

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

Факультет прикладної математики

Кафедра системного програмування і спеціалізованих комп'ютерних систем

«На правах рукопису»
УДК 004.2 : 004.94 : 004.45

До захисту допущено:

Завідувач кафедри

_____ Віталій
РОМАНКЕВИЧ

« ____ » _____ 2024 р.

Магістерська дисертація

на здобуття ступеня магістра

за освітньо-професійною програмою

«Системне програмування і спеціалізовані комп'ютерні системи»

зі спеціальності 123 «Комп'ютерна інженерія»

**на тему: «Модель та метод верифікації користувачів комп'ютерної системи
за зображенням обличчя»**

Виконав:

студент II курсу, групи КВ-22мп
Байдаус Михайло Віталійович _____

Науковий керівник:

Проф.-д.т.н. кафедри СПСКС
Терейковський Ігор Анатолійович _____

Рецензент:

д.т.н., професор кафедри ОТ
Сергієнко Анатолій Михайлович _____

Засвідчую, що у цій магістерській дисертації
немає запозичень з праць інших авторів без
відповідних посилань.

Студент _____

Київ – 2024 року

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

Факультет прикладної математики

Кафедра системного програмування і спеціалізованих комп'ютерних систем

Рівень вищої освіти – другий (магістерський)

Спеціальність – 123 «Комп'ютерна інженерія»

Освітньо-професійна програма «Системне програмування і спеціалізовані комп'ютерні системи»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Віталій РОМАНКЕВИЧ

« ___ » _____ 2024 р.

ЗАВДАННЯ

на магістерську дисертацію студенту

Байдаусу Михайлу Віталійовичу

1. Тема дисертації «Модель та метод верифікації користувачів комп'ютерної системи за зображенням обличчя», науковий керівник дисертації Проф.-д.т.н. кафедри СПСКС Терейковський Ігор Анатолійович, затверджені наказом по університету від «9» 11. 2023 р. №5217-С
2. Термін подання студентом дисертації 27.12.2023
3. Об'єкт дослідження: процес розробки та аналізу верифікації користувача та збору даних за обличчям.
4. Предмет дослідження: методи та алгоритми розпізнавання обличчя людини в режимі реального часу.
5. Перелік завдань, які потрібно розробити: опис предметної області досліджень та порівняльний аналіз існуючих способів верифікації користувачів за зображенням обличчя, розробка власного способу верифікації користувачів.
6. Перелік ілюстративного матеріалу: презентація (кількість аркушів: 20).
7. Перелік публікацій: Терейковський І.А., Байдаус М.В., Крайносвіт А. А. Порівняльний аналіз методів розпізнавання осіб для використання в підсистемі верифікації користувачів комп'ютерної системи. Прикладна математика та комп'ютинг. ПМК, 2023 : шістнадцята наук. конф. магістрантів та аспірантів, 28–30 листопада 2023 р. Аналіз методів розпізнавання облич в бібліотеці OPENCV VI

Міжнародна наукова конференція «Здобутки та досягнення прикладних та фундаментальних наук XXI століття» (08.12.2023; м. Черкаси, Україна).

8. Дата видачі завдання 22.10.2023

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації	Примітка
1	Вивчення літератури за тематикою проєкту	29.09.2023	
2	Розроблення та узгодження завдання магістерської дисертації	03.10.2023	
3	Аналіз існуючих засобів профайлінгу програм	13.10.2023	
4	Підготовка матеріалів першого розділу магістерської дисертації	17.10.2023	
5	Підготовка матеріалів другого розділу магістерської дисертації	21.10.2023	
6	Підготовка матеріалів третього розділу магістерської дисертації	18.11.2023	
7	Підготовка матеріалів четвертого розділу магістерської дисертації	24.11.2023	
8	Підготовка матеріалів доповіді пов'язаної з тематикою магістерської дисертації	26.11.2023	
9	Попередній розгляд магістерської дисертації на кафедрі	22.12.2023	

Студент

Михайло БАЙДАУС

Науковий керівник

Ігор ТЕРЕЙКОВСЬКИЙ

РЕФЕРАТ

Актуальність теми. Захист інформації, один із найважливіших аспектів сучасного світу. В епоху комп'ютерних технологій матеріальні цінності відходять на другий план, хто володіє інформацією, той і головний. Саме тому кожен рік створюється захист із мінімальним ризиком його злому. Існує велика кількість захисту інформації, але в епоху комп'ютерних технологій для більшості з них можна підібрати комбінації і зламати захист за допомогою спеціалізованого коду. У результаті виник захист даних біометричними засобами, завдяки якому для розблокування замість паролів використовуються дані про яку-небудь частину тіла власника. При спробі злому кількість варіантів для підбору буде надзвичайно великою, і результат залежить лише від системи аутентифікації та методу її реалізації. Верифікація особи за допомогою обличчя стала ефективним рішенням для біометричної ідентифікації. У контексті розквіту інформаційних технологій та цифрового суспільства, збір та обробка особистих даних стали неодмінною складовою багатьох сфер нашого життя. Розробники програмних систем зосереджують свою увагу на ефективній верифікації та зборі різноманітних даних за допомогою обличчя людини. Здатність розпізнавання складових обличчя дозволяє вирішити різноманітні завдання, включаючи точне визначення того, що представляє собою людське обличчя. Аналізуючи елементи обличчя, можна визначити настрій, вік, стать та поточний стан особи. Автоматизована ідентифікація емоцій, віку та статі виявляється актуальною у сучасних системах із штучним інтелектом, які постійно розвивають свої можливості. Цей підхід сприяє покращенню та піднесенню на новий рівень взаємодії між людиною та машиною.

Об'єктом дослідження є процес розробки та аналізу верифікації користувача та збору даних за обличчям.

Предметом дослідження є методи та алгоритми розпізнавання обличчя людини в режимі реального часу.

Мета роботи: Мета полягає в розробці та впровадженні моделі та методу верифікації користувачів комп'ютерної системи за зображенням обличчя. Головна

мета полягає в створенні швидкого та легкого у навантаженні інструменту для верифікації та збору даних користувачів, що використовує передові методи обробки та аналізу обличчя. Основні завдання включають в себе розробку ефективної моделі для верифікації осіб, вивчення методів розпізнавання емоцій та інших параметрів за виразом обличчя.

Наукова новизна.

1. Комплексний аналіз емоцій: Розроблена модель враховує не лише основні елементи обличчя для верифікації, але й використовує алгоритми для розпізнавання емоцій. Це дозволяє створювати більш повні та точні профілі користувачів на основі їхнього емоційного стану.
2. Аналіз віку та статі: В роботі використовується метод аналізу обличчя для точного визначення віку та статі користувача. Це розширює можливості системи і забезпечує більш глибокий та повний аналіз користувачів.
3. Додаткові застосування в сучасних технологіях: Розроблений додаток може мати значення для різних сфер, таких як реклама, соціальні мережі, аналітика користувачів та безпека даних. Це відкриває нові перспективи для застосування у сучасному цифровому середовищі.

Практична цінність.

1. Проведено теоретичний аналіз методів та прийомів розпізнавання обличчя, що стало основою для подальшої програмної реалізації та проведення дослідження розробленого методу.
2. Здатність аналізувати емоції дозволяє персоналізувати взаємодію з користувачем.
3. Практична цінність полягає у забезпеченні безпеки, покращенні взаємодії з користувачами та різноманітних можливостях в сферах маркетингу, аналітики та соціальних мереж.

Апробація результатів дисертації. Результати та висновки засобів для профайлінгу програм були представлені та обговорювалися на науковій конференції магістрантів та аспірантів «Прикладна математика та комп'ютинг» ПМК-2023 (Київ, 28-30 листопада 2023 р.) та VI Міжнародна наукова конференція «Здобутки та досягнення прикладних та фундаментальних наук XXI століття» (08.12.2023; м. Черкаси, Україна).

Структура та обсяг роботи. Магістерська дисертація складається з вступу, чотирьох розділів, висновків та додатків. Повний обсяг дисертації – 83 сторінок, у тому числі 80 сторінки основного тексту, 36 рисунка.

У вступі обґрунтовується актуальність роботи, вказуються мета і завдання роботи, визначаються об'єкт і предмет дослідження, обґрунтовується інформаційна база досліджень та подається короткий зміст роботи.

У першому розділі розглядаються види біометричних методів захисту інформації. Досліджуються існуючі підходи до розпізнавання обличчя, емоцій, віку та статі. Розглядаються області застосування та виокремлюються труднощі, що виникають під час розпізнавання. Проводиться аналіз основних понять в сфері розпізнавання, а також здійснюється порівняльний аналіз існуючих систем розпізнавання.

У другому розділі проводиться порівняльний аналіз методів та прийомів розпізнавання обличчя, віку, статі та емоцій людей. Вказані переваги та недоліки методів.

В третьому розділі проводиться аналіз та аргументація вибору засобів для впровадження системної програми.

В четвертому розділі розглядається процес розробки моделей системи, архітектури та алгоритму верифікації користувачів. Описуються мета та функціонал додатка, а також розглядається алгоритм взаємодії з додатком, проводиться перевірка його функціональності та надаються рекомендації щодо його використання.

У висновках представлені результати проведеної роботи.

Ключові слова: розпізнавання обличчя, модель та метод верифікації, верифікація користувачів, збір даних.

ABSTRACT

Actuality of theme. Information security is one of the most crucial aspects of the modern world. In the era of computer technologies, material values take a back seat, and those who possess information are in control. That's why every year, security with minimal risk of breach is developed. There are numerous information protection methods, but in the age of computer technologies, many of them can be cracked. As a result, biometric data protection emerged, allowing the use of owner's body part data instead of passwords for unlocking. In an attempt to breach, the number of possible combinations becomes exceptionally large, and the outcome depends solely on the authentication system and its implementation method. Facial recognition has become an effective solution for biometric identification. In the context of the booming information technologies and the digital society, the collection and processing of personal data have become an essential component of many aspects of our lives. Software developers focus on efficient verification and data collection using facial features. The ability to recognize facial components allows solving various tasks, including accurately determining the nature of a human face. Analyzing facial elements enables the identification of mood, age, gender, and the current state of the individual. Automated identification of emotions, age, and gender becomes relevant in modern artificial intelligence systems that continuously enhance their capabilities. This approach contributes to the improvement and elevation of human-machine interaction to a new level.

The object of the study is the process of developing and analyzing user verification and facial data collection.

The subject of the study includes methods and algorithms for recognizing a person's face in real-time.

Purpose: The goal is to develop and implement a model and method for user verification in a computer system based on facial detection. The main objective is to create a fast and lightweight tool for user verification and data collection that employs advanced facial processing and analysis methods. Key tasks include developing an

efficient model for person verification, exploring methods for emotion recognition, and analyzing other parameters based on facial expressions.

The scientific novelty of the work:

1. Comprehensive Emotion Analysis: The developed model not only considers basic facial elements for verification but also employs algorithms for emotion recognition. This enables the creation of more comprehensive and accurate user profiles based on their emotional state.
2. Age and Gender Analysis: The study utilizes facial analysis methods for precise determination of user age and gender. This expands the system's capabilities and ensures a deeper and more thorough analysis of users.
3. Additional Applications in Modern Technologies: The developed application can be valuable in various domains such as advertising, social networks, user analytics, and data security. This opens up new perspectives for application in the contemporary digital environment.

The practical value:

1. Theoretical analysis of facial recognition methods and techniques has been conducted, serving as the foundation for the subsequent software implementation and investigation of the developed method.
2. The ability to analyze emotions allows for personalized interaction with the user.
3. The practical value lies in ensuring security, enhancing user interactions, and various opportunities in the fields of marketing, analytics, and social networks.

Work approbation. The results and conclusions of the software profiling tools were presented and discussed at the scientific conference of undergraduate and postgraduate students "Applied Mathematics and Computer Science" PMC-2023 (Kyiv, 27-30 November 2023) and the VI International Scientific Conference "Achievements and achievements of applied and fundamental sciences of the XXI century" (08.12.2023; Cherkasy, Ukraine).

Structure and scope of the work The master's thesis consists of an introduction, four chapters, conclusions and appendices. The total volume of the thesis is 83 pages, including 80 pages of the main text, 36 figures.

The introduction justifies the relevance of the work, outlines the goals and objectives, defines the object and subject of the research, substantiates the informational foundation of the study, and provides a brief overview of the work.

The first section discusses various types of biometric information protection methods. Existing approaches to facial, emotional, age, and gender recognition are investigated. Application areas are discussed, and challenges during recognition are identified. An analysis of fundamental concepts in the recognition field is conducted, along with a comparative analysis of existing recognition systems.

The second section presents a comparative analysis of methods and techniques for facial, age, gender, and emotion recognition. The advantages and disadvantages of these methods are specified.

The third section analyzes and justifies the selection of tools for implementing the system program.

The fourth chapter describes the development of system models, system architecture, and user verification algorithms. It outlines the goals and functionality of the application, examines the application's workflow, verifies its functionality, and provides recommendations for its usage.

The conclusion presents the results of the work carried out.

Keywords: facial recognition, model and verification method, user verification, data collection.

ЗМІСТ

РЕФЕРАТ	4
ABSTRACT	8
ЗМІСТ	11
ПЕРЕЛІК СКОРОЧЕНЬ ТА ТЕРМІНІВ	12
ВСТУП.....	14
РОЗДІЛ 1	
ОГЛЯД ПРЕДМЕТНОЇ ОБЛАСТІ ДОСЛІДЖЕНЬ	16
1.1 Історія розвитку комп'ютерного зору	16
1.2 Сфери застосування	19
1.3 Основи методів розпізнавання обличчя	20
1.4 Оцінка ефективності систем розпізнавання	24
1.5 Проблематика розпізнавання	26
1.6 Сучасні системи розпізнавання обличчя та їх порівняльний аналіз	28
1.6.1 Система FaceVACS від компанії Cognitec Systems	28
1.6.2 Система NEC's Face Recognition від компанії NEC	29
1.6.3 Система VeriLook SDK від компанії Neurotechnology	30
1.6.4 Система Microsoft Emotion Recognition	32
1.6.5 Система FaceReader від компанії Noldus	33
1.6.6 Система Betaface від компанії IRIS	34
1.6.7 Система EmoVu від компанії Eyeris Technologies	36
1.7 Висновки до першого розділу	38
РОЗДІЛ 2	
ОСНОВНІ МЕТОДИ ТА АЛГОРИТМИ ОБРОБКИ ЗОБРАЖЕНЬ	40
2.1 Метод Віоли Джонса	40
2.2 Метод головних компонент	43
2.3 Лінійний дискримінантний аналіз	44
2.4 Метод гнучкого порівняння на графах	45
2.5 Згорткова нейронна мережа	46
2.6 Метод ключових точок	48
2.7 Алгоритм Local Binary Patterns Histograms (LBPН)	48
2.8 Бібліотеки	49
2.8.1 OpenCV	49
2.8.2 TensorFlow	51
2.9 Висновки до другого розділу	52

РОЗДІЛ 3**ОБҐРУНТУВАННЯ ОБРАННЯ ЗАСОБІВ РЕАЛІЗАЦІЇ ДОДАТКА 53**

3.1 Обґрунтування мови програмування Python..... 53

3.2 Обґрунтування бібліотеки OpenCV 53

3.3 Обґрунтування методу Віоли - Джонса та алгоритму LBPН 54

3.4 Обґрунтування бібліотеки Keras 56

3.5 Висновки до третього розділу 57

РОЗДІЛ 4**РЕАЛІЗАЦІЯ ДОДАТКУ ВЕРИФІКАЦІЇ ТА ЗБОРУ ДАНИХ****КОРИСТУВАЧА ЗА ОБЛИЧЧЯМ..... 58**

4.1 Пояснення роботи програми 58

4.2 Розробка додатку..... 59

4.3 Огляд додатку..... 65

4.4 Висновки до четвертого розділу..... 77

ВИСНОВКИ 79**СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ 80**

ПЕРЕЛІК СКОРОЧЕНЬ ТА ТЕРМІНІВ

МГК - метод головних компонент

БД - база даних

OpenCV - бібліотека комп'ютерного зору

TensorFlow- бібліотека машинного навчання

LBP - Local Binary Patterns

ROC - Receiver Operating Characteristic

HOG - Гістограма орієнтованих градієнтів

FAR - Коефіцієнт хибного доступу

FRR - Коефіцієнт хибного відмовлення

EER - рівень рівних помилок

ЛДА - Лінійний дискримінантний аналіз

CNN - Згорткова нейронна мережа

ВСТУП

На сьогоднішній день існує багато фізичних характеристик людини, які можна використовувати для її ідентифікації, але широке застосування цих можливостей тривалий час обмежувалося відсутністю засобів, які автоматизували біометричну ідентифікацію. Однак за останні роки з'явилися нові технології, які вирішують це питання, і їх швидке здешевлення робить біометричні системи ідентифікації широкодоступними. Раніше вони використовувалися лише обмеженим колом осіб з особливими повноваженнями. У цій роботі, під час аналізу сучасного стану біометричних технологій, використовується термін "ідентифікація" як складова частина сталого вислову "biometric identification". Це термін описує біометричну перевірку в широкому контексті, включаючи також аутентифікацію.

За визначенням, розпізнавання обличчя - це технологія, здатна ідентифікувати або верифікувати суб'єкта за зображенням, відео або будь-яким аудіовізуальним елементом його обличчя. Зазвичай така ідентифікація використовується для доступу до програми, системи або сервісу і працює як сканер обличчя. Це метод біометричної ідентифікації, який використовує виміри тіла, в даному випадку обличчя та голови, для підтвердження особи людини за допомогою біометричного шаблону та даних її обличчя. Технологія збирає набір унікальних біометричних даних кожної людини, пов'язаних з її обличчям і виразом обличчя, для ідентифікації, верифікації та/або автентифікації особи. Процедура ідентифікації за обличчям вимагає будь-якого пристрою з цифровою фотографічною технологією для створення та отримання зображень і даних, необхідних для створення та запису біометричного шаблону обличчя особи, яку потрібно ідентифікувати. На відміну від інших рішень для ідентифікації, таких як паролі, верифікація електронною поштою, селфі або зображення, або ідентифікація за відбитками пальців, біометричне розпізнавання обличчя використовує унікальні математичні та динамічні

шаблони, які працюють як сканер обличчя, що робить цей метод одним з найбезпечніших і найефективніших.

Дана магістерська робота присвячена створенню ефективного та легкого інструменту для верифікації та збору даних користувачів, який використовує передові методи обробки та аналізу обличчя.

РОЗДІЛ 1

ОГЛЯД ПРЕДМЕТНОЇ ОБЛАСТІ ДОСЛІДЖЕНЬ

1.1 Історія розвитку комп'ютерного зору

Комп'ютерне зорове сприйняття виникло в середині минулого століття як окрема галузь під час дослідження штучного інтелекту. У той час, коли обговорювалося можливість створення "мислячих" машин та штучного інтелекту, розпізнавання образів призвело до виникнення комп'ютерного зорового сприйняття. Визначаючи його математичну специфіку, зрозуміло, що зображення (фотографія) є двовимірною проекцією тривимірного світу, що потребує додаткової інформації для правильної і повної інтерпретації.

Порівнюючи комп'ютерне і людське зорове сприйняття, визначено кілька методів отримання додаткової інформації про об'єкти. Ці методи включають активний огляд, використання стереопар, апріорних моделей побаченого, змістовних, семантичних і геометричних. Однак розглядання цих методів було переважно теоретичним, оскільки відповідна технічна база ще не існувала, і математичні моделі виявилися надто складними на той момент.

Використання фотозображень для вимірювань почалося задовго до поширення комп'ютерної техніки, і цей метод, відомий як фотограмметрія, рекомендувався та використовувався для аерофотозйомки та космічної навігації. Однак для вирішення інших практичних завдань в різних галузях використовувалися інші методи отримання додаткової інформації, такі як штучна організація середовища, спеціальне підсвічування, сигнальні вогні і маятники для маркування деталей чи триангуляційні вежі під час аерофотозйомки, а також стандартне розташування предметів.

У 70-х роках минулого століття були визначені завдання та отримані перші результати в галузі комп'ютерного зору. На першому етапі вирішення цих завдань було надзвичайно складним, зокрема, важким було введення зображень в обмежену пам'ять електронно-обчислювальної машини, що мала обмежений обсяг (64 Кб для БЭСМ-6). Виведення зображень на екран було

неможливим до винайдення системи ГРАФОР, яка дозволяє формувати на папері контурні малюнки або графіки. Ситуація була трохи кращою для ЕОМ, які могли виводити зображення на дисплей (М-6000), проте навіть на них виводились лише контурні або точкові зображення. Обробка цих зображень та моделювання тривимірних об'єктів вимагало використання саморобного програмного забезпечення, що робило комп'ютерне моделювання і технічне лягання на той час досить спрощеними, хоча й вимагає значних зусиль.

У 1990-2000 роках відбувся значний прогрес у галузі комп'ютерного зору, спричинений введенням нових потужних комп'ютерів, а саме суперкомп'ютерів. Проте в цей період не спостерігалось суттєвих нововведень у сфері "перевороту" в усталеній дисципліні; розвиток був обумовлений постійним зростанням інтересу до комп'ютерного зору і, відповідно, збільшенням спроб використання методів з інших галузей, таких як математична статистика, штучний інтелект, фотограмметрія та інші, у сфері комп'ютерного зору.

Важливо відзначити, що внаслідок цього прориву першим видимим результатом став прогрес у розробці методів опису зображень. Застосування згортки зображень через використання вейвлетів і фільтрів Габора дозволяє стисло описувати зображення та зменшувати їхню розмірність. Це відкрило нові можливості, зокрема у напрямку пошуку в базах даних зображень. Зокрема, оскільки тепер описи цікавих областей зображення стали більш вдосконаленими, наприклад, завдяки вдосконаленому дескриптору та його аналогам, тепер можливо реалізувати щільну тривимірну конструкцію. Точність цієї конструкції приблизно рівна точності моделей, отриманих за допомогою активного лазерного сканування.

Вже на початку 21-го століття досягнення в галузі фотограметрії дозволяли практично в режимі реального часу створювати тривимірні моделі в області медицини. Також широко використовуються тривимірні моделі для розробки і конструювання комп'ютерних моделей міст. Один із застосунків

цього - додаток PhotoSynth, який обробляє численні фотографії пам'яток і намагається створити їхні тривимірні моделі. Навіть якщо математичні моделі цих об'єктів можуть не бути абсолютно точними, візуально отримані тривимірні об'єкти передають набагато більше інформації, ніж окремі фотографії, і створюють враження "ефекту присутності". Це підтверджує, що у цьому контексті синтетичний світ є ближчим, ніж може здаватися на перший погляд.

Протягом довгого часу обробка відео мало невелике значення через обмежену потужність комп'ютерів на початку 2000-х років. Однак ця область стала активно розвиватися в межах комп'ютерного зору. Використання комп'ютерів у системах відеоспостереження та біометрії стало розповсюдженим, а також широко впроваджується програмне забезпечення для розпізнавання обличчя. На сьогоднішній день проводяться дослідження з аналізу розпізнавання поведінки та дій на основі відеоінформації. Об'єднання комп'ютерного зору та комп'ютерної графіки веде до створення "розширеної реальності", яка поєднує відчуття реального світу, зокрема зображень, з доданими до них "уявними" об'єктами допоміжно-інформаційного характеру.

Наприклад, уявіть тривимірну модель міста, такого як Помпеї, яка створена на основі зображень і фотографій. До цієї моделі додані графічні деталі, такі як віртуальні дерева, мости та персонажі. Таким чином, розширена реальність дозволяє вводити штучні елементи для моделювання реалістичних об'єктів. Важливо відзначити, що розширена реальність відрізняється від віртуальної тим, що вона не створює зовсім нового світу. Цей підхід можна використовувати як для розваг, так і в науці, особливо там, де потрібно об'єднати інформацію з різних джерел, уніфікувати її та представити у компактній формі.

В даний час провідні пошукові системи, такі як Google та інші, дозволяють користувачам не лише переглядати карти міст, але й віртуально "прогулятися" по їхніх кварталах, отримуючи враження, ніби вони є місцевими

мешканцями, що спостерігають за навколишнім середовищем. За допомогою доповненої реальності в майбутньому можливо отримувати інформацію про будівлі, включаючи їх візуалізацію з фасадів, і створювати вид з вікна будь-якого поверху.

В сучасній добі теорія комп'ютерного зору, яка є окремим розділом кібернетики, належно сформована і базується на великому обсязі практичних та наукових знань. Регулярно проводяться конференції та симпозіуми з цієї тематики, розробляються різноманітні програмні та апаратні комплекси, видавництва видають численні книги, а також публікуються наукові статті. Існують наукові організації, такі як SPIE (Міжнародне товариство оптичного інжинірингу), IEEE Computer Society (Інститут інженерів електротехніки та електроніки) та інші, які активно вивчають та підтримують дослідження у сфері сучасних інформаційних технологій, зокрема в області комп'ютерного зору.

1.2 Сфери застосування

У нашому сучасному світі використання технології розпізнавання обличчя набуває значущості в різних сферах життя, впливаючи на нові можливості та підходи до вирішення різноманітних завдань. Здатність аналізувати унікальні риси обличчя робить цю технологію ефективною та перспективною.

Однією з основних областей використання розпізнавання обличчя є сектор безпеки та криміналістики. Застосування цієї технології значно полегшує ідентифікацію та відстеження осіб, що є критичним для проведення розслідувань злочинів та забезпечення громадської безпеки.

У фінансовому секторі системи розпізнавання обличчя використовуються для забезпечення безпеки та аутентифікації користувачів в банківських системах та платіжних системах. Це допомагає у запобіганні фінансовим шахрайствам та недозволеним доступу до особистих рахунків.

В галузі освіти системи розпізнавання обличчя використовуються для автоматизації реєстрації та контролю навчального процесу. Розпізнавання статі та віку може сприяти індивідуалізації навчання та адаптації програм до потреб студентів.

Технологія розпізнавання емоцій за виразом обличчя виявляє застосування у виявленні шахраїв в банках, казино та фінансових установах. Казино в Лас-Вегасі вперше використовували такі системи для виявлення шахраїв за гральними столами.

Аналіз емоцій через розпізнавання обличчя стає ефективним інструментом для рекламодавців та маркетологів. Рекламні кампанії можуть бути налаштовані на основі реакцій цільової аудиторії, що забезпечує більш ефективний вплив та взаємодію.

Для оптимізації процесів навчання також можливо використовувати системи розпізнавання емоцій. Застосування цих систем дозволяє адаптувати навчальний процес, роблячи його більш ефективним, цікавим та корисним для учнів.

1.3 Основи методів розпізнавання обличчя

Технологія розпізнавання обличчя - це система, яка може визначити або перевірити особу на цифровому зображенні чи відеокадрі.

Методи розпізнавання обличчя, віку, статі та емоцій людей:

Розпізнавання обличчя:

- Методи згорткових нейронних мереж: Застосовують глибокі нейронні мережі для визначення особливостей обличчя та ідентифікації осіб.
- Методи з використанням особливостей: Визначають ключові точки обличчя (очі, ніс, рот) для подальшого аналізу.
- Метод Віюли-Джонса [1]: Використовує каскади класифікаторів та шаблонні зображення для виявлення основних ознак обличчя, таких як очі, ніс і рот. Дуже швидкий та ефективний для реального часу.

- Гістограма орієнтованих градієнтів (HOG) [2]: Аналізує градієнти зображення для визначення напрямку та інтенсивності пікселів. Використовується для виявлення контурів та форм обличчя.
- Методи використання глибинних камер: Використовують інформацію про глибину для створення точних тривимірних моделей обличчя. Забезпечують високу точність в аналізі геометричних особливостей.
- Ансамблеві методи: Комбінують декілька методів для підвищення точності розпізнавання. Використовуються для покращення результатів та зменшення помилок.

Розпізнавання віку:

- Статистичні моделі: Використовують статистику для визначення типових змін у фізичних ознаках з віком.

Розпізнавання статі:

- Методи на основі візуальних ознак: Використовують форму обличчя, волосся та інші фактори для визначення статі.

Розпізнавання емоцій:

- Аналіз виразу обличчя: Використовує зміни у виразі для визначення емоцій, використовуючи нейромережі та аналітичні методи.
- Використання глибоких нейронних мереж: Тренується на великій кількості даних для точного розпізнавання емоцій.

Для покращення ефективності розпізнавання та якості виділення обличчя, на етапі перед обробкою вхідних зображень в системах розпізнавання застосовується додатковий процес. Після виявлення обличчя на кадрі та визначення його параметрів, таких як розмір, поза, положення голови та ключові риси, зазвичай проводиться нормалізація. Цей процес включає кодування, масштабування та перетворення обличчя до горизонтального положення, де лінія, що з'єднує центри очей, стає горизонтальною. Також на етапі попередньої обробки застосовуються різноманітні фільтри для зменшення рівня шуму, такі як медіанні фільтри та фільтри Гаусса[3].

Методи попередньої обробки можуть бути різноманітними і залежать від конкретних завдань досліджень, включаючи використання нелінійних фільтрів для видалення імпульсного шуму та застосування фільтра Гаусса для розмиття дрібних деталей на зображенні. Медіанні фільтри також можуть використовуватися для збереження контурів та пригнічення імпульсних шумів.

Методи екстракція ознак:

1. Габорові фільтри: Визначають текстурні ознаки, використовуючи габорові хвилі.
2. Кутові ознаки: Оприлюднюють точки відмінностей, що допомагають визначити структуру зображення.

Методи трансформації:

Вейвлет-трансформація: Аналізує різні рівні деталізації зображення для виявлення ключових особливостей.

Класифікація:

Машинне навчання: Використовує класифікаційні алгоритми для розпізнавання та категоризації об'єктів на зображеннях.

Системи розпізнавання обличчя у форматі 2D опрацьовують двовимірні зображення та визначають особливості обличчя, аналізуючи його текстуру та області з високою контрастністю. Такий підхід призводить до ускладнень у випадку недостатнього освітлення чи зміненого положення обличчя.

Натомість системи розпізнавання у форматі 3D є більш стійкими до таких змін, оскільки вони створюють модель обличчя, враховуючи його структурні особливості та череп. Незважаючи на ці переваги, системи 3D поки що не отримали широкого застосування через обмежену можливість обслуговування великої кількості користувачів у режимі ідентифікації та низьку швидкість роботи. Додатково технології 3D-розпізнавання вимагають значних обчислювальних ресурсів, і вартість відповідного обладнання суттєво перевищує витрати на системи 2D.

У сучасних системах відеоспостереження широко застосовуються кілька ефективних алгоритмів розпізнавання, які ґрунтуються або на аналізі значень пікселів, або на визначенні характерних точок.

Основний алгоритм розпізнавання складається із таких етапів (Рис.1)



Рис. 1. Загальний процес обробки зображення обличчя при розпізнаванні

Отримання зображення може відбуватися шляхом сканування фотографії, завантаження зображень або відео або просто підключення до камери. Попередня обробка зображення включає початковий етап обробки для усунення шумів і колірних перетворень, таких як недоліки отримання зображення, погане освітлення, неправильне положення об'єкта чи перешкоди при передачі інформації. Виділення особи на зображенні включає визначення області, де розташована людина, та ігнорування навколишніх об'єктів. Виділення ознак особи передбачає отримання релевантних даних зі зразка і поділяється на цілісні та локальні категорії. Також на етапі виділення ключових точок на обличчі визначаються такі аспекти, як губи, ніс, очі, брови та вуха. На завершальному етапі класифікації дані з ключових точок порівнюються з базою даних облич, емоцій аналізується, і система визначає емоційний стан людини на зображенні.

1.4 Оцінка ефективності систем розпізнавання

Спеціалізовані та любительські системи розпізнавання відрізняються за рядом параметрів. Любительські системи привабливіші, хоча й дорожчі, і вимагають менше витрат. Через те, що їх помилки мають зазвичай невеликі наслідки, вартість таких систем значно менша, ніж вартість професійних, до яких ставлять серйозні вимоги.

Наприклад, технологія розпізнавання обличчя у соціальних мережах є прикладом любительської системи, тоді як професійна система призначена для пошуку розшукуваних осіб. Перша працює з фотографіями, де люди зазвичай дивляться в камеру, а друга - з фотографіями людей у громадських місцях, де вони можуть приховувати своє обличчя від камер відеоспостереження. Важливо відзначити, що помилка в першій системі несприятлива, але не критична, тоді як помилки в другій можуть мати серйозні наслідки, оскільки пропущена особа може виявитися злочинцем.

Отже, вибір системи повинен враховувати не лише ціль, але і критерії ефективності, такі як придатність до роботи в різних умовах та пороги допустимих помилок. Для оцінки ефективності систем розпізнавання використовують параметри, такі як Коефіцієнт хибного доступу (FAR) та Коефіцієнт хибного відмовлення (FRR).

Коефіцієнт хибного відмовлення вказує на ймовірність того, що система неправильно відкидає обличчя, тобто не впізнає особу в базі даних. Чим менший цей коефіцієнт, тим вища точність розпізнавання. Навпаки, коефіцієнт хибного доступу вказує на ймовірність того, що система помилково визнає особу правильною.

Також існує коефіцієнт EER (рівень рівних помилок), який дозволяє порівнювати біометричні системи. Цей коефіцієнт визначає той рівень, при якому обидві помилки (помилка прийняття і помилка відхилення) є еквівалентними. Чим нижчий коефіцієнт EER, тим вища точність біометричної системи.

На Рис. 2 відображені взаємозв'язки характеристик FAR, FRR та EER.[4]

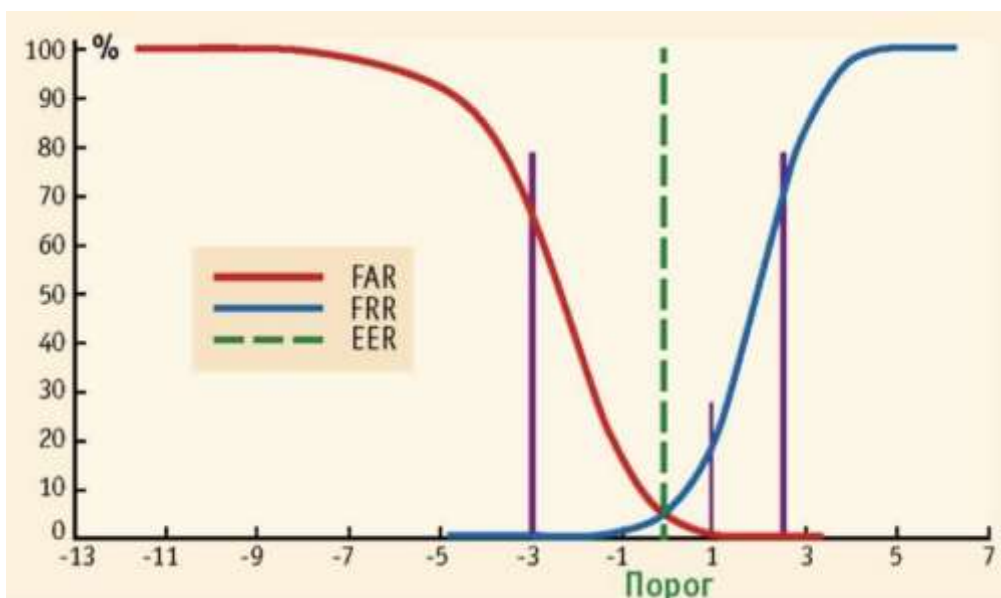


Рис.2 Характеристики біометричних систем

Графік ROC (Receiver Operating Characteristic) використовується для порівняння систем та візуалізації взаємозв'язків між помилками. Ця крива показує, наскільки подібний зразок повинен бути до стандарту, щоб система визнала його збігом. Під час налаштування порогів помилок важливо усвідомлювати, що значення FAR і FRR є взаємно оберненими: зменшення одного з них призводить до зменшення помилкових невідповідностей, але може збільшити кількість помилкових прийняття.

ROC-крива дозволяє аналізувати та оптимізувати точність біометричних систем. Це інструмент, що допомагає збалансувати значення коефіцієнтів FAR і FRR для досягнення оптимальної ефективності системи розпізнавання. При підборі порогів помилок важливо узгоджувати їх так, щоб забезпечити оптимальний баланс між високою точністю та мінімізацією помилок.

1.5 Проблематика розпізнавання

Однією з головних проблем у цій сфері є забезпечення високої точності та надійності систем розпізнавання обличчя. Незважаючи на значний прогрес у

цьому напрямі, існують виклики, пов'язані з різноманітністю осіб, освітленням, а також можливістю обхідних методів, які ускладнюють завдання розпізнавання. Можуть виникати труднощі, пов'язані з положенням голови та наявністю різних аксесуарів на обличчі, таких як окуляри, маски, пірсинг, борода та головні убори, як це зображено на Рис. 3



Рис.3 Приклад проблеми розпізнавання у вигляді затулення ключових елементів

Додатковою проблемою є етичний аспект використання технологій розпізнавання обличчя. Виникають питання щодо конфіденційності та захисту особистої інформації громадян, а також можливості зловживання такими технологіями у владних чи комерційних цілях.

Крім того, існують технічні виклики, пов'язані з обробкою великої кількості даних при високій швидкості, зокрема в області відеоспостереження та розпізнавання в реальному часі. Це вимагає вдосконалення алгоритмів та обчислювальних ресурсів.

Також під час тренування алгоритмів для розпізнавання ключових елементів або рис обличчя передусім використовуються бази даних, в яких зображені людські емоції у вираженні з найвиразнішими виявами, особливо в

піковій формі. У піковій формі алгоритм швидко визначає ключові елементи обличчя та виявляє виражену на зображенні емоцію. Однак такий метод навчання може призводити до труднощів при аналізі менш виражених або неповних емоцій, як показано на рис. 4. Ця ситуація суттєво збільшує ймовірність невірної розпізнавання емоційного стану людини.

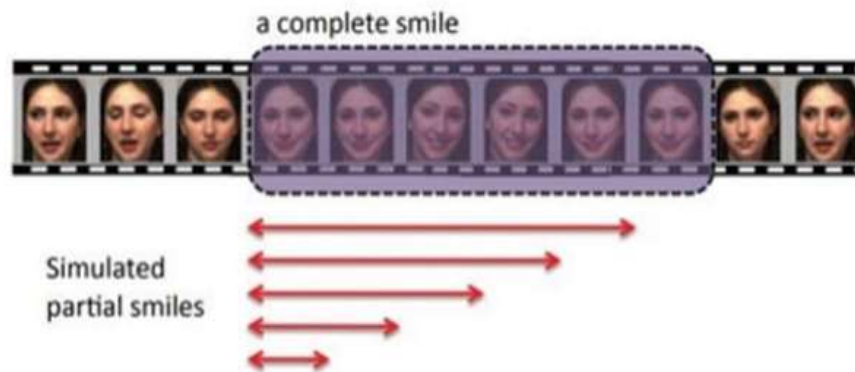


Рис.4 Приклад неповних та пікових форм посмішки

Подолання зазначених проблем в сфері розпізнавання обличчя має потенціал перенести цю технологію на абсолютно новий рівень і розширити її застосування в різноманітних галузях, включаючи безпеку, логістику, медицину та соціальні служби. Активний розвиток сучасних технологій свідчить про постійні зусилля у вирішенні цих викликів, спрямованих на отримання оптимальних результатів у сфері розпізнавання обличчя людини. Нові можливості та вдосконалення технологічних рішень відкривають перспективи для ефективного впровадження цієї технології в різноманітні сфери життя, сприяючи подальшій інноваційній та функціональній еволюції розпізнавання обличчя.

1.6 Сучасні системи розпізнавання обличчя та їх порівняльний аналіз

На сьогоднішній день у світі існує множина різноманітних систем розпізнавання обличчя. Перспективними та інноваційними прикладами сучасних розпізнавальних систем є:

1.6.1 Система FaceVACS від компанії Cognitec Systems

FaceVACS-VideoScan[5] - це програмне забезпечення для розпізнавання обличчя в реальному часі, яке легко використовувати та налаштовувати. Це продукт компанії Cognitec Systems.

Система FaceVACS-VideoScan включає в себе декілька ключових компонентів: відеосервер для управління потоками відео, сервер відео сканування, який координує всі компоненти та виконує основні біометричні операції, обчислювальний вузол для розподілу обчислювального навантаження, користувацький інтерфейс, диспетчер сигналів для отримання сповіщень та обслуговування мобільних пристроїв, операційна база даних і комплект інтеграторів. На сьогодні технологія FaceVACS використовує алгоритм розпізнавання обличчя V10T9. Особливості системи FaceVACS включають можливість одночасного відстеження кількох облич, порівняння обличчя в реальному часі, відображення та відправлення статистики про потоки, підтримку інтерактивної реєстрації з непорушного, використання C++ API та Web Services API.

Недоліки цієї системи включають проблеми з фоном та низьку завадостійкість, а також обмежену кількість ключових точок. Дотриманням вказаних особливостей можна значно покращити ефективність та точність системи в різних умовах.

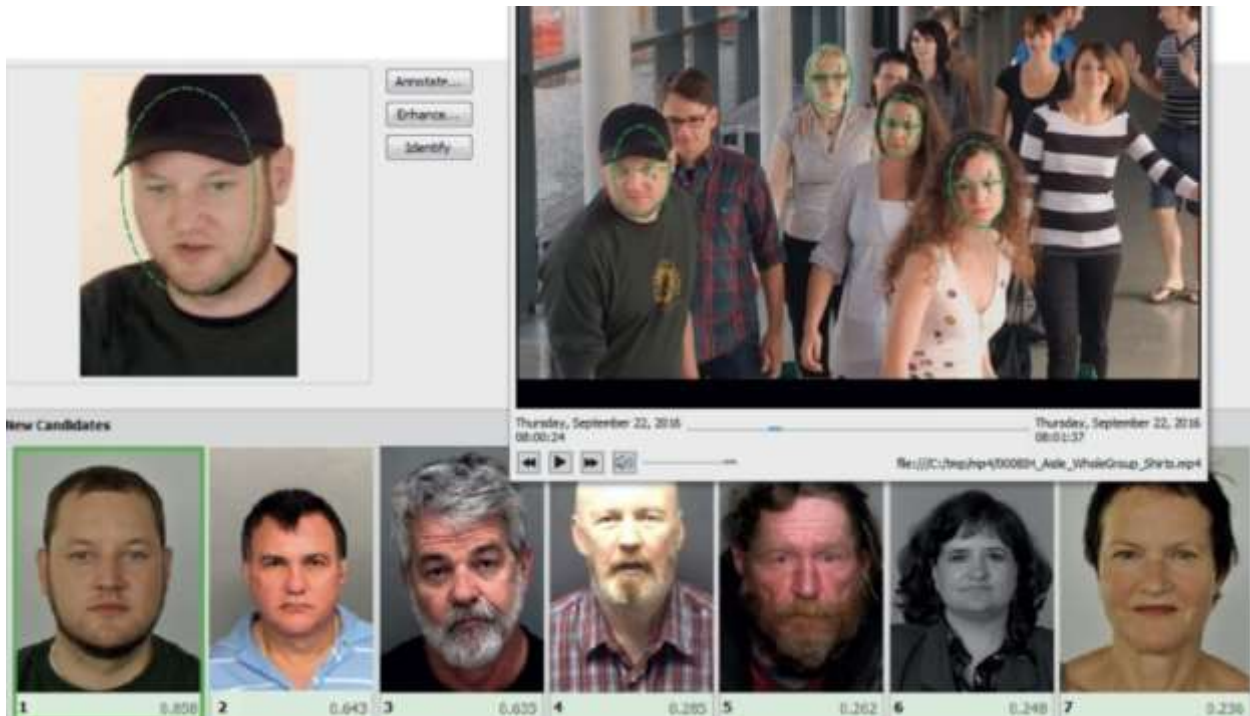


Рис. 5 – Приклад роботи розпізнавання системи FaceVACS

1.6.2 Система NEC's Face Recognition від компанії NEC

NEC Face Recognition[6] є передовою системою визначення осіб, створеною японською компанією "NEC". Ця технологія ефективно розпізнає людей навіть на старих фото або в ситуаціях, коли обличчя захищені окулярами чи виражальністю. Інформація про визнані обличчя надійно зберігається в базі даних, що дозволяє докладно переглядати історію відеозаписів і отримувати дані про дату та час кожного розпізнаного зображення. Технологія NEC виділяється високою точністю та швидкістю порівняно з іншими системами розпізнавання, проявляючи вражаючу продуктивність при обробці відео низької якості та сильно стиснених зображень. Система аналізує різні характеристики обличчя, такі як розмір, форма зіниць, лінії носа і рота, а також їх просторові взаємозв'язки, забезпечуючи ефективний пошук в базі даних.

Структура системи включає різні модулі, які реалізують ефективні алгоритми. Метод узагальненого виявлення облич (GMFD) заснований на нейронних мережах та призначений для швидкого та точного виявлення та

розпізнавання обличчя. Алгоритм Perturbation Space Method (PSM)[7] вирішує проблеми, пов'язані з позиціонуванням обличчя в кадрі, навіть у випадках, коли воно нахиле чи повернуте під певним кутом. Метод адаптивного регіонального зіставлення (ARBM) спрямований на зменшення впливу малих змін на обличчя, таких як зміни виразу, наявність окулярів або головного убору, для забезпечення точного розпізнавання. Система розпізнавання осіб NeoFace від NEC відзначається спостереженням і контролем в режимі реального часу, ідентифікацією на основі індивідуальних характеристик обличчя, багатозадачністю розпізнавання та можливістю пошуку подій у базі даних.



Рис. 6 – Приклад роботи розпізнавання системи NEC's Face Recognition

1.6.3 Система VeriLook SDK від компанії Neurotechnology

VeriLook SDK [8] - це технологія розпізнавання облич, створена Neurotechnology. Ця система дозволяє виявляти обличчя та одночасно розпізнавати осіб на кадрах зі швидкістю до 100000 облич за секунду. VeriLook SDK є комплектом для розробки програмного забезпечення і сумісний з різними пристроями на Windows, Linux, Mac OS X, iOS та Android.

Алгоритм VeriLook використовує глибокі нейронні мережі для локалізації обличчя на цифрових зображеннях. Система володіє численними перевагами:

1. Одночасна обробка кількох облич.
2. Гендерна класифікація, яка може визначати стать для кожної людини на зображенні.
3. Живе розпізнавання обличчя, що визначає, чи є обличчя "живим" на відеопотоці.
4. Розпізнавання емоцій за шістьма основними категоріями.
5. Визначення атрибутів обличчя, таких як посмішка, відкритий рот, закриті очі, окуляри, борода чи вуса.
6. Визначення якості обличчя з використанням порогу якості.
7. Можливість додавання декількох зразків одного обличчя з різних джерел та часів.
8. Ідентифікаційна здатність для порівнянь 1-к-1 та режиму 1-до-багатьох.

Ця технологія є високоефективною та гнучкою, що робить її ідеальним вибором для розпізнавання облич у різних сценаріях та застосуваннях.

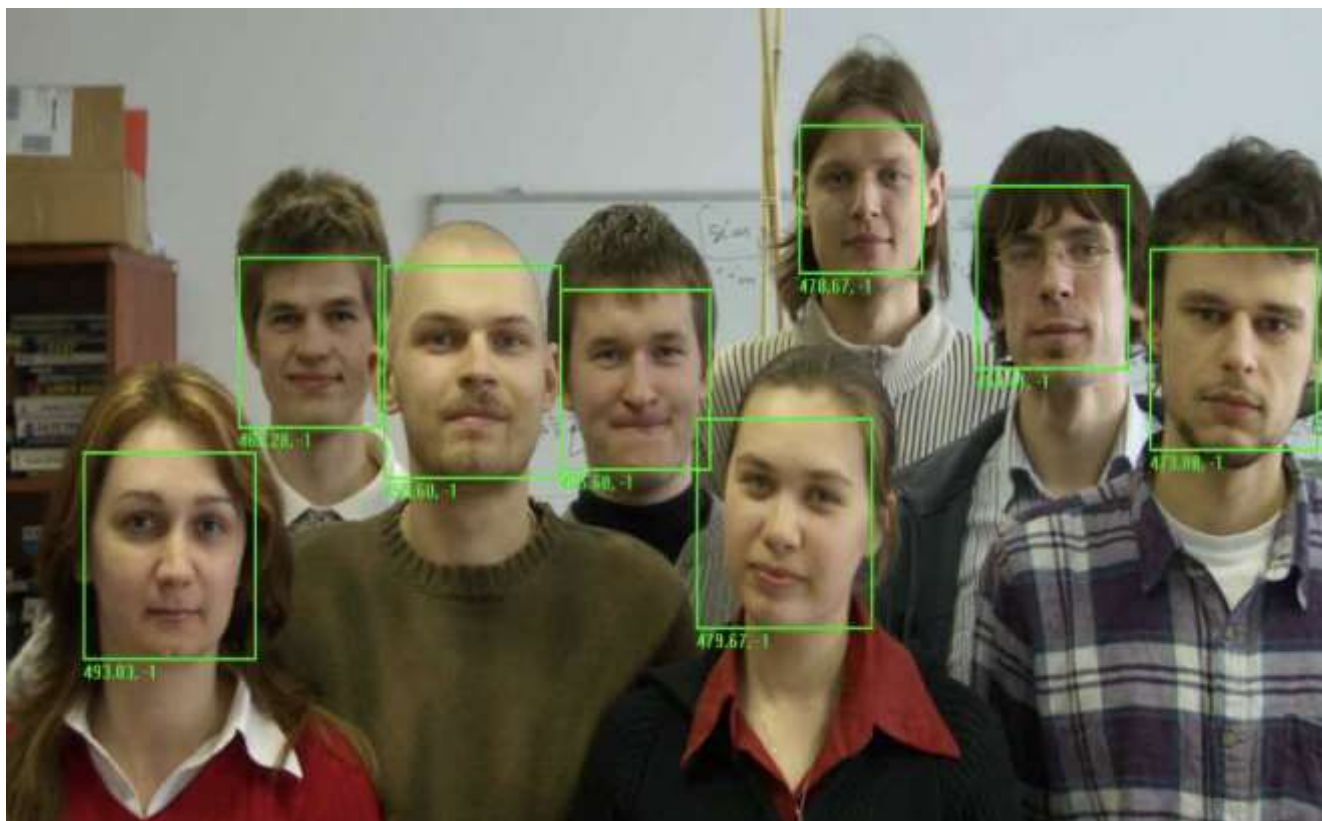


Рис. 7 – Приклад роботи розпізнавання системи VeriLook SDK

1.6.4 Система Microsoft Emotion Recognition

Emotion Recognition дозволяє використовувати нейронні мережі без потреби їх навчання. Ця система є простою у використанні та постійно вдосконалюється, знаходячи застосування в таких додатках, як Microsoft Cortana та Skype. Переваги включають постійний прогрес та вдосконалення, можливість інтеграції у робочі процеси, зручний інтерфейс та невисокі вимоги до обчислювальної потужності. З іншого боку, недоліки включають обмежену здатність працювати зі статичними зображеннями та проблеми з проекцією, кутом нахилу та огляду.

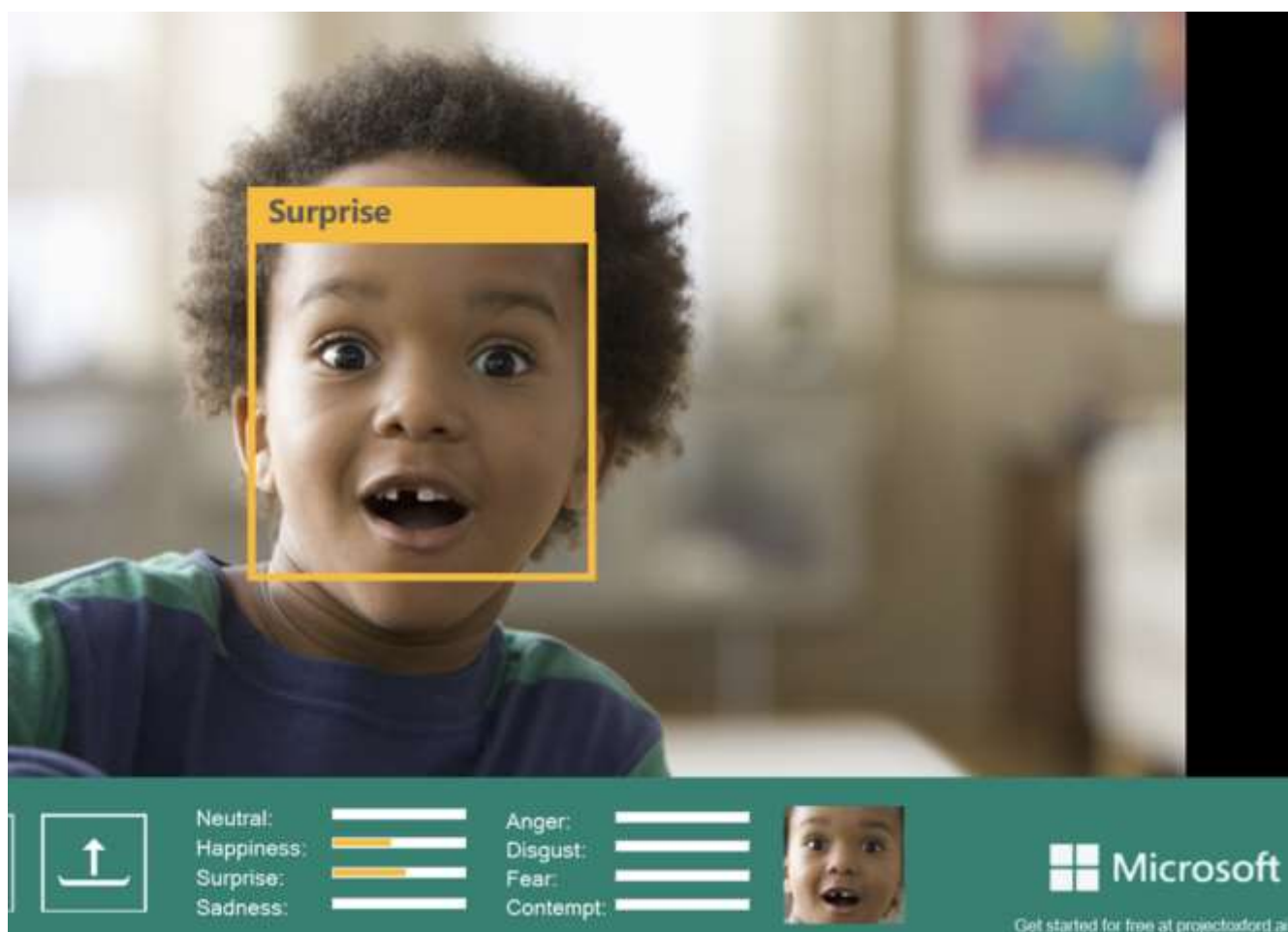


Рис. 8 – Приклад роботи розпізнавання системи Microsoft Emotion Recognition

1.6.5 Система FaceReader від компанії Noldus

Нідерландська компанія Noldus розробила FaceReader, інноваційний додаток з розширеними функціональними можливостями. FaceReader вміє впізнавати шість основних емоцій, а також розпізнавати нейтральний стан та зневагу.

Переваги додатка включають відображення та інтеграцію даних, можливість онлайн використання, високий рівень розпізнавання (80%) та розширений функціонал.

З іншого боку, недоліки додатка полягають у високій ціні за розширений функціонал, відсутності можливості роботи офлайн, неповноцінному розпізнаванні дитячих емоцій та проблемах з частково прихованими ділянками обличчя.

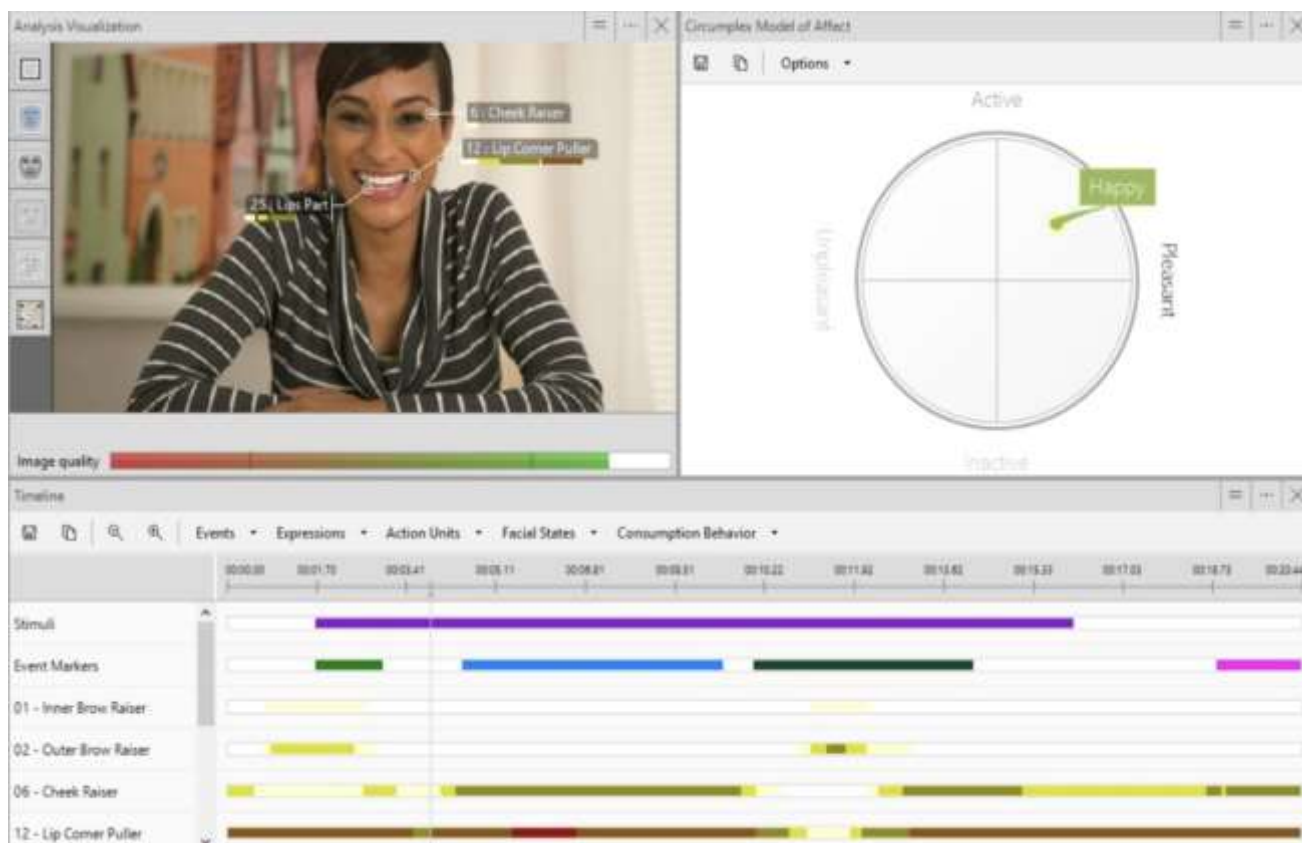


Рис. 9 – Приклад роботи додатка FaceReader

1.6.6 Система Betaface від компанії IRIS

Betaface - це технологія розпізнавання облич, розроблена білоруською компанією Image Recognition Integrated Systems (IRIS). Представлена як потужний інструмент для розпізнавання та верифікації осіб, Betaface здобула визнання завдяки своїй точності та універсальності.

Betaface був створений командою IRIS, компанії, що спеціалізується на технологіях розпізнавання зображень та облич. Розробники прагнули надати комплексне рішення для додатків, пов'язаних з розпізнаванням облич, від виявлення до верифікації.

Betaface використовує складні алгоритми для аналізу рис обличчя та створення унікального підпису для кожної людини. Процес складається з декількох ключових етапів:

Виявлення обличчя: Система ідентифікує та ізолює обличчя на зображенні або відеопотоці, фокусуючись на критично важливих рисах обличчя.

Виділення рис обличчя: Betaface виділяє ключові риси обличчя, такі як очі, ніс, рот і загальну структуру обличчя. Потім він аналізує просторові зв'язки між цими рисами.

Створення підпису обличчя: Витягнуті риси сприяють створенню унікального підпису обличчя, по суті, цифрового представлення обличчя людини.

Зіставлення та перевірка: Отримавши нове зображення, Betaface порівнює підпис обличчя з базою даних зі збереженими підписами. Це забезпечує точне зіставлення та верифікацію обличчя.

gПереваги:

Точність: Висока точність розпізнавання обличчя робить Betaface ідеальним для точних застосувань.

Універсальність: Технологія може використовуватися в різних областях, включаючи безпеку, контроль доступу та верифікацію особи.

Реальний час: Розпізнавання обличчя в реальному часі забезпечує швидкі та ефективні результати.

Інтеграція: Легка інтеграція з різними платформами та додатками для широкого застосування.

Недоліки:

Вартість: Великі витрати на впровадження, що може стати обмежуючим для деяких компаній.

Конфіденційність: Підвищені обліки щодо конфіденційності, як у більшості систем розпізнавання обличчя.

Залежність від бази даних: Точність значно залежить від якості та розміру бази даних, що може впливати на результати розпізнавання.

Отже, Betaface - це надійне рішення для розпізнавання облич, розроблене IRIS, що поєднує в собі точність і універсальність. Хоча воно має значні переваги з точки зору точності і обробки в реальному часі, при його впровадженні слід ретельно враховувати питання вартості, конфіденційності та екологічних факторів.



Рис.10 – Приклад роботи додатка Betaface

1.6.7 Система EmoVu від компанії Eyuris Technologies

EmoVu, технологія розпізнавання емоцій, стала відомою завдяки своїй здатності розшифровувати людські емоції за виразом обличчя.

EmoVu був розроблений компанією Eyuris Technologies, що спеціалізується на рішеннях у сфері штучного інтелекту. Заснована Модаром Алауї в 2013 році, Eyuris знаходиться в авангарді розвитку технології розпізнавання емоцій обличчя.

EmoVu використовує передові алгоритми комп'ютерного зору та машинного навчання для аналізу виразу обличчя та вилучення емоційної інформації. Технологія розбиває риси обличчя на ключові моменти, зіставляючи їх з певними емоціями. Розпізнаючи тонкі зміни у виразі обличчя, EmoVu може ідентифікувати цілу низку емоцій - від радості та здивування до гніву та смутку.

Процес включає в себе виявлення орієнтирів на обличчі, таких як рух очей, вигин губ і положення брів. Потім ці особливості аналізуються, щоб визначити емоційний стан людини. EmoVu працює в режимі реального часу, забезпечуючи швидку і точну емоційну оцінку.

Переваги:

Точність: Висока точність розпізнавання обличчя робить Betaface ідеальним для точних застосувань.

Універсальність: Технологія може використовуватися в різних областях, включаючи безпеку, контроль доступу та верифікацію особи.

Реальний час: Розпізнавання обличчя в реальному часі забезпечує швидкі та ефективні результати.

Інтеграція: Легка інтеграція з різними платформами та додатками для широкого застосування.

Недоліки:

Вартість: Великі витрати на впровадження, що може стати обмежуючим для деяких компаній.

Конфіденційність: Підвищені обліки щодо конфіденційності, як у більшості систем розпізнавання обличчя.

Залежність від бази даних: Точність значно залежить від якості та розміру бази даних, що може впливати на результати розпізнавання.

EmoVu є свідченням безперервного розвитку технології розпізнавання емоцій. Хоча його точність і адаптивність роблять його потужним інструментом, міркування щодо вартості, конфіденційності та екологічних

факторів підкреслюють необхідність збалансованого та етичного підходу до його розгортання. Оскільки Eyeris Technologies продовжує вдосконалювати і покращувати EmoVu, ландшафт емоційного ШІ, безсумнівно, стане свідком подальшого прогресу в прагненні зрозуміти і реагувати на людські емоції.

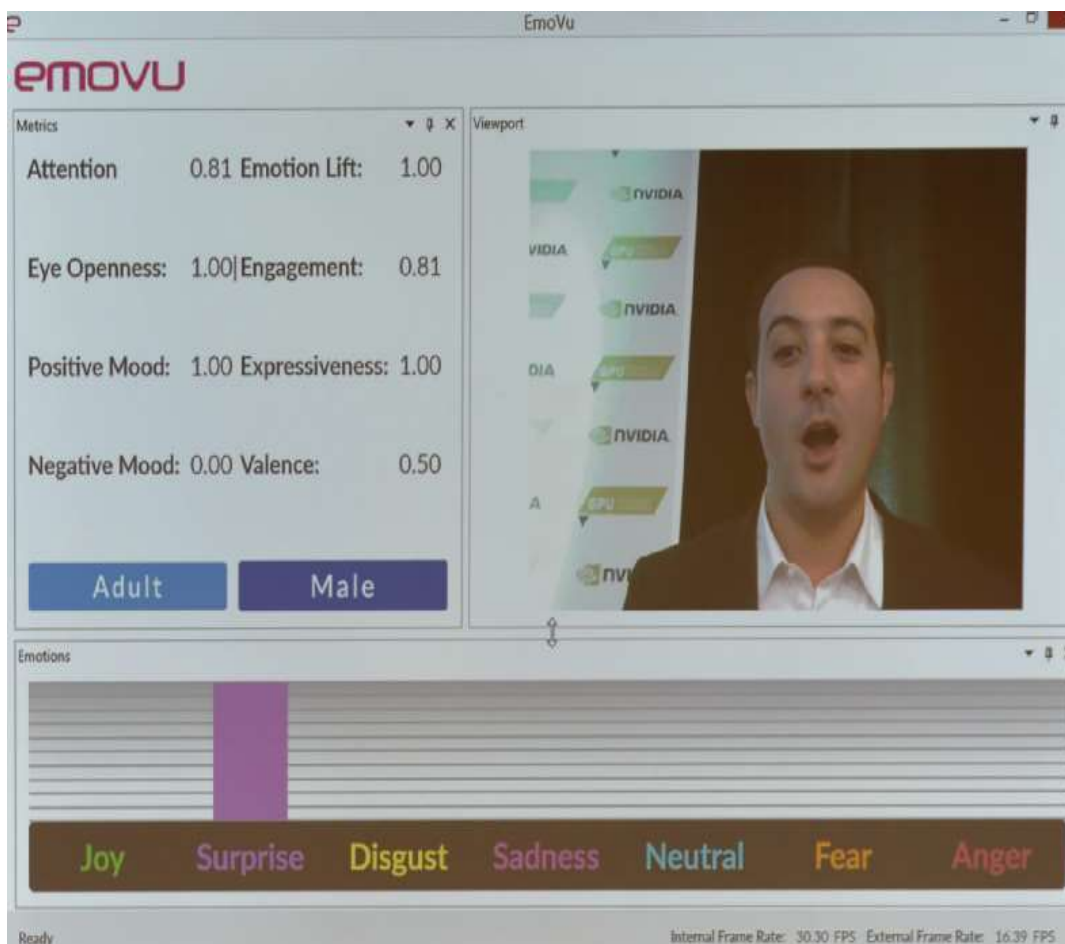


Рис.11 – Приклад роботи додатка EmoVu

1.7 Висновки до першого розділу

В цьому розділі виконано аналіз об'єкта дослідження, що включає вивчення історії еволюції комп'ютерного зору. Додатково розглянуті різні сфери застосування систем верифікації особи та розпізнавання емоцій за обличчям, кожна з яких супроводжується конкретним прикладом.

Під час дослідження були ретельно вивчені основи методів розпізнавання обличчя. Також була розглянута оцінка ефективності систем розпізнавання, визначено основні поняття.

Був проведений порівняльний аналіз існуючих систем розпізнавання обличчя людей. Окрема увага приділялася висвітленню основних переваг і недоліків кожної програми в контексті її використання в конкретних областях застосування.

РОЗДІЛ 2

ОСНОВНІ МЕТОДИ ТА АЛГОРИТМИ ОБРОБКИ ЗОБРАЖЕНЬ

На сьогоднішній день існує безліч різноманітних методів розпізнавання ключових елементів та рис обличчя людини, кожен з яких розроблений для роботи з певним видом зображень. Для виконання конкретної задачі важливо обирати найбільш оптимальний і ефективний метод. У даному розділі детально розглядаються методи розпізнавання ключових елементів обличчя, які

використовуються найчастіше та є в основі більшості систем автоматичного розпізнавання.

2.1 Метод Віоли Джонса

Метод представлений у 2001 році Полом Віолою та Майклом Джонсоном [10, 11], продовжує залишатися ефективним для реального часу пошуку об'єктів на зображеннях та відео послідовностях. Важливо відмітити, що цей детектор характеризується низькою ймовірністю помилкового виявлення обличчя.

Основні принципи методу включають інтегральне представлення зображення. Тобто воно подається у вигляді інтегралу, що дозволяє ефективно обчислювати необхідні об'єкти [12]. Інтегральне представлення зображення представляє матрицю, розміри якої співпадають із вихідним зображенням. Кожен елемент містить суму яскравостей пікселів, розташованих ліворуч і вище даного елемента. Елементи матриці розраховуються за формулою

$$L(x, y) = \sum_{i=0, j=0}^{i \leq x, j \leq y} I(i, j),$$

де $L(x, y)$ - інтегральне зображення; $I(i, j)$ - вихідне зображення. Приклад обчислення інтегрального зображення на рис. 10.

0	1	1	1
1	2	2	3
1	2	1	1
1	3	1	0

0	1	2	3
1	4	7	11
2	7	11	16
3	11	16	21

Вихідне зображення $I(i, j)$

Інтегральне зображення $L(x, y)$

Рис. 12 – Приклад інтегрального зображення

Нехай наявний прямокутник АБВГ (рис. 2). Тоді суму пікселів всередині прямокутника виражається різницю суміжних прямокутників за формулою

$$S(АБВГ) = L(Б) + L(Г) - L(А) - L(В)$$

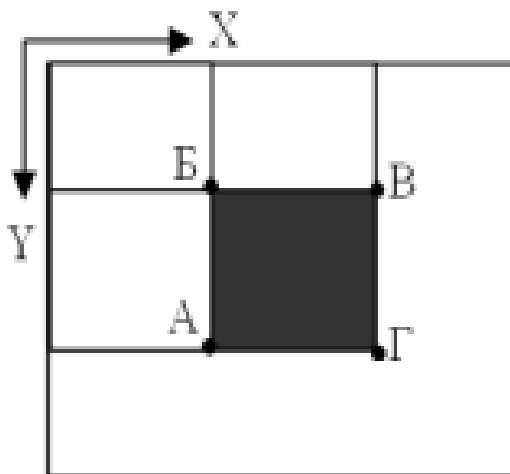


Рис. 13 – Обчислення суми яскравостей пікселів довільного прямокутника.

Ознаки Хаара. Застосовуючи ознаки Хаара [13], проводиться виявлення потрібного об'єкта, яким у цьому випадку є обличчя та його характеристики.

Ознака - відображення області зображення на множину допустимих значень, тобто

$f : X \Rightarrow D$, де X - область зображення для обчислення ознаки; D_f - множина допустимих значень ознаки.

У стандартному методі Віулі - Джонса використовують прямокутні ознаки, подані на рис.12

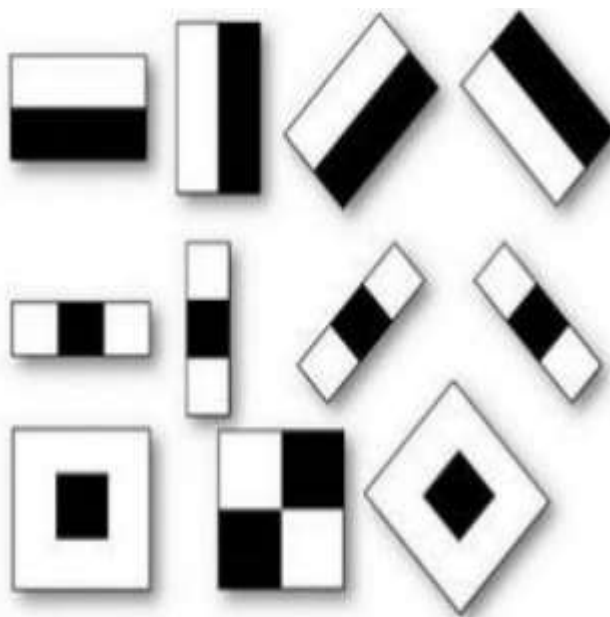


Рис. 14 – Приклад ознак Хаара

Бустинг. Для вибору найбільш підходящих ознак для шуканого об'єкта на частині зображення застосовується бустинг.

Бустинг - це набір методів, спрямованих на підвищення точності аналітичних моделей. У загальному розумінні, бустинг означає посилення "слабких" моделей, що включає у себе послідовне покращення композиції алгоритмів машинного навчання. Кожен новий алгоритм у цьому процесі спрямований на усунення недоліків попередніх алгоритмів у композиції.

Метод комбінування класифікаторів у каскадну структуру. Для швидкого відкидання вікон, де не знайдено обличчя, застосовуються каскади ознак. Метод підвищує швидкість виявлення, фокусуючи свою роботу на найбільш інформативних областях зображення. Структуру каскадного детектора наведено на рис. 13. Каскад складається з шарів, які являють собою класифікатори, навчені за допомогою процедури бустингу.

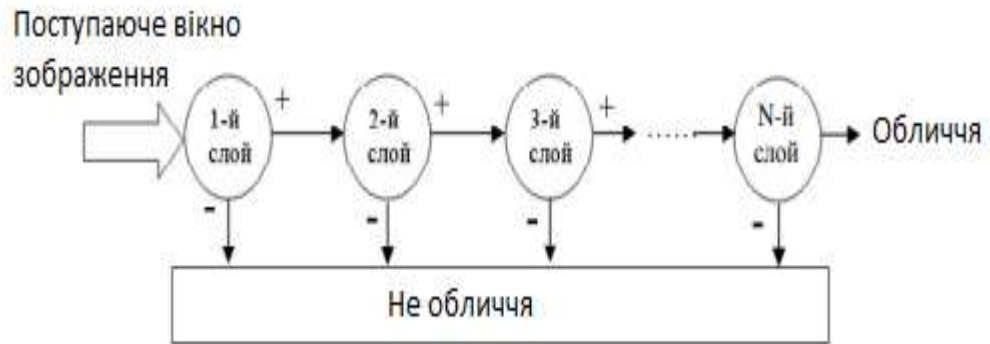


Рис. 15 – Структура каскадного детектора

Метод Віоли-Джонса є дуже популярним завдяки своїм перевагам:

1. Він може виявляти всі обличчя на зображенні;
2. Завдяки класифікатору він працює дуже швидко у відношенні обробки та розпізнавання;
3. Хоча метод був розроблений для розпізнавання облич, його можна також використовувати для ідентифікації інших об'єктів;
4. Він може застосовуватися в режимі реального часу для відеопотоку.

Проте є певні недоліки у цьому методі:

1. Вимагає великої кількості баз даних з урахуванням їхніх особливостей для ефективного навчання;
2. Для досягнення високої точності розпізнавання необхідно витратити значний час на алгоритм навчання.

2.2 Метод головних компонент

Метод головних компонент (Principal Component Analysis, МГК) дає змогу перейти від простору ознак великої розмірності до простору ознак малої розмірності так, щоб ознаки стали некорельованими. Цей підхід у завданні розпізнавання обличч було вперше використано в 1991 році М. Турком і А. Пентлендом і отримав назву "Власні обличчя" (від англ. Eigenfaces)[14]

Основна мета цього методу полягає в представленні зображень обличч у вигляді набору головних компонентів, які називаються "власні обличчя".

Кожен вектор відповідає конкретному зображенню з обличчям подібної форми. Для знаходження цих головних компонентів використовується розрахунок власних векторів і власних значень коваріаційної матриці, що визначається з зображень. Реконструкція зображення відбувається шляхом сумування головних компонент, помножених на відповідні власні вектори.

Для представлення обличчя проводиться розрахунок основних компонентів, вибираючи їх кількість від 5 до 200. Процес розпізнавання включає порівняння основних компонент невідомого зображення з відомим. Зображення облич однієї і тієї ж особи групуються в кластери. З бази даних беруться зображення-кандидати, відстані від яких до вхідного зображення є мінімальними.

Переваги розглянутого методу :

1. У наборі зображень осіб за наявності варіацій, таких як раса, стать, емоції, освітлення, створюються компоненти, величина яких загалом визначається цими факторами. У зв'язку з цим за значеннями відповідних головних компонент можна визначити, наприклад, расу або стать людини;
2. Робота з великими базами даних.

Недолік методу :

- Вимоги до умов зйомки зображень надто високі. Фотографії повинні бути зроблені в схожому освітленні та з подібним ракурсом. Потрібна попередня обробка.

При ідеальних умовах точність такого методу може перевищувати 90%.

2.3 Лінійний дискримінантний аналіз

Лінійний дискримінантний аналіз[15], дискримінант Фішера (Linear Discriminant Analysis, Fisherfaces, ЛДА), - метод статистики та машинного навчання, який застосовують для знаходження лінійних комбінацій ознак, що якнайкраще розділяють два або більше класи об'єктів або подій. Отриману

комбінацію використовують як лінійний класифікатор або для скорочення розмірності простору ознак перед подальшою класифікацією.

Суть цього методу полягає в виборі проекції простору зображень на простір ознак з метою мінімізації внутрішньокласової відстані та максимізації міжкласової відстані в просторі ознак. Проектування на власний простір призводить до змішування класів, ускладнюючи процес розпізнавання, у той час як лінійний дискримінант обирає таку проекцію на простір ознак, яка чітко розділяє різні класи.

Переваги методу ЛДА є висока точність розпізнавання (близько 94 %) та метод не впливає на розпізнавання умови освітленості, різні вирази обличчя і наявність або відсутність перешкод (окулярів, борід).

Недоліком цього методу є те, що не проводяться експерименти зі зміною ракурсу, тому працездатність методу не відома.

2.4 Метод гнучкого порівняння на графах

Метод гнучкого порівняння на графах відноситься до методу перебору. Він використовує порівняння з базою даних, де для кожного типу об'єктів подані різноманітні модифікації відображення.

Основне завдання методу полягає в гнучкому порівнянні графів, які описують зображення облич. Обличчя представлені у формі графів з вагованими вузлами та ребрами. Під час розпізнавання один із графів залишається незмінним (еталонним), тоді як інший змінюється для найкращого відповідності першому.

Обчислення значень ознак відбувається в вузлах графа за допомогою комплексних фільтрів Габора[16] або їх впорядкованих наборів в конкретній локальній області. Вага ребер графа визначається відстанями між сусідніми вузлами, а різниця між графами оцінюється функцією деформації. Ця функція враховує відмінності в значеннях ознак в вузлах та ступінь деформації ребер графа.

Створення графа включає зсув кожної вершини на визначену відстань у певних напрямках відносно її початкового положення та вибір оптимальної позиції. Це дозволяє мінімізувати відмінність у значеннях ознак між деформованим та еталонним графами. Процес деформації застосовується до всіх вузлів графа, поки не досягається мінімальна різниця між характеристиками деформованого та стандартного графів. Показник функції деформації вказує на відмінність між вхідним зображенням та еталонним графом. Цей процес повторюється для всіх еталонних облич, що входять до бази даних системи. Результатом розпізнавання є еталон із найкращим значенням функції деформації.

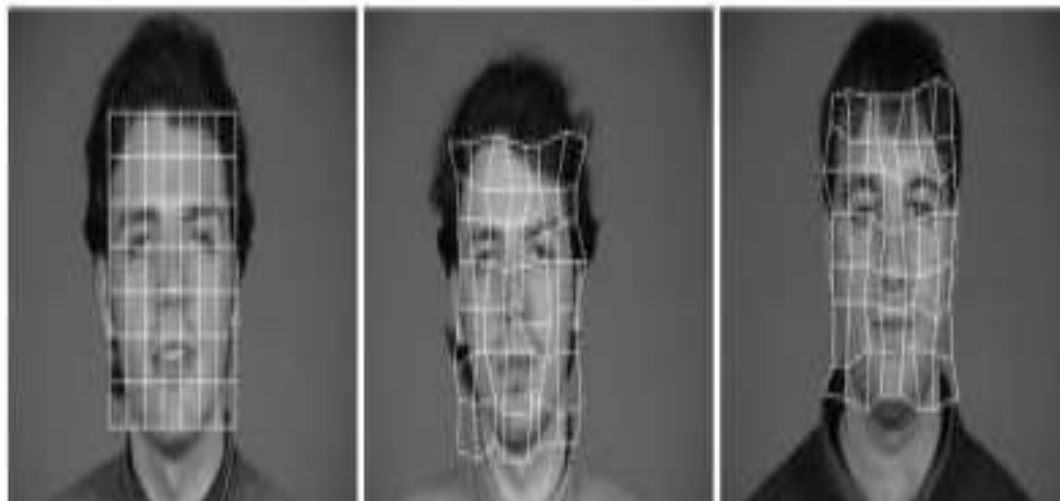


Рис. 16 – Приклад деформації графа у вигляді регулярної решітки

Метод гнучкого порівняння на графах демонструє високу ефективність, розпізнавши різні емоційні вирази та зміни ракурсу обличчя до 15 градусів на рівні 90–95%. Його переваги включають високу точність, стійкість до зміни ракурсу та міміки. Однак, маючи високу обчислювальну складність та лінійну залежність часу роботи від розміру бази даних, метод має свої недоліки, також виявивши низьку технологічність при запам'ятовуванні нових еталонів.

2.5 Згорткова нейронна мережа

Згорткова нейронна мережа (CNN)[17], розроблена Яном Лекуном у 1988 році, представляє собою спеціалізовану архітектуру штучних нейронних мереж, спрямовану на ефективне аналіз зображень у глибокому навчанні. Існують два типи клітин: прості, що реагують на прямі лінії під різними кутами, і складні, що активуються в результаті взаємодії конкретного набору простих клітин.

Згорткова нейронна мережа має багат шарову та одно спрямовану архітектуру. Для навчання використовують різноманітні методи, зокрема метод зворотного поширення помилки. Важливою особливістю є можливість вибору функції активації нейронів (передавальної функції), яка залежить від вибору дослідника.

Ця архітектура дозволяє впроваджувати концепції зорової кори та допомагає в ефективному аналізі зображень у різних завданнях, таких як розпізнавання облич, класифікація об'єктів та багато інших.

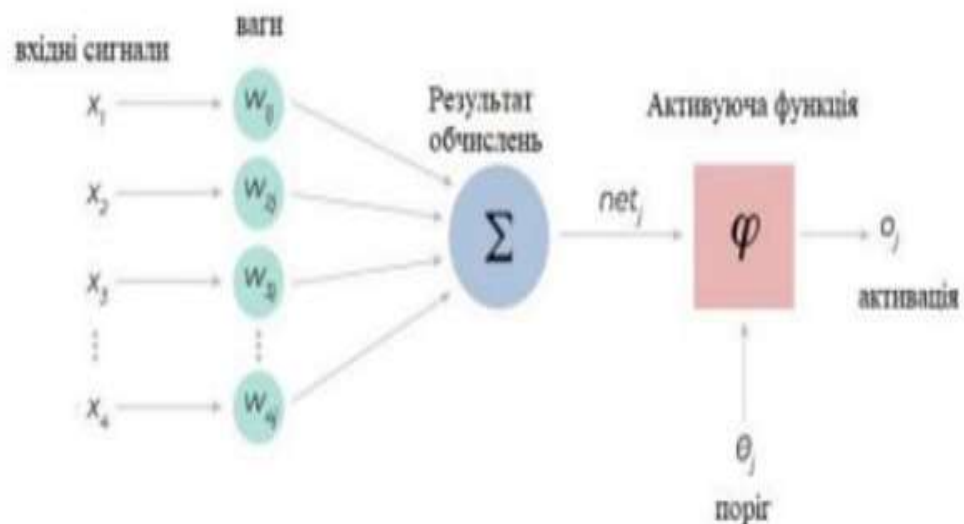


Рис. 17 – Структура нейронної мережі

Ця мережа отримала свою назву через використання операції згортки, де кожен фрагмент зображення перемножається на матрицю (ядро) згортки поелементно, а потім результати сумуються та записуються у відповідні позиції

вихідного зображення. Вона відзначається високою точністю розпізнавання, перевищуючи 90%, і стійкістю до шумів у вхідних даних.

Однак важкість реалізації та перенавчання, особливо при додаванні нового обличчя до бази даних, визначає недоліки цієї архітектури. Важливо розглядати ці аспекти при впровадженні та вдосконаленні системи для забезпечення її ефективності та легкості використання.

2.6 Метод ключових точок

Після виявлення обличчя на зображенні застосовують метод пошуку ключових точок. Для визначення особливих точок використовують два компоненти: детектор та дескриптор [18].

Детектор визначає особливі точки на обличчі, тоді як дескриптор виділяє ці особливі точки серед великої кількості інших і ідентифікує їх.

Емоційний стан людини відображається у міміці обличчя через зміни в положенні та формі губ, очей і брів. За допомогою цих даних визначають ключові точки елементів обличчя та відслідковують зміни їх положення. Після аналізу даних з точок можна визначити емоційний стан людини рис.16.

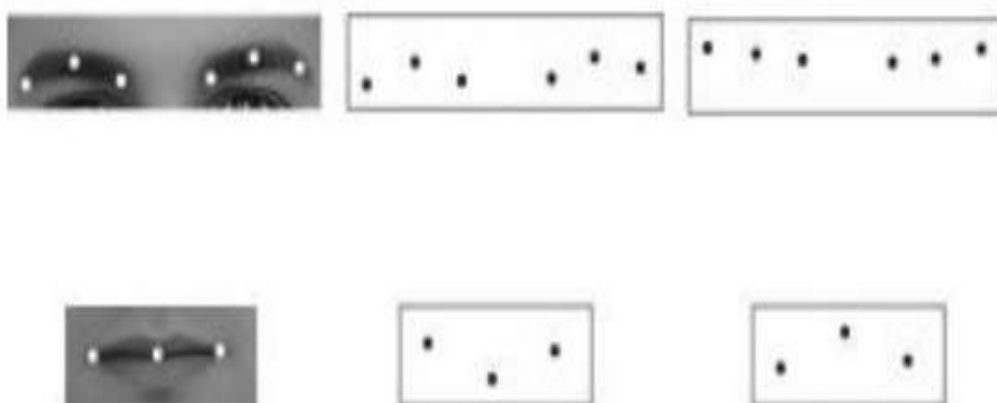


Рис. 18 – Приклад виявлення ключових точок.

Аналізуючи положення точок брів та губ, можна визначити, чи людина щаслива чи вона виражає гнів. Так можна визначити основні види емоцій.

2.7 Алгоритм Local Binary Patterns Histograms (LBP)

Алгоритм Local Binary Patterns Histograms (LBP) був вперше представлений Т. Ояла в 1996 році. Він конвертує кожен піксель зображення в бінарне число в залежності від яскравості його сусідніх пікселів. Цей метод знаходження локальних бінарних шаблонів є ефективним з обчислювальної точки зору, оскільки він операцію проводить тільки з цілими числами і залишається нечутливим до змін яскравості на зображенні.

LBP - це метод опису області навколо кожного пікселя на зображенні за допомогою восьми розрядного двійкового коду. Спочатку зображення розділяється на блоки, і для кожного блока обчислюється гістограма кодів. Процес включає в себе вибір центрального пікселя, визначення його інтенсивності та порівняння її з сусідніми пікселями. В результаті отримується восьмирозрядний двійковий код, який відображає область навколо даного пікселя(рис.17).

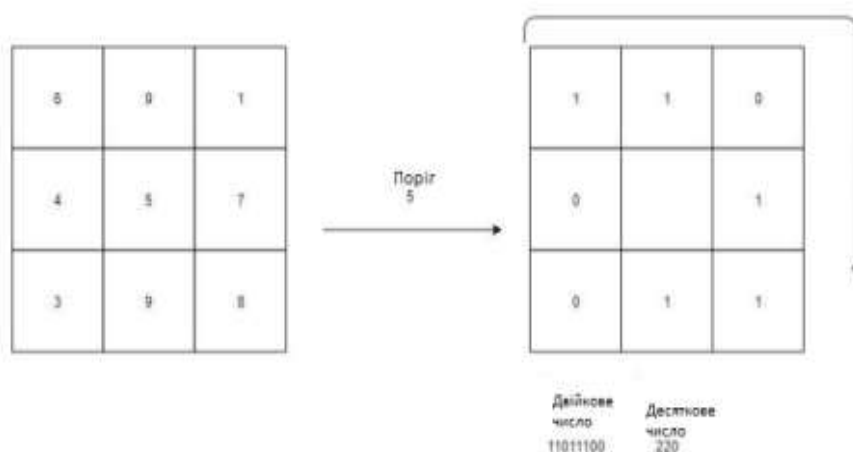


Рис. 19 – Алгоритм роботи LBP

Загалом, метод дії алгоритму такий: зображення поділяється на блоки, і для кожного блоку обчислюється власна гистограма. Шляхом об'єднання цих гистограм формується загальна гистограма, яка враховує як локальні, так і глобальні особливості зображення.

2.8 Бібліотеки

2.8.1 OpenCV

У 1999 році компанія Intel розпочала реалізацію проекту OpenCV[19], що призвело до створення бібліотеки баз даних OpenCV. Розробка бібліотеки виконана на мовах програмування C та C++ і може інтегруватися з іншими популярними мовами, такими як Python, C++, Java. OpenCV є сумісною з операційними системами Windows, Linux, MacOS, iOS та Android.

Бібліотека OpenCV має відкритий код і включає в себе різноманітні дані для застосування в системах комп'ютерного зору, навчальні ресурси для автоматичних систем розпізнавання об'єктів та функції обробки алгоритмів. Загальна кількість алгоритмів в базі складає 2500, включаючи класичні та сучасні оптимізовані алгоритми обробки даних.

Основною метою OpenCV є забезпечення високої швидкості автоматичного машинного розпізнавання для систем, що використовують комп'ютерний зір для розпізнавання об'єктів та обличчя людей. Бібліотека має різноманітні застосування, такі як класифікація об'єктів на зображеннях, розпізнавання та відстеження рухомих об'єктів, знаходження ключових елементів обличчя, а також використання при розпізнаванні пейзажів.

OpenCV володіє численними перевагами, що роблять його популярним та зручним інструментом для роботи з зображеннями.

Активна Спільнота: Широко використовується по всьому світу, у тому числі в компаніях, таких як Google та Microsoft, завдяки активній спільноті для різних мов програмування.

Документація та Відкритий Код: Наявність документації на різних мовах та відкритий вихідний код дають можливість як початківцям, так і досвідченим розробникам ефективно використовувати OpenCV.

Вільний Доступ: Бібліотека поширюється під вільною ліцензією, дозволяючи використовувати її як для освітніх, так і комерційних цілей.

Багатофункціональність: З понад 2500 інструментів і алгоритмів, OpenCV відмінно впорається з складними завданнями обробки та розпізнавання зображень.

Висока Швидкість: Бібліотека працює швидше, ніж інші важкі програми для математичних обчислень, забезпечуючи ефективну обробку зображень.

Робота в Реальному Часі: Завдяки високій швидкості та кросплатформенності, OpenCV підходить для використання в системах розпізнавання в режимі реального часу.

2.8.2 TensorFlow

Розробленою Google Brain та представлено у 2015 році, TensorFlow[20] – це відкрита бібліотека для числових обчислень і розгортання масштабних моделей машинного навчання. TensorFlow об'єднує в собі різноманітні моделі та алгоритми машинного і глибокого навчання (також відомих як нейронні мережі) і робить їх доступними за допомогою спільних програмних метафор. Зручний API дозволяє розробникам створювати додатки з використанням Python або JavaScript, під час коли основна платформа виконує ці додатки за допомогою високопродуктивної мови C++. TensorFlow також надає бібліотеки для інших мов, хоча Python переважає.

У порівнянні з іншими фреймворками, TensorFlow може використовуватися для навчання та застосування глибоких нейронних мереж у різних завданнях. Важливо відзначити, що TensorFlow забезпечує можливість прогнозування в масштабі, використовуючи ті ж самі моделі, що і для процесу навчання.

TensorFlow має свої переваги та недоліки. Серед переваг можна виділити відкритий код, підтримку різних мов програмування, ефективність та швидкість обробки даних, а також постійний розвиток. Проте, до недоліків слід віднести потребу в значній кількості програмного коду, можливі проблеми при внесенні змін в архітектуру, відсутність можливості роботи в реальному часі та обмежену кількість баз даних з особливостями.

2.9 Висновки до другого розділу

В другому розділі розглянуті різноманітні методи та алгоритми для виявлення облич на зображеннях та класифікації емоційного стану, використовуючи виявлені ключові елементи та риси обличчя. Подано опис роботи алгоритмів, розглянуті їх переваги та недоліки. Наразі існує значна кількість методів розпізнавання елементів обличчя, проте не існує єдиного універсального підходу. Вибір конкретного методу залежить від численних чинників.

Таким чином, завдання розпізнавання ключових елементів обличчя на зображеннях може бути вирішено за допомогою комбінації кількох методів розпізнавання для точного та ефективного виконання завдання виявлення облич. Комбінування різних методів підкреслює їх сильні сторони та компенсує слабкі. Це дозволяє отримувати високоякісні результати при розв'язанні різноманітних проблемних сценаріїв.

РОЗДІЛ 3

ОБҐРУНТУВАННЯ ОБРАННЯ ЗАСОБІВ РЕАЛІЗАЦІЇ ДОДАТКА

3.1 Обґрунтування мови програмування Python

Python[21] - популярна мова програмування. Її створив Гвідо ван Россум і випустив у 1991 році.

Python є популярною мовою програмування загального призначення, яку можна використовувати для широкого спектру додатків. Вона включає високорівневі структури даних, динамічну типізацію, динамічне зв'язування та багато інших можливостей, які роблять її настільки ж корисною для розробки складних додатків, як і для написання сценаріїв або "клейового коду", що з'єднує компоненти разом. Її також можна розширити, щоб здійснювати системні виклики майже до всіх операційних систем і запускати код, написаний на C або C++. Завдяки своїй повсюдності та здатності працювати майже на будь-якій системній архітектурі, Python є універсальною мовою, яку можна знайти в найрізноманітніших додатках.

Python ідеально підходить для написання додатка по ідентифікації користувача за обличчям. Він ефективно співпрацює і втілює переваги обраного набору засобів та алгоритмів.

3.2 Обґрунтування бібліотеки OpenCV

Таблиця 1 – Порівняння бібліотек OpenCV та TensorFlow

Програмне обладнання	TensorFlow	OpenCV
Автор	Google Brain	Intel

Платформа	Windows, Linux, Mac OS, IOS, Android	Windows, Linux, Mac OS, IOS, Android
Opensource	Так	Так
Мови програмування	C++, Python, Java, C#, JavaScript, Go	C++, Python, Java
Підтримка спільнотою	Так	Так
Низькорівневі оператори	Так	Так
Можливість відлагодження	Так	Так
Робота в реальному часі	Ні	Так
Можливість покращення	Так	Так
Швидкість роботи	Висока	Середня
Бази даних для навчання	Велика кількість простих але мала кількість з особливостями	Велика кількість будь яких
Середня точність	97%	97.5%

Після аналізу можливостей OpenCV та TensorFlow прийнято рішення взяти на озброєння OpenCV.

Окрім широкого спектру різноманітних баз даних для навчання, OpenCV має здатність працювати в режимі реального часу. Також його використання спрощує і швидше дозволяє додавати нових користувачів до системи..

3.3 Обґрунтування методу Віоли - Джонса та алгоритму LBPН

Зазвичай кожний метод має фундамент, без якого він не міг би існувати взагалі, і на цьому фундаменті будується вся інша структура. У методі Віоли-Джонса цей фундамент представлений примітивами Хаара, які розбивають задану

прямокутну область на набори різнотипних прямокутних підобластей. Ці примітиви Хаара є ключовим елементом, оскільки вони дозволяють аналізувати різні частини області та виділяти особливості, які можуть вказувати на наявність об'єктів чи певних ознак. Подальший алгоритм роботи ґрунтується на цих примітивах, дозволяючи ефективно виявляти обличчя або об'єкти у зображенні.

В бібліотеці OpenCV наявний каскадний алгоритм розпізнавання облич Хаара - це алгоритм типу "ковзного вікна", який виявляє об'єкти на основі їхніх особливостей.

Модель Хаара-Каскаду використовує різні типи розпізнавання ознак, які включають в себе розмір і розташування певних рис обличчя. Зокрема, перенісся, лінія рота та очі. Область навколо очей темніша, ніж верхня частина щоки. Область перенісся яскравіша за область очей. В OpenCV наявна попередньо навчену модель, яка була створена в результаті тривалого навчання і завантажена Райнером Ліенхартом від імені компанії Intel. Модель Райнера використовує алгоритм адаптивного прискорення (AdaBoost) для отримання кращих результатів і точності. Це закодований список дерев рішень, в якому кожна вершина тестує одну ознаку Хаара, і кожен список стверджує "це не обличчя" або "це може бути обличчям". Це може бути використано для перевірки того, що частина зображення є обличчям. У поєднанні з алгоритмом ковзного вікна це дає детектор облич. Для обробки великої кількості облич це є одним із ліпших рішень з точки зору надійності та швидкості.

Метод LBPН (Local Binary Patterns Histograms). У загальному випадку беруться значення яскравості пікселя зображення і ще восьми пікселів, що оточують вихідний. Виходить таблиця 3x3 зі значеннями яскравості пікселів. Потім у таку саму таблицю записуються 0 або 1. Якщо значення яскравості крайнього пікселя перевищує значення яскравості центрального, ставиться 1, інакше - 0. Потім отриманий код зчитується за годинниковою стрілкою з лівого верхнього осередку, перекладається в десяткове число, а це число записується в матрицю, що має розміри зображення, у відповідну позицію. І так для кожного

пікселя. Матриця ділиться на певну кількість сегментів (за замовчуванням це сітка 8×8), для кожного сегмента будують гістограму, і наприкінці шляхом конкатенації гістограм отримують результуючу, що характеризує все зображення. Під час розпізнавання для досліджуваного зображення будується така сама гістограма, яку потім порівнюють із тренувальними даними.

LVRN створений для впізнавання осіб при різних рівнях освітлення та різноманітних фонових умовах. Відмінністю є те, що для його ефективної роботи не потрібно великої кількості зображень однієї і тієї ж самої особи в різних умовах. У цьому алгоритмі конкретні особливості людини виражені простим вектором даних. Цей підхід дозволяє уникнути впливу варіацій у рівнях освітлення.

3.4 Обґрунтування бібліотеки Keras

Keras представляє собою бібліотеку для мови програмування Python, спрямовану на глибоке машинне навчання. Ця бібліотека дозволяє швидко створювати та налаштовувати моделі - схеми для обчислення та розповсюдження інформації під час навчання. Не виконуючи складних математичних обчислень, Keras використовується як надбудова над іншими бібліотеками. Далі описані ключові функції Keras, такі як зручна побудова моделей, конфігурація шарів у моделях для досягнення точності, обробка введення та виведення інформації, перетворення вхідних даних, візуалізація моделі, підбір датасетів для навчання, підготовка моделі до роботи з визначенням функцій помилки та оптимізаторів, а також етапи навчання, тестування та збірка програми машинного навчання.

Keras надав можливість використовувати навчені моделі, такі як `mini_xception`. `Mini_XCEPTION` це модель навчена на 28 тис. фото і здатна в реальному часі класифікувати емоції: `happiness` (щастя); `sadness` (смуток); `scare` (страх); `anger` (гнів); `disgust` (відраза); `contempt` (презирство); `surprise` (здивування); `neutral` (нейтральність - відсутність прояву емоцій). Основною

метою запропонованої моделі є автоматичне виявлення емоцій та прогнозування емоційного стану з високою точністю.



Рис. 20 – Приклад роботи Xception

3.5 Висновки до третього розділу

У третьому розділі ретельно розглянуто причини, які визначили вибір конкретних бібліотек та рішень у процесі втілення задачі. Детальний аналіз виправдовує обрані засоби і підкреслює їх ефективність у контексті реалізації поставленої задачі.

РОЗДІЛ 4

РЕАЛІЗАЦІЯ ДОДАТКУ ВЕРИФІКАЦІЇ ТА ЗБОРУ ДАНИХ КОРИСТУВАЧА ЗА ОБЛИЧЧЯМ

4.1 Пояснення роботи програми

Система розпізнавання облич у спрощеному вигляді складається з двох основних модулів:

1. Модуль виявлення облич у кадрі.
2. Модуль розпізнавання за ключовими ознаками.

Загальна схема роботи такої підсистеми показана на рис. 19. Вхідною інформацією для неї є відеопотік. З нього виділяється послідовність кадрів, до кожного з яких застосовується алгоритм виявлення облич. Отримавши область кадру з об'єктом, необхідно провести порівняння ключових ознак отриманого об'єкта з наявними в базі даних. В результаті буде отриманий список користувачів, які відповідають умовам відбору.

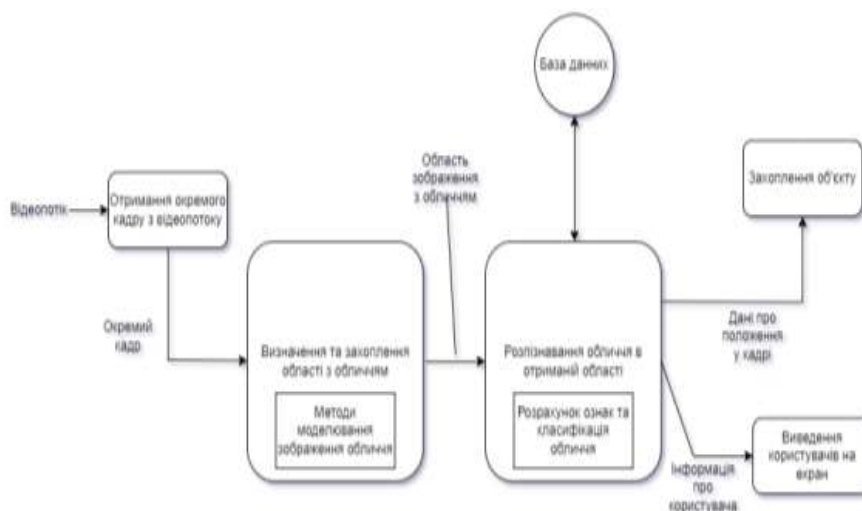


Рис. 21 – Узагальнена схема роботи підсистеми розпізнавання облич.

Для створення додатка верифікації користувача за обличчя потрібно:

1. Отримати права доступу до веб камери.
2. Зробити алгоритм для знаходження обличчя в кадрі.
3. Зібрати навчальну вибірку для кожного нового користувача.
4. Створити алгоритм для розпізнавання користувачів.

- Створити алгоритм для визначення емоційного стану людини, її віку та статі.

4.2 Розробка додатку

Перш за все потрібно отримати доступ до відеопотоку з веб камери. Для цього використовується функція VideoCapture

```
cap = cv2.VideoCapture(0)
pred = 0
while True:
    ret, frame = cap.read()
```

Далі нам потрібно знайти обличчя в відеопотоці і зібрати набір фотографій користувачів для подальшої їх верифікації.

```
vid = cv2.VideoCapture(0)
while True:

    ret, img = vid.read()
    new_img = None
    graying = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
    face = detector.detectMultiScale(image=graying, scaleFactor=1.1, minNeighbors=5)
    for x, y, w, h in face:
        cv2.rectangle(img, (x, y), (x+w, y+h), (0, 0, 0), 2)
        cv2.putText(img, "Face Detected", (x, y-5), cv2.FONT_HERSHEY_SIMPLEX, 0.8, (0, 0, 255))
        cv2.putText(img, str(str(num_of_images)+" images captured"), (x, y+h+20), cv2.FONT_HERSHEY_SIMPLEX, 0.8, (0, 0, 255))
        new_img = img[y:y+h, x:x+w]
    cv2.imshow("FaceDetection", img)
    key = cv2.waitKey(1) & 0xFF
```

Отримані фотографії користувачів потрібно перевести в numpy формат для подальшого тренування та створення класифікаторів для верифікації користувачів.

```

# Store images in a numpy format and ids of the user on the same index in imageNp and id lists

for root,dirs,files in os.walk(path):
    | pictures = files

for pic in pictures :

    | imgpath = path+pic
    | img = Image.open(imgpath).convert('L')
    | imageNp = np.array(img, 'uint8')
    | id = int(pic.split(name)[0])
    | #names[name].append(id)
    | faces.append(imageNp)
    | ids.append(id)

ids = np.array(ids)

```

Після цього потрібно навчити та зберегти класифікатори кожного користувача.

```

#Train and save classifier
clf = cv2.face.LBPHFaceRecognizer_create()
clf.train(faces, ids)
clf.write("./data/classifiers/"+name+"_classifier.xml")

```

Далі потрібно зробити детектор що буде використовувати готові класифікатори користувачів і саме розпізнавати їх.

```

for (x,y,w,h) in faces:
    roi_gray = gray[y:y+h,x:x+w]
    id,confidence = recognizer.predict(roi_gray)
    confidence = 100 - int(confidence)
    pred = 0
    if confidence > 50:
        #if u want to print confidence level
        #confidence = 100 - int(confidence)
        pred += +1
        text = name.upper()
        font = cv2.FONT_HERSHEY_PLAIN
        frame = cv2.rectangle(frame, (x, y), (x + w, y + h), (0, 255, 0), 2)
        frame = cv2.putText(frame, text, (x, y-4), font, 1, (0, 255, 0), 1, cv2.LINE_AA)
        frame = cv2.putText(frame, str(confidence), (x,y-20), font, 1, (0,255,0), 2)

    else:
        pred += -1
        text = "UnknownFace"
        font = cv2.FONT_HERSHEY_PLAIN
        frame = cv2.rectangle(frame, (x, y), (x + w, y + h), (0, 0, 255), 2)
        frame = cv2.putText(frame, text, (x, y-4), font, 1, (0, 0,255), 1, cv2.LINE_AA)

```

```

if cv2.waitKey(20) & 0xFF == ord('q'):
    print(pred)
    if pred > 0 :
        dim =(124,124)
        img = cv2.imread(f"\\data\\{name}\\{pred}{name}.jpg", cv2.IMREAD_UNCHANGED)
        resized = cv2.resize(img, dim, interpolation = cv2.INTER_AREA)
        cv2.imwrite(f"\\data\\{name}\\50{name}.jpg", resized)
        Image1 = Image.open(f"\\2.png")

        # make a copy the image so that the
        # original image does not get affected
        Image1copy = Image1.copy()
        Image2 = Image.open(f"\\data\\{name}\\50{name}.jpg")
        Image2copy = Image2.copy()

        # paste image giving dimensions
        Image1copy.paste(Image2copy, (195, 114))

        # save the image
        Image1copy.save("end.png")
        frame = cv2.imread("end.png", 1)

        cv2.imshow("Result",frame)
        cv2.waitKey(5000)
break

```

Наступним кроком, для верифікованих користувачів потрібно розробити систему розпізнавання емоційного стану, віку, статі та виділення ключових точок на обличчі людини.

Для визначення ключових точок на обличчі користувача використано додаток Dlib. Dlib - це популярний інструментарій для машинного навчання, який використовується переважно для задач комп'ютерного зору та обробки зображень, таких як розпізнавання облич, виявлення орієнтирів на обличчі, виявлення об'єктів тощо. Він написаний на C++, але має прив'язки до Python, що робить його легко доступним з коду Python. .

Детектор міток обличчя dlib повертає об'єкт `shape`, що містить 68 (x, y)-координат областей міток обличчя. Потім ми використовуємо утиліту `face_utils.shape_to_np` для перетворення результату в масив NumPy, що дозволяє легко роздрукувати виявлені орієнтири обличчя на зображенні для візуалізації.

```

def landmark():
    cap = cv2.VideoCapture(0)

    hog_face_detector = dlib.get_frontal_face_detector()

    dlib_facelandmark = dlib.shape_predictor("./data/shape_predictor_68_face_landmarks.dat")

    faces = hog_face_detector(gray)
    for face in faces:

        face_landmarks = dlib_facelandmark(gray, face)

        for n in range(0, 68):
            x = face_landmarks.part(n).x
            y = face_landmarks.part(n).y
            cv2.circle(frame, (x, y), 2, (0, 255, 255), 1)

    cv2.imshow("Face Landmarks", frame)

```

Для того щоб скористатися моделями для розпізнавання емоційного стану, віку, статі нам потрібно додати Keras. Keras - це API для глибокого навчання, написаний на Python і здатний працювати поверх JAX, TensorFlow або PyTorch. Моделі Keras пропонують простий і зручний спосіб визначити нейронну мережу, яка потім буде побудована для вас за допомогою TensorFlow.

```

.....
from keras.models import load_model
from keras.preprocessing.image import img_to_array

```

Для визначення емоцій використовується навчена модель mini_XCEPTION. Ця модель, яку навчено на 28 тис. фото, може в реальному часі класифікувати емоції, такі як щастя, смуток, страх, гнів, відраза, презирство, здивування та нейтральність. Модель автоматично визначає емоції та прогнозує емоційний стан з високою точністю.

```

emotion_model = "./data/_mini_XCEPTION.106-0.65.hdf5"

```

```
Emotions = ["angry", "disgust", "scared", "happy", "sad", "surprised", "neutral"]
```

```
emotion_classifier = load_model(emotion_model, compile=False)
```

```
def emotion():
    while True:
        ret, img = frame.read()
        gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
        face = face_cascade.detectMultiScale(image=gray, scaleFactor=1.3, minNeighbors=5)

        for x, y, w, h in face:
            roi = gray[y:y + h, x:x + w]
            roi = cv2.resize(roi, (48, 48))
            roi = roi.astype("float") / 255.0
            roi = img_to_array(roi)
            roi = np.expand_dims(roi, axis=0)
            preds = emotion_classifier.predict(roi)[0]
            emotion_probability = np.max(preds)
            label = Emotions[preds.argmax()]
            cv2.rectangle(img, (x, y), (x + w, y + h), (0, 255, 0), 2)
            cv2.putText(img, f"{label}", (x, y - 5), cv2.FONT_HERSHEY_SIMPLEX, 0.8, (0, 255, 0))
            #cv2.face.drawFacemarks(img,)

        cv2.imshow("Emotion Prediction", img)
```

Такі ж самі дії виконано для реалізації розпізнавання емоцій із відео.

```
def emotionvideo():
    f = open('result.txt', 'w')
    while cap.isOpened():
        ret, img = cap.read()
        gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
        face = face_cascade.detectMultiScale(image=gray, scaleFactor=1.3, minNeighbors=5)

        for x, y, w, h in face:
            roi = gray[y:y + h, x:x + w]
            roi = cv2.resize(roi, (48, 48))
            roi = roi.astype("float") / 255.0
            roi = img_to_array(roi)
            roi = np.expand_dims(roi, axis=0)
            preds = emotion_classifier.predict(roi)[0]
            emotion_probability = np.max(preds)
            label = Emotions[preds.argmax()]
            cv2.rectangle(img, (x, y), (x + w, y + h), (0, 255, 0), 2)
            cv2.putText(img, f"{label}", (x, y - 5), cv2.FONT_HERSHEY_SIMPLEX, 0.8, (0, 255, 0))
            #cv2.face.drawFacemarks(img,)
            print(label)
            f.write(label + '\n')

        cv2.imshow("Emotion Prediction", img)
```

Щоб визначити вік та стать використовуються натреновані моделі від фреймворку Caffe. Caffe (Convolutional Architecture for Fast Feature Embedding)

- це фреймворк для глибокого навчання, який підтримує різні архітектури глибокого навчання, такі як CNN, RCNN, LSTM і повністю зв'язані мережі. Завдяки підтримці графічних процесорів (GPU) і готовим шаблонам, які спрощують налаштування і навчання моделі, Caffe є найбільш популярним для задач класифікації і сегментації зображень. Gender_net.caffemodel: попередньо навчені ваги моделі для визначення статі. Deploy_gender.prototxt: архітектура моделі для моделі визначення статі. Age_net.caffemodel: попередньо навчені ваги моделі для визначення віку. Deploy_age.prototxt: архітектура моделі для моделі визначення віку.

```
ageProto="./data/age_deploy.prototxt"
ageModel="./data/age_net.caffemodel"
genderProto="./data/gender_deploy.prototxt"
genderModel="./data/gender_net.caffemodel"

MODEL_MEAN_VALUES=(78.4263377603, 87.7689143744, 114.895847746)
ageList=['(0-2)', '(4-6)', '(8-12)', '(15-20)', '(20-32)', '(38-43)', '(48-53)', '(60-100)']
genderList=['Male', 'Female']

def ageAndgender():
    while True:
        net, img = frame.read()
        default_img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
        face = face_cascade.detectMultiScale(image=default_img, scaleFactor=1.3, minNeighbors=5)
        for x, y, w, h in face:
            roi = default_img[y:y + h, x:x + w]
            blob = cv2.dnn.blobFromImage(roi, 1.0, (227, 227), MODEL_MEAN_VALUES, swapRB=False)
            genderNet.setInput(blob)
            genderPreds = genderNet.forward()
            gender = genderList[genderPreds[0].argmax()]
            ageNet.setInput(blob)
            agePreds = ageNet.forward()
            age = ageList[agePreds[0].argmax()]
            cv2.rectangle(img, (x, y), (x + w, y + h), (255, 0, 0), 2)
            cv2.putText(img, f"{gender}, {age} year", (x, y - 5), cv2.FONT_HERSHEY_SIMPLEX, 0.8, (0, 0, 255))

        cv2.imshow("Gender and Age Prediction", img)
```

Для реалізації графічного інтерфейсу обрано Tkinter[23]. Python пропонує декілька варіантів розробки графічного інтерфейсу користувача (GUI). З усіх методів розробки графічного інтерфейсу tkinter є

найпоширенішим методом. Це стандартний інтерфейс Python до інструментарію Tk GUI, що постачається з Python. Python tkinter - це найшвидший і найпростіший спосіб створення GUI-додатків.

```
import tkinter as tk
from tkinter import font as tkfont
from tkinter import messagebox, PhotoImage
```

```
class MainUI(tk.Tk):
    def __init__(self, *args, **kwargs):
        tk.Tk.__init__(self, *args, **kwargs)
        global names
        with open("nameslist.txt", "r") as f:
            x = f.read()
            z = x.rstrip().split(" ")
            for i in z:
                names.add(i)
        self.title_font = tkfont.Font(family='Helvetica', size=16, weight="bold")
        self.title("Face Recognizer")
        self.resizable(False, False)
        self.geometry("500x250")
        self.protocol("WM_DELETE_WINDOW", self.on_closing)
        self.active_name = None
        container = tk.Frame(self)
        container.grid(sticky="nsew")
        container.grid_rowconfigure(0, weight=1)
        container.grid_columnconfigure(0, weight=1)
        self.frames = {}
        for F in (StartPage, PageOne, PageTwo, PageThree, PageFour):
            page_name = F.__name__
            frame = F(parent=container, controller=self)
            self.frames[page_name] = frame
            frame.grid(row=0, column=0, sticky="nsew")
        self.show_frame("StartPage")
```

Повний код реалізації програми надано в додатку А.

4.3 Огляд додатку

Після відкриття програми наявне вікно, де буде запропоновано додати нового користувача, увійти в додаток вже існуючим користувачем або закрити програму. Також доступна вибіркова можливість залучення нових користувачів чи використання існуючих облікових записів.

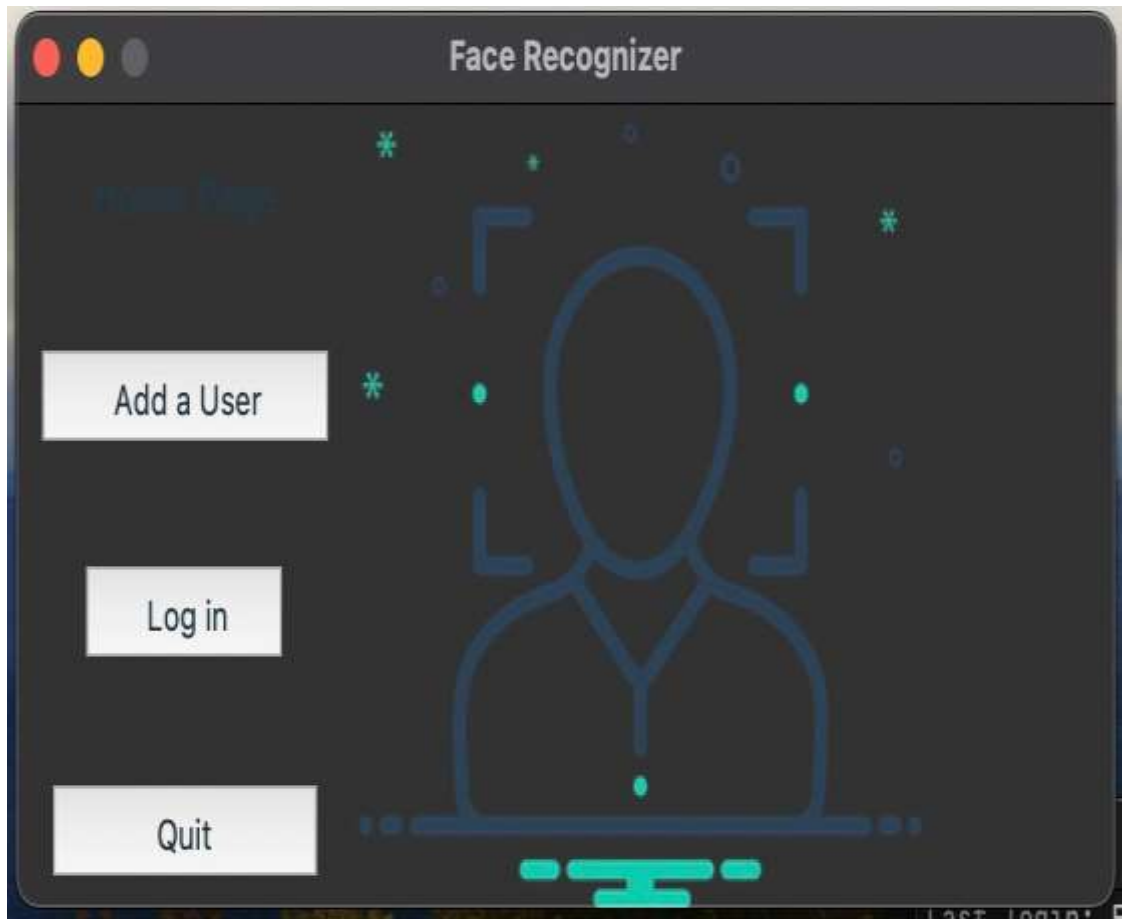


Рис. 22 – Початкове вікно

При натисканні на кнопку "Add User" буде запропоновано ввести ім'я для нового користувача. Після введення імені надається можливість продовжити налаштування профілю та використання функцій додатку.

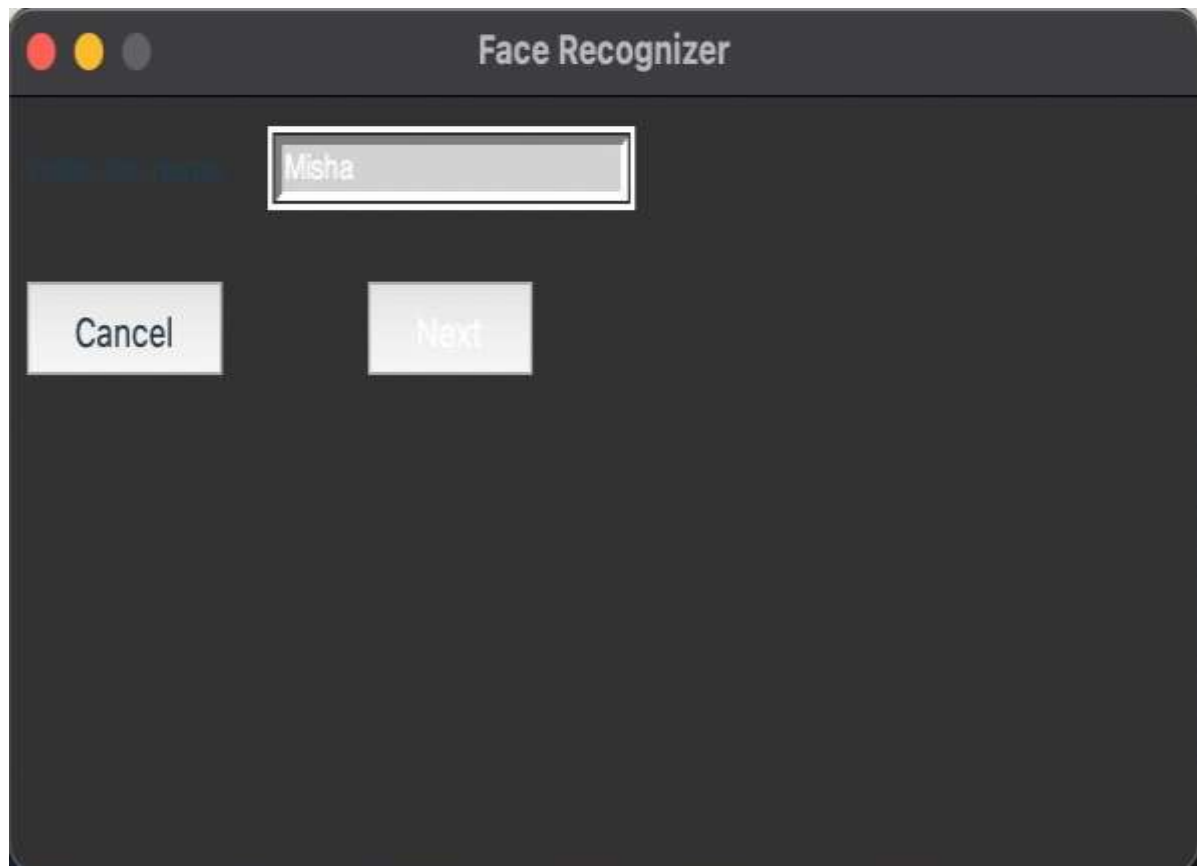


Рис. 23 – Додавання нового користувача

Після вказання імені користувача, буде запропоновано зібрати набір фотографій для тренування моделі. Якщо ж такий набір знімків вже наявний, можна скористатися навчанням моделі за допомогою цих фотографій.

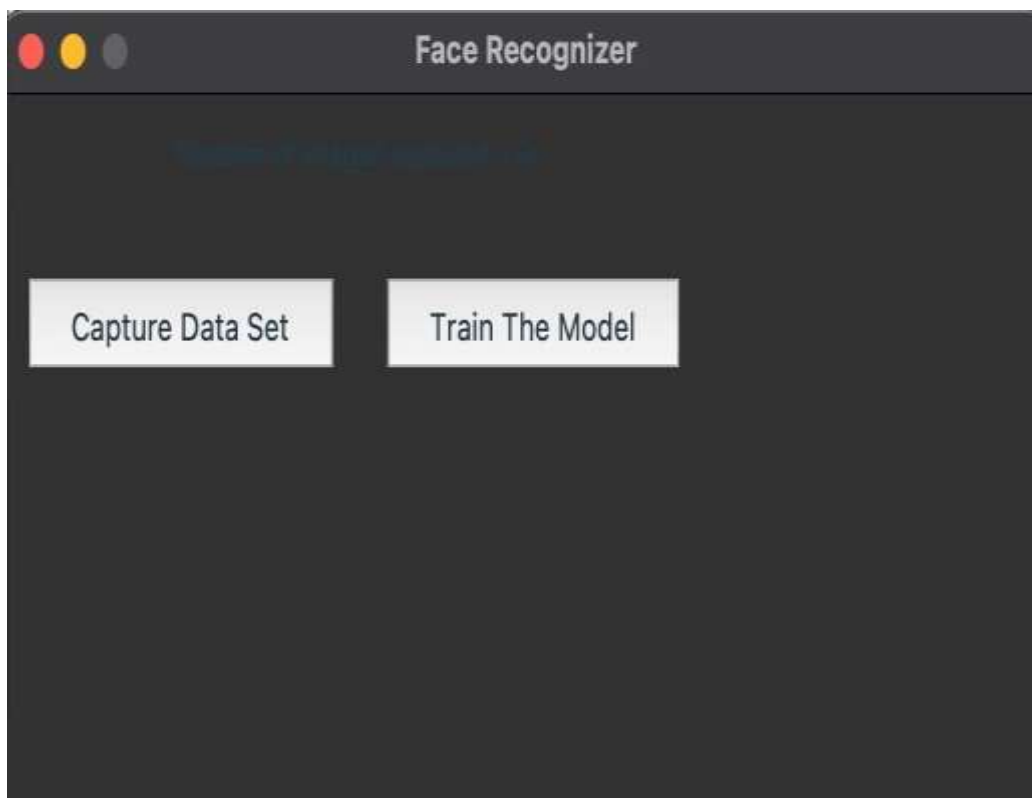


Рис. 24 – Фото нового користувача або навчання моделі

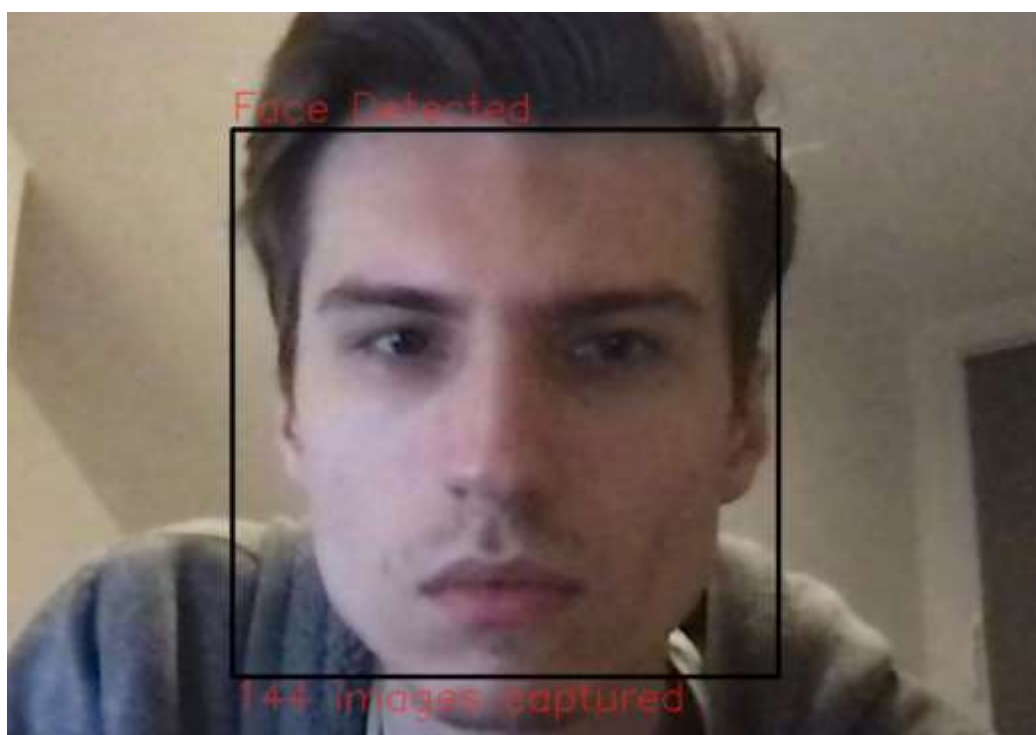


Рис. 25 – Процес збору фотографій користувача

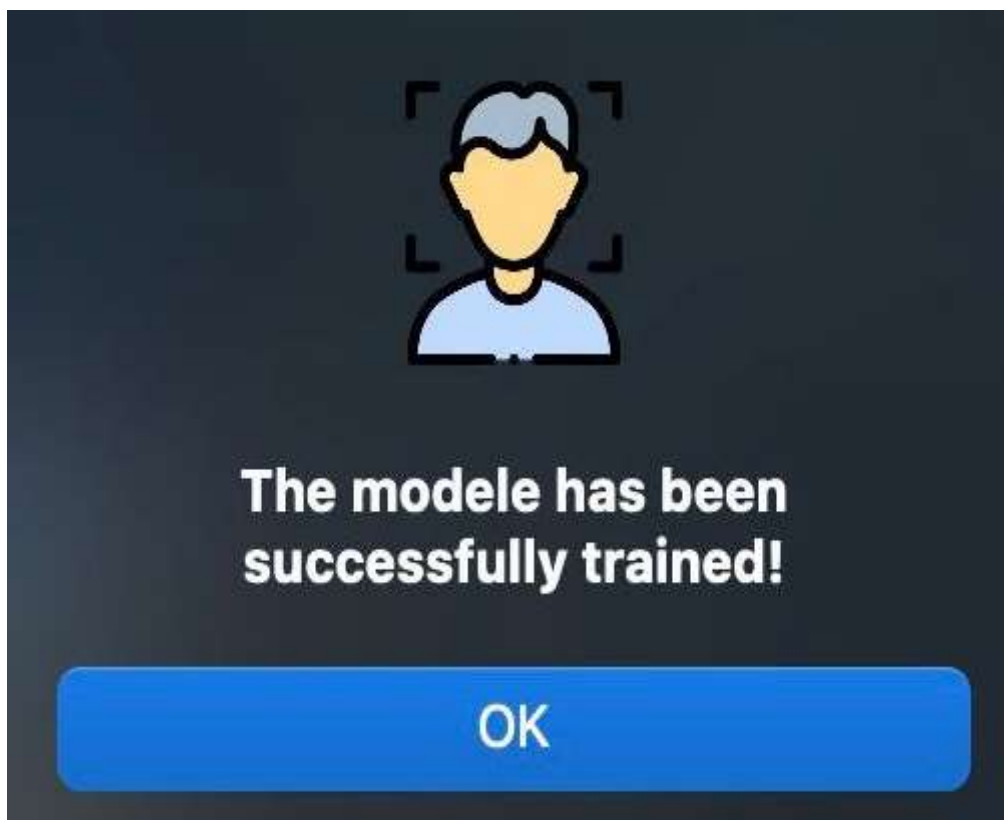


Рис. 26 – Повідомлення про успішно навчану модель

Завдяки успішно навченій моделі, користувач може увійти в додаток, просто натиснувши кнопку "Log in".

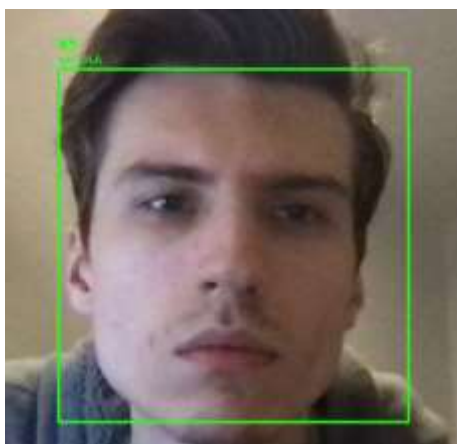


Рис. 27 – Процес верифікації користувача

Після успішної верифікації, користувач отримує повідомлення, підтверджуюче успіх проходження процедури верифікації.

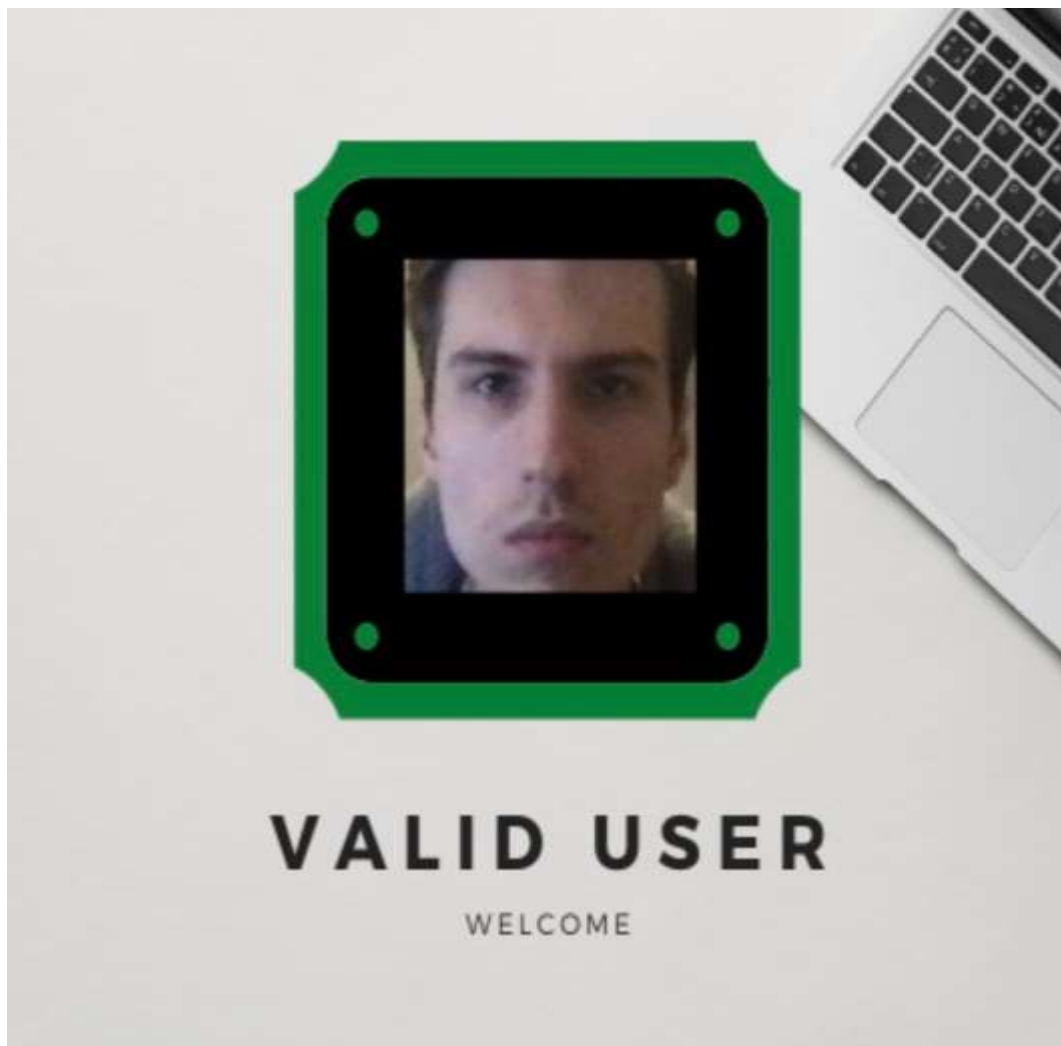


Рис. 28 – Повідомлення про успішний вхід в систему

Якщо спроба не вдається, користувач отримує відповідне повідомлення, що допоможе йому розібратися в ситуації (див. рис. 27).

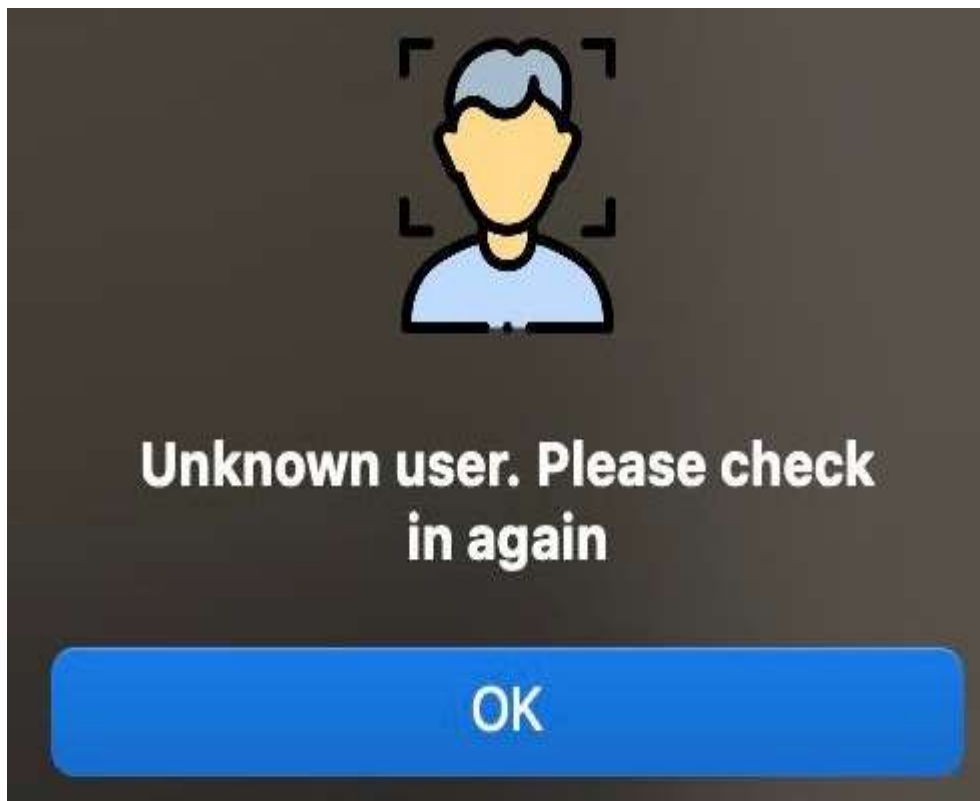


Рис. 29 – Невдала спроба верифікації

У разі успішної верифікації користувача, що підтверджує його ідентичність, відкривається головне меню програми. Тут користувач може не лише перевірити свої настрій в режимі реального часу, але й взяти участь у визначенні віку та гендерної приналежності за допомогою аналізу обличчя. За бажанням, він може використовувати функціонал для пошуку ключових точок на обличчі або переглядати свій емоційний стан на окремому відеозаписі. Крім того, доступна кнопка, що дозволяє знову перейти до початкового екрану програми, щоб користувач міг зручно використовувати різноманітні можливості додатку.

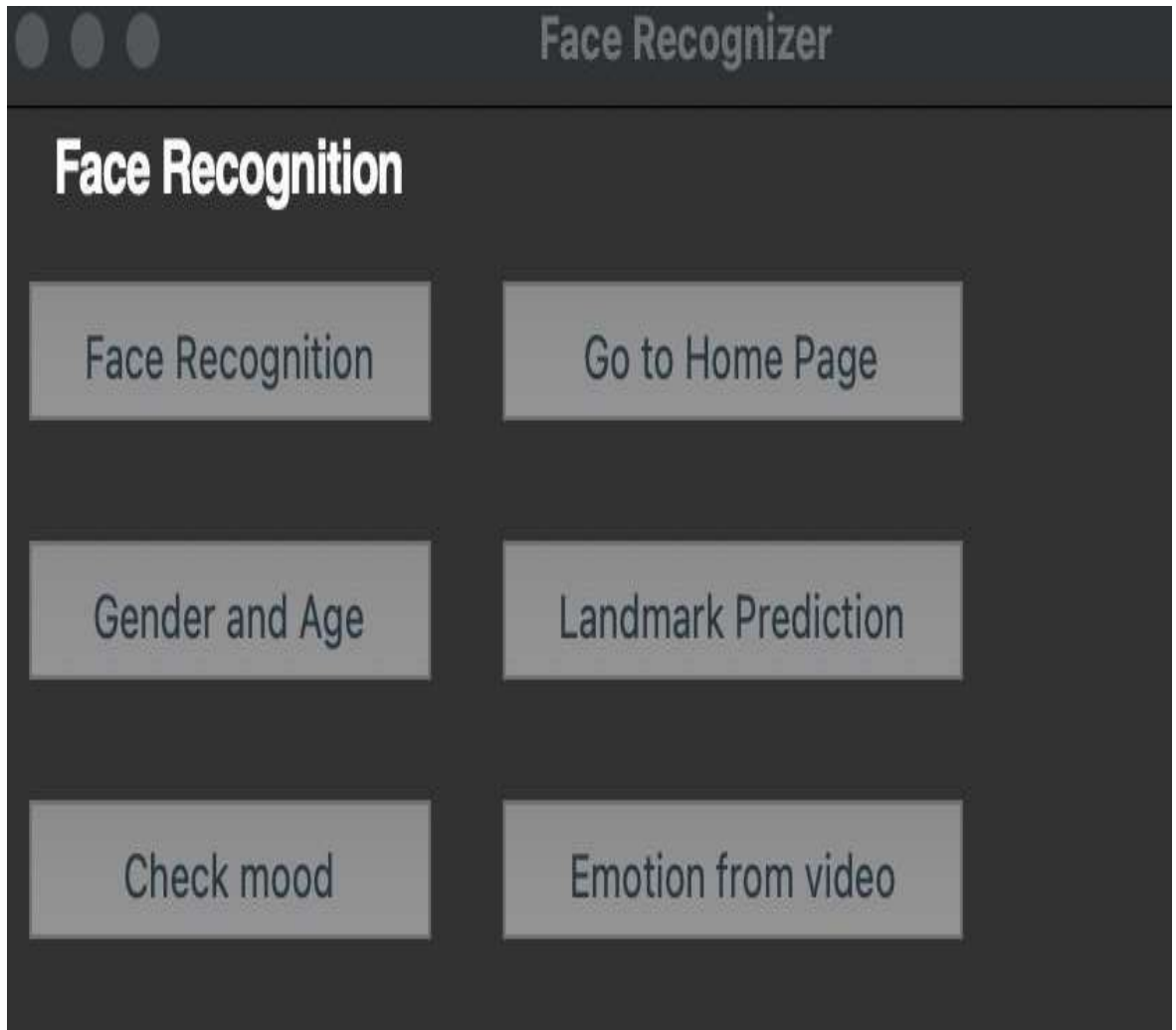


Рис. 30 – Функціонал що надається користувачу після верифікації

Коли користувач натискає кнопку "Check mood", відкривається можливість визначити його емоційний стан. На прикладах, представлених на рис. 29-31, демонструється цей процес.

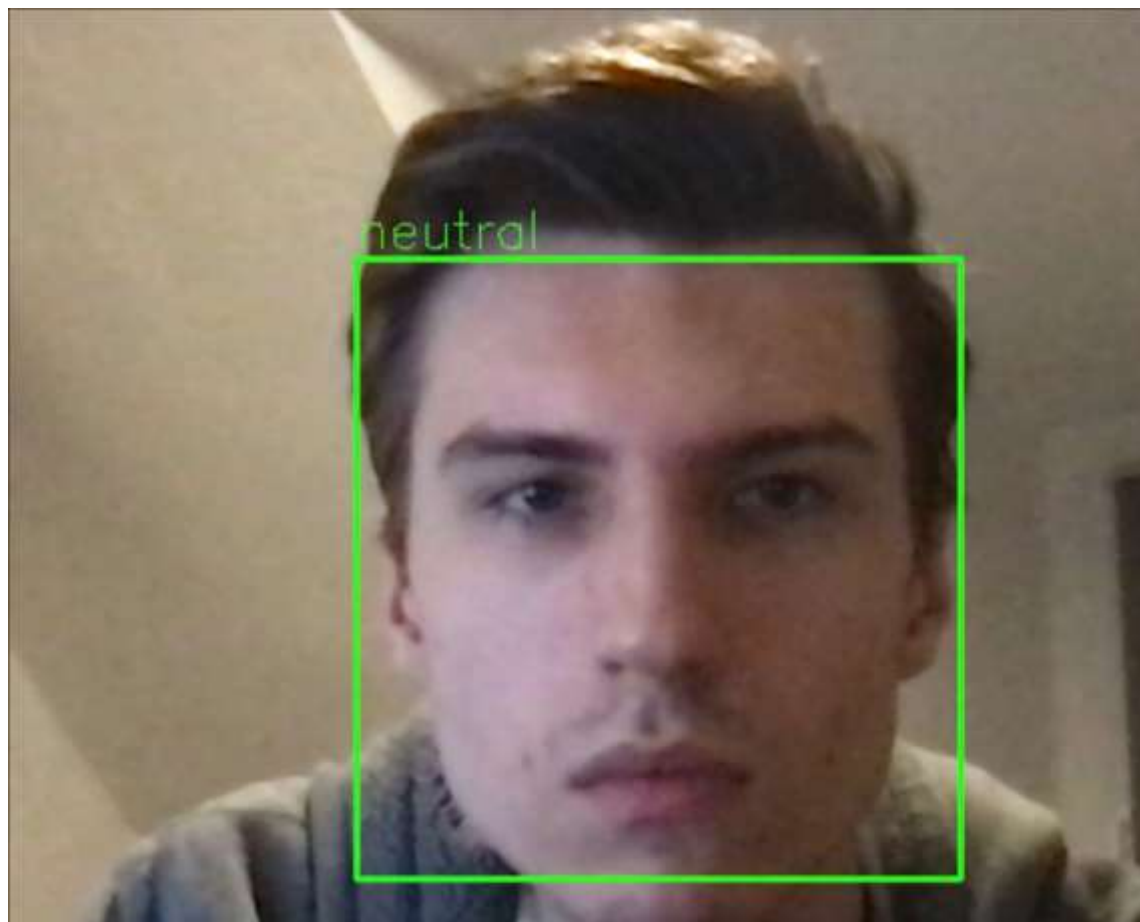


Рис. 31 – Приклад роботи розпізнавання емоцій(нейтральний стан)

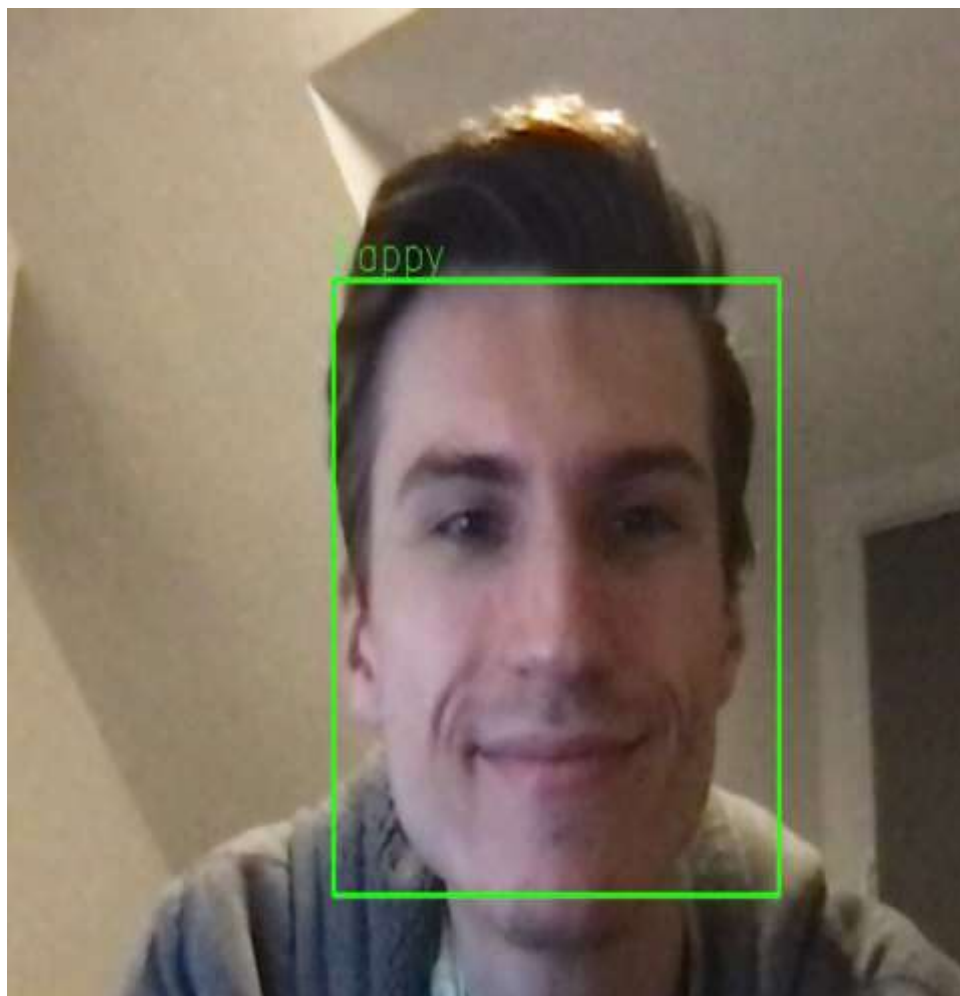


Рис. 32 – Приклад роботи розпізнавання емоцій(щасливий)

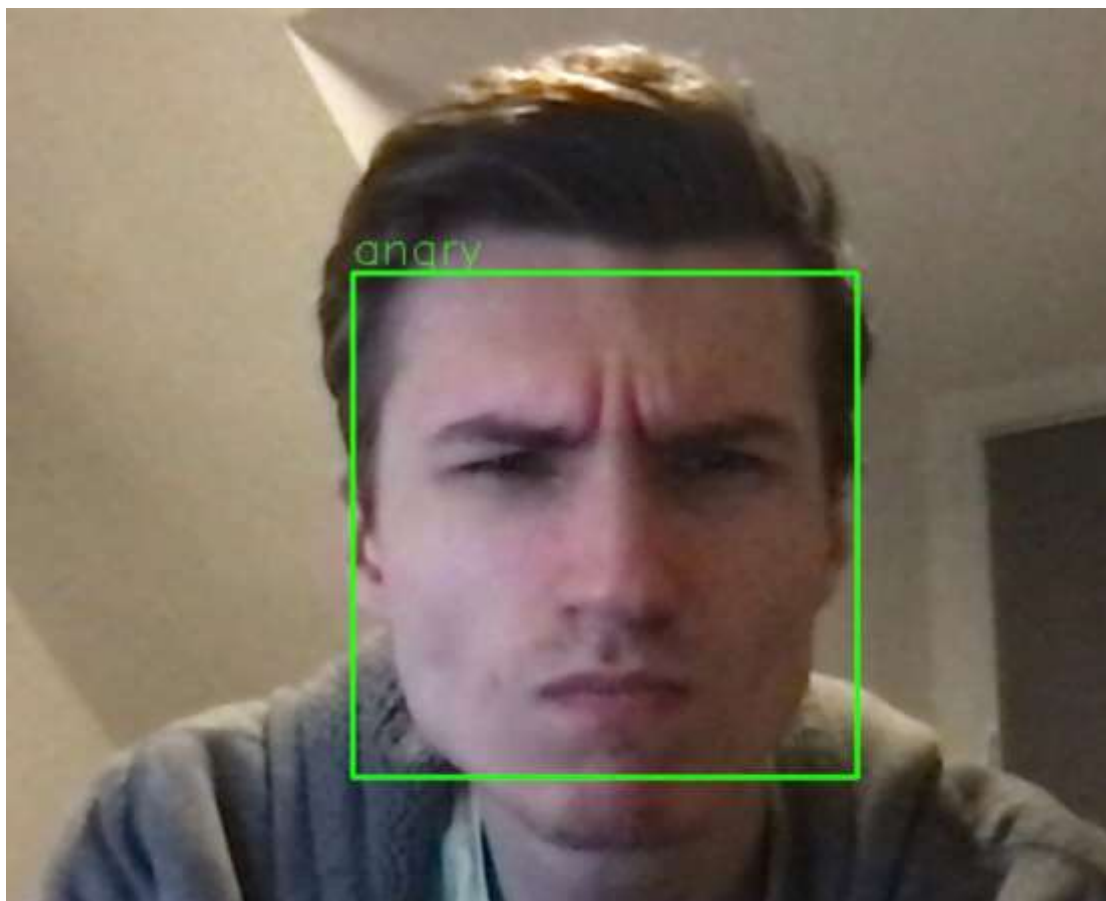


Рис. 33 – Приклад роботи розпізнавання емоцій(розгніваний)

У випадку вибору опції перевірки віку та статі, система відкриває камеру та проводить аналіз зображення користувача, як показано на рис.32.

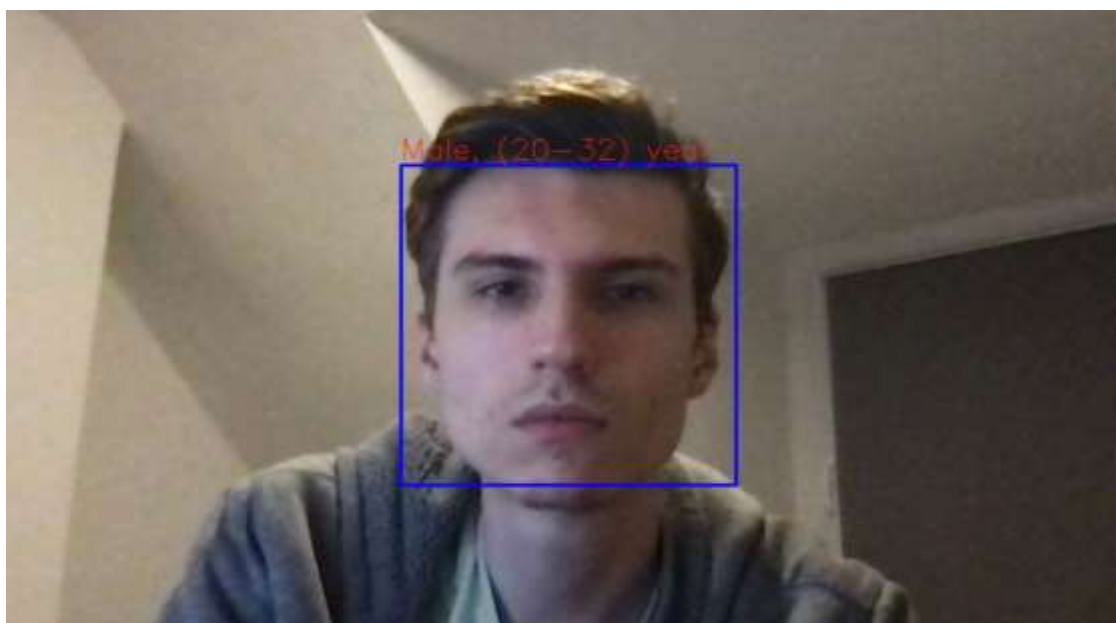


Рис. 34 – Приклад розпізнавання віку та статі користувача

Користувач може також спостерігати за розташуванням ключових точок обличчя в режимі реального часу, просто взаємодіючи з функцією "Landmark Prediction" натисканням відповідної кнопки.



Рис. 35 – Приклад пошуку ключових точок на обличчі

Так само, як користувач може відслідковувати свій емоційний стан у реальному часі, він може також використовувати цю можливість з відеозапису, просто натисканням кнопки "Emotion from video".

Після успішної верифікації та введення імені користувача, в програмі автоматично реєструються інформація про користувача, включаючи ім'я, вік, стать та емоційний стан. Ці дані зберігаються в окремий файл для подальшого використання та можуть бути виведені в консоль із заданим інтервалом в одну секунду. Деталі цього процесу можна знайти на рис. 34.

A screenshot of a terminal window with a dark background. The window title bar shows three window control buttons on the left and a document icon with the text 'result' on the right. The terminal output is as follows:

```
Username: Misha

Gender and age: Male (20-32)

Mood:

neutral      1s
neutral      2s
neutral      3s
neutral      4s
neutral      5s
happy        6s
happy        7s
happy        8s
happy        9s
happy       10s
happy       11s
happy       12s
happy       13s
happy       14s
neutral     15s
neutral     16s
neutral     17s
neutral     18s
neutral     19s
neutral     20s
```

Рис. 36 – Приклад отриманих даних після використання програми

4.4 Висновки до четвертого розділу

У четвертому розділі викладено реалізацію конкретних аспектів додатку, супроводжуючи це кодовими прикладами для кращого розуміння. Окрім того, здійснено докладний огляд функціоналу програми, і наведено приклади її роботи для кращого розуміння.

ВИСНОВКИ

В результаті виконання магістерської роботи розроблено програму верифікації і збору біометричних даних користувача за обличчям.

Під час ретельного аналізу предметної області вивчались різні сфери використання систем верифікації та збору біометричних даних людини. Особлива увага була приділена актуальності даної теми. Проведено дослідження різноманітних методів та алгоритмів для розпізнавання та верифікації осіб за обличчям, враховуючи виклики та проводячи аналіз наявних рішень.

У контексті порівняльного аналізу систем розпізнавання та верифікації осіб були розглянуті як їх переваги, так і недоліки. Вибір методів та інструментів для розробки додатку був обґрунтований, і мова програмування Python разом із бібліотеками OpenCV, Keras, Tkinter, Dlib, а також використанням певних навчених моделей для збору біометричних даних стали обраними для реалізації проекту.

Розроблений додаток включає кілька методів розпізнавання, відзначається високою надійністю та простотою в користуванні. Результати тестування підтвердили правильну функціональність всіх методів. Зазначений програмний продукт проявив себе як універсальний та здатний впоратися з різноманітними завданнями.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Метод Віоли-Джонса [Електронний ресурс]

https://uk.wikipedia.org/wiki/%D0%A1%D0%B8%D1%81%D1%82%D0%B5%D0%BC%D0%B0_%D0%92%D1%96%D0%BE%D0%BB%D0%B8_%E2%80%94%D0%94%D0%B6%D0%BE%D0%BD%D1%81%D0%B0_%D0%B2%D0%B8%D1%8F%D0%B2%D0%BB%D1%8F%D0%BD%D0%BD%D1%8F_%D0%BE%D0%B1%27%D1%94%D0%BA%D1%82%D1%96%D0%B2

2. Гістограма орієнтованих градієнтів (HOG) [Електронний ресурс]
https://uk.wikipedia.org/wiki/%D0%93%D1%96%D1%81%D1%82%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%B0_%D0%BD%D0%B0%D0%BF%D1%80%D1%8F%D0%BC%D0%BB%D0%B5%D0%BD%D0%B8%D1%85_%D0%B3%D1%80%D0%B0%D0%B4%D1%96%D1%94%D0%BD%D1%82%D1%96%D0%B2
3. Фільтр Гаусса [Електронний ресурс]
https://uk.wikipedia.org/wiki/%D0%A4%D1%96%D0%BB%D1%8C%D1%82%D1%80_%D0%93%D0%B0%D1%83%D1%81%D0%B0
4. FAR, FRR та ERR [Електронний ресурс] <https://www.bayometric.com/false-acceptance-rate-far-false-recognition-rate-frr/>
5. Система FaceVACS [Електронний ресурс]
<https://www.biometricupdate.com/201601/cognitec-exec-discusses-facevacs-facial-recognition-technology>
6. NEC's Face Recognition [Електронний ресурс]
<https://www.nec.com/en/global/solutions/biometrics/face/index.html#:~:text=NEC's%20technology%20can%20match%20a,recognition%20in%20poorly%20lit%20areas.>
7. Алгоритм PSM [Електронний ресурс]
https://en.wikipedia.org/wiki/Propensity_score_matching
8. Система VeriLook SDK [Електронний ресурс]
https://visagetechologies.com/face-recognition/?ppc_keyword=face%20recognition%20sdk&gclid=Cj0KCQiAtOmsBhCnARIsAGPa5yZlJvWlqnkalrzbKJAa_FEYI9IAJ-6ksD8wVn-p91-c5ZFx0eKuV_QaAtNuEALw_wcB
9. Система FaceReader [Електронний ресурс]
https://www.noldus.com/facereader?utm_term=facereader%20noldus&utm_campaign=OTM%7CHuman%7CFaceReader%7CExact&utm_source=adwords&utm_medium=ppc&hsa_acc=5401040478&hsa_cam=1382615190&hsa_grp=54748186277&hsa_ad=627400602672&hsa_src=g&hsa_tgt=kwd-466715666314&hsa_kw=facereader%20noldus&hsa_mt=e&hsa_net=adwords&hsa_ver=3&gad_source=1&gclid=Cj0KCQiAtOmsBhCnARIsAGPa5yZ20kP

[hdH_oh3j6Mgzj2oibSF5-7Yo-p5WEAYHm_bJ2_YhmDC6OShwaArAvEALw_wcB](#)

10. Viola P., Jones M.J. Rapid object detection using a boosted cascade of simple features // IEEE Conf. on Computer Vision and Pattern Recognition. – Kauai, Hawaii, USA, 2001.
11. Viola P., Jones M.J. Robust real time face detection // International Journal of Computer Vision. – 2004.
12. Bradski G., Kaebler A. Learning OpenCV. Published by O'Reilly Media, 2008. P. 495–512.
13. Gonsales P., Vuds R. Tsifrovaia obrabotka izobrazhenii [Digital processing of the image]. Tekhnosfera Publ., 2005. P. 1072
14. Goncharov A.V., Karkishchenko A.N. Vliyanie osveshchennosti na kachestvo raspoznavaniya frontal'nykh lits [Influence of illumination on a quality of frontal face recognition]. Izvestiya Yuzhnogo federal'nogo universiteta. Tekhnicheskie nauki – Izvestiya Southern Federal University. Engineering sciences, 2008, no. 4 (81), pp. 88–92
15. Лінійний дискримінантний аналіз [Електронний ресурс]
[https://en.wikipedia.org/wiki/Linear_discriminant_analysis#:~:text=Linear%20discriminant%20analysis%20\(LDA\)%2C,classes%20of%20objects%20or%20events](https://en.wikipedia.org/wiki/Linear_discriminant_analysis#:~:text=Linear%20discriminant%20analysis%20(LDA)%2C,classes%20of%20objects%20or%20events).
16. Фільтр Габора [Електронний ресурс]
https://uk.wikipedia.org/wiki/%D0%A4%D1%96%D0%BB%D1%8C%D1%82%D1%80_%D2%90%D0%B0%D0%B1%D0%BE%D1%80%D0%B0
17. Згорткова нейронна мережа [Електронний ресурс]
https://uk.wikipedia.org/wiki/%D0%97%D0%B3%D0%BE%D1%80%D1%82%D0%BA%D0%BE%D0%B2%D0%B0_%D0%BD%D0%B5%D0%B9%D1%80%D0%BE%D0%BD%D0%BD%D0%B0_%D0%BC%D0%B5%D1%80%D0%B5%D0%B6%D0%B0
18. A.K. Jain, Technology: biometric recognition. Nature 449, 38–40 (2007).
19. OpenCV [Електронний ресурс] <https://www.geeksforgeeks.org/opencv-overview/>
20. Tensorflow [Електронний ресурс]
<https://www.infoworld.com/article/3278008/what-is-tensorflow-the-machine-learning-library-explained.html>
21. Python (programming language). [Електронний ресурс]
[https://en.wikipedia.org/wiki/Python_\(programming_language\)](https://en.wikipedia.org/wiki/Python_(programming_language))
22. Keras [Електронний ресурс] <https://blog.skillfactory.ru/glossary/keras/>

23. Tkinter [Электронный ресурс] <https://www.geeksforgeeks.org/python-gui-tkinter/>