

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Західноукраїнський національний університет
Факультет комп'ютерних інформаційних технологій
Кафедра комп'ютерної інженерії

Горішній Ярослав Ігорович

**«Алгоритм сегментації цифрових зображень на
основі аналізу характерних ознак однорідних
областей / Algorithm for segmentation of digital images
based on the analysis of characteristic features of
homogeneous areas»**

спеціальність: 123 - Комп'ютерна інженерія
освітньо-професійна програма - Комп'ютерна інженерія
Кваліфікаційна робота

Виконав студент групи КІм-21
Я.І. Горішній

Науковий керівник:
к.т.н., доц. Ю.М. Батько

Кваліфікаційну роботу допущено
до захисту:

" ____ " _____ 20__ р.

Завідувач кафедри
_____ Л. О. Дубчак

Тернопіль – 2023

РЕЗЮМЕ

Кваліфікаційна робота на тему “Алгоритм сегментації цифрових зображень на основі аналізу характерних ознак однорідних областей” зі спеціальності 123 «Комп’ютерна інженерія» освітнього ступеня «магістр» написана обсягом 85 сторінок та містить 30 ілюстрацій, 4 таблиць, 2 додатки та 60 джерел за переліком посилань.

Метою роботи є розробка алгоритму сегментації цифрових зображень на основі аналізу характерних ознак однорідних областей.

Методи досліджень. Для розв’язання поставлених задач у кваліфікаційній роботі використано методи: обробки цифрових зображень (для попередньої обробки цифрових зображень); сегментації зображень (для попередньої розмітки зображень); об’єктно-орієнтованого програмування (для проектування програмного додатку обробки та аналізу цифрових зображень).

Результати дослідження: алгоритм сегментації цифрових зображень на основі однорідності областей, програмний додаток обробки та аналізу цифрових зображень.

Результати роботи можуть бути використані в створенні нових систем обробки цифрових зображень, системах комп’ютерного зору, для наукових досліджень та в навчальному процесі.

Орієнтовні напрямки розвитку досліджень: розроблення алгоритмів сегментації складних цифрових зображень, створення нових програмних засобів та моделей для навчання.

КЛЮЧОВІ СЛОВА: ЦИФРОВІ ЗОБРАЖЕННЯ, СЕГМЕНТАЦІЯ ЗОБРАЖЕНЬ, ХАРАКТЕРНІ ОЗНАКИ, ТЕХНОЛОГІЇ ОБРОБКИ ЦИФРОВИХ ЗОБРАЖЕНЬ.

RESUME

Graduate qualification work on “ Algorithm for segmentation of digital images based on the analysis of characteristic features of homogeneous areas ” specialty 123 – Computer Engineering is 85 pages long and contains 30 illustrations, 4 tables, 2 appendices and 60 references.

The aim of the work is to develop an algorithm for segmentation of digital images based on the analysis of characteristic features of homogeneous areas.

Research methods. To solve the tasks in the qualification work, the following methods were used: digital image processing (for pre-processing of digital images); image segmentation (for preliminary marking of images); object-oriented programming (for designing a software application for processing and analyzing digital images).

Research results: an algorithm for the segmentation of digital images based on the unidirectionality of areas, a software application for the processing and analysis of digital images.

The results of the work can be used in the creation of new digital image processing systems, computer vision systems, for scientific research and in the educational process.

Indicative directions of research development: development of algorithms for segmentation of complex digital images, creation of new software tools and models for learning.

KEY WORDS: DIGITAL IMAGES, SEGMENTATION OF IMAGES, CHARACTERISTICS, TECHNOLOGIES OF DIGITAL IMAGE PROCESSING.

ЗМІСТ

Вступ.....	7
1 Системи та засоби опрацювання цифрових зображень	10
1.1 Цифрове зображення, класифікація та основні характеристики	10
1.2 Цифрова обробка зображень особливості та сфери застосування.....	18
1.3 Програмні засоби обробки та опису цифрових зображень	27
1.4 Постановка задач дослідження.....	30
1.5 Висновки до розділу	31
2 Методи та алгоритми обробки кольорових цифрових зображень	32
2.1 Методи та алгоритми опису цифрових зображень.....	32
2.2 Алгоритми сегментації цифрових зображень	40
2.3 Алгоритм сегментації на основі аналізу однорідних областей.....	48
2.4 Висновки до розділу	51
3 Програмний додаток аналізу та опису цифрових зображень	52
3.1 Структура програмного додатку опрацювання цифрових зображень	52
3.2 Модулі програмного додатку аналізу та оцінки програмного коду	61
3.3 Тестування та аналіз реалізованого програмного додатку	66
3.4 Висновки до розділу	72
Висновки	73
Список використаної літератури	74
Додаток А Лістинг реалізації алгоритмів цифрової обробки зображень	79
Додаток Б Світлокопії виданих публікацій	82

ВСТУП

Актуальність роботи. Технологія цифрової обробки зображень широко використовується в нашому житті, забезпечуючи велику зручність для нашого життя, виробництва та досліджень. Сьогодні за участю штучного інтелекту цифрова обробка зображень просувається вперед у сферу дослідження алгоритмів, і процес стандартизації відповідної технології обробки зображень буде здійснюватися відповідно. Під впливом розвитку штучного інтелекту та глибокого навчання поступово пропонуються алгоритми покращення зображення на основі нейронної мережі глибокої згортки. Цей тип методу може задовольнити потреби обробки цифрових зображень як у швидкості, так і в точності.

У світі цифрової обробки зображень настала ера швидкого технологічного прогресу, цифрова обробка зображень стала важливою сферою, яка відіграє фундаментальну роль у різноманітних додатках у різних галузях. Від покращення візуальної якості зображень до вилучення важливої інформації, методи обробки цифрових зображень революціонізували спосіб взаємодії та інтерпретації візуальних даних.

Обробка цифрових зображень стосується обробки цифрових зображень за допомогою алгоритмів та існуючих обчислювальних методів. Вона передбачає трансформацію зображень для покращення їх якості, вилучення цінної інформації та полегшення подальшого аналізу. Ці методи виконуються на цифрових зображеннях, які представлені у вигляді сітки пікселів, причому кожен піксель містить інформацію про колір та інтенсивність.

Іншими словами, технологія обробки цифрових зображень використовується майже в усіх аспектах життя людей, створюючи велику соціальну цінність. З безперервним зростанням людських потреб обробка цифрових зображень все ще має багато нових аспектів для вивчення, і вона продовжуватиме вдосконалюватися та розвиватися в кращому напрямку.

Тому обрана задача розробки та реалізації алгоритму сегментації цифрових зображень на основі аналізу характерних ознак однорідних областей є актуальною.

Метою роботи є розробка алгоритму сегментації цифрових зображень на основі аналізу характерних ознак однорідних областей.

Для досягнення даної мети ставились наступні завдання:

- провести класифікацію та виділити основні характеристики цифрових зображень;
- провести аналіз технологій обробки цифрових зображень;
- провести аналітичний огляд програмних систем обробки цифрових зображень;
- проаналізувати існуючі алгоритми сегментації цифрових зображень;
- розробити алгоритм сегментації цифрових зображень на основі аналізу характерних ознак однорідних областей;
- реалізувати програмний додаток обробки цифрових зображень.

Об'єкт дослідження – процес обробки цифрових зображень.

Предмет дослідження – методи і алгоритми сегментації цифрових зображень

Наукова новизна одержаних результатів визначається наступним чином:

- проведено комплексний аналіз та класифікацію алгоритмів сегментації на основі дослідження результатів їх роботи, що дозволило групу характеристик для проектування алгоритму сегментації цифрових зображень;
- розроблено алгоритм сегментації цифрових зображень на основі аналізу однорідних областей, що дозволило підвищити точність виділення однорідних областей.

Практична цінність одержаних результатів полягає в тому, що:

- розроблено структуру додатку обробки цифрових кольорових зображень та проведено її моделювання, що дозволило програмно реалізувати та дослідити запропоновані алгоритми обробки цифрових зображень;

- реалізовано програмне рішення обробки та опису цифрових кольорових зображень на основі об'єктно-орієнтованого підходу та з використанням алгоритмів попередньої обробки.

Публікації та апробація до випускної кваліфікаційної роботи. За результатами наукових досліджень, проведених у випускній кваліфікаційній роботі, підготовлено тези доповіді «Алгоритм сегментації цифрових зображень на основі аналізу характеристик однорідних областей» обсягом 1 сторінка на VIII Науково-практичній конференції молодих вчених і студентів «Інтелектуальні комп'ютерні системи та мережі», а також «Алгоритм опрацювання та аналізу об'єктів на рухомих зображеннях» обсягом 1 сторінка на VIII Науково-практичній конференції молодих вчених і студентів «Інтелектуальні комп'ютерні системи та мережі».

1 СИСТЕМИ ТА ЗАСОБИ ОПРАЦЮВАННЯ ЦИФРОВИХ ЗОБРАЖЕНЬ

1.1 Цифрове зображення, класифікація та основні характеристики

Комбінація кінцевої кількості пікселів, $f(x,y)$ це зображення, де x та y є просторовими координатами, а f це рівень сірого зображення. Говорячи комп'ютерною мовою, картинка – це двовимірний масив із розміром x на y , а самі пікселі розташовані у вигляді рядків і стовпців. Є два типи зображень які розрізняються по типу фізично сприйняття та відтворення (рисунок 1.1):



Рисунок 1.1 – Класифікація методів оцінки якості програмного коду

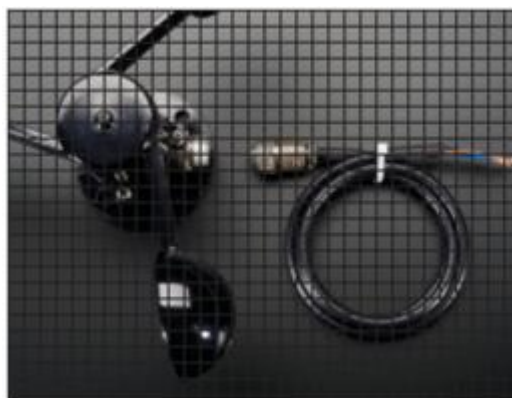
- Аналогові зображення: ці зображення мають безперервний діапазон значень f, x, y .
- Цифрові зображення: ці зображення мають окремі кількості f, x, y .

Цифрове зображення містить кінцевий набір елементів, які називаються пікселями. Вони також відомі як елементи зображення або точки зображення. Зображення представляється у вигляді матриці, де кожен елемент складається з окремих пікселів. Кожен піксель – це найменша точка на зображенні, яка набуває певної непрозорості, кольору або відтінку. Припустимо, у нас є зображення розміром 200 x 200 (ширинаХвисота). Загальна кількість пікселів на зображенні

становить 400 000. На рисунку 1.2 наведено приклад бачення зображення у аналоговому (а) та цифровому представленні (б).



a)



b)

Рисунок 1.2 – Приклад зображення у аналоговому (а) та цифровому представленні (б)

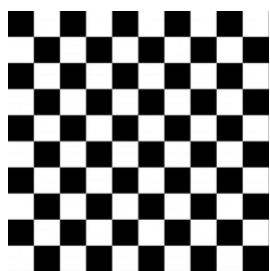
Для відображення цифрових зображень необхідно використовувати техніку/об'єкти, які випромінюють світло, наприклад, на дисплеї комп'ютера. Комп'ютерний монітор складається з прямокутного масиву маленьких точок, які випромінюють світло. У більшості технологій кожна точка насправді складається з трьох менших точок, і кожна з цих менших точок випромінює червоне, зелене та синє світло. Якщо кількість червоного, зеленого та синього змінюється, наш мозок об'єднує світло від трьох маленьких джерел світла та сприймає світло різних кольорів. Таким чином можна керувати кольором кожного набору з трьох точок, і із загальної кількості точок можна побудувати кольорове зображення. Важливо розуміти, що можна створити більшість, але не всі кольори, змішуючи червоний, зелений і синій. Крім того, різні комп'ютерні монітори використовують трохи різні червоний, зелений і синій кольори, і якщо це не враховувати, кольори виглядатимуть по-різному на двох моніторах. Це також означає, що деякі кольори, які можна відобразити на одному моніторі, можуть не відображатися на іншому моніторі. Принтери використовують той самий принцип побудови зображення з маленьких точок. Однак на більшості

принтерів малі точки не складаються з менших крапок різних кольорів. Замість цього до потрібного кольору змішується 7–8 різних чорнил (або подібних речовин). Це дає змогу виробляти широкий діапазон кольорів, але не всі, і проблема підбору кольору з іншого пристрою, наприклад монітора, принаймні така ж складна, як і підбір різних кольорів на різних моніторах. Відеопроєктор створює зображення, яке проектується на стіну. Отже, кінцеве зображення є відбитим зображенням, і важливо, щоб поверхня була білою, щоб вона однаково відображала всі кольори. Якість пристрою тісно пов'язана з щільністю точок.

Двома найпоширенішими способами отримання цифрових зображень є цифрова камера або сканер. Сканер фактично робить фотографію документа у формі прямокутного масиву (можливо, кольорових) точок. Що стосується принтерів, то важливим показником якості є кількість точок на дюйм. Роздільна здатність сканера зазвичай коливається в діапазоні від 75 dpi до 9600 dpi, а колір представлено до 48 біт на точку. Для цифрових камер немає сенсу вимірювати роздільну здатність у точках на дюйм, оскільки це залежить від способу друку зображення (його розміру). Натомість роздільна здатність вимірюється кількістю записаних точок. Кількість пікселів, записаних цифровою камерою, зазвичай коливається в діапазоні від 320×240 до 6000×4000 з 24 бітами інформації про колір на піксель. Загальна кількість пікселів коливається в діапазоні від 76 800 до 24 000 000 (0,077 мегапікселя до 24 мегапікселя). Для сканерів і камер легко подумати, що чим більше точок (пікселів), тим краща якість. Хоча в цьому є частка правди, є багато інших факторів, які впливають на якість. Основна проблема полягає в тому, що виміряна інформація про колір дуже легко забруднюється шумом. І, звичайно, висока роздільна здатність також означає, що отримані файли стають дуже великими; нестисне зображення 6000×4000 створює файл розміром 72 МБ. Перевага високої роздільної здатності полягає в тому, що ви можете значно збільшити зображення, зберігаючи прийнятну якість.

Після отримання цифрового зображення від фотофіксуючої апаратури воно може бути в одному з основних форматів зберігання даних. Для класифікації таких зображень одним з критеріїв є кількість інформації

яка використовується для задання значення кольору в одній точці цифрового зображення. На рисунку 1.3 наведено різні типи цифрових зображень



а)



б)



в)

Рисунок 1.3 – Види цифрових зображень: бінарне (а), в градаціях сірого (б) та кольорове (в)

В загальному випадку, значення $f(x,y)$ дає інформацію про колір у точці (x,y) . Важливо пам'ятати, що для цього існує багато форматів. Найпростішим випадком є звичайні чорно-білі (бінарні) зображення, у яких $f(x,y)$ дорівнює або 0, або 1.

Зображення в градаціях сірого. У сірому зображенні піксель є цілим числом від 0 до 255. Нуль повністю чорний, а 255 білий. Сірий падає близько 127, що призводить до білого кольору. Ці зображення також відомі як зображення 8-бітного кольорового формату. Двійкові зображення та зображення у відтінках сірого є двовимірними масивами, тому маємо одне значення для представлення пікселя в кожному місці.

Зображення RGB. Ці зображення представлені у вигляді 3-D матриці або 3-вимірного масиву. Ці зображення використовують формат високого кольору, відомий як 16-бітний формат кольору. RGBA є розширенням RGB з новим полем альфа, яке використовується для представлення непрозорості зображення.

У зображеннях RGB піксель складається з трьох цілих чисел, а діапазон значень становить від 0 до 255. Ці зображення складаються з трьох основних кольорів RGB. R означає червоний, G означає зелений, а B означає синій. Ці зображення RGB також відомі як кольорові зображення.

Іншим критерієм поділу цифрових зображень є принцип за яким вони зберігаються в пам'яті комп'ютера. За таким поділом зображення бувають растрові та векторні. Приклад збереження одного і того ж зображення у різних форматах наведено на рисунку

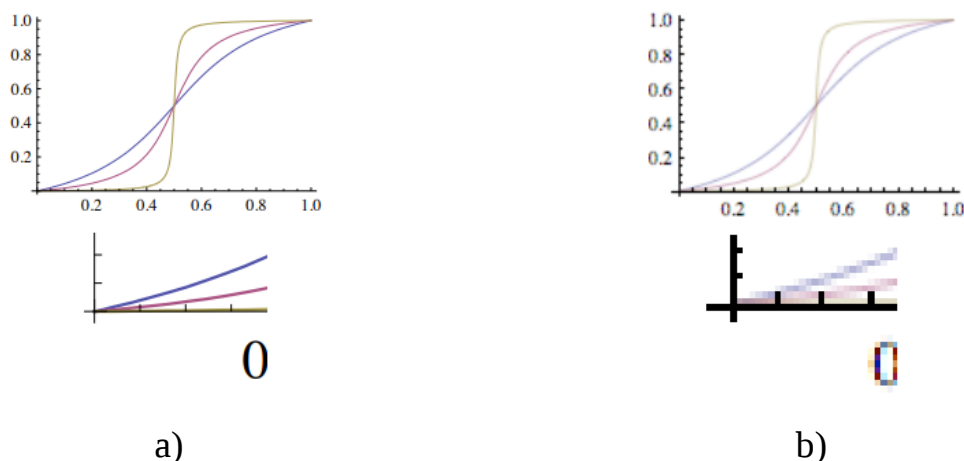


Рисунок 1.4 – Приклад зображення у векторному форматі (а) та растровому форматі (б)

Все що виводиться на моніторі комп'ютера або на принтері, складається з маленьких точок. Це ідеальний варіант для цифрових зображень, які також складаються з великої кількості маленьких точок. Однак, коли збільшуємо зображення, крапки на зображенні стають видимими. Крім зображень, текст і різні види штрихового мистецтва (малюнки) також відображаються на моніторах і друкуються на принтерах, і тому повинні бути представлені у вигляді точок. Однак існує велика різниця в тому, як зберігаються такі види графічних зображень. Для такого випадку використовується інший підхід для збереження інформації про елементи на зображенні. Різниця полягає в тому, що в другому випадку інформація представлена в термінах математичних примітивів, таких як лінії та криві. Такий спосіб називають векторним зображенням або векторною графікою. Перевага векторної графіки полягає в тому, що фактичні точки, які будуть використовуватися, не визначаються, поки не буде намальовано фігуру. У форматах векторної графіки всі елементи малюнка представлені в термінах

математичних примітивів. Це включає всі лінії та криві, а також текст. Лінія зазвичай представлена двома кінцевими точками та шириною. Криві фігури представлені у вигляді коротких з'єднаних відрізків ліній або гладко з'єднаних поліноміальних відрізків кривої. Кожного разу, коли зображення виводиться на моніторі чи принтері, фактичні крапки для друку визначаються з математичного представлення. Зокрема, це стосується шрифтів (графічних форм символів), які зазвичай представляють у вигляді квадратичних або кубічних кривих багаточленів (так звані криві Без'є), дивіться приклад на рисунку 1.5.

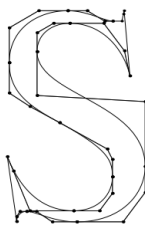


Рисунок 1.5 – Приклад векторизації зображення літери

Двома найпоширенішими форматами векторної графіки є Postscript і PDF, які є форматами для представлення двовимірної графіки. Існують також стандарти для тривимірної графіки, але вони не є загальноприйнятими.

Postscript – мова програмування, розроблена Adobe Systems на початку 1980-х років. Її основне застосування – представлення зображень сторінок, тобто інформації, яка може бути відображена на моніторі або надрукована на принтері. Основними примітивами в Postscript є прямі сегменти та кубічні багаточленні криві, які часто з'єднуються (плавно) разом, щоб утворити більш складні форми. Шрифти Postscript складаються з програм Postscript, які визначають контури форм усіх символів у шрифті. Щоразу, коли програмі Файли Postscript – це стандартні текстові файли, тому програму, яка створює сторінку, можна перевіряти (і редагувати) у стандартному редакторі. Недоліком цього є те, що файли Postscript кодуються неефективно та вимагають багато місця для зберігання. Файли Postscript мають розширення .eps або .ps. Оскільки багато

сторінок містять зображення, Postscript також підтримує включення растрової графіки на сторінку.

Іншим більш поширеним форматом є PDF. Portable Document Format – це стандарт для представлення зображень сторінок, також розроблений Adobe. На відміну від Postscript, якому для правильного відображення сторінки може знадобитися зовнішня інформація, наприклад бібліотеки шрифтів, pdf-файл містить усю необхідну інформацію в собі. Він підтримує ті самі математичні примітиви, що й Postscript, але кодує інформацію в компактному форматі. Оскільки сторінка може містити зображення, можна також зберегти цифрове зображення у форматі pdf. Pdf-файли можуть бути заблоковані, щоб їх неможливо було змінити. Pdf широко використовується на комп'ютерних платформах і є кращим форматом для обміну документами.

Для групи растрових зображень існують власні формати, які дозволяють зберігати зображення з зазначеною точністю. Існує багато форматів для представлення цифрових зображень. У більшості кольорових зображень колір пікселя представлений кількістю червоного, зеленого та синього або (r, g, b) . Кожне з цих чисел зазвичай представлене вісьмома бітами і може приймати цілі значення в діапазоні 0–255. Іншими словами, інформація про колір кожного пікселя вимагає трьох байтів. Коли кольорові зображення та монітори стали загальнодоступними, розмір файлу 24-бітного зображення був дуже великим порівняно з розміром жорстких дисків і доступної пам'яті комп'ютера. Тому замість того, щоб зберігати всі 24 біти інформації про колір, було прийнято створити таблицю з 256 кольорів, за допомогою якої дане зображення може бути представлено досить добре. Замість того, щоб зберігати 24 біти, можна просто зберегти таблицю на початку файлу, а в кожному пікселі вісім бітів, що відповідають правильному запису в таблиці. Зазвичай це називають восьмибітним кольором, а таблицю називають пошуковою таблицею або палітрою. Для великих фотографій 256 кольорів далеко недостатньо для отримання розумного відтворення кольорів. Зображення можуть містити велику кількість даних і мати великий потенціал для стиснення як без втрат, так і з

втратами. Для стиснення з втратами використовуються стратегії, подібні до тих, що використовуються для стиснення аудіо. Це означає, що дані перетворюються за допомогою dct або вейвлет-перетворення (ці перетворення легко узагальнюються на зображення), малі значення встановлюються на нуль, а отримані дані кодуються за допомогою алгоритму кодування без втрат. Як і аудіоформати, формати зображень зазвичай містять таку інформацію, як роздільна здатність, час запису зображення та подібну інформацію на початку файлу.

Формат Graphics Interchange Format (GIF) був представлений компактне представлення кольорових зображень. Він використовує палітру щонайбільше з 256 кольорів, взятих із 24-розрядної моделі кольорів, як пояснено вище. Це означає, що він не підходить для кольорових зображень із безперервними кольоровими тонами, але він досить добре працює для менших зображень із великими ділянками постійного кольору, як-от логотипи та кнопки на веб-сторінках. Gif-файли кодуються без втрат за допомогою варіанту алгоритму Лемпеля-Зіва-Велча. Розширення GIF-файлів - .gif.

Формат файлу зображень із тегами (TIFF) – це гнучкий формат зображень, який може містити кілька зображень різних типів в одному файлі за допомогою так званих «тегів». TIFF підтримує стиснення зображень без втрат за допомогою стиснення Лемпеля-Зіва-Велча, але також може містити зображення, стислі JPEG. TIFF спочатку був розроблений як формат для сканованих документів і підтримує зображення з одnobітними піксельними значеннями (чорно-білі). Він також підтримує розширені формати даних, наприклад більше восьми біт на компонент кольору. TIFF-файли мають розширення .tiff.

Joint Photographic Experts Group (JPEG) – це формат зображення, затверджений як міжнародний стандарт у 1994 році. JPEG, як правило, має втрату, але також може бути без втрат і став популярним форматом для представлення зображень в Інтернеті. Стандарт визначає як алгоритми кодування та декодування, так і формат зберігання. Jpeg ділить зображення на блоки 8×8 і перетворює кожен блок за допомогою дискретного косинусного

перетворення. Ці значення, що відповідають вищим частотам (швидкі зміни кольору), потім встановлюються на 0, якщо вони не є досить великими, оскільки людське сприйняття не помічає цього. Збурені значення dct потім кодуються різновидом кодування Хаффмана. Jpeg також може використовувати арифметичне кодування, але це збільшує час як кодування, так і декодування, лише приблизно на 5 % покращення коефіцієнта стиснення. Рівень стиснення зображень jpeg вибирається користувачем і може призвести до помітних артефактів, якщо встановлено занадто високо. Jpeg особливо схильний до артефактів у областях, де інтенсивність швидко змінюється від пікселя до пікселя. Jpeg-файл має розширення .jpg або .jpeg.

Portable Network Graphics (PNG) – це формат зображень без втрат. Png не створювався для професійного використання, а скоріше для передачі зображень в Інтернеті та підтримує лише зображення з рівнем сірого та зображення rgb (також кольорові зображення на основі палітри). Png було створено, щоб уникнути патенту на алгоритм lzw, який використовується в gif, а також обмеження gif восьмибітовою інформацією про колір. Для ефективного кодування png може (це опція) передбачити значення пікселя зі значення попередніх пікселів і відняти прогнозоване значення від фактичного значення. Потім він може кодувати ці значення помилок за допомогою методу кодування без втрат під назвою deflate, який використовує комбінацію алгоритму lz77 і кодування Хаффмана. Це схоже на алгоритм, який використовується в аудіоформатах без втрат, таких як Apple Lossless і flac.

1.2 Цифрова обробка зображень особливості та сфери застосування

Технологія обробки цифрових зображень – це метод перетворення сигналів зображення в цифрові сигнали з подальшим використанням комп'ютерної обробки для досягнення певної мети модифікації зображення.

Швидкий розвиток обчислювальної техніки та математики призвів до вдосконалення та вдосконалення технології обробки цифрових зображень. Покращення якості зображення, аналіз зображення та реконструкція зображення є його трьома компонентами (рисунок 1.6).

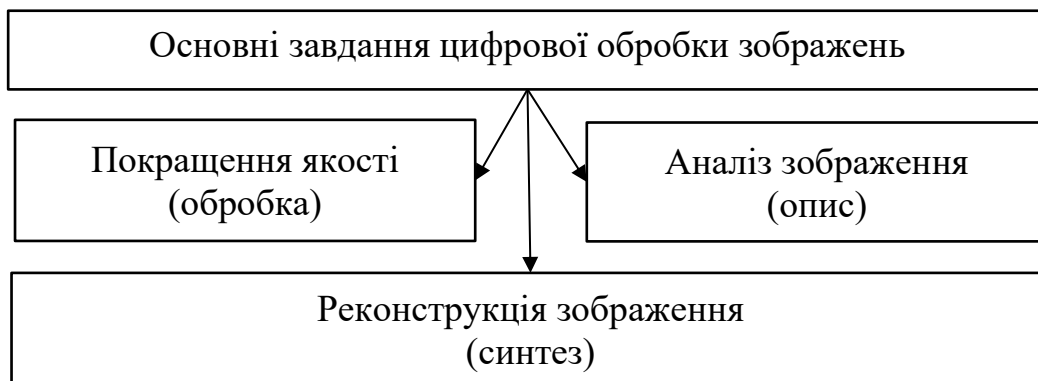


Рисунок 1.6 – Основні задачі цифрової обробки зображень

Оскільки цифрові технології можуть обробляти лише лінійні процеси, оптичну обробку зображень можна застосовувати лише в обмеженій кількості сценаріїв. Цифрова обробка зображень також здатна реалізувати нелінійну обробку, дозволяючи представляти всі операції над математичними формулами, а також логічні зв'язки процедур обробки зображень. Тепер обробку зображень більш високого рівня можна реалізувати за допомогою алгоритмів штучного інтелекту, таких як нейронні мережі. Обробка цифрових зображень в основному складається з наступних чотирьох аспектів. По-перше, оцифровка зображення, яка є основним кроком перед обробкою зображення комп'ютером, полягає в перетворенні реального зображення у формат зберігання, прийнятний комп'ютером. Процес оцифрування зображення ділиться на два етапи: дискретизація та квантування. Вибірка стосується процесу зміни безперервних величин у часовій області або просторовій області на дискретні величини, тоді як квантування стосується процесу наближення безперервних значень сигналів до обмеженої кількості дискретних значень. По-друге, кореляція цифрових зображень є однією з двох фундаментальних характеристик стиснення цифрових

зображень. Іншими словами, часто існує значний зв'язок між сусідніми пікселями та їхніми відповідними пікселями в суміжних кадрах. Усунення або зменшення цих кореляцій також призведе до зменшення надмірності в інформації зображення, що означає, що реалізується стиснення цифрових зображень. Інший – візуальні психологічні особливості людини. Людський зір нечутливий до слабкої роздільної здатності кольорів і різких змін країв. Використання цих властивостей може належним чином знизити точність кодування у відносних частинах, через що люди не відчують погіршення якості зображення візуально, щоб реалізувати мету стиснення цифрового зображення. По-третє, це покращення зображення, яке використовує різні види математичних засобів і засобів перетворення для покращення коефіцієнта контрастності та чіткості зображення, щоб оброблене зображення було більш придатним для візуальних характеристик людини або машинних систем розпізнавання за допомогою цього методу. По-четверте, сегментація зображення полягає в тому, щоб розділити зображення на кілька окремих областей з унікальними властивостями, кожна з яких є безперервним набором пікселів.

У розвитку цифрової обробки зображень існує багато факторів, які обмежують її розвиток. Наприклад, обробка цифрових зображень має великий обсяг інформації, займає обчислювальну швидкість комп'ютера та має певні вимоги до ємності для зберігання. При цьому важко обробити смугу частот. Оскільки розвиток цифрової обробки зображень має характеристики зображень у високочастотному діапазоні, це особливо важливо для обробки, зберігання та відображення в діапазоні частот. По-друге, оскільки зображення є двовимірною проекцією тривимірного об'єкта, двовимірне зображення не може відновити всю стереоскопічну інформацію об'єкта, тому це ускладнює розпізнавання зображення на двовимірному етапі. Крім того, звичайно, є багато проблем в цифровій обробці зображень, які повинні бути вирішені дослідниками. Основна причина цих проблем полягає в тому, що часто обробляють і аналізують зображення не люди, а машини чи комп'ютери. Цей несуб'єктивний спосіб мислення за допомогою коду змушує комп'ютери часто не мати можливості

оцінювати зображення через людське сприйняття, що також є однією з проблем, з якими стикається сфера обробки цифрових зображень і штучного інтелекту. Досліджуючи сучасні наукові праці можна виділити ряд напрямів (рисунок 1.7) в яких цифрова обробка зображень досягла основних здобутків та продовжує активно розвиватись.

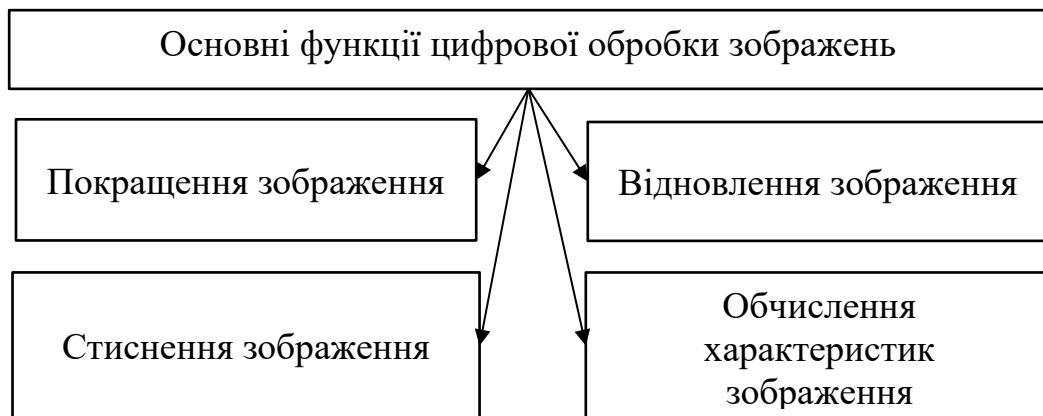


Рисунок 1.7 – Основні функції цифрової обробки зображень

Покращення зображення. Методи покращення зображення спрямовані на покращення візуальної якості зображень шляхом регулювання яскравості, контрасту, колірного балансу та різкості. Вирівнювання гістограми, розтягування контрасту та фільтрація просторової області є поширеними методами, які використовуються для покращення зображення.

Відновлення зображення. Відновлення зображення зосереджується на видаленні або зменшенні деградації, викликаній такими факторами, як шум, розмиття або артефакти стиснення. Методи відновлення використовують методи фільтрації, алгоритми деконволюції та статистичні моделі для відновлення деталей вихідного зображення.

Стиснення зображення. Стиснення зображень зменшує розмір цифрових зображень, щоб заощадити місце для зберігання та забезпечити ефективну передачу через мережі. Використовуються методи стиснення без втрат і

стиснення з втратами, причому останній приносить певну шкоду якості зображення для досягнення вищого коефіцієнта стиснення.

Витяг характеристик. Метою виділення ознак є витяг значущої інформації із зображень, щоб полегшити завдання розпізнавання образів і класифікації. Різноманітні математичні та статистичні методи використовуються для визначення відповідних характеристик, таких як текстура, колір, форма та краї.

На сьогоднішній день методи та алгоритми використовуються для вирішення практичних задач в найрізноманітніших галузях (рисунок 1.8). Обробка цифрових зображень у техніці дистанційного зондування Технологія дистанційного зондування – це метод, який збирає інформацію про електромагнітне випромінювання об'єктів на штучних супутниках, а потім визначає середовище та ресурси землі. Сучасна технологія дистанційного зондування в основному включає в себе отримання, передачу, зберігання та обробку інформації, серед яких обладнання для обробки інформації включає синтезатори кольорів, зчитувачі зображень і цифровий процесор зображень. Серед них дистанційне зондування з гіперспектральною роздільною здатністю розвинулося в спектроскопії зображень.

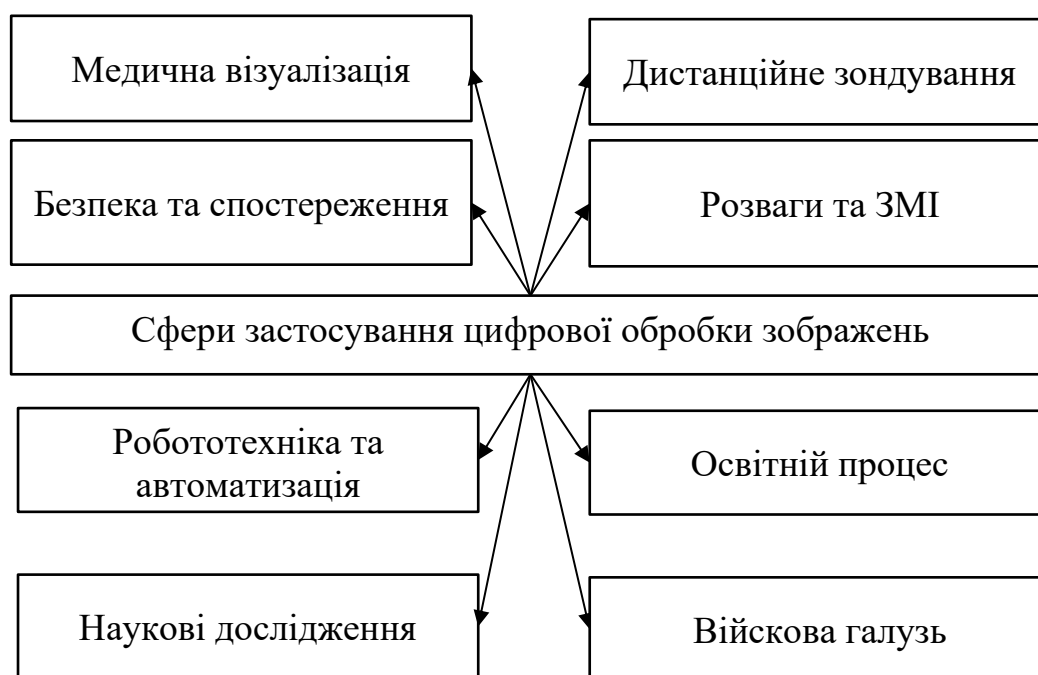


Рисунок 1.8 – Основні сфери застосування цифрової обробки зображень

У порівнянні з традиційним дистанційним зондуванням, спектрометр зображення з гіперспектральною роздільною здатністю забезпечує дуже вузьку смугу зображення для кожного пікселя зображення, і вона приблизно безперервно розподіляється в певному спектральному інтервалі, що може збільшити кількість просторової інформації спектру наземного об'єкта. В даний час радар, який може створювати зображення з високою роздільною здатністю, - це в основному радар із синтетичною апертурою (SAR), який може отримувати радіолокаційні зображення високої роздільної здатності, подібні до оптичної фотографії в метеорологічних умовах надзвичайно низької видимості. Менша справжня апертура антени синтезується в радарі з більшою еквівалентною апертурою антени з обробкою даних за допомогою відносного руху між радаром і ціллю. Зараз розробка sar в основному зосереджена на створенні зображень і видаленні шумів. Щоб отримати корисну інформацію з цього насиченого інформацією зображення та відновити атрибути його фізичного поля, дані зображень індикатора гіперспектральної горизонтальної ситуації (HSI) [5] і цифрової моделі поверхні (DSM) [6] зазвичай мають нижчу просторову роздільну здатність і більш складні типи покриття, а також різні смуги HSI сильно корельовані та надлишкові, тому важко розширити традиційну модель обробки зображень RGB незалежного ідентичного розподілу (IID) до зображень HSI. Jin та ін. запропонували метод обробки та отримання інформації з даних HSI за сотнями каналів [13]. Використовуючи сферичне гармонічне освітлення, вперше була запропонована зручна модель візуалізації HSI та DSM, яка представляє взаємодію відбивної здатності матеріалу, геометричного розподілу та освітлення навколишнього середовища. Введення попередньої теорії локальної та нелокальної відбивної здатності для забезпечення локальної плавності та глобальної узгодженості відновленої відбивної здатності забезпечить автентичність відновлення гіперспектрального зображення. Крім того, через чутливість детектування гіперспектральні зображення часто мають шум, який заважає сигналу. Щоб усунути ці перешкоди запропонували шумопоглинач із хорошою стійкістю та ефективністю виявлення. Завдяки

попередній обробці вихідної інформації можна проаналізувати багато видів характерної інформації із зображення. Основними типами аналізу є аналіз рельєфу [5], аналіз рослинності [6] та аналіз хмар [7], які є похідними від широкого використання цифрової моделі рельєфу з високою роздільною здатністю. У той же час завдяки корекції форми та покращенню стійкості до кольору ці зображення з високою роздільною здатністю все ще можуть зберігати більшу частину корисної інформації після повторної обробки.

Обробка цифрових зображень у технологіях медичної візуалізації Завдяки інтеграції численних нових технологій із сфер інформатики, фізики, біомедицини та інших галузей технологія медичної візуалізації є важливим інструментом для медичних досліджень і клінічної медицини. За останні 20 років вона перетворилася на галузь медичних технологій, яка швидко розвивається. Одним із найактивніших напрямків досліджень є використання комп'ютерної обробки зображень для допомоги в медичній діагностиці та лікуванні. Це значно підвищило точність діагностики та лікування. Відповідно до фактичного застосування клінічної медицини, технологію медичних зображень можна розділити на дві категорії: технологію структурної візуалізації та технологію функціональної візуалізації. Технологія структурної візуалізації в основному застосовується для отримання внутрішніх структурних зображень тканин, які потрібно розрізати в організмі людини. Застосування цієї технології дозволяє лікарям зручно спостерігати за реальним станом пацієнтів. Технологія функціональної візуалізації використовує зображення як носій інформації та відображає роботу деяких органів у пацієнтів через зміни зображень. Використання цифрових зображень у лікуванні в основному відображається у вилученні зображення [18], покращенні зображення [19], сприйнятті зображення тощо, що має на меті забезпечити зручність для лікарів у діагностиці та певною мірою підвищити точність діагностики. Сегментація зображень, реєстрація медичних зображень і злиття медичних зображень – це приклади технології обробки медичних зображень, яка використовується в процес діагностики. Технологія сегментації зображення – це процес сегментації ураженої тканини з

подальшим пошуком. Цей процес вимагає, щоб машина завершила пошук і аналіз ураженої тканини, щоб її можна було впорядкувати відповідно до кореляції. Щоб переконатися, що цей пошук також є точним і надійним для зображення рухомих об'єктів, запропоновано нову багаторакурсну мережу оцінки руху для точної оцінки стану руху серцевого м'яза, яка об'єднує двовимірні зображення з відхиленням загального режиму (CMR), отримані на площинах короткої осі та довгої осі [11]. У цьому методі за допомогою злиття зображень із кількома видами створюється 2D-3D гібридна мережа для створення щільного тривимірного спортивного поля. Щоб забезпечити узгодженість оцінки 3D-руху, вони запровадили модуль регуляризації форми в процесі навчання та використовували інформацію про форму в мультипроекційних зображеннях для слабкого контролю за оцінкою 3D-руху, щоб забезпечити найкращу точність алгоритму. Правильне збереження та конфіденційність електронної медичної інформації також поступово входить у поле зору людей. Точність верифікації медичного зображення може бути покращена, що є перевагою для забезпечення приховування та автентичності.

Обробка цифрових зображень у системі розпізнавання рис людини

Технологія розпізнавання рис людини загалом відноситься до процесу ідентифікації унікальних фізіологічних характеристик людини для визначення пароля. Зараз основна технологія розпізнавання включає в себе розпізнавання відбитків пальців і розпізнавання обличчя. Для розпізнавання відбитків пальців його складність текстури може забезпечувати потребу в розпізнаванні, але щоб компенсувати дефекти якості зображення, часто необхідно використовувати технологію покращення зображення, щоб покращити його надійність. Використання технології цифрової обробки зображень може усунути шум, підвищити контрастність ліній хребта та долини та покращити якість зображення для досягнення кращих результатів розпізнавання. Для розпізнавання обличчя він більше зосереджується на алгоритмі розпізнавання, який, у свою чергу, є розпізнаванням обличчя на основі алгебраїчних ознак, розпізнаванням обличчя з огляду на геометричні особливості та розпізнаванням обличчя на основі

механізму з'єднання. У цьому процесі технологія цифрової обробки зображень в основному відображається на доповненні яскравості зображення, покращенні контрастності та різкості та забезпеченні правильної швидкості розпізнавання машиною. Зараз із швидким розвитком різноманітних технологій розпізнавання з'являється все більше нових технологій, заснованих на розпізнаванні ознак. Маніпулюючи зображеннями обличчя, особливо обміном обличчям, деформацією та модифікацією, шаблон інформації про зображення розпізнавання людини можна додатково проаналізувати, що корисно для вдосконалення алгоритму розпізнавання обличчя. На додаток до розпізнавання обличчя та інших модулів, шкіру також можна безпосередньо застосовувати для вивчення стану людини.

Обробка цифрових зображень у інтелектуальному транспорті У дослідженні інтелектуальних транспортних систем важливим способом покращення реалізації інформації про дорожній рух є розробка обладнання для збору інформації про дорожній рух, яке може ефективно отримувати робочу інформацію на дорозі, включаючи потік транспортних засобів, швидкість транспортних засобів, класифікація транспортних засобів, інтенсивність руху та інше інформації. Крім того, у технології без драйверів цифрова обробка зображень також є особливо важливою. Це не тільки для ідентифікації транспортних засобів, але також має бути точним щодо номерних знаків, під'їзних шляхів і пішоходів. З одного боку, розпізнавання цих аспектів залежить від штучного інтелекту. З іншого боку, це також потрібно застосовувати до цифрової обробки зображень, наприклад, периферійних обчислень, які зараз є гарячою точкою досліджень. У той же час, з огляду на інформаційну та інтелектуальну тенденцію розвитку систем управління дорожнім рухом, шляхом дослідження характеристик номерних знаків і технології визначення місця розташування, також пропонується система розпізнавання номерних знаків транспортних засобів. Окрім базового розпізнавання та обчислення транспортних засобів, тепер технологія обробки цифрових зображень також використовується для виявлення несправностей на дорогах. Також цифрова

обробка використовується для обробки зображень телевізійних камер замкнутого циклу для визначення умов дорожнього руху із зображень дорожнього руху на дорозі, щоб користувачі могли заздалегідь перевірити умови дорожнього руху в певний час і здійснювати планування дорожнього руху або контроль світлофорів на перехрестях.

Розваги та засоби масової інформації. Індустрія розваг використовує цифрову обробку зображень для спеціальних ефектів, редагування відео та корекції кольорів у фільмах, телешоу та відеоіграх. Це збагачує візуальний досвід для аудиторії в усьому світі.

Робототехніка та автоматизація. У робототехніці обробка зображень полегшує розпізнавати такі завдання як навігація та планування шляху. Це дозволяє роботам інтерпретувати візуальну інформацію з навколишнього середовища, що веде до прогресу в автоматизації та автономних системах.

Цифрова обробка зображень вийшла за межі свого раннього коріння і стала невід'ємною частиною сучасних технологій. Його застосування охоплює різноманітні сфери, від медицини та дистанційного зондування до розваг і робототехніки. Оскільки технологія продовжує розвиватися, обробка цифрових зображень, безсумнівно, відіграватиме ще більшу роль у формуванні нашого майбутнього, дозволяючи нам взаємодіяти з візуальними даними новими та цікавими способами. Оскільки сучасні дослідники та інженери продовжують розширювати межі цієї галузі, потенціал для інновацій та відкриттів залишається безмежним.

1.3 Програмні засоби обробки та опису цифрових зображень

Для обробки цифрових зображень реалізовано велику кількість різноманітних програмних засобів та цифрових бібліотек. Кожна з них має свої переваги, недоліки та сфери застосування. Проведемо аналіз програмних засобів які найчастіше використовуються при реалізації програмних систем.

OpenCV. Мультиплатформенна бібліотека широкого використання, проста у використанні. Він чудово працює з C++ і Python і охоплює всі ключові методи й алгоритми для виконання кількох робіт із обробки зображень і відео.

CUDA. Основу для паралельних обчислень забезпечує NVIDIA, яка проста в програмуванні, високоефективна та швидка. Використовуючи потужність GPU, він забезпечує фантастичну продуктивність. Бібліотека NVIDIA Performance Primitives, яка включає вибір операцій обробки зображень, сигналів і відео, включена в його набір інструментів.

SimpleCV. Система для створення програм комп'ютерного зору називається SimpleCV. Він надає доступ до кількох технологій комп'ютерного зору на будь-яких платформах OpenCV, pygame та інших. Це інструмент, який вам потрібен, якщо ви не хочете вивчати тонкощі обробки зображень і просто хочете виконати роботу. SimpleCV – найкращий варіант, якщо вам потрібно швидко створити прототип.

Keras. Бібліотека глибокого навчання Python під назвою Keras інтегрує компоненти з кількох інших бібліотек, таких як Tensorflow, Theano та CNTK. Оскільки Keras побудовано на основі Tensorflow, він має перевагу перед такими конкурентами, як Scikit-learn і PyTorch. TensorFlow, Microsoft Cognitive Toolkit, Theano або PlaidML можна використовувати з Keras. Він зосереджений на зручності, кількісно визначеній якості та розширюваності та розроблений для швидких експериментів із глибокими нейронними мережами. Найкращі методи зменшення когнітивного навантаження дотримуються Keras, який забезпечує стабільні та базові API та обмежує кількість взаємодії з користувачем, необхідну для типових випадків використання.

EmguCV – це кросплатформне доповнення .Net для обробки зображень для OpenCV. Він працює з мовами, сумісними з .NET, такими як IronPython, C#, VB і VC++. Він сумісний з Windows, Linux, Mac OS, iOS і Android. Він також сумісний з Unity, Xamarin Studio та Visual Studio.

GPUImage. Це структура, побудована на OpenGL ES 2.0, яка дає змогу додавати ефекти та канали, прискорені GPU, до живого відео, нерухомих

зображень і відео. Потрібно написати багато коду, щоб налаштувати та підтримувати роботу користувацьких каналів на GPU.

YOLO. Метод виявлення об'єктів під назвою «You Just Look Once» (YOLO) був розроблений спеціально для обробки в реальному часі. Дослідники Вашингтонського університету Джозеф Редмон і Алі Фархаді розробили YOLO, складну систему виявлення об'єктів у реальному часі. Їхня техніка накладає нейромережу на все зображення, розбиваючи його на сітку та залишаючи відбитки районів з виявленими об'єктами.

VXL – це група безкоштовних бібліотек C++. Функціональні можливості цього засобу обробки зображень такі:

- Відкривайте, зберігайте та редагуйте фотографії в різних широко використовуваних типах файлів, включаючи великі зображення.
- Геометрія в 1, 2 або 3 вимірах для точок, кривих та інших основних об'єктів.
- Фізика камери.
- Відновлення стабільності після руху.
- Створення візуального інтерфейсу користувача.
- Аналіз цифрових зображень.
- Створення шаблонів обробки зображень.
- 3D зображення топології.

BoofCV BoofCv — це ліцензований Apache 2.0 фреймворк Java з відкритим вихідним кодом для робототехніки та програм комп'ютерного бачення в реальному часі, який можна використовувати в комерційних і академічних цілях. Його можливості включають вирівнювання камери, структуру від руху, відстеження функцій і оптимізовані алгоритми низькорівневої обробки зображень.

Для проведення порівняння існуючих програмних засобів було обрано ряд узагальнюючих критеріїв та проведено їх оцінювання за п'ятибальною шкалою. Результати порівняльного аналізу було згруповано та наведено в таблиці 1.4.

Таблиця 1.4 – Програмні засоби для обробки цифрових зображень

Назва програмного засобу	Тип технології	Зручність користування	Оцінка отриманого програмного коду	Підтримка від розробників	Загальна оцінка
OpenCV	бібліотека	5	5	5	5
CUDA	технологія	4	4	4	4
Keras	бібліотека	4	4	4	4
SimpleCV	бібліотека	4	5	4	4,6
EmguCV	бібліотека	4	4	4	4
GPUImage	технологія	4	4	4	4
YOLO	технологія	5	4	3	4
VXL	бібліотека	5	5	4	4,8

1.4 Постановка задач дослідження

В даному розділі було Д першому розділі було проведено дослідження та аналіз пробметної області, об'єкту дослідження та огляд відомих програмних рішень, які використовуються для вирішення завдань цифрової обробки зображень. Проведене дослідження підходів до вводу та виводу цифрових зображень у комп'ютерні системи показало основні переваги та недоліки використання саме такого типу даних. Використання цифрових зображень збільшити інформаційну складову повідомлень та дозволить отримувати інформацію, яку не можна отримати за допомогою людського ока. Проведений аналіз основних завдань та сфер використання технологій обробки показав високу поширеність даних технологій в сучасних програмно-апаратних системах. Аналіз програмних засобів для обробки , аналізу та опису цифрових зображень показав поширеність та різноманітність засобів даного типу у сучасних технологічних процесах. Програмні системи використовуються як для простих задач моніторингу та навчання так і для складних – медицина, космічні дослідження, військові завдання.

Для досягнення поставленої мети необхідно розв'язати наступні задачі.

- провести класифікацію та виділити основні характеристики цифрових зображень;
- провести аналіз технологій обробки цифрових зображень;
- провести аналітичний огляд програмних систем обробки цифрових зображень;
- проаналізувати існуючі алгоритми сегментації цифрових зображень;
- розробити алгоритм сегментації цифрових зображень на основі аналізу характерних ознак однорідних областей;
- реалізувати програмний додаток обробки цифрових зображень.

1.5 Висновки до розділу

Проведено дослідження та класифікацію цифрових зображень на основі технологій їх формування та збереження, що дозволило виділити основні характеристики які містять максимальну кількість інформації про цифрове зображення;

Здійснено аналіз технологій обробки цифрових зображень на основі сфер їх застосування, що дозволило виділити основні вимоги до алгоритмів сегментації й цифрових зображень;

Проведено аналітичний огляд програмних засобів обробки та аналізу цифрових зображень на основі дослідження їх основних складових та функціональних можливостей, що дозволило визначити основні структурні елементи які реалізуються в сучасних програмних системах обробки цифрових.

2 МЕТОДИ ТА АЛГОРИТМИ ОБРОБКИ КОЛЬОРОВИХ ЦИФРОВИХ ЗОБРАЖЕНЬ

2.1 Методи та алгоритми опису цифрових зображень

Анотація зображень – це практика призначення міток зображенню або набору зображень. Людина-оператор переглядає набір зображень, визначає релевантні об'єкти на кожному зображенні та коментує зображення, вказуючи, наприклад, форму та мітку кожного об'єкта. Ці анотації можна використовувати для створення навчального набору даних для моделей комп'ютерного зору. Модель використовує анотації людини як свою основну істину та використовує їх, щоб навчитися виявляти об'єкти або самостійно позначати зображення. Цей процес можна використовувати для навчання моделей таким завданням, як класифікація зображень, розпізнавання об'єктів і сегментація зображень.

Кількість міток, призначених зображенню, може змінюватися залежно від типу та обсягу проекту. У деяких випадках однієї мітки достатньо для представлення всього зображення. В інших випадках анотатори ідентифікують конкретні об'єкти, сегментують зображення на відповідні регіони або визначають орієнтири, які є конкретними точками інтересу на зображенні. Щоб забезпечити точність маркування, прийнято дозволяти кільком анотаторам позначати те саме зображення, при цьому більшість голосів вибирає найвірогіднішу правильну позначку. Проекти анотації зображень включають широкомасштабну анотацію зображень командами анотаторів. Анотатори повинні бути добре навчені вимогам проекту та вміти точно виконувати необхідні анотації. Робота з анотуванням зображень зазвичай включає такі завдання:

- Підготовка набору даних зображення.
- Визначення класів об'єктів, які анотатори використовуватимуть для позначення зображень.
- Присвоєння міток зображенням.

- Позначення об'єктів на кожному зображенні за допомогою обмежувальних рамок.
- Вибір міток класу об'єктів для кожного боксу.
- Експортування анотацій у форматі, який можна використовувати як навчальний набір даних.
- Постобробка даних для перевірки правильності маркування.
- У разі неузгодженого маркування система має увімкнути другий або третій раунд маркування з голосуванням між анотаторами.

Існує кілька відкритих і безкоштовних інструментів для анотування зображень. Поширеним інструментом з відкритим кодом, який використовується в багатьох масштабних проектах, є Computer Vision Annotation Tool (CVAT). Інструменти анотації зображень підтримують сам процес анотації (наприклад, вони дозволяють малювати складні форми на зображенні) і забезпечують структуровану систему позначок, щоб анотатори могли застосовувати правильні позначки до артефактів зображення.

Основні функції та можливості платформ анотацій зображень включають:

- Ефективний користувальницький інтерфейс повинен підтримувати швидке маркування, зменшувати кількість людських помилок і бути інтуїтивно зрозумілим для робочої сили без тривалого навчання.
- Підтримка онтології та таксономії має бути налаштована відповідно до структури міток, необхідної для моделі машинного навчання, включаючи класифікації, ієрархічні зв'язки та спеціальні змінні.
- Якість анотації це точність анотації зображень має першорядне значення. Платформа анотацій повинна підтримувати кілька методів вимірювання якості, включаючи порівняльний аналіз (порівняння анотаційної роботи із золотим стандартом) і консенсус (порівняння міток двох або більше анотаторів, які працюють над одним завданням).
- Керування користувачами має підтримувати віддалене керування користувачами, можливість визначати керівників, які можуть переглядати

роботу анотаторів, і функції співпраці, щоб дозволити керівникам надавати відгуки про роботу.

- Автоматизація передові платформи анотацій можуть зменшити людські помилки та зробити роботу анотацій ефективнішою за рахунок автоматизації складних завдань анотації. Автоматизовані піксельні карти та пропозиції міток можуть стати відправною точкою для анотаторів.

- Підтримує загальні формати інструмент має експортувати дані анотацій у простому форматі, який користувачі можуть легко зрозуміти та використовувати в моделях машинного навчання.

Анотація зображення передбачає призначення міток, щоб допомогти моделям штучного інтелекту (ШІ) виявити певні аспекти у візуальному представленні. Різні типи анотацій зображень допомагають представити різні аспекти зображення. Ці поняття використовуються в багатьох типах анотацій зображень:

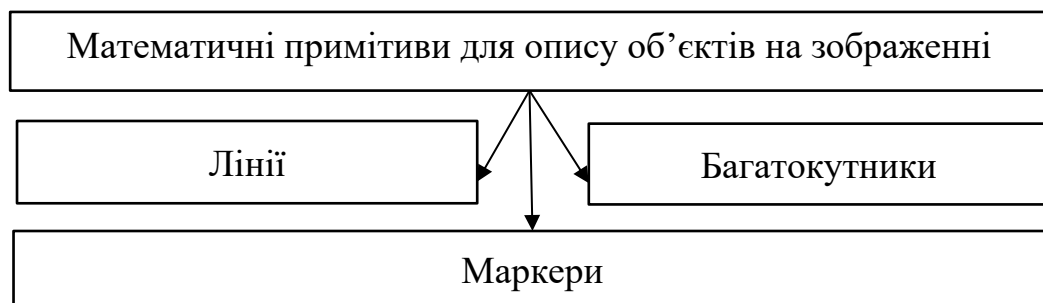


Рисунок 2.1 – Основні елементи для опису об'єктів на зображенні

Лінії – лінії можуть допомогти анотувати об'єкти на зображенні, щоб дозволити машинам визначити межі.

Багатокутники – багатокутники допомагають анотувати об'єкти, які не є ні симетричними, ні правильними. Це включає в себе розміщення крапок поперек розмірів об'єкта та вручну малювання ліній уздовж периметра чи кола об'єкта.

Маркери – деякі анотації включають розміщення маркерів на координатах зображення, які мають особливе значення.

Класифікація зображень дозволяє машинам виявляти об'єкти на зображеннях і в усьому наборі даних. Класифікатор навчається на наборі даних із мітками, навчаючись класифікувати нові невидимі зображення за тим самим набором міток.

Процес підготовки зображень для класифікації зображень широко відомий як анотація або тегування. Це передбачає додавання тегів, які описують об'єкти або сцени на зображенні. Наприклад, ви можете позначити зображення зовнішнього вигляду будівлі мітками на зразок «паркан» або «сад», а внутрішні зображення будівлі – «ліфт» або «сходи».

Розпізнавання об'єктів, або виявлення об'єктів, дозволяє машинам: Визначте певний об'єкт на зображенні та нанесіть точну позначку та визначте наявність кількох об'єктів, включаючи кількість екземплярів і розташування, і застосуйте точні мітки.

Можливо повторити цей процес, використовуючи різні набори зображень, щоб навчити модель машинного навчання автономно ідентифікувати та позначати ці об'єкти на нових зображеннях. Методи, сумісні з розпізнаванням об'єктів, наприклад багатокутники або обмежувальні рамки, можуть допомогти вам позначити різні об'єкти на одному зображенні. Наприклад, можна додавати анотації до автомобілів, велосипедів і пішоходів окремо на одному зображенні.

Орієнтир дозволяє моделям машинного навчання ідентифікувати риси обличчя, вирази, емоції та жести. Ця техніка також може служити для позначення положення та орієнтації тіла людини.

Наприклад, можна використовувати мітки даних, щоб позначити певними цифрами певні місця на обличчі, наприклад губи, брови, очі та лоб. Ваша модель машинного навчання використовує ці позначки, щоб вивчати різні частини людського обличчя.

Розпізнавання меж дає змогу машинам ідентифікувати межі або лінії об'єктів на зображенні. Ці межі можуть включати області топографії, присутні на зображенні та краї певного предмета.

Анотоване зображення може допомогти навчити ваші моделі ідентифікувати схожі шаблони на немаркованих зображеннях. Розпізнавання меж особливо корисно для забезпечення безпечної роботи безпілотних автомобілів.

Існує два основні методи анотації даних – анотація, зроблена людиною, і автоматична анотація. Анотування людиною зазвичай займає більше часу та коштує дорожче, ніж автоматичне, а також потребує навчання для анотаторів, але забезпечує більш точні результати. Для порівняння, автоматизована анотація є більш економічно ефективною, але може бути важко визначити рівень точності результатів.

Моделі машинного навчання потребують узгоджених даних високої якості, щоб робити точні прогнози. Однак особи, які розміщують дані, можуть інтерпретувати суб'єктивні дані по-різному через свої переконання, культуру та особисті упередження. Якщо дані позначені неузгоджено, результати моделі машинного навчання також будуть спотвореними.

Існує багато платформ і інструментів для анотацій зображень, кожна з яких надає різні можливості для різних типів анотацій. Різноманітність пропозицій може ускладнити вибір найбільш підходящих інструментів для кожного проекту. Також може бути складно вибрати правильний інструмент, який відповідатиме навичкам вашої робочої сили.

Позначення зображення – це тип позначення даних, який зосереджується на ідентифікації та позначенні тегами певних деталей зображення.

У комп'ютерному зорі маркування даних передбачає додавання тегів до необроблених даних, таких як зображення та відео. Кожен тег представляє клас об'єкта, пов'язаний з даними. Моделі керованого машинного навчання використовують мітки, коли навчаються ідентифікувати конкретний клас об'єктів у некласифікованих даних. Це допомагає цим моделям пов'язувати значення з даними, що може допомогти в навчанні моделі.

Анотація зображень використовується для створення наборів даних для моделей комп'ютерного зору, які поділяються на набори для навчання, які

використовуються для початкового навчання моделі, і набори для тестування/перевірки, які використовуються для оцінки продуктивності моделі. Фахівці з обробки даних використовують набір даних для навчання та оцінки своєї моделі, а потім модель може автоматично призначати мітки невидимим, непозначеним даним.

Маркування зображень є ключовим компонентом розробки контрольованих моделей із можливостями комп'ютерного зору. Це допомагає навчити моделі машинного навчання маркувати цілі зображення або ідентифікувати класи об'єктів на зображенні. Нижче наведено кілька способів позначення зображень.

Розробка функціональних моделей штучного інтелекту (AI) – інструменти та методи маркування зображень допомагають виділити або зафіксувати певні об'єкти на зображенні. Ці мітки роблять зображення читабельними машинами, а виділені зображення часто служать навчальними наборами даних для моделей ШІ та машинного навчання.

Покращення комп'ютерного зору – маркування та анотації зображень допомагають підвищити точність комп'ютерного зору, дозволяючи розпізнавати об'єкти. Навчання штучного інтелекту та машинного навчання за допомогою міток допомагає цим моделям ідентифікувати шаблони, поки вони не зможуть розпізнавати об'єкти самостійно.

Позначення зображень є основною функцією в алгоритмах комп'ютерного зору. Ось кілька способів позначення зображень системами комп'ютерного зору. Кінцевою метою алгоритмів машинного навчання є автоматичне маркування, але для того, щоб навчити модель, їй знадобиться великий набір даних попередньо позначених зображень.

Алгоритми класифікації зображень отримують зображення як вхідні дані та можуть автоматично класифікувати їх за однією з кількох міток (також відомих як класи). Наприклад, алгоритм може класифікувати зображення транспортних засобів за такими мітками, як «автомобіль», «поїзд» або «корабель». У деяких випадках одне зображення може мати кілька міток. У

наведеному вище прикладі це може статися, якщо одне зображення містить кілька типів транспортних засобів. Щоб створити навчальний набір даних для класифікації зображень, необхідно вручну переглянути зображення та додати до них мітки, які використовує алгоритм. Наприклад, навчальний набір даних для транспортних зображень включатиме велику кількість зображень, що містять транспортні засоби, і людина відповідатиме за перегляд кожного зображення та застосування відповідної мітки – «автомобіль», «потяг», «корабель» тощо. Крім того, можна створити такий набір даних за допомогою методів синтетичних даних.

У семантичній сегментації зображення алгоритм комп'ютерного зору має завдання відокремити об'єкти зображення від фону чи інших об'єктів. Зазвичай це передбачає створення піксельної карти зображення, де кожен піксель містить значення 1, якщо він належить відповідному об'єкту, або 0, якщо не належить. Якщо на одному зображенні є кілька об'єктів, зазвичай підхід полягає у створенні кількох піксельних об'єктів, по одному для кожного об'єкта, і їх об'єднання по каналах. Щоб створити навчальний набір даних для набору даних семантичної сегментації, необхідно вручну переглянути зображення та накреслити межі відповідних об'єктів. Це створює піксельну карту, перевірену людиною, яку можна використовувати для навчання моделі. Крім того, можна генерувати піксельні карти шляхом створення синтетичних зображень, у яких уже відомі межі об'єктів.

Завданням алгоритму виявлення об'єктів є виявлення об'єкта на зображенні та його розташування в кадрі зображення. Розташування об'єктів зазвичай визначається за допомогою обмежувальних рамок. Обмежувальна рамка – це найменший прямокутник, який містить увесь об'єкт зображення.

Технічно обмежувальна рамка – це набір із чотирьох координат, присвоєних мітці, що визначає клас об'єкта. Координати обмежувальних рамок і їхні мітки зазвичай зберігаються у файлі JSON у форматі словника. Номер або ідентифікатор зображення є ключем у файлі словника.

Позначення зображення – це тип позначення даних, який зосереджується на ідентифікації та позначенні тегами певних деталей зображення. Поширеним способом позначення зображень є анотація вручну. Це процес ручного визначення міток для всього зображення або малювання областей на зображенні та додавання текстових описів кожної області. Анотація зображення встановлює стандарт, на якому намагається навчитися алгоритм комп'ютерного зору. Це означає, що будь-які помилки в мітках будуть прийняті алгоритмом, знижуючи його точність. Це означає, що точне маркування зображень є критичним завданням у навчанні нейронних мереж.

Анотації вручну зазвичай допомагають інструменти, які дозволяють операторам обертати велику кількість зображень, малювати області на зображенні та призначати мітки, а також зберігати ці дані в стандартизованому форматі, який можна використовувати для навчання даних.

Анотація зображення вручну викликає кілька проблем:

- Мітки можуть бути непослідовними, якщо є кілька анотаторів, і щоб вирішити це, зображення потрібно позначати кілька разів більшістю голосів.
- Ручне маркування займає багато часу. Анотатори повинні бути ретельно навчені, а процес потребує багатьох ітерацій. Це може затримати час виходу на ринок проектів комп'ютерного зору.
- Ручне маркування є дорогим і його важко масштабувати для отримання великих наборів даних.

Анотація зображення вручну є трудомістким завданням, а для деяких алгоритмів комп'ютерного зору це може бути важко виконати людям. Наприклад, деякі алгоритми вимагають створення піксельних карт із зазначенням точних меж кількох об'єктів на зображенні.

Інструменти автоматизованого анотування можуть допомогти ручним анотаторам, намагаючись виявити межі об'єктів на зображенні та забезпечуючи початкову точку для анотатора. Алгоритми автоматичного анотування не є абсолютно точними, але вони можуть заощадити час анотаторам, надаючи принаймні часткову карту об'єктів на зображенні.

Синтетичне маркування зображень є точним і економічно ефективним методом, який може замінити ручні анотації. Він передбачає автоматичне генерування зображень, схожих на реальні дані, відповідно до критеріїв, встановлених оператором. Наприклад, можна створити синтетичну базу даних об'єктів реального життя або людських облич, які схожі, але не ідентичні реальним об'єктам.

Основна перевага синтетичних зображень полягає в тому, що мітки відомі заздалегідь, наприклад, оператор автоматично генерує зображення столів і стільців. У цьому випадку алгоритм, що генерує зображення, може автоматично надати обмежувальні рамки столів і стільців на кожному зображенні.

2.2 Алгоритми сегментації цифрових зображень

Сегментація зображення передбачає поділ зображення на кілька областей або об'єктів для полегшення розпізнавання та аналізу об'єктів. Для ідентифікації меж і розрізнення об'єктів використовуються такі методи, як порогове значення, сегментація на основі регіону та визначення країв.

Сегментація зображення - це метод, що дозволяє розбити цифрове зображення на окремі групи, відомі як сегменти. Це значно полегшує аналіз та обробку зображення, фокусуючись на конкретних об'єктах або областях інтересу. Технічно сегментація включає в себе призначення міток пікселям для ідентифікації об'єктів, що робить її ключовим елементом комп'ютерного зору та обробки зображень.

Алгоритми сегментації дозволяють виявляти об'єкти на зображенні, зменшуючи складність аналізу. Замість обробки всього зображення вони визначають окремі області, які потрібно подальше досліджувати. Наприклад, алгоритм сегментації може виділити об'єкти, і детектор об'єктів вже буде

працювати з цими областями, збільшуючи ефективність та зменшуючи час обробки.

Сегментація зображення застосовується в різних галузях, включаючи медичну діагностику, автономні транспортні засоби, розпізнавання облич, відеоспостереження та обробку супутникових зображень. Вона є важливим етапом в алгоритмах комп'ютерного зору, дозволяючи аналізувати та розпізнавати об'єкти на зображеннях. Сегментація зображення функціонує, отримуючи вхідні дані та генеруючи вихідні дані у вигляді маски, де кожен піксель отримує мітку, що визначає його клас чи об'єкт. Такий результат полегшує подальший аналіз та розпізнавання об'єктів на зображенні. Традиційні методи сегментації використовують евристичні або високорівневі функції, такі як кластеризація, обробка країв та гістограма. Однак вони можуть вимагати налаштувань та не завжди точні для складних зображень. Сучасні методи використовують машинне та глибоке навчання для покращення точності та гнучкості сегментації. Алгоритми сегментації дозволяють виявляти об'єкти на зображенні, зменшуючи складність аналізу. Замість опрацювання всього зображення, вони визначають окремі області, які потрібно докладніше досліджувати. Наприклад, алгоритм сегментації може виділити об'єкти, а детектор об'єктів вже працюватиме з цими областями, збільшуючи ефективність та зменшуючи час обробки.

У сучасних підходах до сегментації, використовуючи машинне навчання, моделі навчаються визначати важливі функції для поліпшення здатності програми. Глибокі нейронні мережі особливо ефективні для завдань сегментації.

Існує різноманітні конструкції та варіації нейронних мереж, спроектованих для завдань сегментації зображень. Безперечно, їхня спільна основа включає такі ключові компоненти:

Кодер є першим етапом у нейронній мережі для сегментації зображень. Цей шар або серія шарів використовують фільтри для виділення важливих особливостей та деталей зображення. Часто використовуються згорткові та пулінгові шари для поетапного зменшення розмірності та збереження ключових

особливостей. Кодер може бути попередньо навчений на інших завданнях, що сприяє ефективній екстракції функцій.

Декодер відповідає за перетворення вихідних даних кодера в сегментаційну маску. Це зазвичай включає транспоновані згорткові та підвищення розмірності шари для відновлення деталей та структури. Декодер дозволяє моделі відтворити зображення у високій роздільності, щоб отримати точний результат сегментації.

Ці компоненти узгоджено працюють для створення нейронних мереж, здатних ефективно та точно сегментувати різноманітні зображення. Використовуючи ці елементи, модерні архітектури, такі як U-Net, SegNet, та DeepLab, здатні досягти високої ефективності в сегментації об'єктів на зображеннях.

Для проведення класифікації методів сегментації було обрано декілька критеріїв, зокрема алгоритми сегментації можна розділити на декілька груп в залежності від отриманих результатів (рисунок 2.2)



Рисунок 2.2 – Класифікація алгоритмів сегментації в залежності від кінцевого результату обробки

Сегментація зображення дозволяє машинам визначати межі та об'єкти на зображенні. Ця методика дозволяє досягти вищої точності для завдань класифікації. Одним із підтипів цього методу є семантична сегментація, яка включає в себе класифікацію пікселів на основі семантичних класів, тобто присвоєння кожному пікселю конкретного семантичного значення.

У семантичній сегментації кожен піксель призначається певному класу, і модель сегментації спрощено розглядає зображення як набір ізольованих класів без врахування контексту чи взаємодії між ними. Наприклад, якщо використовувати семантичну сегментацію для зображення, на якому є дерева та транспортні засоби, кожен піксель буде призначений конкретному класу залежно від його семантичного призначення.

Такий підхід призводить до того, що постановка проблеми часто погано визначена, особливо якщо в одному класі згруповано кілька екземплярів. Наприклад, зображення багатолюдної вулиці може сегментувати всю область натовпу як клас «люди». Семантична сегментація не забезпечує глибокої деталізації таких складних зображень. Приклад роботи алгоритмів семантичної сегментації наведено на рисунку 2.3.



Рисунок 2.3 – Приклад семантичної сегментації цифрового зображення

Сегментація екземпляра є методом, який передбачає класифікацію пікселів на основі конкретних екземплярів об'єктів, а не їхніх класів. Основна відмінність від семантичної сегментації полягає в тому, що алгоритми сегментації екземплярів не класифікують області на підставі їхнього семантичного значення,

а намагаються виділити окремі екземпляри об'єктів, навіть якщо вони частково накладаються один на одного.

Однак на відміну від семантичної сегментації, яка може класифікувати всі області як "людина", сегментація екземплярів може виділити кожну особу окремо, не передбачаючи конкретного класу ("людина") для кожного екземпляра. Приклад роботи алгоритмів сегментації екземплярів наведено на рисунку 2.4.

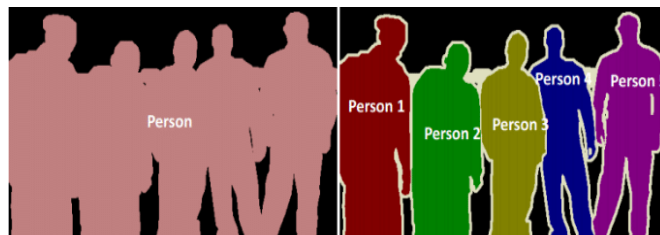


Рисунок 2.4 – Приклад сегментації екземплярів цифрового зображення

Паноптична сегментація – це новий тип сегментації, який часто виражається як поєднання семантичної та екземплярної сегментації. Вона передбачає ідентичність кожного об'єкта, відокремлюючи кожен екземпляр кожного об'єкта на зображенні. Приклад роботи алгоритмів паноптичної сегментації наведено на рисунку 2.5.



Рисунок 2.5 – Приклад паноптичної сегментації цифрового зображення

Паноптична сегментація корисна для багатьох продуктів, які потребують величезних обсягів інформації для виконання своїх завдань. Наприклад,

безпілотні автомобілі повинні швидко й точно знімати та розуміти навколишнє середовище. Вони можуть досягти цього, подаючи живий потік зображень на алгоритм панорамної сегментації.

Іншим критерієм для проведення сегментації є критерій характеристик на основі яких проводиться процес розділення вхідного зображення на однорідні області. За даним критерієм алгоритми сегментації можна розділити на такі окремі класи (рисунок 2.6)

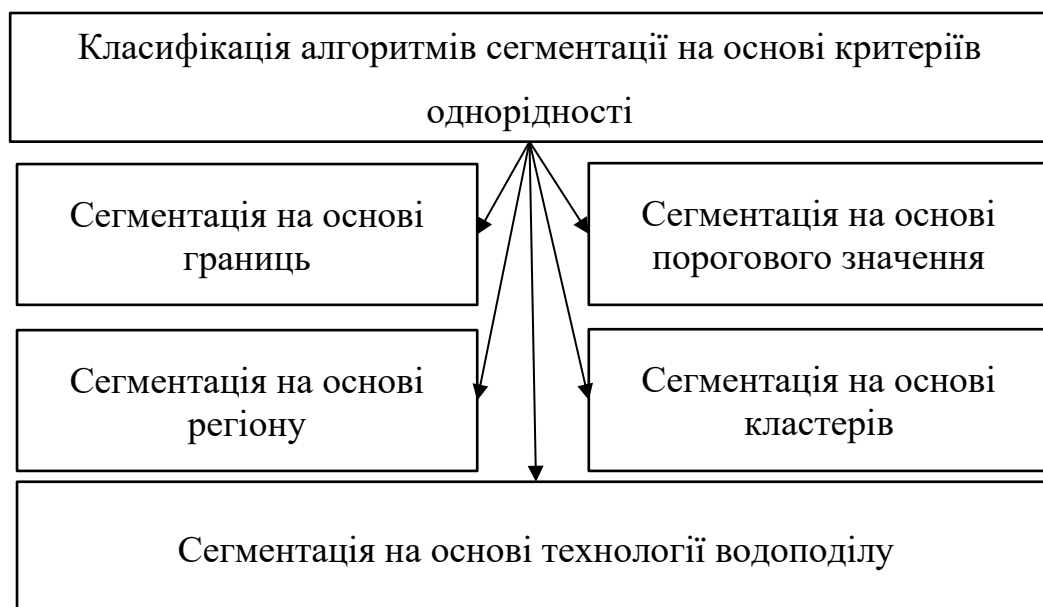


Рисунок 2.6 – Класифікація алгоритмів сегментації

Техніка сегментації на основі країв є ефективним методом обробки зображень, який фокусується на ідентифікації та виділенні країв різних об'єктів на зображенні. Цей підхід дозволяє точно визначати межі об'єктів, використовуючи інформацію про їх краї. Виявлення країв має кілька переваг. По-перше, воно спрощує зображення, видаляючи зайву інформацію та залишаючи лише ключові аспекти, такі як контур об'єктів. Це полегшує аналіз та обробку зображення. По-друге, використання різних параметрів, таких як контраст, текстура, кольорові та насиченість, дозволяє більш гнучко адаптувати алгоритми до різних типів зображень і сцен.

Приклад роботи алгоритмів на основі виділення границь наведено на рисунку 2.7.



Рисунок 2.7 – Приклад роботи алгоритмів на основі виділення границь

Сегментація на основі порогових значень. Метод порогового значення є найпростішим способом сегментації зображення, в основі якого лежить розподіл пікселів в залежності від їхньої інтенсивності в порівнянні з певним значенням або порогом. Цей метод особливо ефективний для виділення об'єктів з вищою інтенсивністю в порівнянні з іншими об'єктами або фонами на зображенні.

Порогове значення, позначене як T , може використовуватися як константа в зображеннях з низьким рівнем шуму. Однак в деяких ситуаціях застосування динамічних порогів може бути ефективнішим. Динамічні пороги адаптуються до локальних змін в інтенсивності, що дозволяє покращити результати сегментації на зображеннях з високим рівнем змішаного шуму або нерівномірною інтенсивністю. Приклад сегментації на основі порогових значень наведено на рисунку 2.8.

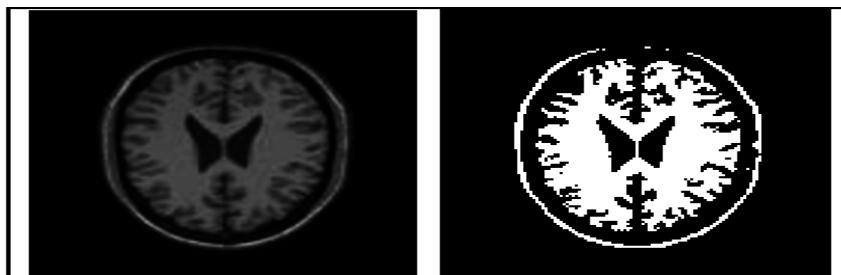


Рисунок 2.8 – Приклад роботи алгоритмів на основі порогових значень

Сегментація на основі регіонів передбачає поділ зображення на регіони зі схожими характеристиками. Кожна область — це група пікселів, які алгоритм визначає через вихідну точку. Коли алгоритм знаходить вихідні точки, він може розширювати регіони, додаючи більше пікселів або звужуючи та об'єднуючи їх з іншими точками. Приклад роботи алгоритмів на основі виділення регіонів наведено на рисунку 2.9.

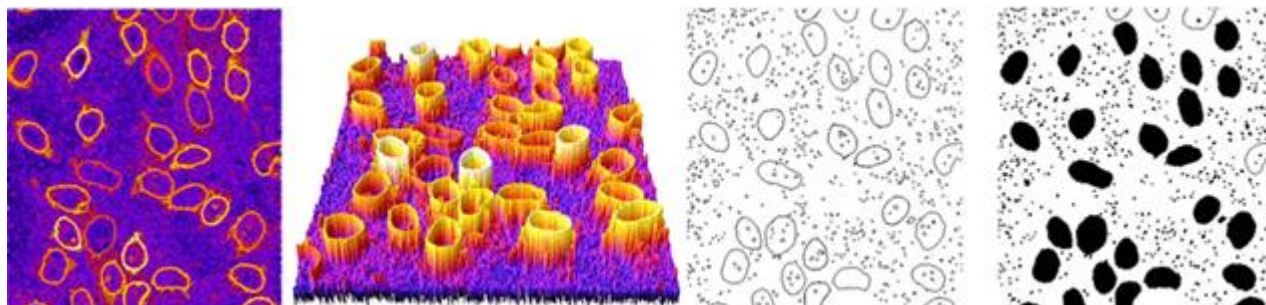


Рисунок 2.9 – Приклад роботи алгоритмів на основі виділення регіонів

Сегментація на основі кластерів використовує алгоритми кластеризації, які є неконтрольованими методами класифікації, спрямованими на ідентифікацію прихованої інформації на зображеннях. Ці алгоритми покращують сприйняття людиною, виділяючи кластери, тіні та структури на зображенні. Принцип дії полягає в тому, що алгоритм розбиває зображення на кластери пікселів зі схожими характеристиками, відокремлюючи елементи даних і групуючи подібні елементи в кластери.

Щодо сегментації водоподілу, вона використовує техніку перетворення сірого зображення, розглядаючи зображення як топографічну карту, де яскравість пікселів визначає їхню висоту над рівнем моря. Алгоритми сегментації водоподілу виявляють лінії, представляючи хребти та западини, і виділяють ділянки між лініями водоподілу. Ця техніка дозволяє розділити зображення на області на основі висоти пікселів, групуючи пікселі з однаковим значенням сірого. Використання техніки водоподілу має важливі застосування, зокрема у медичній сфері. Наприклад, вона може бути корисною для визначення відмінностей між світлішими та темнішими ділянками на зображеннях.

2.3 Алгоритм сегментації на основі аналізу однорідних областей

Алгоритми сегментації розділяють вхідне зображення на деяку кількість однорідних областей, які в подальшому можуть об'єднуватись в більші групи. Об'єднання може відбуватись на основі декількох критеріїв: зменшення порогу відмінності або додавання моделей опису об'єктів. Розроблений алгоритм сегментації базується на твердженні, що окремі точки зображення будуть мати подібні характеристики при використанні різних алгоритмів сегментації. Тоді поєднання результатів сегментації різними підходами дозволить провести процес об'єднання різних областей в більші. На основі проведеного аналізу алгоритмів сегментації для роботи запропонованого алгоритму було обрано три основних. Їхній перелік, переваги та недоліки наведено в таблиці 2.1.

Таблиця 2.1 – Алгоритми попередньої сегментації

Назва алгоритму	Переваги алгоритму	Недоліки
Алгоритм порогової сегментації з адаптивним порогом	• Простота реалізації; • Висока швидкість роботи; • Адаптація порогових значень.	• Проблема вибору оптимальних порогових значень; • Вплив шумів на результати роботи.
Алгоритм водоподілу	• Незалежність від складності зображення; • Висока точність роботи.	• Складність реалізації; • Висока трудоемність.
Алгоритм К-середніх	• Обробка складних зображень; • Можливість адаптації до особливостей вхідного зображення.	• Залежність від початкових даних; • Складність обробки зображень з великою кількістю незначних перепадів яскравостей.

Обраний набір алгоритмів дозволяє отримати на етапі початкової обробки зображень масив однорідних областей, що базуються на різних критеріях однорідності. Даний набір масивів необхідно об'єднати на основі значень координат розташування кожної з областей. Області точки які залишаються в одній області при різних однорідних областях повинні бути згруповані в кінцеві однорідні області. Для аналізу запропонованого алгоритму його основні етапи відображені за допомогою блок-схеми алгоритму сегментації цифрових зображень на основі аналізу однорідностей області (рисунок 2.10).

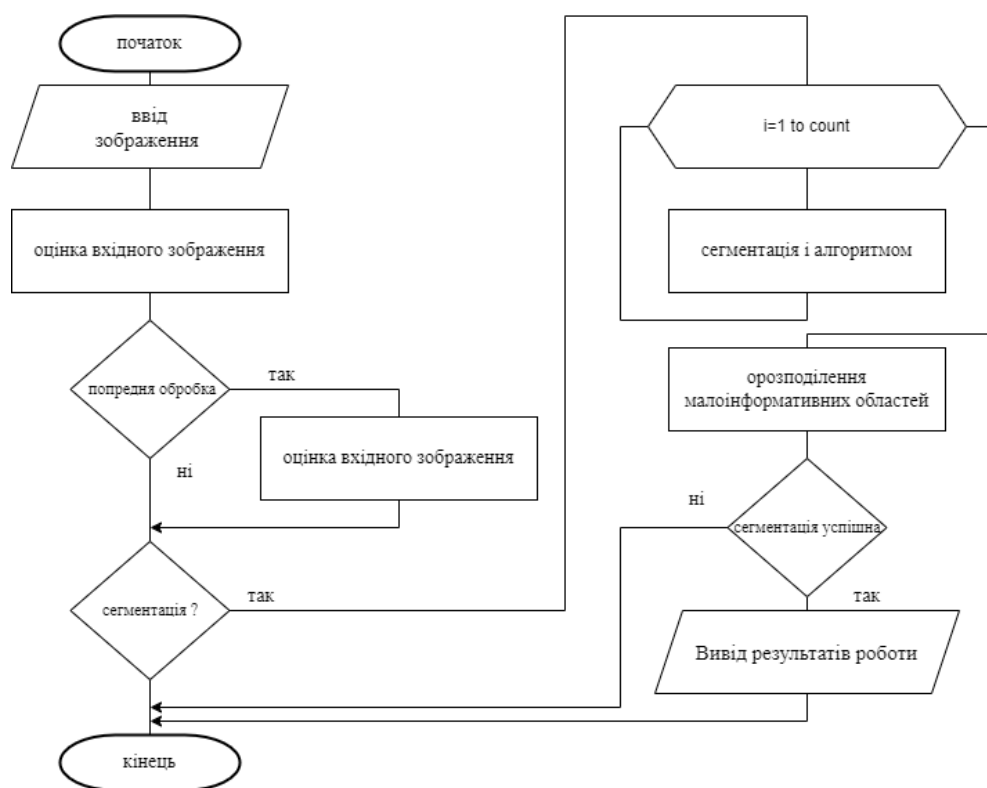


Рисунок 2.10 – Блок-схема алгоритму алгоритму сегментації цифрових зображень на основі аналізу однорідностей області

Проведення сегментації з використання додаткових алгоритмів попередньої обробки дозволяє провести більш точну оцінку точок вхідного зображення, що в кінцевому результаті підвищує точність сегментації. Запропонований алгоритм розбитий на декілька кроків:

- 1) Завантаження цифрового зображення.

- 2) Проведення первинної оцінки вхідного зображення.
- 3) Якщо вхідне зображення містить невелику кількість шумів або взагалі шуми на зображенні відсутні, то переходимо на крок 5.
- 4) Попередня обробка вхідного зображення (переконвертація в градації сірого, видалення шумів, вирівнювання гістограми яскравості).
- 5) Визначення параметрів порогових значень для алгоритму порогової сегментації.
- 6) Оримання масиву однорідних областей на основі порогової сегментації.
- 7) Оримання масиву однорідних областей на основі алгоритму сегментації водоподілом.
- 8) Оримання масиву однорідних областей на основі алгоритму сегментації K-середніх.
- 9) Аналіз масивів однорідних областей та формування підсумкового набору однорідних областей.
- 10) Оцінка якості результатів сегментації на основі одноріжних областей. Якщо оцінка негативна повернення на крок 5.
- 11) Вивід результатів робота алгоритму сегментації.

Запропонований в роботі алгоритм дозволяє здійснити розділення вхідного цифрового зображення на одноріжні області на основі групи різних критеріїв, що дозволить значно підвищити точність кінцевого результату. Перевагами запропонованого алгоритму можна вважати наступні:

- Запропонований алгоритм може опрацьовувати зображення різної складності, форматів та типів;
- Для роботи необхідний рівень втручання оператора, що робить його перспективним під час проектування програмних автоматизованих систем та систем комп'ютерного зору;
- Алгоритм може просто адаптуватись під різні вхідні цифрові зображення шляхом налаштування параметрів своєї роботи або зміни/додавання алгоритмів первинної обробки.

Недоліки запропонованого алгоритму:

- складність реалізації та висока трудомність роботи у зв'язку з використання великої кількості математичних обчислень та операцій порівняння;
- необхідність встановлення коректних параметрів роботи алгоритмів початкового аналізу цифрового зображення.

В загальному, проаналізувавши усі переваги та недоліки можна зробити висновок, що запропонований алгоритм буде працювати довший час проте результати роботи будуть більш точніше відповідати експертним оцінкам. Даний алгоритм можна використовувати в програмних системах у яких важливіша точність результатів, а не сам час роботи.

2.4 Висновки до розділу

Проведено аналітичне дослідження алгоритмів сегментації цифрових зображень на основі критеріїв поділу на однорідні області, що дозволило виділити групу алгоритмів результати яких максимально подібні до результатів отриманих за допомогою експертів.

Розроблено алгоритм сегментації цифрових зображень на основі аналізу однорідних областей, що дозволило спроектувати структуру програмного додатку обробки та аналізу цифрових зображень.

3 ПРОГРАМНИЙ ДОДАТОК АНАЛІЗУ ТА ОПИСУ ЦИФРОВИХ ЗОБРАЖЕНЬ

3.1 Структура програмного додатку опрацювання цифрових зображень

Розробка програмних систем обробки зображень є комплексним і багатогранним завданням, що вимагає врахування різноманітних аспектів. Існує кілька ключових особливостей, які визначають успішність та ефективність таких систем. По-перше, вибір та оптимізація алгоритмів обробки зображень відіграють вирішальну роль. Розробники стикаються з викликом забезпечити ефективність та точність алгоритмів в умовах різноманітних типів зображень та їх дозволів. Це включає роботу з колірними просторами, фільтрацію, сегментацію і виділення об'єктів. Інший важливий фактор - це продуктивність. Продуктивність має критичне значення для систем обробки зображень, особливо під час роботи у реальному часі. Це вимагає оптимізації коду, використання паралельних обчислень та ефективного керування пам'яттю. Програми, призначені для обробки відео або великих обсягів зображень, повинні бути оптимізовані для мінімізації затримок. Також необхідно враховувати варіативності вхідних даних. Розробники повинні передбачити підтримку різних форматів зображень, дозволів та умов освітлення. Адаптивність алгоритмів до різноманітних сценаріїв використання, таких як медична діагностика, аналіз відеоспостереження або обробка фотографій є невід'ємною частиною розробки. Ще однією ключовою особливістю є взаємодія з користувачем. Інтерфейс обробки зображень має бути інтуїтивно зрозумілим, гнучким та функціональним. Це включає можливість візуалізації результатів обробки, регулювання параметрів алгоритмів та інтеграції з іншими програмами. Важливим аспектом у сучасних системах обробки зображень є забезпечення безпеки та конфіденційності даних. Програмні продукти, які працюють із медичними зображеннями або чутливою інформацією, повинні відповідати високим стандартам безпеки та шифрування.

Слід також відзначити вплив технологічних трендів, таких як машинне навчання та штучний інтелект, на розробку систем обробки зображень. Використання нейронних мереж для розпізнавання образів, класифікації об'єктів та семантичної сегментації відкриває нові горизонти у сфері аналізу зображень.

Розробка програмних систем обробки зображень вимагає глибокого технічного розуміння, творчого підходу та обліку безлічі факторів. Сучасні тенденції в галузі комп'ютерного зору та технології можуть значно збагатити функціональність таких систем, роблячи їх більш точними, адаптивними та застосовними у різних сценаріях використання.

Для створення програмного додатку обробки та аналізу цифрових зображень на основі запропонованого алгоритму необхідно спроектувати внутрішню архітектуру. При розробці архітектури було враховано результати аналізу програмних засобів роботи з цифровими зображеннями які присутні на сучасному ринку. Запропонована структура наведена на рисунку 3.1.

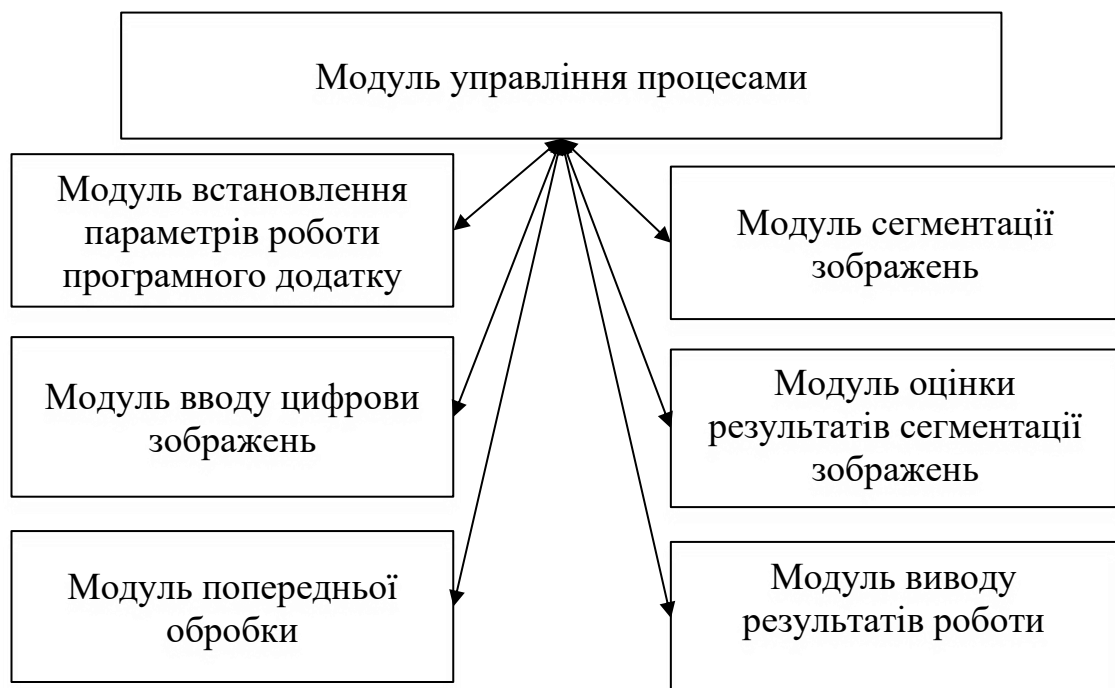


Рисунок 3.1 – Узагальнена структура додатку обробки та аналізу цифрових кольорових зображень

Запропонований підхід дозволив розробити децентралізовану архітектуру програмного додатку, що дозволить в процесі реалізації приділити більшу увагу окремим елементам, а не цілій системі. При такому підході в систему додається модуль управління процесами який дозволить автоматизувати процес обміну інформацією між різними програмними модулями, а самі програмні модулі будуть більш незалежними та менше піддаватимуться впливу помилкових ситуацій на інших етапах обробки зображення.

Модуль управління процесами призначений для організації основних технічних функцій в середині програмного додатку та моніторингу проходження окремих процесів. Серед функцій реалізованих в даному модулі присутні функції перевірки цілісності програмного додатку, наявності підключення апаратних засобів вводу/виводу цифрових зображень, переконвертація зображень та перевірка коректності отриманих результатів роботи на кожному з етапів. Додатково в даному модулі організована система повідомлень про виникнення критичних помилок та логування їх у відповідних файлі.

Модуль вводу цифрового зображення. Даний модуль був виділений для групування функцій для завантаження цифрових зображень в програмний додаток. Серед можливих шляхів отримання цифрового зображення реалізовані наступні можливості:

- отримання файлу з фотокамери;
- отримання одного файлу з носія даних;
- отримання групи файлів з сховища даних.

Даний набір механізмів дозволяє користувачеві отримувати інформацію у зручний для нього спосіб. При отриманні зображення з фотокамери користувач має можливість самостійно обрати розмір області інтересу, формат та додаткові технічні характеристики цифрового зображення. При завантаженні групи зображень в програмі формується внутрішня черга і файли завантажуються по мірі опрацювання попередніх.

Модуль попередньої обробки. Для автоматизації роботи програмного додатку важливим етапом є попередня оцінка вхідного зображення та при

можливості усунення факторів які будуть негативно впливати на кінцевий результат роботи. Окрім того, для універсальності проходження процесів сегментації вхідні зображення повинні відповідати деяким критеріям, зокрема всі кольорові зображення повинні бути переконвертовані в відтінки сірого. Саме на даному етапі приймається рішення чи вхідне зображення зможе бути опрацьовано за допомогою розроблених алгоритмів сегментації.

Модуль встановлення параметрів роботи програмного додатку. В програмному додатку будуть встановлені стандартні початкові параметри роботи. Проте набір даних параметрів є рекомендаційним і для отримання більш точних результатів роботи необхідно проводити їх корекцію. Параметри роботи програмного додатку розділені на дві основні групи: параметри роботи програми та параметри роботи алгоритмів обробки зображень. В першу групу параметрів були виділені параметри які відповідають за візуальний вигляд програми (розмір, тип, корір шрифтів, колір самого програмного додатку тощо). До другої групи віднесені параметри алгоритмів та функцій обробки зображень (порогові значення, коефіцієнти для перетворення між різними форматами тощо). Параметри першої групи не впливають на результати обробки зображень, вони необхідні для підвищення зручності роботи з програмним додатком. Тому їх корекція дозволена усім користувачам. Неправельне встановлення параметрів другої групи може призвести до виникнення критичних ситуацій та відмов при обробці цифрових зображень, тому корекція цих параметрів потребує певного рівня знань та досвіду обробки зображень за допомогою програмних засобів.

Модуль сегментації. Функції для проведення попередньої обробки цифрових зображень та розбиття їх на одноріжні області були згруповані в даному модулі. Оскільки етап первинної сегментації має основний вплив на отримані результати роботи, тому для його реалізації та тестування їх було виділено, що полегшить роботу над ним в кінцевому результаті. Для попередньої сегментації були обрані алгоритми порогової сегментації з адаптивним порогом, алгоритми на основі водоподілу та k-середніх. Як показали результати

дослідження запропонованого алгоритму саме ці алгоритми дозволяють в повній мірі та за мінімальний час розділити вхідне зображення на одноріжні області.

Модуль оцінки результатів сегментації зображення. В результаті роботи попередні модулів в програмному додатку будуть знаходитись декілька масивів з однорідними областями, але для отримання кінцевого результату їх необхідно об'єднати. Функції даного модуля проводять оцінку попередньо обробленого зображення та проводять процес формування однорідних областей отриманих на основі різних критеріїв. Для процесу зіставлення як додатковий параметр використовується положення точки/області на зображенні, точки які знаходяться в однорідній області та розташовані в одних координатах формують кінцеву одноріжну область. Якщо в результаті роботи функції об'єднання не вдалось провести остаточну сегментацію зображення, то програма повернеться на етап попередньої обробки зображень та змінивши параметри роботи алгоритмів попередньої сегментації спробує коректно завершити свою роботу.

Модуль виводу результатів роботи. На останньому етапі роботи з програмною системою вона має можливість показати результати у двох варіаціях, або вивести результати сегментації на екран, у відповідному вікні, або зберегти їх в файні на носій даних. Спосіб збереження та візуалізації обирає користувач на основі поставленого завдання.

Перевагами використання запропонованої структури є:

- запропонована модульна структура з незалежними програмними модулями має ряд переваг:
- можливість реалізації, тестування та налагодження кожного з модулів незалежно один від одного;
- можливість простої заміни кожного з модулів без втрати цілісності всього проекту;
- можливість додавання нових модулів для розширення функціональних можливостей програмного додатку при цьому основне навантаження на його інтеграцію буде покладатись на модуль організації процесів;

- можливість додавання/видалення окремих алгоритмів в кожний з модулів без необхідності переформатування системи в цілому.

Для виявлення критичних місць у запропонованій структурі програмного додатку та виправлення їх до початку етапу реалізації проведено процес моделювання та аналізу запропонованої структури за допомогою інструментарію універсальної мови моделювання.

Використання Unified Modeling Language (UML) діаграм у процесі розробки програмного забезпечення приносить значні переваги, покращуючи комунікацію та полегшуючи розуміння проекту. Спочатку UML забезпечує стандартизований та візуальний підхід до представлення структури та поведінки системи, зменшуючи ризик непорозуміння між членами команди.

Також важливим аспектом є можливість проведення аналізу та проектування на вищому рівні абстракції. UML дає можливість команді швидко зосереджуватись на ключових аспектах системи, уникаючи зайвої деталізації на початкових етапах. Цей підхід сприяє створенню чистої, ефективної системи, що легко підтримується.

В процесі моделювання розроблена структура буде перевірятись на два основних параметра:

- доступність функціональних можливостей програми для користувачів;
- аналіз проходження внутрішніх процесів програмного додатку з метою виявлення внутрішніх колізій та суперечностей.

Дані параметри є одними з основних та від рівня продуманості та якості їх протікання буде залежити функціональна привабливість програмного додатку. У випадку, якщо користувачі не будуть отримувати доступ до усіх необхідних функцій, вони можуть відмовитись від використання програми або отримані результати будуть низької якості. У випадку несумісності внутрішніх процесів під час експлуатації програмного додатку можуть виникати критичні ситуації, що в цілому негативно вплине на рівень задоволення програмою серед користувачів. Для першого моделювання використаємо діаграму прецедентів. Приклад самої діаграми наведено на рисунку 3.2.

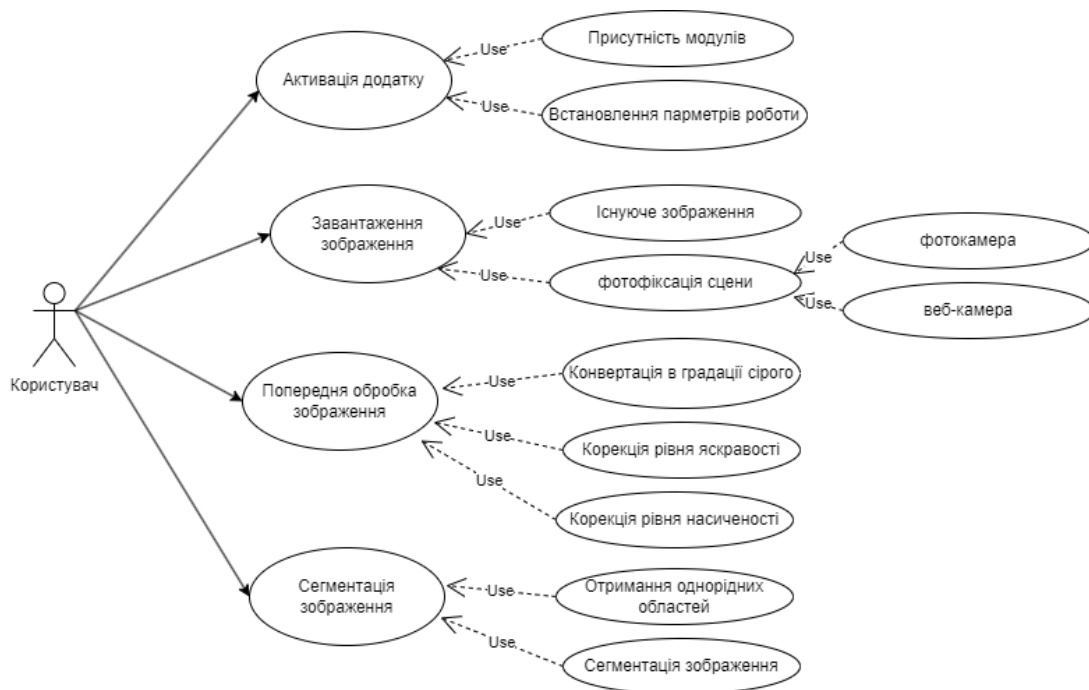


Рисунок 3.2 – Діаграма прецедентів програмної системи

Діаграма прецедентів дозволяє оцінити рівні доступності окремих груп користувачів до функцій програмного додатку. В випадку програмного додатку обробки та аналізу цифрових зображень було виділено одну групу акторів «Користувачі». Програмний додаток призначений для широкого використання та надає однакові функціональні можливості без поділу користувачів на групи. Результати моделювання показали, що обрана група акторів має повний доступ до усіх функцій програмної системи та може використовувати усі запрограмовані можливості по обробці цифрових зображень. Окремо було проаналізована ситуація з наданням прав доступу до корекції параметрів процесів сегментації, оскільки внесення неправильних даних може призвести до негативних результатів, проте під час програмної реалізації функцій корекції параметрів був реалізований механізм при якому користувачі мають можливість встановлювати нові параметри тільки в наперед визначених межах. Це дозволило одразу надати повний доступ до усіх функцій програми та прискорити час навчання роботи з програмою для користувачів з невеликим досвідом роботи.

Для проведення другої оцінки запропонованої структури було використано діаграму послідовностей, яка дозволяє проаналізувати послідовності активації

окремих процесів в середині програмного додатку та виявити місця можливих колізій або розривів процесу обробки цифрового зображення. Результат моделювання наведено на рисунку 3.3.

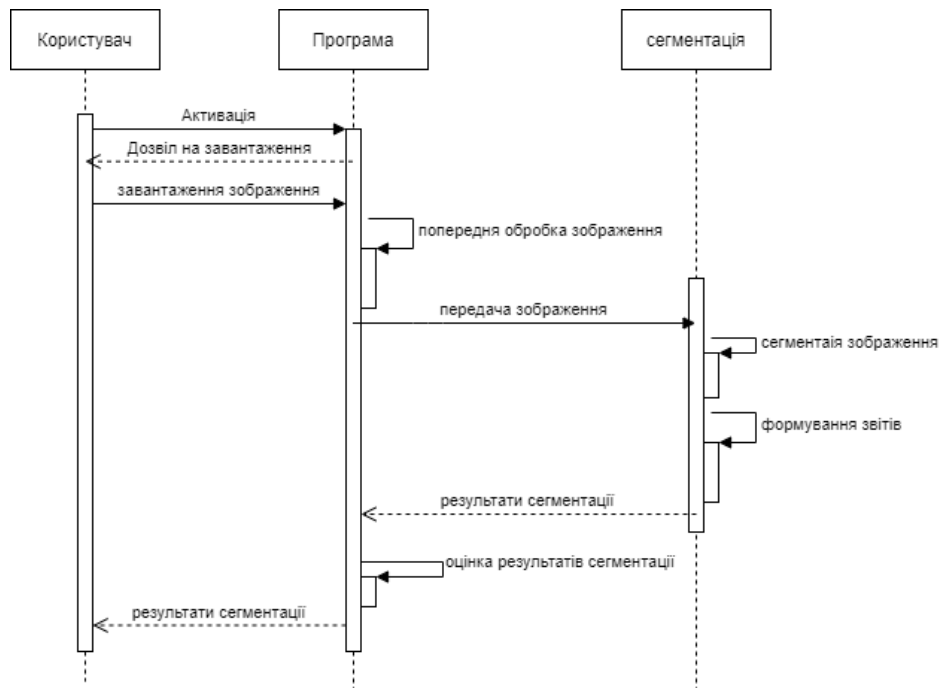


Рисунок 3.3 – Діаграма послідовності програмної системи

Отримані в результаті моделювання дані показують, що під час проектування структури модулі будуть активуватись в наперед передбаченій послідовності, дані між різними модулями будуть передаватись послідовно, що не призведе до виникнення критичних чи помилкових ситуацій. Окрім того додавання модуля управління внутрішніми процесами надало можливість реалізувати систему внутрішнього моніторингу та корекції помилок.

Результати моделювання продемонстрували, що запропонована структура є повною, дозволяє вирішувати усі поставлені перед програмою завдання та на її основі можна реалізовувати програмний додаток обробки та аналізу цифрових кольорових зображень.

Останній етап проектування програмного додатку обробки та аналізу цифрових зображень є розробка дизайну графічного інтерфейсу користувача. Під час розробки зовнішнього вигляду програмного додатку були використані

елменти які є звичними для програмних додатків під управлінням операційної системи Windows. Даний хід був зобумовлений тим, що такі візуальні компоненти є звичними для більшості користувачів, а це в свою чергу призведе до зменшення часу адаптації та наведення роботи розробки. Приклад запропонованого дизайну наведений на рисунку 3.4.

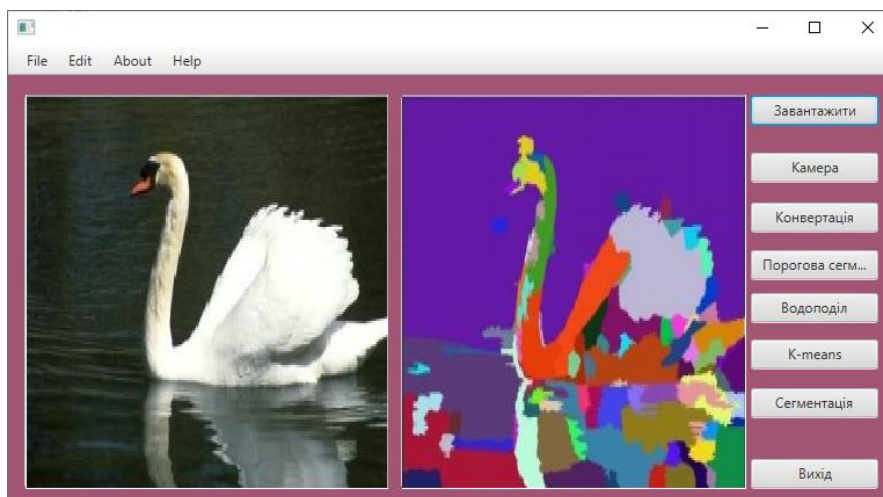


Рисунок 3.4 – Графічний інтерфейс програмного додатку

Запропонований інтерфейс складається з трьох візуально розділених елементів дизайну. Більшу частину робочої області програмного додатку займають віна для візуального відображення цифрових зображень, при цьому користувач може одночасно бачити зміни які були здійснені в процесі роботи програми. Розміри даний вікон можуть корегуватись та масштабуватись для більшої зручності відображення результатів. В верхній частині вікна було виділено область для системного меню програмного додатку. Дане меню є подібним до меню більшості офісних програм, та дублює стандартні операції при роботі з програмними додатками. В даному меню реалізовані команди завантаження/збереження цифрового зображення, встановлення параметрів роботи програми, система підказок для полегшення процесу опанування програмним додатком. З правої сторони було виділено людасть гарячих кнопок. Дані кнопки дозволяють активувувати найбільш частовикористовувані функції напряму, без використання елементів головного меню.

3.2 Модулі програмного додатку аналізу та оцінки програмного коду

Програмна реалізація напряду залежить від обраної мови програмування та сторонніх цифрових бібліотек для прискорення реалізації другорядних завдань. Як основна мова програмування було обрано мову C++. Дана мова програмування має великі технічні можливості для роботи з інформацією представленою у табличній формі (зображення – це двовимірні матриці) та цифрові бібліотеки які постійно оновлюються та вдосконалюються. Для обробки файлів з цифровими зображеннями було використано бібліотеку функцій EasyBMP. EasyBMP - це легка та проста використанні бібліотека для роботи з зображеннями у форматі BMP на мові програмування C++. Заснована на простих структурах даних та операціях, бібліотека розроблена для забезпечення легкості використання та швидкості обробки BMP-зображень в програмах.

Основні задачі, які виконує дана бібліотека це:

Забезпечує можливість читати зображення BMP з файлу та записувати його назад у файл. Це включає зчитування та запис інформації про пікселі, колір, розмір, заголовки тощо.

Дозволяє здійснювати операції з пікселями, такі як отримання та встановлення значень кольору, маніпуляції альфа-каналом та інше.

Легко інтегрується в програми завдяки простим функціям та структурам даних. Вона особливо підходить для початківців, які шукають простий спосіб роботи з BMP-зображеннями.

Робота з кольоровими зображеннями та можливість використання 24-бітного та 32-бітного кольорових форматів.

Забезпечує можливість роботи на низькорівневому рівні, що дозволяє виконувати різноманітні операції, такі як зчитування і запис байтів.

Бібліотека EasyBMP створена з урахуванням простоти та легкості використання. Це робить її ідеальним інструментом для тих, хто шукає швидке вирішення для роботи з BMP-зображеннями без зайвого складнощів. EasyBMP

має відкритий вихідний код, що дозволяє користувачам переглядати, змінювати та розповсюджувати код з урахуванням ліцензії, що надається. Забезпечує спрощений інтерфейс для доступу до пікселів, кольорів та розмірів зображення, що полегшує взаємодію з BMP-зображеннями. Підтримка роботи з зображеннями у 24-бітному та 32-бітному кольорових форматах, що розширює можливості роботи з різними типами зображень. Завдяки своїм простим інструментам та можливостям EasyBMP знаходить широке застосування в різних областях програмування та розробки.

Для конвертації зображення з формату gif в формат bmp було використано засоби бібліотеки EasyBMP. Нижче наведено приклад функції на мові програмування C++, яка використовує EasyBMP для зчитування gif-зображення та збереження його у форматі bmp (рисунк 3.5):

```
void convertGifToBmp(const char* gifFileName, const  
char* bmpFileName) {  
    BMP image;  
    image.ReadFromFile(gifFileName);  
    image.WriteToFile(bmpFileName);  
}
```

Рисунок 3.5 – Функція переконвертації цифрового зображення з формату gif в формат bmp

Для виконання даної задачі використовується тип даних для роботи з цифровими зображеннями BMP, що реалізований в структурі цифрової бібліотеки. Він дозволяє самостійно представити один формат для зберігання цифрового зображення в інших за допомогою використання методів ReadFromFile() та WriteToFile(). У цьому коді функція convertGifToBmp приймає ім'я вхідного GIF-зображення та ім'я вихідного BMP-зображення. Вона використовує методи бібліотеки EasyBMP для зчитування та запису зображень.

Інша задача яка може бути виконана на етапі попередньої обробки зображення є бінаризація вхідного зображення за заданим пороговим значенням T . Бінаризація цифрового зображення у форматі BMP вимагає розуміння

структури файлу BMP та здатність зчитувати і записувати дані у цей формат. Нижче подано приклад C++-функції для бінаризації зображення BMP з використанням бібліотеки EasyBMP та використання параметра порогу який може бути адаптивно підібраний на основі додаткового аналізу цифрового зображення та використання допоміжних функцій (рисунок 3.6).

```
void binarizeBMPImage(const char* inputFileName, const char*
outputFileName, int threshold) {
    BMP image;
    image.ReadFromFile(inputFileName);
    for (int i = 0; i < image.TellWidth(); ++i) {
        for (int j = 0; j < image.TellHeight(); ++j) {
            RGBAPixel pixel = image.GetPixel(i, j);
            pixel.Red = pixel.Green = pixel.Blue = (pixel.Red
< threshold) ? 0 : 255;
            image.SetPixel(i, j, pixel);
        }
    }
    image.WriteToFile(outputFileName);
}
```

Рисунок 3.6 – Функція бінаризації цифрового зображення за заданим пороговим значенням

Алгоритм визначення нового значення для кожного пікселя цифрового зображення оснований на використанні тернорногооператора та заміни значення для всіх кольорових каналів передачі значення кольору.

```
RGBAPixel pixel = image.GetPixel(i, j);
pixel.Red = pixel.Green = pixel.Blue = (pixel.Red
< threshold) ? 0 : 255;
image.SetPixel(i, j, pixel);
```

Серед допоміжних функцій реалізований в програмному додатку слід відмітити функцію для автоматичного визначення та побудови гістограми розподілу рівнів сірого. Дана інформація воєористовується для обчислення значення адаптивного порогу при використанні алгоритмів порогової сегментації. Для побудови гістограми розподілу рівня яскравості на цифровому зображенні у форматі BMP воєористовуються технічні можливості бібліотеки EasyBMP для зчитування зображення та обчислення гістограми. Нижче подано функцію на мові програмування C++, яка реалізує цю задачу (рисунок 3.7).

```
void computeHistogram(const char* fileName, std::vector<int>&
histogram) {
    BMP image;
    image.ReadFromFile(fileName);
    histogram.assign(256, 0);
    for (int i = 0; i < image.TellWidth(); ++i) {
        for (int j = 0; j < image.TellHeight(); ++j) {
            RGBAPixel pixel = image.GetPixel(i, j);
            int brightness = (int)(0.299 * pixel.Red + 0.587
* pixel.Green + 0.114 * pixel.Blue);
            histogram[brightness]++;
        }
    }
}
```

Рисунок 3.7 – Функція побудови гістограми розподілу рівня сірого

У цьому коді функція computeHistogram обчислює гістограму розподілу рівня яскравості для зображення, а функція printHistogram виводить гістограму на екран. Гістограма обчислюється на основі значень яскравості, які визначаються за формулою:

$$brightness=0.299\times Red+0.587\times Green+0.114\times Blue.$$

Обрані коефіцієнти для формули перетворення є загально прийнятими і дозволяють максимально точно перекодувати кольорове значення в обраній точці зображення в сіре. Проте в програмі реалізована можливість корекції цих значень. Дана функція призначена для більшої адаптації роботи програмного додатку до поставлених завдань.

Окрім функції попередньої обробки, які носять додаткове навантаження на програмний додаток для виконання основної задачі по сегментації зображення реалізовано ряд функцій. Зокрема реалізовано функцію сегментації на основі підходу водоподілу. Програмний код даної функції наведено на рисунку 3.8.

```
void watershedSegmentation (const Mat& inputImage, Mat&
segmentedOutput) {
    Mat grayscaleImage, markerMask;
    cvtColor(inputImage, grayscaleImage, COLOR_BGR2GRAY);
    threshold(grayscaleImage, markerMask, 0, 255,
    THRESH_BINARY | THRESH_OTSU);
    Mat markers;
    connectedComponents(markerMask, markers, 8, CV_32S);
    watershed(inputImage, markers);
    segmentedOutput = Mat::zeros(inputImage.size(),
inputImage.type());
    for (int i = 0; i < markers.rows; ++i) {
        for (int j = 0; j < markers.cols; ++j) {
            int index = markers.at<int>(i, j);
            if (index == -1) {
                segmentedOutput.at<Vec3b>(i, j) = Vec3b(0, 0, 0
            } else {
                segmentedOutput.at<Vec3b>(i, j) =
inputImage.at<Vec3b>(i, j);
            }
        }
    }
}
```

Рисунок 3.8 – Функція сегментації на основі алгоритму водоподілу

Дана функція проводить розмітку всіх точок зображення при цьому значення міток обираються виходячи з того на якій висоті в гистограмі розподілу знаходиться та чи інша точка. Гістограма розподілу береться на основі аналізу вхідного цифрового зображення яке попередньо було переконвертоване в градації сірого. Дана функція дозволяє отримати масив одноріжних областей, який в подальшому буде підлягати додатковій процедурі об'єднання.

3.3 Тестування та аналіз реалізованого програмного додатку

Обробка зображень широко використовується фотографами, науковцями та медичними працівниками. Існують різні типи програм обробки зображень. Програмне забезпечення для обробки зображень фіксує та обробляє цифрові зображення. Воно відіграє важливу роль у перетворенні численних зображень у різні якості, носії та роздільну здатність. Частина цього процесу включає перетворення зображення в детальний цифровий формат. Найважливіше те, що обробка зображень є різким контрастом із редагуванням зображень. Щоб краще зрозуміти ідею обробки зображень, можна візуалізувати зображення як життєвий цикл із трьох фаз: обробки, редагування та публікації. Обробка – це початок, коли зображення є найбільш сирим. Редагування та публікація відбувається, коли зображення переходить у фазу, в якій можна маніпулювати. Тому очевидним є той факт, що для програмних додатків обробки цифрових зображень важливими є технічні характеристики робочих станцій на яких вони використовуються. Для проведення тестування реалізованого програмного додатку було обрано персональний комп'ютер із технічними характеристиками більшими за стандартні параметри офісного ПК. Це було зроблено для того, щоб мінімізувати вплив апаратної складової на процес тестування. Технічні параметри наведено на рисунку 3.9.

Назва параметра	Значення
Процесор (CPU):	Intel Core i5-9400 (6 ядер, 2.9 ГГц, до 4.1 ГГц Turbo Boost, 9 МБ кешу)
Оперативна пам'ять (RAM):	8 ГБ DDR4
Жорсткий диск:	SSD 512 ГБ
Графічний процесор (GPU):	Вбудована графіка Intel UHD Graphics 630
Операційна система:	Windows 10 Professional (64-bit)
Материнська плата:	MicroATX
Бездротові технології:	Wi-Fi 802.11ac, Bluetooth 5.0
Порти та роз'єми:	USB 3.0/3.1, HDMI, DisplayPort, Ethernet (RJ-45)
Корпус:	Mid Tower
Блок живлення:	400 Вт

Рисунок 3.9 – Технічні параметри робочої станції для тестування

Обрана конфігурація дозволяє проводити обробку цифрових зображень різної складності, розмірів та форматів. При цьому виконання повного циклу обробки та аналізу проводиться в режимі реального часу та дозволяє опрацювати велику кількість вхідних зображень для отримання більш об'єктивних результатів тестування.

Тестування програмного додатку проводилось в декілька етапів. На першому етапі здійснювалось ручне тестування оцінки зручності використання програмного додатку. Користувачі оцінювали інтерфейс та функціонал який доступний з головного вікна програмного додатку. Результати тестування наведено в таблиці 3.1.

Таблиця 3.1 – Результати тестування програмного інтерфейсу

№ тесту	Опис тесту	Очікуваний результат	Фактичний результат	Примітки
1	Загальна оцінка інтерфейсу	Високий рівень привабливості інтерфейсу	Графічний інтерфейс виявився зручним та привабливим	Тест пройдено
2	Швидкість опанування програмою	Мінімальні терміни для навчання роботи з програмою	Терміни навчання нових користувачів в залежності від досвіду роботи в не перевищують 1 год	Тест пройдено
3	Зручність інтерфейсу	Мінімальний час доступу до усіх функцій програмного додатку	Активація усіх запропонованих функцій програмного додатку відбувається за допомогою 1-2 натискань елементів програмного меню	Тест пройдено

Аналіз результатів проведеного тестування показав, що спроектований та реалізований графічний інтерфейс користувача є простим та інтуїтивно зрозумілим, не викликає великої кількості запитань та дозволяє використовувати програмну розробку користувачам з різним досвідом роботи.

Для отримання оцінки продуктивності програмного додатку процес тестування передбачав декілька етапів обробки цифрового зображення. Проміжні результати обробки наведено на рисунку 3.10.

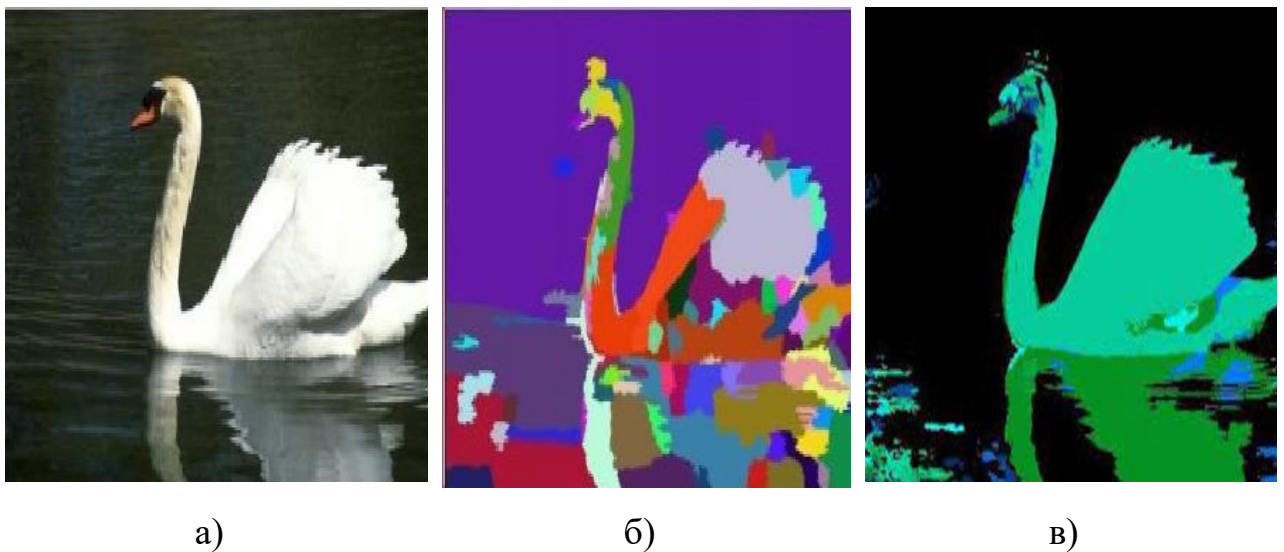


Рисунок 3.10 – Приклад результатів роботи програмного додатку: вхідне зображення (а), масив однорідних областей (б), результати сегментації (в)

Зокрема на рисунку показано приклад вхідного зображення, яке має наступні характеристики: кольорове зображення з невеликою кількістю однорідних областей які відповідають об'єкту. Також на наведено проміжний результат роботи програмного додатку на якому виділені однорідні області, проте вони не об'єднані в загальну область яка відповідатиме об'єкту. Даний результат отримується на основі використання алгоритмів попередньої сегментації. Також на зображенні проілюстровано кінцевий результат сегментації на якому виділено окремі елементи на вхідному зображенні.

Для отримання більш об'єктивних результатів роботи програмного додатку було проведено ряд тестів на основі груп зображень. Для класифікації вхідних

зображень було обрано ряд критеріїв які безпосередньо впливають на результати роботи алгоритму сегментації . Серед обраних критеріїв:

- розмір зображення;
- кількість окремих об'єктів на зображенні;
- рівень відмінності між сусідніми об'єктами;
- рівень зашумленості зображення;
- кількість кольорів на зображенні.

Базуючись на обраних критеріях всі вхідні зображення були розділені на чотири основні групи:

- прості зображення – мінімальна кількість різних кольорів та об'єктів (1-3 шт), відсутні шуми, невеликі за розміром, чіткі границі між сусідніми областями;
- звичайні зображення – кількість об'єктів на зображенні невелика (3-5 шт), середня кольорова гама, можливі присутні сторонні артефакти або шуми, розмір зображень від малого до середнього, сусідні області чітко розділяються;
- складні зображення – велика кількість об'єктів на зображенні (5+ шт), різноманітна кольорова гама, присутні сторонні артефакти або шуми, розмір зображень від середнього до великого, сусідні області не завжди чітко розділяються;
- штучно ускладнені. Дана група зображень була сформована на основі групи складних зображень для перевірки можливостей програмного додатку при опрацюванні зображень з низькою якістю. На зображення було додатков додано шуми та зображення модифікувались за допомогою фільтрації (розмиття).

Для підсумкової оцінки запропонованого алгоритму сегментації зображень на основі однорідності областей та на основі отриманих результатів тестування груп цифрових зображень можна зробити висновки, що оптимальними для сегментації є зображення з кількістю однорідних на них областей від 10 до 45 (рисунк 3.11). Така кількість дозволяє з високим рівнем точності описати присутні на зображенні об'єкти, при цьому загальний час обробки зображення є мінімальним. Якщо на зображенні виділялось менше

областей, то границі отриманих підсумкових об'єктів мають похибку в порівнянні з експертними результатами. Велика кількість областей з одого буку збільшує час обробки зображення, а з іншого ускладнює процес об'єднання невеликих областей, оскільки характеристики можуть мало відрізнятись.



Рисунок 3.11 – Порівняння результатів сегментації цифрових зображень на основі кількості виділених однорідних областей

Аналіз результатів за іншими показниками показав, що розмір вхідного зображення не впливає на результати етапу сегментації, він лише збільшує час на опрацювання одного зображення. Зашумленість зображення має прямий вплив на результати сегментації, проте якщо рівень зашумленості є невеликим, то даний фактором можна знехтувати. Кількість об'єктів на цифровому зображенні мало впливає на результати сегментації, більший вплив мають фактори розміру об'єктів та їх взаєморозташування. Сегментація зображення з невеликими об'єктами має нижчий рівень точності ніж результати отримані при аналізі зображень з такою самою кількістю об'єктів, проте вони мають великі розміри. Кількість кольорів які використовуються на зображенні має відносний вплив, у більшій мірі на результати сегментації впливають рівень відмінності

між різними кольорами. Інші результати отримані в процесі тестування були згруповані на основі обраних критеріїв та наведені в таблиці 3.2

Таблиця 3.2 – Узагальнена таблиця результатів тестування

Тестові групи	Точність сегментації	Вплив шумів	Кількість об'єктів на зображенні	Кольорова гама	Чіткість границь
Прості зображення	Висока точність	Відсутній	Незначний вплив	Незначний вплив	Незначний вплив
Звичайні зображення	Висока точність	Низький	Присутній вплив на результат	Присутній вплив на результат	Присутній вплив на результат
Складні зображення	Середня точність	Високий	Значний вплив на результат	Значний вплив на результат	Значний вплив на результат
Штучно ускладнені	Низька точність	Високий	Значний вплив на результат	Значний вплив на результат	Значний вплив на результат

На основі аналізу моделювання запропонованого алгоритму та результатів тестування програмного додатку обробки та аналізу зображень можна виділити наступні переваги алгоритму сегментації на основі однорідності областей:

- можливість використання для обробки та аналізу цифрових зображень різного типу та формату;
- простота реалізації та підтримки програмного додатку;
- зручний та зрозумілий інтерфейс, що підвищує зручність використання розробки;
- висока точність отриманих результатів;
- робота в реальному часі, при використанні великих груп зображень процес опрацювання не великий.

Недоліками розробленого алгоритму можна вважати:

- вплив на результати роботи рівня зашумленості цифрового зображення;
- необхідність використання додаткової первинної обробки для отримання масивів однорідних областей при попередній обробці зображення.

3.4 Висновки до розділу

Спроековано внутрішню архітектуру програмного додатку обробки цифрових зображень та проведено її моделювання на основі біаграм послідовності та прецедентів, що дозволило реалізувати додаток для опрацювання цифрових зображень.

Здійснено функціональне та практичне тестування реалізованого програмного продукту на основі сегментації цифрових зображень різної складності, що підтвердило високий рівень подібності отриманих результатів в порівнянні з експертними результатами.

ВИСНОВКИ

На основі аналізу існуючих систем аналізу та оцінки цифрових зображень, а також запропонованому алгоритмові та структурі програмної системи можна зробити такі висновки:

1. Проведено дослідження та класифікацію цифрових зображень на основі технологій їх формування та збереження, що дозволило виділити основні характеристики які містять максимальну кількість інформації про цифрове зображення;

2. Здійснено аналіз технологій обробки цифрових зображень на основі сфер їх застосування, що дозволило виділити основні вимоги до алгоритмів сегментації й цифрових зображень;

3. Проведено аналітичний огляд програмних засобів обробки та аналізу цифрових зображень на основі дослідження їх основних складових та функціональних можливостей, що дозволило визначити основні структурні елементи які реалізуються в сучасних програмних системах обробки цифрових.

4. Проведено аналітичне дослідження алгоритмів сегментації цифрових зображень на основі критеріїв поділу на однорідні області, що дозволило виділити групу алгоритмів результати яких максимально подібні до результатів отриманих за допомогою експертів.

5. Розроблено алгоритм сегментації цифрових зображень на основі аналізу однорідних областей, що дозволило спроектувати структуру програмного додатку обробки та аналізу цифрових зображень.

6. Здійснено функціональне та практичне тестування реалізованого програмного продукту на основі сегментації цифрових зображень різної складності, що підтвердило високий рівень подібності отриманих результатів в порівнянні з експертними результатами.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Горішній Я.І., Шевчук В.В.. Алгоритм сегментації цифрових зображень на основі аналізу характеристик однорідних областей. Збірник тез VIII Науково-практична конференція молодих вчених і студентів «Інтелектуальні комп'ютерні системи та мережі», Тернопіль, 05 грудня 2023 р. с. 11.
2. Гриненко А.А., Горішній Я.І. Алгоритм опрацювання та аналізу об'єктів на рухомих зображеннях. Збірник тез VIII Науково-практична конференція молодих вчених і студентів «Інтелектуальні комп'ютерні системи та мережі», Тернопіль, 05 грудня 2023 р. с. 12.
3. Xiao, P., Guo, W., Liu, M., & Liu, B. A Three-Step Imaging Algorithm for the Constellation of Geostationary and Low Earth Orbit SAR (ConGaLSAR). IEEE Transactions on Geoscience and Remote Sensing, 60, - (2021).- 1-14.
4. Sun, Y., Hua, Y., Mou, L., Conditional GIS-Aware Network for Individual Building Segmentation in a VHR SAR Image. In 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. – 2021. - pp. 432-435.
5. Jin, X., Gu, Y., Intrinsic Hyperspectral Image Decomposition With DSM Cues. IEEE Transactions on Geoscience and Remote Sensing, 60, - 2021.- 1-13.
6. Zhuang, L., Gao, L., Zhang, B., Fu, X., & Bioucas-Dias, J. M. Hyperspectral image denoising and anomaly detection based on low-rank and sparse representations. IEEE Transactions on Geoscience and Remote Sensing, 60. – 2020 - 1-17.
7. A. Arnab and P. H. Torr. Pixelwise instance segmentation with a dynamically instantiated network. In CVPR, 2017.
8. M. Bai and R. Urtasun. Deep watershed transform for instance segmentation. In CVPR, 2017.
9. H. Caesar, J. Uijlings, and V. Ferrari. COCO-Stuff: Thing and stuff classes in context. In CVPR, 2018.

10. L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. PAMI, 2018.
11. M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele. The cityscapes dataset for semantic urban scene understanding. In CVPR, 2016.
12. C. Desai, D. Ramanan, and C. C. Fowlkes. Discriminative models for multi-class object layout. IJCV, 2011.
13. P. Dollar, C. Wojek, B. Schiele, and P. Perona. Pedestrian detection: An evaluation of the state of the art. PAMI, 2012.
14. M. Everingham, S. A. Eslami, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman. The PASCAL visual object classes challenge: A retrospective. IJCV, 2015.
15. A. Fathi, N. Kanazawa, and K. Murphy. Places challenge 2017: instance segmentation, G-RMI team. 2017.
16. A. Fathi, K. Yang, and K. Murphy. Places challenge 2017: scene parsing, G-RMI team. 2017.
17. J. Fu, J. Liu, L. Guo, H. Tian, F. Liu, H. Lu, Y. Li, Y. Bao, and W. Yan. Places challenge 2017: scene parsing, CASIA IVA JD team. 2017.
18. B. Hariharan, P. Arbelaez, R. Girshick, and J. Malik. Simultaneous detection and segmentation. In ECCV, 2014.
19. K. He, G. Gkioxari, P. Dollar, and R. Girshick. Mask R-CNN. In ICCV, 2017.
20. J. Hosang, R. Benenson, P. Dollar, and B. Schiele. What makes for effective detection proposals? PAMI, 2015.
21. J. Hosang, R. Benenson, and B. Schiele. Learning non-maximum suppression. PAMI, 2017.
22. A. Kirillov, R. Girshick, K. He, and P. Dollar. Panoptic feature pyramid networks. In CVPR, 2019.
23. A. Kirillov, E. Levinkov, B. Andres, B. Savchynskyy, and C. Rother. InstanceCut: from edges to instances with multicut. In CVPR, 2017.

- 24.I. Kokkinos. UberNet: Training a universal convolutional neural network for high-level vision using diverse datasets and limited memory. In CVPR, 2017.
- 25.A. Krizhevsky, I. Sutskever, and G. Hinton. ImageNet classification with deep convolutional neural networks. In NIPS, 2012.
- 26.J. Li, A. Raventos, A. Bhargava, T. Tagawa, and A. Gaidon. Learning to fuse things and stuff. arXiv:1812.01192, 2018.
- 27.Q. Li, A. Arnab, and P. H. Torr. Weakly-and semi-supervised panoptic segmentation. In ECCV, 2018.
- 28.Y. Li, X. Chen, Z. Zhu, L. Xie, G. Huang, D. Du, and X. Wang. Attention-guided unified network for panoptic segmentation. arXiv:1812.03904, 2018.
- 29.Y. Li, H. Qi, J. Dai, X. Ji, and Y. Wei. Fully convolutional instanceaware semantic segmentation. In CVPR, 2017.
- 30.T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, and C. L. Zitnick. Microsoft COCO: Common objects in ' context. In ECCV, 2014.
- 31.C. Liu, J. Yuen, and A. Torralba. SIFT flow: Dense correspondence across scenes and its applications. PAMI, 2011.
- 32.H. Liu, C. Peng, C. Yu, J. Wang, X. Liu, G. Yu, and W. Jiang. An end-to-end network for panoptic segmentation. arXiv:1903.05027, 2019.
- 33.S. Liu, J. Jia, S. Fidler, and R. Urtasun. SGN: Sequential grouping networks for instance segmentation. In CVPR, 2017.
- 34.S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia. LSUN'17: insatnce segmentation task, UCenter winner team. 2017.
- 35.J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In CVPR, 2015.
- 36.R. Luo, B. Jiang, T. Xiao, C. Peng, Y. Jiang, Z. Li, X. Zhang, J. Sun. Places challenge 2017: instance segmentation, Megvii (Face++) team. 2017.
- 37.J. Malik, P. Arbelaez, J. Carreira, K. Fragkiadaki, R. Girshick, ' G. Gkioxari, S. Gupta, B. Hariharan, A. Kar, and S. Tulsiani. The three R's of computer vision: Recognition, reconstruction and reorganization. PRL, 2016.

- 38.D. R. Martin, C. C. Fowlkes, and J. Malik. Learning to detect natural image boundaries using local brightness, color, and texture cues. PAMI, 2004.
- 39.I. Misra, A. Shrivastava, A. Gupta, and M. Hebert. Cross-stitch networks for multi-task learning. In CVPR, 2016.
- 40.G. Neuhold, T. Ollmann, S. Rota Bulò, and P. Kotschieder. The `mapillary vistas dataset for semantic understanding of street scenes. In CVPR, 2017.
- 41.P. O. Pinheiro, R. Collobert, and P. Dollar. Learning to segment ´ object candidates. In NIPS, 2015.
- 42.S. Ren, K. He, R. Girshick, and J. Sun. Faster R-CNN: Towards realtime object detection with region proposal networks. In NIPS, 2015.
- 43.O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. IJCV, 2015.
- 44.J. Shotton, J. Winn, C. Rother, and A. Criminisi. Textonboost: Joint appearance, shape and context modeling for multi-class object recog. and segm. In ECCV, 2006.
- 45.M. Sun, B. Kim, P. Kohli, and S. Savarese. Relating things and stuff via object property interactions. PAMI, 2014.
- 46.J. Tighe and S. Lazebnik. Finding things: Image parsing with regions and per-exemplar detectors. In CVPR, 2013.
- 47.J. Tighe, M. Niethammer, and S. Lazebnik. Scene parsing with object instances and occlusion ordering. In CVPR, 2014.
- 48.Z. Tu, X. Chen, A. L. Yuille, and S.-C. Zhu. Image parsing: Unifying segmentation, detection, and recognition. IJCV, 2005.
- 49.R. Vaillant, C. Monrocq. Original approach for the localisation of objects in images. IEE Proc. on Vision, Image, and Signal Processing, 1994.
- 50.P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In CVPR, 2001.
- 51.D. B. West. Introduction to graph theory, volume 2. Prentice hall Upper Saddle River, 2001.

52.Y. Xiong, R. Liao, H. Zhao, R. Hu, M. Bai, E. Yumer, and R. Urtasun. UPSNet: A unified panoptic segmentation network. arXiv:1901.03784, 2019.

53.T.-J. Yang, M. D. Collins, Y. Zhu, J.-J. Hwang, T. Liu, X. Zhang, V. Sze, G. Papandreou. DeeperLab: Single-shot image parser. arXiv:1902.05093, 2019.

54.Y. Yang, S. Hallman, D. Ramanan, and C. C. Fowlkes. Layered object models for image segmentation. PAMI, 2012.

55.J. Yao, S. Fidler, and R. Urtasun. Describing the scene as a whole: Joint object detection, scene classification and semantic segmentation. In CVPR, 2012.

56.F. Yu and V. Koltun. Multi-scale context aggregation by dilated convolutions. In ICLR, 2016.

57.Y. Zhang, H. Zhao, and J. Shi. LSUN'17: semantic segmentation task, PSPNet winner team. 2017.

58.H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia. Pyramid scene parsing network. In CVPR, 2017. 8 [55] B. Zhou, H. Zhao, X. Puig, S. Fidler, A. Barriuso, and A. Torralba. Scene parsing through ADE20K dataset. In CVPR, 2017. 2, 3, 5, 6 [56] Y. Zhu, Y. Tian, D. Mexatas, and P. Dollar. Semantic amodal segmentation. In CVPR, 2017.

59.Березький О.М., Дубчак Л.О., Мельник Г.М. Методичні рекомендації до виконання кваліфікаційної роботи з освітнього ступеня “Магістр”. Спеціальність: 123 - Комп’ютерна інженерія. Магістерська програма - Комп’ютерна інженерія". Тернопіль: ЗУНУ, 2022. 32 с.

60.Гураль І.В., Дубчак Л.О. Методичні вказівки до оформлення курсових проектів, звітів про проходження практики, випускних кваліфікаційних робіт для студентів спеціальності «Комп’ютерна інженерія» Тернопіль: ТНЕУ, 2019. 33 с.