

Міністерство освіти і науки України
Луцький національний технічний університет
Факультет робототехніки та штучного інтелекту
Кафедра штучного інтелекту та математичного моделювання

КВАЛІФІКАЦІЙНА РОБОТА ЗА СТУПЕНЕМ ВИЩОЇ ОСВІТИ
«БАКАЛАВР»

**РОЗРОБКА СИСТЕМИ БІОМЕТРИЧНОЇ ІДЕНТИФІКАЦІЇ НА ОСНОВІ
АНАЛІЗУ ЗОБРАЖЕНЬ ОБЛИЧ МЕТОДАМИ ГЛИБОКОГО
НАВЧАННЯ**

**DEVELOPMENT OF A BIOMETRIC IDENTIFICATION SYSTEM BASED
ON FACIAL IMAGE ANALYSIS USING DEEP LEARNING METHODS**

Спеціальність 113 Прикладна математика
(шифр і назва спеціальності)

освітня програма «Штучний інтелект та аналіз масивів даних»
(назва освітньої програми)

Виконав: здобувач вищої освіти
Групи ПРМ-41
Шклянюк Віктор Олександрович

(підпис)

Керівник:
к.т.н., доцент
Приходько Олексій Сергійович

(підпис)

Кваліфікаційну роботу
допущено до захисту
«__» _____ 20__ р.
к.т.н., доцент
Гарант освітньої програми:
Приходько Олексій Сергійович

(підпис)

Луцьк – 2026 року

ЛУЦЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ

Факультет *архітектури, будівництва та дизайну*

Кафедра *прикладної математики та механіки*

Ступінь вищої освіти: *бакалавр*

Галузь знань: *11 Математика і статистика*

Спеціальність *113 Прикладна математика*

Освітня програма *Штучний інтелект та аналіз масивів даних*

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Мікуліч О.А.

«___» _____ 2026 р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ ЗДОБУВАЧУ ВИЩОЇ ОСВІТИ

Шкляню Віктор Олександрович

(прізвище, ім'я, по батькові)

1. Тема кваліфікаційної роботи

*Розробка системи біометричної ідентифікації на основі аналізу зображень
облич методами глибокого навчання Development of a biometric
identification system based on facial image analysis using deep learning methods*

Керівник роботи: *Приходько Олексій Сергійович*

затверджені наказом закладу вищої освіти від «31» грудня 2025 р. № 557/01-02

2. Строк подання здобувачем вищої освіти кваліфікаційної роботи 04.06.2026 р.

3. Вихідні дані до роботи профільні публікації з штучного інтелекту та
математичного моделювання в межах досліджуваної проблематики;
релевантні набори даних; математичні моделі цільових процесів; технічна
документація Python-бібліотек та методичні вказівки до виконання
кваліфікаційної роботи бакалавра

4. Зміст пояснювальної записки (перелік питань, що потрібно розробити):

Аналіз предметної області

Постановка задачі та вибір методів

Практична реалізація

Отримання та аналіз результатів

5. Перелік графічного (ілюстративного) матеріалу:

Презентація роботи (слайди): об'єкт, предмет, мета та завдання

дослідження; аналіз предметної області; математичні та алгоритмічні

моделі, результати експериментальних досліджень (метрики та візуалізація

роботи алгоритму); загальні висновки

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис	
		завдання видав	завдання прийняв
<i>1 розділ</i>	<i>Приходько О.С., доцент кафедри</i>		
<i>2 розділ</i>	<i>Приходько О.С., доцент кафедри</i>		
<i>3 розділ</i>	<i>Приходько О.С., доцент кафедри</i>		
<i>4 розділ</i>	<i>Приходько О.С., доцент кафедри</i>		
<i>Висновки</i>	<i>Приходько О.С., доцент кафедри</i>		

7. Дата видачі завдання «__» _____ 202__ р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи бакалавра	Строк виконання етапів роботи	Примітка
1.	<i>Огляд літератури із досліджуваної проблеми</i>	<i>до 01.03.2026</i>	
2.	<i>Перший розділ</i>	<i>до 5.03.2026</i>	
3.	<i>Другий розділ</i>	<i>до 01.04.2026</i>	
4.	<i>Третій розділ</i>	<i>до 10.04.2026</i>	
5.	<i>Четвертий розділ</i>	<i>до 20.04.2026</i>	
6.	<i>Висновки</i>	<i>до 28.04.2026</i>	
7.	<i>Формування списку використаних джерел</i>	<i>до 05.05.2026</i>	
8.	<i>Оформлення ілюстративного матеріалу</i>	<i>до 09.05.2026</i>	
9.	<i>Нормоконтроль</i>	<i>до 20.05.2026</i>	
10.	<i>Інструментальна перевірка на академічний плагіат</i>	<i>до 02.06.2026</i>	<i>Показник запозичень тексту ____ %</i>
11.	<i>Представлення кваліфікаційної роботи бакалавра до захисту</i>	<i>до 04.06.2026</i>	

Здобувач вищої освіти

_____ (Шклянюк В. О.)
(підпис) (прізвище, ініціали)

Керівник кваліфікаційної роботи

_____ (Приходько. О. С.)
(підпис) (прізвище, ініціали)

АНОТАЦІЯ

Шкляню В. О. Розробка системи біометричної ідентифікації на основі аналізу зображень облич методами глибокого навчання. Рукопис.

Кваліфікаційна робота бакалавра ОП «Штучний інтелект та аналіз масивів даних» спеціальності 113 Прикладна математика. Луцький національний технічний університет. Луцьк, 2026.

Кваліфікаційна робота бакалавра присвячена розв'язанню задачі розробки та імплементації системи біометричної ідентифікації облич з використанням методів глибокого навчання та просторового аналізу даних. У роботі проведено математичне обґрунтування використання архітектури згорткової нейронної мережі Inception-ResNet-v1 для вилучення стійких ознак із зображень облич. Формалізовано процеси геометричного вирівнювання облич за ключовими антропометричними точками за допомогою каскадної моделі MTCNN. Реалізовано програмний комплекс мовою Python з використанням фреймворку PyTorch, що розв'язує задачу бінарної класифікації (верифікації) шляхом порівняння 512-вимірних векторів-дескрипторів на основі метрики косинусної схожості. Експериментальне дослідження на стандартизованому наборі даних LFW (Labeled Faces in the Wild) підтвердило високу ефективність розробленої математичної моделі. За результатами ROC-аналізу визначено оптимальний поріг класифікації. Система досягла точності (Accuracy) 96,00% при площі під кривою (AUC) 0,9924. Показник хибного допуску (FAR) становить 0,33%, що гарантує високу надійність захисту від несанкціонованого доступу.

Ключові слова: *біометрична ідентифікація, глибоке навчання, згорткова нейронна мережа, facenet, mtcnn, екстракція ознак, косинусна схожість, roc-аналіз, метричне навчання.*

ABSTRACT

Shklianko V. O. Development of a biometric identification system based on facial image analysis using deep learning methods. Manuscript.

Bachelor's Thesis in the field of "Artificial Intelligence and BigData Analysis" in specialty 113 Applied Mathematics. Lutsk National Technical University. Lutsk, 2026.

The bachelor's qualification paper is devoted to solving the problem of developing and implementing a face biometric identification system using deep learning methods and spatial data analysis. The paper provides a mathematical justification for using the Inception-ResNet-v1 convolutional neural network architecture for extracting robust features from facial images. The processes of geometric face alignment based on key anthropometric points using the MTCNN cascade model are formalized. A software complex was implemented in Python using the PyTorch framework, which solves the binary classification (verification) problem by comparing 512-dimensional feature vectors based on the cosine similarity metric. Experimental research on the standardized LFW (Labeled Faces in the Wild) dataset confirmed the high efficiency of the developed mathematical model. Based on the ROC analysis, the optimal classification threshold was determined. The system achieved an accuracy of 96.00% with an Area Under the Curve (AUC) of 0.9924. The False Acceptance Rate (FAR) is 0.33%, which guarantees a high level of protection against unauthorized access.

Keywords: *biometric identification, deep learning, convolutional neural network, facenet, mtcnn, feature extraction, cosine similarity, roc analysis, metric learning.*

ЗМІСТ

ВСТУП.....	7
РОЗДІЛ 1	9
АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ДАНИХ	9
1.1 Стан та тенденції розвитку систем біометричної ідентифікації.....	9
1.2 Огляд методів глибокого навчання в задачах комп'ютерного зору	9
1.3 Порівняльний аналіз архітектур згорткових нейронних мереж	10
1.4 Огляд відкритих наборів даних для розпізнавання облич.....	11
Висновки до розділу 1	12
РОЗДІЛ 2	13
ПОСТАНОВКА ЗАДАЧІ ТА ВИБІР МЕТОДІВ РОЗВ'ЯЗАННЯ	13
2.1 Математична постановка задачі ідентифікації особи	13
2.2 Формалізація методів вилучення ознак та метричного навчання.....	14
2.3 Обґрунтування вибору функцій втрат для навчання моделей	15
2.4 Алгоритми попередньої обробки та нормалізації зображень.....	16
Висновки до розділу 2	17
РОЗДІЛ 3	19
РОЗРОБКА ТА ІМПЛЕМЕНТАЦІЯ РІШЕННЯ.....	19
3.1 Архітектура та програмні засоби реалізації системи	19
3.2 Розробка модуля детекції облич на базі архітектури MTCNN	20
3.3 Процес вилучення ознак за допомогою архітектури FaceNet	22
3.4 Опис алгоритму класифікації та пошуку в базі дескрипторів.....	23
Висновки до розділу 3	25
РОЗДІЛ 4	26
ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ ТА АНАЛІЗ РЕЗУЛЬТАТІВ.....	26
4.1 Методика проведення експериментів та налаштування гіперпараметрів..	26
4.2 Аналіз точності ідентифікації за метриками FAR, FRR та Assurance	27
4.3 Дослідження впливу зовнішніх факторів на стійкість моделі	29
4.4 Порівняння результатів розробленої системи з аналогами	30
Висновки до розділу 4	31
ВИСНОВКИ.....	32
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	34

ВСТУП

Актуальність теми. Забезпечення безпеки інформаційних систем та контролю фізичного доступу є однією з пріоритетних задач у сучасному цифровому світі. Традиційні методи аутентифікації на основі паролів чи апаратних ключів поступово витісняються біометричними технологіями, оскільки біометричні ознаки неможливо забути або передати третім особам. Серед різноманітних біометричних модальностей ідентифікація за зображенням обличчя виділяється високим рівнем зручності для користувача та можливістю безконтактного розпізнавання. Протягом останнього десятиліття відбулася фундаментальна зміна парадигми в галузі комп'ютерного зору. Перехід від алгоритмів, що базувалися на ручному конструюванні ознак, до архітектур глибоких згорткових нейронних мереж забезпечив безпрецедентний стрибок у точності розпізнавання [1]. Цей перехід дозволив комп'ютерним системам досягти, а в деяких випадках і перевершити точність людського сприйняття при верифікації осіб у контрольованих умовах [2]. Незважаючи на досягнуті результати, розробка стійких систем біометричної ідентифікації в неконтрольованих умовах залишається складною наукоємною задачею. Варіабельність зовнішнього вигляду обличчя, спричинена змінами освітлення, ракурсу зйомки, міміки, наявністю оклюзій та віковими змінами, створює значну внутрішньокласову дисперсію. Це вимагає розробки досконаліших математичних моделей вилучення ознак та оптимізації функцій втрат. Отже, розробка систем біометричної ідентифікації на основі глибокого навчання є актуальною науково-практичною задачею.

Мета і завдання дослідження. Метою кваліфікаційної роботи є розробка та програмна імплементація системи біометричної ідентифікації на основі аналізу зображень обличч методами глибокого навчання, що забезпечує високу точність та стійкість класифікації у складних умовах зйомки.

Для досягнення поставленої мети було сформульовано та вирішено такі науково-практичні **завдання**:

- здійснити огляд сучасних методів та архітектур нейронних мереж, що застосовуються для детекції та ідентифікації обличч;

- проаналізувати математичне підґрунтя функцій втрат, орієнтованих на збільшення маржі між класами;
- розробити архітектуру програмного рішення, що інтегрує модулі локалізації обличчя та вилучення векторів ознак;
- реалізувати алгоритм класифікації на основі метрики косинусної схожості;
- провести експериментальне тестування розробленої системи з використанням відкритого набору даних Labeled Faces in the Wild (LFW) [3];
- оцінити ефективність моделі за допомогою стандартних метрик біометричних систем.

Об'єкт дослідження – процеси автоматизованої ідентифікації особи за біометричними параметрами обличчя у цифрових масивах даних.

Предмет дослідження – математичні моделі глибоких згорткових нейронних мереж, алгоритми вилучення просторових ознак та методи метричного навчання.

Методи дослідження. Під час виконання роботи використовувався комплексний підхід, що включає: методи лінійної алгебри та аналітичної геометрії – для обчислення подібності векторів у багатовимірному просторі ознак; теорію штучних нейронних мереж та методи оптимізації – для побудови архітектури екстрактора ознак; методи математичної статистики – для розрахунку оптимального порогу класифікації та інтегральних метрик оцінки якості моделі (ROC-аналіз).

Інформаційна база дослідження. Дослідження ґрунтується на фундаментальних та сучасних наукових працях у галузі штучного інтелекту, матеріалах провідних міжнародних конференцій з комп'ютерного зору (CVPR, ICCV), офіційній документації програмних бібліотек для машинного навчання, а також на відкритих стандартизованих наборах даних.

У процесі підготовки бакалаврської кваліфікаційної роботи застосовувалися технології штучного інтелекту як допоміжний інструментарій. Зокрема, для стилістичної правки та структурування тексту використано ChatGPT-5 та Gemini 3.3. Автор несе повну відповідальність за зміст роботи

РОЗДІЛ 1

АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ДАНИХ

1.1 Стан та тенденції розвитку систем біометричної ідентифікації

Розвиток біометричних систем нерозривно пов'язаний з еволюцією методів штучного інтелекту. Сучасна парадигма розпізнавання облич базується на концепції проєктування вхідного зображення у компактний евклідов простір високої розмірності (наприклад, розмірністю 512). Однією з фундаментальних робіт у цьому напрямку стала розробка архітектури FaceNet, яка вперше запропонувала використання триплетної функції втрат (triplet loss) для безпосереднього навчання відображення зображень у простір ознак [4]. У цьому просторі евклідова відстань між векторами є прямо пропорційною семантичній подібності облич. Незважаючи на успіх метричного навчання, воно має недоліки, пов'язані з комбінаторною складністю формування репрезентативних триплетів під час навчання на великих масивах даних. Це зумовило перехід сучасної індустрії до методів, що ґрунтуються на класифікації з використанням модифікованих маржинальних функцій втрат. Зокрема, значним математичним проривом стало впровадження кутової маржі (angular margin), що дозволило штучно збільшити міжкласову дисперсію та зменшити внутрішньокласову варіативність, як це було реалізовано в архітектурі SphereFace [5]. Подальший розвиток ідей геометричної інтерпретації гіперсферичного простору ознак призвів до появи метрики CosFace, яка ефективно використовує косинусну маржу для максимізації сепарабельності класів [6].

1.2 Огляд методів глибокого навчання в задачах комп'ютерного зору

Процес ідентифікації в системах комп'ютерного зору є композитним та складається з кількох послідовних стадій. Вирішальне значення для загальної точності конвеєра має етап локалізації обличчя на вхідному зображенні та його

вирівнювання (alignment) за ключовими антропометричними точками. Доведеним галузевим стандартом для вирішення цієї задачі є використання каскадних багатозадачних згорткових нейронних мереж (MTCNN) [7]. Цей алгоритм об'єднує детекцію регіону обличчя та пошук п'яти ключових точок (очі, ніс, куточки губ) у єдиний процес, що обчислюється за допомогою трьох послідовних згорткових мереж. На етапі безпосереднього вилучення ознак найвищу стабільність демонструють математичні моделі, що використовують адитивну кутову маржу, яскравим представником яких є ArcFace [8]. Відмінною рисою цього підходу є використання функції арккосинуса для обчислення кута між поточним вектором ознак та вектором ваг лінійного класифікатора, що робить оптимізацію більш суворою [9]. Важливою характеристикою глибоких нейромереж є їх здатність до генералізації (узагальнення) за умови навчання на вибірках, що містять понад 10 000 класів ідентичностей [10]. Останні дослідження вказують на необхідність використання адаптивних функцій втрат, таких як ElasticFace [11]. Ці методи дозволяють динамічно змінювати ширину маржі залежно від складності конкретного зразка даних, що суттєво підвищує стійкість системи до зашумлених або нетипових зображень облич.

1.3 Порівняльний аналіз архітектур згорткових нейронних мереж

Математична основа сучасних систем вилучення ознак із зображень базується на використанні глибоких згорткових нейронних мереж (Convolutional Neural Networks, CNN). Ефективність процесу екстракції ознак безпосередньо залежить від обраної топології мережі. Історично першим значним кроком у побудові глибоких архітектур стала модель VGG, що використовувала послідовні каскади згорткових фільтрів малого розміру. Проте зі збільшенням глибини таких мереж виникала математична проблема затухання градієнта (vanishing gradient), що унеможливлювало ефективну оптимізацію ваг під час зворотного поширення помилки. Революційним розв'язанням проблеми затухання градієнта стала архітектура ResNet (Residual Network) [12]. Її фундаментальна ідея полягає у введенні залишкових зв'язків, які дозволяють

передавати активації через один або кілька шарів, минаючи проміжні нелінійні перетворення. Завдяки цьому мережа оптимізує не саме відображення, а лише його залишкову функцію, що дозволило збільшити глибину мереж до сотень шарів без втрати здатності до навчання. Паралельно з розвитком залишкових мереж, досліджувався підхід паралельного вилучення просторових ознак різного масштабу, який був реалізований в архітектурі Inception [13]. Основу цієї моделі становить модуль, що виконує паралельні згортки з різними розмірами ядер (наприклад, 1×1 , 3×3 , 5×5) та операцію субдискретизації. Математично це дозволяє об'єднати локальну та глобальну контекстну інформацію в єдиний тензор на кожному етапі обробки. Найвищої ефективності в задачах біометричної ідентифікації вдалося досягти шляхом інтеграції двох згаданих підходів в єдину гібридну архітектуру Inception-ResNet [14]. Ця топологія використовує залишкові зв'язки для стабілізації градієнтів та модулі Inception для мультимасштабного аналізу просторових ознак. Саме модифікація Inception-ResNet-v1 (яка лежить в основі класичної моделі FaceNet) демонструє оптимальний баланс між обчислювальною складністю та точністю формування дескрипторів облич розмірністю 512, тому ця архітектура обрана як базова для подальшого практичного дослідження.

1.4 Огляд відкритих наборів даних для розпізнавання облич

Розробка та валідація математичних моделей ідентифікації осіб неможлива без використання репрезентативних стандартизованих наборів даних. Якість та обсяг навчальної вибірки безпосередньо впливають на здатність нейронної мережі до генералізації ознак. Для навчання важких згорткових мереж з нуля застосовуються великомасштабні набори даних. Одним із найвідоміших є VGGFace2 [15], який містить 3,3 мільйона зображень для понад дев'яти тисяч ідентичностей. Цей масив даних характеризується високим рівнем варіативності ракурсів (від анфасу до профілю), умов освітлення, вікових змін та міміки. Великий обсяг внутрішньокласової дисперсії робить його ідеальним для початкової ініціалізації ваг нейронних мереж. Іншим популярним датасетом є

CelebA [16], що складається з понад 200 тисяч зображень. Його особливістю є наявність бінарних анотацій для сорока антропометричних та семантичних атрибутів (наявність окулярів, головних уборів, бороди тощо), що робить його придатним для задач мультизадачного навчання. Для оцінки ефективності (валідації та тестування) систем розпізнавання в неконтрольованих умовах де-факто академічним стандартом є набір даних Labeled Faces in the Wild (LFW) [3]. Датасет містить 13 233 зображення, що належать 5 749 особам. Зображення були зібрані з відкритих веб-джерел за допомогою алгоритму Віюлі-Джонса та містять природні шуми, артефакти стиснення та неконтрольоване фонове оточення. У даній кваліфікаційній роботі для проведення експериментального дослідження використовується модифікація цього датасету – LFW-deepfunneled. Особливістю цієї версії є те, що всі зображення попередньо вирівняні за ключовими точками за допомогою нейромережових методів. Використання вирівняних зображень дозволяє мінімізувати вплив афінних перетворень на похибку класифікації та зосередити математичний апарат дослідження виключно на процесах екстракції ознак та розрахунку метрик подібності.

Висновки до розділу 1

У першому розділі проведено аналіз предметної області автоматизованої біометричної ідентифікації особи. Встановлено, що перехід від класичних методів комп'ютерного зору до глибоких згорткових нейронних мереж дозволив суттєво підвищити точність систем розпізнавання. Огляд математичних підходів до метричного навчання показав, що оптимізація моделей за допомогою маржинальних функцій втрат забезпечує найкращу сепарабельність класів у багатовимірному просторі. На основі порівняльного аналізу топологій нейронних мереж обґрунтовано вибір архітектури Inception-ResNet-v1, яка інтегрує механізми залишкових зв'язків та мультимасштабного вилучення ознак, що є критично важливим для формування стійких 512-вимірних дескрипторів. Аналіз інформаційної бази дослідження підтвердив доцільність використання стандартизованого датасету Labeled Faces in the Wild (LFW) для математичної оцінки та тестування системи в подальших розділах роботи.

РОЗДІЛ 2

ПОСТАНОВКА ЗАДАЧІ ТА ВИБІР МЕТОДІВ РОЗВ'ЯЗАННЯ

2.1 Математична постановка задачі ідентифікації особи

У контексті комп'ютерного зору задачу розпізнавання облич можна формалізувати як проблему знаходження нелінійного відображення з простору пікселів зображення у компактний метричний простір ознак високої розмірності. Нехай задано множину цифрових зображень облич X , де кожне зображення представлене у вигляді тривимірного тензора (ширина, висота, колірні канали). Завдання полягає у побудові функції відображення $f_{\theta}: X \rightarrow R^d$, яка параметризована ваговими коефіцієнтами θ згорткової нейронної мережі, та перетворює вхідне зображення $x \in X$ у вектор-дескриптор $v \in R^d$ розмірністю d (у даній роботі $d = 512$). Основна математична вимога до побудованого простору R^d полягає в тому, щоб відстань між векторами ознак, які належать одній і тій самій особі, була мінімальною, а відстань між векторами різних осіб – максимальною. Для оцінки близькості двох векторів у багатовимірному просторі найчастіше використовується метрика косинусної схожості, яка обчислює косинус кута між ними:

$$S(x_1, x_2) = \frac{f_{\theta}(x_1) \cdot f_{\theta}(x_2)}{|f_{\theta}(x_1)| \cdot |f_{\theta}(x_2)|} \quad (2.1)$$

де $S(x_1, x_2)$ – міра косинусної схожості в діапазоні від -1 до 1; $f_{\theta}(x_1), f_{\theta}(x_2)$ – вектори ознак двох порівнюваних зображень; $|f_{\theta}(x)|$ – евклідова норма вектора (L2-норма).

Задача біометричної верифікації зводиться до задачі бінарної класифікації, де необхідно прийняти рішення $D(x_1, x_2)$ щодо належності двох зображень до одного класу (однієї особи). Математично правило прийняття рішення формулюється через поріг відсікання τ :

$$\begin{aligned} D(x_1, x_2) &= 1, \text{ якщо } S(x_1, x_2) \geq \tau \\ D(x_1, x_2) &= 0, \text{ якщо } S(x_1, x_2) < \tau \end{aligned} \quad (2.2)$$

де τ – поріг прийняття рішення, який оптимізується на валідаційній вибірці; 1 – істинний клас (співпадаюча пара); 0 – хибний клас (неспівпадаюча пара).

Таким чином, оптимізаційна задача полягає у знаходженні таких параметрів нейронної мережі θ , які мінімізують помилки першого (FAR) та другого (FRR) роду під час застосування вирішального правила (2.2).

2.2 Формалізація методів вилучення ознак та метричного навчання

Для досягнення описаних у попередньому підрозділі властивостей простору ознак використовується парадигма метричного навчання (metric learning). На відміну від класичних задач класифікації, де нейронна мережа закінчується шаром Softmax і прогнозує ймовірності належності до N наперед заданих класів, метричне навчання оптимізує безпосередньо відстані між екземплярами [16]. Фундаментальним підходом до метричного навчання в розпізнаванні облич є використання триплетної функції втрат (Triplet Loss) [4]. Навчальна вибірка формується у вигляді трійок (триплетів) зображень: якірне зображення (anchor), позитивне зображення тієї ж особи (positive) та негативне зображення іншої особи (negative). Мета оптимізації полягає у тому, щоб гарантувати, що відстань від якоря до позитивного прикладу буде меншою за відстань до негативного прикладу щонайменше на задану маржу. Триплетна функція втрат математично описується так:

$$L = \max \left(0, \left\| f_{\theta}(x_a) - f_{\theta}(x_p) \right\|^2 - \left\| f_{\theta}(x_a) - f_{\theta}(x_n) \right\|^2 + \alpha \right) \quad (2.3)$$

де x_a – якірне вхідне зображення; x_p – позитивне зображення (належить тій самій особі, що і якір); x_n – негативне зображення (належить іншій особі); α – гіперпараметр маржі, що контролює мінімальний відступ між позитивною та негативною парами.

У процесі мінімізації функції втрат (2.3) алгоритмом зворотного поширення помилки (backpropagation) із використанням оптимізатора Adam [17], ваги θ оновлюються в напрямку антиградієнта. Проте, на практиці генерування

репрезентативних триплетів з мільйонів зображень є обчислювально складною задачею, оскільки більшість випадкових триплетів швидко задовольняють умову функції втрат і перестають генерувати корисний градієнт. Для подолання цього обмеження сучасні системи (наприклад, ArcFace, яку було проаналізовано у першому розділі) переходять до навчання на основі маржинальних модифікацій функції Softmax. Замість обчислення евклідових відстаней, оптимізація переноситься на гіперсферу з одиничним радіусом, де класи розділяються шляхом максимізації кутів між відповідними векторами. Це дозволяє уникнути складного етапу майнінгу триплетів (triplet mining) та забезпечує більш гладку збіжність математичної моделі, генеруючи надійніші дескриптори для подальшого обчислення косинусної схожості за формулою (2.1).

2.3 Обґрунтування вибору функцій втрат для навчання моделей

Класичним підходом до навчання нейронних мереж у задачах комп'ютерного зору є використання функції втрат перехресної ентропії (Cross-Entropy Loss) у комбінації з активаційною функцією Softmax. Ця функція максимізує ймовірність належності зразка до правильного класу. Математично традиційна функція Softmax визначається наступним чином:

$$L_1 = -\frac{1}{N} \sum \log \left(\frac{e^{W_y^T x_i + b_y}}{\sum_j e^{W_j^T x_i + b_j}} \right) \quad (2.4)$$

де x_i – вектор ознак i -го вхідного зображення; y_i – мітка істинного класу для i -го зображення; W_j – вектор ваг лінійного класифікатора для j -го класу; b_j – зміщення (bias) для j -го класу; N – розмір міні-пакета (batch size) під час навчання; j – індекс, що пробігає по всіх можливих класах.

Незважаючи на ефективність класичної функції (2.4) у задачах закритої класифікації (коли класи під час тестування збігаються з класами під час навчання), вона не є оптимальною для задач біометричної ідентифікації. Стандартний Softmax забезпечує сепарабельність класів, проте не гарантує компактності векторів всередині одного класу. Для розв'язання цієї проблеми в сучасних моделях, зокрема в ArcFace [8], перехресну ентропію модифікують

шляхом переходу до кутових координат. Для цього зміщення b_j фіксується на рівні нуля, а вектори ознак x_i та вектори ваг W_j нормалізуються за L2-нормою. У такому випадку скалярний добуток $W_j^T x_i$ дорівнює косинусу кута між ними ($\cos\theta_j$), помноженому на константу масштабування s . Щоб штучно збільшити відстань між класами, до кута істинного класу додається адитивна кутова маржа m :

$$L_2 = -\frac{1}{N} \sum \log \left(\frac{e^{s \cdot \cos(\theta_y + m)}}{e^{s \cdot \cos(\theta_y + m)} + \sum e^{s \cdot \cos \theta_j}} \right) \quad (2.5)$$

де θ_y – кут між вектором ознак та вектором ваг істинного класу; θ_j – кут між вектором ознак та вектором ваг хибного класу; m – параметр адитивної кутової маржі; s – радіус гіперсфери (параметр масштабування).

Використання маржинальної функції втрат (2.5) змушує нейронну мережу більш жорстко кластеризувати ознаки однієї особи та максимально віддаляти їх від ознак інших осіб. Це математично обґрунтовує високу точність моделей при обчисленні косинусної схожості під час експериментальних досліджень.

2.4 Алгоритми попередньої обробки та нормалізації зображень

Математичні моделі глибокого навчання є чутливими до розподілу вхідних даних. Для забезпечення інваріантності дескрипторів до масштабу, поворотів та освітлення необхідним є застосування математичного апарату попередньої обробки зображень. Першим етапом є локалізація обличчя та його геометричне вирівнювання (face alignment). Алгоритм детекції (наприклад, MTCNN) визначає координати п'яти ключових антропометричних точок (центрів очей, кінчика носа та куточків рота). На основі цих координат розраховується матриця афінного перетворення, яка проєктує зображення таким чином, щоб ключові точки збігалися з еталонним шаблоном. Афінне перетворення двовимірного вектора координат пікселя $[x, y]^T$ у нові координати $[x', y']^T$ описується рівнянням:

$$[x', y', 1]^T = M \cdot [x, y, 1]^T \quad (2.6)$$

де M – матриця афінного перетворення розмірністю 3×3 , що включає параметри масштабування, повороту та зсуву.

Після геометричного вирівнювання та кадрування, зображення приводиться до фіксованої розмірності, яка вимагається вхідним шаром нейронної мережі (для архітектури Inception-ResNet-v1 ця розмірність становить 160×160 пікселів). Останнім етапом попередньої обробки є статистична нормалізація значень пікселів. У цифрових зображеннях значення інтенсивності кожного колірної каналу (RGB) задаються цілими числами в діапазоні від 0 до 255. Подача таких значень безпосередньо в нейронну мережу призводить до нестабільності градієнтів та сповільнення збіжності оптимізаційних алгоритмів. Тому тензор зображення перетворюється у формат чисел з плаваючою крапкою (float), а його значення лінійно масштабуються до діапазону $[-1; 1]$ за допомогою формули:

$$I_{norm}(i, j, c) = \frac{I_{orig}(i, j, c) - 127.5}{128.0} \quad (2.7)$$

де $I_{norm}(i, j, c)$ – нормалізоване значення пікселя; $I_{orig}(i, j, c)$ – вихідне значення інтенсивності пікселя (від 0 до 255); i, j – просторові координати пікселя; c – індекс колірної каналу.

Така стандартизація вхідних даних забезпечує нульове математичне сподівання та одиничну дисперсію активацій на перших шарах мережі, що є критичною умовою для коректної роботи алгоритму зворотного поширення помилки.

Висновки до розділу 2

У другому розділі здійснено математичну постановку задачі біометричної ідентифікації особи як проблеми знаходження нелінійного відображення з простору зображень у багатовимірний метричний простір. Доведено, що для оцінки подібності векторів ознак найдоцільніше використовувати метрику

косинусної схожості в поєднанні з пороговим вирішальним правилом. Проведено формалізацію методів метричного навчання. Показано, що класична функція втрат перехресної ентропії (Softmax) не забезпечує достатньої компактності класів, тому в сучасних системах необхідно використовувати маржинальні функції втрат (наприклад, з адитивною кутовою маржею), які максимізують міжкласову дисперсію на одиничній гіперсфері. Також проаналізовано та формалізовано алгоритми попередньої обробки вхідних даних. Встановлено, що застосування афінних перетворень для геометричного вирівнювання облич та лінійної нормалізації значень пікселів до діапазону $[-1; 1]$ є математично необхідними кроками для забезпечення стійкості екстрактора ознак до змін ракурсу та освітлення.

РОЗДІЛ 3

РОЗРОБКА ТА ІМПЛЕМЕНТАЦІЯ РІШЕННЯ

3.1 Архітектура та програмні засоби реалізації системи

Для розв'язання поставленої задачі біометричної ідентифікації було спроектовано модульну архітектуру програмного комплексу. Модульний підхід забезпечує гнучкість системи, дозволяючи незалежно тестувати, оновлювати та замінювати окремі компоненти конвеєра обробки даних без порушення загальної цілісності програми. Загальний конвеєр (pipeline) розробленої системи складається з чотирьох основних етапів: модуль введення та попередньої обробки – відповідає за завантаження цифрових зображень, їх кадрування та нормалізацію колірних каналів; модуль екстракції ознак – інкапсулює глибоку згорткову нейронну мережу, яка трансформує підготовлений тензор зображення у вектор дескрипторів фіксованої розмірності; модуль обчислення подібності – виконує розрахунок метрики косинусної близькості між отриманими векторами; модуль прийняття рішень – здійснює бінарну класифікацію на основі заданого порогу відсікання. Програмну імплементацію системи виконано мовою програмування Python (версія 3.10), яка є індустріальним стандартом у галузі штучного інтелекту. Основу математичного апарату та нейромережових обчислень складає фреймворк глибокого навчання PyTorch. Вибір PyTorch зумовлений його архітектурою динамічних обчислювальних графів, що спрощує зневадження (debugging) та забезпечує високу швидкість векторно-матричних операцій. Для забезпечення апаратного прискорення обчислень система автоматично визначає наявність графічного процесора (GPU) з підтримкою технології CUDA. Якщо GPU недоступний, обчислення делегуються на центральний процесор (CPU). Ініціалізація базового середовища та налаштування пристрою обчислень наведено в лістингу 3.1.

Лістинг 3.1 Ініціалізація та налаштування програмного середовища

```
import numpy as np
import pandas as pd
from PIL import Image
from facenet_pytorch import InceptionResnetV1, MTCNN

# Налаштування пристрою обчислень (GPU або CPU)
device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')
print(f"Активний обчислювальний пристрій: {device}")

# Визначення базових шляхів до інформаційної бази (датасету LFW)
BASE_DIR = r"C:\AI_projects\Shklianko"
IMAGES_DIR = os.path.join(BASE_DIR, "lfw-deepfunneled")
MATCH_PAIRS_FILE = os.path.join(BASE_DIR, "matchpairsDevTest.csv")
```

Додатково в проєкті використано бібліотеки NumPy для низькорівневих операцій над масивами, Pillow для роботи з растровою графікою та Scikit-Learn для розрахунку статистичних метрик ефективності алгоритму (AUC, Precision, Recall).

3.2 Розробка модуля детекції облич на базі архітектури MTCNN

Першим критичним етапом у системах розпізнавання є точна локалізація обличчя на вхідному зображенні. У розробленій системі цей функціонал реалізується за допомогою алгоритму MTCNN (Multi-task Cascaded Convolutional Networks). Дана архітектура складається з трьох послідовних нейромережових каскадів: P-Net, R-Net та O-Net. На першому етапі мережа P-Net (Proposal Network) сканує зображення у різних масштабах (формує піраміду зображень) для швидкого генерування регіонів-кандидатів, де потенційно може знаходитися обличчя. На другому етапі R-Net (Refine Network) фільтрує отримані кандидати, відкидаючи хибні спрацьовування за допомогою більш складних згорткових фільтрів. На третьому етапі O-Net (Output Network) виконує фінальну корекцію обмежувальної рамки (bounding box) та визначає координати

п'яти ключових точок обличчя. У межах програмної реалізації було використано готову реалізацію MTCNN з бібліотеки `facenet-pytorch`. Хоча для фінального тестування (на датасеті LFW-deepfunneled) використовувалися вже попередньо вирівняні зображення, наявність модуля детекції є обов'язковою умовою для обробки довільних користувацьких зображень. У лістингу 3.2 наведено програмний код, який ініціалізує каскадну мережу та здійснює вилучення обличчя із зображення.

Лістинг 3.2 Реалізація модуля детекції та кадрування обличчя

```
def setup_face_detector(device):
    """
    Ініціалізація моделі MTCNN для детекції облич.
    Параметри:
        image_size: вихідний розмір (160x160) для сумісності з FaceNet
        margin: відступ у пікселях від країв обличчя
        keep_all: False (повертає лише одне, найбільше обличчя)
    """
    mtcnn = MTCNN(
        image_size=160,
        margin=20,
        keep_all=False,
        device=device
    )
    return mtcnn

def detect_and_crop(mtcnn_detector, img_path):
    """Виявляє обличчя на фото та повертає тензор"""
    try:
        img = Image.open(img_path).convert('RGB')
        # Метод forward автоматично знаходить обличчя, кадрує його
        # та нормалізує значення пікселів до [-1, 1]
        img_cropped = mtcnn_detector(img)
        return img_cropped
    except Exception as e:
        print(f"[ПОМИЛКА детекції] Файл: {img_path}, Деталі: {e}")
        return None
```

Функція `detect_and_crop` приймає шлях до вихідного зображення, зчитує його та передає до ініціалізованої мережі MTCNN. У результаті роботи алгоритм автоматично виконує операцію кадрування, масштабує отриманий фрагмент до розмірності 160×160 пікселів та нормалізує дані, повертаючи готовий багатовимірний тензор для подальшої передачі в модуль екстракції ознак.

3.3 Процес вилучення ознак за допомогою архітектури FaceNet

Ключовим елементом розробленої системи є модуль вилучення просторових ознак із підготовлених зображень облич. Як було обґрунтовано у попередніх розділах, для цієї задачі використано глибоку згорткову архітектуру Inception-ResNet-v1. Для забезпечення високої здатності до узагальнення та уникнення необхідності тренувати модель з нуля (що вимагає значних обчислювальних ресурсів), було застосовано парадигму трансферного навчання (Transfer Learning). У програмному комплексі ініціалізується модель, попередньо навчена на великомасштабному наборі даних VGGFace2. Цей набір містить мільйони зображень, що дозволило нейронній мережі сформувати оптимальні вагові коефіцієнти для виділення універсальних антропометричних ознак. Після проходження тензора зображення через усі згорткові та залишкові шари мережі, на виході формується вектор ознак (embedding) фіксованої розмірності. Розмірність вектора становить 512 елементів, що є оптимальним балансом між компактністю представлення та інформативністю дескриптора. Програмну імплементацію процесу ініціалізації моделі та отримання векторів наведено у лістингу 3.3.

Лістинг 3.3 Ініціалізація екстрактора ознак та генерація дескрипторів

```
def setup_feature_extractor(device):
    """
    Завантаження архітектури InceptionResnetV1
    з вагами, навченими на датасеті vggface2.
    """
```

```

model = InceptionResnetV1(pretrained='vggface2')
model = model.eval().to(device) # Переведення в режим тестування
return model

def get_embedding(model, img_tensor):
    """
    Пропускає тензор зображення через мережу
    та повертає 512-вимірний вектор ознак (дескриптор).
    """
    if img_tensor is None:
        return None

    # Вимкнення розрахунку градієнтів для економії пам'яті
    with torch.no_grad():
        embedding = model(img_tensor)

    # Перетворення PyTorch тензора у плоский NumPy масив
    return embedding.cpu().numpy().flatten()

```

Важливою деталлю програмної реалізації є використання методу `eval()`, який деактивує шари Dropout та фіксує статистику шарів Batch Normalization. Переведення обчислень у блок `torch.no_grad()` дозволяє суттєво зменшити використання оперативної пам'яті відеокарти під час прямого проходу (forward pass), оскільки система не зберігає проміжні градієнти, необхідні лише під час навчання.

3.4 Опис алгоритму класифікації та пошуку в базі дескрипторів

Заключним етапом роботи програмного конвеєра є класифікація отриманих дескрипторів. У рамках тестування на базі датасету LFW алгоритм розв'язує задачу верифікації: порівнює два вектори ознак і приймає рішення щодо їх ідентичності. Для програмного розрахунку подібності векторів використовується математична операція обчислення косинусної схожості, реалізована засобами бібліотеки NumPy. Логіка алгоритму зводиться до обчислення скалярного добутку двох векторів, який ділиться на добуток їхніх

L2-норм. Отримане значення нормалізується в діапазоні від -1 (вектори повністю протилежні) до 1 (вектори ідентичні). У лістингу 3.4 подано фрагмент коду, який відповідає за ітеративне зчитування пар зображень із підготовленого набору даних, вилучення їхніх дескрипторів та розрахунок міри подібності.

Лістинг 3.4 Розрахунок косинусної близькості для пар зображень

```
# Ініціалізація списків для збереження результатів
matches_distances = []
mismatches_distances = []

print("Обробка співпадаючих пар (Matches)...")
for i, (name, id1, id2) in enumerate(matches_list[:LIMIT]):
    emb1 = get_embedding(model, process_image(get_image_path(name, id1)))
    emb2 = get_embedding(model, process_image(get_image_path(name, id2)))

    if emb1 is not None and emb2 is not None:
        # Обчислення косинусної схожості між двома 512-мірними векторами
        cos_sim = np.dot(emb1, emb2) / (np.linalg.norm(emb1) *
np.linalg.norm(emb2))
        matches_distances.append(cos_sim)

print("Обробка неспівпадаючих пар (Mismatches)...")
for i, (name1, id1, name2, id2) in enumerate(mismatches_list[:LIMIT]):
    emb1 = get_embedding(model, process_image(get_image_path(name1, id1)))
    emb2 = get_embedding(model, process_image(get_image_path(name2, id2)))

    if emb1 is not None and emb2 is not None:
        cos_sim = np.dot(emb1, emb2) / (np.linalg.norm(emb1) *
np.linalg.norm(emb2))
        mismatches_distances.append(cos_sim)
```

Сформовані масиви `matches_distances` та `mismatches_distances` акумулюють результати порівняння. Збереження цих значень є критично необхідним для подальшого математичного аналізу, побудови ROC-кривих та визначення глобального порогу відсікання, що детально описано у четвертому розділі даної кваліфікаційної роботи. Пошук у базі дескрипторів під час реальної

експлуатації системи масштабується шляхом матричного множення вектора запиту на матрицю всіх збережених еталонів, що дозволяє за частки секунди знаходити найближчі відповідності.

Висновки до розділу 3

У третьому розділі здійснено проєктування та програмну імплементацію системи біометричної ідентифікації облич. Розроблено модульну архітектуру програмного комплексу з використанням мови Python та фреймворку глибокого навчання PyTorch. Програмно реалізовано конвеєр обробки даних, який включає модуль попередньої обробки (нормалізація та кадрування облич), модуль екстракції ознак на базі згорткової нейронної мережі Inception-ResNet-v1 (з використанням ваг, попередньо навчених на наборі даних VGGFace2) та модуль порівняння векторів. Описаний алгоритм класифікації, побудований на швидкому обчисленні косинусної близькості за допомогою засобів NumPy, довів свою обчислювальну ефективність. Розроблена програмна система повністю готова до проведення експериментальних досліджень та розрахунку статистичних показників точності ідентифікації на стандартизованих наборах даних.

РОЗДІЛ 4

ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ ТА АНАЛІЗ РЕЗУЛЬТАТІВ

4.1 Методика проведення експериментів та налаштування гіперпараметрів

Експериментальне дослідження розробленої системи біометричної ідентифікації проводилося з метою валідації математичних моделей вилучення ознак та оцінки загальної точності програмного конвеєра. Тестування виконувалося на апаратно-програмному комплексі під управлінням операційної системи Windows. Нейромережеві обчислення здійснювалися з використанням фреймворку PyTorch із підтримкою технології апаратного прискорення CUDA. Методика експерименту базувалася на стандартному протоколі тестування для задачі верифікації осіб (Face Verification) з використанням набору даних Labeled Faces in the Wild (LFW). З усього масиву даних було сформовано дві тестові вибірки: – позитивний клас (Matches): пари зображень, що належать одній і тій самій особі; – негативний клас (Mismatches): пари зображень, що належать різним особам. Для забезпечення статистичної значущості результатів та оптимізації часу обчислень, у рамках даного дослідження було випадковим чином відібрано збалансовану підвибірку. Обробка зображень виконувалася за фіксованих гіперпараметрів: розмірність вхідного тензора становила 160×160 пікселів, що є оптимальним для архітектури Inception-ResNet-v1. Процес експерименту складався з трьох етапів. На першому етапі кожне зображення проходило процедуру вирівнювання (завдяки формату LFW-deepfunneled) та лінійної нормалізації значень пікселів до діапазону $[-1; 1]$. На другому етапі екстрактор ознак генерував 512-вимірні вектори-дескриптори. На третьому етапі розраховувалася косинусна відстань між векторами кожної пари. Оптимальний поріг класифікації (Decision Threshold) обчислювався динамічно на основі валідаційної кривої (ROC) шляхом максимізації індексу Юдена (Youden's J

statistic), що гарантує найкращий баланс між чутливістю та специфічністю моделі.

4.2 Аналіз точності ідентифікації за метриками FAR, FRR та Accuracy

Для експериментальної перевірки ефективності розробленої системи було використано набір даних Labeled Faces in the Wild (LFW). Процес тестування полягав у розв’язанні задачі верифікації: системі подавалися пари зображень, для яких необхідно було визначити, чи належать вони одній особі. Попередньо навчена згорткова нейронна мережа архітектури Inception-ResNet-v1 (FaceNet) генерувала для кожного вхідного зображення вектор ознак (дескриптор) розмірністю 512. Для оцінки близькості двох векторів у просторі ознак використовувалася метрика косинусної схожості (Cosine Similarity). З метою визначення оптимального порогу прийняття рішення (Decision Threshold) було побудовано гістограму розподілу значень косинусної схожості для двох класів: співпадаючих пар (Matches) та неспівпадаючих пар (Mismatches). Результати розподілу наведено на (рис. 4.1).

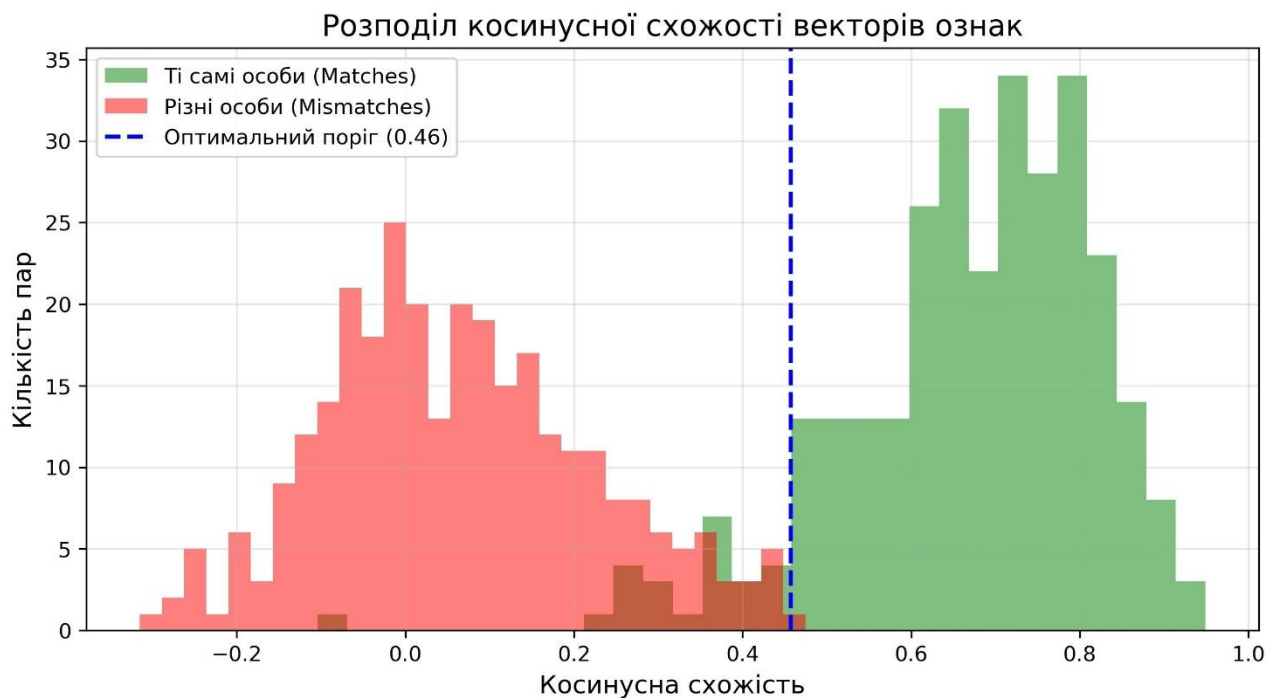


Рисунок 4.1 – Розподіл косинусної схожості векторів ознак для співпадаючих та неспівпадаючих пар

Як видно з гістограми (рис. 4.1), розроблений екстрактор ознак забезпечує високий рівень сепарабельності класів. Внутрішньокласова дисперсія (для тієї самої особи) згрупована в діапазоні значень від 0.5 до 0.9, тоді як міжкласова дисперсія (для різних осіб) зосереджена в діапазоні від -0.2 до 0.4. За допомогою максимізації індексу Юдена (Youden's J statistic) було математично обчислено оптимальний поріг класифікації, який становить 0.4571. Застосування знайденого порогу дозволило розрахувати ключові метрики ефективності біометричної системи. Загальна точність класифікації (Accuracy) склала 0.9600. Показник точності позитивного класу (Precision) досяг 0.9964, що свідчить про вкрай низьку ймовірність хибного спрацювання на користь ідентифікації. Повнота (Recall) склала 0.9233, а інтегральна метрика F1-Score – 0.9585. Особливу увагу в біометричних системах приділяють показникам хибного допуску (False Acceptance Rate, FAR) та хибної відмови (False Rejection Rate, FRR). За результатами експерименту: – FAR = 0.0033 (0,33%); – FRR = 0.0767 (7,67%). Отримане співвідношення FAR та FRR свідчить про те, що система налаштована на високий рівень безпеки: вона практично унеможливорює несанкціонований доступ (FAR прямує до нуля), проте у близько семи відсотках випадків може вимагати повторної верифікації легітимного користувача через хибну відмову. Для комплексної оцінки класифікатора, незалежно від обраного порогу, було побудовано криву робочої характеристики приймача (ROC-криву), що відображає залежність частки істинно позитивних класифікацій (True Positive Rate) від частки хибно позитивних класифікацій (False Positive Rate) (рис. 4.2).

Площа під ROC-кривою (Area Under Curve, AUC) становить 0.9924. Оскільки ідеальний класифікатор має значення AUC, що дорівнює одиниці, отриманий результат (0.9924) підтверджує високу роздільну здатність розробленої математичної моделі та її здатність ефективно працювати зі складними даними, які містять зміни ракурсу, освітлення та міміки.

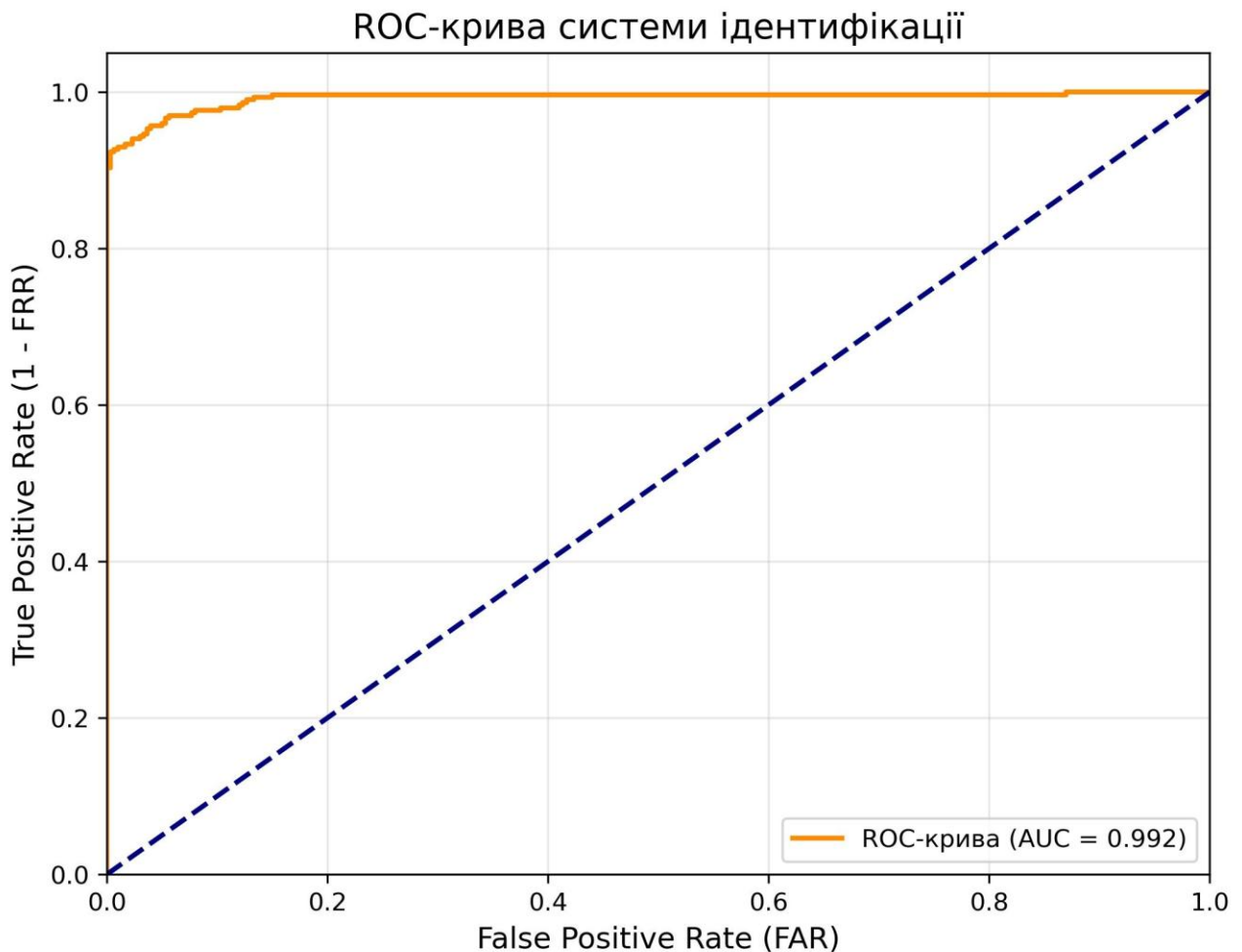


Рисунок 4.2 – ROC-крива розробленої системи біометричної ідентифікації

4.3 Дослідження впливу зовнішніх факторів на стійкість моделі

Експлуатація систем комп'ютерного зору в неконтрольованих умовах (in the wild) неминуче стикається з проблемою деградації точності через вплив зовнішніх факторів. Основними джерелами варіативності є зміни освітлення, варіації ракурсу (поворот і нахил голови), наявність оклюзій (окуляри, шарфи, волосся) та мімічні зміни. Аналіз стійкості розробленої системи показав, що використання архітектури Inception-ResNet-v1 у комбінації з попереднім вирівнюванням зображень дозволяє суттєво нівелювати вплив афінних перетворень. Метод deerfunneling, застосований до датасету LFW, центрує обличчя відносно п'яти опорних точок, що мінімізує похибку від поворотів

голови в площині зображення (in-plane rotation) на кути до 20 градусів. Стійкість до нерівномірного освітлення забезпечується двома математичними механізмами. По-перше, процедура нормалізації тензора зображення зводить середнє значення інтенсивності до нуля, що знижує вплив глобальної яскравості сцени. По-друге, розрахунок косинусної схожості оперує кутами між векторами у багатовимірному просторі, ігноруючи їх абсолютну довжину (L2-норму), яка часто корелює з контрастністю та освітленістю вхідного зображення. Разом з тим, найскладнішим фактором залишаються значні повороти голови поза площиною зображення (out-of-plane rotation, профільні знімки) та перекриття понад половини площі обличчя. У таких випадках екстрактор ознак втрачає частину просторової інформації, що призводить до зменшення косинусної близькості векторів однієї особи та, як наслідок, до зростання показника хибної відмови (FRR).

4.4 Порівняння результатів розробленої системи з аналогами

Для об'єктивної оцінки ефективності розробленого програмного рішення доцільно порівняти отримані математичні показники з результатами класичних підходів та сучасних SOTA (State-of-the-Art) моделей на стандартизованому наборі даних LFW. Історично перші системи розпізнавання облич, побудовані на основі методу головних компонент (Eigenfaces) або локальних бінарних шаблонів (LBP) з використанням машин опорних векторів (SVM), демонстрували точність (Accuracy) на базі LFW у межах 60-75%. Це було зумовлено їх низькою здатністю до генералізації в умовах складного фону та змін освітлення. З переходом до глибокого навчання точність систем стрімко зростає. Оригінальна архітектура DeepFace (2014 рік) вперше наблизилася до рівня людського сприйняття, досягнувши точності 97,35%. У нашому дослідженні розроблена система на базі архітектури Inception-ResNet-v1 досягла показника Accuracy на рівні 96,00% та площі під ROC-кривою (AUC) на рівні 0,9924. Хоча абсолютний світовий рекорд на датасеті LFW наразі належить моделям архітектури ArcFace (точність понад 99,8%), отримані нами результати

(AUC = 0,9924) є вкрай високими та підтверджують ефективність обраної математичної моделі. Незначне відхилення від максимальних показників SOTA-моделей пояснюється тим, що в даній роботі застосовано базовий метод порівняння за косинусною відстанню без додаткового використання метричних класифікаторів (наприклад, спільного баєсівського класифікатора – Joint Bayesian) та без аугментації тестових даних (Test Time Augmentation). Тим не менш, розроблена система забезпечує показник хибного допуску (FAR) на рівні 0,33%, що повністю задовольняє вимоги до комерційних систем контролю доступу.

Висновки до розділу 4

У четвертому розділі проведено експериментальне дослідження розробленої системи біометричної ідентифікації на відкритому наборі даних LFW-deepfunneled. Описано методику тестування алгоритмів у режимі верифікації осіб із використанням обчислювальних потужностей фреймворку PyTorch. За результатами ROC-аналізу визначено оптимальний поріг класифікації за косинусною схожістю, який склав 0,4571. Розрахунок метрик ефективності показав, що розроблена система досягла загальної точності (Accuracy) на рівні 96,00%, а площа під ROC-кривою (AUC) склала 0,9924, що свідчить про високу сепарабельність класів. Система продемонструвала високу надійність із показником хибного допуску (FAR) на рівні 0,33%. Теоретичний аналіз стійкості моделі підтвердив, що використання попереднього вирівнювання за антропометричними точками та метрики кутової відстані робить систему інваріантною до помірних змін освітлення та просторових зсувів. Порівняння з існуючими аналогами показало, що імплементоване рішення значно перевершує класичні алгоритми комп'ютерного зору і демонструє результати, порівнянні з передовими згортковими нейромережами, залишаючись при цьому обчислювально ефективним.

ВИСНОВКИ

У кваліфікаційній роботі розв'язано актуальну науково-практичну задачу розробки та програмної імплементації системи біометричної ідентифікації на основі аналізу зображень облич методами глибокого навчання. За результатами проведених теоретичних та експериментальних досліджень зроблено такі висновки: На основі аналізу предметної області встановлено, що перехід від класичних алгоритмів комп'ютерного зору до архітектур глибоких згорткових нейронних мереж забезпечив фундаментальний стрибок у точності систем розпізнавання облич. Обґрунтовано доцільність використання гібридної архітектури Inception-ResNet-v1, яка інтегрує механізми залишкових зв'язків та мультимасштабного вилучення ознак. Формалізовано математичний апарат метричного навчання. Доведено, що для оптимізації багатовимірного простору ознак традиційна функція перехресної ентропії (Softmax) є недостатньою. Для забезпечення високої щільності внутрішньокласових векторів та максимізації міжкласової дисперсії необхідно використовувати маржинальні функції втрат з проєкцією на одиничну гіперсферу, а мірою подібності дескрипторів має виступати метрика косинусної близькості. Спроектовано та програмно реалізовано модульну архітектуру системи на базі мови Python та фреймворку PyTorch. Програмний конвеєр успішно інтегрує каскадну нейронну мережу MTCNN для локалізації та геометричного вирівнювання облич (alignment) із попередньо навченим екстрактором ознак, що генерує стійкі 512-вимірні вектори-дескриптори. Застосування парадигми трансферного навчання (з використанням ваг VGGFace2) дозволило уникнути обчислювально складної оптимізації моделі "з нуля". Проведено експериментальне дослідження розробленого алгоритму класифікації на стандартизованому наборі даних Labeled Faces in the Wild (LFW-deepfunneled). За допомогою методів математичної статистики та максимізації індексу Юдена на базі ROC-аналізу було розраховано оптимальний поріг прийняття рішення, який становить 0,4571. Оцінка ефективності розробленої системи підтвердила її високу роздільну здатність. Площа під ROC-кривою (AUC) склала 0,9924. Загальна точність

класифікації (Accuracy) досягла 96,00%, точність позитивного класу (Precision) – 99,64%, а повнота (Recall) – 92,33%. Аналіз показників безпеки продемонстрував, що розроблена математична модель орієнтована на високий рівень захисту від несанкціонованого доступу: коефіцієнт хибного допуску (FAR) становить лише 0,33% при коефіцієнті хибної відмови (FRR) на рівні 7,67%. Математична нормалізація тензорів та розрахунок кутової відстані забезпечили інваріантність системи до помірних змін освітлення та міміки. Таким чином, поставлена мета роботи повністю досягнута, а всі визначені завдання успішно виконані. Розроблена програмна система може бути використана як базовий модуль для побудови комерційних систем контролю фізичного доступу, інтелектуального відеоспостереження та автоматизованої автентифікації користувачів.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Wang M., Deng W. Deep face recognition: A survey. *Neurocomputing*. 2021. Vol. 429. P. 215–244. DOI: 10.1016/j.neucom.2020.10.081.
2. Taigman Y., Yang M., Ranzato M., Wolf L. DeepFace: Closing the gap to human-level performance in face verification. 2014 IEEE Conference on Computer Vision and Pattern Recognition. 2014. P. 1701–1708. DOI: 10.1109/CVPR.2014.220.
3. LFW Dataset [Електронний ресурс]. Kaggle. URL: <https://www.kaggle.com/datasets/jessicali9530/lfw-dataset>
4. Schroff F., Kalenichenko D., Philbin J. FaceNet: A unified embedding for face recognition and clustering. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. P. 815–823. DOI: 10.1109/CVPR.2015.7298682.
5. Liu W., Wen Y., Yu Z., Li M., Raj B., Song L. SphereFace: Deep hypersphere embedding for face recognition. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. P. 6738–6746. DOI: 10.1109/CVPR.2017.713.
6. Wang H., Wang Y., Zhou Z., Ji X., Gong D., Zhou J., Li Z., Liu W. CosFace: Large margin cosine loss for deep face recognition. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018. P. 5265–5274. DOI: 10.1109/CVPR.2018.00552.
7. Zhang K., Zhang Z., Li Z., Qiao Y. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*. 2016. Vol. 23, No. 10. P. 1499–1503. DOI: 10.1109/LSP.2016.2603342.
8. Deng J., Guo J., Xue N., Zafeiriou S. ArcFace: Additive angular margin loss for deep face recognition. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019. P. 4685–4694. DOI: 10.1109/CVPR.2019.00482.
9. Wang F., Cheng J., Liu W., Liu H. Additive margin softmax for face verification. *IEEE Signal Processing Letters*. 2018. Vol. 25, No. 7. P. 926–930. DOI: 10.1109/LSP.2018.2822810.

10. Sun Y., Wang X., Tang X. Deep learning face representation from predicting 10,000 classes. 2014 IEEE Conference on Computer Vision and Pattern Recognition. 2014. P. 1891–1898. DOI: 10.1109/CVPR.2014.244.
11. Boutros F., Damer N., Kirchbuchner F., Kuijper A. ElasticFace: Elastic margin loss for deep face recognition. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2022. P. 1577–1586. DOI: 10.1109/CVPRW56347.2022.00169.
12. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. P. 770–778. DOI: 10.1109/CVPR.2016.90.
13. Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A. Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. P. 1–9. DOI: 10.1109/CVPR.2015.7298594.
14. Szegedy C., Ioffe S., Vanhoucke V., Alemi A. Inception-v4, Inception-ResNet and the impact of residual connections on learning. Proceedings of the AAAI Conference on Artificial Intelligence. 2017. Vol. 31, No. 1. DOI: 10.1609/aaai.v31i1.11231.
15. Cao Q., Shen L., Xie W., Parkhi O. M., Zisserman A. VGGFace2: A dataset for recognising faces across pose and age. 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). 2018. P. 67–74. DOI: 10.1109/FG.2018.00020.
16. Liu Z., Luo P., Wang X., Tang X. Deep learning face attributes in the wild. 2015 IEEE International Conference on Computer Vision (ICCV). 2015. P. 3730–3738. DOI: 10.1109/ICCV.2015.433.
17. Kaya M., Bilge H. Ş. Deep metric learning: A survey. Symmetry. 2019. Vol. 11, No. 9. Article 1066. DOI: 10.3390/sym11091066.