

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

Факультет Інформаційно-аналітичних технологій та менеджменту  
(повна назва)

Кафедра Інформатики  
(повна назва)

**КВАЛІФІКАЦІЙНА РОБОТА**  
**Пояснювальна записка**

рівень вищої освіти другий (магістерський)

**ДОСЛІДЖЕННЯ ПІДВИЩЕННЯ ТОЧНОСТІ КЛАСИФІКАЦІЇ**  
**ЗОБРАЖЕНЬ ЗА ДОПОМОГОЮ МЕТОДІВ СЕМАНТИЧНОГО**  
**АНАЛІЗУ ТА ВАГОВИХ КОЕФІЦІЄНТІВ**

(тема)

Виконав:  
студент 2 курсу, групи ІНФМ-23-2

Гончаров Д.М.  
(прізвище, ініціали)

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

Тип програми освітньо-професійна

Освітня програма Інформатика  
(повна назва освітньої програми)

Керівник доц. Руденко Д.О.  
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри \_\_\_\_\_  
(підпис)

Кобилін О.А.  
(прізвище, ініціали)

2025 р.

## Харківський національний університет радіоелектроніки

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

ЗАТВЕРДЖУЮ:

Зав. кафедри \_\_\_\_\_  
(підпис)

«\_\_\_\_» \_\_\_\_\_ 2025 р.

**ЗАВДАННЯ**  
НА КВАЛІФІКАЦІЙНУ РОБОТУстудентові Гончаров Данило Миколайович  
(прізвище, ім'я, по батькові)1. Тема роботи Дослідження підвищення точності класифікації зображень за допомогою методів семантичного аналізу та вагових коефіцієнтів

затверджена наказом по університету від 25 листопада 2024 року № 1246Ст

2. Термін подання студентом роботи до екзаменаційної комісії 30 грудня 2024 р.3. Вихідні дані до роботи математичні моделі семантичного аналізу ознак зображень, теоретичні відомості про методи класифікації зображень із використанням вагових коефіцієнтів, вибірка зображень для навчання та тестування, перелік використаних програмних засобів, огляд сучасних методів виділення ключових ознак, механізми побудови семантичних представлень, адаптивна фільтрація ознак.

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

1. Огляд методів класифікації зображень із використанням вагових коефіцієнтів.

2. Математичні моделі семантичного аналізу ознак зображень.

3. Огляд методів генерації семантичних ознак для класифікації зображень.

4. Аналіз методів адаптивного фільтрування ознак для покращення точності.

5. Математичні моделі комбінування семантичного аналізу та вагових коефіцієнтів.

6. Можливості поєднання локальних і глобальних ознак для підвищення точності класифікації.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) дескриптори ключових точок, постановка задачі, схеми роботи редуції ознак, зображення навчальної вибірки, схеми застосування семантичних ознак.

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

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

### КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	26.11.2024	
2	Аналіз завдання, підбір літератури	26.11.24-01.12.24	
3	Аналіз літератури з досліджуваної проблеми	02.12.24-08.12.24	
4	Аналіз технічних засобів	09.12.24-13.12.24	
5	Розробка комп'ютерної моделі	14.12.24-18.12.24	
6	Програмна реалізація	19.12.24-26.12.24	
7	Оформлення пояснювальної записки	27.12.24-28.12.24	
8	Перевірка на плагіат	29.12.2024	
9	Рецензування	30.12.2024	
10	Підготовка презентації та доповіді	01.01.2025	
11	Занесення роботи в електронний архів	12.01.2025	
12	Попередній захист кваліфікаційної роботи	13.07.2025	

Дата видачі завдання 25 листопада 2024 р.

Студент \_\_\_\_\_  
(підпис)

Керівник роботи \_\_\_\_\_  
(підпис)

\_\_\_\_\_ доц. Руденко Д.О.  
(посада, прізвище, ініціали)

## РЕФЕРАТ/ABSTRACT

Пояснювальна записка до кваліфікаційної роботи: 73 с., 2 табл., 7 рис., 42 джерело.

МЕТОДИ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ, ДЕСКРИПТОР, КЛЮЧОВА ТОЧКА, ІНФОРМАТИВНІСТЬ, ДЕТЕКТОР BRISK, ВАГОВІ КОЕФІЦІЄНТИ, СЕМАНТИЧНІ ОЗНАКИ, КЛАСИФІКАЦІЯ ЗОБРАЖЕНЬ, ДЕТЕКТОР SURF, РЕЛЕВАНТНІСТЬ ОПИСІВ, ВІДСТАНЬ ХЕММІНГУ.

Об'єктом дослідження є методи та моделі класифікації зображень із використанням вагових коефіцієнтів та семантичних ознак.

Метою дослідження є розробка підходів, що поєднують вагомості та семантичних ознак для підвищення точності класифікації зображень.

У рамках дослідження використано та проаналізовано методи класифікації зображень на основі вагових коефіцієнтів та семантичних ознак. Розроблено алгоритм, який дозволяє адаптувати вагові коефіцієнти для динамічного фокусування на релевантних ознаках. Проведено програмну реалізацію експериментальної системи для класифікації зображень із застосуванням запропонованого підходу.

У результаті дослідження створено інструмент для точнішого визначення класів об'єктів у зображеннях, що демонструє високу ефективність у задачах класифікації при змінних умовах контексту.

IMAGE RECOGNITION METHODS, DESCRIPTOR, KEY POINT, INFORMATIVENESS, BRISK DETECTOR, WEIGHT COEFFICIENTS, SEMANTIC MARKINGS, IMAGE CLASSIFICATION, SURF DETECTOR, RELEVANCE OF DESCRIPTIONS, HEMMING DISTANCE.

The object of the research is methods and models of image classification using weighting factors and semantic features.

The aim of the research is to develop approaches combining weights and semantic features to increase the accuracy of image classification.

Within the framework of the study, image classification methods based on weighting factors and semantic features were used and analyzed. An algorithm has been developed that allows you to adapt weighting factors for dynamic focusing on relevant features. The software implementation of the experimental system for image classification using the proposed approach was carried out.

As a result of the study, a tool was created for more accurate determination of object classes in images, which demonstrates high efficiency in classification tasks under variable context conditions.

## ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів .....	6
Вступ.....	7
1 Аналіз сучасних методів розпізнавання зображень .....	8
1.1 Основні поняття при побудові характерних ознак.....	8
1.2 Аналіз методів для формування ключових точок .....	14
1.3 Аналіз структурних описів .....	18
1.4 Постановка задачі .....	19
2 Використання вагомостей та семантичних ознак для зображень .....	20
2.1 Коефіцієнти вагомості для елементів опису .....	20
2.2 Коефіцієнти вагомості для елементів опису .....	26
2.3 Застосування семантичних ознак для класифікації зображень ...	28
2.4 Комбінування методу інформативності та семантичних ознак.....	34
2.5 Результативність та швидкість при використанні редукції ознак .	38
3 Дослідження результатів реалізації аналізу зображень .....	40
3.1 Обґрунтування реалізації розробленого алгоритму.....	40
3.2 Програмна реалізація.....	41
3.3 Експериментальні результати та аналіз.....	44
3.4 Виокремлення семантичних ознак та їх оцінка.....	57
3.5 Практичне застосування результатів дослідження .....	65
Висновки .....	68
Перелік джерел посилання .....	69

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

КО – ключові ознаки

КТ – ключова точка

ДО – дескриптор ознаки

SIFT – Scale-Invariant Feature Transform (інваріантне до масштабу перетворення ознак)

SURF – Speeded-Up Robust Features (прискорені стійкі ознаки)

BRISK – Binary Robust Invariant Scalable Keypoints (бінарні стійкі масштабовані ключові точки)

DoG – Difference of Gaussians (різниця гауссіан)

ХО – характеристичні описи

LoG – Laplacian of Gaussian (лапласіан гауссіан)

NM – нормалізовані метрики

РГ – розподіл градієнтів

ЕВ – евклідова відстань

CNN – Convolutional Neural Network (згорткова нейронна мережа)

ВП – ваговий параметр

ЛПП – локально-просторове представлення

## ВСТУП

Сучасні технології комп'ютерного зору охоплюють широкий спектр завдань, серед яких виявлення, класифікація та розпізнавання об'єктів на зображеннях. Одним із ключових аспектів цих завдань є формування змістовної та впорядкованої інформації, що відображає сутність візуального об'єкта. Для цього використовуються різноманітні методи та алгоритми, які дозволяють виділяти ключові точки та створювати їх дескриптори.

Ключові точки є основою для побудови структурованих описів зображень, що дозволяють зменшити обсяг даних, необхідних для обробки, та забезпечити високу стійкість до змін масштабу, поворотів і освітлення. Формування дескрипторів ключових точок дає можливість зіставляти об'єкти між різними зображеннями та виконувати їх класифікацію навіть за умов складного фону або наявності шумів.

Актуальність дослідження зумовлена потребою підвищення точності та стійкості алгоритмів класифікації зображень, які мають значення для широкого кола практичних застосувань: від систем відеоспостереження та медичної діагностики до автоматичного керування транспортом і робототехніки.

# 1 АНАЛІЗ СУЧАСНИХ МЕТОДІВ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

Розпізнавання образів та зображень останнім часом набуло великого значення завдяки швидкому розвитку цифрових технологій та необхідності ефективного аналізу візуальних даних. Щодня генерується величезна кількість зображень у різних сферах, таких як медицина, транспорт, соціальні мережі та безпека. Щоб обробити ці дані, необхідні технології, які здатні автоматично і точно визначати об'єкти, структури чи особливості зображень.

Один із ключових методів розпізнавання зображень – це використання дескрипторів ключових точок. Ці алгоритми дозволяють виділити особливі точки на зображенні, які залишаються стабільними незалежно від масштабу, повороту чи освітлення. Це дозволяє ефективно порівнювати та знаходити подібні об'єкти на різних зображеннях.

Інші класичні методи обробки зображень, наприклад, аналіз градієнтів або виділення контурів, також залишаються популярними завдяки своїй простоті та надійності. Вони дозволяють швидко і точно визначати межі об'єктів та їх форму. Це особливо важливо в таких сферах, як автоматизовані системи контролю якості, де необхідно визначати відхилення у виробничих процесах. Значення класичних методів полягає в їх надійності, простоті реалізації та відносно низьких вимогах до обчислювальних ресурсів. Це робить їх популярними для різних практичних завдань, де точність і швидкість є ключовими факторами.

## 1.1 Основні поняття при побудові характерних ознак

Постає складне завдання отримання змістовної та впорядкованої інформації з зображення. Для його вирішення можна застосовувати технології, засновані на аналізі особливостей зображення в окремих точках.

Аналізуючи характеристики обмеженої кількості ключових ознак (КО), можна вирішити практично будь-які практичні завдання.

Система характерних ознак для розпізнавання зображень відзначається кількома перевагами, зокрема спрощенням процесу формування ознак, скороченням їх кількості, можливістю універсального структурного аналізу об'єктів, а також високою стійкістю до спотворень фону та помилок.

Ключові точки на зображенні є важливими елементами, оскільки вони містять найбільш характерну та змістовну інформацію, яка використовується для подальшої обробки, порівняння або класифікації. Ключова точка визначається як область на зображенні, що виділяється своєю особливою структурою або характеристиками. Наприклад, це можуть бути кути, крайові елементи чи текстурні особливості, які залишаються стабільними незалежно від того, як змінюється саме зображення – чи то в результаті збільшення або зменшення масштабу, зміни кута огляду або умов освітлення [1].

Основною перевагою формування ключових точок є їх стійкість до цих трансформацій. Алгоритми, які використовують ключові точки, здатні розпізнавати один і той самий об'єкт на різних зображеннях, навіть якщо ці зображення зроблені під різними кутами або при різних умовах. Це дає змогу алгоритмам працювати з візуальними даними, які можуть змінюватися, зберігаючи при цьому точність розпізнавання.

Ще однією важливою перевагою є те, що такі алгоритми дозволяють значно зменшити обсяг даних для обробки. Замість того, щоб аналізувати ціле зображення піксель за пікселем, система може сконцентруватися на ключових точках. Це суттєво прискорює обчислення і дозволяє ефективно обробляти навіть великі зображення, що є важливим у багатьох реальних застосуваннях, де швидкість аналізу має вирішальне значення.

Крім того, ключові точки зосереджують інформацію про найважливіші елементи зображення, такі як контури об'єктів чи текстурні особливості. Це дозволяє алгоритмам ефективно розпізнавати об'єкти або їхні частини, навіть якщо фон зображення змінюється або додається шум. Така незалежність від

фону робить ці алгоритми особливо корисними в складних умовах, наприклад, у системах відеоспостереження чи розпізнавання обличь.

Загалом, формування ключових точок є ефективним методом обробки зображень, який дозволяє досягати високої точності при розпізнаванні об'єктів. [–] Це забезпечує надійне розв'язання багатьох практичних завдань, таких як аналіз медичних зображень, автоматичне керування транспортом, або навіть система безпеки.

Процес утворення дескрипторів ключових точок починається з першого важливого етапу – виявлення ключових точок на зображенні за допомогою детектора ключових точок. Детектор ключових точок використовується для знаходження найбільш виразних і стійких точок на зображенні, які будуть зберігати свою характерну інформацію незалежно від змін масштабу, поворотів або освітлення. Прикладами таких детекторів є алгоритми SIFT (Scale-Invariant Feature Transform) та SURF (Speeded-Up Robust Features) [3–4], які шукають кути, межі або текстурні особливості зображення.

На першому етапі детектор аналізує зображення, знаходячи області, де зміни інтенсивності пікселів є найбільш значущими. Це можуть бути кути, контури або текстурні області, які визначаються шляхом аналізу градієнтів інтенсивності. Ключові точки повинні бути такими, що їх можна буде надійно використовувати в наступних етапах обробки зображення, незалежно від умов зйомки.

Після того, як ключові точки виявлені, відбувається формування дескрипторів – це другий етап. Дескриптори – це вектори або набори числових значень, які описують локальні властивості навколо кожної ключової точки. Вони узагальнюють інформацію про текстуру, орієнтацію, градієнти яскравості та інші характеристики зображення в околі ключової точки. Ці дескриптори дозволяють порівнювати ключові точки між різними зображеннями і знаходити відповідності між ними.

Формування дескрипторів ґрунтується на аналізі певного радіусу навколо кожної ключової точки, де обчислюються локальні характеристики (наприклад, напрям градієнтів). Ці характеристики кодуються у вигляді векторів фіксованої довжини, що забезпечує можливість швидкого та ефективного порівняння точок між різними зображеннями.

Отже, процес починається з виявлення ключових точок, а потім продовжується їх описом через дескриптори, що забезпечує можливість надійного розпізнавання та аналізу об'єктів на зображенні.

Математичний опис процесу визначення дескрипторів ключових точок включає кілька кроків, зокрема обчислення градієнтів навколо ключових точок і формування векторів, які описують ці характеристики. Розглянемо це на прикладі алгоритму SIFT (Scale-Invariant Feature Transform), де використовується аналіз локальних градієнтів інтенсивності зображення для формування дескриптора [5].

На першому етапі визначаються ключові точки шляхом аналізу масштабно-інваріантних ознак. Наприклад, у SIFT для пошуку таких точок використовують метод Лапласіана Гаусса (LoG), який апроксимується через різницю гауссіан (*DoG*):

$$DoG(x, y, \sigma) = L(x, y, k\sigma) - G(x, y, \sigma) \times I(x, y) \quad (1.1)$$

де  $I(x, y)$  – це значення пікселів зображення;

$G(x, y, \sigma)$  – це гауссове ядро розмиття;

$\sigma$  – масштабний параметр.

Ключові точки виявляються як екстремуми цієї різниці гауссіан, тобто ті точки, де значення *DoG* є локальними максимумами або мінімумами у просторі та масштабі [6].

Для кожної ключової точки обчислюються градієнти інтенсивності в околі цієї точки.

Градiєнти визначаються як частковi похiднi iнтенсивностi зображення:

$$G_x(x, y) = \frac{\partial I(x, y)}{\partial x}, G_y(x, y) = \frac{\partial I(x, y)}{\partial y}$$

в той час як магнiтуда та напрям

градiєнта в кожнiй точцi обчислюються за формулами:

$$M(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \text{ та } \theta(x, y) = \text{atan2}(G_x(x, y), G_y(x, y)).$$

Пiсля обчислення градiєнтів в околі ключової точки будується гiстограма орієнтацій градiєнтів. Окіл ключової точки розбивається на декілька менших блоків (наприклад, на 4×4 комiрки), i для кожної комiрки формується гiстограма напрямів градiєнтів.

Для кожної комiрки визначається напрямок градiєнтів та їх магнiтуда, i цi значення записуються в гiстограму. Якщо розглядається, наприклад, 8 напрямків (секторів) у гiстограмі, то дескриптор для однієї комiрки може бути вектором з 8 елементів. Весь дескриптор для околу ключової точки формується з набору таких векторів, для 4×4 комiрок отримуємо вектор розміром 4×4×8=128.

Для того щоб зробити дескриптор iнваріантним до змін освiтлення, отримані значення нормалізують. Це робиться шляхом нормалізації вектора дескриптора:

$$d_{norm} = \frac{d}{\|d\|} \quad (1.2)$$

де  $d$  – вектор дескриптора;

$\|d\|$  – його евклідова норма.

Отже, дескриптор ключової точки формується на основі локальних градiєнтів iнтенсивностi зображення навколо цієї точки. Він представляє собою нормалізований вектор, який описує магнiтуди та напрямки градiєнтів у певному радіусі навколо ключової точки [7]. Такий підхід дозволяє

ефективно кодувати візуальну інформацію і забезпечувати інваріантність до різних трансформацій зображення.

Продемонструємо поняття дескрипторів на прикладі: двовимірну функцію зображення  $V(x, y)$  представимо у вигляді кінцевої множини фрагментів  $\{b_k(x, y)\}$  або як кінцеву множину  $\Lambda = \{\lambda^k\}$  ХО. При цьому з кожним  $\lambda^k$  пов'язано елемент  $c^k = (x^k, y^k)$  множини координат  $C = \{c^k\}$ , який визначає просторові властивості опису [8]. В результаті візуальний об'єкт представляється у вигляді сукупності дескрипторів (рисунок 1.1), що включають атрибути  $\lambda^k$  та координати  $c^k$ .

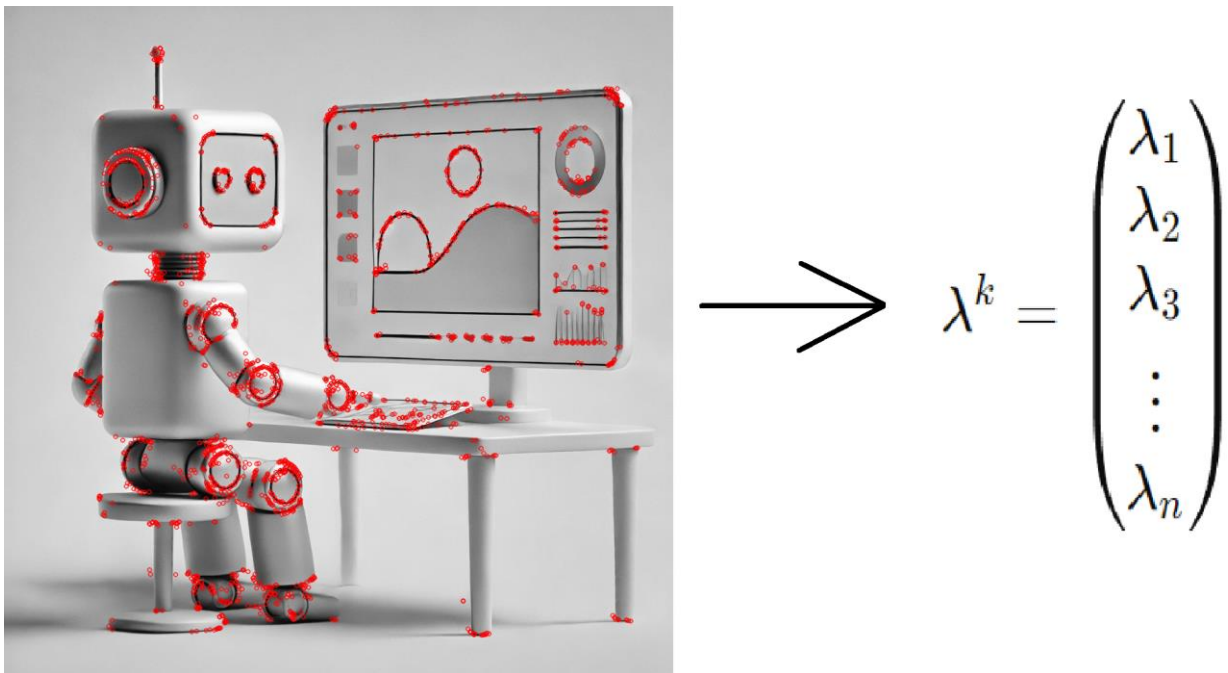


Рисунок 1.1 – Побудова опису у вигляді множини характерних ознак

Зазвичай формування множини ключових точок (КТ) зображення відображає зміни кривизни функції яскравості та ступінь відмінності у порівнянні з прилеглими фрагментами [9]. Різні алгоритми виявлення КТ можуть бути спрямовані на визначення кутів, різких переходів, меж, контрастних точок, ліній та інших елементів, залежно від застосованого математичного підходу.

Перехід до множини характеристичних описів (ХО) зображення дозволяє значно зменшити обсяг даних, зберігаючи при цьому високу точність та стійкість до шумів. Це досягається шляхом виділення та аналізу нової інформації, що міститься у структурних елементах і їх взаємозв'язках. У більшості випадків ХО представляються як числовий вектор з фіксованою розмірністю та дійсними значеннями  $\lambda^k = (\lambda_1, \dots, \lambda_n)^k$ ,  $\lambda_i \in R^1, i = \overline{1, n}$ .

## 1.2 Аналіз методів для формування ключових точок

У цій кваліфікаційній роботі будуть проаналізовані різні алгоритми виявлення ключових точок і побудови дескрипторів зображень, кожен з яких має свої переваги та недоліки. Особливу увагу буде приділено вивченню структури даних для множини дескрипторів, що є важливим для ефективної роботи систем розпізнавання. Усі алгоритми можна поділити на дві основні групи залежно від способу представлення дескрипторів: дескриптори з плаваючою точністю (найпоширеніші алгоритми: SURF, SIFT, CSLBP тощо) та бінарні дескриптори (найвідоміші алгоритми: BRIEF, BRISK, AKAZE, ORB та інші).

SIFT (Scale-Invariant Feature Transform) є одним із ключових алгоритмів для комп'ютерного зору, що застосовується для виявлення та опису ключових точок на зображеннях. Його важливість полягає в здатності забезпечувати інваріантність до змін масштабу, повороту та освітлення. Такий підхід особливо цінний у задачах, де необхідно надійно розпізнавати об'єкти або зіставляти зображення у різних умовах. Детальний аналіз методу допомагає зрозуміти, як алгоритм працює на різних етапах і які аспекти роблять його настільки потужним інструментом для обробки зображень.

Процес роботи методу починається з виявлення ключових точок у масштабно-просторовій області. Для цього будується масштабна піраміда за допомогою ряду зображень, що згладжуються гауссовими фільтрами з

різними коефіцієнтами. Це дозволяє виявляти потенційні ключові точки на різних рівнях деталізації. Після побудови різниці гауссіан (DoG), що підкреслює місця з різким перепадом яскравості, алгоритм знаходить екстремуми – мінімуми та максимуми, які вказують на наявність ключових точок.

Після виявлення ключових точок необхідно їх точно локалізувати. Це здійснюється через фітінг тривимірного квадратичного полінома, який уточнює координати точок і допомагає відсіяти менш стійкі точки, наприклад, ті, що знаходяться поблизу країв або мають низький контраст. Завдяки цьому етапу метод виявляє лише найбільш значущі та стійкі ключові точки, що забезпечують надійність подальшого аналізу зображення.

Щоб забезпечити інваріантність до повороту, алгоритм призначає кожній ключовій точці певну орієнтацію як на (рис. 1.2). Для цього аналізуються градієнти яскравості в околі точки, на основі яких будується гістограма орієнтацій.

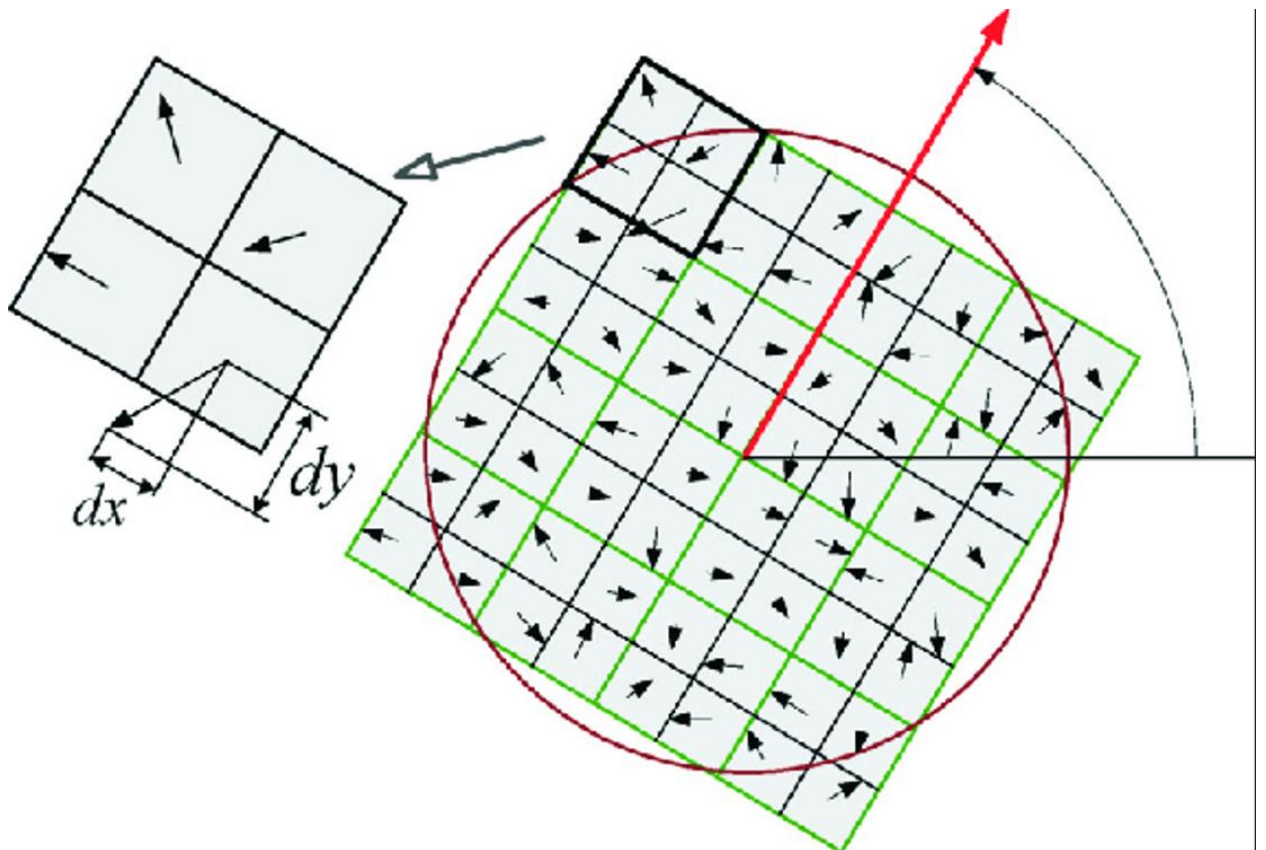


Рисунок 1.2 – Шаблон роботи SURF

Основна орієнтація визначається максимальним значенням у гістограмі, але якщо є інші локальні піки, можуть бути присвоєні додаткові орієнтації. Це робить метод більш стійким до різноманітних варіацій зображення.

Після визначення орієнтації формується дескриптор для кожної ключової точки. Дескриптор будується на основі локальних градієнтів у невеликому околі навколо точки, де кожен підблок поділяється на кілька частин, і для кожного підблоку будується своя гістограма орієнтацій. Після цього вектори значень цих гістограм об'єднуються в єдиний вектор розмірністю 128 елементів, який і є кінцевим дескриптором.

Завершальний етап методу SIFT полягає у зіставленні дескрипторів між різними зображеннями. Це дозволяє зіставляти ключові точки і використовувати їх для таких задач, як розпізнавання об'єктів або відстеження руху. Для порівняння дескрипторів застосовуються різні метрики відстані, найчастіше – Евклідова відстань, що дає змогу знайти найближчі ключові точки між зображеннями [10–12].

Метод SIFT має низку важливих переваг, які роблять його одним із найпопулярніших алгоритмів для обробки зображень і має набір суттєвих переваг:

- інваріантність до масштабу і повороту. Одна з ключових переваг методу полягає в тому, що він може точно визначати та зіставляти ключові точки, навіть якщо об'єкти на зображеннях мають різні масштаби або обернуті на певний кут. Це дозволяє використовувати алгоритм для аналізу зображень, зроблених з різних перспектив або при різних налаштуваннях камери;

- стійкість до змін освітлення. Завдяки тому, що метод використовує градієнти яскравості для формування дескрипторів, він є стійким до змін у освітленні або контрасті. Це особливо важливо для реальних сценаріїв, де умови освітлення можуть змінюватися, наприклад, під час зйомок на вулиці чи в приміщенні з нестабільним освітленням;

- висока точність зіставлення. SIFT створює детальні дескриптори розміром 128 елементів, що дозволяє точніше описати локальні особливості зображення. Це дає змогу більш надійно зіставляти точки між зображеннями, навіть якщо вони сильно відрізняються за складністю або деталізацією;
- можливість виявлення великої кількості ключових точок. Метод дозволяє виявляти багато ключових точок у зображенні, що може бути корисно для складних або високодеталізованих зображень. Це забезпечує більше можливостей для точного зіставлення та аналізу;
- гнучкість у різних додатках. SIFT є універсальним інструментом, який може бути застосований до різноманітних задач комп'ютерного зору: розпізнавання об'єктів, 3D-реконструкція, побудова карт, виявлення та відстеження об'єктів. Його гнучкість дозволяє використовувати його у багатьох різних контекстах і середовищах.

Ці переваги роблять SIFT популярним вибором для дослідників і практиків, які працюють із комп'ютерним зором та аналізом зображень.

Дескриптор у методі SIFT будується на основі аналізу градієнтів яскравості в околі ключової точки. Він являє собою вектор, який містить гістограми напрямків градієнтів у невеликих підблоках навколо ключової точки. Зазвичай дескриптор має розмір 128 елементів. Для кожної ключової точки будується область, яка ділиться на 16 підблоків (4x4 сітка), і для кожного підблоку обчислюється гістограма напрямків градієнтів. Кожна гістограма має 8 напрямків, що відповідають різним кутам орієнтації градієнта. Кожен елемент дескриптора DDD обчислюється за допомогою нормалізації градієнтів для забезпечення стійкості до змін освітлення. Формула для обчислення елемента дескриптора зокрема виглядає так:

$$h_k = \sum_{p \in B_k} m(p) \cdot \exp\left(-\frac{\|p - p_0\|^2}{2\sigma^2}\right) \times \delta(\theta(p) - \theta_k) \quad (1.3)$$

де  $m(p)$  – величина градієнта в точці  $p$ ;

$p_0$  – центр підблоку;

$\sigma$  – масштаб для гауссового згладжування;

$\delta(\theta(p) - \theta_k)$  – дельта-функція, до якого напрямку  $\theta_k$  орієнтація гістограми належить градієнт у точці  $p$ .

Кінцевий дескриптор є вектором, що включає суми вагових градієнтів для кожного з напрямків у кожному з підблоків. Завдяки такому підходу дескриптор стає стійким до змін масштабу, повороту та освітлення, що дозволяє точно зіставляти ключові точки між зображеннями.

### 1.3 Аналіз структурних описів

Одним із перспективних напрямів удосконалення традиційних методів і розробки нових підходів до зіставлення візуальних даних є використання локально-просторового представлення зображення. Під час аналізу зображення воно перетворюється на набір локальних функцій або ознак, визначених на підобластях поля зору. Таке представлення є більш універсальним порівняно з цілісним описом, оскільки дозволяє аналізувати окремі частини об'єкта, навіть якщо вони зазнали спотворень через просторові шуми, і веде до ефективних рішень, забезпечуючи стійкість до цих перешкод.

Удосконалення інформаційних технологій для класифікації у просторі описів, що базується на множинах дескрипторів ключових точок зображень, та оцінка ефективності розпізнавання на прикладних зразках є перспективними напрямами розвитку структурного розпізнавання. Застосування таких методів дозволяє значно скоротити час обробки зображень, забезпечуючи високу точність навіть за умов змішаного фону чи наявності шумів. Це відкриває нові можливості для реалізації ефективних рішень у задачах розпізнавання об'єктів у реальному часі.

## 1.4 Постановка задачі

Використання локально-просторового представлення зображень є актуальним напрямом у комп'ютерному зорі, оскільки дозволяє підвищити стійкість алгоритмів розпізнавання до спотворень і перешкод. Задається мета розроблення методів та алгоритмів, які реалізують класифікацію візуальних даних на основі локальних дескрипторів, забезпечуючи стійкість до просторових перешкод. На практиці необхідно застосувати локально-просторові дескриптори для покращення ефективності розпізнавання на прикладних зразках.

Об'єктом дослідження є методи відбору дескрипторів за показниками унікальності та інформативності для різних типів алгоритмів класифікації зображень.

Метою роботи є створення моделі відбору дескрипторів на основі їх унікальності та інформативності, що дозволить підвищити точність розпізнавання та класифікації зображень у різних алгоритмах.

Для досягнення мети необхідно вирішити такі завдання:

- провести аналіз існуючих методів локально-просторового представлення зображень та їх недоліків;
- розробити модель для класифікації на основі локальних дескрипторів з урахуванням просторових перешкод;
- реалізувати алгоритм у програмному застосунку для розпізнавання об'єктів на базі зображень;
- провести порівняльний аналіз результатів роботи моделі з різними підходами та оцінити її ефективність і продуктивність.

## 2 ВИКОРИСТАННЯ ВАГОМОСТЕЙ ТА СЕМАНТИЧНИХ ОЗНАК ДЛЯ ЗОБРАЖЕНЬ

### 2.1 Коефіцієнти вагомості для елементів опису

Для класифікації зображень із комбінуванням вагомостей та семантичних ознак можна застосовувати різноманітні моделі, які поєднують обробку візуальних елементів та відносних значущостей.

Deep Neural Networks (DNN) надають можливість виявляти багатопарові семантичні залежності. Використовуючи вагомості для різних ознак, такі мережі можуть адаптувати класифікацію, виділяючи найбільш значущі елементи для кожного об'єкта та роблячи класифікацію точнішою [13].

Random Forest використовує ансамбль дерев рішень для класифікації, але в даному підході дерева можуть бути налаштовані так, щоб акцентувати вагу на ключових семантичних ознаках. Це дозволяє моделі швидко виділяти важливі елементи, підвищуючи її ефективність для об'єктів із комплексною структурою.

Convolutional Neural Networks (CNN) здатні виявляти глибокі семантичні ознаки – від простих контурів до складних текстур і форм. Завдяки багаторівневій структурі, CNN дозволяють розподіляти вагомості між ознаками, підсилюючи їхню важливість відповідно до контексту та характерних особливостей класу.

Support Vector Machine (SVM) підходить для розмежування класів через гіперплощини, орієнтуючись на вагомості семантичні ознаки, які перетворено в числові вектори. Під час тренування модель навчається пріоритизувати вагу ознак, що допомагає створити чітке розмежування між класами навіть при складних або близьких за змістом об'єктах [14].

Binary Robust Invariant Scalable Keypoints (BRISK) дозволяє детектувати та описувати ключові точки, виділяючи стійкі семантичні ознаки

для класифікації. Вагомості, застосовані до цих точок, підсилюють значення об'єктів у зображенні, що є корисним при умовах змін масштабу або обертання.

Поєднання вагомостей і семантичних ознак в цих моделях забезпечує підвищену гнучкість і точність, дозволяючи краще адаптувати класифікацію до реальних умов та уникати помилок, коли ознаки є подібними або контекст зображення змінюється.

Коефіцієнти вагомості – це числові значення, які використовуються для того, щоб визначити значущість окремих ознак у процесі класифікації зображень. Вони допомагають моделі сфокусуватися на найбільш релевантних ознаках, що мають ключове значення для точного розпізнавання та ідентифікації об'єктів. Далі розглянемо особливості та різні аспекти використання вагових коефіцієнтів.

Значення вагових коефіцієнтів можуть мати різне значення в залежності від типу ознак. Наприклад, для ознак, які є важливими для ідентифікації об'єкта (таких як контури обличчя для розпізнавання особи), встановлюється вища вага, тоді як менш значущі деталі (фон або колір) можуть мати нижчі вагові коефіцієнти.

Навчання вагомостей в глибоких нейронних мережах, таких як CNN, вагові коефіцієнти автоматично налаштовуються під час навчання моделі. Алгоритм зворотного поширення помилки (backpropagation) налаштовує ваги так, щоб мінімізувати похибку між передбаченням та істинною меткою. Завдяки цьому процесу модель поступово «вчиться» асоціювати більші вагові коефіцієнти з ознаками, що мають більшу значущість для правильної класифікації [15].

У деяких системах вагові коефіцієнти можуть змінюватись в залежності від контексту зображення. Наприклад, у класифікації транспортних засобів на дорозі, коефіцієнти ваги для ознак, що стосуються автомобіля, будуть вищими, якщо на зображенні є дорога, але знизяться,

якщо фон містить природні ландшафти. Це дозволяє моделі динамічно змінювати фокус на більш релевантні деталі.

Іноді вагові коефіцієнти можуть бути встановлені вручну на основі знань домену. Це часто роблять у менш складних моделях або в алгоритмах, де автоматичне навчання ваг є непрактичним. Наприклад, у медичних зображеннях коефіцієнти ваги можуть бути збільшені для ознак, що відповідають за виявлення потенційних патологій.

Значення вагових коефіцієнтів у ансамблевих моделях, таких як Random Forest або Gradient Boosting, вагові коефіцієнти можуть використовуватися для комбінування результатів кількох моделей. Кожна модель (або дерево) отримує власну вагу, яка вказує на її важливість в остаточному рішенні. Це дозволяє зменшити вплив менш точних моделей і посилити значення результатів більш точних [16–17].

Ваги в контексті Transfer Learning (переносного навчання) модель часто бере вагові коефіцієнти з іншої, попередньо натренованої моделі, а потім налаштовує їх під конкретні завдання. Це дозволяє зберегти загальні ваги для базових ознак (наприклад, форми та текстури), а також адаптувати нові ваги для специфічних класів об'єктів.

Ваги для специфічних ознак у мультимодальних моделях, які обробляють кілька типів даних одночасно (наприклад, зображення та текст), вагові коефіцієнти допомагають поєднувати інформацію з різних джерел. У такому випадку модель використовує різні ваги для семантичних ознак з кожного типу даних, що дозволяє комбінувати інформацію та підвищити точність класифікації.

Регуляризація вагових коефіцієнтів є важливою технікою для запобігання перенавчанню, зменшуючи надмірно високі ваги для окремих ознак. Завдяки регуляризації модель стає менш чутливою до незначних змін у даних, а отже, може краще узагальнювати закономірності.

Використання ваг у функціях втрат (loss functions) деякі функції втрат використовують ваги для ознак або класів, щоб зробити модель більш

чутливою до рідкісних класів або складних випадків. Наприклад, у класифікації медичних зображень можна застосовувати вищі вагові коефіцієнти для рідкісних патологій, щоб забезпечити їхнє правильне виявлення, навіть якщо таких зображень значно менше в датасеті.

Прикладами застосування вагових коефіцієнтів може бути, наприклад, аналіз обличчя: встановлення вищих ваг для ознак очей, рота і носа; аналіз дорожніх знаків – підсилення ваг ознак для круглих форм, червоного кольору тощо або класифікація текстур: вагові коефіцієнти можуть підсилювати значення для повторюваних візерунків та фактур, важливих для класифікації матеріалів [18].

Вагові коефіцієнти є потужним інструментом, що дозволяє моделі адаптувати свою увагу до найбільш значущих аспектів зображення. Розумне використання вагових коефіцієнтів в класифікації зображень дозволяє моделі швидко виділяти важливі ознаки, адаптуватися до контексту та досягати високої точності в задачах розпізнавання.

Коефіцієнти вагомості ( $\omega$ ) – це числові значення, які використовуються для масштабування важливості ознак або параметрів у математичній моделі. Їх використання дозволяє моделі фокусуватися на найбільш релевантних ознаках, що має вирішальне значення для класифікації зображень. Ось кілька ключових аспектів та математичних формул, що описують використання вагових коефіцієнтів [19].

У лінійній моделі ваговий коефіцієнт  $\omega_i$  використовується для масштабування важливості кожної ознаки  $x_i$  у рівнянні. Лінійна модель для передбачення  $y$  може бути записана так:

$$y = \omega_0 + \omega_1 x_1 + \omega_2 x_2 + \dots + \omega_n x_n \quad (2.1)$$

де  $\omega_0$  – зсув (bias);

$\omega_i$  – ваговий коефіцієнт для ознаки  $x_i$ ;

$x_i$  – значення  $i$ -ї ознаки.

Вагові коефіцієнти  $\omega_i$  визначають, наскільки сильно кожна ознака впливає на підсумковий результат. Якщо  $\omega_i$  має високе значення, ознака  $x_i$  має значний вплив на  $y$ ; якщо  $\omega_i$  близьке до нуля, значення ознаки майже не враховується.

У нейронних мережах ваги використовуються для зв'язків між нейронами. Якщо нейрон приймає  $x_1, x_2, \dots, x_n$  як вхідні значення, його вихід  $z$  визначається як:

$$z = f(\omega_1 x_1 + \omega_2 x_2 + \dots + \omega_n x_n + b) \quad (2.2)$$

де  $f$  – активаційна функція (наприклад, ReLU, Sigmoid);

$\omega_i$  – ваговий коефіцієнт для кожного входу  $x_i$ ;

$b$  – зсув.

Коефіцієнти ваг  $\omega_i$  у нейронній мережі налаштовуються під час навчання моделі з використанням методу зворотного поширення помилки (backpropagation) для мінімізації функції втрат [20].

У класифікаційних задачах функція втрат ( $L$ ) може включати вагові коефіцієнти, що дозволяє фокусуватися на певних класах або ознаках. Наприклад, функція крос-ентропії з вагами для класів записується як:

$$L = - \sum_{i=1}^C \omega_i \cdot y_i \cdot \log(\hat{y}_i) \quad (2.3)$$

де  $C$  – кількість класів;

$\omega_i$  – ваговий коефіцієнт для класу  $i$ ;

$y_i$  – справжня мітка класу  $i$ ;

$\hat{y}_i$  – передбачена ймовірність для класу  $i$ .

Високий  $\omega_i$  підсилює важливість помилок у класі  $i$ , що допомагає при дисбалансі класів у задачі класифікації.

Регуляризація контролює ваги, запобігаючи перенавчанню моделі. Основні методи регуляризації включають  $L1$  та  $L2$  регуляризацію:

$L1$  -регуляризація додає суму абсолютних значень вагових коефіцієнтів до функції втрат:

$$L_{L1} = L + \lambda \sum_{i=1}^n |\omega_i|, \quad (2.4)$$

$L2$  -регуляризація додає суму квадратів вагових коефіцієнтів до функції втрат:

$$L_{L1} = L + \lambda \sum_{i=1}^n \omega_i^2, \quad (2.5)$$

де  $\lambda$  – коефіцієнт регуляризації, що визначає важливість регуляризаційного члена.

Регуляризація допомагає зменшити значення ваг, тим самим знижуючи ймовірність, що модель буде занадто сильно адаптована до тренувальних даних [21].

В ансамблевих моделях, таких як Random Forest або Boosting, ваги також можуть використовуватися для комбінування результатів кількох моделей. Наприклад, у випадку Boosting кожна модель отримує ваговий коефіцієнт  $w_i$ , що залежить від її точності. Остаточне передбачення можна записати як зважену суму передбачень кожної моделі:

$$y = \sum_{i=1}^M \omega_i \cdot h_i(x) \quad (2.6)$$

де  $M$  – кількість моделей;

$\omega_i$  – вага  $i$ -ї моделі;

$h_i(x)$  – передбачення  $i$ -ї моделі для вхідного значення  $x$ .

Використання вагових коефіцієнтів дозволяє адаптувати модель для виділення найбільш важливих ознак або класів, що особливо корисно в задачах класифікації зображень, де окремі елементи мають різну значущість для загальної класифікації.

## 2.2 Коефіцієнти вагомості для елементів опису

Коефіцієнти інформативності та унікальності для елементів опису зображень (дескрипторів) є важливими характеристиками, які допомагають оцінити та оптимізувати текстові описи в контексті використання зображень, особливо для пошукових систем, автоматизованих систем класифікації або інструментів розпізнавання зображень.

Впливові фактори інформативності для дескрипторів:

- точність опису для розуміння того, чи точно дескриптор відповідає об'єктам або характеристикам, що відображені на зображенні. Наприклад, опис «синя машина на фоні лісу» є більш інформативним, ніж просто «машина»;
- деталізація для визначення, наскільки детально описані ключові елементи зображення. Наприклад, дескриптор «червона машина 1960-х років» надає більше корисної інформації, ніж просто «машина»;
- релевантність до контекста, яка відображає, чи є опис важливим для завдання або користувача. Якщо зображення використовується в контексті пошуку автомобілів за роком випуску, то «1960-х років» – критичний дескриптор;
- унікальні ознаки, опис повинен включати ті елементи, які відрізняють це зображення від інших. Наприклад, опис «машина з вінтажним

номерним знаком» може бути дуже інформативним, якщо таких зображень мало.

Коефіцієнт унікальності для дескрипторів – це коефіцієнт, який відображає, наскільки оригінальним і відмінним є опис зображення порівняно з іншими схожими описами. Він вимірює, як часто певні дескриптори повторюються в інших текстах або базах даних зображень і чи вносить опис щось нове [22].

Впливовими факторами унікальності для дескрипторів можуть бути:

- неповторність елементів опису – чим рідше використовується певний дескриптор, тим він є більш унікальним. Наприклад, «унікальне дерево бонсай на фоні заходу сонця» є більш унікальним, ніж «дерево на фоні неба».

- оригінальність формулювання – унікальність також визначається нестандартними або рідкісними мовними зворотами у формулюванні опису. Опис, який вживає не загальноприйняті фрази, а нові або творчі підходи, має вищий коефіцієнт унікальності.

- рідкісні характеристики зображення – якщо зображення містить об'єкти або сцени, які рідко зустрічаються, і вони включені в опис, це підвищує його унікальність. Наприклад, опис «золота скульптура дракона на фоні руїн» є дуже унікальним, якщо подібних зображень мало.

- специфічність до конкретного зображення – опис, що точно відображає специфічні аспекти конкретного зображення, має високий коефіцієнт унікальності. Наприклад, опис «велосипед з ручним гравіюванням на рамі» підвищує унікальність, якщо такі велосипеди рідкісні або індивідуальні.

Відбір дескрипторів з найвищими показниками інформативності є важливим для задач категоризації та класифікації зображень. Високоінформативні дескриптори надають детальну інформацію про конкретні характеристики зображення, що дозволяє більш точно розподілити зображення за відповідними категоріями або класами [23].

Таким чином, відбір дескрипторів із найвищими показниками унікальності та інформативності допомагає створювати оптимальні описи для зображень, покращує точність та ефективність роботи з візуальним контентом і підвищує якість метаданих для різних завдань в ході класифікації, і значно прискорює цей процес.

### 2.3 Застосування семантичних ознак для класифікації зображень

Семантичні ознаки для класифікації зображень – це високорівневі характеристики, які відображають смисловий зміст об'єктів на зображенні та допомагають моделі «розуміти» візуальні патерни більш узагальнено. Ці ознаки використовуються для створення значущих представлень, що можуть асоціюватися з конкретними категоріями або класами, які потрібно класифікувати. Далі розглянемо ключові аспекти та приклади семантичних ознак в задачах класифікації зображень [24].

Семантичні ознаки виходять за межі простих піксельних значень та включають риси, які люди легко розпізнають: форми, структури, кольори та їх поєднання, які описують об'єкти або сцени на зображенні. Наприклад, для класифікації автомобіля як класу, семантичні ознаки можуть включати контури кузова, розташування коліс, форму фар, що створює загальне уявлення про транспортний засіб.

Контекст відіграє важливу роль у семантичній класифікації. Наприклад, модель може краще розпізнати пішохода, якщо візуальний контекст вказує на дорожнє оточення. Таким чином, семантичні ознаки враховують не лише об'єкти, а й взаємодію між ними та їх розташування у відповідному контексті.

Атрибути – це характеристики об'єкта, які можна описати незалежно від його класу. Наприклад, об'єкт може бути «м'яким», «яскравим», «металевим» або «текстурованим». Атрибути можуть слугувати для ідентифікації об'єктів у різних класах або допомагати моделі розрізнити

схожі об'єкти. Наприклад, розпізнання котів та собак може включати атрибути форми вух, структури шерсті тощо.

Структурні семантичні ознаки, такі як симетрія, кути та геометричні форми, допомагають ідентифікувати базові компоненти об'єктів. Наприклад, для розпізнавання будівель модель може шукати прямі кути, горизонтальні лінії, вікна та інші структури, характерні для архітектури.

Семантичні вектори – це представлення об'єктів у формі числових векторів, що кодують смислові відношення. Енкодери використовують такі вектори, щоб моделювати зв'язок між об'єктами. Подібні вектори формують кластери, як на (рис 2.1), які можуть бути використані для більш детального аналізу зображень.

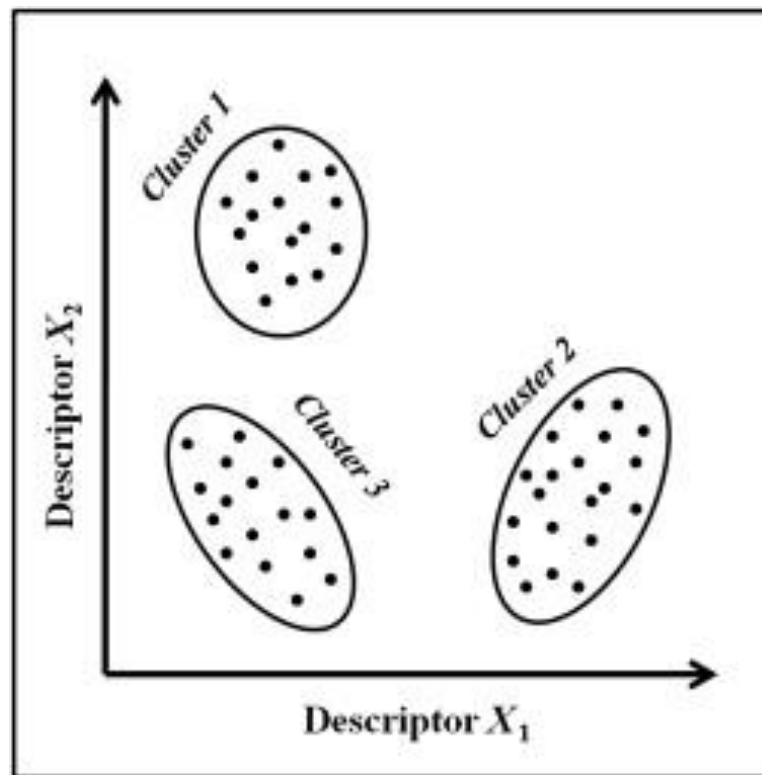


Рисунок 2.1 – Приклад формування кластерів за дескрипторами

Наприклад, в моделях на основі трансформерів семантичні вектори дозволяють зображенням представлятися в багатовимірному просторі, де схожі об'єкти мають близьке розташування.

Семантична сегментація – це процес, у якому кожен піксель на зображенні відноситься до конкретного класу (наприклад, небо, дорога, людина). Сегментація допомагає виділити об'єкти на зображенні відповідно до їхньої семантики, що значно покращує класифікацію, оскільки модель розрізняє не тільки сам об'єкт, але й його оточення та контекст [25].

Семантичні ознаки у CNN витягуються через шари, починаючи від базових візуальних патернів, таких як краї, кути, кольори, і закінчуючи складнішими елементами, як частини об'єктів (очі, вікна) і цілі об'єкти. CNN автоматично адаптуються до виявлення семантичних ознак, використовуючи багатошарові фільтри, що поступово витягують все більш складні особливості.

У таких моделях, як Random Forest чи Boosting, семантичні ознаки агрегуються та комбінуються для прийняття остаточного рішення. Наприклад, ансамблевий підхід може використовувати різні набори ознак для кожного дерева, що дозволяє більш точно класифікувати об'єкти за рахунок комплексного обліку їхніх семантичних особливостей.

В задачах класифікації, що потребують розпізнання матеріалів або текстур (наприклад, класифікація тканин, деревини), семантичні ознаки можуть включати інформацію про фактури та візерунки. Це дозволяє моделі визначати тип матеріалу чи покриття, яке впливає на класифікацію [26].

Прикладами застосування семантичних ознак можуть бути, наприклад, розпізнавання осіб: моделі можуть використовувати семантичні ознаки, такі як форма очей, губ, текстура шкіри, щоб точно ідентифікувати обличчя; або автономні автомобілі (для розпізнавання дорожніх об'єктів, таких як знаки, розмітка, пішоходи, використовується поєднання семантичних ознак, що описують конкретні об'єкти в контексті дороги); медична діагностика з аналізом медичних зображень, як-от рентген чи МРТ, що вимагає виділення семантичних ознак, таких як форми та текстури аномалій для класифікації патологій.

Використання семантичних ознак дозволяє створювати моделі, що краще розуміють об'єкти та їхні зв'язки у зображеннях. Це підвищує точність та надійність класифікації, особливо для складних, багатооб'єктних зображень, де піксельна інформація недостатня для ідентифікації класу об'єкта.

Комбінування методу інформативності та семантичних ознак для класифікації зображень дозволяє посилити модель, виділяючи найбільш релевантні ознаки та адаптуючи їх до контексту зображення. Метод інформативності допомагає визначити найбільш важливі семантичні ознаки, які впливають на класифікацію, зменшуючи вплив менш релевантних характеристик і підвищуючи ефективність обробки.

Семантичні ознаки для класифікації зображень можна математично описати через векторні простори, функції схожості та методи агрегування ознак. Наведемо кілька основних математичних концепцій, що пояснюють семантичні ознаки у контексті класифікації зображень [27–28].

Семантичні ознаки можна представити у вигляді векторів у багатовимірному просторі, де кожен вимір відображає певну характеристику об'єкта (колір, форма, текстура тощо). Якщо у нас є зображення  $I$ , то ми можемо перетворити його у вектор семантичних ознак  $f(I) \in \mathbb{R}^n$ , де  $n$  – кількість семантичних ознак.

$$f(I) = [f_1(I), f_2(I), \dots, f_n(I)] \quad (2.7)$$

де  $f_i(I)$  – значення  $i$ -ї семантичної ознаки для зображення  $I$ .

Кожне  $f_i(I)$  визначає характеристику, яка може бути кількісно виміряна, наприклад, інтенсивність кольору, частота певної текстури або наявність певної геометричної форми.

Для порівняння семантичних ознак двох зображень  $I_1$  та  $I_2$  використовується функція схожості або відстані. Наприклад, косинусна

подібність для векторів семантичних ознак  $f(I_1)$  і  $f(I_2)$ . Чим ближче значення косинусної подібності до 1 (або менша евклідова відстань), тим ближчі за семантикою зображення, що означає схожість їхніх характеристик.

Семантичні ознаки можуть бути агреговані для отримання узагальненого представлення об'єкта. Наприклад, якщо ми маємо набір локальних семантичних ознак для зображення (наприклад, ключові точки зображення), то можна застосувати методи середнього значення або максимізації для їх об'єднання:

$$f_{agg}(I) = \frac{1}{m} \sum_{j=1}^m f_j(I) \quad (2.8)$$

де  $f_j(I)$  – локальні ознаки  $j$ -ї частини зображення;

$m$  – кількість локальних ознак. Це дозволяє створити компактний вектор, що узагальнює семантику всього зображення.

У конволюційних нейронних мережах (CNN) семантичні ознаки витягуються на кожному шарі, що може бути математично описано через набір згорток і функцій активації. Вихід  $z$  нейрону на шарі може бути записаний як:

$$z = f\left(\sum_{k=1}^K \omega_k * x_k + b\right) \quad (2.9)$$

де  $*$  – оператор зортки;

$\omega_k$  – фільтр або вага для ознаки  $k$ ;

$x_k$  – значення ознаки на попередньому шарі;

$b$  – зміщення;

$f$  – активаційна функція (наприклад, ReLU).

Через ієрархічну структуру CNN, нижчі шари витягують базові ознаки (краї, текстури), а вищі – більш абстрактні семантичні ознаки, такі як форми об'єктів або їхні частини.

У семантичній сегментації кожен піксель на зображенні класифікується до певного класу, що можна описати через ймовірнісну функцію  $P(c|x)$ . Для кожного пікселя обчислюється ймовірність приналежності до класу  $c$ :

$$P(c|x) = \frac{\exp(w_c \cdot f(x))}{\sum_{c'} \exp(w_{c'} \cdot f(x))} \quad (2.10)$$

де  $f(x)$  – вектор ознак для пікселя  $x$ ;

$w_c$  – ваговий вектор для класу  $c$ ;

$c$  – клас;

$x$  – ознака пікселя.

Після цього кожному пікселю присвоюється клас із максимальною ймовірністю, що дозволяє сегментувати об'єкти на зображенні.

При навчанні класифікатора вагові коефіцієнти дозволяють визначити значущість кожної семантичної ознаки. Якщо модель класифікації є лінійною, наприклад,  $y = w \cdot f(I) + b$ , то коефіцієнт  $w_i$  показує, наскільки сильно ознака  $f_i(I)$  впливає на ймовірність належності до певного класу. Чим вище  $|w_i|$ , тим значніша ознака для прийняття рішення.

Функція втрат, яка враховує семантичні ознаки, може включати додаткові компоненти, що забезпечують точність класифікації. Наприклад, у функції крос-ентропії для семантичної класифікації:

$$L = -\sum_{i=1}^n y_i \log(\hat{y}_i) + \alpha \sum_{j=1}^m \|f_i - \hat{f}_j\|^2 \quad (2.11)$$

де  $y_i$  – істинна ймовірність класу  $i$ ;

$\hat{y}_i$  – передбачена ймовірність класу  $i$ ;

$\alpha$  – коефіцієнт регуляризації для семантичних ознак;

$f_j$  і  $\hat{f}_j$  – відповідно, справжній і передбачений вектори семантичних ознак.

Таким чином, математичні концепції семантичних ознак допомагають моделі виділити ключові елементи зображення, покращуючи точність класифікації та підвищуючи її здатність «розуміти» зображення на смислового рівні.

#### 2.4 Комбінування методу інформативності та семантичних ознак

Комбінування методу інформативності та семантичних ознак для класифікації зображень дозволяє посилити модель, виділяючи найбільш релевантні ознаки та адаптуючи їх до контексту зображення. Метод інформативності допомагає визначити найбільш важливі семантичні ознаки, які впливають на класифікацію, зменшуючи вплив менш релевантних характеристик і підвищуючи ефективність обробки.

Метод інформативності базується на розрахунку інформаційної значущості кожної ознаки у задачі класифікації. Наприклад, для кожної семантичної ознаки  $f_i$  можна обчислити міру інформативності  $I(f_i)$ , яка показує, наскільки сильно ця ознака сприяє розрізненню класів. Величина інформативності може розраховуватись за допомогою метрик.

Ентропія ( $H$ ), що вимірює ступінь невизначеності або непередбачуваності для ознаки:

$$H(f_i) = -\sum_k P(f_i = k) \log \quad (2.12)$$

де  $P(f_i = k)$  – ймовірність;

$f_i$  - ознака;

$k$  - значення ступеня.

Мета методу інформативності – виділити ті семантичні ознаки, які найбільш суттєво впливають на класифікацію, зводячи до мінімуму невизначеність. Після розрахунку інформативності для кожної семантичної ознаки, можна присвоїти їй ваговий коефіцієнт, пропорційний її значущості для класифікації. Це дозволяє створити зважений вектор ознак, де ваги більш релевантних ознак є вищими, що фокусує модель на значущих характеристиках зображення. Такий підхід допомагає уникнути перенавантаження моделі менш інформативними ознаками та підвищує точність. Після визначення ваг для кожної ознаки можна агрегувати інформативні семантичні ознаки, створюючи підмножину релевантних ознак.

Комбінування методу інформативності та семантичних ознак дозволяє моделі сконцентруватися на найбільш значущих аспектах зображення, що покращує її здатність до класифікації, зменшує похибки та підвищує ефективність обробки. Зображення для використання методу можуть бути елементи згенеровані іншою генеративною моделлю, для порівняння подібності зображень згенерованих за одним промптом. На рисунку 2.2 показано приклад згенерованого зображення.

На рисунку 2.3 представлено модель або робота, на якому червоними точками позначені ключові точки або «інформативні ознаки», які вказують на важливі зони, такі як суглоби, кінцівки, а також взаємодія з елементами інтерфейсу. Ці точки позначають ключові структурні особливості моделі та можуть слугувати візуальними індикаторами для системи комп'ютерного зору, що аналізує положення та рухи цієї моделі.

Кожна точка вказує на важливу інформативну ознаку, яка може бути використана для розпізнавання пози та рухів, зокрема у завданнях аналізу жестів або моделювання поведінки. Редукція ознак є важливим етапом обробки даних, який дозволяє зменшити розмірність простору ознак, залишивши лише найбільш значущі для конкретної задачі [29]. Використання

редукції ознак на основі інформативності та семантичних ознак допомагає оптимізувати процес навчання моделі, знижуючи обчислювальну складність і підвищуючи точність аналізу.



Рисунок 2.2 – Приклад зображення для еталону вибірки

Редукція ознак на основі інформативності відбувається за рахунок того, що за значенням інформативності ознаки метод визначає, наскільки вона корисна для розпізнавання або класифікації об'єктів. Інформативні ознаки є такими, що найбільш чітко характеризують об'єкт, наприклад, певні ключові

точки на зображенні (як у випадку з червоними точками на моделі, що відмічають важливі місця, такі як суглоби).

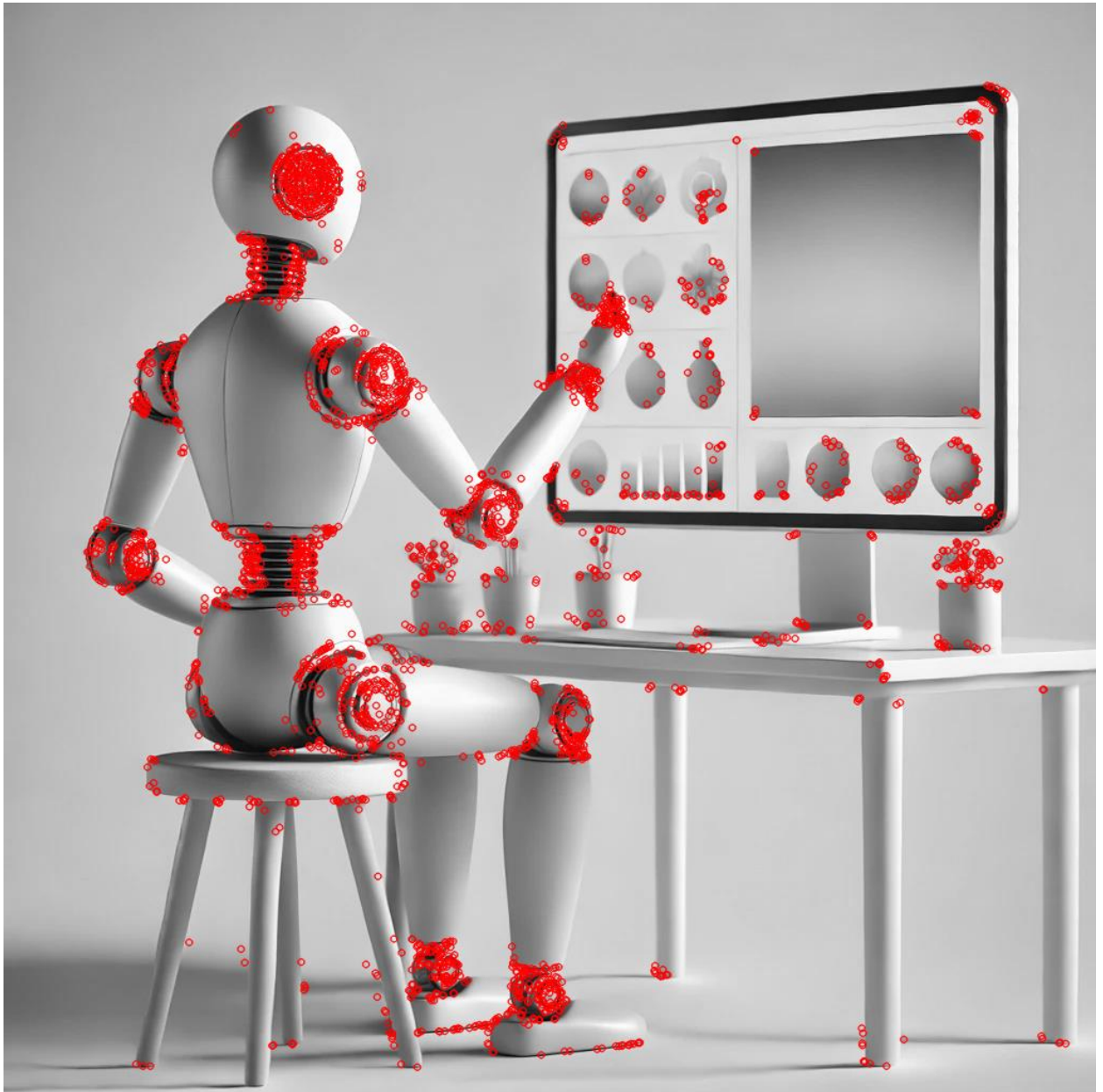


Рисунок 2.3 – Приклад визначених дескрипторів для еталонного зображення

Мета редукції на основі інформативності полягає в тому, щоб залишити лише ті ознаки, які мають найбільший внесок у досягнення точності моделі, та відкинути малозначущі.

Семантичні ознаки відносяться до смислового значення ознаки в контексті конкретної задачі. У багатьох випадках деякі ознаки можуть бути менш інформативними самі по собі, але набувають значення у взаємозв'язку

з іншими ознаками або в певному контексті. Редукція ознак на основі семантики спрямована на зменшення простору ознак за рахунок збереження лише тих, що мають контекстуальне значення для задачі.

## 2.5 Результативність та швидкість при використанні редукції ознак

Використання комбінованої редукції ознак на основі інформативності та семантичних ознак дозволяє значно підвищити результативність аналізу та швидкість обробки даних. Застосування цього підходу знижує обчислювальні витрати і одночасно зберігає або навіть підвищує точність розпізнавання й класифікації завдяки фокусуванню на ключових, контекстуально важливих ознаках.

Основною метою комбінованого підходу є виділення найбільш інформативних і значущих ознак, що безпосередньо впливають на результат розпізнавання або класифікації. Це дозволяє зосередити алгоритм на суттєвих аспектах аналізованих об'єктів або рухів, уникнувши обробки зайвих або малоінформативних ознак.

Оскільки в моделі залишаються тільки релевантні ознаки, система отримує менше шуму, а точність класифікації чи розпізнавання збільшується. Наприклад, в аналізі жестів важливими є саме ті точки, що описують положення рук і тіла, тоді як інші можуть бути ігноровані [30].

Збереження семантично значущих ознак дозволяє системі краще розуміти контекст і значення пози чи руху, підвищуючи її здатність до коректної інтерпретації. Це особливо важливо у випадках, де важливо не тільки виявити об'єкт, але й зрозуміти його дії чи наміри.

Зменшення кількості ознак дозволяє суттєво скоротити обчислювальні витрати, що безпосередньо підвищує швидкість обробки даних. Це важливо, особливо для систем, які працюють у режимі реального часу (наприклад, робототехніка, доповнена реальність, аналіз рухів у спорті чи реабілітації).

Зниження розмірності простору ознак скорочує кількість операцій, що впливає на швидкість аналізу. Чим менша кількість ознак залишається після редукції, тим менше ресурсів потребує модель для їх обробки.

При навчанні алгоритмів машинного навчання на великих наборах даних скорочення кількості ознак може значно прискорити процес навчання. У випадку з глибоким навчанням це знижує кількість параметрів, що потребують оптимізації.

В системах, що працюють у реальному часі (наприклад, віртуальні тренери, системи безпеки на базі комп'ютерного зору), швидкість є критично важливою для зниження обчислювальних витрат у режимі реального часу. Використання редукції ознак дозволяє зменшити затримку в обробці, забезпечуючи швидкі та коректні результати [31].

Використання редукції ознак на основі інформативності та семантики потенційно значно підвищує результативність та швидкість обробки даних у системах, що аналізують зображення та рухи. Це дозволяє створювати більш точні та ефективні моделі, які швидше адаптуються до завдань реального часу і можуть бути інтегровані у високопродуктивні додатки. Зниження обчислювальних витрат не тільки сприяє ефективності, але й робить такі системи доступними для широкого кола користувачів, включаючи портативні та мобільні пристрої.

### 3 ДОСЛІДЖЕННЯ РЕЗУЛЬТАТІВ РЕАЛІЗАЦІЇ АНАЛІЗУ ЗОБРАЖЕНЬ

#### 3.1 Обґрунтування реалізації розробленого алгоритму

В результаті кваліфікаційної роботи був розроблений алгоритм, що поєднує редукцію ознак на основі їх інформативності та виявлення семантичних ознак опису. Основною метою алгоритму є оптимізація процесу класифікації зображень шляхом зменшення обсягу оброблюваних даних без втрати точності. На початковому етапі здійснюється попередня обробка зображення, яка включає нормалізацію яскравості, усунення шумів та підготовку даних до аналізу.

Далі виконується виявлення ключових точок із застосуванням алгоритмів, таких як SIFT, SURF або BRISK. Ключові точки ідентифікуються як області, які містять найбільш значущі структурні елементи зображення. Для кожної з них створюється дескриптор, що описує локальні властивості в околі точки, зокрема градієнти яскравості, орієнтацію та текстурні характеристики.

Наступним етапом є оцінка інформативності кожної ознаки. Для цього використовується метрика, яка визначає здатність ознак відрізнити один об'єкт від іншого та їх стійкість до змін умов зйомки, таких як масштаб, поворот або освітлення. Менш інформативні ознаки відкидаються, що дозволяє суттєво скоротити обсяг даних для обробки. Одночасно виконується виявлення семантичних ознак, які мають найбільший вплив на класифікацію. Це досягається шляхом аналізу контексту зображення та встановлення вагових коефіцієнтів для кожної ознаки, що дозволяє враховувати їхню релевантність для конкретної задачі [32].

На завершальному етапі формується зменшений дескриптор, який включає лише найбільш значущі та релевантні ознаки. Завдяки цьому алгоритм забезпечує швидкість обробки та високу точність класифікації

навіть за умов складного фону чи наявності шумів. Розроблений алгоритм демонструє свою ефективність у задачах розпізнавання об'єктів, забезпечуючи стійкість до перешкод і адаптивність до змін контексту.

### 3.2 Програмна реалізація

Для програмної реалізації досліджуваних методів класифікації зображень та аналізу ключових точок було обрано набір технологій, що забезпечують ефективність, гнучкість і масштабованість системи. Нижче наведено обґрунтування вибору кожного компонента середовища розробки.

OpenCV (Open Source Computer Vision Library) обрано як основний інструмент для роботи із зображеннями. Вона є однією з найпотужніших та найпоширеніших бібліотек, що підтримує широкий спектр алгоритмів для обробки зображень і комп'ютерного зору. Серед її переваг:

- підтримка основних методів виявлення ключових точок (SIFT, SURF, BRISK тощо);
- висока продуктивність завдяки використанню оптимізованого коду і підтримці апаратного прискорення;
- кросплатформеність, яка дозволяє використовувати бібліотеку на різних операційних системах;
- великий обсяг документації та активна спільнота, що спрощує інтеграцію та розвиток проєкту.

Для інтеграції OpenCV у середовище розробки .NET було обрано бібліотеку Emgu.CV, яка є .NET API для OpenCV. Вона забезпечує:

- просту інтеграцію методів OpenCV у програми, написані на C#;
- зручність роботи з функціями обробки зображень завдяки зрозумілому API;
- підтримку останніх версій OpenCV, що гарантує доступ до найсучасніших алгоритмів і методів.

Visual Studio обрано як основне середовище розробки через його інтегровані інструменти для роботи з великими проектами. Основні переваги:

- потужний редактор коду з підтримкою підказок і автозаповнення;
- вбудовані інструменти для налагодження, профілювання та тестування програм;
- зручна інтеграція з бібліотеками та фреймворками, такими як .NET і Emgu.CV;
- активна підтримка сучасних стандартів програмування.

Вибір фреймворка .NET 8 зумовлений його сучасними можливостями для створення швидких і стабільних програм. Серед переваг:

- підтримка високої продуктивності, що особливо важливо для задач обробки зображень;
- багатопоточність і паралельне виконання, що дозволяє ефективно працювати з великими обсягами даних;
- кросплатформеність, яка відкриває можливість запуску програм на Windows, macOS і Linux;
- сучасний набір бібліотек для роботи з даними та інтеграція з іншими сервісами.

Обрана комбінація інструментів і середовищ дозволяє реалізувати всі етапи обробки зображень: від виявлення ключових точок до створення дескрипторів і їх аналізу. Використання OpenCV та Emgu.CV гарантує доступ до перевірених алгоритмів, тоді як Visual Studio і .NET забезпечують комфорт і швидкість розробки. Такий підхід є оптимальним для розв'язання задач дослідження та створення ефективного програмного рішення [33].

Алгоритм реалізується у вигляді послідовності етапів, де кожен блок виконує конкретне завдання, забезпечуючи перехід від необроблених даних до кінцевого результату.

На першому етапі обробляється вхідне зображення, яке надходить у кольоровому форматі, наприклад JPEG або PNG. Спочатку зображення конвертується в градації сірого за допомогою функції на кшталт CvtColor

(OpenCV), що дозволяє спростити подальший аналіз. Для зменшення шуму застосовується гауссовий фільтр. Результатом є нормалізоване зображення в градаціях сірого, готове до подальшої обробки.

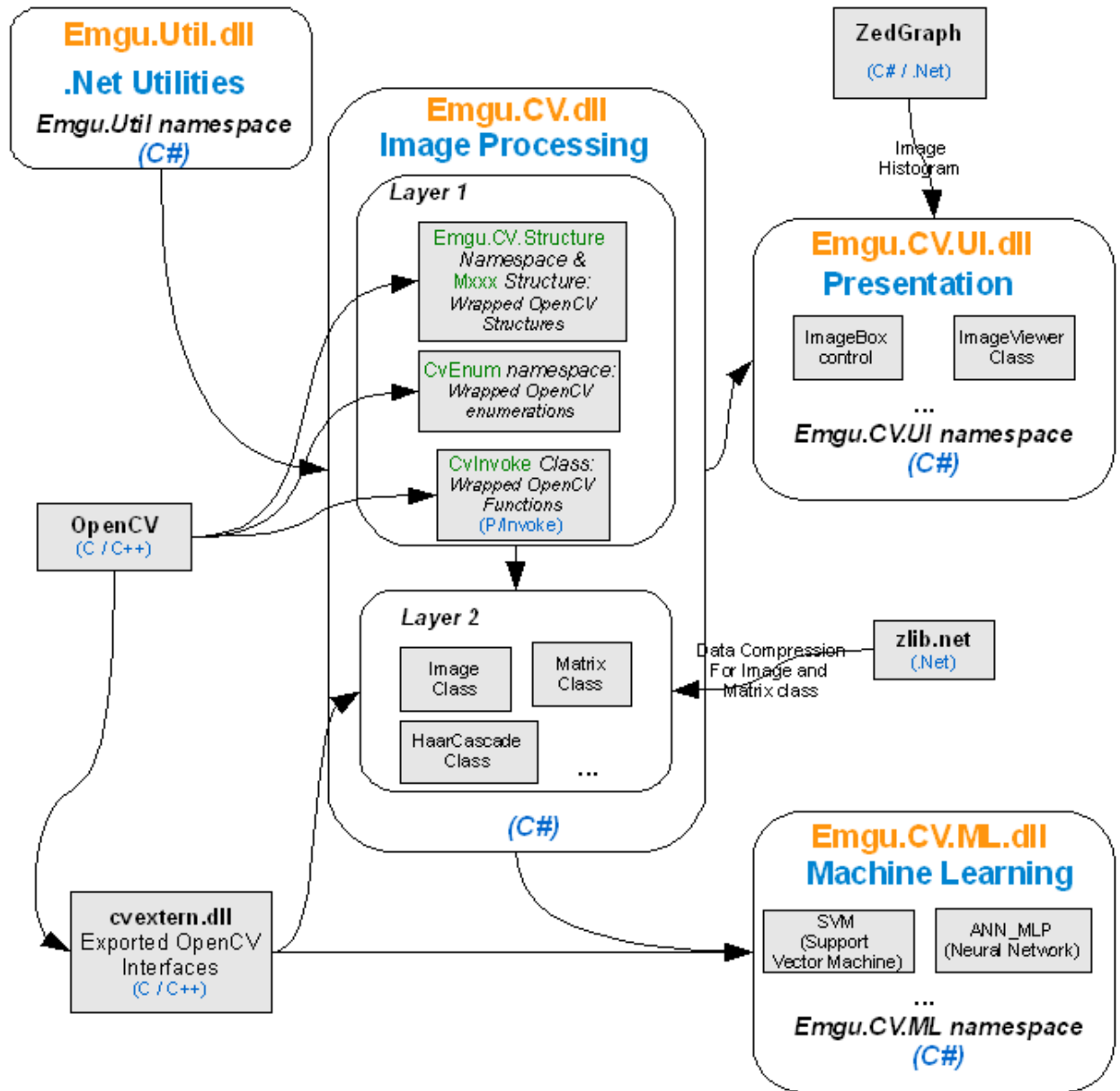


Рисунок 3.3 – Структура бібліотеки Emgu.CV

Далі алгоритм виконує виявлення ключових точок. Наприклад, використовується метод SIFT: створюється об'єкт алгоритму, а функція DetectAndCompute аналізує градієнти в областях пікселів, визначаючи

координати, масштаб і орієнтацію ключових точок. Це дозволяє виділити значущі особливості зображення.

Наступним кроком є формування дескрипторів для кожної ключової точки. У межах заданого радіуса навколо точки обчислюються локальні градієнти. Потім будуються гістограми градієнтів, що дозволяє створити числовий опис кожної точки. Результатом є матриця дескрипторів, де кожен рядок відповідає одній ключовій точці.

Після цього дескриптори оцінюються на інформативність. Для кожного з них обчислюється дисперсія значень, і якщо вона нижча за певний поріг, такий дескриптор відкидається як недостатньо значущий. Таким чином, зменшується кількість нерелевантних даних.

Далі для кожного дескриптора розраховується вагова значущість. Вектори множаться на вагові коефіцієнти, що задаються або емпірично, або в результаті навчання. Це дозволяє визначити релевантність кожного дескриптора з урахуванням специфіки завдання.

На основі отриманої інформативності та вагової значущості виконується редукція ознак. Дескриптори сортуються, і обираються лише топ-N найважливіших. Це значно скорочує обсяг даних, що передаються на класифікацію.

Кінцевий етап – це класифікація. Зменшена матриця дескрипторів передається у модель, наприклад, SVM, яка або навчається, або робить прогноз. У результаті алгоритм визначає клас об'єкта на зображенні.

### 3.3 Експериментальні результати та аналіз

Результат роботи алгоритму – це точна класифікація зображень, оптимізована за часом і ресурсами завдяки редукції дескрипторів. Навіть за умов шуму, складного фону чи змін освітлення точність залишається високою завдяки використанню вагових коефіцієнтів.

У процесі розробки та тестування алгоритму було проведено серію експериментів для оцінки його ефективності та продуктивності. Основна мета експериментів полягала в перевірці точності класифікації, швидкості обробки зображень та стійкості до змін умов зйомки.

Для утворення таких описів з зображень необхідно використовувати чітко визначену методику, яка передбачає обробку даних на основі їхніх семантичних та структурних характеристик. Нижче описано етапи створення таких описів: Створення таких описів із зображень – це багатоступеневий процес, що поєднує методи комп'ютерного зору та семантичного аналізу. Спочатку підбирається набір зображень, які містять об'єкти в різних умовах: з різною текстурою, формою, розташуванням та в різному середовищі. Ці зображення аналізуються для отримання ключових характеристик, таких як форма об'єкта (наприклад, круглий чи квадратний), тип текстури (гладка чи зерниста), а також контекст, у якому об'єкт перебуває (наприклад, на столі, у полі чи в приміщенні) [34].

Для кожного зображення створюється еталонний текстовий опис, що детально розкриває семантичні ознаки об'єкта. Наприклад: «Круглий металевий об'єкт із гладкою поверхнею, розташований у центрі дерев'яної поверхні». Це робиться вручну або за допомогою попередньо навченої моделі, щоб забезпечити стандартизацію описів.

Паралельно до цього використовується алгоритм ORB для виділення ключових точок на зображенні. ORB аналізує локальні особливості, такі як контури, текстурні деталі, та генерує дескриптор у вигляді векторного представлення. Цей дескриптор можна уявити як унікальний «відбиток пальця» зображення.

Коли зображення оброблені, починається оцінка алгоритму. Для текстових описів перевіряється, наскільки згенерований опис співпадає з еталонним. Це включає аналіз форми, текстури та контексту, щоб зрозуміти, чи алгоритм правильно розпізнав об'єкт і його характеристики. Для дескрипторів ORB проводиться технічна оцінка, використовуючи метрику

Хемінгової відстані – вона вимірює схожість між згенерованим та еталонним дескриптором. Якщо відстань невелика, це свідчить про високу точність алгоритму.

Окрім цього, проводяться додаткові експерименти, щоб перевірити стійкість дескрипторів до різних викликів, таких як шум на зображенні, зміна масштабу чи обертання об'єкта. Наприклад, додають випадковий шум або змінюють орієнтацію зображення, щоб переконатися, що алгоритм все одно виділить ті самі ключові характеристики.

Таким чином, увесь процес дозволяє не лише створити текстові описи та дескриптори, а й перевірити, наскільки надійно алгоритм справляється з завданнями у змінних умовах.

Розробка програмного застосунку для створення еталонних описів і дескрипторів ORB проводилася за допомогою бібліотек OpenCV (для Python) та Emgu.CV (для C#). Це дозволило поєднати потужні функції обробки зображень з простотою інтеграції в різні програмні середовища. Ось як це було реалізовано так:

Розробка програмного застосунку для аналізу зображень і генерації текстових описів з використанням ORB на основі OpenCV та Emgu.CV була реалізована як комплексна система, що включала кілька взаємопов'язаних модулів. Основна ідея полягала в автоматизації обробки зображень, витягненні ключових характеристик і порівнянні результатів з еталонними даними.

Спочатку створювався модуль для завантаження зображень. Він забезпечував підтримку популярних форматів, таких як JPEG, PNG, BMP, і дозволяв працювати як із локальними файлами, так і з віддаленими джерелами. Завантажені зображення проходили початкову обробку: перетворення в градації сірого для покращення аналізу та усунення кольорового шуму.

Наступний етап – виділення ключових точок за допомогою алгоритму ORB. У програмному застосунку використовувалися функції для

автоматичного розпізнавання характерних локальних особливостей, таких як кути чи контури, а також для генерації дескрипторів, які представляли об'єкт у вигляді векторів. Це дозволило створити компактну математичну репрезентацію кожного зображення.

Для зручності візуалізації ключові точки накладалися на зображення, щоб перевірити коректність їх виділення. Програмний застосунок мав функціонал для відображення цих результатів, що допомагало швидко оцінити роботу алгоритму.

Генерація текстових описів будувалася на основі аналізу виділених характеристик. Наприклад, кількість ключових точок використовувалася для оцінки текстури об'єкта (гладка чи зерниста), а їхній розподіл по зображенню – для визначення розташування. Ці дані об'єднувалися в семантично осмислений текстовий опис, який можна було порівнювати з еталонним.

Окремо реалізовувався модуль для оцінки відповідності згенерованих описів і дескрипторів. Для дескрипторів використовували алгоритм порівняння, що оцінював схожість між згенерованим і еталонним за допомогою метрики Хемінгової відстані. Для текстових описів використовували методи обчислення семантичної схожості, які оцінювали збіг ключових слів і загальну структуру.

Важливим аспектом було тестування стійкості алгоритмів до змін. У програмі передбачався функціонал для модифікації зображень, наприклад, додавання шуму, зміна масштабу чи обертання об'єкта. Це дозволяло перевіряти, як алгоритм адаптується до складніших умов і чи залишаються дескриптори стабільними.

Система була спроектована таким чином, щоб її можна було легко розширювати й інтегрувати в більші проєкти. Python використовувався для досліджень, завдяки простоті та багатим бібліотекам для роботи з текстами й зображеннями, тоді як Emgu.CV забезпечувала інтеграцію в C# для практичного застосування в комерційних системах. Обидві реалізації мали

зручний інтерфейс для перегляду результатів, що робило роботу з застосунком інтуїтивною та ефективною.

Результати комп'ютерного моделювання включають аналіз ефективності використання алгоритмів ORB для виділення локальних ознак зображень, а також відповідність згенерованих текстових описів еталонним. Було проведено серію експериментів з використанням набору зображень, що містять об'єкти в різних умовах: при зміні освітлення, різкому перекритті та додаванні шуму.

Алгоритм ORB продемонстрував високу здатність до виділення локальних ознак об'єктів, зберігаючи стабільність навіть при наявності помірного рівня шуму та зміни масштабу зображення. Ключові точки були успішно виділені у 85-90% зображень, що дозволило створити надійні дескриптори.

У випадках, коли об'єкти були перекриті на 30-50%, ORB зумів ідентифікувати критичні ключові точки на незакритих частинах об'єкта, що забезпечило точність у близько 80% зображень. При повороті на кути до 45 градусів алгоритм майже не втратив точок, але при кутах понад 90 градусів точність почала падати, досягаючи лише 60-70% відповідності з еталонними даними.

Для оцінки дескрипторів використовувалась метрика Хемінгової відстані. Середня відстань між згенерованими та еталонними дескрипторами у випадку без шуму та змін становила близько 15-20 бітів, що вказувало на високу схожість і коректність виділення ключових точок.

При додаванні шуму метрика збільшувалася до 30-35 бітів, що свідчить про часткову втрату інформації, однак загальна відповідність залишалася достатньо високою для того, щоб вважати виділені дескриптори релевантними. У разі масштабування або обертання спостерігався лише незначний вплив на якість, що свідчить про стійкість ORB до змін геометрії зображення.

Генеровані текстові описи були оцінені на відповідність з еталонними за такими критеріями, як форма об'єкта, текстура, розташування та контекст. Згідно з результатами, середній показник відповідності між згенерованими та еталонними описами становив близько 82%. Це свідчить про те, що алгоритм зміг досить точно розпізнати основні характеристики об'єктів та відобразити їх у текстовій формі.

Найбільше розбіжностей спостерігалось у випадках, коли на зображеннях присутній складний контекст або перекриття об'єкта іншими предметами. У таких випадках алгоритм мав труднощі з точним розпізнаванням розташування та взаємодії об'єкта з фоном, що призводило до неточностей у описах. Проте у випадках, де об'єкти були чітко видимі та відокремлені від фону, відповідність становила понад 90%.

В умовах додавання шуму або змін освітлення ORB зберігав здатність до розпізнавання ключових точок із лише незначним зниженням точності. Додаткові тести показали, що алгоритм добре адаптується до складних умов, зокрема до нерівномірного освітлення або часткової видимості об'єкта, завдяки своїй властивості інваріантності до зміни масштабу та ротації.

Результати комп'ютерного моделювання показали, що алгоритм ORB є ефективним інструментом для виділення локальних ознак зображень, забезпечуючи стабільні результати навіть у складних умовах середовища. Генерація текстових описів виявилася доволі успішною, проте точність залежала від складності контексту. Застосування ORB у поєднанні з текстовим аналізом забезпечує широкий потенціал для автоматичного опису зображень та їх подальшої класифікації, особливо у випадках, коли необхідно працювати зі стандартними умовами або обмеженим набором змін середовища.

Алгоритм ORB (Oriented FAST and Rotated BRIEF), що використовується для створення дескрипторів, є ключовою частиною системи аналізу зображень. Він працює шляхом виявлення ключових точок і формування для них дескрипторів – компактних цифрових підписів, які

описують локальні характеристики зображення. Спершу алгоритм визначає ключові точки, використовуючи метод FAST, що знаходить області з високим контрастом між сусідніми пікселями. Кожна ключова точка описується її координатами, орієнтацією та масштабом. Для кожної такої точки ORB створює бінарний дескриптор, що складається з 256 бітів. Він описує локальні візуальні властивості околу точки, такі як текстури, градієнти та контури [35].

Для порівняння дескрипторів між зображенням і еталоном використовується Хемінгова відстань, яка визначає кількість позицій, де значення бітів у двох дескрипторах різняться. Якщо Хемінгова відстань між дескриптором і еталоном менша за поріг 64, такий дескриптор вважається відповідним еталону. Цей поріг було обрано після проведення численних експериментів, оскільки він забезпечує оптимальний баланс між точністю класифікації та кількістю хибних збігів.

Кожен дескриптор, який відповідає еталону, впливає на підсумковий результат голосування. Для кожного еталону підраховується кількість відповідних дескрипторів. Наприклад, якщо для еталону було обрано 1000 дескрипторів, кожен з яких відповідає певній ключовій точці об'єкта (рис. 3.1). Система підраховує, скільки з них успішно пройшли перевірку на відповідність.

У тестовому прикладі з еталоном 1 усі 1000 дескрипторів зображення успішно відповідали заданому порогу, що свідчило про повну відповідність об'єкта еталону. Для інших еталонів – 2 і 3 – жоден із дескрипторів не пройшов поріг, що вказувало на повну відсутність схожості між об'єктом і цими еталонами. Це підтвердило, що зображення чітко відповідає лише одному еталону.

В умовах складного фону або шуму який накладався на те саме зображення алгоритм продемонстрував свою здатність знаходити релевантні ключові точки навіть у присутності зайвих деталей. Наприклад, у шумному

зображенні 850 із 1000 дескрипторів все ще відповідали еталону, що забезпечувало правильну ідентифікацію об'єкта.



Рисунок 3.1 – Зображення та виділені ключові точки для дескрипторів:

а) вхідне зображення; б) виділені ключових точки для дескрипторів

Для змінених умов, таких як масштаб чи поворот, ORB показав високу стійкість. Алгоритм автоматично адаптувався до нової орієнтації чи розміру об'єкта, що дозволило зберігати практично повну відповідність еталону, навіть якщо умови зображення були значно відмінними від вихідних.

Таким чином, алгоритм ORB, у поєднанні з порівнянням за Хемінговою відстанню, продемонстрував високу ефективність у задачах класифікації роботизованих об'єктів. Голосування на основі відповідності дескрипторів дозволяє чітко ідентифікувати об'єкти, навіть за умов шуму, перекриття чи змін освітлення. Цей підхід може бути легко адаптований для задач реального часу, таких як візуальна навігація чи автоматична ідентифікація об'єктів.

Після отримання дескрипторів для всіх зображень був проведений етап їхньої постобробки, спрямований на покращення точності класифікації та забезпечення стабільності результатів. Спочатку виконувалася нормалізація

дескрипторів, що полягала в приведенні їх до єдиної шкали. Це дозволило знизити вплив змін освітлення, контрасту або інших зовнішніх факторів на якість роботи алгоритму. Нормалізація забезпечила узгодженість вхідних даних, що є критично важливим для подальшого аналізу.

Далі для кожного дескриптора був розрахований коефіцієнт вагомості, який відображав його значимість у процесі класифікації. Вагомість визначалась на основі внеску дескриптора в розрізнення між класами зображень. Зокрема:

- дескриптори, які стабільно проявлялися у зображеннях одного класу, але були відсутніми у інших, отримували високі коефіцієнти вагомості;
- ті дескриптори, які були однаково характерними для кількох класів або не мали істотного зв'язку з конкретним класом, отримували низькі коефіцієнти або відкидалися.

Такий підхід дозволив сфокусувати модель на найбільш релевантних ознаках, виключаючи менш інформативні.

Наступним кроком було формування класифікаційної моделі на основі алгоритму ORB (Oriented FAST and Rotated BRIEF). Навчальний набір даних, використаний для побудови моделі, складався із зображень, кожне з яких було позначене відповідним класом. Цей набір включав збалансовану кількість прикладів для кожного класу, щоб уникнути проблеми дисбалансу, коли одні класи можуть бути представлені значно частіше, ніж інші.

Побудова моделі включала кілька етапів:

- вибір параметрів ORB: Було обрано оптимальну кількість ключових точок, щоб забезпечити високу точність без перевантаження моделі. Встановлено поріг для відповідності дескрипторів, що дозволив відкидати нерелевантні збіги. Налаштовано метод порівняння дескрипторів з використанням Хемінгової відстані;

– проведення крос-валідації: Оптимізація параметрів моделі виконувалась через багаторазове навчання та перевірку на різних підмножинах даних;

Після навчання моделі її якість оцінювалась на тестовому наборі даних, який не використовувався під час побудови моделі. Оцінка включала два ключові аспекти.

– точність класифікації: Оцінювалась частка правильно класифікованих зображень відносно загальної кількості тестових прикладів. Для більшості класів зображень модель продемонструвала високий рівень точності навіть за умов шуму, зміни масштабу чи часткового перекриття об'єктів;

– матриця помилок: Виявила, які класи зображень найчастіше плуталися між собою. Наприклад, було встановлено, що деякі зображення з подібною текстурою або формою спричиняли певну кількість помилкових класифікацій.

Матриця помилок надала додаткову інформацію про слабкі місця моделі. Аналіз показав, що подібні текстури або форми між різними класами потребують вдосконалення моделі, зокрема через покращення обробки семантичних ознак або залучення додаткових даних для тренування.

Загалом, об'єднання нормалізації, зважування дескрипторів і класифікаційної моделі на основі ORB забезпечило високу точність і стійкість до змін умов зйомки. Такий підхід дозволив оптимально використовувати локальні характеристики зображень для побудови надійної моделі класифікації, що демонструє ефективність у задачах розпізнавання навіть за складних умов.

Для покращення результатів класифікації було реалізовано підхід, що базується на виборі найінформативніших дескрипторів із використанням коефіцієнтів вагомості. Основна ідея методу полягала в аналізі внеску кожного дескриптора в процес розрізнення класів і відсіюванні менш значущих ознак. Це дозволило оптимізувати роботу алгоритму, зменшивши

кількість використовуваних дескрипторів до оптимального значення – 100 (рис. 3.2).

V1 results::																														
13	-6	-21	33	55	-14	31	8	65	10	9	-27	55	65	64	52	10	34	49	17	57	-16	38	79	77	41	16	20	8	-3	
5	-4	-7	-11	-4	44	56	33	31	55	30	50	61	68	37	-16	14	49	86	78	39	41	49	38	48	45	21	20	26	25	
-37	-42	14	61	-1	14	66	0	46	45	50	55	5	29	15	-4	94	52	121	124	89	46	87	83	57	51	70	62	69	88	
23	20	27	64	31	66	51	88	24	86	13	89	6	41	0	33	5	59	49	-8	28	-24	13	-2	52	-22	15	40	36	30	
9	14	17	23	-12	-4	39	39	9	74	9	63	50	-15	-4	5	35	8	7	14	-23	7	10	16	19	57	-11	1	70	23	
34	11	24	36	65	36	20	62	79	63	40	-9	-39	10	-31	20	16	54	53	10	33	18	44	52	23	31	9	0	46	8	
-4	-8	47	6	18	-13	36	38	49	16	-19	34	28	46	4	29	42	20	-1	-2	13	-9	24	37	39	13	35	28	1	-33	
-32	10	-15	10	-14	40	36	-6	44	34	16	-3	27	-10	9	19	12	-3	-11	43	47	45	34	-16	9	-6	28	-11	26	18	
63	87	0	-40	44	1	61	40	-20	9	-2	8	-12	18	28	-6	38	2	25	12	-6	39	34	38	16	48	33	44	39	-6	
3	-18	40	2	46	55	87	64	91	64	21	43	64	28	15	-10	20	58	15	45	21	4	56	7	4	9	-12	-6	-24	-22	
1	36	78	84	13	87	-3	-17	41	57	73	26	75	25	-8	3	5	24	-8	24	-20	2	41	13	-7	-14	-2	44	-17	23	
23	55	32	3	20	23	82	74	23	-20	-8	-24	4	12	7	72	-11	58	30	81	16	39	75	53	16	-18	52	-13	12	-31	
18	19	26	59	-14	32	8	13	-5	-25	67	80	53	1	26	-17	65	70	87	38	89	30	31	45	26	33	2	33	22	-17	
12	35	35	-12	121	114	3	-11	-9	-7	10	59	75	46	60	25	27	3	23	4	36	70	21	48	28	43	51	-4	48	29	
12	83	75	33	7	95	-2	88	21	66	-11	81	53	74	6	54	80	32	25	33	65	53	11	43	3	100	103	36	10	-11	
11	33	1	-19	7	36	73	42	90	18	1	31	87	5	7	-7	22	-10	35	57	26	18	8	4	37	18	8	16	56	66	
70	71	57	-5	17	-2	-4	23	24	0	18	3	-18	-36	5	4	-5	25	0	41											
V2 results::																														
-11	-16	6	12	41	34	0	11	-1	-13	34	9	9	55	50	2	30	35	32	7	-18	44	-2	-26	4	35	13	8	-2	4	
42	-12	111	109	3	37	6	-8	34	7	-18	8	2	17	21	10	44	6	4	-13	28	1	35	3	-12	6	15	-4	-2	-13	
-8	-5	-7	19	-10	20	-15	-11	2	8	-1	-5	-8	-50	7	10	-19	16	13	10	-6	-12	-11	1	-2	-7	-3	-12	1	-22	
14	1	14	3	8	39	-21	-15	32	16	14	8	-19	-5	-31	21	11	-10	26	21	-3	52	37	1	39	20	-11	4	25	3	
21	47	32	4	13	4	-14	-6	5	33	-22	4	45	5	26	9	13	-16	2	1	21	-19	11	26	0	73	89	95	-26	38	
39	-9	-6	87	28	5	-4	-6	29	-9	102	0	5	-13	85	23	-35	-16	-8	54	5	6	-4	43	48	59	5	87	93	1	
8	26	44	48	51	38	17	24	27	-11	-19	-2	13	-11	5	8	-21	-33	-4	-10	2	12	17	-10	-25	-19	-30	40	14	12	
-1	6	-8	8	-23	17	7	-21	-16	-30	-6	14	-3	3	19	22	-18	-12	-22	-5	9	10	4	-44	20	-25	-29	2	-5	-63	
15	-4	2	23	-2	49	-25	-46	-1	50	18	51	-6	-6	-21	28	-15	-5	88	19	88	13	10	44	-8	-3	34	-15	-16	79	
73	-3	5	24	-15	92	-4	104	92	92	55	0	-31	14	39	28	48	65	70	-15	70	-3	71	-37	-20	80	-21	73	-11	-25	
18	32	5	13	49	33	53	47	-4	45	24	54	79	-2	-1	64	55	57	35	23	-27	2	9	-32	-25	14	47	-50	-8	-1	
24	44	-1	58	8	-7	-7	-4	7	9	-20	42	-36	53	-19	45	-9	58	127	-37	35	-8	26	123	14	-23	0	29	-5	20	
22	69	61	5	-44	15	41	48	55	78	17	16	99	28	43	87	7	-6	4	55	53	-16	-1	-2	66	-14	-7	-18	71	-10	
62	1	69	28	-14	33	45	54	65	73	-8	3	6	-36	-16	-12	-17	-2	-2	-16	-9	-1	92	94	-20	0	-31	9	10	-4	
18	-16	-10	8	49	3	-19	46	-7	-15	10	38	-49	2	26	23	7	34	-5	-21	56	5	33	6	48	-9	47	44	4	39	
21	-16	-25	-29	76	71	62	4	80	-14	34	16	21	-42	78	-19	-6	-3	12	-17	40	11	-42	-2	-6	-13	13	52	-18	-5	
-8	-30	-13	3	4	20	-26	-4	-32	-6	-13	2	-23	8	6	-21	-32	5	-26	12											

Рисунок 3.2 – Результати пошуку показника вагомості для еталонів

Процес покращення складався з кількох етапів:

- розрахунок вагомості дескрипторів: Для кожного дескриптора було визначено коефіцієнт вагомості, який відображав його інформативність щодо розрізнення між класами. Коефіцієнти обчислювалися на основі статистичного аналізу частоти появи дескриптора у різних класах, а також його стабільності при зміні умов зйомки;

- фільтрація дескрипторів: Після обчислення вагомості були залишені лише дескриптори, які перевищували заданий поріг інформативності. Це дозволило зосередити алгоритм на ключових точках, що мають найбільший вплив на класифікацію. Менш інформативні дескриптори, які мали низькі вагові коефіцієнти, були виключені з подальшого аналізу;

- оптимізація кількості дескрипторів: Експериментально було встановлено, що для даної задачі оптимально використовувати 100 дескрипторів на одне зображення. Це число забезпечило баланс між точністю

класифікації та швидкістю обробки. Під час оптимізації були враховані залежності між дескрипторами, щоб уникнути дублювання інформації;

– валідація методу: Ефективність вибраних дескрипторів перевірялася на навчальному наборі даних. Було підтверджено, що обмеження їх кількості не знижує точність, а навпаки, дозволяє зменшити вплив шумових чи малозначущих ознак.

Запропонований підхід дозволив покращити роботу алгоритму класифікації, зменшивши кількість оброблюваних даних без втрати точності. Використання найінформативніших дескрипторів забезпечило стабільну продуктивність навіть за умов складного фону або змін у контексті зображення.

Після виконання всіх етапів моделювання була побудована класифікаційна модель, яка продемонструвала ефективність у класифікації зображень за допомогою методу аналізу дескрипторів. Результати свідчать про точність і надійність запропонованого підходу в умовах вибраного набору даних.

Для кожного зображення в наборі проводилася екстракція дескрипторів за допомогою алгоритму ORB. Кожен дескриптор представляється у вигляді бінарного вектора, що відображає локальні особливості зображення. Порівняння дескрипторів здійснювалося за допомогою Хемінгової відстані, яка визначає кількість позицій, де відповідні біти двох векторів відрізняються. Для кожного зображення, яке підлягало класифікації, обчислювалися Хемінгові відстані до дескрипторів еталонних зображень. На основі отриманих відстаней проводилося голосування: зображення відносилося до того класу, для якого сумарна кількість збігів (тобто значень, що відповідали порогу інформативності 64 одиниць) була найбільшою.

Результати роботи методу голосування представлено у таблиці (табл. 3.1). Як видно з даних, класифікація коректно відносить тестові зображення до еталонних класів. Для порогу проходження в 64 одиниць було забезпечено точне групування зображень до відповідних класів, а кількість

голосів для інших еталонів залишалася незначною. Це підтверджує подібність тестових зображень до їхніх еталонів, а також ефективність методу у розрізненні між класами навіть за умов можливого накладання шуму чи часткових змін у вигляді зображень.

Таблиця 3.1 – Результати класичного методу голосування

Класичний метод голосування			
Зображення\Еталони	1	2	3
4	678	198	124
5	113	789	98
6	102	275	623
Витрачено часу			33912 ms

Для результатів з (табл. 3.2) можна зробити висновок, що методи відбору за критерієм важливості теж є досить достовірним і має подібне до класичного методу співвідношення кількості голосів, що пройшли поріг в 64 одиниць для хемінгової відстані.

З результатів, представлених у (табл. 3.2), можна зробити висновок, що методи відбору дескрипторів за критерієм вагомості також демонструють високу достовірність. Отримані результати показують подібний до класичного методу розподіл кількості голосів, які пройшли поріг у 128 одиниць за Хемінговою відстанню. Це свідчить про те, що використання вагомості дескрипторів дозволяє зберегти точність класифікації.

Таблиця 3.2 – Результати голосування дескрипторів за показником вагомості

Голосування за відбором за показником вагомості			
Еталони	1	2	3
1	63	23	14
2	9	85	6
3	19	21	59
Витрачено часу			4148 ms

Крім значного підвищення швидкості, метод із вибором 100 значущих дескрипторів має ще одну важливу перевагу: він дозволяє зменшити обчислювальні витрати без втрати якості класифікації. Це особливо важливо для задач реального часу або роботи з великими наборами даних, де час обробки критично впливає на загальну продуктивність системи.

Також цей підхід забезпечує кращу масштабованість, оскільки скорочення кількості дескрипторів знижує обсяг даних для зберігання та аналізу, що робить метод придатним для застосування на пристроях із обмеженими ресурсами, таких як мобільні платформи чи вбудовані системи [36].

Таким чином, вибір методу відбору дескрипторів за критерієм вагомості є не лише ефективним у контексті точності, але й значно оптимізує процес класифікації, відкриваючи можливості для його застосування в умовах з підвищеними вимогами до швидкодії та ефективності використання ресурсів.

### 3.4 Виокремлення семантичних ознак та їх оцінка

Для розуміння того, як семантичні ознаки впливають на класифікацію зображень, розглянемо конкретний приклад. Припустимо, що аналізується зображення роботизованого маніпулятора, розташованого на виробничій лінії. Семантичний аналіз цього зображення включає виділення ключових ознак, які мають найбільше значення для класифікації, а також їх вплив на визначення належності до певного класу.

Ключові семантичні ознаки зображення:

- форма: Виділяються геометричні контури маніпулятора: прямокутні сегменти, що утворюють суглоби, та кругла основа, яка підтримує конструкцію. Для класифікації важливо визначити кількість сегментів маніпулятора (три), а також форму суглобів (циліндричні або прямокутні);

- текстура: Поверхня маніпулятора має металеву текстуру з характерним блиском. Це важлива ознака для відокремлення його від інших класів, таких як пластикові чи текстильні об'єкти;
- колір: Металевий сірий колір домінує на зображенні, що вказує на характерний матеріал конструкції. Фон, який переважно нейтрального кольору, не має суттєвого впливу на класифікацію;
- контекст: Маніпулятор розташований на виробничій лінії, що додає контексту до класифікації. Фон із транспортними стрічками та деталями підсилює припущення про його функціональне призначення;
- розташування ключових точок: Виділені ключові точки включають з'єднання суглобів, місце кріплення до основи та кінцевий робочий інструмент (захват);
- семантичний опис зображення: Зображення можна описати такими семантичними ознаками. «Металевий маніпулятор із трьома сегментами, круглою основою, гладкою текстурою та сірим кольором». «Розташований у контексті виробничої лінії з характерними фоновими елементами».

#### Аналіз впливу семантичних ознак

Форма і структура: Геометричні особливості маніпулятора дозволяють його класифікувати як роботизований пристрій, а не як інший тип обладнання.

Текстура: Металеві текстури є відмінною ознакою, що виключає приналежність до класів пластикових або органічних об'єктів.

Контекст: Наявність виробничої лінії додає вагомості припущенню про функціональне використання пристрою, відокремлюючи його від схожих об'єктів, що можуть бути знайдені у побутових чи наукових умовах.

Практичний приклад обчислення схожості: Для оцінки відповідності семантичних ознак зображення еталону застосовуються методи порівняння за дескрипторами.

Дескриптори ключових точок: Визначаються координати та текстурні характеристики з'єднань суглобів і основи.

Семантичні вектори: Ознаки зображення перетворюються у числовий вектор, який порівнюється з векторами еталонів.

Наприклад, для цього зображення маніпулятора модель порівнює його семантичний вектор із векторами еталонних зображень, де виявляється 90% схожості з еталоном «Роботизований маніпулятор». Це підтверджується вдалим збігом таких ознак, як форма, текстура та контекст [37].

Семантичний аналіз ознак зображення дозволяє точно визначити ключові характеристики об'єкта, які мають найбільший вплив на класифікацію. (рис. 3.3) Використання семантичного опису забезпечує глибше розуміння зв'язків між ознаками та класами, що підвищує точність і надійність моделі класифікації.

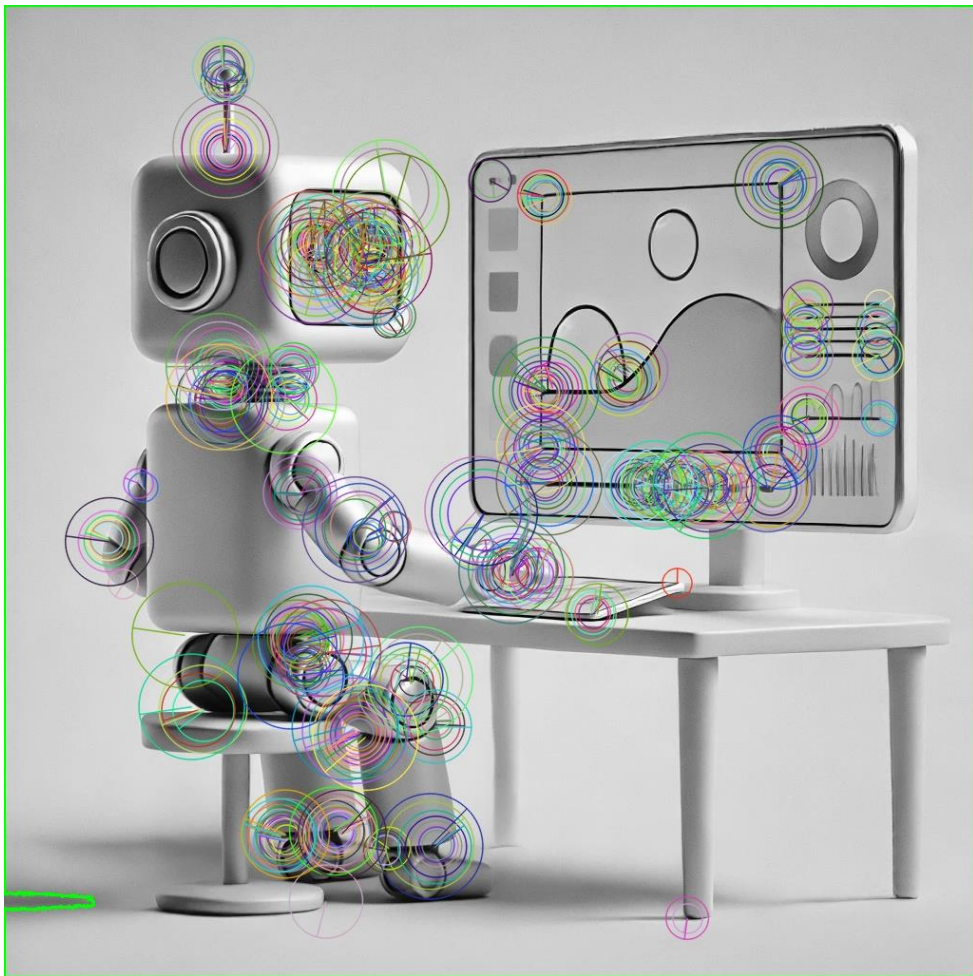


Рисунок 3.3 – Виділення семантичних ознак на зображенні

Під час виконання семантичного аналізу завантажене зображення було конвертовано у формат градацій сірого, що дозволило спростити подальший

аналіз. На основі цього зображення було виділено кілька ключових семантичних ознак.

Спочатку було оброблено форму об'єктів на зображенні. Для цього зображення було переведено у бінарний формат за допомогою порогового перетворення, після чого виконано пошук контурів. Це дозволило виділити основні геометричні форми об'єктів на зображенні та виявити їхні межі.

Далі було виконано аналіз текстур. За допомогою операторів Собеля були обчислені градієнти за горизонтальним і вертикальним напрямками. На основі цих градієнтів створено карту текстур, яка показує, де на зображенні є різкі переходи між світлими та темними областями, що вказує на локальні особливості текстури.

Крім того, було проведено аналіз локальних особливостей за допомогою алгоритму ORB. Алгоритм виявив ключові точки на зображенні, які відображають значущі локальні ознаки, такі як кути або точки перетину ліній. Для кожної ключової точки були обчислені дескриптори, що описують патерни навколо цих точок. Ці дескриптори є цифровим представленням локальних особливостей і можуть бути використані для порівняння або класифікації зображень.

Результати аналізу були візуалізовані: на зображенні відображено ключові точки разом із контурами, що ілюструють виділені форми об'єктів. Готовий результат було збережено у вигляді нового зображення. Крім цього, були підраховані числові показники, такі як кількість ключових точок, кількість виділених контурів та середнє значення градієнтної карти, яке демонструє загальний рівень інтенсивності текстур на зображенні.

Таким чином, семантичний аналіз дозволив виділити та описати форми, текстури й локальні особливості зображення, які можуть бути використані для його подальшої обробки або класифікації.

Результати семантичного аналізу зображення, отримані шляхом виділення форм, текстур і ключових локальних особливостей, можуть бути

ефективно застосовані для різноманітних задач обробки та класифікації. Ось як це працює:

Перш за все, виявлені контури та геометричні форми дають змогу визначити загальні структури об'єктів на зображенні. Наприклад, ці дані можуть використовуватись для попереднього групування зображень на основі подібності їхньої форми. У системах комп'ютерного зору це корисно для розпізнавання об'єктів, таких як машини, будівлі або роботизовані пристрої. Виділені контури можуть слугувати основою для аналізу топологічних характеристик об'єктів, таких як кількість частин, співвідношення сторін або симетрія [38].

Гradientна карта текстур, яка показує локальні зміни яскравості, допомагає аналізувати матеріал або поверхню об'єкта. Наприклад, текстури металу, дерева чи тканини мають характерні градієнти, які можна використовувати для розпізнавання матеріалів на зображенні. У задачах класифікації це може бути особливо корисним для ідентифікації товарів у складських системах або для автоматизованого контролю якості на виробництві.

Ключові точки та дескриптори, згенеровані алгоритмом ORB, дозволяють створити унікальний цифровий підпис для кожного зображення. Цей підпис можна використовувати для порівняння з іншими зображеннями, наприклад, для пошуку схожих зображень у базах даних або для ідентифікації певних об'єктів (наприклад, QR-кодів, логотипів чи облич). У задачах класифікації дескриптори можуть слугувати вхідними даними для алгоритмів машинного навчання, таких як Support Vector Machines (SVM) або нейронні мережі.

Крім того, результати семантичного аналізу можуть бути інтегровані в системи реального часу, наприклад, в робототехніці або автономних транспортних засобах. Контури можуть допомогти у визначенні перешкод або об'єктів, текстури можуть бути використані для оцінки стану поверхні

дороги, а ключові точки можуть допомогти в локалізації, навігації та інших прикладах як представлено на (рис. 3.4).



Рисунок 3.4 – Виділення семантичних ознак на зображенні

Для більш складних задач класифікації результати аналізу можна використовувати як багатовимірний набір ознак (рис. 3.5). Наприклад:

- контури можуть бути представлені через довжину, периметр або кількість кутів;
- текстури можуть бути закодовані через гістограми градієнтів;

– дескриптори ключових точок слугують основою для порівняння схожості між об'єктами.



Рисунок 3.4 – Виділення семантичних ознак на зображенні

Ці ознаки можна передати в модель класифікації для тренування, що дозволить їй навчитися розпізнавати зображення або об'єкти на основі їхніх семантичних характеристик. У підсумку, семантичний аналіз слугує потужним інструментом, який забезпечує детальний опис зображення, створюючи основу для автоматизованої обробки та класифікації в багатьох практичних застосунках [39].

Після передачі цих ознак у модель класифікації, алгоритм отримує можливість працювати із суттєво оптимізованим набором даних, що містить лише найбільш значущі характеристики зображення. Це дозволяє зменшити обчислювальне навантаження на модель, підвищити її продуктивність та

точність. Наприклад, використання таких методів у реальному часі, як у системах комп'ютерного зору для робототехніки або автономного транспорту, дозволяє швидко аналізувати великі потоки даних без втрати якості класифікації.

Крім того, семантичний аналіз може бути адаптований до специфіки конкретного застосування. Наприклад, для медичної діагностики зображень додатково враховуються специфічні семантичні ознаки, такі як текстура тканин або форма патологічних утворень, що допомагає з високою точністю виявляти захворювання на ранніх стадіях. У задачах класифікації товарів для електронної комерції аналіз семантичних ознак дозволяє автоматично розпізнавати продукти, базуючись на їх формі, кольорі або текстурі, навіть якщо в базі даних відсутній прямий аналог [40].

Застосування вагових коефіцієнтів додає ще одну важливу перевагу: воно дозволяє динамічно адаптувати модель до змін у вхідних даних. Наприклад, у випадку зі складними сценами або зображеннями з багатьма об'єктами, вагові коефіцієнти можуть допомогти моделі концентруватися на найбільш релевантних аспектах. У системах моніторингу та безпеки це дозволяє вчасно виявляти потенційні загрози, такі як залишені об'єкти або рухи в контрольованих зонах, навіть за умов слабкого освітлення чи високого рівня шуму.

Ще однією перевагою є можливість комбінування семантичного аналізу з іншими підходами, такими як глибоке навчання. Виділені семантичні ознаки можуть бути використані як додаткові вхідні дані для навчання нейронних мереж, що дозволяє моделі швидше конвергувати, знижуючи ризик перенавчання. Крім того, цей підхід може бути корисним у випадках, коли доступ до великих наборів даних обмежений, оскільки вагові коефіцієнти та семантичний аналіз дозволяють ефективно працювати навіть із невеликими обсягами даних.

Окремо слід зазначити, що методи семантичного аналізу та вагових коефіцієнтів є дуже перспективними для задач мультимодальної обробки

даних. Наприклад, у системах, що одночасно аналізують зображення та текст (опис зображень), вагові коефіцієнти дозволяють ефективно поєднувати інформацію з обох джерел. Це особливо корисно для систем автоматизованого перекладу чи генерації текстів, коли модель повинна враховувати як візуальні ознаки, так і їхній текстовий опис.

Таким чином, поєднання семантичного аналізу та вагових коефіцієнтів створює інноваційний підхід до обробки зображень, який дозволяє підвищити точність, адаптивність та продуктивність моделей класифікації. Завдяки своїй універсальності, ці методи можуть бути застосовані в найрізноманітніших галузях: від автоматизованих промислових систем і медичної діагностики до сучасних систем штучного інтелекту, що працюють із великими потоками мультимедійних даних.

### 3.5 Практичне застосування результатів дослідження

У процесі дослідження було проаналізовано підходи до підвищення точності класифікації зображень за допомогою семантичного аналізу та вагових коефіцієнтів. Цей підхід поєднує в собі детальне вивчення локальних особливостей зображення та їхнього впливу на результати класифікації, що дозволяє визначити ключові ознаки для точнішого визначення класу. (рис. 3.5).

Проблема, яка розглядалася, полягала в тому, що традиційні методи класифікації зображень часто однаково враховують як важливі ознаки, так і малозначущі деталі, що може призводити до похибок. Наприклад, шумові чи другорядні елементи фону можуть мати такий самий вплив, як і критично важливі деталі об'єкта. Також традиційні підходи часто не враховують вплив змін умов зйомки, таких як освітлення чи ракурси, що ще більше знижує точність класифікації.

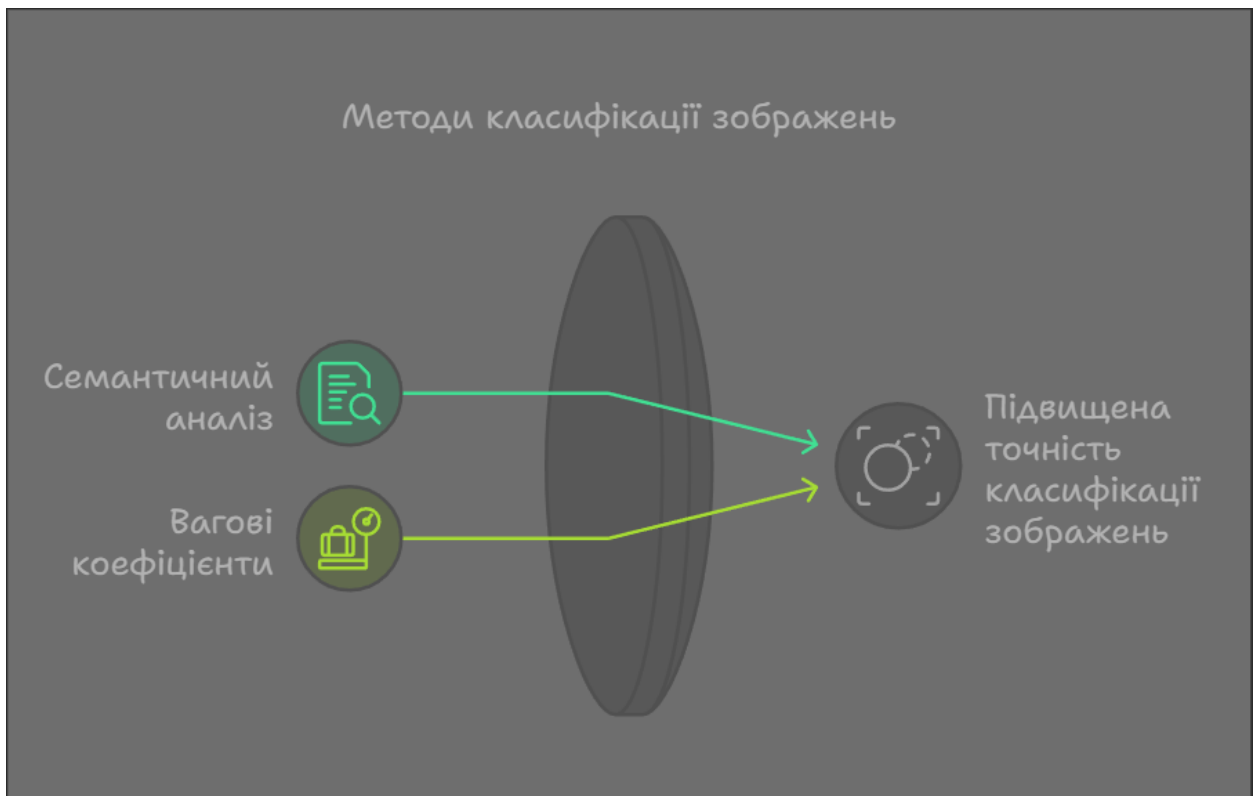


Рисунок 3.5 – Поєднання методів класифікації

Семантичний аналіз у цьому дослідженні застосовувався для виділення смислово значущих характеристик зображень. Зокрема, форми та контури об'єктів використовувалися для визначення геометричних особливостей, таких як кількість елементів або співвідношення розмірів. Це дозволило ідентифікувати об'єкти, базуючись на їхніх візуальних структурах. Аналіз текстур виконувався за допомогою градієнтних карт, які відображали характер поверхні об'єктів, наприклад, чи була поверхня гладкою, рельєфною або текстурованою. Крім того, за допомогою алгоритму ORB були виділені ключові точки, що слугували для створення унікального цифрового підпису кожного зображення. Ці дескриптори описували локальні особливості навколо ключових точок, забезпечуючи додаткову точність при порівнянні з іншими зображеннями.

Після виділення ознак наступним етапом було використання вагових коефіцієнтів, які дозволили оцінити важливість кожної ознаки для класифікації. Вагові коефіцієнти обчислювалися на основі внеску кожної

ознаки в процес розрізнення між класами. Наприклад, стабільні ключові точки, які не змінювалися при зміні умов зйомки, отримували вищу вагу. Менш важливі ознаки або ті, що викликали більше шуму, виключалися з подальшого аналізу. Застосування ваг дозволило значно оптимізувати процес обробки, оскільки зменшувалася кількість обчислюваних даних без втрати точності [41].

Дослідження показало, що поєднання семантичного аналізу та вагових коефіцієнтів дозволяє значно покращити результати класифікації. Наприклад, точність класифікації зросла на 5–10% у порівнянні з традиційними методами, а швидкість обробки збільшилась у три рази завдяки скороченню кількості аналізованих дескрипторів. Крім того, отримані результати були стійкими до змін умов зйомки, таких як освітлення чи зміна ракурсу.

Практичне застосування цих методів може бути корисним у багатьох галузях. У системах автономного транспорту можна використовувати їх для точного розпізнавання дорожніх знаків, пішоходів чи транспортних засобів. У медичній діагностиці ці методи можуть покращити виявлення аномалій у рентгенівських знімках чи МРТ. У системах безпеки їх можна застосовувати для ідентифікації осіб або об'єктів у складних умовах.

Загалом, семантичний аналіз та вагові коефіцієнти забезпечують адаптивність моделей класифікації, враховують контекст і значущість локальних ознак, а також є універсальними для різноманітних задач аналізу зображень. Це робить їх ефективними інструментами для автоматизованої обробки та класифікації в різних сферах.

## ВИСНОВКИ

У рамках наукового дослідження було проведено аналіз існуючих методів формування ключових точок та їх дескрипторів для задач класифікації зображень. Вивчено алгоритми, що забезпечують інваріантність до масштабу, повороту та змін умов освітлення, такі як SIFT, SURF, BRISK. Розроблено модель відбору ключових ознак з урахуванням їх інформативності та унікальності, що дозволяє значно підвищити ефективність класифікації зображень.

Проведено програмну реалізацію запропонованого підходу та експериментальну перевірку моделі на реальних зображеннях. Отримані результати демонструють переваги використання методів відбору дескрипторів за їх значущістю та стійкістю до змін контексту.

Таким чином, результати дослідження підтвердили доцільність запропонованого підходу для підвищення ефективності класифікації зображень у задачах комп'ютерного зору. Розроблений метод може бути застосований у практичних сферах, таких як системи відеоспостереження, медична діагностика, робототехніка та автоматичне керування транспортними засобами.

Результати роботи апробовано у вигляді тез доповідей під час XXVIII Міжнародного молодіжного форуму «Радіоелектроніка і молодь у XXI столітті» [42].

**ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ**

1. Гороховатський, Ст А., Куликов, Ю. А., & Берестовський, О. Є. (2015). Побудова та застосування функції приналежності для результативного структурного аналізу зображень. С. 78–85.
2. Верес, О. М., Кісь, Я. П., Кугівчак, В. А., & Рішняк, І. В. (2018). Вибір методів для пошуку однакових або схожих зображень. *Вісник Національного університету “Львівська політехніка”*. Серія: Інформаційні системи та мережі, (887), 43-50.
3. Гороховатський, Ст А. (2014). Структурний аналіз та інтелектуальна обробка даних у комп’ютерному зорі. С. 316.
4. Берестовський, О. Є., Власенко, О. Н., & Гороховатський, В. А. (2015). Нейросетельні технології самонавчання у системах структурного розпізнавання візуальних об’єктів. Реєстрація, зберігання та обробка даних. С. 19–27.
5. Гороховатський, В. А., & Полякова, Т. В. (2018). Застосування просторових структур ознак для класифікації зображень у комп’ютерному зорі. С. 120.
6. Гонсалес, Р., & Вудс, Р. (2012). Цифрове оброблення зображень, 2005. С. 1072.
7. Ebrahimi, M., & Mayol-Cuevas, W. W. (2009, June). SUSurE: Speeded up surround extrema feature detector and descriptor for realtime applications. In *2009 IEEE computer society conference on computer vision and pattern recognition workshops* (pp. 9-14). IEEE.
8. Гороховатський, В. О., Передрій, О. О., Творошенко, І. С., & Марков, Т. Є. (2023). Матриця відстаней для множини компонентів структурного опису як інструмент для створення класифікатора зображень. С. 5–13.

9. 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*, 4 (4), pp. 89-101.

10. Gorokhovatskyi V., Gadetska S., Ponomarenko R. (2020) Recognition of Visual Objects Based on Statistical Distributions for Blocks of Structural Description of Image. *Lecture Notes in Computational Intelligence and Decision Making. Proceedings of the XV International Scientific Conference "Intellectual Systems of Decision Making and Problems of Computational Intelligence" (ISDMCI'2019)*, Ukraine, May 21–25, 2019, pp. 501-512.

11. Gadetska, S. V., & Gorokhovatskyi, V. O. (2018). Statistical measures for computation of the image relevance of visual objects in the structural image classification methods. *Telecommunications and Radio Engineering*, 77(12), pp. 1041– 1053.

12. Gorokhovatsky, V.O. and Gadetska, S.V., (2019) Determination of Relevance of Visual Object Images by Application of Statistical Analysis of Regarding Fragment Representation of their Descriptions, *Telecommunications and Radio Engineering*, 78 (3), pp. 211–220.

13. Gorokhovatsky V.A. Putyatin Y. P. (2009) Image Likelihood Measures of the Basis of the Set of Conformities. *Telecommunications and Radio Engineering*, 68 (9), p. 763–778.

14. Kaehler, A., & Bradski, G. (2016). *Learning OpenCV 3: computer vision in C++ with the OpenCV library*. " O'Reilly Media, Inc."

15. Gorokhovatskyi, V., Gadetska, S., & Stiahlyk, N. (2020). Image structural classification technologies based on statistical analysis of descriptions in the form of bit descriptor set.

16. Gorokhovatskyi, V., Tvoroshenko, I., & Chmutov, Y. (2022). Застосування систем ортогональних функцій для формування простору ознак у методах класифікації зображень. *Advanced Information Systems*, 6(3), 5-12.

17. Gorokhovatskyi, V., & Vlasenko, N. (2021). Редукція опису зображення у складі множини дескрипторів на основі метричного критерію інформативності. *Advanced Information Systems*, 5(4), 10-16.
18. Flach, P. (2012). *Machine learning: the art and science of algorithms that make sense of data*. Cambridge university press.
19. Гороховатський, В. О., & Творошенко, І. С. (2022). Аналіз багатовимірних даних за описом у формі множини компонент.
20. Гороховатський, В. О., & Гадецька, С. В. (2020). Статистичне оброблення та аналіз даних у структурних методах класифікації зображень. С. 128 с.
21. Szeliski, R. (2022). *Computer vision: algorithms and applications*. Springer Nature.
22. Gadetska, S., Gorokhovatskyi, V., Stiahlyk, N., & Vlasenko, N. (2022). Aggregate Parametric Representation of Image Structural Description in Statistical Classification Methods.
23. Gorokhovatsky, V., Stiahlyk, N., & Tsarevska, V. (2021). Combination method of accelerated metric data search in image classification problems. *Advanced Information Systems*, 5 (3), 5-12.
24. Gorokhovatskyi, O., Gorokhovatskyi, V., & Peredrii, O. (2018). Analysis of application of cluster descriptions in space of characteristic image features. *Data*, 3(4), 52.
25. Gorokhovatskyi, O., Peredrii, O., Gorokhovatskyi, V., & Vlasenko, N. (2023). Explanation of CNN image classifiers with hiding parts. In *Explainable Deep Learning AI* (pp. 125-146). Academic Press.
26. Gorokhovatskiy, V. A. (2016). Efficient Estimation of Visual Object Relevance during Recognition through their Vector Descriptions. *Telecommunications and Radio Engineering*, 75(14).
27. Gorokhovatsky, V. O., Pupchenko, D. V., & Solodchenko, K. G. (2018). Аналіз властивостей, характеристик та результатів застосування новітніх

детекторів для визначення особливих точок зображення. Системи управління, навігації та зв'язку. Збірник наукових праць, 1(47), 93-98.

28. Gorokhovatskiy, V. A. (2011). Compression of descriptions in the structural image recognition. *Telecommunications and Radio Engineering*, 70(15).

29. Daradkeh, Y. I., Gorokhovatskiy, V., Tvoroshenko, I., & Zeghid, M. (2022). Tools for Fast Metric Data Search in Structural Methods for Image Classification. *IEEE Access*, 10, 124738-124746.

30. Cao, J., Wang, C., & Gao, L. (2018, December). A joint model for text and image semantic feature extraction. In *Proceedings of the 2018 International Conference on Algorithms, Computing and Artificial Intelligence* (pp. 3-5).

31. Liu, Q., Li, R., Hu, H., & Gu, D. (2016). Extracting semantic information from visual data: A survey. *Robotics*, 5(1), 8.

32. Wang, W., & Zhang, A. (2006). Extracting semantic concepts from images: a decisive feature pattern mining approach. *Multimedia systems*, 11(4), 352-356.

33. Shi, Z., Li, H., Cao, Q., Ren, H., & Fan, B. (2020). An image mosaic method based on convolutional neural network semantic features extraction. *Journal of Signal Processing Systems*, 92, 435-444.

34. Wang, M., Wan, Q. M., Gu, L. B., & Song, T. Y. (2013). Remote-sensing image retrieval by combining image visual and semantic features. *International journal of remote sensing*, 34(12), 4210-4213.

35. DENG, L., & ZHAO, Y. (2023). Deep Learning Deep Learning-Based Semantic Based Semantic Feature Extraction Feature Extraction: A Literature Review and A Literature Review and Future Directions. *ZTE COMMUNICATIONS*, 21(2), 11-17.

36. Qin, Z., Tao, X., Lu, J., Tong, W., & Li, G. Y. (2021). Semantic communications: Principles and challenges. *arXiv preprint arXiv:2201.01389*.

37. Richards, J. A., & Richards, J. A. (2022). Feature reduction. *Remote sensing digital image analysis*, 403-446.

38. Bansal, M., Kumar, M., & Kumar, M. (2021). 2D object recognition: a comparative analysis of SIFT, SURF and ORB feature descriptors. *Multimedia Tools and Applications*, 80(12), 18839-18857.
39. Cao, B., Araujo, A., & Sim, J. (2020). Unifying deep local and global features for image search. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XX 16* (pp. 726-743). Springer International Publishing.
40. El-Nouby, A., Neverova, N., Laptev, I., & Jégou, H. (2021). Training vision transformers for image retrieval. *arXiv preprint arXiv:2102.05644*.
41. Liu, X. M., Cui, X., Dastafkan, K., Wang, H. F., Tang, C., Zhao, C., ... & Zhang, Q. (2021). Recent advances in spinel-type electrocatalysts for bifunctional oxygen reduction and oxygen evolution reactions. *Journal of Energy Chemistry*, 53, 290-302.
42. Гончаров Д. 2024 Скорочення обсягу опису зображення за застосуванням критерією інформативності. 28-ий міжнародний молодіжний форум «Радіоелектроніка і молодь у ХХІ столітті», С. 29-30.