

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Інформаційно-аналітичних технологій та менеджменту
(повна назва)

Кафедра Інформатики
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

ДОСЛІДЖЕННЯ ТА ПОРІВНЯННЯ МЕТОДІВ ДЕТЕКТУВАННЯ
ЛЮДЕЙ НА ЗОБРАЖЕННЯХ
(тема)

Виконав:
студент 2 курсу, групи ІНФМ-20-1
(прізвище, ініціали)

Спеціальності 122 Комп'ютерні науки
(код і повна назва спеціальності)

Тип програми освітньо-професійна

Освітня програма Інформатика
(повна назва освітньої програми)

Керівник _____
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____
(підпис)

Кобилін О.А.
(прізвище, ініціали)

2021 р.

Харківський національний університет радіоелектроніки

Факультет Інформаційно-аналітичних технологій та менеджменту
(повна назва)Кафедра Інформатики
(повна назва)Рівень вищої освіти другий (магістерський)Спеціальність 122 Комп'ютерні науки
(код і повна назва)Тип програми освітньо-професійнаОсвітня програма Інформатика
(повна назва освітньої програми)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

«_____» _____ 2021 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУстудентові Товстолес Дарині Юріївні
(прізвище, ім'я, по батькові)1. Тема роботи Дослідження та порівняння методів детектування людей на зображенняхзатверджена наказом по університету від «22» жовтня _____ 2021 року №1574Ст.

2. Термін подання студентом роботи до екзаменаційної комісії _____ 2021 р.

3. Вихідні дані до роботи витяги з робіт вчених, які займалися схожими проблемами, опис відомих методів, перелік використовуваних програмних засобів: мова програмування C#, бібліотеки програмного забезпечення, системне програмне забезпечення macOS Catalina.

4. Перелік питань, що потрібно опрацювати в роботі _____

1. Розглянути різні підходи та алгоритми для обробки зображень для розпізнавання облич2. Визначити параметри, що варіюються, методів обраних для дослідження та порівняння.3. Розробити алгоритми формування тестового та навчального наборів для проведення дослідження.4. Провести експериментальне дослідження вибраних методів на підготовленому вихідному матеріалі.5. Проаналізувати та порівняти результати показані різними методами, дати рекомендації щодо їх використання.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) актуальність теми, постановка задачі, способи побудови вибірки (перехресна перевірка), імовірнісні способи, детерміновані способи, метод головних компонент, метод локальних бінарних шаблонів, база ORL, база FERET, створення навчальних та тестових наборів, набори, сформовані для навчання та перевірки класифікаторів, результати, показані на основі бази ORL, результати, показані на основі бази FERET, порівняння результатів методів, висновки.

6. Консультанти розділів роботи (п.6 включається до завдання за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата
Консультант з дотримання діючих стандартів та норм	Доцент Белова Н.В.		

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	22.10.2021	
2	Аналіз завдання, підбір літератури	23.10.21-25.10.21	
3	Аналіз літератури з досліджуваної проблеми	25.10.21-26.10.21	
4	Аналіз технічних засобів	26.10.21-31.10.21	
5	Розробка методу	31.10.21-01.11.21	
6	Програмна реалізація	01.11.21-03.11.21	
7	Оформлення пояснювальної записки	03.11.21-15.11.21	
8	Перевірка на плагіат	16.11.2021	
9	Рецензування	16.11.2021	
10	Підготовка презентації та доповіді	17.11.2021	
11	Занесення роботи в електронний архів	18.11.2021	
12	Попередній захист кваліфікаційної роботи	01.12.2021	

Дата видачі завдання 22 жовтня 2021 р.

Студент _____
(підпис)

Керівник роботи _____
(підпис) _____
(посада, прізвище, ініціали)

РЕФЕРАТ/ABSTRACT

Пояснювальна записка до кваліфікаційної роботи: 76 с., 4 табл., 20 рис., 1 дод., 47 джерел.

КОМП'ЮТЕРНИЙ ЗІР, СИСТЕМА РОЗПІЗНАВАННЯ ОБЛИЧ, РОЗПІЗНАВАННЯ ОБРАЗІВ, ВИЯВЛЕННЯ ОБЛИЧ, РОЗПІЗНАВАННЯ ОБЛИЧ

Об'єктом дослідження є системи розпізнавання облич.

Метою дослідження проведеного в рамках даної роботи є підвищення ефективності та якості роботи систем розпізнавання облич за рахунок вибору найбільш ефективного методу розпізнавання та підбору необхідного в конкретних умовах поєднання параметрів цього методу.

У роботі розглядаються деякі аспекти використання алгоритмів ідентифікації людини за зображенням обличчя, які можуть впливати на ефективність такого розпізнавання.

За результатами роботи були обрані методи дослідження такі, як метод головних компонент та метод локальних бінарних шаблонів. Було виявлено особливості вихідного набору даних, при яких певний метод ідентифікації має найбільші переваги. На підставі результатів дослідження було сформульовано рекомендації щодо використання методів розпізнавання облич.

COMPUTER VISION, FACE RECOGNITION SYSTEM, PATTERN RECOGNITION, FACE DETECTION, FACE RECOGNITION

The object of research is face recognition systems.

The purpose of the study conducted in this work is to improve the efficiency and quality of facial recognition systems by choosing the most effective method of face recognition and selection of the necessary in specific conditions, a combination of parameters of this method.

The paper considers some aspects of the use of algorithms for identifying a person by facial image, which may affect the effectiveness of such recognition.

According to the results of the paper, research methods were chosen as the method of principal components and the method of local binary templates. Features of the original data set were identified, according to which a certain method of identification has the greatest advantages. Based on the results of the research, recommendations were made for the use of face recognition methods.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів	7
Вступ.....	8
1 Огляд існуючих підходів та методів.....	11
1.1 Класифікація методів розпізнавання облич.....	11
1.1.1 Класифікація за основними класами розв’язуваних задач...	11
1.1.2 Класифікація по інваріантності методів до варіацій вихідного зображення та способи її досягнення	12
1.1.3 Класифікація за алгоритмічними особливостями методів розпізнавання.....	14
1.1.4 Класифікація за ознаками об’єктів, що використовуються..	19
1.2 Способи побудови вибірки.....	30
1.2.1 Перехресна перевірка	30
1.2.2 Імовірнісні способи	32
1.2.3 Детерміновані способи.....	34
1.3 Постановка задачі дослідження	35
2 Теоретичний базис: досліджувані методи розпізнавання облич	36
2.1 Метод головних компонент.....	36
2.2 Метод локальних бінарних шаблонів.....	42
3 Експериментальне дослідження на основі методу головних компонент та методу локальних бінарних шаблонів	46
3.1 Вихідні дані: бази зображень облич.....	46
3.1.1 База ORL.....	46
3.1.2 База FERET.....	47
3.2 Формування вибірки	47
3.2.1 Створення навчальних та тестових наборів.....	49
3.2.2 Набори, сформовані для навчання та перевірки класифікаторів	50
3.3 Програмне забезпечення дослідження	51

	6
3.3.1 Програмна платформа	51
3.3.2 Середовище розробки	52
3.3.3 Мова програмування	53
3.3.4 Розширення для роботи із зображеннями	53
3.4 Результати, отримані в ході дослідження	54
3.4.1 Результати, показані на матеріалі бази ORL.....	64
Висновки.....	66
Перелік джерел посилання	70

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

МГК – метод головних компонент

РСА – Principal Component Analysis

ЛБШ – локальний бінарний шаблон

LBP – Local Binary Pattern

НМ – нейронна мережа

ЗНМ – згорткова нейронна мережа

БНМ – багатошарова нейронна мережа

ПММ – приховані марківські моделі

SVM – Support Vector Machine

ORL – Olivetti Research Laboratory

FERET – Face Recognition Technology

DARPA – Defense Advanced Research Projects Agency

CLR – Common Language Runtime

OpenCV – Open Source Computer Vision Library

ВСТУП

Задача виділення обличчя людини в природній або штучній обстановці та подальшій ідентифікації завжди знаходилася в ряді найпріоритетніших завдань для дослідників, які працюють у галузі систем машинного зору та штучного інтелекту.

У багатьох дослідженнях, проведених у межах розв'язання цієї задачі протягом кількох десятиліть, отримані найрізноманітніші результати, як для методик, що відрізняються одна від одної, так і для методик, що використовують одні й ті ж алгоритми, так і для систем, що ґрунтуються на одному й тому ж алгоритмі.

Звичайно, алгоритми та методи, що використовуються для ідентифікації зображень обличчя, постійно вдосконалюються, однак, і серед уже існуючих методів, складно знайти об'єктивне та достовірне порівняння їх ефективності та результатних показників.

Для об'єктивного порівняння різних методів розпізнавання зображень необхідно використовувати єдину методику перевірки та аналогічну вибірку зображень. Крім того, сам спосіб поділу вибірки на тестову та навчальну частини повинен мати мінімальну залежність від пред'явленого набору даних.

Для того, щоб виділити найбільш ефективний метод ідентифікації зображення обличчя необхідно застосувати математично обґрунтований метод оцінки того, наскільки успішні досліджувані методи на незалежному наборі даних. Крім того, слід оцінити вплив кольорової нормалізації зображення та виключити вплив його геометричної нормалізації. А також розглянути залежність результатів від обсягу тестової вибірки для того, щоб виділити найефективніший метод ідентифікації, мінімізувати вплив зовнішніх параметрів на оцінку результатів.

У роботі розглядаються деякі аспекти використання алгоритмів ідентифікації особистості за зображенням обличчя, які можуть впливати на

ефективність такого розпізнавання. Для досягнення більшої об'єктивності було вирішено такі питання:

- розроблено алгоритм формування тестової та навчальної вибірки з неоднорідного вихідного набору даних;
- з використанням цього алгоритму проведено дослідження та порівняння ефективності методів ідентифікації людини за зображенням обличчя;
- виявлено особливості вихідного набору даних, при яких певний метод ідентифікації має великі переваги;
- визначено, які вхідні параметри методів більшою мірою впливають на ефективність та точність розпізнавання методів;
- на підставі результатів дослідження було сформульовано рекомендації щодо використання методів розпізнавання облич.

Сьогодні системи розпізнавання облич застосовують у трьох областях. Перша – боротьба із криміналом. Фоторобот злочинця звіряється із зображеннями в базах даних або проводиться зйомка людського потоку в людних місцях, і обличчя людей в реальному часі порівнюються з обличчями порушників, які перебувають у розшуку. Систему можна використовувати і у боротьбі з банківським шахрайством.

Актуальність дослідження полягає в тому, що розпізнавання працює у сучасних системах контролю доступу. Зазвичай технологія діє у зв'язці з електронними перепустками: зразок фото на них порівнюється з моделлю, отриманою в результаті зйомки людини, що входить до будинка. У цьому випадку фіксація обличчя – найзручніший спосіб ідентифікації. Обличчя завжди доступне для зйомки. Фотографування відбувається миттєво і процедура не вимагає від людини жодних додаткових «дій».

Третя сфера застосування, що останнім часом швидко розвивається, — маркетинг. Людина підходить до інтерактивного рекламного стенду, система фіксує його обличчя, будує тривимірну модель та визначає стать та вік. У результаті клієнт бачить на стенді лише рекламу, яка потенційно йому цікава.

Є й інше маркетингове завдання: удосконалити традиційні програми лояльності. Раніше зрозуміти, що організацію відвідав постійний клієнт, можна було тільки на касі, коли він дістане з гаманця карту знижок. Тепер, помістивши в залі магазину камеру та встановивши систему розпізнавання облич, можна впізнавати кожного.

1 ОГЛЯД ІСНУЮЧИХ ПІДХОДІВ ТА МЕТОДІВ

1.1 Класифікація методів розпізнавання облич

Задача розпізнавання людини по зображенню обличчя поділяється на три великих класи: пошук у великих базах даних, контроль доступу та контроль фотографій в документах. Вони відрізняються за вимогами, що надаються до систем розпізнавання, та за способами вирішення, тому представляють собою окремі класи. У рамках цієї роботи актуальним є перший клас завдань. Оскільки будь-який архів, не так важливо відеоархів це чи просто набір зображень, є насамперед базою даних [1].

1.1.1 Класифікація за основними класами розв'язуваних задач

Висуваються різні вимоги до помилок першого і другого роду для таких класів. Помилка першого роду – ситуація, коли об'єкт заданого класу не розпізнається (пропускається) системою. Помилка другого роду – ситуація, коли об'єкт заданого класу приймається за об'єкт іншого класу. В експериментальній частині роботи основним питанням є факт успішного розпізнавання, тобто ситуація, коли система відносить об'єкт до того класу, якому він відповідає, у всіх інших ситуаціях результат випробування вважається невдачею.

Слід також відзначити відмінність понять верифікації та розпізнавання (ідентифікації). У задачі верифікації невідомий об'єкт заявляє, що належить до деякого, відомому системі, класу. Система підтверджує чи спростовує це. У процесі розпізнавання потрібно віднести об'єкт, що належить невідомому

класу до одного з відомих чи зробити висновок, що цей об'єкт не належить жодному з відомих класів.

При пошуку зображень у базах даних великого обсягу висуваються високі вимоги до помилок першого роду. Система розпізнавання повинна знаходити зображення, що відповідають певній людині, за можливості не пропустивши жодного такого зображення. Допускається наявність у результуючій вибірці невеликої кількості інших людей.

Часто у великих базах даних потрібно знайти такі зображення, які найбільше схожі на ті, що подаються на вхід. Пошук повинен проводитись за прийнятний період часу. Один із варіантів полягає у тому, що в базі даних зберігаються деякі заздалегідь вилучені набори важливих ознак, що визначають зображення з граничною однозначністю [2]. У цій ситуації вимоги до точності не є настільки критичними як у завданнях контролю доступу або контролю документів.

До цього класу, насамперед, слід віднести метод основних компонент (метод «власних облич»). Коефіцієнти, отримані при розкладанні вхідного зображення на головні компоненти, використовуються для порівняння зображень шляхом обчислення Евклідової відстані або на основі метрики Махаланобіса з використанням визначень Гауса.

1.1.2 Класифікація по інваріантності методів до варіацій вихідного зображення та способи її досягнення

Якість, що є у методів розпізнавання, інваріантна до таких варіацій зображення як масштаб, зміщення, умови освітлення тощо. Крім цього, є деякі способи обробки зображення з метою досягнення подібної інваріантності при використанні інших методів.

Неокогнітронам та нейронним мережам високого порядку властива інваріантність до зміни масштабу. Це означає, що ці мережі не вимагають

спеціального навчання на зображеннях з різними масштабами об'єкта, що дозволяє згодом розпізнавати об'єкт на зображеннях з іншими варіантами масштабу. Передбачається, що в інших методах цей параметр апроксимується при геометричній нормалізації зображення, інакше у навчальний набір необхідно включати зображення, що мають різні варіації масштабу.

Для багатьох методів критичним є чинник зміни в освітленні, за його впливу результативність розпізнавання значно знижується. Вирівнювання гістограми яскравостей (світлова нормалізація) дозволяє якоюсь мірою вирішити цю проблему, хоча найчастіше передобробка не дозволяє усунути цей фактор.

У деяких роботах використовується алгоритм виділення лінійних підпросторів, що дозволяє провести успішну класифікацію за різних умов освітлення, але для досягнення цього ефекту необхідно формувати навчальний набір із зображень, на яких присутні різні варіанти освітлення.

Існують також різні евристичні підходи. Наприклад, в одному з варіантів методу основних компонент, перші 2-3 компоненти, в основному, містять інформацію про відмінності в освітленні, тому вони не враховуються при порівнянні [3].

В рамках цієї роботи буде розглянуто вплив світлової нормалізації (вирівнювання гістограми яскравостей) на ефективність розпізнавання досліджуваних методів.

У багатьох методах необхідна попередня обробка, що модифікує положення обличчя так, що воно набуває стандартної орієнтації в просторі (лінія очей вирівнюється горизонтально, а вісь симетрії – вертикально).

Серед методів розпізнавання зображень немає таких, на ефективність яких впливають великі зміни ракурсу [4]. З цією проблемою дозволяє впоратися лише навчання на зображеннях обличчя у різних ракурсах.

Розробляються методи, що дозволяють подолати цей недолік, в основі лежить синтез зображення в інших ракурсах за одним прикладом зображення в довільному ракурсі.

При використанні різних за рівнем геометричної нормалізації баз зображень можна оцінити вплив цих факторів на досліджувані методи.

Алгоритми розпізнавання працюють таким чином, що кожне обличчя окремої людини є певним класом. Внутрішньокласові відмінності, наприклад: окуляри, борода, зачіска чи емоції ускладнюють процес розпізнавання.

Певною мірою описана проблема долається виділенням областей обличчя, не схильних до таких змін. Також це може вирішуватися у процесі навчання.

1.1.3 Класифікація за алгоритмічними особливостями методів розпізнавання

Практично все різноманіття алгоритмів розпізнавання орієнтоване на виділення основних ознак і їх наявності у тестового прикладу. Можна виділити три групи методів, які відрізняються способами порівняння зображень.

У першій групі методів набір ознак (у найпростішому випадку вихідне зображення) представляє собою точку в просторі ознак, де значення кожної ознаки (наприклад, яскравість окремого пікселя, значення головної компоненти, коефіцієнт частотного перетворення тощо) є координатою уздовж деякої осі простору ознак. Процедура порівняння ґрунтується на розділенні простору ознак на області, що належать до однакових класів [5].

Для цього, наприклад, може бути обчислена відстань від невідомого образу до всіх інших образів за допомогою будь-якої метрики. Клас може бути представлений центром кластера, тоді відстань від невідомого образу обчислюється до центрів всіх кластерів, як у методі основних компонент. Лінійний дискримінант будує набір лінійних поверхонь, що розділяють, вважаючи, що класи лінійно розділені в просторі ознак. Різновид нейронних

мереж – багатошарові перцептрони, що дозволяють будувати розділяючі поверхні будь-якої складності.

На рисунку 1.1 показано схему класифікації способів порівняння зображень.

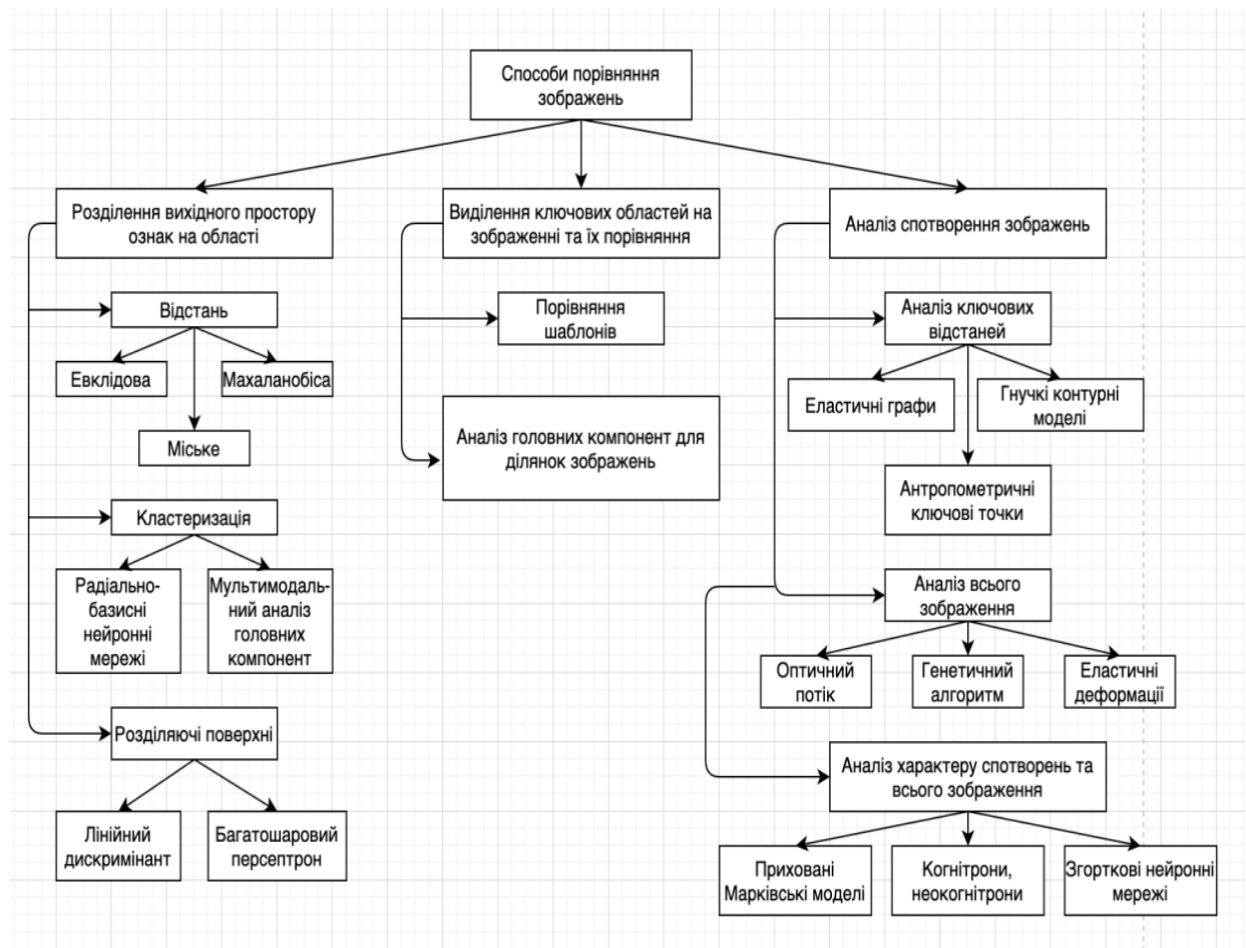


Рисунок 1.1 – Схема класифікації способів порівняння зображень

Незважаючи на те, що найдосконаліші методи можуть бути кращим способом розбити вихідний простір на області, цього недостатньо, оскільки для реальних об'єктів потрібна величезна кількість областей, що розділяють (і навчальних прикладів), щоб врахувати всілякі способи зміни зображень об'єктів. Це пов'язано з тим, що навіть незначна, з людської точки зору, зміна зображення (наприклад, ракурс, освітлення або наявність бороди) може дати становище в просторі ознак, що дуже далеко від вихідного. І система в цьому випадку може зреагувати не на однакову людину, а, наприклад, на однаковий

ракурс, вважаючи зображення іншої людини в тому ж ракурсі найбільш схожим на невідоме [6]. Ні застосування різних перетворень для початкового представлення зображення, ні передобробка не спроможна вирішити цю проблему загального випадку.

Наступна група методів частково долає цей недолік. На зображенні знаходяться важливі області обличчя (наприклад, області очей, брів, носа, губ). Потім за допомогою першого способу проводиться порівняння кожної області, результат кожного такого порівняння робить внесок у остаточний результат. До таких методів відноситься наприклад порівняння за зразками, різні способи аналізу основних компонентів за блоками зображень.

У третій групі враховується топологічне спотворення зображення. У методі порівняння еластичних графів зображення накладається на набір ключових точок, пов'язаних між собою дугами. На невідомому зображенні знаходяться точки, що відповідають вихідним, а потім вимірюється сумарна зміна розмірів дуг. У гнучких контурних моделях витягується контур обрисів обличчя, потім порівнюються форми контурів різних зображень [7]. В інших дослідженнях використовуються нейронні мережі та приховані Марківські моделі для аналізу співвідношень відстаней між областями очей, носа та рота.

У вищеописаних методах після знаходження ключових областей чи контурів решта інформації не враховується, що знижує точність.

Вищенаведені методи порівнюють лише сумарне спотворення, не намагаючись врахувати його характер, у цьому полягає їх недолік. Характер спотворення зображення несе важливу інформацію для розпізнавання зображення, тому методи, що наводяться нижче, мають кращу точність розпізнавання.

До таких методів належать псевдодвовимірні приховані Марківські моделі, згорткові нейронні мережі, когнітрони та неокогнітрони.

В алгоритмах технології FaceIta корпорації Visionics зображення розбивається на блоки та аналізується взаємне розташування таких блоків. Ці методи для порівняння використовують як характеристики ділянок (блоків)

зображень (набір яскравостей пікселів, коефіцієнтів частотних перетворень, головні компоненти блоків зображення тощо), так і взаємне розташування таких ділянок. Причому характер спотворень і вміст ділянок засвоюється у процесі навчання. У процесі розпізнавання спотворення проводиться не до кожного прикладу тренувального набору, а до узагальненої моделі класу [8].

За характером використання навчального набору методи розпізнавання облич можна поділити на два великі класи. У методах першого класу в процесі налаштування не використовують навчальні приклади. У таких методах всі параметри задаються вручну і зазвичай підбираються шляхом різних експериментів. Наприклад, криміналістична ідентифікація людини в особі із застосуванням ключових точок.

На противагу цьому другий клас методів різною мірою для отримання ознак використовує аналіз навчальної вибірки, що є набір типових об'єктів необхідних класів. Тому такі методи обчислювально більш трудомісткі та його розвиток став можливим зі зростанням потужності обчислювальної техніки. Це метод основних компонентів, лінійні дискримінантні методи, метод порівняння еластичних графів [9]. Проте в них основний наголос зроблено на автоматизоване вилучення ознак, а не відносин між ними та визначення їх важливості (зважування). Існують дослідження, де було зроблено спробу обчислити вагові параметри для еластичного порівняння графів. Лінійний дискримінантний аналіз використовувався для аналізу ознак та перетворення їх такого виду, що максимізує міжкласові та мінімізує внутрішньокласові відмінності.

За повнотою використання інформації зображення методи розпізнавання зображення обличчя можна розділити на три класи.

Перший клас використовує вихідне зображення тільки для початкового виділення будь-яких ключових точок, відстаней між такими точками, співвідношення відстаней або інших геометричних характеристик і надалі до вихідного зображення (наприклад, напівтонового) не звертається. Така інформація досить достовірно характеризує конкретну людину, тому

використовується в криміналістиці для впізнання людей. Однак показано, що загалом геометричної інформації недостатньо для розпізнавання людини зображення обличчя.

Другий клас методів використовує все вихідне зображення. Наприклад, метод аналізу основних компонентів, лінійний дискримінантний аналіз оперують вихідним напівтоновим зображенням як багатовимірним вектором. До цього класу також відносяться різні нейронні мережі, що приймають на вхід зображення цілком. Багато джерел сходяться на думці, що для загальних випадків розпізнавання потрібно аналізувати всю інформацію, надану на зображенні (наприклад, напівтонове зображення) [10].

Третій клас є поєднанням двох попередніх. Тут використовуються і геометричні характеристики, і елементи, наприклад напівтонового зображення. Наприклад, метод порівняння еластичних графів. Вершини графа характеризують особливі точки зображення, дуги – відстані між цими точками, крім цього, кожній вершині приписаний набір локальних характеристик зображення (джет). Іншим прикладом є спосіб порівняння стандартів. На зображенні обличчя виділяються особливі області, такі як очі, ніс, рот і зображення порівнюється по цих областях.

Методи розпізнавання характеризуються так само різним співвідношенням у них аналітичних та емпіричних способів розв'язання задачі розпізнавання. Метод головних компонент та лінійні дискримінанти, наприклад, засновані на розв'язанні систем матричних рівнянь та знаходжень відстаней між образами у різних метриках. У методі порівняння еластичних графів ключові точки відзначаються на навчальному наборі зображень облич, виходячи з емпіричних міркувань.

Особливе місце посідають нейромережні методи. Під час налаштування мережі використовуються аналітичні методи градієнтного спуску. Але таке налаштування не дає ні гарантії досягнення мінімуму помилки, ні оцінки відхилення отриманої помилки від оптимальної. На відміну від статистичних методів, рішення, одержувані з допомогою нейронних мереж, носять не

імовірнісний, а правдоподібний характер, який визначається навчальним набором [11].

Схожа ситуація із прихованими Марківськими моделями. Налаштування параметрів кожної моделі здійснюється математичними оптимізаційними методами, але не дають точного рішення.

Методи поділяються на два класи за способом порівняння невідомого зображення з тренувальним набором (або базою зображень).

Перший клас методів використовує тренувальний набір побудови узагальненої моделі кожного класу обличчя (чи всіх класів відразу). Це різні нейронні мережі, лінійні дискримінанти, приховані марківські моделі.

Другий клас методів безпосередньо порівнює невідоме зображення з кожним зображенням з тренувальної вибірки. Це спосіб основних компонент, порівняння стандартів, способи порівняння еластичних графів, аналіз оптичного потоку.

1.1.4 Класифікація за ознаками об'єктів, що використовуються

На початку 2000-х років виділилося дві базові групи методів порівняння облич: «холістичні підходи» (або «глобальні підходи») та «ознакові підходи» (або «структурні підходи») [12]. У першому випадку обличчя розглядаються як цілісні зображення, що порівнюються між собою. У другому випадку, із зображення обличчя виділяються локальні ознаки, такі як інформація про абсолютне або взаємне розташування очей, носа, рота і т.п. Обидві групи мають свої недоліки: холістичні підходи показали в цілому більшу надійність у близьких до ідеальних умов розпізнавання порівняно з ознаковими підходами, проте ознакові підходи краще зарекомендували себе в ситуації, коли окремі частини обличчя чомусь не видно (наприклад, закриті елементами одягу або окулярами).

Структурні підходи до розпізнавання об'єктів засновані на автоматичному підстроюванні до властивостей оброблюваних даних, що дозволяє розробити систему, що розпізнає, з прийнятними характеристиками. До другої категорії належать методи, що характеризуються скороченням розмірності даних.

Складність завдання розпізнавання об'єктів полягає в тому, що характерні ознаки є багатовимірними векторами, що в сукупності описують деякі властивості об'єктів, що розглядаються, наприклад: текстури, набір моментів, гістограми і т. п. Серед адаптивних методів можна виділити наступні різновиди:

- метод опорних векторів;
- штучні нейронні мережі;
- приховані Марківські моделі;
- метод гнучкого порівняння на графах.

Метод опорних векторів (Support Vector Machines, SVM) використовується для розв'язання задач класифікації та регресійного аналізу [13]. Застосування SVM до розв'язання задачі полягає у пошуку гіперплощини в ознаковому просторі, що відокремлює класи зображень «обличчя» та «не обличчя».

Основна ідея методу – переведення вихідних векторів у простір більш високої розмірності та пошук роздільної гіперплощини з максимальним зазором у цьому просторі. Дві паралельні гіперплощини будуються по обидва боки гіперплощини, що розділяє наші класи.

Розділяючою гіперплощиною буде гіперплощина, що максимізує відстань до двох паралельних гіперплощин. Алгоритм працює в припущенні, що чим більша різниця чи відстань між цими паралельними гіперплощинами, тим меншою буде середня помилка класифікатора.

Недоліки: суттєві тимчасові витрати на налаштування, необхідність великого обсягу пам'яті.

Архітектура багат шарової нейронної мережі (БНМ) складається з послідовно з'єднаних шарів, де нейрон кожного шару своїми входами пов'язаний з усіма нейронами попереднього шару, а виходами наступного. НМ з двома вирішальними шарами може з будь-якою точністю апроксимувати будь-яку багатовимірну функцію. НМ з одним вирішальним шаром здатна формувати лінійні розділяючі поверхні, що сильно звужує коло завдань, ними розв'язуваних, зокрема, така мережа не зможе вирішити задачу типу «що виключає або». НМ з нелінійною функцією активації та двома вирішальними шарами дозволяє формувати будь-які опуклі області у просторі рішень, а з трьома вирішальними шарами – області будь-якої складності, у тому числі й невиконувальні. При цьому БНМ не втрачає своєї узагальнюючої здатності. Навчаються БНМ за допомогою алгоритму зворотного розповсюдження помилки, що є методом градієнтного спуску в просторі терезів з метою мінімізації сумарної помилки мережі [14]. При цьому помилки (точніше, величини корекції ваги) поширюються у зворотному напрямку від входів до виходів, крізь ваги, що з'єднують нейрони. Найпростіше застосування одношарової НМ (названої автоасоціативною пам'яттю) полягає в навчанні мережі відновлювати зображення, що подаються. Подаючи на вхід тестове зображення та обчислюючи якість реконструйованого зображення, можна оцінити, наскільки мережа розпізнала вхідне зображення. Позитивні властивості цього методу полягають у тому, що мережа може відновлювати спотворені та зашумлені зображення, але для більш серйозних цілей не підходить. БНМ також використовується для безпосередньої класифікації зображень – на вхід подається або саме зображення у якомусь вигляді, або набір раніше вилучених ключових характеристик зображення, на виході нейрон з максимальною активністю вказує на належність до розпізнаного класу.

Нейрон з максимальною активністю (тут перший) вказує на належність до розпізнаного класу. Якщо ця активність нижча за деякий поріг, то вважається, що поданий образ не відноситься до жодного з відомих класів.

Процес навчання встановлює відповідність образів, що подаються на вхід, з належністю до певного класу. Це називається «навчанням з учителем». У застосуванні до розпізнавання людини зображення обличчя такий підхід хороший для завдань контролю доступу невеликої групи облич. Він забезпечує безпосереднє порівняння мережею самих образів, але зі збільшенням числа класів час навчання та роботи мережі зростає експоненційно, і тому для таких завдань, як пошук схожої людини у великій базі даних, вимагає отримання компактного набору ключових характеристик, на їх основі можна здійснювати пошук [15].

Зокрема, БНМ може використовуватися для класифікації зображень облич з урахуванням таких характеристик, як відстані між специфічними частинами обличчя (ніс, рот, очі). Також існують гібридні системи, наприклад, об'єднання з марківською моделлю. У класичній БНМ міжшарові нейронні сполуки пов'язані, і зображення представлене у вигляді одновимірного вектора, хоча воно двовимірне. Архітектура згорткової НМ спрямована на подолання цих недоліків. У ній використовуються локальні рецепторні поля (забезпечують локальну двовимірну зв'язність нейронів), загальні ваги (забезпечують детектування деяких рис у будь-якому місці зображення) та ієрархічна організація з просторовими підвиборками (spatial subsampling). Згорткова НМ забезпечує часткову стійкість до змін масштабу, зсувів, поворотів, спотворень [16]. Архітектура ЗНМ складається з багатьох шарів, кожен з них має кілька площин, причому нейрони наступного шару пов'язані лише з невеликим числом нейронів попереднього шару з околиці локальної області (як у зоровій корі людини). Ваги в кожній точці однієї площини однакові (згорткові шари). За згортковим шаром слід шар, що зменшує його розмірність шляхом локального усереднення. Потім знову згортковий шар, і так далі. Таким чином досягається ієрархічна організація. Пізніші шари витягують загальніші характеристики, що менш залежать від спотворень зображення. Навчається ЗНМ стандартним методом зворотного поширення помилки [17]. Порівняння БНМ і ЗНМ показало суттєві переваги останньої як

за швидкістю, так і за надійністю класифікації. Корисною властивістю ЗНМ також є те, що характеристики, що формуються на виходах верхніх шарів ієрархії, можуть бути застосовні для класифікації за методом найближчого сусіда (наприклад, обчислюючи евклідову відстань), причому ЗНМ може успішно отримувати такі характеристики і для образів, які відсутні в навчальному наборі. Для ЗНМ характерні висока швидкість навчання та роботи. Ця архітектура є перспективною для подальших розробок у галузі розпізнавання зображень просторових об'єктів. БНМ застосовуються і виявлення об'єктів певного типу. Крім цього, будь-яка навчена БНМ певною мірою може визначати належність образів до «своїх» класів, її можна спеціально навчити надійному детектуванню певних класів. У цьому випадку вихідними класами будуть класи, що належать і не належать до типу образів [18].

Приховані Марківські моделі є потужним засобом моделювання різних процесів та розпізнавання образів. За своєю природою Марківські моделі дозволяють враховувати безпосередньо просторово-часові характеристики сигналів, і тому набули широкого застосування у розпізнаванні мови, а останнім часом – зображень (зокрема, зображень обличчя). Кожна модель є набір N станів між цими станами можливі переходи. У кожний момент часу система перебуває у строго певному стані.

У найпоширеніших Марківських моделях першого порядку вважається, що такий стан залежить лише від поточного стану. При переході в кожен стан генерується символ, що спостерігається, відповідний фізичному сигналу з виходу моделюється системи. Вихід, що генерується моделлю, може бути так само безперервним. Існують також моделі, з однаковим набором символів для всіх станів [19].

Щоб скоротити обчислення, у розпізнаванні мови використовуються лінійні моделі. У таких моделях кожне стан має лише одне наступне, як і перехід можливий назад у той самий стан. Такі моделі враховують тимчасові характеристики мовного сигналу: певний порядок проходження ділянок

сигналу, їх взаємне розташування, можливість локальних розтягувань чи стискань. Це дозволяє їх застосовувати і розпізнавання зображень.

Суть двомірних Марківських моделей полягає в тому, що, на відміну від одновимірних лінійних ПММ, вони дозволяють моделювати спотворення зображення та взаємне розташування ділянок не окремо по горизонталі або вертикалі, а в обох напрямках одночасно. Для зменшення обчислювальної складності застосовуються псевдодвовимірні ПММ. Така модель складається з кількох лінійних вертикальних моделей нижнього рівня та однієї лінійної горизонтальної моделі верхнього рівня, на її вхід надходять виходи моделей нижнього рівня. Кожен стан моделі верхнього рівня включає послідовність станів відповідної моделі нижнього рівня [20]. Моделі нижнього рівня не пов'язані між собою. Спочатку моделі верхнього рівня були вертикальними. У наступних роботах моделі верхнього рівня були зроблені горизонтальними, щоб вертикальні моделі нижнього рівня могли врахувати той факт, що очі можуть знаходитися на різній висоті. Таким чином, псевдодвовимірна модель дозволяє врахувати локальні деформації, а також взаємне розташування ділянок зображень. Але на відміну від оптичних потоків та інших методів зіставлення деформаціями, псевдодвовимірна модель враховує характер деформацій, а те, якими можуть бути можливі деформації, псевдодвовимірні ПММ засвоюють у процесі навчання [21]. Іншими словами, ділянка, що відповідає оку, ніколи не буде зіставлена, наприклад, ділянці на місці рота.

Для ПММ важливе значення має початкова ініціалізація моделі. Як початкова ініціалізація всіх моделей використовуються всі зображення з тренувального набору. Потім модель кожного класу налаштовується на зображення.

Недоліком ПММ є те, що ПММ не має здатності, що розрізняє, тобто алгоритм навчання тільки максимізує відгук кожної моделі на свої класи, але не мінімізує відгук на інші класи і не виділяються ключові ознаки, що відрізняють один клас від іншого. Таким чином, схожі класи можуть

виявитися слабкими і при збільшенні обсягу бази або використання в ширших умовах ПММ може виявитися ненадійними [22].

Суть методу гнучкого порівняння на графах (Elastic graph matching) зводиться до еластичного зіставлення графів, що описують зображення облич. У цьому методі (Elastic Bunch Graph Matching) обличчя представляється у вигляді графа, вершини цього графа розташовані на ключових точках обличчя таких, як контури голови, губ, носи та їх крайніх точках. Кожна грань позначена відстанями між її вершинами. У кожній такій точці обчислюються коефіцієнти розкладання функцій Габора для п'яти різних частот і восьми орієнтацій. Набір таких коефіцієнтів $J = \{J_j\}$ називається джетом (jet). Джети характеризують локальні області зображень і служать для двох цілей: по-перше, для знаходження точок відповідності в заданій області на двох різних зображеннях; по-друге – для порівняння двох відповідних областей різних зображень. Кожен коефіцієнт $J_j = a_j \exp(i\phi_j)$ для точок з однієї області різних зображень характеризується амплітудою a_j , що повільно змінюється зі зміною положення точки, і фазою ϕ_j , яка обертається зі швидкістю, пропорційною частоті базового хвильового вектор базисної функції. Тому в найпростішому випадку для пошуку на новому зображенні точки з аналогічними характеристиками функції подібності фазу не враховують, цей випадок описується формулою (1.1)

$$S_a = \frac{\sum_j a_j a'_j}{\sqrt{\sum_j a_j^2 \sum_j a_j'^2}}. \quad (1.1)$$

Функція подібності з одним джетом у фіксованій позиції та іншим зі змінною позицією є досить гладкою, щоб отримати швидку і надійну збіжність при пошуку із застосуванням найпростіших методів, таких як дифузія або градієнтний спуск. Більш досконалі функції подібності залучають інформацію про фазу [23]. Для різних ракурсів відповідні ключові точки відзначаються вручну на тренувальному наборі. Крім того, щоб для одного й того ж обличчя

представити різні варіації його зображення в тому самому графі, для кожної точки використовуються кілька джетів, що відповідають різним локальним характеристикам даної точки, наприклад, відкритому і закритому оку. Процес розпізнавання невідомої людини полягає в порівнянні графа зображення обличчя G' з усіма іншими графами набору B за допомогою функції подібності, що описується за допомогою формули (1.2):

$$S_b(G', B) = \frac{1}{N} \sum_n \max_m S_\varphi(J_n^I, J_n^{Bm}) - \frac{\lambda}{E} \sum_e \frac{(\Delta x_e^I - \Delta x_e^B)}{(\Delta x_e^B)^2}, \quad (1.2)$$

де N – кількість вершин,

E – кількість граней,

λ – коефіцієнт відносної важливості топографічної інформації.

Ліва сума характеризує подібність джетів, обчислену із застосуванням фазочутливої функції, права – топографічна відповідність, пропорційна квадрату різниці відстаней між відповідними вершинами зображень, що порівнюються.

У наведеному вище вигляді метод здатний досить надійно розпізнавати при змінах ракурсу до 20° ; при великих кутах точність розпізнавання різко зменшується, функція подібності виявляється більш чутливою до ракурсу, ніж міжкласових відмінностей. Подальший розвиток методу полягає у добуванні коефіцієнтів важливості на основі аналізу навчальної вибірки. Для кожного джета симплекс-методом обчислюється коефіцієнт важливості, потім використовується згодом у функції подібності. Коефіцієнти важливості обчислюються з умови максимізації функції подоби для одного й того ж обличчя та мінімізації – для різних облич. Існують також більш ранні різновиди цього методу, які не використовують спочатку певні ключові точки та структури графа. Вони використовуються для порівняння решітки джетів, накладені на зображення [24]. У невідомому зображенні відшуковуються точки відповідності, потім за знайденими точками будується спотворені грати і

вимірюється міра її спотворення визначення найбільш схожого зображення. В інших методах точки вилучення джетів спочатку утворюють ґрати, а потім найменш придатні для розпізнавання точки відсіваються у процесі навчання.

Недоліки: висока обчислювальна складність процедури розпізнавання, низька технологічність запам'ятовування нових еталонів.

У холістичних методах зображення розглядається як матриця, значення елементів цієї матриці дорівнюють інтенсивності відповідних пікселів. Таку матрицю можна подати у вигляді одномірного вектора. Отриманий вектор характеризується великою розмірністю, а аналізу необхідно отримати дані меншої розмірності. Таким чином, основною метою цього розділу методів є знаходження оптимального зменшення розмірності безлічі ознак, що дозволяє представити основні особливості кожного об'єкта [25].

Лінійний дискримінантний аналіз (лінійний дискримінант Фішера) використовує таку проекцію простору зображень на простір ознак, що мінімізує внутрішньокласову та максимізує міжкласову відстань у просторі ознак. У цих методах передбачається, що класи лінійно розділяються.

Матриця W для проектування простору зображення на простір ознак вибирається з наступної умови (1.3):

$$W_{opt} = arg \max_W \frac{W^T S_B W}{W^T S_W W}, \quad (1.3)$$

де S_B – матриця міжкласової дисперсії,

S_W – матриця внутрішньокласової дисперсії.

Може існувати до $c-1$ векторів складових базис простору ознак, де c – загальна кількість класів. За допомогою цих векторів простір зображень перетворюється на простір ознак.

Оскільки робота безпосередньо з матрицею $S_W \in R_{n \times n}$ скрутна через її розмірність, використовується попереднє зменшення розмірності за допомогою методу головних компонент, потім обчислення проводяться в просторі меншої розмірності (1.4):

$$W_{fld} = \arg \max_W \frac{W^T W_{pca}^T S_B W_{pca} W}{W^T W_{pca}^T S_W W_{pca} W}, \quad (1.4)$$

де W_{pca} – матриця для проектування в простір меншої розмірності (простір основних компонентів).

Зазвичай тренувальний набір містить зображення облич за кількох базових умов освітленості, на основі цих зображень за допомогою лінійних комбінацій можна отримати будь-які інші умови освітленості [26]. Цей метод дає високу точність розпізнавання (близько 96%) для широкого діапазону умов освітленості, різних виразів обличчя та наявності чи відсутності окулярів. Однак залишаються нез'ясованими питання, чи застосуємо це метод для пошуку у великих базах даних, чи може метод працювати, коли в тренувальній вибірці для деяких облич є зображення лише за умов освітленості. Для завдання детектування людини за допомогою LDA – класи облич та «не облич» розбивають на підкласи. Вищеописаний метод ґрунтується на припущенні про лінійну роздільність класів у просторі зображень. У випадку таке припущення несправедливо. Інструмент для побудови складних поверхонь, що розділяють, пропонують нейромережеві методи.

Завдання аналізу основних компонентів, має, як мінімум, чотири базові версії:

- апроксимувати дані лінійними різноманіттями меншої розмірності;
- знайти підпростори меншої розмірності, в ортогональній проекції на які розкид даних (тобто середньоквадратичне ухилення від середнього значення) максимальний;
- знайти підпростори меншої розмірності, в ортогональній проекції на які середньоквадратична відстань між точками максимально;
- для даної багатовимірної випадкової величини побудувати таке ортогональне перетворення координат, що в результаті кореляції між окремими координатами обернуться в нуль.

Перші три версії оперують кінцевими множинами даних. Вони еквівалентні і не використовують жодної гіпотези щодо статистичного породження даних. Четверта версія оперує випадковими величинами. Кінцеві множини з'являються тут як вибірки з даного розподілу, а вирішення трьох перших завдань як наближення до «справжнього» перетворення Кархунена-Лоева. При цьому виникає додаткове і не цілком очевидне питання про точність цього наближення.

Суть кластерних підходів для порівняння облич полягає у поділі зображення обличчя на деяку кількість ділянок, до кожної з цих ділянок можна застосувати холістичний підхід та порівнювати їх попарно, тобто незалежно від інших ділянок. У такому випадку можна сподіватися на те, що якщо деякі пари ділянок порівнюваних зображень не збігатимуться між собою в силу вищезгаданих заважаючих факторів на кшталт тіней і перекриттів, пари ділянок, що залишилися, забезпечать потрібний ступінь взаємної кореляції, якої буде достатньо для правильного підсумкового рішення збігу/незбігу облич.

Існує спосіб опису 3D моделі за допомогою локальних бінарних шаблонів. Цей метод визначає локальний текстурний патерн за допомогою бінарного коду. Він побудований на граничній класифікації деякої околиці точки деякого радіусу (зазвичай розглядають 8 сусідніх пікселів). Тобто якщо значення яскравості (у чорно-білому зображенні) сусіднього пікселя більше, ніж у пікселі, що розглядається, то він кодується одиницею, інакше нулем. Так кожному пікселю текстури відповідає набір із 8 біт (кількість сусідів). Потім кожній області стоїть гістограма подібних наборів і в такий спосіб описується обличчя. У разі тривимірного зразка цей алгоритм трохи змінюється, щоб описувати внутрішню геометричну інформацію, знижуючи ефект від зміни виразу обличчя ефективно. На стадії навчання спектральна регресія застосовується на навчання основних компонент кожного тривимірного – зображення. Спочатку цей метод застосовувався для порівняння тривимірних

зразків, проте існує застосування локальних бінарних шаблонів і для порівняння об'єктів, що мають різну кількість вимірів розпізнавання.

1.2 Способи побудови вибірки

Для оцінки ефективності розпізнавання та ідентифікації облич на зображенні необхідно загальний набір зображень розділити на два підмножини. Одна з цих підмножин стане навчальним набором, тобто тими зображеннями, приналежність облич, присутніх на якому однозначно визначена і відповідає тій чи іншій людині. Друге підмножина (зображення, що залишилися, із загального набору) буде використано в якості тестової вибірки. Приналежність облич, присутніх ними визначатиметься досліджуваним алгоритмом. Співвідношення правильних і хибних ідентифікацій у результаті стане головним критерієм оцінки ефективності способу. Чим більше виявиться значення цього параметра, тим ефективніше даний метод справляється з поставленим завданням ідентифікації людини зображення обличчя.

1.2.1 Перехресна перевірка

Крос-валідація, яка іноді називається перехресною перевіркою, – це техніка валідації моделі для перевірки того, наскільки успішно застосований у моделі статистичний аналіз здатний працювати на незалежному наборі даних.

Зазвичай крос-валідація використовується в ситуаціях, де метою є передбачення, потрібно оцінити, наскільки модель здатна працювати на практиці.

Перехресна перевірка – це спосіб оцінити здатність моделі працювати на гіпотетичному тестовому наборі, коли такий набір явно отримати неможливо.

І тут вихідний набір даних розбивається на K однакових за розміром блоку. З блоків K один залишається для тестування моделі, а залишки $K-1$ блоку використовуються як навчальний набір. Процес повторюється K разів, кожен із блоків використовується один раз як тестовий набір. По завершенні, виходить K результатів, по одному на кожен блок, вони усереднюються або комбінуються якимось іншим способом, і дають одну оцінку.

Основна перевага такого способу перед випадковим вибором полягає в тому, що всі спостереження використовуються і для тренування, і для тестування моделі і кожне спостереження використовується для тестування в точності один раз.

На рисунку 1.2 показана схема формування навчального та тестового наборів для перехресної перевірки.

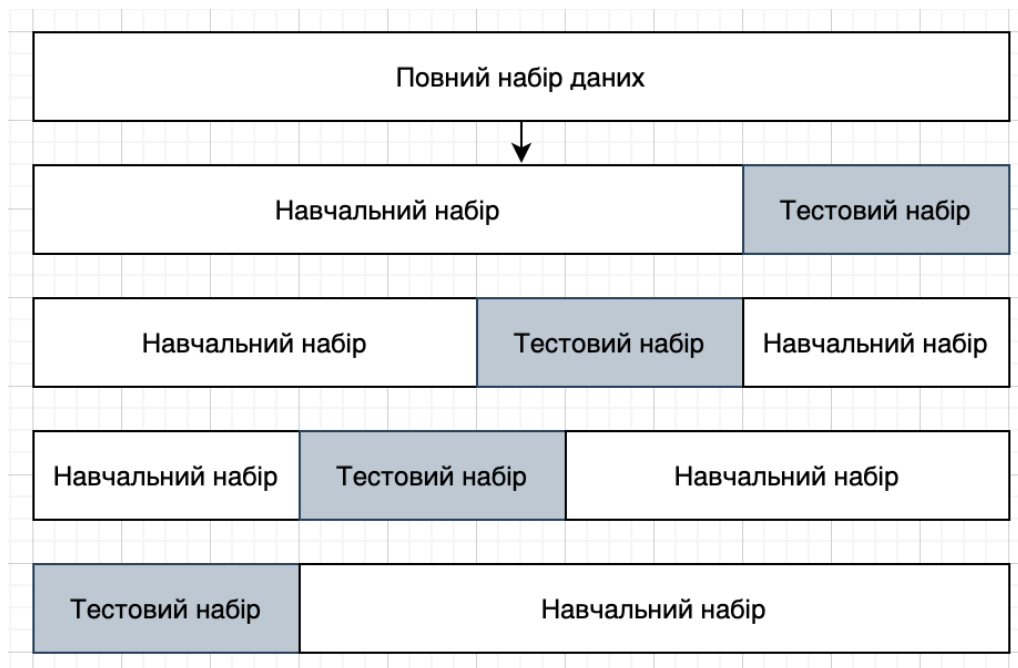


Рисунок 1.2 – Схема формування навчального та тестового наборів для перехресної перевірки

Валідація послідовним випадковим вибором випадково розбиває набір даних на навчальний і тестовий набори. Для кожного такого розбиття модель підганяється під тренувальні дані, а точність передбачення оцінюється на тестовому наборі. Результати потім усереднюються по всіх розбиття. Перевага такого методу перед крос-валідацією на K блоках у тому, що пропорції тренувального та тестового наборів не залежать від кількості повторень (блоків). Недолік методу в тому, що деякі спостереження можуть жодного разу не потрапити до тестового набору, тоді як інші можуть потрапити до нього більш ніж один раз. Інакше кажучи, тестові набори можуть перекриватися [30]. Крім того, оскільки розбиття проводяться випадково, результати відрізнятимуться у разі повторного аналізу.

У пошаровому варіанті цього методу, випадкові вибірки генеруються таким способом, що середня відповідь моделі дорівнює по тренувальному та тестовому набору. Це особливо корисно, коли відповідь моделі бінарна, з нерівними пропорціями відповідей за даними.

У поелементній крос-валідації окреме спостереження використовується як тестовий набір даних, а решта спостережень з вихідного набору – як тренувальний. Цикл повторюється, доки кожне спостереження не буде використано один раз як тестовий. Це те саме, що і K -блочна крос-валідація, де K дорівнює числу спостережень у вихідному наборі даних.

Поділ загального набору на навчальну та тестову вибірку може ґрунтуватися на різних підходах до цього питання.

Для формування вибірки можуть застосовуватися імовірнісні та детерміновані методи.

1.2.2 Імовірнісні способи

Імовірнісні методи включають до свого складу: простий випадковий відбір, систематичний відбір, кластерний відбір і стратифікований відбір.

Простий випадковий (власне-випадковий) відбір припускає, що можливість бути обраним у вибірку відома і є однаковою для всіх одиниць сукупності. Імовірність бути включеним у вибірку визначається ставленням обсягу вибірки до розміру сукупності.

Простий випадковий відбір може здійснюватися за допомогою таких методів: формування вибірки наосліп або за допомогою таблиці випадкових чисел.

При використанні методу формування вибірки наосліп одиниці сукупності з даного загального набору випадковим чином вибирається деяка кількість зображень, що визначається обсягом вибірки. Для визначення ідентифікатора випадкового зображення можна використовувати будь-який програмний генератор випадкових (псевдовипадкових) чисел.

Початкова частина методу систематичного (механічного) відбору відповідає початковій частині методу простого випадкового відбору: потрібно отримати повний перелік одиниць генеральної сукупності.

Однак далі замість присвоєння ідентифікаторів використовується показник інтервал стрибка, розрахований як відношення розміру сукупності до обсягу вибірки.

Особливо широко метод систематичного відбору використовується, коли різних видів сукупностей є різні довідники, списки тощо.

Іншим методом імовірнісного відбору є кластерний відбір, заснований на розподілі сукупності на підгрупи. Базова концепція цього методу дуже схожа на базову концепцію методу систематичного відбору, проте реалізація цієї концепції здійснюється по-іншому. Далі один із кластерів вибирається випадковим чином, визначається сукупність для цієї частини, в ній проводиться відповідне дослідження, а висновки узагальнюються на сукупність всього кластера.

В основі всіх описаних методів є припущення, що будь-яка сукупність характеризується симетричним розподілом її ключових характеристик. Іншими словами, кожна вибірка досить повно характеризує всю сукупність,

різні крайнощі у вибірці врівноважують один одного. Така ситуація практично зустрічається вкрай рідко.

У разі несиметричного розподілу сукупності остання поділяється на різні підгрупи (страти), за деякими незалежними параметрами, і вибірки формуються з цих підгруп. Такий метод зветься стратифікованого відбору.

Далі для кожної страти за допомогою випадкового відбору формується вибірка.

1.2.3 Детерміновані способи

При застосуванні детермінованих методів відбору формування вибірки здійснюється без використання понять теорії ймовірностей, унаслідок чого неможливо розрахувати ймовірність включення до вибірки одиниці сукупності.

Охарактеризуємо такі детерміновані методи відбору: відбір на основі принципу зручності, відбір на основі суджень, формування вибірки в процесі обстеження та формування вибірки на основі квот.

Сенс методу відбору з урахуванням принципу зручності у тому, що формування вибірки здійснюється найзручнішим з позицій дослідника чином, наприклад, з позицій мінімальних витрат часу і зусиль, з позиції доступності зразків.

Формування вибірки на основі судження ґрунтується на використанні думки кваліфікованих експертів щодо складу вибірки.

Формування вибірки з урахуванням квот (квотний відбір) передбачає попереднє, з цілей дослідження, визначення чисельності груп зразків, відповідних певним критеріям.

Рішення про обсяг вибірки є компромісом між теоретичними припущеннями щодо точності результатів обстеження та можливостями їх

практичної реалізації, передусім маються на увазі витрати на проведення дослідження.

1.3 Постановка задачі дослідження

Об'єктом дослідження є системи розпізнавання облич.

Метою дослідження проведеного в рамках даної кваліфікаційної роботи є підвищення ефективності та якості роботи систем розпізнавання облич за рахунок вибору найбільш ефективного методу розпізнавання та підбору необхідного в конкретних умовах поєднання параметрів цього методу.

У роботі розглядаються деякі аспекти використання алгоритмів ідентифікації людини за зображенням обличчя, які можуть впливати на ефективність такого розпізнавання.

Для досягнення більшої ефективності було вирішено такі питання:

- розроблено алгоритм формування тестової та навчальної вибірки з неоднорідного вихідного набору даних;
- з використанням цього алгоритму проведено дослідження та порівняння ефективності методів ідентифікації людини за зображенням обличчя;
- виявлено особливості вихідного набору даних, при яких певний метод ідентифікації має найбільші переваги;
- визначено, які вхідні параметри методів більшою мірою впливають на ефективність та точність розпізнавання методів;
- на підставі результатів дослідження було сформульовано рекомендації щодо використання методів розпізнавання облич.

2 ТЕОРЕТИЧНИЙ БАЗИС. ДОСЛІДЖУВАНІ МЕТОДИ РОЗПІЗНАВАННЯ ОБЛИЧ

2.1 Метод головних компонент

Основною ідеєю застосування методу головних компонент (Principal Component Analysis, PCA) для розпізнавання облич є введення поняття «власних облич» (eigenfaces), що представляють зображення облич у вигляді набору головних компонентів зображень. Процес розпізнавання облич полягає у порівнянні головних компонент невідомого зображення з головними компонентами всіх відомих зображень. Основні переваги застосування PCA виявляються при зберіганні та пошуку зображень у великих базах зображень, а також реконструкції зображень. Недолік PCA: високі вимоги щодо умов зйомки зображень [31].

Метод головних компонент застосовується для стиснення інформації без суттєвих втрат інформативності. Він полягає у лінійному ортогональному перетворенні вхідного вектору x розмірності N у вихідний вектор y розмірності M , $N > M$. При цьому компоненти вектора y є некорельованими і, отже, загальна дисперсія після перетворення залишається незмінною. Матриця X складається з усіх прикладів зображень навчального набору. Розв'язавши рівняння

$$\Lambda = \Phi^T \Sigma \Phi, \quad (2.1)$$

отримуємо матрицю власних векторів Φ ,

де Σ – коваріаційна матриця для x ,

Λ – діагональна матриця власних чисел.

Вибравши з Φ підматрицю Φ_M , відповідну M найбільшим власним числам, отримаємо, що перетворення

$$y = \Psi_M^T X', \quad (2.2)$$

де $X' = X - \bar{X}$ нормалізований вектор з нульовим математичним очікуванням, характеризує більшу частину загальної дисперсії та відбиває найбільш істотні зміни x .

Вибір перших M головних компонент розбиває векторний простір на головний (власний) простір $F = \{\phi_i\}_{i=1}^M$, що містить головні компоненти, та його ортогональне доповнення $F = \{\phi_i\}_{i=M+1}^M$.

Як індикатори приналежності в методі основних компонент використовують:

DIFS – distance in feature space, відстань від образу аналізованого зображення у своєму просторі, до еталонного образу;

DFFS – distance from feature space, відстань від представлення аналізованого зображення у просторі спостереження до проекції еталона у власному просторі.

Застосування задачі розпізнавання людини за зображенням обличчя має такий вид. Вхідні вектори представляють собою відцентровані та приведені до єдиного масштабу зображення облич. Власні вектори, обчислені для набору зображень облич, називаються власними обличчями (eigenfaces). Метод основних компонент у застосуванні до зображень облич називають методом власних облич. За допомогою обчислених раніше матриць вхідне зображення розкладається на набір лінійних коефіцієнтів, які називаються головними компонентами. Сума N перших головних компонентів, помножених на відповідні власні вектори, є апроксимацією зображення порядку N .

До кожного зображення обличчя обчислюються його основні компоненти. Зазвичай береться від 5 до 200 основних компонентів. Інші компоненти кодують дрібні відмінності між обличчями та шум. Процес розпізнавання полягає у порівнянні головних компонент невідомого зображення з компонентами решти зображень. Для цього зазвичай застосовують якусь метрику (найпростіший випадок – Евклідова відстань)

[32]. Додаткове підвищення надійності досягається за рахунок додаткового застосування аналізу основних компонентів до окремих ділянок обличчя таких, як очі, ніс, рот.

Також метод головних компонент застосовується для виявлення обличчя на зображенні. Для облич значення компонент у своєму просторі мають великі значення, а в доповненні до власного простору – близькі до нуля. За цим фактом можна виявити, чи є вхідне зображення обличчям. І тому перевіряється величина помилки реконструкції; чим більша помилка, тим більше ймовірність, що це не обличчя. За наявності у наборі зображень обличчя варіацій таких, як раса, стать, емоції, висвітлення, з'являтимуться компоненти, їх величина в основному визначається цими факторами. Тому за значеннями відповідних головних компонентів можна визначити, наприклад, расу або стать людини. Основні недоліки PCA такі. Метод головних компонентів вимагає для свого застосування ідеалізованих умов таких, як єдині параметри освітленості, нейтральний вираз обличчя, відсутність перешкод на кшталт окулярів і борід. При недотриманні цих умов головні компоненти не будуть відображати міжкласові варіації [33]. Наприклад, за різних умов освітленості метод власних обличчя практично не застосовується, оскільки перші головні компоненти переважно відображають зміни освітлення, і порівняння видає зображення, що мають схожий рівень освітленості. На рисунку 2.1 показано реконструкції перших 20 власних облич.



Рисунок 2.1 – Приклад зображень власних векторів (власні обличчя)

Обчислення набору власних векторів відрізняється високою трудомісткістю. Один із способів – це згортка зображень рядками та стовпцями; у такій формі подання зображення має на порядок менший розмір, обчислення та розпізнавання відбувається швидше, але відновити вихідне зображення вже неможливо.

Проблема представлення зображення полягає у його високій розмірності. Двовимірне $p \times q$ зображення у відтинках сірого відображається у векторний простір $m = pq$, таким чином, зображення 100×100 пікселів уже лежить у 10000-мірному просторі. Питання полягає в тому, чи всі виміри для нас однаково корисні? Ми можемо лише робити висновки про те, чи є якийсь розкид у даних, таким чином, шукаються ті компоненти, що містять більшу частину інформації. Аналіз головних компонентів був запропонований незалежно Карлом Пірсоном у 1901 році та Гарольдом Хотелінгом, для того щоб перетворити набір ймовірно корельованих змінних у менший набір некорельованих змінних. Ідея полягає у тому, що багаторозмірний набір даних

найчастіше описується змінними, що корелюються між собою. І тому менша кількість найбільш значних змінних містить основну частину інформації [34]. Метод аналізу основних компонентів дозволяє вибрати основне у великому різноманітті даних.

Алгоритм роботи методу головних компонент:

Нехай $X = \{x_1, x_2, \dots, x_n\}$ випадковий вектор, у якому:

1. Обчислити середнє μ за формулою:

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i. \quad (2.3)$$

2. Обчислити коваріаційну матрицю S , що обчислюється за формулою:

$$S = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)(x_i - \mu)^{\tau}. \quad (2.4)$$

3. Обчислити власні значення λ_i та власні вектори v_i з S , обчисленого за формулою:

$$S v_i = \lambda_i v_i, \quad (2.5)$$

де $i = 1, 2, \dots, n$.

Впорядкувати власні вектори за спаданням їх власних значень. У основних компонентів це власні вектори, відповідні найбільшим k – власним значенням. K -головних компонент аналізованого вектора задані рівністю:

$$y = W^{\tau}(x - \mu), \quad (2.6)$$

де μ – нормальний розподіл.

Відновлений за допомогою аналізу основних компонентів базис представлений рівністю:

$$x = W_y + \mu, \quad (2.7)$$

де W – внутрішньокласова дисперсія.

Метод головних компонентів виконує розпізнавання облич за допомогою:

- проєціювання всіх навчальних прикладів у підпростір основних компонентів;
- проєціювання оброблюваного зображення в підпростір основних компонентів;
- знаходження найближчого сусіда для зображення, що обробляється серед навчальних прикладів.

Залишається ще одна невирішена проблема. Припустимо, дано 400 зображень розмірністю 100 на 100 пікселів. Методом головних компонентів отримуємо коваріаційну матрицю $S = XX^t$, де розмір виходить рівним 10000×400 . Рішенням буде матриця 10000×10000 , приблизно 0,8 гігабайт, це практично нездійсненно. Тому використовується наступний варіант. Матриця $M \times N$, при $M > N$, може мати лише $N - 1$ ненульових власних значень, тому можна взяти декомпозицію $S = X^t X$, розмірністю $N \times N$ описану формулою:

$$X^t X_{v_i} = \lambda_i v_i, \quad (2.8)$$

і отримати вихідні власні вектор $S = XX^t$, помножені зліва на матрицю даних: $XX^t(X_{v_i}) = \lambda_i(X_{v_i})$ отримані власні вектори ортогональні, щоб отримати ортонормальні вектори вони повинні бути нормалізовані [35].

2.2 Метод локальних бінарних шаблонів

Локальний бінарний шаблон – це певний вид ознаки, що використовується для класифікації у комп'ютерному зорі, і є простим оператором. Локальні бінарні шаблони вперше були запропоновані в 1996 для аналізу текстури напівтонових зображень. При цьому подальші дослідження показали, що ЛБШ інваріантні до невеликих змін в умовах освітлення та невеликих поворотів зображення [36].

ЛБШ представляє собою опис околиці пікселя зображення в двійковому уявленні. Базовий оператор ЛБШ, що застосовується до пікселя зображення, використовує вісім пікселів околиці, приймаючи значення інтенсивності центрального пікселя як поріг (рис. 2.2). Пікселі зі значенням інтенсивності більшим або рівним значенню інтенсивності центрального пікселя приймають значення рівні «1», інші набувають значення рівні «0». Таким чином, результатом застосування базового оператора ЛБШ до пікселя зображення є восьмирозрядний бінарний код, що описує околицю пікселя.

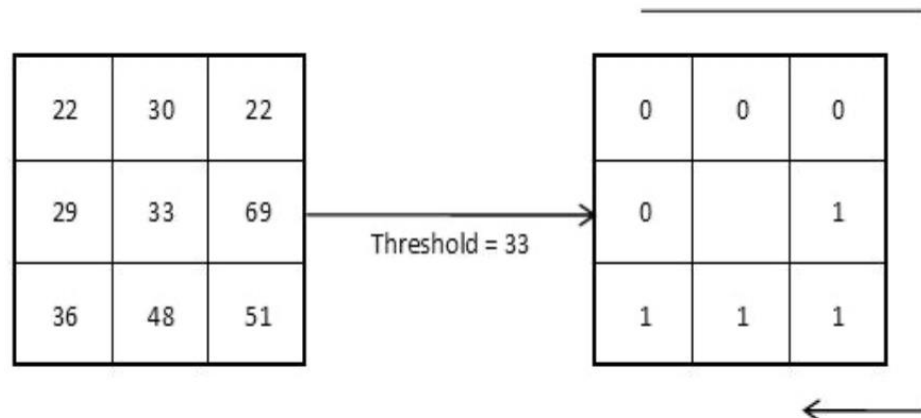


Рисунок 2.2 – Схема базового оператора ЛБШ

Використання кругової околиці та білінійної інтерполяції значень інтенсивностей пікселів дозволяє побудувати локальний бінарний шаблон із довільною кількістю точок P та радіусом R .

Деякі бінарні коди несуть у собі більше інформації, ніж інші. Так, локальний бінарний шаблон називається рівномірним, якщо він містить не більше трьох серій «0» та «1» (наприклад, 00000000, 001110000 та 11100001). По-перше, рівномірні ЛБШ визначають лише важливі локальні особливості зображення, такі як кінці ліній, грані, кути та плями (рис. 2.3). По-друге, вони забезпечують суттєву економію пам'яті ($P(P - 1) + 2$ різних шаблонів замість $2P$).

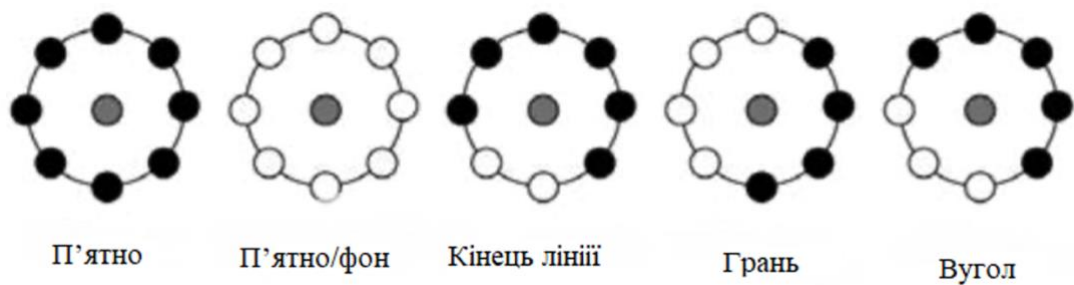


Рисунок 2.3 – Оператори, що відповідають деяким графічним примітивам

Застосовуючи оператор ЛБШ кожного пікселя зображення, ми можемо побудувати гістограму, де кожному рівномірному коду ЛБШ відповідає окремий стовпець. Також є ще один додатковий стовпець, що містить інформацію про всі нерівномірні шаблони [37].

Алгоритм роботи методу локальних бінарних шаблонів

Формальний опис ЛБШ оператора виглядає, як показано у формулі:

$$LBP(x_c, y_c) = \sum_{p=0}^{p-1} 2^p S(i_p - i_c), \quad (2.9)$$

де $(i_p - i_c)$ – інтенсивність центрального пікселя;

i_n інтенсивність сусіднього пікселя.

S – знакова функція, що визначена рівністю:

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{else} \end{cases}. \quad (2.10)$$

Цей опис дозволив позначити 5 різних видів деталі на зображенні. Незабаром після публікації було помічено, що за фіксованих сусідів розпізнавання деталей у масштабі не вдається. Тому почали використовувати змінну кількість сусідів. Ідея у тому, щоб прив'язувати довільну кількість сусідів по колу зі змінним радіусом, що дозволило вловлювати наступні сусідства.

Для цієї точки положення сусідів може бути за наступними формулами:

$$x_p = x_c + R \cos\left(\frac{2\pi p}{p}\right), \quad (2.11)$$

$$y_p = y_c - R \sin\left(\frac{2\pi p}{p}\right), \quad (2.12)$$

де R – радіус кола;

p – кількість точок зразка.

«Оператор» – це розширення початкового ЛБШ коду, Тому його іноді називають розширений ЛБШ (також називається круговий ЛБШ). Якщо координати точки не відповідають координатам зображення, точка інтерполюється. У комп'ютерній науці існує безліч «розумних» інтерполяційних схем, в OpenCV застосовується білінійна інтерполяція, що визначається формулою:

$$f(x, y) \approx [1 - xx] \begin{bmatrix} f(0,0) & f(0,1) \\ f(1,0) & f(1,1) \end{bmatrix} \begin{bmatrix} 1 - y \\ y \end{bmatrix}. \quad (2.13)$$

Залишається додати просторову інформацію до моделі розпізнавання обличчя. У представленні Ахонена пропонується розділити ЛБШ зображення на m ділянок та витягувати гістограму для кожного. Вектори просторових

характеристик виходять шляхом поєднання локальних гістограм (не зливаючи їх). Ці гістограми називаються гістограмами локальних бінарних шаблонів.

Зображення обличч можуть розглядатися як набір різних локальних особливостей, вони добре описуються за допомогою локальних бінарних шаблонів. Однак гістограма, побудована для всього зображення в цілому, кодує лише наявність тих чи інших локальних особливостей, але не містить жодної інформації про їх розташування на зображенні [38]. Для обліку такого роду інформації зображення розбивається на підобласті, у кожному з яких обчислюється своя гістограма ЛБШ (рис. 2.4). Шляхом конкатенації цих гістограм можна отримати загальну гістограму, яка враховує як локальні, так і глобальні особливості зображення.



Рисунок 2.4 – Розділ зображення на підобласті, для обчислення окремої гістограми ЛБШ на різних ділянках

3 ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ НА ОСНОВІ МЕТОДУ ГОЛОВНИХ КОМПОНЕНТ ТА МЕТОДУ ЛОКАЛЬНИХ БІНАРНИХ ШАБЛОНІВ

3.1 Вихідні дані – бази зображень облич

Вихідний набір зображень був сформований з двох баз зображень облич ORL і FERET, що вільно розповсюджуються.

3.1.1 База ORL

База зображень облич, підготовлена в науково-дослідній лабораторії компанії Olivetti (Olivetti Research Laboratory або ORL; після придбання корпорацією AT&T у 1999 р. і до закриття у 2002 р. лабораторія мала назву AT&T Laboratories Cambridge). В основу входять фронтальні зображення облич. База містить 400 фронтальних зображень 40 чоловік. Усі зображення у базі напівтонові, з 256 градаціями яскравості. Розмір кожного зображення – 92×112 пікселів. Файли представлені у форматі PGM. На рисунку 3.1 показані приклади вихідних із зображень бази ORL.



Рисунок 3.1 – Приклади зображень бази ORL

3.1.2 База FERET

База зображень обличчя FERET – це стандартний набір, який використовується для оцінки систем розпізнавання обличчя. Технологія розпізнавання обличчя – FERET – програма під управлінням агентства передових оборонних дослідницьких проєктів DARPA та Національного інституту стандартів та технологій США. Дана база зображень збиралася з грудня 1993 по серпень 1996 У 2003 DARPA випустила базу у високій роздільній здатності і 24-бітної кольоровості. Розмір зображень обличчя бази даних FERET дорівнює 512×768 пікселів. Приклади вихідних зображень з бази FERET показані на рисунку 3.2.



Рисунок 3.2 – Приклади зображень бази FERET

3.2 Формування вибірки

Оскільки мета даної роботи – дослідження ефективності методів розпізнавання та ідентифікації обличчя, весь набір зображень був оброблений за допомогою каскадів Хаара, виявлені каскадом області зображень, виділені як зображення обличчя та стали основою подальшого дослідження [39]. Таким чином, для обробки пропонувалися лише ті частини зображень, де машина виявила обличчя, та ефективність методів виявлення обличчя не вплинула на результати дослідження.

Рисунки 3.3 і 3.4 показують, яка область вихідних зображень виділяється у якості той, що містить зображення обличчя.



Рисунок 3.3 – Результат виділення області обличчя на зображеннях бази ORL



Рисунок 3.4 – Результат виділення області обличчя на зображеннях бази FERET

Після процедури виділення облич у підсумкових наборах виявилось 515 класів, які відповідають обраним критеріям. 477 їх у базі FERET і 38 у базі ORL. Усього було виділено 4705 зображень облич (4366 належали класам бази FERET, 339 – ORL). У кожному класі з бази ORL вийшло від 5 до 10 (3 – 8 у навчальній вибірці) зображень, а класах з бази FERET від 4 до 36 (2–28 для навчальної вибірки).

3.2.1 Створення навчальних та тестових наборів

Після того, як визначені кількісні характеристики для всіх класів обох вибірок, необхідно вирішити питання про оптимальний вибір того, як розділяти набори зображень кожного окремого класу на тестовий та навчальний для проведення перехресної перевірки у кількох варіантах [40]. Такий підхід дозволяє вирішити два завдання. По-перше, сама по собі перехресна перевірка дозволяє уникнути так званої помилки «третього роду» – пов'язаної з параметрами, що присутні у вихідному наборі даних, що враховуються класифікаторами, такими як освітлення та геометричні спотворення зображень обличчя. По-друге, використання різних пропорцій співвідношення тестового та навчального набору дозволяє розглянути варіанти, де кількість зображень, що пред'являються для навчання класифікатора буде різною для одного і того ж класу, але при цьому збережуться переваги перехресної перевірки і практично кожне зображення потрапить у тестову вибірку, але буде відсутнє у навчальному наборі.

Зрештою, обрані варіанти розбиття вибірки кожного класу на 2, 3, 4 та 5 частин. Вибір мінімальної кількості очевидний, оскільки одна частина необхідна для тестового набору, а друга для навчального. Максимальне аналізоване число – п'ять, зумовлено, з одного боку тим, що більшість частин вимагали незрівнянно більше тимчасових і машинних ресурсів щодо відповідної обробки всіма класифікаторами з усіма обраними варіантами параметрів. З іншого боку, характеристики вихідного набору даних не дозволяють включити до обробки понад тридцять відсотків всіх класів, кількість зображень усередині яких було менше 7, що негативно може позначитися на репрезентативності результатів.

Створення тестового та навчального набору для окремого класу зображень

До кожного класу зображень обличчя застосовані вище описані види розбиття загального набору, тестовий набір, за допомогою програмного

генератора випадкових чисел обрані зображення таким чином, що кожне зображення потрапляє в тестовий набір не більше одного разу для даного варіанту розбиття. У навчальний набір відповідно потрапляють всі зображення класу. Схема алгоритму формування тестового та навчального набору для кожного класу у вихідному наборі показана на рисунку А.1

Для кожного варіанта вибірки на навчання класифікаторів збирається загальна для всіх класів навчальна вибірка шляхом конкатенації всіх навчальних наборів зображень кожного класу, що потрапили в число класів, що мають достатню кількість зображень для даного варіанту розбиття загальної кількості зображень у класі на обрану кількість частин. Аналогічним чином збирається тестова вибірка для поточного варіанта розбиття, тобто тестовий набір потрапляють всі зображення кожного класу не потрапили в навчальний набір. На рисунку А.2 показаний алгоритм формування загального тестового та навчального набору для всіх класів.

3.2.2 Набори, сформовані для навчання та перевірки класифікаторів

Після проведення цих процедур були сформовані набори для обробки за допомогою досліджуваних методів розпізнавання, у таблиці 3.1 показано, як розподілилися за вибірками зображення, залежно від того, скільки частин ділився вихідний набір зображень.

Таблиця 3.1 – Об’єм тестового та навчального набору при розбитті вихідного набору зображень класу на різну кількість частин

Частка зображень класу, що вибираються для складання тестового набору	1/2	1/3	1/4	1/5
Кількість зображень, що складають один навчальний набір				
Із бази FERET	2306	3083	3418	3438
Із бази ORL	174	239	270	280
Всього	2480	3322	3688	3718
Кількість зображень, що складають тестовий набір				
Із бази FERET	2060	1283	900	704
Із бази ORL	165	100	69	59
Всього	2225	1383	969	763

3.3 Програмне забезпечення дослідження

Для розробки використовується платформа .NET Framework – програмна платформа, випущена компанією Microsoft у 2002 році. Основою платформи є загальномовне середовище виконання Common Language Runtime (CLR), що підходить для різних мов програмування. Функціональні можливості CLR доступні у будь-яких мовах програмування, що використовують це середовище.

3.3.1 Програмна платформа

Хоча .NET є патентованою технологією корпорації Microsoft і офіційно розрахована працювати під операційними системами сімейства Microsoft

Windows, існують незалежні проекти (передусім це Mono і Portable .NET), дозволяють запускати програми .NET на деяких інших операційних системах.

3.3.2 Середовище розробки

Середовищем розробки стала програма Microsoft Visual Studio.

Лінійка продуктів компанії Microsoft включає інтегроване середовище розробки програмного забезпечення та низку інших інструментальних засобів. Дані продукти дозволяють розробляти як консольні програми, так і програми з графічним інтерфейсом, у тому числі з підтримкою технології Windows Forms, а також веб-сайти, веб-додатки, веб-служби як в рідному, так і в керуваному коді для всіх платформ, що підтримує Windows, Windows Mobile, Windows CE, .NET Framework, Xbox, Windows Phone .NET Compact Framework та Silverlight.

Visual Studio включає редактор вихідного коду з підтримкою технології IntelliSense і можливістю найпростішого рефакторингу коду. Вбудований відладчик може працювати як відладчик рівня вихідного коду, так і як відладчик машинного рівня. Інші вбудовані інструменти включають редактор форм для спрощення створення графічного інтерфейсу програми, веб-редактор, дизайнер класів і дизайнер схеми бази даних. Visual Studio дозволяє створювати та підключати сторонні доповнення (плагіни) для розширення функціональності практично на кожному рівні, включаючи додавання підтримки систем контролю версій вихідного коду (як, наприклад, Subversion та Visual SourceSafe), додавання нових наборів інструментів (наприклад, для редагування та візуального проектування коду предметно-орієнтованими мовами програмування) або інструментами для інших аспектів процесу розробки програмного забезпечення (наприклад, клієнт Team Explorer для роботи з Team Foundation Server).

3.3.3 Мова програмування

Як мову програмування було обрано C#.

Це об'єктно-орієнтована мова програмування. Розроблена в 1998-2001 роках групою інженерів під керівництвом Андерса Хейлсберга в компанії Microsoft як мова розробки додатків для платформи Microsoft .NET Framework і згодом була стандартизована як ECMA-334 та ISO/IEC 23270.

C# належить до сім'ї мов з C-подібним синтаксисом, їх синтаксис найбільш близький до C++ і Java. Мова має статичну типізацію, підтримує поліморфізм, навантаження операторів (у тому числі операторів явного та неявного приведення типу), делегати, атрибути, події, властивості, узагальнені типи та методи, ітератори, анонімні функції з підтримкою замикань, LINQ, винятки, коментарі у форматі XML.

Переїнявши багато від своїх попередників – мов C++, Pascal, Модула, Smalltalk і, особливо, Java – C#, спираючись на практику їх використання, виключає деякі моделі, що зарекомендували себе як проблематичні при розробці програмних систем, наприклад, C# на відміну від C++ не підтримує множинне успадкування класів (між тим допускається множинне успадкування інтерфейсів).

3.3.4 Розширення для роботи із зображеннями

Для роботи із зображеннями використовувалася OpenCV (Open Source Computer Vision Library, бібліотека комп'ютерного зору з відкритим кодом) — бібліотека алгоритмів комп'ютерного зору, обробки зображень та чисельних алгоритмів загального призначення з відкритим кодом. Реалізована на C/C++, також розробляється для Python, Java, Ruby, Matlab, Lua та інших мов.

Щоб скористатися можливостями OpenCV у середовищі .Net знадобилася Emgu CV – крос-платформна .Net обгортка для Open CV,

бібліотеки обробки зображень. Вона дозволяє викликати функції Open CV з .NET сумісних мов програмування, таких як C#, VB, VC++, IronPython і т.п. , Android та Windows Phone.

3.4 Результати, отримані в ході дослідження

Розглянемо та проаналізуємо результати, показані на матеріалах бази ORD та FERET.

За результатами розбиття наборів зображень бази ORL для кожного класу кількість зображень у класі, для різних варіантів розбиття, приймає значення від 3 до 8. Відповідно, кількість спроб ідентифікації зображення, зроблені кожним класифікатором, варіюється в діапазоні від 16 до 398. Відсоток успішних розпізнавань для всіх комбінацій, що вийшли, показаний у таблиці 3.2.

Таблиця 3.2 – Результати, що показані на матеріалі бази ORL

Кількість зображень у класі	Кількість спроб розпізнавання	Доля вдалих розпізнавань, %					
		Алгоритм розпізнавання					
		Головні компоненти			Локальні бінарні шаблони		
		80 ГК	160 ГК	160 ГК та світлова норм.	Радіус = 1	Радіус = 5	Радіус = 5 та світлова нормалізація
3	16	69	69	75	75	56	56
4	96	80	80	80	92	74	74
5	284	86	86	84	92	85	86
6	156	87	88	88	95	87	86
7	251	90	91	92	95	94	94
8	398	93	93	92	98	96	96
	1201	89	89	89	95	90	90

Число випробувань при кількості зображень у класі рівному 3 помітно нижче інших варіантів, тому цей результат можна орієнтуватися лише визначення певної загальної тенденції, основну увагу слід приділити результатам переважають у всіх інших випадках.

В цілому очевидно, що зі збільшенням кількості зображень у класі ефективність розпізнавання помітно зростає, незалежно від класифікатора та його параметрів.

Крім того, метод локальних бінарних шаблонів практично у всіх випадках показує кращі результати, ніж метод головних компонент. У середньому перевага близько 5% (в окремих випадках більше 10%) ефективності [41].

Світлова нормалізація не має значного впливу ефективність обох алгоритмів (відмінності 1–2%).

На рисунках 3.5 та 3.6 показані результати світлової нормалізації для зображень облич для кожної бази. Очевидно, що така обробка не вносить істотних змін до зображень бази ORL, тоді як зміни у зображеннях FERET помітні неозброєним поглядом. Це пояснює відмінності у впливі даного параметра на ефективність методів розпізнавання.



Рисунок 3.5 – Світлова нормалізація зображень облич бази ORL



Рисунок 3.6 – Світлова нормалізація зображень облич бази FERET

Для метода основних компонент, очевидно, збільшення кількості основних компонент з 80 до 160 також дає принципові зміни в результатах. Це можна пояснити тим, що значуща в даній ситуації інформація цілком достатнім чином описується 80 компонентами, і подальше додавання компонент не має сенсу, оскільки інформація, що потрапляє в ці компоненти, не містить особливостей, що дозволяють відрізнити зображення інших класів і віднести зображення до зображення свого класу [42].

На рисунку 3.7 видно, що криві графіків усіх результатів методу головних компонент схожі і дуже близькі один щодо одного, тому практично зливаються в одну лінію.

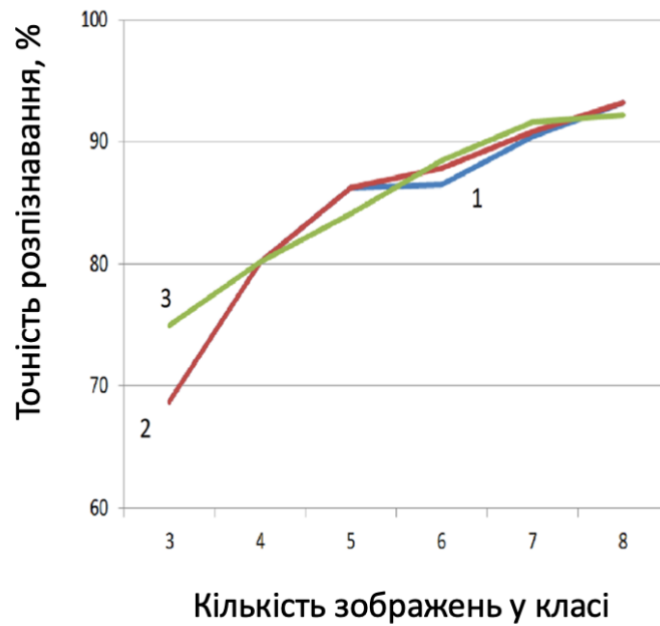


Рисунок 3.7 – Графік результатів метода головних компонент з різними параметрами для бази ORL (1 – 80 головних компонент, 2 – 160 головних компонент без світлової нормалізації, 3 – 160 головних компонент зі світловою нормалізацією)

Для методу локальних бінарних шаблонів зміни параметра радіуса – величини кожної окремої ділянки, що обробляється, зробило негативний вплив на вибірках з малою кількістю навчальних прикладів, проте зі збільшенням цієї кількості вплив зменшується, і криві графіків зближуються (рис. 3.8) Найімовірніше, це не стільки кількістю зображень у класі, скільки збільшенням кількості випробувань, а отже і зниженням значимості кожного окремого випробування, пов'язаним з перебором різних поєднань зображень у тестовому та навчальному наборі, при використанні методу перехресної перевірки.

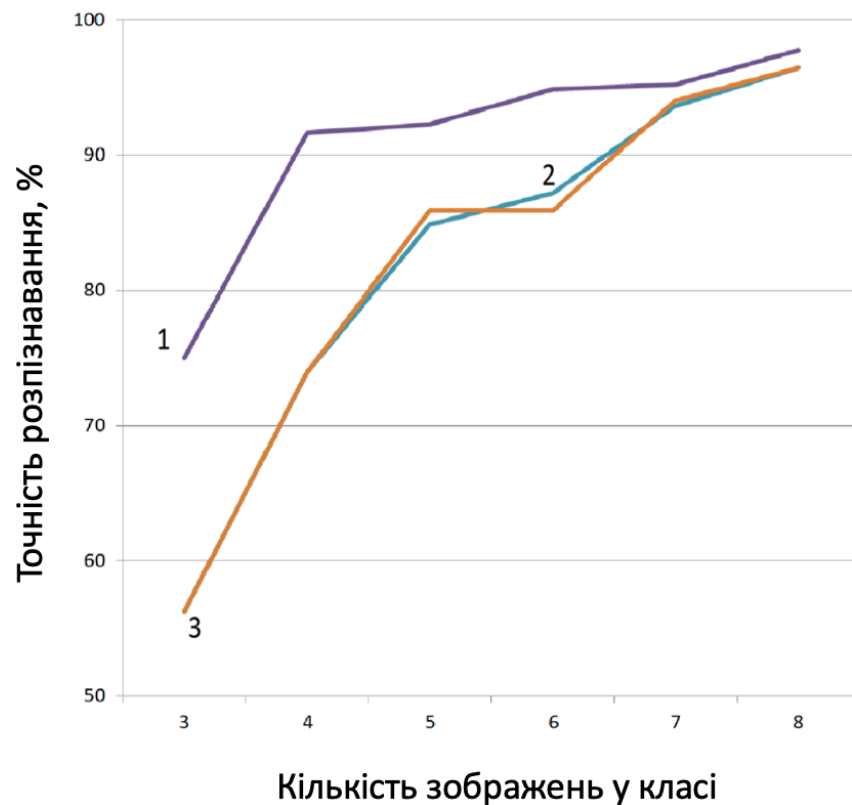


Рисунок 3.8 – Графік результатів методу локальних бінарних шаблонів з різними параметрами для бази ORL (1 – Радіус патерна = 1, 2 – Радіус патерна = 5 без світлової нормалізації, 3 – Радіус патерна = 5 зі світловою нормалізацією)

В цілому, середні значення ефективності для обох алгоритмів на матеріалі бази ORL досить схожі, проте, результати методу локальних бінарних шаблонів у кращому поєднанні параметрів все ж таки трохи вище результатів методу головних компонент.

Графіки найкращих результатів обох методів на матеріалі бази ORL показані рисунку 3.9

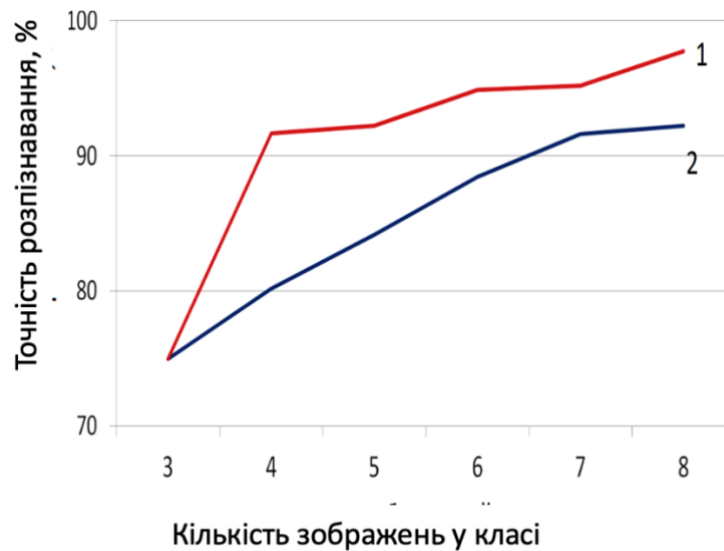


Рисунок 3.9 – Графіки найкращих результатів обох методів для бази ORL (1 – Радіус патерна = 1 без світлової нормалізації, 2 – 160 головних компонент зі світловою нормалізацією)

Зображення облич, представлені в базі FERET за своїми особливостями, істотно відрізняються від зображень бази ORL. По-перше, ракурс зйомки багатьох зображень куди сильніше відхиляється від зйомки анфас, ніж у ORL, як і положення голови людини щодо осей зображення, по-друге, умови освітленості мають набагато більше варіацій розташування щодо джерела світла та його інтенсивності.

Немаловажний також той факт, що в базі FERET варіативність міміки людини на знімку набагато більша, ніж у ORL [43].

Крім того, кількість зображень у класі часто куди більше, ніж для ORL, та й загальна кількість класів на порядок більше. Само собою, такі відмінності у вихідному матеріалі істотно впливають на результативність обробки зображень класифікаторами, ускладнюючи завдання ідентифікації. Кількість класів із бази зображень FERET 477, а кількість зображень у класі від 2 до 28 зображень. А кількість спроб розпізнавання зображень для різної кількості зображень у класі становить мінімум 32 та максимум 2571.

Найбільша кількість випробувань понад 1000 для кожного варіанта проводилася на наборах, де кількість зображень у класі складала від 4 до 9 зображень.

В рамках дослідження ці особливості корисні для того, щоб з одного боку, зафіксувати більш об'єктивні результати роботи класифікаторів зі збільшенням кількості зображень у класі, а з іншого поспостерігати, наскільки зміниться їх ефективність, при роботі в умовах ближчих до реальних практичних завдань подібних систем.

Зрозуміло, всі ці чинники вплинули на підсумкові результати розпізнавання, хоча найзагальніші закономірності, показані на матеріалі бази ORL чітко простежуються й у бази FERET.

Збільшення кількості зображень у класі також позитивно позначається на ефективності розпізнавання, незалежно від алгоритму. Використання кількості головних компонентів більше 80 практично не впливає на результати методу головних компонентів, а світлова нормалізація не впливає на ефективність методу локальних бінарних шаблонів [44]. Результати, показані на матеріалі бази FERET, досліджуваними методами з усіма вибраними варіантами вихідних параметрів наведено в таблиці 3.3.

Таблиця 3.3 – Результати, показані на матеріалі бази FERET

Кількість зображень у класі	Кількість оброблених зображень	Доля вдалих розпізнавань, %					
		Алгоритм розпізнавання					
		МГК			ЛБШ		
		80 ГК	160 ГК	160 ГК та світлова норм.	Радіус = 1	Радіус = 5	Радіус = 5. та світлова норм.
2	256	27	29	29	31	32	32
3	824	32	32	36	44	50	50
4	1502	32	32	38	45	56	55
5	1848	44	45	45	54	70	68
6	2006	39	40	41	50	62	62
7	1509	45	47	47	56	74	73
8	2571	45	47	47	55	75	74
9	1342	47	49	53	64	75	76
10	514	46	47	58	63	71	70
11	363	41	41	52	56	65	65
12	548	44	46	57	58	72	73
13	281	43	44	59	59	73	74
14	230	46	47	59	57	73	69
15	230	44	44	54	63	76	72
16	175	53	52	65	72	74	74
17	158	41	41	55	63	71	72
18	213	34	34	47	63	69	70
19	44	68	75	86	63	82	86
20	123	50	50	72	77	83	82
21	74	51	55	76	74	89	84
22	25	60	60	76	69	88	76
23	61	51	51	62	64	85	85
24	63	56	56	78	75	84	83
25	63	47	47	78	76	91	88
26	32	47	47	66	69	84	88
27	32	50	50	78	75	85	81
28	65	38	38	72	71	85	80
	15089	42	43	47	55	68	68

У той же час, для метода головних компонент світлова нормалізація дає відчутний приріст в ефективності, малопомітний при невеликій кількості зображень у класі, що зростає до 10% зі зростанням кількості зображень до 18. І до 20-30% при подальшому збільшенні кількості зображень у класі. Графіки результатів досліджених варіантів комбінації параметрів класифікатора, що працює методом головних компонентів, представлені на рисунку 3.10.

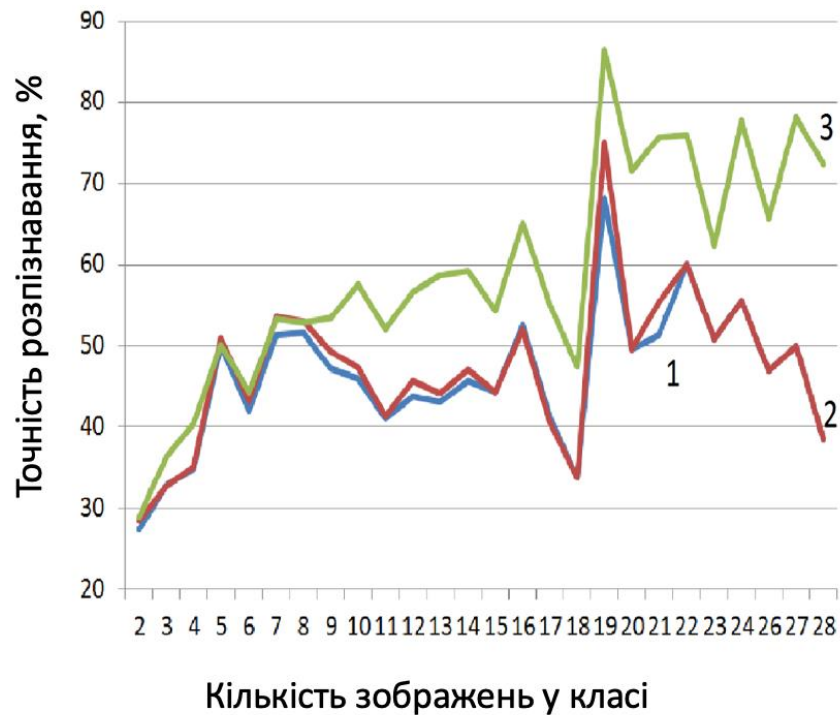


Рисунок 3.10 – Графіки результатів метода головних компонент з різними параметрами для бази FERET (1 – 80 головних компонент, 2 – 160 головних компонент без світлової нормалізації, 3 – 160 головних компонент зі світловою нормалізацією)

Для результатів методу локальних бінарних шаблонів помітно що збільшення радіусу кожного окремого патерну на матеріалі бази FERET позначається позитивно. Так, збільшення радіусу до 5, дає перевагу в результатах 5 – 20% для різної кількості зображень у класі. Ця картина очевидна щодо графіків результатів методу локальних бінарних шаблонів, зображених на рисунку 3.11.

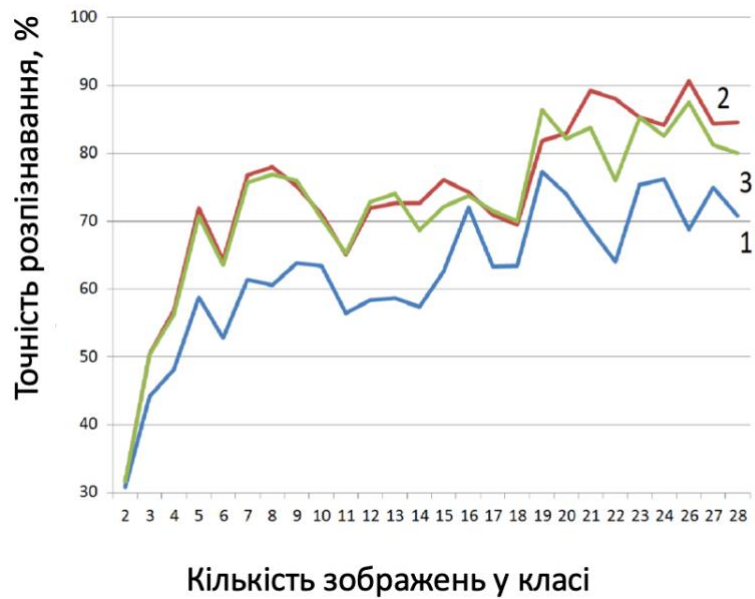


Рисунок 3.11 – Графік результатів методу локальних бінарних шаблонів з різними параметрами для бази FERET (1 – Радіус патерна = 1, 2 – Радіус патерна = 5 без світлової нормалізації, 3 – Радіус патерна = 5 зі світловою нормалізацією)

Таким чином, серед досліджених методів та різними варіантами вихідних параметрів, на базі зображень FERET найкращі результати для методу локальних бінарних шаблонів показав варіант із захопленим радіусом патерну, а для методу головних компонентів визначальним фактором виявилася світлова нормалізація. Серед кращих результатів обох методів, алгоритм локальних бінарних шаблонів, має перевагу над методом головних компонентів, в середньому, понад 10%, в окремих випадках ця різниця варіюється від 5 до 27%. На графіках (рис. 3.12) видно, що найбільша перевага спостерігається при обсязі вибірки в 5 – 8 зображень, де кількість випробувань, які пройшли методи, відчутно вище. Ця обставина дозволяє припустити, що це спостереження можна розцінювати як найбільш об'єктивне і незалежне від випадкових особливостей окремих класів зображень [45].

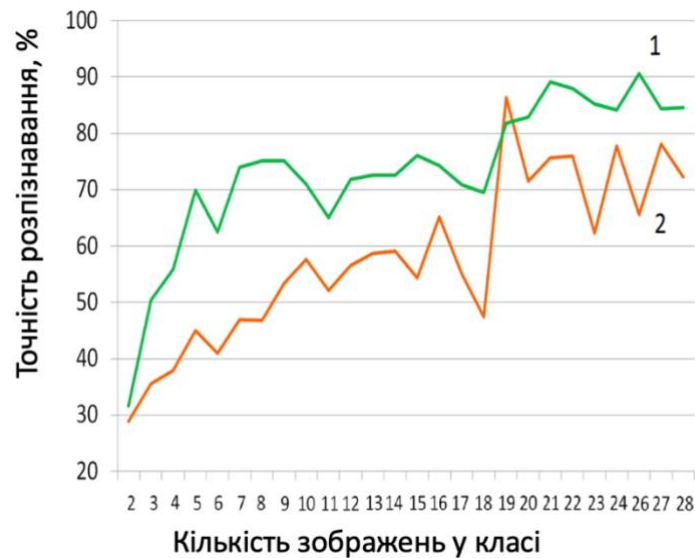


Рисунок 3.12 – Графіки найкращих результатів обох методів для бази FERET (1 – Радіус патерна = 1 без світлової нормалізації, 2 – 160 головних компонент зі світловою нормалізацією)

3.4.1 Порівняння результатів методів для баз ORL та FERET

Особливості результатів, показаних класифікаторами на матеріалі різних баз зображень, багато в чому визначаються специфікою зображень, включених в ці бази. Однак, вони дозволяють зробити висновки про загальні закономірності результативності аналізованих методів. Оскільки залежність ефективності ідентифікації від кількості зображень у класі не викликає сумніву, правомірно порівнювати лише результативність, показану на вибірках з однаковим обсягом тестового набору [46]. У цій ситуації це кількість в межах від 3 до 8 зображень. У таблиці 3.4 наведено отримані результати.

Таблиця 3.4 – Порівняння результатів методів для баз FERET та ORL

Кількість зображень у класі	Метод головних компонент						Метод локальних бінарних шаблонів					
	80 ГК		160 ГК		160 ГК та світлова норм.		Радіус = 1		Радіус = 5		Радіус = 5 та світлова норм.	
	FERET	ORL	FERET	ORL	FERET	ORL	FERET	ORL	FERET	ORL	FERET	ORL
3	32	69	32	69	36	75	44	75	50	56	50	56
4	32	80	32	80	38	80	45	92	56	74	55	74
5	44	86	45	86	45	84	54	92	70	85	68	86
6	39	87	40	87	41	88	50	95	62	87	62	86
7	45	90	47	90	47	92	56	95	74	94	73	94
8	45	93	47	93	47	92	55	98	75	96	74	96

Рисунок 3.13 наочно відображає співвідношення найкращих результатів методу головних компонент і методу локальних бінарних шаблонів на матеріалах баз ORL та FERET.

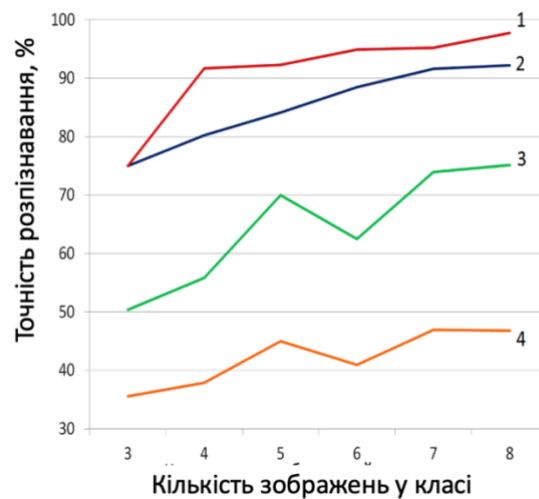


Рисунок 3.13 – Порівняння найкращих результатів методів для баз FERET та ORL (1 – метод локальних бінарних шаблонів на матеріалі ORL; 2 – метод головних компонент на матеріалі ORL; 3 – метод локальних бінарних шаблонів на матеріалі FERET; 4 – метод головних компонент на матеріалі FERET)

ВИСНОВКИ

У рамках кваліфікаційної роботи було досліджено різні підходи до класифікації методів розпізнавання облич, проаналізовано відповідні методи. Розглядаються існуючі підходи щодо побудови вибіркової сукупності.

Були вирішені такі завдання:

- вибрати методи та алгоритми, за допомогою яких буде проводитися ідентифікація та розпізнавання облич, виділених на окремому зображенні;
- визначити варіанти параметрів, що задаються, для даних методів, які можуть вплинути на ефективність розпізнавання;
- підготувати широку вибірку зображень, у різних умовах освітленості у яких присутні зображення облич із різних ракурсів, заздалегідь співвіднесених з певними людьми;
- визначити методику, за допомогою якої наявна вибірка буде найбільш об'єктивним способом розділена на навчальний та тестовий набори;
- провести обробку зображень із застосуванням різних методів, задаючи різні значення параметрів;
- застосувати обрані варіанти обробки вихідного матеріалу для різних поєднань зображень у навчальному та тестовому наборах;
- співвіднести отримані результати між собою, тим самим порівняти ефективність.

Для досягнення більшої ефективності було вирішено такі питання:

- розроблено алгоритм формування тестової та навчальної вибірки з неоднорідного вихідного набору даних;
- з використанням цього алгоритму проведено дослідження та порівняння ефективності методів ідентифікації людини за зображенням обличчя;
- виявлено особливості вихідного набору даних, при яких певний метод ідентифікації має найбільші переваги;

- визначено, які вхідні параметри методів більшою мірою впливають на ефективність та точність методів розпізнавання обличчя;
- на підставі результатів дослідження було сформульовано рекомендації щодо використання методів розпізнавання облич.

За результатами вивчення цих питань були обрані методи для дослідження та порівняння ефективності ідентифікації людини за зображенням обличчя: метод головних компонентів та метод локальних бінарних шаблонів.

Вихідним матеріалом для оцінки та порівняння ефективності методів стали бази даних ORL та FERET. Для об'єктивного порівняння результатів ідентифікації було підготовлено кілька варіантів вибірок зображень з розбиттям наявних наборів зображень у класах на різну кількість частин. Шляхом перехресної перевірки результатів для аналогічних наборів даних була проведена обробка отриманих наборів за допомогою досліджуваних методів із застосуванням різних параметрів.

Порівняння результатів даної обробки дозволило зробити такі висновки:

- в цілому, метод локальних бінарних шаблонів має більшу ефективність, ніж метод головних компонент як на матеріалі бази FERET, так і на матеріалі ORL;
- зі збільшенням кількості зображень у класі ефективність ідентифікації зростає на матеріалі обох баз для обох методів з будь-яким поєднанням вибраних параметрів;
- на матеріалі бази ORL зі збільшенням кількості зображень у класі відмінності між ефективністю методів зменшуються;
- на матеріалі бази FERET зі збільшенням кількості зображень у класі відмінності між ефективністю методів незначно збільшуються;
- на матеріалі бази FERET перевага в ефективності методу локальних бінарних шаблонів над методом головних компонент більша, ніж на матеріалі бази ORL;

- вплив світлової нормалізації на результати методу локальних бінарних шаблонів мінімальний для обох баз;
- метод головних компонент дещо ефективніший при світловій нормалізації на матеріалі бази FERET, вплив цього параметра менш помітний, ніж на матеріалі бази ORL;
- на матеріалі бази ORL збільшення радіусу знижує ефективність методу локальних бінарних шаблонів, на FERET ситуація обернена;
- найнижчий показник ефективності ідентифікації зображень облич у методу головних компонент без світлової нормалізації, при малій кількості зображень у класі;
- кращі результати ефективності ідентифікації зображень облич показує метод локальних бінарних шаблонів з мінімальним радіусом без світлової нормалізації на матеріалі бази ORL при великій кількості зображень у класі.

На підставі зроблених висновків, можна зробити висновок, що метод локальних бінарних шаблонів, практично у всіх досліджених варіантах поєднання вхідних параметрів і вихідних даних, представлених для обробки, перевищує ефективність розпізнавання методу головних компонент. Ця перевага стає очевидною, коли вихідний набір даних має досить великий обсяг, тим самим знижується ймовірність помилки в результатах дослідження, нав'язаної специфікою матеріалу, що обробляється, і фактором випадковості. При використанні у якості вихідного матеріалу зображень спочатку мають меншу «ступінь нормалізації», така перевага куди більш відчутна. З цієї причини в реальних умовах можна рекомендувати використання методу ЛБШ.

У процесі дослідження стало очевидним, що навчання класифікаторів при великому обсязі вихідних даних дуже вимогливе до апаратних ресурсів і займає досить тривалий час. Частково цю проблему вдалося вирішити шляхом розпаралелювання процесу обробки даних та запуску її в кілька потоків. Однак, на окремій робочій станції обмеження продуктивності все ж таки виявилися дуже відчутними. Внаслідок цього для вирішення масштабних

завдань рекомендується використовувати розподілені обчислення безпосередньо для обробки даних та продуктивну систему зберігання даних для запису та зберігання навчених класифікаторів.

Результати кваліфікаційної роботи апробовано у вигляді статті під час Міжнародної науково-практичної конференції «INNOVATIONS AND PROSPECTS OF WORLD SCIENCE» [47].

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Suma, S. L., & Raga, S. (2018). Real time face recognition of human faces by using LBPH and Viola Jones algorithm. *International Journal of Scientific Research in Computer Science and Engineering*, 6(5), 6-10.
2. Rössler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2018). Faceforensics: A large-scale video dataset for forgery detection in human faces. arXiv preprint arXiv:1803.09179.
3. Гончаренко, М. О., & Машталир, С. В. (2012). Детектирование изменений сцены в потоке видеоданных. *Электротехнические и компьютерные системы*, (7), 143-147.
4. Машталир, С. В., Мегель, Ю. Е., & Путятин, Е. П. (2004). Анализ изображений биологических объектов по бинарным сечениям. *Радиоэлектроника и информатика*, (2 (27)).
5. Mashtalir, S., & Mikhnova, O. (2017). Detecting significant changes in image sequences. In *Multimedia Forensics and Security*, pp. 161-191. Springer, Cham.
6. Le, T. H. N., Quach, K. G., Zhu, C., Duong, C. N., Luu, K., Savvides, M., & Center, C. B. (2017, July). Robust Hand Detection and Classification in Vehicles and in the Wild. In *CVPR Workshops*, pp. 1203-1210.
7. Машталир, С. В., & Сакало, Е. С. (2008). Адаптивное нейросетевое сжатие сигналов большой размерности на основе взвешенного информационного критерия.
8. Егоров, А. С., & Машталир, С. В. (2005). Сравнительный анализ методов морфологической нормализации. *Радиоэлектроника и информатика*, (4).
9. Машталир, С. В., & Столбовой, М. И. (2018). Адаптивные матричные модели в задаче контроля потоков видео. *Радіоелектроніка, інформатика, управління*, (4 (47)).

10. Pitsikalis, V., Katsamanis, A., Theodorakis, S., & Maragos, P. (2017). Multimodal gesture recognition via multiple hypotheses rescoring. In *Gesture Recognition*, pp. 467-496. Springer, Cham.
11. Ведмедь, А. Г., Машталир, С. В., & Сакало, Е. С. (2010). Восстановление изображений с использованием анализа главных и независимых компонент. *Системы обработки інформації*, (6), 66-72.
12. Xi, C., Chen, J., Zhao, C., Pei, Q., & Liu, L. (2018, February). Real-time hand tracking using kinect. In *Proceedings of the 2nd International Conference on Digital Signal Processing*, pp. 37-42.
13. Mashtalir, S., & Mikhnova, O. (2017). Detecting significant changes in image sequences. In *Multimedia Forensics and Security*, pp. 161-191. Springer, Cham.
14. Машталир, С. В., Столбовой, М. И., & Яковлев, С. В. (2019). Кластеризация последовательностей видеоданных на основе гармонических k -средних. *Кибернетика и системный анализ*.
15. Машталир, С. В., Столбовой, М. И., & Яковлев, С. В. (2019). Кластеризация последовательностей видеоданных на основе гармонических k -средних. *Кибернетика и системный анализ*.
16. Богучарский, С. И., & Машталир, С. В. (2015). Модифицированный метод кластеризации X -средних в задачах сегментация изображений. *Електротехнічні та комп'ютерні системи*, (20), 106-110.
17. Мантула, Е. В., & Машталир, С. В. (2014). Адаптивная полиномиальная нейросетевая прогнозирующая модель временных рядов и ее обучение. *Восточно-Европейский журнал передовых технологий*, 2(4 (68)).
18. Suresh, M., Sinha, A., & Aneesh, R. P. (2019). Real-Time Hand Gesture Recognition Using Deep Learning. *International Journal of Innovations and Implementations in Engineering*, 1.
19. Богучарский, С. И., Каграманян, А. Г., & Машталир, С. В. (2014). Иерархическая агломеративная кластеризация изображений в больших базах данных. *Системы обработки інформації*, (8), 93-97.

20. Mashtalir, S., Mashtalir, V., & Stolbovyi, M. (2018, August). Representative based clustering of long multivariate sequences with different lengths. In 2018 IEEE second international conference on Data Stream Mining & Processing (DSMP), pp. 545-548. IEEE.

21. Качур, С. А. (2006). Критерий управляемости систем на основе сетей Петри. Радиоэлектроника и информатика, (2).

22. Богучарский, С. И., & Машталир, С. В. (2014). Кластеризация коллекций изображений в больших базах данных на основе рекуррентной оптимизации.

23. Le, T. H., & Vo, T. N. (2012). Face alignment using active shape model and support vector machine. arXiv preprint arXiv:1209.6151.

24. Ложников, П. С., Сулавко, А. Е., Еременко, А. В., & Волков, Д. А. (2016). Экспериментальная оценка надежности верификации подписи сетями квадратичных форм, нечеткими экстракторами и перцептронами. Информационно-управляющие системы, (5 (84)).

25. Ведмедь, А. Г., Машталир, С. В., & Сакало, Е. С. (2010). Восстановление изображений с использованием анализа главных и независимых компонент. Системи обробки інформації, (6), 66-72.

26. Биленко, А. Ю., & Нестругина, Е. С. (2021). РАСПОЗНАВАНИЕ ЛИЦ НА ИЗОБРАЖЕНИЯХ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ. Вестник Донецкого национального университета. Серия Г: Технические науки, (1), 33-44.

27. Нестеров, В. Н., & Ли, А. Р. (2017). Принцип двухканальности в реализации инвариантных измерительных преобразователей и систем. In Проблемы получения, обработки и передачи измерительной информации, pp. 126-131.

28. Богучарский, С. И., Каграманян, А. Г., & Машталир, С. В. (2015). Модификация метода J-средних в задачах фрагментной сегментации изображений.

29. Гонсалес, Р., & Вудс, Р. (2019). Цифровая обработка изображений. Litres.
30. Bodyanskiy, Y. V., Tyshchenko, O. K., & Mashtalir, S. V. (2019, June). Fuzzy clustering high-dimensional data using information weighting. In International Conference on Artificial Intelligence and Soft Computing, pp. 385-395. Springer, Cham.
31. Богучарский, С. И., & Машталир, С. В. (2014). Анализ текстур в последовательности изображений на основе векторного квантования. *Радіоелектроніка, інформатика, управління*, (2 (31)).
32. Кафтанников, И. Л., & Парасич, А. В. (2016). Проблемы формирования обучающей выборки в задачах машинного обучения. *Вестник Южно-Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника*, 16(3).
33. Парпиева, А. Ж. (2017). МЕТОД ГЛАВНЫХ КОМПОНЕНТОВ РАСПОЗНАВАНИЯ ЛИЦ. *Научные горизонты*, (4), 218-223.
34. Егорова, Е. А., Киношенко, Д. К., Машталир, С. В., & Шляхов, Д. В. (2006). Метрическое сравнение результатов сегментации изображений. *Радиоэлектроника и информатика*, (2).
35. Мокеев, В. В. (2011). О повышение эффективности вычислений главных компонент в задачах анализа изображений. *Цифровая обработка сигналов*, (4), 29-36.
36. Катасёв, А. С., Катасёва, Д. В., & Кирпичников, А. П. (2016). Нейросетевая биометрическая система распознавания изображений человеческого лица. *Вестник Казанского технологического университета*, 19(18).
37. Чернодуб А. Генетические алгоритмы в лицах/ URL: <http://habrahabr.ru/post/221137/> (дата звернення 25.10.2021)
38. Library documentation/Microsoft /Developer Network// 2019. URL: <https://msdn.microsoft.com/library/67ef8sbd.aspx> (дата звернення 27.10.2021)

39. Library documentation/Microsoft /Developer Network// 2019. URL: <https://msdn.microsoft.com/library/dd831853.aspx> (дата звернення 27.10.2021)
40. OpenCV Documentation/ Face Recognition with OpenCV /OpenCV API Reference //URL: <http://docs.opencv.org/2.4/modules/contrib/doc/facerec/> (дата звернення 28.10.2021)
41. The ORL Database of Faces, /AT&T Laboratories Cambridge 2018. //URL: <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html> (дата звернення 01.11.2021)
42. The Color FERET Database, / The National Institute of Standards and Technology (NIST) 2009. // URL: <http://www.nist.gov/itl/iad/ig/colorferet.cfm> (дата звернення 05.11.2021)
43. Zafaruddin, G. M., & Fadewar, H. S. (2019). Face recognition using eigenfaces. In Computing, communication and signal processing, pp. 855-864. Springer, Singapore.
44. MAAFIRI, A., & CHOUGDALI, K. (2019, March). Face recognition using wavelets based feature extraction and PCA-L1 norm. In 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), pp. 1-4. IEEE.
45. Král, P., Vrba, A., & Lenc, L. (2019, June). Enhanced local binary patterns for automatic face recognition. In International conference on artificial intelligence and soft computing, pp. 27-36. Springer, Cham.
46. Tuncer, T., Dogan, S., Abdar, M., Ehsan Basiri, M., & Pławiak, P. (2019). Face recognition with triangular fuzzy set-based local cross patterns in wavelet domain. *Symmetry*, 11(6), 787.
47. Tovstolies D. (2021) Comparative analysis of facial recognition methods, Proceedings of III International Scientific and Practical Conference «Innovations and prospects of world science» (November 4 – 6, 2021). Vancouver, Canada, pp.254-256.