

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

Факультет  
Кафедра

Комп'ютерної інженерії та управління  
Комп'ютерних інтелектуальних технологій та систем

## КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти

другий (магістерський)

Система розпізнавання обличчя з використанням згорткових  
нейронних мереж  
(тема)

Виконав:

студент II курсу, групи КІТМ-22-1

Владислав КУРЕНКО

(власне ім'я, прізвище)

Спеціальність 123 Комп'ютерна інженерія

(код і повна назва спеціальності)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма Комп'ютерні  
інтелектуальні технології

(повна назва освітньої програми)

Керівник зав. каф. КІТС Олег РУДЕНКО

(посада, власне ім'я, прізвище)

Допускається до захисту

Зав. кафедри

(підпис)

Олег РУДЕНКО

(власне ім'я, прізвище)

2023 р.

Харківський національний університет радіоелектроніки

|                     |   |
|---------------------|---|
| Факультет           | Комп'ютерної інженерії та управління              |
| Кафедра             | Комп'ютерних інтелектуальних технологій та систем |
| Рівень вищої освіти | другий (магістерський)                            |
| Спеціальність       | 123 Комп'ютерна інженерія                         |
| Тип програми        | освітньо-професійна                               |
| Освітня програма    | Комп'ютерні інтелектуальні технології             |

ЗАТВЕРДЖУЮ:

Зав. кафедри \_\_\_\_\_  
(підпис)

« \_\_\_\_\_ » \_\_\_\_\_ 2024 р.

## ЗАВДАННЯ

### НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові \_\_\_\_\_ Куренку Владиславу Олександровичу  
(прізвище, ім'я, по батькові)

1. Тема роботи Система розпізнавання обличчя з використанням згорткових нейронних мереж

затверджена наказом по університету від “ 03 ” листопада 2023 р. № 1290Ст

2. Термін подання студентом роботи до екзаменаційної комісії 13.01.2024

3. Вхідні дані до роботи (проекту): набір зображень з обличчям

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

1) аналіз предметної області та постановка задачі;

2) огляд засобів вирішення задачі;

3) реалізація рішення;

4) робота системи.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій 15 слайдів

---

---

---

---

---

---

---

---

6. Консультанти розділів роботи и (п.6 включається до завдання за наявності консультантів згідно до наказу, зазначеному 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 проблеми та огляд існуючих рішень     | 06.11.23 – 19.11.23            | Виконано |
| 2 | Вибір засобів розробки                       | 20.11.23 – 26.11.23            | Виконано |
| 3 | Розробка структури системи                   | 27.11.23 – 03.12.23            | Виконано |
| 4 | Реалізація системи                           | 04.12.23 – 17.12.23            | Виконано |
| 5 | Тестування системи                           | 18.12.23 – 31.12.23            | Виконано |
| 6 | Оформлення матеріалів кваліфікаційної роботи | 01.01.24 – 11.01.24            | Виконано |
| 7 | Перевірка кваліфікаційної роботи керівником  | 12.01.24 – 13.01.24            | Виконано |
|   |  |                                |          |
|   |  |                                |          |
|   |  |                                |          |
|   |  |                                |          |
|   |  |                                |          |

Дата видачі завдання 06 листопада 2023 р.

Студент \_\_\_\_\_  
(підпис)

Керівник роботи \_\_\_\_\_ зав. каф. КІТС Олег РУДЕНКО  
(підпис) (посада, ім'я, прізвище)

## РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 71 с., 18 рис., 1 дод., 21 джерел.

КОМП'ЮТЕРНИЙ ЗІР, ВИЯВЛЕННЯ ОБЛИЧЧЯ, РОЗПІЗНАВАННЯ ОБЛИЧЧЯ, ЗГОРТКОВІ НЕЙРОННІ МЕРЕЖІ, TENSORFLOW, KERAS, ІДЕНТИФІКАЦІЯ ОСОБИ, АУТЕНТИФІКАЦІЯ ОСОБИ, ВЕРИФІКАЦІЯ ОСОБИ

Метою кваліфікаційної роботи є реалізація системи розпізнавання обличчя з використанням згорткових нейронних мереж.

У ході виконання кваліфікаційної роботи було проаналізовано методи розпізнавання обличчя та засоби розробки систем комп'ютерного зору. Було оглянуто методи класифікації даних та принцип роботи згорткових нейронних мереж. Було розроблено систему розпізнавання обличчя з використанням згорткових нейронних мереж, яка була успішно протестована.

Об'єктом дослідження є процес обробки зображення в системах комп'ютерного зору.

Предметом дослідження є обробка зображення в системах розпізнавання обличчя.

Розроблене рішення може бути використано в системах контролю доступу та відвідуваності, наприклад у домашніх господарствах чи невеликих підприємствах.

## ABSTRACT

Master's thesis: 71 pages, 18 figures, 1 appendices, 21 sources.

COMPUTER VISION, FACE DETECTION, FACE RECOGNITION, CONVOLUTIONAL NEURAL NETWORKS, TENSORFLOW, KERAS, PERSON IDENTIFICATION, PERSON AUTHENTICATION, PERSON VERIFICATION

The major goal of this thesis is to develop a facial recognition system using convolutional neural networks.

During the qualification work, facial recognition methods and computer vision systems development tools were analyzed. Methods of data classification and the principle of operation of convolutional neural networks were reviewed. A facial recognition system using convolutional neural networks was developed and successfully tested.

The object of research is the process of image processing in computer vision systems.

The subject of research is image processing in facial recognition systems.

The developed solution can be used in access and attendance control systems, for example, in households or small businesses.

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

Факультет  
Кафедра

Комп'ютерної інженерії та управління  
Комп'ютерних інтелектуальних технологій та систем

**АНОТАЦІЯ**  
**КВАЛІФІКАЦІЙНОЇ РОБОТИ**

рівень вищої освіти

другий (магістерський)

Система розпізнавання обличчя з використанням згорткових  
нейронних мереж  
(тема)

Виконав:

студент II курсу, групи КІТМ-22-1

Владислав КУРЕНКО

(власне ім'я, прізвище)

Спеціальність 123 Комп'ютерна інженерія

(код і повна назва спеціальності)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма Комп'ютерні  
інтелектуальні технології

(повна назва освітньої програми)

Керівник зав. каф. КІТС Олег РУДЕНКО

(посада, власне ім'я, прізвище)

2023 р.

## АНОТАЦІЯ

Куренко В. О.. Система розпізнавання обличчя з використанням згорткових нейронних мереж. – Магістерська кваліфікаційна робота.

У магістерській кваліфікаційній роботі вирішено актуальну проблему розробки систем розпізнавання обличчя з використанням згорткових нейронних мереж.

Системи розпізнавання обличчя є досить поширеними у людській діяльності. На цей час, при реалізації таких систем розробники покладаються на великі за обсягом набори даних з зображеннями обличчя. Це призводить до збільшення витрат часу на збір даних та на навчання нейронної мережі. Такі рішення, що використовують відповідні поставленій задачі архітектури нейронних мереж та функції втрат, показують дуже високу точність на складних тестових наборах даних. Однак, системи розпізнавання обличчя можуть використовуватись у більш простих умовах, де не буде такої необхідності у використанні великих наборів даних, складних архітектур нейронних мереж та функцій втрат. Актуальною задачею залишається вдосконалення існуючих рішень для розпізнавання обличчя.

Метою кваліфікаційної роботи є реалізація системи розпізнавання обличчя з використанням згорткових нейронних мереж.

Об'єктом дослідження є процес обробки зображення в системах комп'ютерного зору.

Предметом дослідження є обробка зображення в системах розпізнавання обличчя.

В ході виконання роботи було використано такі методи дослідження: аналіз, синтез, експеримент.

Для отримання векторів ознак із зображення було запропоновано використовувати моделі, що натреновані на наборі даних ImageNet, не дивлячись на те, що в наборі даних ImageNet клас “обличчя людини” не представлений.

Запропоноване рішення зменшує часові витрати на розробку архітектури та тренування нейронної мережі, збір тренувальних даних.

Розроблене рішення може бути використано в системах контролю доступу та відвідуваності, наприклад у домашніх господарствах чи невеликих підприємствах.

У першому розділі було проаналізовано предметну область, відбулась постановка задачі. Комп'ютерний зір полягає в розумінні цифрового зображення на рівні людини. Для цього необхідно отримати корисну інформацію з зображення і проаналізувати її. Зараз активно розвивається глибоке навчання, яке мало дуже сильний вплив на комп'ютерний зір. Його застосування призвело до значного підвищення точності вирішення задач комп'ютерного зору.

Розпізнавання обличчя широко використовується в житлових і нежитлових приміщеннях, для надання доступу авторизованим працівникам, членам сім'ї або попередньо зареєстрованим окремим особам, щоб обмежити доступ неавторизованих осіб в приміщення. Існує безліч важливого обладнання та приміщень, які потребують контролю доступу, відстеження відвідуваності та звітності. Завдяки технології розпізнавання обличчя доступ до обмежених приміщень і обладнання може залишатися контрольованим. Системи відеоспостереження присутні в житлових і нежитлових приміщеннях, і це сприяє добробуту суспільства та підвищенню його рівня безпеки.

У другому розділі було проведено огляд засобів вирішення задачі. Оглянуто засоби реалізації систем комп'ютерного зору, збору та обробки зображень обличчя. Розглянуто принцип роботи згорткових нейронних мереж, методів класифікації даних, проаналізовано існуючі методи розпізнавання обличчя. Python є найбільш відомою мовою програмування для розробки систем штучного інтелекту. Вона проста, але потужна, на ній можна розробляти складні системи для інтелектуального аналізу даних. Обчислення на процесорах загального призначення мають більшу затримку та нижчу пропускну здатність, ніж програми, що виконуються з використанням апаратного прискорення. Окрім

звичайного зберігання великих наборів зображень у файловій системі, є відповідні формати файлів та спеціальні бази даних.

Виявлення об'єктів методом Віоли-Джонса покладається на цінності простих ознак. В такому випадку обчислення відбуваються значно швидше. AdaBoost використовується для вибору ознак, і для навчання класифікатора. Метод гістограми орієнтованих градієнтів полягає в аналізі нормалізованих локальних гістограм орієнтації градієнта зображення, що поділене на комірки. А в бібліотеку DLib нещодавно було додано реалізацію алгоритму виявлення об'єктів з максимальним запасом, і ця функція стала досить затребуваною. Вона дозволяє створити якісну модель для виявлення обличчя з використанням відносно невеликого набору тренувальних даних. Авторами MTCNN було запропоновано глибоку каскадну багатозадачну структуру, яка використовує каскад з трьома етапами з згортковими нейронними мережами, які передбачають розташування обличчя та його орієнтири від грубого до точного. А авторами RetinaFace було запропоновано підхід до виявлення обличчя, що зокрема об'єднує в собі такі функції, як передбачення обмежувальної рамки обличчя та розташування двовимірних орієнтирів обличчя.

Для класифікації даних можна використовувати багатосарові штучні нейронні мережі, які можна описати множиною нейронів та множиною зв'язків між ними.

Багатосарові перцептрони не зовсім підходять для вирішення задач обробки зображень, тому дослідниками було запропоновано згорткові нейронні мережі. Згортковими вони називаються через використання операції згортки для вилучення ознак з зображення.

Для розробки системи для розпізнавання обличчя використовуються три основні етапи: виявлення обличчя, виділення ознак і розпізнавання обличчя. Мета вилучення ознак полягає в тому, щоб описати особливості зображень обличчя, локалізованих на етапі виявлення. Цей крок представляє обличчя вектором ознак, який описує ключові риси зображення обличчя, такі як рот, ніс і очі, з їх геометричними особливостями. Кожне обличчя характеризується своєю будовою,

розміром і формою, які дозволяють його ідентифікувати. Методи EigenFaces, аналіз незалежних компонентів, лінійний дискримінантний аналіз, локальний бінарний шаблон, масштабно-інваріантне перетворення ознак, фільтр Габора, локальне фазове квантування, вейвлети Хаара, перетворення Фур'є широко використовуються для опису ознак обличчя.

Натхненні надзвичайним успіхом на конкурсі ImageNet, такі типові архітектури згорткових нейронних мереж, як наприклад VGGNet, GoogleNet, ResNet, тощо, відомі та широко використовуються як базові моделі для задачі розпізнавання обличчя безпосередньо або з незначними змінами. Окрім класичних типів рішень, деякі типи, наприклад багатозадачні мережі та мережі з декількома входами також використовуються для вирішення цієї задачі. Багато публікацій зосереджено на розробці нових функцій втрат, щоб зробити ознаки не тільки більш роздільними, але й дискримінаційними.

Третій розділ присвячений реалізації рішення. Було наведено структуру системи, визначено засоби її реалізації. Описано процес реалізації системи та наведено результати дослідження.

Програмна реалізація була виконана на мові програмування Python. Для роботи з штучними нейронними мережами використовується бібліотека Keras, яка є високорівневим API платформи TensorFlow. Також у програмній реалізації використовуються й інші бібліотеки. Окрім головного модуля для розпізнавання обличчя, також було розроблено модулі для проведення експериментальних досліджень. Для вирішення задач ідентифікації та верифікації було використано два підходи: з використанням інтелектуального класифікатора, та без нього.

У четвертому розділі було визначено процес підготовки системи та вхідних даних до роботи, описано процес використання системи. На комп'ютері має бути встановлено інтерпретатор Python 3, та необхідні бібліотеки. На пристрої зберігання даних комп'ютера мають бути записані вхідні дані та код програми.

В ході виконання кваліфікаційної роботи було виконано поставлене завдання. Розроблено систему розпізнавання обличчя з використанням згорткових нейронних мереж, і вона була протестована, успішно. В подальшому, можуть

бути вдосконалені використані архітектури нейронних мереж, їх функції втрат та активації.

КОМП'ЮТЕРНИЙ ЗІР, ВИЯВЛЕННЯ ОБЛИЧЧЯ, РОЗПІЗНАВАННЯ ОБЛИЧЧЯ, ЗГОРТКОВІ НЕЙРОННІ МЕРЕЖІ, TENSORFLOW, KERAS, ІДЕНТИФІКАЦІЯ ОСОБИ, АУТЕНТИФІКАЦІЯ ОСОБИ, ВЕРИФІКАЦІЯ ОСОБИ

Публікації здобувача за темою роботи:

1. Куренко В. О. Проблеми використання згорткових нейронних мереж для розпізнавання обличчя. Тренди та перспективи розвитку мультидисциплінарних досліджень : матеріали II Міжнар. студент. наук. конф., м. Хмельницький, 25 листоп. 2022 р. Вінниця, 2022. С. 150–151.

2. Куренко В. О. Особливості використання трансферного навчання нейронних мереж для розпізнавання обличчя. Комп'ютерні інтелектуальні системи та мережі : матеріали XVI Всеукр. науково-практ. WEB конф. аспірантів, студентів та молодих вчен., м. Кривий Ріг, 21–23 берез. 2023 р. Кривий Ріг, 2023. С. 160–162.

Використані публікації керівника та співробітників кафедри, що становлять теоретичну базу роботи:

1. Руденко О. Г., Бодянський Є. В. Штучні нейронні мережі: Навчальний посібник. – Харків: ТОВ «Компанія СМІТ», 2006. – 404 с.

## ЗМІСТ

|  |    |
|--|----|
| Вступ .....  | 14 |
| 1 Аналіз предметної області та постановка задачі .....           | 15 |
| 1.1 Комп'ютерний зір .....                                       | 15 |
| 1.2 Розпізнавання обличчя .....                                  | 16 |
| 1.3 Постановка задачі .....                                      | 19 |
| 2 Огляд засобів вирішення задачі .....                           | 21 |
| 2.1 Засоби реалізації систем комп'ютерного зору .....            | 21 |
| 2.2 Засоби збору та обробки зображень обличчя .....              | 24 |
| 2.3 Методи класифікації даних .....                              | 30 |
| 2.4 Згорткові нейронні мережі .....                              | 32 |
| 2.5 Методи розпізнавання обличчя .....                           | 36 |
| 3 Реалізація рішення .....                                       | 43 |
| 3.1 Структура системи .....                                      | 43 |
| 3.2 Ідентифікація особи без інтелектуального класифікатора ..... | 44 |
| 3.3 Ідентифікація особи з інтелектуальним класифікатором .....   | 48 |
| 3.4 Аутентифікація особи з інтелектуальним класифікатором .....  | 50 |
| 3.5 Верифікація особи з інтелектуальним класифікатором .....     | 52 |
| 3.6 Верифікація особи без інтелектуального класифікатора .....   | 53 |
| 4 Робота системи .....   | 54 |
| 4.1 Підготовка до роботи .....                                   | 54 |
| 4.2 Підготовка вхідних даних .....                               | 54 |
| 4.3 Використання системи .....                                   | 55 |
| Висновки .....   | 61 |
| Перелік використаних джерел .....                                | 62 |
| Додаток А Графічний матеріал кваліфікаційної роботи .....        | 64 |

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ,  
СКРОЧЕНЬ І ТЕРМІНІВ

API – інтерфейс програмування застосунків (англ., Application Programming Interface)

NTFS – файлова система нової технології (англ., New Technology File System)

ReLU – зрізаний лінійний вузол (англ., Rectified Linear Unit)

## ВСТУП

Останніми роками швидко зростає кількість техніки, що має доступ до цифрової камери. Зростає якість зображення з камер, збільшується швидкість обробки даних на комп'ютерах. Це сприяло розвитку систем комп'ютерного зору. На комп'ютерах тепер вирішуються задачі, які 50 років тому було неможливо вирішити.

На сьогодні, розпізнавання обличчя широко застосовується в людській діяльності. Воно використовується в освіті та науці, в державних установах і бізнесі, для ідентифікації та аутентифікації особи.

В цій галузі вже здійснено немало досягнень. Станом на зараз, при реалізації систем розпізнавання обличчя розробники покладаються на великі набори зображень обличчя. Через це збільшуються витрати часу на збір даних та на навчання нейронної мережі. Такі рішення показують дуже високу точність на складних тестових наборах даних. Однак, системи розпізнавання обличчя можуть використовуватись і у більш простих умовах. Подальші дослідження є актуальними, тому що існуючі системи розпізнавання обличчя потребують вдосконалення.

Метою кваліфікаційної роботи є реалізація системи розпізнавання обличчя з використанням згорткових нейронних мереж.

Об'єктом дослідження є процес обробки зображення в системах комп'ютерного зору.

Предметом дослідження є обробка зображення в системах розпізнавання обличчя.

Було проаналізовано фахові джерела. В ході виконання роботи використані такі методи дослідження: аналіз, синтез, експеримент.

# 1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧІ

## 1.1 Комп'ютерний зір

Комп'ютерний зір займається отриманням корисної інформації з цифрового зображення і подальшим аналізом. Тобто, комп'ютер має розуміти, що відбувається на зображенні, як людина. Комп'ютерний зір застосовується в різних галузях, зокрема у промисловості, медицині, військовій справі. При вирішенні задач у комп'ютерному зорі можна виділити такі етапи, як: отримання зображення, обробка зображення, вилучення ознак, аналіз зображення.

Вже до 2000 року було проведено немалу кількість досліджень, яка стала фундаментом для багатьох сучасних розробок. Зараз методи, які були запропоновані у минулому столітті, поступово замінюються новими методами, які базуються на глибокому навчанні. Але, зараз дослідники пропонують і гібридні методи, де методи засновані на ознаках поєднуються з інтелектуальними класифікаторами.

Останніми роками дуже активно розвивається глибоке навчання, яке дуже сильно вплинуло на галузь комп'ютерного зору. Застосування методів, заснованих на глибокому навчанні, призвело до значного збільшення точності роботи алгоритмів. Глибоке навчання продовжує розвиватися, з'являються нові архітектури штучних нейронних мереж, які показують вищу точність, порівняно з попередніми розробками.

Також розвитку галузі комп'ютерного зору сприяв розвиток пристроїв зберігання та обробки даних, що надало технічні можливості для реалізації глибокого навчання, та поява недорогих та якісних цифрових камер, що сприяло полегшенню у процесі збору цифрових зображень.

Серед застосувань комп'ютерного зору можна виділити, такі додатки:

- оцінювання пози;
- розпізнавання мови по губах або жестах;

- доповнена реальність;
- анотування зображень;
- безпілотні апарати;
- пошук зображень на основі вмісту;
- розпізнавання символів;
- медична діагностика;
- аналіз відео;
- розпізнавання обличчя.

Цифрове зображення на комп'ютері можна представити як багатовимірний масив. Наприклад, зображення в градаціях сірого можна представити як двовимірний масив, де кожен елемент описує інтенсивність сірого кольору в конкретному пікселі. Якщо зображення кольорове, то масив описує значення кольору у каналі відповідно до кольорової моделі у певному пікселі зображення. Приклад опису зображення в форматі RGB наведено на рисунку 1.1.

|  |  |  | R   |     |   | G   |     |     | B   |     |     |
|--|--|--|-----|-----|---|-----|-----|-----|-----|-----|-----|
|  |  |  | 255 | 255 | 0 | 255 | 255 | 0   | 0   | 0   | 255 |
|  |  |  | 0   | 0   | 0 | 255 | 255 | 255 | 255 | 255 | 255 |
|  |  |  | 0   | 0   | 0 | 255 | 255 | 255 | 0   | 0   | 0   |

Рисунок 1.1 – Приклад опису кольорового зображення в форматі RGB

Найбільш поширеними є кольорові моделі RGB та HSV, окрім них є й такі: CMYK, LMS, HSL, AHLS, RYB, YUV.

## 1.2 Розпізнавання обличчя

Системи розпізнавання обличчя дозволяють співставити вектор ознак обличчя людини, отриманий з зображення, з базою еталонних векторів, з метою

ідентифікації особи. Враховуючи доступність цифрових камер, обличчя як біометрична характеристика є одним з найбільш підходящих для ідентифікації особи.

Розпізнавання обличчя для людини може ускладнюватися при збільшенні кількості людей. Сучасні комп'ютери можуть мати великий обсяг пам'яті та високу швидкість обчислень, тому порівняння обличчя людини з великою базою еталонів не є проблемою.

Розпізнавання обличчя широко застосовується, зокрема в системах контролю доступу та в системах контролю відвідуваності. Біометричні характеристики для ідентифікації особи мають низку переваг порівняно з іншими. Номер для ідентифікації може бути легко втрачений, вкрадений, або підібраний. Засоби апаратної ідентифікації можуть стати несправними, бути втраченими або вкраденими. Безумовно, зловмисники намагаються атакувати і біометричні методи ідентифікації, однак це важче зробити, і також розробники систем біометричної ідентифікації вдосконалюють свої розробки для запобігання обману.

Розпізнавання обличчя використовується у багатьох сферах. В безпекових цілях воно застосовується для контролю доступу в будівлях, на закритих об'єктах, при перетині кордону, навіть для доступу до комп'ютера або іншого пристрою, наприклад банкомату. Також обличчя як біометрична характеристика може застосовуватися для аутентифікації користувача для доступу в веб-ресурс або додаток, особливо це поширено на мобільних пристроях, у зв'язку з високою якістю цифрових камер у них. Активно поширюється розпізнавання обличчя в правоохоронній галузі. Його використовують для пошуку злочинців або викрадених людей. Камери розташовують в місцях масового скупчення людей, наприклад на людних вулицях, зупинках громадського транспорту, стадіонах, магазинах, у закладах освіти. Також поліція може проводити ідентифікацію людини по фотографії обличчя з певного джерела пошуком у базі даних. По обличчю може проводитись верифікація особи в державних установах або банках, наприклад обличчя людини у банку порівнюється з його документами перед тим як видати йому гроші. Розпізнавання обличчя також використовується для

ідентифікації людей шляхом пошуку їх у базі даних, наприклад для пошуку свого давнього знайомого в соцмережах по фотографії. Розпізнавання обличчя часто використовується в комерційних і житлових приміщеннях, щоб контролювати доступ авторизованих осіб, та обмежити доступ неавторизованих осіб. Існує багато спеціалізованого обладнання та приміщень, де необхідний ретельний контроль доступу. Завдяки системам розпізнавання обличчя можна керувати доступом осіб до певних об'єктів. Системи відеоспостереження присутні в різних типах приміщень, і вони можуть зберігати відеозаписи для їх подальшого перегляду та аналізу. Впровадження технологій розпізнавання обличчя в системи безпеки може значно підвищити рівень захищеності в багатьох галузях людської діяльності. Склади та заводи є важливими у виробництві різних типів товарів. Обладнання з функцією розпізнавання обличчя гарантує, що лише навчений персонал з відповідними дозволами зможе отримати доступ до важливого обладнання та працювати з ним. Системи безпеки з функцією розпізнавання обличчя можуть допомогти знаходити в потоці певних людей і попереджати власника, коли така особа підходить або заходить до приміщення. У магазині може бути відвідувач, який до цього погрожував співробітникам або був спійманий на крадіжці. Систему розпізнавання обличчя можна налаштувати так, щоб сповістити охорону, якщо ця особа заходить у приміщення магазину. Використання технологій розпізнавання обличчя для захисту житла орієнтується на приватних осіб, щоб захистити місце проживання. Якщо система безпеки ідентифікує особу в приміщенні як невідому, вона може негайно сповістити власників житла та правоохоронні органи. Після спалаху пандемії, здоров'я стало для людей пріоритетом. Розпізнавання обличчя використовується для побудови мережі контактів між людьми, для того, щоб запобігти поширенню вірусів. Багато існуючих систем управління робочим часом вимагають від співробітників введення пін-кодів, і навіть біометричних даних, наприклад, відбитків пальців, для фіксації входу та виходу з приміщення. Однак використання пін-кодів має ризик видавання себе за іншу людину, так як співробітники можуть випадково загубити або поділитися своїм кодом. Ідентифікація особи по відбитку пальців

потребує контакту людини з поверхнею, а це підвищує ризик поширення вірусу. Оскільки зараз багато компаній виводить співробітників з дистанційної роботи у офіси, використання розпізнавання обличчя в системах управління робочим часом буде безпечнішим для співробітників. Часто розпізнавання обличчя асоціюється лише з використанням у системах безпеки. Але широкий спектр закладів може отримати більші прибутки, пропонуючи персоналізований підхід до клієнтів. Учасники програми лояльності в магазині можуть зареєструвати своє обличчя, але такі системи з розпізнаванням обличчя можуть навіть забезпечити певний рівень персоналізації для незареєстрованих відвідувачів. Власникам магазинів чи ресторанів, які хочуть підвищити свої прибутки, було б доречно використовувати розпізнавання обличчя. Можна налаштувати систему на ідентифікацію відвідувача як важливого клієнта, і сповістити про його прибуття співробітників, щоб вони надали йому особисте привітання та підвищену увагу. В готелях можна використовувати ідентифікацію особи по обличчю для обліку вподобань клієнтів. Після ідентифікації відомого раніше клієнта, персонал отримує сповіщення про відомі уподобання клієнта, такі як поверх кімнати, бажані продукти в холодильнику, тощо. Розпізнавання обличчя може використовуватись для ідентифікації, або як один із етапів аутентифікації для безконтактної оплати. У правоохоронних органах також активно впроваджуються технології розпізнавання обличчя. Однак у громадськості існує занепокоєння щодо стеження за людьми та виникнення помилок при ідентифікації. І це сповільнило процес впровадження розпізнавання обличчя у правоохоронних органах. Тому регулювання етичних питань все ще потребує вдосконалення [1, 2].

### 1.3 Постановка задачі

Для реалізації поставленої мети роботи необхідно виконати такі завдання:

- проаналізувати та обрати засоби для реалізації системи комп'ютерного зору;
- проаналізувати та обрати засоби для збору та обробки зображень обличчя;

- проаналізувати методи вилучення ознак із зображення;
- проаналізувати методи класифікації даних;
- проаналізувати методи розпізнавання обличчя;
- запропонувати метод розпізнавання обличчя для реалізації в системі;
- розробити програмну складову системи;
- протестувати систему.

Вхідними даними є набір зображень з обличчям.

В програмній складовій треба реалізувати такі функції:

- вилучення ознак зображення або набору зображень;
- запис ознак зображення у файл та їх зчитування;
- розрахунок відстані між векторами ознак;
- розпізнавання обличчя;
- модуль для тренування та тестування моделей глибокого навчання;
- модуль для візуалізації отриманих результатів експериментів.

## 2 ОГЛЯД ЗАСОБІВ ВИРІШЕННЯ ЗАДАЧІ

### 2.1 Засоби реалізації систем комп'ютерного зору

Для реалізації систем комп'ютерного зору необхідні відповідні програмні та апаратні засоби. Активне поширення систем штучного інтелекту, яке відбувається останніми роками, сприяло розвитку програмних засобів розробки таких систем.

Не припиняється розвиток мов програмування, з кожною наступною версією вони отримують новий функціонал. Продовжується активний розвиток бібліотек для розробки систем штучного інтелекту, зокрема, для обробки зображень, комп'ютерного зору, інтелектуального аналізу даних. Створюються нові інструменти, вдосконалюються вже існуючі.

Серед мов програмування, які використовуються для розробки систем штучного інтелекту, найбільшою популярністю користується Python. Це проста, але потужна мова програмування, на якій можна швидко написати простий скрипт, або розробляти складні програмні системи. Найбільш часто Python використовується в таких напрямках: наука та освіта, системи штучного інтелекту, розробка веб-додатків, розробка додатків з графічним інтерфейсом користувача.

Серед її особливостей варто виділити такі:

- динамічна типізація;
- високорівневі структури даних;
- багатоплатформність;
- доступно об'єктно-орієнтоване програмування;
- потужна стандартна бібліотека;
- реалізовано велику кількість бібліотек для інтелектуального аналізу даних;
- достатня кількість безкоштовних середовищ розробки.

OpenCV це бібліотека з відкритим кодом для розробки систем комп'ютерного зору. Містить в собі велику кількість класичних та сучасних

алгоритмів комп'ютерного зору. Має інтерфейси для Python, C++, та деяких інших мов програмування. Бібліотека підтримується на операційних системах Windows, Linux, MacOS, Android. OpenCV орієнтується на паралельні алгоритми, що полегшує обробку зображень в реальному часі.

Dlib це бібліотека з відкритим кодом для вирішення різних класів задач, зокрема машинного навчання та обробки зображень. Бібліотека написана на C++, але може використовуватись і в Python. Бібліотека підтримується на операційних системах Windows, Linux, MacOS. В бібліотеці Dlib реалізовано функціонал для роботи з глибоким навчанням, кластеризацією даних, багат шаровим перцептроном, методом опорних векторів для різних задач інтелектуального аналізу даних. Для обробки зображень доступний функціонал для зчитування зображень, базові операції над зображеннями та для вилучення ознак з зображення, а також реалізовані методи виявлення та розпізнавання обличчя.

Pillow є форком бібліотеки PIL. Бібліотека використовується для обробки зображень. В ній реалізований функціонал для зчитування зображень та їх запису, геометричні перетворення, фільтрації, отримання гістограм, тощо.

Бібліотека scikit-learn, яку ще називають sklearn, є однією з найбільш поширених бібліотек для машинного навчання з відкритим кодом для мови програмування Python. У sklearn реалізовано основні алгоритми машинного навчання, і інструменти для попередньої обробки даних та їх генерації, оцінки моделі.

TensorFlow – ще одна з найбільш поширених бібліотек для машинного навчання з відкритим кодом. Ця бібліотека вважається платформою для створення і розгортання моделей машинного навчання на різних пристроях: настільних комп'ютерах, мобільних пристроях, серверах, мікроконтролерах, програмованих логічних інтегральних схемах, тощо. Також є інструменти для попередньої обробки даних та аналізу моделі під час її розробки, та репозиторій існуючих наборів даних та натренованих моделей.

Keras – високорівневий API платформи TensorFlow. В ній реалізовані необхідні будівельні блоки, що дозволяє швидко створювати моделі глибокого

навчання. Не дивлячись на простоту API, є можливість більш гнучко налаштувати необхідні частини функціоналу, завдяки тісному зв'язку з TensorFlow. Бібліотека підтримується на операційних системах Windows, Linux, MacOS. Є можливість реалізації паралелізму, відповідно можливий запуск моделей на кластерах з центральними, графічними, або тензорними процесорами. В Keras доступні інструменти для роботи з зображеннями та текстом, налаштуванням гіперпараметрів, оптимізації моделей шляхом квантування, тощо.

Прогрес у комп'ютерному зорі обумовлений не тільки застосуванням глибокого навчання та дуже великими наборами даних, але і розвитком апаратних засобів, який дає доступ до паралельної обчислювальної архітектури для ефективного тренування та використання нейронних мереж. Апаратне прискорення покладається на переваги апаратних засобів для виконання обчислень із вищою пропускну здатністю та меншою затримкою, ніж програми, що працюють на центральних процесорах загального призначення. Так склалося, що архітектури в стилі фон Неймана призначаються для послідовних обчислень із складним плануванням завдань, і добре справляються з цим. Вони характеризуються високим споживанням енергії та низькою пропускну здатністю пам'яті при переміщенні даних під час роботи глибоких нейронних мереж, де необхідні паралелізм в обчисленнях, висока пропускна здатність пам'яті і можливість повторного використання даних. Як показує досвід, обчислення можуть бути виконані швидше на апаратних засобах, які спроектовані спеціально під конкретну програму. Апаратне прискорення дає високий ступінь паралелізму. Графічні процесори від самого початку розробили для прискорення обробки графічної інформації на комп'ютерах. Та зараз графічні процесори використовуються ще як паралельні обчислювальні процесори з високою пропускну здатністю пам'яті для виконання обчислень по паралельних алгоритмах. У згортковій нейронній мережі можна виділити переваги алгоритмічного паралелізму в таких операціях: згортка та об'єднання може виконуватись паралельно, активація кожного нейрона в повнозв'язному шарі теж може виконуватись паралельно. Завдяки структурі для паралельних обчислень та

великим можливостям обчислень з числами з плаваючою комою графічні процесори вдало зроблені для прискорення в глибокому навчанні. Графічні процесори забезпечують дуже високу пропускну здатність і часто використовуються для апаратного прискорення обчислень у нейронних мережах, але вони не зовсім підходять для пристроїв з обмеженим енергоспоживанням або обчислювальною потужністю, наприклад це вбудовані системи. Прискорення обчислень в нейронних мережах можливе іншим способом, що базується на використанні програмованих користувачем вентильних матриць. На них можна реалізувати архітектуру апаратного засобу для конкретної програми, і це забезпечує високий рівень гнучкості. Як показує досвід, моделі нейронних мереж часто тренуються на потужних обчислювальних платформах, таких як графічні процесори, а програмовані користувачем вентильні матриці використовуються для виконання вже навченої моделі. При використанні глибокого навчання в комп'ютерному зорі часто виникає високе навантаження з точки зору обчислень та пам'яті, а пропускну здатність пам'яті програмованих користувачем вентильних матриць досить часто у 10 разів менша, ніж у графічних процесорів. Проблему можна вирішити шляхом оптимізації, зокрема стисненням моделей. Тензорний процесор це інтегральна схема специфічного застосування, яку розробили дослідники з Google. Це обчислювальний пристрій, який спеціально спроектований для інтелектуального аналізу даних з використанням штучних нейронних мереж, і адаптований під Tensorflow. Тензорний процесор використовується для виконання великої кількості арифметичних операцій низької точності. Нещодавно тензорні процесори почали з'являтися у мобільних пристроях як співпроцесори, які призначені для роботи з додатками, що використовують штучні нейронні мережі [3].

## 2.2 Засоби збору та обробки зображень обличчя

При розробці систем комп'ютерного зору обов'язковим компонентом є набір даних, що містить зображення. Зображення використовуються для

тренування моделей машинного навчання, оцінки та тестування моделі. Розмір цих наборів даних зазвичай коливається між декількома сотнями та декількома мільйонами зображень. Відповідно, постає питання про формат зберігання даних. Найбільш простим варіантом є зберігання зображень в їх звичних файлових форматах у типовій файловій системі, наприклад NTFS. Такий варіант підходить, якщо у вас невеликий набір даних. Якщо файли займають багато місця, їх можна упакувати в формати .zip або .tar. Також типовою рекомендацією по зберіганню наборів даних зображень є серіалізація даних у двійковий формат, тому що вони займають менше місця, та можуть бути зручно передані між різними операційними системами та мовами програмування. Серед найбільш поширених засобів зберігання великих наборів даних для вирішення задач глибокого навчання є HDF5 та LMDB.

HDF5 це не просто формат файлу, він включає в себе моделі для логічної організації та доступу до даних, а також програмних засобів, зокрема інструментів, бібліотек та інтерфейсів для роботи з форматом. Файл формату HDF5 можна вважати контейнером, в якому містяться різні об'єкти чи набори даних. Набір даних може бути представлений як зображення, таблиця, або навіть документи, наприклад в форматі xlsx. Основними об'єктами в моделі даних виступають групи та набори даних. Групи займаються організацією об'єктів даних. Кожен файл формату HDF5 містить кореневу групу, в якій можуть бути інші групи або можуть бути об'єкти з інших файлів [4].

LMDB це база даних з відображенням у пам'яті, база даних типу ключ-значення, яка базується на B+ деревах. Одна з її особливостей полягає у тому, що за керування сторінками відповідає операційна система, тому що використовується однорівневе сховище на основі файлів мапи пам'яті. Читачі не блокують записувачів, записувачі не блокують читачів, а читачі не блокують читачів. Допускається стільки екземплярів даних у момент часу, скільки є транзакцій у цей момент часу, тобто за принципом багато читає, один записує. І також реалізовано вільний список сторінок, що використовується для відстеження та повторного використання сторінок, щоб кожний раз не виділяти пам'ять [5].

При розробці системи розпізнавання обличчя однією з перших задач, що виникає, є збір зображень та виявлення обличчя на ньому. Існує достатня кількість методів виявлення обличчя, та найбільш поширеними підходами є такі:

- метод Віоли-Джонса;
- гістограма орієнтованих градієнтів і метод опорних векторів;
- з використанням моделей глибокого навчання.

Метод Віоли-Джонса використовують на комп'ютерах з низькою обчислювальною потужністю для виявлення обличчя на зображенні, хоча й метод може бути адаптований під виявлення інших об'єктів. Цей метод характеризується високою швидкістю роботи, та порівняно вищою кількістю помилок. Метод залишається актуальним і на сьогоднішній день. Ключова особливість цього методу – використання каскадного класифікатора та ознак Хаара.

Процес виявлення об'єктів методом Віоли-Джонса полягає в аналізі зображення, базуючись на цінності простих ознак. Достатньо аргументів, чому в даній задачі ознаки краще, ніж пікселі. З використанням певної кількості тренувальних даних, можна визначити цінні знання. Також при використанні ознак обчислення виконуються значно швидше, ніж при використанні пікселів. Алгоритм AdaBoost використовується і для вилучення невеликого набору ознак, і для тренування класифікатора. Кожну ознаку можна ефективно обчислити, але обчислення всіх ознак є досить неефективним. Доведено, що для формування якісного класифікатора достатньо невеликого набору ознак. Та проблема полягає в пошуку цих особливостей. Для кожної ознаки формується така порогова функція класифікації, щоб була неправильно класифікована невелика кількість екземплярів. Як показує практика, жодна ознака не здатна провести класифікацію з низькою похибкою. Головна ідея каскадних класифікаторів це те, що можна сконструювати менші, та більш ефективні посилені класифікатори, які відкидують багато негативних вікон, виявляючи більшість позитивних екземплярів. Класифікатор можна налаштувати так, щоб рівень помилково негативного спрацювання був мінімальним. Прості класифікатори

використовуються щоб відкинути більшу частину вікон. Далі будуть використовуватись більш складні класифікатори для зменшення показника помилково позитивної класифікації. Такий процес виявлення об'єктів на зображенні можна описати як вироджене дерево рішень, яке можна визначити як каскад. Позитивне спрацювання першого класифікатора запускає другий класифікатор, його позитивний результат запускає наступний класифікатор, і так далі. Якщо класифікатор на будь-якому кроці видасть негативний результат, переходу на наступний крок не буде. При використанні одношарового перцептрона необхідно було б принаймні у двадцять разів більше операцій, ніж при обчисленнях для двох класифікаторів ознак у вікні. Сенс використання каскаду полягає в тому, що на зображенні переважна більшість вікон будуть негативними. Тобто, така структура використовується для того, щоб відхилити більшість негативних результатів ще на початку [6].

Гістограма орієнтованих градієнтів та метод опорних векторів використовується для виявлення об'єктів на зображенні, зокрема для виявлення обличчя. Гістограма орієнтованих градієнтів виступає в якості дескриптора зображення, а метод опорних векторів використовують для класифікації.

Використання гістограми орієнтованих градієнтів для обробки зображень базується на аналізі нормалізованих локальних гістограм орієнтації градієнта на зображенні, розділеному на комірки. Об'єкт на зображенні можна описати розподілом градієнтів або напрямків, навіть якщо невідома точна позиція. Для цього необхідно поділити зображення на невеликі комірки, і для кожної з них вираховуються напрямки градієнта або орієнтації країв, і ці дані формують гістограму. Для підвищення якості дескриптора рекомендовано також виконувати нормалізацію в комірках. Зображення описується як об'єднання цих гістограм. Виявлення об'єктів на зображенні цим методом полягає у використанні вектора ознак зображення з використанням дескриптора на базі гістограми орієнтованого градієнта у класифікаторі, що зазвичай базується на методі опорних векторів. Цей підхід має певні позитивні сторони. Такий метод визначає градієнтну або краєву структуру зображення, яка є особливою. Такий підхід сприяє зменшенню

чутливості до зміни освітлення на зображенні та до поворотів та зсувів об'єкта. Цей дескриптор не такий чутливий до таких перетворень, та підходить для визначення особливостей притаманних обличчю. Такий підхід показує високу швидкість обробки, високу точність, і є актуальним підходом для вирішення задачі виявлення об'єктів на зображенні [7, 8].

Найбільш сучасним підходом до виявлення обличчя є використання глибокого навчання. Використання нейронних мереж забезпечує дуже високу точність при вирішенні цієї задачі. Серед реалізацій цього підходу є, зокрема, такі моделі: DLib, MTCNN, RetinaFace.

Кілька років тому в бібліотеку DLib було додано реалізацію алгоритму виявлення об'єктів з максимальним запасом. З того часу цей функціонал став досить популярним. Завдяки йому є можливість створити якісну модель детектора з використанням відносно невеликого набору даних. Є випадки, де цей алгоритм для виявлення об'єктів успішно натреноує модель для виявлення обличчя на тренувальному наборі з чотирьох зображень. На початку, реалізація цього підходу для виявлення об'єктів в DLib використовувала дескриптор з використанням гістограми орієнтованих градієнтів, після чого йшов один лінійний фільтр. Тому такий детектор не міг навчитися виявляти об'єкти, у яких були складні пози або вони мали дуже багато різних варіантів свого вигляду. Для того, щоб подолати цю проблему, зазвичай треба було тренувати окремі детектори під конкретну позу. Часто це хоч не ідеально, але вирішує цю проблему. Та останні декілька років активно розвиваються згорткові нейронні мережі, і вони спроможні вирішувати ці проблеми одразу в одній моделі. Тому було вирішено, що необхідно змінити в реалізації алгоритму виявлення об'єктів з максимальним запасом дескриптор на основі гістограми орієнтованих градієнтів на згорткову нейронну мережу. У нових версіях бібліотеки реалізований саме такий підхід. Розробниками бібліотеки було проведено експеримент на декількох тестових зображеннях. Детектор на основі гістограми орієнтованих градієнтів показує високий результат у виявленні фронтального обличчя, але детектор на основі згорткової нейронної мережі ще й краще справляється з більш складними випадками [9].

Виявлення та вирівнювання обличчя є складною задачею комп'ютерного зору, через варіації поз, зміни освітлення, зашумлення та перешкоди. Нещодавні відкриття в області глибокого навчання доводять, що з використанням згорткових нейронних мереж можна досягти високих результатів у цих двох задачах. Авторами MTCNN було запропоновано глибоку каскадну багатозадачну структуру. Зокрема, їх розробка застосовує каскадну структуру з проходженням трьох етапів згорткових нейронних мереж, які локалізують обличчя та визначають його ключові точки від грубого до точного. Запропонований підхід показує високу точність виявлення, і обробка даних відбувається швидко. Отримуючи вхідне зображення, його розмір змінюється у різних масштабах, для створення піраміди зображення, яка виступає вхідними даними для даної каскадної структури. На першому кроці обробки використовується згорткова нейронна мережа P-Net, щоб запропонувати кандидатів з їх обмежувальними рамками. Після цього, оцінені обмежувальні рамки використовуються для калібрування кандидатів. Потім виконується неадекватне придушення, щоб кандидати, які занадто перекриваються, були об'єднані. На наступному кроці усі ці кандидати передаються в згорткову нейронну мережу під назвою R-Net. Вона відкидає велику кількість хибних кандидатів, і також виконує калібрування обмежувальних рамок та об'єднання кандидатів. Третій крок схожий на другий, але на цьому кроці необхідно виконати більш детальний опис обличчя. Вираховується розташування ключових точок обличчя. Також розробниками цього рішення було зменшено загальну кількість фільтрів і змінено розмір фільтра з 5 на 5, на 3 на 3 [10].

Не дивлячись на досягнуті результати в задачі виявлення обличчя, такі задачі, як двовимірне вирівнювання та тривимірна реконструкція обличчя залишаються актуальними. Дослідниками було запропоновано метод локалізації обличчя RetinaFace, що зокрема об'єднує в собі такі функції, як передбачення обмежувальної рамки обличчя та розташування двовимірних ключових точок обличчя. Автори вручну розмітили ключові точки обличчя в наборі WiderFace і виконали напівавтоматичну розмітку тривимірних вершин на зображеннях

обличчя з використанням декількох тренувальних наборів. Використовуючи додаткову розмітку було реалізовано визначення тривимірних вершин, що проектувалися на площину зображення, обмежену загальною тривимірною топологією. Дослідження показали, що розроблене рішення може одночасно забезпечити виконання таких завдань, як виявлення обличчя, двовимірне вирівнювання обличчя по його ключових точках, і тривимірну реконструкцію обличчя [11].

### 2.3 Методи класифікації даних

Для вирішення задачі ідентифікації, аутентифікації чи верифікації особи можна використовувати інтелектуальні класифікатори. Ці класифікатори, отримуючи в якості вхідних даних вектор ознак зображення, вирішують поставлене завдання класифікації. Серед існуючих методів класифікації варто виділити такі:

- метод опорних векторів;
- метод К-найближчих сусідів;
- наївний байесівський класифікатор;
- дерево рішень;
- метод випадкового лісу;
- багат шаровий перцептрон.

Метод опорних векторів є досить поширеним алгоритмом машинного навчання для класифікації даних. Зазвичай, використовується для бінарної класифікації, але використання цього методу може бути адаптоване і для багатокласової класифікації. Для початку, вектори ознак з тренувального набору даних розміщуються в багатовимірному просторі. А навчання моделі полягає в пошуку гіперплощини, що розділить ці точки двох класів, до того ж, з максимальною відстанню від гіперплощини до найближчої точки. При вирішенні задачі класифікації вхідний вектор ознак розміщується в багатовимірному просторі, та визначається, з якого боку від визначеної гіперплощини він

знаходиться.

Метод К-найближчих сусідів є досить простим та непоганим алгоритмом класифікації. Він досі використовується для вирішення простих задач з невеликим набором даних. В цьому методі процесу тренування немає як такого, тренувальні дані потрібно просто зберегти в пам'яті, а всі розрахунки будуть відбуватися вже під час класифікації. Процес класифікації полягає в розрахунку відстані між вхідними та тренувальними даними, і потім вираховується, точок якого класу найбільше поруч з вхідними даними.

Наївний байесівський класифікатор ґрунтується на теоремі Байеса, та використовується для класифікації даних. Особливість методу полягає в тому, що ознаки вважаються незалежними та рівноцінними. Класифікація полягає у виборі класу з найбільшою розрахованою по формулі на базі теореми Байеса вірогідністю.

Дерево рішень це алгоритм машинного навчання, що можна використовувати для класифікації. Процес прийняття рішення можна візуалізувати у форматі деревоподібної структури. Процес навчання класифікатора використовує стратегію розділай-володарюй, для визначення оптимальних точок для розділення у дереві. Цей процес повторюється рекурсивно зверху донизу до отримання бажаного результату.

Метод випадкового лісу це метод класифікації даних, що ґрунтується на деревах рішень. Кінцевий результат методу випадкового лісу ґрунтується на результатах роботи декількох дерев рішень. Тобто, використовується ансамбль – набір класифікаторів. Кожен екземпляр дерева рішень в ансамблі працює з конкретною вибіркою з набору даних. Точність класифікації у методі випадкового лісу зазвичай вища, ніж в одного дерева рішень.

Якщо штучні нейрони з'єднати між собою, то отримаємо штучну нейронну мережу. Тобто, штучну нейронну мережу можна описати множиною нейронів та множиною зв'язків між ними. Велика кількість досліджень сформувала різні функції активації та кількості шарів нейронів у штучних нейронних мережах. Багат шарові нейронні мережі складаються з вхідного та вихідного шарів, і

прихованих шарів. У галузі штучного інтелекту багатошаровий перцептрон є найбільш відомим, і цей тип штучної нейронної мережі використовується ледь не найчастіше. Перехід від одношарових до багатошарових перцептронів відбувся, щоб мати можливість збільшити складність областей прийняття рішень, бо наявність лише одного шару накладала певні обмеження. Одношаровий перцептрон, який має один вхід, створює області прийняття рішень у формі півплощини. У двошаровій мережі кожен нейрон буде діяти як звичайний перцептрон, тому ця мережа може створити вже опуклі області прийняття рішень, тому що півплощини перетинаються. А тришаровий перцептрон здатен створювати довільні області прийняття рішень. При застосуванні багатошарових штучних нейронних мереж не відбувається підвищення точності порівняно з одношаровими штучними нейронними мережами, якщо вони мають лінійну активаційну функцію. Саме нелінійні функції активації надають вказані переваги багатошаровому перцептроні. На сьогодні, найчастіше використовуються сигмоїдальні функції активації. Багатошаровий перцептрон може бути універсальним апроксиматором, бо навіть з лише одним прихованим шаром він має можливість виконати апроксимацію будь-якої безперервної функції [12, 13].

## 2.4 Згорткові нейронні мережі

Новим етапом розвитку комп'ютерного зору стало використання згорткових нейронних мереж. Застосування глибокого навчання сприяло значному підвищенню точності моделей для класифікації зображень. Останніми роками згорткові нейронні мережі є основним інструментом вирішення багатьох задач комп'ютерного зору.

Багатошаровий перцептрон можна використовувати для вирішення задачі класифікації зображень. Але в більшості випадків набори зображень формують великі обсяги даних. Відповідно, тоді у нейронної мережі буде дуже велика кількість параметрів. А велика кількість параметрів буде сприяти переоснащенню моделі, що є небажаним. Ще одне обмеження полягає в тому, що багатошарові

персептрони є нестійкими до зсуву зображення. Згорткові нейронні мережі були розроблені для ефективного вирішення задач комп'ютерного зору. Згортковою вона називається через те, що для вилучення ознак з зображення використовується операція згортки. Однією з особливостей таких мереж є спільне використання параметрів, тобто одні й ті ж ваги використовуються для обробки різних частин одного зображення. Це дозволяє виявляти особливості, що незалежні від зсуву або повороту. Такий підхід скорочує кількість параметрів моделі, якщо порівнювати з багатошаровим персептроном, і це підвищує ефективність моделі. Згорткові нейронні мережі складаються з двох частин: для вилучення ознак та для класифікації. Частина для вилучення ознак складається з згорткових блоків, які складаються з шарів згортки та шарів об'єднання. Отримані вектори ознак зображення проходять через повнозв'язні шари мережі для класифікації зображення. Шари в мережі мають свою висоту, ширину, та глибину, бо кольорове зображення представляється як тривимірний масив. Для обробки зображення в згорткових шарах використовують фільтри, його типовий розмір складає три на три пікселя. Фільтр переміщається по вхідних даних, як ковзаюче вікно, і на кожному кроці виконується згортка. На виході отримуємо одне число, яке передається через активаційну функцію, і дані з виходу активаційної функції заносяться у карту активації. Вхідні дані після операції згортки можуть мати менший просторовий розмір, ніж були на вході. Так буває, якщо використовується крок більше одиниці, або не використовуються методи заповнення меж зображення. У згортковому шарі зазвичай не один фільтр, а множина таких фільтрів. Ваги у фільтрі ініціалізуються випадковими значеннями та вивчаються в процесі тренування. У перших шарах фільтри вчать виявляти прості елементи зображення – прямі та краї. Однак, у наступних шарах на зображенні виявляються складніші структури. Шари об'єднання використовуються для зменшення просторового розміру карт активації. Це зменшує кількість параметрів, що полегшує обчислення та запобігає переоснащенню нейронної мережі. Об'єднання це операція для зменшення розміру двовимірних даних. Зазвичай використовується об'єднання двох типів:

середнє або максимальнє. Наприклад, при максимальному об'єднанні на вихід буде передано максимальнє значення із вхідних даних, які були оброблені фільтром. У фільтра немає параметрів, які можуть навчатися. Після того, як зображення вилучаються ознаки, отримані дані передаються в повнозв'язні шари. Кількість нейронів у вихідному шарі мережі дорівнює кількості класів, і кожен вихідний нейрон пов'язаний з конкретним класом. Відповідно, значення кожного нейрона на виході мережі є ймовірністю того, що зображення відноситься до відповідного класу [14].

У згорткових нейронних мережах застосовуються різні функції активації. Функція активації визначає, які дані мають бути передані далі. У нейронній мережі як вхідні дані кожен нейрон у приймає вихідне значення нейронів із попереднього шару, а потім передає оброблене значення далі на наступний шар. Сигмоїдальна функція активації може відобразити дійсне число у проміжку від 0 до 1, тому цю функцію використовують для бінарної класифікації. Функція гіперболічного тангенса може відобразити дійсне число у проміжку від -1 до 1. Так як середнє значення виходу гіперболічного тангенса дорівнює 0, може бути виконана нормалізація даних. Це покращує тренувальний процес у штучній нейронній мережі. Крім того, ще однією відомою функцією активації є ReLU. Коли вхідне значення менше 0, на виході функції активації ReLU буде 0, а коли більше або дорівнює 0, то вихідне значення буде дорівнювати вхідному. Перевагою цієї функції порівняно з сигмоїдою або гіперболічним тангенсом є те, що при її використанні прискорюється тренувальний процес. Для задач бінарної класифікації рекомендовано використовувати сигмоїду, для багатокласової класифікації рекомендують використовувати функцію softmax. Для прихованих шарів зазвичай обирають функцію ReLU. Для обчислення відстані між прогнозованим і фактичним значенням обчислюється функція втрат. Ця функція переважно застосовується як критерій навчання у задачі оптимізації. Функції втрат у згорткових нейронних мережах можна використовувати для зменшення помилки при вирішенні задачі класифікації. До відомих функцій втрат відносять середню абсолютну похибку, середню квадратичну похибку, перехресну

ентропію, контрастну та триpletну функції, тощо. Перехресна ентропія є поширеним варіантом функції втрат, зазвичай використовується в згорткових нейронних мережах, у яких на останньому шарі функція активації softmax. При використанні контрастної функції втрат збільшується відстань між різними класами та зменшується відстань між екземплярами одного класу. При використанні tripletної функції втрат зменшується відстань між якірним екземпляром та позитивним екземпляром та збільшується відстань між якірним екземпляром та негативним екземпляром. У згорткових нейронних мережах нерідко виникає необхідність оптимізувати невивуклі функції. Відповідні математичні розрахунки мають великий обсяг та потребують високого рівня потужності апаратних засобів, тому в тренувальному процесі використовуються оптимізатори, для мінімізації значення функції втрат і досягання достатньої точності моделі за прийнятний проміжок часу. Поширеними методами для оптимізації є Adam та RMSprop. Для досягнення балансу між обчислювальними витратами та точністю моделі можна застосовувати міні-пакети. Продуктивність оптимізаторів сильно залежить від розподілу даних, тому можна спробувати різні методи оптимізації. В деяких випадках, може допомогти зміна швидкості навчання нейронної мережі [15].

Згорткові нейронні мережі є актуальним засобом для вирішення задач комп'ютерного зору. На сьогодні, цей засіб вирішує широкий спектр задач. Згортковими нейронними мережами давно послуговуються для вирішення задач класифікації зображень. Якщо порівнювати застосування згорткових нейронних мереж з іншими підходами, то використання згорткових нейронних мереж дозволяє досягти кращої точності класифікації на великих наборах даних. Прорив у задачі класифікації на великих наборах даних зображень відбувся у 2012 році. Тоді була представлена мережа AlexNet, і з її використанням було досягнуто найвищої точності на той момент. Після успішної появи цього рішення ще декілька розробок значно покращили точність класифікації, для цього дослідниками було запропоновано збільшувати глибину мережі або зменшувати розміри фільтра. Виявлення об'єктів на зображенні теж є важливою задачею

комп'ютерного зору. Використання рішень на базі згорткових нейронних мереж для виявлення об'єктів на зображенні було помічено ще навіть до 1990 року. Але тоді через недостатню кількість тренувальних даних і обмежені ресурси комп'ютерів прогрес у виявленні об'єктів на зображенні методами на основі згорткових нейронних мереж був повільним до появи AlexNet. Після появи AlexNet та успіху згорткових нейронних мереж на ImageNet повертається інтерес до виявлення об'єктів на зображенні на основі згорткових нейронних мереж. На початку дослідники поклалися на підходи на основі ковзаючого вікна для обробки вікон, відібраних у певному масштабі та місці. Але зазвичай в зображенні є дуже багато вікон-кандидатів, і такі рішення були обчислювально складними, і було недоцільно їх використовувати у таких наборах даних, як ImageNet, PascalVOC, MSCOCO, тощо. Результат роботи рішень для виявлення об'єктів на зображенні з використанням згорткових нейронних мереж залежить від того, наскільки було подолано вплив поворотів, зміни освітленості, зашумленості зображень та перешкод на ньому, тощо [16].

## 2.5 Методи розпізнавання обличчя

Активне дослідження та розробка методів розпізнавання обличчя триває вже понад тридцять років. Їх розвиток не відстає від інших основних задач комп'ютерного зору, наприклад, класифікації зображень. Зростаюча поширеність систем розпізнавання обличчя сприяє подальшому активному розвитку методів та засобів вирішення цієї задачі.

Процес розпізнавання обличчя можна поділити на три основні кроки: виявлення обличчя на зображенні, вилучення ознак зображення, розпізнавання обличчя. Виявлення обличчя полягає у його локалізації на зображенні. Тобто необхідно визначити, чи є на вхідному зображенні людські обличчя, і якщо є, то визначити координати обмежувальної рамки. Серед труднощів у виявленні обличчя можна виділити різне освітлення на зображеннях, та різні вирази обличчя людини на зображеннях. Вилучення ознак полягає у визначенні особливостей

зображення, зокрема зображень обличчя. Зображення описується вектором ознак, який визначає особливості зображення обличчя. Наприклад, кожне обличчя можна охарактеризувати геометрією його ключових точок, наприклад форма рота, носа, очей, що дозволяє відрізнити його від інших. Широко відомі такі методи для вилучення ознак зображень обличчя: EigenFaces, FisherFaces, аналіз незалежних компонентів, локальні бінарні шаблони, гістограма орієнтованих градієнтів, фільтр Габора, вейвлети Хаара, перетворення Фур'є. В процесі безпосередньо розпізнавання обличчя обробляються ті ознаки, що були вилучені із вхідного зображення, і порівнюються з базою векторів ознак з еталонними обличчями. При ідентифікації особи по зображенню обличчя вхідне зображення порівнюється з базою еталонних векторів ознак, щоб вирахувати найбільш відповідний вхідному зображенню еталон. В задачі ідентифікації особи по зображенню обличчя зазвичай орієнтуються на два протоколи тестування, а саме відкритий і закритий. При відкритому протоколі тестування тренувальний набір даних не може містити зображення для тестування. Щоб визначити точність моделі можуть використовуватись різні метрики, наприклад частота помилково негативних ідентифікацій і частота помилково позитивних ідентифікацій. Частота помилково негативних ідентифікацій вираховує показник помилки, коли позитивні випадки класифіковані як негативні. Частота помилково позитивних ідентифікацій вираховує показник помилки, коли негативні випадки класифіковані як позитивні. Залежно від типу використовуваних методів для вилучення ознак та їх подальшої класифікації, методи розпізнавання обличчя в двовимірному просторі можна розділити на чотири типи: цілісні методи, геометричні методи, методи на основі локальних дескрипторів текстури, та методи, засновані на глибокому навчанні. Концепція цілісних алгоритмів полягає в тому, що будь-який набір зображень обличчя містить надмірні дані, які можна не враховувати. Такі методи формують набір базисних векторів, що представляють простір меншого розміру, який називають підпростором, і зберігається вихідний набір зображень. По набору базисних векторів кожне обличчя можна відновити. Зазвичай, процес класифікації полягає у проектуванні вхідного зображення обличчя та обчислення відстані з

усіма еталонами, що описані в даному підпросторі. EigenFaces та FisherFaces є найбільш відомими методами розпізнавання обличчя цього типу. EigenFaces представляє власні вектори коваріаційної матриці набору зображень обличчя або основні компоненти розподілу. Після цього дослідниками було запропоновано метод FisherFaces, нечутливий до зміни освітлення та виразу обличчя. У ньому кожен піксель зображення вважався точкою у багатовимірному просторі. Зображення лінійно проектувалося в підпростір. Принцип такої проєкції базувався на принципах лінійного дискримінаційного аналізу, який здатен генерувати добре ізольовані класи в маловимірному підпросторі, навіть при різному освітленні та виразах обличчя. Дослідження показали, що в методі FisherFaces нижчий показник помилок, ніж у EigenFaces. Фільтри Габора здатні вилучати ознаки обличчя під певними кутами. Дискретне косинусне перетворення та дискретне вейвлет-перетворення також може вилучати ознаки зображення обличчя. Переважно при використанні цілісних підходів до розпізнавання обличчя, обличчя для розпізнавання із зображення потрібно вирізати вручну. І при цьому, усі зображення обличчя мають бути зіставлені. Якщо зображення обличчя не буде співпадати з іншими, вірогідно це призведе до помилки в розпізнаванні. Ключові точки на обличчі можуть містити корисну інформацію для розпізнавання осіб. Наприклад, лоб і щоки мають одноманітну текстуру без характерної геометрії та візерунків. А очі, ніс та губи мають особливу геометрію, яка відрізняється між людьми. Існує велика кількість різних ключових точок обличчя, але немає рекомендацій, які з них найкращі. Найчастіше використовується: очі, брови, ніс, ніздрі, рот. Геометричні підходи до розпізнавання обличчя вимагають вирівняне зображення обличчя. Якщо обличчя на зображенні не буде вирівняно, вектор ознак буде сформовано некоректно. Зазвичай, вирівнювання зображення обличчя виконується вручну. Підходи до вилучення ознак зображення, що вивчають текстуру, діляться на статистичні та структурні. Вони є стійкими до зміни освітлення. Локальний дескриптор перетворює інформацію, яку містить зображення, у форму, що містить цінні ознаки, які не є чутливими до змін у середовищі, наприклад до змін в освітленні.

Локальні дескриптори аналізують локальні фрагменти зображення, а не усе зображення одразу, як це робиться в глобальних дескрипторах. Локальний бінарний шаблон це відомий в своєму класі дескриптор зображення обличчя. Зображення ділиться на комірки, розподіл ознак локального бінарного шаблону об'єднаний в гістограму, яка використовується як вектор ознак зображення обличчя. Такий оператор, що спроектований в одновимірному просторі, називається одновимірний локальний бінарний шаблон. Вхідне зображення ділиться на комірки однакового розміру. Застосовується вертикальна проекція до кожної комірки в одновимірному просторі. І запропонований дескриптор застосовується до кожної комірки. Вектори ознак, обчислені в кожній комірці, об'єднуються, щоб зробити один вектор. Аналіз головних компонентів використовується для зменшення розмірності даних та збереження лише цінних ознак кожної особи. Оператор локального фазового квантування базується на квантуванні в локальних комірках фази перетворення Фур'є. На цей оператор покладаються при розпізнаванні обличчя, яке розмито на зображенні. Мітки, що розраховані в комірках, використовуються в якості дескриптора, як і в звичайному дескрипторі локальних бінарних шаблонів. Локальне фазове квантування добре показує себе при розмитті зображення та при зміні виразу обличчя чи освітлення на зображенні. Бінарні ознаки статистичного зображення засновані на базі локального бінарного шаблону і локального фазового квантування. Концепція методу полягає в автоматичному формуванні набору фільтрів з певної кількості тренувальних зображень. В локальному бінарному шаблоні і локальному фазовому квантуванні використовуються зроблені вручну фільтри. В методі доступна адаптація та керування розміром дескриптора. Зазначені методи не є чутливими до зміни масштабу та повороту обличчя, але точність розпізнавання знижується при зміні виразу обличчя [17, 18].

Найбільш сучасним підходом до розпізнавання обличчя, що активно розвивається, є розпізнавання обличчя з використанням глибокого навчання. Застосування методів на базі глибокого навчання призвело до підвищення точності моделей не тільки у задачі класифікації зображень чи виявленні об'єктів,

а і у розпізнаванні обличчя.

Такі вагомні архітектури згорткових нейронних мереж, як ResNet, VGGNet, GoogleNet, тощо, широко використовуються як базова модель для розпізнавання обличчя як в оригінальному, так і в модифікованому варіанті. На сьогодні, велика кількість досліджень зосереджена на створенні нових функцій втрат, щоб ознаки були не тільки більш роздільними, але й дискримінаційними. В більшості випадків при вирішенні задачі ідентифікації особи еталонні зображення обличчя неможливо включити в тренувальний набір зображень. Людські обличчя мають схожу загальну геометрію та текстуру, тому знання, отримані на одних прикладах обличчя, можуть бути застосовані на інших обличчях. Враховуючи це, можна зробити висновок, що для підвищення точності розпізнавання треба ввести в тренувальний набір даних якомога більше осіб. Нейронні мережі Deepface і DeepID, що базувались на AlexNet, використовували softmax втрати на основі крос-ентропії. Згодом дослідники дійшли до висновку, що функції втрат softmax недостатньо для вивчення дискримінаційних ознак. Тому почались роботи по створенню нових функції втрат. Раніше важливу роль відігравала функція втрат на основі евклідової відстані, та потім почали набирати популярність функція втрат на основі кутового-косинусного запасу, і нормалізація ознак і ваг. Деякі існуючі функції втрат в дечому схожі одна на одну, але зазвичай нові функції призначені для покращення тренувального процесу мережі. У функціях втрат на базі евклідової відстані зменшується внутрішня дисперсія, та збільшується зовнішня дисперсія. Досить поширеними функціями втрат в штучних нейронних мережах для розпізнавання обличчя є контрастні втрати та триплетні втрати. При використанні контрастних втрат необхідна пара зображень, і позитивні пари “об’єднуються”, а негативні “розсуваються”. Серед труднощів використання цієї функції втрат, можна виділити складність налаштувань параметрів. Триплетні втрати вираховують відносну різницю відстаней між парами зображень, а не абсолютну відстань, як в контрастних. Ця функція втрат була запропонована розробниками FaceNet. Для розрахунку функції необхідні триплети обличчя, тобто позитивний, негативний, і “якірний” зразок. А потім відбувається

мінімізація відстані між “якорем” та позитивним зразком та максимізація відстані між “якорем” та негативним зразком, тобто зображенням іншої особи. Але через особливості даних бувають випадки, коли контрастна та триплетна функції втрачаються з нестабільністю у тренувальному процесі, і дослідження інших, простіших функцій втрачати продовжились. Центрова функція втрачати підходить для зменшення внутрішньої дисперсії, вона визначає центр кожного класу та мінімізує відстані до відповідного центру класу. Дослідники вважали, що зразки треба розділяти суворіше, щоб уникнути неправильної класифікації складних екземплярів. Функція втрачати на основі кутового-косинусного запасу полягає у розділенні особливостей за допомогою більшої кутової чи косинусної відстані. Дослідниками було запропоновано мережу глибокої гіперсферичної згортки SphereNet. Вона використовує гіперсферичну згортку для операції згортки та функцію втрачати на основі кутового запасу. ArcFace і CosFace, відповідно ввели свої функції втрачати на основі кутового-косинусного запасу. Функція втрачати на основі кутового-косинусного запасу здатна досягати високих результатів на чистому наборі даних, але на зашумлених даних вони показують результат гірше, ніж центрова функція втрачати та softmax. Окрім переходу від втрачати softmax до кутово-косинусних функцій втрачати, деякі дослідники займаються нормалізацією характеристик та ваг у функціях втрачати, для підвищення точності моделі. З вдосконаленням архітектур нейронних мереж та такими підходами до тренування, як пакетна нормалізація, нейронні мережі для розпізнавання обличчя поступово вдосконалюються, а їх точність підвищується. DeepFace це перша розробка, яка використовувала згорткову нейронну мережу на дев'ять шарів для розпізнавання обличчя. І вона досягає точності 97,35 відсотків на наборі даних LFW. FaceNet використовувала великий набір даних зображень обличчя та триплетну функцію втрачати для навчання, і досягла точності 99,63 відсотків. Для VGGFace запропонували окремий підхід до збору великого набору зображень обличчя. Базова мережа VGGNet була натренована на цьому наборі даних, і налаштована з функцією триплетних втрачати. Дана розробка досягає точності 98,95 відсотків. SphereFace використовувала архітектуру ResNet як базу, була запропонована

функція втрат, з таким підходом точність досягла 99,42 відсотків. Для зменшення обчислювальних витрат дослідниками було запропоновано легкі мережі. Зокрема, авторами Light-CNN було запропоновано функцію активації max-feature-map, яка зменшує обчислювальні витрати. Останнім часом ідуть дослідження мереж з адаптивною архітектурою, що можуть знайти хороші архітектури штучних нейронних мереж та їх параметрів, так як проектування архітектур вручну людьми займає багато часу та може не дати очікуваний високий результат. Пошук нейронної архітектури є цікавим напрямком, його здобутки вже можуть перевершувати вручну розроблені архітектури в деяких завданнях, наприклад у класифікації або сегментації зображень. Крім пошуку нейронної архітектури, є інші напрямки досліджень адаптивних архітектур. Дослідниками був запропонований підхід до розпізнавання обличчя, що складається з кількох блоків: розпізнавання обличчя, вирівнювання обличчя, тощо. Може бути виконане доповнення навчальних даних, щоб із одного зображення обличчя утворити декілька. Взавши ці кілька зображень в якості вхідних даних, одночасно декілька мереж вилучають та поєднують ознаки різних типів вхідних даних, і вони разом можуть дати кращий результат, ніж окрема мережа. Мережі збираються після вилучення різних частин обличчя, а потім різні частини обличчя відправляються в різні підмережі для вилучення ознак із зображення. Відповідно, якщо поєднати результати різних підмереж, може покращитись загальний результат. На розпізнавання обличчя має вплив повороти і вираз обличчя, освітлення, вік особи. Щоб зменшити цей вплив, дослідниками було запропоновано багатозадачне навчання. У таких мережах основним завданням є розпізнавання, а додатковими завданнями є, зокрема, аналіз пози голови, виразу обличчя, освітлення. Нижчі рівні поширені між усіма цими завданнями, а на вищих рівнях виконуються конкретні завдання на різних підмережах [19].

## 3 РЕАЛІЗАЦІЯ РІШЕННЯ

### 3.1 Структура системи

Програмна реалізація системи виконана на мові програмування Python. Відповідно, це розширює можливості її розгортання на різних апаратних засобах.

Для роботи з штучними нейронними мережами використовується бібліотека Keras, яка є високорівневим API платформи TensorFlow. Також у програмній реалізації використовуються такі бібліотеки: NumPy, SciPy, Matplotlib, Pandas, Scikit-learn.

Запропонована наступна структура системи:

- джерело зображень;
- блок обробки даних;
- блок збереження та відображення результатів.

Програмна частина складається з головного модуля для розпізнавання обличчя. Також розроблено модулі для проведення експериментальних досліджень.

В головному модулі реалізовано такі функції:

- вилучення ознак з зображення або набору зображень;
- зчитування та запис ознак зображення в текстовому форматі;
- розрахунок відстані між векторами ознак;
- розпізнавання обличчя.

В модулях для експериментальних досліджень розроблено функції для вибору базової моделі та її параметрів, тренування та тестування розроблених моделей, оформлення звітів, проведення необхідних розрахунків та їх візуалізації.

Для реалізації системи розпізнавання обличчя, зокрема, було використано такий підхід, як трансферне навчання нейронних мереж. Для вилучення ознак з зображення було використано базові моделі, які попередньо навчені на наборі даних ImageNet. До базової моделі було під'єднано мережу, і вона тренувалася на

наборах даних з обличчями. В останньому шарі мережі для ідентифікації особи використовується функція активації softmax. Якщо ж необхідно провести бінарну класифікацію, тобто аутентифікацію особи, варто використовувати функцію активації sigmoid.

У трансферному навчанні можна використовувати різні архітектури мережі, “розморозувати” для тренування певні шари базової моделі, використовувати різну кількість шарів та їх параметрів, пробувати різні методи регуляризації.

Розпізнавання обличчя включає в себе такі задачі:

- ідентифікація;
- аутентифікація;
- верифікація.

В програмній складовій системи було реалізовано всі ці задачі. Для вилучення ознак із зображення використовувались блоки для вилучення ознак згорткових нейронних мереж, наприклад VGG16, що були натреновані на наборі даних ImageNet. Для тестування рішень було використано декілька наборів даних. Для збору та обробки зображень обличчя було спеціально розроблено програмне забезпечення. Перший набір даних на три класи було зібрано власноруч, використовуючи для цього камеру. Другий і третій набір це вибірка на чотири та двадцять класів відповідно із набору даних LFW.

Для вирішення задач ідентифікації та верифікації було використано два підходи: з використанням інтелектуального класифікатора, та без нього.

### 3.2 Ідентифікація особи без інтелектуального класифікатора

Ідентифікація особи без інтелектуального класифікатора полягає в порівнянні вхідного вектора ознак з набором еталонних векторів. Для отримання векторів ознак із зображення було запропоновано використовувати моделі, що натреновані на наборі даних ImageNet. Зібрати великий набір зображень обличчя для тренування нейронної мережі є непростим завданням, і запропонований підхід може вирішити цю проблему. В наборі даних ImageNet клас “обличчя

людини” не представлений. Запропоноване рішення зменшує часові витрати на розробку архітектури та тренування нейронної мережі, збір тренувальних даних [20].

Для тестування було використано модель VGG16. Модель може бути легко замінена на іншу натреновану модель, що доступна в Keras або загальнодоступних репозиторіях.

Для демонстрації роботи було проведено верифікацію “всіх до всіх”. Результуюча матриця з відстанями між векторами ознак для першого набору даних у вигляді теплової карти наведено на рисунку 3.1.

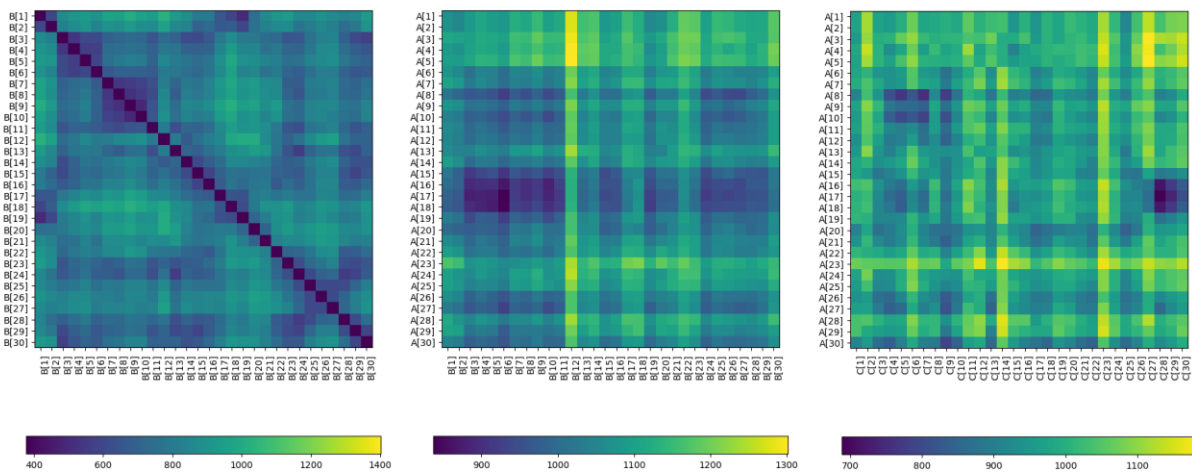


Рисунок 3.1 – Теплова карта порівняння трьох осіб з першого набору даних

Для порівняння векторів використовувалась евклідова відстань, та вона може бути легко замінена на іншу метрику. Розподіл відстаней між векторами однакових та різних людей у першому наборі даних продемонстровано гістограмою на рисунку 3.2.

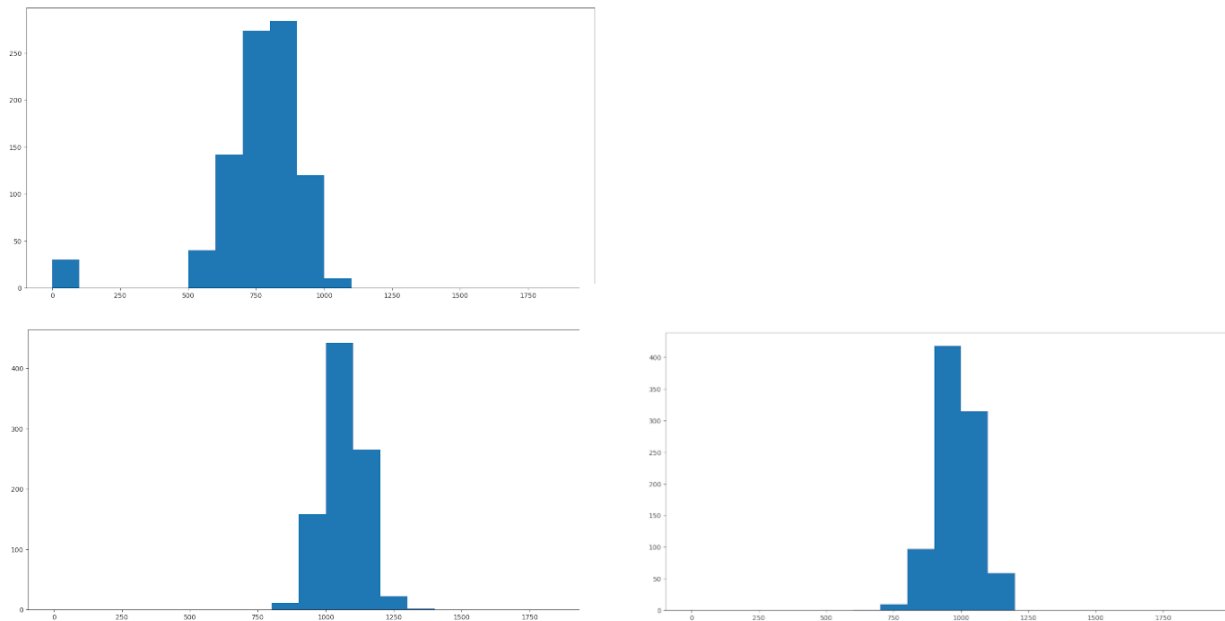


Рисунок 3.2 – Відстані між векторами для першого набору даних

Результуюча матриця з відстанями між векторами ознак для другого набору даних у вигляді теплової карти наведено на рисунку 3.3.

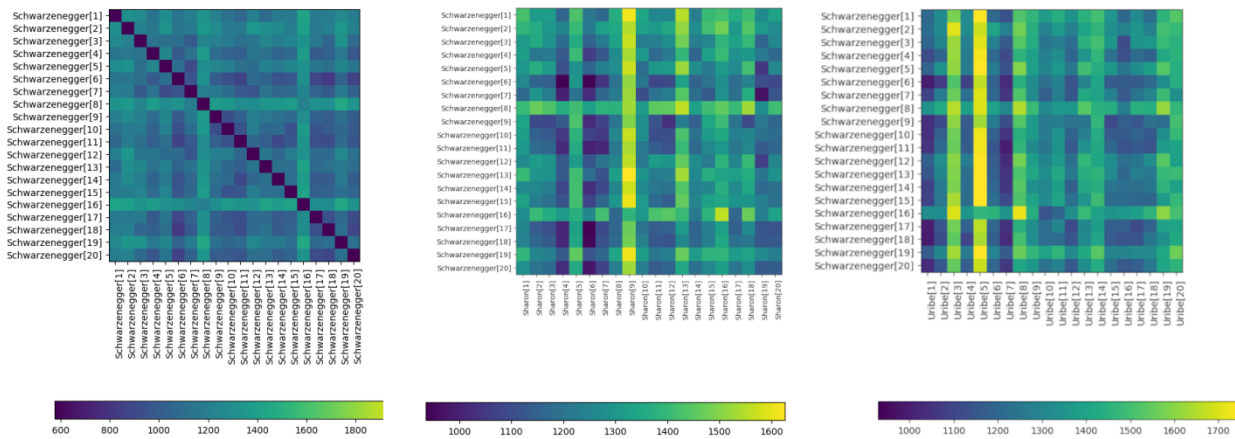


Рисунок 3.3 – Теплова карта порівняння трьох осіб з другого набору даних

Розподіл відстаней для другого набору даних продемонстровано гістограмою на рисунку 3.4.

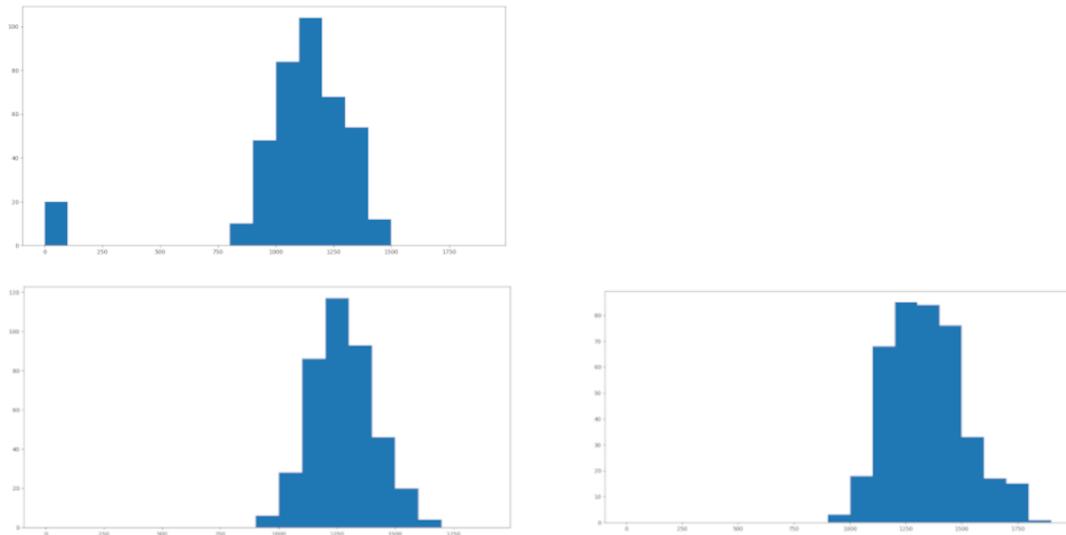


Рисунок 3.4 – Відстані між векторами для другого набору даних

По отриманих теплових картах і гістограмах можна визначити розподіл відстаней, порогову відстань для розпізнавання, та точність розпізнавання.

Згортка нейронна мережа VGG16 має п'ять частин в блоці для вилучення ознак. Блоки на початку розпізнають більш прості особливості зображення, наприклад лінії, тоді як блоки в кінці розпізнають більш складні форми. Було виявлено, що блоки, які ближче до входу, мають більший вектор ознак, відповідно більшу відстань між векторами, більший розрив між відстанню між векторами ознак однакових та різних людей, але це не сприяло підвищенню точності розпізнавання.

Тестування показало, що запропонований підхід має високу точність на першому і другому наборі даних. На “ідеальних” даних точність досягає дуже високого рівня, з ускладненням даних, наприклад оклюзія, значні повороти голови, різниця у віці людини на зображеннях, точність зменшується. На третьому наборі даних точність розпізнавання низька. Це можна пояснити великою кількістю класів у наборі. Велика кількість класів та різноманітність зображень в наборі даних підвищенню точності ніяк не сприяє.

### 3.3 Ідентифікація особи з інтелектуальним класифікатором

Ідентифікація особи з інтелектуальним класифікатором полягає в використанні трансферного навчання. Завдяки такому підходу, за певних умов, можна позбавитись необхідності збирати великі набори даних для тренування моделей для розпізнавання обличчя. Під час трансферного навчання згорткових нейронних мереж як відправна точка використовуються ваги попередньо навчених моделей, наприклад на наборі ImageNet. Попередньо натренована модель також може бути включена як частина нової моделі. Це скорочує час на навчання і може зменшити помилку узагальнення. Задача ідентифікації особи по зображенню обличчя схожа на класифікацію зображень, тому й було запропоновано використовувати трансферне навчання. Такий підхід буде доцільним у системах розпізнавання обличчя, до яких не ставиться вимога мати дуже високу точність. Часові витрати на реалізацію такого рішення менші, ніж при розробці та тренуванні нейронної мережі з нуля [21].

Для тестування було використано натреновану на ImageNet модель VGG16, яка за необхідності може бути легко замінена на іншу натреновану модель. Можна виділити три стратегії застосування трансферного навчання згорткових нейронних мереж, їх наведено на рисунку 3.5.

Перша стратегія полягає в заморожуванні вагів всіх шарів блоку вилучення ознак, класифікатор навчається на оригінальних вихідних даних цього блоку. У другій стратегії окрім класифікатора навчається сусідня частина блоку вилучення ознак. Наприклад, блок вилучення ознак в VGG16 складається з п'яти підблоків згортки та об'єднання. Можна спробувати заморозити ваги лише перших декількох підблоків. За наявності якісного набору даних можна добре адаптувати останні підблоки для вилучення ознак зображення. У третій стратегії виходом блоку вилучення ознак буде не останній підблок згортки та об'єднання, а наприклад, передостанній. Останній підблок, що навчений для вилучення ознак класів ImageNet може видавати менше інформації, на відміну від попередніх підблоків, що вилучають більш загальні ознаки.

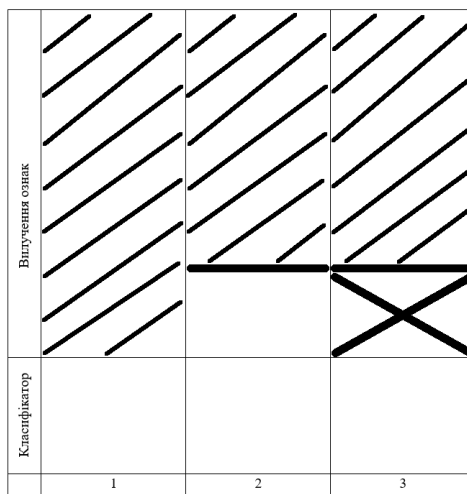


Рисунок 3.5 – Стратегії трансферного навчання

Схеми обраних архітектур нейронних мереж наведено на рисунку 3.6.

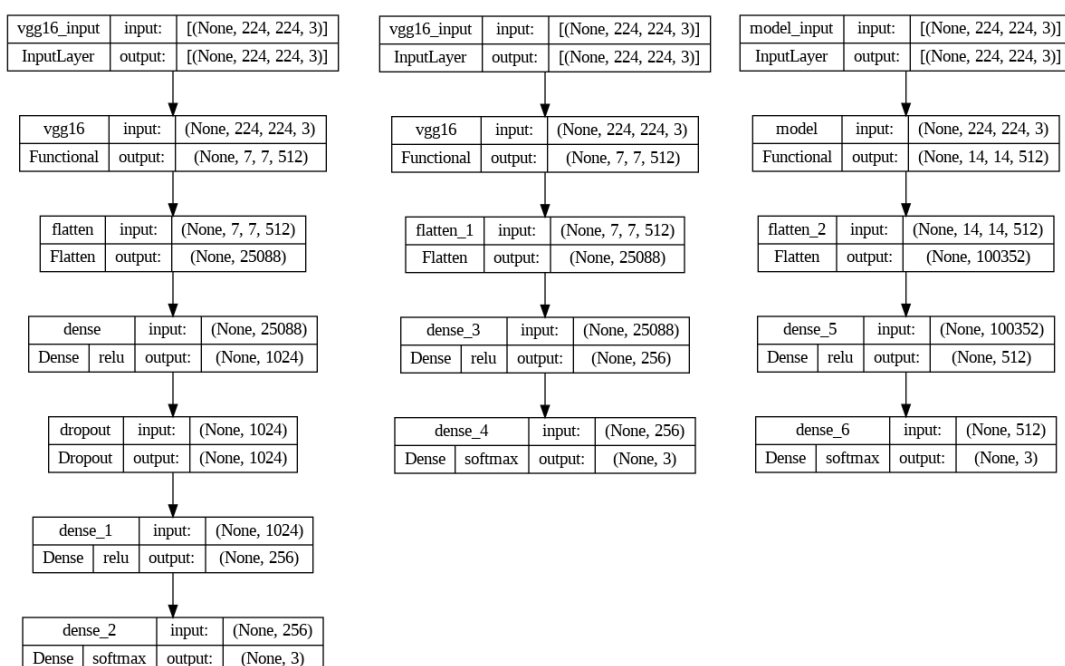


Рисунок 3.6 – Обрані архітектури мереж для ідентифікації особи

Для вилучення ознак зображення використовується VGG16. Багатовимірні вихідні дані цього блоку випрямляються шаром Flatten, і передаються на повнозв'язний шар Dense. Параметри повнозв'язних шарів підбирались

емпіричним шляхом. Також варто зазначити, що використання шару `GlobalAveragePooling2D` після блоку вилучення ознак призводило до значного зниження точності моделі, тому в рішенні використовувався шар `Flatten`.

Графік навчання однієї з мереж на третьому наборі даних для ідентифікації особи наведено на рисунку 3.7.

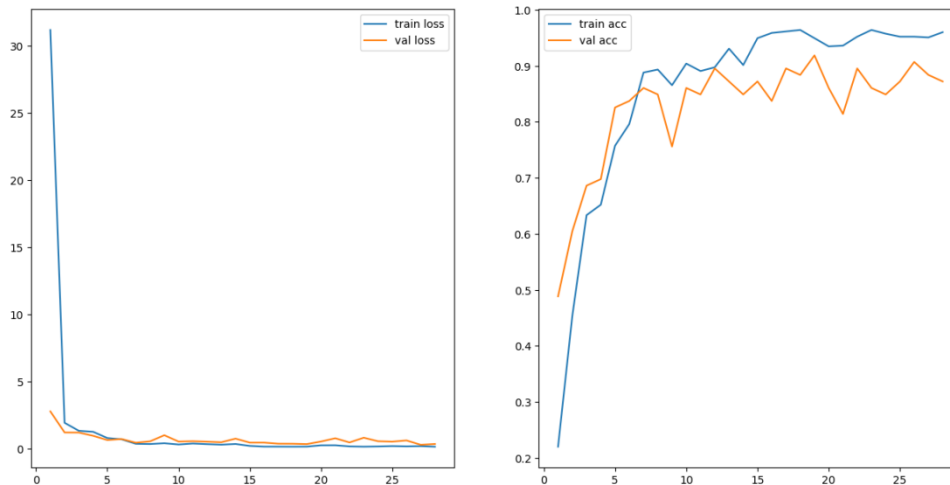


Рисунок 3.7 – Графік навчання моделі

Тестування показало, що запропонований підхід має дуже високу точність на першому наборі даних. Це пов'язано з простотою цього набору. На другому наборі точність теж висока, кожна з трьох розглянутих архітектур дає різні результати. На третьому наборі даних точність розпізнавання трохи нижча, але вона значно вища порівняно з низькою точністю ідентифікації особи без інтелектуального класифікатора на цьому наборі. Однак, все ж велика кількість класів у наборі знижує точність моделі.

### 3.4 Аутентифікація особи з інтелектуальним класифікатором

Для аутентифікації особи з інтелектуальним класифікатором також було використано трансферне навчання. В реалізації рішення було використано

натреновану на ImageNet модель VGG16, яка за необхідності може бути легко замінена на іншу натреновану модель.

Схеми обраних архітектур нейронних мереж наведено на рисунку 3.8.

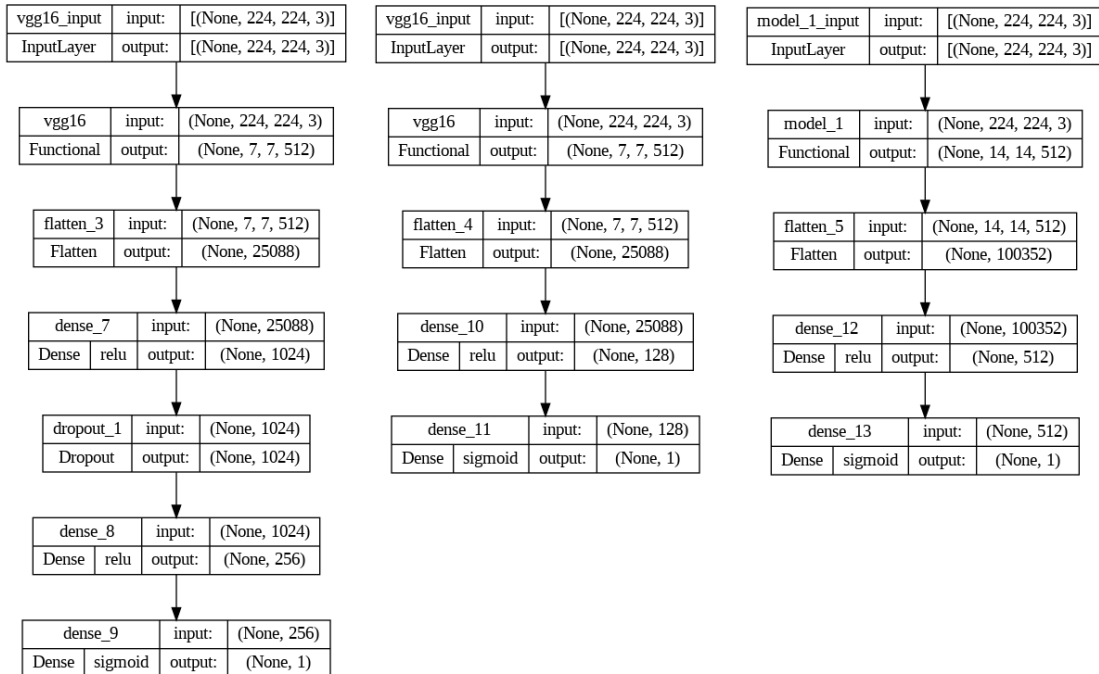


Рисунок 3.8 – Обрані архітектури мереж для аутентифікації особи

Для вилучення ознак зображення використовується VGG16. Багатовимірні вихідні дані цього блоку випрямляються шаром Flatten, і передаються на повнозв'язний шар Dense. В останньому шарі мережі використовується сигмоїдальна функція активації. Для навчання штучної нейронної мережі перший і другий набір даних було адаптовано під задачу аутентифікації.

Графік навчання однієї з мереж для аутентифікації особи на першому наборі даних наведено на рисунку 3.9.

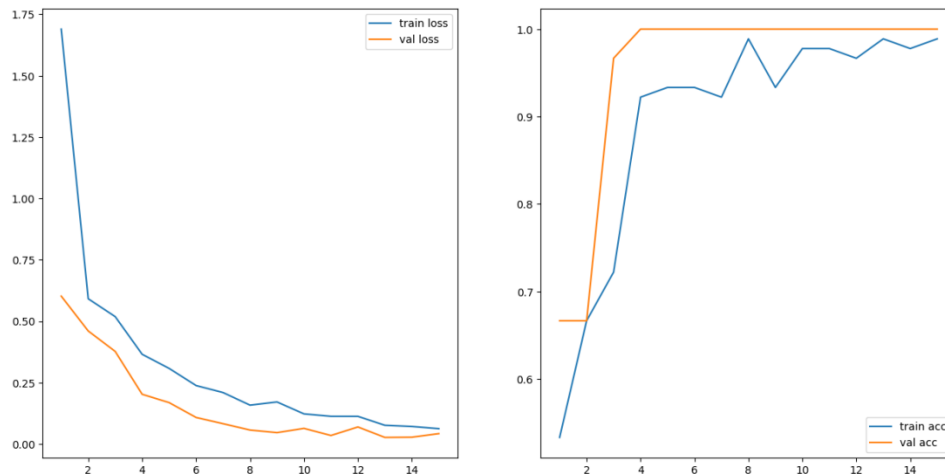


Рисунок 3.9 – Графік навчання моделі для аутентифікації

Тестування показало, що запропонований підхід має високу точність на обох наборах даних. Можна припустити, що це пов'язано з наявністю всього двох класів.

### 3.5 Верифікація особи з інтелектуальним класифікатором

Для верифікації особи з інтелектуальним класифікатором також було використано трансферне навчання. Для вилучення ознак із зображення, яке проводилось завчасно, було використано натреновану на ImageNet модель VGG16, яка за необхідності може бути легко замінена на іншу натреновану модель. Вилучені ознаки зображення подавались у класифікатор. В останньому шарі мережі використовується сигмоїдальна функція активації. Тестування показало, що запропонований підхід має точність нижчу, ніж існуючі рішення, які використовують триплети.

Особливістю запропонованого рішення було використання звичайної бінарної класифікації. Тренувальні дані у першому варіанті виглядали як поелементна різниця векторів ознак однакових та різних людей. У другому варіанті замість розрахунку поелементної різниці відбувалась конкатенація цих векторів.

### 3.6 Верифікація особи без інтелектуального класифікатора

Верифікація особи без інтелектуального класифікатора схожа за принципом з ідентифікацією особи без інтелектуального класифікатора. За допомогою попередньо навченої на наборі ImageNet моделі VGG16 вилучались ознаки зображення, які потім порівнювались між собою. Для розрахунку відстані між векторами було використано евклідову відстань, але її можна замінити на іншу.

По теплових картах або гістограмах розподілу відстаней між векторами можна визначити порогову відстань для розпізнавання. Тестування показало, що запропонований підхід має високу точність на першому наборі даних. У другому наборі даних, з ускладненням даних, точність зменшується.

## 4 РОБОТА СИСТЕМИ

### 4.1 Підготовка до роботи

Щоб система працювала коректно, необхідно дотримуватись рекомендованих системних вимог. Для роботи системи потрібен справний комп'ютер, що підключений до стабільного джерела електроживлення. Для виведення даних рекомендовано мати екран, для введення – мишку і клавіатуру.

В операційній системі має бути встановлено інтерпретатор Python 3, та необхідні бібліотеки. На пристрої зберігання даних комп'ютера мають бути записані файли вхідних даних, необхідні скрипти та файли моделей.

### 4.2 Підготовка вхідних даних

При розпізнаванні обличчя без інтелектуального класифікатора користувачу системи доступний функціонал для вилучення ознак із зображення. Він використовується для обробки вхідного зображення. База еталонних векторів може бути тимчасово сформована в пам'яті під час виклику функції для розпізнавання, але краще зробити базу завчасно, і для цього доступний функціонал для збереження та зчитування еталонних векторів ознак у текстовому форматі з розширенням “json”.

Рекомендовано використовувати зображення з розширенням “jpg”. В програмі потрібно вказати шлях до папки з зображеннями обличчя. В цій папці мають бути підпапки для кожної особи. Така структура папок наведена на рисунку 4.1. Для збору та обробки зображень обличчя може бути використано спеціальне програмне забезпечення.



Рисунок 4.1 – Приклад структури папок

При розпізнаванні обличчя з інтелектуальним класифікатором необхідно мати файл моделі та вхідне зображення. Шляхи до цих файлів вказуються в параметрах в програмі. Файл моделі має розширення “h5”.

#### 4.3 Використання системи

Основні функції системи розташовані в модулі “fr.py”. Ці функції можна використовувати як в головному скрипті, так і в службових модулях. Функція “fe\_image” призначена для вилучення ознак з зображення, а “fe\_images” та “fe\_dir” для вилучення ознак з папки або підпапок з поверненням в формі списку або словника. Ці ознаки можна зберігати в пам’яті комп’ютера під час роботи програми, а можна зберегти в текстовому форматі функцією “fe\_write”, і потім зчитати функцією “fe\_read”.

Для верифікації особи без інтелектуального класифікатора використовується функція “fr\_verification”, яка в якості параметрів приймає функцію відстані між векторами, вектори для порівняння, порогове значення. Результат виконання наведено на рисунку 4.2.

```

150     d = fr_verification(distface, f["f1"][0], f["f1"][3], 800)
151
Run: fr x
2023-12-26 12:36:46.643560: W tensorflow/stream_executor/plat
2023-12-26 12:36:46.643802: I tensorflow/stream_executor/cuda
True
d = 732.8429167100414
t = 800
Process finished with exit code 0

```

Рисунок 4.2 –Верифікація без інтелектуального класифікатора

Для аутентифікації особи з інтелектуальним класифікатором використовується функція “fr\_authentication”, яка в якості параметрів приймає шлях до зображення, файл моделі, порогове значення. Результат виконання наведено на рисунку 4.3.

```

157     model = load_model(r"data\models\model235.h5")
158     d = fr_authentication(r"data\ds\5\test\N\0Angelina_Jolie_0005.jpg", model, 0.6)
Run: fr x
2023-12-26 12:46:18.496431: I tensorflow/core/platform/cpu_feature_guard.cc:193]
To enable them in other operations, rebuild TensorFlow with the appropriate compi
1/1 [=====] - 1s 1s/step
False
d = 0.0
t = 0.6
Process finished with exit code 0

```

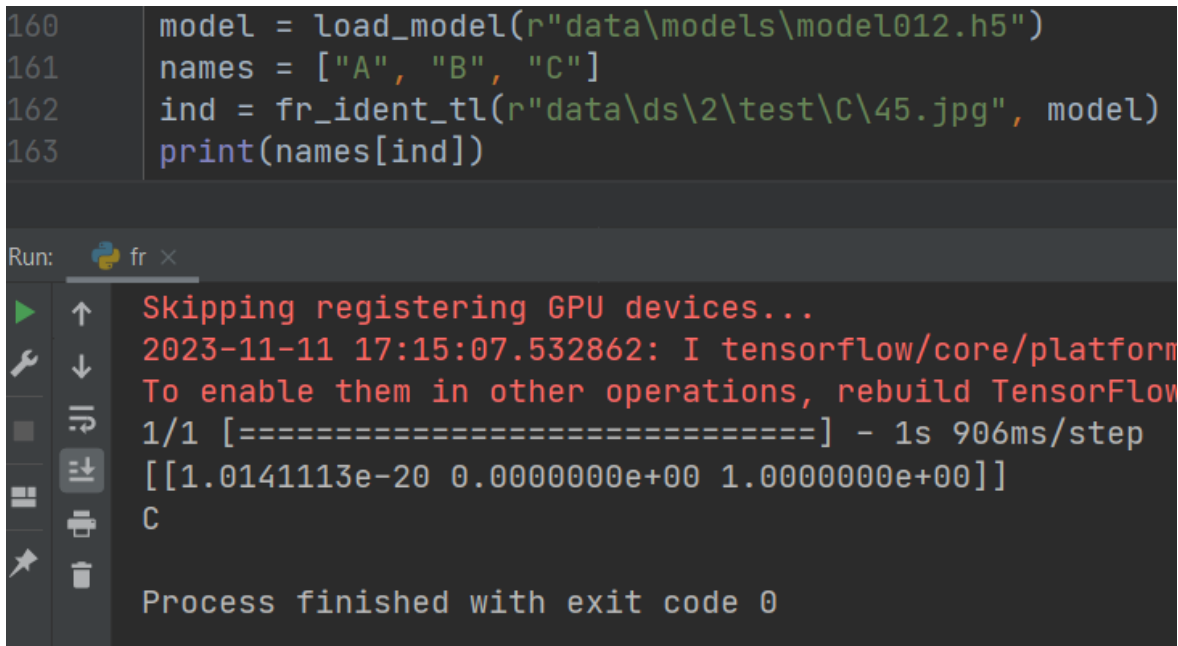
Рисунок 4.3 – Результат аутентифікації

Для ідентифікації особи з інтелектуальним класифікатором використовується функція “fr\_ident\_tl”, яка в якості параметрів приймає шлях до зображення та файл моделі. Результат виконання наведено на рисунку 4.4.

```

160 model = load_model(r"data\models\model012.h5")
161 names = ["A", "B", "C"]
162 ind = fr_ident_tl(r"data\ds\2\test\C\45.jpg", model)
163 print(names[ind])

```



```

Run: fr x
Skipping registering GPU devices...
2023-11-11 17:15:07.532862: I tensorflow/core/platform
To enable them in other operations, rebuild TensorFlow
1/1 [=====] - 1s 906ms/step
[[1.0141113e-20 0.0000000e+00 1.0000000e+00]]
C
Process finished with exit code 0

```

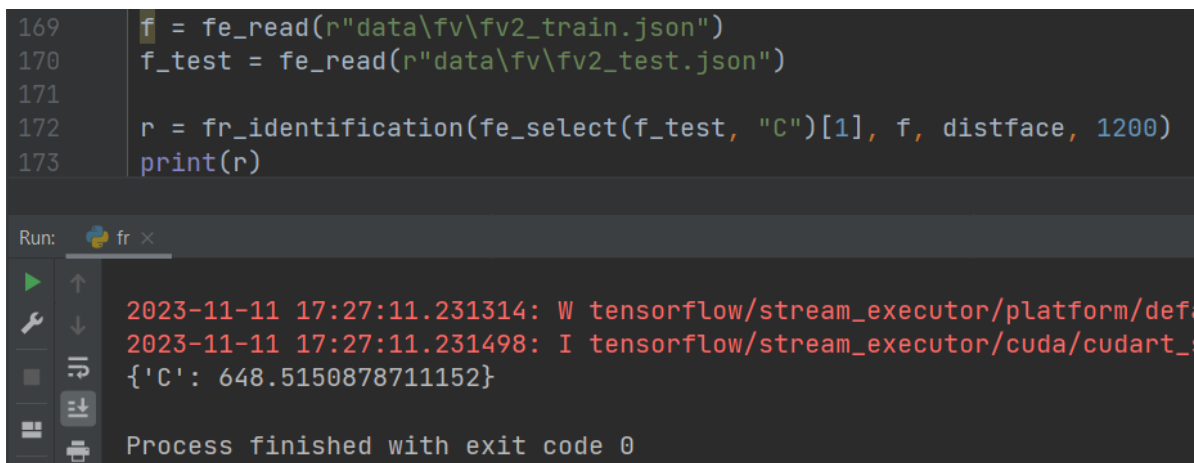
Рисунок 4.4 – Ідентифікація з інтелектуальним класифікатором

Для ідентифікації особи без інтелектуального класифікатора використовується функція “fr\_identification”, яка в якості параметрів приймає вхідний вектор ознак зображення обличчя, словник з еталонними векторами ознак, функцію відстані між векторами, порогове значення. Результат виконання наведено на рисунку 4.5.

```

169 f = fe_read(r"data\fv\fv2_train.json")
170 f_test = fe_read(r"data\fv\fv2_test.json")
171
172 r = fr_identification(fe_select(f_test, "C")[1], f, distface, 1200)
173 print(r)

```



```

Run: fr x
2023-11-11 17:27:11.231314: W tensorflow/stream_executor/platform/def
2023-11-11 17:27:11.231498: I tensorflow/stream_executor/cuda/cudart_
{'C': 648.5150878711152}
Process finished with exit code 0

```

Рисунок 4.5 – Ідентифікація без інтелектуального класифікатора

Реалізовані функції для розпізнавання обличчя можна інтегрувати в системи комп’ютерного зору, що працюють в реальному часі, тобто обробляють

відеопотік з камери. Це можна зробити і при доступі до Tensorflow і Keras в системі, так і без нього. При відсутності цих бібліотек, моделі згорткових нейронних мереж для розпізнавання обличчя можна конвертувати в формат “.onnx”. Тому що використання моделі на базі Tensorflow в OpenCV може працювати некоректно. Модель в форматі “.onnx” була успішно протестована.

Модуль “fr\_lab” містить код для тренування моделей інтелектуального класифікатора. На початку треба вказати шляхи папок з тренувальним, валідаційним та тестовим набором даних, після чого відбувається генерація наборів даних. Далі вказується архітектура моделі, та налаштування процесу тренування. Після успішного тренування можна вивести тренувальні графіки. Також є можливість завантажити модель та протестувати її на тестовому наборі даних з отриманням метрик.

Модуль “fr\_test” у першу чергу призначений для дослідження ідентифікації особи без інтелектуального класифікатора. В модулі реалізовано функції для аналізу матриці відстаней між векторами ознак при верифікації “всіх до всіх”. Є можливість отримати мінімальне ненульове, максимальне, та середнє арифметичне значення відстані. Також можна побудувати гістограму з розподілом значень по діапазонах. Матрицю відстаней можна зручно візуалізувати в форматі теплової карти.

Також було розроблено графічний інтерфейс, для демонстрації можливостей програми. Інтерфейс наведено на рисунку 4.6. Цей модуль знаходиться у файлі “fr\_manager.py”.

В графічному інтерфейсі є ті ж основні можливості, що і в консольному режимі: ідентифікація, аутентифікація, верифікація особи по зображенню обличчя. Розроблений графічний інтерфейс сприяє спрощенню взаємодії з користувачем.

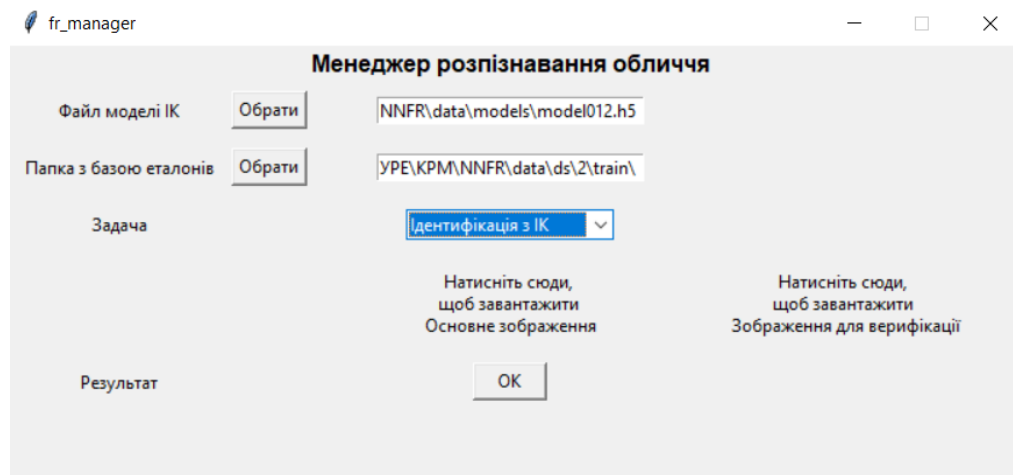


Рисунок 4.6 – Головне вікно

У головному вікні програми є можливість вказати шлях до необхідних файлів і папок, та обрати необхідну задачу. Для того, щоб вказати шлях до файлу моделі інтелектуального класифікатора, необхідно натиснути кнопку “Обрати”, та у системному діалоговому вікні обрати файл, або вручну ввести шлях до файлу у поле для введення. Введення шляху до папки з еталонними зображеннями обличчя, що використовуються як база при ідентифікації без інтелектуального класифікатора, відбувається за схожим принципом. Результат ідентифікації без інтелектуального класифікатора наведено на рисунку 4.7.

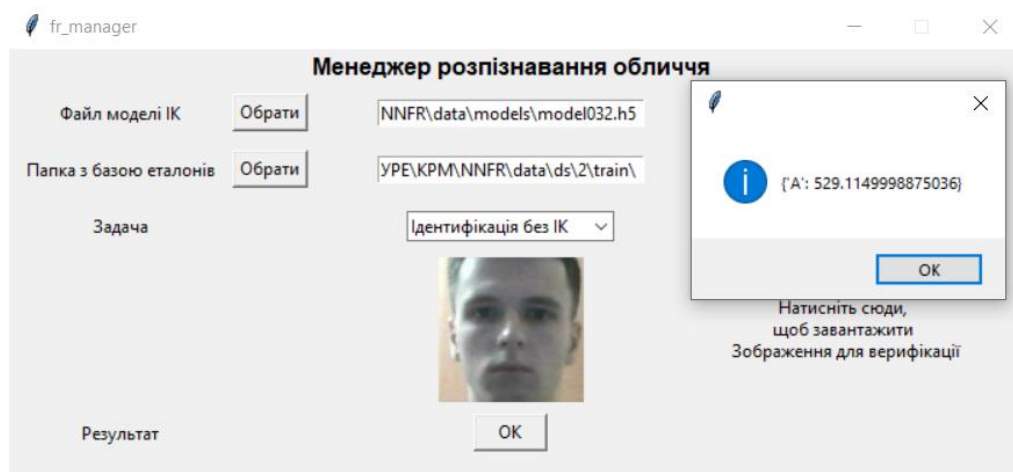


Рисунок 4.7 – Результат ідентифікації без класифікатора

Тип задачі, яку треба вирішити, обирається у випадяючому списку у головному вікні. Потім необхідно вказати вхідне зображення. Якщо виконується

задача верифікації, необхідно також вказати друге зображення. Результат верифікації без інтелектуального класифікатора наведено на рисунку 4.8. Для того, щоб вказати зображення, потрібно натиснути на мітку “Натисніть сюди”, після чого у системному діалоговому вікні вибрати файл, і він відобразиться у вікні програми. Для того, щоб змінити тип задачі або шлях до вхідних даних, перезапускати програму не потрібно.



Рисунок 4.8 – Результат верифікації без класифікатора

Для обробки зображення та відображення результату, треба натиснути кнопку “ОК”. Після цього, буде відображено вікно з повідомленням, яке залежить від обраної задачі.

## ВИСНОВКИ

В ході виконання роботи було проведено аналіз предметної області по фахових джерелах, визначено задачі, які вирішуються комп'ютерним зором, визначено галузі застосування розпізнавання обличчя. Було проведено огляд засобів реалізації систем комп'ютерного зору, засобів збору та обробки зображень обличчя, методів класифікації даних. Було проаналізовано принцип роботи згорткових нейронних мереж та існуючих методів розпізнавання обличчя.

Для отримання векторів ознак із зображення було запропоновано використовувати моделі, що натреновані на наборі даних ImageNet, не дивлячись на те, що в наборі даних ImageNet клас “обличчя людини” не представлений. Запропоноване рішення зменшує часові витрати на розробку архітектури та тренування нейронної мережі, збір тренувальних даних.

Як результат, було розроблено систему розпізнавання обличчя, яка базується на згорткових нейронних мережах. Тестування було проведено успішно, поставлене завдання виконується з достатньою точністю. Недоліком системи є зменшення точності розпізнавання при великій кількості осіб у базі еталонів.

Розроблене рішення може бути використано в системах контролю доступу та відвідуваності, наприклад у домашніх господарствах чи невеликих підприємствах.

Розвиток роботи вбачаємо у підвищенні точності розпізнавання, за рахунок дослідження інших функцій активації та втрат, та вдосконалення вектора ознак шляхом його додаткового аналізу та обробки.

## ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Jafri R., Arabnia H. R. A Survey of Face Recognition Techniques // Journal of Information Processing Systems. – 2009. – Т. 5. – №. 2. – С. 41-68.
2. Top 7 Use Cases for Facial Recognition in 2023 [Веб-сайт]. - URL: <https://www.cyberlink.com/faceme/insights/articles/228/how-to-use-facial-recognition> (дата звернення: 04.10.2023).
3. Feng X. et al. Computer vision algorithms and hardware implementations: A survey // Integration. – 2019. – Т. 69. – С. 309-320.
4. HDF5: Introduction to HDF5 [Веб-сайт]. - URL: [https://hdfgroup.github.io/hdf5/\\_intro\\_h\\_d\\_f5.html](https://hdfgroup.github.io/hdf5/_intro_h_d_f5.html) (дата звернення: 08.10.2023).
5. LMDB - Database of Databases [Веб-сайт]. - URL: <https://dbdb.io/db/lmdb> (дата звернення: 08.10.2023).
6. Viola P., Jones M. J. Rapid object detection using a boosted cascade of simple features // Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2001, Kauai, 08-14 December, 2001
7. Dalal N., Triggs B. Histograms of oriented gradients for human detection // 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). – IEEE, 2005. – Т. 1. – С. 886-893.
8. Mohammed M. G., Melhum A. I. Implementation of HOG feature extraction with tuned parameters for human face detection // International Journal of Machine Learning and Computing. – 2020. – Т. 10. – №. 5. – С. 654-661.
9. Easily Create High Quality Object Detectors with Deep Learning [Веб-сайт]. - URL: <https://blog.dlib.net/2016/10/easily-create-high-quality-object.html> (дата звернення: 08.10.2023).
10. Zhang K. et al. Joint face detection and alignment using Multitask Cascaded Convolutional Networks // IEEE Signal Processing Letters. – 2016. – Т. 23. – №. 10. – С. 1499-1503.
11. Deng J. et al. RetinaFace: Single-Shot Multi-Level Face Localisation in the

Wild // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – IEEE, 2020. – С. 5202-5211.

12. Руденко О. Г., Бодянський Є. В. Штучні нейронні мережі: Навчальний посібник. – Харків: ТОВ «Компанія СМІТ», 2006. – 404 с.

13. Popescu M. C. et al. Multilayer perceptron and neural networks // WSEAS Transactions on Circuits and Systems. – 2009. – Т. 8. – №. 7. – С. 579-588.

14. Understanding Convolutional Neural Network (CNN): A Complete Guide [Веб-сайт]. - URL: <https://learnopencv.com/understanding-convolutional-neural-networks-cnn/> (дата звернення: 22.10.2023).

15. Li Z. et al. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects // IEEE transactions on neural networks and learning systems. – 2022. – Т. 33. – №. 12. – С. 6999-7019.

16. Gu J. et al. Recent advances in convolutional neural networks // Pattern recognition. – 2018. – Т. 77. – С. 354-377.

17. Adjabi I. et al. Past, present, and future of face recognition: A review // Electronics. – 2020. – Т. 9. – №. 8. – С. 1188.

18. Kortli Y. et al. Face recognition systems: A survey // Sensors. – 2020. – Т. 20. – №. 2. – С. 342.

19. Wang M., Deng W. Deep face recognition: A survey // Neurocomputing. – 2021. – Т. 429. – С. 215-244.

20. Куренко В. О. Проблеми використання згорткових нейронних мереж для розпізнавання обличчя. Тренди та перспективи розвитку мультидисциплінарних досліджень : матеріали II Міжнар. студент. наук. конф., м. Хмельницький, 25 листоп. 2022 р. Вінниця, 2022. С. 150–151.

21. Куренко В. О. Особливості використання трансферного навчання нейронних мереж для розпізнавання обличчя. Комп'ютерні інтелектуальні системи та мережі : матеріали XVI Всеукр. науково-практ. WEB конф. аспірантів, студентів та молодих вчен., м. Кривий Ріг, 21–23 берез. 2023 р. Кривий Ріг, 2023. С. 160–162.