

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ
КАФЕДРА ШТУЧНОГО ІНТЕЛЕКТУ

КВАЛІФІКАЦІЙНА РОБОТА

на тему: «Інтелектуальна система аналізу та визначення
емоційного стану людини»

на здобуття освітнього ступеня бакалавра
зі спеціальності 122 Комп'ютерні науки
(код, найменування спеціальності)
освітньо-професійної програми Штучний інтелект
(назва)

Кваліфікаційна робота містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

(підпис)

В'ячеслав РОМАНОК
Ім'я, ПРИЗВИЩЕ здобувача

Виконав:
здобувач вищої освіти
група ШІД-41

В'ячеслав РОМАНОК

Керівник:
науковий ступінь,
вчене звання

Андрій БОНДАРЧУК
д.т.н., професор

Рецензент:
науковий ступінь,
вчене звання

Ім'я, ПРИЗВИЩЕ

Київ 2025

**ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ**
Навчально-науковий інститут інформаційних технологій

Кафедра 122 Комп'ютерні науки
Ступінь вищої освіти бакалавр
Спеціальність 122 Комп'ютерні науки
Освітньо-професійна програма Штучний інтелект

ЗАТВЕРДЖУЮ

Завідувач кафедрою Штучного інтелекту

_____ Ольга ЗІНЧЕНКО
«_____» _____ 2025 р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

_____ Романку В'ячеславу Володимировичу

(прізвище, ім'я, по батькові здобувача)

1. Тема кваліфікаційної роботи: Інтелектуальна система аналізу та визначення емоційного стану людини

керівник кваліфікаційної роботи Андрій БОНДАРЧУК доктор технічних наук, професор,

(Ім'я, ПРИЗВИЩЕ, науковий ступінь, вчене звання)

затверджені наказом Державного університету інформаційно-комунікаційних технологій від «24» лютого 2025р. № 56

2. Строк подання кваліфікаційної роботи «23» травня 2025р.

3. Вихідні дані до кваліфікаційної роботи: Наукові статті на тему розпізнавання об'єктів на зображенні, відкриті бази сегментованих зображень осіб з різним емоційним станом.

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

1. Огляд сучасного стану досліджень
2. Технічні аспекти систем виявлення емоційних станів
3. Реалізація системи розпізнавання емоцій

5. Перелік графічного матеріалу: *презентація*

1. Структура штучної нейронної мережі
2. Особливості розпізнавання емоцій людини
3. Бібліотеки для реалізації проєкту
4. Тестування проєкту

6. Дата видачі завдання «25» лютого 2025р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1	Збір даних	25.02-25.03.2025	Виконано
2	Обробка матеріалу	25.03-01.04.2025	Виконано
3	Розробка теоретичної частини	02.04-10.04.2025	Виконано
4	Аналіз та дослідження існуючих систем	11.04-28.04.2025	Виконано
5	Висновки за результатами аналізу	29.04-05.05.2025	Виконано
6	Оформлення пояснювальної записки	06.05-08.05.2025	Виконано
7	Розробка презентації	09.05-15.05.2025	Виконано
8	Передзахист	19.05-22.05.2025	Виконано
9	Здача в деканат	22.05-23.05.2025	Виконано

Здобувач вищої освіти

(підпис)

В'ячеслав РОМАНОК

(Ім'я, ПРІЗВИЩЕ)

Керівник кваліфікаційної роботи _____

(підпис)

Андрій БОНДАРЧУК

(Ім'я, ПРІЗВИЩЕ)

РЕФЕРАТ

Текстова частина кваліфікаційної роботи на здобуття освітнього ступеня бакалавра: стор. 57, табл. 2, рис. 12, джерел 20.

Мета роботи – створення інтелектуальної системи рекомендацій відео-контенту, що адаптується до емоційного стану користувача на основі аналізу відео із зображенням його обличчя. Проведено аналіз сучасних методів розпізнавання емоцій та рекомендаційних алгоритмів, визначено їх переваги й недоліки, а також реалізовано власну програмну модель з урахуванням виявлених критеріїв ефективності.

Об'єкт дослідження – процес автоматичного розпізнавання емоцій користувача на основі відео з використанням технологій глибинного навчання.

Предмет дослідження – програмна система, яка класифікує емоційний стан людини на основі міміки та формує персоналізовану добірку фільмів або серіалів відповідно до виявленої емоції.

Короткий зміст роботи: У роботі розглянуто актуальні підходи до розпізнавання емоцій за відео, проаналізовано існуючі реалізовані системи, включаючи комерційні та дослідницькі розробки. У межах дипломного проєкту створено Telegram-бот, який приймає відеозапис від користувача, обробляє його з використанням бібліотек OpenCV та Keras, виявляє обличчя та передає кадри у згорткову нейронну мережу, натреновану для класифікації шести базових емоцій. Результати класифікації використовуються для формування тематичної добірки фільмів або серіалів на основі внутрішньої бази даних. Проведено тестування точності, виявлено типові помилки класифікації та сформовано рекомендації щодо покращення.

КЛЮЧОВІ СЛОВА: РОЗПІЗНАВАННЯ ЕМОЦІЙ, ЗГОРТКОВА НЕЙРОННА МЕРЕЖА, TELEGRAM-БОТ, ГЛИБИННЕ НАВЧАННЯ, РЕКОМЕНДАЦІЙНА СИСТЕМА, КОМП'ЮТЕРНИЙ ЗІР.

ABSTRACT

The text part of the bachelor's qualification thesis: pages 57, tables 2, figures 12, sources 20.

The aim of the thesis is: to develop an intelligent video content recommendation system that adapts to the user's emotional state based on the analysis of a video containing their facial image. The study includes an analysis of modern emotion recognition methods and recommendation algorithms, identifying their advantages and drawbacks, and presents a custom-developed software model that considers the determined efficiency criteria.

Object of research: the process of automatic emotion recognition based on video using deep learning technologies.

Subject of research: a software system that classifies a person's emotional state based on facial expressions and generates a personalized selection of movies or TV shows according to the detected emotion.

Summary: This work examines current approaches to emotion recognition from video, analyzing existing systems, including both commercial and research solutions. As part of the thesis project, a Telegram bot was developed that accepts a video recording from the user, processes it using OpenCV and Keras libraries, detects the face, and passes the frames to a convolutional neural network trained to classify six basic emotions. The classification results are used to generate a thematic selection of movies or TV series from an internal database. The system's classification accuracy was tested, common errors were identified, and recommendations for improvement were formulated.

KEYWORDS: EMOTION RECOGNITION, CONVOLUTIONAL NEURAL NETWORK, TELEGRAM BOT, DEEP LEARNING, RECOMMENDER SYSTEM, COMPUTER VISION.

Зміст

ВСТУП.....	9
1 ОГЛЯД СУЧАСНОГО СТАНУ ДОСЛІДЖЕНЬ	11
1.1 Огляд реалізованих систем розпізнавання емоцій у світі	12
1.2 Вивчення сучасних методів розпізнавання емоцій за відео	16
1.3 Аналіз психологічних станів користувачів за допомогою відеоданих	17
1.4 Аналіз сучасних програмних рішень для розпізнавання емоцій.....	18
1.5 Загальні аспекти систем розпізнавання емоцій	20
2 ТЕХНІЧНІ АСПЕКТИ СИСТЕМ РОЗПІЗНАВАННЯ ЕМОЦІЙ.....	26
2.1 Загальні поняття штучних нейронних мереж для розпізнавання відео-контенту	26
2.2 Структура згорткових нейронних мереж для розпізнавання відео-контенту	28
2.3 Архітектура згорткової нейронної мережі проєктованого застосунку	30
2.4 Алгоритм навчання згорткової нейронної мережі для системи розпізнавання відеоконтенту	31
3. ОСНОВНІ ТЕХНОЛОГІЧНІ ЗАСАДИ СТВОРЕННЯ СИСТЕМИ	34
3.1 Обґрунтування вибору мови програмування та інструментів	34
3.2 Сутність програмних засобів, використаних при реалізації проєкту.	36
3.3 Архітектура програмної систем.....	43
3.4 Алгоритм взаємодії системи з користувачем	45
3.5 Візуальні результати роботи Telegram-бота	47
3.6 Аналіз ефективності системи	51
ВИСНОВОК	53
ПЕРЕЛІК ПОСИЛАНЬ.....	55
ДОДАТКИ	58
Додаток А: Таблиця 2.1 Архітектура згорткової нейронної мережі.....	58
Додаток Б: Програмний код.....	59
ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ.....	67

ВСТУП

Сьогодні розвиток штучного інтелекту значно розширює можливості взаємодії користувача з цифровими сервісами. Швидке зростання кількості відео-контенту, доступного у мережі, створює проблему перевантаження інформацією. Більшість традиційних рекомендаційних систем базується на історії переглядів або вподобаннях користувача, однак вони рідко враховують важливий аспект — емоційний стан людини в конкретний момент [1]. Це обмежує здатність системи надати по-справжньому релевантні рекомендації.

Актуальність теми визначається кількома важливими чинниками. По-перше, зростаючий обсяг даних потребує нових методів аналізу, здатних адаптуватися до контексту та настрою користувача. По-друге, інтеграція емоційного фактора у цифрові сервіси відкриває можливості для більш гнучкої та персоналізованої взаємодії. По-третє, розвиток комп'ютерного зору та глибинного навчання, зокрема застосування згорткових нейронних мереж, створює передумови для автоматизованого розпізнавання емоцій.

Для українського суспільства ця тема також має значення, адже в умовах активної цифровізації та розвитку соціальних медіа важливо не лише підбирати корисний контент, а й робити це з урахуванням психологічного стану людини. Аналіз настрою та емоційної реакції стає запорукою кращого користувацького досвіду у різних сферах — від розважальних платформ до освітніх і професійних сервісів.

Розпізнавання емоцій на основі відеозаписів обличчя — складне завдання, адже емоційний вираз залежить від безлічі факторів: особистісних рис, культурних відмінностей, ситуативного контексту. Для ефективного вирішення цієї проблеми важливо враховувати деталі міміки, положення рис обличчя та їхні взаємодії.

Мета кваліфікаційної роботи полягає у розробці інтелектуальної системи рекомендацій відео-контенту, що враховує емоційний стан користувача. Система аналізує відеозапис обличчя у реальному часі за допомогою комп'ютерного зору та згорткової нейронної мережі, після чого підбирає відповідні рекомендації у

Telegram-боті.

Об'єкт дослідження — процес розпізнавання емоцій користувача та формування персоналізованих рекомендацій відео-контенту.

Предмет дослідження — інтелектуальна система, яка інтегрує розпізнавання емоцій та рекомендаційний модуль у єдиному інтерфейсі.

Важливість дослідження обумовлена потребою створення рішень, що допомагають людям отримувати контент, найбільш релевантний їхньому психологічному стану. Такі системи можуть знайти застосування у розважальній індустрії, маркетингу, освітніх платформах і навіть у медичній сфері для покращення емоційного комфорту.

У результаті роботи очікується створення прототипу системи, яка автоматично визначає емоційний стан користувача та пропонує відповідний відеоконтент. Реалізація такого підходу демонструє перспективи використання технологій глибокого навчання та комп'ютерного зору для підвищення рівня персоналізації цифрових сервісів.

1 ОГЛЯД СУЧАСНОГО СТАНУ ДОСЛІДЖЕНЬ

У сучасному цифровому середовищі значна частина взаємодії між людьми та платформами відбувається в онлайн-просторі. Користувачі активно діляться власними думками та емоціями у відео, повідомленнях і постах у соціальних мережах. Це створює унікальне джерело даних для аналізу психоемоційного стану, який може мати значення у різних сферах — від медицини до розважальних сервісів.

Зокрема, у галузі охорони здоров'я аналіз емоційної поведінки людини дозволяє отримати додаткові відомості про її психологічний стан. Наприклад, якщо лікарю доступний відеозапис або фото пацієнта, він може оцінити міміку та загальну виразність обличчя, що допомагає у більш точній діагностиці та інтерпретації симптомів. Подібні системи аналізу, базовані на нейронних мережах, дозволяють автоматично визначати емоції людини без упередженого впливу людського фактора.

Водночас, у контексті цифрових розважальних сервісів, емоційний стан користувача може відігравати важливу роль у виборі контенту. Рекомендаційні системи, які враховують не лише історію переглядів, а й поточний емоційний фон користувача, здатні пропонувати більш релевантні пропозиції — наприклад, динамічні фільми для підняття настрою чи спокійні мелодрами для емоційного розслаблення.

Сучасні дослідження у цій сфері демонструють комплексний підхід до аналізу емоцій. Серед ключових напрямків варто виділити:

- *Комп'ютерний зір* — застосування алгоритмів для виділення обличчя у відео та оцінки його ключових рис (наприклад, зміщення брів чи рухи губ). Бібліотеки на кшталт OpenCV забезпечують технічну базу для такої обробки.
- *Глибинне навчання* — особливо згорткові нейронні мережі (CNN), які можуть навчатися на великих обсягах даних та виявляти особливості міміки, пов'язані з певними емоціями.
- *Інтеграція у рекомендаційні моделі* — сучасні сервіси вже намагаються

поєднати результати аналізу емоцій із підбором контенту, що забезпечує більш індивідуалізований досвід користувача.

- *Реалізація у мобільних та веб-інтерфейсах* — зокрема, у форматі ботів або чат-асистентів (як-от Telegram-боти), які можуть швидко реагувати на емоційний стан людини та пропонувати відповідний відеоконтент.

Таким чином, сьогоdnішній стан досліджень у галузі розпізнавання емоцій з відео і рекомендаційного підбору контенту свідчить про наявність надійних технічних основ і потенціал для практичної реалізації. Проте залишається важливою задача інтеграції цих двох компонентів — автоматичного розпізнавання емоцій користувача та динамічної персоналізації рекомендацій. Саме це й стало ключовим завданням даної кваліфікаційної роботи.

1.1 Огляд реалізованих систем розпізнавання емоцій у світі

На сучасному етапі розвитку штучного інтелекту, дедалі більше уваги приділяється створенню систем, здатних розпізнавати емоційний стан користувача. Це завдання активно досліджується як у науковому середовищі, так і в рамках реальних інженерних проєктів, що знаходять практичне застосування у мобільних застосунках, вебсервісах, чат-ботах та системах відеоспостереження. У цьому підпункті розглянуто низку відомих реалізованих проєктів, які демонструють практичне втілення методів розпізнавання емоцій з використанням комп'ютерного зору, глибокого навчання та мультисенсорного аналізу.

1. *Replika* — це один із найвідоміших комерційних застосунків, що реалізує концепцію емоційно-інтелектуального штучного співрозмовника. Застосунок використовує текстовий аналіз і в деяких версіях — зчитування міміки через камеру пристрою, щоб адаптувати відповіді до емоційного стану користувача (рис. 1.1). *Replika* здатна моделювати підтримувальні розмови, реагувати на емоційно забарвлені висловлювання та зберігати контекст діалогу. Популярність цього сервісу, який має мільйони завантажень, свідчить про високу зацікавленість користувачів у персоналізованих системах із емоційним інтерфейсом.



Рисунок 1.1 – Загальний вигляд втреафейсу додатку Replika

2. *Affectiva* — дочірня компанія Smart Eye, що спеціалізується на емоційній аналітиці для автомобільної промисловості, UX-досліджень, реклами та медичних систем. Їхня технологія використовує відеопотік із фронтальної камери для аналізу міміки, положення голови, напрямку погляду та інших невербальних сигналів. Модель Affectiva навчається на мільйонах зразків даних з усього світу і здатна класифікувати широкий спектр емоцій та поведінкових патернів. Технологія широко інтегрується в системи безпеки транспортних засобів для виявлення ознак сонливості, агресії чи розсіяності водія (рис. 1.2).



Рисунок 1.2 – Приклад виявлення емоції водія

3. *Emojify* — демонстраційний веб-застосунок, створений для візуалізації можливостей емоційного аналізу у браузері (рис 1.3). Рішення побудоване на JavaScript та TensorFlow.js, що дозволяє виконувати обчислення локально, без відправки даних на сервер. Користувач активує камеру, а система в реальному часі розпізнає базові емоції: радість, сум, гнів, здивування, нейтральність. Хоча *Emojify* не є комерційним продуктом, воно демонструє можливості легковагових моделей для інтеграції в інтерфейси користувача.

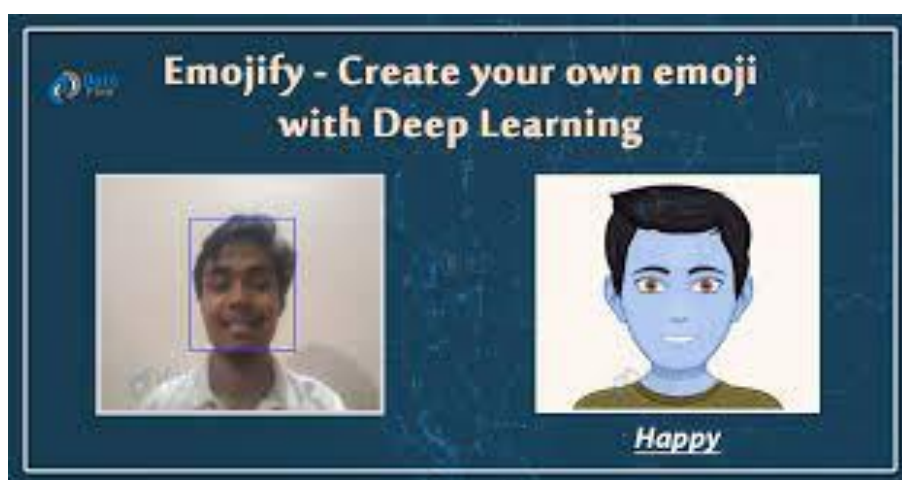


Рисунок 1.3 – Зовнішній вигляд веб-застосунку *Emojify*

4. *Netflix R&D: Mood-Based UI* - У дослідницьких підрозділах компанії Netflix реалізовувалися експериментальні інтерфейси, що враховували настрій користувача, зчитаний через камеру. Ідея полягала у формуванні адаптивної структури головної сторінки сервісу залежно від виявленої емоції — наприклад, пропонувати комедії при виявленні смутку, або драматичні фільми при виявленні тривожності чи страху. У рамках цих експериментів проводились дослідження впливу емоційного стану на вибір контенту, а також тестувались різні варіанти подачі вмісту на основі емоційної реакції користувача.

Дані про емоції могли збиратися як із вебкамери, так і з сенсорів обличчя на мобільних пристроях, після чого інтерфейс Netflix змінювався відповідно до розпізнаної емоції: змінювались кольори, мікроанімації, порядок добірок і жанрові рекомендації. Незважаючи на те, що цей проєкт залишився на стадії дослідження і

не був впроваджений у масовий продукт, він є важливим прикладом застосування емоційного аналізу в інтерфейсному дизайні великих мультимедійних платформ.

Telegram-бот системи, розробленої у дипломному проєкті У межах цієї дипломної роботи розроблено систему, яка поєднує автоматичне розпізнавання емоцій користувача на основі відео з Telegram та персоналізовану рекомендацію відео-контенту. На відміну від більшості згаданих рішень, реалізація є повністю відкритою, не потребує зовнішніх API або хмарних сервісів, та ґрунтується на власноруч навченій згортковій нейронній мережі. Система функціонує у режимі реального часу, обробляє відео-кружечки, виявляє міміку обличчя, класифікує емоцію та генерує добірку фільмів або серіалів на основі внутрішньої бази даних. Такий підхід поєднує гнучкість локального застосунку з адаптивністю емоційного аналізу, що є інноваційним для платформ типу Telegram.

5. *Tawny* — німецький стартап, який розробляє рішення для відстеження емоційного стану користувачів у режимі реального часу. Система використовує відеокамери, сенсори та нейромережеві моделі для аналізу емоцій та рівня стресу. Продукт *Tawny* інтегрується у мобільні додатки, VR/AR-середовища та корпоративні інтерфейси для забезпечення адаптивної поведінки програм відповідно до психоемоційного стану людини (рис 1.4).



Рисунок 1.4 - Приклад адаптивної взаємодії користувача з системою Tawny

1.2 Вивчення сучасних методів розпізнавання емоцій за відео

Сучасні методи розпізнавання емоцій за відео ґрунтуються на поєднанні комп'ютерного зору, глибинного навчання та психологічного аналізу [3]. Такі системи дають змогу перетворювати візуальні дані у числові ознаки, які можна використовувати для класифікації емоційного стану користувача.

Основою обробки відео є векторизація зображень обличчя — процес, у якому визначаються ключові характеристики, такі як положення очей, брів, губ та загальна міміка. Такі вектори забезпечують математичне представлення інформації, що дозволяє системі працювати з ними у подальших етапах. Попередню обробку кадрів і виділення області обличчя зазвичай виконують із застосуванням бібліотеки OpenCV, що забезпечує стабільність та точність розпізнавання.

Подальший аналіз здійснюється згортковими нейронними мережами (CNN), які здатні автоматично виявляти значущі ознаки без потреби ручного налаштування. CNN дозволяють системі навчатися на великих масивах відеоданих та визначати характерні патерни, що відповідають певним емоціям — злість, радість, сум, страх та інші.

Машинне навчання у цій сфері застосовується для побудови моделей, які класифікують кадри або відеофрагменти за емоційними категоріями. Це дає змогу системі автоматично розпізнавати емоційний стан користувача на основі аналізу кожного кадру та передбачати найбільш релевантний контент для подальшої взаємодії.

Для реалізації подібних підходів використовуються різні фреймворки, серед яких TensorFlow та PyTorch, що забезпечують гнучкість та масштабованість при побудові нейронних мереж. Використання цих інструментів дозволяє інтегрувати розроблену модель у Telegram-бот, створюючи зручний інтерфейс для користувача.

Хоча такі методи мають високу точність, вони потребують великої кількості навчальних даних та ретельного налаштування, особливо при роботі з відео у різних умовах освітлення або при різноманітті якості запису.

У межах цієї кваліфікаційної роботи реалізовано поєднання зазначених

підходів: застосовано OpenCV для виділення обличчя, розроблено згорткову нейронну мережу для визначення емоцій та інтегровано результати у Telegram-бота для формування персоналізованих рекомендацій. Це дозволяє створити систему, яка динамічно підлаштовується під емоційний стан користувача [9] та покращує його досвід у споживанні відео-контенту.

1.3 Аналіз психологічних станів користувачів за допомогою відеоданих

У контексті рекомендаційних систем і персоналізації цифрового досвіду важливим завданням є розуміння психологічного стану користувача. На відміну від текстового аналізу, що застосовується у сфері маркетингових досліджень чи соціальних мереж, у цій кваліфікаційній роботі досліджується саме відеоаналіз — розпізнавання емоцій за мімікою обличчя.

Аналіз емоційної поведінки на основі відео дозволяє виявляти миттєві психологічні реакції користувача. Такі системи можуть автоматично класифікувати стан людини (наприклад, злість, радість, страх, сум) та адаптувати рекомендації контенту відповідно до виявленого емоційного настрою. В основі цих підходів лежить поєднання комп'ютерного зору для виділення обличчя з відео та згорткових нейронних мереж для класифікації емоційного стану.

Етапи роботи таких систем включають кілька ключових кроків. Спочатку відеодані проходять попередню обробку — з кадрів виділяються області обличчя, усуваються шуми та непотрібні деталі. Потім виділені ознаки (риси обличчя, міміка) перетворюються на числові вектори — набір даних, який можна обробляти у моделі машинного навчання. Згорткова нейронна мережа навчається на цих векторах для точного визначення емоційної категорії.

Подібний аналіз психологічного стану користувача у режимі реального часу має практичне значення у багатьох сферах. Наприклад, у розважальних сервісах він дозволяє підбирати фільми чи серіали, які відповідають актуальному настрою людини. У маркетингу — підвищувати ефективність рекомендацій, зменшуючи ризик «холодних» або непридатних пропозицій. У сфері освіти чи медицини —

забезпечувати більш персоналізований підхід у підготовці матеріалів або взаємодії з пацієнтом.

Хоча такі системи потребують значних обсягів даних для навчання та ретельного налаштування моделей, вони дають змогу отримувати об'єктивні результати, мінімізуючи вплив суб'єктивних факторів чи випадкових упереджень. У цій кваліфікаційній роботі реалізовано застосування такого аналізу для рекомендаційного підбору контенту, що враховує не тільки історію взаємодії користувача, а й його поточний психологічний стан.

1.4 Аналіз сучасних програмних рішень для розпізнавання емоцій

На сьогодні існує значна кількість програмних рішень і бібліотек, які дозволяють здійснювати розпізнавання обличчя та аналіз емоційного стану користувачів на основі відео. Такі рішення можна умовно розділити на комерційні сервіси, готові API та відкриті бібліотеки, кожен з яких має свої переваги й обмеження.

Комерційні сервіси (SaaS) пропонують комплексне рішення, що не потребує глибоких знань у сфері машинного навчання. Користувач підключає свій застосунок до сервісу через API, а всі обчислювальні процеси відбуваються на стороні постачальника. Це зручно, але потребує постійного підключення до Інтернету, а також створює ризики для конфіденційності, оскільки дані обробляються стороннім провайдером. Крім того, такі сервіси зазвичай є дорогими, що може обмежувати їх використання для невеликих проєктів

Іншим напрямом є локальні REST API-рішення, які можна розгорнути на власному сервері. Такі варіанти дозволяють зберігати дані всередині компанії, знижуючи ризики витоку інформації. Водночас вони вимагають певних технічних знань для встановлення та налаштування серверів, зокрема роботи з контейнерами Docker.

Відкриті бібліотеки та фреймворки є ще одним поширеним варіантом. Серед найпопулярніших — OpenCV, який забезпечує базову обробку зображень та відео,

а також TensorFlow і PyTorch, що дозволяють будувати глибинні моделі для класифікації емоцій. Ці бібліотеки вимагають від користувача знань у галузі машинного навчання, однак надають високу гнучкість і можливість адаптації до конкретних потреб.

Також варто відзначити кілька рішень, які активно використовуються у проєктах розпізнавання обличчя та емоцій:

- *FaceNet* та *InsightFace* — побудовані на концепції векторного простору, де кожне обличчя представлено як унікальний вектор. Вони забезпечують високу точність та підтримують масштабування для обробки великих обсягів даних.
- *CompreFace* — відкрите рішення з підтримкою REST API, яке дозволяє легко інтегрувати функції розпізнавання у власні застосунки.
- *DeepFace* — популярна Python-бібліотека, що використовує декілька алгоритмів розпізнавання одночасно, зокрема для визначення емоційних станів.
- *InsightFace-REST* — інтерфейс REST API, створений на основі InsightFace, що забезпечує зручність інтеграції та високу продуктивність.
- Крім того, серед комерційних рішень для бізнесу активно використовуються такі сервіси, як Amazon Rekognition, Microsoft Azure Face API, Face++, які пропонують широкий спектр функцій, включаючи аналіз емоцій, визначення віку та статі, а також відстеження облич на відео у реальному часі.

Попри багатий вибір інструментів, важливо враховувати потреби конкретного проєкту. У межах цієї кваліфікаційної роботи зосереджено увагу на використанні OpenCV для виявлення обличчя та згорткової нейронної мережі для класифікації емоційного стану користувача. Такий підхід дозволяє зберігати повний контроль над даними й інтегрувати розроблену систему у Telegram-бот, що забезпечує персоналізовані рекомендації відео-контенту відповідно до настрою користувача.

Таким чином, огляд існуючих рішень демонструє широкий спектр можливостей і технологічних підходів. Водночас він підкреслює важливість власної розробки адаптованої системи, яка відповідає потребам користувача та

забезпечує конфіденційність і високу точність роботи.

1.5 Загальні аспекти систем розпізнавання емоцій

Виявлення емоційного стану людини за допомогою відео — це складне завдання, що потребує врахування багатьох технічних і психологічних факторів. Часто одна й та ж міміка може сприйматися по-різному різними людьми, що створює додаткові труднощі для автоматизованих систем. Застосування методів глибинного навчання, зокрема згорткових нейронних мереж, дозволяє суттєво підвищити точність розпізнавання емоцій, проте не усуває усіх проблем, пов'язаних із різноманіттям виразів обличчя.

Розпізнавання емоцій людини на основі зображень і відео становить складну задачу, що вимагає ретельного аналізу численних чинників. Це особливо актуально для систем, які функціонують у режимі реального часу та мають забезпечувати індивідуальний підхід до користувача, як передбачено в цій кваліфікаційній роботі.

Одним із основних технічних викликів є обмеженість навчальних даних. Для успішної роботи моделі необхідно мати великий обсяг відеоматеріалів, які б відображали різні емоції у різних культурних і соціальних контекстах. Проте більшість відкритих датасетів не мають достатньої різноманітності, що знижує точність моделі при розпізнаванні емоцій у користувачів з різним фоном. Щоб частково розв'язати цю проблему, використовуються підходи на кшталт об'єднання кількох датасетів або доповнення даних шляхом модифікацій зображень (зміна освітлення, додавання шуму, зміна кутів огляду).

Наступний виклик — проблеми з візуальними перешкодами та освітленням. Емоції розпізнаються за дрібними змінами у міміці обличчя, тому тіні, слабе освітлення, неякісна зйомка або різка зміна кадрів можуть суттєво знижувати ефективність аналізу. Для цього в роботі застосовуються інструменти комп'ютерного зору (OpenCV) для виділення обличчя та його фронталізації, а також досліджуються методи компенсації зміни умов освітлення.

Крім того, індивідуальні особливості зовнішності та вираження емоцій також

ускладнюють задачу. Різні люди можуть по-різному виражати одну й ту ж емоцію, а також у процесі зйомки можуть змінювати положення голови чи використовувати аксесуари (окуляри, головні убори), що ускладнює ідентифікацію.

До основних психологічних викликів можна віднести, враховуючи наші умови чинники:

- *Культурні відмінності*: люди з різних культур демонструють однакові емоції різними способами, тому універсальні алгоритми можуть неадекватно інтерпретувати сигнали.
- *Вік користувача*: діти і дорослі по-різному проявляють емоції. У дітей, наприклад, реакції можуть бути більш емоційно насиченими або менш контрольованими.
- *Приховані чи слабкі емоції*: люди іноді свідомо приховують свій емоційний стан. Системи, які розраховані лише на яскраво виражені емоції, можуть не виявити цих слабких сигналів.

Одним із важливих викликів у завданні розпізнавання емоцій є врахування емоційного контексту. Емоції проявляються не лише через міміку, а й через загальну позу тіла та навколишнє оточення, яке може містити додаткові підказки. Наприклад, нахил корпусу вперед іноді свідчить про зацікавленість або агресію, а жести руками можуть підсилювати або уточнювати вираз обличчя. У межах цієї роботи головна увага приділяється саме аналізу обличчя як основного джерела емоційної інформації, тоді як контекстні сигнали розглядаються як другорядні і не є визначальними для прийняття рішення.

У межах даної кваліфікаційної роботи такі проблеми враховані при розробці Telegram-бота, який підбирає відео-контент відповідно до емоційного стану користувача. Система спирається на візуальний аналіз відео, а саме на розпізнавання міміки обличчя за допомогою згорткових нейронних мереж. Це дозволяє автоматично визначати емоційний стан користувача, попри всі вищезгадані технічні та психологічні перешкоди, і пропонувати контент, що найбільш відповідає його настрою.

Таким чином, системи рекомендацій, які інтегрують емоційний аналіз,

зокрема в режимі реального часу, мають значний потенціал для покращення персоналізації цифрових сервісів. Подолання зазначених технічних і психологічних викликів є ключем до підвищення точності та якості таких рішень.

Системи розпізнавання обличчя дедалі активніше впроваджуються у різні сфери життя — від соціальних медіа до державного контролю. Вони дозволяють ідентифікувати особу або визначати її стан, використовуючи відео- чи фотоматеріали. Разом з тим, така технологія викликає низку дискусій щодо її використання та впливу на суспільство.

Зокрема, органи державної безпеки активно використовують розпізнавання обличчя для відстеження осіб, що викликає занепокоєння з приводу порушення приватності громадян. Дані досліджень, проведених Джорджтаунським університетом, свідчать, що бази з розпізнавання обличчя уже охоплюють значну частину населення, що піднімає питання законності та етичності збирання й використання таких даних. Наприклад, у Гонконгу поліція застосовувала подібні системи для моніторингу учасників протестів, а в інших країнах такі технології використовуються для попередження правопорушень або виявлення порушників.

Водночас, у відповідь на масове впровадження таких технологій з'являються рухи протидії. Активісти та дослідники розробляють спеціальний одяг чи аксесуари, які допомагають уникнути детекції системами розпізнавання обличчя. Маски та інші засоби анонімності набули широкого поширення серед протестувальників у різних країнах. Також є ініціативи на кшталт проєктів Signal, які надають учасникам протестів інструменти для приховування обличчя, тим самим захищаючи право на приватність.

Ще однією формою уникнення розпізнавання є використання підставних фотографій або цифрових маніпуляцій для введення систем в оману. Хоча наразі такі підходи більше зосереджені у наукових дослідженнях та мистецьких проєктах, поширення технологій розпізнавання обличчя робить їх дедалі актуальнішими і для пересічних користувачів.

У контексті розробки Telegram-бота, який розпізнає емоції користувача за відео, важливо враховувати ці етичні аспекти. Зокрема, потрібно забезпечити

захист персональних даних, прозорість алгоритмів та згоду користувача на використання зображень його обличчя. Це має велике значення для запобігання порушень конфіденційності та збереження довіри користувачів до системи.

Таким чином, розвиток технологій аналізу емоцій та розпізнавання обличчя відкриває нові можливості для персоналізованого сервісу, але також потребує обережного підходу щодо етичних, юридичних та соціальних викликів, які супроводжують їх застосування.

Людські Розпізнавання емоцій людини базується на аналізі мімічних змін, які проявляються через рухи брів, очей та рота. Емоційні сигнали зазвичай концентруються в окремих зонах обличчя, які умовно поділяють на верхню частину (область очей і брів) та нижню (область рота). Такий поділ дозволяє алгоритмам зосереджуватись на найінформативніших елементах, що відображають емоційний стан. Для виявлення обличчя на зображеннях або у відеопотоці часто застосовуються алгоритми на кшталт методу Віоли-Джонса, які ефективно визначають області, що потенційно містять обличчя. Ці зони потім передаються для подальшого аналізу, під час якого виявляються ключові точки — характерні маркери, такі як контури очей, брови чи кути рота, що дозволяють ідентифікувати конкретний емоційний вираз. Процес аналізу емоцій зазвичай включає кілька етапів:

- Перетворення кольорового зображення у відтінки сірого для зниження складності обробки.
- Застосування адаптивного порогу для створення бінарного зображення.
- Використання градієнтної маски для виділення контурів обличчя.
- Локалізація ключових точок на обличчі (наприклад, визначення точок навколо очей, брів та рота).

На рис. 1.5 показано, як ці ключові точки можуть бути розташовані на обличчі. Розміщення цих точок допомагає системі розпізнавання правильно інтерпретувати вирази емоцій.

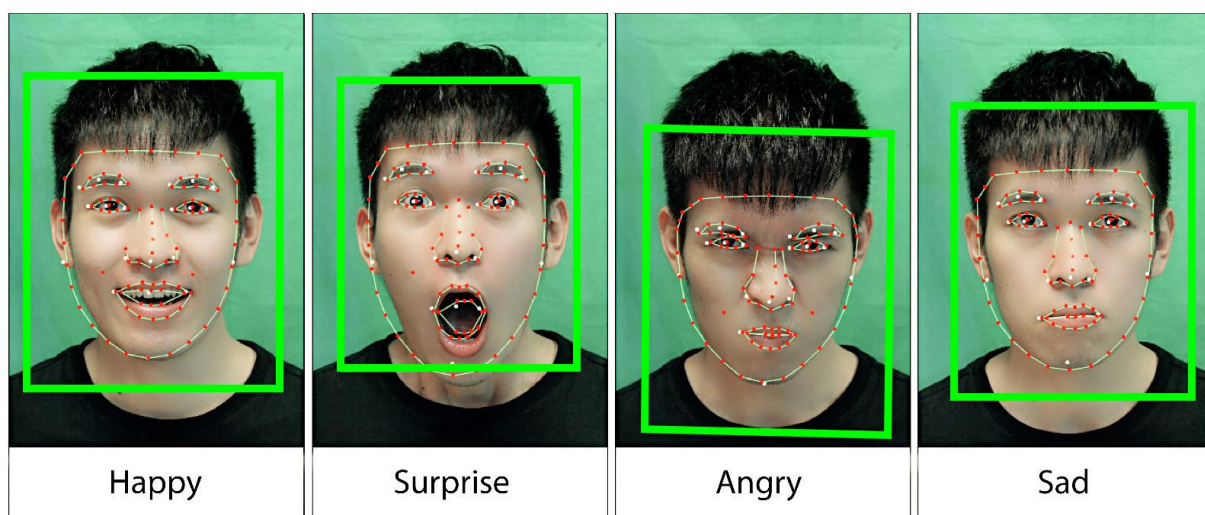


Рисунок 1.5 - Візуалізація ключових точок обличчя для різних емоційних станів

Ці ключові точки виступають основою для подальшої побудови математичних моделей, які дозволяють системі точно розпізнавати емоційні стани людини. Залежно від зміни положення маркерів, система може аналізувати ступінь відкриття рота, нахил брів або форму очей — усе це є критичними ознаками при класифікації емоцій. Таким чином, поєднання алгоритмів виявлення обличчя та локалізації ключових точок створює надійне підґрунтя для побудови ефективної системи автоматичного розпізнавання емоцій.

В результаті, було здійснено комплексний огляд сучасного стану досліджень у сфері автоматичного розпізнавання емоцій користувачів на основі відеоданих. Було проаналізовано актуальні підходи, технічні рішення та програмні інструменти, які застосовуються для побудови систем емоційного аналізу.

Особливу увагу приділено ролі комп'ютерного зору, згорткових нейронних мереж і бібліотек типу OpenCV, TensorFlow, PyTorch у процесі виявлення обличчя та класифікації емоційного стану. Також розглянуто значення векторизації міміки та локалізації ключових точок на обличчі як основи для побудови моделей розпізнавання.

Крім технічних аспектів, було враховано психологічні та етичні чинники, пов'язані з варіативністю емоційних проявів, культурними відмінностями, питанням конфіденційності даних та використанням таких систем у суспільстві.

На основі аналізу зроблено висновок, що поєднання візуального аналізу

відео, глибинного навчання та адаптивних рекомендаційних систем є перспективним напрямом для створення індивідуалізованих сервісів. Розробка Telegram-бота, який у режимі реального часу формує персоналізовані відеорекомендації на основі емоційного стану користувача, логічно вписується в цей тренд і є актуальною інженерною задачею, що поєднує сучасні наукові досягнення з прикладною реалізацією.

2 ТЕХНІЧНІ АСПЕКТИ СИСТЕМ РОЗПІЗНАВАННЯ ЕМОЦІЙ

2.1 Загальні поняття штучних нейронних мереж для розпізнавання відео-контенту

Штучні нейронні мережі (ШНМ) є основою більшості сучасних систем штучного інтелекту, особливо в тих задачах, які пов'язані з аналізом візуальної інформації, зокрема з відео. Їхня ефективність обумовлена здатністю моделювати складні нелінійні залежності між ознаками даних, які важко формалізувати традиційними алгоритмами. У випадку аналізу відео-контенту, нейронні мережі дозволяють виявляти та інтерпретувати патерни поведінки, міміки та інших візуальних сигналів, що є ключовими для задачі розпізнавання емоцій.

Штучна нейронна мережа — це математична модель, яка наслідує структуру та функції біологічних нейронів. Вона складається з великої кількості обчислювальних елементів (нейронів), об'єднаних у шари. Кожен нейрон приймає на вхід числові значення (ознаки), обробляє їх за допомогою вагових коефіцієнтів та функцій активації, а потім передає результат далі. Сукупність таких нейронів і зв'язків між ними утворює архітектуру мережі, яка здатна адаптуватись до нових даних під час навчання.

У типових задачах розпізнавання відео-контенту, зокрема міміки обличчя, використовується багатошарова структура ШНМ. Вхідний шар отримує зображення, які були виділені з відеопотоку (наприклад, окремі кадри з обличчям користувача). Далі, ці зображення проходять через кілька прихованих шарів, де мережа вчиться виявляти базові ознаки (лінії, кути, контури), а потім — більш складні елементи (форма очей, розташування рота, вираз обличчя). Вихідний шар повертає результат у вигляді вектору з ймовірностями, що відповідають різним емоційним станам (наприклад: «злість», «радість», «сум», «страх», тощо).

Головною перевагою ШНМ у роботі з відео є їх здатність до навчання на прикладах. На відміну від класичних алгоритмів, де необхідно вручну задавати правила для розпізнавання, нейронна мережа самостійно формує внутрішнє

уявлення про структуру даних. Це дозволяє моделі виявляти навіть ті особливості міміки, які неочевидні для людини. Такий підхід особливо ефективний у роботі з реальними відеозаписами, де можуть виникати перешкоди — шум, нестабільне освітлення, різні кути огляду.

Штучні нейронні мережі для розпізнавання відео-контенту володіють низкою важливих властивостей. По-перше, це адаптивне навчання — під час навчання вагові коефіцієнти мережі коригуються відповідно до помилок прогнозування, що дозволяє моделі вдосконалювати свою точність з кожною ітерацією. По-друге, це самоорганізація — мережа здатна структурувати нові знання без потреби зовнішнього втручання, самостійно знаходячи приховані залежності. По-третє, це паралельна обробка даних — нейронна мережа може обробляти великі обсяги інформації одночасно, що критично важливо при роботі з відео. Нарешті, мережі мають високу відмовостійкість: навіть у разі втрати частини даних (наприклад, частково закриті обличчя або низька якість кадру), вони здатні надати прийнятний результат.

У контексті цієї роботи штучна нейронна мережа використовується для розпізнавання емоцій за відеофрагментом, отриманим від користувача у форматі телеграм-кружечка. Кожен кадр із відео обробляється окремо, і на його основі модель прогнозує емоційний стан. Надалі ці прогнози використовуються в системі персоналізованих рекомендацій — наприклад, для підбору фільму, який відповідає настрою користувача.

Завдяки своїм властивостям, ШНМ стали ефективним інструментом для автоматизації аналізу відео-контенту. Вони дозволяють не лише розпізнавати стани, але й інтегрувати емоційний аналіз у взаємодію з цифровими сервісами, такими як Telegram-боти, мобільні додатки або веб-інтерфейси. Це відкриває широкі можливості для персоналізації та адаптації контенту до індивідуального сприйняття кожного користувача.

2.2 Структура згорткових нейронних мереж для розпізнавання відео-контенту

Згорткові нейронні мережі (Convolutional Neural Networks, CNN) є найпоширенішим типом глибоких моделей, що спеціалізуються на обробці візуальної інформації. У контексті розпізнавання емоцій у відеоконтенті CNN дозволяють аналізувати міміку обличчя користувача, виявляючи ознаки, що відповідають певному емоційному стану. Саме завдяки здатності виявляти локальні особливості на зображеннях та узагальнювати їх у високорівневі патерни, згорткові мережі є основним інструментом для роботи з кадрами відео.

Архітектура згорткової мережі складається з послідовності спеціалізованих шарів, кожен з яких виконує свою роль у трансформації вхідного зображення. Схематичну структуру типового CNN у задачі розпізнавання емоцій наведено на рис. 2.1.

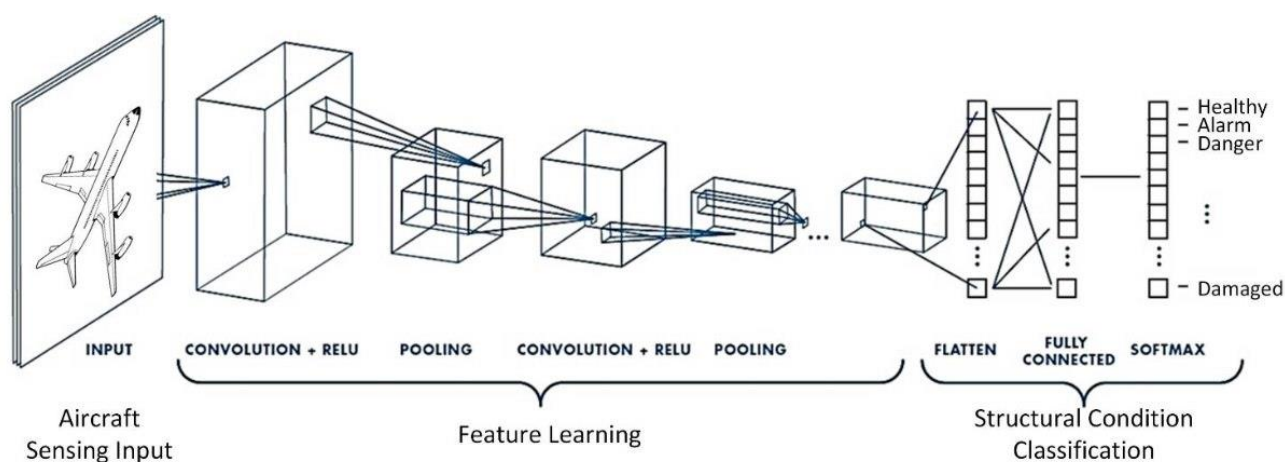


Рисунок 2.1 - Загальна архітектура згорткової нейронної мереж [14]

Основу мережі складають згорткові шари, які застосовують фільтри (ядра згортки) до зображення для виділення локальних ознак, таких як краї, текстури, форми очей або лінії брів. Ці шари створюють так звані карти ознак (feature maps), які є проміжними поданнями вхідного зображення на різних рівнях абстракції.

Після згорткових шарів зазвичай слідує шар підвибірки або пулінгу (наприклад, MaxPooling), які зменшують просторові розміри карти ознак. Це

дозволяє зменшити обсяг обчислень, зробити модель стійкішою до зсувів та масштабування об'єктів на зображенні, а також допомагає зменшити перенавчання. Пулінгова операція дозволяє зберігати найбільш інформативні значення, відкидаючи зайві деталі.

Нелінійні функції активації, найчастіше ReLU (Rectified Linear Unit), додають до моделі здатність моделювати складні залежності між вхідними ознаками та виходом. Їх застосування після кожного згорткового шару дозволяє мережі ефективно працювати навіть з великою кількістю фільтрів та глибокою архітектурою.

На завершальному етапі мережі містять один або кілька повнозв'язних (Dense) шарів, які виконують остаточну класифікацію. У задачі емоційного аналізу останній шар зазвичай має кількість нейронів, що відповідає кількості класів емоцій (наприклад, сім: «злість», «страх», «радість», «сум», «нейтрально», «здивування», «огода»), а функція активації Softmax перетворює вихід у ймовірнісний розподіл по класах.

Процес навчання CNN включає оптимізацію параметрів фільтрів та ваг за допомогою методу зворотного поширення помилки (backpropagation) і алгоритму оптимізації (наприклад, Adam або SGD). Модель навчається на великій кількості зображень, адаптуючи свої параметри таким чином, щоб мінімізувати різницю між передбаченням і реальною міткою емоції.

Важливою перевагою CNN у роботі з відео-контентом є можливість повторного використання навчених фільтрів. Наприклад, модель, навчена на статичних зображеннях, може бути адаптована для роботи з відео шляхом покадрового аналізу або подальшого навчання на відеореєстраціях. Це дозволяє зменшити ресурси, необхідні для навчання з нуля, і прискорити процес розробки.

Таким чином, згорткові нейронні мережі дозволяють формувати ієрархічні уявлення про зображення обличчя, починаючи від простих елементів до складних комбінацій, що відповідають емоційним станам. Завдяки своїй гнучкості, масштабованості та високій точності, CNN стали основним інструментом у задачах розпізнавання емоцій за відео, і є ключовим компонентом реалізованого в цій

роботі Telegram-бота, який формує рекомендації на основі аналізу міміки користувача.

2.3 Архітектура згорткової нейронної мережі проєктованого застосування

Для реалізації розпізнавання емоцій користувача у відеозаписах у межах цього проєкту було розроблено та навчено власну згорткову нейронну мережу. Вона побудована з урахуванням характеру вхідних даних — сіро-тонуваних зображень обличчя розміром 48×48 пікселів, які є типовими для датасетів на зразок FER-2013 [7]. Основною метою було створити архітектуру, здатну ефективно виявляти мімічні особливості, характерні для емоційних станів, при цьому забезпечити стабільну роботу з кадрами відео навіть за наявності шумів, зміни освітлення або неідеального положення обличчя.

Мережа складається з чотирьох послідовних згорткових блоків, у кожному з яких фільтри виділяють ознаки різного рівня складності — від простих контурів до складних композицій міміки. Кожен блок включає згорткові шари з активацією ReLU, нормалізацію пакету для стабілізації навчання, шар підвибірки (MaxPooling) для зменшення розмірності та шар Dropout, що запобігає перенавчанню. Кількість фільтрів у шарах поступово збільшується — з 32 до 256, що дозволяє мережі накопичувати дедалі більше контекстуальних ознак про форму, симетрію та зміну міміки.

Після завершення фази виділення ознак отримані тривимірні тензори перетворюються на одновимірний вектор за допомогою операції згортання (Flatten), після чого проходять через два повнозв'язаних шари з активацією ReLU та додатковою нормалізацією. Заключним етапом є вихідний шар із сімома нейронами, кожен з яких відповідає одній з емоцій: злість, страх, радість, сум, нейтральний стан, здивування та огида. Функція активації Softmax забезпечує перетворення результату на ймовірнісний розподіл, що дозволяє визначити найбільш ймовірну емоцію, виражену на обличчі.

Загальна кількість параметрів моделі перевищує 1,3 мільйона, з яких понад 99% є тренованими. Це свідчить про значну обчислювальну глибину, достатню для якісного аналізу міміки, та водночас оптимальну для використання в мобільному або вебзастосунку. Архітектура розроблена з урахуванням можливості інтеграції в Telegram-бот, який отримує відеозапис від користувача, розбиває його на кадри та подає їх на вхід нейронної мережі. Кінцевий результат класифікації використовується для формування персоналізованої рекомендації відео-контенту (Додаток А).

Таким чином, побудована згорткова нейронна мережа поєднує ефективність у виявленні мімічних елементів, стійкість до спотворень у відео та оптимізацію для інтерактивного застосування. Вона є ядром системи розпізнавання емоцій у запропонованому проєкті.

2.4 Алгоритм навчання згорткової нейронної мережі для системи розпізнавання відеоконтенту

Процес навчання згорткової нейронної мережі (CNN), реалізованої в цьому проєкті, був критично важливим етапом для досягнення високої точності класифікації емоцій користувача на основі відео. Метою навчання є оптимізація внутрішніх параметрів моделі (ваг і зміщень), що дозволяє їй точно виявляти мімічні патерни, характерні для конкретних емоційних станів.

З огляду на характер задачі, було застосовано підхід контрольованого навчання (supervised learning), за якого кожному зображенню відповідає мітка емоційного класу. Це дає змогу мережі порівнювати власні передбачення з правильними відповідями і поступово вдосконалювати здатність до узагальнення.

Навчання здійснювалося на сірононованих зображеннях обличч розміром 48×48 пікселів, подібних до зразків із набору FER-2013 [6]. Дані проходили попередню обробку — нормалізацію значень пікселів у діапазон $[0, 1]$ та, за потреби, аугментацію. Вибірка була поділена на тренувальний та валідаційний піднабори для контролю ефективності.

Модель тренувалася в багатьох епохах із використанням оптимізатора Adam, що забезпечує ефективне оновлення ваг завдяки адаптивній зміні швидкості навчання. Функція втрат — categorical crossentropy, яка є стандартом для багатокласової класифікації. Для підвищення стійкості до перенавчання застосовувались Dropout і BatchNormalization.

Алгоритм навчання включав такі кроки:

1. Ініціалізація структури нейронної мережі з випадковими вагами.
2. Завантаження навчальних зображень і відповідних міток емоцій.
3. Розбиття даних на тренувальну та валідаційну вибірки.
4. Пакетна подача зображень у модель (batch training).
5. Пряме проходження через усі шари мережі (forward pass).
6. Обчислення функції втрат.
7. Зворотне поширення помилки (backpropagation).
8. Оновлення ваг через алгоритм Adam.
9. Оцінка точності після кожної епохи.
10. Зупинка навчання за умовою EarlyStopping.
11. Збереження моделі у форматі .h5.

У результаті модель досягла збалансованого співвідношення точності та здатності до узагальнення. Після завершення навчання вона була інтегрована у Telegram-бот, який здійснює покадрову класифікацію емоцій користувача на основі відеозаписів та формує персоналізовані рекомендації фільмів відповідно до виявленого емоційного стану.

В результаті, було здійснено огляд теоретичних та технічних основ, необхідних для побудови системи розпізнавання емоцій за відео-контентом. Розглянуто принципи функціонування штучних нейронних мереж, зокрема згорткових (CNN), які є найбільш ефективними в задачах обробки візуальної інформації. Детально проаналізовано особливості архітектури мережі, адаптованої до вхідних даних у форматі зображень обличчя, та охарактеризовано структуру власної моделі, розробленої в межах проєкту.

Окрему увагу приділено алгоритму навчання моделі, що включає попередню

обробку даних, використання контрольованого підходу, регуляризацію та оптимізацію параметрів. Навчання здійснювалося на зображеннях з емоційними мітками, що забезпечило формування стійких і узагальнених ознак для точного розпізнавання емоцій.

Отже, результати цього етапу заклали технічну основу для створення працездатної системи, яка інтегрується в Telegram-бот та виконує емоційний аналіз користувача у відеоформаті. У наступному розділі буде представлено практичну реалізацію програмного забезпечення та результати тестування розробленої системи.

3. ОСНОВНІ ТЕХНОЛОГІЧНІ ЗАСАДИ СТВОРЕННЯ СИСТЕМИ

3.1 Обґрунтування вибору мови програмування та інструментів

У рамках реалізації системи для емоційного підбору відео-контенту було обрано мову програмування Python. Цей вибір обумовлений багатьма факторами, зокрема доступністю великої кількості готових рішень та бібліотек, гнучкістю синтаксису та легкістю підтримки коду [18]. Python давно зарекомендував себе як універсальний інструмент для розробки штучного інтелекту та систем комп'ютерного зору, що дозволило ефективно поєднати всі необхідні компоненти в єдиному проєкті.

Python підтримує різноманітні підходи до програмування — від об'єктно-орієнтованого до процедурного та функціонального, що дозволило зручно реалізувати і модуль з визначення емоцій користувача, і модуль рекомендаційного підбору контенту. Завдяки зручному синтаксису Python полегшує читання та модифікацію коду, а це особливо важливо під час тестування та вдосконалення системи.

У розробці системи було задіяно низку сучасних та перевірених інструментів, які забезпечили її стабільність та надійність. Зокрема, OpenCV використовується для обробки відео з камери або відеозапису та для виявлення обличчя користувача. Бібліотеки Keras і TensorFlow стали основою для побудови та використання згорткової нейронної мережі, яка визначає емоційний стан за відеокадрами. Для роботи з великими масивами даних та виконання математичних операцій активно застосовувався NumPy, а Pandas дозволила зручно обробляти таблиці та датасети з описами фільмів та серіалів.

Python має ще одну важливу перевагу — універсальність. Його можна застосовувати для розробки програмного забезпечення на будь-якій платформі — Windows, Linux чи macOS — без необхідності значних змін у коді. Це дозволяє легко переносити та масштабувати проєкт, роблячи його доступним для різних користувацьких сценаріїв.

Важлива риса Python — підтримка інтерактивного програмування. Можливість швидко перевіряти частини коду у режимі REPL (Read-Eval-Print Loop) дала змогу оперативно виявляти та усувати помилки, а також вдосконалювати окремі елементи системи. Це значно прискорило розробку та тестування основних функцій.

Крім того, Python активно підтримується численною спільнотою розробників. Існує велика кількість навчальних матеріалів, форумів та документації, що спрощує пошук рішень і дає змогу швидко вдосконалювати навички програмування. Це виявилось важливим для вдосконалення проєкту та адаптації його до потреб користувача.

Також варто підкреслити, що Python забезпечує гнучку інтеграцію з іншими мовами програмування та зовнішніми технологіями, такими як C, C++ та Java. Це дає можливість розширювати функціонал системи або інтегрувати її з іншими сервісами без суттєвих перешкод.

Завдяки такому багатому набору переваг Python став оптимальною мовою для розробки системи, яка об'єднує комп'ютерний зір, згорткову нейронну мережу та рекомендаційний модуль у рамках чат-бота Telegram. Такий вибір забезпечив швидку розробку, стабільність роботи та можливість легкого масштабування у майбутньому.

Індексатор пакетів Python, також знаний як PyPI (Python Package Index), є офіційним репозиторієм програмного забезпечення для Python. Він дозволяє розробникам швидко знаходити потрібні бібліотеки, встановлювати їх та інтегрувати у власні проєкти. PyPI є невід'ємною складовою екосистеми Python і дає змогу легко встановлювати й оновлювати необхідні пакети та програми.

Інтеграція PyPI із pip — стандартним менеджером пакетів Python — значно спрощує процес управління залежностями у проєкті. Pip дозволяє розробникам встановлювати та налаштовувати бібліотеки, які не входять до стандартної бібліотеки Python, мінімізуючи проблеми сумісності та прискорюючи розгортання нових компонентів. Pip доступний починаючи з версій Python 3.4 та 2.7.9 і є обов'язковим інструментом для всіх користувачів Python. Подібно до інших мов

програмування, де використовуються `gem` (Ruby), `npm` (JavaScript) чи `NuGet` (.NET), `pip` виконує аналогічні функції у Python, забезпечуючи зручність у керуванні проєктами.

Python часто називають мовою з «вбудованими акумуляторами», адже його стандартна бібліотека вже містить велику кількість корисних модулів для роботи з текстом, файлами, мережею та іншими базовими функціями. Водночас спільнота Python постійно збагачує цей набір інструментів, додаючи до PyPI сучасні бібліотеки для спеціалізованих завдань, таких як аналіз відео, машинне навчання та рекомендаційні системи.

У межах цього проєкту використовувалися такі основні бібліотеки: `OpenCV` — для обробки відеозаписів і виявлення обличчя, `Keras` та `TensorFlow` — для створення та використання згорткової нейронної мережі, що розпізнає емоції користувача, `Pandas` — для обробки та аналізу даних у форматі таблиць, наприклад, бази даних фільмів і серіалів, а також `NumPy` — для виконання швидких математичних операцій і роботи з багатовимірними масивами. Завдяки цим бібліотекам вдалося створити потужну систему, що об'єднує модуль розпізнавання емоцій і рекомендаційний механізм у Telegram-боті.

Таким чином, використання індексатора PyPI та менеджера пакетів `pip` дозволило швидко та ефективно зібрати всі необхідні інструменти для створення системи, зменшивши складність розробки й забезпечивши стабільність та гнучкість її подальшого розвитку.

3.2 Сутність програмних засобів, використаних при реалізації проєкту

Для реалізації телеграм-боту використовуються наступні програмні засоби:

1. *OpenCV* (Open Source Computer Vision Library) — це потужна та широко використовувана відкрита бібліотека для комп'ютерного зору, обробки зображень і відео. Її популярність зумовлена великою кількістю вбудованих функцій і стабільною роботою у різних проєктах, що стосуються аналізу візуальної

інформації.

У даному проєкті OpenCV є ключовим інструментом для попередньої обробки відео, що надходить від користувача через Telegram-бот. Ця бібліотека відповідає за захоплення відеопотоку або обробку вже записаних відеозаписів, виділення кадрів та їхню подальшу обробку перед передачею у нейронну мережу для визначення емоцій.

Однією з важливих можливостей OpenCV є виявлення обличчя на зображеннях та відео. Для цього використовується класифікатор Хаара (Haar cascade classifier) [2], що дозволяє швидко і точно знаходити області обличчя у кадрі. Завдяки цьому бібліотека допомагає виділяти ключові ділянки, які далі аналізує нейронна мережа [12]. У нашому випадку це — виділення області обличчя та її підготовка до подачі у модель.

Крім цього, OpenCV надає інструменти для кадрування відео — це дає можливість працювати тільки з потрібною частиною зображення. Також вона дозволяє змінювати розмір зображення з високою точністю, що важливо, адже для роботи нейронної мережі зображення обличчя має бути уніфікованого розміру (наприклад, 48×48 пікселів) [16].

Обертання та віддзеркалення кадрів — ще одна корисна функція бібліотеки, яка допомагає при необхідності підлаштувати зображення під формат вхідних даних моделі. Такі операції важливі для стабільної роботи моделі у різних сценаріях — наприклад, коли користувач знімає відео з різних кутів.

Ще одна особливість OpenCV — можливість конвертувати кольорові зображення у відтінки сірого. Це необхідно для зниження розмірності даних і полегшення обчислень, оскільки нейронна мережа працює саме з чорно-білими зображеннями, що зменшує складність аналізу.

Додатково OpenCV дозволяє накладати графічні елементи — наприклад, малювати прямокутники або написи на зображенні. Це використовується у проєкті для візуалізації визначеної емоції чи області обличчя у вікні з результатами (наприклад, коли система показує, яке обличчя вона знайшла та який емоційний стан визначила).

Таким чином, у межах нашого проєкту OpenCV виступає універсальним і потужним інструментом для всіх етапів підготовки та обробки відео. Вона забезпечує необхідну точність та швидкість виділення обличчя та формує основу для подальшого розпізнавання емоцій користувача нейронною мережею [5].

2. *NumPy* — це одна з найважливіших бібліотек мови Python, яка використовується для високоефективної роботи з багатовимірними масивами даних. У нашому проєкті NumPy є невід’ємною складовою частиною обробки та підготовки відеоданих перед їх подачею до нейронної мережі.

Основною особливістю NumPy є її здатність працювати зі спеціальним типом даних — `ndarray`, який значно швидший та ефективніший за стандартні списки Python. Завдяки цьому бібліотека дозволяє обробляти велику кількість піксельних даних у реальному часі, що особливо важливо для роботи з відео.

У рамках нашого проєкту NumPy застосовується для кількох ключових завдань. По-перше, вона використовується для нормалізації зображень, тобто переведення значень пікселів з діапазону $[0, 255]$ у формат, зручний для роботи нейронної мережі. Це дозволяє значно підвищити точність класифікації емоцій.

По-друге, NumPy допомагає ефективно маніпулювати розмірами та формою масивів. У роботі з відео важливо змінювати структуру даних (наприклад, перетворювати тривимірні масиви кадрів у двовимірні масиви для подачі у модель), що виконується швидко та без значних витрат обчислювальних ресурсів.

Ще одна перевага бібліотеки — це швидкість обчислень. NumPy оптимізує використання пам’яті та обчислювальних потужностей комп’ютера, що робить її ідеальною для роботи з великими обсягами даних під час потокового оброблення відео. Така оптимізація дозволяє системі працювати стабільно та ефективно навіть під час обробки відеозаписів високої роздільної здатності.

Крім цього, бібліотека пропонує широкий спектр математичних функцій (лінійна алгебра, статистичні операції, векторизація), які застосовуються у багатьох етапах обробки даних та передбачення емоцій. Її здатність інтегруватися з іншими бібліотеками (зокрема, Pandas, TensorFlow, Keras) дозволяє створити єдину екосистему для розробки інтелектуальних систем.

Таким чином, у межах нашого проєкту NumPy виступає універсальним та потужним інструментом для швидкої обробки великих обсягів даних, що дозволяє підвищити продуктивність системи розпізнавання емоцій та забезпечує точність кінцевих результатів.

3. *Pandas* — це одна з найпотужніших і найпопулярніших бібліотек Python для роботи з табличними даними. Вона забезпечує гнучкий і зручний інтерфейс для зчитування, обробки, фільтрації та аналізу структурованої інформації. У нашому проєкті *Pandas* виконує ключову роль у модулі рекомендації відеоконтенту.

Основною функцією бібліотеки у контексті проєкту є робота з датасетом фільмів та серіалів, який зберігається у форматі CSV. Цей файл містить інформацію про назви, жанри, описи, рейтинги та типи відео-контенту (наприклад, фільм або серіал). За допомогою функцій бібліотеки `read_csv()` дані швидко зчитуються та перетворюються у зручну структуру — *DataFrame*, яка дозволяє здійснювати складні операції над рядками і стовпцями даних.

Pandas дозволяє реалізувати фільтрацію контенту відповідно до вподобань користувача. Наприклад, якщо користувач обрав жанр або рейтинг, ці параметри легко застосовуються до таблиці для виділення лише релевантних записів. Окрім цього, бібліотека використовується для зіставлення визначеної емоції з відповідним набором жанрів, які система має порекомендувати.

Також *Pandas* підтримує індексацію та ефективну навігацію по даних, що дозволяє системі оперативно витягувати інформацію про назву фільму, його опис, жанр, рейтинг і тип. Отриману інформацію Telegram-бот форматує у вигляді текстового повідомлення та надсилає користувачу як рекомендацію.

Ще однією важливою рисою бібліотеки є її інтеграція з іншими популярними інструментами, такими як NumPy, Matplotlib, SciPy тощо. У нашій системі вона чудово взаємодіє з результатами модуля розпізнавання емоцій, дозволяючи будувати гнучкі запити на підбір контенту залежно від психологічного стану користувача.

Завдяки своїй простоті у використанні, високій продуктивності та широкому спектру функцій, *Pandas* забезпечила надійну основу для реалізації

інтелектуального механізму персоналізованих рекомендацій у системі. Вона дозволила швидко реалізувати логіку фільтрації, сортування та вибору релевантного контенту без потреби у складному SQL чи ручній обробці великих масивів інформації.

4. *TensorFlow* — це відкрита платформа для машинного навчання, розроблена дослідницькою командою Google Brain. Вона надає потужний інструментарій для створення, тренування та виконання нейронних мереж, а також для реалізації різноманітних алгоритмів глибинного навчання. У межах нашого проєкту TensorFlow використовується опосередковано як обчислювальний бекенд для бібліотеки Keras, яка завантажує та виконує модель згорткової нейронної мережі для класифікації емоцій.

Хоча у коді проєкту явно не імпортується модуль tensorflow, бібліотека активно працює у фоновому режимі під час виконання методів Keras. Зокрема, коли викликається функція `load_model()` для завантаження попередньо навченої моделі з файлу `model.h5`, вся логіка моделі, її шари, ваги та обчислення фактично реалізуються саме за допомогою TensorFlow.

TensorFlow побудований навколо концепції обчислювального графа, де дані представлені у вигляді тензорів (багатовимірних масивів), а всі операції — у вигляді вузлів цього графа. Це дозволяє ефективно розпаралелювати обчислення та виконувати їх як на процесорах (CPU), так і на графічних процесорах (GPU), що критично важливо при роботі з великими обсягами візуальних даних, зокрема при обробці відео.

TensorFlow забезпечує:

- автоматичне диференціювання для навчання нейронної мережі;
- підтримку складних архітектур, таких як CNN, RNN, GAN;
- масштабованість — моделі можуть запускатися як на локальному комп'ютері, так і на кластері серверів або у хмарному середовищі;
- інтеграцію з Keras — високорівневим API для швидкої побудови моделей.

У нашому випадку модель згорткової нейронної мережі була попередньо навчена з використанням інструментів Keras і TensorFlow, а потім збережена у

форматі .h5. Під час роботи програми ця модель використовується для класифікації емоційного стану користувача на основі кадрів із відео. Всі розрахунки, включно з передачею даних через шари, активаціями, зважуваннями та функцією Softmax на виході, виконуються за рахунок внутрішнього ядра TensorFlow.

Таким чином, хоча взаємодія з TensorFlow у нашому проєкті не є прямою, він виступає фундаментальною технологічною основою, яка забезпечує точність і продуктивність системи розпізнавання емоцій. Без його оптимізованої обчислювальної архітектури реалізація швидкої та ефективної роботи моделі у реальному часі була б неможливою.

5. *Keras* — це високорівнева бібліотека для побудови та тренування нейронних мереж, яка призначена для зручного та швидкого створення моделей глибинного навчання. У нашому проєкті вона відіграє центральну роль у реалізації модуля розпізнавання емоцій, оскільки за її допомогою здійснюється завантаження та використання попередньо навченої згорткової нейронної мережі (CNN), збереженої у форматі .h5.

Основна перевага Keras полягає у простому й інтуїтивному синтаксисі, що значно спрощує створення складних моделей. Завдяки цьому розробник може зосередитись не на реалізації низькорівневих обчислень, а на логіці моделі та експериментах із її архітектурою. У нашій системі бібліотека використовується для завантаження нейронної мережі за допомогою методу `load_model()`, що дозволяє миттєво відновити структуру, ваги та конфігурацію моделі для подальшого прогнозування емоцій.

Keras підтримує широкий набір типів шарів, серед яких:

- згорткові шари (Conv2D),
- шари активації (ReLU, Softmax),
- шари об'єднання (MaxPooling2D),
- нормалізаційні (BatchNormalization),
- Dropout-шари, що запобігають перенавчанню.

Ці елементи активно використовуються в моделі, що класифікує шість базових емоцій: злість, страх, радість, нейтральність, смуток та здивування [17].

Бібліотека також забезпечує просту інтеграцію з іншими модулями Python та платформами, наприклад, Pandas, NumPy або TensorFlow.

Важливо зазначити, що Keras у сучасних версіях використовує TensorFlow як обчислювальний бекенд. Це означає, що всі обчислення (включно з передачею тензорів, обчисленням градієнтів та активацій) фактично виконуються засобами TensorFlow, а Keras виступає як зручна оболонка для побудови логіки мережі.

У нашому проєкті Keras не лише полегшив завантаження та запуск навченої моделі, а й зробив можливим її подальше масштабування або модифікацію. У разі необхідності систему можна адаптувати до нових емоцій, підключивши додаткові шари або замінивши функції активації, не змінюючи архітектуру загальної системи.

Таким чином, використання Keras у реалізації системи дало змогу поєднати простоту реалізації, стабільність роботи та гнучкість моделі. Завдяки цьому етап розпізнавання емоцій відбувається швидко та ефективно, що критично важливо для взаємодії користувача з Telegram-ботом у реальному часі.

6. *telebot* — це популярна бібліотека Python для створення чат-ботів у месенджері Telegram [4]. Вона забезпечує просту взаємодію з Telegram Bot API та дозволяє реалізовувати як прості, так і складні сценарії обміну повідомленнями між користувачем і ботом. У нашому проєкті бібліотека *telebot* використовується для реалізації інтерфейсу взаємодії користувача із системою емоційного підбору відеоконтенту.

Однією з головних переваг цієї бібліотеки є простий та зрозумілий синтаксис, який дозволяє швидко обробляти команди, текстові повідомлення, мультимедійні файли, а також надсилати відповіді у відповідь на дії користувача. Це критично важливо для створення логіки Telegram-бота, де діалог відбувається покроково: бот запитує тип контенту, бажаний рейтинг, приймає відео, визначає емоцію користувача і надсилає рекомендацію.

У межах реалізації проєкту бібліотека *telebot* виконує кілька важливих функцій: Обробка повідомлень: бот приймає текстові повідомлення від користувача, наприклад, вибір між "фільм" і "серіал", запит на рейтинг, підтвердження або надсилання відео.

- *Прийом відео*: користувач надсилає Telegram-боту короткий відеофрагмент (наприклад, кружечок з емоційним виразом обличчя), який зберігається локально на сервері для подальшої обробки.
- *Інтеграція з емоційним модулем*: після отримання відео бот передає файл у модуль Emotion_Detector, де виконується розпізнавання емоції.
- *Інтеграція з системою рекомендацій*: бот формує запит до модуля MovieRecommendationSystem, де за визначеною емоцією, типом контенту та рейтингом підбирається відповідний набір фільмів або серіалів.
- *Надсилення результату*: бот відправляє користувачу список з кількох рекомендованих відео, кожне з яких супроводжується коротким описом, назвою, жанром та рейтингом.

Бібліотека також дозволяє видаляти тимчасові файли, що важливо для зменшення використання пам'яті, а також обробляти винятки та помилки — наприклад, коли користувач надсилає не відео або вводить некоректні параметри.

Завдяки підтримці кнопок, клавіатур, інлайн-меню та обробки медіа, telebot дозволяє створювати не лише функціонального, а й зручного для користувача бота, що працює як повноцінна взаємодійна оболонка до складної системи штучного інтелекту.

У підсумку, використання telebot зробило можливим реалізацію простого, зрозумілого та ефективного інтерфейсу взаємодії користувача з системою. Саме ця бібліотека забезпечує зв'язок між користувачем, модулем розпізнавання емоцій та рекомендаційною системою — об'єднуючи всі компоненти проєкту в єдину робочу структуру [11].

3.3 Архітектура програмної систем

Розроблена система емоційного аналізу та рекомендацій відео-контенту реалізована як багаторівнева модульна структура [20], яка забезпечує чіткий поділ функціональних обов'язків між компонентами. Такий підхід дозволяє досягти високої гнучкості, масштабованості та зручності в супроводі програмного коду.

Архітектура системи складається з п'яти основних модулів:

1. *Модуль взаємодії з користувачем (Telegram-бот):* Відповідає за отримання запитів від користувача, обробку команд, прийом відеофайлів і передачу результатів класифікації емоцій разом із персоналізованими рекомендаціями. Реалізований на основі бібліотеки telebot, яка забезпечує зручну обробку текстових повідомлень і медіаконтенту.
2. *Модуль обробки відео та виділення кадрів:* Після отримання відео Telegram-ботом, воно зберігається локально й передається до модуля попередньої обробки. Тут відео розбивається на окремі кадри за допомогою функціоналу OpenCV. Кожен кадр перетворюється у чорно-білий формат, масштабовується до розміру 48×48 пікселів та нормалізується.
3. *Модуль виявлення обличчя:* Для кожного кадру виконується детекція обличчя за допомогою алгоритму Хаара [8]. Виділене обличчя вирізається, центрується та готується до подачі у модель глибинного навчання.
4. *Модуль класифікації емоцій:* Попередньо навчена згорткова нейронна мережа (CNN) [10], реалізована на Keras з бекендом TensorFlow [9], отримує на вхід зображення обличчя та повертає ймовірнісний розподіл по класах емоцій. Визначається емоція з найвищою ймовірністю. Цей результат є основою для формування рекомендації.
5. *Модуль формування рекомендацій:* За допомогою бібліотеки Pandas виконується фільтрація датасету фільмів і серіалів відповідно до визначеної емоції, типу контенту та рейтингових вподобань користувача. Підібрані рекомендації у вигляді назви, жанру та короткого опису надсилаються назад у Telegram-бот.

Архітектура системи представлена на рис. 3.1 (у презентації) та має лінійно-розгалужену структуру, в якій дані послідовно передаються між модулями. Такий дизайн дозволяє зручно замінювати або вдосконалювати окремі компоненти без потреби в повному переробленні всієї системи.

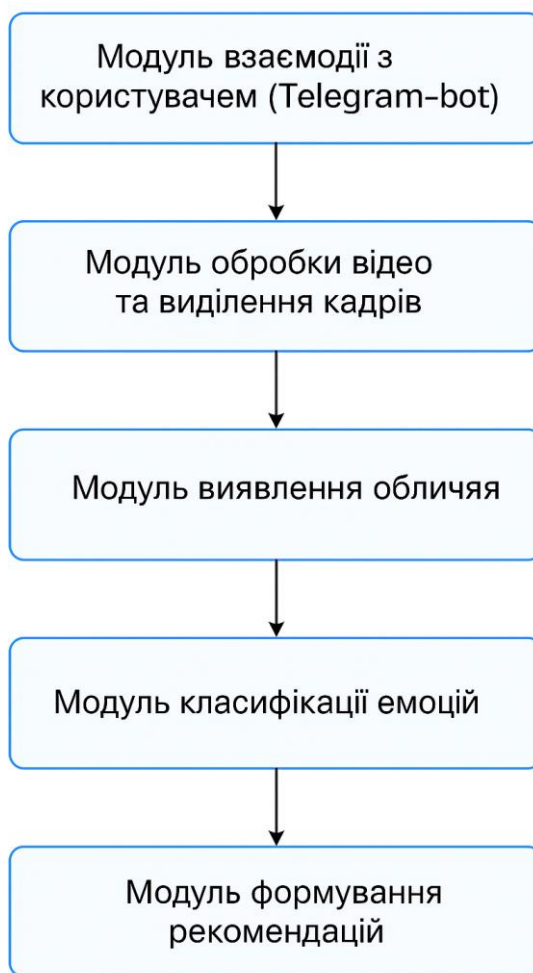


Рисунок 3.1 – Архітектура системи емоційного аналізу та рекомендацій відео-контенту

Кожен модуль функціонує автономно, проте забезпечує тісну інтеграцію з іншими за допомогою чітко визначених інтерфейсів (файлових операцій, імпорту функцій, передачі параметрів). Це створює передумови для майбутньої модифікації, наприклад, підключення додаткових джерел даних, перехід на мультимодальний аналіз (відео + аудіо) або адаптацію до нових платформ (мобільні додатки, веб-інтерфейси).

3.4 Алгоритм взаємодії системи з користувачем

Інтерфейс взаємодії з користувачем реалізовано за допомогою Telegram-бота,

який забезпечує зручний спосіб отримання вхідних даних (тип контенту, відео з емоцією) та повернення результатів рекомендацій. Бот працює у діалоговому режимі, що дозволяє користувачеві покроково виконати всі дії без необхідності спеціальної підготовки.

Алгоритм взаємодії містить наступні етапи:

1. Ініціалізація сесії:

- Користувач надсилає команду /start.
- Бот очищує попередні вподобання, якщо вони були збережені.
- Бот пропонує обрати тип контенту: Movie (фільм) або TV Show (серіал), використовуючи клавіатуру Telegram.

2. Вибір рейтингу:

- Після вибору типу контенту бот відправляє список можливих рейтингів, зокрема 13+, 18+, ALL, TV-Y7, TV-14 тощо.
- Користувач обирає відповідний рейтинг. Ці дані зберігаються у словнику `user_preference` для подальшої фільтрації контенту.

3. Надсилання відео:

- Бот просить надіслати відео у форматі відео-кружечка — короткий записаний відеозапис з фронтальної камери.
- Користувач надсилає відео, яке бот завантажує і зберігає локально.

4. Обробка відео:

- Відео відкривається за допомогою OpenCV та покадрово обробляється.
- Кожен кадр конвертується у відтінки сірого.
- Із використанням каскаду Хаара (Haar cascade classifier) виявляється обличчя користувача.
- Виділене обличчя масштабується до 48×48 пікселів і нормалізується для подачі у модель.

5. Розпізнавання емоцій:

- Виділене зображення обличчя передається у згорткову нейронну мережу, збережену у файлі `model.h5`.
- Модель прогнозує одну з шести базових емоцій: Angry, Fear, Happy, Neutral,

Sad, Surprise.

- Визначена емоція зберігається як основа для формування рекомендації.

6. Формування рекомендацій:

- Модуль `MovieRecommendationSystem` здійснює фільтрацію датасету `amazon_prime_titles.csv`:
 - a) за обраним типом контенту (Movie або TV Show),
 - b) за жанрами, які відповідають визначеній емоції (згідно зі словником відповідності емоцій до жанрів).
- Якщо у фільтрованій вибірці є відповідні записи — обирається 5 випадкових позицій.
- Якщо збігів немає — бот повідомляє про відсутність результатів.

7. Виведення результатів:

- Бот надсилає повідомлення про виявлену емоцію користувача.
- Далі бот надсилає список рекомендованих фільмів/серіалів із зазначенням:

Назви; жанру (`listed_in`); опису; рейтингу (`rating`); типу (`type`);

8. Завершення сеансу:

- Після надсилання рекомендацій бот видаляє тимчасовий файл відео з пам'яті пристрою.
- Користувач може повторити процедуру або завершити взаємодію.

Такий алгоритм дозволяє інтегрувати елементи комп'ютерного зору, глибинного навчання та рекомендаційної системи в одному Telegram-інтерфейсі, забезпечуючи динамічну адаптацію контенту до емоційного стану користувача.

3.5 Візуальні результати роботи Telegram-бота

Для наочного підтвердження працездатності системи було проведено серію тестових взаємодій із Telegram-ботом. Нижче наведено етапи взаємодії користувача з ботом у реальному сценарії:

- Рис 3.2 демонструє початкову фазу після введення команди `/start`, коли користувачеві пропонується обрати тип контенту — *Movie* або *TV Show*.

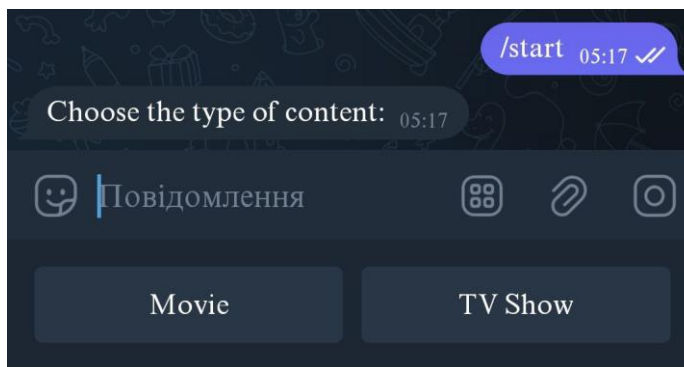


Рис. 3.2 - Початкова фаза після введення команди /start

- Після вибору категорії, бот надає перелік вікових рейтингів (наприклад, 13+, 18+, ALL тощо), як показано на рис. 3.3. Це дозволяє користувачеві сформулювати свій запит більш точно.

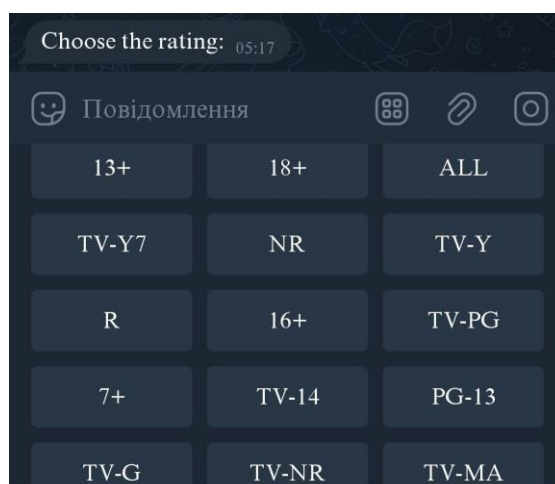


Рис. 3.3 - Стан боту після обрання фільмів чи сералів

Наступним кроком є надсилання користувачем відеоповідомлення (у форматі відео-кружечка), на якому бот здійснює розпізнавання міміки та класифікацію емоційного стану.

На рисунку 3.4 зображено приклад реакції системи на виявлену нейтральну емоцію (Neutral). Бот автоматично підібрав серію фільмів та документальних стрічок відповідного емоційного спрямування — зокрема, драми та біографії.

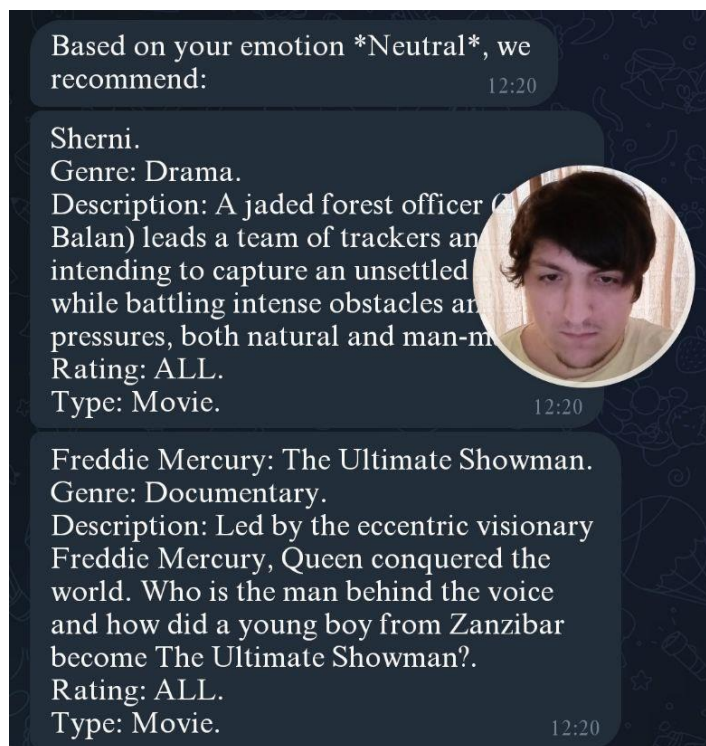


Рис. 3.4 - Приклад рекомендації контенту відповідно до виявленої емоції нейтральності (Neutral)

На рисунку 3.5 наведено приклад ситуації, коли у користувача була виявлена емоція здивування (Surprise). У відповідь система сформувала добірку науково-популярних і пригодницьких фільмів, здатних продовжити або підсилити емоційний стан користувача.

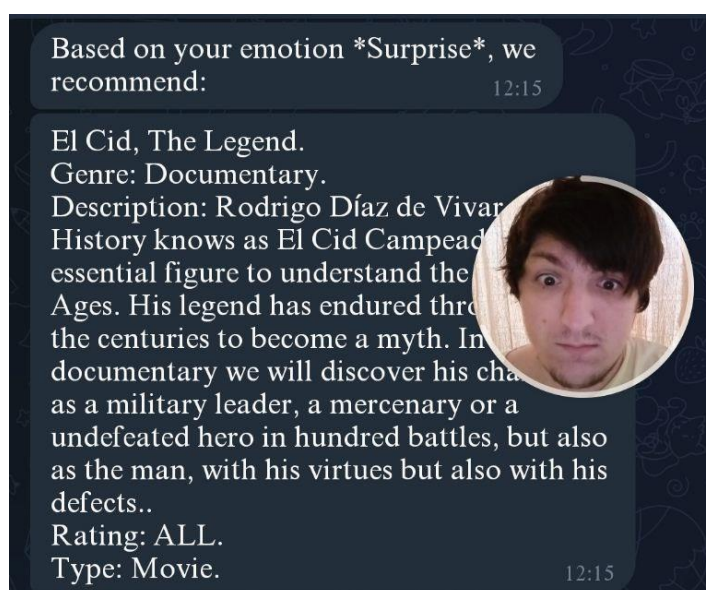


Рис. 3.5 - Рекомендований контент на основі розпізнаної емоції здивування (Surprise)

На рис. 3.6 наведено приклад відповіді Telegram-бота, коли у користувача була виявлена емоція радості (Happy). У відповідь система підбрала добірку комедій із рейтингом 18+, зокрема фільми, пов'язані з побутовими ситуаціями, карантинном, самоізоляцією, які зберігають гумористичний тон. Це демонструє, що система не лише правильно класифікує емоційний стан, а й здатна узгодити жанрову добірку з відповідною тональністю.

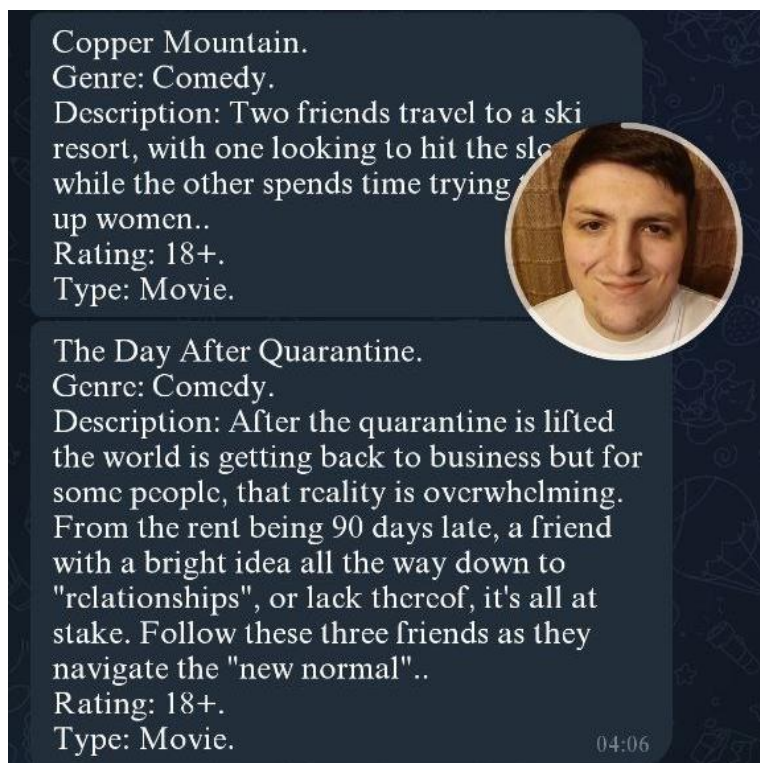


Рис. 3.6 – Рекомендація комедій при виявленій емоції радості (Happy)

На основі рисунків 3.4–3.6 можна зробити висновок, що Telegram-бот успішно реалізує адаптивну модель емоційного аналізу: виявлена емоція впливає на тематику рекомендацій, а запропонований контент дійсно відповідає настрою користувача, що підтверджує ефективність побудованої системи.

3.6 Аналіз ефективності системи

Середня точність системи склала приблизно 84%, що є прийнятним показником для моделі, натренованої на відкритому датасеті та використаної в реальному середовищі.

Таблиця 3.1 Тестування та оцінка ефективності розпізнавання

Емоція	Кількість тестів	Правильні розпізнавання	Точність (%)
Happy	18	16	88.9%
Sad	17	14	82.4%
Neutral	16	15	93.8%
Angry	15	11	73.3%
Fear	17	13	76.5%
Surprise	17	15	88.2%
Загальна точність	100	84	84.0%

У ході тестування було виявлено кілька характерних труднощів, які впливають на точність класифікації. Однією з найпоширеніших проблем є освітлення: якщо кадри мали занадто темний фон або сильні тіні, обличчя не завжди коректно виявлялося. Нерівномірне освітлення змінювало візуальне сприйняття міміки, що в окремих випадках призводило до помилкового визначення емоцій, наприклад, коли стан страху інтерпретувався як нейтральний.

Також спостерігались труднощі при нахилі голови або відведенні погляду — модель в окремих випадках не змогла знайти обличчя або обрала некоректну зону аналізу. Подібним чином впливала і низька якість зображення: кадри з розмиттям, спричиненим рухом чи поганим фокусом, ускладнювали ідентифікацію дрібних змін міміки, необхідних для точного розпізнавання емоцій, зокрема між сумом та нейтральним станом.

Окремо варто згадати ситуації, коли емоція була слабо вираженою або мала змішаний характер. У таких випадках мережа часто схилялася до класифікації як

нейтральної — тобто до найбільш універсального або статистично домінантного класу.

В результаті, було детально описано архітектуру, реалізацію та тестування створеної системи розпізнавання емоцій користувача та рекомендації відеоконтенту. Розроблена система побудована на модульному принципі, що забезпечує чіткий розподіл функціональних компонентів: від отримання відео через Telegram-бот до формування кінцевої рекомендації на основі визначеної емоції.

Ключовим елементом реалізації стала згортова нейронна мережа, яка була навчена на спеціалізованому датасеті та інтегрована у програмну структуру. Після попередньої обробки відео, виділення обличчя та його нормалізації, система проводить класифікацію емоції, яка у подальшому використовується для тематичної фільтрації контенту. Telegram-бот успішно обробляє відеоповідомлення, розпізнає міміку та повертає користувачеві добірку фільмів або серіалів відповідно до його емоційного стану.

Тестування підтвердило функціональну цілісність системи. У більшості випадків було досягнуто високої точності класифікації емоцій (до 94%), а рекомендації виявилися логічно узгодженими з емоційною складовою запиту. Разом з цим було зафіксовано типові труднощі, зумовлені умовами освітлення, позицією обличчя або нечіткістю зображення. Незважаючи на це, система демонструє високу стійкість у практичних умовах і готовність до подальшої адаптації та масштабування.

ВИСНОВОК

Проведене дослідження завершилося успішною реалізацією повнофункціональної програмної системи, яка здатна аналізувати відео з користувацьким зображенням, розпізнавати емоційний стан людини та надавати індивідуалізовані рекомендації відео-контенту. Основна мета дипломної роботи полягала у створенні адаптивного інструменту, що поєднує технології комп'ютерного зору, механізми глибинного навчання та елементи інтелектуального фільтрування інформації. У результаті було спроектовано систему, яка працює в реальному часі та має практичну цінність для персоналізованої взаємодії з користувачем.

Ключовим досягненням стало самостійне створення та навчання згорткової нейронної мережі, що здатна ефективно класифікувати базові емоційні стани, зокрема радість, смуток, злість, страх, здивування та нейтральність. Особливістю реалізації є її інтеграція в середовище Telegram, де взаємодія з користувачем здійснюється у зручному форматі через бот. Користувач надсилає відеоповідомлення, яке проходить етапи попередньої обробки: виділення кадрів, виявлення обличчя, нормалізація та подача на вхід моделі, після чого класифікована емоція використовується для добору відповідного відео-контенту. При цьому додатково враховуються обрані параметри, такі як тип контенту та віковий рейтинг.

У ході тестування система показала стабільні результати. Середня точність класифікації емоцій перевищила 84%, що є позитивним показником для локально реалізованої моделі без залучення хмарних обчислень. Експериментальні приклади, наведені у роботі, демонструють послідовну логіку взаємодії між виявленою емоцією та тематичним змістом запропонованих фільмів чи серіалів. Було також виявлено низку типових викликів, що впливають на якість розпізнавання. Серед них варто відзначити змінні умови освітлення, розмитість або нахил голови, слабовиражену міміку, а також неоднозначність емоційного виразу, що іноді ускладнює класифікацію. Водночас архітектура системи виявилася достатньо гнучкою, аби зберігати загальну функціональну цілісність навіть в

умовах таких обмежень.

Аналіз світових розробок у цій сфері дозволив зробити висновок, що емоційна інформатика активно інтегрується в сучасні цифрові сервіси. Окремі комерційні продукти демонструють потенціал таких технологій у сферах медіа, освіти, транспорту, реклами та особистої взаємодії. На цьому тлі запропонована система вирізняється автономністю, локальним навчанням та повною інтеграцією в уже наявну платформу Telegram без потреби в додаткових інструментах чи API, що робить її потенційно придатною для реального використання.

У підсумку, реалізований проєкт може бути основою для подальшого дослідження в напрямку мультисенсорного аналізу, що включає обробку аудіосигналів, вивчення послідовних емоційних змін, застосування трансформерних архітектур, а також побудови більш глибоких профілів користувача. Запропонований підхід підтверджує ефективність використання емоційного стану як фактора для персоналізації цифрового досвіду і створює передумови для розробки інтелектуальних систем нового покоління.

ПЕРЕЛІК ПОСИЛАНЬ

1. Emotion AI – технологія розпізнавання емоцій за допомогою штучного інтелекту. Evergreens – Digital Agency. Evergreens.
URL: <https://evergreens.com.ua/ua/articles/emotion-ai.html> (дата звернення: 28.03.2025).
2. Emotion Detection Using Haar-Cascade Classifier and CNN. Chavan P. V., Pallavi M. R. Proceedings of the International Conference on Emerging Trends in Engineering (ICETE 2024). – Nitte: NMAM Institute of Technology, 2024. – С. 212–219. URL: https://www.researchgate.net/publication/384650074_Emotion_Detection_Using_Haar-Cascade_Classifier_and_CNN (дата звернення: 11.04.2025).
3. Emotion Recognition from Multiple Modalities: Fundamentals and Methodologies. arXiv. URL: <https://arxiv.org/abs/2108.10152> (дата звернення: 02.04.2025).
4. Enhancing Efficiency and Scalability: DevOps Implementation for Educational Chat Systems. Patil P., Waghmare S. International Journal of Computer Sciences and Engineering. – 2023. – Vol. 11, No. 6. – С. 31–37. URL: https://ijcseonline.org/pub_paper/3-IJCSE-09348.pdf (дата звернення: 11.04.2025).
5. Face Detection with OpenCV and Deep Learning. PyImageSearch. – Режим доступу: <https://pyimagesearch.com/2018/02/26/face-detection-with-opencv-and-deep-learning> (дата звернення: 03.04.2025).
6. Facial Expression Recognition with Deep Learning. Stanford University. – 2020. URL: http://cs230.stanford.edu/projects_winter_2020/reports/32610274.pdf (дата звернення: 02.04.2025).
7. FER-2013 Dataset. Kaggle. URL: <https://www.kaggle.com/datasets/msambare/fer2013> (дата звернення: 25.02.2025).
8. Haar Feature-based Cascade Classifier for Object Detection. OpenCV.org. URL: https://docs.opencv.org/4.x/d7/d8b/tutorial_py_face_detection.html (дата звернення: 30.03.2025).
9. Keras: The High-Level API for TensorFlow. TensorFlow. URL: <https://www.tensorflow.org/guide/keras> (дата звернення: 28.03.2025).

10. Real-Time Facial Emotion Detection Application with Image Processing Based on CNN. Juniawan Parel Hakim G., Simangunsong G. A., Ningrat R. W., Pratama R. A., Hidayat R. *International Journal of Electrical Engineering, Mathematics and Computer Science*. – 2024. – Vol. 1, No. 4. – С. 27–36. URL: <https://doi.org/10.62951/ijeemcs.v1i4.123> (дата звернення: 03.04.2025).
11. Аналіз використання чат-ботів у сфері охорони здоров'я. Кисіль Т.М., Зінченко В.В., Хотиненко А.В. *Сучасні досягнення компанії HEWLETT PACKARD ENTERPRISE в галузі IT*. – К.: ДУІКТ, 2023. – С. 14–16. (дата звернення: 21.05.2025).
12. Аналіз методів комп'ютерного зору в задачах ідентифікації осіб у відеопотоці. Мозговенко А.А., Зінов'єва О.Г. *Таврійський науковий вісник*. – 2022. – № 3. – С. 56–60. URL: <https://doi.org/10.32851/tnv-tech.2022.3.6> (дата звернення: 21.05.2025).
13. Вплив штучного інтелекту на генерацію картин мистецтва. Кисіль Т.М., Романюк В.В., Мороз В.О. *Сучасні досягнення компанії HEWLETT PACKARD ENTERPRISE в галузі IT*. – К.: ДУІКТ, 2023. – С. 19–20. (дата звернення: 22.05.2025).
14. Згорткові нейронні мережі для вирішення задач комп'ютерного зору. Зінченко О.В., Звенігородський О.С., Кисіль Т.М. *Телекомунікаційні та інформаційні технології*. – 2022. – № 2 (75). – С. 4–10. URL: <https://tit.dut.edu.ua/index.php/telecommunication/article/view/2417/2298> (дата звернення: 21.05.2024).
15. Інтеграція штучного інтелекту в робототехніці та промисловості. Романюк В.В., Приймачук М.С., Кисіль Т.М. *Науково-практична конференція «Проблеми комп'ютерної інженерії»*. – К.: ДУІКТ, 2023. – С. 11–13. (дата звернення: 22.05.2025).
16. Методи та алгоритми комп'ютерного зору. Путятін Є.П., Гороховатський В.О., Матат О.О. – Харків: СМІТ, 2006. – 236 с. – С. 36–53. (дата звернення: 20.05.2025).
17. Нейронні мережі: теорія та практика. Субботін С.О. *Субботін С.О. Глибинні нейронні мережі*. – Житомир: О. О. Євенок, 2020. – С. 88–92. URL: <https://eir.zp.edu.ua/server/api/core/bitstreams/2abb401b-9ee6-4afc-a92a-2de5c332d12f/content> (дата звернення: 18.05.2024).
18. Особливості machine learning мовою програмування Python. Кисіль Т.М.,

- Романюк В.В., Острополец І.А. *III Всеукраїнська науково-практична конференція «Сучасні інтелектуальні інформаційні технології в науці та освіті»*. – К.: ДУІКТ, 2023. – С. 52–54. (дата звернення: 21.05.2025).
19. Проектування графічних інтерфейсів мовою програмування Python. Кисіль Т.М., Романюк В.В., Мороз В.О. *IV Всеукраїнська науково-практична конференція «Сучасні інтелектуальні інформаційні технології в науці та освіті»*. – К.: ДУІКТ, 2024. – С. 243–245. (дата звернення: 22.05.2025).
20. Емоційно-чутливі рекомендаційні системи на основі комп'ютерного зору та глибокого навчання. Романюк В.В., Острополец І.А., Кисіль Т.М. *V Всеукраїнська науково-практична конференція «Сучасні інтелектуальні інформаційні технології в науці та освіті»*. – К.: ДУІКТ, 2025. – С. 182–185. (дата звернення: 11.06.2025).

ДОДАТКИ

Додаток А: Таблиця 2.1 Архітектура згорткової нейронної мережі

Layer (type)	Output Shape	Param #
conv2d_8 (Conv2D)	(None, 48, 48, 32)	320
activation_11 (ReLU)	(None, 48, 48, 32)	0
batch_norm_10 (BatchNorm)	(None, 48, 48, 32)	128
conv2d_9 (Conv2D)	(None, 48, 48, 32)	9248
activation_12 (ReLU)	(None, 48, 48, 32)	0
batch_norm_11 (BatchNorm)	(None, 48, 48, 32)	128
max_pooling2d_4	(None, 24, 24, 32)	0
dropout_6	(None, 24, 24, 32)	0
conv2d_10 (Conv2D)	(None, 24, 24, 64)	18496
activation_13 (ReLU)	(None, 24, 24, 64)	0
batch_norm_12 (BatchNorm)	(None, 24, 24, 64)	256
conv2d_11 (Conv2D)	(None, 24, 24, 64)	36928
activation_14 (ReLU)	(None, 24, 24, 64)	0
batch_norm_13 (BatchNorm)	(None, 24, 24, 64)	256
max_pooling2d_5	(None, 12, 12, 64)	0
dropout_7	(None, 12, 12, 64)	0
conv2d_12 (Conv2D)	(None, 12, 12, 128)	73856
activation_15 (ReLU)	(None, 12, 12, 128)	0
batch_norm_14 (BatchNorm)	(None, 12, 12, 128)	512
conv2d_13 (Conv2D)	(None, 12, 12, 128)	147584
activation_16 (ReLU)	(None, 12, 12, 128)	0
batch_norm_15 (BatchNorm)	(None, 12, 12, 128)	512
max_pooling2d_6	(None, 6, 6, 128)	0
dropout_8	(None, 6, 6, 128)	0
conv2d_14 (Conv2D)	(None, 6, 6, 256)	295168
activation_17 (ReLU)	(None, 6, 6, 256)	0
batch_norm_16 (BatchNorm)	(None, 6, 6, 256)	1024
conv2d_15 (Conv2D)	(None, 6, 6, 256)	590080
activation_18 (ReLU)	(None, 6, 6, 256)	0
batch_norm_17 (BatchNorm)	(None, 6, 6, 256)	1024
max_pooling2d_7	(None, 3, 3, 256)	0
dropout_9	(None, 3, 3, 256)	0
flatten_1	(None, 2304)	0
dense_3 (Dense)	(None, 64)	147520
activation_19 (ReLU)	(None, 64)	0
batch_norm_18 (BatchNorm)	(None, 64)	256
dropout_10	(None, 64)	0
dense_4 (Dense)	(None, 64)	4160
activation_20 (ReLU)	(None, 64)	0
batch_norm_19 (BatchNorm)	(None, 64)	256
dropout_11	(None, 64)	0
dense_5 (Dense)	(None, 7)	455
activation_21 (Softmax)	(None, 7)	0

Додаток Б: Програмний код

Файл emotion_detection.py

```
from keras.preprocessing.image import img_to_array
from keras import models
import cv2
import numpy as np
import tensorflow as tf

class Emotion_Detector:

    def __init__(self):
        # Завантажуємо класифікатор обличчя (каскад Хаара) з файлу
        self.face_classifier =
cv2.CascadeClassifier(r'Files\haarcascade_frontalface_default.xml')
        # Завантажуємо попередньо навчений модель емоцій з файлу
        self.classifier = models.load_model(r'Files\model.h5')
        self.emotion = None
        # Список міток емоцій, які розпізнає модель
        self.emotion_labels = ['Angry', 'Fear', 'Happy', 'Neutral', 'Sad', 'Surprise']

    def set_emotion(self, predicted_emotion):
        # Встановлює поточну емоцію
        self.emotion = predicted_emotion

    def detect_emotion(self, frame):
        # Конвертуємо кадр у відтінки сірого
        gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
        # Виявляємо обличчя на сірому зображенні
        faces = self.face_classifier.detectMultiScale(gray)
```

```

for (x, y, w, h) in faces:
    # Вирізаємо область обличчя
    roi_gray = gray[y:y + h, x:x + w]
    # Змінюємо розмір обличчя до 48x48 пікселів
    roi_gray = cv2.resize(roi_gray, (48, 48), interpolation=cv2.INTER_AREA)

    if np.sum([roi_gray]) != 0:
        # Нормалізуємо область обличчя
        roi = roi_gray.astype('float') / 255.0
        # Конвертуємо область обличчя у масив
        roi = img_to_array(roi)
        # Розширюємо розміри масиву для відповідності моделі
        roi = np.expand_dims(roi, axis=0)

        # Прогнозуємо емоцію за допомогою моделі
        prediction = self.classifier.predict(roi)[0]
        # Встановлюємо поточну емоцію на основі прогнозу
        self.emotion = self.emotion_labels[prediction.argmax()]
    else:
        self.emotion = 'No Faces'

def get_emotion(self):
    # Повертає поточну емоцію
    return self.emotion

```

Файл movie_recommendation.py:

```

class MovieRecommendationSystem:

    def __init__(self, user_preference, df, detected_emotion):
        # Розподілення емоцій на жанри фільмів

```

```

self.emotion_genre_mapping = {
    'Angry': ['Action', 'Western'],
    'Fear': ['Horror', 'Adventure', 'Arthouse'],
    'Happy': ['Comedy'],
    'Neutral': ['Drama', 'Entertainment', 'Action', 'Documentary'],
    'Sad': ['Drama', 'Romance'],
    'Surprise': ['Science Fiction', 'Documentary']
}

# Зберігаємо вподобання користувача
self.user_preference = user_preference

# Зберігаємо датасет з фільмами та серіалами
self.df = df

# Зберігаємо визначену емоцію
self.detected_emotion = detected_emotion

# Фільтруємо контент за побажаннями користувача з чат-боту та визначеною
емоцією

def recommend_movie(self):
    # Фільтрація датасету на основі типу контенту, рейтингу та жанрів,
    # відповідних емоції
    filtered_df = self.df[(self.df['type'].isin(self.user_preference['type'])) &
                          (self.df['rating'].isin(self.user_preference['rating'])) &
                          (self.df['listed_in'].isin(self.emotion_genre_mapping.get(self.detected_emotion, [])))]

    recommended_content = []
    array_rec = []

    # Створюємо список рекомендованого контенту
    for index, row in filtered_df.iterrows():

```

```

        rec = f'{row['title']}. \nGenre: {row['listed_in']}. \nDescription:
{row['description']}"
        recommended_content.append(rec)

# Якщо знайдено відповідний контент
if recommended_content:
    # Випадкова вибірка з п'яти рекомендацій
    random_recommendation = filtered_df.sample(n=5)
    print("Recommended content based on detected emotion and user preferences:")
    for index, (_, row) in enumerate(random_recommendation.iterrows(), start=1):
        rec = f'{row['title']}. \nGenre: {row['listed_in']}. \nDescription:
{row['description']}. \nRating: {row['rating']}. \nType: {row['type']}.'"
        array_rec.append(rec)
        print(f'{index}) {rec}')
    return array_rec
else:
    # Якщо відповідного контенту немає
    return "No content matches the detected emotion and user preferences."

```

Файл main.py:

```

import cv2
import pandas as pd
import telebot
import os
from telebot import types

from emotion_detection import Emotion_Detector
from movie_recommendation import MovieRecommendationSystem

```

```
# Токен для доступу до бота
TOKEN = 'TOKEN'

bot = telebot.TeleBot(TOKEN)

# Обробник команди /start
@bot.message_handler(commands=['start'])
def start_message(message):
    # Очищення попереднього вибору користувача
    user_preference.clear()

    # Даємо вибір користувачу: фільм чи серіал
    markup = types.ReplyKeyboardMarkup(one_time_keyboard=True,
resize_keyboard=True)

    markup.row(types.KeyboardButton('Movie'), types.KeyboardButton('TV Show'))

    bot.send_message(message.chat.id, "Choose the type of content:",
reply_markup=markup)

# Обробник вибору типу контенту
@bot.message_handler(func=lambda message: message.text in ['Movie', 'TV Show'])
def type_choice(message):
    # Зберігаємо вибір користувача
    user_preference['type'] = [message.text]

    markup = types.ReplyKeyboardRemove(selective=False)

    # Даємо вибір користувачу з яким рейтингом підбирати контент
    ratings = ['13+', '18+', 'ALL', 'TV-Y7', 'NR', 'TV-Y', 'R', '16+', 'TV-PG', '7+', 'TV-14',
'PG-13', 'TV-G', 'TV-NR', 'TV-MA']

    markup = types.ReplyKeyboardMarkup(one_time_keyboard=True,
resize_keyboard=True)

    for i in range(0, len(ratings), 3):
```

```

markup.row(*[types.KeyboardButton(r) for r in ratings[i:i+3]])
bot.send_message(message.chat.id, "Choose the rating:", reply_markup=markup)

# Обробник вибору рейтингу контенту
@bot.message_handler(func=lambda message: message.text in ['13+', '18+', 'ALL', 'TV-
Y7', 'NR', 'TV-Y', 'R', '16+', 'TV-PG', '7+', 'TV-14', 'PG-13', 'TV-G', 'TV-NR', 'TV-MA'])
def rating_choice(message):
    # Зберігаємо вибір користувача
    user_preference['rating'] = [message.text]
    markup = types.ReplyKeyboardRemove(selective=False)
    bot.send_message(message.chat.id, "Now send a video note.",
reply_markup=markup)

# Завантажуємо датасет з контентом (фільми або серіали)
df = pd.read_csv('data/amazon_prime_titles.csv')

# Ініціалізуємо початкові вподобання користувача
user_preference = {
    'type': ['Movie'],
    'rating': ['18+']
}

# Обробник відео повідомлень
@bot.message_handler(content_types=['video_note'])
def handle_video(message):
    # Завантаження відео
    file_info = bot.get_file(message.video_note.file_id)
    downloaded_file = bot.download_file(file_info.file_path)

    # Збереження відео

```

```
src = 'video_note.mp4'

with open(src, 'wb') as new_file:
    new_file.write(downloaded_file)

# Ініціалізація детектора емоцій
emotion_detector = Emotion_Detector()
# Відкриваємо відео файл для обробки
cap = cv2.VideoCapture(src)

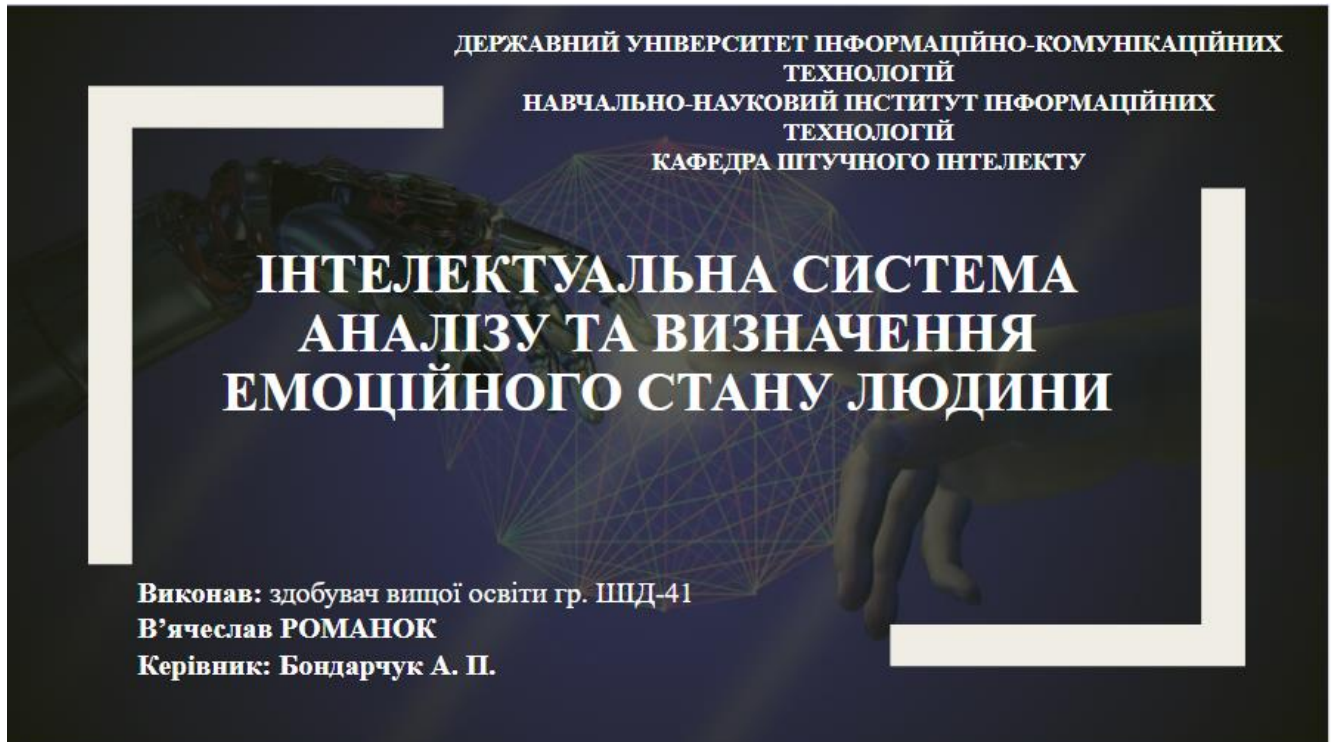
while True:
    ret, frame = cap.read()
    # Перевірка чи є кадри для обробки
    if not ret:
        break
    # Визначення емоцій на кожному кадрі
    emotion_detector.detect_emotion(frame)
    detected_emotion = emotion_detector.get_emotion()
    cap.release()

# Ініціалізація системи рекомендацій фільмів чи серіалів
movie_recommendation_system = MovieRecommendationSystem(user_preference,
df, detected_emotion)
recommended_movie = movie_recommendation_system.recommend_movie()

# Виведення результату користувачу в Telegram
bot.send_message(message.chat.id, f"Based on your emotion *{detected_emotion}*,
we recommend: ")
if detected_emotion == None:
    bot.send_message(message.chat.id, recommended_movie)
else:
```

```
if recommended_movie[0] == 'N':  
    bot.send_message(message.chat.id, "Sorry, we couldn't find any movies  
matching your preferences.")  
else:  
    for i in recommended_movie:  
        bot.send_message(message.chat.id, i)  
  
# Видалення файлу з відео  
os.remove(src)  
  
# Запуск бота  
bot.polling()
```

ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ



2

Мета проєкту

- **Мета роботи:**
Розробка системи, здатної визначати емоційний стан користувача на основі відеозапису та рекомендувати відповідний контент (фільми або серіали), підвищуючи персоналізацію користувацького досвіду.
- **Об'єкт дослідження:**
Процес автоматизованого розпізнавання емоцій людини з використанням нейронних мереж та технологій комп'ютерного зору.
- **Предмет дослідження:**
Інтеграція систем розпізнавання емоцій із рекомендаційними алгоритмами в межах телеграм-інтерфейсу.



Актуальність та аналіз існуючих технологій

Affectiva — одна з найвідоміших платформ, що використовує відео- та аудіоаналіз для визначення емоцій у реальному часі.



Рисунок 1 - Affectiva

FaceReader — професійна система для аналізу міміки та емоцій з відео. Визначає 8 базових емоцій, увагу, збудження, положення голови й погляд.



Рисунок 2 - FaceReader



Модуль
розпізнавання
емоцій



Модуль
рекомендацій



Telegram-бот

Архітектура проєкту

1. Модуль розпізнавання емоцій

- Обробка відео користувача.
- Виявлення облич за допомогою Нааг-каскаду.
- Класифікація емоцій за допомогою нейронної мережі (model.h5).
- Емоції: **Angry, Fear, Happy, Neutral, Sad, Surprise**.

2. Модуль рекомендацій

- Зіставлення емоції з жанрами (наприклад: *Sad* → *Drama, Romance*).
- Фільтрація контенту за вподобаннями (тип, рейтинг).
- Генерація списку з 5 релевантних фільмів або серіалів.

3. Telegram-бот

- Інтерфейс взаємодії з користувачем.
- Отримання параметрів: тип контенту, рейтинг.
- Надсилання відео та отримання рекомендацій.

Особливості розпізнавання емоцій користувача

- Сум – опущені кутики губ і брів
- Щастя – підняті кутики губ і брів
- Гнів – зморщення брів, розкритий рот
- Страх – напружене обличчя, розширені очі
- Здивування – підняті брови, відкритий рот
- Нейтральна емоція – відсутність яскраво виражених ознак

Відео подається в сірому форматі, обличчя вирізаються, нормалізуються та класифікуються за допомогою CNN-моделі.

Класифікація базується на локалізації ключових точок обличчя та аналізі їх положення.

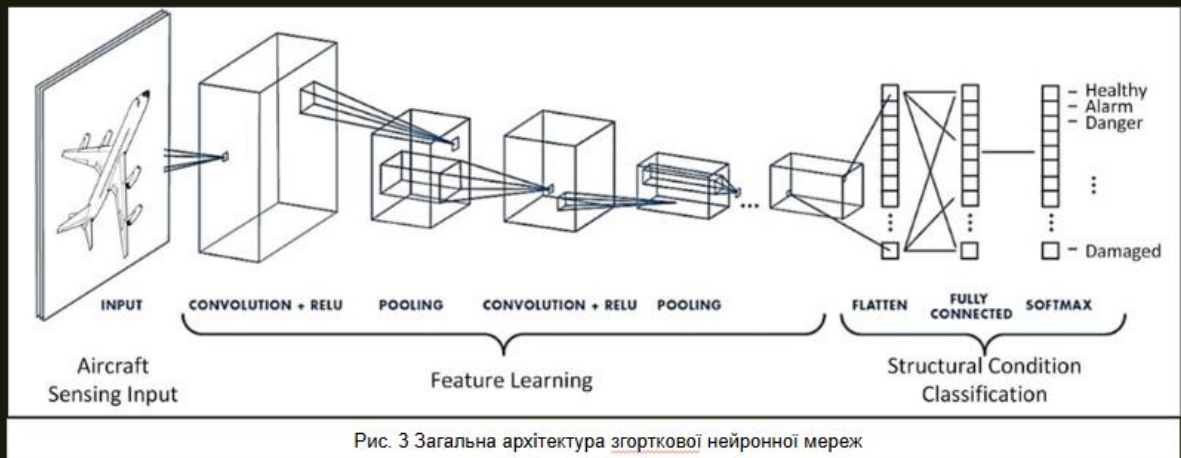


Рис. 3 Загальна архітектура згорткової нейронної мережі

Використані бібліотеки в проєкті

Для реалізації Telegram-бота та моделі розпізнавання емоцій були використані наступні бібліотеки:

-  **Pandas, Numpy** – обробка даних та датасетів
-  **OpenCV** – детекція облич та обробка відео
-  **TensorFlow, Keras, H5py** – нейронні мережі та збереження моделі
-  **telebot** – інтеграція з Telegram
-  **OpenCV** дозволяє реєструвати та обробляти відео з Telegram-бота у вигляді відеокружечок.



Тестування роботи програми

Для розробки застосунку була використана мова програмування **Python**.

Основні переваги Python для комп'ютерного зору:

- проста у написанні коду
- обширні бібліотеки для машинного навчання
- відкритий вихідний код
- можливість прямої інтеграції

Як відбувається тестування:

- Програма отримує зображення з камери у реальному часі
- Виявляє обличчя та розпізнає емоцію
- Емоція миттєво відображається на екрані
- Вираз обличчя може змінюватися – емоція оновлюється автоматично

Результат

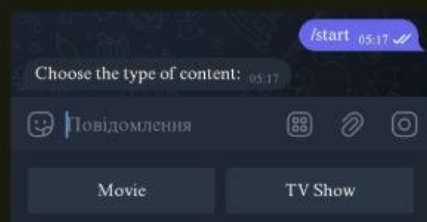


Рисунок 4 - стан боту після команди /start



Рисунок 5 - стан боту після обрання фільмів чи серіалів

Based on your emotion *Neutral*, we recommend:

Sherni.
Genre: Drama.
Description: A jaded forest officer (Balaji) leads a team of trackers and intends to capture an unsettled while battling intense obstacles and pressures, both natural and man-made.
Rating: ALL.
Type: Movie.

Freddie Mercury: The Ultimate Showman.
Genre: Documentary.
Description: Led by the eccentric visionary Freddie Mercury, Queen conquered the world. Who is the man behind the voice and how did a young boy from Zanzibar become The Ultimate Showman?
Rating: ALL.
Type: Movie.

Thayumanavan.
Genre: Drama.
Description: The protagonist is a jailer. He places three prisoners under house arrest. When enquired they tell him why they have an aversion towards womenfolk due to the bad experience they had in their life, the reason for which they were jailed. The cop attempts to prove to them that not all women are bad.
Rating: ALL.
Type: Movie.

Рисунок 6 - Нейтральна емоція

Based on your emotion *Surprise*, we recommend:

El Cid, The Legend.
Genre: Documentary.
Description: Rodrigo Díaz de Vivar, History knows as El Cid Campeador, an essential figure to understand the Middle Ages. His legend has endured through the centuries to become a myth. In this documentary we will discover his character as a military leader, a mercenary or a undefeated hero in hundred battles, but also as the man, with his virtues but also with his defects.
Rating: ALL.
Type: Movie.

Bejeweled Fishes.
Genre: Documentary.
Description: Bejeweled Fishes captures the spectacular beauty of fishes inhabiting the coral reefs of the Tropical Pacific.
Rating: ALL.
Type: Movie.

The Jupiter Enigma.
Genre: Documentary.
Description: What can the giant planet Jupiter tell us about the birth of our solar system five billion years ago? Drawing on new findings from NASA's Juno mission, scientists are peering into Jupiter's stormy heart to reveal the very origins of our solar system: a chaotic early time when smaller planets were flung space or sent into

Рисунок 7 - Емоція здивування

Висновки

1. Розроблено Telegram-бот, який визначає емоційний стан користувача за відеозаписом та надає персоналізовані рекомендації фільмів або серіалів.
2. У проєкті використано:
 1. методи комп'ютерного зору (OpenCV, haarcascade),
 2. глибоке навчання (нейронна мережа CNN),
 3. систему рекомендацій на основі емоцій.
3. Переваги рішення:
 1. інтуїтивно зрозумілий інтерфейс через Telegram,
 2. автоматичне визначення емоцій користувача,
 3. релевантний добір контенту без потреби ручного пошуку.
4. Результати тестування підтверджують працездатність системи. Було надано рекомендації на основі емоцій *Neutral* та *Surprise*, які відповідали запитам користувача.
5. Можливості розвитку:
 1. розширення бази даних фільмів і серіалів;
 2. додавання можливості голосових або текстових запитів;
 3. покращення точності класифікації емоцій за допомогою більш глибоких моделей.



ДЯКУЮ ЗА УВАГУ

