

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ВОЛИНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ЛЕСІ УКРАЇНКИ
Кафедра комп'ютерних наук та кібербезпеки

На правах рукопису

ЮЗВА АННА СЕРГІЇВНА
**НЕЙРОМЕРЕЖЕВІ ТЕХНОЛОГІЇ В МЕДИЧНІЙ ДІАГНОСТИЦІ
ЛЕГЕНЕВИХ ЗАХВОРЮВАНЬ**

Спеціальність: 122 Комп'ютерні науки

Освітньо-професійна програма: Комп'ютерні науки та інформаційні технології
Кваліфікаційна робота на здобуття освітнього ступеня «магістр»

Науковий керівник:
СОБЧУК ВАЛЕНТИН
ВОЛОДИМИРОВИЧ
д.т.н., професор, професор кафедри
комп'ютерних наук та кібербезпеки

РЕКОМЕНДОВАНО ДО ЗАХИСТУ
Протокол № _____
засідання кафедри комп'ютерних наук
та кібербезпеки
від _____ 2024 р.
Завідувач кафедри

(_____) Гришанович Т. О.

Луцьк 2024

ЗМІСТ

ВСТУП	4
РОЗДІЛ 1. Теоретичні основи нейромережових технологій та інструменти медичної діагностики захворювань легень	8
1.1 Теоретичні основи нейромережових технологій	8
1.1.1 Визначення та класифікація нейромереж.....	8
1.1.2 Огляд сучасних методів машинного навчання та нейронних мереж у сфері обробки зображень.....	13
1.1.3 Порівняння ефективності різних нейронних архітектур	16
1.2 Інструменти медичної діагностики легневих захворювань	19
1.2.1 Методи діагностики	19
1.2.2 Методи підготовки наборів даних.....	24
1.2.3 Результат дослідження наборів даних	28
1.3 Методи діагностики за допомогою нейромережових технологій.....	32
1.3.1 Алгоритми згорткових нейронних мереж для аналізу зображень	32
1.3.2 Методи попередньої обробки знімків.....	46
1.3.3 Використання технік навчання та перенавчання нейромереж для медичних задач.....	55
1.3.4 Методи оцінки точності роботи нейромереж на медичних знімках.....	59
РОЗДІЛ 2. Методи діагностики легневих захворювань нейромережовими технологіями та підходи до створення інструментів її автоматизації.....	65
2.1. Постановка задачі, призначення та вимоги.....	65
2.2. Методологія дослідження	66

2.3 Обґрунтування вибору інструментальних засобів	79
2.4 Етапи програмної реалізації.....	83
2.5 Аналіз отриманих результатів дослідження, рекомендації використання та впровадження	86
2.5.3 Аналіз отриманих результатів дослідження	86
2.5.2 Рекомендації використання та впровадження	88
ВИСНОВКИ.....	90
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	91
ДОДАТОК А.....	93
ДОДАТОК Б	97

ВСТУП

Актуальність теми. Сучасний розвиток та досягнення нейромережових технологій дозволяє науковцям та науковицям знаходити практичне застосування для них майже у всіх сферах життя та діяльності. Застосування нейромереж у медицині відкриває нові можливості для оптимізації роботи лікарів та підвищенні ефективності медичних послуг. У сучасному житті під час медичної діагностики використовується безліч знімків, таких як МРТ, КТ, рентгенографія тощо, вивчення та опрацювання яких займає у фахівців багато часу, який вони можуть використовувати доцільніше. Нейромережові технології можуть значно підвищити швидкість та точність опрацювання цих медичних досліджень. Програми, засновані на нейромережових алгоритмах, здатні не тільки знижувати ризик людської помилки, але й істотно зменшувати навантаження на лікарів, звільняючи час для більш складних і важливих задач.

У цей складний для України час особливо важливо пришвидшити та покращити надання медичних послуг, адже зростає потреба в ефективній та швидкій допомозі. Відсутність єдиної національної медичної бази даних створює значні труднощі в обміні інформацією між медичними закладами, що ускладнює доступ до історії хвороб та результатів обстежень пацієнтів. Через це пацієнти часто змушені повторно проходити обстеження або не можуть отримати необхідну допомогу оперативно, що збільшує ризик виникнення ускладнень і ставить під загрозу їх здоров'я. Більшість медичної документації пацієнтів зберігається локально в межах одного медичного закладу або навіть у паперовому вигляді на руках у лікаря, що не тільки уповільнює процес надання медичних послуг, але й підвищує ризик втрати чи помилкового використання інформації. У таких умовах впровадження єдиної медичної бази, доступної для лікарів у будь-якій точці країни, могло б значно покращити координацію між закладами охорони здоров'я та забезпечити безперервність надання медичної допомоги.

Мета дослідження. Дослідити ефективність нейромережових моделей для автоматизованої діагностики у пульмонології; створити інструмент підтримки

прийняття рішень для лікарів, який би підвищив якість діагностики легеневих захворювань та пришвидшив процес обстеження пацієнтів.

Завдання дослідження:

- провести детальний огляд сучасних нейромережових моделей, що використовуються в медичній діагностиці;
- визначити переваги та недоліки існуючих підходів;
- вибрати оптимальну нейромережову архітектуру для вирішення поставленої задачі;
- створити інтерфейс інструменту підтримки прийняття рішень для діагностики легеневих захворювань.

Для виконання завдань дослідження вирішити такі задачі:

- проаналізувати сучасні нейромережові технології для медичної діагностики, зокрема для обробки та інтерпретації медичних знімків, таких як рентген, КТ та МРТ;
- оглянути методи обробки зображень, включаючи сегментацію, класифікацію та виявлення патологій, з метою вибору найбільш ефективних підходів для медичних знімків;
- дослідити алгоритми машинного навчання та нейромереж, що застосовуються для аналізу медичних зображень, таких як згорткові нейронні мережі (CNN) та їх модифікації;
- оцінити ефективність існуючих моделей нейронних мереж для різних типів медичних знімків та аналіз їх результатів у контексті медичної діагностики;
- порівняти методи аналізу зображень за точністю, швидкістю та надійністю виявлення патологій у клінічних умовах;
- розробити рекомендації для медичних фахівців щодо використання нейромережових технологій для покращення якості та швидкості діагностичних рішень.

Об'єктом дослідження є процес медичної діагностики захворювань легень за допомогою нейромережових моделей.

Предметом дослідження використання нейромережових технологій для автоматизації підтримки прийняття рішень та підвищення точності діагностики легеневих захворювань за даними медичних зображень.

Практичне застосування отриманих результатів. Практичне застосування результатів дослідження полягає в удосконаленні медичної діагностики через впровадження нейромережових технологій для автоматизованого аналізу та діагностики медичних знімків та іншого, що знижує навантаження на лікарів та підвищує точність діагностування. Це сприятиме швидшому виявленню патологій, покращить якість медичних послуг та дозволить інтегрувати такі технології в системи телемедицини, особливо в віддалених регіонах. Результати також можуть бути використані для підготовки фахівців у медичній інформатиці та діагностиці.

Дослідження охоплює аналіз сучасних підходів до використання нейромережових технологій у медичній діагностиці, їх можливості та обмеження, а також, на основі виявлених проблем та потреб у медичній практиці, розробку методик для діагностики патологій на рентгенівських знімках. Для навчання нейромережі обрано набори даних PadChest, COVID-19 Radiography Database та MIMIC-CXR Dataset.

У рамках роботи буде проведено аналіз існуючих підходів до діагностики за допомогою нейромереж, визначені переваги та недоліки кожного з них. Будуть розглянуті сучасні нейромережові технології та їх можливості для автоматизації діагностики. На основі цього дослідження буде розроблена та протестована методика класифікації патологій на рентгенівських знімках, що враховує специфіку медичних даних та можливості нейромережових технологій. Результати дослідження можуть бути використані для подальшого вдосконалення систем медичної діагностики.

1. Лаптев, О., Юзва, А. 2024. Дослідження використання технік моделювання для аналізу відгуків клієнтів. Прикладні проблеми комп'ютерних наук, безпеки та математики. 2 (Бер 2024), 4–17. [13]
2. Лаптев О. А., Юзва А. С. (2024). Штучний інтелект: можливості та виклики для суспільства // Тези доповідей XIII Міжнар. наук.-практ. конф. «Математика. Інформаційні технології. Освіта.», 31 травня – 2 червня 2024 р. Луцьк-Світязь: СНУ імені Лесі Українки, 2024. С. 133-135. [10]
3. Омельчук А., Юзва А., Булатецька Л. (2024). Методи високоефективної обробки даних в Oracle за допомогою секціонування таблиць та індексів // Тези доповідей I Міжнар. наук.-практ. конф. «Проблеми комп'ютерних наук, програмного моделювання та безпеки цифрових систем.», 13–16 червня 2024 р. Луцьк–Світязь: табір "Гарт", 2024. [11]
4. Юзва А. С., Собчук В. (2024). В. Автоматизація процесів медичної діагностики із застосуванням нейромереж // Тези доповідей V Міжнар. наук.-практ. конф. «Кібербезпека в сучасному світі: актуальні виклики.» , 22 листопада 2024 р., м. Одеса, Україна. 2024. [12]

РОЗДІЛ 1.

ТЕОРЕТИЧНІ ОСНОВИ НЕЙРОМЕРЕЖЕВИХ ТЕХНОЛОГІЙ ТА ІНСТРУМЕНТИ МЕДИЧНОЇ ДІАГНОСТИКИ ЗАХВОРЮВАНЬ ЛЕГЕНЬ

1.1 Теоретичні основи нейромережевих технологій

1.1.1 Визначення та класифікація нейромереж

Нейронні мережі – це обчислювальні моделі, засновані на принципах біологічних нейронних мереж мозку. Вони використовуються для вирішення завдань штучного інтелекту (ШІ) в різних сферах, таких як розпізнавання образів, класифікація, прогнозування та автоматичне навчання. Основна ідея нейронних мереж полягає в тому, щоб навчити нейронну мережу на основі даних і адаптувати її внутрішні параметри для досягнення певної мети.

Нейронні мережі є частиною галузі штучного інтелекту, яка спрямована на моделювання аналітичних механізмів, що виконуються людським мозком. Завдання, які зазвичай вирішують нейронні мережі, – це класифікація, прогнозування та розпізнавання. Нейронні мережі здатні самостійно навчатися і розвиватися, роблячи висновки з помилок і накопичуючи досвід.

Нейронні мережі – це масиви нейронів, з'єднаних синапсами. Структура нейронних мереж прийшла безпосередньо з біології у світ програмування. Завдяки цій структурі машини можуть отримати здатність аналізувати і навіть зберігати різноманітну інформацію. Нейронні мережі також можуть аналізувати вхідну інформацію та реконструювати її з власної пам'яті.

Нейронні мережі можна класифікувати за кількома критеріями, кожен з яких відображає різні аспекти архітектури, навчання та реалізації. Є декілька основних підходів до класифікації: кількість шарів, за типом зв'язку між нейронами, за типом навчання, за архітектурою, за застосуванням.

За кількість шарів є одношарові нейронні мережі, коли лише один шар між вхідними та вихідними нейронами. Багатошаровий перцептрон (MLP), коли один або більше прихованих шарів між входом і виходом. Багатошарові мережі можуть моделювати більш складні функції завдяки більшій кількості шарів.

За типом зв'язку між нейронами є нейронна мережа прямого поширення (Feed Forward Neural Network, FNN), коли інформація передається тільки вперед – від вхідного шару до вихідного без зворотного зв'язку. Це найпростіша і найпоширеніша архітектура. За типом зв'язку між нейронами є ще рекурентна нейронна мережа (Recurrent Neural Network, RNN), яка включає зворотний зв'язок, враховує попередні стани і може обробляти послідовності даних (текст, часові послідовності тощо).

За типом навчання виділяють навчання з учителем, навчання без учителя та навчання з підкріпленням. Навчання з допомогою вчителя (Teacher Assisted Learning, TAL) полягає в тому, що нейромережа навчається на основі підготовлених даних із відповідями, її завдання — навчитися прогнозувати правильні виходи для нових даних. Навчання без учителя передбачає, що нейромережа аналізує вхідні дані й намагається виявити структури та закономірності без правильної відповіді. Навчання з підкріпленням базується на взаємодії нейромережі з навколишнім середовищем, у ході якої вона отримує винагороду за кожну правильну дію й намагається максимізувати цю винагороду.

За архітектурою виділяють згорткові нейронні мережі (CNN), рекурентні нейронні мережі (RNN), довга короткочасна пам'ять (LSTM), трансформери, автокодери, генеративні змагальні мережі (GANs).

Згорткові нейронні мережі (CNN) спеціалізуються на обробці даних із зображень та інших 2D-структур. Використовують згорткові шари для вилучення ознак. Рекурентні нейронні мережі (RNN) можуть зберігати внутрішні стани (історію попередніх даних), що робить їх корисними для завдань, пов'язаних з безперервними даними (обробка аудіо та відео). Довга короткочасна пам'ять (LSTM) це покращена версія RNN, яка краще обробляє довготривалі

залежності в послідовностях і зменшує проблему «градієнтних спотворень». Трансформери це архітектури, що використовують механізми самофокусування (Рис.1.1) для паралельної обробки великих масивів даних. Широко використовуються в задачах обробки природної мови (NLP) та обробки текстів. Автокодери це спеціалізовані нейронні мережі для зменшення розмірності та навчання ефективного представлення даних. Генеративні змагальні мережі (GANs) складаються з двох мереж, генератора і дискримінатора. Використовуються для генерації нових даних (наприклад, зображень, текстів), подібних до даних у навчальній вибірці.

За застосуванням нейронні мережі використовують для класифікації (нейронні мережі, які виконують завдання класифікації даних на основі вхідних ознак) для регресії (використовується для прогнозування безперервних значень), для сегментації (у комп'ютерному зорі, наприклад, коли зображення потрібно розділити на кілька класів або визначити межі об'єктів), для генерації даних (мережі, які генерують нові дані на основі наявної інформації, наприклад, для створення зображень або тексту).

Інші типи нейронних мереж:

- моделі Больцмана, які використовуються для статистичного навчання та оптимізації;
- спайкові нейронні мережі, що моделюють біологічні нейрони та їхню активність на основі спайкових сигналів;
- моделі Хопфілда, рекурентні нейронні мережі, що використовуються для асоціативної пам'яті.

Ці класифікації ілюструють різні типи нейронних мереж та їх використання для конкретних завдань і типів даних.

Крім базового підходу до класифікації нейронних мереж, існують інші типи класифікацій, які дозволяють більш детально вивчати нейронні мережі залежно від їх характеристик, контексту застосування та архітектури.

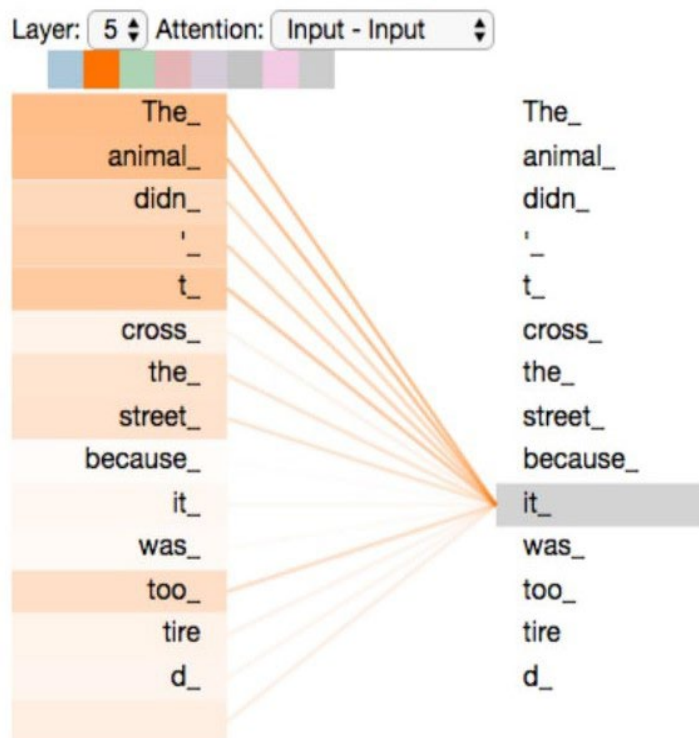


Рисунок 1.1 – Механізм самофокусування на прикладі роботи з текстом

Додаткові критерії класифікації нейронних мереж включають кілька важливих аспектів. За типом переданої інформації або вхідних даних нейронні мережі можуть бути орієнтовані на обробку табличних даних, таких як числові або категоріальні дані в табличній формі (наприклад, дані з баз даних або електронних таблиць Excel). Мережі, створені для роботи із зображеннями, спеціалізуються на розпізнаванні та обробці зображень і часто називаються згортковими нейронними мережами (CNN). Текстові мережі використовуються для обробки текстових даних, наприклад, у завданнях природної мови, за допомогою моделей типу RNN або трансформерів. Мовленнєві та аудіо мережі призначені для обробки аудіо даних, таких як мова або музика, використовуючи RNN чи LSTM. Мережі часових рядів забезпечують прогнозування подій на основі послідовностей часових рядів, як у випадках фінансового прогнозування.

За типом алгоритмів оптимізації та навчання нейронні мережі можуть бути орієнтовані на градієнтний спуск для мінімізації помилок, що підходить для більшості сучасних моделей, або використовувати еволюційні алгоритми для

оптимізації архітектури та ваг, наприклад, на основі генетичних алгоритмів. Байєсівські нейронні мережі застосовують байєсівський підхід для обчислення невизначеності прогнозу.

Класифікація також враховує тип активаційних функцій. Лінійні нейронні мережі працюють із лінійними функціями активації, що обмежує їхню здатність моделювати складні залежності, тоді як нелінійні моделі використовують функції типу ReLU, сигмоїдні або Tanh для кращого відображення нелінійних зв'язків.

За типом вхідних даних нейронні мережі поділяються на класифікаційні моделі, які працюють із дискретними класами, регресійні моделі, орієнтовані на безперервні величини, і генеративні моделі, здатні створювати нові дані, як у генеративних змагальних мережах (GAN).

За ступенем взаємодії з навколишнім середовищем виділяють статичні нейронні мережі, які функціонують у фіксованому середовищі, та адаптивні, здатні змінювати свої параметри в реальному часі. За енергоефективністю традиційні моделі вимагають великих обчислювальних ресурсів, тоді як енергоефективні мережі оптимізовані для мобільних пристроїв та вбудованих систем.

Масштабність моделі визначається кількістю нейронів і шарів. Малі нейронні мережі використовуються для простих завдань, тоді як глибокі мережі здатні вирішувати складні завдання завдяки великій кількості прихованих шарів.

За підходом до моделювання біологічних процесів класичні штучні нейронні мережі є спрощеними математичними моделями, тоді як спайкові нейронні мережі (SNN) краще імітують діяльність біологічних нейронів. Архітектура обробки даних може бути послідовною, як у RNN і LSTM, або паралельною, як у трансформерах, що забезпечує обробку великих обсягів інформації одночасно.

Ці класифікації дають краще розуміння різноманітності нейронних мереж, їхньої адаптивності до різних типів завдань і даних, а також того, як їх можна структурувати для досягнення різних цілей.

1.1.2 Огляд сучасних методів машинного навчання та нейронних мереж у сфері обробки зображень.

Обробка зображень за допомогою машинного навчання та нейронних мереж є одним з найактивніших напрямків досліджень і застосувань штучного інтелекту. Завдяки розвитку обчислювальних потужностей і великим обсягам даних сучасні методи дозволяють досягти високої точності в таких завданнях, як розпізнавання об'єктів, сегментація, формування зображень і покращення зображень. Нижче наведено огляд основних сучасних методів машинного навчання та нейронних мереж, що використовуються для обробки зображень.

Конволюційні нейронні мережі (Convolutional Neural Networks, CNN) є основною архітектурою для обробки зображень. Вони здатні автоматично виділяти ознаки з зображень завдяки використанню конволюційних шарів, що виконують зсуви фільтрів (масок) по зображенню для виділення локальних ознак, таких як краї, текстури та інші важливі елементи.

Наведемо основні підходи.

AlexNet (2012). Одна з перших успішних глибоких CNN, яка виграла конкурс ImageNet і показала потужність глибокого навчання для обробки зображень.

VGG (2014). Архітектура CNN з глибокими шарами, що використовує фільтри 3x3, дозволяючи отримати більшу точність шляхом збільшення кількості шарів.

ResNet (2015). Дозволяє побудову надглибоких мереж завдяки використанню механізму залишкових з'єднань, які допомагають боротися з проблемою затухання градієнта.

Inception (2015, GoogleNet). Використовує шар багаторозмірних конволюцій для покращення ефективності обчислень і точності.

Глибоке навчання для сегментації зображень. Сегментація – це розбиття зображення на кілька регіонів або пікселів, що належать до різних об'єктів. Для

цього застосовують нейронні мережі, які вивчають складну структуру зображень.

Наведемо основні підходи.

U-Net (2015). Архітектура для сегментації зображень, що містить симетричну структуру з низхідною та висхідною частинами (down-sampling та up-sampling), що дозволяє обробляти зображення високої роздільної здатності.

Fully Convolutional Networks (FCN). Мережі, що замінюють всі щільні (fully connected) шари на конволюційні, що дозволяє працювати зі змінними розмірами зображень і покращує ефективність сегментації.

Mask R-CNN (2017). Розширює Faster R-CNN і додає маски для сегментації об'єктів. Застосовується для задач, де потрібно розділити зображення на окремі об'єкти.

Генеративні змагальні мережі (Generative Adversarial Networks, GAN) це тип нейромереж, які використовуються для генерації нових зображень, схожих на наявні. Вони складаються з двох нейромереж: генератора та дискримінатора, які працюють змагально.

Наведемо основні підходи.

Pix2Pix. GAN, що перетворює одне зображення в інше, наприклад, перетворення ескізу в реалістичне зображення або чорно-білого фото в кольорове.

CycleGAN. Дозволяє перетворювати зображення між двома доменами без потреби у парних даних (наприклад, перетворення фотографій у малюнки і назад).

StyleGAN (2018, 2020). Дозволяє генерувати реалістичні обличчя та зображення, контролюючи стиль і дрібні деталі. Застосовується у творчих індустріях, для створення штучних облич і віртуальних об'єктів.

Трансформери в обробці зображень. Хоча трансформери здебільшого застосовуються в обробці природної мови, останнім часом вони успішно інтегруються у сферу комп'ютерного зору.

Наведемо основні підходи.

Vision Transformers (ViT) (2020). Перший успішний трансформер для обробки зображень, який розбиває зображення на невеликі частини (patches) і обробляє їх аналогічно до слів у тексті.

DETR (Detection Transformer). Застосовується для виявлення об'єктів на зображеннях, використовуючи механізм уваги трансформерів.

Swin Transformer. Покращує Vision Transformer, використовуючи зсувні вікна (sliding windows) для обробки локальних ділянок зображення, що знижує складність і збільшує ефективність.

Моделі для покращення якості зображень. Сучасні методи нейромереж використовуються для покращення якості зображень, таких як збільшення роздільної здатності (super-resolution), видалення шуму або артефактів.

Наведемо основні підходи.

SRCNN (Super-Resolution Convolutional Neural Network). Одна з перших CNN для збільшення роздільної здатності зображень.

ESRGAN (Enhanced Super-Resolution GAN). Покращений метод для збільшення роздільної здатності, який використовує GAN для створення реалістичних текстур.

Denoising Autoencoders. Автоенкодера для видалення шуму з зображень, які вчаться відновлювати зображення з шумових варіантів.

Класифікація та розпізнавання об'єктів – основна задача комп'ютерного зору, де потрібно визначити, що зображено на картинці або де розташовані певні об'єкти.

Наведемо основні підходи.

Faster R-CNN. Модель для виявлення об'єктів, яка значно швидша за попередні версії (R-CNN та Fast R-CNN).

YOLO (You Only Look Once). Одна з найшвидших архітектур для реального часу виявлення об'єктів, яка виконує детекцію на всьому зображенні за один прохід.

SSD (Single Shot Multibox Detector). Ще одна модель для швидкого виявлення об'єктів, яка також дозволяє проводити обробку у реальному часі.

Мультизадачне навчання (Multitask Learning) полягає в тому, щоб одна модель виконувала кілька задач одночасно. Наприклад, сегментація, класифікація і виявлення об'єктів можуть бути виконані одночасно однією моделлю.

Наведемо основні підходи.

DeepLab. Глибока неймережа для сегментації зображень, що може використовуватися разом з іншими задачами, наприклад, детекцією об'єктів.

Multi-Task CNN. Мережі, що навчаються одночасно на кількох різних задачах, оптимізуючи ресурси і досягаючи кращих результатів завдяки спільному навчанню.

Сучасні методи машинного навчання та неймережеві методи обробки зображень базуються на використанні згорткових нейронних мереж (CNN), трансформаторів, генеративних моделей та інших передових підходів. Ці методи можуть вирішувати широкий спектр завдань, від класифікації та сегментації до створення та покращення зображень. З розвитком технологій ці методи постійно вдосконалюються і пропонують нові можливості для індустрії комп'ютерного зору.

1.1.3 Порівняння ефективності різних нейронних архітектур

Порівняння продуктивності різних нейронних архітектур, таких як згорткові нейронні мережі (CNN), рекурентні нейронні мережі (RNN), перетворювачі, генеративні змагальні мережі (GAN) та автокодери, залежить від конкретного завдання і типу даних кожної моделі та архітектурних особливостей.

Згорткові нейронні мережі (CNN) найкраще підходять для роботи з зображеннями та іншими даними з просторовою структурою, такими як відео та об'ємні дані; ключовою особливістю CNN є те, що автоматичні CNN мають перевагу над повністю пов'язаними мережами з погляду ефективності обробки зображень, оскільки вони мають менше параметрів і можуть обробляти

зображення різних розмірів. Однак RNN менш ефективні при обробці послідовних або залежних від часу даних і можуть мати труднощі з моделюванням глобального контексту без додаткових механізмів.

Рекурентні нейронні мережі (RNN) спеціалізуються на обробці послідовних даних, таких як текст, мова та часові ряди. RNN використовують циклічні зв'язки для зберігання інформації про попередні стани, що дозволяє їм враховувати часовий контекст. Основною перевагою RNN є моделювання послідовних залежностей, що робить їх ідеальними для обробки природної мови та задач прогнозування часових рядів. Однак класичні RNN стикаються з проблемою градієнтного спотворення, що ускладнює навчання довгих послідовностей; вдосконалені версії, такі як LSTM (Long Short Term Memory) і GRU (Gated Recurrent Unit) розв'язали цю проблему, але RNN все ще є обчислювально інтенсивними й важко розпаралелюються, що обмежує їх ефективність при роботі з великими обсягами даних.

Спочатку розроблені для обробки природної мови, трансформери також показали високу ефективність у комп'ютерному зорі. Трансформатори використовують механізм самоуваги, який дозволяє їм обробляти всі елементи послідовності паралельно і враховувати глобальні залежності між елементами. Це робить перетворювачі масштабованими й дозволяє їм ефективно обробляти навіть великі набори даних: Перетворювачі, такі як GPT і BERT, опанували багато завдань з обробки природної мови, а датчики зору і варіанти перетворювачів, такі як DETR, успішно використовуються в комп'ютерному зорі. Однак перетворювачі вимагають значних обчислювальних ресурсів, особливо при обробці довгих послідовностей або високоякісних зображень, і вони можуть ігнорувати короткочасні залежності без додаткового налаштування.

Генеративні змагальні мережі (Generative Adversarial Networks, GAN) використовуються для створення нових зображень, покращення якості зображень та інших творчих завдань GAN складаються з двох нейронних мереж: генератора, який генерує дані, і дискримінатора, який визначає, чи є зразок справжнім або підробленим StyleGAN-подібні GAN можуть створювати

зображення дивовижної якості, які майже неможливо відрізнити від реальних, і використовуються в творчих індустріях для створення штучних облич, моделей і текстур. Основною перевагою GAN є те, що їм не потрібні марковані набори даних, а також здатність генерувати нові дані в реалістичній манері. Однак GAN мають обмежену універсальність через необхідність балансувати між генераторами і дискримінаторами, що може зробити навчання нестабільним, а також через можливість навчання дуже специфічних шаблонів.

Автокодери використовуються для стиснення даних, зменшення шуму і генерації нових даних. Автокодер складається з двох частин: кодера, який зменшує розмір даних, і декодера, який відновлює вихідні дані. Автокодери корисні для відновлення зображень і зменшення розміру, оскільки вони можуть ефективно відновлювати зображення з пошкоджених версій або зменшувати розмірність даних. Їх також можна використовувати для створення векторних представлень даних, які можуть покращити продуктивність інших алгоритмів машинного навчання. Однак при виконанні складних генеративних завдань автокодери можуть бути менш ефективними, ніж інші архітектури.

Загалом, ефективність різних нейронних архітектур залежить від деталей завдання і типу даних. Згорткові нейронні мережі для обробки зображень, рекурентні нейронні мережі для послідовних даних, перетворювачі для завдань, що вимагають глобального контексту і високої масштабованості, генеративні змагальні мережі для створення високоякісних зображень і автокодери мають вирішальне значення. Кожна архітектура має свої сильні та слабкі сторони, які визначають її придатність для конкретних задач машинного навчання та обробки зображень.

1.2 Інструменти медичної діагностики легневих захворювань

1.2.1 Методи діагностики

Пульмонологія використовує комплексний підхід до діагностики захворювань дихальної системи, комбінуючи традиційні інструментальні методи з сучасними технологіями. Основними традиційними методами є:

Рентгенографія грудної клітки – один з найпоширеніших методів візуалізації, який дозволяє оцінити структуру легень, серця та середостіння. Використовується для виявлення пневмонії, пухлин, пневмотораксу та інших патологій. Рентгенографія працює на основі використання рентгенівських променів — високочастотного електромагнітного випромінювання, яке проходить через тіло пацієнта. Різні тканини тіла мають різну здатність до поглинання цих променів, що дозволяє створити зображення внутрішніх структур організму. Процес починається з того, що рентгенівська трубка генерує пучок променів, які спрямовуються через тіло пацієнта. Під час проходження через тіло рентгенівські промені взаємодіють з різними тканинами. Наприклад, кістки поглинають більше променів, тому на зображенні вони виглядають світлішими, тоді як м'які тканини або порожнини, такі як легені, пропускають більше променів, через що виглядають темнішими. Після проходження через тіло промені досягають плівки або цифрового детектора, де й формується зображення. Це зображення є чорно-білим: щільні структури виглядають світлими, а менш щільні або порожнисті ділянки — темними. Отримане рентгенівське зображення аналізується лікарем-радіологом для виявлення патологій, таких як переломи, пухлини, запальні процеси або інші захворювання. Таким чином, рентгенографія є ефективним інструментом для швидкого діагностування багатьох захворювань та станів, особливо в екстрених випадках. Різні пульмонологічні захворювання відображаються на рентгенівських знімках грудної клітки по-різному, залежно від характеру ураження тканин легень. Ось як виглядають деякі з найпоширеніших патологій.

1. На рентгені пневмонія виглядає як ділянки затемнення (консолідації), що зазвичай виникають у нижніх частках легень. Це зумовлено скупченням рідини та клітин запалення в альвеолах. Зони затемнення є щільними й однорідними, іноді з чіткими межами.
2. Пневмоторакс проявляється на рентгенівському знімку як відсутність легеневої тканини в певній ділянці легень. Замість цього спостерігається тонка лінія, що окреслює колапсовану легеню, і чорна зона — це повітря в плевральній порожнині. Легеня виглядає зменшеною, зсунутою до середини грудної клітки.
3. Туберкульоз зазвичай проявляється на рентгені у вигляді каверн (порожнин) у верхніх частках легень. Можуть також бути видимі затемнені ділянки або вогнищеві ураження через запальні процеси. У випадку хронічного туберкульозу можуть спостерігатися фіброз і рубцювання.
4. Рентгенівські ознаки ХОЗЛ включають надмірну прозорість легень через емфізему (збільшення повітряних просторів) і розширення грудної клітки. Діафрагма зазвичай сплюснена, а легеневі поля виглядають надмірно розширеними.
5. Легеневий фіброз проявляється у вигляді дифузних змін у тканині легень, зокрема появою "сотової легені" (множинних дрібних порожнин) у базальних і периферичних ділянках. Це виглядає як мілкі сітчасті структури на рентгені.
6. Рак легень може виглядати як чітко окреслене ущільнення або маса в легенях, яка має неправильну форму. Якщо пухлина досягає великих розмірів, вона може викликати зміщення структур грудної клітки або призвести до утворення плеврального випоту (рідини в плевральній порожнині).
7. На рентгені саркоїдоз зазвичай проявляється збільшенням лімфатичних вузлів у середостінні та по обидва боки трахеї, що називається двосторонньою гіларною лімфаденопатією. Також можуть спостерігатися дрібні вузлики або ураження в легневих полях.

Кожне із цих захворювань має свої характерні рентгенологічні ознаки, що допомагає лікарям визначити точний діагноз і призначити відповідне лікування.

Комп'ютерна томографія (КТ) – забезпечує більш детальне зображення легень, використовується для діагностики онкологічних захворювань, туберкульозу, інтерстиційних захворювань легень та інших складних патологій. Комп'ютерна томографія (КТ) використовує рентгенівські промені для створення детальних зображень внутрішніх структур тіла шляхом отримання серії пошарових зображень (зрізів). Під час процедури рентгенівська трубка обертається навколо тіла пацієнта, випромінюючи рентгенівські промені, які проходять через тіло під різними кутами. Промені частково поглинаються різними тканинами тіла залежно від їхньої щільності, а детектори фіксують промені, що пройшли через тіло, утворюючи двовимірні зображення зрізів на кожному етапі обертання. Після збору цих зображень комп'ютер реконструює їх у тривимірну модель внутрішніх органів. Завдяки високій роздільній здатності, КТ дозволяє виявляти деталі, які можуть бути невидимі на звичайних рентгенівських знімках. Це дає можливість більш точної діагностики різних захворювань, таких як пухлини, травми, крововиливи або ураження легень. КТ часто використовується для візуалізації мозку, грудної клітки, черевної порожнини та кісток. Окрім цього, КТ можна поєднувати з контрастними речовинами, які вводяться пацієнту, щоб підвищити видимість кровоносних судин, органів чи патологій. Завдяки можливості отримувати зображення в різних площинах та створювати тривимірні реконструкції, комп'ютерна томографія забезпечує високу точність у діагностиці та плануванні лікування.

Різні пульмонологічні захворювання на комп'ютерній томографії (КТ) візуалізуються через характерні зміни в структурі легень та навколишніх тканин. Наприклад, пневмонія на КТ часто виглядає як ділянки консолідації, які з'являються у вигляді затемнених зон, що можуть бути розташовані в нижніх частках легень. У разі запальних процесів можливі також вогнищеві ураження з неоднорідною структурою.

Пневмоторакс проявляється на КТ як відсутність легеневої тканини в певній ділянці, що супроводжується наявністю повітря в плевральній порожнині, яке виглядає як чорна зона на зображенні. У випадку туберкульозу на КТ можна побачити каверни (порожнини) та вогнищеві ураження, які найчастіше локалізуються у верхніх частках легень. Хронічне обструктивне захворювання легень (ХОЗЛ) на КТ може бути представлене розширенням легень, емфіземою та змінами в бронхах. Діафрагма виглядає сплющеною, а міжреберні простори розширеними, що свідчить про надмірну повітряність. Легеневий фіброз характеризується наявністю дрібних, сіточкоподібних структур та змін в альвеолах, які створюють картину "сотової легені". Це може бути супроводжуване фіброзними змінами, які виглядають як щільні зони на фоні нормальної легеневої тканини. Рак легень на КТ може проявлятися як вогнище з неправильною формою, який часто має нерівні краї та може супроводжуватися обструкцією бронхів або наявністю метастазів у регіональних лімфатичних вузлах. Саркоїдоз, своєю чергою, відображається на КТ як збільшення лімфатичних вузлів у середостінні, а також можливі дрібні вогнищеві ураження в легенях. Ці характеристики допомагають лікарям ставити точні діагнози та планувати лікування, орієнтуючись на візуалізацію патологій.

Магнітно-резонансна томографія (МРТ) – рідше використовується у пульмонології через обмежену можливість візуалізації легень, але може бути корисною для дослідження м'яких тканин, зокрема пухлин у грудній клітці. Магнітно-резонансна томографія (МРТ) працює на основі принципів ядерного магнітного резонансу, використовуючи потужне магнітне поле, радіохвилі та комп'ютерну обробку для створення детальних зображень внутрішніх структур тіла. Під час процедури пацієнт розташовується на рухомій платформі, яка вводить його в великого розміру магніт. Це магнітне поле спричиняє орієнтацію протонів, які містяться в тканинах тіла, зокрема у воді. Коли протони орієнтуються в магнітному полі, на них впливають радіохвилі, які відправляються з машини. Ці радіохвилі заважають орієнтації протонів, а після їх вимкнення протони повертаються у своє початкове положення,

випромінюючи радіохвилі. Ці сигналізації захоплюються детекторами, які обробляються комп'ютером для створення двовимірних та тривимірних зображень тканин і органів. МРТ має високу здатність до розрізнення м'яких тканин, що робить його особливо корисним для візуалізації мозку, спинного мозку, суглобів і м'яких тканин, таких як м'язи та зв'язки. Це неінвазивна процедура, що не вимагає рентгенівського випромінювання, тому вона безпечна для більшості пацієнтів. МРТ використовується для діагностики різноманітних захворювань, включаючи пухлини, неврологічні розлади, травми, а також для оцінки стану м'яких тканин та органів.

На магнітно-резонансній томографії (МРТ) різні пульмонологічні захворювання відображаються завдяки здатності цього методу візуалізувати м'які тканини з високою точністю. Наприклад, пневмонія на МРТ може виглядати як зони з підвищеною інтенсивністю сигналу внаслідок запальних процесів у легенях, часто з чіткими межами. Це дозволяє виявити вогнищеві або дифузні зміни в легеневої тканині. Пневмоторакс може бути представлений на зображеннях як область з відсутністю легеневої тканини, що супроводжується наявністю повітря в плевральній порожнині, яка виглядає темною на зображеннях. Відсутність легеневих структур у відповідній ділянці також легко виявити. При туберкульозі на МРТ можуть бути помітні каверни, які виглядають як порожнини з тонкими стінками, а також вогнищеві ураження, які демонструють запальні зміни та гіпертрофію лімфатичних вузлів. Хронічне обструктивне захворювання легень (ХОЗЛ) може візуалізуватися як розширення повітряних просторів і зміни в стінках бронхів, зокрема, зниження їхньої еластичності. Легеневий фіброз на МРТ характеризується наявністю сіточкоподібних структур, які вказують на фіброзні зміни в легеневої тканині. Ці зміни можуть створювати картину "сотової легені", де спостерігається зменшення легеневих обсягів. Рак легень проявляється на МРТ як вогнище з неправильною формою та нерівними краями, що може супроводжуватися метастазами в лімфатичних вузлах, які також добре видно на зображеннях. Саркоїдоз може бути виявлений через збільшення лімфатичних вузлів у

середостінні та можливі вогнищеві ураження, які відрізняються за сигналом від нормальної легеневої тканини. Таким чином, МРТ є важливим інструментом для візуалізації різних пульмонологічних захворювань, що дозволяє лікарям точно оцінити стан легень і вчасно діагностувати патології.

Бронхоскопія – інвазивний метод дослідження, що дозволяє оглянути дихальні шляхи та провести біопсію. Застосовується для діагностики пухлин, запальних процесів і для взяття матеріалу для гістологічного аналізу.

Спірометрія – функціональний тест, який вимірює об'єм та швидкість повітря при вдиху і видиху. Використовується для діагностики обструктивних захворювань, таких як астма та ХОЗЛ.

Аналіз газів крові – дає інформацію про газообмін у легенях, дозволяючи оцінити рівень кисню та вуглекислого газу в артеріальній крові, що важливо для виявлення дихальної недостатності.

1.2.2 Методи підготовки наборів даних

Набір даних (від англ. dataset) – це структурований набір інформації, що використовується для аналізу, моделювання та машинного навчання. Набори даних можуть містити багато різних типів даних, таких як числа, текст, зображення, відео та аудіо. Набори даних можуть бути великими або малими, а їхня структура залежить від їхнього призначення. Навіщо потрібні набори даних: по-перше, для навчання штучного інтелекту та машинного навчання. Набори даних є основним джерелом для навчання алгоритмів машинного навчання. Наприклад, якщо ви хочете побудувати модель для розпізнавання облич на фотографії, вам знадобиться великий набір зображень облич для навчання; по-друге, аналіз даних. Набори даних використовуються для аналізу інформації та отримання корисних висновків. Це може бути економічний аналіз, дослідження в галузі охорони здоров'я або соціальних тенденцій. По-третє, підтримка наукових досліджень. Вчені використовують набори даних для проведення досліджень, перевірки гіпотез і прогнозування. Нарешті, моделі оцінюються і

тестуються. Після того, як модель навчена, її тестують, використовуючи нові дані з набору даних, щоб оцінити її точність і надійність. Таким чином, набори даних є основою для розвитку новітніх технологій, аналізу інформації та прийняття рішень на основі даних.

Підготовка даних для датасетів є важливим етапом у процесі машинного навчання, оскільки від якості даних залежить точність і ефективність моделі. Цей процес охоплює кілька ключових аспектів. Першим кроком є збір даних. Він може здійснюватися з різних джерел, таких як бази даних, вебсайти, сенсори, API або інші системи. Зібрані дані часто бувають неструктурованими або частково структурованими, тому важливо переконатися, що всі дані відповідають стандартам, необхідним для побудови моделі. Наступним кроком є очищення даних. У процесі збору даних часто виникають помилки, дублювання, відсутні значення або некоректні формати. Тому очищення даних передбачає виявлення й усунення цих проблем. Наприклад, відсутні значення можна заповнити середніми або модальними значеннями, а некоректні формати можна виправити за допомогою спеціальних методів обробки. Важливо також позбутися дубльованих даних, оскільки вони можуть викликати спотворення результатів моделі. Після очищення відбувається процес нормалізації або стандартизації даних. Для того, щоб модель працювала ефективно, різні параметри мають бути приведені до єдиного масштабу. Це необхідно, особливо коли дані мають різні одиниці виміру або великі різниці в діапазонах. Для цього застосовуються методи, такі як мінімакс нормалізація або стандартизація за допомогою обчислення середнього значення і стандартного відхилення. Ще один важливий етап — трансформація даних. Це може включати перетворення категорійних даних у числові форми (наприклад, через one-hot encoding) або об'єднання кількох атрибутів у нові характеристики. Іноколи дані можуть бути надто високорозмірними, що ускладнює побудову моделі. У таких випадках використовується зменшення розмірності через методи, як-от PCA (аналіз головних компонент) або LDA (латентний дирихлетовий розподіл). Після того як дані готові, відбувається розділення набору даних на навчальну, тестову та

іноді валідаційну вибірки. Це дозволяє забезпечити об'єктивну оцінку моделі. Зазвичай використовується поділ 70-80% для навчання, а решта для тестування та валідації. Правильний поділ вибірки є критичним для уникнення проблеми перенавчання, коли модель надто добре запам'ятовує навчальні дані, але погано працює з новими. Нарешті, проводиться візуалізація та аналіз підготовлених даних для виявлення можливих трендів, аномалій чи кореляцій. Це допомагає зрозуміти дані краще та прийняти рішення щодо подальшої роботи з моделлю, коригування характеристик або зміни підходів до моделювання. Підготовка даних є фундаментальним процесом, що забезпечує якість моделі та її здатність розв'язувати конкретні задачі.

Процес підготовки зображень, таких як медичні знімки, для формування датасету є складним і багатокроковим. Він потребує уваги до якості даних, етичних аспектів і специфічних вимог обробки медичних зображень.

На першому етапі медичні зображення збираються з різних джерел, таких як лікарні, клініки або дослідницькі установи. Зображення можуть бути отримані за допомогою різних медичних технологій, таких як рентген, МРТ, КТ або УЗД. На цьому етапі важливо дотримуватися законів і етичних норм, таких як захист персональних даних пацієнтів (наприклад, відповідно до HIPAA в США або GDPR в Європі).

Оскільки медичні зображення зазвичай містять особисту інформацію, необхідно провести анонімізацію. Це означає видалення або приховування будь-яких ідентифікаційних даних, таких як ім'я пацієнта, дата народження або номер медичної картки. Анонімізація забезпечує захист конфіденційності пацієнтів.

Медичні зображення часто мають специфічні формати, наприклад, DICOM (Digital Imaging and Communications in Medicine), який є стандартом для зберігання та передачі медичних зображень. На цьому етапі може знадобитися перетворення зображень у формат, зручний для машинного навчання, наприклад, у JPEG, PNG або TIFF. Важливо зберегти якість зображення під час конвертації, оскільки деталі зображення можуть бути критичними для діагностики.

У процесі підготовки датасету зображення можуть містити шуми, артефакти або непотрібні ділянки, які потрібно видалити. Наприклад, можуть бути знімки з поганою якістю, помилками сканування або некоректними позначеннями. На цьому етапі такі зображення відфільтровуються або коригуються, щоб забезпечити однорідність датасету.

Після очищення медичні зображення потребують попередньої обробки для покращення їхньої якості та підготовки до використання в моделях. Попередня обробка може включати:

- нормалізацію яскравості та контрасту для того, щоб зробити зображення більш чітким;
- фільтрацію для зменшення шумів (наприклад, застосування фільтрів Гауса або медіанних фільтрів);
- масштабування зображень до фіксованого розміру, що необхідно для того, щоб усі зображення мали однакову роздільну здатність;
- кадрування або видалення зайвих ділянок, щоб зосередитися на найважливіших частинах зображення, наприклад, на органі, що досліджується.

Для підвищення якості та кількості даних проводиться аугментація зображень. Це метод штучного збільшення кількості зображень шляхом їх обертання, дзеркального відображення, зміни яскравості, контрасту або навіть додавання шуму. Аугментація допомагає уникнути перенавчання моделі, забезпечуючи різноманітність зображень і покращуючи її здатність до узагальнення.

У медичних датасетах дуже важливо правильно анотувати зображення, тобто надати їм правильні мітки або пояснення. Анотація може включати:

- класифікацію зображень на основі типу патології (наприклад, наявність пухлини або інші аномалії);
- сегментацію зображень, де окремі частини, такі як органи або уражені ділянки, виділяються та позначаються для точного аналізу;

- визначення ROI (Region of Interest) — областей інтересу, які мають бути досліджені під час навчання моделі.

Анотацію зазвичай виконують фахівці, такі як лікарі або рентгенологи, оскільки для правильного позначення медичних знімків необхідні глибокі знання в галузі медицини.

Після підготовки та анотації даних важливо розділити датасет на різні частини для навчання, валідації та тестування. Цей поділ гарантує, що модель не тільки навчиться на певному наборі зображень, але й зможе коректно працювати з новими даними. Зазвичай 70-80% датасету виділяється для навчання, а решта використовується для тестування та валідації.

У медичних зображеннях часто буває проблема дисбалансу класів, коли, наприклад, хворих знімків значно менше, ніж здорових. Це може призвести до того, що модель буде схильною робити помилкові передбачення. Тому важливо забезпечити баланс між різними класами даних через методи, такі як аугментація або додатковий збір даних.

Після того як датасет готовий, необхідно перевірити його на наявність помилок або недоліків. Це може включати перевірку якості анотацій, перевірку зображень на відповідність форматам або стандартам, а також перевірку наявності артефактів або проблем із зображеннями.

Медичні зображення повинні використовуватися відповідно до етичних норм і законодавства. Потрібно отримати згоду пацієнтів на використання їхніх даних у наукових або комерційних цілях. Також важливо гарантувати, що датасет відповідає міжнародним і локальним стандартам безпеки та конфіденційності даних.

1.2.3 Результат дослідження наборів даних

У ході проведеного дослідження знайдено декілька релевантних датасетів, що надають можливість аналізу різноманітних захворювань легень. Основними критеріями для вибору цих наборів даних стали їхня доступність, актуальність

для теми роботи та якість представленої інформації. Зібрані набори включають такі типи даних, як зображення (рентгенівські знімки, знімки комп'ютерної томографії), текстові дані.

Chest X-ray14 : цей набір даних містить понад 112 000 рентгенівських знімків грудної клітки, які класифіковані за 14 типами патологій (Рис. 1.2), включно з пневмонією, пневмотораксом, легневим фіброзом та іншими захворюваннями. Він є одним з найбільших загальнодоступних датасетів для вивчення патологій легень, що робить його цінним для тренування нейромереж у завданнях автоматичної діагностики.[1]

COVID-19 Radiography Database. Цей набір даних містить рентгенівські знімки пацієнтів з діагнозом COVID-19, пневмонією, а також здорових людей. Використовується для класифікації зображень та виявлення ознак COVID-19 на рентгенівських знімках. Датасет надає можливість тренувати моделі для специфічних завдань у контексті пандемій та вірусних захворювань легень.[2]

LIDC-IDRI (Lung Image Database Consortium and Image Database Resource Initiative). Цей набір містить понад 1000 КТ-знімків легень, які анотували кілька радіологів для виявлення легневих вузлів. Його застосовують для досліджень, пов'язаних із виявленням раку легень, сегментацією зображень та оцінкою методів візуалізації.[3]

MIMIC-CXR Dataset. Набір даних містить рентгенівські знімки та їхні описи, які отримані з електронних медичних записів пацієнтів. Він містить понад 370 000 знімків та текстові описи, що дозволяє застосовувати нейромережі для аналізу не лише зображень, але й текстової інформації, покращуючи можливості багатомодального навчання.[4]

RadChest – великий набір даних рентгенівських знімків грудної клітки, що містить понад 160 000 зображень від 67 000 пацієнтів з відповідними звітами. Дані зібрані у лікарні Сан-Хуан (Іспанія) з 2009 по 2017 рік і включають 174 рентгенографічні знахідки, 19 діагнозів і 104 анатомічні локації. Близько 27% звітів анотовані вручну лікарями, а інші марковані за допомогою рекурентної нейронної мережі з механізмами уваги, з точністю Micro-F1 0.93.[5]

ChestX-ray8 – це медичний набір даних, що містить 108 948 рентгенівських знімків грудної клітки (фронтальні проєкції) від 32 717 пацієнтів, зібраних з 1992 по 2015 рік. Він включає вісім поширених захворювань, анотації яких отримані за допомогою NLP-аналізу текстових радіологічних звітів.[6]

ChestX-Det – це набір рентгенівських знімків грудної клітки з анотаціями на рівні об'єктів (бокси та маски). Він є підмножиною публічного набору NIH ChestX-ray14 і містить приблизно 3500 зображень із 13 поширеними категоріями захворювань, які були анотовані трьома сертифікованими радіологами.[7]

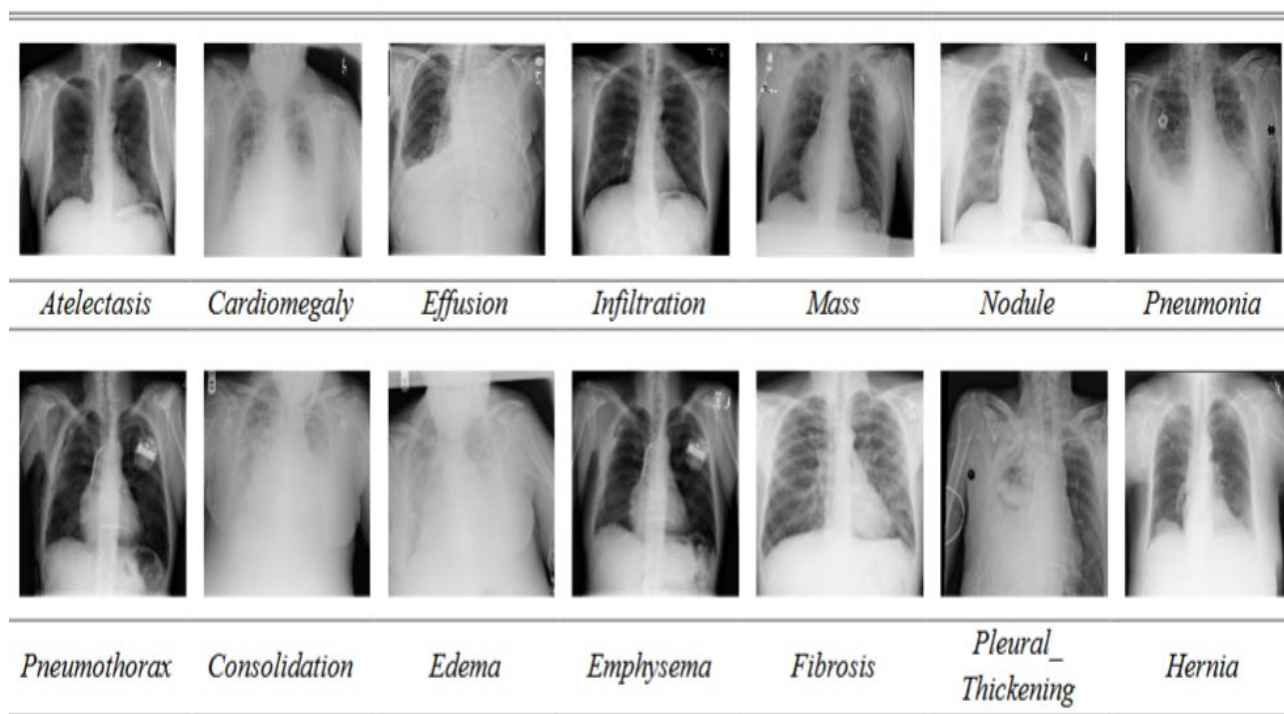


Рисунок 1.2 – Захворювання, зображення яких найчастіше зустрічаються у датасетах

На основі порівняння датасетів, MIMIC-CXR виглядає як найкращий варіант для комплексного аналізу пульмонологічних захворювань завдяки великій кількості зображень та текстових описів, що дозволяє використовувати багатомодальне навчання. Однак, не варто обмежуватися одним набором даних. Наприклад, PadChest має широкі анотації та додаткові діагнози, що можуть

покращити точність моделей. Використання декількох датасетів, таких як LIDC-IDRI для дослідження раку легень чи COVID-19 Radiography для пандемічних захворювань, забезпечить більш глибокий і різноманітний підхід.

Таблиця 2.1

Порівняння датасетів

Датасет	Розмір	Тип	Формат	Збалансованість	Варіативність
ChestX-ray14	112 000 рентгенівських знімків	14 патологій легень	PNG	Незбалансований	Великий набір зображень, фронтальна проєкція
COVID-19 Radiography	16 153 зображень	COVID-19, вірусна пневмонія, здорові легені	PNG	Більше випадків COVID-19, незбалансований	Наявність декількох класів
LIDC-IDRI	1 018 КТ-знімків з анотаціями	КТ-знімки легень, анотації легеневих вузлів	DICOM	Збалансований за вузлами	Детальні анотації вузлів
MIMIC-CXR	370 000 рентгенівських знімків та звітів	Рентгенівські знімки та текстові звіти	DICOM	Різнорманітні випадки	Мультиформатні дані (зображення + текст)
PadChest	160 000 рентгенівських знімків та клінічних звітів	Рентгенівські знімки	JPG	Частково збалансований	Різні проєкції, широкі анотації
ChestX-ray8	108 948 рентгенівських знімків	8 поширених захворювань, фронтальна проєкція	PNG	Незбалансований	Обмежена кількість захворювань
ChestX-Det	3 500 рентгенівських знімків	Анотації на рівні об'єктів (бокси та маски)	PNG	Незбалансований	Анотації на рівні об'єктів (бокси та маски)

1.3 Методи діагностики за допомогою нейромережових технологій

1.3.1 Алгоритми згорткових нейронних мереж для аналізу зображень

Згорткові нейронні мережі (CNN) є основним підходом до аналізу зображень завдяки їх здатності автоматично витягати ознаки, необхідні для розпізнавання об'єктів, класифікації, сегментації та інших завдань. Цей тип нейронних мереж обробляє зображення, виділяючи локальні особливості, що робить їх особливо ефективними у порівнянні з традиційними багатошаровими перцептронами, де кожен нейрон обробляє лише одну точку.

Першим кроком у CNN є використання згорткових шарів, які працюють як фільтри або ядра, що «ковзають» через зображення. Кожен фільтр представляє певну особливість, яку мережа намагається виявити, наприклад, контури, текстури або патерни різних рівнів абстракції. Під час згортки ці фільтри застосовуються до невеликих ділянок зображення, що дозволяє кожному нейрону відгукуватися на локальні риси. Результати цієї обробки створюють «карти ознак», які передають основну інформацію для наступних шарів.

Далі зазвичай додається активаційний шар з функцією, яка вносить нелінійність у модель. Найпопулярнішою є функція активації ReLU (Rectified Linear Unit), яка перетворює всі від'ємні значення на нуль, зберігаючи лише позитивні. Це дозволяє моделі вчитися виявляти не лише лінійні, але і складніші, нелінійні взаємозв'язки між ознаками.

Наступний крок – це шар підвибірки або шар пуолінгу, який зменшує розмірність вхідних даних, зберігаючи водночас найбільш важливу інформацію. Замість того щоб обробляти всі значення, цей шар виділяє максимум кожної ділянки, що значно зменшує кількість обчислень і сприяє узагальненню моделі. Після кількох згорткових і підвибіркових шарів зображення зазвичай стає менш чітким, але містить всі найважливіші ознаки.

У завершальних етапах додаються повнозв'язні шари, де кожен нейрон має зв'язок з усіма нейронами попереднього шару. Тут мережа об'єднує виявлені раніше ознаки і використовує їх для розпізнавання конкретних категорій. Останній шар, якщо це задача класифікації, часто використовує softmax-функцію для перетворення виходу в ймовірності кожного класу.

Функція втрат визначає ефективність моделі, вказуючи, наскільки правильні її прогнози. Основне завдання алгоритму полягає в мінімізації значення функції втрат, що досягається шляхом зворотного розповсюдження помилки. Під час цього процесу ваги коригуються методом градієнтного спуску або іншими оптимізаторами, такими як Adam, що допомагає мережі поступово вдосконалювати свою точність.

Для контролю перенавчання в CNN застосовуються регуляризаційні техніки, коли випадково вимикаються окремі нейрони на етапі тренування. Це запобігає надмірній залежності від певних ознак і підвищує узагальнюючу здатність мережі на нових даних. Таким чином, CNN виявляються надзвичайно ефективними для розпізнавання складних патернів, що робить їх основою багатьох застосувань, таких як комп'ютерне бачення, обробка медичних зображень, розпізнавання облич і багато інших сучасних систем штучного інтелекту.

Архітектура згорткових нейронних мереж (CNN) має свої витоки в 1980-х роках, коли розробники прагнули знайти ефективні методи обробки візуальних даних для автоматизованого розпізнавання об'єктів. Одним із піонерів у цьому напрямі був дослідник Янн Лекун, який створив одну з перших CNN-архітектур під назвою LeNet-5 у 1989 році. Ця мережа призначалася для розпізнавання рукописних цифр у поштових індексах і банківських чеках. LeNet-5 використовувала згорткові шари для автоматичного витягування ознак, що дозволяло ідентифікувати ключові патерни, характерні для кожної цифри, без необхідності ручного налаштування фільтрів. Важливою перевагою LeNet стало зменшення кількості параметрів через використання згорткових і шарів пулінгу, що дало можливість більш ефективно навчати модель і підвищило її

здатність до узагальнення. Попри успіх LeNet-5, до 2010-х років CNN використовувалися обмежено, оскільки на той час не вистачало обчислювальних ресурсів і великих датасетів для ефективного навчання глибоких моделей. Прорив відбувся у 2012 році, коли команда дослідників під керівництвом Джеффри Хінтона, Алекса Крижевського та Іллі Суцкевера створила AlexNet, модель CNN, яка виграла змагання ImageNet Large Scale Visual Recognition Challenge (ILSVRC) з величезним відривом від інших алгоритмів. AlexNet була глибшою і значно складнішою за LeNet, включаючи більше згорткових і пулінгових шарів, а також інноваційні техніки регуляризації, зокрема dropout, що суттєво зменшило ризик перенавчання. AlexNet, по суті, стала точкою відліку для розквіту глибокого навчання. Секрет успіху AlexNet полягав також у використанні графічних процесорів (GPU), які дозволили пришвидшити навчання і обробляти більші обсяги даних. Модель AlexNet продемонструвала, що CNN здатні досягати високої точності при класифікації зображень і показала їх перевагу над традиційними методами машинного навчання. Це надихнуло багатьох дослідників на розробку більш складних моделей, і згодом з'явилися такі мережі, як VGG, ResNet і Inception, які продовжили підвищувати точність і ефективність CNN у різних застосунках. Завдяки AlexNet CNN стали основою сучасного глибокого навчання і стимулювали створення нових підходів для обробки зображень, що сьогодні використовуються в автономних системах, медицині, безпеці й багатьох інших галузях.

В процесі дослідження було вивчено декілька статей на тему кваліфікаційної роботи, де було розглянуто плюси і мінуси використання згорткових та інших мереж.

“CNN ефективно вивчають дискримінаційні ознаки та витягують узагальнюючі предиктори з великих наборів медичних даних, забезпечуючи відмінну продуктивність при виконанні завдань медичної візуалізації, що робить їх невіддільним компонентом сучасних систем медичної візуалізації на основі ШІ. Досягнення в галузі CNN в основному стимулювалися новими архітектурними рішеннями, кращими процедурами оптимізації, наявністю

спеціального апаратного забезпечення (наприклад, графічних процесорів) і спеціально створених бібліотек програмного забезпечення з відкритим вихідним кодом. Ми відсилаємо зацікавлених читачів до вичерпних оглядів, пов'язаних із застосуванням CNNs у медичній візуалізації. Всупереч значному приросту продуктивності, залежність CNNs від великих маркованих наборів даних обмежує їхню застосовність у повному спектрі задач медичної візуалізації. Крім того, підходи на основі CNNs, як правило, складніші для інтерпретації й часто діють як «чорний ящик». Тому в співтоваристві медичної візуалізації спостерігається все більше зусиль, спрямованих на об'єднання сильних сторін ручних методів і методів на основі CNNs, в результаті чого з'являються попередні інформаційно-керовані моделі CNNs. Ці гібридні методи містять спеціальні шари, специфічні для конкретної області, і включають розгорнуту оптимізацію, генеративні моделі і підходи на основі навченого детектора. Попри ці архітектурні та алгоритмічні досягнення, вирішальним фактором успіху CNN є їхня індуктивна орієнтація на конкретне зображення, що забезпечує інваріантність масштабу та моделювання локальних візуальних структур. Хоча ця внутрішня локальність (обмежене рецептивне поле) підвищує ефективність CNN, вона погіршує їхню здатність фіксувати далекі просторові залежності у вхідному зображенні, що призводить до стагнації продуктивності. Це вимагає альтернативного архітектурного дизайну, здатного моделювати далекі взаємозв'язки між пікселями для кращого навчання представлення.” [\[8, с.3\]](#)

Сучасні задачі та застосування згорткових нейронних мереж (CNN) стали надзвичайно актуальними в багатьох сферах завдяки їхній здатності ефективно обробляти та аналізувати візуальну інформацію. Одним з основних напрямків є розпізнавання облич, яке використовується в системах безпеки, соціальних мережах та комерційних додатках. CNN здатні виявляти та ідентифікувати обличчя на зображеннях з високою точністю, що дозволяє застосовувати цю технологію в біометричних системах, що забезпечують доступ до різноманітних ресурсів.

Обробка медичних зображень є ще однією важливою галуззю застосування CNN. Тут ці нейронні мережі використовуються для аналізу рентгенівських знімків, МРТ, КТ та інших медичних зображень, допомагаючи в діагностиці захворювань, таких як рак чи пневмонія. CNN здатні виявляти аномалії та патерни, які можуть бути важко помітними для людського ока, що сприяє підвищенню точності діагнозів та швидкості їхнього встановлення.

Автономні системи, такі як безпілотники і автомобілі, також використовують CNN для аналізу візуальної інформації з камер. Це дозволяє системам здійснювати навігацію, розпізнавати перешкоди та приймати рішення в реальному часі. CNN в поєднанні з іншими технологіями, такими як LiDAR, забезпечують високу точність і надійність у складних умовах навколишнього середовища.

Специфічні задачі, пов'язані з обробкою глибоких карт (depth maps), є важливими в контексті 3D-візуалізації та комп'ютерного зору. CNN можуть використовуватися для аналізу глибоких карт, що дозволяє визначати просторові структури та об'єкти в сцені. Це має велике значення в таких областях, як робототехніка та віртуальна реальність, де точне сприйняття простору є критично важливим.

Обробка відео з використанням 3D-CNN також є актуальною задачею, оскільки ця технологія дозволяє аналізувати не лише просторові, а й часові аспекти візуальних даних. 3D-CNN використовуються для розпізнавання дій у відео, виявлення об'єктів, а також для обробки відеоданих в реальному часі, що має широке застосування в безпеці, спорті та розвагах.

У сучасних згорткових нейронних мережах (CNN) існують різні типи згорткових шарів і функцій активації, кожен з яких має свої особливості і використовується для різних завдань. Вибір конкретного типу згортки чи функції активації може суттєво вплинути на продуктивність моделі і допомогти їй краще узагальнювати дані, адаптуючись до особливостей конкретного завдання.

Наведемо типи згорткових шарів.

1. Dilated (або розширені) згортки: розширені згортки використовують фільтри зі вставленими «пропусками» (дірами) між елементами, що дозволяє збільшувати область сприйняття без збільшення кількості параметрів або розміру фільтра. Це корисно для задач, де необхідно захопити більше контекстної інформації з віддалених областей зображення, наприклад, у сегментації зображень і аналізі часових рядів.
2. Separable Convolutions (роздільні згортки): Цей тип згортки розбиває операцію на дві частини: спочатку виконується окрема згортка для кожного каналу (глибина згортки), а потім відбувається об'єднання каналів у просторовій площині. Це дозволяє суттєво знизити обчислювальні витрати і кількість параметрів, зберігаючи водночас високу точність. MobileNet, наприклад, широко використовує separable convolutions для покращення ефективності.
3. Depthwise Convolutions (глибинні згортки): Це різновид separable convolutions, де кожен канал обробляється окремим фільтром. Глибинні згортки дають змогу покращити обчислювальну ефективність, особливо в мобільних і вбудованих пристроях, оскільки зменшують кількість параметрів.
4. Transposed Convolutions (транспоновані згортки): Такі згортки також називають зворотними або «де-згортками» і використовуються для збільшення розмірності зображень. Вони часто застосовуються в генеративних мережах, таких як GAN, або в задачах семантичної сегментації, де потрібно відновити просторове розширення виводу.
5. Pointwise Convolution (1x1 згортка): Це операція згортки з фільтром розміром 1x1, яка дозволяє зменшувати або збільшувати кількість каналів без зміни просторових розмірів зображення. Pointwise згортка часто використовується у поєднанні з іншими типами згорток для підвищення глибини моделі або зменшення обчислювальної складності.

Наведемо функції активації.

1. Leaky ReLU: Ця функція активації є модифікованим варіантом ReLU, яка дозволяє зберегти невелике значення для негативних вхідних значень (шляхом множення на малий коефіцієнт, наприклад, 0.01). Це розв'язує проблему "вмирання" нейронів, яка виникає в ReLU, де негативні значення перетворюються на нуль і можуть призводити до незворотних втрат нейронів у мережі.
2. ELU (Exponential Linear Unit): ELU забезпечує поступове зростання у негативному діапазоні, що дозволяє моделі мати більш гнучку відповідь для негативних значень. Це допомагає вирівнювати середнє значення виходу нейронів, що може поліпшити стабільність і швидкість навчання.
3. Swish: функція Swish була запропонована дослідниками Google і визначається як $f(x) = x / (1 + e^{-x})$. Вона забезпечує плавний перехід між негативними та позитивними значеннями і дає змогу досягти високої точності на багатьох завданнях, особливо у великих глибоких мережах. Swish може давати кращі результати в порівнянні з ReLU на складних задачах, хоча і потребує більше обчислень.
4. Mish: функція активації Mish має вигляд $f(x) = x \cdot \tanh(\ln(1 + e^x))$. Вона поєднує властивості ReLU і ELU і створює плавний вихід, що сприяє покращенню навчання глибоких мереж на деяких задачах. Mish допомагає мережі навчатися узагальненню ознак завдяки плавному переходу, що робить її актуальною для використання в задачах класифікації і сегментації.
5. Softmax: використовується в останньому шарі CNN для задач багатокласової класифікації. Softmax перетворює вихідні значення в ймовірності кожного класу, що допомагає інтерпретувати вихід мережі в термінах ймовірностей належності до певного класу.

Вибір конкретного типу згортки і функції активації значною мірою залежить від завдання та обчислювальних ресурсів. Різноманітні типи згорток і функцій активації дозволяють налаштовувати CNN для досягнення оптимальних

результатів у різних застосунках, таких як сегментація, розпізнавання облич, обробка медичних зображень тощо.

Підвибірка та зменшення розмірності є важливими методами в згорткових нейронних мережах (CNN), оскільки вони дозволяють знижувати обчислювальну складність, зберігаючи при цьому найважливіші ознаки вхідних даних. Ці методи допомагають уникати перенавчання, підвищують здатність моделі до узагальнення і покращують її продуктивність, зберігаючи при цьому критичну інформацію.

Пулінг є поширеним підходом до підвибірки, що зменшує розмірність простору ознак, виділяючи найзначущі частини зображення. Серед основних підходів виділяють *max pooling*, що обирає максимальні значення у фрагментах зображення для акцентування на найяскравіших ознаках, і *average pooling*, який обчислює середні значення у кожному фрагменті для збереження загальної інформації. *Global average pooling* зменшує просторову розмірність до одного значення для кожного каналу, що особливо корисно для завершальних етапів CNN у задачах класифікації. *Adaptive pooling* дозволяє отримати бажаний розмір виходу, автоматично підлаштовуючи розмір вікна пулінгу, що є особливо зручним для обробки зображень різних розмірів.

Інші методи зменшення розмірності часто використовуються для збереження критичних ознак при одночасному зниженні обчислювального навантаження. *Dropout*, наприклад, є методом регуляризації, що випадковим чином відключає частину нейронів під час навчання і тим самим зменшує складність моделі, підвищуючи здатність до узагальнення. *Principal Component Analysis (PCA)* виконує проєкцію даних на напрямки, які максимально зберігають їхню варіацію, і зазвичай використовується на етапі попередньої обробки для зменшення розмірності простору ознак перед введенням їх до мережі. *Spatial Pyramid Pooling (SPP)* дозволяє пулінг на різних масштабах і забезпечує однаковий вихідний вектор незалежно від розміру вхідного зображення. Це робить метод ефективним для роботи з даними різної розмірності без додаткового масштабування.

Зменшення розмірності в CNN сприяє зниженню обчислювальної складності, покращує ефективність і швидкість навчання, допомагаючи уникати перенавчання і підвищуючи здатність моделі до генералізації. Такий підхід дозволяє оптимально використовувати обчислювальні ресурси, а також підвищує здатність моделі до роботи з різними розмірами зображень.

Сучасні архітектури CNN, такі як ResNet, VGG, Inception та EfficientNet, демонструють різні підходи до покращення продуктивності згорткових нейронних мереж для складних задач, зокрема класифікації зображень. Кожна з цих архітектур має унікальні особливості, що дозволяють оптимально використовувати обчислювальні ресурси, підвищують точність і забезпечують гнучкість у виборі конфігурацій для різних застосунків.

Архітектура VGG (Рис.1.3) має відносно просту структуру, побудовану з послідовних 3x3 згорткових шарів із використанням max pooling для зменшення розмірності зображень, що дозволяє ефективно обробляти просторові ознаки.

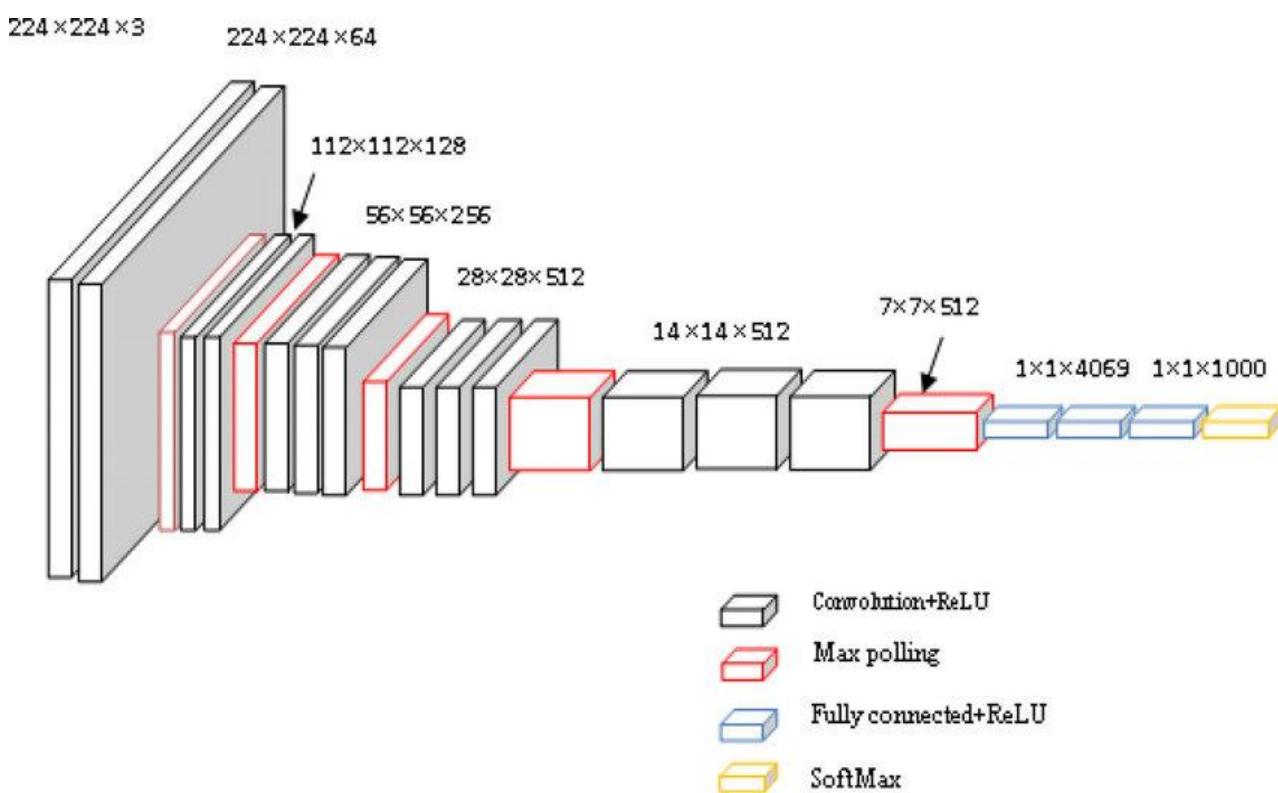


Рисунок 1.3 – Принцип роботи архітектури VGG

Головна перевага VGG полягає в його регулярності та простоті, що полегшує налаштування моделі, але її недоліком є значна кількість параметрів, що призводить до високого обчислювального навантаження.

Архітектура ResNet (Рис.1.4) внесла значне поліпшення в CNN завдяки впровадженню резидуальних блоків, що допомагають подолати проблему зникнення градієнта. Ці блоки містять короткі з'єднання або «шляхи» (skip connections), які дозволяють сигналу градієнта проходити через шари без значного згасання. Це дає змогу успішно тренувати дуже глибокі мережі, зберігаючи стабільність навчання і дозволяючи мережі досягати високої точності. Саме завдяки такій архітектурі ResNet досягла значних успіхів у розпізнаванні зображень та інших задачах комп'ютерного зору.

Інша вдосконалена архітектура, Inception (Рис.1.5), застосовує принцип використання багатомасштабних фільтрів на кожному шарі. Цей підхід дозволяє обирати найрелевантніші просторові ознаки на різних масштабах і поєднувати їх для кращої узагальненості моделі. Архітектура також оптимізує обчислювальні ресурси шляхом використання 1×1 згорток для зменшення розмірності даних, що знижує загальну складність і забезпечує високу ефективність. Завдяки цьому підходу Inception ефективно працює на великих наборах даних і надає модель, яка є гнучкою у налаштуванні для різних видів задач.

EfficientNet – ще одна інноваційна архітектура, яка розвиває ідеї масштабування моделі (Рис. 1.6). Основний принцип EfficientNet полягає в збалансованому масштабуванні глибини, ширини та роздільної здатності мережі. На відміну від традиційного підходу, що просто збільшує кількість шарів або ширину моделі, EfficientNet оптимально обирає баланс між цими параметрами, що дозволяє досягати високої точності з меншою кількістю параметрів і нижчими обчислювальними витратами. Це робить EfficientNet особливо ефективною для мобільних додатків і обчислювально обмежених середовищ, де обробка великих моделей може бути проблематичною.

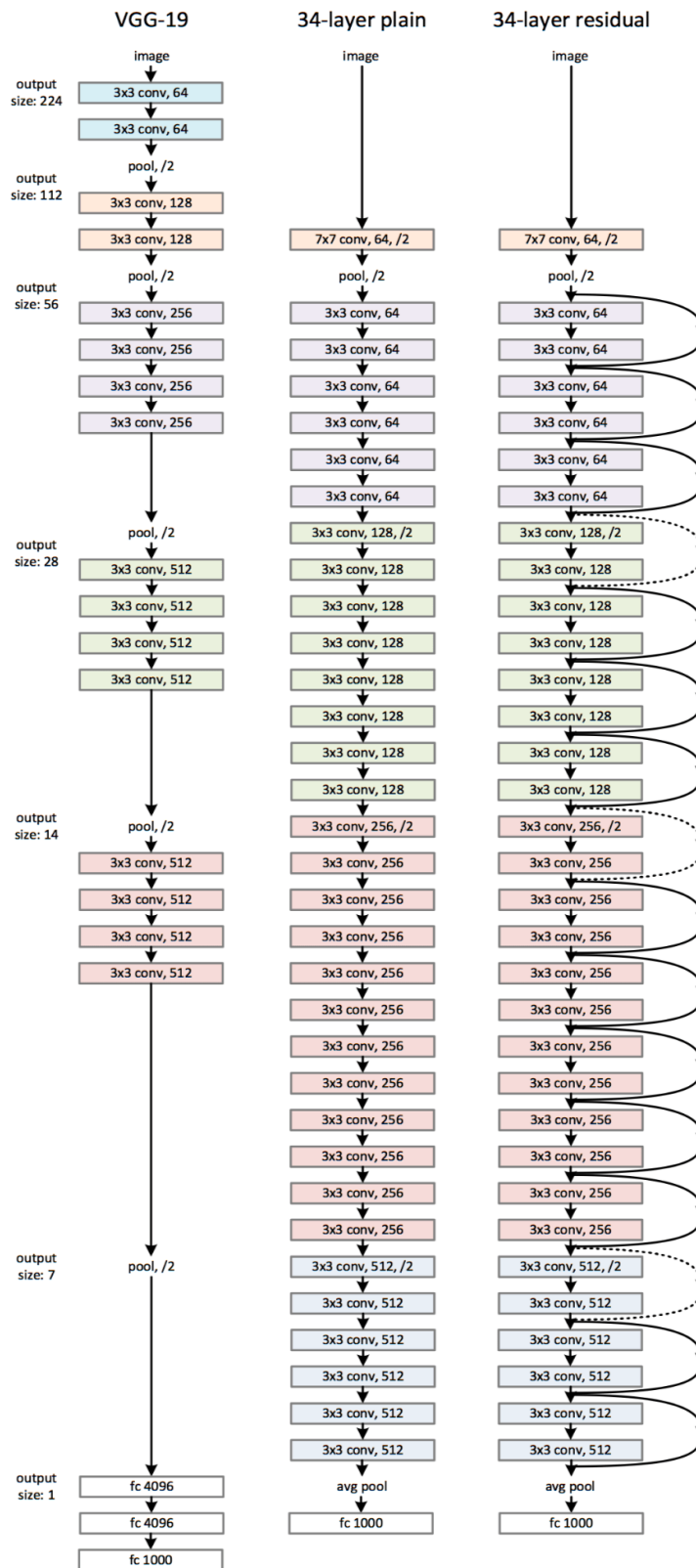


Рисунок 1.4 – Приклад архітектури ResNet. Зліва: модель VGG-19 (19,6 млрд. FLOP) як еталон. Посередині: проста мережа з 34 шарами (3,6 млрд. FLOP). Праворуч: ResNet з 34 шарами (3,6 мільярда FLOP). Пунктирні швидкі з'єднання збільшують розмірність.

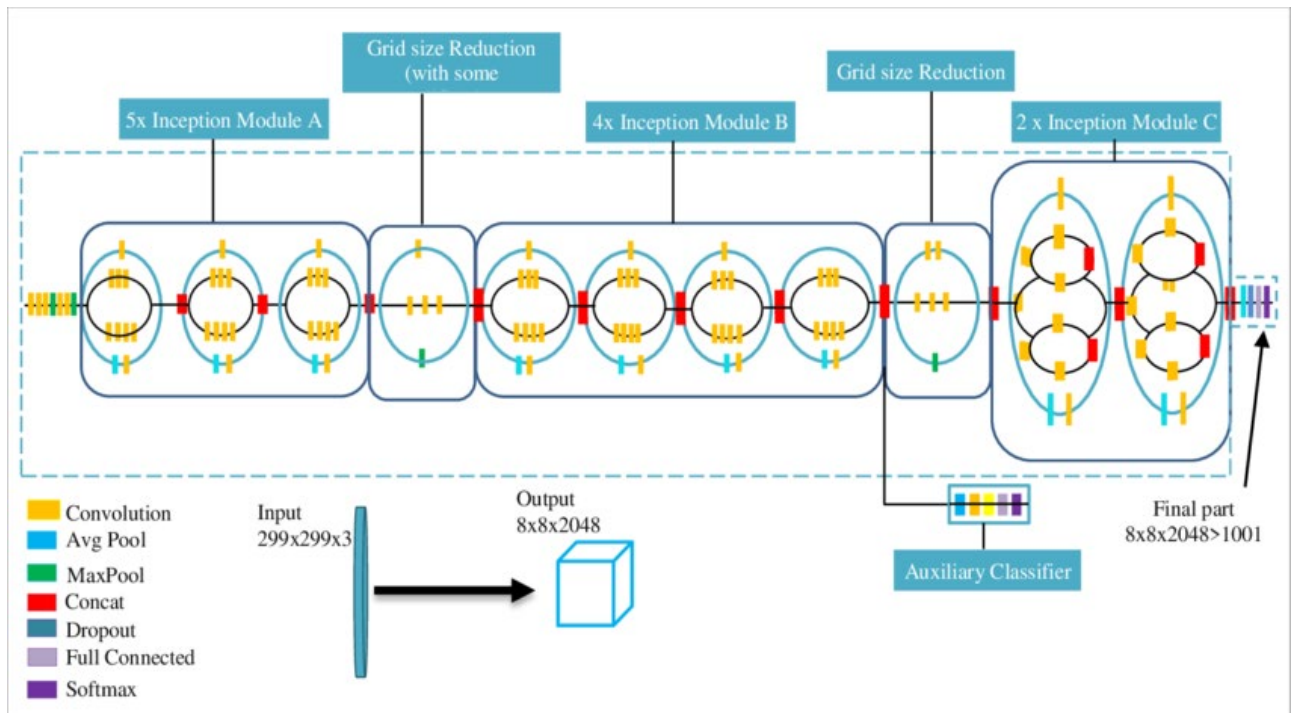


Рисунок 1.5 – Приклад роботи архітектури Inception

Розвиток цих архітектур демонструє, як нові рішення, такі як резидуальні блоки в ResNet, багатомасштабні фільтри в Inception або балансування параметрів в EfficientNet, дозволяють глибоким мережам ефективніше обробляти дані і розв'язувати проблему зникнення градієнта. Кожна з них зберігає різний підхід до обчислювальної ефективності і точності, що дозволяє обирати оптимальну архітектуру для конкретної задачі.

Сучасні згорткові нейронні мережі (CNN) стали основою для багатьох завдань комп'ютерного зору, проте вони також стикаються з рядом проблем і обмежень. Однією з основних проблем є чутливість до шуму та атак. CNN можуть демонструвати значні зниження точності при наявності шуму в даних або в умовах атак, таких як атаки зворотного поширення, коли зловмисники вносять незначні зміни в зображення, що призводить до невірних результатів. Це викликає занепокоєння в застосуваннях, де точність є критично важливою, таких як медична діагностика чи системи безпеки.

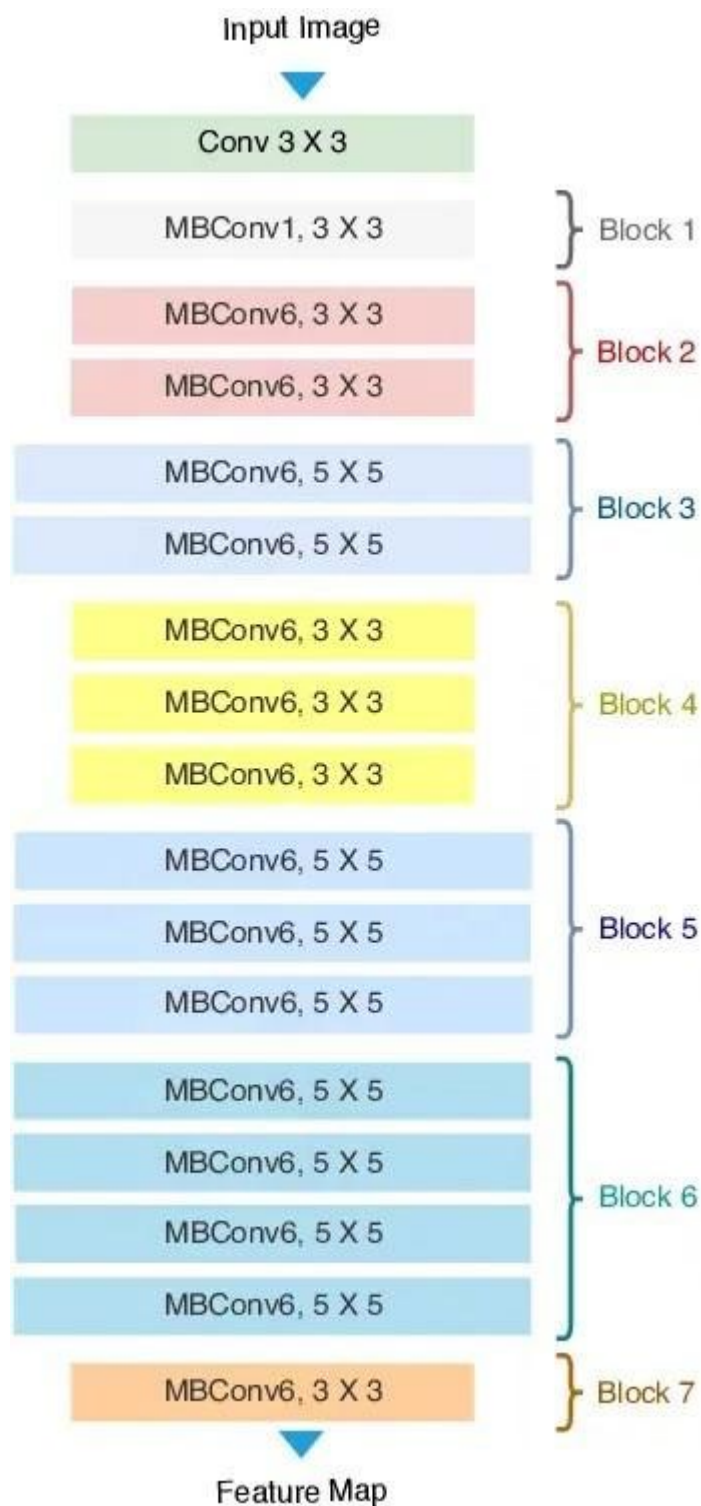


Рисунок 1.6 – Приклад роботи архітектури EfficientNet

“Згорткові нейронні мережі («CNN») вже багато років є невід’ємною частиною досліджень в галузі аналізу медичних зображень. Завдяки згортковим фільтрам, основною функцією яких є навчання та вилучення необхідних ознак з медичних зображень, CNN було присвячено безліч досліджень, починаючи від

виявлення та класифікації пухлин, виявлення уражень шкіри і закінчуючи сегментацією пухлин головного мозку, і це лише деякі з них. CNN також зробили значний внесок в аналіз різних методів візуалізації в клінічній медицині, включаючи рентгенографію, комп'ютерну томографію (КТ), магнітно-резонансну томографію (МРТ), ультразвук (УЗД) і цифрову патологію. Незважаючи на свою видатну продуктивність, CNN страждають від концептуальних обмежень і вроджено не здатні моделювати явні залежності на великих відстанях через обмежене сприйнятливое поле ядер згортки. Крім того, згортковий оператор страждає від того, що під час виведення він застосовує фіксовані значення ваг незалежно від будь-яких змін у візуальному вхідному сигналі. Щоб пом'якшити вищезгадані проблеми, були проведені великі дослідження з інтеграції механізмів уваги, які можна розглядати як процес динамічного коригування ваг на основі вхідних ознак до основоположних структур на основі CNN для покращення можливостей нелокального моделювання.”[9, с. 1]

Інтерпретація результатів, отриманих за допомогою CNN, також є значною проблемою. Багато з цих моделей працюють як "чорні скриньки", що ускладнює розуміння того, як вони приймають рішення. Це особливо важливо в контексті медичних і юридичних застосувань, де важливо не лише отримати правильний результат, а й зрозуміти, як був досягнутий цей результат. Без зрозумілої інтерпретації користувачі не можуть бути впевнені в надійності та точності моделей.

Великі обчислювальні витрати є ще одним обмеженням CNN. Навчання глибоких нейронних мереж вимагає значних обчислювальних ресурсів, що може бути дорогим і вимагати багато часу. Це може обмежувати їх використання на пристроях з обмеженими ресурсами, таких як мобільні телефони або вбудовані системи.

Нові підходи, такі як Capsule Networks (CapsNets) та Vision Transformers (ViT), пропонують потенційні рішення для цих проблем. Capsule Networks намагаються вирішити питання інтерпретації, впроваджуючи концепцію капсул,

які представляють об'єкти або частини об'єктів, дозволяючи моделі розуміти відносини між ними. Це дозволяє зменшити чутливість до трансформацій і покращити стійкість до атак, а також підвищити рівень інтерпретації.

Vision Transformers, своєю чергою, запроваджують новий підхід до обробки зображень, використовуючи механізми уваги, що дозволяє моделі фокусуватися на найважливіших частинах зображення. Це може зменшити потребу в глибоких шарах, що знижує обчислювальні витрати і підвищує ефективність. ViT також демонструють потенціал у контексті узагальнення на нові дані, що робить їх привабливими для різних задач у комп'ютерному зорі.

“Парк та ін. пропонують метод на основі ViT для діагностики COVID-19, використовуючи низькорівневі ознаки CXR, витягнуті з попередньо навченої опорної мережі. Магістральна мережа була навчена в самоконтрольованій манері (з використанням методу SimCLR на основі контрастного навчання) для вилучення аномальних включень CXR-характеристик з великого і добре контрольованого набору даних CXR від CheXpert. Ці вбудовування ознак були використані моделлю ViT для високорівневої діагностики зображень COVID-19. Широкі експерименти на трьох тестових наборах даних CXR, отриманих з різних лікарень, демонструють перевагу їхнього підходу порівняно з моделями на основі CNN.” [\[8, с. 13\]](#)

Отже, хоча CNN мають ряд проблем і обмежень, нові підходи, такі як Capsule Networks та Vision Transformers, відкривають нові перспективи в розвитку технологій, що дозволяють краще справлятися з чутливістю до шуму, інтерпретацією результатів та обчислювальними витратами.

1.3.2 Методи попередньої обробки знімків

Підготовка даних і аугментація є важливими етапами в підвищенні продуктивності та стійкості згортової нейронної мережі (CNN). Попередня обробка, як-от нормалізація та масштабування, полягає в стандартизації діапазону значень пікселів, що стабілізує вхідні дані, роблячи їх однорідними.

Наприклад, значення пікселів можуть масштабуватися у діапазон 0–1 або -1–1, зменшуючи варіативність і сприяючи швидшому навчанні моделі.

Аугментація збільшує обсяг і варіативність навчальних даних за рахунок різних трансформацій. Це допомагає моделі краще узагальнювати інформацію, знижуючи ризик перенавчання. До типових методів аугментації належать обертання зображень для підвищення стійкості до зміни положення об'єктів, масштабування для додання різноманітності розмірів, зміна яскравості та контрасту для імітації різних умов освітлення, а також перевертання та зсув для посилення стійкості до змін у формі об'єктів. Додаткові методи, такі як обрізання чи додавання шуму, імітують реальні умови, коли об'єкти на зображенні можуть бути частково приховані або містити дефекти.

Поєднання попередньої обробки та аугментації підвищує продуктивність CNN, створюючи насичене й різноманітне навчальне середовище. Це допомагає моделі краще узагальнювати знання, стійкіше обробляти нові дані та знижує залежність від специфічних особливостей тренувального набору, роблячи її здатною до роботи в умовах реального світу.

Методи видалення шуму та згладжування є важливими для попередньої обробки зображень, оскільки зменшують вплив випадкових артефактів, які можуть знижувати якість аналізу та перешкоджати виділенню основних ознак. Одним із найбільш розповсюджених методів згладжування є гауссовий фільтр, який використовує принцип згладжування з використанням гауссового розподілу, де пікселі зображення підсумовуються відповідно до їхньої віддаленості від центрального пікселя. Ближчі пікселі мають більшу вагу, тому шум на межах згладжується, зберігаючи чіткість контурів. Гауссовий фільтр є ідеальним для зображень, де важливо зберегти інформацію про краю об'єктів, наприклад, в обробці фотографій об'єктів на фоні.

Медіанний фільтр є ще одним ефективним методом, особливо корисним для зображень із salt and pepper шумом, який виглядає як випадкові темні та світлі точки. На відміну від гауссового фільтра, медіанний фільтр не підсумовує сусідні пікселі, а сортує їх за інтенсивністю і обирає медіанне значення як нове

значення центрального пікселя. Це дозволяє медіанному фільтру зберігати контури країв об'єктів, усуваючи при цьому шумові артефакти. Він добре підходить для медичних або промислових зображень, де чіткість меж є критично важливою.

Інші методи згладжування, такі як середнє усереднення або білінеарне згладжування, також можуть застосовуватися, але вони не так ефективно зберігають контури, як гауссовий або медіанний фільтри. Ці методи можуть створювати розмиття на краях об'єктів, що обмежує їхню застосовність у випадках, коли важлива висока деталізація меж.

Загалом, видалення шуму та згладжування дозволяють зробити зображення більш чітким, зменшивши вплив випадкових перешкод, що особливо важливо для зображень з реальних камер, де рівень шуму може бути значним.

Методи підвищення контрасту дозволяють зробити зображення більш чітким і покращити видимість деталей, особливо в умовах, коли об'єкти мають низький контраст або зображення знято в неоднорідному освітленні. Одним із найбільш поширених методів є ****рівномірний розподіл гістограми****, який вирівнює інтенсивності пікселів на зображенні так, щоб яскравість розподілялася більш рівномірно по всьому діапазону, що надає зображенню більш насичений вигляд. Цей метод працює шляхом перетворення рівнів інтенсивності таким чином, щоб частоти кожного рівня були максимально вирівняні. Це особливо корисно для зображень з темними або світлими областями, оскільки дає змогу покращити видимість деталей в обох діапазонах.

Адаптивне вирівнювання гістограми контрасту (CLaNE) – удосконалена версія стандартного вирівнювання гістограми, що працює більш точно на локальних ділянках зображення, розділяючи його на невеликі блоки і виконуючи вирівнювання окремо для кожного блоку. Це дозволяє зберегти природний вигляд зображення без надмірного підсилення контрасту на більш однорідних ділянках. CLaNE також містить обмеження на максимальне підсилення контрасту, що запобігає появі надмірно яскравих або темних областей. Завдяки адаптивності, цей метод добре підходить для складних умов освітлення або

зображень з деталями різного масштабу, як-от медичні знімки або зображення з аерофотозйомки.

Методи підвищення контрасту, такі як вирівнювання гистограми та CLAHE, особливо корисні в задачах розпізнавання об'єктів і сегментації, оскільки дозволяють виділити важливі деталі, які можуть бути не помітними при початковій яскравості та контрасті.

Методи вирівнювання зображень є ключовими для забезпечення стабільності та однорідності в аналізі, оскільки вони допомагають скоригувати позицію об'єктів, які можуть бути зняті під різними кутами чи зі зміщенням. Це особливо важливо для комп'ютерного зору та нейронних мереж, які потребують чіткого й послідовного представлення об'єктів.

Один із простих підходів — обрізання зображень, яке дає змогу зосередитися на центральній області кадру, що найчастіше містить основний об'єкт. За допомогою цього методу можна вирізати зайві елементи фону, які можуть заважати розпізнаванню та збільшують обчислювальні витрати. Інший метод – обертання зображень, яке дозволяє вирівняти об'єкти до горизонтальної або вертикальної орієнтації, усуваючи випадкові кути, що можуть ускладнювати обробку.

Для автоматичного вирівнювання зображень використовуються більш складні підходи, такі як оптичний потік або відстеження ключових точок. Оптичний потік вимірює зміщення пікселів між двома послідовними кадрами, дозволяючи коригувати зміщення об'єктів або камери. Відстеження ключових точок включає ідентифікацію характерних особливостей, таких як кути або контури, для подальшого вирівнювання на основі відповідності точок між зображеннями. Ці методи корисні для випадків, коли потрібна точна корекція положення або коли об'єкти мають різні перспективи.

Загалом, методи вирівнювання допомагають стандартизувати положення об'єктів, забезпечуючи послідовність при подачі даних до моделей. Це дозволяє зменшити залежність від початкової позиції об'єкта, покращуючи загальну точність та ефективність аналізу.

Нормалізація кольорів і корекція гамми є важливими етапами підготовки зображень для їх подальшої обробки та аналізу. Нормалізація кольорів дозволяє стандартизувати колірну палітру зображень, що сприяє стабільнішому аналізу в різних умовах освітлення або при відмінностях у налаштуваннях камери. Наприклад, за допомогою нормалізації можна перетворити зображення в певний колірний простір (як-от HSV чи YCbCr) і налаштувати інтенсивність кольорів так, щоб уникнути надмірної насиченості або спотворення відтінків, що може вплинути на розпізнавання об'єктів.

Корекція гамми є процесом налаштування яскравості пікселів, що допомагає покращити контрастність зображень. Вона застосовується шляхом піднесення значень інтенсивності пікселів до степеня гамма, що дозволяє змінювати загальний тон зображення: підвищення гамма-корекції робить темні ділянки яскравішими, а зниження — темнішим. Це особливо корисно в умовах низької чи надмірної освітленості, коли зображення може втрачати деталі в тінях або світлих ділянках. Корекція гамми робить зображення більш придатними для аналізу, забезпечуючи краще розпізнавання дрібних деталей.

Застосування нормалізації кольорів і корекції гамми дозволяє отримати більш стабільні дані для аналізу, незалежно від зовнішніх факторів. Це, своєю чергою, покращує точність обробки, зменшує вплив зовнішнього шуму та підвищує загальну якість розпізнавання об'єктів або деталей.

Усунення аномалій та сегментація є важливими етапами обробки складних зображень, особливо коли йдеться про аналіз даних з високим рівнем деталізації, як-от медичні або супутникові знімки. Сегментація дозволяє виділити релевантні області зображення, де містяться ключові об'єкти чи структури, відокремивши їх від фонового шуму, який може бути інформаційно некорисним або навіть заважати точності аналізу.

Застосування методів сегментації забезпечує більш ефективний аналіз зображень, дозволяючи моделі фокусуватися на конкретних ділянках, що містять релевантні деталі. Наприклад, у медичних знімках сегментація може виділити

певні органи чи тканини, зменшуючи обсяг інформації для аналізу та допомагаючи відстежувати патологічні зміни або аномалії. У супутникових зображеннях цей метод дає змогу виділити лише необхідні області, такі як міські зони чи природні утворення, що дозволяє покращити класифікацію або оцінку ресурсів.

Усунення аномалій включає техніки, які допомагають видалити небажані артефакти або шуми, які можуть ускладнити процес розпізнавання. Це особливо важливо при роботі з даними, де можливі значні відхилення в колірній палітрі, структурі або текстурі, як-от у зображеннях з діагностичних апаратів чи супутникових систем. Усунення таких аномалій дозволяє мінімізувати вплив нехарактерних елементів і зосередитися на аналізі необхідної інформації.

Загалом, попередня сегментація та усунення аномалій значно покращують якість обробки, знижують ризик помилок і забезпечують моделі оптимальні вхідні дані для аналізу складних структур. Це створює можливості для точнішої оцінки, більш ефективного використання ресурсів і зменшує час, необхідний для аналізу великих обсягів зображень.

Попередня обробка знімків є критично важливим етапом в процесі аналізу медичних зображень, особливо у контексті нейромережових технологій. Правильна обробка зображень може значно покращити якість діагностики. Колоризація передбачає додавання кольору до чорно-білих зображень для підвищення візуальної інформативності та чіткості. Автоматизовані методи колоризації використовують алгоритми машинного навчання для автоматичного кольорування медичних знімків, що дозволяє зекономити час та ресурси.

Реконструкція деталей включає інтерполяцію, яка використовує математичні алгоритми для заповнення пропусків у зображеннях, які можуть виникати через артефакти чи дефекти сканування. Патерн-реконструкція дозволяє відновити відсутні частини зображення шляхом аналізу наявних елементів, що підвищує точність візуалізації. Контекстуальна реконструкція застосовує алгоритми, які враховують загальний контекст зображення, що дозволяє краще відновити структури органів або тканин.

Вдосконалення якості зображення включає фільтрацію, яка усуває шум і покращує чіткість, що є важливим для точної діагностики. Розширення динамічного діапазону (HDR) також сприяє покращенню контрастності та насиченості кольорів, забезпечуючи більшу чіткість медичних знімків. Застосування зазначених методів попередньої обробки знімків забезпечує покращену якість даних для нейромережових алгоритмів, що своєю чергою веде до підвищення точності медичної діагностики та ефективності аналізу.

Не менш важливим етапом попередньої обробки зображень є виділення контурів, оскільки воно дозволяє виявити і проаналізувати значущі структури в медичних зображеннях. Використання операторів градієнта, таких як Собель і Кана, є одними з найбільш популярних методів для цього.

Оператор Собеля працює шляхом обчислення градієнта зображення в двох напрямках — горизонтальному та вертикальному. Це дозволяє виявити зміни яскравості, які вказують на наявність країв. Результат дії оператора Собеля є двома зображеннями, які представляють градієнти у горизонтальному та вертикальному напрямках. Комбінуючи ці градієнти, можна отримати зображення, що показує контури об'єктів.

Оператор Кана, на відміну від Собеля, спочатку застосовує фільтрацію, щоб зменшити шум у зображенні. Це робить його більш стійким до артефактів, які можуть спотворити результати виділення країв. Кана використовує концепцію градієнта, але також враховує інформацію про напрямок, у якому відбувається зміна яскравості. Результат — зображення, яке чітко демонструє контури об'єктів, що робить його особливо корисним у медичних зображеннях, де важливо виявити аномалії, такі як пухлини або інші патології.

Обидва методи, Собеля і Кана, можуть бути використані в комбінації з іншими техніками попередньої обробки, такими як згладжування або нормалізація, для покращення результатів виділення контурів. У медичній діагностиці виділення контурів дозволяє лікарям та аналітикам зосередитися на важливих деталях, що може сприяти точнішій і швидшій діагностиці.

Приведення всіх зображень до одного розміру дозволяє уникнути численних проблем, які можуть виникнути в подальшому обробці і аналізі даних.

По-перше, стандартні розміри зображень забезпечують однорідність даних, що є важливим для навчання моделей машинного навчання. Багато алгоритмів, включаючи згорткові нейронні мережі (CNN), вимагають, щоб всі вхідні дані мали однакові розміри, оскільки різні розміри можуть призвести до помилок під час обчислення. Це також спрощує процес підготовки даних, дозволяючи моделі швидше обробляти інформацію та підвищуючи ефективність навчання.

По-друге, стандартизація розмірів зображень полегшує аналіз і візуалізацію даних. При роботі з великими наборами медичних зображень, такими як рентгенівські знімки або МРТ, стандартизовані розміри дозволяють легше порівнювати результати, виявляти патерни та аномалії, а також здійснювати статистичний аналіз. Крім того, важливо враховувати аспекти збереження пропорцій зображення. У деяких випадках, наприклад, при обробці анатомічних структур, необхідно зберегти пропорції, щоб уникнути спотворення. Це може бути досягнуто за допомогою технік, таких як додавання полів або зміна розміру зображень з використанням методів, які зберігають співвідношення сторін.

Зберігання та управління метаданими також важливий аспект роботи з медичними зображеннями, оскільки вони надають контекстуальну інформацію, необхідну для ефективного аналізу і використання даних. Метадані можуть включати різноманітну інформацію, таку як параметри зйомки, характеристики обладнання, умови освітлення, а також деталі обробки зображень.

По-перше, метадані забезпечують важливий контекст для розуміння того, як були отримані зображення. Інформація про дату і час зйомки, тип обладнання (наприклад, рентген, МРТ або КТ), а також налаштування, такі як експозиція та дозування, можуть суттєво вплинути на якість та точність зображень. Це допомагає лікарям і аналітикам оцінювати надійність і значимість отриманих даних.

По-друге, зберігання метаданих є критично важливим для повторюваності досліджень. Вчені та медичні працівники можуть покладатися на цю інформацію, щоб повторно відтворити дослідження або порівняти нові результати з попередніми. Це особливо важливо у випадках, коли потрібно провести порівняння між різними зображеннями або дослідженнями, що мають схожі характеристики.

По-третє, метадані можуть бути використані для покращення процесів аналізу даних. Наприклад, при застосуванні алгоритмів машинного навчання важливо знати, які зображення є навчальними, а які — тестовими. Включення метаданих, що вказують на тип зображення, дозволяє краще організувати дані і підвищити ефективність навчання моделей.

Крім того, організація метаданих допомагає в управлінні даними та їх архівації. Вона забезпечує легкий доступ до необхідної інформації для медичних працівників, що може пришвидшити процес прийняття рішень і полегшити роботу з архівами зображень.

Інтеграція медичних зображень з іншими видами даних, такими як геномні чи клінічні дані, є важливим кроком до створення комплексної картини стану пацієнта та підвищення точності діагностики і лікування. Такий підхід дозволяє лікарям та дослідникам отримати глибше розуміння патологій та механізмів хвороб.

По-перше, комбінування медичних зображень з геномними даними може значно покращити можливості діагностики. Наприклад, аналіз генетичних маркерів, пов'язаних з певними захворюваннями, у поєднанні з візуалізацією пухлин на МРТ чи КТ може допомогти виявити специфічні мутації, які впливають на розвиток і прогресування хвороби. Це може дозволити лікарям краще адаптувати лікування під конкретні генетичні профілі пацієнтів.

По-друге, інтеграція медичних зображень з клінічними даними, такими як історія хвороби, результати лабораторних тестів і дані про лікування, може допомогти в оцінці ефективності терапії. Це дозволяє лікарям відстежувати, як зміни в зображеннях, наприклад, зменшення розміру пухлини, корелюють з

покращенням клінічних показників, такими як рівень маркерів у крові або загальний стан здоров'я пацієнта.

По-третє, з використанням алгоритмів машинного навчання та аналітики великих даних, інтеграція різних типів даних може покращити процес прийняття рішень. Такі системи можуть автоматично виявляти патерни, які можуть залишитися непоміченими при традиційному аналізі. Наприклад, комбінація зображень з історією захворювань може виявити нові фактори ризику, які стосуються певних груп пацієнтів.

Крім того, інтеграція даних може сприяти розвитку персоналізованої медицини, де лікування налаштовується відповідно до індивідуальних потреб пацієнта. Це стає можливим завдяки створенню більш повного профілю пацієнта, що включає не лише зображення, а й генетичну інформацію, дані про спосіб життя та клінічні показники.

1.3.3 Використання технік навчання та перенавчання нейромереж для медичних задач

Для успішного застосування нейромереж у медичних задачах важливо враховувати специфічні техніки навчання та перенавчання, які можуть покращити точність, надійність і стабільність моделей.

Машинне навчання поділяється на кілька основних видів, кожен з яких використовує певні підходи для навчання моделі і має свої особливості. Основні види машинного навчання — це навчання з учителем, навчання без учителя, напівкероване навчання і підкріплене навчання.

Навчання з учителем є одним з найбільш поширених підходів. У цьому виді навчання моделі надається набір даних, де кожен зразок має відповідну мітку, тобто відомий правильний результат. Завдання моделі — вивчити залежності між вхідними даними і мітками, щоб згодом коректно передбачати результати для нових, раніше не бачених даних. Такий підхід часто застосовується у задачах класифікації, як-от розпізнавання об'єктів на

зображеннях, або регресії, як-от передбачення числових значень, наприклад цін на акції.

Навчання без учителя не передбачає міток у вихідних даних. Модель аналізує вхідні дані і намагається виявити в них структуру або закономірності. До завдань навчання без учителя належать кластеризація, коли модель групує дані за певними подібностями, і пошук закономірностей для подальшого аналізу. Наприклад, у маркетингу моделі кластеризації використовують для поділу клієнтів на сегменти.

Напівкероване навчання є поєднанням підходів навчання з учителем і без учителя. У такому навчанні модель працює з набором даних, де частина даних має мітки, а інша частина — ні. Напівкероване навчання дозволяє суттєво зменшити потребу в розмічених даних, оскільки моделі здатні використовувати не розмічені приклади для підвищення своєї продуктивності. Цей підхід ефективний у ситуаціях, коли розмічування даних вимагає великих затрат часу і ресурсів, як, наприклад, у медичній діагностиці.

Підкріплене навчання базується на взаємодії агента з навколишнім середовищем. Агент навчається через систему винагород і штрафів, оптимізуючи свої дії для максимального отримання винагород. Підхід підкріпленого навчання зазвичай використовується у складних завданнях, де агент має досягти певної цілі шляхом серії дій, як-от у грі в шахи або в робототехніці, де важливо знаходити оптимальні стратегії поведінки для агентів у різних умовах.

Крім основних видів навчання існують й інші методи, що розширюють можливості машинного навчання.

Само-навчання є підходом, коли модель використовує власні передбачення як основу для подальшого навчання. Спочатку модель навчається на невеликому наборі міток, а потім використовує власні передбачення для позначення нового набору даних, які допомагають розширити обсяг навчальних даних. Це дозволяє підвищити ефективність у ситуаціях, коли обмежені розмічені дані.

Багатозадачне навчання дозволяє моделі навчатися одночасно кільком взаємопов'язаним завданням. Наприклад, у медичних дослідженнях багатозадачні моделі можуть одночасно вивчати кілька аспектів стану пацієнта, як-от розпізнавання типу захворювання і передбачення його перебігу. Цей підхід дозволяє моделям узагальнювати знання і поліпшує їхню здатність до переносу інформації між подібними завданнями.

Навчання з підказками або *weakly supervised learning* — це підхід, коли модель отримує не повністю розмічені дані або використовує інформацію з мінімальною кількістю наданих міток. Моделі працюють на основі більш загальних позначок, які можуть мати форму міток на рівні категорії або на рівні набору даних. Такий підхід часто застосовується для великих наборів даних, де повне маркування є складним.

Метанавчання або навчання навчатися (*learning to learn*) допомагає моделям швидко адаптуватися до нових завдань, навіть з невеликим обсягом даних. Цей підхід націлений на створення моделей, здатних до швидкого навчання нових завдань за допомогою досвіду, отриманого на попередніх завданнях. Метанавчання активно використовується в ситуаціях, де дані можуть швидко змінюватися, і є необхідність у швидкій адаптації моделі.

Крім того, дедалі частіше використовуються підходи ансамблевого навчання, де об'єднується кілька моделей для досягнення кращих результатів. Метод ансамблю дозволяє зменшити похибку та підвищити надійність, оскільки різні моделі взаємно доповнюють і компенсують слабкі сторони одна одної.

Адаптивне та безперервне навчання є важливими підходами, особливо коли йдеться про системи, які потребують динамічної адаптації до нових даних.

Адаптивне навчання дозволяє моделі змінювати свої параметри або структуру у відповідь на нові вхідні дані чи зміну середовища. Це корисно для додатків, де вхідні дані можуть змінюватися з часом. Наприклад, у сфері фінансового прогнозування модель може адаптуватися до нових ринкових умов. Такі адаптивні моделі можуть навчатися на нових даних або навіть на "складних

прикладів" — випадках, де модель раніше давала неправильні передбачення. Це підвищує точність і адаптивність системи.

Безперервне навчання (continual learning) є підходом, коли модель поступово навчається на нових даних, зберігаючи при цьому знання з попередніх етапів навчання. Це дозволяє моделі не лише засвоювати нову інформацію, але й уникати «забування» попередньої. Безперервне навчання є цінним у ситуаціях, де дані змінюються з часом, як-от в охороні здоров'я або в робототехніці. Наприклад, система може постійно оновлювати свої знання про нові медичні дані чи методи лікування, зберігаючи актуальність і точність прогнозів. Обидва підходи роблять моделі більш стійкими до змін та поліпшують їхню здатність до довготривалого застосування в умовах, де статичні моделі швидко втрачають актуальність.

Перенесене навчання (transfer learning) – це підхід, при якому модель, навчена на одному завданні або наборі даних, використовується для іншого завдання, зокрема у випадках, коли є обмеження на кількість даних для нового завдання. Такий підхід дозволяє ефективно перенести знання, отримані від моделі, навченої на великому наборі даних, на інші суміжні або подібні задачі. Це значно скорочує час і обчислювальні ресурси, необхідні для навчання. У глибокому навчанні перенесене навчання зазвичай використовується з попередньо натренованими нейронними мережами, такими як ResNet, VGG або BERT, які були навчені на великих наборах даних, як-от ImageNet (для зображень) або масивів текстів (для обробки природної мови). У процесі перенесеного навчання мережу часто адаптують для нового завдання шляхом тонкого налаштування (fine-tuning), тобто модифікації останніх шарів моделі, щоб оптимізувати її для специфічного завдання. Наприклад, модель, навчена на класифікації об'єктів, може бути адаптована для розпізнавання конкретних медичних аномалій на знімках МРТ або для аналізу фінансових даних. Перенесене навчання особливо корисне у випадках, коли доступні обмежені або коштовні дані, наприклад у медичних зображеннях, де розмічування може вимагати значних зусиль експертів. Цей підхід не лише зменшує обсяг даних,

необхідних для навчання, але також може підвищити точність і ефективність моделі завдяки попередньому навчанню на великих наборах даних.

Сукупне використання цих методів допомагає покращити якість моделей, підвищує їхню точність, узагальнюваність та ефективність при розв'язанні складних медичних задач. Це забезпечує точніші та надійніші прогнози, що особливо важливо для клінічних застосувань, де помилка може коштувати життя пацієнта.

1.3.4 Методи оцінки точності роботи нейромереж на медичних знімках

Оцінка точності роботи нейронних мереж на медичних знімках є критично важливим етапом їх розробки та впровадження в клінічну практику.

Крос-валідація є потужним методом оцінки точності моделей, включаючи нейронні мережі, в медичній діагностиці. Вона передбачає розподіл набору даних на кілька підмножин, званих складками. Зазвичай, дані ділять на k складок, де k може варіюватися залежно від розміру набору даних. Під час процесу крос-валідації модель тренується на $k-1$ складках, використовуючи їх для навчання, а залишена складка використовується для тестування. Цей процес повторюється k разів, кожного разу змінюючи складку, яка слугує тестовою. Завдяки цьому методі, кожна частина даних має можливість бути використаною як для навчання, так і для тестування, що дозволяє отримати більш надійні оцінки точності моделі. Вона зменшує ризик перенавчання, оскільки модель перевіряється на різних підмножинах даних. Це забезпечує більш реалістичну оцінку її здатності до узагальнення на нові, невідомі дані. Результати крос-валідації можуть бути об'єднані в середнє значення, що відображає загальну продуктивність моделі. Цей метод є особливо корисним у медичних застосуваннях, де обмежена кількість даних може призвести до нестабільних оцінок точності. Також це допомагає виявити потенційні проблеми, такі як упередженість даних або дисбаланс класів, що може негативно вплинути на результати. Використання цього методу дозволяє

отримати більш об'єктивну та точну оцінку роботи нейронної мережі, що є важливим для її подальшого впровадження в клінічну практику.

Метрики оцінки якості моделей є критично важливими для визначення ефективності нейронних мереж, особливо у медичних завданнях, де правильність діагнозу має надзвичайне значення. Оцінювання медичних моделей вимагає більш комплексного підходу, ніж просто обчислення точності, оскільки потрібно враховувати помилкові позитивні та негативні прогнози, які можуть мати серйозні наслідки для пацієнтів.

Точність (accuracy) є базовою метрикою, яка визначає відсоток правильних прогнозів відносно загальної кількості. Проте в медичних завданнях з нерівномірно розподіленими класами, наприклад, при діагностиці рідкісних захворювань, точність може бути оманливо високою. Наприклад, якщо модель завжди прогнозує відсутність хвороби, вона матиме високу точність за умов великої кількості здорових випадків, навіть якщо не зможе виявити хворобу.

Повнота (recall) показує, який відсоток справжньо позитивних результатів модель змогла правильно виявити. У медичних застосуваннях, особливо при діагностиці критичних захворювань, важливо мінімізувати ймовірність помилково негативних результатів, тобто випадків, коли хвороба не виявляється у хворих пацієнтів. Висока повнота є бажаною для уникнення пропуску небезпечних станів.

Специфічність (specificity) вимірює здатність моделі правильно визначати здорових пацієнтів. Висока специфічність важлива для зменшення кількості помилкових позитивних результатів, які можуть призводити до необґрунтованих медичних втручань. У задачах, де непотрібне лікування може мати негативні наслідки, специфічність є ключовою метрикою.

Точність передбачення (precision) визначає, скільки з передбачених позитивних результатів є справді позитивними. Ця метрика важлива у випадках, коли критичним є уникнення помилкових позитивних діагнозів, що можуть спричинити зайві діагностичні та лікувальні процедури.

AUC-ROC (Area Under the Receiver Operating Characteristic Curve) є універсальною метрикою, яка відображає здатність моделі розрізняти позитивні та негативні випадки при різних порогах. AUC значення ближче до 1 свідчить про високу точність моделі при розрізненні між класами. ROC-крива дозволяє врахувати компроміс між чутливістю та специфічністю, що є особливо корисним для вибору оптимального порогу для медичної діагностики.

F1-міра є об'єднанням повноти та точності передбачення, дозволяючи балансувати між цими метриками. Ця метрика є особливо цінною при роботі з нерівномірно розподіленими класами, коли важливо знайти рівновагу між точністю та здатністю моделі не пропускати випадки захворювання.

Сукупне використання цих метрик дозволяє отримати більш повне уявлення про ефективність моделі у медичних завданнях. Це забезпечує надійність та безпечність прогнозів, дозволяючи моделі адаптуватися до складних реальних умов, де помилка в діагностиці може суттєво вплинути на здоров'я і життя пацієнтів.

Аналіз помилок є невід'ємною частиною процесу оцінки моделей, що базуються на нейронних мережах, у медичній діагностиці. Дослідження типів помилок, які робить модель, дозволяє глибше зрозуміти, які аспекти зображень призводять до неправильних класифікацій. Цей процес включає вивчення помилок класифікації, наприклад, чи модель плутає одні захворювання з іншими, або ж не виявляє наявності певних ознак на зображеннях.

Одним із методів, що можуть бути використані для візуалізації помилок, є Grad-CAM (Gradient-weighted Class Activation Mapping). Цей інструмент дозволяє визначити, які частини зображення найбільше вплинули на прийняття рішення нейронної мережі. За допомогою Grad-CAM можна створити теплові карти, які відображають області зображення, що сприяють класифікації. Це особливо корисно в медичних застосуваннях, де зрозуміти причини помилок може бути критично важливим для покращення точності діагностики.

Аналіз помилок не лише допомагає виявити слабкі місця моделі, а й відкриває можливості для вдосконалення навчання. Наприклад, якщо модель

часто помиляється в розпізнаванні певних патернів, це може свідчити про потребу в додаткових даних для навчання або в адаптації архітектури моделі. Крім того, розуміння помилок може допомогти лікарям і дослідникам отримати додаткові інсайти про самі зображення та їхні характеристики, що, в свою чергу, може призвести до кращого розуміння медичних станів.

Також не можна забувати за етап дослідження порогів, особливо в медичній діагностиці, де правильне визначення порогів може значно вплинути на результати. Оптимальні пороги для класифікації допомагають збалансувати чутливість (здатність моделі виявляти позитивні випадки) і специфічність (здатність моделі виявляти негативні випадки).

Чутливість та специфічність є ключовими показниками, які визначають ефективність моделі. Висока чутливість важлива, коли необхідно зменшити кількість пропущених позитивних випадків, наприклад, у випадках, коли рання діагностика захворювання може врятувати життя. Проте висока чутливість може супроводжуватися підвищеним рівнем хибнопозитивних результатів, що може призвести до зайвих обстежень або лікування. Специфічність, з іншого боку, допомагає уникнути хибнопозитивних висновків, проте надмірна орієнтація на неї може призвести до того, що деякі позитивні випадки залишаться непоміченими. Визначення оптимальних порогів передбачає аналіз різних значень, за яких модель класифікує зразки як позитивні чи негативні. Це може бути виконано шляхом побудови кривих ROC (Receiver Operating Characteristic) та обчислення площі під кривою (AUC). Криві ROC відображають співвідношення між чутливістю та специфічністю для різних порогів класифікації. Оптимальний поріг може бути обраний в залежності від клінічних потреб, щоб досягти бажаного балансу між чутливістю і специфічністю, що важливо для конкретного випадку використання. Зважаючи на важливість точності діагностики у медицині, дослідження порогів класифікації може значно поліпшити результати, підвищуючи впевненість у діагнозах і, як наслідок, покращуючи догляд за пацієнтами.

Зовнішнє тестування. Використання незалежних тестових наборів, які не були залучені під час навчання моделі, дозволяє отримати об'єктивну оцінку її загальної здатності до узагальнення на нові дані. Цей підхід допомагає виявити, наскільки добре модель може застосовувати свої знання до зразків, які вона раніше не бачила, що є ключовим аспектом її ефективності. Тестування на незалежних наборах даних забезпечує більш реалістичне уявлення про те, як модель може впоратися з різними варіантами даних, які можуть виникнути в клінічній практиці. Це особливо важливо в медичній сфері, де точність діагностики може безпосередньо вплинути на здоров'я пацієнтів. Зовнішнє тестування також допомагає виявити потенційні проблеми, такі як перенавчання моделі, коли вона демонструє високу точність на навчальних даних, але не може успішно класифікувати нові зразки. Це може бути результатом недостатнього представлення певних класів або варіацій у навчальних даних. Оцінка моделі на зовнішніх тестових наборах дозволяє виявити такі недоліки і коригувати стратегії навчання або архітектуру моделі для покращення її загальної продуктивності. Загалом, зовнішнє тестування є необхідним для підтвердження надійності та точності моделей у медичній діагностиці, забезпечуючи впевненість у тому, що результати, отримані за допомогою нейронних мереж, будуть корисними і в реальних умовах. Це дозволяє лікарям та медичним працівникам покладатися на такі технології у прийнятті важливих клінічних рішень.

Клінічна валідація передбачає тестування моделі в реальних клінічних умовах, що дає змогу оцінити її практичну корисність та точність під час роботи з пацієнтами. Клінічна валідація дозволяє перевірити, наскільки добре модель справляється з реальними даними, що можуть варіюватися через різні фактори, такі як демографічні особливості пацієнтів, варіації в обладнанні та умови проведення обстежень.

Оцінка моделі в клінічних умовах є важливою, оскільки лабораторні або контрольовані умови тестування не завжди можуть точно відображати реальні ситуації, з якими стикаються медичні працівники. У реальних клінічних умовах

модель повинна мати змогу адаптуватися до різноманітних і, іноді, непередбачуваних обставин. Клінічна валідація дозволяє виявити такі фактори, як вплив навколишнього середовища, людський фактор у роботі медичного персоналу та можливі зміни в характеристиках пацієнтів, які можуть вплинути на результати діагностики. Крім того, клінічна валідація забезпечує зворотний зв'язок, що дозволяє постійно вдосконалювати модель. Вона може вказувати на області, де модель може бути неефективною або де потрібно внести зміни, щоб покращити точність. Це також важливо для перевірки, чи результати, отримані за допомогою моделі, є зрозумілими та прийнятними для медичних працівників, які повинні приймати рішення на основі цих результатів.

Таким чином, комбінація різних методів оцінки точності, ретельний підбір даних, активна верифікація результатів експертами та клінічна перевірка є ключовими аспектами для досягнення високої точності нейронних мереж у медичній діагностиці.

РОЗДІЛ 2.

МЕТОДИ ДІАГНОСТИКИ ЛЕГЕНЕВИХ ЗАХВОРЮВАНЬ НЕЙРОМЕРЕЖЕВИМИ ТЕХНОЛОГІЯМИ ТА ПІДХОДИ ДО СТВОРЕННЯ ІНСТРУМЕНТІВ ЇЇ АВТОМАТИЗАЦІЇ

2.1. Постановка задачі, призначення та вимоги

Протягом останніх років застосування нейромережових технологій у медичній діагностиці стало важливим напрямком для обміну передовим досвідом серед фахівців у галузі охорони здоров'я, штучного інтелекту та машинного навчання. Медичні заклади, лікарі та науковці виявляють потребу у надійних інструментах для швидкого і точного аналізу медичних зображень, які б полегшували процес діагностики та допомагали у прийнятті рішень.

З огляду на потреби сучасної медичної спільноти та досягнення в галузі нейромережових технологій, можна визначити базові вимоги до програмного забезпечення NeuroMedDiag для підтримки процесу діагностики, яке забезпечить наступне:

- надійне завантаження та обробку медичних зображень (МРТ, КТ, рентген) з використанням нейронних мереж для виявлення патологій і аномалій;
- інтуїтивний інтерфейс, що забезпечує простий доступ до інструментів для завантаження та аналізу даних;
- автоматичне виявлення та класифікація патологій з подальшим наданням діагностичних рекомендацій;
- інтеграцію з базами даних медичних установ та іншими системами обміну медичною інформацією для оперативного отримання додаткових даних про пацієнта;
- адміністративний модуль для налаштування параметрів і моніторингу роботи системи, а також для додавання нових функцій, таких як підтримка різних видів зображень або специфічних вимог лікарень;

- можливість автоматичного генерування звітів, що містять результати аналізу та рекомендації, для подальшого використання лікарями та іншими фахівцями;
- система безпеки, що забезпечує конфіденційність медичних даних пацієнтів та відповідає міжнародним стандартам захисту даних.

Для медичних працівників система забезпечує зручний доступ до діагностичних даних, можливість отримання рекомендацій на основі аналізу зображень, а також підтримку обміну інформацією між різними підрозділами чи медичними закладами. Система розроблена для того, щоб полегшити процес виявлення патологій і аномалій, використовуючи алгоритми нейронних мереж, які здатні аналізувати зображення, такі як МРТ або КТ, та генерувати попередні діагностичні висновки. Це дозволяє лікарям зосередитися на прийнятті обґрунтованих рішень щодо лікування пацієнтів, підвищуючи точність діагностики та оптимізуючи час, витрачений на обробку даних.

2.2. Методологія дослідження

Методологія дослідження у темі "Нейромережеві технології в медичній діагностиці легеневих захворювань" охоплює кілька ключових етапів для вивчення ефективності та практичного застосування нейромереж у медичній діагностиці. Дослідження проводиться шляхом поєднання теоретичних та емпіричних методів, спрямованих на розробку і перевірку моделей для аналізу медичних зображень.

Методи проєктування інформаційних систем є комплексом підходів та принципів, що допомагають системно планувати, організовувати та реалізовувати інформаційні системи відповідно до вимог користувачів і потреб бізнесу.

Одним з найбільш відомих є каскадний метод (Waterfall), де розробка системи проходить через послідовні етапи — від аналізу вимог до обслуговування. Кожен етап виконується один за одним, і перехід до наступного

можливий лише після завершення попереднього. Такий підхід є структурованим і найбільш підходить для проєктів з чітко визначеними вимогами.

Іншим методом є спіральна модель, яка поєднує елементи каскадного підходу та прототипування. Спіральна модель організована у вигляді повторюваних циклів, на кожному з яких відбуваються планування, аналіз ризиків, розробка і тестування. Цей метод дозволяє адаптуватися до змінних вимог і допомагає виявляти потенційні ризики на кожному етапі розробки, що є особливо важливим для складних проєктів.

Метод прототипування орієнтований на швидке створення та тестування попередніх версій системи. Він дозволяє перевіряти функціональність, отримувати відгуки від користувачів та уточнювати вимоги на основі тестових версій, що допомагає краще зрозуміти потреби користувачів і уникнути суттєвих переробок на пізніх етапах розробки. Такий підхід особливо ефективний для проєктів, де вимоги на початку не є чітко визначеними.

Об'єктно-орієнтоване проєктування (OOD) зосереджується на поданні системи у вигляді об'єктів, що взаємодіють між собою. Цей метод підтримує інкапсуляцію, наслідування та поліморфізм, дозволяючи створювати системи, які легше модифікувати та масштабувати.

Аналіз літератури та огляд існуючих технологій є важливим першим етапом дослідження в рамках теми "Нейромеревеві технології в медичній діагностиці легеневих захворювань". Цей етап передбачає глибоке вивчення теоретичних основ нейромеревевих технологій, їхніх сучасних підходів та прикладів застосування для обробки та аналізу медичних зображень, а також огляд методів машинного навчання, які використовуються для діагностики різноманітних захворювань.

Нейромеревеві технології, зокрема глибокі нейронні мережі (Deep Neural Networks, DNN), показали значні досягнення у сфері медичної діагностики. Найбільш поширеним підходом є використання згорткових нейронних мереж (Convolutional Neural Networks, CNN), які відзначаються високою ефективністю в аналізі медичних зображень. CNN використовуються для задач класифікації,

сегментації та виявлення об'єктів у зображеннях, що дозволяє швидко та з високою точністю ідентифікувати патології, наприклад, пухлини, тромби чи інші аномалії.

Дослідження літератури також включає огляд методів для підвищення ефективності роботи моделей нейронних мереж, зокрема, технік аугментації даних, нормалізації та використання різних архітектур, таких як U-Net та ResNet, які показали високу продуктивність при аналізі медичних зображень. U-Net, наприклад, є архітектурою, розробленою для завдань сегментації, що дозволяє виділяти конкретні ділянки на медичних зображеннях з високою точністю, що особливо корисно в онкології та радіології.

Огляд літератури також виявляє сильні та слабкі сторони сучасних методів. Сильні сторони включають високу точність і швидкість діагностики, можливість автоматизованого аналізу великих обсягів даних і виявлення патологій на ранніх стадіях. Однак, слабкі сторони включають потребу в великих обсягах даних для навчання, високу обчислювальну складність, що вимагає потужних ресурсів, а також потенційну вразливість моделей до помилок або упередженості в даних.

Загалом, перший етап дослідження дозволяє отримати цілісне розуміння існуючих технологій, а також визначити перспективи для подальшого вдосконалення методів нейромережевого аналізу медичних зображень. Це формує основу для розробки власної нейромережевої моделі, що відповідатиме вимогам точності та ефективності в діагностичних задачах.

На етапі визначення алгоритмів і архітектур нейронних мереж для медичної діагностики важливо вибрати такі архітектури, які максимально відповідають специфіці аналізу медичних зображень. Згорткові нейронні мережі (CNN) є найпоширенішим вибором для цієї мети, оскільки вони добре підходять для виявлення та класифікації патологій завдяки здатності автоматично розпізнавати важливі візуальні особливості. CNN-архітектури, такі як ResNet (Рис.1.4), VGG (Рис.1.3) та U-Net, зарекомендували себе як ефективні для обробки зображень високої роздільної здатності та здатні виділяти деталі навіть

у складних умовах. Наприклад, архітектура U-Net особливо корисна в задачах сегментації зображень, таких як виділення меж пухлин, що є важливим для аналізу зображень у радіології. ResNet, завдяки особливим «вузлам зворотного зв'язку», зберігає інформацію на всіх етапах навчання, що підвищує її ефективність у класифікації.

Окрім вибору архітектури, критичним є налаштування параметрів обраних моделей, таких як глибина мережі, розмір фільтрів, швидкість навчання та розмір партій (batch size). Від налаштування параметрів залежить здатність мережі до узагальнення та точність її роботи, що забезпечує баланс між точністю та швидкістю.

Оцінка продуктивності моделі проводиться на основі метрик, які дозволяють об'єктивно оцінити результат. Основними метриками є точність, чутливість, специфічність та значення F1. Точність показує загальну ефективність моделі, чутливість вимірює здатність правильно виявляти позитивні випадки, а специфічність — правильно визначати здорові зразки. Значення F1, яке комбінує точність і чутливість, допомагає зберігати баланс між ними і є важливим у випадках, коли є суттєвий дисбаланс між класами.

Збір та підготовка даних є важливим етапом у процесі навчання нейромереж для медичної діагностики. Є декілька основних етапів.

1. *Збір даних.* Для навчання нейромереж використовуються різноманітні медичні зображення, такі як:

- магнітно-резонансні томографії (МРТ), МРТ-сканування надає високоякісні зображення тканин та органів, дозволяючи виявляти зміни на ранніх стадіях захворювань;
- комп'ютерна томографія (КТ), КТ-знімки використовуються для діагностики патологій, які вимагають візуалізації структури органів, кісток та тканин;
- рентген-знімки використовуються для виявлення різноманітних захворювань, таких як переломи, пневмонія, онкологічні захворювання.

Дані можуть бути отримані з відкритих джерел (наприклад, з публічних медичних баз даних, таких як NIH Chest X-rays, LUNA16) або з ліцензованих джерел.

2. *Попередня обробка даних.* Попередня обробка є критично важливою для підготовки даних до навчання нейромережі. Це включає кілька основних етапів:

- нормалізація, приведення пікселів зображення до стандартного діапазону значень (наприклад, від 0 до 1 або від -1 до 1), щоб полегшити навчання нейромережі;
- сегментація, процес розділення зображення на окремі компоненти (наприклад, виділення органу або області інтересу на зображенні, що дозволяє нейромережі фокусуватися на певних ділянках, що важливі для діагностики);
- аугментація, застосування методів штучного збільшення кількості даних, таких як обертання, масштабування, змінення яскравості та контрасту зображень (дозволяє збільшити різноманітність даних і покращити здатність моделі до генералізації на нові зображення).

3. *Розділення на навчальну та тестову вибірки.* Зібрані та попередньо оброблені зображення діляться на навчальні та тестові набори. Зазвичай навчальний набір становить 70-80% від загальної кількості даних, а решта йде на тестування моделі.

4. *Анотовані дані.* Для медичних зображень часто необхідно, щоб дані були анотовані лікарями або спеціалістами, що визначають на зображеннях патології або області інтересу. Це дозволяє нейромережам вчитися на основі правильних міток і здійснювати точнішу діагностику.

Методи статистичної класифікації та кластеризації є важливими інструментами для аналізу великих наборів даних, включаючи медичні дані, оскільки вони допомагають автоматично визначати класи або групи на основі спільних характеристик. Наведемо методи класифікації.

Лгістична регресія використовується для бінарної або багатокласової класифікації. Вона оцінює ймовірність належності до певного класу на основі вхідних змінних і широко застосовується в медичних дослідженнях для прогнозування наявності або відсутності певного стану.

Дискримінантний аналіз, поділяється на лінійний дискримінантний аналіз (LDA) і квадратичний дискримінантний аналіз (QDA), які класифікують об'єкти на основі їхніх характеристик. Цей метод ефективний, коли припускається, що дані нормально розподілені. LDA часто використовується для розділення груп пацієнтів за медичними показниками.

Метод опорних векторів (SVM), це метод класифікації, що шукає гіперплощину, яка оптимально розділяє класи. SVM добре підходить для складних і високовимірних даних і використовується для діагностики на основі зображень, оскільки ефективно розпізнає різні стани або патології.

Байєсівські методи, зокрема, наївний байєсівський класифікатор використовує теорему Байєса для обчислення ймовірностей приналежності до класів на основі відомих характеристик. Цей метод простий, швидкий, і особливо корисний при роботі з текстами або базами даних з обмеженим числом змінних.

Дерева рішень (як, наприклад, алгоритм CART) використовують ієрархічне розбиття даних, де кожна гілка відповідає певній характеристиці. Ансамблеві методи, такі як випадковий ліс (Random Forest) і градієнтний бустинг, поєднують кілька дерев для підвищення точності, що особливо корисно при обробці комплексних медичних даних.

Наведемо методи кластеризації.

Кластеризація K-середніх (K-means). Це один з найпростіших і популярних методів неконтрольованого навчання, який групує об'єкти в певну кількість кластерів, де кожен кластер має свій центр. K-means ефективний для сегментації пацієнтів за певними медичними показниками, дозволяючи виявити підгрупи з подібними характеристиками.

Ієрархічна кластеризація. Виконує кластеризацію шляхом побудови дерева кластерів (дендрограми), що показує ієрархічну структуру об'єктів. Цей

метод може бути корисним для аналізу взаємозв'язків між групами пацієнтів або для виявлення підгруп захворювань.

Метод DBSCAN (Density-Based Spatial Clustering of Applications with Noise). Це метод, що виявляє кластери на основі щільності. DBSCAN добре підходить для кластеризації даних із нерівномірною щільністю і дозволяє виявляти аномалії чи рідкісні випадки, що може бути цінним у медичних застосуваннях.

Метод головних компонент (PCA) для зменшення розмірності. Хоча PCA не є методом кластеризації в класичному розумінні, він використовується для спрощення даних шляхом зменшення розмірності, що допомагає покращити кластеризацію та візуалізацію складних даних.

Кластеризація на основі Gaussian Mixture Models (GMM). GMM моделює дані як суму нормальних розподілів і підходить для кластеризації, коли групи мають складні форми. Цей метод добре підходить для даних з неоднорідною природою, де кластери можуть перетинатися.

Методи статистичної класифікації та кластеризації можуть використовуватися як окремо, так і в комбінації. У медичній діагностиці ці методи дозволяють автоматизувати обробку великих обсягів даних, підтримуючи медичних працівників у діагностиці, прогнозуванні та індивідуалізації підходів до лікування.

Розробка та тренування нейромережевої моделі для медичної діагностики є важливим етапом, який включає кілька ключових процесів для досягнення високої точності та ефективності. Наведемо основні кроки.

1. *Вибір моделі.* Вибір правильної нейромережевої моделі залежить від типу завдання, що розв'язується, і від характеристик наданих медичних зображень. Для задач медичної діагностики часто використовуються наступні моделі: конволюційні нейромережі (CNN), глибинні нейромережі (DNN), рекурентні нейромережі (RNN).

2. *Тренування моделі.* Для навчання моделі використовуються підготовлені набори даних. Тренування включає кілька етапів. Перед навчання (pre-training)

може бути використано попередньо навчене ядро на великій базі зображень, що прискорює навчання (наприклад, transfer learning). Це дозволяє використовувати вже здобуті знання, коли навчальна вибірка є обмеженою. Оптимізація ваг: В процесі навчання ваги нейронів моделі коригуються для мінімізації функції втрат. Для цього використовуються алгоритми оптимізації, такі як стохастичний градієнтний спуск (SGD) та адаптивні методи (наприклад, Adam).

3. *Методи оптимізації для підвищення ефективності.* Для досягнення хороших результатів необхідно використовувати кілька стратегій оптимізації. Регуляризація допомагає уникнути перенавчання (overfitting). Основні методи регуляризації включають L1/L2 регуляризацію, яка додає штраф до функції втрат за великі значення ваг, та Dropout, який випадково "вимикає" частину нейронів під час навчання. Крос-валідація є методом оцінки якості моделі, коли дані діляться на кілька частин (folds), і модель тренується та тестується на різних підмножинах даних. Це дозволяє більш точно оцінити здатність моделі узагальнювати на нових даних. Найпоширеніші методи крос-валідації — це K-fold крос-валідація, де дані діляться на K частин, та Stratified k-fold, що забезпечує рівний розподіл класів у кожній частині.

4. *Оцінка якості моделі.* Після навчання важливо оцінити продуктивність моделі, використовуючи метрики, такі як точність (accuracy), прецизійність (precision), повнота (recall) та F1-міра. Це дає змогу зрозуміти, наскільки добре модель здатна розпізнавати патології в медичних зображеннях і як вона працює в умовах різних клінічних випадків.

В результаті тренування модель повинна бути здатна до точного і швидкого розпізнавання патологій, враховуючи специфіку медичних зображень та різноманітність клінічних випадків.

Методи машинного навчання є ключовими інструментами для аналізу та обробки великих обсягів даних, зокрема для діагностики в медичних застосунках. У цьому контексті основні методи можна поділити на кілька категорій.

Контрольоване навчання (Supervised Learning). Це один із найпоширеніших методів, коли моделі навчаються на даних із відповідними мітками. У медичній діагностиці контрольоване навчання допомагає розпізнавати патології на зображеннях, класифікуючи їх на основі вказаних класів (наприклад, "нормальні" або "патологічні" знімки). Серед поширених алгоритмів:

- лінійна регресія та логістична регресія — прості, але ефективні моделі для передбачення результатів, часто використовувані для класифікації або оцінки ймовірностей;
- метод опорних векторів (SVM) – ефективний для виявлення складних закономірностей у високовимірних даних;
- дерева рішень та ліси рішень (Random Forest) – добре підходять для обробки неоднорідних медичних даних та дозволяють інтерпретувати результати.

Неконтрольоване навчання (Unsupervised Learning). Застосовується для виявлення структур і кластерів у даних без наявних міток. Це корисно для аналізу великих обсягів даних і виявлення закономірностей, які могли б залишитися непоміченими:

- кластеризація (наприклад, алгоритми К-середніх, ієрархічна кластеризація) дозволяє групувати пацієнтів за подібними характеристиками, що може виявити нові підгрупи захворювань;
- метод головних компонент (PCA) — зменшує розмірність даних, що є корисним для попередньої обробки медичних зображень, зберігаючи найбільш інформативні ознаки.

Навчання з підкріпленням (Reinforcement Learning). У цьому підході модель навчається на основі системи винагород, удосконалюючи свою стратегію на основі отриманих результатів. У медичній сфері навчання з підкріпленням корисне, наприклад, для планування процедур лікування або оптимізації дозування ліків.

Глибоке навчання (Deep Learning). Використовується для обробки великих і складних даних, особливо для аналізу медичних зображень, де традиційні методи менш ефективні. Найбільш поширені архітектури:

- згорткові нейронні мережі (CNN) використовуються для виявлення особливостей у медичних зображеннях (КТ, МРТ, рентген);
- рекурентні нейронні мережі (RNN) та довготривала короткочасна пам'ять (LSTM) – застосовуються для аналізу часових рядів, наприклад, для прогнозування змін у медичних показниках;
- генеративні методи включають генеративні змагальні мережі (GAN), що дозволяють створювати нові зображення, подібні до існуючих. GAN є цінним інструментом для збільшення даних і може допомогти покращити точність моделі, створюючи додаткові дані для тренування.

Кожен з методів підходить для різних задач, і їхнє комбінування може забезпечити ефективне рішення для медичної діагностики, підвищуючи точність та надійність результатів.

Методи машинного зору є критично важливими для аналізу медичних зображень, оскільки дозволяють автоматично розпізнавати, класифікувати та сегментувати об'єкти на знімках, що сприяє швидкій та точній діагностиці.

1. *Попередня обробка зображень*. Це початковий етап, що включає методи очищення та поліпшення якості зображень перед їхнім аналізом:

- фільтрація шумів (наприклад, гаусівська та медіанна фільтрація) зменшує кількість небажаних артефактів;
- нормалізація та масштабування дозволяють зробити дані більш однорідними;
- збільшення роздільної здатності (super-resolution) за допомогою згорткових нейронних мереж (SRCNN) покращує деталі на зображеннях, що важливо для виявлення дрібних патологій.

2. *Сегментація зображень*. Це виділення окремих областей на зображенні, таких як органи чи патологічні ділянки:

- U-Net — архітектура нейронної мережі, що спеціалізується на сегментації та показала високу ефективність у медичних застосуваннях;
- Fully Convolutional Networks (FCN) – замінює щільні шари на згорткові, що дозволяє працювати із зображеннями різних розмірів;
- Mask R-CNN – метод, що поєднує сегментацію і виявлення об'єктів, зокрема добре справляється із задачами сегментації в медичних зображеннях.

3. *Класифікація зображень.* Це метод визначення типу об'єкта або патології на зображенні:

- згорткові нейронні мережі (CNN) для класифікації, такі як ResNet і VGG, використовуються для автоматичного розпізнавання захворювань;
- Transfer learning – підхід, який використовує попередньо навчені моделі для адаптації до нових задач, наприклад, при обмеженій кількості медичних зображень для тренування.

4. *Виявлення об'єктів.* Виявлення певних ознак або патологій на зображенні з визначенням їхнього розташування:

- YOLO (You Only Look Once) – швидкий метод для виявлення об'єктів на зображенні в реальному часі;
- Faster R-CNN – модель, що виконує одночасно класифікацію та визначення місця об'єкта, є ефективною для виявлення патологій на медичних знімках.

5. *Аналіз текстур.* Це оцінка закономірностей на зображенні для виявлення особливостей поверхні тканин або органів:

- метод градієнтів (HOG) — використовується для аналізу та розпізнавання текстур на зображеннях;
- GLCM (gray-level co-occurrence matrix) та інші статистичні методи допомагають виділити особливості текстури для подальшого розпізнавання аномалій.

6. *Генерація зображень.* Це створення нових зображень для тренування моделей або збільшення набору даних:

- генеративні змагальні мережі (GAN) дозволяють створювати нові зображення, що мають реалістичні характеристики, і корисні для тренування моделей при нестачі медичних даних;
- CycleGAN – підхід, що допомагає перетворювати зображення між різними доменами (наприклад, перетворення МРТ в КТ).

Ці методи машинного зору дозволяють ефективно обробляти медичні зображення, автоматизуючи процес діагностики, підвищуючи точність та прискорюючи виявлення патологій. Кожен метод може використовуватися як окремо, так і в комбінації, залежно від специфічних вимог медичного застосування.

Тестування та оцінка моделі є важливим етапом у процесі розробки нейромережевих технологій для медичної діагностики, оскільки дозволяє перевірити, наскільки добре модель справляється з реальними даними і порівняти її продуктивність з традиційними методами. Після завершення тренування модель проходить тестування на окремому наборі даних, який не використовувався під час навчання. Це дозволяє оцінити її здатність до узагальнення на нові, невідомі дані та перевірити її продуктивність в реальних умовах. Зазвичай для тестування виділяють 20-30% даних, які не входять до навчального набору. Для оцінки результатів роботи моделі використовуються кілька метрик, які допомагають зрозуміти, наскільки добре модель виконує свою задачу. Основні метрики включають точність, чутливість та специфічність. Точність — це відсоток правильних прогнозів серед усіх прогнозів. Це базова метрика, яка показує загальний рівень успішності моделі. Чутливість, або показник справжніх позитивів, показує здатність моделі правильно ідентифікувати позитивні випадки (наприклад, наявність хвороби). Чутливість важлива в медичній діагностиці, оскільки дозволяє оцінити, як добре модель виявляє наявність патології. Специфічність — це здатність моделі правильно ідентифікувати негативні випадки (наприклад, здорових пацієнтів). Специфічність важлива, щоб уникнути хибнопозитивних результатів, коли здорова людина помилково класифікується як хвора. Оцінка моделі також

включає порівняння результатів, отриманих нейромережею, з результатами традиційних методів діагностики, таких як візуальний огляд зображень лікарем або застосування класичних алгоритмів. Це дозволяє зрозуміти, чи нейромережа може перевершити або доповнити існуючі методи, забезпечуючи швидкість, точність та ефективність. Окрім основних метрик, може використовуватися матриця плутанини для більш детального аналізу, яка показує кількість істинно позитивних, хибно позитивних, істинно негативних та хибно негативних результатів. Також можна використовувати F1-міру, яка є середнім гармонійним між чутливістю та специфічністю і дає збалансовану оцінку моделі, особливо в умовах незбалансованих класів. Після тестування моделі важливо провести аналіз отриманих результатів, зокрема звернути увагу на помилки моделі (наприклад, хибнопозитивні чи хибнонегативні результати), щоб зрозуміти, в яких випадках вона може давати помилки і як покращити її точність. Тестування та оцінка продуктивності є необхідними кроками для забезпечення того, щоб модель була надійною та точною при використанні в реальних медичних умовах, де кожне рішення може мати серйозні наслідки.

Аналіз результатів та формування рекомендацій є важливим завершальним етапом у впровадженні нейромережових технологій у медичну діагностику. На цьому етапі проводиться оцінка продуктивності нейромережі в реальних умовах клінічної практики. Це дозволяє з'ясувати, наскільки ефективно модель працює з медичними зображеннями та даними пацієнтів, і виявити її сильні та слабкі сторони. Оцінка продуктивності включає аналіз точності, чутливості, специфічності та інших метрик, які дозволяють зрозуміти, наскільки правильно нейромережа визначає патології порівняно з традиційними методами діагностики. Також вивчаються випадки, де модель помиляється, щоб зрозуміти, в яких ситуаціях це відбувається, і як можна мінімізувати ці помилки.

Після цього формуються рекомендації для покращення діагностики з використанням нейромережових технологій. Це може включати вдосконалення моделі за допомогою додаткових даних чи нових алгоритмів для підвищення її точності, а також оптимізацію процесу інтеграції з іншими медичними

системами для більш ефективного обміну інформацією. Крім того, важливим є поліпшення взаємодії між медичним персоналом і технологією, що може включати вдосконалення інтерфейсу користувача та забезпечення належного навчання лікарів. Залежно від результатів аналізу можуть бути запропоновані нові напрямки використання нейромереж, наприклад, для автоматизації інших аспектів діагностики або підтримки клінічних рішень на етапах лікування.

2.3 Обґрунтування вибору інструментальних засобів

Для створення сайту, на якому будуть інтегровані нейромережеві моделі для медичної діагностики, HTML та CSS є чудовим вибором для розробки фронтенду, оскільки вони дозволяють створити структуру та стиль сайту. HTML забезпечує базову розмітку сторінок, а CSS дозволяє оформити їх для зручного та привабливого інтерфейсу користувача. Також було використано JavaScript, дозволяє створювати інтерактивні та динамічні веб додатки, що дозволяє користувачам взаємодіяти з графіками, оновлюючи або змінюючи їх у реальному часі. Це дуже важливо для проєктів, де необхідно відображати дані, що змінюються, або надавати можливість взаємодії з інформацією. Бібліотека Chart.js є легкою та потужною, з можливістю створення різноманітних графіків, таких як лінійні, стовпчикові, кругові та інші, з мінімумом коду. Вона дозволяє швидко інтегрувати графічні елементи в веб додатки, що значно спрощує розробку. Однією з ключових переваг є кросплатформенність: оскільки JavaScript підтримується всіма сучасними веб-браузерами, додатки на його основі можуть працювати на різних пристроях і операційних системах без додаткових налаштувань. Chart.js також підтримує анімацію, що дозволяє створювати плавні анімаційні переходи між різними станами графіків, роблячи візуалізацію даних більш захоплюючою для користувачів. Ця бібліотека також відома своєю простотою у використанні та швидким часом розробки, що дозволяє без проблем створювати графіки, налаштовувати їх кольори, підписи і

легенди відповідно до потреб проєкту. Іншим важливим аспектом є документація та активна спільнота Chart.js, що дозволяє швидко знаходити рішення на випадок проблем чи запитів під час реалізації. Гнучкість бібліотеки дає змогу працювати з великими наборами даних, що забезпечує візуалізацію навіть складних наборів інформації, при цьому дозволяючи налаштовувати вигляд і функціональність графіків під специфічні вимоги користувача. Усе це робить використання JavaScript та Chart.js ідеальним для веб-додатків, де необхідно відображати актуальні або змінювані дані.

Що стосується вибору нейромережевої моделі, то для медичних зображень, таких як МРТ, КТ або рентген-знімки, найкращим вибором буде використання Convolutional Neural Networks (CNN) — згорткових нейромереж. Ці моделі здобули популярність завдяки своїй високій ефективності в задачах обробки зображень. Вони є ідеальними для розпізнавання та класифікації медичних зображень, таких як виявлення патологій на МРТ або рентгенівських знімках.

Зокрема, можна використовувати наступні архітектури.

ResNet (Residual Networks) (Рис.1.4). Модель, яка добре працює для класифікації медичних зображень завдяки використанню решіткових зв'язків (skip connections), що дозволяють побудувати дуже глибокі мережі без проблеми згасання градієнтів. Це дозволяє моделі ефективно працювати з великими обсягами медичних зображень.

U-Net. Це специфічна архітектура для сегментації зображень, яка дуже популярна в медичній діагностиці, оскільки дозволяє точніше виявляти межі органів або патологічних утворень на зображеннях. Вона складається з контракційної і розширювальної частин і може використовуватися для виявлення різних аномалій на знімках.

DenseNet. Мережа, де кожен шар взаємодіє з усіма наступними, що покращує потік інформації та допомагає моделі навчатися на більш глибоких даних. Ця модель також демонструє хороші результати при класифікації зображень в медицині.

VGGNet (Рис.1.3). Це одна з класичних архітектур для обробки зображень. Хоча вона застаріла порівняно з новішими моделями, її можна використовувати для простих задач діагностики та класифікації зображень.

Вибір архітектури залежить від специфіки задачі. Якщо необхідно проводити сегментацію, то найкращим вибором буде U-Net. Для більш загальних задач класифікації зображень, таких як виявлення певних захворювань, ResNet або DenseNet будуть відмінними варіантами. Ці моделі можуть бути реалізовані за допомогою таких фреймворків, як TensorFlow або PyTorch, які також дозволяють інтегрувати неймережеві моделі в веб-сайт, якщо це потрібно для онлайн-аналізу медичних зображень через браузер.

Для роботи було обрано навчання на рентгенівських знімках (PadChest, COVID-19 Radiography Database, MIMIC-CXR Dataset) та шлях класифікації патологій. Цей шлях спрямований на класифікацію знімків за наявністю різних патологій (наприклад, пневмонії, COVID-19, фіброзу). Моделі тренуються на знімках із відповідними мітками для кожного захворювання.

Перш за все, рентгенографічні датасети, такі як PadChest, COVID-19 Radiography Database та MIMIC-CXR, містять велику кількість мічених зображень з детальними анотаціями. Це створює можливість побудови надійної класифікаційної моделі, здатної автоматично ідентифікувати різні патології, зокрема пневмонію, пневмоторакс, фіброз, COVID-19 та рак легень.

Рентгенографія є однією з найпоширеніших і доступних методик медичної візуалізації, що робить її ключовим інструментом для первинної діагностики. Автоматизована класифікація патологій на основі рентгенівських знімків значно пришвидшить цей процес і зменшить навантаження на лікарів. Крім того, сучасні глибокі нейронні мережі, такі як згорткові (CNN), демонструють високу ефективність у класифікаційних завданнях на медичних зображеннях, досягаючи високої точності й надійності результатів.

Рентгенографія також є особливо актуальною в умовах пандемій. Наприклад, COVID-19 Radiography Database надає змогу навчити модель розпізнавати патології, пов'язані з COVID-19, що важливо для раннього

виявлення інфекційних захворювань і моніторингу їх поширення. Таким чином, використання рентгенівських знімків і класифікації патологій є оптимальним підходом для вирішення завдань, пов'язаних з діагностикою легень. [16]

Вибір наглядного та трансферного навчання для задачі класифікації патологій на рентгенівських знімках є доцільним і обґрунтованим, враховуючи специфіку медичних даних та поставлені діагностичні цілі.

Наглядове навчання є оптимальним підходом, оскільки медичні датасети, такі як PadChest, COVID-19 Radiography Database та MIMIC-CXR, містять зображення, кожне з яких має певну мітку патології, встановлену експертами. Завдяки наявності мічених даних можна створити модель, яка навчається розрізняти конкретні патології (наприклад, пневмонію, COVID-19, фіброз тощо) на основі раніше визначених медиками прикладів. Наглядове навчання забезпечує високу точність, оскільки модель навчається ідентифікувати відомі класи патологій, що важливо для коректного діагностування захворювань.

Використання трансферного навчання також є надзвичайно корисним у цій задачі. Це зумовлено тим, що медичні датасети можуть мати обмежену кількість прикладів або бути збалансованими нерівномірно за класами, через що навчання з нуля потребуватиме значних обчислювальних ресурсів і часу. Трансферне навчання дозволяє застосувати знання, набуті моделлю на великому загальному датасеті зображень, до специфічного завдання класифікації медичних знімків. Завдяки цьому методологія не тільки підвищує ефективність навчання, але й допомагає уникнути перенавчання, зберігаючи при цьому високу точність та узагальненість результатів для медичних зображень.

Отже, вибір наглядного навчання забезпечує можливість точного розпізнавання патологій завдяки міченим даним, а трансферне навчання зменшує вимоги до обсягу даних, необхідного для тренування, та сприяє стабільності і точності роботи моделі в умовах, коли кількість доступних медичних даних є обмеженою.

2.4 Етапи програмної реалізації

Програмна реалізація системи для класифікації патологій на рентгенівських знімках легень була структурована у кілька послідовних етапів, кожен з яких забезпечував ключові функціональні можливості системи.

На етапі збору та підготовки даних були обрані рентгенівські знімки з датасетів PadChest, COVID-19 Radiography Database та MIMIC-CXR, оскільки вони містять різні види патологій, що дозволяє створити універсальну модель для медичної діагностики. Підготовка даних включала кілька важливих кроків, спрямованих на забезпечення високої якості вхідних зображень і адаптації їх до потреб нейронної мережі. Було здійснено нормалізацію яскравості та контрасту знімків для покращення візуальних характеристик та забезпечення однорідності даних. Масштабування до фіксованого розміру сприяло спрощенню обробки зображень і підвищенню швидкості навчання моделі. Крім того, застосовано методи аугментації даних, зокрема обертання зображень, дзеркальне відображення та регулювання яскравості. Це було зроблено для збільшення розмаїття вхідних даних і покращення стійкості моделі до різних варіантів зображень, які можуть зустрічатися в клінічних умовах. Така аугментація дозволила розширити навчальний набір, покращуючи здатність моделі до узагальнення і, відповідно, підвищуючи її точність при класифікації різних патологій на нових знімках.

На етапі розробки моделі нейромережі для класифікації патологій на рентгенівських знімках була обрана згорткова нейронна мережа (CNN), оскільки вона добре підходить для обробки медичних зображень, виділяючи важливі візуальні ознаки патологій. Враховуючи обмеження ресурсів та наявних медичних даних, було вирішено використати метод трансферного навчання. Для цього основою стала попередньо натренована CNN-модель (наприклад, ResNet, VGG або Inception), яка вже володіє широкими знаннями про зображення завдяки тренуванню на великому датасеті (ImageNet). Метод трансферного

навчання дозволив ефективно адаптувати базову модель до специфічної задачі медичної діагностики. Останні шари мережі, відповідальні за класифікацію, були модифіковані та перенавчені на рентгенівських знімках із специфічними мітками захворювань, таких як пневмонія, COVID-19 та туберкульоз. Це дало змогу моделі краще розрізняти різні патології, зберігаючи при цьому навички, набуті під час початкового навчання. Такий підхід сприяв значному скороченню часу на тренування та підвищенню точності моделі в умовах медичної діагностики.

Розроблено інтерфейс користувача на основі HTML, CSS та JavaScript, що дозволяє користувачам завантажувати рентгенівські знімки для аналізу (Рис.2.1). Інтерфейс підтримує функції відображення ймовірності кожного захворювання на основі зображення. Основні розділи веб-додатку включають інформаційний блок (Рис.2.2), секцію для завантаження зображень (Рис.2.3), блок для відображення результатів (Рис.2.4) і рекомендацій (Рис.2.5).

Для відображення результатів класифікації використовувалася бібліотека Chart.js, що забезпечує графічне представлення ймовірностей різних захворювань. Діаграми відображають імовірності наявності патологій, що надає користувачеві зручне візуальне уявлення про результати аналізу.

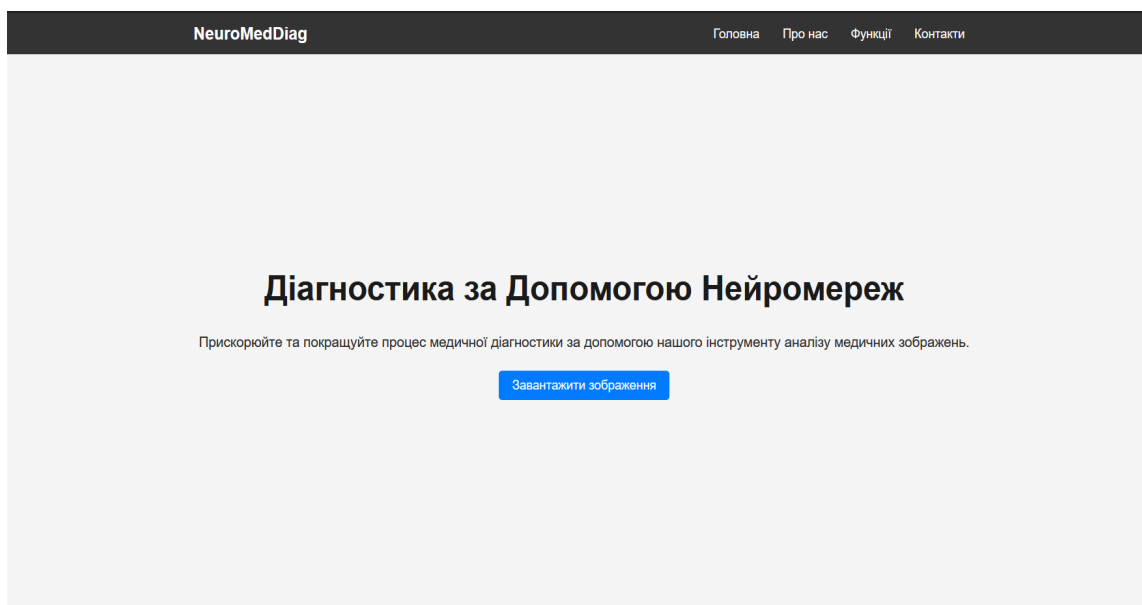


Рисунок 2.1 – Інтерфейс розробки

Про нас

NeuroMedDiag — інноваційне програмне рішення для діагностики легеневих захворювань на підставі аналізу медичних зображень із використанням нейромережових технологій. Ми допомагаємо лікарям швидко й точно визначати патології, покращуючи якість медичної допомоги.

Основні можливості

Аналіз медичних зображень

Автоматичне виявлення патологій на МРТ та КТ знімках.

Класифікація захворювань

Класифікація захворювань на основі оброблених даних.

Діагностичні висновки

Генерація рекомендацій для медичного персоналу.

Рисунок 2.2 – Інформаційний блок інтерфесу користувача

Завантажити зображення

Завантажте свої медичні зображення для аналізу.

Огляд... Файл не вибрано.

Я погоджуюсь на обробку зображення.

Проаналізувати

Рисунок 2.3 – Блок для завантаження медичних зображень

Завантажити зображення

Результати аналізу

Аналіз успішно виконаний!

Найімовірніше захворювання: Бронхіт (87%)

Рисунок 2.4 – Блок результатів обробки

Завантажити зображення

Результати аналізу

Аналіз успішно виконаний!

Найімовірніше захворювання: Бронхіт (87%)

Рекомендації:

Для уточнення діагнозу провести аналіз мокротиння, щоб виявити наявність інфекції, а також спірометрію для оцінки функції легенів. Лікування включає призначення відхаркувальних засобів (амброксол, ацетилцистеїн), бронходилататорів та інгаляторів для полегшення дихання. У разі бактеріальної інфекції застосувати антибіотики. Якщо бронхіт має алергічний або хронічний характер, призначити антигістамінні препарати або інгалятори з кортикостероїдами. Для полегшення симптомів рекомендується рясне пиття, відпочинок і помірне фізичне навантаження після полегшення гострих симптомів. Регулярно проводити моніторинг функції легенів, особливо при хронічному бронхіті.

Рисунок 2.5 – Блок рекомендацій наданих системою

2.5 Аналіз отриманих результатів дослідження, рекомендації використання та впровадження

2.5.3 Аналіз отриманих результатів дослідження

У роботі проведено детальний аналіз теоретичних основ застосування нейромережових технологій у медичній діагностиці, зокрема методів

класифікації патологій на основі медичних зображень. Досліджено фундаментальні принципи роботи згорткових нейронних мереж (CNN), які є основою для класифікації зображень завдяки здатності виділяти та інтерпретувати ключові візуальні ознаки, що характерні для патологій на рентгенівських знімках. Опрацьовано основи трансферного навчання, що дозволяє адаптувати попередньо навчені моделі до специфічних завдань, таких як медична діагностика, з мінімальними витратами на обчислювальні ресурси та час навчання.

Особлива увага була приділена методам обробки зображень, зокрема нормалізації, аугментації та масштабуванню, які дозволяють підвищити надійність і стабільність моделі в умовах клінічної практики. Вивчено різні стратегії аугментації, включаючи обертання, зміни яскравості та контрасту, що сприяють зменшенню залежності моделі від особливостей вихідних даних і підвищують її здатність до узагальнення.

У роботі також було проаналізовано сучасні підходи до створення інтерфейсів для нейромережових систем у медичній діагностиці. Зокрема, опрацьовано найкращі практики щодо інтеграції моделей у веб-інтерфейси, які дозволяють користувачам завантажувати зображення та отримувати результати класифікації в зручному форматі. Ці теоретичні знання були використані для розробки інтерфейсу, який відображає ймовірності діагнозів у вигляді графічних елементів, що підвищує зрозумілість і зручність системи для медичних фахівців.

Окрім цього, було вивчено особливості й виклики використання нейромережових моделей у медичній діагностиці, включаючи питання точності, надійності та інтерпретованості результатів. На основі теоретичних матеріалів були сформульовані рекомендації щодо подальшого розвитку моделі, зокрема щодо регулярного донавчання на актуальних клінічних даних та розширення можливостей аналізу інших типів зображень, що може значно підвищити практичну цінність системи у клінічній практиці.

2.5.2 Рекомендації використання та впровадження

Рекомендації щодо впровадження та використання розробленої системи класифікації патологій на рентгенівських знімках передбачають поступове інтегрування її в медичну практику з урахуванням необхідності подальшого доопрацювання та донавчання. Система може бути корисною для підтримки діагностичних рішень лікарів, особливо на етапах первинного огляду та сортування випадків. Проте, для забезпечення оптимальної точності діагностики, її слід використовувати як допоміжний інструмент, не замінюючи при цьому професійного огляду.

Зважаючи на те, що точність моделі для рідкісних патологій може бути нижчою через обмежену кількість даних у вибірці, рекомендується регулярно оновлювати навчальний набір за рахунок нових клінічних зображень. Це допоможе адаптувати систему до змін у структурі захворювань і підвищить її продуктивність у виявленні малопоширених патологій. Інтеграція системи в існуючі медичні інформаційні системи може полегшити її повсякденне використання в лікарнях і діагностичних центрах, водночас потребує налаштування для забезпечення сумісності із захищеними медичними даними.

Для забезпечення максимальної користі від роботи системи необхідно проводити регулярне перенавчання моделі на нових, актуальних даних, а також постійно відстежувати й аналізувати її точність і надійність. Систему варто розширювати та доповнювати функціональними можливостями для аналізу додаткових типів зображень, таких як КТ або МРТ, щоб охопити ширший спектр діагностичних завдань.

Враховуючи, що робота з медичними зображеннями вимагає обробки конфіденційної інформації, варто дотримуватися вимог безпеки та конфіденційності даних (наприклад, GDPR). Потрібно переконатись, що модель працює з анонімізованими даними або що медична установа використовує надійні механізми захисту інформації.

Для впровадження моделі може знадобитися відповідна апаратна інфраструктура (сервери або потужні комп'ютери) для забезпечення швидкого аналізу зображень. Рекомендовано використання обладнання з підтримкою графічних процесорів (GPU), що прискорить обробку та підвищить зручність використання.

ВИСНОВКИ

У роботі проведено аналіз методів класифікації патологій на рентгенівських знімках із використанням нейромережових технологій. Оцінено ефективність розробленої системи, яка використовує згорткову нейронну мережу (CNN) та метод трансферного навчання для ідентифікації захворювань, зокрема пневмонії, COVID-19 та туберкульозу.

Підтверджено ефективність використання моделей ШІ для медичної діагностики та адаптації до клінічної практики для підвищення точності та ефективності процесу виявлення захворювань.

Вивчено та застосовано передові підходи до підготовки та обробки медичних зображень, що включають нормалізацію, аугментацію та масштабування знімків, з метою підвищення точності моделі та її стійкості до різноманітних варіантів вхідних даних.

Розроблено інтерфейс інструменту підтримки прийняття рішень лікарями для діагностування легеневих захворювань.

Розроблено модель, інтегровану у веб-інтерфейс, яка забезпечує користувачам зручний інструмент для завантаження медичних зображень і отримання результатів діагностики в реальному часі. Запропонований інтерфейс включає візуалізацію ймовірності кожного діагнозу, що спрощує інтерпретацію результатів для лікарів і медичного персоналу. Під час роботи було розглянуто сучасні можливості застосування моделей штучного інтелекту для діагностики легеневих захворювань та адаптації їх до потреб клінічної практики.

Запропоновано рекомендації щодо подальшого вдосконалення системи, зокрема регулярне перенавчання моделі на нових клінічних даних, розширення функціоналу для аналізу інших типів медичних зображень та інтеграцію системи у медичні інформаційні платформи для спрощення діагностичного процесу та підвищення точності виявлення захворювань.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Papers with Code - ChestX-ray14 Dataset. *The latest in Machine Learning | Papers With Code*. URL: <https://paperswithcode.com/dataset/chestx-ray14> (date of access: 02.11.2024).
2. Papers with Code - Covid-19 radiography database Dataset. *The latest in Machine Learning | Papers With Code*. URL: <https://paperswithcode.com/dataset/covid-19-radiography-database> (date of access: 02.11.2024).
3. Papers with Code - LIDC-IDRI Dataset. *The latest in Machine Learning | Papers With Code*. URL: <https://paperswithcode.com/dataset/lidc-idri> (date of access: 02.11.2024).
4. Papers with Code - MIMIC-CXR Dataset. *The latest in Machine Learning | Papers With Code*. URL: <https://paperswithcode.com/dataset/mimic-cxr> (date of access: 02.11.2024).
5. Papers with Code - PadChest Dataset. *The latest in Machine Learning | Papers With Code*. URL: <https://paperswithcode.com/dataset/padchest> (date of access: 02.11.2024).
6. Papers with Code - ChestX-ray8 Dataset. *The latest in Machine Learning | Papers With Code*. URL: <https://paperswithcode.com/dataset/chestx-ray8> (date of access: 02.11.2024).
7. Papers with Code - ChestX-Det Dataset. *The latest in Machine Learning | Papers With Code*. URL: <https://paperswithcode.com/dataset/chestx-det> (date of access: 02.11.2024).
8. Transformers in medical imaging: A survey / F. Shamshad et al. *Medical Image Analysis*. 2023. P. 102802. URL: <https://doi.org/10.1016/j.media.2023.102802> (date of access: 02.11.2024).
9. Papers with Code - Advances in Medical Image Analysis with Vision Transformers: A Comprehensive Review. *The latest in Machine Learning | Papers With Code*. URL: <https://paperswithcode.com/paper/advances-in-medical-image-analysis-with> (date of access: 02.11.2024).

10. Юзва А. С., Лаптев О. А. Штучний інтелект: можливості та виклики для суспільства. *Тези доповідей XIII Міжнар. наук.-практ. конф.* : Міжнар. наук. конф., м. Луцьк-Світязь, 31 трав. – 2 черв. 2024 р. 2024. С. 133–135.
11. Юзва А., Омельчук А., Булатецька Л. Проблеми комп'ютерних наук, програмного моделювання та безпеки цифрових систем. *Тези доповідей I Міжнар. наук.-практ. конф.* : Міжнар. наук.-практ. конф., м. Луцьк-Світязь, 13–16 черв. 2024 р. 2024.
12. Юзва А. С., Собчук В. В. Автоматизація процесів медичної діагностики із застосуванням нейромереж. Тези доповідей V Міжнар. наук.-практ. конф. : Міжнар. наук.-практ. конф., м. Одеса, 22 листоп. 2024 р. 2024.
13. Лаптев, О. і Юзва, А. 2024. ДОСЛІДЖЕННЯ ВИКОРИСТАННЯ ТЕХНІК МОДЕЛЮВАННЯ ДЛЯ АНАЛІЗУ ВІДГУКІВ КЛІЄНТІВ. Прикладні проблеми комп'ютерних наук, безпеки та математики. 2 (Бер 2024), 4–17.
14. Papers with Code - A systematic review: Deep learning-based methods for pneumonia region detection. *The latest in Machine Learning | Papers With Code*. URL: <https://paperswithcode.com/paper/a-systematic-review-deep-learning-based> (date of access: 03.11.2024).
15. Papers with Code - xLSTM-UNet can be an Effective 2D & 3D Medical Image Segmentation Backbone with Vision-LSTM (ViL) better than its Mamba Counterpart. *The latest in Machine Learning | Papers With Code*. URL: <https://paperswithcode.com/paper/xlstm-unet-can-be-an-effective-2d-3d-medical> (date of access: 03.11.2024).
16. Colorectal cancer incidence, mortality, and stage distribution in European countries in the colorectal cancer screening era: an international population-based study / R. Cardoso et al. *The Lancet Oncology*. 2021. Vol. 22, no. 7. P. 1002–1013. URL: [https://doi.org/10.1016/s1470-2045\(21\)00199-6](https://doi.org/10.1016/s1470-2045(21)00199-6) (date of access: 02.11.2024).

ДОДАТКИ

ДОДАТОК А

Технічне завдання на проєктування та розробку програмного засобу для медичної діагностики за допомогою нейромережових технологій

1. Вступ

1.1. Назва продукту

NeuroMedDiag: Медична Діагностика за допомогою Нейромережових Технологій.

1.2. Коротка характеристика програмного продукту

Створений продукт повинен сприяти покращенню та пришвидшенню діагностики легеневих захворювань Користувачі зможуть завантажувати медичні зображення, які автоматично аналізуватимуться програмою для виявлення патологій та аномалій. Результати аналізу надаються у вигляді діагностичних висновків та рекомендацій, що допомагає медичному персоналу в прийнятті інформованих рішень щодо лікування пацієнтів.

2. Підстави для розробки

2.1. Документ, на підставі якого ведеться розробка

Розробка даного програмного забезпечення ведеться на підставі дослідження, яке проводиться у рамках кваліфікаційної роботи.

2.2. Організація, що затвердила документ

Волинський національний університет імені Лесі Українки

3. Призначення розробки

Програмне забезпечення призначене для полегшення медичної діагностики легеневих захворювань.

4. Вимоги до програмного продукту

4.1. Вимоги до функціональних характеристик

- здатність завантажувати та обробляти медичні зображення з різних джерел (наприклад, МРТ, КТ).

- автоматичне визначення аномалій або патологій на зображеннях з використанням алгоритмів комп'ютерного зору;
- можливість класифікації захворювань або станів на основі оброблених зображень;
- надання діагностичних рекомендацій та висновків на основі аналізу медичних даних;
- можливість інтеграції з іншими медичними інформаційними системами для обміну даними та отримання додаткової інформації про пацієнтів;
- генерація звітів та графіків з результатами аналізу для подальшого використання медичним персоналом;
- автоматизація рутинних процесів, таких як імпорт даних, обробка зображень та генерація звітів;
- можливість попередньої обробки даних та виконання аналізу без необхідності втручання користувача.

4.2. Вимоги до надійності

Веб сайт зберігає працездатність незалежно від розмірів вікна браузера. Хороша оптимізація сайту, що забезпечує швидке завантаження сторінок.

4.3. Умови експлуатації

- підтримка операційних систем, на яких працює програмне забезпечення (наприклад, Windows, MacOS, Linux);
- вимоги до обсягу оперативної пам'яті та місця на диску для встановлення програми;
- можливість використання програми на персональних комп'ютерах та інших пристроях (якщо це доцільно);
- забезпечення безпеки та конфіденційності медичних даних, відповідно до вимог законодавства про захист персональних даних (наприклад, GBRT);
- можливість інтеграції з існуючими системами обміну медичною інформацією;
- регулярні оновлення програмного забезпечення для забезпечення безпеки та виправлення помилок;

- наявність технічної підтримки для вирішення технічних проблем та запитань користувачів;
- наявність документації та навчальних матеріалів для користувачів;
- проведення навчальних семінарів або онлайн-курсів для користувачів щодо використання програми;
- визначення моделі ліцензування програмного забезпечення (наприклад, оплата за користування або ліцензійні виплати);
- обговорення вартості використання програми та можливостей щодо різних планів та тарифів для різних типів користувачів (лікарні, клініки, індивідуальні лікарі тощо).

4.4. Вимоги до складу і параметрів технічних засобів

4.5. Вимоги до інформаційної і програмної сумісності

Кросплатформенність.

4.6. Вимоги до маркування і упаковки

Легкий у використанні, є назва та контакти для звернень.

4.7. Вимоги до транспортування і збереження

5. Вимоги до програмної документації

Попередній склад програмної документації: “Технічне завдання”.

6. Стадії та етапи розробки

6.1. Створення 9 вебсторінок.

6.2. Створення структури сайту.

6.3. Створення інтерфейсу сайту.

6.4. Проектування дизайну вебпорталу.

6.5. Підключення внутрішніх та зовнішніх посилань.

6.6. Завантаження контенту на веб сайт.

6.7. Тестування та налагодження.

6.8. Підключення до CMS.

6.9. Тестування та налагодження.

7. Порядок контролю та приймання

7.1. Тестування функціоналу.

7.2. Тестування кросплатформенності.

7.3. Перевірка контенту.

Якщо вказані тести проходять успішно, то програмний засіб передається замовнику.

Керівництво користувачу

Веб сайт (Рис. Б.1) створений для підтримки нейромережевої моделі для медичної діагностики легеневих захворювань.

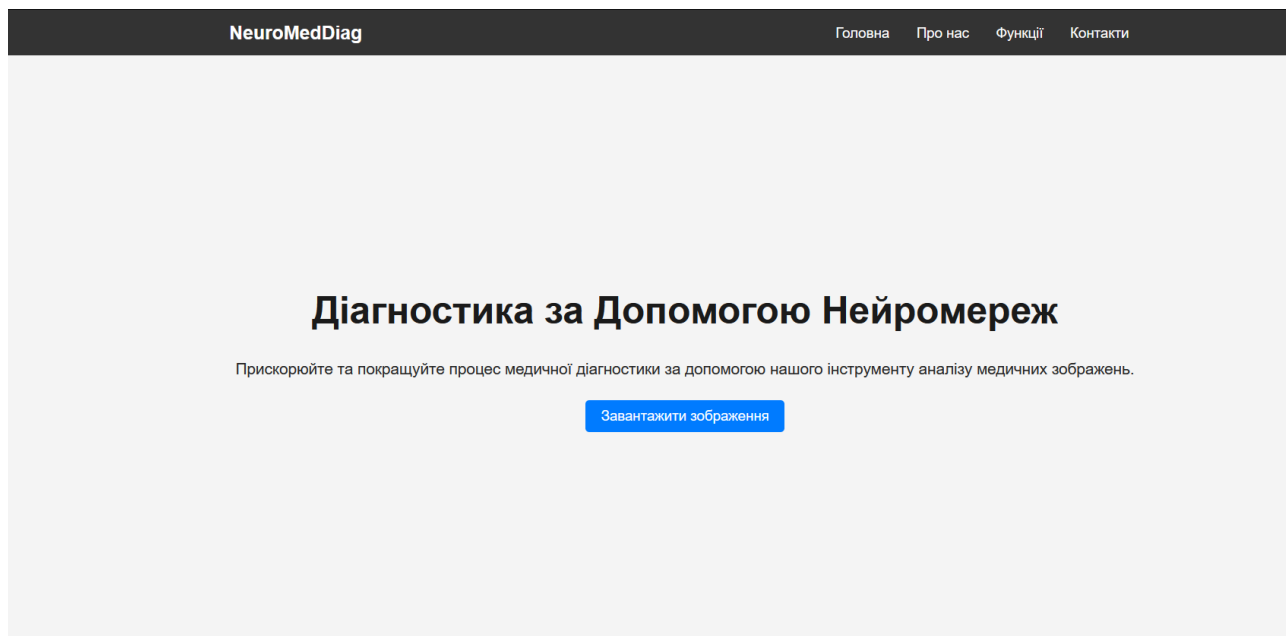


Рисунок Б.1 – Веб сайт для медичної діагностики легеневих захворювань, за допомогою нейромереж

У верхньому лівому куті знаходиться назва сайту «NeuroMedDiag» (Рис. Б.1). У правому верхньому куті (Рис. Б.1) знаходяться посилання на основні блоки сайту «Головна», «Про нас», «Функції», «Контакти».

Для початку роботи потрібно натиснути кнопку «Завантажити зображення» у верхньому блоці сайту (Рис. Б.1) або у блоці «Завантажити зображення» (Рис. Б.2).

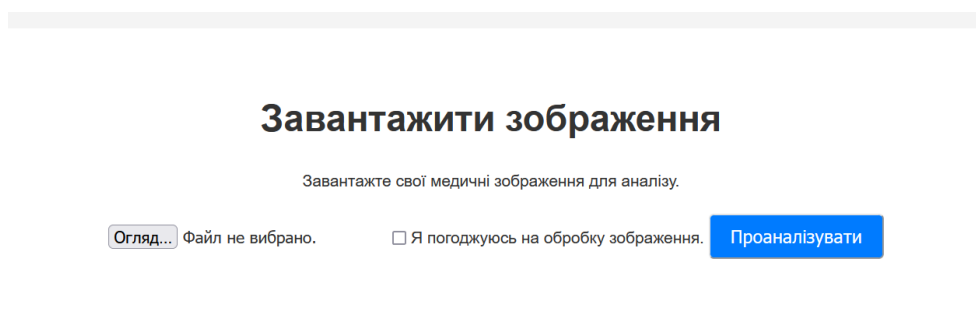


Рисунок Б.2 – Блок «Завантажити зображення» на веб сайті

Після обробки даних користувача перекине до блоку результатів(Рис. Б.3) та рекомендацій(Рис. Б.4).

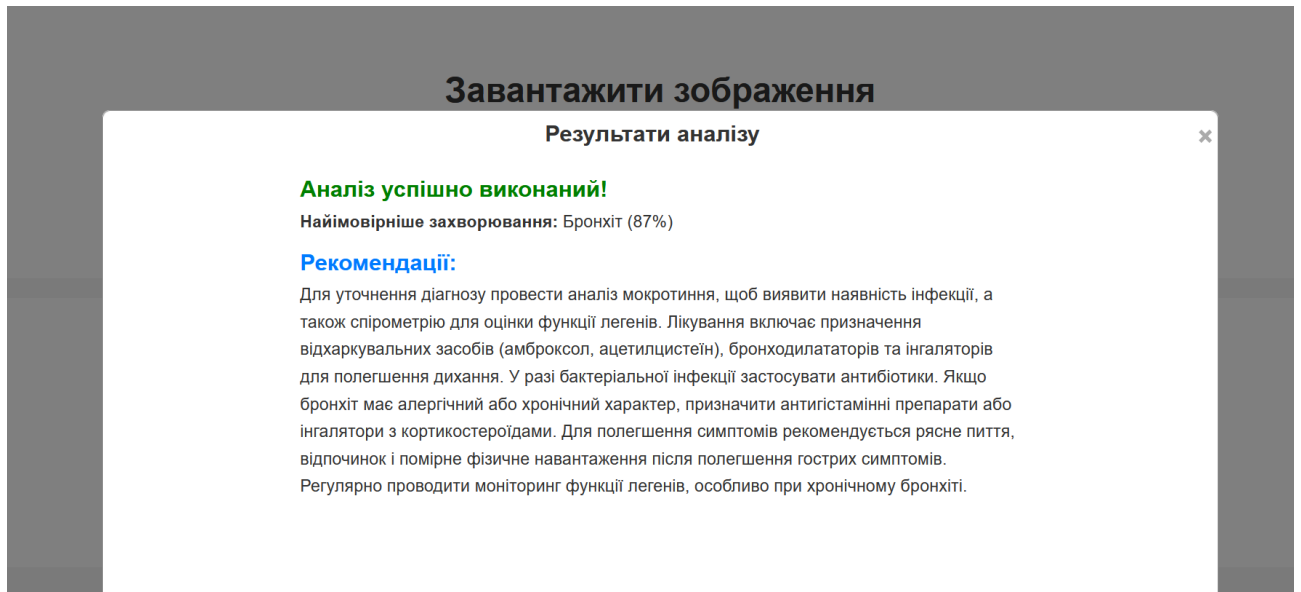


Рисунок Б.3 – Блок результатів роботи нейромережевої моделі

АНОТАЦІЯ

Юзва А.С. Нейромережеві технології в медичній діагностиці легеневих захворювань. Рукопис.

Кваліфікаційна робота на здобуття освітнього ступеня «магістр» за спеціальністю 122 Комп'ютерні науки. Волинський національний університет імені Лесі Українки, Луцьк, 2024 р.

У роботі досліджено ефективність нейромережевих моделей для автоматизованої діагностики легеневих захворювань з використанням сучасних алгоритмів штучного інтелекту. Проаналізовано актуальні підходи до застосування нейронних мереж у медичній практиці, зокрема методи класифікації та сегментації медичних зображень, які базуються на згорткових нейронних мережах (CNN) та їх модифікаціях. Вивчено механізми роботи алгоритмів, що використовуються для аналізу структурних змін у тканинах легень, зокрема виявлення аномалій, таких як пневмонія, туберкульоз та інші патологічні процеси. Особливу увагу приділено методам попередньої обробки медичних зображень, які забезпечують високу якість вхідних даних для навчання моделей. Проаналізовано також методи аугментації, які дозволяють збільшити обсяг та різноманітність навчальних даних, покращуючи тим самим узагальнюючі здібності моделей. Досліджено техніки навчання нейромереж, включаючи перенавчання, використання попередньо навчених моделей, а також їх адаптацію для медичних задач.

Результати дослідження включають створення інструменту підтримки прийняття рішень для лікарів, який підвищує якість діагностики легеневих захворювань. Було розроблено методіку класифікації патологій на рентгеновських знімках, що базується на специфічних особливостях медичних даних. Впровадження цих технологій дозволить автоматизувати діагностичні процеси, знизити навантаження на лікарів і підвищити точність діагностування.

Практичне значення роботи полягає у вдосконаленні процесу медичної діагностики через використання нейромережевих технологій, що сприятиме покращенню медичних послуг, зокрема у віддалених регіонах.

Ключові слова: нейромережеві технології, медична діагностика, легеневі захворювання, рентгенографія, згорткові нейронні мережі, машинне навчання, обробка медичних зображень.

ABSTRACTS.

Neural network technologies in medical diagnostics of pulmonary diseases.

Manuscript.

Qualification work for the educational degree “Master” in specialty 122 Computer Science. Lesya Ukrainka Volyn National University, Lutsk, 2024.

The paper investigates the effectiveness of neural network models for automated diagnosis of pulmonary diseases using modern artificial intelligence algorithms. The current approaches to the use of neural networks in medical practice, in particular, methods of classification and segmentation of medical images based on convolutional neural networks (CNN) and their modifications, are analyzed. The mechanisms of operation of algorithms used to analyze structural changes in lung tissue, in particular, to detect anomalies such as pneumonia, tuberculosis, and other pathological processes, are studied. Particular attention is paid to methods of medical image preprocessing that provide high-quality input data for model training. The augmentation methods are also analyzed, which allows for increasing the volume and diversity of training data, thereby improving the generalization abilities of the models. Techniques for training neural networks, including retraining, using pre-trained models, and adapting them for medical tasks, are investigated.

The results of the study include the creation of a decision support tool for doctors that improves the quality of pulmonary disease diagnosis. A methodology for classifying pathologies in X-ray images based on specific features of medical data was developed. The introduction of these technologies will automate diagnostic processes, reduce the workload of doctors, and increase the accuracy of diagnosis.

The practical significance of the work is to improve the process of medical diagnosis through the use of neural network technologies, which will improve medical services, in particular in remote regions.

Keywords: neural network technologies, medical diagnostics, pulmonary diseases, radiography, convolutional neural networks, machine learning, medical image processing.