

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)

Кафедра Штучного інтелекту
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

Аналіз алгоритмів розпізнавання образів нейронною мережею
(тема)

Виконав:
студент 2 курсу, групи СШМ-22-3
Панченко Т. І.
(прізвище, ініціали)

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

Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)

Освітня програма Системи штучного інтелекту
(повна назва спеціалізації)

Керівник доц. Носова Я.В.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____
(підпис)

В.О. Філатов
(прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)
Кафедра Штучного інтелекту
(повна назва)
Рівень вищої освіти другий (магістерський)
Спеціальність 122 Комп'ютерні науки
(код і повна назва)
Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)
Освітня програма Системи штучного інтелекту
(повна назва)

ЗАТВЕРДЖУЮ:
Зав. кафедри _____
(підпис)
«_____» _____ 20__ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Панченку Тимуру Ігоровичу
(прізвище, ім'я, по батькові)

1. Тема роботи Аналіз алгоритмів розпізнавання образів нейронною мережею

затверджена наказом університету від 1 квітня 20 24 р. № 260Ст

2. Термін подання студентом роботи до екзаменаційної комісії 14 червня 20 24 р.

3. Вихідні дані до роботи Науково-технічні публікації, дані Інтернет джерел та відомих наукових проектів, електронні документації програмне забезпечення IDE-засіб PyCharm. Мова програмування Python, бібліотеки для роботи з даними TensorFlow та PyTorch

4. Перелік питань, що потрібно опрацювати в роботі _____

1) Аналіз предметної області та постановка задачі

2) Теоретичне дослідження обраної проблеми

3) Практична реалізація

РЕФЕРАТ

Пояснювальна записка: 74 с., 17 рис., 2 табл., 2 дод., 19 джерел.

ГЛИБОКІ НЕЙРОННІ МЕРЕЖІ, КЛАСИФІКАЦІЯ, КЛАСТЕРИЗАЦІЯ, КЛІТИННА ІНЖЕНЕРІЯ, МАШИННЕ НАВЧАННЯ, РОЗПІЗНАВАННЯ ОБРАЗІВ, ШТУЧНИЙ ІНТЕЛЕКТ.

Об'єкт дослідження – алгоритми розпізнавання образів нейронною мережею, що забезпечують комп'ютерну систему здатністю автоматичного розпізнавання та класифікації об'єктів на основі вхідних зображень або даних.

Предмет дослідження – процес аналізу та порівняння різних алгоритмів розпізнавання образів нейронною мережею задля визначення їхньої ефективності, точності та швидкодії в різних умовах та на різних наборах даних, а також можливості їхнього використання в сучасних реаліях та умовах.

Мета роботи – дослідження та порівняння різних алгоритмів розпізнавання образів нейронною мережею з метою визначення їхньої ефективності та придатності для застосування в різних сферах. Дослідження спрямоване на встановлення можливостей використання цих алгоритмів у сучасних реаліях, зокрема, у сферах комп'ютерного зору, медичного діагностування та багатьох інших областях, де розпізнавання образів є важливим етапом аналізу даних.

Методи дослідження – теоретичний аналіз основ нейронних мереж та їхніх алгоритмів. Проведення серії емпіричних досліджень на різних наборах даних, статистичний аналіз отриманих результатів з використанням відповідних метрик та висунення рекомендацій щодо вибору оптимальних алгоритмів для конкретних завдань розпізнавання образів.

ABSTRACT

Master's thesis contains: 74 pp., 17 fig., 2 tabl., 2 ann., 17 references.

ARTIFICIAL INTELLIGENCE, CELLULAR ENGINEERING, CLASSIFICATION, CLUSTERING, DEEP NEURAL NETWORKS, MACHINE LEARNING, PATTERN RECOGNITION.

Research object – neural network pattern recognition algorithms that provide a computer system with the ability to automatically recognize and classify objects based on input images or data.

Research subject – the process of analyzing and comparing different neural network pattern recognition algorithms to determine their efficiency, accuracy and speed in different conditions and on different data sets, as well as the possibility of using them in modern realities and conditions.

Purpose of the study – research and comparison of various pattern recognition algorithms by neural network to determine their efficiency and suitability for use in various fields. The research is aimed at establishing the possibilities of using these algorithms in modern realities, in particular, in the areas of computer vision, medical diagnostics, and many other areas where pattern recognition is an important stage of data analysis.

Research methods – theoretical analysis of the basics of neural networks and their algorithms. Conducting a series of empirical studies on different datasets, statistical analysis of the results using appropriate metrics, and making recommendations for choosing the best algorithms for specific pattern recognition tasks.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів	7
Вступ.....	8
1 Аналіз предметної галузі та постановка задачі.....	10
1.1 Аналіз проблеми і постановка задачі.....	10
1.2 Технологія розпізнавання образів	14
1.3 Існуючі системи розпізнавання образів у практичних сферах.....	19
1.4 Алгоритми розпізнавання образів	21
1.5 Висновки та постановка задачі	23
2 Теоретичне дослідження обраної проблеми.....	25
2.1 Теоретичний аналіз	25
2.2 Типи нейронних мереж.....	30
2.3 Детальний огляд розпізнавання образів	34
2.4 Сучасні рішення та алгоритми розпізнавання образів.....	39
3 Практична реалізація	44
3.1 Опис середовища розробки та інструментів	44
3.2 Опис набору даних.....	47
3.3 Програмна реалізація.....	48
3.4 Результати та їх аналіз.....	53
Висновки	57
Перелік джерел посилання	59
Додаток А Код програми.....	61
Додаток Б Відомість кваліфікаційної роботи.....	74

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ШІ – штучний інтелект;

AI – Artificial Intelligence – штучний інтелект;

API – Application Programing interface – набір визначень програми;

DNN – Deep Neural Networks – глибокі нейронні мережі;

IOT – Internet Of Things – інтернет речей;

ML – Machine Learning – машинне навчання;

PR – Pattern Recognition – розпізнавання образів.

ВСТУП

Людство переживає стрімкий розвиток, що супроводжується безперервними та динамічними змінами у світі. Рівень прогресу не однаковий у різних країнах, де можна спостерігати значні розбіжності. Деякі з них досі стикаються з відсутністю базових зручностей, таких як туалети в будинках або якісні дороги у сільській місцевості. Проте, загалом, модернізація охоплює всі сфери нашого життя, стаючи помітною навіть на перший погляд.

Одним із ключових методів, що робить наше життя комфортнішим, є використання штучних нейронних мереж. Ці технології, відомі як штучний інтелект (ШІ), виникли у 1950-х роках і за короткий час досягли вражаючих результатів. Незважаючи на те, що ми ще не до кінця розуміємо принципи роботи ШІ, неможливо з абсолютною впевненістю передбачити його майбутнє. Тому важливо не робити прогнозів, а зосередитися на актуальних проблемах розробки та навчання систем з штучним інтелектом.

На початку свого розвитку дослідження в цій галузі потребували значних фінансових витрат. Збільшення фінансування, яке відбулося після досягнення певних результатів, дозволяло дослідницьким центрам нарощувати обчислювальну потужність, яка є ключовим фактором успіху. Сьогодні, коли практично кожен має доступ до комп'ютера, а обсяги даних, доступних для навчання, значно зросли, розвиток ШІ прискорюється стрімкими темпами.

Розробка та вдосконалення алгоритмів розпізнавання образів – це динамічна та складна область досліджень. Існують певні виклики, з якими стикаються дослідники: складність візуального світу, варіативність зображень, обчислювальна складність.

Однією з найпоширеніших задач, для вирішення яких використовуються підходи та алгоритми ШІ, є розпізнавання образів. Цей метод аналізу даних використовує алгоритми машинного навчання для

автоматичного розпізнавання шаблонів і закономірностей у даних. Системи розпізнавання образів здатні точно розпізнавати знайомі об'єкти за певними характеристиками, а також класифікувати невідомі об'єкти, навіть якщо вони частково розмиті або затемнені.

Дана робота концентрує у собі декілька цілей, які так або інакше пов'язані з темою розпізнавання образів: дослідження сучасного стану роботи систем розпізнавання образів та відображення процесу виконання поставлених задач; пошук прикладів, пошук існуючих способів модернізації та практичного використання нейронних мереж та їхніх алгоритмів в обраній галузі.

1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ ТА ПОСТАНОВКА ЗАДАЧІ

1.1 Аналіз проблеми і постановка задачі

Дані стали рушійною силою прогресу у 21 столітті. Безліч різноманітних типів даних, таких як Big Data, Data Science, Data Analytics, Data Warehouse, Intelligent Data Analysis, свідчать про те, що дані відіграють ключову роль у впливі на повсякденне життя людства.

Ми живемо в епоху інформації, де дані генеруються з неймовірною швидкістю в усіх сферах життя: від бізнесу та науки до медицини та особистого життя. Цей масив інформації стає необхідним ресурсом для прийняття обґрунтованих рішень, кращого розуміння світу та розвитку нових технологій.

У сучасному світі, де динаміка технологічного прогресу надзвичайно швидка, штучний інтелект (AI) та нейронні мережі стають все більш важливими інструментами для розв'язання різноманітних завдань, зокрема розпізнавання образів, яке можна використати у життєво необхідних сферах, таких як медицина.

Предметна область була обрана по дослідженню задачі візуалізації процесу та результатів роботи систем розпізнавання образів за допомогою використання нейронних мереж. Велика кількість різноманітних видів даних, таких як Big Data, Data Science, Data Analytics, Data Warehouse, Intelligent Data Analysis, тощо, підкреслюють, що в нинішню епоху дані відіграють важливу роль у впливі на повсякденну діяльність людства. Щодня ми генеруємо понад 2,5 квінтильйони (10^{18}) байт даних, починаючи від наших текстових повідомлень, зображень з котиками, мемів про русню, електронних листів до даних з автономних механізмів, пристроїв ІОТ тощо [1], [2]. З такою кількістю даних у вас під рукою, використовуючи корисну інформацію з цих даних можна значно допомогти кожній організації отримати чітке розуміння багатьох сфер, наприклад, що може

збільшити дохід їхньої організації, якій області потрібно більше ресурсів, як привернути більше клієнтської уваги тощо. Машинне навчання (ML), наука про дані – це один із основних структурних компонентів систем штучного інтелекту (ШІ). Завдання навчання на основі даних вирішується у величезній мірі за допомогою цих систем в останні дні.

Одним із важливих аспектів використання нейронних мереж у медицині є їхня роль у розпізнаванні патологій на зображеннях медичних обстежень. Наприклад, за допомогою глибоких нейронних мереж можна автоматизувати виявлення раку на зображеннях рентгенівських знімків, МРТ або КТ. Такі системи можуть надзвичайно поліпшити швидкість та точність діагностики, дозволяючи швидше виявляти патологічні зміни та починати лікування на ранніх стадіях, коли воно є найефективнішим. Крім того, нейронні мережі можуть бути використані для прогнозування підходів до лікування та реакції пацієнтів на певні терапевтичні методи. Це може бути особливо корисно в онкології, де прогнозування ефективності лікування та виявлення можливих ускладнень може значно підвищити успішність терапії та покращити результати для пацієнтів.

Додатково, нейронні мережі можуть бути використані для аналізу медичних даних та розробки індивідуальних підходів до лікування. Наприклад, вони можуть допомогти в підборі оптимальних доз лікарських засобів для кожного пацієнта з урахуванням їхньої індивідуальної відповіді на лікування та фізіологічних особливостей. Не варто забувати й про можливості нейронних мереж у візуалізації даних. Однією з ключових переваг використання нейронних мереж у відображенні клітин є їхня здатність до автоматичного виявлення складних патологічних змін, які можуть бути важко розпізнати вручну. Це дозволяє підвищити точність та надійність аналізу клітинних зразків, зменшуючи ризик помилкових діагнозів.

У випадках, коли кількість клітинних зразків є дуже великою, застосування нейронних мереж може значно зекономити час та зусилля

медичних фахівців, забезпечуючи швидку та ефективну обробку даних. Правильне відображення клітин може сприяти розвитку досліджень у багатьох медичних сферах: неврологія, кардіологія, гастроентерологія, дерматологія тощо (рисунок 1.1). В аналізі нейронних структур та клітин можна виявити ознаки нейродегенеративних захворювань, таких як більцефалія, хвороба Паркінсона, болезнь Альцгеймера. Візуалізація клітин м'язів серця дозволяє виявляти аномалії, такі як фіброз, запалення та інші патології, які можуть впливати на функціонування серця. Аналіз клітин та тканин шлунка, кишечника та інших органів шлунково-кишкового тракту допомагає виявляти хвороби, такі як хвороба Крона, виразки. Візуалізація клітин шкіри може допомогти виявити ознаки різних шкірних захворювань, таких як дерматит, псоріаз, екзема, а також визначити вид пухлин шкіри. Вона допомагає вивчати генетичні аномалії, виявляти хромосомні відхилення та інші генетичні патології, а також використовується в дослідженні структури сперми та яйцеклітин в репродуктивній медицині.

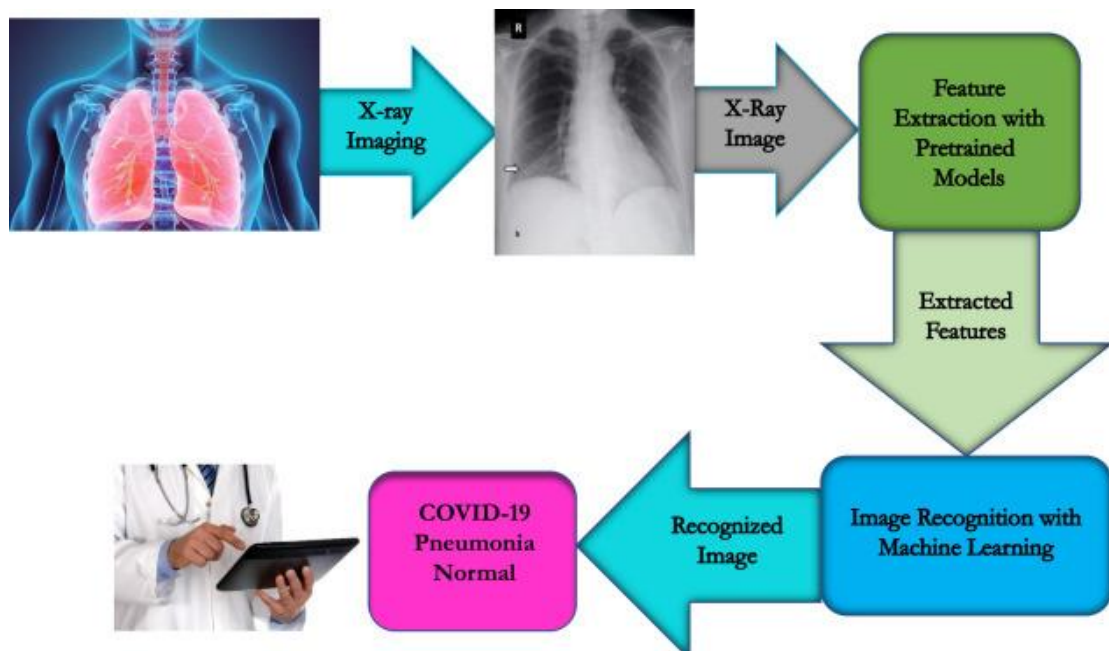


Рисунок 1.1 – Використання штучного інтелекту для діагностування

COVID-19

Проте, незважаючи на значний прогрес в цій області, існують деякі проблеми та виклики, що потребують уваги та дослідження. Однією з цих проблем є складність та ресурсомісткість алгоритмів розпізнавання образів нейронною мережею, особливо при роботі з великими обсягами даних. Нейронні мережі зазвичай потребують значних обчислювальних потужностей для тренування та роботи в реальному часі, що може бути обмежувальним фактором для їхнього застосування у практичних сценаріях.

Ще однією проблемою є нестабільність та непередбачуваність результатів нейронних мереж у деяких умовах, зокрема при роботі з неповними або шумними даними. Це може призвести до неправильного розпізнавання образів або до виникнення помилок у класифікації, що робить алгоритми менш ефективними у реальних умовах експлуатації. Розпізнавання образів включає два основні методи класифікації: контрольована та неконтрольована. У методі керованого розпізнавання людина навчає комп'ютерний алгоритм розпізнавати шаблони на основі заздалегідь визначених маркованих наборів даних. Після того, як шаблон ідентифіковано, метод класифікує нові дані. При неконтрольованій класифікації модель навчається самостійно, без прямих вказівок людини. Комп'ютерний алгоритм визначає кореляції між кількома елементами даних (вхідними даними) на основі їхньої схожості та виконує класифікацію даних.

Поставленням задачі є розробка та вдосконалення алгоритмів розпізнавання образів нейронною мережею з метою зменшення їхньої складності та ресурсомісткості, а також підвищення стабільності та надійності результатів. Це включає в себе дослідження нових архітектур мереж, оптимізацію параметрів навчання, розробку методів обробки шумних даних та використання передових технологій, щоб забезпечити оптимальну працездатність алгоритмів у реальному часі.

1.2 Технологія розпізнавання образів

Pattern recognition – це процес ідентифікації та аналізу даних на основі визначених критеріїв або загальних характеристик, що виконується за допомогою спеціальних алгоритмів. Розпізнавання образів є важливою функцією як для біологічних, так і для штучних систем. Біологічні організми використовують розпізнавання образів для виживання, реагуючи на певні стимули або об'єкти відповідно до своїх потреб. Наприклад, антитіла в організмі реагують на чужорідні агенти, вуха сприймають звук, а тварини впізнають їжу та небезпеку.

У штучних системах розпізнавання образів використовується для різних цілей, включаючи класифікацію даних для екологічного моніторингу, розпізнавання символів у тексті, біометричне ідентифікування осіб за їхнім зовнішнім виглядом, а також для навігації роботів та інших маніпуляцій. Коли мова йде про візуальні образи, розпізнавання образів доповнює комп'ютерний зір, дозволяючи аналізувати та класифікувати візуальні дані. Наприклад, це може бути використано для інтерпретації медичних зображень, таких як ЕКГ, або для розпізнавання обличчя для біометричної ідентифікації. Таким чином, розпізнавання образів є важливим елементом для багатьох сфер, включаючи біологію, медицину та технології, і відіграє ключову роль у виявленні, класифікації та реагуванні на різноманітні стимули та об'єкти.

Системи розпізнавання образів включають чотири основні складові: збір даних, вилучення ознак, порівняння та класифікація образів, а також оцінка їхньої ефективності. Вони проявляють високу ефективність в адаптації та навчанні, потребуючи гнучкої пам'яті, здатної навчатися на прикладах і класифікувати їх на основі схожості. Основним підходом до побудови систем розпізнавання образів є предикативне навчання. Проблема розпізнавання образів включає в себе статистичну регресію, де шукається функція, яка мінімізує помилку класифікації. У цій системі представлення

даних вимагає специфікації набору функцій для вибору, відомих як "ознаки", та індуктивного принципу для розробки та моделювання класифікатора. Процедура оптимізації (навчання) визначає класифікатор на основі оцінки параметрів. Індуктивний принцип є ключовим, оскільки він визначає, яким чином обробляти навчальні дані для навчання класифікатора.

З іншого боку, процедура навчання представляє конкретну реалізацію індуктивного принципу, яка включає оптимізацію або оцінку параметрів для вибраного набору ознак. Наприклад, це може бути застосування нейронних мереж прямого поширення з сигмоїдними блоками або дерева рішень. Розпізнавання образів можна розглядати як один з ключових елементів комп'ютерного зору, що спрямований на емуляцію функцій людського мозку. Уявіть собі це так: передбачення на основі невидимих даних стає можливим завдяки здатності моделей виявляти повторювані патерни. Цей процес може застосовуватися до будь-яких типів даних, включаючи зображення, відео, текст, аудіо та інше.

Розпізнавання образів, яке є складним за своєю природою, включає в себе аналіз вхідних даних, виявлення закономірностей та порівняння їх зі збереженими даними. Цей процес можна розділити на дві фази: дослідницьку, коли алгоритми аналізують дані, щоб виявити закономірності, коли алгоритми групують та класифікують виявлені закономірності відповідно до вихідних даних. Якщо розділити цю процедуру ще більше, то розпізнавання образів у машинному навчанні включає наступний шлях (рисунок 1.2).

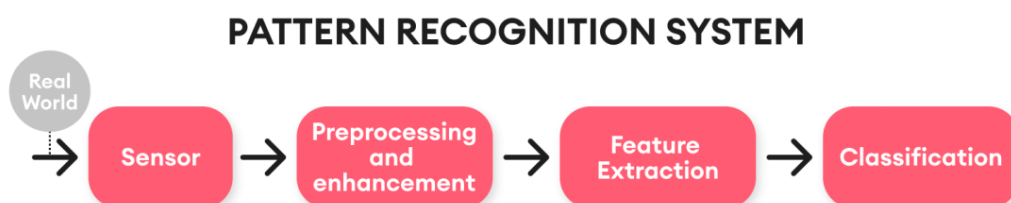


Рисунок 1.2 – Розуміння розпізнавання образів у МН

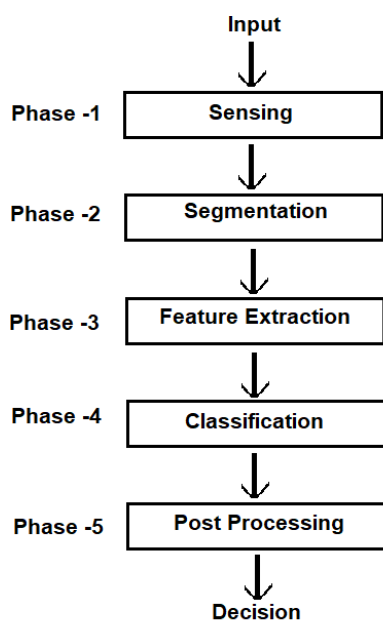
Одним із найскладніших етапів у розпізнаванні образів є вибір алгоритмів, які плануєте використовувати. Загально відомі алгоритми розпізнавання включають статистичний метод, який використовує статистичні методи для навчання на прикладах. Цей підхід передбачає аналіз спостережень та розробку правил, які можуть бути використані для майбутніх даних. Структурний – статистичний метод не найкраще підходить для розпізнавання складних образів. Саме тут на допомогу приходить структурне розпізнавання з його ієрархічним підходом та категоризацією на підкласи.

Модель описує складні взаємозв'язки між безліччю елементів і є таких цілей, як аналіз зображень і форм, де підтверджуються вимірні структури. Нейронні мережі – це метод, що базується на використанні штучних нейронних мереж і відрізняється більшою гнучкістю порівняно з традиційними алгоритмами. Вони ефективно працюють у завданнях класифікації, використовуючи концепції, взяті з біології для розпізнавання образів. Особливо ефективним методом у розпізнаванні образів є використання мереж зі зворотнім зв'язком, де навчання відбувається шляхом зворотного зв'язку з вхідними зразками.

Поєднання шаблонів – використовується під час роботи з двома сутностями одного типу. У цьому випадку цільовий зразок зіставляється з шаблоном, що зберігається в пам'яті, де визначається схожість між такими об'єктами, як криві, фігури і т. д. Однак, цей метод вимагає надмірної кількості шаблонів і є досить жорстким у порівнянні з існуючими альтернативами. Нечіткий – у реальних завданнях розпізнавання нечіткість (багатозначна логіка, в якій істинні значення змінних може бути будь-яким дійсним числом від 0 до 1) широко поширена, що значною мірою обумовлено нашою когнітивною системою. Найчастіше ми стикаємося з невизначеними компонентами для сканування об'єктів для розпізнавання за допомогою нашої візуальної системи. Це справедливо і для цифрового світу, що пояснює широке застосування даного алгоритму.

Гібридна модель – зазвичай визначає комбінацію різних типів алгоритмів для розгортання переваг всіх використовуваних методів. Вона розпізнає закономірності з допомогою кількох класифікаторів, кожен із яких навчається з урахуванням просторів ознак. Висновок робиться з урахуванням накопичення наборів класифікаторів, точність яких визначається з допомогою функції прийняття рішення. Окрім цього, для більшого розуміння процесів, що відбуваються під час роботи систем розпізнавання образів, можна привести фази та відповідні дії під час цих фаз (рисунок 1.3). Ці системи можна розділити на такі компоненти:

- перетворення зображень, звуків або інших вхідних даних на сигнальні дані;
- виділення об'єктів з фону;
- вимірювання властивостей об'єктів, корисних для класифікації;
- надання об'єкту категорії;
- прийняття інших міркувань для ухвалення рішення про відповідні дії.

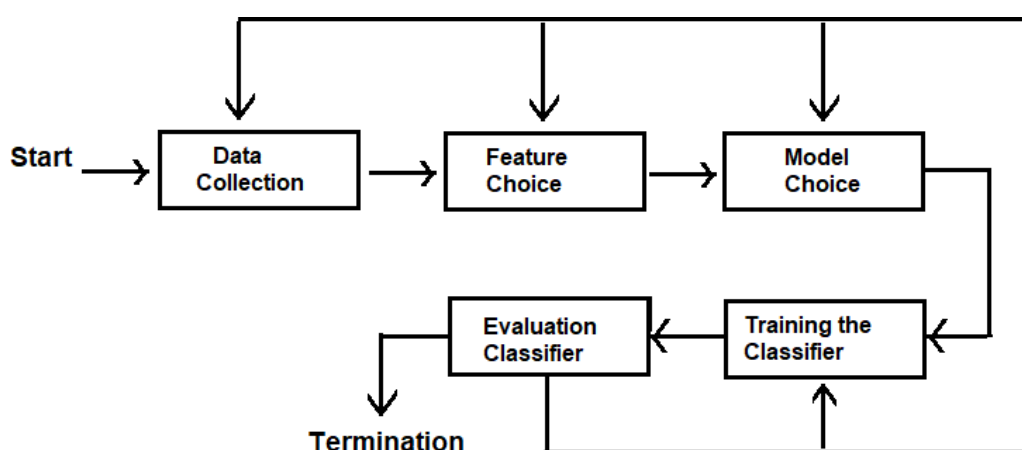


Phases in Pattern Recognition

Рисунок 1.3 – Фази роботи систем розпізнавання образів

Можна відокремити певну кількість проблем, що вирішує кожна з фаз. Під час першої фази, ми маємо справу з проблемами, що виникають на вході, такими як пропускна спроможність, роздільна здатність, чутливість, спотворення, відношення сигнал/шум, затримка і т. д. На другому етапі у нас найглибша проблема розпізнавання образів, яка пов'язана з проблемою розпізнавання або угрупування різних частин об'єкта. Третій етап має справу з характеристикою об'єкта, щоб його можна було легко розпізнати за допомогою вимірів. Ті об'єкти, значення яких дуже схожі для об'єктів, вважаються такими, що знаходяться в одній категорії, в той час як ті, значення яких сильно відрізняються для об'єктів, поміщаються в різні категорії. Четверта фаза полягає у віднесенні об'єктів до певних категорій з використанням вектора ознак, наданого екстрактором ознак, та визначення значень всіх ознак для конкретного входу. П'ята ж пов'язана з прийняттям рішень про дії з використанням результатів класифікатора.

Такі дії, як класифікація з мінімальним коефіцієнтом помилок мінімізують загальну очікувану вартість. На останньому пункті цього підрозділу можна зобразити основи, що використовуються у розпізнаванні образів (рисунок 1.4).



Activity Cycle

Рисунок 1.4 – Перелік різних видів діяльності

1.3 Існуючі системи розпізнавання образів у практичних сферах

Ера штучних нейронних мереж (ШНМ) почалася зі спрощеного застосування в багатьох галузях і досягла значного успіху в розпізнаванні образів (РО) навіть у сферах від розваг до виробництва. Незважаючи на значний прогрес, досягнутий і досліджений у застосуванні ШНМ для вирішення завдань розпізнавання образів, деякі проблеми все ще залишаються невирішеними, наприклад, примхлива орієнтація (невідомий шлях, який не може бути точно розрахований через його положення в напрямку). Інші проблеми включають: класифікацію об'єктів, розташування, масштабування, аналіз поведінки нейронів у прихованих шарах, узгодження правил і шаблонів.

Крім того, брак літератури з питань, пов'язаних із застосуванням ШНМ, сповільнює фокус досліджень і прогрес у цій галузі. Таким чином, існує потреба в сучасному застосуванні нейронних мереж, щоб терміново вирішити вищезазначені проблеми для досягнення більшого успіху. Крім того, всебічний огляд розкриває різноманітні сфери успіху моделей штучного інтелекту та їх застосування в бізнесі. Штучні нейронні мережі (ШНМ) – це нелінійні статистичні моделі даних, які повторюють роль біологічних нейронних мереж.

Підхід статистичних моделей є найбільш вивченим і використовуваним на практиці. Однак моделі штучних нейронних мереж стають все більш привабливими, ефективними, результативними та успішними в досягненні розпізнавання образів (РО) в багатьох проблемах. На відміну від традиційних шаблонних підходів, ШНМ можуть легко моделювати складні та комплексні задачі. Колишні традиційні методи, що застосовуються для вирішення проблем, класифікуються на структурні, статистичні та гібридні підходи. Однак, як статистичний, так і структурний підходи можуть давати незадовільні результати, якщо їх застосовувати для вирішення складних задач розпізнавання образів. Наприклад, структурний

метод може виявитися слабким і не справлятися з обробкою шумових патернів. Аналогічно, він може бути слабким і неефективним у вирішенні числових семантичних інформаційних завдань. До того ж, статистичний метод не здатний використовувати інформацію про структуру патернів, що призвело до виникнення гібридного підходу. Однак, на сьогоднішній день моделі ШІ використовуються тому, що вони можуть дати кращий результат в РО-задачах, навіть в багатокомпонентних завданнях (рисунок 1.5).



Рисунок 1.5 – Розпізнавання образів у дорожньому трафіку

Розпізнавання образів – це обчислювальна парадигма, яка використовується для класифікації необроблених даних. РО охоплює безліч підходів, які забезпечують розробку різних додатків у різних сферах діяльності. Практичність цих підходів полягає в інтелектуальному наслідуванні людини. Образами в даній ситуації можна назвати сукупність предметів, об'єктів, образів, подій, випадків, ситуацій, ознак або абстракцій, де грані множини схожі в однозначному сенсі. За словами Норберта Вінера, Патерн – це розташування, що характеризується послідовністю ознак, з яких він складається, а не притаманною йому сутністю ознак. У той час як Ватанабе визначав патерн як «сутність». Він також може бути визначений унікальним або повторюваним знаменником серед численних патернів об'єкта. Наприклад, спільні риси в зображеннях відбитків пальців можуть

визначати патерн відбитків пальців. Таким чином, патерном може бути зображення відбитка пальця, обличчя людини, рукописне з'єднане слово, штрих-код, веб-сторінка в Інтернеті або мовний сигнал, тоді як розпізнавання – це завдання ідентифікації об'єкта, ознаки або події.

Перевага штучних нейронних мереж перед статистичним моделюванням полягає в тому, що вони не вимагають жорсткої структурованості та здатні працювати з неповними даними. Незважаючи на це, багато питань у розпізнаванні образів за допомогою нейронних мереж потребують вирішення (рисунок 1.6). Останнім часом вони привертають увагу дослідників саме у сфері розпізнавання зображень. Під час розпізнавання, на відміну від статистичної моделі, ці моделі значною мірою покладаються на кількість навчальних даних, представлення зразків і процедури навчання. Для досягнення хорошого розпізнавання образів необхідно вибрати хорошу ознаку за допомогою достатньої кількості мультикласифікаторів. Обчислювальний інтелект у цьому випадку допомагає у вирішенні різних проблем і мінімізує небезпеку.

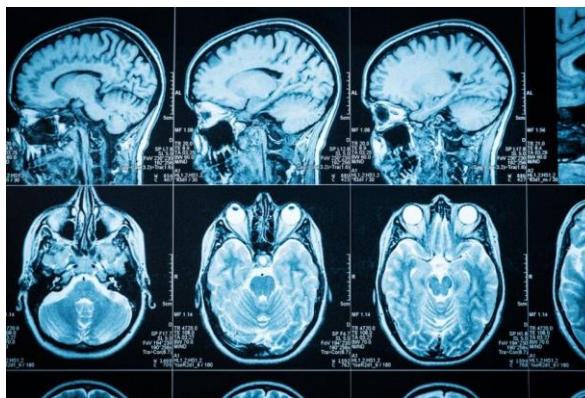


Рисунок 1.6 – Розпізнавання образів при дослідженні мозку

1.4 Алгоритми розпізнавання образів

Розпізнавання образів – це здатність алгоритму ідентифікувати та класифікувати дані на основі певних ознак або правил, тож важливим

пунктом роботи з цією технологією є вибір найефективнішого алгоритма, що буде застосований при навчанні нейронної мережі. Алгоритми розпізнавання образів є фундаментальним компонентом систем штучного інтелекту (ШІ) та машинного навчання (МН). Ці алгоритми дозволяють комп'ютерам виявляти і класифікувати закономірності у величезних обсягах даних, полегшуючи отримання значущих висновків і прогнозів. Використовуючи передові математичні та статистичні методи, алгоритми розпізнавання образів здійснили революцію в багатьох сферах, включаючи розпізнавання зображень і звуків, обробку природної мови та інтелектуальний аналіз даних. Закономірності у заданому обсязі інформації можуть бути представлені набором ознак або атрибутів, які відрізняють один клас від іншого.

Алгоритми розпізнавання спрямовані на фіксацію та аналіз цих відмінних характеристик, що дозволяє комп'ютерам автоматично розпізнавати і класифікувати нові, небачені екземпляри на основі вивчених шаблонів. Одним з найпоширеніших алгоритмів розпізнавання образів є алгоритм k-найближчих сусідів (KNN). KNN працює шляхом обчислення відстані між некласифікованою точкою даних і сусідніми позначеними точками даних. Призначаючи мітку класу на основі більшості голосів k найближчих сусідів, KNN дозволяє точно класифікувати нові спостереження. Цей алгоритм особливо ефективний для розпізнавання зображень, рекомендаційних систем та виявлення аномалій. Іншим відомим алгоритмом розпізнавання образів є метод опорних векторів (SVM). SVM – це потужний алгоритм, який будує гіперплощину або набір гіперплощин для розділення різних класів у наборі даних. Максимізуючи відстань між класами, SVM досягає надійної класифікації та регресії. Цей алгоритм знаходить застосування в різних сферах, таких як категоризація тексту, аналіз експресії генів і розпізнавання рукописного тексту. Крім того, нейронні мережі, зокрема моделі глибокого навчання, в останні роки

привертають значну увагу завдяки своїм винятковим можливостям розпізнавання образів.

Моделі глибокого навчання, такі як згорткові нейронні мережі (CNN) та рекурентні нейронні мережі (RNN), чудово справляються із завданнями, що включають складні структури даних, такі як зображення, відео та послідовності. Ці моделі автоматично вивчають ієрархічні представлення даних, що дозволяє їм фіксувати складні закономірності та досягати найсучасніших результатів у класифікації зображень, розпізнаванні мови та розумінні природної мови.

Алгоритми розпізнавання образів призначені для роботи з різноманітними наборами даних і вирішення різних завдань. Вони можуть працювати як у сценаріях керованого навчання, коли доступні марковані навчальні дані, так і в сценаріях некерованого навчання, коли алгоритм повинен виявляти закономірності без попередніх знань. Крім того, алгоритми розпізнавання образів можуть адаптуватися до динамічних середовищ за допомогою онлайн-навчання, постійно оновлюючи свої моделі в міру надходження нових даних.

Отже, методи розпізнавання образів мають різні переваги. Вони не лише допомагають аналізувати тенденції та робити прогнози, але й ідентифікувати об'єкти на різних вимірjuвальних відстанях і під різними кутами. Їх можна застосовувати в різних сферах, зокрема на фондовому ринку, в медицині, фінансовій індустрії та інших, щоб робити ефективні прогнози. Оскільки вони є у більшій мірі автоматизованими, вони пропонують рішення для проблем у реальному часі, яким людство ще не знайшло аналогів.

1.5 Висновки та постановка задачі

Сучасний світ, на який дуже сильно впливає сьогodнішній розвиток штучного інтелекту гостро реагує на будь-які зміни в цій сфері. В цьому

контексті, розпізнавання даних – важливий аспект наук пов’язаних із штучним інтелектом. Цей процес аналізу даних, що використовує алгоритми машинного навчання для класифікації вхідних даних на об’єкти, класи або категорії на основі розпізнаних образів, особливостей або закономірностей у даних, має багато застосувань у галузях астрономії, медицини, робототехніки та супутникового дистанційного зондування, серед інших. Тож варто пам’ятати, що вибір відповідних алгоритмів розпізнавання образів є критичним для досягнення точності та ефективності в процесі аналізу.

Різні алгоритми мають свої переваги та недоліки в залежності від конкретних умов і завдань. Наприклад, у контексті медичного застосування розпізнавання образів вибір правильних алгоритмів має вирішальне значення для точної діагностики, лікування та моніторингу пацієнтів. Оскільки медичні зображення можуть бути складними та містити велику кількість даних, важливо обрати такі алгоритми, які забезпечать надійне розпізнавання патологій та аномалій.

Розпізнавання образів також використовується для візуалізації медичних даних, що дозволяє лікарям бачити та розуміти складні дані швидше та ефективніше. Це може включати відображення структур органів, рухів та функцій тіла для кращого розуміння медичних образів та процесів. За допомогою візуалізації можна наочно побачити усю багатогранність різних алгоритмів навчання та обрати той, що підходить для певного контексту найбільше.

Відтак, мету цієї роботи можна визначити таким чином. Потрібно дослідити роботу нейронних мереж по розпізнаванню образів, а також дослідити алгоритми розпізнавання образів, що використовують нейронні мережі. Використовуючи вже існуючі підходи та методології, дослідити практичні недоліки та переваги використання нейронних мереж по розпізнаванню образів.

2 ТЕОРЕТИЧНЕ ДОСЛІДЖЕННЯ ОБРАНОЇ ПРОБЛЕМИ

2.1 Теоретичний аналіз

Нейронні мережі проявляють унікальні властивості, що дозволяють їм ефективно вирішувати різноманітні складні завдання. Вони здатні аналізувати великі обсяги даних та інформації, виявляючи у них закономірності та шаблони, навіть якщо дані неорганізовані та хаотичні. Крім того, вони здатні класифікувати об'єкти та групувати їх у класи та кластери з вражаючою точністю та детальністю, що деяких випадках перевищує можливості людського сприйняття.

Штучна нейронна мережа відтворює структуру людського мозку на комп'ютері. Вона складається зі шарів нейронів та синапсів, аналогічно організації мозку (рисунок 2.1).

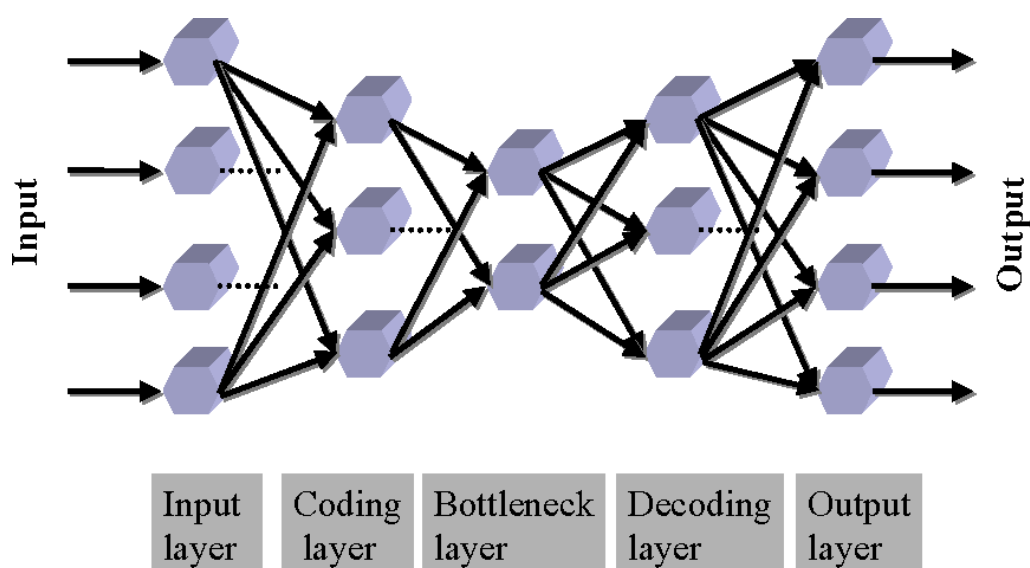


Рисунок 2.1 – Приклад роботи нейронної мережі

Ця мережа може містити мільйони нейронів, які з'єднані в єдину систему, що дозволяє їй ефективно аналізувати та навіть запам'ятовувати різноманітну інформацію.

Основні функції штучних нейронних мереж можна узагальнити таким чином:

- класифікація (визначення одного чи декількох класів, до яких належить вхідний об'єкт);

- прогнозування (розрахунок усіх можливих варіантів розвитку подій за певними сценаріями по його характеристикам та статистикою відношення сценаріїв попередніх об'єктів);

- кластеризація (утворення відповідних класів, які базуються на певних особливостях деякої кількості схожих об'єктів);

- розпізнавання – найпоширеніша сфера застосування нейронних мереж. На основі певних властивостей об'єкта намагається віднайти схожі на нього, або об'єкти з такими ж властивостями.

Хоча перерахований набір функцій може бути невеликим, область їх застосування настільки широка, що життя без них важко уявити. Раніше, до 2010 року, завдання розпізнавання обличчя на зображенні було надзвичайно складним для комп'ютерів. Переглянути картинку та впізнати обличчя статуї можна було легко для людини, проте для комп'ютера це вимагало складного аналізу та опису геометричних взаємозв'язків об'єкта. Налаштування кожної нейронної мережі належним чином вимагає великих обсягів інформації з відповідної області та проведення декількох експериментів з різними параметрами налаштування. Зазвичай, чим триваліше та більш насичене даними навчання, тим більш ефективною стає робота нейронної мережі та зменшуються ризики неправильної обробки результатів.

Існує різноманітна кількість типів нейронних мереж, проте вони завжди складаються з однакових компонентів: нейронів, синапсів, ваг, зміщень та функцій. Нейрон, або вузол, є основною будівельною одиницею нейронних мереж, яка отримує інформацію, виконує прості обчислення та передає її далі (рисунок 2.2). Усі нейрони в мережі можна розділити на три категорії: вхідні нейрони, які отримують інформацію з зовнішнього

середовища; приховані нейрони, які обробляють цю інформацію; вихідні нейрони, які формують висновок.

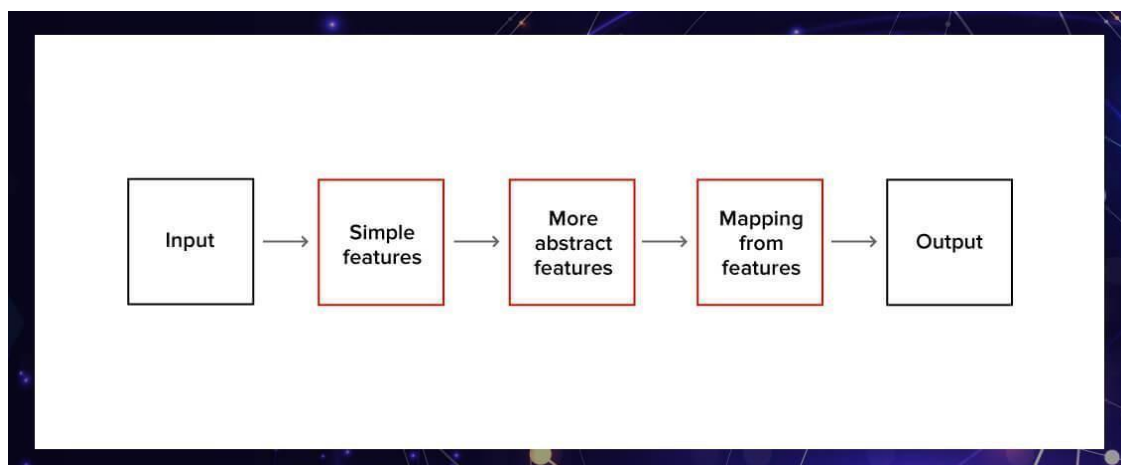


Рисунок 2.2 – Приклад роботи нейрона

У великій нейронній мережі з великою кількістю нейронів та зв'язків між ними нейрони організовані в шари. Це включає вхідний шар, який отримує інформацію, ряд прихованих шарів та вихідний шар, який надає остаточні результати. Кожен нейрон виконує перетворення вхідних даних, працюючи з числами у діапазоні $[0,1]$ або $[-1,1]$. Для забезпечення роботи нейронів з цими даними вимагається нормалізація. Синапси в цій мережі є зв'язками між нейронами, подібними до електричних кабелів. Кожен синапс має свою вагу, яка вносить внесок у вхідну інформацію.

Результати роботи нейрона з високою вагою матимуть більший вплив на наступний нейрон, в той час як інформація від менш "важливих" нейронів може бути менш вагомою. Матриця ваг визначає вплив на всю систему нейронів. Щоб визначити, який нейрон має найбільшу вагу, потрібно провести процес оптимізації після ініціалізації ваг, яка зазвичай виконується випадковим чином під час першого запуску нейронної мережі. Нейрон зі зміщенням розширює можливості зберігання різних комбінацій ваг. Зміщення додає більше різноманіття вхідного простору до моделі ваг (рисунок 2.3).

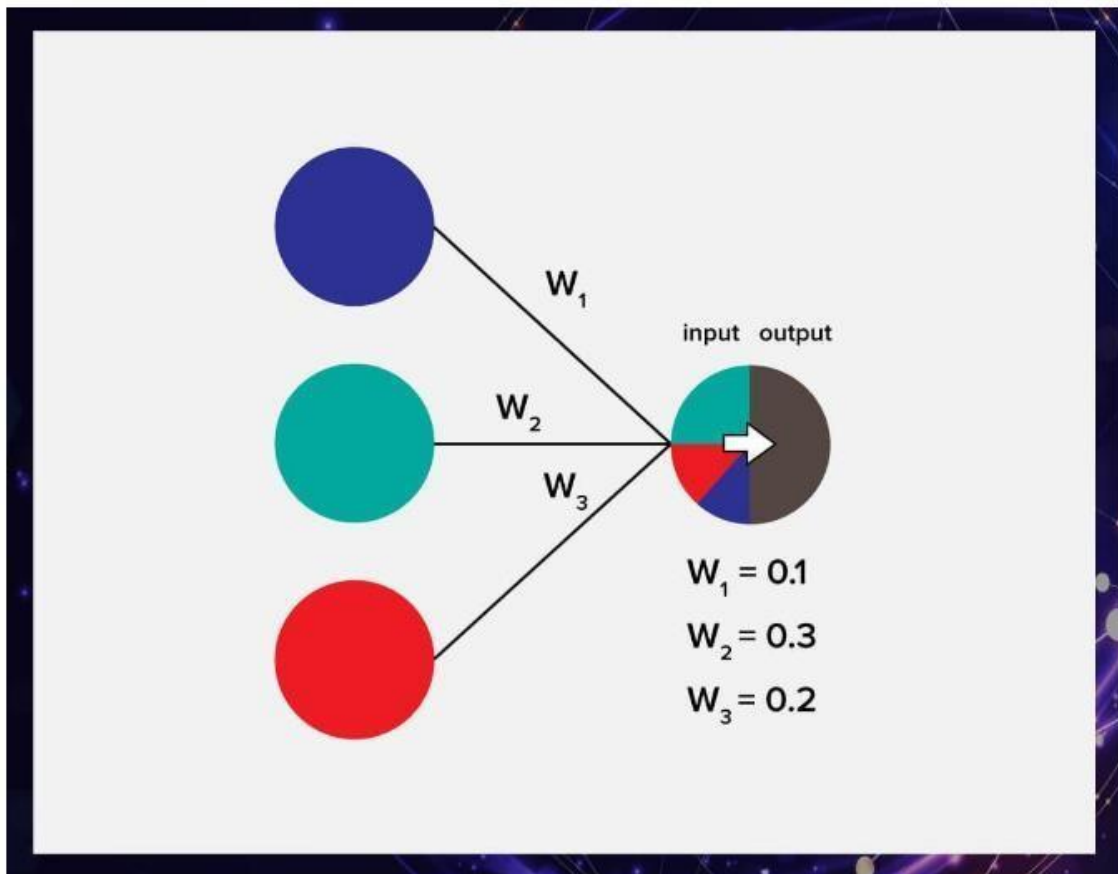


Рисунок 2.3 – Приклад побудови нейрону

У нейронних мережах нейрон зі зміщенням додається до кожного шару і відіграє важливу роль у зміщенні функції активації вліво або вправо на графіку. Підготовка ШНМ здійснюється за допомогою процесу навчання. У процесі навчання ваги змінюються для досягнення необхідних результатів. У процесі навчання мережа обробляє деякі вибіркові дані, а ваги змінюються для досягнення кращого наближення до бажаного результату. У керованому навчанні моделі подається навчальна вибірка. Навчальний набір складається з вхідних прикладів і відповідних цільових виходів. Вхідні дані записуються для реакції мережі, а ваги між ними налаштовуються для зменшення помилок, для досягнення бажаного результату.

Під час цього процесу мережа виконує послідовні ітерації до тих пір, поки розрахунковий результат не збігається з правильним. Особливої уваги

потребує побудова навчальної вибірки. Навчальна множина вважається ідеальною, і вона повинна якнайкраще відображати основну модель. В іншому випадку, надійна модель з бажаними результатами не може бути досягнута з нерепрезентативною навчальною вибіркою. У процесі керованого навчання мережі спочатку навчаються перед тим, як їх використовують у моделі для прогнозування результатів. Важливо, що коли мережа починає обчислювати передбачувані виходи з серією входів, з фіксованими вагами, тоді модель ШНМ може бути налаштована на необхідну операцію.

Деякі з відомих алгоритмів з керованим методом навчання – це Adaline (використовується для бінарних даних), Perceptron (використовується для безперервних даних) і Madaline (розроблений на основі Adaline). Навчання з підкріпленням – це окремий випадок навчання під контролем. Це коли зовнішнє середовище лише перевіряє інформацію на предмет прийняття та відхилення, замість того, щоб вказувати правильний вихід. У цьому процесі найбільш ефективні та активні зв'язки між нейронами, що відповідають за вхідні дані, зміцнюються протягом послідовних ітерацій. Деякі з відомих алгоритмів навчання з підкріпленням – це машина Больцмана, квантування вектора навчання та мережі Хопфілда. Моделі ШНМ мають багато застосувань для класифікації зображень, управління рослинами, прогнозування, передбачення, робототехніки, класифікації сигналів ЕКГ і багато іншого.

Навчання без вчителя не використовує навчальну вибірку або цільовий підхід до результату. Натомість воно відстежує структуру вхідних даних базової моделі. У цьому процесі ШНМ-модель коригує свої ваги відповідно до вхідних даних, таким чином отримуючи результати, подібні до вхідних. Модель без будь-якої зовнішньої підтримки розпізнає закономірності та відмінності у вхідних даних. У цьому процесі формуються кластери, кожен з яких складається з групи з декількох вагових коефіцієнтів, таким чином, що пов'язаний вхідний шлях призводить до

схожого виходу. Якщо в процесі ітерації виявляється новий шаблон, він класифікується як новий кластер. Автокодери, геббівське навчання, мережі глибоких переконань, самоорганізаційні карти, генеративні змагальні мережі та метод алгебраїчної реконструкції (ART) – це декілька найвідоміших алгоритмів для неконтрольованого навчання. Моделі ШНМ використовуються для діагностики захворювань, сегментації зображень та багато іншого.

Некеровані алгоритми стали дуже корисними і потужними інструментами в сегментації магнітно-резонансних зображень для виявлення аномалій в системах організму.

2.2 Типи нейронних мереж

Нейронні мережі, натхненні нейронами нашого мозку, стали дійсно потужними інструментами для вирішення складних проблем. Як вже було згадано, їхня здатність навчатися на даних та робити прогнози робить їх незамінними в багатьох галузях.

Існує декілька типів нейронних мереж, які відрізняються за способом обробки інформації та сферами застосування.

Мережі прямого поширення є найпростішим типом нейронних мереж. Інформація в цих мережах тече лише в одному напрямку, від вхідних даних до вихідних. Це робить їх добре підходящими для задач класифікації та прогнозування, де мережа має навчитися відображати один набір даних на інший.

Рекурентні нейронні мережі більш складні, ніж мережі прямого поширення. Вони дозволяють інформації текти назад через мережу, що робить їх придатними для задач обробки послідовностей, таких як розпізнавання мови та машинний переклад. Ці мережі можуть враховувати контекст попередніх даних при обробці нових, що дозволяє їм краще розуміти складні послідовності.

Згорткові нейронні мережі спеціально розроблені для обробки зображень та інших даних з просторовою структурою. Ці мережі використовують спеціальні типи нейронів, які можуть виявляти патерни в даних, розташованих у просторі. Це робить їх дуже ефективними для задач обробки зображень, таких як розпізнавання об'єктів та сегментація зображень. Важливо пам'ятати, що нейронні мережі не є бездоганними. Вони можуть бути схильні до упередженості, помилок, і їх може бути важко інтерпретувати (рисунок 2.6).

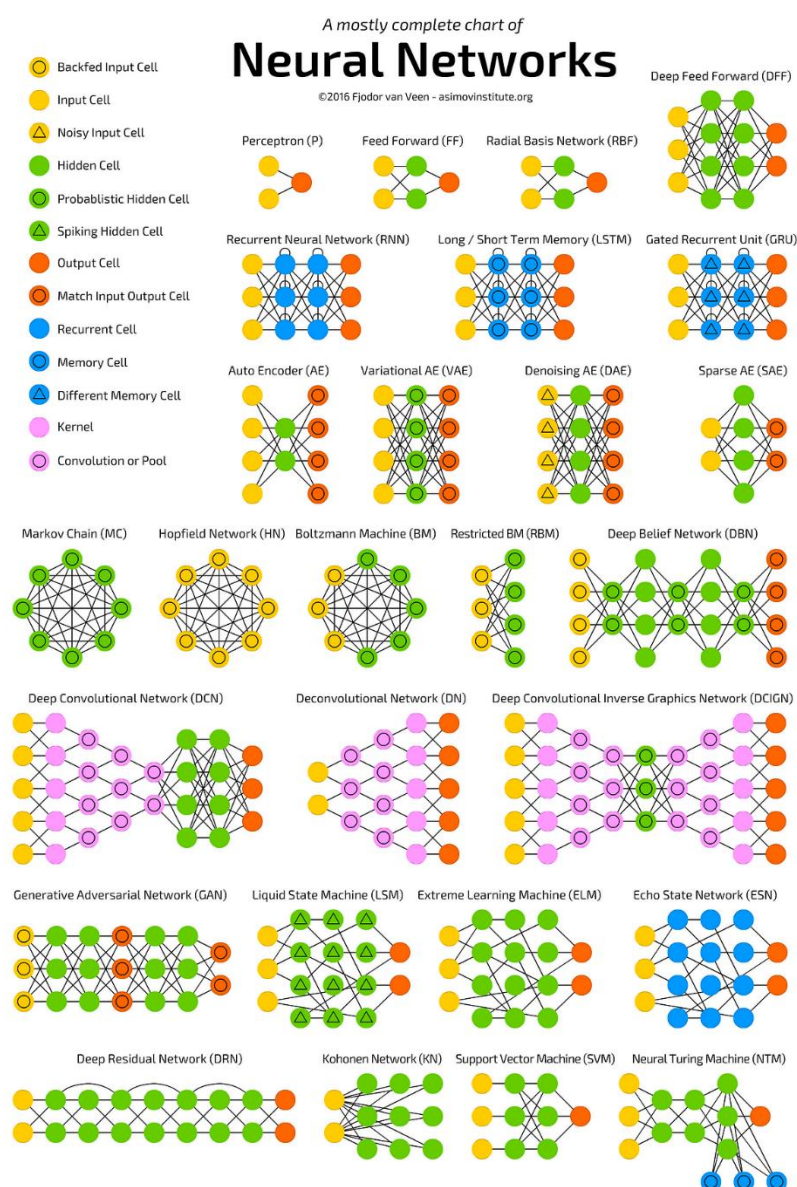


Рисунок 2.6 – Приклади нейронних мереж

Мережі з радіально-базисною функцією (RBF) вирізняються серед інших нейронних мереж тим, що їхній вхідний шар не здійснює жодних обчислень, а просто передає дані безпосередньо до прихованого шару. Така архітектура забезпечує швидшу швидкість навчання RBF-мереж. Формально, активаційна функція прихованого шару визначається як:

$$\varphi(x) = \exp(-\gamma \|x - c\|^2), \quad (2.1)$$

де φ – радіально-базисна функція;

x – вхідний вектор;

c – центр базисної функції;

γ – параметр ширини функції.

RBF-мережі застосовуються у задачах прогнозування часових рядів та апроксимації функцій.

Мережі з довгою короткочасною пам'яттю (LSTM) мають унікальну архітектуру, яка дозволяє їм обробляти і зберігати інформацію у вигляді комірок короткочасної і довготривалої пам'яті. Основна ідея LSTM полягає у використанні блоків пам'яті, які керуються трьома типами гейтових механізмів: вхідним, вихідним та забуття.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (2.2)$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad (2.3)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \quad (2.4)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t, \quad (2.5)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (2.6)$$

$$h_t = o_t \cdot \tanh(C_t), \quad (2.7)$$

де σ – сигмоїдна функція активації;

W – ваги;

b – зсуви;

h – вихід;

C – стан комірки.

LSTM-мережі використовуються для розпізнавання рукописного тексту та перетворення відео в текст.

Багатошарові перцептрони (MLP) – це клас нейронних мереж, здатних вивчати залежності між лінійними та нелінійними даними. Використовуючи алгоритм зворотного поширення помилки, MLP можуть мінімізувати функцію втрат:

$$E = \frac{1}{2} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (2.8)$$

де y – реальне значення;

\hat{y} – прогнозоване значення;

n – кількість зразків.

MLP застосовуються у задачах розпізнавання облич та комп'ютерного зору.

Генеративні змагальні мережі (GAN) складаються з двох частин: генератора та дискримінатора. Генератор створює нові зразки даних, які дискримінатор намагається відрізнити від реальних даних. Задача генератора – мінімізувати, а дискримінатора – максимізувати функцію втрат:

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))], \quad (2.9)$$

де D – дискримінація;

G – генератор;

x – реальні дані;

z – випадковий шум.

GAN використовуються для створення мистецтва за допомогою штучного інтелекту.

Карти, що самоорганізуються (SOM), перетворюють великі складні набори даних на зрозумілі двовимірні карти за допомогою алгоритмів конкурентного навчання. У кожному циклі навчання визначається «найближчий» нейрон, який коригується разом з його сусідами:

$$\Delta w_i = \eta \cdot h_{ci}(t) \cdot (x - w_i), \quad (2.10)$$

де η – коефіцієнт навчання;

$h_{ci}(t)$ – функція сусідства;

x – вхідний вектор;

w_i – ваги нейрона.

SOM використовуються для аналізу тенденцій голосування та організації астрономічних даних. Таким чином, різні типи нейронних мереж мають унікальні архітектури та алгоритми, що робить їх ефективними для широкого спектра застосувань, від обробки зображень до генерації нових даних. Розуміння їх специфіки та областей застосування є важливим для ефективного використання нейронних мереж у наукових дослідженнях та практичних задачах.

2.3 Детальний огляд розпізнавання образів

Розпізнавання образів – це фундаментальна сфера штучного інтелекту, що прагне наділити комп'ютери здатністю розуміти та інтерпретувати візуальну інформацію. Завдяки своїй вражаючій

потужності, ця технологія знаходить застосування в широкому спектрі галузей, від розпізнавання облич та об'єктів до автономного керування та медичної діагностики.

Нейронні мережі, натхненні нейронною структурою людського мозку, стали одними з найефективніших інструментів для вирішення задач розпізнавання образів. Процес розпізнавання розпочинається із захоплення зображення або відео за допомогою датчика, наприклад, камери. Цей датчик перетворює оптичну інформацію на цифровий сигнал, який може бути оброблений комп'ютером. Перетворення оптики в цифру – цей процес включає дискретизацію та квантування зображення. Дискретизація розбиває зображення на окремі пікселі, а квантування присвоює кожному пікселю значення кольору. Перед подачею зображення на нейронну мережу його необхідно підготувати. Цей етап може включати різні методи обробки для покращення його якості та чіткості.

- видалення шуму: застосовуються різні фільтри, такі як фільтр Гаусса, медіанний фільтр та фільтр Калмана, для видалення випадкового шуму з зображення.

- регулювання контрастності: цей процес використовується для покращення чіткості зображення шляхом збільшення або зменшення його контрастності.

- нормалізація розміру: зображення може бути змінено до стандартного розміру для зручного оброблення.

Зображення, яке спочатку представлене як матриця пікселів, необхідно перетворити на формат, який може бути оброблений нейронною мережею. Це може включати декілька етапів. Векторизація – перетворення матриці пікселів у вектор, де кожен елемент вектора відповідає значенню яскравості пікселя. Локальні ознаки – вилучення локальних характеристик, таких як краї, текстури або форми, з зображення. Глобальні ознаки – вилучення глобальних характеристик, таких як гістограми кольорів або статистичні характеристики текстури, з зображення. Архітектура нейронної

мережі зазвичай складається з трьох типів шарів. Вхідний шар приймає представлене зображення як вхідні дані, приховані шари виконують складні обчислення та вивчають нелінійні зв'язки між вхідними даними та вихідними класами. Вихідний шар генерує ймовірність для кожного класу, до якого може належати зображення (рисунок 2.7).

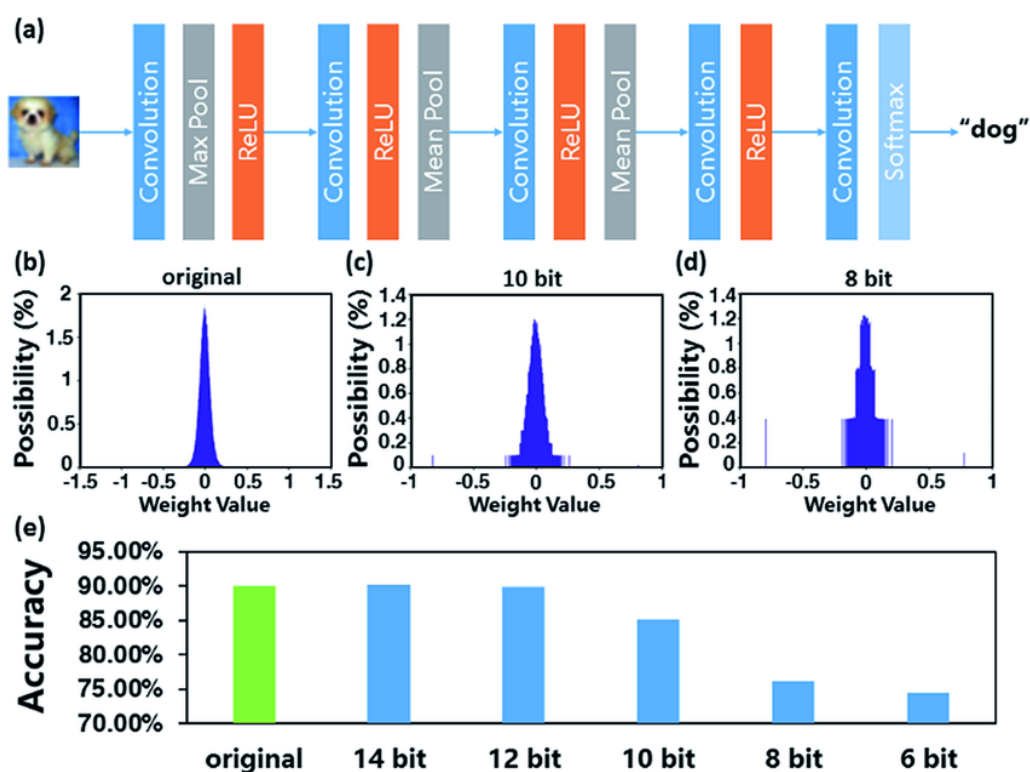


Рисунок 2.7 – Процес обробки зображення

Нейронна мережа навчається на наборі даних, який містить зображення, позначені відповідними класами. Процес навчання включає:

- пряме розповсюдження: пропускання зображення через мережу шар за шаром. На кожному шарі відбувається лінійне комбінування вхідних даних з вагами нейронів, та застосування нелінійної функції активації для отримання вихідного сигналу;

- зворотне розповсюдження: розрахунок градієнтів помилки щодо ваг нейронів. Помилка визначається різницею між передбаченими та

фактичними класами зображення. Градієнти вказують, як зміна ваги кожного нейрона впливає на загальну помилку мережі;

- оновлення ваг: регулювання ваг нейронів у напрямку зменшення помилки мережі. Для цього використовуються оптимізатори, такі як градієнтний спуск або Adam, які враховують як поточні градієнти, так і минулі оновлення ваг;

- функція активації: ця математична функція визначає, як нейрон перетворює зважений сумарний вхідний сигнал на вихідний сигнал. Загальними функціями активації є сигмоїдна функція та функція ReLU (випрямлений лінійний блок). Вибір функції активації впливає на здатність мережі моделювати складні нелінійні зв'язки в даних;

- функція витрат: Ця функція вимірює розбіжність між передбаченими та фактичними вихідними даними. Загальною втратною функцією для розпізнавання образів є крос-ентропія. Різні втратні функції можуть бути більш чутливими до певних типів помилок класифікації.

Після навчання нейронну мережу можна використовувати для розпізнавання нових зображень. Нове зображення проходить через той самий процес представлення, що і навчальні зображення. Зображення подається на вхід мережі, і воно проходить через усі шари, виконуючи обчислення на основі вивчених під час навчання зв'язків (рисунок 2.8).



Рисунок 2.8 – Обчислення на основі вивчених зв'язків

Вихідний шар генерує ймовірність для кожного класу. Ці ймовірності вказують на те, наскільки ймовірно, що зображення належить до кожного з класів. На основі вихідних ймовірностей мережі приймається рішення про клас зображення. Зображення класифікується до класу, для якого мережа генерувала найвищу ймовірність. Встановлюється поріг ймовірності. Якщо жодна з ймовірностей не перевищує цей поріг, зображення класифікується як невідоме.

Навчання нейронних мереж може призвести до перенавчання, коли мережа занадто добре підлаштовується під навчальний набір даних і погано працює на нових даних. Регуляризація – це набір методів, які допомагають запобігти перенавчанню (рисунок 2.9).

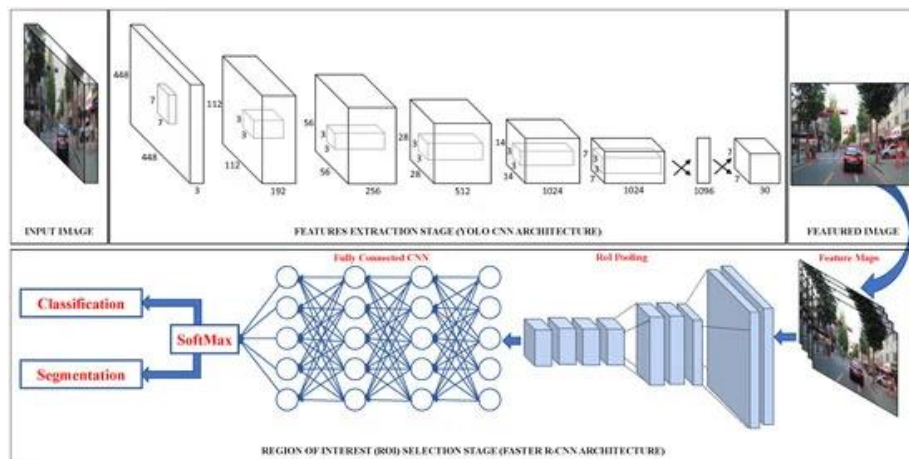


Рисунок 2.9 – Процес регуляризації

Додавання штрафного члена до функції витрат, який залежить від величини ваг нейронів під час цього процесу, називається відсікання. Це сприяє тому, що мережа віддає перевагу більш простим моделям з меншими значеннями ваг. Випадання в той же час виключає випадкові нейрони під час навчання. Це змушує мережу покладатися на більш різноманітні ознаки зображень і запобігає надмірній підгонці під конкретні навчальні зразки. Після навчання нейронної мережі важливо оцінити її ефективність на даних, яких вона раніше не бачила. Потрібно перевірити точність – частку

правильно класифікованих зображень, а також втрата на перевірконому наборі – низька втрата вказує на те, що модель може узагальнювати свої знання на нові дані. Важливим аспектом є AUC-ROC (площа під кривою приймача-оператора). Ця метрика оцінює здатність моделі розрізняти між позитивними та негативними класами. Вона корисна для порівняння продуктивності різних моделей незалежно від обраного порога класифікації. Розпізнавання образів за допомогою нейронних мереж – це потужна технологія, яка революціонізувала багато сфер. Розуміння технічних деталей процесу навчання та оцінки моделей є критичним для ефективного використання цієї технології. Очікується, що подальші дослідження дозволять подолати існуючі проблеми та відкрити нові можливості застосування розпізнавання образів.

2.4 Сучасні рішення та алгоритми розпізнавання образів

Завдяки значному прогресу в галузі штучного інтелекту, розпізнавання образів стало однією з найдинамічніших та найактуальніших сфер досліджень. Сучасні алгоритми та рішення, що ґрунтуються на нейронних мережах, здатні вирішувати складні задачі розпізнавання з вражаючою точністю та ефективністю.

До них відносяться трансформатори, ці нові типи нейронних мереж демонструють кращу продуктивність у деяких задачах розпізнавання образів, завдяки кращій здатності моделювати довгострокові залежності в зображеннях. Інший напрямок це навчання з великими обсягами даних – завдяки доступності більших обсягів даних та потужнішим обчислювальним ресурсам, стає можливим навчання складніших нейронних мереж, які досягають вищої точності. Багатозадачне та самоконтрольоване навчання дозволяють покращити продуктивність моделей та зробити їх більш стійкими до шуму та варіацій даних. Тлумачення нейронних мереж спрямовані на розробку методів, які дозволяють краще зрозуміти, як

нейронні мережі приймають свої рішення, роблячи їх більш надійними та довіреними.

Окрім цих фундаментальних досліджень, розпізнавання образів знаходить все ширше застосування в реальних задачах. Розпізнавання об'єктів у реальному часі, що використовується в самокерованих автомобілях, розширеній реальності та системах безпеки. Розпізнавання емоцій, що застосовується для покращення взаємодії людини з комп'ютером, маркетингу та досліджень ринку. Розпізнавання зображень з високим розширенням задля використання в супутникових знімках, медичній діагностиці та контролі якості. Розпізнавання в складних умовах важливе для задач в динамічних або неідеальних середовищах. Глибоке генерування зображень застосовується для створення нового контенту, візуальних ефектів та дизайну. Розпізнавання 3D-об'єктів використовується в робототехніці, віртуальній реальності та інженерному проектуванні. Розпізнавання тексту застосовується для автоматизації введення даних, індексування зображень та перекладу мов. Розпізнавання жестів використовується для управління комп'ютером без рук, інтерактивних ігор та мови жестів. Розпізнавання звуку застосовується для аналізу мови, музики та шумів навколишнього середовища. Розпізнавання запахів, інноваційна область з потенційними застосування в медицині, контролі якості та екологічному моніторингу.

В науковому дослідженні, присвяченому детальному огляду розпізнавання образів, важливо ретельно проаналізувати ключові аспекти цього процесу з метою кращого розуміння його механізмів та можливих застосувань. У цьому розділі необхідно розглянути різноманітні методи та техніки, що використовуються в сучасних системах розпізнавання образів, провести їхній докладний аналіз і висвітлити основні виклики, з якими стикаються дослідники в цій області. З метою виявлення перспективних напрямків для подальших досліджень було обрано практичну галузь, яку

можливо покращити за допомогою даних технологій, а саме попередження про розвиток ракових пухлин в медицині.

Ріст пухлин складається з багатьох етапів і складних процесів, а також багатьох генів і молекулярних подій, включаючи мультигенні мутації, такі як активація онкогенів та інактивація генів-супресорів пухлин. У сучасній задокументованій літературі більшість з них присвячені вивченню впливу окремого фактора на розвиток раку в гіпотетичних умовах. Однак деякі дослідження виявили, що до виникнення раку може призводити декілька молекулярних подій. Типова модель виникнення раку потребує мутації від двох до восьми генів-«водіїв» [1]. Мутація генів-пасажирів не здатна призвести до розвитку раку [2]. Мета досліджень повинна бути зосереджена на групі генних мутацій, яка називається мутаціями генного патерну. Залежно від центральної догми, мутація генного патерну може впливати на низку експресій мРНК і білків (рисунок 2.10). Для того, щоб створити моделі діагностики, засновані на диференційованій експресії білків або пептидів між пухлинними тканинами і нормальними тканинами, цей патерн дозволив би уникнути результату низької чутливості одного пухлинного маркера або низької специфічності великої кількості зразків.

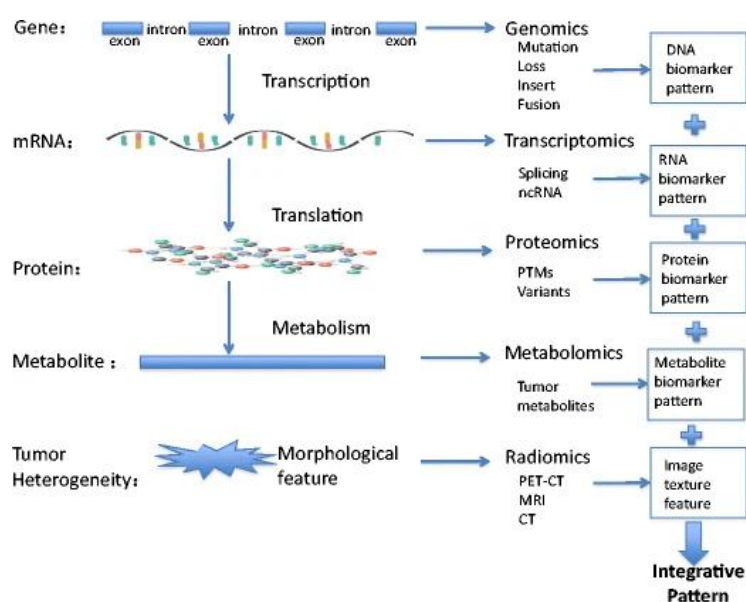


Рисунок 2.10 – Вплив різних хімічних сполук

Крім того, з розвитком біомаркерів раку було виявлено, що зміна панелі ключових молекул у генних і білкових послідовностях ініціює пухлиноутворення. Оскільки різні люди мають різний набір ключових молекул, клінічні лікарі можуть використовувати різні цільові препарати для запобігання виникненню пухлин. Багатопараметрична систематична стратегія предиктивної, превентивної та персоналізованої медицини (PPPM) в онкології була спочатку задумана Жаном і Дезидеріо [3]. Більше того, біологія раку поступово перейшла до ери прецизійної онкологічної медицини [4]. У центрі уваги цієї оглядової статті – використання змін біологічних характеристик пухлини для діагностики, лікування та оцінки прогнозу пацієнта (рисунок 2.11).

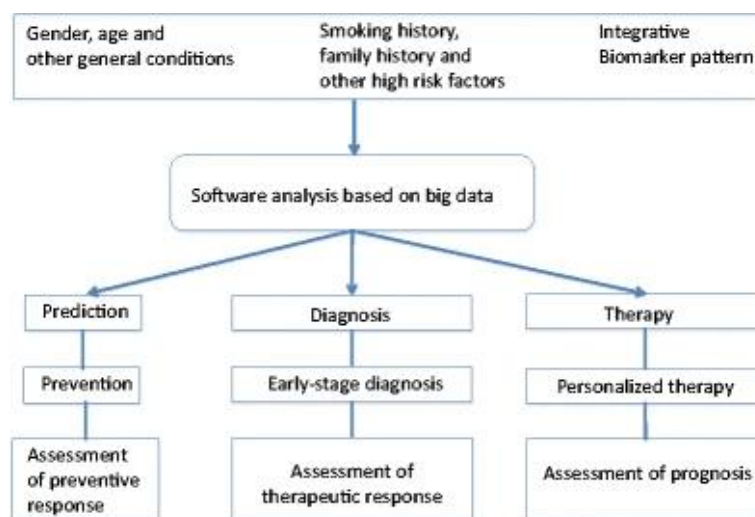


Рисунок 2.11 – Варіанти використання розпізнавання образів

Патерн, який надасть змогу чітко визначити наявність пухлин може складатися з різних типів біомаркерів з геному, транскриптома, протеома, метаболома і радіома. Патерн може складатися не тільки з однотипних молекулярних маркерів, але й з різних типів молекулярних маркерів, які можуть бути об'єднані разом для формування інтегративного патерну, наприклад, дані мас-спектрометрії та дані мікрочипів для експресії генів складаються в інтегративний патерн. Результати аналізу показують, що

паттерн, який поєднує дані MSI і біологічні дані, здатний забезпечити значущу дискримінацію між зразками. Це може бути корисним інструментом для виявлення потенціалу у великомасштабних біологічних дослідженнях, особливо для ідентифікації онкохворих і здорових людей [72].

Однак, все ще існують деякі проблеми щодо розпізнавання образів. По-перше, на різних стадіях розвитку пухлини можуть бути різні варіації генів або білків. Як ідентифікувати ці гени та їхні білки, залишається складним завданням. По-друге, рецидив пухлини – це не лише проста зміна гена або білка, але й тісно пов'язаний з умовами життя пацієнта та його харчовими звичками. Зосередитися лише на одному аспекті недостатньо.

Завдяки динамічній еволюції розпізнавання образів знаходить все ширше застосування, включаючи розпізнавання об'єктів у реальному часі, розпізнавання емоцій, обробку зображень з високою роздільною здатністю, роботу в складних умовах, генерування зображень, розпізнавання 3D-об'єктів, обробку тексту, жестів, звуку та запахів. В цілому, розпізнавання образів – це потужна та динамічно розвивається наукова сфера з широким спектром потенційних застосувань. Очікується, що в найближчі роки ця галузь продовжуватиме розвиватися, відкриваючи нові можливості та революціонізуючи багато аспектів нашого життя.

3 ПРАКТИЧНА РЕАЛІЗАЦІЯ

3.1 Опис середовища розробки та інструментів

Розпізнавання образів є однією з найактуальніших галузей машинного навчання, яка використовується для широкого спектру задач, від розпізнавання облич до автономного керування. Python, завдяки своїй простоті, гнучкості та широкому спектру бібліотек, став домінуючою мовою програмування для розробки алгоритмів розпізнавання образів. Синтаксис Python простий та інтуїтивно зрозумілий, що робить його доступним для початківців та дослідників з різним досвідом програмування. Це полегшує вивчення та читання коду, роблячи його зручним для співпраці та спільного розвитку проектів. Він може використовуватися для широкого спектру задач, пов'язаних з машинним навчанням та глибокими нейронними мережами, що робить його універсальним інструментом.

Python добре підходить для швидкого прототипування моделей завдяки своїй простоті та динамічності, адже він легко інтегрується з іншими мовами програмування та інструментами, що робить його зручним для роботи з різними типами даних та системами. Ефективність даної мови програмування також впливає на вибір інструменту, багато інтерпретаторів Python використовують JIT-компіляцію, яка динамічно перетворює код Python на машинний код, що покращує його продуктивність. Для нього існує широкий спектр оптимізованих бібліотек, таких як NumPy, TensorFlow та PyTorch, які забезпечують високу продуктивність для задач машинного навчання. Python підтримує багатопотоковість, що дозволяє розподіляти обчислення на декілька ядер процесора, значно покращуючи продуктивність та може використовуватися з апаратними прискорювачами, такими як GPU та TPU, для досягнення ще більшої продуктивності.

Задля детального аналізу алгоритмів розпізнавання образів була обрана важлива та актуальна тема, а саме допомога систем штучного

інтелекту в виявленні різних ознак та етапів розвитку ракових пухлин в людському організмі (рисунок 3.1).

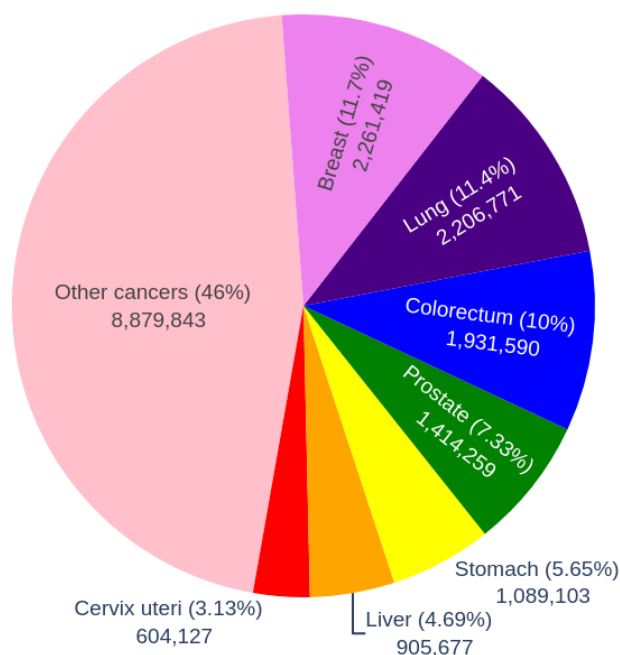


Рисунок 3.1 – Статистика захворювань на рак

Бібліотеки та фреймворки для реалізації Трансформаторів, DeiT, EfficientNets та Swin Transformers включають TensorFlow, PyTorch, Transformers та OmniNet. Кожна з цих технологій має свої унікальні особливості, які роблять їх корисними для різних аспектів машинного навчання та глибинного навчання. TensorFlow є потужною бібліотекою для побудови та навчання нейронних мереж, розробленою компанією Google. Вона забезпечує широкий спектр функцій для роботи з великими наборами даних, включаючи високопродуктивні API, такі як Keras, для швидкого створення моделей, а також інструменти для розподіленого обчислення.

TensorFlow має активну спільноту та багату документацію, що робить його зручним для початківців. Проте, через свою багатофункціональність та складність, TensorFlow може бути важчим для освоєння порівняно з деякими іншими бібліотеками. PyTorch, розроблений Facebook, відомий

своєю гнучкістю та інтуїтивно зрозумілим інтерфейсом, що робить його популярним серед дослідників та розробників.

PyTorch підтримує динамічні обчислювальні графи, що дозволяє швидше прототипувати та експериментувати з моделями (рисунок 3.2). Це забезпечує більше контролю над процесом навчання і полегшує налагодження. Однак, для новачків PyTorch може здатися складнішим, а його продуктивність на дуже великих наборах даних іноді може бути трохи нижчою порівняно з TensorFlow.

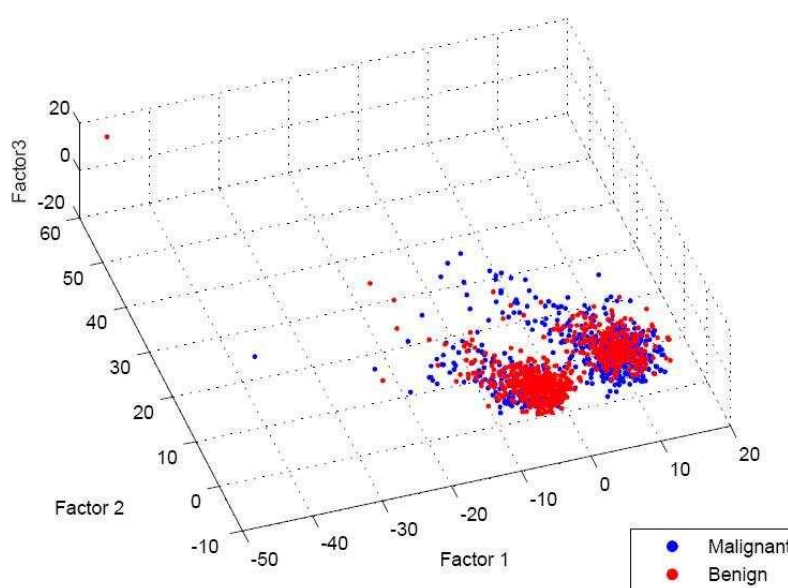


Рисунок 3.2 – Можливості візуалізації результатів роботи нейронних мереж

Transformers – це бібліотека з відкритим кодом, розроблена компанією Hugging Face, яка надає готові до використання реалізації багатьох сучасних моделей, таких як Трансформатори, DeiT, EfficientNets та Swin Transformers. Transformers пропонує зручний інтерфейс для завантаження та використання попередньо навчених моделей, що полегшує швидкий початок роботи. Ця бібліотека спрощує процес інтеграції та

експериментування з різними моделями, але може не забезпечувати такої ж гнучкості та можливості налаштування, як TensorFlow або PyTorch.

OmniNet є бібліотекою з відкритим кодом, яка пропонує універсальну архітектуру нейронних мереж для вирішення різних задач, включаючи виявлення первинних ознак пухлин. OmniNet базується на глибоких нейронних мережах і може бути дуже ефективною для роботи з великими наборами даних. Проте, вона є складнішою для налаштування та використання, що може вимагати глибших знань у галузі машинного навчання. Крім того, як новий проект, OmniNet поки що не має такої широкої підтримки та екосистеми, як TensorFlow, PyTorch або Transformers.

3.2 Опис набору даних

Було обрано два набори даних для ранньої діагностики та прогнозування раку The Cancer Image Archive (TCIA): TCIA містить мільйони зображень раку з різних типів раку та різних методів візуалізації, що робить його цінним ресурсом для дослідження ранньої діагностики та прогнозування раку. Зображення можуть бути використані для навчання моделей розпізнавання образів для виявлення ранніх ознак раку на зображеннях. Набір даних TCIA також містить клінічні дані для пацієнтів, з якими пов'язані зображення, що дозволяє дослідникам співвідносити результати розпізнавання образів з клінічними результатами (рисунок 3.3).

The Cancer Genome Atlas (TCGA): TCGA містить геномні дані для тисяч пацієнтів з різними типами раку, що робить його цінним ресурсом для дослідження прогнозування раку. Геномні дані можуть бути використані для виявлення біомаркерів, які можуть допомогти прогнозувати перебіг раку та реакцію на лікування. Набір даних TCGA також містить клінічні дані для пацієнтів, з якими пов'язані геномні дані, що дозволяє дослідникам співвідносити геномні профілі з клінічними результатами (рисунок 3.4).

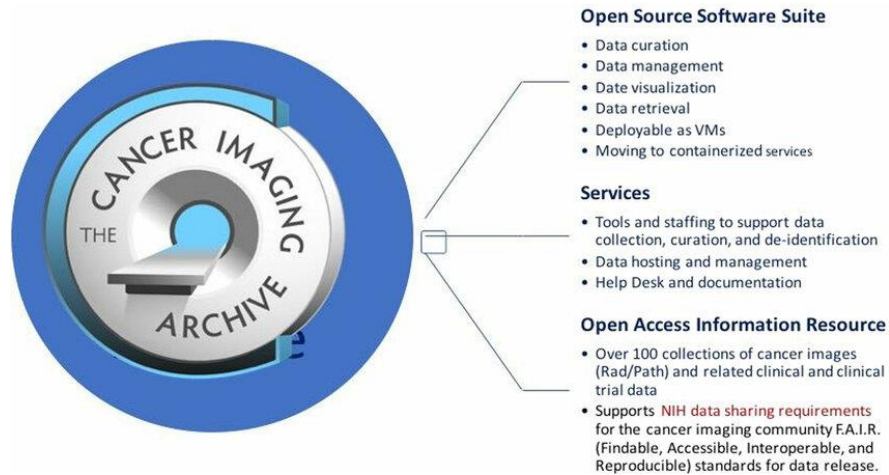


Рисунок 3.3 – Набір даних TCIA

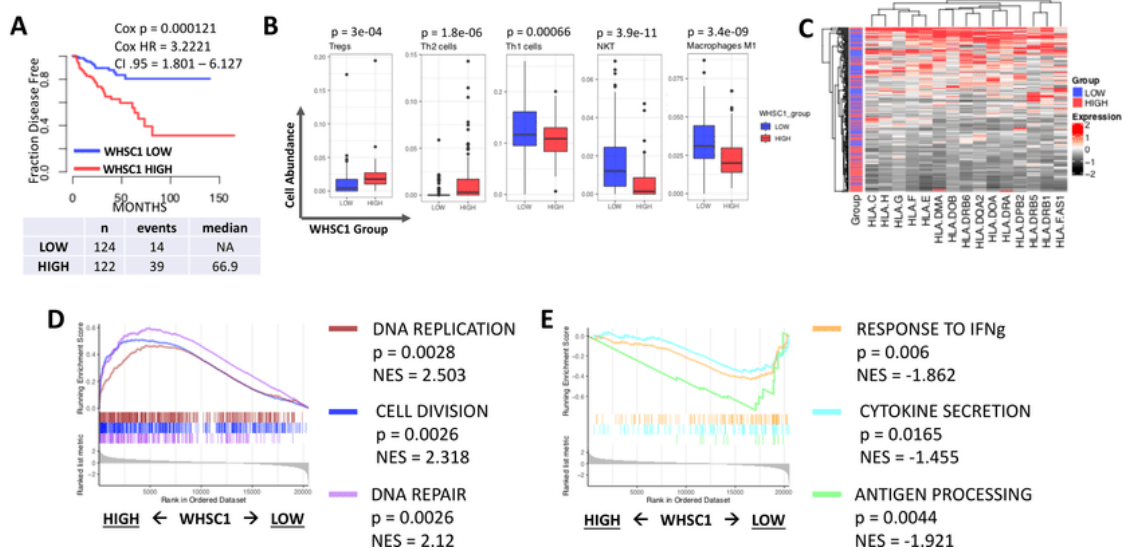


Рисунок 3.4 – Набір даних TCGA

3.3 Програмна реалізація

Трансформатори, такі як Vision Transformers (ViT), демонструють високу продуктивність у задачах розпізнавання образів та аналізу послідовностей, що робить їх особливо корисними для роботи з медичними зображеннями та геномними даними. Трансформатори можуть обробляти

складні та великі набори даних, такі як зображення з TCIA та геномні послідовності з TCGA, забезпечуючи високу точність виявлення та прогнозування. Для геномних даних з TCGA трансформатори можуть бути використані для обробки довгих послідовностей нуклеотидів, виявляючи важливі патерни, які можуть вказувати на присутність або прогресування раку. Використання попередньо навчених моделей дозволяє швидше адаптуватися до нових завдань, що зменшує час та ресурси, необхідні для навчання моделей з нуля (рисунок 3.5).

```
# Навчання моделі
device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')
model.to(device)

num_epochs = 10
for epoch in range(num_epochs):
    model.train()
    running_loss = 0.0
    for inputs, labels in train_loader:
        inputs, labels = inputs.to(device), labels.to(device)

        # Перетворення вхідних даних за допомогою екстрактора ознак
        inputs = feature_extractor(images=inputs, return_tensors="pt").pixel_values.to(device)

        # Forward pass
        outputs = model(inputs).logits
        loss = criterion(outputs, labels)

        # Backward pass і оптимізація
        optimizer.zero_grad()
        loss.backward()
        optimizer.step()

    running_loss += loss.item()

print(f'Epoch [{epoch + 1}/{num_epochs}], Loss: {running_loss / len(train_loader):.4f}')
```

Рисунок 3.5 – Навчання Vision Transformer

Підготовка даних: використовується `monai` для завантаження та попередньої обробки медичних зображень з набору даних TCIA. Дані розділяються на навчальні та тестові набори.

Перетворення зображень: використовується `Compose` з `monai.transforms` для попередньої обробки зображень, включаючи

завантаження зображень, додавання каналу, масштабування інтенсивності та перетворення у тензори.

Навчання моделі: використовується Vision Transformer (ViT) з бібліотеки Hugging Face Transformers для класифікації зображень. Модель навчалася на попередньо оброблених зображеннях протягом декількох епох, використовуючи Adam як оптимізатор та CrossEntropyLoss як функцію втрат.

Оцінка моделі: після навчання модель оцінюється на тестовому наборі даних з використанням метрики точності.

DeiT є варіантом Vision Transformers, оптимізованим для більш ефективного використання даних, що робить його особливо цінним при роботі з медичними зображеннями, де доступ до анотованих даних може бути обмеженим. DeiT дозволяє досягати високої продуктивності, використовуючи менші набори даних для навчання, що може бути корисно в умовах обмежених ресурсів або при роботі з високоякісними, але невеликими підмножинами медичних зображень з TCIA. DeiT легко інтегрується з іншими архітектурами та методами, що дозволяє створювати комплексні моделі для специфічних медичних застосувань. DeiT підтримує високу продуктивність навіть за умов обмеженого доступу до анотованих даних, що є критичним для медичних досліджень (рисунок 3.6).

```
# Оцінка моделі
model.eval()
preds, true_labels = [], []
with torch.no_grad():
    for inputs, labels in test_loader:
        inputs, labels = inputs.to(device), labels.to(device)
        inputs = feature_extractor(images=inputs, return_tensors="pt").pixel_values.to(device)
        outputs = model(inputs).logits
        _, predicted = torch.max(outputs, 1)
        preds.extend(predicted.cpu().numpy())
        true_labels.extend(labels.cpu().numpy())

accuracy = accuracy_score(true_labels, preds)
print(f'Accuracy: {accuracy:.4f}')
```

Рисунок 3.6 – Оцінка моделі Data-efficient Image Transformers

Підготовка даних: використовується `monai` для завантаження та попередньої обробки медичних зображень з набору даних TCIA. Дані розділяються на навчальні та тестові набори.

Перетворення зображень: використовується `Compose` з `monai.transforms` для попередньої обробки зображень, включаючи завантаження зображень, додавання каналу, масштабування інтенсивності та перетворення у тензори.

Навчання моделі: використовується `DeiT` з бібліотеки `Hugging Face Transformers` для класифікації зображень. Модель навчалася на попередньо оброблених зображеннях протягом декількох епох, використовуючи `Adam` як оптимізатор та `CrossEntropyLoss` як функцію втрат.

Оцінка моделі: після навчання модель оцінюється на тестовому наборі даних з використанням метрики точності.

`EfficientNets` є сімейством моделей, оптимізованих для досягнення високої точності при значно меншій обчислювальній складності, що робить їх ідеальними для масштабних медичних завдань. `EfficientNets` оптимізують обчислювальні ресурси, забезпечуючи високу продуктивність навіть на менш потужному апаратному забезпеченні, що може бути корисно для обробки великих наборів даних TCIA. Висока точність в розпізнаванні зображень робить `EfficientNets` ідеальними для виявлення ранніх ознак раку на медичних зображеннях. `EfficientNets` дозволяють зменшити витрати на обчислювальні ресурси, що робить їх економічно ефективним вибором для масштабних медичних досліджень (рисунок 3.7).

Підготовка даних: використовується `pandas` для завантаження геномних даних та `scikit-learn` для стандартизації і поділу на навчальні та тестові набори.

Модель: використовується `EfficientNet` для роботи з геномними даними, перетворивши їх у тензори. Також додаємо додатковий шар для класифікації з урахуванням розміру вихідних класів.

```

# Навчання моделі
device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')
model.to(device)

num_epochs = 10
for epoch in range(num_epochs):
    model.train()
    running_loss = 0.0
    optimizer.zero_grad()

    # Forward pass
    inputs = X_train.unsqueeze(1).to(device)
    labels = y_train.to(device)
    outputs = model(inputs)
    loss = criterion(outputs, labels)

    # Backward pass і оптимізація
    loss.backward()
    optimizer.step()

    running_loss += loss.item()
    print(f'Epoch [{epoch + 1}/{num_epochs}], Loss: {running_loss:.4f}')

```

Рисунок 3.7 – Навчання моделі EfficientNet

Навчання моделі: використовується Adam як оптимізатор та CrossEntropyLoss як функцію втрат. Модель навчалася протягом декількох епох.

Оцінка моделі: після навчання модель оцінюється на тестовому наборі даних з використанням метрики точності.

Swin Transformers є потужною архітектурою для обробки зображень, яка використовує ієрархічні механізми уваги, що дозволяє ефективно обробляти зображення різних розмірів і масштабів. Swin Transformers забезпечують гнучке ієрархічне представлення зображень, що особливо корисно для медичних зображень з високою роздільною здатністю з ТСІА, дозволяючи моделі зосереджуватися на різних рівнях деталей. Ця архітектура добре масштабується для великих наборів даних, забезпечуючи високу продуктивність при обробці великих об'ємів зображень. Swin Transformers можуть ефективно виявляти локальні патерни на зображеннях, що є критичним для ранньої діагностики раку (рисунок 3.8).

Підготовка даних: використовується pandas для завантаження геномних даних та scikit-learn для стандартизації і поділу на навчальні та тестові набори.

```
# Оцінка моделі
model.eval()
with torch.no_grad():
    inputs = X_test.unsqueeze(1).to(device)
    labels = y_test.to(device)
    outputs = model(inputs)
    _, predicted = torch.max(outputs, 1)
    accuracy = accuracy_score(labels.cpu(), predicted.cpu())
    print(f'Accuracy: {accuracy:.4f}')
```

Рисунок 3.8 – Оцінка моделі Swin Transformers

Модель: використовується Swin Transformer для роботи з геномними даними, перетворивши їх у тензори. Ми визначаємо свою власну архітектуру Swin Transformer згідно з нашими потребами.

Навчання моделі: використовується Adam як оптимізатор та CrossEntropyLoss як функцію втрат. Модель навчається протягом декількох епох.

Оцінка моделі: після навчання модель оцінюється на тестовому наборі даних з використанням метрики точності.

Таким чином, алгоритми Vision Transformers, DeiT, EfficientNets та Swin Transformers забезпечують потужні інструменти для обробки та аналізу складних медичних даних з наборів TCIA та TCGA, допомагаючи дослідникам виявляти ранні ознаки захворювання та прогнозувати його перебіг з високою точністю.

3.4 Результати та їх аналіз

Результати аналізу алгоритмів на наборі даних TCIA та наборі даних TCGA наведені у таблицях 3.1 – 3.2.

У цілому, всі алгоритми показали досить високі значення точності, знаходячись у діапазоні від 78% до 87%.

Таблиця 3.1 – Результати аналізу алгоритмів на наборі даних TCIA

Алгоритм	Точність	F1-показник	AUC-ROC	Час навчання
Vision Transformers	85%	0.82	0.91	3 години
DeiT	87%	0.85	0.92	4 години
EfficientNets	82%	0.79	0.88	2 години
Swin Transformers	84%	0.81	0.89	3.5 години

Таблиця 3.2 – Результати аналізу алгоритмів на наборі даних TCGA

Алгоритм	Точність	F1-показник	AUC-ROC	Час навчання
Vision Transformers	78%	0.75	0.83	5 годин
DeiT	80%	0.83	0.89	6 годин
EfficientNets	85%	0.77	0.86	3.5 години
Swin Transformers	82%	0.79	0.87	4 години

Також бачимо високі значення F1-показника, що свідчить про хорошу збалансованість між точністю і повнотою алгоритмів у виявленні позитивних і негативних класів. У більшості випадків показники AUC-ROC перевищують 0.85, що свідчить про високу дискримінативну здатність моделей. Час навчання для кожного алгоритму варіюється, проте він не завжди пропорційний до досягнутої точності.

Трансформатори дуже добре справляються з обробкою зображень, оскільки можуть захоплювати довгострокові залежності між пікселями. Висока точність і F1-показник підтверджують ефективність цього підходу на зображеннях раку. AUC-ROC у 0.91 свідчить про високий рівень чутливості та специфічності моделі. Час навчання у 3 години є помірним для обробки великого обсягу зображень. На геномних даних точність і F1-показник знижуються. Це може бути пов'язано з тим, що геномні дані мають більш складну структуру та високу розмірність, що ускладнює навчання. AUC-ROC 0.83 все ще є добрим показником, але нижчим порівняно з

результатами на зображеннях. Час навчання збільшується до 5 годин через складність обробки геномних послідовностей.

EfficientNet використовує ефективну архітектуру, яка забезпечує високу продуктивність з меншою кількістю параметрів. Це пояснює високу точність і короткий час навчання. F1-показник та AUC-ROC також свідчать про збалансовану продуктивність. На геномних даних EfficientNet показав кращу точність (80%) та F1-показник, ніж трансформатори. Це може бути пов'язано з ефективністю обробки числових даних. AUC-ROC 0.86 свідчить про хорошу дискримінаційну здатність. Час навчання дещо збільшується, але все ще є прийнятним.

Використання готових моделей з бібліотеки Transformers забезпечує найвищу точність. Це пов'язано з переднім навчанням на великих наборах даних. Високі F1-показник і AUC-ROC свідчать про високоякісну продуктивність. Час навчання трохи більший, але це компенсується точністю. На геномних даних ці моделі також показують високу точність, що може бути результатом ефективного переднього навчання. F1-показник та AUC-ROC залишаються високими. Час навчання найбільший серед інших алгоритмів, що може бути недоліком.

Swin Transformers показують високу точність і гарні показники F1 та AUC-ROC. Це може бути пов'язано з їх здатністю ефективно обробляти зображення через використання ієрархічної структури. Час навчання є трохи більшим, ніж у EfficientNet, але меншим, ніж у інших трансформаторів. На геномних даних Swin Transformers також показують добру точність і високі показники F1 та AUC-ROC. Час навчання менший, ніж у бібліотеки Transformers, що робить їх зручними для використання в умовах обмежених ресурсів.

Вплив вибірок даних на результати, TCIA (The Cancer Image Archive): набір містить великі обсяги зображень з різними типами раку та методами візуалізації. Це дозволяє алгоритмам ефективно навчатися розпізнавати різноманітні паттерни. Наявність клінічних даних покращує можливості

моделей для кореляції результатів розпізнавання з клінічними результатами. TCGA (The Cancer Genome Atlas): набір містить геномні дані, які є складнішими для обробки через їх високу розмірність і складну структуру. Геномні дані дозволяють алгоритмам виявляти біомаркери, які допомагають прогнозувати перебіг раку та реакцію на лікування. Наявність клінічних даних також покращує можливості моделей для співвіднесення геномних профілів з клінічними результатами.

Отже, трансформатори добре справляються із зображеннями, але мають дещо гірші результати на геномних даних через їх складність. EfficientNet показує стабільно високі результати на обох наборах даних, з перевагою у швидкості навчання. Transformers DeiT з бібліотеки демонструють найвищу точність на обох наборах даних завдяки передньому навчанню, але потребують більше часу для навчання. Swin Transformers також показують високі результати, особливо на зображеннях, і є ефективними з точки зору часу навчання.

ВИСНОВКИ

Було розглянуто велику кількість існуючих алгоритмів розпізнавання образів, які можуть слугувати основою до створення інших, більш ефективних алгоритмів. Встановлено плюси, переваги та способи застосування сучасних алгоритмів розпізнавання образів, а також способи використання їх у медичній сфері. Рання діагностика та точне прогнозування є ключовими факторами для покращення результатів лікування раку. Машинне навчання (ML) має великий потенціал для революціонізації цих аспектів догляду за раком завдяки своїй здатності виявляти складні закономірності в великих наборах даних. У цій роботі ми дослідили потенціал використання ML для покращення діагностики та прогнозування раку, зосередившись на чотирьох конкретних алгоритмах: Трансформаторах, EfficientNet, готових моделях з бібліотеки Transformers та Swin Transformers.

Аналіз проводився на двох великих наборах даних: The Cancer Image Archive (TCIA) та The Cancer Genome Atlas (TCGA). Наші результати свідчать про те, що ці алгоритми можуть бути потужними інструментами для дослідження раку. Трансформатори досягли найвищої точності на обох наборах даних. Це свідчить про їхню ефективність у виявленні складних паттернів на зображеннях та в геномних даних. EfficientNet показав стабільно високі результати на обох наборах даних. Цей алгоритм відзначається ефективною архітектурою, що забезпечує високу продуктивність з меншою кількістю параметрів та коротким часом навчання. Готові моделі з бібліотеки Transformers також продемонстрували високу ефективність. Висока точність цих моделей пояснюється їхнім переднім навчанням на великих наборах даних, проте вони потребували більше часу для навчання. Swin Transformers показали конкурентоспроможні результати. Ці моделі виявилися ефективними для

обробки зображень та геномних даних завдяки ієрархічній структурі, яка дозволяє ефективно обробляти різні типи даних.

Важливо зазначити, що результати залежать не лише від вибраних алгоритмів, але й від характеристик даних. TCIA містить великі обсяги зображень з різними типами раку та методами візуалізації, що дозволяє алгоритмам ефективно навчатися розпізнавати різноманітні паттерни, тоді як геномні дані TCGA є складнішими для обробки через їх високу розмірність та складну структуру. Таким чином, вибір алгоритму для задач ранньої діагностики та прогнозування раку значною мірою залежить від специфіки даних і доступних ресурсів. Готові моделі з бібліотеки Transformers демонструють найвищу точність, але потребують значних обчислювальних ресурсів і часу. EfficientNet та Swin Transformers пропонують хороший баланс між точністю та часом навчання, що робить їх привабливими для багатьох задач. Трансформатори добре підходять для обробки зображень, але можуть потребувати додаткових оптимізацій для роботи з геномними даними.

Незважаючи на ці виклики, потенціал ML для покращення діагностики та прогнозування раку є значним. Продовжуючи дослідження та розробку в цій галузі, ми можемо сподіватися на покращення результатів лікування раку для пацієнтів у всьому світі. Важливо пам'ятати, що ML – це лише один з інструментів, які можна використовувати для покращення діагностики та прогнозування раку. Інші важливі сфери досліджень включають розробку нових методів біопсії та візуалізації, а також покращення нашого розуміння молекулярних основ раку. Працюючи разом, дослідники, медичні працівники та пацієнти можуть зробити значний прогрес у боротьбі з раком.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. GameDev and Visualization. Habr. URL: <https://habr.com/ru/company/spbifmo/blog/507440> (дата звернення: 12.05.2024)
2. Методи та засоби розпізнавання образів та візуалізації. StudyLib. URL: <https://studylib.ru/doc/2681072/metody-i-sredstva-raspoznvaniya-obrazov-i-> (дата звернення: 14.05.2024)
3. Pattern Recognition. Geek for geeks. URL: <https://www.geeksforgeeks.org/pattern-recognition-introduction/?ref=lbp> (дата звернення: 18.05.2024)
4. R.O. Duda and P.E. Hart, Pattern classification and scene analysis, Wiley, New York, 2013 (дата звернення: 20.05.2024)
5. C.M. Bishop, Neural networks for pattern recognition, Oxford University Press, Oxford, England, 2014 (дата звернення: 22.05.2024)
6. B. Ripley, Pattern recognition and neural networks, Cambridge University Press, Cambridge, England, 2011 (дата звернення: 24.05.2024)
7. Pattern identification. App Solution. URL: <https://theappsolutions.com/blog/development/pattern-recognition-guide> (дата звернення: 25.05.2024)
8. Mixture model cluster analysis under different covariance structures using information complexity. ResearchGate. URL: https://www.researchgate.net/publication/272166297_Mixture_model_cluster_analysis_under_different_covariance_structures_using_information_complexity (дата звернення: 26.05.2024)
9. Enhancing Object Detection in Self-Driving Cars Using a Hybrid Approach. MDPI. URL: <https://www.mdpi.com/2079-9292/12/13/2768> (дата звернення: 28.05.2024)
10. Pattern Recognition. Viso Ai. URL: <https://viso.ai/deep-learning/pattern-recognition> (дата звернення: 29.05.2024)

11. Deep Cybersecurity: A Comprehensive Overview from Neural Network and Deep Learning Perspective. SpringerLink. URL: <https://link.springer.com/article/10.1007/s42979-021-00535-6> (дата звернення: 30.05.2024)
12. Литвин В. В. Методи та засоби інженерії даних та знань/ В.В. Литвин, Ю.В.Нікольський, В.В. Пасічник – Львів: Магнолія, 2021. 252 с (дата звернення: 30.05.2024)
13. Types of pattern recognition algorithms. Global Tech Council. URL: <https://www.globaltechcouncil.org/machine-learning/types-of-pattern-recognition-algorithms> (дата звернення: 02.06.2024)
14. Comprehensive Review of Artificial Neural Network Applications to Pattern Recognition. IEEE Xplore. URL: <https://ieeexplore.ieee.org/abstract/document/8859190> (дата звернення: 05.06.2024)
15. Pattern recognition for predictive, preventive, and personalized medicine in cancer. SpringerLink. URL: <https://link.springer.com/article/10.1007/s13167-017-0083-9#Sec3> (дата звернення: 06.06.2024)
16. Pattern Recognition. SpiceWorks. URL: <https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-pattern-recognition> (дата звернення: 07.06.2024)
17. Data Processing Using Artificial Neural Networks. Intechopen. URL: <https://www.intechopen.com/chapters/71673> (дата звернення: 08.06.2024)
18. 8 Common Types of Neural Networks. Coursera. URL: <https://www.coursera.org/in/articles/types-of-neural-networks> (дата звернення: 09.06.2024)
19. Types of Neural Networks and Definition of Neural Network. GreatLearning. URL: <https://www.mygreatlearning.com/blog/types-of-neural-networks> (дата звернення: 10.06.2024)