури опосередкованої реальності, і ми досліджуватимемо деякі з них разом, щоб краще зрозуміти потенційні варіанти та комбінації реальних і віртуальних об'єктів і речей між ними.

Віртуальна реальність, або VR, — це використання комп'ютерних технологій для створення змодельованого середовища, яке можна досліджувати на 360 градусів. На відміну від традиційних інтерфейсів, VR розміщує користувача у віртуальному середовищі, щоб забезпечити захоплюючий досвід. Щоб створити таке відчуття присутності, використовуються VR-гарнітури, такі як ті, які можна взяти в бібліотеці OISE. Інші помітні приклади гарнітур VR включають Oculus Rift, Samsung Gear VR, HTC Vive, Google Daydream View або Google Cardboard. Ці гарнітури усувають бачення реального світу та

забезпечують відео для кожного ока, що забезпечує глибину зору. Потім ця технологія підтримується відстеженням голови та тіла, щоб зв'язати віртуальний світ із тим, що бачить користувач. Чи знаєте ви, що можете позичити окуляри віртуальної реальності в бібліотеці? Все, що вам потрібно, це мобільний телефон, щоб почати! Ознайомтеся з нашим розділом « Освітні програми », щоб отримати рекомендації щодо програм для використання! У своїй найпростішій формі 360-градусний не є VR. Однак 360-градусний досвід використовується у VR, де користувачі переглядають 360-градусний контент у гарнітурі VR. VR використовує датчики для відстеження рухів голови користувача, щоб створити ілюзію того, що він/вона є альтернативним світом, де немає видимості реального світу. З 360-градусним контентом користувачі не занурюються повністю, тому що вони можуть дивитися вгору, вниз і навколо, але не рухатися вперед.

Змішана реальність має багато практичних застосувань, один з них це інтерактивне управління вмістом продуктів – IPCM – використовується що б дозволити клієнтам переглядати фарби, меблі та інші комплектуючі та розхідні матеріали перед самою покупкою. Інші методи використання MR включають в себе художню виразність, дослідження в галузі штучного інтелекту, оцифровану людську свідомість, військову та медичун підготовку. Також одна із підгалузей як MR, використовується віртуальністю AV. Ця технологія інтегрує об'єкти у віртуальний простір, такий к зелений екран та фони відео чату.

### 1.3.2 Супутня робота точок хмари 3D-екземплярів

Я звернув увагу, що виявлення об'єктів на основі точок хмари тісно пов'язане з алгоритмами сегментації 3D-екземплярів на основі точок хмари. Для точок хмари в 3D-сценах алгоритм сегментації в 3D-екземплярів повинен присвоювати кожній хмарі точки мітку класу та окремі мітки екземплярів, а також відрізняти різні екземпляри одного і того ж самого класу.

Декілька підходів до сегментації 3D-екземплярів базуються на обмежувальних прямокутниках виявлення 3D та додаткової гілки маски для

прогнозування масок для об'єктів в середині боксів. Завдяки появі PointNet надається нова можливість для прямої обробки власних точок хмари, які можуть напряму вивчати ознаки та ідентифікувати об'єкти в хмарах. Підраховано, що експерименти показують, що 3D-сегментація в точці хмари може зробити 3D-локацію більш чіткою та точною. Етап вилучення приймає форму 3D-згортки, а далі використовує 2D-згортку, стискаючи вимір висоти точки хмари.

PointRCNN пропонує новий спосіб створення кадрів-кандидатів 3D-пропозицій з точок хмари. Також пропонує метод прийняття кожної точки переднього плану в якості центру об'єкту, використовуючи сферичну прив'язку для створення пропозицій кандидатів тривимірних областей, ну і шляхом агрегування розріджених точок хмари в більш компактний для прогнозування кадру подальших точних налаштувань.

#### 1.4 Висновки до розділу 1

1. У даному розділі я дійшов до висновку, що реальні програми демонструють, на скільки важливий комп'ютерний зір для діяльності в бізнесі, розвагах, транспорті, охороні здоров'я та повсякденному житті, завдяки експериментам проведеним в даному розділі. Ці дані можуть відігравати важливу роль у діяльності в різних галузях, але сьогодні вони не використовуються. Ця інформація створює тестовий стенд для навчання додатків комп'ютерного зору та панель запуску, щоб вони стали частиною ряду людських дій.

2. Технології 3D-об'єктів та 3D-сегментації пішла далеко вперед, у сучасному світі важко уявити людину, яка не ознайоmlена з поняттям «3D». Віртуальна реальність у даний час замінила та допомагає людству у вирішенні багатьох складних питань, використовується як розваги для населення і також, що не мало важливо, медицина, військова та просто цивільна структури, використовують технології 3D-сегментації у своїх цілях.

## 2. ТЕХНОЛОГІЯ ТА ЗАСОБИ ДОПОВНЕНОЇ РЕАЛЬНОСТІ

### 2.1 Технологія доповненої реальності та використання її у повсякденному житті

Доповнена реальність відкриває нові перспективи для роботи, навчання, ігор та взаємодій з оточуючим світом, це вдосконалена версія нашого реального-фізичного світу, яка досягається за допомогою цифрових візуальних елементів, звуку чи інших сенсорних стимулів і доставляється за допомогою технологій. Це зростаюча тенденція серед компаній, які займаються до прикладу бізнес-додатками. На тлі зростання збору та аналізу даних, одна з головних цілей доповненої реальності полягає в тому, щоб висвітлити особливості фізичного світу, покращити розуміння цих функцій та отримати розумну та доступну інформацію, яку можна застосувати до програм реального світу. Такі важливі дані можуть допомогти компаніям прийняти рішення та отримати уявлення про споживчі звички зокрема.

Технології доповненої реальності:

1. Маркерна.
2. Безмаркерна.
3. Просторова.
4. На основі накладання.

Два інженери Boeing, Томас Коделл і Девід Мізелл, ввели термін «доповнена реальність» у 1992 році. У той час вони працювали над простою гарнітурою, яка допоможе інженерам літаків працювати зі складними схемами електропроводки. Їхньою метою було скоротити витрати та підвищити ефективність процесів, пов'язаних із виробництвом авіакомпаній.

Технологічний прогрес наприкінці 20 століття та в останні два десятиліття розширив горизонт доповненої реальності. Це значною мірою завдяки інвестиціям у технологію, які лише у 2019 році склали 4,1 мільярда доларів. Те, як це працює, також стало складнішим. Нижче я розгляну, як працює AR і які технології потрібні для його функціонування.

Є три технології та програми доповненої реальності, які дозволяють працювати AR. Це одночасне визначення місцезнаходження та картографування (SLAM), відстеження глибини, а також обробка та проектування зображень. У сукупності вони роблять AR реальністю.

SLAM — це технологія, яка дозволяє пристрою накладати цифрові зображення на реальний світ. Цьому допомагають датчики локалізації, щоб отримати повне відчуття фізичної обстановки. З іншого боку, відстеження глибини вимірює відстань до об'єкта або поверхні від камери пристрою AR. Нарешті, додаток або програмне забезпечення AR обробляє зображення, щоб правильно визначити розміри та розташування. Коли ці програми працюють одночасно, вони проектують цифровий об'єкт у реальний простір.

SLAM, обробка зображень і відстеження глибини – це три технології, які забезпечують доповнену реальність. Однак існують інші технології доповненої реальності, які об'єднують усе це разом. Кілька прикладів: датчики, камери, акселерометри, цифровий компас, дисплеї, навігаційні пристрої та гіроскопи. Крім того, для безперебійної роботи їх потрібно поєднати з відповідним процесором.

Пристрій із камерою, як-от мобільний телефон або пара розумних окулярів, необхідний для цього. Крім того, для обробки зображень потрібне програмне забезпечення AR — саме цей інструмент розпізнає об'єкти за допомогою технології комп'ютерного зору. У свою чергу, ця технологія читає або аналізує відеопотік.

Для повноти пристрій отримує інформацію про об'єкт із хмари. Він забезпечує 3D-досвід, який рухається разом з користувачем, чий розмір і орієнтація адаптуються до налаштувань пристрою.

Маркери — це унікальні візерунки, які камери можуть вловити з будь-якого середовища. Користувачам знадобиться додаток, щоб сканувати їх і запускати розширений досвід, який з'являтиметься на пристрої людини. Залежно від дизайну, маркер або камеру можна переміщати для різних ефектів.

Хотілось би розглянути найпопулярніші платформи AV / VR для розробки в США на рисунку 2.1.

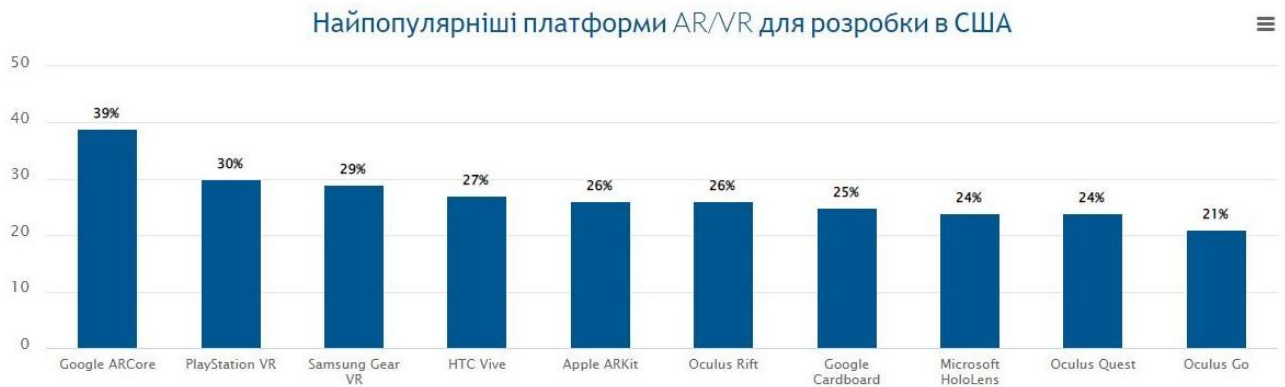


Рисунок 2.1 – Найпопулярніші платформи для розробки в США

Одним із найпоширеніших випадків використання доповненої реальності є ігри. Але ця технологія не просто надає інший аспект іграм. Як видно зі збільшення корпоративних і промислових інвестицій у програмне забезпечення доповненої реальності, воно також широко використовується в інших сценаріях, окрім ігор. Нижче наведено програми та обмеження технологій доповненої реальності.

Медичні установи використовують AR для навчання лікарів проведенню операцій, особливо складніших. Наприклад, у Клівлендській клініці університету Кейс Вестерн Резерв студенти вивчають анатомію за допомогою HoloLens. Студенти повідомили, що 15 хвилин роботи з пристроєм могли б скоротити десятки годин навчання в традиційних лабораторіях.

Також хотілось би сказати про інвестиції в доповнену та віртуальну реальність, до 2024 року тільки в одне навчання хочуть інвестувати 4,1 мільярда доларів, в промислове обслуговування також 4,1 мільярда доларів, а в роздрібну демонстрацію 2,7 мільярди доларів, ну і звісно найбільша інвестиція у розмірі 17,6 мільярдів доларів піде у розвиток AR/VR ігор та VR відео/перегляд функцій

рисунок 2.2.

Існує низка технологічних компаній, які проводять дослідження та розробку програмного забезпечення та продуктів AR. Часто багато з них пов'язані з такими

випадками використання, як наведені вище. Згідно з оглядом літератури, інформатика є провідною категорією, де проводяться дослідження AR. Далі йде сфера медицини, зокрема реабілітації та клінічної неврології. Інші категорії, де AR приділили велику увагу, це інженерія, освіта та освітні дослідження, системи автоматизації та керування тощо.

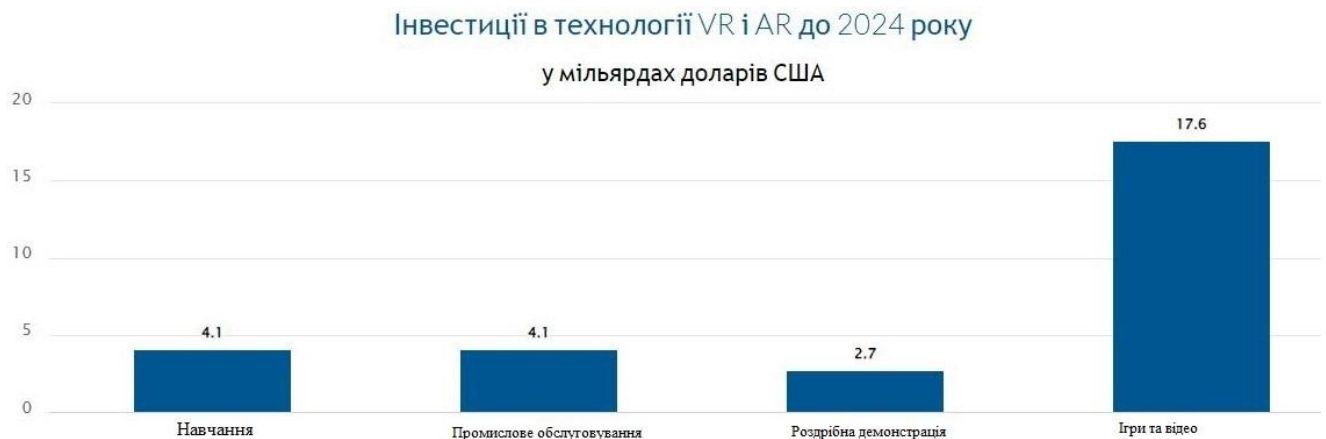


Рисунок 2.2 – Інвестиції в технології VR/AR до 2024 року

Доповнена реальність – це магія. Але в технології все ще є обмеження. Засновники Laan Labs зазначили, що дизайн і розробка обмежені технологіями, доступними сьогодні. Це означає, що самі споживачі також повинні встановлювати реалістичні очікування. Крім того, доповнена реальність все ще певною мірою обмежена, особливо на мобільних пристроях. У статті DAQRI говорилося про камери глибини, які використовуються в мобільній AR. Споживачі можуть бути більше знайомі з Kinect від Xbox 360 — ігровий аксесуар використовує камеру глибини для захоплення рухів користувача. Але ця технологія повернулася в суспільну свідомість завдяки iPhone X від Apple, який інтегрує базову форму.

Однак існують складні вимоги до мобільних камер глибини AR. За словами виробника розумних окулярів, по-перше, датчики мають бути достатньо малими, щоб відповідати обмеженням розміру гарнітур. По-друге, він повинен споживати якомога менше електроенергії, щоб запобігти перегріву. По-третє, він не повинен

бути вимогливим до обробки виходу датчика. Тому що якщо це так, то це буде споживати більше енергії. І, нарешті, вони потребують належного калібрування. Оскільки мобільні камери з часом зазнають навантажень і зношуються, їх заводське калібрування може бути нетривалим.

Але досягнення в інших сферах можуть незабаром запропонувати вирішення таких проблем і просунути технології доповненої реальності на більшу висоту. AR тепер є звичним явищем, і мільйони пристроїв готові його підтримувати. Але все ще є проблеми, які потрібно подолати, щоб створити майже ідеальний досвід AR. По-перше, це питання глибинного зондування. Щоб бути оптимальними, камери глибини мають відповідати трьом важливим критеріям. Також є питання вартості. Незважаючи на те, що доповнена реальність стала доступнішою для звичайного споживача, пристрої з підтримкою доповненої реальності все ще мають високу ціну в інших секторах. Таким чином, медичні установи, будівельні фірми, підприємства та інші організації зможуть дозволити собі широке впровадження ще довго.

І хоча технології доповненої реальності сьогодні є благом, у них є мінуси, такі як, вторгнення в реальний світ цілком ймовірно, що користувачі AR-пристроїв або програм можуть вторгнутися в простір приватних осіб.

Це сталося з Pokemon Go, яка є однією з найкращих ігор у світі, згідно зі статистикою додатку, ризикова поведінка, знову ж таки взяти до прикладу додаток Pokemon Go, двоє чоловіків у Карпатах впали зі скелі під час гри, хоча це не зовсім провина розробника ігор. Це все дуже реальний недолік, оскільки AR використовує середовище, що б надати розширений досвід.

### 2.1.1 Доповнена реальність - Маркерна

Доповнена реальність на основі маркерів, також відома як Image Recognition AR, використовує об'єкт або контрольний маркер як орієнтир для визначення положення або орієнтації камери.

Мобільне сканування спричинить додатковий вміст, підготовлений заздалегідь аби з'явитись у верхній частині маркера (рисунок 2.3).

Використовуючи дві камери, інфрачервона та видима, які розташовані з обох боків холодного дзеркала так, що б їх оптичні центри збігалися один з одним. Ми відстежуємо невидимі маркери за допомогою ІЧ-камери та візуалізуємо у вигляді камери. Для забезпечення надійної роботи системи в умовах захаращеннявикористовується додатковий метод.

Експериментальні результати наведені для демонстрації життєздатності запропонованої системи. У якості застосування запропонованої системи невидимий маркер може діяти як тег ідентифікації та геометрії на основі бачення, що може значно розширити функціональність RFID.

Невидимий тег такий самий як RFID, оскільки він не сприймається, але більш потужний, оскільки інформацію про тег можна представити користувачеві шляхом прямої проекції за допомогою мобільного проектора або шляхом візуалізації AR на екрані мобільного ПК.

Щоб створювати цікавий контент доповненої реальності, ви повинні розуміти, як функціонує технологія, і знати, як плавно інтегрувати реальні ситуації з цифровими шарами. Ось короткий огляд усього, що вам знадобиться для створення власного досвіду доповненої реальності, перш ніж читати деталі та характеристики:

3D-об'єкт або фільм, який відобразатиметься поверх вибраного зображення, статичне тригерне зображення (також відоме як маркер AR) і програмне забезпечення для об'єднання двох фрагментів матеріалу, наприклад надмірне самообслуговування. Щоб отримати доступ до вмісту AR, відскануйте маркер за допомогою мобільного пристрою з відповідним додатком. Немає різниці, чи маєте ви вміст, який потрібно створити, чи хочете об'єднати. Нижче я розгляну деталі та рекомендовані методи переходу для обох випадків. Необхідно спеціально розробити зображення, яке ви використовуєте як маркер AR для свого проекту. Зображення повинно бути оригінальним; ви повинні переконатися. Виберіть дизайн, який тільки ви можете переглядати онлайн,

фотографію, яку ви зробили, або будь-що з бібліотеки вашої компанії. Використання стокових фотографій або зображень Google не є розумною ідеєю, оскільки хтось інший може їх уже використовувати або буде робити це в майбутньому, що може призвести до появи різної інформації на одному маркері. Іншими словами, ви не є власником цього матеріалу.

Використовуйте якомога більше графічних компонентів і контрастів у своїй тригерній фотографії, щоб машинне зір виявило маркер. Ваше зображення маркера – це щось інше, ніж те, щодо чого ви повинні бути мінімалістичними. Система розпізнавання зображень процвітає на різних формах, тінях тощо.

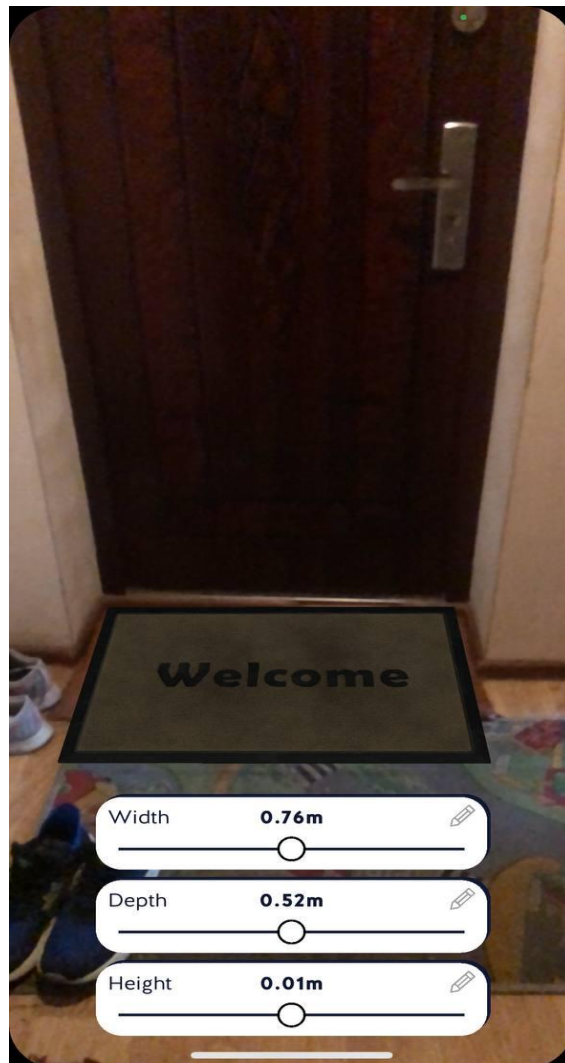


Рисунок 2.3 – Доповнена реальність на основі маркера

Він може стати нестабільним, якщо AR-матеріал поміщається поверх маркерів із занадто великим текстом. Ми пропонуємо використовувати область із графічним елементом як маркер, якщо ви хочете додати життя основному текстовому веб-сайту. Нижче наведено витяг із проспекту Відземського університету прикладних наук. Ми не могли використати праву сторону сторінки як маркер, але ми могли добре використати ліве зображення. Журнали та газети часто використовують цю стратегію, коли оживляють фотографії або об'єкти реального світу на текстовій сторінці.

### 2.1.2 Доповнена реальність - Безмаркерна

Безмаркерна доповнена реальність не використовує маркер. Замість маркера ви запускаєте досвід, вибравши функцію AR у додатку чи на веб-сайті або натиснувши посилання. Хоча доповнена реальність без маркерів поширена в програмах, вона також добре працює з WebAR.

Безмаркерна доповнена реальність сканує реальне оточення та розміщує цифрові елементи на впізнаваному об'єкті, як-от плоскій поверхні. Отже, замість прив'язки до маркера цифрові елементи розміщуються на основі геометрії. Безмаркерна доповнена реальність дуже популярна в іграх, як-от Pokémon Go, де персонажі можуть пересуватися навколо. Він також часто використовується для подій у прямому ефірі та розміщення віртуальних продуктів.

Технологія поєднує в собі програмне забезпечення, аудіо та відео графіку з камерами смартфона або гарнітури, акселерометром, гіроскопом, гаптичними датчиками та службами локації для реєстрації. Нижче на рисунку 2.4[1] наведено приклад, що дає нам змогу зрозуміти, що AR без маркерів не потребує ніяких фізичних маркерів, що б розмістити об'єкт у реальному просторі.

Безмаркерна доповнена реальність використовується для позначення доповненої реальності, якій не потрібні попередні знання про середовище користувача, що б накласти 3D-вміст на сцену та утримувати його у фіксованій

точці простору. Завдяки ARKit безмаркерна AR доступна на сотнях мільйонів смартфонів і планшетів. Роздрібні продавці, освітяни та виробники ігор, це лише деякі групи, які зараз використовують AR без маркерів для створення неймовірних інструментів і контенту. Це кращий метод для розпізнавання зображень для програм AR.

AR без маркерів не потребує ефективної системи відстеження об'єктів і забезпечує стабільну роботу AR. Завдяки численним удосконаленням камер, алгоритмів, датчиків і процесорів мобільних телефонів і планшетів тепер можна створити багатий досвід AR без маркерів.

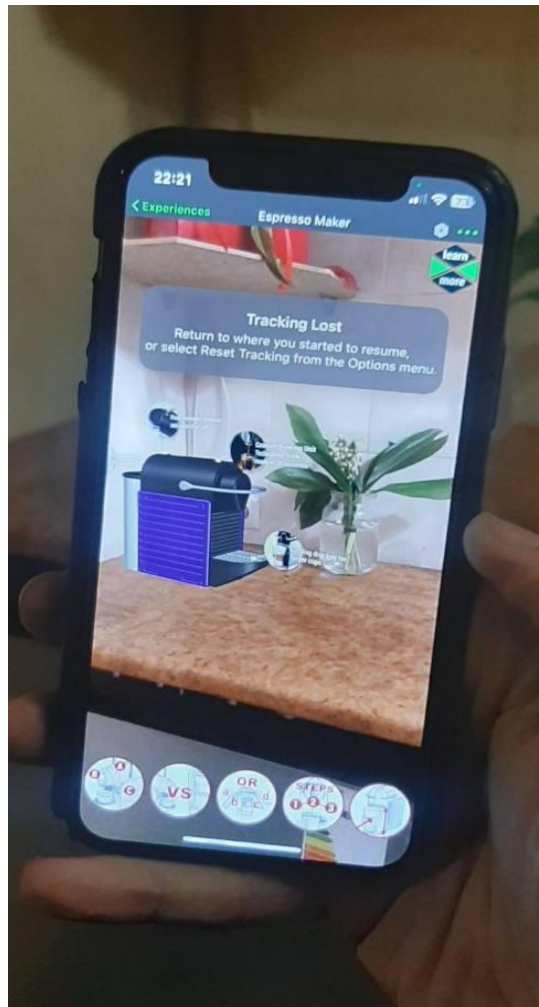


Рисунок 2.4 – безмаркерна доповнена реальність [1]

### 2.1.3 Доповнена реальність - Просторова

Просторова доповнена реальність, це особливий тип технології доповненої реальності, в якій поєднання віртуальних і реальних об'єктів створюється шляхом проектування віртуальних зображень на реальні об'єкти за допомогою проєкційного відображення. Тому монітори, накладні дисплеї або портативні пристрої зазвичай не використовуються в цьому типі AR.

Просторова доповнена реальність (SAR) дозволяє користувачам співпрацювати, не потребуючи прозорих екранів або дисплеїв на голові. Ми досліджуємо природні особисті інтерфейси за допомогою SAR. Просторова доповнена реальність на людині (SARP) використовує самостійні психологічні ефекти, такі як самореферентне кодування (SRE) і володіння, переплітаючи розширені взаємодії тіла з самим собою. Програми, засновані на SARP, можуть надати потужні інструменти в освіті, обізнаності про здоров'я та медичній візуалізації. Мета цієї статті полягає в тому, щоб вивчити переваги та обмеження формування власності та SRE за допомогою техніки SARP. Ми впроваджуємо апаратну платформу, яка забезпечує просторове доповнене ігрове середовище, що дозволяє експериментувати з SARP. Ми тестуємо навчальну гру STEM під назвою «Augmented Anatomy», розроблену для запропонованої нами платформи, з експертами та студентами в США та Китаї. Результати вказують на те, що самостійне вивчення анатомії справді корелює з підвищенням інтересу до STEM і оцінюється як більш привабливе, ефективне та веселе, ніж навчання анатомічних структур лише за підручниками.

Як і віртуальна реальність, доповнена реальність стає новою платформою в нових сферах застосування для музеїв, освітніх розваг, домашніх розваг, досліджень, промисловості та мистецьких спільнот, використовуючи нові підходи, які вивели доповнену реальність за межі традиційних наочних або портативних дисплеїв. У цій книзі автори обговорюють підходи до просторової доповненої реальності, які використовують оптичні елементи, відеопроєктори, голограми, радіочастотні мітки та технологію відстеження, а також алгоритми інтерактивного рендерингу та методи калібрування для вбудовування

синтетичних добавок у реальне середовище або в живе відео реального середовища.

У промисловості AR можна використовувати для чого завгодно: від ідентифікації активів до передачі знань у польових умовах і навчання. Поєднуючи фізичне з віртуальним, щоб покращити спосіб роботи людей, доповнена реальність дає працівникам більше інформації та контексту про продукт або машини, на яких вони працюють, і світ навколо них. AR зазвичай використовується в таких областях, як дизайн та обробка продукту, завдяки доповненій реальності, цифровим близьнюкам та інтернет речей дизайнери продуктів можуть оживляти продукти, тестувати та коригувати їх, перш ніж щось фізичне буде створено, технічне обслуговування, за допомогою AR працівники можуть отримати негайну інформацію про будь-яку машину, з якою вони взаємодіють, навчання співробітників доповнена реальність дозволяє співробітникам навчатись роботі з будь-якою машиною чи обладнанням «на вимогу», перетворюючи їхнє безпосереднє оточення на платформу для постійного навчання, контроль якості, включення AR у контроль та забезпечення якості може допомогти запобігти дефектам під час виробництва, оптимізувати виробничий процес та скоротити час виходу на ринок.

Компанії використовують додатки на базі AR, щоб пропонувати покрокові посібники, документацію, посібники тощо, щоб допомогти навчати нових працівників і підвищувати кваліфікацію існуючих команд. Таке практичне навчання в цифровому середовищі може підтримати або замінити викладачів і зробити робоче місце безпечнішим. І все це можливо, тому що AR пропонує легкий доступ до інформації з будь-якої точки світу в будь-який час завдяки 3D-моделям, покроковим інструкціям і можливостям для співпраці. Уявіть собі створення цифрового двійника вашої фабрики, щоб працівники могли виконувати роботу з технічного обслуговування більш ефективно та віддалено та оптимізувати роботу на виробничих лініях. Цифрові двійники також можна використовувати для створення макетів, які показують, як буде працювати готовий продукт, включати запити клієнтів у проекти та допомагати

керувати прогнозованим обслуговуванням. Інтеграція типу Industry 5.0 із технологією цифрового близнюка, де цифровий двійник діє як клон фізичного об'єкта в парі з IoT, робить це можливим. Завдяки ефекту Amazon сучасні клієнти очікують швидкого отримання своїх замовлень, а оптимізація управління складом і робочого процесу AR за допомогою доповненої реальності може допомогти в цьому. За допомогою складського додатка на базі AR співробітники можуть отримувати інформацію про замовлення та інструкції для виконання кожного замовлення найбільш ефективно. Додаток також сканує штрих-коди для них, заощаджуючи час на ручне сканування штрих-кодів. А після завершення замовлення система інвентаризації автоматично оновлюється.

## 2.2 Інструменти доповненої реальності

Reality Composer — це потужний інструмент, який спрощує створення інтерактивної доповненої реальності без попереднього досвіду 3D. Конвертер реальності швидко перетворює ваші існуючі 3D-моделі в USDZ, щоб бездоганно працювати в наших інструментах і на всіх пристроях iPhone і iPad із підтримкою AR. Індустрія розвитку AR є високо конкурентно спроможною з основними технологічними компаніями, які вкладають кошти у власні комплекси для розробки програмного забезпечення AR – SDK. Розвиток цієї технології AR означає, що 55 % людей, які взяли участь в опитуванні, пов'язаному зі штучним інтелектом, заявили, що використовують голосову допомогу для взаємодії з пристроями, які використовують доповнену реальність у своїх домівках. Як правило, ринок можна розділити на дві області розвитку AR-інструментів:

Доповнена реальність описує досвід користувача, який додає 2D або 3D-елементи до реального перегляду з камери пристрою таким чином, що б ці елементи виглядали як живі у реальному світі, ARKit поєднує в собі відстеження руху пристрою, зйомку сцени камерою, розширену обробку сцени та зручності відображення, що б спростити завдання створення досвіду AR. За допомогою цих технологій можна створювати різноманітні види доповненої реальності за

допомогою передньої чи задньої камери пристрою iOS.

ARKit (рисунок 2.5) допомагає розробникам створити найкращий досвід доповненої реальності, додатки, які використовують ARKit, зазвичай мають більш стабільне відстеження AR і детальніший вміст ніж інші програми. Відстеження навколишнього середовища, положення користувача в просторі завжди можна відобразити, створюючи таким чином основу для дуже стабільного досвіду AR, це робить розумний вибір відповідної пам'яті для збереження даних карти.

У AR або MR маркери — це зображення або об'єкти, зареєстровані в програмі, які діють як тригери інформації у вашій програмі. Коли камера вашого пристрою розпізнає ці маркери в реальному світі (під час роботи програми AR або MR), це запускає відображення віртуального вмісту поверх положення маркера у світі в поданні камери. Відстеження на основі маркерів може використовувати різноманітні типи маркерів, включаючи QR-коди, фізичні світловідбиваючі маркери, цілі зображення та 2D-теги. Найпростішим і найпоширенішим типом маркера в ігрових програмах є Image Target. Об'єкти зображення – це певний тип маркера, який використовується для відстеження на основі маркерів. Це зображення, які ви вручну реєструєте в програмі, і діють як тригери, які відображають віртуальний вміст. Для цілей зображення використовуйте зображення, що містять чіткі форми зі складними контурами. Це полегшує їх розпізнавання алгоритмам розпізнавання зображень і відстеження. Програми, які використовують безмаркерне відстеження, частіше використовують доповнену або змішану реальність на основі розташування або позиції. Ця форма відстеження спирається на такі технології, як GPS, акселерометр, гіроскоп і більш складні алгоритми обробки зображень, щоб розміщувати віртуальні об'єкти чи інформацію в середовищі. Апаратне та програмне забезпечення віртуальної реальності обробляє ці об'єкти так, ніби вони закріплені або пов'язані з конкретними місцями чи об'єктами реального світу. Деякі інструменти Vuforia підтримуються лише на певному обладнанні. У таблиці нижче наведено ці інструменти разом із підтримуваними пристроями та необхідними версіями ОС. Більшість програм AR і MR призначені для мобільних пристроїв, тому цей

посібник зосереджений на розробці для Android та iOS. Перегляньте посібники з увімкнення підтримки збірки для пристроїв Android та iOS у документації «Початок роботи» для Android та iOS . Для компонента Camera на ARCamera за замовчуванням дальні площини відсікання встановлені на 2000.



Рисунок 2.5 – Застосунок Apple ARKit

ARCore — це платформа Google для створення досвіду доповненої реальності. Використовуючи різні API, ARCore дозволяє вашому телефону відчувати навколишнє середовище, розуміти світ і взаємодіяти з інформацією. Деякі API доступні для Android та iOS, щоб забезпечити спільний

досвід AR. ARCore використовує три ключові можливості для інтеграції віртуального вмісту з реальним світом, який видно через камеру вашого телефону:

Фокус ARCore йде на таких напрямках:

1. Відстеження руху.
2. Розпізнавання навколишнього середовища.
3. Оцінка освітлення.

По суті, ARCore робить дві речі: відстежує положення мобільного пристрою під час його руху та створює власне розуміння реального світу.

Технологія відстеження руху ARCore використовує камеру телефону для визначення цікавих точок, які називаються функціями, і відстежує, як ці точки рухаються з часом. Завдяки поєднанню руху цих точок і показників інерційних датчиків телефону ARCore визначає як положення, так і орієнтацію телефону під час його руху в просторі.

Окрім визначення ключових точок, ARCore може виявляти плоскі поверхні, наприклад стіл або підлогу, а також оцінювати середнє освітлення в зоні навколо них. Поєднання цих можливостей дає змогу ARCore будувати власне розуміння навколишнього світу.

Розуміння реального світу за допомогою ARCore дає змогу розміщувати об'єкти, анотації чи іншу інформацію таким чином, щоб бездоганно інтегруватися з реальним світом. Ви можете поставити кошения, що дримає, на кут журнального столика або додати до картини біографічну інформацію про художника. Відстеження руху означає, що ви можете пересуватися та розглядати ці об'єкти під будь-яким кутом, і навіть якщо ви розвернетесь і вийдете з кімнати, коли ви повернетесь, кошения чи анотація будуть там, де ви їх залишили.

Щоб отримати детальнішу інформацію про те, як працює ARCore, перегляньте основні поняття.

ARCore надає SDK для багатьох найпопулярніших середовищ розробки. Ці SDK надають власні API для всіх основних функцій AR, таких як відстеження руху, розуміння навколишнього середовища та оцінка освітленості. Завдяки цим

можливостям ви можете створювати абсолютно нові можливості доповненої реальності або вдосконалювати наявні програми функціями доповненої реальності.

ARCore розпізнає характерні точки будь-якої горизонтальної поверхні (це може бути стіл, зображення чи плитка на тротуарі) і перетворює їх на площину для відображення 3D-вмісту. Як стандартний фреймворк доповненої реальності, ARCore може мати проблеми з виявленням плоских об'єктів без належної текстури та контрасту, наприклад, білої стіни або чорної підлоги. Щоб підвищити рівень реалістичності, ARCore визначає інтенсивність освітлення в сцені, аналізуючи зображення, надане візуальною системою (камерою). Потім структура відображає 3D-контент з такою ж кількістю світла, що й реальні об'єкти у світі. Крім того, пози змінюються під час взаємодії з програмою, і ARCore гарантує, що ваш досвід не буде зіпсований неправильним відображенням 3D-об'єктів. Щоразу, коли ви вирішуєте доповнити реальний світ будь-яким віртуальним вмістом, вам потрібно визначити прив'язку для вмісту, щоб ARCore міг точніше відстежувати позицію. Крім того, це дозволяє стабілізувати вміст відносно реального світу, навіть якщо пристрій рухається.

ARCore від Google (рисунок 2.6) відкриває доповнену реальність для всіх користувачів, дозволяючи компаніям бути більш інноваційними, ніж будь-коли. Невідомо, наскільки ці інновації можуть покращити дизайн UX. Як було зазначено в попередніх статтях, AR чудово підходить для покращення візуалізації продукту. Такі компанії, як Wayfair та IKEA, використовували AR у своїх програмах, щоб допомогти споживачам переглядати цифрові візуалізації продуктів. Це покращує взаємодію з користувачами для споживачів, оскільки їм не потрібно проходити через кропіткий процес фізичного замовлення товарів і перевірки їх ефективності. Декілька інших компаній використовували AR для допомоги у візуалізації продукту.

Geospatial ARCore API був запущений у травні цього року як розширена та більш адаптована версія технології ARCore Cloud Anchors API 2019 року. Розробники можуть використовувати ARCore API для створення більш

захоплюючого, насиченого та цінного досвіду на основі доповненої реальності для користувачів. ARCore також має версію з відкритим кодом, яка дозволяє користувачам і ентузіастам досліджувати варіанти використання технології доповненої реальності. Серед найвідоміших застосувань — Pocket Garden, що дозволяє користувачам створювати барвисті віртуальні сади, щоб прикрашати їхні райони.



Рисунок 2.6 – Google – ARCore

Таблиця 2.1 – Інструменти Vuforia

Додаток	Пристрої	Версія ОС
---------	----------	-----------

Асистент калібрування	Moverio BT-200 ODG R-7	Android 4.0.3
Сканер об'єктів	Samsung Galaxy S8+ Samsung Galaxy S8 Samsung Galaxy S7 Samsung Galaxy S6	Остання підтримувана ОС на пристрої

Vuforia Engine є найпоширенішою платформою для розробки AR, яка підтримує більшість телефонів, планшетів і окулярів. Розробники можуть легко додавати розширені функції комп'ютерного бачення до програм Android, iOS та UWP, щоб створювати доповнену реальність, яка реалістично взаємодіє з об'єктами та середовищем. Vuforia підтримує багато пристроїв сторонніх виробників (наприклад, окуляри AR/MR) і пристрої VR із задньою камерою (наприклад, Gear VR). Повний список підтримуваних пристроїв див. на сторінці Vuforia у розділі «Пристрої». Vuforia Unity розпізнає та відстежує плоскі зображення та 3D-об'єкти в режимі реального часу за допомогою спеціальної технології зору з підтримкою AI. Завдяки цій можливості розробники можуть знаходити та орієнтувати віртуальні об'єкти (3D-моделі) у реальному світі під час перегляду через камеру смартфона. Однією з великих переваг Vuforia Unity є те, що цей SDK підтримує широкий спектр об'єктів, як 2D, так і 3D, а також має можливість установлювати посилання, які діють як маркери, які повністю адресуються. На додаток до цього, ця система розробки додатків Unity AR дозволяє програмістам:

1. Свобода розташування пристроїв у просторі.
2. Локалізоване вивлення оклюзії за допомогою «віртуальних кнопок».
3. Вибір цільового зображення під час виконання.
4. Можливість програмно створювати та налаштовувати параметри цілей під час роботи програми.

Програма пропонує основні функції:

5. Розпізнавання тексту.
6. Розпізнавання зображень.
7. Відстеження цілей, встановлених завдяки SDK, не втрачаються, навіть якщо мобільні пристрої змінюють своє місцезнаходження.
8. Одночасне виявлення та відстеження.
9. Швидке виявлення цілей і визначених цілей.

Переваги використання Vuforia:

1. Програми, розроблені з використанням Vuforia, мають доступ до локальних і хмарних баз даних.
2. Він забезпечує підтримку пристроїв віртуальної реальності.
3. Тестові приклади можливостей включені як частина SDK.
4. Він містить багато конкретних інструкцій і коротких порад.

Незважаючи на те, що він має значні переваги, необхідно також зазначити, що цей SDK не рекомендується людям, які навчаються програмувати доповнену реальність або використовують Vuforia вперше.

### 2.3 Висновок до розділу 2

В цьому розділі мною були описані приклади та створені схеми доповненої реальності, враховуючи всі переваги та недоліки. Було проведено обґрунтування переваг Apple ARKit над Google ARKit, як показала практика, по якості фото знімків або відео зйомки, архітектура компанії Apple дійсно переважає над аналогами. В результаті огляду засобів для користування користувача можна дійти до висновку, що для звичайного користувача серед всіх бібліотек таких як ARKit, OpenCV, Wikitude, краще підходить ARKit.

Також розібрав види доповненої реальності, так як кожна впливає на життя людей і в яких сферах застосовується. В результаті аналізу можливостей використання найпопулярніших віртуальних реальностей, описаних вимог та потреб системи, зі значними перевагами було обрано мобільний додаток Apple

ARKit. Було проведено огляд та порівняння описаних мною видів віртуальної реальності для користувацького інтерфейсу. Попередня обробка хмари відіграє вирішальну роль у підготовці необроблених даних хмари для подальших процесів ідентифікації об'єктів. На цьому етапі використовуються різні методи, спрямовані на підвищення якості даних і отримання значущої інформації.

### 3. АПАРАТНА ЧАСТИНА

#### 3.1 Плата мікроконтролера Arduino Leonardo

Arduino Leonardo - це контролер на базі мікроконтролера ATmega32u4. Він має різні компоненти, такі як 20 цифрових виводів вводу/виводу, кварцовий генератор на 16 МГц, роз'єм micro-USB, роз'єм живлення, роз'єм ICSP та кнопка скидання. Для роботи Arduino Leonardo можна підключити до комп'ютера за допомогою USB-кабелю або жити через адаптер змінного/постійного струму чи батарею. У більшості контролерів Arduino, включаючи Leonardo, немає вбудованого USB-контролера. Замість цього вони використовують перетворювач інтерфейсу COM-USB для зв'язку з комп'ютером через фізичний USB-канал. В результаті, програміст може отримати доступ тільки до внутрішнього UART (універсального асинхронного приймача-передавача) для обміну даними.

Живлення Arduino Leonardo (рисунок 3.1 та 3.2) [4,5] може здійснюватися як через порт micro-USB, так і від зовнішнього джерела живлення, причому тип джерела живлення визначається автоматично. Зовнішнім джерелом живлення може бути мережевий AC/DC адаптер або батарея, штекер адаптера вставляється у відповідний роз'єм живлення на платі. При використанні живлення від батареї дроти слід підключати до виводів Gnd і Vin роз'єму POWER. Leonardo відрізняється від усіх попередніх плат тим, що ATmega32u4 має вбудований USB-зв'язок, що усуває потребу у додатковому процесорі. Це дозволяє Leonardo відображатися на підключеному комп'ютері як миша та клавіатура на додаток до віртуального (CDC) послідовного/COM-порту. Це також має інші наслідки для поведінки правління; вони детально описані в розділі Це означає Serial Peripheral Interface. Ці контакти використовуються мікроконтролерами для ефективного зв'язку з одним або кількома периферійними пристроями. На відміну від Arduino UNO, контакти SPI на платі Leonardo присутні на роз'ємі ICSP, ці контакти підтримують зв'язок SPI за допомогою бібліотеки SPI. Це означає, що навіть якщо у нас є екран, який використовує зв'язок SPI, але НЕ складається з 6-контактного роз'єму ICSP, який можна підключити до 6-контактного роз'єму ICSP Леонардо,

екран не працюватиме. Діапазон напруги зовнішнього джерела живлення зазвичай становить від 6 до 20 В. Однак, якщо напруга живлення падає нижче 7 В, напруга на виводі 5 В може знизитися, що призведе до нестабільної роботи. З іншого боку, використання напруги вище 12 В може призвести до перегріву стабілізатора напруги і потенційного пошкодження плати. Для оптимальної роботи рекомендується використовувати джерело живлення в діапазоні від 7В до 12В.

Цифрові виводи Arduino Leonardo можна налаштувати як входи або виходи за допомогою таких функцій, як `pinMode()`, `digitalWrite()` і `digitalRead()`. Робоча напруга для виводів становить 5 В, і кожен вивід може подавати або споживати максимальний струм 40 мА. Всі виходи підключені до внутрішніх підтягуючих резисторів з номінальними значеннями від 20 кОм до 50 кОм. Для комунікації Arduino Leonardo пропонує кілька варіантів. Він має приймач UART, який забезпечує послідовний зв'язок з іншими пристроями або мікроконтролерами через цифрові виводи 0 (RX) і 1 (TX). Крім того, мікроконтролер ATmega32U4 підтримує послідовний зв'язок через USB. При підключенні до комп'ютера він може бути розпізнаний як віртуальний COM-порт, використовуючи стандартні драйвери USB-COM. Пакет програмного забезпечення Arduino включає програму, яка полегшує зчитування та надсилання простих текстових даних до Arduino. Під час USB-зв'язку світлодіоди RX і TX на платі блимають, вказуючи на передачу даних між комп'ютером і Arduino. Точна ідентифікація об'єктів у тривимірних хмарах має значні переваги для різних областей, зокрема робототехніки, автономної навігації, доповненої реальності та міського планування. Наприклад, під час автономного водіння точна ідентифікація об'єктів у хмарах має вирішальне значення для виявлення та уникнення перешкод. У робототехніці розпізнавання об'єктів у тривимірних хмарах точок має важливе значення для розуміння сцени та завдань маніпулювання. Крім того, у міському плануванні здатність ідентифікувати та аналізувати об'єкти в 3D-хмарах точок полегшує моделювання міської інфраструктури та екологічну оцінку. Для цього він зчитує поточне положення двигуна з двофазного енкодера, підключеного до двигуна постійного струму, і обчислює необхідний вихідний сигнал ПІД для приводу

двигуна до цільового кута. Керувати серводвигунами набагато легше, якщо використовувати лише схему керування з відкритим контуром. Як зазначалося раніше, алгоритми прямої та зворотної кінематики написані в MATLAB і виконуються на комп'ютері-супутнику. Обидва алгоритми повертають набір кутів для кожного з чотирьох з'єднань, які потім надсилаються на вбудований контролер через інтерфейс послідовного зв'язку. Для комп'ютера-супутника я використовую звичайний персональний комп'ютер із встановленою операційною системою Windows 10. Вибір і виділення відповідних ознак із хмари мають вирішальне значення для точної та ефективної ідентифікації об'єкта. Функції на основі форми дають змогу зрозуміти геометричні властивості та локальні варіації об'єктів, а функції на основі кольору фіксують візуальний вигляд і характеристики кольору. Поєднуючи функції на основі форми та кольору, можна досягти більш повного представлення об'єктів у хмарі, забезпечуючи надійну та чітку ідентифікацію об'єктів. Крім того, методи вибору ознак допомагають оптимізувати простір ознак, підвищуючи ефективність і результативність наступних алгоритмів класифікації об'єктів. Для вбудованого контролера, я використовуватиму плату Arduino Leonardo, яка підійде ідеально як для початківців та дозволить легко підключати різноманітні датчики та виконавий механізм. З цією платою не потрібно турбуватись про драйвери низького або середнього рівня та бібліотеки для виконання основних функцій, таких як керування двигуном, послідовний зв'язок або керування зовнішніми перериваннями. Бібліотеку контролерів Arduino можна використовувати, що б додати один або кілька входів до списку пристроїв HID Arduino Leonardo або Arduino Micro може підтримувати, але це не працюватиме з Arduino IDE 1.6.5 або з пристроями Arduino на основі не 32u4. Вирішення цих проблем та перешкод має вирішальне значення для подальшого прогресу в семантичній сегментації хмар. Поточні дослідження спрямовані на розробку надійних алгоритмів, які можуть обробляти розрідженість, шум, оклюзії та різну щільність, а також вирішувати проблеми масштабованості, узагальнення та анотації даних для покращення загальної продуктивності та практичного застосування методів семантичної

сегментації хмари. Вартість і зусилля, пов'язані з отриманням анотованих навчальних даних, можуть бути суттєвою перешкодою, особливо при роботі з великомасштабними або спеціалізованими наборами даних. Цей аспект часто обмежує доступність і різноманітність навчальних даних, що може вплинути на узагальнення та продуктивність моделей семантичної сегментації.

Таблиця 3.2 - апаратна специфікація Arduino Leonardo

Мікроконтролер	ATmega32u4
Робоча напруга процесора	5В
Робоча частота процесора	16МГц
Напруга зовнішнього живлення	7-12В
Кількість цифрових входів/виходів	20
Цифрових виходів з можливістю ШІМ	7
Кількість аналогових входів	12
Допустимий струм на цифрових виходах	40Ма
Обсяг Flash пам'яті програм, всього	32Кб
Обсяг Flash, зайнятої завантажувачем	4кб
Обсяг оперативної пам'яті SRAM	2.5кБ
Обсяг енергонезалежної пам'яті EEPROM	1кБ
Комплектність	Плата в зборі

### 3.2 Камера пристрою

Швидкість обробки має вирішальне значення для багатьох програм ідентифікації об'єктів. Це особливо актуально, коли зір використовується як датчик у системі керування, як це часто зустрічається в робототехніці, машинне бачення та спостереження. Затримка обробки обмежує корисність даних, отриманих шляхом обробки зображень через збільшення затримок у циклі зворотного зв'язку може серйозно вплинути на керованість системи. Як наслідок буде гірша продуктивність, або в крайньому випадку, нестабільність ідентифікації. Мобільна робототехніка накладає додаткові обмеження на обробку зору, Розмір системи і вага часто обмежені, а оскільки мобільні роботи зазвичай працюють від акумулятора, споживана потужність важлива, складність багатьох алгоритмів бачення часто вимагає високопродуктивного обчислення для керування високою швидкістю передачі даних із камер. Це зазвичай передбачає роботу процесорів із високою тактовою частотою лише для підтримки йде висока пропускна здатність пікселів.

Високі тактові частоти призводять до вищих вимог до електроенергії, цей стан речей погіршується зростанням роздільної здатності недорогих цифрових пристроїв камери. Тоді як підвищена роздільна здатність загалом покращує продуктивність багатьох алгоритмів зору, це відбувається за рахунок збільшення обчислювального навантаження. На збільшення вартості обробки зменшує час, виділений на виконання інших завдань, наприклад функція стратегії та контролю, що б подолати деякі з цих проблем, обробка все частіше переноситься в межах камери.

Коли йдеться про аналітику відеоспостереження, виявлення та ідентифікація об'єктів є ключовим компонентом. Ця можливість дає змогу операторам знаходити певний об'єкт, наприклад людину, транспортний засіб або рюкзак, а потім відстежувати його від кадру до кадру. Результат можливість швидко ідентифікувати об'єкт за допомогою багатогодинної зйомки, надаючи тим

самим слідчим служби безпеки та поліції критичні криміналістичні докази; власники бізнесу дуже необхідні знання; і більш ефективні програми для різноманітних користувачів. кщо пояснити, виявлення об'єктів — це просто здатність виявляти об'єкти на відео, проте цей процес далеко не простий, якщо його виконувати за допомогою передової технології. Перший крок, «вилучення об'єкта», виявляє та відстежує об'єкт у будь-який момент часу. Потім додається інша технологія під назвою «розділення фону/переднього плану», яка визначає статичний фон сцени.

Виявлення та розпізнавання об'єктів можна використовувати до, під час і після події. Можливості перегляду дозволяють розслідувати подію, яка вже сталася. Але також можна ввімкнути сповіщення в режимі реального часу, що дозволяє реагувати на подію, коли вона відбувається, щоб ідентифікувати об'єкти інтересу. Виявлення та розпізнавання об'єктів також має застосування не лише для правоохоронних органів. Наприклад, можна встановити сповіщення в режимі реального часу для певної камери на вході на будівельний майданчик, щоб гарантувати, що кожен, хто проходить повз, носитиме каску. Якби хтось не носив позначений предмет, у цьому випадку помаранчеву каску, сповіщення було б надіслано операторам, які потім могли б відповісти в режимі реального часу.

Останніми роками, технологія програмованої вентиляційної матриці дозріла на етапі обробки зображень на практиці. Ці технології були використані в широкому спектрі способів в розумних камерах. Один був через використання SIMD, архітектура для використання просторового паралелізму. Для них ідентичні обробні елементи використовуються паралельно. Інша поширена архітектура для обробки зображення є віконний процесор. Така архітектура дозволяє використовувати широкий спектр фільтрів, які будуть реалізовані, включаючи лінійні фільтри, градації сірого, морфології та суми абсолютної різниці. На іншому кінці спектру знаходиться виділена логіка розроблена для конкретного застосування та реалізована на програмі. Такі підходи як правило є найменшими вимогами до ресурсів, але менш загального призначення. Проте керування пам'яттю ділиться на блоки, інтерфейс камери та деякі основні операції можна

повторно використовувати від одного дизайну до іншого. У програмному забезпеченні більшість зображень шляхом зчитування зображення з пам'яті виконання операції, і запис результатів назад у пам'яті, час який необхідний для такої обробки переважають звернення до пам'яті, а не фактична виконана обробка.

У розумній камері однак обробку можна перемістити з обчислень на основі зображень на обчислення на основі пікселів обчислення.

Процесори загального призначення намагається встигати за вимогами обробки програми обробки зображень, з цієї причини є процесори цифрових сигналів. Зазвичай використовується для впровадження розумних камер. Архітектура DSP була оптимізована для обробки сигналів кількома способами, використана Гарвардська архітектура подвоює пропускну здатність пам'яті шляхом розділення інструкцій і пам'ять даних. ЦП також має функції множення на накопичення за один цикл, що прискорює фільтрацію та інші пов'язані операції. Обробка низькорівневого зображення працює незалежно на окремих пікселях, що дозволяє використовувати асоційований паралелізм. Це призвело до появи паралельних процесорів, які працюють одночасно на кількох частинах зображення.

Для отримання необхідної візуальної інформації використовується USB веб-камера Logitech WebCam C270 (зображена на рисунку 3.3). Ця веб-камера підключається до комп'ютера через USB-роз'єм. Вона дозволяє знімати відео з роздільною здатністю 1280x720 з максимальною частотою кадрів 30 кадрів в секунду. Для виконання заздалегідь визначених завдань системи, які включають горизонтальне обертання камери, використовується двовісний кронштейн (показаний на рис. 3.4). Завдання виявлення об'єктів у хмарі 3D-точок створює складнощі через складну природу даних, що вимагає точної локалізації та всебічного просторового осмислення.

Алгоритми машинного навчання, такі як CNN, SVM і випадкові ліси, пропонують потужні рішення для виявлення об'єктів.

Однак, доступність анотованих навчальних даних і точна анотація хмар створюють певні проблеми. Подолання цих проблем і використання просторових

зв'язків призведе до покращення продуктивності виявлення об'єктів у 3D-хмарах, уможливаючи різні програми, такі як робототехніка, автономне водіння та доповнена реальність.

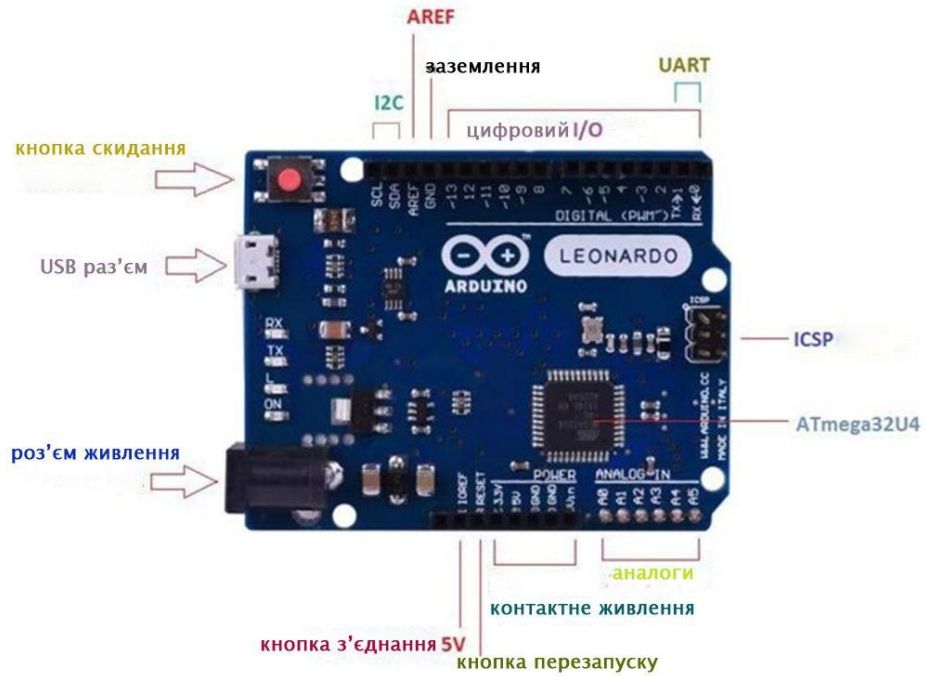


Рисунок 3.1– Arduino Leonardo вигляд зверху [4]



Рисунок 3.2 – Arduino Leonardo вигляд знизу[5]



Рисунок 3.3 – Камера WebCam Logitech C270 [6]

Веб-камера (рисунок 3.3) Logitech WebCam C270 під'єднується до комп'ютера за допомогою USB-з'єднання. Ця веб-камера дозволяє користувачеві

знімати відео з роздільною здатністю 1280x720 з максимальною частотою кадрів 30 кадрів в секунду. Для виконання попередньо визначених завдань системи, які включають горизонтальне обертання камери, був використаний двохвісний кронштейн (зображений на рисунку 3.4)[4].

Сервопривід Tower Pro 9g SG90 (рисунок 3.5) не має потужних характеристик, але має не дорогу цінову політику. Дуже добре підходить для не складних завдань під управлінням контролерів Arduino.

Сервопривід, це такий мотор-редуктор, що здатен повертати вихідний вал в задане положення і утримувати його в цьому положенні, в супереч опорам і збурень. Це потрібно в основному використовуються авіамоделістами для керування положенням різних закрилків, рулів і лопатей гелікоптера.

Підключення контактів сервоприводів відбувається наступним чином:

1. Коричневий дріт: "-" (мінус або земля).
2. Червоний дріт: "+" (позитивний або живлення).
3. Помаранчевий провід: "управління" (сигнал)».



Рисунок 3.4 –Двохосьовий кронштейн [7]

#### Технічні характеристики:

1. Вага: 9 г.
2. Розміри: 21.5 x 11.8 x 22.7 мм.
3. Стандартний роз'єм: futaba/JR універсальний.
4. Довжина кабелю: 15 см.
5. Кут повороту: 180 градусів.
6. Матеріал шестерні: нейлон (пластик).
7. Пусковий момент: 1,2-1,4 кг/см (при 4,8В).
8. Швидкість без навантаження: 60 градусів/0,12 секунди (при 4,8В).
9. Робоча температура: від -30°C до +60°C.
10. Робоча напруга: від 3,5 В до 6,0 В.



Рисунок 3.5 – Сервопривід Tower Pro 9q SG90

У робототехнічних програмах семантична сегментація хмар допомагає в задачах маніпулювання об'єктами.

Наприклад, рука робота, оснащена датчиком глибини, може зафіксувати хмару захищеного столу.

Застосовуючи семантичну сегментацію, робот може ідентифікувати різні об'єкти на столі, такі як чашки, тарілки та столові прибори, і планувати відповідні стратегії хапання для завдань маніпулювання.

Алгоритми ідентифікації об'єктів хмари мають бути розроблені таким чином, щоб бути стійкими до змін навколишнього середовища та датчиків. Це включає обробку змін умов освітлення, погодних умов, шуму датчика та змін у характеристиках датчика.

Дослідницькі зусилля повинні вивчати методи покращення можливостей узагальнення алгоритмів, що дозволяє їм працювати послідовно в різних умовах середовища та датчиків.

Щоб використовувати цей двигун, ви повинні використовувати щось, що може генерувати сигнали ШІМ, які можуть бути будь-якими від 555 Timer IC до інших мікроконтролерів, таких як Arduino, Raspberry Pi, PIC тощо. Живить двигун за допомогою контактів VCC і GND і керуйте рухом валу за допомогою ШІМ-вхід до нього. 1 × Серводвигун Tower Pro SG90 - Міні/мікросерводвигун 9 г.

### 3.3 Схема підключення пристрою

Кабельна розводка починається з хосту, він володіє інтегрованим кореневим концентратором, який надає декілька роз'ємів USB для встановлення зв'язку із зовнішніми пристроями в системі відеоспостереження використовується мережа кабелів, які з'єднують різні USB-пристрої, деякі з яких можуть функціонувати як концентратори для додаткових компонентів.

USB-вихід камери спеціально підключений до персонального комп'ютера (ПК). У системному C++-кодi камера ідентифікується унікальним номером, пов'язаним з USB-портом, до якого вона підключена, як показано на рисунку 3.5.

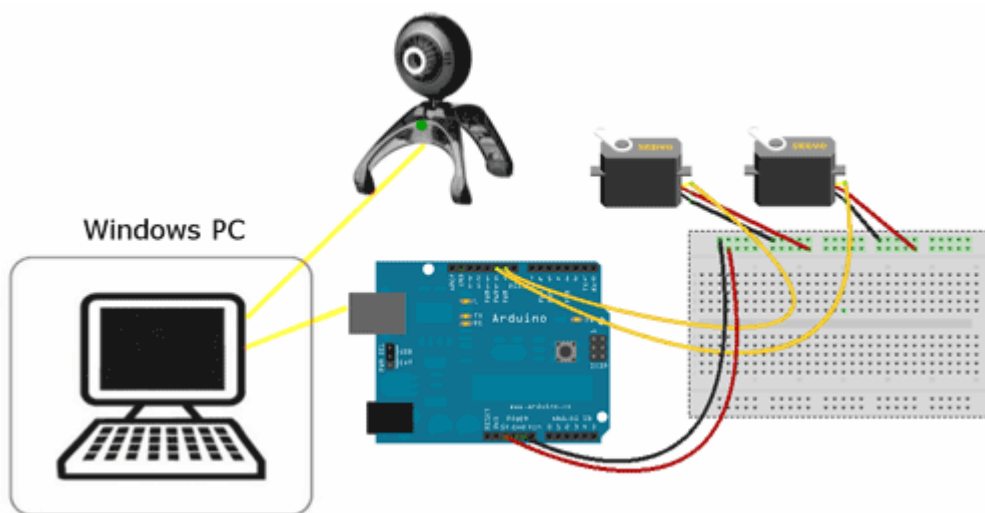


Рисунок 3.5 - схема підключення пристрою

Для сервоприводів функція панорамування використовує жовтий/сигнальний дріт, який підключений до цифрового виходу 9. Аналогічно, функція нахилу використовує жовтий/сигнальний дріт, який підключений до цифрового виходу 10. Червоні дроти VCC обох сервоприводів підключені до виходу 5V Arduino, а чорні дроти GND обох сервоприводів підключені до виходу GND Arduino.

Крім того, контролер з'єднаний з комп'ютером через USB-інтерфейс, як показано на рисунку 3.6. Це з'єднання створює в системі віртуальний COM-порт, який потрібно інтегрувати в програмний код на C++, щоб забезпечити зв'язок і взаємодію з контролером.

Для виявлення об'єктів у тривимірних хмарах можуть бути корисними просторові зв'язки між об'єктами.

Розуміння відносного положення, відстані та орієнтації об'єктів може покращити розуміння сцени та полегшити завдання більш високого рівня, такі як реконструкція сцени або розуміння сцени.

Створення високоякісних анотованих навчальних даних для семантичної сегментації хмар може бути трудомістким і дорогим процесом. Анотація хмар вручну потребує досвіду та ретельного маркування окремих точок. Вартість і зусилля, пов'язані з отриманням анотованих навчальних даних, можуть бути

суттєвою перешкодою, особливо при роботі з великомасштабними або спеціалізованими наборами даних.

Цей аспект часто обмежує доступність і різноманітність навчальних даних, що може вплинути на узагальнення та продуктивність моделей семантичної сегментації.

Виявлення об'єктів у тривимірних хмарах точок є складним завданням через складність даних і потребує точної локалізації та просторового розуміння. Алгоритми машинного навчання, такі як CNN, SVM і випадкові ліси, пропонують потужні рішення для виявлення об'єктів.

Однак, доступність анотованих навчальних даних і точна анотація хмар створюють певні проблеми. Подолання цих проблем і використання просторових зв'язків призведе до покращення продуктивності виявлення об'єктів у 3D-хмарах, уможливлючи різні програми, такі як робототехніка, автономне водіння та доповнена реальність.



Рисунок 3.7 – USB-порт, який представляє собою число ідентифікатора

### 3.4 Висновки до розділу 3

У цьому розділі я заглибився в програмний аспект пристрою спостереження і дослідив сферу комп'ютерного зору за допомогою бібліотеки OpenCV. Вибір OpenCV був зумовлений її великою колекцією алгоритмів, що охоплює понад 2500 варіантів, які поєднують традиційні методи з найсучаснішими досягненнями в галузі комп'ютерного зору та машинного навчання.

Включення OpenCV в систему робить доступними численні функціональні можливості. Однією з таких можливостей є розпізнавання об'єктів, що дозволяє пристрою ідентифікувати і розрізняти об'єкти, присутні як на зображеннях, так і на відео. Крім того, OpenCV полегшує відстеження руху об'єктів, а також руху камери, що дозволяє здійснювати безперервний моніторинг і аналіз динамічних сцен.

Значною перевагою OpenCV є його здатність класифікувати дії на відеозаписах. Це дозволяє пристрою спостереження виявляти та інтерпретувати конкретні дії або поведінку, зафіксовані на відеозаписах.

OpenCV також надає широку підтримку для маніпуляцій із зображеннями та їх обробки, дозволяючи пристрою виконувати широкий спектр операцій, таких як перетворення, покращення та фільтрація зображень. Ці функції сприяють підвищенню якості та ефективності обробки зображень в системі відеоспостереження.

## 4. ДЖЕРЕЛА ТА ЕКСПЕРИМЕНТИ РОЗРОБКИ

### 4.1 Використання бібліотек машинного зору задля ідентифікації динамічних об'єктів

Останніми роками збільшення кількості зображень завдяки прогресу технологій призвело до збільшення кількості немаркованих даних у фоновому режимі. Тому для розробників важливою проблемою стало отримання значущих висновків із цих даних і їх аналіз. Для вирішення цієї проблеми використовуються алгоритми обробки зображень. Виявлення об'єктів – це технологія, яка працює для ідентифікації об'єктів на фото чи відео. Технологія виявлення об'єктів дозволяє нам визначати, де знаходиться об'єкт інтересу на даному зображенні. Визначаючи цей об'єкт, він малює навколо нього прямокутну рамку. Що робить виявлення об'єктів важливим сьогодні, так це те, що воно допомагає нам зрозуміти й проаналізувати сцени на зображеннях. Алгоритми виявлення об'єктів — це технологія штучного інтелекту, пов'язана з комп'ютерним зором і обробкою зображень, яка виявляє моделі семантичних об'єктів певного класу в зображеннях і відео. Конкретним прикладом виявлення об'єктів, який завжди з нами в повсякденному житті, є розблокування наших телефонів за допомогою розпізнавання обличчя. На цьому використанні виявлення об'єктів не закінчується. Від автономних транспортних засобів до систем розпізнавання номерних знаків, від відстеження об'єктів до роботизованого бачення, виявлення об'єктів стало важливою частиною нашого життя. Компанії, які інвестують у ці галузі, сьогодні особливо потребують розробників програмного забезпечення, які працюють над виявленням об'єктів.

Розглянемо сучасні засоби комп'ютерного зору:

1. OpenCV – Бібліотека функцій та алгоритмів для комп'ютерного зору, обробки зображень і числових алгоритмів загального призначення з відкритим кодом.
2. Sxcore – Ядро, містить основні структури даних і алгоритми.
3. CV – Модуль обробки зображень і комп'ютерного зору.

4. Highgui – Модуль для введення/виведення зображень і відео, створення інтерфейсу користувача.
5. Cvaux – Експериментальні та застарілі функції.
6. Tensorflow – Найпопулярніша сьогодні бібліотека машинного та глибокого навчання.
7. CUDA – Архітектура апаратного та програмного забезпечення паралельних обчислень, яка може значно підвищити продуктивність обчислень завдяки використанню графічних процесорів Nvidia.
8. TУАНО – Швидка бібліотека чисел Python.
9. КРАС – Бібліотека глибокого навчання Python, яка поєднує в собі функції інших бібліотек.
10. Scikit – Learn – Популярна бібліотека машинного навчання побудована на NumPy, Scipy та matplotlib.
11. YOLO – Система виявлення об'єктів, для обробки в режимі реального часу.
12. Fastai – Бібліотека глибокого навчання, яка включає функції комп'ютерного зору.
13. IPSDK – Бібліотека розроблена для обробки 2D і 3D зображень.

Ці бібліотеки мають спільні характеристики, такі як безкоштовні ліцензії, швидке створення прототипів на основі Python, використання нейронних мереж для задач розпізнавання та сумісність з веб-платформами для розгортання. Однак розгортання систем машинного зору за допомогою цих бібліотек може стати дорогим, особливо якщо для обробки даних використовуються хмарні сервіси. Наприклад, хмарні фреймворки розпізнавання, такі як Google Cloud Vision, мають такі функції, як розпізнавання облич, розпізнавання тексту, виявлення етикеток, локалізація об'єктів тощо. Ці сервіси тарифікуються на основі використання, і витрати можуть бути значними, що робить їх менш придатними для національних освітніх систем:

1. FACE\_DETECTION - Ідентифікує та виявляє обличчя, присутні на зображенні.

2. OBJECT\_LOCALIZATION - Виявляє та локалізує декілька об'єктів на зображенні.
3. LABEL\_DETECTION - Автоматично додає описові підписи на основі вмісту зображення.
4. DOCUMENT\_TEXT\_DETECTION - Виконує оптичне розпізнавання символів (OCR) для розпізнавання та вилучення щільного друкованого тексту із зображень.
5. WEB\_ВИЗНАЧЕННЯ - виявляє на зображенні релевантні об'єкти, такі як новини, події або знаменитості, і виконує пошук схожих зображень в Інтернеті за допомогою пошуку зображень Google.
6. TEXT\_DETECTION - Виконує оптичне розпізнавання символів (OCR) для розпізнавання та вилучення розрідженого друкованого тексту із зображень.
7. LANDMARK\_DETECTION - ідентифікує та виявляє відомі орієнтири або географічні об'єкти на зображенні.
8. ВЛАСТИВОСТІ ЗОБРАЖЕННЯ - Визначає набір властивостей зображення, включаючи домінуючі кольори, присутні на зображенні.
9. CROP\_HINTS - Допомагає визначити вершини, які визначають відповідну область обрізання на зображенні.

Для нестандартних завдань, зокрема для створення власних моделей машинного навчання для ідентифікації, з використанням платформ на кшталт Vertex AI, можуть виникнути додаткові витрати. Ці кредити дозволяють початківцям вивчати і використовувати їхні послуги як технологічну основу для розробки систем машинного зору. Ця політика заохочує навчання та експерименти в галузі машинного зору і надає можливість навчальним закладам та окремим особам використовувати хмарні платформи в освітніх цілях.

#### 4.2 Розробка прототипу системи машинного зору

"Happy, Sad, Angry Workshop", розроблений J. Bennett, пропонує можливість створити прототип системи машинного зору з використанням Azure Face API. Цей

API слугує програмним інтерфейсом, який дозволяє розпізнавати обличчя на зображенні, визначати 27 орієнтирів обличчя, включаючи положення очей, а також розпізнавати маски.

Для доступу до Azure Face API необхідно отримати ключ підпису. Ключ можна отримати, безкоштовно зареєструвавшись на сайті <https://azure.microsoft.com/en-us/free/students>. Важливо зазначити, що хоча послуги надаються в кредит, для реєстрації не потрібна кредитна картка. Однак, будь ласка, майте на увазі, що поштові адреси в домені другого рівня можуть не розпізнаватися як такі, що належать академічним установам. У таких випадках рекомендується пройти звичайну реєстрацію.

Після отримання ключа його необхідно перевірити або за допомогою тестового коду, або за допомогою веб-форми. Якщо надано неправильний ключ, API поверне повідомлення про помилку з такими кодами стану, як 400, 401, 403, 408, 415 або 429.

При використанні Azure Face API результати повертаються у форматі JSON, що дозволяє отримати доступ до широкого спектру атрибутів, пов'язаних з виявленим обличчям. На хмарі часто впливають розрідженість, шум, оклюзії та різна щільність точок. У розріджених хмарах може бути недостатньо інформації для точного сегментування об'єктів, тоді як шум може внести неточності в процес сегментації.

Оклюзії можуть призвести до неповних даних хмари, через що буде складно сегментувати об'єкти, які частково або повністю приховані. Крім того, різна щільність точок у хмарі може викликати труднощі в досягненні послідовних і надійних результатів сегментації.

Моделі глибокого навчання, такі як FCN або методи на основі графів, можна навчити на позначених наборах даних хмари точок, де кожна точка анується відповідним класом об'єктів. Потім навчену модель можна використовувати для прогнозування міток класів для нових хмар точок.

Це семантичне маркування дозволяє ідентифікувати клас об'єктів і може використовуватися для різних програм, включаючи розпізнавання об'єктів,

розуміння сцени та доповнену реальність. Промислові системи бачення вимагають більшої міцності, надійності та стабільності порівняно з академічними/освітніми системами бачення та зазвичай коштують набагато дешевше, ніж ті, що використовуються в державних/військових програмах. Тому промислове машинне зір передбачає низьку вартість, прийнятну точність, високу стійкість, високу надійність, високу механічну та температурну стабільність.

Системи машинного зору покладаються на цифрові датчики, захищені всередині промислових камер спеціалізованою оптикою для отримання зображень, щоб комп'ютерне обладнання та програмне забезпечення могли обробляти, аналізувати та вимірювати різні характеристики для прийняття рішень.

Системи машинного зору також можуть виконувати об'єктивні вимірювання, такі як визначення зазору свічки запалювання або надання інформації про місцезнаходження, яка спрямовує робота для вирівнювання деталей у виробничому процесі. На малюнку 2 показано приклади того, як системи машинного зору можуть використовуватися для проходження або несправності масляних фільтрів (праворуч) і вимірювання ширини центральної пластини на кронштейні (ліворуч).

Супутня специфікація OPC UA для машинного бачення спрямована на пряму інтеграцію систем машинного бачення в системи управління виробництвом та ІТ. Сфера застосування полягає не лише в доповненні або заміні існуючих інтерфейсів між системою машинного бачення та її середовищем процесу за допомогою OPC UA, а й у створенні неіснуючих можливостей горизонтальної та вертикальної інтеграції для передачі відповідної дати один одному авторизованим учасникам процесу, наприклад прямо до рівня ІТ підприємства.

Таким чином, інтерфейс OPC UA Vision обмінюється інформацією між системою машинного бачення та іншою системою машинного бачення, машинним ПЛК, лінійним ПЛК або будь-якою програмною системою на рівні пристрою керування, що отримує доступ до системи машинного бачення.

Розглянемо результат ідентифікації статичного зображення рисунок 4.3.

Техніка векторного квантування, дуже ефективна технологія стиснення даних, розроблена близько десяти років тому, зараз широко використовується в таких сферах, як кодування зображень. Векторне квантування визначається як набір відображень від вимірного векторного набору до скінченного набору ядром, якого є код книга.

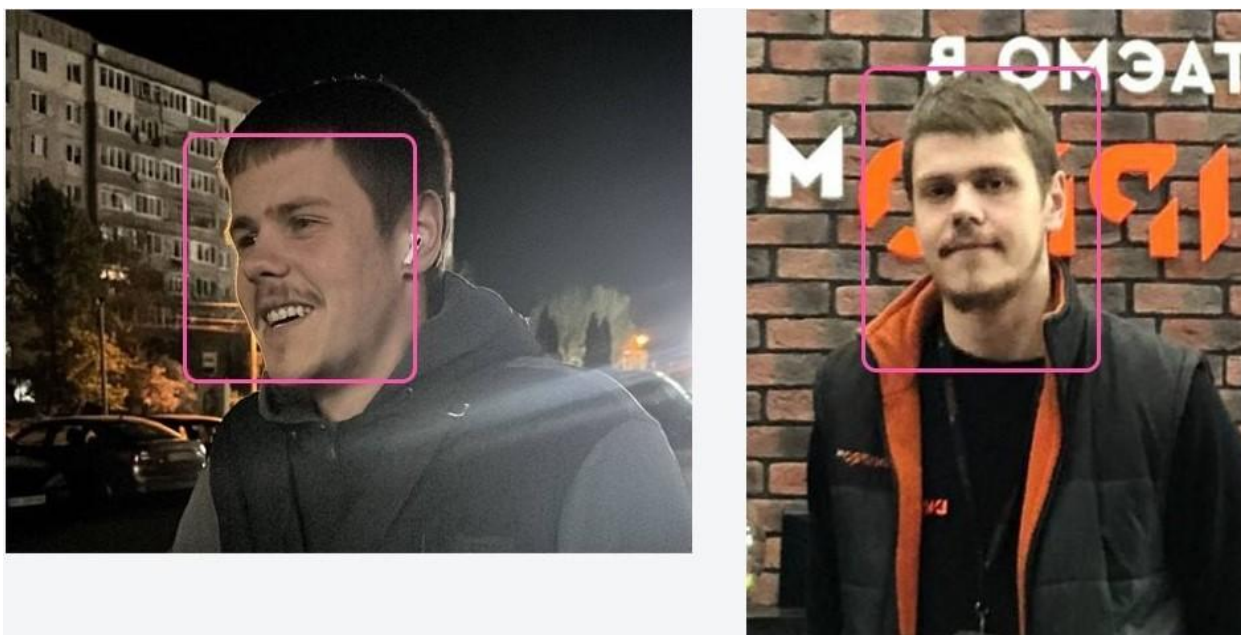


Рисунок 4.3 – Результат ідентифікації статичного зображення

Не зважаючи на те, що цей метод простий і може знайти глобальне оптимальне рішення проблематики, недоліком є відсутність використання кодового слова, що легко призводить до локального оптимального рішення.

Механізм відбору це генетичний алгоритм, однак може добре подолати недоліки, які змушує його уникати локального оптимального рішення, що б знайти глобальне оптимальне рішення. Реставрація зображення полягає у відновленні пошкодженого зображення до його початкового вигляду.

Розглянемо також ідентифікацію динамічних об'єктів з рівнями фокусування на них на рисунку 4.4.

Незважаючи на прогрес, створення пропозицій щодо об'єктів залишається складним. Точність пропозицій зазвичай знижується при роботі зі складними сценаріями бачення, де об'єкти можуть з'являтися в довільних місцях

зображення, різного масштабу, у різних категоріях, а їхня кількість може змінюватись на різних зображеннях. Це пов'язано з тим, що схема знаходиться на попередньому дискримінаційному етапі, тобто алгоритму не дозволено отримати доступ до будь-якої інформації, що стосується категорії об'єкта, і все ж він повинен забезпечити повне покриття: гарантує, що будь-який об'єкт, який буде виявлено, буде закрито одним із вікон-кандидатів.

Таким чином, для традиційних методів, що покладаються на глобальні або локальні особливості зображення, фокус дослідження пропозицій об'єктів належним чином було зосереджено на винаході різних вимірювань, які оцінюють, наскільки ймовірно, що область може містити будь-яку цікаву ціль.



Рисунок 4.4 – Рівні фокусування на динамічних об'єктах

Більшість підходів запропонованих у новітній літературі для виявлення статичних об'єктів, покладається на інформацію відстеження. Недоліком цього підходу є те, що неправильні рішення щодо оновлення можуть згодом призвести

до неправильного виявлення, а незначитий фон може бути виявлений як на передній після видалення статичних об'єктів, навіть якщо вони не поглинаються довгостроковою моделлю, якщо умови освітлення помітно змінились.

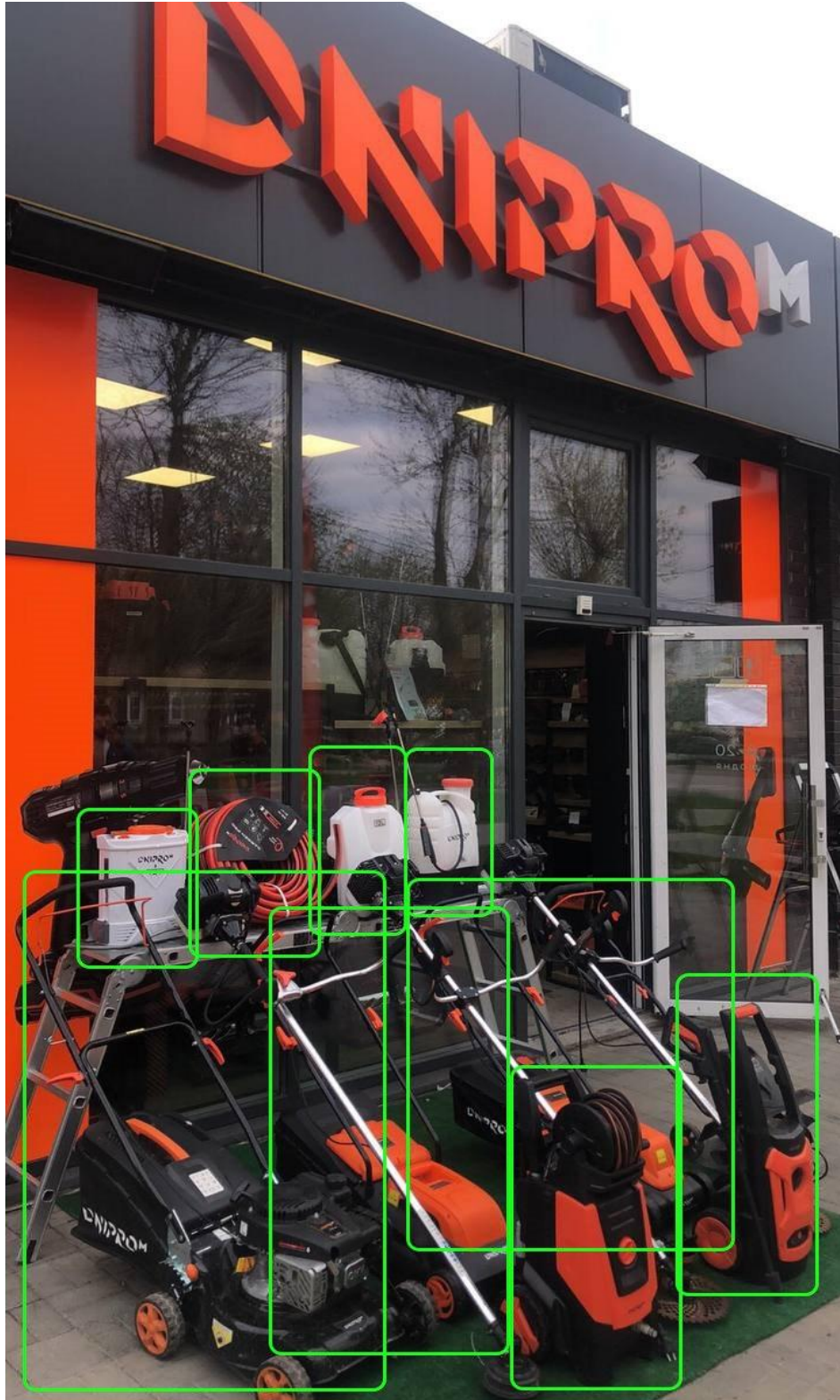


Рисунок 4.4 – Результат ідентифікації статичних об'єктів

Виявлення статичних об'єктів у відеорядах має високу актуальність у багатьох програмах спостереження, до приклада виявлення покинутих об'єктів у громадських місцях. На основі виявлення двох фонових моделей, які навчаються з різною швидкістю, пікселі класифікуються за допомогою кінцевого автомата. Фон моделюється двома сумішами Гаусса з ідентичними параметрами, за винятком швидкості навчання, Кінцевий автомат надає значення для інтерпретації результатів, отриманих від фонового віднімання, його можна реалізувати як таблицю пошуку з незначними обчислювальними витратами і його можна легко розширити, завдяки визначенню станів у кінцевому автоматі система може використовуватися повністю автоматично або інтерактивно, що робить його надзвичайно придатним для застосування в реальному житті, систему було успішно перевірено за допомогою кількох публічних наборів даних, виявлення статичних об'єктів у відеорядах має кілька застосувань у системах відеоспостереження, таких як виявлення незаконно-припаркованих транспортних засобів у моніторингу дорожнього руху або виявлення покинутих об'єктів у системах громадської безпеки. також привернуло увагу широкомасштабних досліджень у галузі відеоспостереження більшість запропонованих методів, спрямованих на виявлення статичних об'єктів, які базуються на виявленні руху, що досягається за допомогою віднімання фону з подальшим певним відстеженням.

Віднімання фону, це широко використовувана техніка для сегментації областей переднього плану у відеопослідовностях, знятих зі статичної камери, яка в основному полягає у виявленні рухомих об'єктів за різницею між поточним кадром і фоною моделлю, що б досягти хороших результатів сегментації, фонова модель повинна регулярно оновлюватися, що б адаптуватись до змінних умов освітлення та стаціонарних змін сцени. Тому методів віднімання фону часто недостатньо для виявлення нерухомих об'єктів і тому доповнюється додатковим підходом.

Дослідження, що вивчають зорову вибірккову увагу, здебільшого вивчали завдання фільтрації, у яких стаціонарні цілі вибираються із стаціонарних відволікаючих засобів за фізичною ознакою, такою як місцезнаходження та ідентичність, що є повідомленим атрибутом. Проте під час багатьох взаємодій із візуальним середовищем цільові стимули вибираються за тим, чим вони є, тоді як дія контролюється тим, де розташований стимул. Це дослідження демонструє, що перешкоди та негативні ефекти праймінгу, які спостерігаються в стандартних завданнях фільтрації, які припускають, що відволікаючі фактори аналізуються та згодом пригнічуються під час вибору, також спостерігаються, коли цілі вибираються на основі ідентичності та просторового розташування є повідомленим атрибутом. Крім того, експериментальні результати показують, що гальмування, пов'язане з відволікаючими об'єктами в цьому новому завданні, є об'єкт-центрованим.

Комбінація масок переднього плану отримана в результаті віднімання двох моделей фону, яка вже використовувалась для швидкої адаптації до змін сцени, одночасно запобігаючи занадто швидкому поглинанню об'єктів переднього плану моделлю фону. Шаблон статичного об'єкта дозволяє додавати 3D-об'єкт до світу, до якого користувач може ходити та довкола нього. Шаблон надає вам доступ до керування маніпуляціями, що дозволяє користувачеві переміщувати, масштабувати та обертати об'єкт.

Після імпортування 3D-об'єкт буде автоматично включено до панелі, а також має бути видимим на панелі. Далі перетягнути щойно імпортований об'єкт, щоб він став дочірнім об'єктом.

Тепер можна видалити об'єкти-заповнювачі шаблону, позначені . На панелі можна позиціонувати, масштабувати, повертати світовий об'єкт відносно початку координат.

У Lens об'єкт світу автоматично розміщуватиметься перед користувачем і прив'язуватиметься до світу.

Об'єкт представляє область, до якої користувач повинен торкнутися, щоб маніпулювати сценою.

На панелі об'єтів обрати об'єкт і масштабувати його так, що б він приблизно відповідав розміру світового об'єкту.

### 4.3 Використання Microsoft Azure для розпізнавання динамічних облич

Почнемо з адаптації прикладу до використання операційної системи Linux, яка часто зустрічається у вітчизняних вищих навчальних закладах. Для роботи з інструментами Azure можна використовувати будь-яке середовище програмування, зокрема спеціалізований текстовий редактор Visual Studio Code. Для налаштування середовища розробки необхідно встановити розширення Python Extension for Visual Studio Code та розширення Azure App Service Extension for Visual Studio Code.

Щоб встановити мікрофреймворк Flask для розробки веб-додатків, ви можете скористатися наступною командою:

```
pip3 install flask.
```

Якщо утиліта pip3 ще не встановлена, ви можете додати її за допомогою менеджера пакетів. Наприклад:

```
sudo apt install python3-pip.
```

Хоча Python 2.7 технічно сумісний, використовувати його не рекомендується. Azure надає можливості для використання різних версій Python, тому бажано використовувати новішу версію.

Щоб налаштувати код Visual Studio для розробки Flask та увімкнути налагодження на стороні сервера, можна слідувати процедурі, описаній J. Bennett. Встановіть змінну FLASK\_APP на повний шлях до системного файлу на стороні сервера, app.py.

Крім того, необхідно внести зміни до конфігураційного файлу Flask у Visual Studio Code. Під час налагодження серверної частини системи ви можете вносити зміни до коду, і Flask автоматично перезапуститься, щоб відобразити оновлення. Клієнтська частина системи представлена одним файлом home.html,

розташованим у каталозі templates.

Аби забезпечити доступність системи, що розробляється, рекомендується розмістити її в хмарі Azure. Одним із способів досягти цього є використання компонента Deploy to Web App служби Azure App Service. Цей процес включає налаштування облікових даних для входу в Azure безпосередньо в коді Visual Studio, вказівку назви системи (яка стає першою частиною доменного імені), вибір потрібної версії Python і відповідного тарифного плану.

Моделі машинного навчання можна розробити двома способами: або з нуля, або використовуючи існуючі моделі, створені іншими. Microsoft пропонує низку попередньо навчених моделей, відомих під назвою Cognitive Services..

Ці моделі включаються розпізнавання зображень, розпізнавання мовлення і переклад між різними мовами. Face API реалізує модель машинного навчання аби знайти обличчя на зображенні. Для доступу до нього, необхідним є ключ підписки, який можна отримати на порталі Azure.

Міркування щодо вартості мають вирішальне значення для освітньої програми. Якщо використання API програми не перевищує 20 викликів на хвилину і менше 30 000 викликів на місяць, ви можете вибрати безкоштовний рівень F0, який дозволяє безкоштовно використовувати один екземпляр кожної служби Azure.

При створенні ресурсу важливо вибрати глобально унікальне ім'я. Це ім'я буде включено до URL-адреси, яка використовується для ідентифікації обличчя на зображенні. Крім того, рекомендується вибрати місце розташування, яке географічно найближче до місця, де ви плануєте запускати код. Це допоможе забезпечити оптимальну продуктивність і зменшити затримку.

Коли аутентифікація пройшла успішно, можна отримати особисті індивідуальні ключі та кінцеву точку для доступу до Face API. Новостворена кінцева точка дозволяє отримати доступ до Face API з різних програм, у тому числі методом. Дж. Беннетт не рекомендує зберігати облікові дані API безпосередньо в програмному коді і пропонує альтернативний підхід - зберігати їх у зовнішніх файлах типу .env. У процесі доопрацювання коду було зроблено

кілька суттєвих змін на стороні клієнта, роздільна здатність була збільшена до 1024 x 768 для покращення розпізнавання облич, Face API вимагає, щоб об'єкти мали мінімальний розмір 36 x 36 пікселів для точного розпізнавання, деталі пов'язані з перешкодами, виявились несумісними з іншими атрибутами, в результаті чого, були реалізовані кнопки для захоплення динамічних зображень, але їх використання прив'язане до обраного тарифного плану. Також слід зазначити, що безперервний доступ до Face API навіть до 15 разів за секунду, може вичерпати місячний ліміт протягом півгодини. Серверна частина також зазнала змін, в додаток до існуючої функції `best_emotion` була введена нова функція `best_color`, вона слугує для визначення кольору волосся людей на оброблених зображеннях, функція `check_result`, що обробляє натискання кнопки «Об'єкт», тепер використовує моделі розпізнавання `detection_03` та `recognition_04`, для отримання атрибутів маски і надання відповідної відповіді за кнопкою «Проаналізувати обличчя» відбудеться отримання таких атрибутів обличчя, як `'accessories'`, `'age'`, `'facialhair'`, `'gender'`, `'hair'`, `'makeup'` і звісно `'smile'`, `'emotion'`.

#### 4.4 Тестування розробленого прототипу

На рисунку 4.5 показано результати тестування системи в режимі розпізнавання масок. З чотирьох проведених тестів три були виконані правильно, а один - частково правильно, що свідчить про те, що система визначила, що рот не був закритий маскою. Проаналізувавши результати тестів і зробивши узагальнення, можна зробити наступні висновки щодо ідентифікації осіб за допомогою навчених моделей Face API:

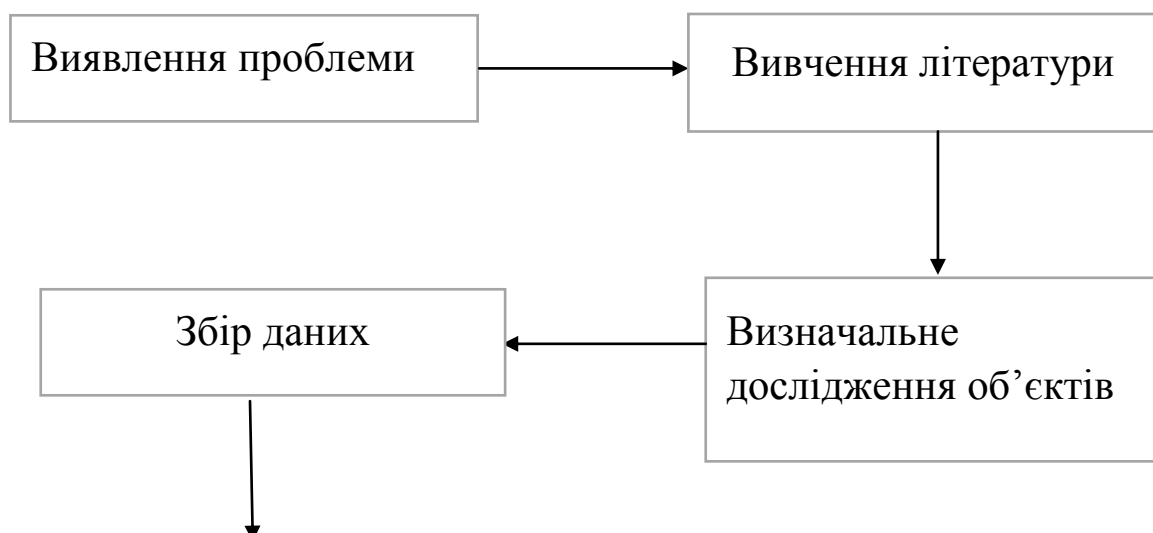
1. На сприйняття віку обличчя сильно впливає настрій людини. Оцінка системою видимого віку суттєво змінюється залежно від емоційного стану людини;
2. Світле волосся створює проблеми з ідентифікацією, оскільки воно може виглядати як "будь-який колір" залежно від умов освітлення. Через ці варіації системі важко точно визначити колір волосся у людей зі світлим

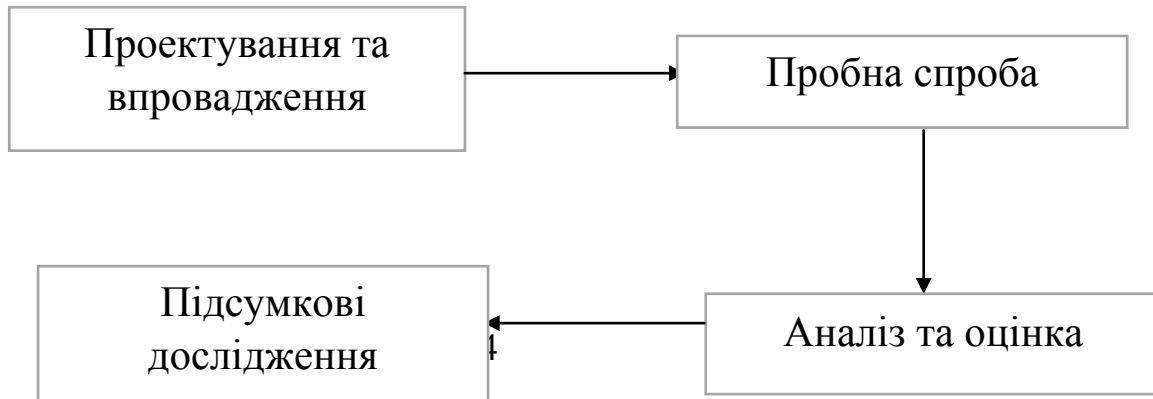
ВОЛОССЯМ.

3. Наявність окулярів для читання або сонцезахисних окулярів з додатковими декоративними елементами іноді може бути неправильно інтерпретована системою, помилково ідентифікуючи їх як окуляри для плавання. Це свідчить про те, що система може мати труднощі з розрізненням певних типів окулярів, особливо якщо вони мають унікальні елементи дизайну. Поширеною є помилкова ідентифікація макіяжу очей обличчя у сонцезахисних окулярах. Декількократне виявлення об'єктів (FSOD) має на меті підвищити продуктивність виявлення нових об'єктів за допомогою кількох позначених зразків. Щоб пом'якшити обмеження кількох зразків, ключову роль відіграє підвищення здатності до узагальнення вивчених ознак для нових об'єктів. Таким чином, процес вивчення ознак FSOD має більше зосереджуватися на внутрішніх характеристиках об'єкта, які є інваріантними при різних візуальних змінах і тому корисні для узагальнення ознак. На відміну від попередніх спроб парадигми метанавчання, у цій статті ми досліджуємо, як покращити властивості об'єкта за допомогою внутрішніх характеристик, які є універсальними для різних категорій об'єктів. Ми пропонуємо новий прототип, а саме універсальний прототип, який вивчається з усіх категорій об'єктів. Крім переваги характеристики незмінних характеристик, універсальні прототипи пом'якшують вплив незбалансованих категорій об'єктів. Після вдосконалення функцій об'єкта за допомогою універсальних прототипів ми втрачаємо узгодженість, щоб максимізувати узгодженість між розширеними функціями та вихідними, що є корисним для вивчення незмінних характеристик об'єкта. Таким чином, ми розробляємо нову структуру виявлення об'єктів із кількома кадрами з універсальними прототипами ( $\{FSOD\}^{\{up\}}$ ), яка має перевагу узагальнення ознак для нових об'єктів. Експериментальні результати на PASCAL VOC і MS COCO показують ефективність  $\{FSOD\}^{\{up\}}$ . Зокрема, для випадку 1-вибуху VOC Split2  $\{FSOD\}^{\{up\}}$  перевершує базовий рівень на 6,8% з точки зору mAP. ми застосовуємо втрату узгодженості, щоб максимізувати узгодженість між розширеними функціями та оригінальними, що є корисним для вивчення

незмінних характеристик об'єкта. Таким чином, ми розробляємо нову структуру виявлення об'єктів із кількома кадрами з універсальними прототипами ( $\{FSOD\}^{\{up\}}$ ), яка має перевагу узагальнення ознак для нових об'єктів. Експериментальні результати на PASCAL VOC і MS COCO показують ефективність  $\{FSOD\}^{\{up\}}$ . Зокрема, для випадку 1-вибуху VOC Split2  $\{FSOD\}^{\{up\}}$  перевершує базовий рівень на 6,8% з точки зору mAP. ми застосовуємо втрату узгодженості, щоб максимізувати узгодженість між розширеними функціями та оригінальними, що є корисним для вивчення незмінних характеристик об'єкта. Таким чином, ми розробляємо нову структуру виявлення об'єктів із кількома кадрами з універсальними прототипами ( $\{FSOD\}^{\{up\}}$ ), яка має перевагу узагальнення ознак для нових об'єктів. Експериментальні результати на PASCAL VOC і MS COCO показують ефективність  $\{FSOD\}^{\{up\}}$ . Зокрема, для випадку 1-вибуху VOC Split2  $\{FSOD\}^{\{up\}}$  перевершує базовий рівень на 6,8% з точки зору mAP. Експериментальні результати на PASCAL VOC і MS COCO показують ефективність  $\{FSOD\}^{\{up\}}$ . Зокрема, для випадку 1-вибуху VOC Split2  $\{FSOD\}^{\{up\}}$  перевершує базовий рівень на 6,8% з точки зору mAP. Експериментальні результати на PASCAL VOC і MS COCO показують ефективність  $\{FSOD\}^{\{up\}}$ . Зокрема, для випадку 1-вибуху VOC Split2  $\{FSOD\}^{\{up\}}$  перевершує базовий рівень на 6,8% з точки зору mAP.

Рисунок 4.5 – Схема робочого процесу дослідження





Що стосується бібліотек машинного зору, то існує кілька варіантів, кожен з яких має свої переваги та придатність для освітніх цілей. Серед них автономні бібліотеки, такі як OpenCV, є особливо придатними завдяки своєму відкритому коду, спільним рисам вільних ліцензій та сумісності з Python, що полегшує швидке створення прототипів та інтеграцію з нейронними мережами для задач розпізнавання.

З іншого боку, комерційні інструменти, такі як MVTec HALCON та Adaptive Vision Library, пропонують ширший спектр функціональних можливостей, тоді як viso.ai орієнтований на непрограмістів, дозволяючи їм розробляти системи машинного зору без глибоких знань програмування. Ці інструменти надають додаткові можливості та гнучкість, хоча часто і за певну ціну.

З точки зору розгортання, можна використовувати різні веб-платформи, при цьому найбільш підходящими варіантами є хмарні сервіси, що надаються Amazon Web Services, Google Cloud і Microsoft Azure. Хоча Microsoft Azure і обробка даних у хмарі розробника є платними опціями, вони пропонують доступ до безлічі попередньо навчених моделей машинного навчання, включаючи Cognitive Services, які можуть бути дуже корисними.

Мій прототип системи машинного зору, створений на основі книги "Happy, Sad, Angry Workshop" Дж. Беннета, інтегрує розпізнавання емоцій учасників освітнього процесу та виявлення порушень масок. Крім того, вона забезпечує надійну ідентифікацію статі, інтенсивності посмішки та окулярів, пропонуючи

помірно надійні оцінки віку, головного убору, макіяжу, кольору волосся та інших атрибутів.

## ВИСНОВКИ

У першому розділі я розглянув концепцію систем відеоспостереження та заглибився в теорію розпізнавання образів і машинного зору. Я також дослідив застосування штучного інтелекту у відеоспостереженні та проаналізував різні існуючі рішення, такі як DeepFace, Face ID від Apple, SmartGate та інші рішення, що використовуються урядами та компаніями по всьому світу. Було виявлено, що хоча ці рішення пропонують численні переваги, вони також мають недоліки. Одним із суттєвих недоліків є те, що жодна з цих систем не може досягти 100% точності через такі фактори, як умови освітлення, вираз обличчя, пози та шум під час захоплення обличчя, які можуть впливати на продуктивність систем розпізнавання облич.

Мета моєї роботи - розробити подібну міні-систему зі значно меншими матеріальними витратами. Малоімовірно, що нам вдасться створити систему з можливостями машинного навчання за допомогою нейронної мережі з нуля. Однак ми можемо спроектувати власний пристрій, який ефективно слугуватиме нашим цілям і може бути розширений для інших розумних систем.

У другому розділі мною були описані приклади та створені схеми доповненої реальності, враховуючи всі переваги та недоліки. Було проведено обґрунтування переваг Apple ARKit над Google ARKit, як показала практика, по якості фото знімків або відео зйомки, архітектура компанії Apple дійсно переважає над аналогами. В результаті огляду засобів для користування користувача можна дійти до висновку, що для звичайного користувача серед всіх бібліотек таких як ARKit, OpenCV, Wikitude, краще підходить ARKit.

Також розібрав види доповненої реальності, так як кожна впливає на життя людей і в яких сферах застосовується. В результаті аналізу можливостей використання найпопулярніших віртуальних реальностей, описаних вимог та потреб системи, зі значними перевагами було обрано мобільний додаток Apple ARKit. Було проведено огляд та порівняння описаних мною видів віртуальної реальності для користувацького інтерфейсу. Попередня обробка хмари відіграє

вирішальну роль у підготовці необроблених даних хмари для подальших процесів ідентифікації об'єктів. На цьому етапі використовуються різні методи, спрямовані на підвищення якості даних і отримання значущої інформації.

У третьому розділі я описав апаратні компоненти системи відеоспостереження. Ці компоненти включають контролер Arduino Leonardo, веб-камеру Logitech WebCam C270 і поворотний механізм на основі двох сервоприводів Tower Pro 9g SG90 і двовісного кронштейна. Перевагою такої установки є те, що всі компоненти системи можна легко розібрати і перепрофілювати для інших застосувань програмній складовій пристрою спостереження, цей пристрій дозволить мені виконувати такі завдання, як розпізнавання об'єктів на фото і відео (наприклад, ідентифікація обличч і людських фігур, розпізнавання тексту), відстеження об'єктів, класифікація дій на відео, конвертація зображень, вилучення 3D-моделей, створення 3D-простору з зображень стереокамер, синтез високоякісних зображень за допомогою злиття зображень, пошук об'єктів за заданим набором елементів, застосування методів машинного навчання, локалізація маркерів і виявлення спільних елементів на різних зображеннях.

В четвертому розділі мною було описано та створено модель, яка считує характеристики обличчя, незалежно від освітлення та кольору волосся, що показує емоційний стан людини, та чи присутні предмети на обличчі. Підсумовуючи, ідентифікація об'єктів у тривимірних хмарах точок за допомогою технологій комп'ютерного бачення має величезний потенціал у різних додатках. Завдяки використанню вдосконалених алгоритмів і методів ідентифікація об'єктів у 3D-хмарах дозволяє зрозуміти сцену, розпізнати об'єкт і просторовий аналіз. У цій роботі ми досліджуємо методи та засоби, що використовуються для ідентифікації об'єктів у тривимірних хмарах точок, а також розглядаємо проблеми та потенційні рішення. Розвиваючи новітні технології в цій галузі, дослідники можуть сприяти розробці більш точних і ефективних систем ідентифікації об'єктів, сприяючи прогресу в робототехніці, автономних системах і міському плануванні.



## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Abandoned Luggage 2021. URL: <https://viso.ai/application/abandoned-luggage-detection/> (Дата звернення: 19.04.2023).
2. Agarwal L. Mukim M., Sharma H., Bhandari A., Mishra A. Face Recognition Based Smart and Robust Attendance Monitoring using Deep CNN. *2021 8th International Conference on Computing for Sustainable Global Development (INDIACom)*. 2021. P. 699-704.
3. Ballard D. H., Dana H. BallKlingler N. Top 8 Applications of Computer Vision in the Education Sector 2021. – URL: <https://viso.ai/applications/computer-vision-in-education/> (Дата звернення: 25.04.2023)
4. Libraries comparison Adaptive Vision. 2021. URL: [https://docs.adaptivevision.com/av1/technical\\_issues/LibrariesComparison.html](https://docs.adaptivevision.com/av1/technical_issues/LibrariesComparison.html) (Дата звернення: 25.04.2023)
5. Mask Detection: Automatically detect unmasked people in public spaces or indoors 2021 URL: <https://viso.ai/application/mask-detection/> (Дата звернення: 25.04.2023)
6. Parking Lot Occupancy URL: <https://viso.ai/application/parking-lot-occupancy-detection/> (Дата звернення: 22.04.2023)
7. Prince S. J. D. Computer Vision: Models Learning, and Inference. *Cambridge University Press*. 2012. P. 331-351.
8. Juliet R. C. Pulliam Cari van Schalkwyk Nevashan Govender Anne von Gottberg Cheryl Cohen Michelle J. Groome Jonathan Dushoff Koleka Mlisana Harry Moultrie Pulliam 6J. R. C. Increased risk of SARS-CoV-2 reinfection associated with emergence of the Omicron variant in South Africa. 2021. P. 251.

9. Raj A. Smart Attendance Monitoring System with Computer Vision Using IOT *Imteyaz Ahmad Journal of Mobile Multimedia*. 2021. Vol. 17(1-3). P. 115-125.
10. Rezaei M. DeepSOCIAL: *Social Distancing Monitoring and Infection Risk Assessment in COVID-19 Pandemi*. *Applied Sciences*. 2020. Vol. 10, no. 21. P. 144.
11. Roberts L. G. Machine perception of three-dimensional solids: Thesis (Ph. D.) Lawrence Gilman Roberts. – Massachusetts Institute of Technology, Dept. of Electrical Engineering, 2013. URL: <https://dspace.mit.edu/bitstream/handle/1721.1/11589/> (Дата звернення: 25.04.2023)
12. Connor Shorten Taghi M. Khoshgoftaar Borko Furht Shorten C. Deep Learning applications for COVID-19. *Journal of Big Data*. 2021. Vol. 8. Article 18. P. 145.
13. Sivakumar S. A. T. J. John G. T. Selvi B. Madhu C. U. Shankar K. P. Arjun IoT based Intelligent Attendance Monitoring with Face Recognition Scheme. *5th International Conference on Computing Methodologies and Communication (ICCMC)*. 2021. P. 349-353.
14. Social Distancing Monitoring 2021 URL: <https://viso.ai/application/social-distancing-monitoring/> (Дата звернення: 29.04.2023)
15. Sutherland I. E. Sketchpad a man-machine graphical communication system. *Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy Ivan Edward Sutherland*; Massachusetts Institute of Technology Department of Electrical Engineering. 2019. 176 p. URL: [http://images.designworldonline.com.s3.amazonaws.com/CADhistory/Sketchpad\\_A\\_Man-Machine\\_Graphical\\_Communication\\_System\\_Jan63.pdf](http://images.designworldonline.com.s3.amazonaws.com/CADhistory/Sketchpad_A_Man-Machine_Graphical_Communication_System_Jan63.pdf) (Дата звернення: 29.04.2023)
16. Tkachuk V. Using Mobile ICT for Online Learning During COVID-19 Lockdown. *Information and Communication Technologies in Education*,

*Research, and Industrial Applications. 16th International Conference, ICTERI 2020, Kharkiv, Ukraine, October 6–10, 2020, Revised Selected Papers / Editors Communications in Computer and Information Science. – Cham : Springer, 2021. – Vol. 1308. – P. 46–67.*

17. Vision API Product Search pricing Google Cloud. 2021. (Дата звернення: 20.04.2023) URL: <https://cloud.google.com/vision/product-search/pricing>

18. Про реалізацію експериментального проекту щодо запровадження першої черги Єдиної державної електронної системи у сфері будівництва: Постанова Кабінету Міністрів України від 01.07.2020 р. № 559. URL: <https://zakon.rada.gov.ua/laws/show/559-2020-%D0%BF#Text> (Дата звернення: 14.04.2023)

19. Склад та зміст містобудівного кадастру – К.: Мінрегіон України, 2018. – 57 с.

20. Географічна інформація. Просторова прив'язка за географічними ідентифікаторами: ДСТУ ISO 19112:2017 (ISO 19112:2003, IDT). К: ДП «УкрНДНЦ»

21. AWS vs. Azure vs. Google: Cloud. URL: <https://www.datamation.com/cloud-computing/aws-vs-azure-vs-google-cloudcomparison.html> (дата звернення: 25.04.2023).

22. Olszewski R. Methodology of creating the new generation of official topographic maps in Poland. *Proceedings of the 26th ICC2013*, Dresden, Germany. (Дата звернення: 29.04.2023) URL: [https://icaci.org/files/documents/ICC2013/\\_extendedAbstract/248\\_proceeding.pdf](https://icaci.org/files/documents/ICC2013/_extendedAbstract/248_proceeding.pdf)

23. Open Location Code: An Open Source Standard for Addresses, Independent of Building Numbers And Street Names URL: [https://github.com/google/open-location-code/blob/master/docs/olc\\_definition.adoc](https://github.com/google/open-location-code/blob/master/docs/olc_definition.adoc). (Дата звернення: 29.04.2023)

24. Pyke C.R. Breaking barriers to interoperability: assigning spatially and temporally unique identifiers to spaces and buildings C.R. Pyke I. Madan *Annals of the New York Academy of Sciences, Issue: The implications of a Data Driven-Built Environment*. 2013. P. 29-53.

25. UUID (Universally Unique Identifier) [Электронный ресурс]. URL: <https://uk.wikipedia.org/wiki/UUID>. (Дата звернення: 25.04.2023)

26. Wang N. Unique Building Identifier: A natural key for building data matching and its energy applications. *Energy Build*. 2019, P. 184–241. URL: <https://doi.org/10.1016/j.enbuild.2018.11.052>. (Дата звернення: 23.04.2023)

27. what3words These examples show how to display a what3words address. (Дата звернення: 25.04.2023). URL: <https://developer.what3words.com/design/what3words-address#overview>

28. Core ML. URL: <https://developer.apple.com/documentation/coreml> (Дата звернення: 25.04.2023).

29. CocoaTouch. URL: <https://developer.apple.com/library/archive/documentation/General/Conceptual/DevPedia-CocoaCore/Cocoa.html> (дата звернення: 25.04.2021).

30. Alamofire. URL: <https://github.com/Alamofire/Alamofire> (дата звернення: 25.04.2021).

31. AWS vs. Azure vs. Google: Cloud. URL: <https://www.datamation.com/cloud-computing/aws-vs-azure-vs-google-cloudcomparison.html> (дата звернення: 25.04.2024).

32. Ian Goodfellow Yoshua Bengio Deep Learning (*Adaptive Computation and Machine Learning series*) *The MIT Press* 2016, p. 621-638.

33. Frank Millstein Deep Learning: 2 Manuscripts *Deep Learning With Keras And Convolutional Neural Networks In Python*, Paperback p.117-129, 2018

34. Apple Developer Documentation: Creating Core ML. URL: <https://developer.apple.com/documentation/coreml> (Дата звернення: 25.04.2023).
35. Apple Developer Documentation: Creating Core ML. URL: <https://developer.apple.com/machine-learning/> (Дата звернення: 25.04.2023).
36. Lawrence J. Introduction to neural networks: design, theory and applications. *California Scientific Software*. 2014. P. 235.
37. Duda R. O Hart P. E D. G. Stork. Pattern classification. Wiley, 2001. P. 502.
38. Cybenko G. V. Approximation by Superpositions of a Sigmoidal function. 2006. 314 с.
39. Krizhevsky A. Sutskever I. Advances in Neural Information Processing Systems. 2012. 1097 с.
40. Hubel D. H. Wiesel D.H. Brain and visual perception: the story of a 25-year collaboration. Oxford University, 2005. 106 с.
41. Meier U. Ciresan D. Multi-column deep neural networks for image classification. New York. 649 с.
42. Visual Geometry. URL: [http://www.robots.ox.ac.uk/~vgg/research/very\\_deep11/](http://www.robots.ox.ac.uk/~vgg/research/very_deep11/). (Дата звернення: 23.04.2023)
43. Visual Geometry. URL: [http://www.robots.ox.ac.uk/~vgg/research/very\\_deep12/](http://www.robots.ox.ac.uk/~vgg/research/very_deep12/). (Дата звернення: 23.04.2023)
44. Belongie S. Wilber M. Viet. A Residual Networks Behave Like Ensembles of Relatively Shallow Networks. 2016, p 107-113.
45. Benchmark Analysis of Representative *Deep Neural Network Architectures*. URL: <https://arxiv.org/pdf/1810.00736.pdf> (дата звернення: 15.04.2021)
46. Richard Hartley Andrew Zisserman "*Multiple View Geometry in Computer Vision*" (2003), Cambridge University Press, P. 674.

47. Erik Solem Programming Computer Vision with Python: Tools and algorithms for analyzing images. 2012. O'Reilly Media P. 408.
48. Kevin P. Murphy Machine Learning: A Probabilistic Perspective 2012. MIT Press, P. 1067.
49. Aurélien Géron Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. 2019. O'Reilly Media, P. 745.
50. Sebastian Raschka and Vahid Mirjalili Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow . 2017. Packt Publishing. P. 612.
51. Gabriel Garrido Calvo Prateek Joshi Michael Beyeler Practical OpenCV 3 Image Processing with Python. 2018. Apress P. 231.
52. Alberto Fernandez Villan Mastering OpenCV 4 with Python: A practical guide covering topics from image processing, augmented reality to deep learning with OpenCV 4 and Python 3.7. 2019. Packt Publishing, P. 401.
53. Joseph Howse Joe Minichino Prateek Joshi Learning OpenCV 4 Computer Vision with Python 3: Get to grips with tools, techniques, and algorithms for solving real-world computer vision problems with OpenCV 4. 2019. Packt Publishing, P. 494.
54. Rami M. S. Elbakoury Mohamed E. Hussein Convolutional Neural Networks in Visual Computing: A comprehensive guide to CNN architectures, learning strategies, and applications. CRC Press. 2019. P. 298.
55. Rawat D. Rodrigues J. Stojmenovic I. Sanfelice R.G. Analysis and Design of Cyber-Physical Systems. A Hybrid Control Systems Approach. *Cyber-Physical Systems: From Theory to Practice*. CRC Press. 2016. ISBN 978-1-4822-6333-6.
56. Fitz T. Theiler M. Smarsly K. A metamodel for cyberphysical systems. *Advanced Engineering Informatics*. 2019. V. 41. Article 100930.
57. Khaitan S. K. McCalley J. D. Design techniques and applications of cyberphysical systems. *A survey IEEE Systems Journal*. 2014. № 9(2). P. 350-

365. URL: <https://doi.org/10.1109/JSYST.2014.2322503>. (Дата звернення: 25.04.2023).

58. Lee Ming-Chang. Software Quality Factors and Software Quality Metrics to Enhance Software Quality Assurance. *British Journal of Applied Science & Technology*. 2014 № 4. P.10.

59. Zhang W. Asiri A. M. Liu D. Nanomaterial-Based Biosensors for Environmental and Biological Monitoring of Organophosphorus Pesticides and Nerve Agents . *TrAC Trends in Analytical Chemistry*. 2014. P. 1–10.

60. Ma H. Internet of Things: Objectives and Scientific Challenges. *Journal of Computer Science and Technology*. 2011. № 26(6). P. 919-924. URL: <https://doi.org/10.1007/s11390-011-1189-5>.

61. Regnier P. Lima G. Massa E. Multiprocessor scheduling by reduction to uniprocessor: an original optimal approach. *Real-Time Syst.* 2013. № 49. С. 436–474. URL: <https://doi.org/10.1007/s11241-012-9165-x>.

62. Furugyan M.G. Scheduling in Multiprocessor Systems with Additional Restrictions. *J. Comput. Syst. Sci. Int.* 2018. № 57. С. 222–229. URL: <https://doi.org/10.1134/S1064230718020077>.

63. Pushkar O. Hrabovskyi Y. Methodology for developing an intelligent user interface for educational publications in the e-learning system. *Development Management*. 2019. V. 17. № 3. P. 23-34

64. Олесків О. Вимірювальна техніка та метрологія. Міжвідомчий науково-технічний збірник. *Видавництво Національного університету «Львівська політехніка»*. 2015. № 76. С. 132– 137.

65. Shorten C. Deep Learning applications for COVID-19 / Connor Shorten Taghi M.Khoshgoftaar Borko Furht . *Journal of Big Data*. 2021. Vol. 8. – Article 18.

66. Rezaei M. DeepSOCIAL: *Social Distancing Monitoring and Infection Risk Assessment in COVID-19 Pandemic*. *Applied Sciences*. 2020. Vol. 10 no. 21. – Article 7514. P.103.

67. Pulliam J. R. C. Increased risk of SARS-CoV-2 reinfection

associated with emergence of the Omicron variant in South Africa. 2021. P. 711.

68. Klingler N. Top 8 Applications of Computer Vision in the Education Sector. 2021. URL: <https://viso.ai/applications/computer-vision-in-education/>.

69. Vision API Product Search pricing. Google Cloud. 2021. URL: <https://cloud.google.com/vision/product-search/pricing>

70. Inception-v3. URL: <https://medium.com/@sh.tsang/review-inception-v3-1st-runner-up-image-classification-in-ilsvrc-2015-17915421f77c>

(дата звернення: 15.04.2021).

71. Belongie S. Wilber M. Viet. A Residual Networks Behave Like Ensembles of Relatively Shallow Networks. 2016.

72. CocoaTouch. URL: <https://developer.apple.com/library/archive/documentation/General/Conceptual/DevPedia-CocoaCore/Cocoa.html> .

73. Machine Learning Proceedings 1991: Proceedings of the Eighth International Workshop (ML91). Elsevier Science. 2014. 364 с.

74. UUID (Universally Unique Identifier). Retrieved from URL: <https://uk.wikipedia.org/wiki/UUID>.

75. An evaluation of Location Encoding Systems.– URL: <https://github.com/google/open-location-code/wiki/Evaluation-of-Location-Encoding-Systems>.ard, Christopher M. Brown. Englewood Cliffs : Prentice Hall, Computer Vision 2022. URL: <https://archive.org/details/computervision0000ball> (Дата звернення: 19.04.2023).

76. Bennett J. Jim Bennett. Happy, Sad, Angry Workshop 2020. URL: <https://github.com/jimbo Bennett/HappySadAngryWorkshop> (Дата звернення: 19.04.2023).

77. Face Recognition: URL: <https://viso.ai/application/face-recognition/> (Дата звернення: 22.04.2023).

78. Facial Emotion Analysis 2021 URL: <https://viso.ai/application/emotion-analysis/> (Дата звернення: 22.04.2023).

79. Gibson J. J. The Perception of the Visual World. Boston : Houghton Mifflin, 2020.
80. Google Ngram Viewer Stanford University 2021. URL: <https://books.google.com/ngrams/graph?content=computer+vision%2C+machine%20+vision> .
81. Grape G. R. Model Based (Intermediate-Level) Computer Vision : PhD Dissertation / Gunnar Rutger Grape. 2010. URL: <https://apps.dtic.mil/sti/pdfs/AD0763673.pdf> .
82. Intrusion Detectio – URL: <https://viso.ai/application/intrusion-detection/> .

## ДОДАТОК А

### Клієнтська частина прототипу (/templates/home.html)

```
<!DOCTYPE html>
<html>
  <head>
    <meta charset="UTF-8">
    <title>Машинний зір у часи пандемії</title>
  </head>
  <body>
    <video id="video" autoplay></video>
    <br/>
    <button id="capture">Чи я у масці?</button>
    <button id="capture2">Проаналізувати обличчя</button>
    <p id="message"></p>

    <script type="text/javascript">
      window.addEventListener("DOMContentLoaded", function() {
        var video = document.getElementById('video');

        if (navigator.mediaDevices &&
navigator.mediaDevices.getUserMedia) {
          const getImage = async () => {
            video.srcObject = await
navigator.mediaDevices.getUserMedia({ video: true })
            video.play();
          }
          getImage()
        }
      })

      var message = document.getElementById('message');
```

```

document.getElementById('capture').addEventListener('click', function() {
    var canvas = document.createElement('canvas');

    canvas.width = 1024;
    canvas.height = 768;

    var context = canvas.getContext('2d');
    context.drawImage(video, 0, 0, canvas.width,
canvas.height);

    var data = {
        'image_base64': canvas.toDataURL("image/png"),
    }

    const getResult = async () => {
        var result = await fetch('result', {
            method: 'POST',

            body: JSON.stringify(data),
            headers: { 'Content-Type': 'application/json' }
        })

        var jsonResult = await result.json()
        message.textContent = jsonResult.message

    }

    getResult()
});

```

```

document.getElementById('capture2').addEventListener('click', function() {var
    canvas = document.createElement('canvas');
    canvas.width = 1024;
    canvas.height = 768;

    var context = canvas.getContext('2d');
    context.drawImage(video, 0, 0, canvas.width
canvas.height);

    var data = {

```

```
        'image_base64': canvas.toDataURL("image/png"),
    }
    const getResult = async () => {
        var result = await fetch('result2', {
            method: 'POST',
            body: JSON.stringify(data),
            headers: { 'Content-Type': 'application/json'
        })
        var jsonResult = await result.json()
        message.textContent = jsonResult.message
    }
    getResult()
});
})
</script>
</body>
</html>
```

**ДОДАТОК В**  
**ТЕЗИ ДО ДИПЛОМНОЇ РОБОТИ**

Міністерство освіти і науки України  
Хмельницький національний університет



**ЗБІРНИК НАУКОВИХ ПРАЦЬ**  
за матеріалами XIV Всеукраїнської науково-практичної конференції  
«Актуальні проблеми комп'ютерних наук АПКН-2022»

*18-19 листопада 2022*

Хмельницький 2022

УДК 004.9

Корольков В.О., Табенський С. М., Свистун С.О., Мельник В.В., Жуковський П.О.

*Хмельницький національний університет*

### МЕТОД ТА ЗАСОБИ ІДЕНТИФІКАЦІЇ ОБ'ЄКТІВ У ТРИВИМІРНИХ ХМАРАХ ТЕХНОЛОГІЯМИ КОМП'ЮТЕРНОГО ЗОРУ ТА МАШИННОГО НАВЧАННЯ

*Важливим є класифікація 3D-об'єктів. У роботі розроблено надійну модель сферичних гармонік для класифікації об'єктів, що підтвердив експеримент. Запропонована модель використовує воксель-сітку концентричних сфер, щоб дізнатися особливості одиничної кулі. Крім того, зберігаються операції згортки в області Фур'є без застосування зворотного перетворення, що використовується в попередніх підходах. В результаті модель здатна до навчання функції, які менш чутливі до неточностей даних. Результати тестування показують, що запропонована модель перевершує сучасні мережі з точки зору стійкості до наслідків неточностей даних.*

*Classification of 3D objects is important. The work developed a reliable model of spherical harmonics for the classification of objects, which was confirmed by the experiment. The proposed model uses a voxel mesh of concentric spheres to learn the features of a unit sphere. In addition, convolution operations in the Fourier domain are preserved without applying the inverse transformation used in previous approaches. As a result, the model is capable of learning features that are less sensitive to data inaccuracies. The test results show that the proposed model outperforms modern networks in terms of resistance to the consequences of data inaccuracies.*

Ідентифікація об'єктів у тривимірних хмарах точок є важливим завданням комп'ютерного зору. Методи, які використовують останні досягнення в глибокому навчанні, досягли значних результатів щодо ефективності. Тривимірні набори вимірювань (хмари точок) створені 3D-сканерами часто порушуються шумом, викидами або відсутніми точками. Експерименти показали, що наявність цих неточностей може мати значний вплив на продуктивність алгоритмів глибокого навчання для класифікації об'єктів. Пропоноване дослідження має на меті сприяти встановленню впливів цих неточностей вимірювання на завдання ідентифікації об'єкта та знайти нові алгоритми класифікації об'єктів, стійкі до впливу цих збурень. Щоб прокласти шлях до відкриття надійних методів класифікації об'єктів, було проведено широкий аналіз продуктивності найкращих доступних алгоритмів розпізнавання об'єктів на даних із зазначеними вище неточностями. Результати показали, що продуктивність більшості методів погіршується навіть за низького збурення на різних рівнях. Результат чітко виправдав дослідження з метою вдосконалення поточної класифікації мережі та запропонувати ті, які стійкі до цих неточностей даних.

Метою роботи є розробка методу ідентифікації об'єктів у тривимірних хмарах точок.

На відміну від попередніх підходів сферичної гармоніки, запропонована структура поєднує в собі наступні три ключові внески: використання концентричних сфер, пряме використання величин коефіцієнтів сферичних гармонік для класифікації об'єктів та обмеження рівня порядку сферичних гармонік для пом'якшення ефекту шуму/викидів для досягнення надійної класифікації. Щоб описати метод, спочатку представити проблему. Хмари точок 3D-моделей, створених або 3D-сканерами, або алгоритмами 3D-реконструкції, часто недосконалі та містять викиди. Щоб перевірити міцність попереднього на таких сценаріях записано 12 сцен за допомогою сканера і він містить такі загальні предмети: стільці, столи, шафи, дивани та столи. Категорії записаних об'єктів були обрані в набір даних, щоб навчати методи на основі даних під час тестування захоплених даних.

Пропонується застосування сферичної згортки на 3D моделях, які декомпонуються в концентричні сфери. Використання концентричних сфер створює рівномірну сферичну сітку, яка дозволяє нейронній мережі сферичної згортки вивчати функції над одиничною кулею (на відміну від навчання лише на одиничній сфері). Використовуємо окремі операції згортання швидкості в кожній концентричній сфері, щоб дозволити мережі навчатися особливості, що стосуються цієї сфери.

При проведенні експерименту порівнювалась розроблена структура з найсучаснішими опублікованими архітектурами сферичної згортки, методами класифікації хмар точок і надійними методами. Було розглянуто викиди, шум і відсутні точки як наші типи неточностей даних у роботі, оскільки пошкодження хмар точок із такими неточностями є звичайним явищем.

Класифікація 3D-об'єктів є важливим завданням у кількох роботах. Таким чином, розроблено надійну модель сферичних гармонік для класифікації об'єктів, що підтвердив експеримент. Запропонована модель використовує воксель-сітку концентричних сфер, щоб дізнатися особливості одиничної кулі. Крім того, зберігаються операції згортки в області Фур'є без застосування зворотного перетворення, що використовується в попередніх підходах. В результаті модель здатна до навчання функції, які менш чутливі до неточностей даних. Результати тестування показують, що запропонована модель перевершує сучасні мережі з точки зору стійкості до наслідків неточностей даних.

#### Перелік посилань

1. D. Dimou and K. Moustakas. Fast 3D Scene Segmentation and Partial Object Retrieval Using Local Geometric Surface Features, pages 79–98. Springer Berlin Heidelberg, Berlin, Heidelberg, 2020. ISBN 978-3-662-61364-1. doi: 10.1007/978-3-662-61364-1 5. URL <https://doi.org/10.1007/978-3-662-61364-1>
2. J. R. Driscoll and D. M. Healy. Computing fourier transforms and convolutions on the 2-sphere. *Advances in applied mathematics*, 15(2):202–250, 1994.
3. C. Esteves, C. Allen-Blanchette, A. Makadia, and K. Daniilidis. Learning so (3) equivariant representations with spherical cnns. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 52–68, 2018.
4. A. Filgueira, H. Gonz'alez-Jorge, S. Lag'uela, L. D'íaz-Vilari'no, and P. Arias. Quantifying the influence of rain in lidar performance. *Measurement*, 95:143–148, 2017.

## *«Метод та засоби ідентифікації об'єктів у тривимірних хмарах технологіями комп'ютерного зору»*

*АВТОР РОБОТИ:*

*СТ. ГР. КІЗМ-21-1 КОРОЛЬКОВ В. С.*

*КЕРІВНИК РОБОТИ:*

*Д. Т. Н. ДОЦЕНТ:*

*БЕРЕЗЬКА К. М.*

Хмельницький 2023

## АКТУАЛЬНІСТЬ

Наш світ, це світ сучасних технологій, які розвиваються з кожним днем дуже швидко. Дана тема дуже актуальна у багатьох галузях життя. А саме в архітектурі, медицині та навіть у військовій сферах. За допомогою 3D технології ідентифікації об'єктів, ми можемо змоделювати до прикладу квартиру своєї мрії та ніби побувати в середині неї. Можемо змоделювати протези для наших поранених героїв. Удосконалити наші безпілотні літальні апарати, завдяки яким ми невпинно наближаємо нашу перемогу. Розвиток технологій віртуальної реальності дає нам змогу поринати у світ фантастики, при цьому відчуваючи почуття максимально наближені до реальних.

- **МЕТА ДОСЛІДЖЕННЯ – В РОБОТІ ДОСЛІДЖЕНІ ОСНОВНІ ЕТАПИ СТВОРЕННЯ ПРИСТРОЮ, ЩО МАЄ РОЗПІЗНАВАТИ ОБЛИЧЧЯ І ФОКУСУВАТИСЬ НА НИХ ЗА ДОПОМОГОЮ КОЛЬОРОВОЇ РАМКИ. ПРО КОЛЬОРИ РАМОК ТА ЯК РОЗУМІТИ ЧИ З ФОКУСУВАВСЯ ПРИСТРІЙ НАМ СКАЖЕ КОЛІР РАМКИ.**
- **ІНСТРУМЕНТ ДОСЛІДЖЕННЯ – ДЛЯ ЗАВАНТАЖЕННЯ І НАПИСАННЯ ПРОГРАМ НА СУМІСНІ ПЛАТИ ВИКОРИСТОВУЄТЬСЯ БІБЛІОТЕКА З ОБРОБКОЮ ЗОБРАЖЕННЯ ТА ВЕБ ПРИСТРОЯМИ, OPENCV, КОНТРОЛЕР ARDUINO LEONARDO ВИСТУПАЄ ІНСТРУМЕНТОМ ДОСЛІДЖЕННЯ.**

## **ОТРИМАНІ НАУКОВІ ТА ПРАКТИЧНІ РЕЗУЛЬТАТИ**

**В РОБОТІ БУЛА ВИКОРИСТАНА ТЕХНОЛОГІЇ КОМП'ЮТЕРНОГО ЗОРУ ДЛЯ ЗАДАЧ ДЕТЕКЦІЇ ТА РОЗПІЗНАННЯ ЛЮДСЬКИХ ОБЛИЧ, ТА РОЗРОБЛЕНО МЕТОД ДЛЯ ІДЕНТИФІКАЦІЇ ОБ'ЄКТІВ В ТРИВИМІРНОМУ ПРОСТОРІ.**

**БІБЛІОТЕКА OPEN CV ЗАБЕЗПЕЧУЄ ОБРОБКУ ЗОБРАЖЕНЬ В РЕЖИМІ РЕАЛЬНОГО ЧАСУ. ТАКОЖ ЗАПРОПОНОВАНО ПРОГРАМНО-АПАРАТНИЙ ЗВ'ЯЗОК ПРОАНАЛІЗОВАНОЇ ВХІДНОЇ ІНФОРМАЦІЇ З МЕХАНІЧНОЮ ЧАСТИНОЮ САМОХІДНОГО ПРИСТРОЮ, ЯКИЙ ДОЗВОЛЯЄ ЙОМУ САМОСТІЙНО ЗМІНЮВАТИ СВОЄ ПОЛОЖЕННЯ, ОСКІЛЬКИ ОБРОБКА ЗОБРАЖЕННЯ ВЕДЕТЬСЯ НА КОМП'ЮТЕРІ, А УПРАВЛІННЯ СЕРВОПРИВОДАМИ ЗДІЙСНЮЄТЬСЯ ЗА ДОПОМОГОЮ КОНТРОЛЕРА.**

## ЗАДАЧІ ДОСЛІДЖЕННЯ:

- АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧІ;
- ДОСЛІДИТИ АРХІТЕКТУРУ ДИСТРИБУТИВНОЇ СИСТЕМИ;
- ШТУЧНІ ІМУННІ СИСТЕМИ ТА НЕЙРОННІ МЕРЕЖІ;
- РЕАЛІЗАЦІЯ АПАРАТНО-ПРОГРАМНОГО КОМПЛЕКСУ ДЛЯ ВИКОНАННЯ НАВИГАЦІЇ І РОЗПІЗНАВАННЯ ЛЮДСЬКИХ ОБРАЗІВ ТА РЕЗУЛЬТАТИ.

## Ідентифікації об'єктів та її види

В СУЧАСНІЙ ІТ-ІНДУСТРІЇ РЕАЛІЗОВАНО БАГАТО ПІДХОДІВ ТА ЗАСОБІВ ДО ФОРМУВАННЯ ГЛОБАЛЬНИХ ІДЕНТИФІКАТОРІВ ОБ'ЄКТІВ РІЗНОГО ПРИЗНАЧЕННЯ ТА ПРЕДМЕТНИХ СФЕР ЗАСТОСУВАННЯ, ТОЧНА ІДЕНТИФІКАЦІЯ ОБ'ЄКТІВ У ТРИВИМІРНИХ ХМАРАХ МАЄ ХНАЧНІ ПЕРЕВАГИ ДЛЯ РІЗНИХ ОБЛАСТЕЙ, ЗОКРЕМА РОБОТОТЕХНІКИ АВТОНОМНОЇ НАВИГАЦІЇ, ДОПОВНЕНОЇ РЕАЛЬНОСТІ ТА МІСЬКОГО ПЛАНУВАННЯ.

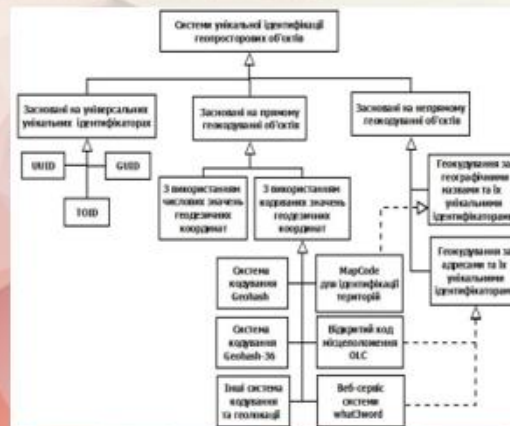
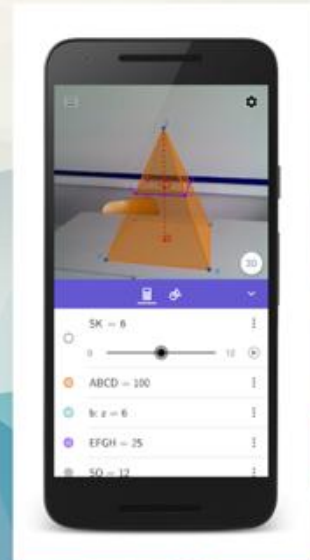


Рисунок 1 - Класифікація систем унікальної ідентифікації геопросторових об'єктів

## ДОПОВНЕНА РЕАЛЬНІСТЬ ТА ЇЇ ВИДИ

*Доповнена реальність – це магія. Це результат введення в поле сприйняття будь-яких сенсорних даних з метою доповнення відомостей про оточення і поліпшення сприйняття інформації, її специфіка полягає в тому, що програмним чином візуально поєднує два спочатку незалежним до простору таких як ,світ реальних об'єктів навколо нас та віртуальний світ, відтворений на комп'ютері. Нове віртуальне середовище утворюється шляхом накладення запрограмованих віртуальних об'єктів на поверхню відеосигналу з камери і стає інтерактивним шляхом, до використання спеціальних маркерів. Доповнена реальність вже роками використовується в медицині, в рекламній галузі, в іграх, для моніторингу об'єктів, військових технологіях і мобільних пристроях.*



*Рисунок 2 - Доповнена реальність на основі маркера*

## АПАРАТНА ЧАСТИНА ІДЕНТИФІКАТОРА ОБ'ЄКТІВ

*Для ідентифікації об'єкту, в роботі було взято контролер Arduino Leonardo, це контролер на базі ATmega32u4, платформа має 20 цифрових входів та виходів, кварцовий генератор 16 МГц, роз'єм мікро-USB, силовий роз'єм, роз'єм ICSP і кнопку перезавантаження.*

*Цей контролер надає ряд можливостей для здійснення зв'язку з комп'ютером.*

*Для виконання цілей, які були наперед закладені в дану систему, камера повинна повертатися по горизонталі. Для цього було використано двоохосьовий кронштейн.*

*Сервопривід, здатен повертати вихідний вал в задане положення і утримувати його в цьому положенні, в супереч опорам, це потрібно в першу чергу моделістам, для управління положеннями різних динамічних об'єктів*

### ТЕСТУВАННЯ СИСТЕМИ В СТАТИЧНОМУ ТА ДИНАМІЧНОМУ РЕЖИМАХ

Мікрофремворк *Flask* для створення веб-додатків встановлюється командою:

`Pip3 install flask`

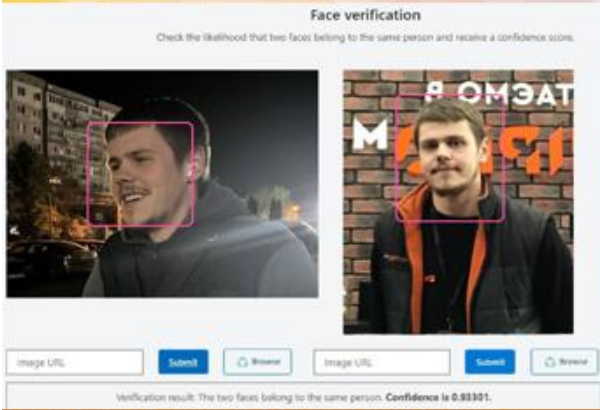


РИСУНОК 7 – РЕЗУЛЬТАТ ІДЕНТИФІКАЦІЇ СТАТИЧНОГО ОБЛИЧЧЯ.



Рисунок 8 – Результат ідентифікації обличчя у динаміці.

### ДОПОВНЕНА РЕАЛЬНІСТЬ ЗА ДОПОМОГОЮ ДОДАТКІВ



РИСУНОК 9 – ЗАСТОСУНОК APPLE ARKIT.



РИСУНОК 10 – GOOGLE ARCORE.

## ВИСНОВКИ

- У першому розділі мною було досліджено поняття систем спостереження, розглянуто теорію розпізнавання образів і машинного зору, досліджено поняття штучного інтелекту для відеоспостереження, а також проаналізовано існуючі рішення, а саме *DeepFace*, *Face ID* від компанії *Apple*, *SmartGate* а також ряд інших рішень, що застосовуються урядами і компаніями з різних країн
- У другому розділі було розглянуто приклади та схеми доповненої реальності, враховуючи всі переваги та недоліки. Було проведено обґрунтування переваг *Apple ARKit* над *Google ARKit*, як показала практика, по якості фото зніmkів або відео зйомки, архітектура компанії *Apple* дійсно переважає над аналогами. В результаті огляду засобів для користування користувача можна дійти до висновку, що для звичайного користувача серед всіх бібліотек таких як *ARKit*, *OpenCV*, *Wikitude*, краще підходить *ARKit*.
- У третьому розділі нами було описано складові елементи системи спостереження, що стосуються апаратної частини. Серед них контролер *Arduino Leonardo*, веб-камера *Logitech WebCam C270* та поворотний механізм для неї на основі двох сервоприводів *Tower Pro 9g SG90* та двохосового кронштейну.
- В четвертому розділі мною було описано та створено модель, яка зчитує характеристики обличчя, незалежно від освітлення та кольору волосся, що показує емоційний стан людини, та чи присутні предмети на обличчі і також ідентифікацію статичних об'єктів.

**ДЯКУЮ ЗА УВАГУ  
СЛАВА УКРАЇНІ!**



Ім'я користувача:  
Кафедра КІ

Дата перевірки:  
25.05.2023 10:42:52 EEST

Дата звіту:  
25.05.2023 10:53:15 EEST

ID перевірки:  
1015241972

Тип перевірки:  
Doc vs Internet + Library

ID користувача:  
100005591

Назва документа: Корольков\_2\_Метод та засоби ідентифікації об'єктів у тривимірних хмарах технологіями ко...

Кількість сторінок: 88 Кількість слів: 17525 Кількість символів: 134980 Розмір файлу: 1.71 MB ID файлу: 1014918034

## 2.44% Схожість

Найбільша схожість: 0.46% з джерелом з Бібліотеки (ID файлу: 1014815572)

2.2% Джерела з Інтернету

116

Сторінка 90

0.69% Джерела з Бібліотеки

100

Сторінка 91

## 0.12% Цитат

Цитати 10

Сторінка 92

Посилання 1

Сторінка 92

## 30.7% Вилучень

Деякі джерела вилучено автоматично (фільтри вилучення: кількість знайдених слів є меншою за 8 слів та 0%)

Немає вилучених Інтернет-джерел

30.7% Вилученого тексту з Бібліотеки

1

Сторінка 92

## Anti-Plagiarism v-15.257

**The maximum coincidence with one document 1.0%**

Dictionary check: en\_US, ru\_RU, ua\_UA. **Errors in the documents: 13%**

ID: 113694 Title: МКР Метод та засоби ідентифікації об'єктів у тривимірних хмарах технологіями комп'ютерного зору Added in a DB: 2023-05-19 Authors: В.О Корольков Heads: К.М Березька Consultants: Opponents:	Document		Sum coincidence on the DB	
	Symbols	Lexemes	Symbols	Lexemes
	91981	704	2789 (3%)	41 (6%)

### Plagiarism sources

ID	Description	Plagiarism presence in the document	
		Symbols	Lexemes

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

РЕЦЕНЗІЯ НА ДИПЛОМНУ РОБОТУ

Дипломник: Корольков Владислав Олександрович

Тема: Метод та засоби ідентифікації об'єктів у тривимірних хмарах технологіями комп'ютерного зору та машинного навчання

Спеціальність: 123 «Комп'ютерна інженерія»

Обсяг дипломної роботи:

Кількість листів креслень —; кількість сторінок записки 70

1. Короткий зміст роботи та прийнятих рішень У роботі запропонована технологія комп'ютерного зору для задач детекції та розпізнання та розпізнання людських облич за допомогою пристроїв зчитування інформації з об'єктів.

2. Висновок про відповідність роботи дипломному завданню \_\_\_\_\_  
Дипломна робота відповідає виданому завданню

3. Характеристика виконання кожного розділу, ступінь використання останніх досягнень науки і техніки і передових методів роботи: У першому розділі було розглянуто концепцію систем відеоспостереження та заглиблення в теорію розпізнавання образів машинного зору. У другому розділі були описані приклади та створені схеми доповненої реальності, враховуючи всі переваги та недоліки. Було проведено обґрунтування переваг Apple ARKit над Google ARCore. В третьому розділі були описані апаратні компоненти системи відеоспостереження за допомогою контролера Arduino Leonardo та веб-камери. В четвертому розділі була описана та створена модель, яка зчитує характеристики обличчя, незалежно від освітлення та кольору волосся, що показує емоційний стан людини та чи присутні предмети на обличчі.

4. Позитивні сторони роботи: Запропонована метод та засоби ідентифікації об'єктів виявились ефективними для використання у різних сферах та галузях життя.

5. Негативні сторони роботи: немає.

6. Оцінка графічного оформлення та пояснювальної записки роботи: —

7. Відгук про роботу в цілому: Робота виконана на належному рівні.

8. Інші зауваження: —

9. Оцінка дипломної роботи:

Розглянувши позитивні та негативні сторони представленої дипломної роботи вважаю, що робота заслуговує оцінки «3,0»

Рецензент (прізвище, ім'я, по батькові, посада, місце роботи) Мартинюк Валерій Володимирович, д.т.н., професор, завідувач кафедри АКІТР ХНУ

“     ”     2023р.



Завідувачу кафедри КПС  
д-р.техн.наук, проф. Говорушенко Т. О.

  
ПІБ здобувача вищої освіти

ФІТ, 2 курсу, групи КІ2М-20-1

### ЗАЯВА

З правилами чинного Положення «Про дотримання академічної доброчесності в Хмельницькому національному університеті» від 26.09.2020 (зі змінами від 26.11.2020), згідно з яким виявлення плагіату є підставою для відмови в допуску кваліфікаційної роботи до захисту та застосування заходів дисциплінарної та академічної відповідальності, ознайомлений (а). Про використання програмно-технічних засобів для перевірки кваліфікаційних робіт здобувачів вищої освіти на плагіатоповіщений (а) та надаю свою згоду на обробку та збереження університетом моєї роботи в інституційному репозитарії університету.

Також надаю університету право на передачу моєї роботи для обробки та збереження в базах даних програмно-технічних засобів (Unicheck та Anti-Plagiarism) та використання роботи для виявлення плагіату в інших роботах, які перевіряються програмно-технічними засобами та користувачами, що мають доступ до цих програмно-технічних засобів, виключно в обмежених цілях для виявлення плагіату в текстах робіт.

Робота для перевірки університетом надається в друкованому та електронному варіанті. Електронна версія моєї роботи збігається (ідентична) з друкованою.

29.05.2023

дата

  
підпис

РІШЕННЯ ЕКСПЕРНОЇ КОМІСІЇ  
КАФЕДРИ КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ ТА ІНФОМАЦІЙНИХ СИСТЕМ  
ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ ДО ЗАХИСТУ

Підтверджуємо ознайомлення з результатом звіту подібності щодо роботи, генерованого системою виявлення текстових збігів/ідентичності/схожості:

Назва: Методи та засоби ідентифікації об'єктів у тривимірних хмарах технологіями комп'ютерного зору

Автор: Корольков Владислав Олександрович

Спеціальність: 123 – Компютерна інженерія

Освітня програма: освітньо-наукова

Науковий керівник: Березька К. М к.т.н., доцент

Після аналізу звіту подібності зроблено такий висновок:

№	Висновок	Позначка про відповідність
1	Запозичення, виявлені в роботі, є законними і не є плагіатом. Робота приймається до захисту.	відповідає
2	Виявлені запозичення не є плагіатом, розміщені в розділах, які не описують безпосередньо авторське дослідження, але кількість цитат перевищує обсяг, виправданий поставленою метою роботи. Робота приймається до захисту, але має бути відкоригована. Відкоригований варіант має бути поданий на кафедру за 2 дні до захисту, разом із заявою щодо самостійності виконання письмової роботи та ідентичності друкованої та електронної версії роботи	
3	Виявлені запозичення не є плагіатом, але частково розміщені в розділах, які описують безпосередньо авторське дослідження, а кількість цитат перевищує обсяг, виправданий поставленою метою роботи. В зв'язку з цим мета роботи та поставлені завдання не були досягнені. Робота може бути допущена до захисту (наступного року) після того як буде відкоригована та допрацьована і успішно пройде повторну перевірку на академічний плагіат.	
4	Робота містить навмисні текстові спотворення, передбачувані спроби укриття запозичень або інші прояви академічного плагіату. Робота містить фабрикацію або фальсифікацію даних. Робота не допускається до захисту.	

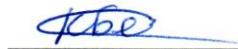
Підтвердження:

Запозичення, виявлені в роботі, є законними і не є плагіатом, оскільки:

- 1) запозичення розміщені в розділах аналізу існуючих аналогів та прототипів, які не описують безпосередньо авторське дослідження і не стосуються результатів роботи;
- 2) усі запозичення фрагментарні, або мають належним чином оформленні посилання;
- 3) окремі виявлені збіги є загальноживаними фразами або виразами, про що свідчить посилання системи на збіг з 10-40 джерелами на один фрагмент речення;
- 4) в якості запозичень в окремих місцях системою зафіксовано послідовності чотирьохрозрядних двійкових кодів, які є вхідними даними до великої кількості задач і не можуть розглядатися як об'єкт авторських прав і, відповідно, їх порушення;
- 5) всі зафіксовані системою ознаки модифікації тексту відносяться до комбінування латинських символів зі україномовними скороченнями індексів в формулах, що не є модифікацією тексту. (Тут текст можна і треба модифікувати)

Сумарний обсяг всіх запозичень, визначений системою виявлення збігів/ідентичності/схожості, складає 2.44% і адресується до 216 першоджерела, що, з урахуванням наведених обґрунтувань, відповідає характеру наукового дослідження і свідчить на користь кваліфікаційної роботи.

Керівник роботи



К. М. Березька

Гарант ОП



О. С. Савенко

Завідувач кафедри КПСч



Т. О. Говорущенко