

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
Західноукраїнський національний університет  
Факультет комп'ютерних інформаційних технологій  
Кафедра інформаційно-обчислювальних систем і управління

КОВАЛЬСЬКИЙ Семен Сергійович

Метод класифікації елементів зображень у середовищі доповненої реальності /  
Method for Classifying Image Elements in Augmented Reality Environment

спеціальність: 122 – Комп'ютерні науки  
освітньо-професійна програма – Комп'ютерні науки

Кваліфікаційна робота

Виконав студент групи  
КНм-21  
С. С. Ковальський

---

Науковий керівник:  
д.т.н., професор А. О. Саченко

---

Кваліфікаційну роботу  
допущено до захисту:  
«\_\_» \_\_\_\_\_ 2022 р.  
Завідувач кафедри  
\_\_\_\_\_ М. П. Комар

ТЕРНОПІЛЬ – 2022

Західноукраїнський національний університет  
Факультет комп'ютерних інформаційних технологій  
Кафедра інформаційно-обчислювальних систем і управління  
Ступінь вищої освіти «магістр»  
спеціальність: 122 "Комп'ютерні науки"  
освітньо-професійна програма - Комп'ютерні науки

ЗАТВЕРДЖУЮ  
Завідувач кафедри

\_\_\_\_\_ “ \_\_\_\_\_ ”  
\_\_\_\_\_ 20 \_\_\_\_ року

### З А В Д А Н Н Я НА КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТУ

Ковальському Семену Сергійовичу

1. Тема роботи «Метод класифікації елементів зображень у середовищі доповненої реальності»

керівник роботи \_\_\_\_\_ д.т.н., професор, Саченко Анатолій  
Олексійович \_\_\_\_\_ ,

затверджені наказом по університету від 31 грудня 2021 року № 606.

2. Строк подання студентом закінченої кваліфікаційної роботи 16 листопада 2022 року.

3. Вихідні дані до роботи: завдання на кваліфікаційну роботу студента, наукові статті, технічна література.

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

- провести опис предметної області доповненої реальності;
- дослідити сучасні методи класифікації елементів зображень;
- сформулювати завдання дослідження та обґрунтувати можливі напрямки його вирішення;
- висвітлити результати дослідження існуючих методів класифікацій елементів зображень у середовищі доповненої реальності;
- визначити сфери застосування методів класифікації елементів зображень;
- сформулювати результати дослідження методів класифікацій елементів зображень у середовищі доповненої реальності;
- описати методіку проведення імітаційних досліджень роботи методів класифікації елементів зображень;
- описати програмну реалізацію і дати характеристику використаного програмного середовища.



## РЕЗЮМЕ

Кваліфікаційна робота на тему «Метод класифікації елементів зображень у середовищі доповненої реальності» на здобуття ступеня вищої освіти «Магістр» зі спеціальності 122 «Комп'ютерні науки» освітньо-професійної програми «Комп'ютерні науки» написана обсягом 102 сторінок і містить 38 ілюстрацій, 1 таблиць, 2 додатки та 63 використаних джерел.

Метою кваліфікаційної роботи є розробка методу класифікації елементів зображень у середовищі доповненої реальності.

Методи досліджень: у роботі використані методи математичного програмування, системного аналізу, машинного навчання.

Результати дослідження: розроблено метод класифікації елементів зображень у середовищі доповненої реальності, що дозволить продуктивніше відображати зображення у середовищі доповненої реальності, а також оптимізовано обсяг вхідних даних, що забезпечить кращу роботу додатків доповненої реальності у яких може використовуватися даний метод.

Результати роботи можуть успішно застосовуватися для використання розробки додатків доповненої реальності, а також класифікації зображень.

Ключові слова: ДОПОВНЕНА РЕАЛЬНІСТЬ, КЛАСИФІКАЦІЯ ЕЛЕМЕНТІВ ЗОБРАЖЕНЬ, ГЛИБОКЕ НАВЧАННЯ, ЗГОРТКОВІ НЕЙРОННІ МЕРЕЖІ, НАБІР ДАНИХ.

## RESUME

The qualification work on the topic " Method for Classifying Image Elements in Augmented Reality Environment " for the degree of Master`s degree in specialty 122 "Computer science" educational and professional program "Computer science" is written in the volume of 102 pages volume and contains 38 illustrations, 1 tables, 2 applications and 63 sources for references.

The purpose of the qualification work is to develop a method for classifying image elements in an augmented reality environment.

Research methods: the work uses methods of mathematical programming, system analysis, and machine learning.

Research results: a method of classifying image elements in an augmented reality environment has been developed, which will allow more productive display of images in an augmented reality environment, and the volume of input data has been optimized, which will ensure better operation of augmented reality applications that can use this method.

The results of the work can be successfully applied to the development of augmented reality applications, as well as image classification.

Key words: AUGMENTED REALITY, CLASSIFICATION OF IMAGE ELEMENTS, DEEP LEARNING, CONVERSIONAL NEURAL NETWORKS, DATASET.

## ЗМІСТ

Вступ.....	7
1 Аналіз методів класифікації елементів зображень у середовищі доповненої реальності.....	10
1.1 Опис предметної області .....	10
1.2 Сучасні методи класифікації елементів зображень .....	20
1.3 Огляд і аналіз існуючих аналогів додатків на основі популярних методів.....	26
1.4 Постановка задачі.....	30
Висновки до розділу 1 .....	32
2 Інформаційне та алгоритмічне забезпечення методу класифікації елементів зображення.....	33
2.1 Технології доповненої реальності.....	33
2.2 Загальна структура класифікації зображень у доповненій реальності.....	42
2.3 Алгоритмічне забезпечення методу класифікації елементів зображень у середовищі доповненої реальності.....	51
Висновки до розділу 2 .....	58
3. Практичне використання методу класифікації елементів зображення у середовищі доповненої реальності.....	59
3.1 Апробація розробленого методу класифікації елементів зображення.....	59
3.2 Експериментальні результати та порівняння .....	65
Висновки до розділу 3 .....	78
Висновки .....	79
Список використаних джерел.....	81
Додаток А Схема алгоритму роботи додатку на базі методу класифікації елементів зображень .....	88
Додаток Б Блок-схема роботи алгоритму SIFT з використанням CNN класифікатора .....	89
Додаток В Апробація отриманих результатів.....	90

## ВСТУП

**Актуальність теми.** У кінотеатрах глядачі можуть повністю поринати у фільми у форматі віртуальної реальності, а перед покупцями відкриється цілий світ актуальних товарів чи послуг, завдяки чому стане можливим випробовування або тестування нових речей у повністю новий спосіб, використовуючи доповнену реальність. Щоб відвідати музей, театр або прогулятися містом, не обов'язково фізично знаходитися у ньому. Всі ці можливості можуть бути реалізовані завдяки сучасним засобам візуалізації. Сьогодні ринок технологій доповненої реальності зростає і стає все більш доступним.

Доповнена реальність (AR) - це інформаційна технологія, яка поєднує реальний світ та цифрові дані. Додатки доповненої реальності все частіше зустрічаються у продукції різних видів. Друкарні все частіше випускають книги, які включають у свої сторінки із зображеннями для зчитування маркерів та відображення доповненої реальності. Традиційні системи доповненої реальності, такі як системи для навчання, дослідження і навіть ігор, раніше були дорогими і вимагали спеціальних пристроїв. За останні роки обчислювальна здатність технологій надзвичайно зросла разом із пропускнуою здатністю передачі, ємністю та швидкістю пам'яті. Цей розвиток технологій дозволив відтворити доповнену реальність на портативних пристроях, наприклад, таких як мобільні телефони. Це, у свою чергу, утворює відкритий ринок з додатками доповненої реальності, оскільки потенційні розробники вже мають відповідну платформу для AR.

Нові мобільні телефони, як і ноутбуки, мають вбудовані камери, потужності яких вистачає для відстеження поверхні, та створення на ній доповненої реальності. Люди використовують соціальні мережі для зустрічей в онлайн режимі, і звикли працювати з веб-камерами.

Технологія хоч і має можливість візуалізації надзвичайно великої кількості об'єктів, як з ознайомчими цілями, з розважальними, з специфічними

по відношенню до роботи завданнями, так і з рекламними цілями, але у всіх цих сферах доповнена реальність буде відображатися з певною затримкою, що у свою чергу утворює проблему і так звані перегони за новими технологіями та методами реалізації прискорення відтворення інформації у середовищі доповненої реальності.

Таким чином виникає нарощування конкуренції за право отримати першими результат, який би зміг покращити швидкість обробки інформації для передачі її структурованої цілісності у використання користувачу певного продукту або технології у вузькоспеціалізованих кругах певних сфер діяльності.

**Метою дослідження** є модифікації методу класифікації елементів зображень у середовищі доповненої реальності. Завдання полягає в зосередженні на аналізі різних методів, які можуть використовуватися для покращення роботи алгоритмічної складової роботи технологій доповненої реальності.

Для досягнення мети необхідно виконати наступні задачі:

- провести опис предметної області доповненої реальності;
- дослідити сучасні методи класифікації елементів зображень;
- сформулювати завдання дослідження та обґрунтувати можливі напрямки його вирішення;
- висвітлити результати дослідження існуючих методів класифікацій елементів зображень у середовищі доповненої реальності;
- визначити сфери застосування методів класифікації елементів зображень;
- сформулювати результати дослідження методів класифікацій елементів зображень у середовищі доповненої реальності;
- описати методику проведення імітаційних досліджень роботи методів класифікації елементів зображень;
- описати програмну реалізацію і дати характеристику використаного



програмного середовища.

**Об'єктами дослідження** є алгоритмічне, інформаційне та програмне забезпечення додатків та методів класифікації елементів зображень у середовищі доповненої реальності.

**Практична цінність** полягає у можливості класифікації елементів зображень у середовищі доповненої реальності. Що в свою чергу збереже робочий час людини, яка буде використовувати додаток. Також за допомогою результату роботи буде виключена можливість помилки у класифікації елементів на зображенні.

**Наукова новизна** у тому, що метод включає можливість використання у багатьох середовищах, де застосовується класифікація елементів зображень, що дозволяє використовувати його не тільки для розробки додатків для середовища доповненої реальності і передачі та обробці великих обсягів даних.

**Апробація результатів** даної роботи була здійснена у двох інтернет конференціях, а саме:

- Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 73)

- Міжнародна науково-практична конференція "Теорія і практика сучасної науки та освіти"

# 1 АНАЛІЗ МЕТОДІВ КЛАСИФІКАЦІЙ ЕЛЕМЕНТІВ ЗОБРАЖЕНЬ У СЕРЕДОВИЩІ ДОПОВНЕНОЇ РЕАЛЬНОСТІ

## 1.1 Опис предметної області

Сьогодні людина кожного дня має діло з графічним дизайном, створеним за допомогою комп'ютерів. У свою чергу графічні форми дають можливість взаємодії з віртуальним світом. Сучасні технології мають змогу відображати ці дані у вигляді користувацького інтерфейсу, який може слугувати у різних цілях та виглядати так як це потрібно певній компанії чи розробнику для досягнення необхідного результату. Визначити їх можна в залежності від потреб певної сфери діяльності. В результаті аналізу і отримавши потрібні розробнику інструменти з'являється змога створення продукту, який буде затребуваний користувачами, а відповідно буде приносити прибуток власнику виробництва. У ході підтримки даних інтерфейсів їх якість виростала з кожним оновленням того чи іншого продукту у різних компаніях, що зародило необхідність кардинальних змін у сфері користувацьких інтерфейсів. Таким чином на думку провідних компаній світу, таких як Google, Apple, Facebook, - у найближчому майбутньому широка популярність набере інноваційний підхід до UI – розробки, а саме відображення даних у середовищі доповненої реальності.

Доповнена реальність – це новітня технологія в сфері комунікації з аудиторією на яку вона націлена, а саме методом накладання на існуючу реальність зображеннями та інформацією з VR.

Типи доповненої реальності визначаються на основі того за допомогою якого методу було викликано створення віртуальних інтерфейсів або об'єктів на уже існуючій площі у реальному світі. Загалом на даний момент є лише два основних типи доповненої реальності, які включають у себе свої підтипи.

Першим типом, який набув популярності є відображення елементів доповненої реальності на основі маркерів, які активують роботу додатку, а отже і відображення закладених у них віртуальних інтерфейсів.

Суть у тому, що при зависанні камери над маркером AR (рис 1.1) споживачу відкривається інтерактивний вміст, тобто клієнт бачить доданий 3D

об'єкт, який у свою чергу може містити у собі різного виду наповнення графічного характеру, а також блок-повідомлення (у відношенні до специфічної події), над яким користувачу надається змога контролю у справжньому світі.

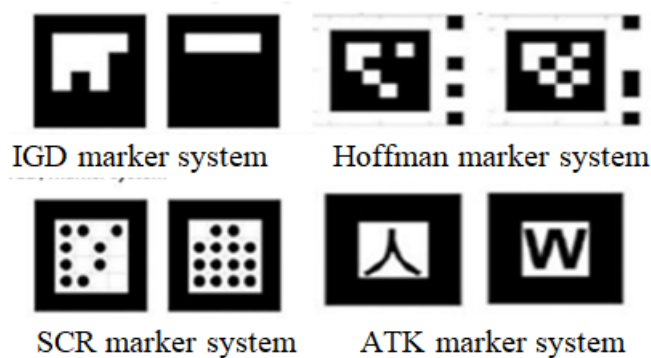


Рисунок 1.1 – Різновиди маркерних систем

Другим основним типом, який є більш перспективним у розвитку і має ширше середовище для застосування є методи, які не використовують ніяких маркерів для відображення елементів доповненої реальності.

Доповнена реальність без маркерів дозволяє позиціонувати віртуальні 3D-об'єкти в середовищі реального зображення, досліджуючи характеристики даних у режимі реального часу. Цей тип вказівок покладається на апаратне забезпечення будь-якого смартфона, зокрема камеру, GPS або акселерометр, а програмне забезпечення доповненої реальності завершує роботу.

У цій моделі немає потреби в системі відстеження об'єктів завдяки останнім технологічним досягненням у камерах, датчиках і алгоритмах ШІ. Таким чином, він працює з цифровими даними, отриманими цими датчиками, здатними записувати фізичний простір у режимі реального часу.

Насамперед безмаркерний аналіз використовує одночасну локалізацію та відображення (SLAM) для сканування середовища та створення відповідних карт для розміщення віртуальних об'єктів. Безмаркерне відстеження зображень SLAM сканує середовище та створює карти розміщення віртуальних об'єктів у 3D, навіть якщо об'єкти не знаходяться в полі зору користувача, вони не рухаються, коли користувач рухається, і користувачеві не потрібно сканувати

нові зображення.

Таким чином, ця технологія може виявляти об'єкти або характерні точки в сцені без попереднього знання навколишнього середовища, наприклад, вона може ідентифікувати стіни або точки перетину. Ця технологія характеризується асоціацією з візуальним ефектом поєднання комп'ютерної графіки з зображеннями реального світу.

Перші системи, які використовували цей тип AR, використовували локацію та апаратні служби пристрою для взаємодії з ресурсами, наданими програмним забезпеченням AR, таким чином, щоб визначали місцезнаходження та орієнтацію користувача в просторі, де він/вона знаходився.

Ще одна особливість цього типу AR полягає в тому, що користувачі можуть збільшити середній діапазон рухів під час досвіду. ARKit від Apple і ARCore SDK від Google зробили AR без маркерів доступним на смарт-пристроях. AR без маркерів є кращим методом розпізнавання зображень для додатків, які використовують цю технологію.

#### AR на основі розташування

Безмаркерна доповнена реальність на основі визначення місця розташування спрямована на об'єднання 3D-віртуальних об'єктів у фізичному просторі, де знаходиться користувач. Ця технологія використовує місцезнаходження та датчики смарт-пристрою, щоб позиціонувати віртуальний об'єкт у потрібному місці чи точці інтересу.

Найбільш репрезентативним прикладом такого типу доповненої реальності є гра для смартфонів Pokémon GO, яка використовує безмаркерну доповнену реальність на основі локації, одразу оживляючи середовище користувача залежно від того, куди він дивиться.

Цей AR пов'язує віртуальне зображення з певним місцем, зчитуючи дані в режимі реального часу за допомогою камери, GPS, компаса та акселерометра. Крім того, оскільки він заснований на доповненій реальності без маркерів, для його роботи не потрібно відстежувати зображення, оскільки він може передбачити підхід користувача, щоб зіставити дані в реальному часі з місцем

розташування користувача.

Крім того, ця типологія дозволяє додавати інтерактивний і корисний цифровий вміст до географічних регіонів, що представляють інтерес, що є дуже корисним для мандрівників у певній місцевості, допомагаючи зрозуміти навколишнє середовище за допомогою 3D віртуальних об'єктів або відео.

#### AR на основі проекцій

Ця методологія використовується для доставки цифрових даних у стаціонарному контексті, тобто AR на основі проекту зосереджується на візуалізації віртуальних 3D-об'єктів у фізичному просторі користувача.

Таким чином, AR дозволяє користувачеві вільно пересуватися по середовищу певної області, де розміщено стаціонарний проектор і камеру стеження. Основним використанням цієї технології є створення ілюзії щодо глибини, положення та орієнтації об'єкта шляхом проектування штучного світла на реальні плоскі поверхні.

Наприклад, AR на основі проекцій підходить для спрощення складних завдань у бізнесі чи промисловості, усуваючи комп'ютери, оскільки інструкції можна розмістити в заданому просторі. Крім того, ця технологія може забезпечити зворотний зв'язок для оптимізації процесів цифрової ідентифікації для виробничих циклів.

#### Накладання AR

Як правило, цей AR використовується для заміни оригінального вигляду об'єкта оновленим віртуальним зображенням цього об'єкта для людського ока. Overlay AR забезпечує кілька переглядів цільового об'єкта з можливістю відображення додаткової відповідної інформації про цей об'єкт.

#### Контурна AR

По суті, ця технологія за допомогою спеціальних камер використовується для того, щоб людські очі окреслювали контури конкретних об'єктів лініями для полегшення певних ситуацій. Наприклад, його можна використовувати для автомобільних навігаційних систем, щоб забезпечити безпечне водіння в умовах поганої видимості.

Системи та додатки доповненої реальності представляють розширену інформацію про реальне середовище, забезпечуючи візуалізацію накладення цифрової інформації, зокрема створеної комп'ютером віртуальної інформації, з візуальними враженнями або зображенням реального середовища. Цифровий інформаційний рух може бути будь-яким типом візуально сприйнятих даних, таких як об'єкти, тексти, малюнки, відео або будь-яка їх комбінація. Реальне середовище знімається, наприклад, камерою, яку тримає користувач, або прикріпленою до пристрою, який тримає користувач. Цифрова інформація має бути накладена на реальне середовище або частину реального середовища на зображенні камери в потрібний час, у потрібному місці та в правильний спосіб, щоб запропонувати користувачам задовольняюче візуальне сприйняття. Відповідний час вимагає, щоб цифрова інформація накладалася на реальне середовище на зображенні тоді і тільки тоді, коли це потрібно або необхідно, наприклад, коли певний реальний об'єкт з'являється на зображенні камери. Відстеження маркерів і калібрування НМД для системи конференц-зв'язку доповненої реальності на основі відео на 2-му міжнародному семінарі IEEE та АСМ з доповненої реальності.

Цифрова інформація має бути накладена на реальне оточення в бажаних позиціях пікселів на зображенні, наприклад, у правильній перспективі, тобто адаптована та отримана з реального середовища, яке переглядається. Щоб досягти цього, має бути відома поза камери, тобто орієнтація та положення щодо реального оточення або його частини (наприклад, See Kato). Зір є незамінним компонентом для обчислення пози камери в додатках доповненої реальності, оскільки зображення камери, яке фіксує реальне середовище, завжди може бути засобом для оцінки пози камери та виявлення реальних об'єктів. Для обчислення положення камери для додатків доповненої реальності було розроблено різні рішення онлайн-відстеження на основі зору.

Правильний спосіб означає, що цифрова інформація має бути вбудована в реальне середовище на зображенні або поданні реального середовища залежно від мети програми. Для додатків доповненої реальності, де віртуальна

інформація використовується для надання інструкцій, привернення уваги користувачів або підтримки розуміння 3D-форм і розмірів, віртуальну інформацію бажано накладати на реальне середовище, щоб вона була яскравою та не відрізнялася від реального середовища. Ця візуалізація є так званою не фотореалістичною візуалізацією. Навпаки, фотореалістична візуалізація є кращою в багатьох програмах, де віртуальну інформацію візуально неможливо відрізнити від реального середовища. Прикладами таких додатків є віртуальний дизайн інтер'єру та доповнений одяг. Для цього було розроблено багато рішень для покращення візуального реалізму шляхом включення оклюзії.

Було розроблено багато методів візуалізації для додатків доповненої реальності, щоб покращити візуалізацію накладання цифрової інформації на візуальні враження або зображення реального середовища, зокрема візуальне сприйняття накладання цифрової інформації та реального середовища. Щоб мати фотореалістичну візуалізацію, Силтанен і Вудворд стирають із зображень штучні маркери, які використовуються для оцінки пози камери, і додають ефекти освітлення й тіні до віртуальних об'єктів. Fischer включає ефекти зображення камери, напр. шум і розмиття у віртуальних об'єктах. основи визначення діапазону даних реального середовища для створення правильного візуального ефекту віртуальної інформації, закритої частиною реального середовища в накладеному зображенні віртуальної інформації та реального середовища.

Ефективність рішень для відстеження на основі зору часто кількісно виражається в термінах помилки повторного проектування або результату вимірювання подібності. Помилка повторної проєкції відповідає відстані в пікселях між спроектованою точкою реальної 3D-точки та виміряною точкою на зображенні. Міра подібності обчислює ступінь відмінності між перетвореною еталонною візуальною характеристикою та візуальною особливістю на зображенні камери. Загальні приклади вимірювань подібності зображень включають суми квадратів відмінностей (SSD), крос-кореляцію та взаємну інформацію. Результатом подібності є дійсне число. Якщо мірою

подібності є різниця суми квадратів між двома фрагментами зображення, чим меншим є результат вимірювання подібності, тим схожішими є дві візуальні характеристики. Якщо мірою подібності є нульова нормалізована крос-кореляція між двома фрагментами зображення, то чим більшим є результат вимірювання подібності, тим більше схожі два візуальні зображення. Запропонований спосіб згідно винаходу, як викладено нижче, може використовувати будь-яку з цих методик.

Параметри (включно з їхніми значеннями) і операційний потік або робочий процес рішень для обробки або відстеження на основі зору завжди налаштовані таким чином, щоб міра подібності та/або помилка повторного проектування були мінімізовані. Однак у додатках доповненої реальності проблема полягає в тому, що жоден із мінімізованих показників схожості та мінімізованої помилки повторної проекції не може вказати або гарантувати задовільне накладання цифрової інформації та реального середовища на зображення камери чи перегляд оптичного прозорого зображення. пристрій відображення. Таким чином, навіть коли Рішення для обробки або відстеження на основі бачення оптимізовано з точки зору вимірювання подібності та/або помилки повторного проектування, часто виникає проблема незадоволеної користувачем візуалізації, наприклад тремтіння, неузгодженість віртуальної інформації з реальним середовищем, і непостійне відображення віртуальної інформації.

Інструменти створення доповненої реальності дозволяють людям, які, наприклад, не мають досвіду розробки програмного забезпечення, обробки зображень і комп'ютерного зору, створювати потужні та гнучкі програми доповненої реальності просто та інтуїтивно зрозуміло. Інструменти розробки відокремлюють використання програми візуалізації та доповненої реальності від рішень на основі зору відстеження, виявлення або локалізації. Користувачі авторських інструментів мають знання про використання візуалізації та властивості цифрової інформації, що накладається, або реального середовища. Під час вибору типу та параметрів алгоритмів відстеження на основі зору



враховується лише еталонний об'єкт, який використовується для локалізації датчика зображення (камери) у середовищі. Однак, як обговорювалося вище, візуалізація доповненої реальності може бути порушена, навіть якщо рішення для відстеження оптимізовано для вимірювання подібності та/або помилки повторної проекції. Тому було б вигідно надати метод обробки зображень у додатку доповненої реальності, який здатний підвищити продуктивність і зручність використання додатків доповненої реальності, зокрема інструментів створення доповненої реальності.

З використанням AR реклама зможе зручно виділити свою важливість, надати споживачу інформацію одразу у момент вибору і у час здійснення вибору можливих рішень про набуття певної продукції у свою власність. Області AR надають змогу споживачам побачити усі функції 3D-моделі, якого завгодно товару або візуально показують повний цикл постачання цієї чи наступної послуги. У такому випадку споживач здобуває надзвичайно реалістичні враження від товару або сервісу, здобуває інформацію, яка засвідчує якість даного продукту.

Актуальність суміжна з:

- наявністю інформації у теперішньому часі;
- взаємодією;
- не звичайним шляхом представлення даних, які надають змогу звертати у свою сторону зацікавленість і підвищувати пізнаваність;
- реалістичністю;
- не традиційність.

Технологія ДР надає максимально інтуїтивну, а також очевидну модель контакту користувача з реальним світом. Для того, щоб отримати змогу спостерігати AR, потрібна камера мобільного пристрою (смартфон, планшет, AR-окуляри) або веб-камера настільного комп'ютера чи ноутбука, а також специфічний програма, яка встановлює цифрові дані (3D моделі, тексти, аудіо, відео) на реальні образи, які приймаються від камери, та доставляє ефект на екран.

Найбільш розповсюджене визначення доповненої реальності у 1997 році дав вчений Рональд Азума [3] визначивши її як структуру, що володіє наступними можливостями:

- додає віртуальність до реальності;
- має можливість взаємодіяти у реальному світі і у теперішній час;
- функціонує в 3D.

Початкові натяки на доповнену реальність, тоді як цього терміну ще не існувало, повсюдно почали застосовуватись у фантастичних літературних публікаціях і поєднаними з ними образотворчій творчості у жанрі альтернативної розповіді, крім того також у виробництві кінофільмів і телебачення, де взаємодіють та змішуються реальні цілі та герої з максимально схожими, але сформовані комп'ютерною анімацією і графікою.

Сьогодні є можливість підкреслити наступні найбільш перспективні ринкові ніші, у яких експлуатація AR проявляє себе найкращим чином:

- сфера розваг і дозвілля;
- сфера медицини;
- навчання;
- маркетингові комунікації та продажі;
- військова сфера.

Вказаний метод надає нескінчений об'єм різної кількості новітніх можливостей у розвитку бренду або медійної персони, а також залученні нових клієнтів. Доповнена реальність дозволяє побудувати взаємодію безпосередньо в момент контакту з рекламним повідомленням. При цьому користувачі зі сторонніх спостерігачів перетворюються в активних учасників процесу, залучаючи і оточуючих теж. Використовуючи ігрові механіки із залучення споживачів, доповнена реальність дозволяє значно збільшити тривалість контакту та час на який запам'ятовується продукт.

Що ж стосується сфери продажу, доповнена реальність – це інструмент, що дозволяє максимально скоротити тривалість циклу продажу для великого спектру товарів.

У сфері навчання на сьогоднішній день традиційна форма навчання все більше поступається місцем сучасним методам, зокрема, дистанційній формі навчання [4]. Для навчального процесу необхідні наочні посібники і проведення лабораторних експериментів, а в разі дистанційної освіти забезпечити всіх студентів посібниками неможливо. І в цьому випадку абсолютно очевидним є застосування програмних продуктів, що використовують доповнену реальність.

У сфері охорони здоров'я технологія доповненої реальності використовується найактивніше. Створюються спеціальні програмні забезпечення для підвищення професійних навичок лікарів, що призводить до зменшення операцій з летальним результатом [5].

Використання даної системи у сфері військового підрозділу теж має велику кількість можливостей. Розробка військових манекенів апаратного типу, покращення розумових здібностей військовослужбовців у питаннях стратегії і тактики щодо проведення оборонних або атакуючих військових дій, створення віртуальних випадків надає змогу сферам збройних сил покращити можливості опанування над різними видами нових для них даних. На теперішній час у апаратному забезпеченні застосовуються індикатори, що передають значиму, корисну і цінну інформацію задля негайного вивчення та оцінки ходу дій, оцінювання лінії і напрямку польоту снаряду. За допомогою таких методів посилюється військова міць країни.

Експлуатація структури додаткової реальності у області проведення вільного часу безмежна. Додатки, які застосовують ДР, є надзвичайно захоплюючими, у деяких випадках чарівними і красивими.

Система даного методу у художніх витворах: при застосуванні структури ДР, полотна з зображеннями будь-чого отримують змогу існувати у реальному світі, архітектурні експозиції починають рухатися, приміщення коригує власний вид. Після таких інноваційних рішень, зацікавленість людей у історичних пам'ятках, а також і у історично відомих людях. Це у свою чергу дає поштовх до інтелектуального піднесення країни.

Отже, доповнена реальність – інноваційний інструмент у всіх областях його використання, тому доцільно буде використати його, щоб створити мобільний додаток для популяризації видатних постатей, а також подати користувачу не тільки бездушні новітні технології, але й разом з тим оживити їх, наділяючи словами та думками великих мислителів.

## 1.2 Сучасні методи класифікації елементів зображень

Сьогодні, із зростаючою мінливістю, необхідністю та застосуванням штучного інтелекту, такі галузі, як машинне навчання та його підмножини, глибоке навчання та нейронні мережі, набули величезного розвитку. Навчання потребує програмного забезпечення та інструментів, таких як класифікатори, які подають величезну кількість даних, аналізують їх і виділяють корисні функції. Метою процесу класифікації є віднесення всіх пікселів цифрового зображення до одного з кількох класів. Зазвичай для виконання класифікації використовуються мультиспектральні дані, і, справді, спектральна картина, наявна в даних для кожного пікселя, використовується як числова основа для категоризації. Мета класифікації зображень полягає в тому, щоб ідентифікувати та відобразити, як унікальний рівень сірого (або колір), особливості, які зустрічаються на зображенні з точки зору об'єкта, який ці особливості фактично представляють на землі. Класифікація зображень є, мабуть, найважливішою частиною аналізу цифрових зображень. Класифікація об'єктів є складним завданням, тому класифікація зображень була важливою задачею в області комп'ютерного зору. Класифікація зображень відноситься до позначення зображень в один із кількох заздалегідь визначених класів. Потенційно існує  $n$  класів, до яких можна класифікувати дане зображення. Ручна перевірка та класифікація зображень може бути виснажливим завданням, особливо коли їх величезна кількість, тому буде дуже корисно, якщо ми зможемо автоматизувати весь цей процес за допомогою комп'ютерного зору. Досягнення в галузі автономного водіння також є чудовим прикладом використання класифікації зображень у реальному світі. Програми включають

автоматизовану організацію зображень, веб-сайти зі стоковими фотографіями та відео, візуальний пошук для кращої видимості продукту, великі візуальні бази даних, розпізнавання зображень і облич у соціальних мережах та багато іншого; Ось чому нам потрібні класифікатори для досягнення максимально можливої точності.

### Структура для виконання класифікації зображень

1. Попередня обробка зображення. Метою цього процесу є покращення даних (функцій) зображення шляхом придушення небажаних спотворень і покращення деяких важливих функцій зображення, щоб моделі комп'ютерного зору могли використовувати ці покращені дані для роботи. Етапи попередньої обробки зображення включають читання зображення, зміну розміру зображення та доповнення даних (відтінки сірого зображення, відображення, розмиття за Гауссом, гистограма, вирівнювання, обертання та переміщення).
2. Виявлення об'єкта: виявлення стосується локалізації об'єкта, що означає сегментацію зображення та ідентифікацію положення об'єкта інтересу.
3. Виділення ознак і навчання: це важливий крок, на якому статистичні методи або методи глибокого навчання використовуються для виявлення найцікавіших моделей зображення, особливостей, які можуть бути унікальними для певного класу і які згодом допоможуть моделі відрізнити різні класи. Цей процес, коли модель вивчає функції з набору даних, називається навчанням моделі.
4. Класифікація об'єкта: на цьому етапі виявлені об'єкти класифікуються за попередньо визначеними класами за допомогою відповідної техніки класифікації, яка порівнює шаблони зображення з цільовими шаблонами.

### Контрольована класифікація

Контрольована класифікація базується на ідеї, що користувач може вибрати зразки пікселів на зображенні, які є репрезентативними для певних класів, а потім наказати програмному забезпеченню обробки зображень використовувати ці навчальні сайти як посилення для класифікації всіх інших пікселів на зображенні. Навчальні сайти (також відомі як набори для тестування або вхідні класи) вибираються на основі знань користувача. Користувач також встановлює межі того, наскільки схожими мають бути інші пікселі, щоб згрупувати їх разом. Ці межі часто встановлюються на основі спектральних характеристик навчальної області. Користувач також визначає кількість класів, до яких класифікується зображення. Після отримання статистичних характеристик для кожного інформаційного класу зображення класифікується шляхом вивчення коефіцієнта відбиття для кожного пікселя та прийняття рішення про те, до якого із сигнатур воно найбільше схоже. Контрольована класифікація використовує алгоритми класифікації та методи регресії для розробки прогнозних моделей. Алгоритми включають лінійну регресію, логістичну регресію, нейронні мережі, дерево рішень, опорну векторну машину, випадковий ліс, наївний Байєс і k-найближчий сусід.

#### Неконтрольована класифікація

Неконтрольована класифікація полягає в тому, що результати (групи пікселів зі спільними характеристиками) базуються на програмному аналізі зображення без надання користувачем зразків класів. Комп'ютер використовує методи визначення пов'язаних пікселів і групує їх у класи. Користувач може вказати, який алгоритм використовуватиме програмне забезпечення, і бажану кількість вихідних класів, але це не допомагає в процесі класифікації. Однак користувач повинен мати знання про область, що класифікується, коли групи пікселів із загальними характеристиками, створені комп'ютером, мають бути пов'язані з фактичними особливостями на землі. Деякі з найпоширеніших алгоритмів, які використовуються в неконтрольованому навчанні, включають кластерний аналіз, виявлення аномалій, нейронні мережі та підходи до вивчення моделей латентних змінних.

### Згорточна нейронна мережа

Згорточна нейронна мережа (CNN або ConvNet) — це особливий вид багатошарових нейронних мереж, призначених для розпізнавання візуальних шаблонів безпосередньо з піксельних зображень з мінімальною попередньою обробкою. Це особлива архітектура штучних нейронних мереж. Згорточна нейронна мережа використовує деякі свої особливості зорової кори і тому досягла найсучасніших результатів у задачах комп'ютерного зору. Згорткові нейронні мережі складаються з двох дуже простих елементів, а саме згорткових шарів і шарів об'єднання. Незважаючи на простоту, існує майже нескінченна кількість способів упорядкувати ці шари для певної проблеми комп'ютерного зору. Елементи згорткової нейронної мережі, такі як згортковий шар і шари об'єднання, відносно прості для розуміння. Складна частина використання згорткових нейронних мереж на практиці полягає в тому, як спроектувати модельні архітектури, які найкраще використовують ці прості елементи. Причина, чому згорточна нейронна мережа надзвичайно популярна, полягає в її архітектурі, найкраще, що немає потреби у вилученні функцій. Система вчиться виділяти ознаки, а основна концепція полягає в тому, що вона використовує згортку зображення та фільтри для створення незмінних функцій, які передаються на наступний рівень. Об'єкти на наступному рівні згорнуті різними фільтрами, щоб створити більш інваріантні та абстрактні об'єкти, і процес триває, доки не буде отримано остаточний об'єкт/вихід, незмінний до оклюзій. Найпоширенішими архітектурами згорткової нейронної мережі є LeNet, AlexNet, ZFNet, GoogLeNet, VGGNet і ResNet.

### Підтримуюча векторна машина

Метод опорних векторів (SVM) — це потужні, але гнучкі керовані алгоритми машинного навчання, які використовуються як для класифікації, так і для регресії. Машини опорних векторів мають унікальний спосіб реалізації порівняно з іншими алгоритмами машинного навчання. Вони надзвичайно популярні через свою здатність обробляти кілька безперервних і категоріальних змінних. Модель опорної векторної машини в основному є представленням

різних класів на гіперплощині в багатовимірному просторі. Гіперплощина буде згенерована ітераційним способом машиною опорних векторів, щоб можна було мінімізувати помилку. Мета полягає в тому, щоб розділити набори даних на класи, щоб знайти максимальну граничну гіперплощину. Він будує гіперплощину або набір гіперплощин у великому просторі, і гарне розділення між двома класами досягається за допомогою гіперплощини, яка має найбільшу відстань до найближчої точки навчальних даних будь-якого класу. Реальна потужність цього алгоритму залежить від використовуваної функції ядра. Найпоширенішими ядрами є лінійне ядро, гауссове ядро та поліноміальне ядро.

#### К-найближчий сусід

K-Nearest Neighbor — це непараметричний метод, який використовується для класифікації та регресії. В обох випадках вхідні дані складаються з  $k$  найближчих навчальних прикладів у просторі ознак. На сьогоднішній день це найпростіший алгоритм. Це непараметричний алгоритм відкладеного навчання, у якому функція апроксимується лише локально, а всі обчислення відкладено до оцінки функції. Цей алгоритм просто покладається на відстань між векторами ознак і класифікує невідомі точки даних, знаходячи найпоширеніший клас серед  $k$ -найближчих прикладів. Щоб застосувати класифікацію  $k$ -найближчого сусіда, нам потрібно визначити метрику відстані або функцію подібності, де загальні варіанти включають евклідову відстань і манхеттенську відстань. Результатом є членство в класі. Об'єкт класифікується більшістю голосів його сусідів, причому об'єкт призначається до класу, який є найпоширенішим серед  $k$  його найближчих сусідів ( $k$  — додатне ціле число, зазвичай невелике). Якщо  $k = 1$ , то об'єкт просто призначається до класу цього єдиного найближчого сусіда. Стислий найближчий сусід (CNN, алгоритм Харта) — це алгоритм, призначений для скорочення набору даних для класифікації  $K$ -найближчого сусіда.

#### Алгоритм випадкового лісу

Випадковий ліс — це контрольований алгоритм навчання, який використовується як для класифікації, так і для регресії. Оскільки ми знаємо,



що ліс складається з дерев, і більше дерев означає більш надійний ліс, подібним чином алгоритм випадкового лісу створює дерева рішень на вибірках даних, а потім отримує прогноз від кожного з них і, нарешті, вибирає найкраще рішення шляхом голосування. Це метод ансамблю, який є кращим, ніж одне дерево рішень, оскільки він зменшує надмірне підгонювання шляхом усереднення результату. Випадковий ліс — це алгоритм класифікації, що складається з багатьох дерев рішень. Він використовує пакетування та випадковість функцій під час побудови кожного окремого дерева, щоб спробувати створити некорельований ліс дерев, передбачення якого комітетом точніше, ніж прогнозування будь-якого окремого дерева.

### Штучна нейронна мережа

Натхненні властивостями біологічних нейронних мереж штучні нейронні мережі є статистичними алгоритмами навчання, які використовуються для різноманітних завдань, від відносно простих завдань класифікації до комп'ютерного зору та розпізнавання мовлення. Штучні нейронні мережі реалізовані як система взаємопов'язаних елементів обробки, званих вузлами, які функціонально аналогічні біологічним нейронам. Зв'язки між різними вузлами мають числові значення, які називаються вагами, і систематично змінюючи ці значення, мережа вресі-ресіт може наближено визначити бажану функцію. Приховані шари можна розглядати як окремі детектори ознак, які розпізнають все більш складні шаблони в даних у міру того, як вони поширюються по мережі. Наприклад, якщо мережі дається завдання розпізнати обличчя, перший прихований шар може діяти як детектор ліній, другий прихований приймає ці лінії як вхідні дані та об'єднує їх, щоб утворити ніс, третій прихований шар приймає ніс і зіставляє його оком і так далі, поки нарешті не буде створено все обличчя. Ця ієрархія дозволяє мережі зрештою розпізнавати дуже складні об'єкти. До різних типів штучної нейронної мережі належать згорточна нейронна мережа, нейронна мережа прямого зв'язку, ймовірнісна нейронна мережа, нейронна мережа із затримкою часу, мережа

глибокого накопичення, мережа радіальної базисної функції та рекурентна нейронна мережа.

1.3 Огляд і аналіз існуючих аналогів додатків на основі популярних методів

Додатки доповненої реальності в туризм є розквітаючою галуззю, яка відіграє обов'язкову роль у визначенні традиційного туризму, і що дані на мобільному екрані туристів стають більш захоплюючими завдяки методу накладання. Додаток має на меті збільшити досвід туризму та максимізувати рівень розваг, а також допомогти туристам отримати доступ до інформації та покращити їхні знання про привабливість місця призначення. Програми AR пропонують невідомі регіони як туристичну діяльність у веселій та інтерактивній формі, а зі збільшенням використання вони надають можливість розвивати досвід відвідувачів у напрямках і туристичних організаціях (Dieck & Jung, 2015). Наголошуючи на необхідності використання сучасних технологій, Хан та ін. (2014) рекомендує, щоб багато напрямків були адаптовані, щоб залишатися конкурентоспроможними та привабливими для туристів, і рекомендує використовувати новий тип (AR), використання якого поступово збільшується. Однак він також зазначає, що він ще не досяг бажаного потенціалу та чисельності населення, але стане основною технологією в майбутньому, і індустрія туризму повинна добре використовувати цю позицію. Бухаліс та ін. (2012) стверджують, що парадигми, які виникають із синергії AR, мають потенціал для покращення туристичного досвіду та підвищення їх рівня, додаючи, що розроблені та застосовні технології, які надають інформацію туристам у будь-який час і за будь-яких умов, більше не є майбутня мрія. Однак є припущення, що це все ще на початковій стадії. Багато застосувань можна назвати в багатьох сферах туризму. Одна з цих програм, *Tuscany AR*, працює як туристичний путівник, розроблений Тосканою, спеціально для регіону Тоска. Вікіпедія співпрацює з багатьма порталами та надає туристам інформацію про житло, їжу та напої, нічне життя та подорожі на природу за допомогою

інформації, взятої з джерел. Layer – це проект, створений для міста Базель. Він надає туристу всю інформацію про ресторани та бари, музеї, торгові центри, події та екскурсії, екскурсії та місця відпочинку в Базелі та його околицях. Усі операційні системи працюють у компетенції одна з одною. Urban Sleuth був розроблений Urban interactive як місто пригод, де користувачі беруть участь у розгадуванні таємниць і виконанні місій під час відвідування міста. Ця програма відіграє важливу роль в інтеграції реального та віртуального світу. Дизайн програми можна зробити так, щоб туристи могли знаходити пам'ятники, музеї та місця культурної спадщини. Нарешті, програма вуличного музею, яка дозволяє туристам візуалізувати особливі та історичні місця Лондона, була розроблена спеціально для Лондонського музею. Люди можуть отримати інформацію, натиснувши кнопку, спрямувавши камери мобільних пристроїв у будь-яку точку, яку вони цікавлять. Додаток також дозволяє туристам відкривати важливі місця перед початком туру (Kesimati et al., 2012).

Очевидно, що і у сфері популяризації та просування сфери туризму існує певний список мобільних додатків.

ArtLens 2.0, Музей мистецтв Клівленда. ArtLens 2.0, одна з найважливіших програм, розроблених і розроблених музеями, використовується в художньому музеї Клівленда. Додаток (рис 1.2) було запущено в результаті 6-місячної пробної версії в 2016 році. Мета додатка – представити вподобання та інтереси відвідувачів у 2D за допомогою програмного забезпечення для розпізнавання зображень. Крім того, програма, яка записує та моделює розмови відвідувачів, створює атмосферу. Додаток також забезпечує велику зручність для відвідувачів, використовуючи можливість відкривати нові дороги за допомогою функції карти та швидше діставатися до потрібного місця. Крім того, він також може попередити їх, коли вони наближаються до дивовижних робіт. Додаток відображає всі твори мистецтва в режимі реального часу за допомогою технології Bluetooth для Collection Wall, 40-метрової інтерактивної стіни MicroTile із підтримкою кількох дотиків. ArtLens 2.0, який містить усі ці функції, можна завантажити та

використовувати приблизно за одну хвилину в Google Play та App Store. Джейн Александер, головний інформаційний директор Музею мистецтв Клівленда, стверджує, що ця програма відрізняється від інших програм, оскільки вона має динамічну структуру, яка оновлює назву, дату, інформацію про надгробки, дидактичні дані, відео та розташування галереї в режимі реального часу.

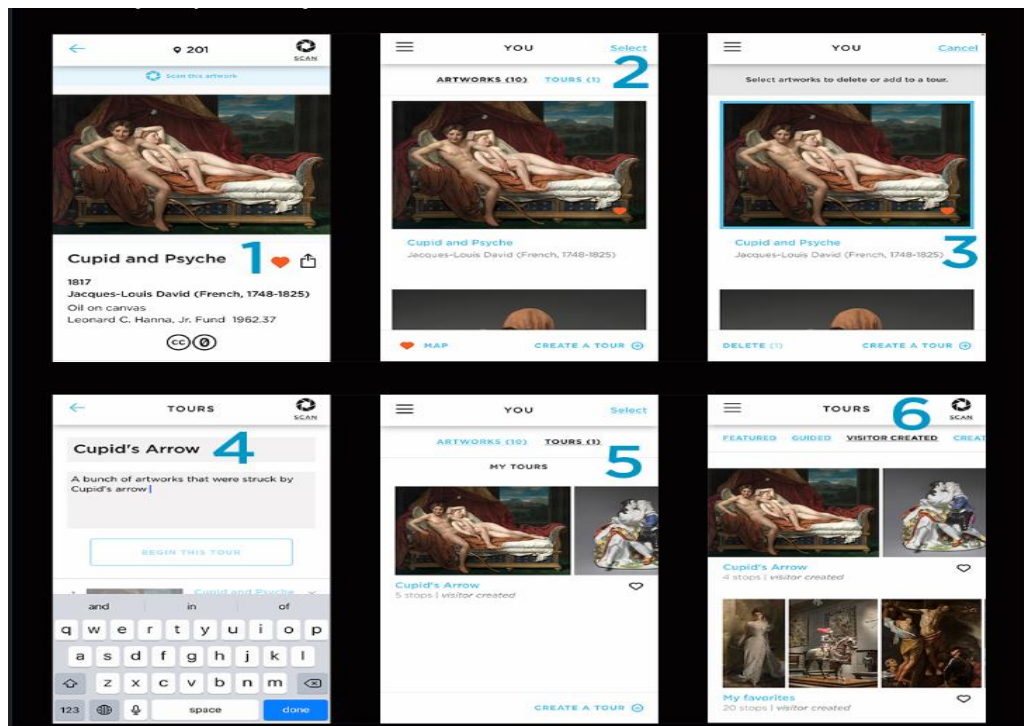


Рисунок 1.2 - ArtLens 2.0, Музей мистецтв Клівленда.

Layar the Blanton Museum 3 грудня 2015 по квітень 2016 року Художній музей Блантона при Техаському університеті в Остіні використовував Layar AR (рис 1.3), щоб надавати відвідувачам пояснювальну інформацію для виставки «The Crusader Bible: A Gothic Masterpiece». Кілька тестів було проведено за 2 години до використання програми, доступної як для iOS, так і для Android. Керівництво музею підкреслило, що артефакти мають бути розкриті ефектно та інформативно. Керівництво, яке вважало, що це буде важко представити, оскільки роботи, виставлені на виставах, містять такі мови, як перська та латинська, досягло дуже хороших результатів, використовуючи безкоштовну програму Layar AR, сумісну з мобільними пристроями та iPad. Під час виставки було встановлено, що в середньому 2000 із 9800 відвідувачів користувалися

програмою Layar. «Додаток додав рівень інтерпретації, який було б майже неможливо забезпечити традиційним способом етикеток. Враховуючи тематику виставки, відвідувачі більше, ніж зазвичай, прагнули залучитися до розповіді самих робіт, і додаток дозволив їм це зробити».



Рисунок 1.3 - Layar the Blanton Museum

Best Western Hotels & Resorts BWVRE Fast Company сьогодні оголосила про свій щорічний рейтинг найбільш інноваційних компаній світу (MIC), відзначаючи Best Western Hotels & amp; Курорти з місцем у 10 найінноваційніших компаній у категорії AR/VR. Бренд отримав визнання за його двосторонній підхід до VR/AR – включаючи Best Western Virtual Reality Experience (BWVRE), який використовує віртуальну реальність, щоб заново винайти спосіб перегляду та покупки готелів у гостей; а також тренінг для співробітників «Я дбаю про кожного гостя щоразу», який використовує віртуальну реальність для трансформації спілкування між персоналом стійки реєстрації та гостями. Ці ініціативи зробили Best Westerns першою компанією такого розміру та масштабу, яка використовує цю передову технологію таким чином. Бренд Best Western прагне слідувати новітнім технологіям, щоб задовольнити потреби туристів. Керівництво задоволене тим, що технології AR і VR принесли новий стандарт у секторі, і прагне зберегти лідерство в інноваціях і інноваційній гостинності. BWVRE пропонує туристам можливість

перевірити громадські зони готелю перед бронюванням або проживанням. Додаток створює інтерактивні можливості для відвідувачів басейну готелю, лобі, фітнес-центру та типів номерів із тривимірними 360-градусними відео, ексклюзивним дикторським текстом і музикою. Клієнти можуть досліджувати все, від дизайну інтер'єру до якості продукції, прогулюючись місцями. Тепер мандрівники можуть відвідувати понад 2200 готелів Best Western у Північній Америці на всіх платформах, що підтримуються HTML5, включаючи Карти Google, Пошук Google, YouTube і Facebook.

Таким чином можна зробити висновок, що на даний момент торгівельні площадки не можуть надати універсальних додатків для розвитку сфери туризму, а в особливості збереження культурної спадщини та не мають можливостей розвитку.

Основними недоліками досліджених додатків стали:

- великий об'єм додатку;
- недостатній об'єм сюжетного наповнення;
- невелика кількість контенту;
- відсутня універсальність контенту.

#### 1.4 Постановка задачі

Мобільні додатки з ДР дозволяють значно підвищити інтерес публіки до певних сфер в яких вони будуть використовуватися, що в свою чергу робить їх більш популярними. Завдяки застосованим передовим мультимедійним технологіям, що дає можливість «оживити» та представити в наочному вигляді найскладніші теми з використанням практично будь-якого сучасного комп'ютера, планшета або смартфона. Переваги використання ДР: візуалізація – полегшує процес запам'ятовування та розвиває абстрактну уяву; наочність – трьох вимірний підхід дозволяє дослідити пристрій або явище у деталях, з різних боків; цікавість – «живі» картинки на сторінках навчального посібника; сучасність – оволодіння інноваційними технологіями в процесі популяризації;

увага студентів; портативні та майже безкоштовні навчальні матеріали – з AR не потрібно вкладати кошти у фізичні матеріали та обладнання.

Для досягнення поставленої мети в процесі роботи необхідно виконати такі задачі:

- провести аналіз можливостей технології доповненої реальності, для яких може бути використаний метод класифікації елементів зображення;
- визначити обмеження та вимоги, необхідні для реалізації методу класифікації елементів зображення на основі технології доповненої реальності;
- дослідивши існуючі методи класифікацій елементів зображень потрібно проаналізувати їх недоліки та переваги;
- на основі визначених рішень імітувати роботу додатку на основі методу класифікації елементів зображень у середовищі доповненої реальності.

Так як метод буде використаний у середовищі доповненої реальності, виникає потреба у задіяні графічного ядра, яке б гарантувало якісний функціонал.

## Висновки до розділу 1

1. У розділі дослідження визначено, що доповнена реальність - це новітня технологія в сфері комунікації з аудиторією на яку вона націлена, а саме методом накладання на існуючу реальність зображеннями та інформацією з VR. Визначено основні типи доповненої реальності. Також досліджено у яких сферах використовується дана технологія.

2. Проведено аналіз існуючих підходів та визначено, що на відміну від аналогів розроблений інтелектуальний метод класифікації елементів зображень, дозволить покращити роботу даних методів для класифікації елементів зображень при великій кількості вхідних даних і підвищить продуктивність роботи додатків де він буде використовуватися

3. Розглянуто, існуючі аналоги на основі сучасних методів класифікацій зображень у середовищі доповненої реальності. Визначено, що дані додатки працюють назвичайно не стабільно через використання великої кількості даних, які додаток не встигає обробляти для повноцінного функціонування, через що користувацький інтерфейс є низької якості.

4. На основі проведеного дослідження сформовано постановку задачі, а саме розробка методу класифікації елементів зображень у середовищі доповненої реальності.



## 2 ІНФОРМАЦІЙНЕ ТА АЛГОРИТМІЧНЕ ЗАБЕЗПЕЧЕННЯ МЕТОДУ

### 2.1 Технології доповненої реальності.

Середовищем розробки програмного забезпечення, що призначене для створення додатку доповненої реальності потрібно вибрати одну з платформ ігрових рушіїв, яка буде максимально відповідати вимогам завдання, а також зазначити ПЗ для розробки 3D моделі.

Розробники використовують ігрові рушії для створення для консолей, мобільних пристроїв і персональних комп'ютерів. Основні функціональні можливості, що зазвичай надаються ігровим рушієм, включають в себе механізм візуалізації (Render або обробка графічного зображення) для 2D або 3D графіки, фізичний рушії або виявлення зіткнень (і реагування на зіткнення), звук, сценарії, анімацію, штучний інтелект, мережеве взаємодія, потокову передачу, пам'ять управління, багатопоточність, підтримка локалізації, графічні сцени і може включати в себе підтримку відео для кінематографії. У багатьох випадках ігрові рушії надають набір інструментів візуальної розробки на додаток до повторно використовуваних програмних компонентів. Ці інструменти можуть надаватися в інтегрованому середовищі розробки, що дозволяє спростити і прискорити розробку ігор на основі даних. Більшість ігрових рушіїв надають засоби, які полегшують розробку, такі як графіка, звук, фізика та функції штучного інтелекту. Ці ігрові рушії іноді називають «проміжним програмним забезпеченням», тому що, як і в діловому сенсі цього слова, вони надають гнучку і багаторазово використовувану програмну платформу, яка надає всі основні функціональні можливості, необхідні для розробки ігрової програми, при одночасному зниженні витрат. Як і інші типи проміжного програмного забезпечення, ігрові рушії зазвичай надають абстракцію платформи, дозволяючи запускати одну і ту ж гру на різних платформах, включаючи ігрові приставки і персональні комп'ютери, з невеликими змінами в вихідному коді гри, якщо такі є. Найчастіше ігрові рушії розробляються на основі компонентної архітектури, яка дозволяє замінювати

або розширювати конкретні системи в рушії більш спеціалізованими (і часто більш дорогими) компонентами ігрового проміжного програмного забезпечення. Деякі ігрові рушії спроектовані як набір слабо пов'язаних компонентів проміжного програмного забезпечення, які можна вибірково комбінувати для створення власного рушія замість більш поширеного підходу до розширення або налаштування гнучкого інтегрованого продукту. Незважаючи на те, що розширюваність досягається, вона залишається високим пріоритетом для ігрових рушіїв через великого розмаїття областей застосування, для яких вони застосовуються.

Незважаючи на специфіку назви, ігрові рушії часто використовуються для інших видів інтерактивних додатків з графічними потребами в реальному часі, такими як маркетингові демонстрації, архітектурні візуалізації, симуляції навчання і середовища моделювання. Деякі ігрові рушії надають тільки можливості 3D-рендеринга в реальному часі замість широкого спектра функцій, необхідних для ігор. Двигуни покладаються на розробника ігор, щоб реалізувати решту цієї функціональності або зібрати її з інших компонентів проміжного програмного забезпечення гри. Ці типи двигунів зазвичай згадуються як «графічний рушій», «рушій відображення» або «тривимірний рушій» замість більш всеосяжного терміна «ігровий рушій». Ця термінологія використовується непослідовно, оскільки багато повнофункціональні ігрові рушії називаються просто «тривимірними двигунами». Сучасні ігрові або графічні рушії зазвичай надають графічні сцени, який є об'єктно-орієнтованим поданням ігрового світу 3D, який часто спрощує ігровий дизайн і може використовуватися для більш ефективного рендеринга великих віртуальних світів. Гральний рушій – це центральний програмний компонент комп'ютерних та відео ігор або інших інтерактивних додатків з графікою, оброблюваної в реальному часі. Він забезпечує основні технології, спрощує розробку і часто дає грі можливість запускатися на декількох платформах, таких як ігрові консолі та 16 настільні операційні системи, наприклад, GNU / Linux, Mac OS X і Microsoft Windows. Основну функціональність зазвичай забезпечує гральний

рушій, що включає рушій рендеринга («відображувач»), фізичний рушій, звук, систему скриптів, анімацію, штучний інтелект, мережевий код, управління пам'яттю і багатопоточність [7].

Хоча серед демонстрації можливостей рушія були вказані лише звичайні програми (не ігри). Також поки відсутня можливість використання нативних бібліотек. А сам API теж вельми бідний на можливості. Зробивши огляд деяких відомих ігрових рушіїв було прийнято рішення використовувати рушій Unity. Таке рішення базується на таких критеріях: - вартість. Використовуючи Unity у своїх цілях, розробник користується їм безкоштовно; - мови програмування. C# і JavaScript являють собою досить розповсюджене використання; - функціонал. Так як Unity підтримує можливість завантаження плагінів, функціонал цього рушія може бути необмеженим; - інтерфейс. Інтерфейс Unity є досить інтуїтивним, що полегшує роботу з ним; - мультиплатформеність. Завдяки Unity можна легко і якісно створювати ігри і додатки на велику кількість відомих платформ.

Unity – це кросплатформене середовище розробки. Unity дозволяє створювати додатки, що працюють під більш ніж 20 різними операційними системами, що включають персональні комп'ютери, ігрові консолі, мобільні пристрої, інтернет-додатки та інші.

Unity має простий Drag & Drop інтерфейс, який легко налаштовувати, що складається з різних вікон, завдяки чому можна проводити налагодження гри прямо в редакторі. Рушій підтримує два скриптових мови: C #, JavaScript (модифікація). Раніше була підтримка Boo (діалект Python), але його прибрати в 5-ій версії. Розрахунки фізики виробляє фізичний рушій PhysX від NVIDIA.

Інтерфейс Unity влючає в себе вікно проекту, вікно сцени, вікно гри, вікно ієрархії, вікно інспектора, панель інструментів, консольне вікно, та опціональні вікна сторонніх плагінів. До об'єктів можна застосовувати колізії (в Unity так звані Колайдери – collider), яких існує декілька типів. Також Unity підтримує фізику твердих тіл і тканини, а також фізику типу Ragdoll (тряпична лялька). У редакторі є система успадкування об'єктів; дочірні об'єкти будуть

повторювати всі зміни позиції, повороту і масштабу батьківського об'єкта. Скрипти в редакторі прикріплюються до об'єктів у вигляді окремих компонентів. При імпорті текстури в Unity можна згенерувати alpha-канал, мір-рівні, normal-map, light-map, карту відображень, проте безпосередньо на модель текстуру прикріпити не можна – буде створено матеріал, з яким буде призначений шейдер, і потім матеріал прикріпиться до моделі.

Unity має компонент для створення анімації, але також анімацію можна створити попередньо в 3D-редакторі та імпортувати разом з моделлю, а потім розбити на файли. Unity 3D підтримує систему Level Of Detail, суть якої полягає в тому, що на далекій відстані від гравця високодеталізовані моделі замінюються на менш деталізовані, і навпаки, а також систему Occlusion culling, суть якої в тому, що у об'єктів, які не потрапляють в поле зору камери не візуалізується геометрія і колізія, що знижує навантаження на центральний процесор і дозволяє оптимізувати проект. При компіляції проекту створюється виконуваний (.exe) файл гри (для Windows), а в окремій папці – дані гри (включаючи всі ігрові рівні і спільні бібліотеки). Рушій підтримує безліч популярних форматів. Моделі, звуки, текстури, матеріали, скрипти можна запаковувати в формат .unityassets і передавати іншим розробникам, або викладати у вільний доступ.

Формат .unityassets використовується у внутрішньому магазині Unity Asset Store, в якому розробники можуть безкоштовно і за гроші викладати в загальний доступ різні елементи, потрібні при створенні ігор. Щоб використовувати Unity Asset Store, необхідно мати акаунт розробника Unity. Unity має всі потрібні компоненти для створення мультиплеєра. Також можна використовувати відповідний користувачеві спосіб контролю версій. Наприклад, Tortoise SVN або Source Gear. У Unity входить Unity Asset Server – інструментарій для спільної розробки на базі Unity, є доповненням, що додає контроль версій і ряд інших серверних рішень.

Як правило, ігровий рушій надає безліч функціональних можливостей, що дозволяють їх задіяти в різних іграх, в які входять моделювання фізичних

середовищ, карти нормалей, динамічні тіні і багато іншого. На відміну від багатьох ігрових рушіїв, у Unity є дві основні переваги: наявність візуальної середовища розробки та міжплатформенна підтримка. Перший фактор включає не тільки інструментарій візуального моделювання, а й інтегровану середу, лінію складання, що направлено на підвищення продуктивності розробників, зокрема, етапів створення прототипів і тестування. Під міжплатформенною підтримкою надається не тільки місце розгортання (установка на персональному комп'ютері, на мобільному пристрої, консолі і т.д.), а й наявність інструментарію розробки (інтегроване середовище може використовуватися під Windows і Mac OS). Третьою перевагою є модульна система компонентів Unity, за допомогою якої відбувається конструювання ігрових об'єктів, коли останні є комбінованими пакетами функціональних елементів. На відміну від механізмів успадкування, об'єкти в Unity створюються за допомогою об'єднання функціональних блоків, а не переміщення в вузли дерева успадкування. Такий підхід полегшує створення прототипів, що актуально при розробці ігор. Як недоліки наводяться обмеження візуального редактора при роботі з багатокomпонентними схемами, коли в складних сценах візуальна робота ускладнюється. Іншим недоліком є відсутність підтримки Unity посилань на зовнішні бібліотеки, роботу з якими програмістам доводиться налаштовувати самостійно, і це також ускладнює командну роботу. Ще один недолік пов'язаний з використанням шаблонів примірників (англ. Prefabs). З одного боку, ця концепція Unity пропонує гнучкий підхід візуального редагування об'єктів, але з іншого боку, редагування таких шаблонів є складним [8].

Наступним кроком потрібно вибрати бібліотеку доповненої реальності серед яких:

- ARCore;
- ARToolKit;
- EasyAR;
- Kudan;

- Maxst;
- Vuforia;
- Xzing;
- Wikitude;
- ThingWorx Studio.

З цього списку можна виключити ARCore, тому що ця платформа підходить для обмеженої кількості смартфонів. Всі представлені інструменти повинні мати сумісність з міжплатформним середовищем розробки Unity. Unity дозволяє створювати проекти як в 2D, так і в 3D, підтримує дві скриптових мови: C #, JavaScript. Це середовище розробки містить в собі редактор сцен, редактор об'єктів і редактор скриптів. Для створення реалістичної поведінки об'єктів в Unity вбудований фізичний рушій, що забезпечує компонентами для симуляції фізичної поведінки об'єктів. При розробці використовується компонентноорієнтований підхід до програмування. Необхідність використання Unity обумовлена тим, що це середовище дозволяє компілювати проекти під операційні системи Android і iOS, спрощує роботу з 3D моделями.

ARToolKit є набором програмних бібліотек, які можуть використовуватися в AR додатках. Головна перевага бібліотеки – відкритий вихідний код. Відсутня WaterMark. WaterMark або водяний знак – це напівпрозоре зображення, що накладається поверх зображення з камери. Найчастіше водяний знак представлений логотипом компанії і присутній в безкоштовних ліцензіях бібліотек доповненої реальності. Є можливість розпізнавати 2D об'єкти. Незважаючи на безкоштовний доступ, документація для розробників вельми обмежена. Відсутня можливість розпізнавання об'ємних об'єктів і мітки з хмари [9].

EasyAR – безкоштовна і проста у використанні бібліотека. Має можливість розпізнавання тільки 2D об'єктів. Відсутня WaterMark. Є хорошою альтернативою Vuforia за якістю розпізнавання міток. Підтримується до 1000

міток для розпізнавання. Відсутні хороші навчальні посібники, вивчення можливо тільки за документацією, що не зручно і не продуктивно.

Kudan – до позитивних можливостей можна віднести відсутність обмежень на кількість розпізнаваних зображень і маленький обсяг пам'яті, необхідний для зберігання файлів на пристрої. Існує можливість розпізнавання 3D об'єктів, безмаркерне відстеження об'єктів. Розробники можуть скористатися базовою документацією при зверненні до бібліотеки. Однак керівництво з використання не дуже докладний і вимагає пошуку додаткової інформації. Відображення доповнень реалізується через окремий компонент-обгортку над OpenGL. Є WaterMark. Але безкоштовна версія призначена тільки для тестування додатків. Для публікації додатка потрібна покупка ліцензії [10].

Maxst – оптимізований саме для мобільних платформ. Розпізнає і відстежує зображення, але одночасно максимально тільки 3 мети. Є можливість відстежити плоску поверхню і розмістити на ній потрібний об'єкт. Безкоштовна версія дозволяє вести розробку тільки для ОС Android і iOS. Є підтримка одночасного створення і збереження тривимірної карти простору. Maxst працює тільки з 32-розрядної версії редактора Unity. Maxst володіє не дуже гарною якістю розпізнавання міток. Xzimg – є три продукти для роботи з додатками на основі доповненої реальності: Augmented Face, Augmented Vision і Magic Face. Xzimg Augmented Face розпізнає і відстежує особи. Xzimg Augmented Vision розпізнає і відстежує плоскі зображення.

Xzimg Magic Face призначений для заміни рис обличчя і нанесення макіяжу. Є можливість розробки проекту під HTML5. Наявність WaterMark. Безкоштовна пробна версія доступна тільки для демонстрації (інвертує колір і змінює зображення) [11].

Wikitude має функцію Extended Tracking для збереження положення мітки на мобільному пристрої, навіть якщо вона за межами огляду камери. Підтримується розпізнавання як 2D, так і 3D форматів. Існує можливість сканування реального об'єкта для подальшого розпізнавання. Є хороша документація, але при цьому мала кількість навчального матеріалу. Але

безкоштовна версія надається тільки для тестування. Для подальшої розробки та публікації необхідно придбання ліцензії [12].

ThingWorx Studio – універсальний інструмент для створення і публікації продуктів доповненої реальності. Дозволяє розробляти програми для промислових підприємств, використовує зрозумілий графічний редактор.

У ThingWorx Studio використовується технологія комп'ютерного зору Vuforia. Мітки для розпізнавання і 3D об'єкти зберігаються і використовуються через хмарну базу даних. ThingWorx Studio єдине з представлених середовищ розробки дозволяє безпосередньо використовувати САПР моделі. Але даний інструментарій має тільки пробну безкоштовну версію, для подальшої розробки та публікації необхідно придбання ліцензії [13].

Vuforia – одна з найпопулярніших в світі платформ доповненої реальності. Існує велика кількість навчального матеріалу як в текстовому, так і в відео форматі. Програмне забезпечення реалізує наступні функції: розпізнавання різних типів візуальних об'єктів (куб, циліндр, площина), розпізнавання тексту і навколишнього середовища, VuMark (комбінація зображення і QR-коду). Процес розпізнавання команд може бути реалізований з використанням локальної або хмарної бази даних. Обмеження безкоштовної версії відносяться тільки до числа VuMark, кількості взаємодій 68 з хмарної базою даних і наявності WaterMark. Vuforia підтримує Extended Tracking. Є можливості розпізнавання 3D об'єктів і підтримки Virtual Button.

Виходячи з наведеного аналізу, зроблений вибір інструменту розробки програмного забезпечення доповненої реальності Vuforia. Дана платформа найбільш популярна, а тому більш доступна для вивчення. На даний момент власником Vuforia є компанія PTC, що займається впровадженням в роботу підприємств доповненої реальності, інтернету речей, 3D-друку, цифрових двійників, що дає надію на подальшу активну підтримку і розвиток ДР SDK (software development kit) Vuforia. Будучи повноцінною SDK для розробників, Vuforia являє собою набір інструментів для створення додатків доповненої реальності.



Vuforia підтримує:

- розпізнавання декількох цілей одночасно (включаючи об'єкти, зображення і текст);
- відстеження цілей;
- розпізнавання 2D і 3D форматів;
- сканування реального об'єкта для подальшого розпізнавання;
- віртуальні кнопки;
- відображення додаткових елементів через OpenGL;
- Smart Terrain™ – можливість реконструювати навколишній ландшафт, створюючи його 3D-карту;
- Extended Tracking – можливість продовжити відображення мети на мобільному пристрої, навіть коли вона знаходиться поза полем зору.

Зокрема, при розпізнаванні зображень Vuforia дозволяє мобільним додаткам використовувати дані, що знаходяться або на пристрої, або в хмарі.

До основних переваг бібліотеки можна віднести підтримку пристроїв віртуальної реальності, а також тестовий додаток з супроводжуючими коментарями, в якому показані можливості бібліотеки. Однак відсутність повноцінного керівництва з використання бібліотеки ускладнює перший досвід роботи з Vuforia. Окремі інструкції і короткі поради представлені у великій кількості, але не впорядковані і тому не замінюють розробнику необхідну документацію [15] [16].

Огляд існуючих додатків показав основні тенденції в розвитку технології доповненої реальності. Ця технологія в наш час дуже бурхливо розвивається і може знайти застосування в багатьох областях. Аналіз додатків показав, що зараз вони мають, в основному, або розважальний, або вузькоспрямований характер. Основна увага в таких додатках приділяється їх можливостям при взаємодії з навколишнім середовищем і користувачем.

Огляд альтернатив та визначення основних критеріїв аналізу зображень [17] та урахування найбільш важливих критеріїв показали, що найоптимальнішим середовищем для розробки системи є Vuforia. Vuforia –

безкоштовна бібліотека, постійно модернізується, дозволяє працювати як з нативними додатками, так і створювати кросплатформені програми за допомогою спеціального ігрового рушія Unity.

## 2.2 Загальна структура класифікації зображень у доповненій реальності

Дещо складнішим є визначення методу, який був би продуктивнішим для обробки інформації на фото, відео, тексті і інших видів даних, які можуть бути відображеними у користувацькому інтерфейсі.

Класифікація між об'єктами є досить легким завданням для живого створіння, але це виявилось складним для машин, тому класифікація зображень є важливим завданням у сфері комп'ютерного зору. Класифікація зображень відноситься до позначення зображень в один із кількох заздалегідь визначених класів.

Потенційно існує невизначена кількість класів, яка може відрізнитися в залежності від задачі і її середовища, до яких можна класифікувати дане зображення. Ручна перевірка та класифікація зображень може бути виснажливим завданням, особливо якщо їх величезна кількість, тому буде дуже корисно, якщо автоматизувати весь цей процес за допомогою комп'ютерного зору.

Класифікації зображень включають:

Позначення рентгенівського знімка як рак чи ні (бінарна класифікація).

Класифікація рукописної цифри (багатокласова класифікація).

Присвоєння назви фотографії обличчя (мультикласова класифікація).

Досягнення в галузі автономного водіння також є чудовим прикладом використання класифікації зображень у реальному світі. Наприклад, це дає змогу створити модель класифікації зображень, яка розпізнає різні об'єкти, такі як інші транспортні засоби, пішоходи, світлофори та покажчики на дорозі.

Класифікація елементів зображень має наступну структуру:

Попередня обробка зображення. Метою цього процесу є покращення даних (функцій) зображення шляхом придушення небажаних спотворень і

покращення деяких важливих функцій зображення, щоб наші моделі комп'ютерного бачення могли використовувати ці покращені дані для роботи.

Виявлення об'єкта. Виявлення означає локалізацію об'єкта, що означає сегментацію зображення та визначення положення об'єкта, що цікавить.

Виділення ознак і навчання. Це важливий крок, на якому статистичні або глибокі методи навчання використовуються для виявлення найцікавіших шаблонів зображення, особливостей, які можуть бути унікальними для певного класу і які згодом допоможуть моделі розрізнити різні класи. Цей процес, коли модель вивчає функції з набору даних, називається навчанням моделі.

Класифікація об'єкта. Цей крок класифікує виявлені об'єкти за попередньо визначеними класами за допомогою відповідної методики класифікації, яка порівнює шаблони зображення з цільовими шаблонами.

#### Попередня обробка зображення

Попередня обробка — це загальна назва для операцій із зображеннями на найнижчому рівні абстракції — і вхід, і вихід є зображеннями інтенсивності. Необхідність попередньої обробки зображення заключається в тому, що комп'ютери здатні обчислювати числа, але не можуть інтерпретувати зображення так, як це роблять люди. Тому при створенні класифікації елемента зображення у середовищі доповненої реальності перш за все потрібно якимось чином перетворити зображення на числа, щоб їх зрозумів комп'ютер.

Метою попередньої обробки є покращення даних зображення, яке пригнічує небажані спотворення або покращує деякі функції зображення, важливі для подальшої обробки. Також попередня обробка на віддалених пристроях зменшує навантаження користувачького приладу відображення даних у доповненій реальності, що робить відображення елемента простішим у використанні так як зменшує час затримки, до його накладання на об'єкти чи суб'єкти реальності.

Етапи попередньої обробки зображення:

Зчитування зображення;

Змінити розмір зображення;

Збільшення даних:

Сірий колір зображення;

Рефлексія;

Гаусове розмиття;

Вирівнювання гістограми;

Обертання;

Переклад.

Далі буде розглянуто покроковий розбір етапів стандарту попередньої обробки зображень для передачі їх даних у функціональний компонент розробника програмного забезпечення додатків, які можуть використовуватися у середовищі доповненої реальності.

Також для більш детального розуміння для кожного з етапів було створено реалізацію цих кроків. Це дасть змогу зрозуміти яким чином відбувається перетворення зображення.

Крок 1

Зчитування зображення

На цьому кроці потрібно зберегти шлях до набору даних зображень у змінній, а потім створити функцію для завантаження папок із зображеннями в масиви, щоб комп'ютери могли працювати з ними.

Крок 2.

Зміна розміру зображення

Деякі зображення, зняті камерою та передані в алгоритм штучного інтелекту, відрізняються за розміром, тому потрібно встановити базовий розмір для всіх зображень, які передаються в алгоритми штучного інтелекту, змінивши їх розмір.

Крок 3

Збільшення даних

Збільшення даних – це спосіб створення нових «даних» з різними орієнтаціями. Переваги цього подвійні: по-перше, це можливість генерувати «більше даних» з обмежених даних, а по-друге, це запобігає переобладнанню.

Методи збільшення даних:

Розширення сірого - зображення буде перетворено в градацію сірого (діапазон відтінків сірого від білого до чорного), комп'ютер призначить кожному пікселю значення на основі того, наскільки він темний. Усі числа поміщаються в масив, і комп'ютер виконує обчислення з цим масивом.

Віддзеркалення - суть заключається у відображенні зображення горизонтально та вертикально. Деякі фреймворки не забезпечують функцію вертикального перевертання. Але вертикальне перегортання еквівалентно обертанню зображення на 180 градусів і подальшому горизонтальному перевертання.

Гаусове - розмиття за Гаусом є результатом розмиття зображення за допомогою функції Гауса. Це широко використовуваний ефект у графічному програмному забезпеченні, як правило, для зменшення шуму зображення.

Вирівнювання гистограми - це ще один метод обробки зображень для збільшення загального контрасту зображення за допомогою гистограми інтенсивності зображення. Цей метод не потребує параметрів, але іноді він призводить до неприродного вигляду зображення.

Обертання - це ще один метод збільшення зображення. Обертання зображення може не зберегти його початкові розміри (залежно від кута, на який був вибраний для його обертання)

Перевертання - передбачає лише переміщення зображення вздовж напрямків X або Y (або обох). Цей метод збільшення дуже корисний, оскільки більшість об'єктів можна розташувати майже в будь-якому місці зображення. Це змушує наш екстрактор функцій шукати всюди.

Усе це у сукупності дає можливість розробнику покращити продуктивність свого продукту, що у свою чергу збільшить зацікавленість користувачів у його подальшому використанні.

Після обробки зображення комп'ютерній програмі необхідно розпізнати його тип. Тому буде логічним детальніше розглянути методи класифікації елементів зображень на основі статистичних класифікаторів машинного

навчання, таких як Support Vector Machine (SVM) і Decision Tree, а також архітектур глибокого навчання, таких як Convolutional Neural Networks.

Характеристики можна виділити за допомогою згорткової нейронної мережі, яка також буде рахуватися як один із класифікаторів. Моделі глибокого навчання досягли найвищих результатів у процесі вилучення ознак.

Для того, щоб краще зрозуміти роботу статистичних класифікаторів машинного навчання, на яких у майбутньому будуть базуватися використання алгоритмів класифікації елементів зображень у середовищі доповненої реальності, потрібно перш за все виділити основні та детальніше їх розглянути та зрозуміти принципи їх використання.

### -SVM

Це контрольований алгоритм машинного навчання, який використовується як для проблем регресії, так і для класифікації. Коли використовується для цілей класифікації, він розділяє класи за допомогою лінійної межі (рис 2.1).

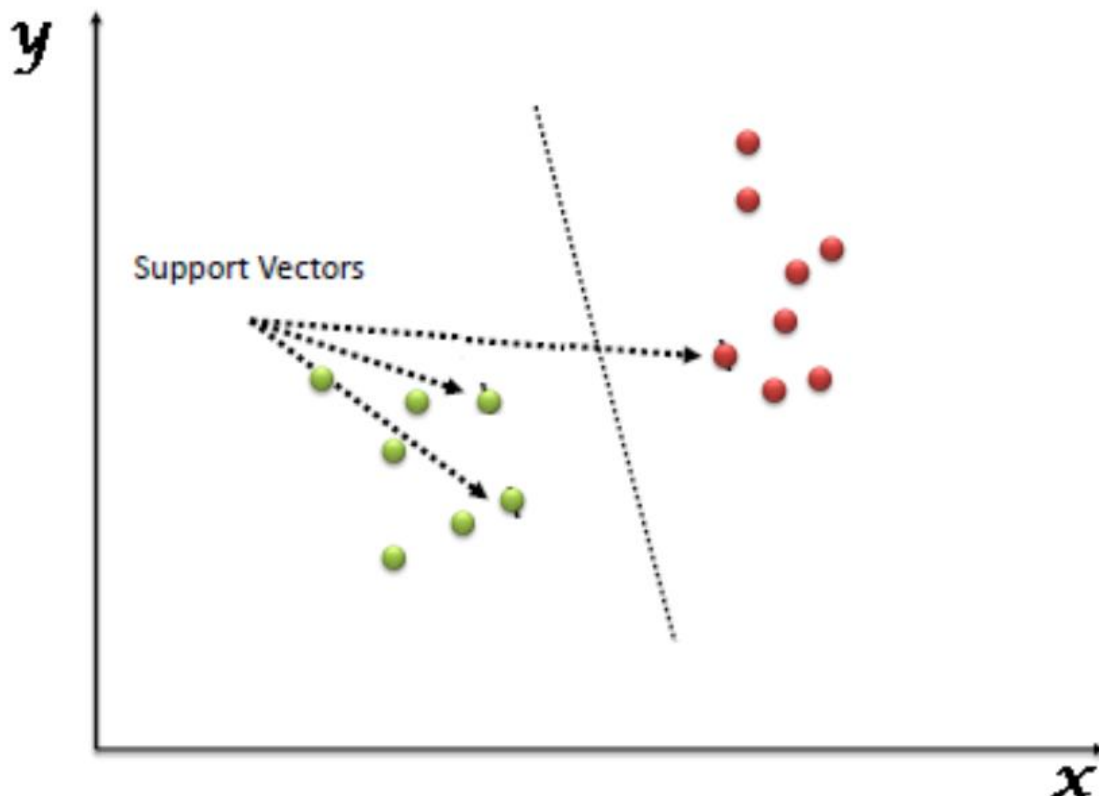


Рисунок 2.1 – SVM. Розділ класів за допомогою лінійної межі.

Класифікатор створює гіперплощину або набір гіперплощин у великому просторі, і розділення між двома класами досягається гіперплощиною, яка має найбільшу відстань до найближчої точки тренувальних даних будь-якого класу. Реальна потужність цього алгоритму залежить від використовуваної функції ядра.

Найбільш часто використовувані ядра:

Лінійне ядро;

Ядро Гауса;

Поліноміальне ядро.

- Дерева рішень

Це також керований алгоритм машинного навчання, який, за своєю суттю, є лише структурою даних дерева, використовуючи пару операторів if/else для вибраних функцій.

Дерева рішень базуються на ієрархічному методі на основі правил і дозволяють приймати та відхиляти мітки класів на кожному проміжному етапі/рівні(рис. 2.2):

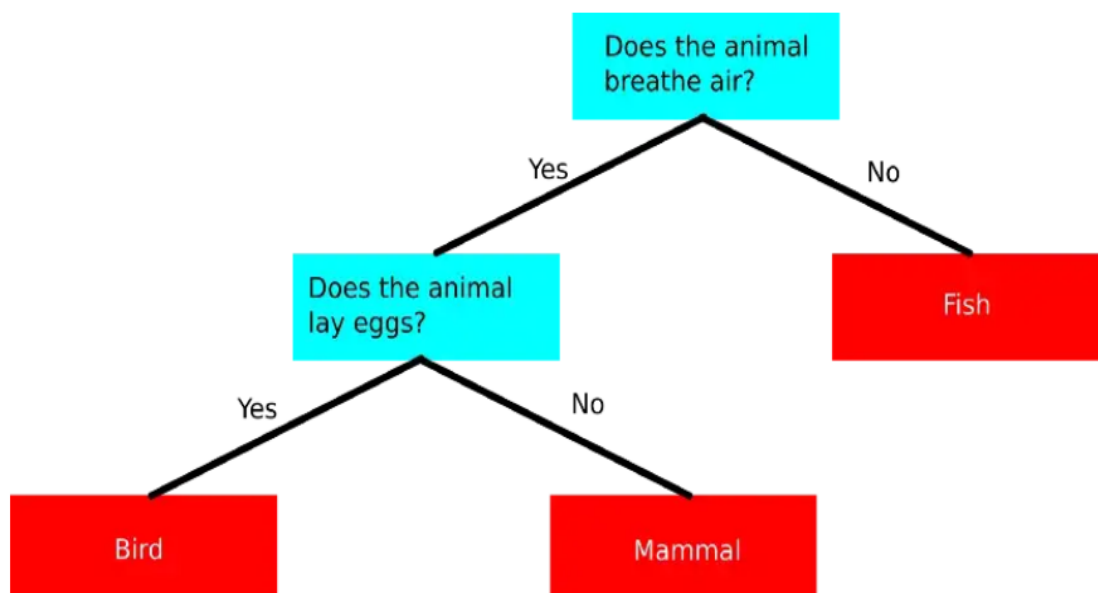


Рисунок 2.2 – Блок схема реалізації дерева рішень.

Цей метод складається з 3 частин

Розбиття вузлів;

Знаходження кінцевих вузлів;

Розподіл мітки класу для термінального вузла.

- KNN

K-найближчий сусід на сьогоднішній день є найпростішим алгоритмом машинного навчання.

Цей алгоритм спирається на відстань між векторами ознак і класифікує невідомі точки даних, знаходячи найпоширеніший клас серед k-найближчих прикладів. Роботу алгоритму зображено на рис. 2.3.

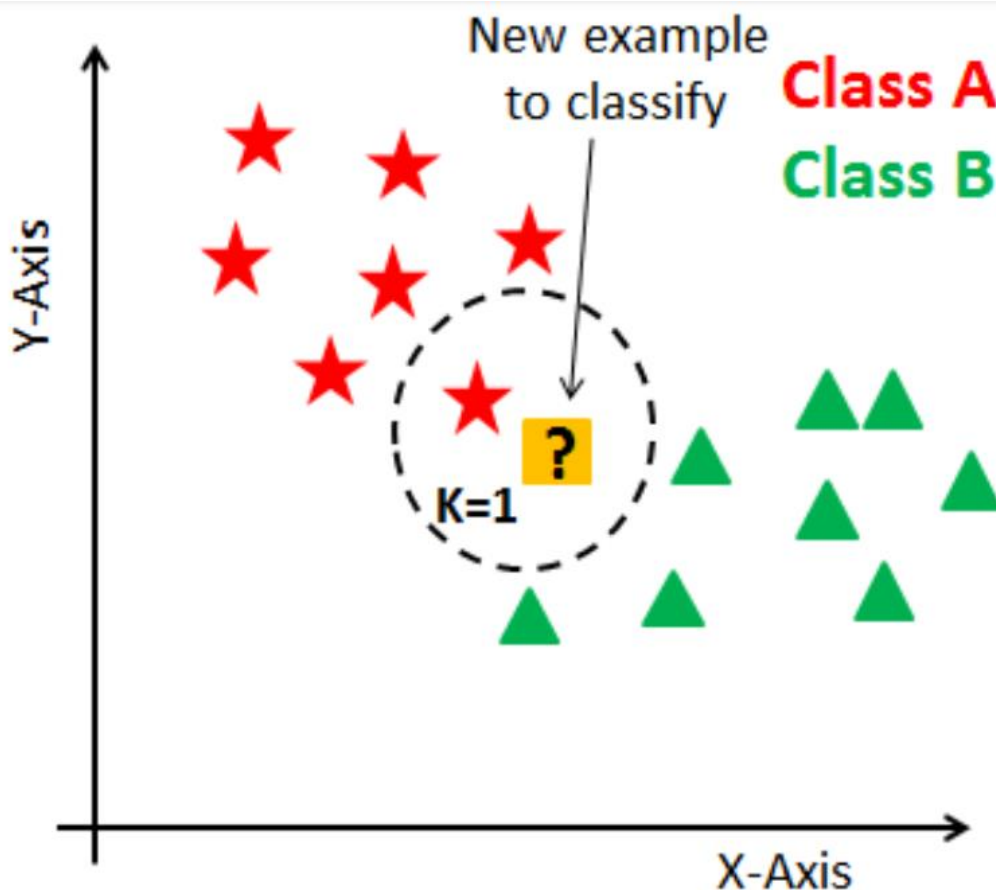


Рисунок 2.3 - Робота алгоритму KNN

Тут ми бачимо дві категорії зображень і що кожна точка даних у кожній відповідній категорії згрупована відносно близько одна до одної в n-вимірному



просторі.

Щоб застосувати класифікацію k-найближчого сусіда, нам потрібно визначити метрику відстані або функцію подібності. Загальні варіанти включають евклідову відстань і манхеттенську відстань.

## ANN

Штучні нейронні мережі є статистичними алгоритмами навчання, які використовуються для різноманітних завдань, від відносно простих завдань класифікації до комп'ютерного зору та розпізнавання мовлення.

Штучні нейронні мережі реалізовано як систему взаємопов'язаних елементів обробки, званих вузлами, які функціонально аналогічні біологічним нейронам. З'єднання між різними вузлами мають числові значення, які називаються вагами, і змінюючи ці значення систематично, мережа зрештою може наближено шукану функцію(рис 2.4).

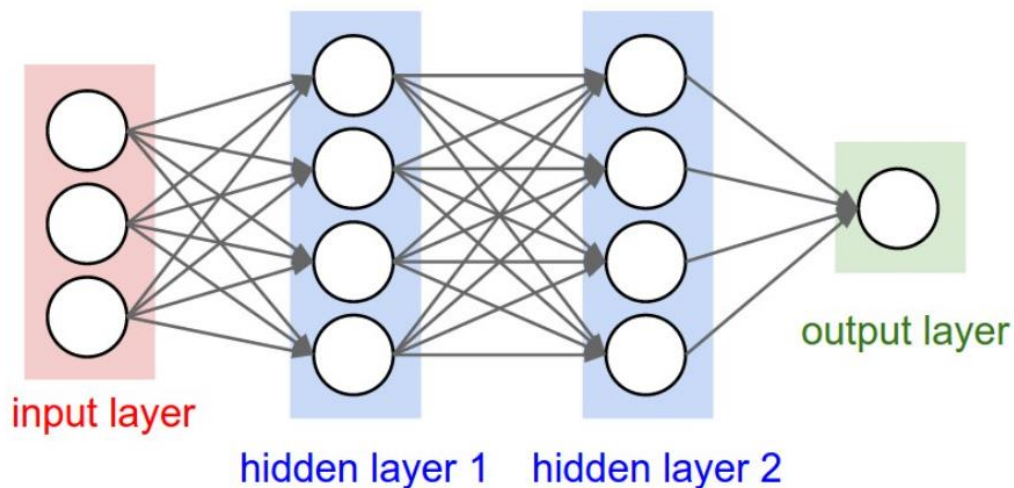


Рисунок 2.4 - Нейронні мережі як нейрони в графах.

Приховані шари можна розглядати як окремі детектори ознак, які розпізнають все більш складні шаблони в даних у міру того, як вони поширюються по мережі. Наприклад, якщо мережі дається завдання розпізнати обличчя, перший прихований шар може діяти як детектор ліній, другий прихований приймає ці лінії як вхідні дані та об'єднує їх, щоб утворити ніс,

третій прихований шар приймає ніс і зіставляє його оком і так далі, поки нарешті не буде створено все обличчя. Ця ієрархія дозволяє мережі зрештою розпізнавати дуже складні об'єкти.

- Згорткові нейронні мережі

Згорткові нейронні мережі (CNN) — особлива архітектура штучних нейронних мереж (рис. 2.5). CNN використовує деякі свої особливості зорової кори головного мозку і тому досяг найсучасніших результатів у задачах комп'ютерного зору.

Згорткові нейронні мережі складаються з двох дуже простих елементів, а саме згорткових шарів і шарів об'єднання.

Незважаючи на простоту, існує майже нескінченна кількість способів упорядкувати ці шари для певної проблеми комп'ютерного зору.

Елементи згорткової нейронної мережі, такі як згортковий шар і шари об'єднання, відносно прості для розуміння.

Складна частина використання згорткових нейронних мереж на практиці полягає в тому, як розробити модельні архітектури, які найкраще використовують ці прості елементи.

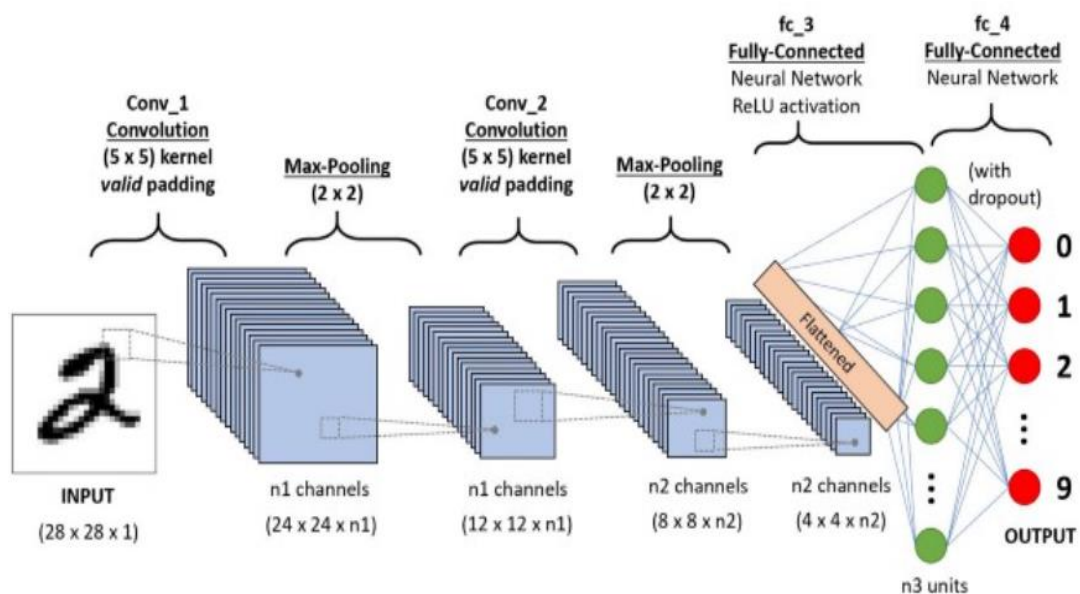


Рисунок 2.5 - архітектура згорткових штучних нейронних мереж

Отже для визначення найкращого методу класифікації елементів

зображень можна провести дослідження, яке покаже продуктивність роботи кожного з методів і зрівняти їх у таблиці.

Таблиця 1 – Порівня продуктивності методів класифікації зображень

Класифікатор	Точність у процентах	Точність	Виклик	ROC
SVM	85.68%	0.86	0.87	0.86
Decision Trees	84.61%	0.85	0.84	0.82
KNN	86.32%	0.86	0.86	0.88
ANN	83.10%	0.88	0.87	0.88
CNN	91.11%	0.93	0.89	0.97

З таблиці продуктивності можна зробити висновок, що згорткові нейронні мережі забезпечують найкращі результати в задачах комп'ютерного зору.

### 2.3 Алгоритмічне забезпечення методу класифікації елементів зображень у середовищі доповненої реальності

Основна можливість відображення додаткових даних на реальних об'єктах є визначення вертикальних та горизонтальних площин фізичного світу.

У задачах популяризації ця технологічна можливість може лежати в основі ДР додатків та використовується в алгоритмах уточненого позиціонування камери з ціллю підвищення продуктивності позначення локальних пересувань комп'ютерної машини.

Аналіз площин розділяється на дві категорії [18].

Markerless підхід

Це метод, який розташовує віртуальний 3D-вміст фізичній реальності шляхом виявлення площини, яка наперед є завантажена у пам'ять.Ця технологія подібна до технології розпізнавання QR - коду.

SLAM - Simultaneous Localization and Mapping.

Другий і більш функціональний метод у аналізі площини. Це модифікація відслідковування площин без застосування маркерів. Для розуміння даної технології проаналізуємо більш детально її модель.

Ціль доповненої реальності полягає у тому щоб вставляти віртуальний контент у фізичний світ, а також зберігати орієнтацію та алокацію даного контенту при переміщенні девайсу в просторі. Для Vuforia та ATK(ARToolKit) головним процесом для цього є зчитування кадрів з камери, для кожного шару обробляється зображення та вилучаються точкові ознаки. Характеристиками можуть бути лише ті точки на даних графічного значення. Це може бути градус геометричного кута об'єкта, сторона текстурованого фрагмента, і так далі. Використовується багато різних способів визначення таких ознак, які уже були розглянуті раніше. Одним з підходів для цього є SIFT. Приклад визначення ознак зображення наведено на рисунку 2.6.



Рисунок 2.6 – Визначення ознак на зображенні

Після визначення всіх особливостей кадру аналізується зміна їх положення на зображенні при русі пристрою. Знаючи вектор і величину зміщення смартфона, можна визначити тривимірне положення всіх знаків. При збільшенні кількості кадрів зі смартфона при русі пристрою значення тривимірних положень знаків наближається до реальних. Знаючи всі тривимірні положення знаків, створюється тривимірна сітка, утворена

з'єднанням сусідніх знаків. Це дозволяє отримати вектори нормалей кожного об'єкта.

Щоб визначити поверхні сучасного світу, необхідно здійснити кластеризацію ознак. Це робиться шляхом поєднання суміжних знаків зі схожими нормалями. Таким чином вибираються ті точки простору, які лежать в одній площині. Процес визначення поверхонь за розрахунковими ознаками продемонстровано на рисунку 2.7.

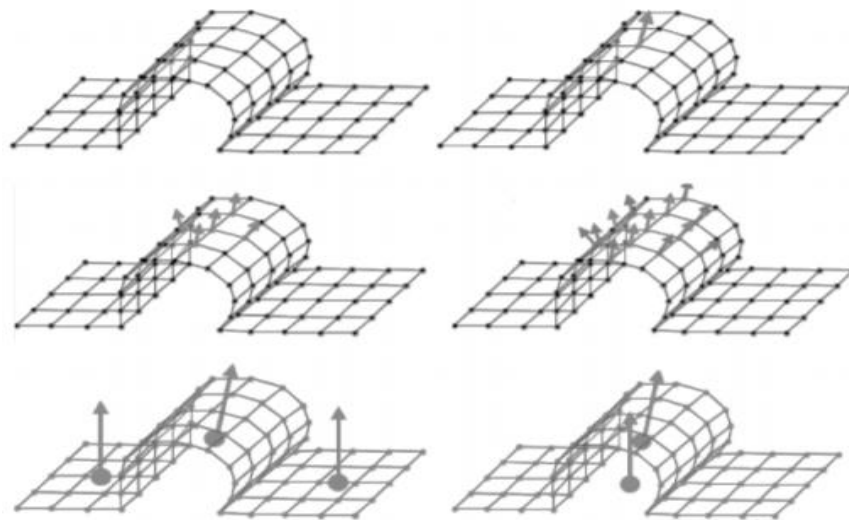


Рисунок 2.7 – Визначення реальних поверхонь за ознаками

Описаний метод дозволяє визначати однорідні поверхні в просторі і представляє їх у вигляді площини з координатами центру і краю. Незважаючи на швидкий розвиток цієї технології на мобільних платформах, вона має певні обмеження та недоліки.

Серед технічних обмежень виділяють необхідність гарного освітлення досліджуваної поверхні та якості її фактури. Тобто поверхня повинна добре розпізнаватися. Головними вимогами до мобільного пристрою для реалізації технології є висока якість камери, оскільки при поганій якості зображення технологія незадовільно визначає ознаки досліджуваного простору.

Таким чином визначено, що визначення ознак зображення є необхідною умовою для технології розпізнавання поверхні. Для цього потрібна реалізація алгоритму SIFT.

SIFT — це алгоритм комп'ютерного зору для виявлення й опису локальних особливостей зображення. Необхідною умовою якісного розпізнавання ознак є можливість однозначного виявлення навіть при зміні масштабу, освітленості та шуму вихідного зображення. Це досягається шляхом вибору лише тих точок, які знаходяться на висококонтрастних ділянках об'єкта, які можуть відображати його межі або зміни текстури на ньому. Іншою необхідною характеристикою цих ознак є постійні взаємні положення між ідентифікованими.

Отже, розпізнаванню підлягають об'єкти, які з часом змінюють свою форму. Найпростіший приклад - двері. Коли лише чотири кути дверей беруться як ознаки, розпізнавання є успішним незалежно від позиції спостереження. Однак, якщо елементи дверної рами включені в масив даних, розпізнавання буде виконано з помилками, якщо двері відчинятимуться та закриватимуться під час виконання алгоритму, оскільки відносне розташування елементів змінюватиметься в реальному світі. Однак на практиці алгоритм SIFT використовує набагато більше можливостей при обчисленнях, що призводить до мінімізації впливу помилок на результат. Оскільки дескриптор ознак алгоритму SIFT стійкий до лінійного масштабування, зміни просторової орієнтації та освітленості, описаний алгоритм успішно застосовується до випадків із зашумленим зображенням або частковим перекриттям досліджуваного об'єкта. Алгоритм SIFT виконується послідовністю наступних операцій:

- 1) масштабно-інваріантне визначення ознак. Ця операція виконується за методом Лоу, коли розташування ключових точок визначається мінімумом і максимумом функції Гауса [19]. Таким чином, вибір ознак вхідного зображення перетворюється на набір векторів ознак. Крім того, кожен із отриманих векторів є інваріантним до переміщень, масштабування, повороту, часткового геометричного перетворення або перекриття об'єктів, які досліджуються;

- 2) порівняння ознак та їх індексація. Цей етап алгоритму необхідний для порівняння ознак, отриманих під час цього циклу обчислень, з індексованими

ознаками, визначеними в попередньому циклі. Таким чином отримують дані про рух кожної ознаки об'єкта в часі. Зіставлення відбувається шляхом визначення найближчих сусідніх об'єктів за допомогою методу Best-bin-first, який дозволяє виконувати обчислення з високою ймовірністю за допомогою обмеженої кількості обчислень. Тобто найкращий кандидат для відповідності в деяких функціях визначається шляхом визначення його найближчого сусіда в базі даних ключових точок із попередніх зображень. Близькість сусідів один до одного вважається евклідовою відстанню. Ймовірність правильного збігу ознак визначається обчисленням відношення відстані до найближчого сусіда до другої секунди.

Для більш коректного рішення описані співвідношення зі значенням більше 0,8 відкидаються. На практиці ця умова усуває до 90% невірних порівнянь, але в той же час усуває 5% правильних рішень. Найбільш оптимальними параметрами для виконання описаного алгоритму є визначення 10 000 ключових точок зображення, причому для кожної ознаки зв'язок завершується після перевірки 200 сусідів.

1) ідентифікація кластерів за допомогою перетворень Гофа. Цей етап алгоритму використовується для об'єднання всіх заздалегідь визначених ознак кластера з подальшим їх порівнянням з конкретною мозаїкою досліджуваного об'єкта. Для цього використовуються перетворення Хофа [20], які виявляють кластери ознак за допомогою процедури голосування. Коли виявлено кластери ознак із голосуванням за ту саму позицію об'єкта, ймовірність правильної інтерпретації кластера зростає. Після цього запис про ймовірне положення, орієнтацію та масштаб зберігається в хеш-таблиці. Таким чином, кожна ключова точка SIFT визначає двовимірну площину, масштаб і орієнтацію, і кожна ключова точка в базі даних має запис із своїми параметрами щодо зображення, на якому вона була знайдена. Виконуючи набір описаних операцій, можна отримати наближення до повного простору позиції з 6 ступенями свободи (три кроки для позиції, три кроки для орієнтації Ейлера);

2) перевірка розрахункової моделі методом найменших квадратів. Цей

етап необхідний для перевірки кожного кластера на правильність його формування. Щоб визначити, чи правильна структура кластера, використовується лінійний метод найменших квадратів.

3) виявлення викидів. Крок в алгоритмі SIFT, який використовується для виявлення та вилучення надлишкових або неправильно сформованих кластерів із моделі. Операцію вилучення можна здійснити, перевіривши відповідність функцій кожного зображення та моделі. Якщо розв'язується метод найменших квадратів, кожен збіг має збігатися не більше ніж на половину інтервалу помилок, який використовувався в обчисленнях у перетвореннях Гафа. Після цього викиди видаляються, рішення методом найменших квадратів знову перераховує всі інші ключові точки, і процес виявлення повторюється. Якщо після виділення викидів у кластері залишилося менше 3 точок, він вилучається. Остаточне рішення про прийняття чи неприйняття гіпотетичної моделі базується на детальній імовірнісній моделі. Цей метод визначення обчислює очікувану кількість помилкових збігів позиційної моделі. Потім байєсівський аналіз визначає ймовірність присутності об'єкта на основі фактичної кількості знайдених збігів ознак. Модель вважається прийнятною, якщо остаточна ймовірність її правильної інтерпретації перевищує 98%.

Описаний розпізнавальний алгоритм об'єктів за моделлю Лоу дає високоточні результати. Істотними перешкодами в правильному обчисленні результату є різка зміна величини освітленості поверхні досліджуваного об'єкта і здатність цього об'єкта виробляти значні зміни в часі. Модель роботи методу класифікації елементів зображень у середовищі доповненої реальності розроблена на основі вище викладеного матеріалу та проаналізувавши його можна зрозуміти, що він підходить для роботи з маркерною системою розпізнавання зображення. Початок роботи додатку відбувається з ініціалізації камери, після чого у він отримує відеопотік з якого визначається поверхня, якщо вона є у базі даних додатку, то він продовжує свою роботу, а якщо ні, то повертається до етапу отримання відеопотку з камери пристрою. Таким чином, після визначення того чи поверхня з відеопоток є у базі даних девайсу,



відбувається розміщення об'єкту на поверхні, тобто на маркері, після того, як операція успішно відбулась починається візуалізація поверх реального світу на екрані пристрою вітуальних моделей, після успішної візуалізації відбувається опрацювання вводу користувача і далі робота буде завершена, якщо ж ні то додаток повертається на стадію ініціалізації камери.

## Висновки до розділу 2

1. Визначено, що існуючі додатки, які мають основні тенденції в розвитку технології доповненої реальності. Ця технологія в наш час дуже бурхливо розвивається і може знайти застосування в багатьох областях. Аналіз додатків показав, що зараз вони мають, в основному, або розважальний, або вузькоспрямований характер. Основна увага в таких додатках приділяється їх можливостям при взаємодії з навколишнім середовищем і користувачем. Також було визначено, що найоптимальнішим середовищем для розробки системи є Vuforia.

2. Визначено, що продуктивність різних сучасних методів класифікації елементів зображень мають різну швидкість обробки, з чого було зроблено висновки, що згорткові нейронні мережі забезпечують найкращі результати в задачах комп'ютерного зору.

3. Визначено алгоритми забезпечення роботи методу за допомогою основних алгоритмів, які б підтримували його повноцінне функціонування у середовищі доповненої реальності, а саме досліджено алгоритм SIFT, який є найкращим варіантом у поєднанні з CNN.

### 3. ПРАКТИЧНЕ ВИКОРИСТАННЯ МЕТОДУ КЛАСИФІКАЦІЙ ЕЛЕМЕНТІВ ЗОБРАЖЕНЬ У СЕРЕДОВИЩІ ДОПОВНЕНОЇ РЕАЛЬНОСТІ

#### 3.1 Апробація розробленого методу класифікації елементів зображення.

Метод класифікації зображень з використанням CNN. Архітектура запропонованого методу CNN містить згорткові шари, максимальний шар об'єднання, повністю зв'язаний шар і двошаровий перцептрон як класифікатор. Експерименти показали, що CNN ефективніші за інші методи. Таким чином, результати, представлені в дослідженні, підтверджують необхідність використання CNN у класифікації зображень. Результати запропонованого методу є задовільними, але велика довжина 4096-вимірному вектору ознак є недоліком і робить запропонований метод є непрактичним у реальній ситуації, особливо коли недостатньо пам'яті. Нові методи виділення ознак і злиття функцій є іншими активними напрямками досліджень, які допоможуть підвищити продуктивність практично будь-яких завдань комп'ютерного зору.

Представлена раніше нова методологія для отримання зображень різних сцен за допомогою масштабно-інваріантного перетворення ознак (SIFT) Метою запропонованого дескриптора є створення високодискримінаційних ознак для опису вмісту зображення. За допомогою цього дескриптора ознаки, які виділяються для опису вмісту зображення, є дуже дискримінаційними. Навіть було показано, що колірний простір, з якого виділяються характеристики зображення, дуже важливий для остаточного виконання завдання класифікації зображення.

Колірний простір Lab є більш інформативним, ніж інші колірні простори, коли необхідно витягти вейвлет-функцію із зображень. Експериментальні результати ілюструють перевагу запропонованого підходу. На основі вищезазначених досліджень ми вирішили використати функцію, отриману CNN, і, щоб отримати кращі результати, ми спробували запровадити новий метод класифікації у середовищі доповненої реальності.

У цьому розділі пояснюється, як запропонований метод працює для класифікації та пошуку зображень. На рисунку 3.1 зображено загальну схему

запропонованого методу. Згідно з цим малюнком запропонований метод включає п'ять основних етапів:

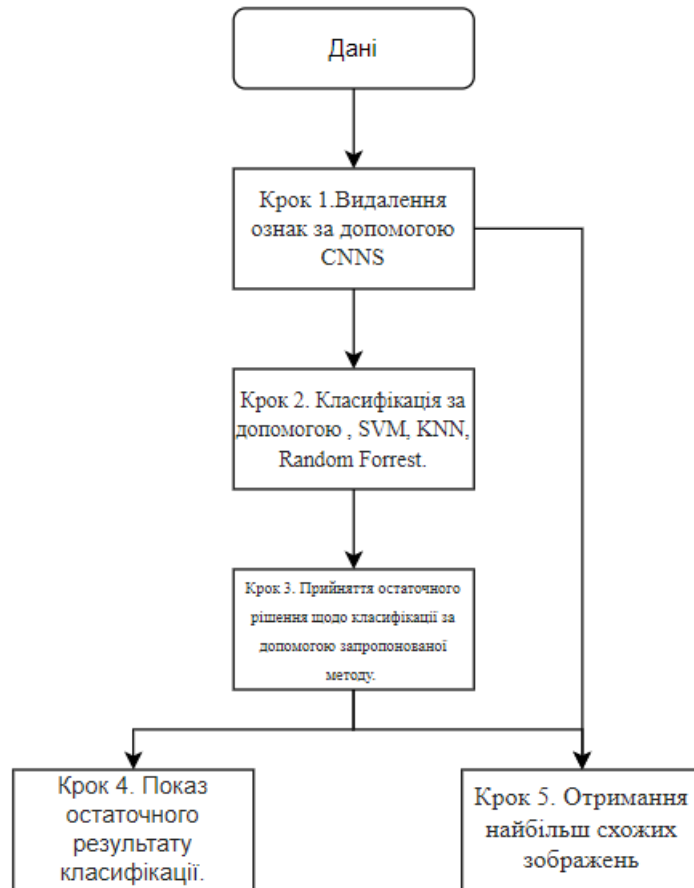


Рисунок 3.1 - Загальна схема запропонованого способу

На першому кроці метод на основі CNN використовується для виділення функцій.

На наступному кроці витягнуті функції класифікуються за допомогою вихідних класифікаторів CNN, SVM, KNN і випадкового лісу.

На третьому та четвертому етапах запропонований метод використовується для прийняття остаточного рішення щодо класифікації.

Потім відображається кінцевий результат класифікації.

На останньому кроці зображення, найбільш схожі на зображення запиту, витягуються відповідно до підходу пошуку зображень на основі вмісту.

Причини, чому запропонований метод працює краще, ніж інші найсучасніші методи, включають пропозицію нової стратегії голосування

комітету для класифікації зображень, використання шарів Softmax замість лінійних шарів для виділення ознак і пропозицію динамічних векторів ознак у термінах розмірів. Далі описані переваги методу. Крім того, порівняння результатів запропонованого методу з іншими сучасними методами, пов'язаними з цими причинами, також потрібно детально навести. На рисунку 3.2 детально зображено графічну форму запропонованого способу. Як показано, він включає п'ять кроків, кожен з яких детально описано в наступних підрозділах.

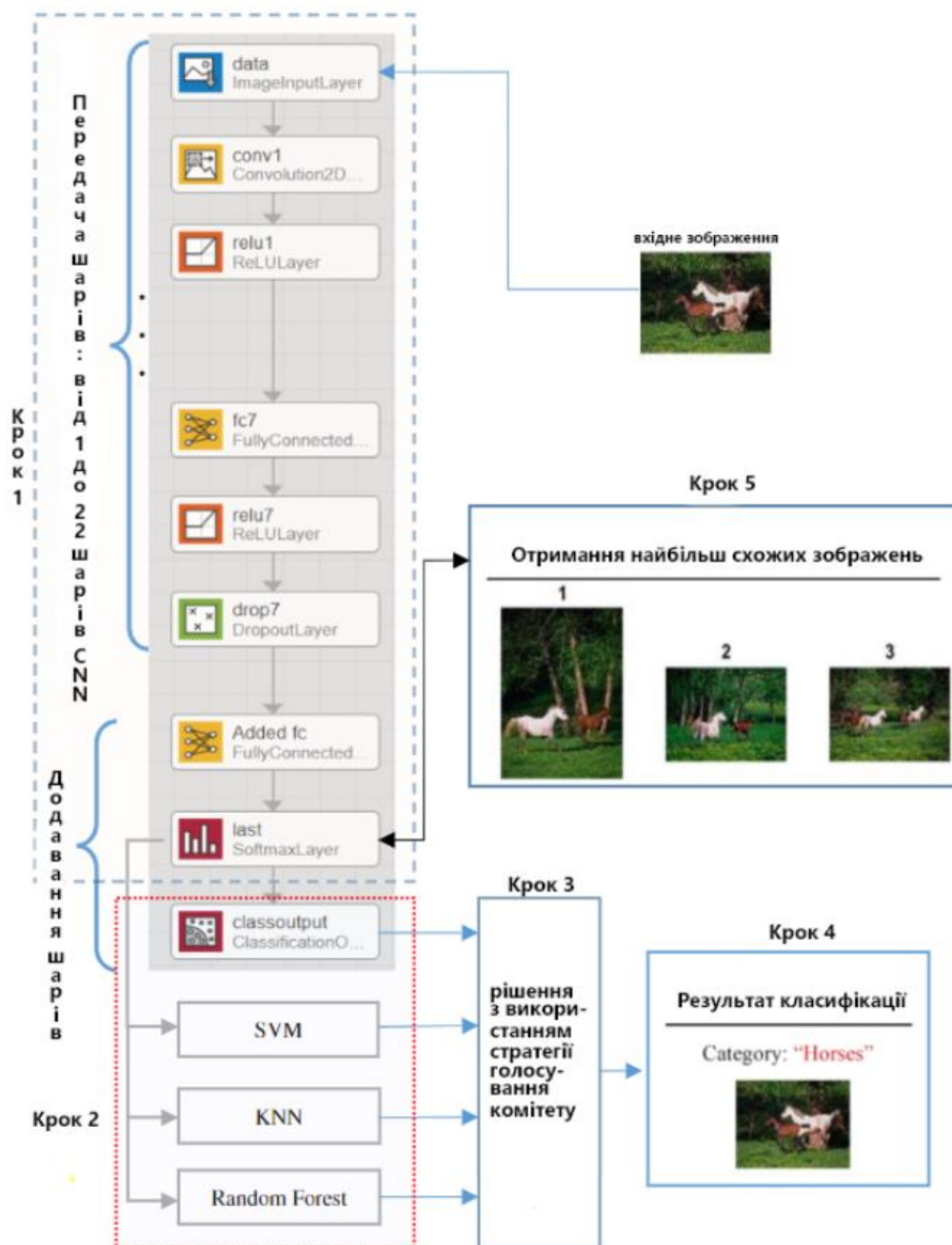


Рисунок 3.2 - Графічний вигляд запропонованого методу в деталях

Для глибоких нейронних мереж у деяких випадках може бути недостатньо даних для навчання мережі або створення даних з мітками може бути дорогим і трудомістким. Відповідно, трансферне навчання може бути застосоване для закріплення знань, отриманих у попередніх умовах. У цьому дослідженні використано трансфертне навчання, щоб виділити особливості зображення та навчити запропонованому методу класифікації. На етапі 1 на малюнку 2, було передано 22 шари CNN, а такі шари, як повне підключення, Softmax і результат класифікації, додано в кінці переданих шарів. Розмір доданих шарів встановлено відповідно до категорій кожного набору даних.

Згідно з кроком 2 на рисунку 3.2, на етапі класифікації SVM, KNN, випадковий ліс і вихідний рівень класифікації CNN класифікують витягнуті характеристики зображень після проходження рівня Softmax. Потім результати класифікаторів надсилаються до стратегії голосування комітету для остаточного прийняття рішення. У наступному підрозділі детально описано стратегію голосування комітету. У просторах ознак, якщо ознаки не надто хаотичні або ми можемо запобігти їх хаотичності, класифікатори в більшості випадків зможуть класифікувати їх належним чином. Крім того, оскільки номер ознаки дуже добре пояснює вхідне зображення серед векторів ознак, класифікатору легше розрізнити таке зображення. Отже, шар Softmax надає нам обидва варіанти, і ми вважаємо їх корисними, оскільки функції нормалізуються шляхом проходження через шар Softmax. Такі нормалізації не тільки запобігають створенню хаотичного простору ознак, але й виділяють жирним шрифтом число вектора ознак. Це корисно для класифікації або створення ефективного набору даних ознак, оскільки вектор ознак перетворюється на відмінний вектор для класифікаторів. Тому ми б віддали перевагу використанню результатів Softmax замість шару aligner. Порівняння результатів у таблицях 8 і 9 показує, що використання результатів Softmax ефективніше, ніж лінійний шар для обох класифікацій зображень. і задоволений пошук зображень.

У більшості випадків вирішення завдань класифікації або розпізнавання за допомогою найсучасніших методів існує один багатокласовий класифікатор, і для даних тестових даних передбачення навченого класифікатора є хибним або істинним. Однак у запропонованому методі ці два випадки розширено на 15 випадків для підвищення точності. Ці 15 випадків наведено в рисунку 3.3. Зелений колір у таблиці 1 означає вірний прогноз, а інші кольори означають помилковий прогноз. Крім того, якщо два кола мають однаковий колір, це означає, що два класифікатори мають однаковий прогноз.

Status	Base classifier	Auxiliary classifier 1	Auxiliary classifier 2	Auxiliary classifier 3	Vote checking	Final prediction
1	(a, ●)	(a, ●)	(a, ●)	(a, ●)	a = 4	(a, ●)
2	(a, ●)	(a, ●)	(a, ●)	(b, ●)	a = 3, b = 1	(a, ●)
3	(a, ●)	(a, ●)	(b, ●)	(b, ●)	a = 2, b = 2	(a, ●)
4	(a, ●)	(a, ●)	(b, ●)	(c, ●)	a = 2, b = 1, c = 1	(a, ●)
5	(a, ●)	(b, ●)	(b, ●)	(b, ●)	a = 1, b = 3	(b, ●)
6	(a, ●)	(b, ●)	(b, ●)	(c, ●)	a = 1, b = 2, c = 1	(b, ●)
7	(a, ●)	(b, ●)	(d, ●)	(c, ●)	a = 1, b = 1, c = 1, d = 1	(a, ●)
8	(a, ●)	(b, ●)	(b, ●)	(b, ●)	a = 1, b = 3	(b, ●)
9	(a, ●)	(b, ●)	(b, ●)	(a, ●)	a = 1, b = 3	(a, ●)
10	(a, ●)	(b, ●)	(b, ●)	(c, ●)	a = 1, b = 2, c = 1	(b, ●)
11	(a, ●)	(b, ●)	(a, ●)	(a, ●)	a = 3, b = 1	(a, ●)
12	(a, ●)	(b, ●)	(a, ●)	(d, ●)	a = 2, b = 1, d = 1	(a, ●)
13	(a, ●)	(b, ●)	(c, ●)	(d, ●)	a = 1, b = 2, c = 1, d = 1	(a, ●)
14	(a, ●)	(a, ●)	(a, ●)	(a, ●)	a = 4	(a, ●)
15	(a, ●)	(d, ●)	(c, ●)	(e, ●)	a = 1, c = 1, d = 1, e = 1	(a, ●)

Рисунок 3.3 - Різні випадки, які трапляються в стратегії голосування комітету

Status	Base classifier	Auxiliary classifier 1	Auxiliary classifier 2	Auxiliary classifier 3	Vote checking	Final prediction
1	(a, ●)	(a, ●)	(a, ●)	(a, ●)	a = 4	(a, ●)
2	(a, ●)	(a, ●)	(a, ●)	(b, ●)	a = 3, b = 1	(a, ●)
3	(a, ●)	(a, ●)	(b, ●)	(b, ●)	a = 2, b = 2	(a, ●)
4	(a, ●)	(a, ●)	(b, ●)	(c, ●)	a = 2, b = 1, c = 1	(a, ●)
5	best: (a, ●)	(b, ●)	(b, ●)	(b, ●) → (a, ●)	a = 2, b = 2	(a, ●)
6	best: (a, ●)	(b, ●)	(b, ●)	(c, ●) → (a, ●)	a = 1, b = 2, c = 1	(a, ●)
7	(a, ●)	(b, ●)	(d, ●)	(c, ●)	a = 1, b = 1, c = 1, d = 1	(a, ●)
8	(a, ●)	(b, ●)	(b, ●)	(b, ●)	a = 1, b = 3	(b, ●)
9	(a, ●)	(b, ●)	(b, ●)	(a, ●)	a = 2, b = 2	(a, ●)
10	(a, ●)	best: (b, ●)	(b, ●)	(c, ●) → (b, ●)	a = 1, b = 3	(b, ●)
11	(a, ●)	(b, ●)	(a, ●)	(a, ●)	a = 3, b = 1	(a, ●)
12	(a, ●)	(b, ●)	(a, ●)	(d, ●)	a = 2, b = 1, d = 1	(a, ●)
13	(a, ●)	(b, ●)	(c, ●)	(d, ●)	a = 1, b = 1, c = 1, d = 1	(a, ●)
14	(a, ●)	(a, ●)	(a, ●)	(a, ●)	a = 4	(a, ●)
15	(a, ●)	(d, ●)	(c, ●)	(e, ●)	a = 1, c = 1, d = 1, e = 1	(a, ●)

Рисунок 3.4 - Запропонована стратегія голосування комітету

Оскільки є чотири класифікатори та багатокласова класифікація, необхідно подбати про важливий випадок. Це відбувається, коли кількість голосів рівна. Перший випадок виникає, коли кожен класифікатор передбачає

різну категорію, а другий випадок виникає, коли два класифікатори передбачають ту саму категорію, а інші два класифікатори також передбачають ту саму категорію. У таких випадках, щоб прийняти остаточне рішення, ми покладаємося на класифікатор, який мав найвищу точність на етапі навчання. У таблиці 1, коли трапляються випадки 5, 6, 8 і 10, запропонована стратегія голосування переглядає знову ситуація. Причина полягає в тому, що в цих випадках передбачення базового класифікатора відрізняється від передбачення допоміжних класифікаторів. Отже, раціонально, що в цих випадках ми не повинні покладатися на базовий класифікатор. Для вирішення цих ситуацій голосування в комітеті буде змінено, як у рисунку 3.4. Результати експериментів у розділі 4 довели, що наша гіпотеза щодо випадків 5, 6, 8 і 10 була вірною. Отже, використовуючи рисунок 3.4, точність нашої стратегії голосування підвищується. Псевдокод запропонованого методу класифікації зображень представлений на рисунку 3.5 наступним чином:

```

Function Decision.
Input: C_labels and C_weights
Return: FinalLabel
1.  argBase ← C_labels.LC ← zeros (length of C_labels);
2.  for i ← 1 until length of C_labels
3.    LC[i] ← Count the number of C_labels [i] events;
4.  end
5.  LC ← dot multiplication (LC, C_weights);
6.  [firstMax, secondMax] ← find first and second ...
7.    maximum number in LC;
8.  if firstMax and secondMax are equal to 2 or ...
9.    firstMax is equal to 1 then
10.   C_weights ← [2, 1, 1, 1];
11. Call Decision(argBase, C_weights)
12. else best_label ← find best label based on all ...
13.   classifiers' training accuracy;
14.   worst_label ← find worst label based on ...

15.   helper classifiers' training accuracy;
16.   worst_label ← best_label; Comment: change voting
17.   LC ← update LC based on change voting;
18. end
19. ind ← find maximum number's index in LC;
20. FinalLabel ← C_labels[ind];
21. Return FinalLabel

```

Рисунок 3.5 – Алгоритм запропонованого методу класифікації зображень представлений

У цьому псевдокодi мітки класу та ваги класу є результатами



передбачення, зробленими класифікаторами, і вагами для голосування кожного класифікатора відповідно. Слід зазначити, що вага класу становить 1 на початку процесу прийняття рішення, а після цього його можна змінити відповідно до таблиці 2 для деяких класифікаторів.

У цьому дослідженні ми використовуємо підхід SIFT для пошуку зображень, найбільш схожих на зображення запиту. Більшість моделей класифікації та пошуку, які базуються на згорткових нейронних мережах, використовують ознаки, отримані останнім повністю підключеним рівнем CNN. Ці вектори ознак мають постійні довжини. У цій роботі ми використали вектори ознак довжиною  $1 \times 10$ ,  $1 \times 100$ ,  $1 \times 102$  та  $1 \times 67$  для опису зображень Corel-1k, Corel-10k, Caltech-101 та Scene-67 відповідно. Результати деяких експериментів, які показують, що такі довжини є оптимальними. Оптимальна довжина зменшує потребу в пам'яті для роботи над об'єктами та створення набору даних об'єктів. Ця проблема може принести переваги.

### 3.2 Експериментальні результати та порівняння

У цьому розділі надано детальні пояснення щодо наборів даних, використаних у статті, і результатів порівняння між запропонованими в статті методами та найсучаснішими методами класифікації та пошуку зображень.

У цьому дослідженні були використані такі набори даних.

Corel-1k: цей набір даних містить 10 різних категорій, включаючи африканських людей, коней, квіти тощо. Кожна категорія містить 100 зображень розміром  $192 \times 128$  або  $128 \times 192$  у форматі JPEG.

Corel-10k: Набір даних містить 100 категорій, включаючи пляж, рибу, квіти тощо. Кожна категорія має 100 зображень розміром  $192 \times 128$  або  $128 \times 192$  у форматі JPEG.

Caltech-101: Цей набір даних містить 914 зображень. у 102 різних категоріях, 101 з яких є об'єктами, а одна категорія містить фонові зображення. Кожна категорія містить від 40 до 800 зображень.

Сцена 67: набір даних містить 67 категорій для приміщень і 15 620

зображень. Кількість зображень залежить від категорії, але є принаймні 100 зображень у кожній категорії.

У всіх експериментах у цій роботі було застосовано 3-кратну перехресну перевірку та витримку. При утриманні дані поділено на 10 частин, 9 частин яких використовуються для навчання, а інші використовуються для тестування. Для оцінки ефективності класифікації зображення та CBIR було виміряно точність, середню точність (mAP) і F1-Score. Крім того, для навчання AlexNet CNN у всіх експериментах було запущено 50 епох. Крім того, для класифікатора випадкових лісів було використано 300 дерев, а для класифікатора KNN значення  $K_{has}$  було встановлено до 1. Оскільки існує велика кількість параметрів, які потрібно встановити, ми знаходимо оптимальні параметри, досліджуючи експерименти на наборі даних Corel-1k.

Швидкість навчання відіграє важливу роль у продуктивності класифікатора нейронної мережі навчання. Вибір малої швидкості навчання призводить до дуже повільної конвергації алгоритму навчання. З іншого боку, вибравши велику швидкість навчання, ми можемо не знайти оптимального рішення.<sup>29</sup> У наших експериментах ми вибрали швидкість навчання з трьома різними значеннями, а саме 0,001, 0,0001 і 0,00001. Рисунки 3.5 і 3.6 демонструють результати класифікації зображень і пошуку зображень з використанням цих коефіцієнтів навчання. На підставі рисунків 3.3 і 3.4 ми бачимо, що найбільш багатообіцяючих результатів досягає швидкість навчання зі значенням 0,0001.

Алгоритм навчання, який використовується для розрахунку оптимальних ваг CNN, має великий вплив на кінцеву продуктивність CNN. З цією метою ми протестували три алгоритми навчання для вимірювання ефективності запропонованого нами методу. Такими алгоритмами є RMSProp,<sup>\*</sup> Adam,<sup>†</sup> і Sgdm.<sup>‡</sup><sup>30</sup> Рисунки 3.7 і 3.8 демонструють результати класифікації зображень і пошуку зображень за допомогою цих алгоритмів навчання. Як показано в рисунках 3.7 і 3.8, найбільш багатообіцяючих результатів досягає алгоритм навчання Sgdm зі значенням моменту, яке дорівнює 0,9.

Іншим важливим ключовим фактором у будь-якій задачі комп'ютерного зору є метод виділення ознак. У наших експериментах ми перевірили наші методи з використанням ознак, витягнутих з різних шарів. Рисунки 3.9 і 3.10 демонструють результати класифікації зображень і пошуку зображень, використовуючи різні характеристики, витягнуті з різних рівнів CNN.

Learning rate	Classifier	Accuracy	
		3-fold (average) cross validation	Hold out
0.001	Classifier	95.60	95
	Random forest	95.70	95
	KNN	95.60	95
	SVM	95.70	95
	Proposed method	95.70	95
0.0001	Classifier	96.60	99
	Random forest	96.50	98
	KNN	96.40	99
	SVM	96.50	99
	Proposed method	96.60	99
0.00001	Classifier	96.10	99
	Random forest	96.30	98
	KNN	96.40	98
	SVM	96.30	98
	Proposed method	96.30	98

Рисунок 3.6 - Результати експериментів з різною швидкістю навчання в класифікації зображень

Learning rate value	Mean average precision	
	3-fold cross validation	Hold out
0.001	96.80	95.40
0.0001	97.40	99.19
0.00001	96.85	97.96

Рисунок 3.7 - Результати експериментів з різною швидкістю навчання при пошуку зображень

Optimizer	Classifier	Accuracy	
		3-fold cross validation	Hold out
RMSProp	Classifier	93.90	94
	Random forest	93.70	94
	KNN	96.40	94
	SVM	93.70	94
	Proposed method	93.70	94
Adam	Classifier	93.80	91
	Random forest	93.80	91
	KNN	93.60	91
	SVM	93.80	91
	Proposed method	93.79	91
Sgdm	Classifier	96.60	99
	Random forest	96.50	98
	KNN	96.40	99
	SVM	96.50	99
	Proposed method	96.60	99

Рисунок 3.8 - Результати експериментів з різними алгоритмами навчання класифікації зображень

Optimizer	Mean average precision	
	3-fold cross validation	Hold out
RMSProp	95.07	95.26
Adam	95.06	92.46
Sgdm	97.40	99.19

Рисунок 3.9 - Результати експериментів з різними алгоритмами навчання

Size of feature vector	Features type	Classifier	Accuracy	
			3-fold cross validation	Hold out
1 × 10	FC (23-th layer in	Classifier	95.80	99
		Random forest	94.50	98
		KNN	96.10	99
		Multi-SVM	95.40	99
		Proposed method	95.80	99
	Softmax (24-th layer	Classifier	96.60	99
		Random forest	96.50	98
		KNN	96.40	98
		Multi-SVM	96.50	98
		Proposed method	96.60	99

Рисунок 3.10 - Результати експериментів з різними ознаками зображення в класифікації зображень

Size of feature vector	Features type	Mean average precision	
		3-fold	Hold out
1 × 10	FC (	91.15	93.25
	Softmax (	97.40	99.19

Рисунок 3.11- Результати експериментів з різними ознаками зображення при пошуку зображень

На основі цих двох таблиць зроблено висновок, що ознаки, отримані з 24-го шару, є більш дискримінантними, ніж ознаки, виділені з 23-го шару. Тому в наших експериментах ми витягуємо ознаки з 24-го шару. В інших експериментах ми працювали з векторами ознак різних розмірів. Тому були досліджені експерименти з векторами ознак довжиною  $1 \times 4096$ ,  $1 \times 200$ ,  $1 \times 50$ ,  $1 \times 10$  та  $1 \times 5$ . Рисунок 3.11 і 3.12 детально описують результати класифікації зображень і пошуку зображень з використанням векторів ознак різних розмірів. У цих таблицях зазначено, що вектори ознак 10-вимірних дають найвищу продуктивність як для завдань класифікації, так і для пошуку. Крім того, у випадку пошуку зображень, щоб виміряти подібність між зображенням запиту та іншими зображеннями, ми досліджували експерименти з використанням евклідових і манхеттенських мір подібності. Рисунок 3.13 показує результати цих експериментів.

Size of feature vector	Classifier	Accuracy	
		3-fold cross validation	Hold out
1 × 4096	Classifier	93.40	95
	Random forest	95.40	96
	KNN	93.70	97
	SVM	94.40	97
	Proposed method	95.40	96
1 × 200	Classifier	84	77
	Random forest	87.50	86
	KNN	74	83
	SVM	49	61
	Proposed method	87.5	86
1 × 50	Classifier	60.90	64
	Random forest	83.60	85
	KNN	69.20	73
	SVM	53.90	64
	Proposed method	83.60	77.47
1 × 10	Classifier	96.60	99
	Random forest	96.50	98
	KNN	96.40	99
	SVM	96.50	99
	Proposed method	96.60	99
1 × 5	Classifier	49.20	46
	Random forest	68.50	66
	KNN	59.70	62
	SVM	42.80	41
	Proposed method	68.51	66

Рисунок 3.12 - Результати експериментів із векторами ознак різних розмірів у класифікації зображень

Size of feature vector	Mean average precision	
	3-fold cross validation	Hold out
1 × 4096	69.17	90.55
1 × 200	59.71	80.56
1 × 50	58.83	67.38
1 × 10	97.40	99.19
1 × 5	48.50	57.09

Рисунок 3.13- Результати експериментів із векторами ознак різних розмірів у пошуку зображень

Size of feature vector	Similarity measure	Features type	Mean average precision	
			3-fold cross validation	Hold out
1 × 10	Manhattan	FC (23-th layer	91.1488	93.2456
		Softmax (24-th layer	97.3952	99.1897
	Euclidean	FC (23-th layer	90.8883	93.8692
		Softmax (24-th layer	97.1199	98.3833

Рисунок 3.14 - Результати запропонованого методу з різною мірою подібності при пошуку зображень

Setting	Parameters	Value
Training	Training algorithm	Sgdm
	Momentum	0.9
	Batch size	60
	Epochs	50
	Initialize learning rate	0.0001
Size of feature vector	Corel-1k	1 × 10
	Caltech-101	1 × 102
	Corel-10k	1 × 100
	Scene-67	1 × 67

Рисунок 3.15 - Оптимальні значення різних параметрів

Як видно з таблиці 11, міра подібності працює трохи краще, ніж евклідова міра подібності. Отже, у запропонованому нами методі ми використовуємо Манхеттенський показник подібності для пошуку зображень. У рисунку 3.14 наведено оптимальні значення всіх параметрів, які ми знайшли в наших експериментах. Слід зазначити, що для того, щоб уникнути переобладнання, у наших експериментах було застосовано техніку збільшення даних.

У всіх експериментах у цьому розділі значення параметрів були встановлені відповідно до рисунку 3.15. Рисунки 3.15-3.19 показують результати класифікації зображень на Corel-1k, Caltech-101, Corel-10k і Scene-

67 відповідно. Рисунок 3.20 була надана для показу середньої середньої точності запропонованого методу для згаданих наборів даних з точки зору пошуку зображень. Розглядаючи рисунок 3.15 детально, можна зрозуміти, що запропонований метод є більш ефективним, ніж інші класифікатори, оскільки він має вищу точність для перехресної перевірки, а також затримати. Крім того, високий показник F1-Score показує, що запропонований метод є успішним з точки зору гармонійного середнього пригадування та точності. Результати цієї таблиці пов'язані з набором даних Corel-1k.

Classifier	Average of 3-fold cross validation		Hold out	
	Accuracy	F1-score	Accuracy	F1-score
Classifier	96.60	93.73	99	95.24
Random forest	96.50	94.17	99	95.25
KNN	96.40	93.34	98	91
SVM	96.50	93.72	99	95.24
Proposed method	96.60	94.17	99	95.24

Рисунок 3.16 - класифікації зображень для Corel-1k

Classifier	Average of 3-fold cross validation		Hold out	
	Accuracy	F1-score	Accuracy	F1-score
Classifier	89.7857	82.35	92.5683	84.62
Random forest	89.1623	82.35	90.7104	86.27
KNN	89.2170	75.06	91.0383	85.15
SVM	88.2170	77.92	89.9454	83.33
Proposed method	89.8075	82.35	92.5683	86.27

Рисунок 3.17 - класифікації зображень для Caltech-101

Classifier	Average of 3-fold cross validation		Hold out	
	Accuracy	F1-score	Accuracy	F1-score
Classifier	88.31	80.53	88.80	81.43
Random forest	88.19	80.39	88.60	84.07
KNN	87.96	75.72	88.70	82.50
SVM	87.63	79.51	87.80	82.50
Proposed method	88.37	80.39	88.90	84.07

Рисунок 3.18- класифікації зображень для Corel-10k

Classifier	Average of 3-fold cross validation		Hold out	
	Accuracy	F1-score	Accuracy	F1-score
Classifier	64.34	59.08	69.03	64.48
Random forest	64.02	59.32	68.65	66.67
KNN	63.21	50.54	68.01	58.29
SVM	61.21	58.55	67.24	63.44
Proposed method	64.37	59.32	69.09	66.67

Рисунок 3.19 - Результати експериментів з використанням оптимальних параметрів у класифікації зображень для Scene-67

Dataset	Mean average precision	
	Average of 3-fold cross validation	Hold out
Corel-1k	97.40	99.19
Caltech-101	87.78	90.22
Corel-10k	90.59	91.12
Scene-67	63.78	68.68

Рисунок 3.20 - Результати експериментів з використанням оптимальних параметрів у пошуку зображень

Dataset	Percentage of memory saving
Corel-1k	$(1 - 10/4096) \times 100 = 99.76\%$
Corel-10k	$(1 - 100/4096) \times 100 = 97.56\%$
Caltech-101	$(1 - 102/4096) \times 100 = 97.51\%$
Scene-67	$(1 - 67/4096) \times 100 = 99.36\%$

Рисунок 3.21- Відсоток збереження пам'яті для кожного набору даних

Детально розглядаючи рисунки 3.16, 3.18 і 3.20, можна побачити, що запропонований метод може класифікувати та відновлювати зображення Caltech-101, Corel-10k і Scene-67 дуже добре. Результати пошуку зображень запропонований метод наведено в таблиці 17. Відповідно до цієї таблиці ясно, що середня середня точність запропонованого методу є значною для всіх чотирьох використаних наборів даних. Згідно з експериментами, деякі важливі проблеми можуть вплинути на ефективність класифікації зображень і пошук зображень, запропонований у цій роботі: 1. Виділені ознаки сильно впливають на кінцеву ефективність запропонованого нами методу. Якщо вони є більш дискримінантними, класифікатор може розпізнати їх більш точно. 2. Хоча CNN



можуть вивчати деталі зображення, їх продуктивність залежить від деталей зображення. Наприклад, якщо категорія зображень містить зображення з простим фоном, ця категорія буде більш ефективно вивчатися запропонованим нами методом. З іншого боку, якщо головний об'єкт категорії важко розпізнати на фоні, запропонований нами метод може зіткнутися з певними труднощами під час класифікації цих категорій. Наявність деяких спільних понять між різними категоріями ускладнює розпізнавання цих категорій одна в одній. Наприклад, деякі поняття, такі як небо, дерево, пісок, річка тощо. поширені в деяких категоріях, таких як гори, пляж, відкрита місцевість і так далі.4. Наші експерименти показали, що класифікація зображень у приміщенні набагато складніша порівняно з класифікацією зображень поза приміщенням, через мінливість між класами та між класами. На рисунках 3.19 і 3.21 показано деякі візуальні результати запропонованого нами методу класифікації та пошуку зображень відповідно. Як показано, запропонований метод правильно класифікував ці зображення. На рисунку 3.21 зображено перші три зображення, які найбільш схожі на зображення запиту. Як було сказано раніше, результати дійсно багатообіцяючі.

У цьому розділі порівнюються результати запропонованих методів із деякими найсучаснішими у сфері класифікації та пошуку зображень.

На рисунках 3.22 і 3.23 видно, що наш запропонований метод досяг кращих результатів у класифікації зображень порівняно з іншими методами для наборів даних Caltech-101 і Scene-67 відповідно. У класифікації зображень внутрішньої сцени (набір даних Scene-67) наш запропонований метод переміг інші методи з відносно високим запасом. На малюнках 3.24 і 3.25 порівнюються результати запропонованого методу з іншими методами для пошуку зображень у наборах даних Corel-1 і Corel-10k відповідно. Як показано, запропонований метод досяг помітної середньої середньої точності (mAP) для пошуку зображення, яка була вищою, ніж AP інших методів.

При використанні підходу SIFT необхідно зберегти набір характеристик. Якщо довжина векторів ознак є оптимальною, необхідна пам'ять буде

мінімальною. Довжина кінцевих векторів ознак дорівнює  $1 \times 4096$ . Тому, використовуючи рівняння (1), можна обчислити відсоток збереження пам'яті для кожного набору даних, враховуючи запропоновані довжини векторів ознак. У цьому рівнянні «Нова довжина» позначає довжини, запропоновані нами в цьому дослідженні, тоді як «Стара довжина» відноситься до довжини вектора ознак CNN.

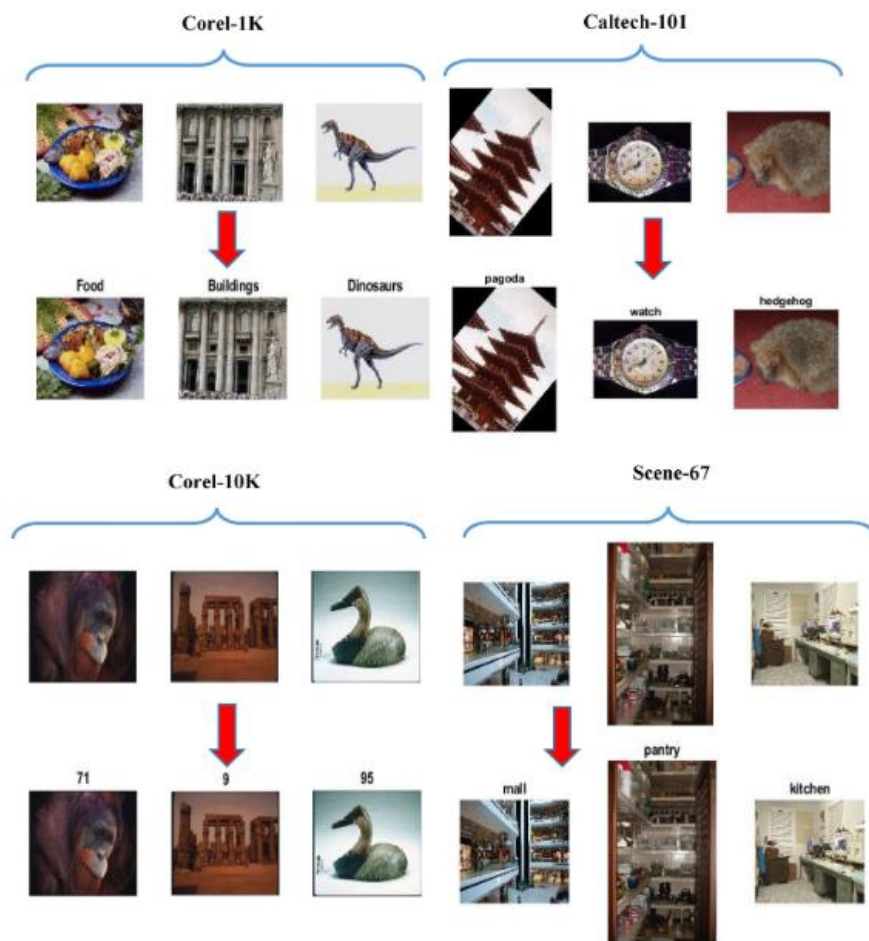


Рисунок 3.22 - Візуальні результати запропонованого нами методу класифікації зображень

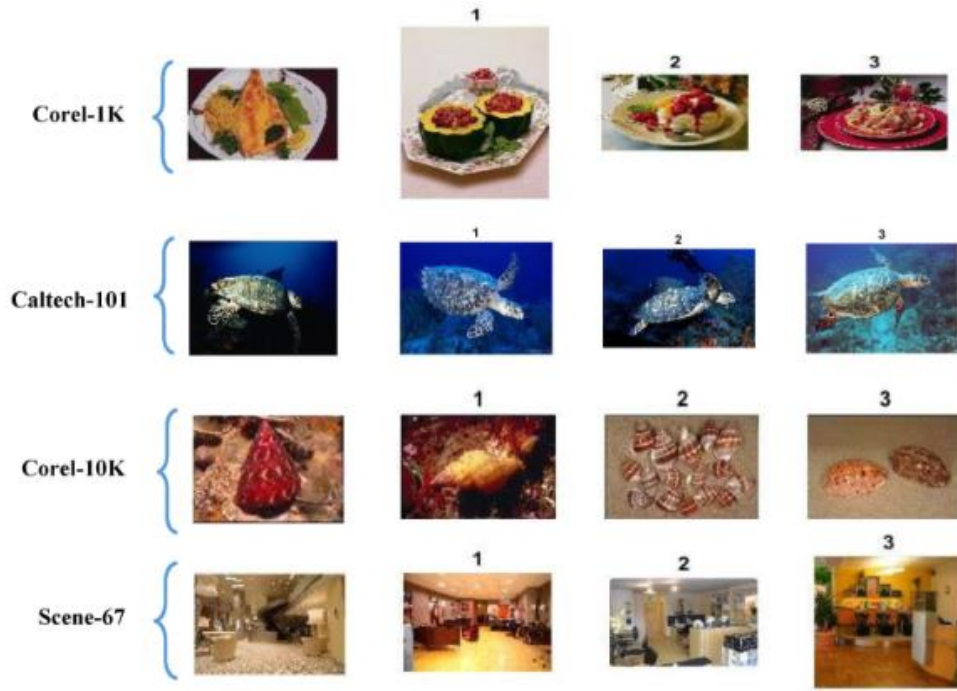


Рисунок 3.23 - Візуальні результати запропонованого нами методу пошуку зображень

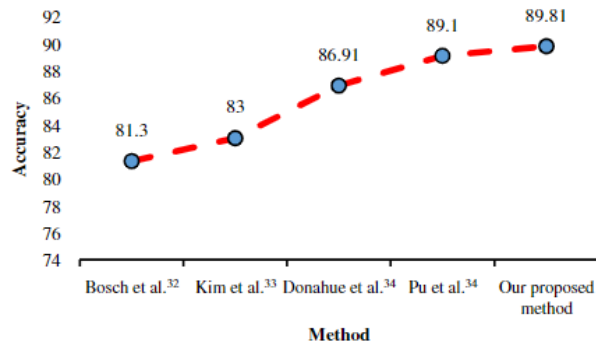


Рисунок 3.24 - Результати порівняння запропонованого нами методу з іншими методами у випадку класифікації зображень для набору даних Caltech-101

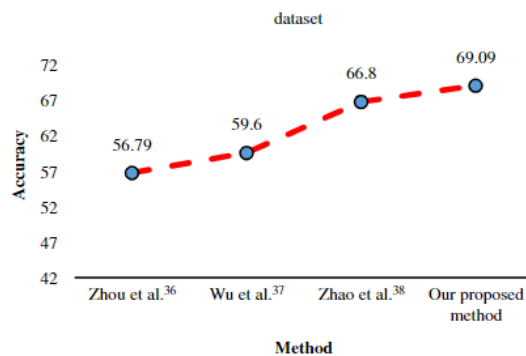


Рисунок 3.25 - Результати порівняння запропонованого нами методу з іншими методами у випадку класифікації зображень для бази даних Scene-67

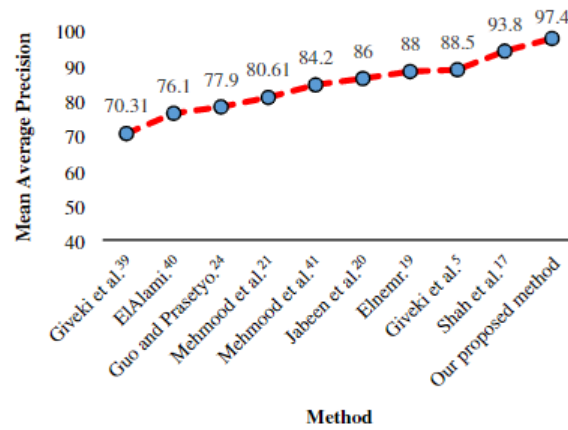


Рисунок 3.26 - Результати порівняння запропонованого нами методу з іншими методами у випадку пошуку зображень для набору даних Corel-1k

$$\text{MemSaving}(\text{NewLength}, \text{OldLength}) = \left(1 - \frac{\text{NewLength}}{\text{OldLength}}\right) \times 100$$

Цікаві спостереження можна зробити з рисунків 3.26 і 3.27. Ці малюнки порівнюють розмірність запропонованого вектора ознак у цьому дослідженні з вектором ознак інших методів для наборів даних Corel-1k і Corel-10k відповідно. Якщо довжина запропонованого нами вектора ознак для Corel-1k дорівнює 1, запропонована довжина дослідження, 40 яка в 2,1 рази більша за наш вектор ознак, дорівнює, наприклад, 2,1 коефіцієнту. Подібним чином, інші методи отримали співвідношення, враховуючи їх довжину ознак, а не запропонований вектор ознак. Як видно з рисунків 3.28 і 3.29, наш запропонований метод потребує меншого простору для роботи над векторами ознак.

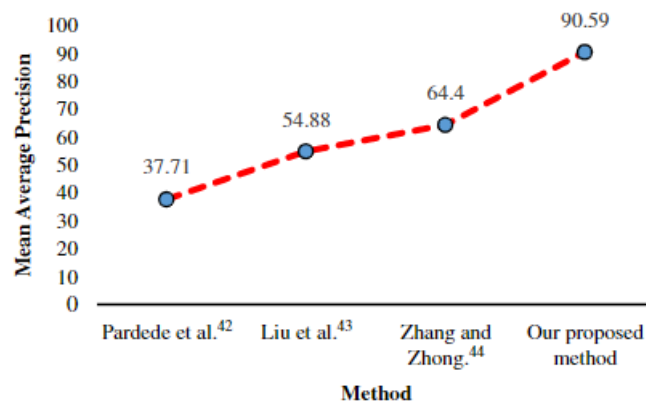


Рисунок 3.27 - Результати порівняння запропонованого нами методу з іншими методами у випадку пошуку зображень для набору даних Corel-10k

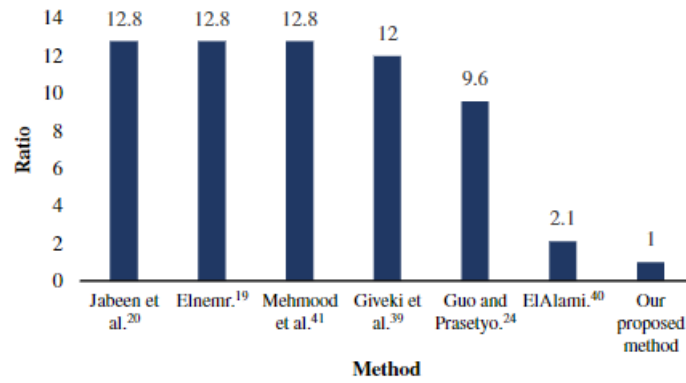


Рисунок 3.28 - Порівняння розміру запропонованого нами вектора ознак з іншими методами для набору даних Corel-1k

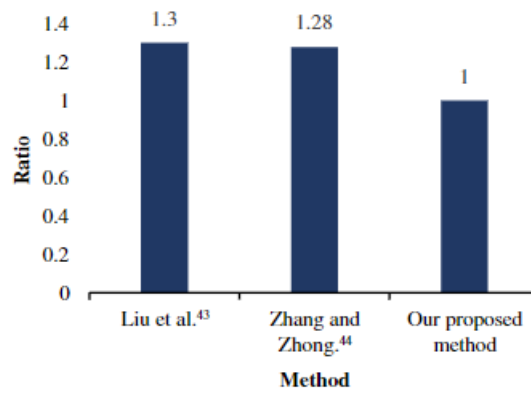


Рисунок 3.29 - Порівняння розміру запропонованого нами вектора ознак з іншими методами для набору даних Corel-10k

## Висновки до розділу 3

1. У цій роботі досліджується новий метод класифікації зображень і SIFT з використанням згорткових нейронних мереж (CNN). Завдяки передачі знань витягнуті функції були застосовані для класифікації та пошуку зображень. Пропонується нова стратегія голосування комітету для прогнозування міток класів. Запропонований метод розумно включає голоси SVM, K-NN, випадкового лісу та класифікатора CNN.

2. Ще однією перевагою запропонованого методу є мала довжина векторів ознак. Як було показано раніше, малі вектори ознак економлять багато пам'яті. Експериментальні результати на різних контрольних наборах даних показали, що запропоновані методи перевершують найсучасніші методи класифікації та пошуку зображень.

3. Результати порівняння показують, що запропонований метод особливо добре працює для класифікації зображень сцени в середовищі доповненої реальності.

## ВИСНОВКИ

1. У кваліфікаційній роботі проведено аналіз методів класифікації елементів зображень у середовищі доповненої реальності, який засвідчив інтенсивність зростання ринку та різноманіття послуг, котрі надаються розробниками доповненої реальності. Сьогодні серед методів класифікацій зображень різні методи пропонують подібні підходи з різними функціями та рівнями продуктивності. Велика кількість різних процесорних обробок ускладнює програмі своєчасне відображення необхідних для користувацького інтерфейсу зображень. Прийняття правильного рішення щодо вибору методу класифікації зображень має вирішальне значення для успіху роботи продукту, і це може заощадити користувачам багато часу і вирішити проблеми у довгостроковій перспективі.

2. При виборі найвідповіднішого методу класифікації елементів зображень деякі з критерії відбору є не простими і двозначними. Деякі методи повторюють функціонал попередніх, в одночас переважаючи в одних показниках і відстаючи у інших, наприклад у продуктивності обробки для малої кількості шарів. Також приходиться шукати компроміс між багатьма з цих методів, такими як продуктивність і кількість інформації для передачі. За таких умов користувачам доводиться або очікувати на відображення, або отримувати недостатньо інформації від інтерфейсу, що не дозволяє розглядати широкий спектр різних послуг постачальника, які могли б стати у пригоді користувачу, а значить принести б прибуток розробнику.

3. Сьогодні розробники додатків доповненої реальності постійно конкурують за створення продукту, яких би водночас був би і достатньо високо функціональним і мав би відповідний інтерфейс. Ця конкуренція спочатку базувалася насамперед на вартості наданих ресурсів, але кількісна оцінка та порівняння фактичних можливостей тепер стають все критичнішими. Для створення покращеного методу класифікації зображень в основу було закладено метод CNN, а для його роботи у доповненій реальності вибрано відповідне середовище Vuforia, також були використанні і інші методи

машинного навчання для класифікації елементів зображення, що зробить метод унікальним і поєднає у собі найкращі частини.

4. У роботі проведено дослідження методів класифікації елементів зображень та проаналізовано алгоритми обробки цих зображень у середовищі доповненої реальності. Показано, що для вирішення проблеми класифікації елементів зображень найкраще підходить метод CNN. Зокрема, обґрунтовано доцільність використання алгоритму SIFT.

5. Модифікація методу CNN полягає у покращенні його алгоритму у плані обробки більшої кількості даних і їх швидкої обробки. Це дає змогу покращити продуктивність методу, а відповідно і відображення зображення. Розроблено детальну схему нового методу.

6. Проведено апробацію запропонованого методу до його реалізації у середовищі доповненої реальності. Для цього показано алгоритм, він включає п'ять кроків, кожен з яких детально описано в наступних підрозділах. Результати апробації підтвердили ефективність розробленого методу.

7. Автором запропоновано модернізовану структуру методу класифікації елементів зображень у середовищі доповненої реальності. Запропонований метод підвищує продуктивність процесорної обробки інформації у великих кількостях, завдяки використанню згорткових нейронних мереж у середовищі доповненої реальності.



## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Види сучасної реклами [Електрон. ресурс]. – Режим доступу: [https://www.marketch.ru/marketing\\_dictionary/marketing\\_terms\\_v/types\\_of\\_advertising/](https://www.marketch.ru/marketing_dictionary/marketing_terms_v/types_of_advertising/)
2. Azuma, R. T. (1997). A survey of augmented reality. *Presence-Teleoperators and Virtual Environments*, 6(4), 355-385.
3. Azuma R., *A Survey of Augmented Reality*. Cambridge University Pres, 2005. 254 p.
4. Rosten E. *Faster and better: a machine learning approach to corner detection*, 2008. 96 p
5. Mikolajczyk K., Schmid C. *An Affine Invariant Interest Point Point Detector // European Conference on Computer Vision*, 2002. P. 128–142.
6. Ronald T. Azuma A. *Survey of Augmented Reality // In Presence: Teleoperators and Virtual Environments*. 1997. No 4. P. 355–385.
7. Tikander M. *Development and evaluation of augmented reality audio systems: Abstract of dissertation for the degree of Doctor of Science in Technology*. Helsinki, 2009. 231 p.
8. Rozier J. *Hear&There: An Augmented Reality System of Linked Audio // Online Proceedings of the ICAD*, 2009. P 92–123.
9. Karahalios K. *GaitAid Virtual Walker for Movement disorder patients Augmented Reality*, 2007. 231 p
10. Benda P., Ulman M. *Augmented Reality As a Working Aid for Intellectually Disabled Persons For Work in Horticulture*. [Електрон.ресурс]. – Режим доступу: [https://www.researchgate.net/publication/298033889\\_Augmented\\_Reality\\_As\\_a\\_Working\\_Aid\\_for\\_Intellectually\\_Disabled\\_Persons\\_For\\_Work\\_in\\_Horticulture](https://www.researchgate.net/publication/298033889_Augmented_Reality_As_a_Working_Aid_for_Intellectually_Disabled_Persons_For_Work_in_Horticulture)
11. Linietsky J. *A Small Defense Of GLTF 2.0 On Its Comparison Against Opengex*. [Електрон.ресурс]. – Режим доступу: <https://godotengine.org/article/small-defense-gltf>

12. Brainville A. Why glTF 2.0 is awesome! [Електрон.ресурс]. – Режим доступу: <https://dev.to/ybalrid/why-glTF-20-is-awesome-2khp>
13. Understanding World Tracking in ARKit [Електрон.ресурс]. – Режим доступу:  
[https://developer.apple.com/documentation/arkit/understanding\\_world\\_tracking\\_in\\_arkit](https://developer.apple.com/documentation/arkit/understanding_world_tracking_in_arkit)
14. Axon S. How ARKit 2 works, and why Apple is so focused on AR. [Електрон.ресурс]. – Режим доступу:  
<https://arstechnica.com/gadgets/2018/06/arkit-2-why-apple-keeps-pushing-ar-and-how-it-works-in-ios-12/>
15. Ronald T. Azuma. A Survey of Augmented Reality. [Електрон.ресурс]. – Режим доступу:  
<https://www.cs.unc.edu/~azuma/ARpresence.pdf>
16. Чалая, Л.Э. Поиск неполных дубликатов в системах анализа цифровых изображений [Текст] / Л.Э. Чалая, П.Ю. Попаденко // Вісник КрНУ імені Михайла Остроградського. Вип. 5/2014 (88). – 2014. – С.42 – 47.
17. Augmented Reality with Surface Recognition [Електрон.ресурс]. – Режим доступу: <http://www.appmechanic.in/augmented-reality-with-surface-recognition>.
18. Різниця гаусіанів [Електронний ресурс] – Режим доступу:[https://uk.wikipedia.org/wiki/Різниця\\_гаусіанів](https://uk.wikipedia.org/wiki/Різниця_гаусіанів)
19. Перетворення Хафа [Електронний ресурс] – Режим доступу:  
[https://ua.wikipedia.org/wiki/Перетворення\\_Хафа](https://ua.wikipedia.org/wiki/Перетворення_Хафа)
20. Alkhamisi, A. O. & Monowar, M. M. (2013). Rise of Augmented Reality: Current and Future Application Areas, International Journal of Internet and Distributed Systems
21. Arth, C. Grasset, R. Gruber, L. Langlotz, T. Alessandro Mulloni, A. & Wagner, D. (2015). The History of Mobile Augmented Reality. Inst. For Computer Graphics and Vision Graz University of Technology, Austria.
22. Azuma, R. (1997). A Survey of Augmented Reality. Presence: Hughes Research Laboratories

23. Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E., & Ivkovic, M. (2011). *Augmented Reality Technologies, Systems and Applications. Multimedia Tools and Applications*
24. Cranmer, E. E. (2017). *Developing an Augmented Reality Business Model for Cultural Heritage Tourism: The Case of Geevor Museum, (Dissertation), (Published Doctoral Thesis), Department of Operations, Technology, Events and Hospitality Management, Manchester Metropolitan University, Manchester.*
25. Dieck, T. M.C. & Jung, T. (2015). *A Theoretical Model of Mobile Augmented Reality Acceptance in Urban Heritage Tourism, Current Issues in Tourism.*
26. Edwards-Stewart, A. & Hoyt, T. & Reger, G. (2016). *Classifying Different Types of Augmented Reality Technology. Annual Review of Cyber Therapy and Telemedicine.*
27. Ferreira, S. Alves, A. P. Quico, C. (2014) *Location Based Transmedia Storytelling in Social Media Peters' TravelPlot Porto Case Study, E Review of Tourism Research (eRTR) ENTER Conference on Information and Communication Technologies in Tourism.*
28. Han, D. I., Jung, T., & Gibson, A. (2013). *Dublin AR: Implementing Augmented Reality (AR) in Tourism, In Z. Xiang, & I. Tussyadiah (Eds), Information and Communication Technologies in Tourism (511-523). Springer Computer Science: New York.*
29. Ding, M. (2017). *Augmented Reality in Museums. Arts Managment & Technology Laboratory, 3-8, Carnegie Mellon Universty, Pennsylvania.*
30. Jacob, J., Da Silva, H., Coelho, A. & Rodrigues, R. (2012). *Towards Locationbased Augmented Reality Games, Procedia Computer Science.*
31. Jung, T. & Han, D. I. (2014). *Augmented Reality (AR) in Urban Heritage Tourism, e-Review of Tourism Research.*
32. Pagani, A., Henriques, J. & Stricker, D. (2016). *Sensors for Location-Based Augmented Reality the Example of GALILEO and EGNOS. The*

International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.

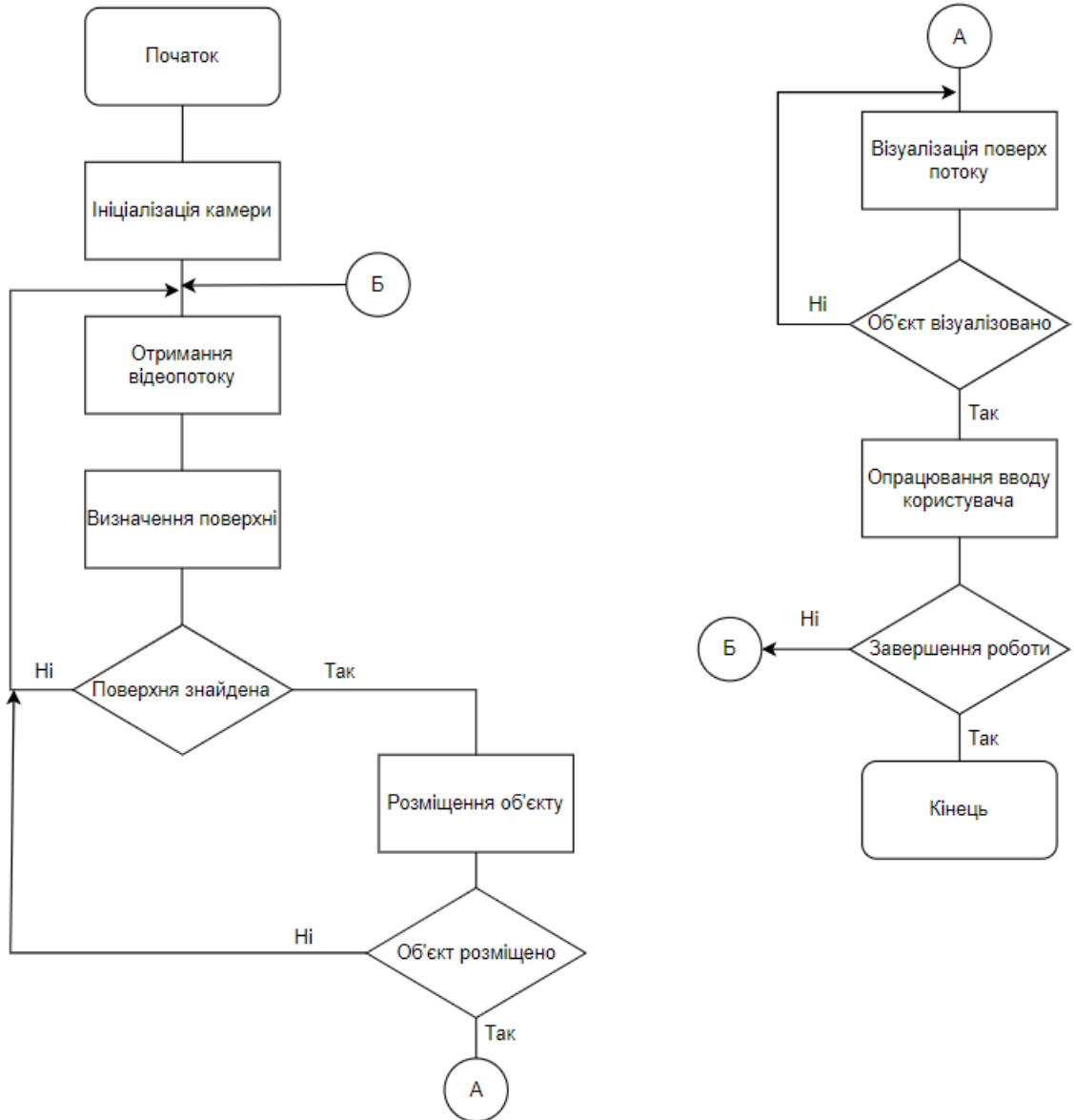
33. Ramos, F., Trilles, S., Torres-Sospedra, J. & Perales, F. J. (2018), New Trends in Using Augmented Reality Apps for Smart City Contexts, *International Journal of Geo-Information*.
34. Reitmayr, G. & Schmalstieg, D. (2003) Location Based Applications for Mobile Augmented Reality. Proc. 4th Australasian User Interface Conference, Adelaide, Australia.
35. Shabani, N., Munir, A. & Hassan, A. (2018). E-Marketing via Augmented Reality: A Case Study in the Tourism and Hospitality Industry.
36. Siltanen, S. (2012). Theory and Applications of Marker-Based Augmented Reality. Science Technology Research Highlights Visions. VTT Technical Research Centre, Finland.
37. Silva, R. Giraldi, G. & Jauvane, C. O. (2003). Introduction to Augmented Reality. Technical Report: 25/2003, National Laboratory for Scientific Computer.
38. Van Krevelen, D.W.F. (2007). Augmented Reality: Technologies, Applications, and Limitations. Department of Computer Science. Vrije University, Amsterdam.
39. Van Krevelen, D. W. F. & Poelman, R. (2010). A Survey of Augmented Reality Technologies.
40. Bosch, A., Zisserman, A., & Munoz, X. Image classification using random forests and ferns. Paper presented at: Proceedings of the 2007 IEEE 11th International Conference on Computer Vision (ICCV 2007); October 2007:1–8; IEEE.33.Kim, S., Min, D., Ham, B., Jeon, S., Lin, S., & Sohn, K. FCSS: Fully convolutional self-similarity for dense semantic correspondence. Paper presented at: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); February 2017;1(2):581–595.
41. Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., & Darrell, T. DeCAF: a deep convolutional activation feature for generic visual recognition. Paper presented

- at: Proceedings of the International Conference on Machine Learning; January 2014:647–655.
42. Pu, Y., Yuan, W., Stevens, A., Li, C., & Carin, L. A deep generative deconvolutional image model. Paper presented at: Proceedings of the 19th International Conference on Artificial Intelligence and Statistics; May 2016:741–750.
  43. Zhou, B., Lapedriza, A., Xiao, J., Torralba, A., & Oliva, A. Learning deep features for scene recognition using places database. Paper presented at: Proceedings of the Advances in Neural Information Processing Systems; 2014:487–495.
  44. Wu, J., Yu, Y., Huang, C., & Yu, K. Deep multiple instance learning for image classification and auto-annotation. Paper presented at: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2015:3460–3469.
  45. Zhao Y, You X, Yu S, et al. Multi-view manifold learning with locality alignment. *Pattern Recognit.* 2018;78:154-166.
  46. Khalifa, F. A., Semaary, N. A., El-Sayed, H. M., & Hadhoud, M. M. (2016). Markerless Tracking for Augmented Reality Using Different Classifiers. *The 1st International Conference on Advanced Intelligent System and Informatics*, 28-30
  47. Cui, Z., Wu, Z., Zhou, C., Gao, G., Yu, J., Zhao, Z., & Wu, B. (2016). An Efficient Subscription Index for Publication Matching in the Cloud. *Knowledge-Based Systems*, 110, 110-120.
  48. Cui, Z., Zhu, H., Shi, J., Chi, L., & Yan, K. (2016). Efficient Authorisation Update on Cloud Data. *International Journal of Web and Grid Services*, 12(2), 109-141.
  49. Çalışkan, A. & Çevik, U. (2018). An Efficient Noisy Pixels Detection Model for CT Images using Extreme Learning Machines. *Technical Gazette*, 25(3), 679-686.
  50. Morrison, J. G., Gálvez-López, D., & Sibley, G. (2016). Moarslam: Multiple Operator Augmented RSLAM. *Distributed Autonomous Robotic Systems*, 119-132.

51. Huang, J., Zhou, G., Zhou, X., & Zhang, R. (2018). A New FPGA Architecture of FAST and BRIEF Algorithm for OnBoard Corner Detection and Matching. *Sensors*, 18(4)
52. Erbek, F. S., Ozkan, C. and Taberner, M. 2004. Comparison of maximum likelihood classification method with supervised artificial neural network algorithms for land use activities.. *International Journal of Remote Sensing*, 25.
53. Du, Q. and Chang, C. 2001. A linear constrained distance-based discriminant analysis for hyperspectral image classification.. *Pattern Recognition*.
54. Crosetto, M., Ruiz, J. A. M. and Crippa, B. 2001. Uncertainty propagation in models driven by remotely sensed data.. *Remote Sensing of Environment*.
55. Foody, G. M. 1992. On the compensation for chance agreement in image classification accuracy assessment.. *Photogrammetric Engineering and Remote Sensing*.
56. Foody, G. M. 2002b. Status of land cover classification accuracy assessment.. *Remote Sensing of Environment*.
57. Franklin, J., Phinn, S. R., Woodcock, C. E. and Rogan, J. 2003. "Rationale and conceptual framework for classification approaches to assess forest resources and properties.". In *Remote Sensing of Forest Environments: Concepts and case studies*, Edited by: Wulder, M. A and Franklin, S. E.
58. Jakubaukas, M. E. 1997. Effects of forest succession on texture in Landsat Thematic Mapper imagery.. *Canadian Journal of Remote Sensing*.
59. Jeon, B. and Landgrebe, D. A. 1999. Decision fusion approaches for multitemporal classification.. *IEEE Transactions on Geoscience and Remote Sensing*.
60. Su J, Vargas DV, Sakurai K (2018) One pixel attack for fooling deep neural networks. *arXiv*.
61. Chen H, Zhang Y, Zhang W et al (2017) Low-dose CT via convolutional neural network.

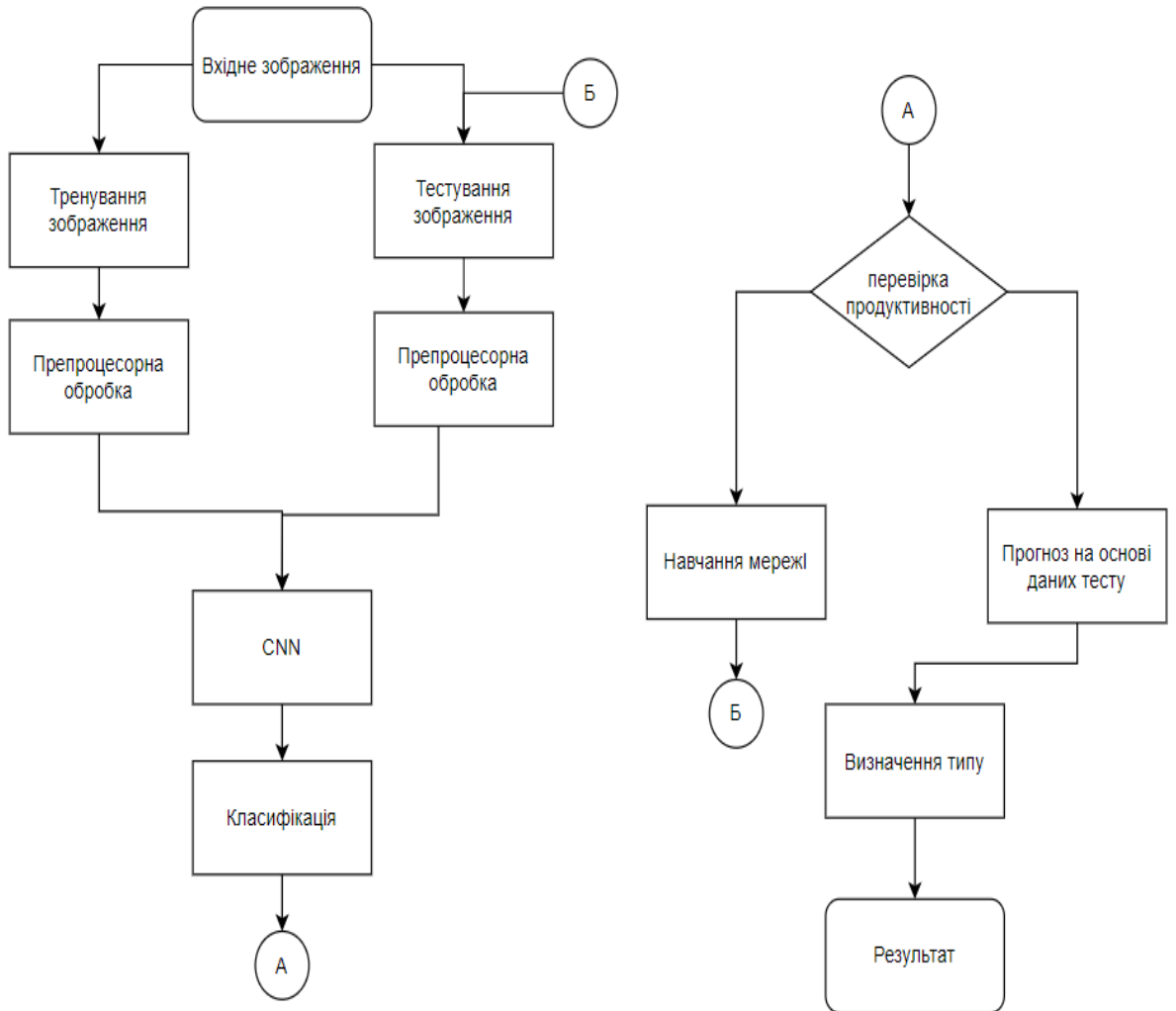
62. Lucchesi FR, Aredes ND (2016) Radiology data from The Cancer Genome Atlas Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma (TCGA-CESC) collection.
63. Setio AA, Ciompi F, Litjens G et al (2016) Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. *IEEE Trans Med Imaging*.

ДОДАТОК А  
СХЕМА АЛГОРИТМУ РОБОТИ ДОДАТКУ НА БАЗІ МЕТОДУ  
КЛАСИФІКАЦІЇ ЕЛЕМЕНТІВ ЗОБРАЖЕНЬ





ДОДАТОК Б  
БЛОК-СХЕМА РОБОТИ АЛГОРИТМУ SIFT З ВИКОРИСТАННЯМ  
CNN КЛАСИФІКАТОРА



## ДОДАТОК В

### АПРОБАЦІЯ ОТРИМАНИХ РЕЗУЛЬТАТІВ

#### Критерії вибору методів класифікації елементів зображень

**Саченко Анатолій Олександрович,**

д.т.н., професор

**Ковальський Семен Сергійович,**

Студент

Західноукраїнський національний університет

м. Тернопіль, Україна

esceyr249@gmail.com

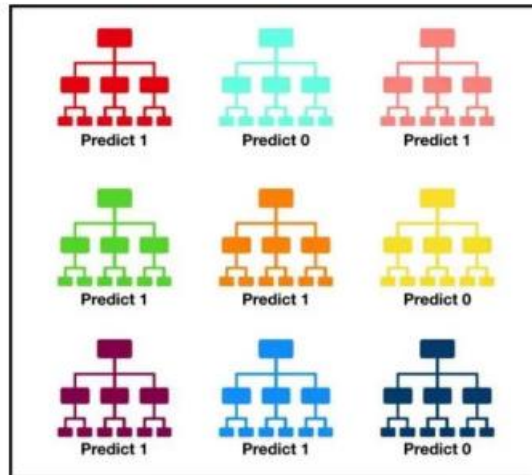
**Анотація:** Однією з основних проблем комп'ютерного зору є проблема класифікації зображень, яка пов'язана з визначенням наявності візуальних структур у вхідному зображенні. Класифікація зображень аналізує числові властивості різних характеристик зображення та впорядковує дані за категоріями. Останніми роками для класифікації зображень широко застосовуються багато передових підходів до класифікації, такі як штучні нейронні мережі, нечіткі множини та експертні системи, але кожен із них має певні проблеми, а рівень їх точності порівняно менший.

**Ключові слова:** Глибоке навчання, Згорткові нейронні мережі, SVM, RF, класифікація елементів зображень.

Спеціалісти з обробки та аналізу даних мають безліч засобів для створення класифікаційних моделей. Один із найпопулярніших і надійніших методів розробки таких моделей полягає у використанні алгоритму «випадковий ліс» (Random Forest, RF). Для того, щоб спробувати покращити показники моделі, побудованої з використанням алгоритму RF, можна скористатися оптимізацією гіперпараметрів моделі.

Він використовує пакування та випадковість функцій під час побудови кожного окремого дерева, щоб спробувати створити некорельований ліс дерев, передбачення якого комітетом точніше, ніж прогнозування будь-якого окремого дерева[1, 278–282].

Випадковий ліс, як випливає з назви, складається з великої кількості окремих дерев рішень, які працюють як ансамбль. Кожне окреме дерево у випадковому лісі видає прогнозований клас, і клас, який набрав найбільше голосів, стає прогнозом моделі (рис. 1).



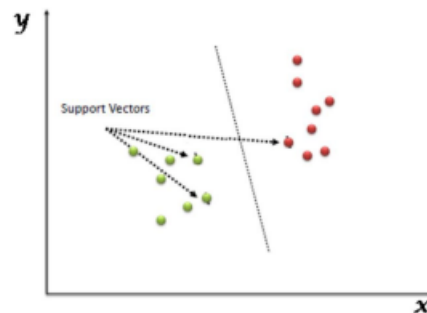
**Рис.1. Візуалізація моделі випадкового лісу, що робить прогноз**

Оскільки ми знаємо, що ліс складається з дерев, і більше дерев означає більш надійний ліс, подібним чином алгоритм випадкового лісу створює дерева рішень на вибірках даних, а потім отримує прогноз від кожного з них і, нарешті, вибирає найкраще рішення шляхом голосування. Це метод ансамблю, який є кращим, ніж одне дерево рішень, оскільки він зменшує надмірне підгонювання шляхом усереднення результату.

Метод опорних векторів (SVM) — це потужні, але гнучкі керовані алгоритми машинного навчання, які використовуються як для класифікації, так і для регресії[2, 125–137]. Машини опорних векторів мають унікальний спосіб реалізації порівняно з іншими алгоритмами машинного навчання. Вони надзвичайно популярні через свою здатність обробляти кілька безперервних і категоріальних змінних. Модель опорної векторної машини в основному є

представленням різних класів на гіперплощині в багатовимірному просторі. Гіперплощина буде згенерована ітераційним способом машиною опорних векторів, щоб можна було мінімізувати помилку. Мета полягає в тому, щоб розділити набори даних на класи, щоб знайти максимальну граничну гіперплощину. Він будує гіперплощину або набір гіперплощин у великому просторі, і гарне розділення між двома класами досягається за допомогою гіперплощини, яка має найбільшу відстань до найближчої точки навчальних даних будь-якого класу. Реальна потужність цього алгоритму залежить від використовуваної функції ядра. Найпоширенішими ядрами є лінійне ядро, гауссове ядро та поліноміальне ядро.

У випадку використання для цілей класифікації, він розділяє класи за допомогою лінійної межі (рис 2.2)



**Рис. 2. SVM. Розділ класів за допомогою лінійної межі**

Розмірність гіперплощини залежить від кількості ознак. Якщо кількість вхідних елементів дорівнює двом, то гіперплощина є просто лінією. Якщо кількість вхідних елементів дорівнює трьом, то гіперплощина стає двовимірною площиною. Стає важко уявити, коли кількість ознак перевищує три.

Згорточна нейронна мережа (CNN або ConvNet) — це особливий вид багатосарових нейронних мереж, призначених для розпізнавання візуальних шаблонів безпосередньо з піксельних зображень з мінімальною попередньою

обробкою. Також у глибокому навчанні (deep learning) це клас штучної нейронної мережі (ANN)[3, 232–243]. Згорткова нейронна мережа використовує деякі свої особливості зорової кори і тому досягла найсучасніших результатів у задачах комп'ютерного зору. Згорткові нейронні мережі складаються з двох дуже простих елементів, а саме згорткових шарів і шарів об'єднання. Незважаючи на простоту, існує майже нескінченна кількість способів упорядкувати ці шари для певної проблеми комп'ютерного зору. Елементи згорткової нейронної мережі, такі як згортковий шар і шари об'єднання, відносно прості для розуміння. Складна частина використання згорткових нейронних мереж на практиці полягає в тому, як спроектувати модельні архітектури, які найкраще використовують ці прості елементи. Причина, чому згорткова нейронна мережа надзвичайно популярна, полягає в її архітектурі, найкраще, що немає потреби у вилученні функцій. Система вчиться виділяти ознаки, а основна концепція полягає в тому, що вона використовує згортку зображення та фільтри для створення незмінних функцій, які передаються на наступний рівень.

Робота відбувається в так званому шарі згортки. Для цього визначається фільтр, який визначає, наскільки великими повинні бути часткові зображення, які переглядаються, і довжину кроку, яка визначає, скільки пікселів продовжують використовуватися між обчисленнями, тобто наскільки близько розташовані часткові зображення одне до одного. Зробивши цей крок, значно зменшиться розмірність зображення.

Наступний крок — шар об'єднання. З чисто обчислювальної точки зору, тут відбувається те саме, що й у шарі згортки, з тією різницею, що береться лише середнє або максимальне значення результату, залежно від програми. Це зберігає дрібні особливості в кількох пікселях, які є вирішальними для вирішення завдання.

Нарешті, є повністю зв'язаний рівень, який використовується у звичайних нейронних мереж. Тепер, коли ми значно зменшили розміри зображення, ми

можемо використовувати шари з щільною сіткою. Тут окремі підзображення знову пов'язані, щоб розпізнати зв'язки та виконати класифікацію.

Тепер, коли ми маємо базове розуміння того, що приблизно роблять окремі шари, ми можемо детально розглянути, як зображення стає класифікацією. Для цього ми намагаємося по зображенню  $4 \times 4 \times 3$  розпізнати, чи є на ньому собака.

На першому кроці ми хочемо зменшити розміри зображення  $4 \times 4 \times 3$ . Для цього ми визначаємо фільтр розміром  $2 \times 2$  для кожного кольору. Крім того, ми хочемо, щоб довжина кроку дорівнювала 1, тобто після кожного кроку обчислення фільтр потрібно переміщати вперед рівно на один піксель. Це не зменшить розмір, але деталі зображення зберуться. Якщо ми переносимо матрицю  $4 \times 4$  на  $2 \times 2$  і просуваємо на один стовпець або рядок на кожному кроці, наш згортковий рівень матиме матрицю  $3 \times 3$  як вихід. Окремі значення матриці обчислюються шляхом скалярного добутку матриць  $2 \times 2$  (рис. 3).

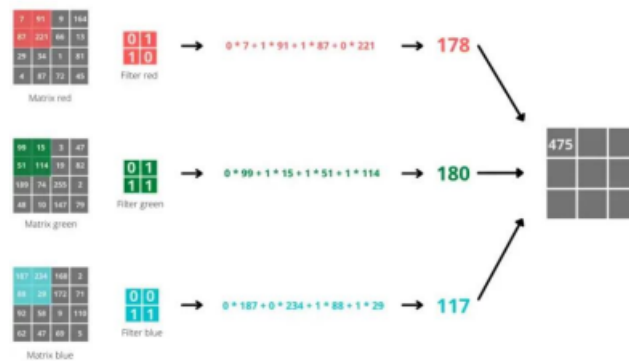


Рис.3. Шар згортки

Об'єкти на наступному рівні згорнуті різними фільтрами, щоб створити більш інваріантні та абстрактні об'єкти, і процес триває, доки не буде отримано остаточний об'єкт/вихід, незмінний до оклюзій. Найпоширенішими архітектурами згорткової нейронної мережі є LeNet, AlexNet, ZFNet, GoogLeNet, VGGNet і ResNet.



Отже проаналізувавши найпопулярніші методи класифікацій елементів зображень можна зробити висновок, що найкращим серед наведених методів є CNN, тому що:

- Згорткові нейронні мережі використовуються в обробці зображень і мови та базуються на структурі зорової кори людини.
- Вони складаються з шару згортки, шару об'єднання та повністю зв'язаного шару.
- Згорткові нейронні мережі ділять зображення на менші області, щоб уперше побачити їх окремо.
- Згорткові нейронні мережі можна запрограмувати всього за кілька кроків за допомогою Tensorflow.
- Можливо налаштувати розташування шарів згортки та максимального об'єднання для кожного окремого випадку використання.

Також його перевагу можна отримати зрівнявши ідентичні дані різними методами(табл. 1).

Таблиця 1

Оцінка ефективності методів класифікації елементів зображень.

LASSIFIER	ACCURACY	PRECISION	RECALL	ROC
SVM	85.68%	0.86	0.87	0.86
Decision Trees	84.61%	0.85	0.84	0.82
KNN	86.32%	0.86	0.86	0.88
ANN(for 100 epochs)	83.10%	0.88	0.87	0.88
CNN(for 300 epochs)	91.11%	0.93	0.89	0.97

З таблиці продуктивності можна зробити висновок, що згорткові нейронні мережі забезпечують найкращі результати в задачах комп'ютерного зору.

#### Список літератури

1. Tin Kam (1995). Random Decision Forests. Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. 278–282.
2. Ben-Hur, Asa; Horn, David; Siegelmann, Hava; Vapnik, Vladimir N. "Support vector clustering" (2001);". Journal of Machine Learning Research. pp 125–137.
3. Valueva, M.V.; Nagornov, N.N.; Lyakhov, P.A.; Valuev, G.V.; Chervyakov, N.I. (2020). "Application of the residue number system to reduce hardware costs of the convolutional neural network implementation". Mathematics and Computers in Simulation. Elsevier BV. 177: pp 232–243.

## ОБРОБКА ЗОБРАЖЕНЬ З ДОПОВНЕНОЮ РЕАЛЬНІСТЮ (AR)

**Саченко Анатолій Олексійович,**

д.т.н., професор

**Ковальський Семен Сергійович,**

Студент

Західноукраїнський національний університет

м. Тернопіль, Україна

[esceyp249@gmail.com](mailto:esceyp249@gmail.com)

**Анотація.** У цьому проекті мета — обговорити та сформулювати намір створити програму Android на основі зображень. Основою цього дослідження є виявлення та обробка зображень у реальному часі. Це новий зручний засіб, який дозволяє користувачам отримувати інформацію про зображення прямо на місці. У минулих дослідженнях були виявлені спроби створення програм на основі зображень, але вони дійшли лише до створення засобів пошуку зображень, які працюють лише з зображеннями, які вже зберігаються в певній формі бази даних. Платформа Android швидко поширюється світом і на сьогоднішній день є найбільш інтерактивною та технічною платформою для смартфонів. Ось чому було важливо базувати дослідження та дослідження на цьому. Доповнена реальність у цьому проекті дозволяє користувачеві маніпулювати даними та може додавати розширені функції (відео, теги GPS) до зробленого зображення.

**Ключові слова:** мобільний додаток, доповнена реальність, мобільна (Android) AR, обробка зображень.

**Вступ.** Стаття включає в себе повний аналіз і покроковий макет запропонованої програми. Ця програма працюватиме лише зі смартфонами Android із підтримкою камери. Користувач фокусує смартфон, потім робить знімок вивіски, рекламного щита чи плаката, і зроблене зображення



виявляється та обробляється, а за допомогою цієї програми виводяться або відображаються результати щодо зображення. Основна мета полягає в тому, щоб відображені результати відображали відповідний зміст на цьому конкретному знаку. Після перегляду статей дійшло загального розуміння, що багато експертів у минулому, які намагалися створити подібний проект, головним чином були стурбовані визначенням країв, кодуванням кольорів і перетином. Запропонована програма допоможе користувачам офлайн отримувати інформацію про вивіски та плакати без необхідності переглядати веб-сторінки чи обов'язково перенаправляти на інші сторінки.

Попереднє дослідження Android від Google — це нова платформа для смартфонів, яка має на меті спростити завдання для користувача шляхом мобілізації завдань за допомогою незліченних програм, що робить його ще зручнішим для нього. Це дослідження охоплює рецензовані статті з технологічних конвенцій [1]. Японія виявила, що їхній внесок у виявлення зображень полягав у позначенні певних точок на зображенні та отриманні даних про нього шляхом акцентування уваги на гістограмі моделі та розмірах пікселів. Інші дослідження [2] включали відеознімки, які були інтегровані з аудіознімками для створення кадрів або зображень у формі кадрів. Виявлення за допомогою алгоритму [3] також проілюстровано на рисунку 1.



Рисунок 1 - Інтенсивність змінюється за допомогою зображень у реальному часі з метро.

Було кілька спроб щодо обробки зображень. Одне з них включає дослідження Чуань Цінь, який відвідував дослідницьку групу систем і мереж в Університеті Дьюка. В основі його дослідження лежить головним чином «TagSense» [4], який, на його думку, легше уявити, ніж обробку зображень. Його аргумент полягав у тому, що теги на зображенні можуть забезпечити швидший зворотний зв'язок і що метод менш технічний. Він також зазначив, що Google Goggles надає погані теги. Google придумав засіб для боротьби з пошуком зображень. Він представив Google Goggles, програмне забезпечення візуального розпізнавання для Android, яке додає функцію створення нотаток, допомагаючи користувачам мати історію пошуку зображень, які вони зробили раніше. Програма все ще досить обмежена в можливостях розпізнавання. Наразі він може розпізнавати лише типові зображення чи об'єкти, які вже зберігаються в Інтернеті. Він покладається на теги, коли Goggles не може знайти подібне зображення.

Постановка проблеми Було запроваджено пошук за зображеннями, але точність результатів низька та дає багато небажаних записів. Тому він не може надати спеціалізовані точні ресурси зображення. Крім того, мобільним телефонам важко обробляти зображення, на що користувачі витрачають значну кількість часу, не отримуючи те, що їм потрібно. Технологія обробки зображень мобільного телефону все ще дуже неефективна і все ще не може забезпечити відповіді в реальному часі.

3.1 Вирішення проблеми з виявленням зображення Цілі цього проекту;

- Щоб інтегрувати високоефективну програму на основі архітектурних зображень у реальному часі.
- Уніфікувати ефективну технологію обробки, зберігання та аналізу зображень.
- Щоб покращити використання доповненої реальності (відео, GPS) у телефонах Android, дозволяючи покращити результат, коли йдеться про зображення.

4.0 Огляд літератури Багато людей розробляють і впроваджують програми на різних рівнях з різних причин і мають намір оптимізувати результати для передбачуваного користувача. Оскільки комп'ютерна обробка стає швидшою та дешевшою, можливості для програм на основі зображень стають ще

цікавішими. Ця технологія широка, оскільки охоплює питання безпеки за допомогою розпізнавання обличчя, перевірки людини або навіть виявлення емоцій [5]. Проте були зроблені інші спроби для успішного виявлення, це включає зусилля, зроблені платформою Windows, у якій вони використовують редактор зображень із рівнем сірого, де вводять зображення низької якості та перетворюють їх на зображення високої якості. Це може включати зміну розміру зображення, перевертання зображень або навіть обертання [6]. В інших дослідженнях дослідники використовували метод виявлення зображень у зв'язку з сонячними зображеннями [7]. Це був більш екологічний підхід, який в основному використовував камеру, сонячні промені та відповідні кути для обробки та розрахунку зібраного зображення для розробки напрямку і висота. Є докази того, що виявлення зображень можна використовувати як вирішення проблем зі здоров'ям [8]. Це передбачає виявлення на ультразвукових зображеннях шляхом фокусування на регіонах інтересу (ROI). Це може полегшити раннє виявлення та діагностику захворювань. У нас є інші нові методи, які повністю відрізняються від традиційних [9]. Зображення виявляються блоками, а не звертають увагу на зображення в цілому. У дослідженні використовується «присвоєння нечіткої функції», де модель може бути виражена у змінних, але це лише техніка (вектор). У таблиці 1 нижче показано типи виявлення, як описано в розділі 4.0.

Таблиця 1 - типи виявлення

Тип виявлення	Опис виявлення
Розпізнавання обличчя	Неточно (потрібна база даних)
Редактор зображень (Windows)	Забезпечує мінімум функцій (відтінки сірого)
Сонячні зображення	Залежить від кліматичних змін
Ультразвукові зображення	Обмеження форми виявлення. Тільки для конкретного завдання
Блокувати зображення (змінні)	технічна форма

Мобільна платформа (Android, Palm OS, Windows Mobile, Linux і Symbian) Мобільні пристрої обіцяють працювати зараз і в майбутньому. Завдяки таким пристроям ми отримуємо ефективність і зручність. Мобільні платформи розширюють бренди завдяки спрощенню перегляду та допомагають деяким збільшити дохід. Зростання важливості мобільних пристроїв викликало конкуренцію між програмними гігантами, такими як Google, Apple, Palm і Microsoft. З моменту запуску яблучного Ios і Android від Google ринок вибухнув або виріс експоненціально. Однак ця частина дослідження має на меті охопити всі сфери, включаючи платформи Nokia (Symbian) і платформу Blackberry, яка швидко поширюється (ОС від RIM). Проведено розбивку на рисунок 2 нижче.

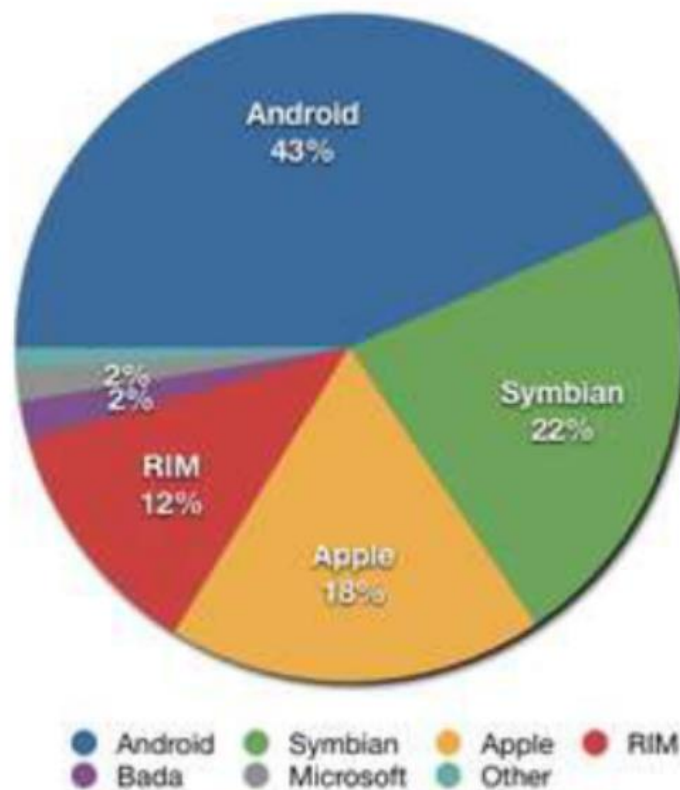


Рисунок 2 Кругова діаграма Розподіл різних ОС

Доповнена реальність на основі мобільних пристроїв MobileAR — це нова сфера мобільних додатків, яка швидко розвивається. Це дозволяє інтегрувати віртуальну реальність на основі інформації та фізичну реальність

реального світу. У цьому документі розглядається, як здійснюється це з'єднання та які практичні та приємні застосування можна зробити за допомогою цієї технології. Наразі Google Android і iPhone мають можливість накладати інформацію на зображення, надаючи додаткові функції.

### Методологія

Методологія розробки. Методом розробки для цього дослідження є інкрементний оборотний життєвий цикл розробки програмного забезпечення (IR-SDLC). Цей метод було обрано, оскільки він тісно відповідає намірам дослідження шляхом покращення роботи в режимі реального часу. Рішення дозволяє фіксувати організаційні події, які можуть перерости в інциденти, оцінювати критичність інцидентів і призначати реагування на основі впливу та нормативних вимог. Таким чином користувачі також можуть консолідувати процедури реагування, керувати наскрізними розслідуваннями та звітувати про тенденції та пов'язані з ними інциденти. Крім того, IR-SDLC підтримує функцію обслуговування, яка дозволяє розробникам виправляти недоліки та підвищувати ефективність програми. IR-SDLC визначено на малюнку 5 нижче.

Методологія тестування приймається пізніше в проєкті. Тип методу, який буде використовуватися в цьому проєкті

це прототип програми. Проєкт не може бути реалізований або схвалений, якщо він не кваліфікується як безпомилковий або нульовий недоліки. Фаза тестування проєкту буде зосереджена в основному на дизайні програми та взаємодії. Зворотній зв'язок буде збиратися відповідно. Прототип буде переглянуто, як показано нижче (прототип програми, як на рисунку 3):

- Створення прототипу;
- Оцінка прототипу;
- Доопрацювання прототипу.

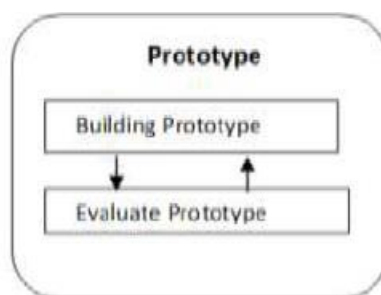


Рисунок 3 - Прототип програми

#### Список використаних джерел

- [1] V V Vinod, 1996, Focussed Color Intersection with Efficient Searching for Object Detection and Image Retrieval, Retrieved on October 4, 2011
- [2] Michael A. Smith, 1997, Video Skimming and Characterization through the Combination of Image and Language Understanding Techniques, Last visited on September 18, 2011
- [3] Stefan Huwer, 2000, Adaptive Change Detection for Real-Time Surveillance Applications, Lastvisited on September 21, 2011
- [4] ChuanQinyx, June 28–July 1, 2011, TagSense: A Smartphone-based Approach to Automatic Image Tagging, Retrieved September 12, 2011
- [5] Zahir Larabi, 2009, Efficient data access management for FPGA-Based image processing SoCs, Reviewd on October 4th 2011
- [6] Omer M. Soysal, 2010, An Image Processing Tool for Efficient Feature Extraction in Computer-Aided Detection Systems, last visited on October 4th 2011
- [7] Luciano Godoy Fagundes, 2010, Development of Computer Graphics and Digital Image Processing Applications on the iPhone, Reviewd on October 4th 2011
- [8] Hua Zhang, 2011, Design and Implementation of Digital Image Processing System, Last visited on October 4th, 2011
- [9] Lu Guiming, 2011, a New Additive Fuzzy System for Image Processing, retrieved on 5th October 2011