

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Інформаційних радіотехнологій і технічного захисту інформації
(повна назва)

Кафедра Радіотехнологій інформаційно-комунікаційних систем
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

ГЮІК.XXXXXX.001ПЗ

(позначення документа)

АНАЛІЗ МЕДИЧНИХ ЗОБРАЖЕНЬ ДЛЯ ДІАГНОСТУВАННЯ ХВОРОБИ
(тема)

Виконав:

студент II курсу, групи ІКТм-20-1
Коваль Б.В.
(прізвище, ініціали)

Спеціальність 122 Комп'ютерні науки

(код і повна назва спеціальності)

Освітня програма Інформаційно-комунікаційні технології
(повна назва освітньої програми)

Керівник проф. Кузьомін О.Я.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри

(підпис)

О.І. Цопа
(прізвище, ініціали)

2021 р.

Не містить відомостей заборонених для відкритого публікування.

Керівник _____ Кузьомін О.Я.

Студент _____ Коваль Б.В.

Харківський національний університет радіоелектроніки

Факультет Інформаційних радіотехнологій та технічного захисту інформацій

Кафедра Радіотехнологій інформаційно-комунікаційних систем

Рівень вищої освіти другий (магістерський)

Спеціальність 122 Комп'ютерні науки

Тип програми Освітньо-професійна

Освітня програма Інформаційно-комунікаційні технології

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

«_____» _____ 2021 р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

студентові КОВАЛЮ Богдану Віталійовичу

1. Тема роботи Аналіз медичних зображень для діагностування хвороби затверджена наказом по університету від 5 листопада 2021 р. № 1648Ст
2. Термін подання студентом проекту (роботи) 5 грудня 2021 р.
3. Вихідні дані до проекту (роботи) розглянути сучасні системи для діагностування хвороби та розробити нейронну мережу для діагностування хвороби за допомогою аналізу медичних знімків, нейронна мережа має сегментувати уражені ділянки органів та правильно класифікувати їх.
4. Перелік питань що потрібно опрацювати в роботі
Вступ. 4.1 Аналіз предметної області. 4.2 Аналіз і обґрунтування вибору технологій і мов програмування. 4.3 Дослідження використання нейронних мереж у медицині. 4.4 Розробка нейронної мережі. Висновки. Перелік посилань. Додатки
5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) загальні схеми нейронних мереж, приклади сегментації на зображеннях, комп'ютерна презентація проекту

6. Консультанти розділів роботи

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата
Основна частина	проф. Кузьомін О.Я.		

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Ознайомлення із завданням. Уточнення ТЗ	01.09.21	Виконано
2	Підбір літератури за темою роботи	13.09-22.09.21	Виконано
3	Теоретичний розділ	23.09-10.10.21	Виконано
4	Проектний розділ	11.10-25.10.21	Виконано
5	Оформлення презентаційного матеріалу,	26.10-30.11.21	Виконано
6	Підготовка до захисту у ЕК	1.12-11.12.21	Виконано
7	Подання роботи на кафедру	13.12.21	Виконано

Дата видачі завдання 1 вересня 2021 р.

Студент _____
(підпис)

_____ Коваль Б.В.
(прізвище та ініціали)

Керівник роботи _____
(підпис)

_____ проф. Кузьомін О.Я.
(посада, прізвище та ініціали)

РЕФЕРАТ

Пояснювальна записка: 71 сторінок, 22 рисунка, 1 таблиця, 2 формули, 2 додатки.

ШТУЧНИЙ ІНТЕЛЕКТ, НЕЙРОННА МЕРЕЖА, ГЛИБОКЕ НАВЧАННЯ, МАШИННЕ НАВЧАННЯ, ШАР, МЕДИЧНІ ЗОБРАЖЕННЯ, ЗГОРТКОВІ НЕЙРОННІ МЕРЕЖІ, ДІАГНОЗ, ХВОРОБА, РОЗМІТКА ДАНИХ.

Предмет дослідження: вибір методів аналізу та технологій.

Об'єкт дослідження: медичні зображення.

Методи дослідження:

Системний аналіз – для виділення окремих блоків для роботи з розробкою системи та дослідження алгоритмів розробки.

Аналогія – для дослідження роботи та визначення недоліків подібних систем прогнозування.

Емпіричний метод – ознайомлення із експертними оцінками, даними провідних вчених та дослідниками теорій прогнозування, статистики, аналізу часових рядів та алгоритму його створення.

Систематизація – для узагальнення прикладів розробки систем прогнозування.

У першій частині було розглянуто сучасне місце нейронних мереж у мереж у медицині, галузі у яких домоглися значних результатів за допомогою використання нейронних мереж.

У другій частині описано обґрунтування технічних засобів та розробку нейронної мережі для діагностування хвороби.

ABSTRACT

Explanatory note: 71 pages, 22 figures, 1 table, 2 formulas, 2 appendices.

ARTIFICIAL INTELLIGENCE, NEURAL NETWORK, DEEP LEARNING, MACHINE LEARNING, LAYER, MEDICAL IMAGES, ROLLED NEURAL NETWORKS, DIAGNOSIS, DIAGNOSIS, DATA MARKING.

Subject of research: the choice of methods of analysis and technology.

Object of research: medical images.

Research methods:

System analysis - to select individual blocks to work with system development and research development algorithms.

Analogy - to study the work and identify the shortcomings of such forecasting systems.

Empirical method - acquaintance with expert assessments, data of leading scientists and researchers of theories of forecasting, statistics, time series analysis and the algorithm of its creation.

Systematization - to generalize examples of forecasting systems development.

The first part examines the current place of neural networks in medical networks, industries in which significant results have been achieved through the use of neural networks.

The second part describes the rationale for technical means and the development of a neural network for diagnosing the disease.

ЗМІСТ

ВСТУП	6
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ	7
1.1 Огляд сучасного стану проблеми	7
1.2. Застосування медичної візуалізації	9
1.3 Підтримка прийняття рішень у невідкладній медичній допомозі	13
1.4 Загальний вигляд систем комп'ютерної обробки медичних зображень для постановки діагнозу	14
1.5 Огляд існуючих систем аналізу медичних зображень	16
1.6 Проблема розмітки медичних даних та їх кількість	17
2 АНАЛІЗ І ОБГРУНТУВАННЯ ВИБОРУ ТЕХНОЛОГІЙ РОЗРОБКИ ТА МОВ ПРОГРАМУВАННЯ	19
2.1 Огляд технічних та програмних засобів	19
2.2 Нейронна мережа	22
2.3 Огляд бібліотеки Tensorflow	29
2.4 Метод зворотного розповсюдження помилки	30
2.5 Постановка задачі	43
3 РОЗРОБКА НЕЙРОННОЇ МЕРЕЖІ	45
3.1 Сегментація	45
3.1.1 Оптимізація моделей згорнутих нейронних мереж	50
3.2 Класифікація та кластеризація	51
3.3 Ансамблювання нейронних мереж	53
ВИСНОВКИ	57
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	58
ДОДАТОК А	Ошибка! Закладка не определена.
ДОДАТОК Б	Ошибка! Закладка не определена.

ВСТУП

Штучний інтелект (ШІ) та машинне навчання успішно застосовуються в медицині і вирішують широке коло завдань, поступово перетворюючись із допоміжного інструменту на хороших помічників медичного персоналу. В основі роботи ШІ лежить аналіз медичних даних та їх обробка по заданим алгоритмам. В даний час аналізуються не лише дані об'єктивного огляду та анамнезу пацієнта, а й результати аналізів та обстежень на медичному обладнанні. Застосування подібних інструментів підвищує ефективність лікаря, позбавляючи його виконання низки рутинних операцій, таких як ведення частини медичної документації та опис норми під час проведення обстежень.

Одна з важливих проблем застосування ШІ в медицині – підготовка коректних медичних даних для навчання алгоритмів, тому що для цього потрібна велика кількість часу спеціалістів вузького профілю. Можливим рішенням є створення об'єднаної платформи зберігання медичних даних, де лікарі зможуть готувати дані для застосування ШІ у своїй спеціальності. Це дозволить у майбутньому підвищити ефективність застосування машинного навчання в медицині завдяки аналізу різнопланових даних різних джерел.

Потенціал цифрових технологій, штучного інтелекту, медичних баз даних, нових математичних методів обробки та аналізу великих даних починає все більш активно використовуватися в медичній практиці: діагностування онкологічних захворювань на основі комп'ютерного аналізу знімків, навчання та тренінг у хірургії, діагностування захворювань очей із цифрового зображення сітківки в офтальмології та ін.

Метою даної роботи є розробка нейронних мереж для аналізу медичних зображень, що допоможе швидше та точніше встановлювати діагноз.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Огляд сучасного стану проблеми

В даний час штучний інтелект (ШІ) та машинне навчання застосовуються в різних галузях медицини та дозволяють вирішувати широке коло завдань: від визначення патології на рентгенологічних знімках до постановки діагнозу та складання плану лікування на підставі даних з історії хвороби пацієнта. У цій роботі розглянемо деякі шляхи застосування ШІ до аналізу різних типів медичних даних та основні можливості, які подібні рішення надають лікарю та пацієнту.

Радіологія була одним із перших напрямів у медицині, де ШІ почали застосовувати для аналізу зображень. Лікарі-рентгенологи отримали можливість обробляти зображення в автоматичному режимі, відфільтровувати норму та робити висновки лише для знімків, помічених програмою як патологічні. Пізніше, завдяки накопиченню нових даних, системи ШІ навчилися класифікувати знімки з різних патологій, визначаючи той чи інший стан із високою ймовірністю.

Сьогодні відбувається перехід від ретроспективного аналізу до прогностичної моделі – алгоритми навчаються визначати ризик розвитку таких патологій, як інсульт чи поширення метастазів при новоутвореннях голови та шиї, до їх фактичної появи. У перспективі це може сприяти наданню допомоги на ранньому етапі захворювання та зменшенню потенційного ризику для здоров'я.

Штучний інтелект активно застосовується у лабораторній діагностиці, мікроскопії, мікробіології. За допомогою фото та відеофіксації можна використовувати алгоритми аналізу зображень та машинного навчання для автоматичної ідентифікації тканин, сегментування окремих клітин, визначення та вимірювання патологічних вогнищ безпосередньо в процесі дослідження. Подібні

результати знайшли широке застосування в таких галузях медицини, як онкологія, кардіологія, вносячи свій внесок у підвищення якості діагностики та лікування. Сучасні системи аналізу зображень здатні генерувати текстові описи результатів своєї роботи, що позбавляє фахівця необхідності писати висновок вручну.

Перспективним напрямом для досліджень та розробок у галузі аналізу зображень є аналіз зображень низької якості та відсканованих з паперових носіїв. Інформація, що зберігається на паперових носіях (амбулаторні карти, історії хвороби, роздруківки досліджень), після оцифрування становитиме повнішу картину стану здоров'я пацієнта. З технологічної точки зору подібні розробки дозволять більш точно навчати існуючі системи ІІ.

Крім аналізу зображень в медицині також застосовується обробка природної мови (англ. natural language processing, NLP) – наукова галузь на стику інформатики, ІІІ та лінгвістики, що вивчає проблеми комп'ютерного аналізу та синтезу текстів та мовлення природними мовами. У медицині NLP застосовується у кількох напрямках. Одне з них – семантичний аналіз медичних текстів та інформаційний пошук, коли система аналізує записи лікаря та категоризує їх за заданими параметрами. Це дозволяє машині правильно фільтрувати різні формулювання одного і того ж діагнозу у різних фахівців. Інший напрямок пов'язаний із вилученням інформації, тобто обробкою великого обсягу текстових даних. Зокрема, під час аналізу результатів клінічних досліджень алгоритм виділяє у тексті терміни і теми, і навіть формує короткий звіт.

В даній роботі буде розглянуто обробку медичних зображень за допомогою штучного інтелекту[1].

1.2. Застосування медичної візуалізації

Ринок медичної візуалізації розширюється. За даними Zion Market Research, його розмір становив близько 34 мільярдів доларів у 2018 році і, як очікується, досягне приблизно 48,6 мільярдів доларів до 2025 року. Медичні зображення становлять основне джерело даних у сфері охорони здоров'я, що становить щонайменше 90% усіх медичних даних згідно з GE Healthcare. Ця сума є величезною для ручного огляду та діагностики — радіологам і патологоанатомам доводиться щодня переглядати тисячі зображень за різними методами. Величезний потік різноманітних зображень, що суперечать застарілим процесам ручного перегляду, збільшує ймовірність медичних помилок і помилкового діагнозу.

Безумовно, ті, хто займається розробкою програмного забезпечення для охорони здоров'я, усвідомлюють цю проблему і надають індустрії автоматизовані методи аналізу медичних зображень на основі комп'ютерного зору. Комп'ютерний зір охоплює кілька методів отримання, обробки та аналізу джерел зображення та відео для отримання певних рішень щодо об'єктів у них. У сфері охорони здоров'я комп'ютерний зір може доповнити рутинну діагностику та оптимізувати робочі процеси рентгенологів і патологоанатомів.

Застосування комп'ютерного зору в охороні здоров'я. З експоненційним зростанням продуктивності обладнання комп'ютерний зір поступово стає поширеним інструментом підтримки прийняття рішень у сфері охорони здоров'я.

Комп'ютер «бачить» зображення не так, як людина. У той час як люди використовують контекстну інформацію, що зберігається в їхньому мозку, щоб ідентифікувати певні елементи, комп'ютери розглядають зображення як серію нулів і одиниць. Здатність машини міркувати про ці нулі та одиниці залежить від даних, на яких вона була навчена. Якщо навчання проходить успішно, комп'ютери, на відміну від людських очей, можуть вловити дуже дрібні деталі.

Типові набори даних медичних зображень, оброблені за допомогою комп'ютерного зору, отримують під час таких процедур:

Магнітно-резонансна томографія (МРТ) найкраще підходить для виявлення проблем у м'яких тканинах, таких як суглоби та кровоносна система.

Arterys спеціалізується на розшифровці зображень МРТ серця, виконанні різноманітних завдань, таких як аналіз перфузії міокарда, оцінка пізнього збільшення гадолінієм, виявлення контурів шлуночків і функції серця тощо. Компанія є піонером у візуалізації та розрахунку кровотоку. Їхній алгоритм аналізує зображення лише за 15 секунд, у порівнянні з 30 хвилинами для спеціаліста з людини.

Комп'ютерна томографія (КТ) в основному використовується для виявлення пухлин, внутрішніх мозкових кровотеч та інших небезпечних для життя станів. Zebra Medical Vision використовує аналіз зображень для ряду систем людського організму, а також пропонує постачальникам послуг автоматично розраховувати показники коронарного кальцію за допомогою базової КТ грудної клітки. Оскільки кальцій коронарних артерій є біомаркером ішемічної хвороби серця, його кількісна оцінка дозволяє прогнозувати серцеві напади та інсульти у пацієнтів з високим ризиком.

Ультразвукове дослідження (УЗД) - ця методика використовується для сканування органів, щоб перевірити їх правильне функціонування та наявність будь-яких аномалій. Це найменш інвазивна методика, і вона використовується для вагітних жінок для спостереження за розвитком плода.

Bay Labs створює алгоритми для інтерпретації ехокардіограм з особливим акцентом на застосуванні технології в країнах, що розвиваються, де мало висококваліфікованих лікарів. Система може аналізувати відео ультразвукового дослідження та виявляти ознаки серцево-судинних захворювань, таких як ревматична хвороба серця, що дозволяє навіть молодшим діагностам виконувати ехокардіограми та гарантувати правильний діагноз.

Рентгенівський знімок використовується для виявлення відхилень або пошкоджень органів. Комп'ютерний зір можна використовувати для класифікації сканувань так само, як людські рентгенологи, і він може визначити всі потенційні проблеми за один прийом. Це більш здоровий підхід, оскільки він обмежує вплив радіації.

Microsoft розробила інструмент під назвою InnerEye, який здатний ідентифікувати можливі аномалії, такі як пухлини, на рентгенівських знімках. Після того, як рентгенологи завантажують 3D-сканування в InnerEye, програмне забезпечення розпізнає потенційні пухлини та забарвлює підозрілі ділянки, щоб радіолог уважно оглянув.

Ядерна візуалізація використовується для візуалізації структури тканин і органів, а також функцій. Під час ядерної томографії виділяється дуже мала кількість радіоактивної речовини (радіонукліду). Радіонукліди поглинаються тканинами організму. Спостерігаючи за поведінкою радіонукліду в організмі, рентгенолог може оцінити різні стани, такі як пухлини та інфекції. За допомогою ядерної томографії пухлини можна виявити на ранніх стадіях.

Завдяки широкому охопленню, аналіз медичних зображень, досягнутий за допомогою комп'ютерного зору, дозволяє лікарям виявляти злоякісні зміни в організмі пацієнтів, відстежувати розвиток пухлин, оцінювати затвердіння артерій, а також вимірювати органи і кровоток з більшою точністю, ніж у спеціалістів.

Неінвазійна діагностика раку. Google, IBM, клінічні дослідники з університетів і понад 100 стартапів вкладають свій час і зусилля в використання комп'ютерного зору для діагностики раку лише за допомогою цифрових зображень, без інвазійної біопсії. Ці віртуальні біопсії замінюють інвазивні процедури за точністю, економічною ефективністю, комфортом пацієнта та часом отримання результату.

Оскільки алгоритми аналізу медичних зображень дуже чутливі до наявності ознак раку, машини можуть допомогти фахівцям у галузі охорони здо-

ров'я виявити хворобу на самих ранніх стадіях, щоб підвищити рівень виживаності та запропонувати пацієнтам найкращі шанси на безперерйне одужання.

Звичайно, алгоритми не є досконалими, і помилкові спрацьовування не є рідкістю. Тим не менш, радіологи та патологоанатоми є там, щоб повторно перевірити діагноз апарату та підтвердити чи відхилити його. Процес діагностики все ще знаходиться під контролем людини; він лише доповнюється чутливістю комп'ютерного зору, щоб забезпечити найкращий результат.

Рак легенів. Mindshare Medical розробила програмне забезпечення для комп'ютерної діагностики RevealAI-Lung для використання разом із КТ для швидшого та легшого виявлення раку легенів. Цей продукт був схвалений для продажу в Європі в липні 2019 року.

Окрім допомоги в діагностиці, RevealAI-Lung може запропонувати рекомендації щодо індивідуальних подальших обстежень. Він добре інтегрується з системами архівації зображень і комунікацій. Mindshare Medical продемонструвала здатність свого продукту зменшувати кількість помилкових результатів і час, необхідний для постановки діагнозу. Це може зменшити вплив на пацієнта радіоактивних речовин і додаткових біопсій.

Підвищення точної медицини. Precision medicine пропонує індивідуальне лікування відповідно до детальних профілів пацієнтів і на основі всіх доступних даних щодо здоров'я, навколишнього середовища та соціально-економічних даних. Відповідно, точна медицина – це вимоглива область, яка потребує значної технічної підтримки для обробки та аналізу величезних наборів даних.

Комп'ютерний зір входить до складу технологій точної медицини разом з аналітикою великих даних і штучним інтелектом, що дозволяє лікарям вилучати кількісні точки даних з кожного зображення в будь-якому способі.

1.3 Підтримка прийняття рішень у невідкладній медичній допомозі

MaxQ-AI створює набір діагностичних інструментів із 3D-зображеннями, специфічними даними для пацієнтів, глибоким баченням та когнітивною аналітикою в партнерстві з GE Healthcare, IBM і Samsung. Усі партнери зосереджуються на використанні даних у реальному часі в відділенні невідкладної допомоги для оцінки пацієнтів з підозрою на гостру травму голови чи інсульту та виявлення внутрішньочерепних крововиливів. У такому випадку пацієнти можуть отримати своєчасне лікування та уникнути довгострокових наслідків хронічного захворювання.

Інша система такого роду, AIDoc, побудована на основі технології глибокого навчання, яка виявляє аномалії високого рівня в медичних зображеннях у областях хребта, живота, голови та грудей. Зокрема, система може виявити гіпощільність кісток, вільну рідину в черевній порожнині, внутрішньочерепну гіперщільність, вільне повітря в грудній клітці тощо. AIDoc можна вбудувати в робочий процес лікаря за допомогою PACS і віджетів, розставляючи пріоритети випадків з підозрілими висновками.

Комп'ютерний зір покращує охорону здоров'я. Сфера діагностики знає збою з такими технологіями, як комп'ютерний зір. Існують певні суперечки, чи призведе це до нестачі робочих місць, чи покращить точність і допоможе радіологам швидше виконувати свою роботу.

Проте навіть сьогодні включення комп'ютерного зору в аналіз медичних зображень має безсумнівні переваги:

- Це покращує якість діагностики: в той час як діагности покладаються на свій досвід, і в деяких випадках неможливо уникнути людських суджень, алгоритми забезпечують вищу точність і можуть вловлювати деталі, які можуть вислизнути з очей людини.
- Це економить час, а отже, і життя: комп'ютерний зір може виявити небезпечні для життя стани на ранніх стадіях.

– Це зменшує витрати: неправильний діагноз може призвести до втрати тисяч доларів на неправильне лікування як для пацієнта, так і для медичної системи. Комп'ютерний зір має тенденцію працювати точно і підтверджує дуже вірогідний діагноз з самого початку, який має бути перевірений експертом

1.4 Загальний вигляд систем комп'ютерної обробки медичних зображень для постановки діагнозу

Сучасні інструментальні діагностичні системи комп'ютерна та магнітно-резонансна томографія (МРТ), електроенцефалографія та інші пов'язані з виробництвом великої кількості зображень та даних, аналіз яких виконується, як правило, вручну. Лікар переглядає їх, та на підставі свого досвіду та знань робить висновки та встановлює діагноз. Основне завдання комп'ютерних систем обробки зображень – автоматично виявляти відхилення від норми. Такі системи будуються переважно на штучних нейронних мережах.

Розроблено багато реалізацій нейронних мереж. Для вирішення простих завдань використовуються одношарові нейронні мережі. У таких мережах вхідні сигнали потрапляють на входи штучних нейронів шару, кожен з яких обробляє сигнали, що надійшли до нього, і видає результат. Сукупність цих результатів є виходом одношарової мережі. Каскади таких мереж утворюють багатошарові мережі. В даний час для розпізнавання зображень найчастіше застосовуються згорткові нейронні мережі (convolutional neural networks). Це багатошарові мережі, в яких чергуються шари згортки, що фільтрують вхідні сигнали, та шари пулінгу або об'єднання, що реалізують скорочення розмірності. Так, перший шар згортки послідовно, один за одним, обробляє всі фрагменти вихідного зображення величиною, наприклад, 3x3 пікселя і «згортає» ці фрагменти, формуючи результуючу ознаку, яка сприймається функцією активації,

що імітує реакцію нейрона на сигнал, що надійшов. Шар пулінгу аналізує фрагменти даних із попереднього шару розміром зазвичай 2×2 і передає на наступний шар середнє, максимальне або по-іншому обчислене значення. Параметри згортки налаштовуються у процесі машинного навчання (тренування) нейронної мережі.

Машинне навчання полягає в тому, щоб на основі машинного аналізу завантажених вихідних даних налаштувати параметри нейронної мережі таким чином, щоб надалі мережа з найбільшою ефективністю виконувала аналіз нових даних, що надходять на її вхід.

Існують два основні методи машинного навчання: навчання з учителем та самонавчання. У першому випадку для навчання використовується набір так званих розмічених даних, в якому кожен вихідний зразок містить інформацію про відповідність або невідповідність його нормі. Розмітка даних вимагає залучення експертів та пов'язана з великими витратами їхнього часу та праці. Самонавчання полягає у самостійному виявленні програмою закономірностей у вихідних даних. Проте, самонавчання на основі глибокого машинного навчання вимагає, щоб кількість об'єктів була дуже великою: сотні тисяч та мільйони. Тому більшість медичних систем аналізу зображень були навчені на розмічених даних.

Результатом роботи мережі є багатовимірний вектор ознак. Він подається на вхід класифікатора, який відповідно до свого алгоритму відносить його до тієї чи іншої категорії (їх може бути дві або більше). Якщо категорій лише дві, наприклад, «норма» та «патологія», то система віднесе аналізований знімок до однієї з них. Якщо комп'ютер навчений розрізняти кілька категорій, відповіддю буде диференційований діагноз. Отриманий у результаті обробки результат матиме імовірнісний характер: комп'ютер, зазвичай, видає відсоток своєї «впевненості» в результаті[2].

1.5 Огляд існуючих систем аналізу медичних зображень

У 2016 році була створена система PredictND, яка аналізує МРТ знімки для виявлення деменції. На 2016 рік точність системи була 69%, що більше ніж точність яку дали експерти – 51%. В оновленій системі PredictND обробляються не лише дані МРТ, а й результати нейропсихологічних тестів та біомаркери цереброспінальної рідини. Середня точність діагностики під час використання всіх джерел даних становила 82.3%. Система PredictND зараз тестується в декількох європейських клініках. Важливо також, що в ході тестування буде накопичено великий обсяг даних, які можуть бути використані надалі для покращення якості роботи системи[3].

Інші вчені запропонували восьмишарову мережу для виявлення хвороби Альцгеймера по МРТ. Були протестовані мережі з різними видами функцій активації та з різними методами пулінгу, серед них обрано мережу з оптимальною структурою. Були використані дані МРТ 98 пацієнтів із хворобою Альцгеймера та такої ж кількості здорових людей. Недостатність обсягу вихідних даних було подолано шляхом використання методу аугментації - створення навчання мережі додаткових даних із набору наявних зображень. З цією метою використовується поворот зображення, масштабування, зашумлення, зміна насиченості зображення та інші прийоми. У роботі обговорюються переваги та недоліки цього методу та показано, що аугментація може бути ефективно використана при аналізі зображень. Створена в результаті мережа із 8 шарів продемонструвала точність понад 97%, проте класифікація була обмежена двома класами.

Створення діагностичних систем засновано на використанні нових математичних методів обробки великих даних. Прикладом такої системи є побудована на суперкомп'ютері та оснащена штучним інтелектом система IBM Watson for Oncology, для навчання якої використовувався архів одного з провідних онкологічних центрів Memorial Sloan Kettering Cancer Center, що постійно поповнюється. Вкрай важливий як обсяг даних, та й їх якість. Ці дані

повинні бути досить повними та докладними та включати інформацію як про лікування, так і про подальше спостереження.

Для успішного аналізу та навчання нейронної мережі, дані які подаються на вхід, а саме зображення які ми аналізуємо, мають бути якісними, та великим плюсом буде якщо ці зображення будуть зроблені приладами, що видають однаковий формат зображень.

1.6 Проблема розмітки медичних даних та їх кількість

Для навчання алгоритмів необхідно, щоб дані, які вони аналізують, були правильно розмічені (наприклад, на рентгенівському знімку фахівець описує область затемнення та позначає її для програми як патологічний осередок). Наразі більшість медичних даних розмічаються лікарями у ручному режимі. Цей підхід має ряд недоліків: децентралізація – відсутність єдиної платформи, де лікарі можуть розмічати медичні дані; дорожнеча – висока вартість робочого часу кваліфікованого спеціаліста; відрив від роботи – фахівець під час розмітки даних не може займатися лікувальною діяльністю; необхідність достатньої кваліфікації – якісну розмітку даних може проводити лише фахівець вузького профілю.

Можливим вирішенням проблеми є створення зазначеної платформи, що збирає медичні дані з різних джерел і дозволяє медичним працівникам розмічати їх у форматі краудсорсингу (мобілізація ресурсів великої кількості людей з метою вирішення певних завдань).

Для навчання нейронних мереж потрібні або розмічені дані (навчання з учителем) або велика кількість даних. Недоліки розмічених даних вже перераховано. Якщо в нас їх нема то можна використати велику кількість медичних зображень для навчання нейронної мережі. Проблема в тому що, на даний час, кількість оцифрованих медичних даних (зокрема зображень) мала. Тому для

подальшого розвитку обробки медичних даних комп'ютерними системами потрібно щоб всі дані переводилися до цифрового вигляду з урахуванням формату. Наприклад всі медичні зображення зберігаються у форматі PNG, щоб системи розроблялися з використанням стандарту, що полегшить роботу розробникам. Та що робити коли даних мало а дослідження та навчання нейронних мереж вже робити потрібно? Для збільшення кількості зображень проводиться аугментація.

Аугментація даних це важливий етап навчання моделей машинного навчання. Під аугментацією даних розуміється збільшення вибірки даних навчання через модифікацію існуючих даних. Використання методів аугментації даних показало себе добре на завданні класифікації зображень. Враховуючи ресурси на розмітку зображень, для завдання розпізнавання об'єктів аугментація може бути більш кориснішою та значно швидшою.

Методи аугментації поділяються на операції з кольором (зміна кольорових каналів, збільшення контрасту, яскравості), геометричні операції (повернути зображення на певну кількість градусів, віддзеркалити горизонтально, вертикально) та операції з об'єктами на зображенні (віддзеркалити, повернути). Тож з кожного зображення ми можемо отримати ще мінімум 5, і наш датасет з 100 зображень (що не дуже багато для задачі) збільшиться до 500 зображень і точність натренованої нейронної мережі зросте. Чим більше даних ми подаємо на вхід нашої нейронної мережі тим більш точно вона буде ідентифікувати хворобу[4].

2 АНАЛІЗ І ОБГРУНТУВАННЯ ВИБОРУ ТЕХНОЛОГІЙ РОЗРОБКИ ТА МОВ ПРОГРАМУВАННЯ

2.1 Огляд технічних та програмних засобів

Для розробки нейронних мереж потрібно обрати мову програмування та середовище розробки. В даний момент більшість об'єктно – орієнтованих мов програмування мають бібліотеки для розробки нейронних мереж.

Велику швидкість роботи мають мережі написані на мові C++, тому що в ній можна керувати пам'яттю та більш ефективніше керувати ресурсами системи.

В даній роботі для розробки було використано мову програмування Python, тому що вона дозволить швидко розробити нейронну мережу для глибокого навчання.

Python - це інтерпретована мова високого рівня програмування для загального використання. Вона забезпечує конструкції, які дозволяють чітко програмування на малих і великих масштабах [5].

Python має динамічний тип системи та автоматичне керування пам'яттю. Він підтримує декілька парадигм програмування, включаючи об'єктно-орієнтовану, важливу, функціональну та процедурну, і має велику та всеосяжну стандартну бібліотеку [5].

NumPy - це фундаментальний пакет для наукових обчислень з Python. Вона містить, серед іншого:

1. Потужний об'єкт N-розмірного масиву.
2. Складні (трансляція) функції.
3. Інструменти для інтеграції C / C++ і Fortran коду.
4. Корисна лінійна алгебра, перетворення Фур'є і можливості випадкових чисел.

Для розробки нейронних мереж також можна застосувати Keras.

Keras - це високорівневий API нейронних мереж, написаний на Python і здатний працювати на вершині TensorFlow, CNTK або Theano (таблиця 2.1). Він був розроблений з урахуванням можливості швидкого експерименту, що дає можливість перейти від ідеї до результату з мінімальною затримкою.

Keras дозволяє:

- легко і швидко створювати прототипи (за допомогою зручності; модульності та розширюваності).
- підтримує як згорткові мережі, так і рекурентні, а також комбінації обох;
- працює без проблем на звичайному та графічному процесорах.

Keras сумісний з: Python 2.7-3.6.

Основна структура даних Keras - це модель, спосіб організації шарів. Найпростіший тип моделі – це sequential лінійний стек шару. Для більш складних архітектур, ви повинні використовувати функціональний API Keras, який дозволяє створювати довільні графи шарів. Нижче наведено sequential модель:

Таблиця 2.1 - Порівняння бібліотек глибинного навчання

с	Theano	Caffe	Torch	TensorFlow
Мова програмування	Python	Python, C++	Lua	Python
Спеціалізація на методах машинного навчання	Різноманітні види регресій, нейронні мережі	Нейронні мережі	-	Різноманітні види регресій, нейронні мережі
Можливість глибинного навчання	+	+	+	+
Швидкість роботи	Середня	Висока	Низька	Висока

Продовження таблиці 2.1.

Назва бібліотеки	Theano	Caffe	Torch	TensorFlow
Якість позпаралелювання між декількома пристроями	-	-	-	+
Допоміжні функції (автоматичне обрахування градієнта та ін.)	-	+ допоміжні бібліотеки Python	+ допоміжні бібліотеки Python	+ допоміжні бібліотеки Python

Для прискорення розробки машинного навчання буде використано бібліотеку Tensorflow, яка спеціально створена для побудови та тренування нейронних мереж, з метою автоматичного знаходження та класифікації образів, що досягає якості людського сприйняття. У цій бібліотеці вже є необхідні функції які можна використовувати для знаходження образів.

Тепер потрібно обрати технічні засоби для розробки нейронної мережі. Так як навчання нейронної мережі потребує великої обчислювальної потужності, то для таких задач потрібно використовувати потужні комп'ютери, з багатоядерним центральним процесором та потужним графічним процесором який має достатній об'єм відеопам'яті.

Одним з варіантів є придбання власного потужного комп'ютера з відповідними характеристиками. Для наших задач а саме розробка нейронних мереж глибокого навчання які будуть аналізувати зображення нам потрібно як можна потужніші графічні процесори. Для нашої задачі потрібно використати відеокарту RTX 3090. При нинішній ситуації на ринку її вартість сягає приблизно 4000 долларів. Звичайно вона в нас залишиться в подальшому, але зараз ця покупка не рентабельна.

Ще одним варіантом є оренда обчислювальної потужності на різних хмарних сервісах, наприклад Amazon Web Services. За допомогою обчислювальних потужностей, які надає користувачам цей сервіс, можна проводити глибоке навчання. Так як в нас кількість медичних зображень у цифровому вигляді зростає нам потрібно обробляти дуже багато зображень. У Amazon Web Services є тарифи які підходять під різні типи задач. Для нашої задачі є спеціальний тариф при якому можна обробляти до 1000000 зображень в місяць за 1000 доларів.

Сервіс від компанії Google – Colaboratory дозволяє розробляти нейронні мережі прямо у браузері та запускати навчання на потужностях компанії. Є безкоштовний тариф який дозволяє запускати свої розробки на потужностях компанії, але ті комп'ютери що виділяються не надто потужні. Якщо обрати максимальних тариф то за 50 доларів на місяці наш проект зможе виконуватися на потужних серверах, та навчання нейронної мережі буде проходити швидше. Мінусом такого сервісу є те що він доступний не у всіх країнах.

2.2 Нейронна мережа

Штучна нейронна мережа (ANN) – ефективна обчислювальна система, центральна тема якої запозичена з аналогії біологічних нейронних мереж. ANN також називаються "штучними нейронними системами" або "паралельними розподіленими системами обробки" або "системами з'єднання". ANN набуває великого набору блоків, які пов'язані за деяким шаблоном для забезпечення зв'язку між блоками. Ці одиниці, також звані вузлами чи нейронами, є простими процесорами, які працюють паралельно.

Кожен нейрон пов'язані з іншим нейроном через канал зв'язку. Кожне посилення на з'єднання пов'язане з вагою, яка має інформацію про вхідний сигнал. Це найбільш корисна інформація для нейронів для вирішення конкретної

проблеми, тому що вага зазвичай збуджує або пригнічує сигнал, що передається. Кожен нейрон має внутрішній стан, який називається сигналом активації. Вихідні сигнали, що виробляються після об'єднання вхідних сигналів та правила активації, можуть надсилатися на інші пристрої.

Загалом структура штучної нейронної мережі схожа з біологічною, але різниця в тому що замість живих клітин у нас пристрої які створили люди.

Далі будемо розглядати тільки штучні нейронні мережі, тому будь-яке згадування нейронних мереж розуміється як штучна нейронна мережа.

Структура штучного нейрона зображена на рисунку 2.1.

Умовні позначення які використані у структурі штучного нейрону:

- 1) $X_1 X_2 \dots X_m$ – входи нейронів;
- 2) $W_1 W_2 W_m$ – ваги нейронів;
- 3) Y_{in} – сума на вході нейронів;
- 4) F – функція активації нейрона;
- 5) Y – вихід нейрону.

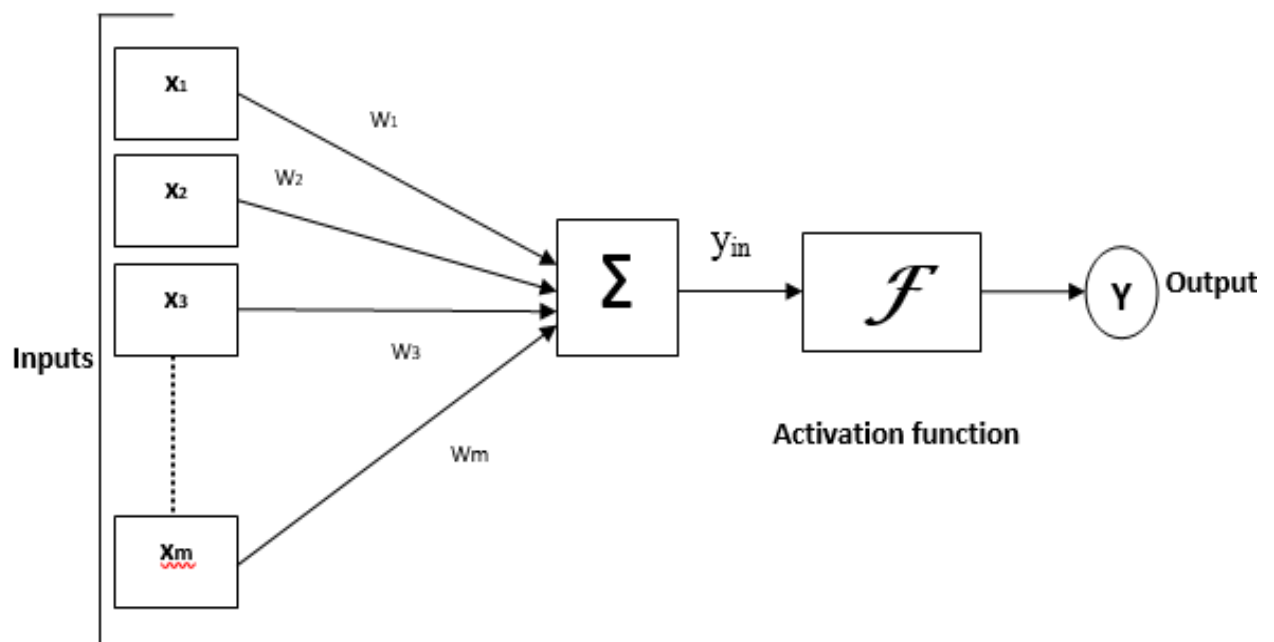


Рисунок 2.1 – Структура штучного нейрона

Далі буде приведено ряд термінів для кращого розуміння предметної області та роботи нейронних мереж.

Так як у нас мережа то нейрони об'єднані між собою в різній послідовності. Це називається архітектурою мережі.

Архітектура – опис числа шарів у нейронній мережі, передавальних (активаційних) функцій кожного шару, числа нейронів у кожному шарі та зв'язків між шарами.

Шар - безліч нейронів (вузлів), що мають спільні вхідні або вихідні сигнали.

Вхідний шар - шар нейронів, утворений сукупністю елементів, на вході яких подаються вхідні дані нейронної мережі.

Прихований (проміжний) шар – шар, що містить нейрони, на які не надходять вхідні дані та з яких не зчитуються вихідні дані нейронної мережі. Цей шар отримує вихід з попереднього шару (може бути вхідний або інший прихований) та передає дані на наступний шар (може бути або прихований або вихідний).

Вихідний шар – шар нейронів, вихідні сигнали яких утворюють вихідний сигнал нейронної мережі.

Структура нейронної мережі представлена на рисунку 2.2.

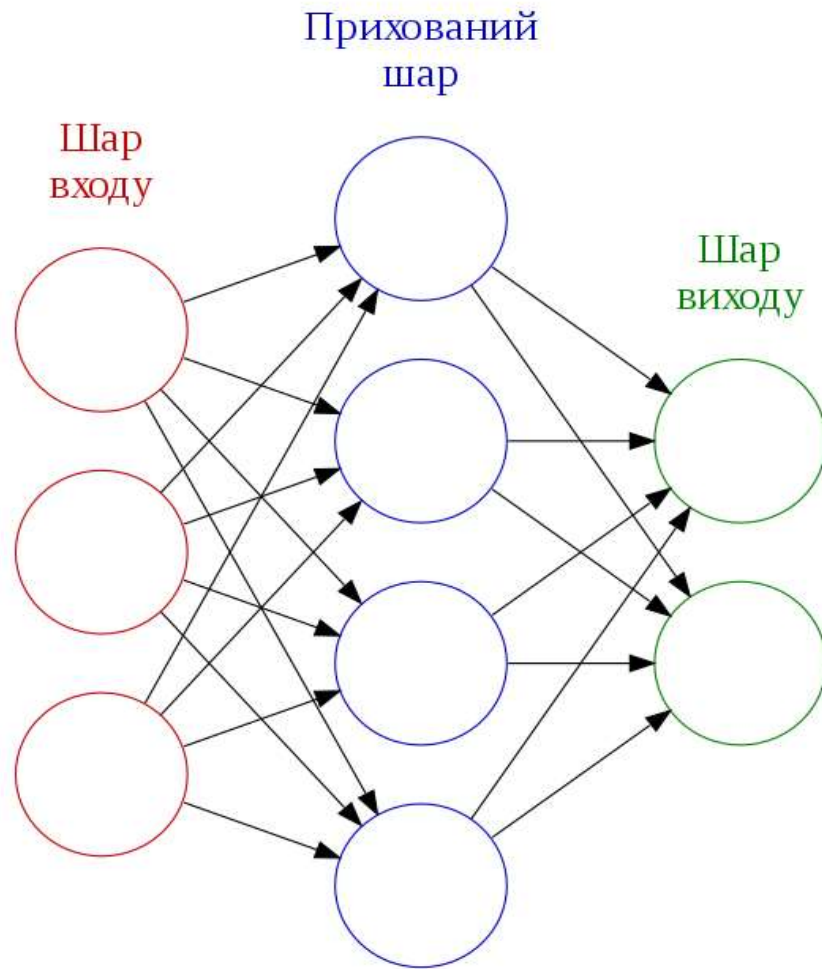


Рисунок 2.2 – Структура нейронної мережі

Навчання (тренування) – етап функціонування нейронної мережі, у процесі якого на її вхід по черзі надходять дані з навчального набору з метою коригування вагових коефіцієнтів синаптичних зв'язків для отримання найбільш адекватного сигналу на виході нейронної мережі. Цей етап займає найбільше часу при достатній складності нейронної мережі та об'єму даних.

Навчання у попередньому режимі (off-line training) – різновид алгоритму навчання, коли корекція ваги і зміщень проводиться один раз за період навчання – після пред'явлення всіх векторів навчальної послідовності («пачки» векторів, batch). Цей метод не дуже зручний в нашому випадку, тому що в нас можуть поповнюватися дані і система повинна підвантажувати їх та обробляти для вдалої класифікації.

Навчання в режимі реального часу (on-line training) – модифікація алгоритму навчання, коли ваги та усунення мережі коригуються після пред'явлення кожного нового образу (вектора, в нашому випадку зображення) навчальної послідовності. Такий режим роботи більш продуктивний, нові зображення додаються та відразу нейрона мережа їх підхоплює та навчається.

Навчання з учителем – процес навчання нейронної мережі, неодмінною вимогою якого є існування готового навчального набору даних. Готовий набір даних це розмічений набір, про проблеми розмітки даних було написано раніше у пункті 1.4.

Навчання без вчителя – процес навчання нейронної мережі, у якому наявність набору еталонів відсутня. Саме таким методом будемо користуватися для тренування нашої нейронної мережі.

На рисунку 2.3 представлено загальна схема роботи навчання з учителем на розмічених даних (machine learning) та навчання без вчителя за допомогою глибокого навчання(deep learning).

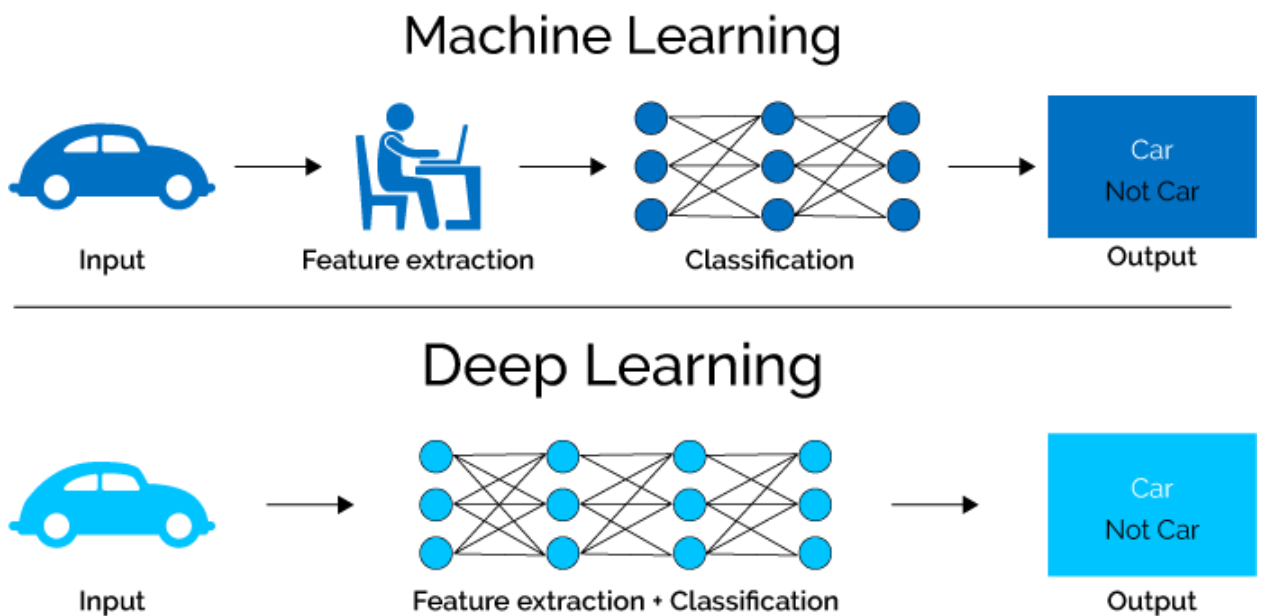


Рисунок 2.3 – Загальна схема машинного (з вчителем) та глибокого (без вчителя) навчання

Зворотне поширення – найбільш застосовуваний спосіб керованого навчання, при якому сигнал помилки на виході нейронної мережі поширюється у зворотному напрямку: від нейронів вихідного шару до нейронів вхідного шару з подальшим коригуванням синаптичних ваг нейронної мережі для досягнення мінімальної вихідної похибки.

Тестування – етап перевірки працездатності нейронної мережі, протягом якого на вхід мережі подаються дані, які не були використані у процесі навчання, але вхідні зразки котрих відомі, що дозволяє оцінити вихідну похибку навченої нейронної мережі. У процесі тестування вагові коефіцієнти нейронної мережі не змінюються.

Глибоке навчання - сукупність широкого сімейства методів машинного навчання, заснованих на імітації роботи людського мозку в процесі обробки даних та створення патернів, що використовуються для прийняття рішень. Як правило, глибоке навчання призначене для роботи з великими обсягами даних та використовує складні алгоритми для навчання моделі. На великих датасетах глибоке навчання демонструє більш високу точність результатів у порівнянні з традиційним машинним навчанням (тобто навчанням з учителем, на розмічених даних)[6].

Одним із методів глибокого навчання є згорткові нейронні мережі.

Для нашого задання з обробки зображень також підходять згорткові нейронні мережі. На рисунку 2.4 представлено принцип її роботи. Наше зображення розміром 5×5 згортається в зображення розміром 2×2 , таким чином ми враховуємо всі ознаки які були на великому зображенні, та комп'ютеру легше обробляти менше за розміром зображення. Тобто вхідні дані будуть меншого розміру та швидше обробляться[7].

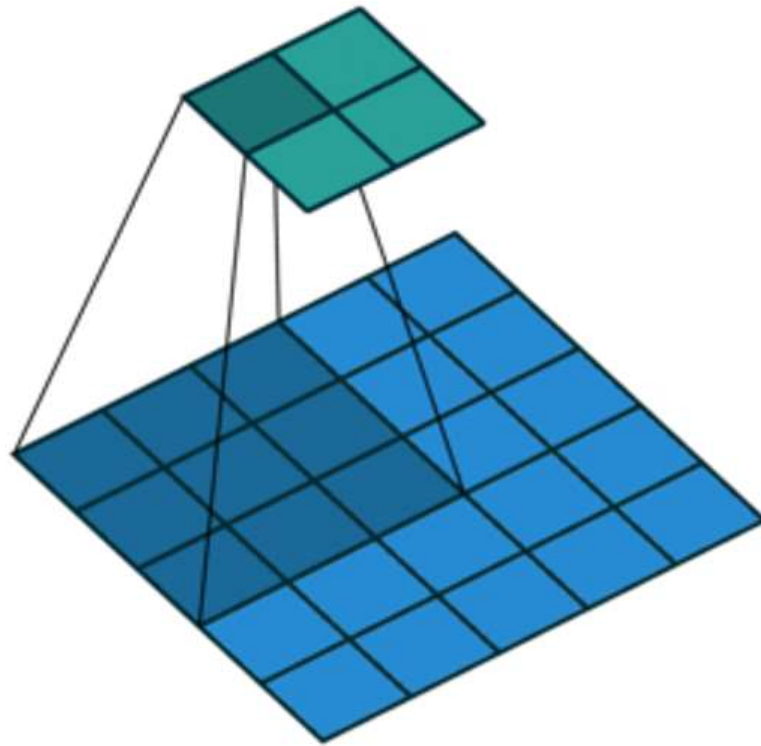


Рисунок 2.4 – Принцип роботи згорткової нейронної мережі

На рисунку 2.5 показано графік залежності продуктивності нейронних мереж від кількості даних. Як бачимо глибоке навчання (deep learning) продуктивніше при великій кількості вхідних даних, в той час як звичайне машинне навчання має деякий поріг після якого не відбувається покращення продуктивності.

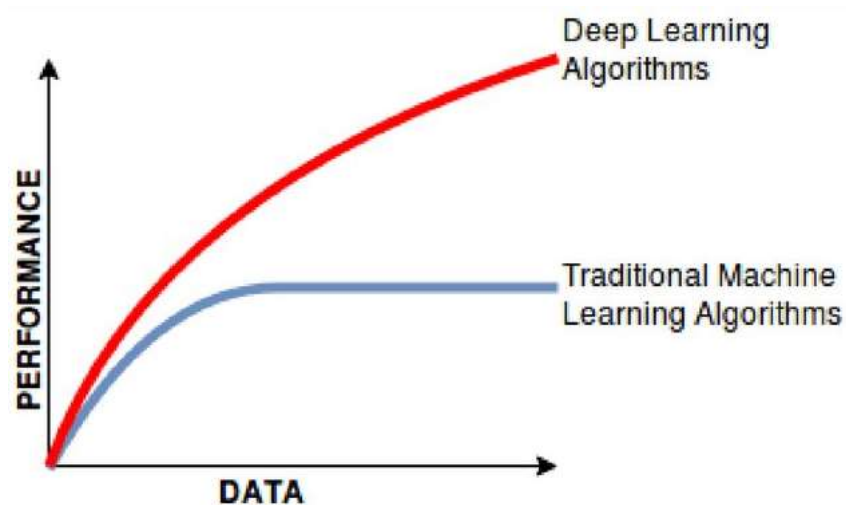


Рисунок 2.5 - Графік залежності продуктивності нейронних мереж від кількості даних

2.3 Огляд бібліотеки Tensorflow

TensorFlow – це фреймворк від Google, який призначений для проектування, створення та вивчення моделей глибокого навчання. Глибоке навчання - це область машинного навчання, алгоритми в якій були натхненні структурою та роботою мозку. Ви можете використовувати TensorFlow, щоб робити чисельні обчислення. Саме це не здається специфічним, проте ці обчислення виробляються за допомогою data-flow графів. У цих графах вершини являють собою математичні операції, у той час як ребра є даними, які зазвичай представляються у вигляді багатовимірних масивів або тензорів, які повідомляються між цими ребрами.

Тензор – математичний об'єкт, що узагальнює такі поняття як скаляр, вектор, ковектор, лінійний оператор і білінійна форма. Вивченням тензорів займається тензорне числення. В деякому базисі тензор представляється у вигляді багатовимірної таблиці (число співмножників збігається з валентністю тензора), заповненої числами (компонентами тензора). При заміні базису компоненти тензора змінюються певним чином, при цьому сам тензор не залежить від вибору базису.

Якщо проводити розрахунки з тензорами на центральному процесорі, то це займе багато часу, то процесор зроблений для вирішення складних дій, тому краще використовувати графічний процесор, бо його архітектура більш підходить для розв'язання таких задач. Операції з тензорами добре паралеляться, тому використання відеооядер більш доцільне.

У 2016 році компанія Google представила тензорний процесор, що пришвидшує обчислення з використанням тензорів. Тензорний процесор це інтегральна схема специфічного застосування (ASIC) призначена для прискорення розрахунків штучного інтелекту, що була розроблена компанією Google спеціально для машинного навчання нейронних мереж для бібліотеки TensorFlow. Зараз вже розроблено 3 покоління таких процесорів.

Але тензорні процесори є тільки в компанії Google, та користуватися можна ними якщо використовувати сервіс Colaboratory. Компанія Nvidia у своїх графічних процесорах також додає тензорні ядра для прискорення машинного навчання. На рисунку 2.6 зображено архітектуру тензорних ядер декількох поколінь від компанії Nvidia.

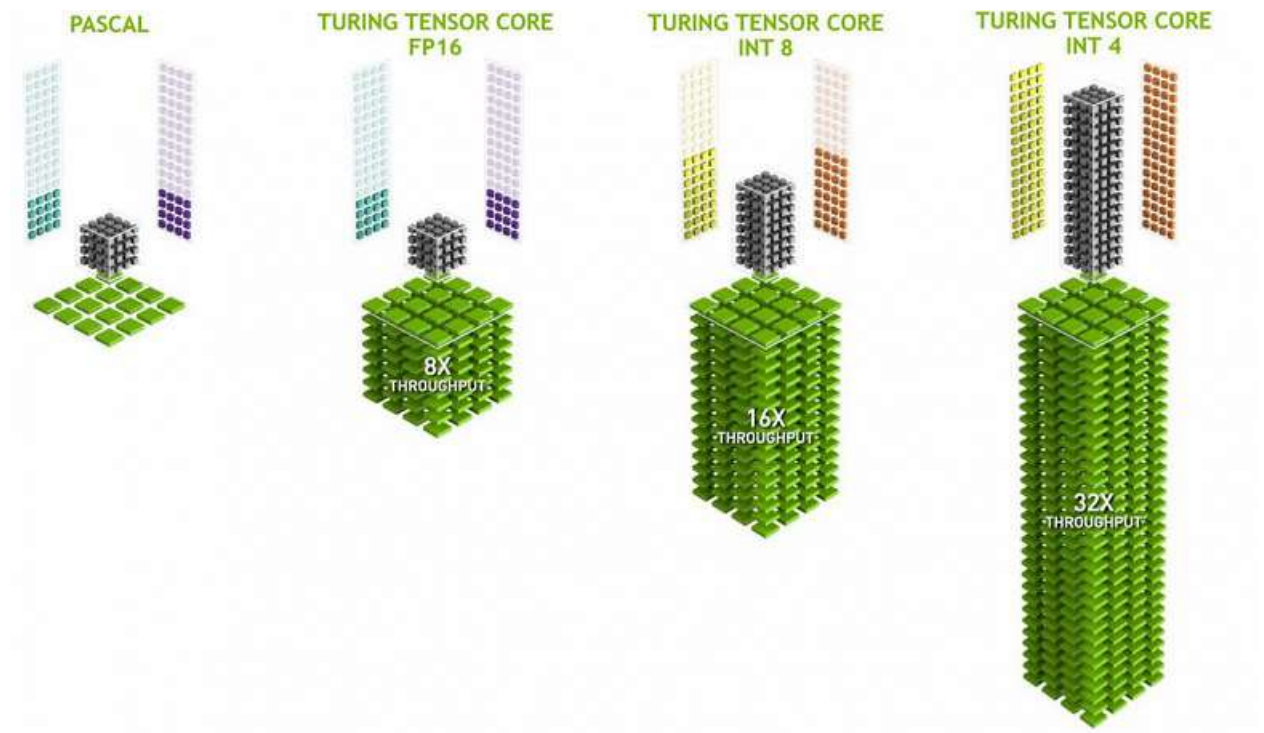


Рисунок 2.6 – архітектура тензорних ядер від компанії Nvidia

Як можемо бачити на рисунку, потужність обчислювальних ядер значно зросла за останні декілька років[8].

2.4 Метод зворотного розповсюдження помилки

Навчання нейронних мереж відбувається в два етапи, перший етап це пряме розповсюдження помилки, другий – зворотне розповсюдження помилки. Під час прямого розповсюдження помилки робиться прогноз відповіді.

При зворотному розповсюдженні помилка між фактичною відповіддю і передбаченим результатом мінімізується. Схема етапу прямого розповсюдження помилки наведена на рисунку 2.7.

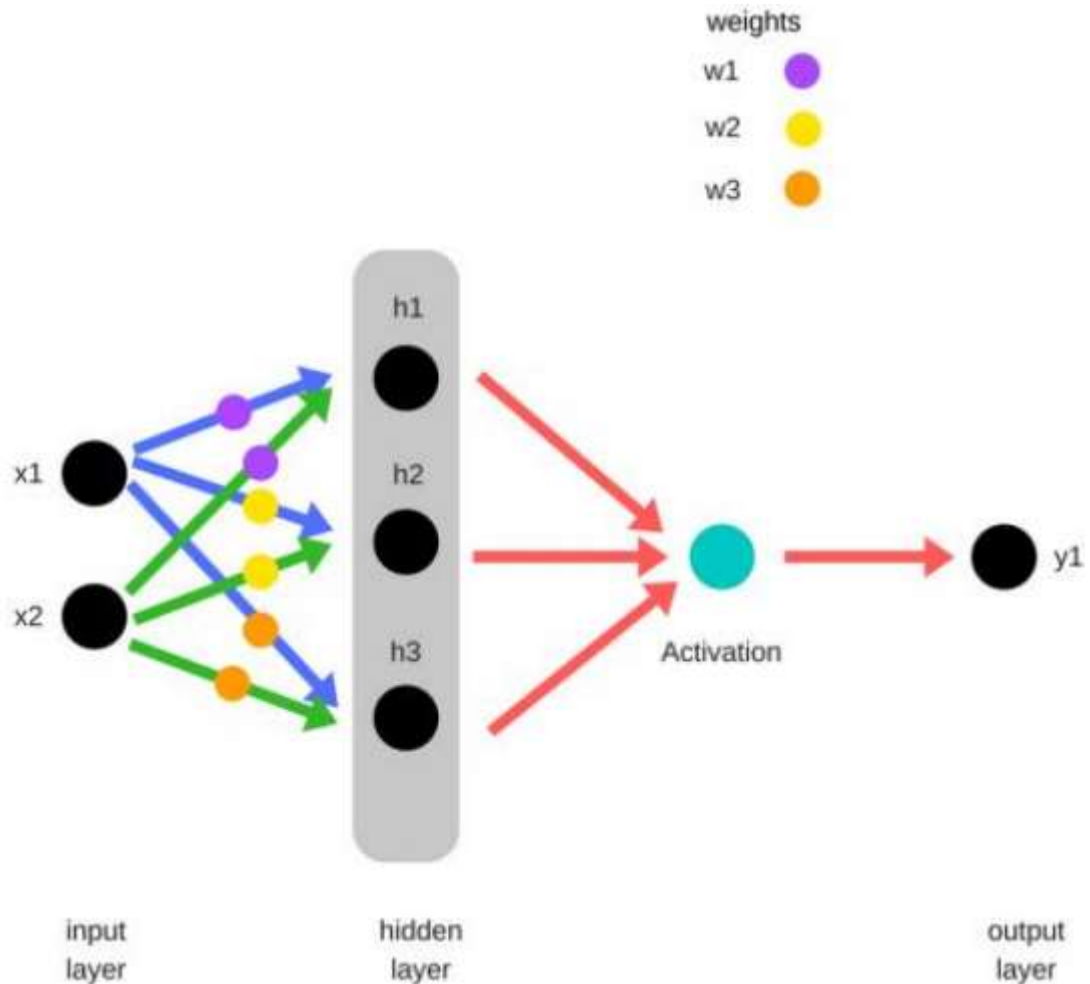


Рисунок 2.7 – Схема прямого розповсюдження помилки

Перший етап прямого розповсюдження працює наступним чином: спочатку задаються довільні, початкові значення ваг w_1 , w_2 , w_3 . Потім ці ваги множаться на вхідний вектор, для формування прихованого шару. Далі, данні з прихованого шару передаються через деяку нелінійну функцію активації для отримання першого виходу мережі. Після цього настає час другого етапу, етапу зворотного розповсюдження помилки, загальна схема якого наведена на рисунку 2.8.

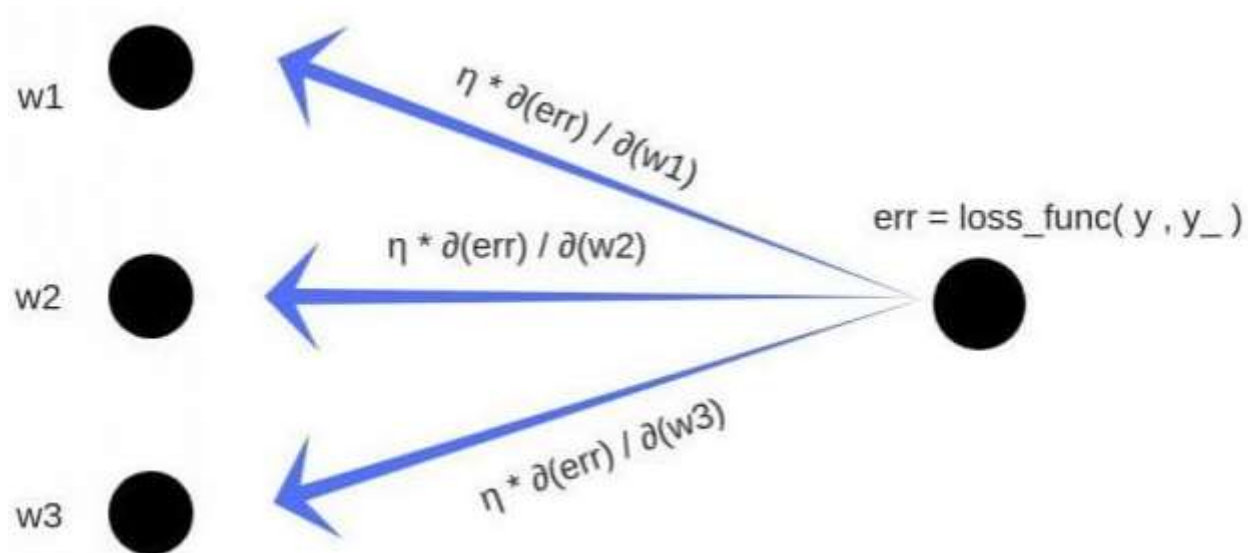


Рисунок 2.8 – Схема зворотного розповсюдження помилки

Основним методом для навчання нейронної мережі у бібліотеці TensorFlow є метод зворотного розповсюдження помилки. Це такий метод навчання багатошарового перцептрону(вид нейронної мережі). Це ітеративний градієнтний алгоритм, який використовується з метою мінімізації помилки роботи нейронної мережі та отримання бажаного виходу. Основна ідея цього методу полягає в поширенні сигналів помилки від виходів мережі до її входів, в напрямку, зворотному прямому поширенню сигналів у звичайному режимі роботи. Для можливості застосування методу зворотного поширення помилки функція активації нейронів повинна бути диференційованою[9].

Навчання нейронних мереж залежить декількох параметрів, а саме:

- функція активації;
- швидкість навчання;
- функція втрат.

Ці параметри обов'язково потрібно налаштовувати, бо від цього залежить результативність навчання.

Функція активації визначає вихідний сигнал, який визначається вхідним сигналом. Тобто ця функція показує як вихідний сигнал залежить від вхідного. Більшість нейронних мереж використовують сигмоїдальні функції активації.

Метод зворотного розповсюдження помилки потребує щоб функція активації була неперервною, нелінійною, монотонно зростаючою і диференційованою.

Нейронні мережі без функцій активації втрачають велику частину своєї здатності до навчання. Якщо функції нелінійні то ступінь свободи підвищується, коли потрібно узагальнити проблеми високої розмірності в більш низьких вимірах[10].

На рисунку 2.9 зображено приклади поширених функції активації та їх графіки.

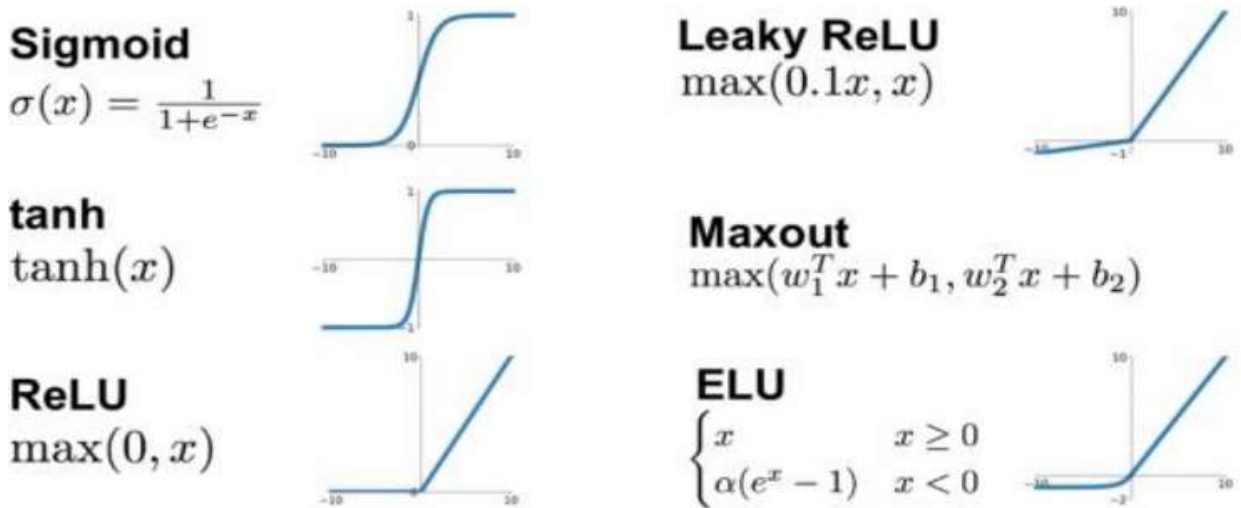


Рисунок 2.9 – Приклади функцій активації

Швидкість навчання також дуже важлива для результату. По-перше, давайте прояснимо, що ми маємо на увазі під «навчанням». У контексті нейронних мереж «вчитися» за значенням більш-менш еквівалентно «тренуватися», але з іншого боку це різні речі. Інженер навчає нейронну мережу, надаючи навчальні дані та виконуючи процедуру навчання. У той час як це відбувається, мережа вчиться - або, більш конкретно, вона вчиться апроксимувати зв'язок вхід-вихід, що міститься в навчальних даних. Проявом навчання є зміна ваги, а швидкість навчання впливає на спосіб цієї зміни.

Існує концепція чаші помилки, тобто тривимірної поверхні, яка допомагає нам візуалізувати процес, за допомогою якого помилка вузла поступово зменшується до нуля у міру того, як його вхідні коефіцієнти ваги змінюються під час навчання. Графічне представлення чаші похибки представлено на рисунку 2.10.

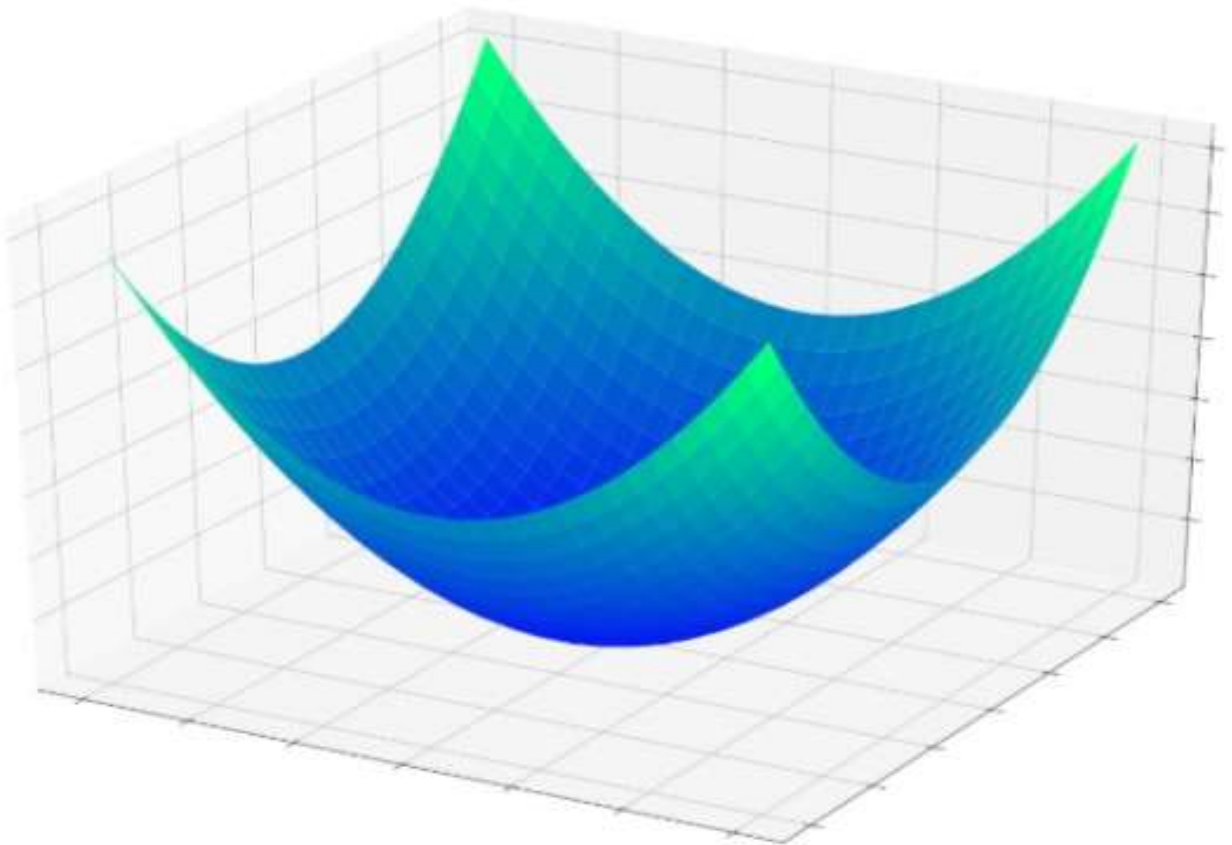


Рисунок 2.10 – Чаша похибки

Якщо у нас є точка, розташування якої визначається значеннями двох ваг і вихідною помилкою вузла, кожна модифікація ваг змушує точку перескакувати в інше місце десь на цій поверхні помилки. Ці стрибки прагнуть досягнути нижню частину чаші, де помилка зводиться до мінімуму; вони не призводять безпосередньо до мінімальної помилки, тому що кожна навчальна вибірка – це лише маленький шматочок математичної головоломки.

Швидкість навчання впливає на розмір стрибків, які ведуть до нижньої частини чаші. Зараз я маю намір перейти до двовимірної вистави, тому що

буде легше створювати та інтерпретувати зображення. Наша двовимірна функція помилки представлена на рисунку 2.11.



Рисунок 2.11 – Двовимірне представлення функції помилки

Для оновлення ваг ми використовували таке правило навчання

$$w_{\text{новий}} = w + (\alpha \times \delta \times \text{вхід}), \quad (2.1)$$

де α – швидкість навчання,

δ – різниця між очікуваним вихідним сигналом та розрахованим вихідним сигналом (тобто похибка).

Щоразу, коли ми застосовуємо це правило навчання, вага переходить у нову точку на кривій помилки. Якщо δ велика, ці стрибки також можуть бути досить великими, і нейромережа може навчатися неефективно, тому що ваги поступово не сходяться до мінімальної помилки. Натомість вони хаотично стрибають, як показано на рисунку 2.12.

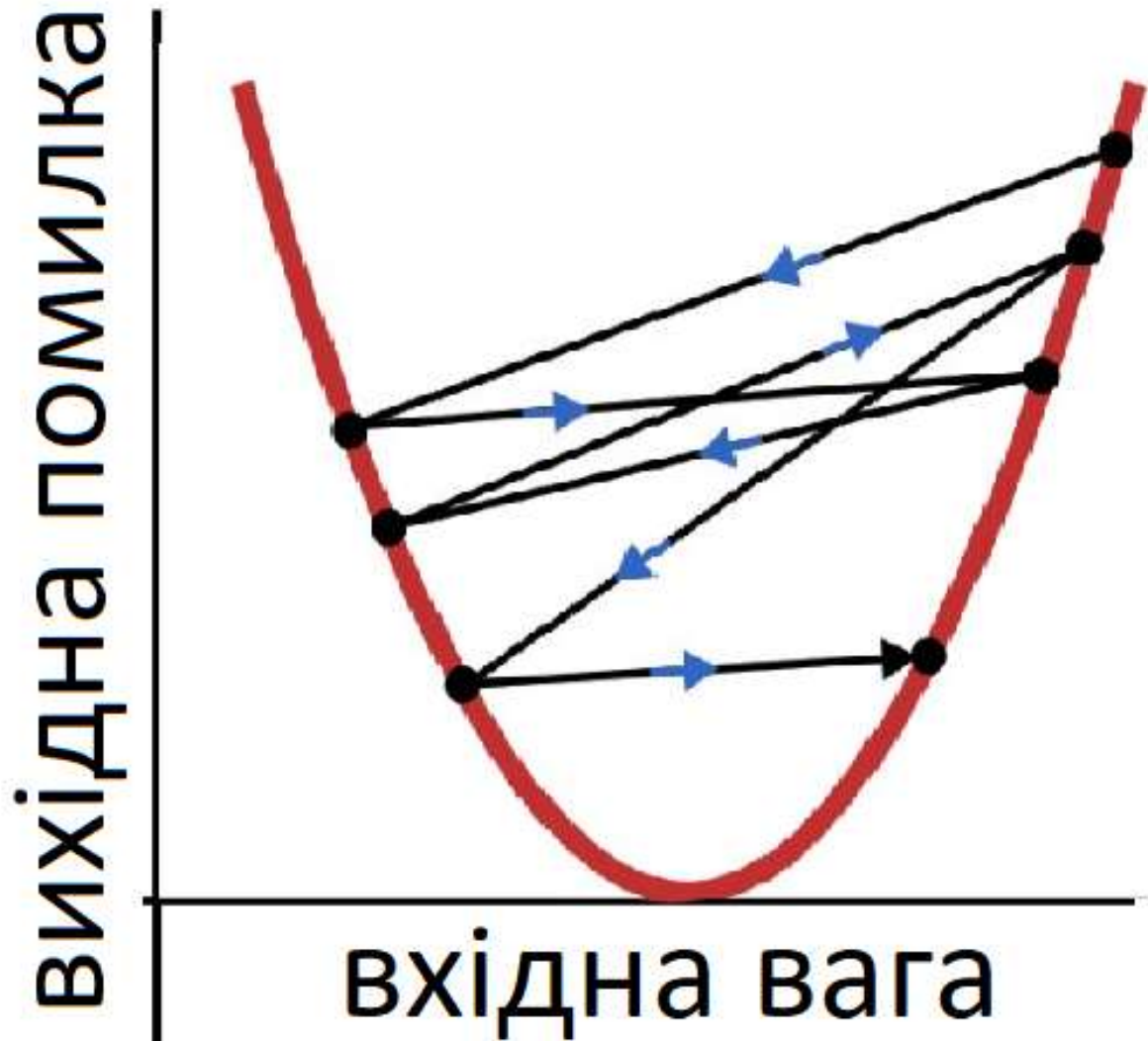


Рисунок 2.12 – Точка розташування помилки при стрибках

Оскільки модифікація застосовується до ваги, δ множиться на швидкість навчання, ми можемо зменшити розмір стрибків, вибравши $\alpha < 1$. Ціль – використувати швидкість навчання для забезпечення помірно швидкої та послідовної збіжності.

Тип навчання який нам потрібен може виглядати так як показано на рисунку 2.13.

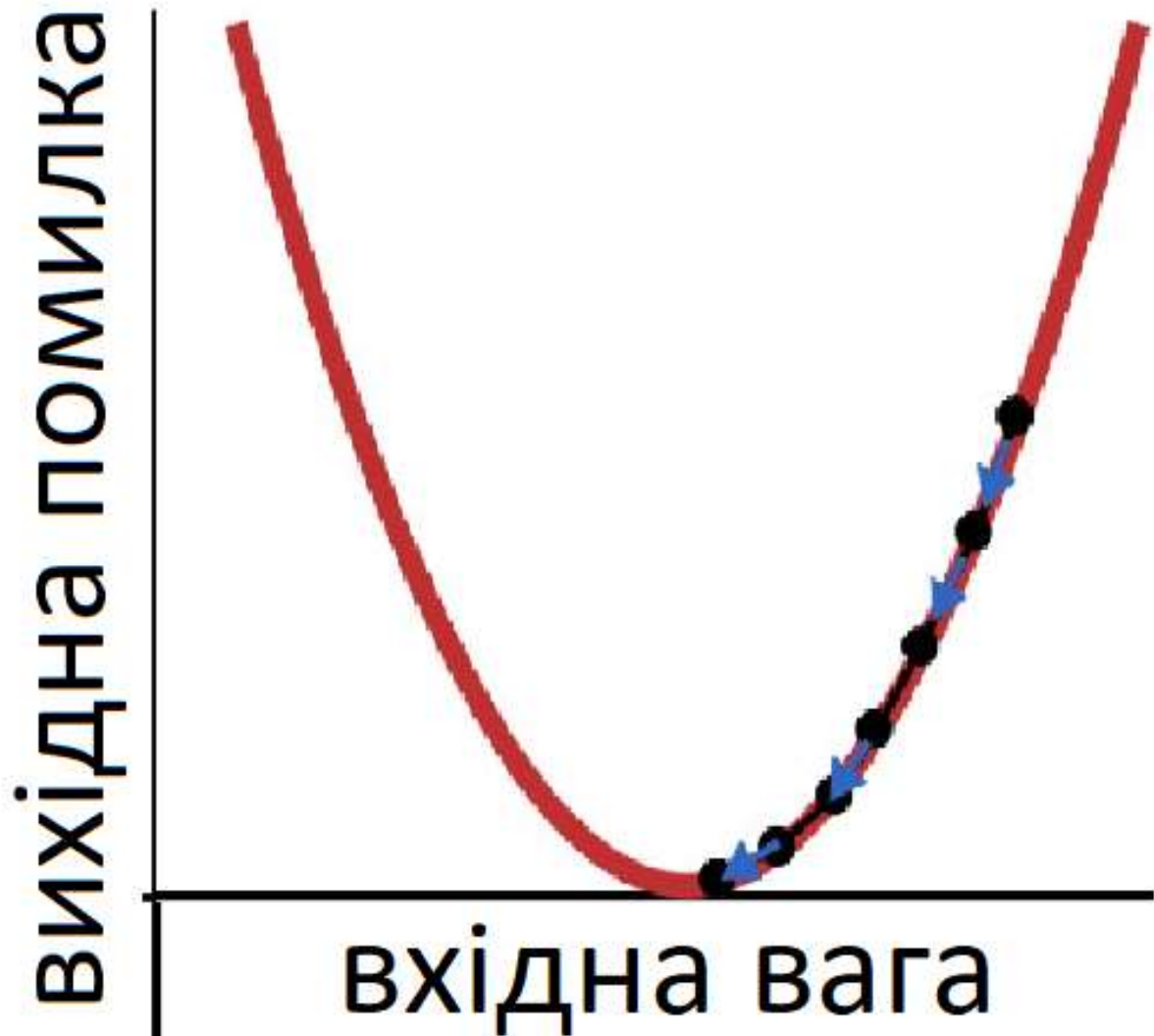


Рисунок 2.13 – Відповідна швидкість навчання допомагає знайти мінімальну помилку

Як обрати швидкість навчання? Не існує універсального правила, яке говорить вам, як вибрати швидкість навчання, і немає навіть точного та акуратного способу визначити оптимальну швидкість навчання для заданої програми. Навчання – це складний та мінливий процес, і коли справа доходить до швидкості навчання, ви повинні покладатися на інтуїцію та експерименти. Якщо ваша нейронна мережа може швидко обробляти навчальні дані, ви можете просто вибрати кілька різних швидкостей навчання та порівняти отримані ваги (якщо ви знаєте, якими вони мають бути) або ввести свіжі дані та оцінити взаємозв'язок між швидкістю навчання та точністю класифікації.

Більш складний підхід, який був би практичнішим для нейромереж, що вимагають тривалого часу навчання, полягає в аналізі змін у помилках під час навчання мережі. Помилка повинна зменшуватися в напрямку мінімуму, а зміни в помилці повинні бути достатньо малими, щоб уникнути стрибання поведінки, показаного вище, але не настільки маленькими, щоб нейронна мережа навчалася вкрай повільно. Як і багато інших речей у житті, швидкість навчання залежить від балансу.

Перше, що треба визнати, це те, що швидкість навчання не обов'язково має бути постійною протягом усієї процедури навчання. Швидкість навчання використовується щоразу, коли за допомогою правила навчання оновлюються ваги; таким чином, якщо швидкість навчання змінюється під час навчання, то відразу буде змінено і еволюційний шлях нейронної мережі до її остаточної форми.

Одним із способів скористатися цим є зниження швидкості навчання під час навчання. Це називається відпалом (annealing) швидкості навчання. Існують різні способи реалізувати це, але зараз важливо зрозуміти, чому це допомагає.

Коли нейронна мережа починає навчання, помилка, ймовірно, буде великою. Вища швидкість навчання допомагає нейронній мережі робити великі кроки у напрямку мінімальної помилки. Однак коли мережа наближається до нижньої частини кривої помилки, ці довгі кроки можуть перешкоджати збіжності, подібно до того, як людині, яка робить довгі кроки, може бути важко потрапити прямо в середину невеликого кола, намальованого на підлозі. Коли швидкість навчання зменшується, довгі кроки стають меншими кроками і, зрештою, нейронна мережа наближається до центру кола[11].

Функція втрат використовується для розрахунку помилки між реальними і отриманими відповідями. Головна наша мета це мінімізація цієї помилки. Таким чином, функція втрат ефективно наближає навчання нейронної мережі до цієї мети. Функція втрат вимірює «наскільки гарна» нейронна мережа

щодо даної навчальної вибірки і очікуваних відповідей. Вона також може залежати від таких змінних, як ваги і зміщення.

Функція втрат одномірна і не є вектором, оскільки вона оцінює, наскільки добре нейронна мережа працює в цілому. Найбільш широковідомі функції втрат це:

- квадратична (середньоквадратичне відхилення);
- крос-ентропія;
- експоненціальна (AdaBoost);
- відстань Кульбака - Лейблера або приріст інформації.

Функція втрат в нейронній мережі повинна відповідати двом головним умовам:

- функція втрат повинна бути записана як середнє;
- функція втрат не повинна залежати від будь-яких активаційних значень нейронної мережі, крім значень, які видаються на виході.

Функція втрат розраховується на кожній епосі і повинна йти на спад протягом процесу навчання[12]. Крім того, функція втрат може виступати у якості індикатору для уникнення такого явища як перенавчання, або оверфіт – негативного явища, що виникає, коли алгоритм навчання виробляє передбачення, які занадто близько або точно відповідають конкретному набору даних і тому не підходять для застосування алгоритму до додаткових даних або майбутнім спостереженнями. Порівняння випадків недонавчання, добре навченої моделі і перенавчання на прикладі класифікації за допомогою логістичної регресії наведено на рисунку 2.14.

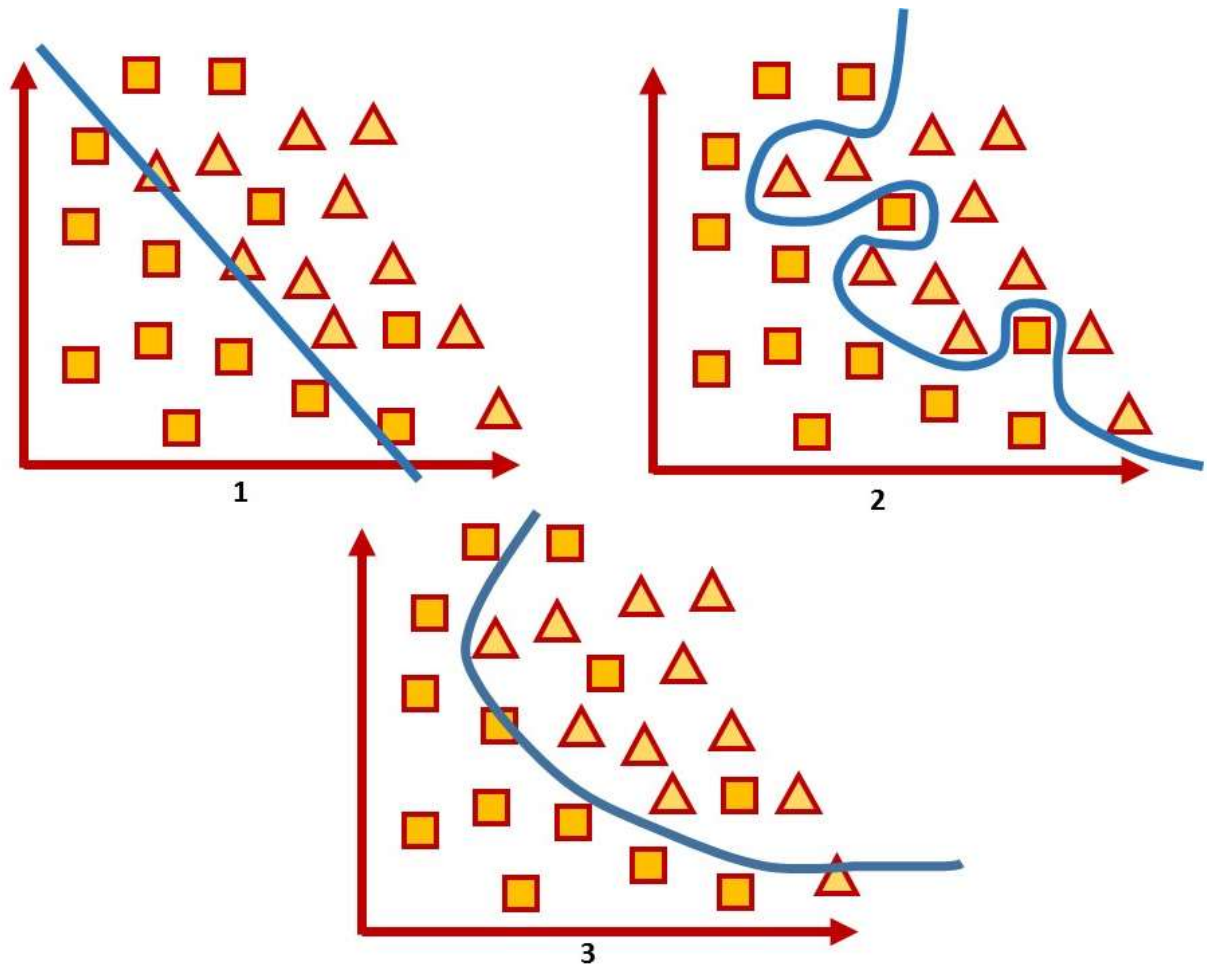


Рисунок 2.14 – 1) недонавчання, 2) перенавчання, 3) добре навчена модель

Перенавчання (англ. *overfitting*) — негативне явище, що виникає, коли алгоритм навчання виробляє передбачення, які надто близько чи точно відповідають конкретному набору даних і тому не підходять для застосування алгоритму додаткових даних чи майбутніх спостережень[13].

Недонавчання (англ. *underfitting*) - негативне явище, при якому алгоритм навчання не забезпечує достатньо малої величини середньої помилки на навчальній вибірці. Недонавчання виникає при використанні недостатньо складних моделей[13].

Для більш повної картини треба розглянути поняття кривої навчання.

Крива навчання – графічне уявлення про те, як зміна міри навченості (по вертикальній осі) залежить від певної одиниці виміру досвіду (по горизонтальній осі). У прикладах нижче представлена залежність середньої помилки від обсягу датасету[13].

При перенавченні невелика середня помилка на навчальній вибірці (training error) не забезпечує такої ж малої помилки на тестовій вибірці (test error). На рисунку 2.15 представлено графік кривих навчання при перенавченні.

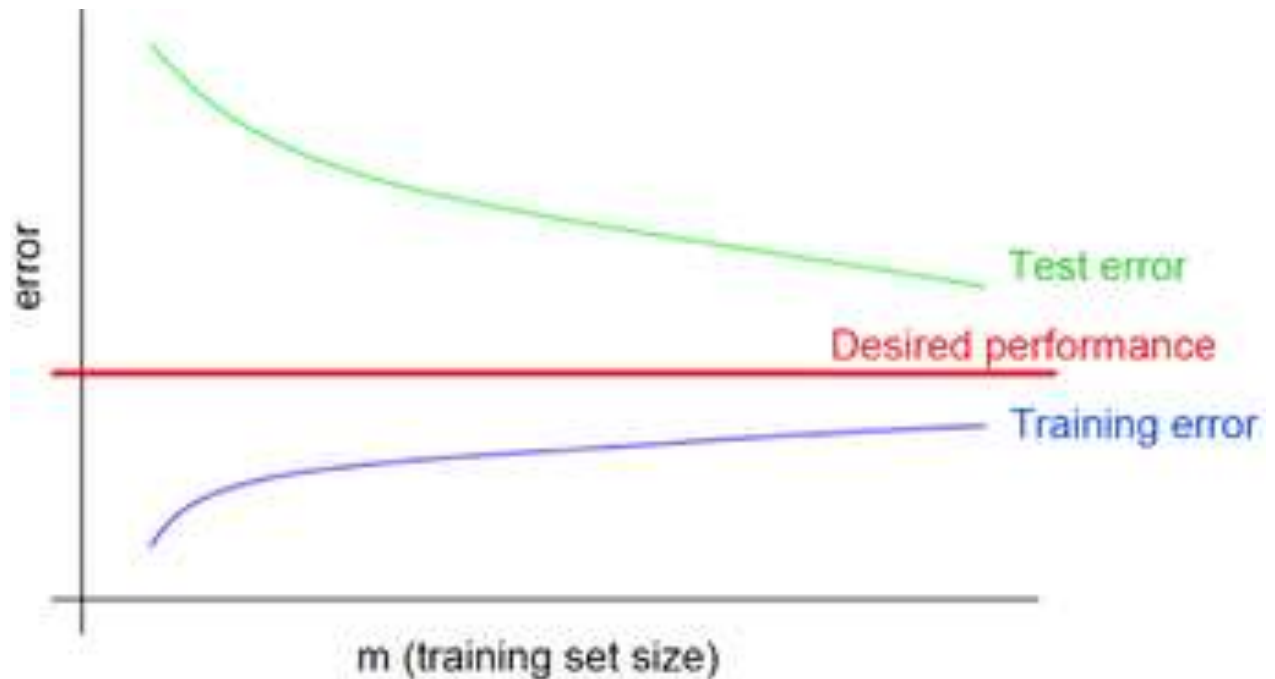


Рисунок 2.15 – Графік залежності середньої помилки для навчальної та тестової вибірок від обсягу датасету під час перенавчання

При недонавчанні незалежно від обсягу навчального датасета як на навчальній вибірці, так і на тестовій вибірці невелика середня помилка не досягається. На рисунку 2.16 представлено графічне зображення кривих навчання при недонавчанні.

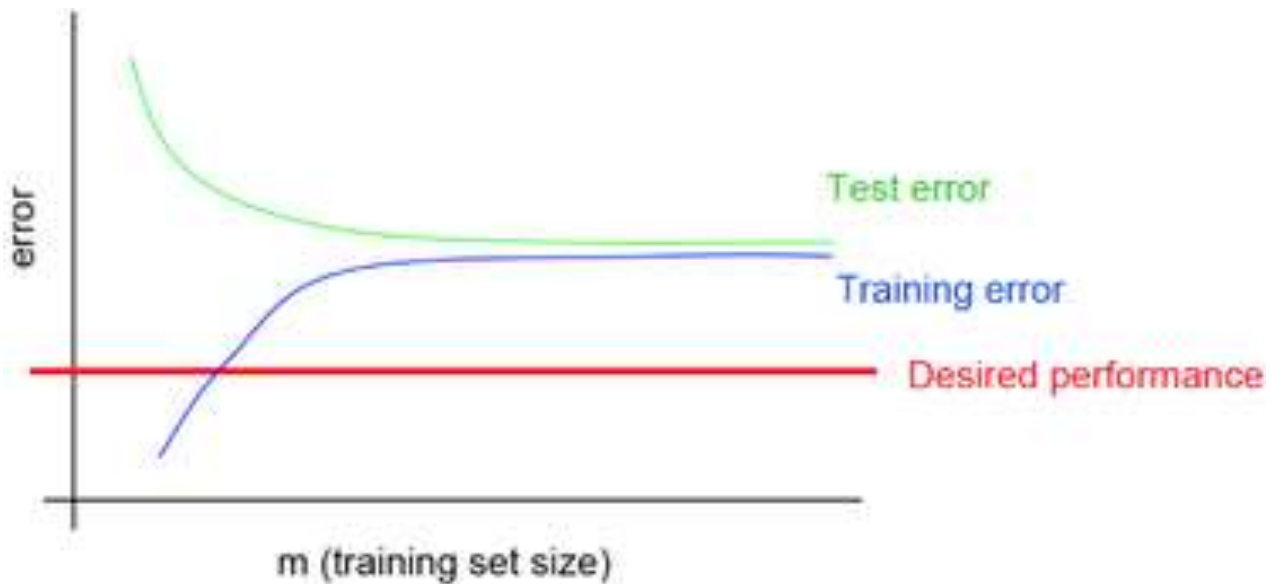


Рисунок 2.16 – Графік залежності середньої помилки для навчальної та тестової вибірок від обсягу датасету при недонавчанні

Перенавчання та недонавчання це проблеми з якими зтикаються всі розробники нейронних мереж. При підборі не коректних параметрів такі результати можуть бути отримані та при неправильній інтерпретації призведуть до неправильних дії тих людей які їх використовують. Стосовно нашої області лікарі можуть подумати що хвора людина здорова та навпаки, і якщо не зробити додаткових досліджень та аналізів можна отримати дуже трагічні наслідки[13]..

Для подолання перенавчання можна використовувати такі методи:

- збільшення кількості даних в датасеті;
- зменшення кількості параметрів моделі;
- додавання регуляризації/збільшення коефіцієнта регуляризації (додавання обмежень).

Для подолання недонавчання можна використовувати такі методи:

- додавання нових параметрів моделі;
- використання для опису моделі функцій більш високого ступеню;
- зменшення коефіцієнту регуляризації.

2.5 Постановка задачі

Створення системи класифікації зображень потрібно для підтримки прийняття рішень у постановці діагнозу лікарем. Користувачем системи буде лікар який завантажує зображення аналіз якого необхідно провести.

Мета створення такої системи – спростити постановку діагнозу лікарям які використовують у своїй роботі різні медичні зображення, зокрема зображення МРТ, КТ, рентгенографії, тощо.

Основні функціональні можливості для користувачів:

1. Завантажити зображення для діагностування хвороби.
2. Переглянути результати аналізу зображення, які повідомляють що за хвороба є у пацієнта та/або ймовірність того чи іншого захворювання.

Нефункціональні вимоги:

1. Система має стабільно працювати при одночасному використанні 1000 клієнтів;
2. Система при завантаженні прогнозу має показувати користувачу що вона активна, має бути присутній індикатор завантаження;
3. Система має бути захищена від несанкціонованого доступу.

При вирішенні завдань машинного навчання за допомогою нейронних мереж можемо виділити наступні кроки:

- визначити завдання яке буде вирішувати нейронна мережа;
- визначити гіперпараметри (швидкість, точність);
- визначити параметри вхідних даних (тип зображення, формат, розмірність) та параметри вихідних даних (вихідні класи, для кожного типу дослідження будуть свої класи);
- визначити оптимальну структуру нейронної мережі для вирішення поставленої задачі.

Задача створюваної нейронної мережі це сегментація та класифікація зображень. Швидкість відповіді має становити не більше 10 секунд. Загальна то-

чність розпізнавання не менше 90%. Вхідними даними можуть виступати зображення отримані в результаті медичних досліджень (МРТ, КТ), у форматі .png та розмір від 255*255 пікселів до 2048*2048 пікселів. На виході ми маємо сегменти які визначила нейронна мережа та класи здоровий пацієнт або є хвороба.

У результаті аналізу предметної області та огляду існуючих систем для класифікації зображень було розроблено план для реалізації завдання. План розробки елементів системи має такі етапи:

- проведення теоретичних досліджень і розгляд існуючих методів машинного навчання у контексті застосування їх до обраної галузі;
- пошук і підготовка вхідного набору даних що буде використаний для навчання розробленої штучної нейронної мережі;
- проведення аналізу і вибір за його результатами найбільш прийнятних алгоритмів і методів для вирішення задачі класифікації зображень, а саме сфокусуватись на використанні штучних згорткових нейронних мереж для цієї мети;
- проведення необхідних експериментів, здійснення експериментального моделювання та навчання моделі, її тестування;
- проведення порівняльного аналізу роботи різноманітних алгоритмів;
- проведення аналізу отриманих результатів

В результаті даної роботи потрібно розробити нейронну мережу для сегментації та класифікації медичних зображень.

3 РОЗРОБКА НЕЙРОННОЇ МЕРЕЖІ

3.1 Сегментація

Сегментація з точки зору комп'ютерної обробки зображень це процес розділення зображення на окремі частини. Ці частини також називають суперпікселями. Головна мета сегментації це спрощення та/або зміна представлення зображення для полегшення аналізу або передачі каналами зв'язку. Зазвичай сегментація використовується для виділення об'єктів та меж (лінії, криві, тощо) на зображеннях. Сегментація зображень – це процес присвоєння міток кожному пікселю зображення, таким чином пікселі з однаковими мітками мають спільні візуальні характеристики. В результаті сегментації отримуємо множину сегментів, які разом покривають все зображення. Пікселі в сегменті схожі за деякими характеристиками, а пікселі в сусідніх сегментах мають істотні відмінності[14].

U-Net вважається однією зі стандартних архітектур згорткової нейронної мережі для задачі сегментації зображень, коли потрібно визначити клас зображення цілісним, але і сегментованим його областям по класу, тобто створити маску, яка буде розкладати зображення на кілька класів. Архітектура складається зі стягуючого шляху для захоплення контексту і симетричного розширювального шляху, який дозволяє здійснити точну локалізацію.

На рисунку 3.1 зображено приклад сегментації зображення пухлини мозку, у лівій частині зображено сегментацію за допомогою експерта, а на правій частині за допомогою нейронної мережі.

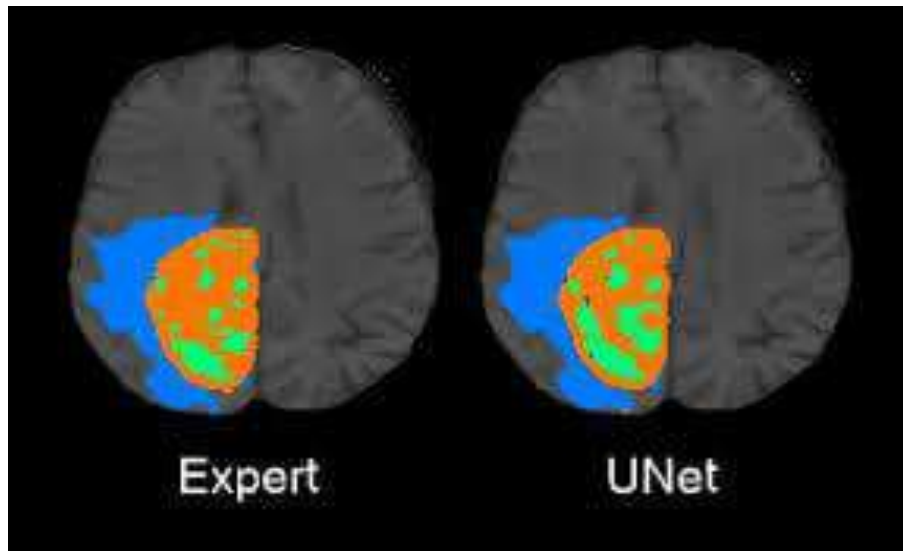


Рисунок 3.1 – Приклад сегментації

Навчена на невеликій кількості зображень. Окрім цього мережа працює відносно швидко. Сегментація зображень 512×512 займає менше секунди на сучасному графічному процесорі (рис. 2.18). Для U-Net характерно досягнення високих результатів у різних реальних задачах, особливо для біомедичних, і використання невеликої кількості даних для досягнення точних результатів.

На рисунку 3.2 представлено архітектуру U-Net.

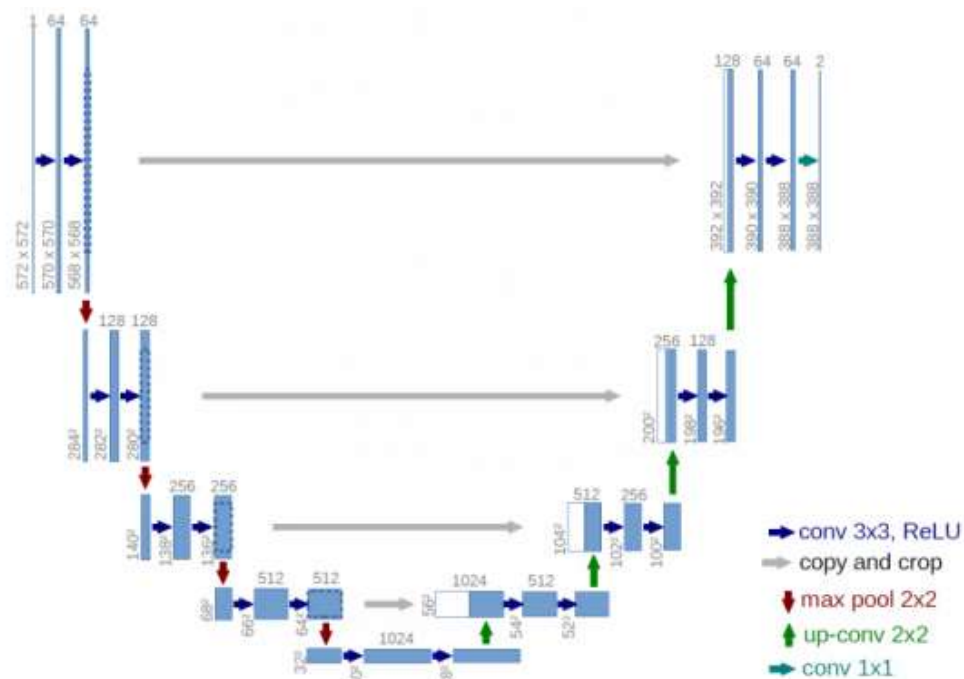


Рисунок 3.2 – Архітектура U-Net

Архітектура мережі складається зі звужуючого шляху (ліворуч) і розширеного шляху (праворуч). Звужений шлях - типова архітектура згорткової нейронної мережі. Шлях складається з повторного застосування двох згорток 3×3 , за якими слідує ініціалізація ReLU і операція максимального об'єднання (2×2 ступеня 2) для зниження розмірів.

На кожному етапі понижувальної дискретизації канали властивостей подвоюються. Кожен крок у розширенні шляху складається з операції підвищення дискретизації карти властивостей, за якою слідує: згортка 2×2 , яка зменшує кількість каналів властивостей; об'єднання відповідним чином обрізаної карти властивостей зі стягуючим шляхом; дві 3×3 згортки, за якими слідує ReLU. Обрізка необхідна через втрату граничних пікселів.

Схема роботи згорткової нейронної мережі представлена на рисунку 3.3.

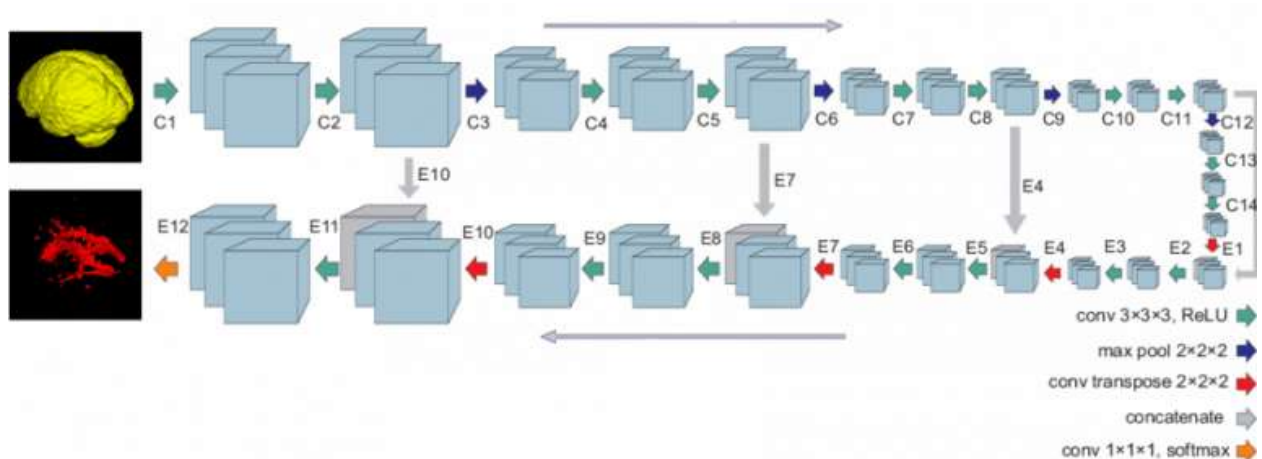


Рисунок 3.3 – Схема роботи згорткової нейронної мережі

На останньому шарі використовується згортка 1×1 для зіставлення кожного 64-компонентного вектора властивостей з бажаною кількістю класів. Всього мережа містить 23 згортальні шари[15].

3D U-Net - це широко відома архітектура для сегментації зображень, вперше запропонована в роботі [16]. При цьому вона може бути легко модифікована для сегментації тривимірних зображень [17]. Відповідно в експериментах була використана і тривимірна версія UNet.

У всій моделі, за винятком перших двох згортальних шарів, всі згортки були замінені на Residual Blocks [18], так як це значно підняло якість сегментації в експериментах. Також, через великі розміри зображень, було незначно зменшено кількість фільтрів в шарах.

При дослідженні алгоритмів сегментації було розроблено власну нейронну мережу, та вона показує гірший результат ніж U-Net, тому використовувати для сегментації будемо його.

Зараз розглянемо алгоритм сегментації зображення за допомогою нейронних мереж.

Для того, щоб покращити якість сегментації був сформований нейромережевий ансамбль. Більшість ансамблевих алгоритмів при створенні однорідних базових моделей використовують єдиний базовий метод навчання. Це призводить до однорідності комбінацій, але існують алгоритми, які використовують гетерогенні моделі (моделі різних типів). В результаті утворюються гетерогенні комбінації. Щоб ансамблі були більш точними, ніж будь-який їх окремий член, базові моделі повинні бути максимально різноманітними. Іншими словами, чим більше інформації надходить до об'єднувача від базових класифікаторів, тим вище точність комбінації. Деякі ансамблеві методи не визначають об'єднувача, але для тих методів, які це роблять існує три алгоритми об'єднання.

Один з алгоритмів працює по принципу "більшості голосів" (majority voting). Ідея полягає в наступному - для кожного вокселя вхідних знімків, обчислюється індивідуальний результат для кожної нейронної мережі на основі натренованих раніше моделей, потім воксель класифікується частиною пухлини тоді і тільки тоді, коли більшість нейромереж класифікує його як пухлину, в іншому випадку воксель є здоровою клітиною мозку. Наприклад, якщо в ансамблі 2 нейронні мережі, то воксель буде класифікований як частина пухлини, якщо хоча б одна з нейронних мереж визначить його як нездорову клітину. Якщо в ансамблі 3 нейронних мереж, тоді необхідно, щоб 2 нейромереж

режі класифікували воксель як клітину пухлини, щоб в результаті роботи ансамбля воксель також вважався враженою клітиною. Але, якщо навчити об'єднувач порівнювати вагові коефіцієнти, то ансамбль працюватиме по принципу “зваженої більшості голосів” (weighted majority voted) і Naive Bayes, а також підхід "classifier selection", в якому рішення по даному об'єкту приймає один класифікатор ансамблю. Очевидно, що нейронні мережі матимуть різні значення вагового коефіцієнта, що рівним чином формує різні узагальнення на будь-якій з ітерацій. Така випадковість диференціює помилки нейронних мереж, тому що вони будуть допускати їх на різних ітераціях вводу. Тому запропонований алгоритм, виходи базових класифікаторів якого, розглядаються як вхідні дані для нового класифікатора, який і стає об'єднувачем. Цей підхід використовує “складне узагальнення”, “узагальнення через навчання”, або просто stacking [19].

Використано набір відкритих даних, який включає в себе МРТ знімки 100 пацієнтів з пухлиною, як з високим рівнем злоякісності, так і з низькою [20, 21]. Вхідні данні - зображення МРТ, зроблені в режимах - T1, T2, T1Gd, Flair. Для всіх пацієнтів було виділені границі пухлини на кожному знімку. Усі інтенсивності на знімках були попередньо приведені до гаусівського розподілу із середнім 0 і стандартним відхиленням 1. Таким чином, для кожного вокселя (елемент об'ємного зображення, що містить значення інтенсивності) знімка було обчислено нове значення за такою формулою:

$$\frac{img\ mean}{std} \quad (3.1)$$

де *img* - значення інтенсивності для даного вокселя;

mean - середнє значення інтенсивності;

std - стандартне відхилення інтенсивностей.

Такий спосіб значно підвищує якість роботи нейронної мережі. Також для кожного знімка була обчислена маска зображення головного мозку. Маска подається на вхід нейронної мережі, щоб під час навчання мережа навчалася тільки на зображенні мозку, не враховуючи фон[22].

3.1.1 Оптимізація моделей згорнутих нейронних мереж

Використано 10 моделей згорнутих нейронних мереж – WNET, Multi-class WNET, Multi-class WNET + TTA, 3D UNet, 3D UNet + TTA, Cascade networks, Cascade networks + TTA, DEEPMEDIC (T1), DEEPMEDIC (T1+T2) експериментально визнані кращими для вирішення завдання сегментації пухлини головного мозку.

В Tensorflow [23] реалізовано 3D UNet [24], багаторівневі WNet і каскадні мережі [25] з використанням NiftyNet [26]. Адаптивна оцінка моменту використовувалася для навчання (ADAM) [27] з початковою швидкістю навчання 10^{-3} , вага розпад 10^{-7} і максимальну ітерацію 20k. Розмір патча для підготовки становив $96 \times 96 \times 96$ для 3D UNet, і $96 \times 96 \times 19$ для багаторівневих WNet. Розмір групи становив 2 і 4 для цих двох мереж відповідно. Для каскадних мереж наслідували стандартні конфігурації [25]. Навчальний процес був реалізований на GeForce RTX 2080. Для навчання використовувалася функція THE DICE LOSS [28].

Для нейронної мережі WNET [29] використовувалася стандартна конфігурація і параметри, описані в оригінальній статті [30]. Модель була навчена на наборі даних з використанням всіх чотирьох режимів для кожного пацієнта. В оригінальну конфігурацію DEEPMEDIC [31] були внесені наступні зміни в архітектуру мережі. Як функції активації була обрана функція SELU[32] замість PReLU. Це дозволило прискорити процес навчання даної мережі, а так само поліпшити якість розпізнавання пухлин. Також, для DEEPMEDIC було навчено дві моделі. Перша версія моделі була навчена тільки на знімках в режимах T1 і T2. Даний вибір обумовлений тим, що T1 і T2 є найбільш поширеними режимами для МРТ. Такі режими, як Flair і T1Gd використовуються не у всіх клініках і лікарнях, тому використання тільки двох режимів робить модель більш універсальною для подальшого використання. Друга версія DEEPMEDIC була навчена виключно з використанням T1 знімків. Дана модель була навчена, щоб вивчити поведінку моделі в тому випадку, коли на

вході є тільки один режим МРТ. Для всіх трьох мереж вихідний набір даних був розділений на 3 частини: навчання (близько 70%), валідація (20%) і тестова множина (10%).

В наступних дослідках плануються провести інші експерименти з варіаціями нейронних мереж, наприклад, зі збільшенням числа епох і розміру навчальної вибірки (для чого необхідно значно більше пам'яті), або з застосуванням RESNET-блоків замість згорнутих нейронних мереж. Іншими перспективними експериментами може бути багатократний стекінг мереж, який ефективно зарекомендував себе в сегментації зображень.

3.2 Класифікація та кластеризація

Після того як ми сегментували наші зображення потрібно класифікувати їх. Так як у нас можуть бути різні хвороби одних і тих самих органів то в нас буде декілька класів куди наша нейронна мережа може віднести конкретне зображення. Для того щоб провести класифікацію потрібно знати яка кількість класів в нас є, кількість класів буде відповідати кількості хвороб які можуть бути в пацієнта та діагностуватись по медичним зображенням.

Класифікація – це система розподілення об'єктів по групах відповідно до наперед визначених ознак. Тобто ми заздалегідь знаємо кількість класів (хвороб) які можуть бути виявлені за допомогою медичного зображення. Деякі хвороби можуть мати різний стан протікання, та для нашої нейронної мережі це буде виглядати як різні хвороби[33].

Для визначення кількості хвороб на наборі даних ми можемо використувати розмічені дані. Але як зазначалося раніше, розмітка даних потребує часу спеціаліста, тому є надто дорогою. Звичайно навчати нейронну мережу на розмічених даних добре, вона відразу буде знати які ознаки притаманні конкретним хворобам. Але частіше за все в нас будуть не розмічені медичні зображення та потрібно буде самим визначати кількість хвороб.

Тож коли ми не знаємо точну кількість хвороб ми можемо обробити наші зображення за допомогою кластеризації.

Кластеризація – це процес розбиття заданої вибірки об'єктів (в нашому випадку медичних зображень) на кластери. Об'єкти одного кластера може схожі між собою, а об'єкти різних кластерів значно відрізняються. Таким чином ми можемо відразу зрозуміти які пацієнти мають хворобу, а які ні, та визначити скільки кластерів, тобто хвороб присутнє в даному датасеті.

Класифікація відноситься до навчання з вчителем, тобто ми заздалегідь знаємо як виглядає об'єкт який ми відносимо до певного класу. Кластеризація відноситься до навчання без вчителя, коли ми не знаємо як будуть виглядати об'єкти які належать різним кластерам, тому нейронна мережа сама визначить схожість об'єктів та розподілить їх між кластерами.

Кластеризація має декілька проблем які призводять до неоднозначності результатів. Не існує однозначно кращого критерію якості кластеризації. Існує ряд критеріїв та алгоритмів що не мають чітко вираженого критерію, але здійснюють достатньо якісну кластеризацію. Застосування однакового алгоритму для різних даних може приводити до різних результатів кластеризації. Також, як правило, заздалегідь невідомо число кластерів і алгоритм сам підбирає оптимальне число кластерів, на які будуть поділятися об'єкти.

Результат кластеризації значною мірою залежить від метрики, вибір якої лягає на експерта.

Кластери – це групи об'єктів об'єднані за якоюсь ознакою чи ознаками. Після кластеризації головною задачею є правильна інтерпретація. Нейронна мережа може неправильно розподілити наші зображення, наприклад, Різні стадії хвороби рознести по різних кластерам, а в завданні було вказано щоб ми визначили два кластери, коли пацієнт здоровий та хворий.

Якщо в нас більше ніж 2 кластери, то виникає питання стійкості кластерів. Якщо віднесення наших зображень до різних відбулося добре, тоді при

невеликій зміні параметрів кластеризації, об'єкти не змінять свій кластер. Теорема неможливості Клейнберга показує що не існує однозначно найкращого алгоритму кластеризації[34].

3.3 Ансамблювання нейронних мереж

Для того щоб наші нейронні мережі працювали злагоджено потрібно щоб вихід однієї нейронної мережі (яка сегментує зображення) подавався на вхід іншої нейронної мережі(яка ці зображення класифікує).

Згорткові нейронні мережі продемонстрували вражаючі показники в класифікації зображень. При налагодженні складних моделей з невизначеними цілями для навчання мережі, все залежить від моделі стохастичної процедури навчання, тобто кінцеву мережу тренують з градієнтом на спуск, що обумовлено такими факторами як порядок даних у кожній епосі, ініціалізація, швидкість навчання тощо. На рисунку 3.4 приведена загальна схема роботи алгоритму.

Навчання ансамблю - це метод генерації декількох версій мереж для прогнозу і використання їх для отримання агрегованого прогнозу.

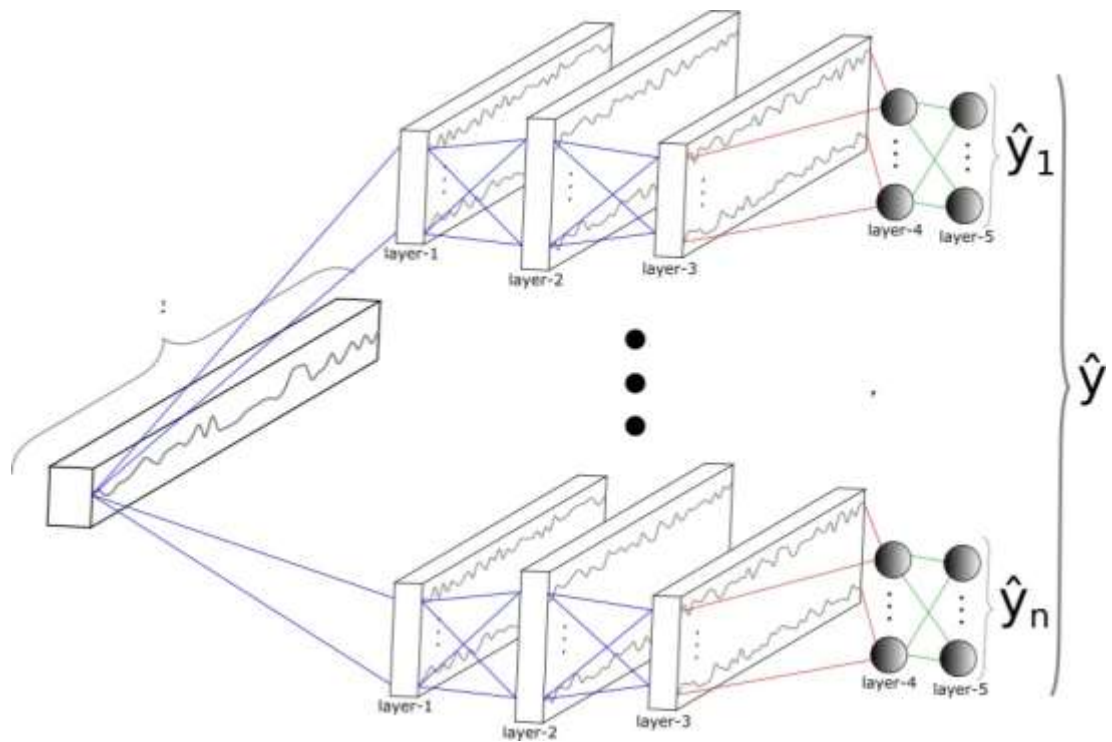


Рисунок 3.4 – Загальна схема роботи алгоритму

Протягом останніх років досліді показали, що ансамблевий метод може забезпечити відмінну ефективність у зниженні помилки тестування. Більшість методів показали, що на базі класифікації ImageNet 2012, їх ансамблева модель з п'яти класифікаторів досягла високого рівня: коефіцієнт помилок – 38,1%, у порівнянні з найвищим коефіцієнтом помилок - 40,7%, наданим єдиною моделлю. Крім того інші методи показали, що ансамбль з 6 класифікаторів може зменшити верхню помилку з 40,5% до 36,0%. У 1994 році Брейман представив концепцію мішків, що допомогло зрозуміти те, чому ансамбль дерев класифікації і дерев регресії працюють, коли вони навчалися випадковими вибірками з цілого набору даних. Проте досі немає чіткого розуміння, чому ансамбль згорткових нейронних мереж виконує свою задачу так добре, які відношення між кількістю моделей в ансамблі та кількістю зменшених помилок або інше [35].

З [35] відомо, що чим більше предикторів залучаються в ансамбль, тим менше помилка. Однак, оскільки Q – кількість усереднених моделей переходить до нескінченності, гранична кількість помилок прямує до 0. Це узгоджується з поведінкою графіка помилки на рис. 3.5.

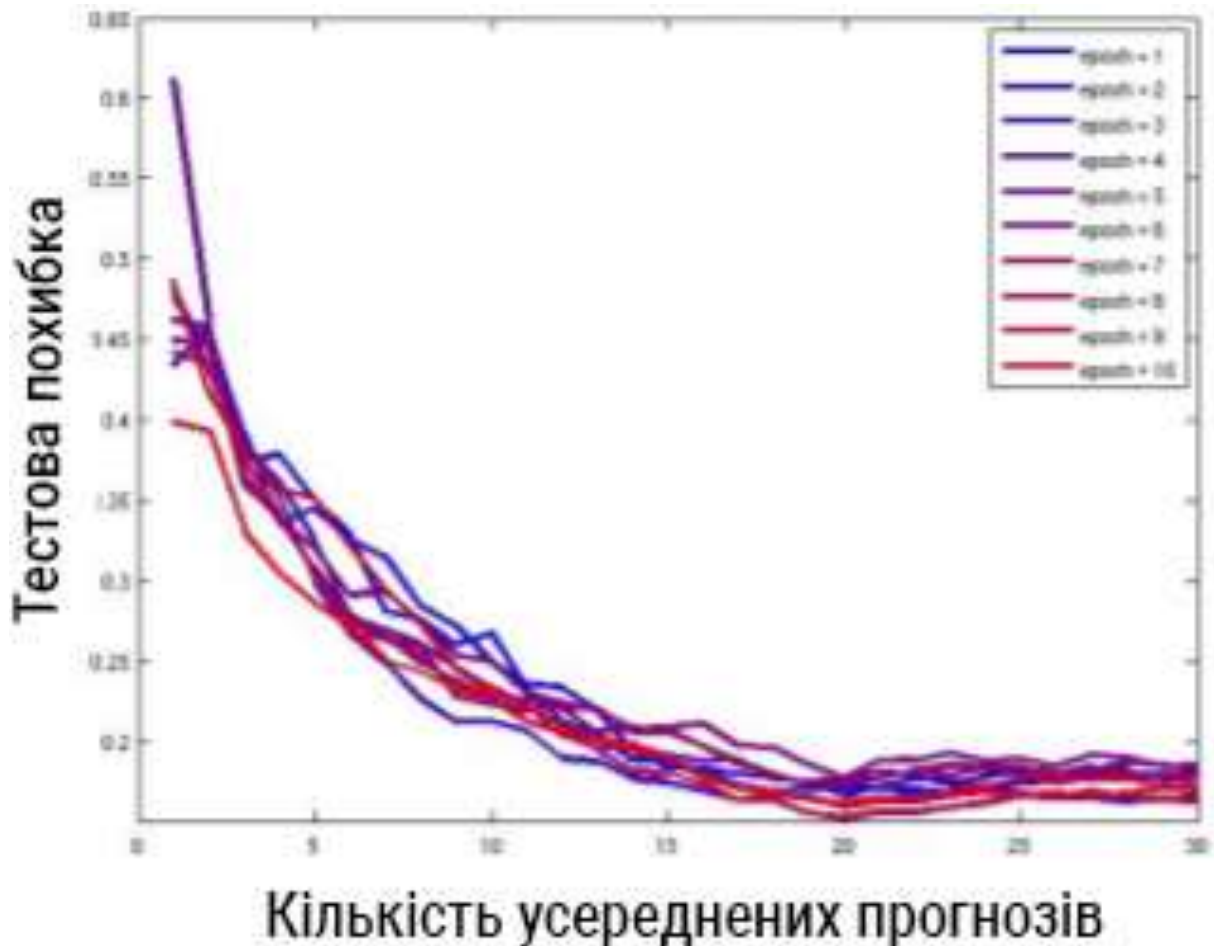


Рис. 3.5 – Тестова похибка згорткових нейронних мереж при прогнозуванні з різною кількістю усереднених прогнозів

Ансамбль - це потужна процедура, яка покращує ефективність однієї мережі.

Як можна бачити на рисунку 3.6 при збільшенні епох навчання тестова похибка поступово зменшується, тож в цьому випадку буде краще збільшити кількість епох, щоб досягти меншої похибки.

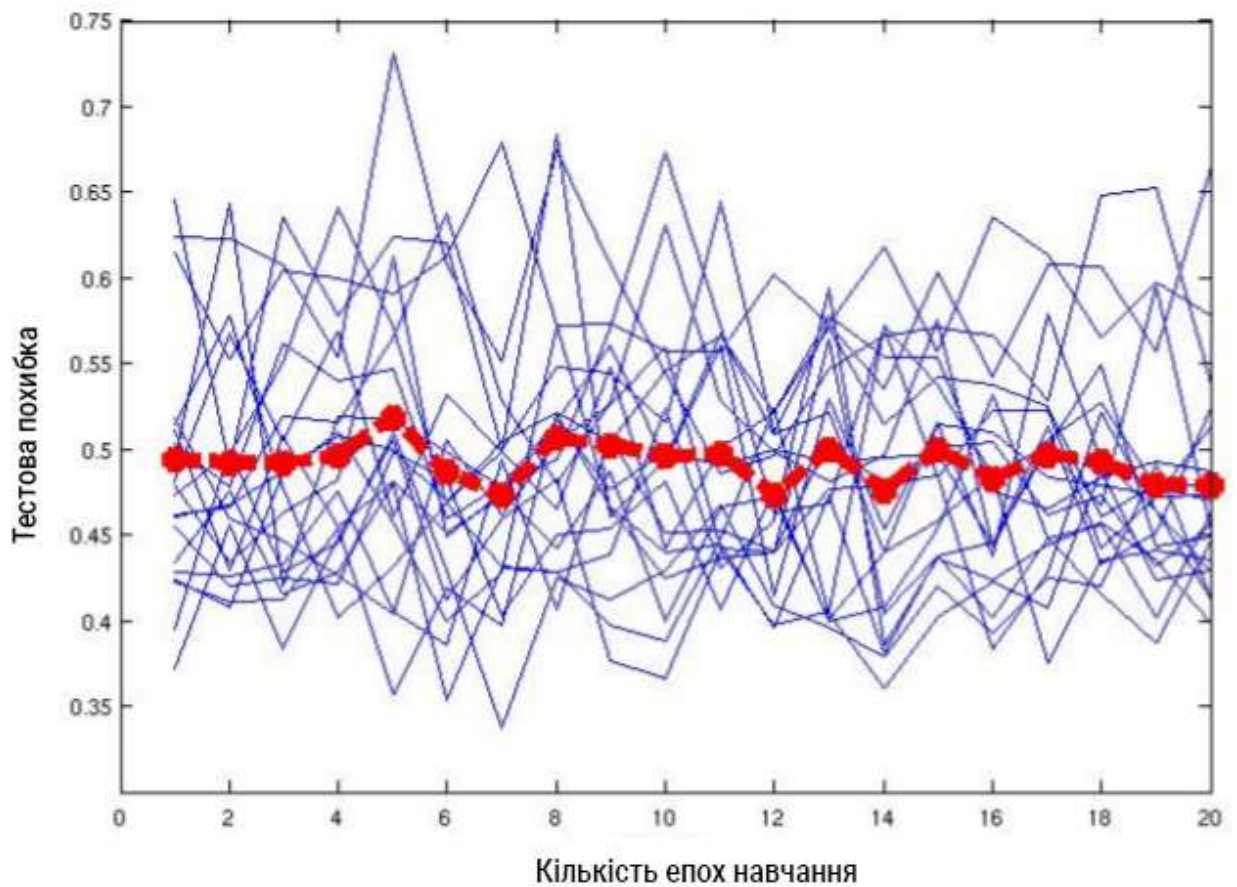


Рис. 3.13 – Тестова похибка згорткових нейронних мереж при прогнозуванні з різною епохою навчання

У даній роботі проєкспериментувано з різними методами ансамблю, які, як правило, сприяють різкому скороченню помилок. Окрім того, має місце думка про зв'язки між кількістю моделей та їх складністю. Існують підтвердження того, що навчання може привести до виграшу точності разом із скороченням часу навчання [35].

ВИСНОВКИ

В результаті виконання роботи було створено нейронну мережу для аналізу медичних зображень, яка допомагає лікарям у встановленні правильного діагнозу.

В рамках даної кваліфікаційної роботи було досліджено використання штучних нейронних мереж в задачі діагностики респіраторних захворювань за допомогою аналізу рентгенівських знімків, що передбачає можливість за рентгенівським знімком пацієнта автоматично визначити відсутність патологій або наявність на ньому вірусної або бактеріальної пневмонії. Було проведено попередній аналіз предметної області, а також огляд актуальності і сучасного стану проблематики респіраторних захворювань як і у світі загалом, так і в Україні. Були розглянуті вже існуючі системи націлені на вирішення подібних задач. У ході проведення теоретичних досліджень були розглянуті загальні принципи роботи штучних нейронних мереж і методи їх навчання, їх можливості щодо вирішення поставленої задачі класифікації зображень. Після загального огляду різноманітних типів штучних нейронних мереж було прийняте рішення про використання згорткових нейронних мереж як найбільш відповідних і придатних до вирішення поставленої задачі. Був здійснений огляд згорткових нейронних мереж, їх особливостей і принципів роботи. В ході виконання роботи була досліджена предметна галузь розробки системи підтримки прийняття рішень та аналіз медичних зображень для встановлення діагнозу. Відповідно було досліджено галузь аналізу зображень, нейронні мережі, машинне та глибоке навчання, згорткові нейронні мережі.

Досліджено методи глибокого навчання нейронних мереж. Проведено порівняльний аналіз методів навчання нейронних мереж та зроблено висновки який метод краще застосувати в нашій системі.

Для покращення результативності системи можна додати систему щоб вона враховувала додаткові дані про пацієнта(вік, стать, самопочуття, тощо).

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. <https://www.ocutri.com/medical/> (дата звернення 21.09.2021)
2. Хвостиков А. В “МАТЕМАТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ И АНАЛИЗА СЛАБОТЕКСТУРИРОВАННЫХ МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ” – 2019
3. <https://www.rscf.ru/news/math/uchenye-s-pomoshchyu-neyrosetey-nauchilis-prognozirovat-bolezn-altsgeymera-s-tochnostyu-v-90/> (дата звернення 27.09.2021)
4. Конспект лекцій з дисципліни «Інноваційні технології» для студентів для студентів кваліфікаційного рівня магістр спеціалізації «**Методи та технології обробки зображень в біомедицині**» за спеціальністю 122 Комп’ютерні науки та інформаційні технології / Упоряд. О.Я. Кузьомін. – Харків: ХНУРЕ, 2020. – 160 с
5. <https://ru.wikipedia.org/wiki/Python> (дата звернення 27.09.2021)
6. https://uk.wikipedia.org/wiki/Штучна_нейронна_мережа (дата звернення 27.09.2021)
7. https://uk.wikipedia.org/wiki/Згорткова_нейронна_мережа (дата звернення 27.09.2021)
8. <https://uk.wikipedia.org/wiki/TensorFlow> (дата звернення 05.10.2021)
9. https://uk.wikipedia.org/wiki/Метод_зворотного_поширення_помилки (дата звернення 05.10.2021)
10. https://uk.wikipedia.org/wiki/Передавальна_функція_штучного_нейрона (дата звернення 09.10.2021)
11. https://ru.wikipedia.org/wiki/Стохастический_градиентный_спуск (дата звернення 09.10.2021)
12. https://uk.wikipedia.org/wiki/Функция_врат (дата звернення 09.10.2021)

13. <https://neerc.ifmo.ru/wiki/index.php?title=Переобучение> (дата звер-
нення 29.10.2021)
14. https://uk.wikipedia.org/wiki/Сегментація_зображення (дата звер-
нення 29.10.2021)
15. U-Net: нейросеть для сегментации изображений [Электронный ре-
сурс]. – 2018. – Режим доступа до ресурсу: [https://neurohive.io/ru/vidy-
nejrosetej/u-net-image-segmentation/](https://neurohive.io/ru/vidy-nejrosetej/u-net-image-segmentation/).
16. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolu-
tional networks for biomedical image segmentation. In International Conference on
Medical Image Computing and Computer-Assisted Intervention, pages 234–241.
Springer, 2015.
17. Ozg'un ,Ci,cek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox,
and Olaf " Ronneberger. 3d u-net: learning dense volumetric segmentation from
sparse annotation. In International Conference on Medical Image Computing and
Computer-Assisted Intervention, pages 424–432. Springer, 2016.
18. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual
learning for image recognition. In Proceedings of the IEEE conference on computer
vision and pattern recognition, pages 770–778, 2016.
19. Menze B.H., Jakab A., Bauer S., Kalpathy-Cramer J., Farahani K., Kirby
J., Burren Y., Porz N., Slotboom J., Wiest R., Lanczi L., Gerstner E., Weber M.A.,
Arbel T., Avants B.B., Ayache N., Buendia P., Collins D.L., Cordier N., Corso J.J.,
Criminisi A., Das T., Delingette H., Demiralp C , , Durst C.R., Dojat M., Doyle S.,
Festa J., Forbes F., Geremia E., Glocker B., Golland P., Guo X., Hamamci A.,
Iftekharuddin K.M., Jena R., John N.M., Konukoglu E., Lashkari D., Mariz J.A.,
Meier R., Pereira S., Precup D., Price S.J., Raviv T.R., Reza S.M., Ryan M.,
Sarikaya D., Schwartz L., Shin H.C., Shotton J., Silva C.A., Sousa N., Subbanna
N.K., Szekely G., Taylor T.J., Thomas O.M., Tustison N.J., Unal G., Vasseur F.,
Wintermark M., Ye D.H., Zhao L., Zhao B., Zikic D., Prastawa M., Reyes M., Van
Leemput K. The multimodal brain tumor image segmentation benchmark (BRATS).
// TMI 34(10)— 1993–2024 (2015).

20. Bakas S et al. Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. // *Nature Scientific Data*.— 2017.

21. Jarrett K., Kavukcuoglu K., Ranzato M., LeCun Y. What is the best multi-stage architecture for object recognition? // 2009 IEEE 12th International Conference on Computer Vision.— 2009. — pp. 2146–2153.

22. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X., Brain, G.: TensorFlow: A system for large-scale machine learning. In: *USENIX Symposium on Operating Systems Design and Implementation*. pp. 265–284 (2016).

23. Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 424–432 (2016).

24. Wang, G., Li, W., Ourselin, S., Vercauteren, T.: Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. In: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 178–190. Springer International Publishing (2018).

25. Gibson, E., Li, W., Sudre, C., Fidon, L., Shakir, D.I., Wang, G., Eaton-Rosen, Z., Gray, R., Doel, T., Hu, Y., Whyntie, T., Nachev, P., Modat, M., Barratt, D.C., Ourselin, S., Cardoso, M.J., Vercauteren, T.: NiftyNet: A deep-learning platform for medical imaging. *Computer Methods and Programs in Biomedicine* 158, pp. 113–122 (2018).

26. Kingma, D.P., Ba, J.L.: Adam: A method for stochastic optimization. In: *International Conference on Learning Representations* (2015).

27. Milletari, F., Navab, N., Ahmadi, S.A.: V-Net: Fully convolutional neural networks for volumetric medical image segmentation. In: *International Conference on 3D Vision*. pp. 565–571 (2016).

28. Kamnitsas K., Ledig C., Newcombe V.F.J., Simpson J.P., Kane A.D., Menon D.K., Rueckert D., Glocker B. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. // Medical Image Analysis. – 2017. — Vol. 36. — pp. 61–78 .

29. Wang G., Li W., Ourselin S., Vercauteren T. Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. // Proc. Multi- modal Brain Tumor Segmentation (BRATS) Challenge 2017 - MICCAI workshop.— arXiv:1709.00382. — 2017.

30. Günter Klambauer, Thomas Unterthiner, Andreas Mayr, Sepp Hochreiter. Self-Normalizing Neural Networks. // arXiv:1706.02515. — 2017.

31. P. K. Sahoo, S. Soltani, A. K. C. Wong, and Y. C. Chen, "A survey of thresholding techniques", CVGIP, vol. 41, no. 2, pp. 233-260, Feb.1988.

32. W. K. Pratt, "Digital Image Processing", New York: Wiley, 1991, pp. 597-625

33. <https://uk.wikipedia.org/wiki/Класифікація> (дата звернення 19.11.2021)

34. https://neerc.ifmo.ru/wiki/index.php?title=Классификация#.D0.A2.D0.B5.D0.BE.D1.80.D0.B5.D0.BC.D0.B0_.D0.BD.D0.B5.D0.B2.D0.BE.D0.B7.D0.BC.D0.BE.D0.B6.D0.BD.D0.BE.D1.81.D1.82.D0.B8_.D0.9A.D0.BB.D0.B5.D0.B9.D0.BD.D0.B1.D0.B5.D1.80.D0.B3.D0.B0 (дата звернення 19.11.2021)

35. Liran C. Learning Ensembles of Convolutional Neural Networks / Liran Chen – Chicago, 2016. – 9 с.