

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ
Факультет аеронавігації, електроніки та телекомунікацій
Кафедра авіаційних комп'ютерно-інтегрованих комплексів

ДОПУСТИТИ ДО ЗАХИСТУ
Завідувач випускової кафедри
_____ Віктор СИНЕГЛАЗОВ
“ ___ ” _____ 2023 р.

КВАЛІФІКАЦІЙНА РОБОТА
(ПОЯСНЮВАЛЬНА ЗАПИСКА)
ВИПУСКНИКА ОСВІТНЬОГО СТУПЕНЯ
“МАГІСТР”

Спеціальність 151 «Автоматизація та комп'ютерно-інтегровані технології»
Освітньо-професійна програма «Комп'ютерно-інтегровані технологічні процеси і виробництва»

Тема: Система оптимального вибору методів комп'ютерного зору для обробки зображень

Виконавець: студент групи КП-226М Хейлик Едуард Володимирович

Керівник: проф., д.т.н. Синеглазов Віктор Михайлович

Консультант розділу «Охорона навколишнього середовища» _____ Ольховик Ю.О.
(підпис)

Консультант розділу «Охорона праці» _____ Козлітін О.О.
(підпис)

Нормоконтролер: _____ Філяшкін М.К
(підпис)

Київ - 2023

НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ
Факультет аеронавігації, електроніки та телекомунікацій
Кафедра авіаційних комп'ютерно-інтегрованих комплексів

Освітній ступінь: магістр

Спеціальність 151 «Автоматизація та комп'ютерно-інтегровані технології»

Освітньо-професійна програма «Комп'ютерно-інтегровані технологічні процеси і виробництва»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Віктор СИНЄГЛАЗОВ
“ ____ ” _____ 2023 р.

ЗАВДАННЯ

на виконання кваліфікаційної роботи студента

Хейлика Едуарда Володимировича

1. **Тема роботи :** «Система оптимального вибору методів комп'ютерного зору для обробки зображень»
2. **Термін виконання роботи:** з 19.09.2023р. до 15.12.2023р.
3. **Вихідні дані до роботи:**
4. **Зміст пояснювальної записки:**
5. **Перелік обов'язкового графічного матеріалу:** схеми, рисунки, таблиці.
6. **Календарний план-графік:**

Етапи виконання дипломної роботи	Терміни виконання	Підпис керівника
1. Огляд та обробка літератури за темою кваліфікаційної роботи	19.09.2023 –29.09.2023	
2. Дослідження особливостей штучного інтелекту	13.10.2023 - 19.10.2023	

3. Дослідження розв'язку задач сегментації	19.10.2023 - 27.10.2023	
4. Дослідження розв'язку задач класифікації	01.11.2023 - 30.11.2023	
5. Розв'язання задач сегментації, визначення ознак, класифікації.	05.11.2023 - 29.11.2023	
6. Висновки	10.12.2023 - 15.12.2023	

7. Консультанти з окремих розділів роботи:

Розділ	Консультант (посада, П.І.Б.)	Дата, підпис	
		Завдання видав	Завдання прийняв
Охорона праці	Старший викладач Ольховик Ю.О.		
Охорона навколишнього середовища	Канд. біолог. наук, доцент Козлітін О.О.		

8. Дата видачі завдання _____

Керівник: _____ Синеглазов В.М.

Завдання прийняв до виконання _____ Хейлик Е.В.

_____ ” _____ 2023 р.

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи «Система оптимального вибору методів комп'ютерного зору для обробки зображень» 111с., 26 рис., 7 табл, 23 джерела.

ШТУЧНИЙ ІНТЕЛЕКТ, РОЗВ'ЯЗАННЯ ЗАДАЧ СЕГМЕНТАЦІЇ,
КОМП'ЮТЕРНИЙ ЗІР, НЕЙРОННА МЕРЕЖА, ПОРОГОВА ОБРОБКА.

Об'єкт дослідження – система оптимального вибору методів комп'ютерного зору.

Предмет дослідження – методи комп'ютерного зору та їх застосування в різних сценаріях.

Мета кваліфікаційної роботи – дослідження та порівняння різних методів комп'ютерного зору з точки зору їхньої точності, швидкості та зручності застосування .

Метод дослідження – порівняльний аналіз, обробка літературних джерел , експериментальне дослідження.

У дипломній роботі проводиться аналіз методів комп'ютерного зору. Робота включає огляд сегментації зображень, методи її проведення.

Ця робота визначає важливість вдосконалення системи вибору методів комп'ютерного зору для обробки зображень. Прогнозується розвиток систем, які можуть автоматично адаптуватися до змінних умов та вибирати оптимальні методи в залежності від конкретного завдання.

ЗМІСТ

Вступ.....	7
РОЗДІЛ 1	8
ШТУЧНИЙ ІНТЕЛЕКТ ТА ЙОГО ОСОБЛИВОСТІ	8
1.1 Розвиток штучного інтелекту в Україні	8
1.2 Складові штучного інтелекту	11
1.2.1 Нейронні мережі	11
1.2.2 Класифікатори в штучних нейронних мережах	13
1.3 Класифікація нейронних мереж.....	19
1.4 Методи машинного навчання	22
РОЗДІЛ 2	32
ЗАДАЧА СЕГМЕНТАЦІЇ ЗОБРАЖЕНЬ	32
2.1 Огляд робіт.....	32
2.2 Критерії для сегментації при обробці зображень	42
РОЗДІЛ 3	49
Огляд методів сегментації для обробки зображень.....	49
3.1 Класифікація технік сегментації зображень	51
3.2 Підходи до сегментації зображень з використанням порогової обробки.....	52
3.2.1 Глобальна порогова обробка	52
3.2.2 Місцеве порогове значення	64
3.2.3 Динамічна бінаризація	70
3.3 Техніки сегментації на основі країв	72
3.3.1 Метод сіро-гістограми	73
3.3.2 Градієнтний підхід.....	73
3.4 Методи сегментації на основі регіону	88
3.4.1 Методи збільшення регіону.....	89
3.4.2 Методи поділу та злиття областей.....	92
3.5. Методи кластеризації.....	93
3.5.1 Жорстка кластеризація.....	94
3.5.2 М'яка кластеризація	95
3.6. Методи глибокої нейронної мережі	97
3.6.1 Нейронна мережа згортки.....	97
3.6.2 Повторювані нейронні мережі (RNN)	99

3.6.3 Моделі кодера-декодера та автокодера.....	101
3.6.4 Машини опорних векторів (SVM)	102
3.7. Гібридні методи.....	104
3.7.1 Еволюційні підходи.....	104
3.7.2 Нечітка логіка.....	104
3.7.3 Ройовий інтелект (SI)	105
РОЗДІЛ 4	110
4.1 Вступ.....	110
4.2 Аналіз умов праці.....	110
4.3 Розробка заходів з охорони праці.....	114
4.4. Пожежна безпека.....	115
4.5 Перевірочний розрахунок штучного освітлення виробничого або офісного приміщення	118
4.6 Висновки	119
РОЗДІЛ 5	120
5.1 Важкі метали і їх негативний вплив на людину і довкілля	120
5.2 Кількість металів у гаджетах і їх загальний вплив	122
5.3 Утилізація важких металів у гаджетах та екологічні виклики.....	125
ВИСНОВКИ.....	109
СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ	109

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

CNN - Згорткова нейронна мережа

RNN- Повторювана нейронна мережа

SRG – Метод зростання регіону

FCM – Fuzzy C-Means, нечіткі C-середні

SVM – Метод опорних векторів

Вступ

В сучасному інформаційному суспільстві, де величезна кількість даних генерується щоденно, обробка зображень є ключовим елементом багатьох наукових та практичних застосувань. За останні десятиліття значно зросла активність у галузі комп'ютерного зору, де вивчається взаємодія комп'ютерів із зображеннями для вирішення різноманітних завдань, таких як розпізнавання об'єктів, визначення патернів, аналіз поведінки і багато інших.

Ця дипломна робота присвячена дослідженню та розвитку системи оптимального вибору методів комп'ютерного зору для ефективної обробки зображень. Однією з найважливіших задач у цьому контексті є визначення оптимального методу залежно від конкретного завдання, що дозволяє досягти максимальної точності та швидкості обробки. У даній роботі розглядаються сучасні методи комп'ютерного зору, їхні переваги та обмеження, а також пропонується система автоматизованого вибору методу в залежності від поставленого завдання.

Основною метою цього дослідження є створення ефективної та гнучкої системи, яка забезпечить оптимальний вибір методів комп'ютерного зору для різноманітних застосувань, починаючи від медичної діагностики та закінчуючи виробничими процесами та автономними системами.

РОЗДІЛ 1

ШТУЧНИЙ ІНТЕЛЕКТ ТА ЙОГО ОСОБЛИВОСТІ

1.1 Розвиток штучного інтелекту в Україні

Штучний інтелект (ШІ) (англ. Artificial intelligence) – розділ комп'ютерної лінгвістики та інформатики, що опікується формалізацією проблем та завдань, які подібні до дій, що виконує людина. Це поняття ввів у 1956 р. професор Дартмутського коледжу Джон МакКарті, який цікавився, чи можна навчити машину, як і дитину – абстрактним поняттям, використовувати мову і самостійно вдосконалюватись методом спроб і помилок. Зі ШІ ми стикаємось кожного дня. Це голосовий пошук – Siri і Alexa, які доступні на iOS, Android і Windows, відеоігри – персонажі, яких можуть непередбачувано поводитись для гравця, автономні авто які можуть самі аналізувати ситуацію на дорозі і діяти, онлайн підтримка клієнтів на сайтах, рекомендації товарів, які вас можуть зацікавити, що створюється внаслідок аналізу відвідуваних вами Інтернет-сторінок. На новинних порталах роботи створюють фінансові звіти, спортивні репортажі і замітки тощо.

Компанія Deep Knowledge Analytics склала рейтинг Artificial Intelligence Industry in Eastern Europe 2018 за кількістю компаній, що працюють у сфері ШІ [2]. Україна входить до трійки лідерів серед країн Східної Європи (1-е місце посіла Росія – 133 компаній, 2-е – Польща – 110 компаній, в Україні працює 57 компаній в галузі ШІ). Наша країна налічує 26 аутсорсинг-компаній, а у світі їх лише 226. Згідно з даними ресурсу LinkedIn, в країні понад 2 тис компаній-розробників у сфері ШІ. Більша частка розробок відведена на програмне забезпечення, інформаційні технології, чатботи та віртуальні асистенти, розважальні продукти тощо.

Проблемами ШІ в Україні активно займаються Інститут кібернетики ім. В.М. Глушкова, Інститут проблем штучного інтелекту НАН України, Харківський національний університет радіоелектроніки (кафедра штучного інтелекту) та інші наукові установи. Багато наукових розробок учених Інституту проблем ШІ НАН України належать до пріоритетних напрямів, вони виконуються, зокрема, в межах цільової НТП НАН України «Дослідження і розробки з проблем підвищення

обороздатності і безпеки держави». Для потреб оборонної сфери створені: комп'ютерні системи, здатні знищувати безпілотні літальні апарати; спеціальні шоломи для корекції психофізіологічного стану людини після стресової ситуації (в т. ч. в бойових умовах); довгострокова вогнева точка, обладнана інтелектуальною системою для автоматичної ідентифікації та (за потреби) ліквідації рухомих об'єктів [3].

Соціальний аспект. У 2018 р. Інститут Горшеніна спільно з компанією Everest провів дослідження «Штучний інтелект: український вимір», у якому взяло участь 1000 респондентів віком від 16 до 65 років [4]. 84,7 % українців чули термін «штучний інтелект». На питання, чи відчуваєте ви вплив технологій ШІ на ваше життя вже сьогодні, 74,1 % відповіли «так» і «скоріше так». 80,2 % опитаних вважають, що ефект ШІ на наше суспільство є позитивним. У 22,8 % людей поява ШІ викликає страх і тривогу. 53,9 % – зацікавленість. 20,1 % вважають, що ШІ – це загроза для людства. Соціологічне опитування чітко показує, що українці не просто зацікавлені в інноваціях – вони готові до конкретних дій з боку держави на їх повсюдне впровадження. Від того, якими будуть ці дії, безпосередньо залежить рішення багатьох ключових проблем у сферах освіти, економіки, медицини, бізнесу, науки та суспільства загалом.

Проблеми правового регулювання. З технологічним розвитком ШІ загострилися питання правового регулювання. Найбільш практичні заходи в цій сфері, вживаються в Європейському Союзі. Європейський Парламент ухвалив Резолюцію 2015/2103(INL) щодо цивільно-правового регулювання робототехніки з рекомендаціями для Європейської Комісії (англ. European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)) [5]. Аналіз змісту документу, дозволяє виділити кілька основних блоків питань. Це, зокрема: основні суспільні, економічні, етичні та правові проблеми та виклики, які необхідно врегулювати у зв'язку із розвитком робототехніки та ШІ; питання регулювання розробки та використання робототехніки на сучасному етапі; вимоги щодо стандартизації при розробці відповідних технологій; питання контролю за прийняттям рішень при використанні технологій робототехніки та ШІ; створення

інституціолізованої системи контролю в сфері робототехніки та ШІ; питання цивільно-правової відповідальності у зв'язку з розробкою та використанням робототехніки та ШІ; забезпечення захисту персональних даних при використанні цих технологій. Підсумовуючи, варто зазначити, що ця Резолюція є одним з перших реальних кроків на шляху до законодавчого закріплення стандартів розробки та використання ШІ. Попри те, що вона має рекомендаційний характер її положення дають можливість сформулювати уявлення про те, що буде лежати в основі норм, що регламентуватимуть відповідну діяльність в найближчій перспективі. Наступним важливим кроком є закріплення цих аспектів на законодавчому рівні, як це зараз роблять на Заході, де давно є розуміння прямого впливу інновацій на загальні показники розвитку держави.

Штучний інтелект (ШІ) в сучасному світі визначає новий етап розвитку технологій, який активно впроваджується в різні галузі людської діяльності. Однією з ключових складових штучного інтелекту є штучні нейронні мережі (ШНМ), які стали основою для досягнень у багатьох сферах, таких як розпізнавання образів, мовленнєві технології, медицина, та багато інших.

1.2 Складові штучного інтелекту

1.2.1. Нейронні мережі

Штучні нейронні мережі - це система обчислень, інспірована роботою людського мозку, яка може вирішувати різноманітні завдання завдяки здатності нейронів взаємодіяти між собою. Одна з ключових переваг ШНМ полягає в їхній здатності навчатися на прикладах та самостійно підлаштовувати свої ваги для оптимального вирішення завдань.

Нейронні мережі є ключовим елементом штучного інтелекту, відіграючи важливу роль у вирішенні різноманітних завдань, від зображення до мови та вирішення складних проблем. Ці штучні структури, інспіровані біологічними нейронними системами, дозволяють моделювати та розуміти складні залежності в даних. У цьому розділі розглянемо принципи дії нейронних мереж, їхню структуру, види та важливість у сучасному штучному інтелекті.

Основні принципи роботи нейронних мереж:

1. Нейрони:

- Нейронні мережі складаються з імітованих нейронів, які виконують функцію обробки інформації. Кожен нейрон приймає вхід, обчислює внутрішній стан та генерує вихід.

2. Ваги та зсуви:

- Кожен вхід нейрона має вагу, що регулює його вплив на роботу нейрона. Зсув визначає базовий внесок нейрона.

3. Функція активації:

- Функція активації визначає, як сума вагованих входів нейрона перетворюється в його вихід. Це дозволяє нейрону навчатися взаємодіяти з різними видами вхідних даних.

Структура нейронних мереж:

1. Шари:

- Нейронні мережі складаються зі шарів: вхідного, прихованого та вихідного. Вхідний шар приймає вхідні дані, прихований оброблює інформацію, а вихідний генерує результат.

2. Пов'язаність:

- Нейрони в одному шарі пов'язані з нейронами попереднього та наступного шарів. Ці зв'язки мають ваги, які навчаються під час тренування.

Види нейронних мереж:

1. Прямий прохід (Feedforward Neural Networks):

- Інформація рухається від вхідного до вихідного шару без циклічного зв'язку. Використовується для класифікації та регресії.

2. Зворотні мережі (Recurrent Neural Networks - RNN):

- Мають циклічні зв'язки, що дозволяють зберігати та використовувати інформацію з попередніх кроків. Використовуються для обробки послідовних даних.

3. Згорткові нейронні мережі (Convolutional Neural Networks - CNN):

- Спеціалізуються на обробці зображень та матриць.

Використовують згорткові шари для виявлення локальних особливостей.

Застосування нейронних мереж:

1. Обробка зображень та відео:
 - CNN використовуються для розпізнавання об'єктів та обробки великих обсягів візуальних даних.
2. Мовний аналіз та обробка природної мови:
 - RNN та LSTM використовуються для розуміння та генерації тексту.
3. Прогнозування та регресія:
 - Прямі прохідні мережі використовуються для прогнозування та аналізу часових рядів.

Важливість нейронних мереж у сучасному штучному інтелекті:

Нейронні мережі є основою багатьох сучасних застосувань штучного інтелекту.

З їх допомогою досягаються значні успіхи у вирішенні завдань, які раніше вважалися складними, включаючи розпізнавання образів, розуміння природної мови, аналіз даних та автоматичне прийняття рішень. Нейронні мережі дозволяють створювати моделі, які самостійно навчаються і покращують свою продуктивність з часом, що робить їх потужним інструментом у сфері штучного інтелекту.

1.2.2 Класифікатори в штучних нейронних мережах

У штучних нейронних мережах (НМ) класифікатори використовуються для призначення вхідного об'єкта до конкретного класу чи категорії. Класифікація є однією з основних задач, які вирішуються за допомогою нейронних мереж. Нижче розглянемо кілька типів класифікаторів, які широко використовуються у штучних нейронних мережах:

1. Древа рішень (Decision Trees)

Древа рішень використовуються в ШНМ для класифікації об'єктів за допомогою послідовного дерева вирішень. Кожен вузол представляє рішення на основі конкретного атрибуту даних.

2. Випадковий ліс (Random Forest)

Випадковий ліс - це ансамбль дерев рішень, які об'єднуються для зниження перенавчання та покращення точності прогнозів. Використання випадкового лісу в ШНМ може поліпшити його загальну ефективність.

3. Глибокі нейронні мережі (Deep Neural Networks)

Глибокі нейронні мережі складаються з багатьох шарів нейронів та використовуються для розв'язання складних завдань, таких як розпізнавання зображень чи мовленнєва розпізнавання.

4. Метод опорних векторів (Support Vector Machines, SVM)

SVM є ефективним методом класифікації в ШНМ, особливо в тих випадках, коли дані не лінійно роздільні. Вони базуються на пошуку оптимального розділяючого гіперплощини в просторі ознак.

5. Логістична регресія (Logistic Regression)

Логістична регресія використовується для бінарної та мультикласової класифікації. Вона стежить за тим, як ймовірність належності об'єкта до певного класу змінюється відносно значень його ознак.

6. Наївний Баєсівський класифікатор (Naive Bayes)

Цей класифікатор ґрунтується на теорії ймовірностей та використовується для класифікації з урахуванням ймовірності виникнення певних ознак у класі.

7. К-найближчих сусідів (k-Nearest Neighbors, k-NN)

k-NN визначає клас об'єкта на основі класів його найближчих сусідів у просторі ознак. Цей метод використовується як для класифікації, так і для регресії.

8. Градієнтний бустінг (Gradient Boosting)

Градiєнтний бустінг є ансамблевим методом, який комбiнує декілька слабших моделей для отримання сильної моделі. Його можна використовувати як для класифікації, так і для регресії.

9. Метод головних компонентів (Principal Component Analysis, PCA)

PCA використовується для зменшення розмірності даних, зберігаючи при цьому максимальну частку їхньої варіації. Використовується як попередній етап перед використанням інших класифікаторів.

Ці класифікатори представляють лише невеличкий перелік доступних методів, і кожен з них має свої унікальні характеристики та використовується в певних ситуаціях залежно від конкретних вимог завдань.

1.2.3. Розглянемо методи оптимізації нейронних мереж. Розуміння цих аспектів є важливим для досягнення оптимальних результатів при застосуванні штучного інтелекту в реальних завданнях.

1.2.3.1 Генетичний алгоритм – це ефективний метод оптимізації, який базується на принципах природного відбору та еволюції. У контексті оптимізації нейронних мереж, генетичний алгоритм може бути застосований для пошуку оптимальних параметрів та ваг, що дозволяє досягати кращої продуктивності та точності.

Ключові етапи генетичного алгоритму

1. Початкова популяція:
 - Створення початкової популяції, де кожен індивід (рішення) представляє нейронну мережу з випадково обраними параметрами та вагами.
2. Оцінка пристосованості:
 - Оцінка пристосованості кожного індивіда за допомогою функції втрат на основі його вирішення завдань.
3. Селекція:
 - Обрання індивідів для наступного покоління з урахуванням їхньої пристосованості. Індивіди з вищою пристосованістю мають більше шансів потрапити в нове покоління.
4. Хромосомний перехресний та мутація:
 - Здійснення хромосомного перехресування та мутації для створення нових індивідів. Ці операції сприяють розвитку нових конфігурацій, адаптованих до специфіки завдань.
5. Створення нового покоління:
 - Створення нового покоління на основі обраних індивідів, яке замінює попереднє.
6. Повторення ітерацій:

- Повторення процесу селекції, хромосомного перехрещування, мутації та оцінки пристосованості до досягнення задовільного рівня продуктивності.

Застосування генетичного алгоритму в оптимізації нейронних мереж

Використання генетичного алгоритму в оптимізації нейронних мереж дозволяє автоматизувати процес підбору оптимальних параметрів, забезпечуючи швидке знаходження ефективних конфігурацій для вирішення конкретних завдань. Генетичний алгоритм стає потужним інструментом у розробці та налаштуванні нейронних мереж, підвищуючи їхню точність та продуктивність.

1.2.3.2 Ройові алгоритми є ефективним підходом до оптимізації, який базується на ідеях взаємодії між індивідами у групі (рої). У контексті оптимізації нейронних мереж, ройові алгоритми можуть бути використані для пошуку оптимальних параметрів та ваг.

1. Рой частинок

Метод рою часток, МРЧ (англ. Particle Swarm Optimization, PSO) — метод чисельної оптимізації, для використання якого не потрібно знати точного градієнта оптимізованої функції. МРЧ був доведений Кеннеді, Еберхартом і Шіі, спочатку призначався для імітації соціальної поведінки. Алгоритм був спрощений, і було зауважено, що він придатний для виконання оптимізації. Книга Кеннеді й Еберхарта описує багато філософських аспектів МРЧ і так званого ройового інтелекту. Велике дослідження застосувань МРЧ зроблене Полі[9][10].

МРЧ оптимізує функцію, підтримуючи популяцію можливих розв'язків, називаних частками, і переміщаючи ці частки в просторі розв'язків згідно із простою формулою. Переміщення підпорядковуються принципу найкращого знайденого в цьому просторі положення, що постійно змінюється при знаходженні частками вигідніших положень.

2. Бджолиний алгоритм

Штучний алгоритм бджолої сім'ї (ABC) — алгоритм рою на основі мета-евристичного алгоритму було введено Карабогом у 2005 році.[8] Він імітує поведінку кормових медоносних бджіл. Алгоритм ABC складається з трьох етапів: робочої

бджоли, бджоли-наглядача, та бджоли-розвідника. Бджоли використовують алгоритм локального пошуку в околі рішення, що вибрані на основі детермінованого відбору робочими бджолами і імовірнісного відбору бджолами-наглядачами. Бждола-розвідник виконує відмову від виснажених джерел живлення в кормовому процесі. За цією аналогією рішення, які не корисні більше для пошуку розв'язку відкидаються, і додаються нові рішення (за аналогією з дослідженням нових регіонів у пошуку джерел).

3. Оптимізація мурахиним алгоритмом (Ant Colony Optimization, ACO):

- Мурахиний алгоритм моделює поведінку мурах, які залишають слід феромонів для позначення шляхів до джерела їжі. Інші мурахи обирають шляхи з більшою концентрацією феромонів.

4. Оптимізація розподіленим роєм (Distributed Particle Swarm Optimization, DPSO):

- DPSO розширює ідею рою частинок, дозволяючи частинкам працювати в розподілених системах. Вони обмінюють інформацію і координуються для знаходження оптимальних рішень.

5. Оптимізація ройовим алгоритмом вогню (Firefly Algorithm):

- Цей алгоритм імітує поведінку світлячків у пошуку партнерів. Світлячки рухаються в просторі, приваблюючи інших світлячків своїм світінням. Тим самим, вони намагаються досягти оптимальних рішень.

Ці ройові алгоритми різняться в своїх підходах та використовують різні ідеї з природи для досягнення оптимізації у різноманітних задачах.

1.2.3.3 Градієнтні алгоритми широко використовуються в оптимізації нейронних мереж. Вони базуються на використанні градієнту (похідної) функції втрат відносно параметрів мережі для оновлення їхніх значень. Ось деякі сучасні градієнтні алгоритми:

1. Метод Adam:

- Принцип роботи: Adam комбінує ідеї з інших градієнтних методів, використовуючи експоненціально зважені середні для оцінки градієнту та його

квадрату. Він включає механізми корекції зміщення та адаптивну швидкість навчання для кращої продуктивності.

2. Алгоритм Нестерова (Nesterov Accelerated Gradient, NAG):

- Принцип роботи: NAG використовує ідею передчасного вирахування градієнту перед оновленням параметрів. Це дозволяє враховувати прискорення в русі та уникати непотрібних зміщень.

3. Adagrad:

- Принцип роботи: Adagrad адаптує швидкість навчання для кожного параметра, враховуючи історію квадратів градієнтів. Великі градієнти отримують меншу швидкість навчання, що дозволяє ефективніше навчати модель на параметрах з великим градієнтом.

4. RMSprop (Root Mean Square Propagation):

- Принцип роботи: RMSprop подібний до Adagrad, але використовує експоненціальне згладжування для квадратів градієнтів. Це дозволяє алгоритму адаптуватися до змінюючоїся структури градієнтів.

5. Адамакс:

- Принцип роботи: Адамакс є варіацією Adam, яка використовує не тільки експоненціально згладжені середні, але й безмежну норму (max) для оцінки градієнту. Це робить алгоритм більш стійким до великих градієнтів.

1.3 Класифікація нейронних мереж

Нейронні мережі можна класифікувати за різними критеріями, такими як архітектура, типи шарів, функції активації та завдання, які вони вирішують. Нижче наведено загальну класифікацію нейронних мереж:

За архітектурою:

1. Прямі нейронні мережі (Feedforward Neural Networks, FNN):

- Опис: Це базовий тип нейронної мережі, де дані рухаються лише в одному напрямку - від вхідного шару до вихідного, без циклічних зв'язків.

- Використання: Зазвичай використовується для задач класифікації та регресії.

2. Рекурентні нейронні мережі (Recurrent Neural Networks, RNN):
 - Опис: Ці мережі мають циклічні зв'язки, що дозволяють їм використовувати інформацію з попередніх кроків в часі. Це корисно для роботи з послідовними даними.
 - Використання: Застосовуються для обробки мови, музики та інших послідовних даних.
3. Згорткові нейронні мережі (Convolutional Neural Networks, CNN):
 - Опис: Ці мережі спеціалізуються на обробці зображень та матриць, використовуючи згорткові та пулінгові шари для виявлення локальних особливостей.
 - Використання: Використовуються для задач класифікації та розпізнавання образів.

За типами шарів:

1. Повністю з'єднані шари (Fully Connected Layers):
 - Опис: Всі нейрони в одному шарі пов'язані з усіма нейронами в попередньому та наступному шарах.
 - Використання: Зазвичай входять в склад прямих нейронних мереж.
2. Згорткові шари (Convolutional Layers):
 - Опис: Використовують фільтри для виявлення локальних особливостей в зображеннях чи інших матрицях даних.
 - Використання: Основний компонент згорткових нейронних мереж.
3. Рекурентні шари (Recurrent Layers):
 - Опис: Зберігають стан та можуть використовувати інформацію з попередніх часових кроків.
 - Використання: Використовуються для моделювання послідовних залежностей у часових рядів та послідовних даних.

За завданням:

1. Класифікаційні нейронні мережі:

- **Опис:** Призначені для призначення вхідних об'єктів до конкретних класів чи категорій.
 - **Використання:** Застосовуються у задачах розпізнавання облич, текстової та звукової класифікації.
2. Регресійні нейронні мережі:
- **Опис:** Використовуються для передбачення числових значень на виході.
 - **Використання:** Ефективні для задач, де потрібно передбачити числові результати, наприклад, ціни на нерухомість чи час сервісу.
3. Генеративні нейронні мережі (Generative Neural Networks, GAN):
- **Опис:** Здатні генерувати нові дані, які схожі на навчальні.
 - **Використання:** Використовуються для створення мистецьких зображень, текстів та інших синтезованих даних.

Ця класифікація допомагає визначити різні типи нейронних мереж в залежності від їхнього призначення, структури та використання. Кожен тип мережі має свої переваги та обмеження і підходить для вирішення конкретних завдань.

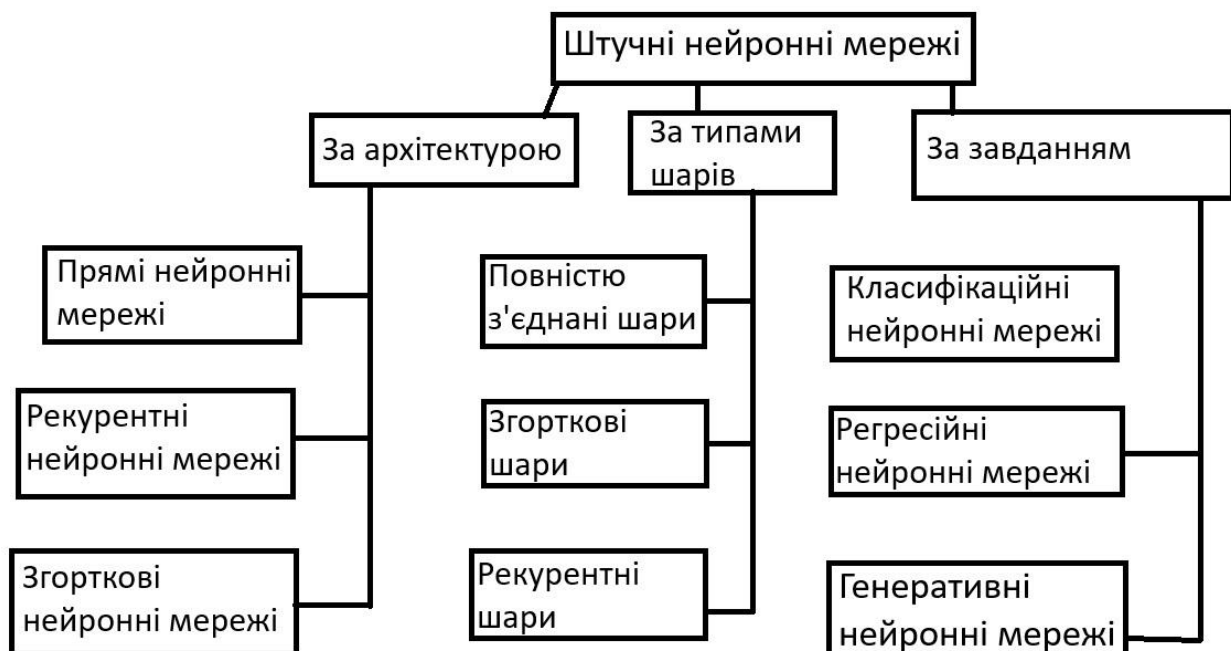


Рис. 1.3.1 Класифікація нейронних мереж.

1.4 Методи машинного навчання.

Машинне навчання (МН) відзначається трансформацією парадигми комп'ютерного програмування, віддавши перевагу системам, які можуть навчатися та пристосовуватися до змін без явного програмування. Ця галузь штучного інтелекту відкриває широкі можливості для автоматизації та оптимізації вирізняючихся за складністю завдань, починаючи від класифікації об'єктів до прогнозування поведінки систем.

Машинне навчання передбачає можливість систем навчатися на основі даних та досвіду, а не через явне програмування. У своїй суті це відображається у здатності алгоритмів покращувати свою продуктивність та точність прогнозів з кожним новим введенням інформації. Важливість цієї галузі стає особливо очевидною в умовах зростання обсягів даних та потреби в ефективному їх аналізі.

Машинне навчання включає різні методи, що розподіляються за завданнями, які вони вирішують: від класифікації та регресії до кластеризації та підкріпленого навчання. В кожному випадку важливо розуміти призначення та обмеження кожного методу для ефективного використання в конкретних умовах.

У цьому розділі ми розглянемо ключові методи машинного навчання, розкриваючи їх принципи роботи, призначення та сфери застосування. Починаючи з основ, ми вивчимо, як ці методи взаємодіють з даними, розвиваючи розуміння їхнього впливу на сучасні технології та наукові дослідження. Ми розглянемо основні методи машинного навчання, поділяючи їх на категорії з вчителем, без вчителя, напівкероване та з підкріпленням.

1. Навчання з вчителем (Supervised Learning)

Навчання з вчителем є одним з основних підходів у машинному навчанні, де модель навчається на основі міткованих прикладів, тобто даних, де відомі правильні відповіді. Основна ідея полягає в тому, щоб алгоритм "учив" себе шляхом порівняння прогнозованих результатів з реальними мітками.

Ключові аспекти:

1. Мітковані дані:

- Для тренування моделі важливо мати доступ до даних, де для кожного вхідного прикладу визначена правильна відповідь.

2. Задачі навчання з вчителем:

- Класифікація: Розділення об'єктів на категорії або класи.
- Регресія: Прогнозування числових значень на основі вхідних даних.

Методи навчання:

1. Логістична регресія:

- Використовується для бінарної та мультикласової класифікації.
- Моделює ймовірність належності об'єкта до конкретного класу.

2. Метод опорних векторів (SVM):

- Визначає оптимальну гіперплощину для розділення класів у просторі.
- Використовується для класифікації та регресії.

3. Нейронні мережі:

- Моделюють нейронну активність для вирішення складних завдань.
- Використовуються для різноманітних задач, включаючи зображення, мову та розпізнавання об'єктів.

Процес навчання:

1. Розділення даних:

- Дані розділяються на тренувальний та тестовий набори для оцінки продуктивності моделі.

2. Визначення функції витрат:

- Функція витрат визначає, наскільки відмінно модель прогнозує відповіді від реальних міток.

3. Оптимізація параметрів:

- Модель намагається оптимізувати свої параметри, коригуючи їх так, щоб мінімізувати функцію витрат.

4. Оцінка продуктивності:

- Модель оцінюється за допомогою тестового набору для визначення її ефективності та готовності до прогнозування нових даних.

Застосування:

1. Медицина:
 - Прогнозування діагнозів на основі клінічних даних.
2. Фінанси:
 - Класифікація та прогнозування ризиків.
3. Рекомендаційні системи:
 - Прогнозування інтересів користувачів на основі їхньої діяльності.

Навчання з вчителем є невід'ємною частиною багатьох сучасних застосувань машинного навчання, дозволяючи системам ефективно та точно працювати з різноманітними видами даних та завдань.

2. Навчання без вчителя (Unsupervised Learning):

Навчання без вчителя (Unsupervised Learning)

Навчання без вчителя - це підхід в машинному навчанні, де модель працює з неміткованими даними, намагаючись знайти в них приховані структури або залежності. У цьому методі немає попередньо визначених правильних відповідей, і завдання полягає в тому, щоб модель самостійно виявляла закономірності чи кластери, які існують у навчальних даних.

Ключові аспекти:

1. Немітковані дані:
 - Дані не містять міток чи класифікацій, і модель самостійно визначає їх структуру.
2. Задачі навчання без вчителя:
 - Кластеризація: Групування схожих об'єктів.
 - Розмірність: Зменшення кількості ознак у даних.
 - Знаходження асоціацій: Виявлення зв'язків між елементами даних.

Методи навчання:

1. Кластеризація (K-Means):
 - Групує об'єкти в кластери на основі їх схожості.
 - Використовується для виявлення природних груп в даних.
2. Метод головних компонент (PCA):
 - Зменшує розмірність даних, зберігаючи при цьому їхню варіативність.
 - Використовується для візуалізації та спрощення аналізу даних.
3. Асоціативні правила (Apriori, Eclat):
 - Виявляє зв'язки між елементами набору даних.
 - Використовується у відділїй та рекомендаційних системах.

Процес навчання:

1. Обробка та нормалізація даних:
 - Підготовка даних для аналізу та використання алгоритмів.
2. Вибір методу:
 - Вибір відповідного методу залежно від завдань, які потрібно вирішити.
3. Виявлення структури:
 - Застосування обраного методу для виявлення структур чи груп у даних.
4. Інтерпретація результатів:
 - Аналіз та інтерпретація отриманих відповідей для отримання інсайтів.

Застосування:

1. Маркетинговий аналіз:
 - Групування клієнтів за їхнім поведінкою та інтересами.
2. Сегментація зображень:
 - Розділення зображення на області зі схожими характеристиками.
3. Зменшення розмірності даних:

- Спрощення аналізу великої кількості ознак.

Навчання без вчителя відіграє важливу роль у ситуаціях, де мітки чи класифікації важко або неможливо отримати. Цей підхід дозволяє отримати інсайти та виявляти закономірності в даних, що може бути корисним у різних галузях від аналізу даних до виявлення нових знань.

3. Напівкероване навчання (Semi-Supervised Learning):

Напівкероване навчання (Semi-Supervised Learning)

Напівкероване навчання - це метод машинного навчання, що поєднує елементи навчання з вчителем та без вчителя. У цьому підході модель тренується на невеликій кількості міткованих даних та великій кількості неміткованих даних. Використання невеликої кількості міткованих прикладів дозволяє зменшити залежність від дорогого етапу міткового тренування.

Ключові аспекти:

1. Мітковані та немітковані дані:
 - Використання обох типів даних для тренування моделі.
2. Задачі напівкерованого навчання:
 - Використовується для задач класифікації, де є обмежена кількість міткованих даних.

Методи навчання:

1. Метод підсилення (Bootstrapped Ensembles):
 - Використовується для комбінування інформації від міткованих та неміткованих даних.
 - Спрямований на поліпшення якості представлення моделі.

Процес навчання:

1. Вибір міткованих даних:
 - Вибір невеликої кількості міткованих даних для тренування.
2. Використання неміткованих даних:
 - Використання великої кількості неміткованих даних для поліпшення універсальності моделі.
3. Тренування з використанням обох типів даних:

- Модель тренується на міткованих та неміткованих даних, змішуючи їх для покращення загальної продуктивності.

Застосування:

1. Обробка обмежених міткованих даних:

- Використовується в ситуаціях, де мітковані дані обмежені або дорогою ціною їхнього отримання.

2. Великомасштабна класифікація:

- Застосовується для класифікації великих наборів даних, де мітковані приклади обмежені.

3. Зображення та відеоаналітика:

- Використовується для вдосконалення процесу розпізнавання об'єктів у великих наборах зображень чи відео.

Напівкероване навчання виступає як компромісний варіант між навчанням з вчителем та без вчителя, дозволяючи покращити продуктивність моделі в умовах, коли мітковані дані обмежені чи важкодоступні.

4. Навчання з підкріпленням (Reinforcement Learning):

Навчання з підкріпленням (Reinforcement Learning)

Навчання з підкріпленням - це підхід у машинному навчанні, в якому модель, називана агентом, навчається взаємодіяти з оточенням, здійснюючи дії та отримуючи відповіді у вигляді винагороди або штрафу. Мета агента полягає в тому, щоб вивчити оптимальні стратегії дій для максимізації нарахованої винагороди в довгостроковій перспективі.

Ключові аспекти:

1. Взаємодія з оточенням:

- Агент діє в середовищі та сприймає його стан.

2. Винагорода та штраф:

- Система нагороди або штрафу визначає, наскільки ефективно агент виконує свої дії.

3. Довгострокова стратегія:

- Мета агента - розвивати стратегії, що максимізують нараховану винагороду в довгостроковій перспективі.

Методи навчання:

1. Q-навчання (Q-Learning):

- Агент вчиться оцінювати якість своїх дій у конкретних станах та вибирати ті, які призводять до найбільшої винагороди.

2. Глибоке підкріплене навчання (Deep Reinforcement Learning):

- Використовує глибокі нейронні мережі для навчання агента з великою кількістю параметрів.

Процес навчання:

1. Визначення стану та дій:

- Визначення простору станів та доступних дій для агента.

2. Визначення функції винагороди:

- Встановлення системи винагород та штрафів для оцінки агента.

3. Тренування агента:

- Агент проводить ітеративні цикли взаємодії з оточенням, намагаючись оптимізувати свою стратегію.

4. Оптимізація стратегії:

- Агент вивчає оптимальні стратегії дій в залежності від стану оточення.

Застосування:

1. Гра в комп'ютерних іграх:

- Навчання агентів грати в ігри та досягати максимальних результатів.

2. Автономне водіння:

- Розвиток систем, які навчають автомобілі приймати рішення на дорозі.

3. Управління роботами:

- Навчання роботів виконувати завдання у різних середовищах.

Навчання з підкріпленням знаходить широке застосування в сферах, де важко визначити жорсткі правила та стратегії, і де агент може взаємодіяти з оточенням для вдосконалення свого поведінкового рівня.

Машинне навчання - це важлива галузь, яка забезпечує розширені можливості аналізу та обробки даних, а також ефективне розв'язання завдань у багатьох сферах життя. Кожен метод має свої унікальні застосування та вимоги, і вибір конкретного методу залежить від характеристик завдання та властивостей даних.

РОЗДІЛ 2

Задача сегментації при обробці зображень.

2.1. Огляд робіт.

Сегментація зображення — це один із основних етапів обробки зображення, під час якого будь-яке зображення поділяється на кілька сегментів. Кожен сегмент представлятиме певну інформацію для користувача у вигляді кольору, інтенсивності або текстури. Тому важливо виділити межі будь-якого зображення у вигляді його сегментів [4]. Цей процес сегментації призначає одне значення для кожного пікселя зображення, щоб полегшити розрізнення різних областей будь-якого зображення. Ця диференціація між різними сегментами зображення здійснюється на основі трьох властивостей зображення, тобто кольору, інтенсивності та текстури цього зображення. Тому вибір будь-якої техніки сегментації зображення здійснюється після спостереження проблемної області [5].

Важливість сегментації зображення не можна нехтувати, оскільки вона використовується майже в усіх галузях науки, наприклад, для видалення шуму із зображення, медичних зображень [6]-[10], супутникових зображень, машинного зору, комп'ютерного зору, біометрії, військовий, пошук зображення [11]-[12], вилучення ознак і розпізнавання об'єктів із заданого зображення. [13]-[15]

Помічено, що не існує ідеального методу сегментації зображення, оскільки кожне зображення має свій власний тип. Також дуже складним завданням є пошук техніки сегментації для конкретного типу зображення. Оскільки метод, застосований до одного зображення, може не бути успішним для іншого типу зображень, тому методи сегментації були розділені на три типи, тобто методи сегментації, засновані на класичному методі, методи ШІ та гібридні методи [16]. Деякі з найвідоміших методологій сегментації зображень, включаючи сегментацію на основі країв, сегментацію на основі нечіткої теорії, сегментацію на основі часткових диференціальних рівнянь (PDE), сегментацію на основі штучної нейронної мережі (ANM), сегментацію зображень на основі порогів і сегментацію зображень на основі

регіонів, виділено на рис. 2.1.1. На рисунку 2.1.1 представлені важливі та відомі методи сегментації зображення, які використовуються для сегментації зображення.



Рис. 2.1.1 Різні техніки сегментації зображення.

ТЕХНІКА СЕГМЕНТАЦІЇ ЗОБРАЖЕНЬ

Багато методів сегментації зображень було розроблено дослідниками та науковцями, деякі з найбільш важливих і широко використовуваних методів сегментації зображень показано на рис. 2.1.1. Останні дослідження методів сегментації зображень, виділені на рис. 2.1.1, обговорюються та оцінюються нижче.

А. Сегментація зображення на основі порогового значення

Порогове значення гистограми використовується для сегментації даного зображення; існують певні методи попередньої та постобробки, необхідні для порогової сегментації [17]. Основними пороговими методами, запропонованими різними дослідниками, є метод середнього значення, метод Р-плитки, метод,

залежний від гістограми, метод максимізації країв і візуальний метод. У цьому розділі обговорюються кілька нових підходів за останні п'ять років щодо сегментації зображення на основі порогових значень.

Салем Салех Аль-Амрі [18] застосував методику Mean, Pile, HDT і EMT на трьох супутникових зображеннях, щоб вибрати найкраще сегментоване зображення з усіх вищевказаних методів. Експерименти та порівняльний аналіз методів показали, що HDT (Метод, залежний від гістограми) та EMT (Метод максимізації країв) є найкращими методами порогового визначення, які перевершують усі інші методи визначення порогового значення.

Кайпінг Вел [19] виявив, що поточні методи сегментації зображень займають багато часу та вимагають великих обчислювальних витрат для виконання сегментації зображення. Це велика проблема для програм реального часу. Вони запропонували новий метод сегментації на основі порогових значень з використанням оптимізації роєм частинок (PSO) і 2-d алгоритму Оцу (TOPSO). Алгоритм TOPSO використовував техніку PSO для пошуку оптимального порогу для процесу сегментації. Вони реалізують запропонований гібридний метод на Matlab 7.0. Результати показали, що алгоритм TOPSO займає в 25 разів менше часу порівняно з традиційним алгоритмом Оцу. Це добре для додатків у реальному часі.

В. Сегментація зображення на основі регіону

Сегментація на основі регіонів проста порівняно з іншими методами, а також стійка до перешкод. Він ділить зображення на різні області на основі попередньо визначених критеріїв, наприклад кольору, інтенсивності або об'єкта. Методи сегментації на основі регіону поділяються на три основні категорії, тобто зростання регіону, поділ регіону та злиття регіону [20]. У цьому розділі обговорюються кілька нових підходів до сегментації зображень на основі регіону за останні п'ять років.

Карої [21] запропонував новий метод сегментації зображення без нагляду з використанням методів набору рівнів і статистики текстур. Вони стверджують, що їхній метод відрізняється від інших методів, оскільки він не передбачає незалежної змінної та не обмежується елементами сірого першого порядку. Реалізація включає етап вибору функції для повторного налаштування ваги кожної функції для

отримання сегментації. На етапі експерименту гістограма відгуку фільтра використовується для обчислення кількості розподілів; Вейвлет Гаара використовується для обчислення енергії вейвлета зображення кожної смуги. PDE використовується для повторної ініціалізації наборів рівнів. Результати показали правильну сегментацію для зображення зебри.

Йонг-мей Жоу [22] представив нову техніку сегментації зображення на основі регіону за допомогою алгоритму кластеризації середнього зсуву. По-перше, їхній метод виділяє особливості кольору, текстури та розташування кожного пікселя зображення, по-друге, створює кластери на основі цих ознак, використовуючи підхід кластеризації середнього зсуву, позначає кожну область i , нарешті, створює сегменти зображення на основі цих міток. Вони використовували Matlab 7.0 для реалізації свого алгоритму. Експеримент показує, що їхній метод дає кращі результати щодо швидкості та сегментації.

Чівахір Цігла [23] представив новий теоретико-графовий метод сегментації кольорового зображення та намагається вдосконалити нормалізований метод сегментації розрізаного зображення. Вони використовували зображення зі зваженим неорієнтованим графом, тоді як вузли представляють регіони, а ваги між вузлами представляють відповідність інтенсивності сусідніх регіонів. Їх модифікований нормалізований метод вирішення вирішив проблему надмірної сегментації, у якій для зображення створюються додаткові області. Експерименти проводяться на зображеннях корови, мозаїки та NC-зображенні з різною роздільною здатністю та результати порівнюються з алгоритмом NCIS на основі критеріїв MSE. Результати показали, що запропонований метод покращує алгоритм NCIS.

С. Сегментація зображення на основі країв

Виявлення країв є основним етапом процесу сегментації зображення [24]. Він ділить зображення на об'єкт і його фон. Виявлення країв ділить зображення, спостерігаючи за зміною інтенсивності або пікселів зображення. Сіра гістограма та градієнт є двома основними методами виявлення країв для сегментації зображення [25]. Декілька операторів використовуються методом визначення краю, наприклад, класичні детектори краю, перехід через нуль, лапласіан Гуасіана (LoG) [26], кольорові

детектори краю тощо [27]. У цьому розділі обговорюються кілька нових підходів до сегментації зображення на основі виявлення країв за останні десять років.

Ю Ксяохан [28] запропонував нову техніку сегментації зображення, засновану на методах вирощування області та визначення країв. Їх гібридний метод допомагає процесу сегментації уникнути помилок, коли обидва методи використовуються окремо. Збільшення регіону використовується для знаходження крайових пікселів на зображенні, тоді як похідна 2-го порядку використовується для виявлення країв. Експерименти проводяться на даних зображень 3D МРТ. Техніка Гауса використовується для згладжування після виявлення краю. Результати показали, що їхня техніка є кращою для збереження більшої кількості периферійної інформації.

Весолковськ [29]-[30] використовували випадкові поля Маркова для гібридної сегментації кольорових зображень на основі країв і областей. По-перше, лінійний процес реалізується за допомогою алгоритму визначення краю. Векторна міра кута використовується як міра відстані між пікселями для визначення країв. Основна проблема з їхньою технологією полягає в тому, що це модель піксельних сусідів і має ті самі недоліки, що й метод зростання області. Для оцінки моделі MRF використовується методика оцінки параметрів.

Юінг-Танг Ксяо [31] запропонував нову техніку сегментації зображення шляхом поєднання морфологічного оператора з технікою вирощування області. Спочатку вони використали морфологічну замкнуту операцію для покращення зображення, а потім виконали краєвиділення за допомогою детектора країв залишків розширення. Після цього вони розгортають насіння, що ростуть, і виконують процес вирощування області для сегментації магів, після чого на зображеннях виконується об'єднання областей і визначення країв. Вони проводять експерименти з настільного тенісу, дівчини та зображення МРТ. Метод граничних умов змійки [32] використовується для отримання кращих результатів виявлення країв. Усі експерименти проводяться у Visual C++.

Амжад Заїм[33] виявив, що сегментація меж передміхурової залози на ультразвукових зображеннях є складним завданням для хірургічних процедур. Вони запропонували нову техніку сегментації на основі країв для ультразвукового

зображення простати. Фазова симетрія використовується для визначення краю на ультразвукових зображеннях. Для зменшення шуму використовується медіанний фільтр. Вилучення країв і зв'язування країв використовуються для створення остаточного сегментованого зображення на основі країв. Головна перевага полягає в тому, що їх метод не вимагає втручання людини. Результати контуру, отриманого за їхнім методом, порівнюють із сегментованими вручну контурами, і встановлено точність 87%.

D. Сегментація зображення на основі нечіткої теорії

Теорія нечітких множин використовується для аналізу зображень і надання точної інформації з будь-якого зображення. Функцію фазифікації також можна використовувати для видалення шуму із зображення [34]. Зображення в градаціях сірого можна легко перетворити на нечітке за допомогою функції фазифікації. Різні морфологічні операції можна комбінувати з нечітким методом для отримання кращих результатів [35]. Нечіткі k-середні та нечіткі C-середні (FCM) є широко використовуваними методами обробки зображень [36]. У цьому розділі представлено декілька нових підходів до сегментації зображення з використанням теорії нечіткості.

Гоур Чандра Кармакар [37] представив нову техніку сегментації зображення на основі нечітких правил, яка може інтегрувати просторове співвідношення пікселів. Використовуються три типи функцій приналежності, тобто функція приналежності для розподілу пікселів регіону, для вимірювання близькості регіону та визначення просторового відношення між пікселями. У їх техніці немає необхідності визначати параметри, як у алгоритмі FCM. Нечіткі правила використовують вищезазначені три функції належності та нечітку структуру правила IF-THEN для виконання сегментації зображення. FCM і запропонована методика реалізована в Matlab 5.3.1 на рентгенівському зображенні та зображенні голосового тракту людини. Результати показали, що GFRIS перевершує FCM і точно ізолює об'єкт від фону.

Амол С. Педнекар [38] запропонував нову техніку сегментації зображення на основі нечіткої зв'язності з використанням динамічних ваг. Автор виявив, що традиційні схеми сегментації не можуть вирішити проблеми нечітких медичних зображень. Вони запроваджують алгоритм DuW, який динамічно регулює лінійні

ваги в нечіткій зв'язності. Початковий алгоритм DuW успішно застосовується до зображень різних модальностей, тоді як множинне початкове значення застосовується до інфрачервоної сегментації обличчя. Виявлено, що алгоритм сегментації зображення DuW дає на 99% більшу точність у порівнянні з іншими методами.

Ліу Яджу [39] запропонував новий алгоритм сегментації нечіткого кольорового зображення, заснований на розбіжності ознак і нечіткій несхожості. Їх алгоритм стверджує, що покращує якість сегментації. Їхній алгоритм виділяє підзображення, що містять eagen-вектор, використовуючи метод вододілу. По-перше, кольорове зображення перетворюється на зображення рівня сірого, гістограма створюється на другому кроці, кластер створюється на наступному кроці, FCM застосовується до кожного кластера, потім вони застосовують ерозію, розширення та розширення області на результуючому зображенні. Після цього в кінці створюється зображення сегментованої області. Зображення зроблено зі складним фоном, тобто фотографічними зображеннями. Результати показали, що нечіткі підходи дають кращі результати.

Е. Сегментація зображення на основі ШНМ

У штучній нейронній мережі кожен нейрон відповідає пікселю зображення. Зображення відображається в нейронній мережі. Зображення у вигляді нейронної мережі навчається за допомогою навчальних зразків, а потім виявляється зв'язок між нейронами, тобто пікселями. Потім нові зображення сегментуються з навченого зображення [40]. Деякі з нейронних мереж, які найчастіше використовуються для сегментації зображень, це Hopfield, BPNN, FFNN, MLFF, MLP, SOM і PCNN. Сегментація зображення за допомогою нейронної мережі виконується в два етапи, тобто класифікація пікселів і визначення країв [41]. У цьому розділі обговорюються кілька нових підходів ШНМ, які використовуються для сегментації зображень за останні п'ять років.

Ксуеджі Жанг[42] запропонував новий підхід до сегментації кольорових зображень на основі швидкого навчання штучної нейронної мережі (FLANN) для простору кластерів R-G-B-S-V (тобто RGB і HSV). На першому кроці шум

видаляється за допомогою фільтра усереднення 3×3 , щоб зменшити розбіжність у розподілі кольорів. На другому кроці пікселі перетворюються на простір RGBSV за допомогою перетворень HSV. Кластеризація FLANN виконується для створення кластерного результату зображення. Далі відокремлюються пікселі однакового кольору. Кожному сегменту зображення присвоюється номер сегмента. Спостерігається вплив толерантності та розміру сусідства. Результати показали, що запропонований алгоритм створює ідеальні сегменти для кольорів зображення.

Фархад Мохамад Каземі [43] запропонував швидке навчання нечіткої нейронної мережі Хопфілда [44] на основі C-середніх для того, щоб застосувати її до сегментації зображення. Цільова функція використовується на основі 2-f Fuzzy HNN. Ця цільова функція знайшла середню відстань між пікселями зображення та центроїдами кластера. За словами автора, Fuzzy HNN забезпечує кращу сегментацію порівняно з іншими методами. Спочатку вони створюють кластери із заданих даних, потім виконують нормалізацію, тобто зображення рівня сірого, обчислюють центроїди, потім обчислюють відстані, знаходять нові центроїди та обчислюють нове значення функції належності за допомогою нечітких C-середніх [45]. Результати показали, що FHNN забезпечує вищу швидкість порівняно з іншими техніками ANN.

F. Сегментація зображення на основі PDE

PDE (диференціальні рівняння в частинних похідних) або моделі PDE широко використовуються в обробці зображень і, зокрема, у сегментації зображень. Вони використовують модель активного контуру для цілей сегментації. Модель Active Contour або Snakes перетворюють проблему сегментації в PDE. Деякі відомі методи PDE, що використовуються для сегментації зображення, це Snakes, Level-Set і метод Mumford shah [46]. У цьому розділі обговорюються кілька нових підходів до сегментації зображень на основі PDE.

Глорія Буено [47] представляє новий метод сегментації анатомічної структури на медичних зображеннях. Для сегментації зображень використовуються адаптивні PDE-моделі, тобто нечітка PDE-контурна модель і PDE-геометрична контурна модель з класифікацією Fuzzy C Means. Адаптивні моделі PDE допомогли знайти область інтересу. Тривимірне МРТ-зображення головного мозку використовується як набір

даних. Модель Fuzzy PDE сегментує МРТ-зображення мозку за допомогою підходу Fuzzy Clustering. Модель перевершила модель «Snakes» і зменшила деякі недоліки моделі Snakes.

Схеми виділення ознак у [48]-[49] здатні обробляти геометричну складність, швидкість зміни та орієнтацію зображення. Також представлена нова схема сегментації на основі PDE, яка підвищує критерії контрастності інформації про текстуру. PDE використовуються для моделювання схеми сегментації. Метод вододілу [50] розширено за допомогою моделей PDE. Вони порівнюють запропоновану схему з методом сегментації вододілу, і виявлено, що поєднання текстурної інформації та моделювання за допомогою PDE призводить до високоякісного процесу сегментації зображення та перевершує алгоритм сегментації вододілу.

2.2 Критерії для сегментації при обробці зображень.

2.2.1 Функція втрат

Окрім поліпшення швидкості та точності сегментації за допомогою розробки основної мережі та блоку функцій, розробка нових функцій втрат також призвела до покращень у точності сегментації під час подальшого виведення результатів. Таким чином, було зроблено значну кількість робіт щодо розробки відповідних функцій втрат для завдань сегментації медичних зображень.

2.2.2 Втрата перехресної ентропії

Для завдань сегментації зображень перехресна ентропія є однією з найпопулярніших функцій втрат. Ця функція порівнює піксельно передбачений вектор категорії з реальним вектором результату сегментації. У випадку бінарної сегментації, нехай $P(Y = 1) = p$, і $P(Y = 0) = 1 - p$, тоді передбачення задається сигмоїдною функцією, де $P(\hat{Y} = 1) = 1/(1 + e^{-x}) = \hat{p}$, і $P(\hat{Y} = 0) = 1 - 1/(1 + e^{-x}) = 1 - \hat{p}$, де x - вихід мережі нейронів. Функція втрати перехресної ентропії визначається як

$$CE(p, \hat{p}) = -(p \log(\hat{p}) + (1-p) \log(1-\hat{p})) \quad (2.1)$$

2.2.3 Взважена функція втрати перехресної ентропії

Функція втрати перехресної ентропії враховує кожний піксель зображень однаково і, таким чином, виводить середнє значення, що ігнорує дисбаланс класів і

призводить до проблеми, коли функція втрати залежить від класу з максимальною кількістю пікселів. Таким чином, функція втрати перехресної ентропії часто показує низьку ефективність для сегментації малих цільових об'єктів.

Для вирішення проблеми дисбалансу класів, Лонг та ін. [32] запропонували взважену функцію втрати перехресної ентропії (WCE) для компенсації дисбалансу класів. У випадку бінарної сегментації взважена функція втрати перехресної ентропії визначається як

$$WCE(p, p^{\wedge}) = -(\beta p \log(p^{\wedge}) + (1-p) \log(1-p^{\wedge})) \quad (2.2)$$

де β використовується для налаштування пропорції позитивних та негативних зразків і є емпіричним значенням. Якщо $\beta > 1$, кількість помилкових від'ємних результатів буде зменшено; навпаки, кількість помилкових позитивних результатів буде зменшено, коли $\beta < 1$. Функція перехресної ентропії є спеціальним випадком взваженої функції втрати перехресної ентропії, коли $\beta = 1$. Для одночасного налаштування ваги позитивних і негативних зразків можна використовувати збалансовану функцію втрати перехресної ентропії (BCE), яка визначається як

$$BCE(p, p^{\wedge}) = -(\beta p \log(p^{\wedge}) + (1-\beta)(1-p) \log(1-p^{\wedge})) \quad (2.3)$$

У [7], Ронебергер та ін. запропонували мережу U-Net, в якій функцію втрати перехресної ентропії поліпшено додаванням функції відстані. Покращена функція втрати може покращити здатність моделей до вивчення відстані між класами. Функція відстані визначається

$$D(x) = \omega_0 - e^{\frac{-(d_1(x) + d_2(x))^2}{2\sigma^2}} \quad (2.4)$$

де $d_1(x)$ і $d_2(x)$ позначають відстань між пікселем x і межами перших двох найближчих клітинок. Отже, кінцева функція втрат визначається як

$$L = BCE(p, p^{\wedge}) + D(x) \quad (2.5)$$

2.2.4 Втрати Dice

Dice є популярною метрикою для оцінки сегментації медичних зображень. Ця метрика суттєво вимірює ступінь перекриття між результатом сегментації та відповідною дійсною сегментацією. Значення Dice знаходиться в діапазоні від 0 до 1. "1" означає, що результат сегментації повністю перекривається з реальним результатом сегментації. Формула обчислення виглядає наступним чином:

$$Dice(A, B) = \frac{2 * |A \cap B|}{A + B} * 100\% \quad (2.6)$$

де A - це передбачений результат сегментації, а B - реальний результат сегментації.

Для сегментації об'ємних медичних даних Мілетарі та ін. [35] запропонували V-Net, яка використовує втрату Dice.

$$DL(p, \hat{p}) = 1 - \frac{2 \langle p, \hat{p} \rangle}{\|p\|_1 + \|\hat{p}\|_2} \quad (2.7)$$

де $\langle p, \hat{p} \rangle$ представляє скалярний добуток значень кожного каналу між матрицею реальних даних та передбаченими результатами.

Слід відзначити, що втрата Dice підходить для нерівномірних вибірок. Однак її використання легко впливає на зворотне поширення та може призводити до труднощів у тренуванні. Крім того, втрата Dice має низьку стійкість для різних моделей, таких як середня відстань до поверхні або Гаусдорфівська відстань до поверхні, через неправдоподібні значення градієнта. Наприклад, градієнт функції softmax може бути спрощений до $(p - t)$, де t - це цільове значення, а p - передбачене значення, але значення втрати Dice складає $2t^2 / (p + t)^2$. Якщо значення p і t дуже малі, то значення градієнта змінюватиметься різко, що призведе до труднощів у тренуванні.

2.2.5 Втрата Tversky

Салехі та ін. [93] запропонували втрату Tversky (TL), яка є регуляризованою версією втрати Dice для контролю внеску як помилкових позитивних, так і помилкових від'ємних значень у функцію втрати. Втрата Tversky визначається як

$$TL(p, \hat{p}) = \frac{p, \hat{p}}{p, \hat{p} + \beta(1 - p, \hat{p}) + (1 - \beta)(p, 1 - \hat{p})} \quad (2.8)$$

де $p \in \{0, 1\}$, $0 \leq \hat{p} \leq 1$. p і \hat{p} відповідають реальним та передбаченим результатам сегментації відповідно. Втрата Tversky еквівалентна (7), якщо $\beta = 0,5$.

2.2.6 Узагальнений програш кубиків

Хоча втрата Dice може певною мірою вирішити проблему класового дисбалансу, вона не працює для серйозного класового дисбалансу. Наприклад, маленькі цілі страждають від помилок передбачення деяких пікселів, що легко спричиняє значні зміни для значень Dice. Sudre та ін. [94] запропонував узагальнений програш у кості (GDL), GDL визначається як

$$GDL(p, \hat{p}) = 1 - \frac{1}{m} \frac{2 \sum_{j=1}^m \omega_j \sum_{i=1}^n p_{ij} \hat{p}_{ij}}{\sum_{j=1}^m \omega_j \sum_{i=1}^n (p_{ij} + \hat{p}_{ij})} \quad (2.9)$$

де вага $\omega = [\omega_1, \omega_2, \dots, \omega_m]$ призначається кожному класу, а $\omega_j = 1/(\sum_{i=1}^n p_{ij})^2$. GDL перевершує втрати Dice, оскільки різні області мають однакові внески у втрати, а GDL більш стабільний і міцний під час процесу навчання.

2.2.7 Втрати на межі

Щоб вирішити проблему дисбалансу класів, Кервадець та ін. [95] запропонували нову втрату на межі, яку використовують для сегментації уражень мозку. Ця функція втрат ставить за мету мінімізацію відстані між сегментованими межами та позначеними межами. Автори провели експерименти на двох несбалансованих наборах даних із мітками. Результати показують, що комбінація втрати Dice та втрати на межі перевершує кожен окремо. Складена втрата визначається як

$$L = \alpha L_{GD}(\theta) + (1 - \alpha) L_B(\theta) \quad (2.10)$$

де перший частину є регуляризованою втратою Dice, яка визначається як

$$L_{GD}(\theta) = 1 - \frac{2(\omega_G \sum_{p \in \Omega} g(p) s_{\theta}(p) + \omega_B \sum_{p \in \Omega} (1 - g(p))(1 - s_{\theta}(p)))}{((\omega_G \sum_{p \in \Omega} [g(p) + s_{\theta}(p)]) + (\omega_B \sum_{p \in \Omega} (2 - g(p) - s_{\theta}(p))))} \quad (2.11)$$

а друга частина – це гранична втрата, яка визначається як

$$L_B(\theta) = \sum_{p \in \Omega} G(p) s_{\theta}(p), \quad (2.12)$$

де, якщо $\sum_{p \in \Omega} G(p) = 1$, в іншому випадку $\sum_{p \in \Omega} G(p) = \sum_{p \in \Omega} G(p)^2$. Крім того, використовується $\sum_{p \in \Omega} g(p) f(s_{\theta}(p))$ для переднього плану та $\sum_{p \in \Omega} (1 - g(p))(1 - f(s_{\theta}(p)))$ для тла. Вага $L_{GD}(\theta)$ це $\omega_G = 1/(\sum_{p \in \Omega} g(p))^2$, а вага $L_B(\theta)$ - $\omega_B = 1/(\sum_{p \in \Omega} (1 - g(p)))^2$.

Символ Ω представляє набір пікселів в усьому просторовому домені.

2.2.8 Експоненційно-логарифмічна втрата

У (9) втрата з вагою Dice фактично полягає в тому, що отримане значення Dice ділиться на суму кожної мітки, що досягає балансу для об'єктів різних розмірів. Таким чином, поєднуючи фокус-втрата [96] і втрату Dice, Вонг та ін [97] запропонували експоненційно-логарифмічну втрату (EXP loss), яку використовують для сегментації мозку для вирішення проблеми серйозного дисбалансу класів. З введенням

експоненційної форми нелінійність функцій втрат можна подальше контролювати для покращення точності сегментації. Функція втрат EXP визначається як

$$L_{EXP} = \omega_{dice} \times L_{dice} + \omega_{cross} \times L_{cross}, \quad (2.13)$$

де дві нові параметри ваги позначаються як ω_{dice} та ω_{cross} відповідно. L_{dice} - це експоненційна логарифмічна втрата Dice, а L_{cross} - втрата перехресної ентропії.

$$L_{dice} = E[(-\ln(Dice_i))^{\gamma_{Dice}}], \quad (2.14)$$

$$L_{cross} = E[\omega_l (-\ln(p_l(x)))^{\gamma_{cross}}], \quad (2.15)$$

$$Dice_i = \frac{2(\sum_x \sigma_{il}(x)p_i(x)) + \epsilon}{\sum_x (\sigma_{il}(x) + p_i(x)) + \epsilon} \quad (2.16)$$

$$\omega_l = \left(\frac{\sum_k f_k}{f_l}\right)^{0.5} \quad (2.17)$$

де x - це позначення пікселя, i - це мітка, а l - значення істинності на позначенні x . $p_i(x)$ - це ймовірнісне значення, виведене з softmax.

У (17), f_k - це частота появи мітки k , цей параметр може зменшити вплив більш часто зустрічаючихся міток. Як γ_{Dice} , так і γ_{cross} використовуються для підвищення нелінійності функції втрат.

2.2.9 Удосконалення функції втрат

Для сегментації медичних зображень вдосконалення функцій втрат в основному спрямоване на проблему сегментації малих об'єктів на великому тлі (проблема дисбалансу класів). Чен та ін. [98] запропонували нову функцію втрат застосувавши традиційну енергетичну мінімізацію активного контура до згорткових нейромереж, Лі та ін. [99] запропонували новий термін регуляризації для поліпшення функції втрати перехресної ентропії, і Карімі та ін. [100] запропонували функцію втрат на основі відстані Гаусдорфа (HD). Крім того, існують ще багато робіт [101] [102], які намагаються вирішити цю проблему, додаючи штрафи до функцій втрат або змінюючи стратегію оптимізації відповідно до конкретних завдань. У багатьох задачах сегментації медичних зображень часто є лише одна чи дві цілі на зображенні, і відсоток пікселів для цілей іноді невеликий, що ускладнює тренування мережі. Отже, для поліпшення тренування мережі та точності сегментації, зручніше зосередитися на менших цілях, змінюючи функції втрат, ніж змінювати структуру мережі. Однак розробка функцій втрат є завданням, що сильно залежить від

конкретного завдання, тому потрібно ретельно аналізувати вимоги задачі і вибирати розумні та доступні функції втрат.

2.2.10 Поглиблений нагляд

Загалом, збільшення глибини мережі може покращити репрезентацію ознак мережі до певної міри, але водночас воно викликає нові проблеми, такі як зникаючий градієнт та вибух градієнту. З метою ефективного тренування глибоких мереж Лі та ін. [72] запропонували глибокі мережі з навчанням (DSNs), додаючи кілька допоміжних класифікаторів до деяких шарів нейромережі. Доу та ін. [103] запропонували 3D DSN для сегментації серця та печінки, яка включає механізм глибокого моніторингу в 3D повністю згортковану мережу для вивчення та інференсу обсягу, що дозволяє уникнути зайвих обчислень і зменшити ризик перенавчання в умовах обмеженого обсягу навчальних даних. Також Доу та ін. [104] представили метод сегментації кортикальної пластини мозку на зображеннях МРТ плода, використовуючи архітектуру повністю згорткованої нейромережі з глибоким наглядом та залишковим з'єднанням, і отримали високу точність сегментації кортикальної пластини на зображеннях МРТ мозку. Фактично, глибокий нагляд не тільки обмежує дискримінацію та стійкість вивчених ознак на всіх етапах, але й поліпшує ефективність тренування мережі.

РОЗДІЛ 3

Огляд методів сегментації для обробки зображень

Сегментацію зображень можна визначити як важливий етап цифрової обробки зображень. Крім того, вона використовується в різноманітних застосуваннях, таких як спільна сегментація об'єктів, завдання визначення, медичне зображення, пошук зображень на основі вмісту, виявлення об'єктів, машинне бачення та відеоспостереження. Було створено багато підходів до сегментації зображень. Крім того, основною метою сегментації є полегшення та зміна представлення зображення на щось, що є більш важливим і простим для аналізу. Підходи до сегментації зображень включають розділення зображень на кілька частин на основі характеристик зображення, таких як текстура, колір, інтенсивність пікселів та інше. Щодо представленого дослідження, розглянуті і обговорені багато підходів до сегментації зображень. Техніки сегментації можна класифікувати на шість класів: По-перше, техніки сегментації порогу, такі як глобальна сегментація порогу (ітеративна сегментація порогу, порог мінімальної помилки, метод Оцу, оптимальна сегментація порогу, аналіз вогкої гистограми та сегментація порогу на основі ентропії), локальна сегментація порогу (підхід Савола, підхід Т.Р. Сінгха, підхід Ніблака, підхід Бернсена, метод Брукстайна і Яновіца та локальна адаптивна автоматична бінаризація) і динамічна сегментація. По-друге, техніки сегментації на основі країв, такі як техніка сірого гистограми, підхід на основі градієнту (лапласіан гауссіана, підхід диференційного коефіцієнта, підхід Кенні, підхід Превітта, підхід Робертса та підхід Собеля). В третьому, техніки сегментації на основі регіонів, включаючи техніки росту регіону (зростання регіону за насічкою (SRG), статистичне зростання регіону, незасіяне зростання регіону (UsRG)), а також техніки злиття та розділення регіонів. В четвертому, техніки кластеризації, включаючи м'яку кластеризацію (розмитий к-середній кластеризації (FCM)) та жорстку кластеризацію (к-середній кластеризації). По-п'яте, техніки глибокого нейронного мережевого навчання, такі як згортокова нейронна мережа, рекурентні нейронні мережі (RNN), моделі кодер-декодер та автоенкодерів, а також метод опорних векторів. Нарешті, гібридні техніки,

такі як еволюційні підходи, нечітка логіка та розумні інтелектуальні методи (PSO та ABC техніки), розглядаються з врахуванням переваг та недоліків кожного методу.

Одним з найвикликовіших і важливих процесів, пов'язаних з обробкою зображень, є сегментація зображення. Підходи до сегментації зображень використовуються для розділення зображень на значущі частини. У цифровій обробці зображень основні застосування включають машинне бачення, пошук зображень на основі вмісту, виявлення об'єктів, медичне зображення, відеоспостереження, завдання визначення та спільна сегментація об'єктів. У процесі сегментації основною метою є отримання більше інформації з області інтересу зображення і дозвіл на анотацію об'єктної сцени [1]. Сегментація зображень є складною в обробці зображень, вона є активною областю досліджень, яка дозволяє виділення об'єктів на зображенні. Сегментація розглядається як ненавчальне навчання. Техніки сегментації зображень можна класифікувати на три типи на основі властивостей зображення: підходи до сегментації на основі країв, підходи до сегментації на основі регіонів і порогові підходи до сегментації [2], а також підходи глибокого нейронного мережевого навчання, гібридні підходи і підходи кластеризації.

Представлене дослідження охоплює останню літературу з сегментації зображень та обговорює різні методи сегментації, запропоновані до 2020 року. Ми розглядаємо шість підходів, які були використані для сегментації зображень. Найважливіші техніки, використовувані для сегментації зображень і представлені в цій статті, включають:

По-перше, техніки сегментації порогу, такі як глобальна сегментація порогу (ітеративна сегментація порогу, порог мінімальної помилки, метод Оцу, оптимальна сегментація порогу, аналіз вогкої гістограми та сегментація порогу на основі ентропії), локальна сегментація порогу (підхід Савола, підхід Т.Р. Сінгха, підхід Ніблака, підхід Бернсена, метод Брукстайна і Яновіца та локальна адаптивна автоматична бінаризація) і динамічна сегментація.

По-друге, техніки сегментації на основі країв, такі як техніка сірого гістограми, підхід на основі градієнту (лапласіан гауссіана, підхід диференційного коефіцієнта, підхід Кенні, підхід Превітта, підхід Робертса та підхід Собеля).

В третьому, техніки сегментації на основі регіонів, включаючи техніки зростання регіону (зростання регіону за насичкою (SRG), статистичне зростання регіону, незасіяне зростання регіону (UsRG)), а також техніки злиття та розділення регіонів.

В четвертому, підходи кластеризації, включаючи м'яку кластеризацію (розмитий к-середній кластеризації (FCM)) та жорстку кластеризацію (к-середній кластеризації).

По-п'яте, техніки глибокого нейронного мережевого навчання, такі як згортова нейронна мережа, рекурентні нейронні мережі (RNNs), моделі кодер-декодер та автоенкодер, а також метод опорних векторів.

Нарешті, гібридні техніки, такі як еволюційні підходи, нечетка логіка та розумні інтелектуальні методи (PSO та ABC техніки).

3.1 КЛАСИФІКАЦІЯ ТЕХНІК СЕГМЕНТАЦІЇ ЗОБРАЖЕНЬ

У літературі існують багато використовуваних методів, кожен з них відрізняється від іншого за використовуваним підходом до сегментації. КЛАСИФІКАЦІЯ методів сегментації зображень зображена на рис.3.1.1. [3].

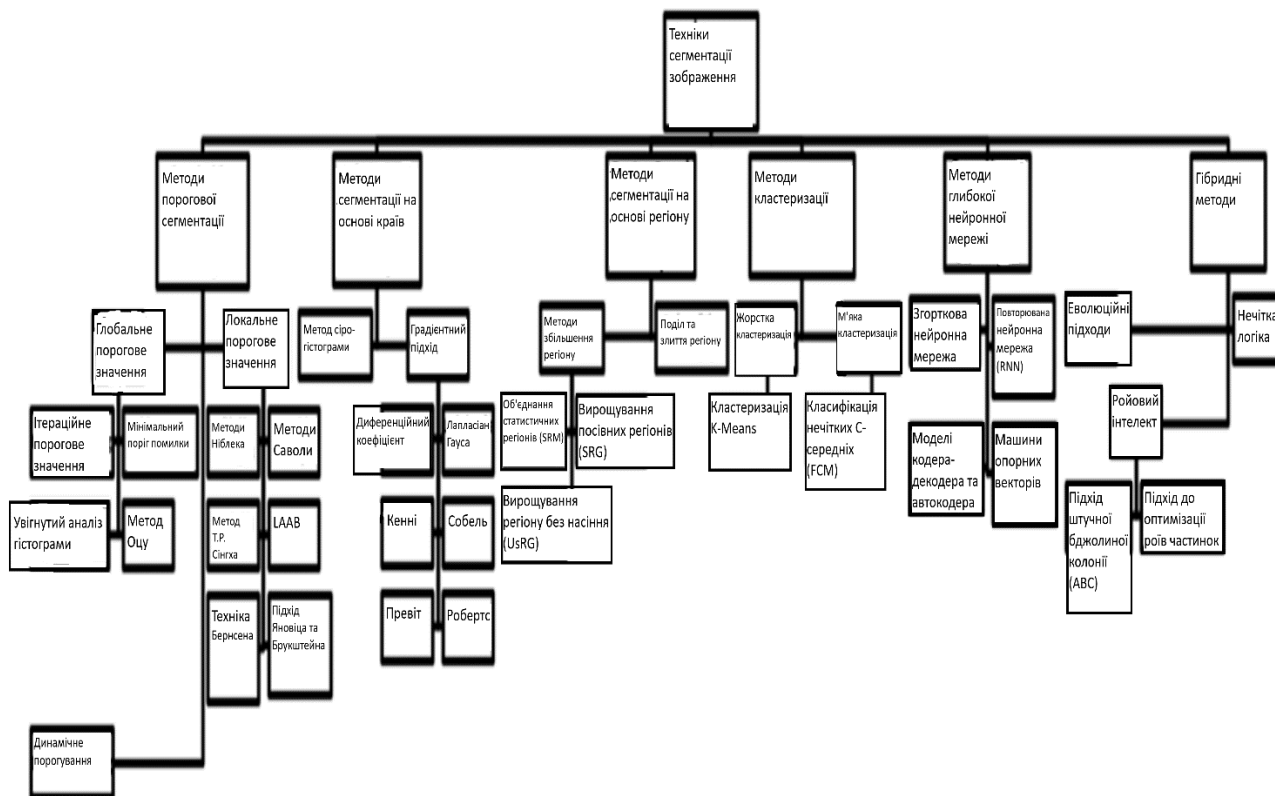


Рис. 3.1.1 Класифікація методів сегментації зображення

3.2 ПІДХОДИ ДО СЕГМЕНТАЦІЇ ЗОБРАЖЕНЬ З ВИКОРИСТАННЯМ ПОРОГОВОЇ ОБРОБКИ

Один із значущих методів сегментації зображень - це порогова обробка. Це простий метод, який використовується для сегментації зображень, і представляє собою чіткий та швидкий спосіб виділення об'єктів на передньому плані від фону. Методи порогової обробки витягують об'єкти з фону, встановлюючи розумний сірий поріг (T) для пікселів зображення [2]. Техніки порогової обробки базуються на особливостях зображень. Вони вибирають відповідні пороги T_n для поділу пікселів зображення на класи та відокремлюють передній план від фону. За одним порогом T "точка об'єкта" визначалася як будь-яка точка (x, y) , для якої $f(x, y) > T$, і "точка фону" визначалася як будь-яка точка (x, y) , для якої $f(x, y) < T$ [4]; це вказує на те, що порогова обробка перетворює сірий рівень на бінарне зображення, перетворюючи кожен піксель зображення, який знаходиться нижче певного порога, в нуль, тоді як кожен піксель, що перевищує певний поріг, стає одиницею. На рисунку 3.2.1 інтенсивний гістограма відображає зображення $f(x, y)$, що складається з освітлення об'єкта на темному фоні, пікселі, пов'язані з об'єктом та фоном, мають інтенсивність, яка групується з двома домінуючими режимами [5].

Техніки порогової обробки були класифіковані на глобальні, динамічні та локальні методи порогової обробки.

3.2.1 Глобальна порогова обробка

У випадку, коли T базується на значеннях сірого рівня, такий метод порогової обробки буде називатися "глобальною пороговою обробкою" [4]. Якщо вибрати відповідне єдине значення порогу T для вхідного зображення, то пікселі зображення будуть поділені на 2 класи (передній план і фон), глобальна порогова обробка вказана у рівнянні (1):

$$g(x, y) = \begin{cases} 1, & \text{if } f(x, y) \geq T_{global} \\ 0, & \text{if } f(x, y) < T_{global} \end{cases} \quad (3.1)$$

У цьому виразі:

$g(x, y)$ представляє бінарне зображення,

$f(x,y)$ представляє оригінальне зображення,

T представляє значення порогу [5].

Існують кілька глобальних методів порогової обробки, включаючи метод мінімальної помилки, ітеративну порогову обробку, аналіз вогкої гістограми, метод Оцу, оптимальну порогову обробку, порогову обробку на основі ентропії та метод утримання порогу Мому [4].

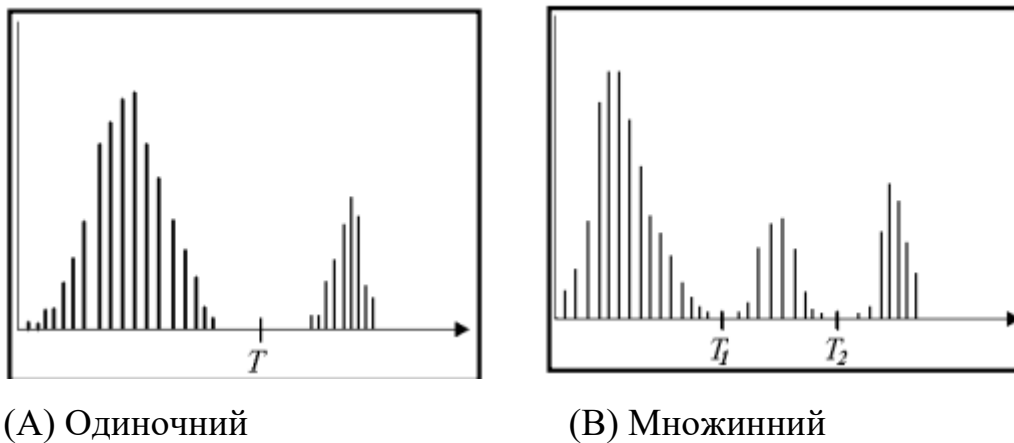


Рис. 3.2.1: Порогова обробка. [4][6]

А) Ітеративна порогова обробка

Це вважається глобальним методом порогової обробки. Це визначається як досить простий підхід, який не вимагає великої кількості знань про зображення, а також показав хороші результати в боротьбі з шумом на зображенні [7]. У дослідженні, проведеному [8], процедура автоматично обчислює поріг інтенсивності T , який вказаний Вудсом та Гонсалесом в (2002 р.) [6]. Нижче подано узагальнення цієї техніки:

Зробіть початкову оцінку, пов'язану з T , наприклад, середній інтенсивності.

Сегментація гістограми зображення на групу $1G1$ інтенсивностей $< T$ та групу $2G2$ інтенсивностей $\geq T$ та обчислення середніх значень інтенсивності μ_1 та μ_2 пікселів $P(I)$ у двох групах за наступним чином:

$$\mu = \sum_{I=0}^{N-1} IP(I) \quad (3.2)$$

$$P(I) = nI / N \quad (3.3)$$

де N представляє загальну кількість пікселів, а nI представляє кількість пікселів із інтенсивністю I .

Обчислення нового порогу:

$$T=0.5(\mu_1+\mu_2) \quad (3.4)$$

Повторення кроків 2 та 3 до тих пір, поки різниця між значенням T та T_n не перевищує попередньо визначений ліміт: $|T-T_n|<0.1$ (3.5)

(В) Мінімізація помилки порогової обробки

Іллінгворт та Кіттлер (1986 р.) запропонували алгоритм мінімізації помилок порогової обробки, який мінімізує ймовірність помилкової класифікації, шляхом врахування виразу помилки. Крім того, гістограма в градаціях сірого в цьому підході вказується як оцінка ймовірності щільності функції $p(g)$ змішаної популяції, що складається з градацій сірого тла та об'єктних пікселів [9][10]. Зазвичай вважається, що два компоненти змішання $p(g|i)$ розподілені нормально з μ_i , стандартним відхиленням σ_i та апіорною ймовірністю p_i , тобто:

$$p(g) = \sum_{i=1}^2 p_i p(g|i) \quad (3.6)$$

$$p(g|i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{g-\mu_i}{2\sigma_i^2}\right) \quad (3.7)$$

Порогове значення може бути вибране за квадратичним рівнянням:

$$\frac{(g-\mu_1)^2}{\sigma_1^2} + \log_e \sigma_1^2 - 2\log_e P_1 = \frac{(g-\mu_2)^2}{\sigma_2^2} + \log_e \sigma_2^2 - 2\log_e P_2 \quad (3.8)$$

Параметри μ_i , σ_i^2 та p_i (де $i=1,2$) щодо щільності змішання $p(g)$, яка пов'язана з зображенням, яке слід пороговувати, зазвичай не відомі. Щоб подолати складність оцінки невідомих параметрів, в дослідженні, проведеному [10], Іллінгворт та Кіттлер надали критерійну функцію $J(t)$, яка визначається так:

$$J(t) = 1 + 2\{P_1(t)\log_e \sigma_1(t) + P_2(t)\log_e \sigma_2(t)\} - 2\{P_1(t)\log_e \sigma_1(t) + P_2(t)\log_e \sigma_2(t)\} \quad (3.9)$$

де L - максимальне значення інтенсивності пікселя на зображенні.

$$P_1(t) = \sum_{g=0}^t h(g), \quad P_2(t) = \sum_{g=t+1}^{t-1} h(g) \quad (3.10)$$

$$\mu_1(t) = \frac{\{\sum_{g=0}^t h(g)g\}}{P_1(t)} \quad \mu_2(t) = \frac{\{\sum_{g=t+1}^{t-1} h(g)g\}}{P_2(t)} \quad (3.11)$$

$$\sigma_1^2(t) = \frac{\{\sum_{g=0}^t (g-\mu_1(t))^2 h(g)\}}{P_1(t)} \quad (3.12)$$

$$\sigma_2^2(t) = \frac{\{\sum_{g=t+1}^{t-1} (g-\mu_2(t))^2 h(g)\}}{P_2(t)} \quad (3.13)$$

Оптимальне порогове значення визначається шляхом мінімізації $J(t)$ із знаходженням:

$$t^* = \text{Arg Min } I \in G J(t)$$

(В) Аналіз вогкої гістограми

Дослідження, проведене [11], вказує, що точки вогкості, пов'язані з максимальною "глибиною", є відмінними кандидатами на порогови. Також такі точки можна ефективно знаходити шляхом проектування опуклої оболонки гістограми. Припускаючи, що задана гістограма H вказана над множиною градацій сірого K, \dots, L з висотами стовбців $h(K), \dots, h(L)$, де $h(K)$ та $h(L)$ ненульові, а всі h - невід'ємні. H представляє 2D-регіон, обмежений ліворуч, знизу та праворуч прямими лініями $(K, 0), (K, h(K)), (K, 0), (L, 0), (L, h(L))$ відповідно (рис. 3).

Щоб знайти вогкості оболонки H , розроблялась мінімальна опукла багатокутник H , який містить H , і вогкості визначалися шляхом взяття множинно-теоретичної різниці $H - H$. (Вогкість можна визначити як зв'язаний компонент, пов'язаний з $H - H$.) Показано, що три прямі лінії, які обмежують H ліворуч, праворуч та знизу, є сторонами H . Таким чином, потрібно побудувати частину H , вказану верхами стовбців гістограми. Був використаний простий алгоритм побудови H за допомогою сканування зліва направо [12].

1. Почніть з точки $(K, h(K))$, обчисліть кут нахилу θ_i відрізка $(K, h(K))$ для $i, K+1 \leq i \leq L$, де $-90^\circ < \theta_i < 90^\circ$.

2. Позначте θ_{K_1} як найбільший з кутів нахилу, а K_1 як праворуч найбільша точка з таким кутом.

3. $(K, h(K)) - (K_1, h(K_1))$ є стороною опуклої оболонки (див. рис. 4).

4. Процес повторюється заміною K на K_1 , що означає обчислення кутів нахилу відрізків для $(K_1, h(K_1))$ до $(i, h(i))$, $K_1+1 \leq i \leq L$.

5. Знаходження найбільшого кута нахилу θ_{K_2} та праворуч найбільшої точки K_2 з таким кутом, що призводить до $(K_1, h(K_1)) - (K_2, h(K_2))$ як сторони опуклої оболонки, і так далі, до досягнення L .

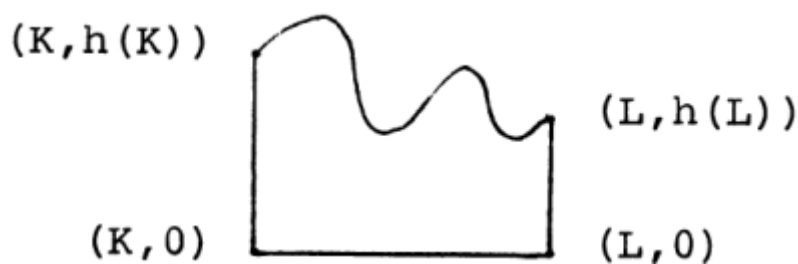


Рис. 3.2.2: Гістограма, визначена як 2D область

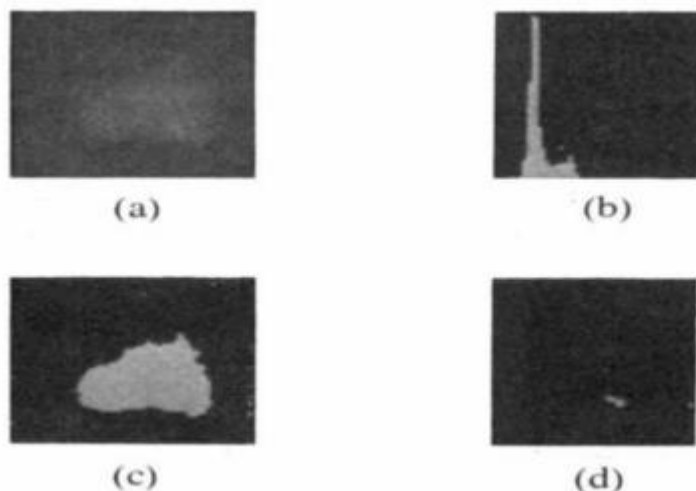


Рис.3.2.3: Тестове зображення. (a) Зображення (b) Гістограма (c) & (d) Порогові зображення

D) Метод Оцу

Метод Оцу може бути визначений як один із ненаглядних та непараметричних підходів до автоматичного вибору порогу для сегментації зображень. Цей підхід використовується для уникнення недоліків ітеративної порогової обробки та оцінює середнє значення на кожному з етапів. Метою методу Оцу є знаходження оптимального значення в термінах глобального порогу [13].

Метод Оцу включає всі можливі значення порогу та обчислює рівні пікселів по обидві сторони значення порогу. Значення порогу розділяє пікселі переднього та заднього плану. Цей алгоритм категоризує зображення на дві категорії пікселів: всередині категорії та міжкатегорійну дисперсію [14-15].

Формулювання цього методу буде представлено наступними рівняннями. Припускаючи, що пікселі зображення представлені у L відтінках сірого $[1, 2, \dots, L]$, кількість пікселів на рівні i позначалася через n_i , а загальна кількість пікселів

позначалася як $N=n_1+n_2+\dots+n_L$. З метою спростити обговорення, гістограма відтінків сірого була нормалізована та визначена як розподіл ймовірностей:

$$p_i = \frac{n_i}{N}, p_i \geq 0, \sum_{i=1}^L p_i = 1 \quad (3.14)$$

Пікселі розділені на дві категорії - C_0 та C_1 (об'єкти та фон, або навпаки) за допомогою порогу на рівні k . C_0 представляє пікселі з рівнями $[1, \dots, k]$, і C_1 представляє пікселі з рівнями $[k+1, \dots, L]$. Крім того, ймовірності класу та середні рівні класу, відповідно, подаються наступним чином:

$$\omega_0 = P_r(C_0) = \sum_{i=1}^k p_i = \omega(k) \quad (3.15)$$

$$\omega_1 = P_r(C_1) = \sum_{i=k+1}^L p_i = 1 - \omega(k) \quad (3.16)$$

та

$$\mu_0 = \sum_{i=1}^k i P_r(i|C_0) = \sum_{i=1}^k \frac{i p_i}{\omega_0} = \mu(k)/\omega(k) \quad (3.17)$$

$$\mu_1 = \sum_{i=k+1}^L i P_r(i|C_1) = \sum_{i=k+1}^L \frac{i p_i}{\omega_1} = \frac{\mu_T - \mu(k)}{1 - \omega(k)} \quad (3.18)$$

$$\text{Де } \omega(k) = \sum_{i=1}^k p_i \quad (3.19)$$

$$\text{Та } \mu(k) = \sum_{i=1}^k i p_i \quad (3.20)$$

Перше та нульове накопичені моменти, пов'язані з гістограмою до рівня k , відповідно:

$$\mu_T = \mu(L) = \sum_{i=1}^L i p_i \quad (3.21)$$

Загальний середній рівень оригінального зображення можна виразити наступним способом для будь-якого вибору k :

$$W_0 \mu_0 + W_1 \mu_1 = \mu_T, \quad W_0 + W_1 = 1 \quad (3.22)$$

Дисперсії класів визначаються як:

$$\sigma_0^2 = \sum_{i=1}^k (1 - \mu_0)^2 \Pr(i|c_0) = \sum_{i=1}^k (i - \mu_0)^2 p_i / W_0 \quad (3.23)$$

$$\sigma_1^2 = \sum_{i=k+1}^L (1 - \mu_1)^2 \Pr(i|c_1) = \sum_{i=k+1}^L (i - \mu_1)^2 p_i / W_1 \quad (3.24)$$

З метою оцінки "якості" порогу (на рівні k), надається наступний критерій дискримінанту (або показники віддільності класу), який використовується в аналізі дискримінанту.

$$\lambda = \frac{\sigma_B^2}{\sigma_W^2}, \quad K = \frac{\sigma_T^2}{\sigma_W^2}, \quad \eta = \frac{\sigma_B^2}{\sigma_T^2} \quad (3.25)$$

де

$$\sigma_W^2 = \omega_0 \sigma_0^2 + \omega_1 \sigma_1^2$$

$$\sigma_B^2 = \omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2 = \omega_0\omega_1(\mu_1 - \mu_0)^2 \quad (3.26)$$

$$\sigma_T^2 = \sum_{i=1}^L (i - \mu_T)^2 p_i \quad (3.27)$$

Техніка Оцу є простою в реалізації, але основним недоліком методу вибору порогу Оцу є припущення, що гістограма є бімодальною. Крім того, вважається, що метод Оцу є ефективним у випадку, коли два класи мають різний розмір та змінюються освітленістю, але він не надає відмінних результатів у випадку, коли гістограма зображення має більше двох піків [13].

Е) Оптимальне порогове значення

Пун (1980, 1981) запропонував оптимальний критерій для порогової обробки зображень. Критерій був виправлений та вдосконалений Капуром та ін. (1985). Дослідники виправили та вдосконалили алгоритм Пуна, враховуючи 2 ймовірнісні розподіли для фону та об'єктів і максимізували ентропію зображення для отримання оптимального порогу [17]. Крім того, оптимальне порогове значення визначається як гістограма зображення, яка є зваженою сумою принаймні двох ймовірнісних щільностей. Потім поріг встановлюється як відтінок сірого, який призводить до мінімальної кількості неправильно класифікованих пікселів, що погано розділяються, що вказує на те, що пікселі фону розділені навпаки із пікселями переднього плану. На рис. (3.2.4) показано шумне зображення, а на рис. (3.2.5) показано оптимальний поріг, а також звичайний поріг, визначений на його гістограмі. Рис. (3.2.6) демонструє, що поріг, визначений оптимально, дає кращий результат порівняно з пороговою обробкою за звичайним пороговим значенням [16].

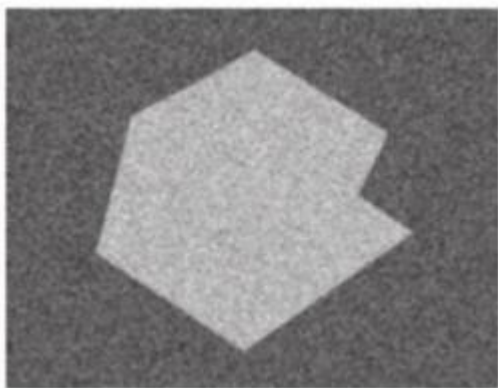


Рис. 3.2.4: Шум зображення

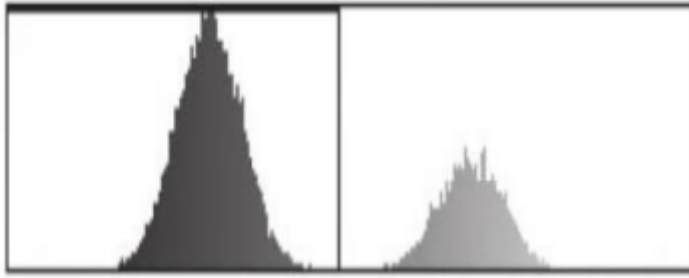


Рис. 3.2.5: Оптимальний поріг і звичайний пороговий виступ



Рис. 3.2.6: Порогове значення на оптимальному пороговому значенні

$$\mu(k) = \sum_{i=k+1}^{L-1} p_i \quad (3.28)$$

Де L - це кількість відтінків сірого (наприклад, 256 для зображення з 8 бітами).

За визначенням:

$$\omega(k) + \mu(k) = \sum_{i=0}^{L-1} p_i = 1 \quad (3.29)$$

Знаходження K відбувається для максимізації різниці між $\omega(k)$ та $\mu(k)$. Це можна виконати, визначивши спочатку середнє зображення як:

$$\mu_T = \sum_{i=0}^{L-1} i p_i \quad (3.30)$$

а потім знаходження значення K , яке максимізує вираз:

$$\frac{(\mu_T \omega(k) - \mu(k))^2}{\omega(k) \mu(k)} \quad (3.31)$$

Це ефективно максимізує варіантність між класами або мінімізує внутрішню варіантність. K вибирається для максимізації відокремлення між двома класами "передній план" та "фон" або альтернативно для мінімізації їхнього перекриття. Крім того, оптимальний поріг, який він знаходить, є стабільним, оскільки він базується на інтеграції гистограми відтінків сірого (тобто глобальних характеристик), а не на її диференціації (тобто локальній властивості, такій як яра). Оптимальну порогову

обробку можна проводити ітеративно за допомогою методу ітеративного аналізу даних самоорганізації (isodata), такий підхід пояснюється наступним чином:

- 1- Порогова обробка зображення за використанням середнього значення для 2 піків або середнього значення пікселів.
- 2- Обчислення середнього значення для пікселів нижче такого порогу μ_1 і середнього значення для пікселів вище такого порогу μ_2 .
- 3- Порогова обробка зображення при новому порозі $T_i = (\mu_1 + \mu_2) / 2$.
- 4- Повторення кроків (2, 3, і 4) до тих пір, поки $T_i - T_{i-1} \leq \Delta$, де зміна Δ може бути визначена різними способами, або оцінюванням відносних змін у значенні порогу, або відсотком пікселів, які змінюють сторони (фон на передній план або навпаки) між ітераціями.

F) Порогова обробка на основі ентропії

Ентропія, яка використовується як міра нечіткості, розглядається в аналогії з ентропією теорії інформації, але з незначними різницями в визначенні. Дослідження, проведене Fan et al. (1996), запропонувало швидкий ентропійний підхід для автоматичного отримання глобального порогу шляхом зменшення обчислювальної складності. Дослідження, проведене Portes de Albuquerque et al. (2004), запропонувало ентропійний підхід до порогової обробки, який був адаптований з концепції ентропії Tsallis. Дослідження, проведене Xiao et al. (2008), запропонувало ентропійний підхід до порогової обробки на основі гістограми просторової кореляції рівнів сірого (GLSC). Дослідники виправили та розширили алгоритм Karur et al.'s (Karur et al. 1985) [17]. Представлене дослідження розглядає порогову обробку ентропії, запропоновану Xiao et al. (2008). Гістограму GLSC оцінюється наступним чином. Припустимо, $g(x,y)$ - це кількість пікселів до невеликих вікон, що містять $N \times N$ пікселів:

$$g(x, y) = \sum_{i=-\frac{N-1}{2}}^{\frac{N-1}{2}} \sum_{j=-\frac{N-1}{2}}^{\frac{N-1}{2}} v(|f(x + i, y + j) - f(x, y)| \leq \zeta) \quad (3.32)$$

Значення сірого кольору пікселів, $g(x,y)$ та $f(x,y)$, використовувалися для створення гістограми GLSC:

$$h(k,m) = \Pr ob (f(x,y)=k \text{ і } g(x,y)=m), \quad (3.33)$$

де $k \in G$, а $m \in \{1, 2, \dots, N \times N\}$.

Функція ймовірності $p(k, m)$ визначається як

$$p(k, m) = h^{(k, m)}.$$

Розподіли ймовірностей, що стосуються фону та об'єкта, були отримані наступним чином:

$$G_{obj} = \left\{ \frac{P(0,1)}{P_{obj}}, \dots, \frac{P(0, N \times N)}{P_{obj}}, \dots, \frac{P(0,1)}{P_{obj}}, \dots, \frac{P(TH, N \times N)}{P_{obj}} \right\},$$

$$P_{obj} = \sum_{k=0}^{TH} \sum_{m=1}^{N \times N} P(k, m),$$

$$G_{obj} = \left\{ \frac{P(0,1)}{P_{obj}}, \dots, \frac{P(0, N \times N)}{P_{obj}}, \dots, \frac{P(0,1)}{P_{obj}}, \dots, \frac{P(TH, N \times N)}{P_{obj}} \right\} \quad (3.34)$$

$$P_{obj} = \sum_{k=0}^{TH} \sum_{m=1}^{N \times N} P(k, m) \quad (3.35)$$

I

$$G_{bg} = \left\{ \frac{P(TH+1,1)}{P_{bg}}, \dots, \frac{P(TH+1, N \times N)}{P_{bg}}, \dots, \frac{P(255,1)}{P_{bg}}, \dots, \frac{P(255, N \times N)}{P_{bg}} \right\} \quad (3.36)$$

$$P_{obj} = \sum_{k=TH-1}^{255} \sum_{m=1}^{N \times N} P(k, m) \quad (3.37)$$

Розрахунок ентропії для елементів у гістограмі GLSC здійснювався з використанням нелінійної функції, зв'язаної з m та N і отриманою за допомогою:

$$\omega(m, N) = \frac{1 + e^{-\frac{9m}{N \times N}}}{1 - e^{-\frac{9m}{N \times N}}} \quad (3.38)$$

Отже, ентропії, пов'язані з розподілами об'єкта та фону, мають вигляд:

$$H_{obj}(TH, N) = - \sum_{k=0}^{TH} \sum_{m=1}^{N \times N} \frac{P(k, m)}{P_{obj}} \ln \left[\frac{P(k, m)}{P_{obj}} \right] \omega(m, N) \quad (3.39)$$

$$H_{bg}(TH, N) = - \sum_{k=TH-1}^{255} \sum_{m=1}^{N \times N} \frac{P(k, m)}{P_{bg}} \ln \left[\frac{P(k, m)}{P_{bg}} \right] \omega(m, N) \quad (3.40)$$

3.2.2 Місцеве порогове значення

Для локального порогового методу необхідно вибрати декілька порогів сегментації та поділу зображення на кілька цільових областей та фонів за кількома порогоми [18]. У випадку, коли поріг T залежить від значення сірого кольору точки та локальної властивості певної точки, такої як середнє значення сірого кольору для сусідніх точок (x, y) , поріг називається "локальним порогом" [4]. Підхід визначає значення порогу локально для кожного з регіонів зображення, як показано у рівнянні (41) [19]. Крім того, одиничний поріг не працюватиме ефективно в разі нерівномірної

освітленості через тіні або напрям освітлення. Основний принцип полягає в розділенні зображення на підзображення розміром $m \times m$ і подальшому виборі порогу.

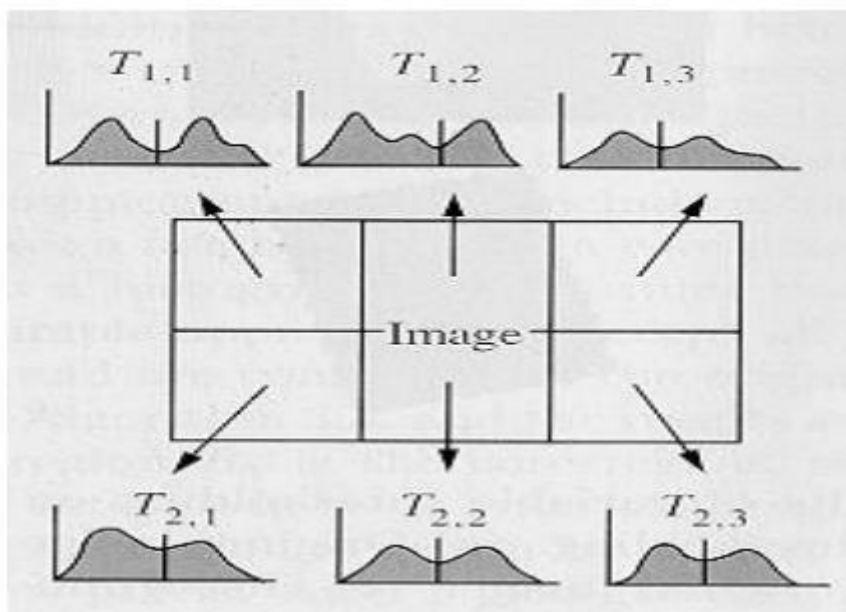


Рис.3.2.7: Підзображення зображення [13]

$$g(x, y) = \begin{cases} 1, & \text{if } f(x, y) \geq T_{local} \\ 0, & \text{if } f(x, y) < T_{local} \end{cases} \quad (3.41)$$

У такому випадку можуть використовуватися багатозначні локальні пороги для багатьох областей зображення, такий процес відомий як багаторівневе порогове визначення. Цей процес можна виразити рівнянням (42) [19].

$$g(x, y) = \begin{cases} a, & \text{if } f(x, y) > T_2 \\ b, & \text{if } T_1 < f(x, y) \leq T_2 \\ c, & \text{if } f(x, y) \leq T_1 \end{cases} \quad (3.42)$$

Гістограма зображення визначається трьома управляючими режимами (приклад двох типів світлих об'єктів на темному тлі). Якщо $T_1 < f(x, y) \leq T_2$, то точку (x, y) можна класифікувати як один клас об'єктів, у випадку, коли $f(x, y) > T_2$, вона класифікується як інший клас об'єктів, і якщо $f(x, y) \leq T_1$, то вона класифікується як фон, як показано на рис. (2). Багато аспектів впливають на придатність гістограми для вибору порогу [5][19]:

- Шум зображення.
- Відносний розмір фону та об'єктів.
- Регулярність освітлення.
- Регулярність відбиття.

А) Методи Niblack

Пороги Niblack визначаються як локальні методи визначення порогу, які мають велике значення для зображень, на яких фон не є однорідним, особливо для розпізнавання тексту. На рис. (9) показано багато методів сегментації, таких як поріг Niblack. Щодо цього методу, локальне значення порогу $T(x, y)$ у точці (x, y) обчислюється вікном розміром $w \times w$ наступним чином:

$$T(x,y)=m(x,y)+k\cdot\delta(x,y)$$

де $m(x, y)$ та $\delta(x, y)$ - це локальне середнє та стандартне відхилення щодо пікселів у локальному вікні, а k представляє зміщення. Встановлено $k = -0,20$, а розмір локального вікна - $w=15$. Стандартне відхилення $\delta(x, y)$ та локальне середнє $m(x, y)$ адаптують значення порогу в залежності від контрасту в локальному сусідстві пікселя. Також зміщення k контролює рівень адаптації, змінюючи значення порогу. Тут піксель = (піксель $>$ середнє + $k \cdot$ стандартне відхилення): об'єкт : фон [20].

Цей підхід дає хороші результати, коли зміна інтенсивності відрізняється вздовж положення пікселя, але недолік полягає в тому, що значення порогу ґрунтується на виборі значення зсуву [21].

В) Метод Sauvola

Цей метод, запропонований Sauvola, також є методом локального визначення порогу, який може бути корисним для зображень, на яких фон нерівномірний, особливо для розпізнавання тексту, так само, як і поріг Niblack. Поріг $T(x, y)$ в методі Sauvola обчислюється з використанням стандартного відхилення $\delta(x, y)$ та середнього значення $m(x, y)$ інтенсивностей пікселів у вікні розміром $w \times w$, центрованому навколо пікселя (x, y) , і його можна представити у вигляді:

$$T(x, y) = m(x, y) \left[1 + k \frac{\delta(x,y)}{R} \right] \quad (3.44)$$

Перевагою методу Sauvola є те, що він дає задовільні результати там, де є зміна між фоном та варіацією переднього плану на зображенні; однак недоліки полягають в тому, що значення порогу залежить від значення зміщення та максимального значення стандартного відхилення (SD) для зображення [21].

R представляє максимальне значення SD (R=128 для зображень в градаціях сірого), а k представляє параметр, який може приймати значення від [0,20 до 0,50]. Де $\text{pixel} = (\text{pixel} > \text{mean} \times (1 + k \times (\text{SD} / r - 1))) ? \text{object} : \text{background}$ [20].

У випадку високого рівня контрасту в певній області зображення, $\delta(x, y) \sim R$, що призводить до $T(x, y) \sim m(x, y)$, що представляє схожий результат, як і в методі Niblack. Однак відмінність виникає у випадку, коли контраст місцевого оточення досить низький. У такому випадку поріг $T(x, y)$ зменшується менше, ніж середнє значення, що призводить до успішного вилучення досить темних областей фону [22]. Для обчислення значення порогу $T(x, y)$ потрібно обчислити місцевий SD та середнє для кожного пікселя. Його обчислювальна складність в $O(n^2 \times w^2)$ у наївний спосіб для зображення розміром $n \times n$ [23]. Це вказує на те, що його обчислювальна складність залежить від розміру вікна. Т. Сінгх запропонував підхід до порогової обробки, яка не залежить від розміру вікна, використовуючи інтегральне сумування зображення як попередній процедури. Рис. (9) показує різні методи сегментації, включаючи метод Sauvola.

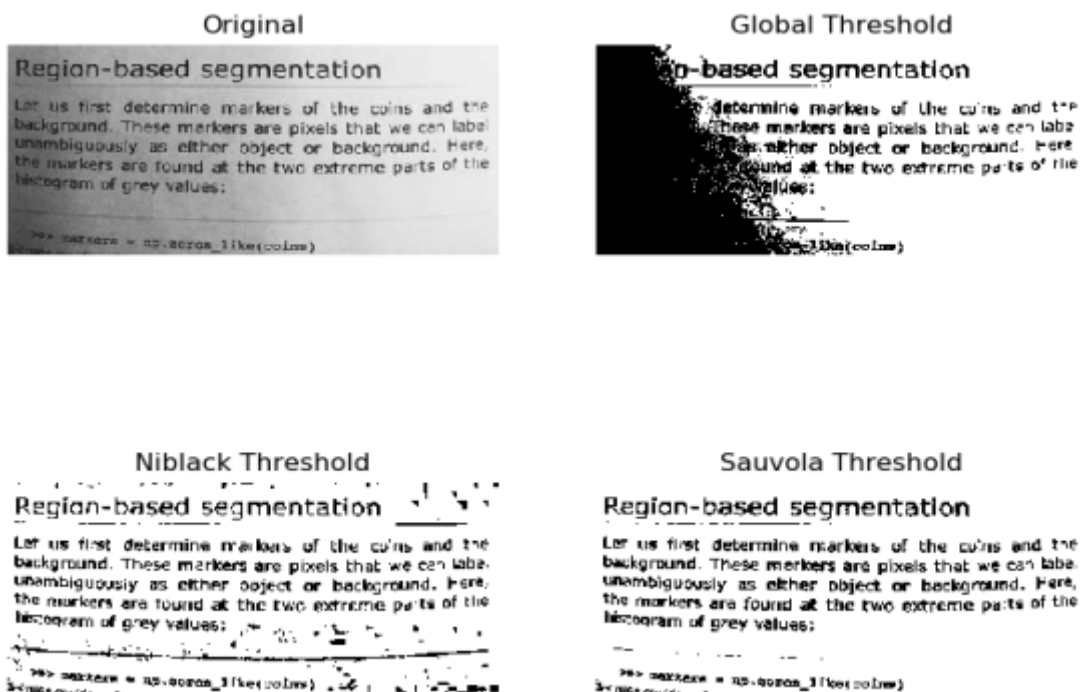


Рис. 3.2.8: Різні методи сегментації

С) Метод Т.Р Сінгха

З метою мінімізації обчислювального часу обчислення локального порогу, Сінгх запропонував достатній спосіб визначення локального порогу з використанням інтегрального сумування зображення як однієї з попередніх процедур для визначення локальної суми [20]. Він використовує лише місцеве середнє і є досить зручним за використання інтегрального сумування, тоді як інші методи, такі як Niblack і Sauvola, не є зручними через використання як стандартного відхилення, так і місцевого середнього. Внаслідок використання інтегрального сумування зображення цей метод не залежить від розміру локального вікна. Цей підхід представлений рівнянням:

$$T(x, y) = m(x, y) \left[1 + k \frac{\delta(x, y)}{1 - \delta(x, y)} - 1 \right] \quad (3.45)$$

Де

$\delta(x, y) = I(x, y) - m(x, y)$ представляє місцеве середнє і є зсув, який може контролювати ступінь адаптації змінності значення порогу. Значення k може суттєво впливати на визначення значення порогу. Зниження значення k призводить до збільшення значення порогу і навпаки [20].

D) LAAB (Локальна адаптивна автоматична бінаризація)

Автоматична бінаризація, яку запропонував Сінгх, може бути представлена як процедура перетворення зображення в градаціях сірого $I(x, y)$ у бінарне зображення $b(x, y)$ автоматичним чином без використання будь-якого порогового значення $T(x, y)$ шляхом адаптації пікселів в області локального середовища. Це автоматизована бінаризація з локальними адаптаціями [24]. Локальні адаптації виконуються в локальному вікні розміром $w \times w$ за допомогою локального середнього $m(x, y)$ значень інтенсивностей пікселів в локальній області. Автоматизована бінаризація розроблена на основі наступного рівняння:

$$b(x, y) = |1 - 2v| - |1 - 2v| / 2 |1 - 2v| \quad (3.46)$$

де $v = k(1 + \delta) / (1 - \delta)$, а k представляє значення зсуву таким чином, що $0.50 < k < 0.60$.

$\delta = \{g(x, y) - m(x, y)\} \{1 - m(x, y)\}$, а $g(x, y)$ представляє оригінальний піксель у точці (x, y) . Зсув відповідає за контроль рівня адаптації пікселя в локальній області під час перетворення на бінаризоване зображення $b(x, y)$. Збільшення значення k призводить до збільшення фону і зменшення області переднього плану і навпаки [20].

Е) Техніка Бернсена

Цей підхід, запропонований Бернсенем, є методом локальної бінаризації, використовуючи значення локального контрасту для визначення значення локального порогу для кожного з пікселів (x, y) , що обчислюється за рівнянням:

$$T(x,y)=(I_{\max}+I_{\min})/2$$

Де I_{\max} та I_{\min} вказують на максимальне і мінімальне значення відтінку сірого в вікні розміром $w \times w$, яке центроване відповідно в точці (x, y) [25]. Однак присвоєння порогу базується на значенні локального контрасту, тому його можна представити як: якщо значення $(I_{\text{max}} - I_{\text{min}}) > L$, тобто в тому випадку, коли зображення в градаціях сірого нерівномірне: Тоді

$$T(x,y)=(I_{\max}+I_{\min})/2$$

У іншому випадку $T(x,y)=GT$ У цьому випадку значення порогу обчислюється з методу глобальної бінаризації.

Де L представляє поріг контрасту, а GT - значення глобального порогу.

Ф) Підхід Яновіца та Брукштейна Брукштейн та Яновіцз запропонували використання значень відтінку сірого в областях з високим градієнтом як відомих даних для інтерполяції поверхні порогу текстурних характеристик зображення [24]. Основні операції цього підходу можна узагальнити наступним чином:

1. Згладжування зображень за допомогою використання фільтрів середнього значення.
2. Отримання величини градієнту.
3. Застосування алгоритму зтонування для визначення граничних точок об'єкта.
4. Вибір відтінків сірого на згладженому зображенні в граничних точках. Ці точки можуть бути визначені як опорні точки для інтерполяції на кроці 5.
5. Знаходження поверхні порогу $T(x,y)$, яка дорівнює значенням зображення в опорних точках і задовольняє рівнянню (3.47) за допомогою методу послідовної релаксації Саутвелла.

$$\frac{\partial^2 p(x,y)}{\partial x^2} + \frac{\partial^2 p(x,y)}{\partial y^2} \quad (3.47)$$

6. Використовуючи отриманий $T(x,y)$, сегментація зображення.

7. Застосування підходу після обробки для підтвердження сегментованого зображення.

3.2.3 Динамічна бінаризація

Локальна адаптивна бінаризація виконується за допомогою аналізу локальних ознак зображення для визначення порогового значення для кожного пікселя; такий метод іноді називають динамічним порогом [26-27]. У випадку, коли в зображенні є кілька об'єктів, які знаходяться в різних відтінках сірого, зображення слід розділити різними динамічними пороговими значеннями (T_1, T_2, \dots, T_n), які залежать від $p(x, y)$, $f(x, y)$, та просторових координат x та y . Загалом методи динамічних порогів включають в себе методи Watershed, бінаризацію зображення, інтерполяційну бінаризацію тощо [4]. Один із методів динамічної бінаризації (Моррін / 1974) використовує градієнт краю для відхилення значення порогу від локального середнього відтінків сірого для отримання чіткого зображення. Поточний піксель встановлюється в чорний колір, якщо поточне значення скану x менше порогового значення $T(x)$, де $T(x) = A + G + b$ де

A - середнє значення оточуючих пікселів,

G - градієнт для поліпшення,

b - фіксована від'ємна константа для відхилення фону в бік білого.

Бажаною операцією градієнту в динамічній бінаризації є чутливість до країв об'єктів, але нечутливість до випадкового шуму та півтонового фону. Міні-макс кутова – сума градієнтів у матриці розміром (3×3) та кластеризація скан-пікселів поліпшують відкидання випадкового шуму в документі (Вонг / 1978). У динамічному алгоритмі бінаризації, розробленому для застосувань OCR (бінарний оптичний розпізнавач) (Уайт та Рорер / 1983), для оновлення рівнів фону в горизонтальному та вертикальному напрямках використовуються два диференціальні рівняння першого порядку. Схема бінаризації подібна до тієї, що використовується у (Моррін / 1974), за винятком того, що вона не використовує підвищення градієнту.

Ми узагальнюємо переваги та недоліки методу порогової сегментації зображення, який був запропонований в літературі.

ТАБЛИЦЯ 1 ПІДСУМОК ПЕРЕВАГ І НЕДОЛІКІВ ТЕХНІК ПОРОГОВОЇ СЕГМЕНТАЦІЇ ЗОБРАЖЕНЬ [18] [28]

Переваги	Недоліки
<p>розрахунки прості</p> <p>Обчислювально недорогий.</p> <p>Не вимагає попередніх знань про зображення.</p> <p>Може використовуватися в програмах реального часу.</p> <p>швидкість операції вище</p> <p>Зокрема, коли фон і ціль мають високий контраст, може бути створений ефект сегментації.</p>	<p>Для зображень із плоскими та широкими западинами або без піків він не працює достатньо.</p> <p>Є труднощі в отриманні точних результатів для проблем сегментації зображення, в яких немає суттєвих відмінностей у градаціях сірого або великих значень градацій сірого, що накладаються на цьому зображенні.</p> <p>Нехтує просторовою інформацією зображення, не може гарантувати, що сегментовані області є суміжними.</p> <p>Через те, що він може розглядати лише сіру інформацію зображення без урахування просторової інформації зображення, він має чутливість до нерівномірності градації сірого та шуму, що зазвичай призводить до поєднання його з іншими підходами.</p> <p>Вибір порогового значення має вирішальне значення, оскільки неправильний вибір може призвести до надмірної або недостатньої сегментації.</p> <p>Висока чутливість до шуму.</p>

3.3 Техніки сегментації на основі країв

Сегментацію зображення можна визначити як підхід, який використовується для відокремлення об'єкта на зображенні від фону та визначення його змісту. Край розглядається як важлива та значуща ознака зображення. Край можна характеризувати як межу, яка розділяє 2 різні області зображення. Операція виявлення краю означає процедуру виявлення та локалізації різких розривів, які можуть бути присутніми на зображенні [29].

Методи виявлення краю є самостійними та добре розвиненими методами обробки зображень. Підходи до сегментації на основі країв базуються на раптовій зміні інтенсивності на зображенні через те, що одне значення інтенсивності не надає достатньо даних щодо країв. Методи виявлення краю відповідальні за визначення положення країв, при цьому або перший похід інтенсивності вищий за певний поріг, або другий похід має перетин з нулем.

У підходах сегментації на основі країв спочатку виявляють краї, а потім їх з'єднують для формування меж об'єктів для сегментації необхідних областей. Існують два основних підходи сегментації на основі країв: сіро-гістограмний підхід і градієнтний підхід [3][4]. Три найбільш використовувані градієнтні підходи - це підхід диференційного коефіцієнта, метод Кенні та Лапласа Гаусіана (LoG) [4].

3.3.1 Метод сіро-гістограми

Якість виявлення краю значно залежить від адекватності порога T . Однак виявлення максимальних і мінімальних відтінків сірого фактично ускладнене через те, що гістограма сірого нерівномірна через вплив шуму. У такому випадку можна приблизно замінити криві об'єкта та фону двома конічними гаусівськими кривими, перетин яких є мінімумом гістограми. Значення порогу T представляє відтінок сірого значення цієї долини.

3.3.2 Градієнтний підхід

Градієнт можна визначити як перший похід зображення $f(x, y)$. У випадку, коли зміна відтінку сірого біля краю достатньо інтенсивна і на зображенні мало шуму, градієнтний підхід працює ефективно, і результат сегментації адаптивний до орієнтації градієнта [4]. Принцип градієнтної сегментації можна описати як

асоціювання меж об'єкта із щільністю гребенів градієнта, помічених на зображенні. Для отримання змістовної сегментації зображень необхідно виконати ряд умов. У випадку низького внутрішнього роздільного здатності пристрою зображення порівняно з розміром вокселів переходи між областями різної активності на зображенні виглядають розмиті. Тому вершини інтенсивності градієнта не є різкими, що ускладнює їх визначення. Ще одне обмеження пов'язане з тим, що шум може бути підсилений на зображенні інтенсивності градієнта порівняно з початковим зображенням. Як перша апроксимація, ефекти шуму та роздільної здатності можуть бути показані так. Спочатку «ідеальне» невідоме зображення (яке не має жодного шуму і має нескінченно велику роздільну здатність) пройшло розмиття через згладжування цього зображення точковою функцією розподілу пристрою зображення. По-друге, це розмите зображення буде викривлене за допомогою статистичного шуму. Реально отримані зображення можна розглядати як результат такої двохетапної моделі. Це навело на думку, що ідеальні зображення можуть бути відновлені за допомогою оберненої моделі, яка була зазначена вище, на першому наближенні [30]. Існують три найбільш використовувані градієнтні підходи, які включають в себе: LoG, метод Кенні та метод диференційного коефіцієнта. Серед них метод Кенні визначається як найбільш представницький метод [4]. Окрім методів Собеля, Превітта та Робертса.

А) Техніка диференційного коефіцієнта

Використовуючи підходи, що базуються на часткових диференціальних рівняннях (ЧДР), і вирішуючи ЧДР за допомогою числової схеми, можна провести сегментацію зображення. Найбільш значущими підходами сегментації зображень, що базуються на ЧДР, є: метод швидкого маршрутування, параметричні методи та методи рівня.

• Параметричні методи

Методи Лагранжа базуються на параметризації контурів за певною стратегією вибірки та подальшого еволюціонування кожного з елементів на основі внутрішніх термінів та зображення. Ці підходи відомі своєю швидкістю та ефективністю; однак оригінальна формула, яка є "чисто параметричною" (заснованою на роботі

Терзопулоса та Касса у 1987 році та відомою як "змії"), отримала значну критику, переважно через її недоліки, такі як вибір стратегії вибірки, внутрішні геометричні характеристики кривої, зміни топології (розділення та злиття кривої), вирішення проблем у деяких вищих вимірах тощо. Сучасні "достатньо дискретизовані" формулювання були розроблені для вирішення цих проблем, зберігаючи при цьому високий рівень ефективності. У обох випадках мінімізація енергії, як правило, виконується за допомогою градієнтного спуску, при цьому похідні обчислюються, наприклад, за допомогою скінчених різниць [31].

- **Методи рівня**

Цей підхід був запропонований спочатку Ошером і Сетіаном в 1988 році для відстеження рухомих інтерфейсів та впроваджений в різноманітні області зображення наприкінці 90-х [32-33]. Він може бути використаний для ефективного вирішення проблем кривих, поверхонь та інших подібних об'єктів, що поширюються неявним способом. Основна ідея полягає в представленні еволюційного контуру за допомогою знакової функції, де його рівень 0 відповідає фактичному контуру. Після цього, на основі рівняння руху контуру, легко отримати еквівалентний потік для неявної поверхні, який, у випадку застосування до рівня 0, відображає поширення контуру. Цей підхід може бути використаний завдяки широкому спектру переваг: він параметричний, неявний, внутрішній, представляє прямий метод для оцінки геометричних характеристик розвиваючихся структур і може змінювати топологію. Крім того, його можна використовувати для визначення фреймворку оптимізації, як це було запропоновано в 1996 році Жао, Ошером і Мерріманом. Таким чином, можна зробити висновок, що це досить зручна модель для вирішення широкого спектру завдань медичного аналізу зображень та комп'ютерного зору. Крім того, дослідження різноманітних структур даних рівня видало високоефективні реалізації цього підходу [31].

- **Швидкі методи маршрутизації**

Цей підхід використовується в сегментації зображень та був удосконалений (що призводить до можливості позитивного та негативного поширення швидкості) в методі, який отримав назву загального методу швидкої маршрутизації [33]. Цей

підхід можна розглядати як числовий метод для вирішення проблем граничного значення ейкональної формули [31][33].

$$|\nabla T(x)| = F(x) = 1 \quad (3.48)$$

Зазвичай такі проблеми описують еволюцію замкненої кривої як функцію часу T зі швидкістю $F(x)$ у нормальному напрямку в точці x на кривій. Функція швидкості ідентифікується, і час, коли контур перетинає точку x , визначається вирішенням формули. Алгоритм схожий на алгоритм Дейкстри і використовує той факт, що інформація може рухатися тільки від області посіву. Ця проблема була визначена як один із спеціальних випадків підходів рівня. Існують ще більш загальні алгоритми, але вони, як правило, більш витратні за часом [31]. Розширення для неплоских (трикутних) областей вирішення:

$$F(x) |\nabla_s T(x)| = 1 \quad (3.49)$$

для поверхні S , та $x \in S$ було введено Реном Кіммеєм та Сетіаном. Метод швидкої маршрутизації підходить лише для проблеми граничного значення. Основна концепція, яка лежить в основі цього підходу, полягає в тому, що проблема граничного значення $|\nabla T| = F = 1$ включає фронт, який завжди зменшується або розширюється.

В) Лапласіан гаусіана (LoG) Лапласіан гаусіана (LoG) використовує гаусів фільтр для розмиття зображення та лапласіан для підсилення контурів. Рідко використовується самостійно для виявлення контурів через його чутливість до шуму. Локалізація контурів виконується за допомогою пошуку нульових перетинів. Також відомий як виявлювач контурів Марра і Гільдрета [34].

Алгоритм LoG пояснюється наступними кроками:

Згладжування зображення 2D гаусівською функцією.

Обчислення лапласіана згладженого зображення L .

Визначення пікселів контуру, які мають нульовий перетин в L .

Радіально-симетрична 2D гаусівська функція:

$$G_0(x, y) = e^{-r^2 / (2\sigma^2)}$$

$$G_0(x, y) = e^{\left(\frac{-r^2}{2\sigma^2}\right)} \quad (3.50)$$

Де

$$r^2 = x^2 + y^2 \quad (3.51)$$

Лапласіан цього:

$$G_0(x, y) = \left(\frac{r^2 - \sigma^2}{\sigma^4}\right) e^{\left(\frac{-r^2}{2\sigma^2}\right)} \quad (3.52)$$

Ця функція має мінімум в своєму центрі, але зазвичай фільтр інвертують, щоб він мав максимум у своєму центрі. Це класична форма "Мексиканського капелюха", яку можна побачити на рисунку (10).

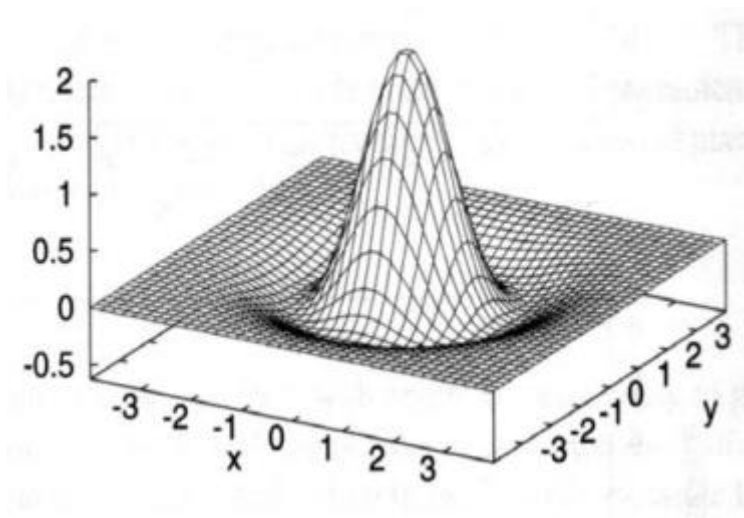


Рис. 3.3.1: Класична мексиканська шапка

Значення σ відповідає за визначення ширини фільтра та контроль за кількістю згладжування, яке виникає внаслідок гаусівського компоненту.

- σ налаштовує фільтр для виявлення контурів на різних масштабах.
- Повинен мати напівширину принаймні 3σ .

$$\begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & 16 & -2 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

Лапласіан гауссового ядра 5×5 . LoG можна апроксимувати шляхом згортання з ядром, яке є різницею двох гаусових ядер із суттєво різними σ . Який відомий як різниця Гаусса (DoG).

С) Техніка Canny

У [35] методи Canny були представлені ДЖОНОМ КЕННІ в 1986 році для виявлення країв. Canny — ефективний і результативний метод виявлення країв, який дає значне зниження вимог до пам'яті, зменшення затримки та збільшення пропускну здатності без будь-яких втрат ефективності виявлення краю [36]. Виявлення країв зображення може бути корисним для сегментації зображення, стиснення даних, а також грати роль у деяких важливих операціях, таких як модернізація зображення [29].

Детектор країв canny можна охарактеризувати як найбільш строго визначений оператор, і він зазвичай використовується серед підходів до виявлення країв. Популярність цього детектора пояснюється тим фактом, що це найкращий підхід для пошуку країв із хорошою локалізацією, хорошим виявленням, а також єдина відповідь на край. Канні може визначити межі шляхом впровадження процесу оптимізації та запропонував наближення до оптимального детектора як максимуми величини градієнта зображень, згладжених за Гаусом [29].

Існує п'ять окремих кроків для виявлення країв методом виявлення країв Canny:

Крок 1. Згладжування: цей процес виконується за допомогою детектора canny, який спрямований на фільтрацію шуму в оригінальному зображенні. Це здійснюється шляхом перетворення вхідного зображення у відтінки сірого за допомогою регулювання яскравості та контрасту, таким чином зображення розмивається з метою видалення шуму. Загалом, для видалення шуму використовується фільтр Гауса.

Крок 2. Пошук градієнтів: пікселі краю – це ті, у яких є раптова зміна значень рівня сірого. Ці пікселі можна охарактеризувати через обчислення градієнта зображення, який можна визначити як одиничний вектор, що вказує на максимальну зміну інтенсивності. Спочатку обчислюються горизонтальні та вертикальні компоненти градієнта, а потім обчислюються напрямок і величина градієнта.

Крок 3. Немаксимальне придушення: по суті, зрідження краю здійснюється в режимі безмаксимального придушення. Під час цієї процедури, на основі величин градієнта, товсті краї зображення перетворюються на майже гострі та тонкі краї, які можна додатково використовувати для цілей розпізнавання. На цьому кроці

зображення скануються вздовж краю, і будь-які значення пікселів, які не вважаються краєм, буде відкинуто, що створить тонку лінію на результуючому зображенні.

Крок 4. Подвійне порогове значення: два значення порогу визначаються в методі виявлення тонкого краю: T_1 , що представляє високий поріг, і T_2 , яке представляє низький. Пікселі, які мають значення рівня градації сірого, які перевищують T_1 , можна вважати сильними пікселями краю, а крайова область є результатом. Пікселі, які мають значення рівня градації сірого $< T_2$, можна вважати слабкими пікселями краю, а результатом є область поза межами. У випадку, коли пікселі мають значення градації сірого в діапазоні від T_1 до T_2 , результат залежатиме від сусідніх пікселів.

Крок 5. Відстеження країв через гістерезис: краї, які не з'єднуються з дуже сильним (тобто сильним) краєм, усуваються в остаточному вихідному зображенні. Сильні краї також можна назвати «Певними краями» та включаються до остаточного зображення краю. Краї, які не є сильними, але з'єднані з сильними краями, були включені у вихідне зображення.

Підхід Canny включає наступні кроки, які виконуються послідовно:

- 1- Зображення фільтрується низькочастотним фільтром за допомогою маски Гауса.
- 2- Обчислення вертикального та горизонтального градієнтів у кожному місці пікселя.
- 3- Розрахунок величини градієнта в кожному місці пікселя.
- 4- Розрахунок нижнього та вищого порогів на основі гістограми градієнта всього зображення.
- 5- Придушення немаксимальних сильних (NMS) країв.
- 6- Виконання порогового значення гістерезису з метою визначення карти краю.

Переваги методу Canny Edge Detection:

- Достатня локалізація.
- Покращене співвідношення сигнал/шум.
- Низькі помилкові відгуки (відповіді з одним фронтом).

- проведення методів обробки зображень на реконфігурованому обладнанні призводить до мінімізації витрат часу виходу на ринок.

- Більш достатнє виявлення в умовах шуму.
- Спрощена перевірка та налагодження.
- Забезпечує можливість швидкого створення прототипів складних методів.

Недоліками Canny Edge Detector є:

- Складні та трудомісткі обчислення, оскільки хитрі методи виявлення країв включають тривалі етапи попередньої та постобробки.

- Важко визначити загальне порогове значення, яке ефективно працює на кожному зображенні.

- Помилковий перехід через нуль.
- Високі вимоги до пам'яті, що призводить до високого ступеня затримки.

С) Метод Собеля

Оператор Sobel можна визначити як підхід, який використовує середнє значення верхнього та нижнього, правого та лівого сусідніх пікселів, а екстремальне значення отримується на краю. Завдяки середньозваженому значенню оператор має лише можливість отримувати інформацію про край, а також здатність захисту від шуму, це звичайний оператор виявлення краю. Значення S_x & S_y оператора Sobel можуть бути обчислені відповідно до шаблону згортки.

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

(а) горизонтальний напрямок (б) вертикальний напрямок

Рис. 3.3.2: Шаблони оператора Sobel [37]

На основі інформації про виявлення зображення можна вважати вигнутою поверхнею, край якої є найбільшою зміною. Інформація про край включає 2 аспекти: один з них - напрямок краю, ізольовані краї незначні, рішенням є пришивання

відповідного краю до лінії краю. Іншим є конкретне розташування, наприклад координати пікселів [38]. На основі диференціальних знань інформацію про положення на поверхні можна судити за допомогою виведення. Виявлення країв на основі диференціального оператора 1-го порядку розглядає зображення як поверхню, а після цього застосовує оператор градієнта до зображення $f(x,y)$, нарешті, форми векторного поля.

$$\nabla(x, y) = \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \quad (3.53)$$

Поле вектора представляє 2 інформаційні аспекти: один з них – це локальна сила градієнта.

$$\|\nabla f(x, y)\| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \quad (3.54)$$

Інший – локальна градієнтна орієнтація.

$$\nabla f(x, y) = \arctan \left(\frac{\partial f}{\partial x} / \frac{\partial f}{\partial y} \right) \quad (3.55)$$

У конкретній експериментальній схемі визначення краю індивідууми зазвичай цікавить спосіб визначення $\|f(x,y)\|$, який є способом отримання інформації про напрямок. А щодо диференціального оператора, різниця або інша подібна технологія зазвичай використовується для дискретного, в кінцевому підсумку еволюціонувала до форми шаблонного процесу; це досить просто. Узагальнений тип використання оператора Собеля для обчислення значення $\|f(x,y)\|$ можна представити на основі рівняння нижче:

$$|T_1 f(i, j)| = \sqrt{(|s_x|^2 + |s_y|^2)} \quad (3.56)$$

S_x і S_y такі:

$$S_x = f(i+1, j+1) + 2f(i+1, j) + f(i+1, j-1) - f(i-1, j+1) - 2f(i-1, j) - f(i-1, j-1) \quad (3.57)$$

$$S_y = f(i+1, j+1) + 2f(i, j+1) + f(i-1, j+1) - f(i-1, j-1) - 2f(i, j-1) - f(i+1, j-1) \quad (3.58)$$

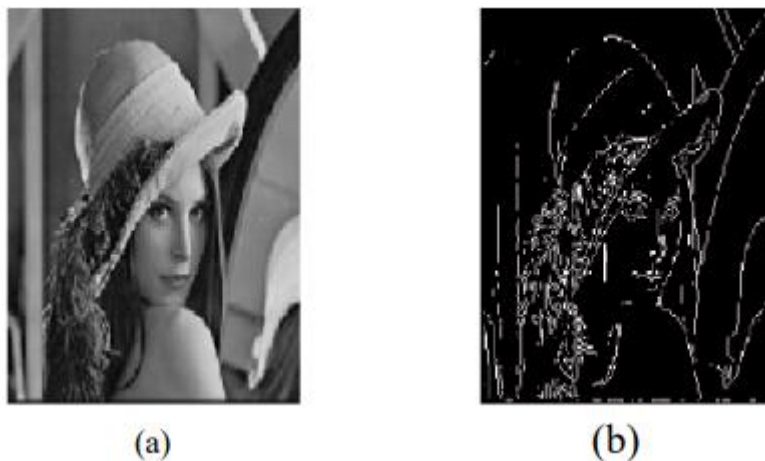


Рис. 3.3.3: (a) оригінальне зображення, (b) зображення, яке було виявлено за допомогою Sobel

Через порівняння рис. (12) оператор Sobel збільшує коефіцієнт ваги центрального пікселя в підлозі шаблону, тому він має здатність виявляти краї без будь-яких великих мутацій градієнта. Експерименти показали: звичайний оператор Собеля, створений за ознакою теорії, згідно з якою функція Грея має здатність отримувати максимум частини на краю, має здатність досягати виявлення краю за допомогою різниць, зважених за сірим. підсумовування результатів із сусідніх пікселів 4 напрямків. Такий тип алгоритму можна легко реалізувати, і він має здатність надавати точну інформацію про напрямок краю. Він має згладжуючий вплив на шум і здатний придушувати шум, особливо у випадку використання в зображеннях великого розміру. Тим не менш, як видно з результатів моделювання, є помилкові краї, точність розташування країв зображення є поганою. У разі використання у зображеннях великого розміру обсяг обчислень великий, а край неочевидний; обсяг розрахунків зростає зі збільшенням напрямку [37].

D) Метод Превітта

Превітта можна визначити як дискретний оператор диференціювання, який обчислює функцію апроксимації градієнта інтенсивності зображення в кожній точці зображення [39]. Результат оператора Превітта може відповідати вектору градієнта норми цього вектора. Крайовий детектор Превітта розглядається як один із придатних способів оцінки орієнтації та величини краю. Незважаючи на те, що виявлення краю диференціального градієнта потребує досить тривалого обчислення для оцінки

орієнтацій за величинами напрямків x і y , виявлення краю компаса отримує орієнтацію прямим способом від ядра з максимальним ступенем відповіді. Оператор Prewitt був обмежений вісьмома варіантами напрямків, тим не менш, досвід показав, що найпряміші оцінки орієнтації не є значно більш точними. Такий градієнтний детектор країв розраховується в околиці 3×3 для 8 орієнтацій, як видно з рис. (13). Кожна з 8 згорткових масок обчислюється [40].

Оператор Превітта використовує ці рівняння як оператори Собеля, за винятком константи $c = 1$. Ось чому: слід зазначити, що на відміну від оператора Собеля, він не зосереджується на пікселях, які знаходяться ближче до центру масок. Оператор Превітта може вимірювати 2 елементи. Вертикальна складова краю обчислюється ядром G_x , а горизонтальна складова краю обчислюється ядром G_y . $|G_x| + |G_y|$ вказує інтенсивність градієнта в поточному пікселі [40]. Одне з ядер просто представляє інше, яке обертається

-1	0	+1
-1	0	+1
-1	0	+1

$G(x)$

+1	+1	+1
0	0	0
-1	-1	-1

$G(y)$

Рис. 3.3.4: Маски, які використовуються оператором Превітта [39].

Їх можна комбінувати для визначення абсолютної величини градієнта в кожній із точок і рівнянь. [39] :

$$|\nabla f| = \sqrt{G_x^2 + G_y^2} \quad (3.59)$$

$$|\nabla f| = |\sqrt{G_x^2}| + |\sqrt{G_y^2}| \quad (3.60)$$

Кут орієнтації країв (відносно сітки пікселя), що призводить до просторового градієнта, було представлено як:

$$\text{Angle of } \nabla f = \tan^{-1} \left(\frac{G_y}{G_x} \right) \quad (3.61)$$

Обчислення градієнта зображення базується на отриманні часткових похідних $\partial f / \partial x$ і $\partial f / \partial y$ в кожному місці пікселя. 3×3 , показані на рисунку (14), є рівнями сірого в околицях зображення.

Ці виведення можуть бути виконані для всього зображення з використанням масок 3×3 , які були проілюстровані на рис. (14) з процедурою згортання.

$$G_x = (z_7 + z_8 + z_9) - (z_1 + z_2 + z_3) \quad (3.62)$$

$$G_y = (z_3 + z_6 + z_9) - (z_1 + z_4 + z_7) \quad (3.63)$$

Зміна ковзання в цих двох рівняннях використовує вагу двох у центральному коефіцієнті. Значення ваги, рівне двом, використовується для досягнення рівня згладжування через надання вищого рівня значущості центральній точці. Рис. (13), який називається оператором Превітта, використовується для реалізації цих двох рівнянь.

Z1	Z2	Z3
Z4	Z5	Z6
Z7	Z8	Z9

Рис. 3.3.5: Область зображення 3×3 [39].

Сильна сторона оператора Превітта полягає в простоті, він визначає ребра та їхню орієнтацію, але його недоліком є неточність і чутливість до шуму [39].

Е) Метод Робертса

Оператор Roberts Cross виконує швидко для розрахунку просте двовимірне вимірювання просторового градієнта зображення. Ось чому він може підсвічувати області високої просторової частоти, які зазвичай відповідають краям. У найпоширенішому застосуванні оператор вводить зображення в градаціях сірого, як і вихід. Значення пікселів у кожній точці на виході є обчисленою абсолютною величиною просторового градієнта вхідного зображення в цій точці [40].

Перехресний оператор Робертса забезпечує просту апроксимацію величини градієнта:

$$G[f[i,j]] = |f[i,j] - f[i+1,j+1]| + |f[i+1,j] - f[i,j+1]| \quad (3.64)$$

Використовуючи згорткові маски, рівняння буде таким:

$$G[f[i,j]] = |G_x| + |G_y| \quad (3.65)$$

Де G_x & G_y можна обчислити за допомогою наведених нижче масок:

1	0
0	-1

G_x

0	-1
1	0

G_y

Рис. 3.3.6: Маски, які були використані оператором Roberts [41]

На рис. 3.3.7 показано вихідне зображення та його зміну за допомогою методів Собеля, Робертса, Превітта, ЛОГ та Кенні.

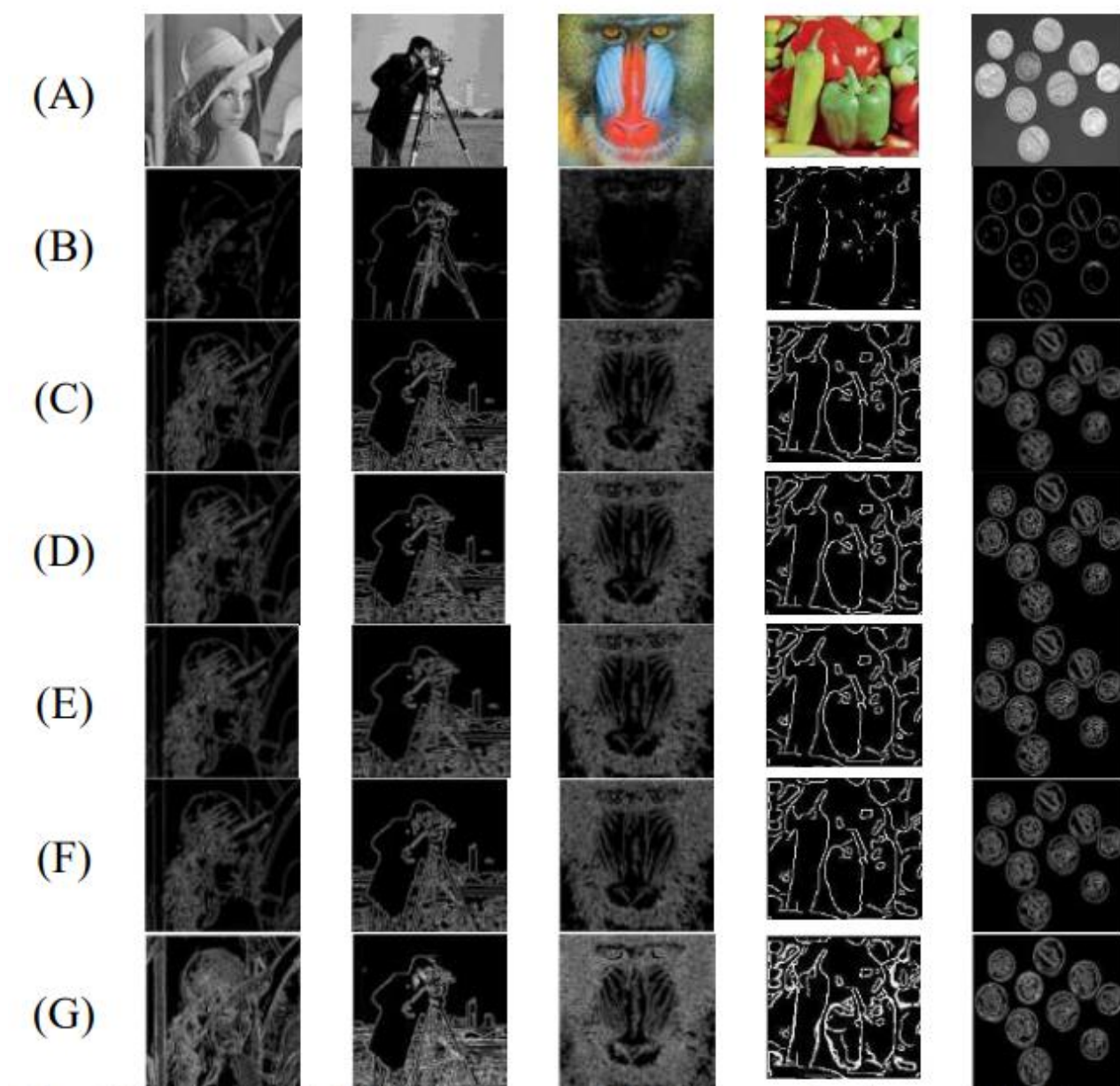


Рис. 3.3.7: (A) оригінальні зображення; (B): виявлені краї з використанням підходу Sobel; (C): виявлені краї з використанням підходу Превітта; (D): отримані ребра з використанням підходу Робертса; (E): отримані ребра з використанням

підходів LOG; (F): отримані ребра з використанням підходу Canny; (G): отримані ребра з використанням запропонованого в [36] підходу.

3.4 Методи сегментації на основі регіону

Регіональні підходи до сегментації – це підходи, які використовуються для сегментації зображень на різноманітні області, які мають подібні властивості [3]. У регіоні області сегментації були створені через асоціацію або дисоціацію сусідніх пікселів. Він працює згідно з принципом однорідності, враховуючи той факт, що сусідні пікселі в межах області мають подібні властивості та не схожі на пікселі в інших областях. Кожен із пікселів порівнюється з сусіднім пікселем для перевірки подібності, як-от колір, рівень сірого, форма та текстура. У разі позитивних результатів цей конкретний піксель додається до пікселя з метою збільшення області.

У випадку, коли повне зображення було представлено як область R , тоді для сегментації складіть його до n непересічних областей $S_1, S_2, S_3, \dots, S_n$ так, щоб

$$\begin{aligned} S_i &= S, & S_i \cap S_j &= \emptyset, & \text{if } i &\neq j \\ \text{Prop}(S_i) &= \text{True}, & f &= 1, 2, 3, \dots, n & & (3.66) \\ \text{Prop}(S_i \cup S_j) &= \text{False}, & f &= 1, 2, 3, \dots, n & & \end{aligned}$$

Де $\text{prop}(S_i)$ було охарактеризовано відповідно до значень ознак через область R . Ці області можуть бути непересічними, пов'язаними та однорідними [28]. Існують 2 основні підходи, які базуються на цьому підході: методи поділу регіонів і зростання регіонів і підходи злиття [3][28].

3.4.1 Методи збільшення регіону

У цьому підході пікселі в межах області позначаються характерною міткою, яка відрізняється від інших міток області. Цей підхід можна класифікувати додатково як SRG (тобто вирощування регіону з посівом) і UsRG (тобто вирощування регіону без посіву), останній є напівавтоматичним підходом, а UsRG — це повністю автоматизований підхід [28].

Підхід регіонального зростання можна визначити як простий регіональний підхід сегментації зображення [42]. Його також класифікували як піксельний підхід до сегментації зображення через те, що він включає вибір точок початкових початкових значень. Цей метод сегментації може досліджувати сусідні пікселі

початкових «точок насіння» та вказувати, чи потрібно додавати околиці пікселя до цієї області. Ця процедура повторюється подібно до загальних алгоритмів кластеризації даних. Основним недоліком визначення регіону на основі гістограми є те, що гістограми не надають жодних просторових знань (тільки розподіл рівнів сірого). Підходи до вирощування області зазвичай забезпечують дуже хороші процеси сегментації, що добре відповідають спостережуваним краям [43].

3.4.1.1 Об'єднання статистичних регіонів (SRM)

SRM можна визначити як метод, який використовується для сегментації зображень. Цей підхід використовується для оцінки цінностей у регіональному діапазоні та комбінується відповідно до критеріїв злиття, що призводить до скорочення списку. Деякі з корисних прикладів створюють набір поколінь в одній із популяцій або під час обробки зображення, об'єднуючи кілька сусідніх пікселів, які базуються на відтінках, що впадають у певне значення порогу [44].

3.4.1.2 Вирощування посівної області (SRG)

SRG запропоновано Р. Адамом [45]. Це напівавтоматичний підхід типу злиття [2][28]. SRG був відомий як швидкий і міцний, і він вільний від параметрів налаштування. Ця процедура починається з вибору вихідного пікселя на зображенні. У цьому підході вкрай необхідний відповідний вихідний параметр, оскільки він зосереджений на загальній якості сегментації.

Основні кроки SRG можна підсумувати нижче [28]:

- 1- Вибір пікселя зерна на зображенні з метою початку процесу сегментації.
- 2- Визначення критеріїв для розвитку регіону.
- 3- Включно з пікселем в області у випадку, якщо його вісім – підключено до мінімум 1 пікселя в області.
- 4- Кожна область позначається після тестування кожного пікселя для розподілу.
- 5- Об'єднання областей у випадку, якщо 2 окремим областям було призначено ідентичні мітки.

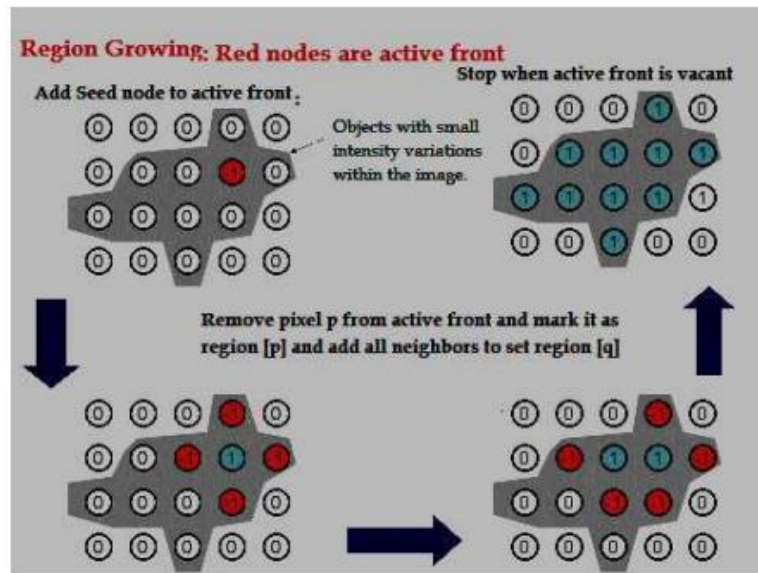


Рис. 3.4.1: .SRG [28]

3.4.1.3 Вирощування регіону без насіння

Цей підхід UsRG можна описати як повністю автоматизований підхід, він базується на подібності пікселя в області. Цей підхід є гнучким, він повністю автоматизований і залежить від параметрів налаштування [28]. Основні етапи методу UsRG можна підсумувати так:

- 1- Ініціалізація процесу сегментації з областю S_1 , яка містить 1 піксель, а потім створює області S_1, S_2, \dots, S_n після завершення.
- 2- Для виділення пікселя реалізовано вимірювання різниці тестових пікселів із середнім статистичним значенням.
- 3- Призначення пікселя для певної області, такої як S_i , у випадку, коли значення різниці менше певного порогового значення; у протилежному випадку цей піксель виділяється в нову область S_j .
- 4- Вищезазначені кроки повторюються для решти пікселів.

3.4.2 Методи поділу та злиття областей

Ці методи були представлені В. Penetal [45], вони працюють на основі квадродререва з фундаментальною метою розрізнення однорідності зображення, де все зображення можна розглядати як повну область, а потім це зображення ділиться до 4 різних квадрантних областей, які були засновані на конкретних заздалегідь визначених критеріях [28]. На малюнку (18) показано цей підхід.

Основні кроки цього підходу можна підсумувати як [28]:

1- Визначення умови однорідності.

2- Створення пірамідної структури даних для зображення.

3- Формування квадродерева з номерами для рівнів і формування номера фрагмента вузла.

4- Повторення цих кроків до моменту, коли більше не буде жодних можливих операцій злиття чи розділення.

ТАБЛИЦЯ 2

ПЕРЕВАГИ ТА НЕДОЛІКИ ТЕХНІКИ ВИРОЩУВАННЯ ОБЛАСТЕЙ З
СЕГМЕНТАЦІЄЮ ЗОБРАЖЕННЯ [42][43]

Переваги	Недоліки
<ul style="list-style-type: none"> - Ці підходи мають здатність правильно відокремлювати області з ідентичними характеристиками. - Ці підходи мають здатність надавати оригінальні зображення з очевидними краями з достатніми результатами сегментації. - Ідея полягає в тому, що необхідна обмежена кількість вихідних точок для представлення бажаних характеристик, після чого регіон буде розширено. 	<ul style="list-style-type: none"> - Ці методи дорогі з точки зору обчислень. - Це локальні підходи без будь-яких глобальних поглядів на проблему. - Чутливий до шуму. - Якщо до зображення не було застосовано порогову функцію, може існувати постійний шлях точок, пов'язаних із кольором, що призводить до з'єднання будь-яких 2 точок зображення.

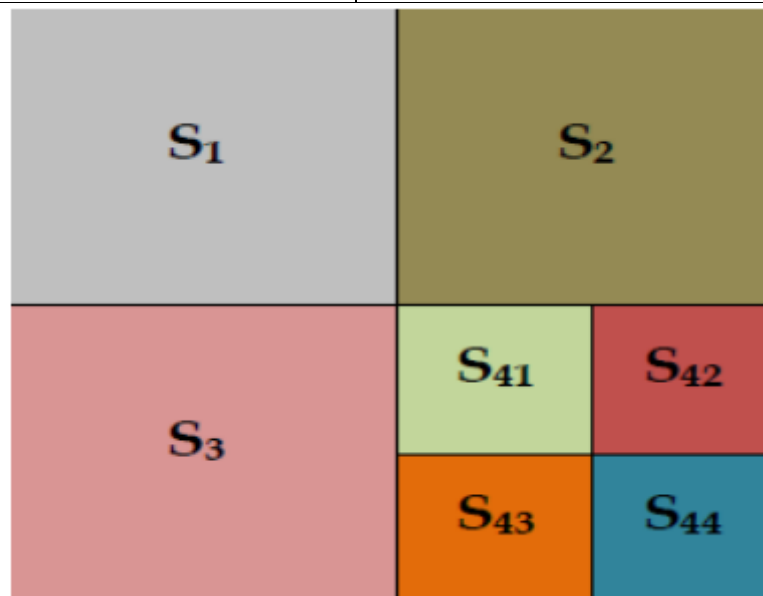


Рис. 3.4.2: Підхід методу поділу та злиття регіонів [28]

3.5 Методи кластеризації

Підходи на основі кластерів можна визначити як методи, які використовуються для сегментації зображення на кластери, які мають пікселі з однаковими властивостями. Кластеризація використовується для поділу елементів даних на кластери таким чином, щоб ті самі елементи кластера були більш схожими один на одного порівняно з іншими. Існує 2 основні класи методів кластеризації: на основі розділів та ієрархічних методів. Останній тип базується на концепції дерева. Підходи на основі розділення використовують підходи оптимізації в ітераційний спосіб для мінімізації цільової функції. Окрім цих двох підходів, існує безліч алгоритмів для кластеризації. Кластеризацію можна класифікувати за 2 основними типами [3][46-47].

3.5.1 Жорстка кластеризація

Ці типи можна визначити як простий метод кластеризації, який працює шляхом поділу зображення на групу кластерів таким чином, що 1 піксель може бути лише частиною одного окремого кластера, тобто можна вважати, що кожен з пікселів можуть входити лише до 1 кластера. Ці підходи використовують функції членства, які мають значення 0 або 1.

Іншими словами, один певний піксель може бути частиною певного кластера або іншим чином. Одним із прикладів підходу на основі жорсткої кластеризації є підхід на основі кластеризації k -середніх, який називають K-M. [3].

Кластеризація K-Means

K-Means — це дуже простий неконтрольований алгоритм навчання, який має здатність вирішувати типові проблеми кластеризації. K-Means — хороша техніка, яка використовується для сегментації зображення. У цьому підході спочатку обчислюються центри, а потім кожен піксель надається найближчому центру. Він робить наголос на максимізації внутрішньокластерної подібності, а також на мінімізації міжкластерної рівності.

Процес працює за допомогою легкого та простого підходу до класифікації певного набору даних за певною кількістю кластерів (припускаючи k кількість кластерів), фіксованих апріорі. Основна концепція полягає у визначенні k центроїдів, 1 для кожного з кластерів. Ці центроїди повинні бути розподілені в хитрий спосіб

через те, що різні місця викликають різні результати. Тому оптимальний варіант – розмістити їх якомога далі один від одного. Наступним кроком буде взято кожен з точок, які належать до вказаного набору даних, і пов'язати їх із найближчим із центроїдів. У випадку, коли жодна з точок не розглядається, виконується початковий крок і буде виконано раннє групування. Тут необхідно перерахувати k нових значень центроїда кластера, які є результатом попереднього кроку. Після обчислення цих k нових значень центроїда потрібно виконати новий зв'язок між ідентичними точками набору даних і новим найближчим центроїдом. Вийшла петля. Завдяки цій петлі можна помітити, що ці k центроїдів поступово змінюють свої позиції до точки, коли більше не відбувається жодних змін, тобто центроїди більше не рухаються. Зрештою, цей підхід має на меті мінімізувати цільову функцію; у даному випадку функція сумарної квадратичної помилки. Цільова функція буде представлена у вигляді:

$$J = \sum_{j=1}^k \sum_{i=1}^n |x_i^{(j)} - c_j|^2 \quad (3.67)$$

$\|x_i^{(j)} - c_j\|^2$ представляє визначену міру відстані між точкою даних $x_i^{(j)}$ і центром кластера C_j , вказує відстані n точок даних до їхніх відповідних центрів кластера. Цей підхід включає наступні кроки [47]:

Загальні кроки в алгоритмі К-середнього:

1- Розміщення K точок у просторі, який представлений кластерними об'єктами.

Ці точки є центроїдами початкової групи.

2- Призначення кожного з об'єктів групі, яка включає найближчий центроїд.

3- Після призначення всіх об'єктів положення K центроїдів буде перераховано.

4- Крок 2 і Крок 3 повторюються до моменту, коли центроїди більше не змінюються. Це призводить до поділу об'єктів на групи, а метрика буде мінімізована та обчислена на основі цього.

3.5.2 М'яка кластеризація

Цей тип кластеризації можна вважати найбільш природним утворенням через те, що в реальному житті точний розподіл неможливий через наявність шуму. Ось чому методи м'якої кластеризації є найбільш вигідними для сегментації зображень, де поділ не є строгим. Одним із прикладів цього методу є підхід кластеризації на основі нечітких середніх [3].

Кластеризація нечітких С-середніх (FCM)

У цьому методі пікселі поділяються на кластери відповідно до часткової приналежності, іншими словами, один з пікселів може бути частиною кількох кластерів і такий ступінь приналежності характеризується значеннями приналежності. Цей підхід має вищу гнучкість порівняно з іншими підходами [46].

FCM виконує розділення групи з n об'єктів $x = \{x_1, x_2, \dots, x_n\}$ у R^d -вимірному просторі на c ($1 < c < n$) нечітких кластерів з $u = \{u_1, u_2, u_3, \dots, u_c\}$ центроїди або центри кластерів [48]. FCM об'єктів можна ідентифікувати за допомогою нечіткої матриці μ , яка має c стовпців і n рядків, де n являє собою кількість об'єктів даних, а c являє собою кількість кластерів. μ_{ij} представляє елемент, який знаходиться в i -му рядку, а j -й стовпець, μ_{ij} означає рівень належності або функції асоціації i -го об'єкта з j -м кластером. Цільовою функцією алгоритму FCM є мінімізація наведеного нижче рівняння [46].

$$J_m = \sum_{j=1}^c \sum_{i=1}^n u_{ij}^m d_{ij} \quad (3.68)$$

Де

$$d_{ij} = \|x_i - y_j\| \quad (3.69)$$

y_j представляє центроїд j -го кластера, який можна обчислити за допомогою рівняння нижче:

$$y_j = \frac{\sum_{i=1}^n u_{ij}^m x_i}{\sum_{i=1}^n u_{ij}^m} \quad (3.70)$$

Алгоритм FCM є одним із ітераційних алгоритмів і може бути представлений такими кроками [46]:

- 1- Виберіть m ($m > 1$); ініціалізація значень функції належності μ_{ij} , $i = 1, 2, \dots, n$; $j = 1, 2, \dots, c$.
- 2- Обчислення центрів кластерів y_j , $j = 1, 2, \dots, c$ на основі рівняння 70
- 3- Обчислення евклідової відстані d_{ij} , $i = 1, 2, \dots, n$

ТАБЛИЦЯ 3
РЕЗЮМЕ ПЕРЕВАГ ТА НЕДОЛІКІВ МЕТОДІВ СЕГМЕНТАЦІЇ ЗОБРАЖЕННЯ
МЕТОДИ КЛАСТЕРИЗУВАННЯ [3]

Переваги	Недоліки
Він використовує часткове членство, яке, таким чином, більш вигідно для вирішення реальних завдань.	Визначення функції членства є складним процесом.

3.6 Методи глибокої нейронної мережі

3.6.1 Нейронна мережа згортки

CNN були визначені як одна з найбільш вигідних і часто використовуваних архітектур у спільноті глибокого навчання, особливо для завдань, пов'язаних із комп'ютерним зором. CNN були спочатку запропоновані Фукусімою в одній із його основоположних робіт про «Неокогнітрон» [49], відповідно до моделі ієрархічного сприйнятливої поля зорової кори.

Згорточна нейронна мережа (CNN) може бути використана для вивчення способів сегментації зображень. CNN виконують виділення характеристик прямим шляхом із піксельних зображень, які мають мінімальну попередню обробку.

Типова CNN зазвичай включає 3 типи шарів: а) шари згортки, в яких фільтр (або ядро) вагових коефіцієнтів піддається згортці з метою виділення ознак; б) нелінійні шари із застосуванням функції активації на картах ознак (зазвичай поелементно) з метою надання можливості моделювати нелінійні функції мережами; с) об'єднання шарів, які замінюють невелику околицю карти об'єктів статистичною інформацією (максимальне, середнє та інше) навколо цієї околиці та зменшують просторову роздільну здатність. Блоки рівня з'єднані локально; це означає, що кожна одиниця може отримувати зважені вхідні значення з невеликого оточення, яке називається сприйнятливим полем, одиниць на попередньому рівні. Завдяки накопиченню шарів для формування пірамід із кількома роздільними здатностями шари вищих рівнів вивчають особливості зі значно ширших сприйнятливих полів. Фундаментальною обчислювальною сильною стороною CNN є той факт, що кожне

сприйнятливим полем на рівні має спільну вагу з іншими, що призводить до значно меншої кількості параметрів порівняно з повністю зв'язаними мережевими мережами. [50].

CNN можна охарактеризувати як орієнтований граф. Вузли відповідають пікселям зображення. Краї відповідають фільтрам. CNN — це багатошаровий перцептрон (MLP), спеціально розроблений для розпізнавання двовимірних форм, які мають високий рівень інваріантності до трансляцій, перекосів, масштабування або будь-яких інших форм спотворення. Архітектура показана на рис. (19) [51].

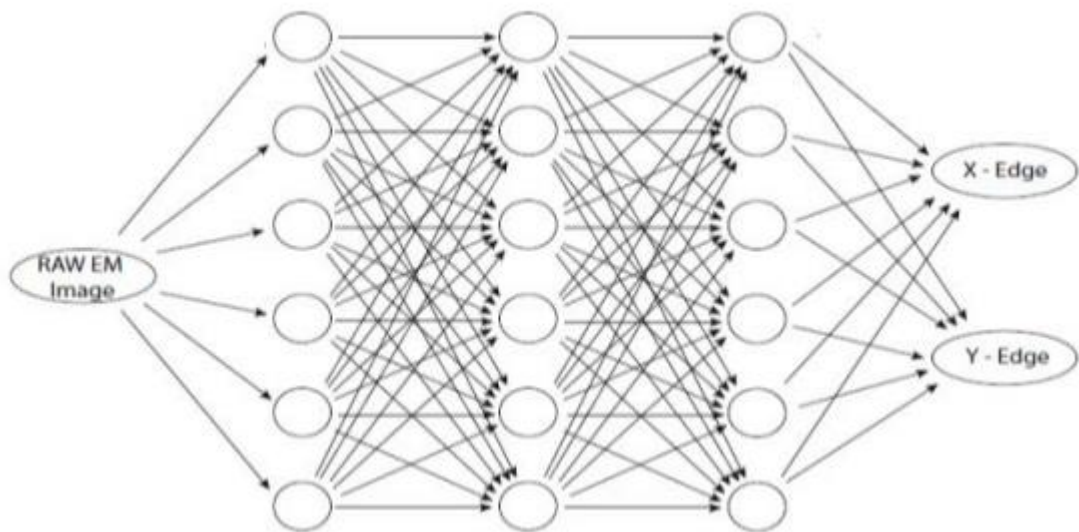


Рис. 3.6.1: Згортова нейронна мережа [51]

Використання розподілу ваги може дозволити паралельну реалізацію CNN. Зменшити кількість вільних параметрів можна за допомогою процесу розподілу ваги. Як наслідок, здатність до машинного навчання знижується, що, як наслідок, призводить до покращення здатності машини до узагальнення. Налаштування вільних параметрів мережі здійснюється за допомогою стохастичного режиму зворотного поширення. Навчання CNN має подвійні переваги. Маючи попередню інформацію про зображення, він має здатність вивчати складне, нелінійне, багатовимірне відображення. На додаток до цього, він має здатність вивчати рівні зміщення та синтаксичні значення ваги [51].

3.6.2 Повторювані нейронні мережі (RNN)

RNN зазвичай використовуються для обробки послідовних даних, таких як текст, мова, часові ряди та відео, у яких дані в певному місці/часі залежать від даних,

які зустрічалися раніше [50]. У кожній із позначок часу модель може збирати вхідні дані з поточного часу X_i і прихованого стану від попереднього процесу, h_{i-1} , і виводити новий прихований стан і необхідне значення, рис. (20).

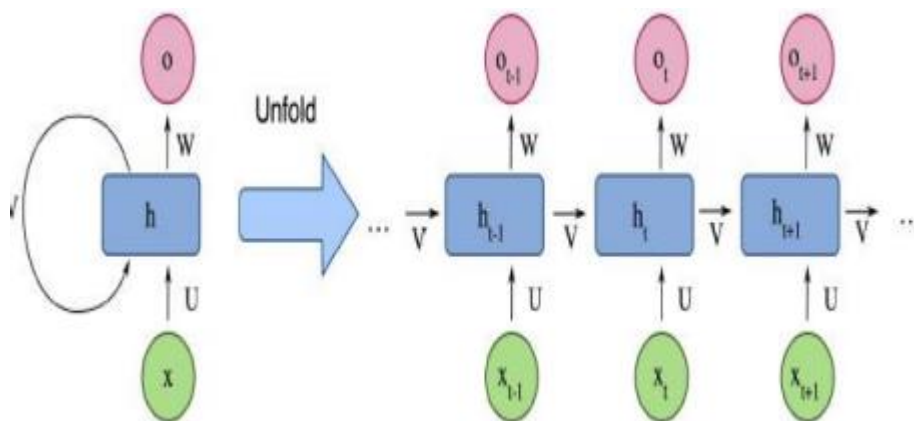


Рис. 3.6.2: Проста архітектура RNN [50]

RNN, як правило, проблематичні у випадку довгих послідовностей через те, що вони не здатні охопити довгострокові рівні залежності в широкому діапазоні реальних реалізацій (хоча вони не демонструють жодних теоретичних обмежень у цьому питанні), і вони зазвичай страждають від проблем вибуху або зникнення градієнтів. Тим не менш, тип RNN, який називають довгостроковою пам'яттю (LSTM), був змодельований для уникнення цих проблем. Архітектура LSTM, яка була проілюстрована на рис. (21), містить 3 вентиля (які є входом, пропуском і вихідним вентилям), які відповідають за регулювання потоку інформації до комірки пам'яті та з неї. для зберігання значень протягом випадкових інтервалів часу.

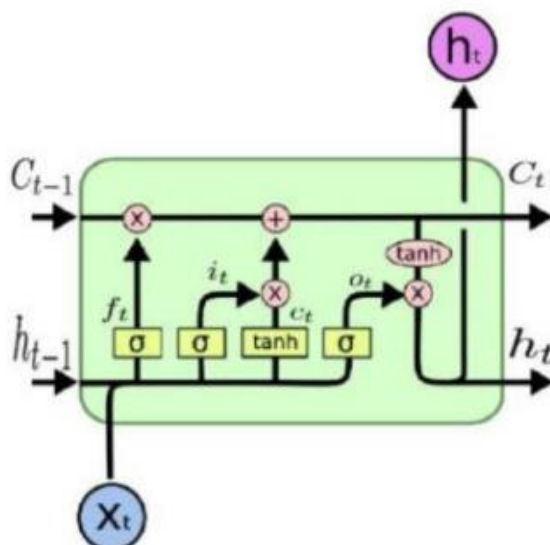


Рис. 3.6.3: Типова архітектура модуля LSTM [50].

Кореляція між входом, прихованими станами та різними воротами може бути представлена як:

$$\begin{aligned}
 f_t &= \sigma (w^{(f)xt} + U^{(f)} + b^{(f)}), \\
 i_t &= \sigma (w^{(i)xt} + U^{(f)ht-1} + b^{(i)}), \\
 o_t &= \sigma (w^{(o)xt} + U^{(o)ht-1} + b^{(o)}), \\
 c_t &= f_t \odot c_{t-1} + i_t \odot \tanh (w^{(c)} xt + U^{(c)} h_{t-1} + b^{(c)}) \\
 h_t &= o_t \odot \tanh(c_t)
 \end{aligned}
 \tag{3.71}$$

де $x_t \in \mathbb{R}^d$ представляє вхідні дані на часовому кроці t , а d представляє розмірність ознаки для кожного зі слів, σ представляє поелементну сигмоїдну функцію (для відображення значень у діапазоні $[0, 1]$), \odot представляє поелементний продукт, а c_t представляє комірку пам'яті, яка була розроблена з метою зниження ризиків вибуху/зникнення градієнта (як наслідок, надаючи можливість вивчати залежності протягом більших періодів часу, які можливі із звичайними RNN). Ворота забуття, f_t , призначені для скидання комірки пам'яті. o_t і він відповідно представляють вихідні та вхідні вентиля, і фактично регулюють вихід та вхід комірки пам'яті [50].

3.6.3 Моделі кодера-декодера та автокодера

Моделі кодера-декодера можна визначити як групу моделей, які вивчають відображення точок даних із домену вхідних даних у домен вихідних даних за допомогою 2-ступеневої мережі: кодера, який характеризується функцією кодування $z = f(x)$, призводить до стиснення вхідних даних до представлення латентного простору; декодер, $y = g(z)$, має на меті передбачити значення виходу з представлення прихованого простору. У цьому випадку приховане представлення фактично вказує на представлення ознак (тобто векторів), яке має здатність фіксувати основну семантичну інформацію вхідних даних, що було корисним для прогнозування результату. Ці моделі набули надзвичайної популярності в проблемах перекладу «від зображення до зображення» на додаток до моделей послідовності НЛП. На рис. (22) зображено схему простої моделі кодера-декодера. Ці моделі, як правило, навчаються через мінімізацію втрат реконструкції $L(y, \hat{y})$, вимірюючи значення різниці між виходом у базовій правді та послідовної реконструкції \hat{y} . у цьому випадку вихідні дані можуть стосуватися покращеної версії зображення (наприклад, у суперроздільній здатності або усуненні розмиття зображення) або карти сегментації [50].

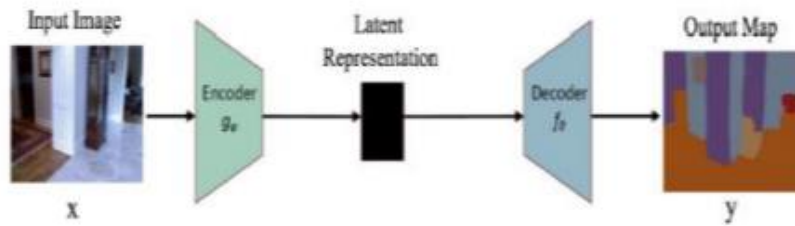


Рис.3.6.4: Проста архітектура моделі кодера-декодера [50]

Автоматичні кодери можна визначити як особливий випадок моделі кодера-декодера, де вихід є таким же, як і вхід. Було запропоновано численні варіації автокодувальника. Одним із найпоширеніших є SDAE (тобто стековий автоматичний кодер із зменшенням шуму), який об'єднує численні автоматичні кодери та використовує їх для цілей усунення шумів зображення. Іншим поширеним типом є VAE (тобто варіаційний автокодер), який накладає більш ранній розподіл на латентні представлення. VAE мають здатність генерувати реалістичні вибірки з певного розподілу даних. Іншим типом є конкурентні автокодери, які вводять змагальні втрати на приховане представлення з метою заохочення їх до наближення попереднього розподілу [50].

3.6.4 Машини опорних векторів(SVM)

SVM мають дуже важливий вплив на підходи до сегментації зображень, зокрема, при обробці медичних зображень. Існують численні модифіковані підходи, які базуються на SVM, наприклад модифіковані підходи SVM у сегментації медичних зображень. Існує багато модифікованих підходів SVM, які були представлені протягом останніх 5 років, зокрема модифіковані SVM у сегментації медичних зображень [52]. У цьому огляді ми покажемо, як використані модифіковані підходи SVM були здатні досягти хороших ефектів на методи сегментації зображення або порогової сегментації медичного зображення SVM. Ми представимо огляд модифікованих підходів SVM за останнє десятиліття, таких як SVM за методом найменших квадратів (LS-SVM), а також генетичний алгоритм-SVM (GA-SVM).

- Векторна машина найменших квадратів (LS-SVM):

Янг [53] запропонував достатній метод сегментації кольорового зображення, який ґрунтується на класифікації пікселів за допомогою LS-SVM. Характеристики кольору на рівні пікселів, а також функції текстури використовувалися як вхідні дані моделі LS-SVM (тобто класифікатора), а класифікатор LS-SVM було навчено шляхом відбору навчальних зразків за допомогою методу порогового значення ентропії

Арімото. LS-SVM — це підхід, який використовується для вдосконалення SVM шляхом модифікації його функції або параметрів. Він не змінює концепції та теорію SVM. Він має здатність покращувати точність сегментації. Тим не менш, недоліком цього підходу є те, що він не може досягти точної сегментації. LS-SVM — це підхід SVM, який був змінений шляхом зміни теорії SVM [53].

- Векторна машина підтримки генетичного алгоритму (GA-SVM): Чжан [54] досліджував GA-SVM для визначення гіперпараметрів SVM (c , g). Ці модифіковані підходи методів SVM мають на меті отримати достатню сегментацію зображення, шукаючи параметри для визначення гіперпараметрів SVM (c , g) з метою досягнення автоматизованого навчання відбору даних, а також автоматизованого вилучення особливостей. GA-SVM є комбінацією двох методів, які: GA та SVM. Підхід до поєднання 2 методів призводить до вдосконалення теорії та принципу SVM. Його перевага полягає в тому, що він нечутливий до шуму, що призводить до того, що він не має можливості точно сегментувати зображення [52]. GA-SVM є одним із оновлених методів SVM, який базується на комбінаціях із генетичними алгоритмами. Такий підхід SVM модифікується з використанням інших підходів [54].

3.7 Гібридні методи

Загалом гібридні підходи можна визначити як комбінацію 2 або більше обчислювальних підходів, що забезпечує більше переваг порівняно з окремими підходами та покращує аналіз даних. Гібридний метод сегментації виконує інтеграцію підходів сегментації на основі регіону та меж, підвищуючи силу та одночасно зменшуючи слабкі сторони двох методів. Ми покажемо деякі гібридні методи, як-от еволюційні підходи, нечітка логіка та інтелектуальний рій, які використовуються для сегментації.

3.7.1 Еволюційні підходи

Еволюційні алгоритми використовуються в широкому діапазоні інженерних застосувань для оптимізації проблем, які зазвичай важко вирішити за допомогою традиційних підходів. Однією з таких проблем є сегментація зображення. Це завдання використовується для виділення об'єктів (контурів) із зображень з метою створення розумного представлення зображення [55]. Цей тип обчислень є частиною біоінспірованих обчислень та області штучного інтелекту, що забезпечує підходи, які використовуються для оптимізації проблем безперервного простору [55]. Це один із придатних інструментів для гнучкої та ефективної сегментації часових рядів

3.7.2 Нечітка логіка

Загалом нечітку логіку можна визначити як один із підходів до обчислень, заснований на «ступенях істинності» замість звичайних «істина чи хибність» (тобто 1 або 0), які називають булевою логікою, яка сучасна комп'ютери були засновані на. Може бути корисно розглядати нечітку логіку як один із способів міркувати, що насправді працює, а булева або двійкова логіка є просто окремим випадком цього.

Теорія нечіткої логіки (FL) все частіше використовується для вирішення проблем обробки зображень, таких як проблеми сегментації. У [57] розроблено програмну систему з використанням комбінації методів нечіткої та нейронної мережі для реалізації алгоритму сегментації зображення. Таким чином, теорію нечіткої логіки з методами нейронної мережі можна об'єднати та створити гібридні методи, які використовуються для сегментації.

З іншого боку, нечітка логіка з генетичними алгоритмами (GA) має значний вплив на вирішення численних проблем в обробці зображень і розпізнаванні образів. У [58] було запропоновано гібридний підхід FL і GA з метою використання для сегментації та виділення характеристик із зображень інтенсивності та діапазону.

3.7.3 Ройовий інтелект (SI)

Ройовий інтелект можна визначити як колективну поведінку самоорганізованих, децентралізованих систем, штучних або природних. На сьогоднішній день підходи на основі SI використовуються в багатьох додатках і демонструють високу ефективність [59]. У [60] представлено два інтелектуальні підходи для визначення порогів зображення: метод штучної бджолиної колонії та метод PSO, які були використані для вирішення цього питання.

А. Підхід штучної бджолиної колонії (ABC).

Алгоритм ABC включає 3 фундаментальні елементи, а саме: кількість нектару, позицію джерела їжі та кілька класів медоносних бджіл [60-61]. У [62] ABC використовується з MRF (випадковими полями Маркова) для цілей отримання найкращої сегментації для сегментації зображення мозку. ABC — це стійкість, особливо в дискретних задачах оптимізації з кількома змінними. Таким чином, це використовується з MRF з метою покращення якості сегментації зображення [62].

Дослідження довели, що APPS добре підходить для сегментації зображень [60]. У [63] запропоновано алгоритм оптимізації (ABC), який використовувався як один із методів кластеризації для сегментації печінки на КТ-зображеннях.

В. Підхід до оптимізації роїв частинок (PSO).

PSO — це один із найпотужніших підходів оптимізації, натхненний соціальною поведінкою тварин, які живуть або пересуваються зграєю, як-от зграя риби чи зграя птахів. У [64] POS використовувався для сегментації медичних зображень і вирішення проблем, що виникають у разі сегментації медичних зображень.

PSO є одним із нових класів метаевристик, які були представлені Кеннеді та Еберхартом у 1995 році [65]. Цей підхід вивчався кількома іншими дослідженнями [66-70]. На цей алгоритм надихнула соціальна поведінка тварин, які пересуваються зграями як зграйні риби або зграйні птахи. ККД рою вище в порівнянні з сумою ККД його частин.

Цей підхід до оптимізації базується на співпраці між окремими особами. Особина рою знає лише швидкість і положення найближчих сусідів. Кожна з частинок коригує свою поведінку відповідно до свого досвіду, а також досвіду своїх сусідів для побудови вирішення проблеми. Використовуючи прості правила переміщення (у просторі рішення), частинки можуть поступово зближуватися до вирішення проблеми.

Формально кожна з частинок i має розташування $x_i(t)$ у момент часу t у можливому просторі розв'язків, що змінюється в момент часу $t+1$ на швидкість $v_i(t)$, на яку впливає оптимальне положення $y_i(t)$, яке відвідується сама по собі (тобто власний досвід), а z_i t представляє оптимальне положення, яке відвідує кожна частинка (яке можна назвати глобальним оптимальним). Позиції були виміряні за допомогою функції відповідності f , яка залежить від задачі оптимізації, а K представляє розмірність простору.

$$x_i(t)=(x_{i1},x_{i2},\dots, x_{ij}, \dots, x_{ik}) \quad (3.72)$$

$$v_i(t)=(v_{i1},v_{i2},\dots, v_{ij}, \dots, v_{ik}) \quad (3.73)$$

$$y_i(t)=(y_{i1},y_{i2},\dots, y_{ij}, \dots, y_{ik}) \quad (3.74)$$

$$z_i(t)=(z_{i1},z_{i2},\dots, z_{ij}, \dots, z_{ik}) \quad (3.75)$$

y_i — оновлений час на основі такого рівняння:

$$y_i(t + 1) = \begin{cases} y_i(t) & \text{if } f(x_i(t+1)) \geq f(y_i(t)) \\ x_i(t+1) & \text{if } f(x_i(t+1)) < f(y_i(t)) \end{cases} \quad (3.76)$$

Найкраще положення, яке відвідала кожна частинка до часу t , z t , буде обчислено в момент часу t за допомогою рівняння нижче:

$$Z(t) \in (y_1(t), y_2(t), \dots, y_s(t) = \min\{f(y_1(t)), f(y_2(t)), \dots, f(y_s(t))\} \quad (3.77)$$

де s представляє кількість частинок (розмір рою). Швидкість v_i $t = (v_{i1}(t), v_{i2}(t), \dots, v_{ij}(t), \dots, v_{iK}(t))$ частинку i в момент часу t оновлюється з використанням наведеного нижче рівняння:

$$v_{ij} t+1 = w * v_{ij} t + c1 * r1j * y_{ij} t - x_{ij} t + c2 * r2j * z_j t - x_{ij} t \quad (3.78)$$

w називають інерційною вагою; $c1$ & $c2$ представляють константи прискорення; $r1j$ і $r2j$ представляють випадкові змінні в іклах від 0 до 1; швидкість v_{ij} залежить від значення V_{\max} для забезпечення конвергенції. Розташування x_i частинки i оновлюється за допомогою наведеного нижче рівняння:

$$X_i(t+1) = x_i(t) + v_i(t+1) \quad (3.79)$$

На рис. (3.7.1) показано МРТ сканування та зображення сегментації за допомогою POS.

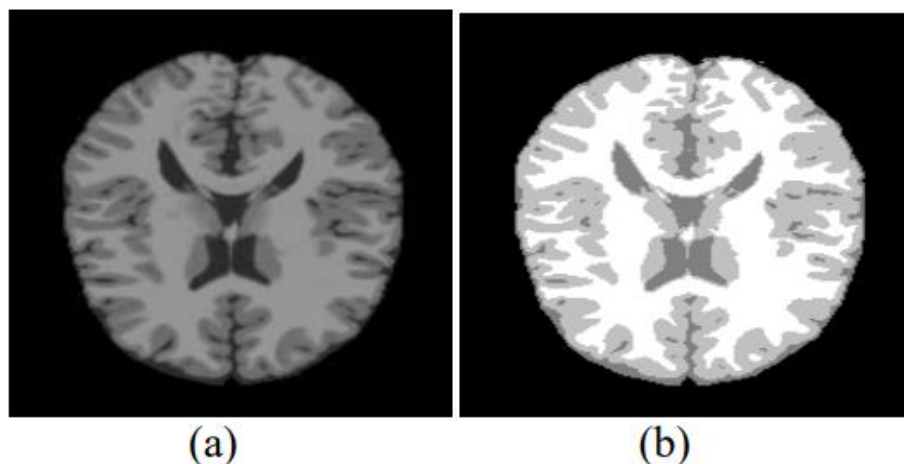


Рис. 3.7.1: (а) МРТ (б) PSO

РОЗДІЛ 4

Охорона праці

4.1 Вступ

В контексті розвитку інтелектуальних систем обробки зображень застосування комп'ютерного зору стає невід'ємною складовою для різноманітних галузей, починаючи від промисловості і закінчуючи медициною. Спеціалізовані системи, які використовують комп'ютерний зір, виявляються важливим інструментом у сучасному світі, забезпечуючи широкий спектр застосувань, від автоматизації виробничих процесів до покращення систем діагностики та моніторингу.

Метою даної дипломної роботи є створення системи оптимального вибору методів комп'ютерного зору для обробки зображень. У фокусі дослідження знаходиться не лише вдосконалення самого алгоритму обробки, але й вивчення аспектів охорони праці під час розробки та застосування таких систем. Забезпечення безпеки та ефективності використання інтелектуальних технологій є важливою умовою для їх успішного впровадження у різні сфери.

У цьому розділі ми розглянемо ключові аспекти охорони праці, пов'язані із застосуванням методів комп'ютерного зору для обробки зображень. Аналіз ризиків, визначення заходів безпеки та впровадження сучасних стандартів стануть невід'ємною частиною розгляду, спрямованого на створення ефективних та безпечних інтелектуальних систем.

4.2 Аналіз умов праці

Приміщення, в яких планується установка та подальша робота з комп'ютером, повинні відповідати проектній документації будинку, погодженій з уповноваженими державними органами. Крім того, роботодавець повинен враховувати санітарні нормативи освітлення, вимоги до параметрів мікроклімату (температура, відносна вологість), ступеня і сили вібрації, звукового шуму і вогнестійкості приміщення, а також характеристики електромагнітного, ультрафіолетового та інфрачервоного полів. Конкретні показники зазначених санітарних норм див. в Державних санітарних правилах і нормах роботи з візуальними дисплейними терміналами електронно-

обчислювальних машин ДСанПіН 3.3.2.007-98, затверджених Постановою Головного державного санітарного лікаря України №7 від 10 грудня 1998 року. Правила поширюються на умови й організацію праці при роботі з візуальними дисплейними терміналами (ВДТ) усіх типів вітчизняного та зарубіжного виробництва на основі електронно-променевої трубки (ЕПТ), що використовуються в електронно-обчислювальних машинах (ЕОМ) колективного використання та персональних ЕОМ (ПЕОМ). Так, наприклад, роботодавцю заборонено встановлювати комп'ютери в приміщеннях, розташованих у підвалах будинків. Для уникнення можливих аварій та замикань, поряд з приміщеннями, де вестиметься робота з комп'ютером (над чи під ними), також не дозволяється проведення робіт, що потребують здійснення надмірно вологих технологічних процесів. Відповідне приміщення повинно бути укомплектоване системами центрального або індивідуального опалення, кондиціонування чи вентиляції повітря. Але при установці зазначених систем, необхідно переконатись, що батареї опалення, водопровідні труби, вентиляційні кабелі тощо, надійно сховані під захисними щитками, які перешкоджатимуть можливому потраплянню робітника під напругу.

У кожній кімнаті, де обладнуватимуться робочі місця співробітників, що працюватимуть на комп'ютері, повинні бути наявні елементи природного та штучного освітлення. При цьому, на вікнах слід встановити легко регульовані жалюзі чи штори, які дозволять працівникам коригувати рівень освітлення в приміщенні. Бажано розмістити комп'ютери в кімнаті таким чином, щоб світло потрапляло на екрани моніторів з півдня чи північного сходу. З метою досягнення максимального рівня безпеки і охорони праці при роботі з комп'ютером, виробничі приміщення необхідно обладнати аптечками першої медичної допомоги, системами автоматичної пожежної сигналізації і вогнегасниками. В приміщенні, в якому разом працюють 5 або більше комп'ютерів, на видимому місці встановлюється службовий вимикач, який у разі потреби дозволить повністю відключити електричне живлення кімнати.

Вимоги до особистого робочого місця

Робоче місце має розміри 4,5x2,5x3 метри (довжина, ширина, висота), площа 11,25 м² та об'єм приміщення 33,75 м³. Відповідно з ДСанПіН 3.3.2.007-98 на одне

робоче місце з використанням ПК повинні задовольнятися такі норми: – площа – не менше 6,0 кв. м; – об'єм – не менше 20,0 куб. м. При необхідності, суміжні робочі місця співробітників, що працюють з комп'ютером, слід розділити перегородками висотою до 2 метрів. При визначенні достатнього розміру приміщення і робочого місця на одну особу необхідно додатково враховувати шафи, сейфи, тумби або інші предмети меблів чи обладнання, які знаходяться в кімнаті. На столі працівника можливо розмістити допоміжні для роботи пристрої (принтери, колонки, сканери), а також місця для зберігання документів, за умови, що це не обмежуватиме видимість екрану і не заважатиме працівнику. У разі надмірного шуму чи вібрації технічного обладнання, роботодавець повинен забезпечити працівників антивібраційними килимками. Робочий стілець співробітника має бути підйомно-поворотним, легко регульованим за висотою та забезпечувати належну підтримку та зручне положення спини і хребта особи. Щодня необхідно проводити вологе прибирання приміщення, та очищати робоче місце та безпосередньо монітор комп'ютера від запиленості. На підприємстві забороняється: проводити ремонт та технічне обслуговування комп'ютера за робочим місцем працівника; самочинно ремонтувати або намагатись здійснити технічне налагодження комп'ютера без залучення компетентних спеціалістів; складувати на робочому місці зайві документи, деталі та предмети, що не потрібні для роботи; використовувати монітори з нечітким зображенням та монітори, у яких наявні поламки екрану; працювати з матричним принтером без антивібраційного покриття та зі знятою кришкою. Допускати до роботи осіб, які не пройшли затверджений на підприємстві курс охорони праці для роботи з комп'ютером, не дозволяється.

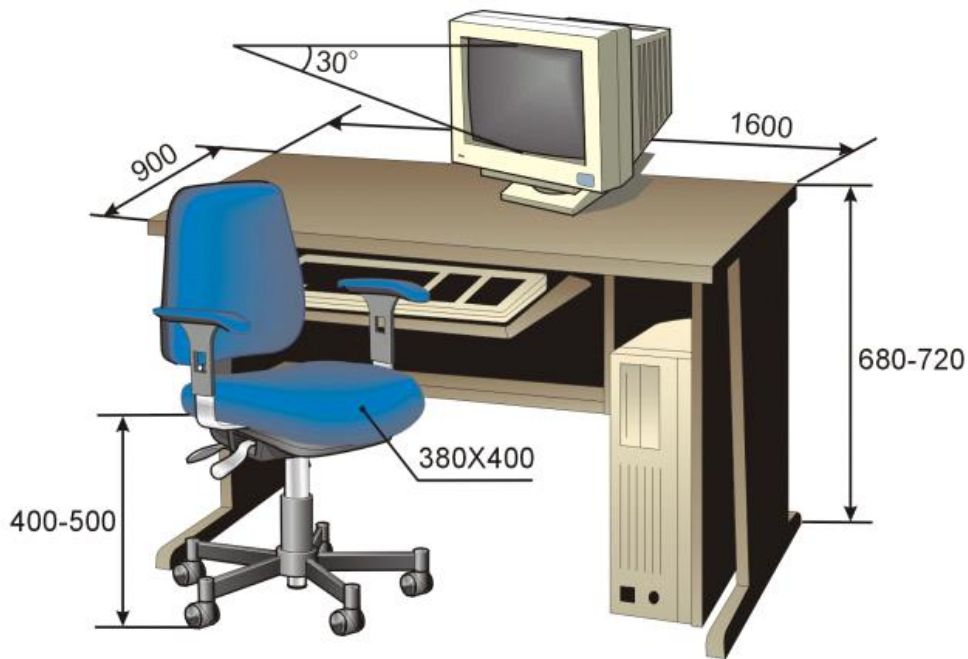


Рис 4.2.1 Вимоги до робочого місця

При прийнятті на роботу кожна особа має пройти лікарський огляд. Окрім того, при подальшій трудовій діяльності в компанії, така особа підлягає регулярному лікарському огляду не рідше ніж раз на 2 роки. Обов'язковим є проходження таких лікарів як терапевта, невропатолога та офтальмолога. В компанії мають бути чітко встановлені перерви для відпочинку працівників (окрім обідньої), як правило, тривалістю 10-15 хвилин раз на годину або дві, в залежності від складності роботи. В будь-якому випадку, роботодавець повинен передбачити такий розпорядок роботи на підприємстві, щоб час неперервної роботи з комп'ютером був не більше ніж 4 години. Додатково, для збереження належного рівня здоров'я та професійної придатності робітників, рекомендується виділити на підприємстві окреме побутове приміщення для перепочинку працівників і зняття ними нервово-емоційного напруження, що виникає при роботі з комп'ютером.

4.3 Розробка заходів з охорони праці

Охорона праці на робочому місці є важливою складовою для забезпечення безпеки працівників та досягнення високих результатів у дослідженні. Нижче наведено заходи з охорони праці, які повинні бути реалізовані в даному приміщенні:

Таблиця 4.1 – Заходи та засоби для попередження та захисту від електротравм

№	Категорії заходів для захисту	Вид заходу	Критерій вибору
1	Технічні	<ul style="list-style-type: none"> - встановлення захисних заземлень; - застосування захисного розділення електромереж; - підтримка сухого, незапиленого приміщення з вологістю не вище 75%; - нанесення ізолюючого покриття на підлогу; - забезпечення повної та надійної ізоляції струмовідних частин. 	Уникнення пробою, витоків струму та попередження контакту зі струмопровідними частинами.
2	Організаційні	<ul style="list-style-type: none"> - проведенні необхідних інструктажів з правил електробезпеки. 	Забезпечення наявності у персоналу знань щодо правильної та безпечної експлуатації приладів.
3	Режимні	<ul style="list-style-type: none"> - перевірка та усунення несправності приладів тільки за умови, що вони знаходяться у відключеному стані. 	Попередження контакту з елементами, що знаходяться під напругою

4	Експлуатаційні	регулярний моніторинг стану пристроїв (температура поверхні, рівень заряду); - регулярне технічне обслуговування; вчасна заміна виявлених пошкоджених елементів.	Забезпечення безпечної роботи з об'єктом, зменшення ризику контакту лікаря з електричним струмом, попередження виникнення вибуху внаслідок несправної роботи елементів.
---	----------------	---	---

4.4. Пожежна безпека

Джерела пожежонебезпеки у лабораторії: ноутбук, стіл робочий дерев'яний. У кабінеті присутні наступні горючі речовини:

- Дерево.
- Тканина (матеріал жалюзі, штори).
- Плати (комп'ютер).

У таблицях 4.2-4.4 визначено джерела пожежі, категорію приміщення, зону класу та клас можливої пожежі а також наведено перелік засобів та заходів захисту від вибухонебезпечності та пожежі та способи подолання пожежної ситуації при її виникненні.

Таблиця 4.2 – Джерела пожежної небезпеки

№	Найменування	Джерело небезпеки	Причини небезпек	Наслідки небезпеки
1	Ноутбук	Блок живлення, деталі під напругою	Коротке замикання	Виникнення пожежі та опіків, що можуть

2	Матеріали і речовини, схильні до займання	що Загорання до матеріалів	Зовнішнє загорання	призвести до серйозних наслідків для здоров'я працівників і пацієнтів. Крім того, можливе виникнення збитків обладнання, оснащення та особистих речей працівників.
---	---	----------------------------	--------------------	--

Таблиця 4.3 – Характеристика вибухонебезпечності та пожежної небезпеки

№	Назва	Значення	Опис
1	Клас пожежі	А,Е	Виникнення горіння внаслідок загорання твердих речовин та електроустановок під напругою
2	Клас зони приміщення по пожежній безпеці	Клас II-Па	Горючі рідини, тверді горючі та важкогорючі речовини, матеріали, здатні при взаємодії з водою, киснем повітря або одне з одним
3	Категорія пожежної небезпеки	В	горіти лише за умов, що приміщення, в яких вони знаходяться або використовуються, не належать до категорій А та Б [джер].

Таблиця 4.4 – Засоби та заходи захисту від пожежної небезпеки

№	Заходи	Реалізація	Критерій вибору
1	Технічні	Розташування у кімнаті порошкового вогнегасника ВП-5, встановлення кондиціонера, а також	Усунення пожеж та первинних осередків займання, підтримка оптимальних показників

		забезпечення наявності температури та вологи у пожежного крана та рукава у коридорі.	кабінеті.
2	Організаційні	Проведення протипожежних навчань, інструктажів з пожежної безпеки. Створення та розміщення на видному місці планів евакуації.	Навчання з питань безпеки при пожежі, надання алгоритму дій для запобігання людських жертв, надання доступної інформації для швидкої та безпечної евакуації.
3	Режимні	Заборона застосування відкритого вогню у кабінеті, недопущення сторонніх осіб у нього, попередження появи вибухонебезпечних предметів у кабінеті.	Попередження виникнення пожежі внаслідок непередбачуваних факторів.
4	Експлуатаційні	Регулярний моніторинг та своєчасна перевірка стану обладнання, ремонт та заміна у разі потреби.	Запобігання появленню пожеж, які можуть виникнути внаслідок появи технічної несправності обладнання.

4.5 Перевірочний розрахунок штучного освітлення виробничого або офісного приміщення.

Перевірочний розрахунок для штучного освітлення робочого приміщення буде виконано за формулою:

$$F = \frac{E * S * K * Z}{N * n * \eta}$$

де F - світловий потік лампи, лм;

E - мінімальне нормоване освітлення, лк, для заданого розряду зорової роботи;

S – площа приміщення, м² ;

N – число світильників; n - число ламп в кожному світильнику;

K - коефіцієнт запасу;

Z - коефіцієнт нерівномірності освітлення;

η - коефіцієнт використання світлового потоку.

Розрахуємо коефіцієнт світлового потоку за допомогою формули:

$$i = \frac{A * B}{(A + B) * h} = \frac{4.5 * 2.5}{(4.5 + 2.5) * 3} = 0.53 \approx 0.5$$

Освітлення відноситься до першої групи та за даними формули визначено значення коефіцієнту потоку у розмірі 22.

Після цього переходимо до розрахунку світлового потоку лампи:

$$F = \frac{E * S * K * Z}{N * n * \eta} = \frac{300 * 11,25 * 5 * 1,5}{2 * 4 * 22} = 143 \text{лм}$$

Для задоволення вимог штучного освітлення робочого кабінету, площа якого становить 11,25 м², достатньо встановити 2 світильника по 4 лампи.

4.6 Висновки до розділу

У цьому розділі розглянуто планування робочого місця та проаналізовано можливі фізичні та біологічні ризики. Проаналізувавши умови праці в кабінеті, було виявлено, що площа та об'єм приміщення на одного працівника відповідають нормативним вимогам, а робочі місця організовані належним чином. Офіс відповідає вимогам електробезпеки, пожежної безпеки та обладнаний для забезпечення біологічної безпеки. Загалом приміщення відповідає загальним вимогам безпеки, що застосовуються в даному контексті.

РОЗДІЛ 5

ЗАХИСТ НАВКОЛИШНЬОГО СЕРЕДОВИЩА

5.1 Важкі метали і їх негативний вплив на людину і довкілля

В сучасному технологічному світі важкі метали знаходять широке застосування у виробництві різноманітних пристроїв, включаючи електроніку та батареї. Однак, разом із зростанням виробництва та використання гаджетів, виникає серйозна загроза для здоров'я людини і стану довкілля через випуск важких металів у навколишнє середовище.

Основні важкі метали та їх властивості

Один із основних аспектів, що варто розглянути, - це властивості основних важких металів, які часто використовуються у виробництві електроніки. Найпоширеніші серед них – свинець, ртуть, кадмій, хром та інші. Ці метали відомі своєю токсичністю та здатністю накопичуватися в організмі.

Вплив на здоров'я людини

- **Загроза для нервової та ендокринної систем.** Важкі метали можуть проникати в організм людини через різноманітні шляхи, включаючи вдихання, шляхи крізь шкіру та споживання забрудненої їжею чи водою. Вони можуть викликати серйозні порушення в роботі нервової та ендокринної систем, що призводить до ряду хвороб, включаючи нейродегенеративні захворювання та проблеми репродуктивної системи.

- **Вплив на дітей та вагітних жінок.** Діти та вагітні жінки вважаються особливо вразливими до впливу важких металів. Навіть невеликі дози можуть призводити до серйозних порушень в розвитку мозку у дітей та викликати вроджені вади.

Вплив на довкілля

- **Забруднення ґрунту та води.** Відпрацьовані гаджети, які містять важкі метали, після утилізації можуть потрапляти на полігони захоронення, де вони розкладаються та вилигають метали в ґрунт та воду. Це може призвести до серйозного забруднення природних ресурсів та загрози для біорізноманіття.

- **Негативний вплив на екосистему.** Важкі метали, накопичуючись в навколишньому середовищі, можуть мати шкідливий вплив на екосистему, включаючи рибу та інших водних організмів, що може вплинути на здоров'я та життєздатність цих видів.

Важкі метали у платах гаджетів

- **Свинець (Pb):** Свинець використовується у паяльних сплавах для з'єднання елементів плат гаджетів. Завдяки своїм фізичним властивостям, цей метал залишається невід'ємною складовою при виготовленні мікроелектроніки.

- **Ртуть (Hg):** Ртуть може бути знайдена у компактних електронних пристроях, які використовуються для відображення індикації або освітлення. Однак, з огляду на токсичні властивості ртуті, важливо докладати зусиль для заміни її більш безпечними альтернативами.

- **Кадмій (Cd):** Кадмій може зустрічатися в платах гаджетів, особливо в акумуляторах, де він забезпечує високий рівень ефективності. Проте його використання стикається із труднощами через його токсичність, що ставить під сумнів його екологічну придатність.

Важкі метали у батареях гаджетів

- **Кобальт (Co):** Кобальт використовується у літій-іонних акумуляторах, які є основою для живлення багатьох сучасних гаджетів. Використання кобальту забезпечує високу енергоефективність, але одночасно викликає проблеми в етичному аспекті та проблеми забезпечення добуванням.

- **Нікель (Ni):** Нікель також є необхідним для виробництва літій-іонних акумуляторів, хоча деякі виробники шукають альтернативи через його потенційно шкідливий вплив на здоров'я та навколишнє середовище.

5.2 Кількість металів в гаджетах та їх загальний вплив

- **Масштаби використання:** Сучасна споживча культура призводить до експоненційного зростання кількості виготовлених гаджетів. З цим ростом виникає загроза значного використання важких металів і пов'язаного з цим впливу на довкілля та здоров'я людей.

- **Утилізація та ефективність:** Проблема видалення гаджетів, що містять важкі метали, визначає важливість розробки та впровадження більш ефективних систем утилізації та переробки, які дозволять використовувати вторинні ресурси та зменшувати тиск на природні джерела.

Детальний розрахунок кількості важких металів у конкретному гаджеті залежить від його типу, моделі та призначення. Проте, можна надати загальний оцінковий розгляд на прикладі типового сучасного смартфона. Візьмемо до уваги основні важкі метали, які можуть міститися в смартфонах:

1. **Свинець (Pb):** Одна батарея смартфона може містити приблизно 5-10 грамів свинцю в паяльних сплавах на платах.
2. **Ртуть (Hg):** Лампи підсвічування на екранах можуть містити невелику кількість ртуті (порядку 1 грама або менше).
3. **Кадмій (Cd):** Акумулятори зазвичай містять кадмій у кількості приблизно 0.1-0.2 грама.
4. **Кобальт (Co) та Нікель (Ni):** Літій-іонні акумулятори в смартфонах містять кобальт та нікель у різних пропорціях. Наприклад, кобальту може бути приблизно 5 грамів, а нікелю – 10 грамів.

Загальна оцінкова кількість металів в одному смартфоні:

- Свинець: 5-10 г
- Ртуть: 1 г (навіть менше)
- Кадмій: 0.1-0.2 г
- Кобальт: 5 г
- Нікель: 10 г

Зараз, за даними аналітиків із Strategy Analytics, в світі налічується 3,95 мільярда власників смартфонів, тобто гаджети є у кожного другого жителя нашої планети. Імовірно, до 2030 року їх кількість збільшиться ще на мільярд.

Тепер ми можемо використати ці дані для оцінки загальної кількості важких металів у всіх смартфонах у світі.

1. **Свинець (Pb):** Кількість свинцю= $(5-10 \text{ г}) \times (3.95 \times 10^9)$

Для нижнього обчислення: Мінімальна кількість свинцю= $5 \text{ г} \times (3.95 \times 10^9)$

Для верхнього обчислення: Максимальна кількість свинцю= $10 \text{ г} \times (3.95 \times 10^9)$

2. **Ртуть (Hg):** Кількість ртуті= $(1 \text{ г або менше}) \times (3.95 \times 10^9)$

Оскільки точна кількість ртуті невідома, ми використовуємо діапазон "1 г або менше".

3. **Кадмій (Cd):** Кількість кадмію= $(0.1-0.2 \text{ г}) \times (3.95 \times 10^9)$

4. Для нижнього обчислення:

Мінімальна кількість кадмію= $0.1 \text{ г} \times (3.95 \times 10^9)$

Для верхнього обчислення: Максимальна кількість кадмію= $0.2 \text{ г} \times (3.95 \times 10^9)$

5. **Кобальт (Co):** Кількість кобальту= $(5 \text{ г}) \times (3.95 \times 10^9)$

6. **Нікель (Ni):** Кількість нікелю= $(10 \text{ г}) \times (3.95 \times 10^9)$

Давайте розрахуємо ці значення.

1. **Свинець (Pb):**

- Мінімальна кількість свинцю: $5 \text{ г} \times (3.95 \times 10^9) = 1.975 \times 10^{10} \text{ г} = 1.975 \times 10^7 \text{ кг}$

- Максимальна кількість свинцю: $10 \text{ г} \times (3.95 \times 10^9) = 3.95 \times 10^{10} \text{ г} = 3.95 \times 10^7 \text{ кг}$

2. **Ртуть (Hg):**

- Кількість ртуті: $1 \text{ г або менше} \times (3.95 \times 10^9)$ (зазначено, що менше 1 г)

3. **Кадмій (Cd):**

- Мінімальна кількість кадмію: $0.1 \text{ г} \times (3.95 \times 10^9) = 3.95 \times 10^8 \text{ г} = 3.95 \times 10^7 \text{ кг}$

- Максимальна кількість кадмію: $0.2 \text{ г} \times (3.95 \times 10^9) = 7.9 \times 10^8 \text{ г} = 7.9 \times 10^7 \text{ кг}$

4. **Кобальт (Co):**

- Кількість кобальту: $5 \text{ г} \times (3.95 \times 10^9) = 1.975 \times 10^{10} \text{ г} = 1.975 \times 10^8 \text{ кг}$

5. **Нікель (Ni):**

- Кількість нікелю: $10 \text{ г} \times (3.95 \times 10^9) = 3.95 \times 10^{10} \text{ г} = 3.95 \times 10^8 \text{ кг}$

Отже, ці значення представляють оцінкову кількість важких металів у смартфонах, враховуючи розмір потенційного діапазону для кожного металу.

Загальна оцінована кількість важких металів у смартфонах складає від 1.975×10^7 до 3.95×10^7 кілограм в залежності від діапазону для кожного металу.

5.3 Утилізація важких металів у гаджетах та Екологічні Виклики

Потреба у системі ефективної утилізації

Розмаїття важких металів у смартфонах, таких як свинець, ртуть, кадмій, кобальт та нікель, створює важливу потребу у системі ефективної утилізації. Оскільки смартфони стають неодмінною частиною нашого повсякденного життя, важливо враховувати великий обсяг важких металів, які входять до їх складу.

Частковий відтік на полігони захоронення

Певна частина використаних смартфонів неправильно утилізується і потрапляє в побутові відходи, які потім відправляються на полігони захоронення. Цей невірний підхід призводить до того, що важкі метали, які могли б бути ефективно відновлені чи вилучені, викидаються і забруднюють навколишнє середовище.

Ризики для довкілля та здоров'я

Вилуговування важких металів на полігоні захоронення створює серйозні ризики для довкілля та здоров'я. Свинець, ртуть, кадмій та інші метали можуть потрапляти в ґрунт та воду, спричиняючи токсичні ефекти для рослин, тварин та, в кінцевому підсумку, для людей, які використовують ресурси цього довкілля.

Невідповідальність утилізації

Невідповідальна утилізація гаджетів підштовхує проблему невірного викидання важких металів на полігони захоронення. Це може бути особливо актуально в країнах з менш розвиненими системами утилізації електронних відходів, де не існують ефективні методи обробки та вилучення цих матеріалів.

Заходи для подолання проблеми

Системи рециклінгу та вторинного використання: Запровадження та підтримка ефективних систем рециклінгу та вторинного використання, які дозволяють вилучати важкі метали та зменшувати кількість електронних відходів.

Свідоме споживання: Підвищення рівня усвідомленості споживачів про важливість правильної утилізації та вибору пристроїв, які легше рециклуються.

Лобіювання та освіта: Проведення лобіювання для посилення вимог до виробників щодо розробки пристроїв, які легко утилізуються, та освітніх програм для споживачів щодо правильної утилізації гаджетів.

Заклучення

Ефективна утилізація важких металів у смартфонах — це не просто обов'язок, але і важлива ланка в реалізації сталого та екологічно свідомого розвитку. Забезпечення високого стандарту утилізації та усвідомлення небезпеки невідповідального викидання важких металів на полігони захоронення є критичним для збереження чистого довкілля та охорони здоров'я.

Висновок

Розкрито проблематику використання важких металів у сучасних електронних пристроях, зокрема смартфонах, та виокремив необхідність ретельної утилізації для збереження навколишнього середовища та здоров'я людини. У цьому висновку ми підсумуємо ключові аспекти розділу та визначимо шляхи розв'язання проблем, пов'язаних із використанням та утилізацією важких металів у гаджетах.

1. Важкі метали у гаджетах: Розглянуті різні види важких металів, такі як свинець, ртуть, кадмій, кобальт та нікель, які активно використовуються у сучасних технологічних пристроях, зокрема у платах та акумуляторах смартфонів.

2. Вплив на довкілля та здоров'я: Відзначено негативний вплив важких металів на навколишнє середовище та здоров'я. Токсичні властивості цих металів можуть призводити до серйозного забруднення ґрунту та води, а також викликати загрозу здоров'ю людини.

3. Потреба у ефективній утилізації: Висвітлено важливість впровадження систем ефективної утилізації для максимального вилучення важких металів та зменшення негативного впливу на довкілля.

4. Частковий відтік на полігони захоронення: Зазначено, що частина гаджетів, невірно утилізованих, потрапляє на полігони захоронення, де вилуговуються та забруднюють навколишнє середовище.

5. Ризики та невідповідальність: Виділено ризики, пов'язані із забрудненням, та невідповідальність при утилізації, особливо в країнах із слабкою інфраструктурою управління відходами.

6. Заходи для подолання проблем: Представлено конкретні заходи, такі як впровадження систем рециклінгу та вторинного використання, свідоме споживання та освіта для ефективного управління важкими металами в гаджетах.

Узагальнюючи, утилізація важких металів у смартфонах — це необхідний крок для створення сталого, екологічно чистого та здоров'ясберігаючого суспільства. Здатність виробників, споживачів та влади до спільних зусиль у напрямку впровадження ефективних систем утилізації та усвідомлення важливості правильної обробки електронних відходів буде вирішальним фактором для забезпечення сталого розвитку та охорони навколишнього середовища для майбутніх поколінь.

ВИСНОВКИ

У цій дипломній роботі була проведена глибока аналіз системи оптимального вибору методів комп'ютерного зору для обробки зображень. Результати дослідження вказують на великий потенціал впровадження подібних систем у різноманітних галузях, зокрема в медичній діагностиці.

Система, розроблена у цьому дослідженні, демонструє високий рівень ефективності та гнучкості при виборі методів комп'ютерного зору залежно від конкретного завдання. Використання цієї системи у медичних дослідженнях, зокрема у виявленні та класифікації пухлин мозку, може сприяти покращенню точності діагностики та розробці персоналізованих методів лікування.

Прогнозується, що подальший розвиток подібних систем призведе до революції в області комп'ютерного зору та обробки зображень, забезпечуючи високу швидкість та точність аналізу. Застосування цих технологій в медичній сфері може значно підвищити якість діагностики та, в кінцевому підсумку, покращити результати лікування пацієнтів.

Об'єктивність результатів, отриманих у цьому дослідженні, вказує на перспективність подальших досліджень в цьому напрямку та необхідність розгляду розробленої системи у реальних клінічних умовах.

Загалом, дана робота ставить під сумнів традиційні підходи до вибору методів комп'ютерного зору, представляючи інноваційну систему, яка має потенціал перетворити сучасні методи обробки зображень та поліпшити якість медичної діагностики.

Список Використаної Літератури

1. R. B. Fisher, T. P. Breckon, K. Dawson-Howe, A. Fitzgibbon, C. Robertson, E. Trucco, and C. K. Williams, “Dictionary of computer vision and image processing,” NIST Special Publication, vol. 7, no. 4, 2013, pp. 22.
2. D. Venkateshwar Rao, S. Patil, N. Anne Babu and V. Muthukumar , “Implementation and evaluation of image processing algorithms on reconfigurable architecture using C-based hardware descriptive languages,” International Journal of Theoretical and Applied Computer Sciences, vol. 1, no.1 2006, 9-34.
3. D. Kaur and Y. Kaur, “Various image segmentation techniques: A Review,” International Journal of Computer Science and Mobile Computing, vol. 3, no. 5, 2014, pp. 809-814.
4. W.-X. Kang, Q.-Q. Yang, and R.-P. Liang, “The comparative research on image segmentation algorithms,” [in Proceedings of the 1st International Workshop on Education Technology and Computer Science, IEEE, pp. 703-707], 2009.
5. R. F. Yaser, “Multiobject extraction from color image, ” PhD Thesis, College of Science, University of AlMustansirya, 2016.
6. R. C. Gonzalez and R. E. Woods, “Digital image processing,” Prentice Hall, Upper Saddle River, New Jersey 07458, 2002.
7. Deepa M, “Wavelet and curvelet based thresholding techniques for image denoising,” International Journal of Advanced Research in Computer Science and Electronics Engineering, vol. 1, no. 10, 2012, pp. 77-81.
8. V. Zharkova, “Artificial intelligence in recognition and classification of astrophysical and medical images,” vol. 46. Springer Science & Business Media, 2007.
9. P. K. Sahoo, S. Soltani, and A. K. Wong, “A survey of thresholding techniques,” Computer Vision, Graphics, and Image Processing, Elsevier, vol. 41, no. 2, 1988, pp. 233-260.
10. F. Deravi and S. K. Pal, “Grey level thresholding using second-order statistics,” Pattern Recognition Letters, Elsevier, vol. 1, no. 5-6, 1983, pp. 417-422.
11. Li, W.: Automatic segmentation of liver tumor in ct images with deep convolutional neural networks. J. Comput. Commun. 3(11), 146–151 (2015)
12. Vivanti, R., Ephrat, A., Joskowicz, L., Karaaslan, O., Lev Cohain, N., Sosna, J.: Automatic liver tumor segmentation in followup CT studies using convolutional neural networks. Sci. Rep. 2, 15497 (2015)

13. Menze, B.H., Jakab, A., Bauer, S., Kalpathy Cramer, J., Farahani, K., Kirby, J., et al.: The multimodal brain tumor image segmentation benchmark (brats). *IEEE Trans. Med. Image.* 34(10), 1993–2024 (2014)
14. Cherukuri, V., Ssenyonga, P., Warf, B.C., Kulkarni, A.V., Monga, V., Schiff, S.J.: Learning based segmentation of ct brain images: application to postoperative hydrocephalic scans. *IEEE Trans. Bio-Med. Eng.* 65(8), 1871– 1884 (2017)
15. Cheng, J., Liu, J., Xu, Y., Yin, F., Wong, D.W.K., Tan, N.M., et al.: Superpixel classification based optic disc and optic cup segmentation for glaucoma screening. *IEEE Trans. Med. Image* 32(6), 1019–1032 (2013)
16. Fu, H., Cheng, J., Xu, Y., Wong, D.W.K., Liu, J., Cao, X.: Joint optic disc and cup segmentation based on multi-label deep network and polar transformation. *IEEE Trans. Med. Image* 37(7), 1597–1605 (2018)
17. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 234–241. (2015)
18. Song, T.H., Sanchez, V., EIDaly, H., Rajpoot, N.M.: Dual-channel active contour model for megakaryocytic cell segmentation in bone marrow trephine histology images. *IEEE Trans. Bio-Med. Eng.* 64(12), 2913– 2923 (2017)
19. Wang, S., Zhou, M., Liu, Z., Liu, Z., Gu, D., Zang, Y., et al.: Central focused convolutional neural networks: developing a data-driven model for lung nodule segmentation. *Med. Image Anal.* 40, 172–183 (2017)
20. Onishi, Y., Teramoto, A., Tsujimoto, M., Tsukamoto, T., Saito, K., Toyama, H., et al.: Multiplanar analysis for pulmonary nodule classification in CT images using deep convolutional neural network and generative adversarial networks. *Int. J. Comput. Assist. Radiol. Surg.* 15(1), 173–178 (2020)
21. Wu, F., Zhuang, X.: CF distance: a new domain discrepancy metric and application to explicit domain adaptation for cross-modality cardiac image segmentation. *IEEE Trans. Med. Imag.* 39(12), 4274–4285 (2020)
22. Chen, C., Qin, C., Qiu, H., Tarroni, G., Duan, J., Bai, W., et al.: Deep learning for cardiac image segmentation: a review. *Front. Cardiovas. Med.* 7, 25 (2020)
23. Yu Qian, Z., Wei Hua, G., Zhen Cheng, C., Jing Tian, T., Ling Yun, L.: Medical images edge detection based on mathematical morphology. *Proc. IEEE Eng. Med. Biol. Soc.* 2005, 6492–6495 (2006)