Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя
(повне найменування вищого навчального закладу)
Факультет комп'ютерно-інформаційних систем і програмної інженерії
(назва факультету)
Кафедра комп'ютерних систем та мереж
(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

магістра
(освітній ступінь))

на тему: **Методи та програмно-апаратні засоби біометричної аутентифікації у реальному часі користувачів комп'ютерної системи**

Виконав: студент (ка) 6 курсу, групи СІм-61
спеціальності 123 «Комп'ютерна інженерія»
(шифр і назва спеціальності)

	(підпис)	Гаврада Д.М. (прізвище та ініціали)
Керівник	(підпис)	Луцків А.М. (прізвище та ініціали)
Нормоконтроль	(підпис)	Луцик Н.С. (прізвище та ініціали)
Завідувач кафедри	(підпис)	Осухівська Г.М. (прізвище та ініціали)
Рецензент	(підпис)	(прізвище та ініціали)

Тернопіль
2025

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя
(повне найменування вищого навчального закладу)
Факультет комп'ютерно-інформаційних систем і програмної інженерії
Кафедра комп'ютерних систем та мереж

ЗАТВЕРДЖУЮ
Завідувач кафедри *Осухівська Г.М.*

«*17*» *листопада* 2025 р.

З А В Д А Н Н Я

НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня *магістр*
(назва освітнього ступеня)

за спеціальністю *123 «Комп'ютерна інженерія»*
(шифр і назва спеціальності)

студенту *Гавраді Дмитру Миколайовичу*
(прізвище, ім'я, по-батькові)

1. Тема проекту (роботи) *Методи та програмно-апаратні засоби біометричної аутентифікації у реальному часі користувачів комп'ютерної системи*

Керівник проекту (роботи) *Луцків Андрій Мирославович, к.т.н., доц.*
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «*14*» *листопада* 2025 року №*4/7-987*

2. Термін подання студентом завершеної роботи

3. Вихідні дані до роботи *Види біометричної аутентифікації людини, алгоритми машинного навчання розпізнавання образів, методи розпізнавання голосу, характеристики і види мікроконтролерів*

4. Зміст роботи (перелік питань, які потрібно розробити)
Вступ. 1. Аналіз методів та алгоритмів біометричної аутентифікації особи. 2.Проектування архітектури системи мультимодальної біометричної аутентифікації користувача комп'ютерної системи. 3. Апаратне і програмне забезпечення комплексу мультимодальної біометричної аутентифікації користувачів комп'ютерної системи 4. Охорона праці та безпека в надзвичайних ситуаціях. Висновки

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)
1. Актуальність і мета дослідження. 2. Задачі дослідження, об'єкт і предмет, наукова новизна і практична цінність дослідження. 3. Класифікація методиів біометричної ідентифікації та аутентифікації 4. Архітектура біометричної системи аутентифікації. 5. Алгоритми опрацювання модальностей. 6. Структура нейромереж біометричної системи. 7. Архітектура програмного забезпечення. 8. Висновки

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці	*Осухівська Г.М., зав.каф. КС*	*04.12.2025*	
Безпека в надзвичайних ситуаціях	*Стручок В.С., ст. викладач каф. ОХ*		

7. Дата видачі завдання *17.11.2025 р.*

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	*Огляд літературних джерел. Постановка завдання на кваліфікаційну роботу.*	*17.11.2025р. – 23.11.2025р.*	*Викон.*
2.	*Проектування архітектури системи мультимодальної біометричної аутентифікації користувача комп'ютерної системи*	*24.11.2025р. – 02.12.2025р.*	*Викон.*
3.	*Апаратне і програмне забезпечення комплексу мультимодальної біометричної аутентифікації користувачів комп'ютерної системи*	*03.12.2025р. – 10.12.2025р.*	*Викон.*
4.	*Охорона праці та безпека в надзвичайних ситуаціях*	*04.12.2025р. – 10.12.2025р.*	*Викон.*
5.	*Оформлення пояснювальної записки і графічного матеріалу*	*11.12.2025р. – 19.12.2025р.*	*Викон.*
6.	*Попередній захист кваліфікаційної роботи магістра*	*16.12.2025р.*	*Викон.*
7.	*Захист кваліфікаційної роботи магістра*	22.12.2025 р.	*Викон.*

Студент _____ *Гаврада Д.М.*
(підпис) (прізвище та ініціали)

Керівник роботи _____ *Луцків А.М.*
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Гаврада Д.М. Методи та програмно-апаратні засоби біометричної аутентифікації у реальному часі користувачів комп’ютерної системи: робота на здобуття кваліфікаційного ступеня магістра: спец. 123 — комп’ютерна інженерія / наук. кер. Луцків А.М. Тернопіль: Тернопільський національний технічний університет імені Івана Пулюя, 2025.

Ключові слова: метод, засіб, біометрична аутентифікація, комп’ютерна система, реальний час.

У роботі визначено структурні компоненти апаратної та програмної частини комплексу біометричної мультимодальної аутентифікації користувачів комп’ютерної системи, а також встановлено вимоги до їх сумісності, що дало змогу побудувати узагальнену архітектуру системи та забезпечити обробку зображення обличчя і голосового сигналу в режимі реального часу. Розроблено формалізований опис процесу мультимодальної аутентифікації, який включає визначення простору вхідних даних, моделей відображення для кожної модальності, метрик подібності та критеріїв прийняття рішень.

Розроблено архітектуру аитьпаратного забезпечення апаратно-програмного комплексу мультимодальної біометричної аутентифікації на базі одноплатного комп’ютера Raspberry Pi 4 Model B із використанням модуля камери, USB-мікрофона та дисплейного інтерфейсу.

Реалізовано нейронні мережі для розпізнавання обличчя та голосу, зокрема MTCNN для детекції та вирівнювання облич, MobileFaceNet для формування векторів ознак обличчя та ECAPA-TDNN для отримання векторів ознак голосу для досягнення високої точності аутентифікації при обмежених апаратних ресурсах одноплатного комп’ютера.

ABSTRACT

Havrada D.M. Methods and hardware–software tools for real-time biometric authentication of computer system users. Master's Graduation Thesis: speciality 123 - Computer engineering/supervisor Lutskiv A.M. Ternopil: Ternopil Ivan Puluj National Technical University, 2025.

Keywords: method, tool, biometric authentication, computer system, real-time.

In this work, the structural components of the hardware and software parts of a multimodal biometric user authentication system for a computer system are identified, and the requirements for their compatibility are established. This made it possible to develop a generalized system architecture and to ensure real-time processing of facial images and voice signals. A formalized description of the multimodal authentication process is developed, including the definition of the input data space, mapping models for each modality, similarity metrics, and decision-making criteria.

The hardware architecture of the multimodal biometric authentication system is developed based on the Raspberry Pi 4 Model B single-board computer, using a camera module, a USB microphone, and a display interface.

Neural networks for face and voice recognition are implemented, including MTCNN for face detection and alignment, MobileFaceNet for generating facial feature vectors, and ECAPA-TDNN for extracting voice feature embeddings, which makes it possible to achieve high authentication accuracy under limited computational resources of the single-board computer.

ЗМІСТ

ВСТУП

Актуальність роботи. У сучасних умовах цифровізації суспільства питання безпечної аутентифікації користувачів набуває особливого значення. Традиційні методи ідентифікації та аутентифікації такі, як паролі, картки доступу чи PIN-коди мають суттєві недоліки. Це пов’язано з можливістю їх втрати, підробки або компрометації записаних у них даних. Сучасні біометричні технології, що базуються на унікальних фізіологічних або поведінкових характеристиках людини, забезпечують більш високий рівень безпеки, зручності та надійності.

Розвиток обчислювальних засобів, мікроелектроніки та штучного інтелекту дозволяє реалізовувати біометричні системи аутентифікації в реальному часі, інтегруючи їх у комп’ютеризовані комплекси контролю доступу, фінансові сервіси, системи моніторингу чи персоналізовані інформаційні середовища. Особливої актуальності розробка і вдосконалення методів і засобів біометричної ідентифікації та аутентифікації набуває у сфері кіберфізичних систем, де швидкість прийняття рішень і достовірність розпізнавання критично важливі.

В Україні дослідженнями у галузі біометричної ідентифікації та аутентифікації займаються А. Луцків, С. Лупенко, І. Коваль, В. Терещенко, А. Іванов, які досліджують питання алгоритмів розпізнавання образів, обробки зображень та інтеграції біометричних систем у комп’ютерно-інтегровані комплекси. Серед зарубіжних учених значний внесок зробили Anil K. Jain, Josef Bigun, Zhenan Sun, David Zhang, John Daugman, які розробили фундаментальні моделі й алгоритми біометричного розпізнавання за відбитками пальців, геометрією обличчя та райдужкою ока.

З огляду на тенденції розвитку комп’ютеризованих інтелектуальних систем безпеки, розробка ефективних методів і апаратно-програмних комплексів для біометричної аутентифікації користувачів у реальному часі є актуальною науково-практичною задачею.

Метою кваліфікаційної роботи є розробка та дослідження методів і засобів побудови апаратно-програмного комплексу, які забезпечують ідентифікацію та

аутентифікацію користувачів у реальному часі на основі біометричних ознак (обличчя та голосу) з використанням сучасних методів комп’ютерного зору та глибинного навчання.

Завдання кваліфікаційної роботи:

– провести аналіз сучасних методів і технологій біометричної аутентифікації користувачів та визначити їхні переваги й недоліки;

– побудувати узагальнену архітектуру апаратно-програмного комплексу біометричної аутентифікації в реальному часі;

– запропонувати та обґрунтувати математичну модель процесу біометричної аутентифікації, яка формалізує перетворення вхідного сигналу у простір ознак та прийняття рішення за критерієм подібності;

– визначити критерії якості функціонування системи;

– обґрунтувати вибір апаратної та програмної платформи, що забезпечує реалізацію системи у режимі реального часу;

– розробити програмні модулі для обробки, аналізу та порівняння біометричних ознак, використовуючи методи комп’ютерного зору й глибинного навчання.

– реалізувати прототип апаратно-програмного комплексу біометричної аутентифікації;

Об’єкт дослідження: процес біометричної аутентифікації користувачів у комп’ютеризованих системах безпеки.

Предмет дослідження: методи, алгоритми і технічні засоби реалізації апаратно-програмного комплексу біометричної аутентифікації в реальному часі.

Методи дослідження: Поставлені у роботі задачі розв’язано з використанням наступних методів:

– аналіз та узагальнення – при дослідженні сучасних методів і алгоритмів біометричної аутентифікації за фізіологічними та поведінковими ознаками, а також при обґрунтуванні вибору мультимодального підходу до аутентифікації особи;

– методи машинного навчання та глибинного навчання – при побудові моделей розпізнавання обличчя та голосу на основі нейронних мереж MTCNN, MobileFaceNet, ECAPA-TDNN;

– методи математичної статистики – при оцінюванні якості біометричної аутентифікації, аналізі показників точності та при визначенні порогових значень прийняття рішень;

– методи проєктування апаратно-програмних систем – при розробці архітектури апаратно-програмного комплексу на базі одноплатного комп'ютера Raspberry Pi, проєктуванні структури програмного забезпечення, бази даних та веб-орієнтованої адміністративної панелі;

– програмування та моделювання – при реалізації алгоритмів обробки зображень і аудіосигналів, інтеграції нейронних мереж у програмне середовище, а також при розгортанні системи в реальних умовах експлуатації;

– експериментальні дослідження – при тестуванні розробленого апаратно-програмного комплексу, оцінюванні його швидкодії, точності аутентифікації та стійкості до змін умов зйомки і запису голосу.

Наукова новизна дослідження. Наукова новизна результатів дослідження полягає в наступному:

– уперше для заданих обчислювальних обмежень Raspberry Pi 4 обґрунтовано оптимальний спосіб поєднання двох незалежних біометричних ознак на основі моделей MobileFaceNet та ECAPA-TDNN, що дало змогу забезпечити підвищення точності при збереженні вимоги до низького рівня затримки опрацювання сигналу та забезпечити стабільність роботи системи у реальному часі навіть за наявності шумів і спотворень;

– уперше визначено та формалізовано критерії оцінювання роботи системи, включно з метриками FAR, FRR, EER та вимогами до максимально допустимої затримки обробки, що дає змогу забезпечити баланс між точністю і швидкодією системи.

Практичне значення результатів кваліфікаційної роботи. Практична цінність результатів кваліфікаційної роботи полягає у розробці та реалізації

архітектури апаратного і програмного забезпечення біометричної системи аутентифікації користувачів комп’ютерної системи на основі мінікомп’ютера Raspberry PI 4 і стеку технологій на базі Python.

Публікації. Результати кваліфікаційної роботи апробовані на XIV міжнародній науково - технічній конференції молодих учених і студентів «Актуальні задачі сучасних технологій» (11-12 грудня 2025 р.) Тернопільського національного технічного університету імені Івана Пулюя та на XIII науково-технічній конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі, системи та технології» (17-18 грудня 2025 року) як тези конференцій.

Структура роботи. Робота складається з пояснювальної записки та графічної частини. Пояснювальна записка складається з вступу, 4 розділів, висновків, списку використаних джерел та додатку.

РОЗДІЛ 1
АНАЛІЗ МЕТОДІВ ТА АЛГОРИТМІВ БІОМЕТРИЧНОЇ АУТЕНТИФІКАЦІЇ ОСОБИ

1.1. Основні поняття та класифікація методів біометричної аутентифікації

Біометрична ідентифікація – це процес розпізнавання особи на основі її унікальних фізіологічних або поведінкових характеристик [1]. Вона є одним із найперспективніших напрямів розвитку систем інформаційної безпеки комп'ютеризованих систем, оскільки дозволяє виконувати аутентифікацію без необхідності запам'ятовування паролів чи носіння фізичних носіїв.

В основі біометричної аутентифікації лежить принцип, за яким біометричний шаблон користувача формується з набору вимірюваних ознак, отриманих під час реєстрації, і згодом порівнюється з поточними вимірюваннями [2].

Розвиток мікропроцесорних систем, сенсорних технологій і методів штучного інтелекту зробив можливим використання біометрії не лише у великих корпоративних структурах, а й у побутових пристроях, зокрема, смартфонах, банкоматах, системах «розумний дім» тощо.

Методи біометричної аутентифікації поділяють за двома основними критеріями [3]:

– за природою ознак – фізіологічні та поведінкові;

– за способом аналізу – одноканальні (монобіометричні) та багатоканальні (мультимодальні).

До фізіологічних характеристик при біометричній аутентифікації людини належать характеристики, які пов'язані з її анатомічними властивостями. До них входять: геометрія обличчя, відбитки пальців, сітківка, райдужна оболонка ока, форма вуха, структура вен [4].

Поведінкові характеристики відображають динаміку дій, які можуть включати почерк, ходу, голос, ритм натискання клавіш тощо.

На рис. 1.1 показано деталізовану класифікацію біометричних методів аутентифікації та відповідних систем за типами ознаками людини.

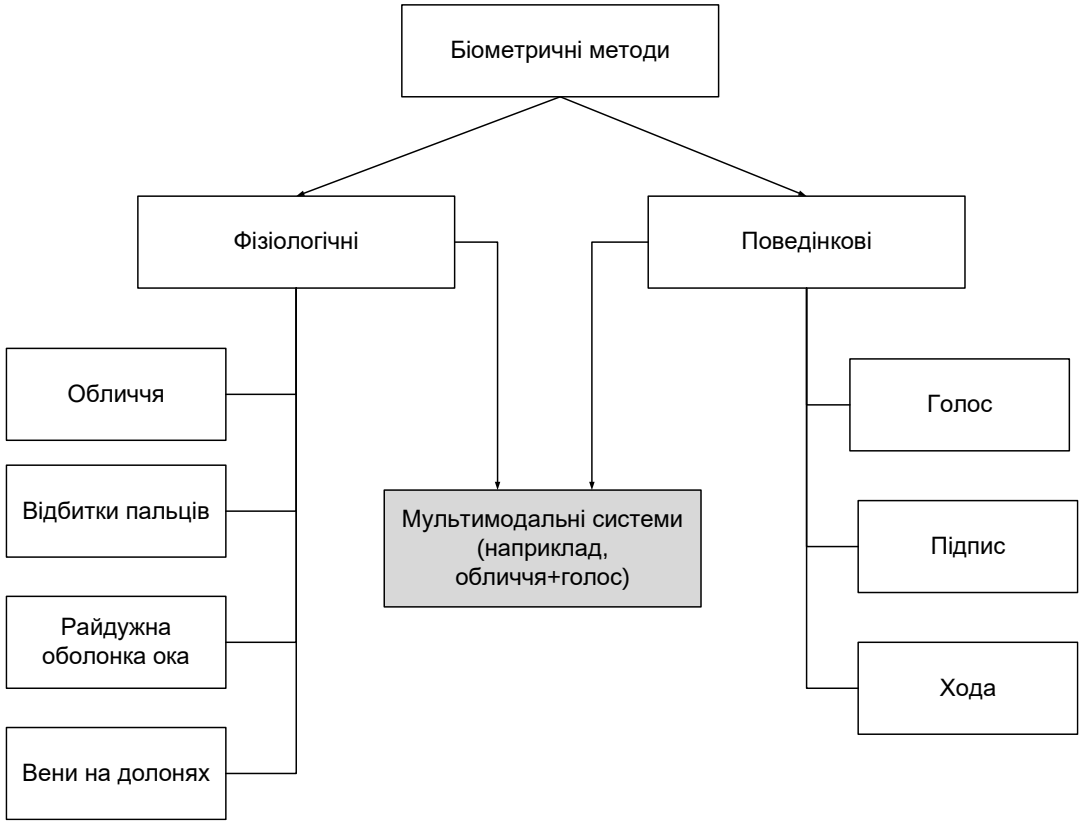


Рис. 1.1. Класифікація біометричних систем за типами ознак

На сьогодні розроблено досить багато різних методів, які дозволяють проводити біометричну ідентифікацію особи на основі її фізіологічних властивостей.

Один з таких методів дає змогу проводити ідентифікацію на основі аналізу відбитків пальців. Він є найдавнішим і найпоширенішим методом, що використовує аналіз ліній папілярного узору. До переваг такого методу належить висока точність розпізнавання та невеликий об’єм даних, необхідних для проведення процедури аутентифікації. Найбільшим недоліком цього методу є висока чутливість до забруднення або пошкодження шкіри [5].

Більш сучасні методи аналізу фізіологічних властивостей базуються на розпізнаванні людини за геометрією її обличчя. При цьому враховуються риси обличчя, зокрема, відстані між очима, носом, ротом, підборіддям.

Цей метод став базовим у системах відеоспостереження та смартфонах. Типова схема архітектури системи біометричної аутентифікації на основі аналізу обличчя показана на рис. 1.2.

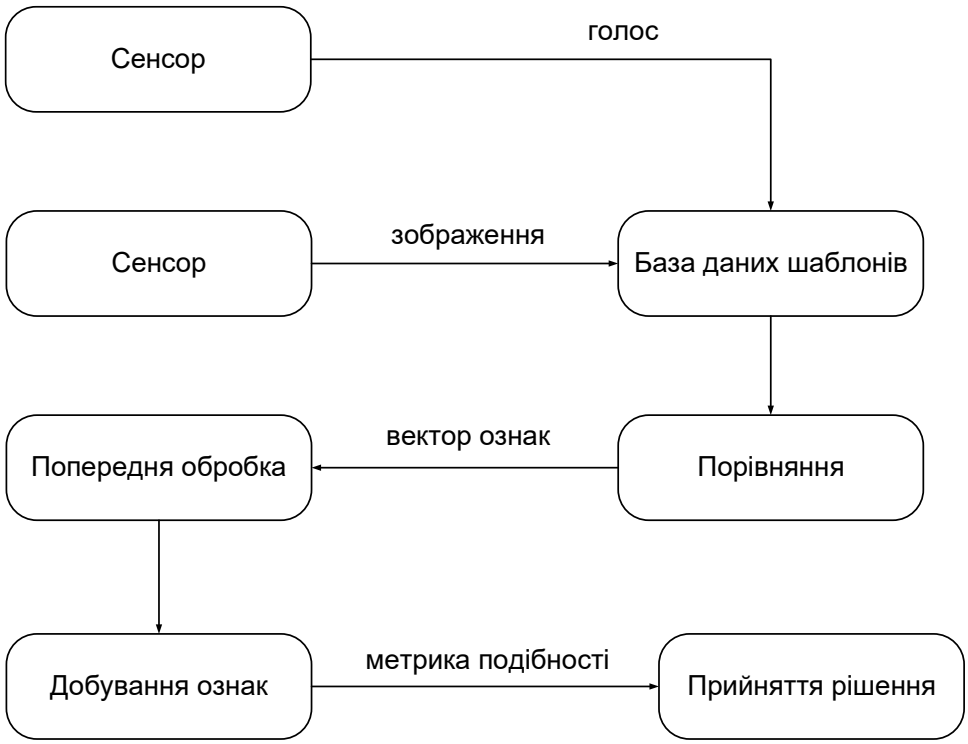


Рис. 1.2. Схема архітектури біометричної системи на основі розпізнавання обличчя

Ще один сучасний метод біометричної аутентифікації людини пов’язаний з розпізнаванням райдужної оболонки ока. Даний метод використовує унікальний візерунок цієї оболонки, оскільки він залишається незмінним протягом життя. Метод характеризується надзвичайною точністю, але потребує спеціалізованих камер з інфрачервоним підсвічуванням [6].

Розроблені також методи розпізнавання фізіологічних властивостей людини на основі аналізу венозного малюнку долоні або пальців При його застосуванні оптичні сенсори фіксують унікальну структуру судин, що не піддається фальсифікації. Метод активно застосовується у банківських системах.

Найбільш поширеними методи поведінкової аутентифікації людини базуються на розпізнаванні голосу, підпису та ходи.

Методи розпізнавання голосу аналізують спектральні та тембральні особливості мовлення і комбінуються з алгоритмами обробки природної мови для підтвердження особи під час голосових запитів.

Розпізнавання за почерком або підписом використовується в електронних планшетах і графічних екранах та аналізують не лише форму символів, а й швидкість, натиск і траєкторію написання.

Методи розпізнавання за ходою використовується у системах відеоаналітики для розпізнавання на відстані. Алгоритми, які реалізують ці методи, проводять аналіз послідовності рухів ніг, темп руху, довжину кроку та інші характеристики.

У табл. 1.1 наведено порівняльну характеристику методів біометричної аутентифікації людини.

Таблиця 1.1

Порівняльна характеристика методів

№	Метод	Тип ознак	Переваги	Недоліки	Точність
1	Відбиток пальця	Фізіологічний	Простота, висока точність	Вразливість до пошкодження	~99%
2	Обличчя	Фізіологічний	Безконтактність, швидкодія	Залежність від освітлення	~95%
3	Райдужна оболонка ока	Фізіологічний	Висока унікальність	Висока вартість сенсорів	~99.5%
4	Голос	Поведінковий	Дистанційність, зручність	Чутливість до шуму	~90%
5	Підпис	Поведінковий	Природність	Нижча точність при стресі	~85%

Для підвищення надійності все частіше застосовуються мультимодальні системи, які комбінують кілька незалежних ознак, наприклад, обличчя та голос або відбиток пальців та райдужна оболонка ока. Такі системи знижують ймовірність

помилок і підвищують стійкість до підробок, проте вимагають більшої обчислювальної потужності.

Мультимодальні системи мають низку переваг:

– зниження ймовірності помилок біометричної аутентифікації;

– підвищення достовірності результатів розпізнавання за двома типами ознак;

– можливість роботи за відсутності одного з каналів, наприклад, при поганому освітленні, ідентифікація проводиться за голосом;

– гнучкість налаштування вагових коефіцієнтів у комбінованій моделі прийняття рішення.

Останніми роками спостерігається злиття класичних методів біометрії з технологіями глибинного навчання (Deep Learning) та нейронних мереж [8]. Моделі типу FaceNet, ArcFace, VGGFace, DeepID навчаються на мільйонах зображень і формують вектори ознак із високою роздільною здатністю у просторі ознак. Завдяки цьому досягається точність понад 99.8% навіть за складних умов освітлення чи пози. Окрім цього, додатково впроваджуються підходи 3D-біометрії, що аналізують тривимірну структуру обличчя, детекції живої людини (liveness detection), яка запобігає використанню фотографій або відео для обману системи, а також федеративного навчання (federated learning), що дозволяє тренувати біометричні моделі без централізованого збирання даних, забезпечуючи конфіденційність.

Проведений аналіз показує, що найперспективнішими є мультимодальні системи, які базуються на розпізнаванні обличчя та голосу в режимі реального часу, оскільки вони поєднують безконтактність, зручність та достатню точність.

1.2. Аналіз моделей і алгоритмів розпізнавання обличчя

Методи розпізнавання обличчя є одним із найінтенсивніше досліджуваних напрямів комп’ютерного зору. Вони пройшли довгий шлях розвитку, починаючи від простих геометричних моделей до складних глибинних нейронних архітектур,

що працюють у реальному часі та перевершують людську точність у деяких задачах.

Перші системи розпізнавання обличчя, створені у 1960-1970-х роках, базувались на вимірюванні геометричних співвідношень між рисами обличчя, зокрема, відстаней між очима, довжини носа, ширини рота тощо. Такі методи вимагали ручного маркування ключових точок і не були стійкими до освітлення, міміки чи поворотів голови. Узагальнений алгоритм розпізнавання обличчя на основі ручного маркування ключових його точок представлено на рис. 1.3.

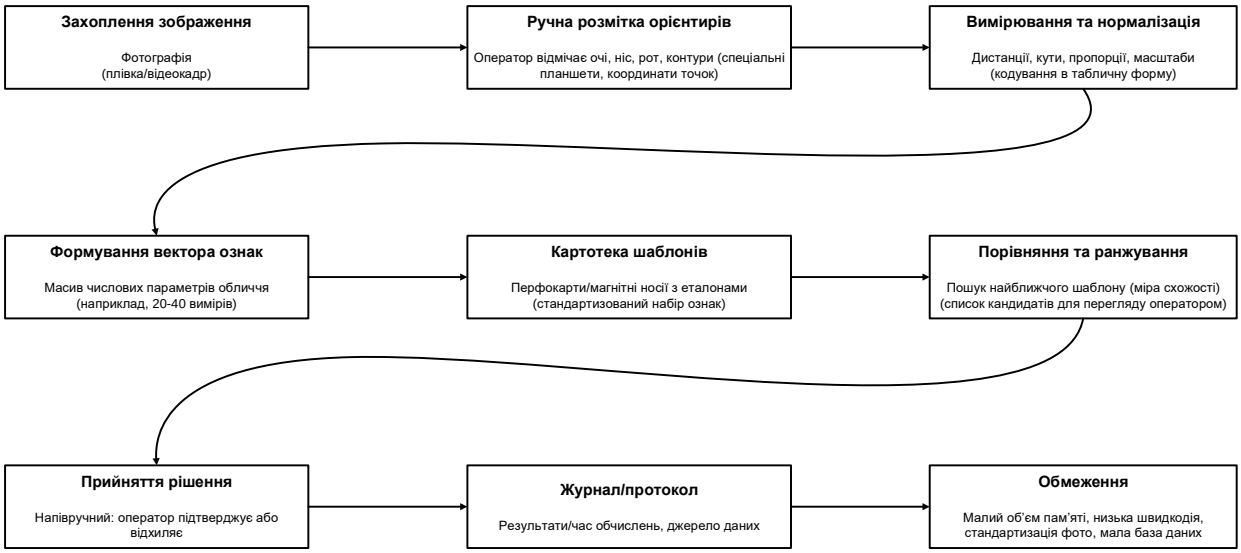


Рис. 1.3. Розпізнавання обличчя на основі аналізу геометричних співвідношень між рисами

Справжній прорив стався з появою методів головних компонент (PCA). Алгоритм Eigenfaces (1991 р.) представляв обличчя як лінійну комбінацію базових «власних облич», отриманих у результаті статистичного аналізу великої вибірки зображень. Хоча метод давав задовільні результати для контрольованих умов, він був чутливий до освітлення й пози. Результати застосування цього алгоритму показано на рис. 1.4.

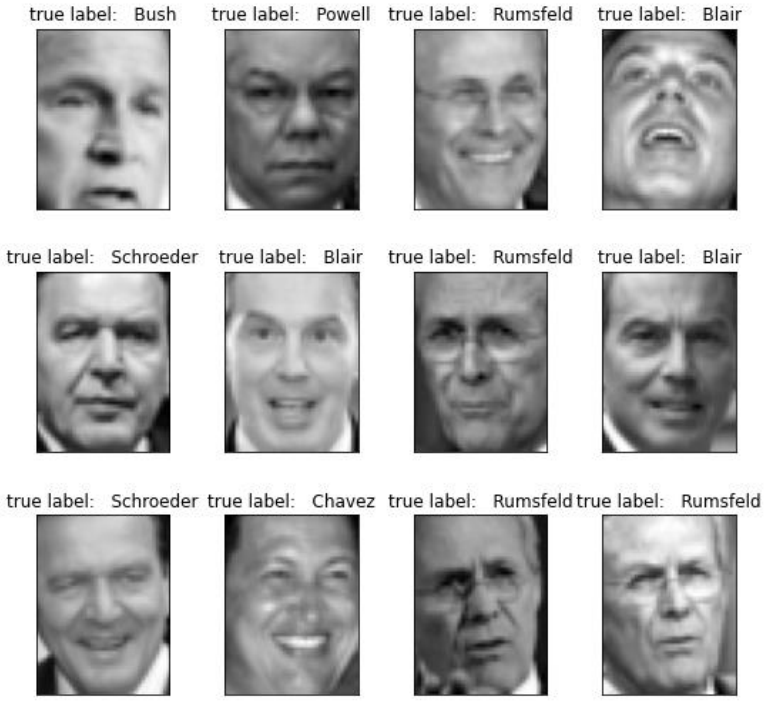


Рис. 1.4. Результат застосування алгоритму PCA

Подальше вдосконалення алгоритму PCA отримало в методі Fisherfaces, який використовував лінійний дискримінантний аналіз (LDA) для максимального відокремлення класів. Цей підхід покращив точність у складніших умовах розпізнавання обличчя. Основні відмінності PCA та LDA показано на рис. 1.5.

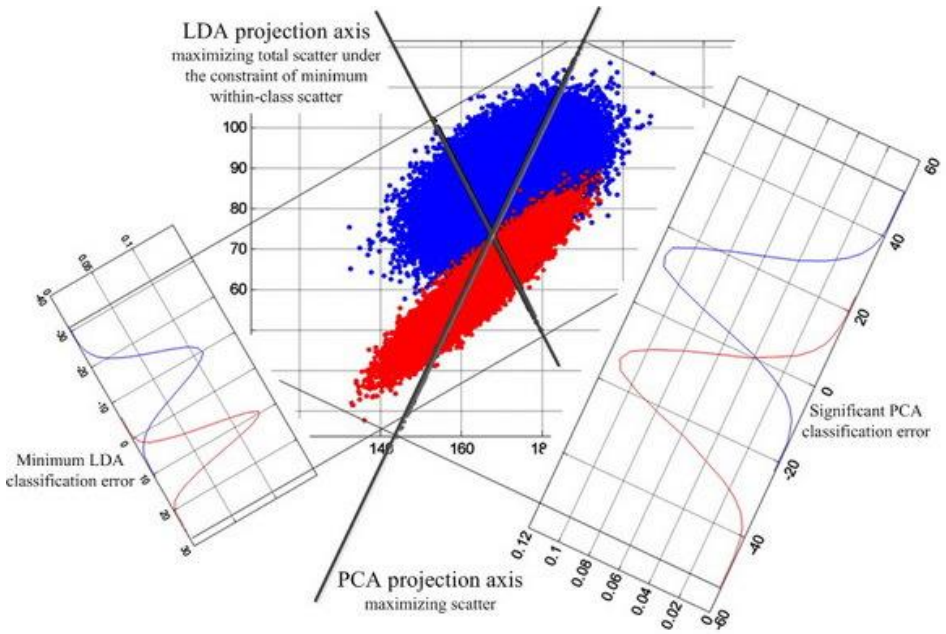


Рис. 1.5. Відмінності між PCA та LDA

Паралельно розвивались методи локальних дескрипторів, зокрема Local Binary Patterns (LBP) і Histogram of Oriented Gradients (HOG). LBP представляє обличчя як набір локальних текстурних візерунків, інваріантних до освітлення. HOG описує напрямки градієнтів яскравості, що добре підходить для виявлення облич у зображенні. Комбінації HOG + SVM або LBP + Adaboost стали базою для популярних детекторів облич у 2000-х роках.

З початку 2010-х років розвиток графічних процесорів і доступ до великих баз даних зображень (LFW, CelebA, MS-Celeb-1M) спричинили революцію у біометрії. Основою сучасних систем стали згорткові нейронні мережі (CNN), здатні автоматично виділяти ознаки без ручного проєктування дескрипторів. Загальна структура згорткової нейронної мережі показана на рис. 1.6.

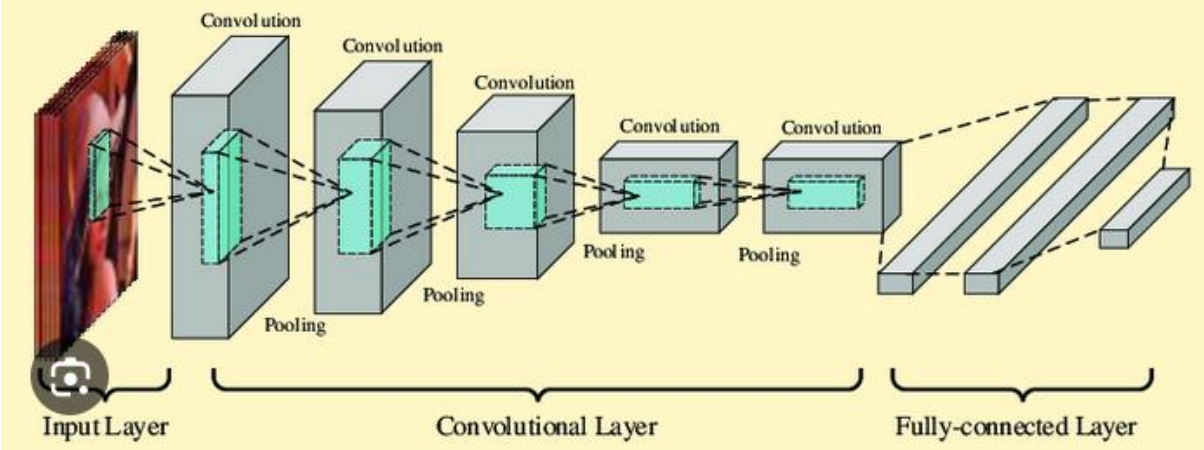


Рис. 1.6. Типова архітектура згорткової нейронної мережі

Однією з перших масштабних систем була DeepFace, розроблена компанією Meta (Facebook) у 2014 році. Вона мала вісім шарів, включно з трьома згортковими, та використовувала 3D-вирівнювання облич для компенсації поворотів. Точність розпізнавання на базі LFW досягла 97,35 %, що наблизило систему до рівня людини.

Далі з’явилась архітектура VGGFace (2015), побудована на глибокій мережі з 16 шарів за аналогією з VGG-16 для класифікації зображень. Вона дозволяла

отримувати embeddings високої роздільної здатності, придатні для порівняння різних облич.

Значного поширення набула модель FaceNet, яка вперше запропонувала навчання з використанням функції втрат, що мінімізує відстань між зображеннями однієї особи та максимізує між різними. Це дозволило формувати універсальний вектор ознак незалежно від кількості користувачів у базі.

Подальші розробки призвели до появи ArcFace, яка ввела нормалізовану функцію втрат з кутовим відхиленням. Завдяки цьому модель краще розділяє класи в гіперсферичному просторі й досягає точності понад 99,8 % на відкритих тестах.

Окремо слід згадати MobileFaceNet спрощену архітектуру, оптимізовану для мобільних і вбудованих пристроїв. Вона використовує глибокі сепарабельні згортки та меншу кількість параметрів, що дозволяє досягати прийнятної точності при низькому енергоспоживанні.

Сучасні системи розпізнавання облич виконують кілька послідовних етапів:

– детекція обличчя – використовується каскад Хаара, MTCNN або RetinaFace;

– вирівнювання (alignment) – корекція положення очей і рота для приведення зображення до єдиної орієнтації;

– нормалізація освітлення і масштабування;

– екстракція ознак – через CNN, зазвичай отримується вектор довжини 128–512;

– порівняння ознак – косинусна подібність або евклідова відстань;

– прийняття рішення – порівняння з порогом достовірності.

Для підвищення стійкості використовуються механізми liveness-detection, тобто виявлення моргання, мікродинаміки, глибинної текстури або теплового підпису.

Методи розпізнавання обличчя забезпечують високу швидкодію та зручність користувача. Вони придатні для дистанційного контролю, відеоспостереження й мобільних пристроїв. Основні обмеження пов’язані з освітленням, зміною виразів обличчя, кутом огляду та частково з етичними питаннями приватності.

Нинішня тенденція передбачає перехід до мультимодальних систем, які поєднують обличчя з іншими ознаками, зокрема, поведінковими, і використовують трансформерні архітектури (наприклад, Vision Transformer або Hybrid CNN–Transformer). Такі підходи підвищують узагальнюючу здатність моделі, дозволяють ефективніше інтегрувати просторову й контекстну інформацію та адаптувати систему до реальних умов роботи.

1.3. Аналіз моделей та алгоритмів розпізнавання голосу

Ідентифікація за голосом, або voice biometrics, є одним із ключових напрямів поведінкової біометрії. Вона спирається на припущення, що голос кожної людини унікальний завдяки анатомічним особливостям голосового тракту та індивідуальним навичкам мовлення.

Перші алгоритми розпізнавання голосу ґрунтувалися на мел-частотних кепстральних коефіцієнтах (MFCC), які представляють компактний числовий спектр мовлення, що імітує слухову шкалу людини. Ці ознаки використовувалися як вхід для статистичних моделей Gaussian Mixture Model (GMM) для опису розподілу параметрів голосу кожного користувача. Приклад візуалізації короткого голосового повідомлення та його спектрограма показані на рис. 1.7.

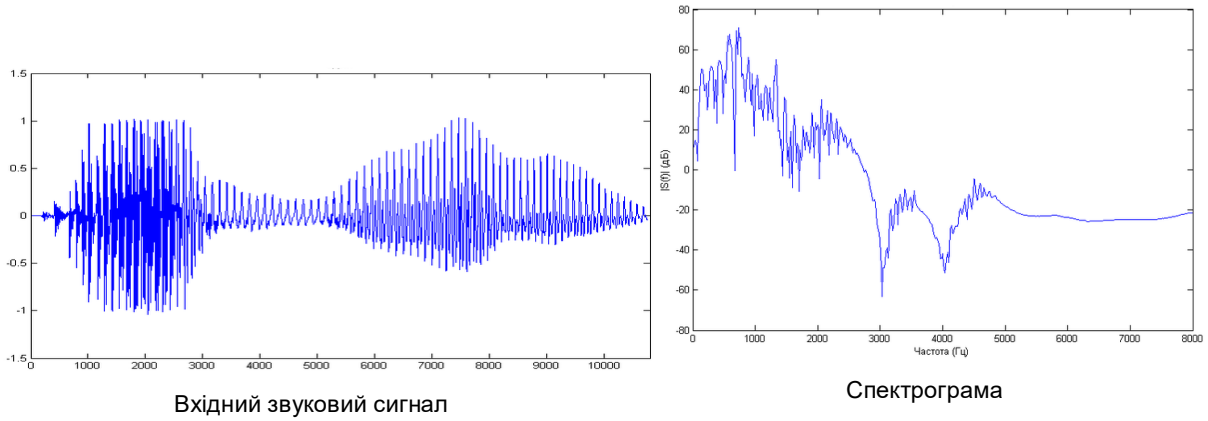


Рис. 1.7. Звуковий сигнал та його спектр

Далі розвиток отримали Hidden Markov Models (HMM), здатні моделювати часову послідовність фонем. Такі системи широко застосовувались у 1990-х – 2000-х роках у телефонних службах розпізнавання мовлення. Проте вони потребували великої кількості даних і не завжди коректно працювали у шумному середовищі. На рис. 1.8 показано типовий приклад імовірнісних параметрів прихованої марковської моделі.

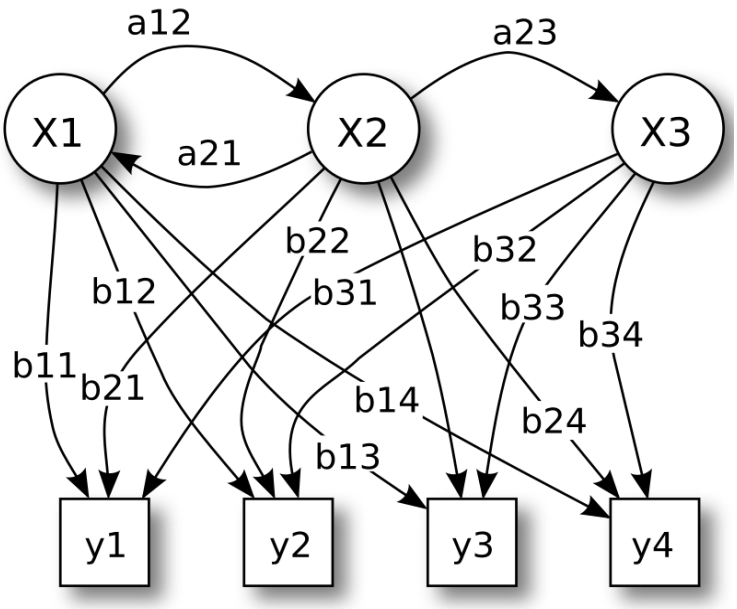


Рис. 1.8. Параметри прихованої марковської моделі

Ще один напрямок – i-vector моделі (2009 р.), які узагальнювали ознаки голосу у вектор фіксованої довжини. Цей підхід використовував техніку фактор-аналізу, що дозволяло ефективно порівнювати записи різної тривалості.

Комбінація MFCC, i-vector та PLDA довгий час залишалася промисловим стандартом і до сьогодні використовується у деяких банківських рішеннях.

З 2017 року з поширенням глибинного навчання ідентифікація голосу перейшла на новий рівень. Модель x-vector [10] запропонувала екстрагувати ознаки за допомогою глибинної нейронної мережі TDNN.

Мережа приймає послідовність фреймів і через статистичний пулінг формує вектор ознак постійної довжини, який описує акустичний стиль мовця. Принцип роботи та архітектуру мережі TDNN показано на рис. 1.9.

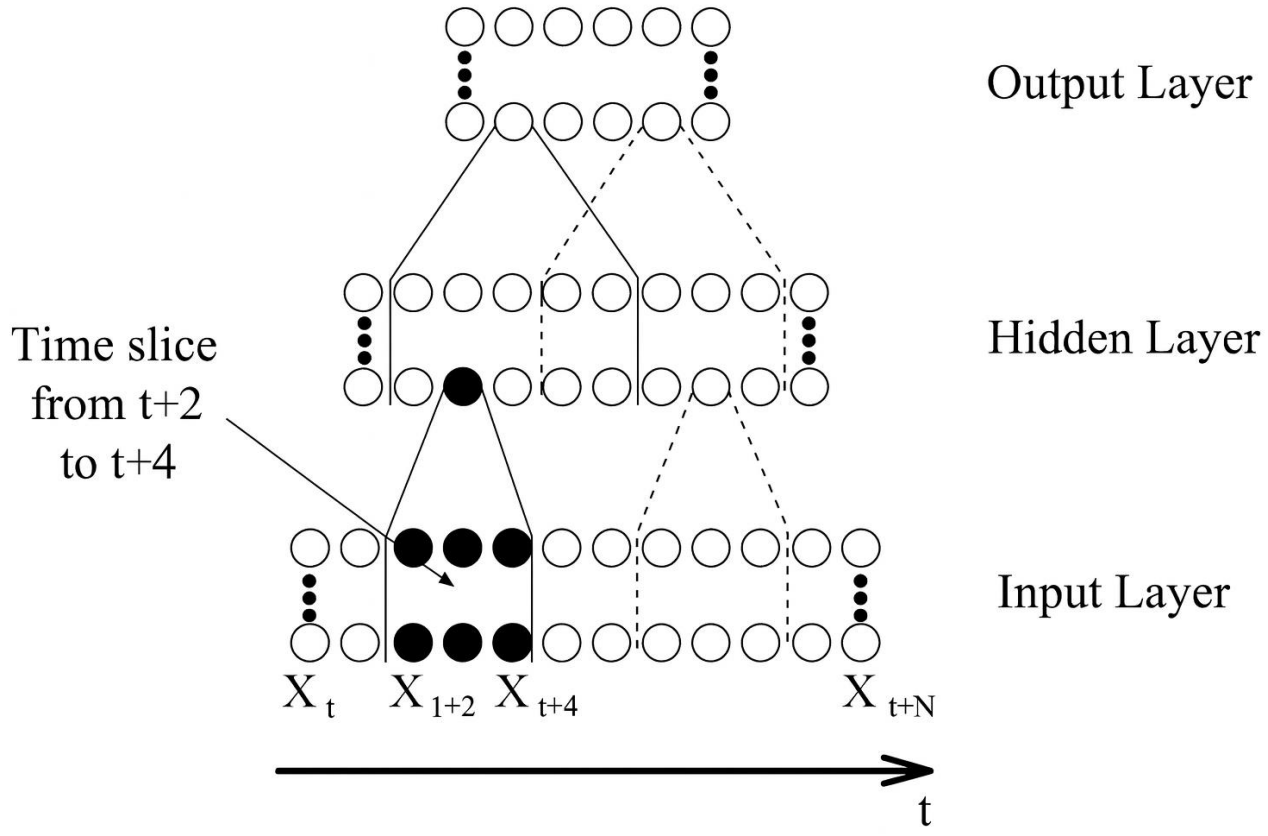


Рис. 1.9. Архітектура типової нейромережі TDNN

Подальшим розвитком став ECAPA-TDNN (Emphasized Channel Attention, Propagation and Aggregation TDNN). Модель запровадила механізми уваги (channel attention) і розширила пропускну здатність мережі завдяки резидуальним блокам. ECAPA демонструє надвисоку точність на базах VoxCeleb 1/2 та стала де-факто стандартом у сучасних дослідженнях.

Паралельно розвиваються рекурентні (LSTM, GRU) та трансформерні архітектури, зокрема wav2vec 2.0, яка навчається у самоконтрольованому режимі на великих аудіокорпусах без розмічених даних. Вона здатна формувати універсальні звукові представлення, придатні як для розпізнавання мовлення, так і для верифікації мовців.

Іншим прикладом є SpeechBrain та ResNet-based speaker encoders, що використовують 2D-спектрограму як зображення й обробляють її за допомогою згорткових мереж.

Типова система розпізнавання голосу виконує такі етапи:

– захоплення сигналу (16 кГц, 16-bit, моно);

– попередня обробка – фільтрація шумів, нормалізація гучності, детекція мовної активності (VAD);

– перетворення сигналу у спектрограму – швидке перетворення Фур'є, мел-масштабування;

– добування ознак через CNN, TDNN або трансформер;

– формування векторного представлення (embedding) фіксованої довжини;

– порівняння embeddings за косинусною відстанню або PLDA;

– прийняття рішення щодо достовірності користувача.

Для підвищення стійкості до шуму додаються алгоритми Noise Suppression, Spectral Subtraction і Speech Enhancement. У промислових системах реалізується також anti-spoofing, що виконує виявлення відтвореного голосу за спектральними артефактами або фазовими характеристиками. У табл. 1.2 наведено порівняльну характеристику моделей архітектур для виявлення та розпізнавання голосу.

Таблиця 1.2

Порівняння моделей і практичні аспекти

Архітектура	Тип	Переваги	Недоліки	Похибка
GMM + MFCC	Статистична	простота, мала потреба в ресурсах	низька точність у шумі	~8–10 %
i-vector + PLDA	Факторна	ефективна, стабільна	складне навчання	~3–5 %
x-vector (TDNN)	нейромережева	автоматична екстракція ознак	чутливість до довжини запису	~1–2 %

Продовження табл. 1.2

Архітектура	Тип	Переваги	Недоліки	Похибка
ECAPA-TDNN	нейромережева	висока точність, увага до каналів	більші вимоги до GPU	< 1 %
wav2vec 2.0	трансформер	самонавчання, універсальність	великі обчислювальні ресурси	~0.5–1 %

Сучасні моделі демонструють майже людську точність, однак мають різні вимоги до обчислювальних ресурсів. Для вбудованих пристроїв застосовують полегшені версії типу Tiny-x-vector, ECAPA-Small, DeepSpeech-lite, що дозволяє досягти компромісу між швидкодією та точністю.

Основною перевагою голосової біометрії є можливість дистанційної автентифікації без спеціального обладнання. Система може працювати на будь-якому пристрої з мікрофоном, що робить її привабливою для мобільних і веб-сервісів. Недоліки полягають у вразливості до фонового шуму, емоційних змін, захриплості або імітації голосу.

У сучасних мультимодальних рішеннях голос поєднують з іншими біометричними ознаками, зокрема, обличчям, відбитками пальців чи жестами. Це підвищує загальну точність системи та зменшує кількість помилкових рішень.

Аналіз сучасних алгоритмів показує, що як у сфері розпізнавання обличчя, так і голосу домінують глибинні нейронні мережі, які значно перевершують класичні статистичні методи. Архітектури типу ArcFace і ECAPA-TDNN забезпечують високу точність і добре адаптуються до різних умов.

Для розробки мультимодальної системи аутентифікації, яка працюватиме на Raspberry Pi 4 B, доцільно використовувати спрощені варіанти цих моделей, що зберігають основні принципи глибинного навчання при меншій складності. Це забезпечить прийнятну швидкодію, високу точність і можливість автономної роботи системи в реальному часі.

1.4. Висновки до розділу

Проведено аналіз основних понять та класифікацій біометричної аутентифікації, включаючи фізіологічні та поведінкові характеристики, а також встановлено відмінності між простими та мультимодальними системами, що дало змогу обґрунтувати вибір комплексного підходу до аутентифікації особи, який поєднує кілька незалежних ознак для підвищення надійності та точності системи.

Проаналізовано сучасні моделі та алгоритми розпізнавання обличчя, починаючи з класичних статистичних методів до глибинних нейронних мереж, що дало можливість визначити переваги використання глибинного навчання, зокрема високу стійкість до варіацій освітлення, пози та міміки, а також сформувати вимоги до обчислювальних ресурсів для подальшого проектування системи.

Досліджено сучасні методи аутентифікації голосу, включаючи MFCC-ознаки, статистичні моделі (GMM, i-vector), та передові нейромережеві архітектури, що дало змогу визначити ефективні підходи до добування голосових ознак та порівняння їх за точністю, стійкістю до шумів і вимогами до обчислень, що є критично важливим для створення мультимодальної системи на вбудованих пристроях.

Встановлено, що найперспективнішими є мультимодальні рішення, які поєднують кілька біометричних характеристик. Це дало змогу сформувати концептуальну основу майбутньої апаратно-програмної системи, спрямованої на підвищення точності, стійкості та здатності до роботи в реальному часі.

РОЗДІЛ 2
ПРОЕКТУВАННЯ АРХІТЕКТУРИ СИСТЕМИ МУЛЬТИМОДАЛЬНОЇ БІОМЕТРИЧНОЇ АУТЕНТИФІКАЦІЇ КОРИСТУВАЧА КОМП'ЮТЕРНОЇ СИСТЕМИ

2.1. Побудова узагальненої архітектури апаратно-програмного комплексу мультимодальної біометричної аутентифікації особи

У роботі пропонується архітектура системи, яка будується за модульним принципом і охоплює як апаратні, так і програмні компоненти. Ці елементи взаємодіють між собою через локальні інтерфейси та внутрішні програмні API. Основною ідеєю є об'єднання двох незалежних каналів біометричної аутентифікації за обличчям і за голосом в єдину інтегровану платформу, яка забезпечує надійне підтвердження особи у режимі реального часу. Запропонована архітектура апаратно-програмного комплексу мультимодальної біометричної аутентифікації особи представлена на рис. 2.1.

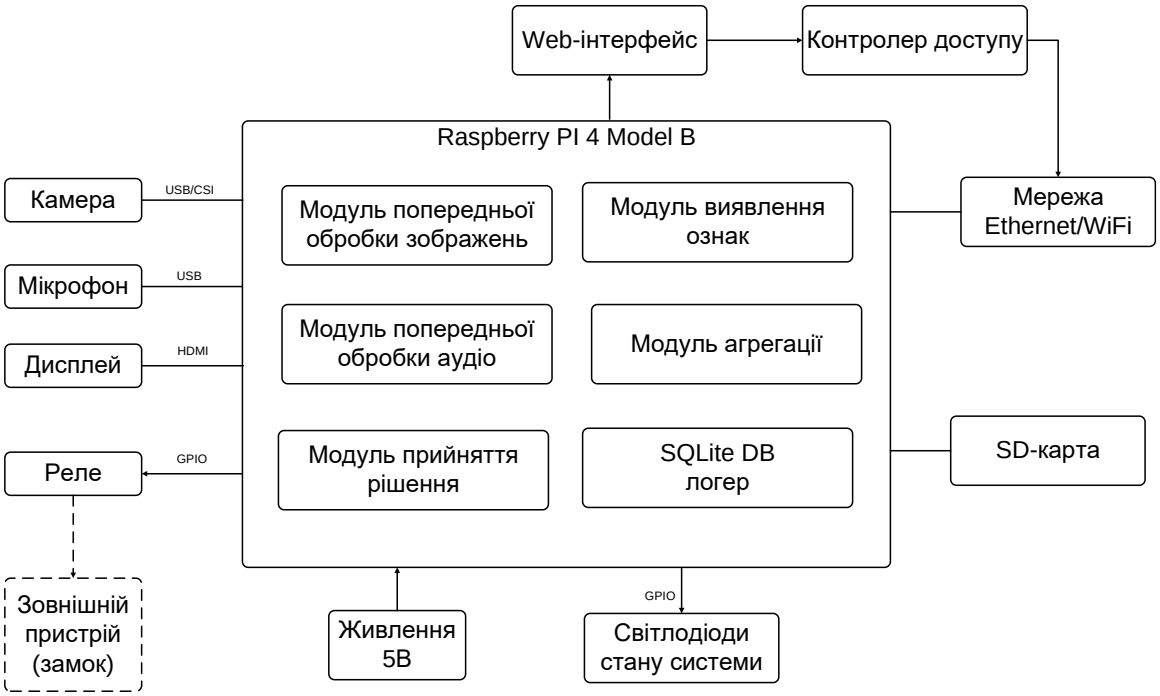
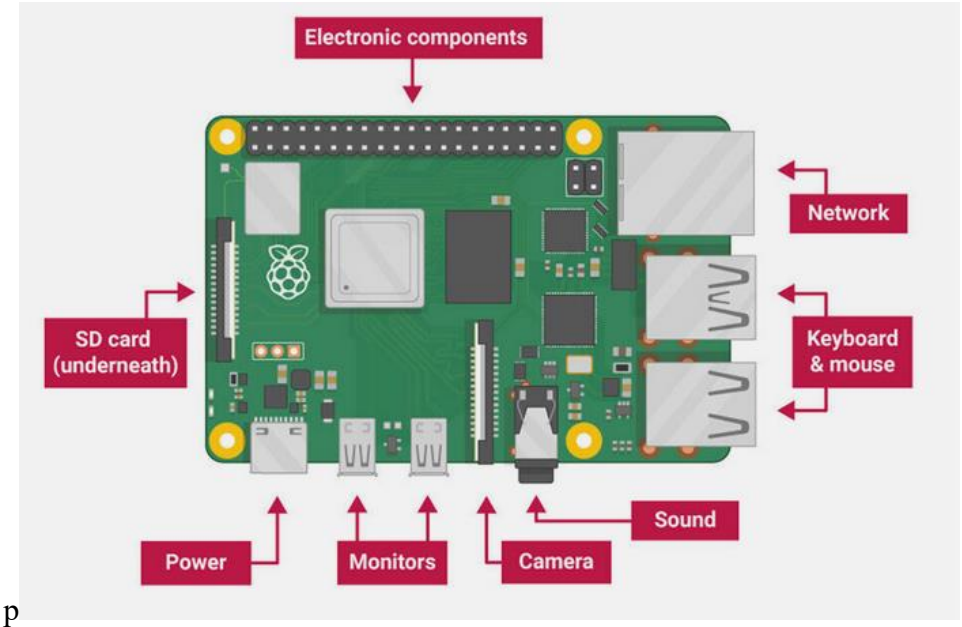


Рис.2.1. Архітектура апаратно-програмного комплексу біометричної аутентифікації

Як видно з рис. 2.1, центральним елементом комплексу виступає одноплатний комп’ютер Raspberry Pi 4 Model B, який координує роботу всіх підсистем.

Базовий вузол системи розміщено на Raspberry Pi, що поєднує у собі достатню обчислювальну потужність, мале енергоспоживання та можливість роботи у бездисковому автономному режимі. Основні компоненти Raspberry Pi 4 Model B представлено на рис. 2.2.



p

Рис. 2.2. Основні компоненти Raspberry Pi 4 Model B

До мінікомп’ютера підключено USB-камеру, яка здійснює захоплення відео з обличчям користувача, та мікрофон, який формує звуковий потік для аналізу мовлення.

Для зворотного зв’язку може використовуватися невеликий дисплей або сенсорний екран, де відображаються результати аутентифікації та стан системи.

Через GPIO-порти Raspberry Pi керує реле або електронним замком, що забезпечує фізичну реакцію системи, яка включає:

– розблокування доступу;
– світлового сигналу за допомогою світлодіодів.

У конструкції передбачено модуль живлення з резервним акумулятором (UPS HAT), який гарантує коректне завершення роботи навіть при раптовому відключенні електромережі. Це запобігає пошкодженню даних у локальній базі. Усі сигнали з керуючих портів проходять через оптоелектронні розв'язки, що підвищує електричну безпеку пристрою.

Програмне забезпечення системи складається з двох основних частин: backend-сервісу та веб-інтерфейсу користувача. Обидві частини функціонують на Raspberry Pi та не потребують підключення до зовнішнього сервера.

Backend пропонується реалізувати на Python із використанням фреймворку FastAPI. Цей компонент відповідає за взаємодію з периферійними пристроями, обробку даних, зберігання інформації та логіку прийняття рішень.

Для зберігання даних використовується локальна база SQLite, що не потребує додаткових служб і добре підходить для вбудованих систем. У базі даних передбачається створення таблиць користувачів, векторних описів облич та голосу, журналів подій і системних параметрів.

Модуль відеоаутентифікації виконує захоплення кадрів з камери, корекцію зображення, нормалізацію яскравості та виявлення області обличчя. Далі активується нейромережева модель, навчена для екстракції ознак обличчя, яка формує компактне числове представлення користувача. Графічно алгоритм роботи даного модуля можна представити у вигляді структури, показаної на рис. 2.3.

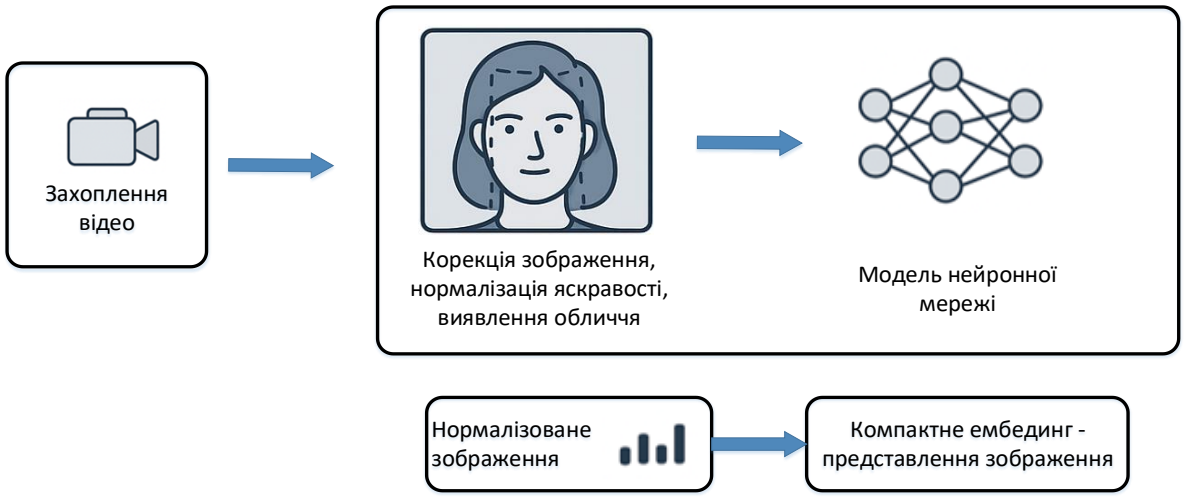


Рис. 2.3. Функції модуля відеоаутентифікації

Аналогічно, модуль голосової аутентифікації аналізує аудіосигнал, фільтрує шум, виділяє мовні ділянки та обчислює спектральні характеристики. У результаті отримується вектор голосових ознак, який у подальшому порівнюється з тими зразками, які є в базі даних. Алгоритм модуля голосової аутентифікації працює за алгоритмом представленим на рис. 2.4.

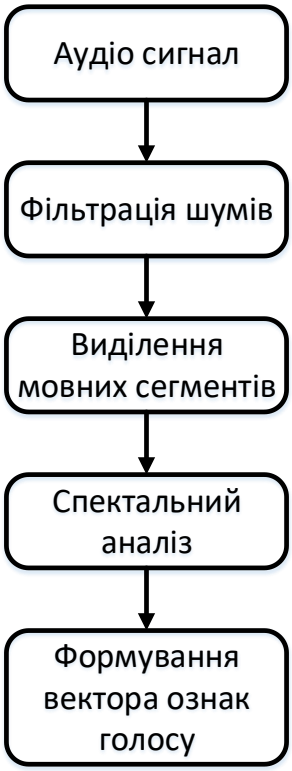


Рис. 2.4. Послідовність кроків у процесі розпізнавання голосу

Одержані вектори ознак надходять до модуля агрегації, який поєднує інформацію від двох каналів і формує єдину оцінку достовірності. На основі цього результату система ухвалює рішення щодо автентичності користувача. У випадку позитивного результату формується команда на реле керування замком та запис у журнал подій.

Веб-інтерфейс адміністратора пропонується створити засобами HTML 5 та JavaScript із використанням шаблонів FastAPI. Через цей інтерфейс адміністратор може реєструвати нових користувачів, здійснювати первинне навчання системи,

переглядати журнали, змінювати параметри роботи, а також тестувати камеру та мікрофон.

Веб-інтерфейс працює у межах локальної мережі або через захищене з’єднання HTTPS, що дозволяє виконувати адміністрування без підключення до хмарних сервісів.

Взаємодію між компонентами системи планується реалізувати через внутрішній API, що базується на обміні HTTP-запитами. Після вмикання пристрою ініціалізується камера та мікрофон, перевіряється наявність з’єднання з базою даних і доступність мережевих інтерфейсів.

Модуль моніторингу стану постійно відстежує температуру процесора, використання пам’яті, кількість активних з’єднань і при потребі надсилає повідомлення в лог.

Під час спроби аутентифікації користувач подаватиме аудіовізуальний зразок, тобто його обличчя повинно потрапляти у камеру і він повинен вимовити фразу, яка відобраєаьбся на дисплеї. Обидва канали активуються паралельно, що скорочує час очікування.

Після завершення обробки система порівнює отримані ознаки з наявними шаблонами в базі. Якщо збіг перевищує заданий поріг, подається сигнал успішного розпізнавання. У випадку відмови на дисплеї з’являється повідомлення про помилку та рекомендація повторити спробу.

Робота системи супроводжується веденням журналу: у базу заносяться дата й час події, ідентифікатор користувача, результати перевірки, оцінка схожості та службова інформація. Такий підхід забезпечує можливість подальшого аудиту та аналізу статистики помилок.

Оскільки система орієнтована на автономну роботу, значну увагу приділена її стабільності. Для захисту даних доцільно використати журналювання транзакцій SQLite (режим WAL), що запобігає пошкодженню бази при непередбаченому завершенні процесу. Автоматичне резервне копіювання бази даних повинно виконуватися у фоновому режимі з можливістю відновлення останньої коректної версії.

При реалізації системи необхідно передбачити синхронізацію годинника через NTP-службу, що забезпечує точність часових міток у журналах подій. У разі втрати живлення UPS-модуль подає сигнал про вимкнення та зберігає усі відкриті файли, після чого пристрій безпечно вимикається.

Безпека доступу до адміністративного інтерфейсу може бути реалізована через локальну авторизацію та контроль IP-адрес. Додатково може бути встановлено зворотний проксі-сервер Nginx, який забезпечує шифрування трафіку за протоколом TLS і базовий захист від зовнішніх запитів. Передбачається, що усі паролі та конфіденційні дані будуть зберігатися у зашифрованому вигляді.

Система спроєктована таким чином, щоб її можна було розширювати без значних змін архітектури. Зокрема, можна додати додаткові біометричні модулі, наприклад, ідентифікацію за відбитком пальця або аналіз ходи через додаткові API. Програмний код має відкриту структуру, тому може бути легко перенесений на інші апаратні платформи або інтегрований у хмарне середовище.

Передбачена можливість централізованого збору даних від кількох Raspberry Pi-вузлів до сервера моніторингу, що дозволяє формувати єдину систему контролю доступу з розподіленими точками аутентифікації. Для спрощення розгортання програмних компонентів застосовуються конфігураційні файли .env та сценарії автоматизації, які забезпечують однакові параметри роботи на різних пристроях.

Запропонована архітектура є збалансованим рішенням для побудови мультимодальної біометричної системи аутентифікації. Вона поєднує компактність апаратного забезпечення, гнучкість програмних компонентів і достатній рівень захищеності даних. Використання Raspberry Pi 4B дає змогу створити повноцінну автономну систему, яка може працювати без підключення до зовнішніх серверів, а при потребі інтегруватися у більшу корпоративну мережу.

Завдяки модульній побудові кожен елемент системи може розвиватися незалежно, що відкриває можливості для подальшого вдосконалення алгоритмів розпізнавання, впровадження нових сенсорів і оптимізації продуктивності. Така архітектура забезпечує не лише виконання основної функції аутентифікації

користувача за двома біометричними ознаками, а й створює основу для подальших досліджень і практичного розширення системи в умовах обмежених ресурсів.

2.2. Формалізація процесу біометричної аутентифікації особи

Апаратно-програмний комплекс мультимодальної біометричної аутентифікації є інтегрованою системою, що забезпечує розпізнавання особи за двома незалежними модальностями: зображенням обличчя та голосовим сигналом. Система складається з апаратної підсистеми збору до якої входять камера, сенсори, мікроконтролер або одноплатний комп'ютер та програмної підсистеми обробки, аналізу й прийняття рішень у реальному часі.

На концептуальному рівні процес аутентифікації описується множиною функціональних модулів:

$$S = \{A, P, F, M, D\}, \tag{2.1}$$

де A – модулі захоплення зображення та голосу;

P – модуль попередньої обробки даних за двома модальностями;

F – модуль виділення біометричних ознак обличчя та голосу;

M – модуль порівняння ознак;

D – модуль прийняття рішень.

На відміну від класичних унімодальних систем, мультимодальна архітектура включає додатковий компонент у вигляді модуля інтеграції, який поєднує інформацію з різних біометричних потоків.

Оскільки, у системі біометричної аутентифікації передбачено використання двох модальностей, то для забезпечення процесу аутентифікації необхідно визначити рівні злиття даних. Існує два варіанти злиття модальностей: на рівні даних та на рівні ознак. Злиття на рівні даних використовується рідко і передбачає об'єднання «сирих» сигналів, отриманих з камери та мікрофону. У роботі

пропонується застосувати злиття на рівні ознак, що передбачає формування комбінованого вектору:

$$e = [e_f || e_v], \tag{2.2}$$

де e_f – вектор ознак при розпізнавання обличчя;

e_v – вектор ознак при розпізнавання голосу.

Завдяки злиттю на рівні ознак система зберігає працездатність навіть за умов часткової втрати або спотворення одного з каналів даних.

Кожен модуль функціонує автономно, але дані передаються у вигляді потоку через стандартизовані інтерфейси, що дозволяє легко масштабувати систему або інтегрувати її з іншими сервісами, наприклад, системами контролю доступу.

Нехай X_f – простір вхідних зображень обличчя, а X_v – простір голосових сигналів. Тоді $x_f \in X_f$ – зображення користувача, а $x_v \in X_v$ – аудіозапис голосу. Завдання аутентифікації особи полягає у побудові двох відображень:

$$\Phi_f : X_f \to \mathbb{R}^{n_f},\ \Phi_v : X_v \to \mathbb{R}^{n_v}. \tag{2.3}$$

Відображення (2.3) переводять відповідні модальності у вектори ознак:

$$\mathrm{e}_f = \Phi_f(x_f),\ \mathrm{e}_v = \Phi_v(x_v). \tag{2.4}$$

Для кожного зареєстрованого користувача i формується пара еталонів:

$$\mathrm{F}_i = \Phi_f(x_f^i),\ \mathrm{V}_i = \Phi_v(x_v^i). \tag{2.5}$$

Враховуючи (2.5), база еталонних шаблонів системи представляється у наступному вигляді:

$$B = \{(F_1, V_1), (F_2, V_2), \dots, (F_m, V_m)\}, \quad (2.6)$$

де m – кількість зареєстрованих у системі користувачів.

2.2.1. Оцінювання якості розпізнавання зображень

Для оцінювання якості розпізнавання образів найчастіше використовують Евклідову метрику або метрику косинуса кута між векторами, які відповідно обчислюються за формулами (2.7) і (2.8):

$$S_f^{(E)}(e_f, F_i) = \|e_f - F_i\|_2, \quad (2.7)$$

де $S_f^{(E)}(e_f, F_i)$ – метрика Евклідової відстані.

$$S_f^{(\mathrm{C})}(e_f, F_i) = \frac{e_f \cdot F_i}{\|e_f\| \cdot \|F_i\|}, \quad (2.8)$$

де $S_f^{(\mathrm{C})}(e_f, F_i)$ – метрика косинуса кута між векторами.

Задавши порогове прийнятне значення за допомогою цих метрик можна імплементувати алгоритм прийняття рішення щодо успішного розпізнавання обличчя.

2.2.2. Оцінювання якості розпізнавання голосового сигналу

Сьогодні при розпізнаванні голосу найбільш ефективними є метрики імовірнісного лінійного дискримінантного аналізу (PDLA) та косинус кута між векторами.

PDLA є сучасним статистичним методом моделювання біометричних ознак, який широко застосовується у системах розпізнавання голосу на основі i-vector та x-vector підходів. На відміну від класичного лінійного дискримінантного аналізу

(LDA), який виконує лише лінійне перетворення ознак, PLDA вводить ймовірнісну модель, що описує як міжкласову, так і внутрішньокласову варіативність.

Математично голосовий вектор представлення e описується моделлю:

$$e = \mu + F_y + \epsilon, \tag{2.9}$$

де μ – середній вектор ознак;

F_y – компонента, що моделює особливість голосу конкретної особи;

ϵ – випадкова похибка, що відображає зміну каналу запису, шум, емоційний стан та інші фактори.

PLDA дозволяє обчислити ймовірність того, що два голосові вектори-ембединги належать одному й тому самому користувачу:

$$score = log \frac{P(e_1,e_1|same\ speaker)}{P(e_1,e_1|different\ speaker)}. \tag{2.10}$$

Таким чином, PLDA є більш інформативною порівняно з евклідовими або косинусними метриками, оскільки порівняння здійснюється у просторі ймовірностей із врахуванням статистичних властивостей вибірки. У сучасних голосових біометричних системах PLDA є стандартом для оцінювання подібності та демонструє високу точність, особливо у поєднанні з векторами-ембедингу, сформованими глибинними нейронними мережами (x-vector, ECAPA-TDNN).

При розпізнаванні голосу у запропонованій системі біометричної аутентифікації особи метрика PLDA та метрика косинуса кута між векторами набудуть наступного вигляду:

$$S_v(e_v, V_i) = PLDA(e_v, V_i), \tag{2.11}$$

$$S_v^{(C)}(e_v, V_i) = \frac{e_v \cdot V_i}{\|e_v\| \cdot \|V_i\|}. \tag{2.12}$$

Інтегральна міра подібності ознак у мультимодальній системі визначається як комбінована міра:

$$S_{fusion}(i) = \alpha S_f(i) + (1 - \alpha) S_v(i), \quad (2.13)$$

де $\alpha \in [0,1]$ – ваговий коефіцієнт;

$S_f(i), S_v(i)$ – оцінки подібності по окремих модальностях.

Варто відмітити, що в якості альтернативних схем злиття можуть використовуватися логістичне комбінування, нормалізація по Z-score, softmax-агрегування або нейронні мережі для fusion-рівня.

2.3. Імовірнісна модель прийняття рішення при розпізнаванні користувача комп'ютерної системи

Для прийняття рішення про наявність збігу між біометричним зразком особи та еталонним шаблоном система порівнює значення міри подібності S з певним пороговим значенням T. Порогове значення встановлюється таким чином, щоб забезпечити баланс між безпекою системи та зручністю користування.

Процес аутентифікації формулюється як задача статистичного розпізнавання гіпотез:

$$\begin{cases} H_0, \text{якщо користувач відсутній в базі даних,} \\ H_1, \text{існує } i \in [1, m], \text{для якого } S_{fusion}(i) \geq T. \end{cases} \quad (2.14)$$

Значення T впливає на ризик хибного доступу, тобто коли неавторизовану особу прийнято за справжню та на ймовірність хибної відмови, коли система не розпізнала користувача з бази даних шаблонів.

Ймовірність того, що система біометричної аутентифікації помилково надасть доступ невідомому або неавторизованого користувача визначається за критерієм False Acceptance Rate (FAR):

$$FAR = P(S \geq T|H_0), \quad (2.15)$$

де H_0 – гіпотеза, що користувач комп'ютерної системи не належить жодній особі у базі даних.

Імовірність того, що система помилково відхилить авторизованого користувача визначається за критерієм False Rejection Rate (FRR):

$$FRR = P(S < T|H_1), \quad (2.16)$$

де H_1 – гіпотеза, що запис про користувача комп'ютерної системи наявний у базі даних.

Інтегральним критерієм моделі прийняття рішень є критерій Equal Error Rate (EER), який інтерпретує значення помилки при такому порозі T, коли FAR та FRR є рівними за значенням:

$$ERR = FAR(T^*) = FRR(T^*), \quad (2.17)$$

де T^* – порогове значення, що забезпечує рівність двох типів помилок.

Критерій EER важливим, оскільки він є універсальним критерієм якості біометричної системи. Кількісне його значення інтерпретується так, що чим нижче значення EER, тим краща загальна точність системи, а також даний критерій дає змогу об'єктивно порівнювати різні алгоритми або модальності незалежно від їхнього порогу.

Типові значення критерію EER представлені у табл. 2.1.

Таблиця 2.1

Типові значення критерію EER

Тип системи	Значення EER
Біометрична системи аналізу голосу (x-vector + PLDA)	1–2 %

Продовження табл. 2.1

Система	Значення EER
Біометрична система розпізнавання обличчя (ArcFace)	0.1–0.5 %
Біометрична мультимодальна система (обличчя + голос)	< 0.1 %

У кваліфікаційній роботі задача проектування біометричної системи полягає у виборі таких моделей ознак, метрик подібності і порогового значення T, які мінімізують EER за умови дотримання вимог реального часу.

Формально оптимізацію системи біометричної мультимодальної аутентифікації користувача комп'ютерної системи можна представити у наступному вигляді:

$$\min_{T,\Phi} EER(T,\Phi) \text{ за умови } Latency \leq L_{max}, \qquad (2.18)$$

де $Latency$ – час відгуку системи у процесі аутентифікації користувача;

L_{max} – максимально допустимий час відгуку системи.

У формулі (2.18), EER залежить від двох параметрів: T і Ф, тобто від порогового значення прийняття рішення та від функції або моделі перетворення вхідних даних. Отже, система шукає такі їх значення, які мінімізують EER.

Умова у формулі (2.18) означає, що система не повинна працювати повільніше, ніж певний максимально допустимий час відгуку L_{max}. Наприклад, якщо система опрацьовує 1 кадр за 300 мс, то це занадто повільно. Для системи реального часу значення L_{max} повинно бути на рівні 50 – 100 мс. Це означає, що навіть якщо дуже ефективна нейромережа дає низький EER, її не можна використовувати, оскільки вона працює повільно. Тому для забезпечення ефективності пропонованої системи мультимодальної біометричної системи аутентифікації користувача комп'ютерної системи необхідно знайти баланс між точністю розпізнавання, швидкодією системи та обчислювальною складністю.

2.4. Обґрунтування вибору алгоритмів та моделей розпізнавання користувачів у біометричній мультимодальній системі аутентифікації

2.4.1. Алгоритми і моделі розпізнавання обличчя

У межах даної роботи реалізується мультимодальна система біометричної аутентифікації користувачів комп’ютерної системи, у якій модуль розпізнавання обличчя передбачає виконання послідовності кроків, які наведені на рис. 2.5.

Рис. 2.5. Алгоритм розпізнавання обличчя користувача комп’ютерної системи

Послідовність кроків, наведених на рис. 2.5, реалізується окремими алгоритмами, які відповідають критеріям оптимальності з точки зору точності, продуктивності та можливості реалізації на Raspberry Pi 4. Кожен з кроків враховує особливості запропонованої архітектурою розроблюваного комплексу.

Нижче подано опис конкретних методів і моделей, які будуть реалізовані у кваліфікаційній роботі.

Для розв'язання задачі детекції обличчя обрано згорткову нейронну мережу MTCNN (Multi-Task Cascaded Convolutional Neural Network). Це обумовлено тим, що дана нейромережа забезпечує стійку точність при роботі з реальними відеопотоками, визначає не тільки координати обличчя, а й такі ключові точки, як очі, ніс, кути рота. Окрім цього, застосування MTCNN не потребує значних обчислювальних ресурсів і може бути реалізована за допомогою бібліотеки facenet-pytorch. Використання саме цієї нейронної мережі забезпечує коректне її функціонування на Raspberry Pi при використанні оптимізованих модулів. Структура нейронної мережі MTCNN наведена на рис. 2.6.

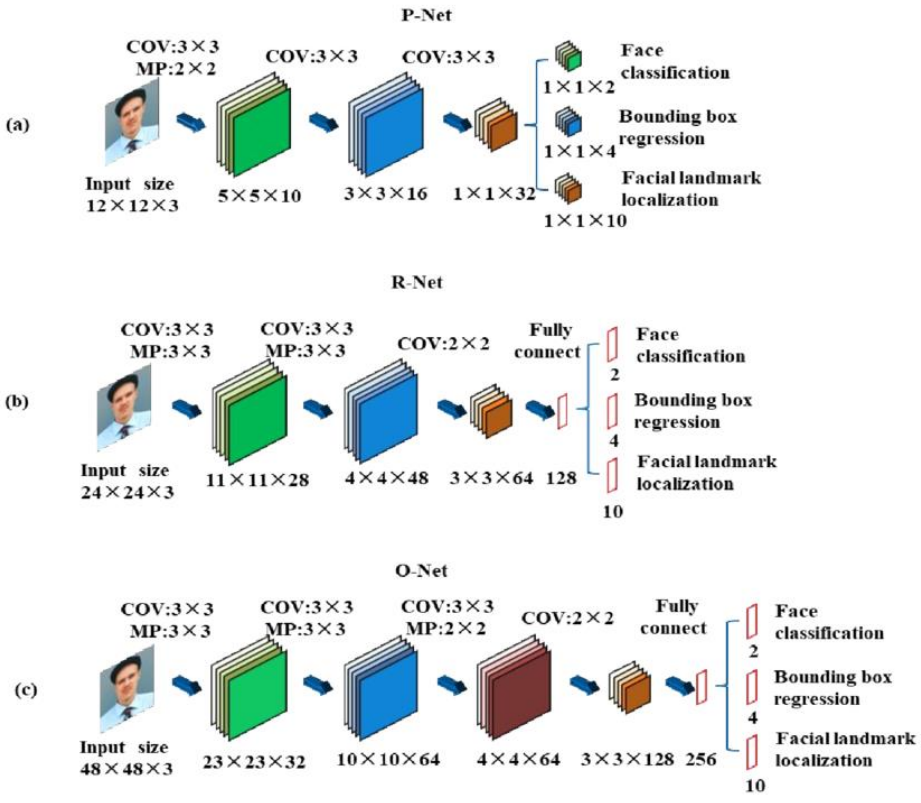


Рис. 2.6. Структура нейронної мережі MTCNN

Як видно з рис. 2.6, MTCNN складається з трьох каскадних згорткових нейронних мереж:

– P-Net (Proposal Network) – призначена для визначення на зображенні потенційних областей з обличчям;

– R-Net (Refine Network) – виконує уточнення меж обличчя;

– O-Net (Output Network) – надає точні координати обличчя на зображенні та 5 ключових анатомічних опорних точок: ліве око, праве око, кінчик носа, лівий кут рота та правий кут рота (рис. 2.7).

Рис. 2.7. Ключові анатомічні точки при аутентифікації обличчя користувача

Математично процес виявлення обличчя на зображенні з кадру відеокамери описується як композиція:

$$B = O(R(P(x_f))), \tag{2.19}$$

де B – результат детекції обличчя у вигляді рамки, накладеної на обличчя та 5-ти ключових ознак користувача;

O – результат нейромережі з точними координатами обличчя та 5-ма ключовими точками;

R – результат нейромережі при уточненні меж обличчя;

P – результат, який повретає нейромережа, що визначає потенційні контури обличчя.

На рис. 2.8 показано приклад результат виявлення обличчя в процесі аутентифікації користувача комп'ютерної системи.

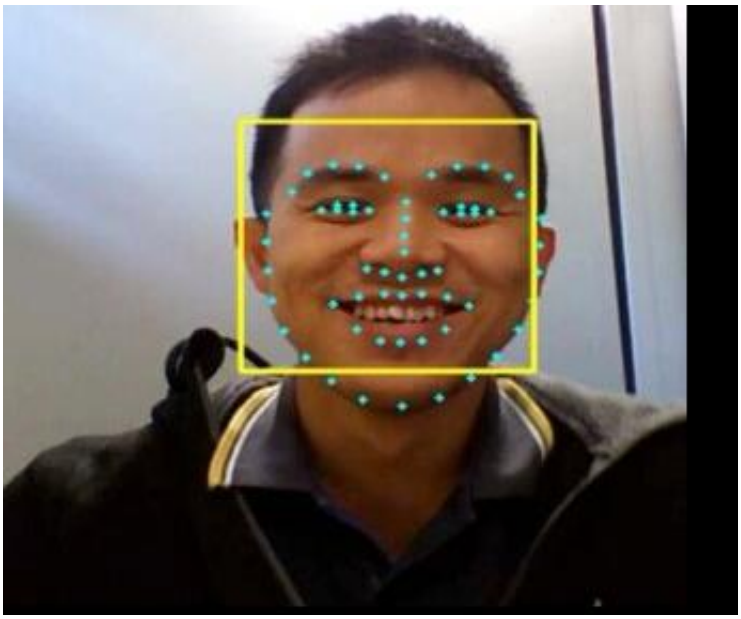


Рис. 2.8. Результат опрацювання обличчя із кадру, захопленого з відеокамери

Згідно алгоритму, продемонстрованого на рис. 2.5 результат (2.19) подається у модуль нормалізації, що забезпечує вирівнювання зображення. Для зменшення варіативності поз та положення голови використовується афінне перетворення:

$$x_{aligned} = T(x_f, L), \tag{2.20}$$

де L – 5 ключових ознак, отриманих з MTCNN;

T – трансформація на основі афінного перетворення.

Вирівняне зображення є стандартизованим вхідним сигналом для моделі добування ознак.

Добування ознак обличчя пропонується виконувати за допомогою нейронної мережі MobileFaceNet, оскільки вона спеціально розроблена для пристроїв із обмеженими ресурсами (мобільні SoC, Raspberry Pi). Окрім цього, дана нейронна

мережа на виході формує ембединг-вектор високої якості, який представлений у вигляді L2-нормованого вектора з розмірністю 128–256 пікселі, а також має низьку латентність: приблизно 3–7 мс на CPU середнього класу. Окрім цього, у роботі запропоновано застосовувати ArcFace не як нейромережу, а як функцію втрат, на основі якої модель була попередньо навчена.

До складу MobileFaceNet (рис. 2.9) входять:

– Convolution + PReLU – виконує первинну обробку зображення, добуває базові локальні ознаки: краї, кути, текстури і зменшує просторовий розмір, формуючи перше представлення обличчя;

– Depthwise Separable Blocks – це оптимізована форма згорткових шарів, що мінімізує кількість параметрів і обчислень;

– Bottleneck Residual Blocks – формує інваріантні до умов зйомки високорівневі ознаки обличчя;

– Global Average Pooling – забезпечує стиснення всіх просторових ознак в один компактний, інваріантний вектор, який буде використовуватися для формування вектора-ембедингу;

– Fully Connected Layer (128-d) – створює кінцеве представлення користувача (біометричний шаблон), яке буде зберігатися у базі даних і використовуватися для порівняння.

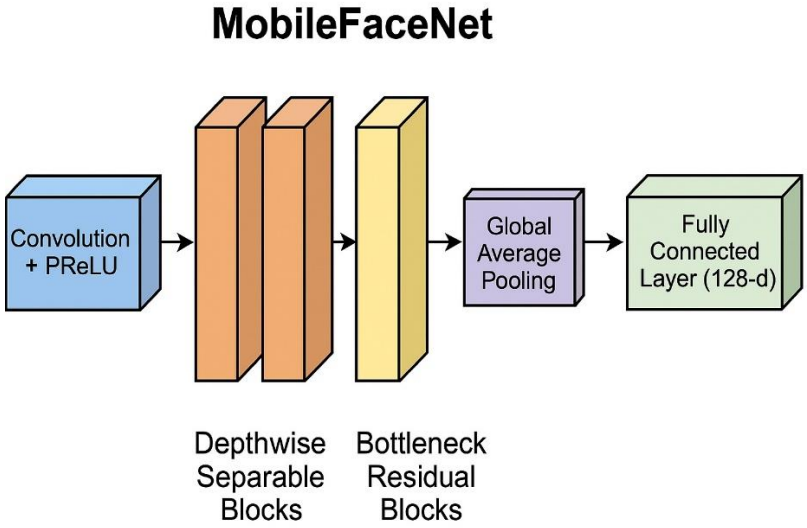


Рис. 2.9. Архітектура нейромережі MobileFaceNet

На практиці за допомогою Python досить просто реалізувати ONNX-модель MobileFaceNet, використати onnxruntime для високої швидкодії. Вектори-ембединги можуть зберігатися у SQLite у вигляді numpy-масиву. Ця модель відповідає вимогам швидкодії та точності, які накладає архітектура мультимодальної біометичної системи аутентифікації користувача комп'ютерної системи.

Оскільки MobileFaceNet генерує L2-нормовані вектори, доцільно використовувати косинусну міру подібності. Це дозволяє обчислювати подібність надзвичайно швидко, що критично при роботі в реальному часі.

Формалізація алгоритму розпізнавання обличчя передбачає виконання наступної послідовності кроків:

– Одержання RGB-зображення – x_f;

– Детекція обличчя – $(B, L) = MTCNN(x_f)$;

– Нормалізація (вирівнювання) обличчя – $\widehat{x_f} = T(x_f, L)$;

– Добування ознак – $e_f = \Phi_f(\widehat{x_f})$;

– Порівняння з даними з бази даних шаблонів – $S_f(i) = sim(e_f, F_i)$;

– Вибір найімовірнішого користувача – $i^* = \underset{i}{\operatorname{argmax}} S_f(i)$.

2.4.2. Алгоритми і моделі розпізнавання голосу

Модуль розпізнавання голосу аналізує другу модальність у мультимодальній системі біометричної аутентифікації. Він складається з таких етапів (рис. 2.10):

– попередня обробка аудіосигналу;

– побудова спектрограми/Mel-спектру;

– екстракція ознак вектора-ембедингу голосового сигналу;

– порівняння вектора-ембедингу голосового сигналу із наявними зразками у базі даних.

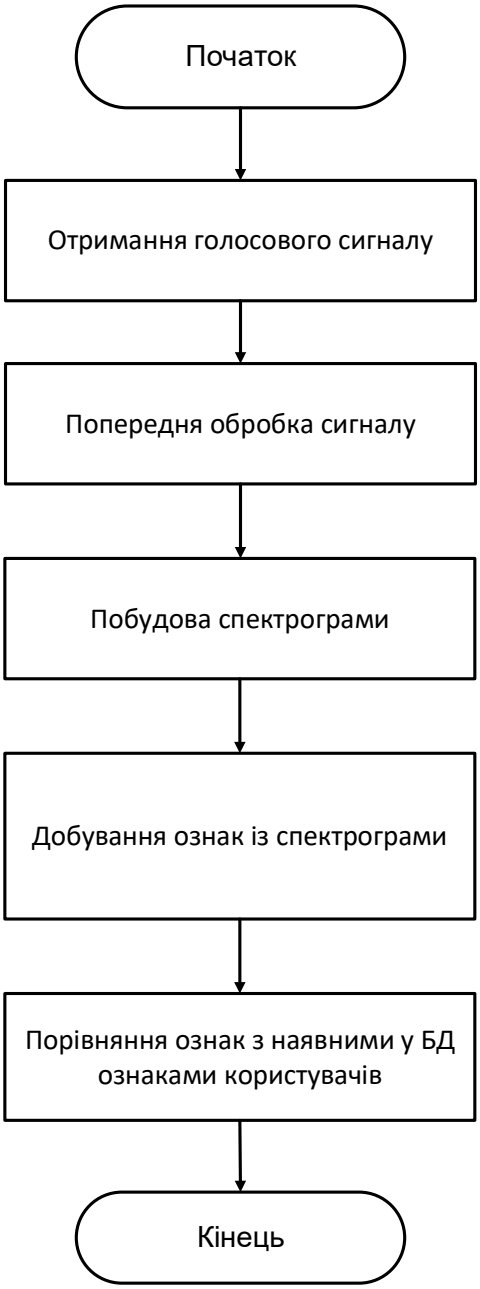


Рис. 2.11. Алгоритм розпізнавання голосу користувача комп’ютерної системи

У кваліфікаційній роботі для реалізації модуля розпізнавання голосу запропоновано використати ECAPA-TDNN, в якості основної моделі для формування векторів-ембедингу голосу та x-vector (TDNN), як резервну або спрощену модель, яка може використовуватися на пристроях з обмеженими апаратними ресурсами. В якості метрики при порівнянні ознак голосу доцільно використовувати косинусну міру подібності.

Ці моделі забезпечують баланс між точністю, стійкістю до шумів, швидкодією та можливістю практичної імплементації.

Аудіосигнал перед подачею до нейромережі проходить кілька етапів обробки. Перша стадія передбачає нормалізацію та фільтрацію. Нормалізація проводиться з метою приведення амплітуди звукового сигналу до стандартного діапазону. Фільтрація аудіосигналу забезпечує підсилення високочастотних компонентів мови шляхом компенсації згасання високих частот у мікрофоні, збільшення відношення сигнал/шум у високочастотній області. Це дає змогу отримати більш точнішу Mel-спектрограму.

Оскільки мовний сигнал є нестаціонарним у часі, його обробку здійснюють на коротких інтервалах (фреймах), протягом яких сигнал можна вважати квазістаціонарним. Типова тривалість фрейму становить 20–30 мс, з перекриттям між сусідніми фреймами 10–15 мс. Такий підхід дозволяє відслідковувати зміну спектральних характеристик у часі та формувати послідовність кадрів для подальшої обробки нейромережею.

До кожного фрейму застосовується віконна функція (найчастіше вікно Геммінга). Це дозволяє згладити краї фрейму, зменшити спектральні витікання та артефакти, які виникають при «жорсткому» обрізанні сигналу. У результаті на виході формується послідовність коротких відрізків сигналу, кожен з яких має згладжені краї і готовий до спектрального перетворення.

Після застосування фільтрації, фреймування та віконування кожен фрейм перетворюється у спектральне представлення. У розроблюваній системі для цього використовується Mel-спектрограма (рис. 2.12), яка відображає енергетичний розподіл сигналу за частотними смугами, які близькі до сприйняття людиною.

Таке представлення є стандартом для сучасних систем розпізнавання голосу та служить вхідними даними для нейромережевої моделі ECAPA-TDNN та x-vector, яка будує компактні вектори-ембедінги голосу.

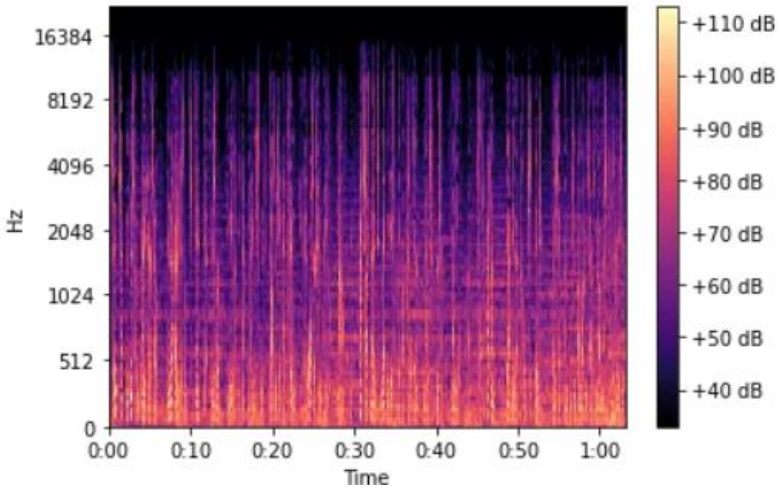


Рис. 2.12. Mel-спектрограма

Узагальнену схему попереднього опрацювання голосового сигналу візуально представлено на рис. 2.13.

Рис. 2.13. Попередня обробка голосового сигналу

В якості основної моделі для розпізнавання голосу користувача комп’ютерної системи використовується нейронна мережа ECAPA-TDNN, альтернативна модель – x-vector.

Архітектура x-vector складається з кількох TDNN-шарів (time-delay neural networks), які обробляють фрейми голосу у часовому контексті. Після frame-level обробки застосовується statistics pooling, що формує глобальні статистики (середнє та дисперсію), а далі – кілька dense-шарів для отримання вектора-ембедингу розмірністю 128 або 512.

Схематично структуру моделі x-vector показано на рис. 2.14.

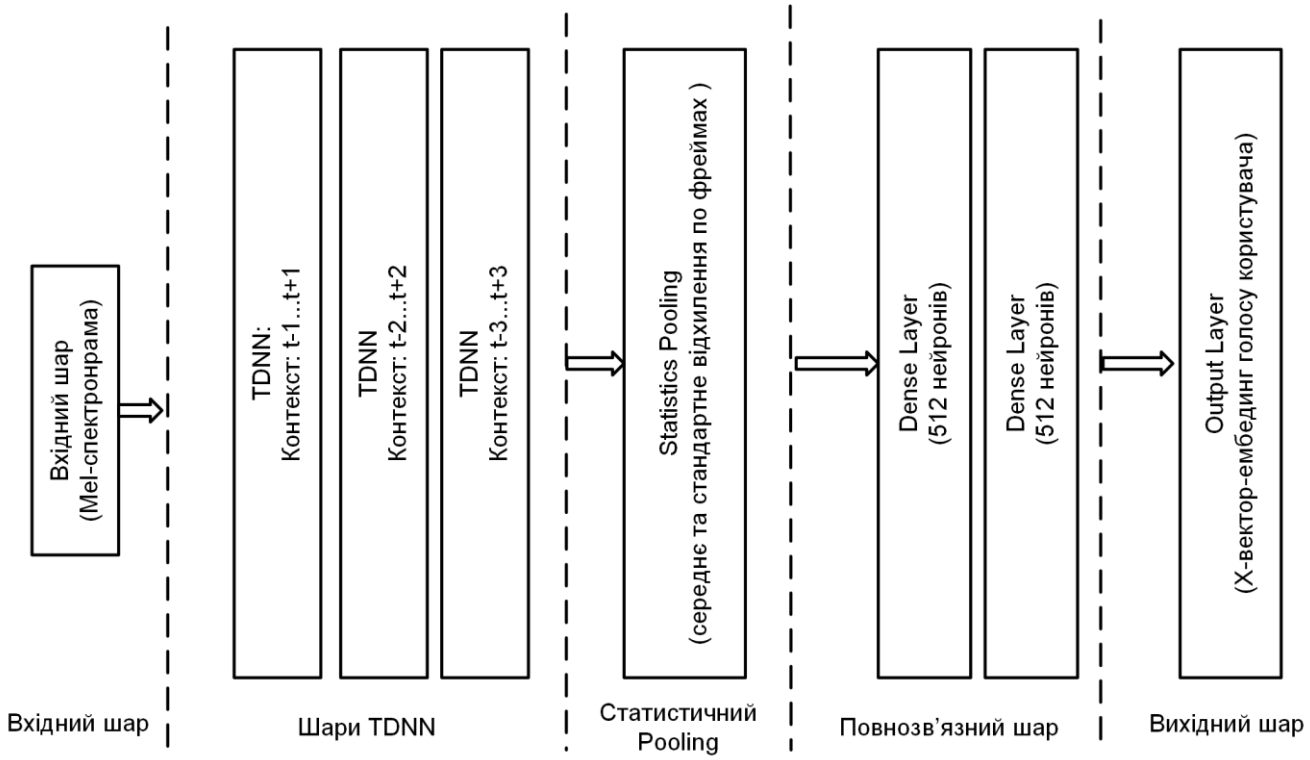


Рис. 2.14. Архітектура нейромережі x-vector

До структури ECAPA-TDNN входять:

– SE-Res2Blocks – блок розширення TDNN з багатоканальною фільтрацією;

– Channel Attention – блок підсилення найбільш інформативних частотних ознак;

– Attentive Statistics Pooling – блок зважування фреймів замість простого усереднення;

– Final Embedding Layer (192D / 256D) – вектор ознак голосу користувача комп'ютерних систем.

На рис. 2.15 наведено архітектуру нейронної мережі ECAPA-TDNN.

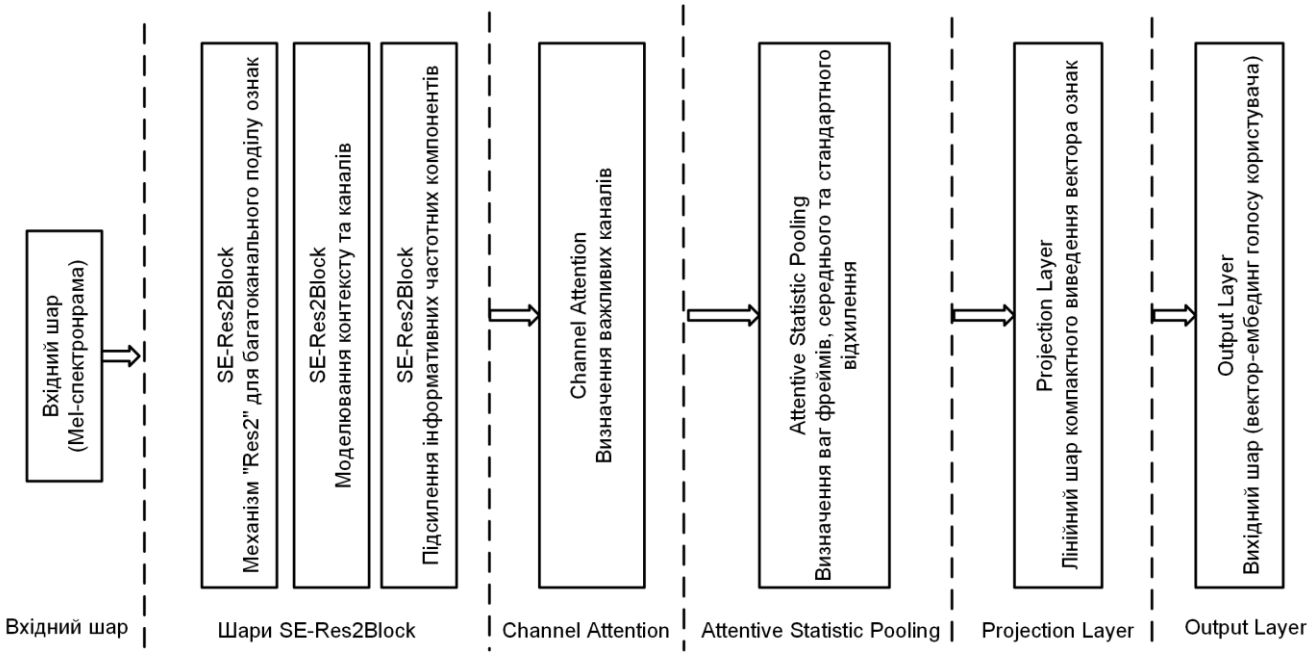


Рис. 2.15. Архітектура ECAPA-TDNN

Ця архітектура є більш точною та стійкою до шумів, тому вона використовується як основна модель голосової біометрії у проектованій біометричній системі аутентифікації користувачів комп’ютерної системи.

2.5. Висновки до розділу

Визначено структурні компоненти апаратної та програмної частини комплексу біометричної мультимодальної аутентифікації користувачів комп’ютерної системи, а також встановлено вимоги до їх сумісності, що дало змогу побудувати узагальнену архітектуру системи та забезпечити обробку зображення обличчя і голосового сигналу в режимі реального часу.

Розроблено формалізований опис процесу мультимодальної аутентифікації, який включає визначення простору вхідних даних, моделей відображення для кожної модальності, метрик подібності та критеріїв прийняття рішень, що дало можливість кількісно виражати якість розпізнавання користувача комп’ютерної системи.

Обґрунтовано вибір алгоритмів і моделей для детекції та розпізнавання обличчя, зокрема застосування MTCNN для виявлення облич та ключових точок і MobileFaceNet з функцією втрат ArcFace для формування високоякісних векторів-ембедингів. Такий вибір забезпечує високу точність і можливість реалізації на Raspberry Pi, що є критичним для апаратно-програмного комплексу.

Обґрунтовано вибір моделей для розпізнавання голосу, де основними стали ECAPA-TDNN як високоточна модель та x-vector як ресурсоефективна альтернатива. Використання цих моделей дає змогу отримати стійкі до шуму та варіативності мовлення голосові ембединг-вектори і реалізувати розпізнавання у реальному часі на вбудованих системах

Визначено та формалізовано критерії оцінювання роботи системи, включно з метриками FAR, FRR, EER та вимогами до максимально допустимої затримки обробки.

РОЗДІЛ 3
АПАРАТНЕ І ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ КОМПЛЕКСУ МУЛЬТИМОДАЛЬНОЇ БІОМЕТРИЧНОЇ АУТЕНТИФІКАЦІЇ КОРИСТУВАЧІВ КОМП'ЮТЕРНОЇ СИСТЕМИ

3.1. Архітектурні особливості системи біометричної аутентифікації на апаратному рівні

Апаратно-програмний комплекс мультимодальної біометричної аутентифікації особи базується на сучасних вбудованих компонентах, які забезпечують достатню продуктивність для обробки зображень та голосових сигналів у режимі реального часу. У цьому підрозділі описано елементну базу системи, її структурну схему, характеристики компонентів та їх інтеграцію у єдиний комплекс.

До складу апаратного забезпечення комплексу входять такі основні елементи:

– одноплатний комп'ютер Raspberry Pi 4 Model B;

– камера для візуальної аутентифікації (Raspberry Pi Camera Module v3) або як альтернатива, USB-камера FullHD;

– мікрофон з USB-інтерфесом Fifine K669B для збирання голосових даних;

– дисплей 7" Raspberry Pi Touch Display;

– карта пам'яті microSD 32-64 ГБ класу A1;

– блок живлення 5.1 В / 3 А;

– корпус із системою охолодження.

Усі вибрані компоненти є сумісними з Raspberry Pi OS та можуть працювати у середовищах Python, ONNXRuntime та бібліотеках обробки сигналів.

Raspberry Pi 4 був обраний як основна обчислювальна платформа завдяки наявності 4-ядерного ARM Cortex-A72 процесора (1.5 ГГц), підтримки до 8 ГБ RAM, що дозволяє виконувати запуск нейронних мереж MobileFaceNet та ECAPA-TDNN.

Окрім цього, даний центральний вузол керування системи біометричної аутентифікації за двома модальностями має апаратну підтримку USB 3.0, що необхідна для високошвидкісних камер та мікрофонів, а також наявна функція апаратного декодування відео. Raspberry PI 4 Model B підтримує Gigabit Ethernet / Dual-band Wi-Fi для передачі даних через комп'ютерні мережі.

Пропонована версія Raspberry Pi забезпечує достатню продуктивність для вирішення задач детекції облич (MTCNN) у 10–15 FPS та генерації face-embedding MobileFaceNet швидше ніж 50 мс. Окрім цього, обробка голосового сигналу тривалістю 1-2 секунд виконується ECAPA-TDNN менш ніж за 250 мс (ONNXRuntime).

У системі використовується Raspberry Pi Camera Module v3 12 Мп з роздільною здатністю 1920×1080 і частотою кадрів 30 FPS. Дана камера підтримує автоматичний баланс білого, має низькі шуми та підтримує роботу з такими Python-бібліотеками, як picamera2, OpenCV.

У біометричній системі аутентифікації користувачів комп'ютерної системи камера використовується для знімання обличчя в реальному часі з подальшою детекцією MTCNN та вирівнюванням обличчя за 5 ключовими точками, що в результаті дає змогу отримати вектори-ембедінги MobileFaceNet. Raspberry Pi Camera Module v3 оптимально працює на відстані 0.4-1.2 м.

Для збирання аудіоданих обрано USB-мікрофон Fifine K669B, який працює у частотному діапазоні 20 Гц – 20 кГц, частота дискретизації 48 кГц. Даний мікрофон має спрямованість кардіоїда та низький рівень власного шуму.

Мікрофон Fifine K669B забезпечує якість голосового сигналу, яка достатня для розпізнавання мовлення користувача комп'ютерної системи моделями x-vector та ECAPA-TDNN. Розміщення мікрофона передбачає спрямованість на користувача, на відстані від 20 до 50 см.

Для локальної взаємодії з комплексом використовується дисплей Raspberry Pi 7" Touch Display з ємнісним сенсором та роздільною здатністю 800×480.

Основне призначення дисплея у системі мультимодальної біометричної аутентифікації полягає у виведенні інформації про хід аутентифікації, перегляд

логів і статистики. Наявність сенсорного інтерфейсу дозволяє створити повноцінну локальну панель керування.

Структурну схема апаратної частини системи біометричної аутентифікації представлено на рис. 3.1.

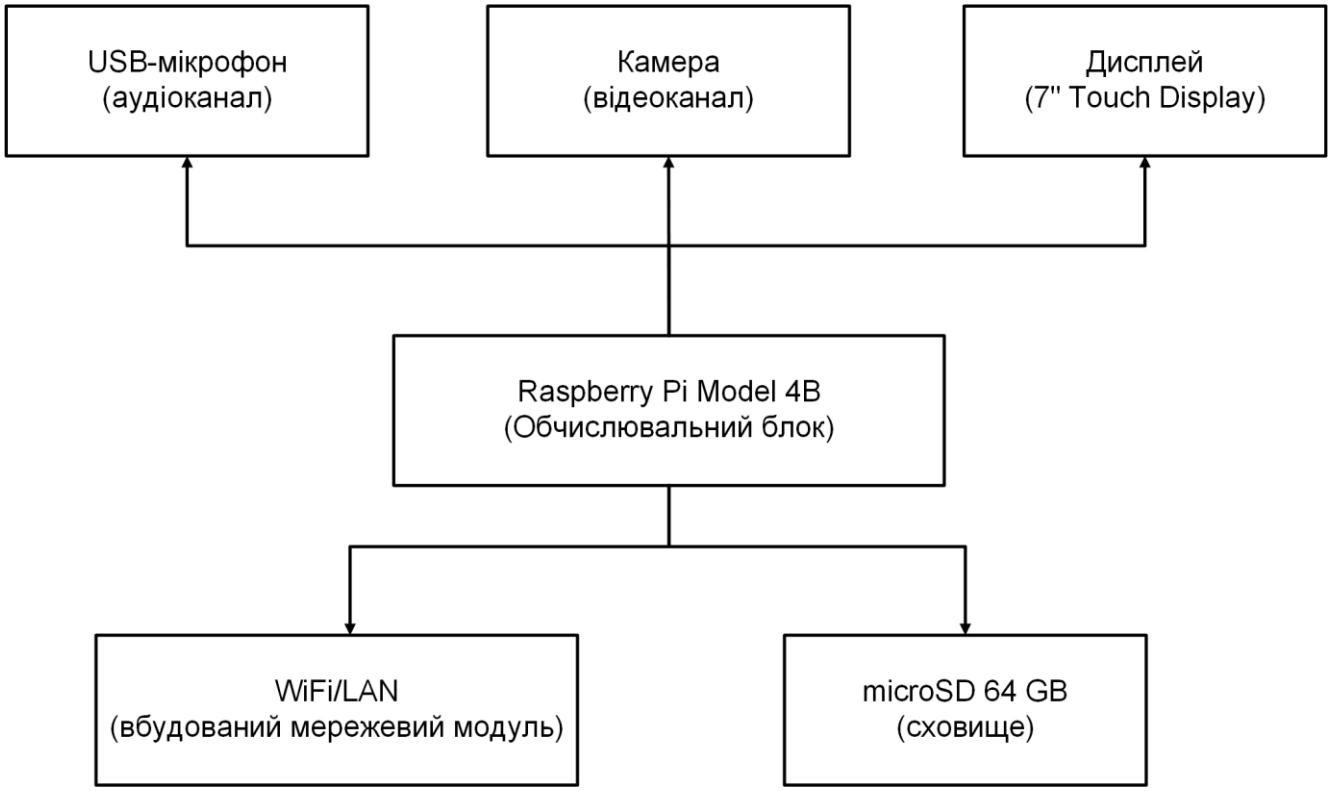


Рис. 3.1. Структурна схема апаратної конфігурації системи

Вибір Raspberry Pi 4 обумовлений факторами оптимального балансу продуктивності та енергоспоживання, можливості запуску моделей у форматі ONNX з апаратними оптимізаціями. Окрім цього, ризики подальшого супроводу системи є доволі низькими, оскільки наявна розвинута спільнота підтримки бібліотек для комп’ютерного зору й аналізу мовлення. Така архітектура біометричної системи забезпечує можливість масштабування до кластерної обробки та підтримує локальний інтерфейс адміністратора без додаткових пристроїв

Модулі камери та мікрофона відповідають вимогам до якості сигналів для стабільної роботи MTCNN, MobileFaceNet та ECAPA-TDNN.

3.2. Налаштування програмного забезпечення на системному рівні

Для коректного функціонування біометричної мультимодальної системи аутентифікації користувачів комп'ютерної системи необхідно спочатку сконфігурувати програмне забезпечення на системному рівні. Налаштування виконуються в операційній системі Raspberry Pi OS (64-bit), яка забезпечує повну сумісність із бібліотеками Python, драйверами камер та аудіопристроїв, а також середовищем виконання нейронних мереж у форматі ONNXRuntime.

Базові інструкції щодо встановлення операційної системи на Raspberry PI можна знайти на офіційному сайті виробника [18]. Дотримуючись таких інструкцій на карту пам'яті microSD було встановлено Raspberry Pi OS (Bookworm, 2023). Після першого запуску операційної системи оновлено системні пакети до останньої доступної версії (рис. 3.2.).

```
sudo apt update
sudo apt upgrade -y
```

Рис. 3.2. Оновлення системних пакетів

Оскільки, у роботі біометрична ідентифікація передбачає застосування двох модальностей, зокрема, аналізу зображення з камери, то відповідно необхідно активувати інтерфейс камери та I2C/USB. Це забезпечується шляхом виконання команди показаної на рис. 3.3.

```
sudo raspi-config
```

Рис. 3.3. Активація інтерфейсу камери

Після активації камери, у меню активуються наступні пункти:

– Interface Options → Camera → Enable;

– Interface Options → I2C → Enable;

– Performance Options → GPU Memory = 256 MB для стабільної роботи OpenCV.

Після налаштування камери необхідно встановити бібліотеки Python для функціонування нейромереж, які описані у розділі 2. Для цього виконується команда, проілюстрована на рис. 3.4.

```
sudo apt install python3-pip python3-opencv python3-numpy python3-pyaudio \
libatlas-base-dev portaudio19-dev ffmpeg -y
```

Рис. 3.4. Встановлення необхідних Python-бібліотек

У системі використовується камера Raspberry Pi Camera Module v3. Для виявлення камери використовується скрипт, наведений на рис. 3.5.

```
libcamera-hello
```

Рис. 3.5. Скрипт виявлення камери

У випадку успішної аутентифікації камери, необхідно отримати зображення (кадр) з відеопотоку. Для цього використовується скрипт, показаний на рис. 3.6.

```
from picamera2 import Picamera2
import cv2

cam = Picamera2()
cam.configure(cam.create_preview_configuration(main={"size": (640, 480)}))
cam.start()

frame = cam.capture_array()
cv2.imshow("Frame", frame)
cv2.waitKey(0)
```

Рис. 3.6. Скрипт отримання зображення з камери

У системі біометричної аутентифікації камера застосовується для захоплення і в подальшому опрацювання модальності (зображення) при розпізнаванні обличчя. Вона забезпечує:

– отримання потокового зображення;
– передавання кадру до модуля MTCNN;
– вирівнювання та нормалізація обличчя;
– передавання в MobileFaceNet для формування вектора-ембедингу.

Мікрофон підключається через USB і визначається системою як ALSA-пристрій. Для перевірки підключення застосовується команда, яка показана на рис. 3.7.

```
arecord -l
```

Рис. 3.7. Перевірка підключення мікофону

Для роботи з аудіо сигналом потрібно проінсталювати бібліотеку librosa (рис. 3.8).

```
pip install sounddevice librosa
```

Рис. 3.8. Інсталяція бібліотеки для роботи з аудіо

Запис голосу користувача комп’ютерної системи виконається за допомогою імплементованого програмного коду, який показаний на рис. 3.9.

У розроблюваній системі біометричної аутентифікації користувачів комп’ютерної системи мікрофон та відповідне програмне забезпечення реалізує функції отримання голосового сигналу, нормалізації та препроцесингу, перетворення у Mel-спектрограму та передавання у модель ECAPA-TDNN.

```
import sounddevice as sd
import numpy as np
import librosa

duration = 2.0
sample_rate = 16000

audio = sd.rec(int(duration * sample_rate), samplerate=sample_rate, channels=1)
sd.wait()

# Попередня обробка
audio = librosa.util.normalize(audio)
```

Рис. 3.9. Скрипт для записування голосу користувача

Для того, щоб забезпечити функціональну працездатність пристрою виводу інформації 7" Touch Display необхідно проінсталювати відповідну бібліотеку. Команда для інсталяції виглядає так, як показано на рис. 3.10.

```
sudo apt install xserver-xorg xinit lightdm
```

Рис. 3.10. Інсталяція бібліотеки для роботи з 7" Touch Display

Далі у файлі /boot/config.txt необхідно активувати функції, які показано на рис. 3.11.

```
dtoverlay=vc4-kms-v3d
lcd_framerate=60
```

Рис. 3.11. Активація функцій файлу конфігурації екрану

Виведення інформаційних повідомлень на екран забезпечує скрипт приведений на рис. 3.12.

```
import tkinter as tk

root = tk.Tk()
root.title("Система ідентифікації")
label = tk.Label(root, text="Очікування користувача...", font=("Arial", 24))
label.pack()
root.mainloop()
```

Рис. 3.12. Скрипт виводу повідомлень на екран 7" Touch Display

Для коректної роботи локального бекенду та адміністративної панелі виконується статична адресація. Конфігурація мережі представлена на рис. 3.13.

```
interface wlan0
static ip_address=192.168.1.50/24
static routers=192.168.1.1
static domain_name_servers=8.8.8.8
```

Рис. 3.13. Налаштування конфігурації мережі для передачі даних

3.3. Архітектура програмного забезпечення прикладного рівня

Програмне забезпечення мультимодальної біометричної системи складається з кількох логічних модулів, які взаємодіють через внутрішні API та забезпечують повний цикл обробки візуальної та аудіоінформації. Архітектура програмного забезпечення прикладного рівня розроблена за принципом модульності, що підвищує гнучкість, масштабованість та можливість незалежного оновлення кожного компонента.

До складу програмної частини входять такі компоненти системного рівня:

– модуль роботи з камерою ;
– модуль роботи з мікрофоном ;
– модуль детекції та нормалізації обличчя (MTCNN);

– модуль отримання векторів-ембедингів обличчя (MobileFaceNet);

– модуль отримання векторів-ембедингів обличчя (ECAPA-TDNN);

– модуль мультимодального злиття результатів.

До вищого програмного рівня відносяться такі компоненти як:

– база даних та модуль роботи з користувачами;

– веб-інтерфейс адміністратора;

– серверна частина (FastAPI backend).

Запропонована загальна архітектура програмного забезпечення прикладного рівня проілюстрована на рис. 3.14.

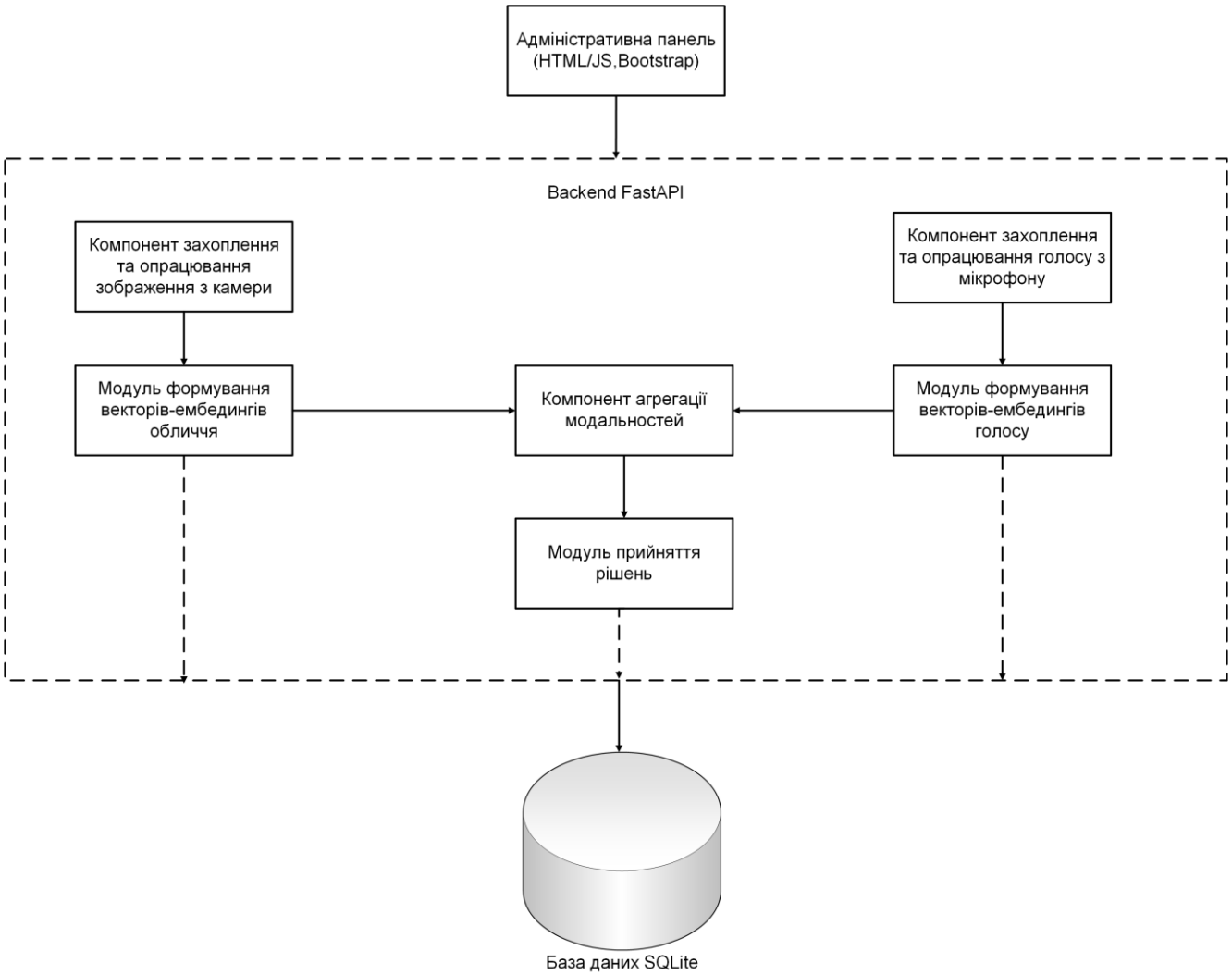


Рис. 3.14. Архітектура програмного забезпечення біометричної системи на прикладному рівні

Серверна частина написана на Python 3.11 + FastAPI, що забезпечує легку побудову REST API та асинхронну обробку.

Структура директорії backend/ представлена на рис. 3.15.

```
backend/
├── main.py
├── routers/
│   ├── users.py
│   ├── face.py
│   ├── voice.py
│   ├── fusion.py
│   └── logs.py
├── models/
│   ├── mobilefacenet.onnx
│   ├── ecapa.onnx
│   └── mtcnn/
├── db/
│   ├── database.py
│   └── schema.sql
└── utils/
    ├── preprocessing.py
    ├── alignment.py
    ├── similarity.py
    └── fusion.py
```

Рис. 3.15. Структура backend-директорії

Фрагмент головного файлу FastAPI показано на рис. 3.16.

```python
from fastapi import FastAPI
from routers import face, voice, users, fusion, logs

app = FastAPI(title="Biometric System API")

app.include_router(face.router)
app.include_router(voice.router)
app.include_router(users.router)
app.include_router(fusion.router)
app.include_router(logs.router)
```

Рис. 3.16. Фрагмент файл FastAPI

Адміністративна панель реалізована як веб-інтерфейс і доступна за локальною IP-адресою Raspberry Pi – 192.168.1.50:8000/admin.

Функціональність адміністратора включає:

– реєстрацію нового користувача (обличчя + голос);
– перегляд списку користувачів;
– перегляд журналу доступів;
– видалення чи оновлення шаблонів;
– проведення тесту аутентифікації;
– налаштування порогів та ваг мультимодальної агрегації.

JavaScript-запит до backend-частини програмного забезпечення біометричної аутентифікації користувачів наведено на рис. 3.17.

```
async function loadUsers() {
    const response = await fetch("/users/list");
    const users = await response.json();

    let rows = "";
    users.forEach(user => {
        rows += `<tr>
                    <td>${user.id}</td>
                    <td>${user.name}</td>
                    <td>${user.face_count}</td>
                    <td>${user.voice_count}</td>
                    <td><button onclick="deleteUser(${user.id})">Видалити</button></td>
                </tr>`;
    });

    document.querySelector("#usersTable tbod innerHTML = rows;
}
```

Рис.3.17. JavaScript-запит до backend-частини

Адміністративна панель побудована з використанням Bootstrap 5, Vanilla JS, REST API та візуальних карток та таблиць. На рис. 3.18 наведено інтерфейс головної сторінки адміністративної панелі.

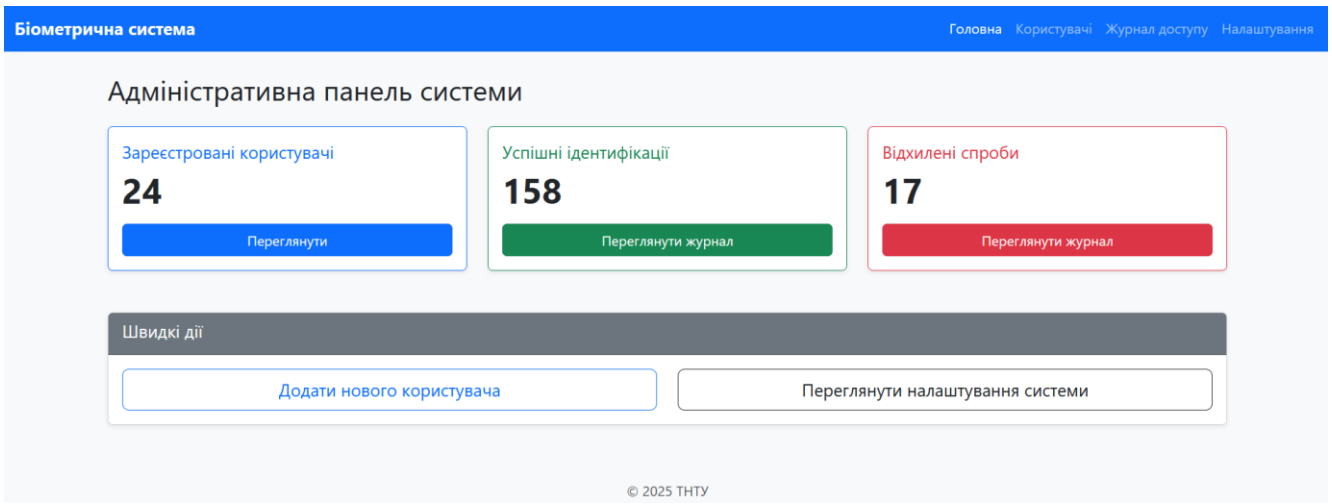


Рис. 3.18. Головна сторінка адміністративної панелі

На головній сторінці адміністративної панелі реалізовані функції переходу до меню для роботи з користувачами, доступу до журналу логування та налаштування біометричної системи. Окрім цього, на головній сторінці відображається статистична інформація щодо кількості зареєстрованих користувачів, кількості успішних та відхилених ідентифікацій. Панель швидких дій забезпечує можливість додавання нових користувачів та перегляду налаштувань системи.

Сторінка адміністративної панелі, що відповідає за роботу з даними про користувачів комп’ютерної системи представлена на рис. 3.19.

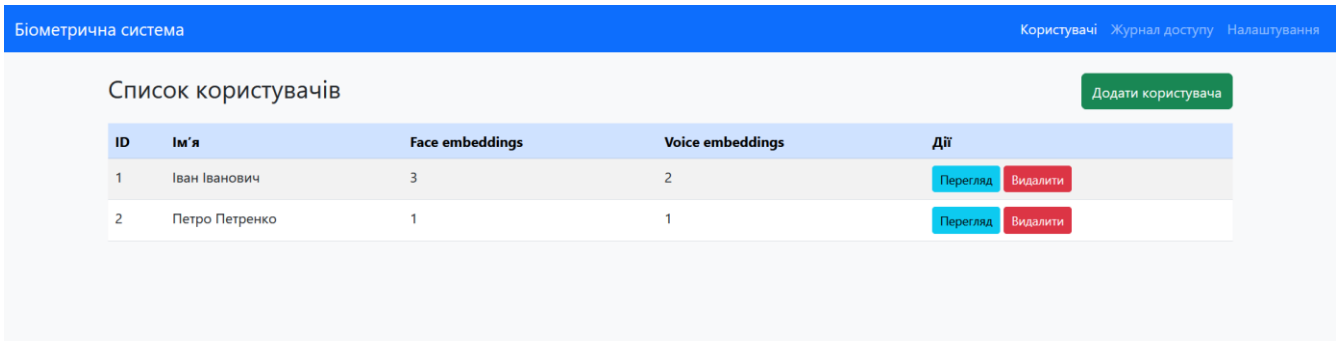


Рис. 3.19. Сторінка для роботи з даними про користувачів системи

Як видно з рис. 3.19, основними функціями, які доступні для роботи з даними про користувачів є:

– перегляд існуючих у БД записів про користувачів;
– видалення даних про користувачів;
– запис інформації про нових користувачів.

Форма запису інформації про користувачів біометричної системи представлена на рис. 3.20.

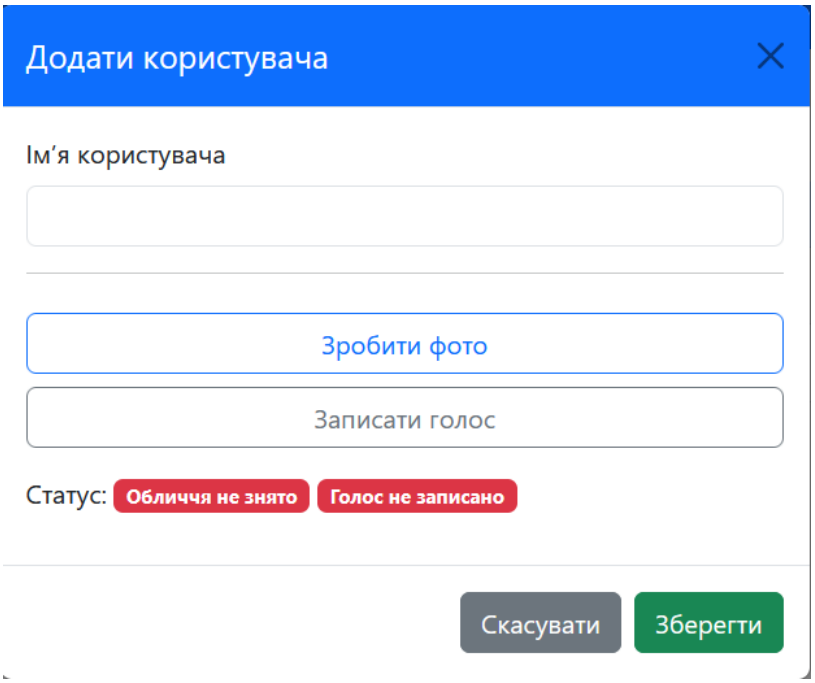


Рис. 3.20. Форма реєстрації користувача біометричної системи

Форма, показана на рис. 3.20, забезпечує можливість отримання кількох зразків обличчя з камери, а також отримання голосового сигналу з мікрофону. На формі передбачено поле для вказання особистих даних користувача, а також реалізовано функцію, що показує статус успішності зчитування біометричних даних.

Усі дані щодо входу та стану аутентифікації користувачів у системі відображаються у відповідному меню «Журнал доступу». Журнал доступу представляється у вигляді таблиці, де відображається дата і час спроби аутентифікації, значення точності розпізнавання за кожною модальністю та інтегральний показник точності розпізнавання користувача. У випадку успішної аутентифікації у таблиці відображається також ідентифікатор користувача. На

рис. 3.21 показано сторінку журналу доступу у систему мультимодальної біометричної аутентифікації.

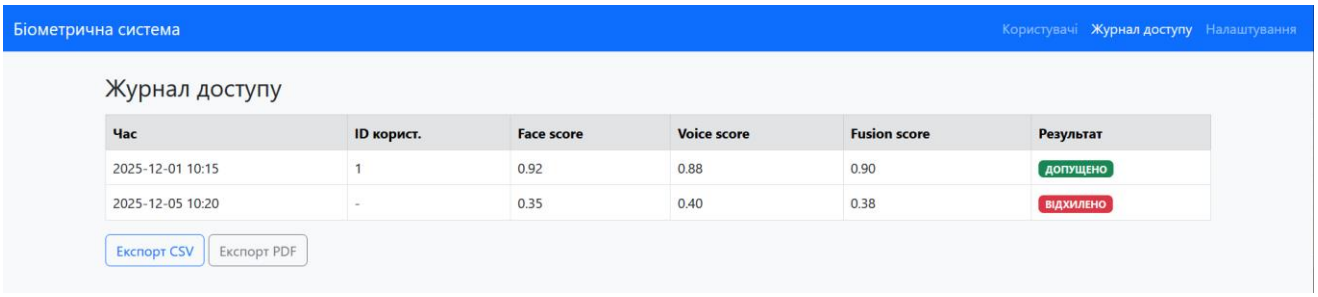


Рис. 3.21. Журнал логування біометричної аутентифікації користувача

Як видно з рис. 3.21, доступними також є функції експорту журналу у форматах CSV та PDF.

Налаштування параметрів біометричної системи розпізнавання на основі аналізу обличя та голосу корситувачів реалізовано у відповідному однойменному меню. Сторінка з налаштуваннями показана на рис. 3.22.

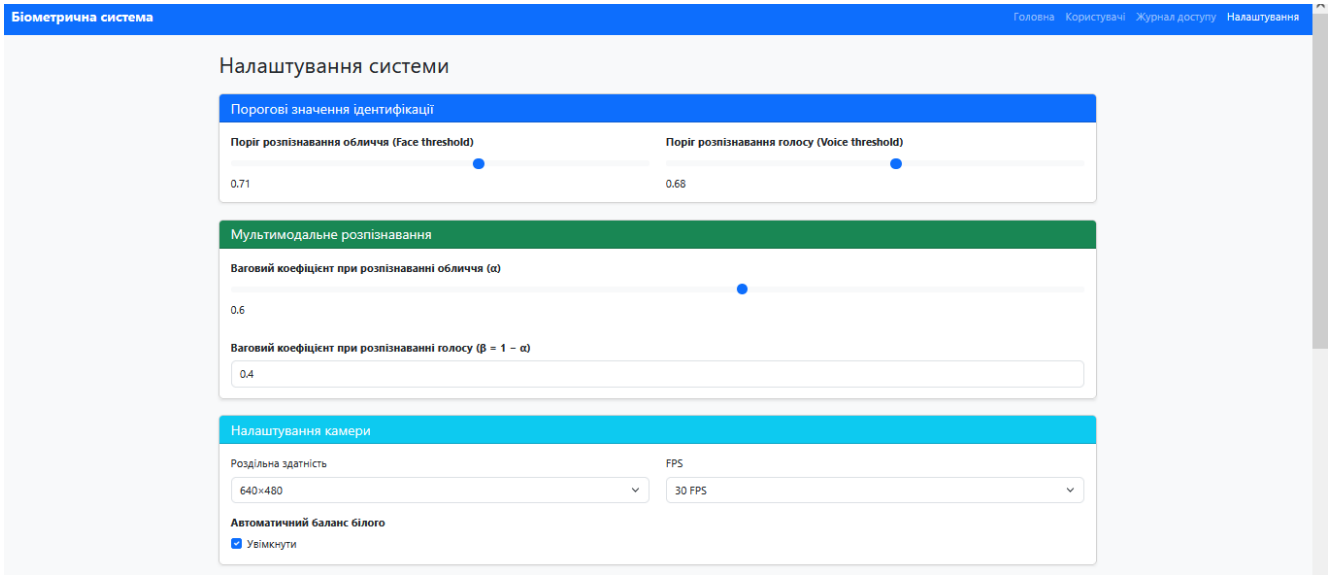


Рис. 3.22. Налаштування параметрів біометричної системи

Сторінка з налаштуваннями біометричної мультимодальної системи складається з кількох панелей:

– панель встановлення порогових значень при аутентифікації користувача окремо за обличчям та голосом;

– панель мультимодального розпізнавання;

– панель налаштування параметрів камери;

– панель налаштування аудіомодуля;

– панель системних параметрів;

– функції очистки журналу логування;

– функції збереження змін та відновлення стандартних параметрів.

Нижня частини сторінки налаштувань біометричної системи проілюстрована на рис. 3.23.

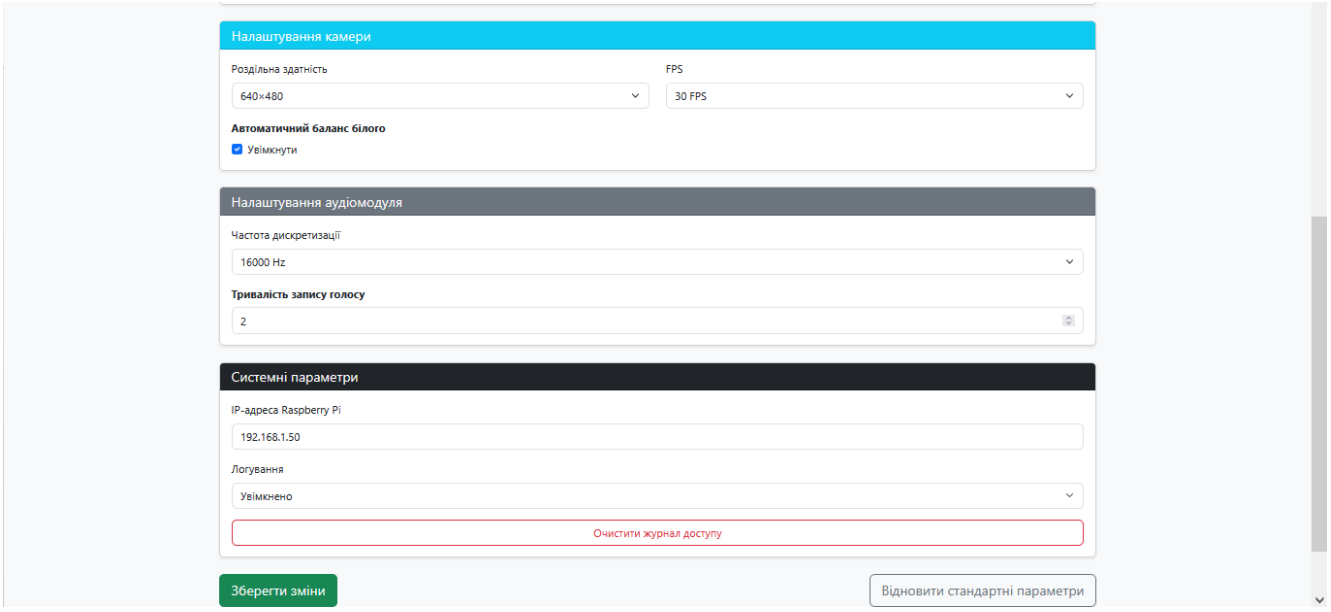


Рис. 3.23. Нижня частина сторінка налаштування параметрів біометричної аутентифікації користувачів

Налаштування камери передбачає встановлення доступних параметрів роздільної здатності, кількості кадрів за секунду та фільтрації (автоматичний баланс білого). Налаштування параметрів мікрофону передбачає встановлення частоти дискретизації та тривалості запису у секундах.

Системні параметри дозволяють встановити IP-адресу Raspberry PI та увімкнення та вимкнення журналу логування.

3.4. Проектування та реалізація схеми бази даних біометричної системи аутентифікації користувачів

У роботі, для зберігання даних, запропоновано реалізувати реляційну схему бази даних за допомогою SQLite, оскільки дане середовище забезпечує легкість, мобільність і не вимагає окремого серверу. Для ефективної роботи системи біометричної аутентифікації користувачів достатньою реалізувати 4 таблиць. Таблиця «User» зберігає інформацію про особисті дані користувача і дату реєстрації у системі. У табл. 3.1 наведено структуру «Users».

Таблиця 3.1

Структура таблиці «Users»

Поле	Тип	Опис
user_id	INTEGER (PK)	Унікальний ідентифікатор
name	TEXT	Ім’я користувача
created_at	DATETIME	Дата реєстрації

Для зберігання даних про зображення обличчя користувача створено таблицю «Face_Embeddings», структуру якої подано у табл. 3.2.

Таблиця 3.2

Структура таблиці «Face_Embeddings»

Поле	Тип	Опис
emb_id	INTEGER (PK)	Ідентифікатор
user_id	INTEGER (FK)	Зовнішній ключ
embedding	BLOB	128D MobileFaceNet
quality	REAL	Якість кадру

Таблиця «Face_Embedding» містить інформацію про вектор-ембедингу конкретного користувача, ідентифікатор користувача, що є зовнішнім ключем та якість отриманого з камери зображення. Для зберігання інформації про модальність, яка відповідає за голос користувач створено таблицю «Voice_Embeddings». Її структуру представлено у табл. 3.3.

Таблиця 3.3

Структура таблиці «Voice_Embeddings»

Поле	Тип	Опис
emb_id	INTEGER (PK)	Ідентифікатор
user_id	INTEGER (FK)	Зовнішній ключ
embedding	BLOB	192D ECAPA
snr	REAL	Signal-to-noise ratio

У базі даних створено таблицю логування процесу аутентифікації користувачів «Access_Logs». До складу цією таблиці входять значення точності розпізнавання зображення обличчя користувача та його голосу, а також зваженне значення по двох модальностях. Структура «Access_Logs» представлена у табл. 3.4.

Таблиця 3.4

Структура таблиці «Access_Logs»

Поле	Тип	Опис
log_id	INTEGER (PK)	Ідентифікатор
user_id	INTEGER (FK)	Зовнішній ключ
face_score	REAL	Точність розпізнавання обличчя
voice_score	REAL	Точність розпізнавання голосу
fusion_score	REAL	Точність за двома модальностями
timestamp	DATETIME	Часова мітка

ER-діаграма створеної бази даних наведена на рис. 3.24.

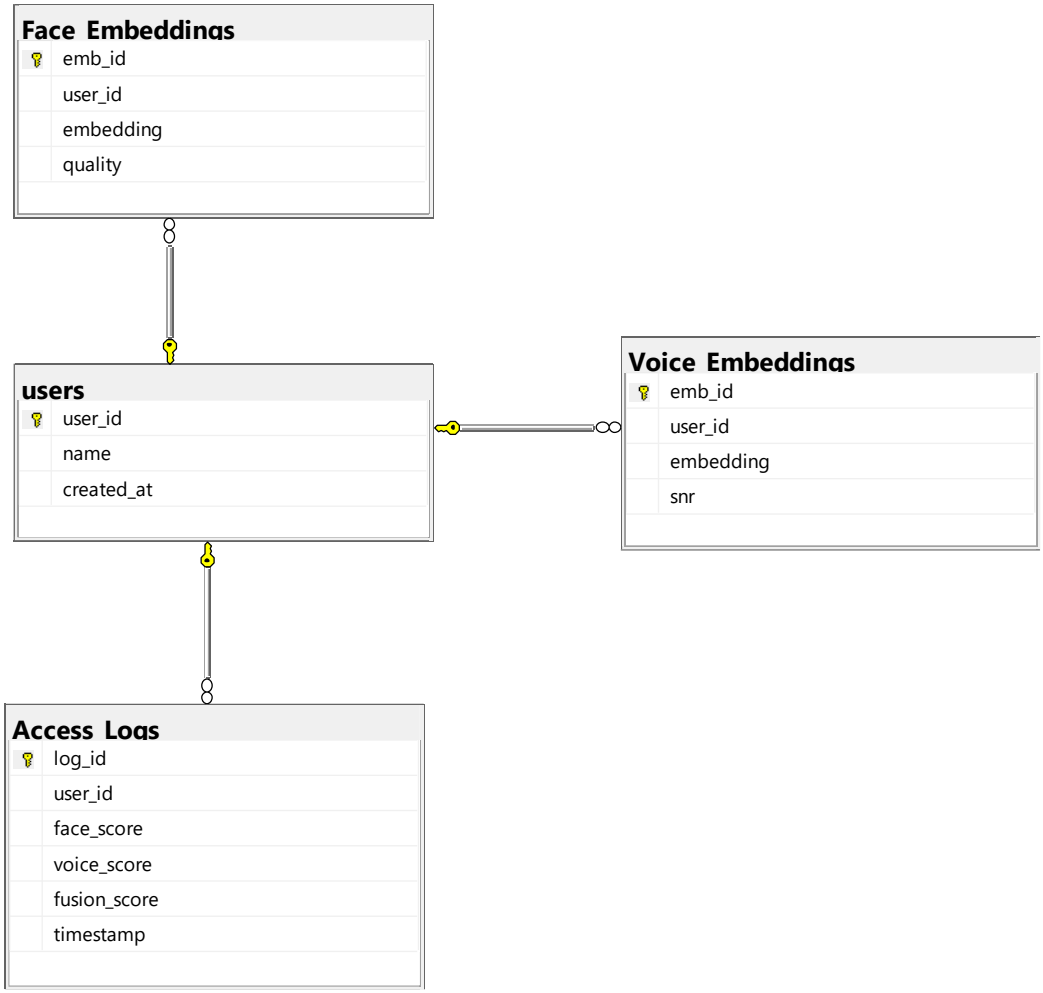


Рис. 3.24. ER-діаграма БД системи біометричної аутентифікації

База даних є нормалізованою і реалізує зв’язки типу «один-до-багатьох». В перспективі можна розширити функціональність системи і в тому числі структуру БД, зокрема, щодо зберігання налаштувань системи та інших необхідних функцій.

3.5. Реалізація нейронних мереж мультимодальної біометричної аутентифікації

У цьому підрозділі наведено практичну реалізацію нейронних мереж, що використовуються в апаратно-програмному комплексі мультимодальної біометричної аутентифікації користувачів комп’ютерної системи.

Реалізація базується на моделях, обґрунтованих у розділі 2, а саме: MTCNN для детекції облич, MobileFaceNet з функцією втрат ArcFace для формування векторів ознак обличчя, а також ECAPA-TDNN для отримання голосових ембединг-векторів користувачів. Усі моделі інтегровані у програмне середовище Raspberry Pi 4 з урахуванням обмежень обчислювальних ресурсів.

Для виявлення облич на зображеннях, отриманих з камери, у системі використовується багатокаскадна нейронна мережа MTCNN (Multi-task Cascaded Convolutional Neural Network). Її вибір обумовлений високою точністю локалізації облич і ключових точок, а також можливістю роботи у режимі реального часу на вбудованих платформах.

MTCNN складається з трьох послідовних мереж: P-Net, R-Net та O-Net, які виконують грубу детекцію, уточнення меж та визначення п’яти ключових точок обличчя (очі, ніс, кути рота). Отримані властивості використовуються для геометричного вирівнювання обличчя перед подачею у модель MobileFaceNet.

Фрагмент програмної реалізації детекції обличчя представлено на рис. 3.25.

```
from mtcnn import MTCNN
import cv2

detector = MTCNN()

image = cv2.imread("frame.jpg")
results = detector.detect_faces(image)

for face in results:
    x, y, w, h = face['box']
    keypoints = face['keypoints']
```

Рис. 3.25. Фрагмент реалізації детекції обличчя користувача

На цьому етапі система отримує рамку обличчя та координати ключових точок, які надалі використовуються для нормалізації положення обличчя. Перед формуванням ембединг-вектора виконується геометричне вирівнювання обличчя на основі 5 ключових точок. Метою цього етапу є приведення обличчя до канонічного положення, що зменшує вплив нахилів голови, поворотів та перспективних спотворень.

Вирівнювання реалізується через афінне перетворення з використанням еталонних координат ключових точок, а програмне реалізація представлена на рис. 3.26.

```python
import numpy as np
import cv2

def align_face(image, src_pts, dst_pts):
    M = cv2.estimateAffinePartial2D(src_pts, dst_pts)[0]
    aligned = cv2.warpAffine(image, M, (112, 112))
    return aligned
```

Рис. 3.26. Функція вирівнювання зображення обличчя

Після вирівнювання зображення масштабується до розміру 112×112 пікселів та нормалізується за значеннями яскравості. Для побудови компактного вектора ознак обличчя у системі використовується нейронна мережа MobileFaceNet, оптимізована для роботи на пристроях з обмеженими ресурсами. Модель реалізує глибину мережу з використанням depthwise separable convolutions та residual-зв’язків, що значно зменшує кількість параметрів.

У системі використовується попередньо навчена модель, експортована у формат ONNX, що дозволяє забезпечити виконання за допомогою ONNX Runtime.

На рис. 3.27 показано програмний код реалізації даної нейронної мережі.

```python
import onnxruntime as ort
import numpy as np

session = ort.InferenceSession("mobilefacenet.onnx")

def extract_face_embedding(face_img):
    input_blob = face_img.astype(np.float32)
    input_blob = np.expand_dims(input_blob, axis=0)
    embedding = session.run(None, {"input": input_blob})[0]
    return embedding
```

Рис. 3.27. Реалізація нейронної мережі розпізнавання обличчя користувача

У результаті виконання програмного коду, представленого на рис. 3.27, одержується 128-вимірний ембединг-вектор, який зберігається у базі даних та використовується для порівняння з еталонними зразками.

Для розпізнавання мовця у системі використовується модель ECAPA-TDNN, яка забезпечує високу точність навіть за наявності шумів. Перед подачею аудіосигналу у нейромережу виконується стандартний ланцюг попередньої обробки: нормалізація, pre-emphasis, фреймування, віконування та формування Mel-спектрограми. На рис. 3.28 наведено програмну реалізацію препроцесингу голосового сигналу.

```python
import librosa
import numpy as np

audio, sr = librosa.load("voice.wav", sr=16000)
audio = librosa.util.normalize(audio)

mel = librosa.feature.melspectrogram(
    y=audio, sr=sr, n_mels=80, fmax=8000
)
mel_db = librosa.power_to_db(mel)
```

Рис. 3.28. Препроцесинг голосового сигналу

Підготовлена Mel-спектрограма подається на вхід ECAPA-TDNN, яка також реалізована у форматі ONNX і показана на рис. 3.29.

```
session = ort.InferenceSession("ecapa_tdnn.onnx")

def extract_voice_embedding(mel_spec):
    mel_spec = np.expand_dims(mel_spec, axis=(0,1))
    embedding = session.run(None, {"input": mel_spec})[0]
    return embedding
```

Рис. 3.29. Одержання вектора ознак голосу користувача

Отриманий вектор ознак голосу має розмірність 192 або 256 та характеризує голос користувача незалежно від конкретного тексту мовлення.

Порівняння біометричних шаблонів здійснюється на основі косинусної подібності, яка є стандартною метрикою для ембединг-просторів. Реалізація метрик подібності показана на рис. 3.30.

```
from numpy.linalg import norm

def cosine_similarity(a, b):
    return np.dot(a, b) / (norm(a) * norm(b))
```

Рис. 3.30. Порівняння біометричних шаблонів на основі метрики косинуса кута

Обчислені значення подібності використовуються як вхідні дані для модуля мультимодальної агрегації.

Таким чином, у роботі реалізовано мультимодальну біометричну систему аутентифікації користувачів комп’ютерної системи, яка функціонує на пристрої з обмеженими ресурсами Raspberry PI 4.

3.6. Висновки до розділу

Розроблено архітектуру апаратного забезпечення апаратно-програмного комплексу мультимодальної біометричної аутентифікації на базі одноплатного комп’ютера Raspberry Pi 4 Model B із використанням модуля камери, USB-мікрофона та дисплейного інтерфейсу, що дало змогу забезпечити збалансоване поєднання продуктивності, енергоефективності та компактності системи при роботі в режимі реального часу.

Реалізовано системне програмне забезпечення апаратних компонентів, включаючи налаштування операційної системи Raspberry Pi OS, конфігурацію камерного та аудіомодулів, а також мережеві параметри взаємодії, що дало змогу організувати стабільне середовище збору мультимодальних біометричних даних та їх передачі до програмних модулів обробки.

Спроєктовано архітектуру програмного забезпечення системи з використанням клієнт-серверного підходу, що включає бекенд-логіку, веб-орієнтовану адміністративну панель та локальну базу даних SQLite, що дало змогу забезпечити централізоване управління користувачами, збереження біометричних шаблонів та журналювання результатів аутентифікації.

Реалізовано нейронні мережі для розпізнавання обличчя та голосу, зокрема MTCNN для детекції та вирівнювання облич, MobileFaceNet для формування векторів ознак обличчя та ECAPA-TDNN для отримання векторів ознак голосу, що дало змогу досягти високої точності аутентифікації при обмежених апаратних ресурсах одноплатного комп’ютера.

Запропоновано та реалізовано програмну інтеграцію мультимодальних нейронних моделей у єдиний обчислювальний цикл, який поєднує результати розпізнавання обличчя та голосу на рівні векторів ознак, що дозволяє підвищити надійність аутентифікації користувачів у порівнянні з монобіометричними підходами.

РОЗДІЛ 4
ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1. Охорона праці

Кваліфікаційна робота магістра присвячена дослідженню та розробці методів і програмно-апаратних засобів біометричної аутентифікації користувачів комп'ютерної системи у режимі реального часу. Процес виконання роботи пов'язаний із тривалою експлуатацією комп'ютерної техніки, периферійних пристроїв, камерних і аудіомодулів, а також з використанням електронних компонентів та програмних засобів обробки даних. У зв'язку з цим особливої актуальності набуває дотримання вимог охорони праці, техніки безпеки, пожежної безпеки та санітарно-гігієнічних норм.

Основними небезпечними та шкідливими факторами під час виконання робіт з розробки та експлуатації системи біометричної аутентифікації є електричний струм, підвищене зорове навантаження, статичні навантаження опорно-рухового апарату, несприятливі параметри мікроклімату, а також можливі електромагнітні випромінювання від комп'ютерної та периферійної техніки [21]. Тому при організації робочого процесу необхідно враховувати вимоги чинних нормативно-правових актів з охорони праці.

Базовим нормативним документом, який регламентує вимоги безпеки при роботі з комп'ютерною технікою, є НПАОП 0.00-7.15-18 «Вимоги щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями», що є чинним на момент виконання кваліфікаційної роботи. Даний документ встановлює вимоги до організації робочих місць, параметрів освітлення, ергономіки, режимів праці та відпочинку, а також до безпечної експлуатації обладнання.

При організації робочого місця розробника системи біометричної аутентифікації передбачено поєднання природного та штучного освітлення. Природне освітлення забезпечується через віконні прорізи і повинно створювати коефіцієнт природної освітленості не менше 1,5 %. Найбільш доцільною

вважається орієнтація вікон на північ або північний схід, що дозволяє уникнути прямих сонячних променів та відблисків на поверхні екранів. Штучне освітлення організовується за допомогою системи загального рівномірного освітлення з використанням люмінесцентних або світлодіодних джерел світла, які забезпечують стабільний світловий потік і не створюють пульсацій, що негативно впливають на зір.

Вимоги до освітлення регламентуються ДБН В.2.5-28-2018 «Природне і штучне освітлення», який є чинним нормативним документом. Рівні освітленості у приміщеннях з комп'ютерною технікою повинні відповідати встановленим нормам, а розрахунок коефіцієнта природного освітлення виконується згідно з методиками, наведеними у даному нормативі.

Приміщення, у яких здійснюється розробка та тестування програмно-апаратного комплексу, не повинні межувати з приміщеннями, де рівень шуму або вібрації перевищує допустимі значення. Підлогове покриття має бути матовим, з коефіцієнтом відбиття в межах 0,3–0,5, що зменшує рівень відблисків. Для внутрішнього оздоблення приміщень рекомендується застосування дифузно-відбивних матеріалів, коефіцієнти відбиття яких становлять 0,7–0,8 для стелі та 0,5–0,6 для стін. Такі вимоги сприяють створенню комфортного світлового середовища та зменшенню втоми користувачів [22].

Важливу роль відіграє дотримання вимог ергономіки при організації робочого місця. Відповідно до НПАОП 0.00-7.15-18, висота робочого столу для роботи з ПК повинна знаходитися у межах 680–800 мм, а його ширина і глибина — 600–1400 мм та 800–1000 мм відповідно. Конструкція столу повинна забезпечувати достатній простір для ніг та можливість зміни положення тіла користувача протягом робочого дня.

Стілець робочого місця повинен бути підйомно-поворотним і мати регулювання за висотою сидіння, кутом нахилу сидіння та спинки. Це дозволяє адаптувати робоче місце до індивідуальних антропометричних особливостей користувача і зменшити навантаження на хребет. Екран монітора повинен розташовуватися на відстані 600–700 мм від очей користувача, при цьому його

верхній край має бути на рівні очей або трохи нижче, а кут огляду у вертикальній площині не повинен перевищувати +30°.

Особливу увагу приділено вимогам електробезпеки. Електроживлення комп'ютерної системи, периферійних пристроїв, камер та аудіомодулів організовується за магістральною схемою. Штепсельні розетки прокладаються у спеціальних кабель-каналах або під знімною підлогою в металевих трубах чи гнучких металевих рукавах. Це дозволяє знизити ризик механічного пошкодження кабелів та підвищує рівень безпеки експлуатації обладнання.

При роботі з комп'ютерною системою біометричної аутентифікації дотримуються загальні правила техніки безпеки. Перед початком роботи здійснюється огляд обладнання та очищення екранів від пилу і забруднень. Після завершення роботи комп'ютер та периферійні пристрої відключаються від електромережі. У разі виникнення аварійної ситуації необхідно негайно знеструмити обладнання. Забороняється виконувати ремонтні, налагоджувальні або монтажні роботи безпосередньо на робочому місці без відповідного дозволу та кваліфікації.

Вимоги пожежної безпеки у приміщеннях, де здійснюється розробка та експлуатація програмно-апаратного комплексу, забезпечуються відповідно до ДБН В.1.1-7-2016 «Пожежна безпека об'єктів будівництва» та ДСТУ Б.В.1.1-36:2016, які регламентують загальні вимоги до пожежного захисту будівель і приміщень. У приміщеннях передбачено наявність первинних засобів пожежогасіння, вільний доступ до евакуаційних виходів та дотримання правил експлуатації електрообладнання.

Крім того, при експлуатації комп'ютерної техніки та периферійних пристроїв дотримано вимог ДСТУ EN 60950-1:2015 «Обладнання інформаційних технологій. Безпека. Частина 1. Загальні вимоги», який регламентує безпечну конструкцію та експлуатацію інформаційного обладнання. Даний стандарт є чинним і застосовується для оцінки відповідності обладнання вимогам електричної та пожежної безпеки.

У приміщеннях, де виконуються роботи з розробки системи, забезпечено оптимальні параметри мікроклімату, зокрема температуру, відносну вологість повітря та швидкість його руху, відповідно до чинних санітарних норм. Регулярне вологе прибирання, провітрювання приміщень та наявність аптечок першої медичної допомоги сприяють підтриманню належних умов праці та збереженню здоров’я користувачів [23].

Таким чином, дотримання чинних нормативних вимог з охорони праці, техніки безпеки, пожежної безпеки та санітарно-гігієнічних норм дозволило мінімізувати негативний вплив комп’ютерної техніки і програмно-апаратних засобів на здоров’я виконавця, забезпечити безпечні умови розробки та експлуатації методів і програмно-апаратних засобів біометричної аутентифікації користувачів комп’ютерної системи у режимі реального часу.

4.2. Забезпечення стійкості роботи вибухопожежнонебезпечного об'єкту. Засоби захисту промислово-виробничого персоналу

Під вибухопожежонебезпечним об’єктом розуміють виробниче або допоміжне приміщення, в якому використовуються, зберігаються або можуть утворюватися горючі речовини, пари, гази чи пил, здатні за певних умов призвести до виникнення пожежі або вибуху [24]. Забезпечення стійкості функціонування таких об’єктів є одним із ключових завдань системи цивільного захисту та охорони праці, оскільки аварійні ситуації на них можуть призвести до значних матеріальних збитків, загрози життю та здоров’ю персоналу, а також порушення роботи підприємства в цілому [24].

Стійкість роботи вибухопожежонебезпечного об’єкта визначається його здатністю зберігати або оперативно відновлювати працездатність у разі виникнення пожежі, вибуху або супутніх небезпечних факторів. До основних факторів, що впливають на стійкість функціонування таких об’єктів, належать конструктивні особливості будівель і споруд, надійність інженерних систем, рівень

автоматизації протипожежного захисту, організація технологічних процесів, а також рівень підготовки персоналу до дій у надзвичайних ситуаціях.

Забезпечення стійкості роботи вибухопожежонебезпечних об’єктів починається з правильного планування та проєктування будівель і приміщень. Важливу роль відіграє застосування вогнестійких будівельних матеріалів, поділ приміщень на протипожежні відсіки, наявність протипожежних перегородок і перекриттів, а також раціональне розміщення технологічного обладнання. Такі заходи дозволяють локалізувати можливі осередки пожежі або вибуху та запобігти їх поширенню на суміжні зони.

Не менш важливим є забезпечення надійності інженерних систем об’єкта. До них належать системи електропостачання, вентиляції, кондиціонування, газопостачання та водопостачання. Для підвищення стійкості роботи вибухопожежонебезпечних об’єктів передбачають резервування джерел електроживлення, використання автоматичних вимикачів і пристроїв захисного відключення, а також застосування вибухозахищеного електрообладнання. Вентиляційні системи повинні забезпечувати своєчасне видалення горючих парів і газів, що знижує ймовірність утворення вибухонебезпечних концентрацій [25].

Значну роль у забезпеченні стійкості функціонування об’єкта відіграють автоматизовані системи пожежної сигналізації та пожежогасіння. Сучасні системи дозволяють оперативно виявляти ознаки загоряння, автоматично передавати сигнал тривоги та запускати відповідні засоби пожежогасіння. Це дає змогу мінімізувати час реагування та зменшити масштаби можливих наслідків аварії. Крім того, застосування систем димовидалення сприяє зниженню концентрації токсичних продуктів горіння та покращує умови евакуації персоналу.

Важливою складовою забезпечення стійкості роботи вибухопожежонебезпечного об’єкта є організація безпечного технологічного процесу [25]. До таких заходів належать дотримання регламентів зберігання та використання горючих речовин, контроль температурних режимів, запобігання утворенню іскр та джерел відкритого вогню, а також регулярне технічне обслуговування обладнання. Систематичний контроль стану технологічного

обладнання дозволяє своєчасно виявляти потенційні несправності та запобігати аварійним ситуаціям.

Окрему увагу необхідно приділяти питанням організації управління та планування дій у разі надзвичайних ситуацій. На вибухопожежонебезпечних об’єктах розробляються плани локалізації та ліквідації аварійних ситуацій, інструкції з пожежної безпеки, а також схеми евакуації персоналу. Проведення регулярних навчань і тренувань дозволяє підвищити рівень готовності працівників до дій у разі виникнення пожежі або вибуху, що безпосередньо впливає на стійкість роботи об’єкта.

Забезпечення безпеки промислово-виробничого персоналу є невід’ємною частиною системи захисту вибухопожежонебезпечних об’єктів. Засоби захисту персоналу поділяються на колективні та індивідуальні. Колективні засоби захисту включають системи пожежної сигналізації, автоматичного пожежогасіння, вентиляції, димовидалення, аварійного освітлення та оповіщення. Їх основне призначення полягає у зменшенні впливу небезпечних факторів на групу працівників та створенні умов для безпечної евакуації [25].

Індивідуальні засоби захисту призначені для безпосереднього захисту кожного працівника. До них належать спеціальний вогнестійкий одяг, захисне взуття, каски, рукавиці, засоби захисту органів дихання та зору. Використання таких засобів є обов’язковим при виконанні робіт у зонах підвищеної пожежної небезпеки або під час ліквідації наслідків аварійних ситуацій. Особливе значення мають засоби захисту органів дихання, оскільки продукти горіння є однією з основних причин ураження персоналу під час пожеж [25].

Важливим аспектом захисту персоналу є також організація належного рівня підготовки та інформування. Працівники повинні проходити первинні, повторні та позапланові інструктажі з охорони праці та пожежної безпеки, знати порядок дій у разі виникнення пожежі або вибуху, а також вміти користуватися засобами пожежогасіння та індивідуального захисту. Регулярне навчання сприяє формуванню навичок швидкого та правильного реагування в умовах небезпеки.

Таким чином, забезпечення стійкості роботи вибухопожежонебезпечного об’єкта є комплексним завданням, що включає технічні, організаційні та навчальні заходи. Раціональне проєктування об’єкта, надійність інженерних систем, застосування сучасних засобів пожежного захисту та високий рівень підготовки промислово-виробничого персоналу дозволяють знизити ймовірність виникнення аварійних ситуацій, мінімізувати їх наслідки та забезпечити збереження життя і здоров’я працівників, а також стабільність функціонування підприємства в цілому.

4.3. Оцінка впливу електромагнітного імпульсу (EMI) ядерного вибуху на елементи біометричної аутентифікації користувачів і методи захисту від нього

Уражаючими факторами ядерного вибуху є: ударна хвиля, світлове випромінювання, проникаюча радіація, радіоактивне зараження, електромагнітний імпульс. Для забезпечення стійкості промислового об’єкту до впливу електромагнітних імпульсів ядерного вибуху необхідно провести відповідні розрахунки. При цьому вихідними даними для визначення міри впливу електромагнітних імпульсів і вжиття заходів щодо зменшення їх негативного впливу на промисловий об’єкт є наступні: промисловий цех розташовується на відстані $R = 10 км$ км від ймовірного ядерного вибуху. Очікувана потужність ядерного боєзапасу $q = 1500 кт$, вибух наземний (табл. 4.1). Елементи системи, схильні до дії EMI:

Живлення апаратних: напруга 380 В по наземних неекранованих кабелях $L_1 = 80$ м. Кабелі мають вертикальне відхилення заввишки $l_1 = 1$м. Допустимі коливання напруги мережі ±5%, коефіцієнт екранування кабелю $\eta = 2$.

Система організаційного обліку та керування підприємства складається з 35 ПК та активного і периферійного обладнання, а також з 5 систем біометричної аутентифікації. Згадані пристрої виконані на мікросхемах, що мають струмопровідні елементи заввишки $l_3 = 0{,}05$м . Робоча напруга мікросхем 5 В. Живлення від загальної мережі напруги 220 В через трансформатор. Допустимі

коливання напруги мережі ±5%. Комп'ютерна мережа управління, на якій функціонує система має горизонтальну лінію $L_2 = 500$ м і вертикальні відгалуження заввишки $l_2 = 2$м. Робоча мінімальна напруга живлення 4В. Допустимі коливання напруги мережі ±5%, коефіцієнт екранування комп'ютерної мережі $\eta = 2$.

Таблиця 4.1

Вхідні дані для оцінки дії ЕМІ на стійкість промислового об'єкту

Відстань, км.	Потужність, кт	Довжина, м		Допуск
		L1	L2	%
10	1500	80	500	5

Розрахуємо очікувані на об'єкті промислового цеху максимальні значення вертикальної (ЕВ) і горизонтальної (ЕГ) складових напруженості електричного поля [25]:

$$E_B = 5 \cdot 10^3 \cdot \frac{(1+2 \cdot R)}{R^3} \cdot \lg(14{,}5 \cdot q), \quad (4.1)$$

$$E_\Gamma = 10 \cdot \frac{(1+2 \cdot R)}{R^3} \cdot \lg(14{,}5 \cdot q), \quad (4.2)$$

де R – відстань об'єкту від вірогідного ядерного вибуху;

q – очікувана потужність ядерного боєзапасу.

$E_B = 5 \cdot 10^3 \cdot \frac{(1+2 \cdot 10)}{10^3} \cdot lg(14{,}5 \cdot 1500) = 455{,}43$ В/м.

$E_\Gamma = 10 \cdot \frac{(1+2 \cdot 10)}{10^3} \cdot \lg(14{,}5 \cdot 1500) = 0{,}91$ В/м.

Визначимо максимальну очікувану напругу наводок [25]:

а) у системі електроживлення:

$$U_B = \frac{E_B \cdot l_1}{\eta}, \quad (4.3)$$

$$U_{\Gamma} = \frac{E_{\Gamma} \cdot L_1}{\eta}, \quad (4.4)$$

де l_1 – висота вертикального відхилення кабелю до верстатів;

L_1 – довжина екранованого кабелю;

η – коефіцієнт екранування кабелю.

$U_B = \frac{455{,}43 \cdot 1}{2} = 227{,}72\text{В}$

$U_{\Gamma} = \frac{0{,}91 \cdot 80}{2} = 36{,}43\text{В}$

б) у комп'ютерній мережі промислового об'єкту:

$$U_B = \frac{E_B \cdot l_2}{\eta}, \quad (4.5)$$

$$U_{\Gamma} = \frac{E_{\Gamma} \cdot L_2}{\eta}, \quad (4.6)$$

де l_2 – висота вертикального відгалуження комп'ютерної мережі управління;

L_2 – довжина горизонтальної лінії комп'ютерної мережі управління;

η – коефіцієнт екранування кабелю.

$U_B = \frac{455{,}43 \cdot 2}{2} = 455{,}43\text{В}$

$U_{\Gamma} = \frac{0{,}91 \cdot 500}{2} = 227{,}72\text{В}$

в) в комп'ютерах та периферійних пристроях:

$$U_B = \frac{E_B \cdot l_3}{\eta}, \quad (4.7)$$

де $l_3\, l_з$ – висота струмопровідних елементів;

η – коефіцієнт екранування кабелю.

$U_B = \frac{455{,}43 \cdot 0{,}05}{2} = 11{,}39\text{В}$

Визначимо допустиму максимальну напругу наводок [25]:

а) у мережі живлення промислового об’єкту:

$$U_{Д} = U + U(5\%), \tag{4.8}$$

де U – напруга живлення щитових;

$U_{Д_1} = 380 + \frac{380 \cdot 5}{100} = 399В$

б) у комп’ютерній мережі:

$U_{Д_2} = 4 + \frac{4 \cdot 5}{100} = 4{,}2В$

в) у комп’ютерах та периферійних пристроях:

$U_{Д_1} = 5 + \frac{5 \cdot 5}{100} = 5{,}25В$

Розрахуємо коефіцієнт безпеки [16]:

$$K = 20 \cdot \lg\left(\frac{U_{Д}}{U_{е}}\right), \tag{4.9}$$

де $U_{Д}$ – допустима максимальна напруга наведень в пристрої введення, ЕОМ, блоці управління

$U_{е}$ – очікувана максимальна напруга наведень в комп’ютерах та периферійних пристроях.

$K = 20 \cdot \lg\left(\frac{5{,}25}{11{,}39}\right) = -6{,}72дБ$

Зведемо отримані дані в таблицю (табл. 4.2).

Таблиця 4.2

Результати оцінки стійкості промислового об’єкту до дії ЕМІ

Елементи системи	Допустима напруга мережі, В	Напруженість електричного поля, В/м		Наводки напруги в струмопровідних елементах, В	
		$E_В$	$E_Г$	$U_В$	$U_Г$
Електропостачання апаратної	399	455,43	0,91	227,72	36,43
Комп'ютери, периферійні пристрої	5,25	455,43	0,91	11,39	—
Комп'ютерна мережа, активне обладнання	4,2	455,43	0,91	455,43	227,72
Коефіцієнт безпеки К =-6,72дБ << 40 дБ.					

Локальна комп’ютерна мережа та обладнання може опинитися в зоні дії ЕМІ наземного ядерного вибуху. Нестійкими елементами об'єкту є: комп’ютери та периферійне обладнання, а також локальна комп’ютерна мережа.

Об'єкт не стійкий до дії ЕМІ, оскільки коефіцієнт безпеки значно менше задовільного значення, що становить до 40 дБ. Для підвищення стійкості роботи об'єкту до дії ЕМІ ядерного вибуху необхідно провести наступні заходи:

- кабель живлення щитових екранувати;
- комп’ютери і периферійні пристрої заземлити;
- для комп’ютерів встановити відключаючі пристрої.

Такі заходи дозволяють забезпечити працівників і службовців від уражаючої дії ядерного вибуху та радіаційного забруднення при збереженні працездатності інформаційної інфраструктури підприємства.

ВИСНОВКИ

Основні наукові та практичні результати полягають в наступному.

1) Проведено аналіз основних понять та класифікацій біометричної аутентифікації, включаючи фізіологічні та поведінкові характеристики, а також встановлено відмінності між простими та мультимодальними системами, що дало змогу обґрунтувати вибір комплексного підходу до аутентифікації особи, який поєднує кілька незалежних ознак для підвищення надійності та точності системи.

2) Проаналізовано сучасні моделі та алгоритми розпізнавання обличчя, починаючи з класичних статистичних методів до глибинних нейронних мереж, що дало можливість визначити переваги використання глибинного навчання, зокрема високу стійкість до варіацій освітлення, пози та міміки, а також сформувати вимоги до обчислювальних ресурсів для подальшого проектування системи.

3) Досліджено сучасні методи аутентифікації голосу, включаючи MFCC-ознаки, статистичні моделі (GMM, i-vector), та передові нейромережеві архітектури, що дало змогу визначити ефективні підходи до добування голосових ознак та порівняння їх за точністю, стійкістю до шумів і вимогами до обчислень, що є критично важливим для створення мультимодальної системи на вбудованих пристроях.

4) Визначено структурні компоненти апаратної та програмної частини комплексу біометричної мультимодальної аутентифікації користувачів комп’ютерної системи, а також встановлено вимоги до їх сумісності, що дало змогу побудувати узагальнену архітектуру системи та забезпечити обробку зображення обличчя і голосового сигналу в режимі реального часу.

5) Обґрунтовано вибір алгоритмів і моделей для детекції та розпізнавання обличчя, зокрема застосування MTCNN для виявлення облич та ключових точок і MobileFaceNet з функцією втрат ArcFace для формування високоякісних векторів-ембедингів. Такий вибір забезпечує високу точність і можливість реалізації на Raspberry Pi, що є критичним для апаратно-програмного комплексу.

6) Обґрунтовано вибір моделей для розпізнавання голосу, де основними стали ECAPA-TDNN як високоточна модель та x-vector як ресурсоефективна альтернатива. Використання цих моделей дає змогу отримати стійкі до шуму та варіативності мовлення голосові ембединг-вектори і реалізувати розпізнавання у реальному часі на вбудованих системах

7) Визначено та формалізовано критерії оцінювання роботи системи, включно з метриками FAR, FRR, EER та вимогами до максимально допустимої затримки обробки.

8) Розроблено архітектуру апаратного забезпечення апаратно-програмного комплексу мультимодальної біометричної аутентифікації на базі одноплатного комп’ютера Raspberry Pi 4 Model B із використанням модуля камери, USB-мікрофона та дисплейного інтерфейсу, що дало змогу забезпечити збалансоване поєднання продуктивності, енергоефективності та компактності системи при роботі в режимі реального часу.

9) Реалізовано системне програмне забезпечення апаратних компонентів, включаючи налаштування операційної системи Raspberry Pi OS, конфігурацію камерного та аудіомодулів, а також мережеві параметри взаємодії, що дало змогу організувати стабільне середовище збору мультимодальних біометричних даних та їх передачі до програмних модулів обробки.

10) Спроєктовано архітектуру програмного забезпечення системи з використанням клієнт-серверного підходу, що включає бекенд-логіку, веб-орієнтовану адміністративну панель та локальну базу даних SQLite, що дало змогу забезпечити централізоване управління користувачами, збереження біометричних шаблонів та журналювання результатів аутентифікації.

11) Реалізовано нейронні мережі для розпізнавання обличчя та голосу, зокрема MTCNN для детекції та вирівнювання облич, MobileFaceNet для формування векторів ознак обличчя та ECAPA-TDNN для отримання векторів ознак голосу, що дало змогу досягти високої точності аутентифікації при обмежених апаратних ресурсах одноплатного комп’ютера.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Паламар М.І., Стрембіцький М.О., Паламар А.М. Проектування комп’ютеризованих вимірювальних систем і комплексів. Навчальний посібник. Тернопіль: ТНТУ. 2019. 150 с.

2. Нечипоренко О. В., Корпань Я. В. Біометрична ідентифікація і автентифікація особи за геометрією обличчя. Вісник Хмельницького національного університету. 2016. №4(236). С. 26–31.

3. Луцків А. М., Калинюк А. М. Інтеграція підсистем біометричної аутентифікації у веб-сервісів. Матеріали Міжнародної наук.-техн. конф. «Актуальні задачі сучасних технологій». Том 2. Тернопіль: ТНТУ, 2017. С. 22.

4. Заєць І. С., Бридінський В. А., Сабодашко Д. В., Хома Ю. В., Руда Х. С., Швед М. Є. Використання ембедингів голосу в інтегрованих системах для діаризації та виявлення зловмисників. Вісник НУЛП «Комп’ютерні системи та мережі». 2024. Т. 6, №1. С. 54-66.

5. Schroff F., Kalenichenko D., Philbin J. FaceNet: A Unified Embedding for Face Recognition and Clustering. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. PP. 815–823.

6. Desplanques B., Thienpondt J., Demuynck K. ECAPA-TDNN: Emphasized Channel Attention, Propagation and Aggregation in TDNN Based Speaker Verification. Interspeech 2020. PP. 3830–3834.

7. Loweimi E., Brown G., Huckvale M. A Comparative Study of x-vector and ECAPA-TDNN Models for Speaker Verification in the Wild. Interspeech 2024. PP. 1025–1032.

8. Snyder D., Garcia-Romero D., Sell G., Povey D., Khudanpur S. X-vectors: Robust DNN Embeddings for Speaker Recognition. ICASSP 2018. PP. 5329–5333.

9. Wu Z., Nagrani A., Chung J.S., Zisserman A. VoxCeleb: Large-Scale Speaker Recognition Dataset and Benchmark. Computer Speech & Language. 2021. Vol. 72. PP. 101255.

10. Wang H., Zheng S., Chen Y., Cheng L., Chen Q. CAM++: A Fast and Efficient Network for Speaker Verification Using Context-Aware Masking. ArXiv preprint arXiv:2303.00332, 2023.

11. Zhu W., Kong T., Lu S. SpeechNAS: Towards Better Trade-off between Latency and Accuracy for Large-Scale Speaker Verification. ArXiv preprint arXiv:2109.08839, 2021.

12. Zhang J., Wang Y., Yu Z. Speaker Recognition Based on Deep Learning: An Overview. ArXiv preprint arXiv:2012.00931, 2020.

13. Луцик Н. С., Луцків А. М., Осухівська Г. М., Тиш Є. В. Методичні рекомендації до виконання кваліфікаційної роботи магістра для студентів спеціальності 123 «Комп'ютерна інженерія» другого (магістерського) рівня вищої освіти усіх форм навчання. Тернопіль. ТНТУ. 2024. 44 с.

14. Луцків А.М., Гаврада Д.М. Архітектура апаратно-програмного комплексу мультимодальної біометричної аутентифікації особи. Матеріали XIV міжнародної науково - технічної конференції молодих учених і студентів «Актуальні задачі сучасних технологій» (11-12 грудня 2025 р.) Тернопільського національного технічного університету імені Івана Пулюя. Тернопіль: ТНТУ. 2025. С. 293.

15. Луцків А.М., Гаврада Д.М. Аналіз та класифікація біометричних систем за типами ознак. Матеріали XIII науково-технічної конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі, системи та технології» (17-18 грудня 2025 року). Тернопіль: ТНТУ. 2025. С. 128.

16. Schmidhuber J. Deep learning in neural networks: An overview. Neural Networks. Vol. 61. N. 1. January 2015. pp. 85–117.

17. Fischer and C. Igel. Training Restricted Boltzmann Machines: An Introduction. Pattern Recognition. Vol. 47, N. 1. January 2014. pp. 25–39,

18. Golovko, V.A. Deep learning: an overview and main paradigms. Optical Memory and Neural Networks (Information Optics). Vol. 26, № 1. 2017. pp. 1–17.

19. Liu, J.N., Hu, Y., You, J.J., Chan, P.W.: Deep neural network based feature representation for weather forecasting. In: International Conference on Artificial Intelligence (ICAI). 2014. p. 10.

20. Dalto, M.: Deep neural networks for time series prediction with applications in ultra-short-term wind forecasting. In: IEEE ICIT.2015. pp. 257-268.

21. НПАОП 0.00-7.15-18 «Вимоги щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями». Київ. 2018.

22. ДБН В.2.5-28-2018 «Природне і штучне освітлення». Київ : Мінрегіон України. 2018.

23. ДБН В.1.1-7-2016 «Пожежна безпека об'єктів будівництва». Київ : Мінрегіон України. 2016.

24. Бедрій Я. Основи охорони праці користувачів персональних комп'ютерів: навчальний посібник для студентів ВНЗ та інженерів-практиків. Навчальна книга-Богдан. 2014. 144 с.

25. Стручок В.С. Методичний посібник для здобувачів освітнього ступеня «магістр» всіх спеціальностей денної та заочної (дистанційної) форм навчання «Безпека в надзвичайних ситуаціях». Тернопіль: ФОП Паляниця В. А.156 с.

26. Стручок В.С. Навчальний посібник «ТЕХНОЕКОЛОГІЯ ТА ЦИВІЛЬНА БЕЗПЕКА. ЧАСТИНА «ЦИВІЛЬНА БЕЗПЕКА»». Тернопіль: ФОП Паляниця В. А., – 156 с.

27. Луцик Н.С., Луцків А.М., Осухівська Г.М., Тиш Є.В. Програма та методичні рекомендації з проходження практики за тематикою кваліфікаційної роботи для студентів спеціальності 123 «Комп'ютерна інженерія» другого (магістерського) рівня вищої освіти усіх форм навчання. Тернопіль: ТНТУ. 2024. 45 с.

28. Варавін А.В., Лещишин Ю.З., Чайковський А.В. Методичні вказівки до виконання курсового проєкту з дисципліни «Дослідження і проєктування комп'ютерних систем та мереж» для здобувачів другого (магістерського) рівня вищої освіти спеціальності 123 «Комп'ютерна інженерія» усіх форм навчання. Тернопіль: ТНТУ, 2024. 32 с.

Додаток А

Тези конференцій

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Тернопільський національний технічний університет імені Івана Пулюя
Маріборский університет (Словенія)
Технічний університет у Кошице (Словаччина)
Вільнюський технічний університет ім. Гедимінаса (Литва)
Краківський економічний університет (Польща)
Вроцлавський економічний університет (Польща)
Університет «Опольська Політехніка» (Польща)
Національний університет «Полтавська політехніка імені Юрія Кондратюка»
Вінницький національний аграрний університет
Львівський національний університет ім. І. Франка
Головне управління Пенсійного фонду в Тернопільській області
Наукове товариство ім. Шевченка
Тернопільський обласний комунальний інститут післядипломної педагогічної освіти
Сумський державний педагогічний університет
Запорізький національний університет

АКТУАЛЬНІ ЗАДАЧІ СУЧАСНИХ ТЕХНОЛОГІЙ

Збірник

тез доповідей

XIV Міжнародної науково-технічної конференції молодих учених та студентів

11-12 грудня 2025 року

УКРАЇНА
ТЕРНОПІЛЬ – 2025

УДК 004.93
А.М. Луцків канд. техн. наук, доцент, Д.М. Гаврада
Тернопільський національний технічний університет імені Івана Пулюя

АРХІТЕКТУРА АПАРАТНО-ПРОГРАМНОГО КОМПЛЕКСУ МУЛЬТИМОДАЛЬНОЇ БІОМЕТРИЧНОЇ ІДЕНТИФІКАЦІЇ ОСОБИ

A.M. Lutskiv PhD., Assoc. Prof., D.M. Havrada

ARCHITECTURE OF A HARDWARE-SOFTWARE COMPLEX FOR MULTIMODAL BIOMETRIC PERSON IDENTIFICATION

Біометрична ідентифікація – це процес розпізнавання особи на основі її унікальних фізіологічних або поведінкових характеристик. Вона є одним із найперспективніших напрямів розвитку систем інформаційної безпеки комп'ютеризованих систем, оскільки дозволяє виконувати аутентифікацію без необхідності запам'ятовування паролів чи носіння фізичних носіїв.

В основі біометричної ідентифікації лежить принцип, за яким біометричний шаблон користувача формується з набору вимірюваних ознак, отриманих під час реєстрації, і згодом порівнюється з поточними вимірюваннями.

У роботі пропонується архітектура системи, яка будується за модульним принципом і охоплює як апаратні, так і програмні компоненти. Ці елементи взаємодіють між собою через локальні інтерфейси та внутрішні програмні API. Основною ідеєю є об'єднання двох незалежних каналів біометричної ідентифікації за обличчям і за голосом в єдину інтегровану платформу, яка забезпечує надійне підтвердження особи у режимі реального часу. Запропонована архітектура апаратно-програмного комплексу мультимодальної біометричної ідентифікації особи представлена на рис.1.

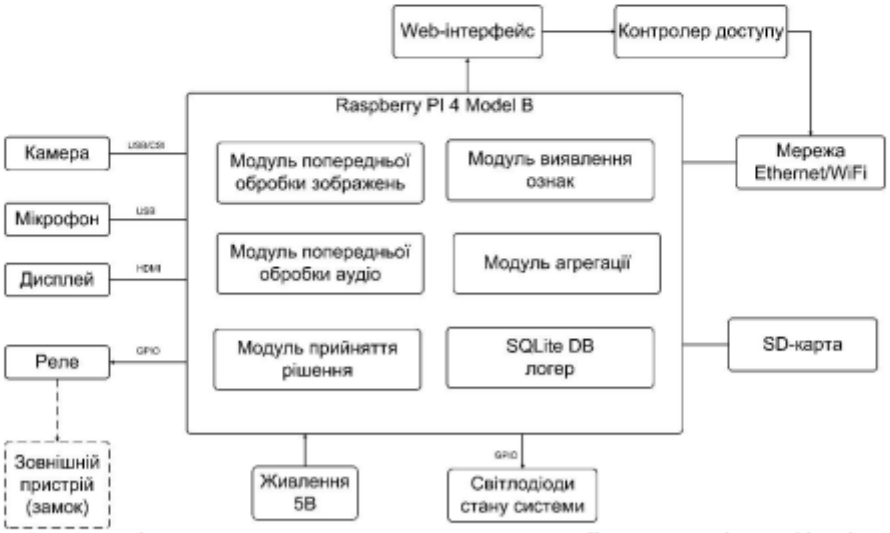


Рис.1. Архітектура апаратно-програмного комплексу біометричної ідентифікації

Як видно з рис.1, центральним елементом комплексу виступає одноплатний комп'ютер Raspberry Pi 4 Model B, який координує роботу всіх підсистем.

Базовий вузол системи розміщено на Raspberry Pi, що поєднує у собі достатню обчислювальну потужність, мале енергоспоживання та можливість роботи у бездисковому автономному режимі. До нього підключено USB-камеру, яка здійснює захоплення відео з обличчям користувача, та мікрофон, який формує звуковий потік для аналізу мовлення.

Для зворотного зв'язку може використовуватися невеликий дисплей або сенсорний екран, де відображаються результати ідентифікації та стан системи.

Через GPIO-порти Raspberry Pi керує реле або електронним замком, що забезпечує фізичну реакцію системи, яка включає:

– розблокування доступу;
– світлового сигналу за допомогою світлодіодів.

У конструкції передбачено модуль живлення з резервним акумулятором (UPS HAT), який гарантує коректне завершення роботи навіть при раптовому відключенні електромережі. Це запобігає пошкодженню даних у локальній базі. Усі сигнали з керуючих портів проходять через оптоелектронні розв'язки, що підвищує електричну безпеку пристрою.

Програмне забезпечення системи складається з двох основних частин: backend-сервісу та веб-інтерфейсу користувача. Обидві частини функціонують на Raspberry Pi та не потребують підключення до зовнішнього сервера.

Backend реалізовано на Python із використанням фреймворку FastAPI. Цей компонент відповідає за взаємодію з периферійними пристроями, обробку даних, зберігання інформації та логіку прийняття рішень.

Для зберігання даних використовується локальна база SQLite, що не потребує додаткових служб і добре підходить для вбудованих систем. У базі даних передбачається створення таблиць користувачів, векторних описів облич та голосу, журналів подій і системних параметрів.

Модуль відеоідентифікації виконує захоплення кадрів з камери, корекцію зображення, нормалізацію яскравості та виявлення області обличчя. Далі активується нейромережева модель, навчена для екстракції ознак обличчя, яка формує компактне числове представлення користувача.

Аналогічно, модуль голосової ідентифікації аналізує аудіосигнал, фільтрує шум, виділяє мовні ділянки та обчислює спектральні характеристики. У результаті отримується вектор голосових ознак, який у подальшому порівнюється з тими зразками, які є в базі даних.

Одержані вектори ознак надходять до модуля агрегації, який поєднує інформацію від двох каналів і формує єдину оцінку достовірності. На основі цього результату система ухвалює рішення щодо автентичності користувача. У випадку позитивного результату формується команда на реле керування замком та запис у журнал подій.

Веб-інтерфейс адміністратора можна релізувати засобами HTML 5 та JavaScript із використанням шаблонів FastAPI. Через цей інтерфейс адміністратор може реєструвати нових користувачів, здійснювати первинне навчання системи, переглядати журнали, змінювати параметри роботи, а також тестувати камеру та мікрофон. Веб-інтерфейс працює у межах локальної мережі або через захищене з'єднання HTTPS, що дозволяє виконувати адміністрування без підключення до хмарних сервісів. Взаємодію між компонентами системи покладено на внутрішній API, що базується на обміні HTTP-запитами.

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ІВАНА ПУЛЮЯ

М А Т Е Р І А Л И

XIII НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

«ІНФОРМАЦІЙНІ МОДЕЛІ, СИСТЕМИ ТА ТЕХНОЛОГІЇ»

17-18 грудня 2025 року

ТЕРНОПІЛЬ
2025

УДК 004.93
А. Луцків к. т. н., доц.; Д. Гаврада
(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

АНАЛІЗ ТА КЛАСИФІКАЦІЯ БІОМЕТРИЧНИХ СИСТЕМ ЗА ТИПАМИ ОЗНАК

UDC 004.93
A. Lutskiv PhD., Assoc. Prof.; D. Havrada

ANALYSIS AND CLASSIFICATION OF BIOMETRIC SYSTEMS BY FEATURE TYPES

Розвиток мікропроцесорних систем, сенсорних технологій і методів штучного інтелекту зробив можливим використання біометрії не лише у великих корпоративних структурах, а й у побутових пристроях, зокрема, смартфонах, банкоматах, системах «розумний дім» тощо.

Методи біометричної ідентифікації поділяють за двома основними критеріями:

– за природою ознак – фізіологічні та поведінкові;

– за способом аналізу – одноканальні (монобіометричні) та багатоканальні (мультимодальні).

До фізіологічних характеристик при біометричній ідентифікації людини належать характеристики, які пов'язані з її анатомічними властивостями. До них входять: геометрія обличчя, відбитки пальців, сітківка, райдужна оболонка ока, форма вуха, структура вен.

Поведінкові характеристики відображають динаміку дій, які можуть включати почерк, ходу, голос, ритм натискання клавіш тощо.

На рис. 1 показано деталізовану класифікацію біометричних методів ідентифікації та відповідних систем за типами ознаками людини.

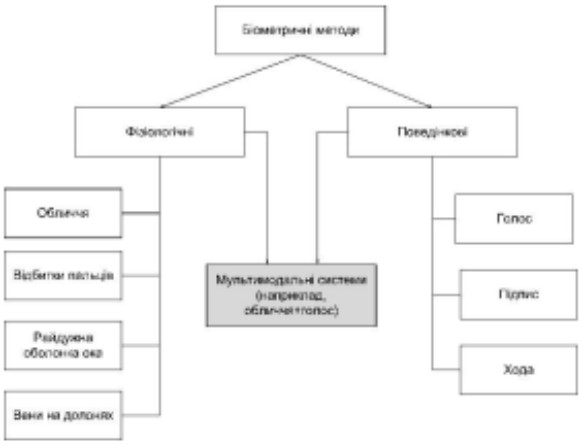


Рисунок 1. Класифікація біометричних систем за типами ознак

На сьогодні розроблено досить багато різних методів, які дозволяють проводити біометричну ідентифікацію особи на основі її фізіологічних властивостей. Однак найбільш перспективним з них є мультимодальний підхід, що поєднує як фізіологічні, так і поведінкові особливості людини.