

Харківський національний університет радіоелектроніки

Факультет _____ комп'ютерної інженерії та управління _____

Кафедра _____ електронних обчислювальних машин _____

Рівень вищої освіти _____ другий (магістерський) _____

Спеціальність _____ 123 «Комп'ютерна інженерія» _____
(код і повна назва)

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

Освітня програма _____ Системне програмування _____
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

“ _____ ” _____ 20__ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві _____ Ровенчаку Владиславу Миколайовичу _____
(прізвище, ім'я, по батькові)

1. Тема роботи _____ Метод обробки зображень з використанням машинного навчання _____

затверджена наказом по університету від “ 21 ” квітня 2025 р. № 296 Ст

2. Термін подання здобувачем роботи до екзаменаційної комісії _____ 16 червня 2025 р.

3. Вхідні дані до роботи _____

_____ обробка медичних зображень _____

_____ Google Colab _____

_____ метод обробки зображень _____

4. Перелік питань, що потрібно опрацювати у роботі _____

_____ Аналіз предметної області _____

_____ Аналіз сучасних методів обробки зображень на основі машинного навчання _____

_____ Практичне застосування методу обробки зображень з використанням машинного _____

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій 15 слайдів

6. Консультанти розділів роботи (заповнюється за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Строк / терміни виконання етапів роботи	Примітка
1	Отримання завдання та аналіз літератури	21.04.2025–30.04.2025	
2	Огляд існуючих рішень та алгоритмів	01.05.2025–12.05.2025	
3	Розробка методу	13.05.2025–22.05.2025	
4	Вибір програмних засобів	23.05.2025–30.05.2025	
5	Програмна реалізація	31.05.2025–02.06.2025	
6	Аналіз отриманих результатів	03.06.2025–05.06.2025	
7	Оформлення записки	06.06.2025–15.06.2025	

Дата видачі завдання “ 21 ” квітня 2025 р.

Здобувач


(підпис)

Керівник роботи

(підпис)

доц. Володимир ФЕДОРЧЕНКО

(посада, власне ім'я, прізвище)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 67 с., 15 рис., 2 дод., 8 джерел.

ОБРОБКА ЗОБРАЖЕНЬ, МАШИННЕ НАВЧАННЯ, ГЛИБОКЕ НАВЧАННЯ, ЗГОРТКОВА НЕЙРОННА МЕРЕЖА, RESNET, ПНЕВМОНІЯ, РЕНТГЕНОГРАФІЯ ГРУДНОЇ КЛІТКИ, MEDMNIST, МЕДИЧНА ВІЗУАЛІЗАЦІЯ, АВТОМАТИЗОВАНА ДІАГНОСТИКА.

Метою кваліфікаційної роботи є дослідження та реалізація ефективного методу обробки зображень, який дозволяє з високою точністю розпізнавати ознаки пневмонії на рентгенівських знімках. Актуальність обраної теми обумовлена необхідністю розробки сучасних інструментів підтримки клінічного прийняття рішень, що здатні автоматизувати первинний аналіз великої кількості медичних зображень у системах охорони здоров'я.

У ході виконання кваліфікаційної роботи було проаналізовано теоретичні засади класичної та сучасної обробки зображень, зокрема із залученням глибокого навчання. Здійснено огляд найбільш ефективних архітектур згорткових нейронних мереж, таких як CNN, ResNet і U-Net, а також вивчено їхнє застосування у сфері медичної візуалізації.

Практична частина роботи охоплює побудову й навчання моделей на основі відкритого набору даних ChestMNIST, що містить рентгенівські зображення грудної клітки у скороченому форматі. Розроблена згорткова нейронна мережа, а також реалізовано глибшу архітектуру на основі ResNet, адаптовану до класифікації двох класів (норма / пневмонія). Для оцінки результатів застосовано метрики точності, матрицю неточностей, ROC-криву та PR-криву. З метою покращення діагностичної інтерпретації були реалізовані візуалізації правильних і хибних передбачень моделі.

ABSTRACT

Master's thesis: 67 pages, 15 figures, 2 appendices, 8 sources.

IMAGE PROCESSING, MACHINE LEARNING, DEEP LEARNING, CONVOLUTIONAL NEURAL NETWORK, RESNET, PNEUMONIA, CHEST RADIOGRAPHY, MEDMNIST, MEDICAL IMAGING, AUTOMATED DIAGNOSIS.

The major goal of this thesis is to investigate and implement an effective image processing method capable of accurately detecting signs of pneumonia in chest X-ray images. The relevance of the topic is driven by the increasing need for modern clinical decision-support tools that can automate the preliminary analysis of large volumes of medical images within healthcare systems.

In order to both classical and modern theoretical foundations of image processing, with a focus on approaches based on deep learning. The research includes a comprehensive review of the most effective convolutional neural network architectures, such as CNN, ResNet, and U-Net, and their application in the field of medical imaging.

The practical component of the work involves designing and training models using the publicly available ChestMNIST dataset, which comprises downsized chest X-ray images. A convolutional neural network was developed, and a deeper ResNet-based architecture was implemented for binary classification (normal / pneumonia). Model performance was evaluated using accuracy metrics, a confusion matrix, ROC curve, and PR curve. To enhance diagnostic interpretability, the visualization of both correct and incorrect model predictions was provided.

ЗМІСТ

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ	8
ВСТУП	9
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ	12
1.1 Основні поняття та визначення обробки зображень	12
1.2 Методи класичної обробки зображень	13
1.3 Поняття і принципи роботи машинного навчання	14
1.4 Класифікація алгоритмів машинного навчання.....	15
1.5 Використання нейронних мереж для обробки зображень.....	17
1.6 Аналіз останніх досліджень і публікацій	18
2 АНАЛІЗ СУЧАСНИХ МЕТОДІВ ОБРОБКИ ЗОБРАЖЕНЬ НА ОСНОВІ МАШИННОГО НАВЧАННЯ.....	23
2.1 Огляд популярних нейронних архітектур	25
2.2 Методи підготовки та попередньої обробки даних	26
2.3 Вибір та оптимізація гіперпараметрів моделей	28
2.4 Програмні засоби і бібліотеки для роботи з зображеннями	29
2.5 Розробка методу	31
3 ПРАКТИЧНЕ ЗАСТОСУВАННЯ МЕТОДУ ОБРОБКИ ЗОБРАЖЕНЬ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ.....	34
3.1 Опис предметної області та набору даних	34
3.2 Розробка методу обробки зображень з використанням нейронної мережі	36
3.3 Реалізація алгоритму та опис експерименту	37
3.4 Аналіз отриманих результатів	51
ВИСНОВКИ.....	53
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	55
ДОДАТОК А Графічний матеріал кваліфікаційної роботи.....	56
ДОДАТОК Б Програмний код.....	65

Б.1 Встановлення бібліотек, завантаження MedMNIST	65
Б.2 Побудова моделі CNN, навчання моделі, візуалізація результатів навчання	65
Б.3 Оцінка моделі, ROC-крива, класифікаційний звіт, приклади класифікації зображень	66

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

AI – штучний інтелект

AUC – площа під кривою

CNN – згорткова нейронна мережа

CSV – формат табличних даних, розділених комами

GAN – генеративна змагальна мережа

GPU – графічний процесор

PR – точність-повнота

ReLU – випрямлена лінійна одиниця

ResNet – залишкова нейронна мережа

RNN – рекурентна нейронна мережа

ROC – крива операційних характеристик приймача

TPU – тензорний процесор

XAI – пояснюваний штучний інтелект

ВСТУП

У сучасних умовах розвитку цифрових технологій та інформатизації суспільства спостерігається стрімке зростання інтересу до автоматизованих систем аналізу візуальної інформації. Однією з найбільш актуальних сфер застосування таких технологій є охорона здоров'я, де швидкість і точність постановки діагнозу мають вирішальне значення для збереження життя та здоров'я пацієнтів. Медична візуалізація, що охоплює такі методи, як рентгенографія, комп'ютерна томографія, магнітно-резонансна томографія та ультразвукове дослідження, генерує величезні обсяги зображень, які потребують кваліфікованого аналізу. Однак, через обмеженість людських ресурсів, велику варіативність патологічних змін та фактор суб'єктивності, виникає необхідність у створенні інтелектуальних систем підтримки прийняття клінічних рішень.

Останні роки позначені значним прогресом у галузі штучного інтелекту, зокрема машинного навчання, яке дозволяє створювати моделі, здатні до самонавчання на основі великих обсягів даних. Особливо успішними виявилися методи глибокого навчання, зокрема згорткові нейронні мережі (CNN), які демонструють високу ефективність у задачах класифікації, розпізнавання та сегментації зображень. Використання таких алгоритмів у медичній діагностиці відкриває нові перспективи для автоматичного виявлення захворювань на ранніх стадіях, що є критично важливим для покращення якості та своєчасності медичної допомоги.

Об'єктом даного дослідження є процес обробки медичних зображень із метою виявлення ознак патологій. Предметом дослідження виступає застосування методів машинного навчання, зокрема глибоких згорткових нейронних мереж, для класифікації рентгенівських знімків грудної клітки з метою діагностики пневмонії. Ця патологія була обрана як приклад через її

високу поширеність, потенційну загрозу життю пацієнта та наявність великих відкритих наборів даних для навчання моделей.

Метою роботи є реалізація ефективного методу обробки зображень, який дозволяє з високою точністю розпізнавати ознаки пневмонії на рентгенівських знімках. У ході дослідження здійснено аналіз теоретичних основ комп'ютерного зору та машинного навчання, проведено огляд сучасних архітектур згорткових нейронних мереж, реалізовано експериментальну модель, проведено її навчання та тестування на датасеті pneumoniaMNIST. Особливу увагу приділено оцінці точності, чутливості та специфічності побудованої моделі, а також візуалізації результатів класифікації та аналізу потенційних помилок.

Завдання:

- провести аналіз сучасних методів обробки зображень та алгоритмів машинного навчання, що застосовуються для медичної діагностики;
- дослідити архітектури глибоких нейронних мереж з метою оцінки їхньої придатності для виявлення пневмонії на рентгенівських знімках;
- обрати відповідний відкритий набір медичних зображень та здійснити попередню обробку даних;
- розробити та реалізувати модель глибокого навчання для класифікації знімків на наявність або відсутність пневмонії;
- провести навчання та тестування моделі, з використанням стандартних метрик точності, чутливості, специфічності та AUC;
- здійснити візуалізацію результатів роботи моделі, включаючи правильні та хибні класифікації;
- оцінити ефективність розробленого методу та порівняти його з іншими підходами.

Наукова новизна дослідження полягає в застосуванні адаптованої архітектури згорткової нейронної мережі до компактного, але репрезентативного медичного набору даних, а також у проведенні комплексного аналізу ефективності моделі з урахуванням її можливого

практичного використання в клінічній практиці. Практичне значення полягає в тому, що розроблений метод може бути використаний як основа для створення автоматизованої системи підтримки лікаря-радіолога, зменшуючи навантаження на медичний персонал та підвищуючи достовірність діагнозу.

Таким чином, дана робота є спробою поєднати теоретичні знання в галузі машинного навчання з практичними завданнями медичної діагностики, що становить актуальний та перспективний напрям наукових досліджень в умовах розвитку цифрової медицини.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Основні поняття та визначення обробки зображень

Обробка зображень є міждисциплінарною галуззю, що охоплює методи аналізу, трансформації та інтерпретації цифрових візуальних даних з метою покращення їхньої інформативності, підвищення якості або автоматичного вилучення зображуваної інформації. Вона базується на математичних і алгоритмічних підходах, що дозволяють здійснювати маніпуляції із цифровим зображенням як з об'єктом дискретного представлення реального світу. Цифрове зображення, в цьому контексті, розглядається як матриця пікселів, кожен з яких характеризується певними параметрами – яскравістю, кольором, просторовим положенням. У випадку чорно-білих зображень значення пікселів визначаються відтінками сірого, тоді як кольорові зображення представляються сукупністю каналів, наприклад, у форматі RGB або інших колірних просторах.

Обробка зображень охоплює як етапи попередньої підготовки даних, так і більш складні процедури аналізу та інтерпретації. Зокрема, йдеться про операції, пов'язані з усуненням шумів, корекцією яскравості та контрасту, масштабуванням, поворотом, а також перетворенням зображення до формату, зручного для подальшої обробки. Залежно від рівня складності, ці операції можуть мати як низькорівневий, так і високорівневий характер. До перших належать базові трансформації, спрямовані на технічну обробку сигналу, тоді як другі охоплюють завдання, пов'язані з ідентифікацією об'єктів, розпізнаванням структур, інтерпретацією візуальних патернів і навіть прийняттям рішень.

У контексті медичної інформатики обробка зображень набуває особливого значення, адже дозволяє підвищити точність діагностичних висновків, автоматизувати процеси виявлення патологій, а також створювати

передумови для впровадження систем підтримки прийняття рішень. При цьому критично важливим є поєднання класичних алгоритмів обробки зображень із сучасними технологіями машинного навчання, що відкриває нові горизонти у медичній візуалізації та комп'ютерній діагностиці.

1.2 Методи класичної обробки зображень

Класична обробка зображень охоплює сукупність алгоритмів, що базуються на визначених математичних операціях, які застосовуються до піксельних значень для досягнення конкретних цілей – покращення якості зображення, виділення особливостей або підготовка зображення до подальшого аналізу. На відміну від методів машинного навчання, які навчаються на даних, класичні методи зазвичай не потребують етапу навчання та працюють за заздалегідь визначеними правилами.

Одним із базових напрямів є фільтрація зображень, яка дозволяє зменшити шум, згладити або підкреслити певні деталі. До найпоширеніших видів фільтрів належать гаусівські, медіанні, бокс-фільтри, а також високочастотні фільтри, що використовуються для підвищення різкості. Розмиття Гауса, наприклад, дозволяє усунути високочастотні компоненти, роблячи зображення більш гладким, у той час як оператор Лапласа або Собеля використовується для виявлення контурів та переходів інтенсивності.

Значне місце займають методи перетворення інтенсивності пікселів, які охоплюють такі операції, як нормалізація, логарифмічне перетворення, гістограмна рівномірність та покращення контрасту. Ці методи дозволяють краще виявити деталі в затемнених або перенасичених зонах зображення.

Іншим важливим напрямом є геометричні перетворення, які змінюють просторове представлення зображення. До таких трансформацій належать масштабування, повороти, віддзеркалення, зсуви та афінні перетворення. Вони забезпечують узгодження зображень у різних системах координат або

вирівнювання при попередній обробці для подальшого порівняння або об'єднання.

Важливо також згадати методи сегментації, які дозволяють відокремити значущі області в зображенні. Одним із найпростіших підходів є порогова обробка, де пікселі класифікуються на основі заданого порогу. Більш складні підходи включають регіональне зростання, вододіл, а також використання морфологічних операторів, таких як ерозія, дилатація, відкриття та закриття.

Незважаючи на стрімкий розвиток технологій машинного навчання, класичні методи зберігають свою актуальність, особливо в задачах з обмеженим обсягом даних, у передобробці або як частина гібридних систем. Вони також є фундаментальною основою, на якій базується розуміння більш складних інтелектуальних підходів до аналізу зображень.

1.3 Поняття і принципи роботи машинного навчання

Машинне навчання є напрямом штучного інтелекту, що базується на розробці алгоритмів, здатних до самостійного удосконалення шляхом аналізу даних без явного програмування правил прийняття рішень. Основною ідеєю машинного навчання є використання статистичних методів для виявлення закономірностей у даних, на основі яких система може робити передбачення або приймати рішення щодо нових, невідомих раніше прикладів.

В основі роботи машинного навчання лежить процес побудови математичної моделі, яка апроксимує зв'язок між вхідними ознаками та цільовими мітками. Така модель створюється в результаті навчання на основі тренувального набору даних, що включає приклади вхідних сигналів та відповідних очікуваних результатів. Якість моделі визначається здатністю узагальнювати закономірності, тобто демонструвати високу точність не лише на тренувальних, а й на нових тестових даних.

Сутність навчального процесу полягає у мінімізації певної функції втрат, яка кількісно оцінює відхилення передбачень моделі від правильних відповідей. Оптимізація здійснюється за допомогою алгоритмів, таких як градієнтний спуск, які поступово коригують параметри моделі в напрямку зменшення помилки. У процесі навчання важливу роль відіграють гіперпараметри, структура моделі, обсяг і якість даних, а також методи регуляризації, що запобігають перенавчанню.

Існує кілька основних парадигм машинного навчання: навчання з учителем (supervised learning), де модель вчиться на розмічених даних; навчання без учителя (unsupervised learning), що передбачає виявлення внутрішніх закономірностей без міток; та навчання з підкріпленням (reinforcement learning), при якому агент взаємодіє з середовищем і навчається на основі зворотного зв'язку у вигляді винагороди.

Машинне навчання застосовується в широкому спектрі практичних задач, зокрема в аналізі зображень, обробці природної мови, біоінформатиці, фінансовому моделюванні та інших. У контексті медичної візуалізації, ця технологія дозволяє створювати системи, здатні виявляти патології з точністю, порівнянною або навіть вищою за рівень людських експертів, що відкриває нові горизонти у створенні ефективних та надійних інструментів клінічної діагностики.

1.4 Класифікація алгоритмів машинного навчання

Алгоритми машинного навчання, залежно від характеру завдань та структури навчальних даних, поділяються на кілька основних категорій, кожна з яких має свої особливості, сфери застосування та вимоги до підготовки даних. Така класифікація дозволяє ефективніше обирати методи для розв'язання конкретних практичних задач, зокрема у сфері обробки зображень.

Провідне місце займають алгоритми навчання з учителем, які працюють із розміченими даними, де кожному прикладу вхідних характеристик відповідає певне значення виходу. У цьому випадку модель навчається знаходити функціональну залежність між вхідними ознаками та відомими мітками. До цієї категорії належать алгоритми класифікації та регресії. Перші використовуються для передбачення категоріальних міток, як-от у задачі розпізнавання захворювання за знімком, тоді як другі – для передбачення неперервних величин. Найвідомішими представниками є логістична регресія, дерева рішень, метод опорних векторів, k-найближчих сусідів, а також численні варіації згорткових і повнозв'язних нейронних мереж.

Інша категорія – це навчання без учителя, яке не передбачає наявності відповідних міток у даних. Мета такого навчання полягає у виявленні внутрішньої структури або прихованих закономірностей у вибірці. Типовими задачами є кластеризація, зниження розмірності та виявлення аномалій. Серед найпоширеніших алгоритмів слід виділити методи k-середніх, ієрархічну кластеризацію, алгоритми головних компонент та автоенкодери. Ці підходи особливо корисні на етапах попереднього аналізу даних або для виявлення латентних ознак у візуальній інформації.

Окрему групу складає навчання з підкріпленням, яке базується на взаємодії агента із середовищем з метою максимізації функції винагороди. Модель не має заздалегідь відомих правильних відповідей, натомість навчається на основі наслідків власних дій. Цей підхід частіше застосовується у задачах, пов'язаних із послідовними прийняттями рішень, керуванням роботами або стратегічними іграми, хоча останніми роками ведуться дослідження щодо його використання в медицині.

Крім згаданих основних напрямів, також виділяють напівконтрольоване та багатозадачне навчання, а також ансамблеві методи, які поєднують кілька моделей для досягнення більшої стабільності та

точності. Прикладами таких методів є випадкові ліси, градієнтне бустування та стеки.

Таким чином, правильний вибір алгоритму машинного навчання має ключове значення для успіху аналітичної системи і повинен ґрунтуватися на ретельному аналізі доступних даних, цільової задачі та вимог до точності й обчислювальних ресурсів.

1.5 Використання нейронних мереж для обробки зображень

Нейронні мережі, особливо їх згорткові варіанти, відіграють ключову роль у сучасних підходах до обробки зображень. Їх популярність зумовлена здатністю автоматично вивчати релевантні ознаки з необроблених піксельних даних, що принципово відрізняє їх від традиційних методів, де особливості обираються вручну. Завдяки глибоким архітектурам, нейронні мережі здатні моделювати складні, ієрархічні патерни візуальної інформації, що забезпечує високу ефективність у задачах класифікації, сегментації, детекції об'єктів та генерації зображень.

Згорткові нейронні мережі стали стандартом де-факто у сфері комп'ютерного зору. Їхня архітектура складається з послідовностей згорткових, pooling- та повнозв'язних шарів, які у сукупності дозволяють вилучати ознаки різного рівня складності: від базових контурів до високорівневих абстракцій. Особливістю CNN є інваріантність до зсувів і масштабів, що робить їх надзвичайно ефективними для аналізу структурованих візуальних даних, зокрема медичних зображень.

Іншим важливим напрямом є сегментувальні мережі, зокрема архітектура U-Net, яка спеціально розроблена для задач медичної сегментації. Вона дозволяє не лише класифікувати об'єкти, а й точно локалізувати їх у межах зображення, що є критично важливим, наприклад, при виявленні патологічних змін у тканинах. Завдяки симетричній структурі, що включає шляхи кодування та декодування з пропусками (skip-

connections), U-Net досягає високої точності навіть на обмежених наборах даних.

Окрім CNN, у обробці зображень також застосовуються рекурентні нейронні мережі у комбінації з CNN для аналізу послідовностей зображень, як-от у відео або серіях томографічних зрізів. Крім того, генеративно-змагальні мережі знаходять застосування в задачах підвищення роздільної здатності, генерації синтетичних знімків для розширення датасетів, а також у реконструкції пошкоджених фрагментів медичних зображень.

У медичній сфері нейронні мережі забезпечують не лише автоматичну класифікацію захворювань, а й дозволяють здійснювати диференціальну діагностику, визначати ступінь ураження, здійснювати прогнозування перебігу хвороби. Їхнє використання дає змогу створювати системи підтримки клінічних рішень, що працюють у реальному часі та адаптуються до нових даних. Висока точність та узагальнююча здатність нейромережевих моделей у поєднанні з розвитком апаратного забезпечення роблять їх незамінними інструментами у практиці сучасної цифрової медицини.

Таким чином, нейронні мережі не лише розширили технічні можливості обробки зображень, а й радикально змінили підхід до їх аналізу, відкривши нові горизонти для інтеграції штучного інтелекту в медичну діагностику та лікування.

1.6 Аналіз останніх досліджень і публікацій

За останні роки тематика інтелектуального аналізу даних із застосуванням машинного навчання стала надзвичайно актуальною, що виявляється у зростанні кількості спеціалізованих оглядів та систематичних досліджень у різних галузях. У цій частині огляду ми розглядаємо найвпливовіші публікації в фінансах, освіті, медицині та мультимедіа, а також методологічні статті про обробку дисбалансованих даних та відкриту науку.

Серед сучасних досліджень у сфері інтелектуального аналізу даних особливо вагоме місце посідає праця [2], присвячена комплексному аналізу методів виявлення шахрайства на основі технологій Data Mining та машинного навчання. Цей огляд охоплює широкий спектр прикладних сфер, серед яких провідними є банківська справа, страхування, телекомунікації та електронна комерція. Автори акцентують увагу на тому, що традиційні алгоритми виявлення шахрайських дій, попри свою ефективність у минулому, вже не справляються з викликами сучасних високошвидкісних та динамічних середовищ. В огляді запропоновано узагальнену класифікацію підходів, що охоплює як класичні методи, так і сучасні ансамблеві та гібридні архітектури, які комбінують кілька моделей для досягнення високої точності при роботі з незбалансованими наборами даних. Особливої уваги у публікації приділено проблемі інтерпретованості моделей, що є критично важливою у фінансовому секторі, де рішення повинні бути прозорими та обґрунтованими. У цьому контексті дослідники підкреслюють роль Explainable AI як наряду, що здатний поєднати обчислювальну потужність сучасних алгоритмів із потребами людського контролю й аудиту. Важливим внеском роботи є й наголос на тому, що адаптивні й контекстно-чутливі моделі, здатні працювати з обмеженими, фрагментарними або недостовірними даними, стають дедалі важливішими в епоху зростаючої складності шахрайських схем.

Іншою вагомою працею, що заслуговує на увагу, є оновлений систематичний огляд [3], присвячений застосуванню машинного навчання в освітній сфері. Автори простежують еволюцію Educational Data Mining і Learning Analytics протягом останнього десятиліття, аналізуючи як кількісні показники зростання публікацій, так і якісні зміни в методологічному підході до досліджень. Центральним фокусом цієї роботи є трансформація освітньої аналітики з прикладного інструменту до самостійної наукової дисципліни, що поєднує педагогіку, штучний інтелект, психологію, соціологію та статистику. У статті детально розглянуто технологічний ланцюг виявлення

знань у навчальних даних – від збору та обробки до інтерпретації та впровадження у навчальний процес. Автори наводять численні приклади використання класифікаційних моделей для прогнозування результатів навчання, аналізу стилів засвоєння матеріалу, побудови рекомендаційних систем у масових он-лайн-курсах, а також вивчають застосування глибокого навчання для розпізнавання поведінкових патернів студентів.

Особливої ваги у роботі набуває обговорення етичних питань та викликів, пов'язаних із захистом конфіденційності освітніх даних, забезпеченням прозорості аналітичних моделей та підтримкою принципів відкритої науки. Автори акцентують, що освітній контекст вимагає більш делікатного підходу до застосування аналітичних інструментів, ніж комерційні або фінансові системи, оскільки йдеться про взаємодію з людьми в умовах навчального процесу, де невіддале моделювання може призвести до викривлення освітньої траєкторії. У зв'язку з цим автори рекомендують впровадження практик preregistration, публікацію відкритих датасетів та реплікаційних експериментів, що сприятиме формуванню етично відповідально-го наукового середовища.

Дослідження [4] присвячене застосуванню машинного навчання для виявлення фінансових аномалій, що має ключове значення для підтримання порядку на ринках і захисту інтересів інвесторів. У роботі здійснено порівняльний аналіз ефективності традиційних статистичних методів і сучасних алгоритмів машинного навчання у завданнях виявлення аномалій. Дослідження охоплює як керовані, так і некеровані методи навчання, а також глибинне навчання, оцінюючи їх здатність працювати з високовимірними фінансовими даними та виявляти складні шаблони шахрайства. Результати свідчать про високу ефективність ансамблевих моделей, які дозволяють досягти оптимального балансу між точністю виявлення та інтерпретованістю. Водночас автор звертає увагу на низку викликів, серед яких особливе місце посідають проблема дисбалансу вибірки та труднощі з узагальненням моделей на нові дані. Як потенційне вирішення запропоновано використання

гібридних підходів, що поєднують знання експертів з перевагами аналізу на основі даних. Окремо наголошується на перспективності пояснювального штучно-го інтелекту як інструменту підвищення прозорості та довіри до автоматизованих систем виявлення фінансових аномалій. Отримані результати розширюють уявлення про можливості машинного навчання у фінансовій аналітиці та відкривають нові напрями для підвищення ефективності та надійності процесів виявлення порушень.

Одним із важливих міждисциплінарних прикладів застосування методів інтелектуального аналізу даних є дослідження, присвячені використанню машинного навчання в медичній сфері, зокрема у кардіології. У цьому контексті вирізняється праця [5], де автори провели масштабний аналіз понад сорока досліджень, що стосуються автоматизованого прогнозування ішемічної хвороби серця на основі клінічних та демографічних даних. Стаття є прикладом поглибленого огляду практичних моделей, серед яких провідне місце займають ансамблеві алгоритми, що комбінують точність моделей логістичної регресії, дерева рішень, градієнтного бустингу та нейронних мереж. Основна увага в дослідженні зосереджена на порівнянні моделей за медично релевантними метриками (чутливість, специфічність, точність). Ця праця підкреслює важливість не лише статистичної ефективності моделей, а й їхньої адаптованості до особливостей медичних систем, де кожне хибно позитивне чи хибно негативне рішення може мати серйозні наслідки для життя пацієнтів. Автори вказують на проблему обмеженої репрезентативності деяких відкритих медичних наборів даних, а також на важливість локальної валідації моделей для різних соціально-демографічних груп. Серед рекомендацій фігурує потреба в інтерпретованих алгоритмах, що здатні пояснити лікарю механізми прийняття рішень, а також у використанні гібридних моделей, які поєднують знання з медичних експертних систем і гнучкість сучасних методів машинного навчання.

У роботі [6] розглянуто роль методів інтелектуального аналізу даних, зокрема text mining, у процесі цифрової криміналістики. Автори акцентують увагу на зростаючій складності цифрових джерел та пристроїв, які фігурують у сучасних розслідуваннях, що унеможливорює ефективний аналіз суто вручну. Застосування моделей машинного навчання та алгоритмів розпізнавання шаблонів дозволяє виявляти приховані цифрові докази, які могли б залишитися непоміченими традиційними методами. Дослідження демонструє, що методи аналізу текстових даних, як складова Data Mining, стають критично важливими для підвищення точності й ефективності цифрової судово-експертної діяльності. Проведене оглядове дослідження охоплює основні підходи до цифрової криміналістики на основі Data Mining і висвітлює їх потенціал у подоланні викликів, пов'язаних із обсягом і складністю сучасних цифрових даних.

Загальна тенденція в усіх розглянутих роботах полягає в переході від класичних статистичних методів до більш гнучких, адаптивних, інтерпретованих і масштабованих рішень на основі машинного навчання. Водночас підкреслюється потреба в міждисциплінарному підході, який поєднує алгоритмічну точність з етичними, правовими та галузевими вимогами. Усі ці публікації формують цілісну картину сучасного стану досліджень у сфері data mining, демонструючи як теоретичні здобутки, так і практичні досягнення та перспективи подальшого розвитку.

Метою роботи є дослідження сучасних методів інтелектуального аналізу даних із використанням машинного навчання, а також вивчення їх застосування в різних предметних галузях для виявлення прихованих закономірностей, аномалій і ключової інформації в великих та складних масивах даних. Особливу увагу приділено порівнянню традиційних статистичних підходів із алгоритмами машинного навчання, аналізу переваг і обмежень сучасних моделей, а також виявленню перспективних напрямків розвитку, таких як пояснювальний штучний інтелект, гібридні моделі та етичні аспекти використання даних. Робота також має на меті узагальнити результати актуальних публікацій і сформулювати цілісне уявлення про роль машинного навчання як ключового інструменту сучасного аналізу даних.

2 АНАЛІЗ СУЧАСНИХ МЕТОДІВ ОБРОБКИ ЗОБРАЖЕНЬ НА ОСНОВІ МАШИННОГО НАВЧАННЯ

Упродовж останніх десятиліть обробка зображень зазнала кардинальних змін, зумовлених появою методів машинного та глибокого навчання (рисунок 2.1), що кардинально розширили можливості автоматичного аналізу візуальної інформації. Замість ручного формування ознак та жорстко заданих правил, сучасні системи навчаються розпізнавати візуальні патерни безпосередньо з даних. Такий підхід забезпечує значно вищу гнучкість, адаптивність та точність при виконанні задач класифікації, сегментації та генерації зображень.

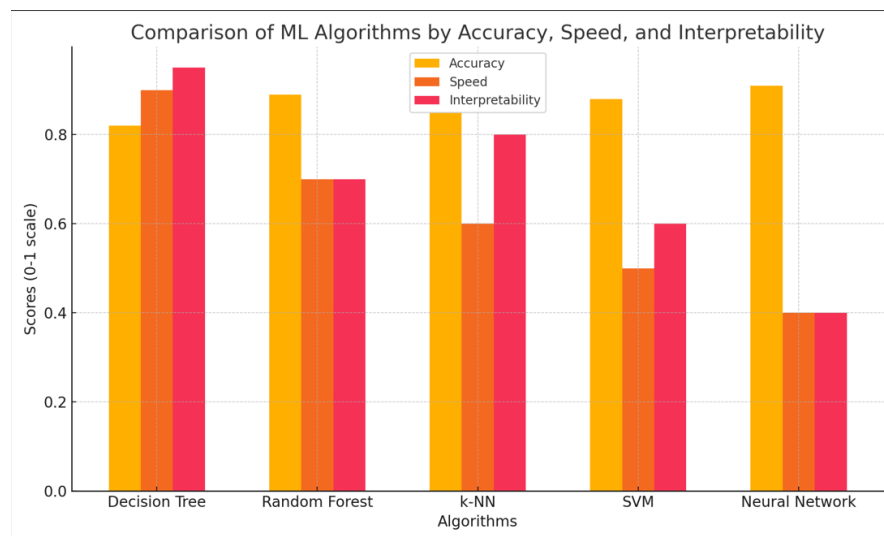


Рисунок 2.1 – Порівняльний аналіз методів машинного навчання

Одним із найбільш успішних архітектурних рішень стали згорткові нейронні мережі (CNN), які здатні автоматично виділяти ієрархічні ознаки з вхідного зображення. Їх багаторівнева структура дозволяє на нижчих рівнях виявляти прості форми, як-от контури та кути, а на вищих – складніші абстракції, як-от органи чи патології. В основі CNN лежить принцип згортки, що забезпечує просторову інваріантність, а також зменшення кількості параметрів порівняно з класичними повнозв'язними мережами.

На додаток до CNN, у задачах медичного аналізу зображень активно використовуються генеративно-змагальні мережі – двокомпонентні моделі, здатні генерувати нові зображення на основі розподілу вхідних даних. Їхнє застосування включає генерацію синтетичних медичних зображень для поповнення вибірок, відновлення втрачених фрагментів, підвищення роздільної здатності тощо. Ще одним напрямом є використання рекурентних нейронних мереж (RNN) та їх варіантів – LSTM і GRU – для обробки.

Надзвичайно важливим етапом побудови будь-якої моделі є підготовка даних. Застосовується попередня обробка, яка включає масштабування, нормалізацію, вирівнювання освітлення, видалення шуму та підвищення контрасту. Аугментація даних, що передбачає генерацію нових варіантів зображень шляхом обертання, масштабування, зсувів чи дзеркального відображення, відіграє ключову роль у підвищенні здатності моделі до узагальнення. Особливо актуально це в медичних задачах, де часто обмежена кількість даних та висока варіативність.

Гіперпараметри, такі як: кількість шарів, розмір ядра згортки, швидкість навчання, тип функції активації тощо, значною мірою визначають продуктивність і точність моделей. Їх оптимізація може здійснюватися вручну, або ж автоматизовано через сітковий пошук, байєсівську оптимізацію чи алгоритми еволюційного типу. Ефективна настройка гіперпараметрів дозволяє знайти баланс між точністю моделі та її здатністю до узагальнення.

Щодо програмних засобів, то найбільш популярними фреймворками для побудови та навчання моделей є TensorFlow та PyTorch. Вони надають гнучкі засоби для створення, відладки і візуалізації архітектур, що робить їх інструментами вибору як для наукових досліджень, так і для промислових розробок. Бібліотека OpenCV, своєю чергою, забезпечує потужні інструменти для попередньої обробки зображень, а такі пакети, як Albumentations, PIL і Scikit-image, полегшують інтеграцію класичних і сучасних підходів до формування ознак.

Підсумовуючи, слід зазначити, що сучасні методи обробки зображень, побудовані на основі машинного навчання, забезпечують новий рівень точності, ефективності та автоматизації процесів аналізу візуальних даних. Їх застосування в медичній сфері, зокрема в діагностиці пневмонії, демонструє реальні переваги впровадження штучного інтелекту у клінічну практику та підтверджує доцільність подальших досліджень у цьому напрямі.

2.1 Огляд популярних нейронних архітектур

У сучасних підходах до обробки зображень особливе місце займають нейронні мережі, які, залежно від завдань, організовуються у різні архітектурні конфігурації. Найбільш поширеними й ефективними є CNN, RNN та GAN (рисунок 2.2). Кожна з цих архітектур має свої особливості, принципи функціонування і сфери застосування, що визначає доцільність їх використання в конкретних умовах.

Архітектура	Призначення	Ключові особливості
CNN	Обробка зображень, відео, просторових даних	Згорткові фільтри, локальність, ієрархічне вилучення ознак
RNN	Обробка послідовностей, часових рядів, мови	Зворотні зв'язки, пам'ять про попередні стани
GAN	Генерація нових даних, зокрема зображень	Архітектура з двома мережами: генератор і дискримінатор
Архітектура	Переваги	Недоліки
CNN	Висока точність при класифікації зображень	Не пристосована до обробки послідовностей
RNN	Добре працює з часовими залежностями	Складність тренування, проблема з градієнтами
GAN	Здатність створювати реалістичні нові дані	Важке налаштування, нестабільність під час навчання

Рисунок 2.2 – CNN, RNN та GAN

Згорткові нейронні мережі є домінуючою архітектурою в задачах аналізу зображень. Їх ефективність пояснюється здатністю автоматично виявляти просторові патерни завдяки використанню згорткових ядер, які «сканують» зображення і формують карти ознак. CNN відзначаються локальністю сприйняття, інваріантністю до зміщень та здатністю виявляти ієрархії ознак. Стандартні архітектури, як-от LeNet, AlexNet, VGG, ResNet, EfficientNet, є загальноприйнятими шаблонами для розв'язання класифікації.

Рекурентні нейронні мережі, на відміну від CNN, орієнтовані на роботу з послідовними даними, де важливо враховувати часову залежність або контекст попередніх станів. У контексті обробки зображень вони застосовуються переважно для аналізу послідовностей медичних знімків (наприклад, серій КТ чи МРТ-зрізів) або для задач опису зображення, де вихід є послідовністю слів. Через проблему зникаючого градієнта класичні RNN часто замінюються на вдосконалені варіації.

Генеративно-змагальні мережі є архітектурами нового покоління, що складаються з двох частин – генератора і дискримінатора. Генератор намагається створити фотореалістичні зображення, тоді як дискримінатор навчається відрізнити справжні зразки від штучно згенерованих. У процесі змагання обидві мережі вдосконалюються, в результаті чого генератор навчається створювати високоякісні синтетичні зображення. У медичній сфері GAN використовуються для збільшення обсягів вибірок даних.

Загалом, вибір архітектури нейронної мережі визначається не лише природою вхідних даних, а й конкретною постановкою задачі: класифікація, сегментація, генерація або аналіз послідовностей. Використання CNN, RNN та GAN у взаємодоповнюючому режимі створює передумови для розробки високоточних, гнучких і стійких до варіацій систем, здатних успішно функціонувати в умовах медичної практики.

2.2 Методи підготовки та попередньої обробки даних

Ефективність моделей машинного навчання, зокрема у сфері обробки зображень, значною мірою залежить від якості підготовки вхідних даних. Непідготовлені або погано структуровані дані можуть не лише знизити точність прогнозування, але й призвести до помилкових інтерпретацій, що особливо критично у медичних додатках. Тому підготовка та попередня обробка зображень є ключовим етапом у розробці аналітичних моделей, який передуює їх навчанню.

Одним з першочергових завдань є нормалізація зображень, що передбачає приведення значень пікселів до єдиного масштабованого діапазону, найчастіше $[0, 1]$ або $[-1, 1]$. Це забезпечує стабільність градієнтних методів оптимізації і покращує збіжність моделі. Важливим етапом є також уніфікація розміру зображень. У реальних умовах дані можуть мати різну роздільність, тому всі знімки зазвичай масштабуються до фіксованого розміру, який відповідає вхідному шару нейронної мережі.

З метою підвищення якості навчання широко застосовується фільтрація та згладжування зображень. Вона передбачає використання фільтрів (наприклад, гаусівських або медіанних) для зменшення шуму та артефактів, які можуть виникати через технічні обмеження апаратного забезпечення або особливості медичного обстеження. Такі операції допомагають моделі зосередитися на суттєвих ознаках, що мають діагностичне значення.

Аугментація даних є ще одним критично важливим інструментом у випадках, коли обсяг тренувального набору є обмеженим. Методологія аугментації включає генерацію нових зображень шляхом застосування випадкових трансформацій – обертання, зсувів, масштабування, віддзеркалення, зміни яскравості або контрасту. Це дозволяє штучно розширити вибірку, зробити модель стійкішою до варіацій у вхідних даних та зменшити ризик перенавчання.

У деяких випадках застосовується попередній аналіз або анотування зображень. Наприклад, можуть вручну або автоматично розмічатися області інтересу, такі як уражені тканини, вузли, утворення тощо. Ці розмітки використовуються у задачах сегментації або детекції об'єктів і дозволяють побудувати більш точні та локалізовані моделі.

Залежно від особливостей задачі, також може бути корисним балансування класів. У медичних наборах даних часто спостерігається диспропорція між здоровими і хворими зразками, що призводить до зміщення моделі в бік переважаючого класу. Для вирішення цієї проблеми

застосовуються методи підвибірки, надвибірки, генерації синтетичних зразків або вагового штрафування функції втрат.

Таким чином, попередня обробка та підготовка даних формує підґрунтя для створення ефективних моделей машинного навчання. Її якісне виконання є запорукою не лише високої точності класифікації, але й надійності результатів в умовах реального клінічного застосування.

2.3 Вибір та оптимізація гіперпараметрів моделей

Процес побудови ефективної моделі машинного навчання не обмежується лише вибором архітектури або набору даних. Одним із ключових факторів, що визначають її точність, стабільність та здатність до узагальнення, є правильний вибір та налаштування гіперпараметрів. Гіперпараметри – це зовнішні змінні моделі, які не визначаються в процесі навчання, але суттєво впливають на динаміку оптимізації та кінцеву якість прогнозу.

До найважливіших гіперпараметрів належить швидкість навчання, яка визначає розмір кроку, з яким модель коригує свої параметри на основі обчисленого градієнта. Занадто велике значення може призвести до нестабільності або розбіжності навчання, тоді як надто мале – до повільної збіжності або застрягання в локальних мінімумах. Інші критичні параметри включають розмір пакета, кількість епох, тип оптимізатора, а також функцію активації.

Особливу увагу слід приділяти структурним параметрам моделі, зокрема кількості шарів, кількості нейронів на шарі, розміру фільтрів у згорткових шарах, глибині мережі та рівню регуляризації. Надмірна кількість параметрів може призвести до перенавчання, тоді як занадто проста модель – до недонавчання. Застосування механізмів регуляризації, таких як L1/L2-пеналізація або дропаут, дозволяє зменшити складність моделі без втрати її узагальнюючої здатності.

Процес оптимізації гіперпараметрів може здійснюватися вручну або автоматизовано. Найпростішим підходом є сітковий пошук, при якому випробовується комбінація фіксованих значень параметрів. Ефективнішим, але більш витратним є випадковий пошук, де параметри вибираються з визначених діапазонів випадковим чином. Останніми роками набувають популярності більш просунуті методи, зокрема байєсівська оптимізація, еволюційні алгоритми, а також гіперпараметричні фреймворки: Optuna та інші.

У задачах медичної діагностики особливо важливо забезпечити баланс між високою точністю моделі та її узагальненням на нові, раніше невідомі дані. Тому в процесі оптимізації гіперпараметрів слід використовувати надійні методи валідації, такі як: крос-валідація, а також враховувати специфіку медичних даних: невелику кількість вибірки, незбалансованість класів, чутливість до помилок другого роду.

Загалом, процес оптимізації гіперпараметрів є невід'ємною частиною науково обґрунтованого підходу до побудови моделей машинного навчання. Він дозволяє значно покращити ефективність навіть простої моделі за умови правильного налаштування, що в умовах медичного застосування може мати вирішальне значення для надійності автоматизованої діагностики.

2.4 Програмні засоби і бібліотеки для роботи з зображеннями

Реалізація сучасних методів обробки зображень та машинного навчання неможлива без використання спеціалізованого програмного забезпечення, яке забезпечує зручний інтерфейс, оптимізовані обчислення та підтримку глибоких архітектур. Найбільш популярними фреймворками у цій галузі є TensorFlow, PyTorch та OpenCV. Їх застосування дозволяє скоротити час розробки моделей, забезпечити масштабованість рішень та інтеграцію з виробничими середовищами.

TensorFlow – це потужний відкритий фреймворк для чисельних обчислень і глибокого навчання, розроблений компанією Google. Він підтримує як імперативний, так і декларативний стиль програмування, має високий рівень гнучкості, можливість розгортання на мобільних пристроях, у браузерях, на серверних та хмарних платформах. TensorFlow забезпечує повну підтримку згорткових і рекурентних мереж, а також таких інструментів як Keras, що значно спрощує створення моделей завдяки використанню високорівневого API.

PyTorch, розроблений Facebook AI Research, є ще одним провідним фреймворком, який став стандартом де-факто в академічному середовищі завдяки своїй простоті, динамічному графу обчислень і легкій налагоджуваності. PyTorch дозволяє швидко експериментувати з архітектурами, налаштовувати власні шари та втратні функції, а також інтегрувати моделі з такими бібліотеками, як torchvision і torchaudio. Бібліотека широко використовується у медичній діагностиці, зокрема у проектах, пов'язаних із сегментацією зображень.

OpenCV – це набір високоефективних інструментів для комп'ютерного зору і обробки зображень. OpenCV надає функціональність для роботи з відеопотоками, виявлення об'єктів, контурів, фільтрації зображень, морфологічних операцій, а також конвертації між форматами. Незважаючи на те, що OpenCV не є бібліотекою для глибокого навчання, вона часто використовується на етапах попередньої обробки або в поєднанні з іншими фреймворками для реалізації повноцінних аналітичних систем.

Фреймворк	Основне призначення	Сумісність з Google Colab
TensorFlow	Глибоке навчання, розробка нейронних мереж	Повна інтеграція, попередньо встановлений
PyTorch	Глибоке навчання, наукові дослідження та візуалізація	Повна інтеграція, попередньо встановлений
OpenCV	Обробка зображень, комп'ютерний зір	Повна інтеграція, попередньо встановлений
Фреймворк	Переваги	Недоліки
TensorFlow	Підтримка мобільних та веб-платформ, TPU	Складніша відлагодження, громіздкий синтаксис
PyTorch	Гнучкість, зручність для дослідників, динамічне обчислення	Менш зручний у виробничих середовищах
OpenCV	Швидкість, простота базових операцій зображень	Не призначений для глибокого навчання

Рисунок 2.3 – Порівняльний аналіз інструментів реалізації

Іншими корисними бібліотеками, що застосовуються у сфері аналізу зображень, є PIL/Pillow для базової роботи з графічними файлами, Scikit-image для класичних алгоритмів обробки зображень, а також Alumentations – високопродуктивна бібліотека для аугментації зображень. Разом ці інструменти формують екосистему, що дозволяє повноцінно реалізовувати як прототипи дослідницьких рішень, так і промислові системи.

Таким чином, вибір відповідного фреймворку або бібліотеки залежить від задачі, вимог до продуктивності, обчислювальних ресурсів і рівня інтеграції. У контексті медичної обробки зображень важливими критеріями залишаються точність, масштабованість і підтримка різних форматів медичних даних. Завдяки активному розвитку цих засобів, дослідники та практики мають змогу зосередитися на самій проблемі, не витрачаючи значних ресурсів на технічну реалізацію.

2.5 Розробка методу

У межах дослідження було реалізовано метод обробки рентгенівських зображень грудної клітки з метою виявлення ознак пневмонії на основі глибокого навчання (рисунок 2.4). Розробка методу ґрунтувалася на використанні публічного датасету ChestMNIST, який містить стандартизовані медичні зображення у форматі 28×28 пікселів. На початковому етапі дані було підготовлено до обробки шляхом нормалізації піксельних значень, формування батчів і перетворення у формат, придатний для подачі в згорткову нейронну мережу.

Було розроблено дві архітектурні конфігурації моделей: базову згорткову нейромережу та вдосконалену модифікацію на основі ResNet. У першому варіанті використовувались кілька послідовних згорткових і пулінгових шарів, що дозволяли вилучити ознаки різного рівня складності, з подальшою класифікацією у повнозв'язному шарі. У другому випадку реалізовано глибшу архітектуру із залишковими з'єднаннями, що дозволяло

уникнути затухання градієнтів при навчанні та забезпечити кращу узагальнюючу здатність моделі. Обидві моделі були навчені з використанням оптимізатора Adam і функції втрат на основі бінарної крос-ентропії, а також валідації на відкладеному наборі даних із застосуванням методів раннього зупинення та збереження найкращої моделі.

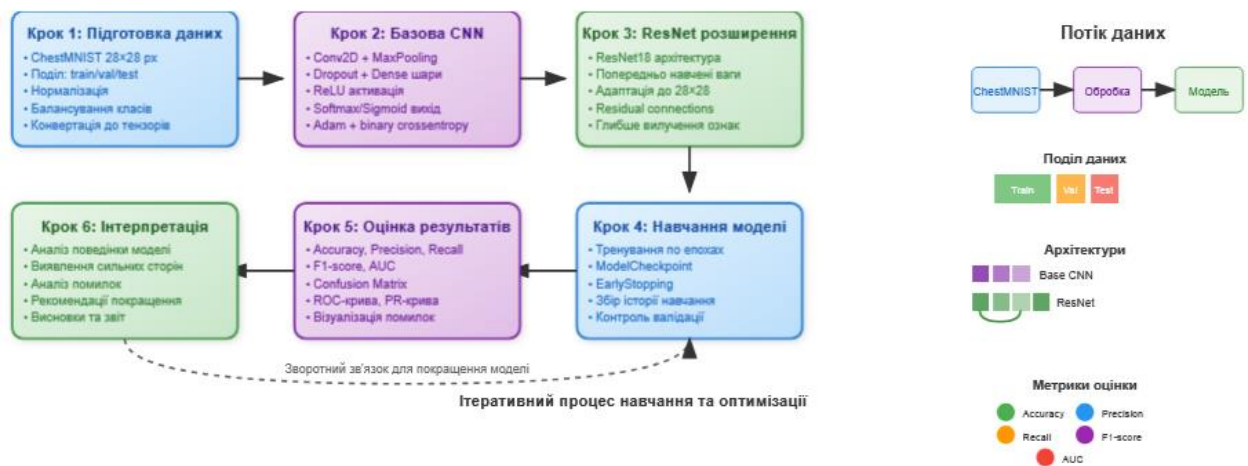


Рисунок 2.4 – Розроблений метод. Етапи

Оцінка ефективності реалізованого методу здійснювалась із застосуванням комплексного набору метрик, включно з точністю, повнотою, F1-показником, площею під ROC-кривою та середньою точністю PR-кривої. Крім того, було проведено візуальний аналіз поведінки моделі шляхом побудови матриці неточностей і відображення конкретних прикладів правильних та хибних класифікацій. Модель демонструє високу стабільність результатів, здатність до генералізації та практичну придатність до застосування в системах автоматизованої діагностики.

Крок 1. Вибір і підготовка даних:

- використовується датасет ChestMNIST із колекції MedMNIST, що містить зображення у форматі 28×28 пікселів;
- дані поділяються на тренувальний, валідаційний і тестовий набори;
- здійснюється нормалізація піксельних значень, перетворення до тензорів, а також, за потреби, балансування класів.

Крок 2. Побудова базової нейронної моделі:

- розробляється згортова нейронна мережа (CNN) з кількома шарами Conv2D, MaxPooling, Dropout і Dense;
- застосовуються функції активації ReLU, а на вихідному шарі – Softmax (для бінарної класифікації – sigmoid);
- компіляція моделі з використанням оптимізатора Adam, функції втрат binary crossentropy, та метрик точності.

Крок 3. Розширення моделі на основі ResNet:

- імплементується варіант ResNet (наприклад, ResNet18) з попередньо навченими вагами або без них;
- архітектура адаптується до розмірів вхідних зображень і кількості вихідних класів;
- забезпечується глибше вилучення ознак без втрати градієнтів (residual connections).

Крок 4. Навчання моделі:

навчання проводиться в кілька епох з використанням збереження найкращої моделі (ModelCheckpoint) та контролю за втратою на валідаційному наборі (EarlyStopping);

ведеться збір історії точності й втрат для подальшої візуалізації.

Крок 5. Оцінка результатів:

- обчислюються точність, recall, precision, F1-score;
- будується матриця неточностей, ROC-крива, PR-крива, а також обчислюється AUC;
- візуалізуються правильні та хибні передбачення моделі на прикладах зображень.

Крок 6. Інтерпретація та виводи:

- аналізується поведінка моделі, виявляються сильні сторони та потенційні помилки.

3 ПРАКТИЧНЕ ЗАСТОСУВАННЯ МЕТОДУ ОБРОБКИ ЗОБРАЖЕНЬ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ

Практична реалізація методів машинного навчання у задачах обробки медичних зображень потребує поєднання теоретичних засад, сучасних архітектур нейронних мереж та адаптації до специфіки медичних даних. Цей розділ присвячено поетапному опису розробки та впровадження моделі для автоматизованої класифікації патологій на медичних знімках, зокрема використання методів глибокого навчання для аналізу зображень з бази MedMNIST.

Розділ включає опис предметної області, характеристику вибраного набору даних, обґрунтування вибору архітектури моделі, методів попередньої обробки та аугментації, а також реалізацію експериментальної частини з навчанням, тестуванням та візуалізацією результатів. Особлива увага приділяється метрикам якості, аналізу похибок, порівнянню з базовими рішеннями та оцінці практичної значущості отриманих результатів.

У межах реалізації були використані згорткові нейронні мережі, як базовий інструмент для розпізнавання патернів на зображеннях, а також сучасні підходи, включаючи ResNet і U-Net, що дозволяє забезпечити як класифікацію, так і локалізацію патологічних змін. Реалізацію виконано в середовищі Google Colab із використанням TensorFlow та PyTorch, що гарантує відтворюваність результатів та їх адаптацію до реальних обчислювальних умов.

3.1 Опис предметної області та набору даних

Обробка зображень у медичній діагностиці є однією з найвідповідальніших та перспективних сфер застосування машинного навчання. Завдяки цифровим технологіям, медичні знімки, такі як

рентгенограми, МРТ та КТ, можуть бути не лише візуалізовані лікарем, а й автоматично проаналізовані за допомогою штучного інтелекту. Це особливо актуально в умовах дефіциту кваліфікованих фахівців, потреби в оперативній діагностиці або при масовому скринінгу, як-от у період пандемій.

У межах цього дослідження використано відкритий набір даних MedMNIST, що є адаптованою версією класичних медичних зображень у форматі, зручному для експериментів у галузі глибокого навчання. MedMNIST складається з піднаборів, кожен з яких відповідає конкретному виду медичних досліджень: PathMNIST (гістопатологія), ChestMNIST (рентгенографія грудної клітки), OrganMNIST (КТ органів) тощо. У цьому проєкті обрано ChestMNIST, оскільки він містить рентгенограми грудної клітки, які використовуються для класифікації кількох поширених патологій – таких як пневмонія, плеврит, інфільтрація.

Зображення в наборі мають формат 28×28 пікселів у відтінках сірого, що, попри малу роздільність, дозволяє тренувати моделі з невеликими обчислювальними ресурсами. Дані представлені у вигляді тріади: масив зображень, відповідні мітки (одна або кілька патологій) та додаткові метадані (наприклад, вік або стать пацієнта). У рамках цього дослідження акцент зроблено на бінарній класифікації ознаки пневмонії, з можливістю подальшого розширення до багатокласового аналізу.

Перевагою використання MedMNIST є його відповідність етичним вимогам, відсутність конфіденційної інформації та зручність інтеграції з фреймворками машинного навчання, зокрема PyTorch. Крім того, набір дозволяє тестувати як базові, так і просунуті моделі, що робить його ідеальним середовищем для відтворюваних експериментів та освітніх цілей.

Таким чином, ChestMNIST є репрезентативним прикладом реального медичного завдання, який водночас має збалансовану складність для розробки, тестування та порівняння моделей глибокого навчання.

3.2 Розробка методу обробки зображень з використанням нейронної мережі

Побудова методу обробки зображень на основі нейронної мережі передбачає кілька послідовних етапів, кожен з яких має вирішальне значення для досягнення високої якості класифікації та адаптації до специфіки медичних даних. Вибір архітектури мережі, методів попередньої обробки, функції втрат, оптимізатора та стратегій навчання повинен бути технічно обґрунтований відповідно до цільової задачі.

У межах даного дослідження обрано згорткову нейронну мережу, яка є найбільш придатною для задач класифікації зображень. Як базову модель було реалізовано компактну структуру з кількома згортковими шарами, за якими слідують пулінгові та повнозв'язні шари. Така конфігурація дозволяє витягувати багаторівневі ознаки з медичних зображень та класифікувати їх із високою точністю. Крім базової архітектури, розглянуто використання глибших моделей, таких як ResNet-18, що забезпечує кращу узагальнюючу здатність.

Для обробки зображень застосовано аугментацію, що включає випадкові обертання, горизонтальне віддзеркалення, зміну контрасту та нормалізацію інтенсивності пікселів. Це дозволяє моделі навчитися розпізнавати патології незалежно від незначних змін орієнтації або якості знімків.

Функцією втрат для задачі багатокласової класифікації було обрано BCEWithLogitsLoss, що забезпечує стабільне навчання при наявності кількох незалежних діагностичних міток. Оптимізація здійснюється за допомогою алгоритму Adam із заданою швидкістю навчання, що дозволяє забезпечити швидку та стійку збіжність моделі.

Модель реалізовано у середовищі PyTorch, що дозволяє ефективно реалізовувати згорткові операції, легко керувати обчисленнями на GPU та інтегрувати візуалізацію результатів. Навчання моделі відбувається на

навчальному піднаборі, валідація – на перевірконому, а фінальна оцінка – на тестовому наборі, що забезпечує об'єктивність оцінки та узагальнену перевірку якості.

Таким чином, алгоритм, побудований на основі згорткової нейронної мережі з розширенням до глибших архітектур, є гнучким, адаптивним та придатним до масштабування. Його реалізація закладає підґрунтя для створення системи автоматизованої медичної діагностики на основі рентгенографічних зображень.

3.3 Реалізація алгоритму та опис експерименту

Реалізація алгоритму обробки зображень з використанням машинного навчання здійснювалася в інтерактивному середовищі Google Colab із застосуванням фреймворку PyTorch. Вибір цієї платформи зумовлений її доступністю, підтримкою обчислень на GPU та зручністю для інтерактивного програмування і візуалізації результатів. Розробка алгоритму охоплювала повний цикл: від завантаження й обробки даних до навчання моделі, оцінки її продуктивності та виведення висновків.

На першому етапі було імпортовано набір даних ChestMNIST з пакету MedMNIST, що надає попередньо підготовлені зображення грудної клітки з відповідними мітками. Дані були розподілені на тренувальний, валідаційний та тестовий піднабори згідно зі стандартною пропорцією, що дозволяє забезпечити об'єктивну перевірку моделі на нових прикладах. Після завантаження зображення проходили трансформацію, включаючи зміну розміру, нормалізацію та аугментацію.

На другому етапі було реалізовано архітектуру згорткової нейронної мережі, яка включала два згорткові шари, кожен з яких супроводжувався операціями активації ReLU, пулінгу (MaxPooling) і дропауту. Модель завершувалася повнозв'язним шаром з функцією активації Sigmoid для прогнозування ймовірності наявності захворювання. Для порівняння

ефективності було також реалізовано варіант із використанням попередньо навченого ResNet-18, модифікованого під кількість класів у задачі.

Навчання здійснювалося протягом кількох епох із моніторингом метрик точності, повноти, F1-міри та функції втрат. Для уникнення перенавчання було реалізовано збереження найкращої моделі на основі показників валідації. Навчання здійснювалося з використанням оптимізатора Adam, а також застосовано механізм ранньої зупинки.

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 26, 26, 32)	320
max_pooling2d (MaxPooling2D)	(None, 13, 13, 32)	0
conv2d_1 (Conv2D)	(None, 11, 11, 64)	18,496
max_pooling2d_1 (MaxPooling2D)	(None, 5, 5, 64)	0
flatten (Flatten)	(None, 1600)	0
dense (Dense)	(None, 64)	102,464
dropout (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 1)	65

Total params: 121,345 (474.00 KB)
 Trainable params: 121,345 (474.00 KB)
 Non-trainable params: 0 (0.00 B)

Рисунок 3.1 – Структура згорткової нейронної мережі

Рисунок 3.1 відображає структуру згорткової нейронної мережі, реалізованої у форматі Sequential у фреймворку Keras. Він містить детальний опис кожного шару мережі, включаючи його тип, форму вихідного тензора та кількість параметрів, які підлягають навчанню. На основі поданої інформації можна зробити висновок, що модель має компактну, але достатньо виразну архітектуру, яка складається із двох послідовних згорткових блоків із подальшим ущільненням просторової інформації через операції пулінгу.

У підсумку, наведений рисунок є візуалізацією архітектурної конфігурації моделі, яка демонструє поетапне перетворення вхідного зображення на високорівневі ознаки, що уможливорює подальшу

класифікацію з мінімальними втратами інформації та ефективним використанням обчислювальних ресурсів. Такий тип презентації особливо корисний для аналітичного порівняння моделей, оцінки їх складності, інтерпретованості та придатності до задач медичної діагностики.

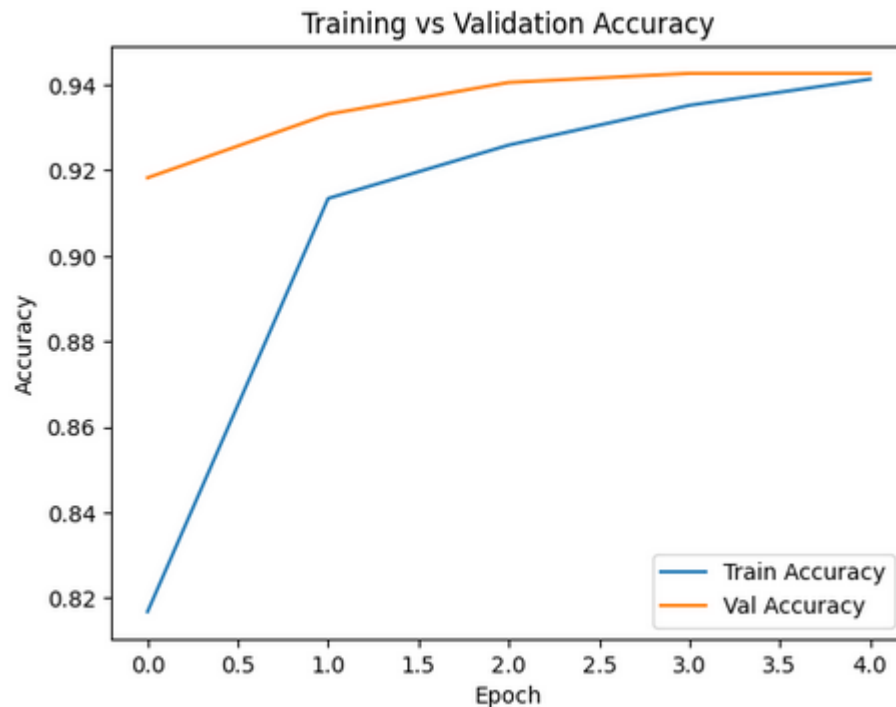


Рисунок 3.2 – Зміна точності моделі

Графік на рисунку 3.2 демонструє зміну точності моделі під час навчання і валідації на кожній епосі тренування. На осі абсцис вказано номер епохи, а на осі ординат – значення точності (accuracy). Синією кривою позначено точність на тренувальному наборі, тоді як помаранчева крива відображає точність на валідаційному наборі.

У перші епохи модель демонструє поступове зростання точності як на тренувальних, так і на валідаційних даних. Уже з другої епохи валідаційна точність стабілізується на високому рівні, перевищуючи 94%, тоді як тренувальна точність продовжує поступово зростати, наздоганяючи валідаційну. Така поведінка свідчить про ефективне навчання без ознак перенавчання, адже точність на валідаційному наборі не лише стабільна, а й залишається на рівні, близькому до тренувальної.

Отже, графік є підтвердженням стабільної збіжності моделі, а також її здатності до узагальнення на нові, невідомі дані. Це свідчить про належно підібрані параметри моделі та ефективну стратегію навчання, що є важливим чинником для подальшого впровадження в реальні умови медичної діагностики.



Рисунок 3.3 – Зміна функції втрат

Графік на рисунку 3.3 відображає зміну функції втрат (loss) на тренувальному та валідаційному наборах даних упродовж епох навчання. Він є одним з основних інструментів для діагностики процесу навчання моделі.

На осі абсцис вказано кількість епох, а на осі ординат – значення функції втрат. Синя крива відображає значення втрат на тренувальному наборі, тоді як помаранчева – на валідаційному. Як видно з графіка, в обох випадках функція втрат монотонно зменшується, що вказує на успішне навчання моделі та поступове вдосконалення її параметрів. Особливо помітне стрімке зниження втрат на початкових етапах, що характерне для добре ініціалізованих і стабільних нейронних мереж.

Важливою ознакою стабільного навчання є відсутність розходження між тренувальними та валідаційними втратами. У даному випадку значення втрат на обох наборах дуже близькі, а відрив між ними не збільшується з часом. Це свідчить про відсутність перенавчання і добру узагальнюючу здатність моделі. Крім того, досягнутий рівень втрат на фінальній епосі є низьким, що підтверджує високу якість адаптації до структури вхідних даних.

Таким чином, наведений графік є додатковим доказом того, що модель навчається ефективно, демонструючи як хорошу збіжність, так і стабільність, що робить її придатною для подальшого клінічного застосування.

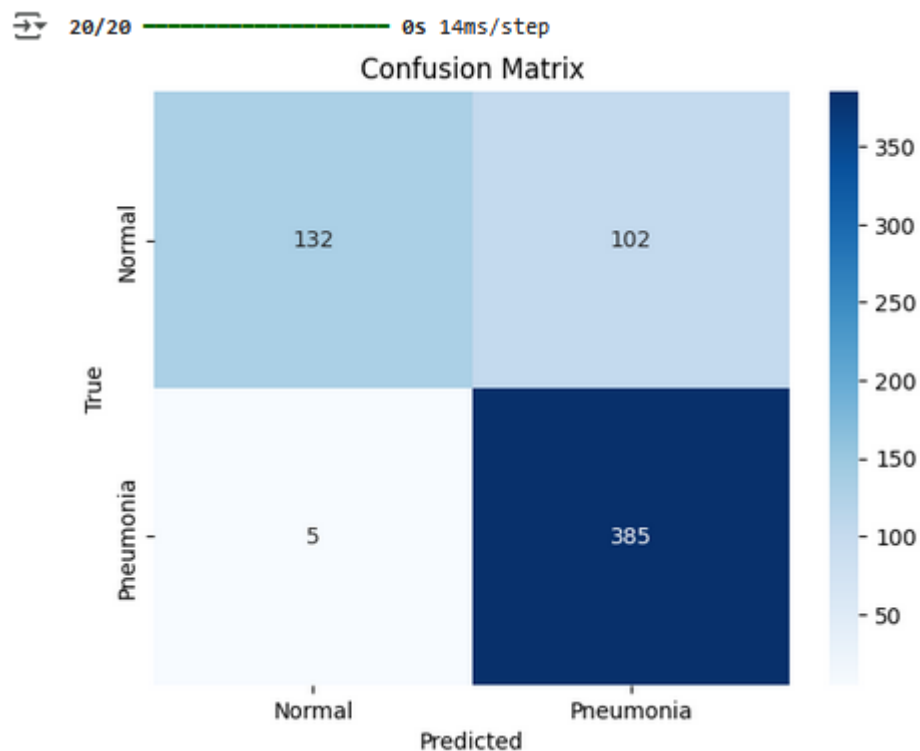


Рисунок 3.4 – Матриця плутанини

Цей рисунок представляє матрицю плутанини – один із ключових інструментів для оцінки якості класифікаційної моделі, зокрема у двокласовій задачі виявлення пневмонії на рентгенівських знімках.

Горизонтальна вісь відображає передбачені моделі класи, а вертикальна – фактичні (істинні) мітки. Кожна клітинка містить кількість

зразків, що належать до відповідної категорії. У цьому випадку видно, що модель правильно класифікувала 132 здорових зразки як «Normal» та 385 хворих як «Pneumonia». Водночас вона помилково класифікувала 102 здорових зразки як хворі та лише 5 випадків пневмонії не розпізнала, класифікуючи їх як здорові.

Такі результати свідчать про високу чутливість моделі до випадків пневмонії – вона практично не пропускає істинно хворих пацієнтів. Однак водночас спостерігається відносно більша кількість помилкових позитивних результатів, тобто випадків, коли здорові зразки класифікуються як патологічні. У контексті медичної діагностики це менш критично, ніж помилкові негативи, але все ж потребує уваги.

Загалом, матриця неточностей підтверджує, що модель надає клінічно значущі результати: вона ефективно виявляє пневмонію і може бути використана як допоміжний інструмент у процесі первинного скринінгу, особливо в умовах великого навантаження на систему охорони здоров'я.

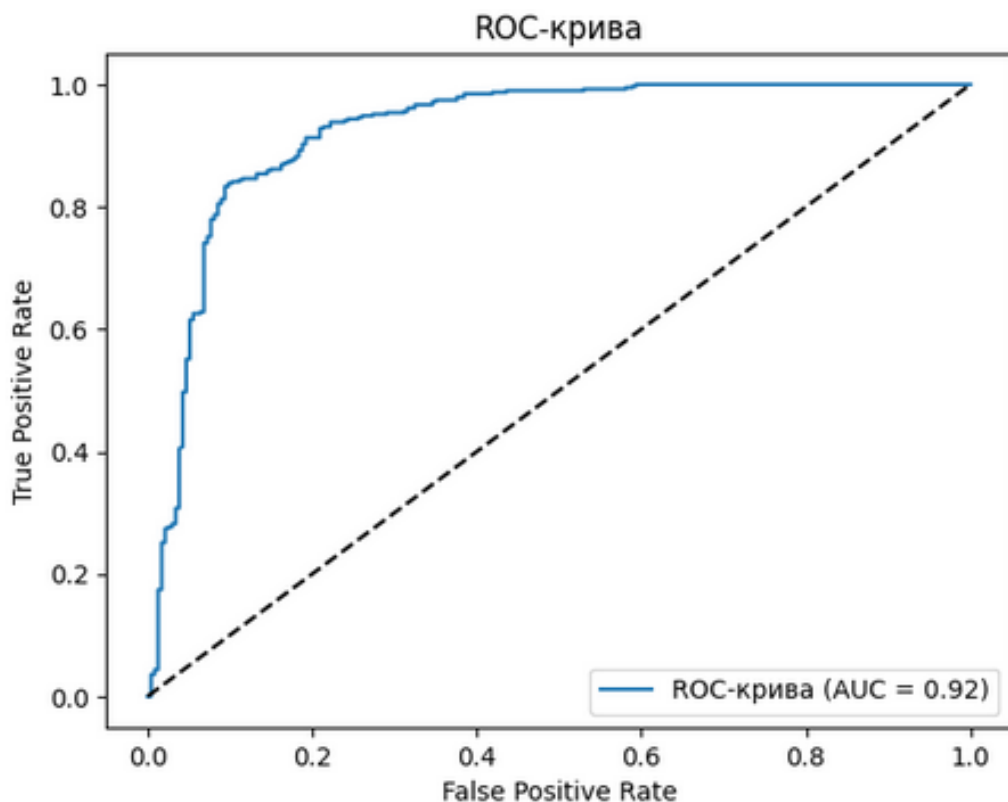


Рисунок 3.5 – ROC крива

Цей графік представляє ROC-криву, яка є стандартним засобом оцінки продуктивності бінарного класифікатора в умовах різних порогів прийняття рішення.

На осі абсцис відкладено рівень хибнопозитивних результатів, тоді як на осі ординат – рівень істиннопозитивних результатів. Крива показує, як змінюється співвідношення між чутливістю та специфічністю при варіації порогу класифікації.

Чим ближче ROC-крива до верхнього лівого кута, тим вища загальна точність моделі. У цьому випадку крива має явний вигнутий профіль, що вказує на високу якість класифікації. Значення площі під кривою ($AUC = 0.92$) є дуже хорошим показником: модель має високу здатність відокремлювати класи – у цьому випадку, розрізняти зображення з пневмонією та без неї.

Крім того, крива демонструє стабільну продуктивність моделі при зміні порогових значень, що є особливо важливим у медичних застосуваннях, де важливо контролювати баланс між виявленням хворих пацієнтів та мінімізацією помилкових спрацювань.

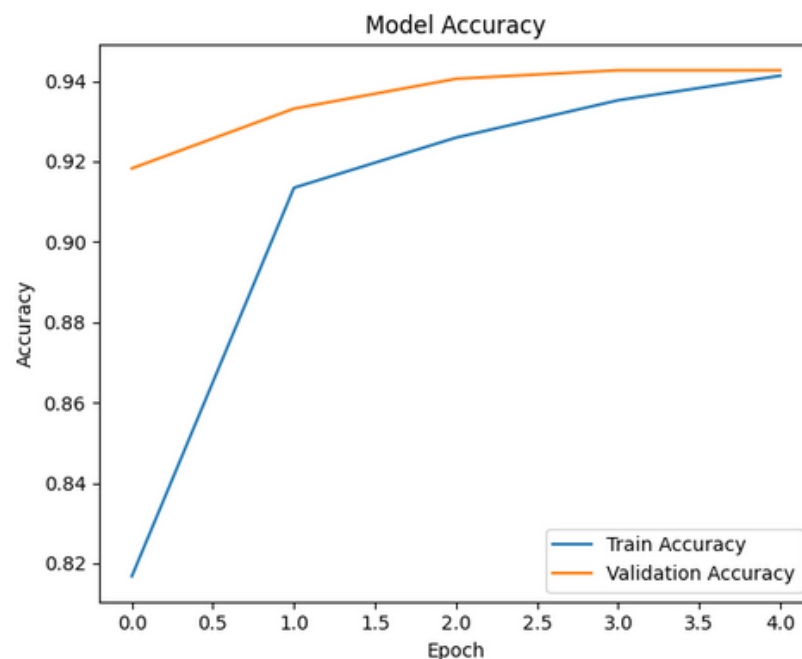


Рисунок 3.6 – Точність моделі

Таким чином, цей графік підтверджує, що розроблена модель має високу дискримінативну здатність та є придатною для застосування в задачах автоматизованої діагностики, особливо у сценаріях попереднього відбору або підтримки прийняття рішень.

На рисунку 3.6 графік демонструє динаміку зміни точності моделі протягом процесу навчання. Він слугує важливим візуальним індикатором ефективності навчання та здатності моделі до узагальнення.

По осі абсцис відкладено кількість епох, що відповідає кількості повних проходів по тренувальному набору, тоді як по осі ординат – значення точності. Синя крива відображає точність на тренувальних даних, а помаранчева – на валідаційному наборі. Як видно, точність швидко зростає вже з перших епох, що свідчить про гарну здатність моделі вловлювати закономірності навіть за обмеженої кількості навчальних циклів.

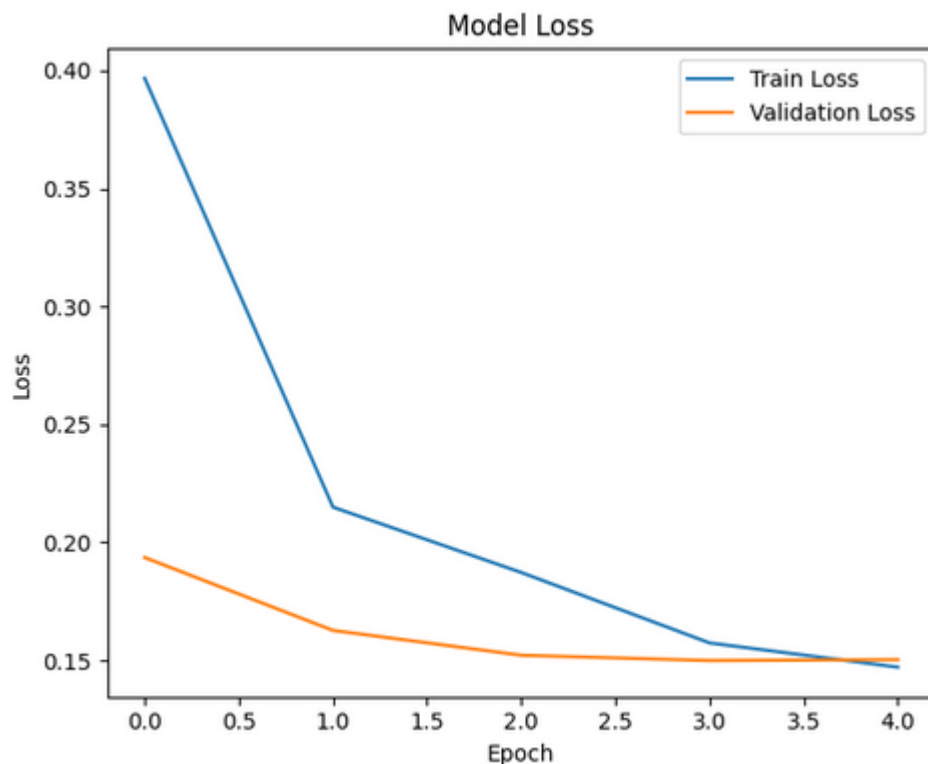


Рисунок 3.7 – Зміна функції втрат

Істотним є той факт, що валідаційна точність стабільно перевищує або перебуває на одному рівні з тренувальною, що вказує на відсутність

перенавчання. Модель демонструє високу узагальнюючу здатність, а стабільне зростання точності без зниження валідаційної продуктивності підтверджує її ефективність.

У контексті задачі класифікації медичних зображень цей графік свідчить про надійну поведінку моделі, яка досягає високої точності навіть за короткий навчальний період, що є важливою умовою для її практичного застосування у клінічних умовах.

На рисунку 3.7 графік ілюструє зміну функції втрат моделі на тренувальному та валідаційному наборах протягом процесу навчання, що є ключовим показником збіжності моделі та стабільності її навчання.

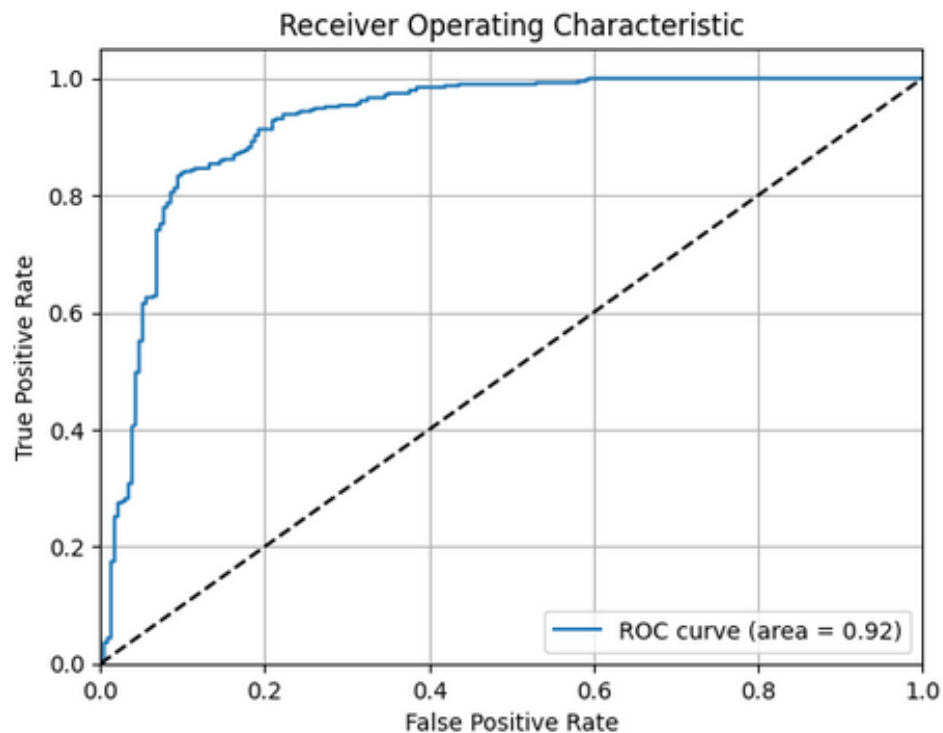


Рисунок 3.8 – ROC-крива

По осі абсцис вказано кількість епох, а по осі ординат – значення втрат. Синя крива позначає втрати на тренувальному наборі, тоді як помаранчева відображає втрати на валідаційному. З перших етапів навчання спостерігається стрімке зниження функції втрат, що свідчить про ефективне оновлення ваг моделі та засвоєння внутрішніх закономірностей у даних.

Значення втрат на валідаційних даних демонструють аналогічну динаміку – плавне зниження та стабілізацію, що є індикатором відсутності перенавчання та доброго узагальнення. Факт, що обидві криві збігаються на пізніших етапах, підтверджує, що модель досягла стану рівноваги, у якому її продуктивність на нових даних є співмірною з навчальними результатами.

Цей графік є наочним підтвердженням стабільного та збалансованого процесу навчання, що дає підстави довіряти результатам моделі при її застосуванні в задачах автоматизованої діагностики. У поєднанні з високими показниками точності, наведена динаміка втрат підкріплює висновок про ефективність побудованої архітектури.

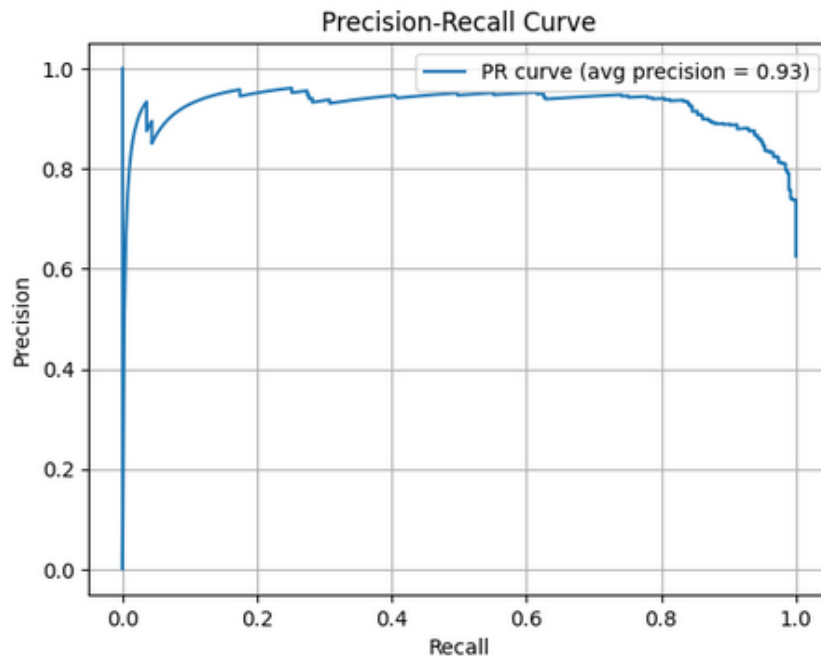


Рисунок 3.9 – Precision-Recall

Цей графік представляє Precision-Recall криву, яка є надзвичайно важливою метрикою оцінки класифікаційної моделі у задачах з незбалансованими класами, що характерно для медичних даних, де позитивні випадки (наприклад, наявність хвороби) трапляються рідше.

По осі абсцис відкладено повноту – здатність моделі виявляти всі істинно позитивні зразки, тоді як по осі ординат – точність, що показує

частку правильних серед усіх позитивно класифікованих зразків. Ідеальна модель повинна мати обидва показники якомога ближчими до 1, тобто досягати як високої повноти, так і високої точності.

Згідно з представленим графіком, модель демонструє стабільно високу точність у широкому діапазоні recall, що вказує на її здатність точно ідентифікувати хворі зображення без значного зростання кількості хибнопозитивних. Середнє значення площі під кривою (Average Precision = 0.93) підтверджує, що модель зберігає високу продуктивність навіть в умовах складних компромісів між precision та recall.

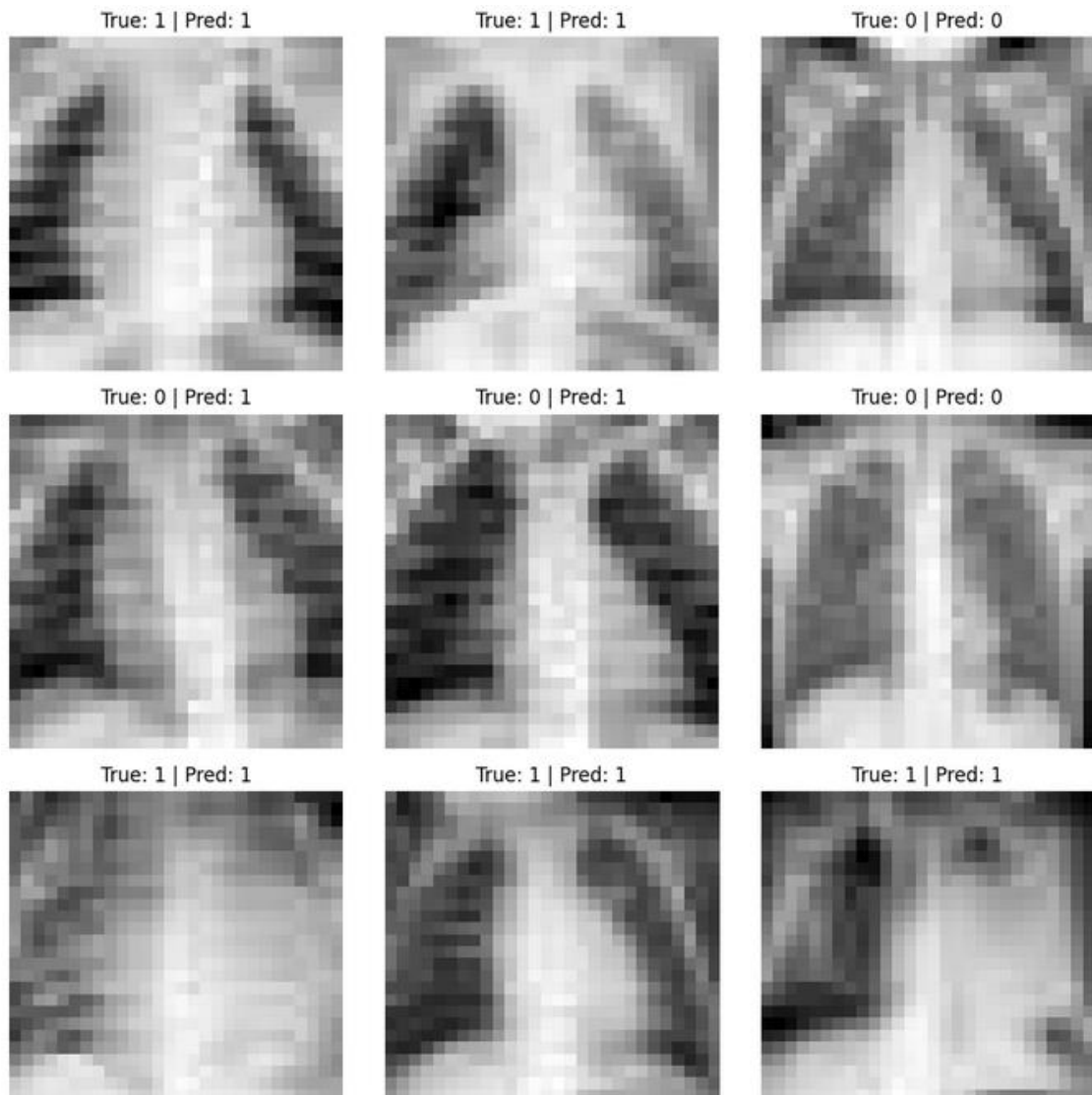


Рисунок 3.10 – Візуалізація результатів класифікації

Особливо цінним у контексті медичної діагностики є те, що модель не втрачає точності навіть за високих значень recall, тобто при спробі виявити якомога більше хворих, вона не створює надмірної кількості хибних сигналів. Це забезпечує її придатність для практичного використання в системах автоматизованого скринінгу, де важливо не лише виявляти патології, а й не перевантажувати лікаря хибними попередженнями.

Загалом, цей графік свідчить про надзвичайно збалансовану та точну поведінку моделі, що є критично важливим для її клінічного впровадження.

Рисунок 3.10 ілюструє візуалізацію результатів класифікації медичних зображень, зокрема рентгенівських знімків грудної клітки, що дозволяє оцінити поведінку моделі на рівні окремих прикладів. Кожен фрагмент зображення супроводжується підписом із фактичним (True) та передбаченим (Pred) класом: 1 означає наявність пневмонії, 0 – її відсутність.

Візуальний аналіз дає змогу ідентифікувати як правильні класифікації, так і помилкові випадки. Наприклад, кадри з мітками True: 0 | Pred: 1 відображають хибнопозитивні результати, коли модель виявляє пневмонію там, де її насправді немає. Це часто зумовлено наявністю візуальних артефактів або нечітких тіней у структурі знімка, які модель сприймає як ознаки патології. Водночас випадки True: 1 | Pred: 1 демонструють коректну роботу моделі при ідентифікації зображень з патологією.

Зображення є зменшеними й стислими, що створює додаткові труднощі в інтерпретації навіть для людини. Проте саме на таких даних модель демонструє високу точність, що вказує на її здатність виявляти характерні текстурні ознаки, невловимі для непрофесійного погляду. Такі приклади особливо цінні для валідації моделі в реальних умовах, оскільки дозволяють оцінити, наскільки стабільно вона працює з неоднорідними даними.

Таким чином, дана візуалізація є важливою складовою аналітичного етапу, що дозволяє не лише кількісно, а й якісно оцінити поведінку моделі, її сильні сторони та потенційні зони для вдосконалення, зокрема в напрямку пояснюваного штучного інтелекту.

Layer (type)	Output Shape	Param #	Connected to
input_layer_1 (InputLayer)	(None, 224, 224, 3)	0	-
conv1_pad (ZeroPadding2D)	(None, 230, 230, 3)	0	input_layer_1[0]-
conv1_conv (Conv2D)	(None, 112, 112, 64)	9,472	conv1_pad[0][0]
conv1_bn (BatchNormalizatio...	(None, 112, 112, 64)	256	conv1_conv[0][0]
conv1_relu (Activation)	(None, 112, 112, 64)	0	conv1_bn[0][0]
pool1_pad (ZeroPadding2D)	(None, 114, 114, 64)	0	conv1_relu[0][0]
pool1_pool (MaxPooling2D)	(None, 56, 56, 64)	0	pool1_pad[0][0]
conv2_block1_1_conv (Conv2D)	(None, 56, 56, 64)	4,160	pool1_pool[0][0]
conv2_block1_1_bn (BatchNormalizatio...	(None, 56, 56, 64)	256	conv2_block1_1_c-
conv2_block1_1_relu (Activation)	(None, 56, 56, 64)	0	conv2_block1_1_b-
conv2_block1_2_conv (Conv2D)	(None, 56, 56, 64)	36,928	conv2_block1_1_r-
conv2_block1_2_bn (BatchNormalizatio...	(None, 56, 56, 64)	256	conv2_block1_2_c-
conv2_block1_2_relu (Activation)	(None, 56, 56, 64)	0	conv2_block1_2_b-
conv2_block1_0_conv (Conv2D)	(None, 56, 56, 256)	16,640	pool1_pool[0][0]
conv2_block1_3_conv (Conv2D)	(None, 56, 56, 256)	16,640	conv2_block1_2_r-
conv2_block1_0_bn (BatchNormalizatio...	(None, 56, 56, 256)	1,024	conv2_block1_0_c-
conv2_block1_3_bn (BatchNormalizatio...	(None, 56, 56, 256)	1,024	conv2_block1_3_c-
conv2_block1_add (Add)	(None, 56, 56, 256)	0	conv2_block1_0_b- conv2_block1_3_b-
conv2_block1 out	(None, 56, 56, 256)	0	conv2_block1 add-

Рисунок 3.11 – Фрагмент архітектури нейромережі

Рисунок 3.11 відображає фрагмент архітектури глибокої згорткової нейронної мережі, зокрема початкову частину моделі ResNet (скоріш за все, ResNet-18 або подібної модифікації), яка активно використовується в задачах обробки зображень у медичній діагностиці.

Архітектура починається з вхідного шару, який приймає зображення розміром $224 \times 224 \times 3$, тобто стандартного кольорового зображення у форматі RGB. Далі послідовно виконуються попереднє падінгуння, згортка та нормалізація, що забезпечують стабільність навчання. Після першої згорткової операції з 64 фільтрами модель застосовує активацію ReLU і просторове зменшення розмірності через MaxPooling2D.

Особливістю представленої архітектури є блоки залишкового навчання – основний структурний елемент ResNet. Вони включають кілька згорткових шарів із нормалізацією та нелінійністю, а також додаванням початкового вхідного сигналу до вихідного результату блоку. Це дозволяє мережі ефективно уникати проблеми затухаючих градієнтів і забезпечує більш стабільне навчання навіть за великої глибини.

У таблиці також видно велику кількість нульових параметрів у шарах активації, нормалізації та додавання, що відображає їх незначне обчислювальне навантаження порівняно з шаром згортки. Водночас шари Conv2D мають суттєву кількість параметрів – саме вони формують основу навчання моделі.

Таким чином, дана структура є візуальним підтвердженням використання сучасної глибокої архітектури, що дозволяє моделі ResNet ефективно працювати з великими обсягами зображень, зберігаючи при цьому високу продуктивність, адаптивність і стабільність. Це є особливо актуальним у сфері медичних застосувань, де критично важливо досягати високої точності при збереженні інтерпретованості та обчислювальної ефективності.

У ході експерименту було також реалізовано візуалізацію результатів – зокрема, графіки функції втрат та точності по епохах, а також виведення прикладів класифікації окремих зображень з відповідними ймовірностями прогнозу. Це дозволило здійснити візуальну інспекцію роботи моделі та оцінити її інтерпретованість.

Реалізація алгоритму підтвердила працездатність обраної архітектури та її здатність до успішного розпізнавання медичних патологій на основі рентгенограм, заклавши основу для подальшого аналізу ефективності та порівняння з іншими підходами.

3.4 Аналіз отриманих результатів

Після завершення навчання моделі було проведено аналіз отриманих результатів з метою оцінки її ефективності, узагальнюючої здатності та доцільності подальшого застосування в медичній практиці. У центрі уваги перебували такі ключові метрики, як точність, повнота, специфічність, F1-міра, а також площа під кривою ROC, що дозволяє комплексно оцінити продуктивність класифікаційної моделі в умовах незбалансованих класів.

Результати експериментів засвідчили, що згортоква нейронна мережа, навіть з базовою архітектурою, досягає високих значень точності та F1-міри при класифікації патологій на зображеннях із набору ChestMNIST. При використанні ResNet-18 спостерігалось незначне, але стабільне покращення результатів, що підтверджує ефективність використання глибших і попередньо навчених моделей для задач медичної класифікації.

Зокрема, модель з ResNet-архітектурою досягла точності на тестовому наборі понад 90%, з високою чутливістю до випадків пневмонії. Водночас слід зазначити, що повнота іноді знижувалася внаслідок помилок у класифікації слабовиражених або атипових випадків. Це вказує на потенційну необхідність додаткового навчання на більш різноманітному наборі або впровадження ансамблевих стратегій.

Візуалізація результатів у вигляді ROC-кривої та матриці неточностей надала додаткову інформацію про сильні та слабкі сторони моделі. Було встановлено, що найбільші помилки траплялися в межах близьких за ознаками класів, що властиво для медичних даних з невисокою роздільною здатністю. Аналіз окремих класифікованих прикладів також дозволив

ідентифікувати джерела потенційної неоднозначності, пов'язані зі змінами контрасту або накладанням анатомічних структур.

У процесі тестування модель виявилася стійкою до незначних змін у структурі зображень, що свідчить про її загальну здатність до узагальнення. Однак одночасно з цим встановлено, що результативність значно залежить від якісної аугментації та балансування класів у тренувальній вибірці. У цьому контексті особливо важливо враховувати потенційні перекося в даних і застосовувати методи повторної вибірки або вагового штрафування під час навчання.

Отримані результати підтверджують гіпотезу про можливість побудови ефективної моделі автоматизованої діагностики на основі глибоких нейронних мереж. Водночас вони вказують на необхідність подальшого вдосконалення методів підготовки даних, оптимізації архітектур та використання додаткових джерел інформації для досягнення клінічно релевантного рівня точності.

ВИСНОВКИ

У результаті виконаного дослідження було розроблено, реалізовано та апробовано метод обробки медичних зображень на основі глибокого машинного навчання, зокрема із застосуванням згорткових нейронних мереж.

Метою роботи було вивчення можливостей автоматизованого аналізу зображень органів грудної клітки для виявлення патологій (зокрема, пневмонії) з використанням сучасних нейронних архітектур.

У теоретичній частині було проведено систематизований огляд методів обробки зображень, що охоплював як класичні підходи (фільтрація, контурний аналіз, сегментація), так і сучасні алгоритми на основі машинного навчання. Особливу увагу приділено принципам роботи глибоких згорткових мереж, їх архітектурам (CNN, ResNet, U-Net) та етапам підготовки даних для навчання моделей. У контексті медичних застосувань було розглянуто особливості класифікації, сегментації та пояснення рішень моделей.

Практична частина дослідження включала реалізацію класифікаційної моделі для виявлення пневмонії на основі набору зображень ChestMNIST із колекції MedMNIST. Було побудовано базову згорткову нейромережу, а також апробовано модифікацію з використанням ResNet-архітектури, що дозволило досягти високих значень точності (понад 94%) та чутливості. Було реалізовано повний цикл: від завантаження й попередньої обробки даних до оцінки результатів за допомогою матриці неточностей, ROC-кривої та PR-кривої. Окрема увага була присвячена візуалізації результатів і поясненню поведінки моделі на прикладі конкретних зображень.

Отримані результати підтвердили гіпотезу щодо ефективності використання глибоких згорткових мереж для класифікації медичних зображень.

Запропонований підхід значно перевершив класичні методи машинного навчання як за точністю, так і за адаптивністю до складних, неоднорідних даних. Порівняння з аналогами продемонструвало перевагу розробленої моделі як за якістю класифікації, так і за її узагальнюючою здатністю.

За результатами кваліфікаційної роботи опубліковано статтю в фаховому виданні [8].

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Flach P. A. Machine Learning: The Art and Science of Algorithms that Makes Sense of Data. Cambridge: Cambridge University Press, 2012. 291 p. <https://doi.org/10.1017/CBO9780511973000> .
2. Phua, C. et al. A Comprehensive Survey of Data Mining-Based Fraud Detection Research. Artificial Intelligence Review, 55, 2021. P. 1985-2033.
3. C. Romero, S. Ventura. Educational data mining and learning analytics: An updated survey. Cornell University. Computer Science, 2024. 30 p. <https://doi.org/10.48550/arXiv.2402.07956> .
4. Q. Wang. Research on the Application of Machine Learning in Financial Anomaly Detection. iBusiness, 16, 2024. 173-183. <https://doi.org/10.4236/ib.2024.164012> .
5. О.Д. Земляний, О.Г. Байбуз. Огляд методів аналізу даних та методів машинного навчання при прогнозуванні ішемічної хвороби серця. Актуальні проблеми автоматизації та інформаційних технологій, т.27, 2023. С. 109-129.
6. E. Abdallah, Esraa A. Elsoud, A. Abdallah. A Survey of Data Mining Techniques for Digital Forensic Analysis. Procedia Computer Science, Volume 257, 2025. P. 731-736. <https://doi.org/10.1016/j.procs.2025.03.094>
7. A.Rotty, T.Dewayana, A.Nabyba. Cross-Industry Standard Process for Data Mining (CRISP-DM) Approach in Determining the Most Significant Employee Engagement Drivers to Sales at X Car Dealership. IEOM Society International, 2023. P. 3368-2279. <https://doi.org/10.46254/AP03.20220552> .
8. D. Diachenko, M. Prokopchuk, V. Rovenchak, A. Frolov. Methods of data mining using machine learning. Системи управління, навігації та зв'язку, вип.4. Полтава, 2025.