МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ В. Н. КАРАЗІНА
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ КОМП'ЮТЕРНОЇ ФІЗИКИ ТА ЕНЕРГЕТИКИ

Кафедра інформаційних технологій в фізико-енергетичних системах

ПОЯСНЮВАЛЬНА ЗАПИСКА

до кваліфікаційної роботи
за освітньо-кваліфікаційним рівнем магістр

«Метод покращення зображення в умовах слабкого освітлення на основі глибокого навчання»

Виконав: студент 2 курсу, групи НФ-63
Цехмейструк А. М.
спеціальність:
105 Прикладна фізика та наноматеріали
освітня програма: Прикладна фізика енергетичних систем
Керівник: доцент Нємченко Є. К.

Рецензент:

Харків – 2023

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Харківський національний університет імені В. Н. Каразіна

Навчально-науковий інститут комп'ютерної фізики та енергетики
Кафедра інформаційних технологій в фізико-енергетичних системах
Рівень вищої освіти (освітньо-кваліфікаційний рівень) магістр
Спеціальність 105 Прикладна фізика та наноматеріали
Освітня програма: Прикладна фізика енергетичних систем

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Руслан СУХОВ

підпис

"____"_____2023 року

З А В Д А Н Н Я
НА КВАЛІФІКАЦІЙНУ РОБОТУ

Цехмейструк Антон Миколайович

1. Тема роботи «Метод покращення зображення в умовах слабкого освітлення на основі глибокого навчання» керівник роботи: Нємченко Єгор Костянтинович, кандидат фізико-математичних наук, доцент. Затверджено наказом по університету від 12 вересня 2023 року № 4201–5/2491.

2. Строк подання студентом роботи 07 грудня 2023

3. Перелік питань, які потрібно вирішити:

1) Ознайомитись з сучасними алгоритмами використання штучного інтелекту при визначенні параметрів зображень;

2) Дослідити основні способи визначення діагностичних параметрів з результатів обробки зображень;

3) Запропонувати методи поліпшення якості обробки зображень в умовах високих шумів;

4) Дослідити використання стандартних нейронних мереж для задач визначення форми і розмірів об’єктів на зображеннях;

5) Розробити власну методику послідовного використання різних нейронних мереж з ціллю покращення діагностичних властивостей алгоритму;

6) Провести аналіз результатів використання запропонованого алгоритму послідовного використання нейронних мереж на реальних зображеннях.

4. План роботи

№ з/п	Назви етапів роботи
1	Вибір теми.
2	Затвердження теми на кафедрі
3	Робота над літературою та іншими джерелами
4	Збір фактичного матеріалу для створення вступу та першого розділу
5	Підготовка матеріалів для другого розділу
6	Постановка задач, їх розв’язання та представлення результатів. Формування третього розділу
7	Формулювання висновків та формування списку цитованої літератури
8	Підготовка усної доповіді та представлення диплому науковому керівнику
9	Ознайомлення з відгуком, рецензією, внесення остаточних виправленню, підготовка до захисту
10	Представлення роботи на попередньому захисті, остаточне корегування презентації та докладу.
11	Захист дипломної роботи

5.Дата видачі завдання **21 вересня 2023 року**

Студент _____ Антон Цехмейструк
підпис

Керівник роботи _____ Єгор Нємченко
підпис

ПЕРЕЛІК СКОРОЧЕНЬ ТА АБРЕВІАТУР

ШІ – штучний інтелект;

МН – машинне навчання;

НМ – нейронна мережа;

ГН – глибоке навчання;

ШНМ – штучна нейронна мережа;

КЗ – комп’ютерний зір;

РЕФЕРАТ

Антон Цехмейструк «Метод покращення зображення в умовах слабкого освітлення на основі глибокого навчання»

Кваліфікаційна робота магістра з прикладної фізики. – Х. Харківський національний університет імені В. Н. Каразіна, 2023. – 67 с.

Мета дипломної роботи полягала в розробці алгоритму використання ансамблю нейронних мереж в послідовному порядку з ціллю обробки зображень, які здобути при низької інтенсивності випромінювання. В роботі проведено ознайомлення з сучасними алгоритмами використання штучного інтелекту при визначенні параметрів зображень. Зокрема, було досліджено використання стандартних нейронних мереж для задач визначення форми і розмірів об’єктів на зображеннях. Це дозволило визначити основні недоліки використання нейронних мереж для обробки зображень з великим рівнем шумів. В роботі було досліджено основні способи визначення діагностичних параметрів з результатів обробки зображень. В результаті цього дослідження було запропоновано нові методи поліпшення якості обробки зображень в умовах високих шумів і розроблено власну методику послідовного використання різних нейронних мереж з ціллю покращення діагностичних властивостей алгоритму. Запропоновано методику було втілено на практиці, а також проведено аналіз результатів використання запропонованого алгоритму послідовного використання нейронних мереж на реальних зображеннях.

КЛЮЧОВІ СЛОВА: ШТУЧНИЙ ІНТЕЛЕКТ, МАШИННЕ НАВЧАННЯ, ГЛИБОКЕ МАШИННЕ НАВЧАННЯ, НЕЙРОНА МЕРЕЖА, АНСАМБЛЬ НЕРОННИХМЕРЕЖ

ABSTRACT

Anton Tsehmeistruk "A method of image improvement in low light conditions based on deep learning" Master's thesis in applied physics. – Kh. Kharkiv National University named after V. N. Karazin, 2023. – 67 p.

The aim of the thesis was to develop an algorithm for using an ensemble of neural networks in a sequential order for processing images obtained at low radiation intensity. The paper introduces modern algorithms for the use of artificial intelligence in determining image parameters. In particular, the use of standard neural networks for the tasks of determining the shape and size of objects in images was investigated. This made it possible to determine the main disadvantages of using neural networks for processing images with a high level of noise. The main methods of determining diagnostic parameters from the results of image processing were investigated in the paper. As a result of this study, new methods of improving the quality of image processing in conditions of high noise were proposed and a proprietary method of sequential use of different neural networks was developed in order to improve the diagnostic properties of the algorithm. The proposed methodology was put into practice, and the results of using the proposed algorithm of sequential use of neural networks on real images were also analyzed.

KEY WORDS: ARTIFICIAL INTELLIGENCE, MACHINE LEARNING, DEEP MACHINE LEARNING, NEURON NETWORK, NEURON NETWORK ENSEMBLE

ЗМІСТ

ВСТУП…….. 8
1. АНАЛІТИЧНИЙ ОГЛЯД ……... 11
1.1 Методи глибокого навчання…… 11
1.2 Глибоке навчання і зображення…….. 16
2. ШТУЧНИЙ ІНТЕЛЕКТ І ОБРОБКА ЗОБРАЖЕНЬ…….. 25
2.1. Поняття штучних нейронних мереж …….. 25
2.1.1. Перцептрон і його побудова……. 31
2.1.2. Основні методи побудови нейронних мереж……. 34
2.1.3. Особливості використання МН при обробці зображень… 38
2.2. Методи глибокого навчання…….. 39
2.3.Особливості використання ШТ при обробці зображень…….. 42
2.4. Параметри визначення якості обробки ……. 47
3. ЗАСТОСУВАННЯ РОЗРОБЛЕНОГО АЛГОРИТМУ……... 49
3.1. Дослідження використання однорівневої НМ…… 49
3.1.1. Машинне навчання…….. 49
3. 2. Використання ансамблю НМ…..……...…… 51
3.2.1. Збір та підготовка даних…… 51
3.2.2. Порівняння нейромереж …… 52
3.2.3. Аналіз мережі U-Net ……... 53
3.2.4. Аналіз мережі Vgg 16…….. 55
3.2.5. Аналіз мережі MobileNet 16…… 56
3.3 **Реалізація основної ідеї і обговорення результатів……** 57
ВИСНОВКИ……... 58
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ……... 59

ВСТУП

Актуальність теми. Сьогодні важко представити область діяльності, у якій можна обійтися без комп'ютерної обробки зображень. При комп'ютерній обробці зображень вирішується широке коло завдань, таких як поліпшення якості зображень; вимірювання параметрів зображення; спектральний аналіз багатомірних сигналів; розпізнавання зображень; стиск зображень.

Методи покращення зображень є надзвичайно важливими для сучасної науки, вони вважаються одними з таких, які безперервно розвиваються та вдосконалюються. При цьому під обробкою зображень розуміють не лише поліпшення зорового сприйняття зображень, але й класифікацію об'єктів, що виконується при аналізі зображень.

Багато галузей техніки, що мають відношення до отримання, обробки, зберігання та передачі інформації, в даний час орієнтуються на розвиток систем, в яких інформація може надаватися у вигляді зображень. Зображення, яке можна розглядати як двовимірний сигнал, є значно більш ємним носієм інформації, ніж звичайний одновимірний, найчастіше тимчасової, сигнал. Разом з тим, рішення наукових і інженерних задач при роботі з візуальними даними вимагає особливих підходів, що спираються на знання специфіки зображень, оскільки традиційні способи обробки та аналізу одновимірних сигналів мало придатні в цих випадках. Однією з основних проблем, що виникають при обробці зображень, є поліпшення якості зображення, зокрема, підвищення розрізнення окремих фрагментів.

До причин, що знижує якість зображень, можна віднести:

- технічні характеристики яскравості перешкоди шумового характеру;
- недостатня або надмірна освітленість об'єктів зйомки;
- відсутність різкості при отриманні зображення;
- занадто дрібні розміри деталей, які необхідно розрізнити.

Основною метою комп'ютерної обробки зображень є знаходження методів, результат роботи яких виявився б більш підходящим з точки зору конкретного застосування.

Метою роботи є дослідження методів та алгоритмів покращення зображень. Об'єктом дослідження є системи та алгоритми які використовуються для покращення вхідних зображень. У даній роботі запропоновано використання методу покращення вхідних зображень що ґрунтується на основі нейронної мережі GAN. Перевага цього методу в тому що він може використовуватися для поліпшення якості нечітких або частково зіпсованих зображень. Проведено дослідження результатів покращення зображень на базі нейронної мережі GAN. У результаті роботи здійснена програмна реалізація системи для покращення вхідних зображень з використанням нейронної мережі GAN.

Запропонований у роботі алгоритм послідовного використання нейронних мереж дозволяє поступово поліпшувати якість алгоритму визначення параметрів об’єктів на зображенні з великим рівнем шумів. Для досягнення поставленої **мети** – визначення форми та розмірів об’єктів на зображеннях, які получено з малою інтенсивністю випромінювання, було визначено такі **задачі**:

1) Ознайомитись з сучасними алгоритмами використання штучного інтелекту при визначенні параметрів зображень;

2) Дослідити основні способи визначення діагностичних параметрів з результатів обробки зображень;

3) Запропонувати методи поліпшення якості обробки зображень в умовах високих шумів;

4) Дослідити використання стандартних нейронних мереж для задач визначення форми і розмірів об’єктів на зображеннях;

5) Розробити власну методику послідовного використання різних нейронних мереж з ціллю покращення діагностичних властивостей алгоритму;

6) Провести аналіз результатів використання запропонованого алгоритму послідовного використання нейронних мереж на реальних зображеннях.

Об’єкт дослідження.

Використання ансамблю нейронних мереж задачах покращення зображення в умовах слабкого освітлення на основі глибокого навчання.

Предмет дослідження.

Нові методи покращення зображення в умовах слабкого освітлення на основі глибокого навчання.

Методи дослідження.

Послідовне використання нейронних мережа для визначення об’єктів на зображеннях з великим рівнем шумів

Наукова новизна одержаних результатів.

Було представлене використання методу використання ансамблю нейронних мереж для обробки зображень, які були здобуті в умовах слабкого освітлення.

Практичне значення отриманих результатів:

Розроблений в дипломі алгоритм був використаний в практичних застосуваннях для визначення форми об’єктів на реальних рентгенівських зображеннях.

Структура і обсяг роботи. Магістерська робота складається зі вступу, трьох розділів, висновків, списку використаних джерел та додатків. Загальний обсяг роботи становить 61 сторінку, із них 55 сторінки основного тексту сторінок, 8 рисунків за текстом, 1 таблицю, список використаних джерел із 28 найменування на 3 сторінках.

1. АНАЛІТИЧНИЙ ОГЛЯД

1.1. Методи глибокого навчання

Покращення зображення (Image Enhancement) - процес поліпшення якості картинки без втрати інформації до отримання бажаного візуального результату (дозвіл, колір і стиль) або підготовка фото для подальшого аналізу у різних програмах комп'ютерного зору: розпізнавання об'єктів, класифікації, інтерпретації зображень. Підвищення якості зображення зазвичай включає ряд перетворень: шумозаглушення, поліпшення розмитого фото, підвищення розширення, контрастності, освітлення темної фотографії, усунення оптичних спотворень тощо.

Deep Photo Enhancer — покращення зображення поганої якості за допомогою нейромереж.

Deep Photo Enhancer використовує GAN (Generative Adversarial Networks - генеративно-змагальні нейронні мережі) і непарні алгоритми навчання. Deep Photo Enhancer пропонує метод поліпшення зображень на основі вивчення фотографій. Нейрони вчаться знаходити загальні характеристики в наборі зразків зображень (наприклад, рівень контрастності, баланс білого, колірна гамма) і потім застосовувати ці характеристики до поліпшеного зображенню так, щоб зберігався сенс оригінальної картинки. Цей метод вимагає використання вихідних зображень високої якості, і його можна у подальшому персоналізувати.

Алгоритми глибокого навчання дозволяють навчити нейронну мережу збільшувати зображення в 2 або навіть 4 рази, що дозволяє підвищити якість знімків із низьким розширенням. Глибокі нейронні мережі (GAN, глибокі рекурсивні згорткові нейронні мережі (DRCN)) здатні відновлювати фотореалістичні текстури із зображень із дуже низьким розширенням: нейронна мережа "домальовує" деталі за результатами навчання на прикладах зображень. Це не на 100% точна реконструкція фотографій у HD, однак результати вражають.

IBM/MAX Image Resolution Enhancer —нейромережа для відновлення зжатих фото.

Ця модель, що розгортується, дозволяє збільшувати розмір пікселізованого зображення в 4 рази, одночасно генеруючи фотореалістичні деталі, використовуючи GAN (SRGAN-tensorflow), навчений на 600 тисячах знімків OpenImages V4. Ідеальним вхідним зображенням повинен бути PNG-файл розміром від 100x100 до 500x500 пікселів, бажано без постобробки. Модель може генерувати деталі з пікселізованого зображення, але вона не підходить для корекції розмитих зображень [1].

Глибо́ке навча́ння (англ. *deep learning*, також іноді **глибинне навча́ння**) — це частина ширшого сімейства методів машинного навчання, яка ґрунтується на штучних нейронних мережах із навчанням подань. Прикметник «глибоке» (англ. *"deep"*) у глибокому навчанні стосується використання декількох шарів у мережі. Використовувані методи можуть бути керованими, напівкерованими, та некерованими.

Архітектури глибокого навчання, такі як глибокі нейронні мережі, глибокі мережі переконань, глибоке навчання з підкріпленням, рекурентні нейронні мережі, згорткові нейронні мережі та трансформери, застосовували в таких галузях як комп'ютерне бачення, розпізнавання мовлення, обробка природної мови, машинний переклад, біоінформатика, конструювання ліків, аналіз медичних зображень[en], кліматологія, перевірка матеріалів і програми настільних ігор, де вони дали результати, порівнянні, а в деяких випадках і кращі за продуктивність експертів — людей.

Штучні нейронні мережі (ШНМ) було натхненно обробкою інформації та розподіленими комунікаційними вузлами в біологічних системах. ШНМ мають різноманітні відмінності від біологічного мозку. Зокрема, штучні нейронні мережі мають тенденцію бути статичними та

символьними, тоді як біологічний мозок більшості живих організмів динамічний (пластичний) та аналоговий [2].

Глибоке навчання намагається імітувати людський мозок, дозволяючи системам кластеризувати дані та робити неймовірно точні прогнози. Це підмножина машинного навчання, яка навчає комп'ютер виконувати завдання, подібні до людських, наприклад розпізнавання мовлення чи ідентифікації зображень. Завдяки глибокому навчанню системи можуть покращити свої здібності класифікувати, розпізнавати, виявляти й описувати за допомогою даних.

Глибоке навчання відіграє велику роль у багатьох сучасних технологіях, таких як Alexa та Siri. Він передбачає навчання даних комп'ютера за допомогою глибоких алгоритмів для автономного навчання шляхом розпізнавання шаблонів за допомогою рівнів обробки. Алгоритми глибокого навчання також використовують процеси градієнтний спуск і зворотне поширення, щоб стати більш точним. Це також дозволяє їм робити прогнози на основі нових даних, з якими вони ніколи не стикалися.

Моделі глибокого навчання можуть виконувати різні типи методів навчання. Наприклад, вони можуть пройти непідконтрольне навчання, для якого не потрібні мічені набори даних. Цей метод навчання дає змогу моделям виявляти шаблони в даних і групувати їх за певними характеристиками без допомоги людини [3].

VCLabs Photo Enhancer AI — це спеціальний інструмент для підвищення якості зображення, який використовує глибоке навчання для підвищення роздільної здатності зображення на 400% і покращення деталей, а також усуває дефекти та артефакти стиснення. Його було навчено на мільйонах зображень різних типів, включаючи портрети, дику природу, аніме, пейзажі, архітектуру тощо, що дозволяє природним чином підвищувати роздільну здатність різних зображень із збереженням правильних і відповідних детале

Інструмент пропонує функцію інтелектуального видалення шуму, яка може автоматично видаляти шум із зернистих фотографій і покращувати якість.

Модель AI Face Refinement створена спеціально для покращення портретної фотографії, не турбуючись про складний процес редагування. Завдяки штучному інтелекту ви можете одним клацанням миші отримувати високоякісні приголомшливі портрети, покращуючи тон шкіри, видаляючи плями або покращуючи такі деталі, як очі та губи.

Upscale — це інструмент штучного інтелекту для зображень, який підвищує роздільну здатність зображення без зниження якості. Він здатний збільшити ваше зображення до будь-якого бажаного розміру до 4 разів. Завдяки алгоритму AI він може додавати деталі до ваших зображень, щоб вони виглядали професійно та зберігали природні деталі зображення.

Однією з головних функцій Upscale є його здатність видаляти артефакти JPEG. Коли ви редагуєте та зберігаєте зображення JPEG кілька разів, це може призвести до пікселізації та погіршення якості. Цей процес, який називається артефактингом, може зробити ваші зображення нечіткими. Upscale усуває стиснення, щоб згладити краї.

BigJPG — чудовий інструмент, розроблений, щоб допомогти вам вирішити проблему нечітких зображень та ілюстрацій аніме. Це ще один інструмент для збільшення масштабу аніме-зображень штучного інтелекту, який може обробляти аніме-зображення та ілюстрації перед їх збільшенням без втрати якості.

Процес покращує кольори, деталі, краї тощо. Але важливо зазначити, що інструмент має обмеження 3000 x 3000 пікселів для масштабування [4].

Глибоке навчання: Визначення глибокого навчання та його зв’язок з нейромережами.

Глибоке навчання — це розвиток і розширення класичних нейромереж, що забезпечує здатність розпізнавати, аналізувати та класифікувати більш складні шаблони та ієрархії. Глибокі нейромережі

складаються з багатьох шарів нейронів, які взаємодіють між собою, передаючи інформацію від входів до виходів.

Переваги та обмеження глибокого навчання.

Перевагами глибокого навчання є його висока точність та здатність обробляти великі обсяги даних. Глибокі нейромережі демонструють надзвичайну ефективність в розпізнаванні образів, обробці мови та інших задачах, які класичні алгоритми вирішують із труднощами. Завдяки глибокому навчанню стали можливі прориви в таких областях, як комп'ютерний зір, автономні автомобілі та машинний переклад.

Однак глибоке навчання має свої обмеження.

По-перше, воно потребує величезних обсягів даних та обчислювальної потужності для навчання та оптимізації моделей.

По-друге, глибокі нейромережі можуть бути складними для інтерпретації та пояснення, що ускладнює зрозуміння того, як вони приймають рішення.

По-третє, може виникнути проблема перенавчання, коли модель надто добре "запам'ятовує" навчальні дані, але погано узагальнює нову інформацію.

Попри ці обмеження, глибоке навчання продовжує розвиватися та вносити значні зміни у багатьох галузях науки та техніки. Дослідники активно працюють над вдосконаленням алгоритмів глибокого навчання, зменшенням вимог до обчислювальної потужності та поліпшенням зрозумілості та інтерпретації моделей.

Однією з таких тенденцій є використання технік передачі навчання (transfer learning), які дозволяють використовувати знання, отримані під час навчання однієї моделі, для швидкого навчання інших моделей у схожих задачах. Це може сприяти ефективному використанню ресурсів та підвищенню загальної продуктивності глибокого навчання.

Також з'являються нові архітектури нейромереж, які намагаються відтворити більш точні моделі людського мозку та його роботи, такі як

капсульні мережі (capsule networks) та спайкові нейронні мережі (spiking neural networks). Вони можуть привести до створення ще більш потужних та ефективних систем глибокого навчання.

Під час навчання нейромережі слід дотримуватися кількох основних принципів. По-перше, розділіть дані на навчальний, валідаційний та тестовий набори, щоб мати можливість оцінити ефективність моделі та уникнути перенавчання. По-друге, використовуйте методи оптимізації, такі як градієнтний спуск або адаптивні методи, які допомагають знаходити оптимальні параметри моделі. По-третє, під час налаштування гіперпараметрів, таких як швидкість навчання, кількість шарів або кількість нейронів, використовуйте методи, такі як перехресна перевірка (cross-validation) або пошук на сітці (grid search), щоб знайти оптимальні значення для вашої моделі.

Також важливо контролювати процес навчання нейромережі, використовуючи такі інструменти, як візуалізація функції втрат (loss function) та метрик точності (accuracy metrics). Це дозволяє вам слідкувати за успішністю навчання та виявляти проблеми з перенавчанням або недонавчанням. За потреби можна зупинити навчання або змінити гіперпараметри для поліпшення результатів.

Враховуючи всі ці аспекти роботи з нейромережами, ви зможете побудувати ефективні та надійні моделі, які допоможуть вирішити складні задачі та принести цінні результати у різних областях застосування. З практичним досвідом та застосуванням кращих практик навчання та налаштування нейромереж, ви зможете вдосконалити свої навички та зробити значний внесок у розвиток штучного інтелекту [5].

1.2. Глибоке навчання і зображення

Глибоке навчання — це один з багатьох підходів до машинного навчання, що включає в себе кілька технологій. Однією з основних таких

революційних технологій при роботі з зображеннями є згорткові нейронні мережі (convolutional neural networks, CNN).

Пояснити, що таке згорткові нейронні мережі, досить складно. Це нове покоління алгоритмів розпізнавання зображень, поява яких була завдяки дослідженням в тій області кори головного мозку тварин, яка відповідає за зір. Ці мережі імітують те, як мозок отримує абстрактну інформацію із сигналів, що генеруються оптичними нервами ока. Ця інформація може мати форму квадратів, трикутників і т. д., а також силуету людини, автомобіля або будь-якого іншого об’єкта.

Як і людський мозок, згорткові нейронні мережі навчаються на власному досвіді. Таким чином, розробникам більше не потрібно вручну створювати шаблони або елементи, які потрібно шукати в зображенні. Навпаки, за наявності великої кількості мічених зображень — набору для навчання — мережа самостійно освоює риси, які в сукупності складають будь-який об’єкт.

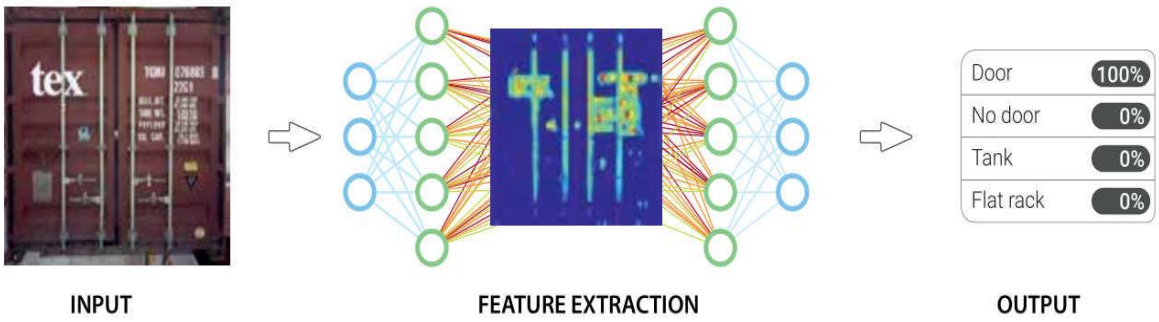


Рис. 1: Приклад класифікації контейнерів, який ілюструє принцип наскрізного навчання. Зображення знаходиться на вході, результат його обробки – на виході. У проміжку між ними мережа самостійно вивчає риси певного класу, які мають значення, на основі тисяч завантажених зображень. У разі дверцят контейнера це логотип компанії, стрижні і петлі, номер контейнера та інша текстова інформація. Теплова карта центральної частини контейнера показує, що отримана мережа дуже добре навчилася отримувати ці визначальні риси

Які саме ці функції, залежить від набору даних і архітектури мережі. Наприклад, автомобільний детектор можна навчати даних про окремі частини, з яких складається автомобіль, таких як шини, радіатор, і навчити його розпізнавати форми і текстури кожної з цих частин і комбінувати їх.

Процес керованого навчання вимагає, щоб Camco забезпечувала мережі великою кількістю високоякісних зображень і приміток до них (об’єктів, що містяться в зображенні, і їх розташувань в зображенні). Процеси збору високоякісних даних, а також оптимізації як даних, так і моделей і процедур навчання – це складні процеси, що вимагають досвіду і ресурсів. Але результати їх реалізації досить корисні [6].

У глибокому навчанні кожен рівень вчиться перетворювати свої вхідні дані на дещо абстрактніше й складніше подання. У застосунку для розпізнавання зображень сирий вхід може бути матрицею пікселів; перший шар подання може абстрагуватися від пікселів і кодувати контури; другий шар може складати та кодувати розташування контурів; третій шар може кодувати ніс та очі; а четвертий шар може розпізнавати, що зображення містить обличчя. Важливо, що процес глибокого навчання може *самостійно* вчитися, які ознаки оптимально розмістити на якому рівні. Це не звільняє від необхідності ручного налаштування; наприклад, різна кількість та розміри шарів можуть забезпечувати різні ступені абстракції.

Слово «глибоке» у «глибокому навчанні» позначує кількість шарів, крізь які перетворюються дані. Точніше, системи глибокого навчання мають значну глибину *шляху розподілу внеску* (ШРВ, англ. *credit assignment path, CAP*). ШРВ — це ланцюг перетворень від входу до виходу. ШРВ описують потенційно причинно-наслідкові зв'язки між входом та виходом. Для нейронної мережі прямого поширення глибина ШРВ дорівнює глибині мережі та є кількістю прихованих шарів плюс один (оскільки шар виходу також параметризовано). Для рекурентних нейронних мереж, у яких сигнал може поширюватися крізь шар понад один раз, глибина ШРВ потенційно необмежена.

Деякі архітектури глибокого навчання демонструють проблематичну поведінку, наприклад, впевнене класифікування невпізнанних зображень як належних до знайомої категорії звичайних зображень (2014) та неправильне класифікування незначних збурень правильно класифікованих зображень (2013). Ґьорцель припустив, що така поведінка зумовлена обмеженнями у їхніх внутрішніх поданнях, і що ці обмеження перешкоджатимуть інтегруванню до гетерогенної багатокомпонентної архітектури загального штучного інтелекту (ЗШІ). Можливо, ці проблеми можна розв'язати за допомогою архітектур глибокого навчання, які внутрішньо утворюють стани, гомологічні розкладам граматики зображень спостережуваних об'єктів та подій. Виведення граматики (візуальної чи мовної) з тренувальних даних було би рівнозначним обмеженню системи міркуваннями здорового глузду, які оперують поняттями в термінах граматичних породжувальних правил, і є основною метою як засвоєння мови людиною, так і штучного інтелекту (ШІ).

Більшість систем глибокого навчання покладаються на тренувальні та контрольні (англ. *verification*) дані, породжувані та/або розмічувані людьми. У філософії засобів масової інформації стверджують, що для цієї мети регулярно використовують не лише низькооплачувану клікпрацю (наприклад, на *Amazon Mechanical Turk*), а й неявні форми людської мікропраці, які часто не визнають як таку. Філософ Райнер Мюльхоф розрізняє п'ять типів «машинного заволодівання» людською мікропрацею для породжування тренувальних даних:

1) ігрофікація (вбудовування розмічування або обчислювальних завдань у потік гри),

2) «захоплювання та відстежування» (наприклад, *CAPTCHA* для розпізнавання зображень, або відстежування кліків на сторінках результатів пошуку *Google*),

3) використання соціальних мотивів (наприклад, позначування облич у *Facebook* для отримування позначених зображень обличчя),

4) розроблення інформації (наприклад, за допомогою пристроїв самооцифровування[en], таких як відстежувачі активності),

5) клікпрацю [7].

Серед уже відомих архітектур глибокого навчання можна відзначити згорткові нейронні мережі (CNN) для аналізу зображень, рекурентні нейронні мережі (RNN) для послідовного опрацювання даних і трансформаторні моделі для розуміння природної мови.

Тенденції розвитку глибокого навчання формувалися під впливом цілої низки факторів, включаючи:

Доступність великих масивів даних. Наявність великомасштабних помічених наборів даних, таких як ImageNet для комп'ютерного зору або Common Crawl для опрацювання природної мови, стала вирішальним фактором для навчання моделей глибокого навчання.

Збільшення обчислювальної потужності. Наявність високопродуктивних графічних процесорів (GPU) і спеціалізованого обладнання, такого як TPU (Tensor Processing Units), дала змогу дослідникам і практикам навчати моделі глибокого навчання у більших масштабах і швидшими темпами.

Пошук нейронної архітектури. Алгоритми NAS у поєднанні з обчислювальними ресурсами дали змогу дослідникам автоматично знаходити нові та оптимізовані мережеві архітектури для розв'язання відповідних задач .

Ці фактори сприяли швидкому зростанню і розвитку напрямків глибокого навчання. Це призвело до прориву в таких задачах, як розпізнавання зображень і мови, розуміння природної мови, генеративне моделювання і навчання з підкріпленням.

Інтеграція глибокого навчання з іншими технологіями

Моделі глибокого навчання будуть все частіше поєднуватися з технологіями AR і VR для створення імерсивного досвіду та

інтелектуальних віртуальних середовищ. Ця інтеграція дасть змогу створювати такі додатки, як:

- розпізнавання і відстеження об'єктів у реальному часі
- розуміння того, що відбувається, і взаємодія з урахуванням контексту в AR і VR

Об'єднання глибокого навчання і блокчейна також набиратиме обертів у 2024 році. Цьому сприяє активний процес регуляції блокчейна в традиційний сектор фінансів, медицини, поставок тощо.

Інтеграція гібридних моделей

Інтеграція гібридних моделей - це процес об'єднання різних типів, моделей та архітектур глибокого навчання для використання їхніх індивідуальних переваг і підвищення загальної продуктивності.

Подібні тенденції глибокого навчання у 2024 році сприяють масштабуванню, популярності та створенню більш ефективних підходів у цій галузі. Гібридні моделі доповнюють одна одну для вирішення складних завдань і досягнення кращих результатів.

Розглянемо кілька прикладів інтеграції гібридних моделей у глибокому навчанні:

- Додавання моделей. Цей підхід передбачає навчання декількох моделей глибокого навчання незалежно одна від одної і подальше об'єднання їхніх результатів у вигляді кластера. Кожна модель може мати різну архітектуру або бути навчена на різних підмножинах даних.
- Попередньо навчені моделі та тонке налаштування. Ці моделі мають загальні візуальні характеристики, які можуть бути перенесені на нове завдання. Тонке налаштування або їхня адаптація дає змогу ефективно використовувати вивчені уявлення і прискорити збіжність.
- Об'єднання архітектур. Різні архітектури глибокого навчання можуть бути об'єднані для використання їхніх унікальних характеристик. Наприклад, об'єднання рекурентної нейронної мережі (RNN) і згорткової

• нейронної мережі (CNN) дає змогу моделі вловлювати просторові та часові залежності в послідовних даних.

Інтеграція гібридних моделей у глибокому навчанні вимагає ретельного врахування поставленої задачі, характеристик наявних моделей, а також ресурсів і обмежень системи.

Глибоке навчання на основі нейронауки

Глибоке навчання на основі нейронауки - це тип ML, у якому для навчання штучних нейронних мереж використовуються дані нейронаукових експериментів. Це дає змогу дослідникам розробляти моделі, які засновані на роботі людського мозку.

Вона передбачає використання принципів і концепцій, заснованих на вивченні мозку і нейронних систем, для поліпшення архітектури, алгоритмів і загальної продуктивності моделей глибокого навчання.

Розглянемо кілька ключових аспектів глибокого навчання на основі нейронауки:

• Архітектура нейронних мереж. Технологія спрямована на розробку архітектур нейронних мереж, які відображають різні структури зв'язків, що спостерігаються в мозку. Наприклад, конволюційні нейронні мережі натхненні ієрархічним опрацюванням даних у зоровій корі, а рекурентні нейронні мережі - рекурентними зв'язками в мозку.

• Алгоритми навчання. Нейронаука дає уявлення про те, як мозок навчається і обробляє інформацію. Включаючи ці принципи в алгоритми глибокого навчання, дослідники прагнуть підвищити його ефективність.

• Когнітивні та поведінкові аспекти. Розуміння того, як мозок сприймає, обробляє і взаємодіє з навколишнім середовищем, може допомогти в розробці моделей глибокого навчання для таких завдань, як розпізнавання образів, розуміння природної мови або навчання з підкріпленням.

Глибоке навчання на основі нейронауки - це міждисциплінарна галузь, яка долає розрив між нейронаукою і глибоким навчанням [8].

Переваги та обмеження глибокого навчання.

Перевагами глибокого навчання є його висока точність та здатність обробляти великі обсяги даних. Глибокі нейромережі демонструють надзвичайну ефективність в розпізнаванні образів, обробці мови та інших задачах, які класичні алгоритми вирішують із труднощами. Завдяки глибокому навчанню стали можливі прориви в таких областях, як комп'ютерний зір, автономні автомобілі та машинний переклад.

Однак глибоке навчання має свої обмеження.

По-перше, воно потребує величезних обсягів даних та обчислювальної потужності для навчання та оптимізації моделей.

По-друге, глибокі нейромережі можуть бути складними для інтерпретації та пояснення, що ускладнює зрозуміння того, як вони приймають рішення.

По-третє, може виникнути проблема перенавчання, коли модель надто добре "запам'ятовує" навчальні дані, але погано узагальнює нову інформацію.

Попри ці обмеження, глибоке навчання продовжує розвиватися та вносити значні зміни у багатьох галузях науки та техніки. Дослідники активно працюють над вдосконаленням алгоритмів глибокого навчання, зменшенням вимог до обчислювальної потужності та поліпшенням зрозумілості та інтерпретації моделей.

Однією з таких тенденцій є використання технік передачі навчання (transfer learning), які дозволяють використовувати знання, отримані під час навчання однієї моделі, для швидкого навчання інших моделей у схожих задачах. Це може сприяти ефективному використанню ресурсів та підвищенню загальної продуктивності глибокого навчання [9].

Google Research презентували технологію зменшення шуму на зображенні, яка може суттєво покращити якість фотографії в умовах слабкого освітлення. Інструмент під назвою RawNeRF (Raw Neural Radiance Fields) розроблений на базі штучного інтелекту. Він здатний

реконструювати темну сцену з усуненням шумів, значно перевершуючи аналоги. Та найбільш дивовижним є те, технологія прибирає шуми з мінімальною втратою якості та набагато меншою кількістю видимих дефектів зображення, що виникають при стисненні знімка. Секрет нової технології в тому, що штучний інтелект навчається на необроблених даних зображення, а не на JPEG з постобробкою. Технологія здатна комбінувати зображення, зроблені з різних точок огляду камери, щоб реконструювати сцену та подивитися на неї під різними кутами.

Сцени реконструюються в лінійному колірному просторі HDR, який може опрацьовувати такі деталі, як зміна експозиції, використання показу тонів і зміна фокуса [10].

Висновок до розділу 1

У розділі було проведено аналіз структури хребтового стовпа та основних функцій опорно-рухової системи людини, приділено увагу поняттю хребців як ключових будівельних блоків хребта. Досліджено проблему остеопорозу, як прогресуючого захворювання сучасності, включаючи механізм його розвитку, симптоматику та рекомендації щодо профілактичних заходів.

Також розглянуті різні методи обстеження хребтового стовпа та внутрішньої структури людського тіла. Підкреслено вивчення рентгенівської морфометрії, визначено принцип роботи цього методу та розглянуті показники ступеня деформації тіл хребців за двома критеріями.

2. ШТУЧНИЙ ІНТЕЛЕКТ І ОБРОБКА ЗОБРАЖЕНЬ

2.1. Поняття штучних нейронних мереж

Шту́чні нейро́нні мере́жі (**ШНМ**, англ. *artificial neural network*), які зазвичай просто називають нейронними мережами (НМ, англ. *neural networks, NN*) або *нейромережами* (англ. *neural nets*), це обчислювальні системи, натхнені біологічними нейронними мережами, які складають мозок тварин.[2]

ШНМ ґрунтується на сукупності з'єднаних вузлів (англ. *units, nodes*), які називають штучними нейронами, які приблизно моделюють нейрони біологічного мозку. Кожне з'єднання, як і синапси в біологічному мозку, може передавати сигнал до інших нейронів. Штучний нейрон отримує сигнали, потім обробляє їх і може сигналізувати нейронам, з якими його з'єднано. «Сигнал» у з'єднанні це дійсне число, а вихід кожного нейрона обчислюється деякою нелінійною функцією суми його входів. З'єднання називають *ребрами* (англ. *edges*). Нейрони та ребра зазвичай мають *вагу* (англ. *weight*), яка підлаштовується в процесі навчання. Вага збільшує або зменшує силу сигналу на з'єднанні. Нейрони можуть мати такий поріг, що сигнал надсилається лише тоді, коли сукупний сигнал перевищує цей поріг.

Як правило, нейрони зібрано в шари (англ. *layers*). Різні шари можуть виконувати різні перетворення даних свого входу. Сигнали проходять від першого шару (шару входу) до останнього (шару виходу), можливо, після проходження шарами декілька разів.

Пошук нейронної архітектури (ПНА, англ. *neural architecture search, NAS*) використовує машинне навчання для автоматизування побудови ШНМ. Різні підходи до ПНА побудували мережі, добре порівнянні з системами, розробленими вручну. Основним алгоритмом цього пошуку є пропонувати модель-кандидатку, оцінювати її за набором даних, і

використовувати результати як зворотний зв'язок для навчання мережі ПНА. Серед доступних систем — АвтоМН та *AutoKeras*.

До проблем побудови належать визначення кількості, типу та з'єднаності рівнів мережі, а також розміру кожного, та типу з'єднання (повне, агрегувальне, …).

Гіперпараметри також слід визначати як частину побудови (їх не навчаються), керуючи такими питаннями як кількість нейронів у кожному шарі, темп навчання, крок, крок фільтрів (англ. *stride*), глибина, рецептивне поле та доповнення (для ЗНМ) тощо. Великі й ефективні нейронні мережі вимагають значних обчислювальних ресурсів. У той час як мозок має апаратне забезпечення, ідеально пристосоване для задачі обробки сигналів графом нейронів, імітація навіть спрощеного нейрону на архітектурі фон Неймана може споживати величезну кількість пам'яті та дискового простору. Крім того, розробникові часто потрібно передавати сигнали багатьма цими з'єднаннями та пов'язаними з ними нейронами, що вимагає величезної обчислювальної потужності та часу ЦП.

Шмідхубер зазначив, що відродження нейронних мереж у двадцять першому сторіччі значною мірою обумовлено досягненнями в апаратному забезпеченні: з 1991 до 2015 року обчислювальна потужність, особливо забезпечувана ГПЗП (на ГП), зросла приблизно в мільйон разів, зробивши стандартний алгоритм зворотного поширення придатним для навчання мереж, на кілька рівнів глибших, ніж раніше. Використання прискорювачів, таких як ПКВМ та ГП, може скорочувати тривалість тренування з місяців до днів.

Нейроморфна інженерія або фізична нейронна мережа розв'язує проблему апаратного забезпечення безпосередньо, створюючи мікросхеми, відмінні від фон неймановых, для безпосереднього втілення нейронних мереж у схемах. Ще одна мікросхема, оптимізована для обробки нейронних мереж, зветься тензорним процесором або ТП (англ. *Tensor Processing Unit, TPU*) [11].

В основі нейромереж лежать штучні нейрони, які об’єднуються в графові структури і передають сигнали один одному через ваги зв’язків. Завдяки процесу навчання, під час якого ваги та зміщення між нейронами оптимізуються, нейромережі стають здатними до виявлення закономірностей та залежностей у вхідних даних. Нейромережі активно використовуються в таких галузях, як комп’ютерний зір, машинний переклад, розпізнавання мови та інше. Нейромережа (Штучна нейронна мережа або Нейронка) — це комп’ютерна система, яка вчиться та вдосконалюється, намагаючись стати схожою на людський мозок у вирішенні різних задач.

Структура нейромережі складається з трьох основних типів шарів: вхідний шар, приховані шари та вихідний шар.

Шар у нейромережі — це група нейронів, які працюють разом і виконують певну функцію в мережі. Уявіть собі нейромережу як мультиповерховий будинок, де кожен поверх містить кімнати (нейрони). Шари з’єднані між собою, як сходи між поверхами. Кожен шар має свою роль у процесі обробки інформації.

Є три основні типи шарів у нейромережі:

Вхідний шар: Це “перший поверх” будинку, де ми починаємо. Вхідний шар приймає дані ззовні, наприклад, зображення або текст, і передає їх у наступні шари. Вхідний шар не змінює дані, а лише служить точкою входу для них.

Приховані шари: Це “середні поверхи” будинку. Приховані шари забезпечують обробку вхідних даних і передачу інформації між шарами. Вони називаються “прихованими”, оскільки їхні результати не відображаються напряму на виході мережі. Кількість прихованих шарів та нейронів у них може варіюватися в залежності від складності задачі та архітектури мережі.

Вихідний шар: Це “останній поверх” будинку. Вихідний шар формує результат, який нейромережа передбачає на основі вхідних даних. Результат може бути класифікацією, числовим значенням або іншою інформацією, залежно від типу задачі.

Отже, шари в нейромережі – це групи нейронів, які працюють разом і відповідають за різні етапи обробки інформації. Вони забезпечують нейромережі можливість адаптуватися до різних задач.

Усі нейрони в шарах з’єднані між собою через ваги зв’язків. Ваги відіграють важливу роль у навчанні нейромережі, оскільки вони визначають силу впливу одного нейрона на інший. В процесі навчання, ваги оптимізуються, щоб мінімізувати помилку передбачення мережі.

Крім ваг зв’язків, кожен нейрон має так зване зміщення (bias), яке дозволяє регулювати активацію нейрона незалежно від вхідного сигналу. Зміщення допомагає нейромережі легше адаптуватися до різних даних та виконувати більш гнучкі перетворення на вхідних даних.

Ще одним ключовим компонентом нейромережі є функція активації. Вона застосовується до кожного нейрона у прихованих та вихідних шарах, щоб визначити його активність на основі суми вхідних сигналів, помножених на відповідні ваги та додавання зміщення. Функція активації може бути лінійною або не лінійною, в залежності від типу задачі та архітектури мережі. Деякі з найпопулярніших функцій активації включають сигмоїду, гіперболічний тангенс, ReLU (Rectified Linear Unit) та Softmax.

Роль нейромереж у вирішенні проблем.

Нейромережі відіграють важливу роль у вирішенні проблем, оскільки вони можуть вивчати та адаптуватися до різних типів даних і вирішувати складні завдання. Їх головна мета полягає в тому, щоб знайти шаблони в даних і застосовувати ці знання для передбачення або класифікації нових даних.

Типи завдань, які виконують нейромережі.

Нейромережі можуть виконувати різні типи завдань, залежно від архітектури мережі та навчальних даних. До основних видів завдань належать:

Класифікація: Визначення категорії, до якої належить певний об’єкт або подія, на основі його характеристик.

Регресія: Прогнозування числового значення на основі вхідних даних.

Генерація тексту**:** Створення тексту на основі навчальних даних, зазвичай використовується для автоматичного створення описів, статей або відповідей на запитання [12].

Обробка зображень: Розпізнавання об’єктів, тексту або осіб на зображеннях.

Через постійну динаміку нейрональної активності реалізується здатність людини сприймати та обробляти емоції, виконувати розумові операції та інші важливі когнітивні функції. Нейрони, як базові структурні одиниці нервової системи, встановлюють інформаційний обмін за допомогою нейромедіаторів через синапси, які з'єднують відомі аксони з дендритами сусідніх нейронів.

У біологічному контексті, ця взаємодія нейронів може бути порівняна із витонченою хореографією, де кожен елемент ємно впливає на динаміку системи в цілому. З точки зору штучних нейронних мереж, біологічне уявлення нейрона може бути абстраговане до моделі, що включає інтеграцію сигналів через вагові коефіцієнти, що визначають силу зв'язків між нейронами.

Розглядаючи нейрон в контексті штучної нейронної мережі, важливим є розуміння взаємодії між вхідними та вихідними сигналами, а також принципів моделювання, які відтворюють основні аспекти функціонування біологічного нейрона. Такий підхід дозволяє досліджувати та аналізувати роботу штучних нейронних мереж з погляду їхньої взаємодії та інформаційного обміну, узагальнюючи принципи, визначені в біологічних системах. (Рис. 2.1.) [13-17].

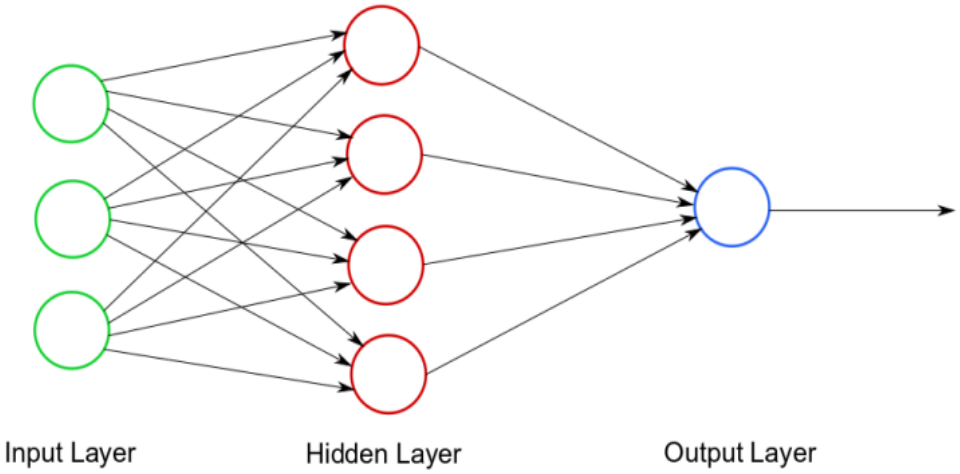


Рис. 2.1. Будова нейронної мережі

Нейрон, як математична модель та основна обчислювальна одиниця мережі, виконує математичні операції над середньозваженою вхідною інформацією та застосовує зміщення. Після цих операцій результат пропускається через функцію активації, як, наприклад, нелінійна функція сигмоїда. Цей підхід часто використовується в бінарній класифікації.

Нейронна мережа складається з:

- вхідного шару (одного)
- прихованого шару (з різноманітною кількістю)
- вихідного шару (одного)

Кожен шар взаємодіє з попереднім та наступним шарами, утворюючи структуру, що забезпечує ефективну обробку інформації.

Така організація нейронної мережі дозволяє здійснювати складні обчислення та вирішувати завдання, такі як класифікація, з високою точністю.

Кожен нейрон, або, інакше кажучи, нейронний вузол, утворює зв'язок з іншими вузлами та характеризується своєю вагою та порогом. Вхідний шар є початком нейронної мережі, отримуючи вхідну інформацію і, через функцію активації, оброблює її, передаючи до наступного шару нейронів та

порівнюючи з встановленим порогом. Прихований шар, якщо їх декілька, отримує інформацію від вхідного шару, або може взаємодіяти з попереднім прихованим шаром. Отримана вхідна інформація проходить через нелінійну функцію активації перед тим, як передати вихідну інформацію далі.

Важливими параметрами нейрона є вага та поріг. Наприклад, при змаганні 10 нейронів, перемагає той, чий ваговий коефіцієнт вищий. Підбір вагових коефіцієнтів допомагає оптимізувати навчання штучної нейронної мережі. Вихідний шар слугує кінцевою логікою мережі, оброблюючи вихідну інформацію від попередніх шарів та надаючи очікуваний результат. Залежно від мети та типу мережі, цей вихід може бути кінцевим результатом або подаватися на вхід до іншої нейронної мережі (каскад). [13-17].

2.1.1. Перцептрон і його побудова

Одношарова нейронна мережа прямого зв’язку була представлена наприкінці 1950-х років Френком Розенблатом. Це був початковий етап глибокого навчання та штучних нейронних мереж. Протягом цього часу для прогнозування використовується статистичне машинне навчання або традиційне кодове програмування. Персептрон — одна з перших і найпростіших моделей штучних нейронних мереж. Незважаючи на те, що перцептрон є простою моделлю, було доведено, що він успішний у вирішенні конкретних проблем категоризації [13 – 17].

Персептрон є однією з найпростіших архітектур штучних нейронних мереж. Він був представлений Франком Розенблатом у 1957 році. Це найпростіший тип прямої нейронної мережі, що складається з одного шару вхідних вузлів, які повністю з’єднані з шаром вихідних вузлів. Він може вивчати лінійно роздільні шаблони. він використовує дещо інші типи штучних нейронів, відомих як порогові логічні одиниці (TLU). вперше він був представлений Маккалохом і Волтером Піттсом у 1940-х роках.

Розглянемо два основних типів персептрона

Одношаровий персептрон: цей тип персептрона обмежений вивченням лінійно роздільних шаблонів. ефективний для завдань, де дані можна розділити на окремі категорії через пряму лінію.

Багатошаровий персептрон: багатошарові персептрони мають розширені можливості обробки, оскільки вони складаються з двох або більше рівнів, вправно обробляючи складніші моделі та зв’язки всередині даних.

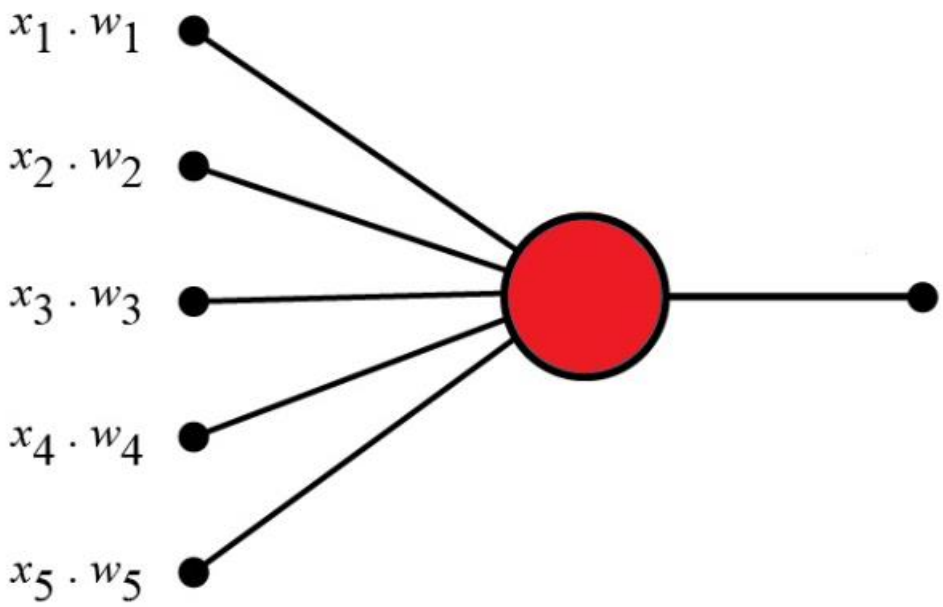


Рис. 2.2. Типовий приклад побудови персептрона

Тепер можна перейти до розглядання основних компонентів персептрона.

Персептрон, основна одиниця нейронної мережі, містить важливі компоненти, які взаємодіють в обробці інформації.

Вхідні функції: Персептрон використовує кілька вхідних функцій, кожна вхідна функція представляє характеристику або атрибут вхідних даних.

Вагові коефіцієнти: кожна вхідна функція пов’язана з вагою, що визначає значення кожної вхідної функції у впливі на вихідні дані

персептрона. Під час навчання ці ваги коригуються, щоб засвоїти оптимальні значення.

Функція підсумовування: перцептрон обчислює зважену суму своїх вхідних даних за допомогою функції підсумовування. Функція підсумовування поєднує вхідні дані з відповідними вагами, щоб створити зважену суму.

Функція активації: потім зважена сума передається через функцію активації. Perceptron використовує ступінчасті функції Хевісайда. які приймають підсумкові значення як вхідні дані та порівнюють із порогом і надають на виході 0 або 1.

Вихід: кінцевий вихід персептрона визначається результатом функції активації. Наприклад, у задачах двійкової класифікації вихід може представляти прогнозований клас (0 або 1).

Зміщення: термін зміщення часто включається в модель персептрона. Зміщення дозволяє моделі робити коригування, які не залежать від вхідних даних. Це додатковий параметр, який вивчається під час навчання.

Алгоритм навчання (правило оновлення ваги): під час навчання перцептрон навчається, регулюючи свої ваги та зміщення на основі алгоритму навчання. Поширеним підходом є алгоритм навчання персептрона, який оновлює ваги на основі різниці між прогнозованим виходом і справжнім виходом.

Ці компоненти працюють разом, щоб дозволити перцептрону навчатися та робити прогнози. Хоча один персептрон може виконувати двійкову класифікацію, складніші завдання вимагають використання кількох персептронів, організованих у шари, утворюючи нейронну мережу.

Тепер розглянемо алгоритм роботи перцептрона.

Кожному вхідному вузлу персептрона призначається вага, що вказує на значення цього входу для виходу. Вихідні дані персептрона є зваженою сумою вхідних даних, які пройшли через функцію активації,

щоб вирішити, чи спрацює персептрон. він обчислює зважену суму своїх вхідних даних як:

$$\sum_{i=1}^{m} x_i w_i \,, \tag{2.1}$$

Функція кроку порівнює цю зважену суму з порогом, який виводить 1, якщо вхідні дані більші за порогове значення, і 0 в іншому випадку. Це функція активації, яку персептрони використовують найчастіше. Найбільш поширеною ступінчастою функцією, яка використовується в перцептроні, є ступінчаста функція Хевісайда.

Після отримання зваженої суми, вона проходить через функцію активації для нормалізації даних. Це необхідно, оскільки нейронна мережа оперує у визначеному діапазоні чисел, який може бути (0,1) або (-1,1), залежно від обраної функції активації. Для забезпечення правильної роботи мережі та уникнення виходу за межі цього діапазону, дані піддаються процедурі нормалізації.

$$Y = f\left(\sum_{i=1}^{m} x_i w_i\right), \tag{2.2}$$

де f – функція активації функції, m – кількість входів до перцептрону, x_i – вектор вхідних значень, w_i – вектор дійсних ваг.

2.1.2. Основні методи побудови нейронних мереж

Почнемо з поняття багатошарового персепрона. У багатошаровому персептроні (MLP) персептрони розташовані у взаємопов’язаних шарах. Вхідний рівень збирає вхідні шаблони. Вихідний рівень має класифікації або вихідні сигнали, на які можуть відображатися вхідні моделі. Наприклад, шаблони можуть містити список величин для технічних індикаторів цінного паперу; потенційні результати можуть бути «купити», «тримати» або «продавати».

Приховані шари точно налаштовують вхідні ваги, доки межа похибки нейронної мережі не стане мінімальною. Існує гіпотеза, що приховані шари екстраполюють помітні характеристики у вхідних даних, які мають передбачувану силу щодо вихідних даних. Це описує вилучення ознак, яке виконує корисність, подібну до статистичних методів, таких як аналіз головних компонентів.

Різні способи поєднання шарів дозволяють створити велику кількість різних НМ. Розглянемо деякі з них.

Нейронні мережі прямого зв’язку

Нейронні мережі прямого зв’язку є одним із найпростіших типів нейронних мереж. Він передає інформацію в одному напрямку через вхідні вузли; ця інформація продовжує оброблятися в цьому єдиному напрямку, поки не досягне режиму виведення. Нейронні мережі прямого зв’язку можуть мати приховані рівні для функціональності, і цей тип найчастіше використовується для технологій розпізнавання обличчя.

Рекурентні нейронні мережі

Більш складний тип нейронних мереж, рекурентні нейронні мережі отримують вихідні дані вузла обробки та передають інформацію назад у мережу. Це призводить до теоретичного «навчання» та вдосконалення мережі. Кожен вузол зберігає історичні процеси, і ці історичні процеси повторно використовуються в майбутньому під час обробки.

Це стає особливо критичним для мереж, у яких передбачення є неправильним; система намагатиметься дізнатися, чому стався правильний результат, і відповідно коригуватиметься. Цей тип нейронної мережі часто використовується в програмах перетворення тексту в мовлення.

Згорткові нейронні мережі

Згорткові нейронні мережі, також звані ConvNets або CNN, мають кілька рівнів, на яких дані сортуються за категоріями. Ці мережі мають вхідний рівень, вихідний рівень і приховану безліч згорткових шарів між ними. Шари створюють карти функцій, які записують ділянки зображення,

які далі розбиваються, доки вони не отримають цінні результати. Ці рівні можуть бути об’єднані або повністю з’єднані, і ці мережі є особливо корисними для програм розпізнавання зображень.

Деконволюційні нейронні мережі

Деконволюційні нейронні мережі просто працюють навпаки згорткових нейронних мереж. Застосування мережі полягає у виявленні елементів, які могли бути визнані важливими в згортковій нейронній мережі. Ці елементи, ймовірно, були б відкинуті під час процесу виконання згорткової нейронної мережі. Цей тип нейронної мережі також широко використовується для аналізу або обробки зображень.

Модульні нейронні мережі

Модульні нейронні мережі містять кілька мереж, які працюють незалежно одна від одної. Ці мережі не взаємодіють одна з одною під час процесу аналізу. Натомість ці процеси виконуються для більш ефективного виконання складних, складних обчислювальних процесів. Подібно до інших модульних галузей, таких як модульна нерухомість, мета незалежності мережі полягає в тому, щоб кожен модуль відповідав за певну частину загальної більшої картини.

Тепер є можливість перейти до різних методів застосування НМ.

Нейронні мережі широко використовуються з додатками для фінансових операцій, планування підприємства, торгівлі, бізнес-аналітики та обслуговування продуктів. Нейронні мережі також отримали широке застосування в бізнес-додатках, таких як рішення для прогнозування та маркетингових досліджень, виявлення шахрайства та оцінка ризиків.

Нейронна мережа оцінює дані про ціни та виявляє можливості для прийняття торгових рішень на основі аналізу даних. Мережі можуть розрізняти тонкі нелінійні взаємозалежності та шаблони, які інші методи технічного аналізу не можуть. Згідно з дослідженнями, точність нейронних мереж у прогнозуванні цін на акції відрізняється. Деякі моделі прогнозують правильні ціни акцій у 50-60% випадків, тоді як інші є точними в 70% усіх

випадків. Деякі стверджують, що підвищення ефективності на 10% — це все, чого інвестор може вимагати від нейронної мережі.

Що стосується фінансів, то нейронні мережі можуть обробляти сотні тисяч бітів даних про транзакції. Це може призвести до кращого розуміння обсягу торгів, торгового діапазону, кореляції між активами або встановлення очікувань волатильності для певних інвестицій. Оскільки людина може не мати змоги ефективно переглядати багаторічні дані (іноді зібрані через секундні інтервали), нейронні мережі можуть бути розроблені для визначення тенденцій, аналізу результатів і прогнозування майбутніх змін вартості класу активів.

Тепер розглянемо осноні переваги та недоліки НМ.

Переваги нейронних мереж

Нейтральні мережі, які можуть працювати безперервно та є більш ефективними, ніж люди або простіші аналітичні моделі. Нейронні мережі також можна запрограмувати на навчання на попередніх виходах для визначення майбутніх результатів на основі подібності до попередніх вхідних даних.

Нейронні мережі, які використовують хмарні онлайн-сервіси, також мають переваги зменшення ризиків порівняно з системами, які покладаються на апаратне забезпечення локальних технологій. Крім того, нейронні мережі часто можуть виконувати кілька завдань одночасно (або, принаймні, розподіляти завдання, які повинні виконуватися модульними мережами одночасно).

Нарешті, нейронні мережі постійно розширюються в нові програми. Хоча ранні теоретичні нейронні мережі були дуже обмежені в застосуванні в різних галузях, сьогодні нейронні мережі використовуються в медицині, науці, фінансах, сільському господарстві чи безпеці.

Недоліки нейронних мереж

Хоча нейтральні мережі можуть покладатися на онлайн-платформи, для створення нейронної мережі необхідний апаратний компонент. Це

створює фізичний ризик для мережі, яка залежить від складних систем, вимог до налаштування та можливого фізичного обслуговування.

Хоча складність нейронних мереж є сильною стороною, це може означати, що на розробку певного алгоритму для конкретного завдання потрібні місяці (якщо не більше). Крім того, може бути важко помітити будь-які помилки або недоліки в процесі, особливо якщо результати є оцінками або теоретичними діапазонами.

Нейронні мережі також може бути важко перевірити. Деякі процеси нейронної мережі можуть виглядати як «чорна скринька», де вводяться вхідні дані, мережі виконують складні процеси та повідомляють про результати. Людям також може бути важко проаналізувати слабкі сторони в процесі обчислення або навчання мережі, якщо мережі бракує загальної прозорості щодо того, як модель вивчає попередні дії.

Таким чином, можна стверджувати, що нейронні мережі — це складні інтегровані системи, які можуть виконувати аналітику набагато глибше та швидше, ніж здатна людина. Існують різні типи нейронних мереж, які часто найкраще підходять для різних цілей і цільових результатів. У фінансах нейронні мережі використовуються для аналізу історії транзакцій, розуміння руху активів і прогнозування результатів фінансового ринку.

2.1.3. Особливості використання МН при обробці зображень

Розпізнавання зображень (або класифікація зображень) — це завдання ідентифікації зображень і класифікації їх в одному з кількох заздалегідь визначених окремих класів. Отже, програмне забезпечення та програми для розпізнавання зображень можуть визначити, що зображено на зображенні, і відрізнити один об’єкт від іншого.

Галузь дослідження, спрямована на забезпечення машин цією здатністю, називається комп’ютерним зором. Будучи одним із завдань комп’ютерного зору (КЗ), класифікація зображень служить основою для вирішення різних проблем КЗ, включаючи:

2.2. Методи глибокого навчання і КЗ

Дослідники можуть використовувати моделі глибокого навчання для вирішення завдань комп'ютерного зору. Глибоке навчання – це техніка машинного навчання, яка спрямована на навчання машин навчатися на прикладі. Оскільки більшість методів глибокого навчання використовують архітектури нейронних мереж, моделі глибокого навчання часто називають глибокими нейронними мережами.

Розпізнавання зображень є одним із завдань, у яких чудово працюють глибокі нейронні мережі (ГНМ). Нейронні мережі — це обчислювальні системи, призначені для розпізнавання шаблонів. Їхня архітектура натхненна структурою людського мозку, звідси й назва. Вони складаються з трьох типів шарів: вхідні, приховані шари та вихідні. Вхідний рівень отримує сигнал, прихований рівень обробляє його, а вихідний рівень приймає рішення або прогноз щодо вхідних даних. Кожен рівень мережі складається з взаємопов'язаних вузлів (штучних нейронів), які виконують обчислення.

Що робить нейронну мережу глибокою? Кількість прихованих шарів: у той час як традиційні нейронні мережі мають до трьох прихованих шарів, глибокі мережі можуть містити їх сотні.

Давайте розберемося, як нейронні мережі вчаться розпізнавати шаблони. Як зрозуміти, хто проходить на вулиці знайомий чи незнайомий (таких ускладнень, як короткозорість без урахування)? Ми дивимося на них, підсвідомо аналізуємо їхню зовнішність, і якщо якісь притаманні риси — форма обличчя, колір очей, зачіска, тип статури, хода чи навіть модний вибір — збігаються з конкретною людиною, яку ми знаємо, ми впізнаємо цю людину. Ця мозкова робота займає лише хвилину.

Отже, щоб мати можливість розпізнавати обличчя, система повинна спочатку дізнатися їхні риси. Його потрібно навчити передбачати, чи є об'єкт X чи Z. Моделі глибокого навчання вивчають ці характеристики

іншим способом, ніж моделі машинного навчання (ML). Тому й підходи до модельного навчання різні.

Тепер перейдемо до того , як проходить навчання моделей глибокого навчання (таких як нейронні мережі)

Щоб побудувати модель МН, яка може, наприклад, передбачити відтік клієнтів, дослідники даних повинні вказати, які вхідні характеристики (властивості проблеми) модель враховуватиме при прогнозуванні результату. Це може бути освіта клієнта, дохід, етап життєвого циклу, функції продукту або використовувані модулі, кількість взаємодій зі службою підтримки та їх результати. Процес побудови функцій із використанням знань предметної області називається розробкою функцій.

Якби ми навчили модель глибокого навчання, щоб побачити різницю між собакою та котом за допомогою розробки функцій… Ну, уявіть собі, що ми збираємо характеристики мільярдів котів і собак, які живуть на цій планеті. Ми не можемо побудувати точні характеристики, які працюватимуть для кожного можливого зображення, враховуючи такі ускладнення, як мінливість об’єкта, що залежить від точки огляду, безлад фону, умови освітлення або деформація зображення. Має бути інший підхід, і він існує завдяки природі нейронних мереж.

Нейронні мережі вивчають функції безпосередньо з даних, з якими вони навчаються, тому фахівцям не потрібно витягувати функції вручну.

«Потужність нейронних мереж полягає в їхній здатності вивчати представлення ваших навчальних даних і те, як найкраще пов’язати їх із вихідною змінною, яку ви хочете передбачити. У цьому сенсі нейронні мережі вивчають відображення. З математичної точки зору вони здатні вивчати будь-яку функцію відображення, і було доведено, що вони є універсальними алгоритмами наближення», – зазначає Джейсон Браунлі в прискореному курсі багаторівневих нейронних мереж персептронів.

Навчальні дані, у цьому випадку, є великим набором даних, який містить багато прикладів кожного класу зображень. Коли ми говоримо про великий набір даних, ми дійсно це маємо на увазі. Наприклад, набір даних ImageNet містить понад 14 мільйонів зображень, анотованих людиною, що представляють 21 841 поняття (набори синонімів або синсети відповідно до ієрархії WordNet), у середньому 1000 зображень на поняття.

Кожне зображення анотовано (позначено) категорією, до якої воно належить – кішка чи собака. Алгоритм досліджує ці приклади, дізнається про візуальні характеристики кожної категорії та, зрештою, дізнається, як розпізнавати кожен клас зображення. Цей модельний стиль навчання називається навчанням під наглядом.

Кожен шар вузлів тренується на виході (наборі функцій), створеному попереднім шаром. Таким чином, вузли в кожному наступному шарі можуть розпізнавати більш складні, детальні характеристики – візуальні представлення того, що зображено на зображенні. Така «ієрархія зростаючої складності та абстракції» відома як ієрархія функцій.

Отже, чим більше рівнів має мережа, тим більша її здатність передбачити.

Провідною архітектурою, яка використовується для задач розпізнавання та виявлення зображень, є згорточні нейронні мережі (CNN). Згорткові нейронні мережі складаються з кількох шарів з маленькими колекціями нейронів, кожен з яких сприймає невеликі частини зображення. Результати всіх колекцій у шарі частково перекриваються таким чином, щоб створити повне представлення зображення. Потім шар нижче повторює цей процес на новому представленні зображення, дозволяючи системі дізнатися про композицію зображення.

Історія глибоких CNN сягає початку 1980-х років. Але тільки в 2010-х роках дослідникам вдалося досягти високої точності вирішення завдань розпізнавання зображень за допомогою глибоких згорткових нейронних мереж. як? Вони почали навчати та розгортати CNN, використовуючи

графічні процесори (GPU), які значно прискорюють складні системи на основі нейронних мереж. Обсяг навчальних даних – фотографій чи відео – також збільшився, оскільки камери мобільних телефонів і цифрові фотоапарати почали швидко розвиватися і стали доступними.

Тепер ми знаємо про розпізнавання зображень та інші завдання комп'ютерного зору, а також про те, як нейронні мережі вчаться призначати мітки зображенню або кільком об'єктам на зображенні. Перейдемо до реальних застосувань цієї технології.

2.3. Особливості використання ШТ при обробці зображень

Розглянемо приклад виявлення логотипу в аналітиці соціальних мереж. Бренди відстежують текстові дописи в соціальних мережах із згадуванням свого бренду, щоб дізнатися, як споживачі сприймають, оцінюють, взаємодіють з їхнім брендом, а також що вони говорять про нього та чому. Це називається соціальним прослуховуванням. Тип соціального прослуховування, який зосереджується на моніторингу візуальних розмов, називається (барабанна дроб, будь ласка)... візуальне прослуховування.

Той факт, що понад 80 відсотків зображень у соціальних мережах із логотипом бренду не мають назви компанії в підписі, ускладнює візуальне прослуховування. Як отримати інформацію про цю справу? З розпізнаванням логотипу.

Стартап Meerkat провів експеримент, щоб показати, як розпізнавання логотипу може допомогти візуальному прослуховуванню. Протягом півроку стартапери збирали твіти зі словами, які зазвичай використовуються в контексті пива, наприклад, пиво, цервеза, барбекю, бар та інші. Вони навчили систему виявляти логотипи популярних пивних брендів: Heineken, Budweiser, Corona, Bud Light, Guinness і Stella Artois. І вони використовували його для аналізу зображень з твітів, що містять логотипи брендів.

Фахівці проіндексували метадані твітів, щоб отримати уявлення про частку ринку кожного бренду та його споживачів.

Спочатку вони порівняли кількість постів з логотипами кожного бренду з їхньою часткою ринку і виявили, що ці два параметри не пов’язані між собою. Далі фахівці витягли геокоординати для майже 73 відсотків зображень у твіттері, щоб оцінити присутність бренду в регіонах. Потім вони побудували графік відсотка кожного сорту пива для п’яти найкращих країн у наборі даних. Наприклад, Bud Light є найпопулярнішим у США, тоді як Heineken має прихильників у різних країнах із найбільшими частками в США та Великобританії. Команда також проаналізувала зображення, які містили обличчя, щоб визначити стать любителів пива. Різниця була незначною: фотографії опублікували на 1,34% більше чоловіків.

Це не лише вимірювання впізнаваності бренду. Підприємства використовують функцію виявлення логотипів, щоб обчислити рентабельність інвестицій від спонсорства спортивних подій або визначити, чи не було використано їхній логотип.

Іншим важливим прикладом застосування ГНМ є аналіз медичного зображення

Програмне забезпечення на базі моделей глибокого навчання допомагає радіологам справлятися з величезним навантаженням із інтерпретації різноманітних медичних зображень: комп’ютерної томографії (КТ) та ультразвукового сканування, магнітно-резонансної томографії (МРТ) або рентгенівських променів. IBM наголошує, що радіолог відділення невідкладної допомоги повинен оглядати щонайменше 200 випадків щодня. Крім того, деякі медичні дослідження містять до 3000 зображень. Не дивно, що медичні зображення складають майже 90 відсотків усіх медичних даних.

Радіологічні інструменти на основі штучного інтелекту не замінюють клініцистів, а допомагають їм приймати рішення. Вони позначають гострі

аномалії, ідентифікують пацієнтів з високим ризиком або тих, хто потребує термінового лікування, щоб радіологи могли визначити пріоритети своїх робочих списків.

Дослідницький підрозділ IBM у Хайфі, Ізраїль, працює над Cognitive Radiology Assistant для аналізу медичних зображень. Система аналізує медичні зображення, а потім поєднує цю інформацію з інформацією з медичних записів пацієнта та представляє результати, які рентгенологи можуть взяти до уваги під час планування лікування.

Вчені з цього відділу також розробили спеціалізовану глибоку нейронну мережу для позначення аномальної та потенційно ракової тканини молочної залози.

Aidoc пропонує ще одне рішення, яке використовує глибоке навчання для сканування медичних зображень (зокрема комп'ютерної томографії) і визначення пріоритетів у списках пацієнтів. Розчин отримав дозвіл Управління з контролю за продуктами й ліками США (FDA), Therapeutic Goods of Australia (TGA) і Європейського Союзу, отримавши маркування CE для позначення трьох небезпечних для життя станів: легеневої емболії, перелому шийного відділу хребта та внутрішньочерепного крововиливу.

Клієнтами компанії є Меморіальний медичний центр UMass у Вустері, штат Массачусетс, лікарня Монтефіоре Найак в окрузі Рокленд, штат Нью-Йорк, і центр візуалізації Global Diagnostics Australia.

Ще одним, незвичайним застосуванням ГНМ є програми для розпізнавання творів мистецтва. Наприклад, Magnus — це програма для розпізнавання зображень, яка веде любителів мистецтва та колекціонерів «крізь мистецькі джунглі». Коли користувач робить фотографію витвору мистецтва, програма надає такі відомості, як автор, назва, рік створення, розміри, матеріал і, що найважливіше, поточну та історичну ціну. Додаток також має карту з галереями, музеями та аукціонами, а також виставленими на даний момент творами мистецтва.

Магнус отримує інформацію з бази даних, що містить понад 10 мільйонів зображень творів мистецтва; інформація про речі та ціни є краудсорсингом. Цікавий факт: Леонардо Ді Капріо інвестував у додаток, повідомляє Магнус на своїй сторінці в Apple Store.

Відвідувачі музеїв можуть вгамувати свій голод знань за допомогою таких програм, як Smartify. Smartify – це музейний путівник, який можна використовувати в десятках відомих місць мистецтва світу, таких як Метрополітен-музей у Нью-Йорку, Смітсонівська національна портретна галерея у Вашингтоні, Лувр у Парижі, Рейксмузеум в Амстердамі, Королівська академія мистецтв у Лондон та інші.

Щоб розкрити подробиці про твір мистецтва, додаток порівнює відсканові твори мистецтва з цифровими зображеннями в базі даних, яка містила майже 50 000 творів мистецтва станом на 2017 рік. Співзасновник Smartify Анна Лоу пояснює, як працює додаток: «Ми скануємо твори мистецтва за допомогою фотографії або цифрові зображення, а потім створюють цифрові відбитки художнього твору, тобто його зводять до набору цифрових крапок і ліній».

Щоб розкрити подробиці про твір мистецтва, додаток порівнює відсканові твори мистецтва з цифровими зображеннями в базі даних, яка містила майже 50 000 творів мистецтва станом на 2017 рік. Співзасновник Smartify Анна Лоу пояснює, як працює додаток: «Ми скануємо твори мистецтва за допомогою фотографії або цифрові зображення, а потім створюють цифрові відбитки пальців ілюстрації, тобто вони зводяться до набору цифрових крапок і ліній."

Одни з найбільш важливих застосувань ШІ є розпізнавання облич. Розпізнавання облич стає мейнстрімом серед авіакомпаній, які використовують його для покращення посадки та реєстрації. Є два основних напрямки цих оновлень: слідувати тенденціям самообслуговування та цієї біометричної технології та зробити роботу в аеропорту безпечнішою та швидшою. Чим менше кроків повинні зробити

пасажири та персонал, щоб продовжити передпольотні процедури, тим краще. Обладнання для посадки сканує обличчя мандрівників і зіставляє їх із фотографіями, які зберігаються в базах даних прикордонного контролю (наприклад, митної та прикордонної служби США), щоб підтвердити їх особу та дані про рейс. Це можуть бути фотографії з посвідчень особи, візи чи інших документів.

American Airlines, наприклад, почали використовувати розпізнавання облич біля виходу на посадку в терміналі D міжнародного аеропорту Даллас/Форт-Ворт, штат Техас. Замість використання посадкових талонів мандрівникам сканують обличчя. Єдине, що не змінилося, це те, що для проходження перевірки безпеки як і раніше потрібно мати паспорт і квиток. У 2018 році американці тестували біометричні дані протягом 90 днів у терміналі міжнародного аеропорту Лос-Анджелеса з ідеєю розширити використання технології, якщо випробування пройде добре. Численні авіакомпанії також використовують розпізнавання облич як додатковий варіант посадки: JetBlue, British Airways, AirAsia, Lufthansa або Delta.

Тепер можна розглянути таке застосування, як візуальний пошук товару. Відтоді, як у гру з’явився візуальний пошук, межі між онлайн- та офлайн-покупками зникли. Наприклад, у додатку Urban Outfitters є функція Scan + Shop, завдяки якій споживачі можуть відсканувати предмет, який вони знайшли у звичайному магазині або надрукований у журналі, отримати його детальний опис і миттєво замовити його. Візуальний пошук також покращує досвід онлайн-покупок.

Програми з цією можливістю працюють на основі нейронних мереж. NN обробляють зображення, завантажені користувачами, і генерують описи зображень (теги), наприклад, тип одягу, тканину, стиль, колір. Описи зображень зіставляються з елементами на складі разом із відповідними тегами. Результати пошуку представлено на основі оцінки подібності.

2.4. Параметри визначення якості обробки

Функція втрат у сфері нейронних мереж — це функція, яка оцінює відхилення між передбаченнями, зробленими мережею, та реальними значеннями спостережень, які використовуються під час процесу навчання. Чим нижчий результат цієї функції, тим ефективніша нейронна мережа. Метою навчання нейронної мережі є спроба мінімізувати цю функцію, наскільки це можливо, щоб мінімізувати відхилення між прогнозованим значенням і фактичним значенням. Однією з найбільш часто використовуваних функцій втрат у випадку сегментації зображення є коефіцієнт «Dice». Цей коефіцієнт використовується для визначення ступеня перекриття між прогнозованою маскою та позначеною маскою. Коефіцієнт Dice має діапазон значень між [0,1], де 1 означає загальне перекриття. Тому, якщо ми хочемо максимально мінімізувати цю функцію, значення (1-коефіцієнт Дайс) можна використовувати як функцію втрат.

Коефіцієнт подібності DICE або КПД є основним показником перевірки індексу просторового перекриття [25]. КПД вимірює просторове перекриття між двома сегментаціями, цільовими областями A та B, і визначається як КПД (A,B)= 2(A∩B)/(A+B), де ∩ є перетином (рис. 2.3). У двійковій ручній сегментації цей коефіцієнт може бути отриманий із таблиці непередбачуваності ймовірностей класифікації сегментації 2 на 2 (таблиця 1). Концептуально КПД також є окремим випадком каппа-статистики, яка зазвичай використовується в аналізі надійності, коли існує набагато більша кількість фонових вокселів, ніж кількість цільових вокселів, як було показано раніше [26].

Крім того, оскільки DSC має обмежений діапазон [0,1] і часто близький до значення 1, виявилjcm корисним прийняти логіт-перетворення для цілей статистичних висновків, де logit(DSC)=ln{DSC /(1−DSC)}. Це монотонне перетворення відображає область [0,1] на необмежений діапазон (−∞,∞), а logit(0,5) = 0.

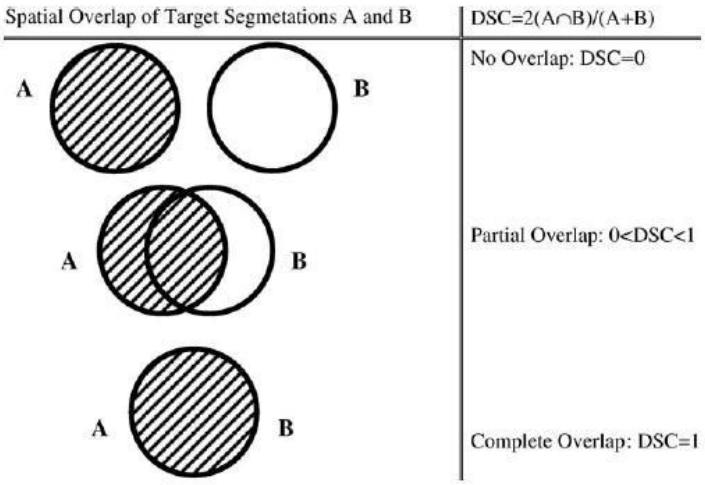


Рис.2.3. Пояснення к визначенню коефіцієнта подібності DICE [25]

Agresti [27] показав, що для великого розміру вибірки (у нашому випадку кількість вокселів), розподіл логіт-перетворених пропорцій і швидкості слідував нормальному розподілу, асимптотично з великою кількістю вокселів. Статистичний тест нормальності може бути проведений за допомогою z-тесту за припущенням незалежності від вокселів. Як рекомендує [26], в літературі з перевірки зображень гарне перекриття відбувається, коли DSC >0,700 або, еквівалентно, logit(DSC) >0,847. Виходячи з аналізу каппа-статистики, чудова згода відбувається, коли k >0,75, як рекомендує [29].

Висновок до розділу 2

У розділі було докладно розглянуто основні концепції, такі як нейронні мережі та комп'ютерний зір, разом із їхнім практичним використанням та принципами функціонування. Були проведені дослідження щодо можливостей та завдань, які можна вирішити за допомогою згорткових нейронних мереж. Також проведений аналіз алгоритмів сегментації зображень з метою їхнього подальшого використання у дипломній роботі. Детально вивчено архітектуру згорткової нейронної мережі U-Net, а також проаналізовано поняття коефіцієнта Dice.

3. ЗАСТОСУВАННЯ РОЗРОБЛЕНОГО АЛГОРИТМУ

3.1. Дослідження використання однорівневої нейронної мережі

Хвороби хребта поширені по всьому світу. За даними ВООЗ, остеопороз посідає 4-е місце серед причин інвалідності та смерті [23]. Існує велика кількість хвороб хребта, які можуть спричинити значний біль та вести до різноманітних симптомів. Біль у спині - це найпоширеніша скарга, яку лікарі чують. Це може бути різні травми, перевантаження м'язів, неправильна постава при сидінні тощо. Приблизно 80 відсотків громадян Великої Британії зіткнулися з проблемами хребта протягом свого життя [24]. Остеохондроз, вибухнуті міжхребцеві диски, спонділоз, сколіоз, спондилолістез - і це лише невелика частина хвороб хребта. З такою великою кількістю хвороб хребта на допомогу приходить машинне навчання. Для вирішення медичних проблем використовуються алгоритми штучного інтелекту, які здатні обробляти великі обсяги даних швидше та точніше. Основні важливі застосування машинного навчання в медицині можна визначити: діагностика захворювань, прогнозування захворювань, аналіз медичних зображень, управління медичними даними, медична робототехніка. Цей перелік демонструє, що використання машинного навчання в медичній галузі має великі перспективи і є неот'ємною частиною прогресу. Завдяки цьому ми можемо покращити діагностику, обробку та управління медичними даними. Однак, з іншого боку, слід розуміти, що використання машинного навчання в медицині повинно мати високий рівень глибоких досліджень, перевірки та стандартизації.

3.1.1. Машинне навчання

Машинне навчання є галуззю штучного інтелекту, яка фокусується на використанні даних та алгоритмів для моделювання способу вчення людини та поступового покращення його точності [25]. Машинне навчання подолало довгий шлях від комп'ютерного шаху до алгоритму рекомендацій

YouTube та автомобілів Tesla без водія. Прогрес у сфері обчислювальних технологій зробив такі інноваційні проекти можливими. Машинне навчання використовується в багатьох галузях, таких як: комп'ютерне зорове сприйняття, маркетинг, реклама, робототехніка, обробка мови, виробництво, автономні транспортні засоби, медицина, фінансовий сектор, військовий сектор. Основні алгоритми машинного навчання включають наступне:

- Лінійна регресія - це алгоритм, який використовується в навчанні з учителем і моделює відносини між скалярною та векторною змінною. Цей алгоритм може бути використаний для прогнозування цін на автомобілі за рік виробництва. Загальна модель має формулу:

$Y=\beta_0+ \beta_1 x+ \varepsilon$, де β_0 і β_1 - невідомі константи, Y - залежна змінна, x - незалежна змінна, а ε - помилка.

- Логістична регресія - це алгоритм, який використовується в навчанні з учителем і застосовується, коли залежна змінна є бінарною, тобто має дві варіанти значень. Широко використовується у випадках, які вимагають відповіді - правда чи брехня.

- Кластеризація - це алгоритм навчання без учителя, в якому точки даних розділяються на кілька груп. Так, щоб точки даних в одній групі були схожі між собою, а водночас відрізнялись від точок інших груп.

- Дерева рішень - це алгоритм навчання з учителем, який використовується для прогнозування числових значень, а також їх категоризації. Дерево складається з трьох елементів: внутрішнього вузла - це ознака набору даних, гілки - це метод прийняття рішень, і листка - це результат.

- Випадковий ліс відноситься до сімейства контрольованих алгоритмів. Використовується для побудови кількох дерев рішень і комбінує результат з кількох дерев.

3. 2. Використання ансамблю НМ

Для дослідження було обрано вісім різних моделей: Fcn8, Fcn32, MobileNet_Segnet, MobileNet_Unet, Segnet, Unet, Vgg_Segnet, Vgg_Unet. Ці мережі використовуються для різних цілей, хоча деякі були спеціально розроблені для використання в медицині, наприклад, Unet [26]. Вибір нейромереж здійснювався відповідно до конкретної задачі та обсягу даних. Для медичних застосувань, де найважливішою є точність, безпека та надійність, використовуватимуться різні комбінації моделей для досягнення найвищої точності. Тому для цього краще підходять мережі з архітектурою, що базується на структурі згорткового енкодера-декодера. Вони здатні вирішувати проблему "вибування градієнту", яка виникає в складних задачах сегментації. Кожна з використовуваних моделей має свої переваги та недоліки, тому будуть використовуватися не лише стандартні архітектури, але й модифікації та комбінації цих архітектур.

3.2.1. Збір та підготовка даних

Основною частиною усього процесу навчання є дані, на яких модель буде подальше тренуватися. Кінцевий результат і успіх будуть залежати від їх якості та кількості. На першому етапі збирають рентгенівські зображення. На другому етапі потрібно видалити дублікати, привести всі зображення до одного розміру, у деяких випадках буде необхідно обрізати або заповнити зображення. На третьому етапі відбувається кодування маски, де кожному пікселю хребця присвоюється значення 1, а кожному пікселю фону - значення 0. На останньому етапі дані розділяються на дві групи - навчальний зразок та тестовий зразок. Для тренування всіх моделей використовувався набір з 183 зображень розміром 256x256 пікселів. Усі зображення були в градаціях сірого. Розширення даних не відтворювалося. Для кожного хребця було створено чотири бази даних: Th8, Th9, Th10, Th11.

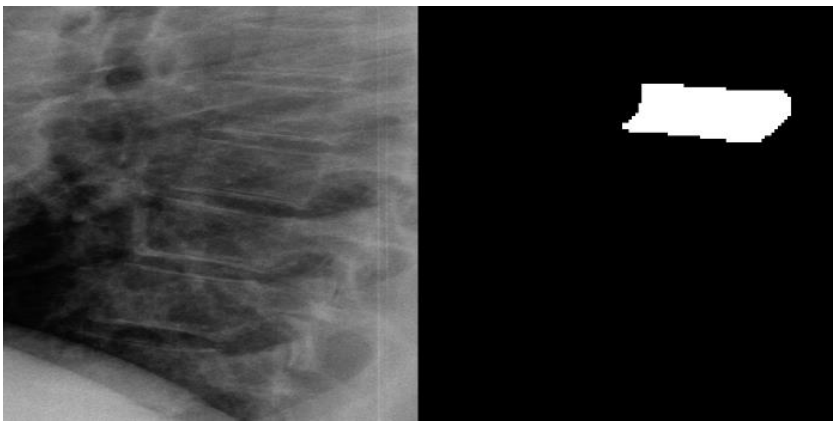


Рис.3. 1 Приклад рентгенівського зображення та його маски

3.2.2. Порівняння нейромереж

Після успішної підготовки даних прийшов час для тренування моделей. Всі мережі навчалися за однаковими налаштуваннями, після кожної ітерації тестові дані валідувалися для запобігання перенавчанню. Процес навчання був розділений на чотири етапи: спочатку тренування проводилося на хребці Th8, потім на Th9, Th10 і Th11. На кожному з чотирьох етапів було виконано 10 тренувань кожною моделлю для кожного хребця, після чого обирався та табулювався кращий результат в кожній моделі. Як можна бачити з результатів у таблиці, лідером є MobileNet_Unet, він перемагає у сегментації 8, 9, 11 хребців. І модифікація Vgg_Unet також перемагає у сегментації 10 хребця. Для визначення точності використовувалися наступні метрики - FW IoU (Частотно зважена Інтерсекція над Об'єднанням), Mean IoU (Середня Інтерсекція над Об'єднанням), IoU (Інтерсекція над Об'єднанням).

Таблиця 1. Передбачення форми і розміру об’єкту

	Fcn8	Fcn32	MobileNet _Segnet	MobileNet _Unet	Segnet	Unet	Vgg_Segnet	Vgg_Unet
FW IoU	99.18	98.96	99.36	99.55	97.95	99.19	98.61	99.51
Mean IoU	95.69	94.58	96.68	97.64	89.12	95.72	92.75	97.41
IoU	91.82	89.72	93.70	95.51	79.32	91.87	86.24	95.09

Формула, що використовувалася для вимірювання точності:

$$\mathrm{IoU} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

де TP - правильні позитивні результати, TN - правильні негативні результати, FP - помилкові позитивні результати і FN - помилкові негативні результати.

Згідно з результатами навчання, мережі, такі як Vgg_Unet та MobileNet_Unet, показали кращі результати в порівнянні з іншими. Ми пропонуємо розглянути їх докладніше. Ці мережі є модифікаціями мережі U-Net, яка була спеціально розроблена для медичних досліджень.

3.2.3. Аналіз мережі U-Net

U-Net - це згорткова нейромережа, яка була спеціально створена для сегментації біомедичних зображень на факультеті комп'ютерних наук Університету Фрайбурга. Однією з важливих змін у цій мережі є те, що в частині розширення ми також маємо велику кількість функціональних каналів, що дозволяють мережі розподіляти контекстуальну інформацію на вищому рівні роздільної здатності. У результаті розширюючий шлях більш-менш симетричний шляху звуження і формує архітектуру у формі букви U. Архітектура цієї моделі використовує структуру кодувальника та декодувальника з пропуском з'єднань. Її шари кодування та декодування симетричні один одному.

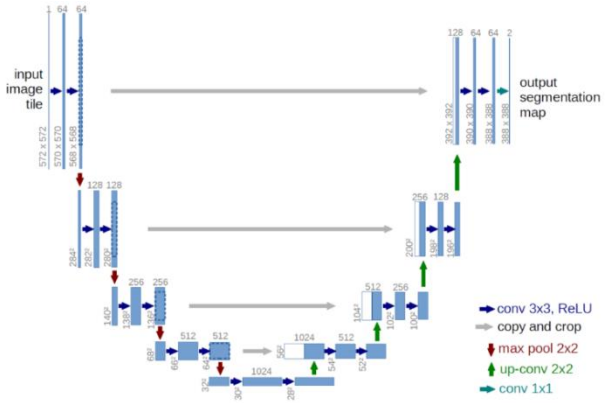


Рис.3. 2 Архітектура *U-Net*

Енергетична функція розраховується за допомогою нормалізованої експоненційної функції над кінцевими картами властивостей у поєднанні з функцією втрати кросс-ентропії [26]. Нормалізована експоненційна функція визначається як:

$$p_k(x) = \exp\big(a_k(x)\big) / \left(\sum_{k'=1}^{K} \exp\big(a_{k'}(x)\big)\right) \tag{2}$$

де a_k(x) позначає активацію в каналі функції k на позначці пікселя x ∈ Ω з Ω ⊂ Z^2.

K - це кількість класів, а p_k(x) - наближена функція максимуму.

Потім крос-ентропія покарає за кожне відхилення позначення p_l(x) від 1, використовуючи:

$$E = \sum_{x \in \Omega} w(x) \log\big(p_{l(x)}(x)\big) \tag{3}$$

де l: Ω →{1,…,K} - це справжній позначення кожного пікселя, а w: Ω→R - це карта ваг, введена для надання деяким пікселям більшого ваги під час тренування.

де w_c : Ω →R - це карта ваг для балансування частот класу, d_1 : Ω →R позначає відстань до межі найближчої клітини, а d_2 : Ω →R - відстань до межі другої найближчої клітини.

3.2.4. Аналіз мережі Vgg 16

Vgg 16 - це згорткова нейромережа, яка є вдосконаленою версією AlexNet. Під час навчання розмір вхідних даних становить 224x224 пікселів RGB. Єдина операція, яку виконують автори цієї мережі, - це віднімання середнього значення RGB, яке було розраховано на навчальному наборі, від кожного пікселя. Спочатку зображення проходить через стек згорткових шарів, де використовуються фільтри розміром 3x3. У одній з конфігурацій автори використовують фільтри розміром 1x1, які можна розглядати як лінійне перетворення вхідних каналів. Після цього крок згортки зафіксований на 1 піксель. Просторове заповнення вхідного конвеєра шару таке, що просторова роздільна здатність зберігається навіть після згортки. Просторове пулінг виконується за допомогою п'яти рівнів максимального пулінгу, які слідують за деякими шарами. Максимальний пулінг виконується над вікном 2x2 пікселя з кроком 2. З стеку згорткових шарів слідують три з'єднаних шари: перші два мають по 4096 каналів на шар, третій виконує класифікацію на 1000 каналів. Останній шар - це нормалізована експоненційна функція. Всі приховані шари оснащені ректифікацією [27].

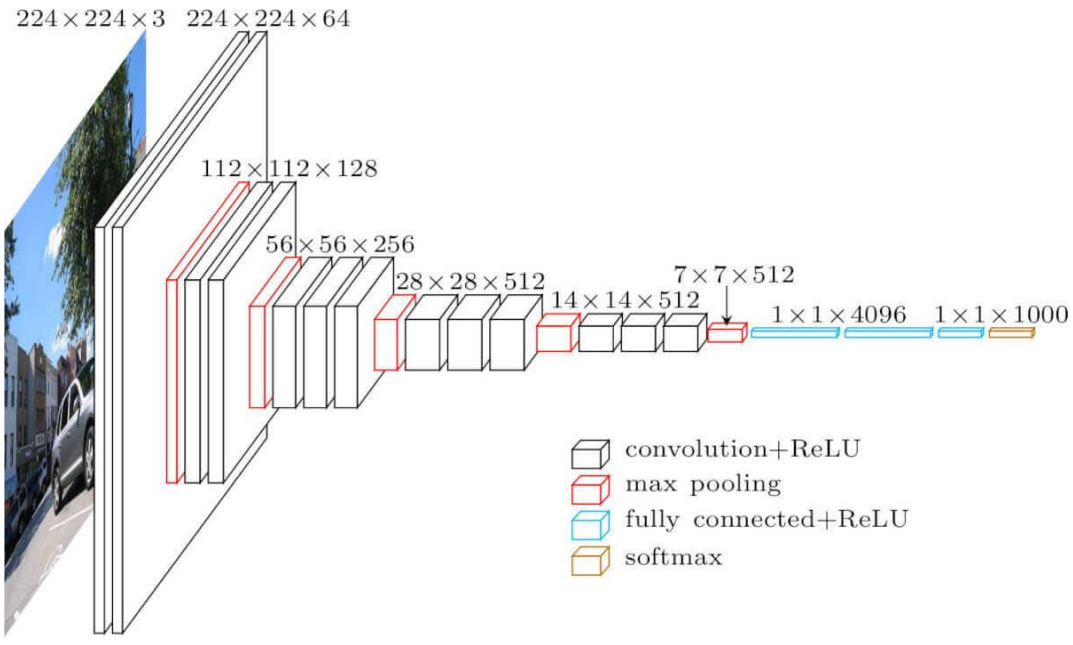


Рис. 3.3 Архітектура Vgg 16

3.2.5. Аналіз мережі MobileNet 16

Це ефективна модель для мобільних і вбудованих застосувань у сфері комп'ютерного зору. Вона ґрунтується на простій архітектурі та використовує глибокі згорткові шари для створення високоефективних мобільних нейромереж. Першою перевагою цієї моделі є висока швидкість та невелика кількість параметрів, що в першу чергу дозволяє використовувати її ефективно на пристроях з обмеженими обчислювальними ресурсами.

Ця архітектура ґрунтується на двох концепціях: глибоких згорткових шарах та параметрі, що дозволяє регулювати кількість каналів. У цій архітектурі модель конструюється таким чином, що вона використовує згорткові шари, які можуть бути розділені за глибиною. Це само по собі є конструкцією факторизованих згорток, яка розкладає звичайну згортку на згортку за глибиною та згортку 1x1, яку називають точковою згорткою. Кожна згортка за глибиною застосовує один фільтр до кожного вхідного каналу. Після цього точкова згортка повинна застосувати згортку 1x1, щоб подальше об'єднати виходи глибокої згортки. Звичайна згортка повинна фільтрувати та об'єднувати входи в новий набір виходів за один крок. З свого боку, згортка за глибиною розбиває її на два шари, де перший - це шар для фільтрації, а другий - шар для об'єднання [28].

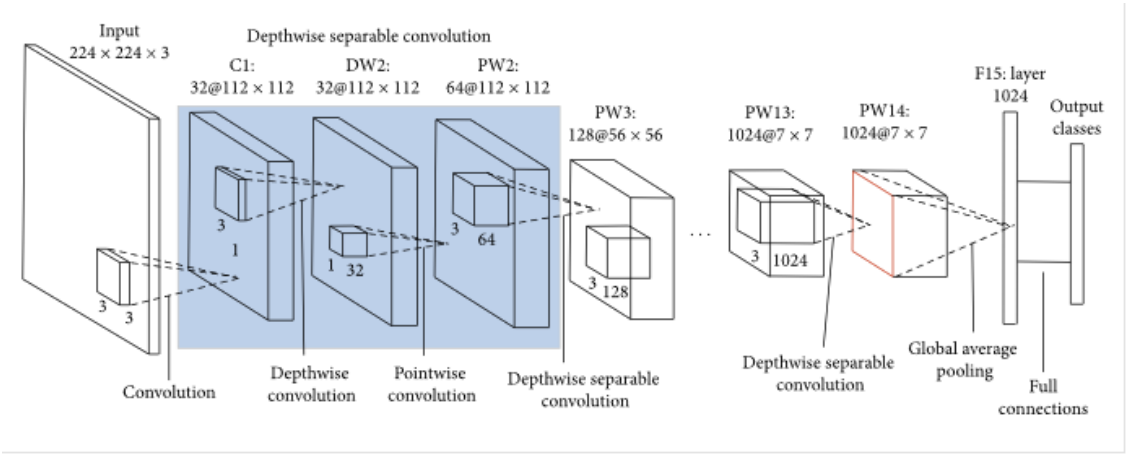


Рис. 3.4 Архітектура MobileNet

3.3. Реалізація основної ідеї і обговорення результатів

Проведений аналіз дозволив втілити основну ідею роботи. Ця ідея складається в використанні не однієї мережи, а цілого ансамблю мереж. Після тренування кожної з мереж на однаковому наборі даних зі всієї множини вибирається певна кількість, наприклад, чотири з восьми, які дають найкращий результат. Після цього тільки ці чотири мережи використовуються при роботі з реальними зображеннями. При цьому вони використовуються для визначення об’єктів незалежно і паралельно. Після знаходження об’єкта кожною з цих мереж проводиться усереднення знайденого зображення по цьому набору. Це усереднення проводилось таким чином. Для кожного пикселя визначалась кількість разів коли він належав до шуканого зображення на передбаченнях кожної з мереж. Потім це число порівнювалось з певною, наперед заданою границею і якщо воно перебільшувало цю границю, то цей піксель позначався як той, що належить до шуканого об’єкту. В результаті такого підходу вдалося суттєво покращити якість передбачення місця знаходження об’єктів, їх форми та розмірів.

Остаточно, можна зробити висновок, що використання ансамблю мереж безумовно покращує якість розв’язання задачі, і дозволяє здобути результати, кращі, ніж в кожної окремої нейронної мережі.

Висновки до розділу

У розділі зосереджено увагу на актуальності використання машинного навчання для сегментації об’єктів медичних зображень. Розглянуті ключові етапи формування бази зображень для подальшого навчання нейромережі. Проведено дослідження архітектур мереж, які продемонстрували найкращі результати у сегментації. Здійснено експериментальне порівняння восьми різних мереж, що дозволило виділити оптимальні та проаналізувати їхні переваги та недоліки. Застосування машинного навчання дозволяє автоматично визначати та виділяти об’єкти на медичних зображеннях, забезпечуючи високу швидкість обробки та точність.

ВИСНОВКИ

У ході написання дипломної роботи було ретельно вивчено основні алгоритми використання нейронних мереж для сегментації зображень з великим рівнем шумів. Особлива увага приділялася методам визначення місця знаходження та розмірів певних об`єктів, зокрема рентгенівському морфометричному методу. Розроблено алгоритм, який дозволяє порівнювати результати, отримані з різних нейронних мереж і визначати найбільш достовірні результати.

Проведено дослідження використання двоетапної згорткової нейронної мережі для сегментації зображень з великим рівнем шумів. Також розроблено власну схему дворівневого алгоритму, що використовує нейронну мережу для визначення форми об`єктів. Подальше дослідження включало в себе використання двохрівневого алгоритму нейронної мережі для сегментації зображень. Всі ці етапи спрямовані на розширення розуміння та вдосконалення процесів визначення та лікування остеопорозу за допомогою передових технологій машинного навчання.

Застосування машинного навчання дозволяє автоматично визначати та виділяти задані певні об’єкти, забезпечуючи високу швидкість обробки та точність. Використання цієї технології, зокрема, в медицині сприяє автоматизації обробки великих обсягів даних. Важливо відзначити, що впровадження машинного навчання в обробці зображень вимагає додаткових досліджень, особливо в питаннях безпеки та точності.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Штучний інтелект для покращення якості фото. Https://evergreens.com.ua/ua/articles/image-enhancement-solutions.html

2. Глибоке навчання. https://uk.wikipedia.org/wiki/Глибоке_навчання#:~:text=image%20reconstruction)%20%20це%20відбудова%20зображень,наприклад%2C%20спектральних%20та%20ультразвукових%20зображень.

3. Глибоке навчання проти нейронних мереж. Https://www.unite.ai/uk/deep-learning-vs-neural-networks/

4. 10 найкращих інструментів штучного інтелекту для покращення зображення та масштабування (вересень 2023 р.). Https://www.unite.ai/uk/best-ai-image-upscaler-tools/

5. Нейромережа — що це таке, як працює та навіщо потрібна. **https://termin.in.ua/neyromerezha/#Gliboke_navcanna_Viznacenna_glibokogo_navcanna_ta_jogo_zvazok_z_nejromerezami**

6. Технології, що змінюють правила гри. Https://www.railway.supply/uk/tehnologiї-shho-zminyuyut-pravila-gri/

7. Глибоке навчання. Https://www.wikiwand.com/uk/глибинне_навчання#розпізнавання_зображень

8. Тренди глибокого навчання і нейронних мереж. Https://merehead.com/ua/blog/deep-learning-trends-in-2024/

9. Нейромережа — що це таке, як працює та навіщо потрібна. Https://termin.in.ua/neyromerezha/#gliboke_navcanna_viznacenna_glibokogo_navcanna_ta_jogo_zvazok_z_nejromerezami

10. Google розробила нейромережевий алгоритм, який здатен суттєво покращувати якість нічних знімків. Https://itc.ua/ua/novini/google-rozrobila-nejromerezhevij-algoritm-yakij-zdaten-suttyevo-pokrashhuvati-yakist-nichnih-znimkiv/

11. Штучна нейронна мережа. Https://uk.wikipedia.org/wiki/штучна_нейронна_мережа

12. Нейромережа — що це таке, як працює та навіщо потрібна. Https://termin.in.ua/neyromerezha/

13. The basics of neural networks. URL: https://towardsdatascience.com/the-basics-of-neural-networks-neural-network-series-part-1-4419e343b2b

14. Neurons in neutral networks. URL: https://www.baeldung.com/cs/neural-networks-neurons

15. History of perceptron. URL: https://home.csulb.edu/~cwallis/artificialn/History.htm

16. Activation functions in neutral network. URL: https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6

17. Deep learning: overview of neurons and activation functions. URL: https://srnghn.medium.com/deep-learning-overview-of-neurons-and-activation-functions-1d98286cf1e4

18. What is Unsupervised Learning. URL: https://www.ibm.com/topics/unsupervised-learning#:~:text=the%20next%20step-,What%20is%20unsupervised%20learning%3F,the%20need%20for%20human%20intervention

19. Introduction to Convolutional Neutral Networks. URL: https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/

20. CNN Basic architecture for classification & segmentation. URL: https://vitalflux.com/cnn-basic-architecture-for-classification-segmentation/#Basic_CNN_architecture_for_Segmentation

21. An Introduction to Image Segmentation: deep learning vs. Traditional. URL: https://www.v7labs.com/blog/image-segmentation-guide#h1

22. Image Segmentation algorithms with implementation in Python. URL: https://www.analyticsvidhya.com/blog/2021/09/image-segmentation-algorithms-with-implementation-in-python/

23. Міжнародний день боротьби з остеопорозом: разом до перемоги над хворобою! // Новини медицини та формації. — 2010. — № 18. — 6С.

24. Living Room Health Regenerative Treatments, (2019). 80% UK adults experience back pain [онлайн]. Living Room Health Regenerative Treatments. [Дата

звернення 02 серпня 2023]. Режим доступу: https://www.theregenerativeclinic.co.uk/news/80-uk-adults-experience-back-pain/

25. IBM, (2023). What is machine learning? [онлайн]. IBM. [Дата звернення 02 серпня 2023]. Режим доступу: https://www.ibm.com/topics/machine-learning

26. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. (2015). “U-Net: Convolutional Networks for Biomedical Image Segmentation”. arXiv:1505.04597

27. Karen Simonyan, Andrew Zisserman. (2015). “Very Deep Convolutional Networks for Large-Scale Image Recognition”. arXiv:1409.1556

28. Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam. (2017). “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications”. arXiv:1704.04861