Хмельницький національний університет
Факультет інформаційних технологій
Кафедра комп'ютерних наук

КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА

на тему *Розпізнавання мімічних проявів емоцій засобами штучного інтелекту*

Галузь знань *12 – Інформаційні технології*
Шифр і назва галузі знань

Спеціальність *122 – Комп'ютерні науки*
Шифр і назва спеціальності

Освітня програма *Комп'ютерні науки*
Назва освітньої програми

Виконав: *студент 4 курсу, група КН-19-2* — Курс, група виконавця | Підпис | *А.В. Смотрицький* — Ініціали, прізвище

Керівник: *док. філ., ст. викладач кафедри КН* — Науковий ступінь, посада | Підпис | *П.М. Радюк* — Ініціали, прізвище

Нормоконтроль: *к.т.н., доцент кафедри КН* — Науковий ступінь, посада | Підпис | *Р.О. Багрій* — Ініціали, прізвище

До захисту допускаю:
Зав. кафедри КН, д.т.н., професор — Підпис | *О.В. Бармак* — Ініціали, прізвище

01 06 2023 р.

Хмельницький 2023

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет *інформаційних технологій*
Кафедра *комп'ютерних наук*
Освітній ступінь *бакалавр*
Галузь знань *12 – Інформаційні технології*
Спеціальність *122 – Комп'ютерні науки*
Освітня програма *освітньо-професійна програма підготовки бакалавра*

ЗАТВЕРДЖУЮ
Завідувач кафедри *комп'ютерних наук*

_____
(підпис)
д.т.н., професор О.В. Бармак
«06» 03 2023 року

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ БАКАЛАВРА

1. Тема кваліфікаційної роботи бакалавра: *«Розпізнавання мімічних проявів емоцій засобами штучного інтелекту»*

2. Завдання видано студенту *Смотрицькому Антону Валентиновичу*
(прізвище, ім'я, по батькові)

3. Керівник роботи *ст. викладач кафедри КН Радюк Павло Михайлович*
(посада, прізвище, ім'я, по батькові)

4. Затверджено наказом університету від «01» 03 2023 р. № 5

5. Дата видачі завдання студенту: «03» 03 2023 р.

6. Зміст пояснювальної записки (перелік задач) та вихідні дані:

Провести аналіз способів та підходів до ідентифікації змін емоційного стану людини за мімічними проявами та обрати найкращий; застосувати обраний спосіб до розв'язання задачі автоматизованого розпізнавання мімічних проявів емоцій; реалізувати обраний спосіб у вигляді інформаційної системи автоматизованого розпізнавання мімічних проявів емоцій; провести експериментальне тестування інформаційної системи за еталонними наборами даних; вихідними даними є ідентифіковані емоційні стани людини за її фото- та відеоматеріалом.

7. Календарний план виконання кваліфікаційної роботи бакалавра:

№	Назва етапів (розділів) кваліфікаційної роботи бакалавра	Термін виконання	Примітка
1	*Вибір напряму дослідження та узгодження тематики кваліфікаційної роботи бакалавра з керівником*	*грудень 2022*	виконано
2	*Ознайомлення з предметною областю, формулювання мети та задач дослідження, визначення об'єкта та предмета дослідження*	*січень 2023*	виконано
3	*Робота над розділом 1 – Дослідження сучасних способів та підходів до розпізнавання мімічних проявів емоцій*	*січень 2023*	виконано
4	*Робота над розділом 2 – Спосіб розпізнавання мімічних проявів емоцій засобами штучного інтелекту*	*березень 2023*	виконано
5	*Робота над розділом 3 – Програмна реалізація інформаційної системи розпізнавання мімічних проявів емоцій*	*квітень 2023*	виконано
6	*Оформлення пояснювальної записки згідно вимог*	*травень 2023*	виконано
7	*Попередній захист кваліфікаційної роботи бакалавра*	*травень 2023*	виконано
8	*Захист кваліфікаційної роботи бакалавра на засіданні Екзаменаційної комісії*	*червень 2023*	виконано

Виконавець: *студент 4 курсу, група КН-19-2* (Курс, група виконавця) _____ (Підпис) *А.В. Смотрицький* (Ініціали, прізвище)

Керівник: *док. філ., ст. викладач кафедри КН* (Науковий ступінь, посада) _____ (Підпис) *П.М. Радюк* (Ініціали, прізвище)

Анотація

Тема кваліфікаційної роботи бакалавра: *Розпізнавання мімічних проявів емоцій засобами штучного інтелекту*

Виконавець кваліфікаційної роботи бакалавра: *студент групи КН-19-2 Смотрицький Антон Валентинович*

Керівник кваліфікаційної роботи бакалавра: *доктор філософії, ст. викладач кафедри КН Радюк Павло Михайлович*

Кваліфікаційна робота бакалавра містить:

Пояснювальна записка				Кількість додатків
Сторінок	Рисунків	Таблиць	Джерел інформації	
65	32	2	19	5

Метою кваліфікаційної роботи бакалавра є підвищення точності ідентифікації змін емоційного стану людини за мімічними проявами за цифровими фото- та відеоматеріалами.

Зміст пояснювальної записки (перелік задач) та вихідні дані: провести аналіз способів та підходів до ідентифікації змін емоційного стану людини за мімічними проявами та обрати найкращий; застосувати обраний спосіб до розв'язання задачі автоматизованого розпізнавання мімічних проявів емоцій; реалізувати обраний спосіб у вигляді інформаційної системи; провести експериментальне тестування інформаційної системи за еталонними наборами даних; вихідними даними є ідентифіковані емоційні стани людини за її фото- та відеоматеріалом.

Досягнення мети кваліфікаційної роботи бакалавра полягає у створенні інформаційної системи, використання якої дасть змогу покращити ідентифікацію змін емоційного стану людини за мімічними проявами за цифровими фото- та відеоматеріалами.

Ключові слова: ідентифікація змін емоційного стану, мімічні прояви, штучний інтелект, згорткова нейронна мережа.

Виконавець: *студент 4 курсу, група КН-19-2* (Курс, група виконавця) _____ (Підпис) *А.В. Смотрицький* (Ініціали, прізвище)

Зміст

Перелік скорочень ... 3

Вступ ... 4

Розділ 1 Дослідження сучасних способів та підходів до розпізнавання мімічних проявів емоцій ... 6

1.1 Актуальність застосування інформаційних технологій у системах розпізнавання візуальних образів ... 6

1.2. Сучасний стан розпізнавання змін емоційних станів людини ... 9

1.3 Аналіз засобів та підходів штучного інтелекту до виявлення мімічних проявів людини ... 13

1.4 Мета, завдання та вимоги до реалізації інформаційної системи ... 17

Розділ 2 Спосіб розпізнавання мімічних проявів емоцій засобами штучного інтелекту ... 18

2.1 Модель розпізнавання мімічних проявів емоцій ... 18

2.2 Спосіб розпізнавання мімічних проявів емоцій ... 23

2.3 Спосіб класифікації емоційного стану за мімічними проявами ... 24

2.4 Функціональна структура інформаційної системи ... 30

2.5 Проєктування інтерфейсу інформаційної системи ... 37

2.6 Висновки до розділу 2 ... 39

Розділ 3 Програмна реалізація інформаційної системи розпізнавання мімічних проявів емоцій ... 40

3.1 Структура та функціональне призначення програмних складових інформаційної системи ... 40

3.2 Особливості реалізації програмних складових інформаційної системи ... 44

3.3 Експериментальне тестування інформаційної системи ... 47

3.3.1 Тестування системи за валідаційним набором даних ... 47

3.3.2 Тестування системи за реальних умов використання ... 50

3.4 Вимоги до розгортання інформаційної системи та інструкція користувача 54

3.5 Висновки до розділу 3 ... 62

Висновки ... 63

Перелік посилань ... 64

Додатки

Перелік скорочень

Скорочення, термін, позначення	Пояснення
КН	Комп’ютерні науки
ШІ	Штучний інтелект
ІТ	Інформаційні технології
FACS	Facial Action Coding System
AU	Action Units
EMFACS	Emotional Facial Action Coding System
CNN	Convolutional Neural Network
DT	Decision Tree
SVM	Support Vector Machine
ЗНМ	Згорткова нейронна мережа
API	Application Programming Interface
ПЗ	Програмне забезпечення

Вступ

Кваліфікаційна робота бакалавра присвячена розв’язанню задачі розпізнавання мімічних проявів емоцій засобами штучного інтелекту.

Актуальність. Розпізнавання мімічних проявів емоцій за допомогою штучного інтелекту (ШІ) актуальне для різноманітних застосувань, включаючи взаємодію людини з комп’ютером, діагностику психічного здоров’я та емоційну аналітику в маркетингу та рекламі. Допомагає покращити користувацький досвід в технологічних інтерфейсах. У сфері психічного здоров’я розпізнавання мімічних проявів емоцій можна використовувати для оцінки та моніторингу емоційних станів, що може допомогти в діагностиці та лікуванні розладів психічного здоров’я. У маркетингу та рекламі емоційну аналітику можна використовувати для розуміння реакцій та настроїв споживачів і підвищення ефективності рекламних кампаній.

Таким чином, розпізнавання мімічних проявів емоцій має важливе значення для емоційного аналізу, досліджень психологічних станів та настроїв, маркетингу, розвитку технологій та покращення охорони здоров’я.

Об’єкт дослідження – процес ідентифікації змін емоційного стану людини за мімічними проявами з використанням систем фото- та відеоспостереження.

Предмет дослідження – методи, способи та підходи до ідентифікації змін емоційного стану людини за мімічними проявами за цифровими фото- та відеоматеріалами.

Мета кваліфікаційної роботи бакалавра – підвищення точності ідентифікації змін емоційного стану людини за мімічними проявами за цифровими фото- та відеоматеріалами.

Завдання кваліфікаційної роботи бакалавра – для досягнення поставленої мети визначено наступні завдання:

1. Провести аналіз методів, способів та підходів до ідентифікації змін емоційного стану людини за мімічними проявами та обрати найкращий.

2. Застосувати обраний спосіб ідентифікації змін емоційного стану людини за мімічними проявами до розв’язання задачі автоматизованого розпізнавання мімічних проявів емоцій.

3. Реалізувати обраний спосіб у вигляді інформаційної системи автоматизованого розпізнавання мімічних проявів емоцій.

4. Провести експериментальне тестування інформаційної системи за еталонними наборами даних.

Розділ 1 Дослідження сучасних способів та підходів до розпізнавання мімічних проявів емоцій

1.1 Актуальність застосування інформаційних технологій у системах розпізнавання візуальних образів

Візуальне розпізнавання зображень – це галузь, яка швидко розвивається, і в останні роки вона досягла значного прогресу. Ця технологія дозволяє комп'ютерам розпізнавати, розуміти та класифікувати зображення так само, як це роблять люди. Інформаційні технології, зокрема машинне навчання та алгоритми комп'ютерного зору, відіграють вирішальну роль у системах візуального розпізнавання зображень, дозволяючи системі автоматично аналізувати та розуміти зображення. На рисунку 1.1 приведено статистику поширення та росту попиту в системах розпізнавання об'єктів [1].

Рисунок 1.1 – Зріст кількості публікацій з теми розпізнавання об'єктів [1]

Візуальне розпізнавання зображень має довгу історію, перші розробки сягають 1960-х років. На початку системи розпізнавання зображень були обмежені у своїх можливостях і значною мірою покладалися на методи ручного виділення ознак. Однак із розвитком IT та появою алгоритмів машинного

навчання та нейронних мереж точність і ефективність систем розпізнавання візуальних зображень різко підвищилися.

Інформаційні технології, такі як алгоритми машинного навчання та нейронні мережі, значно вплинули на точність та ефективність систем розпізнавання візуальних зображень [2, 3]. Алгоритми машинного навчання використовуються для навчання систем візуального розпізнавання зображень на великих наборах даних, що дозволяє системам вивчати та ідентифікувати шаблони в зображеннях. Нейронні мережі, з іншого боку, забезпечують більш складний підхід до розпізнавання зображень, дозволяючи системам зрозуміти базову структуру зображення та робити прогнози на основі цього розуміння. Тренд ринку технологій розпізнавання обличчя за регіоном наведено на рисунку 1.2

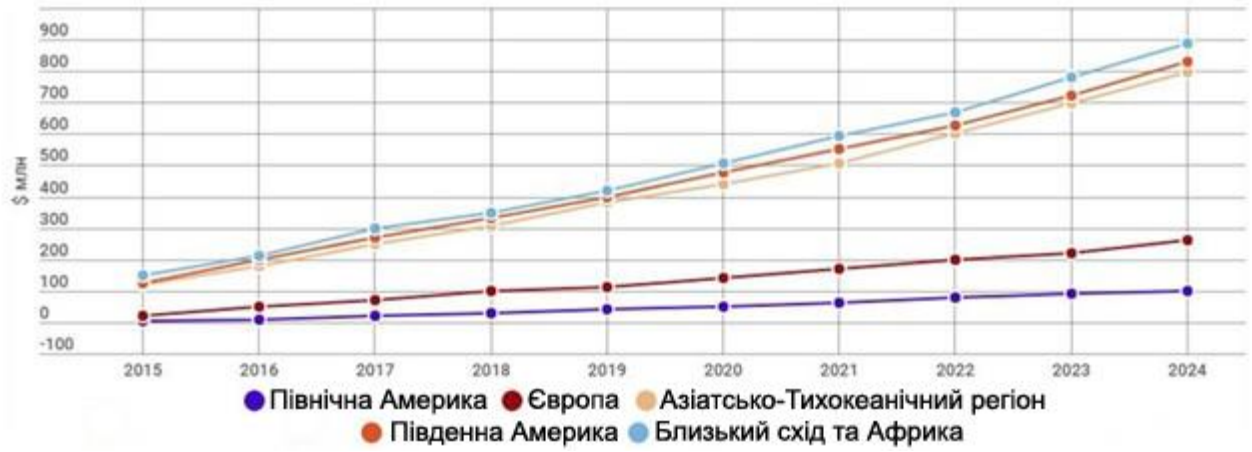


Рисунок 1.2 – Зростання ринку технологій розпізнавання обличчя за регіоном [4]

Наприклад, алгоритми машинного навчання можна використовувати, щоб навчити систему розпізнавати на зображенні конкретні об’єкти, такі як автомобілі, тварини чи люди. По-іншому це називають виявленням об’єктів або образів, таке завдання є фундаментальним для систем розпізнавання зображень. Таким чином алгоритми комп’ютерного зору дозволяють системі витягувати із зображення такі характеристики, як краї, текстури та кольори, які можна використовувати для класифікації зображення за різними категоріями.

Системи візуального розпізнавання зображень широко використовуються в різному програмному забезпеченні, включаючи розпізнавання обличчя, виявлення об’єктів і класифікацію зображень. Системи розпізнавання обличчя використовуються в системах безпеки, а системи виявлення об’єктів використовуються в таких сферах, як безпілотні автомобілі та навігація безпілотників. Системи класифікації зображень використовуються для класифікації зображень у певні класи, наприклад тварин, транспортних засобів або людей.

Крім того, із збільшенням доступності високопродуктивних обчислювальних ресурсів використання інформаційних технологій у системах візуального розпізнавання зображень і надалі залишатиметься важливою сферою досліджень і розробок. Доступність великих обсягів даних дозволяє навчати складніші моделі, а високопродуктивні обчислювальні ресурси дозволяють запускати ці моделі за розумний проміжок часу.

Попри численні переваги, використання інформаційних технологій у системах візуального розпізнавання зображень також має свої обмеження та проблеми. Однією з головних проблем є проблеми конфіденційності, пов’язані з використанням систем розпізнавання облич. Іншою проблемою є точність систем, оскільки вони все ще можуть допускати помилки та схильні до неточностей. Крім того, обмеження існуючих алгоритмів, такі як нездатність точно розпізнавати зображення за певних умов, також створюють складнощі.

Ось деякі статистичні дані щодо актуальності використання інформаційних технологій у системах візуального розпізнавання зображень:

1. Зростання ринку: згідно зі звітом MarketsandMarkets, розмір світового ринку розпізнавання візуальних зображень у 2020 році оцінювався в 10,2 мільярда доларів США та, як очікується, досягне 39,9 мільярда доларів США до 2025 року, зростаючи на 30,7% протягом прогнозованого періоду [1].

2. Рівень впровадження: Звіт Grand View Research показує, що використання технологій візуального розпізнавання зображень стрімко зростає, з показником впровадження 63,8% у 2020 році [4].

3. Галузеві застосування: використання інформаційних технологій у системах розпізнавання візуальних зображень має різні застосування в кількох галузях, включаючи роздрібну торгівлю, охорону здоров'я, транспорт і розваги. Згідно зі звітом Allied Market Research, очікується, що галузь роздрібної торгівлі займатиме найбільшу частку ринку розпізнавання візуальних зображень до 2027 року.

4. Точність і ефективність: дослідження Університету Оулу у Фінляндії показало, що використання алгоритмів глибокого навчання в системах візуального розпізнавання зображень значно підвищує точність і ефективність завдань розпізнавання зображень.

5. Економія: у звіті Accenture було виявлено, що використання технологій візуального розпізнавання зображень може допомогти організаціям заощадити кошти за рахунок автоматизації ручних процесів і зменшення потреби в ручній праці.

Ці статистичні дані підкреслюють зростаючу актуальність і важливість використання інформаційних технологій у системах візуального розпізнавання зображень, а також значний вплив, який вони мають на різні галузі та процеси.

Підсумовуючи, використання інформаційних технологій у системах візуального розпізнавання зображень є дуже актуальним, як для практичних, так і для дослідницьких цілей. Ці технології значно підвищили точність і ефективність цих систем і мають широкий спектр застосувань. Однак також важливо визнати обмеження та проблеми, пов'язані з їх використанням, і постійно працювати над пошуком рішень цих проблем.

1.2. Сучасний стан розпізнавання змін емоційних станів людини

Розпізнавання змін в емоційному стані людини є складним завданням, яке передбачає інтерпретацію різноманітних невербальних сигналів, включаючи вираз обличчя, мову тіла та тон голосу. Ці сигнали можуть передавати різні емоції, а їх поєднання може допомогти визначити більш складні емоційні стани.

Використання таких сигналів поширюється на такі галузі, як психологія, неврологія та взаємодія людини з комп’ютером.

Вивчення емоцій має довгу історію, що сягає корінням у стародавні цивілізації, а розпізнавання емоційних станів було предметом інтересу протягом століть. Перші підходи зосереджувалися на фізіологічних параметрах, таких як вираз обличчя, частота серцевих скорочень і мова тіла. Однак розвиток обчислювальних моделей і алгоритмів машинного навчання призвів до появи більш складних методів розпізнавання емоційних станів. Основні етапи процесу розпізнавання емоцій зображені на рисунку 1.3.

Рисунок 1.3 – Основні етапи процесу розпізнавання емоцій [5]

Фізіологічні показники, такі як вираз обличчя та частота серцевих скорочень, безпосередньо передають інформацію про емоційний стан людини і їх відносно легко отримати. Однак на них можуть впливати такі фактори, як стрес і тривога, що може вплинути на їхню точність. Обчислювальні моделі, такі як алгоритми машинного навчання, пропонують складніший підхід, враховуючи безліч показників та інтегруючи "розумний" одяг. Тим не менш, вони все ще схильні до помилок, упереджень і неточностей.

Метод розпізнавання мімічних проявів емоцій передбачає використання комп’ютерних алгоритмів для виявлення та аналізу різних рис обличчя, таких як рухи очей, положення брів і вигин губ людини з метою визначення її емоційного стану [5]. Переваги розпізнавання обличчя включають його універсальність,

можливості в реальному часі та здатність виявляти ледве помітні зміни у виразі обличчя. Недоліки включають вплив таких факторів, як умови освітлення та ракурси камери, різне розуміння виразу обличчя в різних культурах та окремих людей, а також можливість упередженості в алгоритмах.

Для класифікації виразів обличчя людини створювались систематичні підходи та системи, які описують та кількісно оцінюють всі можливі вирази обличчя. Основним з таких підходів є The Facial Action Coding System (FACS) [6]. Він був розроблений Полом Екманом і Уоллесом Фрізеном у 1970-х роках і широко використовується в галузях психології, неврології, соціології та інформатики. FACS починається з поділу обличчя на певні ділянки та визначення різних м’язів, які беруть участь у рухах обличчя. Потім кожен м’яз або група м’язів пов’язується з певним кодом, відомим як «Одиниця дії» (AU). У посібнику FACS перераховано 44 AU, які можна використовувати для опису широкого діапазону виразів обличчя, включаючи емоції, рухи мови та іншу невербальну поведінку.

Пізніше було визначено підсистеми, які обмежуються емоціями та мімічними проявами на обличчі людини (EMFACS) [7]. Приклади емоцій цієї системи наведено у таблиці 1.1.

Таблиця 1.1 – Основні емоції EMFACS

Емоція	Action units (AU)
Задоволення	6+12
Сум	1+4+15
Здивування	1+2+5B+26
Страх	1+2+4+5+7+20+26
Злість	4+5+7+23
Відраза	9+15+17
Зневага	R12A+R14A
Емоція	Action units (AU)

Кожна з емоцій виражається своїм характерним проявом та зміною рис на обличчі людини, приклад емоційних виразів обличчя зображено на рисунку 1.4.

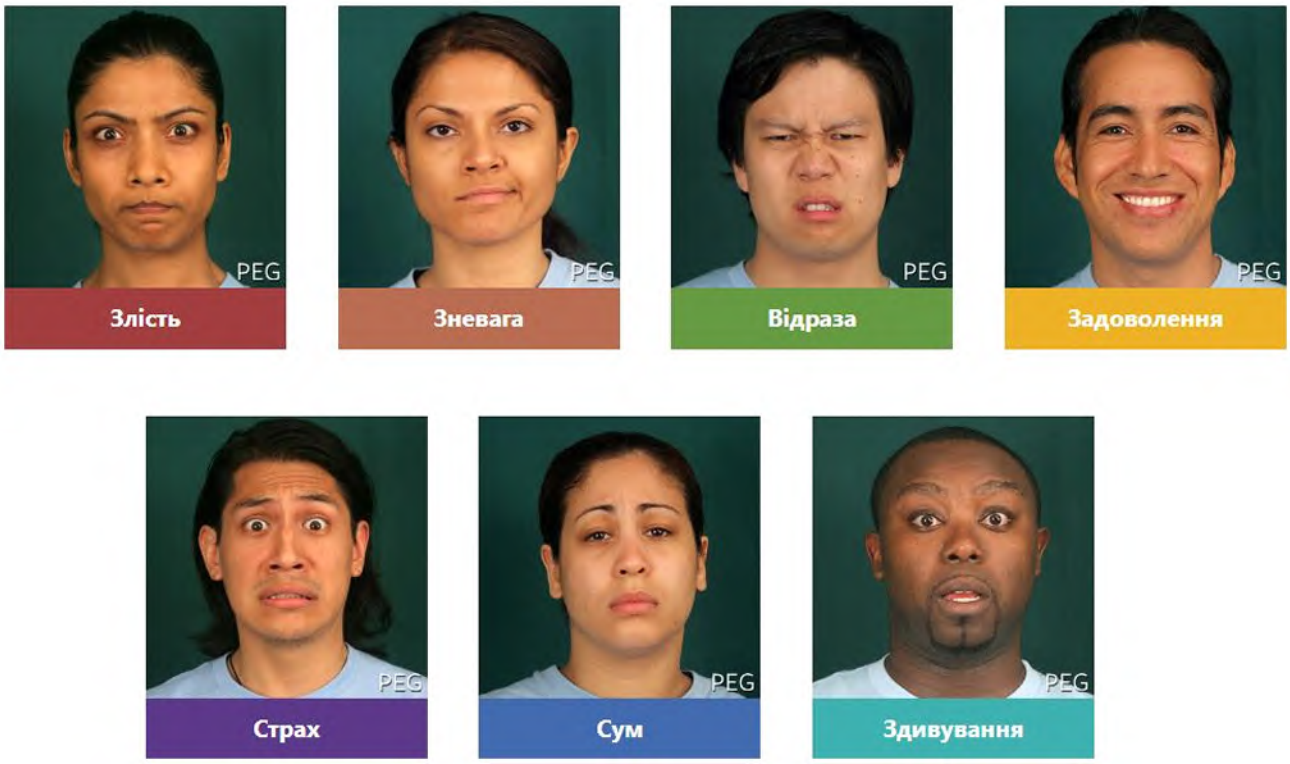


Рисунок 1.4 – Приклад виразів обличь характерних для емоцій [8]

Таким чином, для емоції задоволення є характерною посмішка, яку можна помітити за такими ознаками: очі звужені, навколо очей з’являються зморшки, щоки підняті, губи відтягнуті назад, а зуби оголені в посмішці.

Смуток характеризується дуже надійною ознакою, яку дуже важко імітувати: людина підіймає і зводить куточки брів. Також, верхні повіки опущені, а очі дивляться донизу та кути губ відтягуються вниз.

Для страху та здивування характерні однакові риси: очі, рот та брови, через що їх часто плутають. Емоція здивування від страху відрізняється вище більшим кутом підйому брів, верхні повіки та щелепа більш розслаблені.

Для страху є характерними підняті, прямі брови, високо підняті повіки, більше видно склеру. Також, губи напружені і розтягнуті, мляві від страху.

При гніві брови зведені та опущені вниз, очі широко відкриті та відображають пильний погляд, а губи щільно стиснуті.

Найпростішою ознакою відрази є зморщування носа. Також можна виділити опущені брови, зморшки на лиці та переніссі, верхня губа віднята у вигляд перевернутої букви «U», а нижня губа також піднята і трохи виступає вперед.

Нарешті, зневага – це єдина емоція, яка виділяється одностороннім виразом обличчя, тобто, вираз не симетричний. Зневага може бути з натяком на посмішку або гнів. Основною ознакою є підтягнутий і піднятий куточок губ з одної сторони обличчя.

Однією з ключових переваг FACS є його об’єктивність. Використовуючи стандартизовану систему кодування, FACS дозволяє дослідникам порівнювати й аналізувати вирази обличчя різних людей і культур у точний і логічний спосіб. Це робить його важливим інструментом для вивчення емоційного та комунікативного значення виразу обличчя, а також для розуміння зв’язку між виразом обличчя та емоціями.

Підсумовуючи, сучасний стан розпізнавання змін в емоційному стані людини характеризується розвитком складних алгоритмів і технологій. Кожен спосіб має свої переваги та недоліки, і вибір способу залежатиме від конкретних потреб і цілей особи чи організації. Однак, незважаючи на всі досягнення, розпізнавання змін в емоційному стані людини все ще є активною сферою досліджень, і значна частина поточної роботи зосереджена на покращенні точності та надійності цих алгоритмів. Крім того, розпізнавання емоцій за невербальними ознаками є складним завданням, яке часто залежить від контексту та індивідуальних особливостей, наприклад, деякі люди можуть мати більш виразне обличчя, а інші – ні.

1.3 Аналіз засобів та підходів штучного інтелекту до виявлення мімічних проявів людини

Розпізнавання виразу людського обличчя вже давно є сферою дослідження та розробки у галузі штучного інтелекту (ШІ). З розвитком технологій штучного

інтелекту за останні роки точність визначення емоційного стану згідно виразу обличчя значно покращилася. Існує кілька основних засобів і підходів, які використовуються для вирішення цього завдання:

Комп'ютерний зір. Цей підхід використовує методи обробки зображень для аналізу візуальних особливостей обличчя, таких як форма, текстура та колір, щоб розпізнавати мімічні прояви на обличчі. Наприклад, алгоритм комп'ютерного зору може визначити розташування очей, брів, рота та інших орієнтирів обличчя, а потім проаналізувати їх відносне розташування та рухи, щоб виявити емоційний стан.

Машинне навчання. Алгоритми машинного навчання, такі як опорні векторні машини та дерева рішень, можна навчити на великому наборі даних виразів обличчя, щоб розпізнавати закономірності та робити прогнози [9]. Алгоритми вчяться на позначених прикладах облич із різними виразами, а потім можуть аналізувати та класифікувати нові, незнайомі обличчя.

Глибоке навчання. Це напрям штучного інтелекту, який використовує глибокі нейронні мережі для моделювання складних шаблонів у даних. Згорткові нейронні мережі (CNN) [10, 11] особливо корисні для розпізнавання виразу обличчя, оскільки вони призначені для аналізу візуальних моделей на зображеннях.

Глибоке навчання є одним із найпопулярніших підходів до розпізнавання виразу обличчя, оскільки воно може ефективно моделювати складні шаблони згідно до візуальних даних. CNN – це тип алгоритму глибокого навчання, який особливо добре підходить для завдань розпізнавання зображень [11]. Вони використовують згорткові шари, щоб витягти особливості з вхідного зображення, а потім використовують повністю пов'язані шари, щоб робити прогнози. У контексті розпізнавання мімічних проявів людини CNN можна навчити розпізнавати риси на обличчі, які відповідають різним виразам.

Також, окрім CNN, для вирішення цієї задачі можуть бути використані і інші алгоритми. Досить популярними є: дерево рішень (DT), метод опорних векторів (SVM) і метод KNN. Це класичні алгоритми машинного навчання, які

вимагають створених вручну функцій для виявлення виразу обличчя людини. Вони мають відносно нижчу точність і можуть не підходити для розпізнавання складних виразів. Порівняння точності алгоритмів зображено на рисунку 1.5.

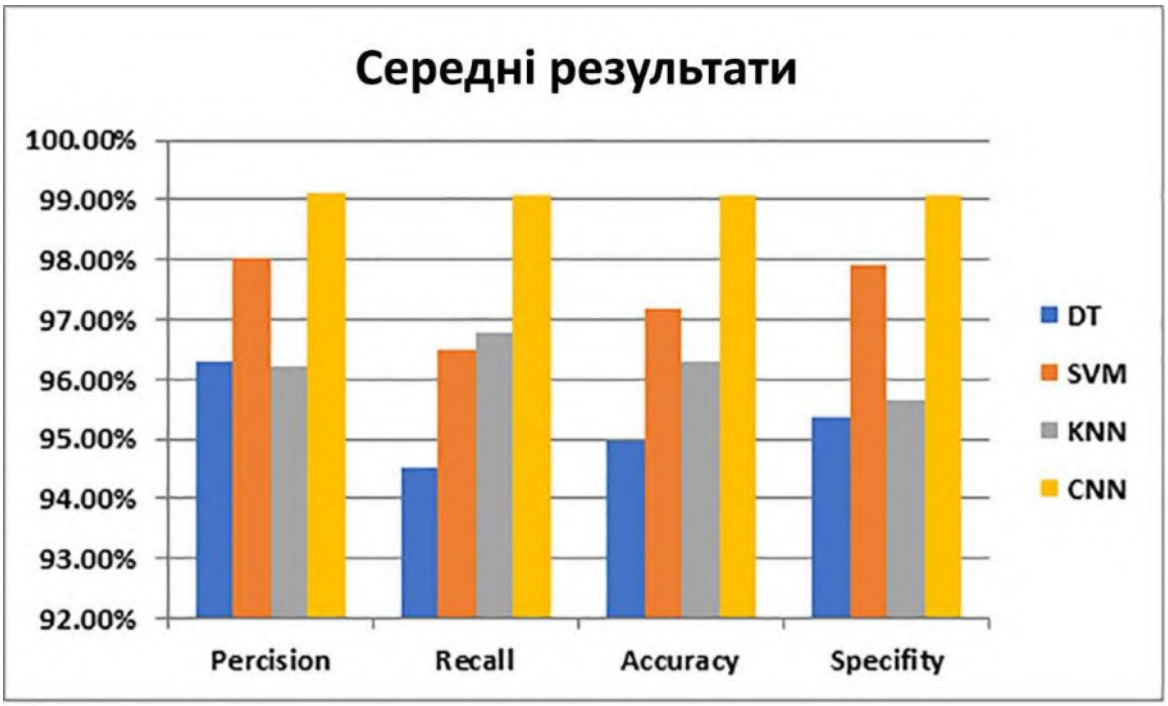


Рисунок 1.5 – Порівняння показників алгоритмів у завданні розпізнавання виразів обличь [12]

Кожен із алгоритмів має свої особливості:

1. Дерево рішень [13]: Це контрольований алгоритм навчання, який використовується для вирішення проблем класифікації. Він створює деревовидну структуру шляхом рекурсивного розбиття даних на основі значення ознак, щоб максимізувати приріст інформації. Його можна використовувати для визначення важливих орієнтирів обличчя, таких як куточки рота, носа та очей, і створення правил для класифікації виразів обличчя на основі їхніх геометричних властивостей. Однак дерева рішень можуть страждати від перенавчання та мати нижчу точність, ніж більш просунуті методи.

2. Метод опорних векторів (SVM) [14]: Це алгоритм двійкової класифікації, який розділяє дані, знаходячи найкращу гіперплощину, яка максимально розділяє два класи. SVM можна навчити на таких ручних функціях,

як локальні двійкові шаблони, фільтри Габора або вейвлети Хаара, щоб розпізнавати вирази обличчя. SVM мають хорошу продуктивність узагальнення, але потребують конструювання ознак, що може зайняти багато часу.

3. K-найближчі сусіди (KNN) [15]: Це простий непараметричний алгоритм, який використовується для задач класифікації або регресії. Він класифікує дані, знаходячи найпоширеніший клас серед K найближчих сусідів у просторі ознак. KNN можна використовувати з різноманітними дескрипторами ознак, такими як гістограми орієнтованих градієнтів, локальні бінарні шаблони або реберні функції для розпізнавання виразів обличчя. Однак KNN чутливий до вибору K і може мати проблеми з масштабованістю. У додатку А наведено таблицю А.1 з характеристикою основних алгоритмів.

CNN мають свої особливості та переваги. До переваг можна віднести такі особливості:

– Висока точність: показано, що CNN досягають високої точності в задачах розпізнавання виразу обличчя, особливо при навчанні на великих наборах даних. Завдяки цьому вони добре підходять для використання у сферах, де точність є ключовою вимогою.

– Автоматизація. CNN можуть автоматизувати процес визначення виразу обличчя, зменшуючи потребу в ручній роботі та налаштуванні, завдяки цьому підвищується ефективність.

– Стійкість до особливостей зображень. CNN можна навчити розпізнавати вирази обличчя, незважаючи на зміни в освітленні, позах обличчя та інших факторах, які можуть вплинути на вигляд обличчя. Це робить їх особливо придатними для додатків, де такі особливості є поширеними.

Недоліки:

– Вимоги до обчислень. Навчання CNN розпізнаванню виразу обличчя може бути дорогим з обчислювальної точки зору, вимагаючи значних обчислювальних ресурсів. Це може бути недоліком для систем з обмеженими ресурсами або для програм, де обробка в реальному часі не потрібна.

– Перенавчання. CNN можуть бути схильні до перенавчання, особливо коли навчаються на невеликих наборах даних. Перенавчання відбувається, коли модель стає занадто спеціалізованою для попередніх навчальних даних і погано узагальнює нові, невідомі дані.

– Недостатність інтерпретації. CNN може бути важко інтерпретувати, тому важко зрозуміти, чому саме модель робить певні прогнози.

Порівнюючи CNN з іншими методами розпізнавання виразу обличчя, стає зрозуміло, що CNN пропонують ряд переваг у плані точності, автоматизації та стійкості до особливостей зображення. CNN може автоматично вивчати риси обличчя із зображень і демонструє найкращу продуктивність у розпізнаванні виразів обличчя. Порівняно з іншими методами, CNN більше підходить для розпізнавання складних виразів і може обробляти вхідні зображення з різними особливостями.

1.4 Мета, завдання та вимоги до реалізації інформаційної системи

Метою кваліфікаційної роботи бакалавра є підвищення точності ідентифікації змін емоційного стану людини за мімічними проявами за цифровими фото- та відеоматеріалами.

Для досягнення поставленої мети необхідно виконати наступні завдання:

1. Провести аналіз способів та підходів до ідентифікації змін емоційного стану людини за мімічними проявами та обрати найкращий.

2. Застосувати обраний спосіб ідентифікації змін емоційного стану людини за мімічними проявами до розв’язання задачі автоматизованого розпізнавання мімічних проявів емоцій.

3. Реалізувати обраний спосіб у вигляді інформаційної системи автоматизованого розпізнавання мімічних проявів емоцій.

4. Провести експериментальне тестування інформаційної системи за еталонними наборами даних.

Розділ 2 Спосіб розпізнавання мімічних проявів емоцій засобами штучного інтелекту

2.1 Модель розпізнавання мімічних проявів емоцій

При перетворенні зображення в математичну модель або матрицю використовують процес, який називається "дискретизація". Він передбачає поділ безперервного зображення на кінцевий набір дискретних точок, які потім можна представити у вигляді матриці. Щоб дискретизувати зображення, спочатку потрібно визначитися з роздільною здатністю або розміром пікселів. Це визначає розмір кожної дискретної точки на зображенні. Наприклад, якщо роздільна здатність – 1 піксель на квадратний міліметр, то кожна точка на зображенні буде відповідати квадратному міліметру вихідного неперервного зображення.

Після визначення роздільної здатності, роблять вибірки зображення через регулярні інтервали, щоб створити матрицю дискретних точок. Кожна точка в матриці представляє колір або інтенсивність відповідної точки на оригінальному зображенні.

Існують різні способи представлення кольору в матриці, але одним з найпоширеніших є використання значень RGB (червоний, зелений, синій). Кожен піксель у матриці представлений триплетом значень від 0 до 255, що відповідає інтенсивності червоного, зеленого та синього каналів. Для зображень у відтінках сірого використовують одне значення від 0 до 255 для представлення інтенсивності кожного пікселя (рисунок 2.1).

Загалом, перетворення зображення в математичну модель або матрицю передбачає дискретизацію зображення і представлення кожної точки зображення як значення в матриці. Це дозволяє нам виконувати різні математичні операції над зображенням, такі як фільтрація, покращення або сегментація.

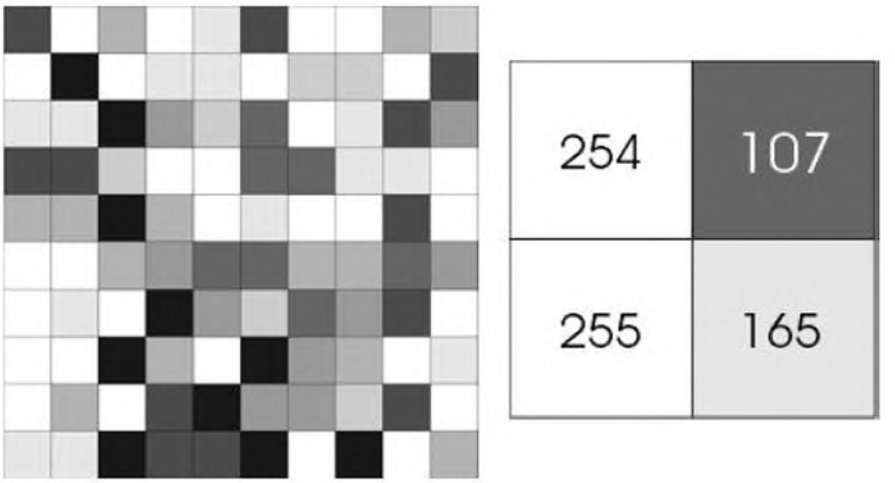


Рисунок 2.1 – Приклад значень матриці

Щоб краще зрозуміти притаманні зображенням властивості та технічні процедури, що використовуються для обробки зображень зображення подають з окремих пікселів, як функцію *f*. Кожен піксель також має власне значення. Для зображення у відтінках сірого кожен піксель має інтенсивність від 0 до 255, де 0 – чорний колір, а 255 – білий (рисунок 2.1). *f(x,y)* тоді дасть інтенсивність зображення у позиції пікселя *(x,y),* якщо припустити, що вона задана у вигляді прямокутника з обмеженим діапазоном: *f: [a,b]* x *[c,d]* → [0, 255].

Кольорове зображення є простим розширенням цього. *f(x,y)* тепер є вектором трьох значень замість одного. На прикладі RGB-зображення кольори будуються з комбінації червоного, зеленого та синього (RGB). Тому кожен піксель зображення має три канали і представляється у вигляді вектора 1x3 (рисунок 2.2). Оскільки три кольори мають цілі значення від 0 до 255 [16], існує загалом 256*256*256 = 16,777,216 комбінацій або варіантів кольорів.

Приклад функції для перетворення кольорового зображення:

$$f(x,y)=\begin{bmatrix} r(x,y) \\ g(x,y) \\ b(x,y) \end{bmatrix}, \tag{2.1}$$

де $r(x,y)$, $g(x,y)$, $b(x,y)$ – кольорові канали, відповідно до [Red, Green, Blue].

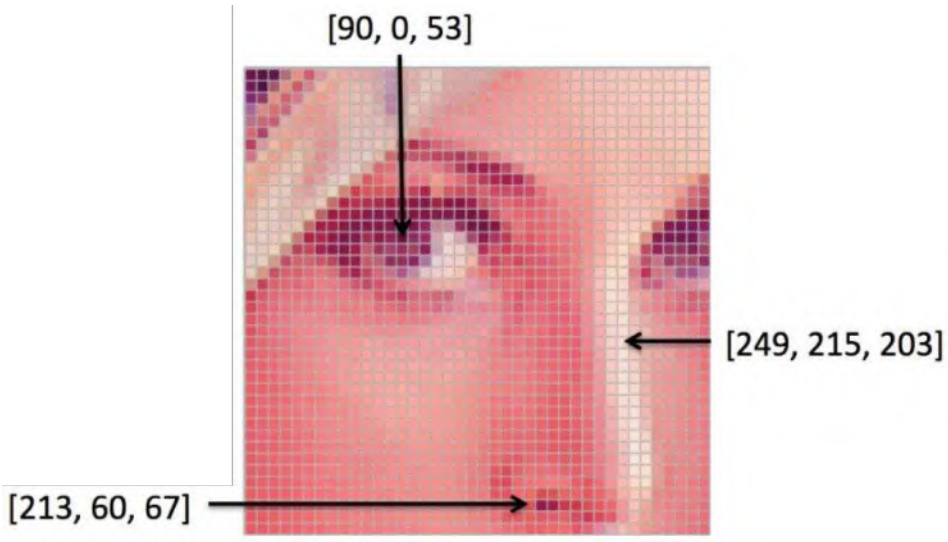


Рисунок 2.2 – Значення функції для пікселів з фотографії обличчя

Таким чином, зображення за формулою (2.1) можна представити як матрицю значень пікселів (рисунок 2.3).

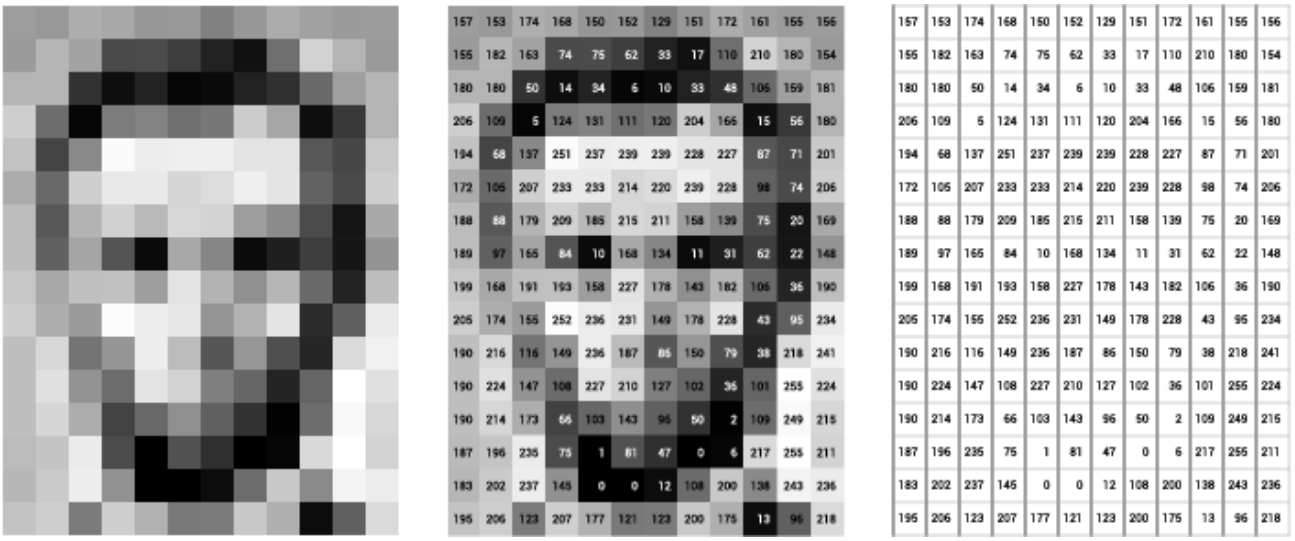


Рисунок 2.3 – Матричне представлення спрощеного зображення обличчя

Попередня обробка зображень і відеофайлів – важливий крок у підготовці даних до використання в нейронній мережі для задач розпізнавання. Мета попередньої обробки – очистити і нормалізувати дані, щоб нейронній мережі було легше вивчати закономірності і робити точні прогнози.

Етапи попередньої обробки зображень і відео можуть відрізнятися залежно від конкретного завдання і характеристик даних. Однак є деякі загальні кроки попередньої обробки:

1. Зміна розміру: Зображення та відео можуть мати різний розмір і співвідношення сторін, що може ускладнити процес навчання нейронній мережі. Зміна розміру зображень і відео до стандартного розміру може зробити дані більш узгодженими і легшими для обробки.

2. Обрізання: У деяких випадках важлива інформація на зображенні або відео може бути сконцентрована в певній області. Обрізання зображення або відео, щоб зосередитися на цій області, може підвищити точність нейронної мережі.

3. Нормалізація: Нормалізація передбачає масштабування значень пікселів у зображенні або відео до стандартного діапазону. Це може бути важливо для того, щоб нейронна мережа однаково обробляла всі дані, незалежно від відмінностей у яскравості або контрастності.

4. Перетворення колірного простору: Різні колірні простори можуть бути більш придатними для різних завдань. Перетворення колірного простору зображення або відео в більш відповідний формат може полегшити нейронній мережі вивчення шаблонів.

5. Доповнення (Аугментація): Доповнення передбачає створення нових даних шляхом застосування перетворень до наявних даних. Це може допомогти збільшити розмір набору даних і підвищити надійність нейронної мережі. Прикладами доповнення є обертання, перевертання або зсув зображень чи відео.

6. Кодування: Кодування передбачає перетворення зображення або відеоданих у формат, який може бути використаний нейронною мережею. Найпоширенішими форматами кодування зображень є JPEG, PNG і BMP, а відео – H.264 і MPEG-4.

В даній роботі важливою є задача навчити нейронну мережу розпізнавати різні емоції за зображеннями облич, такі як радість, сум, злість або здивування.

Спершу робота ведеться з набору даних зображень людських облич, які можуть бути різного розміру, орієнтації та умов освітлення.

Першим кроком у попередній обробці даних є приведення зображень до стандартного розміру, наприклад, 64x64 пікселів. Це можна зробити за допомогою бібліотеки OpenCV.

Далі виконується виявлення обличчя на кожному зображенні та вирівняти його за стандартною орієнтацією. Варто використовувати алгоритми виявлення облич, такі як каскади Хаара, або методи на основі глибокого навчання, такі як MTCNN (багатозадачні каскадні ЗНМ), щоб ідентифікувати обличчя на кожному зображенні та обрізати його до стандартного розміру та орієнтації.

Після обрізання нормалізується значення пікселів на зображенні до стандартного діапазону, наприклад, від -1 до 1. Це може допомогти гарантувати, що нейронна мережа однаково обробляє всі зображення, незалежно від відмінностей у яскравості або контрастності.

Для зображень обличчя колірний простір перетворюють у відтінки сірого, щоб спростити дані і зменшити розмірність вхідних даних. Це може допомогти покращити продуктивність нейронної мережі, зменшивши кількість даних, які їй потрібно обробити.

Щоб збільшити розмір набору даних і підвищити надійність нейронної мережі використовують методи доповнення. Наприклад, випадковим чином повертають зображення, щоб створити нові варіації вихідних даних.

Нарешті, нам потрібно закодувати зображення у формат, який може бути використаний нейронною мережею. Поширеними форматами кодування зображень є JPEG, PNG та BMP. Залежно від фреймворку, який використовується для навчання нейронної мережі, нам може знадобитися конвертувати зображення в певний формат, наприклад, формат TFRecord від TensorFlow.

Загалом, попередня обробка даних зображень обличчя для використання в нейронних мережах включає низку кроків для очищення та нормалізації даних. Як наслідок покрокового виконання, суттєво покращується продуктивність

нейронної мережі та підвищується її точність у розпізнаванні різних емоцій на зображенні обличчя.

2.2 Спосіб розпізнавання мімічних проявів емоцій

Процес розпізнавання обличчя зазвичай включає створення копії вхідного зображення, а потім додавання графічних елементів для відображення виявлених орієнтирів обличчя. Наприклад, у OpenCV [17] функція circle може бути використана для малювання кола в кожній виявленій точці орієнтира. Центр кола розташовується в точці-орієнтирі, а радіус кола можна встановити на потрібне значення, виходячи з розміру обличчя на зображенні. Схема виділення обличчя на зображенні подана на рисунку 2.4.

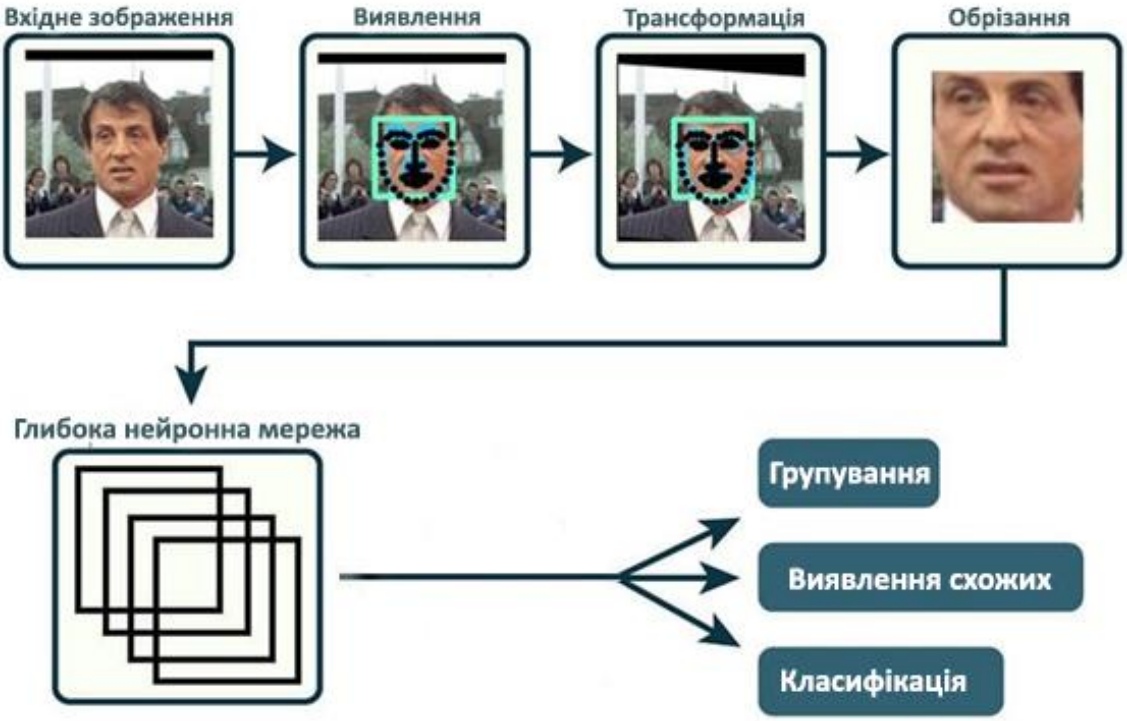


Рисунок 2.4 – Кроки виділення області інтересу за допомогою OpenCV [18]

Накладання точок і ліній може бути корисним у багатьох випадках. Наприклад, його можна використовувати для візуалізації та налагодження, дозволяючи розробникам побачити точність алгоритму виявлення орієнтирів обличчя і виявити будь-які помилки або неточності. Він також може бути

використаний для навчання і тестування моделей машинного навчання для розпізнавання емоцій, надаючи стандартизований набір ознак, які можуть бути вилучені з кожного зображення і використані в якості вхідних даних для моделі.

2.3 Спосіб класифікації емоційного стану за мімічними проявами

Розпізнавання виразу обличчя є складною проблемою в галузі комп'ютерного зору, оскільки вимагає точної ідентифікації та класифікації різних емоційних станів на основі рис обличчя людини. Згорткові нейронні мережі (CNN) показали чудові результати у вирішенні цієї проблеми, використовуючи свою здатність витягувати важкі ознаки із зображень і вивчати складні патерни.

Метою проекту є розробка моделі на основі CNN, яка може точно класифікувати емоційні стани на основі виразу обличчя. Це передбачає навчання моделі на великому наборі даних з маркованих зображень обличчя, де кожне зображення відповідає певному емоційному стану, такому як щастя, смуток, гнів або здивування.

Модель на базі CNN буде призначена для вивчення основних закономірностей у рисах обличчя, характерних для кожного емоційного стану. Це передбачає навчання моделі розпізнавати і класифікувати різні патерни в м'язах обличчя, такі як розташування і рух брів, рота і очей.

Отримана модель матиме широкий спектр застосувань, включаючи розпізнавання емоцій у соціальній робототехніці, діагностику психічного здоров'я та аналіз настроїв клієнтів. Вона має потенціал докорінно змінити спосіб взаємодії з машинами та надати нові знання про людську поведінку та психологію.

На рисунку 2.5 показано основний процес технології розпізнавання емоцій на обличчі.

Рисунок 2.5 – Процес розпізнавання емоції

Виділення рис обличчя (Feature extraction) – це метод виділення рис компонентів обличчя, таких як очі, ніс, рот і т.д., із зображення людського обличчя. Для ініціалізації таких методів, як виявлення обличчя, розпізнавання виразу обличчя або розпізнавання емоцій, виділення рис обличчя є обов'язковим. Серед усіх рис обличчя важливо знайти очі та визначити, де розташовані інші риси обличчя.

Простіше кажучи, алгоритм розпізнавання емоцій – це набір правил, які машина використовує для ідентифікації емоцій на обличчі на зображенні, а потім відповідно їх класифікує. Для кожного зображення обличчя вибирається вибірка точок властивостей в областях з найбільшими відхиленнями від стандартних емоцій, щоб точно класифікувати емоції, зображені на обличчі. CNN піксельні зображення, вивчаючи ієрархічні представлення зображення за допомогою серії згорткових та об'єднувальних шарів.

Перші кілька шарів CNN зазвичай складаються зі згорткових шарів, які вивчають локальні шаблони та особливості зображення, такі як краї, кути та текстури. Кожен згортковий шар складається з декількох фільтрів або ядер, які являють собою невеликі матриці, що ковзають по зображенню і обчислюють точкові добутки з локальними значеннями пікселів. Точкові добутки підсумовуються і пропускаються через нелінійну функцію активації для отримання карти особливостей.

Оскільки вхідне зображення згортається за допомогою декількох фільтрів, мережа вчиться розпізнавати все більш складні патерни та особливості. Вихід першого шару згортки пропускається через шар об'єднання, який зменшує просторову роздільну здатність карти ознак, вибираючи максимальне або середнє значення в кожній невеликій області. Об'єднання допомагає зменшити кількість параметрів у моделі і робить її більш стійкою до невеликих просторових змін у вхідному зображенні.

Процес згортки та об'єднання повторюється кілька разів, при цьому кожен наступний шар вивчає все більш складні та абстрактні ознаки. Нарешті, вихід

останнього об’єднуючого шару перетворюється на одновимірний вектор і пропускається через один або кілька повністю пов’язаних шарів, які використовують вивчені ознаки для класифікації вхідного зображення в один з декількох класів.

CNN, яка буде використовуватись в інформаційній системі складається з ряду шарів, кожен з яких містить набір нейронів. Структура CNN зображена на рисунку 2.6.

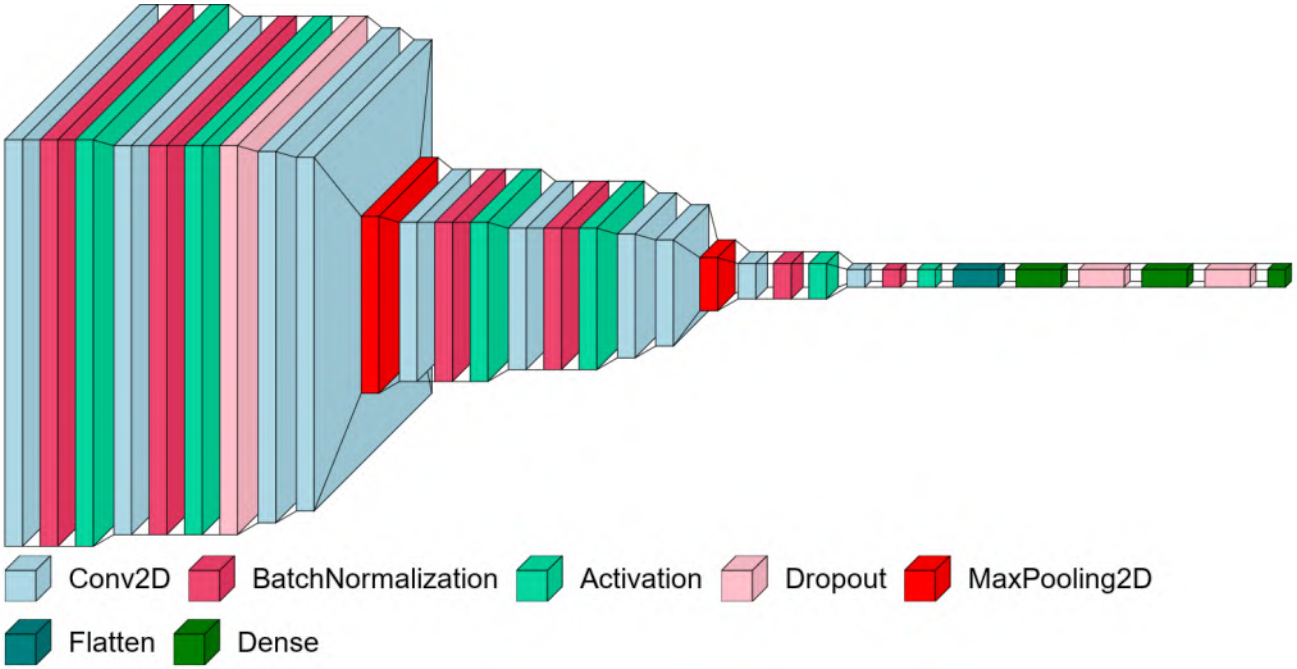


Рисунок 2.6 – Архітектура нейронної мережі, яка, власне, виконуватиме розпізнавання та класифікацію емоції людини

Визначення шарів CNN:

Conv2D: Цей шар є 2D-згортковим шаром, який виконує згортку двовимірних даних зображення. Він застосовує фільтр або ядро до кожної частини вхідного зображення, щоб створити карту особливостей. Шар Conv2D має визначену кількість фільтрів, розмір ядра (3, 3) та активацію ReLU.

BatchNormalization: Цей шар нормалізує вихідні дані попереднього шару, що допомагає стабілізувати та пришвидшити навчання. Це відбувається шляхом стандартизації середнього значення та дисперсії кожної карти ознак.

Activation: Цей шар застосовує функцію активації до результатів попереднього шару. Також використовується функція активації ReLU (Rectified Linear Unit), яка широко використовується в нейронних мережах і відома своєю ефективністю у вирішенні проблеми зникаючого градієнта.

Dropout: Цей шар є технікою регуляризації, яка випадковим чином відсіває частину нейронів під час навчання, щоб запобігти надмірному пристосуванню. Коефіцієнт відсіву встановлено на 0.5, що означає, що половина нейронів попереднього шару буде випадковим чином відсіяна під час навчання.

MaxPooling2D: Цей шар виконує 2D максимальне об'єднання вхідних даних. Максимальне об'єднання – це техніка зменшення вибірки, яка зменшує просторові розміри карт ознак, зберігаючи при цьому найважливішу інформацію. Буде використано розмір об'єднання (2, 2) і крок (2, 2).

Flatten: Цей шар вирівнює вхідні дані в одномірний масив. Це необхідно для того, щоб з'єднати вихідні дані згорткових шарів з повністю з'єднаними шарами.

Щільний (Dense): Цей шар є повністю пов'язаним шаром нейронної мережі. Він з'єднує кожен нейрон попереднього шару з кожним нейроном поточного шару. Використовується два щільних шари, кожен з яких має 1024 одиниці та активацію ReLU. За першим щільним шаром слідує шар відсіву з коефіцієнтом відсіву 0.2.

Останній щільний шар має стільки ж одиниць, скільки класів у завданні класифікації, і використовує активацію softmax для виведення прогнозованих ймовірностей класів.

Для задачі розпізнавання емоцій спочатку збирається і попередньо обробляється набір даних з мічених зображень облич, що виражають різні емоції (наприклад, щастя, смуток, гнів, страх, огиду і здивування). Це включає зміну розміру зображень до однакового розміру, нормалізацію значень пікселів та розділення набору даних на навчальний, валідаційний та тестовий набори.

Далі на навчальному наборі даних за допомогою тренувального набору навчається модель CNN. CNN складається з декількох шарів фільтрів, які

навчаються ідентифікувати та виокремлювати ознаки з вхідних зображень. Вихідні дані CNN проходять через один або кілька повністю пов’язаних шарів, які вчяться зіставляти вилучені ознаки з передбачуваною емоцією.

Під час навчання модель CNN оптимізується, щоб мінімізувати різницю між передбаченою емоцією та істинною міткою емоції для кожного зображення в навчальній вибірці. Для цього використовується функція втрат, така як категоріальна перехресна ентропія або середньоквадратична помилка. Оптимізація виконується за допомогою зворотного поширення та градієнтного спуску, які коригують ваги та упередження моделі для покращення її роботи на навчальній вибірці.

Після того, як модель CNN навчена, вона оцінюється на валідаційному наборі для визначення її продуктивності та точного налаштування її гіперпараметрів (таких як швидкість навчання, кількість фільтрів та рівень відсіву). Потім модель з найкращими показниками тестується на тестовому наборі, щоб виміряти її точність та ефективність узагальнення.

Щоб розпізнати емоції на новому вхідному зображенні, попередньо оброблене зображення подається на навчену модель CNN, яка виводить передбачувану мітку емоції. Потім передбачена мітка порівнюється з набором можливих міток емоцій, щоб визначити найбільш ймовірну емоцію, виражену на вхідному зображенні.

Точність системи розпізнавання емоцій залежить від якості набору даних, дизайну моделі CNN та параметрів, що використовуються під час навчання. Її можна покращити, використовуючи такі методи, як доповнення даних, навчання з переносом та ансамблювання.

Загалом, процес розпізнавання емоцій на обличчі за допомогою CNN включає збір і попередню обробку маркованого набору даних, навчання CNN-моделі на цьому наборі даних, оцінку і точне налаштування моделі, а також використання моделі для прогнозування емоцій, виражених на нових вхідних зображеннях.

Навчити згорткову нейронну мережу (CNN) розпізнавати емоції за виразом обличчя – складне завдання, яке вимагає великого та різноманітного набору даних. Набір даних повинен містити достатню кількість зображень з мітками різних емоцій, щоб покрити широкий спектр виразів обличчя. Щоб вибрати найбільш підходящий набір даних для навчання CNN, потрібно врахувати кілька ключових характеристик.

По-перше, розмір набору даних має вирішальне значення для визначення продуктивності CNN. Більший набір даних дає змогу краще узагальнити модель, особливо коли йдеться про варіації виразу обличчя, умов освітлення та положення голови. Отже, набір даних має бути достатньо великим, щоб забезпечити різноманітний набір зображень, які відображають вищезгадані варіації.

По-друге, набір даних має бути достатньо різноманітним, щоб охопити широкий спектр емоцій та виразів обличчя. Він повинен включати зображення людей різних етнічних груп, культур і гендерів. Така різноманітність допомагає моделі вивчити загальні риси виразу обличчя і точніше розпізнавати емоції.

По-третє, набір даних повинен бути точно позначений відповідними емоціями. Маркування має бути послідовним, а анотації – надійними, що дозволить моделі навчатися і розпізнавати емоції з більшою точністю.

По-четверте, якість зображень у наборі даних має вирішальне значення для роботи моделі. Зображення повинні бути високої якості, з достатньою роздільною здатністю та чіткістю. Зображення також не повинні містити шуму, розмиття або інших артефактів, які можуть вплинути на якість роботи CNN.

Нарешті, набір даних має бути легкодоступним і загальнодоступним, щоб забезпечити відтворюваність дослідження.

Враховуючи ці характеристики, набір даних FER+ є найбільш придатним для навчання CNN розпізнаванню емоцій. Набір даних FER+ [19] – це загальнодоступний набір даних, який містить зображення облич людей, що виражають сім основних емоцій, включаючи нейтральну, радість, смуток, гнів, страх, здивування та огиду. Набір даних містить понад 35 000 зображень облич, що робить його одним з найбільших наборів даних, доступних для розпізнавання

емоцій. Зображення в наборі даних FER+ різноманітні, з варіаціями виразу обличчя, умов освітлення та положення голови. Набір даних має чітке маркування, надійні анотації та високу якість. Набір даних FER+ є покращеним порівняно з набором даних FER, оскільки він містить додаткові анотації та більш збалансований розподіл емоцій.

Нарешті, дуже важливо оцінювати ефективність CNN за допомогою відповідних метрик, таких як точність, достовірність, відтворення та оцінка F1. Ці метрики забезпечують об’єктивну оцінку ефективності моделі та дозволяють порівнювати її з іншими моделями. Крім того, методи перехресної перевірки, такі як k-кратна перехресна перевірка, можуть допомогти в оцінці продуктивності моделі та запобігти перенавчанню, гарантуючи, що модель добре описує непередбачувані дані. Вибір відповідних метрик оцінювання та методів перехресної перевірки має вирішальне значення для досягнення надійних результатів у поставленій задачі.

Таким чином, при навчанні CNN розпізнаванню емоцій за виразом обличчя важливо мати великий, різноманітний, точно маркований, високоякісний і легкодоступний набір даних. Набір даних FER+ задовольняє всім цим вимогам і є найбільш придатним.

2.4 Функціональна структура інформаційної системи

Для розуміння схематичної структури інформаційної системи, варто створити UML-діаграму діяльності для інформаційної системи. Діаграма забезпечує візуальне представлення робочого процесу системи і допомагає розробникам зрозуміти різні етапи, пов’язані з обробкою вхідних даних. Розроблену діаграму зображено на рисунку 2.7.

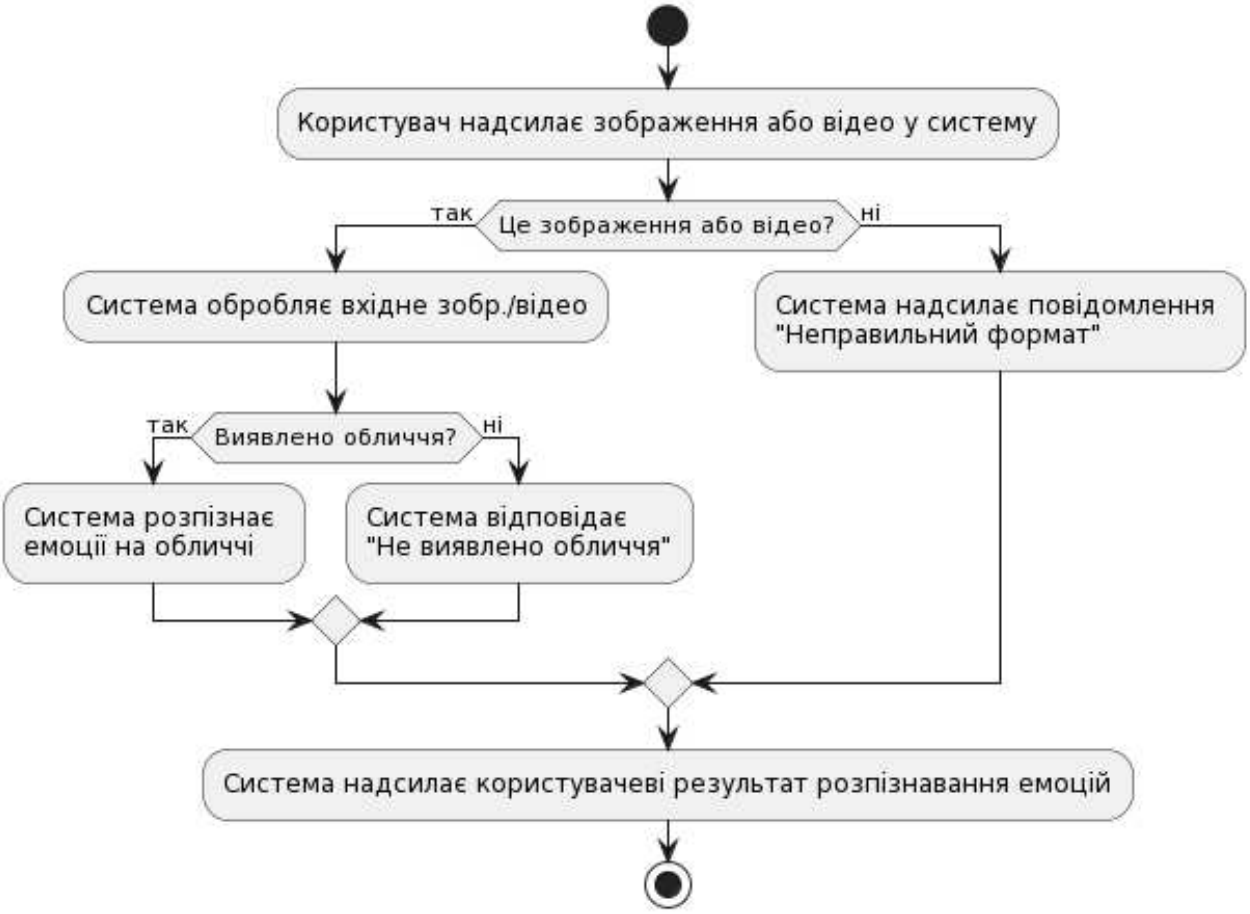


Рисунок 2.7 – Діаграма роботи системи з користувачем

Використовуючи UML-діаграму діяльності, розробники можуть переконатися, що інтерфейс відповідає вимогам системи, виявити потенційні проблеми та оптимізувати роботу системи. Крім того, діаграма може слугувати цінним довідником для інших членів команди, які працюють над проектом, таких як тестувальники та менеджери проектів, щоб отримати чітке розуміння того, як працює система, і переконатися, що вона відповідає очікуванням кінцевих користувачів.

Для вдалого формування структури функціональної системи для систем розпізнавання емоційного стану – важливо об’єднати функції інформаційної системи в окремі модулі, оскільки це дасть змогу краще організувати систему та забезпечити її модульність. Розбиваючи систему на окремі функціональні модулі, кожен модуль може зосередитися на певному наборі пов’язаних між собою завдань, що полегшує розробку, тестування та підтримку системи.

Для інформаційної системи було розроблено такі функціональні модулі системи (рисунок 2.8):

1. Модуль розпізнавання зображень.
2. Модуль розпізнавання відео.
3. Модуль інтерфейсу чат-бота.

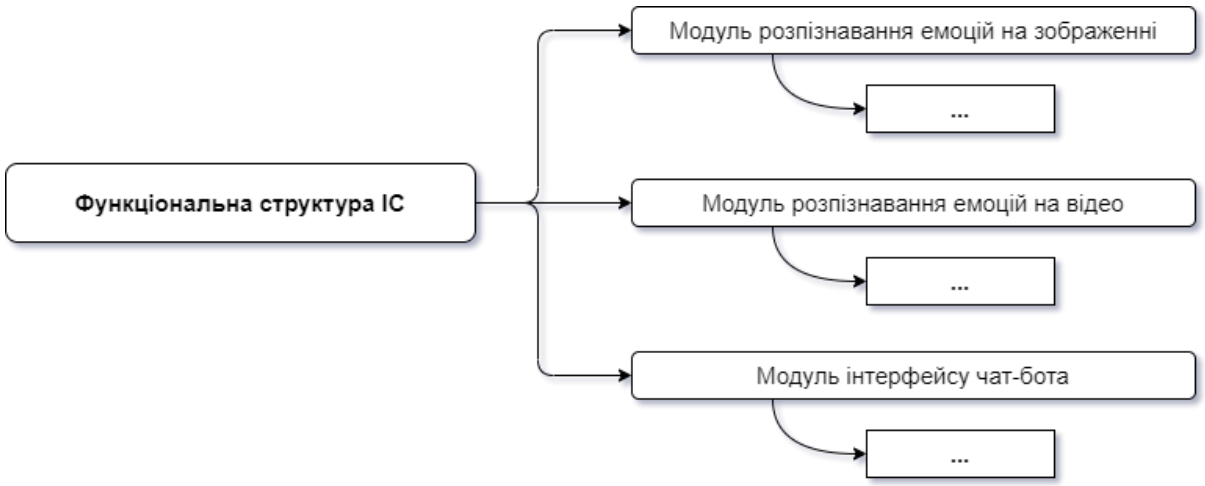


Рисунок 2.8 – Схема модулів інформаційної системи

Модуль розпізнавання зображень

Цей модуль (рисунок 2.9) буде відповідати за функції розпізнавання емоцій людини на зображеннях з обличчям.

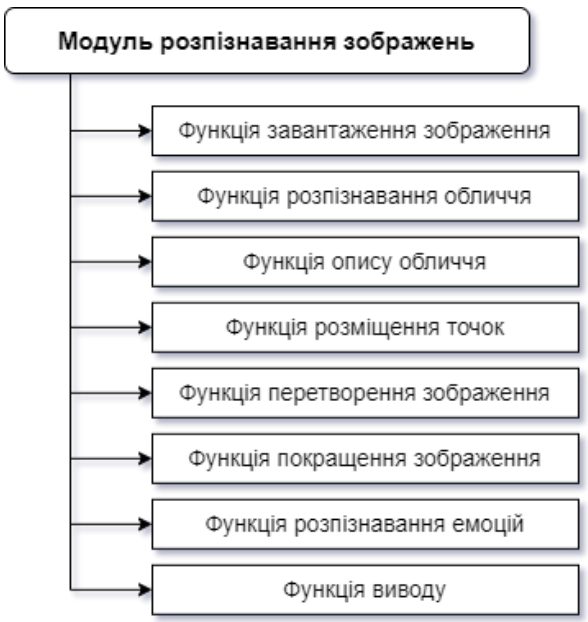


Рисунок 2.9 – Схема модулю розпізнавання емоцій на зображенні

Нижче описано функції модуля розпізнавання зображень:

1. Функція завантаження зображення:

Ця функція дозволяє користувачеві завантажити зображення в систему, надіславши його чат-боту. Функція повинна мати можливість отримувати зображення та зберігати його в системі.

2. Функція розпізнавання обличчя:

Ця функція визначає наявність обличчя на завантаженому зображенні за допомогою алгоритму розпізнавання облич. Функція повинна бути здатна ідентифікувати положення обличчя на зображенні.

3. Функція опису обличчя:

Ця функція описує риси обличчя виявленого обличчя, такі як очі, ніс, рот тощо. Функція повинна бути здатна розпізнавати і описувати положення і розмір цих рис.

4. Функція розміщення точок:

Ця функція дозволяє користувачеві вибирати та розміщувати точки на рисах обличчя виявленого обличчя, таких як очі, ніс, рот тощо. Функція повинна мати можливість отримувати точки і зберігати їх в системі.

5. Функція перетворення зображення:

Ця функція перетворює завантажене зображення в необхідний формат для розпізнавання емоцій. Функція повинна вміти змінювати розмір зображення, нормалізувати значення пікселів і конвертувати його у відповідний формат файлу.

6. Функція покращення зображення:

Ця функція покращує якість зображення шляхом застосування фільтрів, регулювання яскравості/контрастності або інших методів обробки зображення. Ця функція має підвищити точність алгоритму розпізнавання емоцій.

7. Функція розпізнавання емоцій:

Ця функція використовує попередньо навчену модель машинного навчання для розпізнавання емоцій на завантаженому зображенні на основі вибраних рис обличчя. Функція повинна мати можливість аналізувати вираз виявленого обличчя та визначати найбільш ймовірну емоцію.

8. Функція виводу:

Ця функція надсилає розпізнану емоцію та іншу відповідну інформацію назад користувачеві.

Модуль розпізнавання відео

Цей модуль (рисунок 2.10) відповідатиме за функції розпізнавання людських емоцій на відео яке містить обличчя людини.

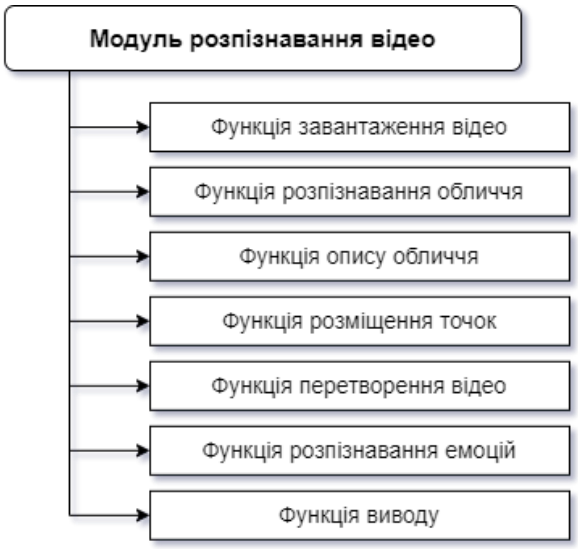


Рисунок 2.10 – Схема модулю розпізнавання емоцій на відео

Нижче описано функції модуля розпізнавання відео:

1. Функція завантаження відео дозволяє користувачеві завантажити відео в систему, надіславши відео чат-боту. Функція повинна мати можливість отримувати відео та зберігати його в системі.

2. Функція розпізнавання обличчя виявляє наявність обличчя на завантаженому відео за допомогою алгоритму розпізнавання обличчя. Функція повинна бути здатна ідентифікувати положення обличчя у кожному кадрі відео.

3. Функція опису обличчя описує риси обличчя виявленої особи, такі як очі, ніс, рот тощо. Функція повинна бути здатна розпізнавати і описувати положення і розмір цих особливостей в кожному кадрі відео.

4. Функція розміщення точок дозволяє користувачеві вибирати та розміщувати точки на рисах обличчя виявленої особи у кожному кадрі відео,

таких як очі, ніс, рот тощо. Функція повинна мати можливість отримувати точки і зберігати їх у системі.

5. Функція перетворення відео перетворює завантажене відео в необхідний формат для розпізнавання емоцій. Функція повинна мати можливість змінювати розмір відео, нормалізувати значення пікселів і конвертувати його у відповідний формат файлу.

6. Функція розпізнавання емоцій використовує попередньо навчену модель машинного навчання для розпізнавання емоцій у завантаженому відео на основі виділених рис обличчя в кожному кадрі. Функція повинна бути здатна аналізувати вираз обличчя виявленого виразу обличчя людини у кожному кадрі та визначати найбільш ймовірну емоцію.

7. Функція виводу надсилає розпізнану емоцію та іншу релевантну інформацію назад користувачеві.

Модуль інтерфейсу чат-бота

Цей модуль (рисунок 2.11) буде відповідати за обробку функцій для взаємодії користувача з чат-ботом.

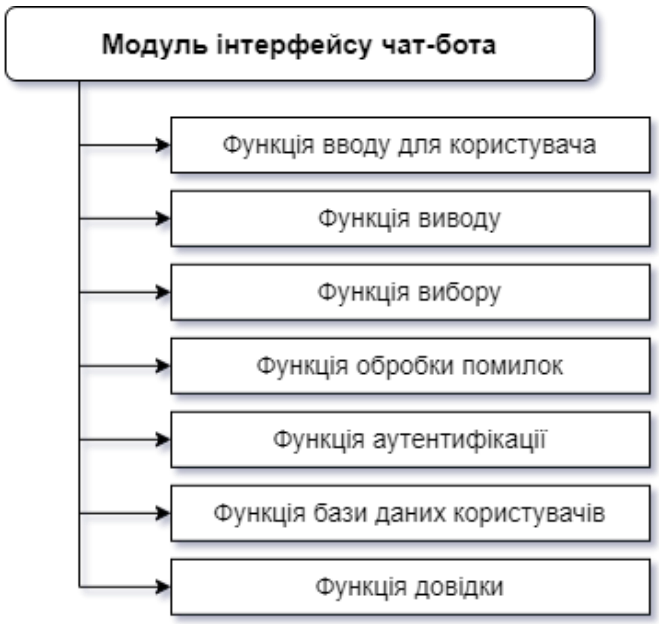


Рисунок 2.11 – Схема модулю для інтерфейсу чат-бота

Нижче описано функції модуля інтерфейсу чат-бота:

1. Функція вводу користувача прослуховує введення користувача та аналізує його для визначення намірів користувача. Функція повинна розуміти команди природною мовою, такі як "завантажити зображення" або "проаналізувати відео", і перенаправляти користувача до відповідного модуля.

2. Функція виводу генерує та надсилає відповідь користувачеві на основі результатів роботи модуля розпізнавання зображень або відео. Функція повинна мати можливість виводити розпізнану емоцію та будь-які відповідні метрики або статистику.

3. Функція вибору дозволяє користувачеві вибирати між аналізом зображення або відео. Функція повинна надавати користувачеві меню вибору і перенаправляти його до відповідного модуля на основі зробленого вибору.

4. Функція обробки помилок обробляє будь-які помилки або винятки, що виникають в процесі взаємодії з користувачем. Функція повинна визначати причину помилки та надавати користувачеві чітке повідомлення про помилку.

5. Функція аутентифікації перевіряє особу користувача і гарантує, що тільки авторизовані користувачі можуть отримати доступ до системи. Функція повинна мати можливість аутентифікувати користувачів за допомогою безпечного процесу входу в систему або інших механізмів аутентифікації.

6. Функція бази даних користувачів зберігає та витягує дані користувача, такі як уподобання, історія та налаштування. Функція повинна мати можливість створювати профілі користувачів та керувати ними, а також дозволяти користувачам налаштовувати свій досвід.

7. Функція довідки надає допомогу та вказівки користувачеві. Функція повинна надавати інструкції про те, як користуватися системою, і відповідати на найпоширеніші запитання.

У результаті було розроблено функціональну структуру інформаційну системи для розпізнавання емоцій людини на зображеннях та відео з обличчями, реалізовану у вигляді чат-бота в Telegram. Система складається з трьох функціональних модулів: Модуль розпізнавання зображень, Модуль розпізнавання відео та Модуль інтерфейсу чат-бота. Модуль розпізнавання

зображень та модуль розпізнавання відео відповідають за розпізнавання емоцій на зображеннях та відео відповідно, шляхом виявлення облич, ідентифікації рис обличчя та аналізу міміки. Модуль інтерфейсу чат-бота відповідає за взаємодію користувача з чат-ботом, включаючи введення, виведення, вибір, обробку помилок, автентифікацію, управління базою даних користувачів і довідкові функції. Включення функцій виводу в кожен модуль гарантує, що розпізнані емоції та інша відповідна інформація надсилається назад користувачеві. Розбивши систему на окремі модулі, було створено модульний і масштабований дизайн, який сприяє повторному використанню коду, гнучкості та модульності.

2.5 Проєктування інтерфейсу інформаційної системи

Інтерпретація емоцій у візуальних медіа може використовуватись, як для особистого використання, наприклад, для виявлення емоційних станів на фотографіях і відео, так і для професійного використання, наприклад, для аналізу емоційних реакцій у маркетингових кампаніях або в онлайн-оглядах. Однак розпізнавання емоцій у візуальних медіа може бути складним завданням для тих, кому бракує технічних знань чи інструментів. У такому випадку доцільно розробити інтерфейс інформаційної системи.

Екран/форма для входу: При першій взаємодії з ботом користувачеві неохідно запустити бота за допомогою свого облікового запису Telegram.

Інтерфейс завантаження медіа: Після початку роботи із системою, користувач зможе завантажити зображення або відео для розпізнавання емоцій скориставшись функціоналом додатку.

Результати розпізнавання емоцій: Після того, як користувач завантажить свої медіа файли, бот виконає розпізнавання емоцій і покаже користувачеві результати. Результати можуть відображатися у вигляді міток або тегів, прикріплених до медіа файлу, із зазначенням емоцій, виявлених у медіа (наприклад, щастя, смуток, гнів, страх тощо).

Функція завантаження: Користувач матиме можливість завантажити позначений медіа файл із доданими тегами емоцій. Загальна схема інтерфейсу інформаційної системи зображена на рисунку 2.12:

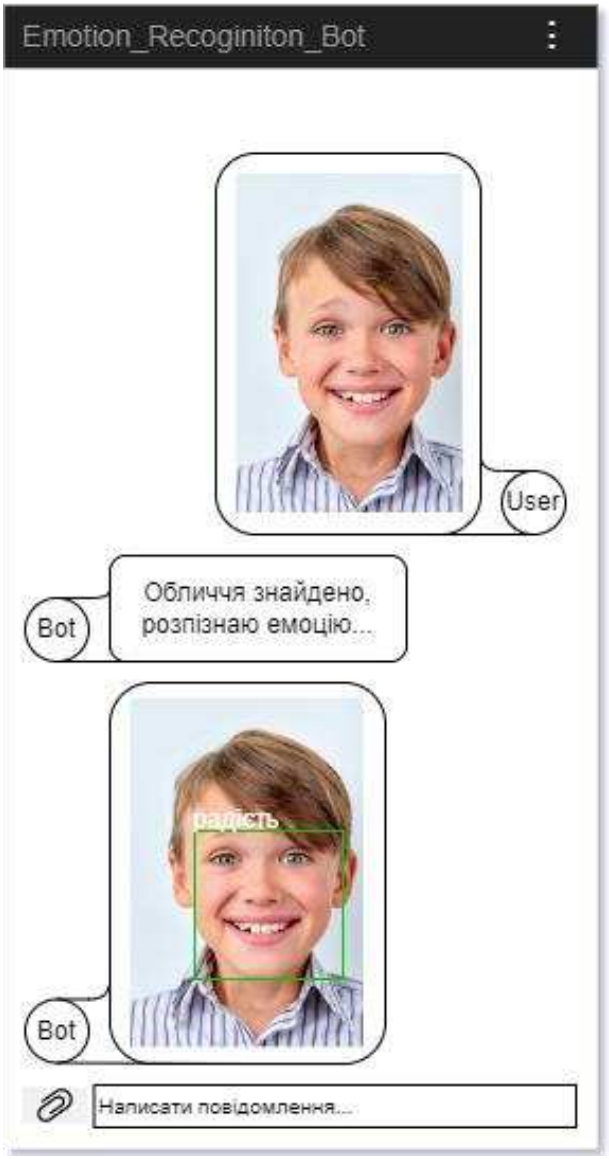


Рисунок 2.12 – Приклад інтерфейсу для інформаційної системи

Загалом, користувацький інтерфейс повинен бути простим та інтуїтивно зрозумілим, з чіткими підказками та інструкціями для користувача на кожному кроці процесу. Меню вибору мови та результати розпізнавання емоцій повинні бути легко читабельними та зрозумілими, з чіткими позначками та тегами, що вказують на емоції, виявлені в медіа. Функції завантаження та вивантаження медіа файлів мають бути легкодоступними та зручними для користувача, щоб

користувач міг швидко та ефективно завантажувати та вивантажувати медіа файли.

2.6 Висновки до розділу 2

У розділі було досліджено різні аспекти, пов’язані з розпізнаванням мімічних проявів емоцій за допомогою згорткових нейронних мереж.

Проаналізовано та визначено архітектуру згорткової нейронної мережі для класифікації емоційних станів на основі виразу обличчя. Було визначено кроки для попередньої обробки вхідних даних та обрано набір даних для подальшого навчання нейронної мережі.

Також було розглянуто функціональну структуру інформаційної системи, яка включала аналіз різних модулів програми. Модулі були розроблені для визначення та виконання таких завдань, як обробка зображень та відео, виділення ознак, класифікація та візуалізація результатів.

Було спроектовано інтерфейс інформаційної системи. Інтерфейс користувача відіграє вирішальну роль у полегшенні взаємодії та забезпеченні безперешкодної роботи для користувачів, забезпечуючи ефективне використання системи. Для реалізації обраних методів та кроків було обрано мову програмування Python, платформу TensorFlow та бібліотеку OpenCV. Реалізація інтерфейсу буде здійснюватися за допомогою бібліотеки python-telegram-bot.

У підсумку, розділ 2 розкриває різні компоненти та кроки, необхідні для реалізації системи розпізнавання емоційних виразів обличчя за допомогою згорткових нейронних мереж. Інтегрування обраного способу обробки зображень, геометричної інтерпретації мімічних проявів, архітектури CNN, функціональних модулей та інтерфейсу було запропоновано інформаційну систему, що є перспективною для точної класифікації емоційних станів за мімічними проявами.

Розділ 3 Програмна реалізація інформаційної системи розпізнавання мімічних проявів емоцій

3.1 Структура та функціональне призначення програмних складових інформаційної системи

Було створено інформаційну систему, яка може розпізнавати мімічні прояви емоції людини, чи групи людей за зображеннями та відео. Система складається з двох основних модулів, один із, використовується для навчання моделі згорткової нейронної мережі (CNN) за набором даних, а інший для надання користувацького інтерфейсу через додаток-бот Telegram, який завантажує навчену модель і розпізнає емоції людей на вхідному відео або зображенні та видає результат розпізнавання. Додаток телеграм-бота є основним компонентом системи, який надає користувачеві інтерфейс для доступу до функції розпізнавання мімічних проявів емоцій.

Модуль 1: Навчання моделі CNN

Перший модуль системи відповідає за навчання CNN-моделі, яка може розпізнавати мімічні прояви емоцій на зображеннях або відео, які містять обличчя людей. Навчальний модуль складається з декількох підмодулів, включаючи попередню обробку даних, побудову моделі, навчання та оцінювання моделі.

Попередня обробка даних передбачає збір, очищення та підготовку даних для моделі CNN. За це відповідають функції preprocess_data та clean_data_and_normalize. Спочатку набір даних завантажується за допомогою функції read_csv() Pandas, а значення пікселів витягуються і перетворюються у відтінки сірого 48x48 за допомогою функції fromstring() NumPy. Мітки також завантажуються і перетворюються на ймовірності від 0 до 1 за допомогою виразу y[:, :-2] * 0.1, а клас "презирство" об’єднується з класом "нейтралітет".

Потім дані розбиваються на навчальну, валідаційну та тестову вибірки за допомогою функції train_test_split() Scikit-Learn. Збільшення даних виконується за допомогою функції ImageDataGenerator() Keras, яка обертає, перевертає та зсуває зображення для збільшення розміру набору даних.

Підмодуль побудови моделі передбачає створення моделі CNN у функції define_model, включаючи її архітектуру та параметри, які були визначені у п. 2.3 другого розділу (рисунок 2.7). Модель визначена за допомогою Keras Sequential API і складається з п’яти етапів згортки та об’єднання шарів, за якими слідують два повністю з’єднані шари. Перший етап має два згорткових шари і шар, що відсівається. Другий етап складається з двох згорткових шарів і максимального об’єднаного шару. Третій і четвертий етапи мають по два згорткових шари. Нарешті, п’ятий етап має два згорткові шари, і всі етапи мають шари пакетної нормалізації. Докладна архітектура використаної моделі продемонстрована у додатку Б на рисунку Б.1.

Модель навчається за допомогою оптимізатора Адама і категоріальної функції втрат перехресної ентропії, а точність і втрати будуються за епохами навчання за допомогою бібліотеки pyplot з Matplotlib. Навчена модель CNN зберігається у вигляді набору ваг і зсувів, які будуть завантажені та використані в другому модуль системи.

Модуль 2: Додаток Telegram-бота (Модуль розпізнавання)

Модуль системи відповідає за надання користувацького інтерфейсу для доступу до функції розпізнавання мімічних проявів емоцій. Він складається з підмодулів, які включають інтеграцію з телеграм-ботом, завантаження навченої моделі, розпізнавання вхідних зображень та відео, видачу результату.

Для підключення до телеграм-боту використовується клас ApplicationBuilder, який проводить ініціалізацію та підключення боту за вказаним токеном. Інтеграція з телеграм-ботом дозволяє користувачам надсилати вхідні зображення або відео до програми через чат з телеграм-ботом. Після ініціалізації на стороні сервера бот завантажує в пам’ять раніше навчену модель CNN. Тоді система ініціалізує об’єкти з класів Handler подій у боті та очікує вхідні запити.

Загалом, бот використовує один об’єкт класу CommandHandler та три об’єкти класу MessageHandler:

– start_handler – об’єкт класу CommandHandler, який обробляє початкову команду «start» та видає привітальне повідомлення з інструкціями для користувача, коли він вперше його запускає.

– echo_handler – об’єкт класу MessageHandler, обробляє повідомлення та події, які не передбачені у системі, викликає функцію echo та надсилає користувачеві нагадування, коли користувач надсилає боту повідомлення замість того, щоб надіслати зображення або відео.

– image_handler – об’єкт класу MessageHandler, обробляє вхідні повідомлення, які містять зображення, обробляються відповідним фільтром та передають вхідне зображення до функції rec_image для подальшого розпізнавання.

– video_handler – об’єкт класу MessageHandler, обробляє вхідні повідомлення та відео повідомлення, які містять відеофайл, обробляються відповідним фільтром та передають вхідне відеофайл до функції rec_video для подальшого розпізнавання.

Коли користувач надсилає повідомлення, бот фільтрує його за типом та передає до відповідної функції. В залежності від файлу, зображення чи відео – система передає файл у одну з двох функцій: rec_image або rec_video.

Функція rec_image відповідає за розпізнавання емоцій на зображенні, надісланому користувачем. Вона отримує ідентифікатор файлу зображення з API Telegram і завантажує його у вказану директорію. Далі використовується метод detectMultiScale з класу cv2.CascadeClassifier для виявлення облич на зображенні. Метод застосовує каскадний класифікатор Хаара до версії зображення у відтінках сірого. Якщо класифікатор виявляє обличчя, скрипт застосовує до нього раніше завантажену навчену нейронну мережу для розпізнавання мімічних проявів емоції людини. Розпізнана емоція записується на зображення, і зображення надсилається назад користувачеві разом із повідомленням про те, чи було виявлено обличчя.

Функція rec_video відповідає за розпізнавання мімічних проявів емоцій у відео, надісланому користувачем. Вона працює подібно до функції rec_image, але замість того, щоб обробляти все відео відразу, вона обробляє відео кадр за кадром

за допомогою класу cv2.VideoCapture і записує розпізнану емоцію на кожен кадр. Нарешті, кадри із розпізнаними емоціями об’єднуються для створення нового відео, яке надсилається назад користувачеві. Схема структури програмних складових інформаційної системи наведена на рисунку 3.1.

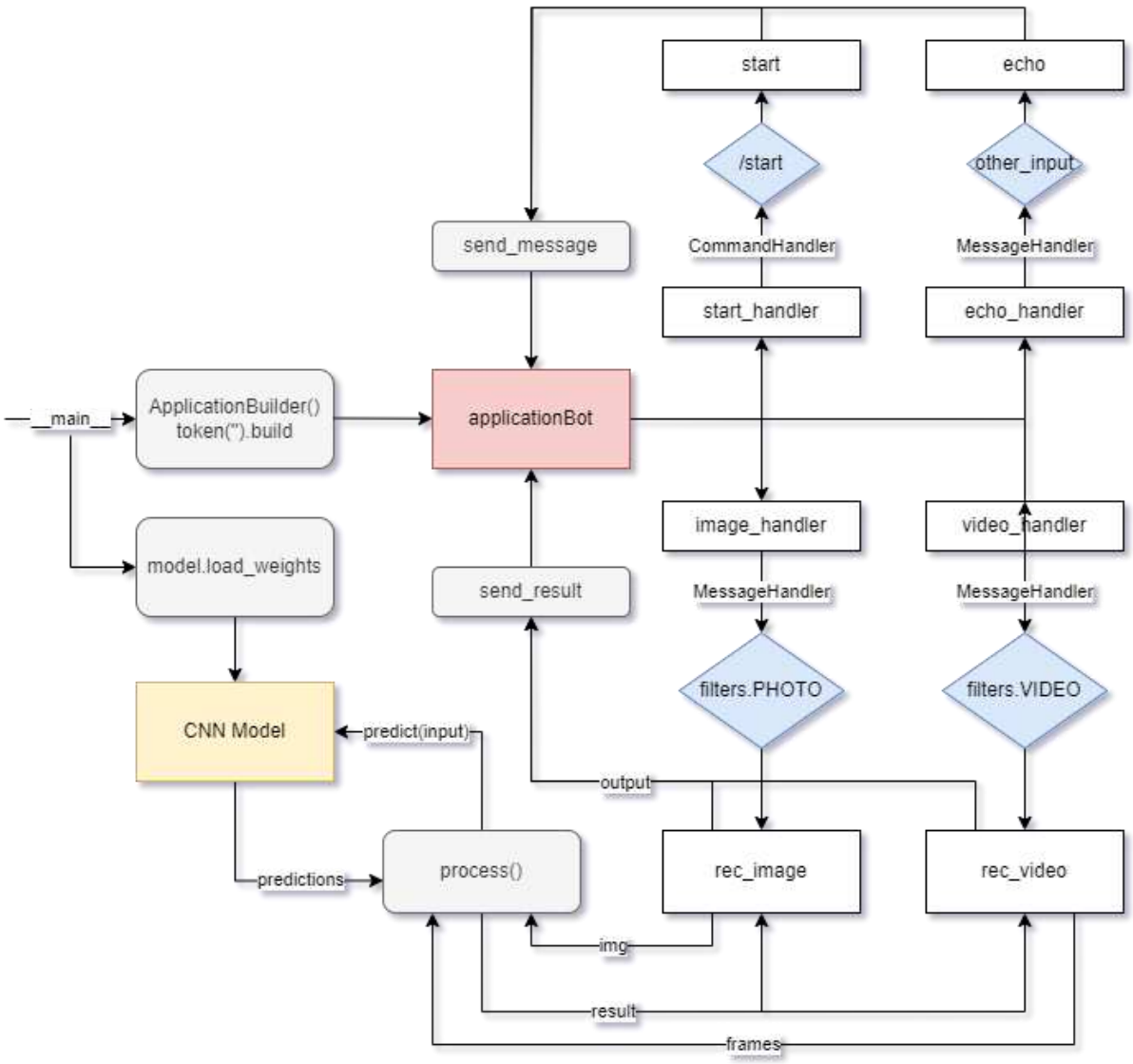


Рисунок 3.1 – Схема структура програмних складових інформаційної системи

Додаток телеграм-бота реалізовано за допомогою мови програмування Python і використовує сторонню бібліотеку python-telegram-bot для взаємодії з API Telegram.

Розроблена інформаційна система може бути корисною в різних сферах, таких як маркетингові дослідження, охорона здоров’я та безпека, де аналіз

людських емоцій має вирішальне значення. Точність і швидкість системи роблять її ефективним інструментом для аналізу емоцій на основі великих наборів даних у режимі реального часу.

3.2 Особливості реалізації програмних складових інформаційної системи

Основний модуль інформаційної системи – розпізнавання мімічних проявів емоцій за зображенням або відео, яке містить обличчя людини. Для цього було реалізовано функції, які виконують попередню обробку зображення та відео. Після отримання файлу системою і передачі його ідентифікатора до функції розпізнавання, в залежності від типу вхідного файлу система за фільтрами обирає необхідну функцію. Схему процесу обробки та розпізнавання вхідних даних подано на рисунку 3.2.

Коли програма отримує на вході зображення, вона звертається до функції rec_image, яка отримує вхідний файл за ідентифікатором. Лістинг програмного коду подано у додатку В. Вхідний файл проходить наступні кроки:

1. Збереження на диск за допомогою функції download_to_drive, у папку «in_image/».

2. Завантаження у пам’ять за допомогою cv2.imread. Функція imread завантажує зображення з вказаного файлу і повертає його.

3. Після завантаження зображення, використовується функція cv2.cvtColor(img, cv2.COLOR_BGR2GRAY) для перетворення кольорового простору з RGB до градацій сірого.

4. Далі, використовується detectMultiScale з класу cv2.CascadeClassifier для виявлення обличчя на зображенні.

5. Проводиться перевірка на наявність обличчя на зображенні за допомогою «if faces_detected.any():», якщо обличчя не знайдено – система надсилає попередження користувачу і переходить у очікування.

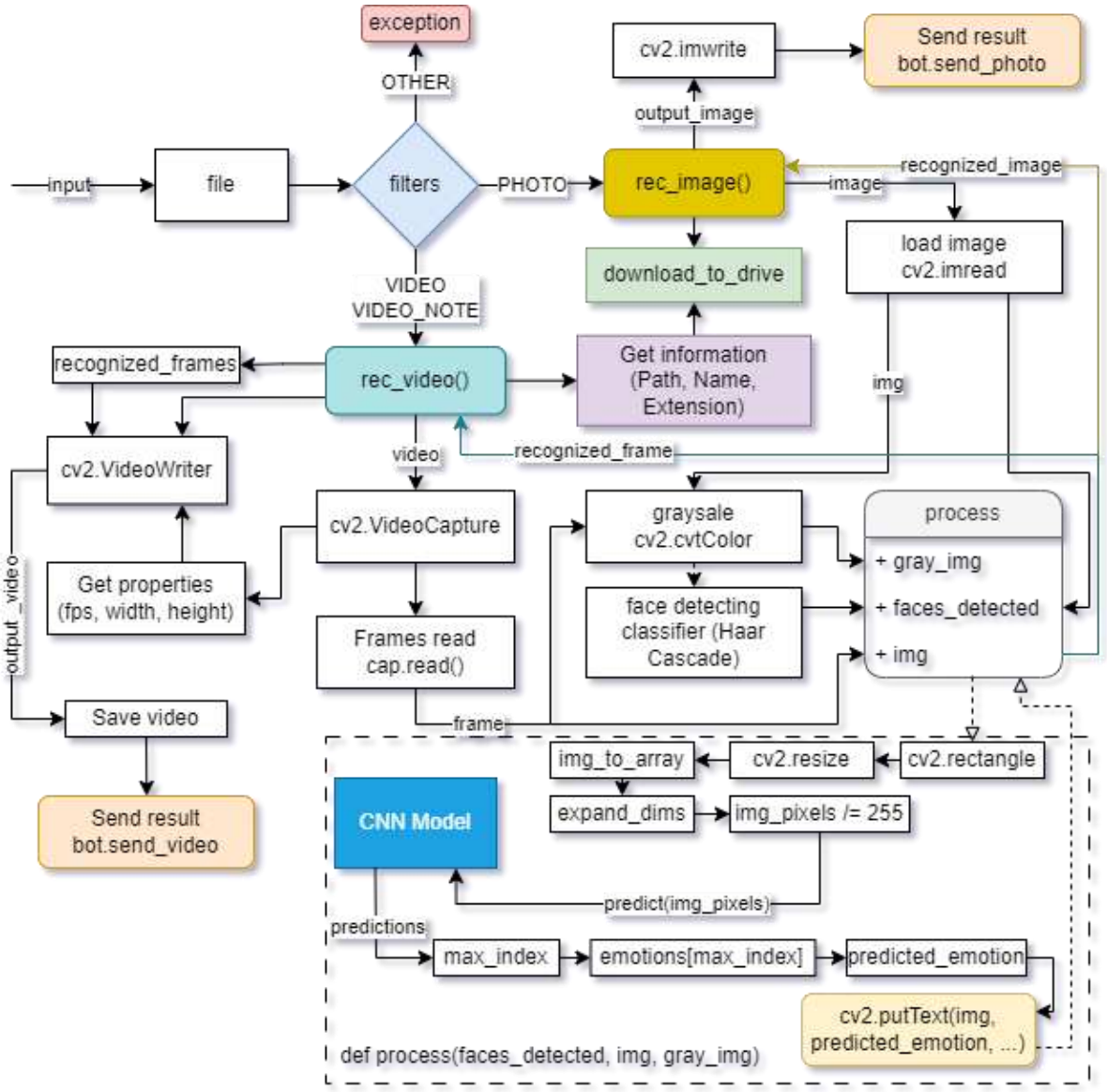


Рисунок 3.2 – Схема процесу обробки та розпізнавання вхідних даних

6. Коли обличчя знайдено – система переходить до попередньої обробки зображення. Для кожного виявленого обличчя (заданого координатами x, y, w, h) код наносить зелений прямокутник на вхідне зображення за допомогою функції cv2.rectangle() з OpenCV. Після цього код виділяє область зображення, що містить виявлене обличчя, за допомогою оператора зрізу масиву NumPy – gray_img[y:y + w, x:x + h]. Далі, використовуючи функцію cv2.resize(), область зображення зменшується до розміру 48x48 пікселів, що є стандартним розміром для зображень, які використовуються для розпізнавання мімічних проявів емоцій.

7. Зменшене зображення конвертується до масиву NumPy за допомогою функції image.img_to_array() з бібліотеки keras.preprocessing.image. Після цього

масив зображення розширюється до 4D за допомогою функції np.expand_dims(), що дозволяє передати його в нейромережу Keras для подальшої обробки.

8. Насамкінець, масив зображення нормалізується, ділячи кожен піксель на 255 за допомогою операції ділення, що дає значення від 0 до 1. Це необхідно для відповідальності до параметрів моделі.

9. Після цього predictions = model.predict(img_pixels) – пропускає нормалізовані значення пікселів через модель CNN для отримання прогнозу емоції в ROI. Далі max_index = int(np.argmax(predictions)) витягує індекс найвищої передбачуваної емоції з результату моделі CNN.

10. Далі використовується список, який містить різні емоції, які може передбачити модель CNN «emotions = [‘neutral’, ‘happiness’, ‘surprise’, ‘sad’, ‘anger’, ‘disgust’, ‘fear’]», тоді, використовується рядок predicted_emotion = emotions[max_index], який витягує передбачену емоцію на основі індексу, отриманого на кроці 9.

11. Передбачену емоцію необхідно накласти на вхідне зображення, вирівнюючи його за виявленим обличчям, для цього використовується функція cv2.putText.

12. Після розпізнавання, зображення з виявленими обличчями та передбаченими емоціями зберігається на диск за допомогою функції cv2.imwrite(path_2,img) у файл. Також, за допомогою функції send_photo відправляє розпізнане зображення користувачу.

Для розпізнавання мімічних проявів емоцій на відео використовуються аналогічні кроки, з деякими змінами відповідно до відеофайлів. Таким чином, код ініціалізує об’єкт VideoCapture з бібліотеки OpenCV для зчитування кадрів з відео. Він також отримує частоту кадрів, ширину та висоту відео за допомогою методів get(cv2.CAP_PROP_FPS), get(cv2.CAP_PROP_FRAME_WIDTH) та get(cv2.CAP_PROP_FRAME_HEIGHT) відповідно.

Тоді, обробляється кожен кадр відео. Зчитується кожен кадр з об’єкта VideoCapture, використовуючи ті ж кроки, як і з зображенням, конвертує його у відтінки сірого і виявляє обличчя на зображенні у за допомогою класифікатора

Haar Cascade. Потім він обрізає обличчя з зображення і змінює їх розмір до 48x48 пікселів. Далі, обрізані зображення пропускаються через модель згорткової нейронної мережі (CNN) для прогнозування емоції, відображеної на обличчі. Прогнозована емоція записується на кадрі у вигляді тексту за допомогою методу cv2.putText.

Отримані кадри зберігаються у форматі JPEG у тимчасовому каталозі temp_video. Після того, як всі кадри оброблені, вони об’єднуються у нове відео за допомогою об’єкта VideoWriter з OpenCV. Отримане відео зберігається у каталозі out_video.

Нарешті, код видаляє тимчасові JPEG-файли і надсилає оброблене відео користувачеві за допомогою методу send_video або send_video_note API бота, залежно від того, чи було початкове повідомлення відео або відео-нотаткою.

3.3 Експериментальне тестування інформаційної системи

3.3.1 Тестування системи за валідаційним набором даних

У розділі представлено результати експериментального тестування інформаційної системи, призначеної для розпізнавання виразів емоцій на зображеннях людей. Метою експерименту є оцінка точності та продуктивності системи в прогнозуванні емоцій на основі вхідних зображень. Результати тестування аналізуються я, виділяючи загальну оцінку правильності та точність для кожної категорії емоцій.

Для тестування моделі була використана частина набору даних ADFES. Скомпонований набір, що складається з 154 відзнятих емоційних виразів.

– Набір містить сім базових емоцій (гнів, відраза, страх, радість, смуток і здивування і спокійний стан).

– Вирази обличчя демонструють 22 моделі (10 жіночих, 12 чоловічих).

– До набору входять як північноєвропейські, так і середземноморські моделі.

Емоції були позначені у форматі "GenderNumberEmotion.png", де "Emotion" – це власне категорія емоції, "Gender" – стать людини, а "Number" – порядковий номер моделі на зображенні. Система використовувала попередньо навчену модель, завантажену з JSON-файлу, та відповідні ваги. Модель розпізнавала емоції, аналізуючи мімічні прояви на обличчі, витягнуті з вхідних зображень. Потім передбачені емоції порівнювалися з мітками, щоб оцінити точність системи.

Результати експериментального тестування показують, що інформаційна система досягла загального показника правильності у 80,52%. Цей показник відображає відсоток правильно передбачених емоцій серед усіх протестованих зображень. Результати розпізнавання наведені на рисунку 3.3.

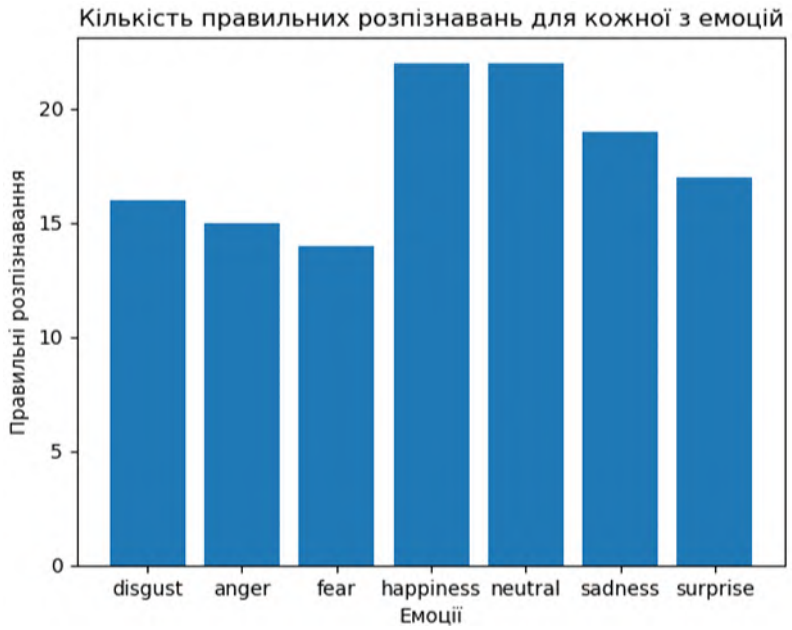


Рисунок 3.3 – Кількість правильних розпізнавань для кожної з емоцій

Продуктивність навченої моделі аналізується за допомогою наступних оціночних метрик:

1. Точність (Precision): Точність вимірює частку правильно передбачених випадків серед усіх передбачених випадків для певної емоції.

2. Повнота (Recall): Відтворення вимірює частку правильно передбачених випадків серед фактичних випадків певної емоції.

3. F-Norm: F-норма, також відома як F1-score, є середнім гармонійним значенням точності та пригадування.

Значення точності, повноти та F-норми для кожної емоції та загальні показники наведені у таблиці 3.1.

Навчена модель, продемонструвала гарну ефективність у розпізнаванні мімічних проявів емоцій. Із загальною точністю 80,52% модель продемонструвала вміння розрізняти основні типи емоцій. Зокрема, вона продемонструвала чудову точність і здатність розпізнавання щастя, досягнувши показників точності та повноти у 100%, що свідчить про її здатність точно ідентифікувати радісний вираз обличчя. Крім того, модель продемонструвала задовільну точність і повноту розпізнавання відрази, страху, смутку і здивування – від 64,71% до 94,20% точності і від 63,64% до 86,36% повноти розпізнавання.

Таблиця 3.1 – Значення метрик для оцінки моделі

Emotion	Precision	Recall	F-Norm
Disgust	94.12%	72.73%	82.05%
Anger	78.95%	68.18%	73.17%
Fear	87.50%	63.64%	73.68%
Happiness	100.00%	100.00%	100.00%
Neutral	64.71%	100.00%	78.57%
Sadness	73.08%	86.36%	79.17%
Surprise	80.00%	72.73%	76.19%
Загалом	80.52%	80.52%	80.40%

Однак були певні області, де продуктивність моделі можна було б покращити. Зокрема, точність для нейтральних виразів була відносно низькою – 64,71%, що свідчить про більшу кількість помилкових спрацьовувань. Крім того, модель намагалася досягти вищої точності та пригадування для таких емоцій, як гнів, де точність досягла 78,95%, але повнота залишилася на рівні 68,18%. Це свідчить про те, що моделі було складно правильно ідентифікувати випадки гніву в наборі даних.

Отримані результати дають цінну інформацію про сильні та слабкі сторони моделі. Подальші дослідження та вдосконалення моделі можуть призвести до покращення її продуктивності та полегшити її інтеграцію в різні ПЗ.

3.3.2 Тестування системи за реальних умов використання

Проведене тестування навченої моделі на тестовому наборі даних показало, що модель може розпізнавати мімічні прояви на обличчі людини. Для остаточного підтвердження повного функціонування розробленої інформаційної системи необхідно провести тестування модулів, які було визначено в п. 2.4, за реальних умов.

Для перевірки коректної роботи системи, необхідно обрати вхідні дані, з якими буде проводитись тестування. Розпочати перевірку варто з фото, які містять одну людину, яка виражає своєю мімікою якийсь з емоційних станів. Було надіслано боту кілька зображень з певними емоціями. Результат розпізнавання зображено на рисунку 3.4.

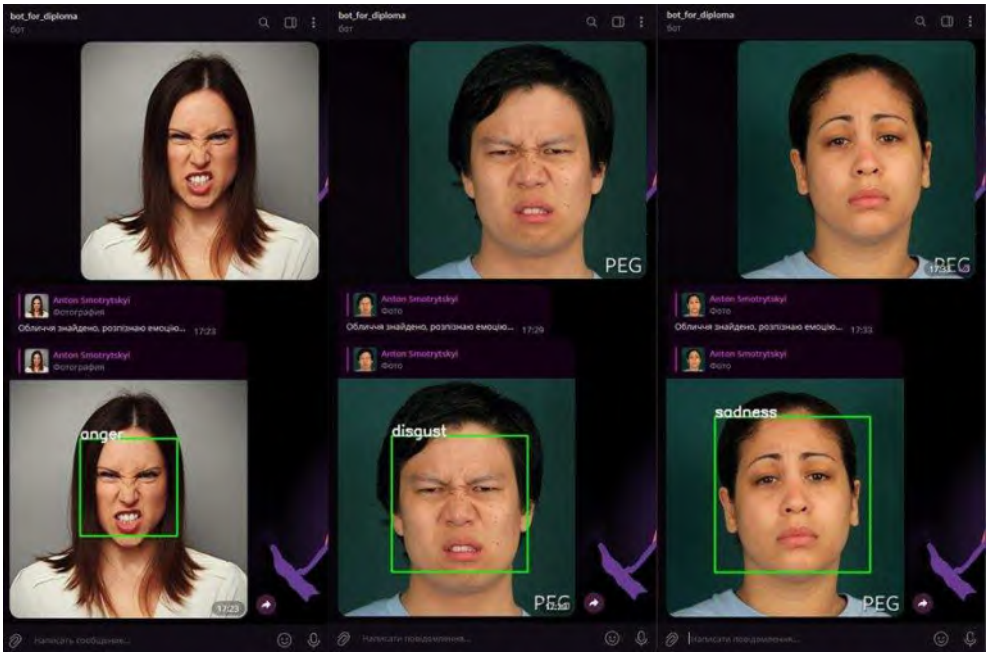


Рисунок 3.4 – Розпізнавання зображень за допомогою телеграм-боту

У додатку Ґ на рисунку Ґ.1 продемонстровано розпізнавання кожної з можливих емоцій, які можна розпізнати за мімічними проявами.

Також, при розпізнаванні мімічних проявів на зображенні, доцільно перевірити роботу виявлення кількох обличь на зображенні, для випадків великої кількості людей на зображенні (групові фото, фото натовпу, тощо). На рисунку 3.5 зображено випадок групового фото з великою кількістю людей.

Рисунок 3.5 – Розпізнавання облич на груповому фото

Результати розпізнавання зображень показують, що система може виявляти велику кількість облич на зображені та розпізнавати мімічні прояви для кожного із них.

Далі, перевіримо функціонування модулю для обробки та розпізнавання мімічних проявів на відео файлах. Для розпізнавання було надіслано коротке відео, на якому зображені люди, які спілкуються між собою. Результат роботи програми зображено на рисунку 3.6.

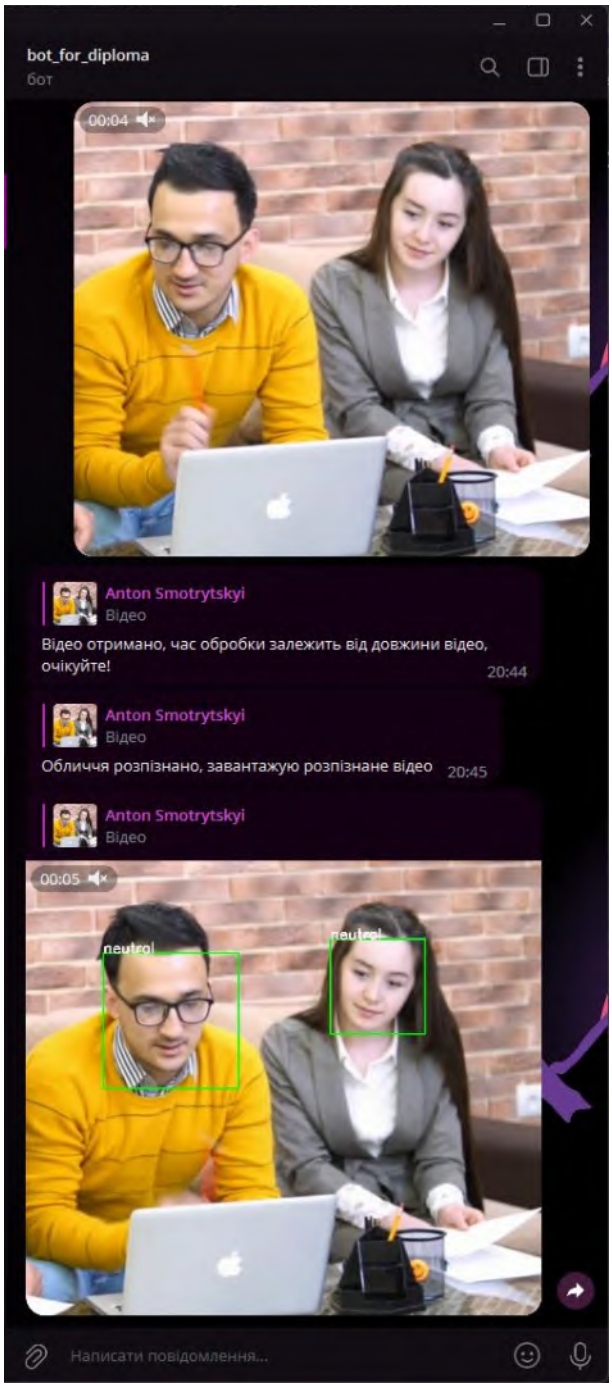


Рисунок 3.6 – Результат обробки та розпізнавання облич на відео

На рисунку 3.6 показано, що система успішно обробила та розпізнала вираз обличчя у наданому відеофайлі. Додатково, було протестовано розпізнавання

вхідного відео повідомлення від звичайного користувача, записане з мобільного пристрою. На рисунку 3.7 показана імітація кількох емоційних станів у вхідному повідомленні та їх розпізнавання.

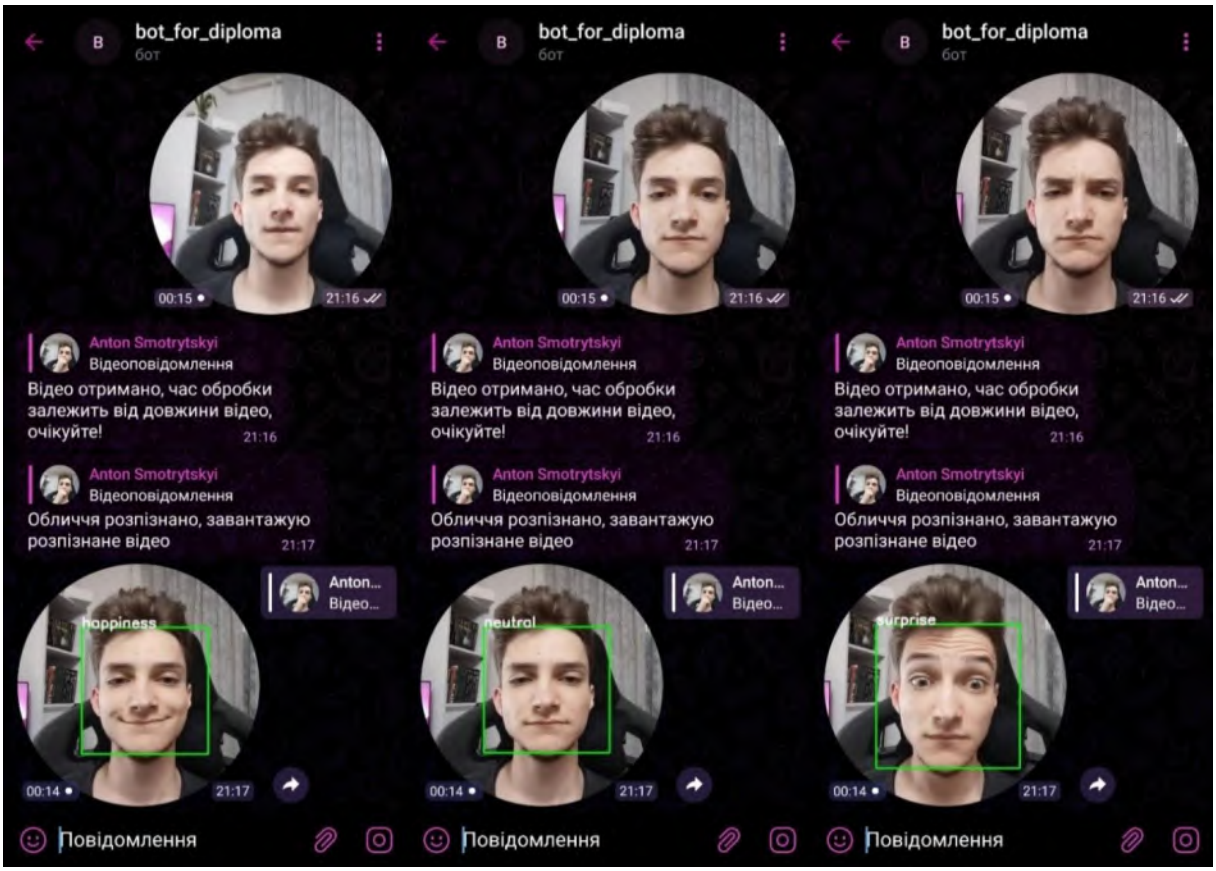


Рисунок 3.7 – Розпізнане відео повідомлення

Система чудово впоралась із обробкою та розпізнаванням відео та відео повідомлень. Це свідчить про те, що розроблена інформаційна система здатна точно виявляти та інтерпретувати мімічні прояви емоцій в сценаріях реального часу.

Функціональне тестування модулів, підтвердило повноцінне функціонування розробленої інформаційної системи. Система демонструє здатність точно розпізнавати вираз обличчя як на статичних зображеннях, так і на відеофайлах. Після успішного завершення тестування можна зробити висновок, що розроблена інформаційна система ефективно виконує свою функцію розпізнавання мімічних проявів емоцій.

3.4 Вимоги до розгортання інформаційної системи та інструкція користувача

Для розгортання інформаційної системи на іншому комп’ютері необхідно внести деякі зміни у файл run_bot.py. Якщо необхідно створити нового бота, для цього потрібно звернутись до бота «BotFather» у телеграм та отримати токен для доступу як показано на рисунку 3.8.

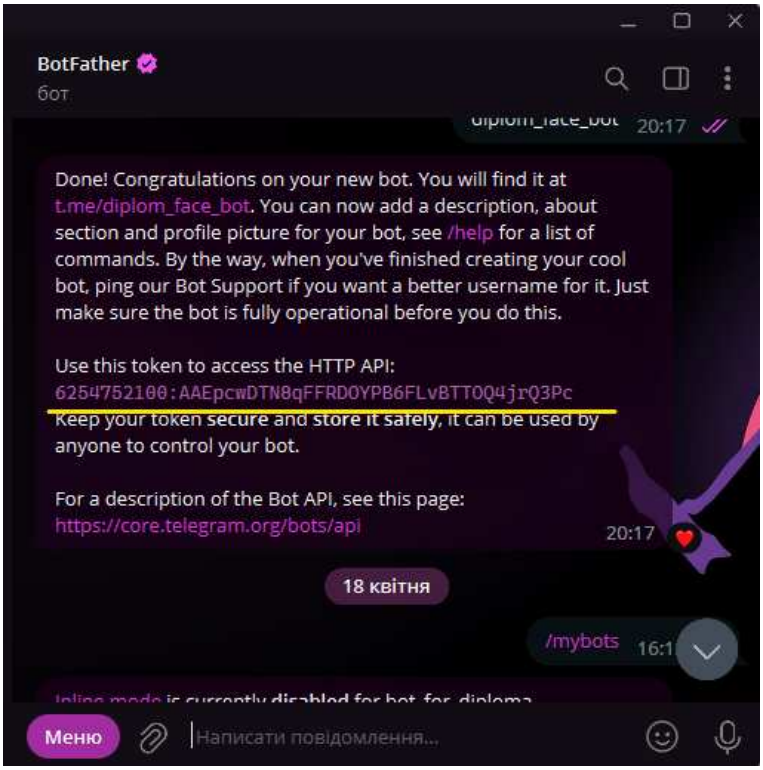


Рисунок 3.8 – Приклад отримання токену

Отриманий токен необхідно вставити у рядок:

```
application = ApplicationBuilder().token('Ваш токен').build()
```

Після цього, за необхідністю, потрібно оновити шляхи path_1 та path_2 для збереження файлів у функціях rec_image та rec_video, та шлях frame_dir для зберігання тимчасових кадрів для відео. Після внесення цих змін та виконання файлу бот запуститься і можна перейди до використання.

Для початки роботи з ботом, необхідно знайти його за іменем, яке було визначене у BotFather, або за навчальним іменем, яке використовувалось для тестування «@diplom_face_bot». Далі, продемонстровано роботу із ботом використовуючи мобільний пристрій.

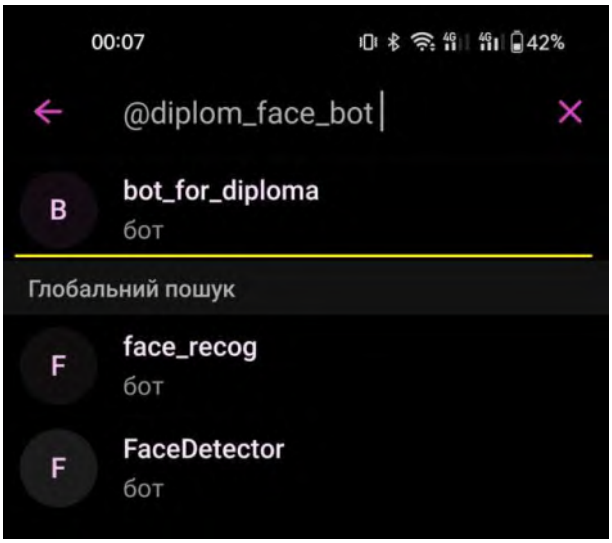


Рисунок 3.9 – Пошук бота у Telegram

Коли бота знайдено у телеграм, необхідно відкрити із ним чат та надіслати команду «/start» для початку роботи. Тоді, користувач отримає привітальне повідомлення з інструкціями (рисунок 3.10)

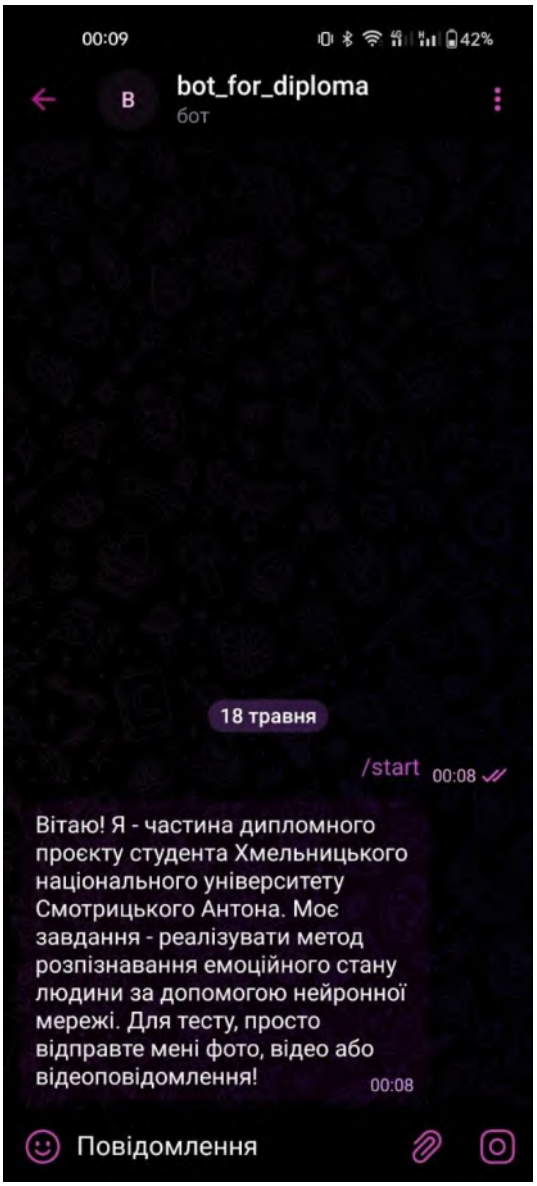


Рисунок 3.10 – Початок роботи із системою

Наступним кроком для користувача, буде відправка файлу. Він може надіслати зображення, натиснувши на значок скріпки біля поля вводу повідомлення. Та у вікні вибору, обрати зображення або зробити його на камеру пристрою (рисунок 3.11).

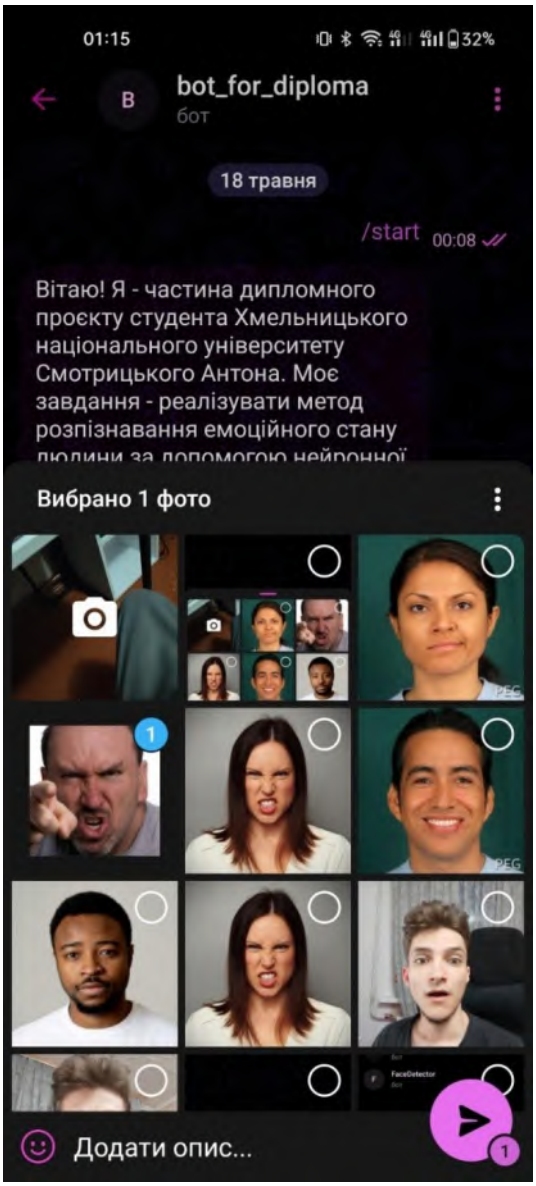


Рисунок 3.11 – Меню вибору зображення для відправки

Коли користувач обрав зображення він відправляє його для розпізнавання та отримує від бота повідомлення, яке інформує про те, чи наявне обличчя людини на зображені, якщо ні – система попросить інше зображення відповідним повідомленням. Якщо ж зображення містить хоча б одне зображення – система розпізнає емоцію та відправить результат користувачу (рисунок 3.12).

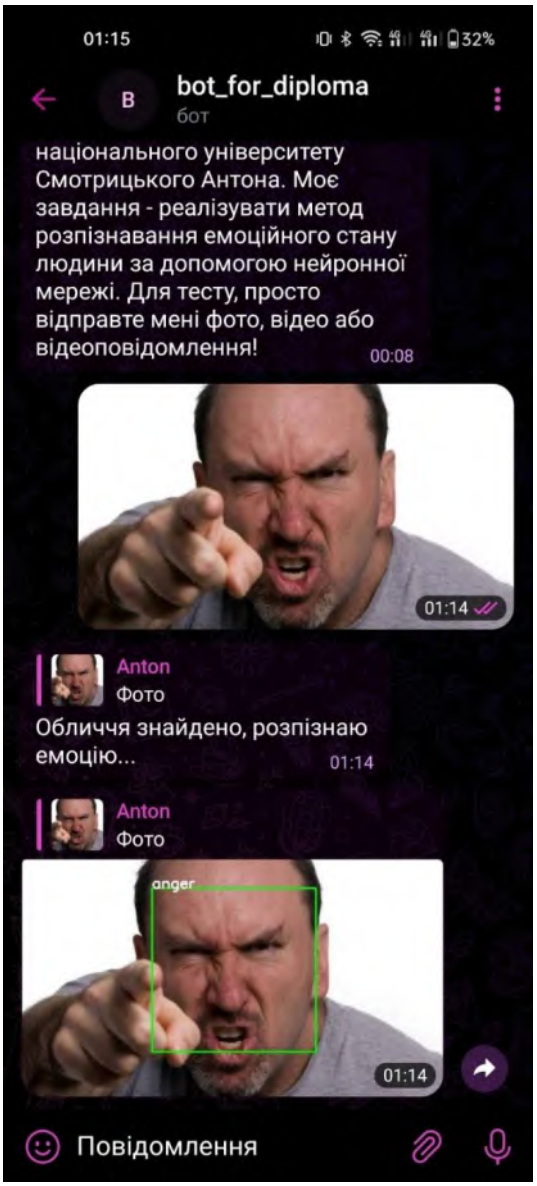


Рисунок 3.12 – Розпізнане зображення надіслане користувачем

Після розпізнавання користувач отримує у відповідь те ж зображення, але вже з виявленим обличчям та розпізнаною емоцією. Також користувач може скористатись завантаженням відео, аналогічно до вибору зображення. Або ж правіше від скріпки натиснути кнопку для запису відео повідомлення та надіслати його як показано на рисунку 3.13.

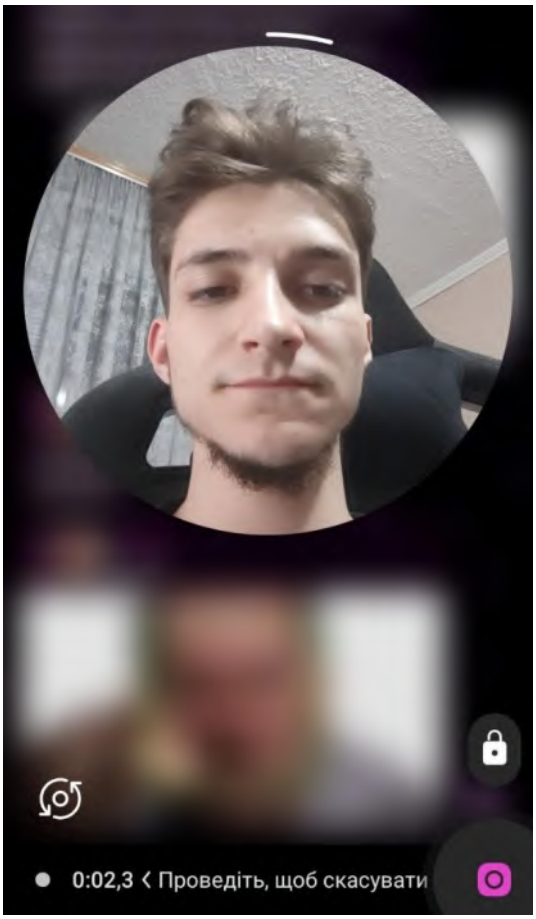


Рисунок 3.13 – Процес запису відео повідомлення

Коли система отримує вхідне повідомлення він повідомить користувача про початок обробки та попросить зачекати, оскільки обробка та розпізнавання відео файлу займає трохи більше часу. Коли відео буде розпізнано, користувач також отримує про це сповіщення та, через деякий час, розпізнане відео (рисунок 3.14)

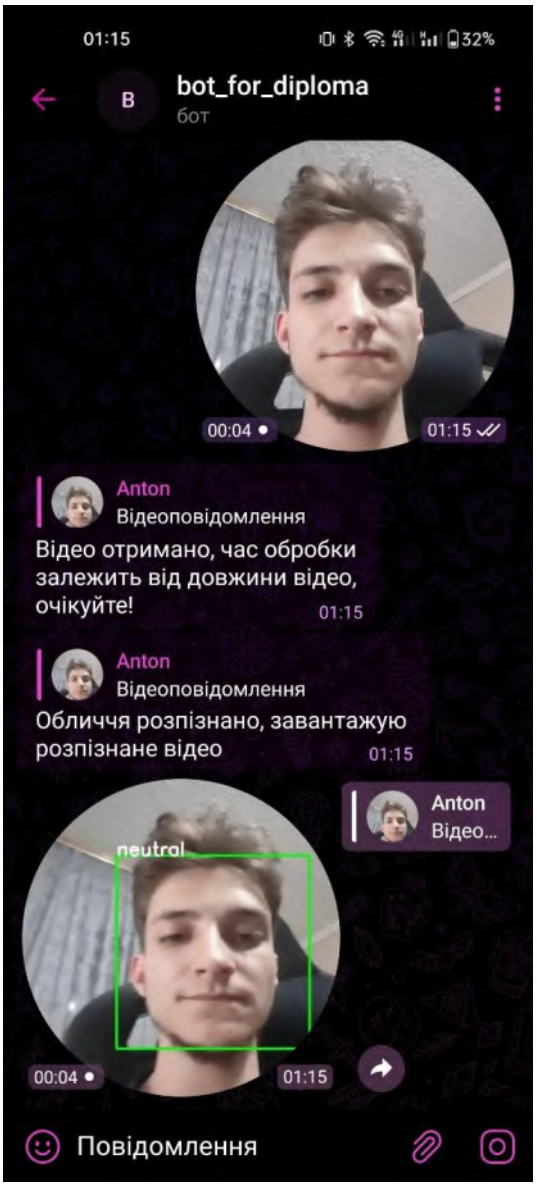


Рисунок 3.14 – Результат розпізнавання відео повідомлення

Після отримання результату, користувач може переглянути розпізнане зображення або відео та зберегти їх собі на пристрій для подальших дій. Оскільки, окрім функцій розпізнавання зображень та відео файлів не передбачено, будь які дії, які не стосуються роботи системи будуть повернені з інформуванням про функціонал. Приклад такого випадку зображений на рисунку 3.15.

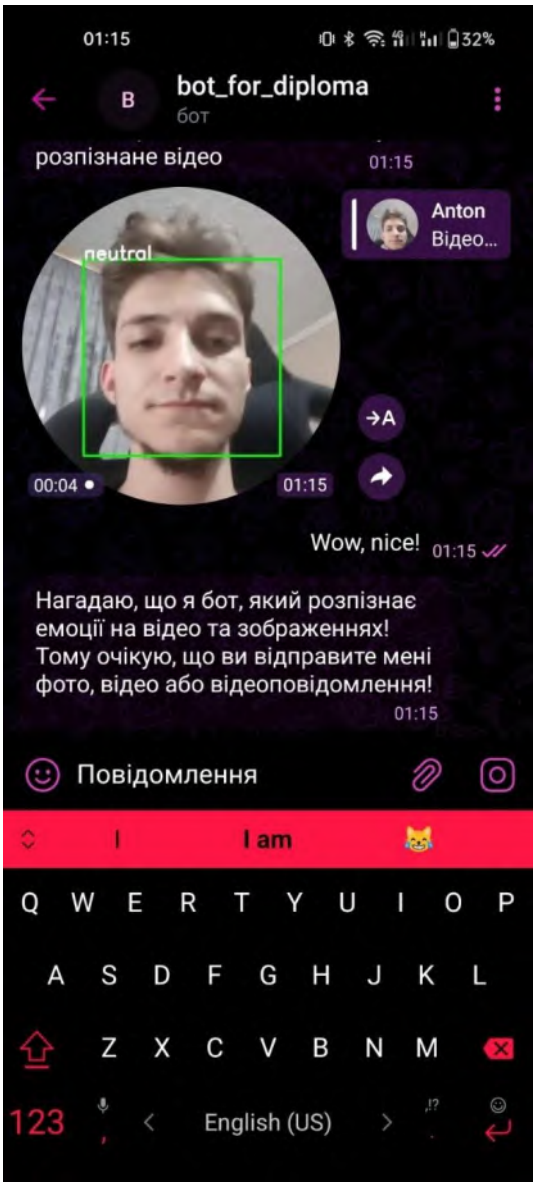


Рисунок 3.15 – Приклад повідомлення, яке не передбачено системою

Після отримання результатів користувач може надіслати інший медіа файл або завершити роботу із системою.

Інформаційну систему було розроблено з використанням мови програмування Python. Для розгортання та функціонування телеграм боту необхідні такі бібліотеки:

– Python >= 3.8.x.
– python-telegram bot >= v20.x.
– OpenCV >= 4.6.
– NumPy >= 1.20.x.
– TensorFlow >= 2.x.

Рекомендована операційна система Windows 10 або новіше. Linux та MacOS також підтримуються.

3.5 Висновки до розділу 3

Загалом, розроблена інформаційна система продемонструвала свій потенціал як ефективний інструмент для аналізу людських емоцій. Програмні компоненти системи були розроблені для виконання конкретних функціональних завдань, основним модулем яких є розпізнавання мімічних проявів емоцій з зображень або відео. Під час експериментального тестування система показала хороші результати, продемонструвавши високий рівень точності в прогнозуванні емоцій як на зображеннях, так і на відео.

Впровадження програмних компонентів, включаючи функції попередньої обробки зображень і відео, дозволило використовувати навчену модель для розпізнавання мімічних проявів емоцій у вигляді телеграм боту. Завдяки використанню фільтрів і вибору відповідної функції на основі типу вхідного файлу, система оптимізує свою роботу і підвищує ефективність. Експериментальне тестування та статистичне оцінювання інформаційної системи продемонстрували високу точність її роботи.

Таким чином, інформаційна система здатна розпізнавати мімічні прояви емоцій із точністю у 80,52%. найкраще система показує себе із емоцією радості та демонструє 100% точність та повноту. Для таких емоцій як відраза, страх та здивування показники точності досить хороші, але їх повнота поступається у показнику повноти. Окремо варто виділити нейтральний стан, де повнота показує себе ідеально, а ось точність дещо гірше, через хибні спрацювання. Але, у загальному, система в реальних умовах використання здатна показувати чудові результати у розпізнаванні мімічних проявів емоцій.

Точні прогнози системи в розпізнаванні емоцій на зображеннях і відео ще раз підтверджують її надійність і потенційні можливості застосування в різних сферах.

Висновки

Результати виконання кваліфікаційної роботи бакалавра підкреслюють актуальність і важливість інформаційних технологій у системах розпізнавання мімічних проявів емоцій, особливо в контексті розпізнавання емоційного виразу обличчя. Ці технології значно підвищили точність і ефективність цих систем, що призвело до широкого спектру практичних і дослідницьких застосувань.

Використання передових технологій, таких як згорткові нейронні мережі, відіграло значну роль у підвищенні продуктивності систем розпізнавання візуальних зображень. Ці технології уможливлюють точне виявлення та вилучення рис обличчя, що дозволяє точно ідентифікувати та інтерпретувати вираз обличчя. Інтеграція методів обробки зображень та CNN дало змогу системі класифікувати емоційні стани високим рівнем точності.

Запропонована інформаційна система, яка включає в себе різні програмні компоненти, показала великий потенціал як ефективний інструмент для аналізу емоцій людини в реальному часі. Модулі системи, що включають функції попередньої обробки зображень та відео, сприяють точному розпізнаванню емоційних проявів. Експериментальне тестування та результати оцінки системи продемонстрували гарну загальну точність у 80,52%, що підкреслює її надійність та високий рівень точності у прогнозуванні мімічних проявів на основі зображень та відео. Для розробки інформаційної системи було використано мову програмування Python, платформу TensorFlow, бібліотеки python-telegram bot та OpenCV.

В цілому, отримані результати підкреслюють безперервний розвиток інформаційних технологій та їх трансформаційний вплив на системи розпізнавання візуальних образів. Інтеграція згорткових нейронних мереж та методів обробки зображень та відео забезпечує комплексний підхід до точної класифікації емоційних станів на основі виразу обличчя.

Перелік посилань

1. Object detection in 20 years: A survey / Z. Zou et al. *Proceedings of the IEEE*. 2023. Vol. 111, No. 3. Pp. 1-20.

2. Бармак О.В., Радюк П.М. Інформаційна технологія візуального аналізу рентгенівських зображень для інтерпретації результатів діагностування пневмонії. *Вісник Хмельницького національного університету. Технічні науки*. 2021. № 295(2). С. 52-55.

3. Pavlova O., Radiuk P., Kravchuk S., Kulbachnyi V. Information system for public places and institutions visualization with opportunities of inclusive access and optimal routing. *Computer systems and information technologies*. 2022. Vol. 1, No 6. Pp. 62-68.

4. GrandViewResearch: Image Recognition Market Size. *Grand View Research, Inc*. 2019. URL: https://www.grandviewresearch.com/industry-analysis/image-recognition-market

5. Dalvi C., Rathod M., Patil S., Gite S., Kotecha K. A survey of AI-based facial emotion recognition: Features, ML and DL techniques, age-wise datasets and future directions. *IEEE Access*. 2021. Vol. 9. Pp. 165806-165840.

6. Ekman P., Friesen W. V., Hager J. C. The facial action coding system: The manual. Salt Lake City : UT Research Nexus eBook, 2002. 527 p.

7. What is EMFACS, and how can I obtain it? *Paul Ekman Group LLC*. 2023. URL: https://www.paulekman.com/helpie_faq/what-is-emfacs-and-how-can-i-obtain-it-2/

8. Universal Emotions. *Paul Ekman Group LLC*. 2023. URL: https://www.paulekman.com/universal-emotions/

9. Method of facial geometric feature representation for information security systems. *The 3rd International Workshop on Intelligent Information Technologies & Systems of Information Security (IntelITSIS-2022)* : CEUR-Workshop Proceedings. Vol. 3156. (Khmelnytskyi, Ukraine, 23-25 March 2022). CEUR-WS.org, Aachen, 2022. Pp. 319-328.

10. Barmak O., Radiuk P. Web-based information technology for classifying and interpreting early pneumonia based on fine-tuned convolutional neural network. *Computer systems and information technologies*. 2021. Vol. 3, No 1. Pp. 12-18.

11. Radiuk P., Pavlova O., Avsiyevych V., Kovalenko V. Convolutional neural network for parking slots detection. *The 3rd International Workshop on Intelligent Information Technologies & Systems of Information Security (IntelITSIS-2022)* : CEUR-Workshop Proceedings. Vol. 3156. (Khmelnytskyi, Ukraine, 23-25 March 2022). CEUR-WS.org, Aachen, 2022. Pp. 284-293.

12. A deep facial recognition system using computational intelligent algorithms / D. Salama AbdELminaam et al. *PLOS ONE*. 2020. Vol. 15, No. 12. P. e0242269.

13. Decision Tree *GeeksforGeeks.org*. 2023. URL: https://www.geeksforgeeks.org/decision-tree/

14. Support Vector Machine — Introduction to Machine Learning Algorithms | *Published in Towards Data Science by Rohith Gandhi*. 2018. URL: https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47

15. Weighted K-NN. *GeeksforGeeks.org*. 2021. URL: https://www.geeksforgeeks.org/weighted-k-nn/?ref=gcse

16. Tutorial 1: Image Filtering. *Stanford Artificial Intelligence Laboratory*. 2019. URL: https://ai.stanford.edu/~syyeung/cvweb/tutorial1.html

17. Gollapudi S. OpenCV with Python. Learn computer vision using OpenCV: With deep learning CNNs and RNNs. Berkeley, CA : Apress, 2019. Pp. 31-50.

18. OpenCV Face Recognition - PyImageSearch. *PyImageSearch*. 2022. URL: https://pyimagesearch.com/2018/09/24/opencv-face-recognition/

19. FERV39k: A large-scale multi-scene dataset for facial expression recognition in videos / Y. Wang et al. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR-2022)* : Proceedings. (New Orleans, LA, USA, 18-24 June 2022). IEEE, Inc., 2022. Pp. 20890-20899.

20. Wingenbach T.S.H., Ashwin C., Brosnan M. Validation of the Amsterdam Dynamic Facial Expression Set – Bath intensity variations (ADFES-BIV): A set of

videos expressing low, intermediate, and high intensity emotions. *PLOS ONE*. 2016. Vol. 11, No. 1. P. e0147112.

21. Analysis of deep learning methods in adaptation to the small data problem solving / I. Krak et al. *Intellectual Systems of Decision Making and Problem of Computational Intelligence (ISDMCI-2022)* : Lecture Notes in Data Engineering, Computational Intelligence, and Decision Making. 2022. Vol. 149. (Zalizniy Port, Ukraine, 23-27 May 2022). Springer, Cham. Pp. 333-352.

22. Hu Z. A CNN-based facial expression recognition system. *Highlights in Science, Engineering and Technology*. 2023. Vol. 39. Pp. 496-507.

23. A study of LBPH, Eigenface, Fisherface and Haar-like features for face recognition using OpenCV / A.M. Jagtap et al. *2019 International Conference on Intelligent Sustainable Systems (ICISS-2019)* : Proceedings. (Palladam, Tamilnadu, India, 21-22 February 2019). IEEE, Inc., 2019. Pp. 219-224.

24. Радюк П.М. Підхід до прискорення навчання згорткової нейронної мережі за рахунок налаштування гіперпараметрів навчання. *Комп'ютерні системи та інформаційні технології*. 2020. № 2(2). С. 32-35.

25. Telegram Bot API. *Telegram APIs*. URL: https://core.telegram.org/bots/api

26. python-telegram-bot v20.3. *python-telegram-bot.org*. URL: https://python-telegram-bot.org

27. Python 3.8 | Python is a programming language that lets you work quickly and integrate systems more effectively. *Python.org*. URL: https://www.python.org/

28. Python™ Package Index. Requests 2.31.0. PyPI. URL: https://pypi.org/project/requests/

29. PyTorch: An imperative style, high-performance deep learning library / A. Paszke et al. *Advances in Neural Information Processing Systems* : Proceedings. (Vancouver, BC, Canada, December 8-14, 2019). 2019. Vol. 32. Curran Associates, Inc., 2019. Pp. 8024-8035.

30. torchvision – Torchvision 0.15 documentation. *PyTorch.org*. URL: https://pytorch.org/vision/stable/index.html

ДОДАТКИ

Додаток А

Характеристика основних алгоритмів, які використовуються для розпізнавання мімічних проявів

Таблиця А.1 – Характеристика основних алгоритмів, які використовуються для розпізнавання мімічних проявів

Алгоритм	Опис	Переваги	Недоліки
Дерево рішень (DT)	Деревоподібна модель, яка рекурсивно розбиває набір даних на найбільш значущі атрибути, приймаючи рішення на основі результатів кожного розбиття.	1. Легко розуміти та інтерпретувати. 2. Здатний обробляти пропущені значення та зашумлені дані. 3. Потребує мінімальної попередньої обробки даних.	1. Схильний до переналаштування. 2. Може погано працювати з безперервними даними. 3. Погана продуктивність на незбалансованих даних.
Метод опорних векторів (SVM)	Модель машинного навчання, яка використовує гіперплощину для розділення класів, максимізуючи відстань між гіперплощиною та найближчими точками в кожному класі.	1. Хороша продуктивність на даних високої розмірності. 2. Добре працює з нелінійно розділеними даними завдяки функціям ядра. 3. Стійкий до перенавчання.	1. Обчислювально дорогий, особливо для великих наборів даних. 2. Не підходить для багатокласової класифікації без модифікацій. 3. Чутливий до вибору функції ядра.
Метод k-найближчих сусідів (KNN)	Простий алгоритм, який класифікує точку даних за класом її найближчих сусідів у навчальній вибірці.	1. Непараметричний, тобто не робить припущень щодо розподілу даних, які лежать в основі методу. 2. Простий у розумінні та застосуванні.	1. Повільний при виконанні, особливо для великих наборів даних. 2. Продуктивність може погіршитися, якщо кількість об’єктів велика.

		3. Стійкий до зашумлених даних.	3. Чутливий до вибору метрики відстані.
Згорткова нейронна мережа (CNN)	Архітектура глибокого навчання, яка використовує згорткові шари для вилучення ознак із зображень та об’єднання шарів для зменшення розмірності ознак.	1. Найсучасніші результати в задачах класифікації зображень. 2. Здатність до навчання ієрархічних представлень ознак. 3. Стійкість до змін орієнтації та освітлення зображення.	1. Потребує великих обсягів навчальних даних та обчислювальних ресурсів. 2. Може бути піддатись перенавченню, якщо не впорядкований належним чином. 3. Інтерпретованість може бути складною.

Додаток Б

Архітектура використаної згорткової нейронної мережі

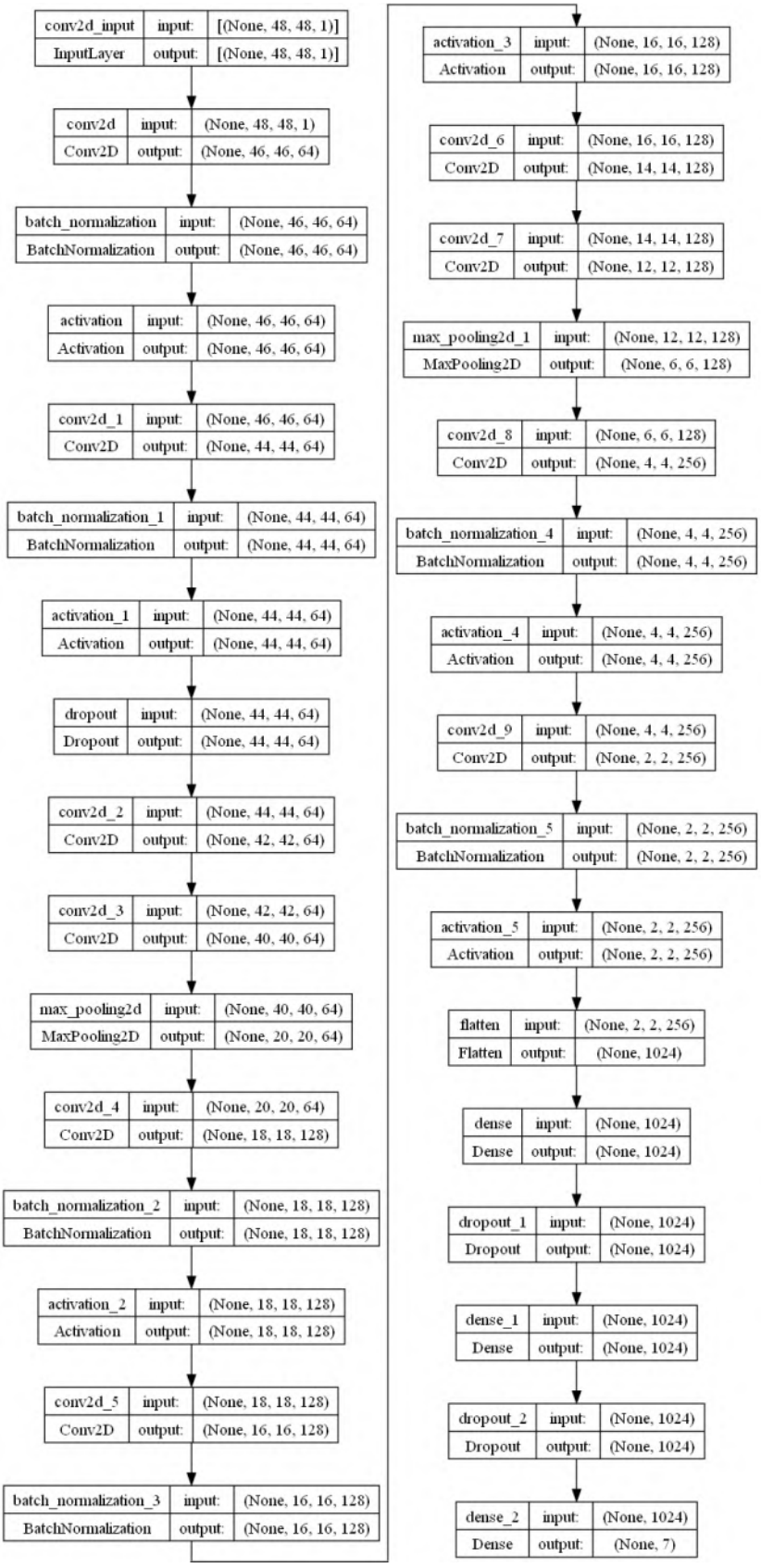


Рисунок Б.1 – Архітектура використаної згорткової нейронної мережі

Додаток В

Лістинг програмного коду

Модуль оброблення та розпізнавання зображень:

```
async def rec_image(update: Update, context: ContextTypes.DEFAULT_TYPE):
    file_id = update.message.photo[-1].file_id
    photo1 = await context.bot.get_file(file_id)
    file_name = 'img_{}.png'.format(update.message.message_id)
    path_1 = 'in_image/' + file_name
    path_2 = 'out_image/' + file_name
    await photo1.download_to_drive(path_1)

    img = cv2.imread(path_1)
    gray_img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
    faces_detected = classifier.detectMultiScale(gray_img, 1.18, 5)

    img = process(faces_detected, img, gray_img)

    cv2.imwrite(path_2,img)

def process(faces_detected, img, gray_img):
    for (x, y, w, h) in faces_detected:
        cv2.rectangle(img, (x, y), (x + w, y + h), (0, 255, 0), 2)
        roi_gray = gray_img[y:y + w, x:x + h]
        roi_gray = cv2.resize(roi_gray, (48, 48))
        img_pixels = image.img_to_array(roi_gray)
        img_pixels = np.expand_dims(img_pixels, axis=0)
        img_pixels /= 255.0

        predictions = model.predict(img_pixels)
        max_index = int(np.argmax(predictions))
          emotions = ['neutral', 'happiness', 'surprise', 'sadness', 'anger',
'disgust', 'fear']
        predicted_emotion = emotions[max_index]

                    cv2.putText(img,    predicted_emotion,    (int(x),    int(y)),
cv2.FONT_HERSHEY_SIMPLEX, 0.75, (255, 255, 255), 2)
    return img
```

Модуль оброблення та розпізнавання відео:

```
async def rec_video(update: Update, context: ContextTypes.DEFAULT_TYPE):

    if update.message.video:
        file_id = update.message.video.file_id
    elif update.message.video_note:
        file_id = update.message.video_note.file_id

    video1 = await context.bot.get_file(file_id)

    video_path = video1.file_path
    video_extension = video_path.split('.')[-1]
    file_name = 'video_{}.{}'.format(update.message.message_id,video_extension)
```

```
    file_name_out =
'video_{}.mp4'.format(update.message.message_id,video_extension)

    path_1 = 'in_video/' + file_name
    path_2 = 'out_video/' + file_name_out
    await video1.download_to_drive(path_1)

    cap = cv2.VideoCapture(path_1)

    fps = cap.get(cv2.CAP_PROP_FPS)

    width = int(cap.get(cv2.CAP_PROP_FRAME_WIDTH))
    height = int(cap.get(cv2.CAP_PROP_FRAME_HEIGHT))

    count = 0
    while True:
        ret, img = cap.read()
        if not ret:
            break

        gray_img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
        faces_detected = classifier.detectMultiScale(gray_img, 1.2, 6)

        img = process(faces_detected, img, gray_img)

        cv2.imwrite('temp_video\\frame%d.jpg' % count,img)
        count += 1

        if cv2.waitKey(1) & 0xFF == ord('q'):
            break

    cap.release()
    await context.bot.send_message(chat_id=update.effective_chat.id,
reply_to_message_id=update.message.message_id, text='Обличчя розпізнано,
завантажую розпізнане відео')

    frame_dir = 'temp_video'
    frame_paths = [os.path.join(frame_dir, f) for f in os.listdir(frame_dir)]
    frame_paths = sorted(frame_paths, key=lambda x:
int(x.split('frame')[1].split('.jpg')[0]))

    fourcc = cv2.VideoWriter_fourcc(*'mp4v')

    out = cv2.VideoWriter(path_2, fourcc, fps, (width, height))

    for frame_path in frame_paths:

        frame = cv2.imread(frame_path)
        frame = cv2.resize(frame, (width, height))
        out.write(frame)

    out.release()

    for filename in os.listdir(frame_dir):
        file_path = os.path.join(frame_dir, filename)
        try:
            if os.path.isfile(file_path):
```

```
                os.unlink(file_path)
            except Exception as e:
                print(f"Failed to delete {file_path}. Reason: {e}")

    if update.message.video:
        await context.bot.send_video(chat_id=update.effective_chat.id,
reply_to_message_id=update.message.message_id, video=path_2)
    elif update.message.video_note:
        await context.bot.send_video_note(chat_id=update.effective_chat.id,
reply_to_message_id=update.message.message_id, video_note=path_2)
```

Модуль реалізації способу розпізнавання мімічних проявів емоцій:

```
def preprocess_data():
    data = pd.read_csv('fer2013/fer2013.csv')
    labels = pd.read_csv('fer2013/fer2013new.csv')

    orig_class_names = ['neutral', 'happiness', 'surprise', 'sadness', 'anger',
'disgust', 'fear', 'contempt',
                        'unknown', 'NF']

    n_samples = len(data)
    w = 48
    h = 48

    y = np.array(labels[orig_class_names])
    X = np.zeros((n_samples, w, h, 1))
    for i in range(n_samples):
        X[i] = np.fromstring(data['pixels'][i], dtype=int, sep=' ').reshape((h,
w, 1))

    return X, y


def clean_data_and_normalize(X, y):
    orig_class_names = ['neutral', 'happiness', 'surprise', 'sadness', 'anger',
'disgust', 'fear', 'contempt',
                        'unknown', 'NF']
    y_mask = y.argmax(axis=-1)
    mask = y_mask < orig_class_names.index('unknown')
    X = X[mask]
    y = y[mask]

    y = y[:, :-2] * 0.1

    y[:, 0] += y[:, 7]
    y = y[:, :7]

    X = X / 255.0

    return X, y
```

```
def split_data(X, y):
    test_size = ceil(len(X) * 0.1)

    x_train, x_test, y_train, y_test = model_selection.train_test_split(X, y,
test_size=test_size, random_state=42)
    x_train, x_val, y_train, y_val = model_selection.train_test_split(x_train,
y_train, test_size=test_size,
                                                                      random_sta
te=42)
    return x_train, y_train, x_val, y_val, x_test, y_test


def data_augmentation(x_train):
    shift = 0.1
    datagen = ImageDataGenerator(
        rotation_range=20,
        horizontal_flip=True,
        height_shift_range=shift,
        width_shift_range=shift)
    datagen.fit(x_train)
    return datagen


def show_augmented_images(datagen, x_train, y_train):
    it = datagen.flow(x_train, y_train, batch_size=1)
    plt.figure(figsize=(10, 7))
    for i in range(25):
        plt.subplot(5, 5, i + 1)
        plt.xticks([])
        plt.yticks([])
        plt.grid(False)
        plt.imshow(it.next()[0][0], cmap='gray')
        # plt.xlabel(class_names[y_train[i]])
    plt.show()


def define_model(input_shape=(48, 48, 1), classes=7):
    num_features = 64

    model = Sequential()

    # 1st stage
    model.add(Conv2D(num_features, kernel_size=(3, 3), input_shape=input_shape))
    model.add(BatchNormalization())
    model.add(Activation(activation='relu'))
    model.add(Conv2D(num_features, kernel_size=(3, 3)))
    model.add(BatchNormalization())
    model.add(Activation(activation='relu'))
    model.add(Dropout(0.5))

    # 2nd stage
    model.add(Conv2D(num_features, (3, 3), activation='relu'))
```

```
    model.add(Conv2D(num_features, (3, 3), activation='relu'))
    model.add(MaxPooling2D(pool_size=(2, 2), strides=(2, 2)))

    # 3rd stage
    model.add(Conv2D(2 * num_features, kernel_size=(3, 3)))
    model.add(BatchNormalization())
    model.add(Activation(activation='relu'))
    model.add(Conv2D(2 * num_features, kernel_size=(3, 3)))
    model.add(BatchNormalization())
    model.add(Activation(activation='relu'))

    # 4th stage
    model.add(Conv2D(2 * num_features, (3, 3), activation='relu'))
    model.add(Conv2D(2 * num_features, (3, 3), activation='relu'))
    model.add(MaxPooling2D(pool_size=(2, 2), strides=(2, 2)))

    # 5th stage
    model.add(Conv2D(4 * num_features, kernel_size=(3, 3)))
    model.add(BatchNormalization())
    model.add(Activation(activation='relu'))
    model.add(Conv2D(4 * num_features, kernel_size=(3, 3)))
    model.add(BatchNormalization())
    model.add(Activation(activation='relu'))

    model.add(Flatten())

    # Fully connected neural networks
    model.add(Dense(1024, activation='relu'))
    model.add(Dropout(0.2))
    model.add(Dense(1024, activation='relu'))
    model.add(Dropout(0.2))

    model.add(Dense(classes, activation='softmax'))

    return model


def save_model_and_weights(model, test_acc):
    # Serialize and save model to JSON
    test_acc = int(test_acc * 10000)
    model_json = model.to_json()
    with open('Saved-Models\\model' + str(test_acc) + '.json', 'w') as
json_file:
        json_file.write(model_json)
    # Serialize and save weights to JSON
    model.save_weights('Saved-Models\\model' + str(test_acc) + '.h5')
    print('Model and weights are saved in separate files.')
```

```
def run_model():
    fer_classes = ['neutral', 'happiness', 'surprise', 'sadness', 'anger',
'disgust', 'fear']

    X, y = preprocess_data()
    X, y = clean_data_and_normalize(X, y)
    x_train, y_train, x_val, y_val, x_test, y_test = split_data(X, y)
    datagen = data_augmentation(x_train)

    epochs = 100
    batch_size = 64

    print("X_train shape: " + str(x_train.shape))
    print("Y_train shape: " + str(y_train.shape))
    print("X_test shape: " + str(x_test.shape))
    print("Y_test shape: " + str(y_test.shape))
    print("X_val shape: " + str(x_val.shape))
    print("Y_val shape: " + str(y_val.shape))

    # Training model from scratch
    model = define_model(input_shape=x_train[0].shape, classes=len(fer_classes))
    model.summary()
    model.compile(optimizer=Adam(learning_rate=0.0001),
loss='binary_crossentropy', metrics=['accuracy'])
    history = model.fit(datagen.flow(x_train, y_train, batch_size=batch_size),
epochs=epochs,
                        steps_per_epoch=len(x_train) // batch_size,
                        validation_data=(x_val, y_val), verbose=2)
    test_loss, test_acc = model.evaluate(x_test, y_test, batch_size=batch_size)

    plot_acc_loss(history)
    save_model_and_weights(model, test_acc)
run_model()
```

Додаток Г

Демонстрація розпізнавання всіх емоцій на зображенні

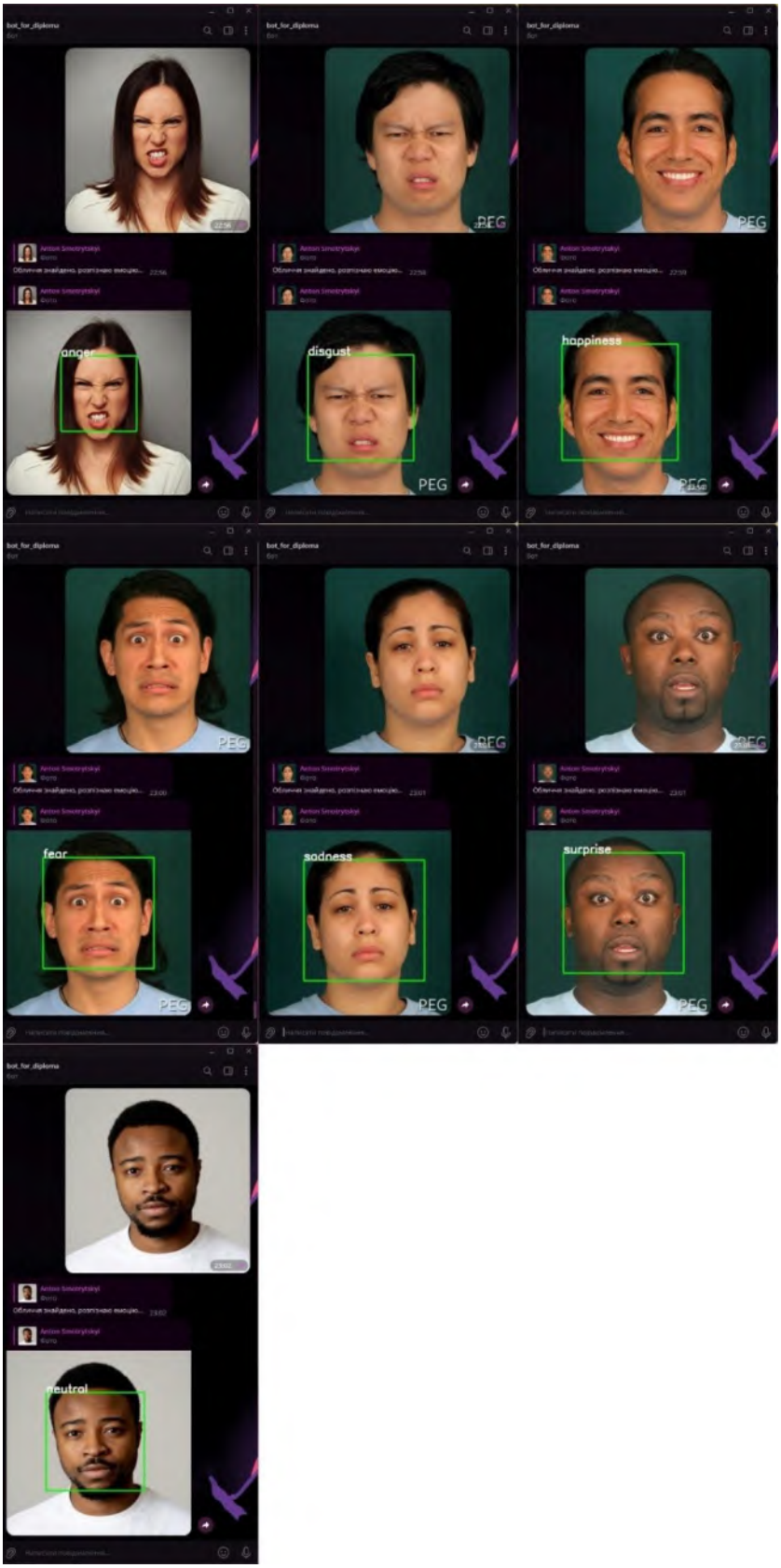


Рисунок Г.1 – Розпізнавання всіх емоцій на зображенні

Додаток Ґ

Презентаційний матеріал

КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА

РОЗПІЗНАВАННЯ МІМІЧНИХ ПРОЯВІВ ЕМОЦІЙ ЗАСОБАМИ ШТУЧНОГО ІНТЕЛЕКТУ

Виконав:
студент 4 курсу, групи КН-19-2
Смотрицький Антон Валентинович

Керівник:
старший викладач кафедри КН
Радюк Павло Михайлович

2

Актуальність

Розпізнавання виразу обличчя та емоцій має вирішальне значення для різних аспектів людської взаємодії та взаєморозуміння. Ось чому це важливо:

- Емоційне благополуччя: Це допомагає нам підтримувати людей, розуміючи їхній емоційний стан.
- Психологічні дослідження: Дозволяє вивчати емоційні реакції та психічне здоров'я.
- Маркетинг і реклама: Покращує маркетингові дослідження, аналізуючи реакцію споживачів.
- Взаємодія між людиною та комп'ютером: Покращує користувацький досвід в технологічних інтерфейсах.
- Персоналізована охорона здоров'я: Допомагає діагностувати та лікувати психічні розлади, покращуючи догляд за пацієнтами.

3

Мета і задачі роботи

- **Об'єкт дослідження** – процес ідентифікації змін емоційного стану людини за мімічними проявами з використанням систем фото- та відеоспостереження.
- **Предмет дослідження** – методи та підходи до ідентифікації змін емоційного стану людини за мімічними проявами за цифровими фото- та відеоматеріалами.
- **Мета кваліфікаційної роботи бакалавра** – підвищення точності ідентифікації змін емоційного стану людини за мімічними проявами за цифровими фото- та відеоматеріалами.

4

Мета і задачі роботи

Завдання кваліфікаційної роботи бакалавра. Для досягнення поставленої мети визначено наступні завдання:

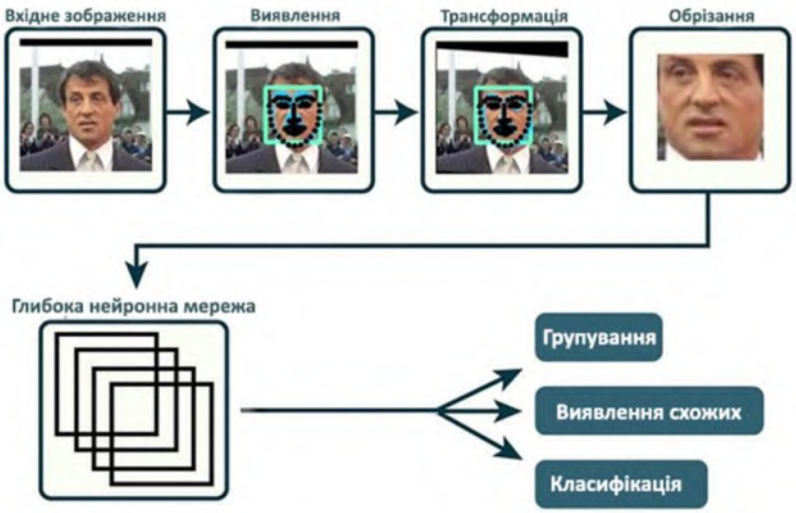
1. Провести аналіз методів та підходів до ідентифікації змін емоційного стану людини за мімічними проявами та обрати найкращий.
2. Застосувати обраний метод ідентифікації змін емоційного стану людини за мімічними проявами до розв'язання задачі автоматизованого розпізнавання мімічних проявів емоцій.
3. Реалізувати обраний метод у вигляді інформаційної системи автоматизованого розпізнавання мімічних проявів емоцій.
4. Провести експериментальне тестування інформаційної системи за еталонними наборами даних.

Розроблена програмна реалізація способу автоматизованого розпізнавання мімічних проявів емоцій в вигляді інформаційної системи з використанням штучного інтелекту має виконувати наступні основні групи функцій:

- Реалізація архітектури згорткової нейронної мережі та навчання на обраному наборі даних;
- Попередня обробка вхідних даних (зображення та відео матеріалів);
- Розпізнавання та прогнозування мімічних проявів на вхідних даних. Відправка результату користувачу.

5

Спосіб розпізнавання мімічних проявів емоцій засобами штучного інтелекту



6

Вхідні дані:
– Користувацьке зображення або відео;

Крок 1:
Аналіз та попередня обробка вхідного файлу:
1. Завантаження вхідного файлу як набору даних за допомогою OpenCV
2. Перетворення кольорових каналів до градацій сірого.
3. Виявлення облич за допомогою каскадного класифікатора Хаара.
4. Виділення області, що містить виявлене обличчя та кадрування області до розміру 48x48 пікселів.
5. Конвертація зображення до масиву та його нормалізація, що дозволить передати дані до нейронної мережі

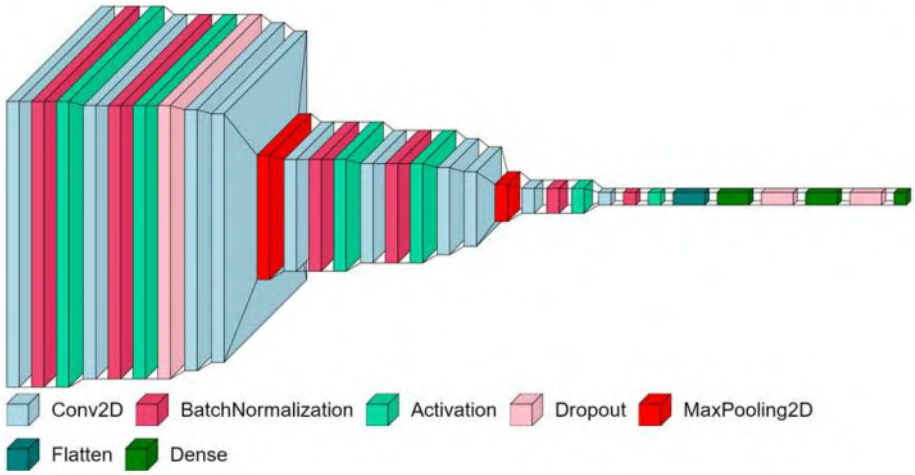
Крок 2:
Використання навченої моделі згорткової нейронної мережі для розпізнавання мімічних проявів емоцій та їх класифікації:
1. Масив пікселів пропускається через навчену модель для отримання індексу найвищої передбачуваної емоції
2. Класифікація емоцій за отриманим індексом

Вихідні дані:
– Відправка результату із розпізнаною емоцією користувачу.

Схема способу розпізнавання мімічних проявів емоцій засобами штучного інтелекту

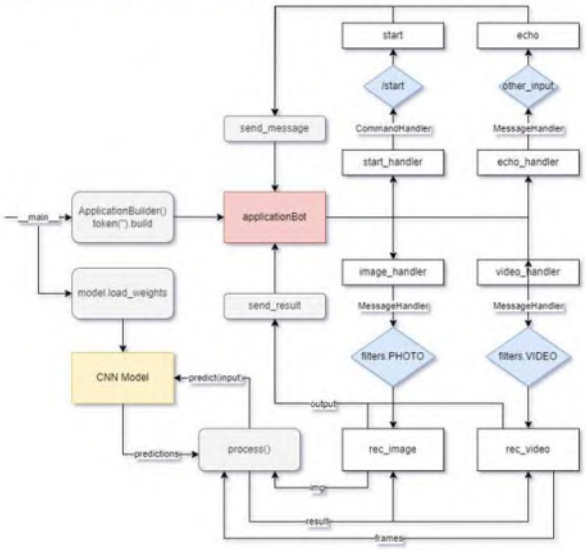
7

Архітектура нейронної мережі, яка, виконуватиме розпізнавання та класифікацію емоції людини



8

Діаграма взаємодії складових програмної реалізації розпізнавання мімічних проявів емоцій засобами штучного інтелекту



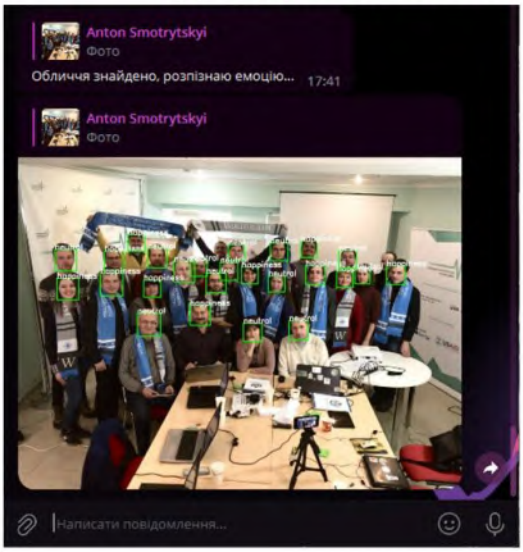
9

Програмна реалізація способу розпізнавання мімічних проявів емоцій засобами штучного інтелекту

Розпізнавання мімічних проявів на відео-повідомленні у телеграм

10

Програмна реалізація методу автоматизованого підбору асоціативної відповіді за семантичною подібністю



Розпізнавання багатьох облич на груповому фото

Висновки

У рамках виконання кваліфікаційної роботи бакалавра було виконано **розробку й апробацію способу** розпізнавання мімічних проявів емоцій засобами штучного інтелекту. Зокрема, було проведено аналіз предметної області й досліджено сучасні підходи до розпізнавання мімічних проявів на зображеннях та відео, розглянуто існуючі програмні реалізації за цим напрямком.

Розроблена **програмна реалізація** способу Розпізнавання мімічних проявів емоцій засобами штучного інтелекту з використанням мови програмування Python виконує наступні основні функції:

- Завантаження та попередня обробка вхідного файлу.
- Розпізнавання мімічних проявів емоцій на зображенні та відео, використовуючи попередньо навчену модель.
- Відправка результату користувачу.

У якості засобів розробки було використано платформу TensorFlow, мову програмування Python, редактор програмного коду Visual Studio та Jupyter Notebook.

Sat May 27 17:39:04 EEST 2023, Петровський Сергій Степанович, Хмельницький національний університет, ХНУ

Anti-Plagiarism v-15.257

Максимальне співпадіння з одним документом 2.0%

Словники перевірки: en_US, ru_RU, ua_UA. **Помилок в документах: 7%**

	Документ		Сумарний збіг по Базі Даних	
ID: 114155 Назва: КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА Додано в БД: 2023-05-27 Автора: А.В. Смотрицький Керівники: П.М. Радюк Консультанти: Опоненти:	Символи	Лексеми	Символи	Лексеми
	68898	1060	1835 (3%)	31 (3%)

Джерело плагіату

ID	Опис	Наявність плагіату в документі	
		Символи	Лексеми

UNICHECK

Ім'я користувача:
Кафедра КН

Дата перевірки:
27.05.2023 18:32:35 EEST

Дата звіту:
27.05.2023 18:43:48 EEST

ID перевірки:
1015286933

Тип перевірки:
Doc vs Internet + Library

ID користувача:
100005671

Назва документа: **КН-19-2 Смотрицький**

Кількість сторінок: 63 Кількість слів: 10881 Кількість символів: 82324 Розмір файлу: 2.22 MB ID файлу: 1014959355

Виявлено модифікації тексту (можуть впливати на відсоток схожості)

4.71%
Схожість

Найбільша схожість: 2.35% з джерелом з Бібліотеки (ID файлу: 1014959353)

3.8% Джерела з Інтернету	391	Сторінка 65
2.97% Джерела з Бібліотеки	85	Сторінка 67

0.31% Цитат

Цитати	2	Сторінка 68

Не знайдено жодних посилань

0%
Вилучень

Немає вилучених джерел

Модифікації

Виявлено модифікації тексту. Детальна інформація доступна в онлайн-звіті.

Підозріле форматування	10 сторінок

РІШЕННЯ ЕКСПЕРНОЇ КОМІСІЇ КАФЕДРИ КОМП'ЮТЕРНИХ НАУК
ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ ДО ЗАХИСТУ

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

Назва: *Розпізнавання мімічних проявів емоцій засобами штучного інтелекту*

Автор: *студент групи КН-19-2 Смотрицький Антон Валентинович*

Спеціальність: *122 – Комп'ютерні науки*

Освітня програма: *освітньо-професійна*

Науковий керівник: *док. філ., ст. викл. Радюк П.М.*

Після аналізу звіту подібності зроблено такий висновок:

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

Підтвердження:

Запозичення, виявлені в роботі Смотрицького А.В., не є плагіатом, оскільки: запозичення розміщені в розділі огляду існуючих підходів, не описують безпосередньо авторську роботу і не стосуються її результатів; усі запозичення фрагментарні; до запозичень входять фрагменти програмного коду, що не мають авторства і містять поширені конструкції; поміж запозичень знаходяться загальновідомі терміни та скорочення.

Обсяг запозичень, визначений системами виявлення збігів/ ідентичності/схожості, складає:

- за системою Anti-Plagiarism: 2.0%;

- за системою Unicheck: 4.71 %.

Отже, запозичення є допустимими та відносяться до описаних вище і адресуються до першоджерел, що, з урахуванням наведених обґрунтувань, свідчить на користь кваліфікаційної роботи.

Керівник роботи _____ *Павло РАДЮК*

Гарант ОП _____ *Олександр МАЗУРЕЦЬ*

Завідувач кафедри КН _____ *Олександр БАРМАК*

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МОН УКРАЇНИ

Кафедра комп'ютерних наук

ВІДГУК НАУКОВОГО КЕРІВНИКА

на кваліфікаційну роботу бакалавра

студента *гр. КН-19-2 Смотрицького Антона Валентиновича*

за темою *Розпізнавання мімічних проявів емоцій засобами штучного інтелекту*

1. Актуальність теми

Людські емоції відіграють ключову роль у міжособистісному спілкуванні. Проте автоматизоване розпізнавання та інтерпретування емоцій видається складною задачею. Непорозуміння через емоційну неоднозначність можуть призвести до напружених стосунків, неефективної комунікації та втрачених можливостей для співпереживання. Відповідно виникає потреба в розробленні нового рішення для автоматизації процесу розпізнавання мімічних проявів емоцій засобами штучного інтелекту.

2. Відповідність роботи предметній області Стандарту спеціальності 122 Комп'ютерні науки

Відповідно до стандарту бакалавра вищої освіти України спеціальності 122 – Комп'ютерні науки, описом предметної галузі, об'єктом та предметом вивчення є математичні, інформаційні та імітаційні моделі реальних явищ, об'єктів, систем і процесів та методи й технології отримання, зберігання, обробки, передачі та використання інформації. Метою поданої кваліфікаційної роботи бакалавра є підвищення точності ідентифікації змін емоційного стану людини за мімічними проявами за цифровими фото- та відеоматеріалами. Мету роботи досягнуто внаслідок використання методів, способів та алгоритмів розв'язання теоретичних і прикладних задач, що виникають у процесі проєктування та розроблення інформаційних технологій. Отже, результати виконання кваліфікаційної роботи відповідають стандарту бакалавра спеціальності 122 – Комп'ютерні науки.

3. Професійні та особистісні якості бакалавра

Під час виконання кваліфікаційної роботи бакалавра студент Смотрицький Антон Валентинович проявив себе кваліфікованим фахівцем та дисциплінованим студентом, вчасно виконуючи поставлені перед ним завдання. Як у процесі написання пояснювальної записки, так і під час розроблення прикладного програмного забезпечення студент Смотрицький успішно засвоїв компетентності та результати навчання. Загалом Смотрицький А.В. опанував професійні навички та компетентності, що

відповідають виконанню освітньо-професійної програми рівня вищої освіти «Бакалавр» за спеціальністю 122 – Комп'ютерні науки.

4. Ступінь самостійності під час виконання кваліфікаційної роботи

Одержані в роботі результати та обґрунтована їхня практична значущість слугують наслідком особистої діяльності студента, який самостійно виконував усі поставлені перед ним завдання.

5. Ступінь оволодіння методами дослідження

У процесі аналізу способу розпізнавання мімічних проявів емоцій засобами штучного інтелекту та проєктування і розроблення на його основі інформаційної системи студент Смотрицький А.В. продемонстрував відмінний рівень компетентностей та володіння необхідними інструментами й обладнанням, методами, методиками та технологіями галузі інформаційних технологій.

6. Повнота та якість розкриття теми роботи

Тема роботи повністю обґрунтована й розкрита, актуальність предметної галузі та відомі дослідження щодо обраної тематики проаналізовано достатньо. Усі поставлені завдання в роботі виконані. Розроблене програмне забезпечення для валідації та верифікації розробленої інформаційної системи відповідає технічним вимогам спеціальності 122 – Комп'ютерні науки.

7. Логічність, послідовність, аргументованість, літературна грамотність викладення матеріалу

Матеріал кваліфікаційної роботи Смотрицького А.В. подано логічно, послідовно, аргументовано та є таким, що відповідає поставленій меті. Мова і стиль викладення роботи відповідають стандартам, що забезпечує доступність сприймання матеріалу і відповідає вимогам до сучасних кваліфікаційних робіт.

8. Можливість практичного застосування кваліфікаційної роботи бакалавра, окремих її частин

Розроблена в роботі інформаційна система може бути використана під час проведення аналітики маркетингових кампаній для розуміння реакцій та настроїв споживачів і, як наслідок, підвищення ефективності рекламних кампаній.

9. Висновок про можливість допуску кваліфікаційної роботи бакалавра до захисту, на яку оцінку заслуговує робота

*З огляду на рівень виконання та забезпечення всіх необхідних вимог, вважаю, що кваліфікаційна робота Смотрицького Антона Валентиновича може бути допущена до захисту. Рекомендована оцінка – **«відмінно»**.*

Керівник _____ *док. філ., ст. викл. каф. КН Павло РАДЮК*

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МОН УКРАЇНИ

Кафедра комп'ютерних наук

РЕЦЕНЗІЯ

на кваліфікаційну роботу бакалавра

студента *гр. КН-19-2 Смотрицького Антона Валентиновича*
за темою: *Розпізнавання мімічних проявів емоцій засобами штучного інтелекту*

1. Актуальність обраної теми

Розпізнавання емоційного виразу обличчя за допомогою штучного інтелекту є дуже актуальною темою в сучасному технологічно розвиненому суспільстві. Використовуючи алгоритми штучного інтелекту та методи комп'ютерного зору, дослідники та розробники можуть точно аналізувати та інтерпретувати людські емоції на основі міміки. Ця технологія має величезний потенціал у різних галузях, таких як психологія, взаємодія людини з комп'ютером і навіть маркетингові дослідження. Розуміючи емоційний вираз обличчя, системи штучного інтелекту можуть допомогти поліпшити підтримку психічного здоров'я, розробити емпатичні користувацькі інтерфейси та покращити клієнтський досвід завдяки персоналізованій взаємодії.

2. Повнота розкриття мети та завдань роботи

У процесі виконання кваліфікаційної роботи бакалавра були ретельно розкриті мета та завдання. Автором проведено глибоке дослідження предметної області, визначено структуру системи та розроблено відповідне системі програмне забезпечення. В результаті виконаної роботи поставлена мета та завдання повністю розкриті.

3. Зміст кожного розділу роботи

У першому розділі кваліфікаційної роботи проведено дослідження сучасних способів та підходів до розпізнавання мімічних проявів емоцій. Другий розділ присвячений проєктуванню способу розпізнавання мімічних проявів емоцій засобами штучного інтелекту. У третьому розділі наведено та описано програмну реалізацію інформаційної системи розпізнавання мімічних проявів емоцій. Насамкінець оформлено висновки до виконаних у роботі завдань.

4. Оцінка розробленої інформаційної системи, її практична цінність

Розроблена система автоматизованого розпізнавання мімічних проявів емоцій та її програмна реалізація можуть бути широко застосовані для оцінювання емоційних станів та покращення точності аналізу психологічного стану людей за виразом обличчя. Запропонована автором система обчислює комплексну оцінку емоційного виразу обличчя людини, що дає змогу за фото- або відеоматеріалами, яке містить обличчя людей, оцінювати їхній емоційний стан.

5. Якість оформлення кваліфікаційної роботи бакалавра

Оформлена автором записка має чудову якість, демонструє логічний виклад матеріалу та надає сильні, переконливі аргументи. Написання роботи демонструє літературну компетентність та доречність.

6. Недоліки кваліфікаційної роботи бакалавра

Робота також містить кілька недоліків. У тексті трапляються скорочення, які не наведено у переліку скорочень. В деяких пунктах не вистачає висновкового речення.

7. Загальний висновок (допускається чи не допускається до захисту), та оцінка на яку заслуговує кваліфікаційна робота.

З огляду на рівень виконання та забезпечення всіх необхідних вимог, вважаю, що подана кваліфікаційна робота бакалавра може бути допущена до захисту. Рекомендована оцінка – «Відмінно»

Рецензент _____ доц., к.ф.-м.н. Наталія Ярешко