

**«Система підтримки прийняття рішень при діагностуванні за
допомогою МРТ-знімків»**

Шифр: «Інтерпретація МРТ»

ЗМІСТ

Вступ.....	3
1 Дослідження предметної області.....	5
1.1 Передпроектне обстеження.....	5
1.1.1 Опис та аналіз предметної області	5
1.1.2 Огляд та аналіз існуючих методів обробки медичних даних	6
1.2 Актуальність та мета розробки.....	9
2 Теоретичні аспекти нейронних мереж.....	10
2.1 Обробка зображень за допомогою нейронних мереж.....	10
2.2 Опис шарів згорткової нейронної мережі	11
2.3 Оцінка ефективності моделі.....	12
3 Моделювання системи.....	17
3.1 Опис та попередня обробка даних	17
3.2 Опис методу і проектування моделей.....	19
3.3 Тестування моделей і аналіз результатів.....	23
3.4 Практичне застосування і перспективи розвитку.....	27
Висновки і перспективи розвитку	29
Перелік джерел посилання	31
Додаток А Навчання і тестування моделей.....	34

ВСТУП

В медичній галузі дуже велике значення має використання правильних методів діагностики, які давали б найбільш повне уявлення про стан здоров'я людини. При цьому потрібно обирати діагностичні засоби, які могли б забезпечувати не тільки точність показань, а й безпеку пацієнта. Методи обстеження прийнято розділяти на інвазивні (з проникненням всередину організму або підшкірного покриву) і неінвазивні, такі як комп'ютерна томографія (КТ), магнітно-резонансна томографія (МРТ), ультразвукове дослідження (УЗД), рентгенографія тощо. Інтерпретація результатів вищезазначених методів є досить складною процедурою, що потребує знання висококваліфікованих експертів (лікарів-діагностів).

Одним із найкращих сучасних засобів для визначення стану організму людини є магнітно-резонансна томографія. Цей метод спрямований на те, щоб на ранніх стадіях відстежувати новоутворення, зміни структури органів та інші потенційні загрози для здоров'я. Колінний суглоб (КС) є одним із найбільших і найскладніших суглобів людського організму. Згідно зі статистикою, приблизно третина населення старше тридцяти років хоча б раз у житті зазнали на травму колінного суглобу. Обробка результатів, що отримані під час діагностики захворювань колінного суглоба, методами штучного інтелекту є актуальною проблемою, вирішення якої дозволить значно оптимізувати процес інтерпретації магнітно-резонансної томографії фахівцями.

В даній роботі запропоновано та програмно реалізовано інтелектуальну систему підтримки прийняття рішень для діагностики патологій колінного суглоба. Буде представлено метод діагностики, заснований на технології глибокого навчання, а саме – згорткових нейронних мережах зі специфічним розподілом зрізів магнітно-резонансних зображень. Метод передбачає поділ кожного сканованого зображення МРТ на декілька частин, що уможливить цілковите дослідження пошкодженого органу. Наукова новизна даної роботи полягає в тому, що запропонований підхід ще не був використаний для

діагностики 3D медичних зображень раніше. На отриманих вибірках буде натреновано декілька нейронних мереж для кожної площини та об'єднано за допомогою використання середньозваженого ансамблю класифікаторів.

Інтерфейс системи підтримки прийняття рішень буде запроектовано за допомогою Experience Design (UX) / User Interface Design (UI) технологій.

Додаток має надавати можливість швидко отримати повну індивідуальну картину необхідних площин (коронарних, сагітальних, аксіальних) та зрізів для конкретного пацієнта, генерувати анімацію для перегляду динаміки (анімовані gif-файли). Система має включати інтелектуальну складову, а саме: функціонал інтерпретації МРТ-зображень з імовірнісними показниками наявності патології, що реалізовано за допомогою використання вищезгаданих технологій штучного інтелекту. Інтелектуальний додаток допоможе швидше та більш точно визначити характер патології на зображенні, знайти неявні ознаки порушення, виявити малий розмір новоутворень, що не можуть бути виявлені вручну лікарем-діагностом, а також запобігти випадковій помилці в рутинній великій кількості зображень.

1 ДОСЛІДЖЕННЯ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Передпроектне обстеження

1.1.1 Опис та аналіз предметної області

Метод магнітно-резонансної томографії був сформульований у 1973 році професором університету Іллінойсу, Полом Лотербуром. Тридцять років по тому разом із британським ученим Пітером Менсфілдом він отримав Нобелівську премію в галузі фізіології та медицини за винахід магнітно-резонансного томографа [1]. Дослідження може бути проведено як на м'яких тканинах, так і кісткових структурах. В основі методу лежить явище ядерно-магнітного резонансу. Це властивість протонів водню специфічним чином реагувати на індуковане апаратом поле. Заряджені частинки змінюють траєкторію руху і повертаються в первісний стан. Томограф збирає сплески енергії, комп'ютерна програма перетворює дані у зображення. Таким чином, в процесі томографії отримують серії монохромних знімків анатомічної ділянки в аксіальній, сагітальній і фронтальній площинах, що приведено на рисунку 1.1 [2].

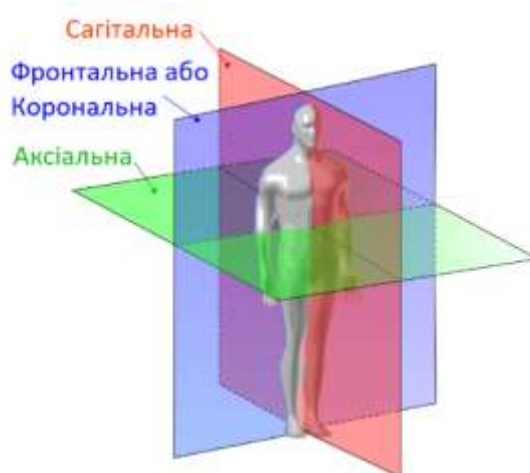


Рисунок 1.1 – Анатомічні площини

Результат дослідження формується у вигляді медичного зображення, наприклад, у форматі DICOM (Digital Imaging and Communications in Medicine), NIfTI (Neural Imaging Technology Initiative in Informatics), Analyze, Minc1 (Medical Imaging Network Common Data Format), Minc2 або іншого відповідного формату. Файли в будь-якому з цих форматів можна легко конвертувати в розширення .pru, тобто стандартний двійковий формат файлів, який є універсальним рішенням для подальшої обробки даних методами машинного навчання.

Під час вивчення зображень лікар бачить контури різних елементів зчленування. Розшифровка МРТ-знімка колінного суглоба передбачає оцінку сканів у порівнянні з нормою. МР-томограми називають зрізами, оскільки зображення роблять пошарово, з кроком 1-2 мм. У процесі розшифровки фото звертають увагу на якість візуалізації елементів, чіткість контурів, присутність зон зміненого МР-сигналу в різних режимах сканування. Сучасні апаратні можливості технології МРТ забезпечують візуалізацію, необхідну для отримання зображень з високою роздільною здатністю зображення для оцінки менісків, зв'язок і сухожилів [3], [4].

1.1.2 Огляд та аналіз існуючих методів обробки медичних даних

Ідея використання принципу зорового процесу в людському мозку для виявлення об'єктів була запропонована Фукушімою в 1980-х роках [5], і вдосконалена Яном ЛеКунном [6]. Завдяки їхній пропозиції, комп'ютерні моделі можуть виділяти ознаки об'єкта і намагатися зіставити їх з ознаками невідомих об'єктів.

Методи штучного інтелекту мають широкий спектр застосувань, включаючи інженерію, будівництво та медицину. Медичні застосунки є основною сферою інтересу даної роботи. Автори статті «Deep Learning for Health Informatics» [7] зосереджуються на ключових застосуваннях глибокого навчання

в галузях трансляційної біоінформатики, медичної візуалізації, проникаючого зондування, медичної інформатики та охорони здоров'я.

Інший продемонстрований підхід полягає в класифікації уражень шкіри за допомогою згорткової нейронної мережі (CNN), яка навчається безпосередньо на зображеннях, використовуючи лише пікселі та мітки захворювань як вхідні дані [8]. Специфічну комп'ютерну діагностику (CAdE), а саме виявлення торако-абдомінальних лімфатичних вузлів та інтерстиціальних захворювань легень, що було описано в роботі [9], також реалізовано з використанням згорткових нейронних мереж. У статті [10] зроблено огляд існуючих методів обробки медичних зображень, окремо виділено методи CNN.

На відміну від рентгенографії, магнітно-резонансна томографія генерує проекцію органу у 3D-вимірі, тому найчастіше для роботи з подібними зображеннями використовують 3D згорткові нейронні мережі [11], тренування яких потребує набагато більших обчислювальних потужностей.

Часто технологія трансферного навчання, що базується на попередньо навчених нейронних мережах, використовується для підвищення точності моделі. Набір даних МРТ-зображень, що буде досліджено у даній роботі, був використаний в роботі [12]. У статті досліджено вплив структурних варіацій, аугментації (доповнення) даних та різних реалізацій трансферного навчання на продуктивність глибокої нейронної мережі в задачі класифікації МРТ колінного суглоба. CNN на основі AlexNet з використанням шарового заморожування для пошуку аномалій колінного суглобу досягла результату AUC-метрики 0.913, що є кращим показником у порівнянні з аналогічним методом без заморожування шарів – 0.896. Класифікатор на основі ResNet18 показав AUC-метрику, що дорівнює відповідно 0.843. Класифікатор на основі VGG16 досяг AUC, рівну 0.728.

Система виявлення ураження хряща на основі глибокого навчання була розроблена з використанням сегментації та класифікації CNN. У роботі [13] за допомогою методу глибокого навчання проаналізовано набори даних

«T2-weighted fast» МРТ колінного суглоба у 175 пацієнтів з різними болями в коліні.

У роботі [14] було представлено результати наукового огляду робіт виявлення травм колінного суглоба (передньої хрестоподібної зв'язки, меніска та хряща) з використанням глибокого навчання.

Набір даних, що буде розглянуто у даній роботі, було використано у дослідженні [3] з метою побудови глибокої нейронної мережі MRNet для діагностики колінного суглобу. Щоб врахувати змінну шкалу інтенсивності пікселів у серії МРТ, до зображень було застосовано алгоритм стандартизації інтенсивності на основі гістограми. Основним структурним елементом даної системи прогнозування є згортована нейронна мережа, яка аналізує 3-вимірну серію МРТ й обчислює ймовірність захворювання. Вхідні дані для MRNet мають розмірність $S \times 3 \times 256 \times 256$, де S – кількість зображень у серії МРТ. Спочатку кожен фрагмент двовимірного МРТ-зображення пропускали через функцію вилучення ознак на основі AlexNet, щоб отримати тензор $S \times 256 \times 7 \times 7$, що містить ознаки для кожного фрагмента. У дослідженні зазначено, що кількість зрізів S варіюється від пацієнта до пацієнта, тому неможливо навчати мережу МРТ партіями. Запропонований підхід передбачає встановлення параметру розміру батча (англ. batch), що дорівнює 1. Проблемою даного підходу є те, що зазвичай подібне значення параметру розміру батча призводить до флуктуацій функції втрат.

Складність глибокого навчання для медичної візуалізації полягає в обробці великих наборів даних, що містять приклади найрізноманітніших патологій, які можуть траплятися під час певного візуального обстеження. Завданням класифікації можна визначити відмінність «нормального значення» від «аномального», оскільки, якщо модель може вивчити діапазон норми для даного набору даних МРТ колінного суглоба, то теоретично будь-яка аномалія, незалежно від того, наскільки рідкісною вона є, може бути виявлена за допомогою моделі [3].

1.2 Актуальність та мета розробки

Медицина сприяє розвитку науки та технологій, вимагаючи новітні та ефективні методи діагностики, лікування та запобігання хворобам. Останнім часом значною мірою на її розвиток впливає впровадження передових досягнень інформаційних технологій, оскільки використання науково-технічного та виробничого потенціалу для удосконалення і розробки інноваційних методів і засобів медичного призначення, є стратегічно важливою задачею в галузі біомедичної інженерії та інтелектуальних технологій для формування та підтримки діагностичних рішень [16].

Через появу методів і технічних можливостей інвазивної діагностики, таких як МРТ, КТ, ультразвукова діагностика, рентгенографія, виникла необхідність обробки великої кількості даних. Задля автоматизації та оптимізації процесів діагностування впроваджують системи підтримки прийняття рішень, які є інтелектуальними помічниками в інтерпретації та встановленні діагнозу.

За статистикою, кожен третій житель планети страждає від патологій та травм колінного суглобу, що може призводити до втрати рухливості, працездатності та навіть інвалідності. Діагностика захворювань КС методами штучного інтелекту є актуальною проблемою, вирішення якої дозволить значно оптимізувати процес інтерпретації магнітно-резонансної томографії фахівцями.

Метою даної роботи є проектування і створення системи підтримки прийняття рішень для діагностування патологій на МРТ-зображеннях колінного суглобу засобами штучних нейронних мереж. Додаток прискорить діагностику МРТ-знімків кваліфікованими радіологами і може запобігти випадковій помилці. Він також може бути використаний у навчальних цілях студентами-медиками.

Новизна даної роботи полягає в тому, що запропонований підхід не був використаний для діагностики 3D медичних зображень раніше. Запропоновано специфічний поділ зрізів площин, через який натреновано декілька нейронних мереж для кожної площини та об'єднано за допомогою використання середньозваженого ансамблю класифікаторів.

2 ТЕОРЕТИЧНІ АСПЕКТИ НЕЙРОННИХ МЕРЕЖ

2.1 Обробка зображень за допомогою нейронних мереж

Згорткова нейронна мережа, також відома як CNN, є типом нейронних мереж, який спеціалізується на обробці даних з сітчастою структурою, таких як зображення. Цифрове зображення – це двійкове представлення візуальних даних, може бути представлене у вигляді тривимірного масиву цілих чисел. Зазвичай, перший та другий вимір масиву представляють ширину та висоту пікселів зображення відповідно, тоді як третій вимір містить інформацію про глибину кольору (рисунок 2.1). На зображенні дані кольору нормовано у шкалу сірого, таким чином, що білий представлений 0, а чорний – 256.

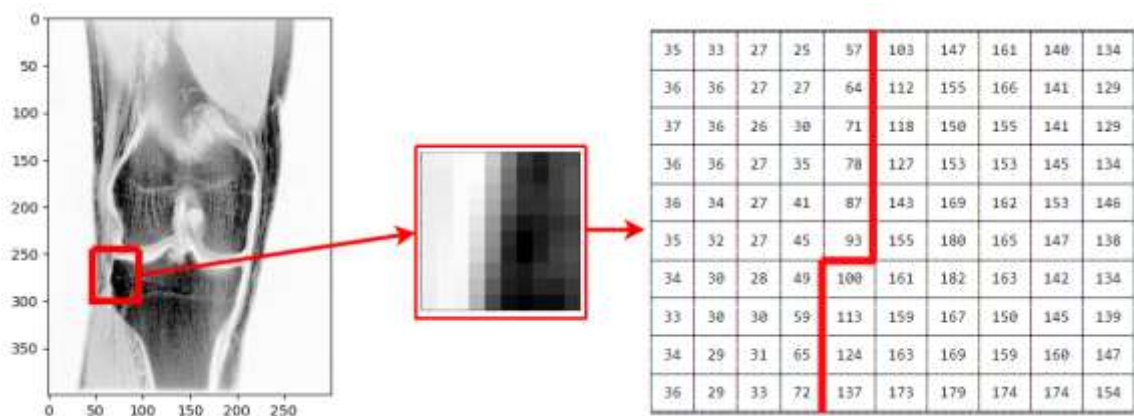


Рисунок 2.1 – Представлення цифрового зображення, визначення контурів

Згорткова нейронна мережа є частиною технологій глибокого навчання та архітектурою штучних нейронних мереж (ШНМ), спрямованою на ефективне розпізнавання образів. Вона використовує деякі особливості зорової кори головного мозку, в якій виявлено так звані прості та складні клітини. Прості клітини реагують на прямі лінії під різними кутами, а відповідь складних клітин пов'язана з активацією певного набору простих клітин.

Структура CNN є односпрямованою, або прямою, багат шаровою. У структурі нейронної мережі прямого поширення сигнал рухається в напрямку від вхідного шару до вихідного [17]. На сьогоднішній день розробки цього типу ШНМ широко розповсюджені і успішно вирішують задачі розпізнавання образів, прогнозування та кластеризації. Найбільш поширеним стандартним методом навчання ШНМ є метод зворотного поширення помилки (backpropagation). Метод використовується для ефективного навчання нейронної мережі за допомогою так званого ланцюгового правила. CNN зазвичай має три типи шарів: шари згортки, об'єднувальні та повнозв'язані шари.

2.2 Опис шарів згорткової нейронної мережі

Відмінною особливістю згорткових нейронних мереж є наявність так званого згорткового шару. Згортковий шар використовується для вилучення ознак з вхідних зображень. У цьому шарі виконується математична операція згортки між вхідним зображенням і фільтром, також відомим як ядро. Основним завданням шару згортки є виявлення локальних поєднань об'єктів з попереднього шару та відображення їх вигляду на карту ознак (feature map) [18], яка містить інформацію про зображення (наприклад, кути і межі). Пізніше ця карта ознак надсилається до інших шарів для того, щоб дізнатися інші особливості вхідного зображення.

У більшості випадків за шаром згортки слідує шар об'єднання або субдискретизації. Основна мета цього шару – зменшення просторового розміру вхідного зображення, прискорення обчислень за рахунок зменшення обчислювальних витрат, а також намагання зробити ШНМ більш загальною. Основною метою цього шару є зменшення розміру згорнутої ознаки для зменшення обчислювальних витрат. Зазвичай це виконується шляхом мінімізації кількості зв'язків між шарами та окремої роботи з кожною картою ознак.

Останній шар у CNN називається повнозв'язаним (fully-connected layer). Його призначення полягає в тому, щоб зробити класифікацію. Даний шар

моделює складну нелінійну функцію, яка буде оптимізована з метою покращення якості розпізнавання.

Після перетворення вхідного зображення у форму, придатну для багаторівневого перцептрона, ми повинні згладити зображення у вектор-стовпець. Тип вихідного шару впливає на вибір функції активації. Функції sigmoid та softmax, зазвичай, використовуються як функції активації вихідного шару мережі для розв'язання задачі класифікації, тобто коли необхідно віднести описаний об'єкт до одного з декількох класів. Softmax використовується для задачі класифікації з $N > 2$ класів і моделюванням значення ймовірності належності до класу. У даній роботі розглянуто задачу бінарної класифікації, тому функцію сигмоїди буде використано у якості функції активації вихідного шару. Однією з переваг сигмоїди для бінарної класифікації є те, що вона дає змогу інтерпретувати вихідне значення як імовірність належності до одного з двох класів. Тобто, якщо вихідне значення близьке до 1, то об'єкт класифікується як такий, що належить до першого класу, а якщо близьке до 0 – як такий, що належить до другого класу.

Вибір структури згорткової нейронної мережі залежить від конкретного завдання, розміру вхідних даних, доступних ресурсів для навчання та інших факторів. Зазвичай ядро згортки має розмір від 3×3 до 7×7 пікселів. Більш дрібні ядра можуть бути більш ефективними для розпізнавання деталей, в той час як більші ядра можуть бути ефективними для розпізнавання більших об'єктів.

2.3 Оцінка ефективності моделі

При розгляді і порівнянні різних моделей виникає питання оцінки якості мережі. Для моделей класифікації в машинному навчанні в якості таких критеріїв використовуються різні типи оціночних метрик, серед яких необхідно виділити наступні: точність (accuracy), достовірність (precision) та чутливість (recall) моделі. Використання різних метрик оцінки моделі машинного навчання є важливим для забезпечення повної оцінки ефективності та точності моделі.

Точність A є відсотком правильних прогнозів для тестових даних. Точність не підходить в якості міри ефективності моделі у випадку задачі класифікації при незбалансованих даних. Основною причиною цього є те, що велика кількість прикладів з мажорного класу (або класів) перебиває кількість прикладів у мінорному класі; це означає, що навіть неефективні моделі можуть досягати точності до 95%, залежно від того, наскільки серйозним є дисбаланс класів.

Релевантність даних є основою як для достовірності, так і для чутливості моделі.

Достовірність (precision) P оцінює частку правильно класифікованих екземплярів серед тих, що класифіковані як позитивні [19], визначено за формулою:

$$P = \frac{N_{TruePositive}}{N}, \quad N = N_{TruePositive} + N_{FalsePositive}, \quad (2.1)$$

де $N_{TruePositive}$ – кількість відповідних прогнозів;

N – кількість всіх прикладів, які за прогнозом належать до певного класу;

$N_{FalsePositive}$ – кількість спостережень, які модель помилково віднесла до певного класу.

У свою чергу чутливість (recall) R моделі визначено за формулою:

$$R = \frac{N_{TruePositive}}{M}, \quad M = N_{TruePositive} + N_{FalseNegative}, \quad (2.2)$$

де $N_{TruePositive}$ – кількість елементів, які були правильно передбачені за приналежністю до класу;

M – кількість всіх елементів, які дійсно належать до класу;

$N_{FalseNegative}$ – кількість спостережень, класифікованих як не належних до класу, тоді як вони насправді належать.

У незбалансованих наборах даних метою є покращення достовірності без шкоди для чутливості. Дані цілі, однак, часто є суперечливими, оскільки для

того, щоб підвищити $N_{TruePositive}$ мінорного класу, часто збільшують кількість $N_{FalsePositive}$, що призводить до зниження достовірності [20].

Для роботи з незбалансованими даними краще використовувати F - міру, яка об'єднує достовірність та чутливість в одну метрику. В залежності від контексту конкретної задачі вона надає змогу змінювати вплив вище зазначених метрик оцінки моделі.

F -міру можна визначити:

$$F = \frac{2 * R * P}{R + P}, \quad (2.3)$$

де R – чутливість моделі;

P – достовірність моделі.

Для незбалансованого набору даних, коли більший вплив мають класи з більшою кількістю екземплярів у наборі даних, перевага надається середньозваженому значенню F -міри. Причина полягає у впливі кожного класу у F -середнє, яке зважується за його розміром.

ROC-крива (Receiver Operator Characteristic) – це оціночна метрика, що використовується для задач бінарної класифікації. Це крива ймовірності, яка будує графік $N_{TruePositive}$ в залежності від $N_{FalsePositive}$ при різних порогових значеннях, відокремлюючи «сигнал» від «шуму». Площа під кривою (AUC) є мірою здатності бінарного класифікатора розрізнити класи. Вона використовується як узагальнення ROC-кривої. Чим вище AUC , тим краще модель розрізняє позитивні та негативні класи.

Тепер введемо функцію втрат. Це функція, яка порівнює цільове та прогнозоване вихідні значення і вимірює якість навчання. Однією з найбільш використовуваних функцій втрат у машинному навчанні є функція перехресної ентропії, відома також як функція логістичних втрат. Відомі наступні типи перехресної ентропії: бінарна та категоріальна, їх використання залежить від кількості передбачуваних класів у завданні [21]. У даній роботі буде використано бінарну крос-ентропію, що математично виражається наступним чином:

$$L_i = -(y_i * \log(p(x_i)) + (1 - y_i) * \log(1 - p(x_i))), \quad (2.4)$$

де для L_i екземпляра $p(x_i)$ – прогнозована ймовірність того, що екземпляр x_i належить до класу;

y_i – булева мітка для прикладу i [22].

Втрати перехресної ентропії мінімальні, коли різниця між прогнозованою ймовірністю та істинним значенням мала, і набуває великих значень у протилежному випадку.

Може статися так, що модель зазнає складнощів під час апроксимації навчальних даних. Такий випадок відбувається, наприклад, коли лінійна модель навчається на експоненціальному наборі даних. Даний тип ситуації називається недостатньою пристосованістю (underfitting) або моделлю з високим зміщенням. Недостатня пристосованість виникає, коли модель не здатна зменшити похибку ані для тестової, ані для навчальної вибірки. Модель недостатньо потужна для того, щоб відповідати складності розподілу даних. З точки зору функції втрат, це означає, що модель дає недостатньо малу помилку як на навчальній, так і на валідаційній вибірках.

Тому метод, застосований для зменшення недостатньої пристосованості, зазвичай полягає у збільшенні коефіцієнту швидкості навчання (learning rate) [23]. Часто архітектуру ШНМ потрібно модифікувати, тобто збільшити кількість шарів у моделі або збільшити кількість нейронів у кожному шарі. Також може допомогти зміна порядку шарів.

Інша проблема, яка може виникнути під час навчання моделі, називається перенавчанням (overfitting). Це відбувається в тому випадку, коли модель машинного навчання починає захоплювати не тільки сигнал, але і шум в даних. Дана ситуація призводить до коливань в моделі, які не відображають справжній тренд. Таку проблему ще називають високою дисперсією моделі. Іншими словами, модель машинного навчання не здатна добре узагальнювати (помилка узагальнення зростає) через те, що занадто багато уваги приділяється навчальним даним без подальшої адаптації до нових даних. У цьому випадку

помилка на навчальній вибірці досить мала, тоді як помилка на валідаційній вибірці набагато більше.

Тому основним завданням є знаходження точки балансу між недостатнім та надмірним пристосуванням моделі, як проілюстровано на рисунку 2.2.

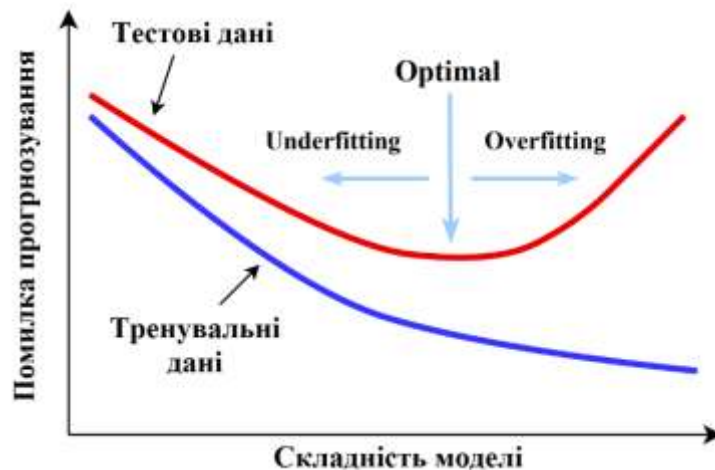


Рисунок 2.2 – Баланс між недостатньою пристосованістю і перенавчанням

Найпоширенішими методами зменшення високої дисперсії є доповнення даних, регуляризація та зменшення архітектурної складності. Для того, щоб отримати найкращі результати узагальнення моделі, необхідно мати більше даних, включаючи різні варіації. Тобто, необхідно штучно збільшити розмір вхідної вибірки, це можна зробити за допомогою аугментації даних. Для створення необхідної кількості прикладів для навчання можна взяти кілька «ідеальних» прикладів і застосувати до них різні деформації.

Найпоширеніший метод, який використовується для регуляризації нейронних мереж, називається методом виключення (Dropout).

3 МОДЕЛЮВАННЯ СИСТЕМИ

3.1 Опис та попередня обробка даних

Набір даних, що було обрано для даного дослідження, складається з 1370 МРТ-зображень колінного суглоба, виконаних у Медичному центрі Стенфордського університету [24]. Набір даних містить 1104 (80,6%) патологічних зразка, з яких 319 (23,3%) – розриви передньої хрестоподібної зв'язки та 508 (37,1%) розривів менісків. Маркування були отримані шляхом ручного вилучення з клінічних звітів.

Розглянутий набір даних містить три папки для кожної площини (корональна, сагітальна та аксіальна) з файлами .pru кожного з 1130 пацієнтів. Цей набір використовується для навчання моделі. В той час як дані, використані для тестування моделі, містять файли 120 пацієнтів.

Для подальшої роботи дані було програмно конвертовано у файли формату .jpg, роздільною здатністю 256×256 пікселів для кожного зрізу. Вигляд отриманих даних проілюстровано на рисунку 3.1. Перетворення зображень у формат JPG не погіршує якість зображення завдяки стисненню. Для цього було використано модуль бібліотеки PIL (Python Imaging Library) із фільтром Image.ANTIALIAS, що використовує високоякісний алгоритм на основі згорток.

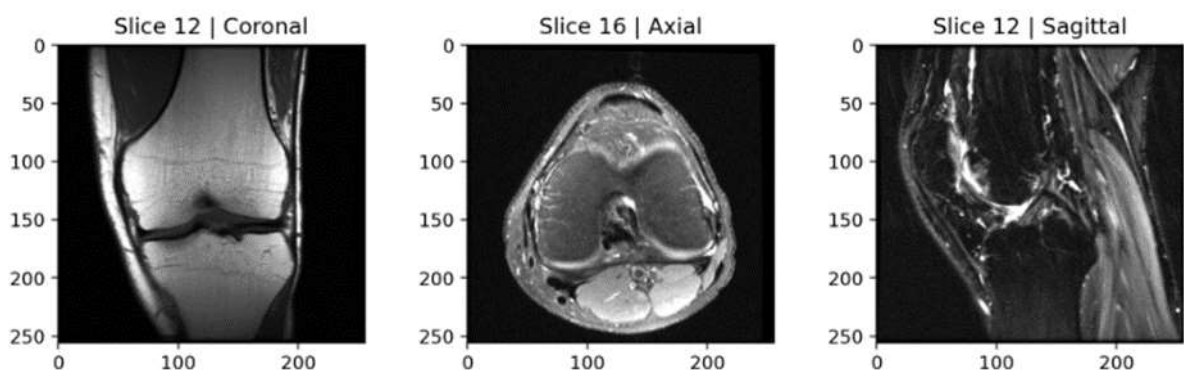


Рисунок 3.1 – Центральні зрізи корональної, аксіальної та сагітальної площин засобами бібліотеки Matplotlib

Слід зазначити, що кількість зрізів для кожної площини відрізняється для різних пацієнтів, і, більш того, кількість зрізів одного пацієнта є різною для кожного з трьох площин, структуру приведено на рисунку 3.2. Обидва випадки пов'язані з фізичною індивідуальністю людини і роблять задачу навчання моделі більш складною. Для того, щоб вирішити цю проблему, необхідно нормалізувати дані.

У даному дослідженні буде розглянута задача бінарної класифікації з метою пошуку аномалій та патологій колінного суглобу. Тому також слід звернути увагу на збалансованість навчальних і тестових даних. Було помічено, що позитивне цільове значення переважає в обох наборах даних. Пропорція становить 80,8% до 19,2% (913/217 випадків) для навчання і 79,1% до 20,9% (95/25 випадків) для тестування.

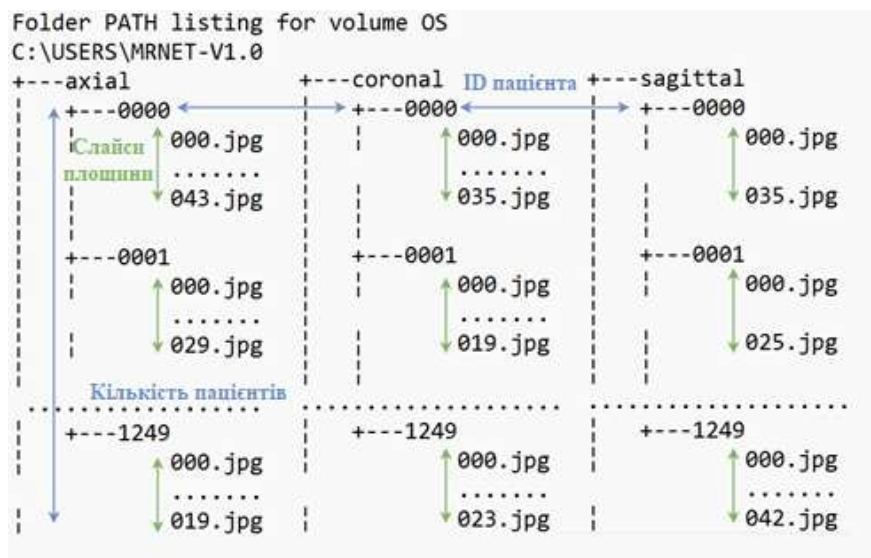


Рисунок 3.2 – Структура вхідних даних

Незбалансований набір даних може спричинити перенавчання через нестачу навчальних даних. Дана ситуація є прикладом того, коли необхідно використовувати додаткові метрики (F -міра та AUC), щоб отримати об'єктивну оцінку роботи моделі. У деяких випадках проблему незбалансованих даних

можна вирішити за допомогою стратифікованої К-кратної перехресної перевірки [25] або регуляризації.

Найбільш поширеними способами вирішення проблеми незбалансованості даних при класифікації зображень є аугментація даних та використання стохастичних градієнтних оптимізаторів [26], наприклад, Adam. Враховуючи індивідуальність колінного суглоба кожної людини та похибки при проведенні магнітно-резонансної томографії були запропоновані наступні типи методів аугментації даних (візуалізацію приведено на рисунку 3.3):

- поворот на 15 градусів;
- збільшення/зменшення масштабу на 10%;
- зсув по вертикалі та горизонталі на 0,05%.

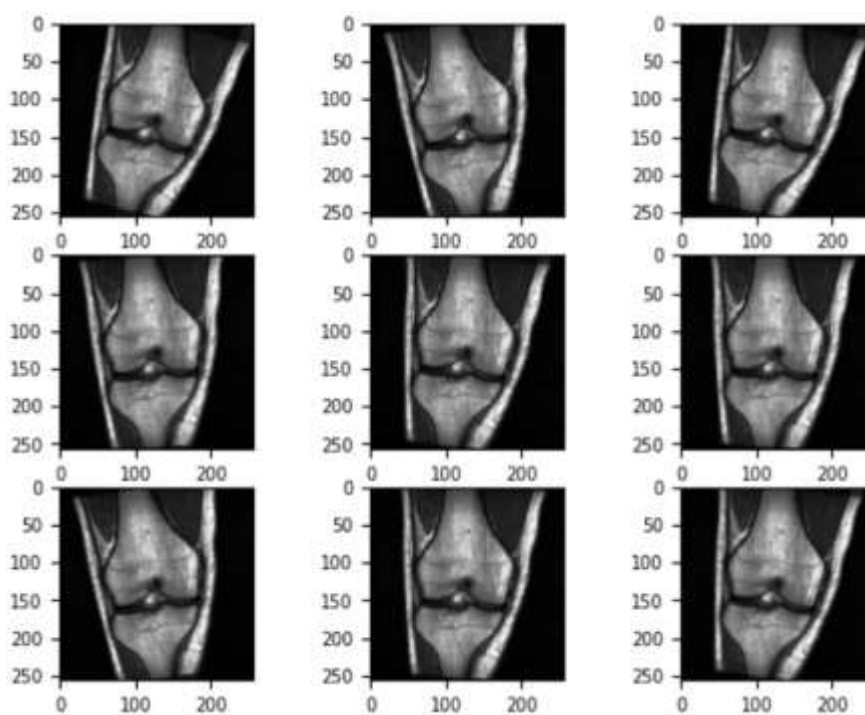


Рисунок 3.3 – Аугментація даних засобами tensorflow.keras.preprocessing.image

3.2 Опис методу і проектування моделей

Зазвичай для тривимірних моделей навчання, як МРТ і КТ, використовують архітектуру 3d-CNN або U-Net, але вони не завжди враховують

індивідуальність конкретного органу. У даній роботі було запропоновано поділити кожне скановане зображення на декілька областей для того, щоб отримати і дослідити тримірну модель органу, пошкодження якого видно на МРТ-зображенні. Найбільш прийнятними були центральне сканування і зсув на 10% в кожную сторону. Запропонований поділ приведено на рисунку 3.4.

На відміну від рентгенографії, де отримують двомірне зображення, магнітно-резонансна томографія вивчає окремі площини. Це необхідно через анатомічну структуру органу (тканини мають різний ступінь «прозорості», видимий на МРТ). Запропонований поділ враховує дані особливості. У подальших дослідженнях заплановано розглянути інші поділи в залежності від органу, що досліджується, якості даних та обчислювальних можливостей комп'ютеру, на якому буде проведено навчання моделей.



Рисунок 3.4 – Запропонований у дослідженні поділ зрізів

Таким чином, необхідно навчити 9 моделей ШНМ, тобто по 3 моделі для кожної з трьох площин. У подальших дослідженнях планується розглянути інші поділи в залежності від індивідуальних особливостей органу.

З метою структурування та розділення даних на навчальні/валідаційні/тестові набори (пропорція навчання/валідації 70/30, навчальні/тестові 90/10) було використано функцію `flow_from_dataframe()`

модуля `keras.preprocessing`. Набір даних був перемішаний і розділений з генератором випадковості, що необхідно для якісного порівняння продуктивності моделей.

Структуру моделі згорткової нейронної мережі продемонстровано на рисунку 3.5.

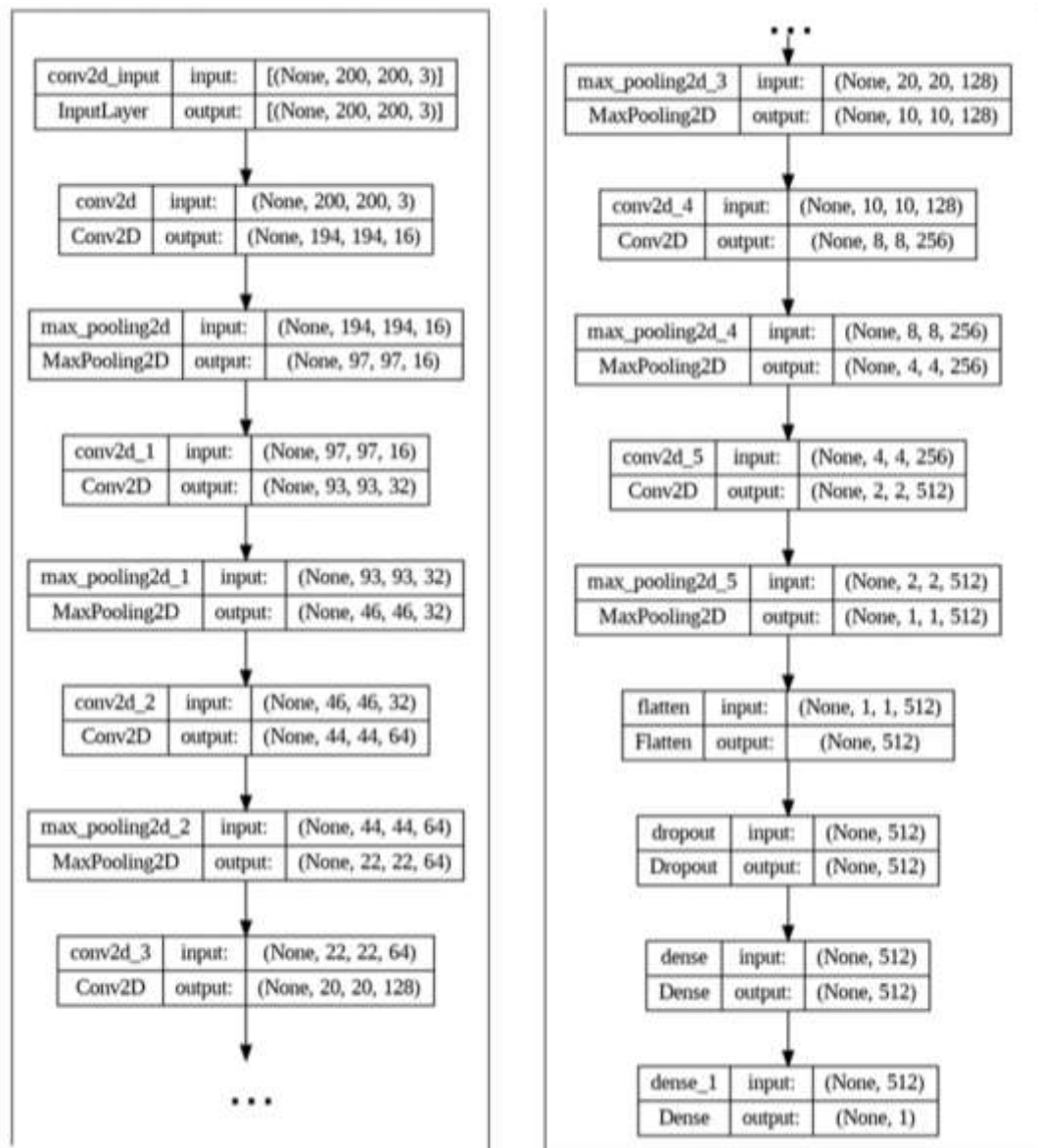


Рисунок 3.5 – Структура запропонованої згорткової мережі

Запропонована послідовна (`sequential`) модель має 6 шарів згортки з різною кількістю фільтрів та розміром фільтрів; кількість фільтрів збільшується зі

збільшенням глибини нейронної мережі. Збільшення кількості фільтрів у кожному шарі при збільшенні глибини мережі дає змогу моделі навчатися більш складним та абстрактнішим функціям, що може покращити точність ШНМ. Також збільшення кількості фільтрів допомагає компенсувати зменшення розміру вхідних даних (зменшення розміру вхідних даних призведе до втрати інформації) у кожному шарі, що відбувається під час використання операцій згортання та об'єднання.

У якості активаційної функції було використано ReLU. Розміри зображень було змінено для зменшення обчислювальної складності під час навчання, тому вхідне зображення має розмір $200 \times 200 \times 3$. Після кожного шару згортки розташовано шар об'єднання (pooling layer).

Навчання мереж було проведено за допомогою оптимізатора Adam з параметром швидкості навчання $5 * 10^{-6}$. Таку малу швидкість навчання було обрано згідно з [23], навчання займало більшу кількість епох, але збіжність була гарантована. Як було зазначено в [27], метод оптимізації Адам є обчислювально ефективний, не вимагає багато пам'яті, інваріантний до діагонального масштабування градієнтів і добре підходить для задач з великими обсягами даних та параметрів.

Розмір батча дорівнює 8. Малий розмір пакету було обрано для того, щоб мати більше ефектів регуляризації. Вплив розміру батча на продуктивність обумовлено тим, що великі розміри батча можуть прискорити процес навчання шляхом зменшення кількості оновлень ваг на кожній ітерації навчання. Однак, занадто великий розмір батча може вимагати більше пам'яті та процесорного часу, що може призвести до проблем із продуктивністю.

Функцію втрат задано як бінарну функцію перехресної ентропії через вирішення задачі бінарної класифікації. Щоб досягти найбільш об'єктивних результатів, необхідно ретельно обирати метрики відповідно до завдання й аналізувати їхні результати в контексті обраної архітектури моделі, її гіперпараметрів і процесу навчання. Обрано наступні метрики: точність та AUC пакета `keras.metrics`, F-міру пакета `tensorflow_addons`.

3.3 Тестування моделей і аналіз результатів

Експерименти було розпочато з навчання моделей для центрального зрізу. Процедуру навчання/валідації, а саме значення функцій точності та втрат в залежності від епохи для центральних зрізів кожної площини проілюстровано на рисунках 3.6–3.8.

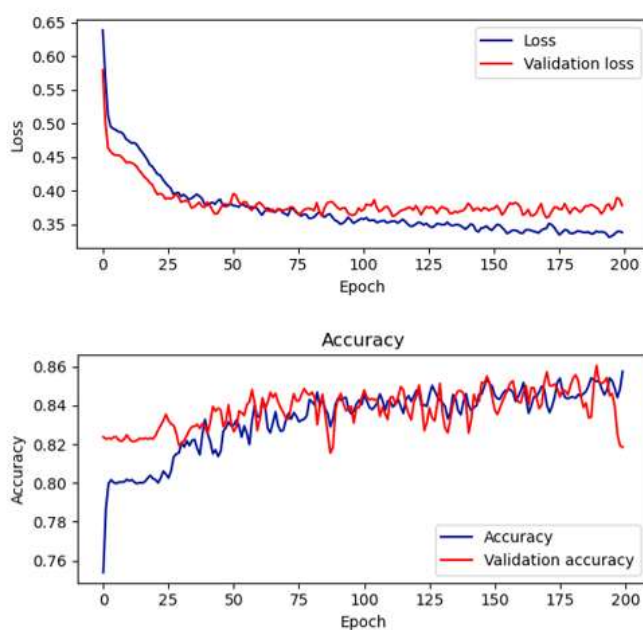


Рисунок 3.6 – Процедура навчання/валідації. Залежність точності та втрат від епохи для аксіальної площини

Дані графіки дають змогу оцінити процес навчання моделі та зрозуміти, наскільки успішно вона навчилася. Існує декілька ознак того, що модель навчилася: збіжність графіків та відстань між функціями тренування та валідації. Наявність плато в графіку втрат і точності на валідаційних даних може вказувати на те, що модель досягла оптимального значення. Якщо відстань між графіками функцій збільшується з кожною епохою навчання, це може бути ознакою перенавчання моделі. З іншого боку, якщо відстань зменшується або залишається постійною, це може вказувати на поліпшення узагальнюючої здатності моделі.

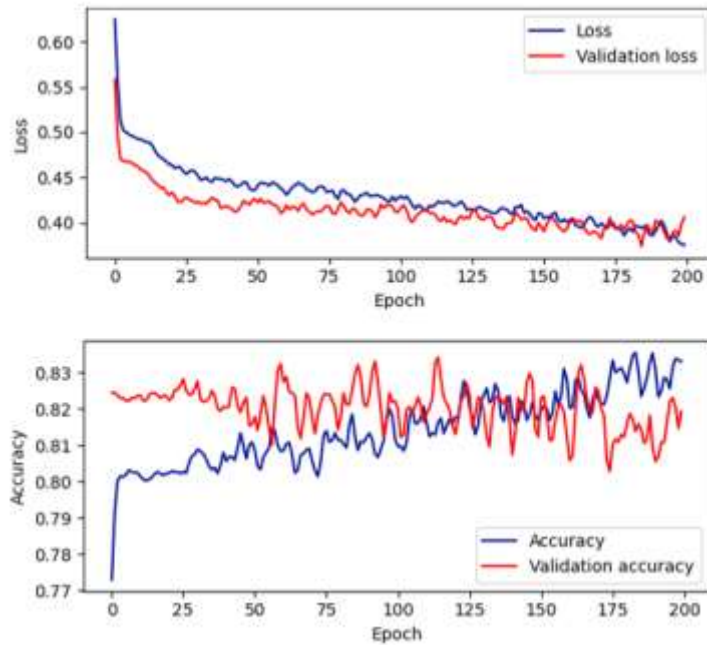


Рисунок 3.7 – Процедура навчання/валідації для корональної площини

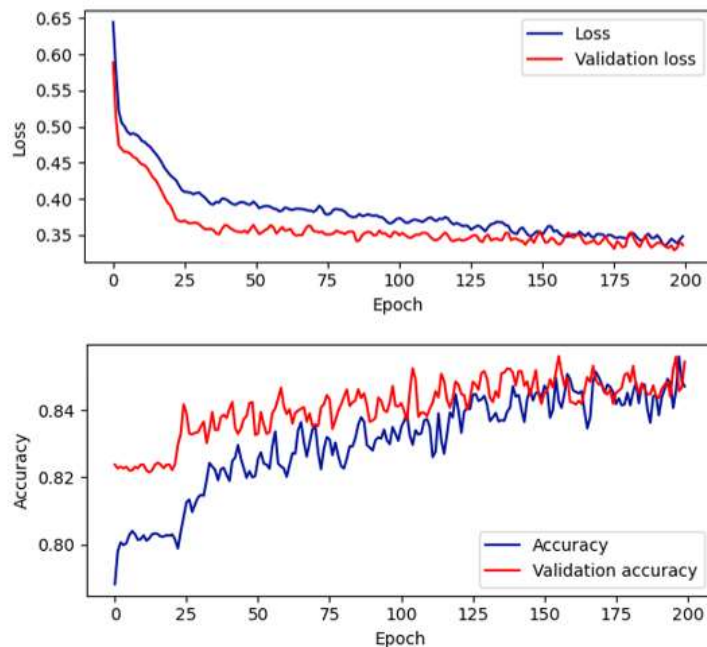


Рисунок 3.8 – Процедура навчання/валідації для сагітальної площини

Як було зазначено вище, вибірка містить зміщені дані, тому для отримання об'єктивної оцінки ефективності моделі були використані додаткові метрики. Втрати, точність, а також AUC та F-метрику проілюстровано в таблиці 3.1.

Таблиця 3.1 – Результати роботи моделей центрального зрізу

	Аксіальна	Корональна	Сагітальна
Test loss	0.352	0.510	0.450
Точність	0.866	0.808	0.850
AUC	0.856	0.657	0.807
F-міра	0.883	0.884	0.884

Макро- та середньозважені показники достовірності, чутливості та F-метрики для кожного цільового класу приведено у додатку А. Аналізуючи результати, слід звернути увагу на показники F-міри та AUC, оскільки точність не дасть реалістичної оцінки для незбалансованих даних.

Матриці помилок приведено на рисунку 3.9. Бувають випадки, що модель «пропускає» хворобу, коли вона присутня (прогнозоване значення цільової змінної дорівнює 0, тоді як фактичне значення дорівнює 1). Це називається хибно-негативним прогнозом, або помилкою II типу. У медичних додатках таку ситуацію слід мінімізувати.

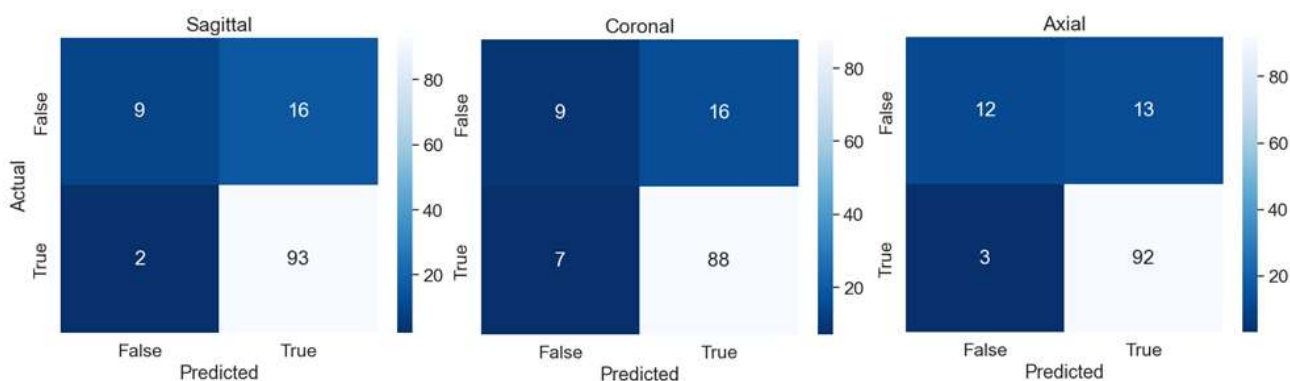


Рисунок 3.9 – Матриці помилок, отримані при тестуванні центральних зрізів сагітальної, коронарної та аксіальної площин відповідно

Аналогічні експерименти були проведені на 40% і 60% зрізів відповідно до розподілу, згаданого вище (рисунок 3.3). Процедура навчання/валідації на основі точності та втрат, а також результати метрик для моделей продуктивності та матриці помилок для кожної моделі наведено в додатку А.

З метою отримання результату з більшою ефективністю, було запропоновано об'єднати всі моделі. У даній роботі було використано середньозважений ансамбль класифікаторів. Це ансамблевий підхід в машинному навчанні, який об'єднує прогнози декількох моделей, причому внесок кожної моделі зважується відповідно до її продуктивності. Остаточний результат отримується через скалярний добуток вектору вагових коефіцієнтів на вектор прогнозів моделі; після цього береться сума або найбільш ймовірний клас [28]. Середньозважене значення AUC для 9 моделей було нормалізовано і взято у якості вектору вагових коефіцієнтів (0.862, 0.606, 0.829, 0.857, 0.658, 0.808, 0.800, 0.692, 0.844). Таким чином, кожна модель має різний вплив на кінцевий результат, пропорційний її індивідуальній продуктивності.

Точність 85,83% і середньозважена F-міра 91,79% були досягнуті завдяки поєднанню всіх попередньо навчених моделей, представлено в таблиці 3.2, що є значним покращенням у порівнянні з показниками окремих моделей. У свою чергу, AUC став 81,22% (рисунок 3.10).

Таблиця 3.2 – Результати ансамбля класифікаторів

Метрики	Результати роботи ансамбля класифікаторів
Точність	85.83%
Середньозважена достовірність	84.82%
F-міра	91.79%

Крім того, слід звернути увагу на кількість хибно-негативних прогнозів (помилка II типу), яка тепер дорівнює 0 (рисунок 3.11). Цей факт є надзвичайно важливим при розробці системи підтримки прийняття рішень у медичній галузі. Таким чином, модель показала кращий результат, ніж розглянуті вище існуючі методи. Так запропонована модель має краще значення оцінки ефективності, ніж згортова нейронна мережа VGG16 з AUC 71%, і майже досягла продуктивності ResNet18 (84%).

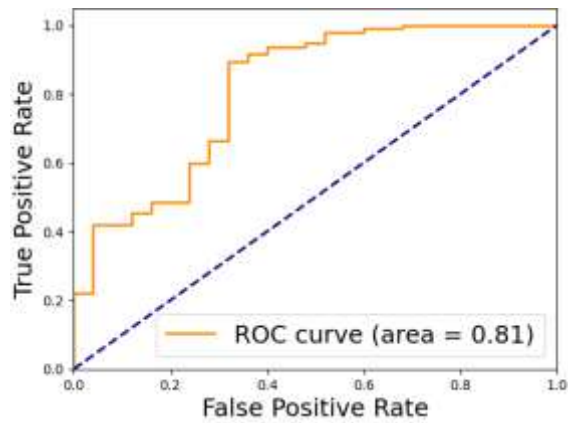


Рисунок 3.10 – AUC-крива ансамблевої моделі

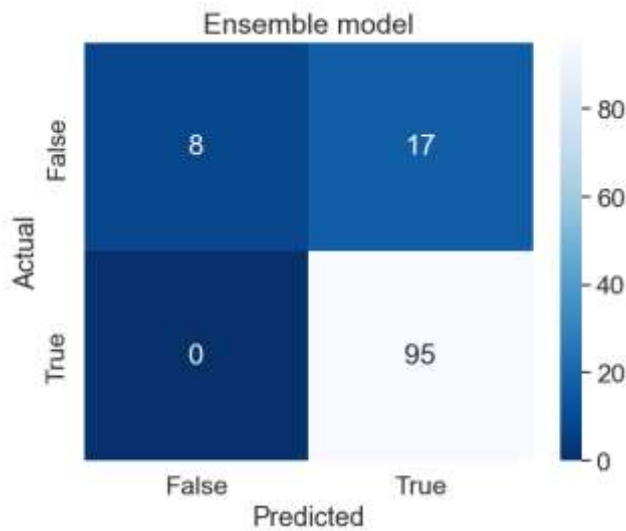


Рисунок 3.11 – Матриця помилок ансамблевої моделі

3.4 Практичне застосування і перспективи розвитку

Веб-інтерфейс системи підтримки прийняття рішень реалізовано на основі технологій User Experience Design (UX) / User Interface Design (UI). Додаток, що проведено на рисунку 3.12, надає можливість швидко отримати повну індивідуальну картину необхідних площин та зрізів для конкретного пацієнта, генеруючи анімацію для безперервного перегляду сканів. Система також включає в себе функцію автоматизованої інтерпретації з імовірнісними

показниками наявності патології, що виконується за допомогою використання технологій штучного інтелекту. Процес інтерпретації МРТ в локальній мережі займає від 5 до 7 секунд залежно від вхідних даних. Розроблений додаток сумісний з останніми версіями браузерів Google Chrome, Firefox, Microsoft Edge, Opera та Safari.

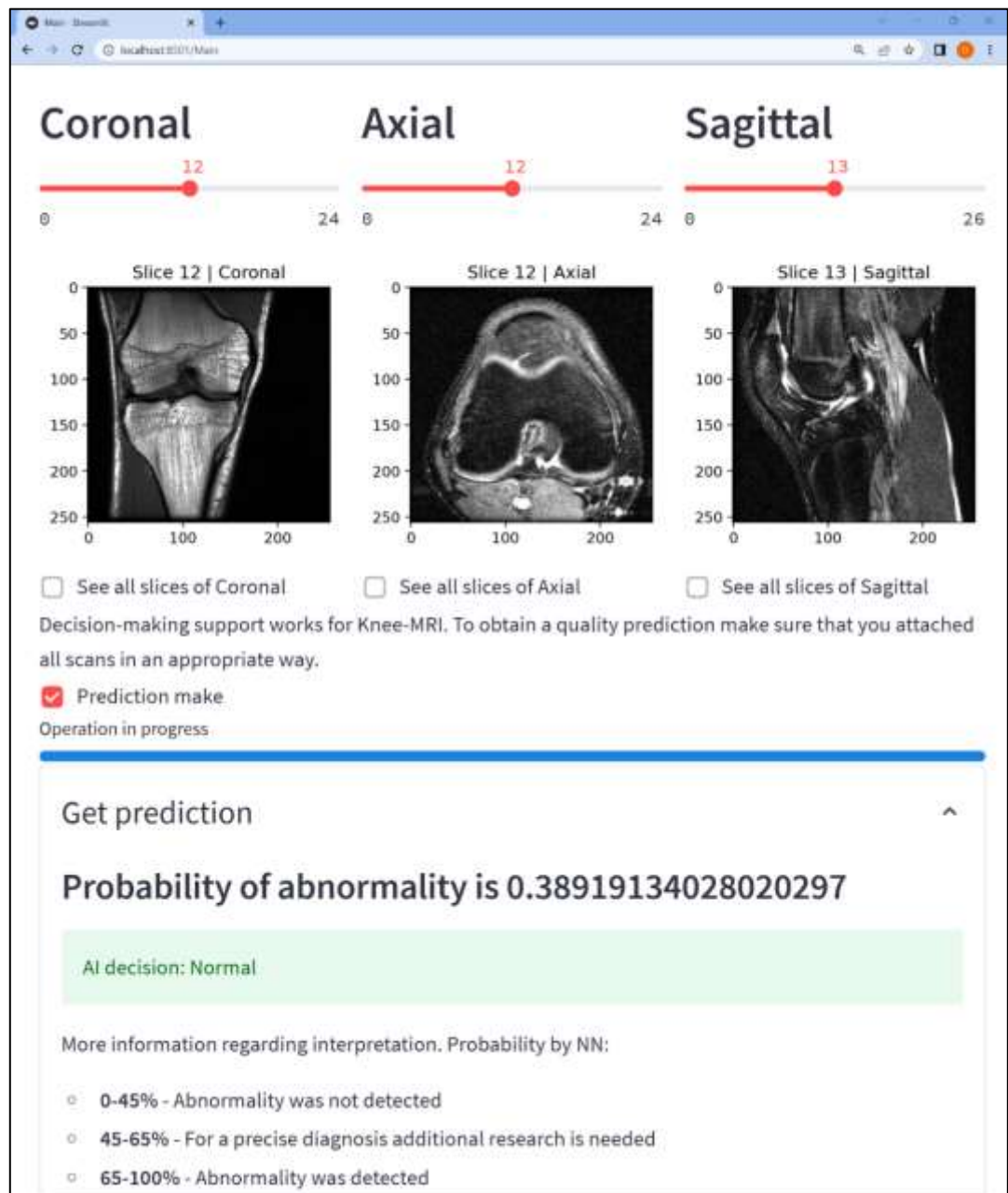


Рисунок 3.12 – Інтерфейс головної сторінки системи

ВИСНОВКИ І ПЕРСПЕКТИВИ РОЗВИТКУ

Основною задачею дослідження були аналіз і вирішення проблеми діагностики зображень магнітно-резонансної томографії засобами штучного інтелекту. Було проведено порівняльний аналіз сучасних методів інтерпретації медичних зображень, аналіз та підготовка навчальної виборки шляхом аугментації даних.

В результаті дослідження запропоновано новий метод діагностики на основі технології глибокого навчання, а саме – згорткових нейронних мережах зі специфічним розподілом зрізів зображень магнітно-резонансної томографії. Даний метод передбачає поділ кожного сканованого зображення на декілька частин, що уможливить цілковите дослідження органу, чиє пошкодження видно на МРТ. Було запроєктовано та програмно реалізовано систему підтримки прийняття рішень для діагностики патологій колінного суглоба. Наукова новизна даної роботи полягає в тому, що запропонований підхід ще не був використаний для діагностики 3D медичних зображень раніше. Найбільш придатними виявились центральний зріз та зсув на 10% в обидві сторони для аксіальної, корональної та сагітальної площин. На отриманих вибірках було натреновано низку нейронних мереж для кожної площини та проведено синтез за допомогою використання середньозваженого ансамблю класифікаторів.

Під час попередньої обробки даних було виявлено, що набір даних проблеми є незбалансованим, що ускладнює задачу тренування моделей машинного навчання. Дану проблему було вирішено шляхом доповнення даних (аугментації) та використання стохастичного градієнтного оптимізатора. Таким чином, дев'ять нейронних мереж були навчені та інтегровані з використанням середньозваженого ансамблю класифікаторів. Було проведено тестування моделі. Для отримання об'єктивної оцінки ефективності системи були використані додаткові метрики (F-міра та AUC). Точність 85,83% і середньозважена F-міра 91,79% були досягнуті завдяки поєднанню всіх попередньо навчених нейронних мереж, що є значним покращенням у

порівнянні з показниками окремих моделей. У свою чергу, AUC став 81,22%, що означає, що запропонована модель сягнула вищого значення оцінки ефективності, ніж згортова нейронна мережа VGG16 з AUC 71%, і майже досягла продуктивності відомої ШНМ ResNet18 (84%).

Крім того, слід зазначити, що кількість хибно-негативних прогнозів (помилка II типу) дорівнює 0. Цей факт є надзвичайно важливим при розробці системи підтримки прийняття рішень у медичній галузі. Таким чином, модель показала кращий результат, ніж розглянуті вище існуючі методи.

Запропонований додаток є гнучким завдяки обраним технологіям проектування. Серед подальшого розвитку заплановано підвищення точності моделі підтримки прийняття рішень. Це може бути досягнуто шляхом поповнення бази даних зображень МРТ для навчання моделі, використання трансферного навчання та нечіткої логіки, а також зміни запропонованого поділу зрізів МРТ. Також передбачено розширення діагностування ШІ на інші області організму. Унаслідок воєнних дій в Україні велика кількість травмованих людей потребує на своєчасну діагностику пошкодження. За допомогою представленого у роботі метода за наявності відповідних маркірованих даних може бути запроєктовано та реалізовано системи діагностування МРТ хребта, окремих кісток, таза, грудної клітини тощо.

Інтерфейс системи підтримки прийняття рішень було запроєктовано за допомогою Experience Design (UX) / User Interface Design (UI) технологій. Додаток надає можливість швидко отримати повну індивідуальну картину необхідних площин (коронарних, сагітальних, аксіальних) та зрізів для конкретного пацієнта, генерувати анімацію для перегляду динаміки (анімовані gif-файли). Система включає інтелектуальну складову, а саме: функціонал інтерпретації МРТ-зображень з імовірнісними показниками наявності патології, що реалізовано за допомогою використання вищезгаданих технологій і методів штучного інтелекту.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Dawson M. J. Paul Lauterbur and the Invention of MRI. *The MIT Press*, 2013. URL: <https://doi.org/10.7551/mitpress/9146.001.0001> (date of access: 07.02.2023).
2. Human anatomy planes. URL: https://commons.wikimedia.org/w/index.php?title=File:Human_anatomy_planes.svg&oldid=448399308 (дата звернення: 17.03.2023).
3. Deep-learning-assisted diagnosis for knee magnetic resonance imaging: Development and retrospective validation of MRNet / N. Bien et al. *PLOS Medicine. Public Library of Science*. 2018. Vol. 15, no. 11. URL: <https://doi.org/10.1371/journal.pmed.1002699> (date of access: 19.03.2023).
4. Nacey N. C., Geeslin M. G., Miller G. W., Pierce J. L. Magnetic resonance imaging of the knee: an overview and update of conventional and state of the art imaging. *Journal of Magnetic Resonance Imaging. Wiley Online Library*. 2017. Vol. 45, no. 5. P. 1257–1275.
5. Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*. 1980. Vol. 36. P. 193–202.
6. Learning algorithms for classification: A comparison on handwritten digit recognition / Y. LeCun et al. *Neural networks: the statistical mechanics perspective*. 1995. P. 261–276.
7. Deep Learning for Health Informatics / D. Ravi et al. *IEEE Journal of Biomedical and Health Informatics*. 2017. Vol. 21, no. 1. P. 4–21. DOI: 10.1109/JBHI.2016.2636665.
8. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* / A. Esteva et al. *Nature Publishing Group*. 2017. Vol. 542, no. 7639. P. 115–118. DOI: 10.1038/nature21056.
9. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning / H.-C. Shin et al. *IEEE*

Transactions on Medical Imaging. 2016. Vol. 35, no. 5. P. 1285–1298. DOI: 10.1109/TMI.2016.2528162.

10. A survey on deep learning in medical image analysis / G. Litjens et al. *Medical image analysis*. Elsevier. 2017. Vol. 42. P. 60–88.

11. Laiton-Bonadiez C., Sanchez-Torres G., Branch-Bedoya J. Deep 3D Neural Network for Brain Structures Segmentation Using Self-Attention Modules in MRI Images. *Sensors*. 2022. Vol. 22, no. 7. P. 2559. URL: <https://doi.org/10.3390/s22072559> (дата звернення: 17.03.2023).

12. Haddadian J., Balamurali M. Transfer Learning and Data Augmentation in the Diagnosis of Knee MRI. In: Long, G., Yu, X. and Wang, S., ed. *AI 2021: Advances in Artificial Intelligence*. Cham: Springer International Publishing. 2022. P. 452–463.

13. Deep learning approach for evaluating knee MR images: achieving high diagnostic performance for cartilage lesion detection / F. Liu et al. *Radiology. Radiological Society of North America*. 2018. Vol. 289, no. 1. P. 160.

14. Knee Injury Detection Using Deep Learning on MRI Studies: A Systematic Review / A. Siouras et al. *Diagnostics*. 2022. Vol. 12, no. 2. ISSN 2075-4418. URL: <https://www.mdpi.com/2075-4418/12/2/537> (дата звернення: 17.03.2023).

15. Deep learning reconstruction enables prospectively accelerated clinical knee MRI / P. M. Johnson et al. *Radiology. Radiological Society of North America*. 2023. P. 220425.

16. Інформаційні технології підтримки прийняття рішень при визначенні порушень носового дихання : монографія / О. Г. Аврунін та інші. Харків : «Інтерпретація МРТ», 2018. 132 с.

17. Руденко О. Г., Бодянський Є. В. Штучні нейронні мережі: навч. посібник. Харків : СМІТ, 2006. 404 с.

18. Abraham A., Kumar A., Melin P. and Gandhi N. Intelligent Systems Design and Applications. *Springer Cham*. Vellore, India, 2020.

19. Learning from imbalanced data sets / A. Fernández et al. *Springer*. 2018. Vol. 10. P. 978-3.

20. Ma Y., He H. Imbalanced learning: foundations, algorithms and applications. *John Wiley & Sons*. 2013.

21. Obi J. A Comparative Study of Several Classification Metrics and Their Performances on Data. *World Journal of Advanced Engineering Technology and Sciences*. 2023. Vol. 8., no. 1. P. 308–314. DOI: 10.30574/wjaets.2023.8.1.0054.

22. Wali R. Xtreme Margin: A Tunable Loss Function for Binary Classification Problems. URL: <https://doi.org/10.48550/arXiv.2211.00176> (дата звернення: 17.03.2023).

23. Smith L. N. A disciplined approach to neural network hyper-parameters: Part 1--learning rate, batch size, momentum, and weight decay. URL: <https://doi.org/10.48550/arXiv.1803.09820> (дата звернення: 17.03.2023).

24. StanfordMLGroup.MRNet Dataset. URL: <https://stanfordmlgroup.github.io/competitions/mrnet/> (дата звернення: 17.03.2023).

25. «Інтерпретація МРТ»

26. Valova I., Harris C., Mai T., Gueorguieva N. Optimization of Convolutional Neural Networks for Imbalanced Set Classification. *Procedia Computer Science*. 2020. Vol. 176. P. 660–669. URL: <https://doi.org/10.1016/j.procs.2020.09.038> (дата звернення: 17.03.2023).

27. Kingma D., Ba J. Adam: A Method for Stochastic Optimization. *International Conference on Learning Representations*. 2014. URL: <https://doi.org/10.48550/arXiv.1412.6980> (дата звернення: 17.03.2023).

28. Cao Y., Geddes T. A., Yang J. Y. H., Yang P. Ensemble deep learning in bioinformatics. *Nature Machine Intelligence*. 2020. Vol. 2, no. 9. P. 500–508. URL: <https://doi.org/10.1038/s42256-020-0217-y> (дата звернення: 17.03.2023).

29. Deep Learning Reconstruction Enables Prospectively Accelerated Clinical Knee MRI / P. M. Johnson et al. *Radiology*. 2023. URL: <https://doi.org/10.1148/radiol.220425> (дата звернення: 17.03.2023).

ДОДАТОК А

Навчання і тестування моделей

```
Test loss: 0.45079952478408813
Test accuracy: 0.8500000238418579
Auc: 0.8067368268966675

120/120 [=====] - 2s 13ms/step
      precision    recall  f1-score   support

   class 0       0.82     0.36     0.50         25
   class 1       0.85     0.98     0.91         95

 accuracy                   0.85         120
 macro avg       0.84     0.67     0.71         120
 weighted avg    0.85     0.85     0.83         120
```

Рисунок А.1 – Метрики продуктивності сагітальної моделі для центральних зрізів

```
Test loss: 0.5104283094406128
Test accuracy: 0.8083333373069763
Auc: 0.6568421125411987

120/120 [=====] - 2s 12ms/step
      precision    recall  f1-score   support

   class 0       0.56     0.36     0.44         25
   class 1       0.85     0.93     0.88         95

 accuracy                   0.81         120
 macro avg       0.70     0.64     0.66         120
 weighted avg    0.79     0.81     0.79         120
```

Рисунок А.2 – Метрики продуктивності корональної моделі для центральних зрізів

```

Test loss: 0.35231342911720276
Test accuracy: 0.8666666746139526
Auc: 0.8564209938049316

120/120 [=====] - 1s 11ms/step
      precision    recall  f1-score   support

   class 0       0.80     0.48     0.60        25
   class 1       0.88     0.97     0.92        95

 accuracy
macro avg       0.84     0.72     0.76       120
weighted avg    0.86     0.87     0.85       120

```

Рисунок А.3 – Метрики продуктивності аксіальної моделі для центральних зрізів

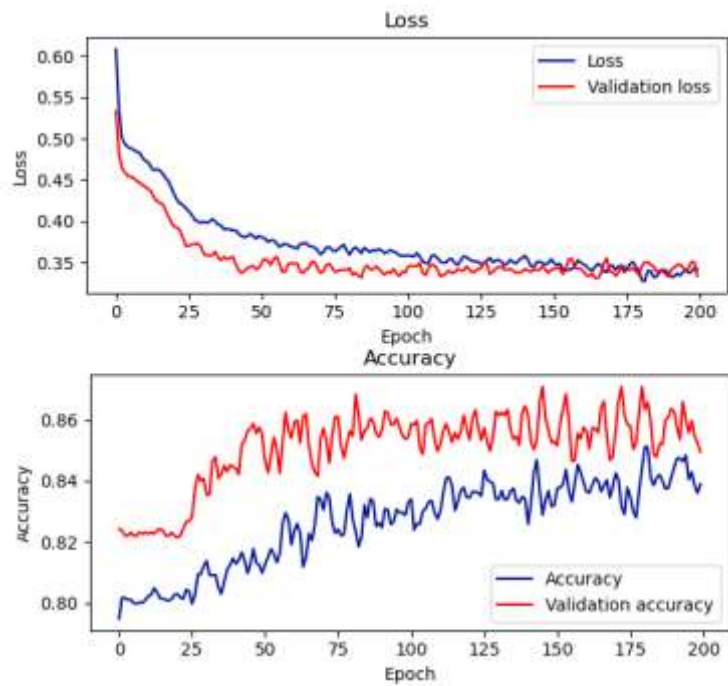


Рисунок А.4 – Процедура навчання/валідації на основі точності та втрат для 40% зрізу аксіальної моделі

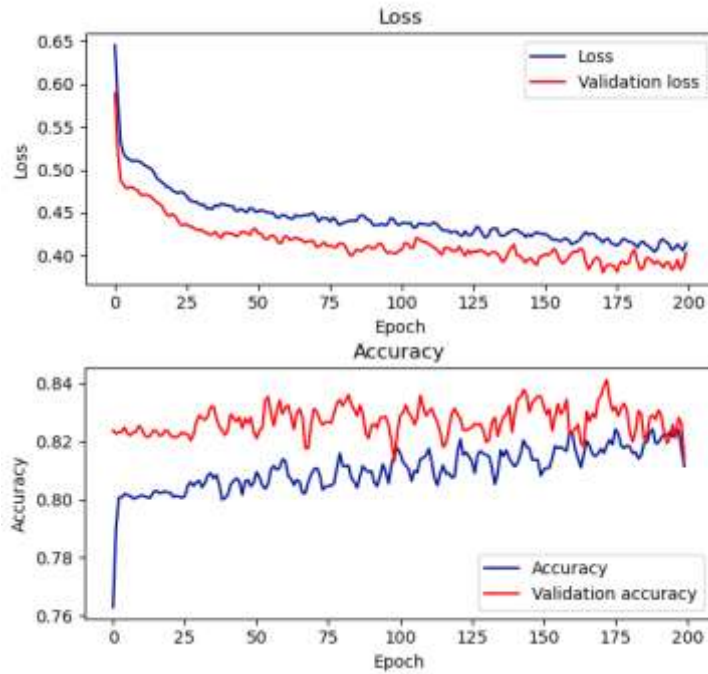


Рисунок А.5 – Процедура навчання/валідації на основі точності та втрат для 40% зрізу корональної моделі

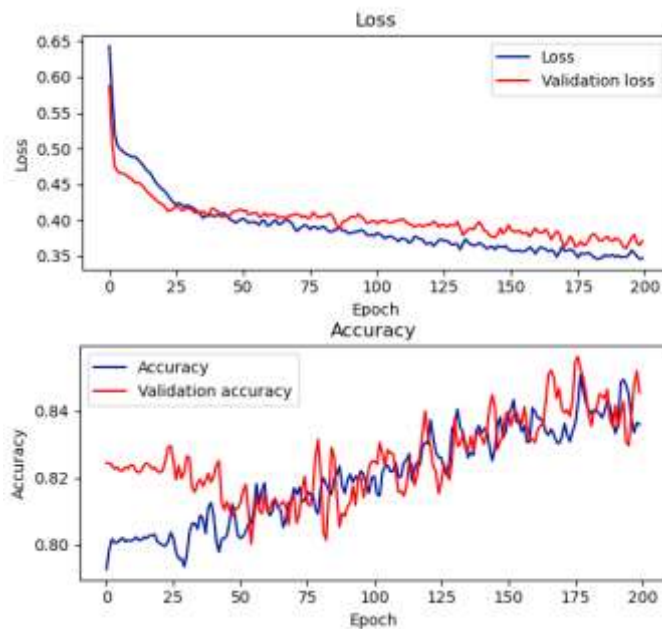


Рисунок А.6 – Процедура навчання/валідації на основі точності та втрат для 40% зрізу сагітальної моделі

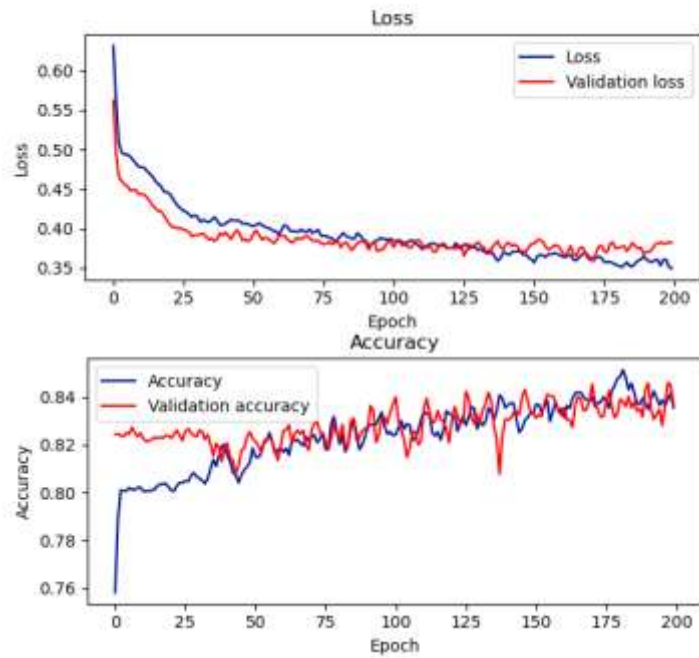


Рисунок А.7 – Процедура навчання/валідації на основі точності та втрат для 60% зрізу аксіальної моделі

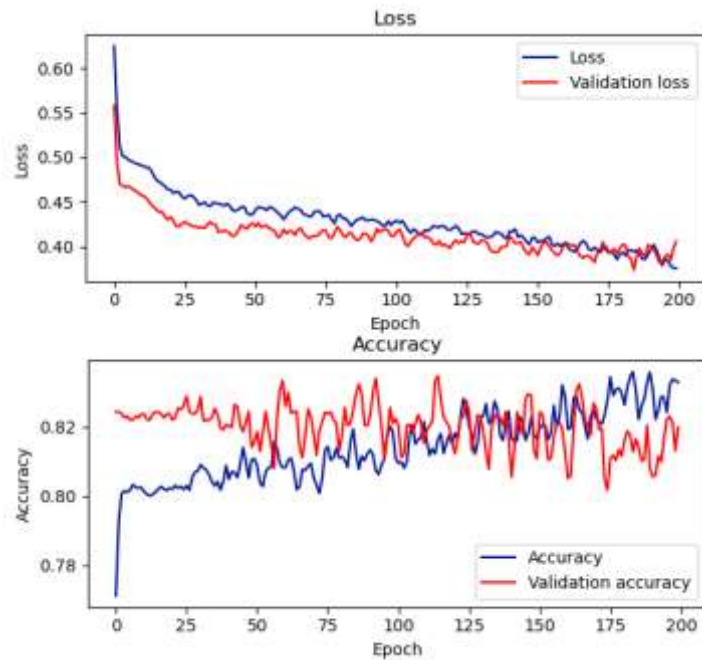


Рисунок А.8 – Процедура навчання/валідації на основі точності та втрат для 60% зрізу корональної моделі

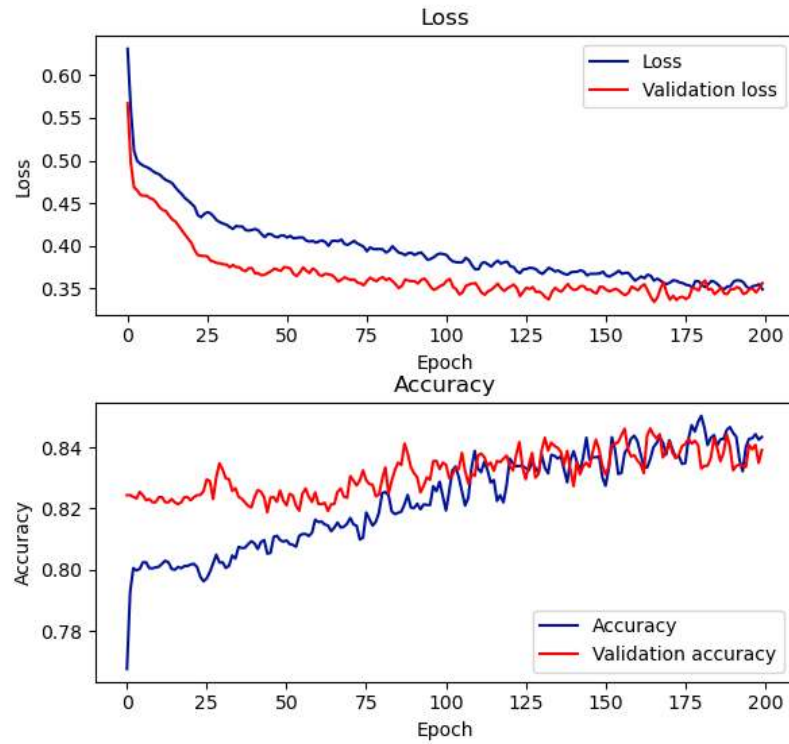


Рисунок А.9 – Процедура навчання/валідації на основі точності та втрат для 60% зрізу сагітальної моделі

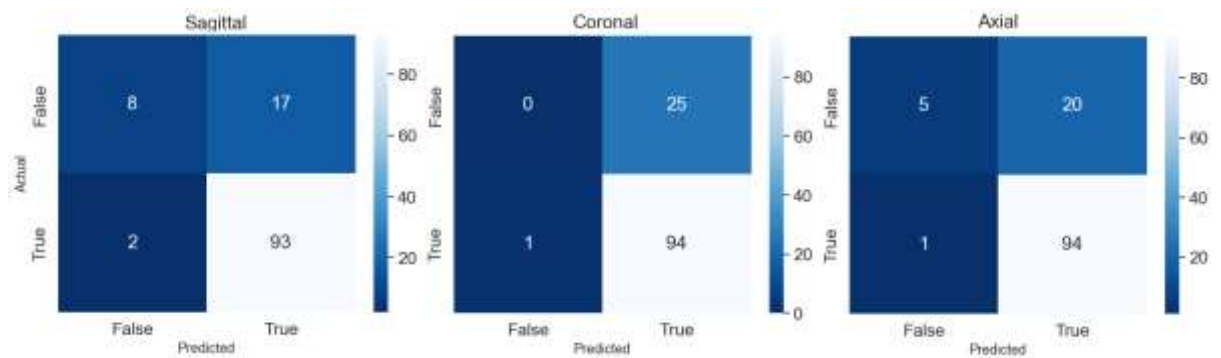


Рисунок А.10 – Матриці помилок, отримані при тестуванні 40% зрізів сагітальної, коронарної та аксіальної проекції відповідно

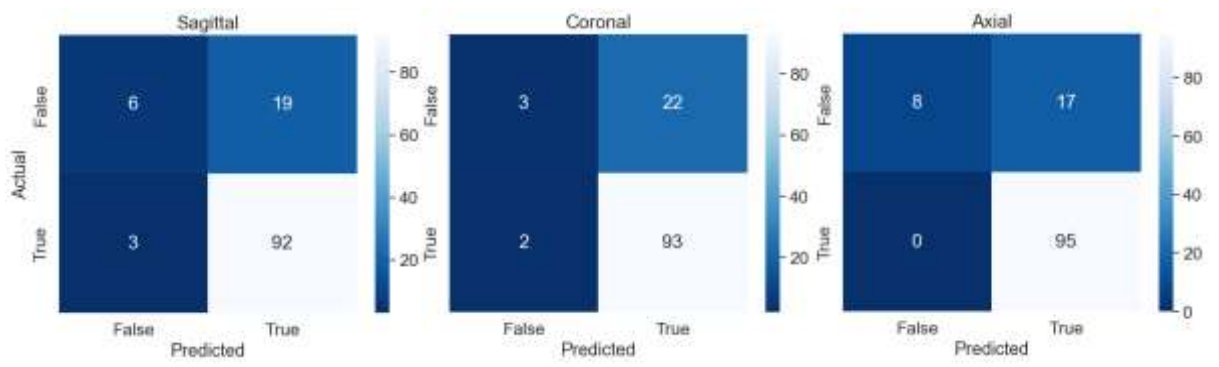


Рисунок А.11 – Матриці помилок, отримані при тестуванні 60% зрізів сагітальної, коронарної та аксіальної проекції відповідно