

**ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ**

КВАЛІФІКАЦІЙНА РОБОТА
на тему: «МОДУЛЬ РОЗПІЗНАВАННЯ ЕМОЦІЙНИХ СТАНІВ
ЛЮДИНИ НА ОСНОВІ ШІ»

на здобуття освітнього ступеня магістра
зі спеціальності 123 Комп'ютерної інженерії
(код, найменування спеціальності)
освітньо-професійної програми Комп'ютерні системи та мережі
(назва)

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

(підпис)	Владислав ДРАГУН Ім'я, ПРІЗВИЩЕ здобувача

Виконав:	здобувач(ка) вищої освіти гр. <u>КСДМ-61</u> <u>Владислав ДРАГУН</u> Ім'я, ПРІЗВИЩЕ
Керівник: науковий ступінь, вчене звання	<u>Анастасія ВСЧЕРКОВСЬКА</u> Ім'я, ПРІЗВИЩЕ
Рецензент: науковий ступінь, вчене звання	 Ім'я, ПРІЗВИЩЕ

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

Кафедра Комп'ютерної інженерії

Ступінь вищої освіти Магістр

Спеціальність 123 «Комп'ютерна інженерія»

Освітньо-професійна програма Комп'ютерні системи та мережі

ЗАТВЕРДЖУЮ

Завідувач кафедри Наталія ЛАЩЕВСЬКА

_____ Ім'я, ПРІЗВИЩЕ
«_____» _____ 2024р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

Драгун Владислав Петрович

(прізвище, ім'я, по батькові здобувача)

1. Тема кваліфікаційної роботи: «Модуль розпізнавання емоційних станів людини на основі ІІІ»

керівник кваліфікаційної роботи Анастасія ВСЧЕРКОВСЬКА, канд.техн.наук, доцент,
(Ім'я, ПРІЗВИЩЕ, науковий ступінь, вчене звання)

затверджені наказом Державного університету інформаційно-комунікаційних технологій від
«15» жовтня 2024р. №320,

2. Строк подання кваліфікаційної роботи «29» грудня 2024р.

3. Вихідні дані до кваліфікаційної роботи: Наукова та технічна література, робочі параметри структур машинного навчання, принципи попередньої обробки та доповнення даних, методи налаштування автоматизації, вимоги до етичного штучного інтелекту та безпеки, критерії для порівняння існуючих систем розпізнавання емоцій.

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

1. Основи та значення розпізнавання емоцій

2. Технології та методи розпізнавання емоцій

3. Розробка та вдосконалення модуля розпізнавання емоцій

5. Перелік ілюстративного матеріалу: презентація

1. Мета, об'єкт, предмет дослідження

2. Актуальність роботи

3. Проблематика та необхідність впровадження систем розпізнавання емоцій

4. Архітектура запропонованого рішення у вирішенні та покращенні взаємодії «людина-комп'ютер»

5. Блок-схема роботи моделі в різних сценаріях: промисловість, транспорт, медицина, громадські місця

6. Віртуальне середовище та інструмент для розробки моделі розпізнавання емоційних станів людини

7. Бібліотеки та залежності для розробки моделі розпізнавання емоційних станів людини

8. Обраний набір даних та попередньо натренована модель

9. Процес тренування моделі розпізнавання емоційних станів людини
 10. Результат натренованої моделі (через зображення)
 11. Результат натренованої моделі (через веб-камеру)
 12. Висновки
 13. Продовження висновків
 14. Публікації та апробація роботи
6. Дата видачі завдання «01» вересня 2024р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1	Збір даних	10.10-12.10.2024	
2	Обробка матеріалу	12.10-15.10.2024	
3	Розробка теоретичної частини	15.10-18.10.2024	
4	Технології та методи розпізнавання емоцій	18.10-22.10.2024	
5	Розробка та вдосконалення модуля розпізнавання емоцій	22.10-26.11.2024	
6	Висновки за результатами аналізу	29.11-30.11.2024	
7	Розробка презентації	30.11-08.12.2024	
8	Передзахист	09.12-11.12.2024	
9	Здача в деканат	20.12-29.12.2024	

Здобувач вищої освіти

_____ **Владислав ДРАГУН**
(підпис) (Ім'я, ПРІЗВИЩЕ)

Керівник кваліфікаційної роботи

_____ **Анастасія ВСЧЕРКОВСЬКА**
(підпис) (Ім'я, ПРІЗВИЩЕ)

РЕФЕРАТ

Текстова частина кваліфікаційної роботи на здобуття освітнього ступеня магістр: 154 стор., 60 рис., 2 табл., 69 джерел.

Мета роботи – розробити модуль розпізнавання емоційних станів людини на основі технологій штучного інтелекту для покращення комунікації між користувачем та комп'ютером.

Об'єкт дослідження – емоційні стани людини.

Предмет дослідження – процес розпізнавання та аналізу емоційних станів людини за допомогою методів штучного інтелекту.

Короткий зміст роботи: У роботі описано розробку модуля для розпізнавання людських емоцій та його потенціал у покращенні взаємодії в цифровому світі. Використовуючи алгоритми штучного інтелекту та машинного навчання, досліджено методи визначення емоцій за мімікою, голосом та іншими біометричними даними.

Архітектура модуля забезпечує високу точність і швидкість роботи, а результати демонструють перспективи застосування технології в психологічному консультуванні та інтуїтивних взаємодіях людини з машиною. У висновках запропоновано рекомендації для розвитку систем із розширенням діапазону розпізнаваних емоцій.

КЛЮЧОВІ СЛОВА: РОЗПІЗНАВАННЯ ЕМОЦІЙ, ШТУЧНИЙ ІНТЕЛЕКТ, МАШИННЕ НАВЧАННЯ, БІОМЕТРИЧНІ ДАНІ, МІМІКА, ІНТОНАЦІЯ ГОЛОСУ, ПСИХОЛОГІЧНЕ КОНСУЛЬТУВАННЯ, ІНТЕРАКТИВНІ СИСТЕМИ, ВЗАЄМОДІЯ ЛЮДИНА-МАШИНА, ТОЧНІСТЬ РОЗПІЗНАВАННЯ, ШВИДКІСТЬ РОЗПІЗНАВАННЯ, СТРУКТУРА МОДУЛЯ, АРХІТЕКТУРА МОДУЛЯ, АНАЛІЗ ДАНИХ, ТЕСТУВАННЯ МОДЕЛІ, ІНТЕГРАЦІЯ ТЕХНОЛОГІЙ, ЕМОЦІЙНІ СТАНИ.

ABSTRACT

The text part of the qualification work for obtaining a master's degree: 154 pages, 60 figures, 2 tables, 69 sources.

The purpose of the work is to develop a module for recognizing human emotional states based on artificial intelligence technologies to improve communication between the user and the computer.

The object of the research is the emotional states of a person.

The subject of research is the process of recognizing and analyzing human emotional states using artificial intelligence methods.

Summary of the work: The paper describes the development of a module for recognizing human emotions and its potential for improving interaction in the digital world. Using artificial intelligence and machine learning algorithms, methods for determining emotions from facial expressions, voice, and other biometric data have been investigated.

The architecture of the module provides high accuracy and speed of operation, and the results demonstrate the prospects for using the technology in psychological counseling and intuitive human-machine interactions. The conclusions offer recommendations for the development of systems with an expanded range of recognized emotions.

KEY WORDS: EMOTION RECOGNITION, ARTIFICIAL INTELLIGENCE, MACHINE LEARNING, BIOMETRIC DATA, MIMICALS, VOICE INTONATION, PSYCHOLOGICAL COUNSELING, INTERACTIVE SYSTEMS, HUMAN-MACHINE INTERACTION, ACCURACY RECOGNIZED NO, RECOGNITION SPEED, MODULE STRUCTURE, MODULE ARCHITECTURE, DATA ANALYSIS, MODEL TESTING, INTEGRATION TECHNOLOGY, EMOTIONAL STATES.

ЗМІСТ

ВСТУП.....	9
РОЗДІЛ 1 ОСНОВИ ТА ЗНАЧЕННЯ РОЗПІЗНАВАННЯ ЕМОЦІЙ.....	11
1.1 Поняття емоційного стану та його визначення.....	11
1.2 Еволюція підходів до аналізу емоційного стану.....	25
1.3 Важливість розпізнавання емоцій в сучасних технологіях	31
1.4 Соціальні та психологічні аспекти емоційного аналізу	39
1.5 Основні виклики у вивченні емоцій та їх обробці.....	47
РОЗДІЛ 2 ТЕХНОЛОГІЇ ТА МЕТОДИ РОЗПІЗНАВАННЯ ЕМОЦІЙ.....	56
2.1 Штучний інтелект та машинне навчання для розпізнавання емоцій	56
2.2 Комп'ютерний зір та аналіз міміки і жестів.....	64
2.3 Аналіз голосу та мовлення у визначенні емоцій	70
2.4 Порівняння алгоритмів і підходів до розпізнавання емоцій	79
РОЗДІЛ 3 РОЗРОБКА ТА ВДОСКОНАЛЕННЯ МОДУЛЯ РОЗПІЗНАВАННЯ ЕМОЦІЙ.....	90
3.1 Конфігурація середовища та засоби розробки.....	90
3.2 Підготовка набору даних і методи попередньої обробки	97
3.3 Розробка моделі та навчання через трансферне навчання.....	105
3.4 Виявлення та візуалізація емоцій у реальному часі	110
3.5 Перспективи розвитку, етичні аспекти та аспекти безпеки.....	117
ВИСНОВКИ.....	131
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	133
Додаток А. Порівняльний аналіз алгоритмів і підходів до розпізнавання емоцій.....	139
Додаток Б. Конфігурація середовища та засобів розробки.....	140
Додаток В. Підготовка набору даних для тренування моделі.....	143
Додаток Г. Розроблений модуль розпізнавання емоційних станів людини (Python).....	144
Додаток Д. ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ (Презентація).....	147

ВСТУП

Штучному інтелекту дуже важливо розуміти глибину та різноманітність людських емоцій, оскільки люди переживають широкий спектр емоційних станів: від гніву до радості, від смутку до хвилювання, які динамічні та постійно змінюються. На відміну від простих, статичних висловлювань, людські емоції охоплюють багаті, складні реакції, які штучний інтелект повинен досягнути, щоб справді взаємодіяти так, щоб відчувати себе людським.

Наразі ШІ не має повністю людського розуміння емоцій. Більшість сучасних моделей штучного інтелекту інтерпретують емоції як суто точковий набір даних або його текстове представлення, у яких відсутні нюанси справжнього «людського». Однак передова система розпізнавання емоцій може дозволити штучному інтелекту інтерпретувати тонкі сигнали, такі як вираз обличчя, жести та голосовий тон, так само, як це роблять ми, люди. Наприклад, аналізуючи рухи обличчя та голосові якості, штучний інтелект може краще визначити, чи хтось спокійний, розчарований чи радісний, навіть без візуальних підказок, як під час телефонної розмови.

Такий рівень емоційного розуміння зробить взаємодію штучного інтелекту більш природною та чуйною, перетворюючи штучний інтелект із чисто механічного інструменту на помічника, який може віддзеркалювати та, можливо, навіть передбачати людські потреби значущим чином. Емоційно інтелектуальний штучний інтелект може покращити творчі результати, роблячи його творіння схожими на людські не лише зовнішнім виглядом, але й емоційною глибиною, що має вирішальне значення для встановлення справжніх зв'язків.

У робототехніці, особливо в гуманоїдних конструкціях, це емоційне розуміння може бути трансформаційним. Оскільки штучний інтелект продовжує розвиватися, подолання ефекту «моторошної долини» стає все більш можливим за допомогою емоційно обізнаної моделі, які відображають людські емоції та реагують на них у спосіб, який здається природним і доступним. Ця модель могла

б використовувати комп'ютерне бачення, аналіз голосу та розпізнавання жестів, щоб повністю охопити та відобразити складність людського досвіду.

Розробка та впровадження моделі штучного інтелекту з надійним розпізнаванням емоцій дозволяє їй вийти за рамки «просто текстових відповідей», дозволяючи їй обробляти та демонструвати спектр людських емоцій під час взаємодії. Такий штучний інтелект не тільки розумітиме, як емоції проявляються зовні, але й визнаватиме їхнє значення, надаючи йому притаманну людську гнучкість, яка більше походить на реальний людський досвід. Тоді такий штучний інтелект міг би служити більш чуйним партнером, збагачуючи собі людську ідентичність, відображаючи наші власні емоційні ландшафти, і таким чином ставши невід'ємною частиною людського життя та творчості.

РОЗДІЛ 1 ОСНОВИ ТА ЗНАЧЕННЯ РОЗПІЗНАВАННЯ ЕМОЦІЙ

1.1 Поняття емоційного стану та його визначення

Емоції – це складне багатовимірне явище, яке важко описати одним визначенням. Кожна людина має свій унікальний спектр емоцій, і хоча деякі емоційні переживання можуть бути схожими у різних людей, все ж таки кожен відчуває емоції по-своєму. Це пов'язано з тим, що емоційні стани є індивідуальними і можуть суттєво різнитися залежно від обставин, які їх викликають, а також особливостей особистості кожної людини.

Емоції є нетривалими: вони змінюються протягом дня під впливом різних подразників, таких як зовнішні події, взаємодії з іншими людьми чи внутрішні думки. Ця мінливість емоцій робить їх відмінними від таких постійніших психічних процесів, як пам'ять. Емоційний стан людини зазвичай є тимчасовим, і на нього можуть впливати різноманітні фактори, як зовнішнього, так і внутрішнього середовища [1].



Рисунок 1.1 – Стандартні емоційні стани, притаманні будь-якій людині

Залежно від переживань, емоційні реакції можуть бути дуже різноманітними (див. рис. 1.1) [1]:

- Щастя змушує нас відчувати позитивну енергію, і навіть складні завдання можуть виконуватись легше та із задоволенням;
- Смуток або печаль, навпаки, можуть пригнічувати, знижуючи нашу мотивацію й роблячи деякі справи важчими;
- Гнів може виникати через ситуації, де ми відчуваємо несправедливість або коли хтось завдає нам прикроців. У таких випадках ми схильні реагувати імпульсивніше, що також може вплинути на наших близьких;
- Страх є емоцією, що може як активізувати наші рішення й реакції, так і паралізувати нас.

Емоції також мають значний вплив на нашу поведінку. Наприклад, коли ми щасливі, то більш відкриті до нових ідей і дій. У стані смутку або роздратування, навпаки, ми можемо випадково передати цей настрій іншим людям, що впливає на їхній емоційний стан. Таке передання емоцій може відбуватися по «ланцюжку»: коли один відчуває негатив, він мимоволі передає його іншим, а ті – наступним.

Оскільки емоції швидко змінюються, ми можемо миттєво перейти від одного емоційного стану до іншого, наприклад, з печалі до радості, якщо трапиться щось позитивне. Таке непостійне, швидкоплинне та мінливе явище є важливою частиною нашого життя, яке впливає на наші рішення, комунікацію та сприйняття світу [1].

Емоційний стан людини – складне явище, яке потребує вивчення з різних аспектів. Хоча емоції часто сприймаються як щось швидкоплинне і миттєве, їхнє дослідження важливе для багатьох галузей, адже вони можуть розкрити багато інформації про людину. Розуміння емоційного стану важливе, наприклад, у психології, нейронауці та штучному інтелекті. Далі розглянемо, як кожна з цих галузей використовує знання про емоції та їхнє значення (див. рис. 1.2) [1, 2].

Психологія. Психологи допомагають людям зрозуміти та впоратися з емоціями. Людина може звернутися до психолога з емоційними проблемами, які

довго не проходять або навіть переростають у депресію чи тривожність. Вивчення емоційних станів у психології дозволяє фахівцям класифікувати емоції, розуміти їхній вплив на психічне здоров'я, і таким чином допомагати людині знайти способи подолання труднощів, що потребує уважного аналізу й часу, адже іноді важко точно визначити, що відчуває людина в той чи інший момент [1, 2].

Нейронаука. У цій галузі вивчають мозок та нервову систему для розуміння того, як емоції проявляються на фізіологічному рівні. Нейрони, взаємодія яких впливає на емоції, реагують на різні подразники, і наука аналізує ці реакції, щоб зрозуміти, як виникають емоції, які саме ділянки мозку задіяні в цьому процесі. Завдяки цьому ми можемо краще зрозуміти, як емоції формуються та як впливають на поведінку людини [1, 2].

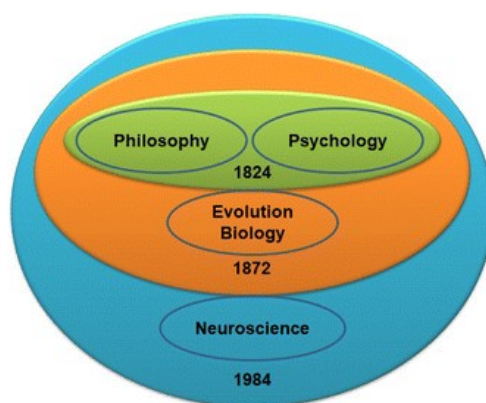


Рисунок 1.2 – Основне коло галузей, що вивчає емоційні стани

Штучний інтелект. Використання емоцій в штучному інтелекті (ШІ) (див. рис. 1.3) є дуже перспективним напрямком. Хоча ШІ здатен обробляти інформацію і генерувати тексти чи зображення, сучасні системи ще далекі від того, щоб зрозуміти емоції на глибокому рівні. Відтворення емоційної глибини може зробити ШІ ближчим до людини, оскільки, розуміючи емоції, він міг би більш точно реагувати на потреби людей. Наприклад, ШІ, який розпізнає емоційний стан людини, може бути корисним у телефонній розмові, надаючи реакцію, схожу на поведінку психолога. Також, розвиток ШІ з емоційними навичками може бути

застосований у творчих завданнях: такі системи могли б створювати художні твори або навіть нові винаходи, що відображали б людські емоційні переживання [1, 3].

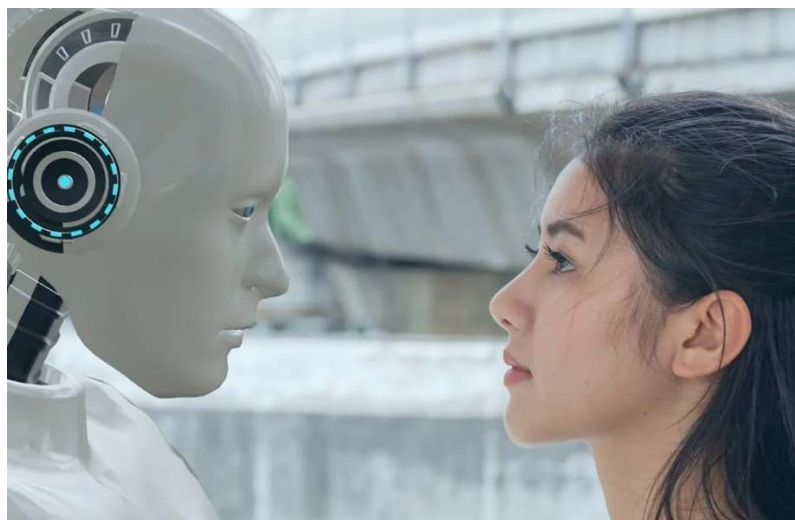


Рисунок 1.3 – ШІ та його зв'язок з людиною

Проблеми та виклики ШІ в розумінні емоцій. На сьогодні ШІ досяг значного прогресу у створенні текстів та зображень, однак його «розуміння» емоцій ще залишається поверхневим [1, 3, 4]. Навіть якщо ШІ навчиться розпізнавати емоції за виразами обличчя на зображеннях чи за інтонацією голосу, йому бракує здатності розуміти, що саме змушує людину відчувати радість, тривогу або натхнення. Справжнє «розуміння» емоцій потребує інтеграції різних аспектів людського досвіду, і це є викликом для розробників.

Проблеми впливу ШІ на ринок праці. Хоча ШІ може покращити продуктивність у багатьох сферах, швидке впровадження технології викликає побоювання, що робочі місця можуть замінюватися автоматизованими системами. Наприклад, роботи, які раніше виконувались тільки людьми, зараз можуть бути автоматизовані, що призводить до втрати робочих місць [4].

Вивчення емоційних станів – це не лише психологічний, а й технологічний виклик. Чим краще ШІ зможе розуміти емоції, тим ближчим до людського він ставатиме, що може покращити якість нашого життя, якщо ШІ допомагатиме у складних ситуаціях або прийматиме на себе рутинні завдання. Однак, важливо, щоб

ІІІ використовувався етично та не завдавав шкоди людям, наприклад, позбавляючи їх роботи [1, 2, 3, 4].

Емоційний стан людини – це складне явище, яке поєднує в собі фізіологічні й психологічні компоненти, а також залежить від попереднього досвіду та індивідуального сприйняття [1, 2]. Розглянемо детальніше, як ці компоненти впливають на наші емоції (див. рис. 1.4) [1, 2, 5].

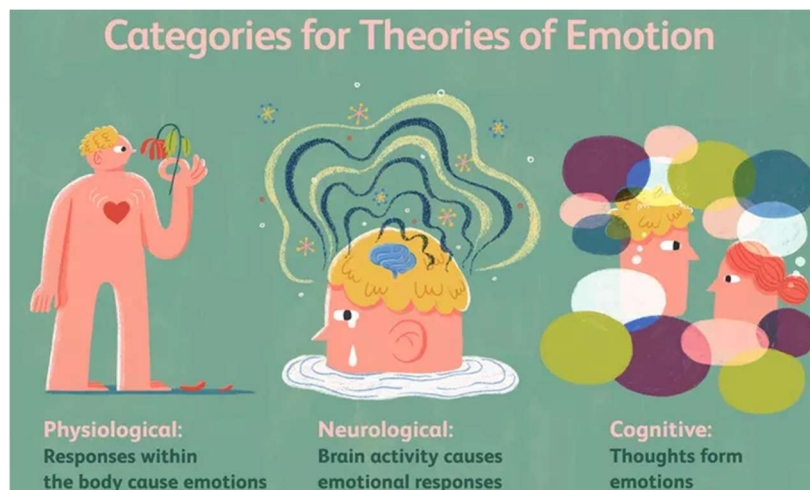


Рисунок 1.4 – Категорії вираження емоційних станів

Фізіологічний аспект емоцій. Емоційний стан людини значною мірою регулюється фізіологічними реакціями організму, зокрема нервовою та ендокринною системами [2, 5]:

- Нервова система забезпечує реакції на емоційні подразники, які іноді виявляються мимоволі, що можуть бути як зміни частоти серцебиття, ритму дихання, підвищення або зниження артеріального тиску. Наприклад, під впливом страху частота серцевих скорочень зростає, дихання стає швидшим, і зростає загальна готовність організму до дій у стресовій ситуації;
- Гормональна регуляція. Емоційні тригери спричиняють виділення різних гормонів. Так, під впливом щастя в мозку виділяється дофамін — гормон, що відповідає за почуття задоволення. Кортизол, навпаки, вивільняється у відповідь на стрес, а окситоцин посилює зв'язок та

довіру між людьми. Такі реакції є результатом складної роботи ендокринної системи.

Таким чином, фізіологічні процеси безпосередньо впливають на те, як ми відчуваємо емоції та як реагуємо на різні події, відображаючи наш емоційний стан через тілесні зміни.

Психологічний компонент емоцій. Окрім фізіологічних змін, емоції мають психологічний компонент, пов'язаний з когнітивною оцінкою ситуацій, тобто тим, як ми сприймаємо і обробляємо події навколо [2, 5]:

Кожна людина унікально інтерпретує одні й ті ж події на основі:

- Попереднього досвіду. Емоції значною мірою залежать від минулих подій. Наприклад, якщо хтось у минулому переживав складну ситуацію, пов'язану зі страхом, то в майбутньому подібні обставини можуть викликати у нього схожий емоційний відгук, що пояснює, чому на однакові ситуації люди реагують по-різному: для когось новий виклик може бути захоплюючою можливістю, тоді як для іншого – джерелом тривоги чи страху;
- Очікування. Іноді емоційний стан людини залежить від того, чого вона очікує. Наприклад, якщо хтось сподівається на дуже позитивний результат, а отримує щось інше, то це може призвести до розчарування;
- Особистісні особливості. Люди мають різний рівень самовпевненості, стійкості до стресу і загальне бачення світу, що також впливає на їхні емоції. Наприклад, одна людина може побачити в складній задачі шанс на успіх, тоді як інша сприйме це як загрозу і буде уникати ситуації.

Тобто психологічний компонент емоцій відображає складні ментальні процеси, що залежать від внутрішнього світу людини. Саме це і робить емоційний стан неповторним і важко прогнозованим для кожної людини.

Універсальні та унікальні реакції. Попри індивідуальні відмінності, певні емоції мають універсальні прояви [1, 2, 5]:

- Страх часто супроводжується характерними рисами обличчя, як от напруженість м'язів обличчя чи розширення очей;

— Щастя часто виражається через посмішку, проте деякі люди відчують таку сильну радість, що навіть можуть плакати.

Такі універсальні риси можуть бути корисними для створення моделей, наприклад, у штучному інтелекті. Проте повне розуміння емоційного стану людини вимагає більшої глибини, ніж просто спостереження за виразом обличчя, адже кожна людина реагує «по-своєму».

Виклики для моделювання емоцій. При моделюванні емоцій важливо розуміти складність і багатогранність емоційного стану. Хоча сучасні системи ШІ можуть розпізнавати базові емоційні реакції за виразом обличчя або інтонацією, повне відтворення людських емоційних переживань залишається складним завданням. ШІ здатен надавати лише поверхнєве відображення емоцій, не беручи до уваги минулий досвід, глибину сприйняття чи особистісні особливості кожної людини [1, 3, 4, 5].

Отже, емоційний стан людини – це більше, ніж просто набір реакцій. Він складається з фізіологічних змін, когнітивного сприйняття і є продуктом особистісного досвіду, що вимагає складного підходу для вивчення та аналізу, що особливо важливо враховувати у сфері штучного інтелекту, щоб навчитися ефективно взаємодіяти з людьми та розуміти їхні потреби на глибшому рівні [1, 3, 4, 5].

Для детальнішого аналізу та категоризації емоційних станів науковці виділяють три основні виміри: валентність, збудження і домінування (див. рис. 1.5) – ця система вимірів допомагає більш структуровано підходити до дослідження емоцій, роблячи зрозумілішою їхню інтерпретацію [1, 6]. Розглянемо ці категорії детальніше.

Валентність. Цей вимір відображає емоційний стан як позитивний або негативний. Наприклад, такі емоції, як радість чи любов, мають позитивну валентність, оскільки вони асоціюються із задоволенням і сприятливими відчуттями. Навпаки, емоції типу смутку або гніву мають негативну валентність, яка пов'язана з дискомфортом чи стражданням. Ця класифікація дозволяє зрозуміти, як людина суб'єктивно оцінює певний досвід або ситуацію. Наприклад,

людина з позитивною валентністю в своїй емоції, як правило, налаштована на соціальну взаємодію, готова до співпраці, тоді як негативна валентність може спричинити бажання дистанціюватися або уникати певної ситуації.

Збудження. Цей вимір визначає рівень енергії чи інтенсивності, пов'язаний з емоцією, і відображає її енергетичний компонент. Збудження може варіюватися від високого (наприклад, хвилювання чи тривога), що супроводжується посиленням серцебиття, до низького (спокій чи нудьга), що відображає розслаблення і низьку енергійність. Фізіологічно збудження регулюється нервовою системою, зокрема симпатичною гілкою, яка підвищує активність організму в разі високих емоційних навантажень, таких як страх чи гнів, забезпечуючи готовність до швидкої реакції. Наприклад, страх може викликати підвищене серцебиття та напруження м'язів, що є характерними для емоцій з високим збудженням.

Домінування. Цей вимір описує відчуття контролю чи підпорядкованості в певному емоційному стані. Високий рівень домінування характеризується відчуттям впевненості та контролю (як-от гордість чи задоволення від досягнення), тоді як низький рівень може супроводжуватися почуттям безпорадності або уразливості (наприклад, у випадку страху або смутку). У контексті соціальних взаємодій домінування відображає також внутрішнє ставлення людини до ситуацій: деякі люди можуть проявляти сміливість та самовпевненість навіть у важких обставинах, тоді як інші можуть втрачати здатність діяти – відчуття контролю важливо враховувати, оскільки воно впливає на те, наскільки впевнено людина бере участь у взаємодії або виконує завдання.

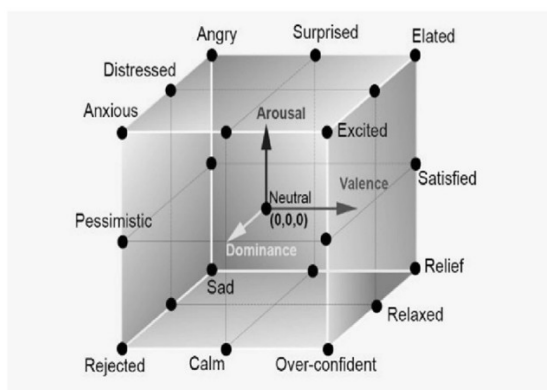


Рисунок 1.5 – Зображення трьох вимірної осі вимірювання емоційних станів

Об'єднання цих трьох категорій дозволяє не тільки структурувати емоційні стани, але й передбачати поведінку людини в певних ситуаціях. Наприклад, гнів характеризується високим рівнем збудження, негативною валентністю та помірним чи високим рівнем домінування, що зазвичай призводить до підвищеної імпульсивності та конфліктності. Смуток, навпаки, характеризується низьким збудженням, негативною валентністю та низьким рівнем домінування, що відображає стан втрати енергії і внутрішнього замкнення.

Загалом, такі виміри надають можливість дослідникам та розробникам штучного інтелекту більш точно класифікувати та прогнозувати емоційні стани людини, що є цінним у створенні систем розпізнавання емоцій [1, 6].

Процес розпізнавання емоційних станів людини є комплексним завданням, яке передбачає аналіз різноманітних сигналів, таких як вираз обличчя, тон голосу, міміка та фізіологічні показники (див. рис. 1.6) [1, 2, 6, 7, 8]. Мета цього процесу – зрозуміти емоційний стан, який людина відчуває у певний момент. Далі розглянемо детально, як це працює.

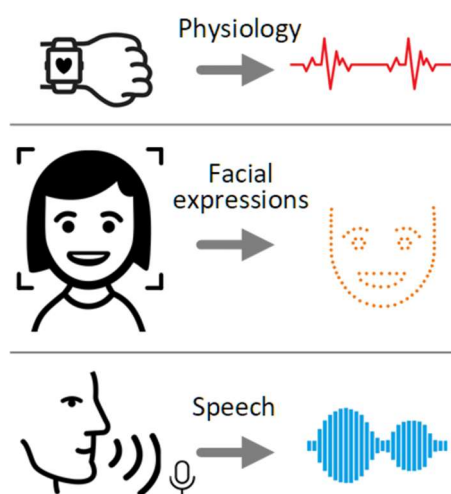


Рисунок 1.6 – Розпізнавання емоційних станів за трьома напрямками: фізіологічні, звукові показники та міміка

Набір даних для розпізнавання. У розпізнаванні емоцій ключову роль відіграє збір даних із різних джерел. Наприклад, аналіз виразу обличчя може дати уявлення про емоції завдяки технологіям розпізнавання міміки, які слідкують за рухами

м'язів обличчя. Коли людина злиться, вона може піднімати голос або змінювати його тон, що також є індикатором її внутрішнього стану. Фізіологічні дані, як-от частота серцебиття, ритм дихання чи електропровідність шкіри, також дають змогу вловлювати зміни, пов'язані з певними емоціями, навіть якщо ці зміни не видно зовні [1, 8].

Аналіз емоційних станів за допомогою технологій. Технології штучного інтелекту (ШІ) дозволяють проводити цей аналіз з високою точністю завдяки алгоритмам машинного навчання. Так, ШІ моделі навчаються на великих наборах даних, що містять приклади різних людських емоцій, що включає дані про вирази обличчя, голосові інтонації та фізіологічні реакції, які зібрані в різних емоційних станах. Завдяки цьому ШІ здатен розпізнавати певні шаблони емоцій, що дозволяє йому виявляти емоційні стани людей у майбутньому [1, 3, 4, 8].

Одним із прикладів є детектор брехні (він же поліграф) (див. рис. 1.7) – він базується на аналізі фізіологічних факторів, таких як зміни частоти серцебиття або потіння, що дозволяє визначити, чи людина обманює, хоча це не зовсім так. Коли людина бреше, її тіло реагує стресом, що може проявлятися у фізіологічних показниках, навіть якщо зовнішніх ознак немає [1, 8, 9]. Такі детектори можна використовувати не тільки у юридичній сфері, а й для покращення ефективності спілкування між людьми та системами ШІ.



Рисунок 1.7 – Розпізнавання емоційних станів за допомогою поліграфа

Взаємодія з користувачем та емоційна адаптивність III. У процесі розвитку, системи штучного інтелекту можуть удосконалювати свої алгоритми розпізнавання емоцій, що дозволяє створювати більш людяні та адаптивні відповіді. III, який навчився розпізнавати емоції, здатний надавати поради чи відповідати на запити користувача так, щоб це більше нагадувало спілкування з людиною, а не машиною [1, 3, 4]. Наприклад, він може підлаштовувати свій стиль спілкування під настрій людини, що підвищує довіру та комфортність у взаємодії.

Розпізнавання емоцій допомагає краще зрозуміти внутрішній стан людини за допомогою аналізу її поведінкових і фізіологічних реакцій [1, 2, 8]. Завдяки цьому, штучний інтелект не лише точно ідентифікує емоційні стани, але й покращує рівень взаємодії з людиною, сприяючи створенню комфортнішого та більш персоналізованого спілкування.

Фізіологічні показники емоційних станів (див. рис. 1.8) – це реакції тіла на емоції, які людина переживає. Вони відображаються через різноманітні фізіологічні зміни, що відбуваються мимоволі та є тимчасовими. Важливо розуміти, що ці реакції не контролюються людиною безпосередньо і можуть слугувати індикаторами емоційних станів. Ось кілька основних фізіологічних показників, які пов'язані з емоційними станами [1, 3, 4, 8, 13]:

1. Частота серцебиття. Емоції, такі як страх або хвилювання, часто викликають збільшення частоти серцевих скорочень, що є показником збудження та інтенсивності емоцій. Наприклад, коли людина відчуває сильний страх, її серце може почати битися швидше – це прискорене серцебиття є важливим фізіологічним маркером, який дозволяє визначити, наскільки інтенсивно людина переживає ту чи іншу емоцію.

2. Частота дихання. Ще один фізіологічний показник емоцій — зміна частоти дихання. Наприклад, страх або тривога часто супроводжуються прискореним диханням. Коли людина сильно хвилюється, вона може дихати швидше, іноді до такої міри, що це може призвести до запаморочення або навіть втрати свідомості. Такий стан можна спостерігати, коли людина раптово відчуває небезпеку чи бачить щось тривожне, наприклад кров.

3. Електродермальна активність (потовиділення). Під час емоційного збудження шкіра починає виділяти більше поту, що підвищує її електропровідність. Таке явище називається електродермальною активністю. Наприклад, під час сильного хвилювання або стресу людина може почати потіти. Такий ефект використовується в детекторах брехні, які вимірюють зміни провідності шкіри для визначення емоційної напруженості.

4. Розширення зіниць. Зіниці також реагують на емоції. Коли людина відчуває хвилювання або переживає сильні емоції, її зіниці розширюються. Натомість, коли людина нудьгує або не зазнає особливого збудження, зіниці можуть звужуватися – ще один спосіб визначити емоційний стан за фізіологічними показниками.

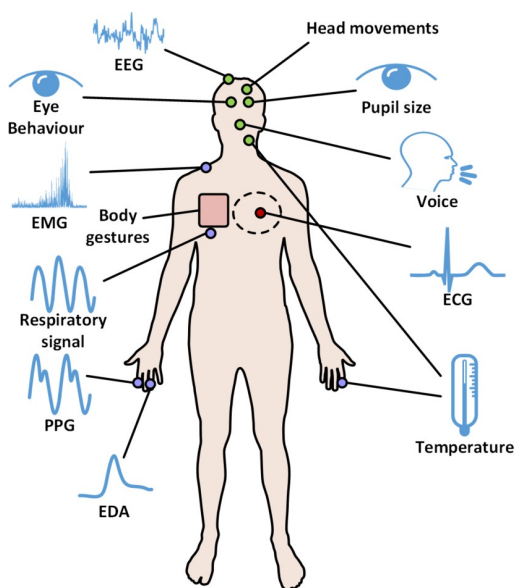


Рисунок 1.8 – Розпізнавання емоційних станів за допомогою фізіологічних показників

Фізіологічні показники дають важливу інформацію про емоційний стан людини, і розуміння цих даних є надзвичайно корисним для технологій, що використовують моделі штучного інтелекту (ШІ). Вони дозволяють алгоритмам машинного навчання більш точно визначати емоції людини в реальному часі [3, 4, 8, 13].

Використання фізіологічних показників у психології та медицині. Ці показники широко застосовуються в психології та медицині. Наприклад, зміни у частоті серцебиття або дихання можуть допомогти медикам виявити емоційні або фізіологічні проблеми, навіть якщо людина не може або не хоче розповісти про свої переживання, що особливо корисно у випадках, коли людина може не усвідомлювати свій стан або не контролювати свою поведінку.

Роль ІІІ в аналізі фізіологічних показників. Штучний інтелект може бути навчений розпізнавати фізіологічні маркери, які супроводжують емоційні стани, й інтерпретувати їх відповідно. Завдяки цьому медичні або психологічні системи на базі ІІІ зможуть надавати більш інформовані відповіді або попередження, наприклад, якщо людина відчуває стрес або панічну атаку, що сприятиме створенню інструментів, які допомагатимуть у клінічному застосуванні, зокрема в діагностиці та моніторингу станів пацієнтів.

Психологічні інтерпретації емоційних станів. Емоційні стани можна розглядати з психологічної точки зору, і тут акцент робиться не стільки на фізіологічні зміни, скільки на суб'єктивні переживання й особисті інтерпретації подій [1, 2, 7].

Емоції можна визначити як реакції, що залежать від внутрішнього світу людини, її особистісного досвіду та очікувань. Ці реакції, хоч і можуть проявлятися через певні фізіологічні зміни (наприклад, прискорення серцебиття), насамперед відображають сприйняття та інтерпретацію ситуації [1, 2, 8].

Серед основних психологічних теорій, що пояснюють виникнення емоцій, є когнітивна та оцінювальна. Згідно з когнітивною теорією, емоції виникають як результат оцінки ситуації, яка базується на минулому досвіді, сформованих очікуваннях та наявних уявленнях. Наприклад, якщо людина переживала певну ситуацію у минулому і має негативний досвід, то подібна ситуація в майбутньому викликати страх або тривогу. Водночас очікування відіграють важливу роль: якщо людина сподівається на позитивний результат, емоції можуть бути більш оптимістичними [1, 2, 5].

Оцінювальна теорія, у свою чергу, стверджує, що емоції є наслідком оцінки індивідом подій у контексті його цілей і бажань. Коли події збігаються з очікуваннями чи бажаннями, виникають позитивні емоції; якщо ні – негативні, що показує, що емоції глибоко індивідуальні, і схожі обставини можуть викликати різні емоційні реакції у різних людей, залежно від їхнього життєвого досвіду та цілей [1, 2, 5].

Роль емоційних станів у соціальних взаємодіях і прийнятті рішень. Емоції мають ключову роль у соціальних взаємодіях і прийнятті рішень. Вони впливають на те, як люди сприймають один одного, будують відносини, формують групову динаміку та вирішують різноманітні завдання. Наприклад, емоції співчуття та гніву можуть змінити наше ставлення до інших: співчуття зближує, тоді як гнів може ускладнити взаємодію та створити конфлікт [1, 2, 5].

Емоційні стани також значно впливають на процес прийняття рішень. Часто рішення, ухвалені під впливом емоцій, є імпульсивними й необдуманими. Наприклад, люди можуть купувати речі, в яких насправді не мають потреби, через миттєве бажання чи вплив реклами. Пізніше, коли емоції вщухають, виникає розуміння, що покупка була зайвою. Усвідомлення цього механізму є цінним, оскільки дозволяє людині приймати більш зважені рішення та уникати надмірних витрат чи емоційно забарвлених імпульсів.

Ці знання важливі для побудови програм обслуговування клієнтів і психологічної підтримки. Наприклад, усвідомлення емоційного стану клієнта дозволяє реагувати більш чуйно та індивідуально, надавати підтримку, яка може допомогти людині почуватися зрозумілою та підтриманою.

Застосування розуміння емоційного стану в технологіях. Розуміння емоційних станів активно застосовується у технологіях, особливо в галузі штучного інтелекту. Системи розпізнавання емоцій дедалі частіше впроваджуються для покращення взаємодії з користувачами. Наприклад, у сфері обслуговування клієнтів віртуальні помічники й чат-боти, здатні розпізнавати емоції, допомагають налаштовувати відповіді, що відповідають настрою користувача, забезпечуючи таким чином індивідуальний підхід [3, 4, 6].

Крім цього, розпізнавання емоцій використовується в охороні здоров'я, де аналіз емоцій пацієнтів може бути корисним у терапії, оцінці їхнього психічного стану та підтримці процесу лікування. Така інформація допомагає лікарям та терапевтам точніше діагностувати стан пацієнта та планувати лікування.

У маркетингу розуміння емоцій використовується для аналізу реакцій клієнтів на продукти чи рекламу, що допомагає покращити маркетингові стратегії. Реакція клієнтів (навіть коментарі) можуть нести емоційний заряд, а аналіз цих емоційних відгуків дозволяє компаніям створювати більш привабливі пропозиції.

У сфері освіти емоційний аналіз допомагає адаптивним навчальним платформам підлаштовувати навчальні матеріали відповідно до емоційних реакцій учнів, що надає можливість зробити процес навчання більш цікавим і продуктивним, надаючи учням стимули до навчання, допомагаючи їм легше освоювати матеріал і зацікавлюючи в навчанні.

1.2 Еволюція підходів до аналізу емоційного стану

Аналіз емоційних станів привертав увагу науковців з різних сфер, таких як психологія, біологія та філософія, протягом багатьох століть, відбувається тому, що розуміння емоцій має ключове значення для вивчення природи людини, соціальної поведінки та мотивації. Вчені намагалися зрозуміти, що таке емоції, як вони виникають і яку роль вони відіграють у житті людей та тварин [2, 5, 7].

Перші теоретичні підходи розглядали емоції як необхідні реакції для виживання, що були частиною еволюційної адаптації. Згідно з цими теоріями, емоції та емоційні стани розвивалися для того, щоб допомагати організмам виживати в мінливих умовах навколишнього середовища. Дослідники спостерігали за поведінкою та фізичними реакціями на різні стимули, припускаючи, що емоції мають біологічне підґрунтя і забезпечують адаптацію до середовища.

Чарльз Дарвін та Вільям Джеймс, видатні теоретики свого часу, запропонували різні підходи до розуміння емоцій. Дарвін у своїй роботі підкреслив роль емоцій в еволюційному процесі, розглядаючи їх як адаптації, що допомагають людям і тваринам реагувати на виклики довкілля. Він вважав, що емоційні реакції є результатом еволюційного розвитку та сприяють виживанню [5].

З іншого боку, Вільям Джеймс звернув увагу на фізіологічні зміни, які супроводжують емоційні переживання. Його теорія припускала, що емоції виникають як реакція на фізичні зміни, спричинені певними подразниками. Таким чином, Джеймс розглядав емоції з точки зору фізіології, підкреслюючи їхню тісну взаємодію з тілесними реакціями [5, 7].

Ідеї Дарвіна та Джеймса стали фундаментом для подальшого дослідження емоційних станів. Їхні підходи заклали основи для розвитку сучасних теорій емоцій, що охоплюють як біологічний, так і психологічний аспекти [5].

Перші біологічні та фізіологічні теорії. На самому початку наукового дослідження емоцій переважали біологічні та фізіологічні підходи. Вчені прагнули створити теоретичні моделі, щоб спочатку пояснити, а потім і експериментально досліджувати механізми виникнення емоцій. Одна з перших структурованих теорій розглядала емоції як біологічні та фізіологічні феномени, що мають еволюційне підґрунтя [5].

Чарльз Дарвін, у своїй відомій праці «Вираження емоцій у людини і тварин» (1872), висунув гіпотезу, що емоції є еволюційними адаптаціями. За його припущенням, емоції допомагають передавати внутрішній стан та наміри, підтримуючи таким чином соціальні зв'язки та сприяючи виживанню [5, 10]. Дарвін вважав, що однакові емоційні прояви у різних видів, зокрема у людей і тварин, свідчать про їхню спільну еволюційну історію. Його спостереження надихнули багатьох дослідників вивчати емоції як універсальні реакції, що притаманні різним живим організмам.

Дещо пізніше Вільям Джеймс та Карл Ланге розробили спільну теорію емоцій, що отримала назву «теорія Джеймса-Ланге». Вони запропонували, що емоції є прямим наслідком фізіологічних реакцій на зовнішні стимули. Відповідно

до їхньої теорії, емоції не виникають самі собою, а з'являються як відповідь на фізичні зміни в організмі [5, 8, 10]. Наприклад, такі реакції, як прискорене серцебиття чи потовиділення, мозок сприймає як ознаки гніву, страху чи радості. Такий підхід дав новий імпульс для дослідження фізіологічних процесів, пов'язаних з емоціями, і став основою для подальших експериментів, які вивчають зв'язок між тілесними та емоційними реакціями [5].

Поява психологічних теорій. З середини XX століття дослідження емоцій почало виходити за межі виключно біологічних та фізіологічних підходів. Психологи почали розглядати емоції як складний досвід, що включає когнітивні, соціальні та ситуаційні фактори. Одним із найвпливовіших дослідників цього напрямку був Пол Екман. Спираючись на роботи Дарвіна, Екман вивчав універсальність виразів обличчя та виділив шість базових емоцій: щастя, сум, гнів, страх, подив і огиду [1, 2, 5]. Він припустив, що ці емоції є універсальними і однаково розпізнаються у різних культурах. Дослідження Екмана створили основу для аналізу виразу обличчя як індикатора емоційного стану та визначення емоційної «мови», яка зрозуміла у всьому світі без слів [2, 5].

У цей же період з'являються й когнітивні теорії емоцій, що базуються на припущенні, що емоції є автоматичною реакцією на інтерпретацію або оцінку певної ситуації. Наприклад, дослідник Річард Лазарус стверджував, що емоції виникають внаслідок оцінки подій, пов'язаних з особистими цілями та благополуччям. Ця модель когнітивної оцінки значно розширила розуміння емоцій, включаючи в них особисте сприйняття, соціальний контекст і когнітивні процеси, що додають багатовимірності емоційним переживанням [1, 2, 5].

Обидва ці напрями – біологічний і когнітивний – дали фундамент для сучасного розуміння емоцій. Вони допомогли виявити, що емоції є складною комбінацією фізіологічних, когнітивних та соціальних факторів, які разом формують людський досвід [5].

Розвиток емпіричних та експериментальних досліджень емоцій. З розвитком психології як науки дослідження емоційних станів перейшло від теоретичних підходів до більш практичних емпіричних досліджень. Науковці

почали активно застосовувати експериментальні методи, що дозволяло досліджувати емоції у контрольованих умовах. Виникла можливість не лише аналізувати теорії емоцій, а й на практиці вивчати, як емоції викликаються, вимірюються та проявляються у різних ситуаціях [1, 2, 5].

Проведення лабораторних експериментів дозволило науковцям досліджувати фізіологічні реакції, вирази обличчя та голосові сигнали, що супроводжують певні емоції. Такі дослідження дали змогу збирати кількісні дані про емоційні реакції, що стало основою для об'єктивного аналізу емоційних станів. Наприклад, розробка системи кодування дій обличчя (FACS), яку запропонував Пол Екман разом із колегами, надала дослідникам інструмент для декодування м'язових рухів обличчя, що відповідають різним емоціям, що стало важливим кроком у напрямку точного аналізу емоцій, оскільки дозволило стандартизувати та об'єктивізувати дослідження емоційних проявів [5, 8].

Контрольовані експерименти не лише покращили точність результатів, але й зробили дослідження повторюваними, що забезпечує науковцям можливість підтверджувати отримані дані. Завдяки таким підходам стало можливим робити науково обґрунтовані висновки про природу емоцій та способи їх вираження у поведінці людини, що є важливим для подальшого розвитку психології як науки.

Технологічний прогрес і перехід до обчислювальних підходів. Поява комп'ютерів та розвиток цифрових технологій наприкінці XX століття спричинили значні зміни у дослідженні емоцій. Дослідники почали використовувати математичні моделі та статистичний аналіз для вивчення емоцій, що відкрило шлях до створення автоматизованих систем розпізнавання емоцій. Технології обробки сигналів дозволили аналізувати голос, вирази обличчя та фізіологічні сигнали, такі як частота серцевих скорочень і провідність шкіри, в реальному часі, що дало змогу дослідникам працювати з великими масивами даних та виявляти закономірності емоційних проявів у різних контекстах [1, 2, 8].

На ранніх етапах комп'ютерні системи були здебільшого орієнтовані на аналіз окремих сигналів, таких як вирази обличчя та голос. Однак з розвитком обчислювальних потужностей стало можливим об'єднання даних з різних джерел,

що дозволило створювати мультимодальні системи, які одночасно враховують аудіо, візуальні та фізіологічні сигнали, що суттєво розширило можливості розуміння емоційних станів [2, 8, 9]. Такий підхід забезпечив більш комплексний аналіз, оскільки дозволяє враховувати багатогранність та складність емоцій у різних ситуаціях.

Машинне навчання та штучний інтелект у сучасному розпізнаванні емоцій.

З розвитком сучасних технологій, зокрема машинного навчання та штучного інтелекту, відбулося значне покращення у сфері розпізнавання емоцій. Ці технології дозволяють системам з часом навчатися, вдосконалюватися та аналізувати великі набори даних для виявлення шаблонів. Наприклад, штучний інтелект на основі машинного навчання використовує великі обсяги даних, що містять емоційні вирази і фізіологічні сигнали, для створення моделей, які можуть з високою точністю ідентифікувати різні емоційні стани [1, 3, 4, 8, 11].

Алгоритми, такі як нейронні мережі і моделі глибокого навчання, здатні аналізувати складні характеристики емоційних виразів значно точніше, ніж традиційні статистичні моделі. Наприклад, згорточні нейронні мережі (CNN) (див. рис. 1.9) ефективно розпізнають вирази обличчя на фото та відео, тоді як рекурентні нейронні мережі (RNN) підходять для обробки послідовних даних, таких як аудіозаписи, аналізуючи зміни тону та ритму в голосі [2, 8, 11, 12, 14].

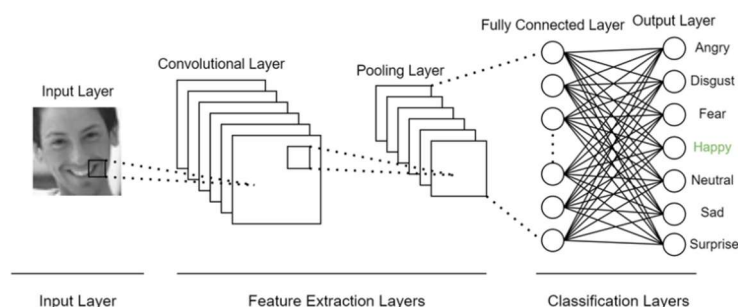


Рисунок 1.9 – Нейронна мережа, що визначає емоційний стан через зображення

Тренуючись на великих наборах даних, штучний інтелект здатен виявляти навіть незначні зміни у виразах обличчя чи тоні голосу, що значно підвищує точність визначення емоційного стану.

Розпізнавання емоцій за допомогою штучного інтелекту пройшло шлях від статичних систем, які використовували заздалегідь визначені правила, до адаптивних моделей, здатних передбачати емоційні стани. Наразі такі технології знаходять широке застосування у різних галузях, включаючи обслуговування клієнтів, охорону здоров'я, освіту та маркетинг, де розуміння емоцій користувачів сприяє покращенню взаємодії та підвищенню якості обслуговування [3, 4, 7, 8, 11, 12, 14].

Перехід від теоретичних досліджень до практичних застосувань. Розвиток технологій розпізнавання емоцій пройшов шлях від теоретичних досліджень до практичних застосувань, що є важливим досягненням для багатьох сфер. Сьогодні системи, які здатні розпізнавати емоції, вже вбудовані у повсякденні програми, що сприяє покращенню взаємодії користувачів та підвищує рівень їхньої безпеки. Наприклад, у сфері психологічного здоров'я такі системи можуть контролювати емоційний стан пацієнтів, що допомагає терапевтам відстежувати прогрес та швидко реагувати на зміни настрою.

У сфері обслуговування клієнтів віртуальні помічники, оснащені технологією розпізнавання емоцій, можуть змінювати свої відповіді, враховуючи емоції користувача, що дозволяє забезпечити більш персоналізовану та чуйнішу підтримку. У галузі автомобільної безпеки розпізнавання емоцій може допомагати стежити за станом водія, відстежуючи ознаки втоми або роздратування, що сприяє підвищенню безпеки на дорогах.

Практична інтеграція аналізу емоцій свідчить про успіх технологічної еволюції у цій сфері, а також її великий потенціал для подальшого розвитку.

1.3 Важливість розпізнавання емоцій в сучасних технологіях

Технології розпізнавання емоцій пройшли значний шлях від теоретичних досліджень до практичного застосування, набуваючи нових можливостей завдяки розвитку штучного інтелекту (ШІ) та машинного навчання (МН). Сьогодні ці технології використовують сенсори та аналітичні алгоритми, щоб розпізнавати, інтерпретувати та навіть реагувати на людські емоції в реальному часі, що досягається за рахунок комплексного збору та аналізу інформації з різних джерел, таких як вираз обличчя, тон та тембр голосу, фізіологічні сигнали (пульс, ритм дихання) та текстові дані, які людина вводить. Разом ці показники дозволяють технологіям ШІ глибше розуміти емоційний стан людини та адаптувати свої реакції відповідно до його особливостей [2, 3, 4, 5, 8, 9, 10, 11, 14].

Здатність ШІ інтерпретувати емоції відкриває широкі можливості для підвищення функціональності технологій, наближаючи їхню поведінку до людських особливостей та роблячи їх більш інтуїтивно зрозумілими і адаптивними для людей, що сприяє створенню технологій, які не просто виконують задачі, але й можуть налаштовувати свою взаємодію, враховуючи емоційні потреби користувачів, що є важливим кроком до створення емпатійних та більш «людяних» систем.

Сьогодні розпізнавання емоцій активно впроваджується в таких сферах, як обслуговування клієнтів, охорона здоров'я, освіта та безпека. У сфері обслуговування клієнтів, наприклад, віртуальні помічники, оснащені емоційним інтелектом, можуть адаптувати свої відповіді залежно від настрою користувача, роблячи спілкування більш природним і сприятливим. В охороні здоров'я системи розпізнавання емоцій допомагають відстежувати психологічний стан пацієнтів, дозволяючи лікарям краще розуміти емоційні зміни та більш ефективно надавати підтримку. У сфері освіти такі системи можуть аналізувати емоційний стан учнів, що дозволяє викладачам коригувати підхід до навчання для підтримки зацікавленості та зменшення стресу.

Технології розпізнавання емоцій відіграють важливу роль у підвищенні рівня емпатії взаємодії між людиною та штучним інтелектом, роблячи системи більш пристосованими до потреб користувачів та підвищуючи ефективність таких технологій [3, 4, 11]. Завдяки цьому вони мають потенціал значно змінити спосіб нашої взаємодії з технологіями та підвищити рівень задоволеності та комфорту в багатьох аспектах повсякденного життя.

Розпізнавання емоцій у взаємодії людини з комп'ютером (HCI). Сучасні технології, що забезпечують взаємодію людини з комп'ютером (HCI) (див. рис. 1.10), поступово розвиваються в напрямі інтеграції здатності розпізнавати емоційний стан користувача, що робить взаємодію більш природною, наближеною до людського спілкування та інтуїтивно зрозумілою для користувача. Технології розпізнавання емоцій використовують штучний інтелект (ШІ) та машинне навчання (МН), щоб аналізувати різні сигнали, такі як вираз обличчя, тон голосу, фізіологічні показники та навіть текстові повідомлення, дозволяючи комп'ютерним системам глибше розуміти емоційний стан користувачів [2, 3, 4, 11, 14, 15].

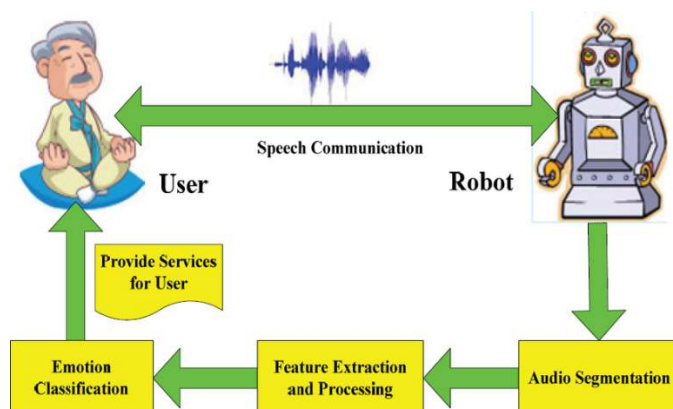


Рисунок 1.10 – Взаємодія людини та ШІ (HCI)

Особливості емоційного інтерфейсу. Традиційні інтерфейси, такі як клавіатура чи текстовий чат, обмежують можливості комп'ютера зрозуміти емоції користувача. Наприклад, під час спілкування через текст чат-бот не може повною мірою зрозуміти емоційні відтінки чи настрій людини. Проте додавання функції

голосового введення дозволяє системі вловлювати тон і тембр голосу, що значно розширює можливості для розпізнавання настрою користувача. Таким чином, ШІ може адаптувати свої відповіді, роблячи їх більш привітними, чуйними або навіть емпатійними [3, 4, 8, 11, 15, 16].

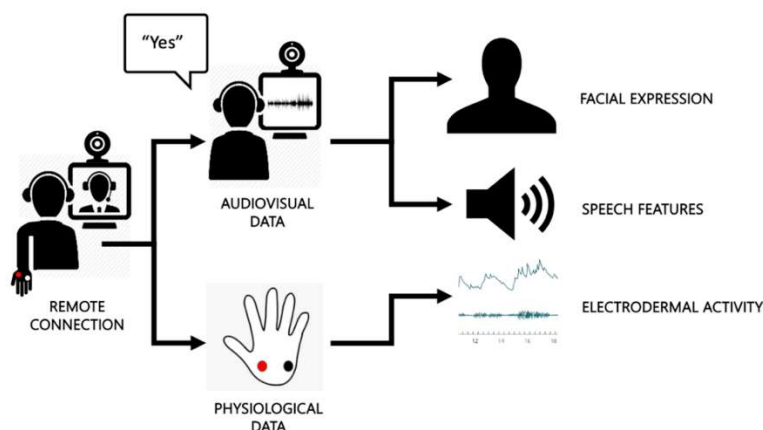


Рисунок 1.11 – Інтерфейс розпізнавання емоційних станів

Наприклад, віртуальні асистенти, такі як Google Assistant, Siri або Alexa, хоч і не мають функції розпізнавання емоційних станів, але якби вони були налаштовані для розпізнавання емоційних ознак у голосі користувача, таких як розчарування чи стрес. Відповідно до цього, асистент зміг би пропонувати допомогу чи надавати підтримку, роблячи це схоже на дії кваліфікованого спеціаліста або навіть психолога, що надає користувачам відчуття підтримки, навіть якщо вони не мають змоги поспілкуватися з реальною людиною.

Застосування для людей з особливими потребами. Розпізнавання емоцій у НСІ має потенціал для допомоги не лише людям, які переживають стресові ситуації, але й користувачам з особливими потребами, включаючи тих, хто має розлади сприйняття емоцій. Для людей з психічними розладами або розладами спектра аутизму комп'ютерні системи, оснащені здатністю до розпізнавання емоцій, можуть стати цінним інструментом підтримки [15, 16, 17]. Такі системи здатні інтерпретувати нетипові соціальні чи емоційні сигнали, що може допомогти користувачу відчувати себе більш комфортно під час взаємодії.



Рисунок 1.12 – Застосування НСІ для особи з обмеженими можливостями

Наприклад, ІІІ-система, навчившись розуміти поведінкові особливості користувача з обмеженими можливостями (див. рис. 1.12), може надавати адаптовані та персоналізовані відповіді, що сприятиме створенню більш комфортного та дружнього досвіду взаємодії, що може допомогти таким людям отримати емоційну підтримку і навіть полегшити їх адаптацію до технологій, що, в свою чергу, сприятиме підвищенню якості їхнього життя [3, 4, 15, 16, 17, 18].

Переваги персоналізованого підходу. Здатність штучного інтелекту розпізнавати емоції сприяє створенню інтерфейсів, які не лише виконують стандартні завдання, але й можуть адаптуватися до індивідуальних емоційних потреб користувачів. Наприклад, комп'ютерна система, яка довго взаємодіє з певним користувачем, може з часом «навчитися» надавати персональні відповіді, враховуючи попередній досвід та емоційні особливості користувача, що дозволяє створити систему, яка є не лише ефективною, але й емпатійною, спроможною надавати комфорт та підтримку в різних ситуаціях [3, 4, 14, 15, 16, 17, 18].

В цілому, технології розпізнавання емоцій у НСІ значно розширюють можливості взаємодії людини з комп'ютером. Вони не лише покращують функціональність техніки, але й сприяють створенню більш гуманних, дружніх та емоційно чуйних систем, які здатні краще відповідати на індивідуальні потреби кожного користувача.

Поліпшення обслуговування клієнтів і маркетингу за допомогою розпізнавання емоцій. Технологія розпізнавання емоцій значно б змінила підхід до обслуговування клієнтів і маркетингу, дозволяючи компаніям краще розуміти емоційний стан своїх клієнтів і відповідно адаптувати комунікацію. Завдяки можливості аналізу емоцій клієнтів у реальному часі компанії можуть підвищити рівень задоволеності клієнтів і сприяти формуванню їхньої лояльності. Використовуючи алгоритми машинного навчання, штучний інтелект здатен розпізнавати емоції, що в умовах традиційного обслуговування зазвичай складно досягти, особливо коли взаємодія ведеться текстовим форматом [3, 4, 14, 15, 16, 19].

Інструменти розпізнавання емоцій дозволяють аналізувати тон голосу, текст та міміку клієнтів, визначаючи такі емоції, як розчарування, радість або незадоволеність [1, 8]. Наприклад, якщо клієнт виявляє ознаки розчарування, система на основі штучного інтелекту може прискорити відповідь або передати запит на більш детальне розглядання фахівцем, що надає агентам обслуговування та маркетологам можливість налаштувати відповіді, що краще відповідають емоційному стану клієнта. У маркетингу розпізнавання емоцій застосовується для оцінки реакцій споживачів на продукти, рекламу та бренд, забезпечуючи цінний зворотний зв'язок для підвищення залученості.

Розпізнавання емоцій у сфері охорони здоров'я та підтримці психічного здоров'я. У сфері охорони здоров'я розпізнавання емоцій б відіграло важливу роль у підтримці психічного здоров'я, дозволяючи медичним працівникам слідкувати за емоційним станом пацієнтів та виявляти ранні ознаки психічних розладів, таких як депресія або тривога. За допомогою пристроїв з датчиками, які моніторять фізіологічні показники, наприклад, частоту серцевих скорочень, електропровідність шкіри чи тембр голосу, можна попередити про зміни, що вказують на можливі психічні кризи [3, 4, 14, 15, 16, 20]. Такий підхід дозволяє своєчасно втрутитися та допомогти уникнути погіршення психічного стану пацієнтів, сприяючи швидшому одужанню.

Крім того, терапевти зможуть використовувати розпізнавання емоцій як інструмент для відстеження прогресу пацієнта, що дозволяє їм ефективніше коригувати лікування. Наприклад, додатки на основі штучного інтелекту можуть аналізувати мовлення або вирази обличчя під час сеансів терапії, виявляючи ознаки покращення або погіршення психічного здоров'я, що допомагає терапевтам об'єктивно оцінювати результати лікування та оперативно адаптувати свої підходи, щоб краще задовольнити потреби пацієнта.

Завдяки додаткам для відстеження емоцій (наразі лише фізичних показників) (див. рис. 1.13) люди зможуть контролювати свій емоційний стан, краще розуміти свої емоційні реакції та формувати стійкість до стресових ситуацій, що надає людям можливість активно управляти своїм психічним здоров'ям, підвищуючи самосвідомість і сприяючи розвитку стійкості [1, 2, 8, 14, 20]. Такі інструменти допомагають користувачам швидше ідентифікувати та вирішувати емоційні проблеми, зменшуючи вплив стресу на їхнє життя.



Рисунок 1.13 – Відстеження фізичних показників за допомогою гаджетів-компаньйонів (додаток Samsung Health Monitor)

Підвищення безпеки в автомобілях або автомобільні системи розпізнавання емоцій. Розпізнавання емоцій на основі штучного інтелекту стає все більш перспективним інструментом для підвищення безпеки в автомобільних системах.

Завдяки інтеграції таких технологій у транспортні засоби можна створити ефективні механізми моніторингу емоційного стану водія, які впливають на його поведінку за кермом і, відповідно, на загальну безпеку на дорозі (див. рис. 1.14) [21].

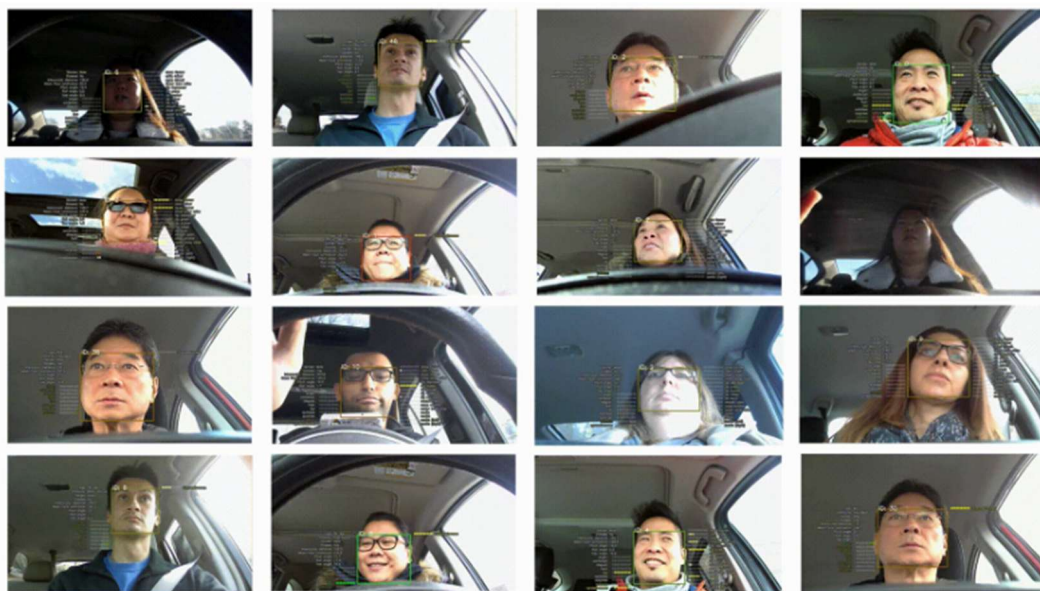


Рисунок 1.14 – Покращення безпеки за кермом під час дороги через розпізнавання емоцій

Системи розпізнавання емоцій в авто, оснащені камерами та датчиками, здатні виявляти ознаки втоми, стресу чи агресії у водія. Ці емоційні стани є одними з ключових факторів ризику для виникнення аварій. Наприклад, система може помітити ознаки сонливості чи занепокоєння водія і видати відповідне сповіщення. В такому випадку, вона може запропонувати зробити паузу для відпочинку або змінити налаштування автомобіля, наприклад, адаптувати рівень освітлення чи увімкнути заспокійливу музику, щоб підвищити рівень концентрації водія [21].

Крім того, розпізнавання емоцій може підвищити комфорт пасажирів. В залежності від настрою пасажирів, система може автоматично регулювати температуру в салоні, положення сидінь або інші параметри для створення приємнішого середовища. Такий підхід сприяє не лише підвищенню безпеки, але й

забезпеченню комфортної подорожі, яка максимально адаптується до потреб водія і пасажирів.

Освіта та навчання або розпізнавання емоцій на адаптивних навчальних платформах. Технології розпізнавання емоцій знаходять широке застосування й у сфері освіти, особливо для створення адаптивних навчальних платформ. Такі платформи здатні реагувати на емоційний стан учнів та студентів у процесі навчання, що може суттєво підвищити ефективність засвоєння матеріалу (див. рис. 1.15) [22].

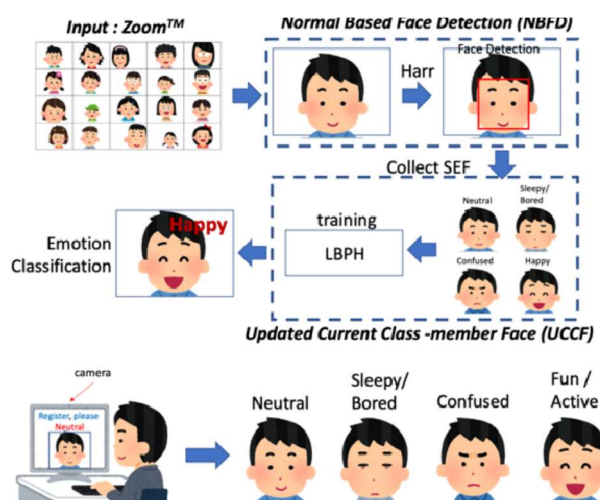


Рисунок 1.15 – Модель адаптивної навчальної платформи, побудованої на розпізнаванні емоцій

В сучасній освіті не завжди враховуються індивідуальні емоційні реакції учнів на навчальні матеріали, що може призводити до зниження мотивації та неефективного навчання. Системи розпізнавання емоцій можуть оцінювати рівень залученості та розчарування студентів на основі міміки, жестів, інтонації голосу. Завдяки цьому алгоритми здатні адаптувати навчальний процес: якщо система помічає, що студент відчуває розчарування або збентеження, вона може надати додаткові підказки або спростити матеріал. А якщо студент демонструє ознаки нудьги, платформа може збільшити складність завдань, щоб утримати його увагу.

1.4 Соціальні та психологічні аспекти емоційного аналізу

Розпізнавання емоцій – це складна міждисциплінарна сфера, яка охоплює не лише технологічні аспекти, але й тісно пов'язана з соціальними та психологічними процесами, що пояснюється тим, що розуміння та інтерпретація емоцій є основою для формування людських стосунків, адже саме емоції визначають, як ми спілкуємося, співпереживаємо та будуємо зв'язки з іншими [1, 2, 5, 7].

З соціальної точки зору, емоції виконують важливу функцію, допомагаючи людям орієнтуватися в динаміці групових стосунків. Вони дозволяють нам зчитувати соціальні сигнали, реагувати на них та формувати міцніші стосунки, зважаючи на потреби й почуття оточуючих. Наприклад, розпізнаючи емоції інших, ми можемо більш ефективно реагувати на їхній стан, проявляти підтримку або, навпаки, стримувати своє ставлення, що сприяє кращому розумінню одне одного та сприяє згуртованості в соціумі.

Психологічний аспект розпізнавання емоцій також є надзвичайно важливим, оскільки здатність ідентифікувати та розуміти власні емоції допомагає людині регулювати свій емоційний стан. Саморегуляція емоцій є важливою складовою саморозвитку, адже дозволяє залишатися стійким у складних ситуаціях, не піддаватися імпульсивним емоційним реакціям та краще адаптуватися до життєвих викликів. Уміння розпізнавати емоції – як власні, так і інших – є основою для емоційної стійкості, яка сприяє психологічному благополуччю [2, 5, 7].

Сучасні технології також знаходять застосування у сфері розпізнавання емоцій, створюючи інструменти, що покращують взаємодію на рівні «людина-комп'ютер» – вони використовують ідеї, закладені в соціально-психологічних дослідженнях, щоб підвищити якість цифрових взаємодій та сприяти особистому благополуччю користувачів [8, 15]. Наприклад, розробляються програми, які можуть аналізувати тон голосу або міміку для кращого розуміння емоційного стану співрозмовника, що допомагає підвищити соціальну згуртованість навіть у дистанційній комунікації.

Проте звести розпізнавання емоцій лише до технологічного аспекту було б надто спрощеним підходом, оскільки емоції мають глибокі психологічні корені, а також залежать від багатьох зовнішніх і внутрішніх факторів, таких як соціальне середовище, життєвий досвід, особисті установки та інші чинники [2, 5, 10]. Тому, щоб створювати ефективні інструменти для розпізнавання емоцій, важливо враховувати всю складність цього явища, де технології лише доповнюють, а не замінюють людське розуміння та емпатію.

Емоційне розпізнавання та емпатія в соціальних взаємодіях. Емоційне розпізнавання відіграє вирішальну роль у розвитку емпатії, тобто здатності розуміти та співпереживати почуттям інших, що є важливою основою для людських стосунків, адже емпатія допомагає формувати більш глибокі й міцні соціальні зв'язки. Коли ми здатні розпізнавати емоції інших, ми можемо краще налаштувати свою поведінку, проявляти розуміння та надавати підтримку, що значно сприяє зміцненню соціальної гармонії (див. рис. 1.16) [2, 5, 10, 23].



Рисунок 1.16 – Взаєморозуміння емоційних станів через емпатію

Наприклад, у близьких стосунках, здатність помічати такі емоції, як розчарування чи смуток, дозволяє нам реагувати чуйно, проявляти розуміння та підтримку, що поглиблює взаємопорозуміння й довіру. В професійних стосунках емпатія також має велике значення, оскільки вона сприяє кращій командній роботі.

Колеги, які розуміють емоційний стан одне одного, можуть більш ефективно підтримувати один одного у складні моменти, що сприяє кращій співпраці.

Емоційне розпізнавання є важливим не лише у міжособистісних та професійних стосунках, але і в аспектах лідерства. Лідери, які вміють розпізнавати емоції своєї команди, можуть мотивувати її членів та швидко орієнтуватися у стресових ситуаціях. Такий лідер може помітити низький моральний дух чи підвищений стрес у команди, вчасно надаючи підтримку, що створює позитивне робоче середовище. Крім того, під час конфліктів здатність лідера «читати» емоції інших дозволяє знайти підхід до вирішення проблем, беручи до уваги емоційний стан кожної зі сторін, що сприяє кращим рішенням [1, 23].

Самосвідомість та емоційний інтелект. Розпізнавання емоцій є центральним елементом самосвідомості, яка є важливою складовою емоційного інтелекту. Здатність розуміти й керувати своїми емоціями допомагає людині краще справлятися з життєвими викликами, зберігаючи самоконтроль та залишаючись стійким до негативних впливів, що також дозволяє визначити джерела своїх емоцій, щоб ефективніше приймати рішення й реагувати на різні ситуації (див. рис. 1.17) [24].



Рисунок 1.17 – Контроль над своїм емоційним станом через самосвідомість

Наприклад, якщо людина усвідомлює, що відчуває тривогу перед публічним виступом, вона може застосувати техніки для зниження рівня стресу, такі як глибоке дихання чи візуалізація, що допоможе впоратися з тривогою і виступити більш впевнено. Самоусвідомлення своїх емоцій також сприяє психологічному

здоров'ю, допомагаючи людям розпізнавати повторювані шаблони емоційних реакцій, що дозволяє краще розуміти, які фактори сприяють виникненню стресу або радості, і дає змогу вибудовувати здорові звички для подолання важких емоційних станів. Такими способами можуть бути фізичні вправи, медитація або пошук соціальної підтримки [2, 5, 10, 23, 24].

Дослідження показують, що люди з розвиненим емоційним інтелектом мають вищу психологічну стійкість, оскільки вони можуть з тверезістю підходити до труднощів та обирати найбільш ефективні способи вирішення проблем. Загалом, розвиток емоційного інтелекту та самосвідомості підтримує емоційне благополуччя, дозволяючи ефективніше управляти стресом, що сприяє особистому розвитку та стійкості у різних сферах життя [23, 24].

Розпізнавання емоцій у технологіях та його вплив на соціальну взаємодію. З розвитком штучного інтелекту технології, що здатні розпізнавати емоції, все більше інтегруються у різні платформи та програми, націлені на поліпшення соціальної взаємодії, що дозволяє адаптувати взаємодію з користувачами та підвищити їх задоволеність від використання цих продуктів [19].

Наприклад, у платформах для відеоконференцій, завдяки функціям розпізнавання емоцій, можна оцінити рівень залученості учасників та їхню зацікавленість, що надає доповідачам можливість коригувати свою презентацію в реальному часі, зосереджуючись на тих моментах, які викликають найбільший інтерес. Якщо ж доповідач бачить, що аудиторія стає менш зацікавленою, він може прискорити або змінити тему, забезпечуючи кращу взаємодію та задоволеність учасників [19, 22, 23, 24].

В освітньому середовищі такі технології допомагають викладачам краще розуміти емоційний стан студентів під час занять (див. рис. 1.18). Наприклад, програми з розпізнаванням емоцій можуть фіксувати, коли учні стають розгубленими або незадоволеними, що дозволяє вчителям вчасно коригувати свій підхід, адаптувати матеріал і методику викладання відповідно до потреб аудиторії. Такий підхід покращує як соціальні, так і академічні результати, створюючи сприятливе навчальне середовище [22, 25].

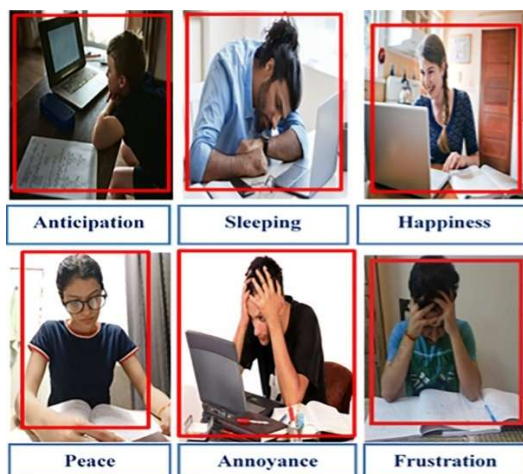


Рисунок 1.18 – Розуміння доповідачем реакцій на викладений матеріал

Технології розпізнавання емоцій корисні також у соціальних мережах та службах підтримки клієнтів. Наприклад, чат-боти, оснащені функціями розпізнавання емоцій, можуть виявити, коли користувач відчуває розчарування, і адаптувати свою відповідь, наприклад, пропонуючи додаткову допомогу чи змінюючи тон повідомлень, що значно підвищує задоволеність користувачів та забезпечує більш персоналізований досвід.

Важливо також враховувати, що такі системи викликають етичні питання, зокрема стосовно конфіденційності та можливості маніпуляцій. Використання технологій розпізнавання емоцій вимагає розробки прозорих політик щодо збору та обробки даних, щоб користувачі були поінформовані про те, як використовуються їхні емоційні реакції, і могли контролювати цей процес [2, 5, 10, 23, 24, 25].

Підвищення особистого благополуччя за допомогою технологій розпізнавання емоцій. Розпізнавання емоцій не обмежується лише вдосконаленням продуктів і послуг, а також може сприяти особистому благополуччю користувачів. Наприклад, застосунки для смартфонів та носимі пристрої на кшталт розумних годинників (див. рис. 1.19), які використовують штучний інтелект і датчики, можуть не лише відслідковувати фізичну активність та стан здоров'я, а й оцінювати емоційний стан користувача.



Рисунок 1.19 – Браслети зі датчиками для аналізу фізичних показників

Технології, що поєднують аналіз емоцій із даними, отриманими від фітнес-трекерів, дозволяють користувачам краще розуміти свій емоційний стан, виявляючи ознаки стресу, тривоги чи щастя, що надає можливість людям отримувати зворотний зв'язок, що допомагає їм своєчасно адаптувати поведінку, покращити настрій або зменшити рівень стресу за допомогою технік релаксації, медитації або соціальної підтримки [2, 25].

Особливо корисними такі технології можуть бути в галузі охорони здоров'я, зокрема психічного. Для терапевтів програми з розпізнаванням емоцій дозволяють отримувати інформацію про емоційний стан пацієнтів між сеансами, що допомагає терапевту краще зрозуміти динаміку їхнього стану, що може сприяти більш ефективній підтримці пацієнтів із хронічним стресом, тривогою або депресією, допомагаючи їм регулювати свій емоційний стан і покращувати стійкість.

Загалом, технології розпізнавання емоцій можуть значно покращити якість життя, підтримуючи емоційне благополуччя, знижуючи рівень стресу та сприяючи особистому розвитку. Інтеграція таких рішень в повсякденне життя дозволяє людям краще усвідомлювати свої емоції, вчасно вносити корективи та покращувати загальний рівень задоволеності життям.

Соціальні та психологічні проблеми у використанні технологій розпізнавання емоцій. Технології розпізнавання емоцій пропонують значні переваги у сфері взаємодії та аналізу емоційних станів, однак одночасно з цим вони породжують

низку соціальних та психологічних проблем. Однією з найвагоміших проблем є ризик неправильної інтерпретації емоцій: алгоритми розпізнавання емоцій ще не є досконалими і можуть допускати суттєві помилки, наприклад, невірно інтерпретувати вираз обличчя або тон голосу, що може призвести до ситуацій, коли емоційний стан користувача буде неправильно зрозумілим системою, а отже — людина не отримає відповідного відгуку чи підтримки [2, 5, 10, 23, 24, 27].

Також варто врахувати культурні та індивідуальні особливості людей, які можуть стати бар'єром для ефективного розпізнавання емоцій. Наприклад, представники різних культур виражають емоції по-різному, а системи часто навчаються на обмежених наборах даних, які можуть не враховувати різноманітність вираження емоцій у представників інших культур чи рас, що може призвести до непорозумінь, коли, наприклад, система, навчена на європейських зразках вираження емоцій, некоректно інтерпретує емоції представників азіатських культур, де невербальні прояви емоцій можуть бути значно стриманішими.

Подібні помилки особливо небезпечні у сферах, де точність розпізнавання емоцій має критичне значення — таких як обслуговування клієнтів або правоохоронна діяльність. У таких випадках алгоритмічні похибки можуть призвести до помилкових рішень та знизити ефективність роботи служб [19, 27].

Ще однією важливою проблемою є те, що постійне спостереження за емоційним станом людей може спричинити підвищення самосвідомості та відчуття тиску (див. рис. 1.20). Люди, знаючи, що їхні емоції відстежуються, можуть почати навмисно контролювати вираз своїх емоцій, підлаштовуючись під очікування системи, що може спричинити психологічний стрес і дискомфорт, коли люди не відчують себе вільними у вираженні власних емоцій, а вимушені підлаштовувати свої реакції під соціальні очікування. Такий контроль емоцій заради «відповідності» очікуванням системи часто негативно впливає на психічне здоров'я і може призвести до тривалого соціального та психологічного напруження [2, 27].



Рисунок 1.20 – Постійний моніторинг емоцій
може привести до психологічного тиску

Етичні наслідки розпізнавання емоцій. Етичні питання в розробці та використанні технологій розпізнавання емоцій стають дедалі важливішими, адже аналіз емоційних даних може призвести до потенційних зловживань та порушень конфіденційності. Наприклад, технології розпізнавання емоцій можуть використовуватись для моніторингу емоційного стану працівників на робочому місці, що може створити відчуття недовіри до керівництва, оскільки працівники можуть відчувати, що їх постійно «спостерігають», і це призводить до тиску й навіть до небажання працювати у такому середовищі [2, 5, 10, 28].

Подібні ризики існують також у сфері маркетингу, де компанії можуть маніпулювати емоціями споживачів для стимулювання покупок. Технології розпізнавання емоцій здатні «зчитувати» емоційний стан клієнтів і адаптувати рекламні повідомлення так, щоб вони мали більший вплив, що породжує питання: чи не є це маніпуляцією, адже таким чином споживачі можуть приймати рішення, що не відповідають їх реальним потребам.

Для запобігання таким зловживанням слід впроваджувати чіткі етичні стандарти та прозорі вказівки щодо відповідального використання технологій розпізнавання емоцій. Перш за все, користувачам мають бути зрозумілі всі процеси збору, зберігання і обробки їхніх емоційних даних. Їм також слід гарантувати повну

анонімність та конфіденційність інформації, аби виключити можливість передачі даних третім особам або їх некоректного використання.

Користувачі повинні мати контроль над своїми даними та чітко розуміти, як саме інформація про їхні емоції буде використовуватись. Для цього важливо надавати користувачам можливість вибору — чи згодні вони на збір емоційних даних, чи ні, щоб забезпечити їхню участь у цьому процесі на добровільних засадах. Такий підхід дозволить мінімізувати етичні ризики та підвищить рівень довіри до подібних технологій.

Таким чином, розробка технологій розпізнавання емоцій вимагає ретельного підходу до розв’язання соціальних, психологічних та етичних проблем. Важливо встановити чіткі межі використання таких систем, щоб технологія справді працювала на благо суспільства, а не викликала додаткове занепокоєння. Прозорі стандарти, конфіденційність та добровільна згода користувачів є ключовими складовими для розвитку та вдосконалення цієї технології у гармонії з правами та інтересами людини.

1.5 Основні виклики у вивченні емоцій та їх обробці

Розпізнавання емоцій — це складна та багатогранна сфера, яка об’єднує психологію, нейробіологію, інформатику та штучний інтелект. Завдяки прогресу технологій, сучасні системи розпізнавання емоцій можуть вже значно впливати на різні сфери життя, показуючи певний рівень успішності. Однак впровадження цих технологій залишається непростим завданням, оскільки постають суттєві проблеми, що впливають на точність та ефективність розпізнавання емоцій, а також на етичні аспекти їхнього використання [1, 2, 5, 10].

Однією з головних складностей є сама природа емоцій — вони суб’єктивні, тимчасові та мінливі, що ускладнює їх точне визначення та класифікацію [5]. Наприклад, емоція, яку певна людина сприймає як радість, може виглядати зовсім

інакше для іншої особи. Також на емоції можуть впливати культурні та соціальні особливості, що створює додатковий бар'єр для розробників таких систем. Крім того, емоції є динамічними та можуть швидко змінюватися в залежності від ситуації, що робить їхнє виявлення непостійним і ускладнює їх надійну класифікацію.

Додатковий рівень складності додають етичні міркування, зокрема питання конфіденційності даних та згоди користувачів. У деяких додатках та програмах емоційні дані можуть збиратися та оброблятися без явної згоди користувача, що викликає серйозні етичні запитання. Для забезпечення прозорості та етичності використання таких технологій необхідно враховувати питання конфіденційності та забезпечення добровільної згоди на обробку емоційних даних [1, 2, 5, 10, 28].

З метою успішного розвитку та впровадження технологій розпізнавання емоцій у різних сферах, важливо знаходити рішення для зазначених проблем, що допоможе ефективно адаптувати технологію до потреб суспільства та створити умови для її подальшого розвитку у майбутньому.

Суб'єктивна та індивідуальна природа емоцій. Вивчення та визначення емоцій є надзвичайно складним завданням, оскільки емоції глибоко суб'єктивні та залежать від індивідуального досвіду людини, що означає, що одне і те саме явище може викликати зовсім різні емоційні реакції у різних людей [28, 29]. Наприклад, ситуація, що приносить радість одній людині, може викликати тривогу у іншій через відмінності у життєвому досвіді, сприйнятті та психічному стані. Люди, що пережили травматичні події або негативний досвід, можуть у схожих ситуаціях відчувати страх чи відчай, тоді як інші, не маючи таких переживань, будуть сприймати ці ж події нейтрально або навіть позитивно.

Ця індивідуальна мінливість у вираженні та сприйнятті емоцій створює серйозні перешкоди для розробки універсальних моделей розпізнавання емоцій [29]. Більше того, кожна людина має свій унікальний спосіб вираження емоцій: деякі люди від природи мають стримані вирази обличчя, тоді як інші більш емоційно виразні, навіть коли відчують спокій. Така різноманітність в емоційних реакціях призводить до того, що штучні інтелектуальні системи, навчені на

узагальнених вибірках даних, можуть некоректно інтерпретувати емоції окремих людей. Наприклад, модель, яка «вагається» між емоціями щастя та смутку, насправді може неточно розпізнавати реальний емоційний стан через те, що не враховує індивідуальні особливості вираження емоцій.

Щоб покращити точність таких моделей, необхідно навчати їх на великих вибірках, які включають різноманітні емоційні прояви від представників різних груп та індивідуумів. Такий підхід допоможе створювати більш надійні та точні системи розпізнавання емоцій, здатні враховувати індивідуальну мінливість у вираженні та сприйнятті емоцій.

Культурні відмінності у вираженні емоцій. Крім індивідуальних особливостей, емоційне вираження значною мірою залежить від культурного контексту (див. рис. 1.21). Культура впливає на те, як люди виражають та сприймають емоції, і цей вплив може бути дуже сильним. Наприклад, у деяких східноазійських культурах прийнято демонструвати менш явні емоційні реакції, у той час як у західних культурах більш відкрите вираження емоцій є соціально прийнятним. Така культурна різниця може спричинити труднощі для моделей розпізнавання емоцій, якщо вони навчені на даних лише однієї культурної групи, а потім застосовуються до представників іншої [2, 5, 10, 30].

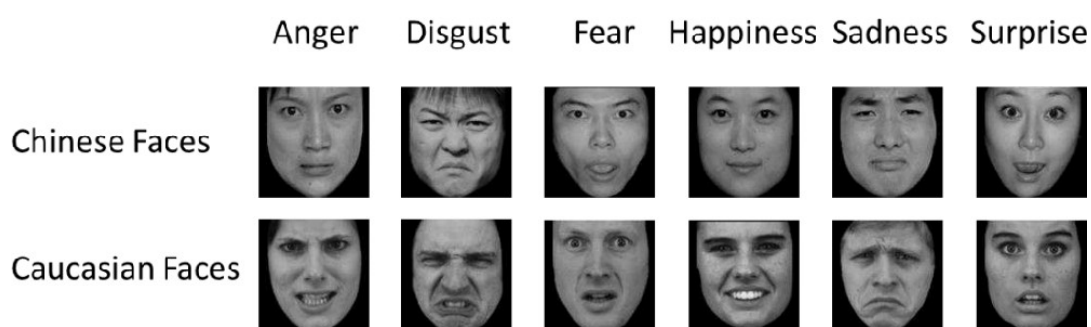


Рисунок 1.21 – Культурні особливості у визначенні емоційних станів

Культурні відмінності можуть стати причиною неправильної інтерпретації емоцій, що особливо важливо у міжнародних або міжкультурних контекстах, де некоректне розпізнавання емоцій може викликати непорозуміння або навіть образу [30]. Наприклад, емоція щастя може виглядати дещо універсальною за своїм

зовнішнім проявом, однак способи її демонстрації можуть істотно відрізнятися в залежності від культури. Тому системи розпізнавання емоцій, які не враховують культурні відмінності, можуть не лише бути неточними, але й створювати ризик культурних непорозумінь.

Для покращення точності та культурної чутливості систем розпізнавання емоцій необхідно навчати їх на наборі даних, що включає представників різних культурних груп і відображає широкий спектр емоційних проявів у різних культурних контекстах. Такий підхід дозволить штучному інтелекту краще враховувати культурні відмінності та точніше розпізнавати емоції в різноманітних соціальних і культурних ситуаціях.

Складність і неоднозначність емоцій. Емоції часто складні й неоднозначні, оскільки включають різні відтінки почуттів і швидкі переходи між ними. Людина може одночасно відчувати змішані емоції, як, наприклад, полегшення та смуток, або ж швидко змінювати емоційний стан — переходити від гніву до спокою за лічені секунди. Така мінливість ускладнює розпізнавання емоцій системами штучного інтелекту, які зазвичай класифікують емоції на основі окремих, загальних категорій, таких як щастя, смуток чи гнів.

Ще однією складністю є виявлення прихованих емоційних відтінків, таких як сарказм або легке роздратування, які складно розпізнати через їхню тонкість у вираженні. Оскільки системи розпізнавання емоцій зазвичай навчаються на узагальнених шаблонах, вони часто не враховують ці тонкощі, що може призвести до неточних результатів.

Щоб вирішити цю проблему, дослідники пропонують моделювати емоції не як окремі категорії, а за допомогою континуальних параметрів, які краще відображають їхню варіативність.

До таких параметрів належать (див. рис. 1.5) [6]:

- Валентність (полярність емоцій) — показує, чи емоція є позитивною або негативною;
- Рівень збудження — визначає інтенсивність або активність емоції;
- Домінування — показує відчуття контролю над емоцією.

Цей підхід дозволяє детальніше описувати емоції та враховувати їхні швидкі переходи, що, у свою чергу, підвищує точність класифікації. Однак створення таких моделей, здатних надійно розпізнавати складні емоційні стани, залишається важливим завданням, яке вимагає додаткових досліджень і вдосконалень у цій галузі.

Технологічні обмеження та вимоги до даних. Ефективність технологій розпізнавання емоцій значною мірою залежить від якості та обсягу вхідних даних, а створення високоточних моделей потребує великих та різноманітних наборів даних, що охоплюють широкий спектр емоційних реакцій. Збирання таких даних є викликом, оскільки емоційні реакції сильно змінюються залежно від контексту, індивідуальних особливостей та культурних відмінностей.

Щоб навчити модель точніше розпізнавати емоції, необхідно включити дані з різних джерел [1, 2, 6, 8, 9]:

- Вираз обличчя — визначення емоцій на основі міміки;
- Тон голосу — виявлення емоційного стану за змінами в голосі;
- Мова тіла — включення даних про жести та пози людини;
- Фізіологічні сигнали — вимірювання параметрів, таких як частота серцебиття.

Інтеграція всіх цих різнотипних даних для створення цілісного «емоційного профілю» є технологічно складним завданням (див. рис. 1.1) [1]. Вона потребує синхронізації та точного інтерпретування різних сигналів, що стає можливим лише за умов точних датчиків і надійних алгоритмів.

Крім того, технологічні обмеження стосуються й можливостей захоплення та обробки емоційних даних. Наприклад, при використанні камер для аналізу обличчя точність розпізнавання може знижуватися в умовах поганого освітлення, а системи, що розпізнають голос, зазнають впливу шуму з навколишнього середовища або варіацій голосів інших людей поруч [8, 27].

Щоб підвищити точність розпізнавання емоцій, необхідні вдосконалені технології сенсорів, алгоритми обробки даних та моделі машинного навчання, здатні обробляти дані з різних джерел у режимі реального часу, що особливо

важливо для застосувань, які вимагають моментального аналізу емоцій, наприклад, у сфері обслуговування клієнтів чи моніторингу психічного здоров'я [1, 2, 3, 4, 5, 9, 19, 20]. Наразі забезпечення такої реальної швидкодії залишається технічною перешкодою, оскільки обробка великого обсягу складних даних вимагає значних обчислювальних ресурсів.

Етичні наслідки та занепокоєння щодо конфіденційності. Технології розпізнавання емоцій піднімають серйозні етичні питання, які стають усе актуальнішими з огляду на те, що такі системи залучаються для збору та аналізу персональних даних. Під час використання цих технологій, особливо в громадських місцях або у сфері обслуговування, наприклад у ресторанах або магазинах, споживачі можуть навіть не здогадуватись, що їхній емоційний стан відстежується. Уявіть ситуацію, коли, заходячи в магазин, ви автоматично потрапляєте під «емоційний мікроскоп». І хоч у цьому немає прямого втручання в приватне життя, такий брак прозорості може викликати занепокоєння і навіть агресію у людей [28].

Занепокоєння також стосується ризику «емоційного спостереження», коли люди почуваються ніби під постійним контролем, і навіть у їхні емоційні стани втручаються [29]. Прикладом може слугувати маркетинг, коли компанії використовують емоційні дані для тонкої маніпуляції рішеннями споживачів, що схоже на підбір правильних «емоційних ключів» для того, щоб заохотити людину до покупки. Маркетологи можуть підлаштувати пропозицію так, щоб вона здавалася споживачу привабливішою саме в момент емоційної відкритості, підштовхуючи до рішення, яке, можливо, без цієї маніпуляції не було б ухвалено. Така практика не тільки порушує особисту свободу вибору, а й викликає занепокоєння щодо можливого маніпулювання емоціями людей.

На робочому місці розпізнавання емоцій також може призвести до етичних проблем. Використання таких технологій для відстеження настрою співробітників може бути сприйняте як надмірний контроль з боку роботодавців, що порушує межу між професійним і особистим простором [20, 29]. Наприклад, працівники можуть відчувати дискомфорт або навіть стрес, якщо знатимуть, що їхній

емоційний стан постійно моніториться для оцінки продуктивності чи лояльності, що породжує проблему вторгнення у приватне життя на робочому місці, і, щоб уникнути таких ситуацій, необхідно мати чіткі політики та правила.

Для зниження ризиків порушення конфіденційності потрібні конкретні механізми, зокрема інформована згода, яка забезпечить людям право вибору і усвідомлення того, як їхні дані будуть використовуватися. Надійні системи захисту, а також зберігання даних в анонімній формі зменшують ризики витоку інформації та сприяють підвищенню довіри до таких технологій. Наприклад, дані, зібрані анонімно, навіть у випадку витоку не можуть бути прив'язані до конкретної особи, що робить їх менш чутливими для користувачів.

Упередження в алгоритмах розпізнавання емоцій. Існує також важлива проблема, пов'язана з тим, як алгоритми розпізнавання емоцій можуть формувати упередження. Подібно до того, як індивідуальні схильності чи стереотипи можуть викривляти наші судження, так і алгоритми, які навчаються на обмежених наборах даних, можуть виявляти нерівність у трактуванні емоційних станів представників різних демографічних груп [30, 31].

Наприклад, якщо алгоритм був навчений переважно на молодих людях, він може неточно інтерпретувати емоції літніх людей, що нагадує ситуацію, коли ми намагаємося інтерпретувати іншу культуру через наші звички — часто виникають непорозуміння. Так само алгоритми, які навчені на даних із переважанням представників однієї культурної групи, можуть неправильно тлумачити емоції представників інших культур.

Це стає особливо проблематичним у серйозних сферах, як-от у правоохоронних органах або в системі охорони здоров'я, де точність та об'єктивність мають вирішальне значення. Для уникнення подібних ситуацій необхідно створювати репрезентативні набори даних, що включатимуть широкий спектр демографічних ознак: вік, стать, етнічне походження та культурний контекст. Крім того, дослідники працюють над розробкою методів виявлення та усунення упередженості в алгоритмах, що можна порівняти з процесом корекції

окулярів, коли ми підлаштовуємо скло, щоб бачити однаково чітко на будь-якій відстані.

Подолання проблем розпізнавання емоцій завдяки міждисциплінарній співпраці. Розв'язання складних проблем у сфері розпізнавання емоцій вимагає міждисциплінарного підходу, який об'єднує знання з різних галузей науки: психології, нейробіології, інформатики, етики та культурології. Такий комплексний підхід дозволяє створювати технології, які є не лише точними та надійними, але й враховують соціокультурні й етичні аспекти [1, 2, 5, 10, 13, 14, 15, 27, 29, 31]:

1. Роль психології та нейробіології. Фахівці з психології та нейробіології забезпечують фундаментальне розуміння природи емоцій. Вони досліджують, як саме емоції формуються та проявляються, а також як їх можна вимірювати за допомогою поведінкових, когнітивних та фізіологічних показників. Наприклад, знання про те, як зміни в тоні голосу або виразі обличчя відображають емоційний стан людини, допомагає розробляти алгоритми, які здатні точніше аналізувати ці дані.
2. Значення етики. Етичні міркування є ключовими для забезпечення захисту конфіденційності користувачів і дотримання їхніх прав. Експерти в галузі етики створюють рекомендації, які регулюють використання емоційних даних, зокрема, з метою захисту автономності користувачів і запобігання маніпуляціям їхніми емоціями. Наприклад, розробка політик для забезпечення інформованої згоди користувачів і впровадження механізмів анонімізації даних є важливими аспектами етичного використання технологій розпізнавання емоцій.
3. Роль культурології. Емоції та їхні прояви значно варіюються залежно від культурного контексту. Залучення експертів з культурології дозволяє враховувати ці відмінності при розробці технологій, забезпечуючи їх чутливість до культурних варіацій. Наприклад,

система повинна враховувати, що в одній культурі посмішка може бути показником радості, а в іншій — засобом приховування незадоволення.

4. Внесок інформатики та машинного навчання. Розвиток алгоритмів машинного навчання та штучного інтелекту підвищує точність і надійність моделей розпізнавання емоцій. Дослідники також працюють над створенням систем, які враховують контекст, у якому перебуває людина. Наприклад, система може аналізувати, чи перебуває людина в людному чи тихому середовищі, і відповідно коригувати свої висновки, що робить аналіз більш точним і адекватним.
5. Контекстно залежні системи. Інтеграція даних про навколишнє середовище та ситуацію стає одним із ключових напрямків у підвищенні точності розпізнавання емоцій. Наприклад, технологія, яка враховує, що людина перебуває у стресовому середовищі (наприклад, в аеропорту або під час публічного виступу), може надати точніший аналіз її емоцій, ніж система, яка враховує лише вираз обличчя чи тон голосу.
6. Інтеграція міждисциплінарних знань. Співпраця між експертами з різних сфер дозволяє створювати технології розпізнавання емоцій, які відповідають сучасним потребам. Такий підхід сприяє розробці систем, що поважають права користувачів, враховують культурне розмаїття та забезпечують точність і надійність аналізу.

РОЗДІЛ 2 ТЕХНОЛОГІЇ ТА МЕТОДИ РОЗПІЗНАВАННЯ ЕМОЦІЙ

2.1 Штучний інтелект та машинне навчання для розпізнавання емоцій

Сучасні технології штучного інтелекту (ШІ) та машинного навчання (МН) відкривають широкі можливості для розпізнавання емоцій, що стало одним із ключових напрямків використання цих технологій. Завдяки поєднанню алгоритмів ШІ та МН, можна швидко аналізувати великі обсяги даних, визначати емоційні стани людей на основі візуальної та аудіальної інформації та робити це з високою точністю.

Розпізнавання емоцій може базуватися на таких даних, як [1, 2, 6, 8, 11, 12, 15]:

- фотографії (наприклад, аналіз виразу обличчя);
- звукові сигнали (наприклад, тон голосу);
- жести та інші невербальні прояви;
- фізіологічні показники (якщо модель включає відповідні сенсори).

Для нашого дослідження основним методом стане аналіз емоцій за фотографіями, що використовуються для навчання системи. Такі технології дають змогу визначати емоційний стан людини в реальному часі та інтегрувати це в різні сфери, зокрема [1, 2, 5, 10, 13, 15, 17, 19, 20, 21, 22, 23]:

- охорона здоров'я, де емоційний стан пацієнтів може впливати на лікування;
- обслуговування клієнтів, щоб покращувати досвід споживачів;
- маркетинг, для аналізу реакцій на рекламу;
- автомобільна безпека, де важливо моніторити емоційний стан водія;
- використання МН у цьому напрямку дозволяє постійно; вдосконалювати алгоритми, роблячи їх точнішими завдяки новим даним.

Існують три основні методи машинного навчання, які використовуються для розпізнавання емоцій [4, 5, 6, 8, 11, 32]:

Навчання під наглядом (Supervised Learning). Такий метод є найпоширенішим і передбачає навчання моделей на заздалегідь підготовлених наборах даних, які містять марковані приклади. Для навчання моделі можуть використовуватись тисячі позначених фотографій, які демонструють різні емоції, такі як радість, смуток, гнів тощо. Аналізуючи такі набори, система вчиться розпізнавати характерні риси кожної емоції (наприклад, підняті брови або кутки рота), що дозволяє створювати високоточні прогнози в реальному часі, оскільки моделі базуються на великих наборах структурованих даних.

Неконтрольоване навчання (Unsupervised Learning). Такий метод застосовується у випадках, коли дані неможливо чітко маркувати або вони є складними для аналізу. Модель аналізує неструктуровані дані, виявляючи приховані шаблони, тобто як приклад можна навести, що система може ідентифікувати групи схожих виразів обличчя чи вокальних тонів, що дає змогу знаходити нові, раніше невідомі емоційні стани або шаблони у даних – це особливо корисно для складних задач, де традиційні марковані набори неможливі.

Навчання з підкріпленням (Reinforcement Learning). Такий метод є менш поширеним у сфері розпізнавання емоцій, але також має свої переваги. Моделі адаптуються на основі отриманого зворотного зв'язку, тобто як приклад можна навести, що чат-бот служби підтримки може використовувати навчання з підкріпленням, щоб краще реагувати на емоційний стан користувача (розчарування, задоволення тощо), що забезпечує постійне вдосконалення моделі в інтерактивних середовищах, таких як обслуговування клієнтів або адаптивні системи.

Для ефективного розпізнавання емоційних станів машинне навчання використовує різноманітні алгоритми та моделі – їх вибір значною мірою залежить від типу даних, які обробляються, та специфіки застосування.

Серед найбільш поширених і результативних методів можна виділити наступні [12, 14, 33]:

Згорткові нейронні мережі (Convolutional Neural Networks, CNN). CNN є однією з найпопулярніших моделей, яка широко застосовується не лише для розпізнавання емоцій, а й у багатьох інших задачах, пов'язаних із обробкою зображень. CNN може аналізувати фотографії, щоб виявити, наприклад, усмішку, підняті брови або інші ознаки емоцій. Завдяки цьому такі мережі досягають високої точності навіть на великих наборах даних.

Особливості та переваги CNN [12, 14, 33]:

- ефективність у роботі з великими обсягами даних, зокрема зображень;
- ідеально підходять для аналізу зображень, оскільки дозволяють моделі визначати особливості, такі як положення та рух рис обличчя;
- розбивають візуальні дані на ієрархічні рівні (сегменти), що підвищує точність виявлення шаблонів.

Повторювані нейронні мережі (Recurrent Neural Networks, RNN). Повторювані нейронні мережі, зокрема їх розширення — LSTM (довготривала короткочасна пам'ять), є ще одним ефективним інструментом, особливо для обробки послідовних даних. LSTM корисні для мультимодальних моделей, які можуть обробляти як зображення, так і голосові дані. У нашому випадку акцент робиться на зображеннях, однак RNN можуть бути інтегровані для аналізу голосових сигналів, що зробить систему більш комплексною [12, 33].

Особливості RNN та LSTM [33]:

- орієнтовані на роботу з даними, що змінюються з часом (наприклад, фізіологічні сигнали або голос);
- дозволяють виявляти складні емоційні особливості, такі як тон голосу, його висота, ритм та тембр.

Підтримуючі векторні машини (Support Vector Machines, SVM). На відміну від нейронних мереж, SVM є класичним алгоритмом машинного навчання, який також активно використовується для розпізнавання емоцій. Також підходять для систем, де необхідно швидко і точно класифікувати емоції на основі невеликих обсягів даних [12, 34].

Особливості та переваги SVM [34]:

- ефективні для роботи з меншими наборами даних;
- потребують менших обчислювальних ресурсів, що важливо у випадках, коли швидкість обробки є пріоритетом;
- класифікують емоції, визначаючи оптимальні межі між класами, наприклад, «щасливі» та «нейтральні».

Ансамблеві методи (Ensemble Methods). Такий підхід, при якому комбінуються кілька моделей для підвищення точності та надійності результатів, що використовують переваги різних алгоритмів. Наприклад, CNN аналізує зображення, а RNN — голосові сигнали, об'єднуючи передбачення різних моделей, що дозволяє створювати мультимодальні системи для розпізнавання емоцій [12, 34, 35].

Особливості цих ансамблевих методів [35]:

- підвищують точність визначення емоцій за рахунок комбінування даних з різних джерел (зображень, голосу тощо);
- дозволяють інтегрувати різні алгоритми в одну систему для отримання більш повних результатів.

Розпізнавання емоцій за допомогою моделей машинного навчання потребує ґрунтовного підходу до вибору джерел даних, оскільки якість і різноманітність цих даних значною мірою визначають успішність результатів. Розглянемо основні типи наборів даних, які використовуються для навчання таких моделей, а також ключові виклики, пов'язані з цим процесом.

Набори даних про вирази обличчя. Одним із найпоширеніших джерел даних для розпізнавання емоцій є зображення та відео, які демонструють вирази обличчя. Ці набори даних забезпечують базові приклади для розпізнавання виразів обличчя, допомагаючи моделям ідентифікувати ключові емоції шляхом аналізу характерних рис, таких як положення брів чи кути губ. Ще такі дані зазвичай включають фото та відео людей з різними емоційними станами, такими як радість, гнів, здивування тощо.

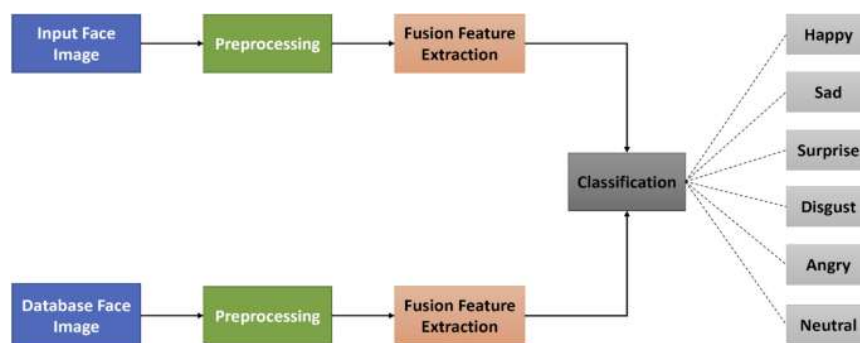


Рисунок 2.1 – Розпізнавання емоційних станів через готові набори даних (FER)

Популярні приклади наборів: FER (Facial Expression Recognition Dataset) та Extended Cohn-Kanade Dataset (СК+), які широко використовуються для аналізу міміки (див. рис. 2.1) [36].

Набори даних про мовлення та голосові емоції. Емоції також можуть бути визначені через аналіз голосу, що включає оцінку висоти, тембру, ритму та інших характеристик мовлення. Набори даних, наприклад, RAVDESS (Ryerson Audio-Visual Database of Emotional Speech and Song) та IEMOCAP (Interactive Emotional Dyadic Motion Capture Database), містять аудіозаписи голосів із різними емоційними станами, іншими словами – спеціальні мітки, що вказують, які емоції виражає мовлення (див. рис. 2.2) [37].

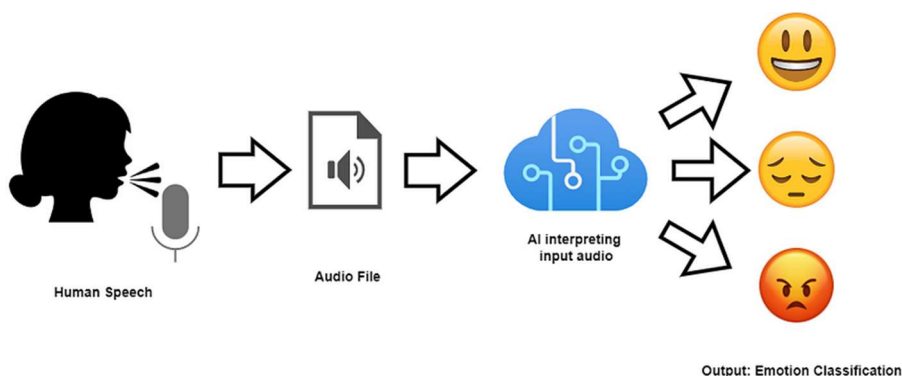


Рисунок 2.2 – Розпізнавання емоційних станів через звукову доріжку

Такі джерела дозволяють навчати моделі розпізнавати емоційні підказки в голосі, що є важливим для додатків, пов'язаних із голосовими помічниками чи аналізом дзвінків у службах підтримки.

Мультимодальні набори даних. Для досягнення високої точності в розпізнаванні емоцій дедалі частіше використовують мультимодальні набори даних, які поєднують кілька джерел – це такі дані, які комбінують інформацію з різних модальностей: зображень, аудіо, жестів і навіть фізіологічних сигналів [31, 38].

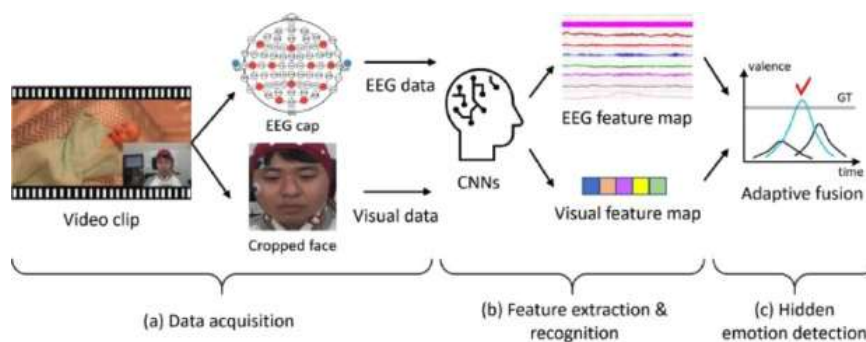


Рисунок 2.3 – Розпізнавання емоційних станів через медіадані

Такі набори допомагають створювати універсальні системи (див. рис. 2.3), здатні враховувати як міміку, так і голос, а також фізіологічні реакції (наприклад, частоту серцевих скорочень), що значно підвищує точність і застосовність моделей у різних реальних сценаріях. Відомі приклади таких наборів даних: SEWA (Sentiment and Engagement Analysis in the Wild) та AMIGOS, які включають відео, аудіо та біометричні дані [38].

Незважаючи на досягнення в галузі машинного навчання, розпізнавання емоцій все ще стикається з низкою серйозних викликів [3, 4, 39]:

Якість і різноманітність даних. Однією з ключових проблем є недостатня репрезентативність даних, які використовуються для навчання моделей:

1. Демографічна нерівність: моделі потребують даних, які враховують такі фактори, як вік, стать, етнічне та культурне походження. Якщо модель навчена переважно на даних однієї групи, вона може працювати некоректно для іншої.

2. Різноманітність емоцій: набори даних повинні охоплювати широкую гаму емоцій, щоб уникнути надмірного спрощення або неточностей.

Складність емоцій. Емоції є багатогранними, і люди часто відчувають декілька одночасно (наприклад, радість і тривогу):

1. Вплив складних емоцій: алгоритми, які покладаються на чіткі категорії, часто не здатні враховувати нюанси або змішані емоційні стани, що може призводити до спотворення результатів.

Вимоги до обробки в режимі реального часу. Багато додатків, таких як чат-боти чи системи автомобільного моніторингу, вимагають миттєвої обробки даних:

1. Обчислювальна складність: обробка великих обсягів мультимодальних даних у реальному часі вимагає значної потужності.

2. Баланс між швидкістю та точністю: для досягнення високої ефективності необхідно вирішувати питання, як забезпечити швидку обробку без втрати точності.

Розпізнавання емоцій на основі штучного інтелекту (ШІ) та машинного навчання (ML) знаходить широке застосування в різних галузях. Ця технологія значно розширює можливості аналізу емоційних станів, що може бути корисним у багатьох сферах. Розглянемо ключові напрямки використання [17, 19, 20, 21, 22]:

Охорона здоров'я та психічне здоров'я. Системи розпізнавання емоцій активно застосовуються в охороні здоров'я для підтримки психічного добробуту пацієнтів:

1. Моніторинг стану пацієнтів: технологія аналізує голосовий тон, вираз обличчя та фізіологічні сигнали, щоб виявляти ознаки стресу, тривоги чи депресії.

2. Допомога терапевтам: такі системи надають лікарям цінну інформацію, що допомагає адаптувати підхід до кожного пацієнта, забезпечуючи індивідуалізовану допомогу.

Обслуговування клієнтів і маркетинг. Розпізнавання емоцій стало потужним інструментом у сфері обслуговування клієнтів та маркетингу:

1. У обслуговуванні клієнтів: технологія дозволяє оцінювати емоційний стан клієнтів (задоволення чи розчарування) для надання більш персоналізованих і точних відповідей.

2. У маркетингу: вивчення реакцій споживачів на продукти чи рекламу допомагає створювати емоційно привабливі кампанії, що резонують із цільовою аудиторією.

Взаємодія людина-комп'ютер (HCI). Системи розпізнавання емоцій удосконалюють взаємодію між людьми та комп'ютерами:

1. Емоційно чутливі помічники: голосові асистенти, такі як Siri або Google Assistant, можуть природно реагувати на емоційний тон користувача, коригуючи свої відповіді відповідно до його настрою.

2. Покращення зручності: така адаптивність сприяє більшій задоволеності користувачів і створює комфортнішу взаємодію з технологіями.

Автомобільна промисловість. В автомобілях системи розпізнавання емоцій спрямовані на підвищення безпеки:

1. Аналіз стану водія: за виразом обличчя, рухами очей та голосовими сигналами визначаються втома, стрес чи роздратування.

2. Автоматичні дії: у разі виявлення небезпечного емоційного стану автомобіль може подавати сповіщення, активувати автопілот чи навіть зупинятися, щоб запобігти ДТП.

Правоохоронні органи та безпека. Системи розпізнавання емоцій можуть підвищити безпеку в місцях з великим скупченням людей.

1. Моніторинг емоцій: камери з вбудованими алгоритмами ШІ аналізують емоції та поведінку людей, щоб виявляти потенційні загрози (наприклад, агресивний стан).

2. Превентивні заходи: такі системи дозволяють правоохоронцям оперативно реагувати на підозрілі ситуації, знижуючи ризики фатальних інцидентів.

2.2 Комп'ютерний зір та аналіз міміки і жестів

Штучний інтелект та комп'ютерне бачення відкрили нову еру у взаємодії між людиною та машиною. Розпізнавання емоцій стало одним із важливих напрямів, який дозволяє комп'ютерам не лише бачити, але й розуміти стан людини через аналіз її міміки, жестів та інших невербальних сигналів.

Ця технологія дозволяє системам у реальному часі зчитувати емоційні сигнали за допомогою візуальної інформації (див. рис. 2.4), наприклад, з камер або датчиків. Завдяки цьому можна створити умови, в яких пристрої здатні більш природно та ефективно взаємодіяти з людьми. Застосування таких рішень є надзвичайно широким: від автоматизованих систем у роздрібній торгівлі до інтерактивного навчання, медичної діагностики та розважальних платформ.

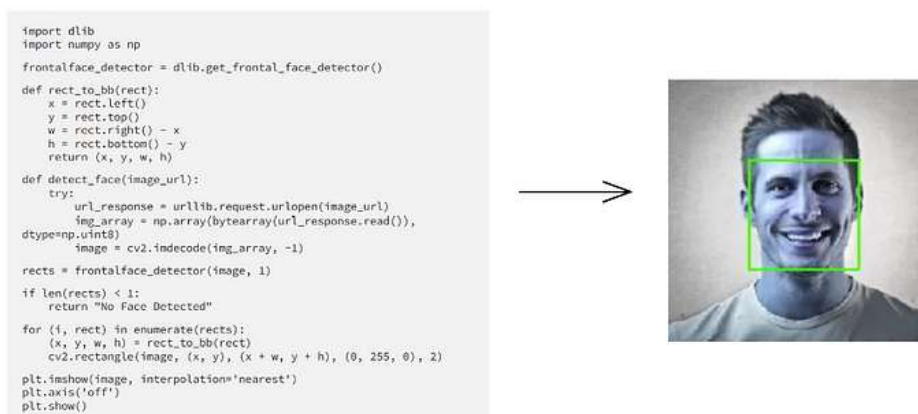


Рисунок 2.4 – Розпізнавання обличчя через Python (бібліотеки Dlib та NumPy)

Обличчя людини — це найінформативніше джерело невербальних сигналів. Навіть найменші зміни в м'язах обличчя можуть відображати складний спектр емоцій: радість, смуток, здивування, страх тощо. Для досягнення цієї мети застосовуються різні методи та підходи:

Підходи, засновані на рисах обличчя. Один із найбільш ранніх і водночас простих способів розпізнавання емоцій. У таких системах аналізуються конкретні характеристики обличчя:

- відстані між ключовими точками (наприклад, між очима, бровами, куточками рота);
- зміни у положенні цих точок, які сигналізують про певний емоційний стан.

Для цього використовуються спеціальні моделі, такі як ASM (моделі активної форми) або AAM (моделі активного вигляду), які вміють відстежувати та фіксувати зміни у формах обличчя [40]. Наприклад, підняття брів або вигин губ вгору може свідчити про здивування або радість відповідно (див. рис. 2.5).

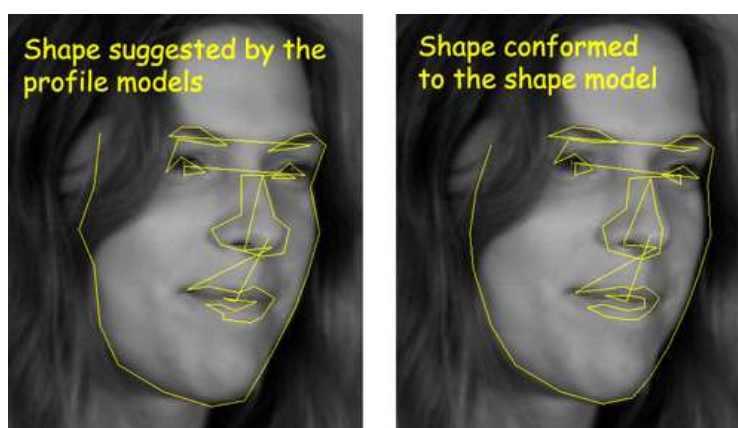


Рисунок 2.5 – Демонстрація визначення моделі активного вигляду та форми

Глибоке навчання та CNN. У наш час такі системи вийшли на новий рівень завдяки використанню глибокого навчання. Згорткові нейронні мережі (CNN) особливо ефективні у розпізнаванні обличчя, оскільки можуть автоматично виділяти ключові особливості з візуальних даних [12, 33]:

1. Вони дозволяють обробляти великі обсяги інформації та досягати високої точності без потреби у ручному налаштуванні ознак.
2. Такі мережі тренуються на величезних наборах даних, які включають тисячі зображень із різними виразами обличчя, що надає можливість системі розпізнавати навіть складні емоції з високою точністю.

Система FACS – деталізація емоційного аналізу. Пол Екман розробив систему кодування дій обличчя (FACS), яка розділяє міміку на окремі одиниці дії

(AU). Кожна одиниця відповідає конкретному руху м'язів [1, 2, 5, 10, 41] (див. рис. 2.6):

- підняття внутрішніх куточків брів (сигналізує про смуток);
- стиснення губ (може свідчити про гнів).

Upper Face Action Units					
AU 1	AU 2	AU 4	AU 5	AU 6	AU 7
Inner Brow Raiser	Outer Brow Raiser	Brow Lowerer	Upper Lid Raiser	Cheek Raiser	Lid Tightener
*AU 41	*AU 42	*AU 43	AU 44	AU 45	AU 46
Lid Droop	Slit	Eyes Closed	Squint	Blink	Wink
Lower Face Action Units					
AU 9	AU 10	AU 11	AU 12	AU 13	AU 14
Nose Wrinkler	Upper Lip Raiser	Nasolabial Deepener	Lip Corner Puller	Cheek Puffer	Dimpler
AU 15	AU 16	AU 17	AU 18	AU 20	AU 22
Lip Corner Depressor	Lower Lip Depressor	Chin Raiser	Lip Puckerer	Lip Stretcher	Lip Funneler
AU 23	AU 24	*AU 25	*AU 26	*AU 27	AU 28
Lip Tightener	Lip Pressor	Lips Part	Jaw Drop	Mouth Stretch	Lip Suck

Рисунок 2.6 – Система FACS (розпізнавання обличчя як код)

Навчені моделі, що використовують FACS, дозволяють не лише розпізнавати основні емоції, але й виявляти мікровирази — швидкоплинні, ледь помітні зміни, які відображають справжні, а не завуальовані емоції людини [41].

Жестикуляція є невід'ємною частиною людської комунікації. Рухи рук, положення тіла та нахили голови також можуть розповісти про емоційний стан. Комп'ютерне бачення дозволяє аналізувати ці жести й інтегрувати їх в загальну систему розпізнавання емоцій:

Розпізнавання на основі скелета. Тіло людини моделюється як набір точок і суглобів, які утворюють умовний «скелет». Алгоритми, як-от OpenPose, відстежують ці точки, аналізуючи жести:

- підняття руки може означати згоду;
- опускання плечей — втому або смуток.



Рисунок 2.7 – Жестикуляція (розпізнавання емоцій через жести тіла)

Такий підхід аналізує рухи між послідовними кадрами відео. Як приклад, різкі рухи рук можуть свідчити про злість, а плавні — про спокій або замисленість.

Рекурентні нейронні мережі (RNN) та приховані моделі Маркова (HMM). Оскільки жести часто є динамічними, для їх аналізу застосовуються алгоритми, які працюють із послідовностями [43]:

- HMM моделюють ймовірності переходів між станами (наприклад, перехід від зімкнених до розведених рук);
- RNN дозволяють враховувати попередні жести для кращого розуміння контексту.

Розвиток систем розпізнавання емоцій відкриває величезні можливості для різних сфер [19, 20, 21, 22, 23]:

1. Охорона здоров'я: аналіз стану пацієнтів для раннього виявлення депресії або тривоги.
2. Освіта: адаптивні навчальні системи, які змінюють підхід залежно від емоцій студентів.
3. Роздрібна торгівля: оцінка задоволеності клієнтів товарами чи обслуговуванням.
4. Розваги: створення ігор чи контенту, які реагують на емоції гравців.

Аналіз виразів обличчя та жестів є одним із важливих напрямків у сучасному розвитку комп'ютерного бачення. Ця технологія дозволяє системам машинного навчання і штучного інтелекту розпізнавати та інтерпретувати емоції, які відображаються на обличчі або передаються за допомогою рухів тіла. Метою цього підходу є створення програм, здатних працювати в реальному часі, що може принести суттєві переваги у багатьох сферах діяльності. Наприклад, це може бути корисним у медичних установах, освітніх платформах, роздрібній торгівлі, взаємодії з роботами тощо [19, 20, 21, 22, 23].

Розглянемо детальніше кілька ключових напрямків застосування технологій розпізнавання емоцій на основі аналізу обличчя та жестів [1, 2, 5, 10, 19, 20, 21, 22, 23]:

Охорона здоров'я та терапія. Використання технологій розпізнавання емоцій у медицині є перспективним і потенційно надзвичайно корисним напрямком. У сфері психічного здоров'я системи, що аналізують міміку та жести, можуть виявляти ознаки стресу, депресії чи тривоги, що надає можливість своєчасно втручатися в кризові ситуації, коли допомога найбільш необхідна.

Одним із прикладів такого використання є телемедицина. Коли пацієнт і лікар не можуть бути фізично в одному приміщенні, аналіз емоційних реакцій може допомогти медикам краще зрозуміти стан пацієнта. Наприклад, система може відстежувати міміку під час розмови й сигналізувати про будь-які емоційні зміни, що особливо актуально для пацієнтів із психічними розладами або під час консультування в кризових ситуаціях.

Освіта та електронне навчання. У сфері освіти розпізнавання емоцій може значно покращити якість навчального процесу, особливо у випадках дистанційного навчання. Наприклад, аналізуючи міміку та жести студентів, система може визначати, наскільки вони залучені до процесу, чи їм складно сприймати матеріал, чи вони розчаровані.

Такі дані можна використовувати для адаптації навчального контенту. Якщо система виявляє, що студенти неуважні або виглядають розгубленими, викладач може змінити підхід до подання матеріалу, знизити рівень складності чи надати

додаткові пояснення. Таким чином, розпізнавання емоцій може допомогти створювати персоналізовані підходи до навчання, що сприятимуть кращому засвоєнню інформації.

Роздрібна торгівля та маркетинг. У роздрібній торгівлі аналіз емоцій може стати потужним інструментом для покращення клієнтського досвіду. Наприклад, системи розпізнавання емоцій, встановлені у фізичних магазинах, можуть аналізувати реакції покупців на певні продукти, оформлення вітрин або рівень обслуговування – ця інформація може бути використана для оптимізації маркетингових кампаній. Розуміння емоцій клієнтів дозволяє створювати рекламні стратегії, які краще задовольняють потреби споживачів. Наприклад, якщо покупець демонструє зацікавленість продуктом, але водночас нерішучість, система може запропонувати йому додаткову знижку або інформацію про переваги товару, щоб мотивувати покупку.

Взаємодія людини з роботами. Сучасні роботи, особливо ті, що розроблені для соціальних або сервісних функцій, починають інтегрувати технології розпізнавання емоцій для кращої взаємодії з людьми. Аналізуючи міміку та жести користувача, роботи можуть відповідати на емоційний стан співрозмовника, демонструючи емпатію або адаптуючи свою поведінку [17], що особливо актуально для соціальних роботів, які працюють із дітьми, людьми похилого віку або пацієнтами. Такий підхід може мінімізувати дискомфорт під час спілкування з машиною, а також допоможе уникнути ефекту «моторошної долини» (див. рис. 2.8) — ситуації, коли робот виглядає і поводить майже як людина, але викликає у співрозмовника відчуття тривоги [44].



Рисунок 2.8 – Ефект моторошної долини (невдала спроба імітації рис людини)

2.3 Аналіз голосу та мовлення у визначенні емоцій

Голос і мова відіграють надзвичайно важливу роль у передачі людських емоцій – це не лише спосіб комунікації, але й інструмент, через який ми можемо краще розуміти емоційний стан людини. Завдяки своєму універсальному характеру, голос здатний не тільки передати зміст сказаного, а й донести до слухача тонкі відтінки настрою. Наприклад, навіть без слів ми можемо відчувати радість, сум чи гнів співрозмовника, просто слухаючи інтонацію його голосу.

Відмінність голосу від інших засобів комунікації, таких як жести чи міміка, полягає в тому, що він дозволяє передавати емоції як вербально, через зміст мовлення, так і невербально, через тон, висоту, ритм чи гучність. Таким чином, аналіз голосу стає ключовим інструментом для визначення емоційного стану.

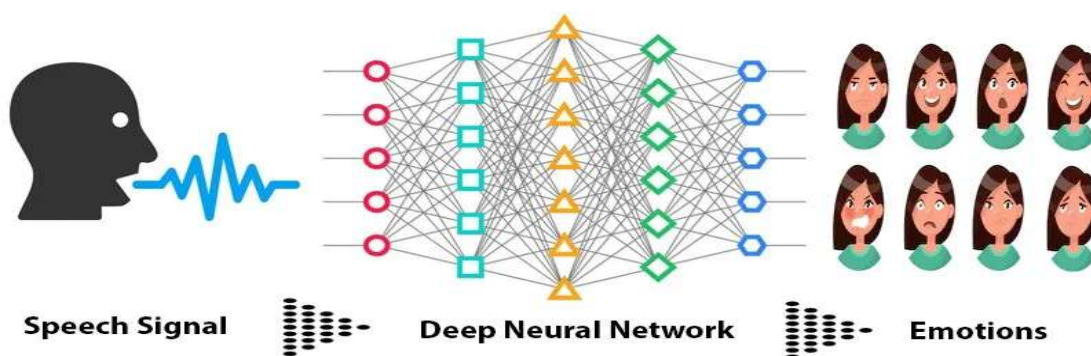


Рисунок 2.9 – Звукові характеристики для визначення емоційного стану

У процесі розпізнавання емоцій голос людини аналізується за кількома основними характеристиками (див. рис. 2.9), які допомагають виявити його емоційний стан [1, 2, 45]:

Висота та тон голосу. Висота звуку є одним із основних показників емоційного стану:

1. Високий тон зазвичай свідчить про хвилювання, радість або страх. Наприклад, коли людина схвилювана, її голос може ставати дзвінкішим і вищим.
2. Низький тон, навпаки, асоціюється зі спокоєм, сумом чи навіть упевненістю. У моменти задумливості чи втоми голос може звучати нижче і спокійніше.

Гучність і інтенсивність. Зміни в гучності голосу можуть багато розповісти про емоційний стан людини:

1. Гучний голос часто сигналізує про гнів, ентузіазм чи навіть радість. Наприклад, підвищений тон може свідчити про обурення або енергійність.
2. Тихий, приглушений голос може бути ознакою смутку, втоми або навіть невпевненості.

Швидкість мовлення. Темп мовлення є ще одним важливим показником емоцій:

1. Швидке мовлення часто супроводжує хвилювання, нетерплячість або нервозність. Коли людина хвилюється, вона може говорити швидко і без пауз.
2. Повільне мовлення, навпаки, може свідчити про сум, втому чи задумливість. Наприклад, людина, яка переживає глибокі емоції, може говорити повільно й з роздумами.

Такі акустичні характеристики є основою для аналізу голосу, особливо коли їх поєднують із вивченням змісту сказаного, що дозволяє глибше зрозуміти емоції мовця.

Техніка аналізу голосу та мови. Сучасний підхід до розпізнавання емоцій через голос включає кілька етапів (див. рис. 2.10), які базуються на передових технологіях, таких як обробка сигналів, машинне навчання та обробка природної мови (NLP) [1, 2, 8, 9, 46].

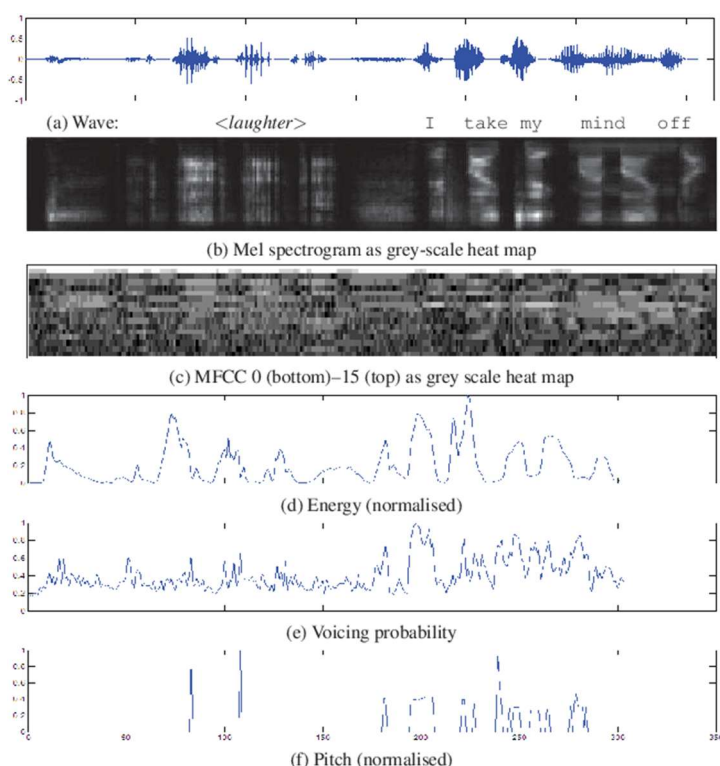


Рисунок 2.10 – Процес аналізу та покращення звукових характеристик

Вилучення акустичних ознак. Аналіз акустичних ознак передбачає вивчення таких характеристик голосу, які надають інформацію про емоційний стан людини.

До основних методів належать [46]:

1. Спектральні характеристики. Аналіз частотного спектру мовлення дозволяє виявити зміни в тоні та висоті голосу. Наприклад, гармонійні структури мовлення можуть відображати емоційні стани.
2. Енергетичні характеристики. Інтенсивність звуку, тобто гучність, може свідчити про силу емоції. Наприклад, різкі коливання енергії можуть вказувати на емоційний стрес або раптову радість.
3. Часові особливості. Вивчення пауз, ритму й темпу мовлення допомагає зрозуміти динаміку емоцій. Наприклад, часті паузи можуть бути ознакою невпевненості чи задумливості.
4. Кепстральні коефіцієнти (MFCC). Цей метод широко використовується для аналізу мовлення, адже він дозволяє стисло й точно описати спектр мовного сигналу. MFCC особливо корисні для виявлення тонких емоційних відтінків.

Підходи до машинного навчання. Для аналізу акустичних ознак і визначення емоцій використовуються різні моделі машинного навчання [12, 33, 37, 46]:

1. Машини опорних векторів (SVM). Цей метод допомагає класифікувати емоційні стани, знаходячи межі між різними класами емоцій.
2. Випадкові ліси та дерева рішень. Використовуються для обробки складних, нелінійних залежностей між даними. Наприклад, ці моделі ефективно працюють із великими наборами різноманітних ознак.
3. Глибоке навчання. Згорточні нейронні мережі (CNN) аналізують великі набори даних, виявляючи шаблони в мовних характеристиках. Рекурентні нейронні мережі (RNN), особливо моделі LSTM, обробляють послідовні дані, такі як мовлення, фіксуючи тимчасові залежності.

Обробка природної мови (NLP). Обробка природної мови зосереджується на змісті сказаного. Основні методи включають [46]:

1. Аналіз настроїв тексту. Визначення емоційного забарвлення слів і фраз. Наприклад, слова «радість» або «горе» самі по собі передають емоцію.

2. Перетворення мовлення в текст. Ця технологія дозволяє перетворити усну мову в текст, що значно полегшує її подальший аналіз.

3. Мультимодальна інтеграція. Поєднання акустичних і текстових ознак для досягнення максимальної точності виявлення емоцій.

Сучасні технології розпізнавання емоцій через голос і мову відкривають широкі можливості для інтеграції у різні сфери життя. Ця здатність аналізувати емоційний стан людини за допомогою її вокальних характеристик стає основою для вдосконалення взаємодії між людиною і комп'ютером, покращення послуг, підвищення ефективності та створення нових інноваційних рішень. Нижче розглянемо, як саме ця технологія впроваджується в різних галузях [1, 2, 45]:

Розмовний штучний інтелект (ШІ) та віртуальні помічники. Розмовні ШІ, такі як Google Assistant, Alexa або Siri, вже стали невід'ємною частиною нашого повсякденного життя. Вони використовуються для автоматизації завдань, пошуку інформації, керування розумними пристроями та багатьох інших речей. Додавання функцій розпізнавання емоцій піднімає взаємодію з такими асистентами на якісно новий рівень.

Віртуальний асистент, якому нададуть здатність аналізувати тон, гучність і ритм мовлення користувача, для того, щоб розпізнати його емоційний стан. Наприклад, якщо в голосі користувача виявляється нотка розчарування або роздратування, асистент може надати спрощене пояснення чи вибачитися за можливі незручності. Якщо користувач, перебуваючи у стресовому стані, просить пояснити функцію пристрою, асистент замість складного технічного опису може запропонувати просту покрокову інструкцію.

Переваги використання віртуальних асистентів:

- підвищення рівня задоволеності користувачів;
- ефективніше вирішення запитів завдяки адаптації до емоційного стану.

Психічне здоров'я та благополуччя. Технології розпізнавання емоцій знаходять дедалі більше застосувань у медицині, зокрема в діагностиці та лікуванні психічних розладів. Голос людини є індикатором її внутрішнього стану, тому

аналіз вокальних характеристик може допомогти виявити певні тривожні сигнали [15, 17, 20].

Емоційні маркери, які можуть бути присутні у голосі:

1. Монотонність: часто асоціюється з депресією або втомою.
2. Високий тон: може свідчити про тривогу чи збудженість.
3. Раптові зміни гучності: можуть бути ознакою стресу.

Програми для віддаленого моніторингу психічного здоров'я можуть аналізувати мовлення користувачів під час їхніх звернень до лікарів. Наприклад, якщо система фіксує стабільну монотонність у голосі протягом кількох тижнів, це може слугувати сигналом для терапевта, що пацієнт перебуває у стані депресії.

Переваги використання віддаленого моніторингу психічного здоров'я:

- раннє виявлення психологічних розладів;
- забезпечення своєчасної допомоги пацієнтам.

Обслуговування клієнтів та кол-центри. У сфері обслуговування клієнтів технології розпізнавання емоцій дозволяють значно підвищити якість послуг, що особливо актуально для роботи кол-центрів, де емоційний стан клієнта може впливати на результативність вирішення його проблеми.

Система аналізує вокальні особливості клієнта, такі як тон, гучність та швидкість мовлення, щоб визначити його настрій [19]. Наприклад, якщо клієнт говорить голосно й емоційно, це може сигналізувати про гнів або роздратування. У таких випадках система може автоматично перенаправити дзвінок до більш досвідченого оператора чи ініціювати протоколи швидкого вирішення проблем. Якщо клієнт скаржиться на недоставлене замовлення у роздратованому тоні, система може запропонувати оператору ввічливо вибачитися та відразу вирішити проблему, надавши бонус чи знижку.

Переваги використання у сфері обслуговування [19]:

- підвищення рівня довіри до бренду;
- зменшення конфліктних ситуацій.

Автомобільні системи. У сучасних автомобілях все частіше впроваджуються інтелектуальні системи, які стежать за станом водія та допомагають забезпечити його безпеку, і розпізнавання емоцій на основі голосу доповнює ці функції [21].

Система може виявляти ознаки втоми, стресу чи роздратування у голосі водія. Наприклад, якщо водій говорить уривчасто або в підвищеному тоні, автомобіль може рекомендувати йому зробити перерву.

Додаткові (можливі) функції:

- автоматичне налаштування музики чи клімат-контролю для зменшення напруження;
- сповіщення про необхідність зупинки для відпочинку.

Переваги використання у транспорті:

- зниження ризику дорожньо-транспортних пригод;
- підвищення комфорту під час поїздки.

Розваги та медіа. Технології розпізнавання емоцій знаходять своє місце і в індустрії розваг. Вони створюють нові можливості для інтерактивного контенту, персоналізації та занурення у віртуальний світ:

1. Інтерактивні відеоігри: у сучасних іграх персонажі можуть реагувати на емоційний стан гравця. Наприклад, якщо система фіксує роздратування гравця, вона може запропонувати спростити завдання чи змінити сценарій гри, щоб уникнути розчарування.

2. Медіа-сервіси: платформи, як-от стримінгові сервіси, можуть аналізувати голос користувача, щоб рекомендувати контент відповідно до його настрою. Наприклад, у разі виявлення смутку можуть пропонуватися надихаючі фільми чи музика.

Переваги використання у сфері розваг та медіа:

- підвищення залученості користувачів;
- унікальний користувацький досвід.

Розпізнавання емоцій голосу та мови є одним із найперспективніших напрямків розвитку технологій, який вже демонструє свою ефективність у багатьох галузях. Інтеграція цих систем не лише покращує якість взаємодії з користувачами,

але й відкриває нові горизонти для інновацій. У майбутньому такі технології можуть стати базовим компонентом багатьох цифрових сервісів, забезпечуючи їхню адаптацію до потреб і емоційного стану кожного користувача.

Проблеми в аналізі голосу та мовлення. Розпізнавання емоцій за допомогою голосу та мовлення — одна з перспективних галузей, що швидко розвивається. Вона обіцяє покращити взаємодію людини з комп'ютером, підвищити ефективність автоматизованих систем і відкрити нові можливості в різних сферах життя. Однак, попри всі перспективи, цей напрямок стикається з низкою проблем, які створюють бар'єри для широкого застосування технологій.

Розглянемо ключові труднощі більш детально [1, 2, 27]:

1. Культурні особливості. В одних культурах емоції можуть проявлятися відкрито, у той час як в інших — пригнічено. Наприклад, у деяких азійських культурах стриманість у вираженні емоцій вважається нормою, тоді як у західних країнах схильність до гучного сміху чи емоційного спілкування є звичною, що ускладнює створення моделей, які могли б коректно працювати в усьому світі.

2. Індивідуальні відмінності. Тембр голосу, швидкість мовлення, манера інтонації — усе це індивідуальні риси, які впливають на вираження емоцій. Наприклад, одна й та сама емоція, така як радість, у молодій дівчини може звучати високим тонким голосом, а у літнього чоловіка — низьким та спокійним.

Неоднозначність та змішані емоції. Людські емоції рідко є однозначними. Часто одна емоція накладається на іншу, створюючи складні стани, які важко класифікувати [27]:

1. Складність одночасних емоцій. Уявіть, що людина переживає нервовий сміх — явище, яке може включати одночасно тривогу, полегшення і навіть нотку радості. Таке багатогранне емоційне забарвлення часто неможливо чітко розмежувати, особливо для автоматизованих систем.

2. Змішані сигнали. Тривога може супроводжуватися підвищеною гучністю голосу або швидкістю мовлення, але це також може бути характерно для збудження. В результаті система може неправильно інтерпретувати тривожний стан як радісний.

Шум навколишнього середовища та якість запису. Одна з основних проблем аналізу голосу в реальних умовах — це фоновий шум та якість технічного обладнання [8]:

1. Вплив фонових звуків. Уявіть, що ви спілкуєтеся в переповненій кав'ярні, де чути розмови інших людей, дзвін посуду або фонову музику. У таких умовах навіть людині важко чітко розчути емоції співрозмовника, не кажучи вже про автоматизовану систему.

2. Проблеми обладнання. Недостатньо чутливий мікрофон або запис низької якості можуть сильно спотворити вокальні сигнали. Наприклад, низькочастотний фон може «маскувати» емоційні аспекти голосу, роблячи їх невпізнаними.

Обмеження набору даних. Для створення точних і надійних моделей розпізнавання емоцій потрібні великі масиви даних, але їх збір і обробка супроводжуються численними труднощами:

1. Недостатність даних. Для тренування моделей потрібні записи, які б відображали реальні емоційні стани. Проте зібрати такі дані складно, адже емоції можуть бути прихованими, зміненими або неправильно інтерпретованими навіть людьми.

2. Труднощі маркування. Емоції не завжди очевидні, і навіть експерти можуть давати різні оцінки одному й тому самому запису, що створює ризик помилкового маркування, що впливає на якість навчання моделей.

Етичні проблеми. Технології аналізу голосу, особливо ті, що спрямовані на розпізнавання емоцій, стикаються з великими етичними викликами:

1. Конфіденційність. Записи голосу містять багато чутливої інформації, яку можна використовувати для ідентифікації особи або навіть для маніпуляцій. Захист даних є критично важливим для забезпечення довіри користувачів.

2. Використання без згоди. Часто користувачі не знають, що їхній голос аналізується на емоційний стан, що може викликати обурення та створює ризики для компаній, які не дотримуються етичних стандартів.

3. Ризик зловживань. Технології можуть бути використані не лише для покращення життя, але й для маніпуляцій чи контролю. Наприклад, комерційні структури можуть використовувати розпізнавання емоцій для впливу на рішення споживачів.

2.4 Порівняння алгоритмів і підходів до розпізнавання емоцій

Розпізнавання емоцій є складним і багатогранним завданням, яке лежить на перетині психології, інформатики та обробки даних. У наш час це завдання стало ще більш актуальним завдяки значному прогресу у сферах штучного інтелекту (AI) та машинного навчання (ML). Саме ці технології відкрили нові горизонти для створення алгоритмів, здатних з високою точністю аналізувати різноманітні дані, що відображають емоційний стан людини.

Для розуміння ефективності алгоритмів важливо враховувати кілька аспектів. Залежно від типу вхідних даних, серед яких можуть бути зорові образи (вирази обличчя), аудіосигнали (інтонація голосу), текст (аналіз слів і контексту), фізіологічні сигнали (серцебиття, потовиділення) або навіть жестикуляція, алгоритми проявляють різну продуктивність. Тому порівняння підходів до розпізнавання емоцій не обмежується лише оцінкою точності. Тут варто враховувати й інші аспекти, такі як обчислювальна ефективність, адаптивність, масштабованість тощо.

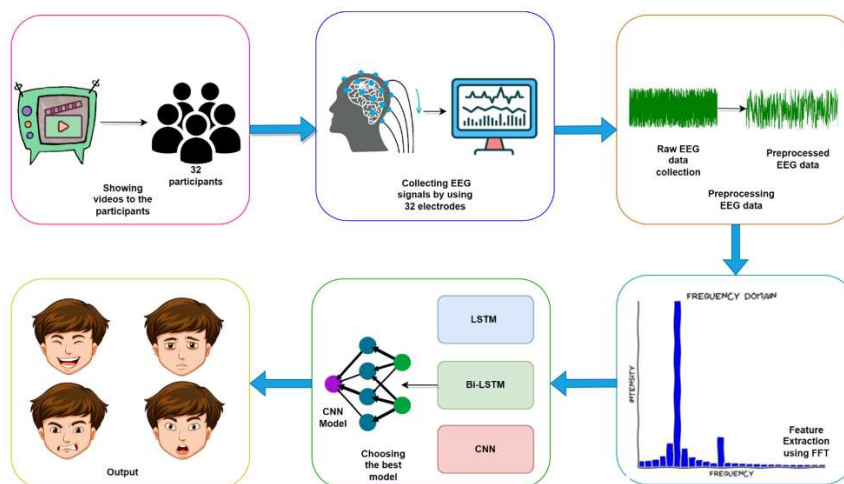


Рисунок 2.11 – Процес глибинного навчання під час визначення емоційних станів

Для детального аналізу алгоритмів розпізнавання емоцій (див. рис. 2.11) варто розглянути кілька ключових критеріїв. Ці параметри допомагають визначити, наскільки алгоритм є придатним для конкретної задачі чи умов використання [47, 48]:

Точність. Один із найважливіших показників, який характеризує здатність алгоритму правильно розпізнавати емоції. Наприклад, для алгоритмів, що працюють з виразами обличчя, точність залежить від їх здатності виявляти дрібні нюанси міміки, такі як мікровирази чи специфічні особливості рис обличчя.

У контексті тексту точність означає здатність коректно інтерпретувати емоційний підтекст слів, навіть у складних багатозначних конструкціях.

Обчислювальна ефективність. Цей параметр визначає, скільки обчислювальних ресурсів (процесорний час, оперативна пам'ять) потребує алгоритм для роботи.

В реальному часі, наприклад, у чат-ботах чи системах моніторингу настрою в роботі, обчислювальна ефективність стає критичною. Складні алгоритми, такі як моделі на основі глибоких нейронних мереж, можуть вимагати значних ресурсів, що ускладнює їхнє застосування на пристроях із низькою продуктивністю.

Адаптивність. Цей критерій показує, наскільки алгоритм здатний працювати з різними типами даних. Наприклад, система повинна однаково ефективно

аналізувати емоції як у людей із різними культурними особливостями, так і в різних умовах, таких як шум у голосових даних або слабе освітлення у відео.

Міцність. Стійкість алгоритму до зовнішніх факторів, таких як фоновий шум, змінні умови освітлення, варіації в акцентах чи індивідуальні відмінності у манері вираження емоцій, але важливо, щоб система залишалася надійною в умовах реального світу, де дані можуть бути «брудними» або неповними.

Масштабованість. Цей критерій відображає здатність алгоритму зберігати продуктивність під час роботи з великими обсягами даних або в умовах обмежених ресурсів. Наприклад, моделі, що застосовуються в глобальних соціальних мережах, повинні обробляти мільйони запитів одночасно.

Традиційні підходи машинного навчання – машини підтримки векторів (SVM). Машини підтримки векторів (SVM) є одним із класичних підходів до розв’язання задач класифікації, зокрема розпізнавання емоцій. Однією з основних переваг цього алгоритму є висока ефективність при роботі з чітко визначеними категоріями емоцій. Оскільки SVM здатні обробляти як лінійні, так і нелінійні дані, вони дуже корисні для завдань, де емоції можна класифікувати на декілька чітко визначених категорій.

Додатково, можливість застосування функцій ядра дозволяє алгоритму працювати з нелінійними даними, що є важливим при розпізнаванні емоцій, оскільки емоції можуть бути виражені через складні візуальні або акустичні патерни, які не можна безпосередньо подати як лінійні дані.

Однак, незважаючи на ці переваги, SVM мають певні обмеження, які можуть вплинути на їх ефективність у складніших задачах. Одним із головних обмежень є погана масштабованість при обробці великих наборів даних. Коли обробляється великий обсяг даних, швидкість обчислень зменшується, що робить SVM менш підходящими для таких ситуацій. Крім того, для ефективного використання цього методу необхідно ретельно налаштовувати параметри, зокрема параметр ядра, що може бути складним і часозатратним процесом.

SVM зазвичай використовуються для голосового розпізнавання емоцій, де важливі такі характеристики, як висота звуку, тембр та інші акустичні ознаки. Вони

також можуть застосовуватися для класифікації міміки обличчя, коли використовуються ручні дескриптори ознак, такі як положення точок на обличчі або форми окремих елементів обличчя.

Традиційні підходи машинного навчання – випадкові дерева рішень. Деревя рішень є інтуїтивно зрозумілими моделями (див. рис. 2.12), що дозволяють візуалізувати і розуміти процес ухвалення рішень. Вони мають високу ефективність при виявленні нелінійних залежностей між ознаками, що особливо важливо при обробці складних або багатовимірних даних, таких як фізіологічні сигнали або поведінкові патерни. Випадкові ліси, у свою чергу, дозволяють підвищити надійність та точність моделі, комбінуючи кілька дерев рішень, що допомагає зменшити вплив переобладнання та покращує стабільність результатів [48].

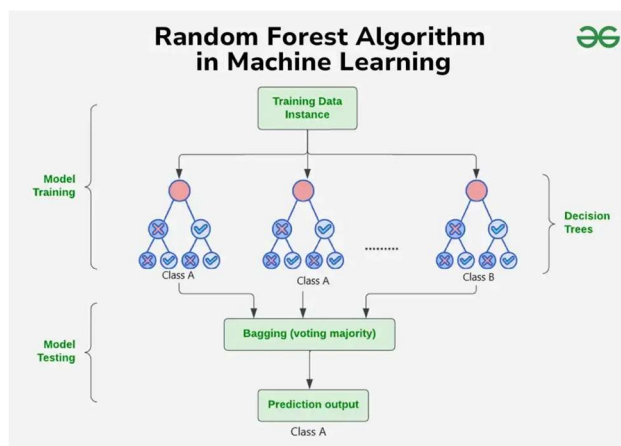


Рисунок 2.12 – Випадкові дерева рішень у глибинному навчанні

Попри їхні переваги, ці методи мають деякі обмеження. Деревя рішень можуть бути схильні до переобучення, особливо коли використовуються окремі дерева без комбінування в ліс, що може стати проблемою при обробці великих, складних наборів даних. Також, ці методи не завжди є ефективними для даних великої розмірності, коли кількість ознак значно більша, ніж кількість зразків [48].

Деревя рішень та випадкові ліси ефективно використовуються для аналізу фізіологічних сигналів, таких як пульс, потовиділення або провідність шкіри. Вони

також підходять для завдань, де необхідно класифікувати емоції за допомогою фізіологічних ознак, які можна спостерігати в реальному часі.

Традиційні підходи машинного навчання – метод k-ближчих сусідів (k-NN). Метод k-ближчих сусідів (див. рис. 2.13) є одним з найбільш простих і легких у виконанні алгоритмів, що робить його привабливим для застосування в невеликих проектах, де не потрібно складних обчислень. Метод k-NN також чудово працює з малими наборами даних, де емоційні кластери чітко визначені, що дозволяє йому ефективно класифікувати емоції, зокрема в простих задачах [49].



Рисунок 2.13 – Алгоритм k-ближчих сусідів (k-NN) у глибокому навчанні

Однак, цей метод має певні недоліки. Він є чутливим до шуму, оскільки класифікація здійснюється на основі відстаней між зразками, і якщо дані містять шум або незначні варіації, точність результатів може значно знизитися. Також, метод має високі обчислювальні витрати на великих наборах даних через необхідність обчислення відстані між кожним елементом в наборі та всіма іншими.

Метод k-NN може бути використаний для локальних досліджень емоційних моделей, таких як аналіз виразів обличчя чи вокальних моделей, де дані обмежені і мають чітку структуру, що дозволяє швидко і ефективно обробляти невеликі набори даних.

Підходи до глибокого навчання – згорткові нейронні мережі (CNN). Згорткові нейронні мережі (CNN) є потужними інструментами для обробки візуальних даних, таких як зображення та відео. Вони ефективно працюють із даними, що мають просторову структуру, і здатні автоматично вилучати ознаки з

таких даних, як текстури, форми та контури. Однією з основних переваг є те, що CNN можуть відкривати ієрархії ознак у даних, починаючи з простих (як-то краю та лінії) і переходячи до більш складних (як-то форми або об'єкти) [12, 33, 48].

З іншого боку, CNN потребують великих обчислювальних ресурсів, що може бути проблемою в умовах обмежених апаратних засобів. Для ефективного навчання таких моделей також необхідні великі марковані набори даних, що є не завжди доступним у реальних умовах.

CNN найкраще підходять для аналізу міміки та жестів, зокрема для розпізнавання емоцій за допомогою відео, де важливі деталі виразу обличчя можуть бути автоматично розпізнані та класифіковані.

Підходи до глибокого навчання – повторювані нейронні мережі (RNN). RNN розроблені для обробки послідовних даних, таких як мовні сигнали або фізіологічні дані. Завдяки здатності утримувати довготривалі залежності, вони є ідеальними для задач, де важлива інформація залежить від контексту або попередніх значень (наприклад, аналіз мовлення). Для удосконалення RNN використовуються модифікації, такі як LSTM та GRU, які ефективно працюють з довготривалими залежностями та зменшують проблему зникнення градієнтів, що характерна для стандартних RNN [12, 33, 48].

Одним із недоліків є те, що навчання моделей RNN може бути досить повільним через їхній послідовний характер обробки даних. Також, ці мережі схильні до переобладнання, особливо на невеликих наборах даних, що може призвести до низької ефективності.

RNN ідеально підходять для аналізу мовлення та фізіологічних даних, таких як серцебиття чи пульс, оскільки можуть обробляти дані, які змінюються з часом, і фіксувати важливі залежності.

Підходи до глибокого навчання – трансформери. Трансформери — це архітектура (див. рис. 2.14), яка революціонізувала обробку текстових даних завдяки механізмам самоуваги, які дозволяють мережам звертати увагу на важливі частини входу в залежності від контексту, що дозволяє їм забезпечити потужні результати при обробці великих обсягів текстових даних. Однією з ключових

переваг трансформерів є їх здатність до мультимодальної обробки даних, що дозволяє комбінувати різні типи вхідних даних, такі як текст, зображення або звук [51].

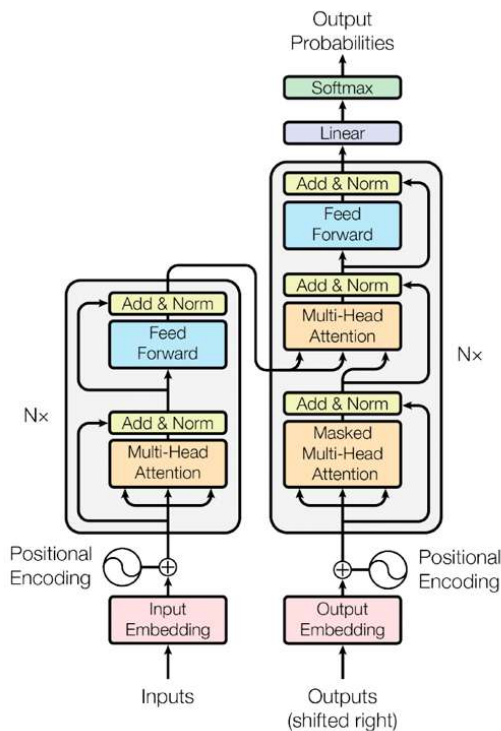


Рисунок 2.14 – Компоненти архітектури трансформерів

Основним обмеженням є високі вимоги до обчислювальних ресурсів та необхідність у великих наборах даних для навчання. Також, хоча трансформери значно підвищують точність, їх застосування може бути дорого коштувати через високу складність навчання.

Трансформери використовуються в таких завданнях, як текстовий аналіз за допомогою моделей на зразок BERT та GPT [51], що дозволяє аналізувати емоції в текстах, а також для мультимодальних задач, де важливо обробляти різні типи даних разом.

Гібридні підходи в машинному навчанні використовують комбінацію кількох алгоритмів для досягнення кращих результатів (див. рис. 2.15), ніж коли застосовуються окремо. Ідея полягає в тому, щоб поєднати переваги різних моделей, щоб компенсувати їхні слабкі сторони. Цей підхід особливо важливий у

складних задачах, де емоції виражаються через різні канали, такі як голос, вирази обличчя, текст, фізіологічні сигнали тощо. Оскільки кожен тип даних має свою специфіку та особливості, комбінування різних алгоритмів дозволяє створювати більш адаптовані та ефективні системи. Нижче розглянемо кілька прикладів таких гібридних підходів [52]:

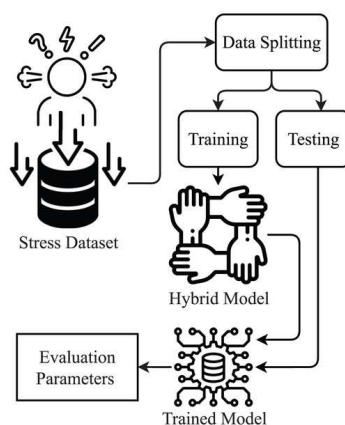


Рисунок 2.15 – Діаграма процесів у гібридній моделі

Гібридні моделі – CNN і RNN. Одним із популярних гібридних підходів є комбінація згорткових нейронних мереж (CNN) та повторюваних нейронних мереж (RNN). Обидва ці методи ефективні, але вони працюють з різними типами даних, тому їх поєднання дозволяє досягти кращих результатів [12, 33, 48, 52].

CNN ефективні у роботі з просторовими даними, такими як зображення або відео. Вони здатні автоматично витягувати важливі ознаки, такі як текстури, форми та контури, що є необхідним для аналізу таких візуальних патернів, як вирази обличчя або жести. Завдяки цьому CNN можуть відокремити різні емоційні стани за допомогою візуальних характеристик, що є важливим для розпізнавання емоцій в реальному часі [12, 33].

RNN, у свою чергу, ідеально підходять для обробки послідовних даних. Вони дозволяють моделі утримувати інформацію про попередні стани, що є важливим для обробки часових залежностей. Наприклад, при аналізі відео або фізіологічних сигналів, таких як пульс або частота серцевих скорочень, RNN можуть відстежувати зміни емоційного стану людини протягом часу [12, 33].

У гібридних моделях CNN можуть витягати просторові характеристики з візуальних даних (наприклад, вираз обличчя або жести), а RNN — обробляти тимчасову інформацію, що дозволяє точно відслідковувати емоції в динамічних умовах. Така комбінація є дуже ефективною для завдань, де потрібно обробляти відео або інші багатовимірні дані, що мають просторову та часову складову, наприклад, у системах для розпізнавання емоцій на основі відео з обличчям або голосу [12, 33].

Гібридна модель – Model Fusion. Ще один важливий гібридний підхід — це Model Fusion, або інтеграція функцій з різних типів даних (див. рис. 2.16), щоб підвищити точність і надійність результатів, що дозволяє комбінувати різні джерела інформації, які можуть доповнювати одна одну. В системах для розпізнавання емоцій, наприклад, можна використовувати такі джерела, як вирази обличчя, голос і текст [53].

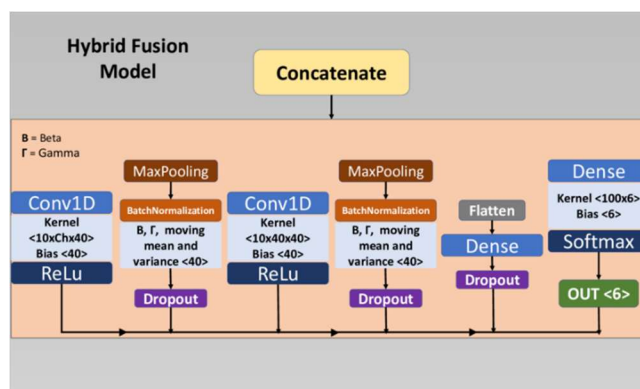


Рисунок 2.16 – Мультимодальна гібридна модель Model Fusion

Вирази обличчя надають інформацію про емоційний стан людини через зміни на обличчі (наприклад, усмішка, підняті брови, нахмурені губи), що є однією з найбільш очевидних форм вираження емоцій.

Голос може передавати емоції через інтонацію, тембр, висоту звуку, швидкість мовлення та інші акустичні характеристики. Голосовий сигнал є особливо важливим, оскільки може передавати емоції навіть у відсутності візуальних ознак.

Текст містить інформацію про емоції в явному чи прихованому вигляді, в залежності від контексту, інтонації та змісту висловленого. Наприклад, конкретні слова або фрази можуть містити явні емоційні маркери.

Комбінування цих різних типів даних за допомогою технік ф'южн дозволяє створити модель, яка враховує всі аспекти вираження емоцій, підвищуючи точність класифікації емоцій, що особливо важливо в реальних програмах, де емоції можуть передаватися через кілька каналів одночасно: наприклад, вираз обличчя, голос та текст можуть вказувати на одну й ту саму емоцію, але кожен канал може мати різний рівень достовірності. Інтеграція всіх цих джерел дає можливість створити більш надійну та адаптивну модель [53].

Висока ефективність гібридних підходів у реальних умовах. Гібридні моделі є особливо ефективними в реальних програмах, де емоції виражаються через кілька різних каналів одночасно [53]. Наприклад, у системах відеоконференцій чи у віртуальних помічниках, де людина може висловлювати свої емоції через вираз обличчя, голос та текст одночасно. У таких ситуаціях кожен канал дає додаткову інформацію про емоційний стан людини, що дозволяє створити більш точну і комплексну картину емоцій.

Наприклад, навіть якщо один з каналів (наприклад, голос) дає неповні або двозначні результати, інформація з інших каналів (наприклад, вираз обличчя або текст) може допомогти вирішити цей парадокс і уточнити, яку саме емоцію переживає людина, що дозволяє значно підвищити продуктивність та адаптованість системи до різних умов і сценаріїв.

Гібридні підходи дозволяють комбінувати сильні сторони різних алгоритмів і використовувати їх для створення високоефективних та точних систем для розпізнавання емоцій, що може бути корисним для широкого спектру застосувань: від інтерфейсів, орієнтованих на емоційне сприйняття, до медичних чи психологічних оцінок емоційного стану [53].

Вибір алгоритму для розпізнавання емоцій значною мірою залежить від специфіки завдання та умов його застосування [12, 33, 47, 48, 49, 51, 52]:

1. Для складних завдань, які вимагають високої точності та здатності адаптуватися до різноманітних джерел даних (таких як відео, аудіо, текст), найбільш ефективними є глибокі нейронні мережі та гібридні моделі. Ці моделі здатні працювати з великими обсягами даних, автоматично виділяючи важливі ознаки і забезпечуючи високу точність розпізнавання навіть в умовах шуму або неповних даних. Наприклад, глибокі нейронні мережі (CNN для зображень, RNN для часових даних) демонструють надзвичайно високі результати в таких застосуваннях, як розпізнавання емоцій через вирази обличчя, голос чи текст.

2. Для простіших завдань, де не вимагається обробка складних або великих наборів даних, а також у випадках, коли важлива висока швидкість роботи з обмеженими обчислювальними ресурсами, традиційні підходи можуть бути достатньо ефективними. Методи, як-от машини підтримки векторів (SVM), дерева рішень або k-ближчих сусідів (k-NN), можуть працювати з меншими наборами даних або в конкретних контекстах, де важливі прості, чітко визначені ознаки.

3. Зокрема, SVM добре працює з чітко розмежованими категоріями емоцій і може бути корисним у випадках, коли вхідні дані мають чітку структуру. У свою чергу, дерева рішень та випадкові ліси дозволяють отримати зрозумілі та інтерпретовані результати, що важливо для задач, що вимагають пояснення або пояснення рішень алгоритму.

4. Для ефективного застосування в реальних умовах важливо також враховувати обчислювальну ефективність, масштабованість та адаптивність моделі до різних умов. Краще вибирати ті алгоритми, що можуть забезпечити баланс між точністю та вимогами до ресурсів, що важливо, особливо у випадках використання на мобільних пристроях або в реальному часі, коли обчислювальні потужності обмежені.

У кінцевому результаті, при порівнянні, оптимальний вибір алгоритму для розпізнавання емоцій значною мірою залежить від конкретних вимог задачі та характеристик даних, з якими працює система. (Додаток А)

РОЗДІЛ 3 РОЗРОБКА ТА ВДОСКОНАЛЕННЯ МОДУЛЯ РОЗПІЗНАВАННЯ ЕМОЦІЙ

3.1 Конфігурація середовища та засоби розробки

Основним етапом у розробці будь-якої системи на основі машинного навчання, особливо такої складної, як система розпізнавання емоцій, є створення стабільного та оптимізованого середовища розробки. Першочергове налаштування такого середовища забезпечує безперебійну роботу модулів, дозволяючи уникнути конфліктів, що можуть виникнути під час виконання завдань.

Середовище розробки виступає фундаментом для реалізації кожного окремого модуля системи, зокрема модуля розпізнавання емоцій за допомогою штучного інтелекту. Саме тому необхідно ретельно налаштувати всі його компоненти, перш ніж переходити безпосередньо до розробки алгоритмів. Такий підхід дозволяє забезпечити стабільну роботу системи як на етапі створення, так і під час подальшої експлуатації.

Вибір віртуального середовища – Anaconda. Anaconda (див. рис. 3.1) — це популярний дистрибутив з відкритим вихідним кодом для мов програмування Python та R, спеціально створений для роботи з даними. Вона включає менеджер пакетів та середовищ Conda, інтеграцію з Jupyter Notebook та широкий набір інструментів для обробки даних. Завдяки цьому Anaconda дозволяє створювати ізольовані віртуальні середовища, що гарантує стабільність роботи проектів і унеможливорює конфлікти між бібліотеками [54] (Додаток А).

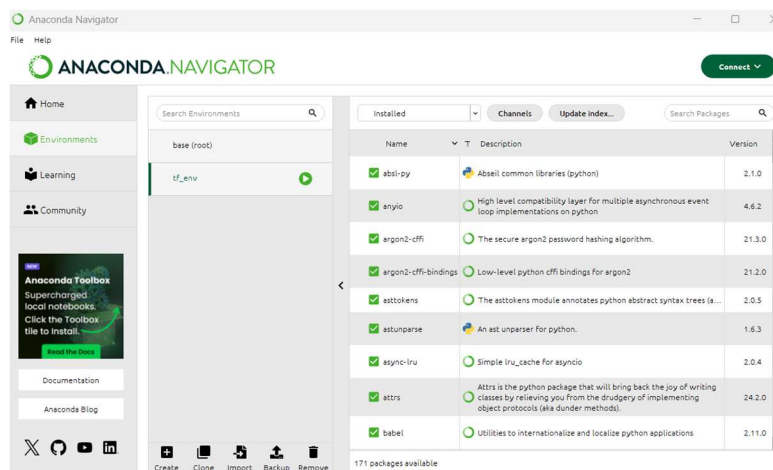


Рисунок 3.1 – ПЗ для створення віртуального середовища – Anaconda Navigator

Віртуальні середовища забезпечують ізольований робочий простір, де можна налаштувати окремі версії Python та необхідні бібліотеки для кожного проекту, що особливо важливо при розробці модуля розпізнавання емоцій за допомогою штучного інтелекту, оскільки:

- уникаються конфлікти між бібліотеками різних проектів;
- спрощується налагодження та управління залежностями;
- гарантується відтворюваність експериментів;
- створення віртуального середовища в Anaconda Navigator.

Для розробки модуля розпізнавання емоцій необхідно створити окреме віртуальне середовище – через графічний інтерфейс Anaconda Navigator або за допомогою командного рядка Anaconda Prompt (див. рис. 3.2) [54] (Додаток Б).

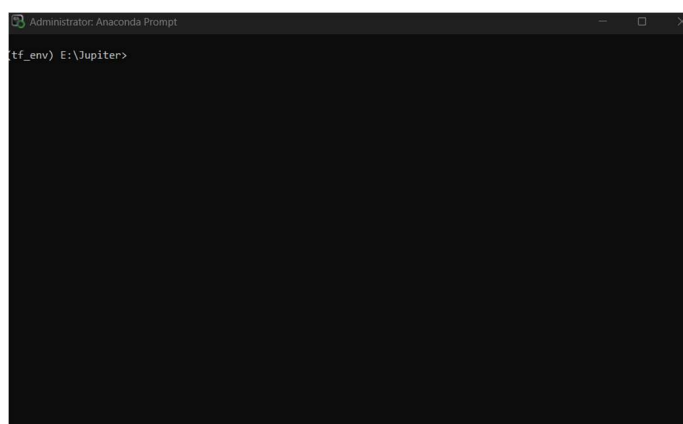


Рисунок 3.2 – Командний рядок Anaconda Prompt

Створення середовища:

```
conda create --name tf_env python=3.10
```

У цьому прикладі середовище називається `tf_env`, але його можна назвати будь-як, наприклад, `emotion_detection_env`.

Активація середовища:

```
conda activate tf_env
```

Встановлення необхідних бібліотек. Після створення та активації середовища потрібно встановити бібліотеки, які знадобляться для роботи з модулем розпізнавання емоцій [56, 57, 58, 59]:

1. TensorFlow та Keras: Для побудови та навчання нейронних мереж.

```
conda install tensorflow keras
```

2. OpenCV: Для роботи з відеопотоком і обробки зображень.

```
conda install -c conda-forge opencv
```

3. NumPy: Для числових обчислень та роботи з масивами даних.

```
conda install numpy
```

4. Matplotlib: Для візуалізації даних.

```
conda install matplotlib
```

5. OS: Вбудований модуль Python для роботи з файловою системою.

Підготовка робочого простору. Після налаштування середовища необхідно організувати проектну структуру (Додаток Б):

- створити папку для зберігання скриптів та моделей.
- налаштувати Jupyter Notebook для зручного тестування коду: `conda install jupyter`

Переваги використання Anaconda:

- легке управління залежностями;
- ізоляція середовищ;
- зручна інтеграція з популярними інструментами.

Вибір засобу розробки – Jupyter Notebook. Для розробки модуля розпізнавання емоційних станів за допомогою штучного інтелекту ми обрали Jupyter Notebook як основний інструмент. Він є частиною середовища Anaconda

Navigator і забезпечує зручність у кодуванні, тестуванні та експериментуванні з моделями (див. рис. 3.3).

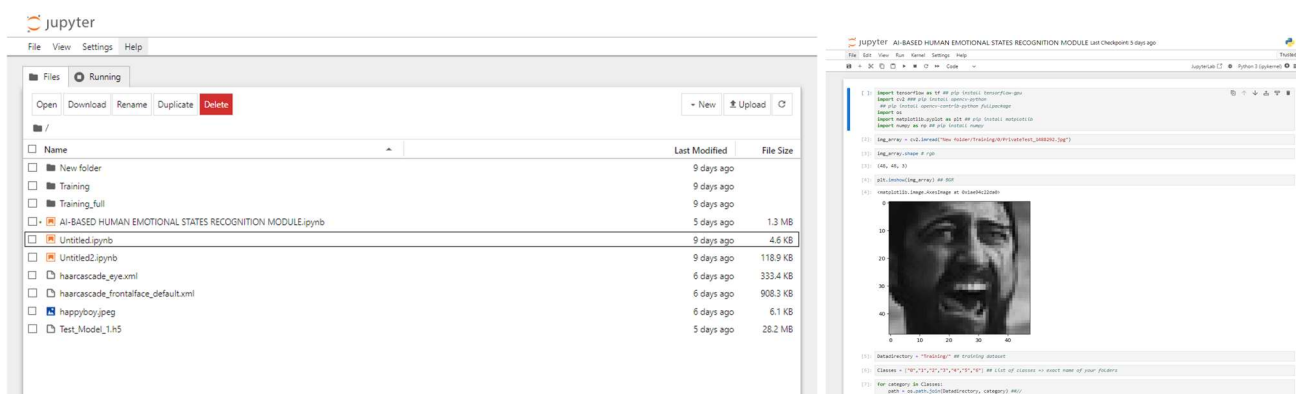


Рисунок 3.3 – Інструмент для розробки – Jupyter Notebook

Jupyter Notebook — це інтерактивне середовище для виконання коду, яке дозволяє розбивати його на окремі блоки (клітинки). Цей формат має кілька ключових переваг [55]:

1. Сегментоване виконання коду. Код можна розбити на логічні частини, які виконуються незалежно, що дозволяє бачити результати негайно після виконання кожного блоку, що значно полегшує налагодження. Помилки також з'являються миттєво, що спрощує пошук і виправлення проблем.
2. Візуалізація даних. Jupyter Notebook дозволяє вбудовувати графіки та інші візуальні елементи безпосередньо в середовище. Наприклад, за допомогою бібліотеки Matplotlib можна створювати графіки для аналізу результатів роботи моделі або попередньої обробки даних.
3. Документація в коді. Підтримка Markdown дозволяє додавати текстові коментарі, заголовки та інструкції безпосередньо до коду, що забезпечує зручне структурування проекту та полегшує розуміння коду іншими розробниками. В результаті, ваші проекти стають більш зрозумілими та легко підтримуваними.

Розробка модуля включає кілька важливих етапів:

1. Підготовка та обробка даних. На початковому етапі потрібно підготувати набір даних, який буде використовуватися для навчання моделі. Попередня обробка даних може включати зміну розмірів зображень, нормалізацію або виправлення аномалій.
2. Навчання моделі. Під час навчання моделі важливо слідкувати за графіками втрат і точності. Вони також можуть бути візуалізовані в Jupyter Notebook, що дозволяє швидко оцінити ефективність навчання та вносити корективи.
3. Налагодження. У процесі розробки можуть виникати різні проблеми, наприклад, пов'язані з неправильною обробкою даних. Завдяки інтерактивному середовищу Jupyter, виправлення таких помилок стає більш гнучким і оперативним.

Інтеграція Jupyter Notebook з Anaconda Navigator. Використання Jupyter Notebook у зв'язці з Anaconda Navigator надає додаткові зручності [54, 55]:

- можна легко керувати середовищами та запускати інструменти через графічний інтерфейс;
- спрощує роботу для тих, хто не хоче використовувати командний рядок, як наприклад, створення та активація віртуальних середовищ можна виконати через інтерфейс Navigator, що значно прискорює процес і знижує кількість помилок.

Таке поєднання засобів дозволяє не лише спростити розробку, а й зробити її більш ефективною та організованою.

Вибір плагінів та залежностей. Для ефективної розробки модуля розпізнавання емоційних станів за допомогою штучного інтелекту необхідно використовувати плагіни та бібліотеки, які доповнюють базові інструменти. Вони забезпечують виконання складних операцій, таких як виявлення облич та створення, навчання і функціонування нейронних мереж.

Основні залежності для нейронних мереж – TensorFlow і Keras. Однією з ключових бібліотек є TensorFlow з інтегрованим API Keras (див. рис. 3.4), яка надає потужний інструментарій для створення нейронних мереж [56]:

1. Створення та навчання нейронних мереж: TensorFlow підтримує як створення моделей «з нуля», так і використання попередньо навчених мереж, таких як MobileNetV2, що дозволяє швидко адаптувати моделі для завдань розпізнавання емоцій.

2. Прискорення роботи: Завдяки оптимізації для GPU, TensorFlow забезпечує швидке навчання моделей та виконання прогнозів, що суттєво скорочує час розробки та дозволяє працювати в реальному часі.



Рисунок 3.4 – Поєднання залежностей – TensorFlow та API Keras

Інструменти комп'ютерного зору – OpenCV та Haar Cascades. Для виявлення облич і попередньої обробки зображень використовується бібліотека OpenCV (див. рис. 3.5). Вона надає широкий набір алгоритмів для роботи із зображеннями та відео [57, 61]:

1. Виявлення облич: Використовується класифікатор Haar Cascade (наприклад, файл haarcascade_frontalface_default.xml), який аналізує зображення у градаціях сірого для ідентифікації облич.

2. Реальний час: OpenCV дозволяє аналізувати відеопотоки, наприклад, з вебкамери, що робить можливим розпізнавання емоцій «на льоту». Виявлені обличчя передаються в модель для подальшого аналізу емоцій.

3. Швидкість і ефективність: Завдяки оптимізації алгоритмів OpenCV забезпечує швидке виконання завдань навіть на обмежених ресурсах.

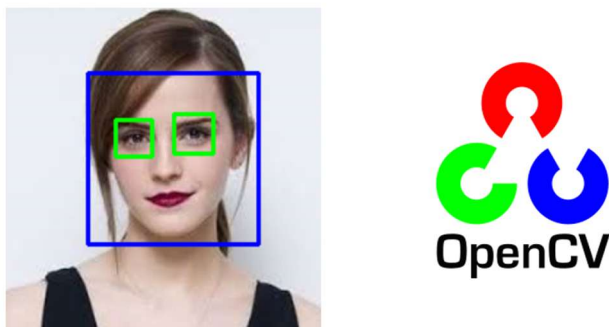


Рисунок 3.5 – Інструменти комп’ютерного зору – OpenCV та Haar Cascades

Підтримуючі бібліотеки для обробки даних та візуалізації [58, 59] (див. рис. 3.6):

1. NumPy. Ця бібліотека є основою для роботи з масивами даних. Її потужні функції обробки матриць дозволяють ефективно виконувати такі завдання, як зміна розміру зображень, нормалізація та пакетування даних для навчання моделей.

2. Matplotlib. Використовується для візуалізації даних і результатів навчання моделі. За допомогою графіків можна оцінити динаміку втрат і точності по епохах, що допомагає вчасно виявити проблеми та вдосконалити модель.



Рисунок 3.6 – Додаткові бібліотеки – NumPy та Matplotlib

3.2 Підготовка набору даних і методи попередньої обробки

Після створення та налаштування віртуального середовища з усіма необхідними бібліотеками, плагінами та залежностями, важливим наступним кроком є підготовка даних для навчання моделі розпізнавання емоційних станів. Цей етап є критичним для забезпечення точності та продуктивності системи штучного інтелекту.

Щоб модель ефективно розпізнавала емоції, необхідно використовувати великий, різноманітний і якісно підготовлений набір даних. Використання обмежених або занадто вузьких наборів, таких як власні фотографії або зображення близьких, не дасть бажаних результатів, що призведе до:

1. Поганої генералізації. Модель буде добре працювати лише з тими даними, на яких вона навчалась, і показуватиме низьку точність на нових обличчях або в різних умовах.

2. Демографічних обмежень. Використання даних лише з одного середовища або етнічної групи може призвести до упередженості моделі. Наприклад, модель, навчена на обличчях однієї етнічної групи, може некоректно працювати з іншими.

Щоб забезпечити широку репрезентативність і покращити якість моделі, доцільно використовувати загальнодоступні набори даних, які вже перевірені часом та ефективно застосовуються в інших проектах. Такі набори містять тисячі або навіть мільйони зображень людей з різними емоціями, що дозволяє моделі [62, 63]:

1. Навчатися на різноманітних прикладах. Це дає можливість враховувати різні вирази облич, освітлення та ракурси.
2. Узагальнювати знання. Модель буде краще адаптована до нових даних і ситуацій, що підвищить її точність на практиці.

Належна підготовка даних передбачає кілька важливих етапів [62]:

1. Очищення. Видалення шумів, некоректних або повторюваних зображень.
2. Нормалізація. Приведення даних до єдиного формату та масштабу, що покращить стабільність навчання.
3. Аугментація. Збільшення різноманітності набору даних шляхом штучного створення нових прикладів (зміна освітлення, обертання тощо).

Правильна підготовка даних забезпечує [62]:

1. Послідовність та чистоту. Всі дані будуть стандартизовані, що покращить процес навчання.
2. Високу продуктивність моделі. Модель буде точніше розпізнавати емоції, навіть на нових або складних даних.
3. Узагальнення. Модель зможе добре працювати з різними типами обличч і в різних умовах, що мінімізує ризик помилок.

Вибір набору даних для розпізнавання емоцій – FER-2013. Набір даних FER-2013 (Facial Expression Recognition 2013) (див рис. 3.7) є широко використовуваним ресурсом для навчання моделей розпізнавання емоцій на обличчі. Завдяки своїй різноманітності та кількості зображень, він надає міцну основу для створення ефективних систем. Використання цього набору даних дозволяє уникнути проблем, які виникають при створенні власних малих датасетів, що можуть бути обмеженими за кількістю та різноманітністю зображень [62].

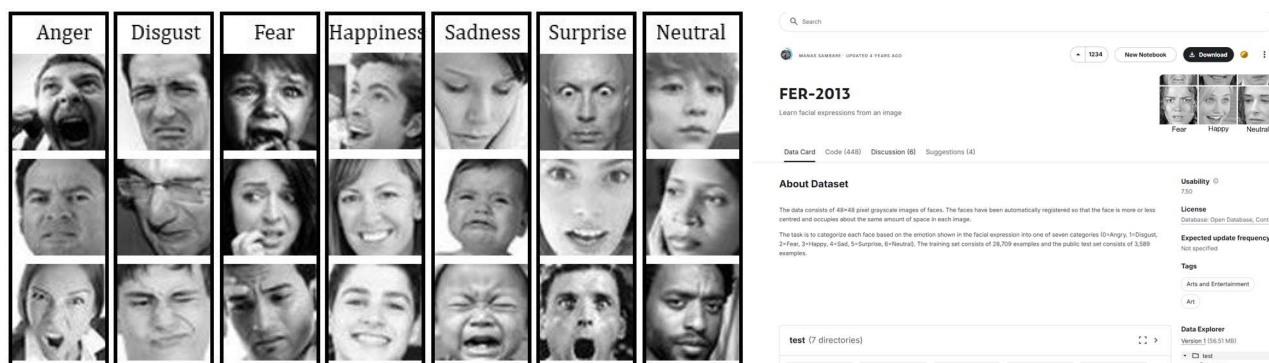


Рисунок 3.7 – Приклад вибірки зображень та ресурс з набором даних – FER-2013

Структура та особливості набору FER-2013 [62]:

1. Кількість зображень. Набір містить 35 887 зображень у градаціях сірого, кожне з роздільною здатністю 48x48 пікселів. Незважаючи на невелику роздільну здатність, цього достатньо для задач розпізнавання емоцій, адже ключовими є мімічні особливості обличчя, а не дрібні деталі.
2. Категорії емоцій (див. рис. 3.7). Всі зображення поділені на сім класів емоцій [1, 62]:
 - Angry (злість);
 - Disgust (огода);
 - Fear (страх);
 - Happy (радість);
 - Sad (смуток);
 - Surprise (здивування);
 - Neutral (нейтральний стан).

Для зручності під час навчання, класи часто позначаються цифрами від 0 до 6.

Розподіл даних відбувається наступним чином:

- 80% даних використовується для навчання моделі;
- 10% — для валідації;
- 10% — для тестування.

Переваги використання FER-2013 [1, 2, 62]:

1. Різноманітність: Набір включає зображення людей різного віку, статі та етнічного походження, що дозволяє створювати моделі, здатні працювати з різними групами населення.
2. Анотація: Всі зображення ретельно розмічені за категоріями емоцій, що полегшує навчання моделі.
3. Доступність: FER-2013 є загальнодоступним і може використовуватися для наукових досліджень та експериментів.

Проблеми та обмеження FER-2013 [1, 2, 27, 62]:

1. Дисбаланс класів. Деякі категорії, як-от Disgust (огода), представлені значно меншою кількістю зображень, ніж інші, такі як Happy (радість) чи Neutral

(нейтральний стан) , що може призводити до перенавчання моделі на частіших класах і помилок під час розпізнавання менш поширених емоцій.

2. Низька роздільна здатність. Зображення мають розмір 48x48 пікселів, що обмежує здатність моделі враховувати дрібні деталі обличчя. Хоча це спрощує обчислення, моделі можуть втрачати важливу інформацію, особливо для більш тонких емоційних станів.

3. Відтінки сірого. Використання чорно-білих зображень може ускладнювати розпізнавання емоцій, пов'язаних з фізіологічними змінами (наприклад, почервоніння обличчя під час злості чи зніяковіння).

Для створення ефективної моделі розпізнавання емоційних станів на основі набору даних FER-2013 необхідна ретельна підготовка зображень, що забезпечує сумісність із моделлю машинного навчання та підвищує якість навчання.

Попередня обробка включає кілька важливих етапів [60, 62]:

1. Зміна розміру зображень. Зображення в наборі FER-2013 мають розмір 48x48 пікселів, що є недостатнім для сучасних моделей. Щоб забезпечити відповідність вимогам до вхідних даних для моделі MobileNetV2, всі зображення масштабуються до 224x224 пікселів.

2. Захоплювати більше деталей. Збільшення розміру сприяє кращому розпізнаванню мікроекспресій і дрібних змін на обличчі.

3. Покращити точність. Великі зображення дозволяють моделі навчитися розпізнавати різні емоції з більшою точністю.

Для зміни розміру використовується бібліотека OpenCV і функція `cv2.resize()`, яка забезпечує рівномірне масштабування всіх зображень. Крім того, дотримується співвідношення сторін, щоб уникнути спотворення облич, що важливо для коректного навчання [57].

Для покращення навчання моделі всі значення пікселів масштабуються в діапазон від 0 до 1, що досягається шляхом ділення кожного пікселя (в діапазоні 0-255) на 255 [57, 62].

Переваги нормалізації:

1. Швидша збіжність. Нормалізовані дані дозволяють мережі швидше навчатися, оскільки усуваються великі різниці між значеннями пікселів.
2. Запобігання чисельним нестабільностям, що зменшує ризик помилок при обчисленні градієнтів і прискорює навчання.

Цей процес особливо важливий для узгодження з попередньо навченими моделями, такими як MobileNetV2, які очікують нормалізовані дані на вході.

Щоб підвищити стійкість моделі та компенсувати дисбаланс класів у наборі FER-2013, застосовується збільшення даних, що дозволяє моделі бути стійкішою до різних варіацій зображень і включає такі техніки [62]:

1. Горизонтальне відображення. Імітує ефект дзеркального відображення, що додає різноманітності даних.
2. Випадкові повороти. Обертання зображень на різні кути дозволяє моделі адаптуватися до різних ракурсів.
3. Масштабування та кадрування. Імітація різної віддаленості камери покращує стійкість до змін відстані між камерою та обличчям.

Для реалізації цих методів використовуються вбудовані функції бібліотек TensorFlow та Keras, що дозволяють виконувати збільшення даних «на льоту» під час навчання [56].

Вибір інструменту для локалізації облич – OpenCV та Haar Cascade. Для ефективного розпізнавання емоцій важливо не просто використовувати повне зображення, а сфокусуватися на обличчі. Локалізація облич (див. рис. 3.8) дозволяє моделі зосереджуватися на релевантних ділянках, що значно покращує точність розпізнавання емоцій, що також допомагає зменшити вплив різних фонових об'єктів, які можуть створювати шум і плутанину для алгоритму.

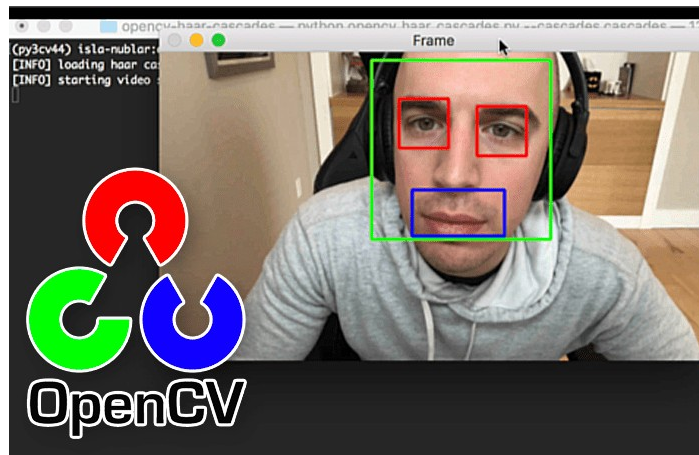


Рисунок 3.8 – Визначення обличчя за допомогою OpenCV та Haar Cascades

Локалізація виконується з використанням методу каскадів Хаара (Haar Cascade), реалізованого в бібліотеці OpenCV. Основні етапи цього процесу [57, 61]:

1. Оскільки каскади Хаара працюють ефективніше з чорно-білими зображеннями, спочатку зображення конвертується за допомогою функції «cv2.cvtColor()».
2. Класифікатор переглядає зображення, шукаючи області, що відповідають шаблонам облич, що робиться за допомогою вікна, яке переміщується по зображенню та оцінює кожен регіон.
3. Після знаходження облич регіони обрізаються та масштабуються до розміру 224x224 пікселів для подальшого введення в модель, що забезпечує сумісність із попередньо навченою моделлю MobileNetV2.

Переваги використання Haar Cascade [61]:

1. Легкість та швидкість. Каскад Хаара є дуже ефективним і легким алгоритмом, що робить його ідеальним для додатків у реальному часі. Він швидко знаходить обличчя навіть на пристроях з обмеженими ресурсами.
2. Оптимізація обчислень. Обробка чорно-білих зображень знижує обчислювальну складність, що дозволяє швидше працювати з даними.

Проблеми та їх рішення [61]:

1. Помилкові спрацьовування. Алгоритм може помилково ідентифікувати як обличчя інші об'єкти. Для зменшення кількості помилок використовуються параметри:

- Масштабний коефіцієнт (scale factor) — контролює, наскільки зменшується зображення під час пошуку облич.
- Мінімальний розмір об'єкта — визначає мінімальні розміри обличчя, які будуть виявлені.

2. Частково видимі або нахилені обличчя. Обличчя, які знаходяться за межами кадру або сильно нахилені, можуть не бути виявлені. Цю проблему частково вирішує розширення даних (data augmentation), яке дозволяє моделі краще адаптуватися до різних положень обличчя.

У процесі попередньої обробки зображень виникають кілька важливих викликів, які необхідно врахувати для покращення роботи моделі розпізнавання емоцій.

Перша проблема – низька роздільна здатність зображень. Більшість навчальних наборів даних для розпізнавання емоцій містять зображення низької якості, наприклад, із роздільною здатністю 48x48 пікселів (див. рис. 3.9), що обмежує здатність моделі отримувати достатньо інформації для точного розпізнавання емоцій [62].



Рисунок 3.9 – Низька роздільна здатність FER-2013 (48x48 пікселів)

Для вирішення цієї проблеми зображення масштабуються до розміру 224x224 пікселів, що відповідає вимогам моделі MobileNetV2. Хоча це не підвищує деталізацію зображення, використання попередньо навчених фільтрів у MobileNetV2 дозволяє моделі адаптуватися до грубих і розмитих даних, що покращує її стійкість до низькоякісних вхідних зображень [60].

Друга проблема – дисбаланс класів емоцій. Зображення можуть нерівномірно представляти різні емоції. Наприклад, емоції на кшталт радості або смутку зустрічаються частіше, ніж менш поширені, як-от огида (disgust), що призводить до того, що модель неефективно навчається на рідкісних класах, що погіршує її точність у визначенні таких емоцій (див. рис. 3.10) [1, 2, 27].



Рисунок 3.10 – Дисбаланс класів емоцій

Для усунення дисбалансу класів застосовується збільшення даних (data augmentation) та надмірна вибірка (oversampling) для рідкісних класів, тобто означає створення саме додаткових зображень із рідкісними емоціями шляхом трансформацій, таких як повороти, дзеркальні відображення, зміна освітлення тощо, що дозволяє моделі краще розпізнавати всі типи емоцій, навіть якщо деякі з них представлені у наборі даних рідше.

Третя проблема – ефективність обробки зображень. Попередня обробка великих обсягів зображень може спричиняти значне навантаження на пам'ять і сповільнювати роботу системи [8].

Для зберігання та обробки зображень використовується бібліотека NumPy, яка дозволяє оптимально працювати з масивами даних. Завдяки цьому мінімізується затримка при обробці зображень, а витрати пам'яті значно скорочуються, що особливо важливо для реального часу або роботи з великими обсягами даних [58].

3.3 Розробка моделі та навчання через трансферне навчання

Розпізнавання емоцій за виразами обличчя є складним завданням, яке потребує ретельно підготовлених шаблонів для обробки вхідних зображень. У цьому контексті було застосовано перехідне навчання, що дозволило ефективніше використовувати функції вилучення ознак за допомогою архітектури MobileNetV2. Ця модель була попередньо навчена на наборі даних ImageNet, а потім адаптована для розпізнавання емоцій. Такий підхід скоротив час розробки модуля, підвищив точність системи завдяки вже наявному досвіду моделі. Для розуміння цього процесу важливо знати, як побудована модель, її конфігурація та основні етапи навчання.

Трансферне навчання як основа підходу. Трансферне навчання — це метод машинного навчання, який передбачає використання попередньо натренованої моделі для виконання нового завдання. У нашому випадку обрано архітектуру MobileNetV2, яка зарекомендувала себе як ефективна та продуктивна у задачах класифікації зображень. MobileNetV2 — це легка архітектура, спеціально оптимізована для мобільних пристроїв та середовищ із обмеженими ресурсами. Вона забезпечує ефективні обчислення без втрати точності, використовуючи техніки, як-от глибинні згортки для зменшення кількості параметрів [60].

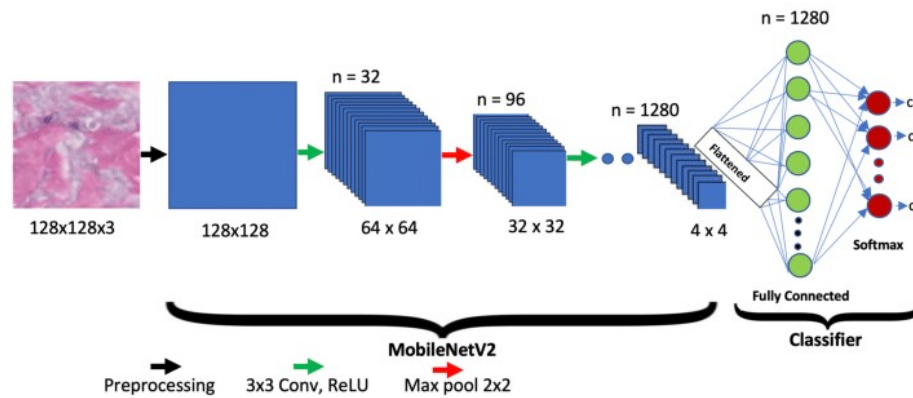


Рисунок 3.11 – Мережева архітектура MobileNetV2

Ця модель, попередньо навчена на наборі даних ImageNet [63], здатна виділяти характеристики, такі як форми, краї та текстури, що робить її придатною для адаптації до різних завдань.

Для завдання розпізнавання емоцій базові шари MobileNetV2, які відповідають за вилучення ознак із низьких та середніх рівнів, були «заморожені», що дозволило зберегти можливості вилучення функцій, отримані на попередньому наборі даних ImageNet (див. рис. 3.12), і спрямувати обчислювальні ресурси на навчання верхніх шарів, які відповідають за специфіку нового завдання [60, 63].



Рисунок 3.12 – Надвеликий набір даних ImageNet

До моделі було додано нові, користувацькі щільні шари [60]:

1. Глобальне середнє об'єднання (GAP) – зменшує розмірність карт функцій базової моделі, зберігаючи просторову інформацію.

2. Два щільні шари: перший шар із 128 нейронами та функцією активації ReLU додає нелінійність у перетворення ознак, а другий шар із 64 нейронами та функцією ReLU додатково уточнює вилучені ознаки.

3. Вихідний шар – щільний шар із 7 нейронами та функцією SoftMax виконує класифікацію за сімома категоріями емоцій.

Щільні шари виконують роль дискримінаторів, уточнюючи загальні ознаки, отримані з MobileNetV2, у представлення, релевантні для класифікації емоцій. Вихідний SoftMax шар перетворює логіти в імовірності, дозволяючи моделі передбачати емоції з високою точністю [60].

Конфігурація навчання. Процес навчання був ретельно спланований для забезпечення оптимального балансу між точністю моделі та ефективністю обчислень (Додаток Г).

а) Функція втрати. Для вирішення задачі багатокласової класифікації використано функцію втрати «sparse_categorical_crossentropy». Вона обчислює різницю між прогнозованими розподілами ймовірностей і реальними класами, представленими як розріджені мітки (цілі числа). Використання розріджених міток мінімізувало споживання пам'яті та обчислювальні витрати, що сприяло оптимізації процесу навчання.

б) Оптимізатор. Модель навчалась із використанням оптимізатора Adam, який забезпечує динамічну адаптацію швидкості навчання для кожного параметра. Завдяки цьому досягається швидка й ефективна конвергенція. Крім того, Adam враховує коефіцієнти імпульсу та масштабування, що знижує ймовірність перевищення глобального мінімуму функції втрати під час оптимізації.

с) Збільшення даних. Для покращення моделі й підвищення її стійкості було впроваджено техніки збільшення даних. У процесі навчання застосовувалися різноманітні трансформації, зокрема:

- випадкове перевертання,
- обертання,
- масштабування,
- зміщення даних на початку кожної епохи.

Ці дії імітували різні сценарії, сприяючи кращому узагальненню моделі та підвищенню її ефективності.

d) Епохи та розмір партії. Навчання моделі проводилося протягом 25 епох, що є відносно невеликою кількістю завдяки використанню попередньо навченої базової моделі. Такий підхід дозволив скоротити кількість необхідних оновлень для тонкого налаштування. Розмір партії було обрано з урахуванням апаратних обмежень, що забезпечило ефективне використання пам'яті та стабільність оновлення градієнтів.

е) Нормалізація входів. Вхідні зображення були нормалізовані до діапазону від 0 до 1 шляхом ділення значень пікселів на 255, що вирівняло розподіл даних із припущеннями моделі MobileNetV2, підвищивши її стабільність і точність під час навчання. Нормалізація також зменшила обчислювальне навантаження, оскільки операції з меншими значеннями є менш ресурсозатратними.

Навчання та результати. Навчальний процес було організовано з урахуванням ключових етапів, спрямованих на ефективність та досягнення задовільної точності моделі [56, 62, 63] (Додаток Г):

1. Розподіл даних. Набір даних FER-2013 розділено на три частини: для навчання, перевірки та тестування, що забезпечило справедливую оцінку продуктивності моделі на кожному етапі.

2. Навчання моделі. Процес навчання проводився за допомогою фреймворку TensorFlow із використанням прискорення на GPU. У реальному часі здійснювався моніторинг ключових показників, таких як точність та втрати, що дозволяло вчасно виявляти потенційні проблеми, наприклад, переобладнання.

3. Перевірка ефективності. Після кожної епохи модель оцінювали на перевірконому наборі, щоб відстежити здатність до узагальнення. У разі стабілізації або погіршення продуктивності застосовували ранню зупинку для оптимізації навчання.

4. Кінцеві результати. Модель продемонструвала задовільну точність у розпізнаванні семи класів емоцій. Журнали навчання та графіки втрат і точності підтвердили стабільність і послідовність процесу навчання.

Переваги та виклики у використанні трансферного навчання [56, 62, 63, 64]:

1. Трансферне навчання. Використання попередньо навченої моделі MobileNetV2 значно скоротило час навчання.
2. Продуктивність на малих наборах даних. Використання FER-2013 дозволило уникнути потреби в масштабних наборах даних.
3. Економія обчислювальних ресурсів. Навчання за допомогою готових екстракторів ознак значно полегшило завдання.
4. Класовий дисбаланс. Обмежена представленість деяких емоцій вимагала застосування стратегій, таких як надмірна вибірка або вагові коефіцієнти.
5. Ризик переобладнання. Незважаючи на техніки збільшення даних (перемішування, масштабування, обертання), існував ризик перенавчання на обмеженому наборі даних, яке пом'якшувалося регулярною перевіркою та налаштуванням гіперпараметрів.

Під час тренування ноутбук із 32 ГБ оперативної пам'яті та сучасним процесором (AMD Ryzen 4800H) дозволяло проводити навчання, проте потребувало 4-5 годин для обробки 25 епох, тобто натренування однієї моделі (див. рис. 3.13.1, 3.13.2).

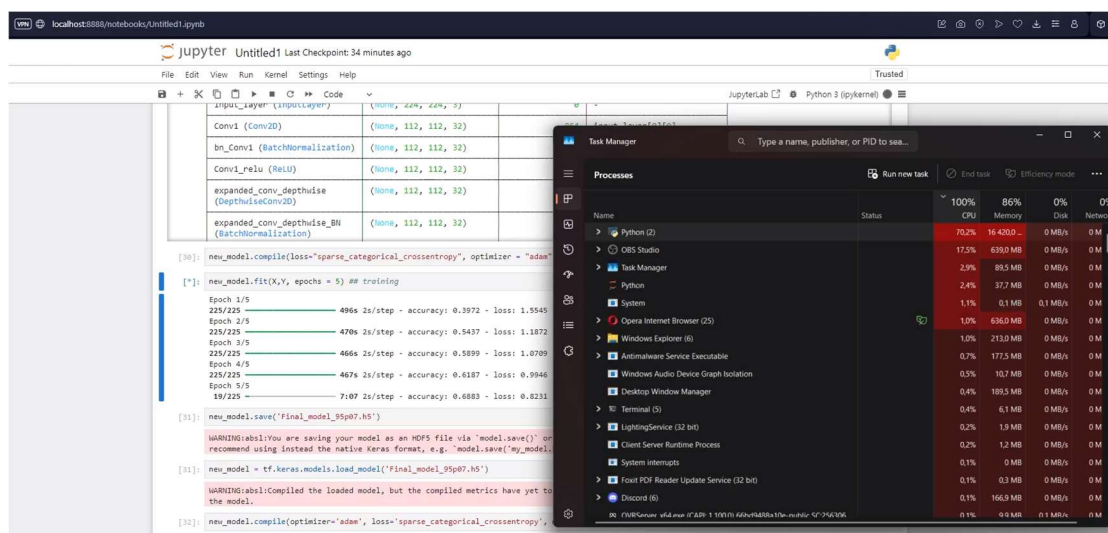


Рисунок 3.13.1 – Тренування моделі займає багато часу та ресурсів ПК

Високі навантаження на GPU, оперативну пам'ять і процесор (див рис. 3.13.2) створювали затримки та викликали помилки при роботі з повним набором FER-2013. Для оптимізації процесу використовували зменшений піднабір даних, що прискорило навчання.

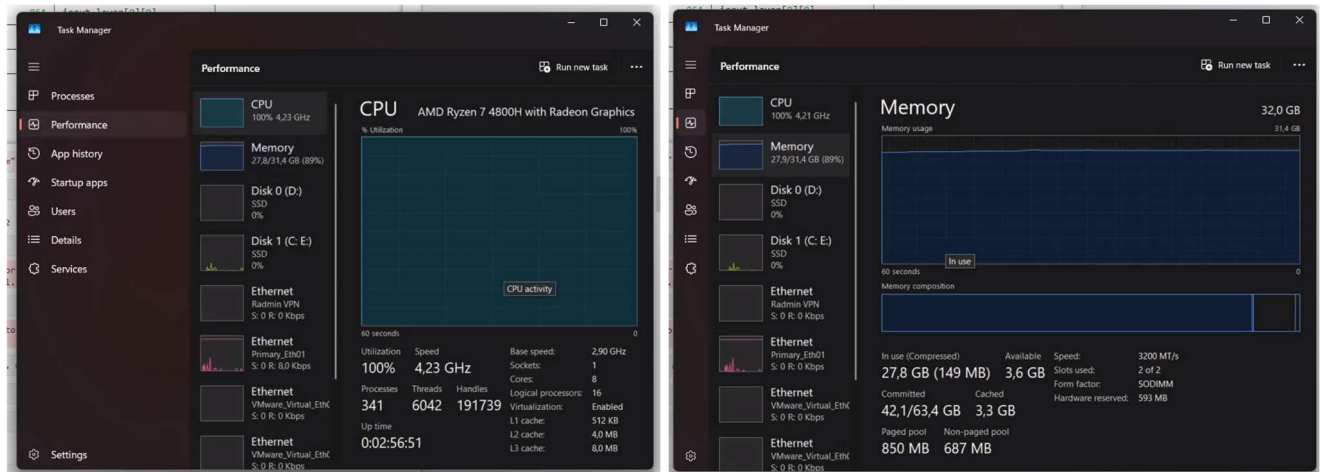


Рисунок 3.13.2 – Тренування моделі займає багато часу та ресурсів ПК

Попри тривалість і ресурсоємність процесу, модель успішно навчалася, підтверджуючи ефективність застосованого підходу.

3.4 Виявлення та візуалізація емоцій у реальному часі

Після успішного навчання та перевірки функціональності моделі, зокрема її тестування на зображеннях, ми можемо переходити до етапу впровадження моделі в практичне використання, де включає, наприклад, інтеграцію моделі з веб-камерою на ноутбуці чи комп'ютері для перевірки її роботи в реальному часі. Головним завданням є впевнитися, що модель функціонує коректно та відповідає очікуванням.

Перед початком практичного використання модель потрібно зберегти, що дозволяє уникнути повторного навчання (що потребує кількох годин), навіть якщо

проект буде закрито або пристрій вимкнено. Збережена модель може бути легко завантажена й використана у майбутньому. Бекапи є необхідністю, адже вони захищають від випадкового перезапису чи втрати результатів роботи.

На практичному етапі модель може бути налаштована для розпізнавання емоцій за допомогою камери, наприклад, визначення виразів обличчя в реальному часі. Також важливо забезпечити можливість використання моделі іншими людьми. Завдяки уніфікації, наявності необхідних бібліотек і плагінів, модель можна легко інтегрувати в інші проекти, що дозволяє ділитися моделлю та кодом з іншими користувачами, які зможуть без проблем її завантажити, налаштувати та протестувати. Таким чином, модель стає універсальним і зручним інструментом для широкого застосування [54-64] (Додатки Г, Г).

Обробка відео. Для розпізнавання емоцій у реальному часі необхідно перейти до обробки відео – перший крок до роботи з динамічними даними, на відміну від статичних зображень. Живе відео, отримане з веб-камери, дає змогу аналізувати емоції за допомогою спеціальних бібліотек та інструментів. Основним інструментом для реалізації цієї задачі є OpenCV, який дозволяє отримувати та обробляти відеопотік [57, 61].

Використовується функція «cv2.VideoCapture», яка забезпечує доступ до камери. Наприклад, команда «cv2.VideoCapture(0)» відкриває основну камеру за замовчуванням. Якщо основна камера недоступна, можна використовувати інші індекси для підключення альтернативних камер, таких як зовнішні пристрої чи смартфони, налаштовані через додатки на зразок DroidCam.

Процес починається з перевірки доступу до камери через «cap.isOpened()», що дозволяє обробляти можливі помилки ініціалізації. У режимі реального часу відеопотік надходить у вигляді послідовності кадрів, представлених масивами NumPy, які обробляються в циклі. Для зчитування кожного кадру використовується метод «cap.read()», що забезпечує мінімальні затримки та оптимальне навантаження на систему (Додатки Г, Г).

Для стабільної роботи враховуються такі фактори:

1. Частота кадрів. У різних країнах можуть використовуватись різні стандарти, наприклад, 25, 30, 50 або 60 кадрів на секунду. Щоб уникнути проблем із продуктивністю, частота кадрів налаштовується відповідно до можливостей системи.
2. Оптимізація кадрів. Використання операцій, таких як зменшення роздільної здатності або перетворення кадрів у градації сірого, допомагає спростити аналіз та зменшити затримки.
3. Налаштування роздільної здатності. Камера повинна мати достатню якість для забезпечення точності аналізу. Для смартфонів, які часто використовуються як камера, цього цілком достатньо. У разі використання бездротового з'єднання можуть виникати затримки, які усуваються за допомогою проводового підключення.

Розпізнавання обличчя. Після інтеграції функцій OpenCV наступним етапом є розпізнавання обличчя та підготовка його до аналізу емоцій. Камера часто фіксує загальну сцену, яка може включати обличчя під різними кутами або в несприятливих умовах освітлення. Тому перед обробкою необхідно визначити зону обличчя та налаштувати зображення [57].



Рисунок 3.14– Розпізнавання обличчя

Для цього використовується каскадний класифікатор Хаара. Метод працює на основі попередньо навченої моделі, збереженої у файлі «haarcascade_frontalface_default.xml». Ця модель дозволяє ідентифікувати обличчя за допомогою аналізу візерунків, таких як краї та текстури, характерні для людського обличчя [61].

Процес включає [57, 61] (Додатки Г, Г):

1. Перетворення зображення. Кожен кадр відео конвертується у градації сірого за допомогою «cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)», оскільки класифікатор працює лише з одноканальними зображеннями.

2. Виявлення облич. Метод «detectMultiScale()» дозволяє ідентифікувати обличчя у кадрі. Параметри, такі як «scaleFactor» і «minNeighbors», налаштовуються для оптимального виявлення без хибних спрацювань. Результатом є координати прямокутників, які позначають обличчя на кадрі.

Виявлені обличчя передаються до модуля розпізнавання емоцій для подальшого аналізу. Проте класифікатор має деякі обмеження [61]:

1. Чутливість до освітлення. У темних умовах можуть виникати шуми, що погіршує точність розпізнавання, що вирішується нормалізацією зображення, наприклад, за допомогою вирівнювання гістограми або гамма-корекції.

2. Обмеження виявлення під кутом. Якщо обличчя частково приховане або повернуте, точність розпізнавання знижується. Для покращення моделі до тренувального набору додаються зображення облич із різних кутів, що дозволяє враховувати подібні випадки.

Після виявлення обличчя здійснюється попередня обробка, яка готує його до аналізу нейронною мережею. Основні етапи [57, 60, 61] (Додатки Г, Г):

1. Зміна розміру. Виявлене обличчя масштабується до розміру 224×224 пікселів за допомогою «cv2.resize()», щоб відповідати вимогам моделі MobileNetV2.

2. Нормалізація пікселів. Значення пікселів масштабуються в діапазон [0, 1] шляхом ділення на 255, що дозволяє моделі ефективніше аналізувати дані та зменшує обчислювальну інтенсивність.

3. Форматування вхідних даних. Дані розширюються для включення розміру партії. Метод `np.expand_dims()` додає вимір, формуючи вхід у вигляді (1, 224, 224, 3), що забезпечує сумісність із вимогами нейронної мережі.

Попередня обробка дозволяє моделі зосередитися на найбільш релевантних деталях обличчя, що підвищує точність аналізу емоцій і забезпечує стабільність обробки навіть у складних умовах.

Для прогнозування емоцій використовується попередньо навчена модель, завантажена через `keras.models.load_model()` із застосуванням TensorFlow. Ця модель обробляє зображення облич, отримані з веб-камери, та розподіляє ймовірності між сімома категоріями емоцій: злість, огида, страх, щастя, сум, здивування, нейтральність. Результати обчислюються швидко завдяки оптимізації GPU та підготовці даних [1, 56, 57, 60, 61].

Для визначення емоції з найвищою ймовірністю використовується `np.argmax()`, яка знаходить індекс максимального значення у масиві ймовірностей, що дозволяє точно визначити домінуючу емоцію, а швидкість обробки кожного кадру забезпечує плавність роботи в реальному часі (див. рис. 3.15).

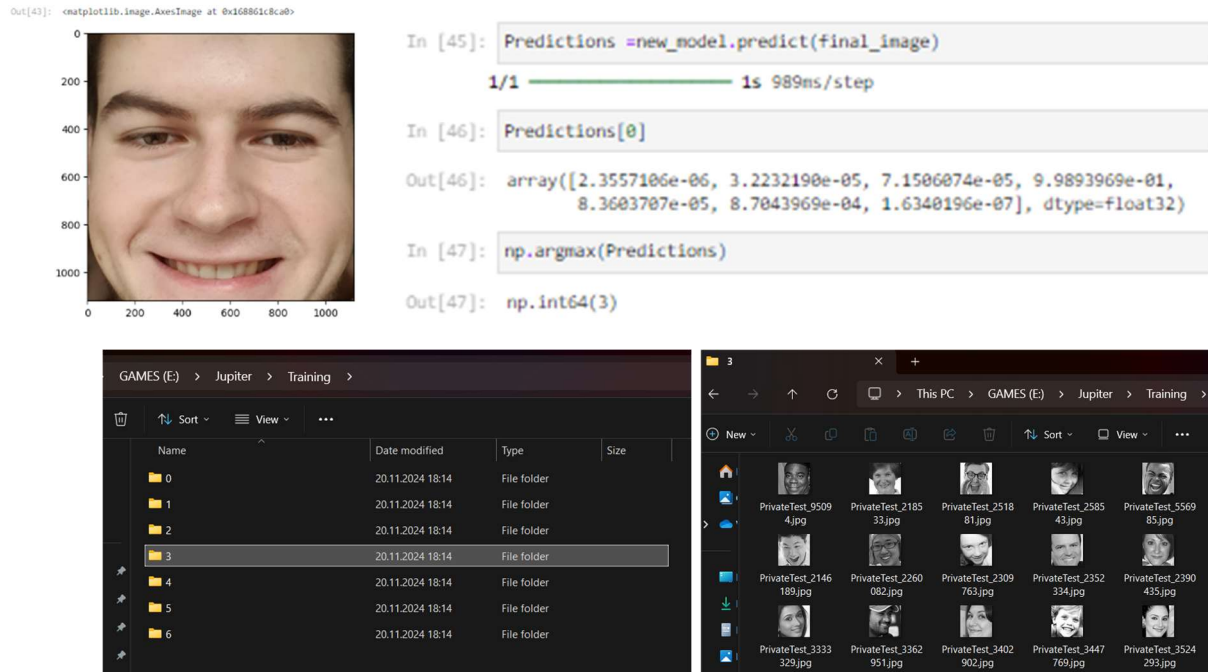


Рисунок 3.15– Коректність відпрацювання тестової моделі

Результати прогнозування візуалізуються в режимі реального часу. Ключові етапи (Додатки Г, І):

1. Накладення рамок на обличчя. Для виділення обличчя у відеокадрі використовується «cv2.rectangle()». Прямокутник відображає область, яку аналізує модель.
2. Відображення емоцій. Прогнозована емоція для кожного обличчя відображається у вигляді тексту, нанесеного поблизу рамки за допомогою «cv2.putText()».

Відеоряд постійно оновлюється, обробляючи до 50 чи 60 кадрів за секунду, що забезпечує динамічне й безперебійне відображення результатів.

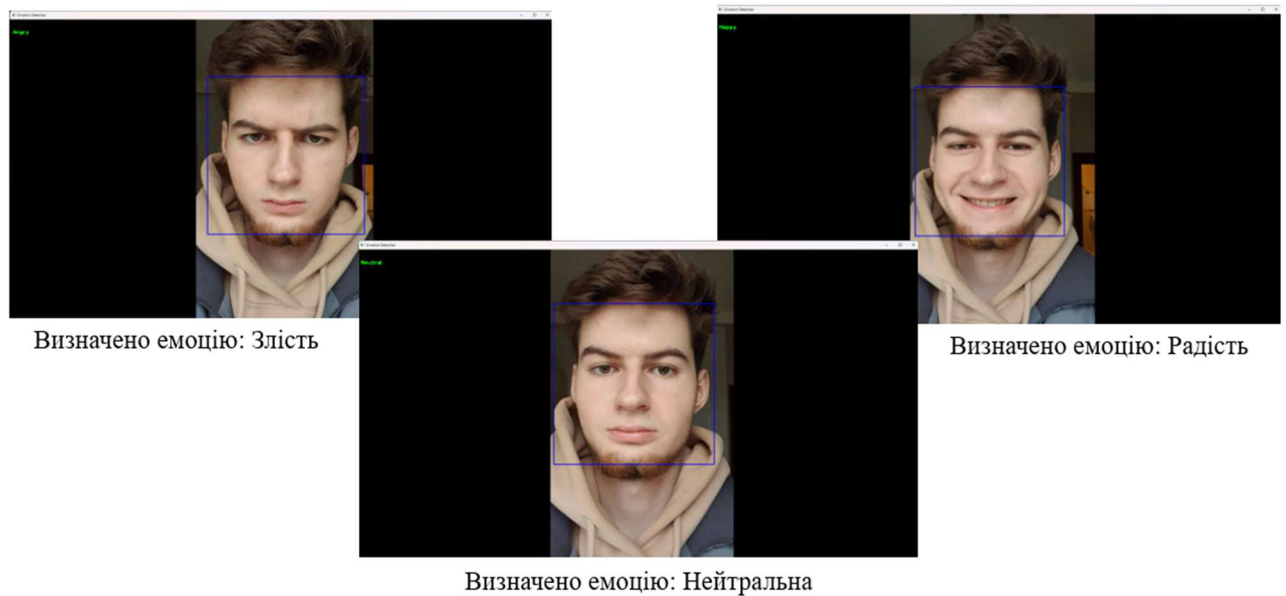


Рисунок 3.16 – Демонстрація моделі розпізнавання емоційних станів людини

Оброблені кадри з накладеними результатами емоційного аналізу виводяться на екран за допомогою «cv2.imshow()», що дозволяє користувачам спостерігати за роботою системи в реальному часі (Додаток Г).

Для зручності передбачений простий механізм завершення роботи програми, наприклад, натискання клавіші «Q» або іншої заданої кнопки. Такий підхід забезпечує користувачам інтуїтивну взаємодію з модулем розпізнавання емоцій.

Для забезпечення максимальної надійності системи розпізнавання емоцій у реальному часі, навіть за умови її стабільної роботи, варто врахувати кілька додаткових аспектів.

Оптимізація продуктивності. Щоб мінімізувати затримки та підвищити швидкість обробки, важливо оптимізувати кількість операцій, а також використовувати можливості прискорення GPU, яке допоможе системі забезпечувати плавну роботу без помітних затримок.

Зменшення хибних спрацювань. Для зниження ймовірності помилкового виявлення облич слід ретельно налаштувати параметри каскаду Хаара. Впровадження додаткових фільтрів для очищення результатів також може значно підвищити точність і стабільність системи.

Контроль порогових значень. Використання порогових значень для визначення емоційних станів дозволяє зробити прогноз більш впевненим і точним, що допомагає не лише уникати невизначеності, але й спрощує розуміння результатів як у текстовому вигляді, так і через візуалізацію в реальному часі.

Усе це сприяє підвищенню надійності системи як для основних користувачів, так і для третіх осіб, які можуть використовувати її у своїх проєктах. Таким чином, система забезпечує стабільність, точність і швидкість роботи, необхідну для практичного застосування.

3.5 Перспективи розвитку, етичні аспекти, аспекти безпеки та приклади застосування

Системи розпізнавання емоцій представляють собою перспективний, хоча й досить чутливий інструмент, який залежить від правильного налаштування та навчання за допомогою нейронних мереж. Цей підхід не лише охоплює технічні аспекти навчання моделей, але й вимагає врахування соціальних, психологічних і етичних факторів.

Для майбутнього розвитку таких технологій важливо створювати моделі, які зможуть опрацьовувати різноманітні демографічні дані та враховувати етичні аспекти, що необхідно для їх відповідального використання. Зокрема, слід забезпечити відповідність систем законам, гарантувати безпеку й конфіденційність даних, а також уникати дискримінаційних ефектів. Без цих заходів розроблені моделі залишатимуться лише прототипами, непридатними для широкого використання.

Система розпізнавання емоцій має потенціал застосування в різних сферах, таких як освіта, охорона здоров'я, психологія, терапія, правоохоронні органи та маркетинг. Її впровадження може принести значну користь як для кінцевих споживачів, так і для фахівців, які працюють із такими даними. Використання цих

технологій дозволить ефективніше взаємодіяти з людьми, розуміти їхній емоційний стан і відповідно адаптувати послуги, що сприятиме покращенню якості життя й загальній ефективності різних галузей.

Технології розпізнавання емоцій продовжують активно розвиватися, незважаючи на багаторічні дослідження та існування численних моделей, адаптованих до різних демографічних груп. Майбутнє цієї галузі полягає у вдосконаленні технологій через усунення наявних обмежень, розширення можливостей моделей і вдосконалення їхньої універсальності.

Підвищення точності та узагальнення. Для ефективного використання цих моделей необхідно постійно підвищувати їхню точність та здатність узагальнювати дані, що досягається за рахунок безперервного навчання на великих і різноманітних наборах даних. Наприклад, інтеграція наборів, які враховують етнічну приналежність, вік, стать і культурні відмінності, дозволяє зменшити упередження моделей і зробити їх більш універсальними.

Крім універсальних рішень, необхідно створювати локальні моделі, адаптовані до специфіки певних регіонів чи груп населення. Наприклад, специфічні дані, отримані від племен чи культурних спільнот, можуть враховувати унікальні емоційні вирази, які не завжди враховуються універсальними моделями.

Для підтримання актуальності систем потрібно створювати динамічні набори даних, що дозволить інтегрувати сучасні контекстуальні елементи, як-от нові культурні тренди чи використання масок, що закривають обличчя. Постійне оновлення забезпечить відповідність моделей сучасним умовам.

Окрім універсальних завдань, моделі мають бути спеціально налаштовані для окремих сфер, таких як охорона здоров'я чи освіта, що допоможе їм працювати ефективніше у вузьких контекстах, уникаючи перетренованості чи зростання похибок.

Правильний вибір даних і методів навчання є критично важливим. Надмірне чи неправильне тренування може призвести до зниження точності моделі навіть у її основній сфері використання. Тому потрібно дотримуватися балансу між

універсальністю та спеціалізацією, щоб забезпечити низький коефіцієнт похибок і високу ефективність у майбутньому.

Розвиток технології розпізнавання емоцій базується на комплексному підході до створення, навчання та впровадження моделей, які будуть не лише точними, але й адаптованими до різноманітних умов використання.

Мультимодальне розпізнавання емоцій. Мультимодальне розпізнавання емоцій передбачає використання не лише зображень, а й додаткових джерел даних, таких як мовні сигнали. Поєднання мимічних виразів із голосовими інтонаціями, тембром чи іншими акустичними особливостями дозволяє отримати більш точну картину емоційного стану, що особливо корисно в ситуаціях, коли зображення недостатньо інформативне через неоднозначність виразів обличчя або погані умови освітлення.

Додаткову точність можна досягти шляхом інтеграції фізіологічних даних, таких як частота серцевих скорочень, провідність шкіри чи інші біометричні показники. Ці дані, отримані за допомогою датчиків чи носимих пристроїв (розумних годинників, фітнес-браслетів тощо), можуть виявляти емоції на основі фізіологічних реакцій. Наприклад, такі реакції, як прискорене серцебиття чи потовиділення, можуть вказувати на стрес, страх або хвилювання. Важливо, щоб використання таких даних здійснювалося за згодою користувача відповідно до умов конфіденційності.

Аналіз групових емоцій. Для зменшення затримок та підвищення ефективності модель розпізнавання емоцій можна адаптувати для роботи на кінцевих пристроях, таких як смартфони чи інші периферійні пристрої. Локальна обробка даних дозволить уникнути передачі інформації на віддалені сервери, що підвищить конфіденційність та зменшить залежність від зовнішніх ресурсів.

Для забезпечення масштабованості та ефективності важливо оптимізувати модель шляхом використання методів, таких як квантування чи скорочення архітектури, що дозволить зменшити обчислювальні ресурси, необхідні для роботи моделі, і зробити її доступнішою для широкого застосування на мобільних та інших пристроях з обмеженими ресурсами.

Мультиmodalьне розпізнавання емоцій і його адаптація до різних умов відкривають перспективи для створення гнучких, точних і масштабованих рішень у цій галузі.

Етичні міркування щодо впровадження моделей розпізнавання емоцій. Моделі розпізнавання емоційних станів часто стикаються з проблемою упередженості, зокрема етнічного характеру. Наприклад, використання набору даних на зразок FER-2013 може не враховувати особливості емоційних виразів населення віддалених регіонів, що призводить до неправильного трактування емоцій і зниження точності моделі в таких умовах, де можна вирішити шляхом застосування більш локалізованих та специфічних наборів даних, адаптованих для окремих груп населення.

Однак забезпечення ефективності моделі має супроводжуватися дотриманням принципів справедливості, прозорості та суспільного схвалення. Важливо не лише впровадити модель, а й підготувати користувачів, споживачів, пацієнтів або інші зацікавлені сторони до її використання, що стосується як комерційних сфер, таких як маркетинг, так і суспільно важливих галузей. Прозорість роботи таких систем є ключовим фактором для здобуття довіри суспільства, незалежно від мети та масштабу їх впровадження.

Прозорість і мета комунікації. Для підвищення довіри користувачів необхідно чітко пояснювати мету, масштаби та обмеження системи. Організації мають інформувати, які дані збираються, як вони використовуються та які результати можна очікувати. Наприклад, зазначати, чи система спрямована на аналіз настроїв клієнтів, підтримку психічного здоров'я або забезпечення безпеки.

Також слід уникати неправомірного використання, як-от стеження або маніпулятивної реклами без згоди користувачів, що знижує довіру та може негативно вплинути на репутацію таких моделей. Прозорість і зрозумілість функціонування системи є ключовими для її прийняття суспільством.

Захист конфіденційності. Необхідно мінімізувати обробку даних, збираючи лише ті, що потрібні для виконання конкретних завдань, наприклад, для навчання або вдосконалення моделі. Дані мають бути анонімними, особливо коли йдеться

про обличчя користувачів. Використання локальної обробки інформації на пристроях замість передачі на сервери допоможе підвищити безпеку.

Важливим є впровадження автоматизованого видалення або приховування ідентифікаційних ознак, щоб дані залишалися конфіденційними. Шифрування інформації під час передачі і зберігання захистить її від стороннього доступу.

Пом'якшення упередженості. Різноманітність у наборах даних для навчання моделі є критично важливою. Відсутність представництва певних груп населення може призводити до упереджених прогнозів. Забезпечення рівності результатів для різних демографічних груп, а також регулярна оцінка точності моделі дозволить уникнути несправедливих рішень.

Періодичний аналіз та оновлення моделей допоможе усунути потенційні упередження, які можуть виникати з часом.

Психологічний вплив і згода. Збір даних і використання системи мають бути узгоджені з користувачами, особливо у сферах охорони здоров'я та безпеки. Психологічний вплив системи може викликати недовіру, особливо якщо результати трактуються некоректно.

Необхідно забезпечити високу точність прогнозів і повідомляти користувачів про рівень невизначеності. Отримання згоди на використання моделі має бути пріоритетним завданням, а комунікація — максимально прозорою, що дозволить уникнути маніпуляцій і забезпечить довіру до технології.

Окрім впровадження функціональних можливостей, у майбутніх моделях особливу увагу слід приділяти питанням безпеки. Безпека займає ключову роль у будь-якій системі штучного інтелекту, особливо коли йдеться про обробку конфіденційних і персональних даних.

Необхідно гарантувати надійність заходів, які забезпечуватимуть захист даних, системи та її користувачів, що включає запобігання витокам інформації, забезпечення стабільної роботи системи й автоматизацію процесів безпеки. Модель повинна бути оснащена механізмами, що дозволять уникати фатальних наслідків, таких як доступ до даних сторонніх осіб, які не мають на це права.

Запобігання несанкціонованому доступу. Для забезпечення захисту від несанкціонованого доступу необхідно впроваджувати сучасні механізми автентифікації. Система повинна підтримувати багатофакторну автентифікацію, яка може включати біометричні дані, такі як сканування відбитків пальців або сітківки ока, а також фізичні ключі або спеціальні картки з передовими технологіями ідентифікації. Лише уповноважений персонал, відповідальний за розробку та підтримку системи, повинен мати доступ до конфіденційних даних.

Крім цього, необхідно застосовувати мережеві захисні інструменти, такі як міжмереві екрани (Firewall), а також системи виявлення вторгнень, які дозволяють моніторити мережеву активність і своєчасно реагувати на кібератаки, що забезпечить високий рівень безпеки як для самої системи, так і для її сховищ даних.

Шифрування даних. Для захисту інформації під час передачі та зберігання важливо використовувати сучасні стандарти шифрування, такі як AES-256, що дозволить знизити ризики несанкціонованого доступу до даних, включаючи дані користувачів, прогнозів емоцій та інші конфіденційні елементи системи.

Також необхідно забезпечити безпечне управління ключами шифрування, зокрема впровадження апаратних модулів безпеки (HSM) або інших захищених сховищ для ключів, що допоможе запобігти несанкціонованому розшифруванню та підвищить рівень загального захисту системи.

Регулярні аудити та оновлення. Регулярна оцінка вразливостей та оновлення компонентів системи є ключовими для її безпеки. Періодичні аудити дозволяють виявляти слабкі місця інфраструктури, що використовує модель, та своєчасно їх усувати.

Оновлення моделей і їх залежностей, таких як фреймворки машинного навчання (TensorFlow, OpenCV тощо), забезпечує актуальність та захищеність системи. Перед впровадженням нових компонентів їх необхідно тестувати в ізолюваному середовищі, щоб уникнути можливих ризиків.

Планування реагування на інциденти. Необхідно розробити план реагування на інциденти, який включатиме детальний алгоритм дій у разі порушень безпеки.

План повинен передбачати локалізацію інциденту, інформування постраждалих користувачів, відновлення даних і заходи для недопущення подібних ситуацій у майбутньому, що забезпечить ефективне управління кризовими ситуаціями та захист репутації системи.

Приклади застосування навченого модуля розпізнавання емоційних станів.

Для застосування модуля розпізнавання емоційних станів можна використати досить малопотужний ПК, наприкладі мініатюрного ПК Raspberry Pi (див. рис. 3.17-3.19), оснащеного навченою моделлю MobileNetV2 для розпізнавання емоцій, демонструє широкий спектр застосувань у чотирьох важливих галузях: промисловості, транспорті, медицині та громадських місцях.

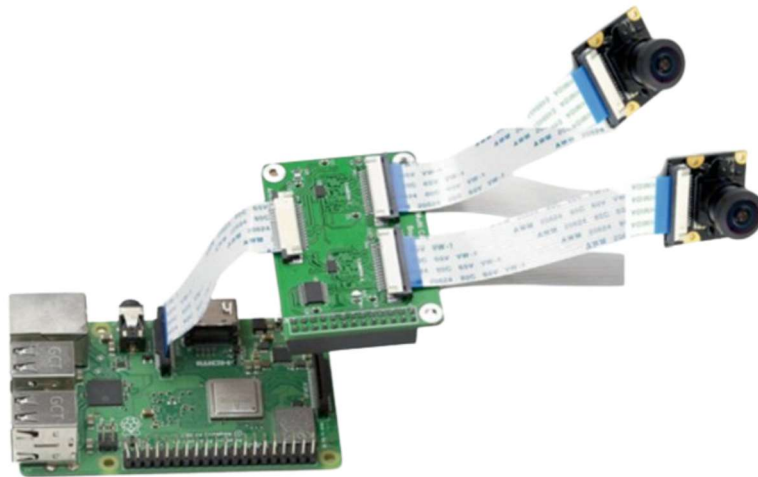


Рисунок 3.17 – Raspberry Pi зі адаптером на дві камери для кращої працездатності та інтеграції модуля розпізнавання емоцій

У кожному випадку система показує здатність вирішувати специфічні задачі, використовуючи відеопотік із камер (див. рис.3.16) у реальному часі для виявлення емоційних станів людей, що можуть впливати на безпеку, ефективність і загальний контроль.

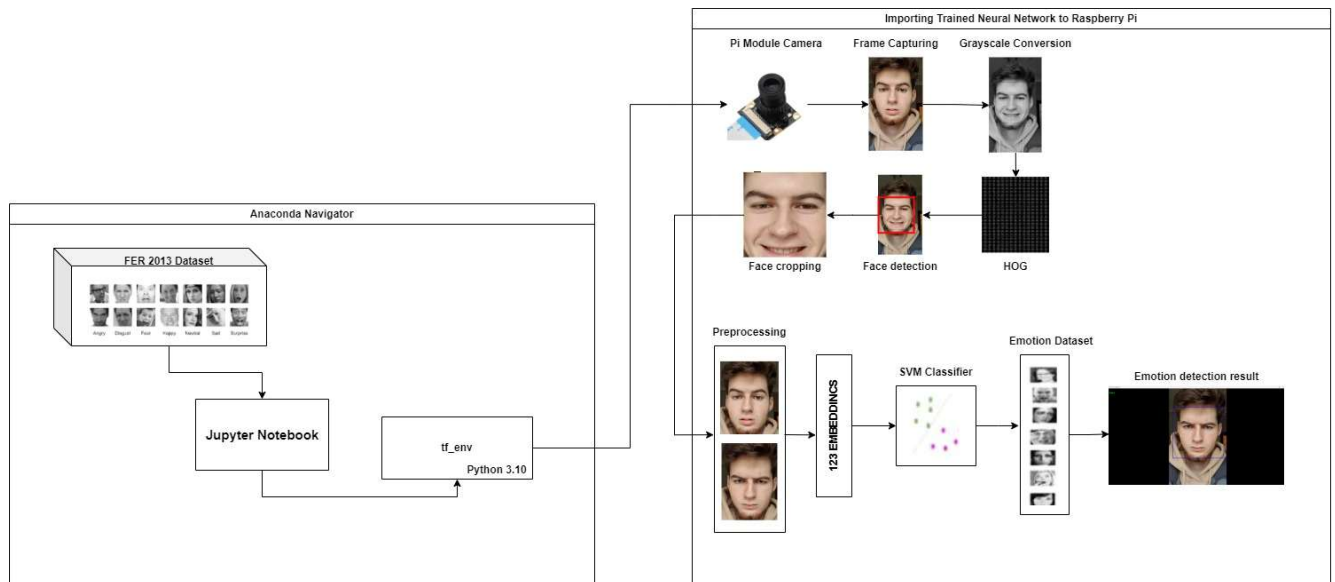


Рисунок 3.18 – Процес інтеграції натренованої моделі у Raspberry Pi

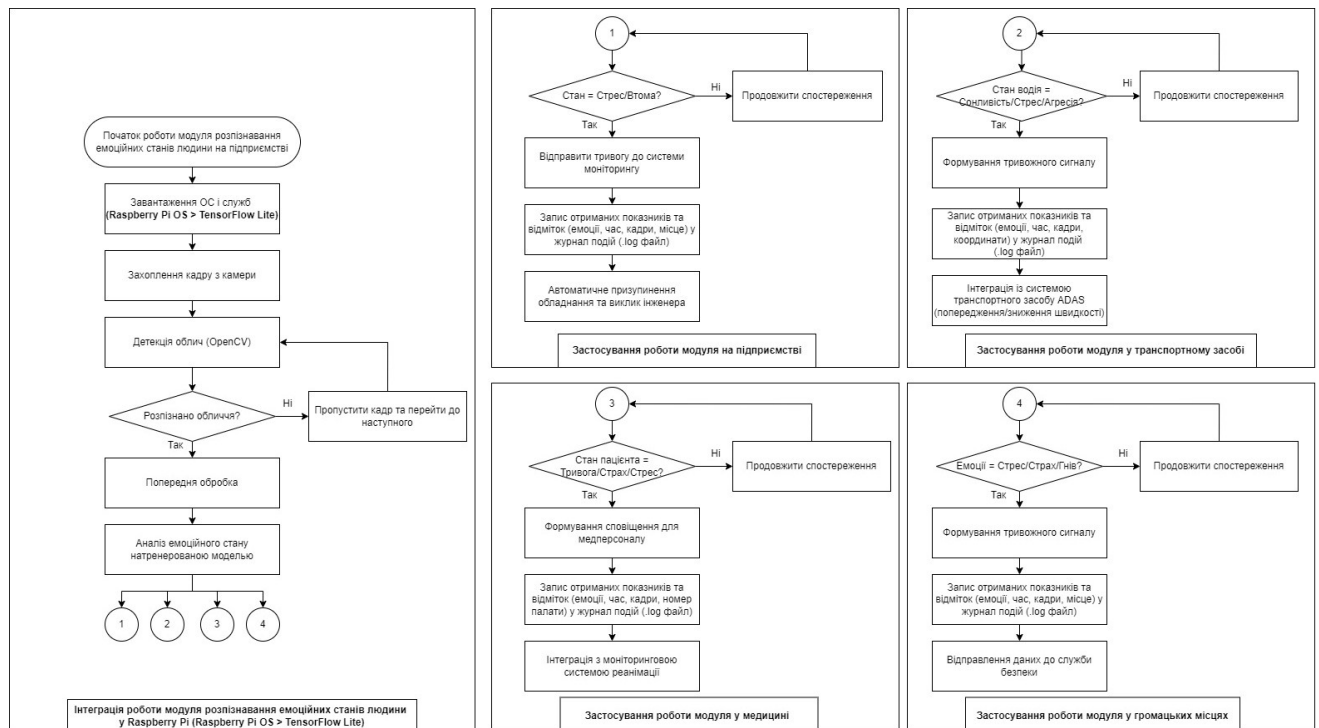


Рисунок 3.19 – Різностороннє практичне застосування натренованої моделі розпізнавання емоційних станів людини

У промисловості (див. рис.3.19) Raspberry Pi інтегрується в системи контролю операторів складного обладнання. Система постійно аналізує вирази обличчя працівників, виявляючи ознаки стресу, втоми або інших станів, які можуть знизити їхню продуктивність або призвести до помилок. Завдяки цьому рішення

дозволяє знизити ризик аварійних ситуацій на виробництві, забезпечуючи своєчасну реакцію на потенційно небезпечний стан оператора. Наприклад, якщо система фіксує ознаки втоми, вона може надсилати повідомлення до диспетчерської або активувати автоматичне тимчасове зупинення роботи обладнання, що дозволить оператору перепочити. Завдяки використанню MobileNetV2 модель здатна працювати швидко та точно навіть на простих обчислювальних пристроях, таких як Raspberry Pi, що робить її доступною для впровадження на середніх і малих підприємствах.

У транспортній галузі (див. рис.3.19) рішення фокусується на аналізі стану водіїв. Камери, встановлені в кабіні транспортного засобу, передають відеопотік на мініатюрний ПК, де модель розпізнає ознаки сонливості, стресу або агресії, що особливо актуально для водіїв вантажівок, автобусів або інших транспортних засобів, які працюють у режимах підвищеного навантаження. Якщо система виявляє ознаки сонливості, вона може активувати звуковий або вібраційний сигнал, що попереджає водія, а також інформувати диспетчерський центр про необхідність зупинки. Завдяки інтеграції з іншими системами безпеки транспортного засобу, такими як автоматичне екстрене гальмування або системи утримання смуги руху, Raspberry Pi дозволяє знизити ймовірність дорожньо-транспортних пригод. Гнучкість цієї моделі дозволяє адаптувати її до різних типів транспортних засобів і умов експлуатації.

У медицині (див. рис.3.19) система є надзвичайно корисною для оцінки емоційного стану пацієнтів у реанімаційних відділеннях. Камери, встановлені біля ліжка пацієнта, передають відео на Raspberry Pi, де модель аналізує вирази обличчя, визначаючи рівень стресу, страху або спокою, що особливо важливо у випадках, коли пацієнти перебувають у несвідомому стані або мають обмежену можливість спілкуватися з медичним персоналом. Виявлення ознак стресу чи дискомфорту дозволяє медичним працівникам оперативно реагувати, коригуючи лікування, знеболення або інші аспекти догляду. Система може працювати автономно, що є критично важливим у випадках, коли реанімаційне відділення обмежене у своїх ресурсах, або в умовах польових госпіталів.

У громадських місцях (див. рис.3.19), таких як аеропорти чи вокзали, система виконує функцію моніторингу поведінки людей у режимі реального часу. Відеопотік із камер спостереження передається на Raspberry Pi, який аналізує емоційний стан кожної особи у кадрі. У критичних зонах, таких як пункти контролю або місця великого скупчення людей, система виявляє ознаки стресу, страху чи агресії, що можуть вказувати на потенційно небезпечну поведінку. У разі виявлення таких емоцій система може надсилати сигнал до служби безпеки для подальшого спостереження за конкретною особою або активації додаткових заходів контролю, що дозволить оперативно запобігати інцидентам, таким як конфлікти, паніка чи можливі терористичні акти. Завдяки низькій вартості та легкості масштабування, рішення може бути впроваджено у будь-яких громадських місцях із мінімальними витратами.

Запропоноване рішення із використанням навченої моделі на основі MobileNetV2, тренуваної на наборі даних FER-2013, має суттєві переваги над існуючими рішеннями завдяки кільком ключовим факторам, які забезпечують її ефективність у зазначених чотирьох напрямках (промисловість, транспорт, медицина, громадські місця). У кожній із цих сфер модель демонструє гнучкість, точність, економічність і здатність до інтеграції, що дозволяє вирішувати завдання з більшою ефективністю, ніж поточні методи або рішення.

Навчання на наборі даних FER-2013 дає моделі перевагу в її здатності розпізнавати широкий спектр емоційних станів, які відповідають реальним сценаріям у критичних умовах. Це включає такі емоції, як стрес, страх, агресія, радість, сум і нейтральний стан. Набір даних FER-2013 є добре збалансованим і широко використовуваним у наукових дослідженнях, що дозволяє моделі адаптуватися до різноманітних ситуацій. Це критично важливо для задач, де емоційний стан людини може стати вирішальним фактором для запобігання аваріям, зниження ризиків або покращення обслуговування.

Використання MobileNetV2 є стратегічним вибором через її архітектурну легкість та високу продуктивність на мобільних і обмежених апаратних пристроях, таких як одноплатні комп'ютери Raspberry Pi. Завдяки своїй оптимізації для

невеликих ресурсів, ця модель забезпечує достатню швидкість обробки для аналізу відеопотоку в реальному часі, навіть у системах із обмеженими обчислювальними потужностями. У порівнянні з важчими моделями, такими як ResNet або VGG, MobileNetV2 дозволяє досягати високої точності без значного збільшення вимог до обчислювальних ресурсів. Це особливо важливо для розгортання в умовах, де необхідна автономність і швидка обробка.

Щодо промисловості, запропоноване рішення має значну перевагу завдяки своїй здатності виявляти стрес або втому у операторів складного обладнання. Існуючі системи моніторингу часто покладаються на непрямі показники, такі як частота серцевих скорочень або рівень активності, які можуть бути ненадійними через індивідуальні особливості. Натомість запропонована модель аналізує безпосередньо емоційний стан, використовуючи відеопотік, що дозволяє виявляти зміни в поведінці оператора в режимі реального часу. Це допомагає запобігти аваріям, викликаним втомою або емоційним перенавантаженням, і дозволяє підвищити ефективність роботи завдяки своєчасному втручанням.

У транспорті рішення перевершує існуючі системи, такі як датчики контролю руху очей чи трекери голови, які іноді дають хибні спрацювання. Завдяки використанню MobileNetV2 і набору FER-2013, система може точно виявляти емоційні стани, такі як сонливість, стрес або агресія, які часто передують аварійним ситуаціям. Ця модель здатна працювати в реальному часі і може інтегруватися з автомобільними системами ADAS для автоматичного запобігання аварій, наприклад, шляхом зниження швидкості або активації автопілота. У порівнянні з існуючими методами, які в основному орієнтовані на фізичну поведінку, запропоноване рішення враховує емоційну складову, що дозволяє більш комплексно оцінювати стан водія.

У медицині запропонована модель надає значний прогрес у порівнянні з традиційними методами оцінки стану пацієнтів. Сучасні системи в основному базуються на моніторингу фізіологічних параметрів, таких як частота серцевих скорочень, артеріальний тиск або рівень кисню в крові. Хоча ці показники важливі, вони не завжди відображають емоційний стан пацієнта, який може бути критичним

у реанімаційних умовах. Використання запропонованої моделі дозволяє медичному персоналу отримувати додаткову інформацію про емоційний стан пацієнта, що може сприяти кращій комунікації, плануванню лікування і швидкому реагуванню на стресові ситуації. Крім того, система може працювати автономно, що особливо корисно у випадках, коли медичний персонал обмежений.

Для громадських місць, таких як аеропорти або вокзали, запропоноване рішення пропонує унікальні можливості для забезпечення безпеки. Сучасні системи безпеки часто покладаються на поведінковий аналіз або технології розпізнавання облич, які орієнтовані на ідентифікацію конкретних осіб. Запропонована модель додає новий рівень функціональності, дозволяючи виявляти стрес, страх чи агресію, які можуть бути ознаками підозрілої поведінки або потенційної загрози. Ця система може інтегруватися з існуючими камерами спостереження і в режимі реального часу надсилати сповіщення службі безпеки, що дозволяє оперативно реагувати на інциденти і запобігати можливим загрозам.

Таблиця 3.1 – Порівняльний аналіз запропонованого рішення з існуючими системами розпізнавання емоцій

Критерій	Запропоноване рішення	Існуючі рішення
Апаратні вимоги	Оптимізоване для малопотужних пристроїв (Raspberry Pi).	Потребують потужних серверів або спеціалізованого обладнання.
Ефективність у реальному часі	Аналіз відеопотоку з мінімальними затримками.	Часто не здатні працювати в реальному часі.
Точність розпізнавання емоцій	Використовує FER-2013, добре збалансований набір даних.	Можуть покладатися на менш точні або застарілі моделі.
Гнучкість у застосуванні	Підходить для різних сфер: промисловість, транспорт, медицина, громадські місця.	Часто обмежені однією сферою або вузьким сценарієм.
Витрати на впровадження	Низька вартість обладнання та обслуговування.	Високі витрати на спеціалізовані системи.
Тип аналізу	Комбінує емоційний стан із контекстом ситуації.	Зазвичай орієнтовані на фізичні показники (рухи очей, пульс).
Автономність	Може працювати без постійного підключення до мережі.	Більшість рішень залежить від хмарних обчислень.

Продовження таблиці 3.1 – Порівняльний аналіз запропонованого рішення з існуючими системами розпізнавання емоцій

Критерій	Запропоноване рішення	Існуючі рішення
Масштабованість	Легка інтеграція з наявними системами та масштабування.	Часто складно інтегрувати або адаптувати до нових потреб.
Інноваційність	Використовує сучасні глибокі нейронні мережі (MobileNetV2).	Спираються на менш ефективні алгоритми або методи.
Розпізнавання складних емоцій	Здатність розпізнавати стрес, агресію, страх тощо.	Орієнтовані лише на базові або загальні емоції.

ВИСНОВКИ

Системи розпізнавання емоцій, створені завдяки прогресу машинного навчання та комп'ютерного зору, мають величезний потенціал для трансформації різноманітних галузей, зокрема охорони здоров'я, транспорту, промисловості та безпеки в громадських місцях. У цій роботі було успішно реалізовано модель розпізнавання емоцій на базі MobileNetV2, яка використовує дані FER-2013 і здатна працювати на обчислювально ефективних пристроях, таких як Raspberry Pi. Використання методів передавання навчання, ретельна попередня обробка даних і реалізація аналізу відеопотоку в реальному часі продемонстрували можливість впровадження таких систем у реальних умовах.

Рішення відзначається масштабованістю, низькими апаратними вимогами та адаптивністю до різних сценаріїв використання. Наприклад, у промисловості система моніторить емоційний стан операторів складного обладнання, запобігаючи аваріям; у транспорті виявляє ознаки стресу, сонливості чи агресії у водіїв, інтегруючись із системами запобігання ДТП; у медицині забезпечує моніторинг стану пацієнтів у реанімаційних відділеннях; у громадських місцях ідентифікує потенційно небезпечну поведінку для забезпечення безпеки. Така гнучкість досягається завдяки оптимізації продуктивності моделі через методи зміни розміру, нормалізації та збільшення даних, що дозволяє моделі добре узагальнювати свої результати для різноманітних умов і середовищ.

Інтеграція можливостей обробки відео в реальному часі є важливою перевагою, яка забезпечує миттєвий зворотний зв'язок. Це дозволяє системі вчасно реагувати на загрози, підтримувати безперервність процесів і підвищувати ефективність у критичних середовищах. Крім того, рішення є економічно доступним завдяки використанню недорогих апаратних засобів, таких як Raspberry Pi, що робить його придатним для впровадження навіть у невеликих організаціях.

Однак етичні та безпекові аспекти залишаються важливими викликами. Для забезпечення довіри та дотримання конфіденційності необхідно впроваджувати заходи анонімізації, шифрування даних, регулярних аудитів та прозорого пояснення мети використання системи. Справедливість моделей залежить від мінімізації упереджень у наборах даних, а також від уважного підходу до їхньої адаптації для різних демографічних груп.

Майбутнє таких систем лежить у поєднанні розпізнавання емоцій із мультимодальними даними, включаючи мовні сигнали, фізіологічні показники та навіть контекстуальні дані. Завдяки цьому можна забезпечити ще більшу точність та розширити спектр застосувань. Успіх цього проекту створює основу для створення ще більш потужних, орієнтованих на користувача систем, які матимуть значний позитивний вплив на суспільство, водночас дотримуючись етичних принципів та гарантій конфіденційності.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Baloglu, Ş., & Kozak, M. (2016). The impact of airport retail environment on passenger emotions and behaviour. ResearchGate. URL: https://www.researchgate.net/publication/303485011_The_impact_of_airport_retail_environment_on_passenger_emotions_and_behaviour (дата звернення: 15.10.2024).
2. Kumar, P., & Lata, S. (2018). Emotion detection from text and speech: A survey. ResearchGate. URL: https://www.researchgate.net/publication/324042746_Emotion_Detection_from_Text_and_Speech_-_A_Survey (дата звернення: 15.10.2024).
3. DMEXCO. (n.d.). Emotion AI: The artificial emotional intelligence. DMEXCO. URL: <https://dmexco.com/stories/emotion-ai-the-artificial-emotional-intelligence> (дата звернення: 15.10.2024).
4. ThinkEQ. (n.d.). Artificial intelligence versus emotional intelligence. ThinkEQ. URL: <https://www.thinkeq.com/artificial-intelligence-versus-emotional-intelligence> (дата звернення: 15.10.2024).
5. Cherry, K. (2021). Theories of emotion. Verywell Mind. URL: <https://www.verywellmind.com/theories-of-emotion-2795717> (дата звернення: 15.10.2024).
6. ResearchGate. (n.d.). Valence, arousal, and dominance: 3D dimensions of emotion. ResearchGate. URL: <https://www.researchgate.net/figure/valence-arousal-and-dominance-are-3D-dimensions-of-emotion> (дата звернення: 16.10.2024).
7. ResearchGate. (n.d.). Recognition of emotional state based on handwriting analysis and psychological assessment. ResearchGate. URL: https://www.researchgate.net/publication/364112989_Recognition_of_Emotional_State_Based_On_Handwriting_Analysis_and_Psychological_Assessment (дата звернення: 16.10.2024).
8. Schuller, B. (2021). Bringing emotion recognition out of the lab into real life: Recent advances in sensors and machine learning. ResearchGate. URL: https://www.researchgate.net/publication/358515968_Bringing_Emotion_Recognition_Out_of_the_Lab_into_Real_Life_Recent_Advances_in_Sensors_and_Machine_Learning (дата звернення: 16.10.2024).
9. Ekman, P. (2022). Do lie detectors actually work? Paul Ekman Blog. URL: <https://www.paulekman.com/blog/do-lie-detectors-actually-work> (дата звернення: 19.10.2024).
10. Darwin, C. (1872). The expression of the emotions in man and animals. ResearchGate. URL: https://www.researchgate.net/publication/364777063_The_Expression_of_the_Emotions_in_Man_and_Animals (дата звернення: 19.10.2024).
11. Viso.ai. (n.d.). Visual emotion AI recognition. Viso AI. URL: <https://viso.ai/deep-learning/visual-emotion-ai-recognition> (дата звернення: 19.10.2024).

12. TELUS Digital. (n.d.). Difference between CNN and RNN. TELUS Digital. URL: <https://www.telusdigital.com/insights/ai-data/article/difference-between-cnn-and-rnn> (дата звернення: 20.10.2024).
13. Ramirez, J., & Vanhara, P. (2020). Review on psychological stress detection using biosignals. ResearchGate. URL: https://www.researchgate.net/publication/334352064_Review_on_psychological_stress_detection_using_biosignals (дата звернення: 20.10.2024).
14. Singh, R. (2023). An efficient approach to face emotion recognition with convolutional neural networks. ResearchGate. URL: https://www.researchgate.net/publication/371658147_An_Efficient_Approach_to_Face_Emotion_Recognition_with_Convolutional_Neural_Networks (дата звернення: 20.10.2024).
15. Pantic, M. (2005). Emotion recognition in human-computer interaction. ResearchGate. URL: https://www.researchgate.net/publication/3321357_Emotion_recognition_in_human-computer_interaction (дата звернення: 21.10.2024).
16. MDPI. (2022). Sensors for emotion recognition. MDPI. URL: <https://www.mdpi.com/2076-3417/12/3/1269> (дата звернення: 22.10.2024).
17. Vejalla, K. (2021). HCI for people with disabilities. Medium. URL: <https://medium.com/@kalyan.vejalla/hci-for-people-with-disabilities-979b6371467b> (дата звернення: 22.10.2024).
18. Sciencedirect. (2020). Emotion recognition technologies. ScienceDirect. URL: <https://www.sciencedirect.com/science/article/pii/S1532046419301686> (дата звернення: 22.10.2024).
19. Dialzara. (n.d.). AI emotion detection for customer service guide. Dialzara. URL: <https://dialzara.com/blog/ai-emotion-detection-for-customer-service-guide> (дата звернення: 25.10.2024).
20. LinkedIn. (2023). Emotion AI and mental healthcare. LinkedIn. URL: <https://www.linkedin.com/pulse/emotion-ai-mental-healthcare-bridging-gap-between-technology-compassionate-o32tc> (дата звернення: 25.10.2024).
21. Affectiva. (2021). Driver emotion recognition and real-time facial analysis. Affectiva Blog. URL: <https://blog.affectiva.com/driver-emotion-recognition-and-real-time-facial-analysis-for-the-automotive-industry> (дата звернення: 25.10.2024).
22. Panichkriangkrai, P., & Silapasuphakornwong, K. (2020). Emotion recognition of students during e-learning. Semantic Scholar. URL: <https://www.semanticscholar.org/paper/Emotion-recognition-of-students-during-e-learning-Panichkriangkrai-Silapasuphakornwong/7ab5147a54b9e066d037006aa7ee0f1744ec7e75> (дата звернення: 25.10.2024).
23. Pentland, A., & Vinciarelli, A. (2010). Social signal processing: The research agenda. ResearchGate. URL: https://www.researchgate.net/publication/254900623_Social_Signal_Processing_The_Research_Agenda (дата звернення: 25.10.2024).
24. Klassen Performance Group. (n.d.). Self-awareness: The foundation of emotional intelligence. Klassen Performance Group. URL:

<https://klassenperformancegroup.com/self-awareness-the-foundation-of-emotional-intelligence> (дата звернення: 26.10.2024).

25. Springer. (2022). Advanced emotion recognition techniques. SpringerLink. URL: <https://link.springer.com/article/10.1007/s11042-022-13558-9> (дата звернення: 26.10.2024).

26. Morphcast. (2023). Emotion AI and wearable technology. Morphcast Blog. URL: <https://www.morphcast.com/blog/emotion-ai-and-wearable-technology> (дата звернення: 26.10.2024).

27. The New Economy. (2023). The problem with emotion detection technology. The New Economy. URL: <https://www.theneweconomy.com/technology/the-problem-with-emotion-detection-technology> (дата звернення: 27.10.2024).

28. Springer. (2023). AI-based emotion recognition models. SpringerLink. URL: <https://link.springer.com/article/10.1007/s43681-023-00307-3> (дата звернення: 27.10.2024).

29. Ekman, P. (2018). The subjective experience of emotion: A fearful view. URL: ResearchGate. https://www.researchgate.net/publication/322864443_The_subjective_experience_of_emotion_a_fearful_view (дата звернення: 27.10.2024).

30. SAGE Journals. (2016). Emotion processing and decision-making. SAGE Journals. URL: <https://journals.sagepub.com/doi/10.1080/17470218.2016.1240816> (дата звернення: 28.10.2024).

31. Zhang, W., & Li, J. (2022). Bias on multimodal emotion detection algorithms. ResearchGate. URL: <https://www.researchgate.net/publication/360334153> (дата звернення: 28.10.2024).

32. Sciencedirect. (2024). EEG, ECG, and GSR in emotion recognition. ScienceDirect. URL: <https://www.sciencedirect.com/science/article/abs/pii/S1746809424010978> (дата звернення: 29.10.2024).

33. Sprh Labs. (2021). Understanding deep learning: DNN, RNN, LSTM, CNN, and R-CNN. Medium. URL: <https://medium.com/@sprhlabs/understanding-deep-learning-dnn-rnn-lstm-cnn-and-r-cnn-6602ed94dbff> (дата звернення: 29.10.2024).

34. IBM. (n.d.). What are SVMs? IBM. URL: <https://www.ibm.com/topics/support-vector-machine> (дата звернення: 29.10.2024).

35. Corporate Finance Institute. (n.d.). Ensemble methods in data science. Corporate Finance Institute. URL: <https://corporatefinanceinstitute.com/resources/data-science/ensemble-methods> (дата звернення: 29.10.2024).

36. Nature. (2022). Advances in emotion recognition. Nature Scientific Reports. URL: <https://www.nature.com/articles/s41598-022-11173-0> (дата звернення: 29.10.2024).

37. Asmai, H. (2023). Speech emotion recognition using CNN. Medium. URL: <https://hami-asmai.medium.com/project-speech-emotion-recognition-using-cnn-with-crema-d-dataset-using-google-colab-8cb4fdbd5044> (дата звернення: 30.10.2024).

38. Garg, M. (n.d.). Multimodal datasets repository. GitHub. URL: <https://github.com/drmuskangarg/Multimodal-datasets> (дата звернення: 30.10.2024).

39. The Guardian. (2024). Emotional AI and algorithmic bias. The Guardian. URL: <https://www.theguardian.com/technology/article/2024/jun/23/emotional-artificial-intelligence-chatgpt-4o-hume-algorithmic-bias> (дата звернення: 30.10.2024).
40. MathWorks. (n.d.). Active shape and appearance models. MATLAB Central. URL: <https://www.mathworks.com/matlabcentral/fileexchange/26706-active-shape-model-asm-and-active-appearance-model-aam> (дата звернення: 30.10.2024).
41. Nguyen, Q. (2021). Real-time feeling detection through facial expression recognition. ResearchGate. URL: <https://www.researchgate.net/publication/340279880> (дата звернення: 04.11.2024).
42. Silva, A. (2022). Generalized zero-shot emotion recognition. ResearchGate. URL: https://www.researchgate.net/publication/355825123_Generalized_zero-shot_emotion_recognition_from_body_gestures (дата звернення: 04.11.2024).
43. IEEE. (2020). Emotion recognition in e-learning environments. IEEE Xplore. URL: <https://ieeexplore.ieee.org/document/8999058> (дата звернення: 04.11.2024).
44. Techtoday. (2024). Uncanny valley effect in robotics. Techtoday. URL: <https://techtoday.in.ua/news/efekt-motoroshtnoyi-dolyny-chy-potribno-robyty-robotiv-shozhymy-na-lyudej-118087.html> (дата звернення: 05.11.2024).
45. Analytics Vidhya. (2022). Speech emotion detection using transcribed text. Medium. URL: <https://medium.com/analytics-vidhya/speech-emotion-detection-using-voice-transcribed-text-part-1-ce1b9e3e3c12> (дата звернення: 07.11.2024).
46. Bachorowski, J. (2005). Voice and speech analysis in emotion detection. ResearchGate. URL: https://www.researchgate.net/publication/224929562_Voice_and_Speech_Analysis_in_Search_of_States_and_Traits (дата звернення: 07.11.2024).
47. Lee, S. (2023). Comparing pre-trained algorithms for facial emotion recognition. ResearchGate. URL: https://www.researchgate.net/publication/375867471_Comparing_pre-trained_algorithms_for_facial_emotion_recognition_An_analysis (дата звернення: 07.11.2024).
48. MDPI. (2022). Facial emotion recognition using AI. MDPI Sensors. URL: <https://www.mdpi.com/1424-8220/22/21/8467> (дата звернення: 08.11.2024).
49. Javatpoint. (n.d.). K-nearest neighbor algorithm. Javatpoint. URL: <https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning> (дата звернення: 08.11.2024).
50. Springer. (2023). Transformers in AI applications. SpringerLink. URL: <https://link.springer.com/article/10.1007/s43681-023-00307-3> (дата звернення: 09.11.2024).
51. AWS. (n.d.). What is Transformers in artificial intelligence? AWS. URL: <https://aws.amazon.com/what-is/transformers-in-artificial-intelligence/> (дата звернення: 09.11.2024).
52. PeerJ. (2022). Advances in deep learning applications. PeerJ Computer Science, 8, Article CS-1154. URL: <https://peerj.com/articles/cs-1154/> (дата звернення: 09.11.2024).
53. Muskangarg, D. (2023). InMyFace: Sensor fusion for wearable facial activity recognition. ResearchGate. URL:

https://www.researchgate.net/publication/368361671_InMyFace_Inertial_and_Mechanomyography-Based_Sensor_Fusion_for_Wearable_Facial_Activity_Recognition (дата звернення: 11.11.2024).

54. Anaconda, Inc. (n.d.). User guide for Anaconda.org. Anaconda Documentation. URL: <https://docs.anaconda.com/anacondaorg/user-guide/> (дата звернення: 11.11.2024).

55. Project Jupyter. (n.d.). Jupyter Documentation. Jupyter Docs. URL: <https://docs.jupyter.org/en/latest/> (дата звернення: 11.11.2024).

56. TensorFlow Developers. (n.d.). TensorFlow GitHub repository. GitHub. URL: <https://github.com/tensorflow/tensorflow> (дата звернення: 11.11.2024).

57. OpenCV Developers. (n.d.). OpenCV GitHub repository. GitHub. URL: <https://github.com/opencv/opencv> (дата звернення: 11.11.2024).

58. NumPy Developers. (n.d.). NumPy GitHub repository. GitHub. URL: <https://github.com/numpy/numpy> (дата звернення: 11.11.2024).

59. Matplotlib Developers. (n.d.). Matplotlib GitHub repository. GitHub. URL: <https://github.com/matplotlib/matplotlib> (дата звернення: 11.11.2024).

60. Keras Developers. (n.d.). MobileNet API documentation. Keras. URL: <https://keras.io/api/applications/mobilenet/> (дата звернення: 11.11.2024).

61. OpenCV Developers. (n.d.). Haar cascades repository. GitHub. URL: <https://github.com/opencv/opencv/tree/master/data/haarcascades> (дата звернення: 11.11.2024).

62. Kaggle. (n.d.). FER2013 dataset. Kaggle. URL: <https://www.kaggle.com/datasets/msmbare/fer2013/data> (дата звернення: 11.11.2024).

63. ImageNet. (n.d.). ImageNet database. ImageNet. URL: <https://www.image-net.org/> (дата звернення: 11.11.2024).

64. TechTarget. (n.d.). Transfer learning definition. TechTarget. URL: <https://www.techtarget.com/searchcio/definition/transfer-learning> (дата звернення: 12.11.2024).

65. Balvak, A., Demchenko, M., Drahun, V., Petrenko, V., Prysiazhniuk, V., Ivanov, O., Korotkov, S., Tsvyk, O., Tonkykh, O., Antonenko, A., Lemeshko, A., Hafanovych, V., & Holubenko, O. (2023). BUILDING CORPORATE NETWORKS FOR FOOD INDUSTRY ENTERPRISES. European Science, 1(sge22-01), 25–47. <https://doi.org/10.30890/2709-2313.2023-22-01-031>

66. Драгун В.П., Лащевська Н.О. Загальні аспекти застосування штучного інтелекту в медицині. «Сучасні інтелектуальні інформаційні технології в науці та освіті»: IV Всеукр. науково-практ. конф., м.Київ, 15 травня 2024 р. С. 25-26. URL: https://duikt.edu.ua/uploads/p_2661_45318838.pdf

67. Драгун В.П., Лащевська Н.О. Застосування автоматизованого навчання в області медицини. «Сучасні інтелектуальні інформаційні технології в науці та освіті»: IV Всеукр. науково-практ. конф., м.Київ, 15 травня 2024 р. С. 294-295. URL: https://duikt.edu.ua/uploads/p_2661_45318838.pdf

68. Драгун В.П., Вечерковська А.С. Розпізнавання емоцій на основі штучного інтелекту за допомогою трансферного навчання та аналізу в реальному часі. «Проблеми комп'ютерної інженерії»: Науково-практ. конф., м.Київ, 3 грудня 2024 р. С.89-92. URL: подано на публікацію

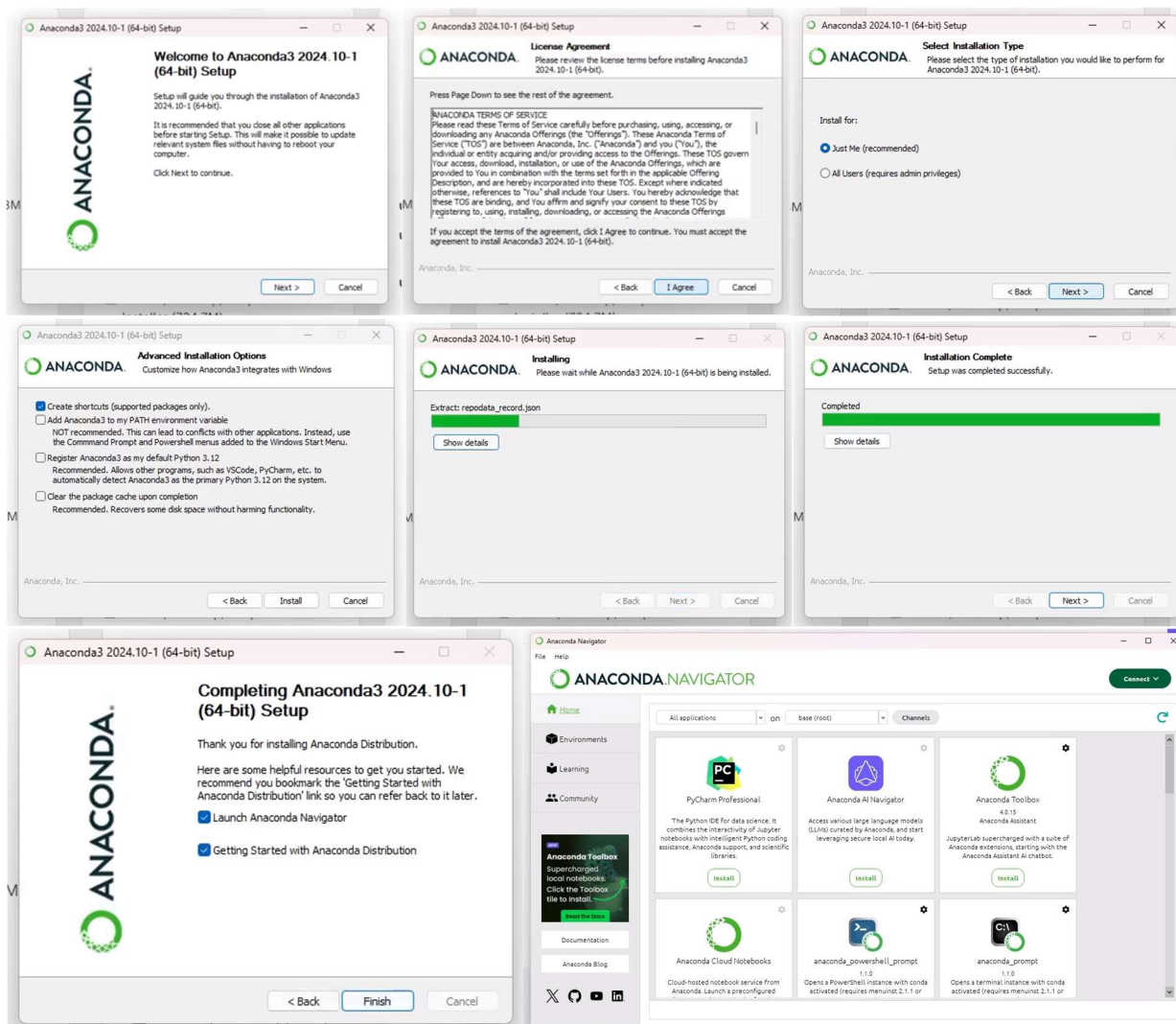
69. Драгун В.П. Вєчерковська А.С. Чому важливо навчити штучний інтелект емоційним станам та як це допоможе уникнути «моторошної долини». «Проблеми комп'ютерної інженерії» : Науково-практ. конф., м.Київ, 3 грудня 2024 р. С.93-95.

Порівняльний аналіз алгоритмів і підходів до розпізнавання емоцій

Таблиця А.1 – Порівняльний аналіз алгоритмів і підходів до розпізнавання емоцій

Алгоритм/Підхід	Сильні сторони	Обмеження	Найкращі програми
Машини підтримки векторів (SVM)	Точність, нелінійна обробка даних	Масштабованість, чутливість до шуму	Невеликі набори даних, аналіз голосу
Дерева рішень та випадкові ліси	Можливість інтерпретації, нелінійне захоплення ознак	Переобладнання, обмежена масштабованість	Аналіз настрою тексту, фізіологічні дані
Згорточні нейронні мережі (CNN)	Візуальне виділення функцій, надійність	Висока вартість обчислення, потреба у великій кількості даних	Розпізнавання виразу обличчя та жестів
Повторювані нейронні мережі (RNN)	Послідовний аналіз даних, часові особливості	Повільне навчання, ризик перенастроювання	Аналіз мови, фізіологічного сигналу
Моделі трансформерів	Мультимодальні можливості, обробка тексту	Обчислювальна вартість, складність	Розпізнавання емоцій тексту та мови

Конфігурація середовища та засобів розробки



Рисунки Б.1-Б.8 – Встановлення ПЗ Anaconda Navigator

Конфігурація середовища та засобів розробки

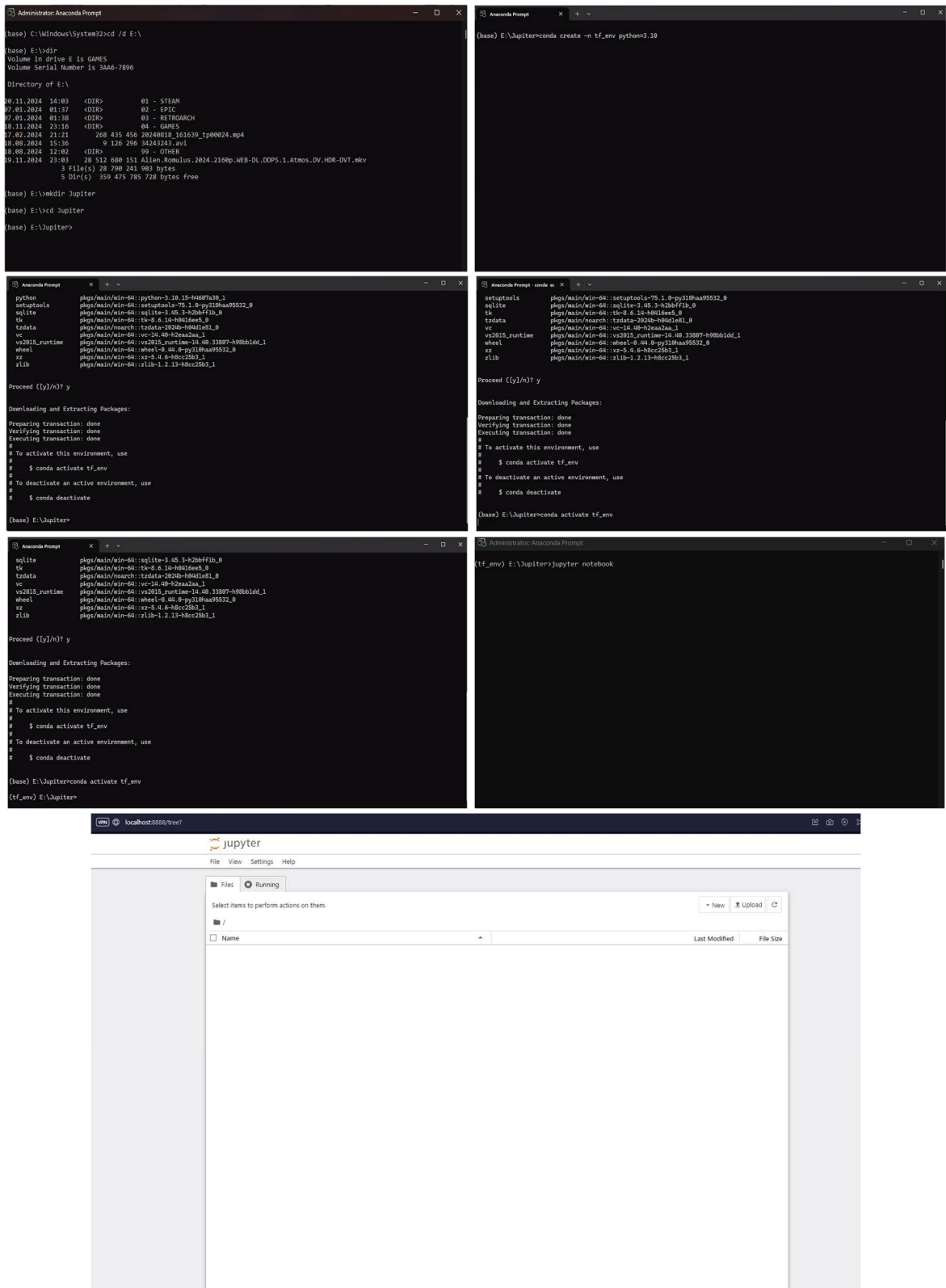


Рисунок Б.9-Б.16 – Створення, активація та запуск нового віртуального середовища та інструменту Jupyter Notebook (через Anaconda Prompt)

Конфігурація середовища та засобів розробки

```

Anaconda Prompt - pip install x + v
zlib      pldg/main/win-64::zlib-1.2.13-h8cc25b3_1

Proceed ([y]/n)? y

Downloading and Extracting Packages:
Preparing transaction: done
Verifying transaction: done
Executing transaction: done
#
# To activate this environment, use
#
#   $ conda activate tf_env
#
# To deactivate an active environment, use
#
#   $ conda deactivate

(base) E:\Jupyter>conda activate tf_env

(tf_env) E:\Jupyter>pip install tensorflow
Collecting tensorflow
  Downloading tensorflow-2.18.0-cp310-cp310-win_amd64.whl.metadata (3.3 kB)
Collecting tensorflow-intel==2.18.0 (from tensorflow)
  Downloading tensorflow_intel-2.18.0-cp310-cp310-win_amd64.whl.metadata (4.9 kB)
Collecting absl-py==0.9 (from tensorflow-intel==2.18.0->tensorflow)
  Using cached absl_py-2.1.0-py3-none-any.whl.metadata (2.3 kB)
Collecting rich (from keras==3.5.0->tensorflow-intel==2.18.0->tensorflow)
  Using cached rich-13.9.0-py3-none-any.whl.metadata (18 kB)
Collecting namee (from keras==3.5.0->tensorflow-intel==2.18.0->tensorflow)
  Using cached namee-0.8.8-py3-none-any.whl.metadata (206 bytes)
Collecting optree (from keras==3.5.0->tensorflow-intel==2.18.0->tensorflow)
  Downloading optree-0.13.1-cp310-cp310-win_amd64.whl.metadata (48 kB)
Collecting charset-normalizer==4.* (from requests==3.*>2.21.0->tensorflow-intel==2.18.0->tensorflow)
  Downloading charset_normalizer-4.0.0-cp310-cp310-win_amd64.whl.metadata (30 kB)
Collecting idna==2.5 (from requests==3.*>2.21.0->tensorflow-intel==2.18.0->tensorflow)
  Using cached idna-3.10-py3-none-any.whl.metadata (10 kB)
Collecting urllib3==2.2.3-py3-none-any.whl.metadata (6.5 kB)
Collecting certifi==2024.4.17 (from requests==3.*>2.21.0->tensorflow-intel==2.18.0->tensorflow)
  Using cached certifi-2024.8.30-py3-none-any.whl.metadata (2.2 kB)
Collecting markdown==2.6.8 (from tensorboard==2.19.*>2.18.0->tensorflow-intel==2.18.0->tensorflow)
  Using cached Markdown-3.7-py3-none-any.whl.metadata (7.0 kB)
Collecting tensorboard-data-server==0.8.0.*>0.7.0 (from tensorboard==2.19.*>2.18.0->tensorflow-intel==2.18.0->tensorflow)
  Using cached tensorboard_data_server-0.7.2-py3-none-any.whl.metadata (1.1 kB)
Collecting werkzeug==1.0.1 (from tensorboard==2.19.*>2.18.0->tensorflow-intel==2.18.0->tensorflow)
  Using cached werkzeug-3.1.3-py3-none-any.whl.metadata (3.7 kB)
Collecting MarkupSafe==2.1.1 (from werkzeug==1.0.1->tensorboard==2.19.*>2.18.0->tensorflow-intel==2.18.0->tensorflow)
  Downloading MarkupSafe-3.0.2-cp310-cp310-win_amd64.whl.metadata (4.1 kB)
Collecting markdown-it-py==2.2.0 (from rich->keras==3.5.0->tensorflow-intel==2.18.0->tensorflow)
  Using cached markdown_it_py-3.0.0-py3-none-any.whl.metadata (6.9 kB)
Collecting pygments==2.18.0-py3-none-any.whl.metadata (2.5 kB)
Collecting mdurl==0.1 (from markdown-it-py==2.2.0->rich->keras==3.5.0->tensorflow-intel==2.18.0->tensorflow)
  Using cached mdurl-0.1.2-py3-none-any.whl.metadata (1.6 kB)
Download tensorflow-2.18.0-cp310-cp310-win_amd64.whl (795.0 MB)
Download tensorflow_intel-2.18.0-cp310-cp310-win_amd64.whl (390.0 MB)
62.4/390.0 MB 176.6 KB/s eta 0:31:03

(tf_env) E:\Jupyter>pip install opencv-python
(tf_env) E:\Jupyter>pip install opencv-contrib-python

```

Рисунок Б.17-Б.22 – Встановлення бібліотек, плагінів та залежностей

Підготовка набору даних для тренування моделі

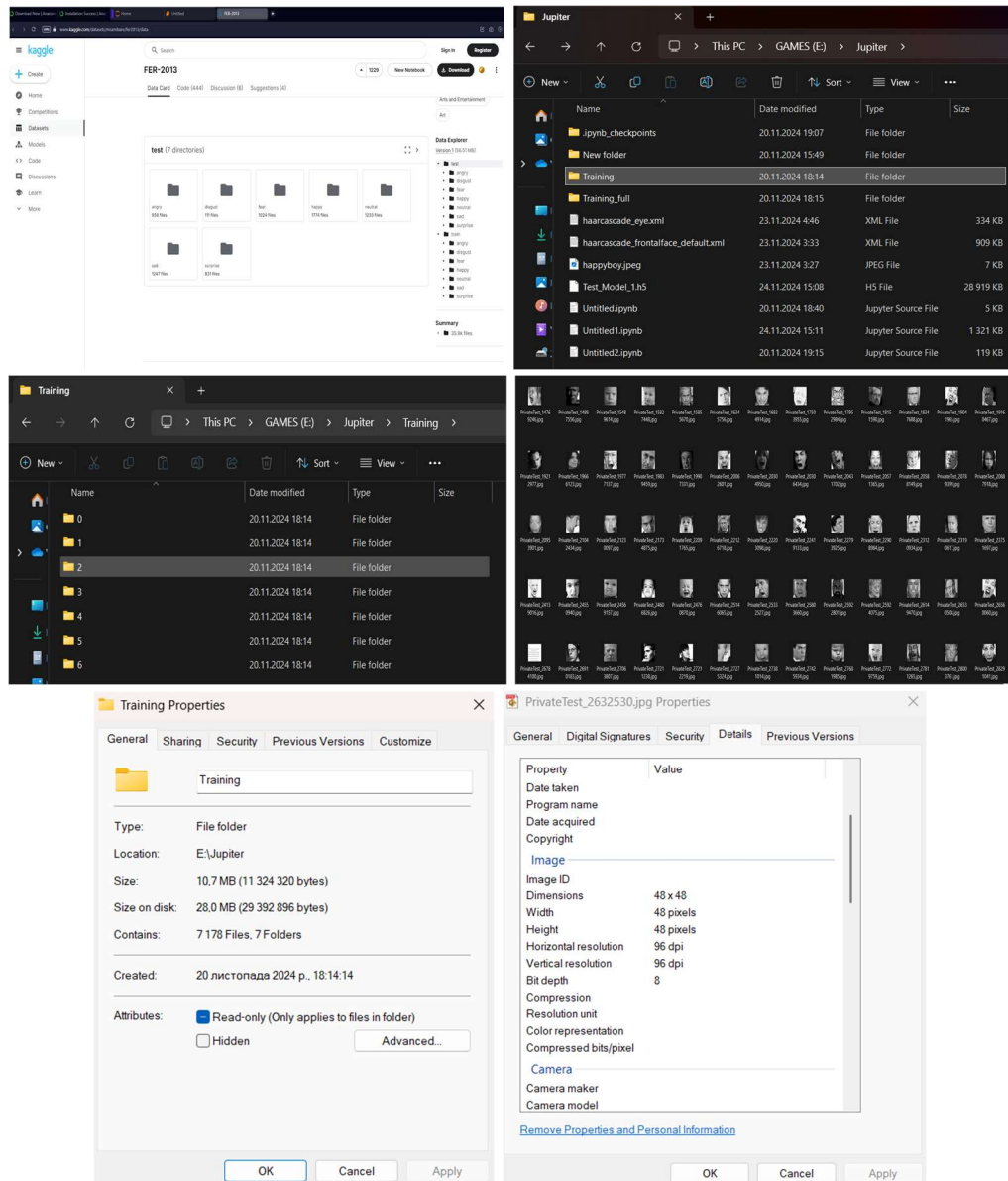
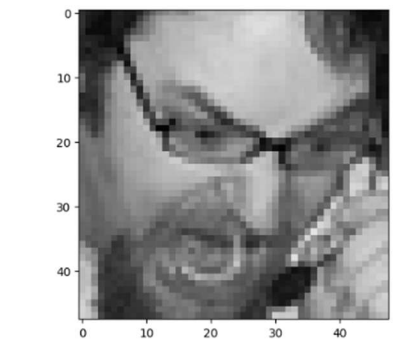


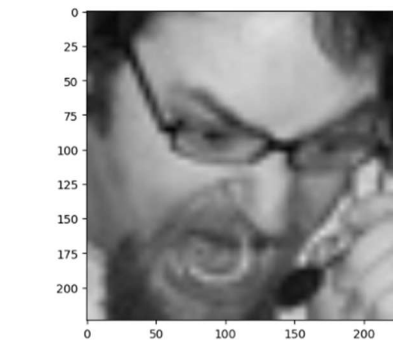
Рисунок В.1-В.6 – Загальнодоступний набір даних FER-2013
(для тренування моделі на основі зображень)

Розроблений модуль розпізнавання емоційних станів людини (Python)



```
In [8]: img_size= 224 ## ImageNet => 224 x 224
new_array= cv2.resize(img_array, (img_size,img_size))
plt.imshow(cv2.cvtColor(new_array, cv2.COLOR_BGR2RGB))
plt.show
```

```
Out[8]: <function matplotlib.pyplot.show(close=None, block=None)>
```



```
In [9]: new_array.shape
```

```
Out[9]: (224, 224, 3)
```

```
In [1]: import tensorflow as tf ## pip install tensorflow-gpu
import cv2 ## pip install opencv-python
## pip install opencv-contrib-python fullpackage
import os
import matplotlib.pyplot as plt ## pip install matplotlib
import numpy as np ## pip install numpy
```

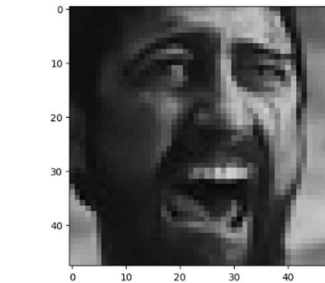
```
In [2]: img_array = cv2.imread("New folder/Training/0/PrivateTest_1488292.jpg")
```

```
In [3]: img_array.shape # rgb
```

```
Out[3]: (48, 48, 3)
```

```
In [4]: plt.imshow(img_array) ## BGR
```

```
Out[4]: <matplotlib.image.AxesImage at 0x1680b0bda6>
```



```
In [5]: Datadirectory = "Training/" ## training dataset
```

```
In [6]: Classes = ["0","1","2","3","4","5","6"] ## List of classes => exact name of your folders
```

```
In [7]: for category in Classes:
    path = os.path.join(Datadirectory, category) ###
    for img in os.listdir(path):
        img_array = cv2.imread(os.path.join(path, img))
        #convert rgb to cv2.cvtColor(img_array, cv2.COLOR_BGR2RGB)
        plt.imshow(cv2.cvtColor(img_array, cv2.COLOR_BGR2RGB))
        plt.show()
    break
```

read all the images and convertin them to array

```
In [10]: training_Data = [] ## data array
```

```
def create_training_Data():
    for category in Classes:
        path = os.path.join(Datadirectory, category)
        class_num = Classes.index(category) ## 0 1, ## Label
        for img in os.listdir(path):
            try:
                img_array = cv2.imread(os.path.join(path, img))
                new_array= cv2.resize(img_array, (img_size, img_size))
                training_Data.append([new_array, class_num])
            except Exception as e:
                pass
```

```
In [11]: create_training_Data()
```

```
In [12]: print(len(training_Data))
```

```
7178
```

```
In [13]: import random
random.shuffle(training_Data)
```

```
In [14]: X = [] ## data /feature
y = [] ## Label
```

```
for features, label in training_Data:
    X.append(features)
    y.append(label)

X = np.array(X).reshape(-1, img_size, img_size, 3) ## converting it in to 4 dimension
```

```
In [15]: X.shape
```

```
Out[15]: (7178, 224, 224, 3)
```

```
In [16]: # normalize the data
X= X/255.0; ## we are normalizing it
```

```
In [17]: type(y)
```

```
Out[17]: list
```

```
In [18]: Y= np.array(y)
```

```
In [19]: Y.shape
```

```
Out[19]: (7178,)
```

deep learning model for training - Transfer Learning

```
In [20]: import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers
```

```
In [21]: model = tf.keras.applications.MobileNetV2() ## Pre-trained Model
```

```
In [22]: model.summary()
```

```
Model: "mobilenetv2_1.00_224"
```

Layer (type)	Output Shape	Param #	Connected to
input_layer (InputLayer)	(None, 224, 224, 3)	0	-
Conv1 (Conv2D)	(None, 112, 112, 32)	864	input_layer[0][0]
bn_Conv1 (BatchNormalization)	(None, 112, 112, 32)	128	Conv1[0][0]
Conv1_relu (ReLU)	(None, 112, 112, 32)	0	bn_Conv1[0][0]
expanded_conv_depthwise (DepthwiseConv2D)	(None, 112, 112, 32)	288	Conv1_relu[0][0]
expanded_conv_depthwise_BN (BatchNormalization)	(None, 112, 112, 32)	128	expanded_conv_depthwise[0]
expanded_conv_depthwise_relu (ReLU)	(None, 112, 112, 32)	0	expanded_conv_depthwise_BN[0]
expanded_conv_project (Conv2D)	(None, 112, 112, 16)	512	expanded_conv_depthwise_relu[0]
expanded_conv_project_BN (BatchNormalization)	(None, 112, 112, 16)	64	expanded_conv_project[0][0]
block_1_expand (Conv2D)	(None, 112, 112, 96)	1,536	expanded_conv_project_BN[0]
block_1_expand_BN (BatchNormalization)	(None, 112, 112, 96)	384	block_1_expand[0][0]
block_1_expand_relu (ReLU)	(None, 112, 112, 96)	0	block_1_expand_BN[0][0]
block_1_pad (ZeroPadding2D)	(None, 113, 113, 96)	0	block_1_expand_relu[0][0]

```
Total params: 3,538,984 (13.50 MB)
```

```
Trainable params: 3,504,872 (13.37 MB)
```

```
Non-trainable params: 34,112 (133.25 KB)
```

Розроблений модуль розпізнавання емоційних станів людини (Python)

Transfer Learning - Tuning, weights will start from last check point

```
In [23]: base_input = model.layers[1].input ## input
In [24]: base_output = model.layers[-2].output
In [25]: base_output
Out[25]: <keras.tensor_shape.Shape object at 0x168bfff34c0>
In [26]: final_output = layers.Dense(128)(base_output) ## adding new layer, after the output of global pooling layer
final_output = layers.Activation('relu')(final_output) ## activation function
final_output = layers.Dense(64)(final_output)
final_output = layers.Activation('relu')(final_output)
final_output = layers.Dense(7, activation='softmax')(final_output) ## my classes are 0? 1x In [47]: final_output ## output
In [27]: final_output ## output
Out[27]: <keras.tensor_shape.Shape object at 0x168bfff34c0>
In [28]: new_model = keras.Model(inputs = base_input, outputs = final_output)
In [29]: new_model.summary()
```

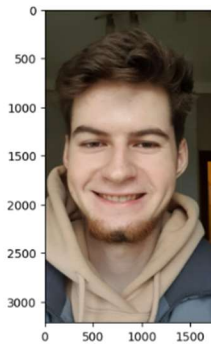
Model: "functional"

Layer (type)	Output Shape	Param #	Connected to
input_layer (InputLayer)	(None, 224, 224, 3)	0	-
conv1 (Conv2D)	(None, 112, 112, 32)	864	input_layer[0][0]
bn_conv1 (BatchNormalization)	(None, 112, 112, 32)	128	conv1[0][0]
conv1_relu (ReLU)	(None, 112, 112, 32)	0	bn_conv1[0][0]
expanded_conv_depthwise (DepthwiseConv2D)	(None, 112, 112, 32)	288	conv1_relu[0][0]
expanded_conv_depthwise_bn (BatchNormalization)	(None, 112, 112, 32)	128	expanded_conv_depthwise[0]
expanded_conv_depthwise_relu (ReLU)	(None, 112, 112, 32)	0	expanded_conv_depthwise_bn[0]
expanded_conv_project (Conv2D)	(None, 112, 112, 16)	512	expanded_conv_depthwise_relu[0]
expanded_conv_project_bn (BatchNormalization)	(None, 112, 112, 16)	64	expanded_conv_project[0][0]
block1_expand (Conv2D)	(None, 112, 112, 96)	1,536	expanded_conv_project_bn[0]
block1_expand_bn (BatchNormalization)	(None, 112, 112, 96)	384	block1_expand[0][0]
block1_expand_relu (ReLU)	(None, 112, 112, 96)	0	block1_expand_bn[0][0]
block1_pad (ZeroPadding2D)	(None, 112, 112, 96)	0	block1_expand_relu[0][0]
block1_depthwise (DepthwiseConv2D)	(None, 56, 56, 96)	864	block1_pad[0][0]

Total params: 2,430,663 (9.27 MB)
Trainable params: 2,390,551 (9.14 MB)
Non-trainable params: 34,112 (133.25 KB)

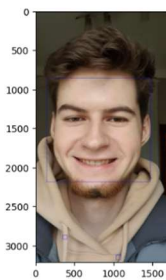
```
In [30]: new_model.compile(loss='sparse_categorical_crossentropy', optimizer = "adam", metrics = ["accuracy"])
In [31]: new_model.fit(X,Y, epochs = 25) ## training
Epoch 1/25
225/225 --- 518s 2s/step - accuracy: 0.3847 - loss: 1.5735
Epoch 2/25
225/225 --- 483s 2s/step - accuracy: 0.5221 - loss: 1.2352
Epoch 3/25
225/225 --- 478s 2s/step - accuracy: 0.5779 - loss: 1.1046
Epoch 4/25
225/225 --- 481s 2s/step - accuracy: 0.6292 - loss: 0.9900
Epoch 5/25
225/225 --- 478s 2s/step - accuracy: 0.6445 - loss: 0.9502
Epoch 6/25
225/225 --- 478s 2s/step - accuracy: 0.6704 - loss: 0.8800
Epoch 7/25
225/225 --- 479s 2s/step - accuracy: 0.6993 - loss: 0.8022
Epoch 8/25
225/225 --- 479s 2s/step - accuracy: 0.7335 - loss: 0.7324
Epoch 9/25
225/225 --- 479s 2s/step - accuracy: 0.7557 - loss: 0.6523
Epoch 10/25
225/225 --- 480s 2s/step - accuracy: 0.7970 - loss: 0.5707
Epoch 11/25
225/225 --- 483s 2s/step - accuracy: 0.8112 - loss: 0.5249
Epoch 12/25
225/225 --- 481s 2s/step - accuracy: 0.8185 - loss: 0.4866
Epoch 13/25
225/225 --- 479s 2s/step - accuracy: 0.8595 - loss: 0.3807
Epoch 14/25
225/225 --- 480s 2s/step - accuracy: 0.8826 - loss: 0.3502
Epoch 15/25
225/225 --- 484s 2s/step - accuracy: 0.8824 - loss: 0.3191
Epoch 16/25
225/225 --- 482s 2s/step - accuracy: 0.9005 - loss: 0.2886
Epoch 17/25
225/225 --- 482s 2s/step - accuracy: 0.9176 - loss: 0.2475
Epoch 18/25
225/225 --- 481s 2s/step - accuracy: 0.9224 - loss: 0.2380
Epoch 19/25
225/225 --- 483s 2s/step - accuracy: 0.9192 - loss: 0.2311
Epoch 20/25
225/225 --- 478s 2s/step - accuracy: 0.9378 - loss: 0.1892
Epoch 21/25
225/225 --- 481s 2s/step - accuracy: 0.9417 - loss: 0.1687
Epoch 22/25
225/225 --- 481s 2s/step - accuracy: 0.9435 - loss: 0.1676
Epoch 23/25
225/225 --- 481s 2s/step - accuracy: 0.9527 - loss: 0.1462
Epoch 24/25
225/225 --- 482s 2s/step - accuracy: 0.9484 - loss: 0.1562
Epoch 25/25
225/225 --- 480s 2s/step - accuracy: 0.9558 - loss: 0.1288
Out[31]: <keras.callbacks.history.history object at 0x168b26cf490>
```

```
In [32]: new_model.save('Final_Model.h5')
WARNING:absl:You are saving your model as an HDF5 file via 'model.save()' or 'keras.saving.save_model()'. This format is not recommended.
In [33]: new_model = tf.keras.models.load_model('Final_Model.h5')
WARNING:absl:Compiled the loaded model, but the compiled metrics have yet to be built. 'model.compile()' must be called before computing metrics during training.
In [34]: new_model.compile(optimizer='adam', loss='sparse_categorical_crossentropy', metrics=['accuracy'])
In [35]: frame = cv2.imread("happy.jpg")
In [36]: frame.shape
Out[36]: (3216, 1888, 3)
In [37]: plt.imshow(cv2.cvtColor(frame, cv2.COLOR_BGR2RGB))
Out[37]: <matplotlib.image.AxesImage at 0x168bfff737c0>
```



we need face detection algorithm (gray image)

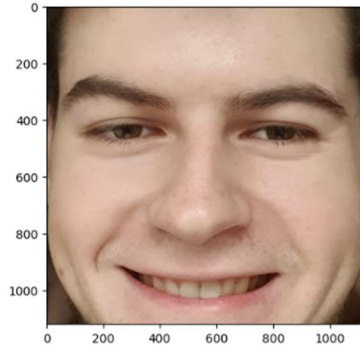
```
In [38]: faceCascade = cv2.CascadeClassifier(cv2.data.haarcascades + 'haarcascade_frontalface_default.xml')
In [39]: gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
In [40]: gray.shape
Out[40]: (3216, 1888)
In [41]: faces = faceCascade.detectMultiScale(gray, 1.1, 4)
for x,y,w,h in faces:
    roi_gray = gray[y:y+h, x:x+w]
    roi_color = frame[y:y+h, x:x+w]
    cv2.rectangle(frame, (x, y), (x+w, y+h), (255, 0, 0), 2) # BGR
    faces = faceCascade.detectMultiScale(roi_gray)
    if len(faces) == 0:
        print("Face not detected")
    else:
        for (ex,ey,ew,eh) in faces:
            face_roi = roi_color[ey:ey+eh, ex:ex+ew]
Face not detected
Face not detected
In [42]: plt.imshow(cv2.cvtColor(frame, cv2.COLOR_BGR2RGB))
Out[42]: <matplotlib.image.AxesImage at 0x168b26cf7580>
```



Розроблений модуль розпізнавання емоційних станів людини (Python)

```
In [43]: plt.imshow(cv2.cvtColor(face_roi, cv2.COLOR_BGR2RGB))
```

```
Out[43]: <matplotlib.image.AxesImage at 0x168861c8ca>
```



```
In [44]: final_image = cv2.resize(face_roi, (224, 224))
final_image = np.expand_dims(final_image, axis=0) ## need fourth dimension
final_image = final_image / 255.0 ## normalizing
```

```
In [45]: Predictions = new_model.predict(final_image)
```

```
1/1 ----- 1s 989ms/step
```

```
In [46]: Predictions[0]
```

```
Out[46]: array([2.3557106e-06, 3.2232190e-05, 7.1506074e-05, 9.9893969e-01,
 8.3603707e-05, 8.7043969e-04, 1.6340196e-07], dtype=float32)
```

```
In [47]: np.argmax(Predictions)
```

```
Out[47]: np.int64(3)
```

```
except Exception as e:
    print(f"Error processing face: {e}")
    status = "Error"

# Display emotion status on the frame
x1, y1, w1, h1 = 0, 0, 175, 75
cv2.rectangle(frame, (x1, y1), (x1 + w1, y1 + h1), (0, 0, 0), -1)
cv2.putText(frame, status, (x1 + 10, y1 + 50), font, 1.2, (0, 255, 0), 2)

# Display the frame
cv2.imshow("Emotion Detection", frame)

# Exit the Loop when 'q' is pressed
if cv2.waitKey(1) & 0xFF == ord('q'):
    break

# Release the video capture and close windows
cap.release()
cv2.destroyAllWindows()
```

WARNING:absl:Compiled the loaded model, but the compiled metrics have yet to be built.

```
1/1 ----- 1s 1s/step
1/1 ----- 0s 41ms/step
1/1 ----- 0s 41ms/step
1/1 ----- 0s 44ms/step
1/1 ----- 0s 44ms/step
```

Realtime Video Demo

```
In [2]: import cv2
import numpy as np

# Load the pre-trained Haar Cascade classifier for face detection
face_cascade = cv2.CascadeClassifier(cv2.data.haarcascades + 'haarcascade_frontalface_default.xml')

# Dummy model for predictions (replace with your trained model)
# Ensure you load your model here (e.g., 'new_model = keras.models.Load_model("model_path")')

from tensorflow.keras.models import load_model

new_model = load_model('Test_Model_1.h5') # Provide the correct path

new_model.compile(loss="sparse_categorical_crossentropy", optimizer="adam", metrics=["accuracy"])

# Set font properties
font_scale = 1.5
font = cv2.FONT_HERSHEY_PLAIN

# Start video capture
cap = cv2.VideoCapture(0)
if not cap.isOpened():
    cap = cv2.VideoCapture(1)
if not cap.isOpened():
    raise IOError("Cannot open webcam")

# Main Loop
while True:
    ret, frame = cap.read()
    if not ret:
        print("Failed to capture frame. Exiting...")
        break

    # Convert the frame to grayscale for face detection
    gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)

    # Detect faces in the frame
    faces = face_cascade.detectMultiScale(gray, scaleFactor=1.1, minNeighbors=4)

    for (x, y, w, h) in faces:
        roi_gray = gray[y:y+h, x:x+w]
        roi_color = frame[y:y+h, x:x+w]

        # Draw rectangle around the face
        cv2.rectangle(frame, (x, y), (x + w, y + h), (255, 0, 0), 2)

        # Resize face for model input
        try:
            face_roi = cv2.resize(roi_color, (224, 224))
            face_roi = np.expand_dims(face_roi, axis=0) # Add batch dimension
            face_roi = face_roi / 255.0 # Normalize

            # Predict emotion (replace with actual model prediction)
            if new_model:
                predictions = new_model.predict(face_roi)
                emotion_index = np.argmax(predictions)
            else:
                emotion_index = -1 # Placeholder for testing without a model

            # Map prediction to emotions
            emotions = ["Angry", "Disgust", "Fear", "Happy", "Sad", "Surprise", "Neutral"]
            status = emotions[emotion_index] if emotion_index >= 0 else "Unknown"
```

ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ



**ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ
ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
Кафедра комп'ютерної інженерії**



**КВАЛІФІКАЦІЙНА РОБОТА
на здобуття освітнього ступеня магістра
НА ТЕМУ: «МОДУЛЬ РОЗПІЗНАВАННЯ ЕМОЦІЙНИХ СТАНІВ
ЛЮДИНИ НА ОСНОВІ ШІ»**

Виконав студент групи КСДМ-61

Драгун Владислав Петрович

Керівник дипломної роботи:

Вечерковська Анастасія Сергіївна,

канд,техн.наук, доцент

Київ – 2025

Мета, об'єкт, предмет дослідження

Мета роботи – розробити модуль розпізнавання емоційних станів людини на основі технологій штучного інтелекту для покращення комунікації між користувачем та комп'ютером

Об'єкт дослідження – емоційні стани людини

Предмет дослідження – процес розпізнавання та аналізу емоційних станів людини за допомогою методів штучного інтелекту

Актуальність роботи

У роботі описано розробку модуля для розпізнавання людських емоцій та його потенціал у покращенні взаємодії в цифровому світі. Використовуючи алгоритми штучного інтелекту та машинного навчання, досліджено методи визначення емоцій за мімікою, голосом та іншими біометричними даними.

Архітектура модуля забезпечує високу точність і швидкість роботи, а результати демонструють перспективи застосування технології в психологічному консультуванні та інтуїтивних взаємодіях людини з машиною. У висновках запропоновано рекомендації для розвитку систем із розширенням діапазону розпізнаваних емоцій.

Проблематика та необхідність впровадження систем розпізнавання емоцій



Рис. 1.21 – Культурні особливості у визначенні емоційних станів



Рис. 2.8 – Ефект моторошної долини (невдала спроба імітації рис людини)



Рис. 3.10 – Дисбаланс класів емоцій

Архітектура запропонованого рішення у вирішенні та покращенні взаємодії «людина-комп'ютер»

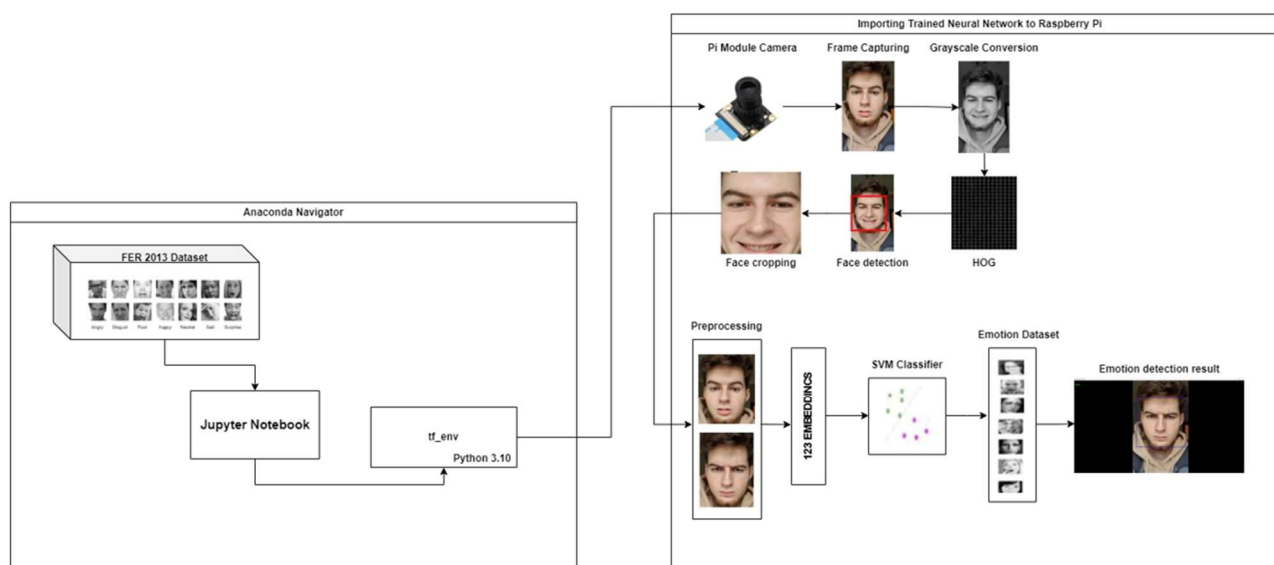


Рис. 3.18 – Процес інтеграції натренованої моделі у Raspberry Pi

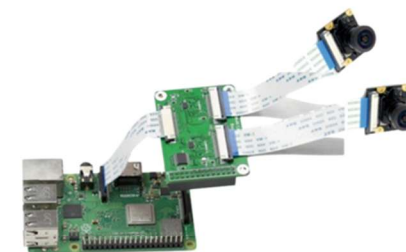


Рис. 3.17 – Raspberry Pi зі адаптером на дві камери для кращої працездатності та інтеграції модуля розпізнавання емоцій

Блок-схема роботи моделі в різних сценаріях: промисловість, транспорт, медицина, громадські місця

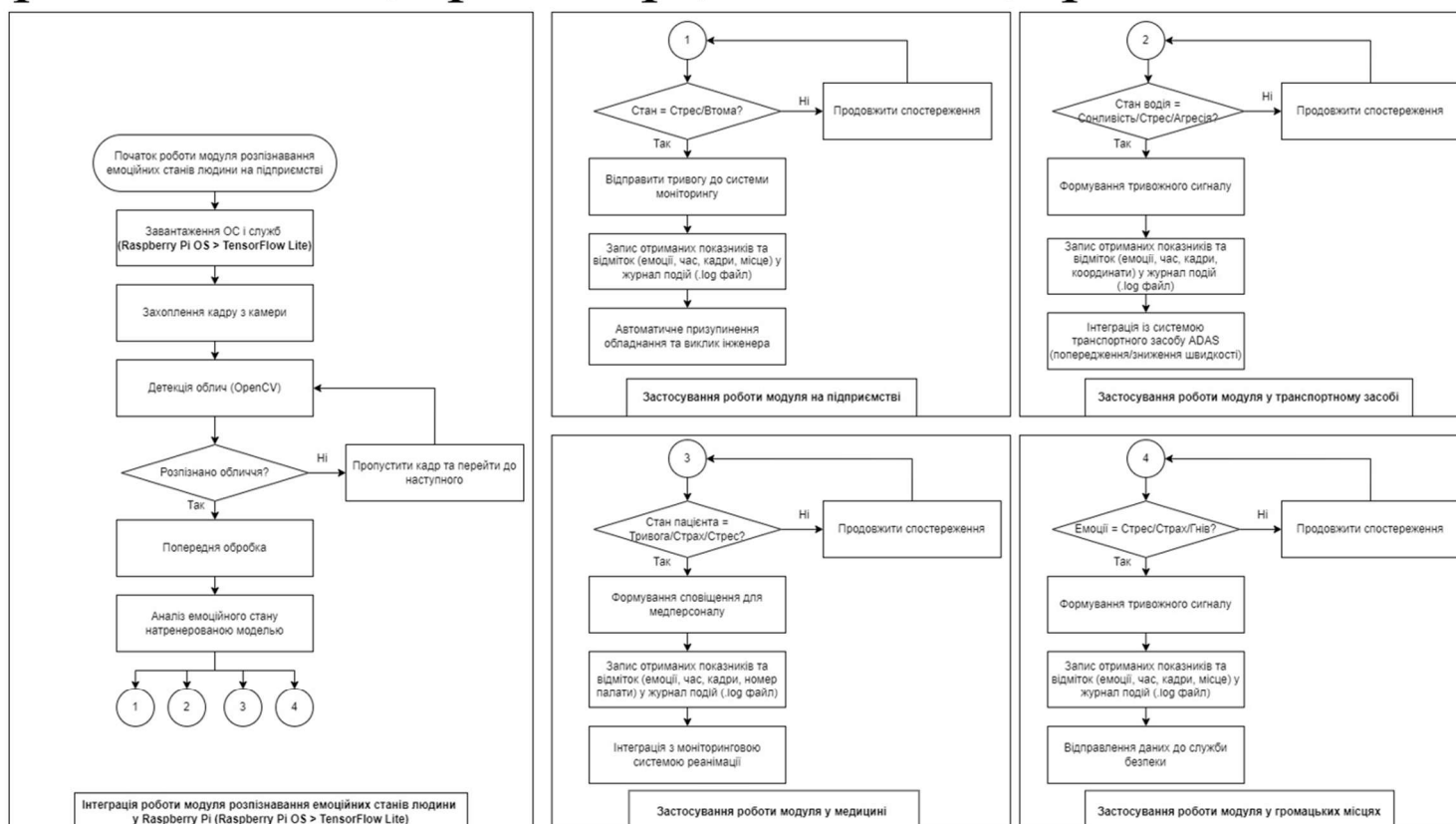


Рис. 3.19 – Різностороннє практичне застосування натренованої моделі розпізнавання емоційних станів людини

Віртуальне середовище та інструмент для розробки моделі розпізнавання емоційних станів людини

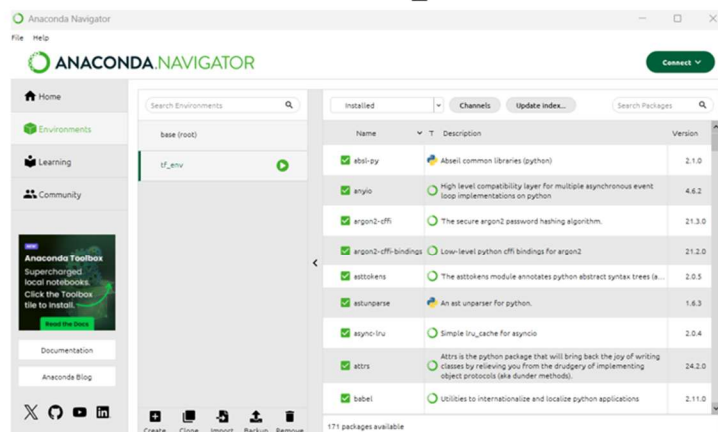


Рис. 3.1 – ПЗ для створення віртуального середовища – Anaconda Navigator

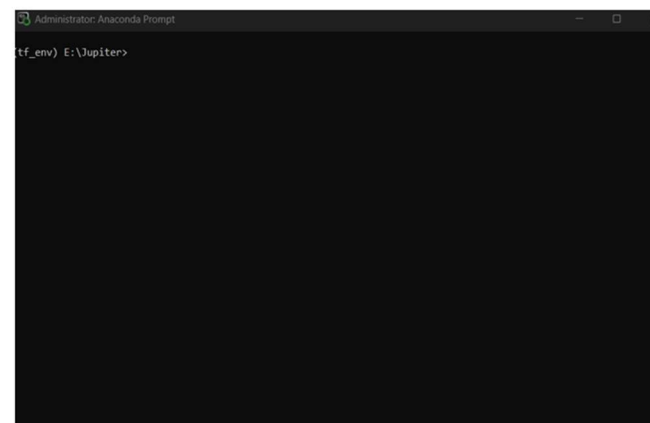


Рис. 3.2 – Командний рядок Anaconda Prompt

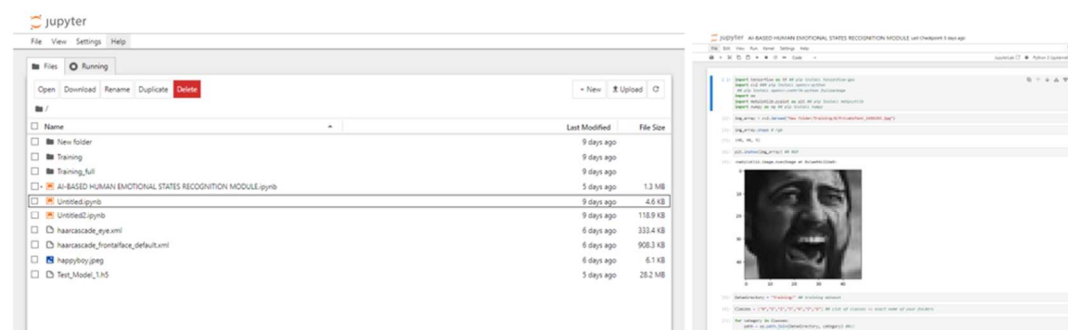


Рис. 3.3 – Інструмент для розробки – Jupyter Notebook

Бібліотеки та залежності для розробки моделі розпізнавання емоційних станів людини



Рис. 3.4 – Поєднання залежностей – TensorFlow та API Keras



Рис. 3.6 – Додаткові бібліотеки – NumPy та Matplotlib

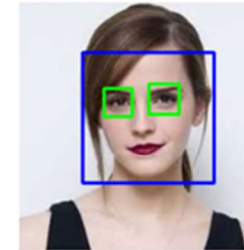


Рис. 3.5 – Інструменти комп'ютерного зору – OpenCV та Haar Cascades

Обраний набір даних та попередньо натренована модель

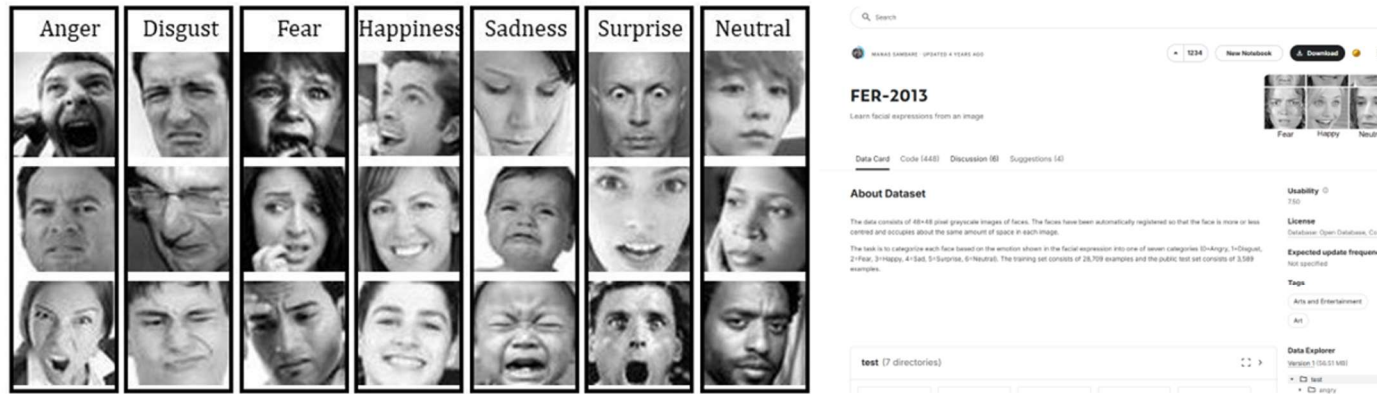


Рис. 3.7 – Приклад вибірки зображень та ресурс з набором даних – FER-2013

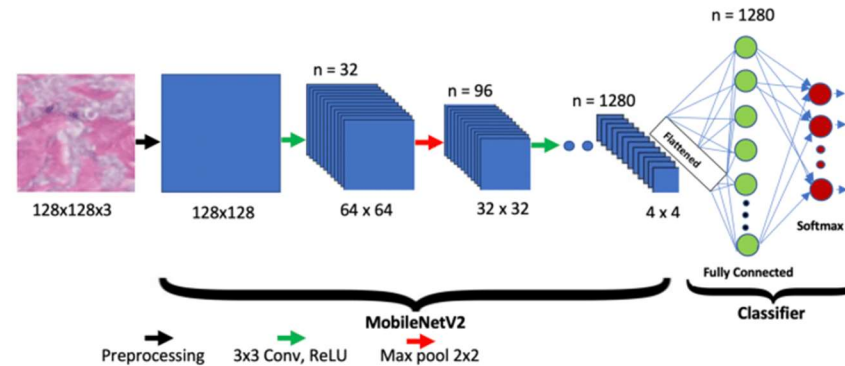


Рис. 3.11 – Мережева архітектура MobileNetV2

Процес тренування моделі розпізнавання емоційних станів людини

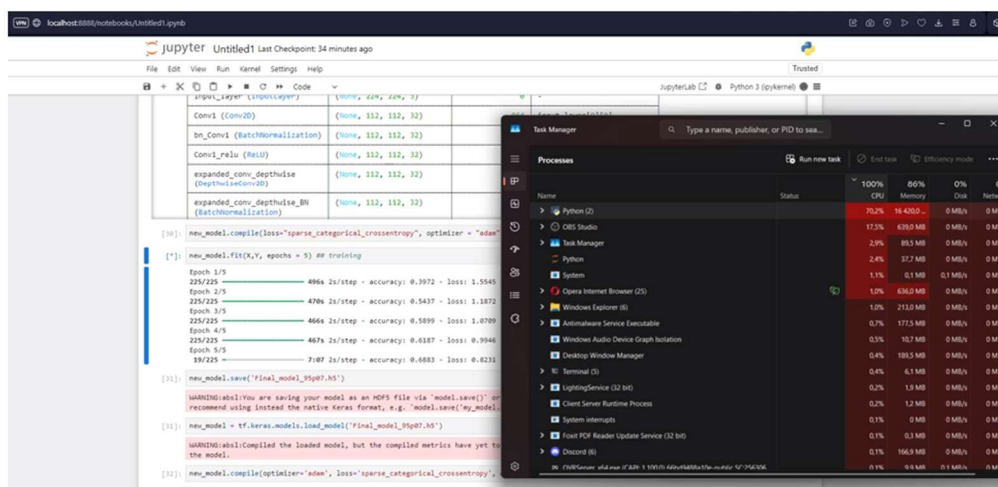


Рис. 3.13.1 – Тренування моделі займає багато часу та ресурсів ПК

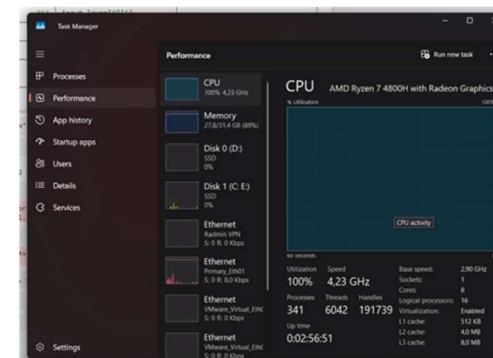
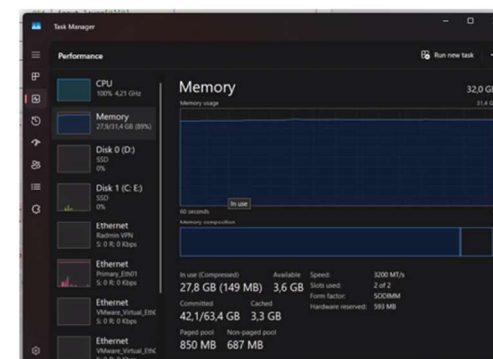


Рис. 3.13.2 – Тренування моделі займає багато часу та ресурсів ПК



Результат натренованої моделі (через зображення)

```
In [42]: plt.imshow(cv2.cvtColor(frame, cv2.COLOR_BGR2RGB))
```

```
Out[42]: <matplotlib.image.AxesImage at 0x16886207580>
```

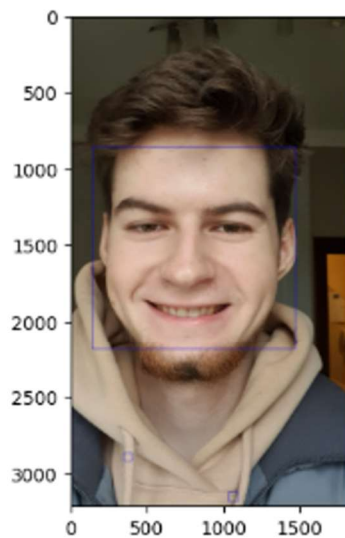


Рис. 3.14 – Розпізнавання обличчя



```
In [45]: Predictions = new_model.predict(final_image)
```

```
1/1 ————— 1s 989ms/step
```

```
In [46]: Predictions[0]
```

```
Out[46]: array([2.3557106e-06, 3.2232190e-05, 7.1506074e-05, 9.9893969e-01,
 8.3603707e-05, 8.7043969e-04, 1.6340196e-07], dtype=float32)
```

```
In [47]: np.argmax(Predictions)
```

```
Out[47]: np.int64(3)
```

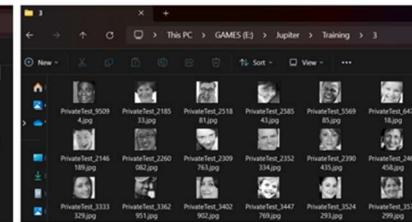
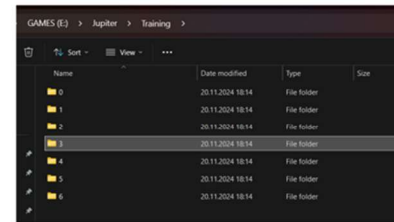


Рис. 3.15– Коректність відпрацювання тестової моделі

Результат натренованої моделі (через веб-камеру)

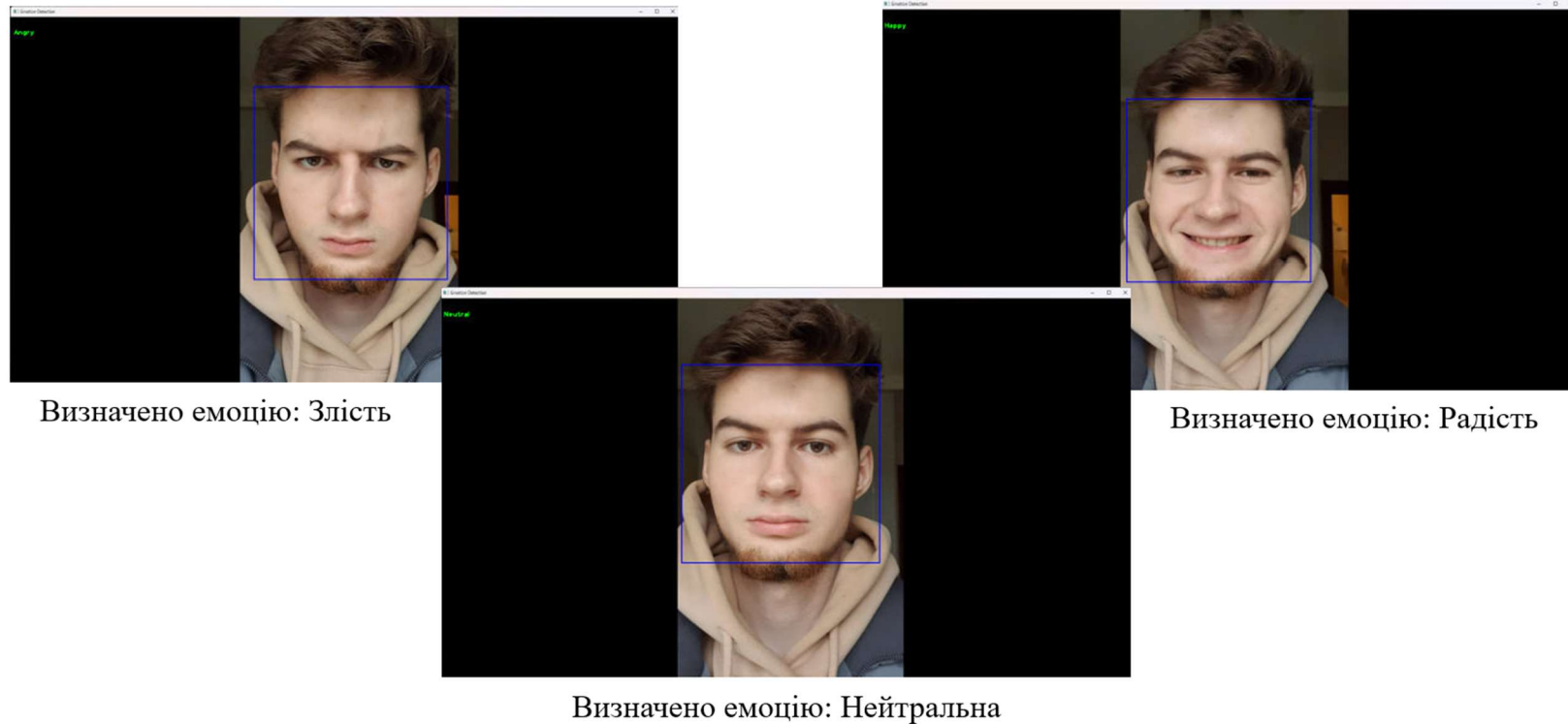


Рис. 3.16 – Демонстрація моделі розпізнавання емоційних станів людини

Порівняння запропонованого рішення з існуючими системами

Критерій	Запропоноване рішення	Існуючі рішення
Апаратні вимоги	Оптимізоване для малопотужних пристроїв (Raspberry Pi)	Потребують потужних серверів або спеціалізованого обладнання
Ефективність у реальному часі	Аналіз відеопотоку з мінімальними затримками	Часто не здатні працювати в реальному часі
Точність розпізнавання емоцій	Використовує FER-2013, добре збалансований набір даних	Можуть покладатися на менш точні або застарілі моделі
Гнучкість у застосуванні	Підходить для різних сфер: промисловість, транспорт, медицина, громадські місця	Часто обмежені однією сферою або вузьким сценарієм
Витрати на впровадження	Низька вартість обладнання та обслуговування	Високі витрати на спеціалізовані системи
Тип аналізу	Комбінує емоційний стан із контекстом ситуації	Зазвичай орієнтовані на фізичні показники (рухи очей, пульс)
Автономність	Може працювати без постійного підключення до мережі	Більшість рішень залежить від хмарних обчислень
Масштабованість	Легка інтеграція з наявними системами та масштабування	Часто складно інтегрувати або адаптувати до нових потреб
Інноваційність	Використовує сучасні глибокі нейронні мережі (MobileNetV2)	Спираються на менш ефективні алгоритми або методи
Розпізнавання складних емоцій	Здатність розпізнавати стрес, агресію, страх тощо	Орієнтовані лише на базові або загальні емоції

Табл. 3.1 – Порівняльний аналіз запропонованого рішення з існуючими системами розпізнавання емоцій

Висновки

Системи розпізнавання емоцій, створені завдяки прогресу машинного навчання та комп'ютерного зору, мають величезний потенціал для трансформації різноманітних галузей, зокрема охорони здоров'я, транспорту, промисловості та безпеки в громадських місцях. У цій роботі було успішно реалізовано модель розпізнавання емоцій на базі MobileNetV2, яка використовує дані FER-2013 і здатна працювати на обчислювально ефективних пристроях, таких як Raspberry Pi. Використання методів передавання навчання, ретельна попередня обробка даних і реалізація аналізу відеопотоку в реальному часі продемонстрували можливість впровадження таких систем у реальних умовах.

Рішення відзначається масштабованістю, низькими апаратними вимогами та адаптивністю до різних сценаріїв використання. Наприклад, у промисловості система моніторить емоційний стан операторів складного обладнання, запобігаючи аваріям; у транспорті виявляє ознаки стресу, сонливості чи агресії у водіїв, інтегруючись із системами запобігання ДТП; у медицині забезпечує моніторинг стану пацієнтів у реанімаційних відділеннях; у громадських місцях ідентифікує потенційно небезпечну поведінку для забезпечення безпеки. Така гнучкість досягається завдяки оптимізації продуктивності моделі через методи зміни розміру, нормалізації та збільшення даних, що дозволяє моделі добре узагальнювати свої результати для різноманітних умов і середовищ.

Продовження висновків

Інтеграція можливостей обробки відео в реальному часі є важливою перевагою, яка забезпечує миттєвий зворотний зв'язок. Це дозволяє системі вчасно реагувати на загрози, підтримувати безперервність процесів і підвищувати ефективність у критичних середовищах. Крім того, рішення є економічно доступним завдяки використанню недорогих апаратних засобів, таких як Raspberry Pi, що робить його придатним для впровадження навіть у невеликих організаціях.

Однак етичні та безпекові аспекти залишаються важливими викликами. Для забезпечення довіри та дотримання конфіденційності необхідно впроваджувати заходи анонімізації, шифрування даних, регулярних аудитів та прозорого пояснення мети використання системи. Справедливість моделей залежить від мінімізації упереджень у наборах даних, а також від уважного підходу до їхньої адаптації для різних демографічних груп.

Майбутнє таких систем лежить у поєднанні розпізнавання емоцій із мультимодальними даними, включаючи мовні сигнали, фізіологічні показники та навіть контекстуальні дані. Завдяки цьому можна забезпечити ще більшу точність та розширити спектр застосувань. Успіх цього проекту створює основу для створення ще більш потужних, орієнтованих на користувача систем, які матимуть значний позитивний вплив на суспільство, водночас дотримуючись етичних принципів та гарантій конфіденційності.

Публікації та апробація роботи

1. Balvak, A., Demchenko, M., Drahun, V., Petrenko, V., Prysiashniuk, V., Ivanov, O., Korotkov, S., Tsvyk, O., Tonkykh, O., Antonenko, A., Lemeshko, A., Hafanovych, V., & Holubenko, O. (2023). BUILDING CORPORATE NETWORKS FOR FOOD INDUSTRY ENTERPRISES. European Science, 1(sge22-01), 25–47. <https://doi.org/10.30890/2709-2313.2023-22-01-031>
2. Драгун В.П., Лашевська Н.О. Загальні аспекти застосування штучного інтелекту в медицині. «Сучасні інтелектуальні інформаційні технології в науці та освіті»: IV Всеукр. науково-практ. конф., м.Київ, 15 травня 2024 р. С. 25-26. URL: https://duikt.edu.ua/uploads/p_2661_45318838.pdf
3. Драгун В.П., Лашевська Н.О. Застосування автоматизованого навчання в області медицини. «Сучасні інтелектуальні інформаційні технології в науці та освіті»: IV Всеукр. науково-практ. конф., м.Київ, 15 травня 2024 р. С. 294-295. URL: https://duikt.edu.ua/uploads/p_2661_45318838.pdf
4. Драгун В.П., Вечерковська А.С. Розпізнавання емоцій на основі штучного інтелекту за допомогою трансферного навчання та аналізу в реальному часі. «Проблеми комп'ютерної інженерії» : Науково-практ. конф., м.Київ, 3 грудня 2024 р. С. 89-92
5. Драгун В.П., Вечерковська А.С. Чому важливо навчити штучний інтелект емоційним станам та як це допоможе уникнути «моторошної долини». «Проблеми комп'ютерної інженерії» : Науково-практ. конф., м.Київ, 3 грудня 2024 р. С.93-95