

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

Факультет Інформаційно-аналітичних технологій та менеджменту  
(повна назва)

Кафедра Інформатики  
(повна назва)

**КВАЛІФІКАЦІНА РОБОТА**  
**Пояснювальна записка**

рівень вищої освіти перший (бакалаврський)

**РОЗРОБЛЕННЯ ЗАСТОСУНКУ ДЛЯ ІДЕНТИФІКАЦІЇ ОБЛИЧ НА**  
**КАДРАХ ВІДЕОЗЙОМКИ**

(тема)

Виконав:  
студент 4 курсу, групи ІТІНФ-17-1

Кухарчук В.А.  
(прізвище, ініціали)

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

Тип програми освітньо-професійна

Освітня програма Інформатика  
(повна назва освітньої програми)

Керівник доц. Творошенко І.С.  
(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 перший (бакалаврський)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 Кухарчуку Василю Андрійовичу  
(прізвище, ім'я, по батькові)1. Тема роботи Розроблення застосунку для ідентифікації облич на кадрах відеозйомки

затверджена наказом по університету від 20 травня 2021 року №663Ст

2. Термін подання студентом роботи до екзаменаційної комісії 24 травня 2021 р.

3. Вихідні дані до роботи науково-методична та науково-технічна література, матеріали конференцій, дані інтернет-мережі, бібліотека комп'ютерного зору з відкритим кодом OpenCV, мова програмування C++, середовище розробки Microsoft Visual Studio.

4. Перелік питань, що потрібно опрацювати в роботі \_\_\_\_\_

1. Аналіз наявних застосунків для ідентифікації облич на зображеннях та кадрах відеозйомки в Україні та за кордоном.2. Розроблення підходу для ідентифікації облич на кадрах відеозйомки.3. Розроблення застосунку для ідентифікації облич на кадрах відеозйомки.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) Актуальність проблеми розпізнавання облич, постановка задачі, сучасні методи розпізнавання, вибір методу для реалізації, навчання моделі, відео-розпізнавання.

---



---



---



---



---



---



---



---

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

### КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	12.04.2021	
2	Аналіз завдання, підбір літератури	13.04.21-16.04.21	
3	Аналіз літератури з досліджуваної проблеми	17.04.21-20.04.21	
4	Аналіз технічних засобів	21.04.21-25.04.21	
5	Розробка методу	25.04.21-03.05.21	
6	Програмна реалізація	04.05.21-18.05.21	
7	Оформлення пояснювальної записки	19.05.21-26.05.21	
8	Перевірка на плагіат	27.06.21	
9	Рецензування	28.05.21	
10	Підготовка презентації та доповіді	29.05.21-30.05.21	
11	Занесення роботи в електронний архів	31.05.21	
12	Попередній захист кваліфікаційної роботи	31.05.21	

Дата видачі завдання 12 квітня 2021 р.

Студент \_\_\_\_\_  
(підпис)

Керівник роботи \_\_\_\_\_ доц. Творошенко І.С.  
(підпис) (посада, прізвище, ініціали)

## РЕФЕРАТ/ABSTRACT

Пояснювальна записка до кваліфікаційної роботи: 100 с., 34 рис., 1 дод., 54 джерела.

КОМП'ЮТЕРНИЙ ЗІР, РОЗПІЗНАВАННЯ ОБРАЗІВ,  
РОЗПІЗНАВАННЯ ОБЛИЧЧЯ, КЛЮЧОВІ ТОЧКИ, МАШИННЕ  
НАВЧАННЯ, МІСЦЕВІ ДВІЙКОВІ ВІЗЕРУНКИ, СИСТЕМА  
РОЗПІЗНАВАННЯ ОБЛИЧ.

Об'єктом роботи є послідовність кадрів відеозйомки із людськими обличчями.

Метою роботи є розробка застосунку для розпізнавання облич на послідовності кадрів відеозйомки.

Використано методи глибокого навчання для знаходження обличчя в кадрі та місцеві двійкові візерунки для виділення текстурних ознак обличчя. Проведено дослідження методів виділення важливих ознак та класифікації об'єктів на зображеннях. Досліджено метод місцевих двійкових візерунків та розроблено алгоритм розпізнавання облич.

У результаті роботи здійснена програмна реалізація системи для розпізнавання облич на кадрах відеозйомки.

Результат роботи апробовано у вигляді 2 тез доповідей під час Міжнародного молодіжного форуму «РАДІОЕЛЕКТРОНІКА І МОЛОДЬ У XXI СТОЛІТТІ» та XXVI Міжнародної науково-практичної конференції «Topical issues of practice and science».

COMPUTER VISION, IMAGE RECOGNITION, FACIAL  
RECOGNITION, KEY POINTS, MACHINE LEARNING, LOCAL BINARY  
PATTERNS, FACIAL RECOGNITION SYSTEM.

The object of the research is the sequence of video frames with human faces.

The aim of the research is to develop an application for face recognition on a sequence of video frames.

Deep learning methods were used to find the face in the frame and local binary patterns to excrete the textural features of the face. The research of methods of selection of important features and classification of objects on images. The method of local binary templates and the algorithm of face recognition.

As a result, the software implementation of the system for face recognition in video footage was implemented.

The result of the work was tested in the form of 2 abstracts of the report during the International Youth Forum «RADIO ELECTRONICS AND YOUTH IN THE XXI CENTURY» and XXVI International Scientific and Practical Conference «Topical issues of practice and science».

## ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів .....	7
Вступ .....	8
1 Аналіз наявних застосунків для ідентифікації облич на зображеннях та кадрах відеозйомки в Україні та за кордоном .....	10
1.1 Сучасний стан розвитку застосунків для ідентифікації облич на зображеннях та кадрах відеозйомки в Україні та за кордоном .....	10
1.1.1 Огляд основних методів фіксації та обробки зображень для ідентифікації осіб .....	12
1.1.2 Використання технології розпізнання облич у системах перевірки посвідчення особи .....	14
1.1.3 Використання технології розпізнавання облич у мобільній індустрії, персональних комп'ютерах та вебзастосунках .....	18
1.1.4 Використання технології розпізнавання облич у муніципальних та приватних системах стеження .....	21
1.2 Аналіз літературних джерел щодо існуючих підходів ідентифікації облич на зображеннях та кадрах відеозйомки .....	27
1.3 Постановка задачі .....	31
2 Розроблення підходу для ідентифікації облич на кадрах відеозйомки .....	32
2.1 Методи визначення особливих точок на зображеннях .....	32
2.1.1 Кути Гарріса .....	33
2.1.2 Детектор ознак FAST .....	37
2.1.3 Детектор ознак SIFT .....	38
2.1.4 Детектор ознак SURF .....	43
2.1.5 Детектор ознак та дескриптор BRISK .....	46
2.2 Класифікація зображень на підставі статистичних характеристик .....	49
2.2.1 Баєсів підхід до прийняття рішень .....	49
2.2.2 Двійкові дерева рішень .....	52
2.2.3 Випадкові ліси .....	55

	6
2.2.4	Метод $k$ -найближчих сусідів ..... 57
2.2.5	Метод опорних векторів ..... 60
2.3	Двовимірні підходи до розпізнавання облич ..... 63
2.3.1	Цілісний підхід..... 64
2.3.2	Місцевий підхід..... 68
2.3.3	Глибоке навчання..... 74
2.4	Розроблення методики ідентифікації облич на кадрах відеозйомки..... 75
3	Розроблення застосунку для ідентифікації облич на кадрах відеозйомки.. 79
3.1	Вибір інструментальних засобів для створення застосунку ..... 79
3.2	Етапи розроблення застосунку для ідентифікації облич на кадрах відеозйомки..... 80
3.3	Тестування розробленого застосунку та аналіз результатів ..... 84
3.4	Перспективи дальшої роботи..... 87
	Висновки..... 89
	Перелік джерел посилання ..... 91

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ  
І ТЕРМІНІВ

СРО – система розпізнавання обличчя

США – Сполучені Штати Америки

БД – база даних

РГ – різниця гаусіанів

LBP – local binary pattern

## ВСТУП

Розпізнавання облич – один із найбільш активно досліджуваних напрямків комп'ютерного зору та розпізнавання образів, а також один з найпопулярніших методів біометричної ідентифікації. На сьогоднішній день, воно має безліч практичних застосувань, серед яких автентифікація, контроль доступу, криміналістика, системи безпеки тощо. Царина використання розпізнавання облич поширилася навіть на ринок послуг та розваг. Наразі рівень розвитку технології вже дозволяє використовувати її в сфері безпеки, наприклад для проведення фінансових операцій та для перевірки посвідчення особи на підприємствах. Системи біометричного розпізнавання на основі обличчя широко використовуються в таких місцях, як аеропорти, митниця, міграційні бюро, а також просто на вулицях міст та навіть мобільних телефонах та комп'ютерах звичайних людей.

Така популярність технології пов'язана з тим, що для її використання часту досить лише звичайної камери, без додаткових спеціалізованих пристроїв. Крім того, технологія є безконтактною, що несе за собою цілий перелік переваг.

За визначенням, біометрична ідентифікація – це метод розпізнавання, який використовує одну або декілька фізичних або поведінкових характеристик людського тіла для розпізнавання особистості. Для цього в техніці біометричної ідентифікації використовуються лише унікальні особливості тіла (у випадку систем розпізнавання облич – унікальні особливості людського лиця).

Біометричні системи, що базуються на даних про обличчя, відбитки пальців, рисунок вен та райдужну оболонку, застосовуються вже багато років, добре досліджені, є дуже точними та ефективними, а також зручними для користувачів. Окрім того, ці дані важко викрасти, порівняно із звичайними методами ідентифікації, що ґрунтуються на картках, посвідченнях, пін-кодах та паролях.

Не зважаючи на нижчу точність порівняно із іншими біометричними підходами, завдяки зусиллям вчених та розробників були запропоновані важливі методи, алгоритми, методології та бази даних для розпізнавання облич в обмеженому та необмеженому середовищах, тому багато досліджень показали, що наразі точність систем розпізнавання обличчя в багатьох випадках порівнянна або перевершує ефективність розпізнавання людини.

Проте, здебільшого, найвища ефективність досягається все ж саме в обмежених середовищах, де контролюються параметри зйомки, такі як освітлення, кут зору та відстань між камерою та об'єктом. У випадку порушення цих умов, точність багато-коли різко погіршується.

З появою глибокого навчання розвиток комп'ютерного зору і розпізнавання облич зокрема дуже пришвидшився, а якість отриманих результатів значно покращилася. Проте глибоке навчання має ряд недоліків і не може бути застосоване в усіх випадках, тому класичні підходи все ще залишаються актуальними і ставлять виклики перед дослідниками.

# **1 АНАЛІЗ НАЯНИХ ЗАСТОСУНКІВ ДЛЯ ІДЕНТИФІКАЦІЇ ОБЛИЧ НА ЗОБРАЖЕННЯХ ТА КАДРАХ ВІДЕОЗЙОМКИ В УКРАЇНІ ТА ЗА КОРДОНОМ**

## **1.1 Сучасний стан розвитку застосунків для ідентифікації облич на зображеннях та кадрах відеозйомки в Україні та за кордоном**

Людське обличчя – це майже унікальна візуальна характеристика кожної людини. Воно завжди дозволяло людині з першого погляду дізнатися, чи вона зустріла знайомого, чи навпаки, цю особу вона бачить вперше. У повсякденному житті обличчя є, мабуть, найпоширенішою та найвідомішою біометричною характеристикою. З винаходом фотографії державні відомства та приватні організації зберігали світлини з обличчям (з особистих документів, що посвідчують особу, паспортів або членських карток). Ці колекції використовувались у розслідуваннях як довідкові бази даних для зіставлення та порівняння зображень обличчя респондента (наприклад, злочинця, свідка чи жертви). Крім того, широке використання цифрових камер та смартфонів полегшило створення зображень обличчя щодня; ці зображення можна легко розповсюджувати та обмінювати за допомогою швидко поширюваних соціальних мереж, таких як фейсбук і твіттер.

За останні десятиліття розвиток технологій в електроніці та інформатиці забезпечив для вагомої частини населення доступ до технологічних пристроїв високого рівня за доступними цінами. Біометричні практики замінюють звичайні рішення, засновані на знаннях, такі як паролі чи PIN-коди, та стратегії, засновані на володінні, такі як посвідчення особи чи бейджі [1]. Різні біометричні системи широко використовуються в реальних застосунках, таких як онлайн-платежі та електронна комерція, автентифікація за допомогою смартфонів, біометричні паспорти та прикордонний контроль.

Людське обличчя не є ідеальним ідентифікатором порівняно з іншими біометричними ознаками; він, як правило, менш точний, ніж інші біометричні

методи, такі як райдужна оболонка ока або відбитки пальців. Однак обличчя має ті переваги, які роблять його однією з найпопулярніших і найперспективніших біометричних характеристик для розпізнавання особистості [2].

На відміну від решти біометричних методів, розпізнавання обличчя – це спосіб, який можна застосовувати без будь-якої взаємодії з користувачем-датчиком і може бути застосовано непомітно в системах спостереження. Окрім того: сенсорний пристрій (тобто камеру) легко монтувати, це не дорого, це не потребує предметної взаємодії, немає гігієнічних проблем, а через мінімальну потребу в будь-яких діях, люди радше віддають перевагу цьому методу [3]. Саме тому з розвитком комп'ютерного зору в розробників та вчених з'явилося бажання втілити в життя таку важливу людську здібність, як розпізнавання обличчя, в програмній формі.

Однак, ця задача виявилася досить складною, адже вона пов'язана з низкою проблем, які швидко вирішує наш мозок, проте для комп'ютерів вони стають непереборною перешкодою. Окрім складності реалізації самої ідеї автоматичного розпізнання, продуктивність таких застосунків стрімко обмежується реальними умовами використання, коли зображення обличчя отримують у неконтрольованому середовищі (зміні освітлення, постави або виразу лиця, часткова оклюзія, маскування чи рух камери).

І хоча зазначені складнощі все ще ставлять перед розробниками певні виклики, а сьогоднішні технології почасти все ще далекі від можливостей зорової системи людини [4], та, на щастя, з початку проведення досліджень у цьому напрямку було здійснено величезний стрибок уперед, тому сучасні системи ідентифікації досягли певної міри зрілості та вже набули значного поширення в багатьох сферах життя й використовуються навіть у побуті звичайними людьми [5]. Згідно з доповіддю аналітичної компанії Mordor Intelligence, ринок розпізнання обличчя у 2019 році оцінювався у 4,4 мільярди доларів, а в 2025 році має перевищити 10,9 мільярдів. Це означає, що дана галузь швидко зростає та є дуже перспективною.

### 1.1.1 Огляд основних методів фіксації та обробки зображень для ідентифікації осіб

На сьогоднішній день існує декілька підходів до аналізу обличчя для його розпізнавання. Традиційний метод ідентифікує окремі риси (ключові точки) обличчя, їх розмір, форму та взаємне розташування, наприклад, відстань між очима, ширину щелепи та носа, форму вилиць. З цих точок створюється числовий код, що називається відбитком обличчя, який заноситься до бази даних. Отримані значення зіставляються з наданими для розпізнання прикладами та, у випадку знаходження відповідності, особистість на зображенні ідентифікується. Традиційний метод нині є найпоширенішим, бо для його застосування не потрібне специфічне обладнання, він працює на відстані (для цього зображення облич з низькою роздільною здатністю покращують за допомогою спеціальних алгоритмів), а необхідні для розпізнання фото значно легше отримати. Проте водночас в даного методу є низка недоліків, таких як чутливість до оклюзії, повороту, зміни виразу обличчя, для подолання яких доводиться вигадувати все новіші алгоритми.

Варто зазначити, що в ранні часи наукові інтереси в основному були зосереджені на розпізнаванні обличчя в контрольованих умовах, коли прості класичні підходи забезпечували непогані результати. Сьогодні ж фокус досліджень зосереджений на необмежених умовах, тому в останні роки практично в усіх застосунках для розпізнання обличчя використовується машинне навчання, яке заступило старіші методи через значно ширші перспективи розвитку та якість отриманих результатів. А з появою глибокого навчання на початку 2010-х років показники у вирішенні багатьох задач комп'ютерного зору, таких як класифікація зображень та розпізнання об'єктів, різко покращилися та саме завдяки йому було досягнуто такої точності розпізнання за наявності факторів, що ускладнюють роботу програми [6]. З'явилося багато спеціалізованих згорткових нейронних мереж, наприклад

DeepFace, DeepIDs, VGG Face, FaceNet, SphereFace та ArcFace, які на сьогодні, згідно з дослідженнями, мають приблизно однакову точність розпізнання.

Іншим методом ідентифікації особистости за обличчям, що зараз стає дедалі популярнішим, є тривимірне розпізнання. Для цього використовуються інфрачервоні датчики, які сканують лице та створюють його тривимірний відбиток, а програма потім виявляє унікальні риси, що не змінюються з часом, наприклад, форма носа та підборіддя, вигини очниць. Завдяки тому, що в цьому методі використовуються інфрачервоні датчики, а вимірюється тільки форма, на яку не впливає освітлення, тривимірне розпізнання можна використовувати навіть у темряві та під різними кутами огляду аж до повороту в 90 градусів (обличчя в профіль).

Хоч 3D-технологія і є більш точною та стійкою до змін у зовнішності та куті нахилу, ніж традиційна, проте вона також не позбавлена своїх значних недоліків. По-перше, для неї необхідне спеціальне обладнання, ІЧ-датчики, які проєктують структуроване світло на обличчя, що збільшує ціну й ускладнює використання цієї технології. По-друге, через зазначені особливості подібні системи менш пристосовані до слідкування за множиною людей на віддаленій відстані (хоча й використовуються у цій сфері). І по-третє, на сьогодні існує недостатня кількість необхідних наборів даних для навчання та дослідження застосунків тривимірного розпізнання обличчя (це зумовлено, зокрема, першим недоліком) [7].

Ще один перспективний напрямок – аналіз текстури шкіри, який використовують рідше окремо, а частіше в купі з іншими методами. Застосунки, що послуговуються цим методом, виділяють із зображення обличчя так звані відбиток шкіри, поділяють його на окремі ділянки й аналізують усі видимі лінії, пори й інші текстурні особливості. Використовуючи таку техніку можливо виявити відмінності між однойцевими близнятами, чого поки не можна досягти за допомогою звичайних програм для розпізнання обличчя. Поєднання традиційного

способу ідентифікації з аналізом текстури шкіри дозволяє відчутно збільшити точність розпізнання, хоча на відстані це, звичайно, зробити неможливо.

Й останній з найбільш відомих і перспективних напрямків, про які хотілося б згадати, – це використання теплових камер. Завдяки ним система ігнорує зовнішні перепони (макіяж, маски й навіть перегородки) та може використовуватися навіть вночі. Головними недоліками цього методу є, як і в тривимірному розпізнанні, необхідність у спеціалізованому обладнанні та нестача навчальних наборів даних. Проте, в 2018 році дослідники з США знайшли спосіб порівнювати знімки облич, отримані за допомогою теплових камер, зі звичайними фото, що відкриває можливості для більш гнучкого використання та розвитку в загалом для цієї технології.

### 1.1.2 Використання технології розпізнання облич у системах перевірки посвідчення особи

Одна зі сфер, в якій застосовується розпізнавання облич, – це системи перевірки посвідчення особи. Саме в цій галузі застосункам для розпізнання обличчя доводиться конкурувати з іншими, поки що більш точними й пристосованими до цього, біометричними системами, такими як сканери відбитків пальців і райдужної оболонки ока. Але через причини, що були описані раніше (відсутність фізичного контакту та мінімальна кількість необхідних дій з боку людини), такі застосунки активно використовуються по всьому світу. Обличчя вже замінили ID-картки та пропуски на вході до багатьох підприємств – досить лише подивитися в камеру і, якщо людина є в базі співпрацівників, система її пропустить.

Аеропорти застосовують системи розпізнавання для ідентифікації облич на міжнародних рейсах. Починаючи з 2018 року, митна та прикордонна служби США почали розміщувати сканери обличчя в американських аеропортах. Замість перевірки документів і посадкового талона на

прикордонному контролю або під час реєстрації на рейс пасажирів просять просто подивитись у камеру. Фотографії, які роблять для мандрівників із громадянством США, протягом 12 годин видаляють. Для аналізу обличчя необхідно 2 секунди, а точність складає 99%. Управління транспортної безпеки США має наміри в майбутньому використовувати цю систему й для внутрішніх повітряних перевезень.

Прикордонна служба Австралії і митна служба Нової Зеландії запровадили автоматизовану систему обробки пасажиропотоку під назвою «SmartGate» (рис. 1.1). Вона використовує розпізнавання обличчя для порівняння особи мандрівника з даними в його електронному паспорті. На початку 2017 року в Канаді, в Міжнародному аеропорту Ванкувера, для тих же цілей було встановлено систему первинної інспекції із застосунком для розпізнання осіб. Потім, у 2018-2019 роках, її розгорнули в решті канадських міжнародних аеропортів.



Рисунок 1.1 – Система SmartGate в аеропорту Австралії

Ідентифікацію особистості використовують і в метро. Наприклад, в японському місті Осака у 2020 році після попереднього тестування почали встановлювати турнікети, що використовують технологію розпізнання облич.

Місцева влада має на меті встановити такі турнікети на всіх станціях метро до початку Всесвітньої Виставки у 2025 році.

До речі, на початку 2020 року в метро столиці Грузії Тбілісі також було додано можливість придбання квитка й швидкого проходу через турнікет, використовуючи біометричні дані обличчя. Розробником встановленої системи стала українська компанія Riddletag. Цього досягли завдяки партнерству з приватним банком Грузії та компанією, що видає кредитні картки VISA. Користувачі метро можуть завантажити безкоштовний застосунок банку та зареєструвати своє обличчя, після чого кошти з їх рахунку списуватимуться автоматично після сканування лиця на проході через турнікет.

Згадана компанія виготовляє продукти й для інших сфер, в яких застосунки для розпізнання особистості наразі чимраз більше користуються попитом. Наприклад, банківська. В банках деяких країн упровадили обов'язкове сканування обличчя для людей, що хочуть отримати кредит. Це гарантує, що вони не причетні до фінансових махінацій та мають хорошу кредитну історію, а також захищає самих клієнтів від неправомірних дій щодо них.

Так само, як і у випадку з метро, можна сплачувати покупки в магазинах, замовлення в закладах харчування тощо. Великі мережі домовляються з банками та встановлюють на касах системи біометричної ідентифікації, щоб в їх клієнтів була можливість зручної та швидкої сплати, без використання терміналу, карток або онлайн-гаманців. Окрім зручності для клієнтів, використання такої технології розширює можливості програм лояльності та дозволяє власникам збирати інформацію про відвідувачів, про їх вподобання, дізнаватися, хто що купує. Такий застосунок цієї технології наразі розвинувся до систем стеження в магазинах, які слідкують за тим, де зупиняється певний покупець, що привертає його увагу тощо. Це відкриває шлях для персоналізованих рекомендацій та надання послуг.

До речі, на початку 2021 року американська ІТ-компанія Intel заявила про випуск власного програмно-апаратного комплексу для розпізнання обличчя під назвою RealSense ID, який можна використовувати в усіх наведених та інших галузях. В її основі лежить камера RealSense, доповнена датчиком глибини та нейромережею. Як повідомляють розробники, біометричні дані користувача шифруються й оброблюються локально, нейронна мережа попереджає спроби оманити за допомогою масок, світлин та відео, плоскі зображення ігноруються за допомогою датчику глибини, а хибні спрацювання можливі тільки в одному випадку з мільйона, що є дуже високим показником.

Найбільшим ринком СРО на сьогоднішній день є Китай. Детальніше китайські системи ідентифікації та їх поширеність будуть розглянуті трохи згодом, коли мова дійде до муніципальних систем стеження, а поки що декілька тез, що стосуються розповсюдженості та специфічного застосування в Китаї саме застосунків перевірки посвідчення особи «один до одного».

Раніше для придбання SIM-картки громадянам КНР досить було пред'явити копію своєї ідентифікаційної карти. Але на початку грудня 2019 року в країні набрав чинності закон, що зобов'язує громадян проходити процедуру сканування обличчя при вкладанні договорів з операторами мобільного зв'язку. В уряді стверджують, що така міра має зробити ідентифікацію людей простіше, а також захистити права й інтереси населення в інтернеті. Сканування особи допоможе уникнути крадіжок особистих даних і перепродажів SIM-карт, додали в міністерстві інформаційних технологій Китаю.

В середині січня 2020 року в шанхайських аптеках з'явилися СРО для боротьби з наркоманами. Люди, які купують в терміналі ліки за рецептами, повинні підтвердити свою особу перед отриманням препарату. Передбачається, що така система не дозволить потенційним правопорушникам отримувати ліки за рецептом, щоб потім вживати в якості наркотиків або використовувати для виготовлення незаконних препаратів. Система сканує як фармацевтів, так і пацієнтів. Крім виявлення осіб, що

зловживають наркотиками, вона повинна запобігати змові медичних працівників із покупцями з метою незаконного продажу ліків, що містять психотропні речовини або транквілізатори.

На вступних іспитах у коледжі по всій країні використовується розпізнавання облич задля гарантії, що екзаменовані є справжніми вступниками.

Після ряду викрадень дітей деякі дитячі садки відкривають двері тільки тим людям, чії особи зареєстровані в системі. В одному з дитячих садків встановили понад 200 камер для забезпечення безпеки.

Компанія Alibaba спільно з виробником готельних інформаційних систем Shiji встановила СРО для реєстрації в більш ніж 50 готелях. Китайські туристи, які вдаються до послуг онлайн-турагентства Fliggy, можуть спочатку забронювати в ньому готель, а потім, використовуючи «маску» свого обличчя швидко заселитися до готелю й оформити депозит.

### 1.1.3 Використання технології розпізнавання облич у мобільній індустрії, персональних комп'ютерах та вебзастосунках

Окрім усього зазначеного раніше, дуже цікавою є тенденція останніх років, а саме вихід застосунків для ідентифікації по обличчю за межі використання лише в бізнесі та муніципальних структурах та задіяння їх у сферах, що значно ближчі до звичайних користувачів, – мобільній індустрії та інтернеті.

Ще в 2011 році відома американська компанія Google додала в свою нову мобільну операційну систему Android 4.0 можливість автентифікації за допомогою обличчя користувача. Проте алгоритм розпізнання був дуже примітивний, тому точність була низька, а розблокувати телефон було можливо навіть за допомогою звичайної світлини власника. В наступній версії

системи, Android 4.1, було додано перевірку життєздатності, що перевіряє кліпання очима, яку втім також було легко обійти.

У 2010 році корпорація Microsoft випустила ігрову консоль Xbox 360, а разом із нею додатковий аксесуар під назвою Microsoft Kinect, який був створений для відстеження руху й положення користувача та здатний розпізнавати його обличчя [8]. Kinect використовує RGB-D камеру (червоний, зелений, блакитний, глибина), яка складається зі звичайної камери, інфрачервоної камери та ІЧ-проектора. Оскільки цей прилад використовує технологію тривимірної ідентифікації, то він мав перевагу в порівнянні зі застосунками, які використовують звичайні камери, але технологія, за якою він працює, не дозволяла проводити автентифікацію в приміщенні з поганим освітленням, а також у випадку значних змін у зовнішності користувача.

За декілька років Microsoft також додала можливість автентифікації по обличчю в операційну систему Windows 10 за допомогою вбудованої системи Windows Hello, для користування якою потрібна камера з інфрачервоними датчиками. Як заявляє розробник, усі біометричні дані користувача зберігаються й оброблюються локально.

Наприкінці 2017 року американський ІТ-гігант Apple презентував свій новий флагманський смартфон iPhone X з революційною технологією розпізнання обличчя Face ID. Ідентифікатор обличчя від Apple вважається найнадійнішим – він використовує інфрачервону камеру, датчик глибини та ІЧ-проектор «Romeo» для відображення 30000 точок на лиці користувача. Після цього інший модуль, під назвою «Juliet», зчитує отриманий відбиток обличчя й порівнює його зі вже встановленим, який зберігається в спеціальній мікросхемі безпеки Secure Enclave. Збережений відбиток обличчя користувача захищається асиметричним шифруванням, тому ніхто ззовні не може отримати доступ до нього. Face ID може розпізнати обличчя навіть в темряві та зі зміною зовнішності господаря (окуляри, вуса й борода тощо), а ймовірність помилки дорівнює один до мільйона. Система не розпізнає обличчя із заплющеними очима задля додаткової безпеки користувачів.

Єдиним значним недоліком Face ID є значне погіршення якості його роботи за наявності маски на обличчі. У квітні 2020 року розробники оголосили, що оновили систему і для розблокування смартфона більше не потрібно знімати маску, але на ділі заява виявилася дещо голослівна, адже система все одно надто погано працювала в таких умовах. На початку 2021 року Apple випустили нове оновлення зі ще одним способом вирішення проблеми: на цей раз користуватися Face ID у масці стало можливо, якщо користувача є смарт-годинник Apple Watch, який необхідно попередньо розблокувати, а потім подивитися на екран смартфона. Це, звичайно, також не можна назвати повноцінним усуненням недоліку.

Окрім розблокування телефону, Face ID найчастіше використовується для проведення транзакцій за допомогою електронного гаманця та для входу до будинків в системах розумних домів. Мобільні застосунки деяких українських банків, наприклад, Privat Bank та Monobank, також підтримують автентифікацію за обличчям.

Система Face ID і дотепер є неповторною в мобільній індустрії, бо програмне забезпечення, що використовує решта виробників, значно простіше; воно створює двовимірну мапу обличчя, іноді в поєднанні зі сканером райдужної оболонки, а отже майже не працює при поганому освітленні, повороті голови, сильному маскуванні та інших чинниках, що ускладнюють роботу. Навіть сьогодні трапляються випадки, коли застосунки ідентифікації у пристроях відомих виробників позитивно спрацьовують, наприклад, у відповідь на показ відео з господарем. Проте, треба визнати, що подібні застосунки демонструють щораз кращі показники якості роботи в нових моделях смартфонів.

У 2015 році в найбільшій у світі соціальній мережі Фейсбук з'явилась функція визначення особистостей людей на завантажених фотографіях. Ця можливість стала доступна завдяки мережі глибокого розпізнання облич DeerFace [9], яку розробили дослідники з Facebook. Ця мережа має дев'ять шарів та понад 120 мільйонів з'єднань, для її навчання було використано

декілька мільйонів фотографій користувачів соцмережі – це все забезпечило людську точність розпізнання (близько 97,5% точности).

Відомо також про розробку Facebook застосунку, який дозволяв дізнаватися особи людей та їхні облікові записи в соцмережі шляхом направлення камери телефону на них. Розробка дійшла до етапу тестування (використання застосунку співпрацівниками Facebook), але в решті-решт була припинена через неоголошені причини.

#### 1.1.4 Використання технології розпізнавання облич у муніципальних та приватних системах стеження

Другою основною цариною використання застосунків розпізнавання обличчя, найбільш масовою, є системи пошуку та стеження, тобто визначення людей в натовпі. Такі системи майже завжди мають складнішу технологію розпізнавання, ніж системи перевірки посвідчення особи, бо вони проводять сканування в зовсім неконтрольованих умовах, на відстані, одночасно для всіх людей, що потрапили в кадр відеозйомки, – ці фактори значно ускладнюють задачу програми. Проте, саме в цій галузі СРО наразі практично незамінні в порівнянні з рештою біометричних засобів, бо лише вони забезпечують таку можливість.

Впровадження розпізнавання облич у системах стеження розпочалося ще з початку століття й на сьогодні існує 109 країн, які або використовують, або затвердили використання СРО для цілей спостереження.

Найбільша мережа камер відеоспостереження із розпізнаванням обличчя діє в Китаї, який також є найбільшим виробником обладнання для ідентифікації. За прогнозами, до 2023 року Китай становитиме 45% світового ринку розпізнавання облич. У 2006 році китайський уряд ініціював проєкт Skynet – впровадження мережі відеоспостереження по всій країні, і станом на 2013 рік було розміщено 20 млн камер із підтримкою розпізнавання облич,

багато з яких здатні проводити ідентифікацію наживо. У 2019 році на материковій частині Китаю було встановлено вже близько 200 млн камер мережі Skynet (не всі обладнані СРО) і це число постійно зростає (рис. 1.2). Місцеві ЗМІ стверджують, що мережа стеження Китаю може просканувати все населення країни за одну секунду, а населення світу за дві секунди, хоча офіційних підсумків досліджень немає.



Рисунок 1.2 – Вуличні камери для розпізнавання обличчя в Китаї

Система здебільшого показує хороші результати розпізнавання із нахилами голови, різними умовами освітлення та деяким маскуванню. Після початку пандемії COVID-19 у 2020 році виникла потреба в поліпшенні якості розпізнавання облич в масках, що підштовхнуло китайських виробників, таких як SenseTime, FaceGo, Minivision, модернізувати програмне забезпечення своїх систем і тепер вони значно краще працюють із напівзакритими обличчями (точність в ідеальних умовах сягає 90%). Компанія SenseTime виступила із заявою, що в процесі розпізнавання їх обладнання зчитує близько 240 точок на обличчі, що дозволяє зменшити похибку до мінімуму, навіть якщо на людині вдягнена маска, шапка та темні окуляри.

Системи стеження в Китаї окрім пошуку людей також автоматизують роботу поліції. Наприклад, після фіксування порушень правил дорожнього

руху (перехід дороги в неналежному місці або на червоний колір світлофора) порушникам, яких було ідентифіковано, автоматично надсилаються штрафи.

На початку 2020 року Китай першим у світі почав використовувати СРО для виявлення осіб, які порушують карантинний режим ізоляції.

З 2018 року китайських правоохоронців обладнують розумними окулярами із вбудованою СРО, яка під'єднана до загальної БД поліції, і дозволяє фіксувати підозрюваних у натовпі наживо (рис. 1.3). Генеральний директор компанії LLVision Technology, яка розробляє окуляри, повідомив, що під час тестування система змогла ідентифікувати обличчя з бази даних на 10000 осіб за 100 мілісекунд, хоча деякі джерела повідомляють, що окуляри працюють тільки якщо ціль стоїть на місці протягом декількох секунд. Потім цю технологію також почали використовувати в деяких інших країнах, зокрема в Бразилії.



Рисунок 1.3 – Поліцейські розумні окуляри в Китаї

Муніципальні системи стеження використовуються по всій Європі, хоча розгортання відбувається дуже повільно й має здебільшого точковий характер. У Британії випробовування систем розпізнавання в реальному часі почалося ще з 2015 року на публічних заходах, але перший арешт злочинця було

здійснено тільки в 2017 році. Це відбулося у Південному Уельсі за допомогою камери, що була встановлена на фургоні поліції.

Спроби пробних упроваджень в перші роки показували невтішні результати. Це доводять декілька досліджень, зокрема, дослідження, здійснене на замовлення Скотленд-Ярду Ессекським університетом й оприлюднене в червні 2019 року, показало, що 81% людей, визначених системою як злочинці, насправді невинні.

В середині липня 2019 року стало відомо про запуск у Британії систем розпізнавання облич для пошуку зниклих людей. Для тестування технології Міністерство внутрішніх справ звернулося до поліції декількох графств, а також до дорожньої поліції. В Міністерстві повідомили, що запроваджену систему будуть використовувати не в режимі реального часу, щоб уникнути хибних дії поліції через помилкові спрацювання, а задля аналізу зображень з камер відеоспостереження.

Починаючи з 2019 року СРО використовують й у британських в'язницях для попередження передавання в'язням заборонених предметів. Встановлені системи сканують відвідувачів та визначають, хто використовує підроблені посвідчення особистості або відвідує різних в'язнів в різний час. Завдяки зібраним даним співпрацівники в'язниць можуть виявити підозрілих осіб та після додаткового розслідування застосувати до них відповідні міри.

Лондонська поліція перейшла від випробувань до масового розгортання та повноцінного використання систем пошуку злочинців у січні 2020 року й незабаром повідомила про перші реальні арешти.

У Нідерландах використання СРО та технологій штучного інтелекту розпочалося у 2016 році, а станом на 2018 рік у базах даних місцевої поліції знаходилися фото більш як 8% населення країни. Німецька поліція в даний час використовує технологію та планує встановити камери розпізнавання обличчя на 134 залізничних станціях і 14 аеропортах. У Чеській Республіці камери розпізнавання обличчя призвели до щонайменше 160 арештів в аеропорту Праги з середини 2018 року.

У 2019 році в Києві в рамках проєкту «Безпечне місто» також почала діяти муніципальна система відеоспостереження з функцією розпізнавання облич від китайського виробника Hikvision. Особливістю аналітичного модуля, що використовується в системі, є те, що він може використовувати для аналізу дані з будь-яких камер спостереження, навіть тих, що початково не були пристосовані до розпізнавання облич. Модуль фіксує обличчя на всіх встановлених у системі камерах та порівнює їх із базою правопорушників, створеною правоохоронними органами, й, у випадку виявлення збігу, система миттєво подає сигнал оператору. Як повідомили в Департаменті інформаційно-комунікаційних технологій, одночасно система може обробляти 1100 облич на секунду. Система вже діє у метрополітені, вокзалах, лікарнях, школах.

Застосування технології розпізнавання облич в Південній Америці призвело до сильного ефекту. В Аргентині в 2019 році система розпізнавання обличчя допомогла поліції зробити 590 позитивних ідентифікацій лише за шість тижнів. У тому ж році використання під час карнавалу лише 16 спеціалізованих камер призвело до 134 арештів. Розпізнавання обличчя також допомогло бразильській поліції заарештувати другу за розшуком в Південній Америці особу Інтерполу. Наразі технологія розпізнавання обличчя використовується у 92% південноамериканських країн, що є найбільшим показником з усіх континентів.

У США використовується одна з найбільших баз даних людей у світі. Станом на 2016 рік туди входило більше 50% населення країни. Фото були зібрані з різних джерел, зокрема, з паспортів та водійських посвідчень. Системи ідентифікації особистості за обличчям використовуються американською поліцією та ФБР вже багато років, але кількість встановлених камер у США в декілька разів менша, ніж у Китаї, як і ефективність і точність застосунків, що використовуються. Тому системи розпізнавання у США, як і в Великобританії, здебільшого не використовуються наживо, а лише для

пошуку людей на записах з камер або для підтвердження особи підозрюваного, якого затримали.

Окрім того, багато досліджень відзначають такий недолік сучасних СРО, як відмінна точність ідентифікації людей різних рас [6]. Для більшості практикованих СРО в ідеальних умовах кількість помилково позитивних або негативних результатів, наприклад, для афроамериканців зростає до 20%-30% у порівнянні з 1%-2% для представників європейської раси. У США це вже багато років призводить до хибних арештів темношкірих громадян. Корінь проблеми полягає, в першу чергу, у наборах даних, які використовують для навчання нейронних мереж, що застосовуються в СРО. Більшість навчальних даних не є достатньо репрезентативними з боку расової різноманітності і практично завжди в них переважають фото людей певної раси (у американських та європейських системах – європейської, а в програмному забезпеченні, яке розробляють у країнах Азії, – азіатської).

Не тільки правоохоронні органи та влада використовують системи стеження з функцією розпізнавання обличчя. Дуже часто в крамницях та супермаркетах трапляються крадіжки товарів. СРО фіксують обличчя злодія та заносять його до чорного списку, тому наступного разу він вже не зможе навіть зайти до магазину (або до решти філіалів, якщо мова йде про мережу), бо система безпеки одразу ж автоматично попередить охорону про потенційну загрозу. Таку технологію, зокрема, вже впровадила мережа супермаркетів Walmart.

Також багато закладів та установ використовують СРО зі штучним інтелектом, щоб аналізувати їх авдиторію: хто й які товари купує, кількість унікальних та повторних відвідувань, вік та стать, контроль дій персоналу тощо. Завдяки СРО працівники заздалегідь можуть дізнатися інформацію про клієнта, який тільки зайшов до приміщення, привітати його по імені, що позитивно впливає на його враження.

Американська компанія Church's займається виробленням систем розпізнавання обличчя для церков. Як тільки обличчя членів церкви додаються

до бази даних, система автоматично відстежує їх відвідування. Вона також збирає демографічні дані про громаду, включаючи вік та стать. Уряд Японії планує зобов'язати усі казино впровадити технологію розпізнавання обличчя, щоб обмежити відвідуваність залежних від азартних ігор або, принаймні, розмір ставок, що вони роблять.

У підсумку, можна сказати, що СРО набули широкого розповсюдження в світі, дуже швидко розвиваються, алгоритми постійно покращуються та конкурують один з одним, але все ще мають кілька суттєвих недоліків та великий простір для покращення їх ефективності та якості розпізнання.

## 1.2 Аналіз літературних джерел щодо існуючих підходів ідентифікації облич на зображеннях та кадрах відеозйомки

Визначення особистости за допомогою розпізнавання обличчя – вкрай містка тема, яка є об'єктом досліджень вже протягом декількох десятиліть. За цей час було написано безліч наукових робіт, що зачіпають різні особливості розробки та проблеми, так чи так пов'язані з цим напрямом комп'ютерного зору. Практично неможливо досягнути навіть усі найцікавіші та найважливіші праці, присвячені цьому питанню, проте в процесі роботи все ж була опрацьована низка актуальних робіт, перелік яких наведено далі.

У статті [6] розглядаються аспекти розпізнавання облич представників різних рас. Авторами дослідження, описаного в даній роботі, було зібрано й анотовано найбільш репрезентативну й збалансовану з боку расової приналежності загальнодоступну базу даних, яка складається зі 175 000 зображень. Окрім того, порівняно ефективність чотирьох найсучасніших згорткових нейронних мереж щодо проблеми расової та етнічної приналежності. Було досліджено вплив різних рис обличчя та їх характеристик на методи класифікації рас методами глибокого навчання. Були проведені великі тести, щоб визначити риси обличчя, до яких чутливі мережі.

Ще в одній статті [10] досліджується проблема нападів презентації обличчя (НПО). У статті ретельно досліджуються методи виявлення НПО, які вимагають лише звичайних RGB-камер. Представляється типологія сучасних методів НПО, орієнтованих на сценарій атаки, та надається огляд понад 50 методів НПО для обличчя за останні два десятиліття та пов'язані з ними проблеми. Авторами описуються основні виклики, еволюція та сучасні тенденції в області НПО та надається розуміння дальших досліджень.

Джерело [11] – ґрунтовне дослідження, яке розглядає питання одиночного розпізнавання обличчя (ОРП), тобто розпізнавання обличчя людей за наявності лише одного зображення із ними. У цій роботі обговорюється доцільність використання оригінального методу вирішення ОРП, який називається багатоблокними кольоровими бінаризованими статистичними функціями зображення (БКБСФЗ), що використовує декілька видів ознак, а саме місцеві, регіональні, глобальні та текстурно-кольорові характеристики. Тестування методу показують, що БКБСФЗ досягає конкурентних показників у неконтрольованих умовах порівняно з рештою сучасних методів, особливо коли йдеться про зміну виразу обличчя, освітлення та частковій оклюзії. Крім того, запропонований метод використовує алгоритми, засновані лише на простих та елементарних операціях обробки зображень, які не передбачають високих обчислювальних витрат, як у цілісних та розріджених методах або методах глибокого навчання, що робить його ідеальним для ідентифікації в режимі реального часу.

Наступна стаття [12] присвячена розпізнанню облич із частковою оклюзією (зокрема вдягнутої маски). У запропонованій методиці спочатку визначаються обличчя із оклюзією і покращується стабільність шляхом обчислення вектору ознак лише тієї області обличчя, яка є відкритою. Згідно з проведеним тестуванням, авторам вдалося дещо покращити точність розпізнавання звичайних алгоритмів у випадках часткової недоступності інформації.

У роботі [7] надається огляд історії технології розпізнавання облич, сучасні методології та напрямки розвитку. Розповідається про новітні спеціалізовані бази даних, дво- та тривимірні методи розпізнавання облич, в особливості приділяється увага підходу глибокого навчання, як найактуальнішого напрямку в цій галузі.

Дослідження [13] розглядає розпізнавання облич у необмеженому середовищі за часткового повороту голови та пропонує так званий модуль профілю для фронтального відображення (ПФВ). Цей модуль здійснює перегляд довільних поз на рівні ознак та перетворює багатоцільові ознаки в приблизне фронтальне представлення для покращення здатності до розпізнавання в уже наявних моделях. Також ПФВ тестується на наборах зображень облич, отриманих в неконтрольованих умовах, таких як Labeled Faces in the Wild (LFW), Celebrities in Frontal Profile (CFP) та IARPA Janus Benchmark A (IJB-A). Результати показують, що обраний метод забезпечує гарні показники.

Автори статті [14], також зосереджуються на проблемі повороту голови, презентуючи новий метод розпізнавання обличчя, який є нечутливим до пози. Він заснований на перетворенні обличчя з вирівнюванням ключових точок на основі генеративних змагальних мереж (FT-GAN). У цьому методі використовується CycleGAN для перетворення пікселів для попередньої, грубої, трансформації обличчя, а потім ці дані уточнюються за допомогою вирівнювання за ключовими точками. Отож, фронтальний синтез обличчя моделюється як процес із двома проміжними завданнями. Наслідки комплексних експериментів довели ефективність FT-GAN.

У наступній праці [15] дослідники надають огляд найновіших та найефективніших методів видобутку орієнтирів обличчя на двовимірних зображеннях та відео, заснованих на глибокому навчанні. Наведено аналіз багатьох підходів із порівнянням їх ефективності. Також пропонується аналіз розповсюджених наборів даних, проблем та напрямків майбутніх досліджень.

Ще в одному дослідженні [16] пропонується схема розпізнавання обличчя із використанням локальних моментів Церніке (LZM), які можна використовувати як для ідентифікації, так і для перевірки. У цій схемі із складних компонентів, отриманих шляхом перетворення LZM, витягуються області навколо ключових точок. Потім у цих областях будуються гістограми фазового розміру для створення дескрипторів зображення обличчя. Піраміда зображення використовується для отримання об'єктів у багатьох масштабах, а дескриптори будуються для кожного зображення в цій піраміді. Результати випробування на наборах зображень, зроблених в неконтрольованому середовищі, показали, що запропонований метод стійкий до таких змін, як освітлення, вираз обличчя та поворот голови. Окрім цього, його можна використовувати із зображеннями обличчя з низькою роздільною здатністю, отриманих у неконтрольованих середовищах або в інфрачервоному спектрі.

Й остання робота, про яку хотілося б згадати, це [17]. У цьому дослідженні запропоновано новий алгоритм вилучення локальних особливостей для гіперспектральних зображень обличчя за допомогою декомпозиції тензора низького рангу на основі локальних областей, що також зберігає інформацію про сусідські зв'язки та розмір спектру. Глобальні контурні характеристики були вилучені за допомогою дискретного алгоритму швидких перетворень Фур'є (PFFT), який вирішує багато проблем із розпізнаванням людського обличчя, таких як освітлення, вираз, орієнтація та зміна віку. Запропонований метод протестували за допомогою бази даних Poly-U та порівняли з іншими гіперспектральними алгоритмами розпізнавання обличчя. Результати показали, що запропонований алгоритм конкурує з найкращими методами CRC\_RLS та PLS.

### 1.3 Постановка задачі

Актуальність даної роботи полягає у важливості систем розпізнання облич, їх широкому спектрі сфер застосування в сучасному світі, а також існуванні ряду пов'язаних із ними проблем та необхідності пошуку оптимальних методів та підходів для розпізнавання.

Об'єктом роботи є послідовність кадрів відеозйомки із людськими обличчями.

Метою роботи є розробка застосунку для розпізнавання облич на послідовності кадрів відеозйомки.

Враховуючи мету роботи, необхідно вирішити такі завдання:

- дослідити методи визначення ключових точок на зображеннях;
- дослідити методи класифікації зображень на підставі статистичних характеристик;
- проаналізувати наявні двовимірні підходи до розпізнавання облич за множиною ознак;
- розробити методіку ідентифікації облич на кадрах відеозйомки;
- визначити необхідні інструментальні засоби для створення застосунку;
- виконати всі етапи розроблення застосунку для ідентифікації облич на кадрах відеозйомки;
- провести тестування розробленого застосунку та проаналізувати результати;
- визначити перспективи дальшої роботи.

## 2 РОЗРОБЛЕННЯ ПІДХОДУ ДЛЯ ІДЕНТИФІКАЦІЇ ОБЛИЧ НА КАДРАХ ВІДЕОЗЙОМКИ

### 2.1 Методи визначення особливих точок на зображеннях

У галузі комп'ютерного зору під поняттям образу розуміють деяку впорядковану сукупність ознак. Сукупність образів, що мають певні спільні властивості, називають класом. При всьому різноманітті методів й алгоритмів розпізнання зображень, практично усі вони полягають в процесі виокремлення вектору ознак, що характеризують певний образ, із загальної маси несуттєвих даних, та віднесення їх до тих чи тих класів, цебто класифікацію, – автоматично або з мінімальним втручанням людини [18]. Наслідки розпізнання відбивають відношення еквівалентності між ознаками на зображенні та ознаками, що зберігаються в БД.

Як відомо, зображення можна представити у вигляді масиву пікселів. При цьому вплив кожного окремого пікселя зображення на наслідки розпізнання є мізерним. Це робить таке представлення зображення надмірним й неекономним. При вирішенні задач розпізнавання образів необхідно значно менше інформації. Таким чином, для ефективного розпізнання об'єктів потрібен деякий стислий та зручний формат представлення зображення [19].

Поняття ключової, або особливої, точки ґрунтується на інтуїтивній ідеї про те, щоб представити зображення або об'єкт в інваріантному вигляді, який залишався б незмінним або, принаймні, схожим в інших подібних зображеннях тієї ж сцени або об'єкту. Особливі точки дають ефективні засоби вирішення цього завдання. Особлива точка – це невелика ділянка зображення, яка зберігає багато місцевої інформації, а тому з великою імовірністю вона буде розпізнана в іншому зображенні [20].

Опис особливої точки представляють у вигляді дескриптора, розмір якого, як правило, значно менший за розмір області пікселів, з якої було сформовано особливу точку. Дескриптор кодує цю область так, щоб її було

можливо розпізнати в іншому зображенні, щоб вона здебільшого була унікальною та щоб допускала параметризацію [21]. Зазвичай (проте не обов'язково) це вектор дійсних чисел, який якийсь описує особливу точку.

Одна з найважливіших характеристик поняття особливої точки полягає у тому, що вона дозволяє створити короткий опис зображення у вигляді скінченного числа вагомих первнів, навіть якщо роздільна здатність дуже велика. У цьому сенсі особливі точки – спосіб перейти від піксельного представлення високої розмірності до більш компактного представлення, якість якого зростає разом зі зростанням розміру зображення, а фактичний розмір залишається сталим [22]. Вчені гадають, що зорова кора мозку збирає окремі сигнали сітківки в інформаційні блоки більш високого рівня, причому, принаймні, в деяких із них дані аналогічні тим, що зберігаються в особливих точках.

Перш ніж перейти до розгляду наявних методів визначення особливих точок, варто зазначити, що їх наразі існує дуже багато й про жоден з них неможливо сказати, що він безумовно кращий за інші. Деякі алгоритми виграють у швидкості, другі віднаходять ознаки, для яких можна побудувати якнайліпші дескриптори, а треті мають певні властивості інваріантності – наприклад, інваріантність відносно обертання, масштабування, невеликих афінних перетворень. Тому в роботі представлено лише декілька найпоширеніших та найзначніших способів виявлення ознак на зображеннях.

### 2.1.1 Кути Гарріса

Визначено, що найважливішими ділянками зображення є ті, в яких спостерігається велика похідна. Проте на практиці цього виявляється недостатньо. Точка, в якій велика похідна, може знаходитися на границі чогось, але при цьому вона також схожа на решту точок, що знаходяться на тій же границі [23].

Але у випадках, коли спостерігаються значні похідні в двох різних напрямках (вздовж різних осей), така точка має значно більше шансів виявитися унікальною та зберігає достатньо інформації для виокремлення в кожному кадрі. Тому багато-які ознаки, що відстежують, називають кутами.

Найуживаніше визначення кута дав Гарріс у своїй праці [24], хоча деякі дослідники пропонували подібні визначення й раніше. Кути Гарріса можна вважати прототипом усіх особливих точок. Їх визначення ґрунтується на понятті автокореляції між пікселями у малому околі.

На рисунку 2.1 наведено приклад кутів Гарріса, які знайдено на світлині з книжковими сторінками (кількість ознак було програмно обмежено). Можна помітити, що знайдені ознаки знаходяться на зарубках або на перетині складових частин літер, а не на границях прямих ліній (як в літерах h або k).

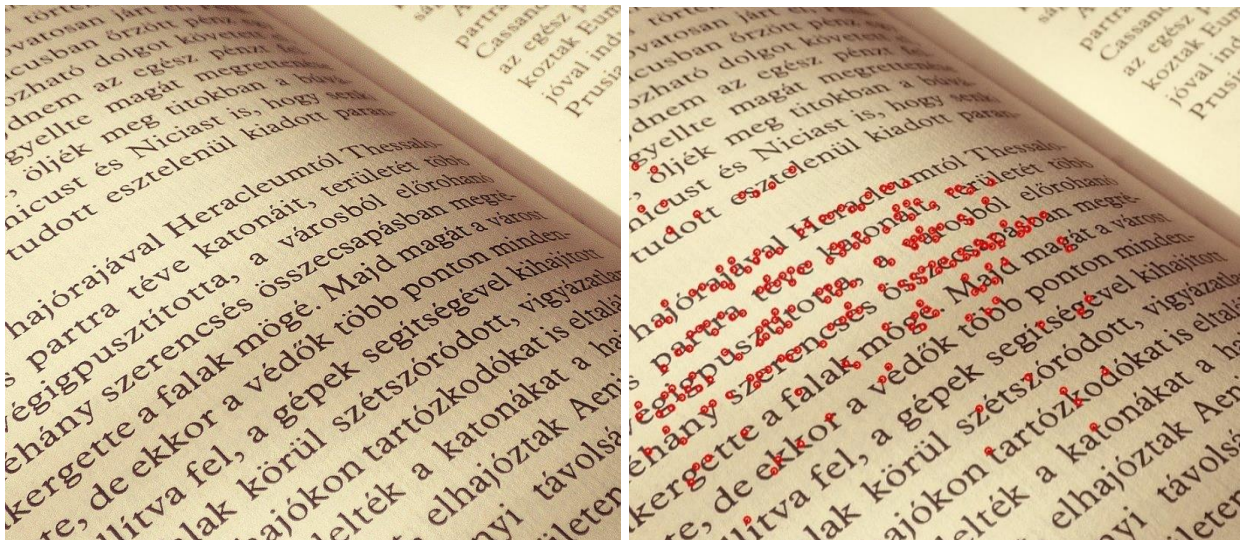


Рисунок 2.1 – Знайдені кути Гарріса

Гарріс розпочав з розгляду наступної функції автокореляції яскравості пікселів  $I(x, y)$ :

$$c(x, y, \Delta x, \Delta y) = \sum_{(i,j) \in W(x,y)} \omega_{i,j} (I(i, j) - I(i + \Delta x, j + \Delta y))^2. \quad (2.1)$$

Це зважена сума квадратів різниць між яскравістю в точці  $(i, j)$  та яскравістю в точці, що зсунута відносно неї на  $(\Delta x, \Delta y)$ , по невеликому вікню навколо точки  $(x, y)$ . Вагова функція  $\omega_{i,j}$  має нормальний розподіл.

Оскільки  $\Delta x$  та  $\Delta y$  припускаються малими, то член  $I(i+\Delta x, j+\Delta y)$  можна апроксимувати виразом  $I(i, j) + I_x(i, j)\Delta x + I_y(i, j)\Delta y$  ( $I_x$  та  $I_y$  – перші часткові похідні  $I(x, y)$  по  $x$  та по  $y$  відповідно). Нове рівняння можна записати у вигляді:

$$c(x, y, \Delta x, \Delta y) = [\Delta x \ \Delta y] M(x, y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}, \quad (2.2)$$

де  $M(x, y)$  – симетрична автокореляційна матриця:

$$M(x, y) = \omega_{i,j} \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}. \quad (2.3)$$

Згідно з визначенням Гарріса, кути – це такі місця на зображенні, де автокореляційна матриця має два великих власних значення. Це означає, що при зміщенні на малу відстань в будь-якому напрямку зображення помітно змінюється. В такого погляду є перевага: власні значення автокореляційної матриці інваріантні відносно обертання, що почасти є вельми корисною властивістю. Окрім того, два власних значення кута Гарріса не просто визначають, чи можна вважати точку хорошою ознакою (тобто особливою точкою), а ще й є ідентифікаційною сигнатурою, цебто дескриптором. Тісний зв'язок з дескрипторами – це типова, хоч і не універсальна властивість особливих точок.

У початковій праці Гарріса вираховувалася різниця між визначником  $M(x, y)$  і квадратом сліду матриці (з деяким ваговим коефіцієнтом):

$$H = \det(M) - \kappa \text{trace}^2(M) = \lambda_1 \lambda_2 - \kappa (\lambda_1 + \lambda_2)^2. \quad (2.4)$$

Для знаходження кутів виконується пошук локальних максимумів цієї функції. Функція  $H$ , що її називають мірою Гарріса, дозволяє використовувати в порівнянні власні значення  $M$  ( $\lambda_1$  та  $\lambda_2$ ), не потребуючи їх видимого обчислення. У порівнянні бере участь параметр  $k$ , що його називають чутливістю, якому можна присвоїти будь-яке значення від 0 до 0,24, але зазвичай беруть значення, близьке до 0,04 (бо за такого значення  $k$  алгоритм оптимально чутливий).

Пізніше Ши та Томасі [25] доповнили метод Гарріса, довівши, що хороші кути виходять за умови, що менше з двох власних значень більше деякого мінімального порогу, який встановлюється в кожному випадку окремо. Метод Ши-Томасі (або Канаде-Томасі) виявився не тільки достатнім, але й у багатьох випадках давав ліпші результати, ніж метод Гарріса, тому зараз найчастіше використовують саме такий варіант детектора кутів (рис. 2.2).



Рисунок 2.2 – Кути Гарріса та Ши-Томасі

### 2.1.2 Детектор ознак FAST

Метод знаходження ознак FAST (Features from Accelerated Segments Test), який запропонували Ростен та Драммонд [26], ґрунтується на прямому порівнянні точки  $P$  із множиною точок всередині невеликого описаного навколо неї кола. Ідея така, що як лише деякі близькі точки схожі на  $P$ , то  $P$  цілком може виявитися хорошою особливою точкою. В ранішній реалізації цієї ідеї, алгоритмі SUSAN (Smallest Univalued Segment Assimilation Nucleus), порівнювалися всі точки в колі з центром  $P$ . Алгоритм FAST, який можна вважати наступником SUSAN, вніс два вдосконалення.

Перша відмінність полягає в тому, що FAST використовує лише точки на обводі з центром  $P$  (обвід визначається за алгоритмом Брезенгема). А друга – що окремі точки обводу розподіляються по трьох класах: темніше  $P$ , світліше  $P$  та схожі на  $P$ . Класифікація заснована на пороговій величині  $t$ : темними вважаються пікселі, яскравість котрих менша за  $I_p - t$ , світлими – пікселі з яскравістю більше за  $I_p + t$ , а схожими – пікселі з яскравістю від  $I_p - t$  до  $I_p + t$ . Після закінчення класифікації детектор FAST вимагає, щоб на обводі була визначена кількість сусідніх точок, які або усі яскравіші, або усі темніші за  $P$ . Якщо загальна кількість точок на обводі дорівнює  $N$ , то кількість тих, що є більш темними або більш світлими, має бути на менше за  $N/2 + 1$  пікселів (тобто більше половини).

В такий спосіб уже можливо досягти високої швидкості роботи визначення особливих точок, але в алгоритмі використовується ще одна оптимізація: перед початком класифікації перевіряються тільки чотири рівновіддалені від  $P$  точки. Було визначено, що коли не існує хоча б двох із них, котрі яскравіші або темніші за  $P$ , то точка  $P$  не може бути FAST-ознакою. На практиці ця оптимізація круто скорочує час пошуку в усьому зображенні. На рисунку 2.3 наведено чергову ітерацію алгоритму. Числами позначено пікселі, що враховуються при класифікації. Пікселі з номерами 1, 5, 9 та 13 перевіряються в першу чергу для попередньої перевірки точки  $P$ .

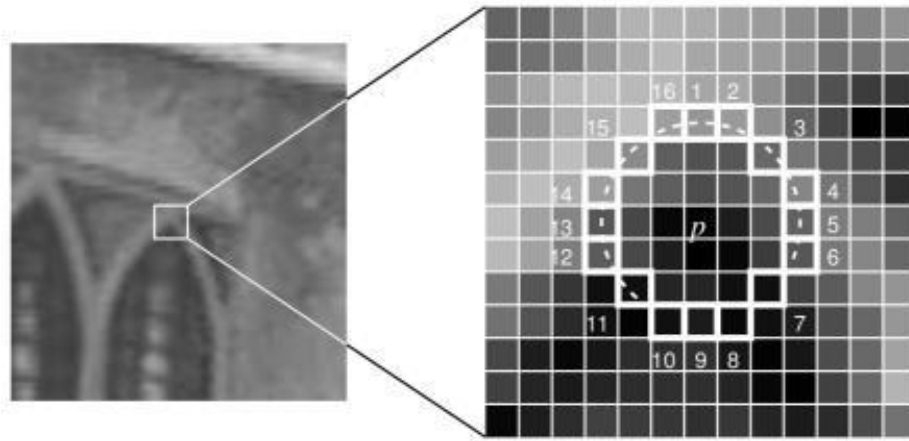


Рисунок 2.3 – Приклад чергової ітерації алгоритму FAST

Єдиним недоліком цього методу є схильність уважати кутами декілька сусідніх пікселів. Для вирішення цієї проблеми алгоритм FAST назначає кожному куту оцінку й відкидає всі особливі точки, поряд з якими знаходиться особлива точка з вищою оцінкою. Для розрахунку оцінки спочатку вираховується сума абсолютних різниць між світлішими та центральним пікселем, потім те ж саме виконується для темніших пікселів і береться максимум:

$$score = \max \left( \sum_{x \in \{brighter\}} |I_x - I_p| - t, \sum_{x \in \{darker\}} |I_x - I_p| - t \right). \quad (2.5)$$

### 2.1.3 Детектор ознак SIFT

Алгоритм SIFT (Scale Invariant Feature Transform), створений Девідом Лоуї [27], сьогодні широко використовується й лежить в основі багатьох розроблених згодом ознак. Вирахування SIFT-ознак є досить коштовним з боку обчислювальних ресурсів, у порівнянні з багатьма іншими, але вони у вищому ступені виразні, а тому добре підходять в задачах супроводу і розпізнання [28, 29] (рис. 2.4).



Рисунок 2.4 – Наслідок роботи детектора SIFT

Властивість інваріантності відносно масштабу, через яку метод SIFT отримав свою назву, здобувається завдяки початковому етапу алгоритму, на якому вираховується набір згорток вихідного зображення з ядрами Гауса щораз більших розмірів. Такий набір згорток називається масштабованим простором зображення (рис. 2.5).

Після цього будується піраміда гаусіанів: увесь масштабований простір розбивається на ділянки, октави. При переході від однієї октави до іншої розмір зображення зменшується вдвічі.

Паралельно з побудовою піраміди гаусіанів, будується піраміда різниць гаусіанів (Difference of Gaussian, DoG), що складається з різниці сусідніх зображень у піраміді гаусіанів (рис. 2.6).

За наявності такого набору зображень, які можна представити у вигляді стосу, кожен піксель кожного зображення порівнюється не тільки з сусідніми пікселями у власному зображенні (котрих усього вісім), але також із собою та сусідами в наступному й попередньому зображеннях у стосі (тобто ще 18 пікселів). Якщо значення пікселя в РГ більше (або менше), ніж в усіх 26

пікселів, з якими його було порівняно, то його називають екстремумом у масштабованому просторі оператору РГ.

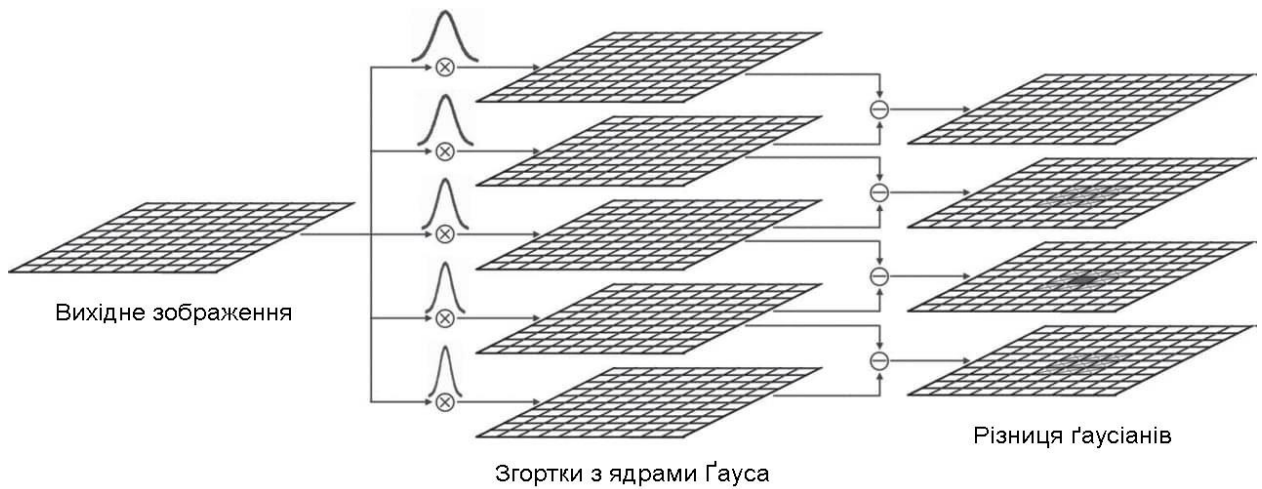


Рисунок 2.5 – Схема побудови піраміди РГ

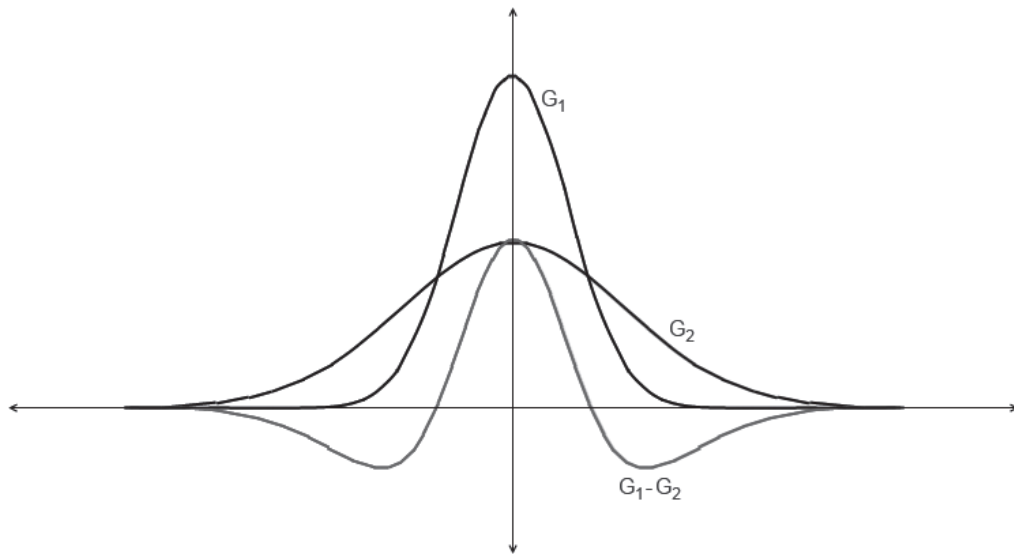


Рисунок 2.6 – Ядра Гауса  $G_1$  і  $G_2$ , та їх різниця  $G_1 - G_2$

Знайшовши набір ознак, алгоритм перевіряє якість кожної ознаки й уточнює оцінку його положення.

По-перше, координати особливої точки уточнюються на субпікельному рівні. Цього досягають завдяки апроксимації функції РГ багаточленом Тейлора другого порядку, який узяли в точці екстремуму:

$$D(x) = D + \frac{dD^T}{dx}x + \frac{1}{2}x^T \frac{d^2D}{dx^2}x, \quad (2.6)$$

де  $D$  – функція РГ;

$x$  – вектор зміщення відносно точки розкладення;

перша похідна РГ – градієнт;

друга похідна РГ – матриця Гессе.

Екстремум багаточлена Тейлора знаходиться шляхом вирахування похідної та прирівнювання її до нуля. У наслідку отримується зміщення точки екстремуму, який початково було вираховано, до його точно положення:

$$\vec{x} = -\frac{d^2D^{-1}}{dx^2} * \frac{dD}{dx}. \quad (2.7)$$

Якщо одна з компонент вектору  $\vec{x}$  більша за половину кроку сітки в цьому напрямку, то це означає, що точка екстремуму була порахована невірно та потрібно посунути до сусідньої точки в напрямку вказаних компонент. Для сусідньої точки все повторюється наново. Якщо таким робом алгоритм вийшов за межі октави, то необхідно виключити дану точку з розгляду.

Коли положення точки екстремуму пораховано, перевіряється на малість саме значення РГ у цій точці по формулі:

$$D(\vec{x}) = D + \frac{1}{2} \frac{dD^T}{dx} \vec{x}. \quad (2.8)$$

Якщо ця перевірка не проходить, то точка виключається, як точка з малим контрастом.

Друга перевірка націлена на виключення особливих точок, що лежать на границі та не є кутами. Ці точки мають великий вигин вздовж границі й малий у перпендикулярному напрямку. Цей великий вигин визначається матрицею Гессе розміром  $2 \times 2$ :

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}. \quad (2.9)$$

Нехай  $Tr(H)$  – слід матриці,  $Det(H)$  – її визначник,  $\alpha$  – великий вигин,  $\beta$  – малий, тоді:

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta, \quad (2.10)$$

$$Det(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta. \quad (2.11)$$

Якщо  $r$  – відношення  $\alpha$  до  $\beta$ , то  $\alpha = r\beta$ , і тоді

$$\frac{Tr(H)^2}{Det(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r\beta + \beta)^2}{r\beta^2} = \frac{(r + 1)^2}{r}. \quad (2.12)$$

Точка розглядається в тому разі, якщо

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r + 1)^2}{r}. \quad (2.13)$$

Після того, як усі екстремуми в масштабованому просторі знайдені, алгоритм SIFT переходить до побудови дескрипторів для кожного об'єкта. Перший крок – визначення орієнтації особливої точки, для чого порівнюються похідні за напрямками в точках з околу особливої точки та обирається напрямок, що відповідає максимальній похідній. Коли головний напрямок визначено, решту властивостей дескриптора вираховуються відносно до нього. Отже, SIFT-ознаки інваріантні не тільки відносно масштабу, а й відносно напрямку.

Нарешті, коли відомі масштаб та орієнтація, вираховується місцевий дескриптор зображення. Він також утворюється з локальних градієнтів, але на цей раз після повороту локального околу на кут напрямку дескриптору. Потім

пікселі об'єднуються в регіони (зазвичай 16 пікселів у квадраті 4×4 навколо особливої точки) і для кожного регіону будується гістограма напрямків у всіх його точках. Зазвичай гістограма складається з восьми інтервалів. Об'єднання восьми інтервалів на регіон з шістнадцятьма регіонами дає 128 компонентів. Цей 128-компонентний вектор і є дескриптором особливої точки в алгоритмі SIFT. Така велика кількість компонентів – основа високої дескриптоспроможности SIFT-ознак.

#### 2.1.4 Детектор ознак SURF

Ознаки SURF (Speeded-Up Robust Features – прискорені стійкі ознаки) [30] є багато в чому розвитком ознак SIFT. Авторів алгоритму SURF цікавило, чи можливо замінити окремі частини алгоритму SIFT більш ефективними з обчислювальної точки зору, при цьому отримавши порівняну або навіть кращу якість вирішення задач розпізнання. Отриманий метод дійсно знаходить ознаки значно швидше, а завдяки деякому спрощенню, в порівнянні зі SIFT, ці ознаки в багатьох випадках виявляються більш стійкими до зміни орієнтації й освітлення.

На декількох етапах віднаходження ознак SURF використовує інтегральне представлення зображення. Це одноразове перетворення всього зображення, яке дозволяє обчислювати суми яскравості в будь-якій прямокутній області, виконав декілька простих арифметичних операцій. Як і багато-які інші детектори, SURF визначає особливу точку в термінах локального гесіану, який в даному випадку має вигляд:

$$\frac{d^2}{dx_i dx_j} G(\vec{x}, \sigma), \quad (2.14)$$

де  $G(\vec{x}, \sigma)$  – нормоване ядро Гауса розміру  $\sigma$ , з яким зображення згортається перед обчисленням похідних.

На відміну від методу SIFT, де використовувалася згортка з ядром Гауса, детектор SURF для обчислення локального гесіану застосовує згортку з прямокутним фільтром, який апроксимує різницю двох ядер Гауса. Головна перевага прямокутних фільтрів – їх можна швидко обчислити, використовуючи інтегральне зображення. На рисунку 2.7 зображено порівняння дискретного ядра ПГ  $9 \times 9$ , що апроксимує другу похідну по вертикалі, та прямокутного фільтру ПГ.

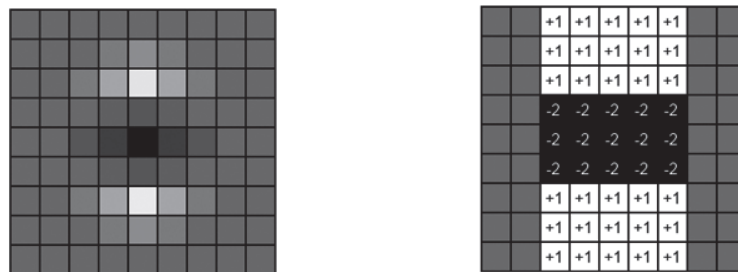


Рисунок 2.7 – Порівнювані фільтри

Оскільки вартість апроксимації прямокутним фільтром не залежить від розміру фільтра (завдяки інтегральному зображенню), немає потреби будувати піраміду масштабованих зображень, як у випадку із SIFT. Замість цього для обчислення гесіану при різних масштабах можна використовувати прямокутних фільтрів дедалі більшого розміру. У SURF особливі точки визначаються як місцеві екстремуми визначника гесіану, що перевищує заданий поріг.

У SURF, як і в SIFT, визначено напрямки ознаки, для вирахування якого також використовується інтегральне зображення, яке дозволяє оцінити максимальний місцевий градієнт в околі ознаки. Для апроксимації місцевих градієнтів використовуються два простих вейвлети Гаара, які застосовуються до різних частин області, в якій був знайдений екстремум в масштабованому просторі. Якщо масштаб ознаки дорівнює  $s$ , то градієнти вираховуються за

допомогою вейвлетів розміром  $4s$  із кроком  $s$  у колі радіусу  $bs$  з центром в особливій точці. Потім оцінки градієнтів поєднуються по ковзному кутовому вікну розміру  $\pi/3$ .

Підсумовуючи усі градієнти в цьому вікні, обирається напрямок максимального результату та вважається, що це і є напрямок ознаки. Знаючи напрямок, стає можливо обчислити векторний дескриптор відносно цього напрямку, що забезпечує інваріантність відносно обертання.

Самі дескриптори обраховуються приблизно так, як і в SIFT. Для особливої точки в масштабі  $s$  прямокутник  $20s \times 20s$  з центром у цій точці спочатку розбиваються на 16 квадратів, що утворюють сітку  $4 \times 4$ . Ця сітка повертається на кут, що дорівнює обчисленому напрямку ознаки. Для кожної комірки цієї сітки застосовуються 25 пар вейвлетів Гаара для апроксимації градієнтів зображення по  $x$  та по  $y$  в кожній з комірок сітки  $5 \times 5$  (в середині кожної комірки сітки  $4 \times 4$ ) (в цьому випадку  $x$  та  $y$  – координати в поверненій системі координат дескриптора, а не в системі координат зображення). Вираховуються суми й суми абсолютних значень 25 згорток з вейвлетами в напрямку  $x$  та 25 згорток у напрямку  $y$ . На виході для кожної комірки сітки  $4 \times 4$  отримуються 4 числа (усього 64 числа). Ці числа і є елементами 64-вимірному вектору дескриптора особливої точки SURF (рис. 2.8).

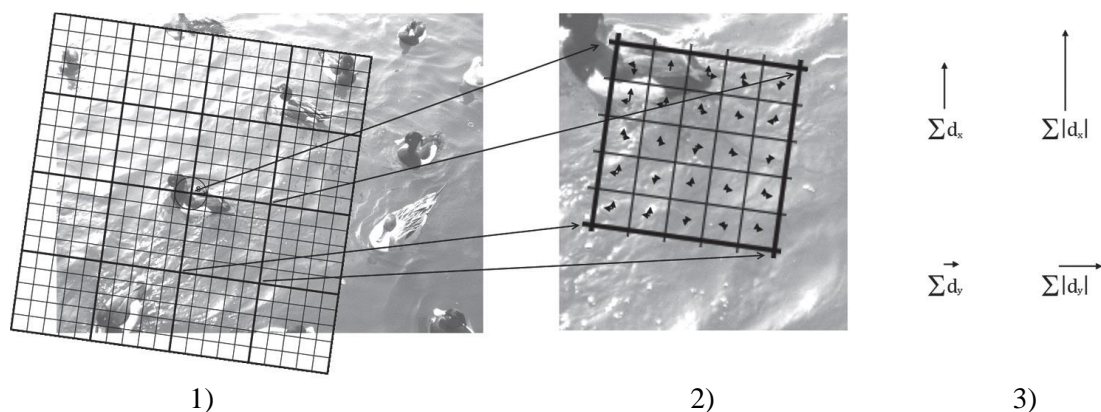


Рисунок 2.8 – Етапи обчислення дескриптора SURF: 1) накладання сітки  $4 \times 4$  на ознаку; 2) розбиття кожної комірки на 25 підкомірок, для яких вираховуються похідні по напрямках; 3) отримані похідні підсумовуються, що дає чотири значення для кожної комірки більшої сітки

Також існують розширені SURF-ознаки, для яких обчислюються не чотири, а вісім згорток. Отримані дескриптори виходять більші за розміром, тому для їх обчислення потрібно більше часу, проте подекуди це поліпшує якість розпізнання.

### 2.1.5 Детектор ознак та дескриптор BRISK

Дескриптор ознак BRISK, який вперше представлено в праці [31], – спроба покращити дескриптор BRIEF, який широко застосовується. Ключові переваги BRISK в порівнянні з BRIEF у тому, що він має власний детектор особливих точок (рис. 2.9) – це перше, ознаки, що описує дескриптор, є більш стійкими – це друге.



Рисунок 2.9 – Наслідок роботи алгоритму BRISK

Детектор ознак у BRISK ґрунтується на схожому на FAST детекторі AGAST, але додатково намагається визначити масштаб та орієнтацію ознаки. Для визначення масштабу BRISK спочатку створює піраміду в масштабованому просторі з фіксованим числом масштабів (сусідні масштаби

відрізняються вдвічі), а потім вираховує фіксовану кількість інтраоктав на кожний масштаб. Перший крок детектора ознак у BRISK – застосування алгоритму FAST для пошуку ознак в усіх цих масштабах. Після чого видаляються немаксимальні ознаки, тобто ті, для котрих оцінка не є найбільшою серед усіх сусідів.

Побудував список ознак, BRISK переходить до обчислення оцінки FAST у відповідних точках зображень попереднього та наступного масштабів. Оцінки AGAST апроксимуються параболою (функцією масштабу) і максимум цієї параболи приймається за істинний масштаб ознаки BRISK. Отже, масштаб ознаки BRISK необов'язково асоційований з одним з дискретних зображень в піраміді. Аналогічна інтерполяція застосовується до координатів пікселів, тож ознаці зіставляються субпіксельні координати. Схематичне зображення алгоритму представлено на рисунку 2.10.

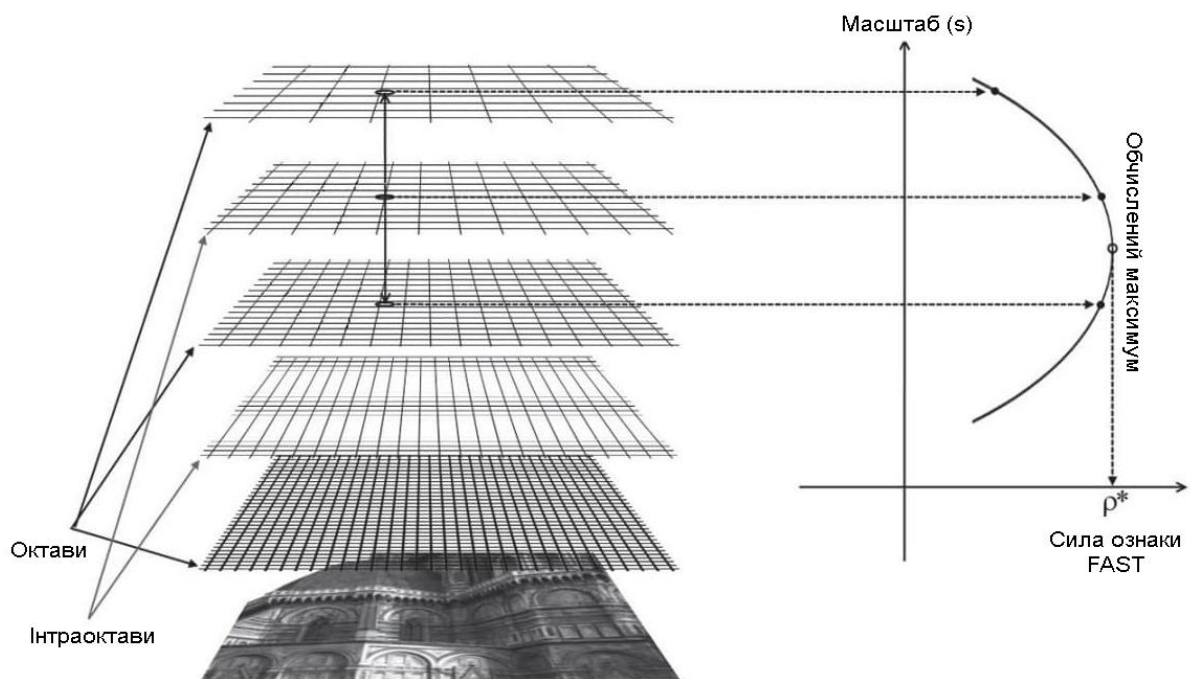


Рисунок 2.10 – Схема пошуку ознак алгоритмом BRISK

Окрім масштабу, в ознак BRISK є ще й орієнтація. Дескриптор BRISK будується у вигляді ряду обводів навколо центральної точки. Кожному обводу назначається  $K_i$  вибірних точок і кожній вибірній точці зіставляється кругова

ділянка, діаметр якої дорівнює довжині околу  $C_i$ , що містить її в собі, поділеної на  $K_i$ . Ця ділянка відповідає вибірці в указаній точці з зображення, згорнутого з ядром Гауса такого ж радіусу ( $\sigma_i = C_i/2K_i$ ).

Порівняння яскравості, результати яких складають бітовий дескриптор, виконуються між парами точок в усіх околах (рис. 2.11). Пари розбиті на дві підмножини: близькі та віддалені. До близьких відносяться пари точок, відстань між якими менше заданого порогу  $d_{max}$ , а до віддалених – більше порогу  $d_{min}$ . Близькі пари створюють дескриптор, а віддалені використовуються для обчислення напрямку, що домінує. Отже, між парами точок вираховуються різниці яскравості, а потім нормуються на відстань між точками, завдяки чому отримується локальний градієнт. Підсумовуючи ці локальні градієнти по всіх віддалених парах вираховується напрямок дескриптора. А ознаки, що визначаються близькими парами, обчислюються відносно цього напрямку, тож дескриптор, який отримано бінарizaцією градієнтів, що вираховані по близьких парах, виявляється незалежним від орієнтації. Шляхом настроювання числа точок на околі й величини  $d_{max}$  можливо отримати дескриптор довільної довжини.

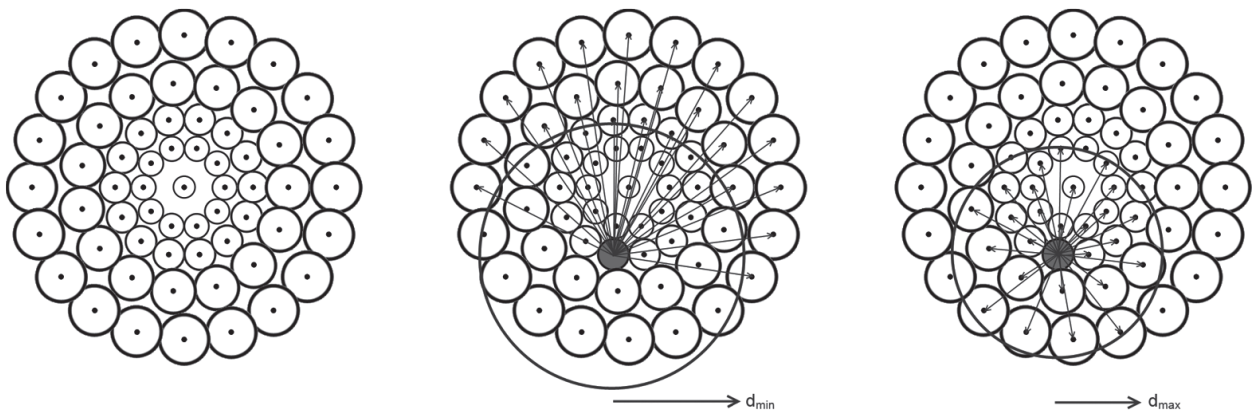


Рисунок 2.11 – Процес порівняння яскравостей між точками

## 2.2 Класифікація зображень на підставі статистичних характеристик

Визначення на певному зображенні ключових точок та розрахунок їх якісних дескрипторів – це лише перший етап вирішення задачі розпізнавання образів, адже далі необхідно правильно обрати ті особливості, що дійсно є цінними у даному випадку, зіставити отримані дескриптори з тими, що зберігаються в БД, і визначити (у випадку наявності необхідних даних), які саме об'єкти знаходяться на зображенні або кадрі відеозйомки. Цей етап, як вже було зазначено раніше, називається класифікацією, а метод, за яким здійснюється класифікація, – класифікатором.

На сьогоднішній день для здійснення класифікації векторів ознак на зображенні майже завжди використовується машинне навчання. Його ідея полягає в тому, щоб шляхом автоматичної обробки великої кількості даних навчитися правильно визначати ознаки та їх характеристики, що асоційовані з окремими образами, а отже, в разі отримання нової множини ознак визначити, до якого образу вони відносяться [32]. Серед алгоритмів машинного навчання окремо виділяють ті, що класифікують зображення на підставі статистичних характеристик. Подібні класифікатори здебільшого дають хороші результати та дуже широко використовуються, як для супроводу об'єктів на відеоряді, так і для ідентифікації. Про деякі найвідоміші з них піде мова в цьому розділі.

### 2.2.1 Баєсів підхід до прийняття рішень

Баєсів класифікатор на основі отриманих ознак відносить об'єкт до класу, до якого цей об'єкт належить з найбільшою імовірністю. Цей класифікатор також називають «найвним» через те, що він припускає статистичну незалежність усіх змінних, хоча на практиці так буває мало коли, бо, частіше за все, ознаки взаємопов'язані й у разі появи однієї з них, скоріше за все, очікується поява інших [33]. Проте, як виявляється, наявності кореляції

почасту можна нехтувати, тому баєсів класифікатор дуже часто показує неочікувано хороші результати. Це найпростіший спосіб широкої дисципліни баєсових мереж або ймовірнісних графічних моделей, що продовжують розвиватися.

Основою баєсового підходу є теорема Баєса. Допустімо, що ми спостерігаємо деяку випадкову величину  $x$ , яка має щільність розподілу  $p(x, \omega)$  з параметром  $\omega$ , і завдяки ній хочемо зробити висновок про іншу випадкову величину  $\omega$ , що має деякий розподіл імовірності  $P(\omega)$ . Нехай у наслідку спостережень отримані статистичні дані про  $x$ . З визначення умовної імовірності витікає:

$$P(\omega|x) = \frac{p(x, \omega)}{P(x)}, \quad (2.15)$$

а також:

$$p(x, \omega) = p(x|\omega) P(\omega). \quad (2.16)$$

Підставивши (2.16) у (2.15) отримаємо формулу Баєса:

$$P(\omega|x) = \frac{p(x|\omega) P(\omega)}{p(x)}. \quad (2.17)$$

Для визначеної задачі є  $s$  класів, тобто  $s$  можливих змінних ( $\omega_1, \dots, \omega_s$ ), тоді:

$$P(\omega_i|x) = \frac{p(x|\omega_i) P(\omega_i)}{\sum_{i=1}^s P(\omega_i) p(x|\omega_i)}. \quad (2.18)$$

Розподіл  $P(\omega)$  називають апіорним розподілом імовірностей можливих значень  $\omega$ . Цей розподіл приймається, заки не отримані статистичні дані.

Розподіл  $P(\omega/x)$  – апостеріорний розподіл значень  $\omega$ , за умови, що статистичні дані вже отримані. При цьому  $\omega_i$  – гіпотеза, а  $x$  – свідчення, що підтримує гіпотезу. Якщо всі класи  $\omega_i$  характеризуються власними несумісними ймовірностями, що охоплюють усі можливі випадки, то можливо застосувати правило Баєса для вирахування апостеріорних імовірностей кожного класу по апіорних імовірностях цих класів і розподілах умовної імовірності для  $x$ .

Величина  $p(x)$  – однакова для всіх функціональних блоків, а отже її можна ігнорувати та приймати класифікаційне рішення про  $\omega_i$  за максимальним значенням  $p(x/\omega_i)P(\omega_i)$ . Якщо для класу  $\omega_i$  існує  $N$  ознак, то отримуємо:

$$P(\omega_i|x) = P(\omega_i) \prod_{k=1}^N p(x_k|\omega_i). \quad (2.19)$$

Для використання цього рівняння в класифікаторі треба навчити модель на об'єктах, що нас цікавлять. В режимі виконання вираховуються ознаки і знаходяться об'єкти, які дають максимум даних у праву частину рівняння. Зазвичай потім перевіряється, що ймовірність об'єкта-переможця більша за заданий поріг. Якщо це так, то остаточно визначається, що об'єкт присутній на зображенні, інакше – що об'єкт не розпізнано.

На практиці необхідно якимось чином забезпечити обчислення умовної щільності  $p(x/\omega_i)$ . Якщо  $x$  схоже на одну з відомих параметричних моделей, то можливо представити розподіл за допомогою невеликої кількості параметрів, що його характеризують. Зазвичай використовують розподіли Пуассона, експоненційний та нормальний.

Безперервна випадкова величина  $x$  має нормальний закон розподілу (Гаусовий) з параметрами  $\mu$  та  $\sigma^2$ , якщо функція її щільності ймовірності має вигляд:

$$p(x) = N(\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (2.20)$$

в якій

$$\mu = \int_{-\infty}^{+\infty} xp(x)dx, \quad (2.21)$$

та

$$\sigma^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 p(x) dx. \quad (2.22)$$

Зазвичай, коли даних небагато, то такі прості методи, як баєсів класифікатор, дають ліпші результати, ніж більш складні моделі.

### 2.2.2 Двійкові дерева рішень

Двійкові дерева рішень, що їх запропонував Лео Брейман у 1984 році, здійснюють розбиття простору об'єктів відповідно до деякого набору правил розбиття [34]. Ці правила є логічними твердженнями щодо тої чи тої змінної та можуть бути дійсними або хибними.

Ключовими тут є наступні обставини:

- правила дозволяють здійснити послідовну дихотомічну сегментацію даних;
- два об'єкти вважаються схожими, якщо вони опиняються в одному й тому ж сегменті розбиття;

– на кожному кроці розбиття збільшується кількість інформації відносно змінної, що досліджується.

Метод дерев рішень також є «наївним», бо вважає всі ознаки статистично незалежними.

Алгоритм ДКТР (Дерева класифікації та регресії, так їх назвав Брейман) рекурсивно поділяє вихідний набір даних на підмножини, які стають дедалі гомогеннішими до певних ознак, в наслідок чого формується деревовидна ієрархічна структура. Поділ здійснюється на основі традиційних логічних правил ЯКЦО-ТО.

На першій ітерації корінний вузол дерева пов'язується з найоптимальнішим умовним судженням і вся множина об'єктів поділяється на дві групи. Від кожного наступного вузла-батька до вузлів-нащадків також може відходити по дві гілки, які в свою чергу пов'язані з граничними значеннями інших, найбільш оптимальних змінних, і визначають правила дальшого розділення. Кінцевими вузлами дерева є «листя», що відповідає віднайденим рішенням і отримує всі поділені на групи об'єкти навчальної вибірки [35].

Процес, що його описано, відноситься до так званих жадібних алгоритмів, які прагнуть побудувати якнайбільш розгалужене дерево (глибоке дерево). Звичайно, що чим розлогіше й розгалуженіше дерево, тим ліпше будуть наслідки його випробування на навчальній вибірці, але не так успішне – на перевірній вибірці. Тому побудована модель має бути ще й оптимальною за розміром, щоб мати інформацію, що покращує якість розпізнання, і нехтувати тією інформацією, котра його не покращує [36]. Для цього зазвичай здійснюють обрізання дерева – відсічення гілок там, де ця процедура не призводить до значного зростання помилки.

Окрім того, у кожному вузлі дерева вираховується «нечистота»  $i(N)$  відносно даних і здійснюється спроба її зменшити. У випадку класифікації для дерев рішень найчастіше використовують один із трьох способів обчислення нечистоти: ентропія, індекс Джині та помилка класифікації. Різні визначення

нечистоти по різному впливають на рішення про розділення. Частіше за все використовують індекс Джині, але усі алгоритми прагнуть зменшити нечистоту у вузлі.

У наведених далі формулах  $P(\omega_i)$  – частка прикладів у вузлі  $N$ , що належать класу  $\omega_i$ .

Ентропія:

$$i(N) = \sum_j P(\omega_j) \log P(\omega_j). \quad (2.23)$$

Індекс Джині:

$$i(N) = \sum_j 1 - \max P(\omega_j). \quad (2.24)$$

Помилка класифікації:

$$i(N) = \sum_j P(\omega_j) P(\omega_j). \quad (2.25)$$

З усіх статистичних методів класифікації дерева рішень отримали найбільше розповсюдження. Пояснюється це простою реалізацією, легкістю тлумачення результатів, гнучкістю роботи з даними різних типів (категоріальні, числові, ненормовані), здібністю обробляти відсутність даних за допомогою сурогатних розділень та природнім способом призначення важливості ознакам у вигляді порядку розділення [37]. Дерева рішень є основою інших алгоритмів, зокрема посилення та випадкових дерев.

### 2.2.3 Випадкові ліси

Дерева рішень виключно корисні, проте самі собою вони нечасто дають найліпшу якість класифікації. У 2001 році творець алгоритму дерев рішень, Лео Брейман, розробив новий метод класифікації – випадкові ліси, який розвивав ідею дерев рішень і значно покращував його [38].

Випадковий ліс (метод також називають випадкові дерева) можливо навчати одразу декільком класам, збираючи віддані за них голоси у листі кожного з багатьох дерев рішень та обираючи у якості переможця клас, що отримав найбільшу кількість голосів [39].

Отже, в основі випадкового ліса лежить дерево рішень. Воно будується до самого кінця, заки не стане чистим. Це означає, що кожне дерево являє собою перенавчений класифікатор з високою дисперсією, який майже досконало відбиває особливості навчальних даних. Щоб компенсувати високу дисперсію алгоритм бере результати пересічно по багатьох деревах (звідси й назва «випадкові дерева»).

Зрозуміло, що усереднення не принесе ніякої користі, якщо всі дерева корельовані. Для вирішення цієї проблеми алгоритм робить дерева статистично незалежними, обираючи в кожному вузлі випадкову підмножину  $m$  із загальної множини ознак  $M$ , на яких дерево може навчатися. Розмір  $m$  часто задається як квадратний корінь із загального числа ознак. Наприклад, якщо усього потенційних ознак 100, то в кожному вузлі буде випадково обрано 10 з них і знайдено найкраще розділення даних по цих ознаках. Зменшення  $m$  зменшує як кореляцію, так і силу класифікаторів. Збільшуючи розмір підмножини, збільшується й те, й інше. Тому існує деякий оптимальний діапазон значень  $m$  – зазвичай досить широкий. Використовуючи коефіцієнт помилок ООВ (про них нижче), потрібне значення  $m$  у діапазоні можна швидко знайти. Це єдиний регульований параметр, до якого випадкові ліси дещо чутливі.

З метою підвищення робастности в методі випадкових лісів застосовується оцінка по ознаках «поза кошиком» (out-of-bag, ООВ), що не використані для перевірки розділення. Тобто в кожному вузлі проводиться навчання на новій підмножині даних, обраній випадково з поверненням, а решта даних – які не ввійшли до навчальної вибірки – використовуються для оцінки якості розділення. Частка ознак поза кошиком зазвичай складає третину від усіх прикладів.

Як і всі методи, що ґрунтуються на деревах, випадкові ліси мають корисні властивості дерев: сурогатне розділення для обробки значень, що відсутні, вміння працювати з категоріальними й числовими ознаками, необов'язковість нормування й простота знаходження ознак, що важливі для прогнозування [40].

Нарешті, за допомогою випадкових дерев можна визначити міру схожості двох прикладів. Для цього алгоритм після проходження прикладів через дерева підраховує, скільки разів ці приклади опинилися в одному й тому ж вузлі, та ділить це число на загальну кількість дерев. Близькість 1 означає, що приклади повністю ідентичні, 0 – що вони зовсім не схожі.

В оригінальній роботі про випадкові ліси було показано, що рівень помилок лісу залежить від двох речей:

- кореляції між будь-якими деревами в лісі. Збільшення кореляції збільшує рівень помилок лісу;
- сили кожного окремого дерева в лісі. Дерево з низьким рівнем помилок є сильним класифікатором. Збільшення сили окремих дерев зменшує рівень помилок у всьому лісі.

Загалом, випадкові ліси показують дуже високу точність та швидкість, пристосовані до роботи з великими базами даних і дають оцінку того, які ознаки важливі для класифікації.

### 2.2.4 Метод $k$ -найближчих сусідів

Одним з найпростіших методів класифікації є метод  $k$ -найближчих сусідів, що спирається на міру близькості між класами.

На початку класифікації задається деяке число  $k$  – кількість найближчих сусідів. Далі необхідно віднайти  $k$  прикладів, які є найближчими до об'єкта класифікації  $y$  і відповідно до того, до якого класу відноситься відносна більшість з цих сусідів, здійснюється рішення про класифікацію  $y$ .

На рисунку 2.12 зображено деякий приклад задачі класифікації. Суцільною лінією виділено окіл із  $k = 3$ , в який потрапили 3 сусідні об'єкти, в цьому випадку зелений об'єкт в центрі, для якого здійснюється класифікація, буде віднесено до класу червоних трикутників. Пунктирною лінією позначено випадок із  $k = 5$ , при якому більшість з найближчих сусідів належать класу блакитних квадратів, тому суб'єкт класифікації буде віднесено саме до цього класу.

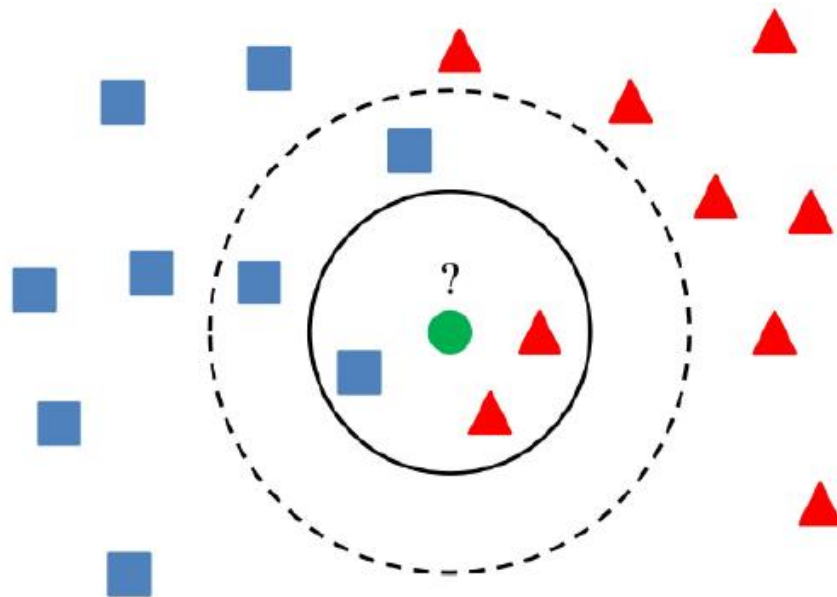


Рисунок 2.12 – Приклад задачі класифікації при  $k = 3$  та  $k = 5$

Розрахунок відстаней виконується наступним чином. Нехай для навчального прикладу  $x$  відомий вектор ознак  $(x_1, x_2, \dots, x_n)$ , для об'єкта  $y$ , що

його класифікують, вектором ознак  $\epsilon (y_1, y_2, \dots, y_n)$ . Тоді для обчислення відстані  $d(x, y)$  між  $y$  та  $x$ , використовуючи Евклідову метрику:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}. \quad (2.26)$$

Для категоріальних ознак, для яких неможливо порахувати відстань таким способом, можна застосувати функцію відмінності, яка має наступний вигляд:

$$dd(x, y) = \begin{cases} 0, & x = y \\ 1, & x \neq y \end{cases}. \quad (2.27)$$

Оскільки різні ознаки мають різний діапазон можливих значень у вибірці, то значення відстані можуть мати значну залежність від ознак з великими діапазонами, тому частіше за все необхідна попередня нормалізація даних. Частіше за все використовуються дві функції нормалізації.

Мінімаксимальна нормалізація:

$$x' = \frac{x - \min(X)}{\max(X) - \min(X)}, \quad (2.28)$$

в такому випадку значення всіх атрибутів будуть знаходитись у проміжку  $[0, 1]$ .

Z-нормалізація:

$$x' = \frac{x - M(X)}{\sigma_X}, \quad (2.29)$$

де  $M(X)$  – середнє значення;

$\sigma_X$  – стандартне відхилення.

Окрім того, при підрахунку відстані іноді враховують значимість атрибутів, яка попередньо суб'єктивно визначається експертом. В такому випадку, при обчисленні відстані кожен  $i$ -й квадрат різниці в сумі помножується на коефіцієнт важливості  $Z_i$ . Наприклад, якщо атрибут А у три рази важливіший за атрибут В ( $Z_A = 3, Z_B = 1$ ), то відстань буде обчислюватися наступним чином:

$$d(x, y) = \sqrt{3(x_A - y_A)^2 + (x_B - y_B)^2}. \quad (2.30)$$

Після визначення  $k$  найближчих сусідів залишається лише порахувати, скільки сусідів відноситься до певних класів (так звана кількість голосів), і віднести об'єкт  $y$  до класу з найбільшою кількістю голосів. Це так зване незважене голосування, в якому не враховується відстань, а лише належність до класу, що може призвести до того, що більше одного класу набрали найбільшу кількість голосів і доведеться обирати (частіше за все випадково або за допомогою певних функцій), до якого саме класу віднести  $y$ . Для уникнення таких ситуацій використовують зважене голосування. Голоси за клас підраховуються за наступною формулою:

$$votes(class) = \sum_{i=1}^n \frac{1}{d^2(y, x_i)}, \quad (2.31)$$

де  $d^2(y, x_i)$  – квадрат відстані від відомого об'єкту  $x_i$  до  $y$ ;

$class$  – черговий клас.

Метод  $k$ -найближчих сусідів – це так званий «лінійний класифікатор», бо при навчанні він не виконує ніяких дій, окрім збереження вхідних даних для дальшого використання, тому навчання виконується дуже швидко. Але в роботі алгоритм споживає багато пам'яті та повільно проводить класифікацію, хоч почасти й дає хороші результати. Дуже часто перед

застосуванням цього методу навчальний набір кластеризують з метою зменшення розміру [41].

### 2.2.5 Метод опорних векторів

Метод опорних векторів (support-vector machines, SVM) – це алгоритм класифікації, що в своїй найпростішій формі застосовується для поділу на два класи, але існують узагальнення алгоритму SVM для багатокласової класифікації [42]. В основі SVM лежить ідея ядра, що дозволяє відобразити множину точок у просторі розмірності  $N_d$  у простір вищої розмірності  $N_{KS}$ , який називають простором ядра. У цьому просторі можна знайти лінійний класифікатор, що поділяє два класи, навіть якщо у вихідному просторі дані не були лінійно розподілені [43, 44] (рис. 2.13).

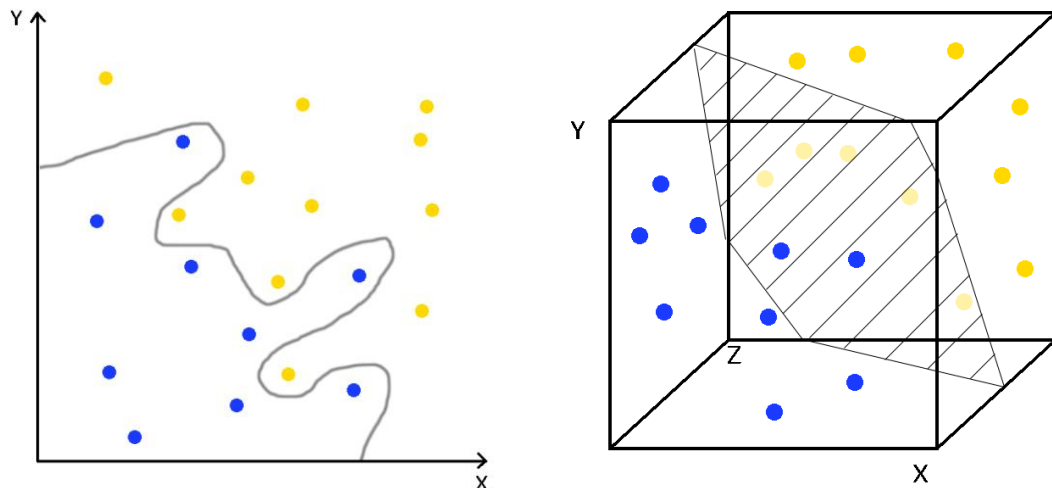


Рисунок 2.13 – Точки у двовимірному просторі та вони ж у просторі ядра

SVM-класифікатор називається класифікатором з максимальним прозором, бо гіперплощина у просторі ядра обирається так, щоб вона не тільки поділяла два класи, але ще й відстань між представниками цих класів, що є найближчими до гіперплощини (прозир), була найбільша. Ці найближчі

представники, які й визначають положення гіперплощини, називають опорними векторами. Важливість віднайдених опорних векторів полягає в тому, що тільки їх треба зберігати для прийняття рішень про класифікацію майбутніх прикладів.

Формально вирішальну гіперплощину в лінійному алгоритмі SVM можна описати рівнянням

$$\vec{\omega} \cdot \vec{x} + b = 0, \quad (2.32)$$

де  $\vec{x}$  – точка у просторі ядра;

$\vec{\omega}$  – вектор, нормальний до гіперплощини.

Вектор  $\vec{\omega}$  усталено нормувати, тож відстань між вирішальною гіперплощиною й опорними векторами дорівнює  $1/\|\vec{\omega}\|$  в кожному напрямку. Значення  $b$  визначає зсув гіперплощини. За такої параметризації кінці опорних векторів належать площинам, що визначені рівняннями  $\vec{\omega} \cdot \vec{x} + b = 1$  та  $\vec{\omega} \cdot \vec{x} + b = -1$  (рис. 2.14).

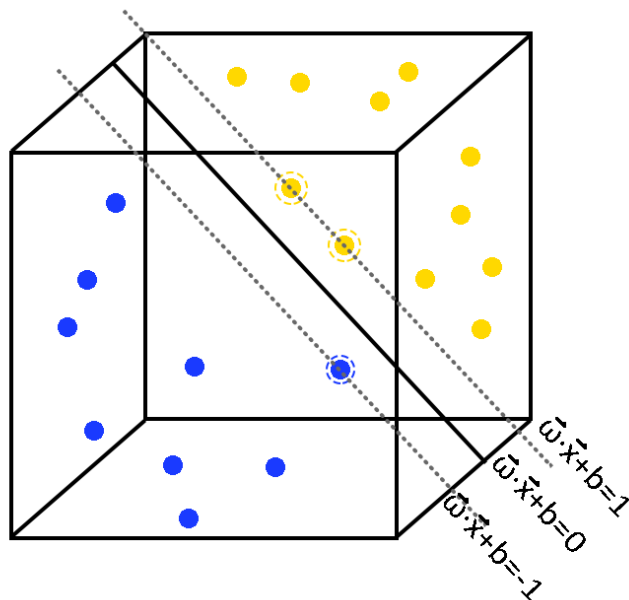


Рисунок 2.14 – Простір ядра з визначеними опорними векторами

Вектор  $\vec{\omega}$  можна визначити у вигляді лінійної комбінації опорних векторів, зсув  $b$  також можна визначити по самих лише опорних векторах. Знаючи ці два параметри, що визначають вирішальну гіперплощину, класифікатор при отриманні нової точки  $x$  може визначити, в якій з чотирьох областей вона знаходиться: глибоко всередині першого класу ( $\vec{\omega} \cdot \vec{x} + b \geq 1$ ), близько до границі першого класу ( $\vec{\omega} \cdot \vec{x} + b > 0$ ), близько до границі другого класу ( $\vec{\omega} \cdot \vec{x} + b < 0$ ) або глибоко всередині другого класу ( $\vec{\omega} \cdot \vec{x} + b \leq -1$ ).

Ядро – це відображення вихідного простору навчальних векторів у простір більшої розмірності, а також зворотне до нього. Ядро записується у вигляді  $K(\vec{x}_i, \vec{x}_j)$ , а пряме та зворотне відображення – у вигляді  $\vec{\varphi}(\vec{x}_i)$  і  $\vec{\varphi}(\vec{x}_j)$  відповідно. За визначенням справедливе наступне рівняння:

$$K(\vec{x}_i, \vec{x}_j) = \vec{\varphi}(\vec{x}_i) \cdot \vec{\varphi}(\vec{x}_j). \quad (2.33)$$

Важливим є те, що в SVM відображення  $\vec{\varphi}(\vec{x})$  ніколи не доводиться вираховувати явно для знаходження роздільної гіперплощини, тому що всі обчислення у просторі ядра складаються тільки зі скалярних векторних добутків. Отже  $\vec{\varphi}(\vec{x})$  завжди входить до складу  $\vec{\varphi}(\vec{x}_i) \cdot \vec{\varphi}(\vec{x}_j)$  для знаходження деякої пари точок  $x_i$  та  $x_j$ , тобто не потрібно обчислювати саме відображення, а лише ядро  $K(\vec{x}_i, \vec{x}_j)$ .

Від вибору ядра сильно залежить, наскільки добре побудована гіперплощина буде розділяти навчальні дані та узагальнюватись на наступні, хоча робота алгоритму у просторах високої вимірності збільшує похибку узагальнення, що компенсується збільшенням кількості зразків.

Найпоширенішими ядрами є:

а) лінійне

$$K(\vec{x}_i, \vec{x}_j) = \vec{x}_i \cdot \vec{x}_j; \quad (2.34)$$

б) поліномінальне

$$K(\vec{x}_i, \vec{x}_j) = (\vec{x}_i \cdot \vec{x}_j + c)^d; \quad (2.35)$$

в) Гаусове радіально-базисне

$$K(\vec{x}_i, \vec{x}_j) = e^{-\gamma \|\vec{x}_i - \vec{x}_j\|^2}; \quad (2.36)$$

г) сигмоїдне

$$K(\vec{x}_i, \vec{x}_j) = \tan(\kappa \vec{x}_i \cdot \vec{x}_j + c); \quad (2.37)$$

д) експоненційне  $\chi^2$ -квадрат

$$K(\vec{x}_i, \vec{x}_j) = e^{-\gamma \frac{\vec{x}_i - \vec{x}_j}{\vec{x}_i + \vec{x}_j}}. \quad (2.38)$$

Існує також метод розпізнання об'єктів, що базується на опорних векторах, який називається HOG-SVM [45]. В ньому застосовується ковзне вікно, яке розбиває на менші ділянки і обчислює гістограму напрямків градієнтів зображення в кожній ділянці. Ці гістограми, що їх називають HOG (histogram of oriented gradients – гістограма орієнтованих градієнтів), поєднують, отримуючи тим самим вектор ознак, який далі передають класифікатору SVM.

### 2.3 Двовимірні підходи до розпізнавання облич

В силу наймовірної актуальності теми розпізнавання облич, в останні десятиліття розробниками та науковцями проводилися активні дослідження в цьому напрямку, в наслідок чого було винайдено безліч методів вирішення проблеми, що засновані на різноманітних підходах. Деякі з них вже сильно

застаріли й не можуть надавати задовільні показники роботи, деякі, незважаючи на свою давнину, все ще демонструють хороші результати, тому застосовуються як такі або як основа для інших, а деякі є надто новими, тому потребують більшого вивчення та розвитку. Існують різні варіанти класифікації цих підходів за тими чи тими ознаками, але загалом їх можна поділити на три основні напрямки: цілісні методи, місцеві методи та методи, що засновані на глибокому навчанні (рис. 2.15).



Рисунок 2.15 – Класифікація підходів до розпізнання облич

Звичайно ж, існують також комбіновані методи (в особливості ті, що об'єднують техніки цілісного та місцевого підходів), які іноді виділяють в окремий напрямок. Далі в роботі розглянуто основні принципи, якими послуговуються двовимірні методи розпізнання, та загальний опис деяких з них.

### 2.3.1 Цілісний підхід

Цілісні або підпросторові методи обробляють усе обличчя в цілому, тобто вони не вимагають вилучення окремих ділянок обличчя або його

ключових точок. Їх можна розділити на лінійні та нелінійні – в залежності від того, який метод застосовується для представлення підпростору.

Найпопулярнішими лінійними методами, що використовуються для систем розпізнавання обличчя, є аналіз основних компонентів (PCA), лінійний дискримінаційний аналіз (LDA) та незалежний аналіз компонентів (ICA).

Аналіз основних компонентів, відомий також як власні обличчя (eigenfaces) [46] ґрунтується на ідеї, що будь-яка колекція з  $M$  зображень містить надмірності, які можна усунути, застосовуючи тензорне розкладання. На початку цей метод застосовувався в статистиці для зниження простору ознак з найменшою втратою інформації. У галузі розпізнавання обличчя його застосовують для представлення зображення лиця вектором малої вимірності (власним вектором), який потім порівнюється з еталонними векторами, що зберігаються в БД. Із власних векторів можливо реконструювати зображення обличчя (рис. 2.16).



Рисунок 2.16 – Реконструйовані з власних векторів обличчя

Набір власних векторів, отриманий одного разу на навчальній вибірці зображень обличчя, використовується для кодування решти зображень обличчя, які становлять зважене сполучення цих власних векторів. За допомогою обмеженої кількості власних векторів можна отримати апроксимацію вхідного зображення лиця, яку потім можна зберігати в БД у вигляді вектору коефіцієнтів, який одночасно є ключем пошуку в БД обличчя.

Класифікація часто проводиться шляхом проектування нового зображення обличчя та обчислення міри відстані до усіх класів, що їх описано в цьому підпросторі.

Метод власних облич показує непогані результати, проте, в тих випадках, коли на зображенні обличчя спостерігаються значні зміни освітлення або виразу, його ефективність значно зменшується. Для вирішення цієї проблеми було запропоновано метод, що використовує лінійний дискримінант Фішера, який називається обличчя Фішера (Fisherfaces) або лінійний дискримінантний аналіз (LDA) [47].

LDA шукає проєкцію даних, за якої класи є максимально лінійно розділеними. Це головна відмінність від PCA, який знаходить таку проєкцію даних, за якої буде максимізовано розсіп по всій БД, без врахування класів.

Для всіх зразків з усіх класів обчислюються матриці розсіпу між класами ( $S_w$ ) та всередині класів ( $S_b$ ) за наступними формулами:

$$S_w = \sum_{i=1}^C \sum_{x_k \in X_i} M_i (x_k - \mu)(x_k - \mu)^T, \quad (2.39)$$

$$S_b = \sum_{i=1}^C M_i (x_i - \mu)(x_i - \mu)^T, \quad (2.40)$$

де  $\mu$  – середній вектор зразків  $i$ -го класу;

$X_i$  – набір прикладів, що належать  $i$ -му класу;

$x_k$  – вектор, що описує  $k$ -й приклад;

$C$  – кількість класів;

$M_i$  – кількість прикладів у  $i$ -му класі.

Головною метою є максимізація відношення  $\det|S_b| / \det|S_w|$ , тобто мінімізувати  $S_w$  і максимізувати  $S_b$ . На рисунку 2.17 зображено приклад шести реконструйованих з власних векторів облич Фішера.



Рисунок 2.17 – Обличчя Фішера

Інший підхід полягає у використанні фільтрів Габора. Фільтри Габора – це просторові синусоїди, розташовані у вікні Гауса, які дозволяють отримувати особливості зображення, лише шляхом зміни значень їх частоти, орієнтації та розміру [48]. Фільтри Габора застосовуються для отримання вирівняних рис обличчя, зображення яких отримано під різними кутами. Крім того, використовується техніка оптимізації вибору ознак для віднайдення оптимального простору ознак. Фільтри Габора часто використовуються в поєднанні з іншими методами, наприклад PCA, що значно підвищує ефективність роботи.

Методи аналізу частотної області пропонують представляти людське обличчя, як функції низькочастотних компонентів, що демонструють високу енергію. Дискретне перетворення Фур'є (DFT), дискретне косинусне перетворення (DCT) та дискретне вейвлет-перетворення (DWT) не залежать від даних, тому не потребують навчання.

У роботі [49] автори використали метод двовимірного дискретного вейвлет-перетворення (2D-DWT) для розпізнавання обличчя з використанням нової стратегії виділення ключових ознак. У роботі пропонується стратегія нерівномірного виділення ознак для низькочастотного підпростору верхнього рівня із використанням техніки інтегральної проєкції для двох високочастотних підпросторів 2D-DWT верхнього рівня на основі середнього зображення всіх навчальних зразків. Ця стратегія виділення ознак зберігає цілісність місцевої інформації та є більш придатною для відображення особливостей структури зображення обличчя.

Дискретне косинусне перетворення – це перетворення, яке представляє кінцевий набір даних як суму ряду косинусних функцій, що коливаються в різних частотах.

### 2.3.2 Місцевий підхід

У контексті розпізнавання обличчя місцеві, що їх відносять до місцевого підходу, розглядають риси обличчя окремо одну від одної. Основна мета цих методів – виявити відмінні риси. До цього напрямку відноситься й розпізнавання облич за допомогою звичайного визначення ключових точок (за допомогою алгоритмів SURF, SIFT, BRISK тощо) та класифікація на їх основі за допомогою статистичних алгоритмів. Проте далі буде розглянуто методи, що ґрунтуються на дещо інших принципах виділення ознак. Вони поділяються на геометричні та текстурні.

Геометричні методи керуються характеристиками орієнтирів, локалізованих у просторі, що розглядається [50]. Ці орієнтири можуть дати корисну інформацію при роботі алгоритмів розпізнавання. Орієнтири на обличчі використовуються для визначення рис обличчя, нормалізації виразу та розпізнавання визначених ділянок на основі геометричного розподілу. Незважаючи на те, що глибокі дослідження, узагальнені в краніології, визначили усі важливі для розпізнавання обличчя точки, в різних алгоритмах використовують різну множину орієнтирів.

Не всі ділянки обличчя на зображенні дають однакову кількість інформації. Наприклад, лоб і щоки мають просту структуру та меншу кількість виразних візерунків порівняно з носом або очима. Найчастіше використовувані орієнтири на обличчі – це кінчик носа, кути очей та рота, брови, середини зіниць, верхня частина вуха, ніздрі та перенісся.

Одним з поширених геометричних методів розпізнавання обличчя є гнучке зіставлення графів (EGM), в якому зіставляються графи, що описують

зображення облич. Обличчя представлені графами зі зваженими вершинами та ребрами. На етапі розпізнання один з графів – еталонний – залишається незмінним, в той час як другий деформується з метою найліпшого підганяння до першого (рис. 2.18).



Рисунок 2.18 – Схематичне зображення накладеного на обличчя графу

На зображення об'єкта накладається граф (може бути прямокутної форми, або форми обличчя), у вершинах якого виконується згортка з набором вейвлетів Габора. Деформація графу відбувається шляхом зміщення кожної з його вершин на деяку відстань в певному напрямку відносно її вихідної позиції та обрання такого її становища, за якого різниця між значеннями ознак (відгуків на вейвлети Габора) у відповідних вершинах графів була найменшою. Ребра графу зважуються відстанню між сусідніми вершинами. Відмінність між двома графами обчислюються за допомогою певної цінової функції деформації, що враховує як відмінність між значеннями ознак, так і ступінь деформації графу.

Описаний процес відбувається для всіх зображень, а потім знаходиться зображення з набору, для якого функція деформації дала найкраще значення.

Згідно з дослідженнями, алгоритм показує високу точність навіть з невеликим нахилом голови, проте чутливий до зміни виразу обличчя та має високу обчислювальну вартість і лінійну залежність швидкості роботи від кількості зразків у БД.

Розвитком алгоритму EGM є алгоритм EBGM – групове гнучке зіставлення графів, в якому граф-еталон обчислюється на основі одразу

декількох зображень обличчя (з різними виразами та поворотами голови). Це дозволяє графу розпізнавати обличчя зі зміненим виразом та нахилами, проте проблема низької швидкості залишається невирішеною.

Ще одним методом, близьким до EGM, є морфологічне гнучке зіставлення графів (MEGM). В цьому алгоритмі характеристики Габора заступаються морфологічними характеристиками, що отримуються за допомогою фільтрування зображення операторами ерозії та розширення.

Активні моделі зовнішнього вигляду (active appearance model, AAM) – це статистичні моделі зображень, які шляхом різноманітних деформацій підлаштовуються до конкретних зображень. Модель складається з 68 орієнтирів, які позначають головні риси обличчя та його форму. Перед використанням модель необхідно навчити на множині заздалегідь розмічених зображень облич. Кожна мітка має власний номер і визначає особливу точку, котру муситиме знаходити модель під час підлаштування до нового зображення.

На початку, методом узагальненого прокрустового аналізу нормалізуються усі форми для компенсації відмінностей масштабу, нахилу та зсуву. Після цього, методом аналізу основних компонентів з кожної точки виділяються власні вектори. Для отриманої форми  $S$  правдиве наступне рівняння:

$$s = s_0 + \Phi_s b_s, \quad (2.41)$$

де  $s_0$  – базова форма, усереднена по всіх елементах навчального набору;

$\Phi_s$  – матриця власних векторів;

$b_s$  – параметри форми.

Таким робом, змінюючи вектор параметрів  $b_s$  можливо деформувати форму для її підлаштування до необхідного зображення. На рисунку 2.19 зображений приклад представлення форми за допомогою базової форми та комбінації власних векторів із заданими параметрами.

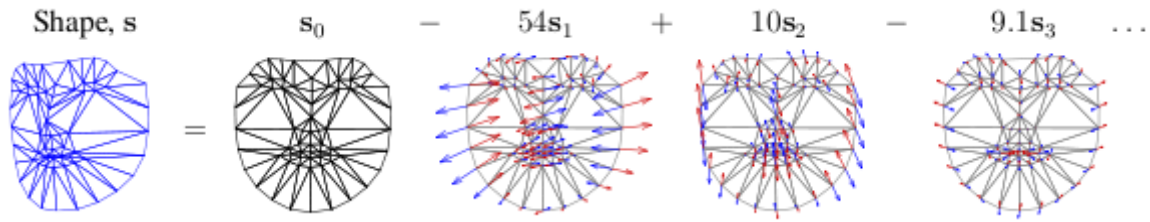


Рисунок 2.19 – Приклад представлення форми

Після отримання базової форми та матриці власних векторів із зображень виділяються і нормалізуються текстури обличчя. Нормалізація виконується для компенсації різних умов освітлення. Потім з текстур формується матриця, в якій кожен стовпчик зберігає значення пікселів відповідної текстури. У випадку багатоканальних текстур вектори пікселів формуються окремо по кожному з каналів, а потім виконується їх об'єднання. Після знаходження головних компонент матриці текстур отримуємо рівняння для синтезованої текстури:

$$t = t_0 + \Phi_t b_t, \quad (2.42)$$

де  $t_0$  – базова текстура;

$\Phi_t$  – матриця власних текстур;

$b_t$  – вектор параметрів текстур.

Як і у випадку з формами, змінюючи вектор  $b_t$  можливо підігнати текстури до конкретного зображення.

Існує також споріднений до ААМ метод – активні моделі форм (active shape models, ASM). Цей метод концентрується на формі обличчя й полягає в обчисленні статистичних зв'язків між позиціями антропометричних точок. Основною відмінністю ASM від ААМ є те, що ASM не враховує текстуру обличчя, що значно підвищує швидкість роботи алгоритму, проте призводить до втрати важливої інформації.

Насправді, головною метою ААМ та АSM є не розпізнавання облич, а їх точна локалізація та отримання важливих особливостей, які потім можуть бути використані іншими алгоритмами. Окрім того, їх використовують для вирівнювання зображення обличчя, що частіше за все необхідно для дальшої роботи.

Ще одним різновидом місцевого підходу є вилучення місцевих текстур лица. Методика вилучення функцій, орієнтованих на знання про текстуру, відіграє значну роль у розпізнаванні облич. Дескриптори локальних структур виразні, стійкі до поганого освітлення, дисперсії яскравості і не потребують сегментації [51].

У 2004 році була представлена робота [52], в якій описується метод зображення обличчя на основі дескриптора місцевої текстури, який називається місцевий двійковий візерунок (local binary pattern, LBP). Зображення обличчя поділяється на кілька блоків, по яких проходить матриця розміром  $3 \times 3$  і будує двійковий код цього блоку. На кожному кроці проходження матриці по блоку центральний піксель  $p$  є порогом, відносно якого обчислюється двійковий код: якщо якийсь піксель, що знаходиться навколо  $p$ , має більшу яскравість, ніж  $p$ , то йому присвоюється 1, інакше – 0. В такий спосіб, отриманий двійковий код зберігає усю інформацію про текстуру обличчя. В кінці, з коду кожного блоку створюється гістограма, усі гістограми об'єднуються в одну загальну гістограму текстури обличчя щоб сформувати вектор ознак.

Згідно з рядом досліджень, LBP демонструє чудові показники ефективності роботи та точності розпізнавання. LBP продемонстрував стійкість до багатьох ускладнень, викликаних зміною освітлення, виразу, старіння людини.

Пізніше були створені модифікації алгоритму LBP, в одній з яких отримана гістограма розглядається як імовірнісний розподіл, а не статистичні дані. Із застосуванням техніки адаптації апостеріорного максимуму отримується специфічна для певного зразка модель.

Інша модифікація ґрунтується на зауваженні, що отримана гістограма є дуже розрідженою, а якщо розміри блоків, на які на початку поділяють зображення, досить великі, то гістограма виходить надто щільна, а тому не є дуже ефективною для відображення місцевих змін обличчя. Отже, для кожного блоку застосовується векторне квантування, щоб отримати корисний вектор ознак, тобто візерунки кожного блоку поділяються на декілька груп і обличчя описується кодовою книгою, що містить лише вагомі мітки LBP та нехтує неефективних міток.

Існує також одновимірний варіант LBP (1DLBP), в якому використовується не двовимірна LBP-матриця, а одновимірний вектор. Після обробки кожного блоку, вектори цього блоку об'єднуються в один глобальний вектор. Потім множину глобальних векторів мінімізують із застосуванням методу PCA. Для визначення ступеня подібності між двома обличчями застосовується відстань  $\chi^2$ -квадрат.

Також був представлений алгоритм, завдяки якому можливо розпізнавати розмиті обличчя, він називається місцеве фазове квантування (LPQ). Оператор LPQ ґрунтується на квантуванні ділянок фази перетворення Фур'є. Фаза вважається нечутливою до розмиття властивістю у специфічних умовах, які дуже часто виконуються. В цьому методі, подібно до LBP, будуються гістограми міток, розраховані в різних блоках, які використовуються як дескриптор зображення обличчя.

Для розпізнання облич з низькою роздільною здатністю був запропонований місцевий дескриптор частоти (LFD). Як і LPQ, пропонований дескриптор ґрунтується на інформації про місцеві частоти, що робить його стійким до низької роздільної здатності та розмитості зображення. Але, на відміну від LPQ, LFD враховує не тільки фазу, а й величину, тим самим забезпечуючи більше інформації. Масштабні тестування показали, що LFD є дійсно надійним дескриптором і корисним для реальних застосунків.

Підсумовуючи, методи, що засновані на місцевих текстурах, характеризуються високою ефективністю, якістю та швидкістю, яка дозволяє використовувати їх в реальному часі.

### 2.3.3 Глибоке навчання

З появою глибокого навчання на початку 2010-х років відбувся значний прорив у царині комп'ютерного зору і класифікації об'єктів зокрема. Сьогодні глибокі нейронні мережі переважають у популярності та якості розпізнання більшість інших методів вилучення ознак і класифікації.

В задачах комп'ютерного зору серед глибоких мереж наразі переважають згорткові нейронні мережі (convolution neural networks, CNN), в яких, окрім звичайних, є також декілька згорткових шарів, пристосованих до вилучення характеристик з матриці зображення (рис. 2.20).

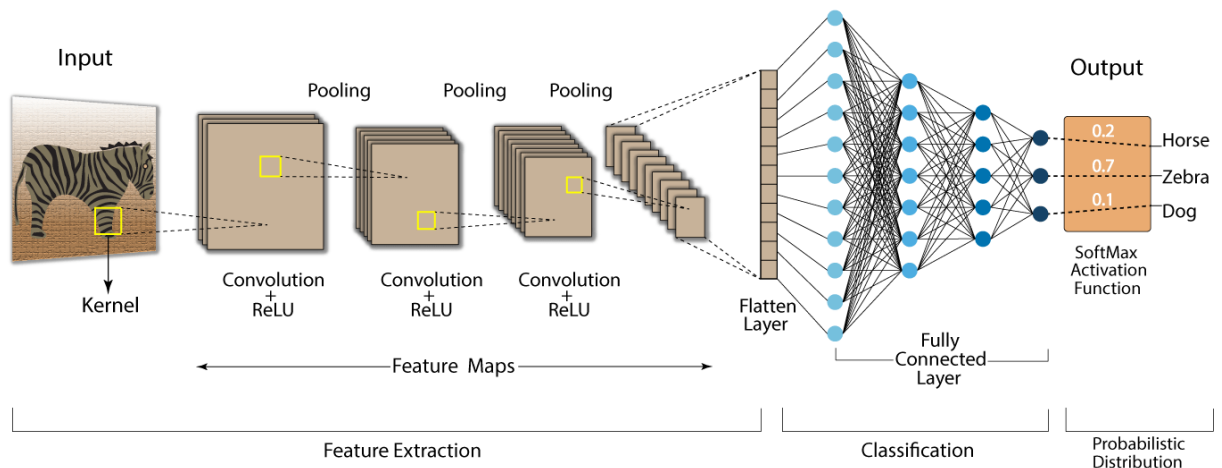


Рисунок 2.20 – Приклад роботи згорткової нейронної мережі

Сьогодні існує безліч згорткових нейронних мереж з різноманітними архітектурами і навчальними наборами. Найвідомішими й найефективнішими мережами наразі є AlexNet, VGGNet, GoogleNet, DeepFace, ResNet та SENet. Усі вони станом на сьогодні мають приблизно однакову точність розпізнання

й порівняно чудово справляються зі змінами виразу обличчя, оклюзією, старінням та поворотами голови.

Серед недоліків згорткових нейронних мереж можна виділити необхідність у дуже великій кількості даних для отримання принаймні непоганих результатів, довгочасність навчання (іноді навчання відбувається впродовж декількох днів), складність початкового визначення та налаштування параметрів мережі, проблема перенавчання та необхідність повного перенавчання мережі у випадку додавання нового еталонного обличчя в БД.

#### 2.4 Розроблення методики ідентифікації облич на кадрах відеозйомки

Будь-який алгоритм розпізнавання облич складається з чотирьох основних етапів, які зображено на рисунку 2.21.

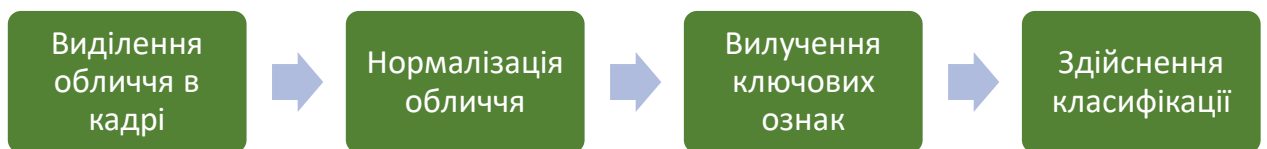


Рисунок 2.21 – Основні етапи розпізнавання обличчя

Для виділення обличчя в кадрі існує багато методів, найвідомішим з яких є алгоритм Віоли-Джонса, що ґрунтується на застосуванні двійкових каскадів Гаара. Даний алгоритм є дуже швидким, що дозволяє використовувати його в реальному часі, і показує непогані результати, проте все ж дає багато хибних спрацьовувань, сильно залежний від нахилів та поворотів і виділяє тільки квадратну область обличчя, залишаючи значну його частину поза рамкою, що може призвести до втрати значної частини важливої інформації про обличчя.

Серед сучасних методів вирішення цієї задачі було вирішено використати глибоке навчання, а саме Single Shot Detector (SSD), в основі якого лежить мережа ResNet. Цей детектор забезпечує високу швидкість і точність визначення обличчя, з мінімальною кількістю хибних результатів, є надзвичайно стійким до поворотів, аж до визначення обличчя в профіль, і виділяє прямокутну область обличчя, завдяки чому всі дані обличчя зберігаються.

Попередня обробка, цебто нормалізація, полягає в очищенні зображення від шуму, вирівнюванні гістограми та приведенні області обличчя до фіксованого розміру. Для очищення від шуму було обрано медіанний фільтр, бо він зберігає чіткість границь, на відміну від фільтрів розмиття, таких як фільтр Гауса.

Для виділення ключових ознак було обрано підхід вилучення текстурних особливостей обличчя, а саме метод LBP. Ідея методу полягає у поділі зображення на певну кількість областей, після чого для кожної області створюється окрема гістограма текстурних ознак, яка будується на основі результатів обробки зображення LBP-оператором. LBP-оператор проходить по всіх пікселях області і порівнює центр оператора з його сусідами, завдяки чому створюється двійковий код. Оригінальний LBP-оператор являв собою вікно розміром  $3 \times 3$ , в якому було 8 сусідів (рис. 2.22).

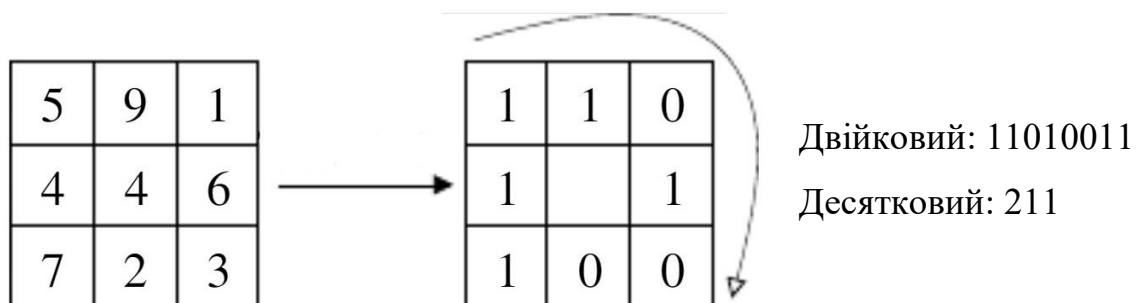


Рисунок 2.22 – Базовий МДФ-оператор

Потім було винайдено декілька розширень цього оператора. Перше розширення полягає у додаванні можливості вибору довільної кількості сусідів  $P$ , які знаходяться на кільці з радіусом  $R$  навколо центру (рис. 2.23). Такий оператор позначається  $LBP_{P,R}$ . Для визначення значень точок, що знаходяться між пікселями використовується дволінійна інтерполяція.

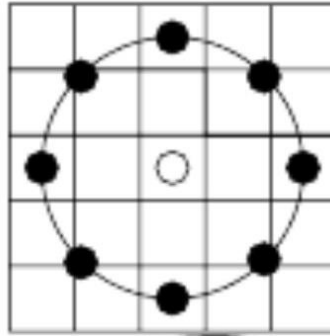


Рисунок 2.23 – Розширений LBP-оператор  $LBP_{8,2}$

Друге покращення полягає у використанні так званих рівномірних візерунків. LBP-оператор називається рівномірним, якщо він містить не більше двох двійкових переходів. Наприклад, 00000000, 00011110 та 10000111 є рівномірними, а 01011001 – ні. Отже, із цим розширенням беруться до уваги тільки рівномірні оператори, а решта відкидаються і не враховуються при побудові гістограм.

Гістограми усіх областей поєднуються в одну спільну гістограму, яку можна обчислити наступним чином:

$$H_{i,j} = \sum_{x,y} I\{f_l(x,y) = i\} I\{(x,y) \in R_j\}, i = 0, \dots, n-1, j = 0, \dots, m-1, \quad (2.43)$$

де  $n$  – кількість різноманітних результатів роботи LBP-оператора;

$m$  – кількість областей;

$x, y$  – координати точки;

$I\{f_l(x,y) = i\}$  – бінарна функція оператора.

Після об'єднання гістограм вони нормалізуються і настає етап класифікації. Коли зразків на кожний клас небагато, використовувати складні класифікатори недоцільно, тому в даному випадку підійде алгоритм найближчого сусіда. Як вже було розглянуто в цій роботі, існує декілька способів обчислення відстані між зразками, серед них було обрано модифіковану метрику  $\chi^2$ -квадрат, яка розроблена саме для порівняння текстур зображень. Її формула:

$$\chi^2(H_1, H_2) = 2 * \sum_i \frac{(H_1(I) - H_2(I))^2}{H_1(I) + H_2(I)}. \quad (2.44)$$

Таким робом, визначення відстані між двома гістограмами обчислюється звичайним підсумовуванням попарно оброблених елементів цих гістограм.

До описаного алгоритму було вирішено додати ще одне розширення. У роботі [11] досліджено, що текстурні методи дають ліпші результати розпізнання, якщо враховувати усі кольорові канали зображення, а не працювати тільки з градаціями сірого, як це зазвичай робиться. Отже, для кожного каналу будується окрема гістограма, а потім вони усі об'єднуються в спільну гістограму зображення.

### **3 РОЗРОБЛЕННЯ ЗАСТОСУНКУ ДЛЯ ІДЕНТИФІКАЦІЇ ОБЛИЧ НА КАДРАХ ВІДЕОЗЙОМКИ**

#### **3.1 Вибір інструментальних засобів для створення застосунку**

У ролі допоміжної бібліотеки для розроблення застосунку було використано OpenCV – бібліотеку комп'ютерного зору та обробки зображень з відкритим кодом, яка має надзвичайно потужне функціональне наповнення. В бібліотеці є безліч функцій та алгоритмів для вирішення різноманітних задач, пов'язаних із роботою з зображеннями, серед яких:

- засоби зчитування, зберігання та запису зображень;
- колірні та геометричні перетворення зображень;
- ядра згортки;
- відстежування руху;
- робота з гістограмами;
- розпізнання об'єктів;
- машинне навчання;
- визначення ключових точок та інше.

Мовою програмування було обрано C++, одну з найпопулярніших мов програмування в світі, через її високу швидкодію, велику стандартну бібліотеку та широкі можливості для оптимізації.

Як середовище розробки було обрано Microsoft Visual Studio – програмний продукт компанії Microsoft, який підтримує OpenCV та має велику кількість функціоналу, що може стати в нагоді під час роботи. Це середовище найчастіше використовується для розробки застосунків мовою C++.

### 3.2 Етапи розроблення застосунку для ідентифікації облич на кадрах відеозйомки

Застосунок, що його розроблювано, фактично має складатися з двох частин – частина, що відповідає за навчання моделі, та частина, в якій відбувається класифікація нових даних. Було вирішено реалізувати можливість роботи як з даними з відеокамери, так і зі звичайними зображеннями, що знаходяться в певному каталозі. Це стосується як процесу навчання, так і процесу розпізнавання. Починається робота програми з того, що користувач обирає необхідні дії через інтерфейс командного рядка, а саме треба обрати режим роботи застосунку (навчання/розпізнавання) та джерело, з якого будуть братися дані для обробки (камера/каталог) (рис. 3.1).

```
C:\Users\Vasya\source\repos\TestOpenCV\x64\Release>TestOpenCV.exe -w
Incorrect arguments
usage: prog.exe [mode] [data_source]
Mods:
  --learn/-l : learning mode
  --work/-w  : working mode
Data sources:
  cam       : using camera for learning or recognition processes
  path      : full path to learning data of some class or person photo for recognition
```

Рисунок 3.1 – Відповідь застосунку на неправильні аргументи

В режимі навчання першим етапом є отримання навчальних зображень з одного або другого джерела. У випадку, якщо користувач увів шлях до певного каталогу, з цього каталогу завантажуються усі зображення, що є в наявності. Якщо в якості джерела даних обирається камера, то вона запускається і після натискання користувачем клавіші Space протягом трьох секунд записуються 30 кадрів відеопотоку, які будуть використані для навчання моделі.

В усіх режимах роботи є необхідність виділяти обличчя на зображеннях, для чого було обрано функціонал модуля глибокого навчання бібліотеки OpenCV. Після перевірки введених аргументів командного рядка здійснюється завантаження будови та ваг глибокої мережі, після чого її можна

застосовувати в процесі завантаження зображень. Спочатку зображення перетворюється на чорно-біле, а його розмір зменшується до розміру 300×300 (на такому розмірі була навчена мережа), потім змінене зображення подається на вхід мережі і в результаті отримуються усі регіони зображення, на яких є людські обличчя. Для відкидання недосить достовірних областей з обличчями використовується поріг `confidenceThreshold`, значенням якого є 0.95, тобто відкидаються усі області, в яких рівень впевненості мережі у правильній ідентифікації обличчя менші за 95%. Варто також зазначити, що задля оптимізації роботи застосунку на кожному зображенні виділяється тільки одне обличчя, яке займає найбільшу площу (очікується, що в такому випадку людина знаходиться найближче до камери). На рисунку 3.2 зображено повний код функції для знаходження облич.

```

cv::Rect findFace(const cv::Mat& img)
{
    extern cv::dnn::Net net;

    cv::Mat inputBlob = cv::dnn::blobFromImage(img, 1.0, cv::Size(300, 300), cv::Scalar(104.0, 177.0, 123.0));

    net.setInput(inputBlob, "data");
    cv::Mat detection = net.forward("detection_out");

    cv::Mat detectionMat(detection.size[2], detection.size[3], CV_32F, detection.ptr<float>());

    float confidenceThreshold = 0.95f;

    cv::Rect faceArea;
    double maxArea = 0.0;
    for (int i = 0; i < detectionMat.rows; i++) {
        float confidence = detectionMat.at<float>(i, 2);

        if (confidence > confidenceThreshold) {
            int x1 = static_cast<int>(detectionMat.at<float>(i, 3) * img.cols);
            int y1 = static_cast<int>(detectionMat.at<float>(i, 4) * img.rows);
            int x2 = static_cast<int>(detectionMat.at<float>(i, 5) * img.cols);
            int y2 = static_cast<int>(detectionMat.at<float>(i, 6) * img.rows);

            if ((x2 - x1) * (y2 - y1) > maxArea) {
                faceArea.x = x1; faceArea.y = y1;
                faceArea.width = x2 - x1; faceArea.height = y2 - y1;
                maxArea = faceArea.area();
            }
        }
    }
}

```

Рисунок 3.2 – Код функції для виділення обличчя на зображенні

Після отримання усіх зображень облич і застосування до них медіанного фільтру для прибирання шуму, користувач вводить номер класу, до якого відносяться усі отримані навчальні зображення, і обирає, додати нові дані до

вже наявних зразків заданого класу (якщо такі є), або повністю оновити їх. В другому випадку всі зразки цього класу видаляються із завантажених старих даних.

Після цього кожне зображення розділяється на три окремих канали і до кожного каналу застосовується функція для вирівнювання гістограми. Далі йде оброблення множини зображень-зразків (кожне з яких поділено на три окремих канали) LBP-оператором і отримання гістограм ознак цих зображень.

Гістограми, що їх отримують в наслідку виокремлення LBP-ознак, можуть бути дуже великими, але при цьому вони також є дуже розрідженими. Такі гістограми займають багато місця в пам'яті і з ними незручно працювати, адже більшість комірок гістограми є нульовими. Тому для роботи з гістограмами використовується зручний тип даних, який надається бібліотекою OpenCV, – SparseMat (розріджені матриці). Тип SparseMat зберігає тільки ненульові значення, тому не потрібно здійснювати зайвих дій при проходженні по нульових значеннях.

Процес отримання гістограми LBP-ознак із зображення поділяється на два основні етапи. Спочатку усе зображення оброблюється двійковим LBP-оператором і перетворюється на матрицю текстурних ознак лиця. Оскільки у програмі реалізовано розширений LBP-оператор, тобто сусіди знаходяться на певній однаковій відстані від центру оператора (творці методу визначили, що оптимальним оператором в більшості випадків є  $LBP_{8,2}$ , тобто 8 сусідів на відстані 2 від центру), значення сусідніх пікселів обчислюється методом лінійної інтерполяції (рис. 3.3).

Отримана матриця ознак ділиться по регіонах і для кожного регіону обчислюється своя частина загальної гістограми обличчя. Спосіб, в який обмежується врахування в гістограмі тільки рівномірних візерунків, тобто тих, в яких не більше двох двійкових переходів, зазначено на рисунку 3.4. Після побудови загальної гістограми вона нормалізується шляхом поділу кожного елемента на максимальне значення в гістограмі.

```

void elbp(Mat& src, Mat& dst, int radius, int neighbors) {
    dst.create(src.rows - 2 * radius, src.cols - 2 * radius, CV_32SC1);
    dst.setTo(0);

    for (int n = 0; n < neighbors; n++) {
        // отримання точних координат чергового сусіда
        float x = static_cast<float>(radius * cos(2.0 * CV_PI * n / static_cast<float>(neighbors)));
        float y = static_cast<float>(-radius * sin(2.0 * CV_PI * n / static_cast<float>(neighbors)));
        // отримання відносних індексів
        int fx = static_cast<int>(floor(x));
        int fy = static_cast<int>(floor(y));
        int cx = static_cast<int>(ceil(x));
        int cy = static_cast<int>(ceil(y));
        // обчислення дрібкової частини
        float ty = y - fy;
        float tx = x - fx;
        // встановлення ваг інтерполяції
        float w1 = (1 - tx) * (1 - ty);
        float w2 = tx * (1 - ty);
        float w3 = (1 - tx) * ty;
        float w4 = tx * ty;

        for (int i = radius; i < src.rows - radius; i++) {
            for (int j = radius; j < src.cols - radius; j++) {
                // обчислення інтерпольованого значення
                float t = static_cast<float>(w1 * src.at<uchar>(i + fy, j + fx) +
                    w2 * src.at<uchar>(i + fy, j + cx) +
                    w3 * src.at<uchar>(i + cy, j + fx) +
                    w4 * src.at<uchar>(i + cy, j + cx));
                // визначення бінарного коду сусіда
                dst.at<int>(i - radius, j - radius) += ((t > src.at<uchar>(i, j)) ||
                    (std::abs(t - src.at<uchar>(i, j)) < std::numeric_limits<float>::epsilon())) << n;
            }
        }
    }
}

```

Рисунок 3.3 – Функція побудови матриці текстурних ознак

```

val = dst.at<int>(r, c);
v = (val ^ (val >> 1)) & 0x7F;
v &= v - 1; v &= v - 1;
if (v > 0) continue;

```

Рисунок 3.4 – Фільтрування рівномірних візерунків

Після сформування гістограм для кожного навчального зображення обличчя ці дані додаються до старих даних і разом записуються у двійковий файл. У двійковому файлі зберігаються тільки ненульові елементи гістограм, завдяки чому досягається значна економія пам'яті.

В режимі розпізнавання є можливість або розпізнавати обличчя в реальному часі, отримуючи потік кадрів з камери, або вказати шлях до якоїсь фотографії, обличчя на якій необхідно класифікувати. Загалом, процес розпізнавання здебільшого складається з тих же кроків, що й навчання, але

після створення гистограми зображення, що є об'єктом класифікації, до неї застосовується класифікатор, який алгоритмом найближчого сусіда знаходить найменшу відстань між двома класами. Код функції-класифікатора зображено на рисунку 3.5.

```
int NNClassifier(cv::SparseMat& hist, std::vector<cv::SparseMat>& data,
               std::vector<int>& classes, double& confidence)
{
    double minDist = 1000000.0;
    int classNum = 0;
    double dist;
    for (int i = 0; i < data.size(); i++) {
        dist = cv::compareHist(hist, data[i], cv::HISTCMP_CHISQR_ALT);
        if (dist < minDist) {
            minDist = dist;
            classNum = classes[i];
        }
    }
    confidence = minDist;
    return classNum;
}
```

Рисунок 3.5 – Функція-класифікатор

У програмі також використовується поріг, що має значення 10, за яким визначається достовірність розпізнання. Якщо відстань між найближчими сусідами більша за цей поріг, то обличчя визначається як невідоме. Розмір порогу було обрано ґрунтуючись на суб'єктивних спостереженнях в процесі розроблення застосунку і в перспективі може бути дещо змінено як в меншу, так і в більшу сторону.

### 3.3 Тестування розробленого застосунку та аналіз результатів

У тестуванні ефективності застосунку загалом взяло участь 10 людей. Обличчя кожного було проскановано методом відео-захоплення тридцяти кадрів, тобто всього в модель було завантажено 300 еталонних зразків облич.

Процес навчання для кожного обличчя відбувався в різних умовах. Декілька знімків процесу розпізнавання наведено в додатку А.

Перше, на що хотілося б звернути увагу, при такій кількості еталонних зразків у моделі процес розпізнавання відбувався дуже швидко, жодного разу не було помічено падіння частоти кадрів відеопотоку, при тому що розпізнавання відбувалося на різних комп'ютерах, як більш, так і менш потужних.

Також, ще під час навчання була перевірена здатність системи визначати, що людини немає в моделі. На рисунку 3.6 наведено приклад результату роботи системи в тому випадку, коли здійснюється спроба розпізнати людину, якої немає в навченій моделі.

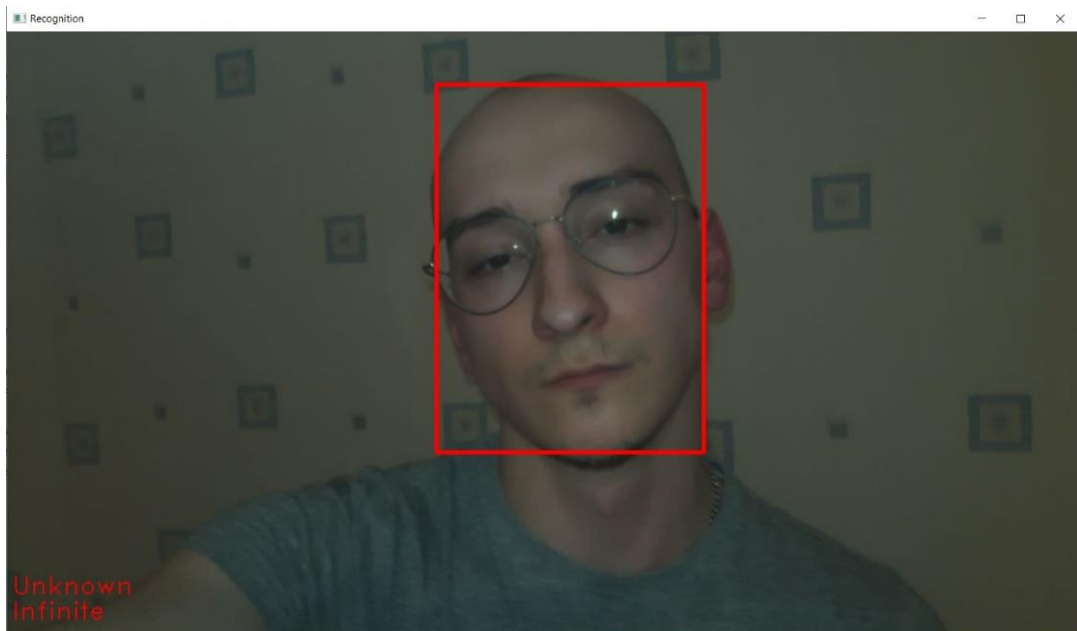


Рисунок 3.6 – Приклад розпізнавання невідомої для системи людини

З наведеного прикладу видно, що коли система отримує на вхід невідоме раніше обличчя, вона правильно визначає, що це невідома особа. Після сканування обличчя, застосунок вірно розпізнає цю особу (рис. 3.7).

Якщо казати про загальні результати роботи, то, проаналізувавши процес випробування застосунку, можна сказати наступне. В тих випадках, коли розпізнавання обличчя відбувалося в тих самих умовах, в яких воно знаходилося під час навчання, застосунок практично завжди вірно визначав

людину в кадрі навіть з різними виразами обличчя та поворотом голови й лише в деяких кадрах давав невірні відповіді. Проте, коли значно змінювались умови освітлення, то точність розпізнання сильно падала й система здебільшого видавала хибні результати, неправильно визначаючи особу в кадрі.

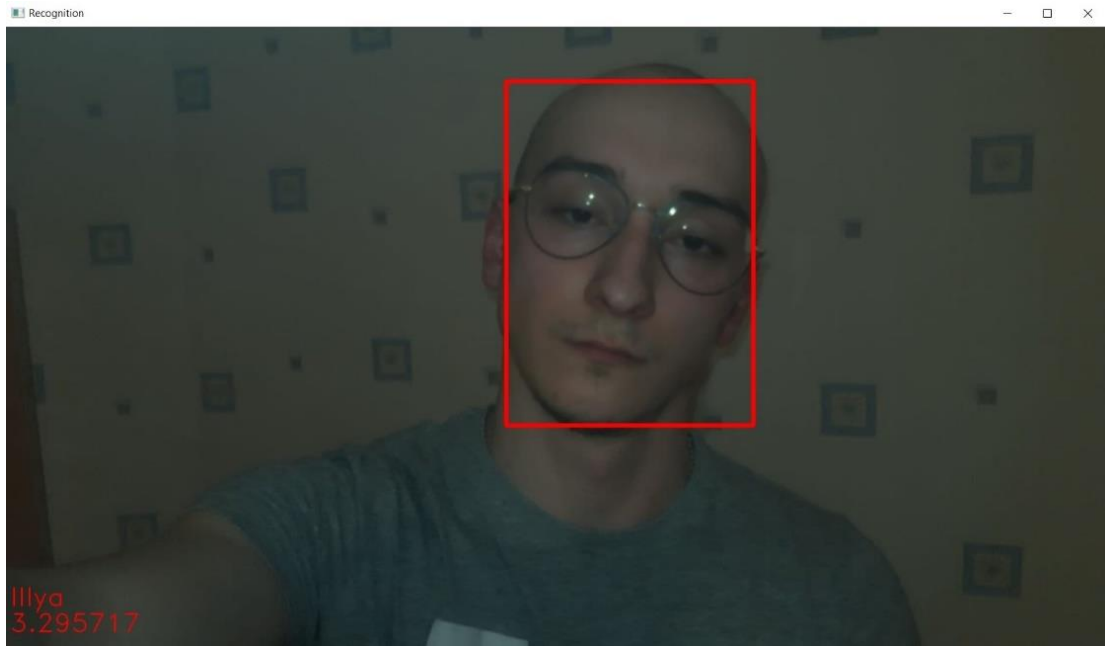


Рисунок 3.7 – Приклад розпізнання людини після сканування її обличчя

Для дослідження цієї проблеми було проведено додаткове випробування, для якого модель була перенавчена на зображеннях облич меншої кількості людей (5 осіб), які під час навчання знаходилися в однакових умовах, з однаковим тлом та освітленням. Під час тестування дана модель показала неочікувано добрі результати. Вона вірно, хоч і з меншою впевненістю, розпізнавала обличчя за умови сильної зміни освітлення та навіть з такою значною оклюзією, як медична маска (рис. 3.8).

З цього можна зробити важливий висновок, що детектор ознак LBP, який було реалізовано, дуже чутливий до умов освітлення, тому навчання для всіх обличч необхідно виконувати в контрольованих умовах, з однаковим освітленням та рівномірним (бажано світлим) тлом. Однакові умови отримання навчальних зображень дозволять максимально зменшити вплив

освітлення і невеликих фрагментів фону, що потрапляють в область з обличчям, що сильно впливає на кінцеві результати в цьому алгоритмі. Краще дослідити це дозволить масштабніше тестування, з більшою кількістю осіб та різним числом зразків на кожну особу.

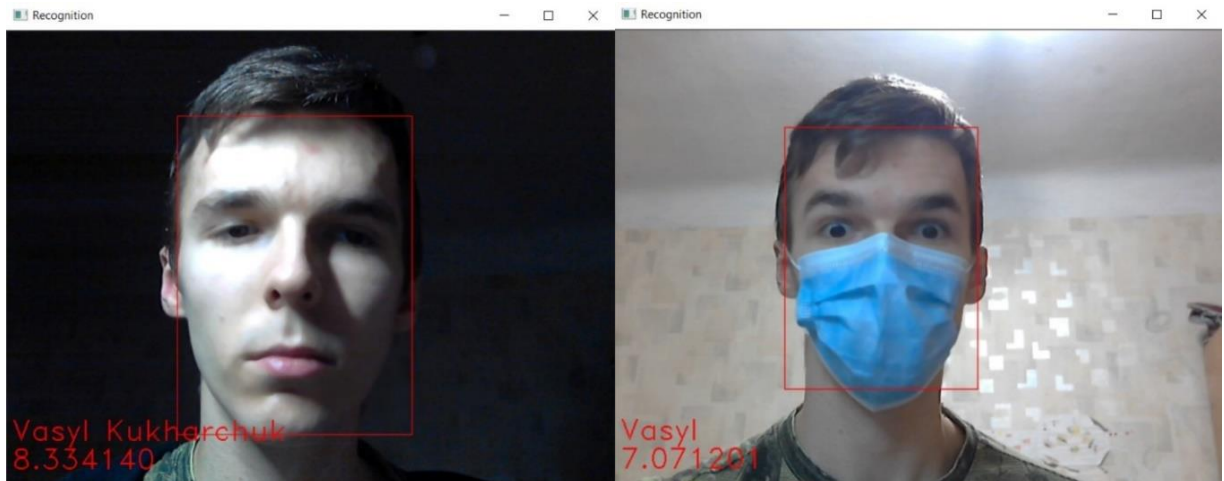


Рисунок 3.8 – Розпізнання обличчя в темряві та з маскою

### 3.4 Перспективи дальшої роботи

Розроблений застосунок задовільно виконує поставлену задачу, модель може навчатися на еталонних зображеннях, отриманих різними шляхами з різною роздільною здатністю, після чого здатна розпізнавати обличчя в реальному часі. Для кожної людини можна в будь-який момент додати нові навчальні дані, що покращує якість роботи програми.

Тестування застосунку показало непогані результати, але й визначило слабкі місця, для дослідження й, в перспективі, усунення яких необхідно проводити масштабніші випробовування.

Отримані результати показують, що реалізований метод в цілому є перспективним і заслуговує на дальший розвиток. Підхід визначення текстурних ознак обличчя дає стійкі ознаки для класифікації і багато можливостей для розширення. Серед можливих покращень розробленої

методики можуть бути ще значніші оптимізації роботи з гістограмами, зокрема використання ймовірнісних характеристик та векторного квантування, введення вектору ваг для областей, на які поділяються зображення, або заміна методу LBP на його розвиненіші версії, наприклад BSIF (бінаризовані статистичні особливості зображення). Окрім того, можливо оптимізувати процес класифікації, виконуючи кластеризацію для зразків кожного класу, або й взагалі спробувати інші методи класифікації.

## ВИСНОВКИ

У рамках кваліфікаційної роботи був розроблений і реалізований застосунок для розпізнання облич на кадрах відеозйомки.

Було виконано всі поставлені задачі, а саме:

- досліджено методи визначення ключових точок на зображеннях, розібрано поняття ключових точок та кутів, надано огляд деяких найвідоміших методів знаходження ключових точок, визначено їх переваги та недоліки, розглянуто проблеми, пов'язані з визначенням ключових точок;

- досліджено методи класифікації зображень на підставі статистичних характеристик, розглянуто декілька різноманітних класифікаторів, їх переваги й недоліки, також описано класифікатор, що ґрунтується на аналізі ймовірнісних характеристиках;

- проаналізовано наявні двовимірні підходи до розпізнавання обличчя, визначено чотири основні підходи до розпізнавання облич, надано загальний опис декількох алгоритмів, що належать кожному з підходів;

- розроблено методику ідентифікації облич на кадрах відеозйомки, в основі якої лежить застосування місцевих двійкових візерунків, визначено етапи розробки застосунку, обрано оптимальні підходи до виконання кожного з етапів;

- визначено необхідні інструментальні засоби для створення застосунку;

- виконано всі етапи розроблення застосунку для ідентифікації облич на кадрах відеозйомки, на кожному етапі здійснено різноманітні програмні й алгоритмічні оптимізації, реалізовано досить гнучкий метод роботи застосунку та його взаємодії з користувачем;

- проведено тестування розробленого застосунку, за результатами якого було визначено особливості роботи реалізованого методу, умови використання та його ефективність, яка виявилася на даному етапі

задовільною, але водночас надавала розуміння необхідності ширших випробувань та можливого доопрацювання;

– визначено перспективи дальшої роботи, які виходять з отриманих результатів тестування та теоретичних знань в досліджуваній області.

Мету роботи досягнуто.

Результати роботи апробовано у вигляді 2 тез доповідей під час Міжнародного молодіжного форуму «РАДІОЕЛЕКТРОНІКА І МОЛОДЬ У XXI СТОЛІТТІ» та XXVI Міжнародної науково-практичної конференції «Topical issues of practice and science» [53, 54].

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Pagnin, E., & Mitrokotsa, A. (2017). Privacy-preserving biometric authentication: challenges and directions. *Security and Communication Networks*, 2017.
2. Daradkeh Y.I., Tvoroshenko I., Gorokhovatskyi V., Latiff L.A., and Ahmad N. Development of Effective Methods for Structural Image Recognition Using the Principles of Data Granulation and Apparatus of Fuzzy Logic. *IEEE Access*. 2021. Vol. 9. pp. 13417-13428.
3. Benzaoui, A., Bourouba, H., & Boukrouche, A. (2012, October). System for automatic faces detection. In *2012 3rd International Conference on Image Processing Theory, Tools and Applications (IPTA)* (pp. 354-358). IEEE.
4. Phillips, P. J., Yates, A. N., Hu, Y., Hahn, C. A., Noyes, E., Jackson, K., ... & O'Toole, A. J. (2018). Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. *Proceedings of the National Academy of Sciences*, 115(24), 6171-6176.
5. Gorokhovatskyi V.O., Tvoroshenko I.S., and Peredrii O.O. Image classification method modification based on model of logic processing of bit description weights vector. *Telecommunications and Radio Engineering*. 2020. Vol. 79(1), pp. 59-69.
6. Darabant, A. S., Borza, D., & Danescu, R. (2021). Recognizing Human Races through Machine Learning – A Multi-Network, Multi-Features Study. *Mathematics*, 9(2), 195. doi:10.3390/math9020195
7. Adjabi, I., Ouahabi, A., Benzaoui, A., & Taleb-Ahmed, A. (2020). Past, Present, and Future of Face Recognition: A Review. *Electronics*, 9(8), 1188. doi:10.3390/electronics9081188
8. Min, R., Kose, N., & Dugelay, J. L. (2014). Kinectfacedb: A kinect database for face recognition. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 44(11), 1534-1548.

9. Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1701-1708).
10. Ming, Z., Visani, M., Luqman, M. M., & Burie, J.-C. (2020). A Survey on Anti-Spoofing Methods for Facial Recognition with RGB Cameras of Generic Consumer Devices. *Journal of Imaging*, 6(12), 139. doi:10.3390/jimaging6120139
11. Adjabi, I., Ouahabi, A., Benzaoui, A., & Jacques, S. (2021). Multi-Block Color-Binarized Statistical Images for Single-Sample Face Recognition. *Sensors*, 21(3), 728. doi:10.3390/s21030728
12. Jang, J., Yoon, H.-S., & Kim, J. (2021). Improvement of Identity Recognition with Occlusion Detection-Based Feature Selection. *Electronics*, 10(2), 167. doi:10.3390/electronics10020167
13. Ruan, S., Tang, C., Zhou, X., Jin, Z., Chen, S., Wen, H., ... Tang, D. (2020). Multi-Pose Face Recognition Based on Deep Learning in Unconstrained Scene. *Applied Sciences*, 10(13), 4669. doi:10.3390/app10134669
14. Zhuang, W., Chen, L., Hong, C., Liang, Y., & Wu, K. (2019). FT-GAN: Face Transformation with Key Points Alignment for Pose-Invariant Face Recognition. *Electronics*, 8(7), 807. doi:10.3390/electronics8070807
15. Bodini, M. (2019). A Review of Facial Landmark Extraction in 2D Images and Videos Using Deep Learning. *Big Data and Cognitive Computing*, 3(1), 14. doi:10.3390/bdcc3010014
16. Basaran, E., Gökmen, M., & Kamasak, M. (2018). An Efficient Multiscale Scheme Using Local Zernike Moments for Face Recognition. *Applied Sciences*, 8(5), 827. doi:10.3390/app8050827
17. Wu, M., Wei, D., Zhang, L., & Zhao, Y. (2018). Hyperspectral Face Recognition with Patch-Based Low Rank Tensor Decomposition and PFFT Algorithm. *Symmetry*, 10(12), 714. doi:10.3390/sym10120714
18. Tvoroshenko I.S., and Gorokhovatsky V.O. (2019) Intelligent classification of biophysical system states using fuzzy interval logic, *Telecommunications and Radio Engineering*, 78(14), pp. 1303-1315.

19. Tvoroshenko I., and Zarivchatskyi R. (2020) Analysis of existing methods for searching object in the video stream, Abstracts of VI International Scientific and Practical Conference «About the problems of science and practice, tasks and ways to solve them» (October 26-30, 2020). Milan, Italy, pp. 500-505.

20. Tvoroshenko I.S., and Gorokhovatsky V.O. (2019) Modification of the branch and bound method to determine the extremes of membership functions in fuzzy intelligent systems, *Telecommunications and Radio Engineering*, 78(20), pp. 1857-1868.

21. Tvoroshenko I., and Tkachenko D. (2020) Mechanisms of image classification based on descriptors of local features, *Abstracts of IV International Scientific and Practical Conference «Integration of scientific bases into practice» (October 12-16, 2020). Stockholm, Sweden*, pp. 443-448.

22. Гороховатський В.О., Творошенко І.С. Методи інтелектуального аналізу та оброблення даних: навч. посібник. Харків: ХНУРЕ, 2021. 92 с.

23. Кобилін О.А., Творошенко І.С. Методи цифрової обробки зображень: навч. посібник. Харків: ХНУРЕ, 2021. 124 с.

24. Harris, C. G., & Stephens, M. (1988, August). A combined corner and edge detector. In *Alvey vision conference (Vol. 15, No. 50, pp. 10-5244)*.

25. Shi, J. (1994, June). Good features to track. In *1994 Proceedings of IEEE conference on computer vision and pattern recognition (pp. 593-600)*. IEEE.

26. Rosten, E., & Drummond, T. (2006, May). Machine learning for high-speed corner detection. In *European conference on computer vision (pp. 430-443)*. Springer, Berlin, Heidelberg.

27. Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2), 91-110.

28. Tvoroshenko I.S., and Gorokhovatsky V.O. (2020) Effective tuning of membership function parameters in fuzzy systems based on multi-valued interval logic, *Telecommunications and Radio Engineering*, 79(2), pp. 149-163.

29. Tvoroshenko I., and Dziubenko M. (2020) Modern methods of analysis of the movement scheme using video detection of vehicles, *Abstracts of V International Scientific and Practical Conference «Study of modern problems of civilization» (October 19-23, 2020). Oslo, Norway*, pp. 422-428.

30. Bay, H., Tuytelaars, T., & Van Gool, L. (2006, May). Surf: Speeded up robust features. In *European conference on computer vision* (pp. 404-417). Springer, Berlin, Heidelberg.

31. Leutenegger, S., Chli, M., & Siegwart, R. Y. (2011, November). BRISK: Binary robust invariant scalable keypoints. In *2011 International conference on computer vision* (pp. 2548-2555). Ieee.

32. Кучеренко, Е. И., Корниловский, А. В., Творошенко, И. С. (2010) О методах настройки функций принадлежности в нечетких системах. *Системы управления, навигации и связи*, (1), 13.

33. I. Tvoroshenko Information technologies for decision-making on the conditions of spatially distributed objects, in *Abstracts of I International Scientific and Practical Conference. Problems and perspectives of modern science and practice*, Austria, 2020. pp. 45-50.

34. Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.

35. Творошенко, И. С., Дехтярь, А. П. (2005, June) Информационные технологии в задачах компьютерной диагностики с использованием интеллектуальных систем. In *Клиническая информатика и Телемедицина. Компьютерная Медицина–2005: материалы междунар. научн.-технич. конф., Харьков* (p. 138).

36. Кучеренко, Е. И., Творошенко, И. С. (2003) Процессы принятия решений в сложных системах на основе нечетких интервальных представлений. *Вісник Національного технічного університету «ХПІ». Тематичний випуск: Системний аналіз, управління та інформаційні технології. – X.: НТУ «ХПІ», 1(7), 79-86.*

37. Kucherenko, Y. I., Filatov, V. A., Tvoroshenko, I. S., & Baidan, R. N. (2005). Intellectual technologies in decision-making technological complexes based on fuzzy interval logic. *East European Journal of Advanced Technologies*, 2, 92-96.
38. Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
39. Tvoroshenko I.S., and Gorokhovatsky V.O. (2019) Modification of the branch and bound method to determine the extremes of membership functions in fuzzy intelligent systems, *Telecommunications and Radio Engineering*, 78(20), pp. 1857-1868.
40. Tvoroshenko Iryna (2019) Development of models of spatial analysis of status of interactive processes of complex systems. International Scientific Conference *Scientific Development of New Eastern Europe: Conference Proceedings* (Riga, April 6th, 2019). Latvia: Baltija Publishing. Part II. pp. 152-156.
41. Gorokhovatskyi V.O., Tvoroshenko I.S., and Vlasenko N.V. (2020) Using fuzzy clustering in structural methods of image classification, *Telecommunications and Radio Engineering*, 79(9), pp. 781-791.
42. Творошенко І.С., Зеленський М.О. Дослідження гібридних методів для класифікації складноструктурованих зображень. *Сучасний рух науки: тези доповідей VIII міжнародної науково-практичної інтернет-конференції* (Дніпро, 3–4 жовтня 2019 р.). Дніпро, 2019. Т. 3. С. 382-387.
43. Кучеренко, Є. І., Творошенко, І. С. (2011) Оперативне оцінювання простору станів складних розподілених об'єктів з використанням нечіткої інтервальної логіки. *Искусственный интеллект*. 2011. № 3. С. 382-387.
44. Gorokhovatskyi V., and Tvoroshenko I. (2020) Image Classification Based on the Kohonen Network and the Data Space Modification, *In CEUR Workshop Proceedings: Computer Modeling and Intelligent Systems (CMIS-2020)*, 2608, pp. 1013-1026.
45. Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05) (Vol. 1, pp. 886-893). Ieee.

46. Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of cognitive neuroscience*, 3(1), 71-86.

47. Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. (1997). Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on pattern analysis and machine intelligence*, 19(7), 711-720.

48. Бодянский, Е. В., Кучеренко, Е. И., Творошенко, И. С. (2004). О синтезе нечетких алгоритмов на основе композиции фрагментов правил и моделей. *АСУ и приборы автоматки*, (128), 19-28.

49. Huang, Z. H., Li, W. J., Shang, J., Wang, J., & Zhang, T. (2015). Non-uniform patch based face recognition via 2D-DWT. *Image and Vision Computing*, 37, 12-19.

50. Kobylin O., Gorokhovatskyi V., Tvoroshenko I., and Peredrii O. (2020) The application of non-parametric statistics methods in image classifiers based on structural description components, *Telecommunications and Radio Engineering*, 79(10), pp. 855-863.

51. Gorokhovatskyi V.O., Tvoroshenko I.S., and Peredrii O.O. (2020) Image classification method modification based on model of logic processing of bit description weights vector, *Telecommunications and Radio Engineering*, 79(1), pp. 59-69.

52. Ahonen, T., Hadid, A., & Pietikäinen, M. (2004, May). Face recognition with local binary patterns. In *European conference on computer vision* (pp. 469-481). Springer, Berlin, Heidelberg.

53. Кухарчук В.А. До питання аналізу сучасного стану розвитку застосунків для ідентифікації особистости на зображеннях та кадрах відеозйомки. *Радіоелектроніка та молодь у XXI столітті: тези доповідей 25-го Міжнародного молодіжного форуму (Харків, 20–21 квітня 2021 р.)*. Харків: ХНУРЕ, 2021. Т. 7, 10. С. 28-29.

54. Tvoroshenko I., and Kukharchuk V. (2021) Current state of development of applications for recognition of faces in the image and frames of video captures, *Abstracts of XXVI International Scientific and Practical Conference «Topical issues of practice and science»* (May 18-21, 2021). London, Great Britain, pp. 685-690.