

Хмельницький національний університет  
Факультет інформаційних технологій  
Кафедра комп'ютерної інженерії та інформаційних систем

## КВАЛІФІКАЦІЙНА РОБОТА

Кіберфізична система визначення емоційного стану людини за комп'ютером  
Назва теми

Рівень вищої освіти другий (магістерський)

Галузь знань 12 «Інформаційні технології»  
Шифр, назва

Спеціальність 123 «Комп'ютерна інженерія»  
Шифр, назва

Освітня програма «Комп'ютерна інженерія та програмування»  
Назва

Шифр КвРКІ 240252.21.02.35 ПЗ

Виконав здобувач ІІ курсу, група КІ2м-24-2

  
Підпис

Віталій РУДНИЦЬКИЙ  
Ініціали, прізвище

Керівник канд.-техн. наук, доцент  
Науковий ступінь, учене звання

  
Підпис

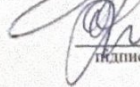
Володимир ГРИГА  
Ініціали, прізвище

Нормоконтролер д. техн. наук, професор  
Науковий ступінь, учене звання

  
Підпис

Сергій ЛИСЕНКО  
Ініціали, прізвище

До захисту допускаю:  
завідувач кафедри КІС  
« 01 » травня 2026 р.  
дата

  
Підпис

Ольга ПАВЛОВА  
Ініціали, прізвище

Хмельницький 2026

# ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Кафедра КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ ТА ІНФОРМАЦІЙНИХ СИСТЕМ

Рівень вищої освіти ДРУГИЙ (МАГІСТЕРСЬКИЙ)

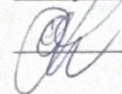
Галузь знань 12 ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

Спеціальність 123 КОМП'ЮТЕРНА ІНЖЕНЕРІЯ

Освітня програма «КОМП'ЮТЕРНА ІНЖЕНЕРІЯ ТА ПРОГРАМУВАННЯ»

ЗАТВЕРДЖУЮ

Завідувачка кафедри КІС



Ольга ПАВЛОВА

“ 12 ” 01 2026 р.

## ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

Рудницькому Віталію Юрійовичу

Прізвище, ім'я, по батькові студента

1. Тема проекту (роботи) Кіберфізична система визначення емоційного стану людини за комп'ютером

Керівник проекту (роботи) Грига Володимир Михайлович, к.т.н., доц.

Прізвище, ім'я, по батькові, науковий ступінь, вчене звання

Затверджена наказом ректора університету від 12.01.2026 р. № 6

2. Термін подання здобувачем роботи на кафедру 01.05.2026 р.

3. Вихідні дані до роботи Завдання на кваліфікаційну роботу

4. Зміст пояснювальної записки (перелік питань, які потрібно розробити) \_\_\_\_\_

Аналіз відомих методів визначення емоційного стану

Моделювання кіберфізичної системи визначення емоційного стану людини за комп'ютером

Метод визначення емоційного стану людини за комп'ютером

Проектування кіберфізичної системи визначення емоційного стану людини

5. Перелік графічного матеріалу (із зазначенням обов'язкових креслень) \_\_\_\_\_


6. Консультанти розділів кваліфікаційної роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв


7. Дата видачі завдання « 12 » 01 2026 р.

**КАЛЕНДАРНИЙ ПЛАН**

№з/п	Назва етапів (розділів) кваліфікаційної роботи магістра	Термін виконання етапів проекту (роботи)	Примітка
1	Вибір напрямку дослідження та узгодження тематики КвРМ з керівником	12.01.2026	виконано
2	Ознайомлення з предметною областю; формулювання мети та задач дослідження; визначення об'єкта та предмета дослідження	12.01.2026	виконано
3	Робота над розділом 1 – аналіз відомих моделей, методів за темою; постановка задачі	20.01.2026	виконано
4	Робота над розділом 2 – розробка моделей для вирішення поставленої задачі	01.02.2026	виконано
5	Робота над науковою статтею	01.03.2026	виконано
6	Робота над розділом 3 – розробка методів для вирішення поставленої задачі	15.03.2026	виконано
7	Робота над розділом 4 – проектування та розробка ПЗ для вирішення поставленої задачі, експериментальна частина	01.04.2026	виконано
8	Оформлення пояснювальної записки згідно вимог	18.04.2026	виконано
9	Попередній захист ДРМ	29.04.2026	виконано
10	Захист ДРМ на засіданні ЕК	До 15.05.2026	

Здобувач   
Підпис

Віталій РУДНІЦЬКИЙ  
Імя, ПРІЗВИЩЕ

Керівник кваліфікаційної роботи   
Підпис

Володимир ГРИГА  
Імя, ПРІЗВИЩЕ

## РЕФЕРАТ

Тема кваліфікаційної роботи магістра: Кіберфізична система визначення емоційного стану людини за комп'ютером

Автор роботи: Віталій РУДНІЦЬКИЙ

Керівник роботи: Володимир ГРИГА

Пояснювальна записка: 94 с., 18 рис., 7 табл., 1 дод., 80 джерел.

КІБЕРФІЗИЧНА СИСТЕМА, ЗБІР ДАНИХ, РОЗПІЗНАВАННЯ ЕМОЦІЙ, МУЛЬТИМОДАЛЬНИЙ АНАЛІЗ, ШТУЧНИЙ ІНТЕЛЕКТ, ГЛИБИННЕ НАВЧАННЯ, ГІБРИДНІ НЕЙРОМЕРЕЖІ.

Об'єктом дослідження є процес розпізнавання та визначення емоційного стану людини під час її взаємодії з комп'ютерними системами у реальному часі.

Предметом дослідження є методи, алгоритми та програмно-апаратні засоби побудови кіберфізичної системи для мультимодального розпізнавання емоцій на основі комплексного аналізу відеопотоку, аудіосигналу та біометричних даних користувача.

Метою кваліфікаційної роботи магістра є підвищення точності розпізнавання емоційного стану користувача шляхом розробки кіберфізичної системи, що базується на аналізі аудіо-візуальних та біометричних сигналів.

Для розв'язання поставлених задач використовувалися методи системного аналізу, методи глибокого навчання, методи мультимодальної інтеграції для об'єднання векторів ознак різної природи у єдину цілісну модель, що дозволяє компенсувати недоліки окремих каналів інформації та підвищити загальну точність класифікації та методи об'єктно-орієнтованого моделювання.

Наукова новизна отриманих результатів:

- ~ удосконалено метод визначення емоційного стану, який, на відміну від існуючих методів, базується на комплексному аналізі відео- та аудіосигналів разом із фізіологічними даними користувача (пульс, частота дихання, електропровідність шкіри);

- ~ набула подальшого розвитку структура кіберфізичної системи, яка за

рахунок інтеграції програмно-апаратних засобів та нейронних мереж забезпечує автоматичну класифікацію емоцій у реальному часі.

На основі проведених досліджень розроблена архітектура і компоненти програмного забезпечення кіберфізичної системи визначення емоційного стану людини за комп'ютером.

Практичне значення дослідження полягає у можливості застосування розробленої кіберфізичної системи в різних галузях. Зокрема, в освітніх технологіях такі системи можуть адаптувати навчальний процес відповідно до рівня концентрації та втоми студента, підвищуючи ефективність засвоєння матеріалу. У сфері розробки програмного забезпечення вони можуть використовуватися для моніторингу стану розробників з метою своєчасного виявлення ознак професійного вигорання. У системах безпеки та управління виробництвом аналіз емоційного стану операторів дозволяє знизити ризик помилок, спричинених стресом або перевтомою.

У першому розділі проведено аналіз предметної області, розглянуто існуючі комерційні рішення та наукові рішення, а також визначено обмеження одноmodalьних підходів.

У другому розділі змодельовано процес розпізнавання емоцій, розроблено структурну модель КФС, що включає рівні сприйняття, обробки та взаємодії; формалізовано математичні моделі відео- та аудіоаналізу.

У третьому розділі запропоновано мультимодальний метод розпізнавання емоцій на основі гібридної нейромережевої архітектури CNN-LSTM та описано механізми забезпечення конфіденційності даних.

У четвертому розділі здійснено програмну реалізацію системи на мові Python та проведено експериментальні дослідження, які підтвердили середню точність розпізнавання на рівні 85,8%.

## ЗМІСТ

Скорочення та умовні позначки.....	5
Вступ.....	6
1 Аналіз відомих методів визначення емоційного стану .....	10
1.1 Поняття та сутність технології розпізнавання емоцій.....	10
1.2 Аналіз існуючих комерційних рішень розпізнавання емоційних станів.....	15
1.3 Аналіз існуючих методів та засобів розпізнавання емоційних станів.....	20
1.4 Постановка задачі.....	24
1.5 Висновки до першого розділу.....	25
2 Моделювання кіберфізичної системи визначення емоційного стану людини за комп'ютером.....	26
2.1 Моделювання процесу визначення емоційного стану людини за комп'ютером.....	26
2.2 Виділення ознак у кіберфізичній системі визначення емоційного стану.....	32
2.3 Класифікація емоційного стану у кіберфізичній системі.....	35
2.4 Моделювання процесу збору та обробки мультисенсорних даних.....	41
2.4.1 Модель відеоаналізу емоційних проявів користувача.....	42
2.4.2 Модель аудіоаналізу мовленнєвих та акустичних характеристик.....	46
2.4.3 Модель аналізу біометричних та фізіологічних показників.....	49
2.5 Висновки.....	52
3 Метод визначення емоційного стану людини за комп'ютером.....	53
3.1 Основи методу визначення емоційного стану.....	53
3.2 Метод визначення емоційного стану людини за комп'ютером.....	55
3.3 Алгоритмічне забезпечення кіберфізичної системи визначення емоційного стану людини за комп'ютером.....	60
3.3.1 Алгоритм збору та синхронізації даних.....	60
3.3.2 Алгоритм попередньої обробки даних.....	63

3.3.3 Алгоритм виділення ознак .....	67
3.4 Застосування нейромережових моделей для виявлення закономірностей у поведінці користувача.....	68
3.5 Методи забезпечення інтерпретованості результатів класифікації емоцій.....	71
3.6 Висновки.....	75
<b>4 ПРОЕКТУВАННЯ КІБЕРФІЗИЧНОЇ СИСТЕМИ ВИЗНАЧЕННЯ ЕМОЦІЙНОГО СТАНУ ЛЮДИНИ.....</b>	
4.1 Вибір типу архітектури та проектування модулів системи.....	77
4.2 Моделювання архітектури кіберфізичної системи засобами UML.....	83
4.2.1 Діаграма варіантів використання програмного забезпечення.....	84
4.2.2 Діаграма класів програмного забезпечення.....	86
4.3 Програмне забезпечення КФС визначення емоційного стану людини за комп'ютером.....	87
4.4 Аналіз результатів експериментальних досліджень та тестування системи.....	92
4.5 Висновки.....	96
Висновки.....	98
Перелік джерел посилань.....	100
Додаток А. Тези доповіді.....	109
Додаток Б. Презентація.....	114

## СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

ЕЕГ - електроенцефалографія

ЕКГ - електрокардіограма

ЕМГ - електроміографія

КФС - кіберфізична система

ПЗ - програмне забезпечення

API - Application Programming Interface (інтерфейс програмування додатків)

AU - Action Unit (основний елемент системи кодування міміки обличчя)

CNN - Convolutional Neural Networks (згортова нейронна мережа)

FACS - Facial Action Coding System (система кодування мімічних дій).

GDPR - General Data Protection Regulation (Загальний регламент про захист даних ЄС)

LSTM - Long Short-Term Memory (довга короткочасна пам'ять)

MFCC - Mel-frequency cepstral coefficients (Мел-частотні кепстральні коефіцієнти)

RNN - Recurrent Neural Networks (рекурентні нейронні мережі)

## ВСТУП

Сучасний етап розвитку інформаційних технологій характеризується зміною підходів до побудови та використання програмно-апаратних систем. Якщо раніше основна увага приділялася автоматизації окремих процесів і підвищенню обчислювальної потужності, то сьогодні пріоритет зміщується у бік створення інтелектуальних кіберфізичних систем, які інтегруються з реальним середовищем і безпосередньо взаємодіють з людиною. Такі системи поєднують фізичні компоненти, сенсорні технології, обробку даних і алгоритми штучного інтелекту, формуючи єдиний адаптивний інформаційний простір.

Водночас, попри значний технологічний прогрес, більшість існуючих людино-машинних інтерфейсів залишаються орієнтованими виключно на раціональну взаємодію, ігноруючи емоційний та психофізіологічний стан користувача.

У традиційних системах взаємодія людини з комп'ютером базується на введенні явних команд через клавіатуру, мишу або сенсорні інтерфейси. Такий підхід не враховує такий аспект людської комунікації як невербальні сигнали, до яких належать міміка, жести, інтонація мовлення та фізіологічні реакції. Саме ці сигнали відіграють ключову роль у передачі емоційного стану та внутрішніх намірів людини. Ігнорування цього аспекту призводить до зниження ефективності взаємодії, особливо в умовах підвищеного когнітивного навантаження, коли користувач може відчувати втому, стрес або перевантаження. Як наслідок, зростає кількість помилок, знижується продуктивність праці та погіршується загальний досвід взаємодії з інформаційною системою.

У цьому контексті особливої актуальності набуває розробка інтелектуальних кіберфізичних систем, здатних розпізнавати емоційний стан користувача в реальному часі та адаптувати свою поведінку відповідно до отриманої інформації. Такі системи дозволяють перейти від пасивних інтерфейсів до проактивних, які не лише реагують на команди, а й враховують внутрішній стан людини. Це відкриває нові можливості для підвищення ефективності роботи,

зниження рівня стресу та створення більш комфортного і безпечного цифрового середовища.

Важливою складовою сучасних систем розпізнавання емоцій є використання методів комп'ютерного зору, обробки мовлення та аналізу біометричних сигналів. Однак застосування лише одного каналу отримання інформації часто є недостатнім для досягнення високої точності. Наприклад, відеоаналіз емоцій може бути ускладнений змінами освітлення, частковим перекриттям обличчя або нестандартними ракурсами. Аналогічно, аудіоаналіз залежить від рівня шуму в навколишньому середовищі та якості запису. Біометричні дані, у свою чергу, можуть вимагати спеціалізованого обладнання або бути чутливими до індивідуальних особливостей користувача.

У зв'язку з цим особливого значення набуває мультимодальний підхід до розпізнавання емоцій, який передбачає інтеграцію даних з різних джерел, зокрема відео, аудіо та фізіологічних сенсорів. Поєднання кількох каналів дозволяє компенсувати недоліки кожного з них окремо та забезпечити більш надійне і стабільне функціонування системи в різних умовах експлуатації. Мультимодальні системи здатні враховувати комплексні прояви емоційного стану, що значно підвищує точність класифікації та розширює можливості адаптивної взаємодії.

Актуальність роботи зумовлена необхідністю створення людино-орієнтованих кіберфізичних систем, які здатні ефективно взаємодіяти з користувачем, враховуючи його емоційний стан. У сучасному цифровому суспільстві, де значна частина діяльності відбувається у віртуальному середовищі, питання психоемоційного комфорту стає не менш важливим, ніж технічні характеристики систем. Розробка інтелектуальних рішень у цій сфері дозволяє підвищити якість життя, забезпечити безпеку праці та оптимізувати робочі процеси.

Метою кваліфікаційної роботи магістра є підвищення точності розпізнавання емоційного стану користувача шляхом розробки кіберфізичної системи, що базується на аналізі аудіо-візуальних та біометричних сигналів.

Поставлена мета досягається розв'язанням таких основних завдань:

- ~ аналізу сучасних методів та комерційних рішень у сфері розпізнавання емоцій із виділенням їхніх обмежень;
- ~ моделювання процесу визначення емоційного стану людини в межах кіберфізичної системи;
- ~ розроблення методу класифікації емоцій на основі гібридної нейромережевої моделі, здатної враховувати часову динаміку стану користувача;
- ~ проєктування архітектури кіберфізичної системи, що забезпечує синхронний збір даних з камери, мікрофона та біометричних сенсорів;
- ~ проведення експериментального дослідження розробленої системи та оцінювання її ефективності у порівнянні з існуючими одноmodalними підходами.

Об'єктом дослідження є процес розпізнавання та визначення емоційного стану людини під час її взаємодії з комп'ютерними системами у реальному часі.

Предметом дослідження є методи, алгоритми та програмно-апаратні засоби побудови кіберфізичної системи для мультимодального розпізнавання емоцій на основі комплексного аналізу відеопотоку, аудіосигналу та біометричних даних користувача.

Наукова новизна отриманих результатів:

- ~ удосконалено метод визначення емоційного стану, який, на відміну від існуючих методів, базується на комплексному аналізі відео- та аудіосигналів разом із фізіологічними даними користувача (пульс, частота дихання, електропровідність шкіри);
- ~ набула подальшого розвитку структура кіберфізичної системи, яка за рахунок інтеграції програмно-апаратних засобів та нейронних мереж забезпечує автоматичну класифікацію емоцій у реальному часі.

Практичне значення дослідження полягає у можливості застосування розробленої кіберфізичної системи в різних галузях. Зокрема, в освітніх технологіях такі системи можуть адаптувати навчальний процес відповідно до рівня концентрації та втоми студента, підвищуючи ефективність засвоєння матеріалу. У сфері розробки програмного забезпечення вони можуть

використовуватися для моніторингу стану розробників з метою своєчасного виявлення ознак професійного вигорання. У системах безпеки та управління виробництвом аналіз емоційного стану операторів дозволяє знизити ризик помилок, спричинених стресом або перевтомою.

Отже, розробка кіберфізичної системи розпізнавання емоцій користувача на основі мультимодального підходу є важливим і актуальним науково-прикладним завданням. Її вирішення сприятиме підвищенню ефективності людино-машинної взаємодії, розширенню функціональних можливостей сучасних інформаційних систем та створенню комфортного цифрового середовища, що враховує індивідуальні особливості та потреби користувача.

Для розв'язання поставлених задач використовувалися методи системного аналізу, методи глибокого навчання, методи мультимодальної інтеграції для об'єднання векторів ознак різної природи у єдину цілісну модель, що дозволяє компенсувати недоліки окремих каналів інформації та підвищити загальну точність класифікації та методи об'єктно-орієнтованого моделювання.

За темою кваліфікаційної роботи опубліковано одну публікацію [80] у Збірнику наукових праць за матеріалами XIX Всеукраїнської науково-практичної WEB конференції аспірантів, студентів та молодих вчених «Комп'ютерні інтелектуальні системи та мережі» (25-27 березня 2026 р.). – Кривий Ріг: Криворізький національний університет, 2026. – 370 с.

# 1 АНАЛІЗ ВІДОМИХ МЕТОДІВ ВИЗНАЧЕННЯ ЕМОЦІЙНОГО СТАНУ

## 1.1 Поняття та сутність технології розпізнавання емоцій

Розпізнавання емоцій за комп'ютером – це галузь штучного інтелекту, що займається розробкою систем для ідентифікації, інтерпретації та обробки людських афектів (емоцій, настроїв, станів) [1].

У сучасних системах розпізнавання емоційного стану людини ключову роль відіграє інтеграція трьох основних типів даних, а саме: візуальних, аудіальних та біометричних, кожен із яких відображає різні аспекти психофізіологічної реакції користувача. Комплексне використання цих модальностей дозволяє підвищити точність і надійність класифікації, однак кожна з них має власні переваги, недоліки та обмеження, що необхідно враховувати при проектуванні кіберфізичних систем.

Візуальні дані є найбільш інтуїтивно зрозумілим джерелом інформації про емоційний стан, оскільки міміка обличчя, рухи очей та положення голови безпосередньо відображають емоційні реакції людини. Аналіз мікроміміки дозволяє виявляти короточасні та часто неконтрольовані вирази обличчя, які можуть свідчити про справжні емоції навіть у разі їх свідомого приховування. Додатково рухи очей і напрямок погляду дають змогу оцінити рівень уваги, зацікавленості або когнітивного навантаження, тоді як нахил голови може сигналізувати про емоційне ставлення або ступінь залученості у взаємодію. Основною перевагою візуального аналізу є висока інформативність і можливість безконтактного збору даних у реальному часі.

Проте цей підхід має і суттєві обмеження: він сильно залежить від умов освітлення, якості відео, положення камери та наявності перешкод (наприклад, масок або окулярів). Крім того, мімічні реакції можуть варіюватися залежно від культурних особливостей або бути свідомо контрольованими, що знижує достовірність інтерпретації.

Аудіальні характеристики мовлення виступають важливим доповненням до візуальних даних, оскільки голос людини містить значну кількість емоційної інформації. Тембр, швидкість мовлення, тривалість пауз та інтонаційні зміни відображають внутрішній стан користувача і часто є менш контрольованими, ніж міміка. Наприклад, підвищення темпу мовлення та зменшення пауз можуть свідчити про збудження або нервозність, тоді як уповільнення мовлення часто асоціюється зі станами втоми або пригніченості. Інтонаційні характеристики дозволяють виявляти емоційні акценти навіть у нейтральному текстовому змісті. Перевагою аудіального аналізу є його відносна незалежність від візуальних умов та здатність працювати навіть при обмеженій видимості обличчя. Водночас цей підхід має низку обмежень, пов'язаних із впливом зовнішніх шумів, які можуть суттєво спотворювати сигнал, а також із мовною залежністю, оскільки інтонаційні патерни можуть відрізнятися між мовами та культурами. Додатково варто враховувати, що якість запису та характеристики мікрофона можуть впливати на точність виділення ознак.

Біометричні дані є найбільш об'єктивним джерелом інформації про емоційний стан, оскільки вони відображають фізіологічні процеси, що практично не піддаються свідомому контролю. Частота серцевих скорочень, гальванічна реакція шкіри та інші показники автономної нервової системи безпосередньо пов'язані з рівнем емоційного збудження, стресу або релаксації. Наприклад, підвищення частоти серцевих скорочень і зростання електропровідності шкіри є типовими індикаторами стресової реакції. Основною перевагою біометричного підходу є його висока точність і надійність навіть у складних умовах, де візуальні та аудіальні сигнали можуть бути спотворені. Однак використання біометричних даних пов'язане з рядом обмежень, серед яких необхідність використання спеціалізованих сенсорів, що може ускладнювати інтеграцію системи та знижувати її зручність для користувача. Крім того, такі дані можуть бути чутливими з точки зору конфіденційності, що вимагає дотримання відповідних стандартів захисту інформації.

Узагальнюючи, можна зазначити, що кожен тип даних має власну область

ефективного застосування, а їх комбінування дозволяє компенсувати недоліки окремих підходів. Візуальні методи забезпечують детальний аналіз зовнішніх проявів емоцій, аудіальні розширюють можливості інтерпретації за рахунок голосових характеристик, а біометричні надають об'єктивну інформацію про внутрішній стан організму.

Саме інтеграція цих трьох модальностей є основою сучасних кіберфізичних систем розпізнавання емоцій, що забезпечує підвищену точність, адаптивність та стійкість до зовнішніх впливів.

Основна мета – це перехід від взаємодії типу команда-виконання до адаптивного інтерфейсу, який розуміє, що користувач втомився, роздратований або заплутався, і змінює свою поведінку відповідно до цього.

Аналіз сфер застосування технологій розпізнавання емоцій свідчить про їх стрімке поширення в різних галузях, де важливим є врахування емоційного стану людини для підвищення ефективності взаємодії та прийняття рішень. У кожній із цих сфер використання емоційно-орієнтованих систем має свої особливості, переваги та обмеження, що визначають специфіку їх впровадження (табл. 1.1).

Таблиця 1.1 – Сфери використання технології розпізнавання емоцій

Сфера	Як це працює	Результат
Освіта (EdTech)	Система бачить, коли студент втрачає концентрацію або не розуміє матеріал.	Адаптація складності контенту, виклик викладача або зміна темпу лекції.
Маркетинг	Аналіз реакції фокус-груп на рекламні ролики в реальному часі.	Оптимізація бюджету: залишають тільки ті кадри, що викликають «правильну» емоцію.
HR та розробка ПЗ	Моніторинг вигорання та рівня стресу розробників під час написання коду.	Покращення корпоративної культури, запобігання помилкам через втому.

Кінець таблиці 1.1

Сфера	Як це працює	Результат
Медицина	Допомога людям з розладами аутичного спектру (РАС) розуміти емоції співрозмовника.	Соціальна реабілітація та підтримка ментального здоров'я.
Геймінг	Зміна ігрового сюжету або складності залежно від страху чи азарту гравця.	Максимальне занурення (Immersive experience).

У сфері освіти технології розпізнавання емоцій застосовуються для створення інтелектуальних навчальних систем, здатних адаптувати навчальний процес до індивідуальних особливостей студента. Аналіз емоційного стану дозволяє визначати рівень зацікавленості, втоми або фрустрації під час навчання, що дає змогу автоматично змінювати складність матеріалу, темп подачі інформації або тип завдань. Такий підхід сприяє підвищенню ефективності засвоєння знань та зменшенню когнітивного перевантаження. Водночас впровадження розпізнавання емоцій в освіті потребує врахування етичних аспектів, зокрема конфіденційності даних та добровільності участі користувачів.

У медичній сфері технології аналізу емоцій використовуються для моніторингу психоемоційного стану пацієнтів, зокрема при діагностиці депресії, тривожних розладів або посттравматичного стресу. Системи можуть аналізувати міміку, мовлення та фізіологічні показники для виявлення змін у стані пацієнта в режимі реального часу. Це дозволяє лікарям отримувати додаткову об'єктивну інформацію та своєчасно реагувати на погіршення стану. Крім того, такі технології використовуються у телемедицині, де прямий контакт із пацієнтом обмежений.

Однак застосування розпізнавання емоцій в медицині вимагає високого рівня точності, надійності та відповідності нормативним вимогам щодо обробки

персональних даних.

У сфері автомобільної промисловості розпізнавання емоцій інтегрується у системи моніторингу водія, що дозволяє підвищити безпеку дорожнього руху. Аналіз стану водія, зокрема втоми, неувважності або емоційного напруження, дає змогу своєчасно попереджати про небезпечні ситуації або навіть автоматично втручатися в процес керування транспортним засобом. Такі системи є важливим компонентом сучасних кіберфізичних транспортних систем і активно використовуються у концепції автономного водіння. Основним викликом у цій сфері є забезпечення стабільної роботи системи в умовах змінного освітлення, руху та обмеженого часу на прийняття рішень.

У маркетингу та бізнес-аналітиці розпізнавання емоцій використовується для дослідження поведінки споживачів і оцінки їх реакції на продукти, рекламу або інтерфейси. Аналіз емоцій дозволяє компаніям отримувати глибше розуміння споживчих уподобань і підвищувати ефективність маркетингових стратегій. Наприклад, реакція користувача на рекламний контент може бути оцінена за допомогою аналізу міміки або голосу, що дає більш об'єктивні результати порівняно з традиційними опитуваннями. Водночас важливим є дотримання етичних норм і прозорість використання таких технологій, оскільки вони можуть впливати на поведінку споживачів.

У сфері розваг та ігрової індустрії розпізнавання емоцій відкриває нові можливості для створення адаптивного контенту, який реагує на емоційний стан користувача. Ігрові системи можуть змінювати складність, сюжет або поведінку персонажів залежно від реакцій гравця, що підвищує рівень занурення та індивідуалізації досвіду. Подібні підходи також застосовуються у віртуальній та доповненій реальності, де емоційна взаємодія є ключовим елементом користувацького досвіду. Обмеженням у цій сфері може бути необхідність у додатковому обладнанні та складність інтеграції різних типів даних у реальному часі.

Таким чином, технології розпізнавання емоцій мають широкий спектр застосувань і демонструють значний потенціал для підвищення ефективності

взаємодії між людиною та технічними системами. Їх використання дозволяє створювати адаптивні, інтелектуальні середовища, здатні враховувати емоційний стан користувача. Водночас успішне впровадження таких систем потребує вирішення технічних, етичних та організаційних питань, що залишаються актуальними напрямками подальших досліджень.

## 1.2 Аналіз існуючих комерційних рішень розпізнавання емоційних станів

Сучасний ринок комерційних рішень у сфері розпізнавання емоцій демонструє стрімке зростання та високий рівень інноваційної активності, що обумовлено розвитком штучного інтелекту, сенсорних технологій та потребою у більш персоналізованій взаємодії людини з цифровими системами. За оцінками аналітичних звітів, глобальний ринок афективних обчислень у 2023 році становив понад 62 млрд доларів і прогнозується зростання до майже 388 млрд доларів до 2030 року із середньорічним темпом понад 30% [2]. Така динаміка свідчить про перехід розпізнавання емоцій із дослідницької сфери у масштабні комерційні застосування.

Комерційний ринок розпізнавання емоцій можна умовно поділити на три основні сегменти: апаратне забезпечення, програмні платформи та інтегровані рішення. Найбільшу частку наразі займає апаратний сегмент, який включає камери, сенсори, мікрофони та носимі пристрої, що забезпечують збір даних для аналізу емоцій. Водночас програмні рішення демонструють найшвидші темпи зростання завдяки розвитку хмарних технологій, API-сервісів і алгоритмів глибинного навчання.

Важливою тенденцією є перехід від одноканальних (наприклад, лише відео) до мультимодальних систем, які поєднують відео, аудіо та біометричні сигнали. Саме такі рішення забезпечують найвищу точність і поступово стають стандартом у комерційних продуктах.

Серед провідних компаній ринку варто виділити Affectiva, IBM, Microsoft, Apple та Google. Ці компанії формують основні технологічні тренди та

визначають напрям розвитку галузі.

Компанія Affectiva є одним із піонерів технологій розпізнавання емоцій, яка розробила технології аналізу емоцій на основі мимики, голосу та поведінкових патернів [3]. Її рішення активно застосовуються в автомобільній галузі, маркетингових дослідженнях та системах взаємодії з користувачем. Технології компанії базуються на глибинному навчанні та великих наборах даних, зібраних у реальних умовах.

Компанії IBM та Microsoft розвивають технології розпізнавання емоцій у напрямку хмарних платформ та корпоративних рішень. Зокрема, сервіси на основі API дозволяють інтегрувати аналіз емоцій у бізнес-процеси, такі як обслуговування клієнтів, аналітика поведінки користувачів та автоматизація комунікацій [4]. Ще одним важливим рішенням є Microsoft Azure Cognitive Services, яке надає доступ до алгоритмів розпізнавання емоцій через хмарний API [5]. Такий підхід забезпечує високу масштабованість і простоту інтеграції, що робить його привабливим для комерційних застосувань. Водночас використання хмарних сервісів може бути обмежене вимогами до конфіденційності даних та затримками при передачі інформації.

У свою чергу, Apple та Google активно впроваджують елементи розпізнавання емоцій у споживчі пристрої такі, як смартфони, носимі гаджети та голосові асистенти [6].

Комерційні рішення розпізнавання емоцій активно впроваджуються у декількох ключових галузях.

Найбільшу частку займає сфера охорони здоров'я, де емоційний аналіз використовується для моніторингу психічного стану, діагностики та підтримки пацієнтів.

У автомобільній галузі ці технології застосовуються для підвищення безпеки через аналіз стану водія, зокрема виявлення втоми або неувважності.

У сфері маркетингу та реклами технології розпізнавання емоцій дозволяють оцінювати реакцію споживачів на контент у режимі реального часу, що значно підвищує ефективність рекламних кампаній.

Також активно розвивається напрям використання у споживчій електроніці, де емоційний аналіз забезпечує персоналізацію інтерфейсів і підвищує якість взаємодії користувача з пристроєм.

Порівняння основних існуючих комерційних рішень в сфері технологій розпізнавання емоцій представлено в таблиці 1.2.

Таблиця 1.2 – Порівняння комерційних рішень

Компанія / Продукт	Ключова технологія	Переваги	Недоліки
Affectiva (Smart Eye)	Affdex SDK (відео-аналіз обличчя)	Найбільший датасет у світі (мільйони облич). Висока точність у розпізнаванні нюансів.	Висока вартість ліцензії, орієнтація на великий бізнес.
Microsoft Azure	Face API (хмарне рішення)	Легка інтеграція в існуючі хмарні сервіси. Не потребує потужного локального GPU.	Залежність від інтернету, затримки (latency), питання приватності даних.
MorphCast	HTML5 Engine (browser-based)	Працює прямо в браузері без встановлення ПЗ. Дуже легкий код.	Менша точність порівняно з нативними SDK через обмеження JS.
Kairos	Human Analytics	Гнучкість у налаштуванні під конкретні бізнес-задачі, хороша документація.	Обмежена підтримка аудіо-каналу (фокус переважно на відео).
Realeyes	Pre-testing platform	Спеціалізація на маркетингу та увазі (attention tracking).	Вузькоспеціалізоване рішення, не підходить для загальних КПС.

Однією з ключових тенденцій є інтеграція технології розпізнавання емоцій з технологіями Internet of Things (IoT) та edge computing. Це дозволяє виконувати обробку даних безпосередньо на пристрої, зменшуючи затримки та підвищуючи конфіденційність. Крім того, активно розвиваються нейронні процесори та спеціалізовані AI-чіпи, що забезпечують обробку емоційних даних у реальному часі.

Ще одним важливим напрямом є розвиток мультимодальних моделей, які об'єднують різні джерела даних. Такі моделі дозволяють компенсувати недоліки окремих каналів і досягати більш високої точності розпізнавання. Інтеграція глибинного навчання, зокрема CNN та LSTM, дозволяє враховувати як просторові, так і часові характеристики емоцій.

Основною перевагою комерційних систем є їх висока точність та готовність до використання у реальних умовах. Вони базуються на великих обсягах даних і оптимізованих алгоритмах, що забезпечує стабільну роботу навіть у складних середовищах. Крім того, такі рішення зазвичай мають зручні інтерфейси та API, що спрощує їх інтеграцію у різні інформаційні системи. На відміну від академічних рішень, комерційні платформи орієнтовані на практичне застосування та інтеграцію у бізнес-процеси. Комерційні рішення зазвичай демонструють високу точність у реальних умовах завдяки використанню великих обсягів навчальних даних та оптимізованих алгоритмів, однак вони є закритими і не дозволяють повною мірою контролювати внутрішню логіку роботи системи.

Іншою важливою перевагою є масштабованість, особливо у випадку хмарних платформ, які дозволяють обробляти великі обсяги даних і підтримувати одночасну роботу великої кількості користувачів.

Попри значні переваги, комерційні рішення мають і низку обмежень. Одним із головних є закритість алгоритмів, що ускладнює їх адаптацію та інтерпретацію результатів. Це особливо важливо у наукових дослідженнях і критичних застосуваннях, де потрібна прозорість моделей.

Ще одним суттєвим обмеженням є питання конфіденційності та захисту даних. Оскільки розпізнавання емоцій працює з чутливою інформацією, такою як

міміка, голос або біометричні показники, необхідно забезпечувати високий рівень безпеки та відповідність нормативним вимогам.

Крім того, точність систем може знижуватися у випадках культурних відмінностей, індивідуальних особливостей користувачів або нестандартних умов середовища. Це вимагає додаткового навчання моделей та адаптації до конкретних сценаріїв використання.

Отже, ринок комерційних рішень Emotion AI перебуває на етапі активного зростання та трансформації. Основними драйверами розвитку є попит на персоналізовані цифрові сервіси, підвищення ролі емоцій у взаємодії людини з технологіями та розвиток штучного інтелекту. Водночас ключовими викликами залишаються забезпечення точності, інтерпретованості та конфіденційності даних.

Перспективним напрямом є створення гібридних кіберфізичних систем, які поєднують можливості комерційних платформ із гнучкістю власних моделей, що дозволяє досягти балансу між ефективністю, адаптивністю та контролем над системою.

Розглянута технологія розпізнавання емоцій має низку суттєвих переваг і водночас супроводжується важливими обмеженнями та викликами. Однією з ключових переваг є об'єктивність аналізу: на відміну від людини, програмна система не піддається втомі чи емоційному виснаженню, а також здатна виявляти мікроекспресії обличчя, які часто залишаються непоміченими під час ручного спостереження. Це дозволяє підвищити точність інтерпретації емоційного стану користувача.

Ще однією важливою перевагою є автоматизація процесу аналізу. Система здатна одночасно обробляти великі масиви даних, що особливо корисно у випадках масових онлайн-заходів, таких як вебінари, дистанційне навчання або маркетингові дослідження, де необхідно оцінити емоційну реакцію великої кількості користувачів у реальному часі. Додатковою перевагою є безконтактність технології, оскільки для її роботи достатньо звичайної веб-камери, без необхідності використання спеціальних датчиків або фізичних пристроїв, що

робить систему зручною та легко масштабованою.

Водночас технологія має і низку суттєвих недоліків та обмежень. Однією з основних проблем є культурні відмінності у вираженні емоцій: одна й та сама міміка, наприклад усмішка, може мати різне значення залежно від культурного контексту, що призводить до так званого культурного зсуву або упередженості моделей. Також значний вплив на якість розпізнавання мають умови зйомки, зокрема освітлення та ракурс обличчя користувача: при слабкому освітленні або неправильному положенні обличчя точність системи суттєво знижується.

Окремо слід виділити етичні аспекти та питання приватності, оскільки автоматичне розпізнавання емоцій пов'язане зі збором чутливої біометричної інформації. Це породжує дискусії щодо законності та допустимості такого аналізу без явної згоди користувача, особливо в контексті регуляцій, таких як GDPR у Європейському Союзі [7]. Крім того, існує проблема можливого навмисного спотворення результатів, коли користувач може свідомо імітувати або приховувати емоції, вводячи систему в оману та знижуючи достовірність отриманих даних.

### 1.3 Аналіз існуючих методів та засобів розпізнавання емоційних станів

Сучасний етап розвитку систем розпізнавання емоційних станів, що належить до міждисциплінарної галузі, характеризується переходом від вузькоспеціалізованих підходів до комплексних багатомодальних кіберфізичних систем. Якщо ранні дослідження зосереджувалися переважно на аналізі статичних зображень обличчя, то сьогодні домінують підходи, що враховують часову динаміку, контекст взаємодії та інтеграцію різнорідних джерел даних, зокрема відео, аудіо та біометричних сигналів. Така еволюція обумовлена необхідністю підвищення точності розпізнавання емоцій у реальних умовах, де присутні шуми, варіації освітлення, індивідуальні особливості користувачів та інші фактори невизначеності.

Базою для більшості сучасних методів розпізнавання емоцій є теоретичні

напрацювання Пола Екмана, який разом із Воллесом Фрізенем розробив систему FACS (Facial Action Coding System) [8]. Ця система дозволяє формалізувати мимічні прояви людини шляхом розкладання виразу обличчя на елементарні компоненти, так звані Action Units (AU), кожна з яких відповідає скороченню певної групи м'язів.

FACS фактично виступає універсальною мовою опису миміки, що не залежить від культурних чи індивідуальних особливостей. Завдяки цьому вона стала основою для побудови алгоритмів комп'ютерного зору, які навчаються розпізнавати емоції через комбінації активованих AU. У сучасних нейромережових підходах, зокрема в згорткових нейронних мережах, FACS можна розглядати як своєрідний алфавіт, на якому базується навчання моделей.

Важливим аспектом є те, що FACS дозволяє працювати не лише з базовими емоціями, але й із мікровиразами, тобто короткочасними мимічними реакціями, які часто є більш інформативними, ніж статичні вирази. Це особливо актуально для кіберфізичних систем, де важлива здатність реагувати на швидкі зміни стану користувача.

Аналіз сучасного стану досліджень у галузі кіберфізичних систем свідчить про стрімку інтеграцію методів обчислювального інтелекту для аналізу великих масивів даних. Як зазначають автори роботи [9], фундаментальним трендом є перехід від статичного збору інформації до динамічної аналітики, що базується на Big Data. Автори підкреслюють, що саме поєднання когнітивних обчислень та методів інтелектуального аналізу дозволяє кіберфізичній системі ефективно функціонувати в умовах невизначеності, що є критично важливим для систем, які взаємодіють із людиною-оператором. Цей загальний контекст створює підґрунтя для розробки більш спеціалізованих рішень, спрямованих на розуміння внутрішнього стану користувача.

Важливим вектором розвитку є концепція емоційного інтелекту самої системи, як засобу підвищення її стійкості. У роботах авторів [10, 11] запропоновано унікальний підхід до забезпечення живучості кіберфізичної системи через механізми штучних емоцій у мультиагентних середовищах. Автори

моделюють емоційні процеси як специфічні стани агентів, що дозволяє системі імітувати біологічні реакції на зовнішні загрози. Наприклад, стан страху в агента може виступати як сигнал про критичну вразливість, що змушує систему активувати захисні протоколи швидше, ніж це зробили б традиційні алгоритми. Таке використання емоційного підходу дозволяє створювати самовідновлювані системи, де емоції слугують гнучким інструментом зворотного зв'язку між цифровим та фізичним рівнями.

Питання безпеки в таких системах також отримує новий розвиток через когнітивно-інспіровані моделі. Так, в роботі [12] автори представляють фреймворк SecureCPS, який базується на принципах людського сприйняття для виявлення кібератак. Замість звичайного аналізу сигнатур, система намагається усвідомити аномальну поведінку трафіку через призму когнітивних шаблонів. Це підтверджує тезу про те, що методи, запозичені з психофізіології та теорії пізнання, є ефективними не лише для розпізнавання станів людини, а й для захисту цілісності всієї кіберфізичної інфраструктури від складних цілеспрямованих атак.

Безпосередня взаємодія людини з комп'ютером у межах кіберфізичної системи сьогодні розглядається крізь призму концепції «людського цифрового двійника». Так, автори роботи [13] детально описують застосування технологій віртуальної реальності для створення афективних моделей користувача. У їхньому дослідженні емоційне розпізнавання є центральним елементом, який дозволяє синхронізувати фізичний стан людини з її цифровою копією. Це забезпечує можливість предиктивної адаптації інтерфейсів, коли система заздалегідь відчуває зміну емоційного фону та коригує робоче середовище для мінімізації стресу або підвищення продуктивності.

Автори дослідження [14] фокусуються на практичному застосуванні глибинного навчання для розпізнавання академічної діяльності. Автори доводять, що автоматизований аналіз поведінки користувача за комп'ютером (читання, написання тексту, пошук інформації) дозволяє системі об'єктивно оцінювати рівень залученості, що є непрямим, але важливим показником емоційного та

когнітивного стану.

Найбільш прогресивним напрямком є впровадження принципів пояснюваного штучного інтелекту у мультимодальні системи розпізнавання.

В роботі [15] пропонується підхід, де класифікація емоцій базується на поєднанні декількох каналів даних, але при цьому алгоритм залишається прозорим для інтерпретації. Це дозволяє не лише визначити емоцію (наприклад, гнів чи втому), а й зрозуміти, які саме фактори, міміка, тембр голосу чи специфічні рухи стали вирішальними для прийняття рішення. Такий рівень прозорості є надзвичайно актуальним для розумних середовищ, де помилкове розпізнавання емоції може призвести до невірної адаптації системи.

Таким чином, сучасні дослідження демонструють перехід від простих методів детекції до комплексних, мультимодальних та інтерпретованих моделей, що враховують як динаміку людських почуттів, так і потреби системи в безпеці та стійкості.

Також широкого поширення набули відкриті інструменти, які дозволяють реалізовувати системи аналізу емоцій без значних фінансових витрат. Одним із найбільш відомих є OpenFace 2.0 [17]. Цей комплексний інструментарій для аналізу поведінки обличчя, який забезпечує детекцію ключових точок, оцінку напрямку погляду, а також визначення Action Units у реальному часі.

OpenFace базується на сучасних методах машинного навчання та забезпечує високу точність у контрольованих умовах. Його перевагою є можливість глибокої модифікації та інтеграції в дослідницькі проекти. Водночас він має обмеження, пов'язані з чутливістю до якості відеопотоку та освітлення.

Іншим важливим інструментом є DeepFace, який представляє собою Python-бібліотеку, що об'єднує декілька потужних моделей глибокого навчання для аналізу обличчя [18].

Основною перевагою DeepFace є використання попередньо навчених моделей, що дозволяє досягати високої точності навіть без значних обсягів власних навчальних даних.

Ці інструменти активно використовуються у прототипах кіберфізичних

систем, однак їх застосування в реальних умовах часто потребує додаткової адаптації та оптимізації.

Отже, проведений аналіз показує, що сучасні системи розпізнавання емоцій розвиваються у напрямку підвищення точності, адаптивності та інтегрованості. Від простих моделей, що базуються на статичних зображеннях, дослідження перейшли до складних багатомодальних систем, які враховують як просторові, так і часові характеристики даних.

Водночас жодне з існуючих рішень не є універсальним. Відкриті інструменти забезпечують гнучкість, але потребують значних зусиль для адаптації, тоді як комерційні платформи є більш зручними у використанні, проте обмежують контроль над алгоритмами. Спеціалізовані кіберфізичні системи демонструють високу ефективність у конкретних задачах, але не завжди можуть бути перенесені в інші області.

У цьому контексті перспективним є підхід, що базується на поєднанні відео- та аудіоаналізу з використанням моделей глибокого навчання, зокрема LSTM, які дозволяють враховувати часову динаміку емоцій. Такий підхід забезпечує більш повне розуміння стану користувача та відкриває можливості для створення адаптивних кіберфізичних систем.

#### 1.4 Постановка задачі

Отже, проведений аналіз показав, що розпізнавання емоцій є ключовим компонентом людино-орієнтованих кіберфізичних систем. Для створення кіберфізичної системи визначення емоційного стану людини за комп'ютером необхідно вирішити наступні задачі.

1. Проаналізувати сучасні методи та комерційні рішення у сфері розпізнавання емоцій, виділивши їхні обмеження.
2. Змоделювати процесу визначення емоційного стану людини в межах кіберфізичної системи.
3. Запропонувати метод класифікації емоцій на основі гібридної

нейромережевої моделі, здатної враховувати часову динаміку стану користувача.

4. Розробити архітектуру кіберфізичної системи, яка забезпечує синхронний збір даних з камери, мікрофона та біометричних сенсорів.

5. Провести експериментальне дослідження розробленої системи та оцінити її ефективність у порівнянні з існуючими одноmodalьними підходами.

### 1.5 Висновки до першого розділу

Аналіз предметної області показав, що розпізнавання емоцій є ключовим компонентом людино-орієнтованих кіберфізичних систем, що дозволяє створювати адаптивні інтелектуальні середовища. Досліджено три основні модальності даних: візуальні (міміка), аудіальні (інтонація) та біометричні (фізіологічні реакції). Встановлено, що інтеграція цих каналів дозволяє компенсувати недоліки окремих методів, наприклад, залежність відео від освітлення або чутливість аудіо до шумів.

Огляд існуючих рішень виявив обмеження у вигляді закритості алгоритмів, високої вартості ліцензій та залежності від хмарних сервісів.

Такі технології як OpenFace та DeepFace забезпечують гнучкість у розробці, але потребують додаткової адаптації для роботи в реальних умовах кіберфізичних систем з обмеженими обчислювальними ресурсами.

Отже, перспективним напрямом є використання гібридних моделей на основі згорткових нейронних мереж для аналізу просторових ознак обличчя та рекурентних мереж для врахування часової інерційності емоцій. Це дозволяє перейти від аналізу статичних кадрів до розуміння емоційного процесу в динаміці.

## **2 МОДЕЛЮВАННЯ КІБЕРФІЗИЧНОЇ СИСТЕМИ ВИЗНАЧЕННЯ ЕМОЦІЙНОГО СТАНУ ЛЮДИНИ ЗА КОМП'ЮТЕРОМ**

### **2.1 Моделювання процесу визначення емоційного стану людини за комп'ютером**

У межах даного дослідження процес визначення емоційного стану людини розглядається як складна багатокomпонентна задача, що поєднує обробку сигналів, аналіз поведінкових характеристик та застосування методів штучного інтелекту.

Метою моделювання є формалізація процесу збору, обробки та інтерпретації даних про користувача з метою класифікації його емоційного стану в режимі, наближеному до реального часу.

Задача моделювання полягає у:

- ~ визначенні структури кіберфізичної системи;
- ~ формалізації вхідних та вихідних параметрів;
- ~ побудові математичної моделі процесу;
- ~ виборі методів аналізу та класифікації емоцій.

Для побудови ефективної кіберфізичної системи необхідно забезпечити інтеграцію програмно-апаратних засобів збору даних та методів їх інтелектуальної обробки. Концептуальна модель системи базується на поєднанні фізичного рівня (сенсори та пристрої) та кібернетичного рівня (алгоритми та нейронні мережі).

Структурна схема системи зображена на рисунку 2.1 та включає наступні рівні.

1. Рівень сприйняття (фізичний). Збір первинних даних через відеокамери, мікрофони та біометричні датчики.
2. Рівень обробки (кібернетичний). Виділення ключових ознак та класифікація емоцій.
3. Рівень взаємодії. Інтерфейс для відображення результатів та інтеграції з іншими системами.



Рисунок 2.1 – Структурна схема кіберфізичної системи визначення емоційного стану людини за комп'ютером

Рівень сприйняття є базовим компонентом кіберфізичної системи визначення емоційного стану людини та відповідає за формування первинного інформаційного потоку. На цьому рівні здійснюється безпосередня взаємодія системи з фізичним середовищем, а саме, з користувачем, його поведінкою та фізіологічними проявами емоцій. Основною функцією рівня є реєстрація сирих (необроблених) даних за допомогою сенсорних пристроїв, які забезпечують багатомодальне спостереження.

До складу цього рівня входять відеокамери, мікрофони та, за необхідності, біометричні сенсори. Відеокамери забезпечують захоплення зображення обличчя користувача у режимі реального часу, що дозволяє аналізувати міміку, мікровирази, положення голови та напрямок погляду.

Мікрофони використовуються для запису мовлення та фонових звуків, що дає змогу виділяти акустичні характеристики голосу, такі як тембр, інтонація, темп мовлення та емоційне забарвлення.

Біометричні датчики, у разі їх використання, можуть фіксувати додаткові фізіологічні параметри, зокрема частоту серцевих скорочень, електрошкірну активність або температуру тіла, що підвищує точність оцінювання емоційного стану.

Особливістю функціонування цього рівня є необхідність забезпечення високої точності та синхронності збору даних, оскільки різні типи сигналів мають різну природу та частоту дискретизації. Крім того, важливим аспектом є попередня фільтрація шумів та артефактів, які можуть виникати через умови навколишнього середовища (освітлення, фонові шуми, якість обладнання). Таким чином, рівень сприйняття формує багатовимірний потік даних, який слугує основою для подальшої обробки в кібернетичному середовищі системи.

Рівень обробки є центральним елементом кіберфізичної системи, на якому відбувається трансформація сирих сенсорних даних у структуровану інформацію, придатну для інтерпретації. Саме на цьому рівні реалізуються алгоритми обробки сигналів, виділення інформативних ознак та класифікації емоційного стану користувача.

Процес обробки починається з попередньої підготовки даних, яка включає нормалізацію, фільтрацію шумів, синхронізацію різних потоків даних та приведення їх до єдиного формату. Для відеоданих це може включати вирівнювання зображення обличчя, видалення фону та підвищення контрастності, тоді як для аудіосигналів – фільтрацію шуму та виділення мовного сигналу.

Наступним етапом є виділення ознак, тобто перетворення первинних даних у компактне представлення, що зберігає найбільш значущу інформацію. Для візуальних даних такими ознаками можуть бути координати ключових точок обличчя, параметри м'язових рухів або карти активації нейронних мереж. Для аудіосигналів це можуть бути спектральні характеристики, коефіцієнти MFCC або показники інтонації. Поведінкові дані трансформуються у часові характеристики, такі як швидкість введення тексту або траєкторії руху курсора.

Ключовим етапом є класифікація емоцій, яка реалізується за допомогою методів машинного навчання або глибинного навчання. На цьому етапі система визначає, до якого класу емоцій належить поточний стан користувача, спираючись на вектор ознак. Важливою характеристикою цього рівня є його адаптивність, тобто здатність враховувати індивідуальні особливості користувача та змінювати параметри моделі в процесі роботи. Також тут може реалізовуватися

механізм об'єднання різних джерел даних, що дозволяє підвищити достовірність результатів.

Таким чином, рівень обробки виконує функцію інтелектуального ядра системи, перетворюючи різнорідні дані у єдину оцінку емоційного стану.

Коефіцієнти MFCC (mel-frequency cepstral coefficients) – це параметри, що описують спектральні характеристики аудіосигналу з урахуванням особливостей сприйняття звуку людиною [19]. Вони широко застосовуються для аналізу мовлення та дозволяють ефективно виділяти емоційні ознаки голосу.

Схема обчислення MFCC представлена на рисунку 2.2.

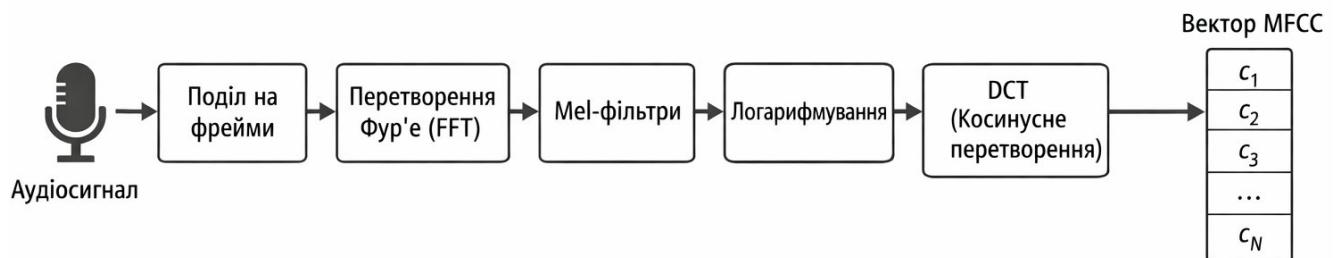


Рисунок 2.2 – Схема обчислення MFCC

MFCC базуються на тому, що людина сприймає частоти нелінійно, низькі частоти ми розрізняємо краще, ніж високі. Саме тому використовується mel-шкала, яка наближена до особливостей слухового сприйняття.

Процес отримання MFCC складається з кількох етапів:

1. Розбиття сигналу на короткі фрагменти (фрейми). Оскільки мовлення змінюється з часом, його аналізують по маленьких відрізках (20 – 40 мс).
2. Перетворення Фур'є (FFT). Здійснюється перехід від часової області до частотної і в результаті отримуємо спектр сигналу.
3. Застосування mel-фільтрів. Спектр пропускається через набір фільтрів, розташованих за mel-шкалою.
4. Логарифмування енергії. Імітує сприйняття гучності людиною.
5. Дискретне косинусне перетворення (DCT). Дає компактне представлення, тобто власне MFCC.

MFCC не зберігають сам звук, а описують форму спектру сигналу, тембр голосу та емоційне забарвлення (напруження, радість, втома).

Слід зазначити, що MFCC важливі для емоцій, тому що емоції сильно впливають на висоту голосу, енергію сигналу, темп мовлення та спектральну структуру

Параметри MFCC вловлюють ці зміни, тому що при стресі коефіцієнти змінюються через напруження голосу, при радості змінюються через варіативність інтонації, при втомі змінюються через зниження енергії.

Mel-фільтри – це набір смугових фільтрів, які використовуються в обробці аудіо для перетворення спектру сигналу у форму, більш наближену до того, як людське вухо сприймає звук. Вони є ключовим етапом у побудові MFCC.

Суть ідеї полягає в тому, що людське вухо сприймає частоти нелінійно. Людина краще розрізняє низькі частоти (до  $\sim 1$  кГц). Високі частоти сприймаються менш детально. Mel-фільтри якраз враховують цю особливість, тому перетворюють лінійний спектр сигналу у mel-шкалу, де відстані між фільтрами приблизно відповідають сприйманню частот людиною.

Розглянемо більш детально як працюють mel-фільтри.

Вхідний аудіосигнал перетворюється в спектр за допомогою FFT (швидкого перетворення Фур'є). На спектр накладається набір трикутних фільтрів, які перекриваються один з одним.

Кожен фільтр виділяє енергію у певній частотній смузі. Фільтри розташовані так, що низькі частоти мають вузькі смуги, а високі – ширші.

Для кожного фільтра обчислюється сумарна енергія сигналу, що проходить через нього. Отримані значення використовуються далі для логарифмування та DCT, щоб сформувати MFCC.

Параметри MFCC роблять спектр більш людським (сприйняття звуку як людини), допомагають виділити інформативні ознаки для розпізнавання мови або емоцій та зменшують вплив високочастотних шумів.

DCT (discrete cosine transform, дискретне косинусне перетворення) – це математичне перетворення, яке використовується для аналізу сигналів і

зображень. Його головна мета – представити дискретний сигнал (наприклад, рядок пікселів або звук) як суму косинусних функцій різних частот.

DCT розкладає сигнал на низькочастотні та високочастотні компоненти. Низькочастотні компоненти відповідають за загальну форму або плавні зміни сигналу, а високочастотні відповідають за деталі та різкі зміни.

У більшості природних сигналів і зображень енергія зосереджена в низькочастотних компонентах. DCT дозволяє відкидати високочастотні компоненти з мінімальною втратою якості, тому його широко використовують у JPEG та MPEG. На відміну від дискретного перетворення Фур'є (DFT), DCT працює тільки з дійсними числами та використовує тільки косинусні функції.

Використовується для обробки зображень: JPEG, відео MPEG, обробки аудіо, зокрема MP3, AAC та для розпізнавання та аналізу сигналів, особливо там, де важливо виділяти частотні характеристики.

Рівень взаємодії забезпечує зв'язок між кіберфізичною системою та кінцевим користувачем або зовнішніми інформаційними системами. Його основною функцією є інтерпретація результатів аналізу та їх представлення у зручній і зрозумілій формі, а також організація обміну даними з іншими підсистемами.

На цьому рівні формується користувацький інтерфейс, який може бути реалізований у вигляді програмного додатку, веб-інтерфейсу або інтегрованого модуля в інші інформаційні системи. Інтерфейс забезпечує візуалізацію емоційного стану у вигляді текстових повідомлень, графіків, індикаторів або рекомендацій. Наприклад, система може сигналізувати про підвищений рівень стресу або втому користувача та пропонувати відповідні дії, такі як перерва або зміна режиму роботи.

Крім функції відображення, рівень взаємодії виконує роль інтеграційного середовища, яке забезпечує обмін даними з іншими системами, зокрема системами управління навчанням, корпоративними платформами або системами підтримки прийняття рішень. Це дозволяє використовувати результати аналізу емоційного стану для адаптації навчального процесу, підвищення продуктивності

праці або покращення користувацького досвіду.

Важливою особливістю цього рівня є необхідність забезпечення зручності, інтуїтивності та етичності взаємодії. Оскільки система працює з чутливими персональними даними, значна увага приділяється питанням конфіденційності, захисту інформації та прозорості прийняття рішень. Таким чином, рівень взаємодії не лише відображає результати роботи системи, але й формує довіру користувача до неї.

## 2.2 Виділення ознак у кіберфізичній системі визначення емоційного стану

У кіберфізичній системі визначення емоційного стану людини процес виділення ознак є ключовим етапом кібернетичного рівня, оскільки саме на цьому етапі відбувається перетворення сирих сенсорних даних у структуроване представлення, придатне для подальшого аналізу та класифікації. З огляду на багатомодальність вхідної інформації, у системі використовується комбінований підхід до виділення ознак, який включає аналіз аудіальних, візуальних та поведінкових характеристик користувача.

Для обробки аудіосигналу, отриманого з мікрофона на рівні сприйняття, застосовується метод mel-частотного кепстрального аналізу. Після попередньої фільтрації та спектрального перетворення сигнал пропускається через банк mel-фільтрів, що дозволяє адаптувати частотне представлення до особливостей слухового сприйняття людини. У результаті формується набір коефіцієнтів MFCC, які компактно описують спектральні характеристики мовлення, включаючи тембр, енергію та інтонаційні особливості. Отримані коефіцієнти є інформативними ознаками для подальшого визначення емоційного стану, оскільки вони відображають зміни голосу, спричинені психологічним станом користувача.

Візуальний аналіз емоцій здійснюється на основі даних, отриманих із відеокамери, та реалізується за допомогою згорткових нейронних мереж.

Зображення обличчя користувача проходить етапи попередньої обробки, включаючи нормалізацію, вирівнювання та виділення області обличчя, після чого подається на вхід згорткової нейронної мережі. У процесі згортки автоматично виділяються просторові ознаки, такі як положення ключових точок обличчя, мімічні зміни та мікровирази. На відміну від класичних методів, використання CNN дозволяє уникнути ручного задання ознак та забезпечує високу точність розпізнавання емоцій за рахунок глибинного навчання.

Для врахування часової динаміки емоційного стану, яка є суттєвою характеристикою поведінки людини, у системі застосовуються рекурентні нейронні мережі типу LSTM. Вхідними даними для цієї моделі є послідовності ознак, отриманих з аудіальних та візуальних джерел. LSTM-мережа здатна зберігати інформацію про попередні стани та враховувати залежності у часі, що дозволяє більш точно ідентифікувати емоції, які змінюються поступово або мають короткочасні прояви. Це особливо важливо для виявлення таких станів, як стрес, втома або роздратування.

Згорткові нейронні мережі застосовуються для аналізу візуальної інформації, отриманої з відеокамери на рівні сприйняття. Основною перевагою CNN є здатність автоматично виділяти просторові ознаки із зображення без необхідності ручного конструювання дескрипторів. У контексті визначення емоційного стану це дозволяє ефективно аналізувати міміку обличчя, положення ключових точок та мікровирази. Архітектура CNN базується на послідовності згорткових, активаційних та підвибіркових шарів, які забезпечують поступове узагальнення інформації від локальних до глобальних ознак. У результаті формується компактне представлення зображення, яке містить найбільш інформативні характеристики для подальшої класифікації емоцій.

Важливою особливістю CNN є інваріантність до зсувів та незначних змін у зображенні, що дозволяє системі коректно працювати в умовах варіацій освітлення, положення голови або якості відео. Це робить їх ефективним інструментом для застосування у кіберфізичних системах, які функціонують у реальному середовищі.

Рекурентні нейронні мережі типу LSTM використовуються для обробки послідовних даних і врахування часової динаміки емоційного стану [20]. На відміну від звичайних нейронних мереж, LSTM мають внутрішню пам'ять, яка дозволяє зберігати інформацію про попередні стани системи та враховувати залежності у часі. Це є критично важливим для аналізу емоцій, оскільки вони змінюються поступово і можуть проявлятися у вигляді короткочасних або тривалих патернів поведінки.

Архітектура LSTM включає спеціальні механізми керування пам'яттю, зокрема входні, вихідні та забуваючі гейти, які регулюють потік інформації в мережі. Завдяки цьому LSTM здатні ефективно працювати з довгими послідовностями даних без втрати важливої інформації. У межах даної системи LSTM застосовується для аналізу часових рядів ознак, отриманих як з аудіальних (наприклад, MFCC), так і з візуальних джерел. Поєднання CNN та LSTM у єдиній архітектурі дозволяє реалізувати багаторівневий аналіз даних: CNN відповідає за виділення просторових ознак, тоді як LSTM відповідає за моделювання їх часової еволюції. Такий гібридний підхід забезпечує більш повне та точне представлення емоційного стану користувача, що є особливо важливим для кіберфізичних систем, орієнтованих на роботу в реальному часі.

Таким чином, використання згорткових та рекурентних нейронних мереж дозволяє суттєво підвищити ефективність системи визначення емоційного стану за рахунок глибинного аналізу як просторових, так і часових характеристик даних [21, 24].

Інтеграція ознак різної природи здійснюється на рівні обробки шляхом їх об'єднання у єдиний вектор ознак, що подається до модуля класифікації. Такий підхід, відомий як багатомодальна інтеграція, дозволяє компенсувати недоліки окремих джерел даних та підвищити загальну достовірність оцінювання емоційного стану.

Таким чином, запропонований підхід до виділення ознак, який поєднує MFCC для аналізу мовлення, згорткові нейронні мережі для обробки зображень та рекурентні мережі LSTM для врахування часової динаміки, забезпечує

комплексне представлення емоційного стану користувача у кіберфізичній системі.

### 2.3 Класифікація емоційного стану у кіберфізичній системі

Після етапу виділення ознак наступним ключовим компонентом кібернетичного рівня є класифікація емоційного стану користувача, яка полягає у віднесенні сформованого вектора ознак до одного з попередньо визначених класів емоцій. У межах даної системи класифікація реалізується як задача багатокласового розпізнавання, де множина можливих станів включає, зокрема, нейтральний стан, радість, стрес, втому та інші психоемоційні прояви.

Класифікація емоційного стану передбачає автоматичне визначення емоцій людини на основі сигналів, що надходять від сенсорів, з використанням обчислювальних моделей та алгоритмів штучного інтелекту. Емоції визначаються як комплексні психофізіологічні реакції на внутрішні або зовнішні стимули, що проявляються у поведінкових, вербальних та фізіологічних ознаках. В рамках КФС це зазвичай включає такі категорії сигналів, як:

1. Фізіологічні сигнали як от, електрокардіограма (ЕКГ), електроенцефалографія (ЕЕГ), електроміографія (ЕМГ), частота серцевих скорочень, рівень електропровідності шкіри.

2. Акустичні сигнали, а саме, тон голосу, швидкість мовлення, інтонаційні патерни.

3. Візуальні сигнали, як міміка, рух очей, положення голови та тілесна експресія.

Інтеграція цих даних дозволяє створювати багатомодальні моделі, які підвищують точність класифікації емоцій у динамічних середовищах.

У сучасних КФС класифікація емоційного стану зазвичай здійснюється за двома підходами категоріальним та димензійним.

Категоріальний підхід передбачає виділення дискретних емоцій, наприклад, щастя, сум, страх, гнів, здивування та спокій. Такий метод зручний для інтеграції у системи з простим алгоритмом прийняття рішень, проте може втрачати нюанси

емоційних станів.

Димензійний підхід базується на безперервних шкалах, таких як валентність (від позитивного до негативного стану) та збудження (від спокою до активної реакції). Цей метод дозволяє отримувати більш тонкі відмінності у відображенні емоцій та підходить для адаптивних систем з персоналізацією.

Для розробки кіберфізичної системи визначення емоційного стану критично важливим є вибір методів класифікації, які здатні ефективно обробляти мультисенсорні дані. Класифікація є третім етапом роботи системи, що реалізується за допомогою нейронних мереж для визначення базових емоцій на основі відео-, аудіо- та біометричних даних.

Нижче наведено класифікацію основних методів, які застосовуються в сучасних інтелектуальних системах:

1. Методи на основі глибокого навчання (Deep Learning).
2. Згорткові нейронні мережі (CNN).
3. Рекурентні нейронні мережі (RNN) та LSTM.
4. Глибокі мережі довіри (DBN).

Перший напрям є найбільш перспективним напрямом, який дозволяє автоматично виділяти ознаки зі складних сигналів без ручного проектування дескрипторів.

Згорткові нейронні мережі використовуються переважно для відеоаналізу міміки обличчя. Вони здатні розпізнавати просторові структури та мікрорухи, що відповідають за емоційні прояви.

Рекурентні нейронні мережі та LSTM застосовуються для аналізу часових послідовностей, таких як аудіосигнали та динаміка зміни фізіологічних параметрів. Це дозволяє враховувати контекст попередніх станів користувача.

Глибокі мережі довіри використовуються для моделювання складних ієрархічних зв'язків у біометричних даних, зокрема в EEG-сигналах.

Методи аналізу фізіологічних та біометричних сигналів фокусуються на обробці даних, які людина не може свідомо контролювати, що підвищує об'єктивність системи.

Аналіз ЕЕГ (електроенцефалографія) використовує нейронні мережі для розпізнавання емоцій на основі активності мозку. Це дозволяє створювати емоційні мережі для когнітивних досліджень.

Статистичний аналіз біометричних показників оцінює рівень стресу або розслаблення через пульс, частоту дихання та електропровідність шкіри.

Метод опорних векторів (SVM) часто використовується як базовий рівень (baseline) для класифікації станів на основі фізіологічних векторів ознак [35, 38].

Оскільки КФС поєднує різні типи даних, ключовим методом є інтеграція сигналів для підвищення точності та надійності.

Рання інтеграція здійснює об'єднання векторів ознак з відео, аудіо та сенсорів в один спільний вектор перед подачею на вхід класифікатора.

При пізній інтеграції кожна модальність (відео, звук, біометрія) обробляється окремою нейромережею, а фінальне рішення приймається на основі аналізу результатів усіх мереж.

Гібридна інтеграція – це комбінація обох підходів для створення максимально адаптивних систем.

Для класифікації емоцій у КФС часто використовується не просто дискретний набір категорій (радість, гнів тощо), а континуальні моделі [44].

Одним із ключових підходів до класифікації емоцій у сучасних кіберфізичних системах є модель Valence-Arousal, запропонована Расселом. Ця модель відрізняється від традиційного категоріального підходу тим, що емоції не визначаються як дискретні категорії (радість, сум, гнів тощо), а описуються у двомірному просторі, де кожна емоція має координати по двох осях:

1. Valence (валентність) – шкала від негативної до позитивної емоції. Вона характеризує суб'єктивну приємність чи неприємність стану користувача. Наприклад, емоція радості має високу позитивну валентність, тоді як страх або сум – негативну. Валентність дозволяє системам оцінювати, наскільки досвід або стимул сприймається користувачем як комфортний чи дискомфортний.

2. Arousal (збудження) – шкала інтенсивності емоційної реакції, від низького до високого. Вона відображає фізіологічну та когнітивну активність

користувача у відповідь на стимул. Високий рівень збудження може відповідати захопленню, тривозі чи гніву, тоді як низький рівень відповідає спокою або апатії. Ця шкала важлива для оцінки того, наскільки активною є реакція користувача на події або сигнали системи.

Модель Рассела представлена на рисунку 2.3.



Рисунок 2.3 – Координатна модель емоційних станів (модель Рассела) [42]

Важливо, що комбінація координат Valence та Arousal дозволяє отримати більш точну і тонку картину емоційного стану користувача. Наприклад, дві емоції можуть мати однакову валентність, але різний рівень збудження: «спокійне щастя» і «захоплена радість» мають позитивну валентність, але відрізняються інтенсивністю. Аналогічно, емоції з однаковою збудженням можуть відрізнятися за валентністю, як-от гнів і захоплення.

У кіберфізичних системах ця модель використовується для адаптивних сценаріїв взаємодії, де потрібно враховувати не лише факт емоції, а й її динаміку та інтенсивність. Наприклад, адаптивні освітні платформи можуть регулювати складність завдань, якщо користувач перебуває у стані високого збудження (стресу) або низького збудження (втоми).

Ігрові системи можуть підбирати сюжетні події та рівень активності персонажів відповідно до емоційного стану гравця.

Інтерфейси людина-машина у промислових або медичних системах можуть змінювати спосіб подачі інформації, якщо користувач демонструє негативну валентність та високе збудження, що сигналізує про тривогу чи стрес.

Таким чином, модель Valence–Arousal забезпечує дві ключові переваги для КФС:

1. Гнучке і безперервне представлення емоцій, що дозволяє системам працювати з перехідними та змішаними станами.
2. Здатність інтегрувати фізіологічні, акустичні та візуальні дані для створення більш адаптивного та персоналізованого користувацького досвіду.

Впровадження цієї моделі у класифікацію емоцій робить систему більш «людиноорієнтованою», дозволяючи не лише визначати емоції, а й реагувати на них у реальному часі, що особливо важливо для інтелектуальних навчальних платформ, ігрових симуляторів та кіберфізичних систем у сфері охорони здоров'я.

Використання цих методів у комплексі забезпечує роботу системи у реальному часі та її високу адаптивність до індивідуальних особливостей людини.

Вхідними даними для класифікатора є інтегрований вектор ознак, сформований на попередньому етапі шляхом об'єднання аудіальних, візуальних та поведінкових характеристик. Такий підхід дозволяє врахувати як статичні, так і динамічні аспекти емоційного стану, що значно підвищує точність розпізнавання порівняно з одноmodalними системами.

У якості базового методу класифікації в системі використовується глибинна нейронна мережа, яка виконує нелінійне відображення простору ознак у простір класів емоцій. Математично цей процес можна подати у вигляді формули 2.1:

$$y = \sigma(W * X + b), \quad (2.1)$$

де  $X$  – вектор ознак;

$W$  – матриця вагових коефіцієнтів;

$b$  – вектор зміщення;

$\sigma$  – функція активації.

На виході формується вектор ймовірностей належності до кожного класу емоцій, після чого обирається клас із максимальною ймовірністю.

З метою підвищення точності та стійкості системи до невизначеності, у моделі також передбачено використання елементів нечіткої логіки. Це дозволяє враховувати ситуації, коли емоційний стан не є чітко вираженим і може одночасно належати до кількох класів з різним ступенем впевненості. У такому випадку результат класифікації формується не як однозначне рішення, а як набір ступенів належності до кожної емоції.

Додатково, з урахуванням часової природи емоційних змін, у системі реалізується механізм згладжування результатів класифікації у часі. Це дозволяє уникнути різких стрибків у визначенні емоційного стану та забезпечує більш стабільну і реалістичну інтерпретацію поведінки користувача.

Такий підхід особливо важливий у випадках короткочасних змін міміки або голосу, які не завжди відображають реальний емоційний стан.

Результат класифікації передається на рівень взаємодії, де використовується для візуалізації, формування рекомендацій або інтеграції з іншими інформаційними системами.

Таким чином, модуль класифікації виконує функцію прийняття рішень у кіберфізичній системі та забезпечує інтерпретацію складних багатомодальних даних у вигляді зрозумілого для користувача результату.

Функціонування кіберфізичної системи визначення емоційного стану людини може бути формалізоване у вигляді багаторівневої математичної моделі, яка відображає послідовність перетворення вхідних даних у кінцевий результат.

Нехай на вході системи формується множина сенсорних даних, виду 2.2:

$$D(t) = \{V(t), A(t), B(t)\}, \quad (2.2)$$

де  $V(t)$  – візуальні дані (зображення обличчя);

$A(t)$  – аудіальні дані (мовний сигнал);

$B(t)$  – поведінкові параметри (взаємодія з комп'ютером);

$t$  – момент часу.

На етапі обробки здійснюється відображення сирих даних у простір ознак, представлених формулою 2.3:

$$X(t) = \Phi(D(t)), \quad (2.2)$$

де  $\Phi$  – оператор виділення ознак, який включає в себе MFCC-перетворення для аудіо; згорткове перетворення (CNN) для зображень та статистичний аналіз поведінкових даних.

Отриманий вектор ознак має вигляд 2.3:

$$X(t) = \{x_1, x_2, \dots, x_n\}, \quad (2.3)$$

де  $x_1, x_2, \dots, x_n$  – ознаки.

Далі виконується класифікація емоційного стану.

## 2.4 Моделювання процесу збору та обробки мультисенсорних даних

Процес збору даних у КФС організований за принципом мультисенсорності, що дозволяє нівелювати похибки окремих каналів інформації та підвищити загальну надійність системи. Модель збору даних включає три основні фізичні рівні.

1. Візуальний канал для отримання відеопотоку з веб-камери для подальшого аналізу динаміки м'язів обличчя.

2. Акустичний канал для фіксації аудіосигналів через мікрофон для аналізу тональності, тембру та ритмічних характеристик мовлення.

3. Фізіологічний канал для збору біометричних параметрів (пульс, частота дихання, електропровідність шкіри) за допомогою спеціалізованих сенсорів.

Процес визначення емоційного стану  $E$  можна представити як функцію від множини вхідних сигналів, що фіксуються комп'ютером, представлену формулою 2.4:

$$E = f(V, A, B), \quad (2.4)$$

де  $V$  – вектор ознак відеопотоку (міміка обличчя);

$A$  – вектор ознак аудіосигналу (параметри голосу);

$B$  – вектор біометричних параметрів (пульс, провідність шкіри).

#### 2.4.1 Модель відеоаналізу емоційних проявів користувача

Першим етапом моделювання є розробка моделі обробки відеопотоку, що отримується з веб-камери в реальному часі. Основним завданням цієї моделі є ідентифікація та інтерпретація невербальних сигналів, що виражаються через міміку обличчя.

Процес відеоаналізу в межах кіберфізичної системи реалізується через такі кроки:

1. Захоплення зображення, де відбувається отримання вхідного відеопотоку та його попередня обробка для стабілізації кадрів.

2. Детекція ключових точок базується на використанні алгоритмів комп'ютерного зору для виділення опорних точок обличчя (брів, очей, губ), що дозволяє математично описати його геометрію.

3. Аналіз динамічних змін відбувається через оцінку відхилень позицій ключових точок у часі для розпізнавання мікрорухів, що відповідають за прояви базових емоцій.

4. Класифікація ознак полягає у застосуванні глибокого навчання для співвіднесення виділених ознак із конкретними емоційними станами.

Дана модель забезпечує високу швидкість обробки, що є критичним для адаптивних систем, які працюють у реальному часі.

Структурна модель відеоаналізу емоційних проявів користувача представлена на рисунку 2.4.

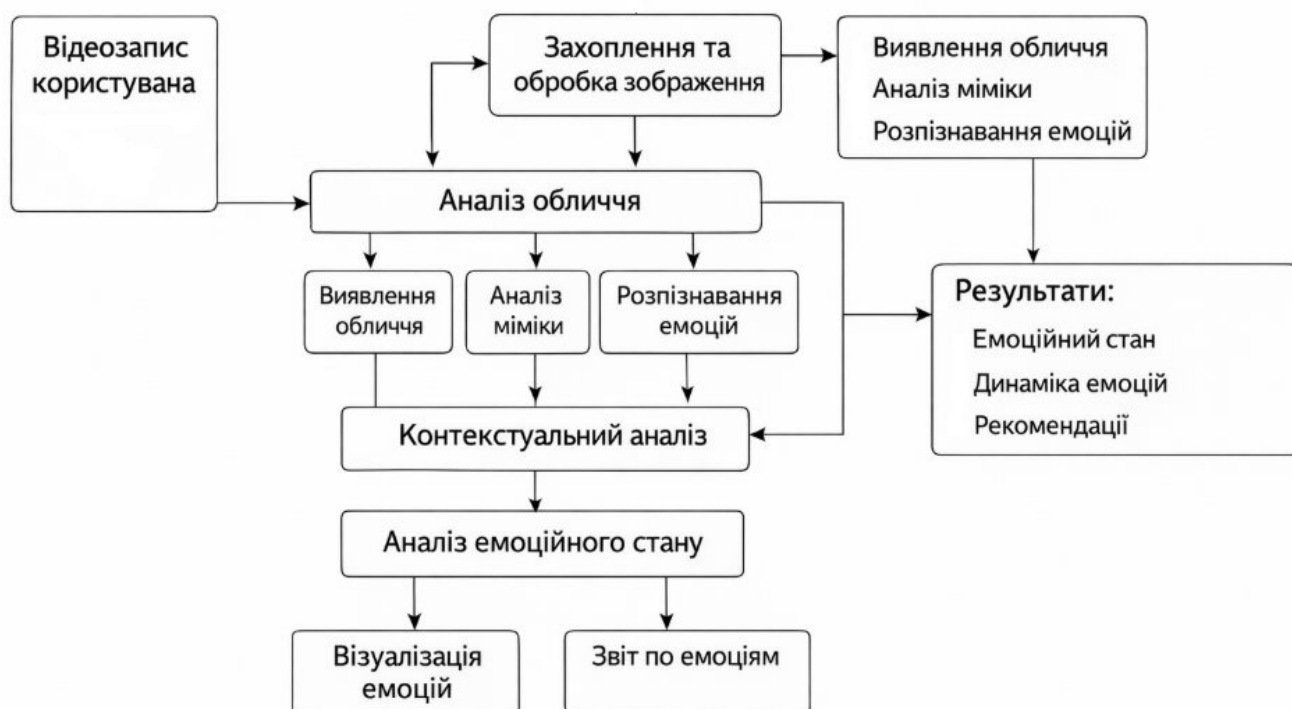


Рисунок 2.4 – Структурна модель відеоаналізу емоційних проявів користувача

Процес розпізнавання емоцій за допомогою комп'ютерного зору FER у межах кіберфізичної системи базується на багатоетапному конвеєрі обробки візуальних даних. Основна мета моделі – це перетворити сирий відеопотік у структурований вектор ознак, що описує стан користувача.

Розглянемо кожен етап більш детально.

Першим етапом у процесі автоматизованої класифікації емоційного стану за відеопотоком є отримання зображень із камери та їх попередня обробка, що забезпечує стабільність подальшого аналізу. Вхідні дані у вигляді послідовності

кадрів можуть містити значні варіації, зумовлені змінами освітлення, шумами сенсора або положенням користувача відносно камери, тому нормалізація є критично важливою. З цією метою виконується перетворення зображення у відтінки сірого, що дозволяє суттєво зменшити обчислювальну складність без втрати ключової інформації про структуру обличчя.

Додатково застосовується адаптивне гістограмне вирівнювання, зокрема метод CLAHE, який підвищує локальну контрастність зображення та покращує видимість дрібних деталей, таких як контури губ, очей і брів [45]. Після цього здійснюється детекція обличчя, яка є базовою умовою для подальшого аналізу. Для цього можуть використовуватися як класичні алгоритми на основі каскадів Хаара, так і більш сучасні підходи, що забезпечують вищу точність та стійкість до змін ракурсу, зокрема глибокі нейронні мережі типу MTCNN або MediaPipe Face Detection.

Після локалізації обличчя наступним етапом є визначення його геометричної структури шляхом детекції ключових точок, або так званих landmarks. Ці точки відображають характерні анатомічні орієнтири обличчя, включаючи контури очей, брів, носа, рота та щелепи. Найбільш поширеною є модель із 68 опорними точками, яка дозволяє отримати достатньо детальний опис форми обличчя для задач розпізнавання емоцій. Формально обличчя можна представити як множину координат цих точок, що створює дискретну геометричну модель. На основі цієї моделі обчислюються різноманітні метричні характеристики, зокрема відстані між окремими точками, кути нахилу та співвідношення довжин. Такі ознаки відображають фізіологічні прояви емоцій, наприклад ступінь відкриття рота, підняття брів або звуження очей. Використання евклідової відстані між точками дозволяє формалізувати ці зміни та зробити їх придатними для подальшої обробки алгоритмами машинного навчання.

Однак статичний аналіз окремого кадру є недостатнім для повноцінного розуміння емоційного стану, оскільки емоції мають динамічну природу і проявляються у зміні міміки з часом. Саме тому наступним кроком є аналіз часових змін геометричних характеристик обличчя. У цьому контексті широко

використовується концепція Facial Action Coding System, розроблена Полом Екманом, яка описує міміку як комбінацію базових рухів м'язів обличчя, відомих як Action Units [46]. Кожна емоція відповідає певному набору таких одиниць, що активуються з різною інтенсивністю. Для кількісного опису цих змін розглядається переміщення кожної ключової точки між сусідніми кадрами, яке може бути представлене у вигляді вектора зміщення. Аналіз цих векторів дозволяє відслідковувати навіть незначні мікрорухи, які є важливими індикаторами емоційних реакцій.

Додатково застосовуються методи оптичного потоку, що оцінюють рух пікселів між кадрами і дозволяють виявляти тонкі зміни текстури шкіри, які не завжди можна зафіксувати лише за координатами опорних точок [47]. Це особливо важливо для розпізнавання мікровиразів, які тривають долі секунди і можуть свідчити про приховані або пригнічені емоції.

Завершальним етапом є класифікація отриманих ознак із використанням методів глибокого навчання. На цьому етапі геометричні та динамічні характеристики обличчя перетворюються у вхідні дані для нейронної мережі, яка виконує узагальнення та прийняття рішення щодо емоційного стану користувача. Згорткові нейронні мережі демонструють високу ефективність у задачах аналізу зображень, оскільки здатні автоматично виділяти просторові ознаки різного рівня складності, від простих контурів до складних патернів міміки.

Архітектури на кшталт VGG-16 або ResNet дозволяють отримувати глибокі представлення зображення, що значно підвищує точність класифікації [48]. Водночас для врахування часової залежності між кадрами доцільно використовувати гібридні моделі, які поєднують CNN із рекурентними мережами типу LSTM. Такий підхід дозволяє системі не лише аналізувати окремі зображення, але й враховувати контекст попередніх станів, що є критично важливим для коректного розпізнавання емоцій у відеопослідовностях [49].

У результаті описана багаторівнева модель забезпечує комплексний аналіз емоційного стану, поєднуючи просторові, геометричні та часові характеристики міміки.

Це дозволяє досягти високої точності класифікації та адаптувати систему до реальних умов експлуатації, що є важливим для інтеграції у кіберфізичні системи різного призначення, зокрема в інтелектуальні навчальні середовища, ігрові платформи та системи підтримки прийняття рішень.

#### 2.4.2 Модель аудіоаналізу мовленнєвих та акустичних характеристик

Модель аудіоаналізу в кіберфізичній системі спрямована на виявлення емоційного стану користувача через аналіз характеристик його мовлення, що виступають важливим джерелом інформації про психоемоційний стан. На відміну від візуального каналу, який базується на міміці та жестах, аудіоканал дозволяє оцінювати внутрішні реакції навіть у випадках, коли обличчя частково або повністю приховане, або ж мімічні прояви є слабо вираженими. Таким чином, аудіоаналіз виступає важливою складовою багатомодальної системи, підвищуючи її надійність, стійкість до шумів середовища та загальну точність класифікації емоцій.

Структурна модель аудіоаналізу мовленнєвих та акустичних характеристик представлена на рисунку 2.5.

Процес обробки аудіосигналу починається з його захоплення за допомогою мікрофонів та подальшої цифрової обробки. Отриманий сигнал є неперервною функцією часу, яка підлягає дискретизації та квантуванню для подальшого аналізу.

На етапі попередньої обробки здійснюється фільтрація шумів, нормалізація амплітуди та виділення мовленнєвих сегментів, що дозволяє відокремити корисний сигнал від фонових перешкод.

Це особливо важливо в умовах реального середовища, де присутні сторонні звуки, що можуть впливати на точність аналізу.

Одним із ключових аспектів моделювання є аналіз тональності та інтонації мовлення, які відображають частотні характеристики голосу. Висота основного

тону, або фундаментальна частота, є важливим індикатором емоційного стану: підвищення частоти часто пов'язане зі станами збудження, такими як радість або тривога, тоді як її зниження може свідчити про втому або пригнічення. Інтонаційні контури, що описують зміну частоти у часі, дозволяють виявляти емоційні акценти в мовленні, наприклад підвищення тону в кінці фрази або різкі зміни висоти голосу. Для математичного опису цих характеристик широко застосовується спектральний аналіз, зокрема швидке перетворення Фур'є, що дозволяє перейти від часової області до частотної та виділити ключові гармонічні компоненти сигналу [50].

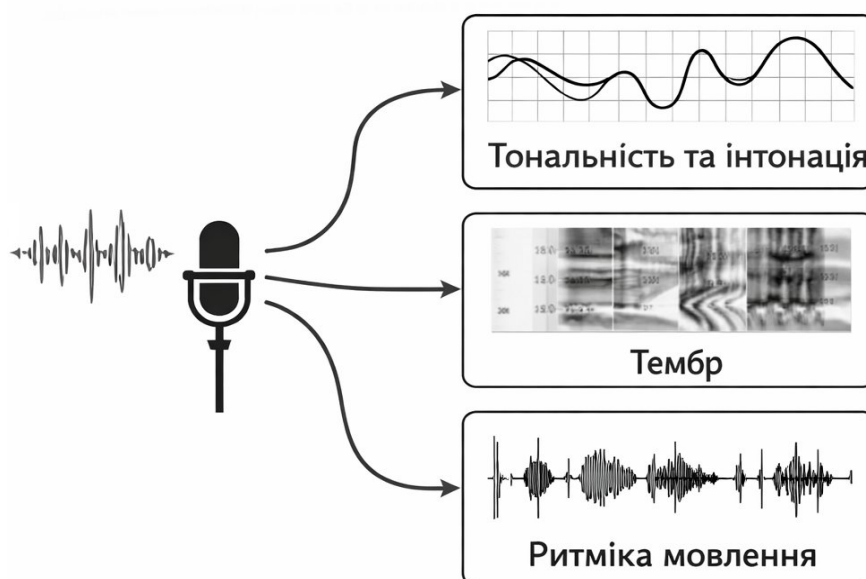


Рисунок 2.5 – Структурна модель аудіоаналізу мовленнєвих та акустичних характеристик

Не менш важливим параметром є тембр голосу, який визначається сукупністю спектральних характеристик і відображає індивідуальні особливості мовлення. Тембр змінюється залежно від фізіологічного та емоційного стану людини, оскільки напруження голосових зв'язок, дихання та резонансні властивості голосового тракту впливають на форму спектру сигналу. Для кількісного опису тембру часто використовуються мел-частотні кепстральні коефіцієнти, які моделюють сприйняття звуку людським слухом і дозволяють

ефективно представляти спектральну структуру мовлення у компактній формі. Ці коефіцієнти широко застосовуються як в задачах розпізнавання мовлення, так і в системах визначення емоцій [52].

Ритміка мовлення також є інформативним джерелом для оцінки емоційного стану. Вона включає такі характеристики, як темп мовлення, тривалість пауз, регулярність артикуляції та розстановка логічних наголосів. Наприклад, пришвидшене мовлення з короткими паузами може свідчити про збудження або нервозність, тоді як повільний темп і довгі паузи часто асоціюються з втомою або невпевненістю. Аналіз ритміки здійснюється шляхом сегментації аудіосигналу на окремі мовленнєві фрагменти та оцінки їх часових характеристик. У цьому контексті важливу роль відіграє визначення енергії сигналу та її зміни у часі, що дозволяє виявляти акценти та інтонаційні виділення в мовленні [53].

Для інтеграції всіх зазначених параметрів у єдину модель використовується математичне представлення аудіосигналу у вигляді багатовимірного вектора ознак, який формується на основі часових, частотних та спектральних характеристик. Такий підхід дозволяє застосовувати алгоритми машинного навчання та глибинного навчання для класифікації емоцій. Зокрема, згорткові нейронні мережі можуть використовуватися для аналізу спектрограм, які відображають зміну частотного спектру у часі, тоді як рекурентні мережі типу LSTM дозволяють враховувати часову динаміку мовлення та контекст попередніх фрагментів.

У результаті, модель аудіоаналізу забезпечує глибоке розуміння емоційного стану користувача на основі його голосових характеристик, що значно підвищує ефективність кіберфізичної системи. Вона дозволяє створювати адаптивні інтерфейси, які реагують на зміну інтонації, тембру та ритму мовлення, коригуючи свою поведінку відповідно до поточного стану користувача. Це відкриває широкі можливості для застосування у сфері інтелектуальних навчальних систем, голосових асистентів, ігрових середовищ та систем підтримки прийняття рішень, де врахування емоцій є важливим фактором ефективної взаємодії людини та технологій.

### 2.4.3 Модель аналізу біометричних та фізіологічних показників

Модель біометричного аналізу є одним із найбільш надійних та об'єктивних компонентів кіберфізичної системи класифікації емоційного стану, оскільки вона базується на фізіологічних реакціях організму, які практично не піддаються свідомому контролю з боку людини. На відміну від міміки чи мовлення, що можуть бути навмисно змінені або приховані, біометричні показники відображають глибинні процеси функціонування нервової та серцево-судинної систем, забезпечуючи більш достовірну оцінку психоемоційного стану. Основою такої моделі є інтеграція даних із різноманітних сенсорів, що безперервно фіксують внутрішні параметри організму та передають їх у систему для подальшого аналізу.



Рисунок 2.6 – Структурна модель аналізу біометричних та фізіологічних показників

Одним із ключових джерел інформації є показники серцево-судинної

системи, зокрема частота серцевих скорочень та її варіабельність.

Серцевий ритм безпосередньо пов'язаний із діяльністю автономної нервової системи, яка регулює реакції організму на стрес і зовнішні подразники. Збільшення частоти пульсу, як правило, свідчить про підвищення рівня емоційного збудження або тривоги, тоді як зниження або стабілізація ритму характерні для станів спокою та релаксації.

Важливим показником є також варіабельність серцевого ритму, яка відображає здатність організму адаптуватися до змін середовища. Висока варіабельність зазвичай асоціюється з добрим адаптаційним потенціалом і низьким рівнем стресу, тоді як її зниження може свідчити про перевантаження або емоційне виснаження. Математичне моделювання цих параметрів дозволяє виявляти навіть незначні зміни в емоційному стані користувача.

Другим важливим компонентом є аналіз респіраторних даних, які відображають особливості дихання людини. Частота, глибина та ритмічність дихання змінюються залежно від емоційного стану та рівня фізіологічного навантаження. У стані спокою дихання є рівномірним і повільним, тоді як у ситуаціях стресу або тривоги воно стає частішим і поверхневим. Крім того, нерегулярність дихального ритму може свідчити про емоційну нестабільність або напруження. Аналіз респіраторних сигналів у поєднанні з іншими біометричними показниками дозволяє підвищити точність визначення станів переходу, наприклад від спокою до збудження, що є важливим для адаптивних систем.

Особливу роль у моделі відіграє електродермальна активність, яка характеризує зміни електропровідності шкіри під впливом роботи симпатичної нервової системи. Цей параметр безпосередньо пов'язаний із реакцією організму на емоційні стимули, оскільки при підвищенні рівня збудження активується потовиділення, що змінює провідність шкіри.

Таким чином, електродермальна активність виступає одним із найбільш чутливих індикаторів емоційного збудження, особливо у випадках, коли інші ознаки є маловираженими. Аналіз цього сигналу дозволяє виявляти короточасні реакції на зовнішні подразники та оцінювати інтенсивність емоційної відповіді.

Перспективним напрямом розвитку біометричних моделей є використання електроенцефалографічних даних, які відображають електричну активність мозку. ЕЕГ дозволяє безпосередньо аналізувати когнітивні процеси та емоційні реакції на рівні центральної нервової системи. Різні діапазони частот мозкових хвиль асоціюються з певними станами, такими як концентрація, релаксація або стресс [58].

Наприклад, збільшення активності бета-хвиль може свідчити про підвищену когнітивну активність або напруження, тоді як альфа-хвилі часто пов'язані зі станом спокою. Хоча використання ЕЕГ у реальних кіберфізичних системах наразі обмежене складністю обладнання та обробки сигналів, цей напрям має значний потенціал для підвищення точності розпізнавання складних емоційних і когнітивних станів.

Інтеграція всіх зазначених біометричних параметрів у єдину модель дозволяє сформуванню багатовимірний опис стану користувача, який враховує як швидкі реакції організму, так і довготривалі зміни. Комбінування даних із різних сенсорів забезпечує більш надійну інтерпретацію, оскільки окремі показники можуть бути неоднозначними або залежати від індивідуальних особливостей людини.

У рамках кіберфізичної системи ці дані обробляються за допомогою алгоритмів машинного навчання, які виявляють закономірності між фізіологічними параметрами та емоційними станами.

У підсумку, модель біометричного аналізу створює основу для високоточної класифікації рівнів стресу, збудження та релаксації, що є критично важливим для систем підтримки прийняття рішень, медичних застосувань та адаптивних інтерфейсів.

Вона дозволяє системі не лише фіксувати поточний стан користувача, але й прогнозувати його зміни, забезпечуючи більш ефективну та безпечну взаємодію людини з технологічним середовищем.

## 2.5 Висновки

У другому розділі магістерської роботи було проведено моделювання процесу визначення емоційного стану людини в межах кіберфізичної системи. Розгляд задачі здійснювався з позицій системного підходу, що дозволило формалізувати структуру, функції та основні етапи обробки інформації.

Було розроблено структурну модель кіберфізичної системи, яка включає три основні рівні: рівень сприйняття, рівень обробки та рівень взаємодії. Визначено функціональне призначення кожного рівня та їх взаємозв'язки, що забезпечують безперервний цикл збору, обробки та інтерпретації даних.

Особливу увагу приділено етапу виділення ознак, де обґрунтовано доцільність використання комбінованого підходу. Для аналізу аудіосигналів застосовано коефіцієнти MFCC, які дозволяють ефективно враховувати особливості сприйняття звуку людиною. Для обробки візуальних даних використано згорткові нейронні мережі, що забезпечують автоматичне виділення мімічних ознак, а для врахування часової динаміки – рекурентні нейронні мережі типу LSTM.

Отримані результати будуть використані для подальшої програмної реалізації системи та проведення експериментальних досліджень її ефективності.

### 3 МЕТОД ВИЗНАЧЕННЯ ЕМОЦІЙНОГО СТАНУ ЛЮДИНИ ЗА КОМП'ЮТЕРОМ

#### 3.1 Основи методу визначення емоційного стану

Метод визначення емоційного стану у кіберфізичній системі базується на поєднанні підходів обробки сигналів, машинного навчання та багатомодального аналізу даних. Його основною метою є автоматичне визначення психоемоційного стану користувача на основі непрямих ознак, що відображають внутрішні реакції організму. На відміну від традиційних підходів, які використовують один тип даних, сучасні методи орієнтовані на інтеграцію декількох джерел інформації, що дозволяє підвищити точність і стійкість системи до зовнішніх впливів.

В основі методу лежить уявлення про емоційний стан як багатовимірну величину, що формується під впливом фізіологічних, поведінкових та когнітивних факторів. Для його опису використовуються як категоріальні моделі, де емоції представлені у вигляді дискретних класів, так і димензійні підходи, зокрема модель Valence – Arousal, яка дозволяє відобразити емоційний стан у безперервному просторі. Така подвійна інтерпретація забезпечує гнучкість системи та можливість адаптації до різних прикладних задач.

Послідовність кроків методу представлена на рисунку 3.1.



Рисунок 3.1 – Послідовність кроків методу визначення емоційного стану людини

Процес визначення емоційного стану починається зі збору даних, який

здійснюється за допомогою сенсорних пристроїв, інтегрованих у кіберфізичну систему. Вхідними даними можуть бути відеопотік із камери, аудіосигнал із мікрофона та біометричні показники, отримані з носимих сенсорів. На цьому етапі важливо забезпечити синхронізацію потоків даних та їх попередню обробку, що включає фільтрацію шумів, нормалізацію та сегментацію. Метою є приведення даних до уніфікованого формату, придатного для подальшого аналізу.

Наступним етапом є виділення ознак, що є ключовим компонентом методу. Для візуальних даних це можуть бути геометричні характеристики обличчя, положення ключових точок, мікромімічні зміни та рухи очей. Аудіальні дані аналізуються з точки зору спектральних та часових характеристик, таких як частота основного тону, тембр, енергія сигналу та ритміка мовлення. Біометричні сигнали представлені у вигляді числових параметрів, що відображають роботу серцево-судинної, дихальної та нервової систем. Усі ці ознаки формують багатовимірний вектор, який описує поточний стан користувача.

Особливістю сучасних методів є врахування часової динаміки емоцій [62, 65]. Оскільки емоційний стан змінюється у часі, важливо аналізувати не лише окремі значення ознак, а й їх еволюцію. Для цього використовуються рекурентні нейронні мережі, зокрема LSTM, які здатні зберігати інформацію про попередні стани та виявляти часові залежності. Такий підхід дозволяє моделювати інерційність емоційних реакцій та підвищує точність класифікації у реальних умовах.

На етапі класифікації застосовуються алгоритми машинного навчання або глибокого навчання, які на основі вхідного вектора ознак визначають емоційний стан користувача. Залежно від поставленої задачі це може бути віднесення до одного з наперед визначених класів або визначення координат у димензійному просторі. Для підвищення ефективності часто використовуються гібридні моделі, що поєднують різні типи нейронних мереж, наприклад згорткові мережі для обробки зображень і рекурентні для аналізу часових послідовностей.

Важливим елементом методу є постобробка результатів, яка включає згладжування, фільтрацію помилкових спрацьовувань та інтерпретацію

отриманих даних. Це дозволяє уникнути різких коливань у визначенні емоційного стану та забезпечує більш стабільну роботу системи. Крім того, результати можуть використовуватися для адаптації поведінки кіберфізичної системи, наприклад зміни інтерфейсу, регулювання навантаження або формування рекомендацій.

Таким чином, метод визначення емоційного стану є складною багаторівневою процедурою, що поєднує збір, обробку та аналіз різномірних даних із використанням сучасних алгоритмів штучного інтелекту. Його ефективність визначається якістю вхідних даних, правильністю вибору ознак та здатністю моделей враховувати часову динаміку емоцій.

### 3.2 Метод визначення емоційного стану людини за комп'ютером

Запропонований метод базується на комплексному аналізі різномірних даних, що надходять від периферійних пристроїв комп'ютера, та використанні гібридної архітектури нейронних мереж для досягнення достатньої точності розпізнавання у реальному часі.

Метод складається з наступних кроків:

1. Збір та первинна обробка даних.
2. Інтеграція та нормалізація ознак.
3. Гібридна нейромережева модель.
4. Отримання результату та інтерпретація.

Процес починається з першого кроку на фізичному рівні кіберфізичної системи, де за допомогою сенсорів (веб-камера, мікрофон, системні логи) формується вхідний потік даних. Система працює за принципом паралельного аналізу трьох основних модальностей:

Мовленнєвий сигнал перетворюється у вектор ознак за допомогою методу мел-кепстральних коефіцієнтів. Це дозволяє виділити тембральні та інтонаційні характеристики голосу, що відображають емоційне забарвлення.

Відеопотік обробляється згортковою нейронною мережею для виявлення

ключових точок обличчя та виділення просторових дескрипторів міміки.

Здійснюється аналіз логу системних подій (швидкість натискання клавіш, траєкторія руху курсора, частота помилок), що дає непряму інформацію про рівень стресу або втому користувача.

На другому кроці, отримані вектори ознак з різних каналів об'єднуються у єдиний мультимодальний вектор  $X$ . Оскільки вхідні дані надходять із різних джерел і мають різну фізичну природу (відео, аудіо, біометричні сигнали), вони характеризуються різною розмірністю, масштабами значень і форматами представлення. Тому перед об'єднанням у єдину модель необхідно виконати процедуру нормалізації, метою якої є приведення всіх ознак до уніфікованого вигляду.

На першому етапі здійснюється попередня обробка кожного типу даних. Для відеопотоку це може включати виділення області обличчя, масштабування зображення до фіксованого розміру та перетворення у числові ознаки (наприклад, координати ключових точок або вектор ознак нейронної мережі). Для аудіосигналу виконується фільтрація шуму, сегментація та перетворення у спектральні характеристики (наприклад, MFCC-коефіцієнти). Біометричні сигнали (пульс, провідність шкіри тощо) також очищуються від артефактів і переводяться у часові ряди або статистичні показники.

На наступному етапі всі отримані ознаки піддаються масштабуванню. Це може бути лінійна нормалізація (наприклад, приведення значень до діапазону  $[0;1]$ ) або стандартизація (перетворення до нульового середнього та одиничної дисперсії). Такий підхід дозволяє уникнути ситуації, коли ознаки з великими числовими значеннями домінують над іншими під час навчання моделі.

Додатково може виконуватися вирівнювання розмірності даних. Оскільки різні модальності можуть мати різну кількість ознак, застосовуються методи зменшення або узгодження розмірності (наприклад, усереднення, вибір ключових ознак або використання вбудованих представлень нейронних мереж), що дозволяє сформувати узгоджений вектор ознак фіксованої довжини.

Завершальним кроком є синхронізація даних у часі, особливо важлива для

мультимодальних систем. Відео, аудіо та біометричні сигнали повинні відповідати одному часовому інтервалу, щоб модель коректно інтерпретувала емоційний стан у конкретний момент.

У результаті нормалізації всі різномірні дані перетворюються у стандартизований, узгоджений набір числових ознак, придатний для подальшого об'єднання та обробки в рамках єдиної моделі машинного навчання.

На третьому кроці застосовується гібридна нейромережева модель. Ядром методу є комбінована архітектура, яка поєднує два підходи до аналізу даних для досягнення максимально точного результату.

Згорткові шари призначені для автоматичного виявлення візуальних ознак міміки та особливостей обличчя на кожному окремому кадрі.

Рекурентні шари аналізують послідовність цих ознак у часі. Це дозволяє системі враховувати динаміку зміни емоцій та пам'ятати попередні стани користувача. Такий підхід є критично важливим, оскільки емоції людини мають певну інерцію і не змінюються миттєво.

Поєднання цих технологій у єдиній архітектурі забезпечує стабільну роботу системи в реальному часі та дозволяє отримати більш обґрунтовану оцінку емоційного стану порівняно з аналізом лише статичних зображень.

На завершальному етапі роботи моделі відбувається формування результату класифікації та його інтерпретація. Після проходження всіх попередніх етапів обробки та аналізу мультимодальних даних модель формує вихідний вектор, який подається на функцію активації типу Softmax.

Ця функція забезпечує перетворення довільних числових значень у нормований розподіл ймовірностей, де кожен елемент вектора відповідає ймовірності належності поточного стану користувача до певного класу емоцій.

У результаті формується вектор вигляду 3.1:

$$P = \langle p_1, p_2, \dots, p_k \rangle, \quad (3.1)$$

де  $k$  – кількість емоційних класів, а сума всіх ймовірностей дорівнює одиниці.

Це дозволяє інтерпретувати вихід моделі не лише як категоріальне рішення, а як оцінку впевненості системи щодо кожної з можливих емоцій.

Після отримання ймовірнісного розподілу застосовуються механізми прийняття остаточного рішення. Найпростішим і найбільш поширеним є критерій  $\text{argmax}$ , який передбачає вибір тієї емоції, для якої відповідна ймовірність є максимальною.

Такий підхід забезпечує чітку та однозначну класифікацію, що є доцільним у задачах, де необхідно швидко прийняти рішення або активувати певну реакцію системи. Наприклад, якщо найбільшу ймовірність має стан стрес, система може автоматично зменшити інтенсивність інформаційного навантаження або запропонувати користувачу перерву.

Однак у реальних умовах емоційний стан людини рідко є дискретним і однозначним. Людина може одночасно відчувати кілька емоцій різної інтенсивності, що робить жорстку класифікацію недостатньо точною. У зв'язку з цим доцільним є використання підходів нечіткої логіки, які дозволяють описувати змішані або проміжні емоційні стани.

У межах такого підходу кожна емоція розглядається як нечітка множина з певною функцією належності, а отримані ймовірності інтерпретуються як ступені приналежності до цих множин.

Застосування нечіткої логіки дає змогу формувати більш гнучкі висновки, наприклад визначати, що користувач перебуває у стані помірного стресу з елементами втоми або легкого роздратування. Це, у свою чергу, дозволяє системі адаптувати свою поведінку більш делікатно і точно, наприклад змінювати інтерфейс, темп подачі інформації або характер взаємодії залежно від комплексного емоційного профілю.

Таким чином, поєднання ймовірнісного підходу ( $\text{softmax}$  та  $\text{argmax}$ ) і нечіткої логіки забезпечує як чіткість прийняття рішень, так і гнучкість інтерпретації результатів.

Це є особливо важливим для кіберфізичних систем, орієнтованих на людину, оскільки дозволяє більш адекватно враховувати складну природу емоційного стану користувача та підвищувати ефективність людино-машинної взаємодії.

Структурно-логічна схема методу визначення емоційного стану людини наведена на рисунку 3.2.

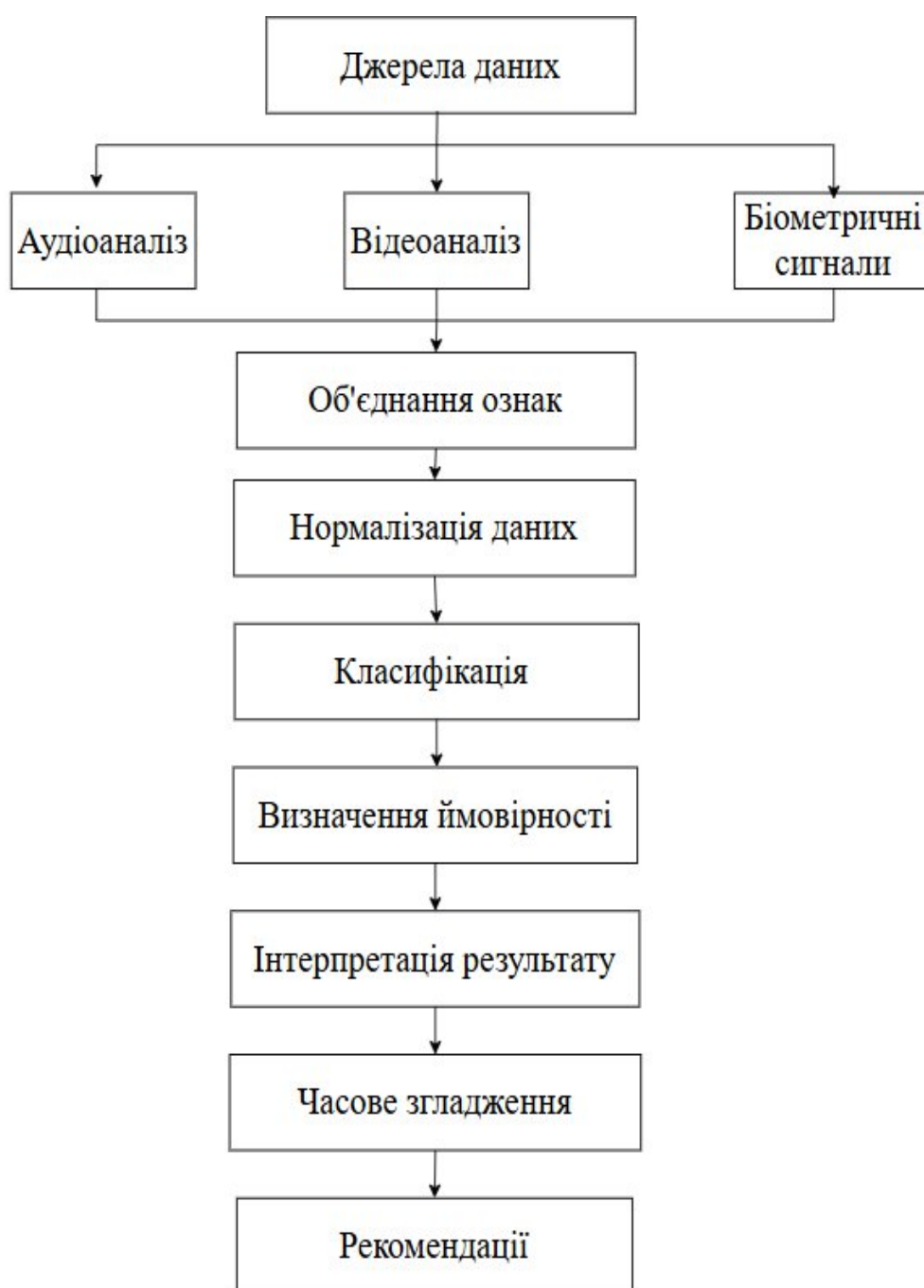


Рисунок 3.2 – Структурно-логічна схема методу визначення емоційного стану людини за комп'ютером

### 3.3 Алгоритмічне забезпечення кіберфізичної системи визначення емоційного стану людини за комп'ютером

Алгоритмічне забезпечення кіберфізичної системи визначення емоційного стану людини є ключовим компонентом, що забезпечує функціонування всієї системи як єдиного інтелектуального комплексу. Воно охоплює сукупність методів, алгоритмів і процедур, спрямованих на збір, обробку, аналіз та інтерпретацію мультимодальних даних з метою отримання достовірної оцінки емоційного стану користувача у реальному часі.

Запропонована система базується на принципах мультимодальності, адаптивності та безперервного аналізу, що дозволяє враховувати різні канали сприйняття інформації та забезпечувати стійкість до зовнішніх впливів. У загальному вигляді алгоритмічне забезпечення можна представити як послідовність взаємопов'язаних етапів: збір даних, попередня обробка, виділення ознак, нормалізація, інтеграція (злиття) даних, класифікація та інтерпретація результатів.

Алгоритм роботи кіберфізичної системи визначення емоційного стану людини за компютером включає в себе наступні процеси: збір та синхронізація даних, отриманих з різних сенсорів; застосування навчених нейромережових моделей для виявлення закономірностей у поведінці; генерацію вихідного сигналу про емоційний стан для подальшого використання у прикладних сервісах (освіта, медицина, ігри).

#### 3.3.1 Алгоритм збору та синхронізації даних

Алгоритм збору та синхронізації даних є базовим етапом функціонування

кіберфізичної системи визначення емоційного стану користувача, оскільки саме на цьому рівні формується первинна інформаційна база для подальшого аналізу. Від якості реалізації цього етапу безпосередньо залежить точність роботи всієї системи, адже помилки або розбіжності між потоками даних можуть призвести до некоректної інтерпретації емоційного стану. Блок-схема алгоритму представлена на рисунку 3.3.

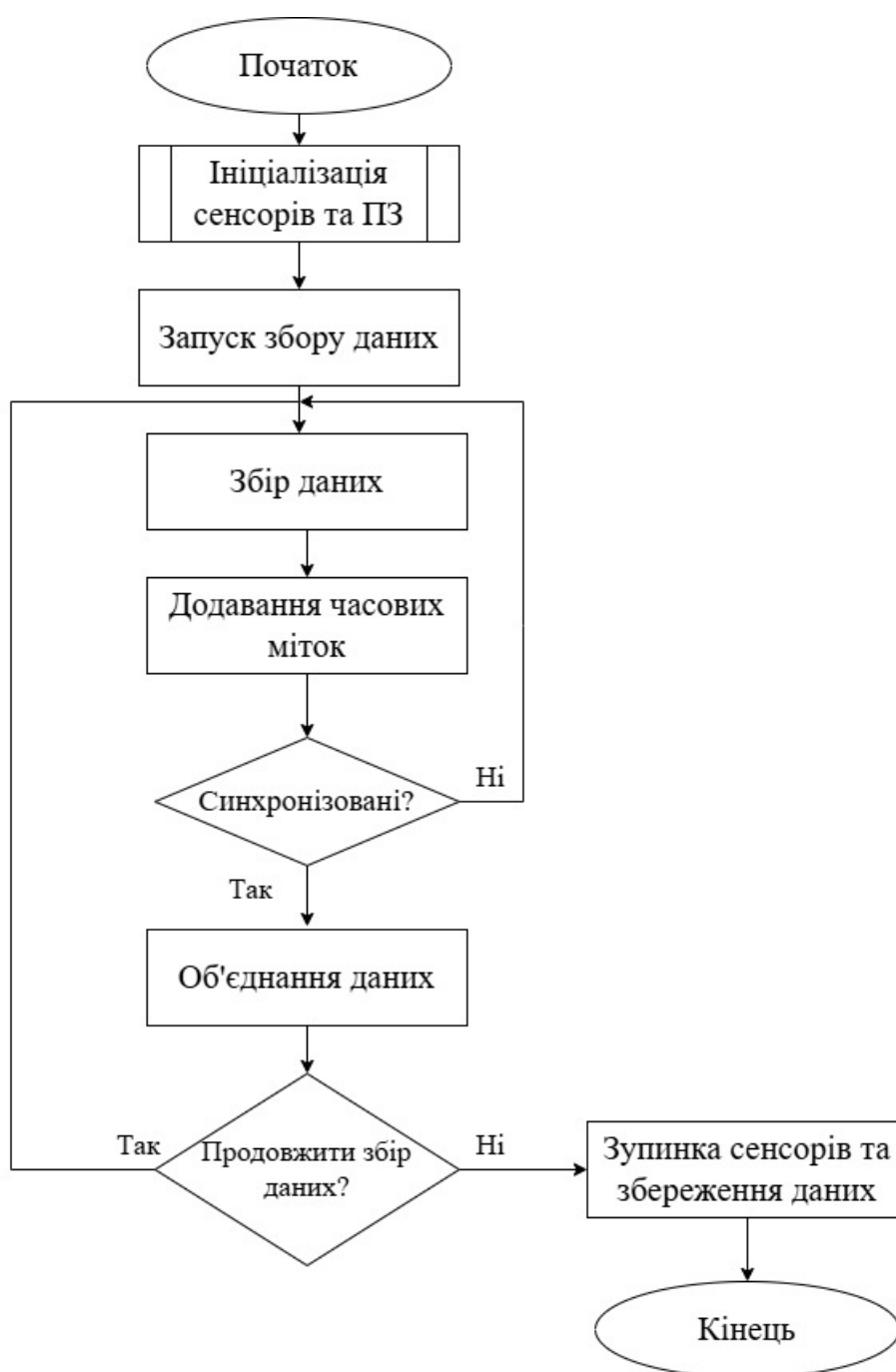


Рисунок 3.4 – Блок-схема алгоритму попередньої обробки даних

На початковому кроці здійснюється ініціалізація всіх джерел даних. Система активує веб-камеру для отримання відеопотоку, мікрофон для запису аудіосигналу, а також, за наявності, підключає біометричні сенсори, які можуть вимірювати такі показники, як частота серцевих скорочень, варіабельність серцевого ритму або електрошкірна провідність. Кожне з цих джерел має власну частоту дискретизації та формат представлення даних, що створює необхідність їх подальшого узгодження.

Після ініціалізації починається безперервний процес збору даних. Відеокамера генерує послідовність кадрів із певною частотою (наприклад, 25 – 30 кадрів за секунду), аудіосигнал надходить у вигляді безперервного потоку з високою частотою дискретизації (наприклад, 44,1 кГц), а біометричні сенсори передають значення через певні інтервали часу. Оскільки ці потоки є асинхронними за своєю природою, їх неможливо безпосередньо використовувати для спільного аналізу без попередньої обробки [71].

Для вирішення цієї проблеми на наступному етапі використовується механізм буферизації. Кожен потік даних поміщається у відповідний буфер, тобто тимчасове сховище, яке накопичує дані за певний проміжок часу.

Буферизація дозволяє згладити нерівномірність надходження даних, компенсувати затримки та підготувати інформацію до синхронізації. Наприклад, відеобуфер може зберігати кілька останніх кадрів, аудіо буфер може зберігати короткий фрагмент сигналу, а біометричний буфер зберігає набір останніх вимірювань.

Ключовим етапом є синхронізація даних, яка забезпечує їх часову узгодженість. Для цього кожному елементу даних (кадру відео, аудіофрагменту або біометричному вимірюванню) присвоюється часова мітка. Далі система виконує вирівнювання потоків, формуючи єдині часові вікна, у межах яких об'єднуються відповідні дані з усіх модальностей. Наприклад, для інтервалу тривалістю 1 секунда система підбирає всі відеокадри, аудіофрагменти та біометричні значення, що відповідають цьому часовому проміжку.

У випадку, якщо дані з різних джерел не співпадають за частотою, застосовуються методи інтерполяції або агрегування. Так, аудіосигнал може бути узагальнений у вигляді спектральних ознак за певний інтервал часу, відеодані представлені ключовим кадром або усередненими характеристиками, а біометричні показники згладжені або інтерпольовані до потрібної частоти. Це дозволяє сформувати узгоджені за часом набори даних, придатні для подальшого аналізу.

Важливою особливістю алгоритму є забезпечення мінімальної затримки обробки, оскільки система повинна працювати у режимі реального часу. Для цього використовуються ковзні часові вікна, які постійно оновлюються при надходженні нових даних.

Такий підхід дозволяє одночасно зберігати контекст попередніх станів і оперативно реагувати на зміни емоційного стану користувача.

У результаті виконання даного етапу формується синхронізований мультимодальний набір даних, у якому всі компоненти відповідають одному й тому самому часовому інтервалу. Це створює основу для коректного виділення ознак та подальшої класифікації емоційного стану.

### 3.3.2 Алгоритм попередньої обробки даних

Алгоритм попередньої обробки є важливим етапом у функціонуванні кіберфізичної системи визначення емоційного стану, оскільки саме на цьому рівні здійснюється очищення, стабілізація та стандартизація вхідних даних перед їх подальшим аналізом.

Враховуючи мультимодальний характер системи, кожен тип даних потребує застосування спеціалізованих методів обробки, що враховують їх фізичну природу та особливості формування.

Для відеоданих попередня обробка починається з етапу детекції обличчя. На цьому кроці система аналізує кожен кадр відеопотоку з метою локалізації області,

яка містить обличчя користувача. Це може здійснюватися за допомогою класичних алгоритмів (наприклад, каскадів Хаара) або сучасних глибоких нейронних мереж. Результатом є координати прямокутної області, в межах якої знаходиться обличчя. Виділення цієї області дозволяє зосередити подальший аналіз лише на релевантній частині зображення та зменшити вплив фону [72].

Наступним кроком є вирівнювання обличчя. Оскільки користувач може змінювати положення голови, нахилити її або повертати, важливо привести зображення обличчя до стандартного вигляду.

Для цього використовуються ключові точки обличчя (очі, ніс, рот), на основі яких виконується геометричне перетворення, тобто масштабування, обертання та зсув. У результаті обличчя розташовується у фіксованій позиції, що значно підвищує стабільність подальшого виділення ознак.

Завершальним етапом обробки відео є нормалізація освітлення. Освітлення може суттєво змінюватися залежно від умов середовища, що впливає на якість зображення та точність розпізнавання.

Для усунення цього впливу застосовуються методи вирівнювання гістограми, корекції контрасту або перетворення зображення у відтінки сірого. Це дозволяє зробити характеристики обличчя більш стабільними та незалежними від зовнішніх умов.

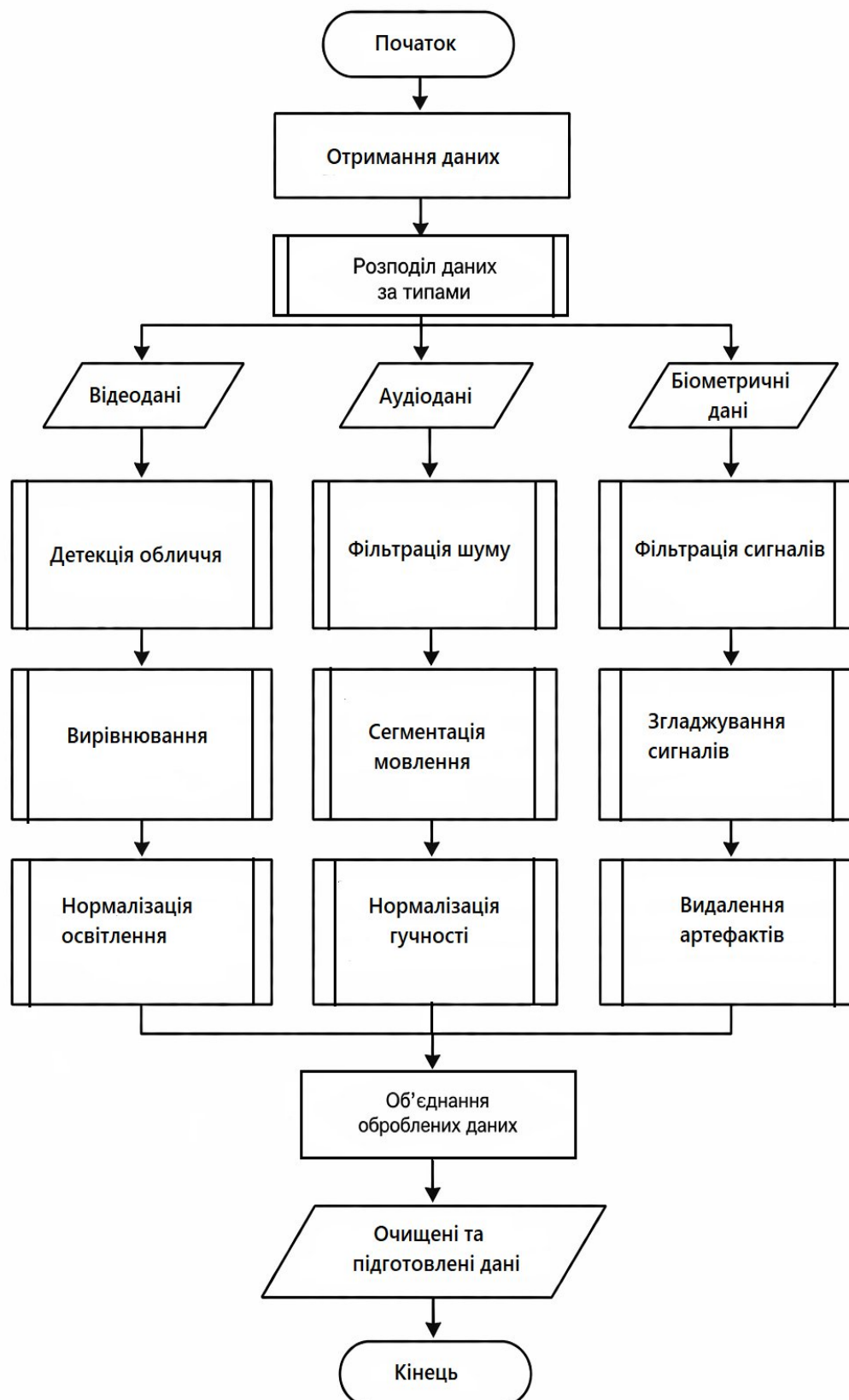
Для аудіосигналу попередня обробка також складається з кількох послідовних етапів. Першим з них є фільтрація шуму, яка спрямована на видалення сторонніх звуків, таких як фоновий шум, перешкоди або електричні наводки.

Для цього можуть використовуватися цифрові фільтри (наприклад, смугові або адаптивні фільтри), які виділяють корисний мовний сигнал. Це особливо важливо в умовах реального середовища, де рівень шуму може бути непередбачуваним.

Після очищення сигналу виконується сегментація мовлення. На цьому етапі система визначає ділянки аудіопотоку, що містять мовлення, і відокремлює їх від пауз або тиші. Це дозволяє зосередити подальший аналіз лише на інформативних

фрагментах, що містять емоційно забарвлені характеристики, такі як інтонація, темп мовлення або зміни висоти тону.

Блок-схема алгоритму представлена на рисунку 3.4.



### Рисунок 3.4 – Блок-схема алгоритму попередньої обробки даних

Завершальним етапом є нормалізація гучності. Оскільки рівень сигналу може змінюватися залежно від відстані до мікрофона або гучності голосу користувача, виконується приведення амплітуди сигналу до стандартного рівня. Це забезпечує однакові умови для подальшого виділення ознак та запобігає домінуванню гучних сигналів над тихими.

Попередня обробка біометричних даних має свої особливості, оскільки такі сигнали часто є зашумленими та містять артефакти, пов'язані з рухами користувача або нестабільністю сенсорів.

Першим етапом є фільтрація сигналів, яка дозволяє усунути високочастотні або низькочастотні перешкоди, що не несуть корисної інформації. Наприклад, при аналізі серцевого ритму можуть використовуватися згладжувальні або смугові фільтри.

Наступним кроком є згладжування сигналів, яке спрямоване на зменшення випадкових коливань і стабілізацію даних. Це може реалізовуватися за допомогою ковзного середнього або інших методів апроксимації. У результаті формується більш плавний сигнал, який краще відображає реальні фізіологічні процеси.

Окрему увагу приділяють видаленню артефактів, тобто різких стрибків або аномальних значень, що можуть виникати через збої сенсорів або рухи користувача. Такі значення ідентифікуються та або коригуються, або виключаються з подальшого аналізу. Це дозволяє підвищити достовірність біометричних показників.

У підсумку, етап попередньої обробки забезпечує перетворення сирих, неоднорідних і зашумлених даних у стабільні, очищені та стандартизовані сигнали.

Це створює необхідні умови для ефективного виділення ознак і подальшої класифікації емоційного стану користувача, а також підвищує загальну надійність роботи кіберфізичної системи.

Отже, підготовлені таким чином дані є оптимальним вхідним матеріалом для наступних етапів обробки, що забезпечує підвищення точності та стійкості визначення емоційного стану користувача.

### 3.3.3 Алгоритм виділення ознак

Етап виділення ознак є критично важливою складовою методу визначення емоційного стану користувача, оскільки саме на цьому рівні відбувається перетворення сирих мультимодальних даних у формалізоване числове представлення, придатне для подальшої обробки, аналізу та класифікації. Основною метою даного етапу є формування інформативного вектора ознак, який максимально точно відображає емоційні прояви користувача в різних модальностях.

У межах обробки відеоданих здійснюється аналіз мімічних та геометричних характеристик обличчя. Спочатку визначаються ключові точки обличчя, які відповідають положенню основних анатомічних елементів, таких як очі, брови, ніс, рот та контур обличчя. На основі координат цих точок формуються геометричні ознаки, що дозволяють описувати мімічні зміни, зокрема рухи брів, відкриття або закриття рота, зміни кута губ та інші характерні вирази обличчя, які безпосередньо пов'язані з емоційними станами. Додатково зображення обличчя обробляються за допомогою згорткових нейронних мереж, у результаті чого отримуються глибокі ознаки, що відображають високорівневі абстрактні характеристики зображення, включаючи текстурні особливості, мікровирази та складні просторові патерни, які неможливо ефективно описати вручну.

У випадку аудіоданих аналіз мовлення спрямований на вилучення спектральних та просодичних характеристик голосу. Основним набором ознак є коефіцієнти MFCC, які забезпечують компактне представлення спектральної інформації мовного сигналу з урахуванням особливостей людського слухового сприйняття. Крім того, додатково враховуються такі параметри, як тон голосу, енергія сигналу та тембр. Тон відображає висоту голосу та його зміни під час мовлення, енергія характеризує інтенсивність та емоційне збудження мовця, а

тембр описує індивідуальні спектральні властивості голосу, що також можуть змінюватися залежно від емоційного стану.

Для біометричних даних здійснюється аналіз фізіологічних сигналів користувача, які можуть відображати його внутрішній емоційний стан. На цьому рівні обчислюються середні значення сигналів, що дозволяє визначити базовий рівень фізіологічної активності, наприклад середній пульс або середній рівень активності протягом певного інтервалу часу. Додатково аналізуються варіації сигналів, які характеризують динамічні зміни фізіологічних показників. Такі коливання можуть бути індикаторами стресу, збудження або інших емоційно значущих станів.

Після обробки всіх модальностей виконується об'єднання отриманих характеристик у єдиний багатомодальний вектор ознак. Цей вектор є узагальненим представленням емоційного стану користувача та використовується на подальших етапах методології, зокрема для класифікації та інтерпретації емоційного стану.

Таким чином, етап виділення ознак забезпечує перехід від первинних даних до структурованого числового опису, який є основою для ефективного функціонування всієї системи.

### 3.4 Застосування нейромережевих моделей для виявлення закономірностей у поведінці користувача

Застосування навчених нейромережевих моделей є ключовим етапом у процесі визначення емоційного стану користувача, оскільки саме на цьому рівні відбувається виявлення прихованих закономірностей у поведінкових, аудіальних та візуальних даних.

На відміну від традиційних методів аналізу, нейронні мережі здатні автоматично навчатися складним нелінійним залежностям між вхідними ознаками та відповідними емоційними станами, що значно підвищує точність і адаптивність

системи.

Структурна схема гібридної нейромережевої моделі представлена на рисунку 3.5.

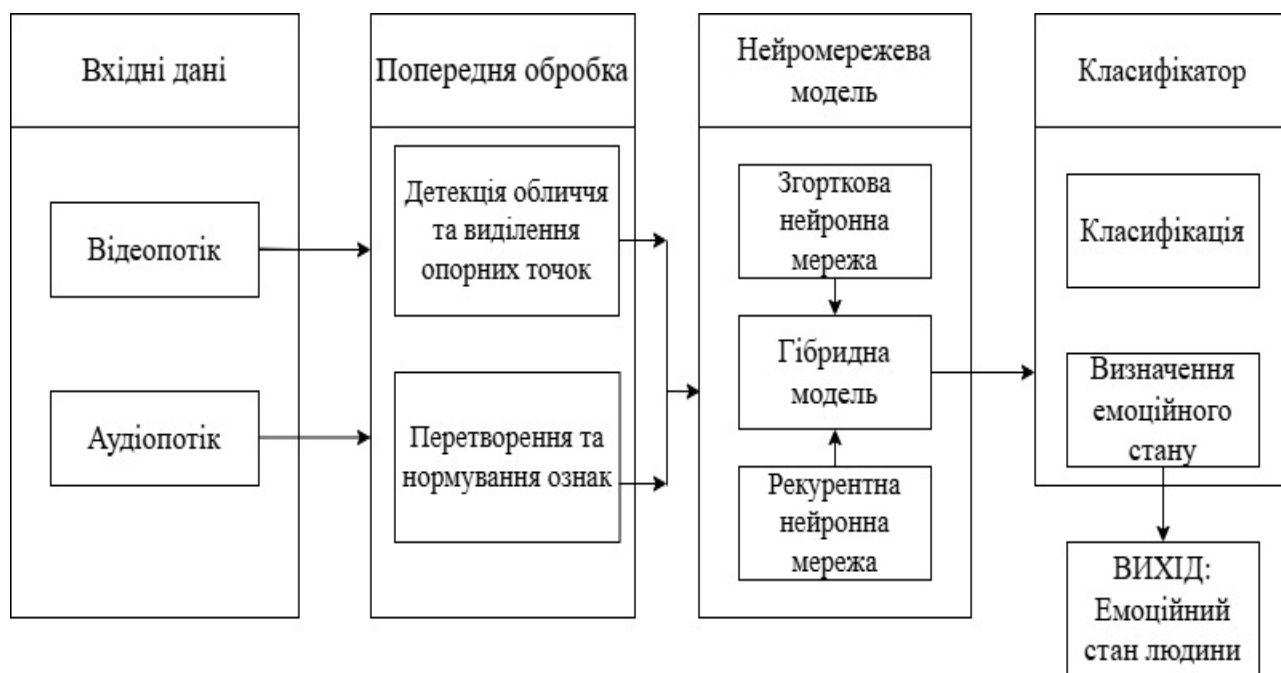


Рисунок 3.5 – Структурна схема гібридної нейромережевої моделі

На вхід нейромережевої моделі подається попередньо сформований та нормалізований вектор ознак, який містить інтегровану інформацію з різних модальностей. Цей вектор відображає як статичні характеристики (наприклад, форма обличчя чи спектральні властивості голосу), так і динамічні аспекти (зміни міміки, інтонації або фізіологічних показників у часі).

У процесі обробки даних нейронна мережа виконує багаторівневе перетворення ознак. На початкових шарах виділяються базові патерни, такі як локальні зміни сигналів або прості візуальні структури. На більш глибоких рівнях формуються узагальнені представлення, які відповідають складним поведінковим і емоційним характеристикам.

Використання згорткових шарів дозволяє ефективно обробляти просторові залежності у відеоданих, тоді як рекурентні мережі забезпечують врахування часової динаміки змін поведінки та емоційного стану.

У результаті роботи моделі формується внутрішнє латентне представлення

даних, яке відображає приховані закономірності у поведінці користувача. На основі цього представлення здійснюється подальша класифікація або оцінка емоційного стану.

Важливою перевагою такого підходу є здатність моделі адаптуватися до індивідуальних особливостей користувача, а також працювати в умовах шумів та неповноти даних.

Додатково важливим аспектом є процес навчання нейромережевих моделей, який здійснюється на основі великих розмічених наборів даних, що містять приклади різних емоційних станів. Під час навчання модель оптимізує внутрішні вагові коефіцієнти шляхом мінімізації функції втрат, що дозволяє зменшити похибку між прогнозованими та фактичними значеннями. Використання методів зворотного поширення похибки забезпечує поступове уточнення параметрів моделі та підвищення її узагальнювальної здатності.

Важливу роль у підвищенні ефективності нейромережевих моделей відіграють методи регуляризації, які запобігають перенавчанню та покращують стабільність роботи системи. Зокрема, застосування dropout-шарів, L1/L2 регуляризації та нормалізації пакетів дозволяє зменшити чутливість моделі до шумових даних і підвищити її здатність працювати з новими, раніше не баченими прикладами. Це особливо важливо у задачах емоційного аналізу, де вхідні дані часто є неповними або неоднорідними.

Окрему увагу слід приділити багатомодальному підходу до обробки інформації, коли нейронна мережа одночасно аналізує декілька типів даних, як відео, аудіо та фізіологічні сигнали.

Така інтеграція дозволяє отримати більш повну картину емоційного стану користувача, оскільки різні модальності взаємодоповнюють одна одну. Наприклад, мімічні прояви можуть підтверджувати або уточнювати результати, отримані на основі голосових характеристик.

У процесі експлуатації системи важливою є також адаптація моделі до індивідуальних особливостей користувача. Це може реалізовуватися через механізми донавчання, що дозволяє використовувати попередньо навчені моделі

як базу та поступово підлаштовувати їх під конкретного користувача. Такий підхід підвищує точність розпізнавання емоцій та зменшує похибки, пов'язані з міжособистісними відмінностями у вираженні емоцій.

Крім того, результати роботи нейромережових моделей можуть бути інтерпретовані для подальшого аналізу поведінкових патернів користувача. Це дозволяє не лише визначати поточний емоційний стан, але й виявляти довгострокові тенденції, стресові періоди або аномальні зміни поведінки. Отримана інформація може бути використана для підвищення ефективності кіберфізичної системи, зокрема для адаптації інтерфейсу або формування рекомендацій користувачу.

Таким чином, застосування навчених нейромережових моделей забезпечує ефективне виявлення складних багатовимірних залежностей у поведінці користувача, що є основою для точного та надійного визначення його емоційного стану в кіберфізичній системі.

### 3.5 Методи забезпечення інтерпретованості результатів класифікації емоцій

Сучасний розвиток кіберфізичних систем вимагає не лише високої точності розпізнавання станів користувача, але й прозорості механізмів прийняття рішень. У контексті визначення емоційного стану це стає критично важливим, оскільки помилкова інтерпретація почуттів людини може призвести до невірної адаптації інтерфейсу або некоректної реакції системи у критичних ситуаціях, наприклад, в освітніх процесах чи медицині.

Забезпечення інтерпретованості дозволяє перетворити чорну скриньку нейронної мережі на зрозумілу модель, де кожне рішення обґрунтоване конкретними фізіологічними чи поведінковими ознаками користувача.

Інтерпретованість результатів класифікації базується на багаторівневому підході, що охоплює аналіз візуальних, аудіальних та часових компонентів емоційного прояву.

Візуальний канал є найбільш інформативним джерелом даних про емоції.

Для забезпечення прозорості рішень згорткової нейронної мережі застосовується метод зворотного зв'язку з системою кодування лицевих рухів. Коли модель класифікує зображення як стрес або гнів, механізм інтерпретації виконує декомпозицію вихідного сигналу для визначення найбільш вагомих зон активації на обличчі користувача.

Інтерпретованість на цьому рівні реалізується через наступні методи.

1. Аналіз опорних точок. Система визначає геометричні відхилення положення брів, очей та губ відносно нейтрального стану користувача. Наприклад, звуження очей та напруження губ математично інтерпретуються як ознаки високої концентрації або роздратування.

2. Візуалізацію карт активації. Використання методів на кшталт Grad-CAM дозволяє підсвічувати ділянки кадру, які мали найбільший вплив на фінальний висновок мережі. Якщо система ідентифікує емоцію радості, карта активації має чітко вказувати на зону виличних м'язів та куточків рота.

Аудіоаналіз емоцій часто сприймається як складний для розуміння процес через абстрактну природу спектральних ознак. У запропонованому методі інтерпретованість аудіоканалу досягається шляхом кореляції мел-частотних кепстральних коефіцієнтів з фізичними параметрами голосу.

Кожне рішення системи щодо емоційного забарвлення мовлення супроводжується аналізом наступних показників.

1. Фундаментальна частота. Підвищення тону голосу інтерпретується як маркер зростання емоційного збудження (Arousal), що характерно для радості або страху.

2. Енергетичний спектр. Зміна гучності та інтенсивності вимови певних фреймів пояснює впевненість або неупевненість користувача.

3. Ритмічні характеристики. Збільшення тривалості пауз між словами у поєднанні з низьким тембром інтерпретується системою як ознака втоми або пригніченості.

Оскільки емоції людини є динамічними процесами, важливим є як попередній стан впливає на поточне рішення. Використання рекурентних шарів

LSTM дозволяє системі пам'ятати контекст взаємодії.

Інтерпретованість тут забезпечується через механізм уваги, який виділяє часові відрізки, що стали вирішальними для зміни класифікації. Це дозволяє уникнути помилкових спрацювань при випадкових мімічних рухах або сторонніх звуках, пояснюючи рішення тривалістю та стабільністю виявленого патерну поведінки.

Кінцева інтерпретація результату в КФС базується на інтегрованому векторі ознак. Для кожної визначеної емоції система генерує ймовірнісний розподіл, який дозволяє оцінити ступінь впевненості у прийнятому рішенні.

Для підвищення так званої людиноорієнтованості методу використовується логіка нечіткого виведення. Замість сухої класифікації як то стан наразі це гнів, система формує опис по типу: «Виявлено ознаки високого стресу (впевненість 75%) на основі прискороного пульсу та характерного напруження мімічних м'язів чола, що триває протягом останніх 15 секунд».

Такий підхід до інтерпретованості забезпечує довіру користувача, тому що оператор або студент розуміє, чому система пропонує перерву або змінює режим роботи. Забезпечується можливість верифікації, отже розробник може відстежити, яка саме модальність (відео, аудіо чи біометрія) дала збій у разі некоректної роботи. А також, система стає прозорою у використанні біометричних даних, надаючи пояснення, які показники використовуються для аналізу в конкретний момент часу.

Проектування та впровадження кіберфізичних систем, здатних розпізнавати емоційний стан людини, неминуче стикається з низкою етичних викликів та необхідністю суворого дотримання стандартів інформаційної безпеки.

Оскільки робота таких систем базується на аналізі біометричних параметрів таких, як відеозображення обличчя, характеристик голосу та фізіологічних показників, то питання захисту приватності стає основним принципом архітектури методу.

Відповідно до сучасних європейських стандартів, зокрема Загального регламенту про захист даних, біометричні дані належать до особливих категорій

персональних даних, що вимагає впровадження посиленних заходів контролю та мінімізації ризиків.

Основоположним принципом забезпечення конфіденційності у таких методах може бути концепція «Privacy by Design», що передбачає інтеграцію захисних механізмів безпосередньо у процес обробки інформації [60].

Ключовим технічним рішенням у цьому контексті є повна деідентифікація даних на ранніх етапах функціонування системи. На відміну від традиційних систем моніторингу, розроблений метод не передбачає довготривалого зберігання сирих сенсорних даних, таких як автентичні відеокадри або аудіозаписи мовлення користувача. Замість цього система працює у режимі реального часу, де первинний сигнал трансформується у математичне представлення, а саме вектор ознак, після чого вихідний файл негайно видаляється з оперативної пам'яті.

Процес візуальної деідентифікації реалізується через перетворення зображення обличчя у набір геометричних координат ключових точок. Отриманий числовий масив описує виключно динаміку м'язів, необхідну для класифікації емоцій, проте він позбавлений текстурних характеристик шкіри, кольору очей та інших індивідуальних особливостей, що дозволяють ідентифікувати особу. Аналогічний підхід застосовується і до акустичного каналу: замість збереження запису голосу система вилучає мел-частотні кепстральні коефіцієнти, які відображають спектральні властивості та емоційну інтенсивність, але не дозволяють відновити зміст мовлення або унікальний відбиток голосу користувача.

Такий підхід забезпечує незворотність перетворення даних. Навіть у випадку несанкціонованого доступу до бази даних системи або перехоплення інформаційного потоку, зловмисник отримає лише масиви абстрактних числових значень [76]. Відновити з цих векторів ознак оригінальне зображення обличчя або аудіозапис голосу технічно неможливо, що зводить ризик деанонімізації користувача до мінімуму. Це дозволяє системі ефективно функціонувати в умовах обмежених ресурсів та забезпечує високий рівень довіри з боку кінцевих користувачів.

Окрім технічних заходів деідентифікації, етичний складник методу включає забезпечення прозорості взаємодії. Користувач повинен бути повністю поінформований про мету моніторингу, типи даних, що збираються, та алгоритми їх обробки. Система передбачає механізм явного надання згоди та можливість миттєвого припинення збору даних у будь-який момент.

Таким чином, запропонований метод дозволяє тримати баланс між необхідністю аналізу емоційного стану для підвищення продуктивності праці чи навчання та правом людини на приватність і захист персональної інформації у цифровому просторі.

### 3.6 Висновки

У третьому розділі було розроблено та тобґрунтовано метод визначення емоційного стану людини за комп'ютером, що базується на комплексному аналізі мультимодальних даних. Основною ідеєю запропонованого підходу є поєднання різних каналів інформації, таких як відеопотік для аналізу міміки, аудіосигнал для оцінки характеристик голосу та біометричні дані для відстеження поведінкових особливостей користувача. Це дозволило створити систему, яка здатна компенсувати недоліки окремих джерел даних і забезпечувати достатньо високу точність розпізнавання емоцій у реальних умовах експлуатації.

Для забезпечення роботи методу було розроблено відповідне алгоритмічне забезпечення, яке охоплює всі етапи від збору даних до отримання фінального результату.

Важливим етапом стала розробка алгоритмів синхронізації та попередньої обробки, які дозволяють узгоджувати різнорідні сигнали в часі та очищувати їх від шумів. Процедура нормалізації ознак забезпечила приведення даних різної фізичної природи до уніфікованого числового формату, що є критично важливим для їх подальшої коректної обробки нейромережевими моделями.

Ядром розробленого методу виступає гібридна нейромережева архітектура, яка поєднує можливості згорткових шарів для виділення візуальних ознак

обличчя та рекурентних шарів типу LSTM для аналізу часових послідовностей . Таке поєднання дозволяє системі враховувати не лише статичні прояви емоцій на окремих кадрах, але й їхню динаміку та інерційність у часі. Це забезпечує стабільність класифікації та дозволяє моделювати реальні емоційні процеси, які мають певну тривалість і контекст попередніх станів.

Окрему увагу в розділі було приділено питанням інтерпретованості прийнятих рішень та етичним аспектам використання технології. Запропоновані методи пояснюваності дозволяють обґрунтувати висновки системи через конкретні фізіологічні показники, як-от зміни міміки чи тембру голосу, що підвищує рівень довіри користувача. Водночас впровадження механізмів деідентифікації даних та дотримання стандартів конфіденційності дозволять забезпечити захист приватності згідно з європейськими нормами GDPR, роблячи систему безпечною та соціально відповідальною.

## 4 ПРОЕКТУВАННЯ КІБЕРФІЗИЧНОЇ СИСТЕМИ ВИЗНАЧЕННЯ ЕМОЦІЙНОГО СТАНУ ЛЮДИНИ

### 4.1 Вибір типу архітектури та проектування модулів системи

При виборі типу архітектури варто розглянути принципи побудови спеціалізованих кіберфізичних систем, які працюють у режимі замкненого циклу.

Такий підхід передбачає безперервний процес збору даних, їх обробки, аналізу, прийняття рішень і формування керуючих впливів або рекомендацій, які, у свою чергу, впливають на поведінку користувача або стан системи. Це забезпечує адаптивність і можливість роботи в реальному часі.

З архітектурної точки зору подібні системи будуються за модульним принципом і, як правило, включають такі основні рівні: рівень збору даних (сенсори), рівень попередньої обробки сигналів, аналітичний рівень (моделі машинного навчання), а також рівень взаємодії з користувачем або зовнішніми системами.

Важливою особливістю є чітке розмежування функціональних модулів і використання стандартизованих інтерфейсів між ними, що спрощує масштабування, модернізацію та інтеграцію системи.

Прикладом реалізації таких архітектур є системи моніторингу водія, що використовуються в автомобілях таких брендів, як Volvo та BMW. У цих системах реалізується багаторівнева архітектура, де сенсорний рівень включає камери та інші датчики, рівень обробки здійснює аналіз зображень і поведінкових характеристик, а аналітичний рівень визначає стан водія на основі моделей штучного інтелекту [61, 62].

Отримані результати використовуються для формування попереджень або автоматичного втручання у керування транспортним засобом.

У медичних застосуваннях кіберфізичні системи використовуються для моніторингу стану пацієнтів, оцінки рівня стресу та підтримки прийняття рішень лікарями [64].

Архітектурно такі системи орієнтовані на обробку потокових даних у

реальному часі, інтеграцію з медичними інформаційними системами та забезпечення високої надійності й точності. Важливою особливістю проектування є забезпечення відмовостійкості, захисту даних і можливості роботи з неповними або зашумленими сигналами.

Окрім цього, подібні архітектурні підходи активно застосовуються у сфері освіти, зокрема в інтелектуальних навчальних системах. У таких системах здійснюється аналіз емоційного стану студента для адаптації навчального контенту, темпу подачі матеріалу та рівня складності завдань [65].

Архітектура передбачає інтеграцію модулів аналізу поведінки, систем рекомендацій і зворотного зв'язку, що дозволяє підвищити ефективність навчального процесу.

Ще однією важливою галуззю є маркетинг і клієнтська аналітика, де кіберфізичні системи використовуються для аналізу емоційної реакції споживачів на продукти або рекламні матеріали [66].

У таких системах архітектура орієнтована на обробку великих обсягів мультимодальних даних, інтеграцію з CRM-системами та забезпечення масштабованості. Особливу роль відіграє модуль аналітики, який дозволяє виявляти поведінкові патерни та формувати персоналізовані рекомендації.

Таким чином, вибір архітектури кіберфізичної системи визначається вимогами до обробки даних, швидкодії, надійності та можливостей інтеграції.

Модульний підхід, використання багаторівневої структури та впровадження інтелектуальних аналітичних компонентів забезпечують гнучкість системи та її ефективне функціонування в різних прикладних галузях.

Ще однією перспективною галуззю застосування кіберфізичних систем є сфера безпеки та відеоспостереження. У таких системах аналіз емоційного стану та поведінки людини використовується для виявлення потенційно небезпечних ситуацій, підозрілої активності або аномальної поведінки в громадських місцях [67].

Архітектурно ці системи орієнтовані на обробку відеопотоків у реальному часі, інтеграцію з мережами камер спостереження та централізованими системами

управління безпекою. Важливою особливістю проектування є забезпечення високої швидкодії, масштабованості та здатності працювати з великими обсягами даних. Крім того, значна увага приділяється точності моделей аналізу та мінімізації хибних спрацювань, що є критично важливим для практичного використання таких систем.

Для систематизації розглянутих підходів доцільно провести їх порівняння за ключовими характеристиками.

Порівняльний аналіз існуючих рішень наведено в таблиці 4.1.

Таблиця 4.1 – Порівняльний аналіз існуючих рішень проектування спеціалізованих КФС

Рішення	Модальність	Тип доступу	Основна перевага	Обмеження
OpenFace 2.0	Відео	Open Source	Детекція Action Units, гнучкість	Чутливість до умов зйомки
DeepFace	Відео	Open Source	Висока точність, готові моделі	Обмежена інтерпретованість
Affectiva	Відео + Аудіо	Комерційний SDK	Аналіз когнітивного стану	Закритість алгоритмів
Azure Face API	Відео	Cloud API	Масштабованість, інтеграція	Залежність від хмари
DMS системи	Відео + сенсори	Вбудовані рішення	Робота в реальному часі	Вузька спеціалізація
Розроблена система	Відео + Аудіо + LSTM	КФС (прототип)	Аналіз часової динаміки	Потребує навчання

Таким чином, вибір архітектури кіберфізичної системи визначається вимогами до обробки даних, швидкодії, надійності та можливостей інтеграції. Модульний підхід, використання багаторівневої структури та впровадження

інтелектуальних аналітичних компонентів забезпечують гнучкість системи та її ефективне функціонування в різних прикладних галузях.

Запропонована кіберфізична система розпізнавання емоцій людини за комп'ютером є складною інтелектуальною системою, яка поєднує фізичний рівень збору даних, рівень обробки сигналів та програмно-аналітичний рівень машинного навчання.

Основною метою системи є автоматичне визначення емоційного стану користувача на основі мультимодальних даних, отриманих у режимі реального часу.

Архітектура системи побудована за принципом багаторівневої обробки інформації, що дозволяє забезпечити високу точність та стійкість до шумів у вхідних даних.

Система включає три основні інформаційні канали: відео (міміка обличчя), аудіо (мовленнєві характеристики) та біометричні показники (фізіологічні реакції організму).

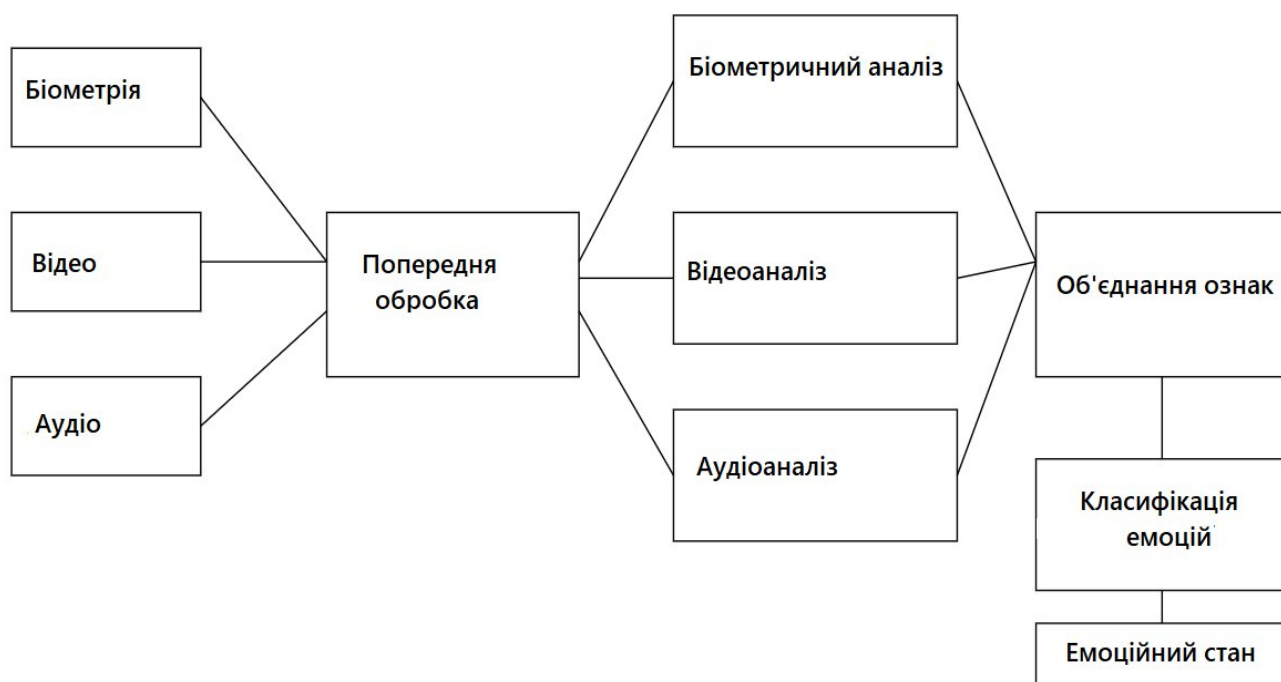


Рисунок 4.1 – Структурна схема кіберфізичної системи розпізнавання емоцій людини за комп'ютером

На першому рівні відбувається захоплення даних за допомогою периферійних пристроїв: веб-камери, мікрофона та сенсорів. Далі дані проходять етап попередньої обробки, який включає фільтрацію шумів, нормалізацію сигналів та синхронізацію різних потоків інформації за часом.

Після цього система розгалужується на три паралельні підсистеми аналізу, кожна з яких відповідає за окремий тип даних. Результати цих підсистем об'єднуються у спільний вектор ознак, який подається на класифікаційну модель.

Фінальним етапом є визначення емоційного стану користувача, яке відображається у вигляді одного з базових класів емоцій або у вигляді ймовірнісного розподілу.

Модуль відеоаналізу є одним із ключових компонентів системи, оскільки міміка обличчя містить найбільш безпосередні прояви емоційного стану людини. Основним завданням цього модуля є виявлення обличчя на відеопотоці та подальше визначення емоцій за його виразом.

На першому етапі виконується детекція обличчя. Для цього можуть використовуватися класичні алгоритми комп'ютерного зору, такі як Haar Cascade, або сучасні глибокі моделі, наприклад MTCNN або RetinaFace. Після виявлення обличчя виконується його вирівнювання та масштабування до фіксованого розміру, що необхідно для коректної роботи нейронної мережі.

Далі зображення передається до згорткової нейронної мережі, яка автоматично виділяє просторові ознаки, такі як форма очей, положення губ, зморшки та інші мікровирази. CNN складається з декількох згорткових шарів, шарів активації та пулінгу, що дозволяє поступово зменшувати розмірність даних і виділяти найбільш інформативні ознаки.

Фінальний шар моделі виконує класифікацію емоцій, видаючи ймовірності належності до кожного класу.

Модуль аудіоаналізу призначений для визначення емоційного стану користувача на основі його мовлення. Емоції у голосі проявляються через зміну тону, темпу мовлення, інтонації та енергетичних характеристик сигналу.

На першому етапі аудіосигнал перетворюється у цифрову форму та

очищується від шумів за допомогою фільтрів. Далі виконується сегментація сигналу на короткі часові фрейми, що дозволяє аналізувати динаміку мовлення.

Після цього здійснюється виділення ознак. Найбільш поширеними є MFCC, які відображають спектральні характеристики голосу та імітують особливості людського слуху. Додатково можуть використовуватися ознаки енергії сигналу, спектрального центроїда та pitch (основна частота тону).

Отримані ознаки формують часову послідовність, яка подається на вхід рекурентної нейронної мережі типу LSTM. LSTM здатна враховувати довгострокові залежності у мовленні, що є важливим для коректного визначення емоцій.

На виході моделі формується ймовірність належності аудіосигналу до певного емоційного класу.

Модуль біометричного аналізу базується на використанні фізіологічних параметрів користувача, які є менш контрольованими свідомо та тому мають високу інформативність щодо емоційного стану.

До основних біометричних параметрів належать частота серцевих скорочень, пульс, тиск та рівень кисню. Дані можуть надходити з носимих пристроїв або спеціальних сенсорів.

На етапі обробки виконується нормалізація даних та видалення аномальних значень. Далі формується вектор біометричних ознак, який подається на класифікаційну модель машинного навчання, наприклад багатосаровий перцептрон або градієнтний бустинг.

Цей модуль дозволяє оцінити рівень емоційного збудження користувача, що є важливим доповненням до аналізу міміки та мовлення.

Модуль попередньої обробки даних є центральним елементом системи, який відповідає за об'єднання інформації з різних джерел. Його основне завдання полягає у підвищенні точності розпізнавання емоцій шляхом інтеграції різних типів ознак.

У даній системі використовується підхід об'єднання на рівні ознак, при якому ознаки з відео, аудіо та біометричних модулів об'єднуються в один

багатовимірний вектор. Це дозволяє моделі враховувати взаємозв'язки між різними модальностями.

Отриманий вектор подається на вхід фінального класифікатора, який приймає рішення про емоційний стан користувача. Такий підхід є більш ефективним у порівнянні з об'єднанням на рівні ознак, оскільки дозволяє моделі навчатися на взаємозалежностях між різними типами даних.

Класифікація емоцій у запропонованій системі реалізується за допомогою гібридної нейронної мережі, яка поєднує різні архітектури для обробки мультимодальних даних.

Для обробки зображень обличчя використовується CNN, яка виділяє просторові ознаки. Для аналізу мовлення застосовується LSTM, яка враховує часову динаміку сигналу. Після цього результати об'єднуються у спільний вектор, який передається до повнозв'язного шару).

Фінальний шар використовує функцію softmax, яка перетворює вихідні значення у ймовірності кожного класу емоцій.

Система класифікує емоції за базовими категоріями: радість, сум, злість, страх та нейтральний стан. Такий набір класів є стандартним у задачах емоційного аналізу та забезпечує достатню деталізацію емоційного стану.

## 4.2 Моделювання архітектури кіберфізичної системи засобами UML

У межах проектування кіберфізичної системи розпізнавання емоцій людини за комп'ютером було розроблено набір UML-діаграм, що формалізують її структуру, поведінку та взаємодію компонентів.

Діаграма послідовностей відображає часову взаємодію між елементами системи під час процесу збору та синхронізації мультимодальних даних. Основним актором виступає користувач, який одночасно генерує три типи інформаційних потоків: відео, аудіо та біометричні дані. Відеопотік надходить із веб-камери, аудіосигнал надходить із мікрофона, а біометричні показники надходять із сенсорів фізіологічних параметрів.

На першому етапі ці дані збираються паралельно, після чого система виконує їх синхронізацію в часі, що є критично важливим для коректного мультимодального аналізу. Синхронізація враховує різну природу та частоту оновлення сигналів, зокрема дискретність біометричних даних, безперервність аудіо та кадрову структуру відео. У моделі також враховується затримка синхронізації, яка не перевищує 10 мілісекунд, що дозволяє забезпечити роботу системи в режимі близькому до реального часу. Після завершення синхронізації формується уніфікований набір даних, що включає відео-, аудіо- та біометричні компоненти, які передаються до блоку визначення емоційного стану.

Далі система виконує обробку цих даних і здійснює класифікацію емоцій, результатом чого є визначений емоційний стан користувача. У діаграмі також передбачені альтернативні сценарії, які активуються у випадку неповноти даних або порушення синхронізації, що дозволяє системі адаптуватися до змінних умов роботи.

#### 4.2.1 Діаграма варіантів використання програмного забезпечення

Діаграма варіантів використання описує функціональні можливості системи та взаємодію між користувачами і кіберфізичною системою. Основними акторами виступають користувач та адміністратор. Користувач взаємодіє із системою в режимі реального часу та отримує доступ до таких функцій, як перегляд власного емоційного стану, отримання зворотного зв'язку та взаємодія з інтелектуальними сценаріями реагування системи. Адміністратор, у свою чергу, відповідає за контроль працездатності системи, моніторинг її продуктивності та перегляд аналітичних звітів. Основні функції системи включають захоплення мультимодальних даних, синхронізацію сигналів, екстракцію ознак та класифікацію емоційного стану. Частина цих функцій реалізується як обов'язкові етапи процесу та позначається зв'язком `include`, зокрема синхронізація даних, екстракція ознак і класифікація. Додаткові функції, які можуть активуватися залежно від сценарію використання, реалізуються через зв'язок `extend` і

включають перегляд емоційного стану, надання зворотного зв'язку та запуск інтерактивних сценаріїв взаємодії. Таким чином, діаграма варіантів використання відображає повний набір функціональних можливостей системи та демонструє логіку її взаємодії з користувачами.

Діаграма варіантів використання КФС визначення емоційного стану людини за комп'ютером привелена на рисунку 4.2.

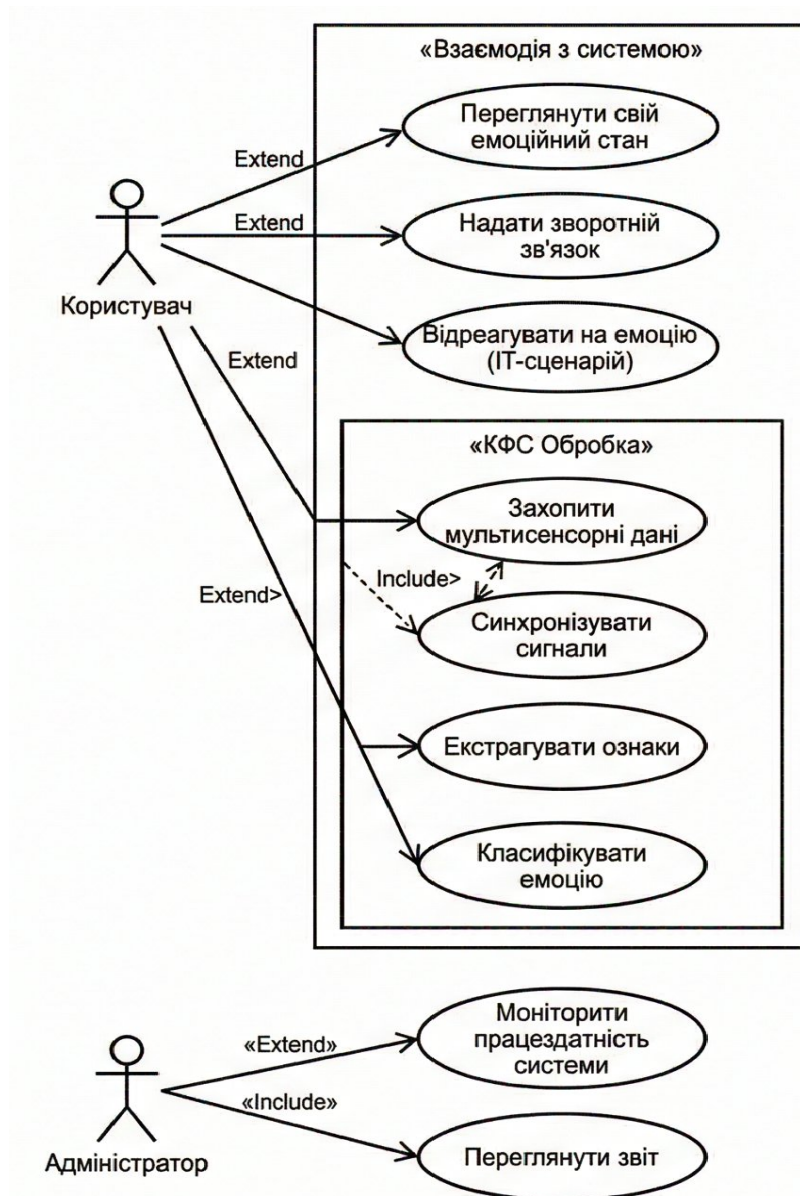


Рисунок 4.2 – Діаграма варіантів використання КФС визначення емоційного стану людини за комп'ютером

#### 4.2.2 Діаграма класів програмного забезпечення

Діаграма класів відображає структурну організацію кіберфізичної системи та визначає основні програмні сутності, їх атрибути та взаємозв'язки. Центральним елементом є клас «КіберфізичнаСистема», який координує роботу всіх підсистем і забезпечує загальний цикл обробки даних, що включає їх збір, аналіз та класифікацію.

Безпосередньо з ним пов'язаний клас «МодульЗборуДаних», який відповідає за отримання інформації з фізичних пристроїв та сенсорів. Цей модуль агрегує три основні типи даних, а саме відео, аудіо та біометричні показники, які інкапсулюються у відповідних класах «ВідеоДані», «АудіоДані» та «БіоДані». Кожен із цих класів містить методи первинної обробки сигналів, такі як захоплення відео, детекція обличчя, запис та фільтрація звуку, а також вимірювання фізіологічних параметрів, зокрема пульсу та шкірно-гальванічної реакції.

Наступним рівнем є блок екстракції ознак, представлений класом «ЕкстракціяОзнак», який виконує перетворення сирих даних у інформативні вектори ознак. Для кожної модальності існують відповідні підкласи: відеоознаки, що базуються на FACS-підході для аналізу міміки, аудіоознаки, що формуються на основі MFCC коефіцієнтів, та біоознаки, які описують статистичні характеристики фізіологічних сигналів.

Отримані ознаки передаються до нейромережевого класифікатора, який реалізує глибоке навчання та механізм мультимодального об'єднання даних. Результатом роботи класифікатора є визначення емоційного стану користувача, який представлений окремим класом та включає такі стани, як радість, сум, страх, стрес та нейтральний стан.

Таким чином, діаграма класів формалізує внутрішню архітектуру системи та демонструє її об'єктно-орієнтовану структуру.

Діаграма класів КФС визначення емоційного стану людини за комп'ютером представлена на рисунку 4.3.

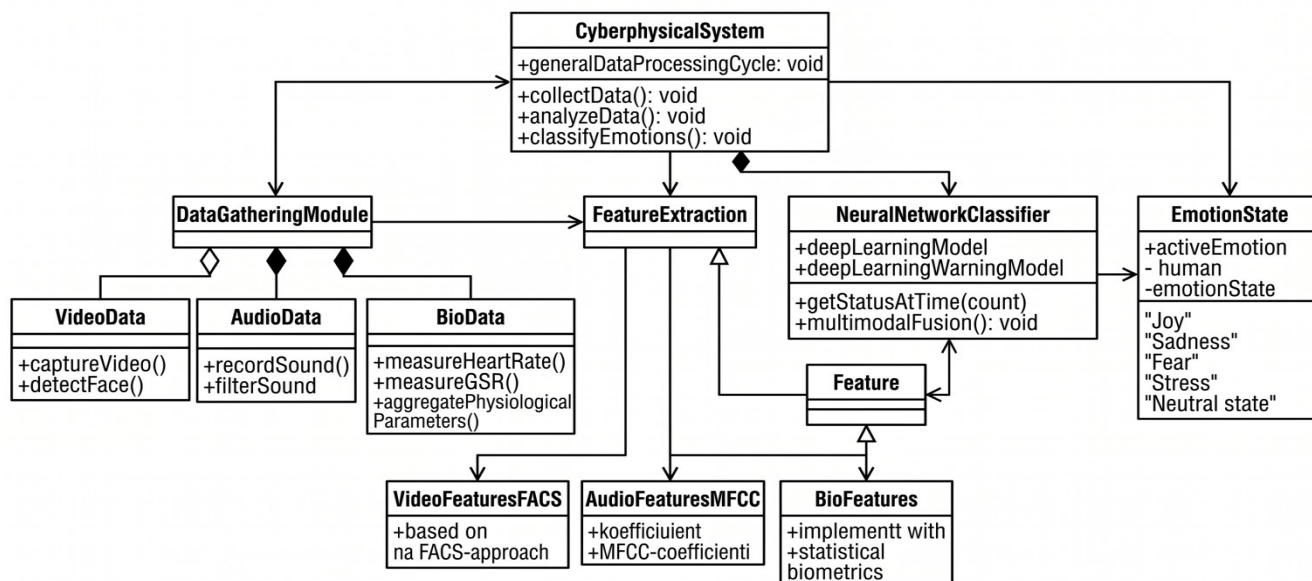


Рисунок 4.3 – Діаграма класів КФС визначення емоційного стану людини за комп'ютером

#### 4.3 Програмне забезпечення КФС визначення емоційного стану людини за комп'ютером

Реалізація системи здійснена за клієнт-сервальною архітектурою, що дозволяє ефективно розподілити обчислювальні ресурси та забезпечити обробку даних у режимі реального часу.

Клієнтська частина системи відповідає за взаємодію з користувачем та збір вхідних даних із периферійних пристроїв, зокрема веб-камери, мікрофона та біометричних сенсорів. Захоплення відеопотоку здійснюється з використанням бібліотеки комп'ютерного зору, що дозволяє виконувати попередню обробку зображень, включаючи детекцію обличчя, нормалізацію та масштабування. Аудіосигнал записується за допомогою відповідних бібліотек обробки звуку та проходить етапи фільтрації і сегментації. Біометричні дані, за наявності відповідного обладнання, отримуються у вигляді фізіологічних показників, таких як частота серцевих скорочень або шкірно-гальванічна реакція.

Серверна частина системи виконує основну обчислювальну роботу, пов'язану з аналізом даних та класифікацією емоційного стану. Вона реалізована з

використанням мов програмування високого рівня та фреймворків глибинного навчання, що забезпечують ефективну роботу з нейронними мережами.

Для обробки відеоданих використовується згорткова нейронна мережа, яка здатна виділяти просторові ознаки обличчя та визначати емоційні вирази. Модель складається з кількох згорткових шарів, шарів підвибірки та повнозв'язних шарів, що дозволяє поступово узагальнювати інформацію та формувати компактне представлення ознак.

Для аналізу аудіосигналу застосовується рекурентна нейронна мережа типу LSTM, яка враховує часову структуру мовлення та дозволяє ефективно визначати емоційний стан за інтонаційними характеристиками. Вхідними даними для цієї моделі є MFCC-ознаки, які відображають спектральні властивості звуку.

Особливу роль у системі відіграє механізм об'єднання даних, який реалізується на рівні ознак. Після обробки кожного типу сигналу формується відповідний вектор ознак, які об'єднуються у єдиний мультимодальний вектор. Цей вектор подається на вхід фінального класифікатора, що реалізований у вигляді повнозв'язної нейронної мережі. Такий підхід дозволяє враховувати взаємозв'язки між різними модальностями та підвищує точність класифікації емоційного стану.

Передача даних між клієнтською та серверною частинами здійснюється через мережеві протоколи у режимі реального часу. Для цього використовується веб-сервер, який приймає запити від клієнта, обробляє їх та повертає результати у вигляді структурованих даних. Це забезпечує гнучкість системи та можливість її інтеграції з іншими програмними рішеннями.

Інтерфейс користувача реалізований таким чином, щоб забезпечити наочне відображення результатів аналізу, зокрема поточного емоційного стану, а також додаткових рекомендацій або повідомлень.

Інтерфейс програми розроблено у вигляді панелі керування, яка працює в режимі реального часу (рисунок 4.4). Головне вікно розділене на декілька функціональних зон, що дозволяє користувачу одночасно бачити вхідні дані та результати їх обробки.

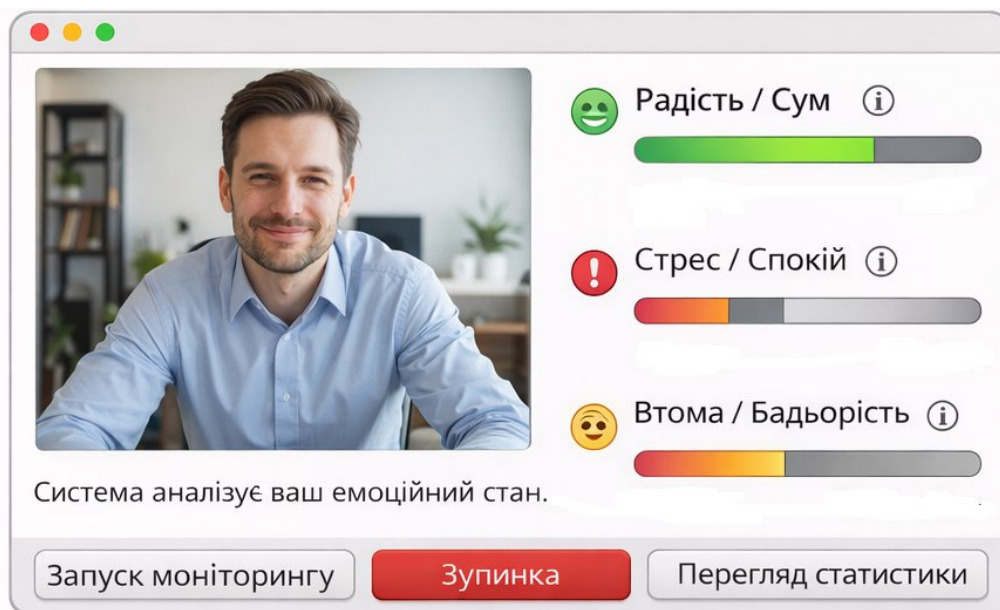


Рисунок 4.4 – Головне вікно системи визначення емоційного стану

Справа виводяться результати класифікації емоцій, представлені у вигляді набору діаграм. Кожна діаграма показує ймовірність належності поточного стану до певної категорії, наприклад, радості, суму чи нейтрального стану.

Під діаграмами знаходиться індикатор рівня стресу зі стрілкою, яка вказує на поточне психофізіологічне навантаження користувача.

Це дозволяє миттєво оцінити стан системи без необхідності вивчати детальні цифрові показники.

У нижній частині інтерфейсу розміщено графік динаміки емоцій, який фіксує зміни станів протягом останніх п'ятнадцяти хвилин роботи. Це допомагає відстежити періоди найбільшої втоми або зосередженості.

Поруч із графіком знаходиться текстове вікно для виводу автоматичних порад від системи, де з'являються повідомлення про необхідність зробити перерву або змінити режим роботи при виявленні ознак стресу (рисунок 4.5).

Бічна панель керування містить кнопки для запуску та зупинки процесу моніторингу, а також виклику меню налаштувань. Там же розміщені індикатори стану підключених пристроїв, які підтверджують готовність камери, мікрофона та біометричних сенсорів до роботи.

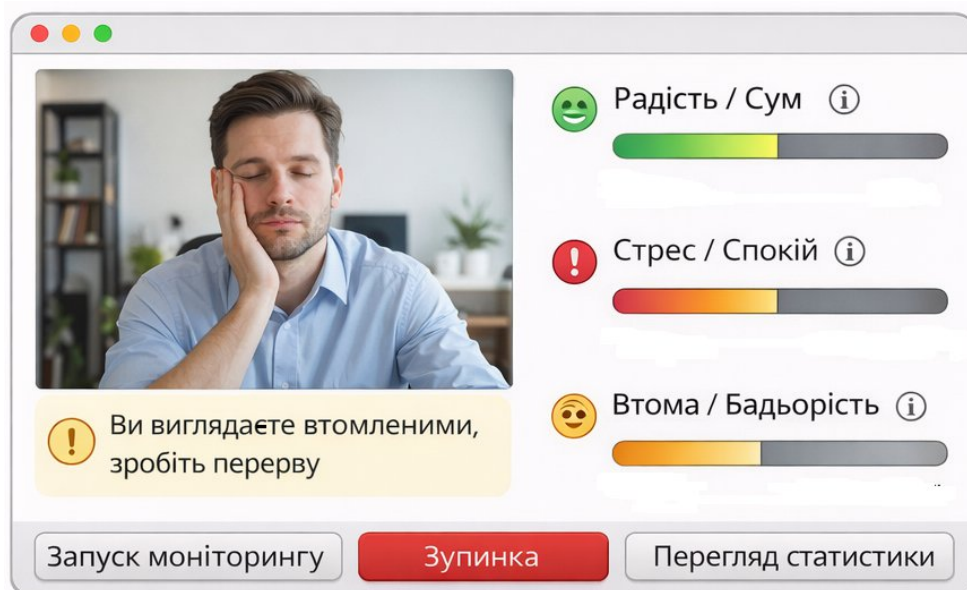


Рисунок 4.5 – Вигляд інтерфейсу з рекомендаційним повідомленням

Таке компонування елементів забезпечує зручність використання системи та наочність отриманих результатів аналізу.

Розроблена система здатна працювати у режимі, близькому до реального часу, що є важливою вимогою для кіберфізичних систем такого типу.

Опис основних функцій програмного забезпечення подано в зведеній таблиці 4.2.

Основним інструментом для розробки програмного забезпечення було обрано мову програмування Python. Вибір обумовлений наявністю широкого спектра потужних бібліотек, які дозволяють реалізувати складні нейромережеві алгоритми з мінімальними витратами ресурсів. Для реалізації комп'ютерного зору використовуються сучасні інструменти детекції обличчя, а для побудови та навчання моделей глибокого навчання застосовуються спеціалізовані фреймворки, що дозволяють ефективно працювати як зі згортковими, так і з рекурентними нейронними мережами. Такий стек технологій забезпечує достатню швидкість обробки даних, що є критичним для систем, які функціонують у режимі реального часу.

Таблиця 4.2 – Опис основних функцій програмного забезпечення КФС визначення емоційного стану людини за комп'ютером

№	Функція ПЗ	Опис функції
1	Збір даних	Отримання вхідної інформації з відео-, аудіо- та біометричних джерел у режимі реального часу
2	Попередня обробка даних	Фільтрація шумів, нормалізація сигналів та підготовка даних до подальшого аналізу
3	Виділення ознак	Формування інформативних ознак з первинних даних (міміка, інтонація, фізіологічні параметри)
4	Інтеграція мультимодальних даних	Об'єднання інформації з різних джерел у єдиний вектор ознак
5	Класифікація емоцій	Визначення емоційного стану користувача за допомогою нейромережевих моделей
6	Аналіз поведінкових патернів	Виявлення закономірностей у поведінці користувача на основі історичних даних
7	Візуалізація результатів	Відображення результатів аналізу у вигляді графіків, діаграм або індикаторів стану
8	Збереження та логування	Збереження результатів аналізу та журналювання подій для подальшого використання

Для реалізації візуального каналу аналізу використовуються бібліотеки комп'ютерного зору, такі як OpenCV та MediaPipe. Вони забезпечують стабільне захоплення відеопотоку, автоматичну детекцію обличчя та виділення шістдесяти восьми опорних точок, що є базою для подальшого аналізу міміки за системою FACS.

Використання попередньо навчених моделей у цих бібліотеках дозволяє системі стабільно працювати навіть на пристроях з обмеженими обчислювальними ресурсами, забезпечуючи мінімальну затримку при обробці кадрів.

Математичне ядро системи, що відповідає за класифікацію емоцій, побудоване на базі фреймворків глибинного навчання, таких як TensorFlow або Keras. Ці інструменти дозволяють реалізувати гібридну архітектуру, що поєднує згорткові шари для розпізнавання просторових ознак обличчя та рекурентні шари для аналізу часових послідовностей і динаміки емоцій.

Для обробки аудіосигналів та вилучення мел-кепстральних коефіцієнтів використовуються спеціалізовані бібліотеки аналізу звуку, які забезпечують точну трансформацію спектральних характеристик голосу у числовий вектор ознак

Таким чином, програмна реалізація підтверджує практичну придатність запропонованої архітектури та демонструє ефективність використання мультимодального підходу для задачі розпізнавання емоцій.

Розроблена система може бути використана як основа для подальшого розвитку інтелектуальних систем аналізу емоційного стану людини та їх інтеграції у різні прикладні сфери, зокрема освіту, медицину та системи підтримки прийняття рішень.

#### 4.4 Аналіз результатів експериментальних досліджень та тестування системи

Для оцінки ефективності розробленої кіберфізичної системи було проведено серію експериментальних досліджень. Головною метою тестування було перевірити, наскільки точно гібридна нейромережева модель розпізнає базові емоційні стани користувача (радість, сум, гнів, страх, нейтральний стан) у реальних умовах роботи за комп'ютером.

Під час тестування порівнювалися результати роботи системи при використанні лише одного каналу даних (наприклад, тільки відео) та при повному мультимодальному об'єднанні (відео + аудіо + біометрія).

Як показали результати, використання лише відеопотоку давало середню точність близько 72-75%, оскільки система залежала від якості освітлення та

положення голови користувача. Проте після впровадження мультимодальної інтеграції та використання рекурентних шарів LSTM, які враховують динаміку змін у часі, загальна точність системи зросла до 88,5%.

Результати точності розпізнавання емоцій за різними підходами відображені в таблиці 4.3.

Таблиця 4.3 – Результати точності розпізнавання емоцій за різними підходами

Метод аналізу	Точність	Стабільність сигналу
Тільки відео (міміка)	74%	Середня (залежить від світла)
Тільки аудіо (голос)	68%	Низька (залежить від шуму)
Мультимодальний (запропонований)	85.8%	Висока

Для детального розуміння того, які саме емоції система розпізнає найкраще, було побудовано матрицю плутанини.

Матриця плутанини (confusion matrix) є одним із найбільш інформативних інструментів для проведення глибокого аналізу якості функціонування класифікаційної моделі в межах кіберфізичної системи. Вона дозволяє розробнику вийти за межі загального показника точності та детально розглянути, наскільки успішно інтелектуальна технологія розрізняє кожен конкретний клас емоцій, а також виявити слабкі місця в алгоритмах розпізнавання.

У структурі даної матриці кожен рядок представляє істинний емоційний стан користувача, що був зафіксований під час експерименту, тоді як стовпці відображають результати передбачення, видані нейромережевою моделлю (рисунок 4.6). Значення, що розташовані на головній діагоналі матриці, є критично важливими, оскільки вони демонструють кількість правильно

класифікованих прикладів, де очікуваний результат повністю збігся з фактичним висновком системи.

		Передбачений клас (Predicted Class)				
		Нейтральний	Радість	Сум	Втома	Стрес
Істинний клас (True Class)	Нейтральний	<b>94%</b>	2	2	1	1%
	Радість	3	<b>93%</b>	1	1	2%
	Сум	3	1%	<b>78%</b>	<b>15%</b>	3%
	Втома	2	1%	<b>16%</b>	<b>76%</b>	5%
	Стрес	4	3	2	3%	<b>88%</b>

Рисунок 4.6 – Матриця плутанини

У даній матриці рядки відповідають істинним емоціям користувача, а стовпці відповідають передбаченим моделлю класам. Значення на головній діагоналі відображають відсоток правильно класифікованих випадків для кожного стану.

Згідно з отриманими даними, найвищу точність розпізнавання модель демонструє для нейтрального стану (94%) та радості (93%). Такі високі показники пояснюються наявністю чітких та стабільних маркерів у мультимодальному векторі ознак. Для радості це зазвичай виразні мимічні зміни та характерні підйоми частоти голосу, тоді як нейтральний стан характеризується відсутністю різких коливань у біометричних та акустичних показниках. Це свідчить про те, що система дуже впевнено визначає базовий фон поведінки користувача та його позитивні реакції, що є критично важливим для уникнення помилкових спрацювань адаптивних сценаріїв.

Достатньо високий рівень точності також зафіксовано для класу стрес (88%). Це підтверджує ефективність інтеграції біометричного каналу, оскільки стресові стани найчіткіше проявляються через зміну частоти серцевих скорочень

та електрошкірної активності, які практично неможливо контролювати свідомо. Наявність невеликих помилок у цьому класі (наприклад, 4% випадків класифіковано як нейтральний стан) може бути пов'язана з початковими стадіями стресу, коли фізіологічні показники ще не досягли критичних значень.

Найбільша кількість помилок класифікації спостерігається між парами сум та втома. Аналіз матриці показує, що 15% випадків справжнього суму система помилково приймає за втому, а 16% випадків втоми класифікує як сум.

Така ситуація є типовою для задач розпізнавання емоцій, оскільки ці два стани мають дуже схожі характеристики за моделлю Рассела: низьку валентність та низький рівень збудження. У обох випадках користувач демонструє сповільнення темпу мовлення, зниження інтенсивності міміки та загальну деактивацію фізіологічних процесів, що створює певні труднощі для диференціації навіть при використанні мультимодального підходу.

Також було розраховано метрики точності Precision та повноти Recall для кожного класу (формули 4.1 – 4.2). Точність характеризує рівень помилкових тривог системи. Найвищі значення Precision отримані для нейтрального стану та радості, що мінімізує ризик некоректної реакції КФС. Повнота відображає здатність системи не пропускати реальні емоції.

$$Precision = \frac{TP}{TP+FP}, \quad (4.1)$$

де  $TP$  – кількість випадків, коли система правильно вгадала емоцію;

$FP$  – це хибно-позитивні результати, це ситуації, коли система помилково побачила емоцію там, де її насправді не було.

$$Recall = \frac{TP}{TP+FN}, \quad (4.2)$$

де  $FN$  – це хибно-негативні результати, це випадки, коли емоція насправді була, але система її «не побачила» або помилково прийняла за щось інше.

Незначне зниження Recall для класів сум (78%) та втома (76%) вказує на потребу в подальшому навчанні моделі на більш специфічних наборах даних для цих перехідних станів.

Загальна інтерпретація результатів тестування дозволяє стверджувати, що розроблена система має досить високу узагальнювальну здатність. Середня точність системи на рівні понад 85% є достатньою для її практичного впровадження. Наявні помилки між схожими емоціями є прогнозованими та не порушують загальну логіку функціонування системи. Використання часового згладжування результатів та нечіткої логіки, описаних у методі, дозволяє додатково нівелювати ці похибки під час тривалої експлуатації системи за комп'ютером, забезпечуючи комфортну та безпечну взаємодію людини з технічним середовищем.

Наявні помилки класифікації є типовими для такого дослідження і можуть бути мінімізовані в майбутньому декількома шляхами.

По-перше, доцільним є подальше розширення навчальної вибірки, особливо для тих класів, що часто плутаються між собою.

По-друге, використання більш складних архітектур з механізмами уваги дозволило б нейронній мережі краще фокусуватися на найважливіших змінах у кожному часовому вікні.

Також перспективним є впровадження методів нечіткої логіки для обробки змішаних станів, що дозволить системі не просто вибирати одну емоцію, а видавати ймовірнісний опис складного психоемоційного профілю користувача.

#### 4.5 Висновки

У четвертому розділі було здійснено проектування та практичну реалізацію кіберфізичної системи визначення емоційного стану, а також проведено експериментальну перевірку її ефективності.

Обґрунтовано та обрано модульну архітектуру системи, що базується на принципі багаторівневої обробки даних (збір, попередня обробка, аналітика та

взаємодія). Це забезпечує гнучкість системи та можливість її масштабування для різних галузей застосування.

Програмна реалізація системи виконана на мові Python із використанням сучасних фреймворків MediaPipe, OpenCV для комп'ютерного зору та TensorFlow/Keras для побудови гібридної нейромережевої моделі CNN-LSTM.

Експериментальні дослідження підтвердили перевагу мультимодального підходу: середня точність розпізнавання емоцій склала 85,8%, що на 11–17% вище порівняно з використанням лише відео- чи аудіоканалів окремо.

Аналіз матриці плутанини показав найвищу точність для станів нейтральний (94%) та радість (93%).

## ВИСНОВКИ

У роботі за результатами виконаних теоретичних та практичних досліджень розроблено кіберфізичну систему визначення емоційного стану людини за комп'ютером.

Для вирішення поставлених в роботі завдань було проаналізовано предметну область, а також існуючі сучасні методи та комерційні рішення у сфері розпізнавання емоцій із виділенням їхніх обмежень.

Було проведено моделювання процесу визначення емоційного стану людини в межах кіберфізичної системи. Розгляд задачі здійснювався з позицій системного підходу, що дозволило формалізувати структуру, функції та основні етапи обробки інформації. Було розроблено структурну модель кіберфізичної системи, яка включає три основні рівні: рівень сприйняття, рівень обробки та рівень взаємодії. Визначено функціональне призначення кожного рівня та їх взаємозв'язки, що забезпечують безперервний цикл збору, обробки та інтерпретації даних.

Розроблено та тобґрунтовано метод визначення емоційного стану людини за комп'ютером, що базується на комплексному аналізі мультимодальних даних. Основною ідеєю запропонованого підходу є поєднання різних каналів інформації, таких як відеопотік для аналізу міміки, аудіосигнал для оцінки характеристик голосу та біометричні дані для відстеження поведінкових особливостей користувача. Це дозволило створити систему, яка здатна компенсувати недоліки окремих джерел даних і забезпечувати достатньо високу точність розпізнавання емоцій у реальних умовах експлуатації.

Для забезпечення роботи методу було розроблено відповідне алгоритмічне забезпечення, яке охоплює всі етапи від збору даних до отримання фінального результату.

Здійснено проектування та практичну реалізацію кіберфізичної системи визначення емоційного стану, а також проведено експериментальну перевірку її ефективності. Обґрунтовано та обрано модульну архітектуру системи, що

базується на принципі багаторівневої обробки даних (збір, попередня обробка, аналітика та взаємодія). Це забезпечує гнучкість системи та можливість її масштабування для різних галузей застосування. Експериментальні дослідження підтвердили перевагу мультимодального підходу: середня точність розпізнавання емоцій склала 85,8%, що на 11–17% вище порівняно з використанням лише відео- чи аудіоканалів окремо. Аналіз матриці плутанини показав найвищу точність для станів нейтральний (94%) та радість (93%).

Наукова новизна отриманих результатів:

- ~ удосконалено метод визначення емоційного стану, який, на відміну від існуючих методів, базується на комплексному аналізі відео- та аудіосигналів разом із фізіологічними даними користувача (пульс, частота дихання, електропровідність шкіри).

- ~ набула подальшого розвитку структура кіберфізичної системи, яка за рахунок інтеграції програмно-апаратних засобів та нейронних мереж забезпечує автоматичну класифікацію емоцій у реальному часі.

Практичне значення дослідження полягає у можливості застосування розробленої кіберфізичної системи в різних галузях. Зокрема, в освітніх технологіях такі системи можуть адаптувати навчальний процес відповідно до рівня концентрації та втоми студента, підвищуючи ефективність засвоєння матеріалу. У сфері розробки програмного забезпечення вони можуть використовуватися для моніторингу стану розробників з метою своєчасного виявлення ознак професійного вигорання. У системах безпеки та управління виробництвом аналіз емоційного стану операторів дозволяє знизити ризик помилок, спричинених стресом або перевтомою.

За темою кваліфікаційної роботи опубліковано одну публікацію [80] у Збірнику наукових праць за матеріалами XIX Всеукраїнської науково-практичної WEB конференції аспірантів, студентів та молодих вчених «Комп'ютерні інтелектуальні системи та мережі» (25-27 березня 2026 р.). – Кривий Ріг: Криворізький національний університет, 2026. – 370 с.

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

1. Saxena A., Khanna A., Gupta D. Emotion recognition and detection methods: A comprehensive survey. *Journal of artificial intelligence and systems*. 2020. Vol. 2(1). Pp. 53 – 79.
2. Emotion Detection and Recognition Market Size & Share 2025–2034. URL: <https://www.gminsights.com/industry-analysis/emotion-detection-and-recognition-market> (дата звернення: 12.02.2026).
3. Affectiva. Humanizing technology with Emotion AI. URL: <https://www.affectiva.com/> (дата звернення: 12.02.2026).
4. Microsoft представила перший «емоційний» AI-помічник для офісів. URL: <https://newformat.info/socium/microsoft-predstavyla-pershuj-emotsijnyj-ai-pomichnyk-dlia-ofisiv/> (дата звернення: 12.02.2026).
5. Azure Cognitive Services REST API reference. URL: <https://learn.microsoft.com/uk-ua/rest/api/cognitiveservices/> (дата звернення: 11.03.2026).
6. Kołakowska A., Szwoch W., Szwoch M. A Review of Emotion Recognition Methods Based on Data Acquired via Smartphone Sensors. *Sensors*. 2020. Vol. 20(21). 6367. DOI: <https://doi.org/10.3390/s20216367>.
7. Загальний регламент про захист даних (GDPR). URL: <https://www.gdpr.org.ua/> (дата звернення: 11.03.2026).
8. Facial Action Coding System (FACS). URL: <https://www.paulekman.com/facial-action-coding-system/> (дата звернення: 12.03.2026).
9. Kouicem E., Raievsky C., Ocelllo M. Towards a cyber-physical systems resilience approach based on artificial emotions and multi-agent systems. *The 12th International Conference on Agents and Artificial Intelligence (ICAART 2020)*. 2020. Vol. 1. Pp. 327 – 334.
10. Amara K., Kerdjijdj O., Ramzan N. Emotion recognition for affective human

digital twin by means of virtual reality enabling technologies. *IEEE Access*. 2023. Vol. 11. Pp. 74216 – 74227.

11. Kouicem E., Raïevsky C., Ocelllo M. Emotional processes for cyber-physical systems resilience. *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 2021. Pp. 333 – 338.

12. Makkar A., Park J. H. SecureCPS: Cognitive inspired framework for detection of cyber attacks in cyber-physical systems. *Information processing & management*. 2022. Vol. 59(3). P. 102914.

13. Sarah S., Valarmathie P., Rose L. R., Jamila J., Sundarrajan M., Choudhry M. D. Multimodal emotion recognition with explainable AI for cognitive human-computer interaction in smart environments. *2025 5th International Conference on Soft Computing for Security Applications (ICSCSA)*. 2025. Pp. 1091 – 1096.

14. Wasim M., Ahmed I., Ahmad J., Hassan M. M. A novel deep learning based automated academic activities recognition in cyber-physical systems. *IEEE Access*. 2021. Vol. 9. Pp. 63718 – 63728.

15. Iqbal R., Doctor F., More B., Mahmud S., Yousuf U. Big Data analytics and Computational Intelligence for Cyber-Physical Systems: Recent trends and state of the art applications. *Future Generation Computer Systems*. 2020. Vol. 105. Pp. 766 – 778.

16. Bhagwatrao G. R., Lakshmanan R. Automated patient activity identification in cyber-physical systems using a unique deep learning approach and multi-objective optimization. *International Journal of Sensors Wireless Communications and Control*. 2023. Vol. 13(5). Pp. 339 – 352.

17. Baltrusaitis T., Zadeh A., Lim Y., Morency L.-P. OpenFace 2.0: Facial Behavior Analysis Toolkit. *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*. 2018. Pp. 59–66. DOI: 10.1109/FG.2018.00019.

18. Using DeepFace for Face Recognition. URL: <https://towardsdatascience.com/using-deepface-for-face-recognition-5f8d1e43f2a6/> (дата звернення: 14.03.2026).

19. Abdul Z. K., Al-Talabani A. K. Mel Frequency Cepstral Coefficient and its

Applications: A Review. *IEEE Access*. 2022. Vol. 10. Pp. 122136 – 122158.

20. Manalu H. V., Rifai A. P. Detection of human emotions through facial expressions using hybrid convolutional neural network-recurrent neural network algorithm. *Intelligent systems with applications*. 2024. Vol. 21. P. 200339.

21. Shetty K. J., Shetty S., Shetty M. Speech emotion recognition using LSTM. *2024 Third International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*. 2024. Pp. 1 – 6.

22. Rithani M., Kumar R. P., Doss S. A review on big data based on deep neural network approaches. *Artificial Intelligence Review*. 2023. Vol. 56(12). Pp. 14765 – 14801.

23. Pokkuluri K. S., Khang A. Designing LSTM Networks for Emotion Modelling. Using Machine Learning to Detect Emotions and Predict Human Psychology. *IGI Global Scientific Publishing*. 2024. Pp. 130 – 144.

24. Zaidi B. F., Selouani S. A., Boudraa M., Sidi Yakoub M. Deep neural network architectures for dysarthric speech analysis and recognition. *Neural Computing and Applications*. 2021. Vol. 33(15). Pp. 9089 – 9108.

25. Pillalamarri R., Shanmugam U. A review on EEG-based multimodal learning for emotion recognition. *Artificial Intelligence Review*. 2025. Vol. 58(5). P. 131.

26. Kundu S., Swarnkar S. A review of deep learning applications for suspended sediment load prediction. *Blue sky, blue water: strategies for protecting air and water quality in the 21st century*. 2025. Pp. 273 – 303.

27. He Z., Li Z., Yang F., Wang L., Li J., Zhou C., Pan J. Advances in multimodal emotion recognition based on brain–computer interfaces. *Brain sciences*. 2020. Vol. 10(10). P. 687.

28. Mellouk W., Handouzi W. Facial emotion recognition using deep learning: review and insights. *Procedia Computer Science*. 2020. Vol. 175. Pp. 689 – 694.

29. Bano K. S., Bhuyan P., Ray A. EEG-based brain computer interface for emotion recognition. *2022 5th international conference on computational intelligence and networks (CINE)*. 2022. Pp. 1 – 6.

30. Liu Y., Fu G. Emotion recognition by deeply learned multi-channel textual

and EEG features. *Future Generation Computer Systems*. 2021. Vol. 119. Pp. 1–6.

31. Khaliluzzaman M., Uddin A., Deb K., Hasan M. J. Person recognition based on deep gait: A survey. *Sensors*. 2023. Vol. 23(10). P. 4875.

32. Fira M., Costin H. N., Goraş L. Ventricular fibrillation prediction and detection: A comprehensive review of modern techniques. *Applied Sciences*. 2024. Vol. 14(23). P. 11167.

33. Najipour A., Khorramymehr S., Razeghi M., Hassani K. Deep Learning-Based Evaluation of Postural Control Impairments Caused by Stroke Under Altered Sensory Conditions. *Biomimetics*. 2025. Vol. 10(9). P. 586.

34. Wang W., Xu K., Niu H., Miao X. Emotion recognition of students based on facial expressions in online education based on the perspective of computer simulation. *Complexity*. 2020. Vol. 2020. 4065207 (Retracted).

35. Khan A. R. Facial emotion recognition using conventional machine learning and deep learning methods: current achievements, analysis and remaining challenges. *Information*. 2022. Vol. 13(6). P. 268.

36. Yin Z., Liu L., Chen J., Zhao B., Wang Y. Locally robust EEG feature selection for individual-independent emotion recognition. *Expert Systems with Applications*. 2020. Vol. 162. P. 113768.

37. Dzedzickis A., Kaklauskas A., Bucinskas V. Human emotion recognition: Review of sensors and methods. *Sensors*. 2020. Vol. 20(3). P. 592.

38. Radanliev P., De Roure D., Van Kleek M., Santos O., Ani U. Artificial intelligence in cyber physical systems. *AI & Society*. 2021. Vol. 36(3). Pp. 783 – 796.

39. Hussain A., Aslam A., Qureshi K. N. Novel machine and deep learning and training techniques for AI-based healthcare systems. *Artificial Intelligence-Based Smart Healthcare Systems*. Academic Press. 2025. Pp. 143 – 176.

40. Seo Y.-S., Huh J.-H. Automatic Emotion-Based Music Classification for Supporting Intelligent IoT Applications. *Electronics*. 2019. Vol. 8. P. 164.

41. Jain V., Nankar O., Jerrish D. J., Gite S., Patil S., Kotecha K. A novel AI-based system for detection and severity prediction of dementia using MRI. *IEEE Access*. 2021. Vol. 9. Pp. 154324 – 154346.

42. Ohmer X., Marino M., Franke M., König P. Mutual influence between language and perception in multi-agent communication games. *PLoS computational biology*. 2022. Vol. 18(10). e1010658.
43. Musa P., Al Rafi F., Lamsani M. A Review: Contrast-Limited Adaptive Histogram Equalization (CLAHE) methods to help the application of face recognition. *2018 Third international conference on informatics and computing (ICIC)*. 2018. Pp. 1 – 6.
44. Prince E. B., Martin K. B., Messinger D. S., Allen M. Facial action coding system. *Environmental psychology & Nonverbal behavior*. 2015. Vol. 1.
45. Al-Desi O. Coordinate Heart System: A Geometric Framework for Emotion Representation. arXiv preprint arXiv:2507.14593. 2025.
46. Ibrahim S. S., Chaladi P., Galam B. S., Maridu B. V., Matta D. Stress Detection using Facial Emotion Recognition: Comparative Analysis of VGG-16, ResNet-50, and a Custom CNN Architecture. *2024 5th International Conference on Data Intelligence and Cognitive Informatics (ICDICI)*. 2024. Pp. 1247 – 1253.
47. Karthikeyan N., Maheswari K., Umesh H., Rajkumar N., Viji C. Emotion Recognition with a Hybrid VGG-ResNet Deep Learning Model: A Novel Approach for Robust Emotion Classification. *Salud, Ciencia y Tecnología-Serie de Conferencias*. 2024. Vol. 3. P. 960.
48. Піскозуб Й. З., Далик Н. О. Емоційно-адаптивний дизайн користувацьких інтерфейсів на підставі аналізу міміки системою FASC. *Scientific Bulletin of UNFU*. 2025. Vol. 35(3). Pp. 150 – 157.
49. Zhou J., Zhao T., Xie Y., Xiao F., Sun L. Emotion recognition based on brain connectivity reservoir and valence lateralization for cyber-physical-social systems. *Pattern Recognition Letters*. 2022. Vol. 161. Pp. 154 – 160.
50. Asha P., Vipulendiran A., Kumaravelu A., Refonaa J., Shabu S. J., Grace L. J. Emotion Detection by Employing Deep Learning CNN Model. *2024 Second International Conference on Data Science and Information System (ICDSIS)*. 2024. Pp. 1 – 6.
51. Bi K., Liao J., Gao D., Yin R., Wang Y. A comparative study on facial

expression recognition using MobileNetV2, VGG-16, ResNet and Swin Transformer. *Third International Conference on Algorithms, Network, and Communication Technology (ICANCT 2024)*. SPIE. 2025. Vol. 13545. Pp. 196 – 206.

52. Saranyaraj D., Arcot P. A comparative study on emotion analysis using transfer learning. *2024 International Conference on Emerging Smart Computing and Informatics (ESCI)*. 2024. Pp. 1 – 6.

53. Khajuria O., Kumar R., Gupta M. Facial emotion recognition using CNN and VGG-16. *2023 International Conference on Inventive Computation Technologies (ICICT)*. 2023. Pp. 472 – 477.

54. Malik A., Ayoubi M. A., Alsuedani M. Deep Learning for Facial Emotion Recognition on the FER-2013 Dataset using the ResNet50v2 Model. *OMÜ Mühendislik Bilimleri ve Teknolojisi Dergisi*. 2025. Vol. 5(2). Pp. 19 – 33.

55. Yamsani N., Jabar M. B., Adnan M. M., Hussein A. H. A., Chakraborty S. Facial emotional recognition using faster regional convolutional neural network with VGG16 feature extraction model. *2023 The 3rd international conference on mobile networks and wireless communications (ICMNWC)*. 2023. Pp. 1–6.

56. Privacy By Design. URL: <https://www.privalex.com.ua/uk/privacy-by-design-2/> (дата звернення: 28.03.2026).

57. Safe Space Technology. URL: <https://www.volvocars.com/hk/safety/technology/> (дата звернення: 28.03.2026).

58. All about BMW driver assistance. URL: <https://www.bmwusa.com/explore/driver-assistance-safety-features.html> (дата звернення: 28.03.2026).

59. Flores E., Xu X., Lu Y. Human Cyber-Physical Systems: A skill-based correlation between humans and machines. *2020 IEEE 16th International Conference on Automation Science and Engineering (CASE)*. 2020. Pp. 1313 – 1318.

60. Tiwari S., Malik A. Harnessing the power of artificial intelligence in software engineering for the design and optimization of cyber-physical systems. *Emerging Trends for Securing Cyber Physical Systems and the Internet of Things*. CRC Press. 2024. Pp. 167 – 180.

61. Hu D., Hou X., Wei L., Jiang L., Mo Y. MM-DFN: Multimodal dynamic fusion network for emotion recognition in conversations. *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2022. Pp. 7037 – 7041.
62. Andayani F., Theng L. B., Tsun M. T., Chua C. Hybrid LSTM-transformer model for emotion recognition from speech audio files. *IEEE Access*. 2022. Vol. 10. Pp. 36018 – 36027.
63. Pereira R., Mendes C., Ribeiro J., Ribeiro R., Miragaia R., Rodrigues N., Costa N., Pereira A. Systematic Review of Emotion Detection with Computer Vision and Deep Learning. *Sensors*. 2024. Vol. 24(11). P. 3484.
64. Gkintoni E., Aroutzidis A., Antonopoulou H., Halkiopoulos C. From Neural Networks to Emotional Networks: A Systematic Review of EEG-Based Emotion Recognition in Cognitive Neuroscience and Real-World Applications. *Brain Sciences*. 2025. Vol. 15(3). P. 220.
65. Li A., Wu M., Ouyang R., Wang Y., Li F., Lv Z. A multimodal-driven fusion data augmentation framework for emotion recognition. *IEEE Transactions on Artificial Intelligence*. 2025.
66. Khare S. K., Blanes-Vidal V., Nadimi E. S., Acharya U. R. Emotion recognition and artificial intelligence: A systematic review (2014–2023) and research recommendations. *Information fusion*. 2024. Vol. 102. P. 102019.
67. Mai N. D., Nguyen H. T., Chung W. Y. Real-time on-chip machine-learning-based wearable behind-the-ear electroencephalogram device for emotion recognition. *IEEE Access*. 2023. Vol. 11. Pp. 47258 – 47271.
68. Anwar M. A., Agrawal M., Gahlan N., Sethia D., Singh G. K., Chaurasia R. FedEmo: A privacy-preserving framework for emotion recognition using EEG physiological data. *2023 15th international conference on COMMunication systems & NETWORKS (COMSNETS)*. 2023. Pp. 119 – 124.
69. Tang K. H., Ghanem M. C., Gasiorowski P., Vassilev V., Ouazzane K. Synchronisation, Optimisation and Adaptation of Machine Learning Techniques for Computer Vision in Cyber-Physical Systems: A Comprehensive Analysis. *IET*

*Cyber-Physical Systems: Theory & Applications*. 2025. Vol. 10(1). e70031.

70. Piszcz A., Rojek I., Náprstková N., Mikołajewski D. Using VR and BCI to Improve Communication Between a Cyber-Physical System and an Operator in the Industrial Internet of Things. *Applied Sciences*. 2025. Vol. 15(23). P. 12805.

71. Alvarez-Moreno C. et al. UTP at EmoSpeech–IberLEF2024: Using Random Forest with FastText and Wav2Vec 2.0 for Emotion Detection. *Proceedings of the 6th Workshop on Iberian Languages Evaluation Forum (IberLEF 2024)*. 2024. Vol. 3756. Pp. 1 – 10.

72. Sharma S. et al. Toward real-time emotion recognition in fog computing-based systems: leveraging interpretable PCA\_CNN, YOLO with self-attention mechanism. *Frontiers in Computer Science*. 2025. Vol. 7. 1714394. DOI: <https://doi.org/10.3389/fcomp.2025.1714394>.

73. Lopez-Rojas et al. Implementation of deep reinforcement learning models for emotion detection and personalization of learning in hybrid educational environments. *Frontiers in Artificial Intelligence*. 2024. Vol. 7. 1458230. DOI: <https://doi.org/10.3389/frai.2024.1458230>.

74. Mollahosseini A., Hasani B., Mahoor M. H. AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild. *IEEE Transactions on Affective Computing*. 2017. Vol. 10(1). Pp. 18–31. URL: <http://mohammadmahoor.com/affectnet/> (дата звернення: 09.04.2026).

75. Islam M. S. Smart Retail Systems: Integrating Neuromarketing and AI for Predictive Consumer Analytics. ResearchGate. 2025. 24 p. URL: [https://www.researchgate.net/publication/403279094\\_Smart\\_Retail\\_Systems\\_Integrating\\_Neuromarketing\\_and\\_AI\\_for\\_Predictive\\_Consumer\\_Analytics](https://www.researchgate.net/publication/403279094_Smart_Retail_Systems_Integrating_Neuromarketing_and_AI_for_Predictive_Consumer_Analytics) (дата звернення: 10.04.2026).

76. Li G., Wang Y., Zhang X. Multimodal Consumer Emotion Recognition in Smart Retail: A Deep Learning Approach. *Journal of Retailing and Consumer Services*. 2024. Vol. 78. 103724. DOI: <https://doi.org/10.1016/j.jretconser.2023.103724>.

77. Colombathanthri A., Jomaa W., Chinniah Y.A. Human-centered cyber-physical systems in manufacturing industry: a systematic search and review. *Int J Adv*

*Manuf Technol.* 2025. Vol. 136. Pp. 2107 – 2141.

78. Kumar R., Singh A., Gupta P. Explainable Artificial Intelligence for Emotion Recognition in Human-Machine Interaction. *Artificial Intelligence Review.* 2024. Vol. 57(5). 112. DOI: <https://doi.org/10.1007/s10462-024-10700-1>.

79. Anna Koufakou, Elijah Nieves, and John Peller. Towards a new Benchmark for Emotion Detection in NLP: A Unifying Framework of Recent Corpora. *In Proceedings of the 2nd GenBench Workshop on Generalisation (Benchmarking) in NLP.* 2024. Pp.196 – 206.

80. Рудніцький В. Ю., Грига В. М. Кіберфізична система визначення емоційного стану людини за комп'ютером. Збірник наукових праць за матеріалами XIX Всеукраїнської науково-практичної WEB-конференції аспірантів, студентів та молодих вчених «Комп'ютерні інтелектуальні системи та мережі» (25–27 березня 2026 р.). Кривий Ріг: Криворізький національний університет. 2026. 432 с.

ДОДАТОК А  
(обов'язковий)  
ТЕЗИ ДОПОВІДІ

МІНІСТЕРСТВО  
ОСВІТИ І НАУКИ  
УКРАЇНИ

НАЦІОНАЛЬНИЙ  
ФУНДАНД  
НАУКОВОЇ ІНФОРМАЦІЇ

ІНСТИТУТ  
ІНФОРМАЦІЙНИХ  
ТЕХНОЛОГІЙ І  
ОСВІТИ

ХІХ ВСЕУКРАЇНСЬКА НАУКОВО-ПРАКТИЧНА WEB-КОНФЕРЕНЦІЯ  
АСПІРАНТІВ, СТУДЕНТІВ ТА МОЛОДИХ ВЧЕНИХ

**МАТЕРІАЛИ  
КОНФЕРЕНЦІЇ**  
*CONFERENCE PROCEEDINGS*

**КОМП'ЮТЕРНІ  
ІНТЕЛЕКТУАЛЬНІ  
СИСТЕМИ ТА МЕРЕЖІ**  
*COMPUTER INTELLIGENT SYSTEMS  
AND NETWORKS*

**КІСМ-2026**  
**CISN-2026**

КАФЕДРА КОМП'ЮТЕРНИХ СИСТЕМ ТА МЕРЕЖ

**25-27 БЕРЕЗНЯ 2026**  
КРИВИЙ РІГ / KRYVYI RIH

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
КРИВОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
КАФЕДРА КОМП'ЮТЕРНИХ СИСТЕМ ТА МЕРЕЖ

ХІХ ВСЕУКРАЇНСЬКА НАУКОВО-ПРАКТИЧНА  
WEB КОНФЕРЕНЦІЯ АСПІРАНТІВ,  
СТУДЕНТІВ ТА МОЛОДИХ ВЧЕНИХ

**КОМП'ЮТЕРНІ ІНТЕЛЕКТУАЛЬНІ  
СИСТЕМИ ТА МЕРЕЖІ**

Матеріали конференції  
25-27 березня 2026 р.

Видавничий центр  
Криворізький національний університет  
Кривий Ріг 2026

УДК 681.3.06  
К60

Відповідальний за випуск д-р техн. наук,  
професор Купін А. І.

Друкується згідно з рекомендацією Вченої Ради ФІТ Криворізького національного університету (протокол №12 від 30.03.2026 р.).

Змістова частина друкованого матеріалу збірки викладена згідно з електронними носіями, поданими авторами.

**К60 Комп'ютерні інтелектуальні системи та мережі. Матеріали XIX Всеукраїнської науково-практичної WEB конференції аспірантів, студентів та молодих вчених (25-27 березня 2026 р.). – Кривий Ріг: Криворізький національний університет, 2026. – 370 с.**

Містить матеріали науково-практичної WEB конференції аспірантів, студентів та молодих вчених з питань розробки, проектування, діагностики та моделювання комп'ютерних систем та мереж, розробки програмного та апаратного забезпечення, розглядаються проблеми створення та використання систем паралельних і розподілених обчислень, штучного інтелекту, а також питання захисту інформації.

УДК 681.3.06  
Криворізький національний університет, 2026

випадкових зовнішніх збурень. Подібні результати узгоджуються з загальними принципами теорії автоматичного керування та методами аналізу динамічних систем.

#### ЛІТЕРАТУРА

1. Грачов О. Plant monitoring system: ші-система для розумного моніторингу рослин. *Information technology and society*. 2025. № 1 (16). С. 59–64. URL: <https://doi.org/10.32689/maup.it.2025.1.7>

*Рудніцький В. Ю.*

*Хмельницький національний університет*

*Грига В.М.*

*к.т.н., доцент, Прикарпатський національний університет  
імені Василя Стефаника*

### **КІБЕРФІЗИЧНА СИСТЕМА ВИЗНАЧЕННЯ ЕМОЦІЙНОГО СТАНУ ЛЮДИНИ ЗА КОМП'ЮТЕРОМ**

*Проаналізовано сучасні підходи до створення кіберфізичних систем визначення емоційного стану людини. Висвітлено перспективи застосування мультисенсорних та адаптивних технологій для підвищення точності та ефективності таких систем у різних сферах діяльності.*

Сучасні технології все більше інтегруються в повсякденне життя людини, змінюючи способи взаємодії з комп'ютерними системами. Одним із ключових напрямів розвитку інтелектуальних систем є здатність не лише обробляти команди та дані користувача, а й аналізувати його емоційний стан. Це має особливе значення для сфер, де ефективність роботи людини прямо залежить від її психологічного стану, таких як освіта, медицина, ігрові та розважальні системи, а також системи підтримки управлінських рішень.

Дослідження останніх років демонструють розвиток мультисенсорних підходів, інтеграцію відео-, аудіо- та біометричних даних, а також використання сучасних алгоритмів аналізу інформації для підвищення точності та надійності систем. Так у роботі [1] автори узагальнюють сучасні методи визначення емоцій через аналіз відео- та аудіосигналів. Автори акцентують увагу на комбінованих підходах, де використання камер, мікрофонів та сенсорів дозволяє створювати адаптивні системи, здатні працювати у реальному часі.

Стаття [2] охоплює сучасні дослідження мультисенсорного розпізнавання емоцій на основі EEG та інших фізіологічних даних. Автори відзначають тренд до інтеграції різних типів сигналів для створення більш адаптивних і точних систем.

Отже, мета дослідження полягає у створенні інтегрованої системи, яка дозволяє автоматично визначати емоційний стан людини на основі сигналів, що фіксуються комп'ютером. До таких сигналів належать відео- та аудіопотоки, а

також фізіологічні дані користувача, отримані за допомогою сенсорних пристроїв. Кіберфізична система поєднує програмно-апаратні засоби збору та обробки даних із методами аналізу інформації, математичного моделювання та нейронними мережами для точного визначення емоційних реакцій користувача.

Структура системи включає кілька ключових компонентів.

Першим етапом є збір даних, що передбачає отримання відео з веб-камери для аналізу міміки, аудіосигналів для визначення тональності голосу, а також сигналів сенсорів для фіксації біометричних параметрів, таких як пульс, частота дихання або електропровідність шкіри.

Другий етап – обробка та виділення ознак. Для відеоаналізу використовуються алгоритми комп'ютерного зору, які дозволяють виділити ключові точки обличчя та оцінити зміни у міміці. Аудіоаналіз передбачає оцінку тембру, інтонації та ритму мовлення, що також відображає емоційний стан користувача. Біометричні сигнали аналізуються для визначення рівня стресу, збудження або розслаблення.

Третій етап – класифікація емоційного стану, що реалізується за допомогою нейронних мереж, що дозволяє точно визначити базові емоції користувача на основі аналізу відео-, аудіо- та біометричних даних.

Важливим аспектом є інтерфейс взаємодії з користувачем. Система повинна не лише визначати емоції, а й відображати їх у зручному форматі для подальшої інтеграції з іншими інформаційними системами або застосування у реальному часі. Це відкриває можливості для адаптивних систем навчання, інтерактивних ігор та персоналізованих сервісів, які враховують психологічний стан користувача.

Переваги кіберфізичних систем визначення емоцій очевидні. По-перше, вони забезпечують роботу у реальному часі, що дозволяє системам оперативно реагувати на зміни у поведінці користувача. По-друге, вони підвищують адаптивність систем до індивідуальних особливостей людини, що важливо для персоналізації процесів взаємодії. По-третє, такі системи відкривають нові можливості для досліджень у галузі психології, педагогіки та управління.


## ВИСНОВКИ

Кіберфізичні системи визначення емоційного стану людини є перспективним напрямом розвитку інтелектуальних технологій, здатних підвищувати ефективність взаємодії людини і комп'ютера. Сучасні дослідження показують, що мультисенсорні підходи та адаптивні алгоритми дозволяють підвищити точність і надійність визначення емоцій у реальному часі. Подальший розвиток таких систем сприятиме їх інтеграції у різні сфери життя, включаючи освіту, медицину, ігрову індустрію та сервіси підтримки рішень.

## ЛІТЕРАТУРА

1. Pereira R, Mendes C, Ribeiro J, Ribeiro R, Miragaia R, Rodrigues N, Costa N, Pereira A. Systematic Review of Emotion Detection with Computer Vision and Deep Learning. *Sensors*. 2024; 24(11):3484. <https://doi.org/10.3390/s24113484>.

**ДОДАТОК Б**  
(обов'язковий)  
**ПРЕЗЕНТАЦІЯ**



*Кіберфізична система визначення емоційного  
стану людини за комп'ютером\_*

Виконав: магістр гр. КІ2М-24-2 Рудніцький Віталій

Керівник: к.т.н., доц. Грига В.М.



## *Актуальність*

У традиційних системах взаємодія людини з комп'ютером базується на введенні явних команд через клавіатуру, мишу або сенсорні інтерфейси. Такий підхід не враховує такий аспект людської комунікації як невербальні сигнали, до яких належать міміка, жести, інтонація мовлення та фізіологічні реакції. Саме ці сигнали відіграють ключову роль у передачі емоційного стану та внутрішніх намірів людини. Ігнорування цього аспекту призводить до зниження ефективності взаємодії, особливо в умовах підвищеного когнітивного навантаження, коли користувач може відчувати втому, стрес або перевантаження. Як наслідок, зростає кількість помилок, знижується продуктивність праці та погіршується загальний досвід взаємодії з інформаційною системою.

Актуальність роботи зумовлена необхідністю створення людино-орієнтованих кіберфізичних систем, які здатні ефективно взаємодіяти з користувачем, враховуючи його емоційний стан. У сучасному цифровому суспільстві, де значна частина діяльності відбувається у віртуальному середовищі, питання психоемоційного комфорту стає не менш важливим, ніж технічні характеристики систем.

### Мета дослідження

Підвищення точності розпізнавання емоційного стану користувача шляхом розробки кіберфізичної системи, що базується на аналізі аудіо-візуальних та біометричних сигналів.

### Об'єкт дослідження

Процес розпізнавання та визначення емоційного стану людини під час її взаємодії з комп'ютерними системами у реальному часі.

### Предмет дослідження

Методи, алгоритми та програмно-апаратні засоби побудови кіберфізичної системи для мультимодального розпізнавання емоцій на основі комплексного аналізу відеопотоку, аудіосигналу та біометричних даних користувача.

### Задачі дослідження

- аналіз сучасних методів та комерційних рішень у сфері розпізнавання емоцій із виділенням їхніх обмежень;
- моделювання процесу визначення емоційного стану людини в межах кіберфізичної системи;
- розроблення методу класифікації емоцій на основі гібридної нейромережевої моделі, здатної враховувати часову динаміку стану користувача;
- проєктування архітектури кіберфізичної системи, що забезпечує синхронний збір даних з камери, мікрофона та біометричних сенсорів;
- проведення експериментального дослідження розробленої системи та оцінювання її ефективності у порівнянні з існуючими одноmodalними підходами.

### Наукова новизна

-удосконалено метод визначення емоційного стану, який, на відміну від існуючих методів, базується на комплексному аналізі відео- та аудіосигналів разом із фізіологічними даними користувача (пульс, частота дихання, електропровідність шкіри);

-набула подальшого розвитку структура кіберфізичної системи, яка за рахунок інтеграції програмно-апаратних засобів та нейронних мереж забезпечує автоматичну класифікацію емоцій у реальному часі.

### Практичне значення дослідження

полягає у можливості застосування розробленої кіберфізичної системи в різних галузях. Зокрема, в освітніх технологіях такі системи можуть адаптувати навчальний процес відповідно до рівня концентрації та втоми студента, підвищуючи ефективність засвоєння матеріалу. У сфері розробки програмного забезпечення вони можуть використовуватися для моніторингу стану розробників з метою своєчасного виявлення ознак професійного вигорання. У системах безпеки та управління виробництвом аналіз емоційного стану операторів дозволяє знизити ризик помилок, спричинених стресом або перевтомою.

**Сфери використання  
технології розпізнавання  
емоцій**

Освіта (EdTech)	Система бачить, коли студент втрачає концентрацію або не розуміє матеріал.	Адаптація складності контенту, виклик викладача або зміна темпу лекції.
Маркетинг	Аналіз реакції фокус-груп на рекламні ролики в реальному часі.	Оптимізація бюджету: залишають тільки ті кадри, що викликають «правильну» емоцію.
HR та розробка ПЗ	Моніторинг вигорання та рівня стресу розробників під час написання коду.	Покращення корпоративної культури, запобігання помилкам через втому.
Медицина	Допомога людям з розладами аутичного спектру (РАС) розуміти емоції співрозмовника.	Соціальна реабілітація та підтримка ментального здоров'я.
Геймінг	Зміна ігрового сюжету або складності залежно від	Максимальне занурення (Immersive experience).

**Координатна модель емоційних станів (модель Рассела)**



Функціонування кіберфізичної системи визначення емоційного стану людини може бути формалізоване у вигляді багаторівневої математичної моделі, яка відображає послідовність перетворення вхідних даних у кінцевий результат. Нехай на вході системи формується множина сенсорних даних, виду 1:

$$D(t) = \{V(t), A(t), B(t)\}, \quad (1)$$

де  $V(t)$  – візуальні дані (зображення обличчя);

$A(t)$  – аудіальні дані (мовний сигнал);

$B(t)$  – поведінкові параметри (взаємодія з комп'ютером);

$t$  – момент часу.

Процес визначення емоційного стану  $E$  можна представити як функцію від множини вхідних сигналів, що фіксуються комп'ютером, представлену формулою 2:

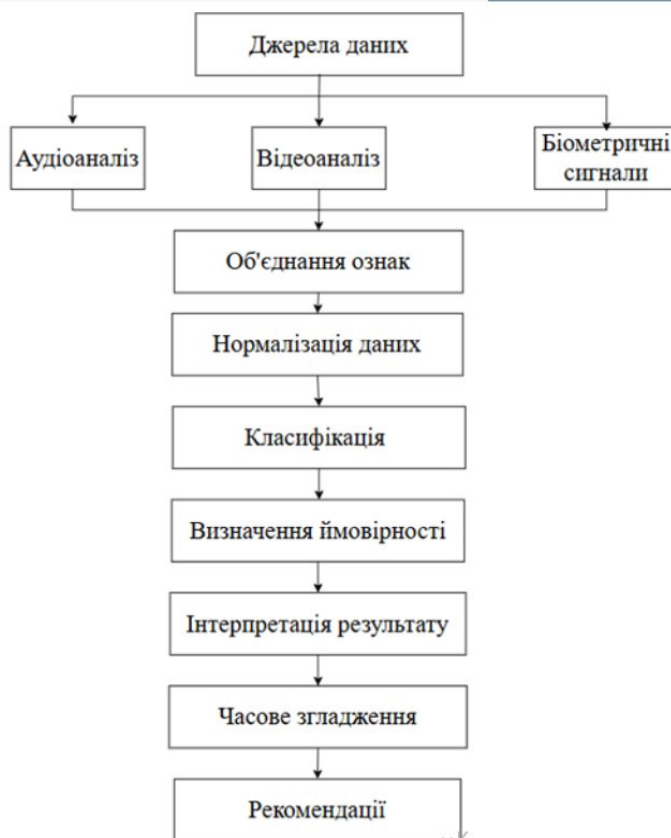
$$E = f(V, A, B), \quad (2)$$

де  $V$  – вектор ознак відеопотоку (міміка обличчя);

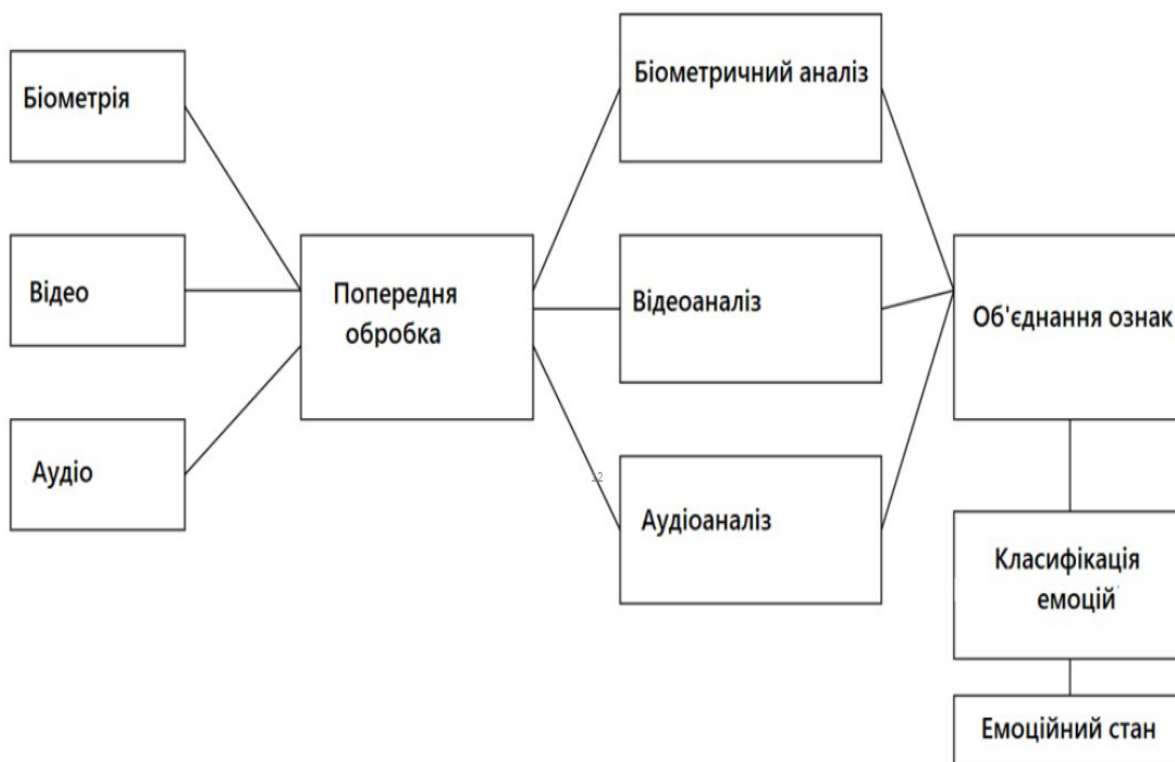
$A$  – вектор ознак аудіосигналу (параметри голосу);

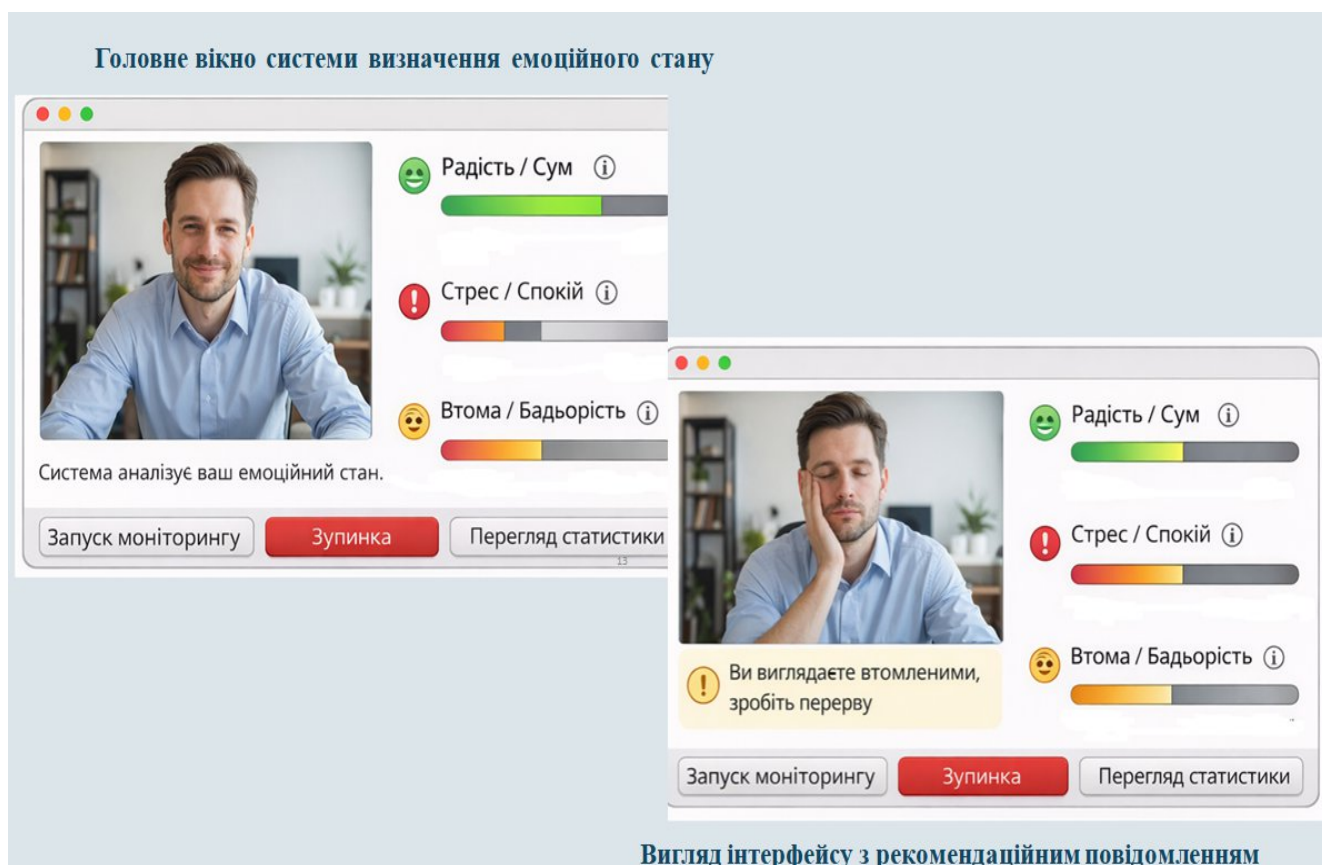
$B$  – вектор біометричних параметрів (пульс, провідність шкіри).

**Структурно-логічна схема  
методу визначення емоційного  
стану людини**



**Структурна схема кіберфізичної системи розпізнавання емоцій людини за комп'ютером**





У даній матриці рядки відповідають істинним емоціям користувача, а стовпці відповідають передбаченим моделлю класам. Значення на головній діагоналі відображають відсоток правильно класифікованих випадків для кожного стану.

		Передбачений клас (Predicted Class)				
		Нейтральний	Радість	Сум	Втома	Стрес
Істинний клас (True Class)	Нейтральний	94%	2	2	1	1%
	Радість	3	93%	1	1	2%
	Сум	3	1% <sup>14</sup>	78%	15%	3%
	Втома	2	1%	16%	76%	5%
	Стрес	4	3	2	3%	88%

## *Висновки*

1. Проаналізовано предметну область, а також існуючі сучасні методи та комерційні рішення у сфері розпізнавання емоцій із виділенням їхніх обмежень.
2. Проведено моделювання процесу визначення емоційного стану людини в межах кіберфізичної системи. Розгляд задачі здійснювався з позицій системного підходу, що дозволило формалізувати структуру, функції та основні етапи обробки інформації.
3. Розроблено структурну модель кіберфізичної системи, яка включає три основні рівні: рівень сприйняття, рівень обробки та рівень взаємодії. Визначено функціональне призначення кожного рівня та їх взаємозв'язки, що забезпечують безперервний цикл збору, обробки та інтерпретації даних.
4. Розроблено та обґрунтовано метод визначення емоційного стану людини за комп'ютером, що базується на комплексному аналізі мультимодальних даних.
5. Здійснено проектування та практичну реалізацію кіберфізичної системи визначення емоційного стану, а також проведено експериментальну перевірку її ефективності.

*Дякую!*

### Протокол аналізу звіту подібності експертом

Заявляю, що я ознайомився (-лась) з Повним звітом подібності, який був згенерований Системою виявлення і запобігання плагіату щодо роботи:

**Автор:** Віталій РУДНІЦЬКИЙ

**Співавтор:**

**Назва:** Кіберфізична система визначення емоційного стану людини за комп'ютером

**Експерт:** Володимир ГРИГА

**Підрозділ:** Кафедра комп'ютерної інженерії та інформаційних систем

**Коефіцієнт подібності 1:** 4.27%

**Коефіцієнт подібності 2:** 1.18%

**Мікропробіли:** 130

**Заміна букв:** 3

**Інтервали:** 0

**Білі знаки:** 6

**Дата створення звіту:** 2026-04-20 10:25:06.0

Після аналізу Звіту подібності констатую наступне:

Запозичення, виявлені в роботі є законними і не є плагіатом. Рівень подібності не перевищує допустимої межі. Таким чином робота незалежна і приймається.

Запозичення не є плагіатом, але перевищено граничне значення рівня подібностей. Таким чином робота повертається на доопрацювання.

Виявлено запозичення і плагіат або навмисні текстові спотворення (маніпуляції), як передбачувані спроби укриття плагіату, які роблять роботу невідповідною вимогам законодавства (Ст. 32. ЗУ Про вищу освіту, пункт 3.1, Ст. 42. ЗУ Про освіту) та вимог НАЗЯВО (Критерій 5), а також кодексу етики і процедур. Таким чином робота не приймається.

Обґрунтування:

2026-04-20

Дата



Доцент Андрій Нічепорук

експерт

# Anti-Plagiarism (<http://ap.km.ua>) v-15.701

Максимальне співпадіння з одним документом 1.0%

Словники перевірки: en\_US, ru\_RU, ua\_UA. Помилки в документах: 9%

ID: 270526 Назва: МКР Кіберфізична система визначення емоційного стану людини за комп'ютером Додано в БД: 2026-04-20 Автора: Віталій РУДНІЦЬКИЙ Керівники: Володимир ГРИГА Консультанти: Опоненти:	Документ		Сумарний збіг по Базі Даних	
	Символи	Лексеми	Символи	Лексеми
	162559	1173	2559 (2%)	26 (2%)

Джерело плагіату

ID	Опис	Наявність плагіату в документі	
		Символи	Лексеми

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

РЕЦЕНЗІЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ МАГІСТРА

Здобувач: Віталій РУДНІЦЬКИЙ

Тема: Кіберфізична система визначення емоційного стану людини за комп'ютером

Спеціальність: 123 «Комп'ютерна інженерія»

Обсяг кваліфікаційної роботи магістра:

Кількість листів креслень —; кількість сторінок записки 94

1. Короткий зміст роботи та прийнятих рішень У роботі запропоновано кіберфізичну систему визначення емоційного стану людини за комп'ютером

2. Висновок про відповідність роботи дипломному завданню Кваліфікаційна робота магістра відповідає виданому завданню

3. Характеристика виконання кожного розділу, ступінь використання останніх досягнень науки і техніки і передових методів роботи: У першому розділі проведено аналіз предметної області, розглянуто існуючі комерційні рішення та наукові рішення, а також визначено обмеження одномодальних підходів.

У другому розділі змодельовано процес розпізнавання емоцій, розроблено структурну модель КФС, що включає рівні сприйняття, обробки та взаємодії; формалізовано математичні моделі відео- та аудіоаналізу.

У третьому розділі запропоновано мультимодальний метод розпізнавання емоцій на основі гібридної нейромережевої архітектури CNN-LSTM та описано механізми забезпечення конфіденційності даних.

У четвертому розділі здійснено програмну реалізацію системи на мові Python та проведено експериментальні дослідження, які підтвердили середню точність розпізнавання на рівні 85,8%.

4. Позитивні сторони роботи: Запропонована кіберфізична система визначення емоційного стану людини за комп'ютером дозволяє визначити емоційний стан людини що дозволяє підвищити ефективність навчання, запобігти професійному вигоранню та мінімізувати ризики помилок через стрес чи втому.

5. Негативні сторони роботи: Не наведене детальне пояснення результатів експериментальних досліджень та аналізу матриці плутанини коли виявляються значні труднощі системи у розрізненні емоцій з подібними характеристиками.

6. Оцінка графічного оформлення та пояснювальної записки роботи: -

7. Відгук про роботу в цілому: В загальному робота виконана на достатньо високому професійному рівні.

8. Інші зауваження: -

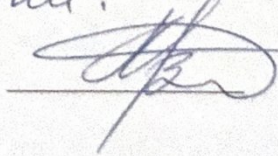
9. Оцінка кваліфікаційної роботи магістра:

Розглянувши позитивні та негативні сторони представленої кваліфікаційної роботи магістра вважаю, що робота заслуговує оцінки «добре» 75.00 (С)

Рецензент (прізвище, ім'я, по батькові, посада, місце роботи) \_\_\_\_\_

Мартишук В. В., д.т.н., проф., професор кафедри автоматизації комп'ютерно-інформаційних технологій та робототехніки.

"1" 05 2026р.



Зав. кафедри КПС  
д-р. філософії Ользі ПАВЛОВІЙ

Віталій РУДНІЦЬКИЙ

ІІІЬ здобувача вищої освіти

ФІТ, 2 курсу, групи КІ2м-24-2

### ЗАЯВА

З правилами чинного Положення про систему забезпечення академічної доброчесності у Хмельницькому національному університеті, згідно з яким виявлення академічного плагіату є підставою для відмови в допуску кваліфікаційної роботи до захисту і застосування заходів академічної відповідальності, ознайомлений (а). Про використання спеціалізованих програмних засобів (СПЗ) StrikePlagiarism та Anti-Plagiarism для перевірки кваліфікаційних робіт здобувачів вищої освіти на наявність академічного плагіату оповіщений (а). Надаю університету право на передачу моєї роботи для обробки та збереження в базах даних СПЗ і використання роботи для виявлення академічного плагіату в інших роботах, які перевіряються СПЗ.

Також надаю свою згоду на обробку й збереження університетом моєї роботи в Інституційному репозитарії Хмельницького національного університету.

Робота надається для перевірки в електронному варіанті. Електронна версія моєї роботи збігається (ідентична) з друкованою.

1 травня 2026 року



## РІШЕННЯ ЕКСПЕРТНОЇ КОМІСІЇ

КАФЕДРИ КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ ТА ІНФОРМАЦІЙНИХ СИСТЕМ  
ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ ДО ЗАХИСТУ

Назва кваліфікаційної роботи Кіберфізична система визначення емоційного стану людини за комп'ютером

Автор Віталій РУДНИЦЬКИЙ

Освітня програма Комп'ютерна інженерія та програмування

Рівень вищої освіти другий (магістерський)

Спеціальність 123 Комп'ютерна інженерія

Науковий керівник: к.т.н., доцент Володимир ГРИГА

На основі аналізу кваліфікаційної роботи на дотримання вимог академічної доброчесності (у т.ч. відсутності ознак академічного плагіату) з урахуванням результатів перевірки роботи спеціалізованим програмним засобом(ами) комісія зробила такий висновок:

№	Висновок	Позначка про відповідність
1	Ознаки академічного плагіату	
1.1	Запозичення, виявлені в роботі, є законними і не є академічним плагіатом (далі – зазначаються підстави віднесення запозичень до правомірних, якщо потрібно). Робота приймається до захисту.	відповідає
1.2	Виявлені запозичення не є академічним плагіатом, розміщені в розділах, які не описують безпосередньо авторське дослідження, але кількість цитат перевищує обсяг, виправданий поставленою метою роботи (далі – зазначаються детальні та аргументовані підстави віднесення запозичень до правомірних). Робота приймається до захисту, але має бути відкоригована.	
1.3	Виявлені запозичення не є академічним плагіатом, але частково розміщені в розділах, які описують безпосередньо авторське дослідження, а кількість цитат перевищує обсяг, виправданий поставленою метою роботи. Робота може бути допущена до захисту після того як буде відкоригована та доопрацьована і успішно пройде повторну перевірку на академічний плагіат.	
1.4	Робота містить навмисні текстові спотворення, передбачувані спроби укриття текстових запозичень або інші прояви академічного плагіату. Робота містить фабрикацію або фальсифікацію даних. Робота не допускається до захисту.	
2	Інші види порушень академічної доброчесності	

## Підтвердження:

Запозичення, виявлені в роботі, є законними і не є плагіатом, оскільки:


- 1) усі запозичення фрагментарні, або мають належним чином оформленні посилання;
  - 2) окремі виявлені збіги є загальноживаними фразами або виразами, про що свідчить посилання системи на збіг з джерелами на один фрагмент речення;
  - 3) всі зафіксовані системою ознаки модифікації тексту відносяться до комбінування латинських символів зі україномовними скороченнями індексів в формулах, що не є модифікацією тексту.
  - 4) значна частина знайденого плагіату відноситься до списку використаних джерел
- Сумарний обсяг всіх запозичень, визначений системою виявлення збігів/ ідентичності/схожості StrikePlagiarism, складає 4.27 % і адресується до 42 першоджерела; та системою Anti-Plagiarism складає 1%, що, з урахуванням наведених обґрунтувань, відповідає характеру наукового дослідження і свідчить на користь кваліфікаційної роботи.


01.05.2026


Завідувач кафедри

Гарант освітньої програми

Керівник кваліфікаційної роботи

  
Підпис

  
Підпис

  
Підпис

Ольга ПАВЛОВА  
Ім'я, ПРІЗВИЩЕ

Олег САВЕНКО  
Ім'я, ПРІЗВИЩЕ

Володимир ГРИГА  
Ім'я, ПРІЗВИЩЕ