Міністерство освіти і науки України

Харківський національний університет імені В.Н. Каразіна

Факультет комп'ютерних наук

Спеціальність 125 «Кібербезпека»

Освітня програма «Безпека інформаційних та комунікаційних систем»

«Допущено до захисту»

В.о. зав. кафедрою БІСТ

Ольга МЕЛКОЗЬОРОВА

_____

« » 2023 р.

Пояснювальна записка

до кваліфікаційної роботи магістра

на тему: «Статистичні показники якості методів біометричної аутентифікації»

оцінка « »

Голова ЕК

Олександр ЛЕМЕШКО _____

Керівник к.т.н. Мелкозьорова О. М.

Рецензент проф., д.т.н. Краснобаєв В. А.

Виконавець : студент групи КБ-61

Мальченко М. С.

Харків – 2023

РЕФЕРАТ

Звіт про виконання дипломної роботи: 62 сторінки, 14 рисунків, 3 таблиці, 1 додаток та 17 джерел.

Метою роботи є дослідження конкретних біометричних технологій, таких як сканування райдужки, відбитків пальців, розпізнавання обличчя та голосу, та аналіз методів їх застосування й інтеграції в системи безпеки і аутентифікації, а також визначення актуальності і ефективності різних методів біометричної ідентифікації в умовах інтернет загроз, які стрімко розвиваються .

У проекті було розглянуто й проаналізовано загальні положення щодо видів біометричної аутентифікації, їх методів, зокрема біометричної аутентифікації за райдужкою ока та методів SOFM та BPNN. На основі проведеного дослідження та аналізу предметної області, було виявлено ключові виклики та можливості для покращення цих технологій. Тестування показало, що інтеграція даних методів біометричної аутентифікації й правильний вибір в контексті їх сфери застосування може значно підвищити рівень безпеки даних та ефективності систем, при цьому мінімізуючи потенційні ризики.

Результати проекту можуть бути використані в якості оглядової роботи для розробників біометричних систем безпеки або поціновувачів новітніх технологій. Надаються цінні й актуальні вказівки щодо вибору та інтеграції біометричних технологій, зокрема, з реалізацією можливості ідентифікації особи за райдужкою ока.

Серед можливих напрямків розвитку роботи можна виділити дослідження новітніх систем забезпечення конфіденційності даних на базі нейронних мереж, зокрема, торкнутися оцінки точності й надійності, аналізу швидкості обробки, використання більшої кількості датасетів. Розробка нових алгоритмів машинного навчання здатне підвищити точність роботи біометричних систем та загальну безпеку даних користувачів.

Ключові слова: FRR, SOFM, РОЗПІЗНАВАННЯ РАЙДУЖКИ, FAR, PYTORCH, BPNN, EER, LBP, ROC, ARTIFICIAL NEURAL NETWORKS, МЕТОДИ БІОМЕТРИЧНОЇ АУТЕНТИФІКАЦІЇ.

ABSTRACT

The thesis report: 62 pages, 14 figures, 3 tables, 1 appendices and 17 sources.

The purpose of the work is the study of specific biometric technologies, such as iris scanning, fingerprints, face and voice recognition, and the analysis of methods of their integration into security and authentication systems, as well as determining the relevance and effectiveness of various biometric identification methods in the conditions of rapidly developing Internet threats.

The purpose of this work is to study specific biometric technologies such as iris scanning, fingerprinting, facial and voice recognition, and to analyze the methods of their application and integration into security and authentication systems, as well as to determine the relevance and effectiveness of different biometric identification methods in the context of rapidly evolving internet threats.

The project considered and analyzed general provisions regarding types of biometric authentication, their methods, particularly iris biometric authentication, and SOFM and BPNN methods. Based on the conducted research and analysis of the subject area, key challenges and opportunities for improving these technologies were identified. Testing showed that the integration of these biometric authentication methods and the right choice in the context of their application can significantly increase the level of data security and system effectiveness, while minimizing potential risks.

The results of the project can be used as a review work for developers of biometric security systems or connoisseurs of cutting-edge technologies. Valuable and current guidelines are provided regarding the choice and integration of biometric technologies, in particular, with the implementation of the possibility of identifying a person by the iris of the eye.

Among the possible directions for the development of the work, one can identify the study of the latest systems for ensuring data confidentiality based on neural

networks, in particular, touching upon the assessment of accuracy and reliability, analysis of processing speed, and the use of larger datasets. The development of new machine learning algorithms can increase the accuracy of biometric systems and the overall data security of users.

Keywords: FRR, SOFM, IRIS RECOGNITION, FAR, PYTORCH, BPNN, EER, LBP, ROC, ARTIFICIAL NEURAL NETWORKS, BIOMETRIC AUTHENTICATION METHODS.

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ, ТЕРМІНІВ

BPNN – Back Propagation Neural Networks
CNN – Convolutional Neural Networks
DNN – Deep Neural Networks
DTW – Dynamic Time Warping
EER – Equal Error Rate
FAR – False Acceptance Rate
FRR – False Rejection Rate
GMM – Gaussian Mixture Models
HMM – Hidden Markov Models
LBP – Local Binary Patterns
MTCNN – Multi-task Cascaded Convolutional Networks
RNN – Recurrent Neural Networks
ROC – Receiver Operating Characteristic
SOFM – Self-Organizing Feature Maps
БД – База даних
ПЗ – Програмне забезпечення
ШНМ – Штучні Нейронні Мережі

ВСТУП

В епоху постійного розвитку цифрових інновацій, зростає важливість засобів для точної та безпечної верифікації особи. У цьому контексті, біометричні системи аутентифікації набувають все більшої популярності та значимості. Завдяки глобальному прогресу в цифровій трансформації, біометрична аутентифікація стає ключовим засобом для забезпечення надійного захисту особистих даних користувачів. Така необхідність є реакцією на зростаючі виклики, пов'язані з ризиками крадіжки особистих даних, несанкціонованого доступу та шахрайства. Біометричні методи, що працюють на базі розпізнавання відбитків пальців, розпізнавання обличчя та сканування райдужки ока є ефективними інструментами для протидії таким поширеним загрозам безпеки.

Об’єктом дослідження є процес впровадження і використання біометричних методів аутентифікації, їх застосування в різних галузях і вплив на системи безпеки.

Предметом дослідження є конкретні біометричні технології, такі як сканування райдужки, відбитків пальців, розпізнавання обличчя та голосу, а також методи їх інтеграції в системи безпеки й аутентифікації.

Метою роботи є дослідження та аналіз, а також визначення актуальності і ефективності різних методів біометричної аутентифікації.

Головними задачами при виконанні кваліфікаційної роботи стали наступні:

- дослідження основних понять біометричної аутентифікації, а саме: аутентифікації, біометрії, сфери застосування біометричної аутентифікації;
- огляд сучасних видів біометричного розпізнавання, особливостей їх застосування та методів біометричної аутентифікації по роботі з ними;

- розглянути архітектури, механізми навчання та використання методів біометричної аутентифікації з використанням райдужки ока;
- провести дослідження статистичних показників ефективності та дослідити їх вплив на методи біометричної аутентифікації;
- реалізувати методи з використанням технологій програмування та готових баз даних з райдужними оболонками;
- виробити висновки й надати практичні рекомендації, щодо застосування досліджених методів в залежності від цілей та умов їх використання.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Основні поняття біометричної аутентифікації

1.1.1 Аутентифікація

Аутентифікація представляє собою метод перевірки особистості людини або ідентифікаційних даних у рамках сучасних інформаційних комп'ютерних систем. Цей процес забезпечує впевненість в тому, що особа або система, яка прагне доступу до певних ресурсів, дійсно є тією, за кого себе видаватиме. Для аутентифікації зазвичай потрібно ввести логін і пароль, але можливі й інші форми, такі як біометричні сканери (наприклад, сканування відбитків пальців або розпізнавання обличчя), використання разових паролів або електронних ключів.

Цей процес має вирішальне значення для забезпечення безпеки даних та систем, допомагаючи уникнути неавторизованого доступу. Зазвичай аутентифікація стає початковим етапом у процедурі контролю доступу, на яку наступає етап авторизації, визначаючи, до яких ресурсів та дій має доступ ідентифікований користувач.

Аутентифікація відіграє ключову роль у забезпеченні безпеки інформаційних систем та мереж. Вона забезпечує встановлення особистості користувача, пристрою чи програми. За факторами, типи аутентифікації можна розділити наступні [1]:

1) Щось, що знаєш - заснована на застосуванні логіну чи паролю й також може включати PIN-коди чи відповіді на заздалегідь налаштовані секретні питання до користувача;
2) Щось, що маєш - володіння такими засобами, як, наприклад, налаштовані фізичні USB ключі дозволяють авторизувати особу користувача;

3) Щось, що ти є - застосування унікальних для кожної людини біометричних даних.

На рисунку 1.1 наглядно зображено типи аутентифікації за факторами.

Рисунок 1.1 – Типи аутентифікації за факторами

Також, серед деяких сучасних ефективних методів аутентифікації виділяють наступні [2]:

1) Біометрична аутентифікація – залучає унікальні фізичні атрибути людини для ідентифікації;
2) Поведінкова аутентифікація - оцінює регулярні дії користувача, як-от стиль введення тексту на клавіатурі чи спосіб використання комп'ютерної миші;
3) Багатофакторна аутентифікація - поєднує декілька з вище розглянутих типів аутентифікації за факторами, задля підвищення загального рівня безпеки системи;

1.1.2 Біометрія

Біометрія являє собою область технологій, котра використовує унікальні характеристики, що пов'язані з фізичною структурою або поведінкою людини, для ідентифікації чи підтвердження особи. Ключові особливості та принципи біометрії охоплюють наступні риси:

1) Індивідуальність - кожна особа має свої неповторні біометричні дані, такі як відбитки пальців, структуру райдужки ока, або вигляд обличчя.
2) Стабільність у часі - біометричні характеристики зазвичай залишаються незмінними протягом тривалого часу. Хоча деякі елементи, як-от вага або зовнішній вигляд, можуть варіюватися, головні біометричні особливості зазвичай не змінюються.
3) Можливість квантифікації - біометричні характеристики можуть бути кількісно оцінені та перетворені в цифровий формат для аналізу та архівації.
4) Співставлення даних - зібрані біометричні інформації можуть бути порівняні з існуючими базами даних для встановлення або підтвердження особи.
5) Застосування у сфері безпеки - біометричні системи часто використовуються для підвищення безпеки, наприклад, у системах контролю доступу чи при аутентифікації у фінансових установах.
6) Легкість використання - сучасні біометричні системи розроблені для зручності та інтуїтивного використання кінцевими користувачами.

1.1.3 Біометрична аутентифікація

Біометрична аутентифікація представляє собою процедуру технологічної ідентифікації особистості, засновану на унікальних атрибутах фізичної або поведінкової природи особи. Такий метод полягає в процедурі верифікації ідентичності особи за допомогою порівняння актуально зібраних біометричних відомостей з даними, які вже збережені в системі.

Центральним елементом біометричної аутентифікації є використання властивих лише конкретній людині біологічних атрибутів для встановлення та підтвердження її ідентичності. Такі атрибути можуть бути як фізичного характеру, наприклад, відбитки пальців, структура обличчя, візерунки сітківки ока, так і поведінкового, як-от особливості голосу або специфіка друку тексту. Застосування цих особливих біологічних параметрів дозволяє створювати міцні та безпечні системи для перевірки автентичності, що складно підробити чи обійти.

Біометричні характеристики охоплюють унікальні атрибути, пов'язані з фізичною або поведінковою ідентичністю особи, використовувані для її розпізнавання або підтвердження [3].

Біометричні характеристики можна побачити на рисунку 1.2.

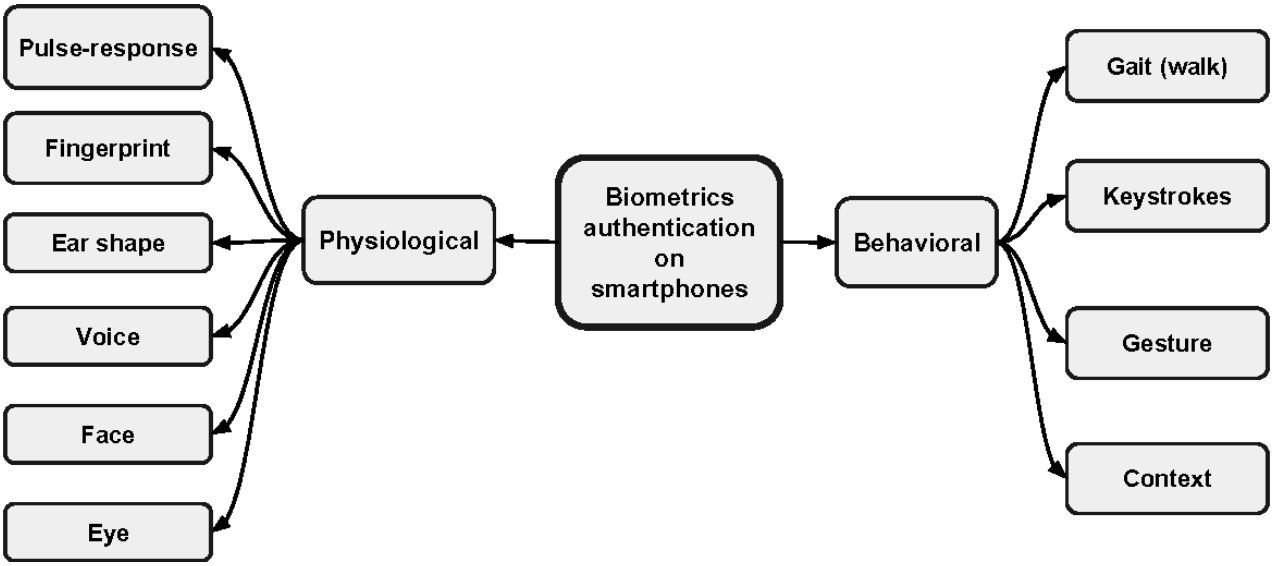

Рисунок 1.2 – Фізичні та поведінкові біометричні атрибути людини

На противагу звичайним способам аутентифікації, таким як використання паролів чи PIN-кодів, біометрична аутентифікація пропонує більш високий рівень захисту, бо базується безпосередньо на фізичних властивостях людини, котрі є унікальними у кожної людини і в той же час їх значно складніше підробити чи скопіювати.

Ключовим елементом біометричної автентифікації є забезпечення захисту та конфіденційності біометричних даних, що включає захист від несанкціонованого доступу, шифрування, спроб фальсифікації, використання штучних атрибутів. Біометричні системи ставлять перед собою високі вимоги до точності, швидкості обробки даних та захисту персональних відомостей користувача, але необхідно відмітити певні труднощі, котрі пов’язані з конфіденційністю й ризиком помилкового визначення особи.

Процес біометричної аутентифікації включає в себе наступні етапи роботи:

1) Відбір біометричних даних - першочерговий крок у процесі біометричної аутентифікації включає збір біометричних відомостей індивіда, що може бути здійснений через процедури сканування біометричних даних людини.
2) Процесування та архівування біометричних даних - отримані біометричні відомості подаються на обробку за допомогою спеціалізованого програмного забезпечення, яке конвертує їх у цифровий формат. З міркувань конфіденційності, ці дані зазвичай зберігаються у вигляді математичного коду, а не як фотографії чи аудіозаписи. Отримані біометричні дані піддаються аналізу для виділення унікальних особливостей, які служать для ідентифікації індивіда.
3) Перевірка з базою даних - визначені характеристики порівнюються з інформацією в базі даних. Якщо система виявляє співпадіння між представленими даними та записами в базі, ідентичність користувача вважається підтвердженою.
4) Рішення про автентифікацію - за результатами порівняння система приймає рішення щодо автентифікації користувача. У разі збігу даних доступ дозволяється; у протилежному випадку - відмовляється.

Серед технічних аспектів біометричної аутентифікації можна виділити наступні:

- Точність аутентифікації залежить від якості використовуваних сенсорів та ефективності алгоритмів обробки, що є критичними факторами для забезпечення надійності системи.
- Захист зібраних біометричних даних є надзвичайно важливим, адже ці дані є чутливими та потребують надійного захисту від несанкціонованого доступу чи витоку.
- Швидкість обробки даних є ключовою для забезпечення оперативності біометричної аутентифікації, оскільки це впливає на загальну ефективність та зручність використання системи.
- Сумісність системи біометричної аутентифікації з існуючими системами та пристроями є важливою для її інтеграції та широкого впровадження, що дозволяє розширити можливості її використання.

Серед сильних сторін біометричної аутентифікації можна виділити наступні:

- Індивідуальність - кожен біометричний параметр унікальний для кожної особи, забезпечуючи високу точність у визначенні та верифікації особи.
- Зручність використання - така форма аутентифікації часто є швидкою та простою, зокрема, методи, як-от ідентифікація по обличчю або сканування відбитків пальців, можуть бути швидшими, ніж введення паролю.
- Складність обходу - фізичні характеристики важче підробити або обійти, на відміну від традиційних паролів.
- Автоматизація процесів - використання біометричних технологій дозволяє автоматизувати процедури ідентифікації, економлячи час та ресурси.

- Зниження ризику крадіжки особистих даних - біометричні дані складно вкрасти або підробити, що зменшує ризик крадіжки особистих даних.

Мінуси систем на базі біометричної аутентифікації виглядають наступним чином:

- Проблеми з приватністю - збір та зберігання біометричних даних може порушувати особисті права на конфіденційність.
- Помилки системи - можливі помилки, такі як неправильне відхилення дійсних користувачів або прийняття несанкціонованих осіб.
- Обмеження доступності - у деяких осіб можуть бути особливі фізичні особливості, які ускладнюють використання біометричної ідентифікації, наприклад, проблеми з відбитками пальців або зором.
- Висока вартість імплементації - установка біометричних систем може бути дорогою, особливо для великих організацій або систем із значною кількістю користувачів.
- Ризик безпеки - незважаючи на складність підробки, біометричні дані все ще можуть бути викрадені або використані шахраями, наприклад, через фотографії, відеозаписи або імітацію голосу.

Враховуючи ці плюси та мінуси, існує реальна необхідність у тому, щоб ретельно виважувати, чи відповідає біометрична аутентифікація вимогам у сфері безпеки користувачів, конфіденційності даних й загальній приватності й безпеки даних розроблюваної системи.

1.1.4 Сфера застосування біометричної аутентифікації

Біометрична аутентифікація, яка використовує унікальні біологічні або поведінкові характеристики людини для її ідентифікації, стає все більш розповсюдженою у різних сферах діяльності. Нижче наводяться основні сфера застосуванні біометрії.

В сучасному світі біометрична аутентифікація відіграє важливу роль у різних сферах, починаючи від мобільних технологій та закінчуючи

транспортною галуззю. У мобільних пристроях і комп'ютерах біометричні дані, такі як відбитки пальців, розпізнавання обличчя або сканування райдужки, використовуються для розблокування пристроїв, забезпечуючи безпеку особистих даних користувачів. Також вони застосовуються для авторизації фінансових операцій у мобільних банківських додатках.

У сфері безпеки та контролю доступу, біометричні системи встановлюються у корпоративних офісах та інших комерційних просторах для контролю доступу через сканування відбитків пальців або облич. Урядові та військові установи також використовують біометричні системи для захисту секретної інформації та стратегічно важливих об'єктів.

У фінансовій сфері біометричні дані застосовуються для ідентифікації клієнтів у банкоматах та при виконанні операцій у відділеннях банків. Для безпечного доступу до онлайн-банкінгу та інших фінансових послуг також використовується біометрична аутентифікація.

У сфері міжнародних переміщень та імміграції біометричні паспорти та системи спрощують процеси перетину кордонів та митного контролю, а також використовуються для ідентифікації осіб при оформленні віз та процесів надання громадянства.

У медичній галузі біометрична ідентифікація обмежує доступ до медичних записів, забезпечуючи конфіденційність інформації пацієнтів, а також використовується для ідентифікації осіб при доступі до медичних послуг.

У роздрібній торгівлі біометричні методи авторизують безготівкові платежі в магазинах та дозволяють надавати персоналізовані пропозиції та послуги клієнтам.

У освітніх установах застосування біометричних систем дозволяє відстежувати відвідуваність студентів та викладачів, а також забезпечує безпеку під час проведення іспитів.

У транспортній галузі біометрична ідентифікація використовується для видачі водійських посвідчень та реєстрації транспортних засобів, а також для контролю доступу до певних транспортних зон та послуг.

1.2 Основні види біометричної аутентифікації

1.2.1 Розпізнавання відбитків пальців

Розпізнавання за допомогою відбитків пальців використовує унікальні особливості відбитків кожної людини для встановлення її особистості або проведення верифікації. Оскільки відбитки пальців залишаються незмінними на протязі багатьох років і відрізняються у кожної людини, цей метод є достатньо надійним. Унікальність відбитків зумовлена їхніми характеристиками, такими як візерунки гребенів та борозен, а також мікроскопічними деталями, як-от точки, де гребені розходяться або закінчуються [4].

На рисунку 1.3 зображено структуру відбитку пальця.

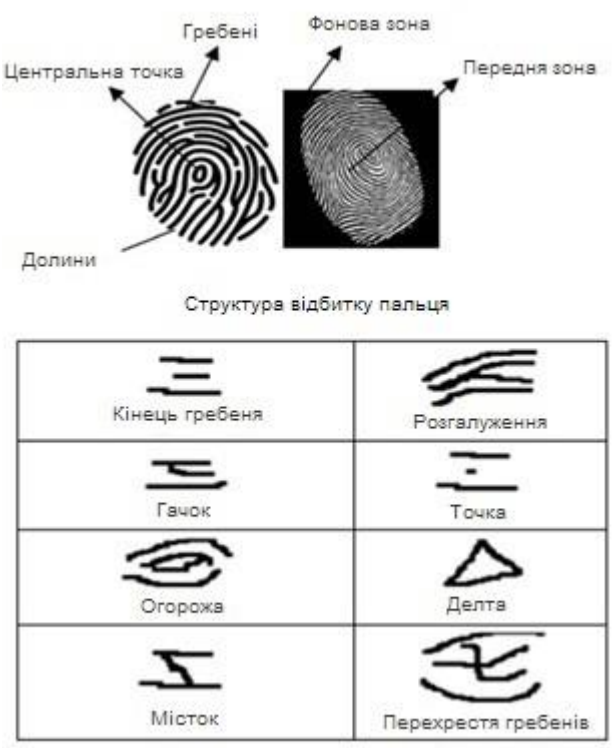

Рисунок 1.3 – Структура відбитку пальця

Використання сканування відбитків пальців поширене завдяки його економічної доступності та широкому спектру можливих застосувань, включаючи використання у смартфонах, ноутбуках, дверних замках та інших пристроях.

Має високу точність, забезпечує швидку обробку даних, легко інтегрувати в різноманітні пристрої, але в той же час можлива чутливість до фізичних ушкоджень пальців та ризик фальсифікації, якість сканування відбитків пальців може бути понижена через ряд зовнішніх та фізичних факторів, включаючи пошкодження шкіри, порізи, подряпини або наявність бруду, що може впливати на ефективність процесу.

Існують також проблеми зберігання відбитків пальців на загальнодоступних поверхнях, що може призвести до їх несанкціонованого копіювання та використання для обходу систем безпеки.

Дані відбитків пальців, зібрані скануючими системами, є дуже чутливими та персоналізованими, що створює ризики витоку інформації. З огляду на їх сталість, відбитки пальців, на відміну від паролів, що можна змінити, виявляються особливо вразливими у разі компрометації, створюючи довготривалу загрозу безпеці.

Серед моделей й методик, що використовуються для ідентифікації за відбитком пальця можна виділити наступні [5]:

1) Алгоритми на основі мінуцій (Minutiae-based algorithms): Цей класичний метод фокусується на виявленні специфічних елементів відбитків, таких як точки розходження та закінчення ліній.
2) Шаблонні алгоритми (Pattern-based algorithms): Такі алгоритми використовують основні візерунки відбитку, наприклад, петлі та хвилі, для ідентифікації особи.

3) Кореляційні алгоритми (Correlation-based algorithms): Вони зіставляють скановані відбитки з уже збереженими в базі даних, враховуючи можливі зсуви та обертання зображення.
4) Спектральний аналіз (Spectral analysis): Цей метод перетворює відбитки в спектральні властивості, дозволяючи їх розпізнавання на основі унікальних спектральних характеристик.
5) Штучні нейронні мережі (ШНМ): Вони використовуються для навчання системи розпізнавання відбитків пальців на основі численних зразків.
6) Тривимірне моделювання: Сучасні методи, які використовують 3D-технології для більш точного виявлення особливостей відбитків.
7) Глибоке навчання (Deep Learning): Методи, такі як згорткові нейронні мережі, ефективні у розпізнаванні відбитків, аналізуючи великі масиви даних для виявлення складних патернів.

1.2.2 Розпізнавання обличчя

Розпізнавання обличчя використовується для ідентифікації людини через унікальні характеристики її обличчя. Технологія базується на аналізі рис обличчя, включаючи форму частин обличчя та відстані між ними. Основна перевага цього методу - відсутність необхідності фізичного контакту, що робить його більш зручним і менше втручаючим у особистісну сферу, порівняно з іншими біометричними технологіями.

Розвиток технології розпізнавання обличчя тісно пов'язаний із прогресом у сферах штучного інтелекту, машинного та глибокого навчання. В рамках глибокого навчання, яке є частиною машинного навчання, використовуються складні нейронні мережі, що дозволяють аналізувати різноманітні аспекти облич. Ці алгоритми сприяють виявленню складних особливостей обличчя, підвищуючи точність процесів ідентифікації та верифікації.

Проте, технологія розпізнавання обличчя стикається з деякими викликами. Серед них - чутливість системи до змін у зовнішності, які можуть включати

освітлення, вираз обличчя, кути зображення, процес старіння, зміни у вигляді, як-от використання окулярів або макіяжу. Всі ці фактори можуть вплинути на точність системи, збільшуючи ризик помилкових результатів [6].

Розпізнавання обличчя використовується для розблокування пристроїв, у системах охорони, на контрольно-пропускних пунктах, в аеропортах й може бути інтегроване у системи відеоспостереження, але це порушує питання конфіденційності та етики, особливо щодо негласного моніторингу осіб. Крім того, законодавство часто не встигає за швидким розвитком і поширенням цих технологій, залишаючи пробіли у захисті прав на приватність і створюючи потенціал для зловживань. Це вимагає розробки строгіших правил та настанов для використання технології розпізнавання обличчя.

Нижче наведено деякі із відомих методів та систем, що застосовуються у розпізнаванні облич:

1) FaceNet від Google: Система, яка створює векторні представлення облич за допомогою нейронних мереж, дозволяючи вимірювати схожість між різними обличчями.
2) Convolutional Neural Networks (CNNs): Ці мережі є одними з найбільш вживаних для ідентифікації осіб, адже вони ефективно розпізнають та класифікують різні особливості обличчя.
3) OpenFace: Це відкрите програмне забезпечення (ПЗ), що базується на глибоких нейронних мережах, відоме своєю точністю у розпізнаванні облич.
4) DeepFace від Facebook: Ця передова система використовує глибоке навчання для ідентифікації осіб, ефективно працюючи з різними кутами та умовами освітлення.
5) OpenCV: Ця бібліотека для обробки зображень містить спеціалізовані інструменти для виявлення облич, використовуючи алгоритми Haar Cascade та LBPH.

6) Multi-task Cascaded Convolutional Networks (MTCNN): Комплексна система для виявлення та розпізнавання облич, що включає кілька етапів нейронних мереж.
7) Алгоритм Viola-Jones: Метод, який використовує набір класифікаторів для виявлення облич, популярний у вбудованих системах завдяки своїй швидкості.
8) Dlib: Ця бібліотека пропонує реалізацію розпізнавання облич на основі глибокого навчання, що використовується для визначення ключових точок обличчя.

1.2.3 Розпізнавання райдужки

Розпізнавання за райдужкою ока використовує процес створення високоякісного образу радужної оболонки, дозволяючи спеціалізованим алгоритмам аналізувати її складні візерунки. Унікальні особливості радужки, включаючи риси такі як кільця, фрактали та інші деталі, перетворюються у цифровий код. Цей код порівнюється з базою даних для швидкої та точної ідентифікації або верифікації особи, оскільки радужна оболонка залишається незмінною протягом усього життя.

На рисунку 1.4 представлено спрощену структуру побудови людського ока.

Рисунок 1.4 – Спрощена структура побудови людського ока

Данний вид біометрії використовується у банкоматах, смартфонах, системах безпеки, завдяки тому що такі біометричні дані дуже важко підробити. Однак, метод ідентифікації за радужкою має свої виклики. Головним серед них

є висока вартість, оскільки необхідність використання високоякісних камер для зафіксування тонких особливостей радужки збільшує витрати порівняно з іншими методами біометричної ідентифікації.

Крім того, ідентифікація за радужкою вимагає від користувачів певної близькості до скануючого пристрою. Це означає, що люди повинні позиціонувати свої очі на певній відстані від сканера та залишати голову нерухомою, що може бути незручним і обмежувати застосування цієї технології у певних сценаріях. Деякі люди можуть також відчувати неприємні відчуття під час сканування їхніх очей, хоча використовуване у сканерах інфрачервоне світло є безпечним [7].

Щоб захистити конфіденційність, біометричні дані, отримані під час сканування радужки, мають бути зберігатися та передаватися захищеним способом, щоб уникнути несанкціонованого доступу. Любі витоки цих даних можуть мати серйозні наслідки для приватності.

Нижче наведені деякі з методів й систем у галузі розпізнавання райдужки:

1) Алгоритм розробки Джона Доуна (John Daugman's Algorithm): Цей алгоритм є ключовим у сфері аналізу райдужок. Він використовує двовимірні перетворення Габора для створення унікального цифрового коду, відомого як ІрисКод, для кожної райдужки.
2) Backpropagation Neural Networks (BPNN): Ці нейронні мережі використовують алгоритм зворотнього розповсюдження помилки для виявлення та класифікації унікальних характеристик райдужки. BPNN ефективні в обробці та вивченні складних шаблонів, що робить їх корисними у розпізнаванні райдужки.
3) Методика Річарда Вайлдеса (Wildes' Algorithm): Розроблена Річардом Вайлдесом, ця техніка використовує обробку зображень та методи виявлення країв для виділення райдужки з усього зображення ока.

4) Локальні бінарні шаблони (Local Binary Patterns, LBP): Цей метод зосереджується на аналізі текстур райдужки, порівнюючи кожен піксель зображення з його навколишніми, щоб сформувати бінарні шаблони.
5) Self-Organizing Feature Maps (SOFM): Ці ШНМ застосовуються для візуалізації та класифікації високовимірних даних. У контексті райдужки, SOFM може використовуватися для виявлення унікальних шаблонів та характеристик.
6) Gaussian Mixture Models (GMM): Цей підхід передбачає моделювання райдужки за допомогою комбінації різних гаусівських розподілів, що допомагає детально аналізувати текстурні характеристики.
7) Фазове кодування: Цей метод залучає фазову інформацію зображень райдужки для формування її унікальної цифрової сигнатури.
8) Інтегральні перетворення: Використовуються для виділення текстурних особливостей райдужки, особливо для зображень з низькою якістю.
9) Глибоке навчання та нейронні мережі: Сучасні підходи використовують згорткові нейронні мережі (CNN) та інші методи глибокого навчання для виявлення та розпізнавання райдужки, особливо у складних умовах освітлення або при частковому закритті ока.

1.2.4 Розпізнавання голосу

Розпізнавання за допомогою голосу базується на використанні індивідуальних акустичних особливостей голосу людини. Цей метод стає все більш популярним у контексті дистанційної аутентифікації, наприклад, у смарт-домах, телемедичних системах або деяких банківських сервісах. Головні переваги голосового розпізнавання полягають у його неконтактності, легкості доступу і можливості використання на відстані через телефонічну чи інтернет-комунікацію.

Процес розпізнавання голосу включає два основних етапи. Перший етап - це екстракція характеристик голосу, де унікальні атрибути, такі як висота голосу, тембр, ритм та акцентуація, перетворюються у цифрові дані. Ці характеристики зумовлені як анатомією голосового тракту мовця, так і їх мовними звичками. Другий етап включає процес порівняння, де записані характеристики порівнюються зі зразками, що зберігаються у базі даних, щоб ідентифікувати або верифікувати особу.

Незважаючи на зручність та потенціал розпізнавання голосу, воно має деякі обмеження. Велика чутливість до фонового шуму є одним з головних недоліків цього методу. Навколишні шуми можуть перешкоджати чіткому захопленню та аналізу голосових характеристик [8]. Крім того, використання різних записуючих пристроїв або комунікаційних каналів з різним рівнем якості може погіршувати результати розпізнавання.

Зміни в голосі, які можуть бути викликані захворюваннями, емоційним станом або старінням, також можуть впливати на точність систем розпізнавання голосу. Емоційні зміни або втома можуть спотворювати голосові характеристики, ускладнюючи ідентифікацію.

Безпека та конфіденційність є ключовими питаннями для систем розпізнавання голосу. Існує ризик імітації голосу, який може використовуватися для обману цих систем, що робить захист від таких атак необхідним для забезпечення безпеки.

Ось деякі з найвідоміших моделей та алгоритмів, які використовуються для реалізації розпізнавання голосу:

1) Hidden Markov Models (HMM): Довгий час були стандартним вибором для розпізнавання голосу. HMM моделює мовлення як стохастичний процес з прихованими станами.
2) Deep Neural Networks (DNN): Поліпшують розпізнавання за рахунок виявлення складних шаблонів у вхідних даних.

3) GMM: Ці моделі відіграють важливу роль у розпізнаванні голосу, особливо коли вони інтегровані з HMM. GMM ефективно відображають характеристики звукових сигналів, розділяючи вхідні дані на декілька гаусівських розподілів, де кожен окремий розподіл представляє певний аспект звуку або мови.
4) Dynamic Time Warping (DTW): Цей метод забезпечує можливість порівняння аудіо послідовностей, які можуть відрізнятися за швидкістю вимови. DTW особливо корисний у розпізнаванні голосу, оскільки люди говорять з різною швидкістю та інтонацією. Методика DTW оцінює ступінь подібності між двома аудіо послідовностями, що змінюються в часі, що робить системи ідентифікації голосу значно більш адаптивними та точними.
5) End-to-End Learning Models: Моделі, які навчаються безпосередньо від звукових хвиль до текстових виводів, мінімізуючи необхідність ручної інженерії функцій.
6) Recurrent Neural Networks (RNN): Особливо ефективні для обробки послідовних даних, наприклад, для моделювання часових залежностей у мовленні.

1.2.5 Розпізнавання за геометрією руки

Біометрична ідентифікація за геометрією руки включає вимірювання та аналіз фізичних характеристик руки особи, таких як розміри пальців, їх ширина, товщина, та загальна площа руки. Ці дані перетворюються в цифровий формат і порівнюються із зразками в базі даних. Метод є відносно точним і менш вимогливим до точності, порівняно з іншими біометричними методами, і застосовується у різних областях, включаючи системи контролю доступу.

Однак, цей метод має декілька обмежень. Перш за все, вимірювальне обладнання для геометрії руки є відносно громіздким та дорогим у порівнянні з пристроями для сканування відбитків пальців або радужки. Це може бути

проблематичним для установ з обмеженим бюджетом чи простором. Крім того, необхідність розміщення руки на сканері може викликати занепокоєння з точки зору гігієни, особливо в місцях з підвищеним ризиком інфекції, як-то медичні установи. Регулярне очищення сканера також є додатковим обтяженням [9].

Люди з фізичними обмеженнями або травмами можуть відчувати труднощі при використанні цієї системи, що може призвести до неточностей у результаті сканування чи невдач у процесі ідентифікації. Оскільки особливості геометрії руки можуть бути менш унікальними, ніж інші біометричні маркери, як-от відбитки пальців чи радужка ока, існує підвищений ризик невірної ідентифікації, що може призвести до несанкціонованого доступу.

Кожен з представлених видів біометричного розпізнавання має свої унікальні переваги та обмеження, вибір яких залежить від конкретних потреб та обставин використання.

Серед методів та систем розпізнавання особи за геометрією руки виділяють наступні:

1) Сканування контуру руки: Цей підхід передбачає вимірювання зовнішнього силуету руки, включаючи довжину та ширину пальців, а також виміри долоні. Для створення точних зображень руки зазвичай використовуються сканери, які працюють на основі світлового випромінювання.
2) Тривимірне моделювання руки: З використанням сучасних технологій можливе створення тривимірних зображень руки, що дозволяє детальніше виявити її геометричні особливості. 3D-сканери оцінюють об'єм та форму руки, надаючи більш глибокий та точний аналіз.
3) Оптичне розпізнавання: Ця техніка використовує оптичні сканери для виявлення різноманітних характеристик руки, включаючи колір шкіри, її текстуру та загальну геометрію.

4) Інфрачервоне сканування: Використовуючи інфрачервоні сканери, можна вимірювати тепловий відбиток руки. Цей метод допомагає виявити особливості руки, які можуть бути неочевидними на звичайних оптичних зображеннях.
5) Вимірювання відстаней між характеристичними точками: Цей метод включає аналіз відстаней між важливими точками на руці, такими як місця з'єднання пальців або основи пальця.

2 ОГЛЯД СТАТИСТИЧНИХ МЕТОДІВ ОЦІНКИ

2.1 Роль статистичних даних в методах біометричної аутентифікації

Процес статистичного аналізу передбачає збір, розгляд, тлумачення та представлення інформації. У сфері біометричної аутентифікації, він сприяє визначенню точності, надійності та ефективності різноманітних біометричних систем.

Різноманіття статистичних методів включає в себе ряд підходів для обробки інформації, включаючи описову статистику (як то середнє, медіану, моду), інференційну статистику, а також багатовимірний аналіз.

Оскільки біометрична аутентифікація базується на унікальних фізичних чи поведінкових атрибутах особи для її ідентифікації, статистичні методи використовуються для аналізу цих атрибутів, оцінки різноманітності, точності ідентифікації, аналізу помилок (наприклад, помилково позитивних та негативних результатів) і підвищення загальної продуктивності системи.

Статистичний аналіз дозволяє оцінити надійність та точність методів біометричної аутентифікації, що критично важливо для їх ефективного застосування.

Роль статистики в аналізі біометричних аутентифікаційних методів включає в себе декілька важливих елементів [10]:

- оцінка точності та надійності;
- аналіз помилок;
- аналіз варіативності даних;
- приспособлення до різноманітності користувачів;
- застосування машинного навчання та оптимізація;
- настроювання параметрів;
- законодавче та етичне регулювання;

- неперервне вдосконалення.

У рамках оцінки точності та надійності, використовуються статистичні аналізи для визначення ефективності біометричних систем у ідентифікації або верифікації особи, що допомагає оцінити рівень точності цих систем. Крім того, статистичні інструменти застосовуються для оцінки консистентності та повторюваності результатів біометричних систем в різних умовах із часом, що гарантує їх надійність.

Аналіз помилок включає використання статистичних методів для вимірювання частоти помилково позитивних та помилково негативних результатів, які можуть виникнути під час ідентифікації осіб. Це охоплює аналіз неправильного ідентифікування несанкціонованих осіб як санкціонованих, а також випадки, коли санкціоновані особи не впізнаються. Також використовуються ROC (Receiver Operating Characteristic)-криві для графічного представлення та аналізу балансу між чутливістю та специфічністю системи.

У процесі аналізу варіативності даних застосовується статистика для дослідження варіацій у біометричних ознаках, таких як відбитки пальців або риси обличчя, між різними людьми. Також оцінюється вплив зовнішніх чинників, таких як освітлення або якість камери, на точність біометричних даних.

В рамках пристосування до різноманітності користувачів, статистика застосовується для оцінки впливу різних демографічних груп, таких як вік, стать, етнічна приналежність, на ефективність біометричних систем.

Застосування машинного навчання та оптимізація включає інтеграцію статистичних методів у машинне навчання для поліпшення точності біометричних систем, що дозволяє алгоритмам адаптуватися та вчитися на основі даних. Також використовується статистика для детальної настройки параметрів системи з метою досягнення оптимальних результатів.

В контексті законодавчого та етичного регулювання, статистичні методи застосовуються для аналізу ризиків, пов'язаних з захистом даних та конфіденційністю користувачів, що допомагає забезпечити їхню приватність та безпеку.

У рамках неперервного вдосконалення, статистичний аналіз використовується для сприяння інноваціям, поліпшенню існуючих методик і розробки нових підходів у галузі біометричної аутентифікації.

У підсумку, статистика є вирішальним фактором у розробці, оцінці та поліпшенні методів біометричної аутентифікації, забезпечуючи точність, надійність, безпеку та справедливість цих систем. Вона не тільки дозволяє аналізувати існуючі системи, але й сприяє їх неперервному розвитку та пристосуванню до змінюваних умов та нових викликів.

2.2 Методи оцінки методів біометричної аутентифікації

Для оцінки ефективності біометричних аутентифікаційних методів використовуються різноманітні техніки, які дозволяють точно виміряти та аналізувати ключові показники якості цих систем. Далі представлені найвідоміші з них:

- тестування за допомогою реальних біометричних даних;
- моделювання різноманітних умов;
- крос-валідація;
- ROC-Криві;
- стрес-тестування;
- аналіз демографічної чутливості;
- лонгітюдні дослідження;
- зворотний зв'язок від користувачів;
- статистичний аналіз;
- оцінка швидкості реакції та пропускної спроможності.

Для оцінки ефективності біометричних аутентифікаційних методів використовуються різноманітні методи та техніки, які дозволяють точно виміряти та аналізувати ключові показники якості цих систем. Одним з таких методів є тестування за допомогою реальних біометричних даних, яке включає використання фактичних даних для перевірки системи з метою визначення її точності, надійності, FRR, FAR в реальних умовах. Цей підхід має перевагу достовірної оцінки ефективності, хоча стикається з викликами зі збором високоякісних даних.

Інший метод - моделювання різноманітних умов, яке полягає у створенні різних сценаріїв, таких як освітлення, відстань, кути огляду для аналізу системи. Це дозволяє перевірити стабільність системи в широкому спектрі умов та виявити її адаптивність і гнучкість, хоча моделювання може не завжди точно відображати реальність.

Крос-валідація, що включає розподіл даних на тренувальні та тестові набори для оцінки здатності системи до узагальнення, знижує ризик перетренування системи, але вимагає великої кількості даних. ROC-криві, використовувані для оцінки балансу між чутливістю та специфічністю, надають інтуїтивно зрозумілу візуальну оцінку, хоча потребують статистичного розуміння для коректного аналізу.

Стрес-тестування включає навантаження системи великою кількістю запитів для визначення її максимальної пропускної спроможності та стабільності. Цей метод демонструє спроможність системи справлятися з великим навантаженням, але може бути складно відтворити реальні умови екстремального навантаження.

Аналіз демографічної чутливості оцінює роботу системи серед різних демографічних груп для виявлення упередженості, гарантуючи справедливість і доступність для всіх користувачів. Цей метод потребує доступу до різноманітних демографічних даних.

Лонгітюдні дослідження, що включають довготривале вивчення роботи системи, аналізують стабільність системи і її адаптацію до змін, надаючи уявлення про довгострокову стабільність, хоча це вимагає тривалого часу та значних ресурсів.

Зворотний зв'язок від користувачів, що збирає відгуки від користувачів, які використовують систему, оцінює загальну зручність та задоволеність користувачів, надаючи реальну інформацію про користувацький досвід, хоча і може стикатися з проблемою суб'єктивності думок.

Статистичний аналіз, який включає застосування статистичних методів, дозволяє об'єктивно порівнювати системи на основі числових даних, забезпечуючи глибше розуміння характеристик систем.

Оцінка швидкості реакції та пропускної спроможності, яка вимірює час реакції на запити та загальну працездатність, оцінює зручність використання та ефективність системи при великому обсязі запитів, будучи критично важливою для систем, що використовуються в умовах з високим рівнем вимог до доступу.

Таким чином, комбінація цих методів дозволяє отримати глибоке розуміння та всебічну оцінку ефективності біометричних аутентифікаційних систем, дозволяючи точно визначити їх сильні та слабкі сторони, а також виявляти потенційні шляхи для їх удосконалення.

2.3 Ключові показників якості методів біометричної аутентифікації

В області біометричної аутентифікації, ключові показники якості, такі як:

- Точність (Accuracy);
- Надійність (Reliability);
- FRR (False Rejection Rate);
- FAR (False Acceptance Rate);
- EER (Equal Error Rate).

В якості додаткових показників можна виділити наступні:

- Точність (Precision);
- Повнота (Recall);
- Час Відповіді (Response Time).

Дані показними мають значний вплив на загальну ефективність системи. Кожен показник відображає різні аспекти продуктивності, і їх взаємодія та баланс є критичними для досягнення оптимальної роботи системи.

Показник Accuracy оцінює загальну здатність системи правильно ідентифікувати або верифікувати користувачів, відображаючи загальну ефективність у правильній ідентифікації або верифікації осіб. Висока точність зменшує ймовірність помилок, таких як неправильні відмови в доступі або неправомірний доступ.

Показник Reliability вимірює стабільність і консистенцію результатів системи незалежно від зовнішніх факторів, таких як освітлення, погодні умови, або зміни в зовнішності користувача. Вона забезпечує, що система працює ефективно в різних сценаріях та умовах.

FRR вимірює частоту випадків, коли справжній користувач помилково відхиляється системою. Високий показник FRR свідчить про надмірну строгість системи, що веде до частого відхилення легітимних користувачів, що може викликати незручності та погіршити користувацький досвід. Низький - вказує на високу точність системи у впізнанні своїх користувачів, забезпечуючи їм легкий доступ. FRR розраховується, як відсоткове співвідношення помилково відхилених спроб до загальної кількості спроб автентифікації легітимними користувачами.

$$FRR = \left(\frac{\text{Кількість помилкових відмов}}{\text{Загальна кількість справжніх спроб}}\right) * 100\% \qquad (2.1)$$

FAR оцінює частоту, з якою система неправильно ідентифікує або верифікує неавторизовану особу як авторизовану. Високий FAR свідчить про недостатню строгість системи, що дозволяє неавторизованим особам отримувати доступ, становлячи серйозну загрозу безпеці. Низький FAR - вказує на ефективність системи в запобіганні несанкціонованому доступу. FAR розраховується як відсоткове співвідношення помилково прийнятих спроб до загальної кількості спроб автентифікації неавторизованими особами.

$$FAR = \left(\frac{\text{Кількість помилково прийнятих спроб}}{\text{Загальна кількість спроб аутентифікації імпостерами}}\right) * 100\% \qquad (2.2)$$

EER надає загальне уявлення про продуктивність системи, об'єднуючи аспекти безпеки FAR та зручності FRR. Вона використовується для порівняння різних біометричних систем, оскільки нижчий EER вказує на більшу точність системи.

Ідеальна біометрична аутентифікаційна система повинна мати низькі показники як FRR, так і FAR. Однак, часто необхідно шукати компроміс, оскільки зниження одного показника може призвести до підвищення іншого.

Важливість FRR та FAR варіюється в залежності від призначення системи. Наприклад, у системах з високими вимогами до безпеки (як-от на військових об'єктах) пріоритет надається низькому FAR, тоді як у системах, орієнтованих на користувацьку зручність (наприклад, у споживчих пристроях), важливіше забезпечити низький FRR. Глибоке розуміння та адекватна інтерпретація FRR та FAR дозволяють розробникам та адміністраторам систем біометричної аутентифікації точно оцінювати та оптимізувати їх продуктивність та рівень безпеки.

EER служить ключовим індикатором при аналізі ефективності методик біометричної аутентифікації, відіграючи вирішальну роль у визначенні рівноваги системи у плані точності та безпеки. Для розрахунку EER важливо

провести низку перевірок системи з використанням реальних та імітованих біометричних даних, визначивши при цьому критичну точку, де рівень помилкових відхилень (FRR) стає еквівалентним рівню помилкових прийнять (FAR). Визначення EER здійснюється шляхом обчислення відсотка, за яким значення FRR та FAR збігаються.

EER розглядається як мірник загальної продуктивності та захищеності системи, при цьому більш низький EER вказує на підвищену ефективність і безпечність. Випадки високого EER свідчать про нахил системи до частіших випадків помилкових відхилень і прийнять, що може представляти ризик для користувацького досвіду та безпеки, водночас низький EER підкреслює знижену частоту помилок обох видів, забезпечуючи більшу продуктивність і безпеку системи.

Важливість EER як метрики полягає у здатності охоплювати обидва види помилок, що дає змогу отримати всебічне уявлення про продуктивність системи. Оптимальний рівень EER значною мірою залежить від контексту застосування системи, при цьому в деяких сценаріях низький EER є критично важливим, тоді як у інших ситуаціях прийнятним може бути компроміс між помилковими відхиленнями та прийняттями. Глибоке розуміння і точний аналіз EER сприяють оптимізації систем біометричної аутентифікації, дозволяючи фахівцям точно настроювати системи відповідно до вимог точності та безпеки, що має ключове значення для їх практичного застосування у різноманітних областях.

Визначення EER, або рівня рівних помилок, не відбувається через пряме математичне рівняння, а шляхом емпіричного аналізу кривої ROC. Процес встановлення EER полягає у декількох етапах [11]:

1) Створення кривої ROC, на якій представлені взаємозв'язки між Частотою помилкових відмов (FRR) і Частотою помилкових прийнять (FAR) при різних рівнях порогових значень системи.
2) Виявлення місця на цій кривій, де значення FRR і FAR стають однаковими.

Цей процедурний підхід можна представити як визначення точки, де дві змінні (FRR і FAR) перетинаються, змінюючись залежно від установлених порогів системи.

EER вважається важливим показником для оцінки загальної надійності системи ідентифікації. Нижчий рівень EER свідчить про краще балансування між безпекою (низький FAR) і зручністю користування (низький FRR).Взаємодія та правильний баланс між цими показниками є вирішальними для створення ефективної, надійної та безпечної біометричної системи аутентифікації. Кожен показник впливає на інший, і оптимальна конфігурація залежить від конкретних вимог та контексту застосування системи.

2.4 Аналіз помилок першого та другого роду

Аналізування помилок першого роду (False Positive Error чи, іншими словами, False Acceptance) та другого роду (False Negative Error або, як іноді називають, False Rejection) відіграє важливу роль у визначенні ефективності методів біометричної ідентифікації. Ці помилки представляють собою різні види невірних висновків, до яких може прийти система, і в залежності від контексту їх застосування, кожна з них має свої особливості та наслідки. Нижче представлений детальний огляд цих помилок:

Помилки першого типу, або помилки позитивного схвалення, виникають, коли біометрична система помилково ідентифікує або верифікує особу, яка не

має авторизації, як таку, що має авторизацію. Це вимірюється за допомогою показника FAR. Наслідки такої помилки можуть бути критичними, особливо з точки зору безпеки, оскільки можуть відкрити доступ до конфіденційних ресурсів недобросовісним особам. Мінімізація цього виду помилок може бути досягнута шляхом удосконалення алгоритмів розпізнавання та підвищення точності біометричних даних.

Помилки другого типу, або помилки відмови, мають місце, коли система помилково відмовляє у доступі авторизованій особі, вважаючи її неавторизованою. Це вимірюється як FRR. Наслідки такої помилки включають затримки та незручності для легітимних користувачів, що може призвести до зниження продуктивності та задоволення користувачів. Зменшення цього виду помилок можливе за допомогою оптимізації порогових значень для ідентифікації та поліпшення здатності системи розпізнавати різноманітні біометричні характеристики.

Важливим аспектом є баланс між FAR та FRR. Існує обернена залежність між цими двома показниками: зниження одного часто призводить до збільшення іншого. Показник EER визначає момент, коли FAR і FRR є рівними, і чим нижчий цей показник, тим вища загальна продуктивність системи. Вибір між низьким FAR та низьким FRR залежить від вимог безпеки та зручності у конкретному застосуванні. Наприклад, у фінансових системах низький FAR може мати пріоритет, у той час як для особистих пристроїв, таких як смартфони, більш важливим може бути низький FRR.

Таким чином, аналіз помилок першого та другого типів допомагає визначити оптимальний баланс між безпекою та зручністю в системах біометричної аутентифікації, роблячи їх ключовими показниками при розробці та оцінці цих систем, особливо в контекстах, де помилки можуть мати серйозні наслідки.

2.5 Вплив умов використання методів біометричної аутентифікації на частоту помилок

Різні умови застосування біометричних аутентифікаційних методів можуть істотно впливати на частоту виникнення помилок, таких як помилки першого (FAR) та другого (FRR) роду, включаючи впливи з боку освітлення, якості обладнання, зовнішніх чинників, а також індивідуальних відмінностей між користувачами. Огляд основних впливів виглядає наступним чином [12]:

1) Якість та тип обладнання - різні біометричні сканери мають різний рівень точності та чутливості. Недоліки в якості або налаштуванні обладнання можуть збільшувати рівень FAR та FRR, як, наприклад, застарілі або малочутливі сканери відбитків пальців можуть частіше помилково ідентифікувати особу.
2) Умови освітлення - критично важливо для технік, що базуються на візуальному розпізнаванні, таких як ідентифікація за обличчям. Неадекватне або надто інтенсивне освітлення може впливати на якість зображення, що призводить до збільшення FRR.
3) Зовнішні впливи та умови навколишнього середовища - фактори, такі як пил, бруд, вологість, температура та інші зовнішні чинники, можуть впливати на ефективність біометричних систем. Брудні або вологі пальці можуть погіршувати роботу сканерів відбитків пальців, що веде до збільшення FRR.
4) Фізіологічні та поведінкові варіації - варіативність біометричних характеристик у населення, як-от розміри та форми облич, кольори очей, відбитки пальців і так далі. Велика варіативність може ускладнювати точне розпізнавання, що збільшує обидва види помилок.
5) Вік користувачів - вікові зміни можуть впливати на фізичні біометричні характеристики. Зокрема, зміни у вигляді обличчя через старіння

можуть впливати на системи ідентифікації за обличчям, збільшуючи FRR.

6) Емоційні та фізичні стани - стрес, втома або хвороба, можуть змінювати біометричні характеристики. Такі зміни можуть впливати на розпізнавання обличчя або голосу, підвищуючи рівень помилок.
7) Контекст використання - різноманітність сценаріїв використання ставить перед системою різні вимоги. Наприклад, у високобезпечних умовах система може бути налаштована таким чином, щоб мінімізувати FAR, що може спричинити підвищення FRR.

Ці фактори підкреслюють необхідність адаптивного підходу при налаштуванні біометричних систем, що має враховувати конкретні умови застосування та індивідуальні характеристики користувачів. Оптимальне налаштування системи залежить від умов її використання та має забезпечувати її ефективність та безпеку. Важливо усвідомлювати, що повне усунення помилок у біометричних системах є малоймовірним, тому головна мета полягає у досягненні оптимального балансу між безпекою та зручністю користувачів.

2.6 Важливість статистичних показників в залежності від сфери застосування

Значення різних критеріїв ефективності у методах біометричної ідентифікації значною мірою змінюється в залежності від області їх використання, оскільки кожен окремий контекст ставить перед системою свої унікальні вимоги щодо продуктивності та безпеки. Нижче наведено розгорнутий аналіз значущості цих показників у різних областях:

У фінансових установах, таких як банки, надзвичайно важливим є показник FAR, адже високий рівень цього показника може спричинити несанкціонований доступ до фінансових ресурсів. Також важлива висока точність, щоб клієнти могли безперешкодно користуватися своїми рахунками.

Для урядових установ, наприклад прикордонного контролю, важливим є низький рівень FRR, оскільки високий FRR може викликати затримки та незручності для громадян, а також важливий FAR для запобігання несанкціонованому в'їзду [13].

У медичних установах критично важливими є надійність та точність для забезпечення правильного доступу до медичних записів та захисту конфіденційності пацієнтів.

У комерційних установах, як-от роздрібній торгівлі, швидкість обробки даних є ключовою для забезпечення позитивного досвіду покупок, а FAR має важливе значення для безпеки транзакцій та запобігання шахрайству.

Для освітніх установ важлива масштабованість для обробки великої кількості студентів та персоналу, а також низький FRR для легкого доступу до освітніх ресурсів.

Високобезпечні об'єкти, такі як військові бази, вимагають низького рівня FAR для забезпечення високого рівня безпеки, а також точності та надійності для стабільної роботи системи в різних умовах.

Для масових заходів, наприклад спортивних подій, пропускна спроможність має велике значення для швидкої обробки великої кількості відвідувачів, а низький FRR важливий для запобігання затримок у легітимних відвідувачів.

У сфері особистого використання, як-от у смартфонах, час відповіді має велике значення для зручності користувача, а FAR та FRR повинні бути збалансовані для забезпечення як безпеки, так і комфорту користування.

Загалом, у всіх областях використання демографічна чутливість має велике значення для забезпечення справедливого та недискримінаційного доступу, а лонгітюдні дослідження важливі для оцінки стабільності та надійності системи в довготривалій перспективі.

В кожному з цих контекстів акцент на конкретних показниках визначається вимогами до безпеки, зручності, швидкості та надійності, а також очікуваннями користувачів і операторів систем. Вибір і оптимізація цих показників залежать від специфічних потреб та цілей, які ставить перед собою кожна конкретна біометрична система.

3 МЕТОДИ БІОМЕТРИЧНОЇ АУТЕНТИФІКАЦІЇ

3.1 Методи біометричного розпізнавання райдужки

3.1.1 Back Propagation Neural Networks

BPNN, або зворотне поширення нейронних мереж, відіграють ключову роль у біометричних технологіях ідентифікації, особливо при аналізі райдужної оболонки ока. Ці мережі, які є однією з розповсюджених форм ШНМ, розроблені на основі створення штучних нейронів, що передають сигнали спочатку вперед, а потім у зворотному напрямку розповсюджують помилки, використовуючи методику зворотного поширення помилок для ефективного навчання.

На рисунку 3.1 зображено внутрішню побудову BPNN.

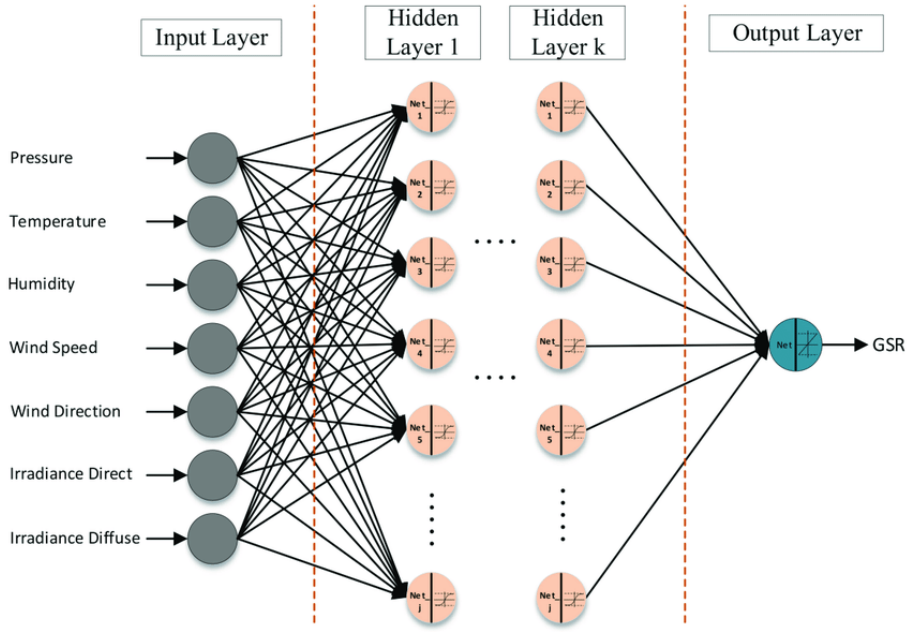

Рисунок 3.1 – Внутрішня побудова BPNN

Структура BPNN має кілька рівнів, включаючи вхідний шар, один або більше прихованих шарів та вихідний шар. Нейрони на кожному рівні з'єднані ваговими коефіцієнтами та зміщеннями, які регулюються протягом навчання для оптимізації реакції мережі на вхідні дані. У сфері ідентифікації райдужки, зображення ока подається на вхідний рівень BPNN, де відбувається первинне оброблення інформації, а потім інформація проходить через приховані шари для

детального аналізу та виділення особливостей райдужки, з акцентом на виявленні специфічних візерунків.

Процес навчання BPNN включає коригування помилок, допущених мережею під час ідентифікації цих візерунків, що дозволяє точно налаштувати ваги та зміщення, використовуючи алгоритм зворотного поширення, який базується на методах контрольованого навчання. Вихідний шар мережі видає результат, який підтверджує або спростовує відповідність зображення райдужки до збереженого в системі еталону. Точність BPNN у виявленні райдужки залежить від її здатності до навчання та адаптації [14]. Однак, налаштування та тренування цих мереж вимагає значних даних та обчислювальних ресурсів.

Таким чином, BPNN є надзвичайно ефективним інструментом у біометричному розпізнаванні райдужки, демонструючи значущість машинного навчання та штучного інтелекту у сфері передових безпекових технологій.

Таким чином, BPNN виявляється надзвичайно ефективним інструментом у біометричному розпізнаванні райдужки, підкреслюючи значущість машинного навчання та штучного інтелекту в області передових безпекових технологій.

Для обчислення вихідних даних кожного рівня мережі задіяна сигмоїдна функція, котра визначається наступним чином:

$$f(x) = \frac{1}{1-e^{-x}} \tag{3.1}$$

Функція помилки, котра представлена у квадратичному вигляді, встановлюється за наступною формулою:

$$E = \frac{1}{2}(y - f(x))^2 \tag{3.2}$$

В цьому контексті, $f(x)$ представляє собою прогноз, який мережа генерує, використовуючи дані з вихідного елемента, тоді як y є маркером класу для

конкретного випадку. Щоб обчислити ваги у нейронній мережі, потрібно визначити похідну від квадратичної функції помилки. Ця похідна для функції помилки, яка відноситься до певної ваги, встановлюється наступним чином:

$$\frac{dE}{dw_i}=(y-f(x))f'(x)a_i \tag{3.3}$$

де w_i - ваги для і-го вхідного змінного,

x - зважена сума входів,

a_i - входи до нейронної мережі.

Цей розрахунок повторюється для кожного навчального інстансу, а зміни, пов'язані з конкретною вагоюw_i, підсумовуються, множаться на швидкість навчання (мала константа) і віднімаються від поточного значення w_i. Це повторюється до тих пір, поки зміни у вагах стануть дуже малими.

Серед переваг BPNN можна сміливо виділити наступні особливості:

1) Здібність до навчання: BPNN ефективно навчаються на базі існуючих даних, що дозволяє їм вирішувати складні задачі, які не піддаються простому алгоритмізуванню.
2) Універсальне застосування: BPNN мають здатність апроксимувати будь-які функції з високою точністю, що робить їх придатними для широкого спектру задач, від аналізу зображень до прогнозування.
3) Використання в реальному часі: Після завершення тренування BPNN можуть оперативно обробляти дані, що є критично важливим для реагування у реальному часі.
4) Паралельна обробка: Використання паралельної обробки даних дозволяє BPNN швидко обробляти великі обсяги інформації.

Серед слабких сторін BPNN варто відмітити наступні:

1) Проблема місцевих мінімумів: BPNN можуть "застрягти" в місцевих мінімумах під час навчання, не знаходячи оптимального рішення.
2) Складність визначення гіперпараметрів: Правильний вибір таких параметрів, як швидкість навчання, кількість епох, та розмір мережі, вимагає значних зусиль та експериментів.
3) Довгий час навчання: Для BPNN характерні довгі періоди навчання, особливо при роботі з об'ємними даними, що ускладнює швидке оновлення моделі.
4) Ризик перенавчання: BPNN схильні до перенавчання при надлишку параметрів порівняно з доступним обсягом даних, що може знизити їх узагальнюючу здатність.
5) "Чорний ящик": BPNN часто критикують за непрозорість рішень, що ускладнює інтерпретацію та знижує довіру до них у відповідальних застосуваннях.
6) Залежність від якості даних: Ефективність BPNN прямо пропорційна якості та об'єму тренувальних даних; неякісні або упереджені дані можуть істотно вплинути на результативність мережі.

3.1.2 Self-Organizing Feature Map

Карта Кохонена, відома як SOFM, є ключовим інструментом у сфері ШНМ, що застосовується для візуалізації та аналізу складних даних з великою кількістю розмірностей у форматі з меншою просторовою складністю. Вона використовує метод неконтрольованого навчання, відомий як неконтрольоване підсилююче навчання (UNSLunsl), що дозволяє мережі вивчати різноманітність шаблонів без необхідності надання специфічної класифікаційної інформації, зберігаючи при цьому топологічну структуру.

Структуру SOFM представлено на рисунку 3.2.

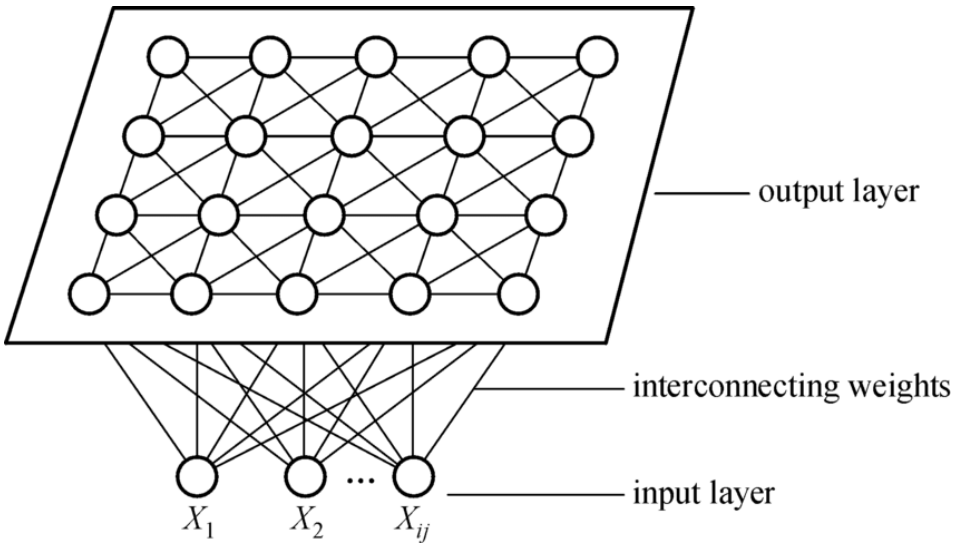

Рисунок 3.2 – Структура SOFM

У процесі навчання активні нейрони змагаються за можливість активації, і згідно з дослідженням, проведеними раніше, лише один нейрон стає активним [15]. В мережі SOFM, вибір переможного нейрона здійснюється через конкурентний шар, і за принципом Кохонена оновлюються не лише переможні нейрони, а й всі нейрони в певній області, що дозволяє адаптувати ваги нейронів відповідно до вхідних даних. Це робить їх корисними для подальшого аналізу. Одного разу, SOFM було використано для класифікації векторів, заснованих на DCT, з метою визначення присутності об'єктів на зображенні.

В ранніх дослідженнях зазначили, що SOFM може мати як одновимірну, так і багатовимірну структуру, залежно від конкретних обставин і доступних даних [15]. Кількість характеристик, які використовуватиме мережа SOM, визначає кількість необхідних вхідних з'єднань. Особливий механізм дозволяє сформувати вхідний вектор, аналізуючи різницю між наданими даними та налаштованими параметрами мережі, визначаючи вхід, який найкраще відповідає заданим критеріям. У конкурентному середовищі мережі цей вхід визнається переможцем, тоді як інші альтернативні варіанти ігноруються.

$$n' = -||IW_{1,1} - p|| \tag{3.4}$$

$$a^1 = C(n^1) \tag{3.5}$$

Коли надходить новий вхідний сигнал p, відбувається адаптація ваг переможного нейрона та його найближчих сусідів, щоб вони більше відповідали цьому сигналу p. Сусідні нейрони навчаються виявляти вектори з подібними характеристиками через постійне повторення цього процесу. Таким чином, мережа SOFM розширює свої навички в класифікації вхідних векторів, які вона обробляє [16]. Реалізація цієї функції досягається завдяки використанню алгоритму SOFM.

- Здійсніть початкове налаштування ваг нейронів W_{jk}^0, встановіть початкову швидкість навчання n^0 та визначте зону впливу для сусідніх нейронів h_{jk}^i.
- Вибір елемента даних для аналізу x^i.
- Вирахуйте нейрон, який найбільше відповідає вибраному елементу, базуючись на мінімальній евклідовій відстані між елементом та вагами нейронів.

$$\left|\left|x^i - w_{jk}^i\right|\right| = min_{jk}\{||x^i - w_{jk}^i||\} \tag{3.6}$$

- Проведіть коригування ваг нейронів, які найбільше відповідають вибраному елементу, для підвищення їхньої точності.

$$w_{ljk}^{i\neq 1} = w_{ljk}^i + n^i(x_l^i - w_{ljk}^i)jk \ \in h_{jk}^i \tag{3.7}$$

- У випадку виявлення значних змін у процесі класифікації, повторіть процес з другого кроку, в іншому випадку продовжіть з четвертого.
- Створіть та використовуйте матрицю ваг, специфічну для SOFM.

Згідно з вже проведенними іншими вченими дослідженнями дослідженнями, системи оброблення інформації в мозку, як природні, так і штучні, істотно залежать від структурних відображень, званих картами. Кохонен у 1995 році описав самоорганізовану карту (SOM) як процес неконтрольованого навчання, що вміє класифікувати шаблони без визначення їхніх класів [16]. В цьому процесі дані відображаються як позиції активних вузлів на карті, які проектуються з простору даних. На відміну від багатьох інших методів класифікації чи групування, SOM забезпечує топологічний порядок, за яким вихідні дані відображаються, зберігаючи при цьому подібність з вхідними шаблонами.

Учені зробили порівняльний аналіз двох розроблених класифікаторів на основі ШНМ – алгоритму зворотного поширення та самоорганізованих карт [16]. Для цього були використані дані експериментального дослідження ефективності цих класифікаторів у різних задачах розпізнавання образів. Виявилося, що в архітектурі зворотного поширення присутня більша кількість елементів обробки на прихованих шарах порівняно з самоорганізованими картами. Варто зазначити, що доповнена підкріпленням самоорганізована карта виявляє точність класифікації, яка залежить від точності застосування класифікації. Критичним аспектом порівняння алгоритмів BPNN та SOFM є відмінності в швидкості їх роботи. Основною проблемою використання мереж зворотного поширення є потреба у великій кількості часу для навчання через велику вибірку даних. Однак, як показали дослідження, використання самоорганізованих карт може бути ефективним рішенням для подолання цієї труднощі.

Серед переваг SOFM варто виділити наступні [16]:

1) Гнучкість у виявленні шаблонів: SOFM здатні виявляти складні взаємозв'язки в даних, що може сприяти визначенню неочікуваних закономірностей та взаємозв'язків.
2) Спрощення складних даних: SOFM вміло перетворюють складні, багатовимірні набори даних на більш прості та зручні для аналізу дво- або тривимірні структури.
3) Інтуїтивність представлення даних: Ці мережі створюють зрозумілі візуалізації, які допомагають в ідентифікації тенденцій та шаблонів у даних.
4) Самостійне навчання: SOFM не вимагають попередньо визначених категорій або міток, що робить їх ідеальними для ситуацій з обмеженими або відсутніми вихідними даними.
5) Збереження відносин між даними: SOFM відображають дані таким чином, що зберігається їх внутрішня структура, розташовуючи подібні точки даних поруч на карті.

Серед недоліків SOFM не можна не згадати наступні:

1) Залежність від налаштувань: Висока чутливість SOFM до вибору початкових параметрів може призвести до необхідності ретельного налаштування для досягнення оптимальних результатів.
2) Брак зворотного зв'язку: Ці мережі не забезпечують прямого роз'яснення того, як вхідні дані пов'язані з результатами класифікації.
3) Вимогливість до обчислювальних ресурсів: Тренування SOFM може вимагати значних обчислювальних зусиль, особливо для великих обсягів даних.
4) Обмеження масштабування: Створення ефективних SOFM для обробки величезних датасетів може бути проблематичним через збільшення часу навчання.

5) Схильність до надмірного навчання: Існує ризик, що SOFM будуть занадто точно відображати особливості тренувальних даних, що може обмежити їх здатність до узагальнення.
6) Труднощі з інтерпретацією: Хоча візуалізації, створені SOFM, можуть бути інформативними, іноді їх може бути важко інтерпретувати та зробити висновки на їх основі.
7) Складність вибору архітектури: Визначення оптимальної структури мережі, що включає кількість і розміщення нейронів, є критичним етапом, що впливає на якість моделі.

Загалом, SOFM відкриває нові перспективи та надає ефективні рішення в області біометричних систем, зокрема у сфері ідентифікації за допомогою райдужки ока. Ця технологія підкреслює важливість використання методів машинного навчання та можливостей штучного інтелекту для покращення сучасних систем безпеки, вказуючи на її значний потенціал у цій галузі [16].

3.2 Застосування методів біометричної аутентифікація

На мові програмування Python на базі PyTorch були застосовані методи SOFM та BPNN, з метою аналізу продуктивності двох відмінних стратегій використання ШНМ. Як інструмент машинного навчання, він надає всі необхідні засоби для проектування, навчання та імплементації нейронних мереж в біометричних додатках, зокрема в системах розпізнавання райдужки.

Моделі застосовувалися з використанням ОС Windows 10, процессора Intel Core i5-12600k, відеокарти AMD Radeon 5700XT та ОЗУ об’ємом 16 гігабайт з частотою 3600 MHz.

PyTorch пропонує потужні засоби для проведення експериментів та оптимізації нейронних мереж, що є ключовим для досягнення високої точності та ефективності в біометричному розпізнаванні райдужки. Для BPNN, що широко використовується в навчанні з вчителем, PyTorch надає зрозумілі та зручні інструменти для їх розробки та тренування, що є важливим для

підвищення точності в розпізнаванні райдужки. У випадку з SOFM, який представляє тип нейронних мереж без учителя, PyTorch підтримує динамічність обчислювальних графів. Це дозволяє модифікувати графи в реальному часі, що є корисним для методів, які потребують ітеративних змін, наприклад, у SOFM.

Процес дослідження включав ряд етапів: від збору інформації, розподілу даних, їх стандартизації, вилучення важливих ознак, зниження розмірності даних до класифікації та порівняння результатів [17]. Для аналізу використовувалася база даних (БД) CASIA-IrisV3, що є однією з найвідоміших та широко використовуваних баз даних у сфері біометричного розпізнавання райдужки. Ця БД містить велику кількість зображень райдужки, які зазвичай використовуються для навчання та тестування алгоритмів розпізнавання райдужки.

На рисунку 3.3 представлено зображення райдужки ока з БД CASIA-IrisV3.

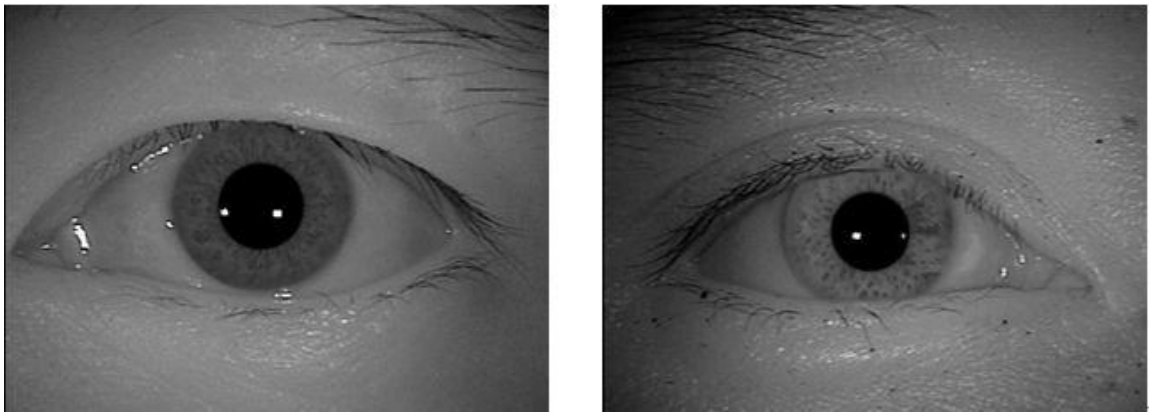

Рисунок 3.3 – Приклад зображення райдужок в БД

Виділення області райдужки ока на зображенні було здійснено за допомогою методу перетворення Хафа. Для стандартизації образів застосовувалися модель «Daugman Rubber Sheet» та вирівнювання гістограм. Було застосовано локальні бінарні шаблони для вилучення ознак та зниження розмірності даних. На заключному етапі проводилася класифікація даних за допомогою систем SOFM та BPNN, які проводилися окремо [14]. Особливу увагу приділяли аналізу точності розпізнавання, частоти помилкових схвалень, частоти помилкових відхилень та рівня помилок, які співпадають (EER) в

контексті використання SOFM та нейронних мереж зворотного поширення (BPNN).

На рисунку 3.4 зображено схема, що ілюструє послідовність кроків у використанні методів біометричної аутентифікації.

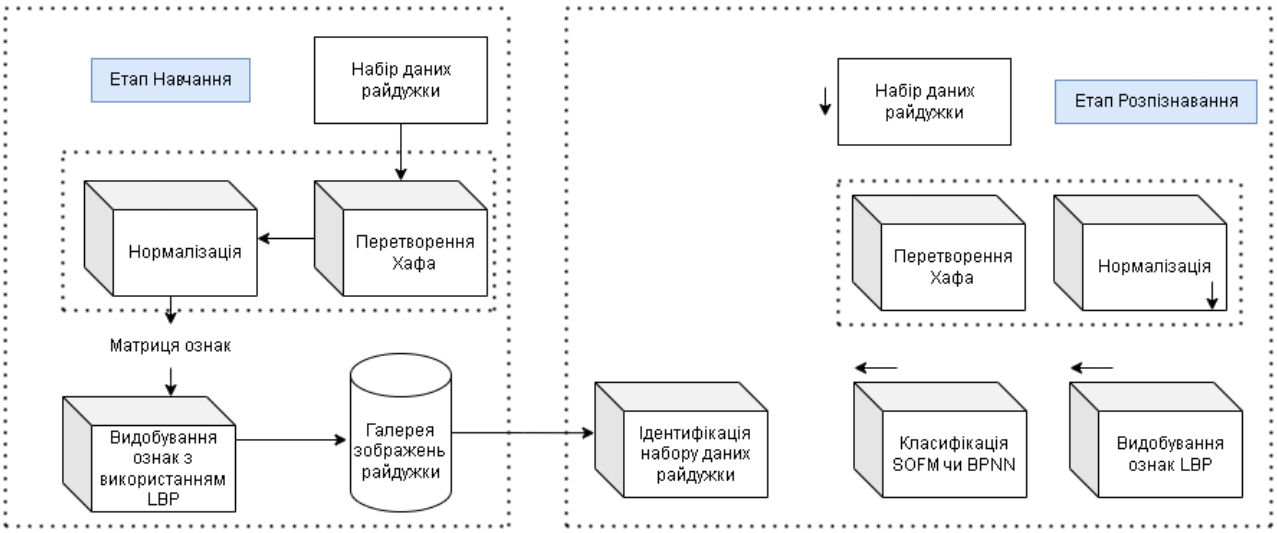

Рисунок 3.4 – Схема, що ілюструє послідовність кроків у використанні методів біометричної аутентифікації

Нижче на рисунку 3.5 приведено графічне зображення процесу, що демонструє етапи роботи з зображеннями райдужної оболонки ока, включаючи їх підготовку та аналіз із використанням методів біометричної аутентифікації райдужки ока.

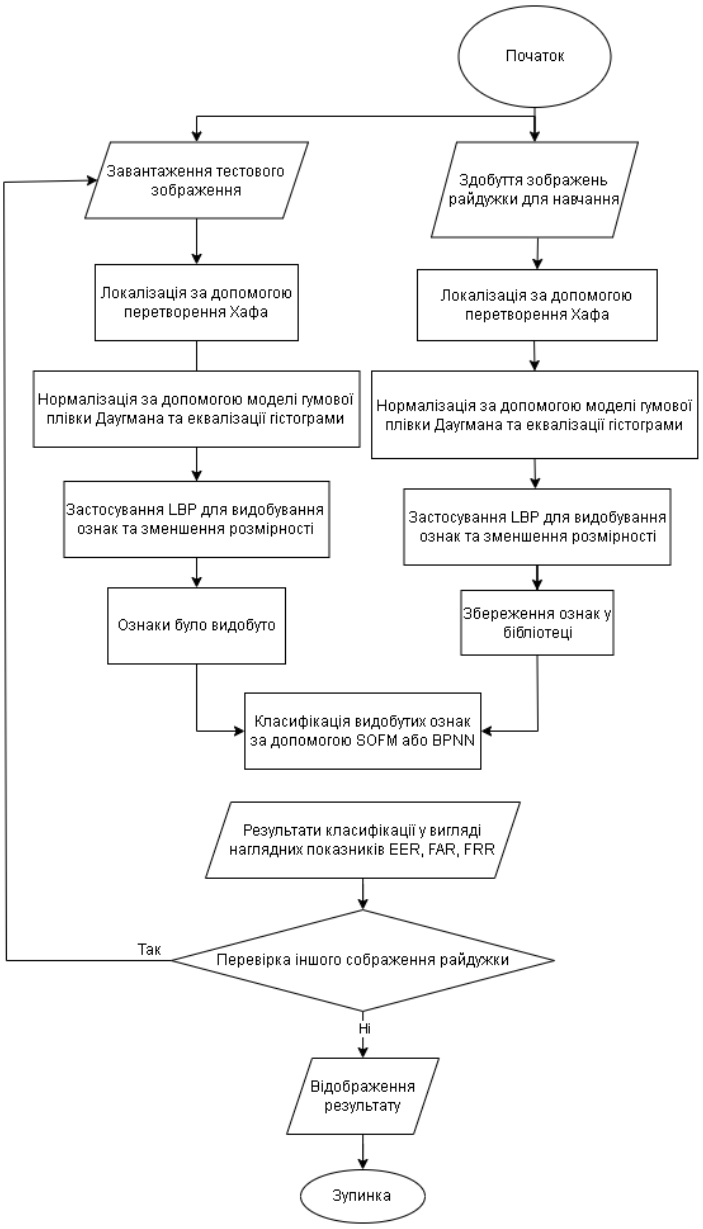

Рисунок 3.5 – Блок-схема з етапами роботи системи розпізнавання

3.3 Аналіз результатів роботи методів розпізнавання райдужки

На представленому рисунку 3.6 ілюструється, що методи BPNN та SOFM при ідентифікації райдужної оболонки показали порівнянну ефективність у контексті використання нейронних мереж із алгоритмом зворотного поширення помилки.

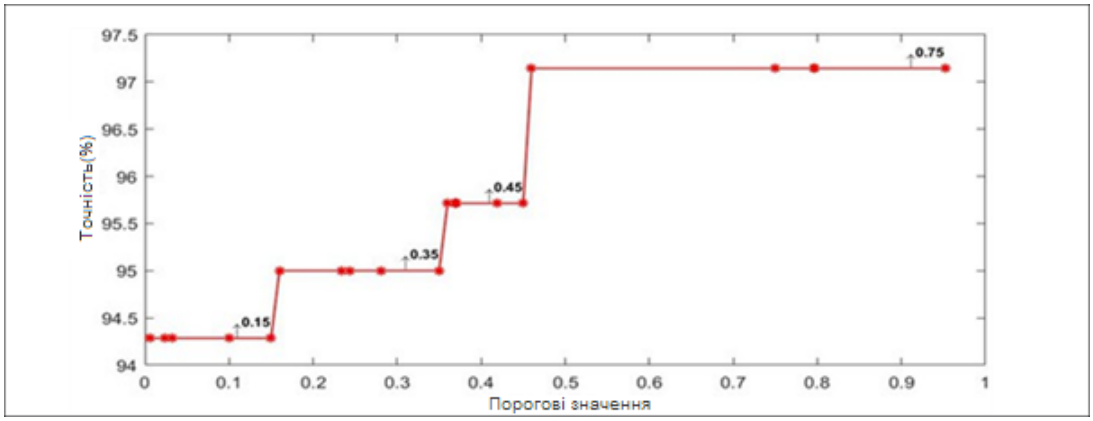

Рисунок 3.6 – Графік залежності порогових величин від точності

Для перевірки системи використовувались порогові значення котрі можна побачити на графіку, котрий зображено на рисунку 3.6. Згідно з рисунком, у цьому дослідженні було обрано порогові значення чотири різні порогові величини, а саме – 0.15, 0.35, 0.45 та, остання – 0.75.

Для забезпечення однакової точності в процесі виконання необхідно було встановити поріг вищий за першу величину. Точність ідентифікації була сталою для порогових значень в діапазонах від обраними пороговими величинами. Для перших трьох груп діапазонів були обрані перші три перші порогові величини, а для останньої групи – останнє значення, тобто 0.75.

Для оцінювання роботи системи були використані такі критерії, як FRR, FAR, EER, точність ідентифікації та час ідентифікації. Для вимірювання цих показників були задіяні різні рівні порогів.

Далі будуть наведені таблиці, в яких порогові величини будуть позначені римськими цифрами від I, що відповідає найменшій величині у 0.15, до IV – 0.75.

В таблиці 3.1, що знаходиться нижче, представлені показники ефективності BPNN за різних визначених порогових величин, результат роботи з якими було зображено вище. Ефективність алгоритму BPNN змінюється в залежності від різних порогових величин. Зі збільшенням порогових величин, зростає FRR та падає точність.

Таблиця 3.1 – Показники ефективності та точності роботи нейронної мережі BPNN

Обраний поріг (номер)	*Показник FAR*	*Показник FRR*	*Точність роботи*	*Час ідентифікації*
I	16.8	3.8	91.6	107
II	11.8	3.8	93.8	106.6
III	8.4	5.1	94.5	108.9
IV	5.1	6.3	95.2	102.8

Зокрема, при найбільшому обраному порозі BPNN показала FRR трохи більший за 6 %, а точність склала 95.2 %, за час ідентифікації у майже 103 секунди. Відповідно, час обчислень варіюється від 102 секунд до 109 секунд, у залежності від виставленої порогової величини.

Методи SOFM були протестовані з використанням таких же самих порогових величин , як зазначено в таблиці 3.2, котру можна бачити нижче.

Таблиця 3.2 – Показники ефективності та точності роботи нейронної мережі SOFM

Обраний поріг (номер)	*Показник FAR*	*Показник FRR*	*Точність роботи*	*Час ідентифікації*
I	11.8	1.3	95.2	83.9
II	8.4	2.5	96	85.6
III	5.1	2.5	97.4	88.5
IV	1.7	3.8	98.1	82.8

Відповідно до отриманих результатів, продуктивність SOFM варіюється в залежності від вибраного порогу. Було виявлено, що FRR зростає, а точність знижується із збільшенням порогу. Додатково в таблиці зазначено, що час обчислень різниться від майже 83 секунд, до, майже 89 секунд, в той же час показуючи тенденцію в підвищенні порогових величин.

Таблиці, що були наведені вище надають нам можливість оглянути результати оцінки всіх критеріїв для ідентифікації райдужної оболонки, включаючи дані, отримані за допомогою методів BPNN та SOFM. За результатами, представленими в таблицях вище, робиться висновок, що підхід SOFM використовує менше обчислювальних ресурсів порівняно з методом BPNN, незалежно від налаштувань порогових значень. Згідно з даними розглянутих вище таблиць, модель SOFM перевищує модель BPNN за

критеріями точності, частоти помилкових відмов (FRR) та частоти помилкових прийнять (FAR) при встановленій максимальній величині порогового значення, що ми можемо спостерігати у таблиці 3.3, котру наведено нижче.

Таблиця 3.3 – Результати роботи методів при пороговій величині 0.75

Обраний метод	*Показник FAR*	*Показник FRR*	*Точність роботи*	*Час ідентифікації*
SOFM	1.7	3.8	98.1	82.8
BPNN	5.1	6.3	95.2	102.8

Метод SOFM продемонстрував нам більше ніж 98 % точності, з показником FAR під 2 % й FRR майже 4 %, у той час як метод BPNN показав 95 % точності, FAR більший за 5 % та FRR у 6 %. Таким чином, стратегія SOFM виявилася більш продуктивною в порівнянні з методом BPNN.

На рисунку 3.7 та рисунку 3.8 зображено діаграми, на яких представлено показник ERR в ході роботи з методами.

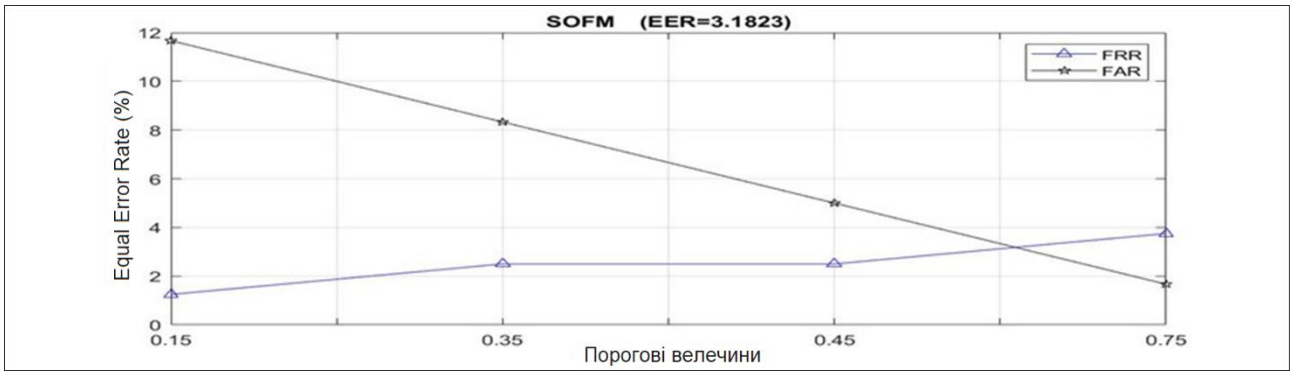

Рисунок 3.7 – Діаграма розрахунку показника ERR для методу SOFM

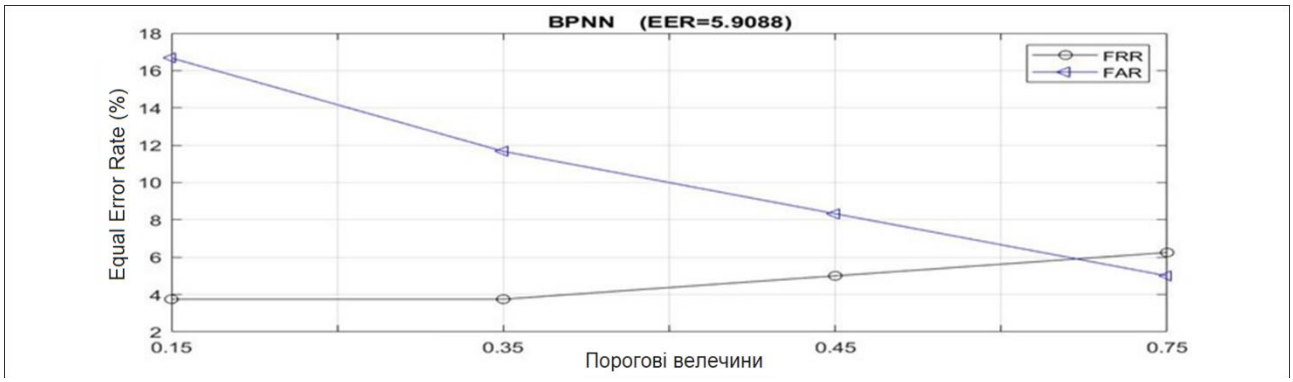

Рисунок 3.8 – Діаграма розрахунку показника ERR для BPNN

Рисунки вище ілюструють рівні показників EER у сфері ідентифікації райдужки ока при використанні методів SOFM та BPNN. Використання методу

SOFM призвело до EER у трохи більше за 3 %, у той час як при використанні методу BPNN EER склав майже 6 %. За результатами EER, метод SOFM підтвердив свою перевагу у зниженні рівня помилок.

У Таблиці 3.3 зазначено, що методика SOFM здійснює процес навчання даних райдужки набагато оперативніше, порівняно з методом BPNN. Це призводить до того, що SOFM вимагає меншої кількості обчислювальних ресурсів у порівнянні з BPNN. На рисунку 3.9 ілюструється взаємозв'язок між тривалістю навчання та числом здійснених спроб.

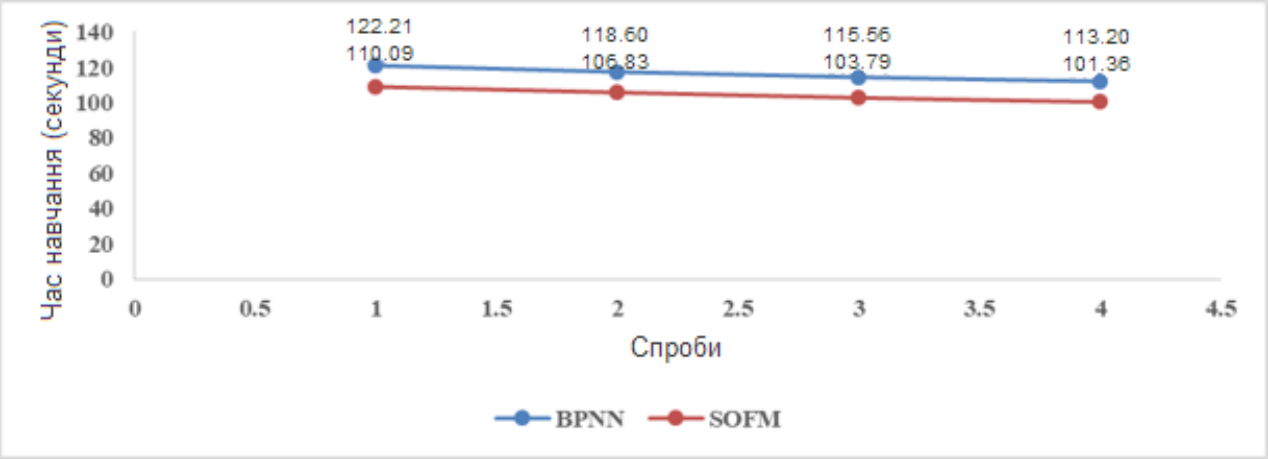

Рисунок 3.9 – Співвідношення спроб до часу навчання моделі

На рисунку відображено час, потрібний для навчання датасету райдужки кожним з методів. Виявлено, що BPNN встигла опанувати набори даних для серій з першої, по четверту всього за середній час 113.2 секунди. Для серій SOFM час навчання становив 110.1 секунди, 106.8 секунди, 103.8 секунди та 101.3 секунди відповідно. По завершенню чотирьох серій з датасетом райдужки, середній час тренування, зареєстрований BPNN, склав 117.4 секунди, тоді як середній час, витрачений SOFM, склав 105.5 секунди. Результати вказують на меншу обчислювальну складність SOFM порівняно з BPNN у контексті тренувального часу.

Аналогічна кореляція між середнім періодом часу, необхідним для розпізнавання, та встановленими пороговими величинами для методів BPNN та SOFM представлена на рисунку 3.10.

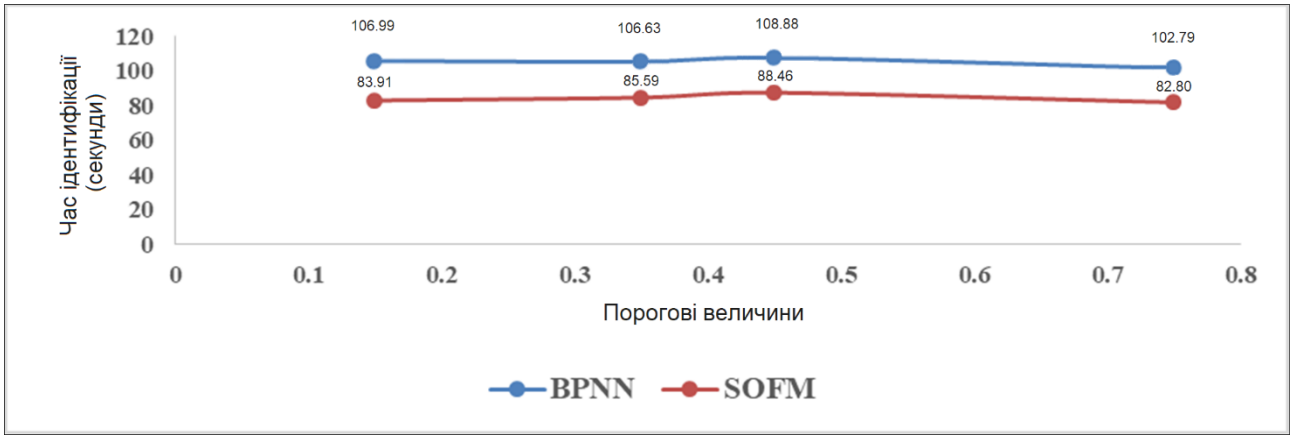

Рисунок 3.10 – Співвідношення між пороговими величинами та часом розпізнання

Аналіз графіка виявив, що залежність між часом розпізнавання (T_R) та пороговими показниками (th) має квадратичний характер з високим рівнем кореляції для обох підходів SOFM та BPNN, що продемонстровано в формулі 3.8 та формулі 3.9.

$$T_R = -43.8th^2 + 38 \quad R^2 = 0.8162 \tag{3.8}$$

$$T_R = -31.467th^2 + 22.132th + 103.02 \quad R^2 = 0.8262 \tag{3.9}$$

Метод SOFM виявляється менш вимогливим до обчислювальних ресурсів у порівнянні з BPNN, особливо коли йдеться про час, необхідний для навчання та тимчасового виявлення. Це висновок випливає з аналізу витраченого часу. Водночас, процес навчання та ідентифікації за допомогою BPNN вимагає більше часу. Можливо зробити передбачення щодо часу навчання, базуючись на встановлених порогових величинах, з урахуванням існуючих кореляцій.

3.4 Рекомендації щодо застосування тестованих методів

В ході проведення дослідження й аналізу методів біометричної аутентифікації радужки ока, будо сформульовано відповідні рекомендації, котрі виглядають наступним чином:

1) При аналізі райдужок ока з бази даних CASIA-IrisV3 або інших, використовуючи Python та, у моему випадку, PyTorch, рекомендується застосовувати методи BPNN та SOFM залежно від конкретних потреб.
2) Найкращі результати для цих методів, включаючи найнижчі показники FRR, FAR та найвищу точність ідентифікації, були зафіксовані при встановленні найбільшої порогової величини. Важливість вибору порогового значення має колосальний вплив на кінцеві результати дослідження.
3) Детальний аналіз виявив, що при застосуванні найвищої порогової величини порогової величини, SOFM продемонстрував зростання точності ідентифікації на 2.9 % та зниження показників FRR та FAR на відповідно 2.5 % і 3.4 % порівняно з BPNN. В той же час, рівень EER для SOFM був зафіксований на рівні 3.2 %, що значно менше ніж 6 % для BPNN, підтверджуючи перевагу SOFM в плані показників FRR, EER та більшої ефективності ідентифікації порівняно з BPNN.
4) Оскільки SOFM застосовує неконтрольований метод навчання, це дозволяє йому автоматично обробляти й групувати дані на основі їх властивостей, без необхідності попередньо встановлених маркерів. BPNN працює на принципах контрольованого навчання з застосуванням заздалегідь визначених маркерів або вихідних значень, що робить його ідеальним для завдань, в котрих доступні чіткі вихідні дані для тренування.
5) SOFM загалом потребує значно менше обчислювальних ресурсів порівняно з BPNN, що робить його більш придатним для опрацювання

великих обсягів даних. BPNN може бути більш обчислювально інтенсивним, особливо при великій кількості шарів і нейронів, що обмежує ефективність для обробки великих датасетів.

6) SOFM зазвичай видає високу якість роботи у виявленні шаблонів та його подальшої класифікації, особливо для завдань, що потребують візуалізації та зменшення розмірності даних. BPNN відрізняються високою точністю у класифікаційних та прогнозних завданнях, зокрема коли дані для навчання добре структуровані та відмічені.
7) SOFM ідеально підходять для завдань, що передбачають неконтрольоване навчання, як-от групування та візуалізація складних наборів даних. BPNN ефективні в широкому спектрі завдань контрольованого навчання , включаючи розпізнавання образів та аналіз часових рядів.
8) SOFM більш сприйнятливий до змін у налаштуваннях, наприклад, до швидкості навчання або наявної кількості шарів. BPNN потребує детальнішого та точнішого налаштування параметрів, включаючи швидкість навчання, тривалість епох та структуру мережі, що може бути досить непростим завданням.
9) SOFM надає більш зрозумілі й інтуїтивно інтерпретовані результати через графічне представлення, спрощуючи аналіз. BPNN часто вважається менш інтуїтивно зрозумілим через його складність й незрозумілу логіку прийняття рішень, тому BPNN часто називають "чорний ящик".
10) При інтеграції з Python та PyTorch обидва методи можна ефективно впровадити, звертаючи увагу на правильне налаштування архітектури нейронної мережі, параметри навчання та оптимізацію. Також важливо використовувати оптимізовані бібліотеки та інструменти PyTorch для підвищення продуктивності обчислень.

11) Для статистичного аналізу та оцінки, рекомендується регулярно перевіряти такі показники, як FRR, FAR та EER для обох методів, щоб забезпечити оптимальну точність та надійність системи ідентифікації. Це дозволяє отримати об'єктивну картину ефективності кожного методу та зрозуміти, як вони працюють у конкретному випадку та на конкретній системі.
12) Використання статистичних методів для аналізу результатів допоможе зробити непоганий порівняльний аналіз ефективності використовуваних методів.

ВИСНОВКИ

Під час дослідження й аналізу методів ідентифікації райдужки ока, які включають застосування SOFM та BPNN, було з'ясовано, що SOFM виявився більш продуктивним, ніж BPNN. Ці висновки ґрунтуються на аналізі трьохсот зображень райдужки з бази даних CASIA-IrisV3, де були випробувані різні варіанти використання цих технік за допомогою програмного забезпечення, розробленого на Python із застосуванням PyTorch.

Під час порівняльного аналізу цих двох підходів виявилось, що SOFM перевершує BPNN за такими аспектами, як точність ідентифікації, статистичні показники як FRR, FAR, EER та час обробки даних у процесі навчання та розпізнавання. Статистичний аналіз підтвердив значні розбіжності у продуктивності між SOFM та BPNN.

Ключовим виявився факт, що SOFM виказує меншу обчислювальну складність і вищу швидкість обробки порівняно з BPNN. Використання топологічної структури в методі SOFM виявилося більш ефективним, особливо у порівнянні з проблемами зворотного процесу навчання в BPNN. Це може бути важливим фактором при виборі методології для конкретних застосувань.

Отже, у контексті безпеки системи ідентифікації райдужки, засновані на SOFM, виявляються більш ефективними, ніж ті, що базуються на BPNN. SOFM пропонує вищий рівень точності класифікації, а також дає можливість більш детального аналізу даних. Важливо враховувати ці аспекти при розробці систем розпізнавання райдужки для забезпечення високої точності та оптимізації обчислювальних ресурсів.

Кваліфікаційна робота надає практичні рекомендації щодо вибору між SOFM та BPNN залежно від конкретних потреб проекту. SOFM є більш придатним для завдань з великими датасетами та неконтрольованим навчанням,

тоді як BPNN може бути більш підходящим для завдань із чітко визначеними вихідними даними.

Дослідження також вказує на важливість правильного налаштування архітектури нейронних мереж, параметрів навчання та оптимізації при використанні Python та PyTorch. Це може бути ключовим для успішної імплементації як SOFM, так і BPNN.

За підсумком кваліфікаційної роботи було:

- розглянуто основні поняття біометричної аутентифікації та сфера її прикладного застосування;
- розглянуто основні види біометричного розпізнавання й ключові методи біометричної аутентифікації за усіма популярними видами;
- поглиблено вивчено архітектури, механізмів навчання та застосування двох типів нейронних мереж – BPNN (Backpropagation Neural Network) та SOFM (Self-Organizing Feature Map), їхніх переваг та недоліків;
- проведено дослідження різних показників ефективності обох методів, включаючи точність ідентифікації, частоту помилкових відмов (FRR), частоту помилкових прийнять (FAR) та рівень помилок, що співпадають (EER);
- використано одну з провідних баз даних у сфері біометричного розпізнавання райдужки для тестування та оцінки ефективності використовуваних нейронних мереж;
- реалізацовано ці методи на мові програмування Python з використанням PyTorch для аутентифікації райдужки ока, включаючи збір даних, обробку, вилучення ознак, зниження розмірності даних та класифікацію;
- створено та перевірено ефективності класифікаторів з використанням зазначених методів для конкретних завдань біометричної аутентифікації;

- надано практичні рекомендації стосовно вибору та використання BPNN та SOFM в залежності від конкретних потреб та умов задачі аутентифікації.

В якості майбутнього розвитку кваліфікаційної роботи пропонується здійснити аналіз SOFM порівняно з іншими методами на базі нейронних мереж, включаючи конволюційні нейронні мережі (CNN), зокрема, торкнутися наступних тем:

- Оцінка точності та надійності: Порівняння точності та надійності SOFM та CNN у завданнях біометричної ідентифікації, зокрема, при розпізнаванні райдужки ока.
- Аналіз швидкості обробки: Вивчення часу обробки та відгуку обох методів, щоб визначити, який з них ефективніший для реального часу.
- Використання різних датасетів: Тестування обох методів на різних базах даних біометричних зображень для оцінки їх універсальності та адаптивності.
- Аналіз витрат обчислювальних ресурсів: Дослідження ефективності використання обчислювальних ресурсів для кожного з методів, з акцентом на оптимізацію та масштабування.
- Впровадження глибокого навчання: Розгляд можливостей інтеграції глибоких нейронних мереж для покращення точності та ефективності біометричної ідентифікації.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Мальченко М., Пожидаєв І., Волосянський О. Сучасні методи біометричної аутентифікації. Ivanov I. Analysis of the phaunistic composition of Ukraine // Innovations and prospects in modern science. Proceedings of the 12th International scientific and practical conference. SSPG Publish. Stockholm, Sweden. 2023. Pp. 21-27. С. 265-272. URL: https://sci-conf.com.ua/xii-mizhnarodna-naukovo-praktichna-konferentsiya-innovations-and-prospects-in-modern-science-20-22-11-2023-stokgolm-shvetsiya-arhiv/ (дата звернення 25.11.2023).
2. What is Biometric Authentication and How Does Is work? *LoginTC*. URL: https://www.logintc.com/types-of-authentication/biometric-authentication/#:~:text=Biometric%20authentication%20refers%20to%20a,%2C%20retinas%2C%20and%20facial%20features (дата звернення: 5.10.2023).
3. Biographical Characteristics: Definition & Examples. *StudySmarter*. URL: https://www.studysmarter.co.uk/explanations/business-studies/organizational-behavior/biographical-characteristics/ (дата звернення 7.10.2023).
4. Fingerprint Technology. *Innovatics.* URL: https://www.innovatrics.com/fingerprint-technology/#:~:text=Fingerprint%20recognition%20is%20the%20process,series%20of%20ridges%20and%20grooves (дата звернення 10.10.2023).
5. Zaeri, N. Minutiae-based fingerprint extraction and recognition. INTECH Open Access Publisher, 2011.
6. Face Recognition. *Electronic Frontier Foundation*. URL: https://www.eff.org/pages/face-recognition (дата звернення 12.10.2023).

7. What is Iris Recognition and how does it word? *NEC New Zaaland*. URL: https://www.nec.co.nz/market-leadership/publications-media/what-is-iris-recognition-and-how-does-it-work/#:~:text=Iris%20Recognition%20is%20a%20biometric,ideal%20form%20of%20biometric%20verification. (дата звернення 15.10.2023).
8. Voice recognition – *An Interwie. AbilityNet*. URL: https://abilitynet.org.uk/factsheets/voice-recognition-overview (дата звернення 18.10.2023).
9. Hand geometry – An Overview. Wikipedia. *The Free Encyclopedia*. URL: https://en.wikipedia.org/wiki/Hand_geometry#:~:text=Hand%20geometry%20is%20a%20biometric,measurements%20stored%20in%20a%20file (дата звернення 21.10.2023).
10. Mitra, S., Savvides, M., Brockwell, A. The Role of Statistical Models in Biometric Authentication. In: Zhang, D., Jain, A.K. (eds) Advances in Biometrics. ICB 2006. Lecture Notes in Computer Science, vol 3832. Springer, Berlin, Heidelberg., 2005
11. Zhu Yong, Tan Tieniu and Wang Yunhong. Biometric personal identification based on iris patterns. In: Proceedings of the IEEE international conference on pattern recognition., 2000
12. Errors in Biometric Systems. *Innovatrics ABIS*. URL: https://abis.innovatrics.com/public/docu/integration/biometrics/errors.html (дата звернення 25.10.2023).
13. Equal Error Rate – an overview. *ScienceDirect topics*. URL: https://www.sciencedirect.com/topics/engineering/equal-error-rate (дата звернення 03.11.2023).
14. *Neural networks and deep learning*.URL: http://neuralnetworksanddeeplearning.com/chap2.html (дата звернення 07.11.2023).

15. Self Organizing Maps – Konohen Maps. *GeeksforGeeks*. URL: https://www.geeksforgeeks.org/self-organising-maps-kohonen-maps/ (дата звернення 12.11.2023).
16. Self-organizing map. *Wikipedia. The Free Emcyclopedia.* URL: https://en.wikipedia.org/wiki/Self-organizing_map (дата звернення 16.11.2023).
17. What is Authentication? | Definition from TechTarget. *TechTarget Security*. URL: https://www.techtarget.com/searchsecurity/definition/authentication (дата звернення: 3.10.2023).

ДОДАТОК А

СПИСОК ПУБЛІКАЦІЙ МАГІСТРА

СУЧАСНІ МЕТОДИ БІОМЕТРИЧНОЇ АУТЕНТИФІКАЦІЇ

Мальченко Максим Сергійович,
Пожидаєв Ілля Дмитрович,
Волосянський Олег Володимирович,
Студенти
Харківський Національний Університет
імені В. Н. Каразіна
м. Харків, Україна

Вступ.

У сфері цифрових технологій, із зростанням потреби в надійних методах ідентифікації та перевірці особистості, біометричні аутентифікаційні системи стали більш поширеними та важливими. З посиленням процесу цифровізації, біометрична аутентифікація стала вирішальною технологією для забезпечення безпеки та приватності даних користувачів. Це обумовлено зростаючою необхідністю в міцних та надійних способах аутентифікації, особливо у контексті високого ризику крадіжки особистості, неавторизованого доступу та фінансових махінацій. Біометричні методи пропонують ефективне рішення для подолання цих безпекових викликів.

Мета роботи.

Дослідити й проаналізувати сучасні методи біометричної аутентифікації, їх сильні та слабкі сторони.

Матеріали та методи.

Основа систем біометричної аутентифікації полягає у використанні унікальних фізіологічних або поведінкових ознак, притаманних кожній людині. Це включає в себе такі риси, як відбитки пальців, характеристики обличчя, голосові відтінки, структуру райдужки та малюнки на сітківці ока та інших. Оскільки ці характеристики є властивими тільки конкретній особі, біометрична аутентифікація вважається одним із найбезпечніших та найзручніших методів встановлення особи.

Використання методів біометричного розпізнавання ставить під загрозу