

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

Факультет комп'ютерної інженерії та управління  
(повна назва)

Кафедра електронних обчислювальних машин  
(повна назва)

**КВАЛІФІКАЦІЙНА РОБОТА**  
Пояснювальна записка

Рівень вищої освіти другий (магістерський)

Методи машинного навчання  
для визначення облич на зображеннях  
(тема)

Виконав:

студент II курсу, групи СПм-22-1  
Педич Д.І.  
(прізвище, ініціали)

Спеціальність 123 «Комп'ютерна інженерія»  
(код і повна назва спеціальності)

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

Освітня програма Системне програмування  
(повна назва освітньої програми)

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

Допускається до захисту  
Зав. кафедри ЕОМ

Коваленко А.А.  
(прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет \_\_\_\_\_ комп'ютерної інженерії та управління \_\_\_\_\_

Кафедра \_\_\_\_\_ електронних обчислювальних машин \_\_\_\_\_

Рівень вищої освіти \_\_\_\_\_ другий (магістерський) \_\_\_\_\_

Спеціальність \_\_\_\_\_ 123 «Комп'ютерна інженерія» \_\_\_\_\_  
(код і повна назва)

Тип програми \_\_\_\_\_ освітньо-професійна \_\_\_\_\_  
(освітньо-професійна або освітньо-наукова)

Освітня програма \_\_\_\_\_ Системне програмування \_\_\_\_\_  
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри \_\_\_\_\_  
(підпис)

“ \_\_\_\_\_ ” \_\_\_\_\_ 20\_\_ р.

## ЗАВДАННЯ

### НА КВАЛІФІКАЦІЙНУ РОБОТУ

студенту \_\_\_\_\_ Педичу Дмитру Ігоровичу \_\_\_\_\_  
(прізвище, ім'я, по батькові)

1. Тема роботи Методи машинного навчання для визначення облич на зображеннях

затверджена наказом по університету від “ 06 ” листопада 2023 р. № 1299Ст

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

3. Вхідні дані до роботи \_\_\_\_\_

Computer Vision

Бібліотека OpenCV

Існуючі технології розпізнавання облич на зображенні

Сучасні реалізації систем визначення облич

Метод Віюли-Джонса

4. Перелік питань, що потрібно опрацювати у роботі \_\_\_\_\_

Аналіз проблем та існуючих рішень розпізнавання обличчя

Аналіз видів алгоритмів розпізнавання

Технології розпізнавання обличчя осіб за допомогою методу Віюли-Джонса

Розробка оптимізації методу

Експеримент та тестування розробленого рішення

Висновки

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) \_\_\_\_\_

Слайд-презентація – 17 слайдів \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

6. Консультанти розділів роботи (заповнюється за наявності консультантів згідно з наказом, зазначеним у п.1 )

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

### КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Аналіз завдання	07.11.2023-09.11.2023	
2	Аналіз науково-технічної літератури	10.11.2023-14.11.2023	
3	Пошук існуючих рішень	15.11.2023-25.11.2023	
4	Розробка і тестування методу гібридизації	26.11.2023-19.12.2023	
5	Оформлення пояснювальної записки	20.12.2023-05.01.2024	
6	Оформлення графічної частини	05.01.2023-08.01.2024	

Дата видачі завдання 06 листопада 2023 р.

Студент \_\_\_\_\_  
(підпис)

Керівник роботи \_\_\_\_\_ ст. викл. Єршоміна Н.С.

## РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 86 с., 32 рис., 2 табл., 1 дод., 28 джерел.

КОМП'ЮТЕРНИЙ ЗІР, ТЕХНОЛОГІЇ РОЗПІЗНАВАННЯ ОБЛИЧЧЯ, OPENCV, АЛГОРИТМИ МАШИННОГО НАВЧАННЯ, МЕТОД ВІОЛІДЖОНСА, ОПТИМІЗАЦІЯ.

Метою кваліфікаційної роботи є розробка метода для створення системи розпізнавання обличчя. Розроблена система буде мати більш гнучкий, швидший та точніший результат роботи з подальшим впровадженням з метою вдосконалення та більш зручнішого і доступного використання.

У ході виконання кваліфікаційної роботи було проведено аналіз існуючих рішень з розпізнавання обличчя на сучасному ринку технологій. Розглянуто принципи роботи, переваги, недоліки, можливі напрямки подальшого розвитку, сфери використання і відповідно розробка системи з використанням методу машинного навчання для визначення облич на зображенні.

## ABSTRACT

Master's thesis: 86 pages, 32 figures, 2 tables, 1 appendices, 28 sources.

COMPUTER VISION, FACE RECOGNITION TECHNOLOGIES, OPENCV, MACHINE LEARNING ALGORITHMS, VIOLA-JONES METHOD, OPTIMIZATION.

The major goal of this thesis is to develop a method for creating a face recognition system. The developed system will have a more flexible, faster and more accurate work result with further implementation for the purpose of improvement and more convenient and accessible use.

In the course of the qualification work, an analysis of existing face recognition solutions on the modern technology market was conducted. The principles of operation, advantages, disadvantages, possible directions of further development, areas of use and, accordingly, the development of a system using the machine learning method for identifying faces in the image are considered.

## ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ .....	8
ВСТУП .....	9
1.1 Загальні відомості.....	11
1.2 Визначення поняття комп'ютерний зір.....	12
1.3 Аналіз технології розпізнавання обличчя.....	15
1.3.1 Характерні завдання для CV.....	15
1.3.2 Структури CV.....	16
1.4 Використання CV .....	18
1.4.1 Amazon Rekognition .....	18
1.4.2 DeepFace.....	20
1.4.3 FaceNet .....	21
1.4.4 FaceID .....	23
1.4.5 Microsoft face API.....	25
1.4.6 Samsung Face Recongition.....	27
1.5 Постановка задачі.....	28
2 ОГЛЯД ПРЕДМЕТНОЇ ОБЛАСТІ. ЗАГАЛЬНІ ТЕОРЕТИЧНІ ВІДОМОСТІ. РОЗГЛЯД СИСТЕМ ВИЗНАЧЕННЯ ОБЛИЧ .....	29
2.1 Основні завдання системи розпізнавання обличчя .....	29
2.2 Аналіз алгоритмів .....	32
2.2.2 Метод Fisherfaces .....	42
2.2.3 Алгоритм LBRH .....	44
2.2.4 Метод Віоли-Джонса.....	46
2.3 Глибоке навчання.....	52
2.5 Порівняння розглянутих методів .....	53
3 ПРИСКОРЕННЯ МЕТОДУ ВИЗНАЧЕННЯ ОБЛИЧ .....	56
3.1 Варіанти прискорення .....	56
3.2 Базова версія алгоритму SIMD .....	57

3.3 Варіанти векторизації алгоритму .....	59
3.3 Scalar, SSE 4.1, AVX2, AVX512 .....	61
4 ДОСЛІДЖЕННЯ МЕТОДУ ВИЗНАЧЕННЯ ОБЛИЧ .....	65
4.1 Розробка програмного забезпечення.....	65
4.2 Тестування методу .....	65
4.2.1 Підготовка вхідних даних .....	65
4.2.2 Результати тестування .....	69
ВИСНОВКИ.....	73
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ .....	74
ДОДАТОК А Графічний матеріал кваліфікаційної роботи.....	77

## ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

CV – комп'ютерний зір(англ. Computer Vision).

PCA – метод головних компонент(англ. Principal component analysis).

LBRH – локальні бінарні шаблони (англ. Local Binary Patterns, LBP) .

ШІ – штучний інтелект.

API – прикладний програмний інтерфейс(англ. Application programming interface).

CNN – згортова нейронна мережа(англ. Convolutional neural network).

ML – машинне навчання (англ. Machine Learning).

СО – системи охолодження.

AVX – розширення системи команд процесору для доповнення інструкцій набору AVX (англ. Advanced Vector Extensions).

SIMD – одиночний потік команд, множинний потік даних (англ. Single instruction, multiple data).

SSE – потокове розширення одиночного потіку команд SIMD (англ. Streaming SIMD Extensions).

ЛДА – лінійний дискримінантний аналіз.

## ВСТУП

Наше обличчя є предметом для вивчення та дослідження вже великий проміжок часу історії людства, оскільки воно є головною та унікальною характеристикою людини та здатне передавати значну частку інформації про її особистість та індивідуальність. Завдяки обличчю, кожен з нас має змогу зрозуміти погляд та емоції іншої людини і здатні їх розрізнити заздалегідь до того, як вони можуть бути показані.

Зі стрімким розвитком інформаційних та комп'ютерних технологій у сучасності, можливості аналізу зображення значно розширилися та мають великий перелік різних інструментів, алгоритмів та підходів для аналізу образів обличчя людини.

Алгоритми розпізнавання образів дозволяють ідентифікувати об'єкти, визначати їх тип, надають необхідну якісну і кількісну інформацію. Задача розпізнати обличчя – одна із перших практичних задач, яка вплинула на розвинення теорії машинного розпізнавання образів об'єктів. Беручи до уваги кількість проведених досліджень і експериментів у даній області, які проводяться у світі протягом останніх десятиліть, станом на зараз досі не знайдено методів, які матимуть надійну якість обчислень – здатність розпізнавати й ідентифікувати людське обличчя зі сталою точністю, не зважаючи на різні вхідні умови.

Комп'ютерний зір – це теорія створення і навчання машин, з метою виконувати виявлення, проводити спостереження та класифікувати об'єкти на зображеннях і відео. Системи комп'ютерного зору все ширше впроваджуються та використовуються для вирішення повсякденних завдань у різноманітних сферах нашого життя.

Важливою проблемою у системах комп'ютерного зору в області розпізнавання облич є вагома залежність якості і точності результату від умов освітленості зображення, масштабу, ракурсу спостереження, виразу безпосередньо самого обличчя, наявності перешкод на обличчі, таких як окуляри або вуса, перекриття сторонніми об'єктами у кадрі і погодні умови.

У світі досі не існує досконалої формули знаходження обличчя і підрахунку кількості учасників на фото, але є кілька методів розпізнавання, оцінки та підрахунку людей на зображенні.

Технологія розпізнавання обличчя на фото і оцінки кількості людей на зображенні разом зі здатністю розрізняти варіанти міміки, емоцій людини та відстеженням фокусування погляду людини є непростим завданням для реалізації у світі технологій, але має величезний потенціал для розвитку та вдосконалення.

Методи розпізнавання облич і підрахунку кількості людей на зображенні мають широке застосування у робототехніці, системах відеоспостереження, у системах безпеки і контролю доступу, у верифікації особи, онлайн банкінгу, в інтерфейсах взаємодії між технічним приладом і людиною, у сфері розваг, криміналістиці та медицині. Завдяки такій практичній затребуваності розпізнавання облич людей на зображенні є дуже актуальною областю досліджень.

# 1 АНАЛІЗ ЛІТЕРАТУРНИХ ДЖЕРЕЛ. ПОСТАНОВКА ЗАДАЧІ ПЕРЕДАТЕСТАЦІЙНОЇ РОБОТИ

## 1.1 Загальні відомості

Штучний інтелект (ШІ) — це галузь інформатики, що вивчає створення систем, які здатні виконувати завдання, які зазвичай потребують людського інтелекту. Ця галузь розвиває та застосовує методи, які дозволяють комп'ютерам аналізувати дані, вирішувати проблеми, навчатися та приймати рішення на основі взаємодії з навколишнім середовищем.

До списку основних аспектів ШІ, які пов'язані з темою дослідження роботи, було виділено для розгляду наступні:

- машинне навчання: використання алгоритмів та моделей, які навчаються на даних, щоб робити прогнози, виявляти патерни та приймати рішення;

- глибоке навчання: підгалузь машинного навчання, яка використовує нейронні мережі з багатьма шарами для вирішення складних завдань, таких як розпізнавання облич, обробка людської мови, симуляція поведінки;

- робототехніка: розробка та використання роботів, які можуть виконувати завдання, використовуючи ШІ для взаємодії з навколишнім середовищем.

Штучний інтелект має застосування у багатьох галузях нашого життя, включаючи сферу медицини, технологій, фінансів, транспорту, виробництва, наукових досліджень та багато інших прикладів застосувань. Він прискорює розвиток нових технологій, сприяє автоматизації процесів та вирішенню складних завдань, що раніше вважалися виключно людськими.

Сучасний стан розвитку ШІ знаходиться у процесі поглибленого вдосконалення вже існуючих та впровадження нових етапів розробки, створення та вбудовування відеоінформаційних технологій, які базуються на використанні та обробці вхідного зображення.

Через потреби впровадження штучних інтелектуальних систем, які здатні самостійно візуально орієнтуватися у просторі, які можуть виконувати візуальний аналіз сцен навколишнього середовища і які можуть знаходити і відстежувати рухомі, або нерухомі об'єкти одночасно з оцінюванням їх характеристик, форм та кількості, можна дійти висновку, що відеоінформаційний напрямок має велику актуальність у сучасному світі.

## 1.2 Визначення поняття комп'ютерний зір

Комп'ютерний зір – це теорія створення і навчання машин, з метою можливості виконання задач виявлення, проведення спостереження та класифікування об'єктів на зображеннях і відео[28].

CV використовує теоретичні та технологічні принципи створення штучних систем, які здатні отримувати вхідну інформацію у графічному вигляді. Дані можуть бути представлені різними шляхами, наприклад у формі зображення з камери, тривимірного зображення зі сканеру, або відеопослідовністю.

Комп'ютерний зір у якості технологічної дисципліни, слугує основою під час створення систем CV. Системи комп'ютерного зору все ширше використовуються для вирішення повсякденних завдань у різноманітних сферах нашого життя, починаючи від промисловості, завершуючи медициною і розвагами. До прикладів структур CV можна віднести наступні:

- механізми управління процесами(автономні транспортні засоби);
- системи відеоспостереження;
- комплекси організації інформації;

- структури відтворення варіацій об'єктів або зовнішньої середовища (процес аналізування медичних знімків, тощо);
- системи взаємодії (одна з найзатребуваніших – система взаємодії людина-машина).

Задача розпізнати обличчя – одна із перших практичних задач, яка вплинула на розвиток теорії розпізнавання образів. Алгоритми розпізнавання образів дозволяють ідентифікувати об'єкти, визначати їх тип, надають необхідну якісну і кількісну інформацію.

Незважаючи на велику кількість досліджень і дослідів у даній області, які проводяться у світі протягом останніх десятиліть, досі так і не було знайдено методів, які матимуть надійну якість роботи – виявляти й розпізнавати обличчя людини з великою сталою точністю, за різних умов.

Важливою проблемою у системах комп'ютерного зору в області розпізнавання обличчя є вага залежність якості і точності результату від умов освітленості зображення, масштабу, ракурсу спостереження, виразу безпосередньо самого обличчя, наявності перешкод на обличчі, таких як окуляри або вуса, перекриття сторонніми об'єктами у кадрі і погодні умови.

Задача розпізнавання обличчя наразі містить у собі два можливих напрямки подальшого розвитку: автоматичне виділення обличчя на зображенні та ідентифікація особистості за допомогою обличчя. Обличчя, яке виділено на зображенні може бути використане у подальшому для аналізу, який включатиме у себе завдання з віртуального моделювання обличчя, відстеження і розпізнавання виразу та емоцій на ньому, розпізнавання віку та статі людини, з метою верифікації і аутентифікації особи.

CV, машинна обробка та аналіз зображень, функція зору у робота і машинний зір – дуже тісно пов'язані між собою області. Але досі немає точної класифікації, чи є дані області розділами однієї галузі, або є представниками окремих напрямків. Під час ретельного аналізу, можна дійти висновку, що це можуть бути тільки відмінні одна від одної назви загальної сфери.

Для того, аби не заплутатися, та не помилятися, прийнято виділяти їх як напрямки, спрямовані кожен окремо на певну дисципліну для подальшого вивчення. Пропоную ознайомитися з деякими найбільш важливими, на мою думку, галузями.

Обробка та аналіз зображень переважно зорієнтовані на роботу з двовимірними графічними зображеннями та спрямовані на зміни в одному зображенні з метою отримання іншого. Даний перелік функцій містить у своєму складі різноманітні попиксельні операції, наприклад, такі операції, як підвищення контрастності, пошук і виділення країв, прибирання з зображення сторонніх шумів, а також геометричні перетворення. Також у літературних джерелах зазначено, що ці операції передбачають обробку та аналіз зображень, незалежно від конкретного змісту самого зображення.

Галузь застосування комп'ютерного зору фокусується на обробці тривимірних сцен, що можуть бути відтворені на основі одного або кількох зображень. Наприклад, це включає відновлення структури або іншої інформації про тривимірну сцену на основі одного або декількох вхідних зображень.

Галузь застосування машинного зору акцентується переважно на застосуваннях у промисловому секторі, наприклад у автономній роботі на підприємстві та системи візуальної перевірки та вимірювання. Звідси, можна дійти висновку, що технічне устаткування датчиків здатних розпізнавати зображення та принципи управління пов'язані з обробкою відеоданих для управління роботами, а обробка даних в реальному часі може здійснюватися як апаратно, так і програмно.

Також важливо зауважити, що технологія розпізнавання образів представляє собою галузь, яка застосовує велику кількість різних підходів для захоплення інформації, яка надходить із потоку відеоданих, переважно ґрунтуючись на статистичному підході.

## 1.3 Аналіз технології розпізнавання обличчя

### 1.3.1 Характерні завдання для CV

Усі сфери із переліку використання CV, що були розглянуті у роботі раніше, безпосередньо суміжні зі списком різних задачами. Багато з цих завдань, пов'язаних з процесом визначення та опрацювання інформації, мають можливість знаходити рішення за допомогою різноманітних методів. Нижче буде приведено декілька прикладів типових задач CV.

Проблема визначення наявності конкретного об'єкта, особливості чи активності у відеоданих є класичною задачею в області комп'ютерного зору, обробки зображень та машинного зору. Хоча людина може легко та надійно впоратися з цією задачею, комп'ютерні системи поки що мають труднощі у її задовільному вирішенні.

У проаналізованому списку джерел наукової літератури, можна виділити наступні проблеми:

- визначення: перевірка відеоданих на відповідність визначеній умові. Наприклад, це може включати виявлення можливих аномалій в клітинах чи тканинах на медичних зображеннях. Методи виявлення, які базуються на відносно легких і ефективних варіантах розрахунках, іноді використовуються для локалізації невеликих областей на аналізованому зображенні. Далі ці області можуть бути піддані більш складному аналізу за допомогою методів, що вимагають більше обчислювальних засобів, з метою здобуття коректного опису;

- проблема розпізнавання об'єктів: відмінне визначення одного або кількох об'єктів, що були попередньо задані або вивчені;

- проблема ідентифікації об'єкту: визначення конкретного екземпляра об'єкта. Наприклад, визначення конкретного обличчя, відбитка пальців або автомобіля.

Існує декілька конкретизованих завдань, що ґрунтуються на процесі розпізнавання, такі як:

- пошук змістових зображень: визначення всього можливого переліку зображень серед значного обсягу графічних матеріалів, які володіють конкретним змістом;

- розгляд розташування: визначення розташування чи орієнтації конкретного об'єкта відносно позиціонування об'єктиву камери;

- ідентифікація символів: розпізнавання позначень на знімках друкованого формату тексту чи власноручнаписанного тексту часто, з метою конвертації його в текстовий формат представлення даних, який є найбільш зручним для подальших дій з коригування чи індексації інформації;

- активність: низка невеликої кількості завдань, пов'язаних із визначенням руху, включають до свого складу обробку послідовності зображень у відеопотоці для визначення швидкості і позиції кожної точки на зображенні, тривимірного простору або навіть самої об'єктиву камери, яка веде зйомку;

- відтворення події: використання двох або більше зображень сцени чи відеоданих з метою створення тривимірного простору навколо об'єкта;

- коригування зовнішнього вигляду зображень: завдання відновлення зображень полягає в усуненні шумів (наприклад, шуму датчика, розмиття об'єкта під час руху тощо). Одним з найбільш простих методів для вирішення цієї проблеми є використання різноманітних типів фільтрів.

### 1.3.2 Структури CV

Впровадження структур CV значно змінюється в залежності від сфери їх застосування. Одні системи функціонують автономно, розв'язуючи конкретні завдання визначення та кількісного аналізу, тоді як інші вбудовуються як підсистеми в більші системи.

Практичне впровадження структури додатково обумовлено насамперед видом її функціональності – чи являє собою вона заздалегідь визначену функціональність, чи перелік деяких її нюансів та функціональність можуть бути перероблені та замінені у процесі її реалізації та виконання.

Проте, існують загальні функції, які характерні великій кількості структур CV:

- захоплення візуальних зразків: цифрова версія вхідного зображення зчитується з відповідних сенсорів таких як світлочутливі камери різних типів, датчики, які вимірюють відстань до об'єкта, спеціальне устаткування у вигляді радарів, камери, що фіксують ультразвукові коливання навколишнього середовища та інші. Дані, які отримує система, насамперед підпорядковані технічному типу устаткування, та можуть представляти звичайне 2D або тривимірне 3D зображення, або послідовність декількох зображень. Значення пікселів вхідного зображення часто корелюють з яскравістю світла, в одному чи кількох спектральних діапазонах, представлені у формі кольорового зображення або зображення в градаціях сірого кольору палітри;

- етап предопрацювання: перед застосуванням методів комп'ютерного зору для видобуття конкретної інформації з відеоданих, необхідно обробити ці дані таким чином, щоб вони відповідали визначеним вимогам, що напряду підпорядковуються обраному методу;

- розпізнавання дрібних фрагментів: із потоку відеоданих виділяються деталі з різним рівнем складності, наприклад, такі як лінії, межі, тощо. Складніші деталі можуть включати в себе структури, форми або рух;

- розпізнавання та виділення сегментів зображення: на відповідному етапі обробки зображення приймається характерне рішення про те, які точки або області на зображенні вважати важливими для подальшого аналізу. Наприклад, це може включати виділення конкретного набору точок, які є об'єктом інтересу.

## 1.4 Використання CV

Для проведення дослідження та аналізу стану систем з розпізнавання обличчя людини, пропонуємо розглянути сучасні технології розпізнавання облич, що використовуються такими великими компаніями як Google, Apple, Amazon, Microsoft, Facebook.

Дані корпорації регулярно розповідають свої теоретичні відкриття в галузях штучного інтелекту, розпізнавання зображень й аналізу обличчя, з метою швидкого донесення принципів роботи технологій ідентифікації та аутентифікації користувачів.

### 1.4.1 Amazon Rekognition

Amazon Rekognition – це сервіс, що розміщений у хмарі, від компанії Amazon, реалізує розпізнавання обличчя людини. Користувачі змогли отримати доступ до цієї послуги вперше у 2016 році.

А вже у 2018 р. компанія Amazon повідомила, що вона просуває для правоохоронних органів власну службу Rekognition з розпізнавання облич людей, принцип роботи якої базується на хмарі. Представлена технологія може розпізнавати до 100 людей на зображенні і може виконувати зіставлення розпізнаних людських облич з БД, у якій містяться фотографії десятків мільйонів людей.

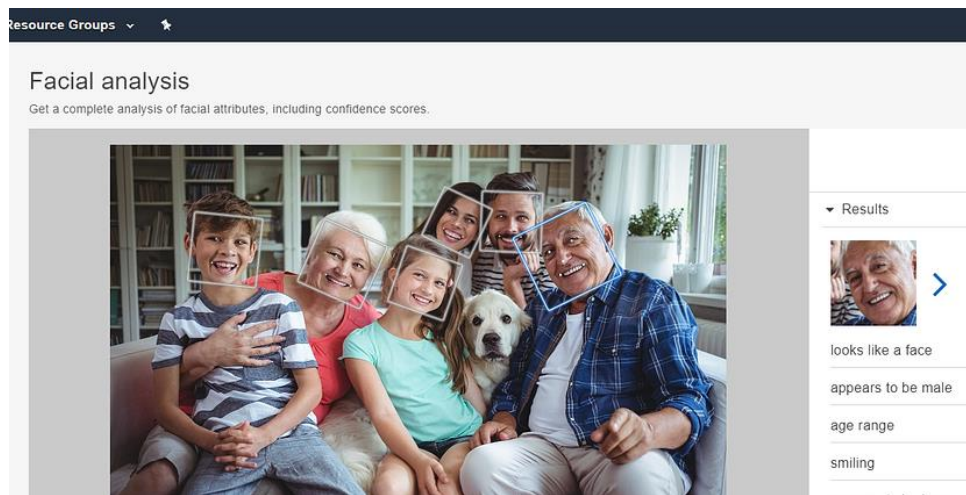


Рисунок 1.1 - Приклад роботи сервісу Amazon Recongnition

Сервіс від компанії Amazon застосовує методи глибокого навчання для аналізу та ідентифікації об'єктів на вхідному зображенні.

Amazon Recongnition використовується деякими урядовими агенціями США, таких як Міграційна та Митна служби, поліція у місті Орlando, а також приватними компаніями. Послуги, які надає сервіс, можна розглядати у двох категоріях: перша включає алгоритми, що опрацьовані на інформації, отриманої компанією Amazon, і остання категорія - методи, які клієнт може навчити за допомогою власного набору даних [8]. У категорії алгоритмів, які можна тренувати включений ряд можливостей:

- розпізнавання відомих людей на зображеннях;
- розпізнавання людського обличчя включає до свого переліку функцій характеризувати стать людини, значення її віку, класифікувати наявну емоцію на обличчі людини, наявність бороди чи вус на обличчі, наявність сонячних чи окулярів для зору та інші характеристики;
- моніторинг за людиною на відеопотоці. Дана можливість дозволяє збирати статистичні дані стосовно гравців на спортивному полі;
- визначення текстового змісту на зображеннях;
- виявлення контенту, який є неприпустимим, небажаним або образливим.

Серед алгоритмів, які можна власноруч навчати, наявні наступні:

- алгоритм SearchFaces: даний алгоритм надає можливість завантаження користувачеві власного датасету, що містить зображення для подальшого навчання і тренування моделі машинного навчання. Одразу після завершення навчання, користувач має можливість завантажити потрібне йому фото з метою одержати інформацію стосовно наявності і ідентифікації обличчя людини на зображенні;
- ідентифікація особи користувачів через розпізнавання обличчя.

#### 1.4.2 DeepFace

У 2014 р. Facebook презентував свою програму, яку назвали DeepFace, яка може визначати, чи належать два сфотографованих обличчя одній людині з рівнем точності у 97%.

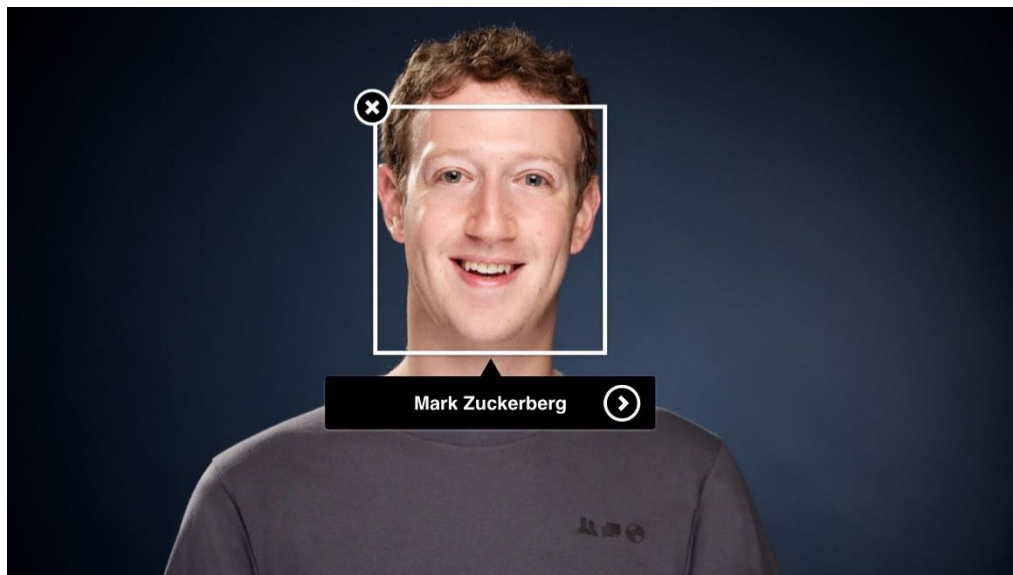


Рисунок 1.2 – Демонстрація роботи програми DeepFace

Технологія DeepFace – це технологічна система розпізнавання облич, яка була розроблена компанією Facebook.

Користувачі вперше отримали доступ до даного сервісу у 2015 році, і вже протягом 9 років система постійно вдосконалюється та розширює свої можливості. Вона базується на алгоритмах глибокого навчання та включає дев'ятишарову нейронну мережу із понад 120 мільйонами вагових з'єднань.

Початкове навчання нейронної мережі DeepFace відбулося на 4 мільйонах фото користувачів безпосередньо соцмережі Facebook. Через те, що користувачі соціальної мережі продовжують активно додавати власну різноманітну кількість фотографій, датасет сервісу DeepFace буде завжди поповнюватися, що обумовить поступове зростання точності обчислень використовуваного алгоритму визначення обличчя на фото. Група співробітників компанії Facebook з технічного відділу повідомила, що станом на зараз, їхня система має обчислювальну точність у 97,35 % на датасеті LFW [11].

Одночасно із цим, здатність розпізнати людське обличчя безпосередньо людиною досягає близько 97%. Порівнявши два результати, можна стверджувати, що технологія DeepFace може ефективніше впоратися із цим питанням.

Технологія аналізує фотографії, викладені у соцмережі Facebook, на яких присутній користувач даної соцмережі, насамперед головне фото його профілю і створює для цього користувача його власне ID - унікальний числовий код. Далі дане ID буде використовуватися для ідентифікації обличчя даного юзера на інших фотографіях його профілю. Система розпізнавання має у своєму складі функцією захисту користувачів від можливого зловживання особистими даними інших осіб [12]. Крім того, DeepFace може надсилати повідомлення особам з порушеннями зору, інформуючи їх, хто зображений на фото чи відео.

### 1.4.3 FaceNet

В 2015 р. Google запропонував технологію FaceNet. FaceNet використовує штучну нейронну мережу та новий алгоритм розпізнавання, завдяки чому вдалося зв'язати обличчя зі своїм власником з майже ідеальною точністю.

Дана технологія використовується у Google Photos, з метою сортування зображень та автоматичного зіставлення їх конкретній особі на основі зображень облич уже визнаних технологією людей.

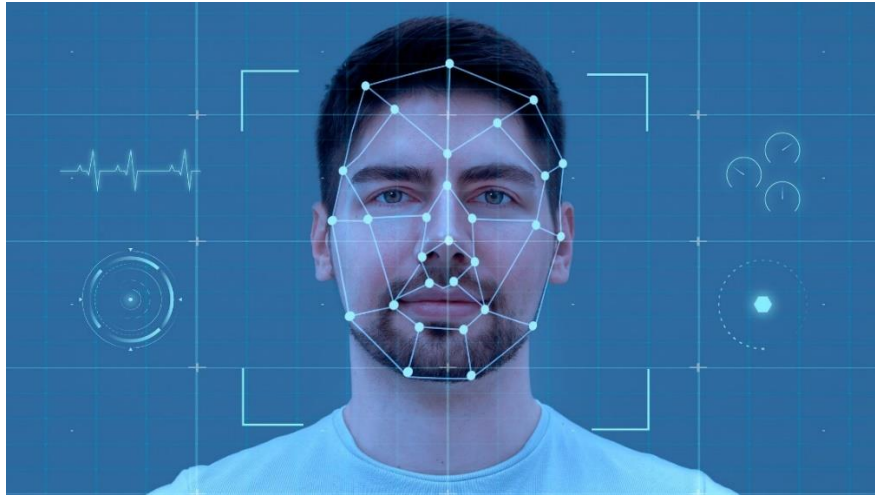


Рисунок 1.3 – Технологія FaceNet

FaceNet - це технологія відома своїми унікальними методами для вирішення завдань розпізнавання облич, розроблена Дейвом та командою вчених з Google, зокрема з групи Google Research.

Основна мета FaceNet полягає в створенні векторних представлень облич, які представляють кожне обличчя у вигляді числового вектора. Це представлення у просторі дозволяє визначати схожість облич, а також робити рішення про ідентифікацію чи аутентифікацію.

Технологія FaceNet базується на глибоких нейронних мережах, зокрема на зглиблених нейронних мережах, що забезпечує її здатність ефективно вирішувати завдання розпізнавання облич навіть у складних умовах, таких як змінне освітлення, кути огляду та інші варіації.

Ключові особливості технології FaceNet:

- векторні представлення: кожне обличчя перетворюється на вектор у просторі, де подібні обличчя розташовані близько одне до одного, що сприяє легкості порівняння та розпізнавання;

- тренування на великому обсязі даних: модель була навчена на великій кількості зображень з різних джерел, що дозволяє їй розпізнавати обличчя різних людей;

- стійкість до змін: FaceNet є стійким до змін в умовах зйомки, таких як фон, освітлення та інші фактори.

FaceNet використовується у системах розпізнавання облич в різних сферах, включаючи безпеку, автоматизований контроль доступу, системи відеоспостереження та інші області, де потрібне точне та ефективне розпізнавання облич для вирішення різних завдань.

#### 1.4.4 FaceID

Face ID – є технологією розпізнавання обличчя людини, розроблена та випущена компанією Apple вперше у 2017 році для мобільних пристроїв.

Зараз, окрім iPhone дану технологію підтримують і iPad. Система надає можливість ідентифікувати обличчя за біометричними показниками з метою розблокування пристрою.

Перший раз ця система мала застосування у iPhone X, який був випущений у 2017 році.

Апаратні потреби технології Face ID включають до свого складу тримодульний датчик, який, в свою чергу, складається з точкового проектора, принцип роботи якого полягає у виведенні сітки інфрачервоних точок на обличчя людини (як свідчить технічна документація даної технології, їх кількість перевищує 30 000), з модуля, що випромінює інфрачервоне світло на обличчя, і з інфрачервоної камери, головна задача якої полягає у тому, що необхідно зафіксувати інфрачервоне зображення, яке випромінює обличчя користувача завдяки інфрачервоному модулю та проаналізувати отриманий шаблон обличчя і створити тривимірну мапу обличчя.

Мапа отриманого обличчя зпівставляється з зареєстрованим обличчям у режимі реального часу за допомогою захищеної підсистеми, і людина отримує доступ, якщо обидва знімки обличчя у вигляді мапи збігаються. Потрібно наголосити на тому факті, що використання інфрачервоної камери у технології значно поліпшує результат роботи системи при низькому рівні освітлення або в темряві.



Рисунок 1.4 – Рисунок прикладу роботи системи FaceID

При першому налаштуванні параметрів безпеки девайсу, у вигляді технології FaceID, обличчя його власника буде двічі проскановано під різними кутами, з метою створення точнішої контрольної мапи обличчя.

Відділ безпеки компанії Apple зауважив, що жодна мапа обличчя не буде записуватися у хмарне сховище, з чого можна дійти висновку, що дані стосовно зображення обличчя будуть зберігатися у зашифрованому вигляді безпосередньо на пристрої користувача[9].

Якщо пристрій залишається неактивним протягом 48 годин, або здійснено 5 невдалих спроб автентифікації, функція FaceID тимчасово вимкнеться, і користувач матиме змогу розблокувати пристрій тільки після вдалого введення паролю.

Розроблена компанією Apple система вміє розпізнавати обличчя людини разом із вдягненими на нього окулярами, маскою, за наявності макіяжу, або бороди і має змогу звикати до змін зовнішнього вигляду користувача із часом. Даний функціонал підтримується нейронним процесором, який був розроблений з метою прискорення виконання завдань машинного та глибокого навчання.

Apple не так давно, оновила свою розробку NeuralEngine у вигляді четвертого покоління нейронних процесорів, за рахунок отримання ще більших показників продуктивності та менших показників енергоспоживання даних процесорів.

Apple виробила вже 4 покоління Neural Engines, щоразу з виходом нового покоління, покращуючи показники продуктивності та ступінь енергоефективності даних процесорів.

Система Face ID дозволяє користувачам використовувати сервіс Apple Pay для здійснення платежів. Крім того, ця технологія забезпечує доволі велику точність захоплення виразу обличчя користувача у реальному часі і може використовувати отриману інформацію стосовно положення обличчя для застосування різноманітних фільтрів і масок.

#### 1.4.5 Microsoft face API

Microsoft Face API - це сервіс від Microsoft, який надає можливість розпізнавання облич та аналізу зображень з використанням штучного інтелекту.

Цей сервіс пропонує набір інструментів для обробки зображень з метою визначення облич та отримання різноманітної інформації про обличчя на фотографіях або відео.

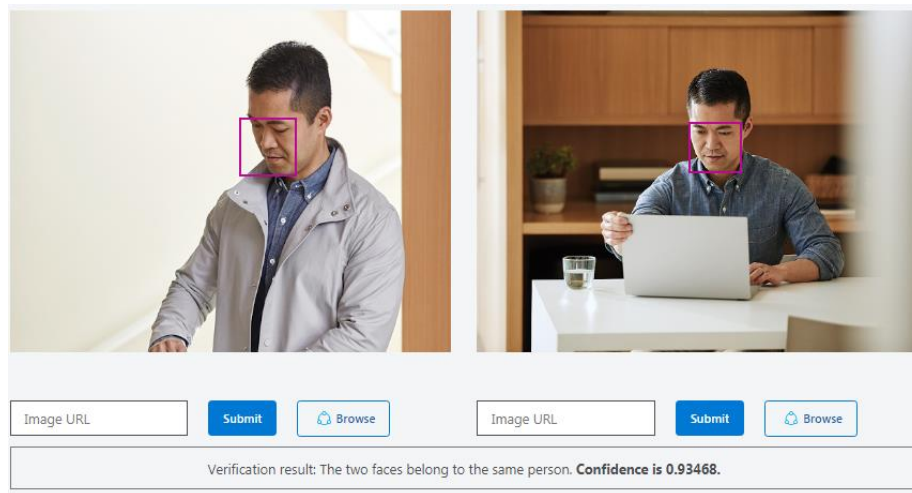


Рисунок 1.5 – Приклад роботи додатку Face API

Основні можливості та функціональність Microsoft Face API включають:

- розпізнавання облич: сервіс може ідентифікувати та розпізнавати обличчя на зображеннях, виявляти основні ознаки та характеристики облич, включаючи очі, ніс, рот та інші деталі;
- визначення емоцій: сервіс може аналізувати вирази облич та визначати емоції, такі як радість, смуток, здивування, гнів тощо;
- вікова та статева група: можливість визначення вікової категорії та статі особи на зображенні;
- визначення схожості: спроможність порівнювати обличчя та визначати, наскільки вони схожі чи відмінні одне від одного;
- визначення унікальних ідентифікаторів: призначення унікальних ідентифікаторів обличам для можливості їх відстеження в різних зображеннях.

Цей сервіс може бути використаний у різних сферах, таких як безпека, маркетинг, розваги та інші галузі, де важливо впровадження функціоналу розпізнавання облич та аналізу зображень для різноманітних цілей.

Дана технологія надає зручний інтерфейс для роботи з цими функціями через API, що дозволяє легко інтегрувати його у різні програмні рішення та додатки.

Це робить розробника незалежним від необхідності мати апаратне забезпечення та займатися навчанням алгоритмів, оскільки всі ці процеси автоматично відбуваються на сервері Microsoft. Однак цей підхід змушує розробника дотримуватися правил і вимог структури та відповідей цього API під час розробки.

#### 1.4.6 Samsung Face Recognition

Samsung Face Recognition – це технологія розпізнавання облич, яка використовується в різних пристроях Samsung, таких як смартфони та планшети, для розблокування екрану та аутентифікації користувачів за допомогою сканування та розпізнавання їх облич.

Перша реалізація технології розпізнавання облич у пристроях Samsung була запроваджена у пристрої Galaxy S8, який був випущений у 2017 році. Face Recognition став одним із методів біометричної аутентифікації користувача, разом з скануванням відбитків пальців та іншими методами розблокування пристрою. Samsung протягом наступних років постійно вдосконалювала цю технологію в своїх наступних моделях смартфонів та планшетів, забезпечуючи більш високу точність та безпеку розпізнавання облич.



Рисунок 1.6 – Технологія Samsung Face Recognition

Основні особливості системи розпізнавання облич Samsung включають:

- обличчя як метод аутентифікації: дозволяє користувачам використовувати своє обличчя для розблокування пристрою, підтвердження особи та доступу до певних функцій або даних на пристрої;
- способи сканування: використовує різні методи, такі як сканування через фронтальну камеру пристрою для захоплення облич користувача та подальше їх аналізу для визначення ідентичності;
- зручний та швидкий доступ: користувачі можуть швидко розблокувати пристрій, просто вказавши на своє обличчя, без необхідності вводити паролі чи використовувати відбитки пальців;
- заходи безпеки: технологія розпізнавання облич Samsung намагається забезпечити високий рівень безпеки для запобігання несанкціонованому доступу до пристрою.

Ця технологія розпізнавання облич використовується для полегшення процесу розблокування та забезпечення безпеки пристроїв Samsung. Однак, варто пам'ятати, що не всі системи розпізнавання облич однаково точні чи надійні, і вони можуть мати свої обмеження або потребувати покращень в плані безпеки.

## 1.5 Постановка задачі

Задачею і метою передатестаційної роботи є ретельний та детальний поглиблений аналіз існуючих рішень методів машинного навчання для вирішення задачі розпізнавання обличчя.

Також головною метою є розгляд переваг, недоліків і можливостей покращення вже існуючих алгоритмів навчання і методів розпізнавання обличчя та обґрунтоване обрання методу за його показниками для розпізнавання обличчя.

## 2 ОГЛЯД ПРЕДМЕТНОЇ ОБЛАСТІ. ЗАГАЛЬНІ ТЕОРЕТИЧНІ ВІДОМОСТІ. РОЗГЛЯД СИСТЕМ ВИЗНАЧЕННЯ ОБЛИЧ

### 2.1 Основні завдання системи розпізнавання обличчя

У загальних випадках, алгоритм, що вирішує задачу розпізнавання обличчя людини на зображенні містить в своєму складі наступні кроки:

- виявлення факту присутності обличчя людини на зображенні і його подальше виділення;
- розпізнавання й опис ключових характеристик людського обличчя, до яких відносяться очі, ніс, вуха, брови, рот, тощо;
- подання обличчя в деякому просторі ознак;
- порівняння виділеної ділянки зображення під обличчя з еталонами обличчя і його подальша ідентифікація.

Розпізнавання обличчя та подальше його виділення на зображенні є провідним етапом усіх методів, які пов'язані з розпізнанням людського обличчя. Мета даного методу полягає в наступних задачах: визначення присутності на вхідному зображенні людського обличчя, і у результаті підтвердження присутності відбувається вказання його розташування і встановлення знаходження межових ліній щодо усіх без винятку розпізнаних облич на вхідному зображенні. Обличчя, яке було ідентифіковано у результаті виконання алгоритму виділення обличчя, повинно бути представлено у вигляді моделі, яку можна використовувати на подальших етапах аналізу[28].

Усі наведені алгоритми ідентифікації людського обличчя за варіантом представлення обличчя можна розгалужити на наступні три класи методів:

- метод зіставлення в цілому: даний метод є найбільш поширеним класом методів для розпізнавання облич. Він розглядає усю площу людського обличчя у ролі вхідних даних для процесу розпізнавання. Алгоритм не аналізує окремо компоненти та сегменти обличчя. Також він не тільки ігнорує ракурс, з

якого представлено обличчя на зображенні, а й інші атрибути, такі як вираз обличчя або емоції на ньому. Найпоширенішим із алгоритмів у даній групі, і в той же час першим, є метод власних облич [2,6]. Всі методи цієї категорії конструюють проекцію вхідного зображення у визначеному підпросторі, і для цього використовують методи головних компонент [2,3,4], лінійний дискримінантний аналіз [5,6], аналіз незалежних компонентів[7], або вейвлет-перетворення[8,9]. Перевагою даних методів є невідкидання інформації безпосередньо в самому зображенні обличчя концентруючи увагу тільки на областях, які обмежені, чи на характерних точках даного зображення. Проте дана технологія обумовлює і наступні недоліки: ці методи є апаратно-затратними, потребують великого рівня взаємозв'язку між тестовими та навчальними зображеннями та відзначаються доволі низькою ефективністю під час значної зміни кількості поз, масштабу всього зображення і ступеню освітленості на зображенні;

- методи зіставлення за ознаками або структурні методи: функції цього методу застосовують характеристики виділення обличчя та його структуру у якості вхідних даних для розпізнавання. В цих функціях насамперед визначаються характеристичні риси людського обличчя, такі як: очі, нос, губи і рот, шия та підборіддя, тощо. Наступним кроком відбувається врахування деяких структурних характеристик обличчя, наприклад таких як геометрично-пропорціональні розміри, коефіцієнт їх співвідношення із шаблонними показниками або наявність різноманітно-можливих зовнішніх особливостей людини). Таким чином, зображення після його опрацювання, виражається у формі вектору структурних ознак. Використані характеристики можуть містити в своєму складі вагової області зображення, вагового контуру чи вагової точки.

В цьому класі методів складність полягає у проблемі відновлення характеристик обличчя, які неможливо виявити на зображенні, особливо у випадку, коли певна частина обличчя прихована. До даної групи методів класифікації належать наступні:

- моделі обличчя з гнучкими контурами: вони використовуються для розпізнавання, ґрунтуючись на принципі порівняння контурів різних частин обличчя, наприклад, таких як лінії голови, вух, губ, носа, брів, очей та ін. частин обличчя;

- аналіз еластичних графів: включає в себе порівняння обличчя, яке представлено у вигляді графа, де вершини останнього розташовані в ключових точках обличчя, таких як контури голови, губ, носа та в їх точках екстремумів;

- методи, що ґрунтуються на геометричних характеристиках обличчя: вони визначають набір ключових точок чи областей на людському обличчі, після чого відбувається визначення характеристик, кожна з яких представляє собою відстань між вагомими точками або представляється відношенням таких відстаней. До головних переваг даних методів можна віднести невразливість до зміни положення позиції обличчя на зображенні, над яким реалізовується опрацювання, а також до масштабу вхідного зображення, його орієнтованості і ступеню освітлення на зображенні. Можна відзначити як недолік складність автоматичного вибору оптимального набору ознак для розпізнавання;

- комбіновані підходи: концепція і принцип роботи даних методів була створена з аналогії людського зору, які розглядають усі можливі із наявних характеристик обличчя, так і весь образ обличчя. Даний тип методів використовує комбінацію порівняння об'єктів взагалі та за конкретними ознаками. Крім того, за підходом до розпізнавання всі ці методи можна класифікувати у дві групи: статистичні та інтелектуальні.

Розглядаючи статистичний підхід, головна мета якого полягає у класифікації вхідних зображень за допомогою порівняння вхідного зображення шляхом обчислення статистичних характеристик порівнюваного зразка з еталонними зображеннями у базі даних.

В свою чергу, інтелектуальні методи підходу цієї групи, які насамперед спрямовані на класифікацію або знаходження облич, застосовують у своєму складі нейронні мережі, спираються на приховані марківські моделі та на існуючі методи машинного навчання[28].

Процес розпізнавання обличчя включає в себе виконання наступних етапів. [13, 14]:

- виявлення обличчя: дана задача виникає незалежно від того, чи в умовах наявне завдання розпізнати особу на фото, відеозаписі чи за допомогою інших методів візуального спостереження;
- орієнтація обличчя у просторі: зараз доволі рідко зустрічається на фотографіях зображення обличчя, де людина стоїть прямо до об'єкту, найчастіше обличчя нахиле. Отже, з'являється проблема, яка полягає у позиціонуванні обличчя так, щоб здавалося, що фотографія зроблена прямо;
- виділення унікальних деталей обличчя: на цьому етапі проводиться аналіз зображення для отримання унікальних цифрових характеристик обличчя;
- встановлення особи: порівнюємо отримані дані з вже існуючими записами у БД, і якщо ступінь схожості є задовільним, то виводимо дані про цю особу.

## 2.2 Аналіз алгоритмів

Обличчя представляє собою найбільш виразну рису характеристики особи. Звідси, можна вважати не дивним той факт, що аналіз обличчя використовується при вирішенні широкого спектру завдань, таких як:

- виявлення облич на зображенні[9];
- розпізнавання емоцій[10];
- класифікувати вираз та емоції обличчя;
- відстеження уваги, с саме напрямок погляду очей людини[11].

Окрім самої фотографії обличчя, широко використовуються його ключові точки[12].

Кількість та розташування цих точок можуть змінюватися залежно від обраної задачі та наявних баз даних.

Зазвичай використовують 68 точок обличчя (Рисунок 2.1).

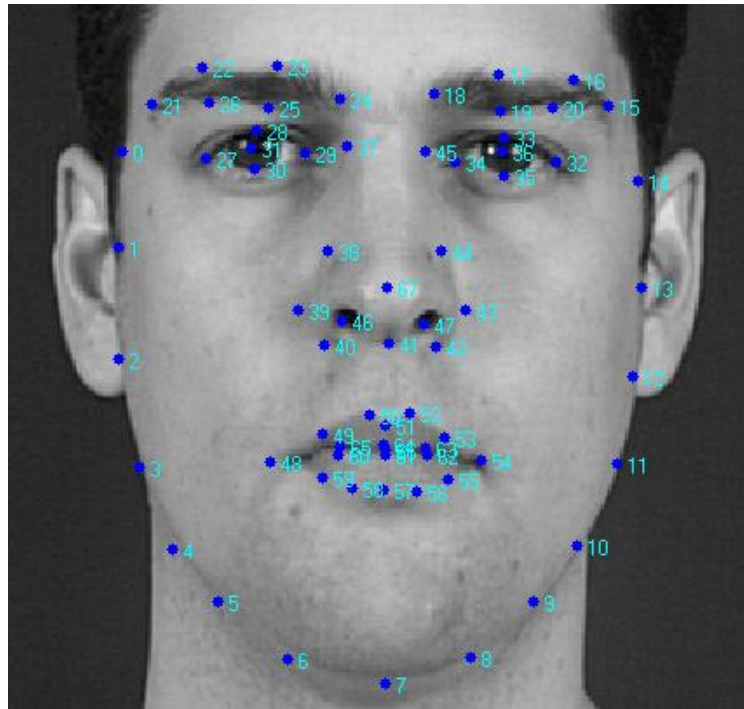


Рисунок 2.1 – Приклад використання 68 КТ

Серед переліку можливих варіацій використання даного підходу, присутній варіант із 56 ключовими точками обличчя. Даний варіант розпізнання обличчя наведено на рисунку 2.2.

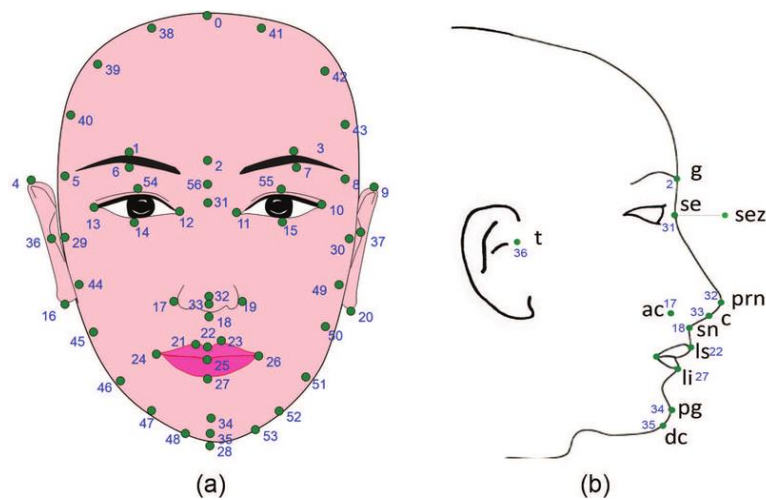


Рисунок 2.2 – 56 ключових точок обличчя

Зараз дедалі частіше починають використовувати підхід із варіантом наявності 98 ключових точок на обличчі людини. Даний варіант розпізнання обличчя наведено на рисунку 2.3.

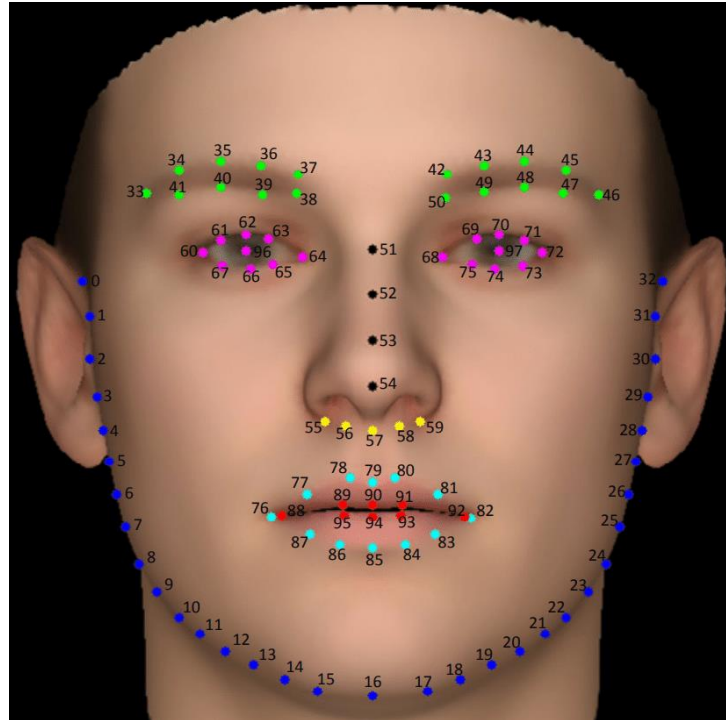


Рисунок 2.3 – 98 ключових точок обличчя

Існують різноманітні способи реалізації ключових точок.

Як вже зазначалося, вибір кількості та розташування точок залежить від конкретної задачі, доступних баз даних і наявних ресурсів. Однак найбільш поширеними є саме 68 і 56 точок.

Але, як зазначено раніше, у найбільш поширених варіантах використання даного набору точок для розпізнавання обличчя, зокрема, у таких бібліотеках, як DLIB і OpenCV, є певні недоліки: вони неефективно визначають точки у випадках, коли голова на фотографії повернута або є дифіцит освітлення, як показано на рисунку 2.4.

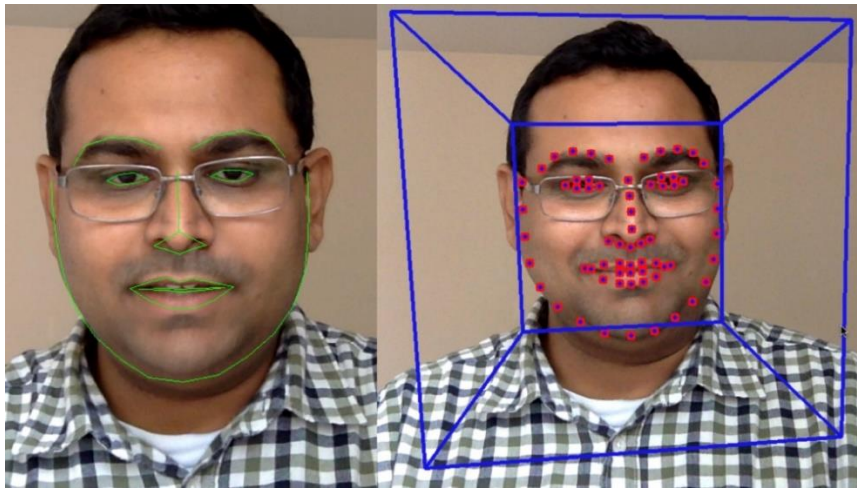


Рисунок 2.4 – Результат знаходження 68 точок dlib при повороті голови

### 2.2.1 Метод Eigenfaces

У 1991 році Меттью Турк та Алекс Пентленд внесли вагомий внесок у вирішення завдань і проблем розпізнавання обличчя людини за допомогою розробленого методу, який отримав назву Eigenfaces[13].

Для часу, початку 1990-их років, метод Eigenfaces був легкий у використанні та мав непоганий ступінь ефективності власної роботи. Основний принцип роботи алгоритму полягає у використанні базових характеристик, що стосуються статистики, а саме математичного очікування і коваріаційної матриці, а також методу головних компонент (РСА). РСА був вперше запропонований науковцем і математиком Карлом Пірсоном у 1901 році, проте він й до цього дня використовується з метою зменшення показника розміру у наборі даних. Графічний матеріал у відтинках сірого розмірністю  $p \times q$  можна виразити у  $p * q$  - вимірному векторному просторі.

Наприклад, якщо надане зображення із розмірами масштабу у  $100 \times 100$  пікселів, воно буде представлене у просторі розмірністю  $100^2$ .

Виконання перетворень з даними значеннями потребує значних апаратних ресурсів, що будуть задіяні у процесі обчислень.

При збільшенні розмірів масштабу вхідного зображення пропорційно збільшується і кількість числа вимірювань, однак далеко не уся кількість вимірів будуть містити інформацію, корисну для подальшого її використання у процесі аналізу. Метод PCA використовується для виконання операцій, які перетворюють колекцію пов'язаних змінних у меншу колекцію непов'язаних змінних. Головна мета заключається у тому, що многовимірні набори даних здебільшого описують за допомогою колекцій пов'язаних змінних, і тому дуже невелика кількість вимірів міститимуть корисну інформацію.

Детальний опис представленого алгоритму:

- створення набору даних із фотографіями, де наявні людські обличчя. Наприклад, ми будемо розглядати чорно-білі версії зображення з представленого набору даних LFW, який містить зображення людей (Рисунок 2.5) [14]. В протилежності від кольорових зображень, що включають три канали (червоний, зелений і синій), градація сірого має тільки один канал;

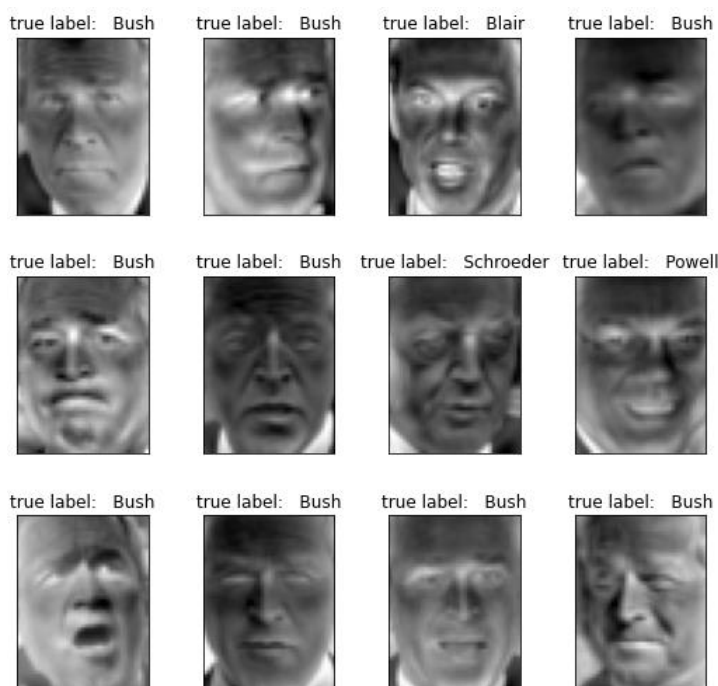


Рисунок 2.5 – Приклад зображення із набору даних LFW

- перетворення обличчя людини на зображенні у векторне представлення обличчя. Будь яке довільне зображення у Grey Scale містить у своєму складі піксельну матрицю, при цьому визначення окремого пікселя лежить у проміжку  $[0, 255]$ , де значення 0 – це значення чорного кольору, а 255 – значення білого кольору на пікселі. Піксельну матрицю у такому вигляді дуже легко перевести у векторне представлення. Наприклад, дано сталу кількість зображень  $p$  у вхідному наборі даних із розмірністю  $[N \times N]$ . За формулою:  $[N^2 \times \text{вектор}]$  перетворимо розмірність наданого зображення  $[x_2, x_2]$  у векторне представлення. Зовнішній вигляд вектору після операцій перетворення матиме наступний вигляд, як на рисунку 2.6;

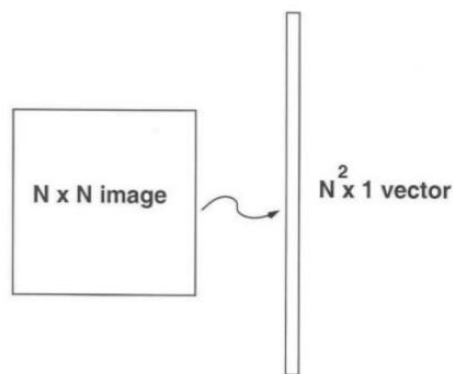


Рисунок 2.6 – Приклад алгоритму перетворення фото у вектор

- нормалізація отриманого завдяки алгоритму перетворення вектора людського обличчя. Даний етап має власну реалізацію у два кроки. Насамперед, відбувається визначення значення усередненого обличчя по наступній формулі:

$$\psi = \frac{1}{p} \sum_{i=1}^m x_i$$

при цьому  $p$  – загальне число кількості зображень у вхідному наборі даних,

$x_i$  – значення вектору обличчя. Усереднене обличчя включає в себе характеристики, які є загальними для всієї кількості значень облич у вхідному наборі даних. Наступний крок – математична операція віднімання значення усередненого обличчя від значення окремо кожного вектору за наступною формулою:

$$a_i = x_i - \psi$$

Далі, після операції віднімання, з отриманих значень векторів, створюється новий вектор  $V$  за наступною формулою:

$$V = [v_1, v_2, \dots, v_p]$$

Даний вектор матиме розмірність  $[N^2 \times p]$  та включатиме до свого складу усі значення векторів із даною розмірністю;

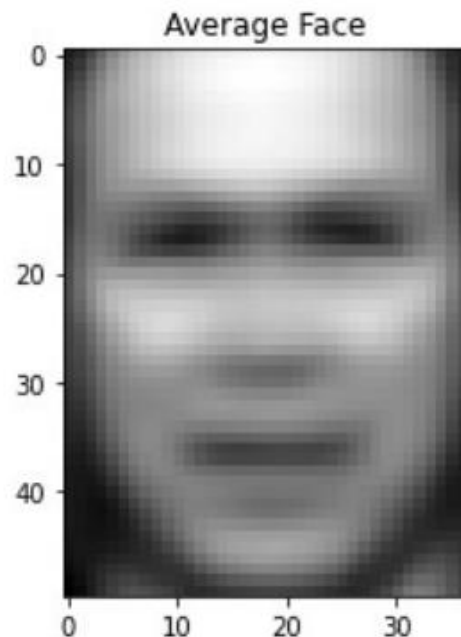


Рисунок 2.7 – Зображення, що містить приклад усередненого обличчя

- наступний крок обчислень – вираження значень матриці коваріацій  $M$  завдяки операції помноження матриці  $V$ , яка містить в собі значення різниці векторів, на матрицю  $V^T$  яка є транспонованою матрицею по відношенню до  $V$ . Розмірність матриці  $V$   $[N^2 \times p]$ , отже розмірність  $V^T$   $[p \times N^2]$ . Звідси, отриманий результат в процесі помноження матриць матиме наступний вигляд значення розмірності:  $[N^2 \times N^2]$ . Зі зростанням значень  $N$ , операція з розрахунку матриць вимагатиме більші апаратні потреби забезпечення. Через це, прийнято знаходити значення матриці коваріації у вигляді операції  $V^T$  і  $V$ . А значення розмірності даної матриці буде визначатися  $[p \times p, p < N^2]$ ;

- ще один етап полягає у визначенні власних значень та значень власних векторів для матриці коваріації  $M'$  за наступною формулою:

$$M'^{u_i} = \lambda_i u_i, i = 1, 2, \dots, p$$

При цьому значення  $M'$  обчислюється як  $V \times V^T$ ;

- упорядкування власних векторів за спаданням відповідних власних значень. На рисунку 2.8 представлена відібрана довільна кількість  $L$  власних векторів – кожне значення якого є власним значенням, що містить найбільше значення ( $L < P$ );

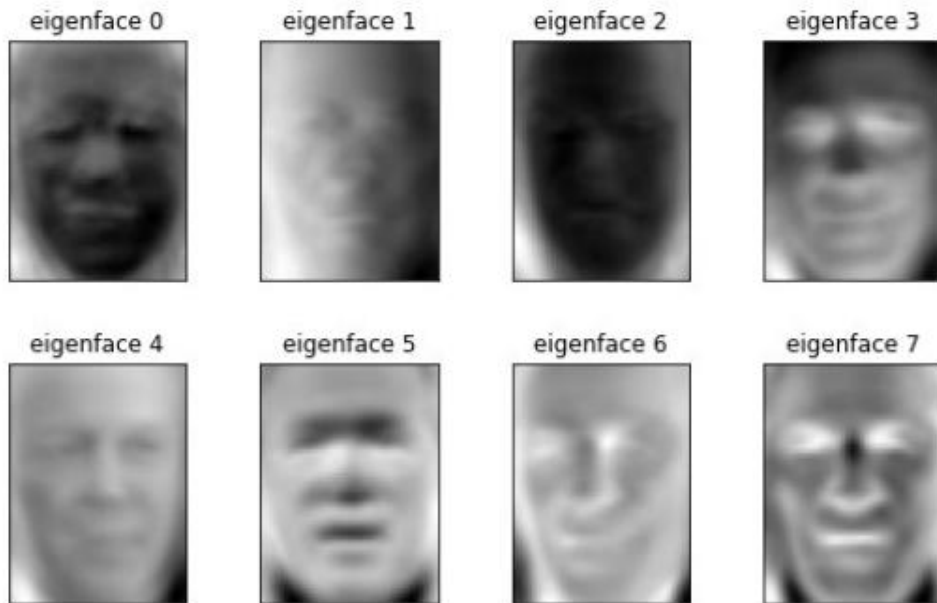


Рисунок 2.8 – Зображення із прикладом вибору найбільших значень векторів  
 - перетворення усієї кількості власних векторів  $L$  із меншею розмірністю до форми представлення початкової розмірності за наступною формулою:

$$[u_i = V \times v_i]$$

- представлення зображення образу кожного обличчя на зображенні, як суми лінійних комбінацій кількості  $L$  власних векторів. Будь яке зображення обличчя, наявного у вхідному наборі даних, можна виразити як математичну операцію суми зважених сум кількості  $L$  і значення усередненого обличчя  $\psi$ :

$$x_i = \sum_{j=1}^L w_j u_j + \psi$$

де  $u_j$  – це вибірка найбільших значень власних векторів, а  $w_j$  – ваги.

Зовнішній вигляд застосування такого перетворення зображено на рисунку 2.9:



Рисунок 2.9 – Людське обличчя, після лінійних перетворень

Після завершення процесу навчання алгоритму, можна переходити до процедури визначення невідомих облич на зображенні. Покроковий алгоритм дій, пов'язаних з розпізнаванням обличчя з використанням технології Eigenfaces виглядає наступним чином:

- потрібно передобробити зображення невідомого обличчя, а саме виконати наступний перелік дій, застосувавши їх до зображення, таких як відцентрування зони лиця на зображенні та адаптування розмірів усього зображення до стандартних розмірів, заданих алгоритмом;
- вираження зображення, що містить обличчя людини, обробленого у кроці №1, у векторну форму преставлення;
- виконання дій нормалізації вектору обличчя за пунктом №3 алгоритму навчання;
- операція з проектування вже отриманого нормалізованого вектору на зображення обличчя людини з метою одержання показників порівняння лінійної комбінації та значень зваженого вектору;
- визначення відмінності між векторами для обличчя вхідного набору даних зображень та вектором невідомого обличчя.

Якщо ця відмінність менша за задану алгоритмом похибку, вважається, що обличчя розпізнане; в іншому випадку обличчя людини на зображенні залишається невідомим[15].

### 2.2.2 Метод Fisherfaces

Алгоритм розпізнавання обличчя Fisherfaces є вдосконаленою версією методу Eigenfaces. Одним із вагомого переліку недоліків методу головних компонент, який використовується в вищевикладеному методі Eigenfaces, безумовно є операція по визначенню складових, що взаємопов'язані між собою. Гарним прикладом цієї проблеми слугує ситуація, коли можуть бути наявні дві, або більше фотографії однієї і тієї ж особи при різних умовах освітлення, а алгоритм Eigenfaces може ідентифікувати обличчя, не дійшовши висновку, що це одна і та ж сама людина.

Як показали тестування на практиці, алгоритм розпізнавання Fisherfaces справляється з поставленою задачею краще, ніж його попередник – метод Eigenfaces завдяки вбудованому у його алгоритм роботи ЛДА.

Головна суть даного методу – розташувати об'єкти, що належать до спільного класу максимально близько один до одного, а об'єкти, що відносяться до різних класів – максимально далеко. Кожна людина посідає місце класу, а зображення, де наявне її обличчя, належить до об'єктів класу.

Загальний вигляд архітектури методу:

- підбір вхідного набору даних;
- представлення вхідних зображень із набору даних у векторній формі.

Наприклад, дано сталу кількість  $K$  класів, звідси вектор  $X$ , що представлятиме усі наявні класи, тоді вектор, що представлятиме кількість у вигляді одного зображення –  $x_i$ :

$$X = \{X_1, X_2, \dots, X_K\}$$

$$X_i = \{x_1, x_2, \dots, x_n\}$$

- наступний етап полягає у операції розрахунку наступних двох середніх значень: універсального середнього значення  $\mu$  та значення середнього показника, що відноситься до кожної категорії класифікації облич  $\mu_i$ :

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i$$

$$\mu_i = \frac{1}{|X_i|} \sum_{x_j \in X_i} x_j$$

Звідси, відбувається операція по розрахунку матриці з розсіювання, яка містить у своєму складі класи  $S_B$  і розрахунок матриці розсіювання у просторі класу  $S_W$ :

$$S_B = \sum_{i=1}^c N_i (\mu_i - \mu)(\mu_i - \mu)^T$$

$$S_W = \sum_{i=1}^c \sum_{x_j \in X_i} (\mu_i - \mu)(\mu_i - \mu)^T$$

- пошук проекції, де відстань між екземплярами класів буде максимальною;

- вираження власних значень, власних векторів та трансформованої матриці.

На рисунку 2.10 наведено приклад, де представлено відповідні матриці з операції розсіювання для наступних трьох класів.  $S_{B1}, S_{B2}, S_{B3}$  містять значення матриці, яке безпосередньо відображає міру розсіювання між даними класами, а в свою чергу, матриці  $S_{W1}, S_{W2}, S_{W3}$  – демонструють ступінь розсіювання в

межах конкретного класу. Метод розпізнавання Fisherfaces сприяє збільшенню відстані між об'єктами різних класів класифікацій обличчя, та групуванню екземплярів класу, які належать до спільного класу[17].

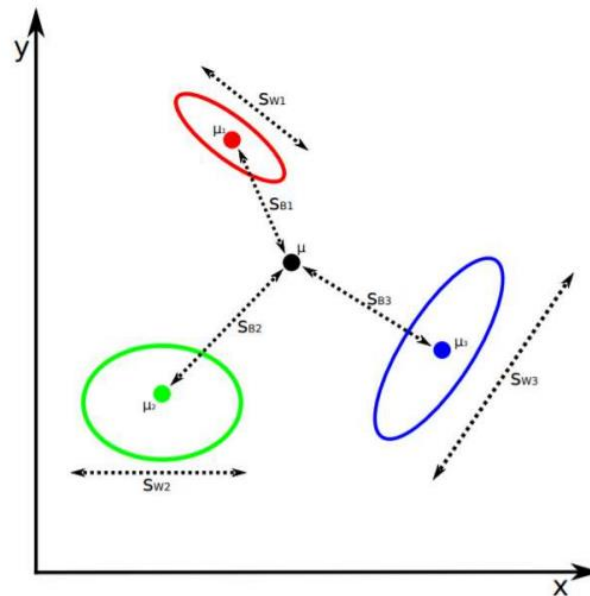


Рисунок 2.10 – Графічне зображення позиціонування матриць з операцій розсіювання

### 2.2.3 Алгоритм LBRH

LBP – є легким і з високим ступенем продуктивності, елементом управління, що застосовується з метою класифікувати надані текстури на зображенні. Принцип роботи алгоритму LBP згідно офіційних джерел, з'явився ще у 1994 році.

Пізніше, під час проведення досліджень та практичних експериментів, було зроблено висновок, що компонування елементу управління з класифікації текстур разом з гістограмою направлених градієнтів значно збільшує показники ефективності роботи алгоритму.

На відміну від розглянутих вище в роботі алгоритмів Eigenfaces і Fisherfaces, які аналізують зображення повністю, даний метод аналізує у вхідному зображенні кожен піксель, що робить його менш чутливим до умов освітленості ніж його попередники[18].

Принцип роботи алгоритму базується на основних чотирьох коефіцієнтах:

- перший коефіцієнт – це радіус, який знаходить своє застосування у проектуванні локального бінарного шаблону, у вигляді кола, всередині якого знаходиться предмет огляду – піксель;
- другий параметр – це кількісна величина точок, для того щоб мати змогу побудувати локальний бінарний шаблон;

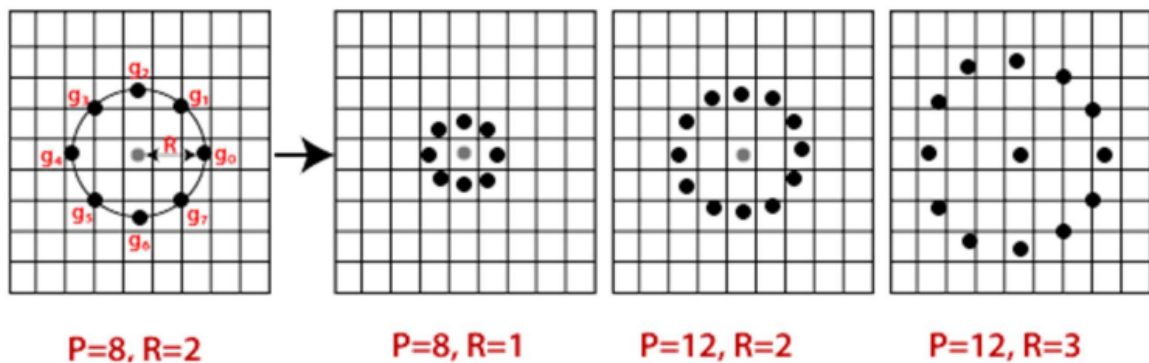


Рисунок 2.11 – Приклад роботи алгоритму LBRN з різними параметрами радіусу та кількістю точок для побудови бінарного шаблону

- третя змінна являє собою суму кількості клітинок осі  $\{OX\}$ , які займає локальний бінарний шаблон;
- четверта змінна аналогічна третій, тільки тепер урахування суми клітинок йде по осі  $\{OX\}$ .

Головним і першопочатковим етапом роботи із алгоритмом обов'язково виступає навчання алгоритму на вхідному наборі даних із зображеннями облич.

Наступний етап – операція перетворення вхідного зображення у тимчасове зображення. Для вирішення цього питання застосовується метод

рухомого вікна, основні величини якого задаються радіусом області, яку необхідно охопити та кількістю сусідів на обраному проміжку після чого визначаються бінарні значення пікселів у цьому вікні. Центральний піксель – є головною областю, з якою буде відбуватися порівняння сусідніх. На рисунку 2.12 рухоме вікно представлено фігурою із площею 9 пікселів.

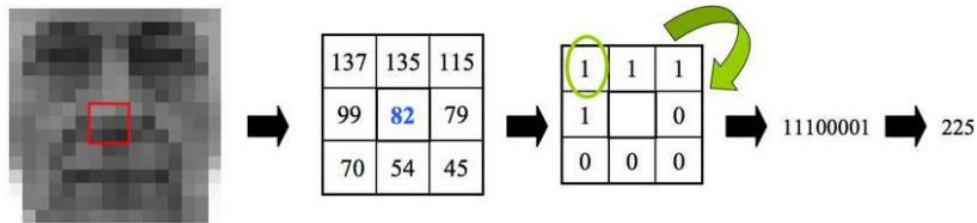


Рисунок 2.12 – Приклад роботи методу рухомого вікна

Версія зображення, яке отримане в результаті роботи даного методу, слугує основою для гістограми. Слід зауважити, що для окремого зображення будується власна гістограма. Гістограми нових невідомих зображень порівнюються з зображеннями вже існуючих, які були створені під час навчання алгоритму. Гістограми, які майже подібні, характеризують, що фото обличчя належить одній людині.

#### 2.2.4 Метод Віоли-Джонса

У 2001 році, П. Віола та М. Джонс розробили власний метод розпізнавання облич на зображенні. Даний алгоритм все ще залишається потужним і доволі ефективним інструментом для виявлення облич, не зважаючи на його вже більш ніж двадцятирічне існування.

Навіть після впровадження сучасних технологічних рішень з розпізнавання облич на зображенні у вигляді глибокого навчання, даний метод складає конкуренцію за показниками точності обчислень, ступеню ефективності і коефіцієнтом швидкості виконання обчислень.

Також, хочу зауважити, що необхідно прийняти до уваги той факт, що алгоритм Віоли-Джонса спрямований на саме виявлення місцезнаходження обличчя на вхідному зображенні.

Загальний вигляд архітектури побудови методу:

- основою для обчислень та з метою зменшення швидкості обчислень застосовуються інтегральні зображення фотознімків;
- щоб проаналізувати наявність людського обличчя на зображенні, застосовуються ознаки Хаара;
- з метою обрання певних видів класифікацій ознак, які найраше будуть задовольняти умови відбору, у методі застосовується технологія бустингу;
- перелік ознак, яким буде встановлено співпадіння, буде завершувати свій шлях на input-ах до класифікатора, який визначатиме, чи приймати даний перелік ознак, чи відкидати його;
- для швидкого відбору вікон, які можуть бути відкинуті з метою оптимізації швидкодії алгоритму, застосовуються каскади ознак.

Інтегральна версія зображення в даному алгоритмі дозволяє швидко виконувати обчислення суми інтенсивностей геометричної фігури, яка задається умовами алгоритму, найчастіше така фігура є прямокутником. Важливою властивістю є те, що незалежно від розмірів даної фігури, час витрат обчислень незмінний.

Ця версія зображення містить у своєму складі матрицю, загальні розміри якої дорівнюють масштабу вхідного зображення. Головною особливістю кожного елемента даної матриці вважається сума усіх показників інтенсивностей кольоровості пікселів зображення, які розшатовуються зліва та вище даного елемента плюс значення інтенсивності безпосередньо самого елемента.

Принцип роботи преведення до інтегральної версії зображення на практиці, представлено на рисунку 2.13, де наявна матриця вхідного зображення та матриця його інтегрального зображення окремої частини зображення С.

30	24	5	20	8	52
17	2	152	0	77	33
5	18	59	89	0	17
34	15	90	104	20	3
9	13	22	44	55	51
72	201	185	104	35	21

30	24	5	20
17	2	152	0
5	18	59	89
34	15	90	104

с

30	24	$s = 30 + 24 + 17 + 2 = 73$	
17	2		

A

30	24	5	20
17	2	152	0

B

$s = 73 + 5 + 20 + 152 + 0 = 250$

30	24	$s = 73 + 5 + 18 + 34 + 15 = 145$	
17	2		
5	18		
34	15		

D

30	24	5	20
17	2	152	0
5	18	59	89
34	15	90	104

$s = 250 + 72 + 59 + 89 + 90 + 104 = 664$

59	89	$s = 59 + 89 + 90 + 104 = 342$	
90	104		

искомый объект

$A + C - B - D = 73 + 664 - 250 - 145 = 342$

Рисунок 2.13 – Приклад преведення зображення до інтегральної версії

Завдяки отриманню інтегральної матриці, задача будь-якого знаходження прямокутника з будь-якою площею, значно спрощується, що має вплив на швидкодню алгоритму пошуку.

В розглянутому алгоритмі, мають застосування з метою здобуття інформації, що стосується різниці значень інтенсивності пікселів зображення, ознаки Хаара (Рисунок 2.14), які обчислюються за наступною формулою:

$$F = X - Y$$

при цьому  $X$  – характеризує суму показників інтенсивності з білою ознакою, а  $Y$  – характеризує суму показників інтенсивності з чорною ознакою.

Наступним кроком відбувається накладання ознак Хаара на отримане раніше інтегральне представлення зображення. В розширеному методі, який використовує бібліотека OpenCV використовуються додаткові ознаки (Рисунок 2.14).

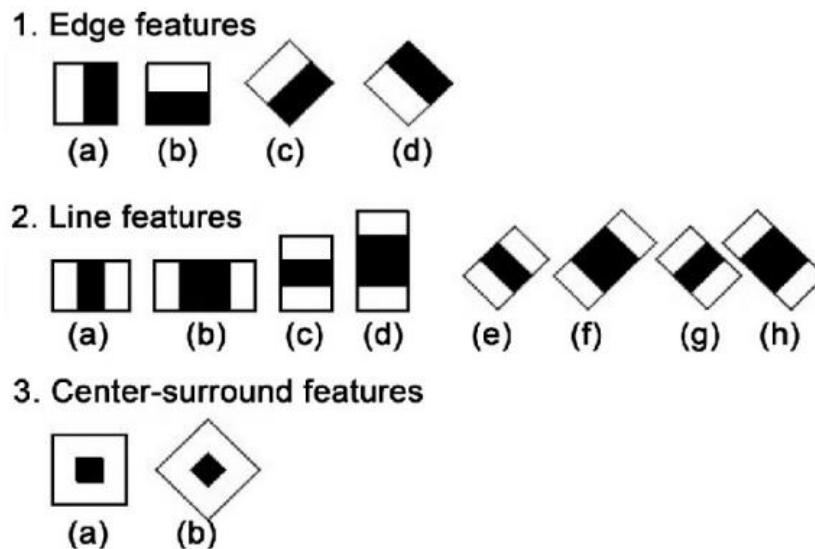


Рисунок 2.14 – Список можливих варіантів класифікацій ознак Хаара

У алгоритмі Віюлі-Джонса в основу закладено метод вікна, яке є сканером, і серед його принципу роботи можна виділити наступні основні етапи:

- насамперед спочатку необхідно обрати зображення, для його подальшого опрацювання, визначити розмір вікна, яке відповідає за процес сканування зображення, та встановити тип класифікації ознак Хаара, які в подальшому будуть накладатися на область сканування з метою визначення об'єктів на ньому. Саме зображення перетворюється у інтегральне представлення;

- аналізуюча область у вигляді скануючого вікна, виконує рух по зображенню із заданим у алгоритмі кроком, наприклад, у одне вікно, розмірами 24x24 пікселі;

- на кожному етапі пересування, у скануючому вікні відбувається процес обчислення відповідно обраних ознак Хаара. Важливо, що під час сканування, позиціонування ознак та їх розмір є змінними величинами;

- кроки 2 і 3 мають повторення застосовуючи різні розміри скануючого вікна;

- після визначення усіх присутніх ознак на зображенні, усі їх значення потрапляють до управляючого рішеннями модулю – класифікатора, котрий прийматиме рішення, чи присутнє у скануючому вікні зображення обличчя[20]. Принцип роботи сканування вікна виглядає приблизно так як показано на рисунку 2.15:

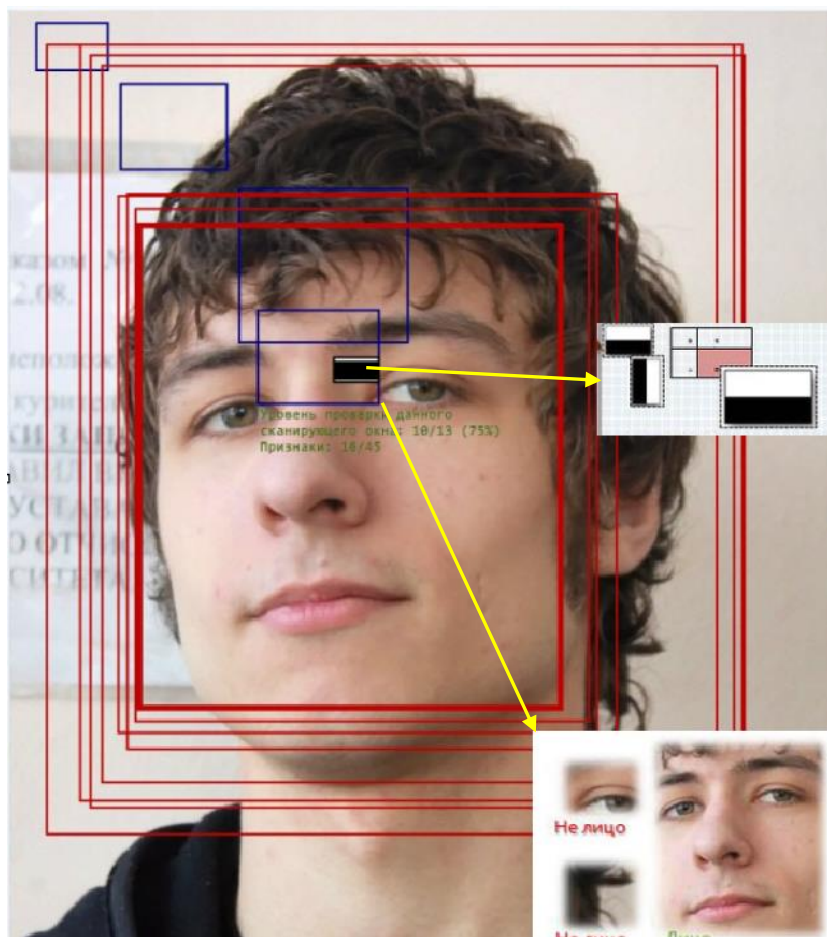


Рисунок 2.16 – Алгоритм сканування вікна з ознаками

Процес використання ознак Хаара до скануючих вікон, які можуть бути різного розміру, вимагає доволі великих апаратних потреб для реалізації обчислень. З метою зменшення апаратних витрат, метод Віоли-Джонса реалізує процес власного навчання, для більш швидкої класифікації ознак, котрі слід застосувати для отримання задовільних результатів.

У цьому випадку метод повинен ідентифікувати конкретні закономірності, за допомогою яких можна розлічити, чи наявне обличчя на вхідному фото, чи воно відсутнє.

Класифікатори, які відповідають за визначення наявності ознак, розділяються на категорію слабких – класифікатори даної категорії можуть робити помилки з великою долею ймовірності і категорію сильних – дані класифікатори фактично можуть визначати приналежність об'єкта до певного класу без ймовірності помилки.

Метод бустингу – функція, яка складає систему класифікаторів, у котрій кожний послідовний класифікатор має призначення виправляти помилки у вигляді поєднання у послідовність усіх його попередніх класифікаторів.

Головне призначення бустингу полягає у тому, щоб з'єднати увесь перелік слабких класифікаторів у єдиний сильний класифікатор. У процесі навчання метод використовує розроблений алгоритм бустингу під назвою Ada Boost. Даний алгоритм здатен поєднувати невизначену кількість класифікаторів за формулою:

$$F(x) = w_1 f_1(x) + w_2 f_2(x) + w_3 f_3(x) + \dots$$

при цьому  $w_1, w_2, w_3$  – це коефіцієнти,  $f_1(x), f_2(x), f_3(x)$  – є переліком ознак, що водночас містять у собі функцію слабого класифікатора. Результат підсумовування слабких класифікаторів за формулою  $F(x)$  – містить у своєму складі потужний, безпомилковий класифікатор.

Каскад – поняття, що характеризує послідовно зв'язані класифікатори.

Явище використання каскадування класифікаторів прискорює процес виявлення обличчя та значно підвищує показники точності розпізнання обличчя на зображенні. Для прикладу, на рисунку 2.16 є класифікатор F1.

На рисунку 2.16 зображено приклад каскадування за алгоритмом Ada Boost, і у випадку, якщо класифікатор  $F_1$  не відпрацьовує успішно операцію знаходження обличчя, то наступний класифікатор  $F_2$  не буде реалізовувати перевірку на наявність обличчя[21].

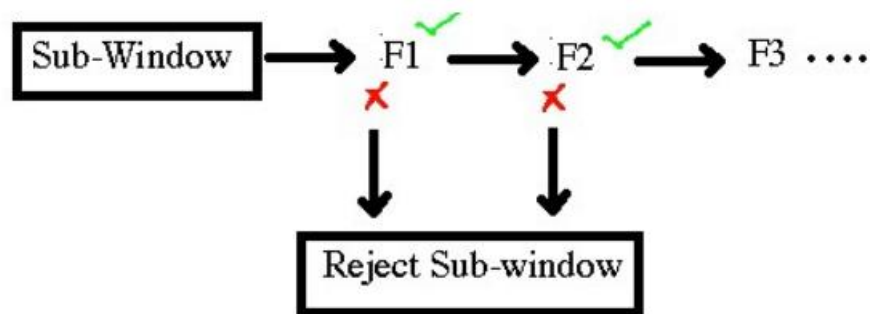


Рисунок 2.16 – Каскад класифікаторів

### 2.3 Глибоке навчання

Глибоке навчання представляє собою метод машинного навчання, який навчає комп'ютери виконувати завдання, що є звичайними для свідомості людини – навчатися з використанням наочних прикладів. Ця технологія є ключовою складовою для керування процесами автономного водіння, без участі людини, як водія, де дана система дозволяє визначати і розрізняти дорожні знаки та відрізняти пішоходів від ліхтарних стовпів. Також глибоке навчання широко використовується для реалізації голосового керування в побутових пристроях нашого життя, на прикладі телефонів, планшетів, комп'ютерів, систем розумного будинку, телевізорів та аудіосистем.

Ключове завдання процесу глибокого навчання полягає в тому, що комп'ютерна модель самостійно навчається вирішувати класифікаційні завдання, аналізуючи дані у формі зображень, відеоматеріалів, текстовому форматі або звуку.

Глибокі навчальні моделі відрізняються високим рівнем точності, іноді перевершуючи показники, які може встановити людина.

Такі моделі для навчання базуються на навчанні нейронних мереж з багат шаровою структурою.

Нові успішні реалізації систем в галузі глибокому навчанні забезпечують вищу точність розпізнавання, ніж та, яку може досягти людина, особливо при класифікації об'єктів на зображеннях.

Проаналізувавши сучасні реалізації технологій глибокого навчання, можна виділити наступні причини, чому тільки зараз глибоке навчання набуло популярності у своєму використанні:

- алгоритм моделі навчання потребує величезних обсягів даних, які зовсім нещодавно почали формуватися у мережі. Мова йде про мільйони зображень та велику кількість годин відеоматеріалу;

- безпосередньо сам алгоритм навчання потребує серйозних апаратних потреб для проведення і забезпечення відповідної обчислювальної потужності. З появою серверних кластерів та можливості проводити обчислення у хмарі, це значно спростило вирішення задачі використання глибокого навчання у своїх комерційних продуктах та скоротило час, витратний на обчислення.

Використання глибокого навчання для аналізу та класифікації зображень, зокрема для процесу виявлення та розпізнавання облич на зображенні, продемонструвало вагомо вищі результати порівняно з традиційними методами застосування комп'ютерного зору[23].

## 2.5 Порівняння розглянутих методів

З метою конструктивного порівняння методів, описаних у кваліфікаційній роботі, перш за все необхідно розглянути переваги і недоліки кожного із них.

Використання функціоналу головних компонентів у методі Eigenfaces спрощує обчислення за рахунок зменшення розмірності матриць. Це дозволяє здійснювати обчислення з меншими потребами на апаратні ресурси і реалізація даного методу легко впроваджується у програмному вигляді.

Через те, що метод реалізує матрицю інтенсивностей кожного пікселя на зображенні, не потребує додаткових операцій над перетворенням зображення, впровадження цього методу у системи є дуже простим.

До головних недоліків у методі Eigenfaces можна віднести дуже велику чутливість до умов освітлення на зображенні, показник його масштабу та кут повороту обличчя.

Найбільш точніші показники визначення облич на зображенні алгоритм Eigenfaces демонстрував на фотознімках з однаковим ступенем освітленості, гарною експозицією та сталим фоном. Через те, що зображення піддаються обробці у цілісному форматі, мінливий задній фон на фото здатен значно погіршувати результат розпізнавання людського обличчя.

Більш досконалий метод Fisherfaces можна характеризувати, як менш вразливий до зміни умов освітлення, куту, під яким знаходиться обличчя, та до змін емоцій на обличчі, що можливо через впровадження до його складу лінійного дискримінантного аналізу. Але, однаковими можна вважати як переваги, так і недоліки обох цих алгоритмів у загальному плані.

Метод LBPН у своїй реалізації використовує аналіз окремих сегментів вхідного зображення для розпізнавання обличчя на цьому зображенні. Це можливо за рахунок пошуку різниці показників інтенсивностей пікселів усього зображення у чорно-білому діапазоні. Дана технологія дозволяє не зменшувати відсоток визначення обличчя на зображенні через зміну умов освітленості. Але така технічна особливість робить цей метод чутливим до наявності сторонніх шумів на фото, що дуже погіршить відсоток визначення обличчя.

Алгоритм Віюли-Джонса Джонса ефективно та швидко справляється з обробкою зображення завдяки використанню каскадів Хаара та алгоритму бустингу навіть зараз, за умов конкуренції серед існуючих рішень на ринку. Вагомою перевагою даного алгоритму є здатність виявлення обличчя навіть у тих частинах фотографії, де воно може бути неочікуваним.

Однак точність визначення облич на вхідному зображенні, які нахилені під кутом більшим за  $30^\circ$ , суттєво менша порівняно з обличчями, які спрямовані прямо в об'єктив камери.

У сучасності, технології глибокого навчання у вигляді згорткових нейронних мереж є широкозатребуваними, і це не випадково, оскільки їхня точність у виявленні та ідентифікації обличчя на зображенні є значною мірою вища, ніж у вищевикладених у кваліфікаційній роботі методах.

Однак для досягнення великої точності модель згорткової нейронної мережі потребує глибокого навчання на великому наборі вхідних даних. Це обумовлює значних апаратних потреб на обчислення і великої кількості часу, витраченого на навчання.

Беручи до уваги усі перелічені недоліки і переваги методів, витрати часу на розробку системи, показник точності обчислень, кількість вхідних зображень, необхідних для коректного навчання системи, коефіцієнт швидкодії системи, проміжок часу, що було спожито під час навчання моделі, можна дійти висновку, що метод Віоли-Джонса та алгоритм LBPН, не беручи до уваги технологію глибокого навчання згорткових мереж, мають найбільший коефіцієнт ефективності в залежності від витрат часу на реалізацію, навчання та обчислення.

## 3 ПРИСКОРЕННЯ МЕТОДУ ВИЗНАЧЕННЯ ОБЛИЧ

### 3.1 Варіанти прискорення

Обраний для дослідження та розглядання алгоритм Віоли-Джонса характеризується великою кількістю позитивних показників, оскільки метод має можливість ідентифікувати обличчя людини на зображенні в будь-якому місці на зображенні. Явище розпізнавання споживає доволі мало часу, але поступається за швидкістю роботи каскадному класифікатору LPB. Визначені каскади ознак Хаара, що лежать у архітектурі побудови алгоритму, мають можливість бути навчені на знаходження обличчя людини безпосередньо в профіль. Зі збільшенням кута повороту більш ніж на 30 градусів, метод стає неефективним у точності визначень під час його використання. Разом із вищевикладеною інформацією, показник точності алгоритму на пряму обумовлюється величиною кількості обраних вхідних зображень. Чим більше зображень на одну людину зібрано для навчання класифікатора, тим точнішими будуть результати визначення.

Основний метод для вдосконалення алгоритму Віоли-Джонса, який я обрав для розгляду у кваліфікаційній роботі – прискорення його швидкості роботи.

Алгоритм Віоли-Джонса не містить в собі жорстко фіксований алгоритм роботи. Його каскади формуються у результат навчання. Через це, швидкість роботи може дуже сильно змінюватися в залежності від складності досліджуваного об'єкту, чи його розміру з метою його знаходження. Найбільш простіший спосіб прискорення часу роботи – власноруч визначити обмеження на мінімальний об'єкт для його знаходження або збільшити крок. Однак ці методи можуть мати відповідні побічні ефекти у вигляді зменшення точності обчислень та у вигляді збільшення помилок під час знаходження об'єкта.

Перенавчання каскадів даного методу може мати значний вплив на показники точності та швидкості роботи.

Проте перенавчання каскадів – дуже витратний процес, який завжди залежить від обсягу навчальної вибірки.

Одним із простих способів прискорення роботи алгоритму є обробка різноманітних частин зображення у паралельних потоках.

Даний метод вже реалізовано у бібліотеці OpenCV, тому нічого у даному випадку робити не доведеться.

Також для значного прискорення роботи можна використовувати можна використовувати графічний прискорювач – GPU. Така реалізація також наявна у OpenCV, який використовується у роботі. Головним недоліком можна назвати умову наявності даного GPU.

До одного із цікавих способів оптимізації методу можна віднести використання векторних інструкцій процесору SIMD. Векторні інструкції наявні у всіх сучасних процесорах, проте можуть бути різними за типом, тому і реалізації мають різні.

### 3.2 Базова версія алгоритму SIMD

Якщо розглянути базову реалізацію даного алгоритму у бібліотеці OpenCV, то можна помітити спроби його прискорення за допомогою різноманітних SIMD інструкцій. Однак судячи зі змісту даних файлів, можна зробити висновок, що прискорення є доволі малим значенням.

Лістинг 3.1 є змістом файлу, що містить скалярну версію алгоритму каскадного Хаар-класифікатору для певного масштабу.

Лістинг 3.1 – Файл зі скалярною версією алгоритму

```
void HaarDetect(const HaarCascade & hid, const Image & mask, const Rect &
rect, int step, Image & dst)
{
    for (ptrdiff_t row = rect.top; row < rect.bottom; row += step)
    {
        size_t p_offset = row * hid.sum.stride / sizeof(uint32_t);
```

### Продовження лістингу 3.1

```

size_t pq_offset = row * hid.sqsum.stride / sizeof(uint32_t);
for (ptrdiff_t col = rect.left; col < rect.right; col += step)
{ if (mask.At<uint32_t>(col, row) == 0)
  continue;
  float norm = Norm(hid, pq_offset + col);
if (Detect(hid, p_offset + col, 0, norm) > 0)
  dst.At<uint32_t>(col, row) = 1; }}}

```

Далі у змісті лістингу 3.2 відбувається сканування зображення за усіма його точками.

Зазвичай при малому масштабі зображення для прискорення роботи використовують сканування зображення із кроком 2. Для кожної точки відбувається розрахування нормувальний коефіцієнт, значення якого залежить від яскравості та контрастності зображення у околиці даної точки.

### Лістинг 3.2 – Файл з алгоритмом сканування точок зображення

```

uint32_t Sum(uint32_t * const ptr[4], size_t offset)
{
  return ptr[0][offset] - ptr[1][offset] - ptr[2][offset] +
  ptr[3][offset];
}

float Norm(const HaarCascade & hid, size_t offset)
{
  float sum = float(Sum(hid.p, offset));
  float sqsum = float(Sum(hid.pq, offset));
  float q = sqsum*hid.windowArea - sum *sum;
  return q < 0.0f ? 1.0f : sqrtf(q);
}

```

А також у Лістингу 3.3 міститься безпосередньо сам алгоритм класифікації зображення.

### Лістинг 3.3 – Файл з алгоритмом класифікації зображення

```

float WeightedSum(const WeightedRect & rect, size_t offset)
{int sum = rect.p0[offset] - rect.p1[offset] - rect.p2[offset] +
rect.p3[offset];
  return rect.weight*sum; }
int Detect(const HaarCascade & hid, size_t offset, int startStage, float
norm)
{const HaarCascade::Stage * stages = hid.stages.data();
  const HaarCascade::Node * node = hid.nodes.data() +
stages[startStage].first;
  const float * leaves = hid.leaves.data() + stages[startStage].first * 2;
  for (int i = startStage, n = (int)hid.stages.size(); i < n; ++i)
  {

```

### Продовження лістингу 3.3

```

const HaarCascade::Stage & stage = stages[i];
const HaarCascade::Node * end = node + stage.ntrees;
float stageSum = 0.0;
for (; node < end; ++node, leaves += 2)
{
    const HaarCascade::Feature & feature = hid.features[node->featureIdx];
    float sum = WeightedSum(feature.rect[0], offset) +
        WeightedSum(feature.rect[1], offset);
    if (feature.rect[2].p0)
        sum += WeightedSum(feature.rect[2], offset);
    stageSum += leaves[sum >= node->threshold*norm];
} if (stageSum < stage.threshold)
    return -i;
return 1; }

```

У даній версії алгоритму класифікації зображення застосовується розрахування ознак Хаара за допомогою попередньо розрахованих інтегральних зображень.

### 3.3 Варіанти векторизації алгоритму

Проведений аналіз принципу роботи даного алгоритму вказує, що основні розрахункові ресурси витрачаються на обчислення інтегральних сум. Тому саме їх потрібно вектаризувати.

Одним із варіантів рішення є векторизація за обчисленням Хаар-ознак, як саме і зроблено у OpenCV, проте проблема полягає у тому, що дані, на якими проводяться обчислення хаотично розкидані у пам'яті. Тому спочатку їх потрібно збирати із пам'яті за допомогою скалярних операцій читання, а потім отримані результати обчислень зберігати завдяки скалярним операціям запису даних.

Наступним і останнім із можливих рішень є векторизація за точками. Дані у інтегральному зображенні для сусідніх точок лежатимуть поруч, що дозволить швидко та ефективно завантажувати їх за допомогою векторних операцій.

Крім цього, даний метод легко масштабується для різної довжини вектору SIMD.

Але у даного методу наявні наступні проблеми роботи:

- ефективність алгоритму Віюлі-Джонса базується на тому, що на кожній стадії роботи відкидаються неправильні точки. Відсоток відкидаємих точок залежить від параметрів навчання, але за більшістю випадків, становить 50%. Якщо ми обчислюємо SIMD-вектор з декількома точками, то ми будемо змушені продовжувати розрахунок стадій доти, доки не будуть відкинуті усі його елементи;

- оригінальний алгоритм із бібліотеки OpenCV у меті оптимізації для невеликого масштабу зображення перевіряє точки із кроком у 2. Це пришвидшує скалярні дані, але знижує ефективність векторних інструкцій, адже нам вже на першій стадії доведеться половину обчислень виконувати дарма.

У лістингу 3.4 розглянемо випадок, коли будуть скануватися усі точки зображення без виключень.

#### Лістинг 3.4 – Файл з алгоритмом сканування без виключень

```
void DetectionHaarDetect32fp(const HidHaarCascade & hid, const Image &
mask, const Rect & rect, Image & dst)
{
    size_t rightAligned = rect.left + Simd::AlignLo(rect.Width(), 4);
    for (ptrdiff_t row = rect.top; row < rect.bottom; row += 1)
    {
        size_t p_offset = row * hid.sum.stride / sizeof(uint32_t);
        size_t pq_offset = row * hid.sqsum.stride / sizeof(uint32_t);
        size_t col = rect.left;
        for (; col < rightAligned; col += 4)
        {
            __m128i result = _mm_loadu_si128((__m128i*)(mask.Row<uint32_t>(row)
+ col));
            if (_mm_testz_si128(result, _mm_set1_epi32(1)))
                continue;
            __m128 norm = Norm32fp(hid, pq_offset + col);
            Detect32f(hid, p_offset + col, norm, result);
            _mm_storeu_si128((__m128i*)(dst.Row<uint32_t>(row) + col), result);
        }
        for (; col < rect.right; col += 1)
        {
            if (mask.At<uint32_t>(col, row) == 0)
                continue;
            float norm = Norm32f(hid, pq_offset + col);
            if (Detect(hid, p_offset + col, 0, norm) > 0)
                dst.At<uint32_t>(col, row) = 1; }}}}
```

Далі у лістингу 3.5 запропоновано варіант проведення сканування зображення у блочному вигляді по 4 точки, де лише крайні точки будуть обчислюватися скалярним методом:

Лістинг 3.5 – Файл з блочним алгоритмом сканування

```

inline __m128 ValidSqrt(__m128 value)
{
    return    __mm_blendv_ps(__mm_set1_ps(1.0f),    __mm_sqrt_ps(value),
__mm_cmpgt_ps(value, __mm_set1_ps(0.0f))); }
inline __m128i Sum32ip(uint32_t * const ptr[4], size_t offset)
{
    __m128i s0 = __mm_loadu_si128((__m128i*)(ptr[0] + offset));
    __m128i s1 = __mm_loadu_si128((__m128i*)(ptr[1] + offset));
    __m128i s2 = __mm_loadu_si128((__m128i*)(ptr[2] + offset));
    __m128i s3 = __mm_loadu_si128((__m128i*)(ptr[3] + offset));
    return __mm_sub_epi32(__mm_sub_epi32(s0, s1), __mm_sub_epi32(s2, s3));
}
inline __m128 Norm32fp(const HidHaarCascade & hid, size_t offset)
{
    __m128 area = __mm_set1_ps(hid.windowArea);
    __m128 sum = __mm_cvtepi32_ps(Sum32ip(hid.p, offset));
    __m128 sqsum = __mm_cvtepi32_ps(Sum32ip(hid.pq, offset));
    return ValidSqrt(__mm_sub_ps(__mm_mul_ps(sqsum, area), __mm_mul_ps(sum,
sum))); }

```

### 3.3 Scalar, SSE 4.1, AVX2, AVX512

У процесорах та комп'ютерних системах термін *Scalar* використовується для опису форми обробки даних. Загалом, існують два основних типи обробки даних: скалярна та векторна.

До скалярної обробки відносяться:

- одинична операція: процесор виконує лише одну операцію за один такт;
- однократний об'єм даних: обробка виконується на одному елементі даних за раз.

Простий приклад скалярної операції - додавання двох чисел. Якщо ви маєте дві змінні, ви додаєте їх одну до одної за один такт.

Що стосується векторної обробки, то до неї можна віднести:

- одна операція, багато даних: процесор може виконувати операцію на багатьох елементах даних одночасно;
- розпаралелюється: операції виконуються паралельно на різних даних.

Наприклад, векторна операція може виконувати додавання не однієї пари чисел, але декількох пар одночасно.

SSE 4.1 це новий набір команд Intel Core архітектури, містить у своєму складі 47 інструкцій, деякі з яких спрощують векторизацію програм.

Інструкції типу AVX є подальшим розвитком інструкцій типу SSE. Головним призначенням даного типу інструкцій є прискорення роботи програмного забезпечення, що працює з фото, відео та 3Д-моделями.

SSE використовує 128 бітні регістри, AVX2 використовує 256 біт, а останній у лінійці AVX512 – виходячи із назви використовує 512 біт.

Слід зауважити, що при використанні нових інструкцій і так не холодні нові процесори типу Alder Lake споживають набагато більше енергії і сильніше розігріваються - температури під навантаженням збільшуються в середньому на 10-15 градусів. При такому використанні прийнятні результати охолодження починають показувати лише дорогі кулери та водяні СО. Те саме було і в попередньому поколінні типу Rocket Lake при залученні AVX-512.

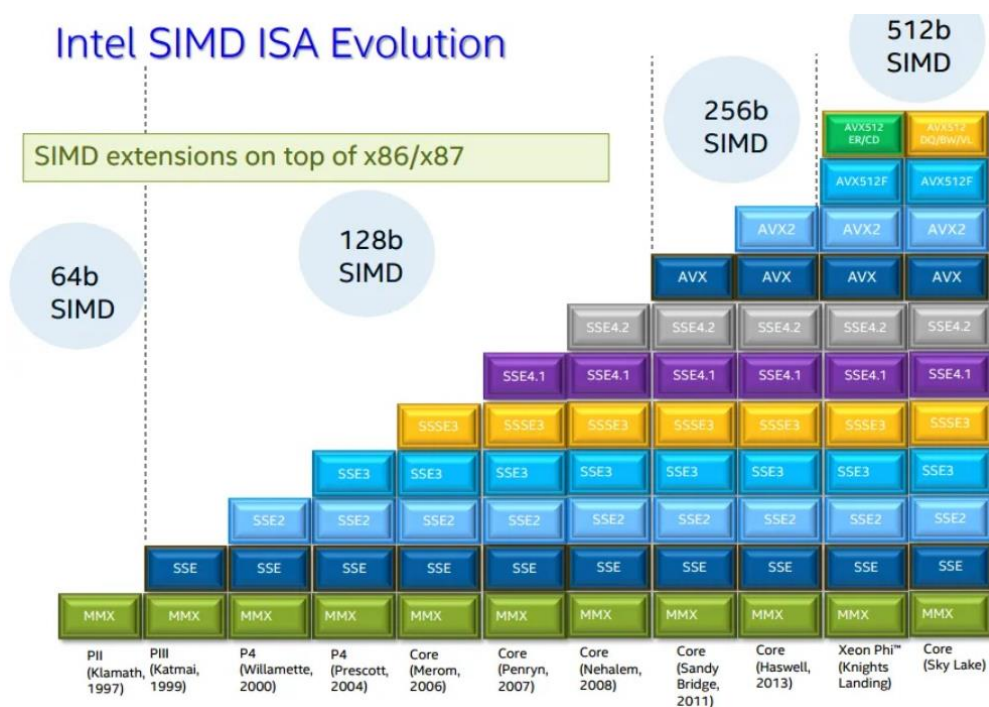


Рисунок 3.1 – Розрядність інструкцій процесорів типу Intel

Також, компанія Intel вважає наявність 512-бітної версії AVX на даний момент необов'язковим в масових платформах через їх малу поширеність в програмному забезпеченні користувача, на відміну від серверного.

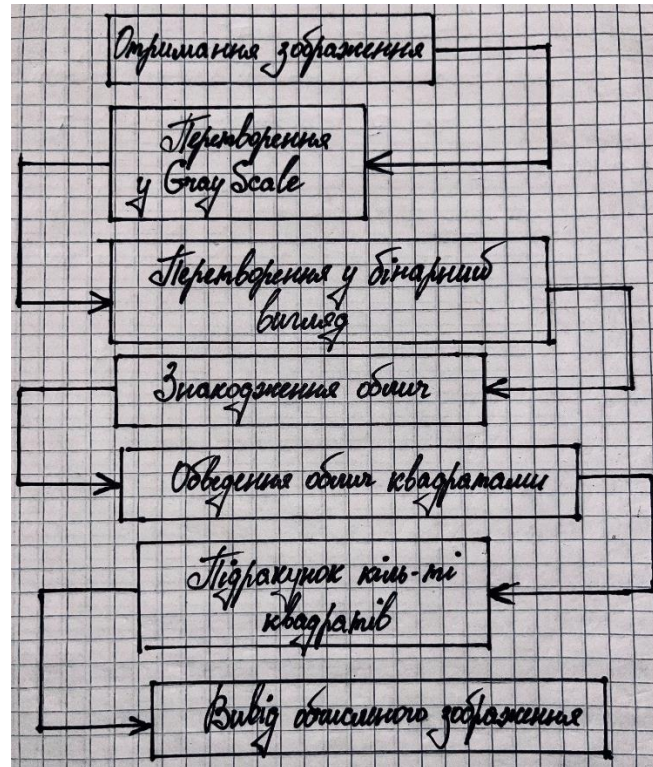


Рисунок 3.2 – Запропонована система для визначення обличчя на зображенні

Користувач самостійно обирає фото, на якому він бажає виділити обличчя на фото та підрахувати їх кількість, завантажує його, після успішної загрузки зображення до додатку, відбувається переведення вхідного зображення у GreyScale з маркуванням контурів (Рисунок 3.3).

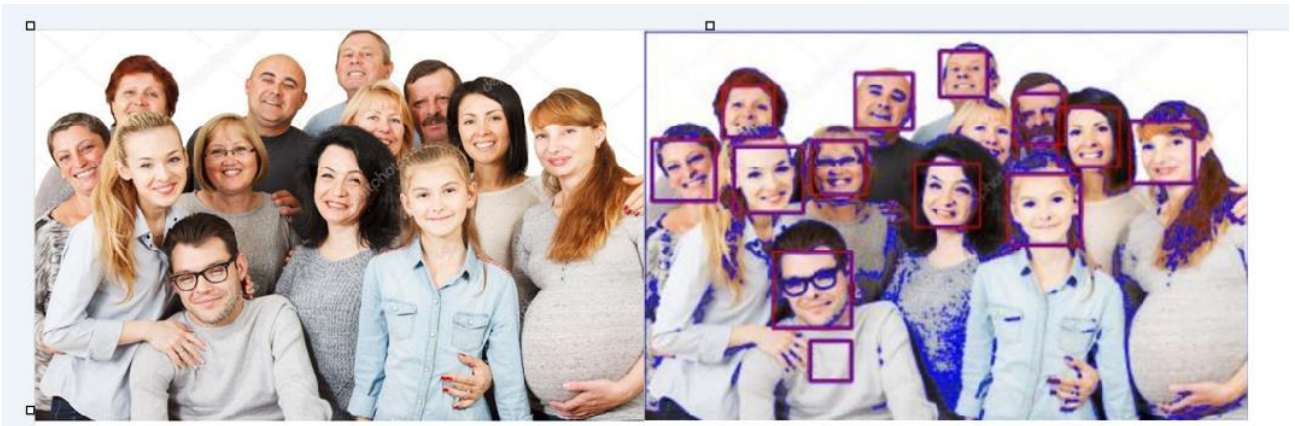


Рисунок 3.3 – Вхідне зображення та зображення переведене у GreyScale з маркуванням контурів обличчя людини

Наступним кроком після отримання зображення у GreyScale, є переведення зображення у бінарний формат також з маркуванням контурів обличчя людини. Після чого відбувається фінальне виділення контурів обличчя, розпізнавання облич за ознаками та підрахунок кількості отриманих областей із обличчями на (Рисунок 3.4).

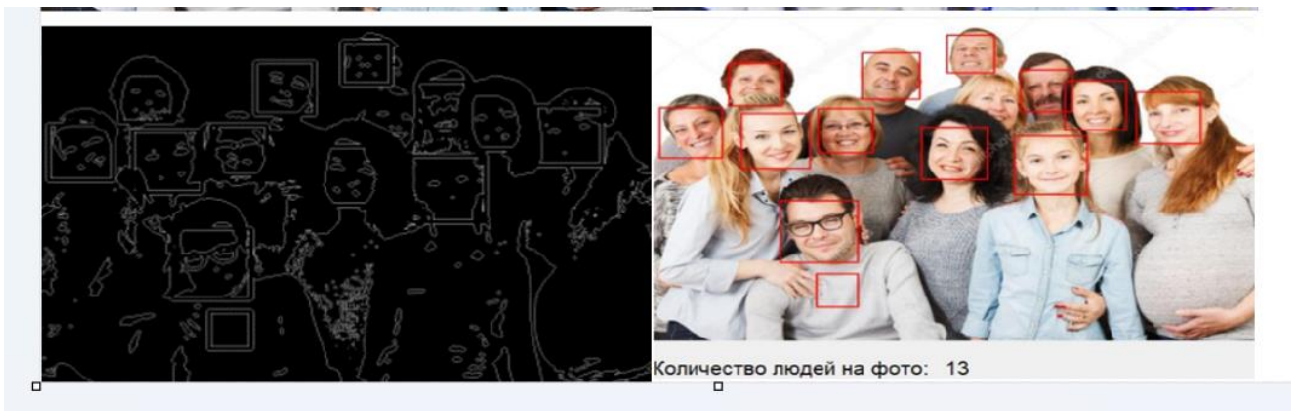


Рисунок 3.4 – Бінарне зображення та отримане в результаті роботи системи зображення із маркуванням знайдених людських облич

## 4 ДОСЛІДЖЕННЯ МЕТОДУ ВИЗНАЧЕННЯ ОБЛИЧ

### 4.1 Розробка програмного забезпечення

Мінімальні технічні вимоги до апаратного забезпечення користувача системи:

- операційна система: Windows 7 та вище;
- оперативна пам'ять: 4 GB;
- процесор: Intel Core i3, 3.06 GHz;
- місце на диску: 1.5 GB.

Для оптимального та продуктивного користування системою необхідна наявність графічного прискорювача (GPU).

### 4.2 Тестування методу

Тестування проводилося на наступних технічних параметрах апаратного забезпечення комп'ютера:

- операційна система: Windows 11;
- оперативна пам'ять: 16 GB;
- процесор: Intel(R) Core(TM) i5-6440HQ, 2.60 GHz із підтримуваним набором інструкцій AVX2, дані про які взято з офіційної документації виробника Intel;
- графічний прискорювач: Intel(R) HD Graphics 530.

#### 4.2.1 Підготовка вхідних даних

Для оцінки точності обчислень було проведено практичний експеримент, за допомогою навчального та практичного датасетів, які містять вибірку різноманітних фотографій, на зображенні яких міститься обличчя людини.

Було проведено два етапи робіт. Під час виконання першого, було реалізовано підготовку вхідних даних для навчання алгоритму.

Підготовка даних для випробування та тестування програми є важливою складовою успіху проекту. Для виконання проекту, було вирішено обрати досить великий «датасет» - колекцію із логічних записів даних, що зберігається у вигляді кортежу. Головна різниця набору даних від файлу заключається у тому, що набір даних водночас є каталогом і файлом файлової системи.

Стосовно набору даних для проекту, обраний «датасет» містить у своєму складі фотографії, на яких зображена невелика група людей. До фотографій додається файл у форматі .txt, з визначеною кількістю людей, інформацією щодо освітленості, місця проведення знімку і додаткових умов фотознімку.

Також, набір даних складається з трьох папок («test»,«train»,«val»), кожна з яких містить більше 2 тисяч фотознімків з різних ракурсів і файл з додатковою інформацією стосовно кожної фотографії.

Щодо змісту «датасету», важливими папками з фотографіями є папки «train» та «val». З папки «train» було проведено першу вибірку даних, для тестування роботи нашої програми. У папці «val» містяться фотографії, які перевіряють кінцеву версію проекту на справну роботу.

Однією з головних складових точності підрахунку людей на фотознімках, має бути добре зображено обличчя людини і гарна або задовільна якість фотографії і добре сфокусоване зображення. Це обумовлено масштабуванням фотознімку, під час алгоритму підрахунку.

Наступною ваговою складовою є задовільна освітленість на фотознімку. Фотознімки вночі, вдень з відблисками сонця, на концертах зі спецефектами, у приміщеннях з великою кількістю відблисків ускладнюють і іноді унеможливають коректний підрахунок людей

Ще слід звернути увагу на погодний стан фотознімку. Під час дощу або снігу, туману або піднятого пилу у повітря, тощо, це все унеможливає коректний підрахунок людей на фотографії.

Другорядною але теж важливою складовою є кількість сторонніх об'єктів на фото. Це можуть бути прапори, машини, рекламні банери, велика кількість піднятих рук, наявність тварин або птахів у кадрі, тощо. Це дуже ускладнює пошук і точність підрахунку. Щоб уникнути такого великого списку нюансів, під час виконання завдання, було проведено ручний відбір фотознімків, спираючись на умови точності підрахунку.



Рисунок 4.1 – Фотознімок, що не задовольняє умовам відбору



Рисунок 4.2 – Фотознімок, що також не задовольняє умовам відбору

Рисунок 4.1 та рисунок 4.2 не відповідають умовам відбору фотознімків, адже під час його аналізу можна знайти порушення наступних пунктів відбору:

- незадовільне освітлення;
- погана фокусування кадру;
- погані погодні умови;
- наявність відблисків та сторонніх об'єктів на фото.



Рисунок 4.3 – Фотознімок, що задовольняє умовам відбору

Рисунок 4.3 задовольняє умови відбору, відповідаючи головним вимогам відбору а саме:

- гарний фокус, невелика відстань до об'єктів;
- задовільне освітлення та погодні умови;
- немає сторонніх об'єктів на фото та відсутні відблиски.

Отже, виходячи із наведених прикладів аналізу зображень для навчання, можна дійти висновку, що процес відбору фотографій за відповідними умовами займає досить багато часу та праці, адже у використовуваному «датасеті» міститься величезна кількість різноманітних фотографій, які слід усі переглянути і власноруч відібрати. Водночас, результат відбору фотознімків демонструє задовільні результати роботи алгоритму.

#### 4.2.2 Результати тестування

Для тестування ефективності роботи алгоритму вхідні зображення були попередньо відібрані, оброблені і з них було взято інформацію з відповідних файлів у вигляді порохованих дескрипторів, про які було викладено вище.



Рисунок 4.4 – Коректний результат роботи алгоритму виявлення облич



Рисунок 4.5 – Неточний результат роботи алгоритму виявлення облич

Після проведених тестів, було проаналізовано отримані результати точності обчислень та встановлено, що загальна точність роботи системи складає для методу виявлення облич – 86.80%;

Також було розглянуто ефективність виконання алгоритму в залежності від використання різних наборів інструментів процесора.

Таблиця 4.1 – Результати витрат часу на сканування зображення розміром 1920x1080. Дані витрат у часі наведено у ms

Функція	Scalar(базовий)	SSE-4.1	AVX2	AVX-512
Середнє геометричне	227.4	111.5	74.95	54.6
DetectionHAAR(0)step2	84.7	45.7	32.7	25.9
DetectionHAAR(1)step2	162.7	79.1	50.9	36.8
DetectionLBP(0)step1	329.9	159.3	109.7	80.8
DetectionLBP(1)step1	588.2	268.2	172.3	114.7

DetectionHAAR(0)step2 – функція, що сканує зображення із кроком 2;

DetectionLBP(0)step1 – функція, що сканує зображення із кроком 1;

(0)– використання каскаду HAAR;

(1)– використання каскаду LBP.

Таблиця 4.2 – Отримані дані в результаті тестів у вигляді коефіцієнту прискорення нових варіантів рішення відносно базової швидкості версії Scalar. Дані витрат у часі наведено у вигляді коефіцієнту відношення

Функція	SSE4.1/Scalar	AVX2/ Scalar	AVX/ Scalar
Середнє геометричне	2.04	3.04	4.17
DetectionHAAR(0)step2	1.85	2.59	3.27
DetectionHAAR(1)step2	2.06	3.19	4.42
DetectionLBP(0)step1	2.07	3.01	4.08
DetectionLBP(1)step1	2.19	3.41	5.13

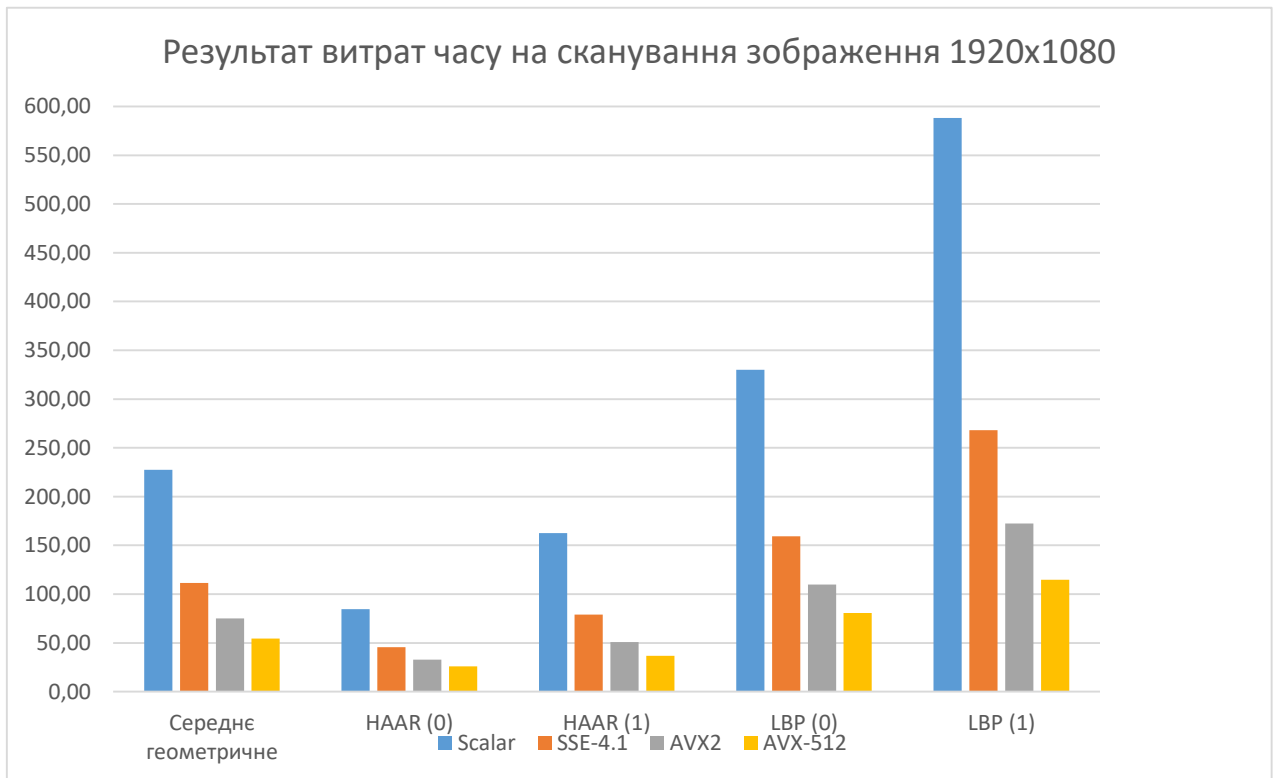


Рисунок 4.7 – Діаграма результатів витрат часу на сканування зображення

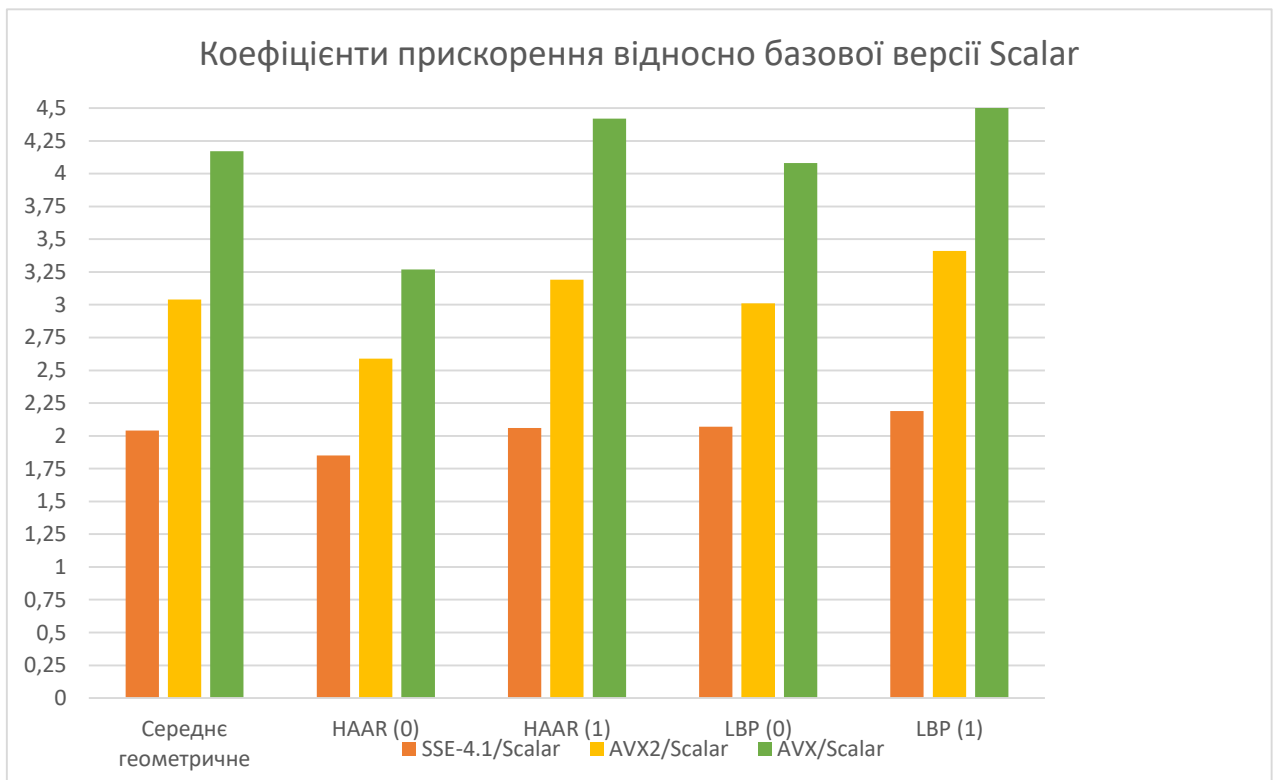


Рисунок 4.8 – Діаграма з результатами коефіцієнтів прискорення

Проаналізувавши отримані результати під час проведення тестування роботи алгоритму за значеннями витрат часу на сканування зображення під час роботи алгоритму та відношенням базового прискорення із прискоренням отриманим для нових варіантів рішень, можна дійти висновку, що із кожним подвоєнням розмірності вектору, корисність роботи на обох каскадах алгоритму збільшується за умови зменшення витрат часу на сканування зображення із сталим масштабом.

## ВИСНОВКИ

В даній кваліфікаційній роботі було проаналізовано сучасні рішення для розв'язання задачі визначення обличчя на зображенні за допомогою методів машинного навчання. Проаналізовано сучасні комерційні продукти на ринку технологій програмного забезпечення для визначення облич, і виділено їх затребуваність та їх алгоритмічні підходи для вирішення задачі.

Розглянуто побудову архітектури розпізнавання облич та класифікацію методів для визначення, виділено їх сильні та слабкі сторони, можливості їх використання та апаратні потреби для коректної роботи.

З наведеного аналізу можна зробити висновок, що всі існуючі методи сучасності на ринку технологій програмного забезпечення мають свої слабкі сторони, велику обчислювальну складність, й вимагають подальшого вдосконалення і розвитку. Наприклад обрані ознаки Хаара, які мають застосування в методі Віюлі-Джонса є дуже простими, але ефективними для фронтального виділення облич на зображенні.

Запропоноване рішення є досить універсальним у використанні, надає можливість використання в різних сферах.

Метою було поліпшення роботи безпосередньо самого методу по розпізнанню облич з різних параметрів перспективи зображення та ще більшу оптимізацію швидкості обчислень під час його роботи. Алгоритм підрахунку облич Віюлі-Джонса дає суттєво вищу точність, але може бути застосований лише для вузького класу зображень.

Новизна кваліфікаційної роботи полягає в створенні більш швидкого методу визначення облич на зображенні на основі алгоритму підрахунку Віюлі-Джонса. Прискорення швидкості роботи є перспективним напрямком у області досліджень показників точності обчислень. Процес може привести до створення більш швидких та точних систем визначення облич на зображенні, враховуючи переваги та виправляючи недоліки вже існуючих методів.

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Yang M.-H. Detecting faces in images: a survey / M.-H. Yang, D.J. Kriegman, N. Ahuja // IEEE Trans. on PAMI. – 2002. – № 24(1). – P. 34–58.
2. Turk M. Eigenfaces for recognition / M. Turk, A. Pentland // Journal of Cognitive Neuroscience. – 1991. – Vol. 13, No. 1. – P. 71–86.
3. Vaswani N. Principal components null space analysis for image and video classification / N. Vaswani, R. Chellappa // IEEE Trans. Image Process. – 2006. – Vol. 15, No. 7. – P. 1816–1830.
4. Suhas S. Face recognition using principal component analysis and linear discriminant analysis on holistic approach in facial images database / S. Suhas, A. Kurhe, Dr.P. Khanale // IOSR Journal of Engineering. – 2012. – Vol. 2, Is. 12. – P. 15-23.
5. Belhumeur P.N. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection / P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman // IEEE Trans. On PAMI. – 1997. – Vol. 19, No. 7. – P. 711–720.
6. Yang M.H. Kernel Eigenfaces vs. Kernel Fisherfaces: face recognition using kernel methods / M.H. Yang // 5th IEEE Int. Conf. on Automatic Face and Gesture Recognition. – 2002. – P. 215-220.
7. Bartlett M.S. Face recognition by independent component analysis / M.S. Bartlett, J.R. Movellan, T.J. Sejnowski // IEEE Trans. Neural Netw. – 2002. – Vol.13, No. 6. – P. 1450–1464.
8. Shen L. A review on Gabor wavelets for face recognition / L. Shen, L. Bai // Journal of Pattern Analysis and Applications. – 2006. – Vol. 9, No. 2-3. – P. 273-292.
9. Imtiaz H. A face recognition scheme using waveletbased dominant features / H. Imtiaz, S.A. Fattah // Signal & Image Processing : An International Journal. – 2011. – Vol.2, No.3. – P. 69-80.

10. Zhao W. Face recognitions literature survey /W. Zhao, R. Chellappa, P.J. Phillips, A. Rosenfeld // ACM Computing Surveys. – 2003. – Vol. 35, No. 4. – P. 399–458.
11. Prince Simon J. D. Computer Vision: Models, Learning, and Inference. [Текст] / Simon J. D. – Видавництво Prince Cambridge University Press, 2012 – 596 с.
12. W. Zhao. <https://dl.acm.org/citation.cfm?id=954342> [Електронний ресурс] / W. Zhao, R. Chellappa, P. J. Phillips – Режим доступу до ресурсу: Face recognition: A literature survey.
13. Raja R. Face Detection Using OpenCV and Python [Electronic resource] / R. Raja. - Access mode: <https://www.superdatascience.com/opencv-face-detection/>
14. Raja R. Face Recognition Using OpenCV and Python [Electronic resource] / R. Raja. - Режим доступу: <https://www.superdatascience.com/opencv-facerecognition/>
15. Machine learning methods for fully automatic recognition of facial expressions and facial actions [Електронний ресурс] / [M. S. Bartlett, G. Littlewort, C. Lainscsek та ін.] – Режим доступу до ресурсу: <https://ieeexplore.ieee.org/abstract/document/1398364>.
16. Z. Zhu. Real-time eye detection and tracking under various light conditions [Електронний ресурс] / Z. Zhu, K. Fujimura, Q. Ji – Режим доступу до ресурсу: <https://dl.acm.org/citation.cfm?id=507100>.
17. Annotated Facial Landmarks in the Wild: A large-scale, real-world database for facial landmark localization [Електронний ресурс] / M.Köstinger, P. Wohlhart, P. M. Roth, H. Bischof – Режим доступу до ресурсу: <https://ieeexplore.ieee.org/abstract/document/6130513>.
18. Turk M. Eigenfaces for Recognition / M. Turk, A. Pentland. // Journal of Cognitive Neuroscience. – 1991. – №3. – С. 71–86.
19. Labeled Faces in the Wild [Електронний ресурс] – Режим доступу до ресурсу: <http://vis-www.cs.umass.edu/lfw/index.html>.

20. ML | Face Recognition Using Eigenfaces (PCA Algorithm) [Електронний ресурс]. – 2020. – Режим доступу до ресурсу: <https://www.geeksforgeeks.org/ml-face-recognition-using-eigenfaces-pcaalgorithm/>.

21. Belhumeur P. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection / P. Belhumeur, J. Hespanha, D. Kriegman. // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1997. – №19. – С. 711–720.

23. Wagner P. Fisherfaces [Електронний ресурс] / Philipp Wagner. – 2012. – Режим доступу до ресурсу: <https://www.bytefish.de/blog/fisherfaces.html>.

24. Rosebrock A. Facial landmarks with dlib, OpenCV, and Python [Електронний ресурс] / Adrian Rosebrock – Режим доступу до ресурсу: <https://www.pyimagesearch.com/2017/04/03/facial-landmarks-dlib-opencv-python/>.

25. Gradient-based learning applied to document recognition / Y. LeCun, L. Bottou, Y. Bengio, P. Haffner. // Proceedings of the IEEE. – 1998. – №86. – С. 2278–2324.

26. An Intuitive Explanation of Convolutional Neural Networks [Електронний ресурс]. – 2016. – Режим доступу до ресурсу: <https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/>.

27. CS231n Convolutional Neural Networks for Visual Recognition [Електронний ресурс] – Режим доступу до ресурсу: <https://cs231n.github.io/convolutional-networks/#fc>.

28. Pedych, D., Методи машинного навчання для визначення обличчя на зображеннях/ Pedych, D. // Міжнародна наукова інтернет-конференція «Proceedings of the IV international scientific and Theoretical Conference». – 2023. – С. 110-115.