

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Інформаційно-аналітичних технологій та менеджменту
(повна назва)

Кафедра Інформатики
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти перший (бакалаврський)

МОДЕЛЮВАННЯ МЕТОДУ ДЛЯ СТИСНЕННЯ ОПИСУ ЯК
МНОЖИНИ ДЕСКРИПТОРІВ У СТРУКТУРНИХ МЕТОДАХ
КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ
(тема)

Виконав:
студент 4 курсу, групи ІТІНФ-18-1

Чефранов І.О.
(прізвище, ініціали)

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

Тип програми освітньо-професійна

Освітня програма Інформатика
(повна назва освітньої програми)

Керівник проф. Гороховатський В.О.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____
(підпис)

Кобилін О.А.
(прізвище, ініціали)

2022 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)

« ____ » _____ 2022 р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**студентові Чефранову Івану Олеговичу
(прізвище, ім'я, по батькові)1. Тема роботи Моделювання методу для стиснення опису як множини дескрипторів у структурних методах класифікації зображеньзатверджена наказом по університету від 16 травня 2022 року № 541Ст2. Термін подання студентом роботи до екзаменаційної комісії 24 травня 2022 р.3. Вихідні дані до роботи Науково-методична та науково-технічна література, матеріали конференцій, дані інтернет-мережі, алгоритми визначення ключових ознак зображень, теоретичні відомості про методи класифікації зображень, відомості про методи знаходження дескрипторів ключових точок, теоретичні відомості про способи редукції множини дескрипторів ключових точок, тестові зображення, середовище розробки PyCharm 2018, відкрита програмна бібліотека для Python.

4. Перелік питань, що потрібно опрацювати в роботі _____

1. Огляд методів розпізнавання зображень за допомогою ключових точок.2. Аналіз методу BRISK пошуку ключових точок на зображенні.3. Математична модель просторової обробки структурних описів.4. Застосування редукції ознак для прискорення класифікації.5. Порівняння якості роботи традиційного методу та структурного методу класифікації.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п. 5 включається до завдання за рішенням випускової кафедри) Актуальність проблеми, мета роботи, постановка задачі, етапи виконання роботи, візуалізація ключових точок, опис методу, опис і аналіз результатів експериментів, висновки.

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

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

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	18.04.2022	
2	Аналіз завдання, огляд літератури	18.04.22–21.04.22	
3	Аналіз літератури з досліджуваної проблеми	22.04.22–25.04.22	
4	Аналіз підходів до розпізнавання та класифікації зображень	26.04.22–03.05.22	
5	Розробка алгоритму	04.05.22–14.05.22	
6	Програмна реалізація	15.05.22–23.05.22	
7	Оформлення пояснювальної записки	24.05.22–26.05.22	
8	Перевірка на плагіат	27.05.22	
9	Рецензування	28.05.22	
10	Підготовка презентації та доповіді	29.05.22–30.05.22	
11	Занесення роботи в електронний архів	31.05.22	
12	Попередній захист кваліфікаційної роботи	01.06.22	

Дата видачі завдання 18 квітня 2022 р.

Студент _____
(підпис)

Керівник роботи _____ проф. Гороховатський В.О.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ/ABSTRACT

Пояснювальна записка до кваліфікаційної роботи: 58 с., 2 табл., 5 рис., 50 джерел.

ПЕРЕТВОРЕННЯ РАДОНА, ПЕРЕТВОРЕННЯ ХАФА, ДЕТЕКТУВАННЯ ЛІНІЙ, ВІДРІЗКІВ, ІНТЕГРАЛЬНЕ ПЕРЕТВОРЕННЯ, ІНВАРІАНТНІ ПЕРЕТВОРЕННЯ, РОБАСТНІ ОЦІНЮВАЧІ.

Об'єктом роботи є структурні методи класифікації зображень.

Метою роботи є підвищення результативності структурних методів класифікації зображень шляхом впровадження засобів стиснення опису.

Використано методи числового моделювання та аналітичного обґрунтування. Проведено дослідження методів пошуку зображення у базах даних та методів погодження різноракурсних зображень однієї сцени, детектування відрізків ліній на основі перетворень Радона і Хафа, та аналіз робастних методів детектування відрізків ліній. Досліджено метод детектування відрізків ліній за допомогою перетворення Радона, розроблено алгоритм детектування відрізків ліній.

У результаті роботи здійснена програмна реалізація системи для пошуку в базах даних.

RADON TRANSFORM, TRANSFORM HOUGH, DETECTION LINE, SEGMENT, INTEGRAL TRANSFORMATION INVARIANT TRANSFORMATION, ROBUST APPRAISERS.

The object of the work is the sequence of multi-point image objects.

The aim of the work is to develop methods based on the use of Radon transforms, which allow to detect features (segments of lines) and to form feature vectors for search in databases.

Methods used numerical simulations and analytical reasoning. Research methods of finding images in databases and methods of coordination image of one scene detection line segments based on Hough and Radon transformation, and analysis of robust methods for detecting line segments. The method of detecting line segments using the Radon transform, the algorithm detecting line segments.

As a result of implemented software implementation of the system to search the database.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів	6
Вступ.....	7
1 Аналіз сучасних методів розпізнавання зображень	9
1.1 Підходи до задач розпізнавання.....	9
1.2 Структурний метод розпізнавання зображень.....	13
1.3 Методи формування ключових точок.....	16
1.4 Постановка задачі	26
2 Метод редукції для множини дескрипторів	28
2.1 Метричні та статистичні критерії інформативності.....	28
2.2 Застосування метричних критеріїв інформативності для множини дескрипторів	30
2.3 Ефективність сукупної інформативності	35
3 Результати комп'ютерного моделювання методу класифікації зображень	39
3.1 Огляд програмних засобів реалізації	39
3.2 Результати комп'ютерного моделювання	42
Висновки	51
Перелік джерел посилання	53

**ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ,
СКОРОЧЕНЬ І ТЕРМІНІВ**

КЗ – комп'ютерний зір

КТ – ключові точки

РО – розпізнавання образів

СЕ – структурні елементи

СМ – структурний метод

ХО – характерні ознаки

CNN – згорткові нейронні мережі

DNN – глибокі нейронні мережі

SIFT – Scale-Invariant Feature Transform

ВСТУП

Для сучасності характерним є швидкий розвиток методів та засобів отримання, обробки, передачі та розпізнавання об'ємних потоків інформації в реальному часі з необхідною точністю та продуктивністю. При цьому значна увага приділяється задачам оброблення та аналізу зображень, адже саме візуальні образи є носієм найбільшого об'єму інформації, яка може бути використана в різних галузях діяльності людини. Саме тому зростає актуальність розробки систем, моделей, методів та засобів обробки зображень, які давали б можливість автоматизації та підвищення точності і швидкодії вказаних процесів [1].

На даний момент комп'ютерний зір є одним із наймолодших та одним з найперспективніших напрямків у сфері штучного інтелекту, а найважливішими його задачами є [2]:

- розпізнавання образів;
- перевірка наявності об'єкта на зображенні;
- локалізація об'єктів (визначення координат);
- відстеження об'єкта у відеопотоці;
- ідентифікація об'єкту;
- розпізнавання інформації, що містить об'єкт на зображенні;
- комбінація або суперпозиція задач для комплексних систем.

Розпізнавання образів – це одна з найбільш фундаментальних задач в галузі інтелектуальних систем та штучного інтелекту, оскільки принципове значення проблеми тісно пов'язане з питанням про те, що може і що не може робити машина чи комп'ютер [3]. З прикладної точки зору вирішення задачі розпізнавання образів дозволяє автоматизувати процеси, які до сьогодні пов'язували лише з діяльністю людини. Сучасна наука пов'язує розвиток інтелекту людини переважно зі слуховим і зоровим сприйняттям інформації. З урахуванням того, що одним з найбільш важливих джерел інформації є

зорове сприйняття, до основних задач, що вирішуються в сучасних інформаційних системах, можна віднести задачі розпізнавання та ідентифікації об'єктів у системах технічного зору, а також розпізнавання літер, цифр, слів або фраз, ідентифікація за фотографією або відбитками пальців, розпізнавання жестів, виявлення цілей.

Задача розпізнавання образів полягає у віднесенні вихідних даних до певного класу за допомогою виділення суттєвих ознак, що характеризують ці дані, із загальної маси несуттєвих даних. Одним з методів розпізнавання (класифікації) зображень є структурний метод, використання якого дозволяє виокремити суттєві та несуттєві ознаки окремих компонентів об'єкта та прийняти рішення за допомогою статистичної обробки компонентних даних [4 – 10].

У роботі здійснено модифікацію методу класифікації зображень на основі аналізу інформативності та оброблення опису як множини дескрипторів ключових точок з метою пришвидшення процесу оброблення даних.

1 АНАЛІЗ СУЧАСНИХ МЕТОДІВ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

1.1 Підходи до задач розпізнавання

Розпізнавання або класифікація образів – це задача, яка весь час вирішуються у професійній та побутовій діяльності людини. Прикладами таких задач можуть бути розпізнавання літер, штрих-кодів на касі, автомобільних номерів, осіб та інших біометричних даних, мови, зображень, локальних ділянок земної кори, багатих на корисні копалини.

На виробництві часто виникає потреба автоматизувати контроль якості деталей. Задача полягає в тому, щоб виявити, чи є деталь дефектною, чи ні. Якщо ж з'ясується, що деталь має дефект, час то потрібно визначити тип цього дефекту.

Системи розпізнавання часто використовуються у медичній практиці. Найтипівіша ситуація полягає в тому, що ті чи інші захворювання діагностуються на основі аналізу кардіограм, рентгенівських знімків і т. п.

Проблема розпізнавання літер має велике значення для комп'ютерних технологій. Системи розпізнавання літер працюють разом зі сканерами – пристроями, які використовуються для введення до комп'ютера друкованих зображень і текстів. При введенні друкованого тексту сканер формує лише графічне зображення; для того щоб створити текстовий документ, з яким може працювати текстовий редактор, необхідно впізнати на цьому зображенні окремі літери. Аналогічно розпізнавання літер є необхідним для підтримки пристроїв рукописного введення. Цими пристроями, зовні схожими на звичайну авторучку, часто комплектуються надпортативні комп'ютери (персональні помічники). Основна мета цих пристроїв – замінити введення з клавіатури, що є незручним для багатьох користувачів [11-14].

Сьогодні інтенсивно розвиваються технології розпізнавання мови, пов'язані, по-перше, з голосовим керуванням комп'ютером, а по-друге – з введенням текстів з голосу.

Застосування методів розпізнавання в робототехніці є абсолютно природним і необхідним, оскільки роботи повинні безпосередньо сприймати зовнішній світ і, відповідно, мати пристрої машинного зору.

Застосування методів розпізнавання в охоронних системах пов'язано в першу чергу з проблемою ідентифікації. Наприклад, потрібно ідентифікувати певну особу, щоб визначити, чи має вона право входити на територію, що охороняється. Розвиваються також системи, які вирішують проблему ідентифікації відбитків пальців тощо.

Питання розпізнавання зображень є одним із важливіших питань у розвитку інтелектуальних систем обробки інформації. У сучасній літературі стосовно цього питання використовують два терміни, що вважають синонімічними: розпізнавання та класифікація [12].

Серед задач класифікації зображень можна виділити основні три етапи: ідентифікація об'єкта, віднесення його до певного класу та кластерний аналіз об'єкта.

Під ідентифікацією розуміється виділення конкретного об'єкта із йому подібних за системою визначених ознак. Наприклад, знайти конкретну людину. Ідентифікація є задачею найнижчого рівня

Прикладом віднесення об'єкта до того чи іншого класу або класифікація є задача розпізнавання літер або виявлення дефектів у технічній деталі. Це найтипівіша задача класифікації.

Кластерний аналіз полягає в поділі заданого набору об'єктів на класи – групи об'єктів, схожі між собою за тим чи іншим критерієм. До того ж класи початково не задаються, тому така задача називається класифікацією без учителя. У задачі класифікації без учителя вибірка ознак образів $x_j, j=1, \dots, N$, розбивається на підмножини, що не перетинаються (кластери), які

складаються із схожих один на одного об'єктів, до того ж вимагається, щоб об'єкти із різних кластерів істотно відрізнялися один від одного.

Методи класифікації зображень охоплює такі завдання як сегментація – виділення на зображенні об'єктів, опис об'єктів і стиснення їх в зручну для комп'ютерної обробки форму, а також класифікація (розпізнавання) окремих об'єктів. Процес обробки зображення буває низького, середнього та високого рівнів. При обробці зображень на низькому рівні на вході і на виході обробки отримують зображення. Для процесів середнього рівня характерна наявність зображень тільки на вході, на виході – ознаки і атрибути, які добувають із цих зображень, такі як кордони областей, лінії контурів, відмінні ознаки конкретних об'єктів, тощо. Тобто основне завдання розпізнавання образів – отримання семантичного опису зображених об'єктів.

У задачах класифікації використовують детерміновані та не детерміновані підходи.

До недетермінованих методів розпізнавання образів відносяться нейронні мережі. Різні спеціальні нейромережеві архітектури, такі як глибокі нейронні мережі (DNN), згорткові нейронні мережі (CNN), були успішно застосовані в таких областях, як комп'ютерний зір, автоматичне розпізнавання мови, обробка природної мови, біоінформатика.

Детерміновані методи мають набір чітких вирішальних правил [2, 8, 15]. В рамках геометричного підходу вирішальні правила задаються геометричними параметрами контейнерів класів розпізнавання, побудованими в просторі ознак. Якщо два вектори-реалізації образу знаходяться на невеликій відстані один від одного, то спостерігається їх топологічна «подібність», або компактність. Детермінованими методами розпізнавання образів є метод еталонів, лінійні вирішальні правила, метод K -найближчих сусідів, метод потенційних функцій, метод опорних векторів, структурні методи.

Метод еталонів спрямований на усунення модельності методу еталонів шляхом наближення до практичних задач, для яких є характерним перетин

класів. Процес навчання за методом еталонів складається із двох етапів. На першому етапі за навчальною вибіркою для кожного класу будується контейнер у вигляді гіперсфери якомога меншого радіуса, до якої входять усі реалізації класу. При цьому формується еталон кожного класу, обчислюється відстань від еталона до всіх реалізацій цього класу. На другому етапі обирається максимальна із цих відстаней. Відтворюється гіперсфера з центром у еталоні й радіусом, що охоплює всі реалізації свого класу. Аналогічна процедура проводиться для всіх класів із заданого алфавіту.

Процес навчання за методом потенційних функцій полягає в побудові вирішального правила за допомогою потенційної функції. Вибір потенційної функції не є простою задачею, яку доводиться розв'язувати розробнику інформаційного забезпечення систем розпізнавання. Від вибору потенційної функції залежить як збіжність алгоритму, так і його точність. При великих обсягах навчальної вибірки доцільно обчислювати не вирішальне правило, а оцінювати розподільну межу для класів розпізнавання. У цьому випадку для підвищення оперативності розпізнавання достатньо лише визначити, де знаходиться реалізація – справа чи зліва від межі. Перевага цього методу полягає у простоті реалізації, що ґрунтується на фізичних законах електростатики (чим далі від заряду, тим вплив його менший). Недоліками є необхідність обґрунтування вибору потенційної функції, що впливає безпосередньо на збіжність алгоритму навчання та на достовірність розпізнавання в режимі екзамену. До того ж метод носить модельний характер, оскільки орієнтований на класи, що не перетинаються у просторі ознак розпізнавання [16].

Глибокі моделі засновані на розподілених уявленнях. Основним припущенням теорії розподілених представлень є те, що спостережувані дані генеруються за допомогою взаємодії факторів, організованих в шари. Глибоке навчання передбачає, що ці шари факторів відповідають рівням абстракції або композиції. Змінюючи кількість шарів і розмірів шарів, ми можемо досягти різних рівнів абстракції.

Глибоке навчання використовує ідею ієрархічних пояснювальних факторів, де більш високий рівень, більш абстрактні поняття, витягуються з нижніх рівнів. Ці архітектури часто будуються шар за шаром. Глибоке навчання допомагає розплутати ці абстракції і виявити, які функції можуть бути корисні для вивчення.

Важлива перевага глибоких моделей – це те, що для них були розвинені алгоритми навчання без вчителя.

1.2 Структурний метод розпізнавання зображень

Структурний метод (СМ) розпізнавання є детермінованим методом, тобто реалізація образу характеризується структурою, а не множиною значень або їх відношень [3, 17, 18]. Структурний підхід застосовується до задач класифікації зображень, в яких важливою є інформація, що описує структуру кожного об'єкта. При такому підході важливо, щоб об'єкт можна було не тільки віднести до певного класу, а й описати ті властивості об'єкта, яка виключають його із іншого класу. Типовим прикладом таких задач є аналіз сцен (розпізнавання образів). Об'єкти, що розглядаються в таких задачах, доволі складні і мають велике число ознак. Тому складний об'єкт розглядається у вигляді ієрархічної структури простіших під образів.

Класичний структурний підхід полягає в наступному [4]. Обирається система базисних структурних елементів і система операцій, що застосовується до системи елементів (СЕ) з метою створення образів та підобразів, які належать певному класу. Для розпізнавання об'єкта необхідно знайти СЕ, задати або виявити відношення між ними, а потім за допомогою структурного аналізу встановити, чи узгоджується опис з правилом, яке могло його створити. У загальному вигляді структурний опис у задачі розпізнавання можна представити у вигляді:

$$\langle \Upsilon_1, \Upsilon_2, \Upsilon_3 \rangle, \quad (1.1)$$

де Υ_1 – множина твірних;

Υ_2 – система правил для побудови конфігурацій;

Υ_3 – правило встановлення відповідності на множині допустимих конфігурацій.

При розв'язуванні практичних задач класифікації СМ застосовується у поєднанні з іншими методами. Наприклад, при розпізнаванні СЕ ефективними є статистичні методи обробки інформації через те, що виміри характеристик СЕ мають шуми. Для встановлення еквівалентності образів використовують методи теорії графів, кореляційні підходи, перетворення Хафа, методи аналізу множин, грануляції тощо [19-25].

Безпосереднє використання класичного СМ при аналізі зображень ускладнюється наступними моментами:

- для формування остаточної абетки СЕ потрібно розв'язувати задачу розпізнавання на локальному рівні, враховуючи значну кількість фрагментів та їх ознак на практиці;

- об'єкти, зокрема хибні, можуть складатися із близьких елементів, які повторюються в межах одного об'єкта;

- побудова інваріантної структурної моделі об'єкта з урахуванням повного аналізу внутрішніх просторових зв'язків, аналогічних праобразу, має високий рівень обчислювальної складності.

Враховуючи попередні зауваження, природним виходом є введення обмежень, пов'язаних з кінцевою множиною еталонів або їх перетворень, заміна задачі розпізнавання СЕ на задачу детектування, а також спрощення правил за допомогою різних апроксимацій для міри подібності структурних описів.

При використанні статистичного та структурного розпізнавання зображень повний опис об'єкта можна представити у вигляді пари функцій

$f = (f_1, f_2)$, де f_1 відображає параметри, які спостерігаються, f_2 – скриті [5]. Специфіка структурного аналізу у такому випадку полягає у розробці засобів для представлення f_2 . Отже структурний аналіз задіяний у тому випадку, коли складна функція із широкою областю визначення задається за допомогою ряду простих функцій менших областей визначення. З математичної точки зору зведення складної функції до системи простих не завжди можливе [5].

Нехай зображення представлено двовимірною функцією $V(x, y)$ [4]. З метою проведення структурного аналізу дану функцію представляють у вигляді скінченої множини фрагментів $\{b_k(x, y)\}$ або як скінчену множину $\Lambda = \{\lambda^k\}$ характерних ознак (ХО). При цьому з кожним λ^k пов'язаний елемент $c^k = (x^k, y^k)$ множини координат $C = \{c^k\}$, який визначає просторові властивості опису. Терміни ХО або ключові точки (КТ) є еквівалентними. У СМ візуальний об'єкт представляється у вигляді сукупності дескрипторів, які описують КТ та включають атрибути λ^k і координати c^k .

Зазвичай ХО відображають кривизну функції яскравості, а також рівень відмінності від сусідніх фрагментів. Залежно від використовуваного математичного апарату, алгоритми пошуку ключових точок зображення можуть бути спрямовані на пошук кутів, перепадів, кордонів, контрастних точок, ліній та інше [5].

Використовуючи перехід від зображення до множини ХО, можна досягти істотного зниження обсягу інформації, в той час як достовірність та перешкодозахищеність при цьому практично не знижуються за рахунок вилучення та обробки нової інформації, що міститься в СЕ і їх зв'язках. У більшості застосувань ХО є числовий вектор фіксованої розмірності з дійсними значеннями, тобто: $\lambda^k = (\lambda_1, \dots, \lambda_n)^k$, $\lambda_i \in R^1, i = \overline{1, n}$.

1.3 Методи формування ключових точок

Існують різні алгоритми пошуку КТ та побудови матриці дескрипторів, у кожного з яких свої переваги та недоліки. Важливим фактором є структура даних для множини дескрипторів, що безпосередньо впливає на показники функціонування систем розпізнавання. Усі алгоритми можна поділити на дві основні групи за способом представлення дескриптору:

- дескриптори з рухомою точкою (найбільш популярні алгоритми: SIFT, SURF, CSLBP, FAST та ін.);
- бінарні дескриптори (найбільш популярні алгоритми: BRIEF, BRISK, AKAZE, ORB та ін.).

Проаналізуємо декілька алгоритмів, які використовують дескриптори з рухомою точкою.

Девід Лоу запропонував алгоритм SIFT (Scale-invariant feature transform) для визначення локальних ознак зображення, які є інваріантними до зміни масштабу зображення, його обертання, та частково інваріантними до змін в експозиції [7, 8].

Фактично цим алгоритмом можна характеризувати й ідентифікувати зображення за ХО. Основні етапи отримання ХО зображення: виявлення просторової шкали екстремуму; локалізація КТ; призначення орієнтації; генерація дескрипторів ключової точки. Першим кроком у використанні SIFT алгоритму є визначення «точок інтересу» [6]. Це реалізується шляхом створення багаторівневої піраміди зображень різниці Гауса (Difference-of-Gaussian).

Обчислення розмиття зображення за Гаусом:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y), \quad (1.2)$$

де x, y – позиція пікселя;

σ – позначення рівня розмиття;

$L(x, y, \sigma)$ – розмите зображення;

$G(x, y, \sigma)$ – функція розмиття Гауса;

$I(x, y)$ – розмите зображення.

Кожен рівень піраміди Гауса складається з:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma), \quad (1.3)$$

де D – зображення, обчислене з різниці двох сусідніх рівнів, розділених постійним мультиплікативним фактором.

Точка буде прийнята за ключову, якщо вона є локальним екстремумом різниці Гауса. Далі серед множини можливих КТ видаляються точки, що потрапили на кордон об'єктів або мають малий контраст (тобто множина уточнюється), а потім обчислюються орієнтації відібраних точок. Для цього будується зважена гістограма градієнтів в околиці, обирається орієнтація, що відповідає максимальній компоненті гістограми. Коли вже призначено місцезнаходження ключової точки, дескриптор цієї точки обчислюється з набору орієнтирів гістограм. Орієнтовні гістограми співвідносяться з орієнтирами ключової точки, дані орієнтири отримують із зображення Гауса, наближаючи їх до рівня ключових точок. Внесок кожної сусідньої точки зважується за величиною градієнта і вікном Гауса [9, 26].

Отже, метод SIFT включає наступні етапи: визначення КТ, локалізація особливостей, обчислення орієнтацій, описання КТ через дескриптор, співставлення дескрипторів.

Алгоритм SURF (Speeded Up Robust Features) [10]. Методика заснована на пошуку КТ і створенні дескрипторів, які інваріантні до масштабування та повороту за допомогою матриці Гессе. Гессіан – детермінант матриці Гессе досягає екстремуму в точках максимальної зміни градієнта яскравості. Для двовимірного випадку Гессіан обчислюється за формулою:

$$\det(H) = \frac{\partial^2 f}{\partial x^2} \cdot \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2, \quad (1.4)$$

де $f(x, y)$ – функція зміни градієнта яскравості.

Після знаходження КТ у методі SURF формуються їх дескриптори. Дескриптор являє собою набір з 64 (або 128) чисел для кожної КТ. Ці числа відображають флуктуації градієнта навколо ключової точки. Отже алгоритм SURF є інваріантним до повороту. При цьому в кожній точці розраховується градієнт максимальної зміни яскравості і коефіцієнт масштабування за матрицею Гессе. Певним чином цей алгоритм є аналогом метода SIFT.

Гессіан інваріантний до зсуву яскравості та повороту зображення, але не до масштабування. Цю проблему вирішують за допомогою послідовного перебору різних масштабів і фільтрів.

Center Symmetric Local Binary Pattern (CSLBP) [10] широко використовується для виявлення текстур і об'єктів, але його використання в хешуванні зображень все ще обмежене. Хешування зображень є потужною технікою для визначення чи зміст зображення змінено або збережено. Дві основні проблеми хешування зображень – це його компактна довжина та його потужна сила дискримінації. Для вирішення цих проблем використовують метод хешування зображень поєднання функцій CSLBP і дискретного вейвлет-перетворення (DWT). Даний алгоритм для визначення КТ виконується у три кроки [12]. На першому кроці обчислюються відхилення на спрощену функцію LoG (Лапласа-Гауса) і малі відкидаються. У результаті цього визначаються межі на зображенні. Другим кроком обчислюються локальні екстремуми. Далі за допомогою метрики Харріса визначаються екстремуми з найбільшим відхиленням. Як дескриптор використовується MU-SURF, який використовує зважені за допомогою функції Гауса з параметром 2,5 вейвлети Хаара для обчислення відхилень у горизонтальному і вертикальному напрямках для областей.

FAST (Features from Accelerated Segment Test) [13]. Вперше був запропонований в 2005 році [14]. Для точки P зображення за допомогою алгоритму Брезенхема будується коло із 16 пікселів радіусом 3. При порівнянні точок кола маємо три можливих виходи: темніше, світліше, схоже. Точка є кутовою, якщо для неї існують N суміжних пікселів на колі, інтенсивності яких не входять в межі інтервалу $(I_p - t, I_p + t)$, де I_p – інтенсивність точки P зображення, t – гранична величина відхилення. Далі необхідно порівняти інтенсивність у вертикальних і горизонтальних точках на колі з інтенсивністю в точці P . При виконанні цієї умови далі досліджується значення яскравості на окружності під номерами 1, 5, 9, 13 (рис. 1.1). Якщо для трьох з цих точок виконується умова $I_{pi} > I_p + t$ або $I_{pi} < I_p - t, i = 1, \dots, 4$, то проводиться повний тест для всіх 16 точок.

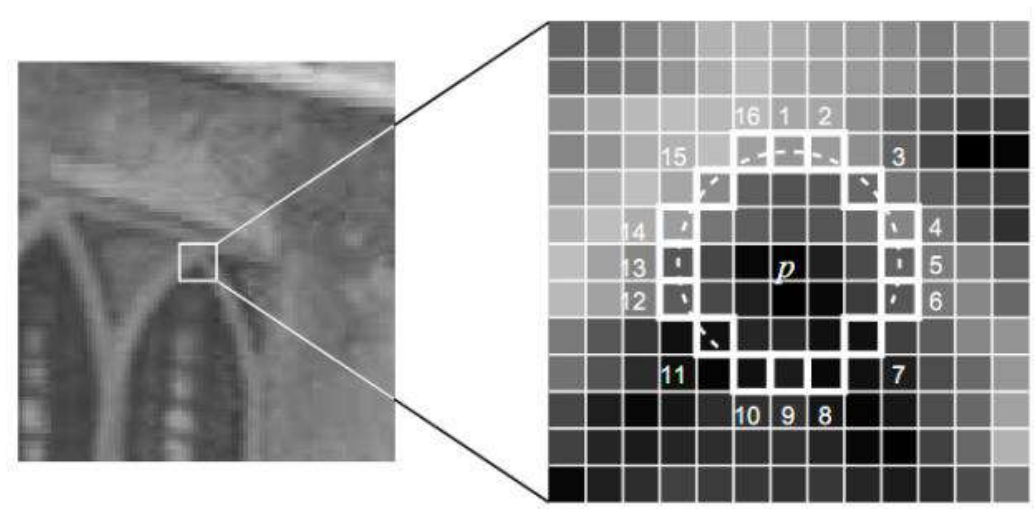


Рисунок 1.1 – Розглядувана околиця точки P детектору FAST

Вибір тільки 4 пікселів на окружності дозволяє швидко відсіяти точки, які не підходять, але в деяких випадках можливе визначення різних особливостей в одному колі.

Аналізуючи ці алгоритми, отримаємо, що дескриптори SURF – це вектор з 128 (або 64) чисел на одну ключову точку. Перевірка на збіг

виконується пошуком найближчої точки (або навіть двох). І чим ближче точка, тим краще. Тобто на зображення з 1 000 ключових точок, буде потрібно 128 000 чисел з плаваючою точкою. Крім того, виявлення точок досить складна операція і вимагає чимало часу. Що не дозволяє ефективно використовувати даний алгоритм на невеликих пристроях.

Алгоритми з використанням бінарних дескрипторів отримують дескриптор особливих точок у вигляді бінарного вектора з розміром, кратним ступені двійки. Бітова природа дескрипторів КТ у просторі V^n бінарних векторів (n – ступінь двійки) дає можливість впровадити подання та аналіз дескриптора як ланцюжка елементів (наприклад, байтів), діапазон значень яких відомий. Бінарне подання значно прискорює процес порівняння дескрипторів за рахунок можливості застосування двійкових операцій та відповідно спрощує апаратну реалізацію системи розпізнавання. Крім того, бінарна арифметика дає можливість застосувати ефективний апарат оброблення бінарних даних та синтезувати нові підходи для визначення подібності дескрипторів при побудові правил класифікації.

Алгоритм BRIEF (Binary Robust Independent Elementary Features) використовує бінарні дескриптори [27], які обчислюються після знаходження особливих точок. BRIEF – евристичний дескриптор, що будується з 256 бінарних порівнянь між яскравістю пікселів на розмитому зображенні. Бінарний тест між точками x і y визначається так:

$$\tau(P, x, y) := \begin{cases} 1: p(x) < p(y) \\ 0: p(x) \geq p(y) \end{cases} \quad (1.5)$$

Найкращий спосіб вибору точок – це випадковим чином за допомогою нормального розподілу навколо центрального пікселя. Ця випадкова послідовність точок обирається один раз і в подальшому не змінюється. Отриманий бінарний дескриптор виявляється стійкий до змін освітлення,

перспективного спотворення, швидко обчислюється і порівнюється, але нестійкий до повороту зображення.

Метод BRISK (Binary Robust Invariant Scalable Keypoints) був представлений у 2011 році. Реалізується детектування КТ за допомогою алгоритму FAST, але дещо оптимізовано [28]. Так, при пошуку КТ, для досягнення інваріантності до масштабу, пропонується вибирати найкращу КТ з максимальним значенням інтенсивності в піраміді, яка складається з 4 октав c_i та 4 внутрішніх октав $d_i, i = 0, \dots, 3$. Октави формуються як стиснення оригінального зображення c_0 в 2^i рази. Внутрішні октави розташовані між c_i та c_{i+1} і представлені у вигляді стислої c_0 в $\frac{2}{3}2^i$ рази (рис. 1.2). А ось пошук особливих точках в октавах здійснюється детектором FAST.

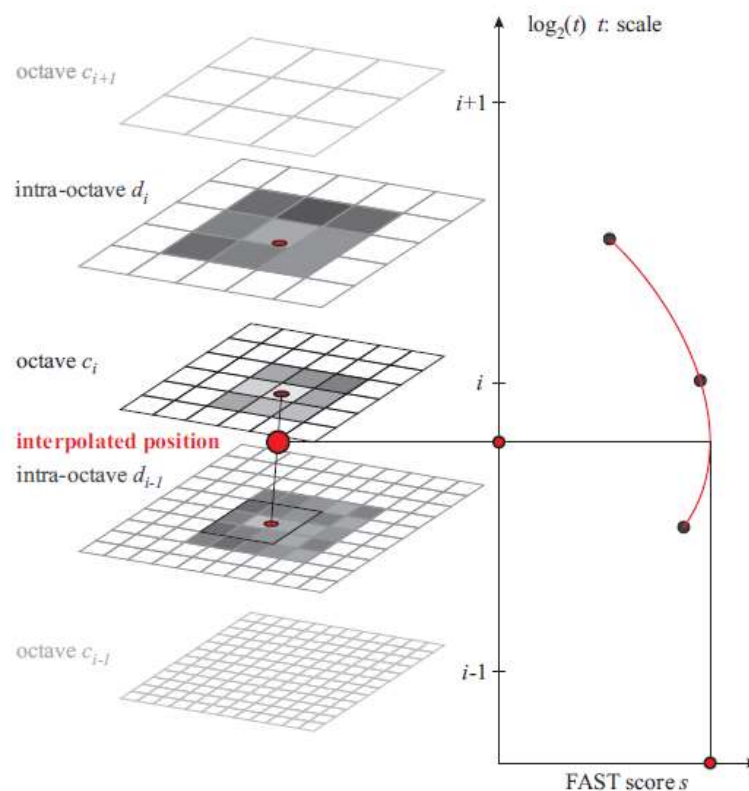


Рисунок 1.2 – Приклад пошуку КТ з максимальним значенням S

Переходячи до етапу обчислення дескриптору, варто зауважити, що зразок вибірки BRISK складається з концентричних кілець. Область навколо КТ розбивається на 60 ділянок (рис. 1.3).

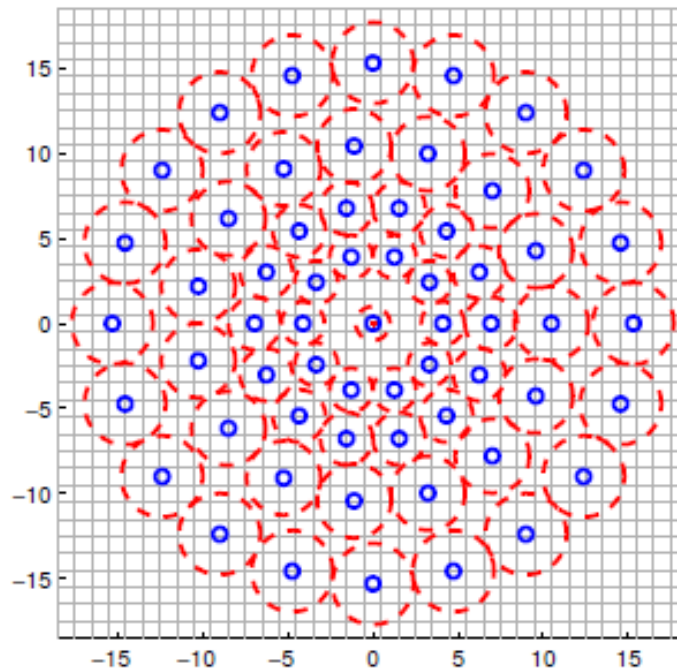


Рисунок 1.3 – Зразок вибірки BRISK

До кругових околів потенційних КТ p_i застосовується Гаусове згладжування. Червоні кола на рисунку 1.3 ілюструють розміри стандартного відхилення фільтра Гауса, що застосовується до кожної точки вибору. Для визначення напрямку КТ використовується сума локальних градієнтів [15].

Дескриптор складається з бінарних рядків довжиною 512, заповнених результатами проведених тестів в множині КТ. BRISK приймає набір коротких пар, що обертаються відповідно до напрямку:

$$b = \begin{cases} 1, I(p_j^\alpha, \sigma_j) > I(p_i^\alpha, \sigma_i) \\ 0, I(p_j^\alpha, \sigma_j) \leq I(p_i^\alpha, \sigma_i) \end{cases}, \quad (1.6)$$

де I – згладжена інтенсивність у колі радіусу σ_i відповідної точки p_i вибірки за відповідним стандартним відхиленням;

$\alpha = \arctg 2(g_y, g_x)$ – кут напрямку градієнта g .

Отже, підводячи підсумок: BRISK було оснащено механізмом виділення орієнтації та стає інваріантним до обертання.

AKAZE: Accelerated-KAZE. Алгоритм був розроблений у 2013 році. При розробці даного методу намагалися отримати високу швидкість роботи як детектора, так і дескриптора. При цьому знайдені КТ та їх дескриптори повинні були задовольняти високим показникам точності при порівнянні зображень.

При пошуку КТ застосовують алгоритм FED – Fast Explicit Diffusion на пірамідальній схемі, який дозволяє побудувати нелінійну багато-масштабну піраміду. При цьому швидкість знаходження потрібної КТ збільшується завдяки застосуванню нелінійного коефіцієнта масштабування в порівнянні з пірамідою Гауса. Обчислення цього коефіцієнта засновано на зміні яскравості зображення при масштабуванні. При застосуванні детектору для кожної октави L_i в піраміді обчислюється Гессіан:

$$L_{Hessian}^i = \sigma_{i,norm}^2 (L_{xx}^i L_{yy}^i - L_{xy}^i L_{xy}^i), \quad (1.7)$$

де $\sigma_{i,norm} = \frac{\sigma_i}{2^i}$ – нормалізований щодо масштабу коефіцієнт, для обчислення

$L_{Hessian}^i$ з урахуванням розміру октави σ_i .

Похідні другого порядку обчислюються за допомогою фільтра Хара. Фільтр дозволяє враховувати орієнтацію особливих точок. За допомогою такого підходу відбувається пошук таких точок в октаві, значення фільтра яких вище заданого порогу і є найбільшим з околиці точки 3×3 пікселів. Далі для кожної точки з потенційних максимумів порівнюється її значення

щодо результатів в сусідніх октавах $i+1$ та $i-1$ у вікні розміром $\sigma_i \times \sigma_i$ відповідно.

Дескриптори AKAZE обчислюються схожим з BRISK чином, однак, через деякі зміни обчислюються швидше. Досягнуто це завдяки тому, що до порівнянь яскравості показників областей додали порівняння значень градієнтів яскравості по осі x та y , тому результат одного тесту складається з трьох бітів замість одного. А додатково, щоб забезпечити інваріантність до обертання та масштабування: вікно дескриптора орієнтується по орієнтації особливої КТ, а розмір вікна дескриптора обирається в залежності від розміру октави σ_i в якій знайдена його особлива точка. Для прискорення обчислення тести проводяться не між середнім значенням всіх пікселів в області, а між заданою їх кількістю, в залежності від розміру σ_i . Підсумковий бінарний дескриптор має довжину 486 по три складових.

У 2011 році був розроблений алгоритм ORB (Oriented FAST and Rotated BRIEF). В цьому алгоритмі була зроблена спроба збільшити продуктивність BRIEF при повороті зображення. В алгоритмі спочатку обчислюється орієнтація точки, а потім проводяться бінарні порівняння у відповідності з цією орієнтацією.

ORB використовує алгоритм FAST для знаходження КТ. За параметр приймається граничне значення інтенсивності між центральним пікселем і описаним навколо нього колом. Радіус кола, яке досліджується навколо кожного пікселя, приймається рівним 9, оскільки саме ця відстань виявилась найбільш ефективною з точки зору продуктивності. Для отриманих точок використовується фільтр Харріса, що відкидає кандидатів, які виходять за межі значень фільтра. Принцип роботи фільтра Харріса в тому, що для зображення розглядається початкове вікно та його зрушення, обчислюється зважена сума квадратів різниць між зрушеним і вихідним вікном. Кут характеризується великими змінами цієї величини та інваріантний до поворотів.

В детекторі ORB максимальну кількість особливих точок за замовчуванням не більше 500, якщо їх більше, то до них застосовується детектор кутів Харріса для виключення найменш значущих, щоб отримати N КТ. Спочатку використовується низький поріг для того, щоб отримати точок, більше N , потім вони упорядковуються за допомогою метрики Харріса і вибираються перші N точок.

Для побудови дескриптора отриманих точок використовується модифікація BRIEF (спрямований дескриптор BRIEF), тобто для інваріантності до масштабування застосовується алгоритм на піраміді Гауса, октавами c_i якої є початкове зображення c_0 стисле з лінійним кроком. Введення параметра кутової орієнтації дозволяє домогтися стійкості детектування при обертанні об'єкта. Він заснований на напрямках градієнта яскравості щодо центру точки, напрямок з найбільшою інтенсивністю призначається орієнтацією особливої точки θ .

Для обчислення кута орієнтації особливої точки обчислюються середні яскравості для околу точки та кут орієнтації особливої точки (центроїд орієнтації):

$$m_{pq} = \sum_{x,y} x^p y^q I(x, y), \quad (1.8)$$

$$\theta = \arctg 2(m_{01}, m_{10}), \quad (1.9)$$

де (x, y) – піксельні координати;

$I(x, y)$ – середня яскравість області.

Дескриптор представляється у вигляді вектора довжиною 256, що складається з результатів бінарних тестів навколо особливої точки. В околі точки порівнюються середні значення яскравостей між x та y :

$$\tau(I; x, y) := \begin{cases} 1: I_x < I_y \\ 0: I_x \geq I_y \end{cases}, \quad (1.10)$$

Всі $n=256$ наборів x_i та y_i формують матрицю S розмірністю $2 \times n$. Далі S за допомогою матриці повороту R_θ орієнтується відповідно до кута θ

$$S_\theta = R_\theta S. \quad (1.11)$$

Більш формально це виглядає так:

$$\begin{pmatrix} x_{i'} \\ y_{i'} \end{pmatrix} = R_\theta \begin{pmatrix} x_i \\ y_i \end{pmatrix}. \quad (1.12)$$

А сам вектор дескриптора записується так:

$$g_n(I, \theta) := f_n(I) | (x_i, y_i) \in S_\theta, \quad (1.13)$$

$$\text{де } f_n(I) := \sum_{i=1}^n 2^{i-1} \tau(I; x_i, y_i).$$

Детектор ORB формує приблизно вдвічі більше дескрипторів, ніж SURF, але час розпізнання для ORB на порядок менше за рахунок скорочення самих дескрипторів.

1.4 Постановка задачі

У прикладних задачах розпізнавання образів одним із ключових завдань є спрощення та пришвидшення процесу оброблення множини дескрипторів КТ, а відповідно, класифікація візуальних образів більш досконалим способом. У роботі поставлено задачу моделювання методу для

стиснення опису як множини дескрипторів у структурних методах класифікації зображень задля прискорення класифікації даних.

Об'єктом роботи є структурні методи класифікації зображень.

Метою роботи є підвищення результативності структурних методів класифікації зображень шляхом впровадження засобів стиснення опису.

У відповідність до мети були поставлені наступні завдання:

- дослідити методи пошуку ключових точок та виявити найбільш оптимальний для вирішення поставленої задачі;
- розробити метод класифікації шляхом зіставлення множини аналізованих зображень з еталоном;
- реалізувати програмну модель алгоритму редукції для прискорення процесу класифікації;
- експериментально перевірити результативність модифікованого методу класифікації із впровадженням засобів редукції.

2 МЕТОД РЕДУКЦІЇ ДЛЯ МНОЖИНИ ДЕСКРИПТОРІВ

2.1 Метричні та статистичні критерії інформативності

Для підвищення ефективності систем класифікації даних шляхом зменшення часових та обчислювальних обсягів обробки інформації серед наявної множини аналізованих ознак часто потрібно вибрати найбільш інформативну підмножину. Візуальні об'єкти та їх розпізнавання відносяться до такого типу задач, де використовується багато-параметричне та об'ємне подання аналізованих даних. Тому задача зменшення множини аналізованих даних є актуальною у класифікації образів.

При розробленні класифікаторів на підґрунті образу як множини дескрипторів КТ зображення ознаками безпосередньо є дескриптори – числові (бінарні) вектори, а у якості основного принципу для визначення класу об'єкта застосовуються метричні співвідношення як для окремих компонентів опису, так і для скінченних множин векторів. Метричні співвідношення у просторі дескрипторів фізично відображають близькість (еквівалентність) відповідних фрагментів зображення із врахуванням допустимих геометричних перетворень об'єктів, що виникають у процесі візуального спостереження [29 – 31]. Отже, задача даної роботи полягає саме в стисненні множини бінарних векторів.

Різноманітні методи класифікації, моделювання та прогнозування, що ґрунтуються на застосуванні дерев рішень, штучних нейронних мереж, генетичних алгоритмів, еволюційного програмування, асоціативної пам'яті, нечіткої логіки складають основу методів Data Science. До методів Data Science також відносяться статистичні методи: описовий аналіз, кореляційний і регресійний аналіз, факторний аналіз, дисперсійний аналіз, компонентний аналіз, дискримінантний аналіз, аналіз часових рядів, аналіз зв'язків. Такі методи припускають певні апріорні уявлення про дані, які

аналізуються. У сучасних методах Data Science при побудові дерев рішень у класифікаторах зображень на основі кластерного подання множин дескрипторів КТ знайшли впровадження критерії прийняття рішень із урахуванням оцінок інформативності [22, 28, 32].

У статті [32] вивчається результативність класифікаторів із використанням дерева рішень на підставі застосування статистичних критеріїв приросту інформації, заснованих на показниках інформаційної ентропії Шеннона, α -ентропії Рен'ї та коефіцієнту Джині.

При виборі методу побудови множини КТ застосовується критерій стійкості до геометричних перетворень (повороту та руху). Критерієм для редукції множини дескрипторів КТ раніше визначалися рівень стійкості відносно дії геометричних перетворень, повторюваність значень на множині дескрипторів власного та решти класів в умовах перетворень, побудова статистичних процентильних відношень для описів на основі визначення медоїду для множини векторів [33, 34].

Одним із результативних засобів редукції даних є здійснення кластеризації для множини компонентів опису із подальшим використанням для класифікації значень визначених центрів кластерів чи кількісних параметрів для кластерів еталонних зображень [16]. Один з таких методів використовує самоорганізуючу карту (SOM). Самоорганізуюча карта (SOM) з її варіантами є найпопулярнішим алгоритмом штучної нейронної мережі в категорії навчання без вчителя. SOM вирішує складні високорозмірні та нелінійні проблеми, такі як виділення ознак і класифікація зображень і акустичних моделей [35].

Крім статистичних співвідношень за критерій інформативності можна обирати метричні співвідношення, які ґрунтуються на величині подібності компонентів опису.

В засобах інтелектуального аналізу даних сьогодні використовують систему критеріїв інформативності, засновану на розташуванні об'єктів різних класів у просторі ознак. Така система метричного типу дозволяє

оцінювати також гуртову інформативність для підмножин ознак при вирішенні задач класифікації та регресії.

Важливо при застосуванні метричних моделей класифікації у таких методах класифікації як, наприклад, метод ближніх сусідів, цілеспрямовано знизити розмірність еталонних описів шляхом відбору із наявної повної множини ознак її підмножини значно меншої потужності за якимось критерієм. Справа в тому, що еталонні описи безпосередньо, гуртом і не однократно приймають участь у процесах розпізнавання [36].

Головна мета при цьому – зниження впливу «прокляття розмірності» при впровадженні класифікаторів у прикладних системах, де кількість еталонів вимірюється десятками та сотнями [37].

Метричні підходи використовуються у задачах кластеризації та кластеризації багатовимірних даних. У якості критерію якості кластеризації можна використати поняття функції силуету, що відображає нормовану різницю між середніми відстанями кожної точки даних до множини точок власного та сусіднього кластерів. Існує ряд методів для формування ефективної стисненої системи ознак на основі метричних критеріїв та статистичного аналізу даних. Одним з них є Relief. Цей метод випадковим чином обирає із бази даних зразки і поновлює значущість кожної ознаки на основі різниці між обраним зразком і двома його найближчими сусідами того ж або протилежного класу.

2.2 Застосування метричних критеріїв інформативності для множини дескрипторів

Розглянемо опис E еталонних зображень з N еталонів $E = \{E_1, E_2, \dots, E_N\}$. Еталон E_k у задачі класифікації представляє окремий клас і є підмножиною E . Кожен окремий клас є скінченною множиною $e_v(k)$

КТ у просторі B^n , де k – це параметр номеру класу, який має кожен дескриптор:

$$E_k = \{e_v(k)\}_{v=1}^s, \quad (2.1)$$

де s – це число дескрипторів у множині E_k .

Загальна кількість дескрипторів у базовій множині E знаходиться як sN .

У просторі B^n введено поняття відстані $\rho(z_1, z_2)$ між двома дескрипторами z_1 і z_2 ($\forall z_1, z_2 \in B^n$)

$$\rho: B^n \times B^n \rightarrow R(B^n), \quad (2.2)$$

де $R(B^n)$ – це простір значень метрики ρ , що встановлює величину метричного відношення відмінності між векторами із B^n .

У випадку бінарних даних відстань $\rho(z_1, z_2)$ може бути метрикою Геммінга:

$$\rho(z_1, z_2) = \sum_{v=1}^n l(z_{1,v}, z_{2,v}). \quad (2.3)$$

Функція $l(z_{1,v}, z_{2,v})$ порівнює біти з однаковими номерами v для двох бінарних векторів та має множину значень $\{0, 1\}$, де 1 – у випадку співпадання бітів, а 0 – у протилежному випадку. Фактично, в такому випадку $\rho(z_1, z_2)$ дає число співпадань.

Інформативність $V(z, E)$ для довільного дескриптора z з множини описів E можна визначити наступним чином:

$$V(z, E) = \rho_{\min}(z, \overline{E_k}) - \rho_{\min}(z, E_k). \quad (2.4)$$

Вектор z є елементом еталонного опису E_k з номером k . За формулою $\rho_{min}(z, \overline{E_k})$ – це мінімальна відстань від z до елемента бази, що не належить класу E_k , а $\rho_{min}(z, E_k)$ – це мінімальна відстань від z до найближчого елемента із класу E_k за виключенням нульової відстані вектора до самого себе:

$$\rho_{min}(z, \overline{E_k}) = \min_{v, i \neq k} \rho(z, e_v(i)), \quad (2.5)$$

$$\rho_{min}(z, E_k) = \min_{v, i=k} \rho(z, e_v(i)). \quad (2.6)$$

Якщо впровадити нормовані відстані $0 \leq \rho \leq 1$, то значення критерію інформативності (2.4) буде знаходитися в інтервалі $-1 \leq V \leq 1$. У методі Relief використовується один із варіантів моделі (2.4), де інформативність оцінюється як різниця квадратів відстаней, а відбір аналізованих ознак здійснюється випадковим чином.

Визначення індивідуальної інформативності V елементу опису окремого класу E_k , яка обчислюється за формулою (2.4), засновується на наступному припущенні: інформативність ознаки є тим вищою, чим в більшій мірі вона розділяє елементи опису на класи. Виходячи з цього припущення, можна зауважити, що чим більшою є відстань між найближчими екземплярами різних класів, тим більш високою є індивідуальна інформативність елементів опису. З іншого боку, що чим більшою є відстань між найвіддаленішими екземплярами одного класу, тим більш низькою є індивідуальна інформативність елементів опису.

Варіації моделі (2.4) дають можливість побудувати спектр способів оцінювання інформативності в залежності від форми подання вхідних даних. За допомогою цих варіацій можна створити більш швидкісний пошук. Наприклад, якщо до множини опису E попередньо ввести кластерне подання. У такому випадку еталонні дані класів E_k можна представити наборами

визначених центрів $\{c_k(j)\}_{j=1}^m$, де m – фіксоване число центрів для кожного класу E_k . Тоді визначення мінімумів у формулі (2.4) можна замінити на пошук серед наборів центрів кластерів, що значно пришвидшує процес пошуку:

$$\rho_{min}(z, \overline{E_k}) = \min_{j=1, \dots, m, i \neq k} \rho(z, c_i(j)). \quad (2.7)$$

Зважаючи на те, що центри $c_i(j)$ кластерів через їх спосіб обчислення взагалі можуть не належати простору дескрипторів, тут потрібно застосовувати більш універсальний тип метрики, наприклад, мангеттенську відстань. Мангеттенська метрика (метрика прямокутного міста, метрика L_1) – метрика, запроваджена Германом Мінковським. За цією метрикою, відстань між двома точками дорівнює сумі модулів різниць їх координат.

У виразі (2.7) замість центрів кластерів можуть бути використані інші представники опису, наприклад, процентильні точки з певним процентильним рангом множини E_k , отримані при медіанному поданні даних цієї множини [38, 39]. Медіанне подання ґрунтується на обчисленні медіан. Медіана є найбільш загальною й фундаментальною характеристикою центра розподілу, оскільки вона базується на принципі симетрії. Процентильними точками є точки з певним процентильним рангом. Процентильним рангом деякого спостереження ряду розподілу називається сума відсотків, що відносяться до спостережень, які стоять у розподілі перед ним, і половини відсотків, які відносяться безпосередньо до нього. Процентильні ранги дають можливість співставляти між собою елементи різних розподілів. Процентильний ранг елемента розподілу часто виявляється змістовнішим, ніж власне значення цього елемента.

Для використання у якості процентильних точок треба сформулювати список дескрипторів із E_k , відсортований за значенням сумарної відстані до решти дескрипторів, що застосовується при визначенні медоїду множини, і

вибрати, наприклад, три представники із номерами, що відповідають дольовим значенням 0,25; 0,5; 0,75 щодо місця розташування у списку [40].

Зауважимо, що критерій (2.4) продуктивно буде враховувати відмінності між елементами свого класу тільки у випадку аналізу множини ознак, що не є значуще метрично близькими між собою всередині класу, так як для таких ознак для другої складової в (2.4) виконується $\rho_{min}(z, E_k) \approx 0$. Для зниження впливу «повторюваності» метрично близьких елементів всередині множини E_k ознак еталонів на попередньому етапі можна додатково обробити дані з виключенням повторюваних елементів або урахуванням числа повторень. Прикладом такої обробки може бути попереднє скорочення множини E_k за допомогою логічної фільтрації з визначеним порогом δ_ρ

$$\rho(z, e_v(i)) \leq \delta_\rho. \quad (2.8)$$

З іншого боку, можна обчислити параметри ознак за виразом (2.4), а потім провести додатковий аналіз отриманих значень інформативності.

Значення критерію інформативності (2.4) буде знаходитися в інтервалі $-1 \leq V \leq 1$. При високому значенні індивідуальної інформативності ($V \rightarrow 1$) ознаки вважаються значущими. Якщо значення індивідуальної інформативності низькі ($V \rightarrow -1$), то вони вважаються малозначущими, що створює умови для виключення цих ознак із системи.

У випадку розпізнавання на множині багатьох класів (більше двох) з метою більш поглибленого аналізу можна у виразі (2.4) враховувати відстань не тільки до найближчого елементу з протилежним значенням класу, а, наприклад, як середню відстань до множини найближчих екземплярів кожного з можливих класів. В прикладному застосуванні таке усереднення іноді зважують ваговими коефіцієнтами класів у конкретній вибірці. Тоді відстань обчислюється з урахуванням зважування на частоту класів у вибірці.

Якщо для фіксованого елементу z структурного опису уже сформовано набір вагових коефіцієнтів $\{\alpha_i\}_{i=1}^N$ щодо віднесення його до множини класів у вигляді

$$z \rightarrow (\alpha_1, \alpha_2, \dots, \alpha_N), z \in E_k, \sum_{i=1}^N \alpha_i = 1, \quad (2.9)$$

то інформативність ознаки z на підґрунті моделі (2.9) можна оцінити величиною різниці власного та максимального коефіцієнтів серед решти класів

$$V(z, E) = \alpha_k - \max_{i \neq k} \alpha_i. \quad (2.10)$$

Оцінювання критеріїв (2.9), (2.10) здійснюється на підставі аналізу метричних співвідношень для значень елементів опису та визначених параметрів $c_i(j)$ центрів еталонних класів.

2.3 Ефективність сукупної інформативності

Критерій сукупної інформативності $V(Z^*)$ обчислюється як середнє значення для набору індивідуальних критеріїв довільної підмножини ознак $Z^* \subseteq Z$. Цей критерій може бути використаний для будь-якої підмножини Z^* , зокрема і для множини описів окремих еталонів E_k , що дасть змогу скоротити обсяги описів.

Критерій сукупної інформативності можна обчислювати не тільки як середнє значення, а застосувати медіанне значення для набору значень інформативності елементів. І медіанне, і середнє значення дозволяють врахувати повторюваність або схожість ознак всередині опису.

Критерії, розглянуті у попередньому розділі, можна застосувати з метою редукції багатоелементної множини наявних ознак шляхом відбору

найбільш інформативної підмножини дескрипторів еталонних описів задля цілеспрямованого скорочення обчислювальних затрат на класифікацію. Медіанне та середнє значення можна застосувати для оцінювання і визначення за допомогою порівняння результативності впровадження різних варіантів стиснення ознак.

Редукцію множини дескрипторів можна проводити різними шляхами:

- заздалегідь обмежити число інформативних ознак та зафіксувати їх кількість;
- відібрати певне число ознак із інформативністю, що не менше зазначеного порога;
- сформулювати таку підмножину ознак, що будуть забезпечувати необхідну результативність.

При впровадженні процедури скорочення системи ознак постає питання, скільки додаткових часових витрат це вимагає. Такі витрати порівняні з обчисленням відстаней між множинами векторів. Крім того процедура зменшення числа елементів множини ознак проводиться до етапу оброблення даних і не впливає на часові затрати під час класифікації.

З іншого боку, потрібно враховувати, що при некоректному скороченні множини ознак може бути втрачена результативність класифікації візуальних образів. Тому редукція вимагає також деякої модифікації самого методу розпізнання.

Для обчислення значень інформативності як варіанту дії з множинами векторів може бути успішно використано методи упорядкування вибірки ознак-векторів за значенням деякого критерію або засобів прискорення пошуку даних.

Аналізовані критерії призначені для безпосереднього застосування у системах відмінних між собою ознак. У разі повторюваності або близькості ознак після визначення індивідуальної інформативності необхідно провести додатковий аналіз із метою подальшого скорочення близьких або однакових

ознак, можливо, із впровадженням кількісних вагових коефіцієнтів повторюваності, що може бути використано для побудови класифікатора.

У обговорюваних моделях для визначення інформативності у ролі ознак використовуються безпосередньо дескриптори КТ. Відмітимо, що ці ж моделі можуть бути впроваджені також і для будь-якої системи ознак, що є функцією від множини дескрипторів.

Використання запропонованих критеріїв для оцінювання ступеня інформативності та відбору системи інформативних ознак дає можливість у задачах розпізнавання візуальних образів знижувати структурну та обчислювальну складність синтезованих моделей класифікації, підвищувати їх інтепрованість (розуміння людиною) і узагальнюючі властивості за рахунок виключення малозначущих, взаємозалежних і надлишкових ознак.

Використання запропонованих критеріїв для оцінювання ступеня інформативності та відбору системи інформативних ознак дає можливість у задачах розпізнавання візуальних образів знижувати структурну та обчислювальну складність синтезованих моделей класифікації, підвищувати їх інтепрованість (розуміння людиною) і узагальнювальні властивості за рахунок виключення малозначущих, взаємозалежних і надлишкових ознак [18-24].

Оцінювання міри релевантності двох множин A, B однотипних векторів, зокрема і редукованих, можна здійснити традиційним шляхом, наприклад, з використанням метрик [2, 13]. Застосуємо для цього відстань Хаусдорфа

$$X(A, B) = \max \left\{ \max_{a \in A} \rho(a, B), \max_{b \in B} \rho(b, A) \right\}, \quad (2.11)$$

де $\rho(a, B) = \min_{b \in B} \rho(a, b)$, ρ – метрика для векторів (наприклад, відстань Хемінга).

Після обчислення відстані Хаусдорфа, застосуємо відстань Танімото (Жаккара), яка означає відношення числа елементів симетричної різниці та об'єднання множин:

$$T(A, B) = \frac{\text{card}(A \Delta B)}{\text{card}(A \cup B)}. \quad (2.12)$$

3 РЕЗУЛЬТАТИ КОМП'ЮТЕРНОГО МОДЕЛЮВАННЯ МЕТОДУ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ

3.1 Огляд програмних засобів реалізації

Розпізнавання візуальних образів на зображенні з використанням методу BRISK реалізовано програмно у середі розробки PyCharm мовою програмування Python з використанням бібліотеки OpenCV.

Python – це мова програмування високого рівня, яка підтримує об'єктно-орієнтовану, процедурну, функціональну та аспектно-орієнтовану парадигми програмування. Вона була розроблена в 1990 році Гвідо ван Россумом.

Спочатку Python задумувався як нащадок мови ABC, що призначений для навчання програмуванню. Це досить важливий факт в історії Python, так як саме він поставив на перше місце лаконічність і простоту синтаксису для майбутнього Python, якими не могла похизуватися переважна більшість мов програмування того часу. Відносний лаконізм мови Python дозволяє створити програму, яка буде набагато коротше свого аналога, написаного на іншій мові. Програми на Python можна запускати на комп'ютерах під різними операційними системами без будь-яких змін вихідного коду та налаштувань. На сьогоднішній день Python використовується в роботі YouTube, Mozilla, Instagram, Facebook, Google, Dropbox тощо.

Однією з переваг Python є чистий і простий синтаксис. У цій мові програмування кодинг максимально наближений до людини, що дозволяє зробити швидкий старт у програмуванні. Для виділення блоків у коді використовуються пробіли та табуляція. До того ж Python вимагає мінімальну кількість коду, необхідного для вирішення задачі, у порівнянні з іншими популярними мовами. Іншою перевагою є кроссплатформеність. Python має інтерпретації під усі відомі порти Microsoft Windows, всі варіанти

UNIX (включаючи FreeBSD та GNU/Linux), Plan 9, Mac OS та Mac OS X, iPhone OS 2.0 і вище, Palm OS, OS/2, Amiga, AS/400 та навіть OS/390[en], Symbian та Android. У міру старіння платформи її підтримка в основній гілці мови припиняється. Наприклад, з версії 2.6 припинена підтримка Windows 95, Windows 98 та Windows ME. Однак на цих платформах можна використовувати попередні версії Python – спільнота активно підтримує версії Python починаючи від 2.3 (для них виходять виправлення). При цьому, на відміну від багатьох портованих систем, для всіх основних платформ Python має підтримку характерних для даної платформи технологій (наприклад, Microsoft COM/DCOM). Існує спеціальна версія Python для віртуальної машини Java – Jython, що дозволяє інтерпретатору виконуватися на будь-якій системі, яка підтримує Java, при цьому класи Java можуть безпосередньо використовуватися з Python й навіть бути написаними на ньому. Також кілька проєктів забезпечують інтеграцію з платформою Microsoft.NET, основні з яких – IronPython та Python.Net.

Дана мова програмування включає в себе велику кількість бібліотек та корисних модулів, зокрема для розробки графічного інтерфейсу. Бібліотеки можуть вільно розповсюджуватися. Інтерпретатор мови може бути розширений функціями та типами даних, розроблених на C або C++.

Python зручний для розв'язування математичних проблем, що робить його широко застосовним у наукових дослідженнях.

До недоліків можна віднести низьку швидкість виконання програм. Причиною цього серед іншого є статична типізація. Цей недолік компенсується зменшенням часу розробки програми.

Крім простих базових типів, у Python існують і базові типи, що представляють структури даних. Так, наприклад, у Python є такі типи даних як: `list` – список, що може зберігати у собі впорядковані набір даних різного типу, `tuple` – кортеж, що як список може зберігати у собі набір даних, але на відміну від списку має фіксовану довжину та не може її змінювати, `dict` – словник, цей тип можна порівняти з так званою хеш-мапою у інших мовах

програмування, цей тип даних дозволяє зберігати попарні відповідності ключзначення, та `set` і `frozenset` – множина і незмінна множина, відповідно, це тип даних дозволяє зберігати неупорядковані набори даних довільної і фіксованої довжини відповідно. Всі типи даних мають відповідний інтерфейс роботи з ними. Наприклад, множини мають методи для виконання операцій алгебри множин, серед яких операції об'єднання, перетину, симетричної різниці, тощо. Таким чином, Python дозволяє дуже легко працювати у межах теорії множин. Існування базових типів, що представляють структури даних робить Python дуже зручним для роботи з наборами даних для обробки. Іноді їх можливостей може бути недостатньо для вирішення деяких задач. Але ця проблема вирішується великою кількістю бібліотек, розроблених спільнотою, згаданою вище. Серед них і бібліотеки, що дозволяють обчислювати статистичні показники масивів, такі як: середнє, дисперсію, середньоквадратичне відхилення, медіану, присутні функції для розрахунку норми, функція побудови нормального розподілу, тощо.

PyCharm – інтегроване середовище розробки для мови програмування Python з повним набором засобів для ефективною розробки. Надає засоби для аналізу коду, графічний зневаджувач, інструмент для запуску юніт-тестів і підтримує веб-розробку на Django. Проста і потужна навігація у кодї, авто доповнення, авто імпорт, шаблон кода, перевірка на сумісність версії інтерпретатора мови, швидкий перегляд документації у вікні середовища. PyCharm розроблена чеською компанією JetBrains в 2010 році. PyCharm працює під операційними системами Windows, Mac OS X і Linux.

OpenCV (англ. Open Source Computer Vision Library, бібліотека комп'ютерного зору з відкритим кодом) – бібліотека функцій та алгоритмів комп'ютерного зору, обробки зображень і чисельних алгоритмів загального призначення з відкритим кодом. Бібліотека надає засоби для обробки і аналізу вмісту зображень, зокрема розпізнавання об'єктів на фотографіях (наприклад, осіб і фігур людей, тексту тощо), відстежування руху об'єктів,

перетворення зображень, застосування методів машинного навчання і виявлення загальних елементів на різних зображеннях.

Бібліотека розроблена Intel і нині підтримується Willow Garage та Itseez. Сирцевий код бібліотеки написаний мовою C++ і поширюється під ліцензією BSD. Біндинги підготовлені для різних мов програмування, таких як Python, Java, Ruby, Matlab, Lua та інших. Може вільно використовуватися в академічних та комерційних цілях.

3.2 Результати комп'ютерного моделювання

Сучасні системи комп'ютерного зору для розпізнавання візуальних образів використовують методи, що працюють подібно зору людини, який фокусується на локальних прикметах зображення. В основі цих методів лежить визначення деякої множини ключових точок (КТ). Їх описують у вигляді дескриптора – бінарний чи числовий вектор, який містить характеристику вмісту функції яскравості для околу кожної КТ.

Для визначення опису об'єкту як множини дескрипторів на зображенні використано детектори КТ. В роботі було використано метод BRISK з бінарним представленням дескриптора, що зменшує вимоги до обчислювальних ресурсів та надає перевагу у швидкості обчислень. У цьому методі детектором FAST визначено координати КТ, а методом BRIEF розраховуються дескриптори. Досягнення інваріантності до масштабу здійснюється за рахунок вибору найкращої КТ з максимальним значенням інтенсивності в піраміді, що складається з 4 октав c_i і 4 внутрішніх октав d_i ($i = 0...3$). Октави формуються шляхом стиснення оригінального зображення c_0 у 2^i рази. Надалі відбувається пошук КТ детектором FAST.

Область, сформована навколо КТ, ділиться на 60 ділянок p :

$$A = \{(p_i, p_j) \in R^2 \times R^2 | i < N \wedge j < i \wedge j \in N\}. \quad (3.1)$$

Множина A розбивається на $2a$ підмножини:

$$S = \{(p_i, p_j) \in A \mid \|p_i - p_j\| < \delta_{max}\} \subseteq A, \quad (3.2)$$

$$L = \{(p_i, p_j) \in A \mid \|p_i - p_j\| > \delta_{min}\} \subseteq A, \quad (3.3)$$

де $\delta_{min} = 13,67t$;

$\delta_{max} = 9,75t$;

t – розмір КТ.

Наступним кроком є обчислення середнього значення градієнта множини A . Дескриптор заповнюється результатами проведених тестів в підмножині S (3.1) – (3.3).

Дескриптори, сформовані методом BRISK, мають вид бінарного вектора розміром 512. Дані у вигляді значень від 0 до 255 зберігаються у вигляді матриці, кількість рядків – число КТ, стовпців – 32.

Розпізнавання візуальних образів на зображенні з використанням методу BRISK реалізовано програмно у середі розробки PyCharm мовою програмування Python з використанням бібліотеки OpenCV.

Результат обробки зображень з нанесеними околами координат отриманих КТ наведено на рисунку 3.1.



Рисунок 3.1 – Зображення з околами КТ

Для перевірки прикладної вагомості теоретичних викладок програмно змодельовані дві скінченні множини бінарних векторів із 100 елементів із компонентами у 32 біти. Елементи множин різнилися фіксованими ймовірностями появи одиниць у наборі із 32 бітів. Для визначення подібності між множинами векторів застосовано метрики (2.11) X Хаусдорфа та (2.12) T Танімото, де у якості внутрішньої метрики між елементами (бінарними векторами) використано відстань Хемінга [41].

Для програмної реалізації було вирішено створити дві функції: 1) для розрахунку супремуму однієї множини відносно іншої; 2) для пошуку максимуму із знайдених супремумів. Для розрахунку супремума на вхід подаються дві множини дескрипторів, а також метрика для розрахунку відстаней, і знаходиться максимін відстаней від кожного з елементів першої множини до елементів другої. Після розрахунку супремумів для першої множини відносно другої і другої відносно першої обирається максимум із них. Таким чином отримуємо відстань між двома множинами.

Для проведення валідації обчислень було взято дві множини дескрипторів і деяку третю множину. Перевірка показала, що метрика Хаусдорфа працює коректно, що дозволяє перейти до реалізації метрики Танімото. Для цього розраховується кардинальність симетричної різниці між цими множинами та кардинальність об'єднання двох множин і перше ділиться на друге. Метрика Танімото також спрацювала. Отже, запропоновані метрики можна використовувати для оцінки подібності наших множин.

Для реалізації алгоритму обчислення інформативності ознак між екземплярами скористалися формулою (2.4). Для знаходження мінімальної відстані від дескриптора оцінюваної множини до дескрипторів іншої була створена функція, що на вході має вектор, множину ознак іншого класу та метрику яку потрібно використовувати для обчислення відстаней між дескрипторами.

Для виявлення повторювальних та схожих елементів однієї множини була створена функція для пошуку мінімальних відстаней у самій множині. Виконується фільтрація, під час якої виключаються нулі, знаходиться мінімальна ненульова відстань у множині дескрипторів. Після реалізації функції пошуку мінімальної відстані обчислюються оцінки інформативності за формулою (2.4).

Наступним був реалізований метод, що приймає на вхід обидва набори дескрипторів та розраховує для них оцінки інформативності. У цьому методі викликаються раніше використані методи розрахунку інформативності, спочатку для першої множини відносно другої, а потім для другої відносно першої, запам'ятовуючи результат всередині об'єкта класу моделі. Це дозволяє у будь-який час повернутися до оцінок інформативності, отриманих під час навчання, та використати їх повторно.

Переходимо до процесу фільтрації ознак за оцінками інформативності. Метод редукції приймає на вхід оцінки інформативності, потім сортує їх і додає до них індекси. Далі застосовується сортування у порядку зменшення інформативності. Отже, маючи відсортовані за інформативністю індекси можна брати ту кількість найбільш значущих, що потрібна. Використовуючи цей метод, реалізується останній, що буде приймати на вхід обидві множини дескрипторів та обирати для кожної лише найбільш інформативні ознаки.

Таким чином була отримана повна модель відбору ознак за критерієм інформативності, яку можна використовувати далі. Оцінки інформативності обох множин до і після редукції представлені у таблиці 3.1.

Таблиця 3.1 – Оцінювання інформативності двох множин дескрипторів

Оцінка	Перша множина		Друга множина	
	До (100)	Після (50)	До (100)	Після (50)
1	2	3	4	5
Мінімум	3	3	5	5

Продовження таблиці 3.1

1	2	3	4	5
Максимум	-11	0	-11	0
Розмах	14	3	16	5
Загальна сума	-228	35	-190	57
Середнє значення	-2,28	0,7	-1,9	1,14

Для оцінки інформативності потрібно проаналізувати наступні показники: мінімальне та максимальне значення оцінок, їх розмах, загальну та середню інформативність. Реалізація оцінки інформативності відбувається за рахунок метрики Хемінга з 50 відібраними ознаками (із 100 початкових), де на вході – дві множини дескрипторів, а на виході оцінки інформативності. В результаті реалізації були отримані наступні оцінки: максимум 3, мінімум -11, розмах 14, загальна сума -228, середнє -2,28. Від’ємні значення ознак свідчать про те, що множина ознак включає ознаки, які відносяться до інших множин і мають низьку інформативність. До того ж за модулем значення мінімуму значно перевищує максимум, що вказує на наявність ознак, що негативно впливають на процес навчання за цими ознаками. Від’ємні значення загальної суми і середнього значення вказують на присутність ознак, що зменшують інформативність усього класу.

Перевіримо, чи позитивно редукція вплинула на оцінки ефективності. В результаті реалізації були отримані наступні оцінки: максимум 3, мінімум 0, розмах 3, загальна сума 35, середнє 0,7. За цими даними очевидно, що в результаті редукції позбулися від’ємних оцінок і, як наслідок, виділенні ознаки не будуть негативно впливати на подальше навчання. Якщо немає від’ємних оцінок, то всі ознаки, які залишилися, повністю інформативні.

Зменшення розмаху вказує на збільшення концентрації оцінок однакової інформативності. Небажана оцінка, від’ємне середнє значення, після редукції отримала додатне значення.

Аналіз таблиці 3.1 стосовно першої множини показав, що процес редукції відносно неї пройшов задовільно.

Розглянемо другу множину. В результаті реалізації програми перевірки інформативності були отримані наступні оцінки: максимум 5, мінімум -11, розмах 16, загальна сума -190, середнє -1,9. Від'ємні значення ознак свідчать про те, що множина ознак включає ознаки, які відносяться до інших множин і мають низьку інформативність. До того ж за модулем значення мінімуму значно перевищує максимум, що вказує на наявність ознак, що негативно впливають на процес навчання за цими ознаками. Від'ємні значення загальної суми і середнього значення вказують на присутність ознак, що зменшують інформативність усього класу.

Перевіримо, чи позитивно редукція вплинула на оцінки ефективності для другої множини ознак. В результаті реалізації були отримані наступні оцінки: максимум 5, мінімум 0, розмах 5, загальна сума 57, середнє 1,14. За цими даними очевидно, що в результаті редукції позбулися від'ємних оцінок і, як наслідок, виділенні ознаки не будуть негативно впливати на подальше навчання. Якщо немає від'ємних оцінок, то всі ознаки, які залишилися, повністю інформативні. Зменшення розмаху вказує на збільшення концентрації оцінок однакової інформативності. Небажана оцінка, від'ємне середнє значення, після редукції отримала додатне значення.

Аналіз таблиці 3.1 стосовно другої множини показав, що процес редукції відносно неї пройшов задовільно.

Використовуючи редуковану систему ознак, оцінимо подібність множин за допомогою Хаусдорфа і Танімото (табл. 3.2).

Таблиця 3.2 – Значення оцінок за метриками Хаусдорфа і Танімото

	Хаусдорфа	Танімото
До	12	0,86
Після	13	1

Як видно з таблиці 3.2, і за метрикою Хаусдорфа, і за метрикою Танімото відстань між класами після редукції збільшилася, що є позитивним і закономірним результатом, тому що були виключені ті ознаки, які несли негативну інформативність для еталонних класів. Залишилися тільки ті дескриптори, які описують виключно класи, яким вони належать [42 – 49].

Перевіримо, яким чином редукція вплинула на просторову та часову складність виконання обчислення. Об'єм пам'яті (просторова складність), необхідний для зберігання об'єкту, до редукції 25,11 Кб, а після 12,06 Кб, тобто зменшився удвічі, але це закономірно, бо об'єм множин дескрипторів також зменшився удвічі.

Для визначення рівня оптимізації часової складності скористаємося утилітою `%timeit` інтерпретатора Python, яка вимірює час виконання поданої на вхід команди. Час виконання для метрики Хаусдорфа для початкових множин склав 26 мс, а для редукованих – 5,04 мс. Час виконання для метрики Танімото для початкових множин склав 1,45 мс, а для редукованих – 0,294 мс. За обома метриками процес пришвидшився у майже 5 разів.

Зробимо остаточний висновок з приводу застосування редукції до двох множин ознак.

Для двох масивів векторів з ймовірностями появи 1: $p_1 = 0,5$ та $p_2 = 0,2$ отримано значення метрик $X = 12$, $T = 0,85$. Як еквівалентність векторів у (2.12) застосовано повне їх співпадіння. Нормована метрика T для нашого прикладу приймає значення із відрізка $[0,1]$, а метрика X – цілі значення із відрізка $[0,32]$. Як бачимо, за метрикою Танімото ці множини суттєво відрізняються, а за метрикою Хаусдорфа – різняться приблизно на третину.

Обчислення значень інформативності показали, що для обох множин інформативність елементів опису змінюється приблизно у однакових межах: -11,...,3 для першої множини і -11,...,5 – для другої (рис. 3.2).

Множина 1

[-10, -3, 0, -1, 0, 1, 0, 0, 0, 1, -11, 0, -1, 1, 0, 0, 0, -2, -10, 0, 0, -8, 0, 1, -10, -10, -1, 1, 0, -1, -3, 1, 0, 0, 0, 3, -10, -10, 2, 0, 0, 0, 0, 0, 2, -1, -10, 2, -10, 2, 1, -7, -9, 1, 1, -9, -2, 2, -9, -1, -1, -7, -1, 0, -10, 1, -9, 0, 1, -7, 0, -1, -2, -3, 0, 3, 1, -1, -1, 0, -1, 2, 0, 0, 0, 0, -10, -11, 0, -11, 0, 2, 0, -10, -1, -9, -8, -1, 3, -9]

Множина 2

[0, 1, -2, 2, 0, -1, -2, 2, -3, -1, 1, 2, 0, 0, 1, 0, 0, 2, -2, 1, 1, 5, -1, 0, 3, 1, -1, 1, 3, 3, -1, 0, 0, 1, 0, 4, 0, 2, -1, 0, -2, 0, 0, 0, 0, 0, 1, 1, 2, 1, 0, 3, 1, -1, 1, 1, -1, -1, 0, 2, 0, 1, 0, 1, -2, 0, 1, 0, 1, 1, 1, 0, 2, 0, 0, -7, -9, -7, -10, -10, -7, -10, -9, -9, -10, -6, -9, -11, -6, -10, -9, -10, -8, -9, -10, -10, -10, -10, -9, -10]

Рисунок 3.2 – Значення інформативності для експериментальних множин векторів

За отриманими оцінками відібрано по 50 найбільш інформативних елементів кожної із множин із загального їх числа 100. Для них отримано такі значення метрик: $X = 13$, $T = 1$.

Бачимо, що за результатами редукції множин (у 2 рази зменшуються обчислювальні витрати) значення метрики Хаусдорфа збільшилось від 12 для повного опису до 13 для редукованого (трохи зменшилась подібність), а значення метрики Танімото зросло від 0,85 до 1 відповідно.

Зрозуміло, що значення оцінок напряму залежать від складу множин. Для інших експериментальних варіантів вхідних даних, отриманих змішуванням векторів із двох розглянутих множин, відстань Танімото за результатом редукування збільшилася від значення 0,70 для повного опису до 1 для скороченого. Як бачимо, в результаті редукції даних відстані між аналізованими описами зростають, що загалом може сприяти підсиленню результативності класифікації.

У той же час швидкодія оброблення редукованих множин має суттєве покращення. У контексті обчислення відстані між описами для метрики Хаусдорфа швидкодія обчислень склала 5,04 мілісекунди для редукованих

множин у порівнянні з 26 мілісекундами для повних, тобто покращилась у п'ять разів! Для метрики Танімото отримані результати виміру швидкодії становлять 0,294 мілісекунди та 1,45 мілісекунди для редукованого і повного описів відповідно.

Як бачимо, після редукування множин швидкодія обчислення відстаней між описами за отриманими ознаками як для метрики Танімото, так і для метрики Хаусдорфа у п'ять разів більша, ніж при обчисленні відстаней у випадку роботи із повними описами.

Проведене пробне моделювання цілковито підтверджує практичну можливість ефективного впровадження метричного критерію інформативності задля значимого скорочення обчислювальних витрат на класифікацію для описів як множини дескрипторів КТ. Об'єм аналізованих даних у експерименті скорочується у 2 рази, швидкодія оброблення зростає у п'ять разів, а рівень відмінностей між описами змінюється незначно, і навіть трохи зростає, всі ці фактори загалом сприяють забезпеченню результативної класифікації.

ВИСНОВКИ

У роботі розроблено та реалізовано метод редукції системи структурних ознак зображення на основі метричного критерію інформативності.

Метричні критерії інформативності даних, що застосовуються на етапі попереднього аналізу, приносять суттєвий вигаш у швидкодії оброблення, наприклад, при класифікації зображень за множиною дескрипторів ключових точок.

Для перевірки результативності методу редукції ознак було використано метрики Танімото і Хаусдорфа для визначення подібності множин ознак до і після редукції. Експериментальним шляхом досліджено роботу методу. У якості вхідних даних було обрано еталоні класи множин бінарних дескрипторів ключових точок. Реалізована модель проведення редукції за метричними критеріями інформативності, яка була використана для вхідних даних. У якості метрики для обчислення інформативності було вирішено обрати метрику Хемінга.

За отриманими результатами дослідження можна побачити, що редуковані множини ознак зменшилися удвічі, а також роботи з ними на прикладі обчислення метрик Танімото і Хаусдорфа пришвидшилась у п'ять разів, що підтверджує ефективність редукції для стиснення опису множини дескрипторів. Множини ознак еталонних класів незначно віддалилися одна від одної за метриками Танімото і Хаусдорфа, що дає можливість зробити висновок про відсіювання сторонніх ознак або ознак з негативним впливом.

На основі цього можна зробити висновок, що досліджуваний метод дозволяє не тільки скоротити витрати ресурсів на обробку зображень, а визначати важливі ознаки класифікації.

Наукова новизна роботи полягає у введенні результативних засобів стиснення еталонних даних, що значним чином прискорює класифікацію зображень з використанням структурних методів.

Прикладна цінність роботи полягає у розробленні програмних моделей для стиснення множини ознак, класифікації з використанням засобів стиснення, а також у експериментальному підтвердженні результативності модифікованих описів у задачі класифікації.

Перспективи роботи полягають у застосуванні запропонованих принципів аналізу даних безпосередньо до експериментальних множин дескрипторів, отриманих для реальних зображень. Використання метричного критерію гуртової інформативності теж потребує подальшого вивчення.

Результати роботи апробовано у вигляді тез доповідей під час Міжнародного молодіжного форуму «РАДІОЕЛЕКТРОНІКА І МОЛОДЬ У ХХІ СТОЛІТТІ» [50].

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Гороховатський В.О., Творошенко І.С. (2022) Аналіз багатовимірних даних за описом у формі множини компонент: монографія. Харків: ХНУРЕ. 124 с.
2. Довбиш А. С., Шелехов І.В. (2015) Основи теорії розпізнавання образів : навч. посібник.
3. Новикова Н. М. (2008) Структурное распознавание образов :учебное пособие для вузов.
4. Гороховатский В. А. (2014) Структурный анализ и интеллектуальная обработка данных в компьютерном зрении: монография. *Харків: ХНУРЕ.*
5. Шлезингер М., Главач В. (2004) Десять лекций по статистическому и структурному распознаванию.
6. Пилип'юк В.В. (2015) Визначення ключових точок у HDR-зображеннях, *Technical sciences*, с. 41 – 45.
7. David G. Lowe (2004), Distinctive Image Features from Scale-Invariant Key points, University of British Columbia press, Vancouver.
8. Lowe G. (1999) Object recognition from local scale-invariant features, *Proceedings of the International Conference on Computer Vision*.
9. Brown M. (2004), Recognizing Panoramas in slides, University of British Columbia press, Vancouver.
10. Пименова М.Б. (2019) Распознавание характерных объектов на местности с использованием метода SURF, *Политехнический молодежный журнал*, с. 1-13.
11. Varsha Patil, Tanuja Sarode.(2019) Image Hashing Using DWT-CSLBP, *Journal of Computers*, pp. 210-223.
12. Гончаренко М.О. (2015) Сравнительный анализ методов формирования дескрипторов изображений в контексте задачи сегментации видеопотока, *Біоніка інтелекту: наук.-техн. Журнал*, 85(2). – pp. 90–94.

13. R. Edward, P. Reid, D. Tom. (2008) Faster and better: a machine learning approach to corner detection, Cambridge University Engineering Department.
14. Edward Rosten. (2005) Fusing Points and Lines for High Performance Tracking/ E. Rosten, T. Drummond, Cambridge University Engineering Department.
15. Гороховатський, В.О., Власенко Н.В. (2021) Редукція опису зображення у складі множини дескрипторів на основі метричного критерію інформативності. Сучасні інформаційні системи, т. 5, №4, с. 10–16.
16. Путятин Е.П., Аверин С.И. (1990) Обработка изображений в робототехнике. Москва: Машиностроение, 320 с.
17. Гороховатський, В.О., Творошенко, І.С. (2021) Методи інтелектуального аналізу та оброблення даних: навч. посібник. Харків: ХНУРЕ.
18. Gorokhovatskiy V.A., Gorokhovatskiy A.V., Peredrii Ye.O. (2018) Hashing of Structural Descriptions at Building of the Class Image Descriptor, Computing of Relevance and Classification of the Visual Objects. pp. 1159–1168.
19. Gorokhovatskiy V.A. (2018) Image Classification Methods in the Space of Descriptions in the Form of a Set of the Key Point Descriptors. Telecommunications and Radio Engineering, 77 (9), pp. 787-797.
20. Гороховатский В.А., Путятин Е.П., Столяров В.С. (2017) Исследование результативности структурных методов классификации изображений с применением кластерной модели данных. Радиоэлектроника, информатика, управление, №3 (42). С. 78–85.
21. Gorokhovatskiy, V.A., (2014), Structural analysis and intellectual data processing in computer vision, SMIT, Kharkiv, p 316.
22. Гороховатский В.А., Передрий Е.О. (2009) Корреляционные методы распознавания изображений путем голосования систем фрагментов. Радиоэлектроника. Информатика. Управление, №1(20), с.74–81.

23. Путятін Є.П., Гороховатський В.О., Матат О.О. (2006) Методи та алгоритми комп'ютерного зору: навч. посібник.
24. Гороховатський В.О., Гадецька С.В., Стяглик Н.І. (2019) Вивчення статистичних властивостей моделі блочного подання для множини дескрипторів ключових точок зображень. *Радіоелектроніка, інформатика, управління*, №2, с. 100–107.
25. Gorokhovatsky V.A. Putyatin Y. P. (2009) Image Likelihood Measures of the Basis of the Set of Conformities. *Telecommunications and Radio Engineering*, 68 (9), p. 763–778.
26. Гороховатський В.О., Пупченко Д.В., Солодченко К.Г. (2018) Аналіз властивостей, характеристик та результатів застосування новітніх детекторів для визначення особливих точок зображення. *Системи управління, навігації та зв'язку*. С. 93–98.
27. L. Stefan, C. Margarita, Y. Roland. (2011) BRISK: Binary Robust Invariant Scalable Keypoints.
28. Alexandre A. (2011) FREAK: Fast Retina Keypoint.
29. Гороховатський, В.О., Власенко, Н.В., Рибалка, М.О. (2021) Застосування засобів хешування даних для прискорення класифікаційних рішень у структурних методах розпізнавання зображень. *Сучасні інформаційні системи*, т. 5, №2, с. 13–20.
30. Gorokhovatskyi O., Gorokhovatskyi V., Peredrii O. (2018) Analysis of Application of Cluster Descriptions in Space of Characteristic Image Features. p. 52.
31. Гороховатский В.А. (2003) Распознавание изображений в условиях неполной информации, Харків: ХНУРЭ, – 112с.
32. Gorokhovatskyi, V., Rusakova, N., Tvoroshenko, I. (2020) The application of image analysis methods and predicate logic in applied problems of magnetic monitoring. *Telecommunications and Radio Engineering*, 79 (20), pp. 1801-1811.

33. Gorokhovatsky, V.O. and Gadetska, S.V., (2019) Determination of Relevance of Visual Object Images by Application of Statistical Analysis of Regarding Fragment Representation of their Descriptions, *Telecommunications and Radio Engineering*, 78 (3), pp. 211–220.
34. M. A. Ahmad, V. Gorokhovatskyi, I. Tvoroshenko, N. Vlasenko, S. Kh. Mustafa (2021) The Research of Image Classification Methods Based on the Introducing Cluster Representation Parameters for the Structural Description, *International Journal of Engineering Trends and Technology*, 69(10), pp. 186-192.
35. Gorokhovatskyi V., Putyatin Y., Gorokhovatskyi O, Peredrii O. (2018) Quantization of the Space of Structural Image Features as a Way to Increase Recognition Performance. The Second IEEE International Conference on DataStream Mining & Processing 21-25 August 2018, Lviv, Ukraine. – pp. 464 – 467.
36. Daradkeh Y.I., Gorokhovatskyi V., Tvoroshenko I., Al-Dhaifallah M., (2022) Classification of Images Based on a System of Hierarchical Features, *Computers, Materials & Continua*, 72(1), pp. 1785-1797.
37. Gadetska, S.V., Gorokhovatskyi, V. O., Stiahlyk, N. I., Vlasenko, N.V. (2021) Statistical data analysis tools in image classification methods based on the description as a set of binary descriptors of key points. *Radio Electronics, Computer Science, Control*, №4, pp. 58-68.
38. Гороховатський В.О., Гадецька С.В., Стяглик Н.І., Власенко Н.В. (2020) Класифікація зображень на підставі ансамблю статистичних розподілів за класами еталонів для компонентів структурного опису. *Радіoeлектроніка, інформатика, управління*, №4, с. 85–94.
39. Gorokhovatskyi V.A., Zamula A.A. (2016) Employment of Intelligent Technologies in Multiparametric Control Systems. *Telecommunications and Radio Engineering*. Vol. 75, No 19, p. 1775–1785.
40. Tvoroshenko I., and Gorokhovatskyi V. (2022) The Application of Hybrid Intelligence Systems for Dynamic Data Analysis, *International Journal of Engineering and Information Systems*, 6(2), pp. 40-48.

41. Tvoroshenko, I.S., 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.
42. Гороховатський В.А., Ересько Ю.Н., Путятин Е.П., Стрельченко В.И. (1990) Локалізація об'єктів на зображеннях візуальних сцен. *Авто-метрия*, 6, с.3–7.
43. Гадецька С.В., Гороховатський В.О., Стяглик Н.І. (2020) Вивчення критеріїв інформативності даних при впровадженні апарату дерев рішень у методах структурної класифікації зображень. *Радіоелектроніка, інформатика, управління*, №3, с. 78–87.
44. Gorokhovatskiy V.A. (2011) Compression of Descriptions in the Structural Image Recognition. *Telecommunications and Radio Engineering*, 70 (15), pp. 1363–1371.
45. Гороховатський, В.О., Гадецька, С.В. (2020) Статистичне оброблення та аналіз даних у структурних методах класифікації зображень (монографія), Харків.
46. Kohonen, T., (2001) *Self-Organizing Maps*, Heidelberg, Berlin: Springer-Verlag, 502 p.
47. Scherer R. (2018) *Computer Vision Methods for Fast Image Classification and retrieval*, Częstochowa, *Springer*, p 137.
48. Gorokhovatskiy V., Gadetska S., Ponomarenko R. (2020) Recognition of Visual Objects Based on Statistical Distributions for Blocks of Structural Description of Image. *Lecture Notes in Computational Intelligence and Decision Making. Proceedings of the XV International Scientific Conference “Intellectual Systems of Decision Making and Problems of Computational Intelligence” (ISDMCI’2019)*, Ukraine, May 21–25, 2019, pp. 501-512.

49. Гороховатский В.А. (2008) Иерархия пространственных отношений структурных признаков в задачах сопоставления визуальных объектов. Системы управління, навігації та зв'язку: зб. наук. пр., 3(7), с. 85–89.

50. Чефранов І.О. (2022) Засоби стиснення опису у методах класифікації зображень. 26-й Міжнародний молодіжний форум «Радіoeлектроніка і молодь в ХХІ столітті».