

ВИПУСКНА КВАЛІФІКАЦІЙНА РОБОТА

на тему:

«Програмна реалізація машинних методів розпізнавання зображень»

Студента 4 курсу, 8 групи,

спеціальності

122 «Комп'ютерні науки»

Мінасяна Дмитра
Андрійовича

підпис студента

Науковий керівник
кандидат технічних наук, доцент

Томашевська Тетяна
Володимирівна

підпис керівника

Гарант освітньої програми
кандидат технічних наук, доцент

Демідов Павло
Георгійович

підпис керівника

Державний торговельно-економічний університет

Факультет інформаційних технологій
Кафедра комп'ютерних наук та інформаційних систем
Спеціальність 122 «Комп'ютерні науки»

Затверджую

Зав. кафедри _____

Пурський О. І.

« 12 » грудня 2022р.

Завдання

на випускну кваліфікаційну роботу (проект) студенту

Мінасяну Дмитру Андрійовичу

(прізвище, ім'я, по батькові)

1. Тема випускної кваліфікаційної роботи (проекту)
«Програмна реалізація машинних методів розпізнавання зображень»
Затверджена наказом ректора від «9» грудня 2022 р. № 3332
 2. Строк здачі студентом закінченої роботи 10 червня 2023 року
 3. Цільова установка та вихідні дані до роботи
Мета роботи: Розробка системи для розпізнавання зображень на основі різних методів та порівняння їх
Об'єкт дослідження: процеси розпізнавання зображень.
Предмет дослідження: методи і технології розпізнавання зображень
 4. Перелік графічного матеріалу _____
-

5. Консультанти по роботі із зазначенням розділів, за якими здійснюється консультування:

Розділ	Консультант (прізвище, ініціали)	Підпис, дата	
		Завдання видав	Завдання прийняв
1	Томашевська Т. В.	15.12.2022 р.	15.12.2022 р.
2	Томашевська Т. В.	15.12.2022 р.	15.12.2022 р.
3	Томашевська Т. В.	15.12.2022 р.	15.12.2022 р.

6. Зміст випускної кваліфікаційної роботи (проекту) (перелік питань за кожним розділом)

ВСТУП

РОЗДІЛ 1. АНАЛІЗ ЗАДАЧ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

1. Представлення зображень в комп'ютері, характеристики зображення
2. Принципи систем розпізнавання зображень

1.3. Постановка задачі

Висновки до розділу

РОЗДІЛ 2. МАШИННІ МЕТОДИ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

2.1. Методи розпізнавання зображень на принципах кластеризації

2.2. Метод опорних векторів

2.3. Використання нейронних мереж для розпізнавання зображень

Висновки до розділу

РОЗДІЛ 3. РЕАЛІЗАЦІЯ СИСТЕМИ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

3.1. Обґрунтування інструментальних засобів для програмної реалізації системи

3.2 Розробка алгоритму роботи програми розпізнавання зображень

3.3. Порівняльний аналіз методів реалізованих методів розпізнавання зображень

Висновки до розділу

ВИСНОВКИ

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

7. Календарний план виконання роботи

№ пор.	Назва етапів випускної кваліфікаційної роботи	Строк виконання етапів роботи	
		За планом	фактично
1	2	3	4
1	Вибір теми випускної кваліфікаційної роботи	04.10.2022	04.10.2022
2	Розробка та затвердження завдання на випускну кваліфікаційну роботу	15.12.2022	15.12.2022
3	Вступ	03.02.2023	03.02.2023
4	Розділ 1. Аналіз задач розпізнавання зображень	28.02.2023	28.02.2023
5	Розділ 2. Машинні методи розпізнавання зображень	06.04.2023	06.04.2023
6	Розділ 3. Реалізація системи розпізнавання зображень	12.05.2023	12.05.2023
7	Висновки	15.05.2023	15.05.2023
8	Здача випускної кваліфікаційної роботи на кафедрі науковому керівнику	30.05.2023	30.05.2023
9	Попередній захист випускної кваліфікаційної роботи	31.05.2023 - 01.06.2023	31.05.2023 - 01.06.2023
10	Виправлення зауважень, зовнішнє рецензування випускної кваліфікаційної роботи	02.06.2023	02.06.2023
12	Представлення готової зшитої випускної кваліфікаційної роботи на кафедрі	05.06.2023	05.06.2023
13	Публічний захист випускної кваліфікаційної роботи	За розкладом роботи ЕК	

8. Дата видачі завдання «15» грудня 2022 р.

9. Керівник випускної кваліфікаційної роботи (проекту)

Томашевська Т. В.

(прізвище, ініціали, підпис)

10. Гарант освітньої програми Демідов П.Г.

(прізвище, ініціали, підпис)

11. Завдання прийняв до виконання студент-дипломник

Мінасян Д.А.

(прізвище, ініціали, підпис)

Анотація

У випускній кваліфікаційній роботі мною було проведено дослідження методів розпізнавання зображень та їх вплив на людське життя. Починаючи із створення комп'ютерної графіки та її використання до використання штучних нейронних мереж та способів їх створення. Також було представлено інші методи розпізнавання зображень, які є досить популярними у наш час та мають широкі можливості, завдяки яким їх застосовують.

Ключові слова: штучні нейронні мережі, дослідження методів, розпізнавання зображень

Anotation

In my graduation thesis, I conducted a study of image recognition methods and their impact on human life. Starting with the creation of computer graphics and their use to the use of artificial neural networks and how to create them. Other methods of image recognition were also presented, which are quite popular nowadays and have wide possibilities due to which they are used.

Keywords: artificial neural networks, method research, image recognition

Зміст

ВСТУП..... 8

РОЗДІЛ 1. АНАЛІЗ ЗАДАЧ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ 11

1.1 Представлення зображень в комп'ютері, характеристики зображення
 11

1.2 Принципи систем розпізнавання зображень 15

1.3 Постановка задачі 18

Висновки до розділу 1..... 20

РОЗДІЛ 2. МАШИННІ МЕТОДИ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ 22

2.1 Методи розпізнавання зображень на принципах кластеризації..... 22

2.2 Метод опорних векторів..... 28

2.3 Використання нейронних мереж для розпізнавання зображень..... 29

Висновки до розділу 2..... 31

РОЗДІЛ 3. РЕАЛІЗАЦІЯ СИСТЕМИ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ..... 34

3.1 Обґрунтування інструментальних засобів для програмної реалізації системи 34

3.2 Розробка алгоритму роботи програм розпізнавання зображень 38

3.3 Порівняльний аналіз реалізованих методів розпізнавання зображень
 40

Висновки до розділу 3..... 50

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ..... 55

ДОДАТОК А 59

ВСТУП

З кожним роком ІТ технології прогресують з неймовірною швидкістю і дуже швидко певні технології приходять на зміну іншим. Це відбувається настільки швидко, що людині яка не поглиблює знання в ІТ сфері на регулярній основі потім дуже важко розуміти що відбувається навколо неї. Адже ця сфера оточує у певній мірі кожну людину та вносить власні корективи в її життя що залишає певні відбитки в подальшому її житті.

Розпізнавання образів та зображень на даний момент має дуже широкий спектр застосування, а саме створення штучного інтелекту для вирішення певних задач, аналіз даних або навіть пошук інформації. Це дозволяє значно зменшити використання людських ресурсів та прискорити вирішення задачі, а в деяких випадках навіть покращити точність виконання задачі.

Методи розпізнавання зображень покращуються з кожним роком і стають все більш складними та досконалими, що дає більше можливостей для вирішення певних задач та розпізнавання більш складних зображень, які мають занадто багато дрібних деталей. Це необхідно для того щоб розпізнати зображення яке зливається з фоном.

Для розпізнавання використовують дві стратегії. Перша - це моделювання фону, щоб відокремити об'єкт від його безпосереднього фону. Друга стратегія полягає у моделюванні об'єкту, який завчасно відокремили від фону, або який є сам по собі досить простим та не потребує відокремлення.

Розпізнавання зображень (образів)- це класифікація даних які входять до системи за певними ознаками та їх фільтрація із загальної маси. Для розпізнавання використовують математичні розрахунки з метою більш точного пошуку та фільтрації з поміж інших.

Розпізнавання образів буває різною, наприклад до цієї сфери можна віднести:

- Розпізнавання штрих-кодів
- Розпізнавання літер
- Розпізнавання автомобільних номерів
- Розпізнавання фотографій

Тому дослідження різних методів розпізнавання зображень є задачею актуальною.

Метою роботи є розробка системи для розпізнавання зображень на основі різних методів та їх порівняння.

Для цього потрібно вирішити такі завдання

1. Дослідити принципи організації обробки зображень.
2. Дослідити алгоритми розпізнавання зображень.
3. Розробити програмне забезпечення для реалізації обраного алгоритму розпізнавання зображень на основі нейронних мереж.
4. Дослідити ефективність обраного методу розпізнавання зображень.

Об'єкт дослідження: задачі розпізнавання зображень.

Предмет дослідження: методи розпізнавання зображень

Методи дослідження. Теоретичною основою дослідження є загальнонауковий аналітичний метод, а також системний підхід і праці провідних вчених з проблем дослідження і оцінювання рівня розвитку соціально-економічних систем. Для практичного вирішення поставлених задач використовувалися такі методи:

- загальнонауковий аналітичний метод (розділ 1);
- методи обробки зображень (розділ 2);
- методи алгоритмічного програмування, для створення програмної розробки для дослідження методів розпізнавання зображень (розділ 3).

Практичне значення. Отримані результати, можуть бути використані для відбору найбільш ефективних методів обробки зображень.

Структура та обсяг випускної кваліфікаційної роботи. Випускна кваліфікаційна робота складається із вступу, трьох розділів, висновків, списку використаних джерел та літератури із 20 найменувань, додатку і містить 42 сторінки основного тексту.



РОЗДІЛ 1. АНАЛІЗ ЗАДАЧ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

1.1 Представлення зображень в комп'ютері, характеристики зображення

Комп'ютерна графіка- це наукова дисципліна (розділ інформатики), яка займається створенням та обробкою графічної інформації. За способами зображення графіку поділяють на растрову та векторну. Також графіка може бути фрактальною та трьохвимірною.

Векторна графіка- використовується для створення зображень за допомогою геометричних фігур, ліній, тощо. Застосовується з метою створення креслень, схем або діаграм. На відміну від растрових зображень, векторні потрібно створювати з нуля. Редагування є більш легким, але це вже залежить від кількості елементів та їх складності, для створення таких зображень відмінно підходить програма Autocad, яка є досить простою у використанні та дозволяє реалізувати весь спектр можливостей які в неї закладені навіть користувачеві, який бачить її 3 раз у своєму житті. Першим векторним редактором був Sketchpad, його створив американський вчений Айвен Сазерленд у 1963 році. Цей редактор дозволяв малювати крапки, лінії і кола, а також робити різні дії з ними (копіювання, переміщення, тощо). У векторній графіці обсяг задіяної пам'яті не залежить від розміру об'єкту, тому що сам об'єкт представлений у вигляді математичної формули, також можна без втрати якості зображення збільшувати або зменшувати об'єкт. Тому векторна графіка не має вагомих недоліків оскільки сама по собі є досить простою.

Растрова графіка- тип графіки, мета якого створити зображення за допомогою кольорових пікселів, які мають прямокутну форму і певну кількість, що напряму впливає на чіткість зображення. Пікселі створюють так звану сітку і потрібні для найбільш чіткого відображення кольору. Такий тип графіки є найбільш звичним у нашому житті, адже використовується для створення

фотографій, фільмів, відео, тощо. У наш час для вводу в комп'ютер растрових зображень використовуючи переважно цифрові фотоапарати і відеокамери, також можна використовувати телефон з цією метою, оскільки в деяких випадках це навіть зручніше. Для редагування растрових зображень використовують фоторедактори, очолює які Adobe photoshop, який має великий спектр можливостей, але є досить важким для недосвідченого користувача. Розмір зображення залежить від роздільної здатності екрану, роздільна здатність може бути будь-якою 640x480, 800x600, 1024x768, тощо. На даний період часу існують екрани з підтримкою роздільної здатності 3840x2160 або навіть більше. Звісно, цей параметр напряму впливає на сприйняття зображення людиною, чим вище цей параметр тим приємніше сприймати картинку. Головною проблемою, яка обмежує створення таких екранів є необхідність купляти дороге комп'ютерне «залізо», а саме відеокарту, процесор та оперативну пам'ять для комфортної обробки зображення такої якості. Таке зображення також потребує все більшого об'єму накопичувача, через це потрібно застосувати досить емним жорстким диском.

Фрактальна графіка оброблюється як векторна, але відрізняється тим, що жодних об'єктів у пам'яті комп'ютера не зберігається[1]. Зображення будується за рівнянням (або за системою рівнянь), тому нічого, крім формули, зберігати не потрібно[1]. Змінюючи коефіцієнти у рівнянні, можна отримати зовсім іншу картину[1].

Тривимірна графіка(3D- від англ. Tree dimesions- «три виміри») оперує з об'єктами в тривимірному просторі. Зазвичай результати являють собою плоску картинку, проекцію. Тривимірні комп'ютерна графіка широко використовуються в кіно, комп'ютерних іграх[1]. У тривимірній комп'ютерній графіці всі об'єкти зазвичай є набором поверхонь або часток. Мінімальну поверхню називають полігоном. Як полігон зазвичай обирають трикутники[1].

Перші обчислювальні машини не мали окремих засобів для роботи з графікою, але використовувалися для отримання і обробки зображень. Програмуючи пам'ять перших електронних машин, побудовану на основі матриці ламп, можна було отримувати візерунки[1].

У 1952 році з'явилася перша наочна комп'ютерна гра- ОХО, або хрестики-нулики, розроблена Олександром Дугласом для комп'ютера EDSAC в рамках кандидатської дисертації як приклад взаємодії людини з машиною[2].

У 1955 році народилося світлове перо. На кінчику пера знаходиться фотоелемент, що випускає електронні імпульси і одночасно реагує на пікове світіння, відповідне моменту проходження електронного променя[2].

У 1957 році для комп'ютера SEAC зразка 1950-го року при Національному бюро стандартів США команда під керівництвом Рассела Кершо розробила барабанний сканер, за допомогою якого була отримана перша в світі цифрова фотографія. Зображення, на якому зображений тримісячний син вченого, вийшла розміром 5x5 см в роздільній здатності 176x176 точок[2].

У 1958 році в Массачусетському технологічному інституті запущений комп'ютер Lincoln TX-2, котрий вперше використовує графічну консоль[2].

У середині 1960-х років з'явилися розробки в промислових додатках комп'ютерної графіки. Так, під керівництвом Т. Мофетта і Н. Тейлора фірма Itek розробила цифрову електронну креслярську машину. У 1964 році General Motors представила систему автоматизованого проектування DAC-1, розробленої спільно з IBM[1].

У 1967 році на базі Університету Юти організований дослідницький центр комп'ютерної графіки світового масштабу[2].

У 1968 році в СРСР знятий мультфільм «Кішечка», що став першим, в якому з'явився анімований комп'ютером персонаж[2].

У 1972 році Буй Фонг, Роберт Макдермот (Robert McDermott), Джеймс Кларк (James Clark) і Рафаель Ром (Raphael Rom) спільними зусиллями під

керівництвом Айвена Сазерленда створили згенероване Комп'ютерною графікою тривимірне зображення, яке вперше в історії виглядало точно так, як його фізичний оригінал-«Фольксваген-жук» (Volkswagen Beetle)[2].

На SIGGRAPH 1975 року відбулася демонстрація об'єкта, якому судилося стати «іконою» комп'ютерної графіки. Ним виявився...звичайний чайник для заварювання, тривимірне зображення якого створено Мартіном Ньюеллом (Martin Newell) з Університету Юти[2].

У 1978 Джеймс Блінн (James Blinn) розробляє техніку реалістичної візуалізації тривимірних об'єктів[2].

Перші комп'ютери були призначені для вирішення задач на виробництві та з метою поліпшити промисловість. Вже згодом за допомогою комп'ютерів почали будувати графіки, креслення та діаграми для візуалізації результату та кращого розуміння робочого процесу.

З плином часу з'явилися системи автоматизації проектування (САПР), цей підрозділ графіки створений для полегшення та прискорення процесу креслення, що згодом дало свої результати, тепер така робота виконується виключно за допомогою комп'ютера та спеціалізованих програм.

Комп'ютерна графіка має такі області застосування:

- Наукова графіка
- Ділова графіка
- Конструкторська графіка
- Ілюстративна графіка
- Художня і рекламна графіка
- Комп'ютерна анімація
- Мультимедія

1.2 Принципи систем розпізнавання зображень

Існує декілька методів розпізнавання зображень, кожен здатен замінити попередній за необхідності. Цих методів досить небагато, але на даний момент саме вони допомагають виконувати завдання з розпізнавання зображень. Використанням цих методів займається штучний інтелект, який завчасно «навчає» людина або який здатен навчатись самостійно.

За певний час утворилась певна процедура розпізнання, вона складається з певних дій, які задіяні послідовно, а саме: сприйняття образу, попередня обробка, виділення характеристик, класифікація або прийняття рішення.

Досить важливим моментом у розпізнанні зображень є автоматична класифікація (кластер-аналіз) даних. Цей аспект складається з множини схожих реалізацій зображення, які можна виділити на фоні інших об'єктів, що зображені. Під час кластеризації виконується пошук подібних реалізацій та формування вирішальних правил для кожного із кластерів.

Методи розпізнавання поділяють на 4 групи, а саме: порівняння із зразком, використання нейронних мереж, використання статистичних методів, структурні та синтаксичні методи.

Метод січних площин (Геометричний)

Полягає у відокремленні образів за допомогою частин гіперплощин, а сам алгоритм складається з двох етапів, навчання та розпізнавання нових об'єктів.

За основу береться 2 точки, які позначають сам об'єкт, потім проводиться довільна пряма лінія, яка відокремлює ці два об'єкти, потім береться 3 об'єкт та перевіряється правильність побудови відносно прямої. Якщо 3 об'єкт належить до тієї ж самої півплощини що і 2, тоді будується ще одна пряма для їх відокремлення. Тим самим площина розбивається частини. Після цього для відокремлення ще 2 точок проводиться ще 2 прямі. В результаті цих дій проводиться видалення зайвих площин. Після цього залишаються порожні

площини, для їх вилучення використовують гіпотезу компактності, але можна також віднести їх до зафарбованих областей, цей процес має назву «Процес вилучення зайвих частин». Метод січних площин застосовують для розпізнавання образів, таких як: букви, цифри, символи, тощо. Може використовуватись у медицині, автомобільній промисловості, розвідці, тощо.

Метод потенціальних функцій (Алгебраїчний метод)

Цей метод полягає в об'єднанні розмитих множин, що описується потенційною функцією, в той же час здійснюється апроксимація дискримінантною функцією. Її значення вираховується шляхом складання значень потенціалу, який поступово зменшується в залежності від його розташування від центру, який вибирається під час навчання. Принцип полягає в тому, що на просторі вхідних векторів X задається функція, що являє собою той самий потенціал. Він залежить від відстані двох точок по відношенню одна до одної. У висновку можна сказати що даний метод являє собою потенційне поле, яке поділяє робочий простір на 2 половини, які складаються з додатних і від'ємних значень, проміжна частина між двома половинами дорівнює нулю і є розділяючою. Даний метод має кілька переваг, таких як простота та швидкість розпізнавання зображень. Але не зважаючи на це, він не підходить для розпізнавання складних об'єктів, що мають велику кількість деталей та особливостей форми.

Евристичний метод

Основою цього методу є досвід людини та її інтуїція, він складається з певного набору спеціальних процедур створених для вирішення певних задач. Для вирішення кожної із задач використовується певна процедура. Цей метод широко використовується під час технічної та медичної діагностики, розпізнавання букв, штрих-кодів, автомобільних номерів або біометричних даних. Головним принципом даного методу є порівняння зображення з еталоном. Також є більш вдосконалений варіант цього методу коли основою класифікації є множина

відібраних вимірів. До відібраних вимірів можна віднести ознаки, вони є малочутливими до змін або спотворень. У такому випадку процес розпізнавання поділяється на дві задачі. Перша задача полягає у визначенні вимірювань, які мають бути у вхідному образі. Розв'язок цієї задачі є дуже суб'єктивним і залежним від обставин. Друга задача полягає у прийнятті розв'язку про належність вхідного образу від певного класу. Перевагою цього методу є те, що він дозволяє використовувати експертні знання та досвід з метою розпізнавання зображень. Однак, цей метод також має декілька недоліків, таких як відсутність точної математичної моделі та невизначеність результатів розпізнавання.

Лінгвістичний метод (Синтаксичний)

Даний метод полягає у використанні принципу загальності властивостей, основним смислом цього методу є те, що група образів, які відносяться до певного класу має певні спільні ознаки та властивості, за якими їх можна об'єднати. А також він базується на здатності людського мозку до асоціацій і на побудові словесних описів зображень. Замість того, щоб спробувати розуміти пікселі зображення, які складаються з чисел і надзвичайно складні для обробки, цей метод використовує вже готову базу знань з метою розуміння всього зображення. Саме через ці властивості можна записати до пам'яті системи ці ознаки та властивості. Якщо зображення не належить до певної групи то відбувається набір ознак, а потім відбувається їх порівняння з тими, що знаходяться у базі даних. Лінгвістичний метод можна поділити на декілька схем, кожна з якої призначена для вирішення різних задач. У схемі «навчання з учителем» система вчиться розпізнавати образи за допомогою різних адаптивних систем. Якщо використовувати «навчання без учителя» потрібно вивчити класи образів. Цю схему використовують у випадках, коли проводиться відслідкування контуру та інших точок. В машинному представленні найкраще використовувати ланцюговий код, який є універсальним інструментом. Його

перевагою є компактність, будь-які дані про зображення він дозволяє подати у стислому вигляді.

1.3 Постановка задачі

Як ми вже знаємо є досить багато методів розпізнавання зображень і кожний з них має свою сферу застосування, а деякі можуть працювати у тандемі з метою підвищення ефективності. Існує декілька типових методів постановки задач розпізнавання зображення: ідентифікація, відношення об'єкту до певного класу задачі та кластерний аналіз.

Ідентифікація полягає у виділенні певного об'єкта серед інших, що дозволяє почати використовувати методи для розпізнавання зображення. Ця задача дозволяє відкинути непотрібні зображення, які не потребують розпізнавання.

Відношення об'єкта до певного класу задачі необхідне для загального розуміння ситуації, щоб можна було відразу задіяти певний метод, який краще за все впорається з цим завданням.

Класифікація об'єкту необхідна для віднесення об'єкту дослідження до певної групи, з метою відображення проблем та специфікацій цієї самої групи. Це необхідно робити для того, щоб зрозуміти який метод застосовувати для роботи з цим зображенням.

Кластерний аналіз займається поділом об'єктів на класи, які об'єднують схожі між собою зображення за певними критеріями. Цю задачу частіше за все називають «класифікація без учителя», тому що на відміну від попередньої задачі, класи з самого початку не задані.

Якщо виникає проблема розпізнавання образу, це легко може вирішити людина, яка ставить безпосередню задачу розпізнавання, оскільки виконання даного роду завдань потребує постійного контролю з боку людини.

Частіше за все розпізнавання зображення проводиться шляхом використання математичних формул, оскільки це гарантує більшу точність на

відміну від використання нейронних мереж, де ми отримуємо результат експериментальним шляхом.

У більшості випадків для розпізнавання образів використовують монохромне зображення, що дозволяє розглядати його як функцію, яка розташована на площині і тоді вже проводити математичні розрахунки.

Розпізнавання образів є незамінним у:

Технічній діагностиці

Однією з основних задач є виявлення дефектів та їх типів у деталях на виробництві автомобілів, побутової техніки, тощо.

Медичній діагностиці

Мова йде про пошук захворювань у тілі людини через кардіограму, флюорографію та інші аналізи

Розпізнаванні літер

Широкого застосування набуває при використанні сканера, який вводить до комп'ютера друкований текст або зображення. Сам сканер під час роботи формує лише графічне зображення, для подальшої роботи потрібно розпізнати текст, який зображений в документі.

Розпізнанні мови

Дуже необхідно при застосуванні голосового вводу інформації. Досить сучасна технологія, яка набула популярності у сьогоднішній час. Голосовий ввід використовується банально навіть під час вводу інформації у пошукову систему якщо немає можливості власноруч надрукувати необхідний запит.

Охоронній системі

Коли потрібно ідентифікувати певну людину, яка зображена на камері або у базі даних та надати або обмежити їй доступ до певної установи або певних даних.

Висновки до розділу 1

У висновку можна сказати що останнім часом розпізнавання зображень набуло дуже великої популярності як у професійній сфері так і у повсякденному житті. Це відбувається через розвиток інформаційних технологій і складно навіть уявити що буде через 5 або навіть 10 років.

Сфера розпізнавання зображення бере свій початок від великих комп'ютерів, які поступово набували популярності на виробництві і були дуже великими та незручними у використанні. На даний момент кожна людина, яка має вдома або на роботі комп'ютер може користуватися засобами розпізнавання зображень. Але через розвиток цих технологій вже навіть комп'ютер не є предметом першої необхідності, оскільки деякі задачі подібного типу можна використовувати і з телефону. Можливо вже через декілька років можна будете повністю замінити комп'ютер телефоном та виконувати всі задачі майже на ходу у будь-якому місці.

Станом на зараз технології розпізнавання образів повністю витіснили своїх конкурентів у певних сферах, а саме: у медичній діагностиці, в охоронній системі, там де потребується розпізнавання літер та символів, у комп'ютерній графіці, при створенні фільмів та відеоігор, тощо. Станом за зараз ці сфери неможливо уявити без технологій розпізнавання зображення, вони є тісно пов'язаними з ІТ технологіями, оскільки саме завдяки їм можна притримуватись точності роботи та своєчасності її виконання.

Буквально нещодавно набули популярності нейронні мережі, які широко використовуються для роботи із зображенням, в тому числі з метою його розпізнавання. Вони допомагають виконувати майже будь-які задачі із зображенням, які раніше потребували використовувати певні програми та володіти навичками роботи з ними. Це звісно непогано, але у той же час знецінює працю людей, які займаються редагуванням фото та довго вчать цим

займатися на професійному рівні. Це лише невеликий приклад того, як нейронна мережа витіснила фахівців певної галузі.

У природі існує декілька стандартних методів розпізнавання зображень, до таких методів можна віднести: математичний, евристичний, лінгвістичний. Всі вони у певній мірі пов'язані та застосовуються за необхідністю у тих чи інших випадках, що допомагає досить зручно оперувати цими методами та об'єднувати їх за необхідності. Не дивлячись на прогрес штучного інтелекту ці методи складають основу розпізнавання зображень.



РОЗДІЛ 2. МАШИННІ МЕТОДИ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

2.1 Методи розпізнавання зображень на принципах кластеризації

Кластерний аналіз- це статистична процедура, методи розпізнавання зображень на принципах якої використовуються для автоматичного аналізу зображень та їх подальшого класифікування. Основна ідея полягає у тому, що схожі зображення групуються в класи (або кластери), які не перетинаються між собою та у яких використовуються алгоритми кластеризації.

Задача кластеризації- це класифікація об'єктів, на основі їх схожих властивостей та характеристик. Цю задачу можна назвати загальною, оскільки для її вирішення використовуються різноманітні підходи. Створювати кластери можна на основі інтервалів між ними або статистичних розподілах. Кластерний аналіз не є повністю автоматичним, оскільки потребує постійного корегування, а саме зміна методів для опрацювання даних та параметрів моделі. Все це відбувається до тих пір, поки не буде отримано потрібного результату.

Існує два типи кластеризації, а саме жорстка кластеризація та м'яка.

Жорстка класифікація полягає у тому, що відбувається точний відбір об'єктів, які належать до певного кластеру. Вони можуть або належати до кластеру або не належати зовсім, через це даний тип і отримав таку назву.

М'яка кластеризація полягає у тому, що один об'єкт може відноситись за своїми ознаками одночасно до двох або більше кластерів, на відміну від жорсткої кластеризації не встановлюються такі точні рамки для процесу кластеризації.

Типи даних що надходять та проблема визначання відстані:

- Кожен об'єкт має певні характеристики та ознаки, вони можуть належати до числових та нечислових.

- Кожен об'єкт має певну відстань по відношенню до інших об'єктів, які використовуються.

Дані, які використовуються у кластерному аналізі поділяються на типи, такі як:

- Інтервальний
- Порядковий
- Категоріальний

Для вимірювання відстані між об'єктами кожного типу використовується певна одиниця вимірювання. У кожного типу вона відрізняється.

Алгоритми кластеризації можна поділити на спираючись на їх кластерну модель, основою якої є принцип кластеризації.

Прикладом таких алгоритмів є:

- Кластеризація за зв'язком полягає у тому, що точки, які знаходяться ближче один до одного у просторі даних мають більше схожих властивостей, ніж ті, які знаходяться на певній відстані один від одного
- Кластеризація за центроїдом полягає у тому, що кластери представляються центральним вектором, він може знаходитись будь де, а сама подібність визначається дистанцією точки даних до цього вектора, тобто центроїда.

Одними з найпопулярніших методів кластеризації є метод k-середніх та ієрархічний метод.

Метод k- середніх

Як вже було описано раніше, цей метод є дуже популярним методом кластеризації, який націлений на мінімізацію середньої квадратичної відстані між точками у межах одного кластеру. Метод k-середніх не розрахований на велику точність, але має велику швидкість та є досить простим, що і є причиною популярності цього методу. Як показав досвід, даний метод є більш

зручним для кластеризації великої кількості спостережень на відміну від ієрархічного методу, у якому дендограми швидко перевантажуються і можуть втрачати свою наочність. До недоліків можна віднести те, що для контролю роботи необхідний спеціаліст, оскільки є шанс, що середнє значення буде спотворене. Метод k -середніх частіше за все використовують з метою дослідження набору даних, але через свою ненадійність його не варто використовувати як автоматичний метод. Основною метою цього методу є визначення кластерів таким чином, щоб їх мінімальна сумарна варіація між кластерами була мінімізована.

Основний принцип методу k -середніх полягає в тому, що програма поділяється на дві частини:

- Перша частина виконується з заданою періодичністю (наприклад, кожні 100 мілісекунд).
- Друга частина виконується за необхідності (наприклад, коли виникає певна подія).

Метод k -середніх дозволяє забезпечити виконання задач в режимі реального часу, зменшуючи кількість механізмів переключення контексту та оптимізуючи використання обчислювальних ресурсів. Також він дозволяє досягти більшої стійкості та надійності системи, забезпечуючи простоту та ефективність виконання коду.

Алгоритм методу k -середніх:

1. Визначити кількість кластерів, які потрібно створити
2. Вибір випадкових k об'єктів з набору даних, які будуть початковими центрами кластерів
3. Призначення спостереження найближчому центроїду
4. Перерахувати для кожного з кластерів центроїд кластера
5. Ітеративно зменшити загальну суму у певних межах

Ієрархічна кластеризація

Ієрархічна кластеризація полягає у побудові системи вкладених розбиттів. Сама побудова відбувається у вигляді дендрограми. Ієрархічна кластеризація може бути двох типів: дивізійна та агломеративна. Ієрархічна кластеризація часто використовується для аналізу даних в таких галузях як: біологія, маркетинг, фінанси, тощо. Основний принцип цього методу полягає в тому, щоб поєднувати більші кластери з меншими до тих пір, поки не буде сформовано один кластер, який містить усі об'єкти.

Дивізійна кластеризація- це метод кластерного аналізу, який ґрунтується на діленні початкового кластера на менші підкластери. Даний метод починається з утворення одного великого кластера, який потім рекурсивно ділиться на менші кластери досягнення потрібного рівня групування. Дивізійна кластеризація застосовується тоді, коли ми маємо один великий кластер і маємо за ціль поділити його на менші за обсягом кластери, які зможемо аналізувати окремо.

Процес дивізійної кластеризації можна розділити на кілька етапів:

- Представлення набору даних у вигляді одного великого кластеру
- Застосування алгоритму, який розбиває великий кластер на менші за допомогою певної міри подібності між об'єктами
- Застосування етапу №2 до кожного кластеру, який був розбитий раніше на менші кластери

Дивізійна кластеризація зазвичай використовується по відношенню до великих наборів даних, де кількість об'єктів може бути дуже великою. Даний тип кластеризації може бути корисною для групування клієнтів за споживчими звичками або для групування продуктів за певними характеристиками.

До переваг цього методу можна віднести:

- Високу точність
- Швидкість

– Ефективність

До недоліків можемо віднести:

– Дивізійна кластеризація може приводити до розбиття на занадто малий кластер

– Може потребувати використання високопродуктивних обчислювальних систем для ефективного аналізу

Використовуватися даний метод може для:

– Сегментації аудиторії: використовується для групування клієнтів на основі їх поведінки та інших характеристик

– Кластеризації об'єктів: використовується для групування об'єктів на основі їх властивостей

– Аналізу генетичних даних: використання методу для групування генів на основі їх функцій та властивостей

– Виявлення аномалій: використовується для виявлення об'єктів, які відрізняються від інших об'єктів в кластері

Агломеративна кластеризація розпочинається з того, що кожен об'єкт належить до власного кластеру, а потім оснований на схожості між об'єктами об'єднуються з іншими кластерами. Сам метод полягає в тому, що на кожному кроці він об'єднує два найбільш близьких кластера в один, до тих пір, поки всі об'єкти не об'єднуються в один великий кластер. Існує багато способів визначити відстань між кластерами, такі як: Евклідова відстань, косинусна відстань, тощо. Після об'єднання об'єктів в один кластер, ми можемо використати метод різких для визначення оптимального числа кластерів, яке дозволяє нам розділити вихідні дані на більш дрібні групи, які мають певні характеристики. Агломеративна кластеризація може використовуватись в багатьох дослідженнях, до прикладів можна віднести: генетику, соціологію, екологію, тощо. Що стосовно генетики, то вона може бути використана для

аналізу подібності генетичних послідовностей та класифікації видів на основі їх генетичної схожості. Якщо брати до прикладу екологію, то вона може допомогти в класифікації рослинних та тваринних видів на основі спільних ознак та їх взаємодії з навколишнім середовищем.

Агломеративна кластеризація має кілька основних принципів:

– Визначення міри відстані між об'єктами: для визначення схожості між об'єктами необхідно використовувати міру відстані, яка відображає відносну відстань між об'єктами

– Об'єднання двох найбільш схожих кластерів: на кожному кроці об'єднуються два найбільш схожих кластери, поки не буде створений один кластер

– Вибір міри схожості: існує багато різних мір схожості, які можуть бути використані для агломеративної кластеризації. Найбільш поширеними є Евклідова відстань, косинусна схожість та кореляційна схожість.

– Визначення порогу: визначення порогового значення, що відокремлює кластери.

– Створення дерева кластерів: в результаті кожного кроку об'єднання створюється новий вузол дерева кластерів, який включає в себе об'єднані кластери

Алгоритм цього методу має такий вигляд:

1. Проводиться обчислення матриці близькості, в якій міститься відстань між шаблонами
2. Знайти найбільш подібні кластери за допомогою цієї матриці, об'єднати ці кластери, оновити матрицю щоб відобразити дану операцію об'єднання.
3. За наявності всіх шаблонів в одному кластері зупинитись, в іншому випадку повторити крок №2

Гістограма орієнтації градієнтів (Histogram of Oriented Gradients-HOG) також має відношення до методів кластеризації зображень використовується у комп'ютерному зорі та у обробці зображень з метою розпізнання об'єктів. Даний метод полягає у підрахунку напрямку градієнтів в локальних точках зображення, є дуже схожим на гістограму орієнтовних границь, але відрізняється тим, що проводить обрахунки в щільній сітці клітин, які рівномірно розташовані та використовує локальну нормалізацію контрасту з метою підвищення своєї точності. До переваг цього методу можна віднести те, що він є нечутливим до геометричних та фотометричних перетворень крім орієнтації об'єкта. Як було зазначено у дослідях сильна локальна фотометрична нормалізація дозволяє не звертати уваги на рух пішоходів, за умови, якщо вони підтримують вертикальне положення тіла. З цього можна зробити висновок, що даний метод досить гарно підходить для виявлення людей на зображеннях.

2.2 Метод опорних векторів

Метод опорних векторів (SVM) використовують для розподілу зображень на декілька класів. Цей метод оснований на побудові гіперплощини, яка відділяє класи один від одного, вона визначається таким чином, щоб максимізувати відстань між класами та самою площиною. Однією з переваг методу опорних векторів є здатність працювати з великими наборами даних, а також здатність до управління перенавчанням, що забезпечує високу точність класифікації. До мінусів можна віднести необхідність попередньої обробки зображень, яка інколи буває складною та затратною у плані часу.

Метод опорних векторів використовується для:

- Розподілу тексту та гіпертексту на категорії
- Класифікації зображень
- Розпізнавання рукописних символів
- Різних наук

Первинний алгоритм було винайдено у 1963 році, винахідниками стали Володимир Вапник та Олексій Червоненкіс. У 1992 році було запропоновано спосіб створювати нелінійних класифікаторів. У 1993 році було створено, а опубліковано у 1995 році м'яке розділення.

Метод опорних векторів займається пошуком лінії, яка розділяє два класи в просторі, після цього максимізує відстань між цією лінією та найближчими векторами ознак, які належать до кожного з класів. Ці найближчі вектори і називаються опорними, а сама лінія- розділяючою гіперплощиною. Після того, як відбувається пошук розділяючої гіперплощини, метод може бути задіяний у пошуку нових зображень, з цією метою вектори ознак нових зображень проходять через розрахункову розділяючу гіперплощину і визначається їх клас. Даний метод використовується для розпізнавання зображень, які мають різні властивості (форма, колір, текстура), також його можна використати по відношенню до зображень, які мають багато шуму, що заважає розпізнати певне зображення.

Навчання у самому методі полягає у вирішенні задач квадратичного програмування. При використанні такої задачі обчислення залишається досить точним навіть при наявності сотні тисяч об'єктів. Незважаючи на це саме проблема вибору ядра є показником гнучкості цього методу, якщо правильно підібрати ядро, то можна підлаштувати метод під вирішення різних задач.

У висновку можна сказати що метод опорних векторів є одним із найдієвіших та найоптимальніших методів, які застосовуються з метою машинного навчання. Цей метод використовується з метою класифікації об'єктів та робить це з найменшою похибкою.

2.3 Використання нейронних мереж для розпізнавання зображень

Нейронні мережі можна віднести до найефективніших методів для розпізнавання зображень, оскільки вони базуються на використанні набору

штучних нейронів, які взаємодіють між собою з метою виконання певних задач, до яких ми можемо віднести і розпізнавання образів.

Під час навчання нейронної мережі зображення передаються через вхідний шар, де кожен піксель представлений окремим вхідним нейроном. В слід за цим дані передаються через декілька прихованих шарів, які можуть містити сотні або тисячі нейронів. Кінцевий шар виконує задачу класифікації зображень відповідно до заданого набору класів. Під час навчання ваги між нейронами налаштовуються для максимізації точності класифікації, це вимагає великої кількості навчальних даних для досягнення оптимальних результатів.

Нейронні мережі можуть виконувати задачу класифікації або сегментації зображення. Під час класифікації вона визначає до якого класу відноситься зображення, за кольором, формою або за будь-якою іншою характеристикою. Під час сегментації мережа призначає кожному пікселю певну мітку або клас. Це використовується для створення маски або для виділення об'єктів, щоб відрізнити пікселі між собою, який з них відноситься до фону, а який до самого об'єкта.

На даний момент досить багато неройнних мереж є у вільному доступі, наприклад:

- [Remove.bg](#) - безкоштовний сервіс, який дозволяє за декілька секунд видалити фон з будь-якої фотографії без використання графічних редакторів, для цього достатньо лише завантажити зображення.

- [Deepart.io](#)- сервіс, який займається створенням картин на основі зображення, яке було до нього завантажено. Найскладнішою частиною процесу є вибір стилю, у якому буде створена картина.

- [Let's Enhance](#)- цей сервіс має дуже широкий спектр можливостей стосовно фотографій, а саме: масштабування фотографій без втрати їх якості, видалення шумів та артефактів, а також він домальовує частини зображення, яких не вистачає. З самого початку у цього сервісу встановлене обмеження в 15

мегапікселів та 15 мегабайт для кожного зображення, щоб прибрати це обмеження потрібно придбати платну підписку і тоді з'явиться можливість завантажувати фото до 30 мегапікселів.

– Colorize- якщо вам необхідно перетворити чорно-біле зображення у кольорове, то можна звернутися до послуг даного сервісу, потрібно лише вказати посилання на файл або завантажити його і цей сервіс зробить всю необхідну роботу.

– CaptionBot- метою цього сервісу є опис об'єктів, які зображені на фото. Своім успіхом він зобов'язаний двом системам штучного інтелекту, які досить непогано працюють у тандемі, а саме Computer Vision та Natural Language Processing.

До переваг нейронних мереж у розпізнаванні зображень можна віднести їх здатність до адаптації до різноманітних зображень, а саме зображень з різними розмірами, орієнтацією, тощо. Також до переваг можна віднести здатність до розпізнання складних зображень, які складаються з багатьох елементів.

Нейронні мережі досить широко застосовуються у наш час, їх використовують для розпізнання товару на полицях магазину, автомобільних номерів, людей, тощо. Незважаючи на це, навчання нейронних мереж може бути досить вимогливим до ресурсів і потребувати дуже потужну обчислювальну систему.

Висновки до розділу 2.

У цьому розділі було представлено декілька важливих тем у сфері розпізнання зображень, а саме метод кластеризації даних, метод опорних векторів та використання нейронних мереж у наші дні з метою розпізнання образів.

Що стосується методу опорних векторів, він має один суттєвий мінус, а саме необхідність у попередній обробці зображення по відношенню до якого застосовується цей метод. Сам же метод використовується для того, щоб

відокремити потрібні об'єкти від необхідних. Це може знадобитись коли мова йде про велику кількість об'єктів, серед яких потрібно відокремити потрібні для роботи та забезпечити високу точність класифікації. Незважаючи на це, метод може бути досить чутливим до зайвих шумів, які заважають нормальній роботі. Саме завдяки цим характеристикам метод опорних векторів є досить актуальним у наш час, крім того, його можна пристосувати до різних задач через його гнучкість завдяки можливості роботи з різними ядрами. Саме через це навіть сьогодні цей метод є дуже актуальним.

Метод кластеризації даних є повною протилежністю вище згаданого методу опорних векторів, він виконує задачу групування об'єктів з метою об'єднання їх у певні категорії або класи за їх спільними ознаками. Для реалізації цього методу використовують певні алгоритми кластеризації, такі як: k-середніх, агломеративну, дендрограми, тощо. Основою для кластеризації можуть бути різні характеристики (колір, форма, текстура, тощо). Окрім того, після розділення зображення на кластери, можна задіяти методи класифікації по відношенню до цих кластерів. Оскільки цей метод є досить простим та універсальним він є досить актуальним на сьогоднішній день. Прикладом його використання є медична діагностика, яка потребує групування зображень органів людини або хвороб. Також до прикладів можна віднести діагностику людського зору, яка теж відноситься до медицини. У цій сфері потребується розподіл зображення на основі кольорів, форм або текстур. Виходячи з цього метод кластеризації досить широко використовується як у медицині так і у інших сферах людської діяльності.

Про нейронні мережі можна багато чого розповісти, але основним є те, що вони стали дуже популярними останнім часом, оскільки виконують роль помічників. Це дозволяє використовувати їх у будь-який час у будь-якому місці та з будь-якого пристрою.

Вони можуть виконати роботу з редагування тексту або фото за лічені хвилини, а саме головне те, що вони не потребують навичок користування на відміну від фоторедакторів або текстових редакторів, які потребують серйозних знань для того, щоб їх використовувати та робити якісну роботу. Прикладом такого використання можна привести те, що ще 5-10 років тому для обробки фото потрібно було мати на комп'ютері Adobe Photoshop та вміти їм досить гарно користуватись, адже це не досить проста програма, але вона має досить багато інструментів для обробки фото. Зараз потреба в цьому досить швидко зникає, тому що можна знайти нейронну мережу, яка задовольняє ваші потреби, завантажити до неї фото та дати їй вказівки щодо виконання завдання з обробки і вона дуже високою вірогідністю все виконає. Звісно, це не дає вам точного результату, але у більшості випадків нейронні мережі задовольняють потреби їх користувачів, у крайньому випадку можна просто самостійно відформатувати отриманий результат.

Принцип роботи таких мереж спирається на використання нейронів, як можна зрозуміти з назви. Спочатку дані подаються на нейрон, слідом за цим обчислюється вага даних, а після цього результати передаються на наступний нейрон і так далі, по колу.

РОЗДІЛ 3. РЕАЛІЗАЦІЯ СИСТЕМИ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

3.1 Обґрунтування інструментальних засобів для програмної реалізації системи

Системи розпізнавання зображень можна реалізувати багато якими методами або технологіями, такими як: нейронні мережі, машинне навчання, комп'ютерне зорове бачення, тощо. Це є досить серйозним показником, адже каже про гнучкість систем розпізнавання зображень.

Існує декілька кроків, які можна виконати з метою програмної реалізації систем розпізнавання зображень:

1. Збір набору зображень для навчання системи, він має містити зображення, які система має розпізнати.
2. Навчання нейронної мережі розпізнанню зображень з використанням певного навчального набору. Існує досить багато архітектур нейронних мереж.
3. Перевірка роботи нейронної мережі на тестовому наборі зображень, обов'язково щоб зображення було новим, відмінним від того, яке мережа бачила під час навчання.
4. Застосування системи до нових зображень, які необхідно розпізнати
5. Перевірка та аналіз ефективності роботи нейронної мережі, вдосконалення алгоритму за необхідності.

Це лише невеликий приклад програмної реалізації за допомогою нейронної мережі, шляхом її навчання та використання по відношенню до зображень.

Для створення засобів розпізнавання зображень було створено велику кількість допоміжних бібліотек, більша їх кількість була розроблена з використанням мови Python, але для більш комфортного використання вони були адаптовані під інші мови програмування, такі як: Java, C#, C++, тощо.

Вибір мови програмування відбувається за певними критеріями, такими як:

- Вимоги проекту: різні проекти потребують різні функції, які можуть бути недоступними у певних мовах програмування.
- Екосистема: певні мови програмування мають більш розвинену екосистему в порівнянні з іншими, що дозволяє швидше та точніше виконувати певні задачі.
- Досвід програміста: це також може вплинути на вибір мови, адже кожен програміст має в одній мові програмування більший досвід ніж в іншій.
- Швидкість роботи: деякі мови програмування можуть забезпечити більшу або меншу швидкість роботи в порівнянні з іншими.

На даний момент найбільш за все набула популярності така мова програмування як Python, через свою простоту та екосистему вона дозволяє з більшим комфортом та швидкістю займатися створенням нейронних мереж ніж з використанням інших мов програмування.

Більшість таких бібліотек розробляються під певні задачі та для кращої оптимізації та більш ефективного вирішення завдань певної області.

Існує дві основні бібліотеки у сфері комп'ютерного зору та розпізнаванні обличчя:

- Dlib - інструменти C++, які містять алгоритми машинного навчання та допомагають створювати складне ПО на C++ для виконання певних задач. Dlib має на борту підтримку розробок для робототехніки, вбудованих пристроїв та великих обчислювальних середовищ з високою продуктивністю, також ця система має відкритий код, що дозволяє безкоштовно використовувати її будь-яких розробках.

- OpenCV (Open Source Computer Vision Library) - ця бібліотека випускається за ліцензією BSD, це означає що вона може бути безкоштовно використана для академічних та комерційних розробок. Має інтерфейси для розробки за допомогою C++, Python та Java. Має підтримку всіх популярних ОС

(Windows, Linux, MacOS, IOS, Android). Ця бібліотека була створена з метою розробки real-time систем та має високу обчислювальну ефективність.

Якщо є мета розробити кросплатформенну систему, то краще за все буде використання мови Java, оскільки код, щоб написаний цією мовою компілюється за допомогою JVM в проміжний бінарний формат, а потім проходить запуск за допомогою JVM.

Для функціонування даної системи є певні необхідні модулі:

- Отримання зображення- відповідатиме за можливість системою отримати зображення для подальшого розпізнання. Може містити інтерфейс, а може працювати безпосередньо з камерами.
- Опрацювання зображення- основний модуль системи, в ньому застосовуються результати раніше проведених досліджень. Основним завданням є опрацювання зображення та виділення обличчя на ньому.
- Розпізнання обличчя- виділення унікальних точок-орієнтирів обличчя і проведення пошуку подібних рис за допомогою SVM або за допомогою бази даних.

Це було представлення системи розпізнання обличчя, яка слугує невеличким прикладом програмної реалізації системи розпізнання зображень.

Для створення нейронних мереж використовується досить багато бібліотек в залежності від потреб та вимог, ось невеликий приклад:

- Caffe- бібліотека для розробки нейронних мереж, яка має відкритий код, та спеціалізується на обробці зображень і використовується в багатьох дослідницьких проектах.
- MXNet- як і попередня бібліотека має відкритий код, дозволяє тренувати та валідувати моделі на кількох пристроях.

- TensorFlow- відкрита платформа для розробки інтелектуальних систем, вона використовується для розробки глибоких нейронних мереж, обробки природних мов та комп'ютерного зору.
- Keras- високорівнева бібліотека, яка базується на TensorFlow, вона дозволяє легко створювати та навчати нейронні мережі.
- PyTorch- ця бібліотека має також відкритий код, використовується вона для розробки та навчання нейронних мереж, а також надає зручний інтерфейс для створення моделей глибокого навчання.
- Pandas- пакет, який використовується для роботи з міченими та реляційними даними. Цей пакет побудований на двох основних структурах даних, а саме «Серії» та «Кадри даних». Також цей пакет підтримує перетворення структур даних в об'єкти DataFrame, роботу з відсутніми даними, додавання або видалення стовпців з DataFrame, візуалізація даних з використанням гістограм або блоків графіків та імпортування відсутніх файлів.
- XGBoost- це набір інструментів, який був розроблений для швидкості, гнучкості та портативності. Це швидкий і точний метод підвищення паралельного дерева, який здатний вирішити широкий спектр проблем науки про дані. Використовуючи фреймворк Gradient Boosting, цю бібліотеку можна використовувати з метою створення алгоритмів машинного навчання. Він включає в себе паралельне підвищення рівня дерев, що допомагає командам вирішувати різноманітні проблеми.
- NLTK (Natural Language Toolkit)- досить популярний пакет Python, який займається тегуванням тексту, токенизацією, семантичним міркуванням та іншими завданнями. На даний момент він використовується як навчальний інструмент та у ролі навчальної платформи для створення прототипів і розробки дослідницьких систем.

3.2 Розробка алгоритму роботи програм розпізнавання зображень

Існує безліч алгоритмів створення програм, які виконують задачу розпізнавання зображень, але можна навести приклад, який зустрічається частіше за все:

- Збір та попередня обробка даних. Проходить збір попередньо оброблених зображень, які використовуються для навчання моделі.
- Навчання самої моделі. Відбувається власне сам процес навчання моделі з використанням попередньо обробленого зображення. З цією метою використовуються досить різні алгоритми, прикладом можна назвати нейронні мережі.
- Тестування моделі. Після навчання самої моделі проходить її перевірка на ефективність та виявлення помилок, проводиться повна оцінка її роботи та виявляються можливі помилки.
- Підготовка до використання. Після всіх попередніх етапів відбувається підготовка моделі до використання в реальних умовах, може відбуватися перетворення зображень у необхідний формат, зменшення їх розмірів, тощо.
- Безпосереднє використання. Після всіх попередніх етапів наша модель повністю готова до роботи, ми можемо її використовувати на повну силу для виконання необхідної нам роботи.

Звісно, це лише невеликий приклад алгоритму для створення такої моделі, вони можуть відрізнитися між собою в залежності від складності та типу моделі яка нам потрібна для виконання роботи.

Також, ми можемо з метою розпізнавання зображень створити нейронну мережу. Вона може бути простою або згортковою. Приведемо приклад створення згорткової нейронної мережі.

Для створення такої нейронної мережі застосовуються 3 шари:

- Згортки
- Підвибірки
- Повнозв'язний шар

Першим кроком є вхідні дані або «input», які мають на борту інформацію про зображення.

Другий крок це шар згортки «Convolution», у якому відбувається поелементне множення, потім всі ці множення підсумовуються та записуються у нову структуру даних.

Особливістю згорткової нейронної мережі є те, що вона має безліч фільтрів через які відбувається пошук характеристик. Кожна позиція є унікальною і вона виробляє число, яке записується у структуру даних.

Третій крок «pool»- полягає у зменшенні розміру зображення (висоти та широти). В цей час виконується нелінійне ущільнення картки ознак. Серед усіх пікселів ми обираємо пікселі з найбільшим значенням.

Четвертий крок- блок нормалізації «Rectified linear units», прибирає ознаки які не потрібні нам для розпізнавання.

П'ятий крок «Fully connected layer» виводить n-вимірний вектор, на цьому етапі дані з попереднього етапу зрівнюються з тими типами об'єктів, які нейронна мережа вже може розпізнати. Такий підхід дозволяє нейронній мережі підвищити власну ефективність та демонструвати високий результат власної роботи.

Можна сказати що кожен шар цієї мережі використовує власне перетворення, що дає можливість правильно класифікувати кожен з об'єктів, які ми бачимо на зображенні. Такі методи ми можемо спостерігати в МРТ-діагностиці або при розпізнаванні обличчя певної людини. Також роботу нейронних мереж ми можемо спостерігати під час візуального пошуку. Спочатку людина робить фото певного предмету, а потім нейронна мережа знаходить цей предмет на просторах інтернету. Простіше кажучи нейронні мережі створені з

метою прискорення певних процесів пошуку, певного об'єкту (фото, текст, тощо).

3.3 Порівняльний аналіз реалізованих методів розпізнавання зображень

Як вже було описано раніше, існує досить багато методів розпізнавання зображень, кожен з них має певні властивості та області використання. Не можна сказати що один метод сильно краще або гірше іншого, оскільки кожен з них має як і свої переваги так і свої недоліки. Тому порівнювати їх з точки зору який краще, а який гірше я не планую, але ніхто не заважає порівняти їх у загальному плані та висвітлити переваги та недоліки кожного з цих методів, а вже на основі цих даних ґрунтувати власну думку та надавати перевагу певному з цих методів.

Почати хочеться з методу штучних нейронних мереж, які останнім часом приймають активну участь як у сфері розпізнавання зображень так і у людському житті загалом і є досить потужним інструментом для виконання багатьох задач. Цей метод використовується з метою навчити комп'ютер розпізнавати зображення, а для цього використовується велика кількість приладів, що є досить незручно, але як показує практика дуже ефективно. Даний метод хоч і є досить потужним, але він і не позбавлений недоліків. Для точного розпізнання необхідного об'єкта на зображенні необхідно щоб він дотримувався певних умов, що стосується цих умов, то вони є індивідуальними передусім для нейронної мережі, адже кожен нейронну мережу навчають індивідуально та за певних умов і с певною ціллю. Через це даний метод не є універсальним за будь-яких умов, інколи потрібно звертатися за допомогою до інших методів. Якщо більше прискіпливо проводити розбір переваг нейронних мереж, то можна виділити такі переваги: висока точність, автоматизація, розширення, незалежність від особистих вподобань, тощо.

Висока точність: що стосовно даної переваги, то все досить неоднозначно, адже для того щоб досягти високої точності роботи нейронної мережі необхідно затратити досить багато ресурсів та даних для її навчання, що потребує також великої кількості часу та людських вмінь та навичок. З цього ми можемо зробити висновок що висока точність може бути досягнута тільки якщо дотримуватись вище описаних умов.

Автоматизація: на мою думку це одна з основних причин за яких створили нейронні мережі, оскільки вони дозволяються повністю автоматизувати процес розпізнавання об'єктів на будь-якому зображенні. Саме ця перевага досить сильно виділяє нейронні мережі в порівнянні з іншими методами, адже станом на сьогоднішній день людині необхідний саме автоматизований помічник у виконанні її буденних задач. Я вважаю що рано чи пізно такий помічник мав би з'явитись, можливо через деякий час з'явиться щось більш досконале, але станом на зараз це можна назвати великим кроком вперед.

Розширення: ця перевага дуже тісно пов'язана з першою у списку, адже важливо як навчити нейронну мережу з самого початку, так і після цього слідкувати за її ефективністю і поступово вчити після цього. Виходячи з цього ще однією перевагою є те, що є здатність навчати мережу після створення, бо якщо не було б такої можливості це ускладнило б роботу її створювачів. Останнім часом я почав спостерігати таку тенденцію, що великою популярністю користуються саме ті засоби, які можна модернізувати у ході їх використання і це дуже чудово, адже це додає гнучкості у роботі з цими засобами і прибирає потребу створення нових, схожих за направленням засобів.

Незалежність від особистих вподобань: дана перевага є досить очевидною, оскільки людина створює нейронну мережу під власні потреби і було б дивно якщо нейронна мережа стала б вибирати зображення з яким вона буде працювати або не буде.

Але не все так гарно як ми могли б уявити, тому хочеться навести більш детальний приклад саме їх недоліків, таких як: необхідність у великому обсязі даних, високі вимоги до обчислювальної потужності, недостатня робота на межі класів та потреба у збалансованих даних.

Необхідність у великому обсязі даних: назва цього методу каже сама за себе, я вважаю що цей недолік є одним з найвагомих, оскільки є показником того, скільки ресурсів потрібно для створення нейронної мережі. Це досить кропітка робота, а до цього додається те, що для її навчання необхідно використовувати досить багато даних щоб досягти високої точності цієї самої мережі.

Високі вимоги до обчислювальної потужності: цей недолік на мою думку тісно пов'язаний з попереднім, адже велика кількість даних потребує для обробки досить потужний пристрій. Через це навряд чи можливо створити якісну нейронну мережу на слабкому пристрої, через це саме ці два недоліки є мабуть самими серйозними по відношенню до інших. Я вважаю що дані недоліки можливо виправити у найближчому майбутньому, адже багато які програми або комп'ютерні додатки на старті потребували досить потужні пристрої для їх створення або запуску, а вже згодом і комп'ютерне залізо стало більш досконалим та доступним і ці додатки стали більш оптимізованими, тому не варто хвилюватись с приводу нейронних мереж, бо це лише питання часу. Особливо я вірю що вже через декілька років кожен хто забажає зможе створити нейронну мережу майже на калькуляторі.

Потреба у збалансованих даних: для коректної роботи нейронної мережі необхідно досить гарно збалансувати дані для того, щоб вона їх точно сприймала і робила подальшу їх обробку. Це досить легко пояснити, адже якщо ввести при навчанні нейронної мережі дані, які є незбалансованими та необробленими, то в подальшому вона буде некоректно працювати та виконувати ті задачі, які їй дадуть.

Інтерпретаційна проблема: сама по собі нейронна мережа є досить складною системою, тому це викликає деякі проблеми з даними, які вона видає користувачеві, що може досить сильно вплинути на співпрацю з цією мережею.

Вразливість до атак: ця проблема є дуже поширеною навіть серед звичайних програм та комп'ютерних мереж, а що стосовно нейронних мереж, то це особисто мене не дуже здивувало, але це факт. Нейронні мережі дійсно вразливі до атак і вірусів, через що робота нейронної мережі порушується і вона видає некоректний результат або зовсім не те, що потребує користувач.

В слід за методом штучних нейронних мереж хочеться розглянути метод опорних векторів, який є також одним з основних коли мова йде про розпізнавання зображень. Його можна назвати одним з базових та мабуть найпростіших якщо порівнювати з попереднім варіантом.

До переваг можна віднести:

Можливість працювати з високо- і низькорозмірними даними: коротко кажучи це відбувається за рахунок того, що цей метод зовсім не залежить від кількості ознак в наборі даних, які використовуються. Це є дуже вагомою ознакою через те, що додає методу універсальності, що допомагає його використовувати з більшим комфортом.

Ефективна робота з великими наборами даних: SVM або метод опорних векторів здатний обробляти великий обсяг даних, це відбувається через те, що він працює безпосередньо з образами, які є пріоритетними коли мова йде про виконання задачі з розпізнавання зображень.

Гнучкий вибір ядра: метод опорних векторів використовує ядро з метою перетворити оригінальні дані в вищорозмірний простір, де класи можуть бути легко розділені. Існує досить багато різновидів ядер, вибір яких може залежати від характеристик потрібним нам даних. На мою думку ця перевага є основою самого методу SVM, оскільки на відміну від нейронних мереж цей метод

використовує зовсім інші інструменти та має дуже гарно з ними працювати і гарно адаптуватися до змін.

Забезпечення універсальності: як і вище описаний метод цей також не позбавлений універсальності, він здатен застосовуватися до широкого спектру задач класифікації та регресії, включаючи задачі з багатокласовою класифікацією, а також можливість вирішувати задачі з несиметричними відносинами помилок класифікації. Ця особливість є об'єднанням всіх попередніх, оскільки вона не виділяє якісь певні можливості або параметри.

До недоліків я можу віднести:

Досить великі вимоги до обчислювальних ресурсів: даний метод може бути дуже вимогливим до обчислювальних ресурсів, особливо коли мова йде про використання великих наборів даних та складних ядер. Цей недолік є спільним як у штучних нейронних мереж так і у методу опорних векторів і це досить закономірно, якщо метод є потужним то він потребує потужних ресурсів для його повноцінного використання.

Погано працює з несиметричними даними: це трохи протирічить вище згаданим перевагам, але не відмінняє того факту, що SVM показує не досить точні результати під час співпраці з несиметричними даними, мається на увазі коли класи мають у собі різну кількість елементів.

Потребує правильного підбору гіперпараметрів: використання даного методу потребує відповідного підбору гіперпараметрів, мається на увазі тип ядра, параметри ядра та параметри регуляризації. При неправильному виборі гіперпараметрів може відбуватися неправильне навчання або перенавчання моделі.

Не рекомендується для задач з багатими етнічними групами: застосування цього методу може бути проблематичним коли мова йде про задачі класифікації які мають багаті етнічні групи, оскільки вони мають дуже різні характеристики, які можуть призводити до хибної класифікації.

Це ми розглянули лише 2 методи, які є досить популярними, але я вважаю що потрібно згадати про методи, в основі яких лежить кластеризація даних, робота цих методів полягає у групуванні схожих між собою зображень за певними ознаками та особливостями.

До переваг можна віднести: незалежність від попереднього навчання, можливість автоматизації, стійкість до шуму та деформації. А до недоліків: висока складність, необхідність правильного підбору параметрів, залежність від вибору метрики подібності, схильність до перенавчання, залежність від кількості зразків, необхідність вибору кластерів для подальшого аналізу. Слід детально їх розібрати.

Незалежність від попереднього навчання: я вважаю що тут і так все зрозуміло, на відміну від методу штучних нейронних мереж ці методи не потребують попереднього навчання. Через це не потрібно витратити час та зайві зусилля.

Можливість автоматизації: методи на основі кластеризації мають можливість автоматизації, що дає можливість розпізнавати зображення в режимі реального часу.

Стійкість до шуму та деформацій: сам кластерний аналіз може бути стійким до деформацій та зайвого шуму які є на зображенні. Це відбувається через те, що він використовує глобальну інформацію про зображення з метою пошуку подібних зразків зображення.

Також маю за ціль розповісти про недоліки:

Висока складність: кластерний аналіз досить витратний з точки зору обчислювальної складності, особливо коли мова йде про велику кількість зображень. На мою думку цей недолік мають майже всі методи розпізнавання зображень які здатні працювати з великим обсягом даних.

Необхідність правильного підбору параметрів: для правильного розподілу зображень між кластерами необхідно правильно підібрати параметри, такі як

кількість кластерів, метрика подібності, тощо. Цю властивість досить складно назвати недоліком, адже якщо фігурує розподіл зображень за певними характеристиками необхідно щоб користувач налаштував все одразу під себе та під свої потреби.

Залежність від вибору метрики подібності: цей недолік є скоріше за все доповненням до попереднього. Через невірний вибір метрики може відбуватися неправильно розподілене кластеризування зображень.

Схильність до перенавчання: цей недолік є досить поширеним серед усіх методів машинного навчання. Схильність до перенавчання проявляє себе коли кількість кластерів є занадто великою і тоді вони можуть представляти одну і ту саму категорію.

Залежність від кількості зразків: може бути така ситуація, коли кількості зображень недостатньо для коректної роботи методу, тоді відбувається неправильна їх класифікація.

Необхідність вибору кластерів для подальшого аналізу: в слід за проведенням кластеризації необхідно вибрати кластери, які відповідають певним категоріям зображень. Цей процес може бути складним, особливо якщо кластери є дуже подібними один до одного.

Крім вище описаних методів існують ще методи розпізнавання зображень, такі як: метод глибокого навчання, метод зіставлення шаблонів та метод заснований на статистичному зображенні. Я вважаю що слід додатково описати ці методи для розуміння повної картини.

Хотілося б почати з методу глибокого навчання. Він є одним з найбільш ефективніших серед тих, що зараз використовуються. Цей метод базується на використанні глибоких нейронних мереж з багатьма шарами для автоматичного визначення відповідних ознак на зображенні та класифікації його відповідно до заданих категорій. Основна перевага методу глибокого навчання полягає в тому, що він може працювати з великим наборами даних, де наявність багатьох різних

характеристик ускладнює задачу класифікації. Також, цей метод здатний автоматично навчитися розпізнати нові ознаки на зображеннях, що робить його гнучкішим та більш універсальним. Недоліком цього методу є висока вимогливість до обчислювальної потужності та велика кількість даних, необхідних для навчання нейронної мережі. Крім того, налаштування параметрів мережі та її навчання можуть бути складними завданнями. Застосування методу глибокого навчання у розпізнаванні зображень включає такі області, як розпізнавання обличчя, емоцій, класифікація зображень за категоріями, тощо. Також даний метод може бути застосований для визначення об'єктів на зображенні, сегментації зображення та багатьох інших завдань, які пов'язані з обробкою зображень.

Метод зіставлення шаблонів є одним з методів розпізнавання зображень, який полягає в порівнянні вхідного зображення зі зразком для визначення ступеня подібності. Цей метод може використовуватися з метою виявлення об'єктів на зображенні або для класифікації зображень на основі їх вмісту.

Основними кроками методу зіставлення шаблонів є:

1. Вибір шаблону або зразка, який необхідно розпізнати на вхідному зображенні.
2. Використання алгоритму порівняння з метою знаходження відповідностей між шаблоном і вхідним зображенням.
3. Встановлення порогового значення для прийняття рішення про те, чи містить вхідне зображення шаблон.
4. Якщо вхідне зображення містить шаблон, то він може бути виділений або класифікований відповідно до потреб користувача.

До переваг можна віднести простоту та швидкість аналізу зображень, особливо якщо шаблони є простими та рідкісними на зображенні. До недоліків належать обмежена точність у випадку, коли шаблони мають складну форму або

знаходяться в різних масштабах та орієнтаціях на зображенні, а також чутливість до змін освітлення та шуму на зображенні.

Метод розпізнавання зображень заснований на статистичному зображенні можна віднести до методів машинного навчання. Даний метод використовує статистичні характеристики зображень (середнє значення, стандартне відхилення, кореляція, щільність) для опису зображень та виявлення закономірностей у даних. Для використання цього методу потрібно спочатку побудувати модель, що описує статистичні характеристики зображень, з використанням відповідного алгоритму машинного навчання, такого як навчання з учителем або навчання без вчителя. Після цього, модель може бути використана для класифікації нових зображень на основі їх статистичних характеристик. Один з підходів до методу розпізнавання зображень на основі статистичного зображення- це метод головних компонент або (PCA). Який полягає у зменшенні кількості змінних, які використовують для опису зображень, шляхом перетворення їх на нові змінні, що представляють основні складові зображень. Потім ці нові змінні можуть бути використані для класифікації зображень. Метод розпізнавання зображень, який заснований на статистичному зображенні має свої переваги та недоліки, як і будь-який інший метод. До переваг належать висока точність та можливість класифікації зображень у великій кількості класів. До недоліків належать складність в обчисленнях та потреба в достатньо великій вибірці зображень для побудови моделі. Також метод може бути чутливий до змін у масштабі або орієнтації зображень, а також до різних типів шуму та артефактів на зображеннях.

Якщо узагальнювати всю отриману інформацію то методи розпізнавання зображень можна поділити на декілька категорій:

1. Методи побудовані на основі характеристик зображення- ці методи базуються на визначенні ключових ознак на зображенні, таких як кути, ребра, кольори, тощо. А після цього використовують дані ознаки для класифікації

об'єктів на зображенні. Прикладами таких методів можна назвати: SIFT (Scale-invariant Feature Transform), SURF (Speeded-Up Robust Features), ORB (Oriented Fast and Rotated BRIEF) і так далі.

2. Методи на основі глибокого навчання- ці методи зазвичай використовують нейронні мережі глибокого навчання для автоматичного визначення ознак та класифікації об'єктів на зображеннях. До прикладів таких методів можна віднести: CNNs (Convolutional Neural Networks), RNNs (Recurrent Neural Networks) та CRNNs (Convolutional Recurrent Neural Networks).

3. Методи на основі геометричних моделей- дані методи призначені на використання геометричних моделей об'єктів та їхніх відносин на зображенні для класифікації об'єктів. Прикладами таких методів є ASM (Active Shape Models), AAM (Active Appearance Models), DPMs (Deformable Part Models).

4. Методи, які базуються на інформації про контекст- такі методи використовують інформацію про контекст зображення, таку як розташування об'єктів на зображенні та їхні відносини з метою подальшої класифікації об'єктів. До прикладів можна віднести: CRFs (Conditional random Fields) та MRFs (Markov Random Fields).

5. Методи на основі здібностей до групування- дані методи використовують алгоритми кластеризації для групування пікселів на зображенні з метою визначення об'єктів на зображенні. Прикладами є k-means clustering та Mean Shift clustering.

6. Методи на основі статистичних моделей- такі методи використовують статистичні моделі для опису об'єктів на зображенні та їхніх відносин. Ці методи зазвичай використовуються для сегментації зображень та класифікації об'єктів. Прикладами таких методів можна назвати HMMs (Hidden Markov Models), GMMs (Gaussian Mixture Models), CRFs (Conditional Random Fields).

На мою думку ці категорії можна назвати основними коли мова йде про методи розпізнавання зображень та їх групування за способами застосування та особливостями роботи.

Висновки до розділу 3

Методи розпізнавання зображень є досить розгалуженою галуззю машинного навчання, яка постійно розвивається та покращується. В залежності від застосування та поставленої задачі, можна обирати різні методи розпізнавання зображень.

Нейронні мережі є одними з найбільш ефективних методів для розпізнавання зображень. У той же час, їх треба тренувати за допомогою великих наборів даних для досягнення високої точності.

Крім того, існують методи, такі як метод головних компонентів або метод опорних векторів, які можуть бути ефективними в деяких випадках.

Також можна використовувати комбінацію різних методів з метою досягнення кращої точності розпізнавання. Наприклад, можна поєднувати нейронні мережі з іншими методами, такими як метод опорних векторів, щоб отримати більш точні результати.

У загальному, для вибору методу розпізнавання зображень необхідно враховувати поставлену задачу, обсяг даних та час, який можна витратити на навчання та розпізнавання. Також важливим фактором є наявність відкритих даних та програмного забезпечення, яке може допомогти в розпізнаванні зображень.

У багатьох випадках, залежно від поставленої задачі, можуть використовувати підхід «transfer learning», коли використовуються попередньо навчені нейронні мережі з інших сфер, які можуть бути застосовані до нової задачі з мінімальними модифікаціями.

Окрім того, важливим фактором є ефективність реалізації методу на різних пристроях та платформах, зокрема на мобільних пристроях з обмеженими обчислювальними ресурсами.

Загалом, для вибору оптимального методу розпізнавання зображень необхідно провести аналіз та порівняння різних методів на конкретній задачі та з врахуванням доступних ресурсів, з метою забезпечення максимальної точності та ефективного розпізнавання зображення.

Незважаючи на це, важливим є процес підготовки даних. Для досягнення високої точності розпізнавання зображень, необхідно мати достатньо великий набір даних для тренування моделі. Важливо також правильно підготувати дані перед тренуванням, зокрема, провести їх нормалізацію.

Також важливим є процес валідації та тестування моделі, який дозволяє оцінити її точність та ефективність на незалежному наборі даних. Для цього можна використовувати метрики, такі як точність, чутливість та специфічність.

Крім того, для досягнення кращих результатів можна використовувати ансамблі моделей, коли декілька моделей розпізнавання зображень використовуються разом для досягнення кращих результатів.

Отже, для досягнення високої точності та ефективності розпізнавання зображень важливо правильно вибрати метод розпізнавання, підготувати та нормалізувати дані перед тренуванням, провести валідацію та тестування моделі, а також можливо використати ансамбль моделей з метою досягнення кращих результатів.

ВИСНОВКИ

Дослідження методів розпізнавання зображень є важливим напрямком у сучасній науці та технологіях. Дані методи використовуються в різних галузях, таких як медицина, автоматизація виробництва, транспорт, безпека, реклама багато інших.

Одним із найбільш поширених методів є нейронні мережі. Вони здатні самостійно навчитися за допомогою великої кількості даних та розв'язувати складні задачі розпізнавання. Нейронні мережі можуть бути використані для розпізнавання обличчя, машин на дорозі, визначення стану здоров'я на основі зображень медичних досліджень та багатьох інших застосувань.

Ще одним методом є метод опорних векторів (SVM). Він використовується для розпізнавання об'єктів на зображеннях та має високу точність розпізнавання. Даний метод може бути використаний для розпізнавання елементів на зображеннях, таких як символи або графіки.

Також існують методи розпізнавання зображень, засновані на аналізі форми та текстурі об'єктів на зображенні. Ці методи використовуються для розпізнавання знаків дорожнього руху, штрих-кодів та інших елементів зображень.

Узагальнюючи, дослідження методів розпізнавання зображень є важливою та перспективною галуззю науки та технології. Використання цих методів може значно полегшити та покращити роботу людей в різних сферах життя.

Незважаючи на це, наразі існують деякі виклики та обмеження, пов'язані з розпізнаванням зображень. Одним з них є нестабільність розпізнавання в умовах зміни освітлення, тіней або перешкод на зображенні. Це може призводити до помилок у роботі систем розпізнавання зображень, що може бути критичним у деяких випадках, наприклад, у медичних дослідженнях.

Крім того, у деяких випадках для досягнення високої точності розпізнавання зображень необхідно використовувати велику кількість даних та потужних обчислювальних ресурсів. Це може бути проблематичним для деяких застосувань, особливо у випадках з обмеженими обчислювальними ресурсами. Це може бути проблематичним для деяких застосувань, особливо у випадках з обмеженими обчислювальними ресурсами.

У зв'язку з цим, постійне вдосконалення методів розпізнавання зображень та розвиток нових підходів є важливою задачею для дослідників та фахівців у галузі. Наприклад, застосування глибинного навчання та штучних нейронних мереж може допомогти вирішити деякі з цих викликів і поліпшити точність та швидкість роботи систем розпізнавання зображень.

Усі ці виклики та обмеження не зменшують значення досліджень методів розпізнавання зображень, а навпаки, стимулюють дослідників до подальшого вдосконалення та розвитку цієї важливої галузі науки та технології.

Одним із сучасних застосувань методів розпізнавання зображень є розвиток систем штучного інтелекту для робототехніки та автономних транспортних засобів. Розпізнавання зображень дозволяє роботам та автономним системам орієнтуватись у навколишньому середовищі та взаємодіяти з ним, що є важливою передумовою для їх ефективної роботи.

Також, методи розпізнавання зображень використовуються у багатьох інших галузях, наприклад, у медицині для діагностики захворювань та моніторингу пацієнтів, у безпеці для виявлення злочинців та відслідковування місць злочину, а також у відеоспостереженні та контролі якості виробництва.

Отже, дослідження методів розпізнавання зображень має велике значення у багатьох галузях технологій та науки, а науки, а його подальший розвиток може допомогти вирішити багато викликів та завдань, що стоять перед сучасним технологічними та науковими галузями.



[D1%96%D1%81%D1%82%D1%96%D0%B2-%D1%96%D0%B7-%D0%BE%D0%B1%D1%80%D0%BE%D0%B1%D0%BA%D0%B8-%D0%B4%D0%B0%D0%BD%D0%B8%D1%85/](https://uk.wikipedia.org/wiki/%D1%81%D1%82%D1%96%D0%B2-%D1%96%D0%B7-%D0%BE%D0%B1%D1%80%D0%BE%D0%B1%D0%BA%D0%B8-%D0%B4%D0%B0%D0%BD%D0%B8%D1%85/)

8) Кластерний аналіз:

https://uk.wikipedia.org/wiki/%D0%9A%D0%BB%D0%B0%D1%81%D1%82%D0%B5%D1%80%D0%BD%D0%B8%D0%B9_%D0%B0%D0%BD%D0%B0%D0%BB%D1%96%D0%B7

9) Використання нейронних мереж для розпізнавання зображень стр.

1/1: <https://conf.ztu.edu.ua/wp-content/uploads/2021/05/72-2.pdf>

10) Розробка програмного модуля розпізнавання зображень:

<http://inmad.vntu.edu.ua/portal/static/5C2FEBDD-5A85-4056-A68A-9BF7FE01995D.pdf>

11) Завдання розпізнавання образів:

https://uk.wikipedia.org/wiki/%D0%97%D0%B0%D0%B2%D0%B4%D0%B0%D0%BD%D0%BD%D1%8F_%D1%80%D0%BE%D0%B7%D0%BF%D1%96%D0%B7%D0%BD%D0%B0%D0%B2%D0%B0%D0%BD%D0%BD%D1%8F_%D0%BE%D0%B1%D1%80%D0%B0%D0%B7%D1%96%D0%B2

12) Ієрархічна кластеризація:

https://uk.wikipedia.org/wiki/%D0%86%D1%94%D1%80%D0%B0%D1%80%D1%85%D1%96%D1%87%D0%BD%D0%B0_%D0%BA%D0%BB%D0%B0%D1%81%D1%82%D0%B5%D1%80%D0%B8%D0%B7%D0%B0%D1%86%D1%96%D1%8F

13) Розробка та створення нейронної мережі:

<https://evergreens.com.ua/ua/development-services/neural-network.html>

14) Згортова нейронна мережа:

https://uk.wikipedia.org/wiki/%D0%97%D0%B3%D0%BE%D1%80%D1%82%D0%BA%D0%BE%D0%B2%D0%B0_%D0%BD%D0%B5%D

https://uk.wikipedia.org/wiki/%D0%B9%D1%80%D0%BE%D0%BD%D0%BD%D0%B0_%D0%BC%D0%B5%D1%80%D0%B5%D0%B6%D0%B0

15) Штучна нейронна мережа:

https://uk.wikipedia.org/wiki/%D0%A8%D1%82%D1%83%D1%87%D0%BD%D0%B0_%D0%BD%D0%B5%D0%B9%D1%80%D0%BE%D0%BD%D0%BD%D0%B0_%D0%BC%D0%B5%D1%80%D0%B5%D0%B6%D0%B0

16) Метод опорних векторів:

https://uk.wikipedia.org/wiki/%D0%9C%D0%B5%D1%82%D0%BE%D0%B4_%D0%BE%D0%BF%D0%BE%D1%80%D0%BD%D0%B8%D1%85_%D0%B2%D0%B5%D0%BA%D1%82%D0%BE%D1%80%D1%96%D0%B2#:~:text=%D0%92%20%D0%BC%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%BE%D0%BC%D1%83%20%D0%BD%D0%B0%D0%B2%D1%87%D0%B0%D0%BD%D0%BD%D1%96%20%D0%BC%D0%B5%CC%81%D1%82%D0%BE%D0%B4%20%D0%BE%D0%BF%D0%BE%CC%81%D1%80%D0%BD%D0%B8%D1%85,%D0%B2%D0%B5%CC%81%D0%BA%D1%82%D0%BE%D1%80%D0%BD%D0%B8%D0%BC%D0%B8%20%D0%BC%D0%B0%D1%88%D0%B8%CC%81%D0%BD%D0%B0%D0%BC%D0%B8%20%D0%9E%D0%92%D0%9C%2C%20%D0%B0%D0%BD%D0%B3%D0%BB

17) Лигун А.О. Комп'ютерна графіка (Обробка та стиск зображень).

Дніпропетровськ: Біла К.О., 2010. 114 с.

18) Веб-сайт OpenCV. URL: <http://opencv.org/>

19) <http://apeps.kpi.ua/neural-networks/en>

20) <https://livingfo.com/shcho-take-nejronni-merezhi-ta-iak-vony-pratsiuiut/>

ДОДАТОК А

Нейронна мережа, яку ми побудували, працює на основі штучних нейронів, які моделюють функціонування нейронів у людському мозку. Основна ідея полягає в тому, що мережа навчається знаходити корисні ознаки або шаблони у вхідних даних, які допомагають їй робити правильні прогнози або класифікацію.

У нашій моделі ми використовуємо згортковий шар (Flatten) для перетворення двовимірного зображення в одновимірний вектор. Цей шар допомагає зберегти просторову структуру зображення та передати його наступним шарам.

Далі ми маємо повнозв'язний (Dense) шар з 128 нейронами та активаційною функцією ReLU (Rectified Linear Unit). Цей шар виконує операцію лінійної комбінації вхідних значень і застосовує нелінійну активаційну функцію, що допомагає моделі набути нелінійність і здатність до складніших відносин між ознаками.

Останній шар має 10 нейронів, що відповідають 10 можливим класам цифр (0-9). Використовується активаційна функція Softmax, яка перетворює вихідні значення в ймовірності розпізнавання кожного класу.

Після побудови моделі ми компілюємо її, встановлюючи оптимізатор (adam), функцію втрат (sparse categorical crossentropy) і метрику (accuracy) для оцінки її продуктивності під час тренування.

Після тренування моделі на тренувальних даних ми оцінюємо її точність на тестових даних, використовуючи метод evaluate.

Наша нейронна мережа в результаті виглядає так:

```
import tensorflow as tf
```

```
from tensorflow.keras.datasets import mnist
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Flatten

# Завантаження датасету MNIST
(x_train, y_train), (x_test, y_test) = mnist.load_data()

# Нормалізація даних
x_train = x_train / 255.0
x_test = x_test / 255.0

# Побудова моделі нейронної мережі
model = Sequential()
model.add(Flatten(input_shape=(28, 28))) # Згладжуємо вхідні зображення
model.add(Dense(128, activation='relu'))
model.add(Dense(10, activation='softmax')) # Вихідний шар з 10 нейронів для
розпізнавання 10 класів цифр (0-9)

# Компіляція моделі
model.compile(optimizer='adam', loss='sparse_categorical_crossentropy',
metrics=['accuracy'])

# Тренування моделі
model.fit(x_train, y_train, epochs=10, validation_data=(x_test, y_test))

# Оцінка точності моделі на тестових даних
test_loss, test_accuracy = model.evaluate(x_test, y_test)
```



```
print("Точність на тестових даних:", test_accuracy)
```

```
Python 3.11 (64-bit)
Epoch 9/10
1875/1875 [-----] - 3s 2ms/step - loss: 0.0208 - accuracy: 0.9936 - val_loss: 0.0877 - val_accuracy: 0.9765
Epoch 10/10
1875/1875 [-----] - 3s 2ms/step - loss: 0.0170 - accuracy: 0.9947 - val_loss: 0.0792 - val_accuracy: 0.9802
<keras.callbacks.History object at 0x000001A101787508>
>>> # Оцінка точності моделі на тестових даних
>>> test_loss, test_accuracy = model.evaluate(x_test, y_test)
313/313 [-----] - 0s 1ms/step - loss: 0.0792 - accuracy: 0.9802
>>> print("Точність на тестових даних:", test_accuracy)
Точність на тестових даних: 0.9801999920507878
>>>
>>> import tensorflow as tf
>>> from tensorflow.keras.datasets import mnist
>>> from tensorflow.keras.models import Sequential
>>> from tensorflow.keras.layers import Dense, Flatten
>>>
>>> # Завантаження даних mnist
>>> (x_train, y_train), (x_test, y_test) = mnist.load_data()
>>>
>>> # Нормалізація даних
>>> x_train = x_train / 255.0
>>> x_test = x_test / 255.0
>>>
>>> # Побудова моделі нейронної мережі
>>> model = Sequential()
>>> model.add(Flatten(input_shape=(28, 28))) # Згладжуємо вхідні зображення
>>> model.add(Dense(128, activation='relu'))
>>> model.add(Dense(10, activation='softmax')) # Вихідний шар з 10 нейронів для розпізнавання 10 класів цифр (0-9)
>>>
>>> # Компіляція моделі
>>> model.compile(optimizer='adam', loss='sparse_categorical_crossentropy', metrics=['accuracy'])
>>>
>>> # Тренування моделі
>>> model.fit(x_train, y_train, epochs=10, validation_data=(x_test, y_test))
Epoch 1/10
1875/1875 [-----] - 4s 2ms/step - loss: 0.2536 - accuracy: 0.9282 - val_loss: 0.1507 - val_accuracy: 0.9519
Epoch 2/10
1875/1875 [-----] - 3s 2ms/step - loss: 0.1117 - accuracy: 0.9672 - val_loss: 0.0981 - val_accuracy: 0.9703
Epoch 3/10
1875/1875 [-----] - 3s 2ms/step - loss: 0.0769 - accuracy: 0.9770 - val_loss: 0.0812 - val_accuracy: 0.9729
Epoch 4/10
1875/1875 [-----] - 3s 2ms/step - loss: 0.0561 - accuracy: 0.9829 - val_loss: 0.0815 - val_accuracy: 0.9751
Epoch 5/10
1875/1875 [-----] - 3s 2ms/step - loss: 0.0435 - accuracy: 0.9867 - val_loss: 0.0791 - val_accuracy: 0.9748
Epoch 6/10
1875/1875 [-----] - 3s 2ms/step - loss: 0.0339 - accuracy: 0.9894 - val_loss: 0.0837 - val_accuracy: 0.9736
Epoch 7/10
1875/1875 [-----] - 3s 2ms/step - loss: 0.0275 - accuracy: 0.9913 - val_loss: 0.0744 - val_accuracy: 0.9787
Epoch 8/10
1875/1875 [-----] - 3s 2ms/step - loss: 0.0212 - accuracy: 0.9935 - val_loss: 0.0771 - val_accuracy: 0.9788
Epoch 9/10
1875/1875 [-----] - 3s 2ms/step - loss: 0.0186 - accuracy: 0.9941 - val_loss: 0.0739 - val_accuracy: 0.9798
Epoch 10/10
1875/1875 [-----] - 3s 2ms/step - loss: 0.0147 - accuracy: 0.9956 - val_loss: 0.0788 - val_accuracy: 0.9795
<keras.callbacks.History object at 0x000001A103088218>
>>>
>>> # Оцінка точності моделі на тестових даних
>>> test_loss, test_accuracy = model.evaluate(x_test, y_test)
313/313 [-----] - 0s 1ms/step - loss: 0.0780 - accuracy: 0.9795
>>> print("Точність на тестових даних:", test_accuracy)
Точність на тестових даних: 0.9794999957884656
>>>
```

Ми використовуємо одношарову нейронну мережу з одним прихованим шаром з 128 нейронами та активаційною функцією ReLU. Останній шар має 10 нейронів (які відповідають 10 класам цифр) і використовує активаційну функцію Softmax для отримання ймовірностей розпізнавання кожного класу. Модель тренується протягом 10 епох, і після тренування оцінюється її точність на тестових даних.

Процес навчання нейронної мережі полягає в оптимізації внутрішніх ваг моделі, з метою зменшити втрату і досягти кращої точності прогнозування. Це досягається шляхом пошуку оптимальних значень ваг за допомогою алгоритму оптимізації (у нашому випадку- алгоритм Adam).

Таким чином, навчання нейронної мережі полягає у підборі внутрішніх параметрів (ваг) моделі таким чином, щоб зменшити втрату на тренувальних

даних і покращити її здатність до розпізнавання чисел на нових, раніше невиданих даних.

Для того, щоб додати зображення для обробки нейронною мережею, нам необхідно мати зображення у форматі, який може бути прочитаний та оброблений бібліотекою машинного навчання (TensorFlow або Keras).

Перед тим, як додавати нове зображення, переконайтесь, що ви перетворили його до формату та розміру, який вимагає модель. У випадку розпізнавання чисел ми використовували розмір зображень 28x28 пікселів та значення пікселів, нормалізовані в діапазоні від 0 до 1.

Одним із способів додавання зображення є завантаження зображення з файлу та обробка його для використання в моделі. Наприклад, якщо у нас є зображення у форматі JPEG, ви можете використати наступний код:

```
from PIL import Image
import numpy as np

# Завантаження зображення з файлу
image = Image.open("Цифри")

# Перетворення розміру зображення до 28x28 пікселів
resized_image = image.resize((28, 28))

# Конвертування зображення в чорно-білий формат
bw_image = resized_image.convert("L")
```



```
# Нормалізація значень пікселів у діапазон від 0 до 1
```

```
normalized_image = np.array(bw_image) / 255.0
```

```
# Приклад: передбачення класу цифри на новому зображенні
```

```
predictions = model.predict(np.expand_dims(normalized_image, axis=0))
```

```
predicted_class = np.argmax(predictions)
```

```
predicted_class_name = str(predicted_class)
```

```
print("Передбачений клас цифри:", predicted_class_name)
```

У цьому прикладі ми використовуємо бібліотеку PIL для завантаження зображення з файлу. Потім ми змінюємо розмір зображення до 28x28 пікселів, конвертуємо його в чорно-білий формат та нормалізуємо значення пікселів до діапазону від 0 до 1. Нарешті, ми передаємо зображення до моделі за допомогою методу «predict» для отриманого передбачення класу

