

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
РАДІОЕЛЕКТРОНІКИ**

**Володимир ГОРОХОВАТСЬКИЙ
Ірина ТВОРОШЕНКО**

**ПРОДУКТИВНІ МОДЕЛІ АНАЛІЗУ
ДАНИХ У МЕТОДАХ РОЗПІЗНАВАННЯ
ЗОБРАЖЕНЬ**

Монографія

**Харків
2026**

УДК 004.932.2:004.93'1

*Рекомендовано до друку Науково-технічною радою ХНУРЕ
(протокол від 10.12.2025 р. № 11/7)*

Рецензенти:

- Є. В. Бодянський,* доктор технічних наук, професор, професор кафедри штучного інтелекту ХНУРЕ, Заслужений діяч науки і техніки України.
- Г. А. Кучук,* доктор технічних наук, професор, професор кафедри комп'ютерної інженерії та програмування НТУ «ХПІ», лауреат Національної премії України імені Бориса Патона.

Гороховатський Володимир, Творошенко Ірина

Продуктивні моделі аналізу даних у методах розпізнавання зображень: монографія / В. Гороховатський, І. Творошенко. – Харків: ХНУРЕ, 2026. – 153 с.

ISBN 978-966-659-347-7

Монографія розвиває структурні технології аналізу багатовимірних даних в інтелектуальних системах комп'ютерного зору. Основна увага приділяється удосконаленню структурних методів розпізнавання зображень за критеріями підвищення швидкодії і скорочення обчислень. Запропоновано ряд модифікацій для класифікаторів зображень на підставі множини дескрипторів ключових точок, проведено їх теоретичне обґрунтування. Досліджуються моделі стиснення опису за критерієм інформативності, вивчаються особливості впровадження мережі Кохонена, кластерного подання даних та метричної грануляції. Запроваджено індексовані структури даних для швидкісного пошуку у процесі класифікації, використано модифікації простору ознак у формі матриці відстаней та апарату редукції із забезпеченням високого рівня точності класифікації. Наведено результати програмного моделювання у порівнянні із традиційними методами. Підтверджено ефективність запропонованих інновацій для прикладних баз даних та зображень.

Рекомендовано для науково-технічних працівників та здобувачів вищої освіти, що розробляють чи вивчають новітні інформаційні технології, а також інтелектуальні методи аналізу багатовимірних даних. Монографія може бути корисною для здобувачів вищої освіти за спеціальністю «Комп'ютерні науки» у галузі «Інформаційні технології».

УДК 004.932.2:004.93'1

ISBN 978-966-659-347-7

DOI: 10.30837/978-966-659-347-7

© Гороховатський Володимир, Творошенко Ірина, 2026
© Харківський національний університет
радіоелектроніки, 2026

ЗМІСТ

ВСТУП.....	5
1 СТИСНЕННЯ ОПИСУ ЗА КРИТЕРІЄМ ІНФОРМАТИВНОСТІ.....	11
1.1 Аналіз проблеми редукції опису.....	11
1.2 Впровадження параметрів значущості для компонентів опису.....	17
1.3 Способи підрахунку голосів класів.....	20
1.4 Здійснення редукції опису.....	22
1.5 Визначення параметрів значущості.....	24
1.6 Критерії ефективності класифікації.....	25
1.7 Результати експериментів та обговорення.....	26
2 ЗАСТОСУВАННЯ НЕЙРОННИХ МЕРЕЖ КОХОНЕНА ДЛЯ НАВЧАННЯ КЛАСИФІКАТОРА ЗОБРАЖЕНЬ.....	38
2.1 Мережа Кохонена у навчанні класифікатора.....	38
2.2 Формалізація класифікатора.....	42
2.3 Застосування мережі Кохонена для отримання центрів даних.....	43
2.4 Класифікація на підґрунті навчання.....	45
2.5 Способи і умови здійснення навчання.....	48
2.6 Аналіз результатів експериментів.....	50
3 ВПРОВАДЖЕННЯ КЛАСТЕРНОГО ПОДАННЯ ОПИСУЗОБРАЖЕНЬ.....	56
3.1 Кластерне подання як індексована структура даних.....	56
3.2 Модель кластерного подання структурного опису.....	60
3.3 Способи класифікації на підставі центрів кластерів.....	61
3.4 Експерименти і обговорення результатів.....	66
4 ІДЕНТИФІКАЦІЯ ВІЗУАЛЬНИХ ОБ'ЄКТІВ ЗА КЛАСТЕРНИМ ПОДАННЯМ.....	74
4.1 Аналіз літературних джерел.....	74
4.2 Математичні моделі ідентифікації опису для зображення запиту.....	76
4.3 Експериментальні результати програмного моделювання.....	79
5 ГРАНУЛЯЦІЙНІ МОДЕЛІ СТИСНЕННЯ ОПИСУ ЗОБРАЖЕНЬ.....	85
5.1 Удосконалення шляхом стиснення опису.....	85
5.2 Моделі метричної грануляції.....	88
5.3 Варіанти кластерного подання.....	92
5.4 Схеми і критерії класифікації.....	93

5.5	План експериментів та критерії результативності.....	97
5.6	Експериментальні результати і обговорення.....	98
6	ШВИДКІСНИЙ МЕТРИЧНИЙ ПОШУК ДАНИХ НА ОСНОВІ ХЕШУВАННЯ.....	104
6.1	Хеш-структури для подання опису.....	104
6.2	Формалізм задачі класифікації.....	106
6.3	Трансформація даних із впровадженням хеш-структури для швидкісного пошуку.....	109
6.4	Класифікатор на основі синтезованої структури.....	112
6.5	Результати комп'ютерного моделювання.....	114
7	ТРАНСФОРМАЦІЯ ОПИСІВ ЗОБРАЖЕНЬ ЗАДЛЯ ПОБУДОВИ КЛАСИФІКАЦІЙНИХ ОЗНАК.....	119
7.1	Матриця відстаней та ортогональні перетворення.....	119
7.2	Подання опису в ортогональному базисі та класифікація.....	123
7.3	Редукція у модифікованому просторі.....	126
7.4	Формалізм матриці відстаней для множини дескрипторів.....	129
7.5	Моделі класифікації на основі матриці відстаней.....	131
7.6	Результати комп'ютерного моделювання.....	134
	ВИСНОВКИ.....	141
	ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ.....	142

ВСТУП

У сучасних системах комп'ютерного зору набули прикладного застосування методи розпізнавання візуальних об'єктів, що засновані на представленні зображення у формі множини багатовимірних векторів.

До них належать структурні методи, що базуються на визначенні множини ключових точок (КТ) зображення та їх описові у вигляді числового чи бінарного вектора – дескриптора, що відображає властивості функції яскравості зображення для локальних околів КТ [1–16]. Працездатність та ефективність цього підходу підтверджена практичними результатами [1–16].

Розпізнавання (класифікація) при застосуванні структурних методів зосереджується на обчисленні міри релевантності між описами об'єкта та еталона, образи яких задані у вигляді множин дескрипторів, та оптимізації значень цієї міри на множині еталонів [12, 14, 15].

Монографія розвиває підхід структурного аналізу в плані покращення основних показників функціонування систем комп'ютерного зору – швидкодії класифікації зображень із забезпеченням достойного рівня точності.

Матеріал монографії продовжує впровадження технологій аналізу багатовимірних даних в інтелектуальних системах [12–16] і ґрунтується на результатах недавніх досліджень авторів [1–11, 17, 18], де детально викладене теоретичне підґрунтя та експериментальне підтвердження ефективності для ряду продуктивних модифікацій структурних методів.

Коротко основний зміст нашого дослідження у межах окремих розділів полягає в наступному.

У розділі 1 основна увага приділяється підвищенню швидкодії процесу встановлення релевантності описів об'єкта та еталону зі збереженням потрібного рівня ефективності класифікації. Розпізнаваний клас представляється нескінченною множиною зображень, отриманих із еталону застосуванням довільних геометричних перетворень у полі зору. Пропонується здійснити редукцію описів для бази еталонів шляхом відбору найбільш значущих компонентів-дескрипторів за критерієм інформативності. Інформативність дескриптора оцінюється різницею найближчих відстаней до свого і альтернативного описів. Розроблений швидкісний спосіб полягає у встановленні ступеня релевантності повного опису розпізнаваного об'єкта із редукованим описом еталонів. Тестова вибірка сформована як сукупність зображень із бази еталонів і поза базою із застосуванням до неї геометричних перетворень масштабу і повороту у полі зору.

Проведене комп'ютерне моделювання виявило практичну можливість редукції описів приблизно у десять разів із повним збереженням точності класифікації. Скорочення описів у двадцять разів у експерименті призводить до незначного зниження точності. Швидкодія аналізу при цьому підвищується пропорційно ступеню редукції. Застосування редукції описів за критерієм інформативності підтвердило можливість отримати найбільш значущу для класифікації підмножину ознак дескрипторів, що гарантує необхідний рівень точності.

У розділі 2 основна увага приділяється впровадженню навчання класифікатора з використанням апарату комітету мереж Кохонена роздільно для кожного опису у базі еталонів. Як результат навчання – кожний еталон компактно подається фіксованою множиною центроїдів даних, що забезпечує високу швидкодію здійснення класифікації. Удосконалення полягає у впровадженні незалежного навчання для кожного еталону, що підвищує точність апроксимації описів набором центроїдів і загалом гарантує збереження ефективності класифікації на високому рівні. Обчислюючи при навчанні центроїди кластерів для кожного класу окремо, ми запобігаємо міжкласовому впливу дескрипторів альтернативних класів. Крім цього, роздільне навчання є дієвим у ситуації з різною потужністю для описів еталонів. Швидкодія класифікації для запропонованого методу на підставі навчання ансамблю мереж Кохонена у порівнянні з традиційним способом лінійного пошуку зростає пропорційно відношенню потужності опису до числа сформованих центроїдів.

Розглянуто декілька моделей навчання мереж і відповідних способів класифікації в умовах геометричних перетворень вхідних зображень. Наведено результати експериментального моделювання пропонуваніх методів для бази, що включає зображення монет. Тестова вибірка сформована як сукупність зображень із бази еталонів із застосуванням до них геометричних перетворень зсуву, масштабу і повороту у полі зору.

Досліджено практичні питання вибору параметрів навчання задля забезпечення необхідного рівня точності класифікації. Тестування показало високий рівень точності класифікації на підставі впровадженого навчання і виявило практичну можливість обрати таку структуру і параметри мережі, що універсально забезпечать потрібний рівень критеріїв точності і швидкодії класифікації у прикладній задачі.

Впровадження удосконалених методів доцільно у прикладних задачах комп'ютерного зору з високими вимогами до швидкодії класифікації візуальних об'єктів.

У розділі 3 представлено результати розроблення ефективних за швидкодією методів класифікації на підставі впровадження і аналізу кластерного подання для множини дескрипторів ключових точок. Запропоновано три варіанти класифікатора за принципами «об’єкт – еталон», «дескриптор об’єкту – еталон» та «векторний опис об’єкту – еталон», які відрізняються ступенем інтегрованості способу аналізу. Застосовано варіанти побудови системи кластерів на усій множині описів бази еталонів, окремо для кожного із еталонів, а також оптимальний спосіб зіставлення наборів кластерних центрів для об’єкту та еталонів.

Здійснено експериментальне оцінювання ефективності розроблених класифікаторів щодо працездатності, необхідного часу обчислень та ефективності класифікації за результатами оброблення прикладної бази зображень. Запропоновані методи класифікують множину еталонів безпомилково. Зроблено висновок про високу продуктивність і результативність технологій класифікації із застосуванням кластерних центрів. Отримано час опрацювання зображень за розробленими методами у сотні разів менший, ніж за традиційним, без зменшення точності.

У розділі 4 запропоновано новий метод пошуку об’єктів на зображенні. Цей метод ґрунтується на ідентифікації кластерного подання описів запиту та поточного зображення вікна із обчисленням міри релевантності. Впровадження кластерного подання для описів запиту і аналізованого зображення всередині вікна дає можливість суттєво підвищити швидкодію ідентифікації чи класифікації візуальних об’єктів. Підхід розвиває апарат структурних методів у аналізі зображень, що використовують опис вхідного зображення у вигляді множини дескрипторів. На підставі розроблених математичних моделей для оброблення множини дескрипторів отримано результативний засіб для ідентифікації зображень візуальних об’єктів.

Розроблений метод може бути застосовано для пошуку одночасно кількох об’єктів. Впроваджено програмну модель аналізу візуальних даних для розробленого методу у програмному середовищі Jupyter Notebook на сервісі Google Colaboratory. Використано спеціалізовані бібліотеки для роботи із зображеннями: Scikit-image та OpenCV. Здійснено оцінювання працездатності та ефективності методу за результатами оброблення конкретних прикладних зображень. Проведено порівняльний експеримент із традиційним методом голосування, де для ідентифікації реалізовано лінійний пошук найближчого дескриптора без застосування попередньої процедури кластерного подання опису.

У експерименті отримано вигаш за показником швидкодії для розробленого методу у порівнянні із традиційним приблизно у 5,2 рази. Таким чином, експериментально підтверджено прикладну результативність розробленого методу пошуку із впровадженням процедури кластерування. Метод може бути використано у прикладних задачах, де час ідентифікації об'єктів є обмеженим. Експеримент показав, що ефективність методу можна посилити за рахунок варіації значень його параметрів та пристосування до властивостей даних. Подальшим етапом досліджень може бути впровадження ієрархічного подання даних у вигляді узагальнених ознак структурного опису.

У розділі 5 пропонуються грануляційні способи для скорочення обчислювальних затрат у процесі класифікації.

Основна увага приділяється впровадженню грануляції, відсіювання та кластеризації для множини елементів еталонних описів. Як результат стиснення, кожний еталон компактно подається скороченою множиною дескрипторів чи центроїдів даних, що забезпечує високу швидкодію класифікації.

Обговорюються кілька варіантів простих схем стиснення даних у порівняльному аспекті з методом лінійного пошуку та двома варіантами кластеризації еталонів – сукупно для усієї бази та роздільно для кожного із образів. Наведено результати програмного моделювання пропонованих підходів для двох експериментальних баз, що включають зображення логотипів футбольних клубів та художніх картин. Тестова вибірка сформована як сукупність зображень із бази еталонів та зображень поза базою із застосуванням до них набору геометричних перетворень зсуву, масштабу і повороту у полі зору.

Досліджено практичні питання вибору параметрів порогів для встановлення еквівалентності дескрипторів та для мінімуму числа голосів класів задля забезпечення необхідного рівня точності класифікації. Тестування підтвердило суттєве прискорення процесу оброблення і достатньо високий рівень точності класифікації з використанням пропонованих процедур стиснення. Конкретно у проведеному моделюванні досягнуто підвищення швидкодії у десятки разів.

Експериментально підтверджено, що застосування апарату кластерування має значно вищий потенціал щодо показників точності та швидкодії класифікації у порівнянні з більш простими схемами відсіювання чи грануляції на основі близькості компонентів.

У розділі 6 обговорюється новий метод класифікації, заснований на впровадженні засобів швидкісного пошуку у створеній на еталонній множині ознак індексованій структурі даних. Метод суттєво переважає у швидкодії оброблення традиційні підходи. Візуальний об'єкт описується набором багатовимірних векторів, а класифікація заснована на обчисленні ступеня релевантності двох багаточисленних за обсягом множин. Класифікатор подано як двох-етапне оброблення, де на першому етапі визначається клас для окремого дескриптора об'єкта, а на другому етапі на підставі отриманої множини локальних рішень статистично визначається результуючий клас об'єкта.

Розроблений метод заснований на попередній побудові індексованих хеш-структур для множини дескрипторів бази еталонних образів. Хешування реалізує перспективну ідею зниження розмірності аналізованих даних шляхом попередньої побудови для еталонної множини деякої індексованої структури, що спрощує оброблення та знижує обчислювальні затрати на класифікацію.

На етапі класифікації опису застосовано узгоджене ланцюжкове поєднання етапів аналізу даних: розбиття на блоки, пошук всередині блоків з обмеженням на значення хеш-функції. Впровадження хеш-подання дає можливість суттєво підвищити швидкодію ідентифікації чи класифікації візуальних об'єктів.

Програмне моделювання виконано у середовищі PyCharm Community Edition із використанням засобів бібліотеки OpenCV та мови програмування Python. Здійснено оцінювання працездатності та ефективності методу за результатами оброблення набору прикладних зображень.

Проведено порівняльний експеримент із традиційним методом голосування, де для ідентифікації реалізовано лінійний пошук без застосування індексованого хеш-подання еталонів. У експерименті отримано виграш за показником швидкодії для розробленого методу у порівнянні із традиційним більше ніж у 10 разів. Виграш у швидкодії пропорційно зростає зі збільшенням числа еталонів та кількості дескрипторів у описах.

Експериментально підтверджено прикладну результативність розробленого методу класифікації із впровадженням індексованої хеш-структури. Метод може бути використано у прикладних задачах, де є обмеження на час ідентифікації об'єктів. Експеримент показав, що ефективність методу можна посилити за рахунок варіації значень його параметрів.

Подальшим етапом досліджень може бути впровадження логічного оброблення розподілів даних всередині хеш-структури, поглиблене вивчення завадостійкості розроблених методів та оцінювання їх прикладної результативності стосовно об'ємних колекцій зображень.

У розділі 7 наведено результати дослідження щодо побудови модифікацій для простору класифікаційних ознак на основі застосування перетворення структурного опису через декомпозицію в ортогональному базисі та реалізації моделі матриці відстаней між компонентами опису.

Застосування системи ортогональних функцій як апарату перетворення опису показало можливість значного підвищення швидкодії оброблення при збереженні високих показників точності класифікації та перешкодостійкості.

Підтверджено ефективність синтезованих ознакових систем за значним збільшенням швидкодії та достатнім рівнем ефективності. Експериментальний приклад показав, що час, витрачений на обчислення релевантності описів відповідно до їх модифікованого представлення, більш ніж у десять разів менший, ніж для традиційних метричних підходів. Розроблені класифікаційні ознаки можуть бути використані в прикладних задачах, де критичним є час візуальної ідентифікації об'єктів.

Автори висловлюють щире подяку аспірантам та студентам кафедри інформатики Харківського національного університету радіоелектроніки, які є співавторами ряду публікацій і протягом останніх років брали дієву участь у розробленні програмних продуктів, здійсненні програмного моделювання та дослідженні результативності розроблених модифікацій методів із впровадженням їх у прикладних базах даних для розпізнавання зображень.

Автори вдячні також керівництву кафедри інформатики за дієву підтримку досліджень, результати яких втілено у матеріалах монографії.

1 СТИСНЕННЯ ОПИСУ ЗА КРИТЕРІЄМ ІНФОРМАТИВНОСТІ

1.1 Аналіз проблеми редукції опису

У задачах розпізнавання зображень, що є зараз досить актуальними для комп'ютерного зору, систему ознак часто формують як множину багатовимірних векторів, що у повній мірі відображають просторові властивості візуального об'єкту для здійснення результативного аналізу [1, 2]. Наприклад, опис візуального об'єкта подається як скінченна множина дескрипторів ключових точок зображення [3–5].

Дескриптор – це багатокomпонентний числовий вектор, що відображає характеристики деякого околу ключових точок на зображенні і формується спеціальними фільтрами – детекторами [6, 7].

Для людського зору різні фрагменти зображення мають різну вагомість у процесі аналізу чи класифікації [8–10]. Людський зір на підґрунті використання розумової діяльності здатен розпізнавати об'єкти навіть за їх окремими невеликими за розміром деталями. Незважаючи на те, що штучні системи комп'ютерного зору зараз досить ефективно розпізнають зображення за дещо іншими принципами, ніж людський зір, класифікаційна важливість складових компонентів образу об'єкту у них теж суттєво різниться і може бути ефективно використана у процесі аналізу.

Фактори, які характеризують важливість (вагомість, значущість, значимість) ознак: розкид значень всередині опису і в межах бази еталонних описів, ступінь стійкості до геометричних перетворень і дії завад, унікальність та цінність присутності на зображенні і т. д. Але конкретний рівень параметру вагомості, як і результативність класифікації загалом, значною мірою визначається базою еталонних зображень, які треба класифікувати.

Введення параметру значущості в процесі класифікаційного аналізу дає змогу перейти від однорідного впливу ознак до врахування їх відносної зваженості, що напряду впливає на такі характеристики класифікації, як продуктивність та результативність [3, 11, 12].

Як правило, врахування вагомості окремих ознак зображення у процесі класифікації здійснюють шляхом уживання вагових коефіцієнтів у числовому вимірі.

Значущість апріорно можна оцінити для наявної множини компонентів еталонних описів, її використання націлено на більшу адаптацію класифікації до аналізованих даних із-за розширення обсягу використовуваної інформації [2, 13]. Зрозуміло, що на ефективність впровадження параметрів значущості впливає також форма подання простору ознак і застосований спосіб класифікації.

Якщо попередньо здійснити аналіз обчислених значень ваг для еталонних компонентів описів бази зображень, можна конструювати описи класів із найбільш цінних для класифікації елементів, відкинувши неінформативну частину опису [2, 14, 15]. Однією із дієвих оцінок класифікаційної значущості є параметр інформативності ознаки [2]. Інформативність ознаки є тим вищою, чим краще ця ознака розділяє екземпляри навчальної вибірки на класи.

Це не тільки пропорційно ступеню скорочення знижує обчислювальні затрати, але й зберігає і часто підвищує ефективність функціонування систем класифікації. На підставі впровадження вагових коефіцієнтів, зокрема, параметру інформативності, з'являється можливість здійснення більш глибокого аналізу даних, що загалом покращує показники точності розпізнавання.

Мета роботи – зменшення обсягу обчислювальних затрат при впровадженні структурних методів класифікації зображень зі збереженням їх результативності. Скорочення опису у формі множини дескрипторів ключових точок досягається шляхом формування підмножини дескрипторів за критерієм класифікаційної значущості.

Завдання дослідження полягають у наступному:

- вивчення впливу параметру інформативності для множини дескрипторів еталонних описів на результативність класифікації зображень;
- здійснення редукції еталонних описів на підставі значення інформативності;
- впровадження параметру інформативності у модель класифікатора;
- експериментальне дослідження результативності розроблених модифікацій класифікаторів за критеріями точності і швидкодії оброблення на підставі імітаційного моделювання для прикладної бази зображень.

Розглянемо множину $E = \{E_i\}_{i=1}^N$ описів бази еталонів $E = \{\{e_v(i)\}_{v=1}^s\}_{i=1}^N$, де N – число класів, $i = 1, \dots, N$ – номер класу, s – потужність множини компонентів окремого опису, $v = 1, \dots, s$ – номер компонента всередині опису класу E_i .

Таким чином, модель даних для бази класифікації E можна представити у формі матриці розміром $N \times s$, де кожний рядок містить опис окремого еталону (класу) у складі s компонентів – дескрипторів ключових точок [15].

Опис $Z = \{z_v\}_{v=1}^s$ розпізнаваного об'єкта подається моделлю, аналогічною опису окремого еталону E_i . У новітніх детекторах ключових точок використовують бінарні подання для дескрипторів ключових точок, що дає суттєві переваги у обсязі обчислень [1, 6]. Будемо вважати $z_v \in B^n$, $e_v \in B^n$, $E_i \subset B^n$, $Z \subset B^n$, де B^n – простір векторів розмірності n з бінарними компонентами.

Скінченна множина бінарних векторів (дескрипторів), отримана детектором ключових точок, створює інваріантний до перетворень опис еталонного чи розпізнаваного зображення [16].

Еталон – це обране зображення, для якого формується опис. Множина еталонів створює базу для класифікації. Формально розпізнаваний клас у такій постановці [1, 4] має вид нескінченної множини зображень, отриманої із зображення конкретного еталону застосуванням до нього всіляких наборів геометричних перетворень зсуву, масштабу, повороту.

Для кожного елемента $e_v(i) \in E$ бази еталонів на підготовчому етапі аналізу даних визначимо величину параметра значущості $\lambda_{i,v}$ у формі числа.

Загалом для бази еталонів E отримаємо матрицю $\Lambda = \{\{\lambda_{i,v}\}_{v=1}^s\}_{i=1}^N$ чисел

$$\Lambda = \left\{ \begin{array}{c} \lambda_{1,1}, \dots, \lambda_{1,s} \\ \lambda_{2,1}, \dots, \lambda_{2,s} \\ \dots \\ \lambda_{N,1}, \dots, \lambda_{N,s} \end{array} \right\}. \quad (1.1)$$

Матриці Λ та E мають ідентичну структуру. Але компонентами E є бінарні вектори, а компоненти Λ – це числа, що характеризують значущість цих векторів для класифікації. У рядках матриці E містяться значення дескрипторів ключових точок для описів окремих класів, а Λ містить кількісний параметр значимості для кожного із дескрипторів.

Зазначимо, що параметр $\lambda_{i,v}$ матриці Λ , як правило, відображає вагомість кожної ознаки $e_v(i)$ всередині окремого класу E_i чи в рамках усієї наявної множини E класів.

У той же час відомі способи [2, 11, 13] визначення елементів $\lambda_{i,v}$ вказують на їх пряму залежність від складу E , так як класифікаційна вагомість компонентів безпосередньо і завжди залежить від набору розпізнаваних класів.

Поставимо задачу розроблення процедури $R: E \rightarrow E^*$ для цілеспрямованого скорочення (редукції) складу опису бази E шляхом відбору найбільш значимих за критерієм матриці Λ елементів, причому $card E^* \ll card E$, $E^* = \{E_i^*\}_{i=1}^N$. Потужність $s^* = card E_i^*$ опису кожного із еталонів E_i при цьому рівноцінно скорочується: $s \rightarrow s^*$, $s^* \ll s$.

Формування новоствореної бази E^* на основі редукції R призначено для здійснення класифікації зі значно меншими обчислювальними затратами із забезпеченням достойного рівня результативності.

Друга нагальна задача полягає у створенні класифікатора F , який для довільного опису $Z = \{z_v\}_{v=1}^s$ об'єкту у формі множини дескрипторів ключових точок прийме рішення $F[Z] \rightarrow [1, 2, \dots, N]$ про належність аналізованого опису Z до одного із N класів із врахуванням значень матриці Λ . У окремих ситуаціях допускається відмова від класифікації.

У обговорюваній постановці класифікація $F[Z]$ здійснюється на основі таких засад: описів Z об'єкту та бази $\{E_i\}_{i=1}^N$, а також визначеної значущості у формі матриці Λ для множини еталонних компонентів. Повна формальна параметризація класифікатора має вид $F[Z, E, \Lambda] \rightarrow [1, 2, \dots, N]$.

При цьому у класифікаторі F допускається використання редукованої бази E^* замість E . Така вимога потребує модифікації для моделі оцінювання релевантності описів Z та E_i^* з нерівнозначним числом елементів.

Відмітимо, що дескриптори ключових точок зображення – це сучасний дієвий апарат для представлення і аналізу описів візуальних об'єктів [6, 7]. Цей апарат ефективно забезпечує інваріантність до поширених на практиці геометричних перетворень візуальних об'єктів на зображеннях, високу швидкодію аналізу даних та достойну результативність класифікації [1, 4].

На відміну від моделей нейронних мереж, що здійснюють узагальнення образу всередині класу, ці методи зосереджуються на виявленні характерних властивостей розпізнаваних зображень, що необхідно у ряді прикладних застосувань [13, 19–22].

Покращення продуктивності структурних методів класифікації, заснованих на зіставленні множин векторів, розвивається у таких аспектах, як пришвидшення процесу пошуку відповідностей компонентів шляхом

кластерування чи хешування даних [4, 19, 23–26], використанням оцінювання для звуження обсягу пошуку [27], формуванням агрегованих ознак у вигляді розподілів [4, 17], а також через визначення найбільш вагомої для класифікації підмножини ознак [2, 14].

Задача формування ефективної підмножини ознак, в тому числі шляхом редукції, постійно знаходиться в полі уваги дослідників комп'ютерного бачення [13, 19, 20] по причині потреби щодо забезпечення продуктивного аналізу і оброблення об'ємних багатовимірних масивів даних у таких системах [18]. У роботах [2, 3, 15] розглядається визначення параметра інформативності щодо набору ознак навчальної множини, а у [15] більш детально вивчаються особливості обчислення інформативності векторів-дескрипторів. У роботах [4, 9, 11] обговорюються способи визначення значущості ознак за принципом унікальності у складі опису та у формі вектора значущості за множиною класів.

У літературі з аналізу даних вивчаються окремі моделі щодо застосування значущості ознак задля підвищення точності класифікації [13, 21]. Для цього при зіставленні аналізованого векторного числового опису $\{a_v\}_{v=1}^s$ з еталонним $\{b_v\}_{v=1}^s$ при встановленій відповідності їх компонент визначається деяка міра $\mu(a, b, \lambda)$, що враховує параметри значущості $\lambda = \{\lambda_v\}_{v=1}^s$ кожної із ознак. Така міра має форму зваженої метрики, наприклад, манхеттенської:

$$\mu(a, b, \lambda) = \sum_{v=1}^s \lambda_v |a_v - b_v|, \quad \sum_{v=1}^s \lambda_v = 1. \quad (1.2)$$

У мірі (1.2) значущість λ_v використовується сумісно із метрикою, причому мається на увазі, що важливіші ознаки вносять більший вклад в прийняття класифікаційного рішення. Зважаючи на те, що метрика в (1.2) при встановленій відповідності дескрипторів a, b налаштована на пошук мінімуму, вектор λ треба адаптувати до такого способу класифікації. Параметр λ повинен підсилювати міру подібності для більш важливих ознак. За визначенням він є більшим для важливіших ознак. Зважаючи на наявну умову $\sum_{v=1}^s \lambda_v = 1$, одним із варіантів моделі (1.2) є використання значення $(1 - \lambda_v)$ замість λ_v . Зауважимо, що при безпосередньому застосуванні (1.2) по схемі лінійного пошуку «кожний з кожним» для описів як множин векторів часто виникає необхідність оперувати із числами з плаваючою комою, що вимагає деякого збільшення обсягу обчислень або здійснення дискретизації значень λ [1, 13].

Вагові коефіцієнти для дескрипторів ключових точок знайшли застосування у ряді ймовірнісних моделей класифікації, де ці коефіцієнти трактуються як значення ймовірності віднесення до класу [4, 20, 28]. За розробленими схемами здійснюється накопичення значущостей тих еталонних елементів, для яких конкурентним способом встановлено відповідність з аналізованою компонентою об'єкта [21]. Впровадження моделі (1.2) за думкою авторів [21] сприяло кращій адаптації процесу аналізу даних у порівнянні з фіксованою сіткою у просторі класифікаційних ознак.

Відмітимо, що апарат вагових коефіцієнтів з успіхом використовуються також у моделях нейронних мереж для класифікації даних [29]. Там вони застосовуються для класифікації як результат навчання мережі [30, 31].

На сьогодні прикладну результативність сучасних нейромережових систем [30, 32] у задачі класифікації зображень дуже важко перевершити. Але ці системи, як відомо, потребують довготривалого навчання на об'ємних наборах уже частково анотованих людиною даних із використанням спеціалізованого програмного забезпечення. При цьому результат і ефективність класифікації нейронними мережами суттєво визначаються їх структурою, а також складом даних, використаних для навчання. Крім того, апарат нейронних мереж традиційно здійснює узагальнення ознак для представників всередині класу [31], що не завжди дає можливість здійснити результативну класифікацію для окремих його представників.

Підходи, що базуються на безпосередньому вимірі ознак зображення у формі множини дескрипторів, мають свої переваги при впровадженні у системах комп'ютерного зору [4, 6, 7, 33–35]. Їх позитивом є відсутність потреби у етапі навчання, а також можливість оперативної і радикальної зміни складу розпізнаваних класів. Найбільш ефективно вони можуть бути впроваджені для ідентифікації чи класифікації стандартизованих зображень (герби, картини, бренди, експонати музеїв) з допустимими довільними геометричними перетвореннями об'єктів у полі зору [3, 23, 33]. Для функціонування таких методів потрібні лише представники класів – еталони. Після нетривалого у часі обчислення описів еталонів метод готовий до використання. Методи універсально підходять для будь-якої обраної множини еталонів, склад якої може оперативно змінюватися під прикладну задачу [1].

Зрозуміло, що апарат дескрипторів ключових точок не має можливості врахувати практично нескінченні варіанти генеративних моделей щодо перетворень при формуванні зображень сучасними нейромережами [36].

Тому ефективність його застосування для класифікації такого різноманіття зображень необхідно вивчати окремо. Можливо, що ці напрямки (нейромережі і структурні методи) досліджень треба розділити за сферами застосування чи прикладними задачами. Наприклад, нейронні мережі важко справляються із мінливістю об'єктів в аспекті геометричних перетворень. А структурні методи чутливі до суттєвої зміни форми об'єктів (морфінг), і при слідуванні за рухомими об'єктами потребують перезапису еталону.

Відомі дослідження [33] щодо визначення числа дескрипторів у описі і вибору детектора ключових точок з найкращими показниками ефективності розпізнавання. Але в такій постановці акцент щодо результату і ефективності класифікації зосереджується виключно на типі і властивостях детектора [37], а не на формі даних і способі їх класифікації.

Таким чином, проведені дослідження вказують на необхідність більш детального вивчення процесу впровадження показників класифікаційної вагомості і оцінювання їх впливу на результативність класифікації за описом у формі множини дескрипторів як на підставі редукції множини ознак, так і безпосереднім використанням у класифікаторах сумісно із метричними відношеннями.

1.2 Впровадження параметрів значущості для компонентів опису

Важливим фактором, що може вплинути на результат класифікації, можна вважати використання індивідуальних величин значущості (1.1) для окремих компонентів опису еталонів [4]. Зважаючи на той факт, що класифікаційне рішення, як правило, ґрунтується на статистичному аналізі для множини значень метричних відношень у просторі дескрипторів, підсилимо його за рахунок введення параметру значущості. Цим фактично розширюється набір параметрів для даних, що аналізуються класифікатором.

Для цього розглянемо одну із практичних схем, в основі якої є первинна класифікація поодиноких дескрипторів об'єкту з наступним накопиченням визначених голосів для їх класів та значущостей [1, 4].

Зважаючи на те, що порядок слідування компонентів у описах еталонів та розпізнаваного об'єкту Z може різнитися із-за можливого впливу геометричних перетворень об'єкту, застосуємо двоетапну схему класифікації, де здійснимо класифікацію $z_v \rightarrow [1, 2, \dots, N]$ для кожного компонента $z_v \in Z$.

Ця схема вигідно відрізняється від інтегрального подання об'єктів [17, 23] тим, що у ній є можливість відбракувати хибні дескриптори фону і завад, що забезпечує кращу завадостійкість [11].

Відповідно до традиційної схеми найближчого сусіда здійснимо класифікацію шляхом обчислення мінімуму

$$k = \arg \min_{v,i} \rho(z, e_v(i)), \quad (1.3)$$

де $\rho(z, e_v(i))$ – відстань між дескрипторами об'єкту та еталонів;

$k = 1, \dots, N$ – визначений клас для дескриптора z , що співвідноситься з класом еталонного елемента з найменшою відстанню.

Процес визначення класу k у (1.3) – це двохпараметрична оптимізація з параметрами номера i класу та номера v еталонного дескриптора всередині класу. Основною метрикою, що використовується для оцінювання відхилення пари бінарних дескрипторів, обрана ефективна в обчислювальному аспекті відстань Хемінга [2, 19], що визначає кількість відмінних бітів.

Модель впровадження значущості багато в чому залежить від обраного способу реалізації (1.3).

У відповідності до традиційної моделі (1.3) для кожного дескриптора $z \in Z$ знаходимо мінімум ρ_m відстані $\rho_m = \min_{v,i} \rho(z, e_v(i))$ на еталонній множині бази E і визначаємо його номер k класу. Важливою умовою для врахування класу і значущості є виконання нерівності $\rho_m \leq \delta_z$, де δ_z – граничний параметр, що апріорно визначає значимість мінімуму метрики для встановлення еквівалентності між дескрипторами [1].

Після здійснення (1.3) для окремого $z \in Z$ інкрементуємо акумулятор $h = (h_i)_{i=1}^N$ у вигляді $h_k = h_k + 1$ для отриманого номеру класу і агрегуємо вектор $val = (val_i)_{i=1}^N$ значущості як $val_k = val_k + \lambda_{k,v}$ для цього класу.

Величина $\lambda_{k,v}$ – це значущість представника еталону з номером k , для якого досягнуто мінімум в (1.3). У результаті оброблення усього опису Z об'єкту отримуємо накопичені вектори h , val що можуть стати основою класифікаційного рішення.

За результатом першого етапу отримано підсумкові кількості h_i голосів та ваг val_i на підставі тих підмножин компонентів об'єкту, що віднесені до кожного із класів E_i , $i = 1, \dots, N$.

Вектор $\{val_i\}_{i=1}^N$ формується як сума чисел $\lambda_{i,v}$ для тих компонентів $e_v(i)$, стосовно яких встановлена відповідність елементам $z_v \in Z$ за процедурою (1.3). У такій схемі основоположним для процесу класифікації залишається метричне відношення, а значущість має другорядну роль. Це можна пояснити тим, що дескриптор є основною ознакою зображення, а значущість $\lambda_{i,v}$ дескриптора є функцією як від змісту бази еталонів, так і від способу її обчислення.

Фінальне класифікаційне рішення F про клас k для об'єкту Z має вид:

$$k = \arg \underset{i=1, \dots, N}{opt} F[\{h_i\}_{i=1}^N, \{val_i\}_{i=1}^N]. \quad (1.4)$$

Модель (1.4) фактично є розвиненням і узагальненням для схеми голосування із використанням вагових коефіцієнтів матриці Λ і потребує детального вивчення і конкретизації для різноманіття варіантів F .

Означимо можливі практичні умови відмови від визначення класу як

$$\max_i \{h_i\} < \delta_h, \quad \max_i \{val_i\} < \delta_{val}. \quad (1.5)$$

Умови (1.5) відповідають ситуації, де накопичені максимуми серед числа голосів та/або значущостей досить низькі, що не дає можливості впевнено визначити клас. Зрозуміло, що порогові значення δ_h , δ_{val} оцінюються експериментально і напряду залежать як від змісту бази еталонів, так і від схеми реалізації (1.3). Наряду із умовами (1.5), що задають абсолютні обмеження, у прикладних системах також використовують відносні обмеження на значимість максимумів (1.5) у порівнянні з найближчим конкурентом для іншого класу [1]. Такі обмеження додатково сприяють підвищенню впевненості класифікаційного рішення і суттєво впливають на результат.

Матриця значущості Λ обчислюється на підготовчому етапі і несуттєво збільшує обчислювальні затрати у процесі класифікації. Зважаючи на те, що дескриптори зображення формуються у довільному порядку, кожний вектор – рядок матриці Λ у процесі аналізу треба розглядати як множину значень.

Як показали результати нашого експериментального моделювання з використанням конкретних моделей (1.4) (розділ експериментів), серед можливих варіантів (1.4) для побудови оптимального двопараметричного рішення більш практичними і ефективними є простіші із них, що оптимізують параметр числа голосів $\{h_i\}_{i=1}^N$.

Використання значущості у формі інформативності [2] значно корелює з числом голосів і може бути використано для підтвердження рішення.

1.3 Способи підрахунку голосів класів

Відповідно до традиційної схеми (1.3) щодо визначення найближчого сусіда оптимізація відстані здійснюється за двома параметрами – номером класу та номером дескриптора всередині класу. У прямому сенсі схема (1.3) реалізує прийняття агрегованого рішення ансамблем простих класифікаторів для сукупної множини дескрипторів об'єкту [4].

Число результативних голосів, за яким приймається рішення, у цій схемі дорівнює потужності аналізованого опису Z .

Зрозуміло, що у класифікації основним результатом є номер класу. Тому на практиці, особливо коли еталони задано, часто застосовують більш продуктивні підходи, що зводяться до поетапного пошуку, де першим етапом є пошук всередині опису для фіксованого класу [23]. В основі таких підходів є основоположний факт, що будь-яка процедура класифікації за описом у формі множини векторів у самому загальному аспекті оцінює ступінь релевантності об'єкт – еталон, що оптимізується на множині еталонів. В свою чергу, така міра релевантності безпосередньо відображає потужність перетину двох скінченних множин векторів (описів об'єкту та еталону). Крім того, поетапні підходи базуються на важливій апіорній інформації про належність еталонного дескриптора до фіксованого класу, що робить класифікацію більш стійкою на основі використання узгодженості з еталонними даними.

Зауважимо, що впровадження кожної із таких модифікацій пов'язано із деякими особливостями (спосіб визначення відповідності дескрипторів, вибір порогу для числа голосів, значущості тощо). Крім того, при незначному числі дескрипторів у описі чи при нерівноцінних потужностях порівнюваних описів виникають додатково труднощі з розрізненням «свій-чужий» [27, 33].

Як один із варіантів розглянемо спосіб класифікації, що полягає у накопиченні оптимальних значимостей для кожного із класів без попереднього встановлення номера класу для окремого дескриптора, як це зроблено у традиційному методі (1.3). Це один із практичних варіантів реалізації (1.3). Його можна вважати модифікацією методу ближнього сусіда.

Знаходимо мінімуми відстаней

$$\rho_{m,v}(i) = \min_z \rho(e_v(i), z), \quad i = \overline{1, N}, \quad v = \overline{1, S} \quad (1.6)$$

окремо для кожного дескриптора $e_v(i)$ класу E_i серед елементів Z . При виконання умови $\rho_{m,v} \leq \delta_z$ накопичуємо вектори голосів $\{h_i\}_{i=1}^N$ та значущостей $\{val_i\}_{i=1}^N$ за адитивною моделлю $val_i = val_i + \lambda_{i,v}$, де $\lambda_{i,v}$ – значущість дескриптора із класу i , для якого досягається оптимум $\rho_{m,v}(i)$. Рішення про клас приймаємо відповідно до (1.4) на підставі накопичених голосів і значущостей.

Цей спосіб більше орієнтований на узгодженість із еталонними даними, коли кожний із еталонів «підшукує своїх» серед компонентів опису об'єкта. Цей спосіб дієвий при різній потужності описів еталону та об'єкту. Максимальне число отриманих голосів за клас у ньому співпадає з кількістю дескрипторів у еталоні.

Ще один практичний підхід до встановлення міри релевантності двох описів здійснює пошук мінімуму з подвійною перевіркою (Cross-Checking [38]): для дескриптора об'єкту, якому відповідає знайдений мінімум, зустрічно визначається мінімум відстані серед дескрипторів еталону; відповідність дескрипторів об'єкту та еталону встановлюється лише у випадку співпадання результату для обох пошуків. Така подвійна перевірка може бути здійснена для процедур зіставлення як у межах усієї бази, так і окремо для кожного із еталонів. Впровадження моделей виду Cross-Checking націлено на підвищення надійності прийняття класифікаційного рішення шляхом зменшення числа можливих викидів.

Загалом схема аналізу даних відповідно до моделей (1.3), (1.4) та (1.6) може бути підсилена шляхом визначення не одного, а кількох найближчих мінімумів, наприклад, трьох [4, 13]. Це відповідає моделі «трьох найближчих сусідів». У такому способі на другому етапі класифікації інкрементацію кількості голосів і агрегацію значущостей отримують зразу три класи (можуть бути одні й ті ж), а схема прийняття рішення (1.4) не зміниться.

У розглянутих моделях клас аналізованого дескриптора з урахуванням параметра значущості фактично визначається на підставі двох критеріїв: величини ρ_m мінімальної відстані до визначеного дескриптора еталону та апріорного набору значимостей для еталонних дескрипторів, отриманих на етапі навчання.

Зауважимо, що оцінка відстані в (1.3), (1.6) може бути здійснена також з використанням сучасних швидкісних методів на підставі ієрархії даних замість чималого за обсягом обчислень лінійного пошуку [1, 27, 38].

1.4 Здійснення редукції опису

Одним із найбільш дієвих способів підвищення продуктивності класифікації є стиснення простору ознак [2, 39]. Для класифікаторів на підставі опису як множини дескрипторів ключових точок цей спосіб може бути реалізовано шляхом скорочення потужностей еталонних описів. Швидкодія класифікації при цьому зростає пропорційно ступеню скорочення. Як критерій для відбору значущих компонентів еталонних описів може бути використано значення матриці Λ . На підставі апріорного аналізу матриці Λ обсяги еталонних описів E_i можуть бути суттєво зменшені за рахунок того, що їх окремі компоненти із значущістю виключаються із процесу класифікації.

Відмітимо, що після скорочення опису самі величини значущостей суттєво змінюються, так як фактично змінюється склад бази. Тому значення Λ треба обов'язково перераховувати, якщо у процесі класифікації із впровадженням значущостей буде використано модифікований опис [2]. За нашими експериментальними дослідженнями можна рекомендувати схему, коли скорочення опису здійснюється поетапно, і на кожному із етапів перераховуються коефіцієнти для зменшеного складу рядків матриці Λ .

Враховуючи особливості застосування методу голосування, де чисельність порівнюваних описів вважається приблизно однаковою, природно вважати рівноцінними і розмірності трансформованих після стиснення описів [40]. Але більш детальне вивчення допускає суттєві відмінності між потужностями описів об'єкту та еталонів, що потребує модифікації процедури підрахунку і аналізу голосів. Наша модифікація методу найближчого сусіда з використанням моделі Cross-Checking дає можливість це реалізувати. За результатом редукції еталонних описів виникає необхідність модифікувати і параметри класифікатора: вирішальне число голосів, відношення глобального і локального максимумів і т. д.

Наш аналіз показав, що прямий відбір фіксованого скороченого числа дескрипторів ключових точок шляхом керування параметрами детектору ключових точок не приводить до покращення результативності, так як точність класифікації при цьому падає пропорційно скороченню опису.

Більш результативним є початкове формування великого числа дескрипторів опису (500 і більше), що відображають усі особливості зображення, з наступним скороченням за критерієм значущості.

Зауважимо, що при істотній відмінності числа дескрипторів для еталонів і вхідного зображення дещо зростає ймовірність виродженої ситуації, коли кілька еталонів можуть бути знайдені у складі вхідного зображення. Таким чином, кожне скорочення має свої межі застосування.

Схема класифікатора з використанням редукції опису наведена на рисунку 1.1. Зеленим кольором на рисунку 1.1 відображено введені нами додаткові блоки оброблення даних з використанням редукції на підставі значущості за параметром інформативності [2, 15].

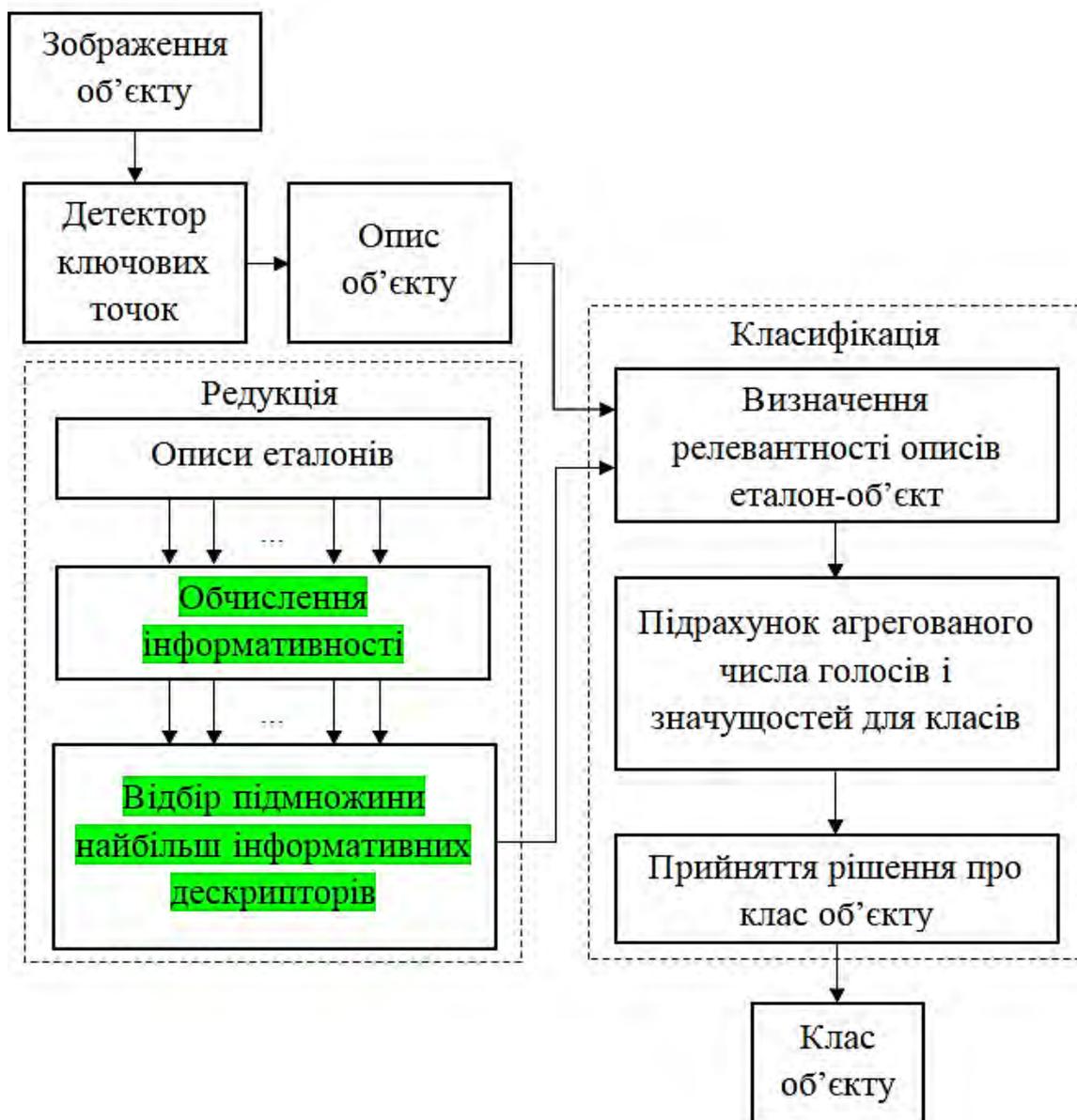


Рисунок 1.1 – Схема класифікації з використанням редукції опису

1.5 Визначення параметрів значущості

Зважаючи на те, що умови розпізнавання у системах комп'ютерного зору часто формує людина, значення матриці Λ апіорно може бути сформовано експертом, а параметр значущості оцінюється людським оком шляхом вибору найбільш важливих ключових точок для класифікації. Але задля забезпечення автоматичного функціонування систем необхідно мати моделі для оцінювання значимості даних. Зауважимо, що всі обговорювані тут параметри значущості обчислюються на навчальній множині ознак.

Застосуємо для обчислення матриці Λ метричний критерій інформативності [2]. Для довільного вектора $z \in E$ у системі класів як складового елемента $z \in E_k$ фіксованого еталонного опису E_k з номером k введемо поняття інформативності $I(z, E)$ у складі бази E

$$I(z, E) = \rho_m(z, \bar{E}_k) - \rho_m(z, E_k), \quad (1.7)$$

де $\rho_m(z, \bar{E}_k) = \min_{v, i \neq k} \rho(z, e_v(i))$ – мінімальна відстань від z до елемента бази, що не належить класу E_k ;

$\rho_m(z, E_k) = \min_{v, i=k} \rho(z, e_v(i))$ – відстань від z до найближчого елемента із класу E_k (за виключенням відстані $\rho(z, z) = 0$ самого до себе, $z \in E_k$).

При впровадженні нормованих відстаней із величиною $0 \leq \rho \leq 1$ оцінювані значення критерію (1.7) лежать в інтервалі $-1 \leq I \leq 1$. У методі Relief використовується один із варіантів моделі (1.7), де інформативність оцінюється як різниця квадратів відстаней, а відбір аналізованих ознак здійснюється випадковим чином [41]. Як бачимо, критерій (1.7) має властивість оцінки даних.

Використання моделі (1.7) для визначення індивідуальної інформативності I для $z \in E_k$ засновано на припущенні про те, що класифікаційна значущість ознаки є тим вищою, чим краще вона розділяє екземпляри навчальної вибірки на класи. Виходячи з цього, чим далі від екземпляра класу знаходиться найближчий до нього екземпляр іншого класу, тим вищою є індивідуальна інформативність. У той же час, чим далі від екземпляра знаходиться найближчий до нього елемент цього ж класу, тим нижчою є індивідуальна інформативність. Тобто впроваджується принцип: «ближче до своїх, подалі від інших».

Таким чином, ознаки (1.7) з високими значеннями I індивідуальної інформативності будемо вважати значущими, тобто суттєво інформативними стосовно результативної класифікації, ознаки з низьким I вважаються малозначущими, тобто є претендентами на виключення із системи ознак.

У статті [11] нами розглянуто інші критерії для оцінювання значущості, зокрема, засновані на принципі унікальності ознаки серед елементів E всередині і поза межами класу. Інші принципи формування критеріїв значущості, засновані на стійкості дескрипторів в умовах геометричних перетворень та з використанням розподілів за визначеними класами, включаючи кластерне подання, розглянуто у роботах [2, 4, 19, 20, 23, 35].

Критерій (1.7) можна вважати більш дієвим за інші, так як він уже частково відображає ступінь метричного розрізнення аналізованих компонентів даних, що є підставою для класифікації зображень за сформованими описами. Крім того, критерій (1.7) характеризується числом, що спрощує оброблення.

Наші пропозиції [1, 5, 15] щодо удосконалення структурних методів класифікації з метою скорочення обчислювальних витрат полягають у такому:

а) у відповідності до схеми класифікації рисунку 1.1 для усієї множини із s дескрипторів у складі кожного еталону обчислюємо (1.7);

б) на підставі найбільших значень показника (1.7) відбираємо у складі кожного еталону фіксоване число із $s^* \ll s$ дескрипторів;

в) використовуємо отриману підмножину із s^* дескрипторів у схемі класифікації (модель (1.6), рис. 1.1).

1.6 Критерії ефективності класифікації

Результативність класифікації будемо оцінювати значенням показника точності pr , який обчислюється відношенням числа правильно класифікованих об'єктів r_p до загального їх числа r , що використовувалося в експерименті [13]

$$pr = r_p / r. \quad (1.8)$$

Ще одним важливим показником для методів голосування є відношення максимуму числа голосів $h_{\max 1}$ чи агрегованої значущості, за якими приймається класифікаційне рішення, до найближчого локального максимуму $h_{\max 2}$ класу-конкурента

$$\Delta = h_{\max 2} / h_{\max 1}. \quad (1.9)$$

Значення Δ характеризує ступінь впевненість рішення. Чим менше Δ , тим надійніше прийняте рішення про клас об'єкту.

Основні порогові значення, що використовуються у дослідженні, стосуються порогу δ_z для значення метрики щодо визнання еквівалентними двох дескрипторів та порогу δ_h для числа голосів дескрипторів, необхідних для прийняття класифікаційного рішення за максимумом набраних голосів. Обидва пороги у роботі обираються експериментально і залежать від обраної бази еталонів. Щодо порогу еквівалентності δ_z є літературні джерела [13], де вважається прийнятним поріг у 25% від максимуму метрики для багатовимірних векторів. Хоча за нашим дослідженням [14, 15] цей поріг можна визначити більш адаптовано і ефективно. У цьому дослідженні поріг δ_h для числа голосів за статистичним експериментом обирається оптимально так, щоб визначений поріг перевищував мінімум серед максимумів голосів перетворених еталонних зображень, але в той же час був вищим від максимуму серед голосів зображень поза базою. Аналітично оптимізувати ці пороги не вдається із-за необмеженого обсягу реальних зображень і їх дескрипторів.

1.7 Результати експериментів та обговорення

Зауважимо, що величина параметру інформативності (1.7) як різниця між двома відстанями напряду не пов'язана з ймовірністю правильної класифікації на відміну від значення подібності чи метрики між множинами даних [13, 28, 33]. Тому результативність його використання для розпізнаваних даних необхідно перевіряти експериментально, що дасть можливість оцінити прикладну точність класифікації з використанням інформативності.

Перший експеримент.

З використанням програмних засобів бібліотеки OpenCV детектором BRISK сформовано опис із множини дескрипторів ключових точок (дескриптор BRISK містить 512 бітів) для аналізованих зображень [6, 42].

У експерименті для трьох різних зображень одного і того ж казкового персонажу (рис. 1.2) з числом $s = 500$ дескрипторів brisk у описах шляхом відбору 100 дескрипторів із найвищим показником інформативності (1.7) прогнозовано досягнуто приблизно п'ятикратний виграш у часі класифікації.

Точність класифікації для повного і редукованого описів при цьому складає $pr = 1,0$, тобто усі зображення класифікуються вірно.

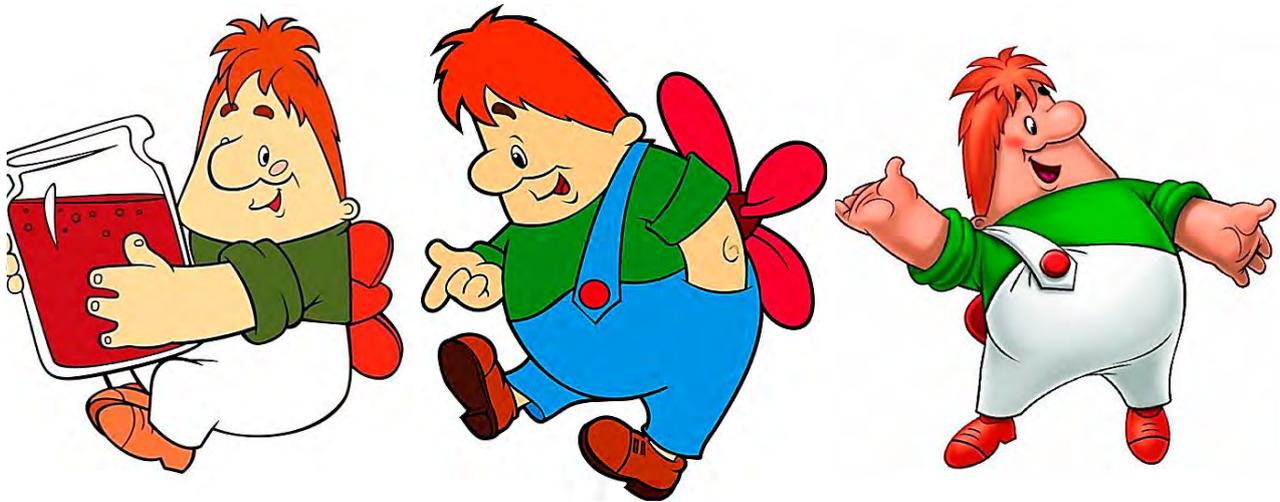


Рисунок 1.2 – Зображення в експерименті редукції за критерієм (1.7)

При цьому показник впевненості класифікації (1.9) для редукованого опису у $s = 100$ дескрипторів склав значення 0,29, у той час як для повного опису у 500 дескрипторів цей показник для аналізованих зображень рисунку 1.2 дорівнює 0,59, що значно гірше.

Отриманий експериментальний результат говорить про те, що навіть для достатньо схожих зображень за рахунок редукції за коефіцієнтом інформативності досягнуто суттєве підвищення надійності в аспекті розрізнення зображень. Редукція опису у п'ять разів за критерієм інформативності дає не тільки істотний вигреш у часі оброблення, але й забезпечує зростання показника впевненості при збереженні точності класифікаційного рішення. Відібрана підмножина найбільш інформативних дескрипторів у повній мірі забезпечує результативність класифікації з істотно нижчими обчислювальними затратами.

Другий експеримент.

З метою більш глибокого вивчення результативності і властивостей редукованих описів для класифікації нами проведено більш масштабний експеримент з використанням об'ємного тестового матеріалу.

Здійснено оцінювання точності класифікації для різних варіантів класифікатора і представлення редукованого опису зі скороченням числа дескрипторів s у описі в широкому діапазоні від 500 до 10.

Для програмного моделювання застосовано мову програмування Python, бібліотеку комп'ютерного зору OpenCV та бібліотеку NumPy для прискорення обробки багатовимірних даних, створено 500 дескрипторів ORB розміром 256 біт для опису кожного зображення [7, 42–44]. Як базу класифікації використано зображення прикрас з Національного музею історії України [42].

Для тестування сформовано і використано базу класифікації із 5 зображень музейних експонатів (діадема (а), пектораль (б), підвіска (в), топелік (г), зулюф-аск (д), рис. 1.3), а також інші 3 зображення колекції прикрас (пряжка (а), підвіски (б), (в), рис. 1.4), що не включені до бази класифікації.



Рисунок 1.3 – Зображення еталонів із бази класифікації

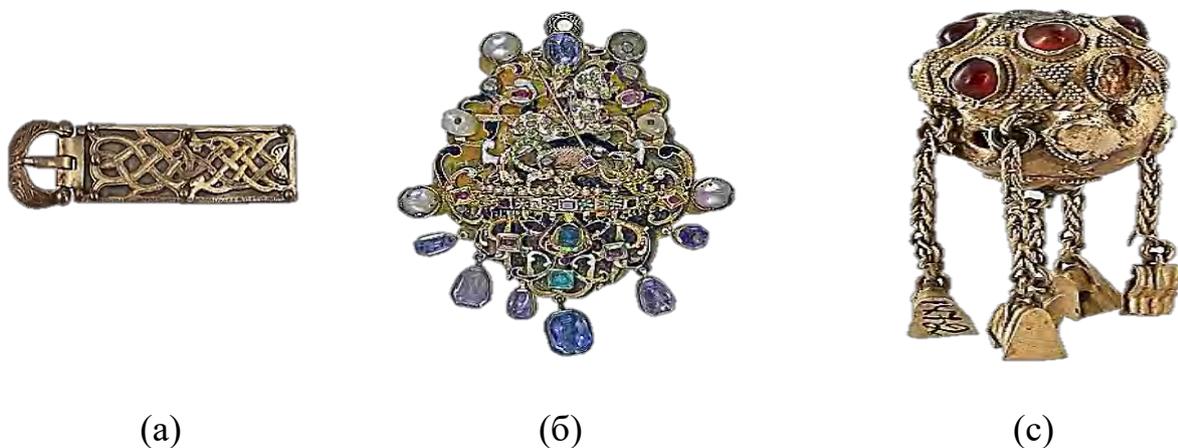


Рисунок 1.4 – Тестові зображення прикрас,
що не включено до бази класифікації

Для проведення тестування сформовано набір вхідних зображень із рисунків 1.3 та 1.4, де кожне зображення додатково трансформовано застосуванням геометричних перетворень для 6 різних комбінацій повороту і масштабу. Таким чином, тестовий набір із 51 зображень включає 30 трансформованих зображень бази і 21 зображення прикрас не із бази.

На рисунку 1.5 показано приклади трансформованого зображення діадеми (поворот 45° , масштаб 80% (а); поворот 30° , масштаб 120% (б)).

Для визначеного тестового набору при числі дескрипторів $s = 500$ усі 51 зображення традиційним методом класифікувались вірно (трансформовані еталони віднесено до «свого» класу, решта – до жодного із класів).

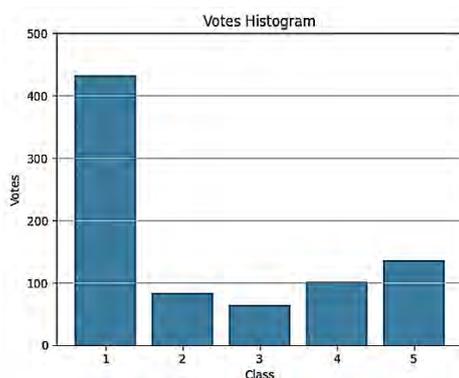


(а)

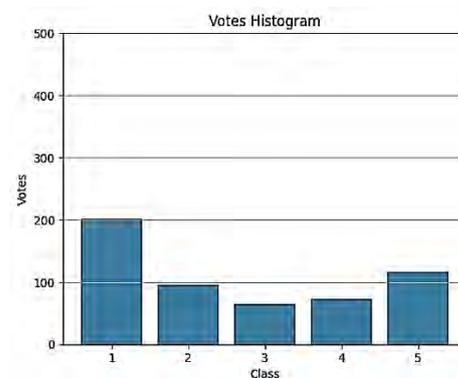
(б)

Рисунок 1.5 – Трансформовані зображення діадеми із координатами ключових точок

При цьому для трансформованих еталонних зображень показник Δ впевненості класифікації не перевищував значення $\Delta \leq 0,1$, а для зображень не із бази спостерігалось досить рівномірне розподілення голосів для класів при значенні $\Delta > 0,45$, що унеможлиблювало їх класифікацію. Типовий приклад числа голосів для зображення не із бази: [212, 107, 80, 105, 139], коефіцієнт $\Delta = 0,66$ (рис. 1.6). Приклад гістограми голосів для наявної множини із 5 класів для трансформованих еталонів (а) та об'єктів не із бази (б) (рис. 1.6).



(а)



(б)

Рисунок 1.6 – Гістограми голосів для перетвореного еталону (а) та об'єкту не із бази (б)

Як бачимо, гістограма для трансформованих еталонів має одномодальну форму, що сприяє здійсненню класифікації, а для об'єкту не із бази голоси розподілені досить рівномірним чином.

Незважаючи на менше порогу значення Δ для деяких зображень не із бази, для жодного з них не отримано число голосів, більше половини (250). Отримані показники для числа голосів загалом вказують на істотну схожість усіх використаних зображень (не із бази і всередині бази) у побудованому просторі ознак.

Далі на підставі обчислення показника (1.7) визначено 50 найбільш інформативних дескрипторів для кожного із еталонів (рис. 1.7 (а)).

Таким чином сформовано редукований опис еталонів із 50 відібраних дескрипторів (рис. 1.7 (б)).



Рисунок 1.7 – Координати 500 (червоний) та редукованих 50 (синій) ключових точок

Зважаючи на нерівноцінне число s дескрипторів ключових точок у редукованому еталоні ($s = 50$) і вхідному зображенні ($s = 500$), застосовано спосіб класифікації, де кожний еталон відбирає на об'єкті «свої» рівноцінні дескриптори, максимально можливе число яких тепер дорівнює обсягу еталонного опису. Отриманий мінімум відстані Хемінга перевіряється за умовою $\rho_m \leq \delta_z$, $\delta_z = 64$ (число 64 як 25% від максимуму значення 256 для метрики Хемінга [13]), і зараховується як голос для класу при її виконанні.

Відмітимо, що використання порогу δ_z для перевірки на значимість обчисленого мінімуму залишається спірним моментом і потребує вивчення і оцінювання для кожної із конкретних баз класифікації, незважаючи на беззаперечне твердження поважних дослідників про суттєву схожість елементів багатовимірного векторного простору [13, 19].

Значення δ_z статистично поділяє дескриптори на «своїх та чужих», залишаючи деякий допуск для відхилень «своїх». Як правило, цей поріг визначається на підставі експериментів класифікації для «своїх», так як «чужі» належать решті фактично нескінченного простору B^n багатовимірних бінарних векторів. У деяких застосуваннях використовують різні значення порогу δ_z для навчання системи чи визначення параметрів і окремо – для класифікації [1, 28].

За результатами наших експериментів цього і інших досліджень [4, 35] можна заявити, що в деяких ситуаціях класифікація без використання порогу δ_z є навіть більш результативною. Тому у прикладних застосуваннях треба досліджувати усі можливі варіанти і обирати найкращий спосіб, вибір якого може залежати від складу розпізнаваних даних.

На основі проведених експериментів статистично обрано поріг для результативного числа набраних голосів як половину від кількості дескрипторів (250 і 25 відповідно). Також експериментально вибрано поріг $\delta_\Delta = 0,62$ для коефіцієнта впевненості Δ , що забезпечує вірне визначення класу для усіх зображень із бази. Класифікаційне рішення формалізовано ситуацією: число голосів для класу – переможця перевищує половину, а значення коефіцієнту впевненості менше 0,62.

Зважаючи на те, що будь-яка прогресуюча система прагне до спрощення, поріг δ_Δ , на наш погляд, не обов'язково включати у процес роботи класифікатора. Особливо це стосується експериментів із повним описом у 500 дескрипторів. Показник δ_Δ в більшій мірі може бути використаний для контролю ситуації, так як характеризує ступінь надійності рішень. Але для невеликих об'ємів опису (50, 25, 10), як показав наш аналіз, використання показника δ_Δ стає нагальним, так як дає можливість відділити «чужі» дескриптори. Для тестового редукованого набору при $s = 50$ усі вхідні зображення класифікувались вірно, але показник Δ впевненості класифікації для трансформованих еталонних зображень передбачувано зріс до рівня $\Delta \leq 0,6$, а для зображень не із бази отримано розподілення голосів для класів при ще більш високих значеннях $\Delta > 0,7$.

Бачимо, що показник Δ у ситуації з редукцією даних дає можливість не тільки оцінити, а й забезпечити надійність класифікації.

Точність pr класифікації оцінювалась як відношення (1.8) кількості правильно класифікованих об'єктів до загального числа експериментів. Експериментальний показник точності склав 1,0 (максимум) як для описів обсягом 500 дескрипторів, так і для редукованої бази у 50 найбільш інформативних еталонних дескрипторів, що ще раз підтверджує високу ефективність методу із впровадженням редукованих описів. При цьому час класифікації одного зображення зменшився приблизно у 10 разів з 0,27 с для 500 дескрипторів до 0,027 с для 50 дескрипторів опису (табл. 1.1).

Таблиця 1.1 – Експериментальні показники ефективності методів

Параметри	Поріг δ_h для числа голосів				Час t_c класифікації, с				Точність pr			
	500	50	25	10	500	50	25	10	500	50	25	10
Число дескрипторів	500	50	25	10	500	50	25	10	500	50	25	10
Модифікація ближнього сусіда	250	25	15	8	0,28	0,027	0,014	0,0061	1	1	0,95	0,93
Cross-Checking	220	25	15	6	0,28	0,026	0,013	0,0056	1	1	1	0,96

Для порівняльної оцінки у нашому окремому експерименті здійснено безпосереднє початкове визначення 50 і 25 (замість 500) дескрипторів ORB для об'єкту та еталонів.

Таке пряме скорочення аналізованих даних і використання їх для класифікації призвело до суттєвого погіршення показника точності pr для 50 дескрипторів – до рівня 0,6...0,7, а для опису із 25 дескрипторів – навіть до рівня 0,4...0,5. На зниження точності pr тут більшим чином вплинули тестові зображення, що не включені до бази класифікації.

Таким чином, тільки спосіб редукції опису на підставі визначення інформативності дескрипторів еталону зберіг високий рівень точності класифікації. При цьому класифікаційна результативність стисненого опису прямо залежить від процедури його формування. Одночасне забезпечення високих показників точності і продуктивності класифікації досягається процедурою поетапного скорочення описів для еталонів бази на підставі оцінювання критерію інформативності (1.7).

Зауважимо, що у процесі відбору дескрипторів за інформативністю її фактичне значення змінюється і для подальшого використання потребує перерахунку, так як на величину інформативності (1.7) напряму впливає склад отриманої скороченої бази.

Ще одне застереження стосується безпосереднього використання значень інформативності у моделях виду (1.2). Впровадження таких моделей може бути ефективне лише в ситуації, якщо обчислені значення інформативності суттєво відрізняються у складі одного еталону чи для різних еталонів. У проведеному нами експерименті значення інформативності для відібраних 50 дескрипторів коливались у межах інтервалу 34...50 для першого еталону і 24...46 для решти чотирьох з приблизними середніми значеннями 38, 28, 31, 31, 29. Як бачимо, ці значення досить близькі між собою, а значить, пряме впровадження моделі (1.2) з цими значеннями не має сенсу, так як воно ніяк не вплине на підсилення класифікаційних показників.

Тому далі на підставі набору із 50 виділених дескрипторів для еталонів нами було перераховано для них значення інформативності. Вони сформували діапазони -19...+59, -28...+49, -9...+49, -27...+57, -21...+60 з середніми значеннями 36, 31, 33, 30, 38. Таким чином, загальний діапазон значень інформативності для модифікованої бази склав -21...+60.

На основі отриманих показників інформативності шляхом відбору найбільших значень сформовано нові описи еталонів обсягом у 25 дескрипторів кожний. Моделювання показало, що для бази описів із 25 дескрипторів визначено тільки 2 хибно позитивних об'єкти (віднесено до певного класу) із числа зображень, які не входять до еталонних. При цьому усі зображення із бази класифіковані вірно! Маємо точність на рівні 0,95.

Аналогічно із 50 відібраних на першому етапі дескрипторів визначено 10 найбільш інформативних дескрипторів. На тестовому наборі експериментально визначено 10 хибно класифікованих об'єктів із 51, отримано точність на рівні 0,81. На зниження точності більше вплинули зображення не із бази, так як на невеликій вибірці ознак схожість усіх зображень суттєво збільшується. Якщо ж тут виключити із правила класифікації умову для не перевищення порогу δ_{Δ} , то точність дещо зростає до $pr = 0,93$, що є досить високим показником.

Зауважимо, що відповідно до принципів data science [19, 20] отримані досить високі показники напряму залежать від тестового набору зображень у складі і за межами обраної бази.

У будь-якому випадку результати проведених експериментів дають можливість суттєво скоротити час класифікації (для 25 інформативних дескрипторів – приблизно у 20 разів!) без суттєвого зниження показника точності.

Найбільш практичний спосіб класифікації полягає у прийнятті рішення за половиною і більше голосів сформованого опису еталону.

Експерименти із накопиченням параметра інформативності дескрипторів як варіант моделі (1.4) у процесі прийняття рішення показали наступне.

Задля спрощення обчислень перераховані для 50 дескрипторів коефіцієнти інформативності ми розбили на приблизно однакові за шириною 4 інтервали, і присвоїли їм вагові інтервальні коефіцієнти 1, 2, 3, 4, щоб разом із числом голосів накопичувати ці вагові значення, а також добуток відповідної ваги на отриману за пошуком відповідності мінімальну відстань.

Аналіз показав, що переважне число інформативних дескрипторів еталонів (39...43 із 50-ти для різних еталонів) отримало інтервальні коефіцієнти 1 та 2, тобто мають суттєвий рівень інформативності. Це можна пояснити тим, що аналізовані дані попередньо уже пройшли процес відбору за критерієм інформативності.

За результатом класифікації модифікацією ближнього сусіда маємо приклад для трансформованого зображення 2-го класу щодо накопичених голосів [13, 43, 3, 4, 7], значущостей [17, 70, 6, 6, 14] та добуток значущості на відстань [938, 2270, 332, 347, 866] з параметром $\Delta = 0,31$ для голосів.

Як бачимо із цього прикладу, отримані класифікатором значення для голосів та значущостей в значній мірі корелюють між собою. Це можна пояснити тим, що застосований критерій інформативності (1.7) визначається через мінімум метрики, а значить, метричні співвідношення домінують у такому класифікаторі. Експериментально у деяких випадках для зображень із бази накопичена значущість правильно вказує на істинних клас об'єкту, але для зображень не із бази вона лише підтверджує факт неправильного віднесення опису до одного із класів за числом голосів.

Задля узагальнення результатів дослідження проведено програмне моделювання класифікатора для тих же тестових даних із використанням засобу подвійної перевірки відповідності дескрипторів за моделлю Cross-Checking. Для 500 дескрипторів точність 1,0 не змінилась, але для її забезпечення треба знизити поріг δ_n для числа голосів до значення 220. Це необхідно для правильної класифікації еталону 2, найбільш схожого з іншими.

Коефіцієнт впевненості Δ для зображень із бази покращився до рівня 0,08, а час t_c обчислень майже не зріс у порівнянні із підходом без подвійної перевірки.

Зауважимо, що експериментальний час класифікації із використанням Cross-Checking коливається у достатньо широких межах для різних тестових зображень у порівнянні з першим підходом, де він фактично постійний.

Для складу із 50 дескрипторів досягнута максимальна точність 1,0 при порозі δ_h у 25 голосів, $\Delta \leq 0,1$ для бази, час класифікації приблизно у 10 разів менше, ніж для повного складу опису. Для 25 дескрипторів досягнута точність 1,0 при порозі δ_h у 15 голосів, $\Delta \leq 0,5$ для бази, час класифікації у 2 рази менше, ніж для складу опису у 50 дескрипторів. Вивчення варіанту для 25 дескрипторів із використанням накопичення значущості val_k фактично підтверджує класифікацію з використанням тільки δ_h .

Важливий результат отримано для невеликого числа дескрипторів у 10 елементів, де більшим чином може впливати коефіцієнт значущості. При виборі порогу $\delta_h = 8$, як у першому методі, точність падає до рівня $pr = 0,84$. Але вже при виборі порогу $\delta_h = 6$ точність значно покращується до $pr = 0,96$ (але при суттєвому збільшенні $\Delta \leq 0,85$). Варіантом для покращення результату при невеликих описах є прийняття оптимального рішення виключно із перевіркою за параметром δ_h .

При цьому класифікація за накопиченою значущістю val_k повністю (навіть при невірному рішенні) підтверджує висновок за підрахунком голосів. При цьому більш важливим є вибір порогу δ_h , ніж використання показника Δ .

При класифікації для 25 дескрипторів за моделлю (1.2), де критерієм є добуток значущості на відстань, отримано такі результати. Для досягнення показника $pr = 1,0$ поріг δ_h треба встановити на рівні 10, при цьому час обчислення незначно зріс. Для опису у 10 дескрипторів точність при виборі $\delta_h = 4$ складає $pr = 0,84$, що дещо гірше, ніж для традиційного класифікатора за числом голосів. Отримані експериментальні дані внесено до таблиці 1.1.

Аналіз проведених експериментів і зміст таблиці 1.1 призводить до таких попередніх висновків;

– накопичена значущість корельована з числом голосів класів і може бути використана для підтвердження рішення. Результативність класифікації за накопиченою значущістю практично співпадає з рішенням за числом голосів;

– впровадження моделі (1.2), де критерієм є добуток метрики на значущість, при $s = 25$ забезпечує точну класифікацію при умові зниження порогу δ_h голосів класів до $\delta_h = 10$. Але при невеликому значенні $s = 10$ модель (1.2) не покращує результативність класифікатора для аналізованих даних;

– Cross-Checking забезпечує більш точну класифікацію з кращим показником Δ за рахунок зменшення впливу викидів;

– метод підрахунку відповідностей із Cross-Checking більш чутливий до вибору порогу δ_h . Швидкодія при застосуванні Cross-Checking коливається у широких межах в залежності від значень даних;

– показник впевненості Δ погіршується (зростає) зі скороченням числа дескрипторів у описі. Значення Δ при $s = 500$ у десять разів краще для методу з Cross-Checking;

– точність класифікації для досліджених методів, особливо при опису у 10 дескрипторів, може бути покращена адаптивним вибором δ_h і спрощенням класифікації шляхом виключення перевірки на значення δ_Δ . Використання коефіцієнтів значущості меншим чином впливає на підвищення точності класифікації для зображень не із баз;

– для невеликих редукованих обсягів опису доцільно застосувати просте правило класифікації на основі тільки числа набраних голосів класів.

Отже, впровадження значущості у формі критерію інформативності даних у процес структурної класифікації сприяє посиленню адаптованості з еталонами класів і забезпечує обґрунтоване прийняття рішень. Застосування параметру інформативності розкриває нові можливості керування процесом аналізу даних в ході класифікації. Ключовим для процесу класифікації залишаються метричні відношення для дескрипторів та еталонів.

Параметр значущості може бути успішно використано лише у тих ситуаціях, де його значення має варіабельність на множині дескрипторів чи еталонів.

Класифікація із впровадженням редукованих описів еталонів у порівнянні із повним описом зберігає ефективність лише у випадку скорочення опису шляхом відбору за критерієм інформативності. Безпосереднє формування опису меншого обсягу суттєво погіршує показник точності класифікації.

Використання у експерименті різних дескрипторів ORB, BRISK підтверджує універсальність пропонованого механізму редукції даних опису незалежно від виду детектора.

Основний результат дослідження полягає у встановленні високої продуктивності і результативності класифікаторів за редукованим складом опису еталонів.

Застосування редукції для трансформації множини дескрипторів опису зображень дає можливість значно прискорити оброблення без суттєвої втрати точності класифікації. Швидкодія оброблення збільшується пропорційно ступеню редукції даних і для проведеного експерименту покращена у 20 разів.

Практично доцільним є використання 10% найбільш інформативних дескрипторів, що забезпечує приріст швидкодії в 10 разів при повному збереженні точності. Якщо ж швидкодія є першочерговим критерієм, то допустимо використовувати навіть 5% від оригінальної кількості дескрипторів, що забезпечує приріст швидкодії майже у 20 разів, але при незначному зниженні точності до 0,95.

Матеріали дослідження опубліковано у статті [42].

2 ЗАСТОСУВАННЯ НЕЙРОННИХ МЕРЕЖ КОХОНЕНА ДЛЯ НАВЧАННЯ КЛАСИФІКАТОРА ЗОБРАЖЕНЬ

2.1 Мережа Кохонена у навчанні класифікатора

Проблема здійснення результативної класифікації зображень візуальних об'єктів в умовах впливу геометричних перетворень є однією із найскладніших для сучасних систем комп'ютерного зору [11, 22, 45, 46]. За останні роки її практичному вирішенню у значній мірі посприяло впровадження апарату нейронних мереж [30, 47, 48]. Як правило, набір зображень окремого класу подається на вхід мережі після попереднього анотування людиною по виділенню меж об'єкта на фоні [49]. Мережа узагальнює дані кожного класу на підставі аналізу і оброблення для системи фрагментів зображень і визначає характеристики класу за процедурою навчання. Потім ці характеристики у формі параметрів, векторів чи вагових коефіцієнтів використовуються в процесі класифікації.

Сьогодні інтенсивно розвиваються також структурні методи класифікації, де прийняття рішення базується на описі зображення у формі множини дескрипторів ключових точок [11, 50, 51]. Ключові точки і їх дескриптори формуються спеціальними фільтрами – детекторами [7, 44].

Сучасні структурні методи реалізують процес класифікації за моделлю «мішок слів», де клас об'єкту визначається за результатом голосування складових компонентів – дескрипторів розпізнаваного об'єкту [11, 52]. При цьому у традиційному підході базовою множиною виступають описи еталонів, а у модифікаціях – центри чи параметри описів [50, 53, 54]. Такі методи мають переваги перед нейромережами, так як базуються на узгодженні з описом чи параметрами еталонів і не потребують довготривалого навчання [51]. Крім того, при їх застосуванні допускається оперативна зміна складу бази еталонів і забезпечується інваріантність до геометричних перетворень.

Виходячи з цього, викликає практичний інтерес здійснити комбінування апарату дескрипторів ключових точок та нейромережових засобів навчання у методах класифікації. Такий підхід у позитивному плані спрощує процес навчання за рахунок попереднього виділення множини дескрипторів для бази еталонів і при цьому забезпечує високі показники як швидкодії оброблення, так і точності класифікації [55]. Крім того, він може бути універсально застосований не тільки для візуальних даних, а і для будь-яких даних, що задані у формі множини багатовимірних векторів [3, 56].

Метою дослідження є покращення результативності структурних методів класифікації шляхом впровадження процедури навчання з використанням комітету мереж Кохонена. Основна увага приділяється підвищенню рівня точності класифікації зі збереженням високої швидкодії на підставі використання результатів навчання.

Задачами дослідження є опрацювання у порівняльному аспекті різних способів навчання мережі, включаючи ансамблі мереж з роздільним навчанням для описів еталонів. Ключовим є вивчення результативності впровадження моделей навчання у методах класифікації, а також проведення експериментального оцінювання запропонованих способів за наслідками аналізу бази зображень.

Під час дослідження запропоновано:

- формальну модель класифікації з використанням результатів роздільного багатомодульного навчання за кортежем еталонних описів. Детально проаналізовано способи впровадження мережі Кохонена у формі комітету. Реалізація поданих ідей забезпечує високі показники результативності та швидкодії класифікаторів;

- способи організації навчання на наборі описів еталонів у формі множини дескрипторів. Обговорюється схеми навчання на множині дескрипторів усієї бази, а також роздільно для описів еталонів. Пропонується схема побудови класифікатора на базі сформованої за результатом навчання системи ознак;

- результати комп'ютерного моделювання. Розроблено програмні засоби для впровадження обговорюваних моделей навчання та методу класифікації. Тестування виконано для бази зображень. За результатами експериментів підтверджено ефективність запропонованого підходу.

Апарат квантування даних і нейронних мереж є на сьогодні потужним інструментом здобування знань [30, 47, 48, 55, 57–59]. Стосовно описів об'єктів у формі множини дескрипторів ці інструменти дають можливість шляхом попереднього групування даних великого об'єму значно прискорити процес класифікації зображень [50, 52, 58]. Це досягається за рахунок здійснення швидкісного пошуку даних при зіставленні з еталонними, що зберігаються у пам'яті системи [50, 53].

Однією із практично ефективних мереж, яку застосовують для широкого спектру даних, є мережа Кохонена [57, 58]. На сьогодні вивчено використання мережі Кохонена для здійснення процедури самонавчання для множини дескрипторів і класифікації зображень на основі отриманих параметрів [55, 60].

При цьому обговорюються різноманітні варіанти для визначення центроїдів квантованих даних як для формування систем ознак, так і в процесі навчання [59, 61].

Але більш глибоке вивчення застосувань мережі для даних у вигляді множини дескрипторів показує [11], що векторні значення центроїдів даних, що отримані для різних зображень, можуть бути дуже близькі між собою. Це пояснюється не тільки багатовимірністю дескрипторів [52], що відображають особливості візуальних образів. Справа в тому, що сигнал зображення має досить великі розміри і практично неосяжне різноманіття, тому реально містить величезний обсяг інформації. Із-за цього навіть візуально не схожі між собою зображення часто можуть складатися із практично однакових чи близьких фрагментів, що призводить і до подібності їх подання у просторі ознак. Близькість характеристик центроїдів, обчислених за результатом навчання для спільного набору класів, знижує точність класифікації у методах голосування, що реалізують технологію «мішок слів» [2, 11].

Щоб уникнути чи знизити вплив цього фактору, необхідно підвищити точність подання даних для навчання і подальшої класифікації. Один із способів – це отримання концентрованого опису роздільно для кожного еталону за результатом навчання у формі подання не одного, а кількох центрів даних [4]. Тоді для навчання фактично маємо комітет мереж, у якому кожна складова незалежно формує фіксовану кількість центрів опису для окремого еталону. Число компонент такого ансамблю співпадає з числом еталонів. При цьому апарат мереж Кохонена виступає засобом для концентрованого подання значного обсягу багатовимірних векторів у формі обмеженого набору центрів даних. Крім того, навчання мережі на окремому еталоні сприяє збільшенню узгодженості центрів з даними в аспекті ефективності подальшої класифікації.

Відмітимо, що при використанні комітету схема класифікації, заснована на визначенні релевантності описів аналізованого об'єкту та еталону залишається незмінною [4, 11]. Підрахунок голосів класів може бути організовано як незалежно для кожного із еталонів [50], так і конкурентним шляхом на множині обчислених центроїдів для усієї бази [4].

Обговорюваний принцип удосконалення шляхом кількісного розширення представництва даних можна універсально впровадити і для інших способів здійснення концентрації даних [53, 61], таких як кластеризація [50], хешування [52], оцінювання [62]. Параметр визначеного числа центрів дієво впливає на результативність класифікації загалом.

Авторська монографія [57] містить результат вивчення різноманіття підходів до побудови та навчання мереж Кохонена, що дає можливість адаптувати структуру мережі до способу оброблення та різновиду аналізованих даних. Загалом застосування мережі Кохонена у системах комп'ютерного зору націлене на виявлення найбільш значущих ознак зображення, які в подальшому застосовуються для розпізнавання [31, 58–60].

Дослідники визначають такі позитивні властивості мережі Кохонена, як здатність розпізнавати стійкі кластери в аналізованих даних, а також можливість встановлювати близькість класів [57]. Таким чином, можна поліпшити розуміння досліджуваної структури даних, щоб потім впровадити ефективні класифікаційні правила у ситуації, де класи вже задані своїми представниками [55].

Мережа Кохонена навчається методом послідовних наближень [69]. Завданнями навчання при цьому є не стільки мінімізація помилки, а в більшій мірі адаптація внутрішніх параметрів мережі задля найбільшої узгодженості з вхідними даними.

У статті [55] нами запропоновано варіант мережі, що здійснює навчання на множині дескрипторів для повного складу бази еталонів. Аналіз метрик для здійснення класифікації на підґрунті інтегрованих ознак зображень наведено у роботі [63].

Останнім часом обговорюються варіанти ансамблевого навчання згорткових нейронних мереж для розпізнавання зображень. Вивчається ефективність різних правил голосування виходів цих мереж: жорстке (мажоритарне) голосування, де меншість голосів підкоряється більшості; жорстке голосування із заданим порогом для мінімуму набраних голосів; зважене і м'яке ймовірнісне голосування. Ці ж підходи актуальні і для здійснення класифікації на підставі голосування дескрипторів опису [4, 51].

Загалом на сьогодні вважається, що ансамблеве навчання із використанням кількох простих моделей мереж сприяє покращенню швидкодії класифікації зі збереженням чи підвищенням показника точності у порівнянні з однією, але більш складною моделлю навчання [64–66].

Як бачимо за аналізом літератури, впровадження нейромережевого навчання у структурні методи класифікації може сприяти не тільки покращенню швидкодії класифікації, а і забезпечує достатній рівень результативності.

2.2 Формалізація класифікатора

Будемо здійснювати класифікацію у рамках бази із N еталонів (представники класів) у формі множини E описів еталонних зображень: $E = \{E_1, E_2, \dots, E_N\}$. E – це навчальна вибірка, яка одночасно є підґрунтям для класифікації шляхом порівняння з еталоном [50]. Кожний еталонний опис $E_k \subset E$ у формалізмі класифікатора репрезентує окремий клас. Клас k з описом E_k визначається як нескінченна множина зображень, отриманих із еталону з номером k шляхом застосування до нього багатопараметричної групи геометричних перетворень, що включає зміщення, поворот, масштабування, дія яких не виводять об'єкт інтересу із поля зору [22, 45, 53].

Справа в тому, що сучасні детектори ключових точок, наприклад ORB, BRISK, AKAZE забезпечують інваріантність отриманого опису до такої групи перетворень [7, 44].

Опис еталону $E_k = \{e_v(k)\}_{v=1}^s$ подаємо як скінченну множину із S дескрипторів ключових точок у просторі B^n бінарних векторів, $e_v(k) \in B^n$, $s = \text{card } E_k$ – число дескрипторів у множині [11]. Кожний дескриптор $e_v(k)$ бази E має параметр k номеру класу, а загалом число ознак – дескрипторів у базовій множині E складає $\text{card } E = sN$.

Класифікатор R_1 , заснований на по-елементному аналізі вхідного опису $Z = \{z_v\}_{v=1}^s$ об'єкта, відносить дескриптор $z_v \in Z$ до одного із класів за правилом

$$R_1 : z_v \rightarrow \{1, \dots, N\}. \quad (2.1)$$

Клас k аналізованого дескриптора z_v об'єкту за правилом (2.1) визначимо як аргумент мінімуму відстані на множині E дескрипторів усіх класів

$$R_1 : k = \arg \min_{i=1, \dots, N; d=1, \dots, s} \rho(z_v, e_d(i)). \quad (2.2)$$

У конкурентній моделі (2.2) доцільно використовувати відстань Хемінга, що обчислює число відмінних бітів у двох бінарних векторах [4, 52].

У виразі (2.2) мінімізується значення відстані за змінною i (номер класу) на множині дескрипторів бази E .

У той же час у формулі (2.2) може бути організовано визначення мінімуму як за повним списком $s \cdot N$ дескрипторів бази, так і роздільно для кожного із N еталонів (з лімітом на величину мінімуму) з наступним визначенням класу-переможця за найбільшим числом голосів [53].

Введемо вектор $\{h_i\}_{i=1}^N$ з цілими значеннями для накопичення голосів класів, отриманих застосуванням моделі (2.2) до усієї множини дескрипторів об'єкту Z . На підставі впровадження R_1 для кожного $z_v \in Z$ відповідно до (2.2) визначимо номер класу k , а потім інкрементуємо акумулятор $h_k = h_k + 1$ для відповідного номеру класу. Фактично вектор $\{h_i\}_{i=1}^N$ є розподілом числа голосів для дескрипторів об'єкту у системі класів.

За результатом оброблення опису Z об'єкту накопичуємо вектор $\{h_i\}_{i=1}^N$. Клас об'єкту визначимо правилом R_2 як аргумент максимуму

$$R_2 : Z \rightarrow E_k \left| (k = \arg \max_{i=1, \dots, N} h_i) \& (h_k \geq \delta_h), \right. \quad (2.3)$$

де δ_h – поріг для мінімально допустимого числа голосів, який визначають експериментально для заданої бази як множини еталонів. Якщо умова $h_k \geq \delta_h$ не виконується, клас об'єкту не встановлюється (відмова від класифікації).

Послідовність правил класифікації R_1 , R_2 реалізує класифікатор на підґрунті рішень локальних класифікаторів, отриманих для компонентів об'єкту [11, 52]. Він має стійкість до викривлень окремих компонентів із-за можливого впливу завад [15].

2.3 Застосування мережі Кохонена для отримання центрів даних

Застосуємо до складу кожного еталонного опису E_i , $i = 1, \dots, N$, незалежно від інших описів еталонів, процедуру квантування даних на M непересічних груп мережею Кохонена (векторне квантування) [55, 57]. Зауважимо, що результат квантування залежить від початкових M значень центроїдів $W = \{w_j\}_{j=1}^M$ мережі [67].

Процес навчання здійснимо за моделлю конкуренції на множині центроїдів $W = \{w_j\}_{j=1}^M$.

Для компонента даних z визначається номер ν найближчого центру як

$$\nu = \arg \min_{j=1, \dots, M} \rho(w_j, z). \quad (2.4)$$

Найбільш часто модель (2.4) застосовується у схемі онлайн навчання, де дані поступають по черзі, а система реагує зміною системи центроїдів на кожному кроці шляхом модифікації нейрона з номером ν :

$$w_\nu(t+1) = w_\nu(t) + \alpha(t)(z(t) - w_\nu(t)), \quad (2.5)$$

де t – крок ітерації навчання;

$z(t)$ – значення навчальної компоненти на кроці t ;

$\alpha(t)$ – коефіцієнт швидкості навчання.

Вирази (2.4), (2.5) є одним із поширених варіантів правила «переможець забирає все», якщо змінюється лише центр, визначений за (2.4) [57]. При здійсненні навчання на множині E_i дескрипторів значення t означає номер дескриптора у описі: $t = \{1, 2, \dots, s\}$.

Зважаючи на бінарний вид аналізованих даних, для забезпечення збіжності та скорочення часу у процесі навчання впровадимо нормування даних шляхом ділення на норму Хемінга $\chi[z]$ для бінарного вектора z (число одиничних бітів) [52, 63]

$$\chi[z] = \sum_{\alpha=1}^n z^\alpha,$$

де z^α – біт з номером α у бінарному векторі $z \in B^n$.

У традиційних схемах навчання коефіцієнт $\alpha(t)$ зменшується від якогось початкового значення зі зростанням t . Найбільш поширена модель $\alpha(t) = \alpha_0 / t$ з деяким апріорно заданим початковим значенням α_0 швидкості навчання.

Помилку квантування даних традиційно обчислюють за формулою [67]:

$$\beta = \sum_{i=1}^M \sum_{\nu=1}^{s(i)} \rho^2(z_\nu, w_i), \quad (2.6)$$

де $s(i)$ – потужність i -го кластера, створеного за результатом навчання;

z_ν – елемент кластера з номером ν .

Зауважимо, що безпосередньо величина (2.6) як характеристика ефективності навчання не впливає на точність класифікації даних, а є лише апостеріорною оцінкою згуртованості даних навколо визначених у результаті навчання центрів.

Важливим моментом процесу навчання, результати якого будуть використовуватися для класифікації, є вибір початкових значень центроїдів [57]. Для цього пропонуємо в кожному із описів еталонів обрати дескриптор із найбільшим значенням параметра інформативності. Обчислимо інформативність довільного дескриптора як різницю між найменшими відстанями до множини дескрипторів із інших еталонів і свого еталону [2, 55].

Формально для вектора $z \in E$ як складового елемента $z \in E_i$ у системі класів E впровадимо критерій інформативності $V(z, E)$

$$V(z, E) = \rho_m(z, \overline{E_i}) - \rho_m(z, E_i), \quad (2.7)$$

де $\rho_m(z, \overline{E_i}) = \min_{v, u \neq i} \rho(z, e_v(u))$ – відстань від z до множини елементів бази, що не належать класу E_i , $e_v(u) \in E \setminus E_i$;

$\rho_m(z, E_i) = \min_{v, u=i} \rho(z, e_v(u))$ – відстань від z до найближчого елемента із класу E_i .

У моделях (2.4)–(2.7) може бути використана метрика Хемінга, якщо центри даних будуть приведені до бінарного виду. Але, зважаючи на необхідність нормування даних, більш практичним при реалізації навчання (2.4), (2.5) є застосування евклідової чи манхеттенської метрик [63].

2.4 Класифікація на підґрунті навчання

Застосування системи отриманих у навчанні центроїдів для процесу класифікації трансформує модель (2.2) до вигляду

$$R_1 : k = \arg \min_{i=1, \dots, N; d=1, \dots, M} \rho(z_v, w_d(i)), \quad (2.8)$$

де пошук мінімуму тепер здійснюється виключно на множині створених центрів даних.

Відмітимо, що множина центрів $W = \{w_j\}_{j=1}^M$ має набагато меншу потужність M у порівнянні з потужністю s множини дескрипторів опису, так як $M \ll s$. Це дає суттєвий вигреш у швидкодії оброблення при переході від лінійного пошуку всередині бази до пошуку на множині центрів [50, 53].

Наприклад, для конкретних значень $s = 500$, $M = 3$ при переході від моделі (2.2) до моделі (2.8) маємо вигреш у швидкодії оброблення більше ніж у 150 разів!

Зауважимо, що у моделі класифікації (2.8) може бути використана як метрика для векторів з дійсними значеннями (наприклад, манхеттенська [63]), так і метрика Хемінга, як у моделі (2.2). У другому випадку компоненти системи $W = \{w_j\}_{j=1}^M$ центроїдів треба трансформувати до бінарного виду.

Якщо застосувати навчання Кохонена для повної бази E [55], не розділяючи E на підмножини E_i , то в результаті отримаємо спільний для бази E набір із L центроїдів $W = \{w_j\}_{j=1}^L$. Тоді модель (2.8) для першого етапу R_1 трансформується у визначення найближчого центру з номером

$$u = \arg \min_{d=1, \dots, L} \rho(z_v, w_d). \quad (2.9)$$

Як результат застосування (2.9) $\forall z_v \in Z$ отримаємо для аналізованого об'єкта Z деякий цілочисловий вектор кластерного подання $g(Z) = [g_1, g_2, \dots, g_L]$ через систему центрів [50]. Тут g_d – число дескрипторів аналізованого опису, що потрапили у кластер з номером d .

Вектор $g(Z)$ є розподілом даних аналізованого опису за набором кластерів на відміну від $\{h_i\}_{i=1}^N$ як розподілу за системою класів. За цією ж схемою кластерне подання $g(E_i)$ можна отримати і для кожного із еталонів E_i . Завершену модель (2.3) класифікації тепер можна реалізувати як

$$R_2 : Z \rightarrow E_k \Big| k = \arg \min_{i=1, \dots, N} \rho(g(Z), g(E_i)), \quad (2.10)$$

де $\rho(g(Z), g(E_i))$ – відстань (наприклад, манхеттенська) між векторами кластерного подання об'єкта $g(Z)$ та еталону $g(E_i)$.

Обмеженням для здійснення класифікації за виразом (2.10) є умова $\rho(g(Z), g(E_k)) \leq \delta_\rho$.

δ_p – найбільше допустиме значення мінімуму відстані між кластерними поданнями.

Пропонована схема класифікації з використанням центрів даних, отриманих мережею Кохонена в процесі навчання, наведена на рисунку 2.1.



Рисунок 2.1 – Схема класифікації з використанням багатомодульного навчання

Запропонований багатомодульний спосіб навчання роздільно для кожного з еталонів у складі бази описів (рис. 2.1) сприяє підвищенню ступеня апроксимації і розділенню даних, що загалом покращує результативність класифікатора. Результативність класифікації будемо оцінювати традиційно значенням показника точності pr , який обчислюється відношенням числа правильно класифікованих об'єктів q до загального їх числа Q , що використовувалися в експерименті [15, 52, 63]

$$pr = q / Q. \tag{2.11}$$

2.5 Способи і умови здійснення навчання

Моделювання запропонованих методів виконано у програмному середовищі Google Colaboratory мовою програмування Python із використанням бібліотеки NumPy для побудови моделі та швидкого обчислення, а також бібліотеки комп'ютерного зору OpenCV [42, 68].

Експериментальним шляхом на підставі програмного моделювання оцінимо і порівняємо результативність методів класифікації з такими різновидами способів навчання:

– спосіб № 1 – спільна мережа Кохонена із використанням даних повної бази E описів з числом $L = N = 5$ центроїдів за моделлю (2.9);

– спосіб № 2 – комітет із N мереж Кохонена з одним нейроном ($M = 1$) для кожного еталона E_i , $i = 1, \dots, N$ (кожна мережа навчається на множині дескрипторів свого еталону);

– спосіб № 3 – комітет із N мереж Кохонена, M центрів для еталону E_i .

Навчання способу № 1 відрізняється від решти тим, що в якості навчального набору даних використовується сукупна база дескрипторів усіх еталонів, які потім конкурентною процедурою квантування розподіляються серед N центрів класів. Таким чином тут реалізується традиційний метод кластеризації, подібний до k -середніх, який фактично є пакетною варіацією алгоритму навчання мережі Кохонена [55, 57, 59].

Способи № 2 та № 3 здійснюють навчання роздільно для кожного із описів еталонів. Це дає можливість створення багатомодульного класифікатора (рис. 2.1), де кожний модуль прив'язано до окремого еталону.

Спосіб навчання № 2 є варіантом більш загального способу № 3 для випадку $M = 1$.

На початку навчання для усіх способів будемо ініціювати нейрони значеннями еталонних дескрипторів, які мають найвище значення інформативності (2.8), тобто є найбільш унікальними представниками свого класу. Здійснено нормування даних для навчання і для класифікації шляхом ділення на норму Хемінга $\chi[z]$ для бінарного вектора [52] (число одиничних бітів). Для забезпечення однорідності процесів оброблення даних у всіх розглянутих способах для класифікації застосована манхеттенська метрика.

У способі № 3 навчання мереж початкові центри обиралися випадковим чином із множини дескрипторів кожного еталону. Потік векторів із опису для здійснення навчання теж формувався випадково.

Оберемо еталонами, які будуть використовуватись для навчання класифікатора, зображення колекційних монет із образами міфічних богів (рис. 2.2) [68]. Розмір зображень становить 770×770 пікселів. Для формування ключових точок використаємо детектор ORB, значення дескрипторів якого є бінарними векторами розміром 256 біт [7, 44].



Рисунок 2.2 – Зображення монет бази еталонів

Для експерименту обрано значення числа класів $N = 5$, для способу № 3 число центроїдів всередині класу $M = 3$. На рисунку 2.3 наведено приклади зображень еталонів із сформованими координатами ключових точок.



Рисунок 2.3 – Приклади множини координат дескрипторів ORB для еталонів

Для оцінювання точності класифікації сформовано і застосовано тестову вибірку зображень, яка складається із еталонів та зображень, отриманих застосуванням до зображень еталонів геометричних перетворень повороту, масштабування, зміщення. Приклади трансформованих зображень із координатами ключових точок наведено на рисунку 2.4.



Рисунок 2.4 – Приклади тестових зображень із координатами точок

Загальна кількість тестових зображень становить 360 екземплярів. Оцінювання точності класифікації здійснювалось за показником (2.11). Клас зображення визначався класифікатором R_2 за максимумом числа голосів, істинність рішення оцінювалась за відповідністю правильному класу.

2.6 Аналіз результатів експериментів

Тестування проводилось із врахуванням впливу кількості епох навчання та початкового коефіцієнту α_0 на ефективність класифікатора із застосуванням моделей (2.5), (2.9). За одну епоху навчання уся множина дескрипторів підлаштовується до системи центрів. Наступна епоха – навчання повторюється зі зменшеним коефіцієнтом α_0 . Кількість епох варіювалась від 5 до 25, а коефіцієнт α_0 – між 0,1 та 0,9. Ансамбль з великою кількістю епох та низьким початковим коефіцієнтом навчання називають ретельною конфігурацією, а з незначною кількістю епох при великому початковому значенні коефіцієнту навчання – швидкою конфігурацією [48, 57].

Більш детально розглянуто такі варіанти: швидка (5 епох із початковим коефіцієнтом $\alpha_0 = 0,75$, помірна (10 епох, $\alpha_0 = 0,25$) та повільна (50 епох і коефіцієнт $\alpha_0 = 0,1$).

Кількість дескрипторів в описах еталонів та вхідних (тестових) зображень зафіксована на рівні $s = 500$.

Проведені експерименти показали наступні результати.

Класифікація на тестовій вибірці при здійсненні навчання способом № 1 показала такі найвищі значення точності:

- $pr = 0,4$ при $\alpha_0 = 0,25$ і числу епох 10;
- $pr = 0,39$ при $\alpha_0 = 0,75$ і числу епох 25.

Класифікація на тестовій вибірці при здійсненні навчання способом № 2 показала такі найкращі значення точності:

- $pr = 0,8$ при $\alpha_0 = 0,1$ і числу епох 5;
- $pr = 0,8$ при $\alpha_0 = 0,75$ і числу епох 25.

Як бачимо, здійснення навчання роздільно для кожного еталону дає суттєво вищу точність 0,8 класифікації у порівнянні з навчанням на повній базі описів (точність 0,4).

Тепер здійснимо порівняння цих показників зі значеннями таблиці 2.1, отриманими при навчанні запропонованим способом № 3 для $M = 3$.

Таблиця 2.1 – Точність (2.11) класифікації при способі № 3 навчання

Коефіцієнт α_0	Число епох			
	5	10	25	50
0,1	0,769	0,800	0,811	0,788
0,25	0,636	0,658	0,802	0,755
0,5	0,580	0,627	0,888	0,802
0,75	0,619	0,660	0,788	0,750
0,9	0,630	0,602	0,683	0,722

Жирним шрифтом у таблиці 2.1 виділено три найбільш ефективні конфігурації навчання за критерієм точності класифікації на тестовому наборі. Отримані показники таблиці 2.1 дають можливість визначити для практичних цілей характеристики навчання, що забезпечують необхідну точність класифікації. За порівнянням отриманих результатів бачимо, що застосування мережі Кохонена для навчання класифікатора призводить до суттєвого підвищення результативності: точність класифікації зростає за найкращими показниками від 0,4 для методу без впровадження комітету до 0,8 при незалежному для еталонів навчанні з одним нейроном і до ще вищого рівня 0,89 з трьома нейронами.

При цьому середнє значення із врахуванням усіх епох і коефіцієнту навчання теж суттєво зросло від 0,11 для способу № 1 та до 0,713 для дослідженого способу № 3.

Метод класифікації на підставі способу № 1 навчання виявив значно нижчу результативність у порівнянні зі способом № 2. Це можна пояснити тим, що у способі № 1 навчання здійснювалося на сукупній базі дескрипторів.

У цьому варіанті дескриптори «чужих» класів суттєво впливають на обчислення центрів, так як природно можуть бути більш наближені до «чужого» центру.

Таблиця 2.2 містить відносне число голосів, отриманих для зображень трансформованих еталонів у процесі класифікації при навчанні способом № 3 (500 дескрипторів, $\alpha_0 = 0,25$, 10 епох).

Таблиця 2.2 – Відносне число голосів класів для способу навчання № 3

Номер еталону	Число голосів				
	1	2	3	4	5
Еталон 1	0,383	0,182	0,141	0,196	0,098
Еталон 2	0,134	0,441	0,117	0,187	0,121
Еталон 3	0,184	0,218	0,269	0,210	0,119
Еталон 4	0,157	0,201	0,154	0,367	0,121
Еталон 5	0,193	0,179	0,139	0,220	0,269

Рядок таблиці 2.1 означає розподіл числа голосів за набором класів.

Із таблиці 2.2 бачимо, що в середньому тестові зображення з усіх класів класифікуються правильно, так як усі максимуми числа голосів «свого» еталону розміщуються на діагоналі.

На жаль, для способів навчання № 1 та № 2 така картина не характерна.

Крім того, із матеріалу таблиці 2.2 можна побачити, що найкраще класифікуються зображення класу 2 (найбільший максимум), а найгірше – зображення класів 3 та 5. Отримані показники дають можливість оцінювати і керувати ефективністю системи класифікації шляхом підбору найбільш ефективних параметрів навчання і еталонних зображень.

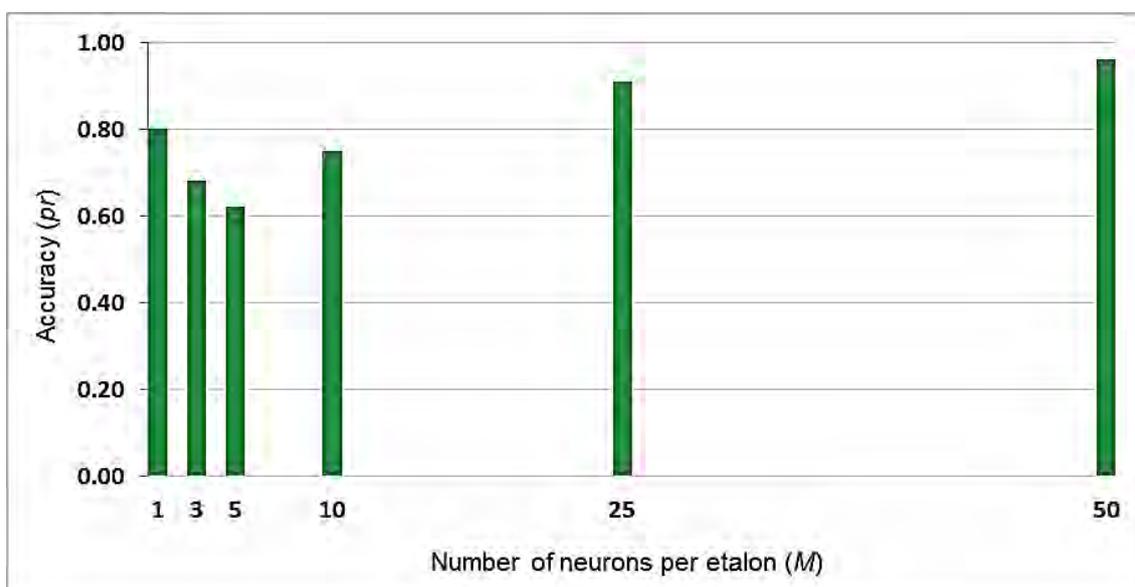
Окремі експерименти, проведені для пропонованого способу № 3 при більшому числі нейронів на еталон, підтвердили подальше зростання показника точності класифікації. При 5 нейронах найвища точність становить $pr = 0,74$, при 10 нейронах – $pr = 0,86$, при 25 нейронах – $pr = 0,94$.

При цьому середні значення із врахуванням епох і коефіцієнту навчання теж суттєво зросли.

Наприклад, середні значення для 10 епох складають:

- 0,65 – для 5 нейронів;
- 0,80 – для 10 нейронів;
- 0,91 – для 25 нейронів на еталон.

На рисунку 2.5 наведена типова залежність показника точності класифікації (2.11) від числа нейронів на еталон для 10 епох навчання при $\alpha_0 = 0,25$.



Рисунку 2.5 – Типова залежність точності класифікації від зміни M

Як бачимо, характерним є деяке зниження точності в точці $M = 5$, а потім продовжується зростання показника (2.11).

Зі зростанням числа M нейронів нелінійно також зростає час на тренування моделі та виконання класифікації. На рисунку 2.5 наведено залежність часу однієї епохи навчання (500 дескрипторів у описі) від числа нейронів.

У таблиці 2.3 видно, що до значення $M = 5$ час на тестування дещо знижується, причому час на навчання є мінімальним саме при 5-ти центроїдах одночасно із найменшою точністю (рис. 2.5). Також у таблиці 2.3 показано час класифікації одного зображення в залежності від використаного числа нейронів. Після цього із перевищенням числа 5 нейронів час на тренування та класифікацію монотонно зростає разом із точністю. Повторні тести підтверджують нелінійну природу цієї залежності.

Таблиця 2.3 – Залежність часу навчання і класифікації від параметра M

Кількість нейронів на еталон	Час, с	
	навчання	класифікації
1	1,68	0,061
3	1,82	0,043
5	0,40	0,010
10	1,70	0,037
25	1,86	0,052
50	2,28	0,086

Отримані експериментальні дані (рис. 2.5, табл. 2.3) та значення таблиць, аналогічних таблицям 2.1 та 2.2 для різноманіття способів навчання, відкривають практичну можливість у прикладній задачі обрати конкретні параметри для здійснення процедури навчання мережі потрібної якості.

Зрозуміло, що спосіб № 1 навчання показав нижчі результати точності в основному із-за того, що тут навчені центри отримані як результат конкуренції між класами. Це призвело до деякого згладжування різниці між центрами за результатом навчання, що вплинуло на точність. Відомо, що у множинах дескрипторів різних класів є багато спільних чи близьких компонентів [11, 52]. Більш чітку різницю центрів класів для цього способу можна було б отримати за рахунок впровадження контрольованого навчання чи навчання з учителем.

Швидкодія класифікації для запропонованого методу на підставі навчання ансамблю мереж Кохонена (модель (2.7)) у порівнянні з традиційним способом лінійного пошуку (модель (2.2)) зростає пропорційно відношенню s/M [50, 53]. Для використаних в експерименті значень $s = 500$, $M = 3$ маємо вигреш у швидкодії оброблення більше ніж у 150 разів. Зі зростанням числа M нейронів точність класифікації природно зростає, зате вигреш у швидкодії зменшується. На підставі проведеного нами дослідження можна обрати таку структуру мережі, яка забезпечить необхідний рівень критеріїв точності і швидкодії класифікації у прикладній задачі з конкретними описами еталонів.

Дане дослідження розвиває ідею поєднати у єдиному способі класифікації переваги нейронної мережі і апарату голосування дескрипторів ключових точок зображення. Комітет мереж шляхом роздільного навчання сприяє якісному формуванню обмеженої за об'ємом множини інтегрованих ознак еталонів, а голосування на підставі узгодженості з еталонними описами спрощує класифікаційне рішення.

У результаті отримано швидкісний метод класифікації, що забезпечує достойний рівень точності.

Комітет мереж у цьому дослідженні застосовано виключно для навчання системи класифікації роздільно для кожного із образів еталонів, що фактично призвело до синтезу багатомодульного класифікатора. Отримані дані у формі системи центроїдів впроваджено у процес класифікації. Навчання мереж у ансамблі та зіставлення із центроїдами можна організувати паралельно. Розділивши навчальні набори та обчислюючи центри кластерів для кожного класу окремо, ми запобігаємо впливу дескрипторів чужих класів.

Крім цього, роздільне навчання може бути здійснено у ситуації, де описи еталонів мають різну потужність.

Значний внесок цього дослідження полягає у впровадженні апарату мереж Кохонена для роздільного за класами еталонів навчання класифікатора зображень за описом у формі множини дескрипторів ключових точок. За результатом багатомодульного навчання отримана система центрів для еталонних описів, застосування якої суттєво скорочує час класифікації. Проведені експерименти підтверджують вигреш у швидкодії із гарантованим рівнем точності класифікації.

Обговорювані підходи до навчання і багатомодульного використання мережі Кохонена можуть бути універсально застосовані для аналізу і класифікації будь-яких об'єктів, описи яких задано множиною багатовимірних векторів.

Перспективи дослідження можуть бути пов'язані із удосконаленням класифікатора шляхом використання засобів навчання з учителем для комітету мереж, так як в обговорюваних підходах вважаються заданими представники класів – еталони. Це могло б сприяти подальшому підвищенню точності, особливо для першого способу навчання.

Матеріали дослідження опубліковано у статті [68].

3 ВПРОВАДЖЕННЯ КЛАСТЕРНОГО ПОДАННЯ ОПИСУЗОБРАЖЕНЬ

3.1 Кластерне подання як індексована структура даних

Ключові завдання комп'ютерного зору пов'язані із формуванням корисних висновків та результативних управлінських рішень щодо фізичних об'єктів реального світу на основі аналізу їх зображень [10, 32, 69–71]. Перевагами структурного опису в аналізі зображень є уявлення візуальних об'єктів у вигляді множини структурних елементів, що дає можливість у процесі розпізнавання здійснювати результативні рішення та забезпечити потрібну стійкість до перешкод на аналізованому зображенні [14, 18, 72–75]. Розвинення структурного розпізнавання пов'язане з удосконаленням технологій класифікації у просторі описів у вигляді множини дескрипторів ключових точок зображень, а також оцінюванням ефективності розпізнавання на зразках.

Досягнення високих показників результативності для методів класифікації у сучасних системах комп'ютерного зору потребує вирішення ряду нагальних проблем, пов'язаних із багатовимірною природою даних [22, 55, 76, 77]. При впровадженні структурних методів класифікації описи візуальних об'єктів подаються у вигляді множини дескрипторів ключових точок (КТ) – числових векторів високої розмірності [4, 5].

Описи із бінарними компонентами з використанням детекторів Oriented FAST and Rotated BRIEF (ORB) [7, 43, 44], Binary Robust Invariant Scalable Keypoints (BRISK) [6] містять відповідно 256 бітів та 512 бітів. Число дескрипторів у складі опису досягає від 500 до 1500 елементів. У такій ситуації трансформація простору ознак шляхом подання у вигляді кластерної системи сприяє значному спрощенню їх прикладного впровадження [4, 77]. Основним інструментом при цьому є апарат кластерного аналізу даних, який базується на встановленні метричних відношень на множині дескрипторів опису і визначенні близьких за значенням груп із опису об'єкта та еталонів [4, 35, 77].

Не менш важливою проблемою є вибір моделей для встановлення відповідності між параметрами кластерного подання різних описів [24, 35, 47]. До таких параметрів відносять центр кластеру, який може визначатися на підставі ряду підходів, основними серед яких є медіана та модифікації середнього значення. Ці дослідження пов'язані з напрямом Content Based Image Retrieval (CBIR) [13, 69] у розвитку систем комп'ютерного зору.

Зазначений підхід націлено на впровадження інтелектуальних технологій розпізнавання на основі вмісту зображення із ієрархією представлення наявної інформації [77]. Формування опису даних у вигляді множини дескрипторів ключових точок, а також кластерне подання опису виступають засобами узагальнення задля синтезу продуктивних рішень високого рівня [16, 39, 51].

Обчислювальна перевага класифікації за множиною центрів кластерів опису у порівнянні з традиційним підходом класифікації безпосередньо за дескрипторами опису обґрунтовується тим, що замість обрахунку релевантності множин визначається релевантність на основі їх обмеженого числа характеристик [9, 22].

Якщо потужність опису із дескрипторів ключових точок сягає від 200 до 300 (а іноді і до 1000) елементів, то кількість центрів може бути зменшена до 2–10 елементів такої ж розмірності. Релевантність обчислюється шляхом аналізу об'ємної множини еталонів, що у прикладних задачах досягає значень від 100 до 1500, тому виконана трансформація суттєво (у тисячі разів) скорочує обсяг необхідних обчислень [35].

Таке прискорення спричиняє необхідність здійснювати кластеризацію даних розпізнаваного об'єкта та деяке зниження ступеня розрізненості. Це відбувається за рахунок формування та застосування узагальнених характеристик – набору центрів опису. Кластеризація еталонних даних бази при цьому виконується на підготовчому етапі і на швидкодію класифікації безпосередньо не впливає [55, 77].

Фундаментом структурного розпізнавання є апріорно заданий набір описів еталонних зображень (алфавіт класів), представлений множиною структурних ознак – дескрипторів.

База зображень подана множиною $E = \{E_i\}_{i=1}^N$ еталонних описів.

Здійснимо на множині E розбиття $E = T = \{T_f\}_{f=1}^k$, $T_f \cap T_d = \emptyset$, $T_f \neq \emptyset$.

Вважаємо елементи $e \in T_f$ еквівалентними (рівноцінними) між собою.

Формально маємо дві моделі для системи класифікації на наявній множині ознак:

– $\{E_i\}$ – для зображень;

– $\{T_f\}$ – для ознак у вигляді кластерного подання.

За рахунок розбиття $\{T_f\}_{f=1}^k$ кожний опис $\{E_i\}$ еталону кількісно трансформується до цілочислового вектору виду

$$h[E_i] = \langle h_{i,1}, h_{i,2}, \dots, h_{i,j}, \dots, h_{i,k} \rangle, \quad (3.1)$$

де $h_{i,f} = \text{card}\{e \mid e \in E_i \ \& \ e \in T_f\}$, k – кількість кластерів.

Подання (3.1) – це образ еталону, що отриманий у результаті кластерного подання усієї бази, а скінченна множина векторів (3.1) бази із N зразків утворює матрицю $H[E] = \{\{h_{if}\}_{i=1}^N\}_{f=1}^k$ опису бази. Рядок матриці H відображає кластерний опис окремого еталону, а стовпець – зміст кластеру у складі рівноцінних елементів різних еталонів.

У загальному випадку для найбільш поширених процедур кластерування, де використовуються векторні дані з нецілими компонентами (k -середніх, ієрархічна класифікація та інші [4, 55]) можна застосувати евклідову або манхеттенську відстані.

Метою дослідження є вивчення ефективності різновидів структурного методу класифікації зображень на підставі формування кластерних систем ознак шляхом впровадження моделей для визначення релевантності трансформованих описів із використанням апарату голосування та оптимального зіставлення.

Задачами є опрацювання моделей даних задля обчислення релевантності описів за показниками кластерного подання, ефективного за швидкістю оброблення даних, а також вивчення результативності впровадження цих моделей за наслідками експериментального оцінювання запропонованих підходів для прикладної бази зображень.

Сучасні структурні методи аналізу візуальної інформації засновані на оцінюванні подібності дескрипторних описів аналізованого та еталонного зображень, представлених у вигляді множин дескрипторів ключових точок [4, 6, 7, 43, 72]. Величина подібності відображає ступінь спрацювання бінарного відношення релевантності для двох інформаційних одиниць як описів візуальних об'єктів.

Множина структурних ознак у вигляді сукупності числових векторів формується застосуванням спеціальних фільтрів – детекторів, у даній роботі використовуються методи ORB та BRISK [6, 7]. Детектори забезпечують отримання векторів-дескрипторів, що інваріантні до геометричних перетворень зсуву, повороту, масштабу [72].

Побудова розбиття у плані отримання моделі еталонів як груп структурних елементів вирішується у різний спосіб [4, 35].

Один з шляхів, що розвивається – застосування кластеризації на множині E . Кластерне розбиття можна налаштувати на структуру даних, що визначається базою [35].

Сучасні методи глибокого навчання з використанням спеціалізованих мереж для оброблення націлені на виявлення значущих відмінностей у класифікованих даних задля досягнення найбільш вагомих показників результативності для фіксованих наборів даних [13, 24, 32, 70].

У той же час розвиваються методи, які за рахунок комбінованого впровадження принципів локального та глобального аналізу даних дають можливість суттєво спростити процес класифікації із забезпеченням потрібного рівня ефективності [4, 5, 22, 55, 73, 77].

Останнього часу набули застосування методи нечіткої кластеризації даних [78], в тому числі і для задач аналізу зображень. Однак проведені нами дослідження [35, 75, 77] в значній мірі зосереджені на засобах традиційної кластеризації з непересічними кластерами, так як для наявних даних опису у вигляді множини багатовимірних векторів-дескрипторів властива значна близькість між собою для різних еталонів. І нечітка кластеризація не дає можливості усунути цю проблему та покращити показники класифікації.

Зараз впроваджується ряд перспективних моделей аналізу зображень, заснованих на об'єднанні переваг і можливостей глобального та локального аналізу даних, що сприяє більш глибокому виявленню розрізнявальних ознак, наприклад, для складних задач класифікації транспортних засобів та ідентифікації людей [79]. У цих системах для генерації ознак використовують спеціалізовані мережі. Зрозуміло, що такі підходи пов'язані зі зростанням обсягів обчислень.

Безперечною перевагою застосування векторної моделі (3.1) при обчисленні релевантності образів в обчислювальному плані є те, що цей спосіб значно (в десятки разів) менш витратний, ніж зіставлення безпосередньо опису зображення із описами еталонних множин E_i [4].

Можна розглядати отриманий образ (3.1) об'єкта як деяку комбінаторну конфігурацію [4, 13], яка організовується з елементів базової множини за допомогою центрів $b = \{b_1, b_2, \dots, b_k\}$, $b_i \in M_i$, утвореної в процесі формування кластерної структури. У такій конфігурації візуальний об'єкт зображуємо вектором (3.1), компонент якого дорівнює кількості входжень базового елемента (із словника ознак) в опис об'єкта. Застосовують термін «рейтинг зустрічальності» для структурних ознак [77].

Процес побудови таких змін можна вважати самоорганізацією. Отримана в результаті кластеризації базова множина b центрів кластерів безпосередньо визначається набором зразків $\{E_i\}$, що служать фокусами в необмеженому просторі зображень, що розпізнаються комп'ютерною системою.

3.2 Модель кластерного подання структурного опису

Розглянемо базу E описів зображень еталонів розмірністю $N : E = \{E_1, E_2, \dots, E_N\}$. Кожен еталонний опис E_i представляє у задачі розпізнавання окремий клас та має вид скінченної множини дескрипторів ключових точок зображення: $E_i = \{e_v(i)\}_{v=1}^s$, де s – число дескрипторів ключових точок у описі [35, 72].

Кожний дескриптор $e_v(i)$ характеризує деякий окіл ключових точок зображення та є елементом векторного простору R^n скінченної розмірності $n : e_v(i) \in R^n$ з дійсними, цілими або бінарними компонентами. Потужності описів еталонів вважаємо рівноцінними для спрощення аналізу:

$$\text{card}(E_1) = \text{card}(E_2) = \dots = \text{card}(E_N) = s. \quad (3.2)$$

Виконання цієї умови завжди можна практично досягти відбором елементів із множини більшого розміру.

Застосуємо відображення $E \rightarrow T$ із простору образів (множини дескрипторів) у множину T непересічних кластерів, сформованих за деяким принципом [4, 35]. Кластер – це підмножина опису. Кожний образ E_i еталону тепер трансформується до M його непересічних підмножин $T_k(E_i)$:

$$E_i = T(E_i) = \bigcup_{k=1}^M T_k(E_i), \quad k = \overline{1, M}, \quad T_k(E_i) \cap T_j(E_i) = \emptyset. \quad (3.3)$$

За множиною $T_k(E_i)$ елементів кожного кластеру визначимо його параметр центру $b_{k,i}$, $k = \overline{1, M}$, який є ключовою характеристикою побудованої кластерної системи для аналізованих даних. Зауважимо, що кластеризація та центри $b_{k,i}$ можуть бути визначені на підставі достатньо широкого різноманіття процедур [4, 13, 35, 39, 77, 80].

У результаті кластерного подання образ E_i еталону сформовано у вигляді M непересічних підмножин-кластерів $T_k(E_i)$ з центрами $b_{k,i}$, $M * N$ – загальне число створених кластерів та центрів для бази еталонів.

Розпізнаваний візуальний об'єкт аналогічно опишемо скінченною множиною $Z = \{z_v\}_{v=1}^s$, де $z_v \in Z$ – дескриптори ключових точок, $s = \text{card } Z$.

Подібно обробленню еталонів застосуємо кластерне розбиття множини Z через відображення $Z \rightarrow T$, у результаті опис зображення об'єкту буде представлено M кластерами:

$$Z = T(Z) = \{T_k(Z)\}_{k=1}^M, T_k(Z) \cap T_j(Z) = \emptyset. \quad (3.4)$$

Для кожного кластеру $T_k(Z)$ визначимо параметри центрів $b_k(Z)$, які застосуємо у процесі класифікації. Для спрощення вважаємо число M кластерів тотожним для вхідного зображення та еталонів.

3.3 Способи класифікації на підставі центрів кластерів

Розглянемо класифікацію виду «об'єкт – еталон» на підставі обчислення значення релевантності їх структурних описів [4, 35, 77]. Замість комплектної множини елементів описів застосуємо прискорену схему класифікації на підставі використання центрів b_k системи кластерів об'єкту і еталону.

Визначення релевантності «об'єкт – еталон» реалізуємо як зіставлення множин центрів об'єкта та еталонів з подальшим визначенням найбільш релевантного представника серед еталонів бази.

Введемо у векторному просторі R^n деяку відстань ρ . Центри b_k кластерів загалом можна теж подати у бінарному виді. У бінарному просторі векторів більш ефективна за обсягом обчислень метрика Хемінга [77]:

$$\rho(b_k, z_v) = \sum_{a=1}^n 1(b_{k,a}, z_{v,a}), \quad (3.5)$$

де функція $1(b_{k,a}, z_{v,a})$ порівняння бітів з номером a для двох векторів дорівнює одиниці, якщо біти не співпадають і нулю у протилежному випадку.

Підрахуємо $M \times M$ усіх відстаней $\rho(T_{k_1}(Z), T_{k_2}(E_i))$, $k_1, k_2 \in [1, 2, \dots, M]$ між кластерними системами еталонів та розпізнаваного об'єкта шляхом обчислення відстаней $q = \rho(b_{k_1}(Z), b_{k_2}(E_i))$ між елементами множин центрів об'єкта та i -го еталону і означимо їх як $\{q_a\}$, $a = 1, 2, \dots, M^2$.

На основі значень множини $\{q_a\}$ можна із застосуванням традиційних підходів обчислити відстані між множинами: середнього зв'язку, ближнього сусіда, далекого сусіда, Хаусдорфа та інші, або їх численних модифікацій, пов'язаних із логічним аналізом чи обробленням значень q_a [39]. Відстань середнього зв'язку тут має вид

$$\rho_{av}(T_k(Z), T_k(E_i)) = \frac{1}{M^2} \sum_{a=1}^{M^2} q_a. \quad (3.6)$$

Одна із модифікацій для визначення відстані обчислюється шляхом додавання трьох найменших елементів попередньо ранжованої вибірки $q^* = q_1 \leq q_2 \leq \dots \leq q_{M^2}$ для набору відстаней [35]:

$$\rho_{3m}(T_k(Z), T_k(E_i)) = \sum_{a=1}^3 q_a^*. \quad (3.7)$$

Відстань (3.7) одночасно має властивості як диференційних, так і інтегральних метрик. Вона отримана у результаті додавання трьох незалежних мінімумів.

Як альтернативу (3.7) розглянемо додавання відстаней найближчого сусіда окремо для кожного з центрів об'єкту

$$\rho_{NN}(T_k(Z), T_k(E_i)) = \sum_{a=1}^M \min_{j=1, \dots, M} \rho(b_a(Z), b_j(E_i)). \quad (3.8)$$

Моделі відстаней (3.7) та (3.8) не гарантують, що окремому центру об'єкту буде відповідати єдиний центр еталону. Визначення центру з такими властивостями спирається на застосування угорського методу для оптимального призначення кожному центру об'єкту найбільш прийнятному центру еталона [13, 35].

Застосуємо угорський метод для оптимального встановлення відповідності між двома наборами кластерних центрів b_1, b_2 .

Вони отримані для описів Z_1, Z_2 з урахуванням можливого впливу завад, що призводять до відхилень від їх ідеальних значень [39].

Результатом впровадження угорського методу є формування максимального паросполучення для елементів двох множин з мінімізацією загальної вартості. Її можна оцінити у вигляді суми відстаней між парами окремих компонентів у b_1, b_2 . Процес зіставлення формально зводиться до оптимізаційної задачі

$$R(x) = \sum_{i=1}^M \sum_{j=1}^M \rho(b_{1i}, b_{2j}) x_{ij} \rightarrow \min, \quad (3.9)$$

де $b_{1i} \in b_1, b_{2j} \in b_2, x_{ij}$ – бінарна ознака;

$x_{ij} \in \{0,1\}$, що дорівнює 1 у випадку відповідності i -го та j -го елементів.

Рішення задачі (3.9) при обмеженні на однозначність відповідності ознак із порівнюваних наборів

$$\begin{cases} \sum_{j=1}^M x_{ij} = 1 \forall i = \overline{1, M}, \\ \sum_{i=1}^M x_{ij} = 1 \forall j = \overline{1, M} \end{cases} \quad (3.10)$$

мінімізує загальну відстань (3.9) між наборами центрів b_1, b_2 .

Загальна кількість кластерних центрів у нашій задачі невелика (3...5), оптимальні методи можна вдало застосувати із незначними вимогами щодо швидкодії оброблення.

Зазначені загальноприйняті відстані та вирази (3.7) та (3.8) для множин векторів можна було б застосувати і безпосередньо до описів E_i та Z , але із суттєво значнішим обсягом обчислень.

Класифікацію об'єкту за описом Z на підставі обчислених відстаней (3.6)–(3.8) між центрами аналізованих даних здійснюємо традиційно шляхом визначення найменшого серед значень для різноманіття еталонів

$$Z \rightarrow E_a : a = \arg \min_{i=1, \dots, N} \rho(T_k(Z), T_k(E_i)). \quad (3.11)$$

Схема такого способу класифікації зображена на рисунку 3.1.

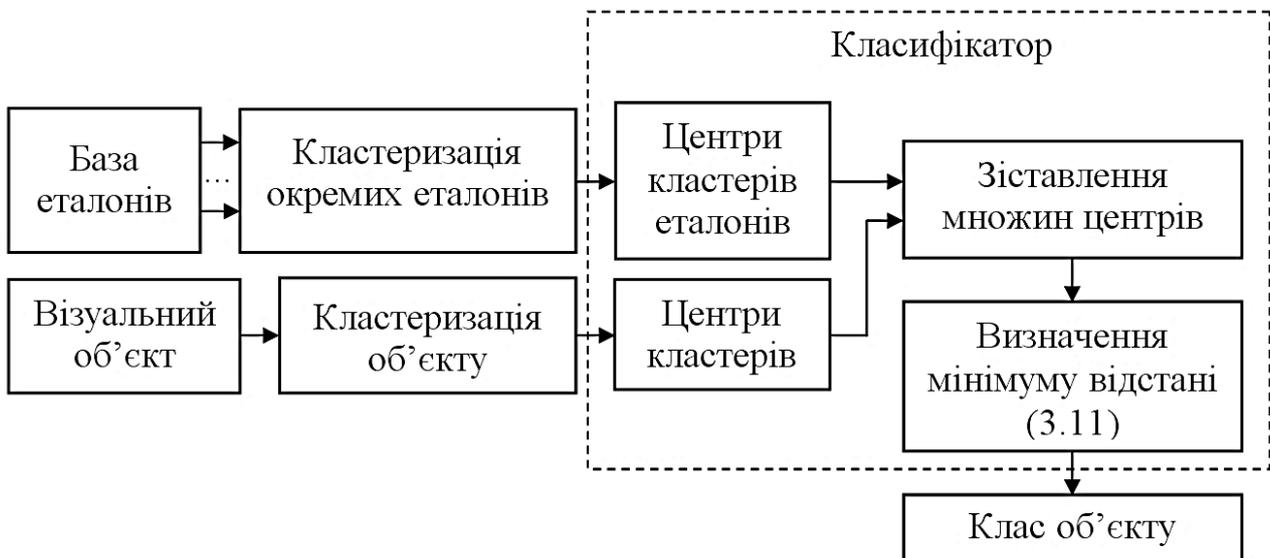


Рисунок 3.1 – Схема класифікації № 1

Незалежне кластерне подання для описів у базі еталонів сприяє також і спрощенню реалізації процедур класифікації виду «deskриптор об'єкту – еталон» стосовно окремих елементів опису розпізнаваного об'єкту. Цей підхід є більш універсальним щодо забезпечення врахування можливої дії завад на зображенні та фільтрації у процесі класифікації появи хибних ключових точок, викликаних впливом завад. При цьому кластеризацію опису об'єкта, як правило, не застосовуємо (це також скорочує обчислювальні витрати).

Кожний deskриптор опису об'єкта знаходить «свій еталонний клас» шляхом конкурентного зіставлення зі сформованою множиною $\{b_j(E_i)\}$ еталонних центрів.

Для кожного deskриптора $z_v \in Z$ об'єкту визначаємо найближчий серед усіх еталонних центрів $\{b_j(E_i)\}$ за процедурою ближнього сусіда

$$d = \arg \min_{j,i} \rho(z_v, b_j(E_i)), \quad d \in \{1, 2, \dots, N\}, \quad (3.12)$$

де ρ – відстань між deskриптором об'єкту та центром $b_j(E_i)$, $i \in \{1, 2, \dots, N\}$, $j \in \{1, 2, \dots, M \times N\}$.

Фактично (3.12) реалізує багатозначну характеристичну функцію $d: R^n \rightarrow \{1, 2, \dots, N\}$, що визначає еталонний клас щодо окремого deskриптора із опису об'єкта.

За результатом (3.12) $\forall z_v \in Z$ підраховуються кількості r_1, r_2, \dots, r_N голосів елементів $z_v \in Z$, віднесених до одного із центрів $\{b_j(E_i)\}$ опису E_i :

$$r_i = \sum_{v=1}^s F [z_v \rightarrow \{b_j(E_i)\}], \quad (3.13)$$

де F – логічна функція, що визначає віднесення елемента z_v до центру з номером j кластера еталону E_i .

Процедура реалізації функції F задля забезпечення фільтрації завд повинна спиратися на значення порогу δ_f для величини мінімуму серед відстаней, обчислених до кожного з наявних центрів $\{b_j(E_i)\}$ кластерів для усіх еталонів [35, 39, 72, 77].

Образ об'єкту класифікується за отриманими значеннями r_1, r_2, \dots, r_N голосів як

$$Z \rightarrow E_j : j = \arg \max_i r_i. \quad (3.14)$$

Вхідне зображення за класифікацією (3.12)–(3.14) буде віднесено до еталону, який набере найбільше число голосів дескрипторів його ключових точок.

Схема класифікації № 2 зображена на рисунку 3.2.

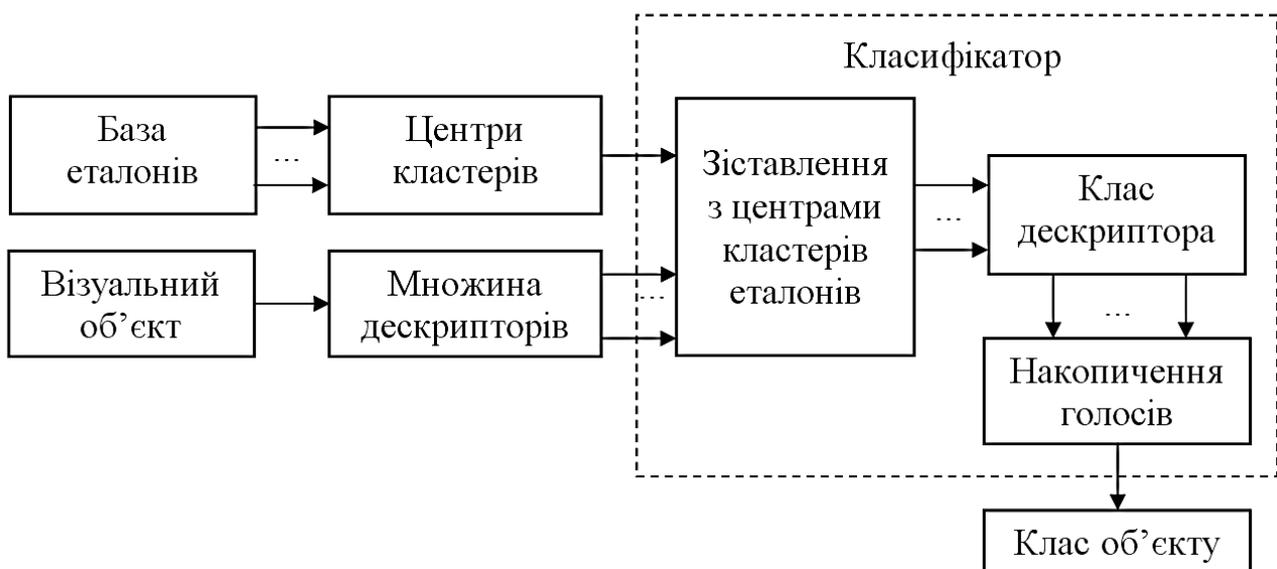


Рисунок 3.2 – Схема класифікації № 2

Розглянемо тепер модель № 3 класифікатора на підставі подання (3.1). Процес класифікації відрізняється від попередніх методів тим, що кластеризація попередньо здійснюється на повній множині еталонних описів і визначаються центри b_f^* , спільні для усієї бази еталонів [35, 77].

Почергово розглядається множина дескрипторів об'єкту $z_v \in Z$, поточний дескриптор відносять до кластеру T_d у відповідності до конкурентного правила

$$z_v \rightarrow T_d \mid \arg \min_f \rho(z_v, b_f^*) = d. \quad (3.15)$$

У результаті аналізу усієї множини Z отримують опис об'єкту $Z = (h_1, h_2, \dots, h_k)$ у вигляді (3.1). Далі у відповідності до моделі (3.11) визначається клас об'єкта з використанням векторної (наприклад, манхеттенської) відстані між кластерними поданнями об'єкту та еталонів у просторі цілочисельних векторів.

3.4 Експерименти і обговорення результатів

Запропоновані методи класифікації на підґрунті кластерного подання пройшли детальну експериментальну перевірку на різних базах зображень [35, 55, 72]. Розроблені моделі класифікаторів застосовано на прикладі зображень покемонів [35, 50]. Застосоване програмне середовище – IntelliJ IDEA 2020 [50] та IDLE із використанням засобів бібліотеки OpenCV [42] та мови програмування Java.

Ілюстрація зображень показана на рисунку 3.3.

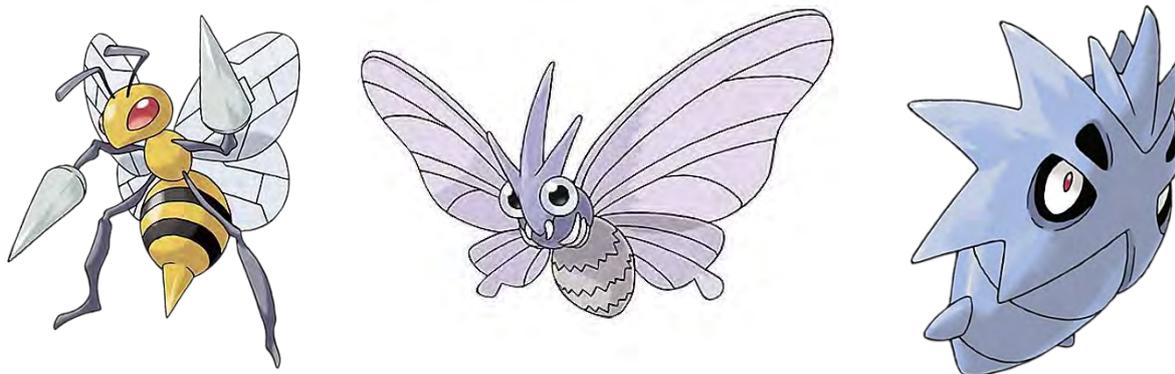


Рисунок 3.3 – Приклади аналізованих зображень

Для моделювання застосовано дескриптори ORB розмірністю $n = 256$.

Суть програмного моделювання зводилася до прикладної реалізації двох видів класифікаторів:

- з використанням голосування дескрипторів (12)–(14);
- на підставі встановлення оптимального співвідношення між центрами кластерів (9)–(11).

Перший метод реалізує принцип класифікації «дескриптор об'єкта – еталон», а другий – «об'єкт – еталон». Обидва підходи спираються на незалежну систему кластерів, сформовану для кожного із еталонів.

Класифікація здійснювалася за схемою, коли множина дескрипторів (перший метод) чи набір центрів кластерного подання одного із еталонів (другий метод) зіставлялися із множиною центрів кластерів для 10 еталонів із бази [35, 50]. Розглядалися варіанти описів із дескрипторів ORB кількістю в 300 ключових точок, 500 ключових точок, 1000 ключових точок та 1500 ключових точок. Результати обчислень нормованої кількості голосів (у відсотках) для вхідних зображень рисунку 3.3 (еталони № 3, № 5, № 6) на множині 10 еталонів при $s = 500$ наведено у таблиці 3.1.

Отримані максимуми виділено маркером у таблиці 3.1.

Таблиця 3.1 – Результати класифікації першим методом

Вхідні зображення	Еталони									
	1	2	3	4	5	6	7	8	9	10
3	7,4	8,8	23,8	9,6	12,6	7	6,8	9,4	6,6	8
5	11	9,8	10,2	7,2	29,8	9,2	3,8	5,8	3,2	10
6	9,6	9,8	8,2	12	9,4	26	6,6	4,8	5,6	8

Перший метод успішно класифікує вхідні зображення, причому перевага максимуму над найближчим значенням у рядку (інші зображення бази) достатньо впевнена для усіх вхідних зображень. Експерименти показали, що для вибраної бази еталонів число використаних дескрипторів ключових точок практично не впливає на результат. Для усіх розглянутих кількостей (300, 500, 1000, 1500) ключових точок отримано схожі показники. Це можна пояснити інтегрованим поданням описів бази еталонів у вигляді фіксованих наборів центрів кластерів.

Результати класифікації у вигляді значення найкоротшої відстані із застосуванням другого підходу на базі угорського методу для визначення оптимального паросполучення між центрами кластерів вхідного зображення рисунку 3.3 (еталони № 3, № 5, № 6) та еталону (еталони № 1–№ 10) для 500 ключових точок наведені у таблиці 3.2.

Таблиця 3.2 – Результати класифікації другим методом

Вхідні зображення	Еталони									
	1	2	3	4	5	6	7	8	9	10
3	212	230	133	242	201	197	224	193	179	198
5	248	182	227	276	133	221	238	193	199	214
6	232	196	227	248	211	49	212	213	183	106

Для опису вхідного зображення кластеризація здійснювалась повторно, інакше результат був би прогнозований із-за повного співпадання центрів. Отримані мінімуми виділено маркером у таблиці 3.2.

За результатом моделювання бачимо впевнену класифікацію для розглянутих еталонів. Значення обчисленої за угорським методом найкоротшої відстані суттєво менше у порівнянні з рештою еталонів. Найкраще при цьому класифікується еталон № 6 (третій рядок таблиці 3.2).

Аналіз експериментальних даних (таблиця 3.1 та 3.2) для розглянутої бази еталонів показує більш впевнену класифікацію першим методом у порівнянні з другим методом. Це можна пояснити більш детальним аналізом компонентів даних у першому методі. Другий метод базується виключно на інтегральних характеристиках описів за результатом кластеризації. Експерименти підтвердили можливість результативного прикладного застосування оптимальних математичних підходів (угорський метод) для задач, де дані представлено у концентрованому кластерному виді. Безпосереднє застосування цих методів до вхідних даних у вигляді кількох сотень векторів дескрипторів неможливо із-за обмежень у часі реалізації.

За результатами програмного моделювання був оцінений обсяг часу для реалізації розглянутих методів. Основну частину роботи програми займає пошук та опис ключових точок за допомогою детектору ORB бібліотеки OpenCV [7, 42, 43].

Параметр числа ключових точок суттєво не впливає на швидкість другого методу (це його перевага), вона визначається виключно числом кластерів [35, 77]. Із зростанням числа ключових точок не критично збільшується час реалізації першого методу.

Час реалізації одного акту класифікації для обох методів коливався в межах 1,1–1,5 секунд для досліджуваних даних зображень, числа ключових точок, методу кластеризації і використаного апаратного забезпечення.

Проаналізуємо особливості запропонованого методу № 3 на базі моделі (3.15) на прикладі експериментальної бази морських ссавців (рис. 3.4) при поданні їх структурних описів у межах 10 кластерів.



Рисунок 3.4 – Зображення морських ссавців

Під час дослідження здійснено визначення ключових точок для еталонних зображень на основі ORB та BRISK (бібліотеки OpenCV та JSFeat), кластеризацію (JavaScript, бібліотека k -Means), обчислення матриці відстаней для бази зображень (рис. 3.4).

Зауважимо, що візуально розглянуті зображення значно схожі між собою, і в природних умовах спостережень їх складно класифікувати навіть людині. Іншою особливістю даної бази є суттєва візуальна подібність 1 та 4-го зображень (кит та китеня).

Розміри зображень – 139×309 пікселів. Кластеризація описів бази здійснена методом k -середніх [55]. Координати визначених ключових точок показано на рисунку 3.5.



Рисунок 3.5 – Зображення координат ключових точок

У результаті отримано опис бази (матриця H кластерного подання) у вигляді таблиці 3.3.

Таблиця 3.3 – Кількість ключових точок еталонів у кластерному поданні

H	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}
Z^1	3	6	2	2	5	3	6	8	0	4
Z^2	11	12	0	15	8	14	9	0	0	0
Z^3	3	2	0	14	5	0	18	13	0	0
Z^4	4	1	2	3	4	5	6	1	1	1

Нормована матриця вагових коефіцієнтів показана в таблиці 3.4.

Таблиця 3.4 – Матриця ваг для бази морських тварин

α^1	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}
Z^1	0,18	0,38	0,42	0,08	0,27	0,17	0,19	0,43	0	0,74
Z^2	0,37	0,43	0	0,34	0,24	0,44	0,16	0	0	0
Z^3	0,13	0,09	0	0,40	0,19	0	0,40	0,50	0	0
Z^4	0,33	0,09	0,58	0,17	0,30	0,39	0,26	0,07	1,00	0,26

За даними таблиці 3.4 можна побудувати різноманіття методів класифікації, що базується на накопиченні ваг як аналогу функції належності до відповідного кластеру [35, 39, 72].

Як можна побачити, отримане у таблиці 3.4 подання виглядає досить ефективним.

У стовпцях 2, 3, 8, 9, 10 спостерігається значне переважання вагових коефіцієнтів для обмеженого числа еталонів. Матрицю подібності для бази еталонів можна обчислити, помноживши стовпець ваг кластера на число структурних ознак еталону, що опинилися в кластері (табл. 3.3).

Матриця відстаней подана у таблиці 3.5.

Таблиця 3.5 – Матриця відстаней класів

Еталони	1	2	3	4
1	0	58	42	21
2	58	0	58	51
3	42	58	0	47
4	21	51	47	0

З таблиці 3.5 бачимо істотні відмінності еталонів у просторі кластерів, тоді як еталони 1 та 4 природно залишаються найбільш схожими.

Ключовим параметром кластерного подання є число кластерів, яке теоретично пов'язане із порогом еквівалентності для структурних ознак. Скорочення обсягу обчислень при розпізнаванні залежить від числа кластерів та від складу аналізованих даних.

Для порівняння матриця кластерного подання у межах 5 кластерів на основі зображень рисунку 3.4 представлена у таблиці 3.6.

Таблиця 3.6 – Кластерне подання для 5 кластерів

H	M_1	M_2	M_3	M_4	M_5
Z^1	10	11	3	7	8
Z^2	19	14	0	20	16
Z^3	7	12	4	26	5
Z^4	8	4	3	6	7

Гістограма подання першого зображення (рис. 3.4) у системі 5 кластерів та 10 кластерів наведена на рисунку 3.6.

Значення стовпчиків (це перші рядки у таблицях 3.3 та 3.6) означають число дескрипторів зображення, віднесених до відповідного кластеру.

Проведене комп'ютерне моделювання показало безпомилкове розпізнавання аналізованих візуальних об'єктів при кількості кластерів від 5 до 10, зі збільшенням кластерів (більше 8) точність не покращується.

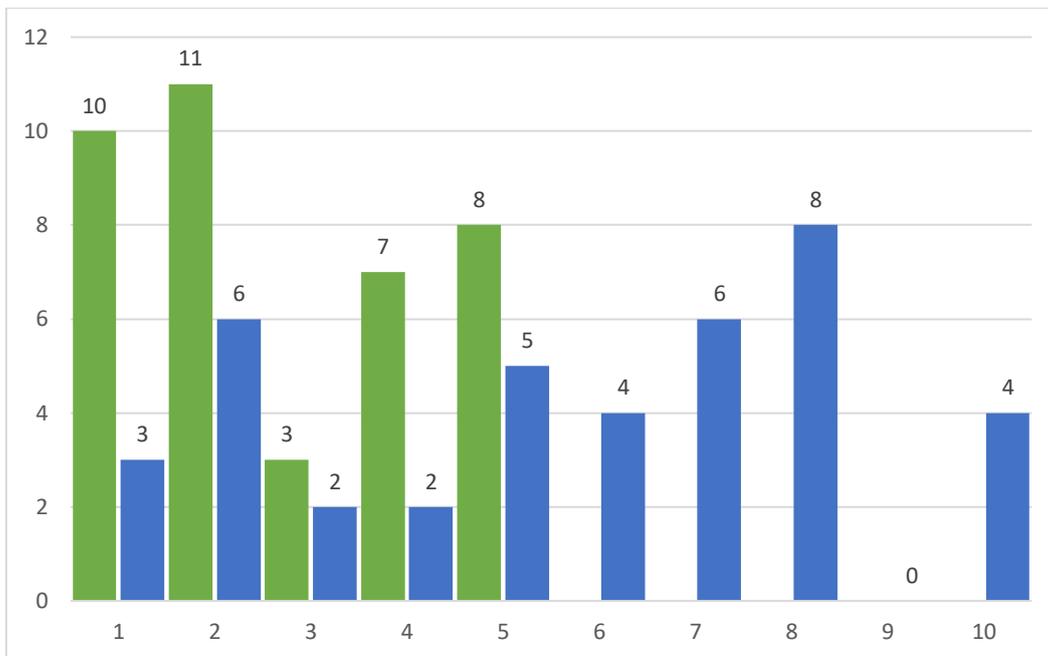


Рисунок 3.6 – Гістограма кластерів зображення рисунку 3.4 (зелений – 5 кластерів, синій – 10 кластерів)

Перехід до векторного виду істотно підвищує швидкодію розпізнавання за рахунок спрощення оброблення. Порівняння схем класифікації при застосуванні різних детекторів (ORB та BRISK) показує, що спосіб № 3 більш точний і чутливий, він розрізняє навіть дуже схожі між собою зображення.

Описані експерименти підтвердили працездатність та результативність запропонованих модифікацій методів на основі кластеризації.

Розроблені підходи показують суттєву перевагу при аналізі та класифікації об'ємних наборів зображень з довільними множинами векторів, дозволяючи прискорити загальний час розпізнавання в сотні разів.

Основна перевага запропонованого у даному дослідженні методу полягає у суттєвому зниженні обчислювальних витрат на класифікацію зображень із забезпеченням достатньо високого показника ефективності.

Впровадження кластерного подання на множині дескрипторів опису сприяє покращенню часових характеристик класифікації за рахунок застосування центрів кластерів із забезпеченням необхідного рівня результативності. Аналіз та оброблення згрупованих даних дає можливість сформувати ієрархічну структуру зі змінюваним параметром деталізації та виявленням значимих для класифікації властивостей опису зображення.

Розроблені методи класифікації на основі кластерного опису забезпечують достатній рівень розрізнення зображень. Кластерне подання у десятки разів скорочує часові витрати на розпізнавання.

Результативність кластерного розпізнавання значно залежить від сформованої системи кластерів у прикладній основі класів, що задані еталонами. На етапі навчання системи доцільно досягати приблизно рівночисельного подання для множин структурних ознак еталонів, щоб отримане кластерне подання призводило до рівноцінного впливу всіх структурних елементів на результат.

Для близьких зображень (рис. 3.4, еталони 1 і 4) доцільно створити єдиний еталонний опис.

За проведеними експериментами рекомендована кількість кластерів у прикладних завданнях комп'ютерного зору має бути в діапазоні 5...10.

Такі параметри методів, як поріг еквівалентності структурних ознак, поріг для значущості ступеня релевантності при визначенні певного класу, а також кількість кластерів визначаються у процесі навчання для конкретних прикладних баз зображень.

Перспективи дослідження пов'язані із впровадженням розроблених методів для об'ємних баз зображень, а також вивченням їх результативності шляхом використання у них сучасних засобів нечіткої та багатовимірної кластеризації.

Матеріали дослідження опубліковано у статті [50].

4 ІДЕНТИФІКАЦІЯ ВІЗУАЛЬНИХ ОБ'ЄКТІВ ЗА КЛАСТЕРНИМ ПОДАнням

4.1 Аналіз літературних джерел

Виявлення, ідентифікація та класифікація об'єктів є ключовими завданнями сучасних систем комп'ютерного зору, що здійснюють встановлення екземплярів візуальних об'єктів певного класу (наприклад, людей, тварин чи автомобілів) у складі цифрових зображень [72, 75, 81, 82].

Такі інтелектуальні завдання вирішуються як у загальному плані дослідження методів для виявлення різних типів об'єктів під уніфіковані рамки задля імітації людського бачення та пізнання [4, 16, 76], так і в прикладному плані під конкретні сценарії застосування, такі як, наприклад, виявлення пішоходів, обличчя, тексту, рухів.

За останні роки стрімкий розвиток методів глибокого навчання вносить новий здобуток у предмет виявлення, що сприяє прориву і просуванню вперед перш за все для прикладних впроваджень [43, 83–86].

Об'єкти виявлення зараз використовуються у багатьох реальних застосунках, таких як, наприклад, автономне водіння, робот-зір, відеоспостереження за рухомими об'єктами.

Структурні методи класифікації зображень набули популярності із-за їх прикладної ефективності у задачах комп'ютерного зору, де здійснюється ідентифікація чи класифікація візуальних об'єктів [72, 75, 86].

При цьому традиційно множина точок розпізнаваного об'єкту формується шляхом аналізу частини зображення, що виділена скануючою рамкою, тобто вікном, яке дає можливість в процесі аналізу частково виключити фонові об'єкти. Для кожного положення рамки приймається рішення про ідентифікацію (два класи) або класифікацію (декілька класів як еталонів) на підґрунті значення релевантності запиту та зображення всередині вікна.

При впровадженні структурних методів класифікації функція яскравості зображення представляється множиною ключових точок, кожна з яких описується вектором ознак – дескриптором ключових точок [81].

Формальна постановка задачі класифікації на підставі опису як множини дескрипторів ключових точок наведена в літературі [72].

Ідентифікація візуальних об'єктів на зображенні сцени може бути успішно реалізована в рамках методу зіставлення опису фрагменту зображення об'єкту та кластерним поданням еталону як запиту для пошуку [75, 81].

Тільки кластерне подання за рахунок суттєвого стиснення опису (як правило, обсяг аналізованого опису – 500 дескрипторів ключових точок і більше) дає можливість забезпечити пошук об'єкту у режимі реального часу. Завдяки переходу від множини до центрів багатовимірних даних значно скорочуються обчислювальні витрати та час прийняття рішення [16, 18, 75, 82, 87].

Метою дослідження є розроблення методу для пошуку візуальних об'єктів на зображенні сцени із використанням кластерного подання для структурного опису зображення запиту.

Задачі дослідження:

- розроблення математичних та програмних моделей інтелектуального аналізу даних при визначенні ступеня релевантності структурних описів вікна та запиту;

- вивчення особливостей застосування моделей для визначення релевантності з впровадженням кластерування даних запиту;

- оцінювання ефективності розробленого методу за результатами аналізу конкретних зображень.

Виділення та класифікація об'єктів на зображенні є головним завданням інтелектуальних систем комп'ютерного зору [72, 85]. Зараз дослідники в основному зосереджуються на методах, що безпосередньо націлені на прикладне впровадження. Зважаючи на багатовимірну і просторову природу сигналу зображення, найбільшу популярність при вирішенні цієї задачі набули статистичні підходи [39, 43, 72, 75, 83–86].

Останнього часу розвиваються спеціалізовані програмні засоби, що спираються на попереднє навчання нейронної мережі в рамках деякої фіксованої бази зображень [39, 81–84]. Наприклад, мережа YOLO поділяє зображення на регіони і передбачає обмежувальні рамки та параметри впевненості для кожного регіону одночасно.

Створена серія удосконалень на основі YOLO та запропоновано версії, які ще більше покращують параметри універсальності, впевненості та точності, зберігаючи високу швидкість виявлення [21]. Однак обмеженнями таких систем є необхідність попереднього тривалого навчання і залежність результатів застосування від конкретної бази, на якій здійснювалося навчання.

Незважаючи на наявність результативних систем на підґрунті машинного навчання, опрацювання та обґрунтування нових методів пошуку об'єктів продовжується [39, 43, 72, 75, 81–84, 86, 88].

Новий перспективний напрямок пов'язаний із використанням описів візуальних об'єктів у вигляді множини дескрипторів ключових точок. Цей апарат забезпечує високу швидкодію аналізу даних та дає можливість з метою класифікації детальніше встановити характеристики об'єкту, що виявляється на зображенні. У той же час допустиме комбінування різних методів задля підвищення результативності. Додаткове впровадження навчання в таких системах сприятиме подальшому покращенню їх характеристик [81, 82, 86].

4.2 Математичні моделі ідентифікації опису для зображення запиту

Універсально опишемо розпізнаваний візуальний об'єкт (запит, еталон) як скінченну множину $Z = \{z_v\}_{v=1}^s$, де $z_v \in Z$ – дескриптори ключових точок, $s = \text{card } Z$ – її потужність [81]. Для бінарних дескрипторів $\text{ORB } Z \subset B^n$, B^n – простір бінарних векторів розмірністю n [43, 83]. Застосуємо кластерне розбиття множини Z через відображення $Z \rightarrow T$, у результаті опис вхідного зображення об'єкту буде представлено M непересічними кластерами:

$$Z = T(Z) = \{T_k(Z)\}_{k=1}^M, T_k(Z) \cap T_j(Z) = \emptyset, \quad (4.1)$$

де $T_k(Z)$ – множина елементів фіксованого кластеру.

Вибір числа M кластерів є виключно прикладною задачею і залежить від змісту аналізованих даних. Зі збільшенням M зростає точність аналізу угруповань даних, зате збільшується і час оброблення.

У наших дослідженнях для описів у вигляді множини дескрипторів застосовано значення $M \in \{3, \dots, 10\}$ [72, 75, 81–83].

За результатом кластерування для кожного кластеру $T_k(Z)$ із опису запиту Z визначимо параметри центрів $T_k(Z)$ та потужностей $c_k(Z)$ кластерів

$$c_k(Z) = \text{card } T_k(Z), k = \overline{1, M}. \quad (4.2)$$

Тепер розглянемо фіксовані числом u вікна W_1, \dots, W_u , $W_i \subset B^n$, що представляють собою окремі фрагменти зображення, всередині яких можуть бути розташовані шукані об'єкти, що представлені множиною дескрипторів ключових точок.

Такі фрагменти можуть бути синтезовані за встановленим порядком огляду зображення в залежності від прикладної задачі [85]. Зрозуміло, що кількість фрагментів безпосередньо впливає на час оброблення. Для забезпечення рівнозначності впливу аналізованих даних на результат аналізу будемо вважати значення параметру s для кожного опису із множини вікон W_1, \dots, W_u однаковим, тобто

$$\text{card}(Z) = \text{card}(W_1) = \dots = \text{card}(W_u) = s. \quad (4.3)$$

Виконання умови (4.3) завжди можна практично досягти фіксацією значення s для запиту Z та відбором s елементів із множин W_1, \dots, W_u більшого розміру. У протилежному випадку необхідне додаткове нормування даних за кількістю елементів опису.

Ідентифікацію зведемо до встановлення ступеня релевантності $\rho(W_i, Z)$ об'єкта W_i та представленого у кластерному вигляді запиту Z із наступним прийняттям рішення на підставі значення $\rho(W_i, Z)$. Для кожного дескриптора $w \in W_i$ конкурентно визначаємо найближчий кластерний центр у наборі векторів $\{b_j(Z)\}$ за процедурою ближнього сусіда

$$d = \arg \min_{j=1, \dots, M} \rho(w, b_j(Z)), \quad d \in \{1, 2, \dots, M\}, \quad (4.4)$$

де ρ – відстань між дескриптором об'єкту та центром b_j із системи кластерів для запиту.

Процедуру оброблення (4.4) іноді називають проєктуванням на множину кластерних центрів [57, 75].

У випадку застосування дескрипторів та центрів бінарного виду у (4.4) може бути застосована відстань Хемінга. У загальному випадку для найбільш поширених процедур кластерування, де використовуються векторні дані з нецілими компонентами (k -середніх, ієрархічна класифікація та інші [21, 39, 57, 89]) можна застосувати манхеттенську відстань.

За результатом оброблення (4.4) $\forall w_a \in W_i$ підраховуються кількості h_1, h_2, \dots, h_M елементів аналізованого опису, віднесених до одного із кластерних центрів $\{b_j\}_{j=1}^M$:

$$h_j = \sum_{a=1}^S f_a[w_a \rightarrow \{b_j\}], \quad (4.5)$$

де f_a – логічна функція, що визначає віднесення елемента опису до відповідного центру j кластера запиту за моделлю конкурентності (4.4).

Процедура реалізації функції f_a задля забезпечення фільтрації завад, які неодмінно присутні на зображеннях, повинна спиратися на значення порогу δ_ρ для величини мінімуму в (4.4) [72]. Рішення $w_b \rightarrow b_j$ приймається при умові $\rho(w_a, b_j) \leq \delta_\rho$, де δ_ρ визначається експериментально, виходячи із складу еталонних зображень аналізованої бази.

На підставі обчислення компонентів вектору (4.5) визначаємо міру релевантності як відстань γ між цілочисельними векторами $h = (h_1, h_2, \dots, h_M)$ для запиту і поточного вікна $\gamma(h[Z], h[W_i])$. У залежності від отриманої величини визначаємося з рішенням про ідентифікацію W_i . У якості γ можна взяти манхеттенську відстань у M – вимірному векторному просторі. Крім того, у даному випадку у якості міри релевантності може бути використана подібність векторів, наприклад, коефіцієнт кореляції.

Зауважимо, що у моделях (4.4) та (4.5) описи W_i оброблюються незалежно один від одного, що дає можливість приймати рішення про наявність кількох пошукових об'єктів на зображенні [44, 55].

Розглянемо покрокову реалізацію запропонованого методу у вигляді етапів попереднього оброблення та ідентифікації.

Етап попереднього оброблення не впливає на час прийняття рішення про ідентифікацію і містить наступні кроки:

Крок 1. Обчислюємо дескриптори ключових точок запиту-еталону.

Крок 2. Здійснюємо кластеризацію структурного опису еталону.

Крок 3. Визначаємо центри та потужності кластерів.

Етап ідентифікації можна розглядати як послідовність дій (рис. 4.1):

– визначимо множину дескрипторів ключових точок розпізнаваного фрагменту (вікна) зображення;

- проєктуємо аналізований опис вікна на структуру кластерного подання запиту (центри кластерів);
- визначимо міру релевантності (відстань, подібність) представлень фрагменту та кластерних центрів запиту;
- за значенням міри релевантності приймаємо рішення відносно ідентифікації запиту та зображення вікна.



Рисунок 4.1 – Схема пошуку об'єкта на зображенні

4.3 Експериментальні результати програмного моделювання

Для досліджень використано програмне середовище Jupyter Notebook на сервісі Google Colaboratory. Програма, що моделює метод пошуку за запитом, написана мовою Python з використанням спеціалізованих бібліотек для роботи із зображеннями: scikit-image та OpenCV [42].

Вхідне зображення містить 4 об'єкти (цуценята), із яких об'єкт № 1 (цуценя зліва) використано у якості запиту.

На рисунку 4.2 наведено вхідне зображення, а рисунок 4.3 містить його напівтоноване подання, що обробляється детектором ключових точок.



Рисунок 4.2 – Аналізоване зображення



Рисунок 4.3 – Зображення рисунку 4.2 у відтінках сірого

Візуально помітна деяка відмінність об'єкта № 4 рисунку 4.2 (щученя справа) від решти, в експерименті теж це виявилось. Зрозуміло, що людський зір легко сприймає цей об'єкт як щученя, так як вмикається образ, створений мозком людини. Штучна інтелектуальна система [88] приймає рішення виключно на підставі множини інформативних точок зображення, для яких обчислено дескриптори ключових точок.

Розмір зображення складає 600×314 пікселів. Пошук ідентичних запити об'єктів виконувався шляхом сканування рамкою в $\frac{1}{4}$ розміру щодо вхідного зображення.

Фрагмент № 1 (перше цуценя зліва) розміром 150×314 взято у якості еталонного запиту (рис. 4.4).

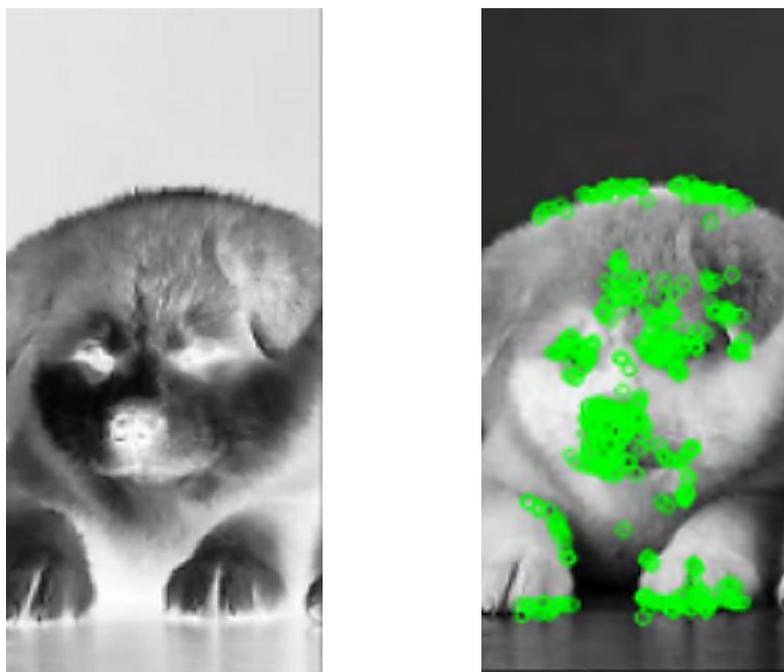


Рисунок 4.4 – Запит (еталон) та його координати ключових точок

Використано детектор ORB, що формує опис у вигляді множини близько 500 бінарних дескрипторів розмірністю 256 бітів [44]. Реалізація методу здійснена в умовах приблизно однакової кількості дескрипторів в описах запиту та аналізованого фрагменту.

Обговорюваний метод передбачає зіставлення рівноцінного числа компонентів. Ця умова легко досягається шляхом зміни використовуваного числа ключових точок аналізованого фрагменту. Кластеризація для опису запиту виконана методом k -means з використанням манхеттенської метрики та числом кластерів $k = 3$.

Спочатку був проведений пробний експеримент, у якому аналізовані фрагменти зображення (поточне вікно сканування) незалежно кластеризувалися у процесі сканування та ідентифікації. Таким чином, навіть для об'єкта № 1 (еталон) була виконана повторна кластеризація. Центри кластерів різнилися для одного і того ж зображення. При такому ускладненому обробленні важко розраховувати на успіх. Але навіть у цьому випадку об'єкти № 1 та № 3 сцени були ідентифіковані. Вони показали достатнє (невелике) значення манхеттенської відстані між гістограмами даних у межах від 46...115 (максимальне значення відстані – 500).

Наступний експеримент полягав у реалізації методу, де кластерне подання здійснювалося одноразово для запиту. Аналізовані структурні описи поточного зображення проєктувалися на зафіксовані центри кластерів запиту шляхом встановлення найближчого з них. Зрозуміло, що обчислювальні витрати для такого підходу значно менші. Адже затратна кластеризація даних здійснюється тільки одноразово для запиту.

Гістограму кластерного подання запиту (об'єкт № 1) для числа 500 дескрипторів ключових точок наведено на рисунку 4.5. Стовпці гістограми містять число дескрипторів зображення, віднесених до відповідного кластеру.

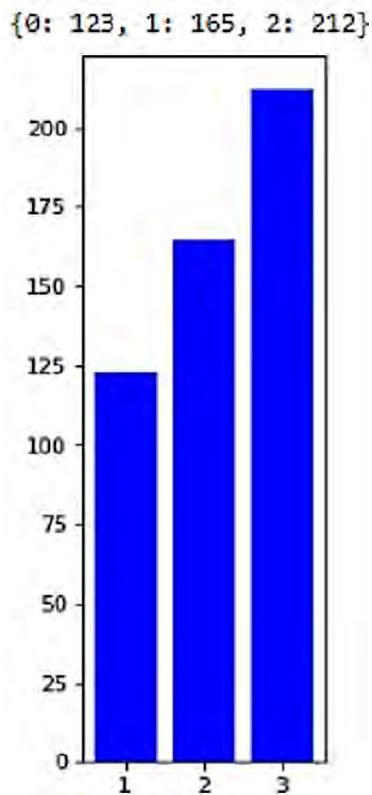


Рисунок 4.5 – Гістограма проєкцій запиту при числі ключових точок 500

Результати експерименту показали, що об'єкти № 1–№ 3 (рис. 4.2) ідентифікуються безпомилково (відстані дорівнюють 0, 80, 90), а об'єкт № 4 за результатом аналізу продемонстрував суттєву відмінність (відстань 212 при максимумі 500). При цьому один із фрагментів, що містить частини двох різних об'єктів, показав відстань 162, що ближче до еталону, ніж об'єкт № 4 (рис. 4.6).

Також проведено експерименти з різним числом ключових точок, що відбиралися випадковим чином із опису в 500 точок. Зрозуміло, що зі зменшенням числа ключових точок обчислювальна ефективність покращується, зате зменшуються розрізнявальні властивості [75].

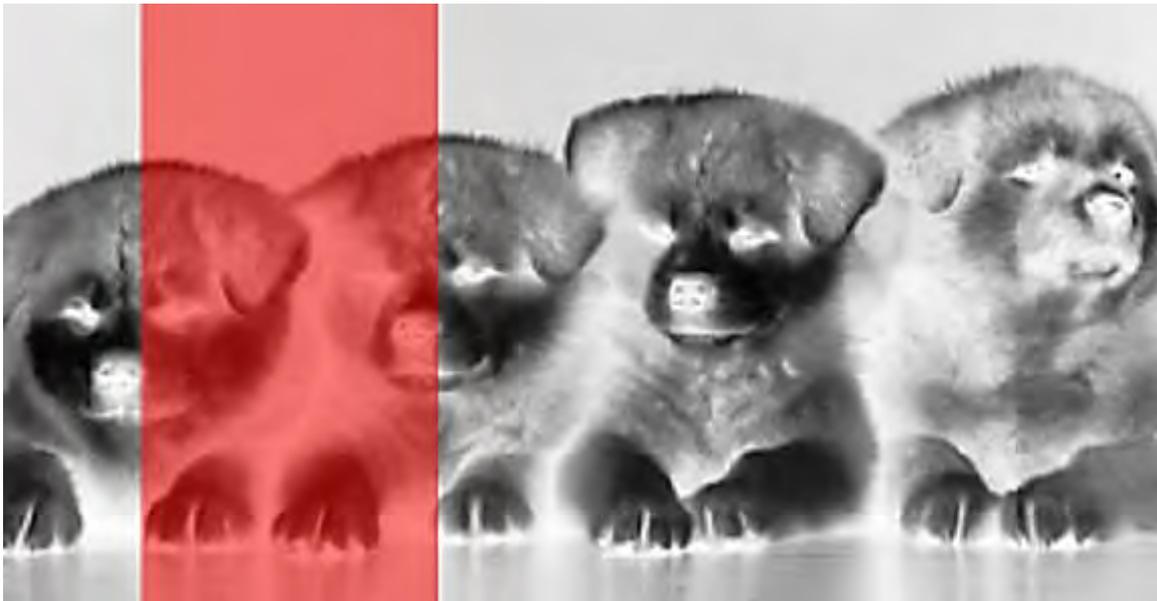


Рисунок 4.6 – Фрагмент з відстанню 162

Кращу прикладну ефективність показав варіант оброблення з числом ключових точок, що дорівнює 200. Конкретно для зображення запиту (рис. 4.4) число дескрипторів у кластерах склало вектор (60, 46, 94).

За результатом моделювання об'єкти-цуценята № 1–№ 4 ідентифіковані вірно (відстані 0, 26, 30, 42 при максимумі 200), в той час як решта аналізованих вікон показала суттєво більші значення відстаней.

Для фіксованого числа у 200 дескрипторів ключових точок нами проведено порівняльний експеримент, де для ідентифікації реалізовано традиційний метод голосування на підставі лінійного пошуку найближчого дескриптора без застосування попередньої процедури кластерного подання опису. Експеримент показав вигреш у швидкодії для розробленого методу у порівнянні з традиційним у 5,2 рази. При цьому величина виграшу залежить від параметру числа кластерів і посилюється зі зростанням кількості кластерів у межах 2...8.

Як бачимо із проведеного експерименту, ефективність методу можна посилити за рахунок зміни значень його параметрів та пристосування до властивостей даних.

Запропоновані методи пошуку об'єктів на зображенні з використанням кластерного апарату відзначаються високою швидкодією оброблення даних та достатньою результативністю. Їх ефективність можна посилити шляхом навчання та вибору таких параметрів методу, як розмір опису, ступінь стиснення множини дескрипторів, відбір інформативної підмножини опису, вибір способу кластеризації.

Розроблений метод можна поширити на багатокласову ситуацію, коли замість ідентифікації «об'єкт – фон» у кожному вікні здійснюється класифікація на кілька еталонних класів.

Наукова новизна дослідження полягає у розробленні методу пошуку об'єктів на зображенні візуальної сцени шляхом впровадження кластерування для даних опису запиту, що сприяє підвищенню швидкодії пошуку і забезпечує достатню результативність.

Практична значущість роботи – підвищення глибини аналізу візуальних даних та швидкодії класифікації, підтвердження результативності запропонованих методів на прикладах зображень, створення прикладних програмних засобів для вивчення та впровадження методів класифікації у новітніх системах комп'ютерного зору.

Подальшими етапами досліджень може бути побудова ієрархічних ознакових систем за ознаками структурного опису, а також врахування при обчисленні релевантності вагових характеристик кластерів, що відображають чисельність їх елементів.

Матеріали дослідження опубліковано у статті [21].

5 ГРАНУЛЯЦІЙНІ МОДЕЛІ СТИСНЕННЯ ОПИСУ ЗОБРАЖЕНЬ

5.1 Удосконалення шляхом стиснення опису

На сьогодні однією із найбільш вагомих і до цих пір не вирішених проблем у системах комп'ютерного зору є забезпечення необхідних показників їх прикладного функціонування, зокрема, точності та швидкодії оброблення даних. Особливо актуальним є вирішення цієї проблеми для задач класифікації візуальних об'єктів, де використовується об'ємна кількість прототипів класів, образи яких подаються у формі множини векторів. Традиційно такі задачі відносять до сфери Big Data [66], вони потребують здійснення об'ємного лінійного пошуку у багатовимірному просторі даних [25, 34, 62, 68, 90].

Одним із поширених у прикладних застосуваннях є неперервно-груповий підхід до розпізнавання, що спирається на детерміновану модель формування аналізованих зображень [27, 34]. Відповідно до цієї моделі необмежений за потужністю клас образів породжується дією апріорі невідомих топологічних перетворень над деякою фіксованою еталонною реалізацією [27, 42, 91].

Сучасні структурні методи класифікації здійснюють аналіз візуального об'єкту у формі множини дескрипторів ключових точок [6, 7, 27, 51, 91]. Дескриптори ключових точок обчислюються спеціальними фільтрами – детекторами. Дескриптор ключових точок подається у формі бінарного вектора розмірністю 256 чи 512 елементів [6, 7, 34].

Структурні методи здійснюють процес класифікації за моделлю «мішок слів», клас об'єкту визначається за результатом голосування для множини дескрипторів розпізнаваного об'єкту. Відмітимо, що структурний аналіз має певні переваги перед сучасними нейромережними технологіями, так як ґрунтується на безпосередньому узгодженні з описом чи параметрами еталонів і не потребує довготривалого навчання [34, 62, 68]. Крім того, при застосуванні структурних методів є можливість оперативної зміни складу бази еталонів, що робить їх універсальними щодо набору розпізнаваних зображень [14]. Ще однією вагомою перевагою впровадження структурних методів є можливість прийняття рішення за окремими частинами чи деталями опису візуального об'єкту. Функціонал апарату дескрипторів ключових точок має необхідну в прикладних застосуваннях властивість інваріантності до геометричних перетворень об'єктів у полі зору [27, 34, 42, 62, 91].

Можна констатувати, що застосування структурних методів для класифікації зображень у порівнянні із апаратом штучних нейронних мереж дає можливість уникнути таких проблем, як складнощі побудови ефективних алгоритмів навчання мережі [30], значні обсяги часу навчання [80], визначення продуктивної розмірності мережі, високі вимоги до обчислювальних ресурсів [16], необхідність забезпечення інваріантності до геометричних перетворень об'єктів, обмеженість кола задач для конкретної мережі виключно за результатом навчання та інші [11, 51].

Останнього часу для задачі класифікації зображень нейромережевими методами достойну ефективність показали методи глибокого навчання, що базуються на статистичних характеристиках текстури зображення. Але результативність цих методів, як і усіх нейромережових підходів, часто залежить від синтезованої структури мережі, яка, в свою чергу, пов'язана з прикладним застосуванням і умовами спостереження. У той же час при впровадженні апарату глибокого навчання задача формування стисненої множини ознак теж не втрачає актуальності.

У той же час структурні методи мають можливості для удосконалення шляхом зниження обчислювальних затрат на класифікацію [52, 53, 68]. Одним із напрямків є розроблення ієрархічних способів пошуку даних на підставі апарату кластерів, хешування чи оцінювання [42, 53]. Відомі також способи зниження розмірності структурного опису шляхом відсіювання дескрипторів за критерієм інформативності [42] чи унікальності. Наприклад, критерій інформативності базується на значенні різниці відстаней дескриптора до свого та інших еталонів у базі класифікації. Але ключовим фактором для результативної класифікації все-таки залишаються гуртові метричні співвідношення між множинами дескрипторів [68].

Підвищення продуктивності структурних методів можна досягти також скороченням числа дескрипторів у описі еталонів [48]. Будь-яке зменшення числа дескрипторів важливе, так як дає можливість уникнути складних для реалізації комбінаторних аспектів у задачі зіставлення багатовимірних множин ознак [22, 92].

Виходячи із цього, перспективним напрямком можна вважати застосування принципів грануляції даних [25, 52, 89, 93, 94]. Заснування грануляції в обчислювальному інтелекті із впровадженням апарату нечітких множин належить Л. Заде [94]. Грануляція – процес отримання значущих у якомусь сенсі агрегованих даних [25, 93].

Зокрема, грануляцію можна здійснити за критерієм нерозрізненості, який стверджує, що компоненти належить згрупувати разом, якщо вони демонструють подібність. Для опису даних можна використовувати різні ступені деталізації, що сприяє створенню ієрархічних систем з кількома рівнями грануляції [25, 52, 68]. Але досягнення потрібного рівня абстракції на основі опису даних і поставленого завдання вимагає експериментального обґрунтування.

Одним із вивчених методів скорочення опису є визначення множини центрів, побудованих на підставі процентильних відношень стосовно медоїду для множини даних [53]. Інші розроблені підходи реалізують видалення підмножини дескрипторів з низьким значенням інформативності [42, 68].

Ефективним прикладом втілення принципу грануляції для інтегрування у просторі ознак зображення є пірамідальне представлення [27, 34, 91]. Пірамідальна структура інтегрує яскравість для просторово близьких координат пікселів зображення і скорочує обсяг багатомасштабної зорової інформації. Загалом концепція грануляції даних набула поширення у системах розпізнавання образів. Вивчаються різні аспекти створення ефективного простору ознак у класифікаторах зображень [2, 34, 62, 90]. Гранулювання застосовують у методах сегментації об'єкт – фон задля покращення результативності [42, 51]. Запропоновано ефективні метрики у просторі гранульованих даних [95]. Отримані результати прикладних досліджень щодо використання стисненого подання описів [42, 53, 96, 97].

Методи кластерування, створення хеш-кошиків чи оцінювання центрів можна вважати метричною грануляцією, так як вони базуються на обчисленні близькості даних між собою з використанням метрики [25, 42, 50, 68, 89]. За рахунок впровадження цих підходів покращується швидкодія оброблення з допустимим зниженням показника точності класифікації. Однак досягнення необхідних значень результативності потребує більш глибокого вивчення усіх можливостей і нюансів прикладного застосування апарату грануляції і відсіювання з необхідністю об'єктивного підтвердження їх універсальності у сфері інтелектуального аналізу багатовимірних даних.

Діапазон методів метричної грануляції досить місткий. Він включає, наприклад, достатньо прості підходи, пов'язані із виявленням пар найбільш подібних дескрипторів опису і відкиданням одного із цієї пари. У той же час він включає і найскладніші багатоетапні методи кластерування у багатовимірних просторах даних [50, 68].

Наприклад, грануляція за подібністю (еквівалентністю) передбачає пошук метрично найближчого дескриптора та його виключення з подальшої обробки. Це дозволяє в разі зменшити обсяги обчислень, зберігаючи при цьому представництво дескрипторів у описі.

Як альтернативу інтелектуальним методам можна розглянути також скорочення опису через прості схеми відсіювання елементів за якимсь законом черговості у повному описі. Відсіювання передбачає виключення, наприклад, кожного другого елемента із множини даних, що дає можливість у два рази скоротити їх обсяг і зменшити об'єм обчислень. Це простіше у виконанні, не вимагає оцінки відстаней і здійснюється швидко. Ще одна із схем полягає у формуванні безпосередньо детектором ключових точок невеликого за обсягом опису, або випадковий відбір скороченої підмножини із повного опису.

Однак зрозуміло, що ці спрощені способи трансформації не враховують значимість і просторові відношення елементів, що може призвести до втрати важливої інформації та зниження точності класифікації.

Дослідження зосереджені на вивченні і експериментальному впровадженні засобів метричної грануляції і відсіювання даних у структурних методах класифікації, а також на аналізі їх ефективності у порівняльному аспекті із традиційним методом без грануляції.

Основною метою є вивчення впливу пропонованих модифікацій на швидкодію та точність класифікації, що є важливим для покращення продуктивності систем комп'ютерного зору у практичних застосуваннях.

Внесок проведеного дослідження полягає у детальному вивченні результативності створення ієрархічного подання опису з кількома рівнями ієрархії. Інший вклад – підтвердження практичних переваг щодо впровадження інтелектуального аналізу шляхом створення гранульованих описів у порівнянні із тривіальними схемами відсіювання і формуванням скорочених описів безпосередньо детекторами ключових точок.

5.2 Моделі метричної грануляції

Представимо структурний опис $Z = \{z_i\}_{i=1}^s$ візуального об'єкту (множина дескрипторів ключових точок) як сукупність $F(Z)$ трансформованих описів f_r , отриманих на послідовних рівнях ієрархічного подання

$$F(Z) = \cup \{f_r\}_{r=1}^R, \quad s(r) = \text{card } f_r, \quad (5.1)$$

де $s(r)$ – потужність опису f_r на рівні ієрархії з номером r .

При $r=0$ маємо найнижчий рівень, що представляє повний опис Z . Найвищим рівнем з номером R будемо вважати трансформацію у формі єдиного вектору даних. Прикладом є медоїд для множини векторів чи центр – вектор середніх елементів множини [42]. Ієрархії f_r будемо формувати послідовно рекурсивним шляхом, починаючи із найнижчого рівня (опис Z).

У результаті отримаємо структуровану послідовність описів

$$Z \rightarrow f_1 \rightarrow f_2 \rightarrow \dots \rightarrow f_R.$$

Мета побудови ієрархії – зниження розмірності опису задля покращення продуктивності класифікації шляхом прискорення процедур оброблення даних. Для створення послідовної ієрархії описів застосуємо на кожному кроці деяку процедуру T трансформації, що зменшує потужність опису при переході до наступного рівня

$$f_{r+1} = T\{f_r\}, \quad s(r+1) \ll s(r).$$

Поставимо задачу формування ієрархії вищого рівня до тих пір, поки класифікатор із використанням стисненого опису f_{r+1} буде результативно здійснювати класифікацію із заданою точністю. Основна задача трансформації T – прискорення процесу класифікації шляхом узагальнення даних опису.

Визначимо T як процедуру грануляції даних за моделлю формування одного вектора наступного рівня на підставі аналізу пари найбільш подібних між собою векторів попереднього рівня. Така дія застосовується у процесі кластерування даних [42, 50, 89].

Рекурсія дає можливість застосовувати одну і ту ж процедуру, що встановлює закон переходу до чергового рівня. Ця процедура продукує один елемент верхнього рівня на основі групи елементів нижнього рівня. Це модель пірамідально-рекурсивних структур. Фактично процедура T здійснює розбиття множини дескрипторів на пари еквівалентних елементів.

Із впровадженням T для опису еталону E_k з числом s дескрипторів, отримуємо трансформацію

$$T(E_k) \rightarrow E_k^*, \text{card}(E_k^*) = s(k) \ll s, \quad (5.2)$$

результатом якої є відфільтрований опис E_k^* , що є підмножиною $E_k^* \subset E_k$ елементів із E_k , відібраних за певним критерієм грануляції. На практиці пропонується знижувати потужність опису приблизно у 1,5–2 рази на кожному етапі обробки.

У результаті впровадження (5.2) база $E = \cup E_k$ набуває скороченого подання як $E^* = \cup E_k^*$.

Критерієм у (5.2) обрано визначення елемента наступного рівня через пару найближчих між собою елементів нижнього рівня. Таким чином, фабула метричної близькості лежить в основі формування гранули наступного рівня.

Принцип аналізу метричної близькості дескрипторів може слугувати основою для створення різних варіантів фільтрації (редукції) даних. Розглянемо конкретні варіанти побудови трансформації T , що базується на значенні метрики ρ для дескрипторів. Зрозуміло, що гранули наступного рівня можна отримати на підставі використання довільної логічної операції.

Грануляція за порогом еквівалентності.

Для кожного з дескрипторів $e_d(k) \in E_k$ визначається найближчий до нього за метрикою ρ дескриптор $e_*(k)$ у множині E_k . Це відповідає моделі лінійного метричного пошуку

$$e_*(k) = \arg \min_{v \neq d} \rho(e_v(k), e_d(k)). \quad (5.3)$$

Після визначення дескриптора з мінімумом метрики перевіряється умова еквівалентності: $\rho(e_d(k), e_*(k)) \leq \delta_\rho$, де δ_ρ – це заданий граничний параметр для значення метрики, що наближено встановлює еквівалентність двох дескрипторів. У просторі багатовимірних векторів величину δ_ρ часто встановлюють як 25% від максимального значення метрики, але її можна змінювати у залежності від поставленої задачі та вхідних даних [42, 52].

Якщо ж еквівалентність за порогом δ_ρ в (5.3) не встановлено, то перший аргумент заноситься до E_k^* , цей елемент виключаємо з подальшої обробки, а другий продовжує брати участь у аналізі. Зауважимо, що порядок слідування дескрипторів у описі та значення δ_ρ прямо впливають на результати грануляції.

Фактично модель (5.3) визначення еквівалентності двох векторів доповнюється предикатом P :

$$P(z_1, z_2) = 1 \mid \rho(z_1, z_2) \leq \delta_\rho. \quad (5.4)$$

Число отриманих векторів у цьому варіанті залежить від порогу δ_ρ .

Альтернативний спосіб реалізації (5.3) полягає у визначенні пари найближчих векторів без застосування умови (5.4). У такому разі обсяг опису від рівня до рівня зменшується приблизно у два рази. Якщо визначається найближчий із порогом – то маємо вигреш менше ніж у 2 рази, зате зберігається вказана межа дистанції між новоствореними гранулами, що підтримує ступінь розрізненості для сформованих даних. Ще більш простий варіант аналізу – пошук із заданим порогом еквівалентності до першого ідентичного елементу.

Грануляція із відсіюванням за найменшою відстанню.

Обчислимо матрицю відстаней між усіма дескрипторами опису E_k [62]. Застосуємо метод відбору починаючи з найменшої відстані. Черговий дескриптор поповнює список E_k^* , а відповідні рядок і стовпець цього дескриптора виключаються зі складу матриці. Потім переходимо до аналізу дескриптора з наступним за величиною більшим значенням відстані. Таким чином, ми послідовно набираємо до множини E_k^* потрібну фіксовану кількість дескрипторів. Цей метод дещо складніший за попередній, оскільки він може включати сортування за відстанню. Затє він не залежить від порядку слідування дескрипторів і не потребує адаптації порогу δ_ρ до змісту даних.

У результаті ознаки нижчого рівня за розглянутими способами аналізу об'єднуються в комплекси і розглядаються далі як одна ознака. У порівнянні із кластеризацією така процедура грануляції поетапно створює модифікований простір даних, що управляється параметром еквівалентності.

Як альтернативу обговорюваним способам грануляції розглянемо скорочення опису через відкидання елементів за фіксованим кроком. Наприклад, застосуємо метод вибіркового виключення, де кожен другий дескриптор видаляється із множини незалежно від своїх характеристик.

Зрозуміло, що цей спосіб має ряд недоліків. Це може бути втрата важливих дескрипторів, які є критичними для класифікації.

Відкидання елементів відбувається механічно, без урахування їх значення або ролі у загальній структурі множини.

Після відкидання можуть залишатися або лише важливі, або лише неважливі дескриптори. Для складних задач, де потрібна висока точність, цей метод не можна вважати ефективним. Він не забезпечує глибокого аналізу даних і може призвести до некоректних результатів.

5.3 Варіанти кластерного подання

Кластеризацію відносять до апарату квантування даних з визначенням системи центрів [42, 50, 89]. Розглянемо специфіку застосування апарату кластеризації на множині дескрипторів.

Здійснимо кластеризацію у базі E еталонних описів.

Будемо розглядати два альтернативних підходи до кластеризації з однаковим числом M кластерів – сукупно для усієї множини E (система центрів W^1) і роздільно для кожного із еталонів E_i , $i=1, \dots, N$ (система центрів W^2 , m центрів для кожного еталону) [42]. У результаті маємо два варіанти центрів

$$W^1 = \{w_j^1\}_{j=1}^M, W^2 = \{w_j^2\}_{j=1}^M. \quad (5.5)$$

Рівноцінне число M кластерів у (5.5) обрано нами виключно задля можливості об'єктивного порівняння результативності роздільного і спільного кластерного подання для еталонних описів. Фактично набір центрів (5.5) є апроксимацією для множини E і створює новий простір класифікаційних ознак.

При цьому для отримання векторного подання для кожного еталону за побудованою системою центрів у першому варіанті (спільна кластеризація для E) треба на підготовчому етапі виконати конкурентний аналіз опису для кожного із еталонів E_k . Для цього для елементів $z \in E_k$ визначимо номер v найближчого центру у множині W^1 , $w_j \in W^1$

$$v = \arg \min_{j=1, \dots, M} \rho(w_j, z). \quad (5.6)$$

Схема (5.6) є втіленням моделі пошуку найближчого кластерного центру.

У результаті застосування (5.6) отримуємо подання для еталону E_k у формі цілочислового вектору $gE_k = [g_1, g_2, \dots, g_M]$ з кількістю компонентів, що дорівнює числу кластерів. Тут g_d – число дескрипторів опису E_k , що потрапили за моделлю (5.6) у кластер з номером d . Вектор g є розподілом даних опису за набором кластерів. У процесі класифікації для аналізованого зображення з описом Z аналогічно також треба отримати вектор $g(Z)$ кластерного подання.

Для системи W^2 вже маємо закріплення фіксованого рівноцінного числа m центрів кластерів за кожним із еталонів, причому $M = mN$, так як кластеризація еталонів виконувалась роздільно. Тому додаткових дій з аналізу чи трансформації еталонного представлення даних не потрібно. Загалом число центрів m є параметром, що залежить від складу даних.

5.4 Схеми і критерії класифікації

Технологію R структурної класифікації здійснимо у два етапи за моделлю $R = R_2 R_1$. Етап R_1 реалізує локальне рішення стосовно класу дескрипторів об'єкту, а етап R_2 на підставі рішень (голосів) компонентів усього складу опису отримує значення класу для аналізованого об'єкту [42, 62].

Класифікатор R_1 реалізує по-елементний аналіз вхідного опису $Z = \{z_v\}_{v=1}^s$ об'єкта і відносить дескриптор $z_v \in Z$ до одного із класів за правилом

$$R_1 : z_v \rightarrow \{1, \dots, N\}. \quad (5.7)$$

Клас k аналізованого дескриптора z_v об'єкту у моделі (5.7) визначимо як аргумент мінімуму для відстані ρ на множині E дескрипторів усіх класів

$$R_1 : k = \arg \min_{i=1, \dots, N; d=1, \dots, s} \rho(z_v, e_d(i)), \quad (5.8)$$

де $\rho : B^n \times B^n \rightarrow [0, \infty]$ – відстань у векторному просторі B^n .

У моделі (5.8) доцільно використовувати відстань Хеммінга, що визначає число відмінних бітів у двох бінарних векторах [42, 62].

У виразі (5.8) мінімізується значення відстані за змінною i (номер класу) на множині дескрипторів бази E . У формулі (5.8) може бути організовано визначення мінімуму як за повним списком $s \cdot N$ дескрипторів бази, так і роздільно для кожного із N еталонів (з встановленим лімітом на величину мінімуму).

У роздільному способі аналізу для кожного із класів підраховується число дескрипторів об'єкту, що знайшли відповідність до фіксованого еталону. У такій моделі маємо множинне врахування голосів, коли один і той же дескриптор об'єкту може знайти відповідність для еталонів різних класів.

Таким чином, модель (5.8) втілює багатокласову класифікацію для окремого дескриптора у векторному просторі B^n на підставі найкращого узгодження на множині E еталонів.

Для втілення R_2 введемо вектор $\{h_i\}_{i=1}^N$ з цілими значеннями для накопиченої кількості голосів класів, отриманих застосуванням (5.8) до об'єкту Z .

На підставі R_1 для кожного $z_v \in Z$ відповідно до (5.8) визначимо номер класу k , а потім інкрементуємо акумулятор голосів $h_k = h_k + 1$ для відповідного номеру класу. Отриманий вектор $\{h_i\}_{i=1}^N$ є розподілом (гістограмою) числа голосів для множини Z у системі класів.

За результатом оброблення опису Z об'єкту обчислюємо вектор $\{h_i\}_{i=1}^N$. Клас Z визначимо правилом R_2 як аргумент максимуму

$$R_2 : Z \rightarrow E_k \mid \left(k = \arg \max_{i=1, \dots, N} h_i \right) \& (h_k \geq \delta_h), \quad (5.9)$$

де δ_h – деякий поріг для мінімально допустимого числа голосів. Якщо умова $(h_k \geq \delta_h)$ не виконується, клас об'єкту вважається не встановленим. Це відмова від класифікації, коли немає підстави віднести об'єкт до жодного із класів.

Значення δ_h визначають експериментально для бази E . Як правило, δ_h обирають як мінімальне число голосів (з допуском), необхідне для впевненої класифікації тестової вибірки на основі еталонів.

Із загальних теоретичних позицій δ_h є компромісом у задачі розрізнення еталонів від решти зображень, що можуть поступити на вхід системи класифікації [42, 68].

Послідовність правил R_1, R_2 реалізує класифікатор на підґрунті рішень, отриманих для множини складових компонентів об'єкту. Він впроваджує принцип структурного аналізу і забезпечує стійкість до просторових викривлень дескрипторів із-за можливого впливу завад.

Традиційний метод класифікації реалізує моделі (5.8) та (5.9) на повному наборі дескрипторів бази. Процес класифікації з використанням гранульованих описів модифікується здійсненням обмеженої процедури пошуку на скороченій базі $E^* = \cup E_k^*$. Треба врахувати той факт, що опис об'єкту залишається повним, а описи еталонів – скорочуються [42]. Будемо при цьому вважати, що повноцінні еталони E_k і трансформовані E_k^* мають приблизно рівноцінні значення потужності для різних класів. У цьому варіанті s елементів об'єкту розподіляють свої голоси за системою N класів, причому в результаті маємо $\sum_{k=1}^N h_k \leq s$.

Значення δ_h визначають експериментально для заданої бази E еталонів. Як правило, його обирають як мінімальне число голосів h_m (з допуском), необхідне для впевненої класифікації тестової вибірки на основі еталонів. Із загальних теоретичних позицій δ_h є компромісом у задачі відділення дескрипторів еталонів від дескрипторів неосязної решти інших зображень, що можуть поступити на вхід системи класифікації [42].

При практичному застосуванні ми рекомендуємо обирати поріг за моделлю $\delta_h = 0,95h_m$, що відповідає допустимому відхиленню в сторону зменшення від мінімуму h_m у 5%.

Класифікатор з використанням спільної кластеризації на основі системи центрів W^1 спочатку отримує за моделлю (5.6) векторне подання $g(Z) = [g_1, g_2, \dots, g_M]$ для розпізнаваного об'єкту, а далі з використанням метрики μ для векторів цілих чисел визначає клас об'єкту як аргумент мінімуму метрики μ на множині кластерних подань для еталонів бази E

$$R_2 : Z \rightarrow E_v \mid v = \arg \min_{k=1, \dots, N} \mu(g(Z), g(E_k)). \quad (5.10)$$

Варіантом метрики μ в (5.10) може бути манхеттенська відстань для простору векторів цілих чисел.

Використання правила (5.10) замість (5.8) та (5.9) для традиційного методу дає вигреш у обсязі обчислень приблизно пропорційний величині $\gamma = s/M$.

Відмітимо, що для відсіювання зображень поза базою оптимальне значення метрики μ у (5.10) повинне задовольняти умові значущості $\mu_v \geq \delta_\mu$.

Класифікатор на підставі роздільної кластеризації на етапі R_1 втілює модель (5.6) на множині центрів $w_j^2 \in W^2$. Клас дескриптора при цьому однозначно визначається визначенням центром кластеру, а клас аналізованого об'єкту оцінюється за моделлю (5.9). Виграш у обсязі обчислень у порівнянні із традиційним методом теж пропорційний значенню γ .

При дії локальних завад класифікація з використанням розглянутих методів фактично здійснюється на підставі невикривленої частини опису аналізованого об'єкту. У цьому випадку число голосів природно знижується, тому параметри δ_ρ, δ_h треба адаптувати до допустимого рівня перешкод.

Схему класифікації за скороченим описом представлено на рисунку 5.1.

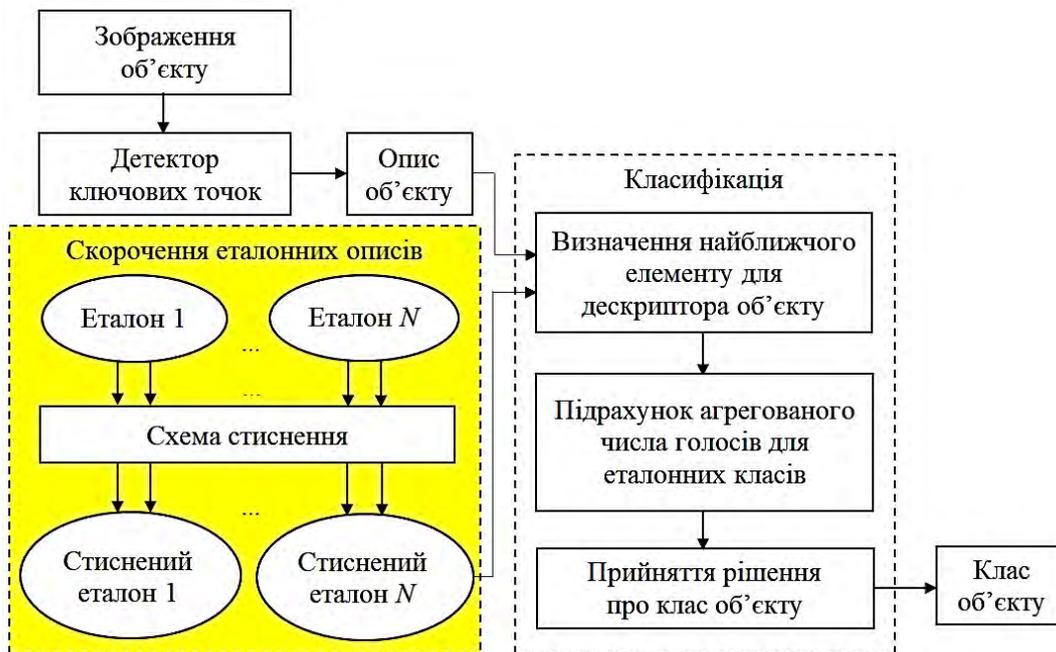


Рисунок 5.1 – Схеми класифікації на підставі скорочення опису еталонів

Зважаючи на те, що при роздільній кластеризації апроксимація здійснюється шляхом формування центрів для еталонних даних, критерієм, що може прогнозувати результативність класифікації, є середня відстань між системами центрів для пар еталонів. Якщо позначити вектори для центрів пари еталонів $A = \{a_i\}_{i=1}^m$ та $B = \{b_j\}_{j=1}^m$, то такий критерій β має вигляд

$$\beta(A, B) = \left[\sum_{i=1}^m \sum_{j=1}^m \rho(a_i, b_j) \right] / m^2. \quad (5.11)$$

Значення (5.11) знаходиться у межах визначення метрики ρ .

Для бінарних векторів розмірністю n і метрики Хемінга показник (5.11) належить $[0, n]$. Чим більше значення β , тим маємо більший розкид між системами сформованих центрів і тим вище ступінь розрізнення класів A, B між собою у просторі кластерних центрів.

5.5 План експериментів та критерії результативності

У порівняльному плані в експериментальній частині проведеного дослідження нами вивчались показники результативності (точність, швидкодія) для наступних типових схем скорочення опису:

1. Класифікація за повними описами без скорочення (традиційна схема).
2. Здійснення грануляції на 5-му рівні з порогом.
3. Здійснення грануляції на 5-му рівні без використання порогу.
4. Формування детектором Oriented FAST and Rotated BRIEF (ORB) [44] фіксованого числа – 15 дескрипторів для кожного еталонного зображення.
5. Просіювання шляхом відбору кожного 33 дескриптора у описі еталону.
6. Просіювання п'ять разів відкиданням через один до досягнення 15-ти дескрипторів у кожному описі.
7. Вибір 15 перших із 500 дескрипторів для кожного еталонного опису.

Параметри у схемах 2–7 обрано таким чином, щоб отримати приблизно рівноцінне число – 15 дескрипторів у скороченому описі.

Результативність класифікації будемо оцінювати критерієм точності pr , який відображає відношення кількості u тестових зображень з правильною оцінкою класу (чи визначення як зображення поза базою еталонів) до загального числа U експериментів [21]

$$pr = u / U . \quad (5.12)$$

Ще одним важливим критерієм є достовірність прийняття класифікаційного рішення за набраним класами числом голосів. Достовірність β характеризує ступінь впевненості щодо прийняття рішення і оцінюється критерієм у формі відношення локального (другого за величиною) максимуму h_{m2} серед голосів $\{h_i\}_{i=1}^N$ до основного максимуму h_{m1} :

$$\beta = h_{m2} / h_{m1}. \quad (5.13)$$

Чим менше значення (5.13), тим вища достовірність класифікаційного рішення.

5.6 Експериментальні результати і обговорення

Основною метою проведення наших експериментів було підтвердження продуктивності і результативності запропонованих способів грануляції (включно з кластеруванням) у задачі класифікації в порівняльному аспекті із традиційним (повний опис) та спрощеним (після відкидання) методами.

Експерименти з грануляцією опису.

Для проведення досліджень використано бібліотеку OpenCV [42]. Застосовано детектор ORB [7] для виявлення та опису ключових точок в зображеннях, визначено множини дескрипторів ключових точок для тестового набору зображень. Повні описи еталонних і тестових зображень включають 500 дескрипторів ORB з можливим невеликим відхиленням.

Тестовий набір у експерименті склав 72 зображення і включав:

– п'ять еталонних зображень логотипів футбольних клубів (база класифікації, алфавіт класів, рис. 5.2);



Рисунок 5.2 – Зображення гербів із бази еталонів

– три зображення поза базою (рис. 5.3);



Рисунок 5.3 – Зображення поза базою

– зображення, отримані із цих восьми зображень застосуванням геометричних перетворень масштабу (10% зменшення чи збільшення), повороту (30 градусів вліво або вправо), комбінацією масштабу та повороту (рис. 5.4).



Рисунок 5.4 – Приклади тестових зображень з координатами ключових точок

Як бачимо, обрані перше і третє зображення поза базою (рис. 5.3) мають суттєву візуальну схожість із деякими еталонами бази. На рисунку 5.4 наведено приклади тестових зображень з координатами ключових точок та під дією геометричних перетворень.

Помилка класифікації фіксувалася у випадку віднесення трансформованого еталонного зображення до «чужого» класу та при віднесенні трансформованого зображення поза базою до одного із наявних класів.

Для виконання грануляції за порогом еквівалентності встановлено поріг $\delta_p = 37,5\%$ від максимуму метрики (256) на кожному рівні. Грануляція виконувалась у декілька етапів, за результатом моделювання вирішено зупинитись на п'ятому рівні ієрархії, де рівень точності починає суттєво знижуватися. Використовуючи аналіз із порогом, досягнуто скорочення числа елементів менше ніж у два рази на кожному з етапів.

У таблиці 5.1 наведено дані, що демонструють кількість дескрипторів бази, що залишилися після виконання грануляції на кожному із рівнів.

Таблиця 5.1 також містить число голосів за правильний клас та оцінку часу виконання класифікації одного еталонного зображення на різних рівнях грануляції. Результати таблиці 5.1 отримані для описів наявної бази еталонів (рис. 5.1).

Як можна помітити із таблиці 5.1, із зростанням рівня грануляції від I до V у порівнянні з повним описом зменшується кількість дескрипторів у скорочених описах еталонів від 2492 до 376 і пропорційно зменшується час класифікації.

Таблиця 5.1 – Параметри класифікації з використанням грануляції

Рівень грануляції	Час, с	Число дескрипторів бази	Число голосів за клас
Повний опис	29,45	2492	500, 500, 495, 497, 500
I	15,19	1279	411, 409, 435, 451, 441
II	10,05	830	304, 331, 404, 417, 368
III	7,27	586	242, 267, 390, 373, 306
IV	5,55	449	212, 237, 368, 318, 256
V	4,48	376	213, 217, 343, 267, 227

Навіть при значному скороченні числа інформативних елементів на V-му рівні грануляції (у шість з половиною разів у порівнянні з повним описом!) отримана в експерименті точність класифікації залишається на найвищому рівні $pr=1$, а максимум голосів у всіх експериментах відповідає «правильному» класу.

Конкретно значення максимумів голосів $\{h_i\}_{i=1}^N$ для одного із варіантів скорочення описів склали 175, 148, 164, 172, 192, значить, за цими даними поріг δ_h дорівнює $\delta_h = 0,95 \times 148 = 141$ голос. Тепер рішення за встановлення класу приймається у випадку, якщо число набраних голосів 141 і вище. При цьому для зображень поза базою число голосів повинно бути меншим 141.

На відміну від грануляції за порогом еквівалентності, підхід із використанням відкидання елементів з фіксованим кроком завжди здійснює скорочення кількості елементів у відповідне число разів. Конкретно число дескрипторів при просіюванні через один на рівнях ієрархії дискретно зменшувалось так: 2492, 1245, 622, 309, 154, 75.

Важливим моментом є вивчення впливу геометричних перетворень на результативність класифікації. Проведення моделювання на тестовому наборі із 72 зображень з втіленням геометричних трансформацій показало наступне.

Для експериментального оцінювання обрано результати грануляції № 2 та № 3 на 5-му рівні, де кількість дескрипторів у описі кожного еталону знизилось до 15, тобто у 33 рази зменшилось у порівнянні з обсягом повного опису. Альтернативами при тих же параметрах виступали схеми відсіювання № 4–№ 7 і традиційний спосіб № 1 з повними описами еталонів.

Отримані значення точності і достовірності містить таблиця 5.2.

Таблиця 5.2 – Показники точності та достовірності для різних схем скорочення

Схема	1	2	3	4	5	6	7
Точність, pr	1,0	0,75	0,63	0,63	0,95	0,85	0,5
Достовірність, β	0,20	0,48	0,60	0,57	0,63	0,62	0,72

Аналізуючи дані таблиці 5.2, бачимо, що традиційний метод (схема № 1) на основі повного опису має найкращі показники точності (найбільше значення) і достовірності (найменше). Зате решта способів (схеми №№ 2–7), що націлені на пришвидшення класифікації, потребують приблизно у 33 рази! менше часу на оброблення, мають відповідно і дещо гірші показники.

Грануляція з порогом (схема № 2) показала дещо кращі результати за обома критеріями у порівнянні із грануляцією без порогу (схема № 3). Це можна пояснити впровадженням додаткового аналізу на значення мінімуму відстані.

Схема № 7, що використовує перші 15 із сформованих 500 дескрипторів із еталонних описів, має найгірші показники і точності і достовірності. Схеми з відсіюванням № 5 та № 6 показали досить високий рівень точності, але з низьким коефіцієнтом достовірності. Цей факт підкреслює, що ефективність процедур відкидання (схеми №№ 4–6) загалом суттєво залежить від послідовності формування дескрипторів у описі і від самого принципу відкидання. Це може потребувати додаткового багатократного оброблення задля досягнення необхідних показників ефективності.

Схема № 4 скорочення повністю базується на властивостях дескриптора ORB, її показники не гарантують високої результативності класифікації.

Експерименти з кластеруванням опису.

У порівняльному аспекті ми провели програмне моделювання класифікатора з використанням двох розглянутих способів кластеризації – для усієї бази і роздільно для кожного із еталонів. Використано базу із п'яти зображень картин української художниці М. Приймаченко [92], а також три її інші картини поза базою класифікації. На рисунку 5.5 наведено картини із використаної бази і координати ключових точок дескрипторів, отриманих детектором Binary Robust Invariant Scalable Keypoints (BRISK) (вектор із 512 бітів) [6]. Як бачимо, художниця використовує оригінальну техніку для своїх картин, що в той же час робить їх візуально схожими між собою.



Рисунок 5.5 – Зображення із бази картин і координати ключових точок

Для кластеризації використано $M = 15$ кластерів для бази, і $m = 3$ для кожного із еталонів. Умови проведення моделювання при дії геометричних перетворень відповідали розглянутим вище способам.

Результати експериментів засвідчили таке. Традиційний метод для зображень картин показав також найвищу точність $pr = 1$, успішно класифікуючи усі зображення із бази і відділивши зображення не із бази. Модифікований метод на основі кластерних центрів для повної бази показав точність $pr = 0,82$, а з роздільною кластеризацією – $pr = 0,90$. Як бачимо, роздільна кластеризація дає дещо вищий рівень точності.

При збільшенні числа кластерів до $m = 5$ для кожного еталону (загалом для бази – 25 центрів) точність методу з кластеризацією повної бази природно зросла до рівня $pr = 0,88$, а з роздільною кластеризацією – до $pr = 0,98$.

Відмітимо при цьому, що значення параметру $m = 5$ у методах з кластеризацією означає використання всього 5 центрів даних для кожного із еталонів, в той час як дані таблиці 5.2 (схеми №№ 2–7 скорочення) отримані для випадку, коли кожний еталон представлений 15-ма дескрипторами.

Таким чином, методи з кластеризацією досягають значно кращої точності (0,88 та 0,98) при менших у три рази обчислювальних затратах.

Заміри часу для класифікації одного зображення у експерименті з кластеризацією показали вигреш у порівнянні з традиційним методом приблизно у 158 разів для варіанту $m = 3$, що майже співпадає з теоретичним розрахунком. При цьому час класифікації із спільною і роздільною кластеризацією приблизно рівноцінний.

Як бачимо, введення кластерування даних у процес класифікації суттєво скоротило обчислювальні витрати за рахунок незначного зниження показника точності.

Зауважимо, що усі отримані експериментально показники точності безпосередньо залежать від обраної бази зображень, а також від набору альтернативних зображень не із бази, і можуть варіюватися. Ключовим фактором є візуальна схожість аналізованих зображень між собою, що безпосередньо відображається і у близькості множин дескрипторів. У той же час значний вигравш у часі оброблення і в показнику точності при застосуванні кластерування у порівнянні з іншими схемами оброблення є значущим і об'єктивним фактором.

Традиційний метод структурної класифікації базується на процедурі голосування і фактично реалізує метод ближнього сусіда на множині дескрипторів для бази еталонних описів. Грануляція і відкидання ознак для еталонних описів успішно здійснюють скорочення обсягу пошуку, що загалом підвищує швидкодію класифікації пропорційно ступеню скорочення.

Незважаючи на те, що вибір найбільш дієвого способу стиснення опису у методах класифікації суттєво залежить від визначеного набору еталонів і допустимого різноманіття вхідних зображень, впровадження інтелектуальних методів кластеризації має значущі переваги над простими схемами скорочення обсягів даних. Зважаючи на те, що стиснення еталонної інформації здійснюється однократно на етапі попереднього оброблення і тому не впливає на збільшення часу класифікації, інтелектуальні підходи за рахунок більш глибокого аналізу даних не тільки спрощують обчислення, але і забезпечують достойний рівень точності. Крім того, при цьому формуються інтегровані центри даних, які у багатьох випадках має сенс впровадити при класифікації.

Результативність же впровадження спрощених підходів має випадковий характер, коливається у широких межах і залежить не тільки від порядку слідування дескрипторів у описі, а і від впровадженої схеми оброблення, що може потребувати багатократної перевірки чи підбору різних варіантів відкидання.

Застосування апарату кластерування має значно вищий потенціал в аспекті кращих показників точності та швидкодії для процесу класифікації у порівнянні з простими схемами відсіювання чи грануляції шляхом визначення близьких компонентів опису.

Матеріали дослідження опубліковано у статті [92].

6 ШВИДКІСНИЙ МЕТРИЧНИЙ ПОШУК ДАНИХ НА ОСНОВІ ХЕШУВАННЯ

6.1 Хеш-структури для подання опису

Структурні методи класифікації зображень, що впроваджуються у сучасних системах комп'ютерного зору, ґрунтуються на даних про образи візуальних об'єктів у вигляді множини дескрипторів ключових точок [4, 35, 44, 86, 98]. Найбільш ефективні такі методи для розпізнавання зображень фіксованої структури [35, 86]. Дескриптор ключових точок – це вектор розміром 64...512 бінарних компонентів, що є апроксимацією для фрагмента функції яскравості зображення [44, 99]. Традиційний класифікатор зображень базується на метричних критеріях релевантності виду множина-множина між образами розпізнаваного об'єкту і еталону та оптимізації значення релевантності у базі еталонів [86].

Клас розпізнаваного візуального об'єкта визначається як нескінченна множина зображень об'єкту з урахуванням його допустимих геометричних перетворень: зсувів, поворотів, масштабування в межах поля зору системи розпізнавання [4, 35, 44, 86]. Представником класу виступає еталонне зображення, за яким комп'ютерна система визначає образ (множину ознак) еталона. Множина обраних користувачем еталонних образів створює скінченну базу зразків, на підставі якої здійснюється класифікація об'єктів. Вона полягає у віднесенні вхідного зображення до одного із еталонних класів або у відмові від класифікації [4, 86].

Більш універсальним у аспекті забезпечення необхідної завадостійкості вважається впровадження по-елементного аналізу для множини компонентів опису, яке на основі формування статистичних рішень для цілісного образу об'єкта дає можливість у процесі класифікації знизити вплив завад [86].

Для способу по-компонентного аналізу традиційний підхід структурної класифікації зводиться до визначення для кожного дескриптора із опису розпізнаваного об'єкта деякої підмножини дескрипторів бази еталонів, релевантних йому у комплектній множині, та підрахунку числа дескрипторів – компонентів об'єкту, віднесених відповідно до кожного із класів. Така процедура реалізує процес голосування. Ступінь релевантності визначається на підставі обчислення значення метрики.

Зважаючи на те, що число дескрипторів у описі окремого еталону досягає 300...500, а класів може бути десятки і навіть сотні, такий підхід призводить до процедури повного перебору та потребує громіздких обчислень [35, 44, 98].

Впровадження засобів хешування даних суттєво скорочує часові затрати у процесі практичної реалізації процедур класифікації на підставі метричного пошуку [4, 25]. При хешуванні дані бази еталонів попередньо розкладають на групи подібних за деякою ознакою елементів, що значно (у тисячі разів) прискорює процес оброблення за рахунок відповідного зростання вимог до об'ємів пам'яті [19, 98, 100, 101]. Хешування реалізує перспективну ідею зниження розмірності аналізованих даних шляхом попередньої побудови для еталонної множини деякої індексованої структури, що спрощує оброблення та знижує обчислювальні затрати на класифікацію.

У статтях [27] досліджено наближені методи метричного пошуку за зразком на підставі створення індексованих структур даних на множині ієрархічних ознак сегментованих зображень. Індексована структура на підґрунті кластеризації для об'ємних масивів багатовимірних даних впроваджена у професійних системах пошуку [31], де задля покращення продуктивності застосована багатоетапна схема кластеризації з моделлю наближеного пошуку релевантних елементів всередині кластеру.

Створення кластерних структур як сучасний підхід інтелектуального аналізу даних дає можливість краще адаптуватися до змісту еталонної інформації, що застосовано нами при визначенні релевантності візуальних об'єктів у технологіях «мішок слів» [75, 86, 101, 102]. Алгоритмічні та програмно-апаратні засоби швидкого пошуку у об'ємних сховищах даних (наприклад, у вебзастосунках) успішно розвиваються і знаходять впровадження у сучасних прикладних пошукових системах [19, 52, 72, 75, 100, 103].

У роботі [19] детально вивчається спектр засобів швидкісного пошуку, серед яких можна виділити застосування принципу комбінаційного поєднання декількох методів у вигляді ланцюжка: хешування, розбиття даних на блоки (проєктування), LSH-оброблення (locality-sensitive hashing, тобто хешування із врахуванням близькості) у вигляді сортування за значенням хеш-функції для блоків даних, логічний аналіз (узгоджена фільтрація) значень хеш-функції для блоків. Ключова ідея LSH-оброблення зосереджує увагу тільки на парах даних, які можуть бути схожими, щоб не перевіряти кожну із пар [19, 25].

Впровадження будь-якого нового засобу аналізу чи оброблення даних вимагає від дослідників уваги в аспекті забезпечення якісних показників для функціоналу класифікатора [18, 84].

Зрозуміло, що для систем комп'ютерного зору ключовим критерієм при впровадженні засобів швидкого пошуку є все-таки гарантоване забезпечення достатніх показників результативності класифікації, таких як ймовірність правильного розпізнавання.

Зважаючи на нагальні вимоги щодо забезпечення високої обчислювальної продуктивності структурних методів класифікації зображень зі збереженням достатнього рівня їх результативності необхідно впровадити в їх структуру сучасні підходи швидкісного пошуку та аналізу даних, що застосовуються у технологіях оброблення об'ємних масивів багатовимірної інформації.

Мета роботи – підвищення обчислювальної продуктивності структурних методів класифікації зображень.

Завдання дослідження полягають у створенні та застосуванні моделей швидкісного класифікатора зображень на основі побудови хеш-структур даних для еталонних описів класів у вигляді множини дескрипторів ключових точок, впровадженні засобів розбиття даних на непересічні блоки, використанні апарату LSH- та логічного оброблення. За результатом програмного моделювання необхідно здійснити оцінювання результативності та досягнутих показників швидкодії для розроблених модифікацій класифікаторів у порівнянні з відомими.

6.2 Формалізм задачі класифікації

Опрацюємо простір B^n багатовимірних бінарних векторів розмірністю n , у якому будемо конструювати образи розпізнаваного об'єкту і еталонів. Зафіксуємо деяку мультимножину векторів $E_i \subseteq B^n$ як опис еталону $E_i = \{e_v(i)\}_{v=1}^s$ у просторі дескрипторів ключових точок, $s = \text{card } E_i$ – потужність множини дескрипторів [4, 86]. Компоненти опису – це вектори $e_k \in B^n$, скінченна множина яких створює опис об'єкту.

Постановка задачі класифікації передбачає наявність деякої бази E описів еталонів розмірністю $N: E = \bigcup_{i=1}^N E_i$. Кожний еталонний опис E_i репрезентує окремий клас, представлений фіксованою скінченною множиною дескрипторів ключових точок.

Подамо опис розпізнаваного об'єкту як множину $Z \subset B^n$, $Z = \{z_w\}_{w=1}^s$.

Поставимо задачу побудови класифікатора K у сенсі відображення $K : Z \rightarrow [1, 2, \dots, N]$ із множини компонентів опису у множину номерів класів на основі попереднього конструювання індексованої структури на множині E .

Класифікацію K за множиною компонентів опису представимо як двоетапний процес

$$K = K_2 K_1, \quad K_1 : B^n \rightarrow [1, 2, \dots, N], \quad K_2 : D \rightarrow [1, 2, \dots, N], \quad (6.1)$$

де на першому кроці $K_1 : B^n \rightarrow [1, 2, \dots, N]$ здійснюємо визначення класу d_w для кожного дескриптора $z_w \in Z$ об'єкта, а на другому кроці $K_2 : D \rightarrow [1, 2, \dots, N]$ із множини $D = \{d_w\}_{w=1}^s$ отриманих рішень компонентів першого кроку формуємо агрегований висновок про клас об'єкту Z . Такий принцип оброблення відповідає узагальненню множини рішень однорідних часткових класифікаторів за моделлю бустінгу [4, 52, 86, 101]. Етап K_1 тут можна розглядати як багатозначну характеристичну функцію для визначення еталонного класу E_i щодо окремого дескриптора із опису об'єкта, а етап K_2 втілює кінцеве класифікаційне рішення.

Етапи K_1, K_2 у класифікаторі (6.1) можна побудувати і в інших варіантах, наприклад, шляхом створення на етапі K_1 деякого ймовірнісного розподілу за класами еталонів, включаючи процедури логічного аналізу чи оброблення таких розподілів [4, 86].

Конструктивно та інформативно реалізація K спирається на апіорні дані наявної бази E , так як належність усіх $e_v(i)$ до відповідного образу E_i всередині бази уже відома на початку класифікації.

Якщо класифікацію K_1 здійснювати традиційним методом лінійного метричного пошуку (повний перебір) шляхом послідовного аналізу кожного із елементів набору E , то застосовують конкурентне правило [86, 101, 102]

$$d_w = \arg \min_{i,v} \rho(z_w, e_v(i)), \quad (6.2)$$

де d_w – номер еталону E_i , до якого буде віднесено дескриптор z_w об'єкту, $d_w \in \{1, \dots, N\}$;

$\rho(z_w, e_v(i))$ – метрика у векторному просторі.

Важливим задля результативної класифікації є впровадження логічного аналізу для значення мінімуму, що досягнуто при здійсненні (6.2) [101].

Значення d_w визначається тільки у випадку, якщо отриманий мінімум відстані у (6.2) не перевищує заданий поріг $\delta_\rho : \rho_{\min} \leq \delta_\rho$. У протилежному випадку клас d_w не визначається, а аналізований дескриптор вважається хибним.

Необхідну кількість Q обчислених значень метрики в (6.1) у базі еталонів лінійним пошуком можна оцінити значенням параметру $Q = N \cdot s^2$, вважаючи обсяги описів еталонів та об'єкта рівноцінними [86]. Для векторів простору B^n в (6.2) може бути застосована проста в обчислювальному сенсі метрика Хемінга χ , що підраховує число не співпадаючих бітів для векторів

$$\chi(z_w, e_v(i)) = \sum_{j=1}^n 1[z_{wj}, e_{vj}(i)], \quad (6.3)$$

у (6.3) логічна функція $1[\dots]$ приймає значення 1, якщо j -ті біти векторів z_w , $e_{vj}(i)$ не співпадають, і значення 0 – у протилежному випадку.

Суть етапу K_2 полягає у тому, що на кожному кроці за правилом (6.2) відповідно до отриманого значення d_w інкрементується і накопичується число r_i голосів елементів, віднесених до i -го класу

$$r_i = \begin{cases} r_i + 1, & d_w = i, \\ r_i, & d_w \neq i, \end{cases} \quad (6.4)$$

а результуючий клас i_0 для цілісного образу Z об'єкта визначаємо за максимумом числа голосів серед усіх класів:

$$i_0 = \arg \max_{i=1, \dots, N} r_i. \quad (6.5)$$

Значення $\{r_i\}_{i=1}^N$ відображають гістограму класів за числом голосів елементів із множини Z . Вирази (6.4) та (6.5) конкретизують етап K_2 , що полягає у обробленні отриманих голосів для компонентів опису Z . При впровадженні етапу K_2 теж допустиме використання логічного аналізу для значення максимального числа голосів $r_i(i_0)$ у виразі (6.5).

Позитивне рішення про клас об'єкту приймаємо за умови $r_i(i_0) \geq \delta_r$, де δ_r – поріг для вирішального числа голосів. Пороги δ_p , δ_r визначають, як правило, умовами класифікації [86].

Запропонована двоетапна процедура класифікації базується на основоположному принципі інтелектуального аналізу, що полягає у підрахунку числа позитивних рішень (підтримка, рейтинг зустрічальності) на аналізованій множині даних [1, 52].

Розглянуті варіанти побудови класифікатора природно можна трактувати в рамках теорії ансамблевих моделей [52, 86], де за рахунок створення та агрегування відгуків компонентних класифікаторів (локальних рішень) синтезується «сильний» класифікатор із гарантовано вищою результативністю прийняття рішень, що відповідає моделі бустінгу [52, 102].

Ключовою обчислювальною проблемою при побудові класифікатора для об'ємних баз даних є реалізація пошуку (6.2), що фактично є підходом ближнього сусіда у багаточисельній множині векторів.

Зосередимося на її вирішенні шляхом застосування у двоетапній процедурі класифікації спеціалізованої структури даних, що націлена на зменшення обсягу обчислень Q .

6.3 Трансформація даних із впровадженням хеш-структури для швидкісного пошуку

На етапі попереднього аналізу даних створимо у еталонній множині E образів спеціалізовану структуру задля організації швидкісної результативної класифікації. Здійснимо розбиття T на множині E дескрипторів бази зображень, отримуємо множину із M непересічних груп $T_k(E)$:

$$E = T(E) = \bigcup_{k=1}^M T_k(E), \quad T_k(E) \cap T_j(E) = \emptyset. \quad (6.6)$$

На сьогодні існує два найбільш поширених прикладних способів розбиття багатовимірних даних: хешування та кластеризація [1, 19, 35, 98, 101]. Впровадження хешування пов'язують із точними засобами оброблення, а кластеризацію відносять до наближених способів самонавчання, так як її результат строго не формалізований.

Трансформація (6.6) практично здійснює попередню класифікацію і повністю зберігає усю сукупність аналізованих даних, яка тепер розподіляється між групами $T_k(E)$. У результаті перетворення (6.6) кожний дескриптор $e_v \in E$ бази еталонів отримує параметр k номеру групи (кошика чи кластеру).

Зважаючи на наявне початкове розбиття $E = \bigcup_{i=1}^N E_i$ даних на еталонні образи, визначимо параметр

$$t_{i,k} = \text{card}\{e_v \mid e_v \in E_i \ \& \ e_v \in T_k\} \quad (6.7)$$

як число дескрипторів i -го класу, що потрапили до кошика T_k . На підставі $t_{i,k}$ маємо вагову кількісну характеристику $b_i(T_k)$ для кожної корзини даних

$$b_i(T_k) = \frac{t_{i,k}}{\text{card } T_k}. \quad (6.8)$$

Вираз (6.8) визначає статистичний розподіл елементів кожного сегменту даних за класами еталонів у вигляді вагових коефіцієнтів класів, причому $\sum_{i=1}^N b_i = 1$. Розподіл b є спільною характеристикою для множини елементів сегменту даних, отриману за результатами аналізу (навчання) для бази E . На етапі навчання ефективним для підвищення результативності класифікації може бути запровадження логічного оброблення для значення вектору b задля використання вагомості впливу коефіцієнтів b_i для найбільш значущих класів у кошику [86].

Для хешування атрибутом окремого кошику даних виступає значення хеш-функції. Важливим параметром, що впливає на швидкодію пошуку, є число кошиків M . Чим менше число кошиків, тим вища швидкодія переходу до кошику. Чим більше число кошиків, тим менше об'єм даних для аналізу всередині кошиків. Граничними ситуаціями є один кошик ($M=1$) та повна відсутність хешування ($M=N \cdot s$), що відповідають лінійному пошуку.

За думкою дослідників, можна формулювати задачу оптимізації числа M кошиків, де критерієм виступає число Q обчислень метрики для компонентів опису [19, 27, 31, 52].

Якщо в одному кошику в середньому число елементів є $\frac{N \cdot s}{M}$, то кількість обчислень (6.2) всередині індексної структури з використанням значень хеш-коду пропорційна $Q_1 = s \cdot M + \frac{N \cdot s}{M}$, що значно менше, ніж $Q_2 = N \cdot s^2$ для традиційного лінійного пошуку. Для конкретних значень $s = 500$, $N = 10$, $M = 10$ вираш $\beta = \frac{Q_1}{Q_2}$ складає приблизно 450 разів і зростає зі збільшенням обсягів N, s даних.

Метод оброблення даних із використанням хешування [19, 98] фактично здійснює попередню класифікацію на окремі кошики та має свої особливості. Число кошиків для цього методу, як правило, задане і пов'язане із діапазоном значень хеш-ключа $h(e_v)$ для хешування, яке задає користувач.

Ключ – це, як правило, ціле число, що приймає скінченну множину із L фіксованих значень $h(e_v)$, $h(e_v) \in H_0$, $H_0 = \{h_i\}_{i=1}^L$ – множина значень хеш-функції [98, 100]. Для векторів із B^n таким ключем може бути, наприклад, число одиничних бітів у складі дескриптору, число вибраних комбінацій бітів довільної довжини та інші. Взагалі ознаку $h(e_v)$ для хешування можна корисно адаптувати до значень описів наявної бази еталонів.

Із впровадженням хешування розбиваємо множину даних E на непересічні групи $H_k(E)$ (кошики) числом L з рівнозначним значенням хеш-ключа

$$E = H(E) = \bigcup_{k=1}^L H_k(E), H_i(E) \cap H_j(E) = \emptyset, H_k(E) = \{e_v \in E \mid h(e_v) \in h_k\}. \quad (6.9)$$

Впровадимо тепер конкретний ланцюжок засобів попереднього оброблення даних задля забезпечення прискореного пошуку (рис. 6.1) [101]. Застосуємо блокове подання даних множини дескрипторів, обчисливши два значення хеш-функції $h(1), h(2)$ для кожного із дескрипторів бази еталонів.

Конкретно для дескриптора BRISK [37] розміром 512 бітів обчислимо значення $h(1), h(2)$ для першої та другої його половини за схемою «256+256».

Далі впровадимо LSH-оброблення [19]: відсортуємо еталонний масив $E = \{e_v\}_{v=1}^{s \cdot N}$ дескрипторів усієї бази за значенням хеш-функції $h(1)$ для 1-го блоку.

Шляхом логічного аналізу здійснимо оброблення діапазонів значень $h(1)$, виключивши із пошуку незначущі для наявної бази еталонних даних діапазони.



Рисунок 6.1 – Схема створення модифікованої структури даних

6.4 Класифікатор на основі синтезованої структури

Застосуємо детектор ключових точок до зображення розпізнаваного об'єкту, формуємо опис $Z = \{z_w\}_{w=1}^s$ об'єкту як множину дескрипторів. Побудуємо класифікатор з використанням швидкісного пошуку у базі E на підставі створеної структури.

Запропонований спосіб із впровадженням ланцюжка засобів включає такі етапи (рис. 6.2):

Етап 1. Для будь-якого дескриптора $z^* \in Z$ вхідного образу обчислюємо значення хеш-ключів $h^*(1), h^*(2)$ як значення хеш-функції для аргументу z^* .

Етап 2. При визначенні у базі дескриптора, ідентичного аналізованому $z^* \in Z$ зі значеннями $h^*(1), h^*(2)$, застосуємо лінійний пошук у «полосі даних» за значенням хеш-функції $h(1)$ для 1-го блоку, що обмежується заданою границею $h^*(1) \pm \Delta$. Результат визначається введеним параметром Δ , який можна назвати точністю пошуку. Крім того, через використання параметру Δ забезпечується ідентифікація варіативних даних, що незначно відрізняються від еталонних.

Етап 3. Логічне узгоджене оброблення. Здійснюємо зіставлення $\rho(z^*, e_v)$ за метрикою Хемінга (6.3) і пошук мінімуму $b = \arg \min_v \rho(z^*, e_v)$ у виразі (6.2) тільки у ситуації, якщо значення хеш-функції для 2-го блоку знаходиться у межах $h^*(2) \pm \Delta$.

Етап 4. За результатом пошуку на Етапі 3 для елемента z^* визначаємо клас d^* .

Етап 5. На підставі аналізу множини Z дескрипторів об'єкту шляхом виконання Етапів 1–3 для усіх її елементів накопичуємо лінійку $\{r_i\}_{i=1}^N$ значень голосів за кожний з наявних еталонних класів.

Етап 6. За виразом (6.5) класифікуємо розпізнаваний об'єкт до класу, що набрав найбільшу кількість голосів компонентів.

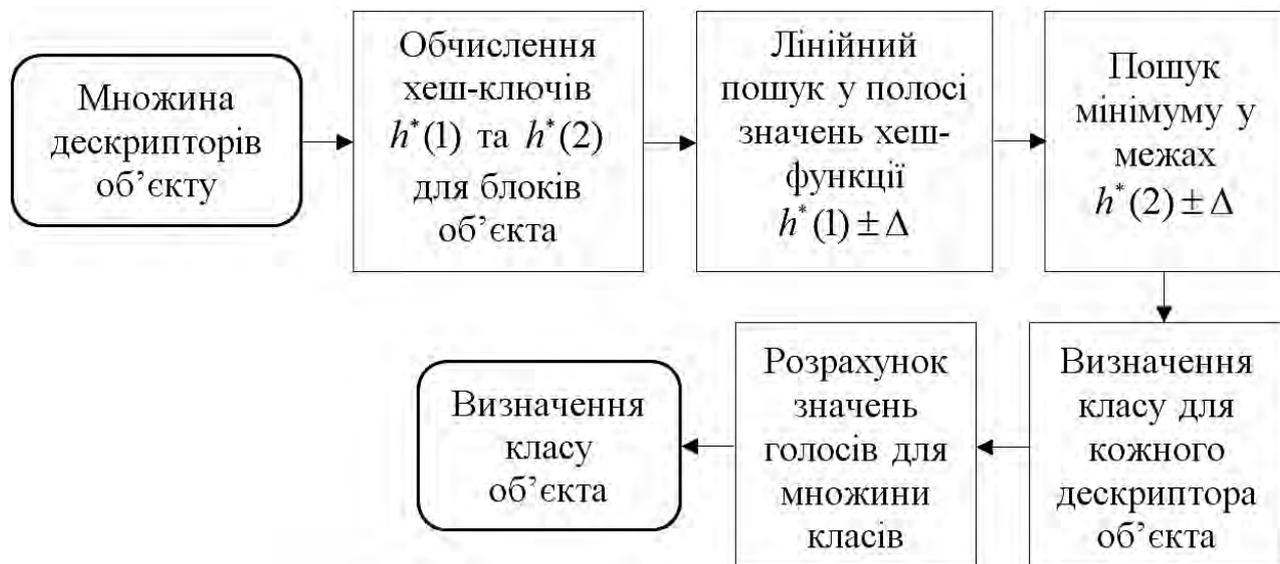


Рисунок 6.2 – Схема класифікації

Запропонований продуктивний за швидкістю спосіб класифікації на підґрунті впровадження хешування може реалізувати як точний, так і наближений види пошуку.

Параметрами є вид хеш-функції, точність Δ , кількість блоків, число кошиків хешу, побудованих для компонентів даних в індексованій структурі. Цими параметрами можна керувати, виходячи із наповненості кошиків для конкретних даних.

Розподіли (6.7) та (6.8) за класами еталонів для створених груп даних теж несуть важливу інформацію для класифікації, на підставі їх значень можна додатково на попередньому етапі взагалі видалити із аналізу окремі кошики (якщо класи всередині них розподілені суто рівномірно), або за значеннями розподілу приймати однозначне класифікаційне рішення (якщо спостерігаються суттєві переваги значень розподілів для окремих класів). Таке логічне оброблення сприяє подальшому спрощенню процедур класифікації.

6.5 Результати комп'ютерного моделювання

Програмне моделювання виконано у середовищі PyCharm Community Edition 2020.2.3 із використанням засобів бібліотеки OpenCV та мови програмування Python. Апаратне забезпечення – ноутбук з процесором Intel Pentium CPU N3540 2.16 ГГц та ОЗП 4 ГБ.

Для визначення дескрипторів ключових точок застосовано детектор ключових точок BRISK [6, 33, 37] з дескриптором розмірністю $n = 512$. Розроблені моделі класифікатора вжито на прикладі образів відомих людей (політик, артист, вчений), розмір зображень масштабовано до 500×500 пікселів.

Ілюстрація класів еталонних зображень та сформовані координати ключових точок показані на рисунку 6.3. Кількість обчислених дескрипторів у описі кожного із еталонів складає $s = 500$.



Рисунок 6.3 – Еталонні зображення з координатами ключових точок

У якості хеш-функції $h(\dots)$ взята кількість одиниць у двійковому коді вектора, параметр точності $\Delta = 13$ для обох блоків використано як $\pm 5\%$ від розміру блока 256, що відображає допустиме відхилення для значення хеш-функції.

Зважаючи на бінарний вид аналізованих даних, для зіставлення векторів застосовано метрику Хемінга (6.2). Всередині полоси фіксованих значень хешу у кошиках у експерименті застосовано традиційний лінійний пошук.

З метою зіставлення характеристик швидкодії програмно змодельований метод лінійного пошуку на усій множині еталонних даних (1500 дескрипторів) без застосування індексованих структур.

Для множини еталонних зображень із застосуванням індексованих хеш-структур на 256 кошиків та для лінійного пошуку отримано максимально точний результат: усі 500 дескрипторів кожного із еталонів правильно віднесено до свого класу.

При цьому час оброблення без індексного хешування склав приблизно 500 с, а із впровадженням комбінаційного хешування – лише 49 с.

Для пошуку полоси даних у відсортованому масиві за значенням $h(1)$ можна замість лінійного пошуку запровадити, наприклад, метод дихотомії [82, 100], що зробить виграш ще суттєвішим. Як бачимо, виграш у часі обчислень для запропонованого підходу (при заданому числі 3 еталонів та $s = 500$) складає більше 10 разів у порівнянні із традиційним. При цьому ясно, що виграш у часі обчислень збільшується при зростанні числа еталонів та дескрипторів у описах. Крім того, час обчислень суттєво залежить від створеної програмної моделі, типу комп'ютера, способу доступу до використаного програмного забезпечення, тому об'єктивними є лише відносні показники.

Хешування належить до точних методів трансформації даних. Природно зрозуміло, що похибка класифікації із впровадженням хешування може виникнути через те, що всередині корзини згрупованих даних можуть зустрітися рівноцінні або близькі за хеш-функцією дескриптори різних класів [51, 69]. Важливим для забезпечення результативної класифікації представляється вплив параметра порогу δ_p для значення мінімальної відстані при прийнятті рішення про клас дескриптора об'єкта за правилом (6.1). Діапазон значень для вибраної метрики складає $\delta_p \in [0, \dots, 256]$.

Експерименти з впливом перетворення повороту (найбільшим чином викривляє зображення, рис. 6.4) для вхідних еталонних зображень показали, що пропонуваній метод успішно класифікує усі еталонні зображення.



Рисунок 6.4 – Зображення з координатами ключових точок при повороті

Гістограми голосів для першого зображення рисунку 6.4 містить значення 415, 52, 33, (для інших двох трансформованих зображень – (104, 330, 66) та (60, 59, 381)), що підтверджує впевнену класифікацію для кожного із класів. Максимум числа голосів за «правильний» еталон більше ніж утричі перевищує найближче значення.

Задля більш детального аналізу кількісного складу кошиків за показниками (6.7), (6.8) проведено обчислення цих значень шляхом розбиття відсортованої множини даних на 8 рівноцінних блоків за значенням хеш-функції $h(1)$ у діапазоні 0...255.

Таблиця 6.1 містить значення показника (6.7) для 8-ми рівнозначних діапазонів значень $h(1)$.

Таблиця 6.1 – Числові показники класів для значень хеш-функції $h(1)$

Класи	Діапазони значень $h(1)$							
	31	63	95	127	159	191	223	255
1	0	0	43	140	224	92	1	0
2	0	0	25	143	257	72	3	0
3	0	0	6	111	254	122	7	0

Із таблиці 6.1 видно, що задля подальшого скорочення обчислень на класифікацію для даного набору еталонів деякі діапазони (0...63, 192...255) можна взагалі виключити із аналізу, для діапазонів 64...95, 160...191 встановити значущу перевагу одного із класів, а для решти діапазонів застосовувати розроблений метод. Це дає можливість додатково скоротити час класифікації.

Таблиця 6.1 розкриває подальші можливості для удосконалення запропонованого методу класифікації шляхом врахування конкретних значень вибраної хеш-функції для фіксованої бази еталонних зображень. Вона містить кількість значень хеш-функції $h(1)$, що потрапили в рівномірні діапазони даних (кратні 31) для різних еталонів.

Результати таблиці 6.1 дають можливість при здійсненні класифікації взагалі не аналізувати окремі діапазони хеш-функції (0...63) або однозначно визначати клас із-за суттєвої переваги для одного із еталонів. Цей попередній аналіз еталонних даних додатково дасть можливість скоротити час класифікації.

У всіх проведених експериментах класифікація здійснювалася вірно, тобто усі вхідні зображення з перетвореннями правильно отримали свій номер класу. Отримані результати підтверджують ефективність розробленого методу.

В основі розробленого методу лежить модель метричного порівняння з еталоном та підрахунок голосів дескрипторів для окремих класів, що у статистичному аспекті відповідає принципу максимальної правдоподібності.

Із введенням індексної структури та хешування у експерименті отримано суттєвий вигаш у часі класифікації більше ніж у 10 разів. Основним обмеженням впровадження хешування традиційно вважається деяке зростання потрібних обсягів пам'яті комп'ютера [19, 27, 31, 98, 100].

У нашій статті [104] описано результати експериментів по застосуванню апарату хешування для класифікації зображень різних порід собак з використанням детектора ORB. Дослідження показало, що час класифікації спадає зі збільшенням числа кошиків через те, що зменшуються кількості дескрипторів всередині кошиків, а після визначення кошика обчислюється відстань Хемінга уже для меншого числа дескрипторів. Пошук кошика за значенням хеш-функції відбувається значно швидше у порівнянні з обчисленням метрики для множини дескрипторів всередині кошиків.

Хешування на 256 кошиків тут застосовано у традиційному сенсі без додаткового оброблення даних, що зменшує обчислювальні затрати у порівнянні із розглянутим методом.

Дослідження для трьох еталонів по 500 дескрипторів у їх описах показало вигаш у швидкодії класифікації приблизно у 65 разів для методу з хешуванням у порівнянні з класичним лінійним пошуком [104]. Точні результати класифікації еталонів отримано і широкого діапазону значень числа кошиків. Ці результати підтверджують доцільність та ефективність впровадження хешування у процедурах структурного розпізнавання зображень.

У запропонованому дослідженні з метою скорочення обсягів метричного пошуку у задачі класифікації зображень структурними методами застосовано поєднання ланцюжка засобів: хешування, розбиття на блоки (проекування), LSH-оброблення у вигляді сортування за значенням хеш-функції, логічний аналіз (узгоджена фільтрація) значень хеш-функції для блоків.

Переваги запропонованого методу – лінійний пошук тільки у обмеженій і параметрично керованій полосі, зниження розмірності даних, пристосування до особливостей прикладної задачі класифікації із врахуванням відхилення значень схожості у певних межах.

Впровадження додаткового логічного аналізу еталонних даних на попередньому етапі оброблення буде сприяти подальшому скороченню часу оброблення при класифікації.

Особливість класифікатора полягає у тому, що на кожному кроці здійснюється не точний пошук релевантного дескриптора опису, а з врахуванням допустимого відхилення. Це покращує завадостійкість.

Розглянуті методи класифікації базуються на принципі «порівняння з еталоном» і можуть бути універсально застосовані до довільних векторів даних бітового виду.

При застосуванні хеш-структур ключовим моментом є вибір результативної та адаптованої до даних хеш-функції, яка здійснює часткову класифікацію, не змінюючи дані. Більш глибокий аналіз значень хеш-функції для конкретних еталонів дає можливість щодо подальшого скорочення необхідних обчислювальних витрат.

Наукову новизну дослідження складає розроблення продуктивного методу класифікації зображень на основі впровадження швидкісного пошуку із використанням ланцюгового способу оброблення індексованих хеш-структур даних, що скорочує потрібний обсяг обчислень в десятки разів.

Практична значущість роботи полягає у побудові моделей класифікації у трансформованому просторі даних, підтвердженні працездатності запропонованих модифікацій на прикладах зображень, створенні програмних застосунків для впровадження розроблених методів класифікації у системах комп'ютерного зору.

Перспективи дослідження можуть бути пов'язані із введенням логічного оброблення розподілів даних всередині хеш-структури, вивченням завадостійкості розроблених методів та оцінюванні їх прикладної результативності стосовно об'ємних колекцій зображень.

Матеріали дослідження опубліковано у статті [1].

7 ТРАНСФОРМАЦІЯ ОПИСІВ ЗОБРАЖЕНЬ ЗАДЛЯ ПОБУДОВИ КЛАСИФІКАЦІЙНИХ ОЗНАК

7.1 Матриця відстаней та ортогональні перетворення

Формування продуктивної системи ознак для забезпечення ефективного розпізнавання зображень у комп'ютерному зорі є ключовим завданням науки про дані та великих даних [1, 52, 56, 103, 105]. Пріоритетним напрямком останніх досліджень є вивчення природи аналізованих даних для встановлення наявних або прихованих закономірностей у їхньому складі [105, 106] та впровадження отриманих знань у процес розпізнавання [4, 107–109].

Подання опису в багатозначних просторах, особливо через його перетворення [50, 53, 72, 83, 108–110], дає змогу зосередити увагу на значущих класифікаційних властивостях даних, що забезпечить достатній рівень ефективності в прикладному аспекті. Опис є відображенням об'єктивної реальності – образним сигналом, відображеним у параметрах нового простору.

Удосконалюються традиційні методи. Ці методи використовують проєкційні значення для перегляду вхідних даних у просторі ортогональних функцій і використання цих значень як ознак для розпізнавання [106, 107]. Ці методи мають суворе формальне обґрунтування, і вони можуть універсально обробляти будь-які дані.

Ефективність цих методів залежить від ступеня розподіленості обраних системних функцій щодо аналізованих даних [24, 55, 85, 110, 111]. При обробці ортогональних систем щодо заданої бази описів зображень можна трохи скоротити повний набір елементів розкладання, щоб створити обмежену підмножину, яка відповідає потребам продуктивної класифікації. У цьому випадку, пропорційно ступеню зниження, є можливість досягти значного зниження обчислювальних витрат [112].

При застосуванні апарату декомпозиції базовий набір для проєктування описів уже заданий, тому його не потрібно формувати в процесі навчання. Тут технологія навчання може сприяти як безпосередній адаптації до бази еталону, так і формуванню ефективної підмножини, що дає змогу виконувати ефективний аналіз за обмежений час [1, 50, 53, 107].

Структуровані методи класифікації мають обмежений набір векторів для опису ключових точок на зображенні, щоб показати, як виглядає об'єкт [7, 33, 44, 113].

Методи, що використовують дескриптори ключових точок, фізично базуються на інформації та характеристиках самого зображення. Використання описів у вигляді набору бінарних дескрипторів особливо ефективно для вбудованих систем: безпілотних літальних апаратів, мобільних пристроїв, роботизованих і супутникових систем [34, 113, 114]. Значення міри релевантності для пари описів зображень оцінюється як близькість або відстань між двома наборами векторів приблизно однакової потужності [1, 53, 109].

Основою для визначення такої міри є матриця відстані, яка містить значення метрики між усіма парами елементів двох множин за правилом «один до одного». Релевантність множин за принципом голосування також розраховується за допомогою матриці відстаней [52, 83, 85].

Матриця відстані відноситься до структурних метричних моделей даних; він виводиться з набору значень проаналізованих даних і містить більш значущу класифікаційну інформацію про відмінності між наборами даних, ніж прямі значення елементів цих наборів. Статистика відстаней, навіть в межах одного набору (опису), суттєво відображає властивості даних, і вони самостійно можуть бути основою для класифікації.

Дослідники систем комп'ютерного зору докладають все більше зусиль для створення нових методів класифікації зображень, які можна використовувати в реальному житті [54, 56, 93, 108].

Ми можемо додати розпізнавання за набором дескрипторів ключових точок [1, 4, 53, 109] до списку методів, які добре працюють, коли є геометричні зміни та збурення. Реалізація цього пристрою забезпечує інваріантність до геометричних перетворень зображень, високу швидкість аналізу даних і гідні результати. Ці методи, порівняно з нейронними мережами, не вимагають тривалого навчання, оскільки вони витягують інформацію про особливості безпосередньо із зображення. Подальшому покращенню їх показників сприяє також можливість реалізації навчальних процедур, у тому числі з викладачем, у таких класифікаторах, де враховуються детальні описи занять.

Прикладом ефективного вирішення проблеми класифікації хвороб рослин у сільському господарстві є гібридний підхід на основі методів глибинного навчання [29, 37].

Оскільки сигнали зображення є багатовимірними та розподіленими в просторі разом із чітко визначеним набором ознак, багато статистичних методів [4, 11, 52, 109] та методи аналізу просторових даних стали популярними. Зазначені методи значно полегшують обробку [50, 103, 105, 107].

Розглянуті методи включають декомпозицію даних в ортогональному просторі [106, 108, 110, 111] і використання моделі матриці відстані [15, 52, 83, 105]. Одним із напрямів розвитку структурних методів є трансформація існуючої системи ознак (набору дескрипторів) шляхом створення модифікованого простору з необхідними класифікаційними ознаками.

Перетворення структурного опису, як ієрархічне представлення [73], кластерне представлення [50], поділ дескрипторів на групи за відстанню до медоїду [53], просторовий аналіз системи непересічних фрагментів [4, 11], і застосування розподілів для компонентів опису запропоновані та успішно реалізовані в задачі класифікації [109]. Перетворення не тільки прискорюють роботу класифікатора, але й забезпечують достатній рівень точності. Розроблено високошвидкісні методи [1, 85, 103], засновані на представленні описів «центрами» даних, але їх ефективність може бути недостатньою, бо багато в чому залежить від координати методу вибір центрів і вмісту даних.

У статті [83] пропонується використовувати матрицю відстаней на множині класів як критерій якості класифікації. Кінцевим результатом є модель для знаходження найкращого порогу, щоб побачити, що частини опису однакові, що є рішенням проблеми оптимізації за двома критеріями.

Порогове значення 0,32, яке було знайдено під час експериментів, є найнижчим рівнем сукупної якості для дескрипторів прискорених надійних функцій (SURF) із 64 вимірами. Метричні відношення між елементами в структурному описі можуть бути джерелом для побудови кінцевих характеристик. Матриця відстаней між компонентами опису є складовою процедур кластеризації в просторі ознак з подальшою побудовою моделі класифікатора в рамках технології «bag of words» [50, 52, 105].

Використання матриці попарних відстаней у задачах візуалізації багатовимірних об'єктів на основі карт Кохонена відоме в інформаційно-пошукових системах [52, 103] при оцінці результатів класифікації [83]. Інструмент матриці відстаней може бути реалізований на наборі класів зображень, які відповідають критеріям матриці неточностей [52].

Варіант ефективною презентації може бути пов'язаний з перетворенням зображення в ортогональному просторі. Застосування системи функцій Уолша та пов'язаних з ними перетворень [110, 111] представляє практичний інтерес завдяки своїй простій програмно-технічній реалізації. Реалізація кусково-постійних базисних функцій Уолша має незначні обчислювальні витрати, бо не пов'язана з операціями множення і реалізована в просторі цілих чисел.

Отриманий набір цілочисельних векторів ознак є міцною основою для побудови класифікатора. Компоненти опису мають характеристики, найважливішими з яких є інформативність і метричні зв'язки всередині класів. Значення цих параметрів оцінюється за допомогою спеціальних критеріїв як на етапі навчання, так і за результатом вимірювання. На основі цього опис можна скоротити для прискорення процесу класифікації [15, 83, 112].

Робимо висновок, що концептуальна декомпозиція набору компонентних дескрипторів і формалізм матриці відстаней для композиції описів класів можуть бути джерелами для побудови нової ефективної системи класифікаційних ознак. Пропонується дослідити можливість та ефективність їх застосування в моделях структурної класифікації.

Метою роботи є розробка структурних методів класифікації зображень в аспекті побудови нових просторів класифікаційних ознак. Перетворення опису в набір дескрипторів шляхом реалізації апарату компонентної декомпозиції на основі системи ортогональних функцій і на основі значень матриці відстаней для елементів опису призводить до скорочення обчислювальних витрат при аналізі численних наборів багатовимірних компонентів.

Завданнями дослідження є реалізація ортогональної декомпозиції еталону та вхідних зображень, використання апарату матриці відстані для отримання інтегральних характеристик опису, побудова метричного класифікатора в перетвореному просторі ознак та дослідження ефективності розроблених класифікаторів. модифікації (продуктивність, швидкість, імунітет) за допомогою імітаційного моделювання.

У просторі структурних ознак опис $Z = \{z_i\}_{i=1}^s$ зображення є скінченною множиною із s векторів Z_i , $s = \text{card } Z$, $z_i \in B^n$, $z_i = \{z_{i,j}\}_{j=1}^n$, B^n – векторний простір розмірності n з бінарними компонентами $\{0,1\}$ [1, 4].

Дескриптор ключових точок зображення розглядаємо як точку n -мірного дискретного сигналу. Будемо трактувати перетворення опису Z як відображення $B^n \rightarrow R$, де R – новостворений простір класифікаційних ознак.

Схема впровадження R подана на рисунку 7.1.

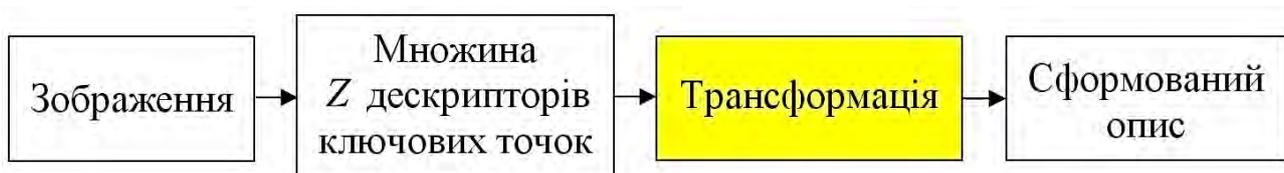


Рисунок 7.1 – Схема застосування трансформації опису

Вивчається два підходи до модифікації простору ознак: ортогональне розкладення і подання у вигляді матриці відстаней між компонентами опису.

Формування нового простору вивчається в аспекті застосування модифікованого опису у задачі класифікації.

7.2 Подання опису в ортогональному базисі та класифікація

У будь-якому скінченному векторному просторі існує хоча б одна ортогональна система векторів $W = \{\{w_{j,b}\}_{j=1}^n\}_{b=1}^n$, така, що будь-який вектор $z \in B^n$, $z = \{z_b\}_{b=1}^n$ може бути поданий у ній кортежем $\alpha = \{\alpha_j\}_{j=1}^n$, де

$$\alpha_j = (z \cdot w_j) = \sum_{b=1}^n z_b w_{j,b}. \quad (7.1)$$

Коефіцієнт (7.1) визначається як скалярний добуток вектора z на вектор w_j ортогональної системи [107, 110, 111, 115]. Параметри α_j є координатами вектора у базисі W . Для кожного $z_i \in Z$ маємо значення вектору розкладання α_i .

Із впровадження моделі векторного простору до класифікації впливає, що будь-який дескриптор $z_i \in Z$ образу зображення можна обернено єдиним способом подати як лінійну комбінацію базисних векторів, а вибраний базис складає при цьому точно n векторів, тобто

$$z_i = (\alpha_i \cdot w_j) = \sum_{b=1}^n \alpha_{i,b} w_{j,b}. \quad (7.2)$$

Зауважимо, що у n -мірному просторі векторів теоретично існує множина різних ортогональних базисів [115]. Зважаючи на те, що функції Уолша формально приймають значення $\{1, -1\}$ та безпосередньо не входять у простір B^n , перетворення (7.1) фактично відображає трансформацію $B^n \rightarrow C^n$ у простір C^n векторів з цілими компонентами.

У той же час простір $B^n \subset C^n$ є підпростором для простору C^n , тому можна вважати, що перетворення (7.1) та (7.2) здійснюються у просторі C^n .

Коефіцієнт нормування для функцій Уолша до ортонормальної системи, який забезпечує зворотність перетворень (7.1) та (7.2), є константою при заданому n , тому при обчисленнях ознак його можна не враховувати.

У результаті застосування по-компонентної трансформації $z_i \rightarrow \alpha_i$ опис Z отримає вид $\alpha(Z) = \{\alpha_i\}_{i=1}^s$, тобто множина $\{Z_i\}_{i=1}^s$, $z_i \in B^n$, дескрипторів ключових точок трансформується у множину $\{\alpha_i\}_{i=1}^s$, $\alpha_i \in C^n$, тієї ж розмірності та чисельності.

Виникає питання про результативність такої трансформації при здійсненні класифікації.

Перетворення (7.1), (7.2) можна записати у векторній формі з використанням матриці A Адамара [110]. Повний набір функцій Уолша розмірності n утворює ортогональну матрицю Адамара A розміром $n \times n$, яка складається із n векторів функцій Уолша W_1, \dots, W_n . Застосуємо розкладання опису у базисі сімейства прямокутних базисних функцій – дискретних функцій Уолша, які є векторами цілих чисел (значення $+1, -1$) скінченної розмірності (ступінь двійки).

Опис Z розглядаємо як прямокутну матрицю $Z = \{z_{i,j}\}$, $i = \overline{1, s}$, $j = \overline{1, n}$ розміром $s \times n$, рядки якої містять s дескрипторів.

Перетворення Z можна подати як множення прямокутної та квадратної матриць

$$U = Z \times A. \quad (7.3)$$

У результаті (7.3) отримаємо прямокутну матрицю $U = \{u_{i,j}\}$, $i = \overline{1, s}$, $j = \overline{1, n}$ розміром $s \times n$, рядки якої містять вектори розкладання α для дескрипторів опису Z .

Зважаючи на те, що у нашому дослідженні набір функцій Уолша розглядається і застосовується у повному складі, ми не будемо аналізувати різноманіття способів упорядкування систем функцій Уолша (Уолш, Адамар, Пелі, Трахтман та інші), кожна з яких має ряд своїх властивостей і відповідні сфери застосування [111].

Традиційно класифікація полягає у встановленні ступеня релевантності $\Theta(Z, E_k)$ між описом $\{Z_i\}_{i=1}^s$ аналізованого зображення та складових $E_k = \{e_i(k)\}_{i=1}^s$ бази $E = \{E_k\}_{k=1}^N$ еталонних описів [1, 53, 108].

Клас m об'єкта визначається відповідно до екстремуму функції $\Theta(Z, E_k)$ як

$$m = \arg \underset{k=1, \dots, N}{extr} \Theta(Z, E_k). \quad (7.4)$$

Після здійснення трансформації процес класифікації буде базуватися на визначенні модифікованої міри релевантності $\Theta^\alpha(\alpha(Z), \alpha(E_k))$ для трансформованих описів $\alpha(Z)$ та $\alpha(E_k)$ у новому просторі ознак. Кожний із перетворених описів є множиною векторів із компонентами із C^n .

Маючи наявну можливість здійснення редукції даних у новоствореному просторі, будемо мати на увазі і деякий редукований простір C^α , отриманий із C^n зменшенням числа застосованих функцій Уолша до $q \ll n$. Зважаючи на обмежений діапазон значень вхідного сигналу та контрольоване перетворення у просторі цілих чисел, при побудові класифікатора є можливість точно оцінити діапазон значень функцій Θ та Θ^α і використати ці знання для класифікації.

Одним із варіантів Θ чи Θ^α є метрика для множин векторів. Це може бути відстань Танімото (Жаккара) для множин A, B , яка містить відношення числа елементів симетричної різниці та об'єднання множин [52]

$$T(A, B) = \frac{\text{card}(A \Delta B)}{\text{card}(A \cup B)}. \quad (7.5)$$

Міра (7.5) відображає кількісні характеристики еквівалентних та відмінних елементів порівнюваних множин. При застосуванні метрики (7.5) важливим є її лімітоване значення, для якого елементи векторного простору вважаються еквівалентними [52, 83]. Ця проблема є однією із ключових у багатовимірних просторах даних. Вибір порогу еквівалентності суттєвим чином впливає на результат класифікації. Для її вирішення запропонуємо два способи. Один із них визначає граничне значення

$$\rho_{\text{lim}}(a, b) = \alpha \rho_{\text{max}}(a, b), \quad (7.6)$$

як відсоток α від теоретично визначеного максимуму метрики. Наприклад, для метрики Хемінга, що для векторів розміром 256 бітів змінюється на інтервалі $[0, 256]$, при відсотку у $\alpha = 25\%$ можна визначити $\rho_{\text{lim}}(a, b) = 0.25 * 256 = 64$.

Більш практичним є спосіб отримання $\rho_{\text{lim}}(a, b)$ за результатом аналізу прикладних експериментальних даних, де ρ_{max} обчислено для набору еталонів. Час обчислення ρ_{max} не впливає на витрати для класифікації і здійснюється на попередньому етапі аналізу даних.

Будемо оцінювати продуктивність методу класифікації показником точності pr , який обчислюється відношенням числа правильно класифікованих об'єктів r_p до загального їх числа r , що використовувалося в експерименті [52]

$$pr = r_p / r. \quad (7.7)$$

Важливим показником результативності методів розпізнавання зображень також є завадостійкість. Вона характеризується значенням точності класифікації в умовах дії завад [56, 85, 108].

Якщо дія адитивної завади на зображення $B(x, y)$ описується моделлю $B_\xi(x, y) = B(x, y) + \xi(x, y)$, а завада $\xi(x, y)$ характеризується середньоквадратичним відхиленням σ з нульовим математичним очікуванням, то відношення сигнал-шум описують як $\mu = B_m / \sigma$, де B_m – амплітудна характеристика (наприклад, середнє значення яскравості).

Показником завадостійкості є залежність $pr(\mu)$ для даного методу.

7.3 Редукція у модифікованому просторі

Зважаючи на те, що визначення трансформованої системи ознак засновано на інтеграційному процесі розкладання даних вхідного опису за ортогональним спектром, у просторі C^n з'являється можливість формування компактного подання $\alpha^*(Z)$ та $\alpha^*(E_k)$ із стисненим обсягом коефіцієнтів. Таким чином здійснюється трансформація до редукованого простору $C^n \rightarrow C^\alpha$.

Стиснення розмірності простору даних реалізує перспективну ідею скорочення обсягу обчислень для Θ^α шляхом зменшення розміру вектору ознак і задля зменшення часу класифікації [106, 108, 110].

Один із традиційних способів полягає у відборі серед множини $\{\alpha_i\}_{i=1}^s$ найбільш інформативної її підмножини $\{\alpha_i\}_{i=1}^q$, $q \ll s$, що в той же час забезпечує достатньо якісну класифікацію.

Таке ж саме оброблення можна було б спробувати здійснити і безпосередньо для вхідної множини $\{z_i\}_{i=1}^s$ ознак, наприклад, за встановленим критерієм інформативності [2]. Однак пряме скорочення може привести до суттєвих втрат інформації пропорційно обсягу модифікованого опису.

У той же час обчислені проєкції $\{\alpha_i\}_{i=1}^s$ не тільки мають багатоспектральний сенс даних, але і містять у більшій мірі інтегровану інформацію за рахунок процесу розкладення (7.1) за системою базисних функцій, тому можуть бути успішно покладені в основу продуктивного стиснення даних.

Формування стисненої підмножини $\{\alpha_i\}_{i=1}^q$ можна здійснити нетрадиційно відбором компактною частини серед функцій системи W , тобто використати для класифікації лише деяку підмножину W_q із W , $W_q \subset W$. У такому випадку відпадає необхідність формування повної системи коефіцієнтів розкладання як для вхідного образу, так і для еталонної множини E . Таке оброблення набуло популярності у спектральних методах, де шляхом фільтрації незначущих компонентів спектру за рахунок деякого збільшення помилки відновлення сигналу зменшують надмірність представлення та скорочують вектор оброблюваних даних [106, 111]. При цьому показники функціонування систем знижуються незначним чином.

Якщо процедурою фільтрації вдасться відібрати високоінформативну частину спектру, яка забезпечує якісну класифікацію, то можна суттєво скоротити обчислювальні затрати на класифікацію із забезпеченням необхідного рівня результативності. Отримана в результаті секвестрування підмножина найбільш інформативної частини спектру функцій Уолша є результатом навчання на заданій множині класів, її формування безпосередньо залежить від аналізованого простору даних і адаптоване до нього. Для інших варіантів колекцій еталонних даних підмножина компонентів спектру може мати зовсім інший склад.

Поставимо задачу шляхом трансформації у новий простір на основі побудови відображення $T: \alpha(Z) \rightarrow \alpha^*(Z)$, або фактично $T: W \rightarrow W_q$, за рахунок застосування ортогонального розкладання за компактною системою W_q функцій Уолша сформувані стиснені описи $\alpha^*(Z)$ та $\alpha^*(E_k)$, що забезпечують достатні показники результативності класифікації. Таке відображення $T = T(E, W)$ є функцією як E , так і складу ортогональної системи функцій W .

Внаслідок наявності значних кореляційних зв'язків між елементами зображення реальної природи, що знаходить своє відтворення у дескрипторах ключових точок, основна енергія у дискретному спектрі має тенденцію концентруватися у відносно невеликій кількості відліків, що відповідають повільно осцилюючим базовим функціям.

Без істотної шкоди для відновлення чи класифікації зображення невеликі за величиною спектральні коефіцієнти можна взагалі обнулити (або відкинути їх аналіз), а значущі елементи спектру використовувати у процесі розпізнавання [111].

Дисперсійний аналіз коефіцієнтів ортогонального перетворення є на сьогодні основним апаратом для оцінювання їх значущості у моделі подання сигналу (7.1), так як середньоквадратична помилка відновлення залежить від дисперсійних характеристик [85, 111]. Тому природним критерієм для вибору множини значимих коефіцієнтів розкладення є дисперсія. Аналізуємо криву нормованих дисперсій в залежності від номера функції Уолша, розміщених за спаданням, і впровадимо у класифікацію підмножину найбільших коефіцієнтів.

Застосування пропонованого апарату в задачі класифікації, де описи зображень подано множиною дескрипторів, має свої особливості та можливості [112]. Справа в тому, що не завжди коефіцієнти з найбільшими дисперсіями забезпечують потрібний рівень розрізнення зображень. Другим фактором є можливість здійснення дисперсійного аналізу даних як у рамках повної бази еталонів, так і в межах окремих її представників, що може вплинути на результативність. Додатковою можливістю є дисперсійний аналіз для системи фрагментів дескрипторного опису, що реалізує віконне оброблення [110].

Для множин E описів еталонів отримаємо матрицю (7.3) розміром $N_S \times n$, обчислимо для кожного стовпця матриці U вектори математичного очікування Mu_i та дисперсії σ_i^2

$$Mu_i = \frac{1}{N_S} \sum_{j=1}^{N_S} u_{i,j}, \quad \sigma_i^2 = \frac{1}{N_S - 1} \sum_{j=1}^{N_S} (u_{i,j} - Mu_i)^2. \quad (7.8)$$

При визначенні показників (7.8) для окремого еталону приймаємо $N = 1$. Здійснимо ранжирування списку функцій Уолша w_i за значенням дисперсії σ_i^2 для компонент спектру: w_1, w_2, \dots, w_n , щоб була виконана умова $\sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_n^2$ спадання. Покладемо отриманий список в основу процедури стиснення.

7.4 Формалізм матриці відстаней для множини дескрипторів

Нехай $A = \{a_i\}_{i=1}^s$, $B = \{b_j\}_{j=1}^s$ – дві скінченні множини n -мірних векторів з рівнозначною потужністю s , $A \subset R^n$, $B \subset R^n$, $\text{card } A = s$, $\text{card } B = s$, R^n – простір числових векторів розмірності n . Вибір рівноцінної потужності дещо спрощує здійснення аналізу множин і може бути досягнутий направленим корегуванням їх потужності.

Множини A, B дескрипторів ключових точок для зображень на практиці відносять до класу мультимножин, так як вони часто містять близькі і навіть еквівалентні елементи [1, 53].

Розглянемо формалізм матриці відстаней M розміром $s \times s$ для складу елементів двох множин A, B :

$$M[A, B] = \{ \{m_{i,j}\}_{i=1}^s \}_{j=1}^s, \quad m_{i,j} = \rho(a_i, b_j), \quad (7.9)$$

де $\rho(a_i, b_j)$ – відстань у просторі R^n , $a_i \in A$, $b_j \in B$.

Для бінарних векторів у якості ρ може бути застосована ефективна у обчислювальному аспекті відстань Хемінга [52]. Формалізм (7.9) задає метричне відношення Ω для елементів пари множин: $\Omega[A, B] \rightarrow M$, де кожній двійці елементів $a_i \in A$, $b_j \in B$ ставиться у відповідність значення ρ . Як відомо, відстань між об'єктами безпосередньо пов'язана із ймовірністю їх еквівалентності у метрично-статистичній теорії класифікації образів [34, 107].

Кожний i -й рядок матриці M відповідає i -й компоненті опису і містить набір відстаней від i -го елемента множини A до елементів із B .

Зважаючи на властивість симетричності для метрики $\rho(a_i, b_j) = \rho(b_j, a_i)$, квадратна матриця M симетрична. Однак у задачі класифікації одну із множин вважають еталоном (наприклад, A), тому доцільно аналізувати рядки чи стовпці M [15]. Окремим випадком для M є відношення $\Omega[A, A]$, коли множини співпадають. Тоді матриця M буде відображати спектр значень відстаней для елементів множини A , що теж може бути значимою ознакою для класифікації.

Важлива для класифікації відзнака полягає в тому, що матриця M зберігає інваріантність до геометричних перетворень зображення. Ця властивість безпосередньо впливає із інваріантності значень дескрипторів.

Якщо значення дескриптора зберігається при геометричних перетвореннях, то і відстані між дескрипторами теж не змінюються. Зрозуміло, що в реальних умовах як інваріантність дескрипторів, так і інваріантність матриці відстаней фіксується і перевіряється наближено.

На підставі матриці M відстаней із введенням граничних значень для метрики щодо еквівалентності двох векторів можуть бути визначені такі експериментальні результати дій над множинами, як перетин, об'єднання, різниця, симетрична різниця та інші [56, 107, 115].

Потужність множин, отриманих у результаті цих дій, також є класифікаційною ознакою. Якщо елементи a_i, b_j у прикладному аспекті вважати еквівалентними ($a_i \approx b_j$) при виконанні деякої граничної умови для відстані між ними

$$a_i \approx b_j \mid \rho(a_i, b_j) \leq \delta_\rho, \quad (7.10)$$

де δ_ρ – задане лімітоване значення для ρ , то перетин $A \cap B$ можна визначити як множину елементів A , для яких у відповідному рядку матриці M виконана хоча б одна умова (7.10). Визначення перетину, різниці та об'єднання множин векторів потрібно, наприклад, для обчислення значення відстані Танімото (7.5).

Одне із безпосередніх застосувань апарату матриці відстаней полягає у здійсненні редукції (відбору) найбільш ефективних для класифікації елементів множин на підставі введеного критерію [52]. При цьому реалізується мета скорочення опису зі збереженням необхідної результативності класифікації. Для бінарних векторів – дескрипторів ключових точок за результатами досліджень можна вважати еквівалентними між собою два вектори у межах відстані Хемінга в 25% від максимуму відстані. Наприклад, для бінарних векторів АКАZE [33] розмірністю 488 два дескриптори вважаються рівноцінними, якщо відстань Хемінга для них не перевищує значення $\delta_\rho = 122$.

Розглянемо деяку базу із N еталонних зображень у вигляді множини E описів: $E = \{E_1, E_2, \dots, E_N\}$. E – це навчальна вибірка, яка є підґрунтям для здійснення класифікації методом порівняння з еталоном [4]. Кожний еталонний опис $E_k \subset E$ у формалізмі класифікатора репрезентує окремий клас.

Опис еталону $E_k = \{e_v(k)\}_{v=1}^s$ – це скінченна множина дескрипторів ключових точок, $e_v(k) \in B^n$, $s = \text{card } E_k$ – число дескрипторів у множині.

Кожний дескриптор $e_v(k)$ бази E має параметр k номеру класу, а загальне число ознак – дескрипторів у базовій множині E складає $card E = sN$.

На підставі значень матриць відстаней $M[E_k, E_k]$ та $M[E_k, E_j]$, $\forall j \neq k$ можна визначити значення критерію інформативності для елементів бази E і ефективно здійснити редукцію еталонних описів задля зменшення обчислювальних витрат (часу класифікації) [2, 53]. Дескриптори з невисоким значенням критерію відсіюються як не значимі. Класифікація здійснюється на основі підмножини інформативних дескрипторів опису. Редукція здійснюється на етапі попереднього оброблення даних, і її впровадження безпосередньо ніяк не впливає на часові затрати під час класифікації.

7.5 Моделі класифікації на основі матриці відстаней

Матриця відстаней між компонентами опису є складовою у процедурах кластеризації у просторі ознак з подальшою побудовою моделі класифікатора в рамках технології «мішок слів» [50, 52, 73]. Як правило, кластеризацію здійснюють для множини E , кожний еталон E_k отримує модифікований опис $E_k = (d_1, \dots, d_u)_k$ у виді кількісного подання за системою із u сформованих кластерів. Класифікація здійснюється шляхом зіставлення вектору кластерного подання об'єкту і еталонів або конкурентним віднесенням елементів опису об'єкту до кластерних центрів бази E [50].

Застосуванням матриці відстаней є обчислення визначних класифікаційних характеристик опису. Таким параметром є медоїд множини, що у задачі класифікації застосовується як деякий центр опису.

Медоїд – це елемент множини з найменшою сумарною відстанню до решти елементів [52, 53]. Значення медоїду та здійснення ранжирування елементів за агрегованою відстанню покладено в побудову модифікацій класифікаторів на підставі стисненого опису [53, 116].

Медоїд $med A$ для множини A на основі матриці $M[A, A]$ визначається як дескриптор з найменшим значенням суми у рядку (стовпці) матриці

$$z = med A | z \in A, z = \arg \min_i \sum_{j=1}^s m_{i,j}. \quad (7.11)$$

Значення матриць $M[A, A]$, $M[A, B]$, $B \neq A$, $B \subset E$ покладемо в основу оцінювання результативності класифікації в рамках бази E . Це можна здійснити з використанням таких параметрів, як точність класифікації відносно складу компонентів.

Точність класифікації (7.7) оцінюється відношенням кількості правильно ідентифікованих компонентів множини A до їх числа s . Факт правильності класифікації для $z \in E_k$ встановлюється предикатом $right(z)$:

$$right(z, M) = 1 \mid z \in E_k, k = \arg \min_{i,j} \{m_{i,j}\}_{j=1}^{Ns}, \quad (7.12)$$

тобто якщо мінімум у рядку матриць відстаней для бази еталонів E досягається для еталону, якому апріорно належить аналізований дескриптор.

На основі такого попереднього критеріального аналізу сукупності матриць відстаней $M[E_k, E_j]$, $j, k = 1, \dots, s$ формально можна встановити гарантований рівень класифікаційної працездатності для будь-якої новоствореної системи ознак, не проводячи детального експериментального дослідження. Заснований на матриці відстаней аналіз даних бази еталонів може бути підставою для необхідності удосконалення вибраної системи ознак задля забезпечення потрібного рівня результативності.

Розглянемо набір матриць відстаней для бази еталонів. Якщо образи $M[Z, Z]$ аналізованого об'єкту Z чи еталонних класів E_k задати у вигляді матриць $M[E_k, E_k]$, $k = \overline{1, s}$ то обчислення ступеня релевантності для пари описів можна буде подати як встановлення нового відношення між метричними відношеннями, які відображаються матрицями $\{M[E_k, E_k]\}_{k=1}^N$ та $M[Z, Z]$. Таке відношення можна створити на підставі статистики значень $M[E_k, E_k]$, $k = \overline{1, s}$. Відомо, що застосування розподілів замість значень даних суттєво знижує обчислювальні витрати на класифікацію [4, 8, 109]. Зважаючи на відомий діапазон величини метрики, побудуємо для матриці гістограму (розподіл) як число елементів, що приймають фіксоване значення x відстані:

$$q_k[x] = \text{card}\{m_{i,j} \mid m_{i,j} = x\}. \quad (7.13)$$

Тепер класифікацію можна здійснити у просторі векторів з цілими компонентами, до яких відносяться значення гістограм $q_k[x]$.

Конкретно для метрики Хемінга, застосованої для бінарних дескрипторів розміром 512, маємо діапазон аргументу гістограми $x \in [0, \dots, 512]$. Здійснимо класифікацію на підставі метрико-статистичного підходу шляхом обчислення і оптимізації на множині еталонів для відстані (наприклад, манхеттенської) між розподілами об'єкту та еталону

$$\rho(q_z[x], q_k[x]) = \sum_x |q_z[x] - q_k[x]|, \quad (7.14)$$

де $q_z[x]$ – гістограма значень матриці відстаней для об'єкту.

Зважаючи на те, що при геометричних перетвореннях порядок слідування дескрипторів у описі змінюється, ступінь релевантності двох матриць можна встановити лише за інтегральними критеріями, в тому числі і за розподілами. Крім того, критерієм класифікації може бути, наприклад, вектор сум $r_j = \sum_{i=1}^s m_{i,j}$, $j = \overline{1, s}$ для значень стовпців (рядків) матриці $\{m_{i,j}\}_{i=1}^s \}_{j=1}^s$, так як його компоненти зберігають інваріантність до зміни порядку елементів [15].

Для класифікатора, заснованого на по-елементному аналізі опису, для довільного дескриптора $z \in Z$ об'єкта отримуємо набір матриць відстаней для повної множини еталонів

$$M(z) = \left\{ \begin{array}{ccc} m_{1,1} & \dots & m_{1,s} \\ m_{2,1} & \dots & m_{2,s} \\ & \dots & \\ m_{N,1} & \dots & m_{N,s} \end{array} \right\}, \quad (7.15)$$

де перший індекс є номер еталону, а другий – номер дескриптора у складі еталону. Список (7.15) може бути основою для побудови класифікатора K , що відносить аналізований дескриптор Z до одного із класів $K : z \rightarrow \{1, \dots, N\}$.

Загалом клас k для дескриптора $z \in Z$ визначимо як аргумент оптимуму на множині класів для функції F матричного аргументу (7.15)

$$k = \arg \underset{i=1, \dots, N}{opt} F[M(z)]. \quad (7.16)$$

Результат визначається моделлю F .

У (7.16) може бути організовано визначення мінімуму або за всім списком розміром $N \times s$, або роздільно для кожного із еталонів (рядок) з наступним визначенням класу-переможця.

Введемо вектор $\{h_i\}_{i=1}^N$ з цілими значеннями для накопичення голосів класів. На підставі впровадження K для кожного дескриптора $z \in Z$ відповідно до (7.16) визначимо номер класу k , а потім інкрементуємо акумулятор $h_k = h_k + 1$ номеру класу. За результатом оброблення опису Z об'єкту накопичуємо вектор $\{h_i\}_{i=1}^N$. Клас Z визначимо як аргумент максимуму числа голосів

$$k = \arg \max_{i=1, \dots, N} h_i \parallel h_k \geq \delta_h, \quad (7.17)$$

де δ_h – поріг для мінімального числа голосів, який встановлюється експериментально для заданої бази. Якщо нерівність у (7.17) не виконується, клас об'єкту не встановлюється (відмова від класифікації).

7.6 Результати комп'ютерного моделювання

Для вивчення продуктивності запропонованих методів класифікації (працездатність, швидкодія, завадостійкість) проведено комп'ютерне моделювання. Для дослідження методу класифікації на основі ознак ортогонального розкладення обрано базу із п'яти зображень гербів футбольних клубів. Зображення мають розмір 325×325 пікселів (формат JPEG, рис. 7.2).



Рисунок 7.2 – Приклад еталону (а) та його зображення з координатами ключових точок (б)

Еталон розташовано на світлому фоні (рис. 7.2 (а)). Із застосуванням системи програмування Python та бібліотеки алгоритмів комп'ютерного зору OpenCV [4, 23, 44, 62, 112] сформовано 500 дескрипторів ORB розміром 256 біт для кожного зображення (рис. 7.2 (б)).

Діапазон значень отриманих трансформованих даних фіксований, що дає можливість оцінити показники класифікації. Номери перших 16 відсортованих функцій Уолша за величиною квадрату дисперсії на множині дескрипторів для бази еталонів мають вигляд: 0, 45, 106, 208, 176, 10, 85, 77, 109, 80, 209, 9, 3, 54, 72, 197. Затрати комп'ютерного часу для обчислення метрики Танімото між коефіцієнтами ортогонального подання склав 29 с, а з поданням у просторі 16 відібраних функцій Уолша – 1,7 с відповідно.

Як бачимо, застосування скороченого простору ознак знижує обчислювальні витрати у 17 разів.

При цьому обидва класифікатори (традиційний та модифікація) забезпечують безпомилкову класифікацію для навчальної вибірки.

Застосуємо метрику Танімото для множин у класифікаторі для трансформованого простору даних з використанням порогу ρ_{lim} як 25% від максимального значення метрики для дескрипторів еталонних даних. При цьому максимум значення метрики, а значить і поріг ρ_{lim} , залежать від впровадженого простору даних.

Досліджено вплив адитивного шуму на показники класифікації. Відмітимо, що відстань Танімото між еталоном і зашумленим еталоном поступово зростає зі збільшенням рівня шуму (зменшенням μ), а це впливає на точність класифікації.

Експериментальне вивчення завадостійкості розроблених модифікацій класифікатора з використанням метрики Танімото (7.5) і вибраного порогу ρ_{lim} під впливом адитивного шуму показало, що показник $pr(\mu)$ для методу з повним набором 256 функцій Уолша, а також для методу із стисненим поданням 16 функцій Уолша дорівнює 1 (класифікація без помилок) при рівні $\mu \geq 1,2$ і знижується до 0,9 при $\mu = 1$.

Для традиційного методу (без застосування апарату функцій Уолша) показник завадостійкості $pr(\mu)$ знаходиться практично в тих же межах. Для методу із стисненим поданням у 8 функцій Уолша значення $pr(\mu)$ знижується до 0,9 при $\mu = 1,2$. Бачимо, що розроблений метод за показником завадостійкості не поступається традиційному.

Загалом маємо досить високі показники завадостійкості для прикладних систем [34, 112], так як запропонований метод працює безпомилково навіть при низькому відношенні сигнал-шум, близьким до 1!

Час класифікації у модифікованих просторах даних при цьому скорочується пропорційно кількості використаних функцій Уолша. Експеримент показав зниження часових витрат у порівнянні з повним набором 256 функцій Уолша для 16 функцій Уолша у 19 разів, для 8 функцій Уолша – у 30 разів (табл. 7.1). При цьому час класифікації для традиційного методу шляхом обчислення метрики Танімото для описів без трансформації отримано у 12 разів вищий, ніж для модифікації із 16 функцій Уолша.

Таблиця 7.1 – Експериментальна оцінка часу класифікації, с

Метод	256 функцій Уолша	16 функцій Уолша	8 функцій Уолша	Традиційний
Час	228	12	7,5	140

Для здійснення моделювання класифікатора з використанням апарату матриці відстаней вибрано детектор KAZE у покращеному варіанті AKAZE, що формує дескриптори розміром 488 бітів [15, 33].

Створено програмне забезпечення на основі мови програмування Java та бібліотеки алгоритмів комп'ютерного зору OpenCV, для кожного зображення сформовано 40 дескрипторів AKAZE типу MLDB, який дає можливість змінювати розмір дескриптора [33, 113]. Експерименти здійснювалися для трьох класів зображень тварин (коней) розміром: 300 × 200 пікселів (рис. 7.3).



Рисунок 7.3 – Приклад зображення (а) і координати сформованих ключових точок (б)

Приклад зображення (рис. 7.3 (а)) і виділених координат ключових точок показано на рисунку 7.3 (б).

Зважаючи на те, що об'єм даних у матриці відстаней залежить від числа сформованих ключових точок, для моделювання взято число ключових точок $s = 40$ в кожному із трьох еталонів. Вибір ключових точок є параметром, що впливає на результат класифікації, так як зазвичай для різних еталонів отримується різне число ключових точок. Конкретно в експерименті випадково взято 40 точок з кожного еталону.

Нами побудовано експериментальні розподіли (7.13) значень відстаней окремо для кожного з еталонів, що показано різними кольорами на рисунку 7.4 (результат для першого еталону показано на рисунку 7.4 (а), для другого еталону – на рисунку 7.4 (б), для третього еталону – на рисунку 7.4 (в)).

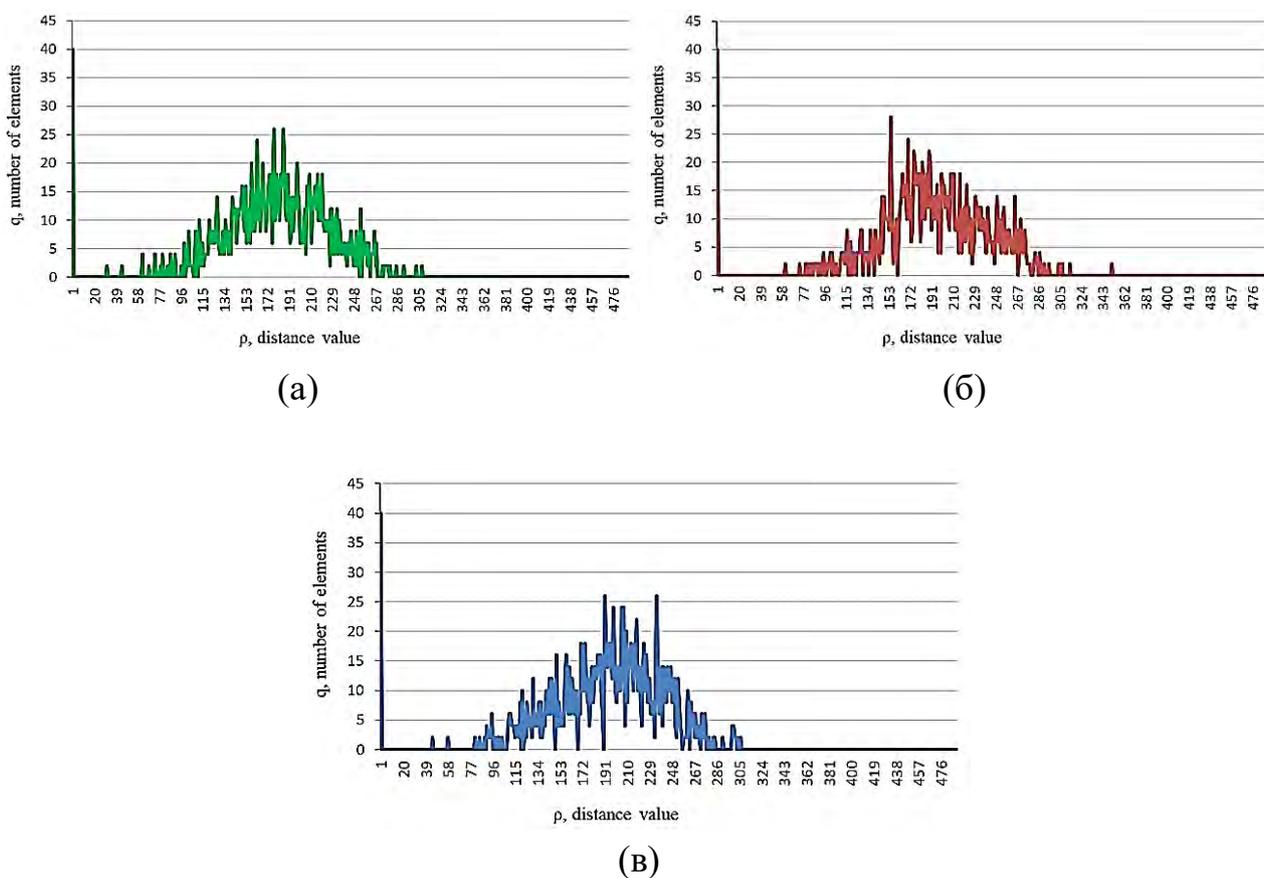


Рисунок 7.4 – Гістограми значень матриць відстаней для першого, другого та третього еталонів

Як бачимо, візуально розподіли у достатній мірі різняться, хоча практично всі їх значення знаходяться у обмеженому діапазоні 0–320.

Манхеттенська відстань (7.14) складає:

- еталони (1-2) – 888;
- еталони (1-3) – 1016;
- еталони (2-3) – 988.

Зважаючи на те, що формально діапазон зміни відстані (7.14) складає $0-3200$ ($2s^2$), можна стверджувати, що інтегральні ознаки розподілів (7.13) можна впровадити для класифікації, так як відмінність між ними для різних еталонів досягає 30% від максимально можливої. У той же час середня відстань всередині кожного з описів має значення 182, 190, 194. Якщо орієнтуватися на ці величини, то відносна вага отриманих відстаней у відсотковому плані буде ще вагомішою.

Тепер проаналізуємо укрупнене квантоване подання для еталонних матриць у інтервальній системі дискретних діапазонів $0-50, 51-100, \dots, 401-488$, отримуємо ефективну стиснену розмірність 10 для кожного із розподілів $q_k[x]$ (рис. 7.5).

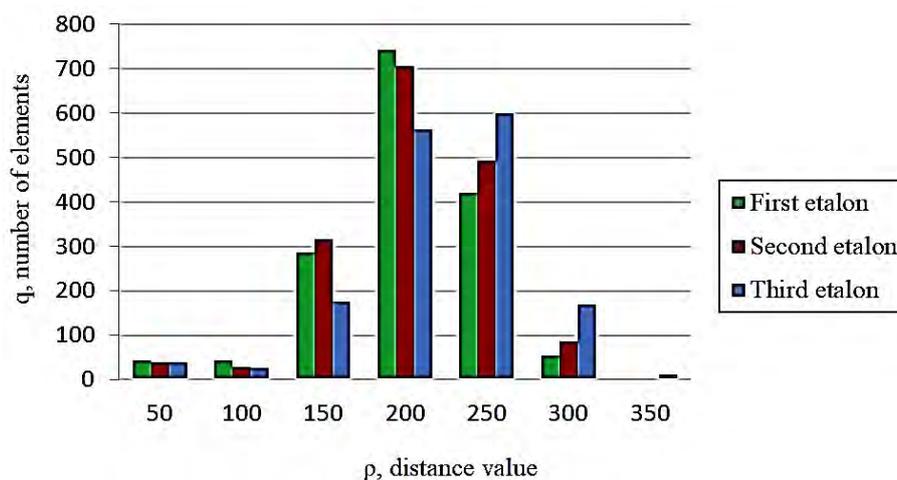


Рисунок 7.5 – Гістограми значень матриць відстаней у дискретних діапазонах

На рисунку 7.5 числа по осі X означають крайню праву межу дискретного діапазону. Гістограми рисунку 7.5 отримані шляхом підрахунку на рисунку 7.4 кількості значень всередині діапазону. У цьому квантованому просторі можна візуально побачити суттєву різницю між аналізованими розподілами в межах діапазону $100-300$ і приблизно однакові значення для класів у решті діапазонів. Це дає можливість зосередитися на здійсненні аналізу і обчислень лише у скороченому діапазоні відстаней $100-300$. Нами експериментально перевірено, що цей факт дійсний і для інших зображень, що розширює можливості спрощення для прискорення обчислень у процесі класифікації.

На рисунку 7.5 є підтвердження відомого факту суттєвої близькості між собою у просторі багатовимірних векторів, що помічено раніше [52, 103, 105].

Крім цього, на рисунку 7.5 візуально можна помітити істотну відмінність опису зображення 3 від інших, в той час як для зображень 1 та 2 спостерігається деяка близькість. У значеннях метрики це склало:

- еталони (1-2) – 216;
- еталони (1-3) – 620;
- еталони (2-3) – 408.

Природно відстані між образами у квантованому просторі дещо зменшилися, але принципова здатність розрізняти зображення кількох класів збереглась. Експериментально обчислені відстані між векторами сум $\{r_j\}_{j=1}^s$ для стовпців еталонних матриць відстаней склали:

- еталони (1-2) – 30524;
- еталони (1-3) – 44452;
- еталони (2-3) – 28694.

Зважаючи на середні експериментальні значення відстаней для еталонів у межах 182–194, можна вважати отримані вектори сум достатньо значимими класифікаційними ознаками. Подальше інтегрування значень відстаней шляхом розширення діапазонів показало, що відмінності між розподілами еталонів природно зменшуються. Для 5 рівноцінних діапазонів класифікація для аналізованих зображень ускладнюється.

Експериментальний підрахунок точності (7.7) класифікації показав, що для множини дескрипторів навчальної вибірки (120 векторів у досліджених еталонах) усі дескриптори класифіковано правильно. Таким чином, показник точності стосовно множини еталонних компонентів дорівнює 1.

У матрицях відстаней нулі зустрічаються тільки для власного еталону. Це можна пояснити великою розмірністю даних (488), що практично виключає випадкові збіги бітів у багатовимірних векторах. У той же час зрозуміло, що для реальних ситуацій під дією завад чи геометричних перетворень зображення показник точності природно може бути дещо нижчим.

Зважаючи на той факт, що експериментально обчислена відстань Танімото для використаних у експерименті описів склала 0,5–0,6, бачимо, що, незважаючи на значну подібність структурних описів аналізованих зображень, запропоновані спрощені моделі класифікації забезпечують достатній рівень їх розрізненості на практиці, а швидкодія оброблення для них у десятки разів вища.

Використання ортогональної функціональної системи для перетворення опису зображення показало потенціал для значного збільшення швидкості обробки, зберігаючи високий рівень точності класифікації та стійкості до перешкод. Критичні фактори для оптимізації ефективності цього представлення включають вибір метрики для відображення змінених описів і визначення порогу для визначення подібності компонентів у щойно згенерованому просторі даних.

Використання вимірювання Танімото разом із пороговим значенням, еквівалентним 25% від максимального метричного значення, продемонструвало свою ефективність в оцінці компонентів опису. Використання апарату функції Уолша призвело не тільки до десятикратного скорочення обчислювальних витрат, але й зберегло відносно хороші показники ефективності класифікації.

Реалізуючи матричну модель відстані, стало можливим створювати ефективні інтегровані ознаки у вигляді одновимірних розподілів даних і векторів, що представляють суму стовпців матриці. Це призвело до зниження обчислювальних витрат при збереженні ефективності класифікації вихідної вибірки даних. Програмне моделювання було проведено для експериментальної оцінки ефективності класифікації зображень у новостворених просторах ознак і часу, необхідного для розрахунку релевантності описів. Це порівнювали з традиційним підходом голосування та обчислення показників на наборі дескрипторів.

Дослідження представляє новий підхід до класифікації зображень шляхом вдосконалення структурного методу. Це досягається шляхом реалізації перетворення опису з використанням ортогональної декомпозиції даних і побудови ознакових моделей на основі матриць відстаней у просторі дескрипторів. Крім того, дослідження пропонує методи зменшення описів у новостворених просторах, що призводить до зменшення обчислювальних витрат на класифікацію.

Практичне значення даної роботи полягає в розробці моделей класифікації в просторі трансформованих даних, перевірці ефективності та стійкості запропонованих модифікацій на прикладах зображень та створенні програмних додатків для реалізації розроблених класифікаторів у системах комп'ютерного зору.

Перспективи дослідження можуть стосуватися розробки ряду запропонованих моделей для створення та вивчення змін в описах зображень у великих базах даних.

Матеріали дослідження опубліковано у статті [62].

ВИСНОВКИ

За результатами викладених у монографії результатів досліджень можна зробити наступні висновки.

Застосування редукції описів за критерієм інформативності підтвердило можливість отримати найбільш значущу для класифікації підмножину ознак, що гарантує необхідний рівень точності у практичних застосуваннях.

Дослідження показало високий рівень точності класифікації на підставі впровадженого навчання і виявило практичну можливість обрати таку структуру і параметри мережі Кохонена, що універсально забезпечать потрібний рівень критеріїв точності і швидкодії класифікації у прикладній задачі. Можна заявити про високу продуктивність і результативність технологій класифікації із застосуванням кластерних центрів. Отримано час опрацювання зображень за розробленими методами у сотні разів менший, ніж за традиційним, без зменшення точності.

Досліджено практичні питання вибору параметрів порогів для встановлення еквівалентності дескрипторів та для мінімуму числа голосів класів задля забезпечення необхідного рівня точності класифікації. Тестування підтвердило суттєве прискорення процесу оброблення і достатньо високий рівень точності класифікації із використанням пропонованих процедур стиснення. Конкретно у проведеному моделюванні досягнуто підвищення швидкодії у десятки разів. Впровадження апарату хешування реалізує перспективну ідею зниження розмірності аналізованих даних шляхом попередньої побудови для еталонної множини деякої індексованої структури.

На етапі класифікації опису застосовано узгоджене ланцюжкове поєднання етапів аналізу даних: розбиття на блоки, пошук всередині блоків з обмеженням на значення хеш-функції. Впровадження хеш-подання підвищує швидкодію ідентифікації чи класифікації візуальних об'єктів.

Застосування системи ортогональних функцій і матриці відстаней як апарату перетворення опису показало можливість значного підвищення швидкодії оброблення при збереженні високих показників точності.

Експериментальний приклад показав, що час, витрачений на обчислення релевантності описів відповідно до їх модифікованого представлення, більш ніж у десять разів менший, ніж для традиційних метричних підходів. Розроблені класифікаційні ознаки можуть бути використані у прикладних задачах, де критичним є час візуальної ідентифікації об'єктів.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Daradkeh Y. I., Gorokhovatskyi V., Tvoroshenko I., and Zeghid M. (2022) Tools for fast metric data search in structural methods for image classification. *IEEE Access*, vol. 10, pp. 124738–124746.
2. Oliinyk A., Subbotin S., Lovkin V., Blagodariov O., and Zaiko T. (2017) The system of criteria for feature informativeness estimation in pattern recognition. *Radio Electronics, Computer Science, Control*, vol. 43, no. 4, pp. 85–96.
3. Kuchuk H., Podorozhniak A., Liubchenko N., and Onischenko D. (2021) System of license plate recognition considering large camera shooting angles. *Radioelectronic and Computer Systems*, vol. 100, no. 4, pp. 82–91.
4. Daradkeh Y. I., Gorokhovatskyi V., Tvoroshenko I., Gadetska S., and Al-Dhaifallah M. (2021) Methods of classification of images on the basis of the values of statistical distributions for the composition of structural description components. *IEEE Access*, vol. 9, pp. 92964–92973.
5. Shen X., Wang C., Li X., Peng Y., He Z. *et al.* (2022) Learning scale awareness in keypoint extraction and description. *Pattern Recognition*, vol. 121, p. 108221.
6. Leutenegger S., Chli M., and Siegwart R. Y. (2011) BRISK: Binary robust invariant scalable keypoints. *In 2011 The International Conference on Computer Vision*, Barcelona, Spain, pp. 2548–2555.
7. Rublee E., Rabaud V., Konolige K., and Bradski G. (2011) ORB: An efficient alternative to SIFT or SURF. *In 2011 International Conference on Computer Vision*, Barcelona, Spain, pp. 2564–2571.
8. Гороховатський В. О., Гадецька С. В. (2020) Статистичне оброблення та аналіз даних у структурних методах класифікації зображень: монографія. Харків: ФОП Панов А. М., 128 с.
9. Wu Q., Li Y., Lin Y., and Zhou R. (2018) Weighted sparse image classification based on low rank representation. *Computers, Materials & Continua*, vol. 56, no. 1, pp. 91–105.
10. Nasirahmadi A., and Ashtiani S.-H. M. (2017) Bag-of-feature model for sweet and bitter almond classification. *Biosystems Engineering*, vol. 156, pp. 51–60.
11. Daradkeh Y. I., Gorokhovatskyi V., Tvoroshenko I., Gadetska S., and Al-Dhaifallah M. (2023) Statistical data analysis models for determining the relevance of structural image descriptions. *IEEE Access*, vol. 11, pp. 126938–126949.

12. Wang Z., Zhao X., and Tao Y. (2023) Integrated algorithm based on bidirectional characteristics and feature selection for fire image classification. *Electronics*, vol. 12, no. 22, p. 4566.
13. Flach P. (2012) Machine Learning: The Art and Science of Algorithms that Make Sense of Data. New York, NY, USA: Cambridge University Press, pp. 298–329.
14. 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*, vol. 79, no. 1, pp. 59–69.
15. Gorokhovatskyi V., Peredrii O., Tvoroshenko I., and Markov T. (2023) Distance matrix for a set of structural description components as a tool for image classifier creating. *Advanced Information Systems*, vol. 7, no. 1, pp. 5–13.
16. Ghahremani M., Liu Y., and Tiddeman B. (2021) FFD: Fast feature detector. *IEEE Transactions on Image Processing*, vol. 30, pp. 1153–1168.
17. Jegou H., Douze M., Schmid C., and Perez P. (2010) Aggregating local descriptors in to a compact image representation. *In 2020 IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 3304–3311.
18. Daradkeh Y. I. and Tvoroshenko I. (2020) Application of an improved formal model of the hybrid development of ontologies in complex information systems. *Applied Sciences*, vol. 10, no. 19, p. 6777.
19. Leskovec J., Rajaraman A., and Ullman J. D. (2020) Mining of Massive Datasets. New York, NY, USA: Cambridge University Press, pp. 73–128.
20. Stańczyk U. (2015) Feature selection for data and pattern recognition. *In Studies in Computational Intelligence*, vol. 584, Berlin-Heidelberg, Germany, pp. 29–44.
21. Biagio M. S., Bazzani L., Cristani M., and Murino V. (2014) Weighted bag of visual words for object recognition. *In 2014 IEEE International Conference on Image Processing (ICIP)*, Paris, France, pp. 2734–2738.
22. Bai Q., Li S., Yang J., Song Q., Li Z. *et al.* (2020) Object detection recognition and robot grasping based on machine learning: A survey. *IEEE Access*, vol. 8, pp. 181855–181879.
23. Gorokhovatskyi V., Tvoroshenko I., Kobylin O., and Vlasenko N. (2023) Search for visual objects by request in the form of a cluster representation for the structural image description. *Advances in Electrical and Electronic Engineering*, vol. 21, no. 1, pp. 19–27.

24. Gorokhovatskyi O., Peredrii O., Gorokhovatskyi V., and Vlasenko N. (2023) Explanation of CNN image classifiers with hiding parts. In: *J. Benois-Pineau, R. Bourqui, D. Petkovic, G. Quenot (eds), Explainable Deep Learning Artificial Intelligence*, pp. 125–146.
25. Fang Y., and Liu L. (2022) Angular quantization online hashing for image retrieval. *IEEE Access*, vol. 10, pp. 72577–72589.
26. Гороховатський В. О., Полякова Т. В. (2018) Застосування просторових структур ознак для класифікації зображень у комп'ютерному зорі: монографія. Харків: ФОП Панов А. М., 120 с.
27. Mashtalir S., and Mashtalir V. (2020) Spatio-temporal video segmentation, in: advances in spatio-temporal segmentation of visual data. In *Studies in Computational Intelligence*, vol. 876, Cham, Switzerland, pp. 161–210.
28. Gorokhovatskyi V., Gadetska S., and Ponomarenko R. (2020) Recognition of visual objects based on statistical distributions for blocks of structural description of image. In *Lecture Notes in Computational Intelligence and Decision Making. Proc. of the XV International Scientific Conference “Intellectual Systems of Decision Making and Problems of Computational Intelligence” (ISDMCI'2019)*, vol. 1020, Cham, Switzerland, pp. 501–512.
29. Wang Y., and Chen P. (2023) A lightweight multi-scale quadratic separation convolution module for CNN image-classification tasks. *Electronics*, vol. 12, no. 23, p. 4839.
30. Carrasco D. P., Rashwan H. A., García M. Á., and Puig D. (2023) T-YOLO: Tiny vehicle detection based on YOLO and multi-scale convolutional neural networks. *IEEE Access*, vol. 11, pp. 22430–22440.
31. Hu Z., Bodyanskiy Y. V., and Tyshchenko O. K. (2017) A deep cascade neural network based on extended neo-fuzzy neurons and its adaptive learning algorithm. In *2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, Kyiv, Ukraine, pp. 801–805.
32. Murthy C. B., Hashmi M. F., Muhammad G., and AlQahtani S. A. (2021) YOLOv2PD: An efficient pedestrian detection algorithm using improved YOLOv2 model. *Computers, Materials & Continua*, vol. 69, no. 3, pp. 3015–3031.
33. Yakovleva O., and Nikolaieva K. (2020) Research of descriptor based image normalization and comparative analysis of SURF, SIFT, BRISK, ORB, KAZE, AKAZE descriptors. *Advanced Information Systems*, vol. 4, no. 4, pp. 89–101.
34. Tymchyshyn R., Volkov O., Gospodarchuk O., and Bogachuk Y. (2018) Modern approaches to computer vision. *Control Systems and Computers*, vol. 6, no. 278, pp. 46–73.

35. Ahmad M. A., Gorokhovatskyi V., Tvoroshenko I., Vlasenko N., and Mustafa S. K. (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*, vol. 69, no. 10, pp. 186–192.
36. Karras T., Laine S., and Aila T. (2021) A style-based generator architecture for generative adversarial networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 12, pp. 4217–4228.
37. Wei P., Ball J., and Anderson D. (2018) Fusion of an ensemble of augmented image detectors for robust object detection. *Sensors*, vol. 18, no. 3, p. 894.
38. Muja M., and Lowe D. (2009) Fast approximate nearest neighbors with automatic algorithm configuration. In *VISAPP International Conference on Computer Vision Theory and Applications*, Lisboa, Portugal, pp. 331–340.
39. Szeliski R. (2010) *Computer Vision: Algorithms and Applications*. London, Great Britain: Springer-Verlag, pp. 655–718.
40. Путятін Є. П., Аверін С. І. (1990) Обробка зображень в робототехніці. *Машинобудування*, 320 с.
41. Robnik-Sikonja M., and Kononenko I. (2003) Theoretical and empirical analysis of ReliefF and RReliefF. *Machine Learning*, vol. 53, pp. 23–69.
42. Daradkeh Y. I., Gorokhovatskyi V., Tvoroshenko I., and Zeghid M. (2024) Improving the effectiveness of image classification structural methods by compressing the description according to the information content criterion. *Computers, Materials & Continua*, vol. 80, no. 2, pp. 3085–3106.
43. Karami E., Prasad S., and Shehata M. (2017) Image matching using SIFT, SURF, BRIEF and ORB: Performance comparison for distorted images. In *Proceedings of the 2015 Newfoundland Electrical and Computer Engineering Conference*, St. John's, Canada, arXiv preprint arXiv:1710.02726.
44. Гороховатський В. О. (2014) Структурний аналіз та інтелектуальна обробка даних у комп'ютерному зорі: монографія. Харків: Компанія СМІТ, 316 с.
45. Putyatin Y., Gorohovatsky V., Gorohovatsky A., and Peredriy E. (2008) Projective methods of image recognition. In *Proceedings XIV International Conference "Knowledge-Dialogue-Solution" (KDS 2008)*, pp. 37–42.
46. Гороховатський В. О. (2003) Розпізнавання зображень в умовах неповної інформації: монографія. Харків: ХНУРЕ, 112 с.

47. Zhang S., Wang J., Tao X., Gong Y., and Zheng N. (2017) Constructing deep sparse coding network for image classification. *Pattern Recognition*, vol. 64, pp. 130–140.
48. Yu T., Chen W., Junfeng G., and Poxi H. (2022) Intelligent detection method of forgings defects detection based on improved EfficientNet and Memetic algorithm. *IEEE Access*, vol. 10, pp. 79553–79563.
49. Lin T.-Y., Goyal P., Girshick R., He K., and Dollár P. (2020) Focal loss for dense object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 318–327.
50. Daradkeh Y. I., Gorokhovatskyi V., Tvoroshenko I., and Zeghid M. (2022) Cluster representation of the structural description of images for effective classification. *Computers, Materials & Continua*, vol. 73, no. 3, pp. 6069–6084.
51. Zhang X., Yu F. X., Karaman S., and Chang S.-F. (2017) Learning discriminative and transformation covariant local feature detectors. *In Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4923–4931.
52. Гороховатський В. О., Творошенко І. С. (2022) Аналіз багатовимірних даних за описом у формі множини компонент: монографія. Харків: ХНУРЕ, 124 с.
53. Gorokhovatskiy V. A. (2011) Compression of descriptions in the structural image recognition. *Telecommunications and Radio Engineering*, vol. 70, no. 15, pp. 1363–1371.
54. Kuchuk H., Kovalenko A., Ibrahim B. F., and Ruban I. (2019) Adaptive compression method for video information. *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8, no. 1, pp. 66–69.
55. 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)*, vol. 2608, pp. 1013–1026.
56. Scherer R. (2020) Computer vision methods for fast image classification and retrieval. *In Studies in Computational Intelligence*, Cham, Switzerland: Springer Nature Switzerland AG, pp. 107–118.
57. Kohonen T. (2001) *Self-Organizing Maps*. Berlin-Heidelberg, Germany: Springer, pp. 245–261.
58. Hu Z., Bodyanskiy Y. V., and Tyshchenko O. K. (2019) *Self-Learning and Adaptive Algorithms for Business Applications*. Leeds, England, UK: Emerald Publishing Limited, pp. 51–77.

59. da Silva I. N., Spatti D. H., Flauzino R. A., Liboni L. H. B., and dos Reis Alves S. F. (2017) *Artificial Neural Networks*. Cham, Switzerland: Springer, pp. 157–172.
60. Aggarwal C. C. (2023) *Neural Networks and Deep Learning*. Cham, Switzerland: Springer Nature Switzerland AG, pp. 73–117.
61. Wang P., Fan E., and Wang P. (2021) Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. *Pattern Recognition Letters*, vol. 141, pp. 61–67.
62. Gorokhovatskyi V., Tvoroshenko I., and Yakovleva O. (2024) Transforming image descriptions as a set of descriptors to construct classification features. *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 33, no. 1, pp. 113–125.
63. Malik F., and Baharudin B. (2013) Analysis of distance metrics in content-based image retrieval using statistical quantized histogram texture features in the DCT domain. *Journal of King Saud University – Computer and Information Sciences*, vol. 25, no. 2, pp. 207–218.
64. Zhou Z.-H. (2021) *Machine Learning*. Singapore, Singapore: Springer, pp. 181–210.
65. Mawane J., Naji A., and Ramdani M. (2022) A cluster validity for optimal configuration of Kohonen maps in e-learning recommendation. *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no. 1, pp. 482–492.
66. Huang Y., Cheng Z., Zhou Q., Xiang Y., and Zhao R. (2020) Data mining algorithm for cloud network information based on artificial intelligence decision mechanism. *IEEE Access*, vol. 8, pp. 53394–53407.
67. Aggarwal C. C., and Reddy C. K. (2014) *Data Clustering: Algorithms and Application*. Boca Raton, Florida, USA: CRC Press, pp. 483–500.
68. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O., Hudáková M., and Gorokhovatskyi O. (2024) Application a committee of Kohonen neural networks to training of image classifier based on description of descriptors set. *IEEE Access*, vol. 12, pp. 73376–73385.
69. Celik C., and Bilge H. S. (2017) Content based image retrieval with sparse representations and local feature descriptors: A comparative study. *Pattern Recognition*, vol. 68, pp. 1–13.
70. Wu H., Liu Q., and Liu X. (2019) A review on deep learning approaches to image classification and object segmentation. *Computers, Materials & Continua*, vol. 60, no. 2, pp. 575–597.

71. Zheng Y., Fan J., Zhang J., and Gao X. (2019) Exploiting related and unrelated tasks for hierarchical metric learning and image classification. *IEEE Transactions on Image Processing*, vol. 29, pp. 883–896.
72. Daradkeh Y. I., Tvoroshenko I., Gorokhovatskyi V., Latiff L. A., and Ahmad N. (2021) Development of effective methods for structural image recognition using the principles of data granulation and apparatus of fuzzy logic. *IEEE Access*, vol. 9, pp. 13417–13428.
73. Daradkeh Y. I., Gorokhovatskyi V., Tvoroshenko I., and Al-Dhaifallah M. (2022) Classification of images based on a system of hierarchical features. *Computers, Materials & Continua*, vol. 72, no. 1, pp. 1785–1797.
74. Gorokhovatskyi V., Rusakova N., and Tvoroshenko I. (2020) The application of image analysis methods and predicate logic in applied problems of magnetic monitoring. *Telecommunications and Radio Engineering*, vol. 79, no. 20, pp. 1801–1811.
75. Gorokhovatskyi O., Gorokhovatskyi V., and Peredrii O. (2018) Analysis of application of cluster descriptions in space of characteristic image features. *Data*, vol. 3, no. 4, p. 52.
76. 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*, vol. 79, no. 2, pp. 149–163.
77. Gorokhovatskyi V. O., Tvoroshenko I. S., and Vlasenko N. V. (2020) Using fuzzy clustering in structural methods of image classification. *Telecommunications and Radio Engineering*, vol. 79, no. 9, pp. 781–791.
78. Zhang Y., Bai X., Fan R., and Wang Z. (2019) Deviation-sparse fuzzy c-means with neighbor information constraint. *IEEE Transactions on Fuzzy Systems*, vol. 27, no. 1, pp. 185–199.
79. Sun W., Chen X., Zhang X., Dai G., Chang P. *et al.* (2021) A multi-feature learning model with enhanced local attention for vehicle re-identification. *Computers, Materials & Continua*, vol. 69, no. 3, pp. 3549–3561.
80. Chen R., Pan L., Li C., Zhou Y., Chen A. *et al.* (2020) An improved deep fusion CNN for image recognition. *Computers, Materials & Continua*, vol. 65, no. 2, pp. 1691–1706.
81. Gorokhovatsky V. A., Gorokhovatsky A. V., and Berestovsky A. Ye. (2016) Intellectual data processing and self-organization of structural features at recognition of visual objects. *Telecommunications and Radio Engineering*, vol. 75, no. 2, pp. 155–168.

82. 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*, vol. 78, no. 20, pp. 1857–1868.
83. Gorokhovatsky V., Gorokhovatsky V. A., Vlasenko A. N., and Vlasenko N. V. (2014) Quality criteria for multidimensional object recognition based upon distance matrices. *Telecommunications and Radio Engineering*, vol. 73, no. 18, pp. 1661–1670.
84. Daradkeh Y. I., and Tvoroshenko I. (2020) Technologies for making reliable decisions on a variety of effective factors using fuzzy logic. *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 5, pp. 43–50.
85. Shapiro L., and Stockman G. (2001) *Computer Vision*. Upper Saddle River, New Jersey, USA: Prentice Hall, pp. 595–607.
86. Gadetska S. V., and Gorokhovatsky V. O. (2018) Statistical measures for computation of the image relevance of visual objects in the structural image classification methods. *Telecommunications and Radio Engineering*, vol. 77, no. 12, pp. 1041–1053.
87. 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*, vol. 79, no. 10, pp. 855–863.
88. Filatov V., and Kovalenko A. (2019) Advances in spatio-temporal segmentation of visual data. *In Studies in Computational Intelligence*, vol. 876, New York, NY, USA: Springer, Cham, pp. 243–274.
89. Xiong H., and Li Z. (2014) *Data Clustering: Algorithms and Application*. Boca Raton, Florida, USA: CRC Press, Chapter 23.
90. Li H., Xie F., Zhou J., and Liu J. (2024) Object detection, segmentation and categorization in artificial intelligence. *Electronics*, vol. 13, no. 13, p. 2650.
91. Lowe D. G. (2004) Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, vol. 60, pp. 91–110.
92. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O., and Hudáková M. (2025) Image description compression in classification structural methods. *IEEE Access*, vol. 13, pp. 43631–43641.
93. Baldini L., Martino A., and Rizzi A. (2019) Stochastic information granules extraction for graph embedding and classification. *In Proceedings 11th International Joint Conference on Computational Intelligence*, vol. 1, pp. 391–402.

94. Zadeh L. A. (1997) Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy Sets and Systems*, vol. 90, no. 2, pp. 111–127.
95. Berehulyak O., and Vorobel R. (2020) The algebraic model with an asymmetric characteristic of logarithmic transformation. *In Proceedings 2020 IEEE 15th International Conference on Computer Science and Information Technologies (CSIT)*, pp. 119–122.
96. Yakovleva O., Kovtunenکو A., Liubchenko V., Honcharenko V., and Kobylin O. (2023) Face detection for video surveillance-based security system. *In CEUR Workshop Proceedings: COLINS-2023*, vol. 3403, pp. 69–86.
97. Crowley J., and Riff O. (2003) Scale space methods in computer vision. *In Proceedings 4th International Conference Scale-Space '03*, vol. 2695, pp. 584–598.
98. Gorokhovatskyi V. A., Gorokhovatskiy A. V., and Peredrii Ye. O. (2018) Hashing of structural descriptions at building of the class image descriptor, computing of relevance and classification of the visual objects. *Telecommunications and Radio Engineering*, vol. 77, no. 13, pp. 1159–1168.
99. Yasser M., Hussain K. F., and Ali S. A. E.-F. (2022) An efficient hyperdimensional computing paradigm for face recognition. *IEEE Access*, vol. 10, pp. 85170–85179.
100. Aho A. V., Hopcroft J. E., and Ullman J. D. (2003) *Data Structures and Algorithms*. Delhi, India: Pearson Education, pp. 135–186.
101. Gorokhovatsky V., Stiahlyk N., and Tsarevska V. (2021) Combination method of accelerated metric data search in image classification problems. *Advanced Intelligent Systems*, vol. 5, no. 3, pp. 5–12.
102. Gorokhovatsky V. A., and Putyatin Ye. P. (2009) Image likelihood measures of the basis of the set of conformities. *Telecommunications and Radio Engineering*, vol. 68, no. 9, pp. 763–778.
103. Manning C. D., Raghavan P., and Schütze H. (2009) *An Introduction to Information Retrieval*. Cambridge, England: Cambridge University Press, pp. 421–441.
104. Gorokhovatskyi V., Vlasenko N., and Rybalka M. (2021) Application of data hashing tools to accelerate classification decisions in structural image recognition methods. *Advanced Information Systems*, vol. 5, no. 2, pp. 13–20.
105. Nong Y. (2013) *Data Mining: Theories, Algorithms, and Examples*. Boca Raton, FL, USA: CRC Press, Taylor & Francis Group, pp. 21–136.
106. Alessio S. M. (2016) *Signals and Communication Technology*. Cham, Switzerland: Springer International Publishing Switzerland, pp. 55–132.

107. Pratt W. K. (2006) *Digital Image Processing*: PIKS Scientific Inside. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 189–216.
108. Vlasenko N. V., and Sytnik O. V. (2013) Classification of video-objects in attribute space of the Walsh functions. *Telecommunications and Radio Engineering*, vol. 72, no. 19, pp. 1777–1785.
109. Gadetska S. V., Gorokhovatskyi V. O., Stiahlyk N. I., and 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*, vol. 4, pp. 58–68.
110. Ahmed N., and Rao K. R. (1975) *Orthogonal Transforms for Digital Signal Processing*. Heidelberg, Germany: Springer Berlin, pp. 99–152.
111. Premkumar K., Twinky H., Naveen S., Kathiresan M., and Aravind R. (2023) A secured communication in watermarking using fast Walsh Hadamard transformation in image processing. *In Proceedings 6th International Conference on Electronics, Communication and Aerospace Technology*, Coimbatore, India, pp. 1186–1191.
112. Gorokhovatskyi V., Tvoroshenko I., and Chmutov Y. (2022) Application of systems of orthogonal functions for formation of sign space in image classification methods. *Advanced Information Systems*, vol. 6, no. 3, pp. 5–12.
113. Gorokhovatskyi V., Chmutov Y., Tvoroshenko I., and Kobylin O. (2025) Reducing computational costs by compressing the structural description in image classification methods. *Advanced Information Systems*, vol. 9, no. 1, pp. 5–12.
114. Tareen S. A. K., and Saleem Z. (2018) A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK. *In Proceedings International Conference on Computing, Mathematics & Engineering Technologies (iCoMET)*, Sukkur, Pakistan, pp. 1–10.
115. Zeidler E. (2012) *Applied Mathematical Sciences*. New York, NY, USA: Springer, pp. 167–223.
116. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O., and Hudáková M. (2026) Feature space quantization and application of metrics in structural methods of image recognition. *IEEE Access*, vol. 14, pp. 17812–17824.

Наукове видання

ГОРОХОВАТСЬКИЙ Володимир
ТВОРОШЕНКО Ірина

**ПРОДУКТИВНІ МОДЕЛІ АНАЛІЗУ ДАНИХ
У МЕТОДАХ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ**

Монографія

За авторською редакцією

Відповідальний випусковий О. Кобилін
Комп'ютерне верстка І. Творошенко

Підп. до друку 10.12.25.
Умов. друк. арк. 6,7.

Формат 60×84 1/16.
Тираж 300 прим.

Спосіб друку – ризографія.
Ціна договірна.

Віддруковано в типографії ФОП Єфименко С.А.
61166, м. Харків, вул. Коломенська, 27
Свідоцтво про державну реєстрацію
Серія ДК № 6869 від 08.08.2019 р.