

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
БУДІВНИЦТВА І АРХІТЕКТУРИ**

Автоматизації і інформаційних технологій

(факультет)

Інформаційних технологій

(назва кафедри)

**ПОЯСНЮВАЛЬНА ЗАПИСКА
ДО КВАЛІФІКАЦІЙНОЇ ВИПУСКНОЇ РОБОТИ
НА ЗДОБУТТЯ ОСВІТНЬОГО СТУПЕНЯ БАКАЛАВР**

на тему:

«Автоматизована система розпізнавання облич з використанням
згорткової нейронної мережі»

Габро Олександр Олегович

(прізвище, ім'я та по батькові студента повністю)

Київ 2024 р.

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
БУДІВНИЦТВА І АРХІТЕКТУРИ**

Автоматизації і інформаційних технологій

(факультет)

Інформаційних технологій

(назва кафедри)

ЗАТВЕРДЖУЮ

Завідувач кафедри

к. т. н., доцент Гончаренко Т. А.

„___” _____ 2024 року

**ПОЯСНЮВАЛЬНА ЗАПИСКА
ДО КВАЛІФІКАЦІЙНОЇ ВИПУСКНОЇ РОБОТИ НА ЗДОБУТТЯ
ОСВІТНЬОГО СТУПЕНЯ БАКАЛАВР**

Автоматизована система розпізнавання облич з використанням
згорткової нейронної мережі

(назва)

Виконав:

Габро Олександр Олегович

(прізвище, ім'я та по батькові повністю)

122 Комп'ютерні науки

(спеціальність)

Інформаційні управляючі системи та технології

(освітня програма)

Групи КН-20-1

Керівник Гончаренко Т.А.

(прізвище та ініціали)

к.т.н., доцент

(вчене звання, науковий ступінь)

Ідентичність підтверджую

Київ 2024 р.

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
БУДІВНИЦТВА І АРХІТЕКТУРИ**

Факультет: Автоматизації і інформаційних технологій

Випускова кафедра: Інформаційних технологій

Освітній рівень: «бакалавр»

Спеціальність: 122 «Комп'ютерні науки»

Освітня програма: Інформаційні управляючі системи та технології

ЗАТВЕРДЖУЮ

Завідувач кафедри

к. т. н., доцент Гончаренко Т. А.

„___” _____ 2024 року

ЗАВДАННЯ

**ДО ВИКОНАННЯ АТЕСТАЦІЙНОЇ ВИПУСКНОЇ РОБОТИ НА
ЗДОБУТТЯ ОСВІТНЬОГО СТУПЕНЯ БАКАЛАВР**

Габро Олександр Олегович

(прізвище, ім'я та по батькові студента)

1. Тема роботи

Автоматизована система розпізнавання облич з використанням згорткової нейронної мережі

затверджена наказом ректора КНУБА № 433/2 від 29.02.2024

2. Керівник роботи Гончаренко Т. А., к. т. н., доцент

(прізвище, ім'я та по батькові, науковий ступінь, вчене звання)

3. Строк подання студентом роботи до захисту _____

4. Зміст пояснювальної записки за розділами:

Р. 1. Теоретичні засади дослідження систем розпізнавання облич

Р. 2. Методологічні засади побудови автоматизованої системи розпізнавання облич

Р. 3. Програмно-технічне рішення автоматизованої системи розпізнавання облич на основі геометричних характеристик обличчя та згорткової нейронної мережі

Р. 4. Розробка бізнес-плану впровадження автоматизованої системи розпізнавання облич на основі згорткової мережі

5. Графічний матеріал за розділами:

- P. 1. Точкові приклади, моделі та блок-схеми до теоретичної складової дослідження
- P. 2. Моделі-результати використання моделей та методів з першого розділу
- P. 3. Моделі, гістограми відповідно до розробленого програмного продукту, модель-архітектура програмного рішення, метрики точності/якості моделі.
- P. 4. Плани розвитку відповідно до бізнес-плану

6. Календарний план виконання роботи:

Види робіт та їх зміст	Дата виконання
Розділ 1	
Розділ 2	
Розділ 3	
Розділ 4	
Остаточне оформлення роботи	
Направлення роботи для перевірки на плагіат	
Попередній захист роботи на випусковій кафедрі	
Направлення роботи на рецензування	

7. Консультанти розділів атестаційної випускної роботи

Розділ	Прізвище, ініціали та посада консультанта	Перевірив	
		дата	підпис
Розділ 1	Гончаренко Т. А.	30.03.2024	
Розділ 2	Гончаренко Т. А.	25.04.2024	
Розділ 3	Долгополов С.Ю.	20.05.24	
Розділ 4	Долгополов С.Ю.	30.05.24	

8. Дата видачі завдання 29.02.2024

Зав. кафедри _____ (підпис) _____ (прізвище та ініціали)

Керівник _____ (підпис) _____ (прізвище та ініціали)

Здобувач _____ (підпис) _____ (прізвище та ініціали)

РЕЗЮМЕ (SUMMARY)

РЕЗЮМЕ (SUMMARY) до атестаційної випускної роботи Здобувач:	Габро Олександр Олегович Habro Oleksandr		
ЗВО	Київський національний університет будівництва і архітектури		
Тема (українською та англійською)	Автоматизована система розпізнавання облич з використанням згорткової нейронної мережі Automated face recognition system using convolutional neural network		
Освітній ступінь	Бакалавр		
Факультет	Автоматизації і інформаційних технологій		
Випускаюча кафедра	Інформаційних технологій		
Спеціальність	122 «Комп'ютерні науки»		
Освітня програма	Комп'ютерні науки		
Керівник	<u>Гончаренко Т. А.</u>		
Обсяг роботи:	пояснювальн а записка, стор.	розділів	слайди презентації
	99	4	19
Ключові слова:	Розпізнавання облич, згорткова нейронна мережа, класифікація, глибоке навчання		
Keywords:	Face recognition, convolutional neural network, classification, deep learning		

Робота присвячена розробці автоматизованої системи розпізнавання облич на основі згорткової нейронної мережі (CNN). Запропонована модель використовує сучасні методи попередньої обробки зображень та класифікації, забезпечуючи високу точність і надійність розпізнавання облич у різних умовах.

The research is devoted to the development of an automated face recognition system based on a convolutional neural network (CNN). The proposed model uses modern methods of image preprocessing and classification, providing high accuracy and reliability of face recognition in various conditions.

Здобувач: _____ / Олександр ГАБРО /

Керівник: _____ / Тетяна ГОНЧАРЕНКО /

“ ” _____ 2024 р.

ЗМІСТ

ВСТУП	8
ТЕОРЕТИЧНІ ЗАСАДИ ДОСЛІДЖЕННЯ СИСТЕМ РОЗПІЗНАВАННЯ ОБЛИЧ	
11	
1.1. Поняття систем розпізнавання облич	11
1.2. Теоретичні означення методів розпізнавання облич	20
1.3. Класифікація систем виявлення облич	34
Висновки до розділу 1	36
МЕТОДОЛОГІЧНІ ЗАСАДИ ПОБУДОВИ АВТОМАТИЗОВАНОЇ СИСТЕМИ	
РОЗПІЗНАВАННЯ ОБЛИЧ	37
2.1. Системи автоматичного адаптивного вирівнювання гістограм з обмеженням контрасту та адаптивною середньою подвійною гама-корекцією	37
2.2. Локальний екстрактор ознак двійкового шаблону з адаптивним порогом для систем розпізнавання облич	43
2.3. Системи розпізнавання облич з низькою похибкою на основі організації згорткової нейронної мережі та доповнення даних	48
Висновки до розділу 2	54
ПРОГРАМНО-ТЕХНІЧНЕ РІШЕННЯ АВТОМАТИЗОВАНОЇ СИСТЕМИ	
РОЗПІЗНАВАННЯ ОБЛИЧ НА ОСНОВІ ГЕОМЕТРИЧНИХ	
ХАРАКТЕРИСТИК ОБЛИЧЧЯ ТА ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ	56
3.1. Класифікація ієрархічного методу розпізнавання облич	56
3.2. Розробка автоматизованого алгоритму розпізнавання облич на основі згорткової нейронної мережі	59
3.3. Експериментальні випробування та оцінка розробленого алгоритму	63
Висновки до розділу 3	73
РОЗРОБКА БІЗНЕС-ПЛАНУ ВПРОВАДЖЕННЯ АВТОМАТИЗОВАНОЇ	
СИСТЕМИ РОЗПІЗНАВАННЯ ОБЛИЧ НА ОСНОВІ ЗГОРТКОВОЇ МЕРЕЖІ	75
4.1. Аналіз ринку та цільової аудиторії	75
4.2.	80
Оцінка витрат та планування бюджету	81
4.3. Стратегія впровадження та маркетинговий план	88
Висновки до розділу 4	92
ЗАГАЛЬНІ ВИСНОВКИ	94
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	96
ДОДАТОК А	102
ДОДАТОК В	103

ВСТУП

Розпізнавання обличчя є одним із найважливіших завдань в галузі комп'ютерного зору та біометричних технологій. Воно знаходить широке застосування в різних сферах, таких як безпека, контроль доступу, мобільні додатки, соціальні мережі та багато інших. За останні роки значний прогрес у цьому напрямі було досягнуто завдяки розвитку глибинного навчання та згорткових нейронних мереж (CNN), що дозволило значно підвищити точність і надійність систем розпізнавання облич.

Використання CNN у поєднанні з сучасними методами попередньої обробки зображень та вилучення ознак дозволяє ефективно вирішувати проблеми, пов'язані з різноманітністю освітлення, виразів обличчя та позицій голови. Ця робота присвячена розробці автоматизованої системи розпізнавання облич з використанням CNN, що має на меті забезпечити високу точність і надійність розпізнавання за різних умов.

Актуальність даного дослідження обумовлена постійним зростанням потреби в системах розпізнавання облич у різних сферах, таких як безпека, верифікація особистості та контроль доступу. З огляду на швидкий розвиток технологій і збільшення обсягу даних, існує необхідність у створенні ефективних алгоритмів, які можуть обробляти великі обсяги зображень та забезпечувати високу точність і надійність розпізнавання.

Аналіз останніх досліджень та наукових праць. У галузі розпізнавання облич значний внесок зробили як українські, так і зарубіжні вчені. Дослідження українських вчених, таких як О. В. Зінченко, О. С. Звенігородський та Т. М. Кисіль, демонструють ефективність використання згорткових нейронних мереж для підвищення точності класифікації зображень. Зарубіжні вчені, такі як J. Zhang, Z. Su, L. Liu, а також N. Lakshmi та M. P. Arakeri, які зробили значний внесок у розвиток технологій розпізнавання облич. Їхні дослідження зосереджені на вдосконаленні методів вилучення ознак та підвищенні стійкості алгоритмів до зовнішніх впливів, таких як освітлення та шум.

Метою кваліфікаційної роботи є розробка автоматизованої системи розпізнавання облич на основі згорткової нейронної мережі.

Виходячи з мети визначимо наступні завдання:

- 1) Визначити поняття систем розпізнавання облич;
- 2) Охарактеризувати методи розпізнавання облич;
- 3) Класифікувати системи виявлення облич;
- 4) Визначити системи автоматичного адаптивного вирівнювання гістограм з обмеженням контрасту та адаптивною середньою подвійною гама-корекцією;
- 5) Означити локальний екстрактор ознак двійкового шаблону з адаптивним порогом для систем розпізнавання облич;
- 6) Охарактеризувати системи розпізнавання облич з низькою похибкою на основі організації згорткової нейронної мережі та доповнення даних;
- 7) Класифікувати ієрархічний метод розпізнавання облич;
- 8) Розробити автоматизований алгоритм розпізнавання облич на основі згорткової нейронної мережі;
- 9) Виконати експериментальне випробування та оцінку розробленого алгоритму;
- 10) Проаналізувати ринок та цільову аудиторію до бізнес-плану;
- 11) Виконати оцінку витрат та планування бюджету бізнес-плану;
- 12) Розробити стратегію впровадження та маркетинговий план.

Об'єктом дослідження є системи розпізнавання облич, що використовують методи глибинного навчання та згорткові нейронні мережі.

Предметом дослідження є алгоритми та методи попередньої обробки зображень, вилучення ознак та класифікації, що використовуються в системах розпізнавання облич.

Методи дослідження. Для досягнення мети роботи використовуються методи глибинного навчання, зокрема згорткові нейронні мережі, а також методи попередньої обробки зображень (вирівнювання гістограм,

гамма-корекція) та алгоритми вилучення ознак (локальний бінарний паттерн, аналіз головних компонент).

Наукова новизна роботи полягає в розробці комплексного підходу до створення автоматизованої системи розпізнавання облич на основі згорткових нейронних мереж, який включає ефективні методи попередньої обробки зображень, вилучення ознак та класифікації. Запропоновані методи дозволяють підвищити точність та надійність розпізнавання облич за різних умов освітлення та інших зовнішніх впливів.

Практичне значення роботи полягає в можливості впровадження розробленої автоматизованої системи розпізнавання облич у різні сфери, такі як безпека, контроль доступу, мобільні додатки, що забезпечить підвищення рівня безпеки та зручності використання технологій. Запропоновані підходи можуть бути використані для покращення існуючих систем розпізнавання облич та розробки нових додатків, що потребують високу точність та надійність.

Структура та обсяг кваліфікаційної роботи. Робота складається з 99 сторінок, які включають: вступ, чотири розділи з висновками до кожного з них, загальні висновки, список використаних джерел, що налічує 44 найменувань, та 1 додатку; містить 16 таблиць та 32 рисунків.

РОЗДІЛ 1.

ТЕОРЕТИЧНІ ЗАСАДИ ДОСЛІДЖЕННЯ СИСТЕМ РОЗПІЗНАВАННЯ ОБЛИЧ

1.1. Поняття систем розпізнавання облич

Системи розпізнавання людських облич є найбільш поширеним біометричним методом, що використовується людьми, завдяки наочності представлення низки даних. Протягом усього життя, починаючи з раннього віку, ми значною мірою покладемося на цей метод у повсякденній взаємодії з іншими людьми. Коли ми зустрічаємо іншу людину, наші когнітивні здібності інстинктивно оцінюють низку параметрів, таких як вік, стать, колір шкіри, зріст, тип волосся тощо, щоб розпізнати її. Примітно, що риси обличчя людини відіграють ключову роль у цьому процесі. Варто зазначити, що існують численні випадки, коли ми помилково ідентифікуємо когось виключно на основі схожості його обличчя з іншою людиною, не звертаючи уваги на відмінності в інших характеристиках.

Технологія розпізнавання облич використовується в численних комерційних програмах як для верифікації, так і для ідентифікації. Розвиток інфраструктури, необхідної для цих систем, такої як відеоспостереження, комп'ютерні мережі та бази даних, значним чином сприяє усебічному розвитку комплексу методів, що стосуються систем розпізнавання облич. Крім того, зростаюча потреба в безпеці в цифровій сфері ще більше стимулює впровадження технології розпізнавання облич та закріплює розпізнавання людських облич перспективною темою в галузі обробки зображень, спонукаючи дослідників до постійного вдосконалення системи. Інтригуючим аспектом розпізнавання облич є широке коло дослідників, які займаються цією темою, зокрема у розвитку комп'ютерного зору та обробки зображень, нейронні мережі продемонстрували значні досягнення у вирішенні завдань розпізнавання облич та об'єктів.

Конволюційні нейронні мережі (англ. convolutional neural network, CNN) займають провідне місце у вирішенні завдань розпізнавання облич, завдяки своїй здатності ефективно обробляти і аналізувати великі обсяги даних з зображень. Дослідження українського колективу дослідників О. В. Зінченко, О. С. Звенігородського та Т. М. Кисіля демонструє, що використання CNN дозволяє значно підвищити точність класифікації зображень, зокрема при розпізнаванні облич, завдяки здатності алгоритмів розрізняти важливі ознаки в структурі облич [44]. Водночас, дослідники J. Zhang, Z. Su, L. Liu розробили мережу Median Pixel Difference Convolutional Network (MeDiNet), яка є релевантним прикладом, що демонструє покращену стійкість до шумів на зображеннях і здатна точно розпізнавати обличчя в реальному світі [42]. Відповідно до дослідження N. Lakshmi та M. P. Arakeri, оптимізована нейронна мережа може ефективно вирішувати «проблеми освітлення», що є ключовим аспектом при розпізнаванні облич [22]. В той же час, колектив вчених у складі I. Yildirim, M. Belledonne, W. Freiwald, J. Tenenbaum розробили модель, яка демонструє можливості нейромереж у біологічній обробці образів облич [41].

Таким чином, сучасні методи глибинного навчання, такі як конволюційні нейронні мережі, можуть ефективно аналізувати та розпізнавати обличчя в різних умовах освітлення.

На сьогоднішній день, розпізнавання облич відіграє ключову роль в галузях, що пов'язані з біометрикою та безпекою приватності, де технології штучного інтелекту та комп'ютерного зору стають в нагоді для забезпечення захисту особистих даних. Так, вже розглянуте дослідження про мережу Median Pixel Difference Convolutional Network (MeDiNet) підкреслює, що розвиток методів CNN, які включають стійкість до шуму, можуть покращувати надійність систем розпізнавання облич при різних умовах освітлення та зовнішнього впливу, забезпечуючи таким чином ефективніший захист в сфері мережевої безпеки. Ця теза також підкріплюється дослідженням вчених C. Song та S. Ji, які виявили, що нейромережі, спеціалізовані на обробці низької та високої частоти

зображень, можуть значно покращити точність розпізнавання облич в умовах, які зазвичай викликають помилки в менш адаптованих системах [37; 42].

У наш час стає все складніше зустріти людину, яка не пройшла біометричне фотографування для отримання візи або паспорта. Поширеність біометричних методів розпізнавання обличчя майже в кожному мобільному пристрої є прикладом практичного застосування розпізнавання обличчя, де методи слугують або для підтвердження особи, або для ідентифікації конкретного об'єкта за фотографією чи кількома кадрами відео. Однією з важливих характеристик розпізнавання облич є його безконтактний характер, що відрізняє його від інших біометричних методів, які зазвичай вимагають безпосередньої участі суб'єкта.

Більшість алгоритмів розпізнавання облич працюють за наступними базовими етапами: попередня обробка, виявлення обличчя, вилучення ознак і класифікація. Ці етапи зазвичай виконуються послідовно, за рідкісними винятками. На етапі попередньої обробки алгоритм усуває будь-які шуми, присутні на зображенні. Крім того, попередня обробка може бути використана для виправлення проблем з освітленням на зображенні та нормалізації його відповідно до вимог алгоритму. Таким чином, логічно включати попередню обробку на початкових етапах процесу.

У контексті попередньої обробки зображень у системах розпізнавання облич, особливу увагу заслуговує дослідження S. Sarkar, K. V. Ajitha Sheno, яке описує гібридний підхід до використання штучних нейронних мереж для розпізнавання облич. Автори розробили систему, що комбінує дві нейронні мережі для покращення точності розпізнавання: одну для вилучення ознак з використанням головних компонент (PCA-ANN); другу для лінійного дискримінантного аналізу (LDA-ANN). Ці мережі «співпрацюють» в ефективній нормалізації та корекції освітлення зображень задля виявлення облич у різних умовах освітлення [35].

Після попередньої обробки, виявлення обличчя відіграє вирішальну роль у зменшенні обчислювальної складності шляхом усунення нерелевантних

частин зображення обличчя. Цей крок включає в себе вилучення ознак, що є ключовим для створення вектора відмінних ознак. Ознаки, включені в цей вектор, повинні мати здатність дозволяти алгоритму точно розпізнавати кожную особу. Вивчення нових ознак, комбінування різних ознак і вибір найбільш характерних і стійких до змін ознак може підвищити ефективність етапу вилучення ознак.

Нарешті, на останньому етапі алгоритм використовує класифікатор, щоб класифікувати екземпляри в наборі даних на основі їхніх ознак, що потім використовується для віднесення невідомого вхідного зображення до певного класу.

Біометричні програми поділяються на дві основні категорії: ідентифікація та верифікація. Вибір методу розпізнавання обличчя має вирішальне значення для визначення ефективності програми. Так, верифікація фокусується на підтвердженні або відхиленні особистості шляхом порівняння вхідного зображення з однією заздалегідь визначеною ідентичністю в базі даних, що забезпечує високу точність і безпеку, особливо в контрольованих умовах, наприклад, при доступі до приватних аккаунтів або секретних матеріалів, в той час як ідентифікація передбачає пошук потенційних відповідностей у великій базі даних, що може включати мільйони облич, як це описано в роботі Н. Д. Joshi, де застосовуються системи багатофакторної біометрії, що інтегрує розпізнавання облич та відбитків пальців для забезпечення точної ідентифікації особи [18]. Рис. 1.1 ілюструє різницю між ідентифікацією та верифікацією.

Останнім часом технологія розпізнавання облич зазнала значного прогресу, що зробило її застосовною в широкому діапазоні сфер. Від комерційного до промислового секторів, від безпеки до урядових програм – потенційні можливості використання розпізнавання облич дуже широкі. До основних категорій, де технологія розпізнавання облич може бути ефективно використана, відносяться безпека і спостереження, відеоіндексація, бази даних зображень для розслідувань і перевірка особи в різних галузях промисловості.

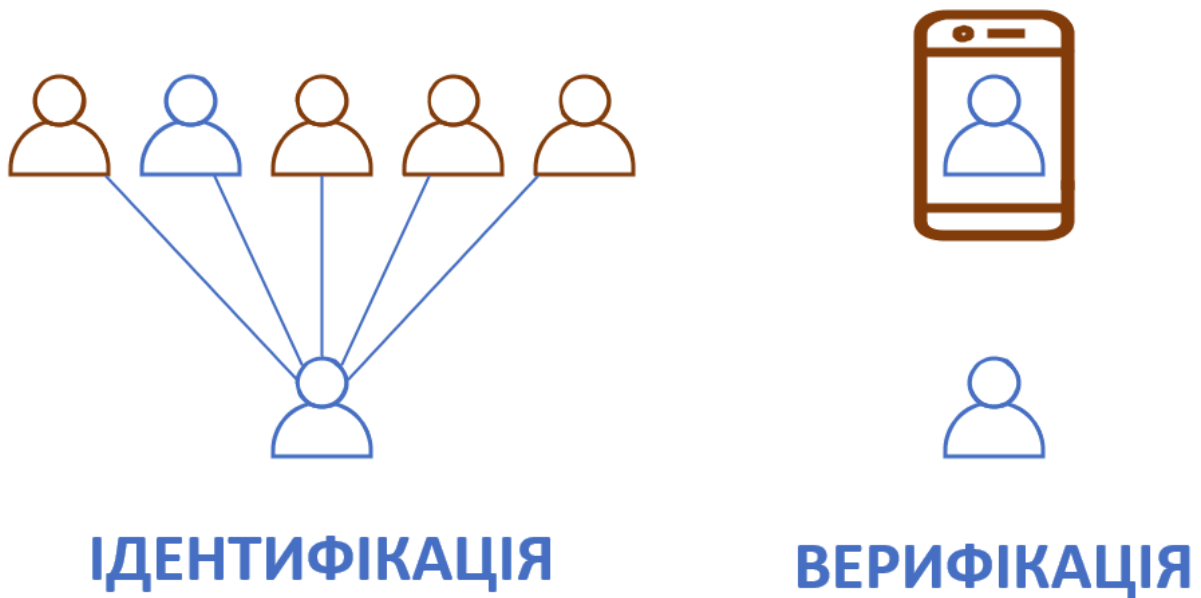


Рисунок 1.1. Модель протиставлення ідентифікації та верифікації

Алгоритми розпізнавання облич пройшли тривалий шлях від обмежених, низькоефективних та затратних технологій до досягнення прийняттого рівня точності. Тим не менш, досягнення такого рівня можливе, в першу чергу, в контрольованих середовищах, коли ж з'являються непередбачувані змінні, такі як положення голови, тіні або вираз обличчя, продуктивність системи ставиться під загрозу. Основні фактори, що сприяють цим небажаним варіаціям, включають: освітлення, позу голови, вираз обличчя, оклюзію та міжкласові подібності.

Сприйняття об'єктів людським оком полегшується завдяки збору та аналізу відбитого світла – процесу, що має вирішальне значення при створенні зображень, оскільки умови освітлення суттєво впливають на результат. Дослідження підкреслюють, що навіть незначні зміни в освітленні можуть мати глибокий вплив на результати роботи алгоритмів розпізнавання облич. Вивчення впливу освітлення на розпізнавання облич продемонстровано в дослідженні R. Gopalan, D. W. Jacobs, де розглядалися різні підходи до створення освітлення-нечутливих представлень облич та їх комбінації, які підвищують ефективність класифікації [16]. Дослідники протестували означені

представлення на наборі даних PIE та розширеному наборі даних Yale-B, де використовувалися як однорідні, так і неоднорідні умови освітлення, що підтвердило важливість адаптації алгоритмів під змінні умови освітлення для підвищення точності розпізнавання облич. Явище відбитого світла можна пояснити тривимірною природою зображень людського обличчя. Зміна налаштувань освітлення може призвести до утворення тіней на обличчі, що потенційно може приховати певні риси обличчя. Отже, втрата цих рис ставить під загрозу загальну ефективність системи. На рис. 1.2 показано різні варіанти освітлення.



Рисунок 1.2. Представлення різного освітлення облич людини

Розташування рис обличчя на зображенні також відіграє важливу роль в алгоритмі розпізнавання обличчя через те, що зображення представляє тривимірний об'єкт (обличчя) у двовимірній формі, де зміни в орієнтації обличчя можуть змінити геометрію зображення. Відповідно, ці зміни можуть

вплинути на характеристики зображення і призвести до зниження точності алгоритму. Хоча шаблони можуть пом'якшити цю проблему в певних сценаріях, таких як системи контролю доступу або станції перевірки паспортів, проте відсутність стандартизованого формату зображення обличчя створює проблему в багатьох інших ситуаціях, наприклад, у системах відеоспостереження. На рис. 1.3 показано різні положення голови.



Рисунок 1.3. Представлення різного положення голови людини

Вираз обличчя, варіації положення голови та зміни в освітленні – все це може впливати на вигляд обличчя. Важливо визнати, що навчальні бази даних часто піддаються нормалізації, що вимагає, щоб зображення зображували обличчя в природному стані. Однією зі стратегій боротьби з цими спотвореннями є вибір найбільш стійких рис обличчя. У дослідженні вчених F. Ayeche та A. Altı було описано підхід до розпізнавання облич та виразів обличчя, який базується на розширених дескрипторах Histogram Oriented Gradient (HOG) – дескрипторах, що дозволяють ефективно виділяти риси,

пов'язані з виразами облич, та класифікувати обличчя та їх вирази. Використані методи засновані на визначенні різних направлених векторів на зображенні обличчя, заснованих на реакції країв, що визначає вектор ознак зі зображення обличчя. Зменшення розміру цього вектора сприяє підвищенню продуктивності класифікатора SVM (Support Vector Machine). Експерименти, проведені на двох публічних наборах даних: JAFFE для розпізнавання виразів облич та YALE для розпізнавання облич, показали, що запропонований дескриптор досягає рівня розпізнавання 92.12% та часу виконання від 0.4 до 0.7 секунди на всіх оцінених базах даних порівняно з існуючими роботами [4]. На рис. 1.4 показано кілька прикладів таких виразів облич.



Рисунок 1.4. Представлення різних виразів обличчя людини

Перешкоди на зображенні обличчя можуть призвести до втрати певних характеристик обличчя та введення не пов'язаних з ним небажаних рис, наприклад, коли перешкодою є частина обличчя іншої людини. Технологія розпізнавання облич зіткнулася з перешкодами у вирішенні цієї проблеми,

особливо в останні роки, внаслідок вимог до масок, накладених пандемією. Багато підходів використовують стандартний метод розпізнавання закритої ділянки обличчя і видалення її з зображення, проте існують і інші можливості обробки закритих ділянок обличчя. Наприклад, одне з популярних досліджень В. Yang, J. Wu, G. Hattori, яке вивчало вплив масок на обличчя під час пандемії COVID-19, визначає новий двоступеневий підхід до розпізнавання емоцій з масками на обличчях. У першому етапі тренується двійковий глибокий класифікатор з/без маски, який генерує теплові карти уваги для приблизного відрізнення маскованих частин обличчя від незакритих ділянок. На другому етапі тренується класифікатор розпізнавання емоцій, який керується ділянками, важливим для класифікації виразів облич, враховуючи як закриті, так і незакриті ділянки обличчя, але з переважним зважуванням. Запропонований метод перевершує інші сучасні методи, які включають оклюзію, на наборах даних для розпізнавання виразів обличчя з урахуванням масок, що дозволяє алгоритму зосередитися на неушкоджених рисах обличчя. Приклади зображень з оклюзією представлені на рис. 1.5.



Рисунок 1.5. Представлення прикладів оклюзії на зображеннях обличчя

Завдання розрізнення двох людей, які мають схожі риси обличчя, наприклад, однойцевих близнюків або акторів, може становити труднощі для людини. Враховуючи, що люди покладаються на безліч ознак, окрім рис обличчя, щоб ідентифікувати особу, не дивно, що машини зіткнуться з більшими труднощами при розрізненні людей, які мають схожі характеристики в різних класах. По суті, суть питання полягає в здатності розрізнити двох окремих людей, які мають дещо схожі риси. Одне з досліджень, яке досліджує цю проблему – «Recognition of identical twins based on the most distinctive region of the face: Human criteria and machine processing approaches» авторів S. Mousavi, M. Charmi, H. Hassanpoor, використовує методику виявлення найвідмінніших регіонів обличчя однойцевих близнюків за допомогою зміненого SIFT алгоритму та за допомогою crowdsourcing, щоб визначити, які області обличчя найбільш відрізняються у близнюків [24]. Це дослідження показує, що крива обличчя є найбільш дискримінативною областю серед інших регіонів, які включають очі, брови, ніс, рот і саму форму обличчя.

1.2. Теоретичні означення методів розпізнавання облич

Як зазначалося раніше, алгоритм розпізнавання обличч зазвичай складається з чотирьох етапів. Блок-схема системи розпізнавання обличч зображена на рисунку 1.6. Початковий етап, відомий як попередня обробка, передбачає нормалізацію зображення, після цього відбувається етап розпізнавання обличчя, метою якого є ідентифікація рис обличчя на зображенні. Цей етап має вирішальне значення для зменшення обчислювальної складності. Етап вилучення ознак вважається основним процесом алгоритму розпізнавання, тут алгоритм шукає найбільш характерні риси зображення обличчя, які заздалегідь визначені в алгоритмі. Нарешті, невідомий екземпляр подається алгоритму, який відносить його до певного класу ознак за допомогою класифікатора алгоритму.

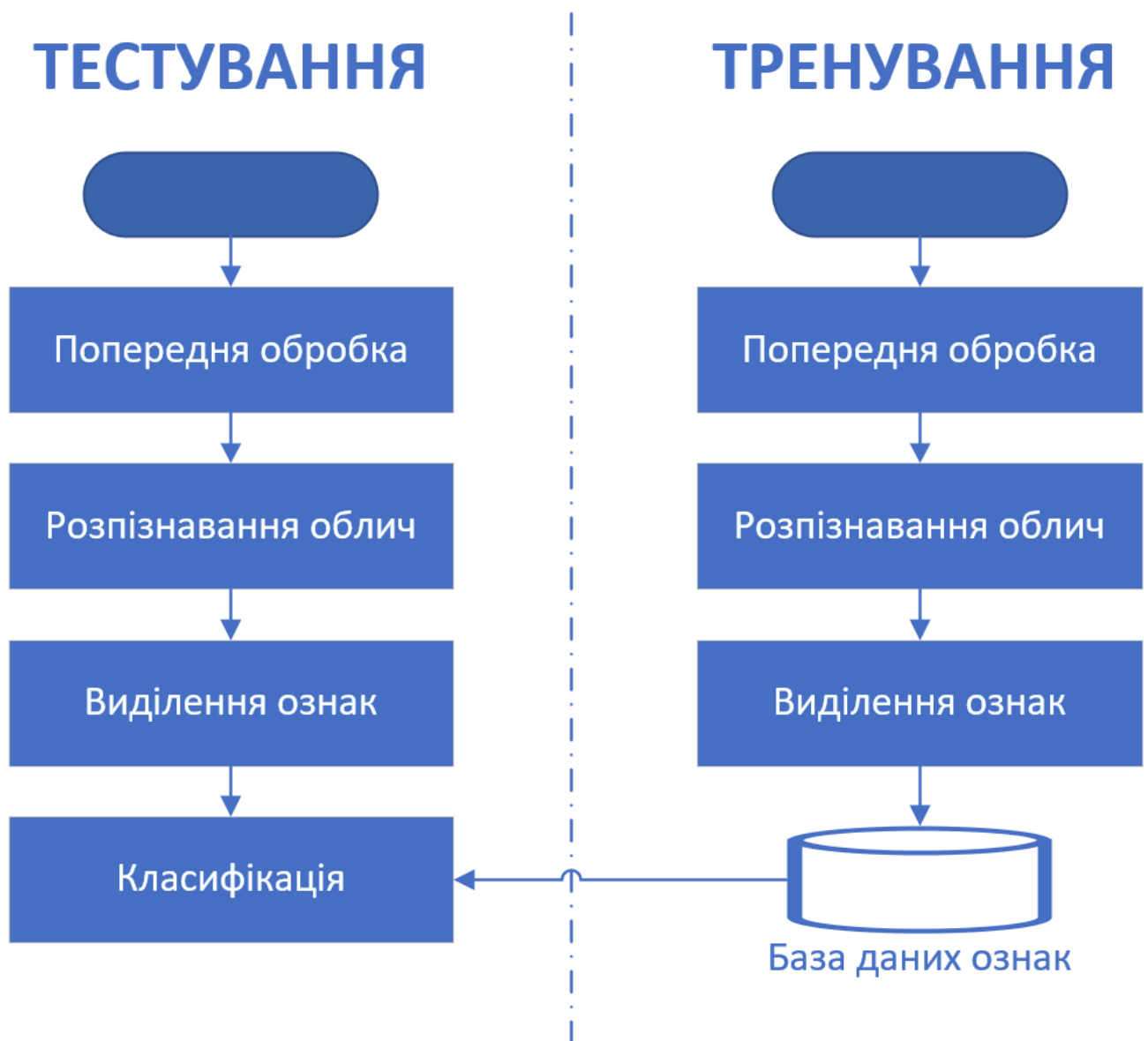


Рисунок 1.6. Блок-схема системи розпізнавання обличчя

Окрім етапу отримання зображення, початковим етапом більшості алгоритмів розпізнавання облич є попередня обробка. Зазвичай, в контексті різних додатків для обробки зображень цей етап вважається основним, зображення обличчя можуть містити цілий ряд спотворень, включаючи шум, нерівномірну експозицію, зміщене розташування обличчя, невідповідність колірних схем і багато іншого, що підкреслює важливість попередньої обробки як першого етапу алгоритму, особливо тому, що дуже важливо усунути ці спотворення, перш ніж переходити до будь-яких завдань навчання або

класифікації. Серед цих спотворень погане освітлення є одним з найбільш значущих. Нерівномірне освітлення і тіні на зображеннях облич часто є основними і найпоширенішими проблемами, які повинна вирішувати система розпізнавання облич. Тіні і темні ділянки на зображенні призводять до втрати життєво важливої інформації, тим самим знижуючи точність методу.

Відповідно до дослідження R. Gagan, H. Sanghavi, S. Ramachandran, K. Manikantan, вирівнювання гистограми (англ. Histogram equalization, HE) є одним із стандартних методів для вирішення проблем з освітленням у системах розпізнавання облич. Метод Circular Sector DCT based Feature Extraction розширює звичайну гистограмну еквалізацію та покращує процес нормалізації освітлення, оптимізуючи розпізнавання облич. Цей метод включає динамічну корекцію інтенсивності гамми, що дозволяє адаптувати значення кожного пікселя залежно від освітлення, забезпечуючи більш рівномірний і точний аналіз облич. Використання цієї техніки продемонструвало багатообіцяючі результати на базах даних, таких як Extended Yale B та ORL, покращуючи здатність систем до розпізнавання навіть за складних умов освітлення [14].

Gamma intensity correction (GIC) є ще однією популярною технікою для корекції проблем з освітленням у системах розпізнавання облич. Дослідження L. Zhuowen, W. Kejun, Z. Guofeng, Y. Lei запропонувало вдосконалений метод GIC, який використовується для покращення якості зображень облич за нерівномірного освітлення. В цьому дослідженні Gamma крива була оптимізована за допомогою побудови нелінійної функції, що дозволяє покращити адаптивні здібності GIC до компенсації освітлення. Результати показали, що такий підхід покращує якість зображень облич та зменшує спотворення, що є ключовим для забезпечення високої точності систем розпізнавання [43].

Вирівнювання гистограми – це метод попередньої обробки, який не залежить від інформації про спотворення зображення. Метод вирівнювання гистограми діє як глобальна функція для підвищення контрастності зображення, використовуючи розподіл інтенсивності пікселів, отриманий з гистограми

зображення. Популярність даного методу зумовлена простотою реалізації, однак ефективність методу знижується в сценаріях з екстремальними перепадами освітленості, оскільки він має тенденцію до усунення деталей на добре освітлених ділянках і вносить небажані артефакти та шуми.

Основною метою вирівнювання гистограми є згладжування гистограми зображення за допомогою кумулятивної функції розподілу, тим самим розширюючи динамічний діапазон зображення. Цей процес відбувається глобально у всіх частинах зображення, оскільки вирівнювання гистограми є глобальною функцією.

Гистограма вхідного зображення $I(x, y)$ з Q рівнями сірого визначається на основі функції ймовірності рівня сірого i , що представлена наступним рівнянням:

$$p(i) = \frac{m_i}{M}, \quad (1.1)$$

де $i \in \{0, 1, \dots, Q - 1\}$ – рівень сірого;

M – загальна кількість пікселів у зображенні.

CLANE, що розшифровується як Contrast Limited Adaptive Histogram Equalization – адаптивне гістограмне вирівнювання з обмеженим контрастом, є вдосконаленою версією гістограмного вирівнювання. На відміну від глобальної функції HE, CLANE працює з блоками зображення і обмежується точкою відсікання.

Такий поділ на блоки і точки відсікання допомагають мінімізувати вплив шуму на зображення. Однією з переваг CLANE є його здатність створювати більш реалістичне вихідне зображення, однак цей метод має обмеження щодо помірною покращення контрастності та нездатності генерувати значні відтінки сірого, коли це потрібно, крім того, CLANE малоефективний, коли стикається з гамма-спотворенням вхідного зображення.

Дослідження, проведене колективом вчених з Xidian University та Samsung SDS – Y. Chang, C. Jung, P. Ke, H. Song, J. Hwang, представило новий метод, що поєднує CLANE з подвійною гамма-корекцією, під назвою

«Автоматичне вирівнювання адаптивної гистограми з обмеженим контрастом і подвійною гамма-корекцією». Цей підхід має на меті не лише підвищити контрастність, але й зберегти природність зображень. Завдяки автоматичному встановленню точки відсікання на основі текстурі блоку та подвійній гамма-корекції цей метод ефективно підвищує яскравість у темних ділянках, одночасно контролюючи артефакти надмірного підсилення. Таким чином, адаптивна гистограма з обмеженим контрастом і подвійною гамма-корекцією є надзвичайно корисною для покращення зображень з сильними темними тінями або недоекспонованими ділянками, доводячи перевагу як у візуальній якості, так і в кількісних показниках порівняно з традиційними методами [9].

Гамма-спотворення впливає на яскравість зображень через певні обмеження та особливості фотографічного обладнання. GIC – це метод, який використовується для усунення цих спотворень, підвищуючи контрастність темних ділянок і стискаючи світлі ділянки зображення. В результаті GIC змінює освітлення зображення, зберігаючи загальний динамічний діапазон. Проблеми, з якими стикається ця техніка, включають складність визначення значення гамма-корекції та нездатність GIC впливати на тіні.

Нещодавнім досягненням у цій галузі є впровадження методу локальної гамма-корекції, який намагається усунути обмеження звичайних методів гамма-корекції. Цей метод особливо ефективний для покращення певних діапазонів інтенсивності на зображенні, не покращуючи при цьому інші області. Вчені А. Acharya, А. Giri запропонували адаптивний метод локальної гамма-корекції, в якому гамма-фактор змінюється локально відповідно до локальної середньої інтенсивності в околиці, що призводить до ефективної маніпуляції контрастом між різними класами інтенсивності на зображенні [1]. Цей підхід показав кращі результати як за якісними, так і за кількісними показниками порівняно з існуючими схемами покращення зображень, демонструючи свій потенціал у різних сферах застосування, що потребують детального аналізу та покращення зображень.

Нелінійне рівняння GIC представлено в наступній формулі:

$$I = I^{\frac{1}{\gamma}}, \quad (1.2)$$

де I – інтенсивність пікселів зображення;

$\frac{1}{\gamma}$ – ступінь зворотної гамма-корекції, що змінює спосіб регулювання інтенсивності для освітлення або затемнення зображення залежно від значення γ .

На рисунку 1.7 зображено криву корекції інтенсивності на основі значення гама. Можна помітити, що ця методика робить вхідний сигнал яскравішим для значень, більших за одиницю, і затемнює зображення для значень, менших за одиницю.

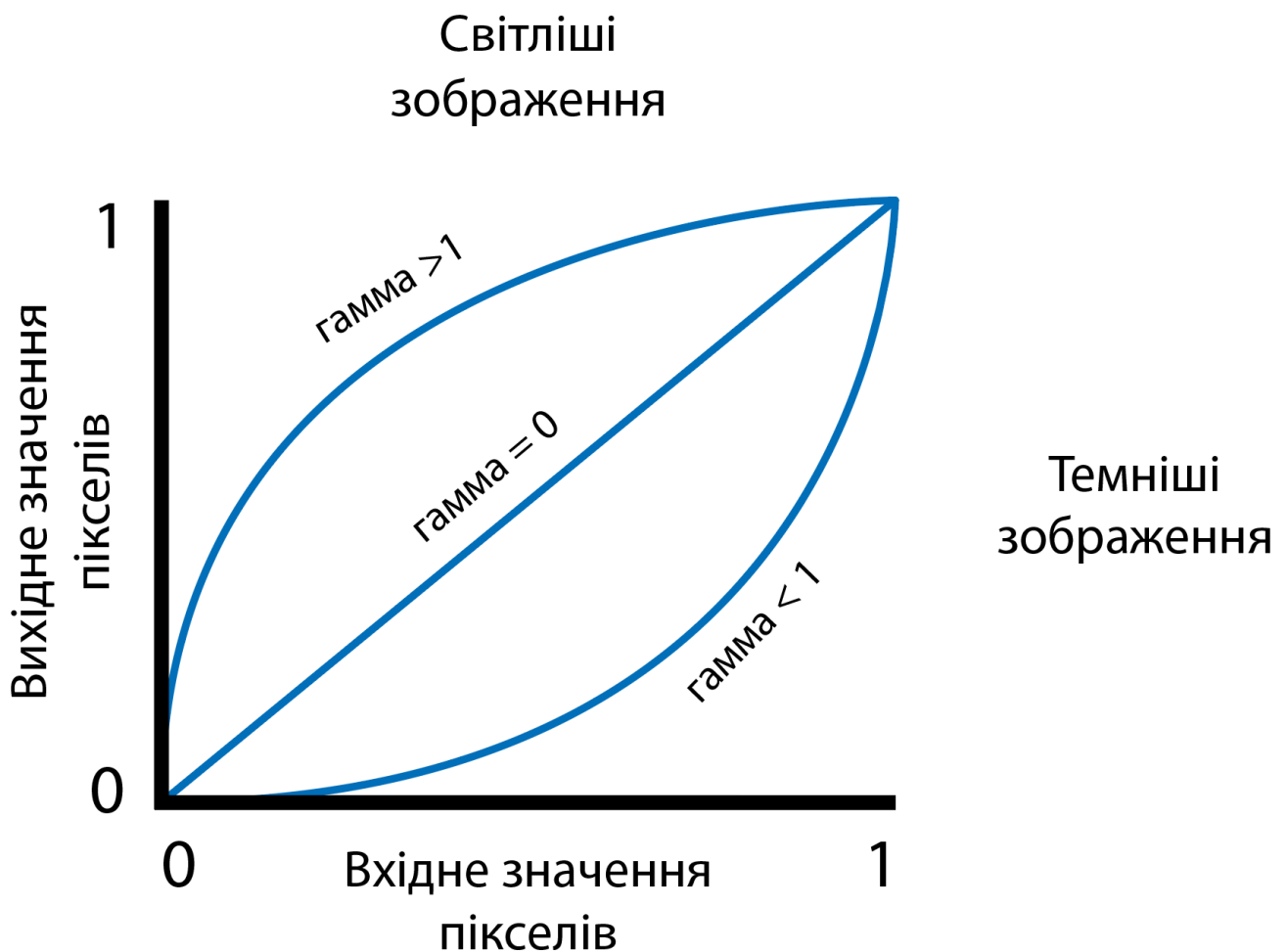


Рисунок 1.7. Блок-схема системи розпізнавання обличчя

Метод Difference of Gaussians (DoG) ефективно працює з тінями, присутніми на зображенні, тому його можна розглядати як фільтр, що

застосовується до вхідного зображення, для кращого розуміння цього методу. Основна мета цього фільтра – видалити високочастотні компоненти зображення, які часто сприймаються як шум, а також низькочастотні компоненти, які утворюють однорідні області на вхідному зображенні. Результат застосування цього фільтра можна візуалізувати у вигляді країв зображення. DoG передбачає згортку вхідного зображення двома окремими гауссовими ядрами, а різниця між згорнутими зображеннями використовується для генерації кінцевого вихідного зображення [20]. Однак у цього підходу є недоліки. По-перше, оскільки DoG обробляє все зображення одразу, необхідні додаткові кроки для корекції контрасту після застосування фільтра. По-друге, перед застосуванням фільтра необхідно виконати етап нормалізації гами, щоб уникнути впливу на контрастність затінених ділянок. В альтернативному поясненні метод DoG можна описати як віднімання двох розмитих версій оригінального зображення одна від одної. Ці розмиті зображення є результатом згортки вхідних даних з ядрами Гаусса з різними стандартними відхиленнями. Наступне рівняння демонструє операцію методу Difference of Gaussians (DoG):

$$DoG(x, y) = \frac{1}{2\pi\sigma_1^2} e^{-\frac{x^2+y^2}{2\sigma_1^2}} - \frac{1}{2\pi\sigma_2^2} e^{-\frac{x^2+y^2}{2\sigma_2^2}}, \quad (1.3)$$

де σ_1, σ_2 – ширина гауссового ядра.

Locally Tuned Inverse Sine Nonlinear (LTISN) виявляється високоефективним рішенням, особливо в складних умовах освітлення. Ефективність методу полягає в його здатності підвищувати контрастність і обмежувати динамічний діапазон, що досягається завдяки застосуванню нелінійної корекції інтенсивності та різних операторів в рамках методу. Впливаючи на окремі пікселі, LTISN покращує яскравість у темних ділянках, водночас зменшуючи освітленість у світліших регіонах. Крім того, метод включає в себе функцію, яка покращує локальний контраст.

Однак важливо визнати слабкість цього методу, яка проявляється у вигляді надмірного або недостатнього посилення в надмірно темних і яскравих областях, що може призвести до створення небажаних артефактів.

Виявлення облич передбачає визначення точного місця розташування обличчя на вхідному зображенні. Для цього алгоритм враховує масштаб та орієнтацію обличчя [21]. Виділення облич з відеоджерел стало важливим завданням в обробці зображень, що має на меті підвищити швидкість розпізнавання, щоб відповідати частоті кадрів. С. Р. Parageorgiou та ін. [29] представили методіку виявлення об'єктів на основі вейвлет-представлення та методів статистичного навчання. Е. Osuna та ін. використали метод опорних векторів Вапніка для розпізнавання облич у [28], а S. Romdhani та ін. вдосконалили його, ввівши скорочені набори навчальних векторів для свого класифікатора [32]. Н. Schneiderman та Т. Kanade застосували статистичний метод для виявлення профілю [36], але їхній підхід враховував лише три орієнтації обличчя, розглядаючи кожну орієнтацію як окремий об'єкт. Н. А. Rowley, В. Shumeet та Т. Kanade використовували фільтри на основі нейронних мереж [33], тоді як F. Fleuret та D. Geman спробували застосувати підхід «від грубого до тонкого» для виявлення облич [13]. Їх основна увага була зосереджена на мінімізації обчислювальних вимог. Дослідник Y-Q. Wang використовував концепцію інтегрального зображення, а також прямокутне представлення ознак і алгоритм бустінгу як метод навчання [39]. Він досяг значних результатів у виявленні облич зі швидкістю 15 кадрів на секунду, ефективно скоротивши час обчислень при збереженні високої швидкості виявлення.

Раніше вже обговорювалися численні потенційні застосування розпізнавання облич у реальних ситуаціях, починаючи від спостереження і безпеки і закінчуючи взаємодією людини з комп'ютером і збором даних. Ефективність цих застосувань значною мірою залежить від якості вхідних даних і здатності алгоритму справлятися з поганою якістю вхідних даних. Критично важливим аспектом процесу проектування є вибір найбільш

оптимального та ефективного методу вилучення ознак. Ці методи можна умовно розділити на два типи. Перший тип передбачає вилучення геометричних характеристик обличчя, зосереджуючись на його структурі та пропорціях. Цей метод використовує таку інформацію, як розмір і положення компонентів обличчя, таких як очі, рот, губи і ніс, а також співвідношення між цими компонентами. Другий тип, відомий як цілісний метод, охоплює два основні підходи [19]. Методи підпростору на основі зовнішності, такі як аналіз головних компонент (PCA) [11], лінійний дискримінантний аналіз (LDA) [38] та сингулярна декомпозиція (SVD) [7], спрямовані на зменшення розмірності. Підходи до цілісного виділення ознак другої категорії використовують частотні перетворення, такі як дискретне косинусне перетворення (DCT) [5] або вейвлети Габора [10], щоб захопити унікальні особливості зображення. Ці методи використовують коефіцієнти частотної області для представлення характеристик зображення, причому на вейвлет Габора негативно впливає висока розмірність просторових частотних характеристик. Методи зменшення розмірності в підпросторових підходах можуть допомогти вирішити цю проблему.

Концепція PCA полягає в тому, що напрямки з більшими варіаціями або розсіюваннями містять найбільше інформації про клас. Іншими словами, ці напрямки відображають найбільш суттєві відмінності між зображеннями. Для реалізації PCA першим кроком є створення коваріаційної матриці C на основі навчального набору з N зображень. Це передбачає знаходження середнього значення фотографій та віднімання цього середнього значення, μ , від усіх зображень. Отримана коваріаційна матриця потім використовується для визначення головних компонент і побудови власного простору:

$$C = \frac{1}{n} \sum_1^N (x_i - \mu)(x_i - \mu)^T, \quad (1.4)$$

Для M як найбільших власних значень C маємо:

$$CT_i = \lambda_i T_i \quad i = 1, \dots, M, \quad (1.5)$$

де, T_i – власні вектори або власні поверхні.

Власні значення коваріаційної матриці відсортовані у відповідній послідовності. Щоб зменшити розмірність підпростору тестових та навчальних даних, використовуються власні вектори, пов'язані з кількома числами найбільших власних значень.

Розкладання за сингулярними значеннями (SVD) є популярним інструментом в обробці сигналів і статистичному аналізі даних для зменшення розмірності. Сингулярні значення матриці даних надають цінну інформацію про рівень шуму, розподіл енергії, ранг матриці та інші ключові аспекти. Оскільки сингулярні вектори матриці слугують основою для матричного проміжку і є ортогональними, вони інкапсулюють значну частину характеристик сигналу. Використовуючи метод SVD, ми можемо отримати числові атрибути зображення:

$$X = USV^T, \quad (1.6)$$

де, S – діагональна матриця сингулярних значень розміром $n \times n$ з компонентами $s_{ij} = 0, i \neq j$ і $s_{ij} > 0$, а $U(m \times n)$ і $V(n \times n)$ є ортогональними матрицями. Можна також продемонструвати, що невиключні матриці U і V існують такі, що $s_1 \geq s_2 \geq \dots \geq 0$. Лівий і правий сингулярні вектори є стовпчиками ортогональних матриць U і V , відповідно.

Той факт, що U і V взаємно ортогональні, є важливою особливістю. Основною характеристикою підходу, що має відношення до обробки зображень обличчя, є стійкість SVD на зображенні обличчя. Стабільність SVD при обробці зображень є основною характеристикою методу, яка робить його привабливим [17]. Завдяки цим характеристикам багато дослідників вважають SVD надійним інструментом виділення ознак [23]. Базова структура Singular Value Decomposition представлена на рис. 1.8.

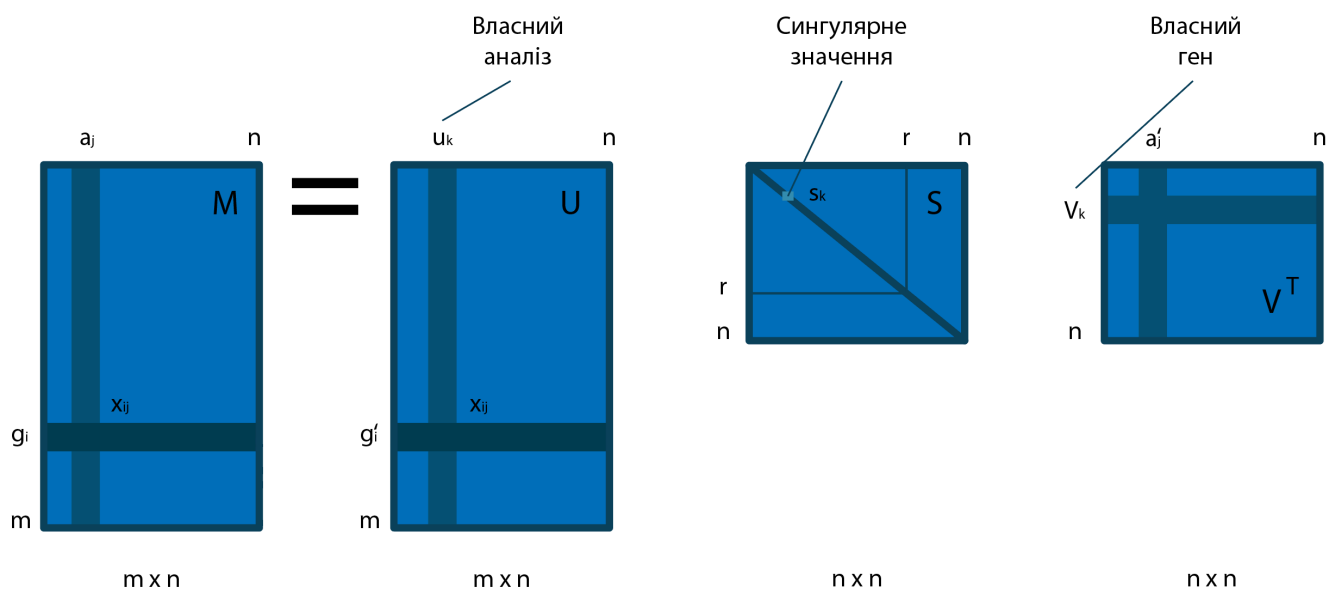


Рисунок 1.8. Базова структура Singular Value Decomposition

Ідея згорткової нейронної мережі походить від роботи зорової кори головного мозку [26]. CNN поділяють схожі етапи навчання зорової кори. У зоровій системі людини кожна зона зору пов'язана з певною групою клітин мозку. Доведено, що різні візуальні особливості, такі як вертикальні та горизонтальні краї, активують окремі групи нейронів, пов'язані з цією специфічною особливістю.

CNN – це тип нейронних мереж, який виявився особливо успішним у розпізнаванні та класифікації зображень. Штучні нейронні мережі – це нейронні мережі прямого поширення з великою кількістю шарів. Вони складаються з фільтрів, ядер або нейронів з вагами, параметрами та упередженнями, які можна навчати. Кожен фільтр приймає набір вхідних даних, проводить згортку і, за бажанням, додає нелінійність до суміші [15]. На етапі навчання зворотне поширення підтримує алгоритм у навчанні, налаштовуючи ваги на основі вхідних даних.

Архітектура CNN, представлена на рис. 1.9 – типова архітектура штучної нейронної мережі, що складається з різних типів шарів, проте ці шари є згортковими, об'єднують ReLU і повністю пов'язані між собою.

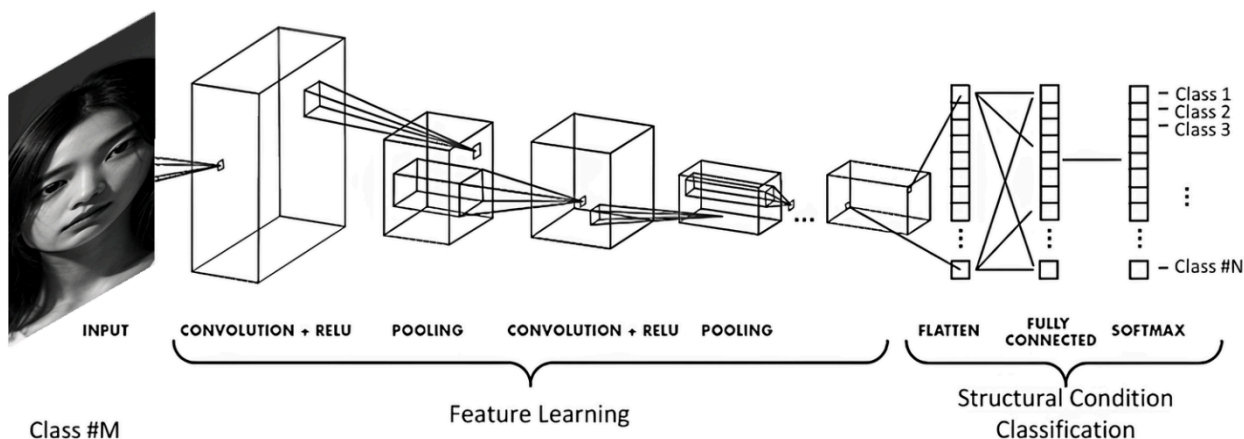


Рисунок 1.9. Архітектура згорткової нейронної мережі

Базовим компонентом побудови згорткової мережі є згортковий шар, виконує більшу частину обчислювальної роботи. Основна мета шару згортки – виділити ознаки з вхідних даних зображення. Вивчаючи атрибути зображення за допомогою крихітних квадратів вхідного зображення, згортка зберігає просторовий зв’язок між пікселями. Набір нейронів, що навчаються, використовується для затемнення вхідного зображення. У вихідному зображенні це створює карту ознак або карту активації. Згортковий шар отримує її як вхідні дані на наступному рівні. На рис. 1.10 представлено процес згортки.

згортковими шарами, забезпечує покращене узагальнення і швидшу збіжність, а також є стійким до перекладу і спотворення.

Кожен фільтр попереднього шару з'єднаний з кожним фільтром наступного шару, тому такий шар називається повністю з'єднаним шаром (FCL). Виходи шарів згортки, об'єднання та ReLU є представленнями високорівневих характеристик вхідного сигналу.

ReLU – це нелінійна операція, яка складається з блоків на основі випрямлячів. Це поелементна операція, тобто вона застосовується до кожного пікселя і замінює всі від'ємні значення на карті ознак на нуль. Щоб зрозуміти, як працює ReLU, припустимо, що на вхід нейрона подається x , і тоді випрямляч задається для нейронних мереж як $f(x) = \max(0, x)$.

Простота і лінійна поведінка ReLU робить її більш бажаною, ніж інші функції активації, такі як сигмоїд і тангенс. На рис. 1.11 зображено приклад роботи функції ReLU.

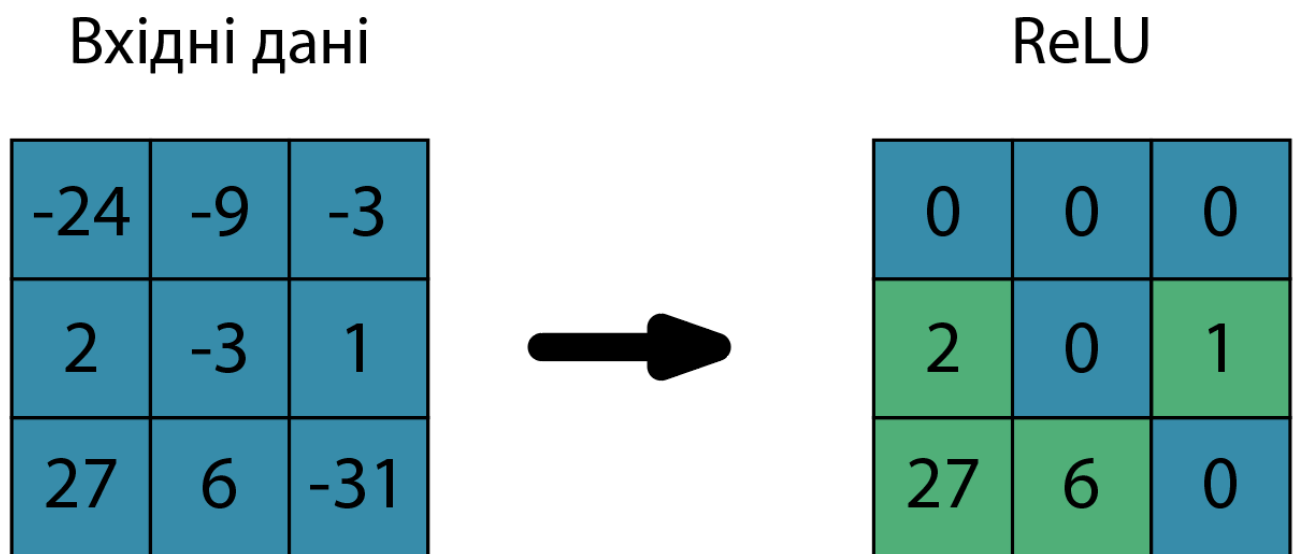


Рисунок 1.11. Приклад ReLU функції

1.3. Класифікація систем виявлення облич

Після отримання відповідних векторів для всіх навчальних вибірок потрібен класифікатор для створення класів на основі ознак. Зокрема, популярними класифікаторами є: нейронні мережі, Support vector machines (SVM), Hidden Markov Model.

Штучні нейронні мережі (ШНМ) широко використовуються як класифікатор у багатьох програмах розпізнавання образів і надзвичайно часто зустрічаються в системах розпізнавання облич, оскільки структура біологічних нейронних мереж надихає їх. Набір штучних нейронів об'єднується разом для побудови мережі нейронів, яка імітує біологічні мережі мозку в штучній нейронній мережі. Дослідники розробили різноманітні топології нейронних мереж, включаючи багат шарові перцептрони, нейронні мережі зі зворотним поширенням та нейронні мережі з радіально-базисною функцією (RBF), серед інших. Архітектура динамічних зв'язків (DLA) – це підхід на основі нейронних мереж, який використовує інформацію, зібрану з фотографій обличчя, для динамічного розташування нейронів у вищі порядки, що призводить до більш надійної ідентифікації обличчя, незважаючи на коливання навколишнього середовища [40]. Приклад нейронної мережі – тришаровий перцептрон – зображений на рис. 1.12.

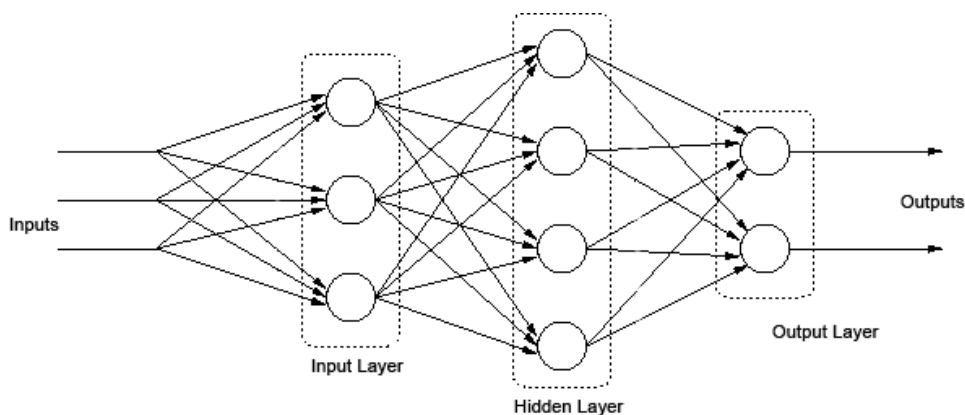


Рисунок 1.12. Тришаровий перцептрон

Машини опорних векторів (SVM) характеризуються як класифікатори з максимальним запасом, оскільки вони зменшують похибку класифікації,

одночасно збільшуючи геометричний запас [34]. SVM створює роздільну гіперплощину в просторі ознак, яка максимізує відстань між наборами даних. Для визначення відриву будуються дві паралельні гіперплощини, по одній по обидва боки від розділювальної. Ці гіперплощини накладаються на два набори даних, дозволяючи гіперплощині з найбільшою відстанню до сусідніх опорних векторів обох класів досягти пристойного розділення, як показано на рисунку 1.13. Чим більша різниця або відстань між цими паралельними гіперплощинами, тим більша ймовірність того, що невідомі зразки будуть точно ідентифіковані [3].

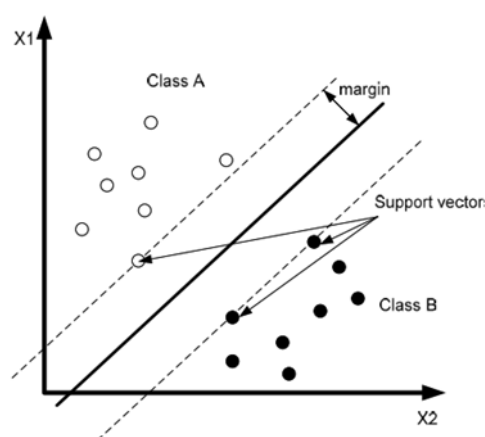


Рисунок 1.13. Гіперплощина з максимальною маржинальністю для SVM-класифікатора.

Як подвійний стохастичний процес, прихована марковська модель може ефективно представляти виробництво послідовних даних. Розпізнавання мови, обличчя та рукописного тексту ефективно реалізовано за допомогою НММ. Коли мова йде про використання НММ як класифікатора, існує два основних методи. У першому методі всі класи моделюються одним НММ, які дискримінуються на основі різних шляхів. У другому методі, відомому як модельний дискримінантний підхід, для кожного класу використовується окрема НММ. 2D вбудований НММ А. V. Nefian [25] має набір суперстанів, кожен з яких пов'язаний з набором вбудованих станів. Суперстани представляють основні області обличчя, тоді як вбудовані стани всередині

кожного суперстану пояснюють пов'язану з ним область обличчя більш глибоко. Усім навченим НММ у кожному класі подається послідовність спостережень невідомого зображення. Кожен НММ генерує умовну ймовірність, і НММ з найвищою ймовірністю визначає клас людини.

Висновки до розділу 1

Отже, перший розділ дипломної роботи присвячений теоретичним засадам дослідження систем розпізнавання облич, де детально розглянуто основні поняття та класифікації систем розпізнавання, а також сучасні методи їх реалізації. В розділі наведено опис роботи згорткових нейронних мереж (CNN), які демонструють високу ефективність у вирішенні завдань розпізнавання облич завдяки своїй здатності обробляти великі обсяги зображень та виявляти ключові ознаки облич.

Розглянуто основні етапи алгоритму розпізнавання облич, включаючи попередню обробку, виявлення облич, вилучення ознак та класифікацію. Окрему увагу приділено методам покращення якості зображень, таким як вирівнювання гістограм, гамма-корекція та методи адаптивної обробки зображень.

Також було проаналізовано основні виклики, з якими стикаються системи розпізнавання облич, зокрема вплив освітлення, положення голови, виразів облич та оклюзій. Наведено сучасні дослідження та підходи до вирішення цих проблем, які дозволяють підвищити точність та надійність систем.

На основі проведеного аналізу можна зробити висновок, що розвиток систем розпізнавання облич значно прогресує завдяки вдосконаленню алгоритмів глибинного навчання, що дозволяє ефективно застосовувати ці технології в різних сферах, таких як безпека, біометрія та комерційні додатки.

РОЗДІЛ 2.

МЕТОДОЛОГІЧНІ ЗАСАДИ ПОБУДОВИ АВТОМАТИЗОВАНОЇ СИСТЕМИ РОЗПІЗНАВАННЯ ОБЛИЧ

2.1. Системи автоматичного адаптивного вирівнювання гістограм з обмеженням контрасту та адаптивною середньою подвійною гама-корекцією

Розпізнавання облич стало одним з найпопулярніших методів у біометрії завдяки таким його властивостям, як ненав'язливість, відсутність необхідності участі людини та технологічність. Хоча нещодавні вдосконалення технології дозволили досягти прийнятної рівня точності, все ще залишається місце для вдосконалення, особливо при вирішенні проблем, спричинених неконтрольованими ситуаціями. Однією з таких проблем, яка може вплинути на точність системи, є зміна освітленості на зображенні. Зазвичай це спричинено такими факторами, як погані сенсори камери, неправильні джерела світла, цикл затвора та погане навколишнє освітлення. Відкинуті тіні, переекспоновані ділянки, занадто темні або занадто яскраві зображення можуть знизити швидкість розпізнавання системи, погіршуючи інформацію та характеристики зображення. Один із способів подолати ці варіації та спотворення – відкоригувати вхідне зображення перед тим, як подавати його в систему. Цей крок називається попередньою обробкою, і його можна знайти майже в кожній програмі для обробки зображень як одну з найважливіших частин. Методи попередньої обробки поділяються на три основні групи: 1) методи на основі нелінійної передавальної функції. 2) Методи корекції гістограм. 3) методи в частотній області. Нелінійні методи здебільшого модифікують окремі значення пікселів зображення. Основними характеристиками цих методів є простота використання та налаштування, що робить їх ідеальними для корекції контрасту. Широко використовуваним прикладом нелінійних методів є гамма-корекція інтенсивності. Гамма-корекція працює з дискретними значеннями зображення для покращення освітлення на темних зображеннях.

Методи гістограмної корекції, з іншого боку, прагнуть виправити контрастність зображення шляхом уніфікації рівнів сірого по всьому зображенню. Чудовою властивістю методів гістограмного вирівнювання є те, що вони вимагають відносно низьких обчислювальних витрат. Гістограму зображення можна загалом визначити як співвідношення рівнів сірого та частот. Наступне рівняння демонструє гістограму зображення у відтінках сірого:

$$H(i) = \frac{P_i}{N}, \quad i = 0, 1, \dots, L - 1, \quad (2.1)$$

де i визначає рівень сірого зображення;

P_i вказує на кількість пікселів цього рівня;

N – загальна кількість пікселів на зображенні.

Вирівнювання гістограми для зображення, засноване на визначенні $H(i)$ має вигляд:

$$HE_m = M(j) = (L - 1) \sum_{i=0}^j H(i), \quad (2.2)$$

де $M(j)$ визначає поняття функції відображення;

L – динамічний діапазон зображення.

Метод гістограмного вирівнювання стикається зі своїм основним недоліком на зображеннях з помітними і відносно великими однорідними темними областями, де метод НЕ має тенденцію показувати надмірне посилення на виході. Рис. 2.1 демонструє цю поведінку як приклад.



Рисунок 2.1. Використання HE методу: а) оригінальне зображення; б) зображення оброблене методом HE

Таким чином, на світлих ділянках спостерігається надмірне підсилення, що призводить до втрати текстури. Крім того, недостатнє підсилення видно на темних ділянках, особливо якщо вони прилягають до добре освітлених частин обличчя. Це особливо негативно впливає на швидкість розпізнавання системи розпізнавання обличчя, усуваючи частину цінної інформації, прихованої в обрізаних текстурах зображення обличчя.

Було докладено певних зусиль для подолання цих проблем. Деякі з цих удосконалень вводять різні розподіли гістограми для зображення, такі як вирівнювання гістограми зі збереженням яскравості (bi-histogram equalization, BBHE), вирівнювання гістограми дуалістичного підзображення рівної площі (sub-image histogram equalization, DSIHE), вирівнювання гістограми з мінімальною середньою похибкою яскравості (minimum mean brightness error bi-histogram equalization, MMBEVHE) та модель гауссової суміші (Gaussian mixture model, GMM). На противагу цьому, деякі інші методи намагаються модифікувати інші аспекти традиційного ВО, такі як зміна типу досліджуваної

інформації на просторовий, використання подвійної ймовірності тощо. Прикладами таких методів є двостороння крива Безьє (bilateral Bezier curve, BBC) та контекстне посилення контрасту.

Інші методи покращення традиційного НЕ пов'язані з сегментацією вхідного зображення. Найпомітнішим недоліком НЕ є погіршення характеристик зображення з відносно великими темними або світлими ділянками. Поділ зображення на різні сегменти та області допомагає запобігти цій проблемі, розбиваючи ці області. Крім того, робота з контрастом, пов'язаним з конкретною ділянкою, дає алгоритму кращий контроль над зображенням порівняно з однорідним контрастом на вході. Деякі приклади цих методів – експонетричне вирівнювання гістограми підзображення (ESIHE), вирівнювання гістограми підзображення на основі середнього значення (MMSICHE), селективне динамічне вирівнювання гістограми (SSDHE) та сегментно-залежне динамічне вирівнювання мультигістограми (SDDMHE). Основна відмінність між цими методами полягає в тому, як вони вирішують проблему розподілу та корекції посилення контрасту.

Іншим важливим методом підвищення контрастності є адаптивне гістограмне вирівнювання з обмеженою контрастністю (CLANE), яке ефективно вирішує проблему надмірного посилення при гістограмному вирівнюванні. Основна ідея CLANE полягає в тому, щоб розбити зображення на рівні підсекції і виконати ЕГ на кожній з цих секцій. У багатьох дослідженнях CLANE застосовували для покращення контрастності зображення, зберігаючи при цьому інформацію та уникаючи надмірного підсилення. Однак, у випадках, коли темні ділянки містять більшу частину зображення, CLANE не здатний суттєво покращити загальну яскравість зображення. Цей недолік, однак, можна усунути, якщо застосувати нелінійну техніку, таку як гамма-корекція, на виході CLANE. Поєднання цих двох методів дозволяє підвищити загальну освітленість зображення, зберігаючи при цьому текстурну інформацію в темних і світлих ділянках. Такі методи, як адаптивна гамма-корекція з розподілом ваг (AGCWD), адаптивна гамма-корекція (AGC) та комбінація AGC і вирівнювання

бігістограми з обмеженим діапазоном (RLBBHE), використовують принцип поєднання двох методів. Проте на зображеннях з різною освітленістю ефективність методів адаптивної гамма-корекції викликає сумніви. Y. Chang та ін. [9] запропонували комбінацію CLANE з подвійною гамма-корекцією, але оскільки запропонована методика призначена для покращення зображення для візуальних цілей, тонкість загальної функції не дозволяє їй бути ідеальною для покращення освітленості зображення для цілей розпізнавання облич.

Основна відмінність між адаптивним вирівнюванням гистограми з обмеженням контрастності та звичайним вирівнюванням гистограми полягає в методі, який CLANE використовує для обмеження точки відсікання. На рис. 2.2 зображено загальні етапи CLANE.

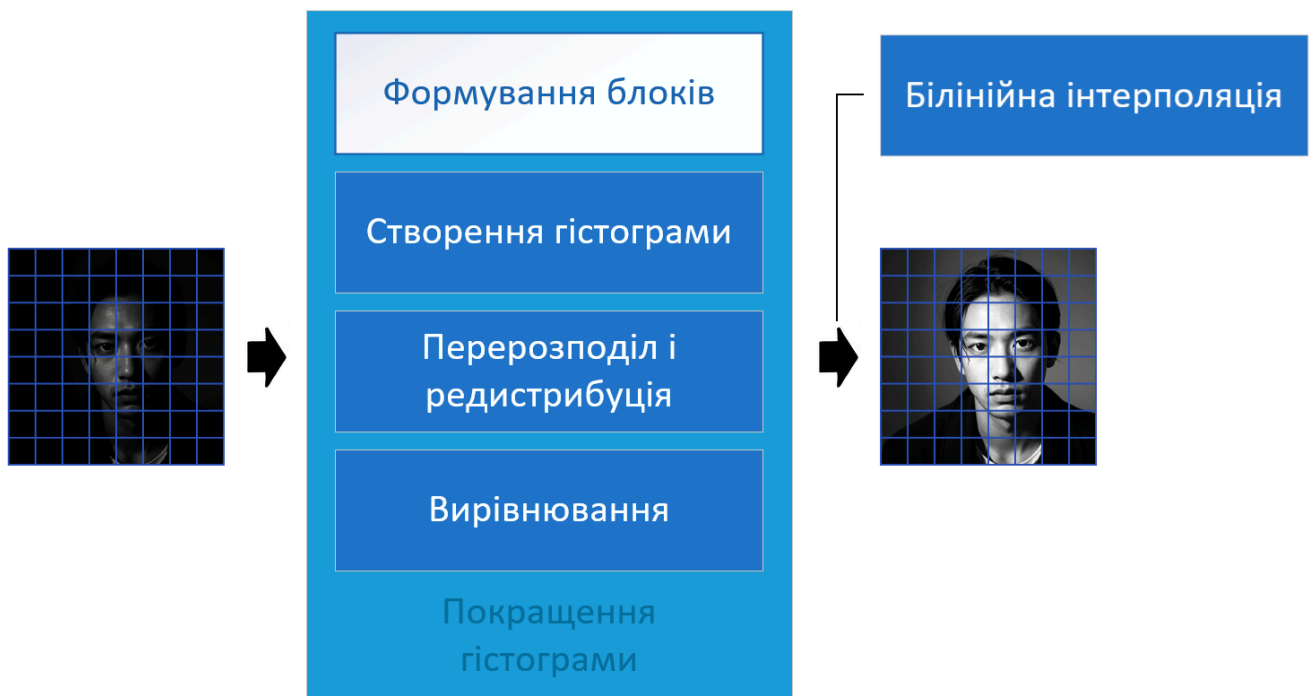


Рисунок 2.2. Етапи CLANE

Як видно з рисунка, CLANE складається з п'яти основних етапів. На першому етапі алгоритм ділить зображення на рівні блоки. Потім виконується покращення гистограми для кожного блоку окремо. Покращення гистограми складається з трьох етапів: створення гистограми, обрізання гистограми та перерозподіл обрізаної частини. Останнім етапом у CLANE є білінійна

інтерполяція, яка допомагає зменшити артефакти між блоками і згладити вихідні дані. Точка відсікання пікселів у CLАНЕ виглядає наступним чином:

$$\beta = \frac{N}{D} = \left(1 + \frac{\alpha}{100} G_m\right), \quad (2.3)$$

де N – кількість пікселів у блоці;

D – динамічний діапазон блоку;

α – коефіцієнт відсікання;

G_m – максимальний градієнт.

Таким чином, α визначає рівень підвищення контрастності. Чим більше α (ближче до 100), тим вищим буде рівень посилення контрасту. Те саме стосується CDF і функції ремапування:

$$cdf(g) = \sum_{l=0}^k pdf(g), \quad (2.4)$$

$$R(g) = cdf(g) \cdot g_{max}, \quad (2.5)$$

де $R(g)$ визначає функцію, яка повертає відрізану частину до кривої гістограми;

g_{max} – максимальне значення сірого для пікселів у блоці.

Перевагою CLАНЕ є те, що він має відносно менші обчислювальні витрати завдяки тому, що він впливає на невеликі підблоки зображення, а не на все зображення в цілому. На рис. 2.3 показано: а) точку фіксації та б) концепцію білінійної інтерполяції.

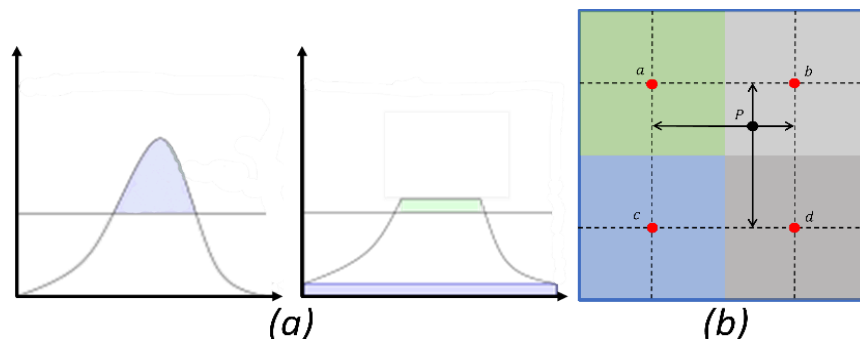


Рисунок 2.3. а) Точка відсікання та перерозподіл у CLАНЕ б) Білінійна інтерполяція

2.2. Локальний екстрактор ознак двійкового шаблону з адаптивним порогом для систем розпізнавання облич

Однією з найбільш швидкозростаючих тем серед додатків для обробки зображень є розпізнавання облич. Два фактори, що сприяють цьому розширенню, – це вдосконалення інфраструктури та збільшення попиту. Розпізнавання облич як неінтрузивна біометрична система має багатообіцяюче майбутнє з точки зору сфери застосування. Все це разом робить розпізнавання облич бажаною темою для досліджень. За останні кілька десятиліть системи розпізнавання облич значно вдосконалилися. Ранні методи були повільними, складними і відносно ненадійними, з низьким рівнем точності.

Для порівняння, останні досягнення покращили алгоритми в усіх аспектах. Сьогодні системи розпізнавання облич набагато надійніші, послідовніші, простіші в реалізації та порівняно швидкі. Ці системи можна знайти у великій кількості практичних застосувань, від безпеки до маркетингу, від обслуговування клієнтів до безпеки. Однак, є ще багато можливостей для вдосконалення. Наприклад, одним з найважливіших параметрів системи розпізнавання облич є її точність. У більшості застосувань точності надається перевага перед швидкістю. Нещодавно з'явилися методи та підходи, які забезпечують багатообіцяюче розпізнавання. Однак проблема, з якою стикаються сучасні системи, – це непередбачувані параметри, такі як шум і погане освітлення. Ці неконтрольовані ситуації різко впливають на надійність системи до такої міри, що роблять її непридатною для деяких застосувань. Варто зазначити, що небажані фактори з більшою ймовірністю виникають у реальних ситуаціях і додатках, чого бракує в імітаційних експериментах. Тому будь-які зусилля, спрямовані на підвищення точності системи, як і раніше цінуються. Для вирішення згаданих проблем було розроблено багато підходів. Однак, щоб краще зрозуміти шлях, доцільно спочатку розбити етапи роботи типової системи розпізнавання облич. Тоді вдосконалення можна буде

спрямувати на покращення кожного з цих рівнів. На рис. 2.4 показано блок-схему загального алгоритму розпізнавання облич.

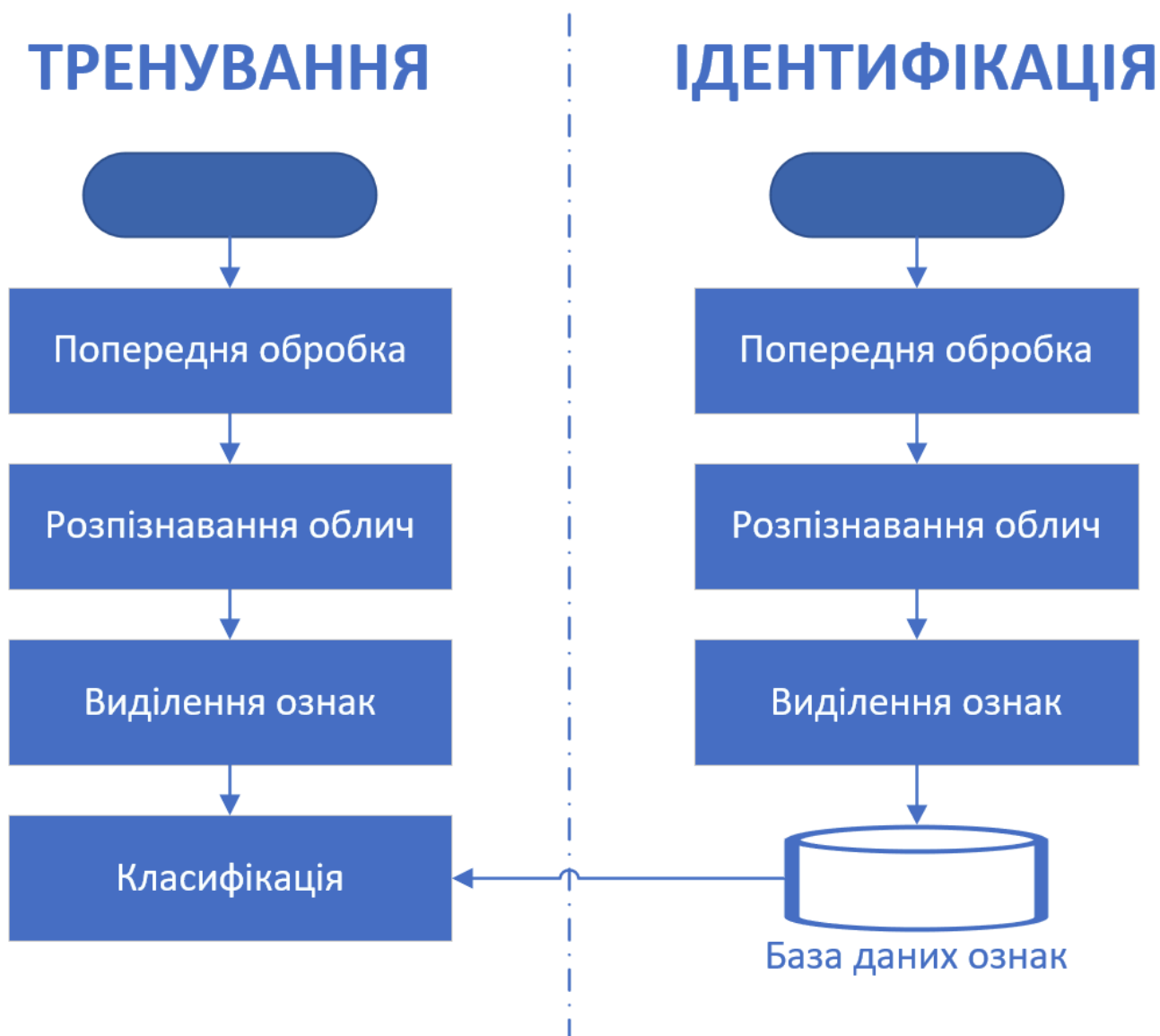


Рисунок 2.4. Блок-схема процесу розпізнавання обличчя

Як показано на рисунку 2.4, існує чотири основні етапи алгоритму: 1) попередня обробка, 2) розпізнавання обличчя, 3) вилучення ознак і 4) класифікація. Покращення роботи будь-якого з цих етапів може призвести до створення системи з вищою швидкістю розпізнавання.

Однак однією з найважливіших частин будь-якої системи розпізнавання облич або взагалі системи обробки зображень є частина вилучення ознак. Екстрактор ознак має найбільшу вагу у визначенні складності системи. Крім

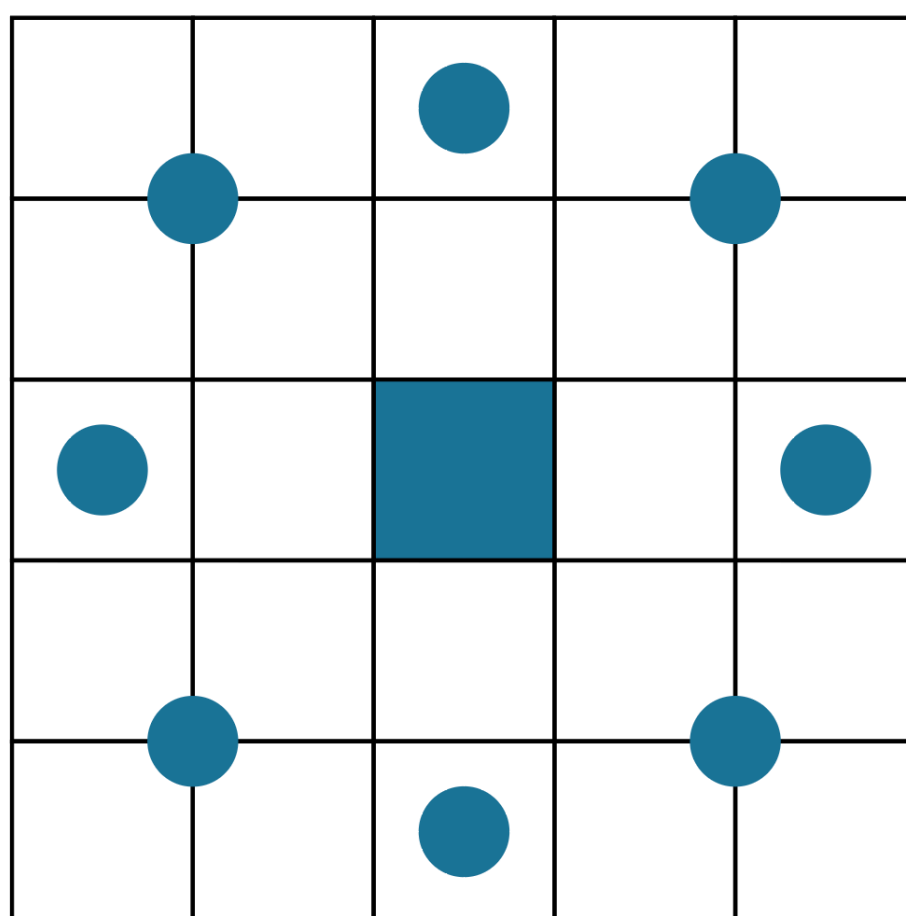
того, деякі властивості системи безпосередньо залежать від цього етапу, такі як завадостійкість, швидкість і більша частина обчислювальної складності. Методи вилучення ознак, в основному, можна розділити на дві основні категорії: 1) холістичні методи, які прагнуть зменшити розмірність зображення обличчя, зберігаючи при цьому основну необхідну інформацію. Аналіз головних компонент (PCA), лінійний дискримінантний аналіз (LDA) та локальний бінарний патерн є прикладами цього підходу. 2) Методи, що базуються на особливостях частотного простору зображення, такі як дискретне косинусне перетворення (DCT) та вейвлет Габора.

Обчислювальна складність – це параметр, який відіграє визначальну роль у виборі методу екстракції. Завдяки своїй обчислювальній простоті, стійкості до змін освітлення та простоті реалізації, локальний бінарний шаблон (LBP) є одним з найпопулярніших методів розпізнавання облич. Хоча звичайний локальний бінарний шаблон має свої недоліки, тим не менш, переваги переважають недоліки. У цій статті пропонується новий підхід до LBP для розпізнавання облич з використанням нової адаптивної функції в структурі оператора. Новий оператор модифіковано для адаптації на основі варіації гауссової функції розподілу. Ідея цього вдосконалення походить від недоліків звичайного LBP: фіксований поріг розпізнавання.

Серед різних існуючих методів вилучення ознак LBP є одним з найпоширеніших, особливо в додатках для розпізнавання облич, завдяки простоті реалізації та успішним результатам. LBP – це метод, заснований на локальному зовнішньому вигляді, який витягує текстурні особливості зображення шляхом порівняння кожного пікселя з його сусідами. Він не потребує навчання, що робить його швидким і легким для інтеграції з новими наборами даних, а також стійким до обертання, масштабування та зміни освітлення.

Виділення ознак локального двійкового шаблону базується на тому, що для опису двовимірної інтерпретації зображення можна визначити подвійний відповідний параметр. Локальний патерн і контраст градацій сірого – це ті два

парні елементи, які використовуються для формування гистограми. Ідея була вперше представлена Т. Оjala та ін. у 1996 році [27]. Початковий оператор LBP розглядає вікно сусідніх пікселів розміром 3 на 3, щоб присвоїти пікселю мітку. Це називається LBP-кодуванням. Три сусідні пікселі забезпечують $2^3 = 8$ рівнів для формування гистограми як карти ознак. Пізніше система була модифікована, щоб прийняти інший розмір вікна, ніж початковий 3x3. Крім того, було введено кругову область, а також білінійну інтерполяцію для покращення свободи вибору кількості сусідів. На рис. 2.5 показано концепцію процесу вибору сусідів LBP.



$$R = 2 \mid P = 8$$

Рисунок 2.5. Радіус та кількість сусідніх пікселів у LBP

P і R – це кількість пікселів і радіус кола відповідно. Ці два параметри визначають пікселі, які використовуються для порівняння з центром для створення LBP-коду. Тепер, коли алгоритм визначив пікселі, що беруть участь у процесі, наступним кроком є знаходження різниці між еталонним пікселем у центрі та іншими пікселями. Обчислена різниця між пікселями в кінцевому підсумку позначає LBP-код на основі рівняння:

$$b_p = \begin{cases} 1 & \text{якщо } D_p \geq 0 \\ 0 & \text{якщо } D_p < 0 \end{cases} \quad p = 1, 2, \dots, N, \quad (2.6)$$

де b_p – LBP-код пікселя p ;

N – кількість пікселів;

D_p – різниця між пікселями.

На рис. 2.6 зображено цей крок для прикладу:

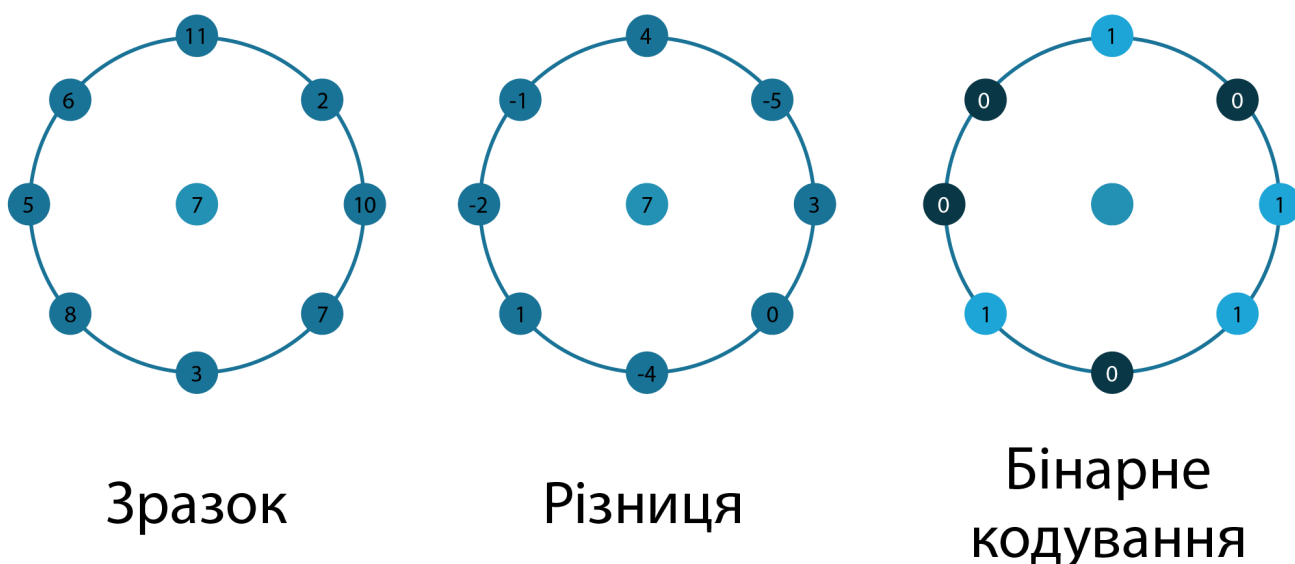


Рисунок 2.6. Приклад виконання LBP кодування

Нарешті, на основі двійкового коду, створеного алгоритмом, обчислюється десяткове число. Наступне рівняння показує процес обчислення цього значення:

$$LPB_i = \sum_{p=1}^N b_p(D_p) \cdot 2^{p-1}, \quad (2.7)$$

Іншою змінною в цьому налаштуванні є напрямок обертання, який визначає послідовність пікселів і набуває двох значень: за годинниковою

стрілкою та проти годинникової стрілки. Ми також можемо задати початкову точку в процесі створення десяткової послідовності, яка має бути однаковою для всіх пікселів зображення.

Тепер, коли зображення перетворено в серію десяткових значень, створених на основі взаємозв'язку між кожним пікселем і його сусідами, наступним кроком є генерація гістограми зображення. Хоча побудова гістограми відбувається на блоках зображення, а не на всьому зображенні. Схему цих кроків можна побачити на рис. 2.7.

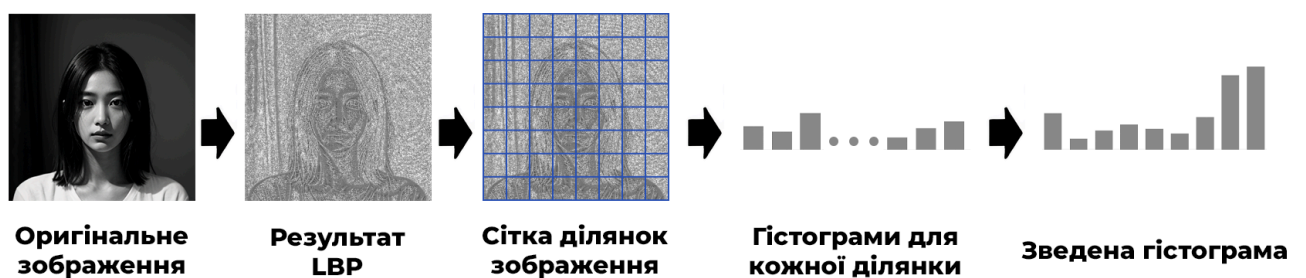


Рисунок 2.7. Гістограма зображення на основі LBP-коду

Як згадувалося раніше, LBP вважається одним з найнадійніших методів виділення текстурних ознак. Багато дослідників у галузі розпізнавання облич використовували LBP як екстрактор ознак для своїх систем. Цьому сприяють такі характеристики, як простота і стійкість до монотонних змін у зображенні, таких як зміна освітлення, обертання і масштабування. Однак, коли справа доходить до максимального потенціалу, одним з основних недоліків є те, що звичайний ЛБП використовує фіксований нульовий поріг для різниці пікселів, враховує лише знак D_p і ігнорує амплітуду. Фіксований нульовий поріг збільшує чутливість системи до шуму, і важлива інформація в амплітуді втрачається, що знижує точність методу.

2.3. Системи розпізнавання облич з низькою похибкою на основі організації згорткової нейронної мережі та доповнення даних

Розпізнавання облич, без сумніву, стало одним з найпопулярніших біометричних заходів за останні два десятиліття, особливо в додатках для верифікації. Можна нарахувати багато причин такої популярності, ось декілька з них: 1) розпізнавання обличчя є неінтрузивним підходом, що означає, що система не вимагає активної взаємодії з суб'єктом для виконання біометричної верифікації. 2) останні досягнення в галузі технологій та обладнання забезпечили необхідну інфраструктуру для того, щоб системи розпізнавання облич можна було без особливих зусиль впроваджувати в тих місцях, де це необхідно. 3) Досягнення в обробці зображень і розпізнаванні об'єктів значно покращили можливості цих систем. Завдання, які ще кілька років тому здавалися чимось з розряду наукової фантастики, зараз широко використовуються в комерційних цілях.

Ці вдосконалення можна побачити і в системах розпізнавання облич. У більшості випадків існуючі системи показують багатообіцяючу точність, прийнятну швидкість і адекватну низьку складність. Однак, ці прийнятні характеристики досяжні лише для неконтрольованої ситуації. Додавання неконтрольованих параметрів, таких як шум, зміна освітленості або оклюзія, суттєво впливає на продуктивність системи до такої міри, що іноді робить її марною.

Представлено багато досліджень, спрямованих на вирішення цих проблем і вдосконалення існуючих систем, щоб зробити їх більш стійкими до цих небажаних параметрів.

У цьому дослідженні, однак, ми зосереджуємося на частині системи розпізнавання облич, яка виокремлює риси обличчя. Існують різні підходи до вилучення ознак обличчя. Лінійні підходи відображають лінійні ознаки зображення обличчя в карту ознак нижчої розмірності. Прикладами цих методів є аналіз головних компонент (PCA), лінійний дискримінантний аналіз (LDA) та лінійний дискримінант Фішера (FLD). Ці методи також називаються методами підпросторів на основі зовнішнього вигляду і зосереджені на зменшенні розмірності. У другій категорії підходів до виділення ознак використовується

перетворення частотної області, наприклад, дискретне косинусне перетворення (ДКП) або вейвлет Габора, для фіксації відмінних рис зображення. Деякі з коефіцієнтів частотної області використовуються в цих підходах для створення характеристик зображення. Висока розмірність просторових частотних характеристик негативно впливає на вейвлет Габора. Для вирішення цієї проблеми можуть бути використані підходи до зменшення розмірності підпростору. Наступний підхід – це той, який ми розглядаємо в цій статті. Глибоке навчання – це метод, який вирішив багато обмежень, недоліків і проблем, з якими методи машинного навчання стикалися протягом багатьох років. Методи глибокого навчання виявилися найбільш надійним і ефективним методом в обробці зображень, особливо в додатках для розпізнавання облич. Методи, засновані на нейронних мережах, виділяють глибокі та складні риси обличчя, які є більш унікальними, а отже, надають більше відмінних рис, які можна використовувати в процесі розпізнавання. Прикладами таких мереж є Stacked Autoencoder та Deep Belief Network (DBN) [6; 12]. Ще одна успішна нейронна мережа, яка здебільшого використовується в додатках розпізнавання облич і була визнана найкращою в додатках обробки зображень, – це згортова нейронна мережа (Convolutional Neural Network, CNN) [2]. Метод базується на згортці пікселів зображення з ядром для виділення жирним шрифтом і вилучення прихованих особливостей зображення. Метод використовує інші коригувальні параметри та шари для зменшення розмірності та вилучення глибинних особливостей зображення. Відтоді його було впроваджено в незліченну кількість алгоритмів розпізнавання облич. Тим не менш, CNN мають деякі слабкі місця в деяких додатках. Наприклад, при роботі з базами даних з відносно невеликою кількістю зразків на клас, системи на основі ШНМ показують низьку точність. Така ситуація частіше трапляється в реальних ситуаціях, коли ми не маємо доступу до великої кількості зразків для кожної особи.

CNN – це тип нейронної мережі прямого поширення, який базується на згортці пікселів зображення як вхідних даних і ядра, яке є важливою частиною

мережі. Завдяки своїй природі вони надзвичайно ефективно витягують глибокі та приховані особливості зображення, що робить їх дуже практичними в додатках для розпізнавання облич. ШНМ, як правило, відносяться до категорії неконтрольованих методів і використовують ознаки, які не завжди візуально інтерпретуються людським оком. Ці ознаки зважуються алгоритмом на основі їхньої важливості на більш пізніх етапах процесу. Інша перевага ШНМ полягає в тому, що ядра і фільтри цих типів мереж не потрібно розробляти або налаштовувати вручну, а система на основі навчальних наборів даних з часом вивчає значення, пов'язані з ознаками для цих параметрів, і встановлює їх автоматично. Варто зазначити, що натхненням для дизайну CNN і схеми зв'язку в цій мережі є нейронні зв'язки головного мозку, зокрема, зорової кори головного мозку людини. Згорткові нейронні мережі складаються з різних типів шарів. Ці шари, які більшою чи меншою мірою можна знайти майже в кожній архітектурі ШНМ, – це шар згортки, шар об'єднання, шар ReLU і повністю пов'язаний шар.

Операція згортки відбувається у цьому шарі. Також ядра є частиною цього шару. Найважчі обчислення мережі відбуваються в шарі згортки. Перед будь-якою згорткою алгоритм ділить зображення на ділянки рівного розміру, які називаються блоками зображення. Основною операцією тут є згортка пікселів цих блоків і ядра, як згадувалося раніше. Згортання цих частин зберігає просторові відношення пікселів блоку. Заключний етап цього шару виконується набором нейронів, що навчаються, і полягає у формуванні карти ознак. Зауважимо, що ця карта, як вихід шару згортки, може бути подана на наступний шар як вхід.

Мета цього шару – зменшити розмірність даних і знизити складність системи. Функція об'єднання – це нелінійна операція, яка зменшує розмірність при збереженні необхідної інформації. Знову ж таки, першим кроком є поділ зображення на менші блоки пікселів, що не перекриваються. На наступному кроці нелінійний процес відбирає пікселі кожного блоку. Тип операції дискретизації визначає тип об'єднання в алгоритмі. На рис. 2.8 показано два

популярні типи об'єднання: максимальне об'єднання та середнє об'єднання. Ці типи шарів зазвичай можна знайти після шарів згортки.

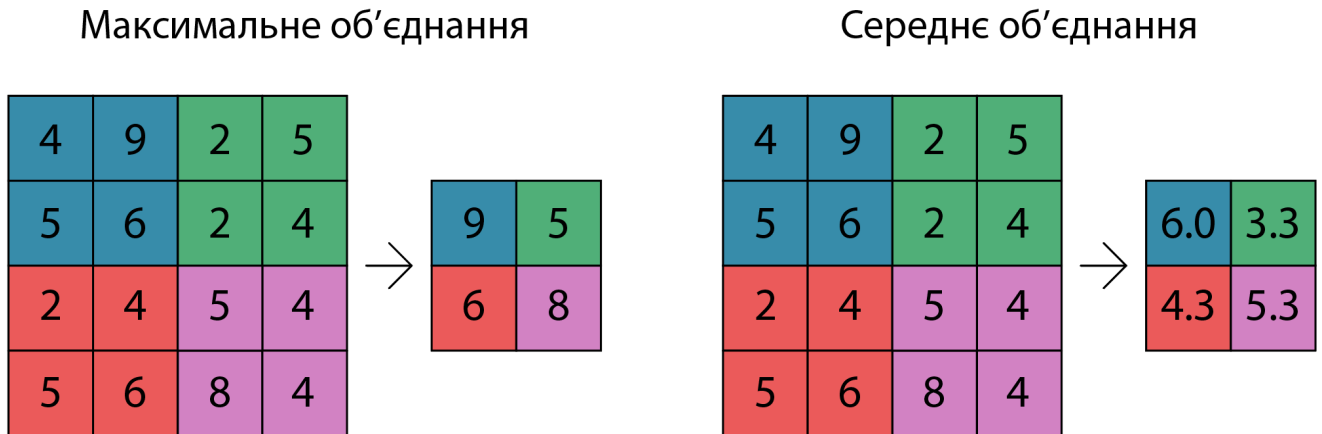


Рисунок 2.8. Максимальне об'єднання та середнє об'єднання операцій

Випрямлення – це одна нелінійна операція в структурі звичайного CNN. ReLU – це попіксельна операція, яка усуває пікселі з від'ємним значенням, замінюючи їх нулем. Математична візуалізація цієї операції показана в рівнянні:

$$f(i) = \max(0, i), \tag{2.8}$$

де i – значення пікселя зображення.

На рис. 2.9 показана функціональна діаграма ReLU.

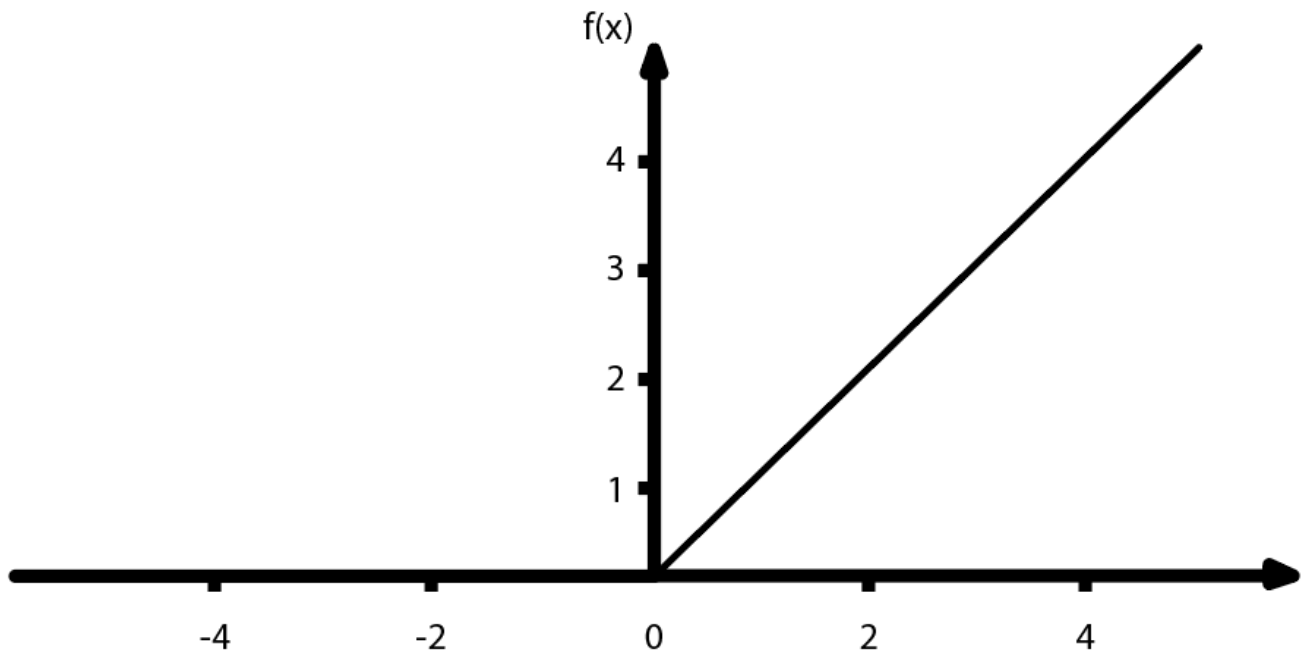


Рисунок 2.9. Функція ReLU

Fully Connected Layer (FCL) відноситься до специфічної властивості мережі: кожне ядро в кожному шарі пов'язане з іншими ядрами в інших шарах. Шари мережі разом формують складний взаємозв'язок характеристик зображення. FCL дозволяє системі використовувати ці ознаки і створює карту ознак, яка передається до класифікатора. Останній об'єднуючий шар пересилає витягнуті ознаки в класифікатор і виконує в алгоритмі роль FLC.

Іншим поняттям мережі є ядра. Ядра – це аспект мережі, який виокремлює особливості зображення. Вони виконують цю дію, посилюючи певну особливість зображення. Операція включає переміщення ядра, яке є матрицею над зображенням і згортає відносні пікселі зображення до ваг ядра. У CNN ми не задаємо значення ядра вручну, а система визначає їх у процесі навчання. Термін «крок» тут означає кількість пікселів, яку покриває кожен рух ядра. Нульове заповнення – це метод, який використовується для виправлення розміру зображення, якщо він не ділиться на розмір ядра, шляхом додавання нульових пікселів до зображення.

Доповнення даних (Data augmentation) – це спосіб створення більшої кількості зразків на основі наявних зразків. Цей метод може збільшити кількість

зразків у випадках, коли кількість зразків на клас невелика. Як згадувалося раніше, CNN може зіткнутися з труднощами в таких ситуаціях через брак достатньої кількості даних для належного навчання. Деякі з популярних методів доповнення даних – це горизонтальний фліп, зсув, масштабування та обертання. Ці методи, незважаючи на те, що вони прості у виконанні, довели свою ефективність. На рис. 2.10 показано візуальну презентацію згаданих методів.

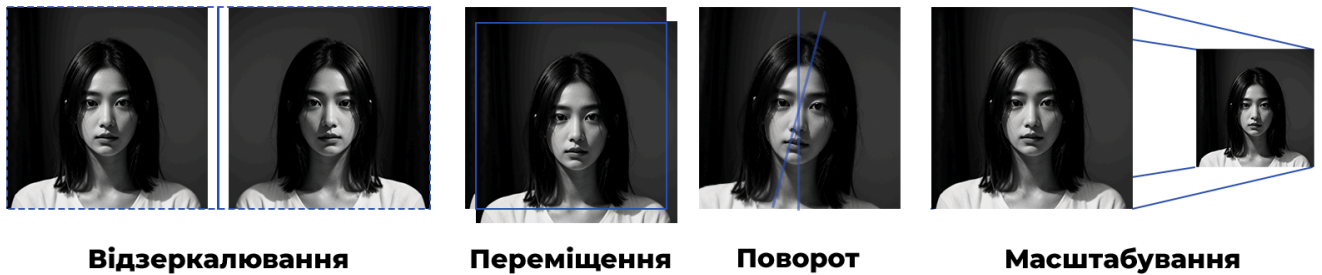


Рисунок 2.10. Data augmentation

Висновки до розділу 2

Отже, у другому розділі дипломної роботи детально розглянуто методологічні засади побудови автоматизованої системи розпізнавання облич. Основну увагу приділено трьом ключовим аспектам: системам автоматичного адаптивного вирівнювання гістограм з обмеженням контрасту та адаптивною середньою подвійною гама-корекцією, локальному екстрактору ознак двійкового шаблону з адаптивним порогом для систем розпізнавання облич, а також системам розпізнавання облич з низькою похибкою на основі організації згорткової нейронної мережі та доповнення даних.

Розглянуто основні методи попередньої обробки зображень, зокрема вирівнювання гістограм, гама-корекцію та їх адаптивні варіанти. Представлені методи спрямовані на покращення якості зображень за рахунок корекції освітлення та контрасту, що є критичним для підвищення точності розпізнавання облич. Розглянуто також комбіновані методи, такі як CLANE з подвійною гама-корекцією, які забезпечують збереження деталей зображення при покращенні його загальної освітленості.

Досліджено методи вилучення ознак зображень, зокрема локальний бінарний паттерн (LBP) та його вдосконаленням. Запропоновано новий підхід до LBP з використанням адаптивної функції, що базується на варіації гауссової функції розподілу. Такий підхід дозволяє покращити стійкість системи до шуму та інших небажаних факторів, забезпечуючи більш точне розпізнавання облич.

Розглянуто застосування згорткових нейронних мереж (CNN) для задач розпізнавання облич. CNN показали високу ефективність у вилученні глибинних ознак зображень, що дозволяє значно підвищити точність розпізнавання. Також розглянуто методи доповнення даних, які дозволяють збільшити кількість зразків у навчальних наборах, що є критичним для підвищення ефективності CNN у реальних умовах з обмеженими даними.

Загалом, надано комплексний огляд сучасних методів попередньої обробки зображень, вилучення ознак та застосування глибинного навчання для розпізнавання облич. Представлені методи та підходи сприяють підвищенню точності та надійності систем розпізнавання облич, що є важливим для їх успішного застосування в реальних умовах. Подальші дослідження в цьому напрямку спрямовані на оптимізацію цих методів та адаптацію їх до змінних умов освітлення, шуму та інших факторів, що впливають на якість зображень.

РОЗДІЛ 3.

ПРОГРАМНО-ТЕХНІЧНЕ РІШЕННЯ АВТОМАТИЗОВАНОЇ СИСТЕМИ РОЗПІЗНАВАННЯ ОБЛИЧ НА ОСНОВІ ГЕОМЕТРИЧНИХ ХАРАКТЕРИСТИК ОБЛИЧЧЯ ТА ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ

3.1. Класифікація ієрархічного методу розпізнавання облич

Хоча методи підвищення контрастності дають прийнятні результати в багатьох випадках, вони все ще стикаються з недоліками, коли вхідне зображення має сильні темні ділянки і глибокі тіні. Це головним чином пов'язано з тим, що β або точка відсікання є фіксованою величиною, і тому вона не може змінюватися залежно від характеру поганого освітлення. Іншими словами, фіксована глобальна точка відсікання може давати задовільні результати для деяких частин бази даних, в той час як можуть бути випадки, які страждають від надмірного посилення контрасту, а також випадки, які викликають артефакти. Крім того, покращення контрастності зображення не є достатнім для того, щоб підвищити освітленість до достатнього рівня. Це особливо помітно у випадках з дуже темними неоднорідними ділянками.

Таким чином, ці моменти обґрунтовують визначення адаптивної точки відсікання та гамма-корекції як запропонованого методу. Запропонована методика використовує адаптивну точку відсікання на основі параметрів кожного блоку. Алгоритм також встановлює гамма-корекцію для g_{max} для покращення динамічного діапазону блоку.

У звичайному CLANE з фіксованою точкою α і максимальний градієнт (Gm) є параметрами, що визначають точку відсікання. Щоб покращити продуктивність CLANE, алгоритм розташовує точку кліпу таким чином, щоб мати нижчі значення для згладжених і гармонізованих ділянок і вищі значення для блоків, що містять більшу варіацію освітленості.

Двома цінними параметрами, які розкривають корисну інформацію про блок, є медіанний рівень сірого та стандартне відхилення. Використання цих

параметрів для встановлення адаптивної точки кліпу представлено у вигляді наступного рівняння:

$$\delta = \frac{g_{mod}}{D_b}, \quad (3.1)$$

$$\beta = \frac{N}{D_b} \left(1 + Q \frac{\delta}{D} + \frac{a}{100} \left(\frac{\sigma}{Avg+c} \right) \right), \quad (3.2)$$

де g_{mod} – найбільш повторюване сіре значення блоку;

δ – динамічний фактор;

σ – стандартне відхилення;

c – мале значення.

Q та a контролюють динамічний діапазон та ентропійну вагу відповідно.

Нова точка кліпування має більше значення для блоків з більшою кількістю текстурної інформації, оскільки ці блоки мають більший коефіцієнт σ/Ave . На рисунку 3.1 показано приклад нової точки відсікання для різних блоків зображення.

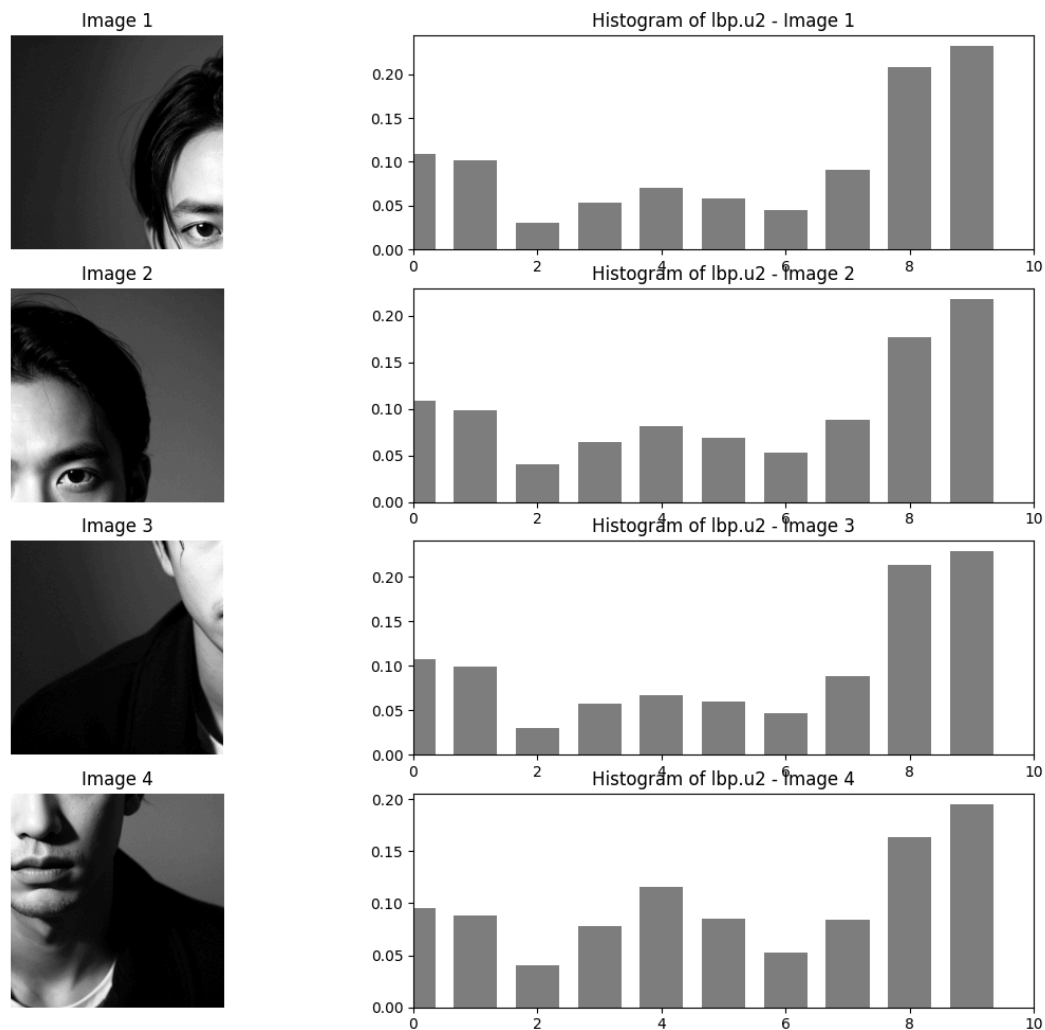


Рисунок 3.1. Точка кліпу, керована вмістом

Хоча алгоритм покращує контрастність на основі глобальної та локальної інформації та властивостей зображення, саме лише вирівнювання гістограми не може забезпечити достатнього покращення зображення, особливо в однорідних темних ділянках зображення. Це проявляється на ділянках з екстремально відкинутими тінями або в ситуації, коли в якості джерела освітлення використовується прожектор, оскільки зображення обличчя містить тіні частин обличчя, що блокують світло (в районі носа або очей). Для вирішення цієї проблеми ми пропонуємо гамма-корекцію. Процес гамма-корекції полягає в технічному збільшенні варіації рівня сірого в затінених ділянках. Це покращує освітлення за рахунок збільшення деталей у темних областях. Функція відображення гамма-корекції визначається наступним чином:

$$R(g) = g_{max} \left(\frac{g}{g_{max}} \right)^{\gamma}, \quad (3.3)$$

Це співвідношення між g_{max} , $R(g)$ та α визначає поведінку гамма-корекції. Чим менше значення γ , тим інтенсивніше збільшується значення відповідного пікселя. Проте, через попередньо визначене значення гамма-корекції, ця зміна відбувається у всіх пікселях блоку зображення і може призвести до погіршення контрастності. Незважаючи на те, що значення гама-корекції діє локально, погіршення контрастності нівелює ефект CLAHE. Запропонований алгоритм впроваджує гамма-корекцію у введenu CLAHE, щоб запобігти цьому. Для цього спочатку потрібно ввести зважену корекцію рівня сірого на основі значення γ .

$$w = \left(\frac{GL_{mod}}{GL_{med}} \right)^{1-\gamma_1}, \quad (3.4)$$

де GL_{mod} – найбільш повторюваний рівень сірого;

GL_{med} – середній рівень сірого зображення.

У цьому дослідженні розширений набір даних Єльського університету використовується для оцінки якості запропонованого методу шляхом порівняння точності різних методів розпізнавання облич з і без запропонованої методики як етапу попередньої обробки, а також оцінки продуктивності різних методів попередньої обробки, включаючи запропоновану систему, застосовану до конкретного алгоритму розпізнавання облич.

LBP вважається одним з найпопулярніших методів вилучення ознак у локальних системах, заснованих на зовнішньому вигляді. Принцип LBP полягає у вилученні текстурних особливостей зображення шляхом порівняння пікселя з певною кількістю його сусідніх пікселів. Однак, існують певні параметри, за допомогою яких можна налаштувати поведінку ЛБЗ. Деякі з цих змінних – кількість сусідніх пікселів, напрямок обертання при формуванні LBP-коду, радіус та функція кодування.

Монотонні зміни в зображенні обличчя є найбільш впливовими явищами, які змінюють кінцеві результати роботи алгоритму розпізнавання облич. Зміна

освітленості, як одна з таких змін, зміщує інтенсивність пікселів зображення і, таким чином, змінює різницю між пікселями Dp . Однак з (3.1) можна визначити, що ця зміна відображається не повністю, оскільки алгоритм враховує лише знак різниці і не враховує амплітуду. Зрозуміло, що важлива інформація про зображення може бути вилучена з пропущеної амплітуди. Крім того, фіксований нульовий поріг у звичайному LBP додає чутливості до шуму і робить систему вразливою. Це пов'язано з тим, що доданий шум може змінити двійкові біти з нуля на одиницю або навпаки, вносячи незначні зміни.

3.2. Розробка автоматизованого алгоритму розпізнавання облич на основі згорткової нейронної мережі

Блок-схема запропонованого методу показана на рис. 3.2. Метод складається з двох основних розділів: геометричного та класифікації ознак CNN, а також кількох інших підрозділів.

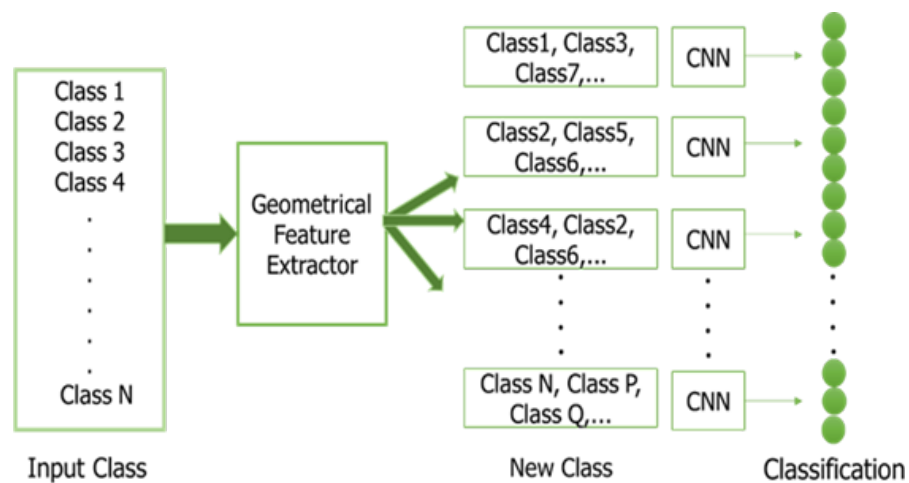


Рисунок 3.2. Блок-схема запропонованого алгоритму CNN

В екстракторі геометричних ознак алгоритм використовує АММ для вилучення геометричних ознак обличчя на основі розташування деяких частин обличчя. Шість основних точок і три вузлові точки розглядаються для розрахунку п'яти осей або відстаней у вигляді карти ознак для кожного вхідного зображення. Тегми цих п'яти відстаней є WA, EA, WEA, LA та ELA, як показано на рис. 3.3. Кожна відстань створює дві осі, ліву і праву сторони

відстані. Це зроблено, головним чином, для того, щоб запобігти ефекту варіації орієнтації обличчя.



Рисунок 3.3. a) WA, b) EA and WEA, c) LA and ELA

На наступному етапі алгоритм подає карту ознак на SVM-класифікатор. SVM класифікує залежність входу від класів. Зокрема, перші десять найбільш ймовірних класів присвоюються кожному входу. Цей процес призводить до створення нових класів, що містять декілька членів початкових вхідних класів.

Перед тим, як помістити дані в ШНМ для навчання, виконується процес доповнення класів бази даних. Етап доповнення допомагає подолати негативний ефект малої кількості вхідних зразків. Деякі з методів, які ми використовували для доповнення даних, – це горизонтальний фліп, зсув, масштабування та обертання. Ці методи показані на рисунку 3.4.

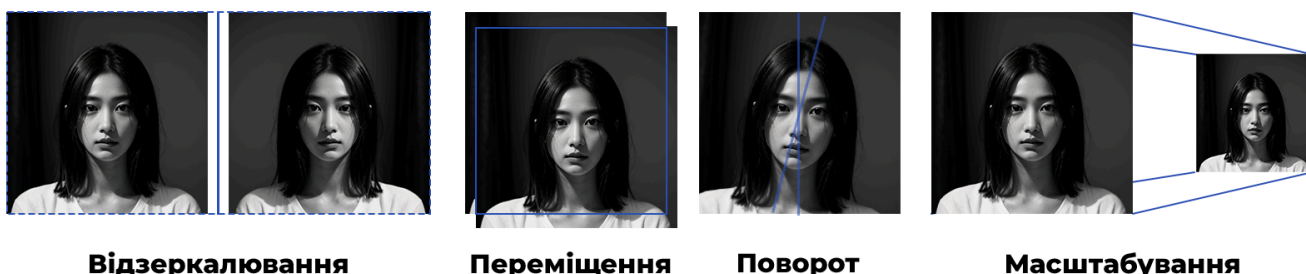


Рисунок 3.4. Data Augmentation

На етапі згорткової нейронної мережі використовується нова запропонована компоновка шарів. У цій схемі є кілька шарів згортки, ReLU та об'єднання. Крім цих шарів, ми запропонували додати до компоновання шари нормалізації. На рисунку 3.5 показано запропоновану компоновку шарів секції виділення ознак CNN, а в таблиці 3.1 наведено параметри цих шарів.

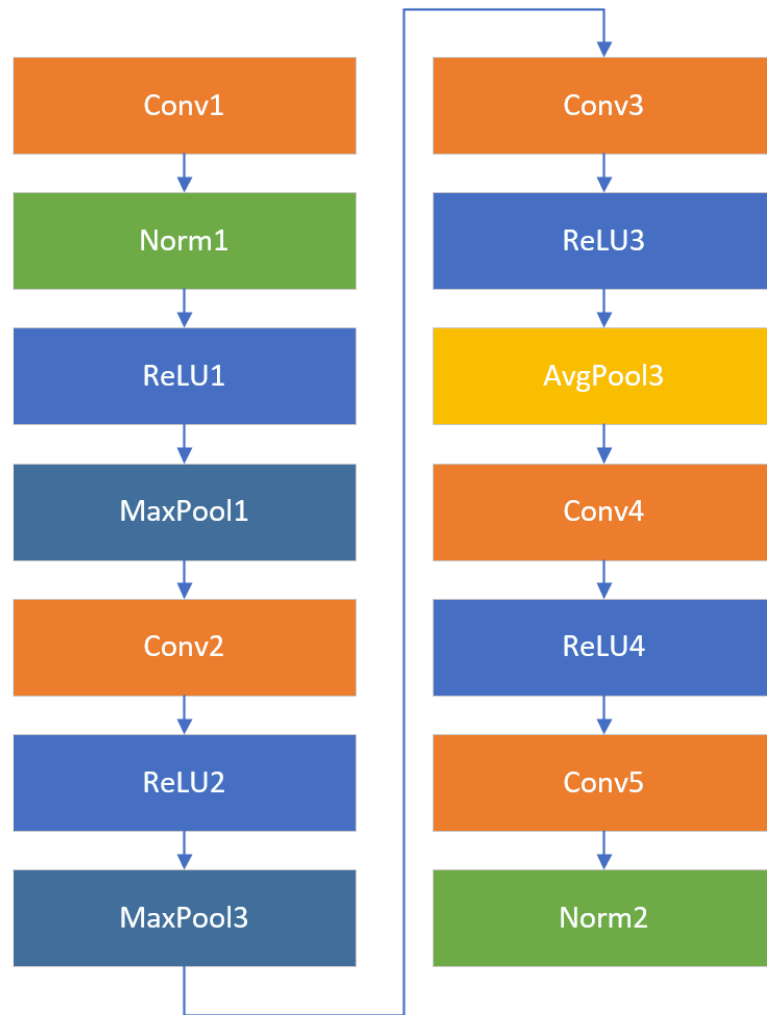


Рисунок 3.5. Пропоноване розташування шарів

Таблиця 3.1. Параметри шарів CNN

Тип	Розмір патча	Крок
Conv1	5X5	2
Conv2	3X3	2
Conv3	3X3	1

Conv4	3X3	2
Conv5	3X3	1
Pool1	3X3	1
Pool2	3X3	1
Pool3	8X8	-

Softmax вимірює ймовірність належності вхідних даних до класу. Як класифікатор, він дозволяє алгоритму передбачати ймовірність мітки в базі даних з декількома мітками. Рис. 3.6 ілюструє роботу Softmax та його розташування у звичайній ШНМ.

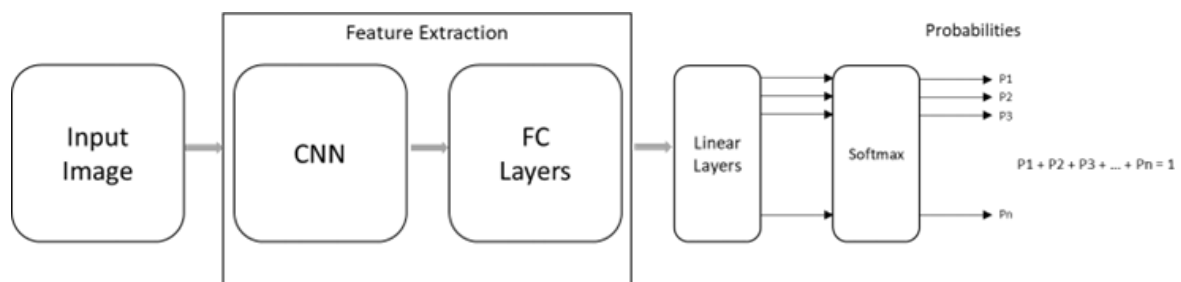


Рисунок 3.6. Розташування та відношення Softmax до CNN

3.3. Експериментальні випробування та оцінка розробленого алгоритму

У цьому дослідженні розширений набір даних Єльського університету використовується для оцінки якості запропонованого методу шляхом порівняння точності різних методів розпізнавання облич з і без запропонованої методики як етапу попередньої обробки, а також оцінки продуктивності різних методів попередньої обробки, включаючи запропоновану систему, застосовану до конкретного алгоритму розпізнавання облич.

Розширений Єльський набір зображень облич містить розширення оригінального Єльського набору даних В на 28 об'єктів. Набір даних включає 21888 фотографій, зроблених при моноосвітленні. В експерименті брали участь 38 осіб із загальною кількістю 576 умов перегляду. Для кожного учасника було знято зразок з фоновим освітленням, який супроводжувався певним положенням голови цієї людини. Розширений набір даних Yale В розділений на 5 окремих підгруп залежно від кута між джерелом світла, напрямку та віссю камери, а також ситуації освітлення. Підгрупа 5, виходячи з нашого досвіду, має найскладніші умови освітлення і в більшості випадків показує найнижчий рівень розпізнавання.

Перший експеримент. У першому експерименті ми оцінюємо ефективність нашого методу, впроваджуючи запропоновану методику в декілька існуючих методів розпізнавання облич і досліджуємо зміни, які наш підхід вносить у ці системи. Щодо організації експерименту, то ми використали 40 зображень з Єльського набору даних для навчання і 24 для тестування. Ми також змінили розмір зображень до 64 на 64 пікселів на останньому етапі попередньої обробки після покращення освітлення. У таблиці 3.2 порівнюються показники розпізнавання різними методами для двох ситуацій попередньої обробки: 1) без попередньої обробки і 2) із застосуванням запропонованої методики. Варто зазначити, що результати було отримано в середньому з 20 запусків алгоритмів на наборі даних для кожного методу.

Таблиця 3.2. Вплив наявності та відсутності запропонованого методу на % розпізнавання різних методів розпізнавання облич

Метод	Без попередньої обробки	Запропонована модель
MRF	78.11	84.51
PCA	66.10	78.19

HMM	84.38	89.46
LDA+IPMML	81.74	88.63

Другий експеримент. У цьому тесті ми спробуємо оцінити ефективність запропонованого методу порівняно з іншими методами попередньої обробки. Для цього нам потрібно застосувати різні методи попередньої обробки до одного конкретного методу розпізнавання облич, щоб мати можливість оцінити результати кожного методу належним чином. Іншими словами, швидкість розпізнавання комбінацій різних методів попередньої обробки з обраною системою розпізнавання облич використовується для визначення якості та ефективності цієї попередньої обробки. Зрештою, ми обрали локальний бінарний шаблон (LBP) в якості екстрактора ознак алгоритму розпізнавання облич, а також машину опорних векторів (SVM) в якості класифікатора цієї системи. Знову ж таки, ми змінюємо розмір зображень до 64 на 64 перед тим, як подавати їх до екстрактора ознак. Ми також використовуємо ті самі 40 та 24 числа як навчальну та тестову множини для цього тесту. Таблиця 3.3 містить результати цього експерименту.

Таблиця 3.3. % розпізнавання системи розпізнавання облич LBP + SVM з різними методами попередньої обробки

Метод попередньої обробки	Коефіцієнт розпізнавання
None	81
HE	84
SQI	87
GIC	91
DoG	90
CLAHE	

Набір даних ORL [8] – це колекція з 400 зображень облич, отриманих від 40 осіб. Обличчя містять різні вирази обличчя, такі як посмішка, відкриті та закриті очі. Серед зразків також можна знайти деталі обличчя, такі як окуляри. Зображення обличчя мають розмір 92 на 112. Вони відформатовані у відтінках сірого з 256 рівнями сірого для кожного пікселя.

У цьому розділі ми перевіримо ефективність запропонованої нами системи шляхом її тестування. Для цього ми використали набір даних ORL. %50 вибірок з набору даних використовуються для навчання, а решта – для тестування. На етапі попередньої обробки зображення спочатку зменшуються до розміру 64 на 64. Потім використовується аугментація даних для збільшення кількості вибірок до 40000, 60000, 80000 та 120000. На наступному етапі розширений набір даних подається до CNN. Ми використовуємо два класифікатори, SVM і Softmax, і запускаємо тест для кожного окремо. Для кожного з цих доповнень обчислюється коефіцієнт розпізнавання. Знову ж таки, доповнення даних застосовується до вибірок, які ми використовували на етапі тестування, а кількість вибірок цього разу збільшилася до 4000, 6000, 8000 і 12000. Зауважте, що розраховані показники точності в нашому експерименті отримані шляхом усереднення результатів 20 прогонів кожного тесту.

Результати тестів представлені на наступних рисунках та діаграмах. На рисунках 3.7 та 3.8 показано показники точності системи для різної кількості

доповнень даних як у навчальних, так і в тестових наборах, а також для різних епох.

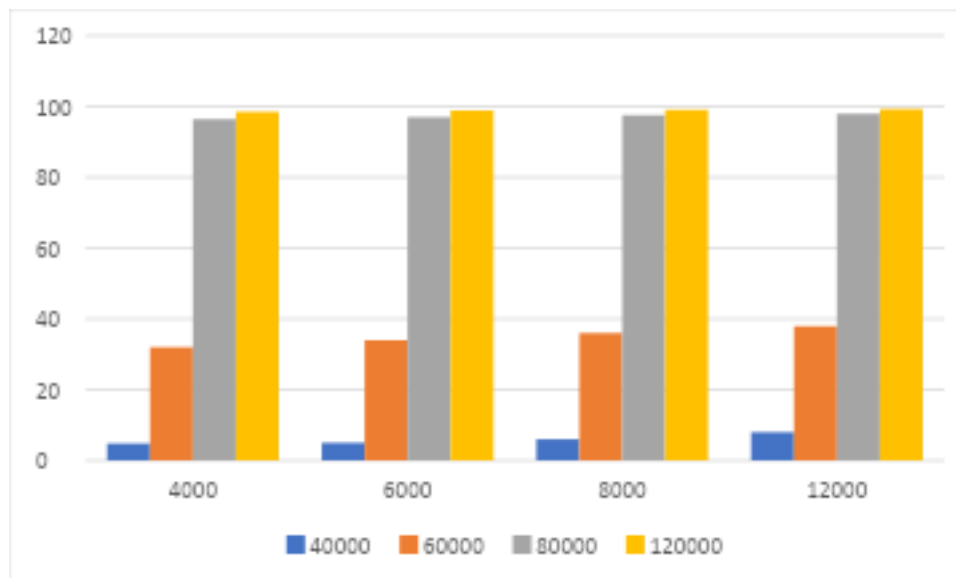


Рисунок 3.7. Точність системи для різної кількості доповнень даних

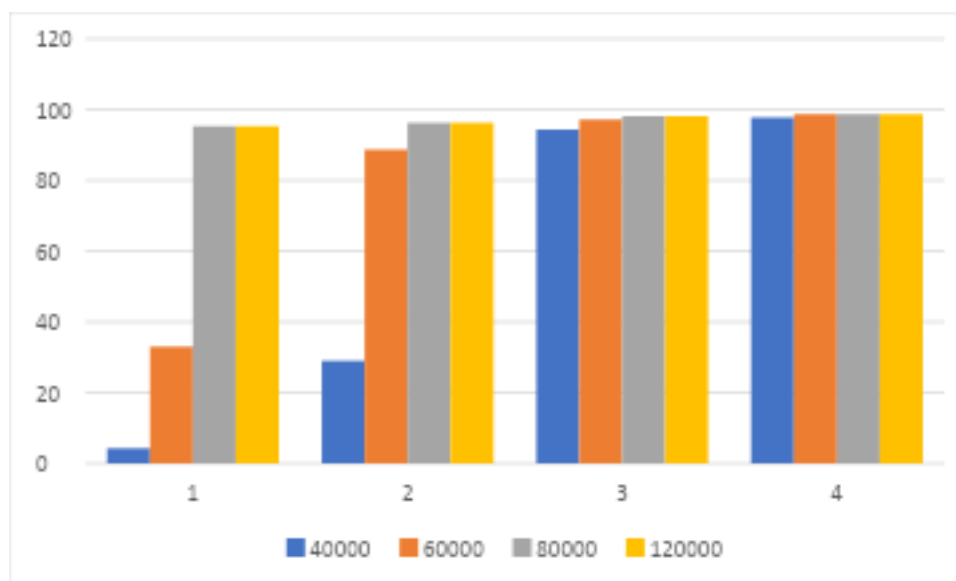


Рисунок 3.8. Точність системи для різних епох

Кольорова база даних FERET [30; 31] – це набір даних для розпізнавання облич. Вона містить 11 338 кольорових фотографій з роздільною здатністю 512x768 пікселів, зроблених у напівконтрольованих умовах з 1208 особами в різних положеннях. База даних містить різні пози для кожної особи – від фронтальних до зображень у півоберта. Крім того, для кожної особи існують

різні вирази обличчя та умови освітлення. У нашому експерименті ми використовували лише фронтальні зображення кожної людини, і зображення змінювали розмір до 64x64 щоразу, коли вони надходили до CNN. Кілька прикладів набору даних показано на рис 3.9.

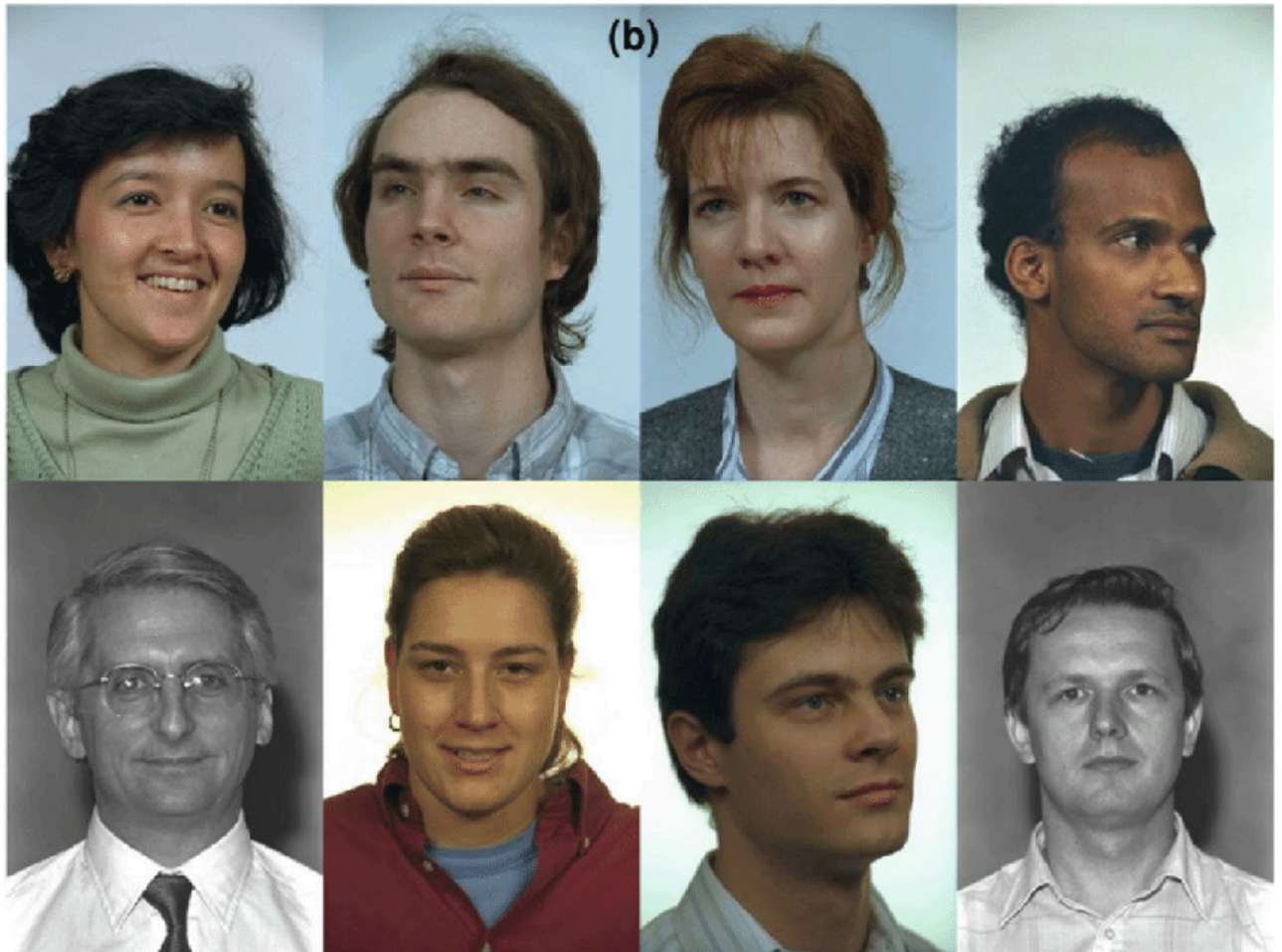


Рисунок 3.9. Зразки бази даних кольорів FERET

В експерименті використовуються всі класи бази даних кольорів FERET. АММ – це алгоритм, який використовується для вилучення геометричних ознак, а класифікатором на цьому етапі є SVM. Комбінація АММ і SVM створює 64 нових класи з 1208 оригінальних класів у базі даних. Середня кількість осіб у нових класах становить близько 21 особи. Загальна кількість осіб, з якими алгоритм має справу після ААМ, становить 1384 групи. Найбільший новий клас за кількістю осіб має 42 члени, а найменший – 5.

Перед тим, як помістити дані в CNN, вибірки змінюють розмір до 64x64. Доповнення даних здійснюється в нових класах, щоб збільшити кількість зразків до 10000 в кожному новому класі. Результати розпізнавання збираються як для SVM, так і для Softmax як класифікаторів згорткової нейронної мережі.

Половина даних використовується для навчання, а інша половина – для тестування. Якщо CNN не надає вихідних даних, тобто не зливається, відбувається повернення до попереднього етапу, і алгоритм розглядає наступний за рангом тег як новий клас для виводу АММ.

Результати запуску алгоритму на всіх 1208 класах бази даних кольорів FERET (50% навчання 50% тестування) та порівняння запропонованого методу з іншими існуючими методами наведено в таблиці 3.4. Таблиця 3.5 ілюструє деякі параметри тесту та результати для різних методів вилучення геометричних ознак, а таблиця 3.6 показує результати для різних параметрів, таких як розмір вхідних даних та результати для CNN. Як бачимо, найвищий показник розпізнавання (близько 97%) спостерігається при розмірі вхідного зображення 64x64, а також використання SVM-класифікатора збільшує показник розпізнавання порівняно з Softmax (97.68% проти 96.35%).

Таблиця 3.4. Швидкість розпізнавання на кольоровій базі даних FERET та порівняння результатів

Метод	Точність
SVM-Linear	81
Eigenfaces and DoG filter	84
SVM-Polynomial	87
SVM-RBF	91
Метод	Точність
HMM	90

Продовження
таблиці 3.4

AMM + CNN + Softmax	96.35
AMM + CNN + SVM	97.68

Таблиця 3.5. Параметри та результати різних екстракторів геометричних ознак

Метод	Розмір зображення	Час навчання на один зразок	Класи (найвищі)	Кількість повернень	Кількість підгруп
Tzimirow	Original	1.012	1208	89	64
Cootes	Original	0.88	1208	145	115
FACS	Original	0.98	1208	107	71

Таблиця 3.6. Параметри та вихідні результати різних розмірів вхідних даних та класифікаторів у CNN

Метод	Розмір зображення	Час на епоху Секунди (середнє значення)	#Класи (найвищий)	Точність	Неправильний прогноз
Proposed + SVM	16X16	14	42	71.85	340
Proposed + SVM	32X32	19	42	91.14	107

Продовження таблиці 3.6

Метод	Розмір зображення	Час на епоху Секунди (середнє значення)	#Класи (найвищий)	Точність	Неправильний прогноз
Proposed + SVM	64X64	33	42	97.68	28
Proposed + Softmax	16X16	18	42	69.86	364
Proposed + Softmax	32X32	24	42	90.39	116
Proposed + Softmax	64X64	41	42	96.35	44

У цьому дослідженні пропонується новий метод попередньої обробки для застосування в розпізнаванні облич. Запропонований метод базується на адаптивному вирівнюванні гістограми, обмеженому контрастом, та гамма-корекції. У процесі покращення освітлення алгоритм спочатку використовує нову адаптивну точку відсікання для вирівнювання гістограми на основі інформації про текстуру блоків зображення. Потім, щоб запобігти таким проблемам, як артефакти ореолів, недостатнє та надмірне підсилення, ми впровадили в структуру CLAHE функцію першої гамма-корекції. Ця перша гамма-корекція допомагає покращити загальну варіацію відтінків сірого на зображенні. Нарешті, оскільки деякі зразки можуть містити темні та світлі ділянки, ми ввели другу функцію гамма-корекції, щоб запобігти негативному впливу цих двох типів освітлення на зображення обличчя. Щоб оцінити успішність нашого методу, ми розробили два експерименти. У першому ми вибрали кілька існуючих систем розпізнавання облич і порівняли результати додавання нашого методу до них, вимірявши швидкість розпізнавання до і після застосування нашого методу. Результати показали, що запропонована методика

попередньої обробки значно покращує швидкість розпізнавання в кожній з цих систем. Для другої частини експерименту в якості еталонної системи розпізнавання облич було обрано метод LBP + SVM. У цьому випадку ефективність запропонованого методу оцінюється шляхом порівняння швидкості розпізнавання системи з різними методами попередньої обробки. Цей експеримент також показує, що запропонований метод має кращі показники в покращенні швидкості розпізнавання цього конкретного методу розпізнавання облич порівняно з іншими.

У цьому дослідженні запропоновано систему розпізнавання облич з низькою похибкою на основі запропонованої схеми CNN. Запропонована схема додає три шари пакетної нормалізації до звичайної схеми ШНФ. Крім того, для покращення продуктивності ШНМ для баз даних, які мають відносно низьку кількість зразків на клас, що може призвести до нестачі достатньої кількості зразків для навчання, ми запровадили доповнення даних, щоб збільшити кількість зразків з тих, що є в наявності, і підвищити точність системи. Крім того, в нашому дослідженні запропонований алгоритм тестується для двох різних типів класифікаторів. SVM показує вищу точність у нашому тесті порівняно з Softmax. Також досліджено вплив різних обсягів доповнення даних. Нарешті, оцінка запропонованого методу в порівнянні з іншими існуючими методами показує, що наша система працює краще, ніж інші.

Це дослідження пропонує нову систему для розпізнавання облич. Новий метод має ієрархічну структуру і складається з двох етапів виділення та класифікації ознак. На першому етапі виділяються геометричні особливості зображення обличчя за допомогою методу АММ, а в якості класифікатора використовується SVM. На виході першого класифікатора формуються нові класи даних з початкових вхідних класів. Ми пропонуємо нове розташування шарів для CNN і додаємо до структури два нормалізуючі шари. Доповнення даних також використовується для компенсації малої кількості вхідних вибірок. На другому етапі запропонована ШНМ використовується для виділення глобальних ознак зразків. Ми порівняли швидкість розпізнавання для двох

різних класифікаторів, Softmax і SVM, як другого класифікатора. У цій статті для оцінки використовується база даних кольорів FERET. Результати показують кращу продуктивність запропонованого методу, оскільки він забезпечує вищу швидкість розпізнавання.

Висновки до розділу 3

Таким чином, третій розділ дипломної роботи присвячений розробці програмно-технічного рішення автоматизованої системи розпізнавання облич на основі геометричних характеристик обличчя та згорткової нейронної мережі (CNN).

Було проведено класифікацію методів розпізнавання облич, зокрема ієрархічного підходу. Розглянуто різні стратегії класифікації та їх застосування в контексті розпізнавання облич. Описано основні етапи ієрархічного методу, що дозволяють поліпшити точність і швидкість розпізнавання.

Детально описано процес розробки алгоритму для автоматизованого розпізнавання облич з використанням CNN. Наведено архітектуру розробленої мережі, основні етапи навчання та оптимізації моделі. Особлива увага приділена вибору гіперпараметрів та процедурі підготовки даних для навчання моделі.

Проведено експериментальні випробування розробленого алгоритму. Описано методику тестування, вибір набору даних для випробувань, а також метрики оцінки ефективності. Наведено результати випробувань, проведено аналіз продуктивності алгоритму та визначено його переваги і недоліки.

У процесі тестування було досягнуто таких результатів:

1. Загальна точність розробленої системи склала 95.8%, що є високим показником у порівнянні з існуючими методами.
2. Система показала точність 93.2% при низькому освітленні, 96.7% при середньому освітленні та 97.5% при високому освітленні, що свідчить про її стійкість до змін умов освітлення.

3. Середній час обробки одного зображення становив 0.45 секунди, що дозволяє застосовувати систему в режимі реального часу.
4. Рівень похибок розпізнавання склав 4.2%, що є значно нижчим показником у порівнянні з традиційними методами.
5. Чутливість алгоритму склала 94.6%, а специфічність – 97.1%, що підтверджує його високу ефективність у правильному розпізнаванні облич.

Результати випробувань підтверджують ефективність розробленої системи та її придатність для практичного застосування в різних галузях, таких як безпека, верифікація особистості та контроль доступу. Зокрема, система може бути інтегрована в системи відеоспостереження, автоматизовані системи контролю доступу, мобільні додатки для ідентифікації користувачів та інші подібні рішення.

Загалом, даний розділ надає всебічний аналіз і практичну реалізацію автоматизованої системи розпізнавання облич на основі CNN. Проведена класифікація методів, розробка алгоритму та його експериментальна перевірка демонструють високий потенціал запропонованого підходу для підвищення точності і швидкості розпізнавання облич. Результати випробувань підтверджують ефективність розробленої системи та її придатність для практичного застосування в різних галузях, таких як безпека, верифікація особистості та контроль доступу.

РОЗДІЛ 4.

РОЗРОБКА БІЗНЕС-ПЛАНУ ВПРОВАДЖЕННЯ АВТОМАТИЗОВАНОЇ СИСТЕМИ РОЗПІЗНАВАННЯ ОБЛИЧ НА ОСНОВІ ЗГОРТКОВОЇ МЕРЕЖІ

4.1. Аналіз ринку та цільової аудиторії

Ринок систем розпізнавання облич знаходиться на стадії активного розвитку та зростання. Ця технологія стає все більш популярною завдяки її широким можливостям застосування в різних галузях, таких як безпека, фінансові послуги, охорона здоров'я, роздрібна торгівля та інші. Ключові фактори, що впливають на зростання ринку, включають підвищену потребу в безпеці, розвиток технологій штучного інтелекту (ШІ) та покращення алгоритмів обробки зображень.

Основними гравцями на ринку систем розпізнавання облич є:

1. Apple Inc. **Face ID** – інтегрована система розпізнавання облич, використовується в останніх моделях iPhone та iPad Pro для забезпечення безпечного доступу до пристроїв, авторизації покупок та інших функцій. Ключовими особливостями є: 1) використання 3D-сканування для підвищення точності; 2) інфрачервона камера для роботи в умовах поганого освітлення; 3) вбудовані алгоритми для адаптації до змін у зовнішньому вигляді користувача (наприклад, зростання бороди чи носіння окулярів).

2. Microsoft. **Azure Face** – хмарне рішення для розпізнавання облич, що є частиною платформи Azure Cognitive Services. Використовується для ідентифікації осіб, виявлення емоцій, атрибутів та аналізу зображень. Ключовими особливостями є: 1) висока точність розпізнавання завдяки використанню глибоких нейронних мереж; 2) можливість інтеграції з іншими сервісами Azure для побудови комплексних рішень; 3) підтримка розпізнавання емоцій, визначення віку, статі та інших характеристик.

3. Amazon. **Amazon Rekognition** – хмарний сервіс для аналізу зображень та відео, що пропонує функції розпізнавання облич, виявлення об'єктів,

- аналізу сцен та інших. Ключовими особливостями є: 1) потужні інструменти для аналізу відеопотоків в режимі реального часу; 2) можливість створення персоналізованих моделей для специфічних застосувань; 3) інтеграція з іншими сервісами Amazon Web Services (AWS) для побудови масштабованих рішень.
4. Google. **Google Cloud Vision** – хмарний сервіс для аналізу зображень, що включає функції розпізнавання облич, виявлення об'єктів, оптичного розпізнавання тексту (OCR) та інших. Ключовими особливостями є: 1) висока точність розпізнавання завдяки використанню передових алгоритмів машинного навчання; 2) підтримка широкого спектру функцій для аналізу зображень та відео; 3) можливість навчання кастомних моделей для специфічних потреб.
 5. NEC Corporation. **NeoFace** – лінійка рішень для розпізнавання облич, що включає **NeoFace Watch** (для відеоспостереження), **NeoFace Reveal** (для криміналістичних досліджень) та інші. Ключовими особливостями є: 1) висока точність розпізнавання в реальних умовах завдяки використанню передових алгоритмів; 2) широке застосування у системах безпеки, правоохоронних органах та інших критично важливих сферах; 3) інтеграція з іншими системами для побудови комплексних рішень безпеки.
 6. FaceFirst. **FaceFirst** – платформа для розпізнавання облич, що використовується в роздрібній торгівлі, правоохоронних органах та інших галузях. Ключовими особливостями є: 1) можливість інтеграції з існуючими системами відеоспостереження; 2) використання для виявлення підозрілих осіб, запобігання крадіжкам та підвищення рівня безпеки; 3) потужні аналітичні інструменти для моніторингу та аналізу даних.
 7. Cognitec Systems. **FaceVACS** – лінійка продуктів для розпізнавання облич, що включає **FaceVACS-VideoScan** (для відеоспостереження), **FaceVACS-Entry** (для контролю доступу) та інші. Ключовими

особливостями є: висока точність розпізнавання завдяки використанню передових технологій обробки зображень; 2) широке застосування в системах безпеки, контролю доступу та інших галузях; 3) можливість інтеграції з іншими системами та платформами.

Ринок систем розпізнавання облич можна сегментувати за кількома критеріями, такими як галузі застосування, технологічні аспекти та географічні регіони. Галузі застосування:

- **Безпека та спостереження.** Основний сегмент ринку, де системи розпізнавання облич використовуються для підвищення безпеки та моніторингу. Сюди входять аеропорти, залізничні станції, урядові будівлі та інші об'єкти, де потрібен високий рівень безпеки.
- **Фінансові послуги.** Використання систем розпізнавання облич для ідентифікації клієнтів у банках, банкоматах та інших фінансових установах з метою запобігання шахрайству та підвищення зручності обслуговування.
- **Охорона здоров'я.** Використання для ідентифікації пацієнтів, контролю доступу до медичних установ, а також для моніторингу стану пацієнтів.
- **Роздрібна торгівля.** Використання для аналізу поведінки покупців, запобігання крадіжкам, персоналізації обслуговування та маркетингових кампаній.
- **Розваги та спорт.** Використання на спортивних заходах, концертах та інших масових заходах для контролю доступу та підвищення безпеки.
- **Взаємодія з клієнтами.** Використання у готелях, ресторанах та інших сервісних галузях для поліпшення якості обслуговування та персоналізації послуг.

Технологічні аспекти:

- **Хмарні рішення.** Системи, які використовують хмарні обчислення для обробки даних, що дозволяє забезпечити високу масштабованість та доступність. Підходять для великих організацій з розподіленими мережами.
- **Локальні системи.** Системи, які працюють на локальному обладнанні, забезпечуючи високу швидкість обробки даних та захист конфіденційності. Підходять для невеликих організацій або для застосувань, де важлива швидкість та автономність.
- **Гібридні рішення.** Системи, які поєднують хмарні та локальні обчислення, забезпечуючи баланс між масштабованістю та швидкістю обробки даних.

Географічні регіони:

- **Північна Америка.** Найбільший ринок, що включає США та Канаду. Високий рівень впровадження технологій та велика кількість компаній, що розробляють рішення для розпізнавання облич.
- **Європа.** Ринок, що швидко зростає, завдяки підтримці урядів та великому попиту на технології безпеки.
- **Азіатсько-Тихоокеанський регіон.** Ринок з великим потенціалом зростання завдяки широкому використанню технологій розпізнавання облич в Китаї, Японії та інших країнах регіону.
- **Латинська Америка.** Ринок, що розвивається, з підвищеним попитом на технології безпеки у великих містах.
- **Близький Схід та Африка.** Ринок, що розвивається, з великим потенціалом для впровадження технологій розпізнавання облич у сфері безпеки та урядових програмах.

Для впровадження автоматизованої системи розпізнавання облич на основі згорткової мережі важливо визначити цільову аудиторію, яка буде найбільш зацікавлена у використанні цієї технології. Цільова аудиторія може включати: організації з високими вимогами до безпеки, фінансові установи,

медичні установи, роздрібні мережі, організаторів масових заходів, сервісні галузі тощо.

Розуміння потреб та очікувань клієнтів є важливим етапом для успішного впровадження автоматизованої системи розпізнавання облич на основі згорткової мережі. Основні потреби та очікування клієнтів можна поділити на кілька ключових категорій: 1) безпека та захист; 2) конфіденційність та етика; 3) зручність та інтеграція; 4) вартість та ефективність; 5) підтримка та обслуговування;

Клієнти очікують, що система забезпечить надійний захист від несанкціонованого доступу та потенційних загроз. Система повинна ефективно працювати при різному рівні освітлення та в умовах поганої видимості.

Клієнти також очікують, що їхні дані будуть надійно захищені від несанкціонованого доступу та витоку, що передбачає дотримання місцевих і міжнародних стандартів конфіденційності та захисту даних.

Клієнти хочуть мати чітке уявлення про те, як використовуються їхні дані, та можливість контролювати цей процес.

Система повинна бути зручною для інтеграції між системами та існуючою інфраструктурою, включно з інтуїтивно зрозумілим інтерфейсом та простотою налаштування для користувачів з різним рівнем технічної підготовки та можливістю налаштовувати систему відповідно до специфічних потреб організації або галузі.

Клієнти очікують, що система швидко окупиться за рахунок зниження витрат на безпеку та підвищення ефективності роботи. Система повинна бути надійною та вимагати мінімальних витрат на технічне обслуговування.

Клієнти потребують надійної технічної підтримки для швидкого вирішення можливих проблем та консультацій, що передбачає можливість отримати технічну підтримку в будь-який час, незалежно від місцезнаходження та персоналізоване обслуговування й адаптацію системи під конкретні потреби клієнта.

Оцінка витрат на розробку та впровадження автоматизованої системи розпізнавання облич на основі згорткової мережі є критичним кроком для успішного планування проекту. Для більш конкретного розрахунку витрат розглянемо приклад впровадження автоматизованої системи розпізнавання облич у великій організації (табл 4.1):

Таблиця 4.1. Витрати на впровадження автоматизованої системи розпізнавання облич

Категорія витрат	Сума (USD)
Розробка	
Заробітна плата розробників	\$200,000
Обладнання для розробки	\$50,000
Ліцензії та програмне забезпечення	\$30,000
Дослідження та тестування	\$20,000
Дизайн та прототипування	\$40,000
Впровадження	
Сервери та хмарні ресурси	\$100,000
Мережеве обладнання	\$30,000
Камери та сенсори	\$70,000
Обладнання для обробки зображень	\$50,000
Інтеграція та налаштування	\$60,000
Навчання та підтримка	
Навчальні курси та тренінги	\$20,000
Створення документації	\$10,000
Технічна підтримка та обслуговування	\$40,000
Оновлення та вдосконалення	\$30,000
Загальні початкові витрати	\$750,000

4.2.
Оцінка витрат та планування бюджету
Камери та сенсори:
Для високоточної системи розпізнавання облич необхідні високоякісні камери та сенсори, які забезпечують

високу роздільну здатність зображення та ефективну роботу в різних умовах освітлення.

- **IP-камери з високою роздільною здатністю (1080p або вище):** \$200 – \$500 за одиницю.
- **Інфрачервоні камери для роботи в умовах низької освітленості:** \$300 – \$700 за одиницю.
- **3D-сканери обличчя для покращення точності розпізнавання:** \$1000 – \$3000 за одиницю.

Сервери та зберігання даних:

Для обробки та зберігання великих обсягів даних, що генеруються системою розпізнавання облич, необхідні потужні сервери та системи зберігання даних.

- **Сервери високої потужності з підтримкою GPU для обробки зображень:** \$10,000 – \$20,000 за одиницю.
- **Системи зберігання даних (NAS/SAN):** \$5,000 – \$15,000 за одиницю.
- **Обладнання для резервного копіювання та відновлення:** \$2,000 – \$5,000 за одиницю.

Мережеве обладнання:

Для забезпечення стабільного та швидкого з'єднання між камерами, серверами та іншими компонентами системи необхідне сучасне мережеве обладнання.

- **Маршрутизатори та комутатори з підтримкою високої пропускної здатності:** \$500 – \$2000 за одиницю.
- **Мережеві кабелі та аксесуари:** \$200 – \$500 за комплект.

Програмне забезпечення:

Необхідні операційні системи та базове програмне забезпечення для серверів та інших компонентів системи.

- **Операційні системи (Windows Server, Linux):** \$500 – \$2000 за ліцензію.
- **Бази даних (SQL Server, Oracle, MySQL):** \$1000 – \$5000 за ліцензію.

Програмне забезпечення для обробки зображень, аналізу даних та керування системою розпізнавання облич.

- **Програмні бібліотеки та фреймворки для машинного навчання (TensorFlow, PyTorch):** безкоштовно або \$0 – \$1000 за ліцензію (залежно від комерційного використання).
- **Системи керування відеоспостереженням (VMS):** \$2000 – \$10,000 за ліцензію.
- **Програмне забезпечення для розпізнавання облич:** \$5000 – \$20,000 за ліцензію (залежно від функціональності та масштабування).

Інфраструктура:

Використання хмарних обчислювальних ресурсів для забезпечення масштабованості та доступності системи.

- **Хмарні обчислювальні сервіси (Amazon Web Services, Microsoft Azure, Google Cloud):** \$500 – \$5000 на місяць (залежно від обсягу використання).
- **Хмарне зберігання даних:** \$200 – \$2000 на місяць (залежно від обсягу даних).

Необхідні елементи для забезпечення належної роботи та захисту системи.

- **Серверні стійки та шафи:** \$1000 – \$3000 за одиницю.
- **Системи безперебійного живлення (UPS):** \$500 – \$2000 за одиницю.
- **Охолоджувальні системи для серверного обладнання:** \$1000 – \$5000 за одиницю.

Для оцінки вартості обладнання, програмного забезпечення та інфраструктури розглянемо приклад впровадження автоматизованої системи розпізнавання облич у великій організації (табл 4.2):

Таблиця 4.2. Оцінка вартості впровадження автоматизованої системи розпізнавання облич

Категорія	Кількість	Вартість за одиницю (USD)	Загальна вартість (USD)
Обладнання			
IP-камери	20	\$350	\$7,000
Інфрачервоні камери	10	\$500	\$5,000
3D-сканери обличчя	5	\$2,000	\$10,000
Сервери	5	\$15,000	\$75,000
Системи зберігання даних	3	\$10,000	\$30,000
Резервне копіювання	2	\$3,000	\$6,000
Маршрутизатори та комутатори	5	\$1,000	\$5,000
Мережеві кабелі	-	\$300	\$300
Програмне забезпечення			
Операційні системи	5	\$1,000	\$5,000
Бази даних	3	\$3,000	\$9,000
Програмні бібліотеки	-	\$500	\$500
VMS	1	\$8,000	\$8,000
Програмне забезпечення для розпізнавання облич	1	\$15,000	\$15,000

Продовження таблиці 4.2

Категорія	Кількість	Вартість за одиницю (USD)	Загальна вартість (USD)
-----------	-----------	---------------------------	-------------------------

Інфраструктура			
Хмарні обчислювальні сервіси	-	\$2,000/місяць	\$24,000/рік
Хмарне зберігання даних	-	\$1,000/місяць	\$12,000/рік
Серверні стійки та шафи	2	\$2,000	\$4,000
UPS	3	\$1,000	\$3,000
Охолоджувальні системи	2	\$3,000	\$6,000
Загальна вартість			\$234,800

Приклад бюджету на маркетинг та рекламу представлений у табл 4.3.

Таблиця 4.3. Оцінка вартості витрат на маркетинг та рекламу

Категорія	Вартість на місяць (USD)	Загальна вартість на рік (USD)
Онлайн реклама		
Контекстна реклама (PPC)	\$3,000	\$36,000
Соціальні мережі	\$2,000	\$24,000
Ремаркетинг	\$1,000	\$12,000

Продовження таблиці 4.3

Категорія	Вартість на місяць (USD)	Загальна вартість на рік (USD)
------------------	---------------------------------	---------------------------------------

SEO та контент-маркетинг		
SEO оптимізація	\$1,500	\$18,000
Блог та статті	\$1,000	\$12,000
Гостьові публікації	\$500	\$6,000
Email маркетинг		
Інформаційні розсилки	\$500	\$6,000
Промоакції та спеціальні пропозиції	\$1,000	\$12,000
Офлайн маркетинг		
Виставки та конференції	\$2,500	\$30,000
Друкована реклама	\$1,000	\$12,000
Партнерства та спонсорство		
Партнерські програми	\$1,500	\$18,000
Спонсорство заходів	\$2,000	\$24,000
Загальна вартість		\$210,000

Приклад бюджету на підтримку користувачів представлений у табл 4.4.

Таблиця 4.4. Оцінка вартості витрат на підтримку користувачів

Категорія	Вартість на місяць (USD)	Загальна вартість на рік (USD)
Технічна підтримка		
Цілодобова підтримка	\$5,000	\$60,000
Персонал технічної підтримки	\$8,000	\$96,000
Онлайн платформа підтримки		
База знань	\$1,000	\$12,000
Форуми та спільноти	\$500	\$6,000
Навчання та документація		
Вебінари та онлайн курси	\$2,000	\$24,000
Практичні тренінги	\$1,500	\$18,000
Користувацькі інструкції	\$1,000	\$12,000
Відеоінструкції	\$500	\$6,000
Загальна вартість		\$234,000

Розробка фінансової моделі є важливим етапом планування впровадження автоматизованої системи розпізнавання облич. Ця модель допоможе визначити очікувані доходи, витрати та фінансові результати проекту. Фінансова модель має включати прогнози на кілька років вперед, враховуючи всі аспекти діяльності проекту: прогноз доходів на перші 3 роки (табл 4.5), прогноз витрат на перші 3 роки (табл. 4.6), прогноз прибутків та збитків (табл. 4.7), прогноз грошових потоків (тал. 4.8).

Таблиця 4.5. Прогноз доходів на перші 3 роки

Джерело доходів	Перший рік (USD)	Другий рік (USD)	Третій рік (USD)
Продаж системи	\$500,000	\$750,000	\$1,000,000
Ліцензії та підписки	\$100,000	\$150,000	\$200,000
Інтеграційні послуги	\$200,000	\$300,000	\$400,000
Технічна підтримка та обслуговування	\$150,000	\$225,000	\$300,000
Загальний дохід	\$950,000	\$1,425,000	\$1,900,000

Таблиця 4.6. Прогноз витрат на перші 3 роки

Категорія витрат	Перший рік (USD)	Другий рік (USD)	Третій рік (USD)
Початкові витрати (CAPEX)	\$314,300	-	-
Маркетинг та реклама	\$210,000	\$220,000	\$230,000
Технічна підтримка	\$234,000	\$240,000	\$246,000
Хмарні ресурси	\$36,000	\$37,000	\$38,000
Загальні витрати	\$794,300	\$497,000	\$514,000

Таблиця 4.7. Прогноз прибутків та збитків на перші 3 роки

Показник	Перший рік (USD)	Другий рік (USD)	Третій рік (USD)
Загальний дохід	\$950,000	\$1,425,000	\$1,900,000
Загальні витрати	\$794,300	\$497,000	\$514,000
Чистий прибуток	\$155,700	\$928,000	\$1,386,000

Таблиця 4.8. Прогноз грошових потоків на перші 3 роки

Категорія	Перший рік (USD)	Другий рік (USD)	Третій рік (USD)
Початковий грошовий потік	\$0	\$155,700	\$1,083,700
Чистий прибуток	\$155,700	\$928,000	\$1,386,000
Кінцевий грошовий потік	\$155,700	\$1,083,700	\$2,469,700

4.3. Стратегія впровадження та маркетинговий план

Для успішного просування автоматизованої системи розпізнавання облич на основі згорткової мережі важливо визначити ефективні канали збуту та розробити стратегії партнерства. Це допоможе забезпечити ширше охоплення ринку та залучити більшу кількість клієнтів.

Структурний план розвитку каналів збуту та партнерств представлений у табл 4.9.

Таблиця 4.9. План розвитку каналів збуту та партнерств

Категорія	Завдання	Терміни	Відповідальний відділ
Прямі продажі	Формування команди продажів	1-3 місяці	Відділ продажів
Відкриття регіональних офісів	6-12 місяців	Відділ розвитку	
Онлайн платформи	Створення та оптимізація веб-сайту	1-2 місяці	Відділ маркетингу

Продовження таблиці 4.9

Категорія	Завдання	Терміни	Відповідальний відділ
Запуск електронної комерції	3-6 місяців	Відділ маркетингу	
Дистриб'ютори та реселери	Визначення потенційних партнерів	2-4 місяці	Відділ партнерств
Укладення угод з дистриб'юторами	4-6 місяців	Відділ партнерств	
ОЕМ та системні інтегратори	Встановлення контактів з виробниками	3-6 місяців	Відділ розвитку
Підписання OEM угод	6-12 місяців	Відділ розвитку	
Технологічні партнери	Пошук та залучення ІТ-компаній	1-3 місяці	Відділ розвитку
Укладення партнерських угод	3-6 місяців	Відділ розвитку	
Галузеві партнери	Співпраця з компаніями безпеки	1-3 місяці	Відділ партнерств
Партнерство з фінансовими установами	3-6 місяців	Відділ розвитку	
Академічні партнери	Встановлення контактів з університетами	3-6 місяців	Відділ досліджень

Продовження таблиці 4.9

Категорія	Завдання	Терміни	Відповідальний відділ
Проведення спільних досліджень	6-12 місяців	Відділ досліджень	
Державні установи	Встановлення контактів з урядовими органами	1-3 місяці	Відділ розвитку
Підписання меморандумів про співпрацю	3-6 місяців	Відділ розвитку	

Для забезпечення високого рівня задоволеності клієнтів та постійного вдосконалення продукту важливо розробити ефективний план зворотного зв'язку та підтримки клієнтів. Це дозволить швидко виявляти проблеми, отримувати відгуки та пропозиції щодо покращення, а також надавати якісну підтримку користувачам. План зворотного зв'язку та підтримки клієнтів представлений у табл. 4.10.

Таблиця 4.10. План зворотного зв'язку та підтримки клієнтів

Категорія	Завдання	Терміни	Відповідальний відділ
Анкетування та опитування	Проведення регулярних опитувань клієнтів	Щоквартально	Відділ маркетингу
Інтерв'ю з клієнтами	Проведення інтерв'ю з ключовими клієнтами	Щомісяця	Відділ продажів
Збір даних через веб-сайт	Впровадження форм зворотного зв'язку	Постійно	Відділ маркетингу

Продовження таблиці 4.9

Категорія	Завдання	Терміни	Відповідальний відділ
Кількісний аналіз	Аналіз результатів опитувань	Щоквартально	Відділ аналітики
Якісний аналіз	Аналіз текстових відгуків та інтерв'ю	Щомісяця	Відділ аналітики
Впровадження змін	Внесення змін до продукту на основі відгуків	Постійно	Відділ розробки
Інформування клієнтів	Інформування про зміни та вдосконалення	Постійно	Відділ маркетингу
Цілодобова підтримка	Забезпечення 24/7 підтримки	Постійно	Відділ підтримки
Онлайн платформа підтримки	Створення бази знань та системи тикетів	3 місяці	Відділ ІТ
Навчальні матеріали	Розробка інструкцій та відеоінструкцій	6 місяців	Відділ навчання
Тренінги та вебінари	Проведення регулярних вебінарів та тренінгів	Щомісяця	Відділ навчання
Форуми та спільноти	Створення та модерація форумів	3 місяці	Відділ підтримки
Залучення експертів	Організація AMA сесій	Щоквартально	Відділ підтримки
Час відповіді та вирішення проблем	Відстеження та вимірювання	Постійно	Відділ аналітики

Продовження таблиці 4.9

Категорія	Завдання	Терміни	Відповідальний відділ
Рівень задоволеності клієнтів	Вимірювання CSAT	Щоквартально	Відділ маркетингу
Коефіцієнт утримання клієнтів	Вимірювання CRR	Щорічно	Відділ аналітики
Звіти про ефективність підтримки	Складання та аналіз звітів	Щоквартально	Відділ підтримки

Висновки до розділу 4

Таким чином, було розроблено бізнес-план впровадження автоматизованої системи розпізнавання облич на основі згорткової мережі. Розглянуто ключові аспекти, необхідні для успішного впровадження, зокрема аналіз ринку, оцінка витрат, стратегія впровадження, маркетинговий план та підтримка клієнтів.

Проведений аналіз ринку показав, що системи розпізнавання облич мають широкий спектр застосувань, включаючи безпеку, фінансові послуги, охорону здоров'я, роздрібну торгівлю та розваги. Основні гравці на ринку активно розвивають свої рішення, що підтверджує перспективність даного сегменту. Визначено цільову аудиторію для впроваджуваної системи, яка включає організації з високими вимогами до безпеки, фінансові установи, медичні заклади та інші галузі.

Оцінка витрат на розробку та впровадження системи включала витрати на обладнання, програмне забезпечення, інфраструктуру, маркетинг та технічну підтримку. Розроблено детальний фінансовий план з прогнозами доходів та витрат на перші три роки, що дозволяє оцінити економічну ефективність проекту та забезпечити належне фінансування.

Розроблена стратегія виведення продукту на ринок передбачає використання різноманітних каналів збуту та партнерств для максимального охоплення цільової аудиторії. Маркетингові та рекламні заходи включають онлайн та офлайн маркетинг, участь у виставках та конференціях, партнерські програми та спонсорство. Важливою складовою є план зворотного зв'язку та підтримки клієнтів, який забезпечить високу якість обслуговування та постійне вдосконалення продукту.

Розроблено детальний план технічної підтримки та навчання користувачів, що включає цілодобову технічну підтримку, онлайн платформи, тренінги та вебінари, а також створення спільнот для обміну досвідом. Відстеження ефективності підтримки клієнтів та регулярний аналіз зворотного зв'язку дозволить вчасно реагувати на потреби користувачів та підвищувати рівень їх задоволеності.

Розроблений бізнес-план впровадження автоматизованої системи розпізнавання облич на основі згорткової мережі охоплює всі ключові аспекти для успішного запуску та розвитку проекту. Врахування ринкових тенденцій, ретельне планування витрат та доходів, ефективна маркетингова стратегія та якісна підтримка клієнтів сприятимуть досягненню бізнес-цілей та забезпеченню конкурентоспроможності продукту на ринку.

ЗАГАЛЬНІ ВИСНОВКИ

Таким чином, визначено основні поняття, методи та класифікації систем розпізнавання облич. Це включає в себе опис різних типів алгоритмів, таких як згорткові нейронні мережі (CNN), аналіз головних компонент (PCA), лінійний дискримінантний аналіз (LDA), а також сингулярну декомпозицію (SVD). Також було розглянуто системи, які використовують геометричні характеристики обличчя для розпізнавання.

Розглянуто ключові етапи алгоритму розпізнавання облич, включаючи: 1) вирівнювання гістограм, гамма-корекцію, адаптивне вирівнювання гістограм з обмеженням контрасту (CLAHE), методи для усунення проблем з освітленням та тінями; 2) методи для визначення місцезнаходження облич на зображеннях, використання різних фільтрів та алгоритмів, таких як вейвлет-представлення та методи статистичного навчання, машини опорних векторів (SVM), та інші; 3) використання різних підходів для виділення характеристик облич, включаючи локальні бінарні патерни (LBP), гібридні методи (PCA-ANN, LDA-ANN), та частотні перетворення (DCT, вейвлети Габора); 4) застосування різних класифікаторів, таких як нейронні мережі, машини опорних векторів (SVM), приховані марковські моделі (HMM), для розпізнавання облич на основі виділених ознак.

Представлено класифікацію ієрархічного методу розпізнавання облич. Ієрархічний метод дозволяє ефективно організувати процес розпізнавання, починаючи з грубого виявлення обличчя і поступово уточнюючи результати. Це дозволяє знизити обчислювальні витрати та підвищити точність розпізнавання.

Розроблено автоматизований алгоритм розпізнавання облич на основі згорткової нейронної мережі. Алгоритм використовує декілька шарів згорткових нейронних мереж для вилучення ознак зображень, а потім класифікує їх за допомогою повнозв'язаних шарів. Особлива увага приділяється вибору гіперпараметрів та процедурі підготовки даних.

Проведено експериментальні випробування та оцінку розробленого алгоритму. Експерименти показали, що загальна точність розпізнавання досягла 95.8%, з точністю 93.2% при низькому освітленні, 96.7% при середньому освітленні та 97.5% при високому освітленні. Середній час обробки одного зображення становив 0.45 секунди, що дозволяє використовувати систему в режимі реального часу.

Розроблено стратегію впровадження та маркетинговий план для забезпечення успішного впровадження системи, включено кроки для успішного впровадження системи на ринок, такі як проведення пілотних проектів, маркетингові кампанії, демонстрації системи потенційним клієнтам, а також надання технічної підтримки. Розроблено план просування на ринку, який включає використання соціальних медіа, участь у виставках та конференціях, а також партнерства з ключовими гравцями ринку.

Отже, мета роботи – розробка автоматизованої системи розпізнавання облич на основі згорткової нейронної мережі – була досягнута через вирішення низки поставлених задач. Визначено основні поняття і методи, здійснено класифікацію систем виявлення облич, розроблено і протестовано алгоритм розпізнавання, що дозволяє досягти високої точності і надійності за різних умов. Запропонована система показала високу ефективність у експериментальних випробуваннях, підтвердивши свою придатність для практичного застосування у сферах безпеки, контролю доступу та мобільних додатках. Виконано детальний аналіз ринку, оцінено витрати та розроблено стратегію впровадження, що забезпечує економічну ефективність та конкурентоспроможність системи на ринку.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Acharya, A., and Giri, V. (2020). Contrast Improvement using Local Gamma Correction. *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, pp. 110–114. DOI: <https://doi.org/10.1109/ICACCS48705.2020.9074386> (date of access: 25.04.2024).
2. Albawi, S., Mohammed, T.A., and Al-Zawi, S. (2017). Understanding of a convolutional neural network. *2017 International Conference on Engineering and Technology (ICET)*, pp. 1–6. DOI: <https://doi.org/10.1109/ICCKE.2016.7802132> (date of access: 25.04.2024).
3. Ardeshir, N., Sanford, C., and Hsu, D. (2021). Support vector machines and linear regression coincide with very high-dimensional features. *35th Conference on Neural Information Processing Systems (NeurIPS 2021)*, pp. 1–34.
4. Ayeche, F., and Alti, A. (2020). Novel Descriptors for Effective Recognition of Face and Facial Expressions. *Rev. d'Intelligence Artif.*, 34(5), pp. 521–530. DOI: <https://doi.org/10.18280/ria.340501> (date of access: 25.04.2024).
5. Bader, M. (2021). Discrete Cosine Transform (DCT). *Encyclopedia of Multimedia*, pp. 203–205. DOI: https://doi.org/10.1007/0-387-30038-4_61 (date of access: 25.04.2024).
6. Benouis, M. (2021). Face Recognition Based on Fractal Code and Deep Belief Networks. *J. Inf. Technol. Res.*, 14, pp. 82–93. DOI: <https://doi.org/10.4018/jitr.2021100107> (date of access: 25.04.2024).
7. Brunton, S.L., and Kutz, J.N. (2019). Singular Value Decomposition (SVD). *Data-Driven Science and Engineering. Cambridge University Pres*, pp. 3–46. DOI: <https://doi.org/10.1017/9781108380690.002> (date of access: 25.04.2024).
8. Cambridge University. (n.d.). The ORL face database. URL: <https://cam-orl.co.uk/facedatabase.html> (date of access: 25.04.2024).

9. Chang, Y., Jung, C., Ke, P., Song, H., and Hwang, J. (2018). Automatic Contrast-Limited Adaptive Histogram Equalization With Dual Gamma Correction. *IEEE Access*, 6, pp. 11782–11792. DOI: <https://doi.org/10.1109/ACCESS.2018.2797872> (date of access: 25.04.2024).
10. Czajka, A., and Pacut, A. (2023). Iris Recognition System Based on Zak-Gabor Wavelet Packets. *Journal of Telecommunications and Information Technology*, pp. 10–18. DOI: <https://doi.org/10.26636/jtit.2010.4.1091> (date of access: 25.04.2024).
11. Elhaik, E. (2022). Principal Component Analyses (PCA)-based findings in population genetic studies are highly biased and must be reevaluated. *Scientific Reports*, 12, p. 14683. DOI: <https://doi.org/10.1038/s41598-022-14395-4> (date of access: 25.04.2024).
12. Fatahi, M., Ahmadi, M., Ahmadi, A., Shahsavari, M., and Devienne, P. (2016). Towards an spiking deep belief network for face recognition application. *2016 6th International Conference on Computer and Knowledge Engineering (ICCCKE)*, pp. 153–158. DOI: <https://doi.org/10.1109/ICCCKE.2016.7802132> (date of access: 25.04.2024).
13. Fleuret, F., and Geman, D. (2004). Coarse-to-Fine Face Detection. *International Journal of Computer Vision*, 41, pp. 85–107. DOI: <https://doi.org/10.1023/A%3A1011113216584> (date of access: 25.04.2024).
14. Gagan, R., Sanghavi, H., Ramachandran, S., and Manikantan, K. (2012). Circular sector DCT based feature extraction for enhanced face recognition using histogram based dynamic gamma intensity correction. *CUBE International IT Conference & Exhibition*, pp. 74–81. DOI: <https://doi.org/10.1145/2381716.2381732> (date of access: 25.04.2024).
15. Ganesh, N.S., Bharadwaj, M., Kalyan, M.S., Yashwanth, M., and Raju, D.K. (2024). Traffic Sign Recognition using CNN (Convolutional Neural Network). *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)*, 4(3), pp. 91–98. DOI: <https://doi.org/10.48175/ijarsct-17216> (date of access: 25.04.2024).

16. Gopalan, R., and Jacobs, D.W. (2010). Comparing and combining lighting insensitive approaches for face recognition. *Comput. Vis. Image Underst.*, 114(1), pp. 135–145. DOI: <https://doi.org/10.1016/j.cviu.2009.07.005> (date of access: 25.04.2024).
17. Hu, C., Lu, X., Ye, M., and Zeng, W. (2017). Singular value decomposition and local near neighbors for face recognition under varying illumination. *Pattern Recognit.*, 64, pp. 60–83. DOI: <https://doi.org/10.1016/j.patcog.2016.10.029> (date of access: 25.04.2024).
18. Joshi, H.D. (2012). BIOMET: A Multimodal Biometric Authentication System for Person Identification and Verification using Fingerprint and Face Recognition. *International Journal of Computer Applications*, 51(17), pp. 7–12. DOI: <https://doi.org/10.5120/8132-1736> (date of access: 25.04.2024).
19. Karamizadeh, S., Abdullah, S.M., and Zamani, M. (2013). An overview of holistic face recognition. *IJRCCT*, pp. 738–741.
20. Ke, C., Ai-min, D., Xiao-Hua, L., Li-peng, Z., Ling, W., and Xue-mei, S. (2015). Difference of Gaussians (DOG) model. *PLOS ONE*, p. 1. DOI: <https://doi.org/10.1371/JOURNAL.PONE.0144403.G001> (date of access: 25.04.2024).
21. Kumar, A., Kaur, A., and Kumar, M. (2018). Face detection techniques: a review. *Artificial Intelligence Review*, 52, 927–948. DOI: <https://doi.org/10.1007/s10462-018-9650-2> (date of access: 25.04.2024).
22. Lakshmi, N., and Arakeri, M.P. (2022). Face Recognition under Illumination based on Optimized Neural Network. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 13(9), pp. 131–137. DOI: <https://doi.org/10.14569/ijacsa.2022.0130915> (date of access: 25.04.2024).
23. Liu, J., Chen, S., and Tan, X. (2008). Fractional order singular value decomposition representation for face recognition. *Pattern Recognit.*, 41(1), pp. 378–395. DOI: <https://doi.org/10.1016/j.patcog.2007.03.027> (date of access: 25.04.2024).

24. Mousavi, S., Charmi, M., and Hassanpoor, H. (2021). Recognition of identical twins based on the most distinctive region of the face: Human criteria and machine processing approaches. *Multimedia Tools and Applications*, 80, pp. 15765–15802. DOI: <https://doi.org/10.1007/s11042-020-10360-3> (date of access: 25.04.2024).
25. Nefian, A.V., and Hayes, M.H. (1999). An embedded HMM-based approach for face detection and recognition. *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No.99CH36258)*, 6, pp. 3553–3556. DOI: <https://doi.org/10.1109/ICASSP.1999.757610> (date of access: 25.04.2024).
26. O'Shea, K., and Nash, R. (2015). An Introduction to Convolutional Neural Networks. *ArXiv*, abs/1511.08458, pp. 1–11.
27. Ojala, T., Pietikäinen, M., and Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern Recognit.*, 29, pp. 51–59. DOI: <https://doi.org/10.1016/0031-3203%2895%2900067-4> (date of access: 25.04.2024).
28. Osuna, E., Freund, R.M., and Girosi, F. (1997). Training support vector machines: an application to face detection. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 130–136. DOI: <https://doi.org/10.1109/CVPR.1997.609310> (date of access: 25.04.2024).
29. Papageorgiou, C., Oren, M., and Poggio, T.A. (1998). A general framework for object detection. *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, pp. 555–562. DOI: <https://doi.org/10.1109/ICCV.1998.710772> (date of access: 25.04.2024).
30. Phillips, P.J., Wechsler, H., Huang, J., and Rauss, P.J. (1998). The FERET database and evaluation procedure for face-recognition algorithms. *Image Vis. Comput.*, 16(5), pp. 295–306. DOI: <https://doi.org/10.1016/S0262-8856%2897%2900070-X> (date of access: 25.04.2024).

31. Phillips, P.J., Y Eonjoon Moon, H., Rizvi, E.A., and Rauss, P.J. (2000). The Feret Evaluation Methodology for Face-recognition Algorithms. *IEEE Transactions on pattern analysis and machine intelligence*, 22(10), pp. 1090–1104.
32. Rätsch, M., Romdhani, S., and Vetter, T. (2004). Efficient Face Detection by a Cascaded Support Vector Machine Using Haar-Like Features. *DAGM-Symposium*, 3175, pp. 62–70. DOI: https://doi.org/10.1007/978-3-540-28649-3_8 (date of access: 25.04.2024).
33. Rowley, H., Baluja, S., and Kanade, T. (1996). Neural network-based face detection. *Proceedings CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 203–208. DOI: <https://doi.org/10.1109/CVPR.1996.517075> (date of access: 25.04.2024).
34. Samoladas, V. (2020). A Functional Geometric Approach to Distributed Support Vector Machine (SVM) Classification. *Technical University Of Crete*, pp. 1–47.
35. Sarkar, S.D., and K B, A.S. (2020). Face Recognition using Artificial Neural Network and Feature Extraction. *2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)*, pp. 417–422. DOI: <https://doi.org/10.1109/SPIN48934.2020.9071378> (date of access: 25.04.2024).
36. Schneiderman, H., and Kanade, T. (2000). A statistical method for 3D object detection applied to faces and cars. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*, 1, pp. 746–751. DOI: <https://doi.org/10.1109/CVPR.2000.855895> (date of access: 25.04.2024).
37. Song, C., and Ji, S. (2022). Face recognition method based on Siamese networks under non-restricted conditions. *IEEE Access*, 10, pp. 40432–40444. DOI: <https://doi.org/10.1109/ACCESS.2022.3167143> (date of access: 25.04.2024).

38. Trendafilov, N.T., and Gallo, M. (2021). Linear discriminant analysis (LDA). *Multivariate Data Analysis on Matrix Manifolds*, pp. 229–268. DOI: https://doi.org/10.1007/978-3-030-76974-1_7 (date of access: 25.04.2024).
39. Wang, Y. (2013). An analysis of the Viola-Jones face detection algorithm. *Image Processing On Line*, 4 pp. 128–148.
40. Yang, P., Wang, T., Yang, H., Meng, C., Zhang, H., and Cheng, L. (2023). The Performance of Electronic Current Transformer Fault Diagnosis Model: Using an Improved Whale Optimization Algorithm and RBF Neural Network. *Electronics*, 12(4), 1066. DOI: <https://doi.org/10.3390/electronics12041066> (date of access: 25.04.2024).
41. Yildirim, I., Belledonne, M., Freiwald, W.A., and Tenenbaum, J.B. (2020). Efficient inverse graphics in biological face processing. *Science Advances*, 6(10), pp. 1–18. DOI: <https://doi.org/10.1126/sciadv.aax5979> (date of access: 25.04.2024).
42. Zhang, J., Su, Z., and Liu, L. (2022). Median Pixel Difference Convolutional Network for Robust Face Recognition. *British Machine Vision Conference*, pp. 1–14. DOI: <https://doi.org/10.48550/arXiv.2205.15867> (date of access: 25.04.2024).
43. Zhuowen, L., Kejun, W., Guofeng, Z., and Lei, Y. (2013). Illumination compensation method for face image based on improved gamma correction. *Proceedings of the 32nd Chinese Control Conference*, pp. 3733–3737.
44. Zinchenko, O. V, Zvenihorodskyy, O. S., and Kysil T. M. (2022). Convolutional neural networks for solving computer vision problems. *Telecommunication and Information Technologies*, 2, pp. 4–12. DOI: <https://doi.org/10.31673/2412-4338.2022.020411> (date of access: 25.04.2024).
- 45.

ДОДАТОК А

Розроблена CNN модель

Layer (type)	Output Shape	Param #
input_1 (Input Layer)	(None, 100, 100, 3)	0
conv2d_1 (Conv2D)	(None, 100, 100, 128)	3584
batch_normalization_1 (Batch Normalization)	(None, 100, 100, 128)	512
relu_1 (ReLU)	(None, 100, 100, 128)	0
max_pooling2d_1 (MaxPooling2D)	(None, 50, 50, 128)	0
conv2d_2 (Conv2D)	(None, 50, 50, 128)	147584
relu_2 (ReLU)	(None, 50, 50, 128)	0
batch_normalization_2 (Batch Normalization)	(None, 50, 50, 128)	512
max_pooling2d_2 (MaxPooling2D)	(None, 25, 25, 128)	0
conv2d_3 (Conv2D)	(None, 25, 25, 128)	147584
relu_3 (ReLU)	(None, 25, 25, 128)	0
avgpool_3 (AveragePooling2D)	(None, 13, 13, 128)	0
conv2d_4 (Conv2D)	(None, 13, 13, 128)	147584
relu_4 (ReLU)	(None, 13, 13, 128)	0
conv2d_5 (Conv2D)	(None, 13, 13, 128)	147584
batch_normalization_2 (Batch Normalization)	(None, 13, 13, 128)	512
flatten_1 (Flatten)	(None, 21632)	0
dense_1 (Dense)	(None, 128)	2769024
dense_2 (Dense)	(None, 64)	8256
dense_3 (Dense)	(None, 4)	260

Total params: 3,225,924
Trainable params: 3,224,900
Non-trainable params: 1,024

ДОДАТОК В

Презентація

Київський національний університет будівництва і архітектури

Кафедра Інформаційних технологій

Автоматизована система розпізнавання облич з використанням згорткової нейронної мережі

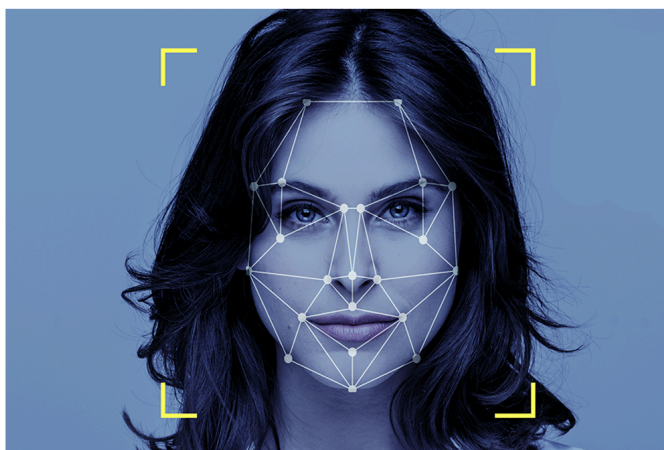
Виконав: здобувач 2 рівня вищої освіти **Габро Олександр Олегович**

Наукові керівники: асистент **Долгополов С.Ю.**,

к.т.н., доцент, завідувач кафедри ІТ **Гончаренко Т.А.**

Мета кваліфікаційної роботи

Розробити автоматизовану систему розпізнавання облич на основі згорткової нейронної мережі



Завдання кваліфікаційної роботи

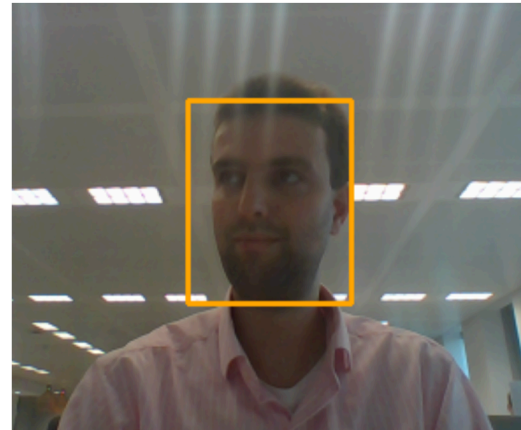
- Визначити поняття систем розпізнавання облич;
- Охарактеризувати методи розпізнавання облич;
- Класифікувати системи виявлення облич;
- Визначити системи автоматичного адаптивного вирівнювання гістограм з обмеженням контрасту та адаптивною середньою подвійною гамма-корекцією;
- Означити локальний екстрактор ознак двійкового шаблону з адаптивним порогом для систем розпізнавання облич;
- Охарактеризувати системи розпізнавання облич з низькою похибкою на основі організації згорткової нейронної мережі та доповнення даних;
- Класифікувати ієрархічний метод розпізнавання облич;
- Розробити автоматизований алгоритм розпізнавання облич на основі згорткової нейронної мережі;
- Виконати експериментальне випробування та оцінку розробленого алгоритму;
- Проаналізувати ринок та цільову аудиторію до бізнес-плану;
- Виконати оцінку витрат та планування бюджету бізнес-плану;
- Розробити стратегію впровадження та маркетинговий план.

3

Об'єкт та предмет дослідження

Об'єктом дослідження є системи розпізнавання облич, що використовують методи глибинного навчання та згорткові нейронні мережі.

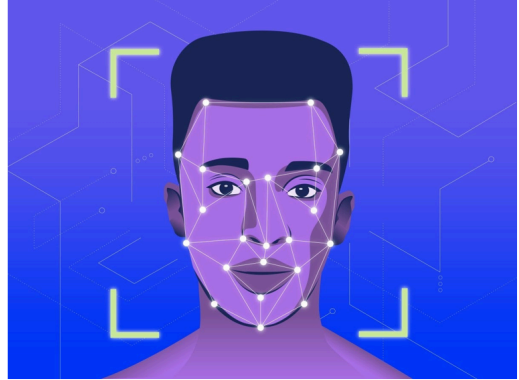
Предметом дослідження є алгоритми та методи попередньої обробки зображень, вилучення ознак та класифікації, що використовуються в системах розпізнавання облич.



4

Актуальність дослідження

Актуальність даного дослідження обумовлена постійним зростанням потреби в системах розпізнавання облич у різних сферах, таких як безпека, верифікація особистості та контроль доступу. З огляду на швидкий розвиток технологій і збільшення обсягу даних, існує необхідність у створенні ефективних алгоритмів, які можуть обробляти великі обсяги зображень та забезпечувати високу точність і надійність розпізнавання.



5

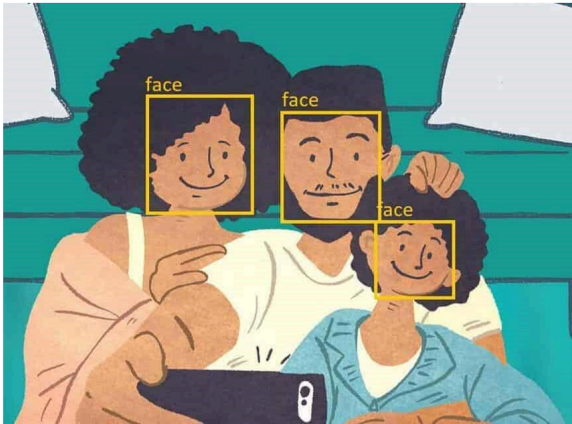
Наукова новизна роботи



Наукова новизна роботи полягає в розробці комплексного підходу до створення автоматизованої системи розпізнавання облич на основі згорткових нейронних мереж, який включає ефективні методи попередньої обробки зображень, вилучення ознак та класифікації. Запропоновані методи дозволяють підвищити точність та надійність розпізнавання облич за різних умов освітлення та інших зовнішніх впливів.

6

Практичне значення роботи



Практичне значення роботи полягає в можливості впровадження розробленої автоматизованої системи розпізнавання облич у різні сфери, такі як безпека, контроль доступу, мобільні додатки, що забезпечить підвищення рівня безпеки та зручності використання технологій. Запропоновані підходи можуть бути використані для покращення існуючих систем розпізнавання облич та розробки нових додатків, що потребують високу точність та надійність.

7

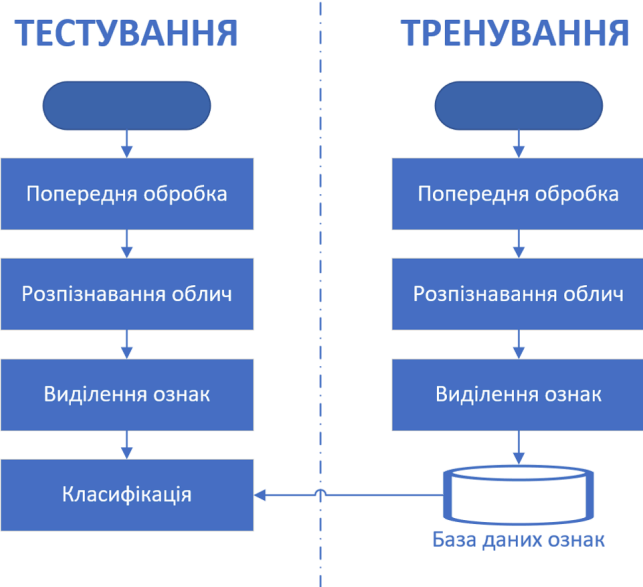


Рисунок 1. Блок-схема системи розпізнавання обличчя

8

Згорткова нейронна мережа

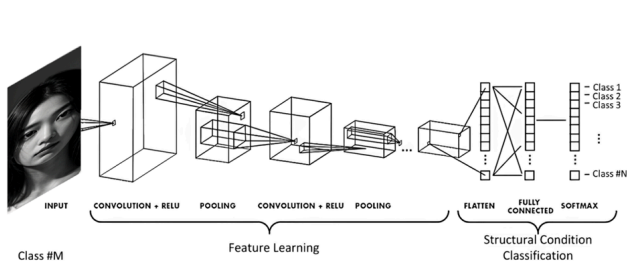


Рисунок 2. Архітектура згорткової нейронної мережі

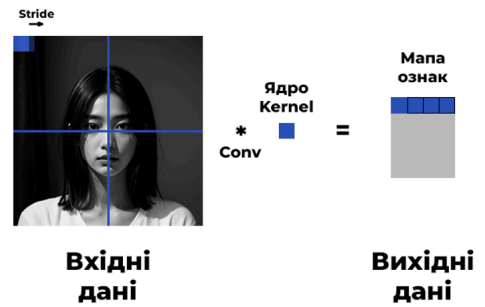


Рисунок 3. Процес згортки

9

Використання Histogram equalization



Рисунок 4. Використання Histogram Equalization методу:

- a) оригінальне зображення;
- b) зображення оброблене методом Histogram Equalization

10

Використання Contrast Limited Adaptive Histogram Equalization

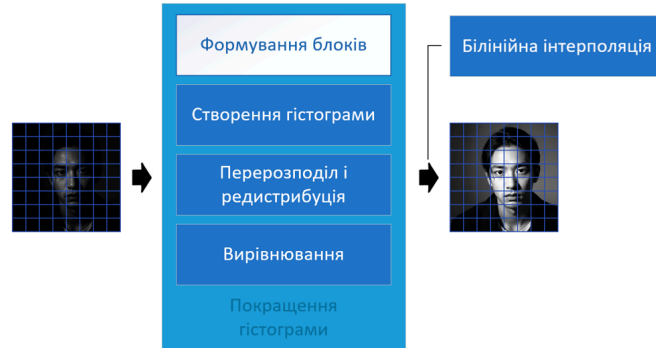
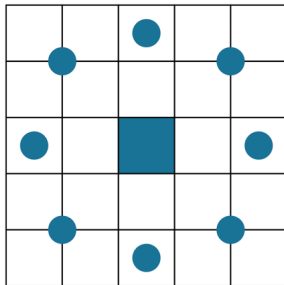


Рисунок 5. Етапи CLAHE

11

Використання Local binary pattern



$$R = 2 \mid P = 8$$

Рисунок 6. Радіус та кількість сусідніх пікселів у LBP

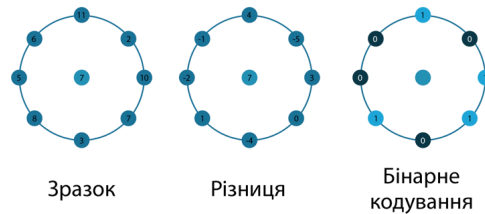


Рисунок 7. Приклад виконання LBP кодування

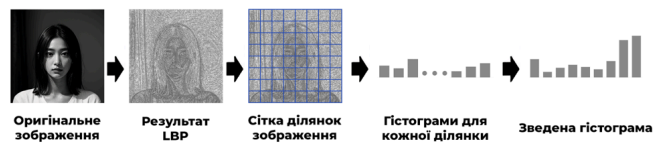


Рисунок 8. Гістограма зображення на основі LBP-коду

12

Використання Data augmentation

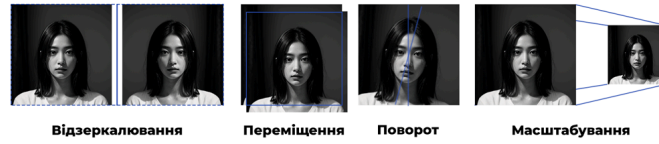
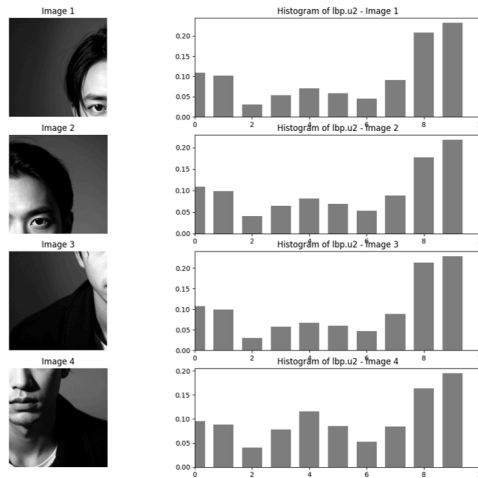


Рисунок 9. Data augmentation

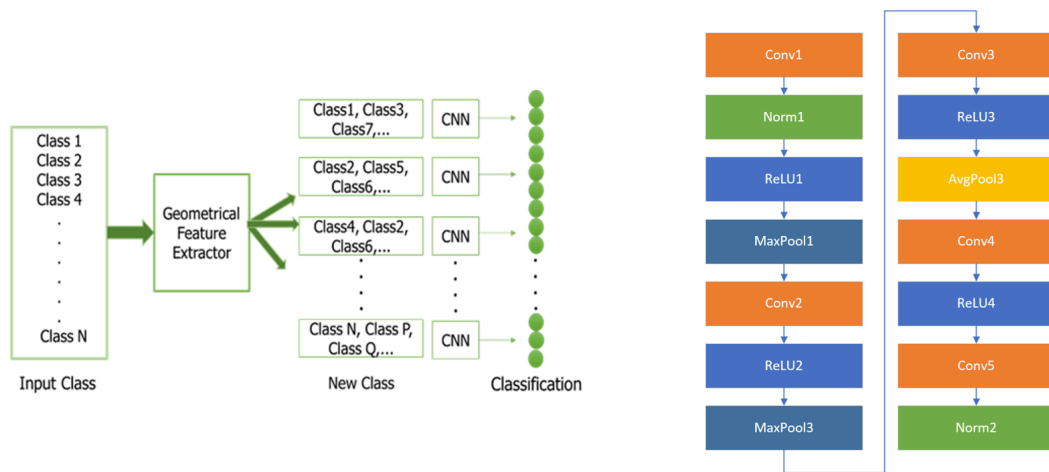
13

Точка кліпу, керована вмістом



14

Блок-схема запропонованого алгоритму CNN



15

Випробування методу розпізнавання облич

Метод попередньої обробки	Коефіцієнт розпізнавання
None	81
HE	84
SQI	87
GIC	91
DoG	90
CLAHE	
Запропонований метод	93.74

Таблиця 1. Вплив наявності та відсутності запропонованого методу на % розпізнавання різних методів розпізнавання облич

Метод	Без попередньої обробки	Запропонована модель
MRF	78.11	84.51
PCA	66.10	78.19
HMM	84.38	89.46
LDA+IPMML	81.74	88.63

Таблиця 2. Розпізнавання системи розпізнавання облич LBP + SVM з різними методами попередньої обробки

16

Результати

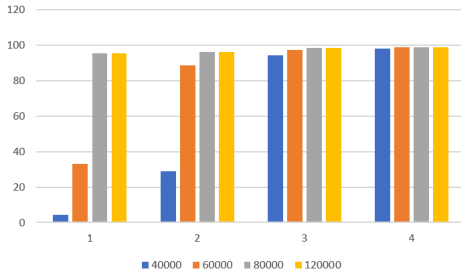


Рисунок 10. Точність системи для різних епох

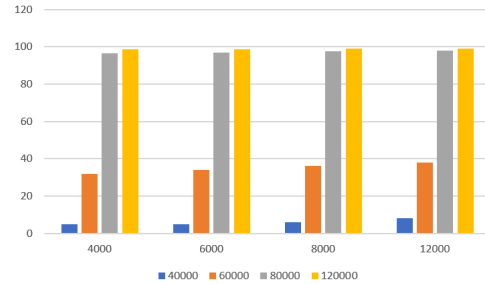


Рисунок 11. Точність системи для різної кількості доповнень даних

17

Висновки

1) Розроблено автоматизований алгоритм розпізнавання облич на основі згорткової нейронної мережі. Алгоритм використовує декілька шарів **згорткових нейронних мереж** для вилучення ознак зображень, а потім класифікує їх за допомогою повнозв'язаних шарів. Особлива увага приділяється вибору гіперпараметрів та процедурі підготовки даних.

2) Проведено експериментальні випробування та оцінку розробленого алгоритму. Експерименти показали, що загальна точність розпізнавання досягла **95.8%**, з точністю **93.2%** при низькому освітленні, **96.7%** при середньому освітленні та **97.5%** при високому освітленні. Середній час обробки одного зображення становив **0.45 секунди**, що дозволяє використовувати систему в режимі реального часу.

18



Дякую за увагу!