

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”

ПУКАЧ ПАВЛО ПЕТРОВИЧ

Кваліфікаційна наукова
праця на правах рукопису
УДК 004.652

**Інформаційна технологія аналізу медичних даних на основі
ансамблю методів обчислювального інтелекту**

122 – комп’ютерні науки

**Дисертація на здобуття наукового ступеня
доктора філософії**

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело



/П. П. Пукач/

Науковий керівник
Шаховська Наталія Богданівна,
доктор технічних наук, професор

Львів – 2023

АНОТАЦІЯ

Пукач П. П. Інформаційна технологія аналізу медичних даних на основі ансамблю методів обчислювального інтелекту. – Кваліфікаційна наукова праця на правах рукопису. Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 122 «Комп'ютерні науки». – Національний університет «Львівська політехніка», Львів, 2023.

Зміст анотації.

Дисертаційне дослідження присвячене побудові методів та засобів для покращення точності та інтерпретовності існуючих систем автоматизованого діагностування ушкоджень та аномалій колінного суглоба на основі зображень магнітно-резонансної терапії (МРТ). В перспективі це дозволить експертам-рентгенологам здійснювати більше обстежень пацієнтів за менший проміжок часу, маючи доволі точну систему підтримки прийняття діагностичних рішень.

У першому розділі «Аналіз алгоритмів автоматизованого діагностування зображень МРТ колінних суглобів» здійснено опис процесу автоматизованого діагностування МРТ колінних суглобів, введено основні поняття для аналізу зображень МРТ. Здійснено огляд літератури, описано ключові досягнення у сфері автоматизованого діагностування колінних суглобів за МРТ. Поставлено задачу покращення точності передбачення існуючої моделі глибинного навчання MRNet шляхом дослідження сучасніших архітектур комп'ютерного бачення в якості каркасних мереж виділення ознак.

У другому розділі «Побудова моделі аналізу ефективності каркасних мереж виділення ознак для задачі автоматизованого діагностування МРТ колінного суглоба» розроблено модель для порівняння ефективності мереж глибинного навчання шляхом підстановки різноманітних сучасних архітектур комп'ютерного бачення в якості каркасів виділення ознак.

У третьому розділі «Покращення ефективності ансамблю моделей. Побудова інтерпретаційної мережі» описано підхід підвищення точності передбачення на нових даних для моделі шляхом аугментації вхідних зображень за допомогою випадкових афінних перетворень для запобігання перенавчанню

моделі глибинного навчання. Також розроблено інтерпретаційну модель на основі методу латентного зсуву для пояснення результатів передбачення з метою підвищення прозорості діагностики та простоти кінцевого аналізу лікарями-рентгенологами.

Четвертий розділ «Побудова архітектури та апробація результатів» зосереджується на проблемі зменшення вартості розгортання складних інформаційних систем для автоматизованого діагностування ушкоджень колінного суглоба за МРТ. Побудовано архітектуру інформаційної системи на основі безсерверних обчислень, здійснено порівняння вартості розгортання для традиційної серверної архітектури та запропонованого рішення.

Основні наукові результати дисертації опубліковано в 6 працях, зокрема: три статті – у наукових фахових періодичних виданнях України; одна – у закордонному науковому періодичному виданні (квартиль Q1 НМБД Scopus); дві публікації – у матеріалах міжнародних науково-технічних конференцій.

Ключові слова: магнітно-резонансна томографія, автоматизоване діагностування, згорткові нейронні мережі, класифікація, глибинне навчання, комп'ютерне бачення, пояснювальне глибинне навчання, лямбда-архітектура.

ABSTRACT

Pukach P. P. Information technology of medical data analysis based on an ensemble of computational intelligence methods. – Qualifying scientific work on manuscript rights. Dissertation for obtaining the scientific degree of Doctor of Philosophy in specialty 122 "Computer Science". – Lviv Polytechnic National University, Lviv, 2023.

Abstract content. The dissertation study is devoted to the development of methods and tools for improving the accuracy and interpretability of existing computer-aided knee joint anomaly and injury diagnosis systems based on MRI images. In the future, this will allow radiologists to perform more examinations of patients in a shorter period of time, having a fairly accurate support system for

making diagnostic decisions.

In the first chapter, "Analysis of algorithms for computer-aided diagnostics of MRI images of knee joints", a description of the process of computer-aided diagnostics of knee joint MRI images was carried out, and basic concepts for the analysis of MRI images were introduced. A review of the literature was carried out, and key achievements in the field of automated diagnosis of knee joints by MRI were described. The goal is to improve the prediction accuracy of the existing MRNet deep learning model by researching more modern computer vision architectures as feature extraction backbone networks.

In the second chapter, "Building a model for the analysis of the effectiveness of backbone networks of feature extraction for computer-aided MRI diagnosis of the knee joint," a model was developed to compare the effectiveness of deep learning networks by substituting various modern computer vision architectures as backbones for feature extraction.

In the third chapter, "Improving the efficiency of the ensemble of models. Building the interpretability network" describes an approach to increase the accuracy of predictions on new data for a model by augmenting input images with random affine transformations to prevent overfitting of a deep learning model. An interpretive model based on the latent shift method was also developed to explain the prediction results in order to increase the transparency of the diagnosis and the simplicity of the final analysis by radiologists.

The fourth chapter, "Architecture development and Results Validation", focuses on the problem of optimizing the cost of deploying complex information systems for the automated diagnosis of knee joint injuries by MRI. The architecture of an information system based on serverless computing was built, the cost of deployment for the traditional server architecture and the proposed solution was compared.

The main scientific results of the dissertation were published in 6 works, in particular: three articles – in scientific specialized periodicals of Ukraine; one – in the foreign scientific periodical (quartile Q1 in scientometric database Scopus); 2

publications – in the materials of international scientific and technical conferences.

Keywords: magnetic resonance imaging, computer-aided diagnostics, convolutional neural networks, classification, deep learning, computer vision, explainable deep learning, lambda architecture.

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Статті у виданнях інших держав:

1. N. Shakhovska, P. Pukach. Comparative Analysis of Backbone Networks for Deep Knee MRI Classification Models. *Big Data and Cognitive Computing*. Vol. 6. 2022 (квартиль Q1 у НМБД Scopus) [1].

Статті у фахових виданнях України:

2. П. Пукач. Аналіз каркасних мереж виявлення ознак в моделях глибинного навчання. *Вісник Національного університету "Львівська політехніка", серія "Інформаційні системи та мережі"*. Том 12, № 5. 2022. С. 169-176 [2].
3. П. Пукач. Огляд та аналіз основних каркасних мереж виявлення ознак для класифікації зображень МРТ в моделях глибинного навчання. *Вісник Хмельницького національного університету. Серія: Технічні науки*. Вип. 6 (135), т. 1. С 183-187 [3].
4. Pukach P. Y., Vovk M., Pukach P. P. On one nonlinear mathematical model of blood circulation with the vessel walls reaction within the hereditary theory. *Прикарпатський вісник НТШ. Серія: Число*. 2022. № 17 (64). С. 31-43 [4].

Матеріали конференцій:

5. Shakhovska N., Darmoriz I., Vyklyuk Y., Kryvenchuk Y., Pukach P. Visualization of the epidemics forecasting results. *CEUR Workshop Proceedings*. 2021. Vol. 3038: Proceedings of the 4th International conference on informatics & data-driven medicine IDDM 2021, Valencia, Spain, November 19 - 21, 2021. P. 283-292 [5].
6. Pukach P.Y., Vovk M., Mylyan Y., Bilushchak H., Pukach P.P. On one nonlinear mathematical model of blood circulation with the vessel walls reaction within the hereditary theory. *CEUR Workshop Proceedings*. 2020. Vol. 2753: Proceedings of the 3rd International conference on informatics & data-driven medicine IDDM 2020, Vaxjo, Sweden, November 19-21, 2020. P. 136-141 [6].

ЗМІСТ

ВСТУП.....	10
РОЗДІЛ 1. АНАЛІЗ АЛГОРИТМІВ АВТОМАТИЗОВАНОГО ДІАГНОСТУВАННЯ ЗОБРАЖЕНЬ МРТ КОЛІННИХ СУГЛОБІВ.....	16
1.1. Автоматизоване діагностування в медицині.....	16
1.2. Штучні нейронні мережі.....	20
Моделі штучних нейроелементів.....	20
Багатошарові нейронні мережі.....	22
Глибинне навчання та глибинні нейронні мережі.....	24
Згорткові нейронні мережі.....	25
1.3. Існуючі набори даних та методи автоматизованого діагностування травм колінних суглобів.....	28
1.4. Каркаси виділення ознак.....	38
1.5. Постановка задачі.....	38
Висновки до розділу 1.....	42
РОЗДІЛ 2. ПОБУДОВА МОДЕЛІ АНАЛІЗУ ЕФЕКТИВНОСТІ КАРКАСНИХ МЕРЕЖ ВИДІЛЕННЯ ОЗНАК ДЛЯ ЗАДАЧІ АВТОМАТИЗОВАНОГО ДІАГНОСТУВАННЯ МРТ КОЛІННОГО СУГЛОБА.....	43
2.1. Порівняльний аналіз сучасних архітектур глибинного навчання для комп'ютерного бачення.....	43
AlexNet.....	43
VGG11 та VGG16.....	45
ResNet.....	45
EfficientNet.....	48

2.2. Методологія побудови моделі аналізу ефективності каркасних мереж виділення ознак.	50
2.3. Стратегії оцінювання каркасних мереж виділення ознак	52
ROC-крива	55
F-оцінка	57
Капа Коена	57
2.4. Побудова моделі порівняння каркасних мереж	60
2.5. Підбір ансамблю каркасних мереж на основі результатів моделі порівняння ефективності.	64
Порівняння оцінок ефективності	64
Оцінка одного блоку для кожного типу каркасної мережі.	67
Висновки до розділу 2	73
РОЗДІЛ 3. ПОКРАЩЕННЯ ЕФЕКТИВНОСТІ АНСАМБЛЮ МОДЕЛЕЙ. ПОБУДОВА ІНТЕРПРЕТАЦІЙНОЇ МЕРЕЖІ.	74
3.1. Стратегії запобігання перенаванчання	74
3.2. Інтерпретації рішень у автоматизованому діагностуванні. Існуючі алгоритми інтерпретації.	78
Підхід латентного зсуву	82
3.3. Побудова інтерпретаційної моделі	85
Висновки до розділу 3.	94
РОЗДІЛ 4. ПОБУДОВА АРХІТЕКТУРИ ТА АПРОБАЦІЯ РЕЗУЛЬТАТІВ.	95
4.1. Побудова архітектури інформаційної системи	95
4.2. Моделювання даних у інформаційній системі	100
4.3. Опис бізнес-процесів всередині інформаційної системи	103
Завантаження файлів DICOM у систему	103
Виконання моделі автоматизованого діагностування на обстеженні.	106

4.4. Зменшення вартості розгортання системи. Підрахунок вартості.....	109
Висновки до розділу 4.....	113
Висновки.....	114
Список використаних джерел.....	116
ДОДАТОК А. АКТИ ВПРОВАДЖЕННЯ.....	127

ВСТУП

Актуальність теми

Зі зростанням населення пропорційно зростає і попит на лікарів, зокрема, на рентгенологів опорно-рухового апарату та суміжних спеціальностей. Останні статистичні дослідження прогнозують значний дефіцит експертів-рентгенологів, а також інших спеціалістів у галузі медицини у розвинутих країнах.

Лише в США нестача лікарів у галузі рентгенології може перевищити 35 тисяч до 2034 року згідно з нещодавно опублікованим щорічним аналізом потреби у лікарських спеціальностях, зробленим Асоціацією американських медичних коледжів [7].

У всіх сегментах медичного обслуговування ця цифра може сягнути й 124 тисяч. Як старіння, так і зростання кількості населення є головною рушійною силою цієї нестачі. У віковому сегменті 65 і вище за прогнозами протягом наступного десятиліття планований дефіцит кваліфікованих медичних спеціалістів зросте більше, ніж на 42%.

У звіті не вказується конкретна цифра суто для рентгенологів. Дані об'єднують цю спеціальність з неврологами, працівниками невідкладної медицини, анестезіологами та іншими. Однак є достатньо підстав вважати частку рентгенологів у звіті суттєвою. За прогнозами у цьому секторі буде нестача від 10300 до 35600 лікарів до 2034 року. Усі вказані вище результати (статистичні дані) є частиною згаданого аналізу, опублікованого 11 червня 2021 року.

Зважаючи на статистичні прогнози цього звіту, підходимо до потреби в надійних системах автоматизованого діагностування, які могли б пришвидшити або спростити процес виявлення патологічних станів рентгенологами під час аналізу знімків магнітно-резонансної томографії пацієнтів. Основним застосуванням таких систем має бути зменшення робочого навантаження на лікарів-спеціалістів відповідного профілю, яке на даний момент постійно зростає.

Розглядаючи проблему нестачі спеціалістів-рентгенологів у контексті задачі діагностування МРТ колінних суглобів, необхідно зазначити, що існуючі дотепер системи для автоматизованого діагностування на базі глибинних нейронних мереж можна вважати досить точними. Однак об'єктивно треба також відзначити той факт, що в якості підґрунтя вони опираються на дещо застарілі мережі комп'ютерного бачення, і порівняльний аналіз використання сучасніших каркасних мереж до цього моменту не проводився.

Щодо питання інтерпретації результатів автоматизованого діагностування, також можна констатувати відсутність актуальних та системних досліджень для зображень МРТ колінного суглоба, в першу чергу через суттєву складність побудови моделей для інтерпретації. Причиною цього є багатовимірність вхідних даних.

Також при побудові прикладних інформаційних систем для автоматизованого діагностування з використанням вищезгаданих методів постає проблема вартості розгортання таких систем. Вона обумовлена тим, що прикладні інформаційні системи вимагають значної кількості обчислювальних ресурсів як для постійного перетренування моделей діагностування на більш свіжих вибірках зображень, так й, власне, для хостингу та налагодження моделей для виконання передбачень.

З огляду на все зазначене вище, задача дисертаційної роботи розроблення методів та засобів інтерпретованого аналізу зображень МРТ є актуальною.

Зв'язок роботи з науковими програмами, планами і темами.

Дисертація виконувалася відповідно до пріоритетних напрямків науково-дослідних робіт Національного університету "Львівська політехніка" та відповідно до координаційних планів Міністерства освіти і науки України. Зокрема, наукові дослідження виконувалися в рамках держбюджетної наукової теми кафедри систем штучного інтелекту «Інформаційна технологія формування психофізичного портрету в умовах стресових ситуацій» (№ держ. реєстру 0119U002257).

Метою дисертаційної роботи є підвищення точності та інтерпретованості аналізу медичних даних, зокрема, серій знімків магнітно-резонансної томографії, на основі ансамблю методів машинного навчання (МН).

Для досягнення поставленої мети у роботі постає необхідність розв'язати наступні **задачі**:

- 1) провести аналіз існуючих методів машинного навчання для аналізу медичних зображень та автоматизованого діагностування на прикладі аналізу МРТ колінного суглоба;
- 2) розробити модель аналізу ефективності каркасних мереж виділення ознак для задачі автоматизованого діагностування МРТ колінного суглоба;
- 3) розробити метод підвищення ефективності ансамблю моделей МН на основі аугментації даних для усунення перенавчання;
- 4) побудувати інтерпретаційну модель діагностування пошкоджень колінного суглоба на основі аналізу серії знімків магнітно-резонансної томографії;
- 5) побудувати архітектуру інформаційної системи діагностування пошкоджень колінного суглоба на основі аналізу серії знімків МРТ із застосуванням безсерверної архітектури для оптимізації вартості розгортання системи; апробувати результати.

Об'єкт дослідження – процеси аналізу серії медичних зображень.

Предмет дослідження – методи, моделі та інформаційні технології аналізу знімків МРТ колінного суглобу, їх автоматичного діагностування та інтерпретації.

Методи дослідження: архітектури комп'ютерного бачення в якості каркасних мереж виділення ознак, а саме: VGG11, VGG16, ResNet, EfficientNet; набори даних KneeMRI та MRNet, методи комп'ютерних наук, що на основі яких побудовані системи автоматизованого діагностування — глибинні нейронні мережі, метод опорних векторів, моделі випадкового лісу; метрики точності (влучність, повнота, чутливість, специфічність, показник каппа Коена,

F1), методи комп'ютерних наук для інтерпретації результатів мереж глибинного навчання – карти помітності, метод зсуву латентної змінної.

Наукова новизна одержаних результатів.

Вперше розроблено модель порівняння ефективності каркасних мереж виділення ознак для задач автоматизованого діагностування МРТ. Вказана модель розроблена з метою виділення найбільш продуктивних каркасних мереж шляхом використання шару виділення ознак як варіативного та різних вхідних шарів класифікаторів. При цьому застосовано метрики капа Коена в рамках моделі надійності різних оцінювачів.

Вперше застосовано метод латентного зсуву до автоматизованого діагностування МРТ коліна, що дало змогу інтерпретувати результати класифікації моделі та забезпечити прозорість діагностичних рішень. Розроблено інтерпретаційну модель діагностування пошкоджень колінного за знімками магнітно-резонансної томографії шляхом побудови карти помітності на основі латентної змінної із заданим значенням зсуву, забезпечуючи цілісність зображень в розрізі. Також розроблена модель дала можливість визначити відхилення точності передбачення при процесах класифікації та пояснення.

Удосконалено метод запобігання перенавчання моделі на основі аугментації наборів даних шляхом застосування послідовності поворотів та випадкових афінних перетворень, що дає змогу застосовувати цей метод до серійних зображень.

Отримав подальший розвиток метод діагностування для МРТ колінного суглоба на основі розроблення моделі на базі ансамблю різнотипних каркасних мереж, що дало змогу підвищити точність прийнятих рішень

Практичне значення одержаних результатів.

Розроблено архітектуру інформаційної системи діагностування на основі знімків МРТ з використанням підходу безсерверних обчислень, в результаті оптимізовано витрати на розгортання системи у 4,7 разів.

Проведено порівняльний аналіз для набору моделей класифікації для побудова алгоритму класифікування знімків МРТ. Порівнювалися моделі, засновані на оригінальній архітектурі MRNet, з різними магістральними мережами як екстракторами функцій для блоків MRNet, включаючи Alexnet, VGG11, VGG16, Resnet і Efficientnet, а також використовуючи метрику Капа Коена для порівняння ефективності. MRNet-VGG16 показує найкращі результати для виявлення розривів ACL. Для виявлення аномалій форми коліна, MRNet-VGG16 домінує над MRNet-Resnet для достовірності від 0,5 до 0,75 та над MRNet-VGG11 для достовірності понад 0,8.

Розроблено методологію інтерпретування для задач автоматизованого діагностування розриву передньої хрестоподібної зв'язки (ACL) з використанням підходу зсуву латентної змінної шляхом параметризованого додавання контрафактних зображень всередину серії. Підібрано оптимальний параметр $F=3$ для такого додавання, що забезпечує найменшу усереднену різницю ймовірностей класифікації у 0.096.

Запропоновані методи і моделі впроваджені у навчальний процес Національного університету «Львівська політехніка» при викладанні дисципліни «Аналіз великих даних» (підтверджено актом впровадження).

Також результати дисертаційної роботи впроваджено в КНП ЛОР «Львівський онкологічний регіональний лікувально-діагностичний центр» (підтверджено актом впровадження).

Отримані у роботі результати використано та впроваджено в ДБ «Інформаційна технологія формування психофізичного портрету в умовах стресових ситуацій» (№ держ. реєстру 0119U002257).

Особистий внесок здобувача. Основні положення та результати дисертаційної роботи одержані автором самостійно. Особисто здобувачеві належать наступні наукові результати: здійснено порівняльний аналіз для набору моделей класифікації для побудова алгоритму класифікування знімків МРТ та побудовано архітектуру глибокої нейронної мережі [1], розроблено нелінійну математичну модель кровообігу з реакцією стінок судин у рамках

спадкової теорії [4], розроблено метод візуалізації на основі латентного зсуву [5], розроблена математична модель задачі медичного діагностування та архітектура динамічної нейронної мережі [6].

Апробація результатів дисертації. Результати дисертаційної роботи доповідались на конференціях: 3-я Міжнародна конференція з інформатики та медицини, керованої даними IDDM 2020, 4-та Міжнародна конференція з інформатики та медицини, керованої даними IDDM 2021. Також результати доповідались на семінарах кафедри систем штучного інтелекту «Національного університету «Львівська політехніка».

Структура та обсяг дисертації. Дисертація складається із вступу, чотирьох розділів, висновків, списку використаних джерел із 94 найменування. Повний обсяг дисертації складає 129 сторінок, основний зміст викладено на 106 сторінках, де наведено 40 рисунків та 7 таблиць, 3 додатки.

РОЗДІЛ 1. АНАЛІЗ АЛГОРИТМІВ АВТОМАТИЗОВАНОГО ДІАГНОСТУВАННЯ ЗОБРАЖЕНЬ МРТ КОЛІННИХ СУГЛОБІВ

У розділі здійснено опис процесу автоматизованого діагностування МРТ колінних суглобів, введено основні поняття для аналізу зображень МРТ. Здійснено огляд літератури, описано ключові досягнення у сфері автоматизованого діагностування колінних суглобів за МРТ. Поставлено задачу покращення точності передбачення існуючої моделі глибинного навчання MRNet шляхом дослідження більш сучасних архітектур комп'ютерного бачення в якості каркасних мереж виділення ознак.

Результати розділу опубліковано у працях автора [2, 4].

1.1. Автоматизоване діагностування в медицині

Автоматизоване діагностування (англ. Computer-aided Diagnosis, CADx) у медицині — це процес використання систем, які допомагають медичним працівникам здійснювати інтерпретацію медичних зображень чи даних. Сучасні методи рентгенографії, МРТ, ендоскопії та ультразвукової діагностики дають досить велику кількість інформації, котру лікар-рентгенолог або інший медичний працівник повинен проаналізувати за короткий час, а також на основі цього сформулювати рішення про встановлення діагнозу для пацієнта.

Системи CADx обробляють вхідні цифрові зображення або відео типових ракурсів частин тіла з метою виділення ділянок, котрі є потенційно значущими для завдання встановлення діагнозу. Серед таких ділянок можуть бути аномалії в частині тіла, деформації, розриви, або будь-які інші індикатори можливих захворювань. Ці дані пропонуються як вхідні для медичного працівника, котрий приймає остаточне рішення щодо діагнозу.

Автоматизоване діагностування як процес є результатом великої кількості наукових здобутків на стику галузей медицини та комп'ютерних наук. Оскільки багато систем автоматизованого діагностування намагаються імітувати процес

прийняття діагностичних рішень медичними експертами, їх можна розглядати як експертні системи в медицині.

Крім того, системи CADx у медицині мають можливість обробки клінічних даних, які можуть бути складними та великими за обсягом, надаючи широкий спектр підтримки медичним працівникам для прийняття рішень.

Оскільки вказані системи дають можливість отримання нових знань з даних та використання цих знань для покращення своєї діагностичної ефективності з часом, то такі системи потрапляють у класифікацію інтелектуальних інформаційних систем (ІС) [8].

CADx фундаментально базується на розпізнаванні дуже складних зразків зображень. Рентгенівські або інші типи зображень скануються на наявність структур, що представляють інтерес. Зазвичай, для оптимальної роботи алгоритму автоматизованого діагностування потрібно від кількох тисяч зображень. Дані зображення копіюються на сервер у форматі DICOM, проводиться їх підготовка та аналіз. Коротко зупинимося на таких етапах.

1. **Підготовка зображень.** Цей етап включає перелічені нижче кроки.
 - a. Виявлення та знешкодження артефактів (помилки) у зображеннях.
 - b. Зменшення шумів у зображенні.
 - c. Вирівнювання, або нормалізація яскравості та контрасту зображення.
 - d. Застосування фільтрів.
2. **Сегментація.** Включає у себе:
 - a. виділення структур (частин тіла, або органів) всередині зображення;
 - b. співставлення структур з сховищем анатомічних даних;
 - c. вибірку значень яскравості з областей інтересу [9].
3. **Аналіз областей інтересу** для виявлення наступних перелічених нижче характеристик.
 - a. Компактність.
 - b. Форма, розмір та розташування.

- c. Зв'язок із найближчими структурами.
- d. Середнє значення яскравості в області інтересу.
- e. Відношення яскравості всередині області до значення на межі області інтересу.

4. Класифікація та прогнозування. Після аналізу кожної області інтересу визначається ймовірність позитивного класу для цієї області. Далі зважуються ймовірності для кожної із областей для кінцевої ймовірності позитивного класу для всього зображення. Серед методів класифікації у цьому випадку, найбільш поширеними є наступні:

- a. метод К-найближчих сусідів;
- b. метод мінімальної відстані;
- c. наївні баєсові класифікатори;
- d. штучні нейронні мережі [10];
- e. метод опорних векторів [11].

В якості джерела вихідних зображень у цій роботі було вибрано магнітно-резонансну томографію (МРТ). Вона є одним з найефективніших методів діагностування травм і аномалій колінних суглобів, що існують донині. МРТ є найкращим способом візуалізації коліна для виявлення можливих патологій з метою скерування пацієнта на лікування [12].

Даний вид діагностування дозволяє робити послідовності зображень, які називаються «зрізами» в трьох різних розрізах – аксіальному, корональному і сагітальному (показані на рисунку 1.1), створюючи повноцінне представлення внутрішньої будови будь-якої частини тіла. Для фахівців у сфері рентгенології, в тому числі рентгенології опорно-рухового апарату, магнітно-резонансна томографія виступає незамінним інструментом для виявлення кінцевого діагнозу та підбору методів лікування.

Серед всеможливих методів класифікації для цієї роботи було вибрано саме штучні нейронні мережі, оскільки станом на сьогодні саме на них зосереджено найбільше уваги серед представників наукового співтовариства.

На перетині галузей штучного інтелекту (ШІ) та рентгенології вже зараз

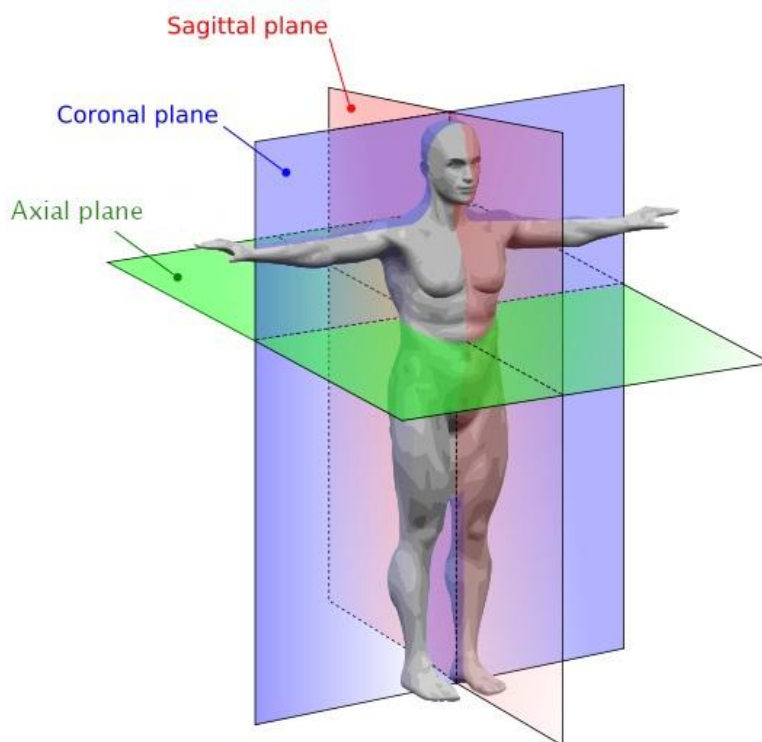


Рисунок 1.1. Розрізи МРТ - аксіальний (зелений), корональний (блакитний) та сагітальний (червоний) [12].

використовуються різноманітні методи комп'ютерних наук, для підтримки прийняття рішень медичними працівниками у завданнях діагностування.

Штучні нейронні мережі все більше використовуються для розв'язання різноманітних задач, в тому числі й для автоматизованого діагностування.

Методи штучного інтелекту можна застосовувати безпосередньо до багатьох різних завдань кістково-м'язової рентгенології, включаючи реконструкцію зображень, створення синтетичних зображень, сегментацію тканин, а також діагностування та виявлення характеристик захворювань опорно-рухового апарату на рентгенограмах, УЗД, КТ та зображеннях МРТ. Серед популярних технологій на сьогодні можна виділити алгоритми глибокого навчання (англ. *deep learning algorithms*), котрі широко застосовуються для розпізнавання та класифікації зображень.

З плином часу прикладні застосування даних методів демонструють покращення результативності. У зв'язку із цим інтерес до них постійно зростає як серед науковців у галузі штучного інтелекту, які можуть досліджувати і

застосовувати все більш сучасні методи машинного навчання (наприклад, штучні нейронні мережі та глибинне навчання), так і від лікарів-рентгенологів, для яких основною цінністю подібних технологій є спрощення діагностування патологій у пацієнтів, і можливість першочергово сфокусуватися на особливих та складних клінічних випадках [13].

У разі їх довершення для практичного застосування, такі системи також покликані автоматизувати діагностування типових патологічних станів для своєї сфери використання. Звичайно, при цьому ефективність діагностування повинна бути з рівнем довіри до результату передбачення, що є аналогічним до комісії експертів-рентгенологів.

1.2. Штучні нейронні мережі

Штучні нейронні мережі (ШНМ) – це сімейство математичних моделей у машинному навчанні, які побудовані за принципами подання й обробки інформації у біологічних нейронних мережах – мережах нервових клітин у живому організмові [14].

Моделі штучних нейроелементів.

Штучний нейрон – це примітивний обчислювальний пристрій (або його модель), що має кілька входів і один вихід, і є основним обчислювальним елементом НМ.

Схему штучного нейрона зображено на рис. 1.2.

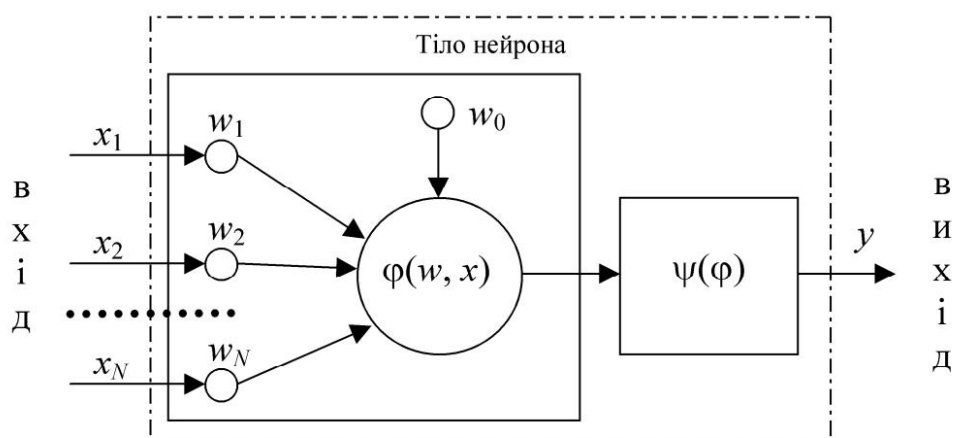


Рисунок 1.2. Штучний нейрон [14].

На вхід одношарового нейрона надходить вхідний вектор – набір вхідних сигналів $x = \{x_j\}, j = 1, 2, \dots, N$, де N – кількість входів. Інколи до вхідного вектора додають також значення x_0 , що є фіксованим входом штучного нейрона, і корелює із значенням зсуву w_0 .

Кожний вхідний сигнал x_j зважується (масштабується) за допомогою відповідної йому ваги зв'язку (вагового коефіцієнта) w_j , яка моделює перетворення сигналу у синапсі (міжнейронному контакті) [14].

Вага є параметром мережі, що відповідає за перетворення вхідних даних в прихованих шарах мережі. Якщо моделювати нейронну мережу як ряд вузлів чи нейронів, то для кожного такого вузла можна виділити набір вхідних даних, вагу, а також та значення зсуву. Саме тому вагу позначають як вектор $w = \{w_j\}$, елементи якого відповідають вагам кожного вхідного сигналу.

Коли вхід надходить у вузол, він множиться на значення ваги, а отриманий результат або спостерігається або передається на наступний рівень нейронної мережі. Часто ваги нейронної мережі містяться в прихованих шарах мережі.

Дискримінантна (вагова, постсинаптична) функція (англ. summing junction) поєднує вхідний сигнал із вектором ваг нейрона та значенням зсуву. Як правило, вона описується як зважена сума:

$$\phi(w, x) = \sum_{j=1}^N w_j x_j + w_0.$$

Математичну модель функціонування штучного нейрона можна описати наступним чином:

$$y = \psi(\phi(w, x))$$

де x – вектор вхідних сигналів; y – значення на виході нейрона; ψ – функція активації; $w = \{w_j\}$ – вектор, що містить значення вагових коефіцієнтів w_j , w_0 – значення зсуву, ϕ – дискримінантна функція [14].

Функція активації $\psi(x)$, де x – аргумент функції активації, повинна бути обмеженою (на інтервалі значень x приймати значення y^0 , на інтервалі x –

приймати значення y^1 , на інтервалі $x \in (a, b)$ приймати значення $y: y^0 \leq y \leq y^1$, де y^0 та y^1 – деякі постійні мінімальне та максимальне значення, a та b – деякі константи, причому: $a \leq b$) і монотонною (на інтервалі $x \in (a, b)$, $\Delta\psi(x) = \psi(x + \Delta x) - \psi(x)$), а також не повинна змінювати знак при $\Delta x > 0$ та $x, x + \Delta x \in (a, b)$ [14].

Функція активації визначає вихід вузла – штучного нейроелемента з урахуванням його входу або набору входів. Нелінійні функції активації дозволяють мережам розв’язувати нетривіальні задачі, використовуючи порівняно невелику кількість нейронів [15].

Багатошарові нейронні мережі.

Нейронну мережу можна задавати за допомогою графа $G = (V, E)$, де V – це множина вузлів та $E \subseteq V \times V$ – множина всеможливих ребер-зв’язків між вузлами мережі на одному шарі [16].

Щоб представити системи, які складаються з таких графових мереж на кількох рівнях або з декількома типами ребер, вводяться структури, які мають шари на додачу до вузлів і ребер. Існує декілька типів розділення на шари для таких типів мереж, використовуючи кольорові графи.

Приклад задання багатошарової нейронної мережі за допомогою графа із кольоровими вершинами подано нижче:

$$G_V = (V, E, C),$$

$$V_\alpha \in V \times C, E_{\alpha\beta} \in V_\alpha \times V_\beta,$$

$$\alpha, \beta \in C.$$

Множина C позначає кількість шарів (кольорів вершин) графа. Цей тип БНМ описує відмінність між типами вузлів для багатошарової нейронної мережі, і різні типи вузлів позначені відповідними кольорами.

Іншим процесом може бути розділення за типом зв’язків між вузлами. Їх теж можна описати за допомогою кольорового графа, в якому з множини кольорів будуть «фарбуватися» ребра [17]:

$$G_E = (V, E, C),$$

$$\{G_\alpha\}_{\alpha=1}^b = \{(V_\alpha, E_\alpha)\}_{\alpha=1}^b,$$

$$\forall \alpha, \beta \in C : V_\alpha = V_\beta,$$

$$E_\alpha \in V_\alpha \times V_\alpha \times C.$$

Граф БНМ вирівнюється по вузлах, тобто один і той же вузол може належати до різних шарів одночасно. Основна увага приділяється динаміці мережі та взаємодії між її компонентами. Такі типи опису нейромереж дуже ефективно використовувати під час вивчення різних аспектів однієї і тієї ж системи. Два вузли мережі можуть з'єднуватися лише за допомогою одного ребра кожного кольору. Іншими словами, між двома штучними нейронами у БНМ не може бути різнотипової передачі даних. Схеми подібних мереж показані на рисунку 1.3.

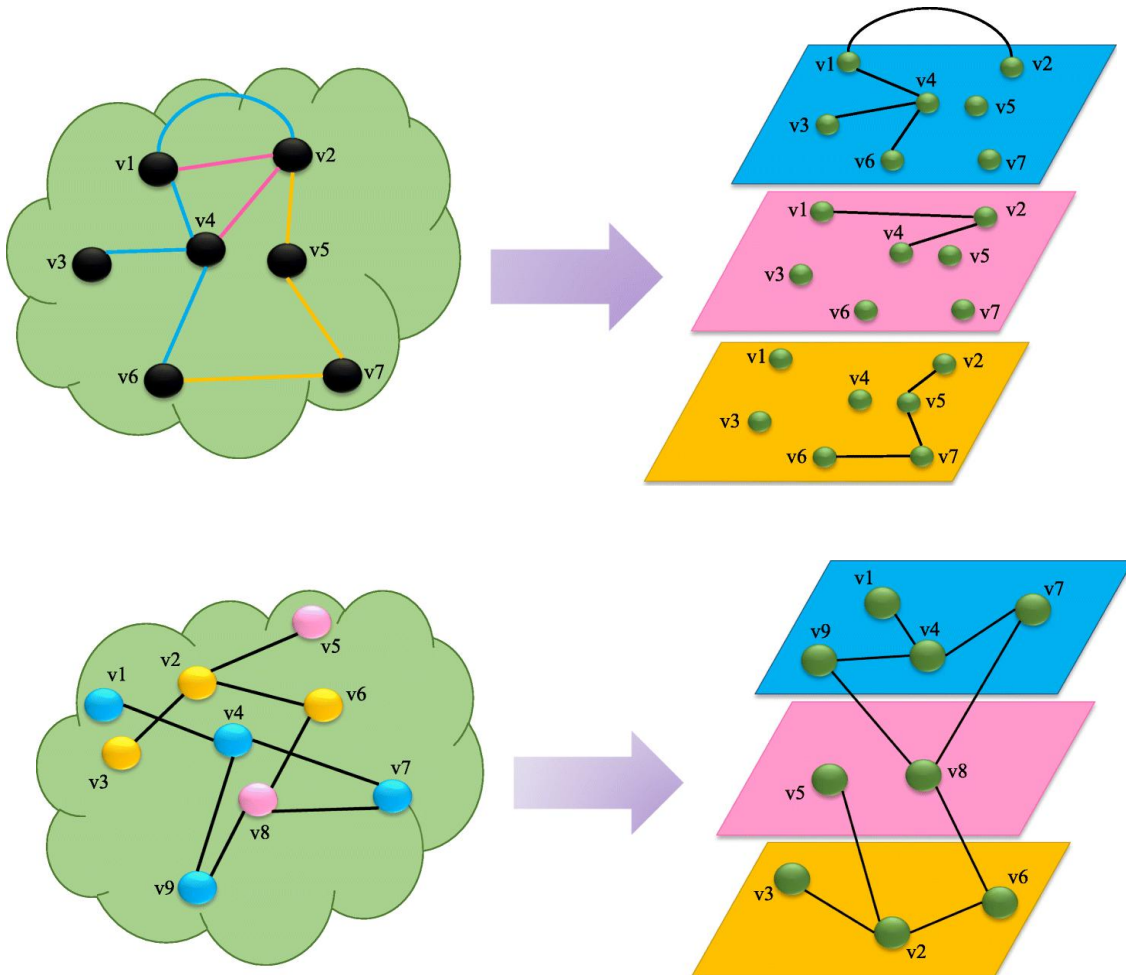


Рисунок 1.3. Схема БНМ. Опис за допомогою графів із кольоровими ребрами(згори), та графів із кольоровими вершинами(знизу) [17].

Глибинне навчання та глибинні нейронні мережі.

Глибинне навчання — це клас алгоритмів машинного навчання, які використовують багато абстрактних шарів для поступового виділення ознак вищого рівня з необроблених вхідних даних. До прикладу при обробці зображень нижчі шари можуть ідентифікувати краї зображення, в той час як вищі шари можуть ідентифікувати ознаки, зрозумілі для людини, такі як цифри, літери чи обличчя [18].

Слово «глибинний» «глибинному навчанні» означає глибину шарів, через які відбувається перетворення даних. Якщо описати точніше, то системи на базі глибинного навчання мають значну довжину ланцюга перетворень даних від входу до виходу. Так краще описуються потенційні причинно-наслідкові зв'язки між входом і виходом системи.

Для прямої нейронної мережі довжина ланцюга перетворення даних дорівнює розміру мережі та є кількістю прихованих шарів плюс один (оскільки вихідний рівень також параметризований). Для рекурентних нейронних мереж, у яких сигнал може поширюватися через один шар неодноразово, така довжина ланцюга є потенційно необмеженою [19].

У центрі глибинного навчання лежать глибинні нейронні мережі та методи їхньої побудови. Глибинною нейронною мережею називають таку штучну нейронну мережу, яка характеризується багат шаровими процесами обробки вхідних даних на вихідні. У глибинній мережі здійснюється агрегування ознак, таким чином підвищуючи рівень абстракції даних [14].

Глибинні нейронні мережі широко застосовуються у медичній сфері для різноманітних задач аналізу зображень [20]. Вперше такий підхід було описано у 1991 році у роботі [21], де автори застосували навчальну мережу для виявлення меж клітини мікрофотографій ендотелію рогівки людини, отриманих за допомогою дзеркальної мікроскопії.

Іншою сферою застосування глибинних нейронних мереж у медичній сфері є аналіз медичних текстів. Він розглядається як в парі з опрацюванням медичних зображень (до прикладу, аналіз рукописної інформації за

зображеннями медичних карток), так і структурування, або екстракція знань із цифрових текстів у більш сучасних медичних інформаційних системах. Прикладами подібних робіт є [22] та [23].

Загалом, прикладів різноманітного застосування методів та технологій на базі глибинного навчання у медицині є чимало, частина із них є наведена на наступних сторінках цього дисертаційного дослідження.

Через свій чіткий поділ на спеціалізовані шари, а також можливість будувати ієрархію таких шарів, глибинні нейронні мережі мають здатність моделювати значно складніші залежності, аніж їх простіші «мілкі» аналоги. На додачу до цього, глибинні мережі мають більшу пропускну здатність для вхідних даних, забезпечену збільшенням кількості штучних нейронів та зв'язків між ними. Наостанок, глибинна архітектура мереж є природно більш зрозумілою для людей, адже кожен шар мережі має своє призначення, і, як правило, фокусується на одній єдиній задачі – через це обробка даних стає більш зрозумілою.

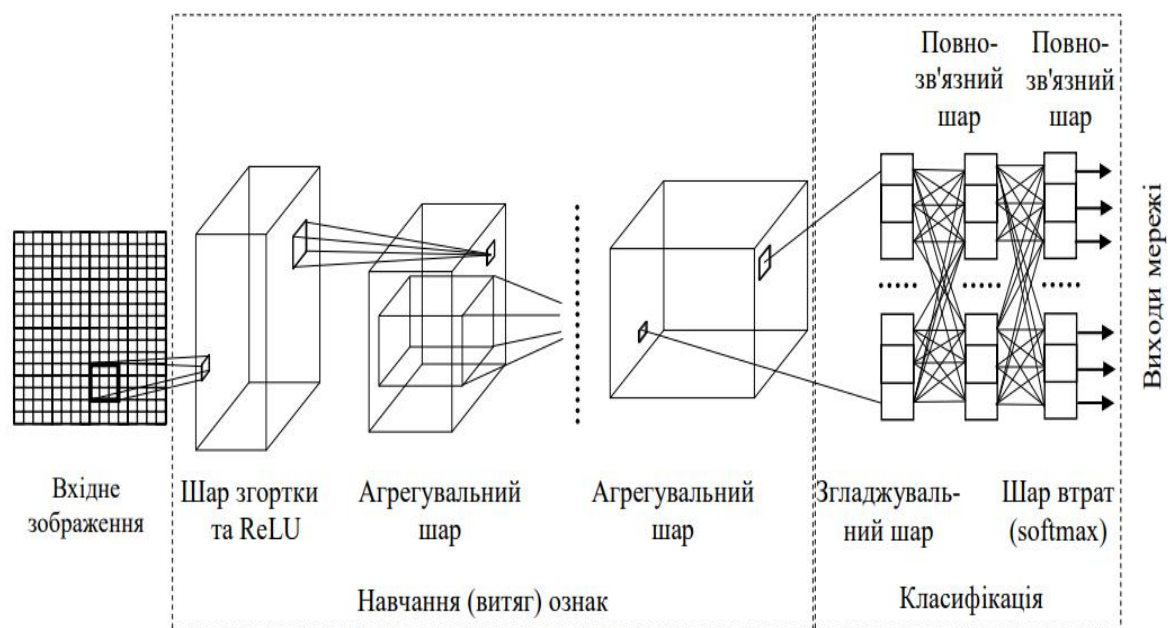
Хорошим прикладом для порівняння спеціалізованості шарів глибинних мереж, та їх легшого розуміння людьми можна вважати модель конвеєру. Якщо уявити вхідні дані мережі за частинами, або каркасом кінцевого виробничого продукту на конвеєрі, то в процесі роботи системи, дані рухаються від одного шару до іншого, трансформуючись або доповнюючись, за допомогою відповідних вузькоспеціалізованих шарів-«фахівців».

Згорткові нейронні мережі.

У глибинному навчанні згорткова нейронна мережа (англ. Convolutional Neural Network, CNN) — це клас штучних нейронних мереж, які найчастіше застосовуються для аналізу зображень [24]. Згорткові нейронні мережі можна вважати глибинними нейронними мережами прямого поширення. Фактично це спеціалізовані БНМ, адаптовані та оптимізовані під задачі опрацювання зображень [14].

Поява нейронних мереж за типом згорткових історично зумовлена будовою оптичних рецепторів у тварин. Нейрони, що знаходяться на сітківці,

реагують на подразники виключно в своєму рецептивному полі. Враховуючи, що такі поля перетинаються для різних нейронів, в результаті досягається повнота сприйняття вхідного зображення. Аналогічно, операція згортки забезпечує своєрідні «рецептивні поля» для кожного штучного нейроелемента,



які перекриваються між собою. Загальна схема згорткової нейронної мережі показана на рисунку (рис. 1.4) [14].

Рисунок 1.4. Згорткова нейронна мережа [14].

Між входом та виходом даних у згортковій нейронній мережі також присутня деяка кількість прихованих шарів. Зазвичай, ці приховані шари є спеціалізованими, і, в свою чергу, вони складаються з шарів згортки, агрегування, повністю з'єднаних шарів, а також нормалізаційних шарів.

Згортковий шар є основним компонентом CNN. Параметрами цього шару є набір навчальних фільтрів (або ядер), які мають відносно невелике рецептивне поле, проте поширюються на всю глибину вхідного об'єму зображення. Під час прямого проходу даних до кожного фільтра застосовується згортка по ширині та висоті вхідного об'єму зображення, обчислюючи скалярний добуток між елементами фільтра та входом шару, таким чином створюючи двовимірну карту активації цього фільтра. У результаті мережа вивчає фільтри, які

активуються, коли виявляється певний тип ознаки у деякій просторовій позиції у об'ємі вхідних даних [25].

Математичний запис опрацювання вхідних даних у згортковому шарі ЗНМ виглядає наступним чином.

$$z^l = h^{l-1} * W^l,$$

де z^l – постсинаптичний (передактиваційний) рівень збудження нейронів l -го шару нейронної мережі, h^{l-1} – активаційний рівень нейронів $(l-1)$ -го шару мережі, W^l – налагоджуваний параметр l -го шару мережі, $*$ – дискретний оператор згортки [14].

Шар об'єднання (агрегування, англ. pooling layer) забезпечує підхід до зменшення роздільної здатності векторів ознак, шляхом узагальнення їх присутності у певних ділянках чи регіонах зображення. Серед найбільш поширених методів агрегування можна навести усереднене агрегування, та максимізаційне агрегування. Ці види агрегування ознак підсумовують відповідно середню присутність певної ознаки на всьому зображенні, або ж найбільш помітну присутність певної ознаки.

Через властивості просторової організації, локальної з'єднаності, спільних параметрів та шарів об'єднання, згорткові нейронні мережі дуже активно використовуються у задачах комп'ютерного бачення. На подібних задачах вони досягають хорошого узагальнення для рішень класифікації на раніше невідомих даних. Також, ЗНМ дозволяють знизити кількість обчислювальних ресурсів, необхідних для навчання мережі, шляхом спільного використання ваг.

.Отриманий набір даних було названо KneeMRI. Для формування цього набору застосовувався метод придушення жиру за протонною щільністю. Із записів обстежень на сканері MPT Siemens Avanto 1.5T у Клінічному лікарняному центрі Рієки, було ретроспективно зібрано значну кількість зображень, датованих від 2006 до 2014 рр. Сам набір KneeMRI містить 917 12-бітних чорно-білих зображень лівого або правого колінного суглоба. Приклади зображень з даного набору даних представлено на рисунку 1.5.

1.3. Існуючі набори даних та методи автоматизованого діагностування травм колінних суглобів

Одна з перших спроб використання методів штучного інтелекту для задачі автоматизованого діагностування колінного суглоба за МРТ, представлена у роботі [26]. В результаті було сформовано загальнодоступний набір даних зображень обстежень МРТ колінного суглоба, зібрані на основі значної кількості клінічних випадків.

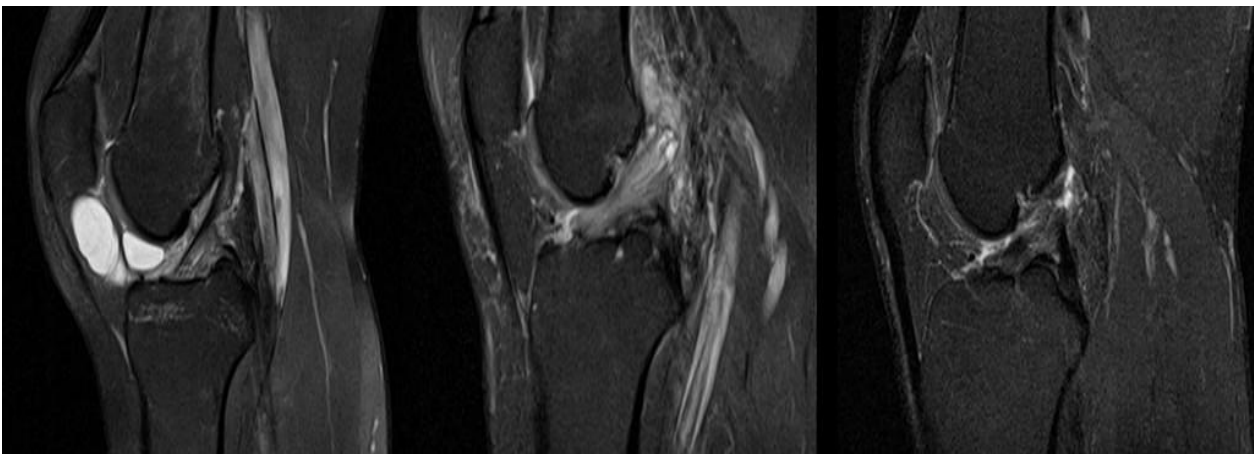


Рисунок 1.5. Приклади зображень коронального розрізу МРТ з набору даних KneeMRI

За допомогою подвійного сліпого методу, кожному зображенню із набору даних було присвоєно відповідне число, що означало загальний стан передньої хрестоподібної зв'язки: (1) здорова, (2) частково ушкоджена або (3) повністю розірвана. Більш широку прямокутну область інтересу (ROI) було виділено вручну з оригінальних зображень, а також анотовано [26]. Цей набір даних було створено, щоб надати простий спосіб роботи з даними науковцям, які займаються прикладним застосуванням методів штучного інтелекту, зокрема, глибинним навчанням та комп'ютерним баченням.

Використовуючи опублікований набір даних KneeMRI, автори створили свій алгоритм класифікації розриву хрестоподібної зв'язки, на основі декількох методів машинного навчання – методу опорних векторів, та моделі випадкового лісу.

Модель випадкового лісу [27] — це сукупність дерев прийняття рішень, які можна використовувати для моделювання як задач класифікації, так і регресії.

Кожне дерево рішень, що утворює ансамбль, вивчається окремо з підмножини екземплярів, випадково відібраних із усього набору. Під час вирощування дерева, розбиття кожного вузла визначається шляхом спостереження лише за випадково вибраною підмножиною доступних функцій і вибору тієї, яка дає найкраще розбиття. Ця комбінація пакетування та випадкового вибору підмножини ознак забезпечує чудові властивості узагальнення моделі, оскільки кількість частин, що слабо навчаються (необрізаних окремих дерев), стає великою. Ілюстрацію моделі випадкового лісу можна побачити на рис 1.6.

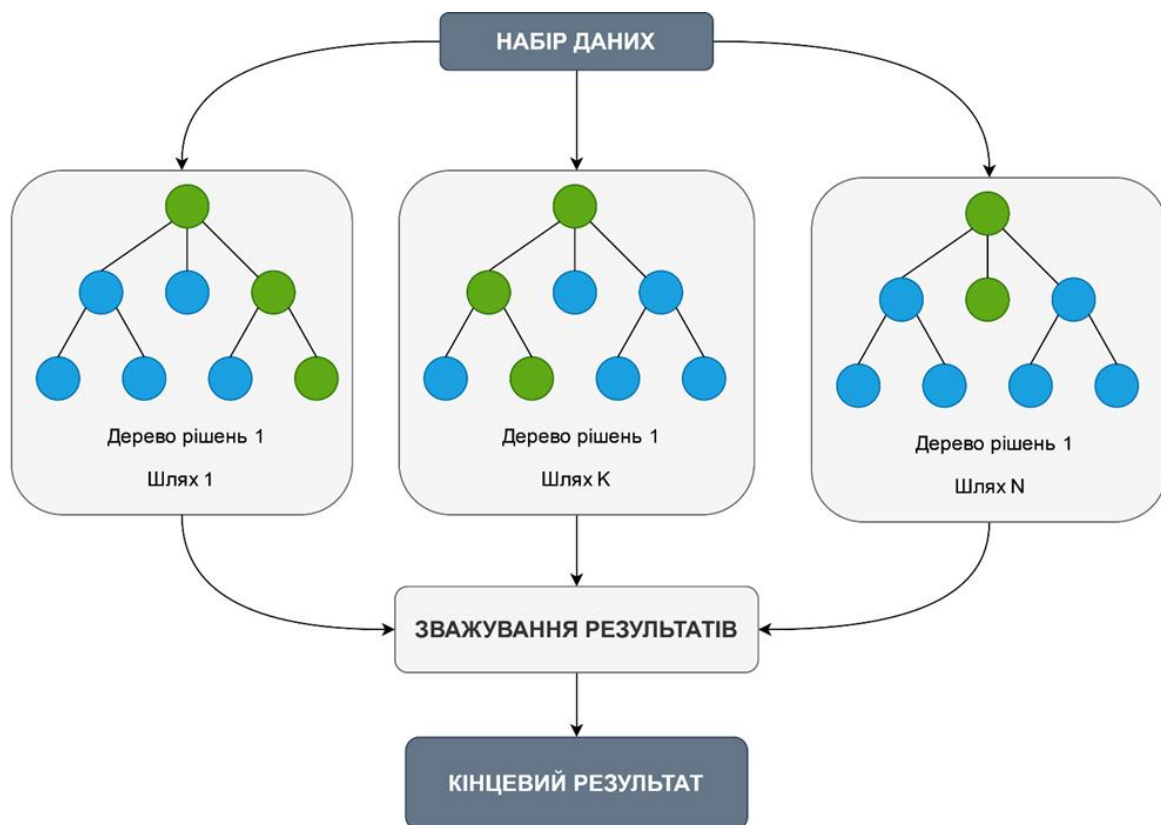


Рисунок 1.6. Візуалізація алгоритму випадкового лісу.

У експериментах кількість ознак, що утворюють підмножину, дорівнює квадратному кореню із загальної кількості ознак, а кількість екземплярів, що утворюють підмножину даних, дорівнює розміру набору даних, який використовується для навчання.

Дерева були вирощені до повного розміру (тобто без обмеження глибини). В такому випадку, часова складність алгоритму побудови моделі випадкового лісу для m дерев, отриманих з n випадків шляхом спостереження за d ознаками, приблизно дорівнює $O(m \cdot n \cdot d \log n)$ [28].

MRNET

Після опублікування набору даних KneeMRI, науковці із Школи медицини Стенфордського університету вирішили створити новий, більш широкий набір даних для досліджень у галузі автоматизованого діагностування колінних суглобів. Цей набір даних був створений для полегшення розробки моделей глибокого навчання, здатних передбачати аномалії в представлених МРТ колінних суглобів.

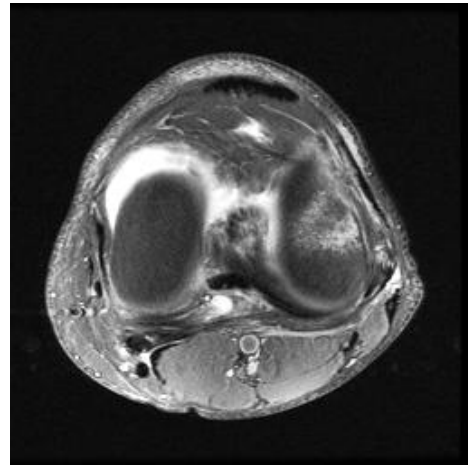
У цій роботі, саме набір даних MRNet був обраний за основу для навчання та порівняння розроблених моделей машинного навчання. В першу чергу, його основною перевагою над набором KneeMRI є наявність зразків зображень у всіх трьох площинах МРТ – аксіальній, корональній та сагітальній. На противагу цьому у KneeMRI були присутні виключно корональні розрізи коліна, що дає значно меншу повноту даних. Маючи зображення усіх трьох розрізів МРТ, можна будувати доволі узагальнені моделі. Крім того, ключовою відмінністю між даними наборами є використання контрасту у MRNet, що також покращує можливість кращого візуального розрізнення жирової тканини від частин кістки, м'язів та зв'язок.

Набір даних MRNet складається з 1370 зразків МРТ колінних суглобів, проведених у медичному центрі Стенфордського університету. Такий набір даних містить 1104 (80,6%) аномальні обстеження, з 319 (23,3%) розривами ACL і 508 (37,1%) розривами меніска. Мітки були отримані шляхом ручного вилучення з клінічних звітів [29]. Приклади МРТ-зображень різних зрізів

колінного суглоба показано на рисунку 1.7.



(a)



(б)



(в)

Рисунок 1.7. Приклади розрізів МРТ із набору даних MRNet: а – сагітальний розріз, б – аксіальний розріз, в - корональний розріз.

Найпоширенішими показаннями до МРТ колінного суглоба в наборі даних були гострий і хронічний біль, контрольне або передопераційне обстеження, а також пошкодження або травми. Обстеження проводили за допомогою сканерів GE (GE Discovery, GE Healthcare, Waukesha, WI, USA) зі стандартною катушкою для МРТ колінного суглоба та звичайним протоколом МРТ колінного суглоба без контрасту, який включав наступні послідовності:

- T1-зважене сканування на корональному розрізі;
- T2-зважене сканування з придушенням жиру на корональному

- розрізі;
- сканування протоної щільності (PD-сканування) на сагітальному розрізі;
- T2-зважене сканування із придушенням жиру на сагітальному розрізі;
- зважене PD-сканування з придушенням жиру на аксіальному розрізі.

Загалом у 775 (56,6%) обстеженнях використовувалося магнітне поле 3,0 Тл; інші використовували магнітне поле 1,5 Тл. Зображення були витягнуті з файлів Digital Imaging and Communications in Medicine (DICOM), масштабовані до 256×256 пікселів і перетворені у формат Portable Network Graphics (PNG) за допомогою мови програмування Python (версія 2.7) [30] і бібліотеки `pydicom` (версія 0.9.9) [31].

Щоб врахувати зміни у шкалі інтенсивності пікселів у серіях МРТ, до зображень було застосовано алгоритм стандартизації інтенсивності на основі гістограми [32]. Для кожної серії еталонний розподіл інтенсивності було отримано з тренувального набору даних. Потім параметри цього розподілу використовувалися для налаштування інтенсивності пікселів у всіх частинах набору даних (для навчання, валідації та тестовому наборі).

Завдяки цим перетворенням пікселі з подібними значеннями інтенсивності відповідають подібним типам тканини. Після стандартизації інтенсивності значення пікселів були обрізані між 0 і 255 як стандартний діапазоном для зображень PNG [29].

При побудові даного набору, було використано описи обстежень кожного пацієнта лікарями – рентгенологами, і з них вручну було витягнуто інформацію про наявність таких перелічених нижче діагнозів для подальшої задачі класифікації.

- Розрив передньої хрестоподібної зв'язки (англ. anterior cruciate ligament, ACL).
- Розрив меніска.
- Аномальна форма коліна.

З моменту публікації набору медичних зображень MRNet, створено велику кількість інтелектуальних інформаційних систем автоматизованого діагностування колінного суглоба за МРТ. Більшість таких систем використовують методи глибокого навчання та згорткові нейронні мережі.

Серед усіх спроб, основною сучасною моделлю для автоматизованого діагностування травм колінного суглоба, розробленою на цьому наборі даних є оригінальна однойменна архітектура глибокого навчання MRNet, реалізована в роботі [29]. У своїй роботі, автори створили комбіновану модель для виявлення травм коліна з окремою моделлю для аналізу кожного зрізу МРТ.

MRNet — це згорткова нейронна мережа (англ. Convolutional Neural Network, CNN), яка приймає як вхідні дані серію МРТ-зображень і видає прогноз класифікації певного діагнозу.

Загальна архітектура моделі показана на рисунку 1.8.

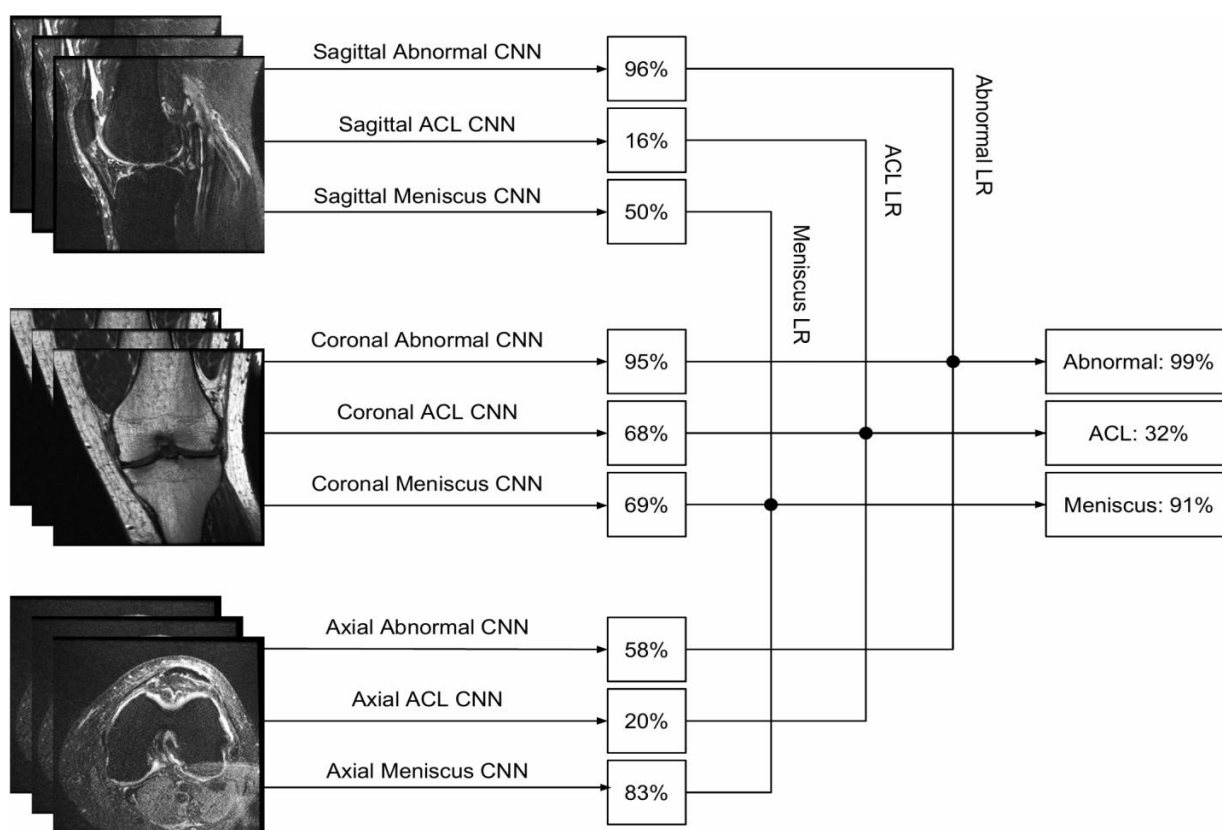


Рисунок 1.8. Загальна структура моделі глибокого навчання MRNet [29].

Кожен із блоків мережі MRNet складається із попередньо натренованого

шару виділення ознак у вигляді мережі AlexNet [33]. Виділені ознаки із шару AlexNet з кожного зрізу серії МРТ об'єднуються за допомогою операції агрегування за максимумом по елементах (англ. max pooling) [34].

Результуючий вектор подається через повністю підключений шар для створення єдиної вихідної ймовірності, використовуючи функцію активації softmax.

У своїй роботі автори здійснили навчання різних MRNet для кожного завдання діагностування (аномальна форма коліна, розрив передньої хрестоподібної зв'язки (ACL), розрив меніска) і розрізу серії МРТ (сагітальний, корональний, аксіальний), у результаті чого було створено 9 різних моделей MRNet.

Також, здійснено зовнішню перевірку натренованих моделей із попереднім набором даних — KneeMRI. Треба однак зазначити, що така перевірка здійснювалася лише на сагітальному розрізі, та виключно для діагностування розриву ACL, оскільки лише такі дані були представлені у KneeMRI).

Для кожного окремого зрізу МРТ – аксіального, коронального та сагітального – окремий блок MRNet навчається класифікувати травми та аномалії. Архітектура одного блоку показана на рисунку 1.9. Насамкінець, використовується підхід логістичної регресії для агрегації результатів роботи окремих блоків MRNet у єдину модель бінарної класифікації. Така модель працює для всіх трьох типів діагнозів у даному дослідженні - розривів ACL, розривів меніска та загальних аномалій колінного суглоба.

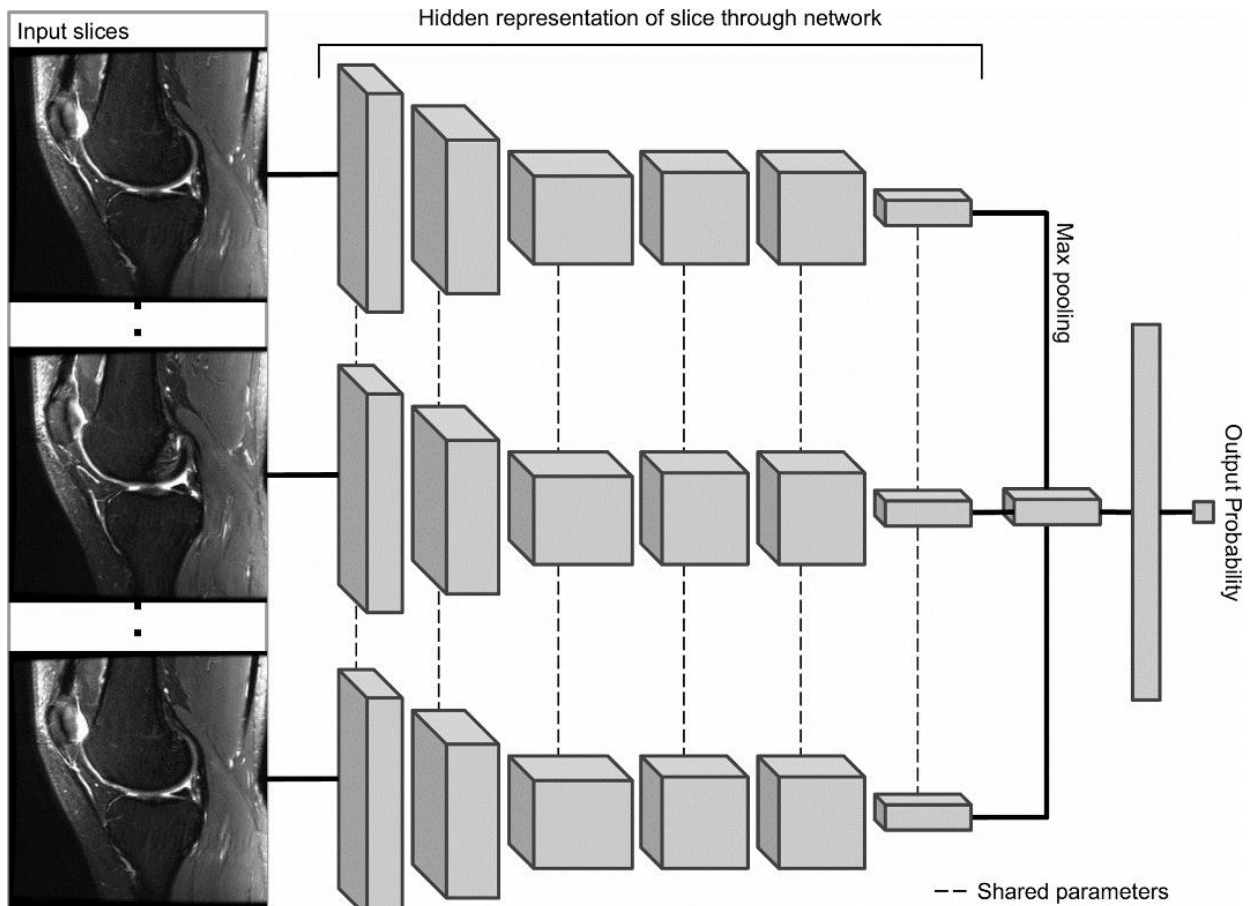


Рисунок 1.9. Схематичне зображення одного блоку мережі MRNet [29].

Для кожної побудованої моделі, кінцевий числовий результат означає ймовірність, яку дана модель асоціює із певним розрізом та певним діагнозом .

Розробники набору даних MRNet зі Стендфорського університету займаються організацією конкурсу, щоб заохотити інших науковців до розробки моделей для автоматизованої інтерпретації МРТ колінного суглоба.

Таблиця учасників, що подавалися на даний конкурс, показана на рисунку 1.10.

Тестовий набір даних, що використовується, (у архіві набору даних, він називається внутрішнім валідаційним набором), має основний набір істинності, використовуючи більшість голосів серез 3х практикуючих сертифікованих

Leaderboard

The leaderboard reports the average AUC of the abnormality detection, ACL tear, and Meniscal tear tasks.

Rank	Date	Model	AUC
1	Jan 09, 2019	mmnet-baseline (single model) <i>Stanford University</i>	0.917
2	May 28, 2019	dc_baseline(single model) <i>Mason High</i>	0.911
3	May 29, 2019	Triple-MRNet (single model) <i>Independent Researcher</i> https://github.com/yashbhalgat/MRNet-Competition	0.904
4	Apr 27, 2019	IL_baseline (single model) <i>Stanford Alum</i>	0.900
5	Oct 31, 2019	MagNet <i>single model</i>	0.860
6	Nov 28, 2019	MRPredNet (single) <i>Anonymous</i>	0.837

Рисунок 1.10. Таблиця лідерів конкурсу моделей для автоматизованого діагностування набору MRNet.

лікарів-рентгенологів. Вони мали досвід професійної діяльності в межах 6–19 років, причому середнє значення складало 12 років.

Рентгенологи мали доступ до всіх оригінальних серій зображень у форматі DICOM, оригінального звіту та історії хвороби, а також до контрольних обстежень під час інтерпретації для кожного пацієнта.

Для оцінки результатів конкурсу MRNet, використовувалась прихована тестова вибірка із набору даних для офіційної оцінки кожної моделі. Команди надсилали свій виконуваний код на Codalab, який потім запускається на тестовому наборі, що не доступний для читання. Така стратегія дозволяє зберігати добросовісність результатів конкурсу.

1.4. Каркаси виділення ознак

В машинному навчанні та розпізнаванні образів **ознака** (англ. feature) — це окрема властивість або характеристика спостережуваного явища, яку можливо виміряти [35]. Ключовим кроком у побудові алгоритмів машинного навчання та розпізнавання образів, а також класифікації та регресії є вибір

інформативних, розрізняваних і незалежних одне від одного ознак.

Зазвичай, ознаки є числовими значеннями, проте інколи, в синтаксичному розпізнаванні образів застосовуються також і структуровані ознаки – стрічки або графи. Саме по собі поняття «ознака» тісно корелює із поняттям «описова змінна» у суміжних методах на кшталт лінійної регресії.

Вектор ознак — це вектор, який зберігає ознаки для певного спостереження в певному порядку. Порядок може бути важливим або ж ігноруватися, залежно від задачі.

У машинному навчанні чисельне представлення об'єктів є переважним, оскільки більшість методів спираються саме на числові характеристики, які до того ж є зручними в обробці та статистичному аналізі. Алгоритми автоматизованого діагностування зображень МРТ також залежать від векторів ознак, оскільки при представленні зображень, значення ознак відповідають пікселям зображення, а саме інтенсивності для кожного пікселя.

Вектори ознак є еквівалентні до векторів описових змінних, що використовують в таких статистичних процедурах, як лінійна регресія. Вектори ознак часто об'єднують з ваговими коефіцієнтами за допомогою скалярного добутку з метою побудови лінійної функції, що застосовують для визначення оцінки для здійснення передбачення.

У даній роботі під **каркасом виділення ознак** (або **каркасною мережею виділення ознак**) розуміється модель глибинного навчання, що розташована на початковому шарі кожного блоку досліджуваної гіпер-моделі MRNet, основним завданням якої є конвертація серії вхідних зображень розрізу МРТ у набір векторів ознак, або ж тензор ознак.

Із часу публікації авторами праці, де вперше було побудовано модель MRNet на основі каркасу виділення ознак AlexNet відбулися суттєві зміни у галузі комп'ютерного бачення. Була представлена велика кількість нових моделей, котрі потенційно можна використати в якості каркасу виділення ознак.

1.5. Постановка задачі

Глобальною метою даного дисертаційного дослідження є побудова інформаційної системи, котра зможе здійснювати більш ефективне автоматизоване діагностування зображень МРТ коліна, ніж існуючі моделі.

При цьому така інформаційна система повинна бути досить компактна щодо використання ресурсів для можливості її майбутнього інтегрування у інформаційні системи лікарень та діагностичних центрів. Для забезпечення мети дисертаційного дослідження виділено наступні перелічені нижче задачі:

- здійснити порівняльний аналіз сучасних архітектур комп'ютерного бачення в якості каркасу виділення ознак для моделі MRNet;
- підібрати оптимальний набір каркасів виділення ознак для кожного діагнозу та для кожного розрізу зображень МРТ, використовуючи додаткові метрики оцінки точності передбачення;
- на основі набору каркасів, сформувану комбіновану модель для автоматизованого діагностування МРТ колінних суглобів із вищою точністю передбачення, аніж моделі, побудовані у роботах авторів, згаданих у цьому розділі, за різними показниками оцінювання точності;
- дослідити алгоритми та методи пояснення результатів передбачення моделей глибокого навчання та підібрати відповідні методи для використання із заданою комбінованою моделлю;
- описати архітектуру та побудувати прикладну інформаційну систему для автоматизованого діагностування зображень МРТ колінного суглоба із функцією пояснення результатів передбачення.

Короткий огляд найбільш суттєвих попередніх досліджень за тематикою автоматизованого діагностування МРТ колінного суглоба для кожної роботи наведено нижче.

Semi-Automated Detection of Anterior Cruciate Ligament Injury from MRI [26] . Оpubлікована у 2014 році, ця робота є одією із перших спроб застосування методів комп'ютерних наук для задачі автоматизованого

діагностування розриву передньої хрестоподібної зв'язки (ACL) на основі зображень МРТ коліна.

Автори провели детальний аналіз 917 клінічних випадків із різноманітним ступенем ушкодження ACL від абсолютно неушкоджених до цілковито розірваних. Кожне обстеження, із відповідним текстом діагнозу, було додатково провалідоване комісією із трьох лікарів-рентгенологів, та одним старшим фахівцем м'язово-скелетної радіології у клінічному центрі Рієки, Хорватія.

Окрім сформованого публічно доступного набору даних для подальшого розвитку технологій для автоматизованого діагностування, авторами було також запропоновано декілька моделей на основі методу опорних векторів для класифікації розриву ACL у кожному із зображень.

Недоліки та нерозв'язані задачі. Попри те, що автори зазначають доволі високий показник ROC-AUC у 0.943, дані моделі показують посередню ефективність на раніше невідомих даних. Автори не згадують проблеми перенавчання та підбору оптимальних значень коефіцієнта регуляризації для моделі.

Щодо самого набору даних, то у кожному обстеженні використовується виділена вручну область інтересу (ROI) і класифікація відбувається з її врахуванням. Це значно ускладнює обробку даних та донавчання моделей у реальних системах, адже кількість даних, що потребуватиме додаткової обробки, є великою.

Новизна моєї роботи порівняно з даною. Враховано та адресовано проблему перенавчання. Моделі у моїй роботі використовують стратегію запобігання перенавчання шляхом штучної випадкової аугментації вхідних даних за допомогою афінних та інших перетворень.

Також, моделі у моїй роботі використовують автоматизований процес виділення ознак, що спрощує подальший неперервний процес навчання після розгортання кінцевих версій систем автоматизованого діагностування.

Deep-learning-assisted diagnosis for knee magnetic resonance imaging:

Development and retrospective validation of MRNet [29].

Це перша базова робота, яка поклала початок мого дисертаційного дослідження. Автори формують новий набір даних для дослідження проблеми автоматизованого діагностування пацієнтів за МРТ колінного суглоба із зовнішньою валідацією зібраних даних комісією експертів-рентгенологів із Стенфордської Школи Медицини. Додатково автори публікують власну архітектуру для такого автоматизованого діагностування на базі мережі глибинного навчання для комп'ютерного бачення – AlexNet.

Недоліки та нерозв'язані задачі. Використовувалась доволі застаріла архітектура комп'ютерного бачення AlexNet в якості шару виділення ознак. Дослідження не фокусувалося на інтерпретованості результатів класифікації моделей глибинного навчання.

Новизна моєї роботи порівняно з даною. Фокус на більшій кількості сучасних архітектур комп'ютерного бачення, а також аугментація вхідних даних для вирішення проблеми перенавчання моделей. Застосування більш складних та результативних метрик оцінки ефективності. Розробка інтерпретаційної моделі для пояснення результатів класифікації.

A Comparative Study of Existing and New Deep Learning Methods for Detecting Knee Injuries using the MRNet Dataset [36].

У даній роботі було проведено порівняльний аналіз методів для автоматизованого діагностування пацієнтів на основі МРТ колінного суглоба. Автори сфокусувалися на порівнянні різноманітних стратегій навчання моделей глибинного навчання, а також способів організації вхідних даних для цих моделей. Досліджувалися як тримодельні архітектури на базі MRNet-baseline, так і одномодельна архітектура, на вхід до якої подаються сконкатеновані зображення по кожному розрізу МРТ.

Недоліки та нерозв'язані задачі. Неможливість відтворити показник точності, застосовуючи стратегію навчання, вказану авторами. Недостатньо детальний опис кінцевих результатів – не вказано, який набір даних використовувався для отримання кінцевих результатів продуктивності моделей.

Дослідження не фокусувалося на інтерпретованості результатів класифікації моделей глибинного навчання.

Новизна моєї роботи порівняно з даною. Більш широкий спектр метрик для порівняння результатів класифікації. Вперше застосовано метрику Капа Коена для порівняння ефективності архітектур на базі різних мереж комп'ютерного бачення в якості каркасних мереж виділення ознак. Капа Коена дозволяє виявити закономірності в продуктивності класифікації, недоступні для інших використаних метрик.

Explaining the Black-box Smoothly - A Counterfactual Approach [37].

У цій роботі автори приділили увагу питанням інтерпретованості моделей глибинного навчання у контексті аналізу медичних даних. Це є важливою тематикою, оскільки сфера застосування у медицині вимагає не тільки безкомпромісно точних методів, але й методів, результати роботи яких можна легко інтерпретувати для безшовної взаємодії із власне, медичними працівниками, котрі апріорі повинні забезпечувати фінальні діагностичні рішення у разі виникнення суперечок між результатами.

Автори використовують підхід латентного зсуву, що дозволяє генерувати зображення на основі автокодувальника, котрі нагадують медичні зображення із вхідного набору даних. Крім цього, даний метод дозволяє «перебільшувати» ознаки, виявлені на згенерованих зображеннях, приводячи до викривлених значень класифікації на згенерованих штучних зображень.

Недоліки та нерозв'язані задачі. Недоліки як такі відсутні, ця робота фокусується на простіших даних, і не вирішує ті проблеми та задачі, що сформульовані у даному дисертаційному дослідженні.

Новизна моєї роботи порівняно з даною. У моєму дисертаційному дослідженні метод, запропонований на базі латентного зсуву, отримав модифікацію для застосування із серійними зображеннями – МРТ.

На відміну від одновимірних рентгенівських знімків грудної клітки у досліджуваній роботі, тут питання цілісності згенерованих серій зображень постає куди більш гостро, адже потрібно не тільки переконатися у тому, що

певне зображення із серії перебільшує лише задалегідь задані ознаки, і не видозмінюється у регіонах, що не впливають на результат роботи класифікатора, а й додатково забезпечити мінімальні відхилення між попереднім та наступним зображеннями у серії.

Усі ці проблеми, пов'язані із додатковою складністю, є сформульовані. Також у моєму дослідженні для їх вирішення застосовані спеціалізовані методики .

Висновки до розділу 1

У цьому розділі було дано визначення поняття процесу автоматизованого діагностування зображень. Здійснено опис процесу автоматизованого діагностування МРТ колінних суглобів, введено основні поняття для аналізу зображень МРТ. Інформація, представлена у цьому розділі, була опублікована у праці автора [3].

Здійснено огляд літератури, описано ключові досягнення у сфері автоматизованого діагностування МРТ колінних суглобів, а саме: набори даних KneeMRI та MRNet, методи комп'ютерних наук, що на основі яких побудовані системи автоматизованого діагностування — глибинні нейронні мережі, метод опорних векторів, моделі випадкового лісу. Здійснено опис сучасних прикладних інформаційних технологій для автоматизованого діагностування МРТ колінних суглобів.

Поставлено задачу покращення точності передбачення існуючої моделі глибинного навчання MRNet, за рахунок дослідження більш сучасних архітектур комп'ютерного бачення в якості каркасних мереж виділення ознак, а саме: VGG11, VGG16, ResNet, EfficientNet.

РОЗДІЛ 2. ПОБУДОВА МОДЕЛІ АНАЛІЗУ ЕФЕКТИВНОСТІ КАРКАСНИХ МЕРЕЖ ВИДІЛЕННЯ ОЗНАК ДЛЯ ЗАДАЧІ АВТОМАТИЗОВАНОГО ДІАГНОСТУВАННЯ МРТ КОЛІННОГО СУГЛОБА

У другому розділі здійснено порівняльний аналіз сучасних архітектур глибинного навчання для комп'ютерного бачення. Розроблено модель для порівняння ефективності мереж глибинного навчання, шляхом підстановки різноманітних сучасних архітектур комп'ютерного бачення в якості каркасів виділення ознак. Розроблено стратегію оцінювання каркасних моделей виділення ознак.

Результати розділу опубліковано у наукових працях [1, 2].

2.1. Порівняльний аналіз сучасних архітектур глибинного навчання для комп'ютерного бачення

У даному розділі представлено порівняльний аналіз сучасних глибинних архітектур комп'ютерного бачення для виділення ознак із зображень МРТ колінного суглоба у задачах класифікації травм та аномалій коліна. Такий аналіз потрібен задля відбору найкращих сучасних моделей комп'ютерного бачення для подальшого створення прикладних архітектур моделей машинного навчання для задачі автоматизованого діагностування травм коліна у медичних приладах та системах.

Для реалізації такого аналізу у цьому розділі дисертаційного дослідження відбувається порівняння продуктивності базової архітектури MRNet для задачі автоматизованого діагностування зображень МРТ колінного суглоба, використовуючи різноманітні сучасні архітектури комп'ютерного бачення в якості каркасних мереж виділення ознак.

AlexNet.

AlexNet – це велика глибинна згортова нейронна мережа, що була розроблена для класифікування 1,2 мільйона зображень із високою роздільною здатністю на конкурсі ImageNet LSVRC-2010 [38], на 1000 різних класів. Згідно

з даними конкурсу автори досягли рівня 37,5% помилок у топ-1 та 17,0% помилок у топ-5, що на момент публікації було значно кращим результатом, аніж у попередніх моделей та методів.

В оригінальній імплементації, дана глибинна нейронна мережа має 60 мільйонів параметрів і 650 000 нейронів, складається із п'яти згорткових шарів, за деякими з яких йдуть шари агрегування по максимуму, і трьох повністю підключених шарів з кінцевим шаром із 1000-канальним softmax в якості активації, що зображено на рисунку 2.1.

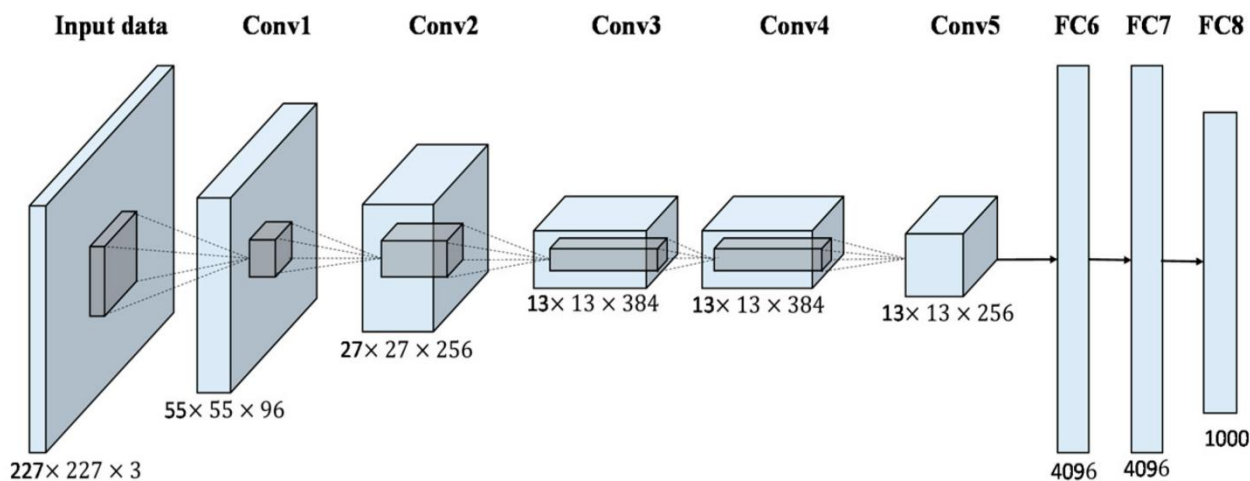


Рисунок 2.1. Архітектура моделі глибинного навчання AlexNet [33].

Побудована архітектура AlexNet стала проривом у сфері згорткових мереж для класифікації зображень. Автори вперше змогли застосувати техніку **випадання**, що дозволяє запобігати перенавчанню моделі, а також зосередили свою увагу на використанні технік аугментації зображення, для синтетичного збільшення вихідного набору даних, що також суттєво впливає на проблему перенавчання.

Випадання (виключення) (англ. dropout) — у машинному навчанні є технікою, що зменшує коадаптацію нейронів, оскільки нейрон не може покладатися на присутність конкретних інших нейронів. Тому кожен нейрон змушений вивчати більш надійні функції, які виявляться корисними в поєднанні

з багатьма різними випадковими підмножинами інших нейронів.

Навіть попри той факт, що дана архітектура була розроблена за 10 років до моменту написання дисертаційного дослідження, її використання у тій же базовій моделі MRNet дозволило досягти суттєвої точності класифікації, із рівнем довіри до результату на рівні комісії експертів-рентгенологів [29].

VGG11 та VGG16.

Успіх архітектури AlexNet популяризував подальші дослідження у галузі комп'ютерного бачення. Велика кількість науковців почала досліджувати шляхи покращення цієї моделі, змагаючись між собою у конкурсах на кшталт ImageNet.

У 2014 році у роботі [39] було зроблено моделі VGG11 та VGG16. Це — дві найефективніші моделі із роботи, які були оприлюднені публічно, щоб сприяти подальшим дослідженням використання глибинних мереж для задач комп'ютерного бачення.

Автори досліджували вплив глибини згорткової мережі на її точність у налаштуваннях розпізнавання великомасштабних зображень. Головним внеском стала ретельна оцінка мереж із зростаючою глибиною з використанням архітектури з дуже малими згортковими фільтрами розміром (3×3), яка показує, що для значного покращення конфігурацій попередніх моделей машинного навчання достатньо збільшивши глибину мережі до 16–19 шарів ваг.

Ці висновки лягли в основу заявки авторів на конкурс ImageNet Challenge 2014, де їх команда зайняла перше та друге місце відповідно в задачах локалізації та класифікації зображень [40].

Також показано, що моделі VGG11 та VGG16 добре узагальнюються на інші набори даних, де вони досягають результатів із приблизно тією ж точністю передбачення, що й інші відповідні сучасні моделі.

ResNet.

Попередні моделі серії VGG-* зробили акцент на збільшенні глибини моделі, задля досягнення кращої точності передбачення. Проте, збільшення

глибини для таких мереж після певного порогу не здатне збільшити точність передбачення, а навпаки — підвищує похибку [41], [42], [43].

Окрім цього, більш глибокі нейронні мережі складніші у процесі навчання. Таким чином, чим більше шарів має звичайна глибинна мережа, тим довше відбувається процес її навчання, і тим більша похибка передбачення виходить у результати. Цей факт проілюстровано на рисунку 2.2.

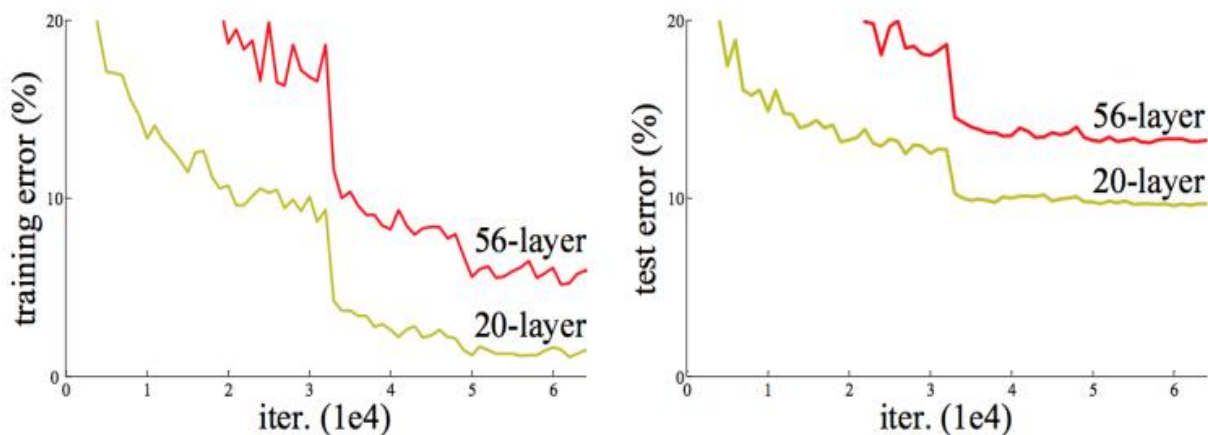


Рисунок 2.2. Тренувальна та тестова похибка для моделей машинного навчання різної глибини, на різній кількості епох (ітерацій) [24].

Для запобігання цьому явищу, у 2015 році було представлено структуру залишкового навчання, щоб полегшити навчання мереж, які є значно глибшими, ніж ті, що використовувалися раніше. Дана структура, або архітектура, отримала назву **ResNet**.

У своїй роботі, автори ResNet чітко переформулюють шари мережі для навчання на залишкових функціях із посиланням на вхідні дані попередніх шарів, на додачу до прямого навчання на вхідних даних шару. Такий підхід називають **з'єднанням швидкого доступу**. Вони здатні пропускати декілька шарів мережі, і співставляють результати до та після цих шарів. Їх виходи додаються до виходів вхідних даних, пропущених через шари мережі. Це проілюстровано на рисунку 2.3.

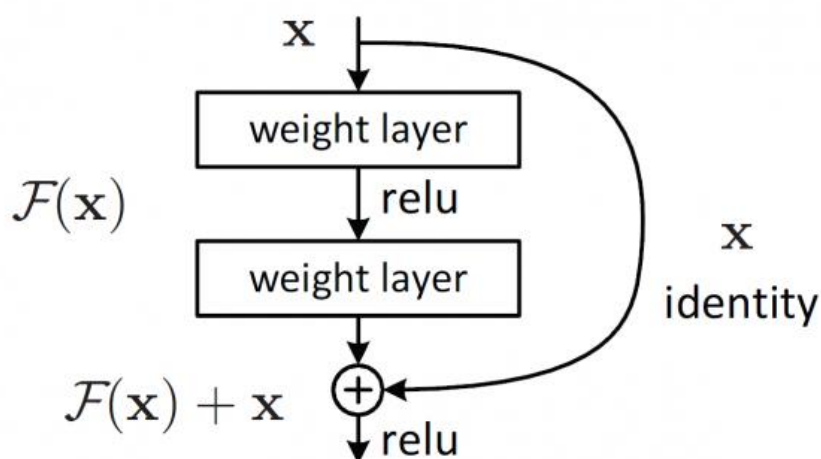


Рисунок 2.3. З'єднання швидкого доступу у мережах залишкового навчання [41].

У роботі надані вичерпні емпіричні докази того, що ці залишкові мережі легше оптимізувати, і вони мають покращену точність завдяки значно збільшеній глибині.

На наборі даних ImageNet автори здійснювали оцінку залишкової мережі із глибиною до 152 шарів - у 8 разів глибше, ніж мережі VGG, проте із меншою складністю та тривалістю навчання. Ансамбль цих залишкових мереж досяг вельми значущого показника похибки передбачення у 3,57% в тестовому наборі ImageNet. Цей результат отримав перше місце на конкурсі ImageNet ILSVRC 2015.

Нижче на рисунку 2.4 показано порівняння тренувальної та тестової похибки передбачення для звичайних глибоких моделей та моделей на базі ResNet. Винайдення ResNet теж стало фундаментальним проривом у галузі комп'ютерного бачення, оскільки надало підґрунтя для побудови дуже глибоких моделей машинного навчання для класифікації зображень, які мають підвищену точність, а також значно спростити розробку нових моделей, оскільки складність процесу тренування моделі та спрощені вимоги до обчислювальних ресурсів дозволили все більшій кількості науковців займатися дослідженнями у цьому напрямі.

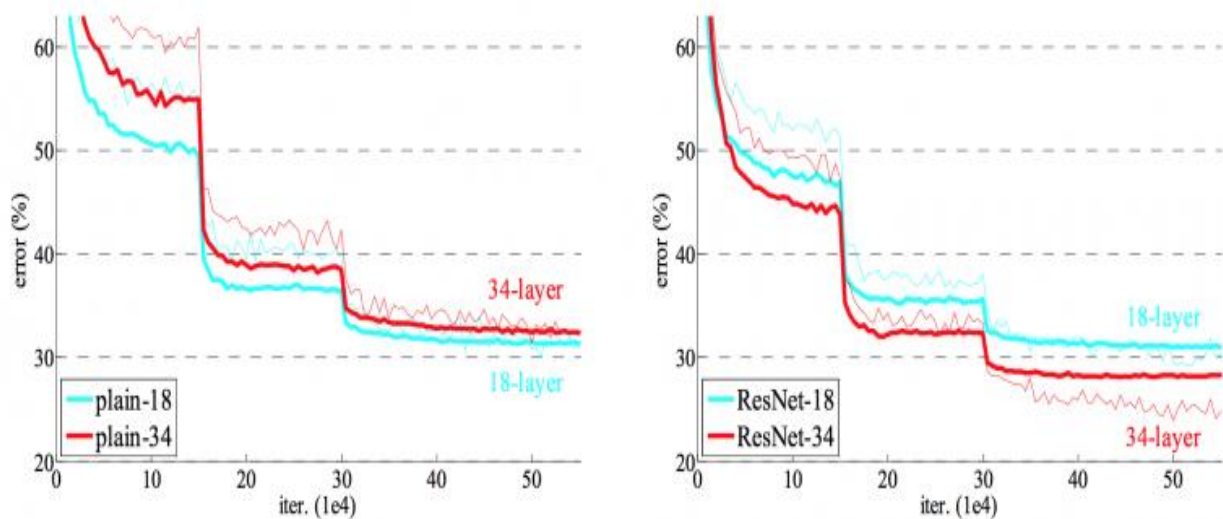


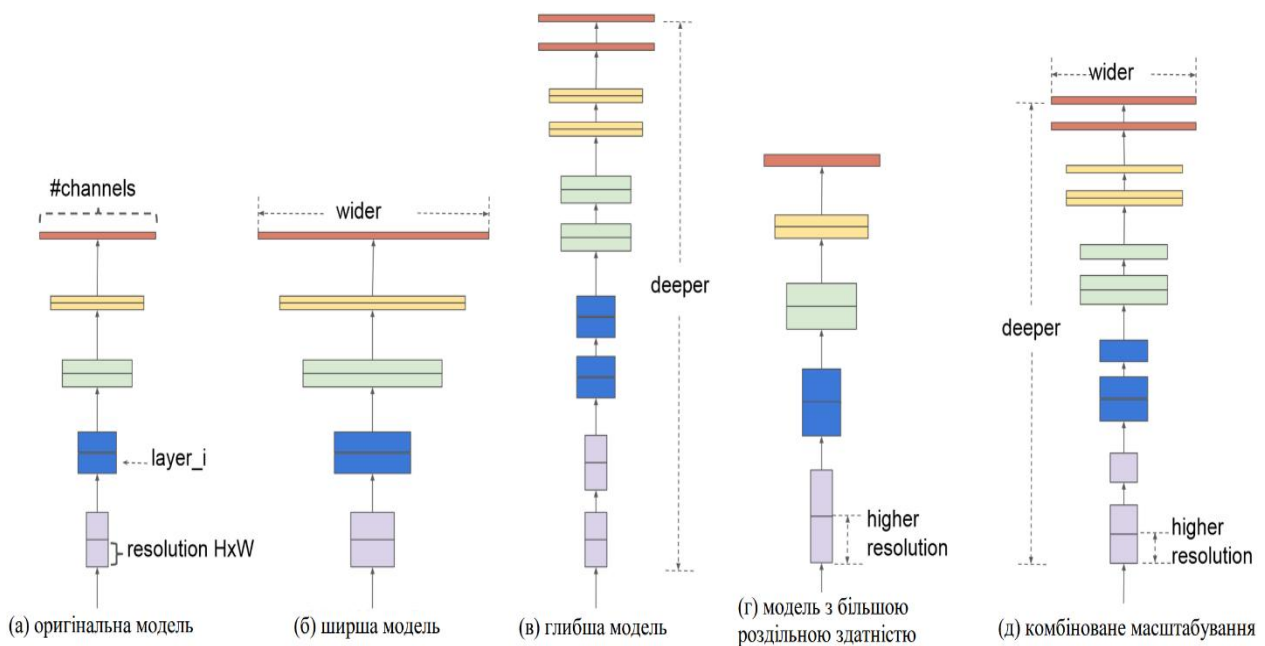
Рисунок 2.4. Тренувальні та тестові похибки передбачення для звичайних глибоких мереж (зліва) та мереж залишкового навчання (справа) [41].

Із винайденням ResNet, наявність кластерів серверів для обчислень, або ж дорогих графічних процесорів для розпаралелення обчислень за допомогою CUDA [44], більше не було критичною вимогою для досліджень.

EfficientNet.

Попередньо згадана архітектура глибокого навчання ResNet робить акцент на покращенні ефективності класифікації моделі за рахунок спрощення масштабування моделі. Проте сам процес масштабування не є достатньо вивченим у відповідній роботі, а масштабування відбувається виключно за рахунок збільшення глибини мережі. Це спонукало до потреби дослідити різноманітні види масштабування мереж глибокого навчання.

Саме тому у роботі автори надають подібний аналіз. Вони роблять припущення, що сама по собі глибина мережі не є ідеальним параметром масштабування, попри інтуїтивне припущення що глибші згорткові мережі можуть охоплювати багатші та складніші ознаки у вхідних даних та краще узагальнюватися на невідомих даних (рисунок 2.5).



Окрім цього, чим більша є глибина мережі, тим складніше її навчати через проблему зникаючого градієнту [45]. Хоча кілька методів, таких як з'єднання швидкого доступу у ResNet або пакетна нормалізація, полегшують проблему навчання, сам приріст точності дуже глибокої мережі зменшується із збільшенням глибини. До прикладу, ResNet-1000 має приблизно таку ж точність, як ResNet-101, хоча перша має майже в десять разів більше шарів. На рисунку 2.5 (посередині) показано наше емпіричне дослідження масштабування базової моделі з різним коефіцієнтом глибини d , що додатково свідчить про зменшення точності для дуже глибоких ConvNets.

Рисунок 2.5. Різноманітні типи масштабування моделей глибокого навчання

Окрім масштабування в глибину, існують також інші способи підвищити точність класифікації моделі (усі вони проілюстровані на рисунку 2.5):

Масштабування роздільної здатності полягає у збільшенні роздільної здатності вхідних зображень. Перші глибокі нейронні мережі для комп'ютерного бачення використовували роздільну здатність у 224x224 пікселі.

Більш сучасні мережі використовують на вхід зображення розміром 299x299, або 331x331 [46], [47]. Чим вища роздільна здатність вхідного зображення, тим більші шанси у глибинної мережі виявити кращі ознаки.

- **Масштабування ширини** – здійснюється за допомогою збільшення кількості нейронів на одному шарі, або ж, у більш релевантній термінології, за допомогою збільшення кількості фільтрів у згортковому шарі.

Автори [48] здійснили емпіричний порівняльний аналіз ефективності комбінування різних типів масштабування на базі оригінальної мережі ResNet, та запропонували свою, більш ефективну архітектуру EfficientNet, базову варіацію якої показано на рисунку 2.6.

Stage i	Operator $\hat{\mathcal{F}}_i$	Resolution $\hat{H}_i \times \hat{W}_i$	#Channels \hat{C}_i	#Layers \hat{L}_i
1	Conv3x3	224 × 224	32	1
2	MBCConv1, k3x3	112 × 112	16	1
3	MBCConv6, k3x3	112 × 112	24	2
4	MBCConv6, k5x5	56 × 56	40	2
5	MBCConv6, k3x3	28 × 28	80	3
6	MBCConv6, k5x5	14 × 14	112	3
7	MBCConv6, k5x5	14 × 14	192	4
8	MBCConv6, k3x3	7 × 7	320	1
9	Conv1x1 & Pooling & FC	7 × 7	1280	1

Рисунок 2.6. Базова архітектура моделі EfficientNet [29]

2.2. Методологія побудови моделі аналізу ефективності каркасних мереж виділення ознак

Майже у всіх попередніх дослідженнях за тематикою дисертаційної роботи використовується базова модель MRNet, що має в якості каркасу виділення ознак попередньо натреновану мережу AlexNet. При цьому використовується трансферне навчання.

У роботі [36] показано, що серед усіх можливих варіацій базової архітектури MRNet саме варіант консенсусу рішення декількох окремих моделей на кожній площині за допомогою логістичної регресії є найбільш продуктивним. Така варіація архітектури має найкращу ефективність, порівнюючи її з AlexNet із самонавчанням, а також композитну модель, натреновану на конкатенації зрізів у всіх трьох площинах МРТ.

Таким чином, у цій дисертаційній роботі прийнято рішення використовувати за основу саме базову архітектуру MRNet для порівняння різних мереж виділення ознак для одного блоку.

У роботі [49] представлено нову архітектуру глибинного навчання, названу Efficiently-Layered Network (ELNet). Дана архітектура є оптимізована для діагностики колінного суглоба за знімками магнітно-резонансної томографії. Основним внеском їхньої роботи є нова мережа виділення ознак, яка включає багатозрізову нормалізацію разом із зменшенням дискретизації BlurPool замість шару агрегування за максимумом чи адаптивного середнього агрегування, які можна побачити в інших пов'язаних роботах.

Автори використовують власний каркас виділення ознак, який, як автори стверджують, є значно ефективнішим за AlexNet, оскільки має в рази меншу кількість параметрів для навчання, в той же час показуючи кращі результати.

Також продемонстровано поступове підвищення точності передбачення цих каркасних мереж, відповідно до еволюції самих моделей, а саме: VGG11, VGG16, Resnet, Efficientnet.

Архітектура Resnet для виділення ознак вже згадувалася в роботі [36], але, враховуючи цей факт, усе одно було прийнято рішення щодо використання цієї архітектури в якості каркасу виділення ознак, незважаючи на її попереднє використання, з декількох причин.

По-перше, у даному дисертаційному дослідженні для порівняння ефективності каркасів виділення ознак використовуються додаткові метрики оцінювання ефективності, які не є присутніми у жодній із робіт, пов'язаних із автоматизованим діагностуванням.

По-друге, навчання моделі із використанням мережі Resnet в якості каркасу виділення ознак додатково слугує для забезпечення прозорості. У даному дослідженні модель на базі Resnet будується з нуля і виключаються такі фактори, як відмінності в параметрах самої моделі або ж параметри навчання. В іншому випадку вплив недетерміністичних факторів міг би завадити адекватній оцінці продуктивності моделей.

Додатковим фактором аби включити модель ResNet до дослідження є відображення поступового підвищення точності передбачення моделей для автоматизованого діагностування, яке співставне із використанням більш нових та прогресивних архітектур комп'ютерного бачення.

Доволі важливим аспектом цього дисертаційного дослідження є той факт, що усі моделі машинного навчання, які розроблені та натреновані в рамках даного експерименту, мають уніфіковану архітектуру за винятком варіативного шару каркасу виділення ознак, і вони усі були навчені з нуля, використовуючи однакові гіперпараметри самої моделі, а також параметри навчання.

Це середовище є ідеально підібраним для порівняння точності передбачення, оскільки недетерміністичні фактори, такі як випадкові початкові значення для ініціалізації ваг у шарі класифікатора моделі, а також параметри навчання (швидкість навчання, спад ваги на кожній епосі, кількість епох навчання) є уніфікованими для кожної окремої моделі, представленої у цій роботі.

На доповнення до вищезгаданого стратегії оцінювання моделей у цій роботі використовують додатковий показник, який ще не використовувався для порівняння точності та якості результатів передбачення у жодній із пов'язаних робіт — а саме метрика Каппа Коена. Ця метрика є значущою, оскільки набір даних MRNet, на основі якого тренуються моделі машинного навчання у даній роботі, не є збалансованим набором даних.

2.3. Стратегії оцінювання каркасних мереж виділення ознак

Для побудови моделі порівняння ефективності каркасних мереж

виділення ознак потрібно мати кількісні характеристики, або метрики, що дозволять порівняти результати бінарної класифікації, отримані цими мережами (рисунок 2.7).

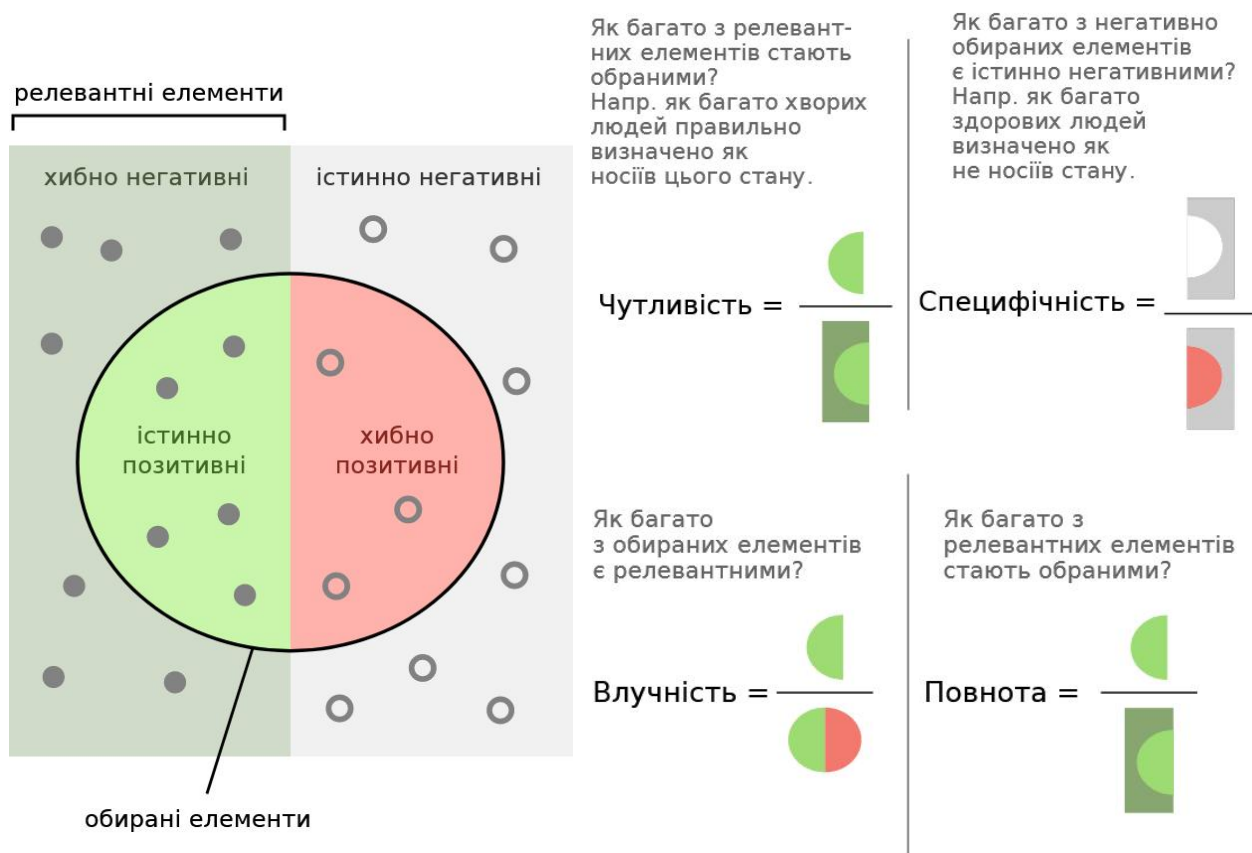


Рисунок 2.7. Візуальна інтерпретація чутливості, специфічності, влучності та повноти (https://en.wikipedia.org/wiki/Sensitivity_and_specificity).

Для вимірювання продуктивності класифікатора, або якості прогнозу є представлено широкий спектр метрик. Як правило, вони не є універсальними і потребують обширного аналізу їх застосування до кожного конкретного типу задачі. В медицині часто використовуються такі поняття як чутливість і специфічність (англ. *sensitivity and specificity*), натомість при видобуванні інформації більш підходящими показниками вважають влучність і повноту (англ. *precision and recall*) [50].

Чутливість вимірює частку позитивних випадків, що були прокласифікованими як істинно позитивні (частка пацієнтів, хто справді має певний стан (уражені), яку було правильно прокласифіковано ураженими).

Специфічність – це частка істинно негативних випадків (наприклад, частка здорових пацієнтів, яких було правильно прокласифіковано як здорових).

Влучність показує релевантність прокласифікованих зразків, в той час, як повнота (яка за визначенням тотожно рівна **чутливості**), теж дорівнює частці правильно прокласифікованих зразків із позитивним класом. Візуальну інтерпретацію даних метрик можна бачити на рисунку 2.7.

У задачах класифікації неможливо самотійно визначити істинність або хибність вхідних даних. Тому, як правило, такі терміни як істинно позитивні результати (англ. true positives), істинно негативні результати (англ. true negatives), хибно позитивні результати (англ. false positives), хибно негативні результати (англ. false negatives) є нічим іншим, аніж надійними зовнішніми судженнями щодо вхідних даних.

«Позитивні» та «негативні» терміни означають результати передбачень з точки зору класифікатора, або ж очікування класифікатора щодо вхідних даних. «Істинно» та «хибно» стосуються зовнішньої надійної оцінки вхідних даних, якій ми довіряємо, і вважаємо істинною для даної задачі.

Суттєвим інструментом у оцінюванні точності бінарних класифікаторів є так звана **матриця невідповідностей**.

Матриця невідповідностей - це таблиця з двома рядками та двома стовпцями, що повідомляє число хибно позитивних, хибно негативних, істинно позитивних, та істинно негативних результатів. Така матриця уможливує більш складний аналіз за просту пропорцію правильних класифікацій (точність). Метрика абсолютної точності може видавати некоректні результати, особливо у випадках незбалансованості вхідного набору даних [51].

Розглянемо на прикладі побудову матриці невідповідностей. Нехай маємо задачу автоматизованого діагностування знімка МРТ коліна на наявність розриву передньої хрестоподібної зв'язки (ACL) та бінарний класифікатор у вигляді моделі глибинного навчання MRNet із заданою каркасною мережею виділення ознак.

Нехай вхідний набір даних складається із 10 серій знімків МРТ у

корональному розрізі і маємо наступні істинні значення діагнозів S у цьому наборі (0 – означає неушкоджену ACL, 1 – розірвану):

$$S = \{0,0,1,1,1,0,1,0,1,1\}.$$

На цьому наборі даних було здійснено передбачення за допомогою вищезгаданого бінарного класифікатора, отримали наступний вектор результатів:

$$X = \{1,0,0,1,1,1,1,0,0,1\}.$$

На основі істинних даних та прогнозованого результату, можна побудувати наступні числові показники: **TP** (істинно позитивні результати), **TN** (істинно негативні результати), **FP** (хибно позитивні результати), що є помилкою першого роду, **FN** (хибно негативні результати), що є помилкою другого роду. Таким чином, можна побудувати матрицю невідповідностей, що узагальнюватиме результати перевірки нашого класифікатора, яка показана у таблиці 2.1.

Таблиця 2.1. Матриця невідповідностей для оцінки класифікатора

		Істинний клас	
		Розірвана ACL	Здорова ACL
Прогнозований клас	Розірвана ACL	TP = 4	FN = 2
	Здорова ACL	FP = 2	TN = 2

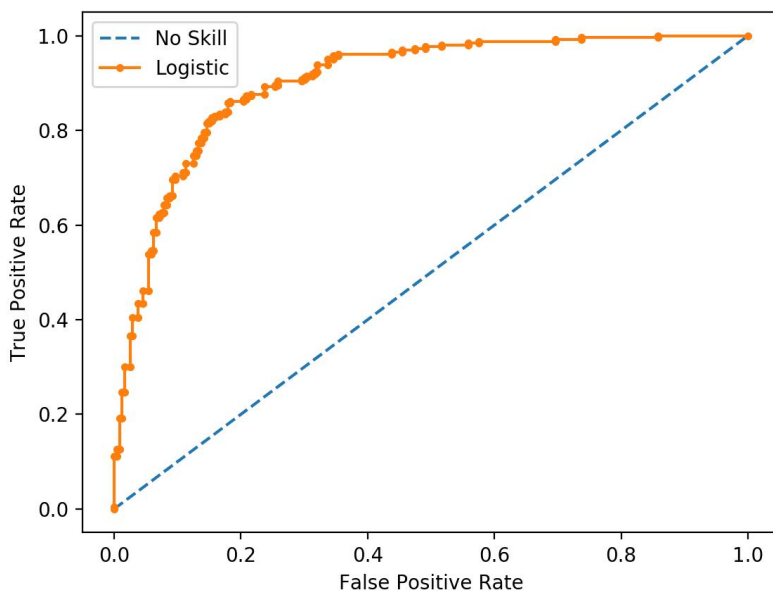
На основі матриці невідповідностей будуються більшість відомих метрик для оцінювання якості бінарного класифікатора.

ROC-крива.

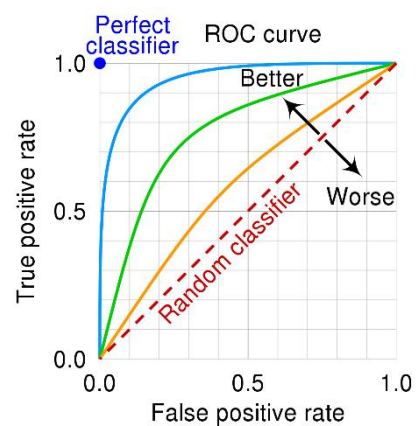
Крива робочих характеристик приймача (англ. receiver operating characteristics curve, ROC), — це графік, який ілюструє діагностичну здатність системи бінарного класифікатора, коли її поріг розрізнення змінюється. Спочатку метод був розроблений для операторів військових радіолокаційних приймачів, починаючи з 1941 року, що призвело до його назви.

Крива ROC створюється шляхом побудови графіка частоти істинних позитивних результатів відносно частоти хибних позитивних результатів, за різних порогових значень перетворення ймовірності класифікатора на власне, конкретний клас. ROC є відома як крива відносної робочої характеристики, оскільки це порівняння двох робочих характеристик (частоти істинних позитивних результатів і частоти хибних позитивних результатів) у міру зміни порогу розрізнення [52].

Приклад кривої ROC, а також характеристики «кращих» та «гірших» класифікаторів за виглядом графіку кривої ROC зображені на рисунку 2.8.



(a)



(б)

Рисунок 2.8. Побудова кривої ROC: а – приклад побудови, б – оцінка класифікаторів за кривою ROC [35].

В якості кількісної характеристики інтерпретації ROC маємо показник AUC (англ. area under ROC curve, площа під ROC-кривою). Цей показник обчислюється як площа, обмежена ROC-кривою та віссю частки хибних позитивних результатів класифікації серед всієї вибірки. Більші показники AUC відповідають якісніше розробленим класифікаторам. При цьому,

класифікатор із значенням ROC-AUC 0,5 відповідає простому випадковому вгадуванню результатів.

F-оцінка.

F-оцінка або **F-міра** є мірою точності результату класифікації. Цей показник розраховується на основі влучності та повноти, які були описані на початку підрозділу [53].

Існує декілька видів F - оцінок. Найбільш простим є показник F_1 , є середнім гармонічним влучності та повноти [54]:

$$F_1 = \frac{2}{recall^{-1} + precision^{-1}} = 2 \cdot \frac{precision \cdot recall}{precision + recall} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}$$

Також є більш загальна F - міра, яка дозволяє масштабувати вплив влучності або повноти на результат оцінки надійності класифікатора. Більш загальний показник F_β , який використовує позитивний дійсний коефіцієнт β , де β вибрано так, що повнота вважається в β разів більш важливою, аніж влучність:

$$F_\beta = (1 + \beta^2) \cdot \frac{precision \cdot recall}{(\beta^2 \cdot precision) + recall}$$

У термінології значень для помилок першого й другого роду, ця формула виглядатиме наступним чином:

$$F_\beta = \frac{(1 + \beta^2) \cdot TP}{(1 + \beta^2) \cdot TP + \beta^2 \cdot FN + FP}$$

Двома значеннями, котрі часто застосовуються для параметра $\beta \in 2$, яке зважає повноту вище, ніж точність, і 0.5, яке зважає повноту нижче за точність.

В якості метрики оцінювання ефективності різних каркасних мереж виділення ознак для архітектури MRNet, у цьому розділі буде використовуватися саме міра F_1 , яку надалі називатимемо F1-мірою, або F1-оцінкою.

Капа Коена.

Під час побудови моделі бінарної класифікації існує багато способів інтерпретації та оцінки продуктивності моделі на основі аналізу результатів її

передбачень. Такі показники оцінювання, як площа під кривою робочих характеристик приймача (ROC-AUC), є добре відомими під час аналізу ефективності класифікаторів. Більшість таких метрик базується на матриці невідповідностей. Однак не всі вони здатні добре себе проявляти на незбалансованих наборах даних.

Незбалансованими можна вважати такі набори даних, де цільовий клас має нерівномірний розподіл спостережень, тобто одна мітка класу має відносно велику кількість спостережень, тоді як інший клас має порівняно низьку кількість елементів у відношенні до загального розміру вибірки.

Набір медичних зображень MRNet, на якому базоване дане дисертаційне дослідження, цілком можна вважати незбалансованим набором даних. При загальній кількості обстежень у 1370 серій знімків МРТ колінного суглоба на всіх трьох площинах, цей набір даних є незбалансованим в сторону позитивних діагнозів, оскільки містить 80,6% таких обстежень. На противагу, кількість «здорових» колінних суглобів у наборі медичних зображень MRNet становить лише 19,4%.

Спитаючись на незбалансованість даних, результати оцінки для різних каркасних мереж виділення ознак, описаних у даній роботі, не можуть покладатися виключно на метрики абсолютної точності прогнозування, оскільки дисбаланс у формі відсутності нормальних, або “здорових” обстежень коліна є значним. Таким чином, з метою подальшої та точнішої оцінки моделей машинного навчання у цій роботі також вводиться розрахунок показника Каппа Коена.

У своїй оригінальній статті [55] Якоб Коен описує свою статистику Капа просто як частку випадково очікуваних розбіжностей між двома оцінювачами, які не виникли, або ж, альтернативно, як частку згоди між оцінювачами після того, як випадкове вгадування було вираховано з розгляду.

Ця статистика часто використовується для перевірки надійності перехресних оцінювань. Перехресні оцінювання полягають у тому, що відбувається класифікація більше, ніж з одного джерела (класифікатора, або

оцінювача) Важливість надійності оцінювача полягає в тому, що цей підхід відображає ступінь кореляції даних, зібраних для аналізу, до вимірюваних змінних, котрі піддаються оцінюванню [56].

З точки зору Коена очікується, що ми матимемо двох незалежних суддів або класифікаторів, причому категорії для класифікації будуть незалежними, взаємовиключними та вичерпними. Для порівняння ефективності каркасних мереж підхід надійності оцінювача потрібно модифікувати для досягнення вищезгаданих умов. Саме тому в якості першого оцінювача позначається модель класифікації для даного діагнозу - розриву ACL, розриву меніска або аномальної форма коліна. Другим оцінювачем є, власне, набір істинних значень із тестового набору даних. Хоча, за термінологією Коена, не існує як таких «істинних» чи «хибних» суджень, але у модифікації для порівняння ефективності каркасних мереж виділення ознак, прогнози, зроблені другим оцінювачем, можна вважати «істинними».

Загалом, значення показника каппа Коена можна розрахувати наступним чином:

$$K = \frac{p_0 - p_e}{1 - p_e},$$

де p_0 - це спостережувана точність, або ж частка рішень оцінювачів, які збігаються, а p_e - це частка передбачень, які очікуються випадково, або ж очікувана точність.

Обчислення метрики капа Коена для моделей класифікації спирається на матрицю невідповідності. Така тенденція присутня і для більшості інших метрик оцінки ефективності класифікації. На противагу абсолютній точності, капа Коена здатна приймати до уваги нерівномірність міток або класів із набору вхідних даних.

Можна виразити p_0 як просто кількість коректно здійснених передбачень, або, із матриці невідповідності:

$$p_0 = \frac{TP + TN}{N}$$

Очікувану точність p_e можна визначити за допомогою відношення

кількості кожного класу передбачення до загальної кількості елементів, для кожного із оцінювачів. Для позитивного класу матимемо відношення R_{pos} , що дорівнює:

$$R_{pos} = \frac{(TP + FP) \cdot (TP + FN)}{N}$$

Аналогічно можемо визначити таке ж відношення і для негачивного класу:

$$R_{-} = \frac{(TN + FP) \cdot (TN + FN)}{N}$$

У кінцевому результаті виразимо очікувану точність p_e через попередньо виведені значення R_{pos} та R_{-} :

$$p_e = R_{pos} + \frac{R_{-}}{N}.$$

Надалі використовуватимемо саме такі формули для обчислення значень p_0 та p_e . Питання інтерпретації значень метрики капа Коена є складним. В першу чергу, дана метрика набуває значень в інтервалі $[-1, 1]$, що вже є поза межами числових значень традиційних метрик. Окрім цього, існує розбіжність щодо трактування від'ємних значень даної метрики. Деякі наукові роботи інтерпретують це як виключно випадкові співпадіння між оцінювачами. Інші ж вважають від'ємні значення цієї метрики як невідповідні неспівпадіння між оцінювачами, тобто що вони невідповідним чином не погоджуються між собою. У даному дисертаційному дослідженні цю метрику обрано як таку, що має значний потенціал для розкриття кращих моделей класифікації шляхом виключення фактора випадкового вгадування.

2.4. Побудова моделі порівняння каркасних мереж

У цій роботі створено та навчено кілька моделей класифікації на оригінальній архітектурі MRNet із використанням більш сучасних каркасних мереж виділення ознак на базі глибинних архітектур комп'ютерного бачення для кожного блоку MRNet, включаючи VGG11, VGG16, Resnet та Efficientnet.

Оригінальна версія MRNet з каркасом AlexNet також була навчена з нуля

порівняння результатів оцінки новішої моделі.

Стратегія оцінювання моделей полягає у вимірюванні показників ефективності класифікації, а саме ROC-AUC, точності передбачення, та міри F_1 . Крім цього, додатково обчислюється показник каппа Коена для кожної моделі та кожного типу діагнозу.

Кожен блок для кожної навченої моделі відповідає вихідній архітектурі MRNet, визначеній у [29]. Блок працює в одній площині МРТ і вчиться виділяти з цієї площини релевантні ознаки в процесі навчання.

У всіх моделях, представлених у цьому документі, єдиним варіативним шаром є рівень виділення ознак, і різні вхідні розміри класифікатора, оскільки не всі каркасні мережі виділення ознак із послідовності вхідних зображень мають однакову розмірність на виході для вектору ознак.

Результуюча архітектура блоків моделі, яка використовується для порівняння ефективності каркасних мереж в цьому дослідженні, представлена на рисунку 2.9.

Для побудови моделі діагностування по кожному із заданих діагнозів, використовувалися три навчальні блоки що відповідають розрізам томографії - аксіальному, корональному та сагітальному відповідно. Для агрегації результатів цих блоків було використано підхід логістичної регресії (LR). Таким чином, здійснюється об'єднання результатів трьох окремих блоків MRNet, і на виході маємо консенсус рішень кожного блоку на кожному розрізі.

Отримана макроархітектура, що включає моделі LR, є такою самою, як і на рисунку 2.9.

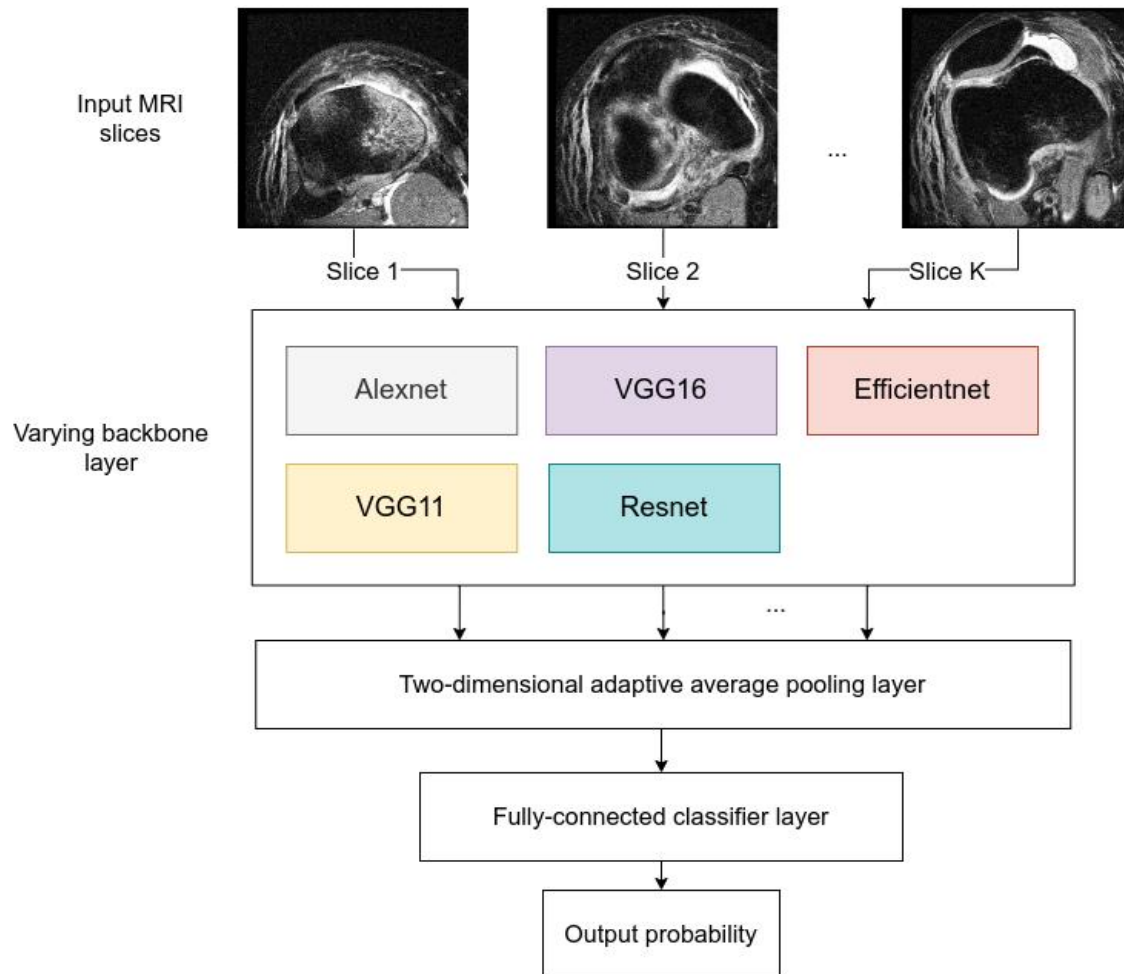


Рисунок 2.9. Результуюча архітектура блоку MRNet для порівняння ефективності каркасних мереж виділення ознак.

Експеримент із порівняння ефективності розпочався із навчання кожного з базових блоків MRNet на кожній площині МРТ. Навчальний процес проводився на платформі Google Cloud Platform, зокрема, з використанням сервісу Vertex AI. Загалом було навчено 45 моделей у хмарі (3 навчальні завдання для кожного типу діагностування, помножені на 3 площини МРТ, помножені на кількість типів каркасних мереж виділення ознак, що розглядалися (5)). Кожне навчальне завдання було прискорено за допомогою двох адаптованих під хмарні обчислення графічних процесорів NVIDIA Tesla P100 GPU, ресурси яких використовувалися через посередництво сервісу Vertex AI.

Кожен з блоків був навчений протягом 10 епох з оцінкою моделі на наборі тестових даних після закінчення кожної епохи. Наприкінці кожної епохи також здійснювався знімок (англ. snapshot) блоку, що завантажувався в хмарний сервіс збереження об'єктів Google Cloud Storage [57]. Схематична діаграма процесу оцінки ефективності однієї каркасної мережі виділення ознак показана на рисунку 2.10.

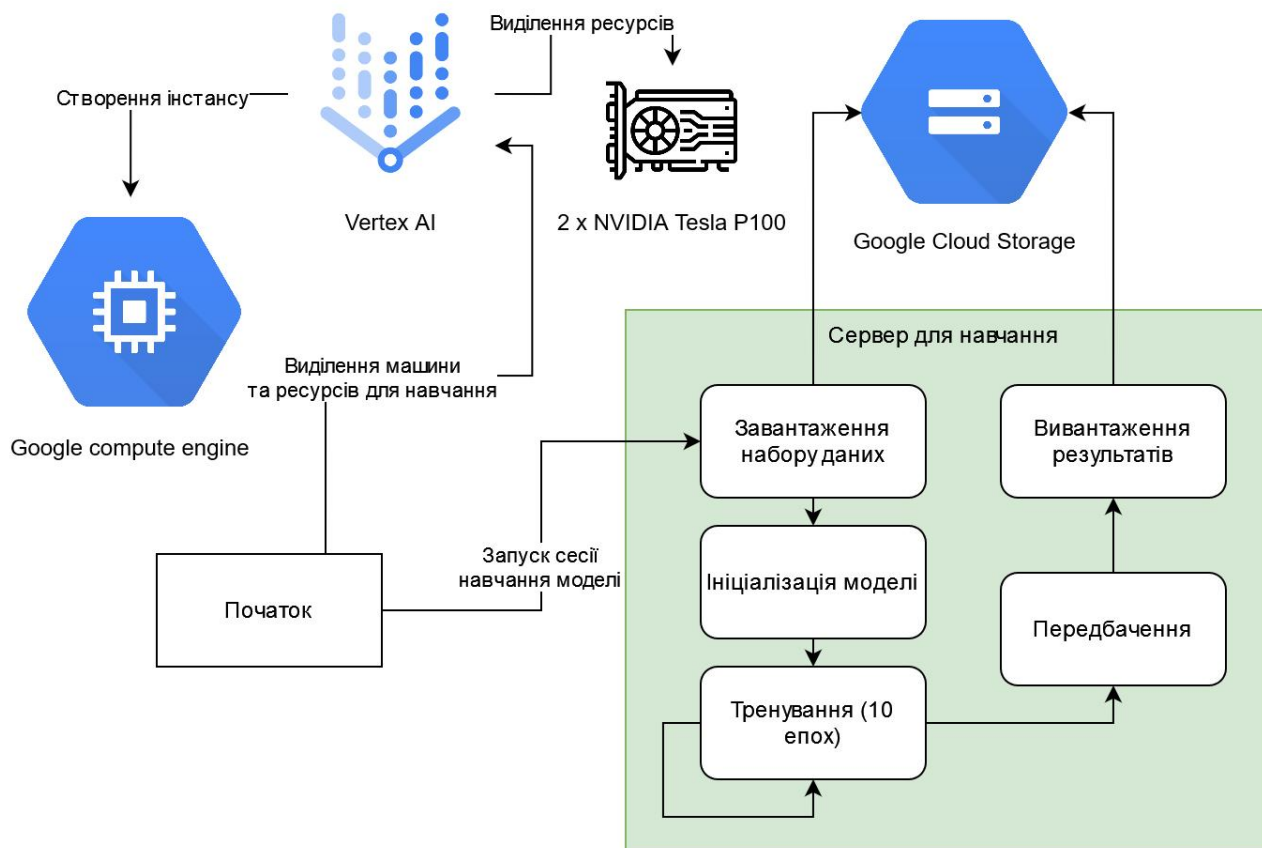


Рисунок 2.10. Схематика процесу порівняння ефективності каркасних мереж виділення ознак

Після тренувального процесу були відібрані лише блоки з найкращими оціночними балами на валідаційному наборі даних. Далі отримані натреновані 45 блоків було зібрано з хмарного сховища та згруповано за типом каркасної мережі виділення ознак.

Потім модель логістичної регресії була навчена на моделях для кожної площини МРТ, щоб зважити рішення щодо класифікації кожної з них відповідно до істинних міток на тестовому наборі даних. Навчання LR

проводилося на локальній машині з використанням графічного процесора NVIDIA GeForce GTX 1080 Ti для прискорення процесу.

Після цього комбіновані прогнози моделі були зібрані в наборі даних перевірки. Відповідні вихідні дані моделі було збережено у файлах CSV для використання в оцінці моделі. Приклад такого передбачення показано у таблиці 2.2. Оціночні показники та графіки були побудовані на основі цих прогнозованих даних.

Таблиця 2.2. Приклад результату передбачення моделі для порівняння каркасних мереж виділення ознак на тестовому наборі даних

ID пацієнта	Ймовірність заданого діагнозу		
	Розрив ACL	Розрив меніска	Аномальна форма коліна
1335	0.06930	0.47787	0.39330
1336	0.14393	0.45038	0.04646
1337	0.36997	0.34413	0.22805

2.5. Підбір ансамблю каркасних мереж на основі результатів моделі порівняння ефективності

Закінчивши тренування базової архітектури MRNet на різних типах моделей комп'ютерного бачення в якості каркасних мереж виділення ознак, було отримано результати передбачення усіх моделей на тестовому наборі даних у вигляді ймовірності діагнозів розриву ACL, розриву меніска, або аномальної форми колінного суглоба.

Порівняння оцінок ефективності.

Як передбачено стратегією оцінювання, для кожної моделі та кожного типу діагнозу було отримано оцінки ефективності класифікатора, а саме тосність передбачення, міру F_1 , показник ROC-AUC, а також було обчислено каппу Коена.

Першим очевидним і значущим показником оцінки для даного завдання

класифікації є загальна точність класифікатора. У таблиці 2.3 показано порівняння точності між моделями, серед усіх трьох наявних діагнозів, додаючи середнє значення точності для них. Максимальні значення підсвічено жирним шрифтом.

Міра F_1 розрахована працювати з конкретними значеннями класу, а не ймовірностями цих класів, тому для отримання із моделей фактичних значень було застосовано перетворення через порогове значення 0.5. Іншими словами, результат діагностування вважався позитивним, якщо ймовірність – результат класифікації – перевищувала значення 0.5.

Таблиця 2.3. Показники точності передбачення, залежно від вибору каркасної мережі виділення ознак, із порогом 0.5

Діагноз	Alexnet	VGG11	VGG16	Resnet	Efficientnet
Аномальна форма коліна	0.825	0.858	0.842	0.858	0.850
Розрив ACL	0.683	0.792	0.808	0.583	0.550
Розрив меніска	0.70	0.733	0.750	0.608	0.583
Середнє	0.738	0.752	0.799	0.683	0.661

Іншим важливим показником оцінки є оцінка $F1$, оскільки вона добре відповідає фактичному дисбалансу набору даних. Даний тип оцінки балансує влучність та повноту для позитивного класу, який є домінуючим у цьому наборі даних.

У таблиці 2.4 наведено показник $F1$, обчислений для кожної навченої моделі та для кожного типу діагнозу. Вихідні ймовірності передбачення по кожному типу діагнозу були перетворені так само, як і для загальної точності моделі, використовуючи порогове значення 0.5. Максимальні значення підсвічено жирним шрифтом.

Таблиця 2.4. F_1 - міра, залежно від вибору каркасної мережі виділення ознак, із порогом 0.5

Діагноз	Alexnet	VGG11	VGG16	Resnet	Efficientnet
Аномальна форма коліна	0.900	0.917	0.909	0.917	0.912
Розрив ACL	0.486	0.713	0.753	0.167	0.000
Розрив меніска	0.673	0.628	0.717	0.299	0.324
Середнє	0.686	0.752	0.793	0.461	0.412

Далі відображаються робочі характеристики приймача. На рисунку 2.11 показано показники ROC-AUC для кожної моделі. Кожна панель відповідає певному типу діагнозу.

Я вважаю, що найбільш значущим показником оцінки ефективності каркасних мереж виділення ознак для цього дослідження є бал Каппа Коена. Через незбалансований характер набору медичних зображень MRNet, існує гіпотеза, що застосування підходу надійності оцінювачів в перспективі може виявити більш продуктивні моделі, котрі недооцінені традиційними метриками.

Знову ж таки, для перетворення прогнозів моделі з ймовірностей на фактичні класи використано порогові значення, або інтервали надійності. Для цього було вибрано поріг в межах від 0,5 до 0,9, що дозволяє оцінити ширший спектр поведінки моделей класифікації, не фіксуючись на одному значенні для порогу.

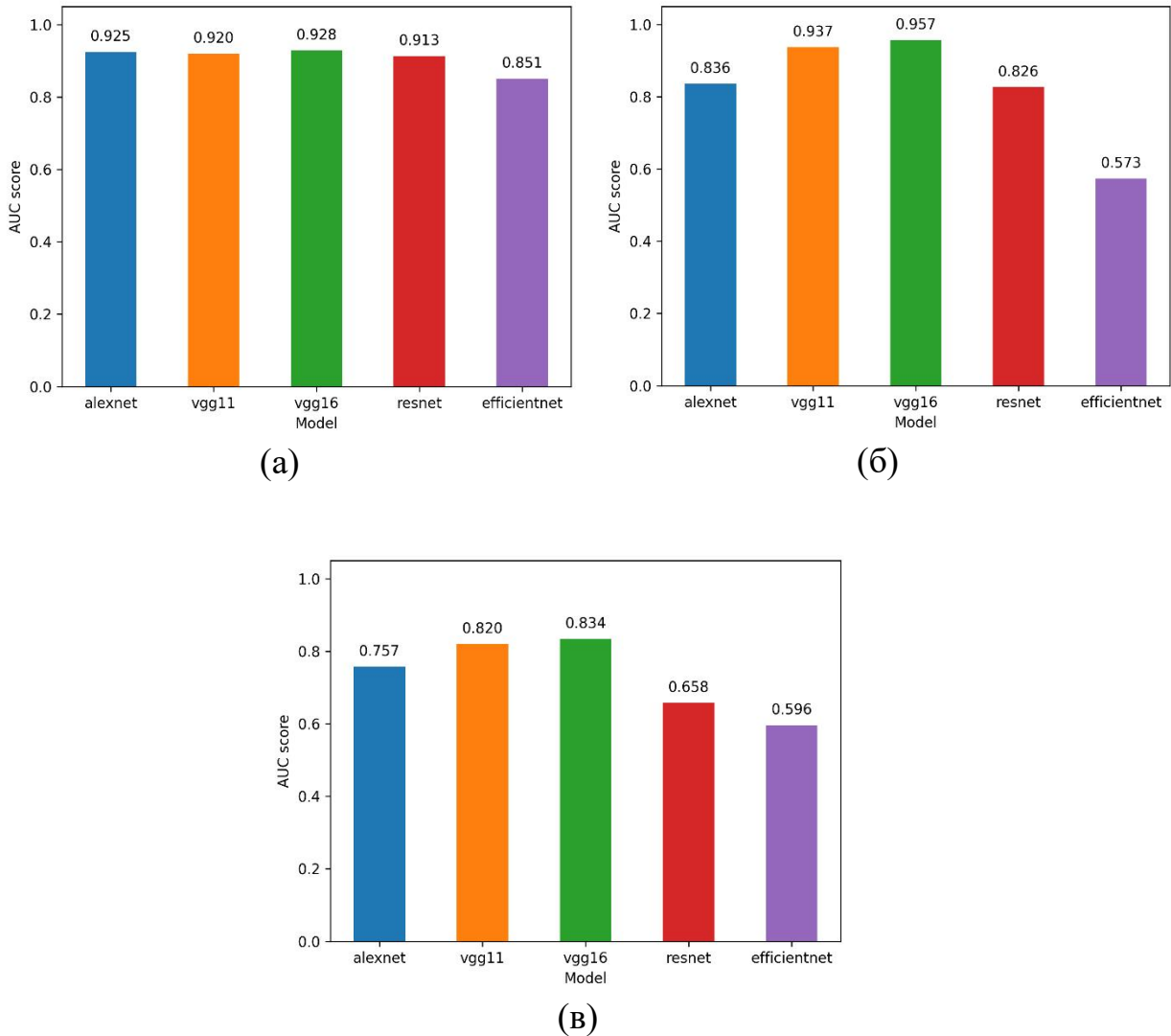


Рисунок 2.11. Показники ROC-AUC оцінки класифікації для кожного типу діагнозу: а – аномальна форма коліна, б – розрив ACL, в – розрив меніска

На рисунку 2.12 зображено графік значень Каппа Коена для заданих порогових значень.

Оцінка одного блоку для кожного типу каркасної мережі.

Ще одним дуже важливим аспектом оцінки цієї моделі є продуктивність кожної магістралі, перерахованої в цьому документі, на площину. Ці результати відображені в таблиці 2.5, і вони показують необроблений бал класифікації ROC-AUC для кожного блоку MRNet, навченого на певних видах каркасних мереж виділення ознак. Шар логістичної регресії не бере участі в цій оцінці, тобто ця таблиця показує саме продуктивність конкретної мережі виділення ознак на усіх трьох зрізах МРТ для всіх трьох завдань класифікації. Найвищий

бал оцінки для кожної площини та типу діагнозу виділено відповідним кольором площини МРТ в таблиці.

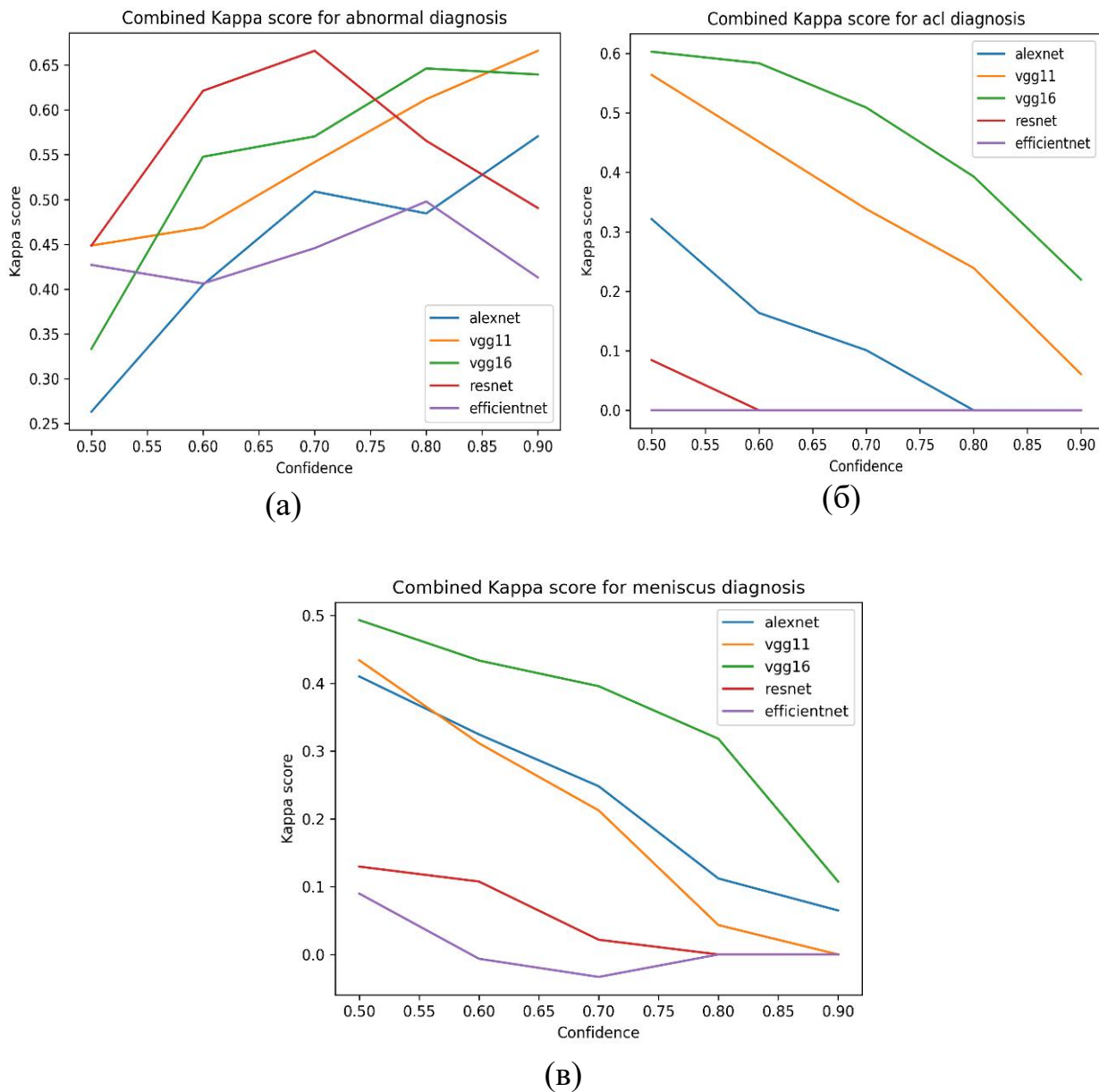


Рисунок 2.12. Показники каппи Коена для різних типів каркасних мереж та задач діагностування, при різних порогових значеннях: а - аномальна форма коліна, б - розрив ACL, в - розрив меніска

Таблиця 2.5. Показники ROC-AUC для кожного окремого блока MRNet, агреговані за типом каркасу виділення ознак, розрізом МРТ та завданням діагностування.

Каркасна мережа	Розріз	Аномальна форма коліна	Розрив ACL	Розрив меніска
Alexnet	axial	0.845	0.738	0.813
	coronal	0.788	0.616	0.633
	sagittal	0.935	0.834	0.707
VGG11	axial	0.892	0.758	0.811
	coronal	0.732	0.926	0.762
	sagittal	0.917	0.903	0.753
VGG16	axial	0.919	0.829	0.759
	coronal	0.840	0.950	0.808
	sagittal	0.909	0.883	0.801
Resnet	axial	0.848	0.763	0.618
	coronal	0.477	0.462	0.659
	sagittal	0.912	0.700	0.632
Efficientnet	axial	0.845	0.586	0.642
	coronal	0.667	0.593	0.535
	sagittal	0.723	0.631	0.591

Як видно з результатів виконання моделі оцінювання ефективності каркасних мереж виділення ознак, спостерігається прогресивне підвищення продуктивності класифікації у послідовності від Alexnet до VGG11 і VGG16. Цей результат означає, що навіть сама заміна мережі Alexnet аналогами VGG-* вже є вигідною з точки зору ефективності класифікації.

Іншим фактом, який впадає в очі, є порівняно низька ефективність класифікації, що показані найбільш сучасними архітектурами комп'ютерного

бачення із представлених.. Efficientnet та Resnet є найновішими моделями з усього набору, на якому проводилися оцінки. Вони, як правило, перевершують як моделі Alexnet, так і моделі VGG за результатами конкурсу ImageNet 2021 [58]. Однак для конкретного випадку класифікації МРТ колінного суглоба вони, як виглядає з наведених вище даних, не змогли використати весь свій потенціал. Ймовірно, це пов'язано з етапами попередньої обробки зображення, і, можливо, цю проблему можна вирішити за допомогою більш надійних методів попередньої обробки зображення.

Ще одним значним результатом є нерівномірний характер продуктивності каркасних мереж виділення ознак залежно від вибору зрізу МРТ.

Замість того, щоб дотримуватися одного типу каркасної мережі, доцільно використовувати ансамбль різних каркасних мереж виділення ознак, які будуть об'єднані за допомогою логістичної регресії: VGG16 на корональному розрізі для всіх завдань класифікації; в аксіальному розрізі для виявлення аномалії коліна та розриву ACL; Alexnet у сагітальному розрізі для виявлення аномалій та аксіальному розрізі для виявлення розриву меніска; VGG11 на сагітальному розрізі для виявлення розривів ACL (таблиця 2.5).

Результати також показують, що оцінка моделі з використанням підходу надійності між оцінювачами, а саме шляхом вимірювання метрики Каппа Коена може показати результати, відмінні від традиційних методів оцінки моделі. Наприклад, у виявленні аномалій, використовуючи ROC-AUC, модель Alexnet має майже таке ж значення, як і найбільш продуктивна модель з каркасною мережею VGG16, але використовуючи оцінку каппа Коена, Alexnet доволі сильно відстає за показником ефективності від всіх моделей, за винятком Efficientnet.

Крім того, ці показники допомагають виявити справжні проблеми застосування каркасної мережі Efficientnet для цього конкретного випадку — для класифікації меніска. Оцінка каппа Коена для моделі Efficientnet стала від'ємною в пороговому значенні 0,7, що можна інтерпретувати як модель, яка приймає неправильні рішення, які не є випадковими.

Проте з цього випливає логічне запитання - чому для порівняння ефективності каркасних мереж виділення ознак між розрізами МРТ не використовувалася показник каппа Коена, який, як зазначено раніше, є більш інформативним, ніж ROC-AUC через незбалансованість набору даних MRNet ? Справа в тому, що ROC-AUC було обрано в якості порівняння у цьому випадку для простоти у виборі найкращої каркасної мережі виділення ознак. Інакше довелось б додавати четвертий вимір для порівняння — порогове значення метрики каппа Коена. У випадку ж із ROC-AUC, цього додаткового показника не потрібно, що значно спрощує можливість відбору.

Серед можливостей майбутніх досліджень в плані порівняння каркасних мереж виділення ознак можна виділити три основні напрямки. По-перше, є можливість дослідити більш складні етапи попередньої обробки зображення. Багатообіцяючим є дослідження застосування техніки зменшення роздільної здатності зображення BlurPool, згаданої в [49] . Є можливість застосувати це в якості шару моделі разом із сучасними мережами комп'ютерного бачення для зменшення кількості параметрів навчання без значного зниження загальної продуктивності моделі.

Другий напрямок полягає у використанні оптимізованих версій більш просунутих моделей комп'ютерного зору для виділення ознак. Існують такі моделі, як Inception-v4 і Inception-Resnet, які можна використовувати як магістралі. Однак вони значно важчі з точки зору навчальних ресурсів, і тому ми ще не змогли застосувати ці новітні моделі до цього завдання класифікації.

Нарешті, але найважливіше, можна експериментувати з використанням блоків MRNet для кожної площини з різними магістралями та комбінувати їх за допомогою логістичної регресії. На нашу думку, одним із найбільш значущих результатів, отриманих у цій роботі, є той факт, що не всі моделі комп'ютерного зору поведуться однаково на різних площинах МРТ. Таким чином, доцільно використовувати найкращу можливу магістральну архітектуру для кожної площини МРТ замість того, щоб дотримуватися єдиної архітектури для всіх площин. Вибір цього параметра має підвищити загальну ефективність

комбінованої моделі класифікації LR.

Враховуючи результати роботи моделі порівняння ефективності каркасних мереж виділення ознак, було підібрано ансамбль із різних типів мереж на різних розрізах МРТ для різних завдань діагностування колінного суглоба. Зображення кінцевої архітектури моделі MRNet для даного дисертаційного дослідження можна побачити на рисунку 2.13.

На рисунку зображено три види блоків – MRNet-Alexnet, MRNet-VGG11 та MRNet-VGG16. Дані блоки відповідають архітектурі одного блоку, попередньо проілюстрованій на рисунку 2.9. Назва кожного блоку також відповідає типу каркасної мережі виділення ознак, що застосовується.

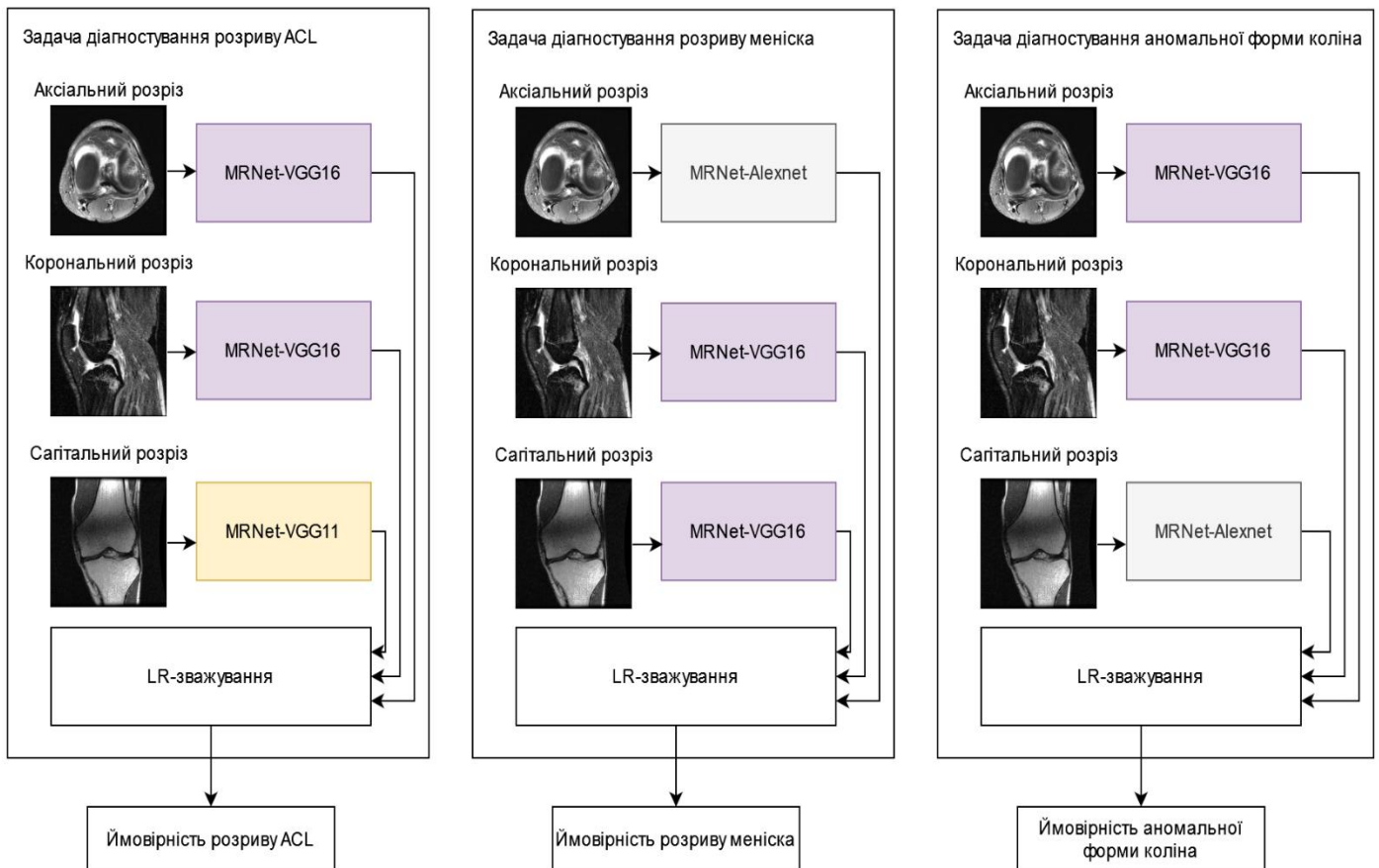


Рисунок 2.13. Результуюча архітектура моделі MRNet з ансамблем різних каркасних мереж виділення ознак

Висновки до розділу 2

У даному розділі було побудовано модель оцінки результатів передбачення мережі MRNet на основі різноманітних сучасній архітектур комп'ютерного бачення в якості каркасних мереж виділення ознак.

На основі зібраних даних було змодельовано кінцеву архітектуру ансамблю мереж-блоків MRNet для задачі автоматизованого діагностування знімків МРТ колінного суглоба у трьох розрізах.

Дослідження із цього розділу були опубліковані у працях автора [1], [2].

РОЗДІЛ 3. ПОКРАЩЕННЯ ЕФЕКТИВНОСТІ АНСАМБЛЮ МОДЕЛЕЙ. ПОБУДОВА ІНТЕРПРЕТАЦІЙНОЇ МЕРЕЖІ

У третьому розділі описано підхід підвищення точності передбачення на нових даних для моделі шляхом аугментації вхідних зображень за допомогою випадкових афінних перетворень, для запобігання перенавчанню моделі глибинного навчання. Також розроблено інтерпретаційну модель на основі методу латентного зсуву, для пояснення результатів передбачення, з метою підвищення прозорості діагностики, та простоти кінцевого аналізу лікарями-рентгенологами.

Результати розділу опубліковано в [3].

3.1. Стратегії запобігання перенавчанню

Як зазначено у роботі [59], загальноновизнаною є думка, що більші набори даних призводять до тренування кращих моделей глибинного навчання. Проте суттєвою є проблема того, що збір та обробка величезних наборів даних потребують колосальної кількості людських ресурсів. Завдання компонування великого набору даних є складним не тільки через рутинні зусилля зі збору та маркування даних, а також враховуючи той факт, що дані у такому наборі повинні бути збалансованими. В опрацюванні медичних зображень, проблема порівняно малих наборів даних набуває особливого поширення, оскільки медичні зображення є надзвичайно складним типом даних як для збору, так і для ручного маркування.

MRNet можна вважати малим набором даних, тому методи запобігання перенавчанню, перелічені у вищезгаданій статті, можуть виявитися дуже корисними для застосування у даному дисертаційному дослідженні. Застосувавши відповідні методи, можна покращити результати передбачення ансамблю моделей із попереднього розділу, допомігши їм краще узагальнити раніше невідомі їм дані.

Щоб запобігти перенавчанню моделей глибинного навчання, ми використовуємо підхід аугментації зображень. Це - метод синтетичного

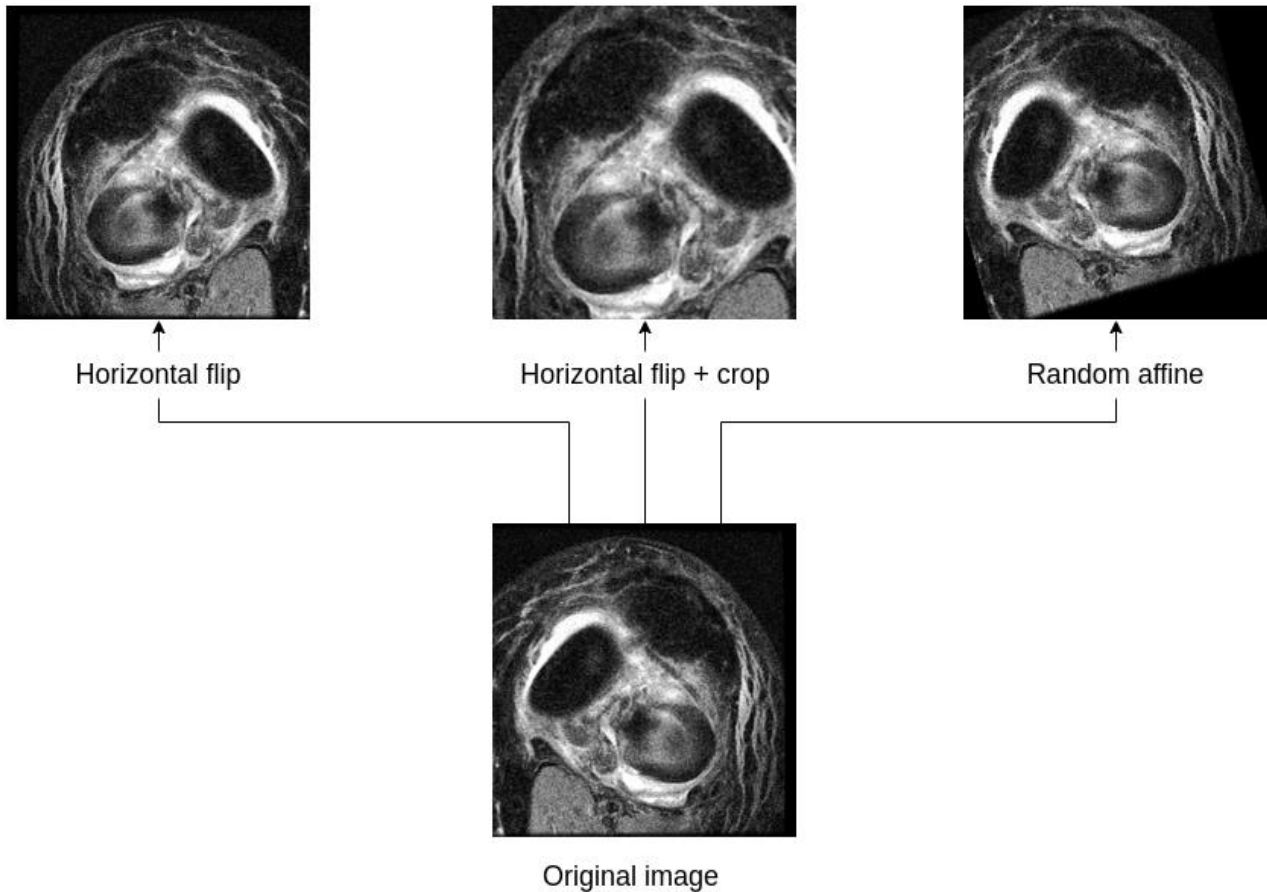


Рисунок 3.1. Алгоритм аугментації вхідних зображень на прикладі аксіального розрізу МРТ

збільшення кількості вхідних зображень шляхом застосування випадкових перетворень зображень.

Для вхідних зображень здійснюється випадкова обробка афінними перетвореннями, що збільшує оригінальний набір вхідних даних. Приклад різноманітних шляхів аугментації одного і того ж зображення МРТ колінного суглоба показано на рисунку 3.1. Це допомагає створити більш широкий і різноманітний вхідний сигнал для моделей машинного навчання, а також допомагає вирішити проблему перенавчання для цього застосування.

В якості основного алгоритму аугментації зображень використовуються афінні перетворення. Позначаючи x та y як координати точки (пікселя) вхідного зображення, їх перенесення до відповідних видозмінених координат x' , y'

відбувається за допомогою афінного перетворення:

$$\begin{aligned}x' &= a_0x + a_1y + a_2, \\y' &= b_0x + b_1y + b_2,\end{aligned}$$

де $a_0, a_1, a_2, b_0, b_1, b_2$ – дійсні коефіцієнти. Ці рівняння можна також виразити за допомогою матриці афінного перетворення:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}.$$

Для різноманітних описаних типів афінних перетворень відповідні матриці виглядатимуть наступним чином.

Горизонтальне віддзеркалення:

$$\begin{pmatrix} -1 & 0 & W \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Тут W – це ширина зображення у пікселях. Оскільки розглядаємо лише першу координатну чверть, і звичайне горизонтальне віддзеркалення посуває координати у другу чверть, додаємо ширину зображення в якості зсуву, аби це компенсувати.

Поворот:

$$\begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Кут повороту θ вибирається випадковим чином в межах від -25 до 25 градусів. Якщо значення пікселя виходить за межі оригінальної ширини чи висоти зображення, воно втрачається.

Показано, що застосування аугментації даних результує у підвищенні продуктивності моделей класифікації. Навіть традиційні способи аугментації зображення на основі набору даних, без залучення спеціальних додаткових мереж за типом Generative Adversarial Network (GAN) [60], які генерують зображення не на основі геометричних перетворень, є дуже ефективними, що описано у працях [61] та [62].

Таблиця 3.1. Порівняння точності класифікації для вхідного та розширеного наборів даних для підмоделі розриву ACL на сагітальному розрізі МРТ. N – загальна кількість зображень, K – метрика капа Коена, F_1 - оцінка, CI – довірчий інтервал

Тип набору даних	N	ROC-AUC	CI = 0,5		CI = 0,8		CI = 0,95	
			K	F_1	K	F_1	K	F_1
Розріз вхідного набору даних	100	0.863	0.566	0.949	0.736	0.965	0.095	0.455
Аугментований розріз вхідного набору даних	147	0.903	0.683	0.96	0.787	0.971	0.131	0.504
Зміна показника	-	1.046	1.207	1.012	1.069	1.006	1.379	1.108

Кут повороту випадкового повороту θ заданий у межах 25 градусів. Імовірність віддзеркалення. Імовірність додавання віддзеркаленого зображення P_{-1} становить 0.5. В процесі аугментації відбулося додавання 47 нових зображень, які є горизонтально перевернутими варіантами випадкових зображень із набору даних. Додатково, кожне із цих 147 зображень було випадково повернуто вліво або вправо на величину кута від 0 до 25 градусів.

Порівняння результатів до та після аугментації відбувалося для моделі, складеної з ансамблю каркасних мереж, визначеного у розділі 2 (рисунок 2.13). Модель тренувалася з нуля на вибраному розрізі набору даних. Передбачення та оцінки здійснювалися на валідаційному наборі даних MRNet.

Результати свідчать про покращення точності передбачення, згідно із метриками ROC-AUC, капи Коена та F_1 для всіх довірчих інтервалів. Середнє значення підвищення точності по всіх показниках становить 11.8%.

3.2. Інтерпретації рішень у автоматизованому діагностуванні. Існуючі алгоритми інтерпретації

Розуміння та інтерпретація класифікаційних рішень автоматизованих систем діагностування в медицині має велике значення, оскільки це дозволяє перевірити міркування системи та надає додаткову інформацію людині – фахівцю, котрому надаються дані для підтримки прийняття рішень. Незважаючи на те, що методи машинного навчання дуже успішно вирішують безліч завдань, у більшості випадків вони мають той недолік, що діють як чорна скринька, тобто не надають жодної інформації про те, яка частина вхідних даних вплинула на конкретне рішення класифікації [63].

Відсутність інтерпретабельності пов'язана з нелінійністю різних відображень, які обробляють пікселі зображення до його представлення як вектору ознак, і надають це представлення до кінцевої функції класифікатора. Це є значним недоліком у застосуванні методів машинного навчання до класифікації у реальному житті, оскільки це заважає фахівцям ретельно перевіряти рішення щодо класифікації. Проста відповідь «так» або «ні» іноді має обмежену цінність у застосуваннях, де питання розташування або структури є більш доречними, ніж проста бінарна оцінка. Крім цього, є багато інших аргументів на користь інтерпретабельного штучного інтелекту [64] , наприклад:

- **Перевірка коректності:** як згадувалося раніше, у багатьох застосуваннях, не можна безумовно довіряти системі типу чорної скриньки. Наприклад, у сфері охорони здоров'я використання моделей, які можуть бути інтерпретовані та перевірені медичними експертами, є абсолютною необхідністю. Існує приклад [65] , де система штучного інтелекту, яка була навчена передбачати ризик пневмонії у людини, приходять до абсолютно хибних висновків. У процесі навчання модель дізнається, що хворі на астму з проблемами серця мають набагато менший ризик смерті від пневмонії, ніж здорові люди. Лікар-фахівець би одразу помітив, що це не може бути правдою, оскільки астма та проблеми з серцем є факторами,

які негативно впливають на прогноз щодо одужання. Однак модель штучного інтелекту нічого не знає про такі явища як астма або пневмонія, а лише робить висновки з наявних даних. У цьому прикладі дані були систематично зміщені, оскільки на відміну від здорових людей більшість хворих на астму та серцево-судинні захворювання перебували під більш суворим медичним наглядом, ніж інші пацієнти.

- **Удосконалення.** Першим кроком до вдосконалення системи ШІ є розуміння її слабких сторін. Очевидно, що важче виконати такий аналіз слабких місць на моделях типу чорної скриньки, ніж на моделях, які можна інтерпретувати. Також легше виявити упередження в моделі або наборі даних (як у прикладі з пневмонією), якщо зрозуміти, що робить модель і чому вона приходить до своїх прогнозів. Можливість інтерпретації моделі може бути корисною при порівнянні різних моделей або архітектур. Деякі моделі можуть бути приблизно однаковими в плані якості класифікації, але їх архітектура передбачає суттєві відмінності між ознаками, що використовуються як основа для прийняття рішень. Це демонструє, що більш повні методи порівняння моделей вимагають інтерпретованості цих моделей. Також можна спробувати прослідкувати тенденцію, що краще розуміння поведінки моделей машинного навчання, дає більше простору до їх удосконалення.
- **Навчання на результатах передбачення:** оскільки сучасні системи штучного інтелекту навчаються на мільйонах прикладів, вони можуть спостерігати закономірності в даних, які раніше не були відомі широкому загалу оскільки люди здатні вчитися на значно меншій кількості прикладів. Використовуючи пояснювані системи штучного інтелекту, можна спробувати використовувати системи штучного інтелекту для отримання нових знань.
- **Відповідність законодавству:** системи ШІ впливають на все більше сфер нашого повсякденного життя. Разом з тим останнім часом підвищена увага приділяється також юридичним аспектам, наприклад, питанню

відповідальності за рішення класифікації, коли система приймає неправильне рішення. Оскільки неможливо знайти задовільні відповіді на ці юридичні питання, покладаючись на моделі типу чорної скриньки, майбутні системи ШІ обов'язково повинні стати інтерпретабельними [66].

На сьогоднішній день існує декілька способів побудови пояснень для систем автоматизованого діагностування. Найпростіші методи опираються на карту помітності, або карту салієнтності (англ. saliency maps). Приклад такої карти можна побачити на рисунку 3.2.



Рисунок 3.2. Приклад побудови карти помітності для вихідного зображення

Карту помітності можна розглядати як приклад сегментації зображення. У комп'ютерному баченні процес сегментації зображення – це поділ цифрового зображення на декілька множин пікселів, відомих як суперпікселі. Мета сегментації полягає в тому, щоб спростити представлення зображення до більш значущого і легшого для аналізу. Зазвичай, сегментація використовується для визначення місцезнаходження об'єктів та їх меж на зображеннях [67].

Згідно із останніми дослідженнями ефективності використання підходів на базі карт помітності для пояснення результатів класифікації зображення глибокими нейронними мережами, дані алгоритми не є достатньо повними, оскільки погано враховують зміни в параметрах чи архітектурі моделей,

результати класифікації яких вони намагаються пояснити.

У роботі [68] п'ять та два методи використання карт помітності (з восьми загалом) не пройшли рандомізований тест моделі на сегментаційному наборі даних та наборі виявлення об'єктів відповідно. Було зроблено припущення, що методи карт помітності не є чутливими до змін параметрів самої моделі. Повторюваність і відтворюваність більшості методів із використанням карт помітності були значно гіршими, аніж спеціально навчені підмережі локалізації, що було показано також на сегментаційному наборі даних, та наборі виявлення об'єктів.

У дослідженні [69] автори оцінювали ефективність кількох популярних методів карт помітності на наборі даних RSNA Pneumonia Detection [70] щодо можливостей локалізації, стійкості до параметрів моделі при рандомізації позначок класів, а також повторюваності та відтворюваності з різноманітними архітектурами моделей. Було виявлено, що метод GradCAM [71] показав чудову чутливість до параметрів моделі та рандомізації міток і був агностичним до архітектури моделі, в той час як традиційні методи на основі карт помітності не змогли показати суттєвих результатів у цьому.

Робота описує техніку для більш прозорих передбачень моделей на основі згорткових нейронних мереж, шляхом візуалізації вхідних областей, які є «важливими» для прогнозів, створюючи візуальні пояснення. Саме цей підхід було названо Gradient-weighted Class Activation Mapping, або ж Grad-CAM. Він використовує інформацію про градієнти класу для локалізації важливих регіонів. Ці локалізації поєднуються з існуючими візуалізаціями піксельного простору, щоб створити нову візуалізацію з високою роздільною здатністю та розрізненням класів під назвою Guided Grad CAM. Ці методи допомагають краще зрозуміти моделі на основі згорткових нейронних мереж, включаючи моделі підписів до зображень і візуальних відповідей на запитання. Grad-CAM пропонує принципово новий спосіб пояснення результатів передбачення. Принцип роботи даного алгоритму зображено на рисунку 3.3.

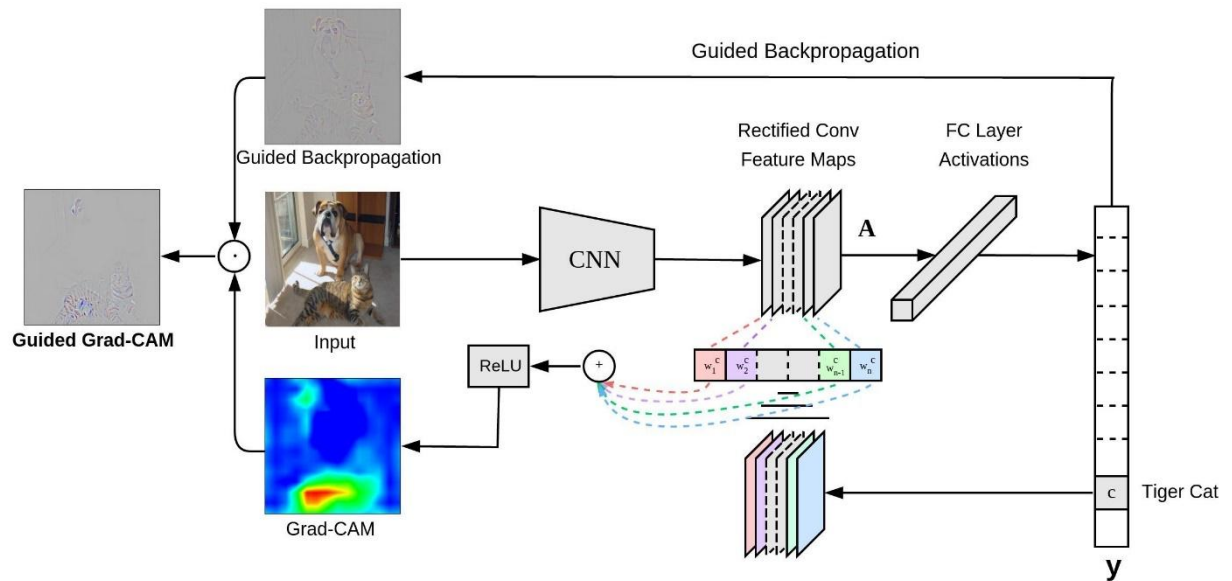


Рисунок 3.3. Огляд та візуалізація роботи алгоритму пояснення Grad-CAM [71].

Алгоритм працює наступним чином – маючи вхідне зображення та мітку категорії («тигровий кіт») як вхідні дані, зображення проганяється через згорткову нейронну мережу, щоб отримати необроблені оцінки кожного класу перед задіювання функції активації softmax. Для градієнту встановлюється початкове значення як нуль для всіх класів, за винятком мітки класу із вхідних даних, яка має значення 1.

Цей сигнал потім поширюється у зворотному напрямку на випрямлену карту ознак згорткової нейронної мережі, де обчислюється груба локалізацію Grad-CAM теплова карта на рисунку 3.3. Нарешті, відбувається поточкове множення теплової карти за допомогою керованого зворотного поширення, щоб отримати візуалізацію Grad-CAM, яка має високу роздільну здатність, і також розрізняє класи.

Підхід латентного зсуву.

Нещодавно ідея візуалізації прогнозів через перебільшення ознак, які впливають на передбачення моделі, була запропонована у роботі [72]. Це перебільшення є результатом здатності нейронної мережі створювати, або

“галюцинувати” нові ознаки [73] , і цей процес, як відомо, може бути контрольованим [74].

Замість того, щоб просто генерувати зображення певного класу, згадані методи перебільшення можуть пояснити конкретні особливості, які використовує класифікатор для кожного прогнозу. Це важливо для прогнозування, коли модель передбачає використання неправильних хибних корелятивів, щоб переконатися, що вона правильна з правильних причин [75]. У той час як більшість моделей прогнозування патології на зображенні передбачають причинно-наслідкові зв'язки, де конкретні області зображення явно призводять до застосування певної класифікаційної мітки (наприклад, збільшене серце \rightarrow кардіомегалія), моделі, що прогнозують ризик у майбутньому (наприклад, прогноз смертності протягом 5 років), не мають заздалегідь відомого причинно-наслідкового зв'язку. У таких сценаріях можна дізнатися, які саме ознаки можуть використовуватися в моделях, що здійснюють такі передбачення через перегляд згенерованого зображення із перебільшеними ознаками.

Однак, існують два основні недоліки існуючих підходів до цього завдання, які обмежують їх застосування. По-перше, більшість із них засновані на мережах GAN, навчання яких може бути дуже важким і трудомістким через підбір стабільної функції втрат і чутливість гіперпараметрів. По-друге, такі моделі є монолітними і вимагають спільного навчання генеративного і дискримінаційного компонентів, що перешкоджає роботі з існуючими попередньо натренованими моделями. Ідеальним варіантом для цього завдання був би модульний підхід, який є максимально простий для реалізації, і який здатен працювати з будь-яким існуючим класифікатором “з коробки”.

У роботі [76] описано підхід латентного зсуву, або модель латентної змінної, такої як простий автокодувальник $D(E(x))$, де E — кодувальник, а D — декодувальник, і також присутній класифікатор f , який робить передбачення цільового показника у наступним чином: $y = f(x)$. Модель латентної змінної та класифікатор навчаються незалежно без будь-яких особливих обмежень, окрім

того, що вони повинні бути диференційовними. Автокодувальник використовується з метою простоти навчання та програмної реалізації.

Коли моделі кодувальника та декодувальника є навчені, пояснення рішення класифікатора можна обчислити наступним чином:

Вхідне зображення x кодується за допомогою $E(x)$, створюючи його представлення у просторі латентної змінної z . Збурення латентного простору обчислюються для класифікатора f у рівнянні 3.1, який потім використовується для отримання λ -зміщених вибірок зображення, показаних у рівнянні 3.2:

$$z_\lambda = z + \lambda \frac{\partial f(D(z))}{\partial z}, \quad (3.1)$$

$$x'_\lambda = D(z_\lambda). \quad (3.2)$$

Очікується, що зображення x'_λ дасть вищий прогноз, такий, що

$$f(x'_\lambda) > f(x).$$

Звідси ми можемо генерувати кілька зображень x'_λ , щоб здійснити перебільшення або видалення ознак на зображенні, які призводять до певного прогнозу. Ці зображення можна з'єднати в короткі послідовності, які допомагають пояснити, чому було зроблено конкретне передбачення та які саме ознаки і в якій мірі використовувалися даним класифікатором.

Візуалізацію методу, описаного в роботі [76] показано на рисунку 3.4. При цьому підході також дуже важливо мати на увазі, що метод є обмежений простором латентного представлення ознак автокодувальника. Якщо декодувальник не є достатньо виразний, він не зможе правильно представити ознаки, які використовує класифікатор. Проте цей підхід дозволяє порівнювати декілька класифікаторів із фіксованим автоматичним кодувальником (або вибором моделі латентної змінної), що дозволяє чітко зрозуміти використання різноманітних ознак із зображення у моделях.

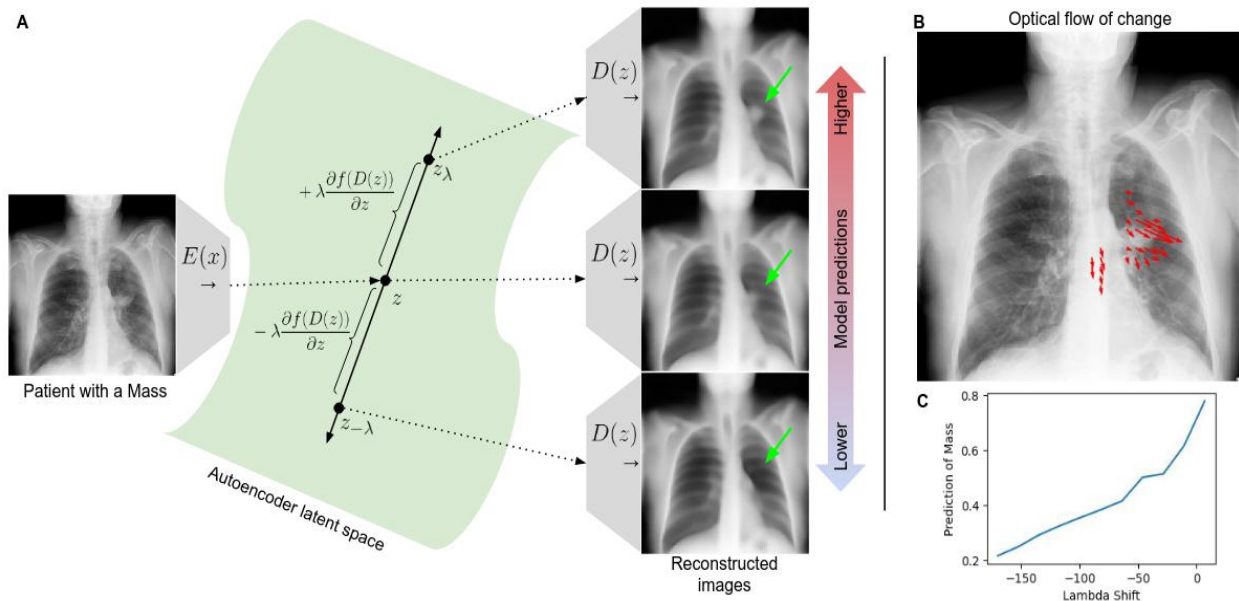


Рисунок 3.4. Візуалізація методу латентної змінної для побудови зображень із перебільшеними ознаками [37].

По суті, цей метод можна вважати протилежним до методу adversarial attack у GAN. Якби здійснювалася проста модифікація зображення за допомогою градієнта $\frac{\partial f(x)}{\partial x}$, що є традиційним підходом у adversarial attack, модифікація ознак була б непомітною та спотворювала б зображення шляхом вибору помилкових пікселів, які мали б вплив на цільову змінну.

У підході латентної змінної, процес є упорядкованим за допомогою фіксованого декодувальника, щоб запобігти зміні незначущих пікселів зображення. Загалом, у даному підході ставиться завдання змінити лише найбільш семантично значущі пікселі ознак, які призводять до конкретного результату класифікації.

3.3. Побудова інтерпретаційної моделі

Які проблеми було виявлено при побудові інтерпретаційної моделі у даному дисертаційному дослідженні ?

MPT можна вважати багатовимірними зображеннями, оскільки кожен розріз складається із набору вхідних зображень. Також, ці зображення розташовуються в строгій послідовності, інакше втрачається сенс обстеження з

точки зору просторових координат. Тому, власне генерація таких складних вхідних даних за допомогою GAN цілком може вписатися в рамки окремого дисертаційного дослідження, оскільки є дуже складною як в плані алгоритмізації, так і в плані наявності ресурсів для обчислень.

Чому? Навіть не враховуючи той факт, що контрафактні зображення повинні абсолютно не відрізнятися від оригіналів у більшості частин за винятком зон із ознаками, додатково потрібно забезпечити їх цілісність в розрізі. Іншими словами, згенероване зображення повинне лише мінімально відрізнятися від попереднього та наступного у розрізах.

Для того, аби мати змогу застосувати підхід генерації контрафактних зображень у даному дослідженні при побудові моделі пояснення було використано наступні спрощення:

1. Модель пояснення може працювати виключно із одним зображенням на одному розрізі МРТ. Зображення повинне знаходитися на відстані не більше ніж F розрізів від умовного центру серії зображень МРТ (у розробленій інформаційній системі передбачається перевірка цієї умови для вибору зображень для запуску моделі пояснення).
2. В якості вхідних даних для класифікатора, використовуються оригінальні дані, за винятком середини серії, де обране зображення викривлюється у просторі ознак дискримінатора із відхиленням λ , яке повторюється $2 \cdot F$ разів у центральній зоні серії.

Тут введено параметр F , котрий означає максимальну відстань від умовного центру серії зображень МРТ, на якій можна вибрати зображення для запуску моделі передбачення. За допомогою валідації моделі передбачення на вхідному наборі даних MRNet, було експериментально підібрано значення параметра $F = 3$. Схематично, принцип підстановки контрафактних зображень на вхід моделі зображено на рисунку 3.5.

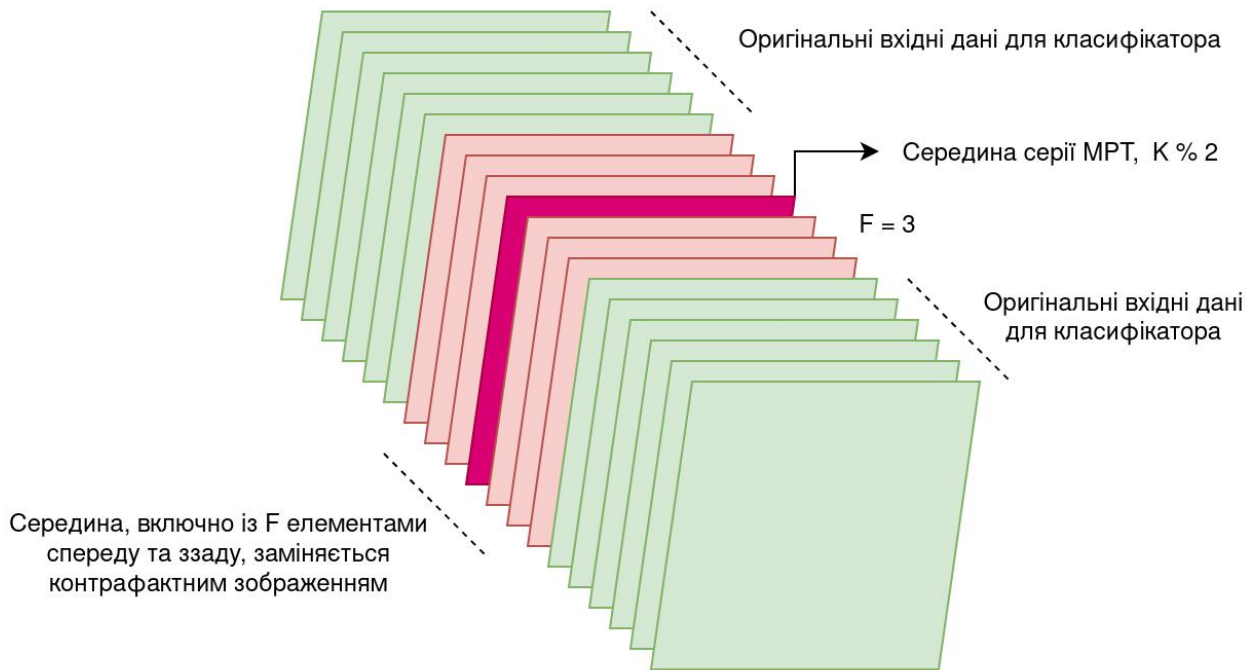


Рисунок 3.5. Схематика підстановки контрафактних зображень на вхід класифікатора для пояснювальної моделі, де K – кількість зображень у серії, F – параметризована максимальна відстань від центру серії до обраного зображення, на основі якого будується пояснення

Отже, оскільки на вхід до класифікатора подається середня зона серії зображень, яка дещо відрізняється від оригінальної, з метою спростити обчислення. Через це також відбувається і внесення похибок в результат передбачення.

Тому, пояснення не будуть ідеальними (якими вони могли би бути, враховуючи всю повноту вхідних даних). Проте у цьому випадку точністю доводиться в деякій мірі пожертвувати, задля можливості побудувати модель пояснення в цілому.

Яким чином можна визначити відхилення точності передбачення при процесах класифікації та пояснення? У даній роботі проведено експеримент, використовуючи в якості вхідних даних валідаційну частину набору даних

MRNet. При фіксації випадкових початкових значень, було обчислено ймовірність діагнозу на завданні класифікації для оригінальних вхідних даних, а також для даних із видозміненою середньою частиною серії зображень, використовуючи центральне зображення замість $F = 3$ суміжних з кожного боку, порівнюючи результат передбачення.

У результаті експерименту, на валідаційному наборі даних MRNet при значенні параметра $F = 3$ було отримано середнє відхилення результату передбачення 0.095623 що вважається допустимим для можливості побудови інтерпретаційної моделі. Параметр F було підібрано таким чином, щоб забезпечити найменшу похибку у разі використання підходу генерації контрафактних зображень, але при цьому маючи змогу виділити бажане центральне зображення із серії МРТ для більшості МРТ із валідаційного набору даних.

Врахувавши та прийнявши можливі неточності у передбаченні при побудові інтерпретаційної моделі, у даній роботі було здійснено тренування моделі на базі cGAN [74], як зазначено у роботі [37]. Аналогічно, як у згаданій роботі, маємо функцію пояснення, яка виглядає наступним чином:

$$L_f(x, \lambda): (X, R) \rightarrow X \quad (3.3)$$

Тут x — це вхідне зображення, для якого потрібно згенерувати пояснення, а λ — це додатній або від'ємний зсув, призначений для класифікатора f . Функція $L_f(x, \lambda)$ діє з декартового добутку простору зображень X та простору дійсних чисел R на простір вхідних зображень X .

Функція пояснення складається із кількох компонентів, включаючи згадані вище елементи автокодувальника — кодувальник $E(x)$ та декодувальник $G(z)$. У даному дисертаційному дослідженні, автокодувальник побудовано на базі ResNet [77]. Автокодувальник було навчено з нуля на базі середніх зображень МРТ із сагітальних розрізів тренувального набору даних MRNet, де найбільш помітно передню хрестоподібну зв'язку (ACL).

Функція втрат для автокодувальника складається із декількох частин, і

покликана задовільнити деякі обмеження на контрафактне зображення. По-перше, контрафактне λ — зміщене зображення повинно належати тому ж простору вхідних зображень X , що й оригінальне зображення. Для цього, зображення повинні мінімально відрізнятися одне від одного. Тому, функція втрат для декодувальника та дискримінатора виглядає наступним чином:

$$L_{cGAN} = E_{x,\lambda P(x,\lambda)} \cdot \quad (3.4)$$

Тут в якості P позначаємо розподіл параметрів для λ, x , а також z . Розподіл zP_z — це шум, взятий із нормального розподілу у просторі латентної змінної z .

Саме z — це латентне представлення вхідного зображення X , яке утворюється в результаті роботи кодувальника $E(x)$. E — це математичне сподівання по заданому розподілу.

На додачу до цього, семантично виділені регіони згенерованих зображень повинні бути співставними з оригіналом. Однієї лише функції втрат на основі подібності двох зображень є недостатньо для генерації досить подібних зображень, тому додатково натреновано модель семантичного виділення хрестоподібної зв'язки $S(x)$ на зображеннях МРТ. Приклад такої моделі взято із роботи [78].

Результат роботи моделі семантичного виділення враховується дискримінатором $D(\cdot)$ при кінцевому відборі кандидатів на зображення x_λ . Результат виконання моделі виділення хрестоподібної зв'язки показано на рисунку 3.6.

Тому, додатково задана функція втрат, що враховує співпадіння регіонів для оригінального та контрафактного зображення. Простими словами, ця частина загальної функції втрат «штрафує» генератор зображень, якщо у заданих зображеннях не співпадають регіони передньої хрестоподібної зв'язки, що виділені моделлю виділення $S(x)$.



Рисунок 3.6. Приклад роботи екстракції зони, де знаходиться хрестоподібна зв'язка із сагітального розрізу зображення МРТ [78].

$$L_{rec} = |p_{ij} - p_{ij}^{\lambda}|, (i, j) \in S(x), \quad (3.5)$$

де p_{ij} – це значення пікселя в двовимірному чорно-білому зображенні x , а p_{ij}^{λ} – у зображенні x_{λ} , відповідно. $S(x)$ повертає множину точок (i, j) , котрі включають в себе передню хрестоподібну зв'язку. Таким чином, маємо функцію втрат, що рахує абсолютні значення різниці пікселів на частині зображення, яка прокласифікована як передня хрестоподібна зв'язка. Інтуїтивно, цей регіон повинен бути розташований у тому самому місці на обох зображеннях.

Нарешті, класифікатор $f(x)$ повинен давати приблизно λ - зміщену ймовірність передбачення для зображень x та x_{λ} . Для цього, використано функцію втрат на основі розходження Кульбака — Лейблера [79]:

$$L_f(D, G) := r(f(x) + \lambda|x) + D_{KL}(f(x_{\lambda})|f(x) + \lambda).$$

Отже, функція пояснення в процесі свого навчання, маючи заданий автокодувальник $E(G)$, а також дискримінатор D , повинна мінімізувати наступну функцію втрат:

$$L = c_{cGAN} \cdot L_{cGAN}(D, G) + c_f \cdot L_f(D, G) + c_{rec} L_{rec}(E, G). \quad (3.6)$$

Загальна архітектура функції пояснення показана на рисунку 3.7.

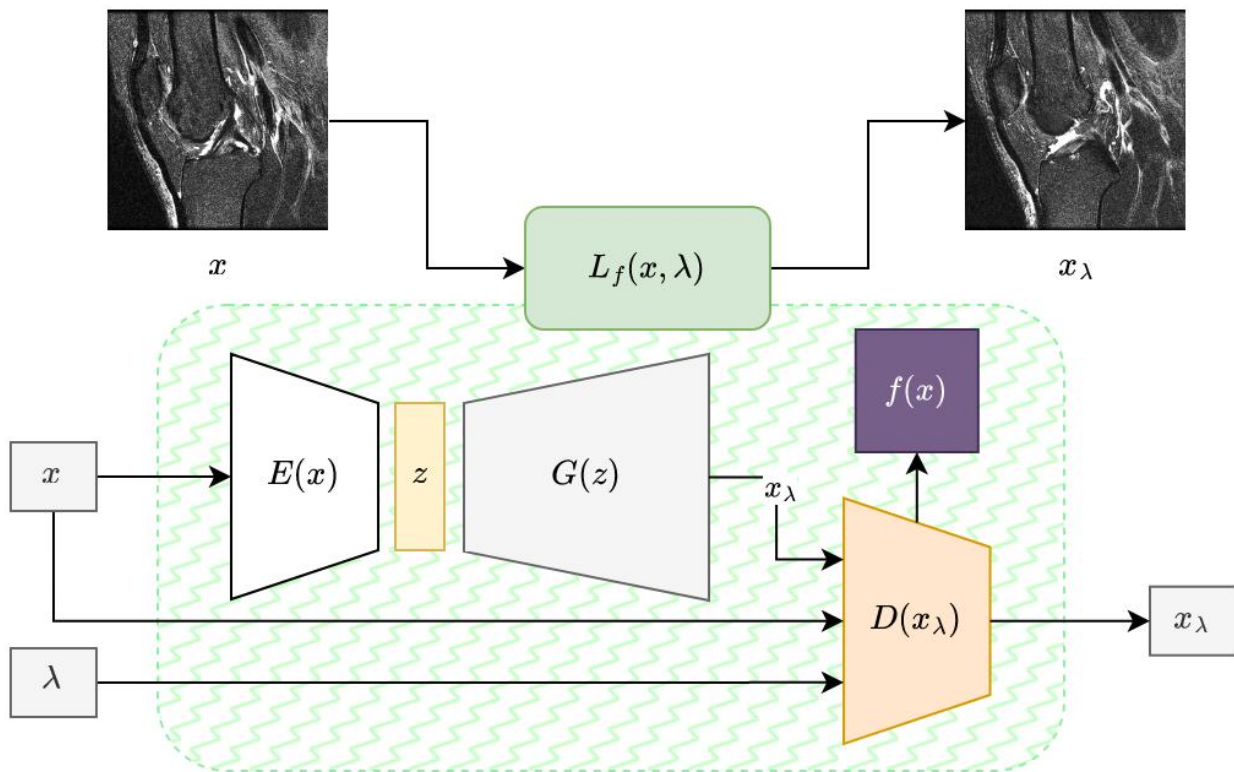



Рисунок 3.7. Архітектура пояснювальної функції $L_f(x, \lambda)$. Вхідне зображення x проходить через кодувальник $E(x)$, і формується його представлення у просторі латентної змінної z . Далі, декодувальник $G(z)$ здійснює реконструкцію вхідного зображення, та подає його на дискримінатор $D(\cdot)$, який вирішує, чи задовільняє згенероване зображення критеріям функцій втрат, використовуючи як саме зображення, так і функцію-класифікатор.

На відміну від мережі для автоматизованого діагностування, тут не розглядалися різноманітні варіанти каркасних мереж для пояснюваного навчання, оскільки це потребує значних затрат часу та обчислювальних ресурсів. Натомість, для елементів інтерпретаційної моделі було задіяно найбільш ефективні та прості в імплементації мережі з наявних.

У таблиці 3.2 показано роботу частини інтерпретаційної моделі, котра займається генерацією зображень на основі оригінального, використовуючи підхід латентної змінної із заданими значеннями зсуву λ .

Таблиця 3.2. Огляд контрафактних зображень для інтерпретаційної моделі, згенерованих на сагітальному розрізі для діагностування розриву ACL.

Зображення	λ	$f(x_\lambda)$	$ F(x_\lambda) - F(x) - \lambda $
	-	0.17803	-
<p>Оригінальне зображення. Середній розріз МРТ, який прокласифіковано моделлю автоматизованого діагностування як малоймовірний з точки зору розриву ACL. На малюнку видно візуально неушкоджену передню хрестоподібну зв'язку.</p>			
	0.2	0.35100	0.02704
<p>Згенероване контрафактне зображення із заданим $\lambda = 0.2$. Видно незначне потоншення лівої частини хрестоподібної зв'язки, на що класифікатор відреагував збільшенням ймовірності розриву.</p>			
	0.5	0.59225	0.08577
<p>Згенероване контрафактне зображення із заданим $\lambda = 0.5$. Видно суттєве видозмінення лівої частини зв'язки, а також незначне видозмінення правої частини. Модель класифікації відреагувала відповідно, хоча із значною похибкою.</p>			
	0.8	0.94476	0.03327
<p>Згенероване контрафактне зображення із заданим $\lambda = 0.8$. Видно, що на зображенні практично відсутня ліва частина хрестоподібної зв'язки.</p>			

Для того, аби побудувати карту помітності на основі цих даних, у цій роботі використано алгоритм піксельної відмінності між двома зображеннями у чорно-білому спектрі.

Беруться два вхідних зображення (оригінальне зображення x , та контрафактне із найбільшим значенням λ - $x_{\lambda=0.8}$ конвертується у матрицю інтенсивності пікселів розміром 256×256 , і для кожного пікселя обчислюється абсолютне значення різниці інтенсивності I_D між даними зображеннями:

$$x = \begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,3} & \cdots & p_{1,256} \\ p_{2,1} & p_{2,2} & p_{2,3} & \cdots & p_{2,256} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{256,1} & p_{256,2} & p_{256,3} & \cdots & p_{256,256} \end{bmatrix}$$

$$x_{\lambda} = \begin{bmatrix} p_{1,1}^{\lambda} & p_{1,2}^{\lambda} & p_{1,3}^{\lambda} & \cdots & p_{1,256}^{\lambda} \\ p_{2,1}^{\lambda} & p_{2,2}^{\lambda} & p_{2,3}^{\lambda} & \cdots & p_{2,256}^{\lambda} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{256,1}^{\lambda} & p_{256,2}^{\lambda} & p_{256,3}^{\lambda} & \cdots & p_{256,256}^{\lambda} \end{bmatrix}$$

$$I_D = (d_{ij}), d_{ij} = |p_{ij}^{\lambda} - p_{ij}|, i = \overline{1,256}, j = \overline{1,256}$$

Абсолютне значення різниці інтенсивності потім кодується у нове зображення розміром 256×256 , і використовується в якості пояснювального зображення.

Приклад даного алгоритму візуалізації відмінностей між оригінальним та контрафактним зображеннями можна бачити на рисунку 3.8.

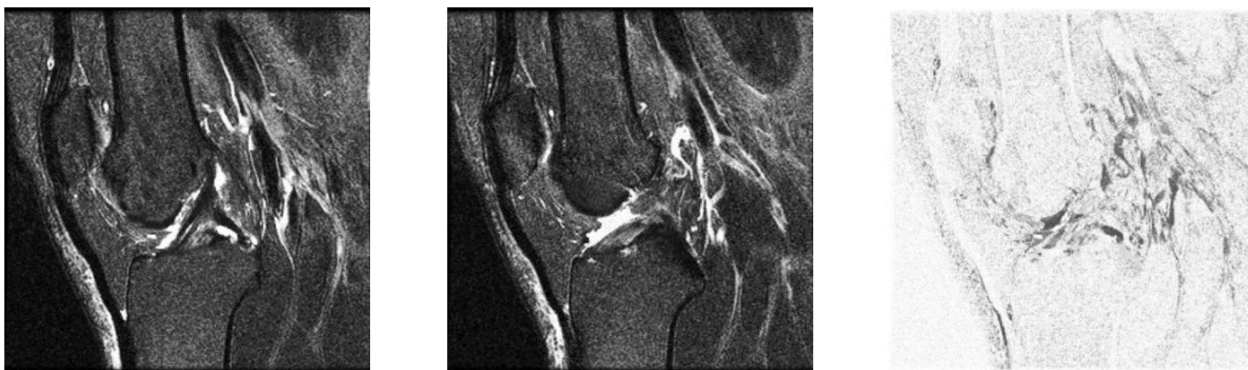


Рисунок 3.8. Візуалізація алгоритму створення пояснювального зображення(праворуч) для вхідного малюнку(ліворуч), на основі максимально контрафактного малюнку(по центру)

Висновки до розділу 3

У даному розділі детально описано розробку методу підвищення точності передбачення та усунення перенавчання для моделей глибинного навчання, що використовуються для задач класифікації у автоматизованому діагностуванні МРТ колінних суглобів. За основу використано метод випадкових афінних перетворень.

Розроблено інтерпретаційну модель для задачі діагностування ушкоджень МРТ коліна на основі методу зсуву латентної змінної. Методологія інтерпретації полягає у генеруванні контрафактних зображень з подальшою їх параметризованою підстановкою у оригінальну серію зображень МРТ.

Було підібрано оптимальні параметри для вставки з метою максимально зберегти вхідну точність класифікатора із максимальною похибкою передбачення між оригінальними та контрафактними серіями в межах 0.096.

РОЗДІЛ 4. ПОБУДОВА АРХІТЕКТУРИ ТА АПРОБАЦІЯ РЕЗУЛЬТАТІВ

У четвертому розділі зосереджено увагу на проблемі зменшення вартості розгортання складних інформаційних систем для автоматизованого діагностування ушкоджень колінного суглоба за МРТ. Побудовано архітектуру інформаційної системи на основі безсерверних обчислень, здійснено порівняння вартості розгортання для традиційної серверної архітектури та запропонованого рішення.

Результати розділу опубліковано у працях [5, 6].

4.1. Побудова архітектури інформаційної системи.

У даному дисертаційному дослідженні передбачається створення інформаційної системи для аналізу знімків МРТ. Було обрано найпоширеніший на даний момент тип ІС – веб-застосунок, із застосуванням прикладного інтерфейсу програмування (API).

Першочерговою перевагою такого підходу є крос-платформовість. Даний тип застосунку позбавляє проблем із портуванням інформаційної системи на різні типи операційних систем (MS Windows, Linux, MacOS). На додачу до цього, наявність вищезгаданого прикладного інтерфейсу програмування дозволяє в майбутньому інтегрувати дану систему на рівень кінцевих споживачів – а саме комерційних медичних інформаційних систем у якості розширюваного модуля.

Для побудови архітектури, потрібно врахувати елементи та бізнес-процеси, що будуть відбуватися в межах нашої інформаційної системи. Їх можна поділити на ключові, або основні – це ті, що невіддільно пов'язані із функціонуванням самої системи, а також другорядні – це додаткові вимоги, що вписуються у контекст її розгортання та функціонування у виробничому середовищі.

Серед **ключових вимог** можна виділити :

1. Користувач системи має мати змогу завантажувати файли МРТ-обстежень

колінного суглоба пацієнтів.

2. Система повинна вміти перетворювати вхідні файли у форматі DICOM на серії зображень по кожному розрізу МРТ, зробивши коректне сортування на розрізах.
3. Користувач повинен мати змогу запустити автоматизоване діагностування для попередньо завантаженого набору файлів DICOM, отримавши результати роботи моделі передбачення у вигляді прогнозованих ймовірностей трьох діагностичних завдань – аномальної форми коліна, розриву передньої хрестоподібної зв'язки (ACL) та розриву меніска.
4. Одночасно із результатами прогнозування по кожному діагнозу, користувач повинен мати змогу побачити результат виконання моделі пояснювання для передбачень вихідної моделі, аби при потребі сфокусуватися на ділянках зображення, що вплинули на передбачення моделі для кожного діагнозу.

Серед **другорядних вимог** до системи можна виділити :

1. Користувач системи повинен мати змогу автентифікуватися та авторизуватися.
2. Користувач повинен мати доступ до попередніх обстежень та переглядати вихідні файли та результати автоматизованого діагностування для цих обстежень, а також результати роботи пояснювальної моделі.

Враховуючи відносну складність даної системи, а також необхідність масштабування лише окремих її складових за підвищеного навантаження на систему, як архітектурну парадигму для побудови такої системи було обрано мікросервіси.

Мікросервіси — архітектурний стиль, або парадигма, за яким єдиний застосунок будується як сукупність невеличких сервісів, кожен з яких працює у своєму власному процесі та спілкується з рештою, використовуючи прості та швидкі протоколи передачі даних, зазвичай HTTP. Ці сервіси будуються навколо бізнес-потреб і розгортаються незалежно один від одного з використанням повністю автоматизованого середовища [80].

Існує абсолютний мінімум централізованого керування цими сервісами. Самі по собі вони можуть бути написані з використанням різних мов програмування і технологій зберігання даних.

Для програмної реалізації інформаційної системи аналізу медичних даних було обрано мікросервісну архітектуру із декількох причин. В довільний момент часу навантаження на вузли системи не буде однорідним. API для інтерфейсу користувача споживатимуть значно менше системних ресурсів, ніж, скажімо, запуск моделі глибокого навчання для аналізу даних, або ж процес парсингу вихідних файлів DICOM та конвертація зображень у формат, необхідний для роботи моделі передбачення.

Також слід врахувати фактор неоднорідності запитів до такої системи. Сам процес екстрації зображень та класифікації МРТ буде викликатися значно рідше, ніж перегляд історії сканувань для певного пацієнта. Тому також слід врахувати неоднорідність навантаження на кожен вузол при плануванні бюджету розгортання даної системи.

Враховуючи задані вимоги, у цій роботі було виділено наступні архітектурні компоненти, або сервіси для інформаційної системи:

- Сервіс **SystemManager** для підтримки автентифікації та авторизації користувачів, а також збереження налаштувань системи.
- Сервіс для завантаження та перетворення файлів DICOM у серії зображень МРТ по кожному розрізу – **ImageRepository**.
- Сервіс **PatientScans**, що відповідає за зберігання інформації про наявні сканування пацієнтів у системі, а також виступає контролером логіки передбачення та пояснення для кожного сканування.
- Сервіс **MRNet-Combined**, що надає API – доступ до моделі автоматизованого діагностування колінного суглоба, що була розроблена у розділі 2.
- Сервіс **ExplAI** – що надає API-доступ до пояснювальної мережі глибокого навчання, що була розроблена у розділі 3.
- Сервіс **Dashboard** – це власне, веб-сторінка, що виконує роль інтерфейсу

користувача для інформаційної системи.

- Супутні сервіси та технології:
 - **API Gateway** [81], або «Ворота API» – вхідна точка для запитів до інформаційної системи. Цей сервіс було використано для спрощення розробки власне сервісів. Таким чином можна позбавити внутрішні сервіси від необхідності перевіряти автентифікацію користувача, делегуючи це завдання до «Воріт».
 - **Amazon SQS** [82] – сервіс черги для забезпечення надійності процесів передбачення та екстракції зображень із DICOM. Дані процеси можуть вважатися порівняно довготривалими, тому для більшої надійності варто розглядати їх у контексті патерну проектування Master-Worker [83].
 - **Amazon RDS** [84] – сервіси баз даних для зберігання інформації про результати передбачень для кожного зображення, а також мета-інформації про завантажені файли DICOM.
 - **Amazon S3** [85] – сервіс для зберігання вхідних файлів, розпакованих зображень, а також згенерованих пояснювальною мережею зображень.

На дуже загальному рівні, архітектуру інформаційної системи зображено на рисунку 4.1.

Враховуючи неоднорідність навантаження на вузли системи, було прийнято рішення в рамках побудови інформаційної системи використовувати сервіс AWS Lambda.

AWS Lambda — безсерверна обчислювальна платформа, що керована подіями. Вона надається Amazon як один із веб-сервісів. Це обчислювальна служба, яка запускає код у відповідь на події і автоматично керує обчислювальними ресурсами, необхідними для виконання цього коду. Сервіс був представлений у листопаді 2014 року [86].

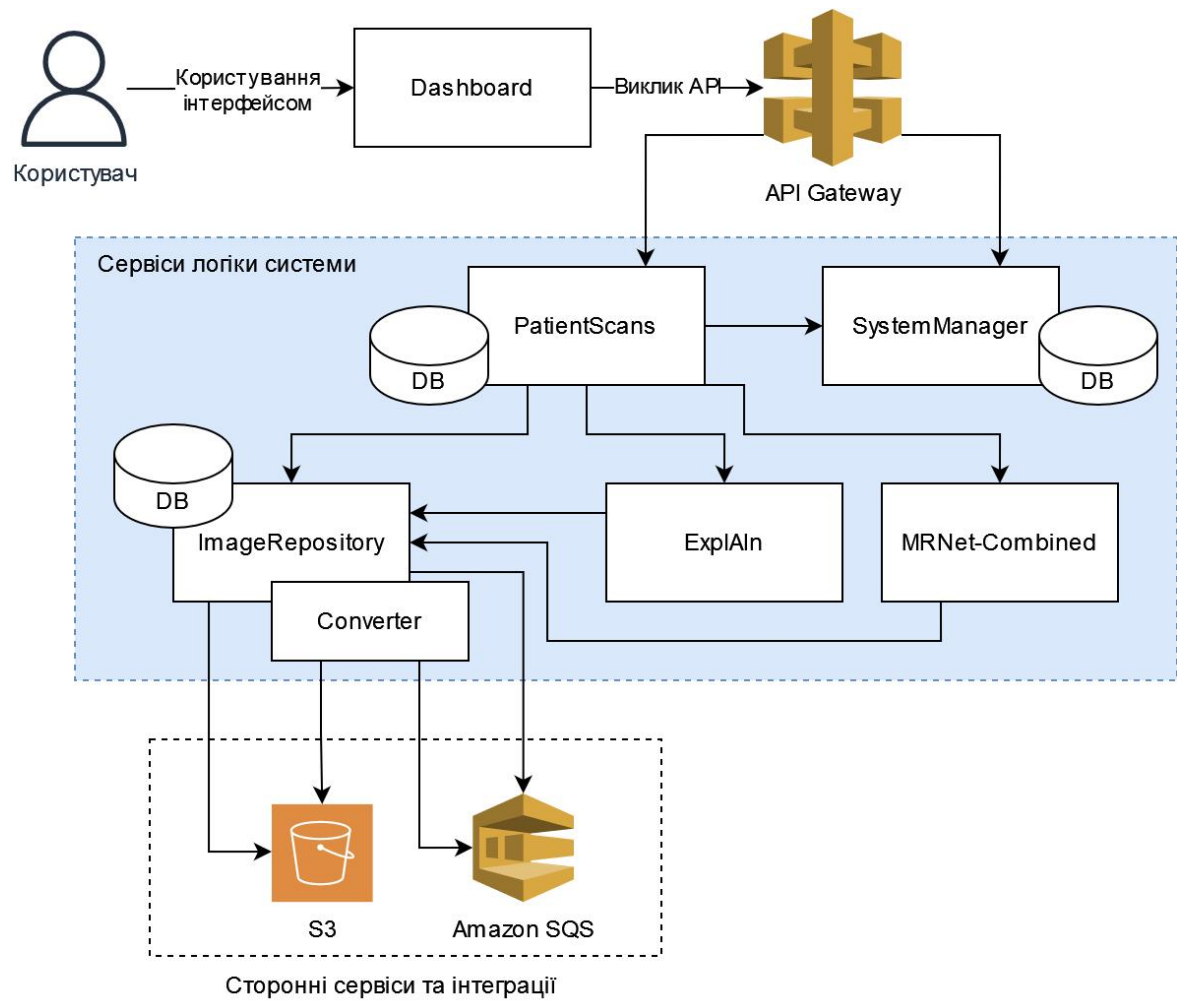


Рисунок 4.1. Схематичне зображення архітектури інформаційної системи автоматизованого діагностування зображень МРТ колінного суглоба із позначенням окремих сервісів

Порівняно з сервісом для традиційного хостингу багатоцільових серверів AWS EC2 призначення сервісу Lambda — спростити створення невеликих додатків, які виконуються на вимогу, як реакція на події або надходження нової інформації. AWS виконує запуск екземпляра Lambda протягом кількох мілісекунд після того, як сталася подія.

Архітектурне планування відбувалося з урахуванням універсальності кожного компоненту системи. Ключовими вимогами для кожного компоненту системи є незалежність та універсальність, як в плані роботи із даними, так і в плані розгортання. Це дуже важливо, оскільки не всі потенційні клієнти даної

системи захочуть використовувати її цілком.

Першочерговою категорією потенційних користувачів системи є медичні заклади. Багато медичних закладів уже використовують готові медичні інформаційні системи (МІС) [87] для обліку пацієнтів та інших процесів всередині закладу. Саме тому при розробці інформаційної системи автоматизованого діагностування, принципово не планується функціоналу, пов'язаного із обліком пацієнтів, або інших властивостей, котрі вже є на ринку медичних інформаційних систем в Україні. Такий підхід покликаний спростити інтеграцію із МІС медичних закладів в майбутньому.

4.2. Моделювання даних у інформаційній системі

Під час моделювання даних у системі, потрібно врахувати декілька ключових вимог до функціоналу, що вимагають зберігання інформації.

По-перше, потрібно мати місце для зберігання власне зображень МРТ. Найкраще для цього застосування підходить хмарне сховище S3 – частина пакету сервісів Amazon Web Services. Воно дає можливість зберігати будь-які види файлів, а також надає послуги хостингу цих файлів та контролю доступу на вимогу користувача. Перевагами використання S3 у такому випадку будуть:

1. Відсутність потреби у наперед зарезервованому дисковому просторі, адже S3 є динамічно розширюваним.
2. Ефективність використання коштів на розгортання, оскільки плата за користування сервісом S3 стягується лише за ту кількість дискового простору, що використовується. Окрім цього, можна розгорнути автоматизоване розділення даних на робочі та архівні. Плата за дані, що не потребують частого доступу, є значно нижчою. Автоматизовану процедуру зменшення використання ресурсів можна підключити через технологію Amazon S3 Intelligent-Tiering [88].

3. Відсутність потреби розробки та розгортання додаткового шару для хостингу зображень. S3 дозволяє не тільки розподіляти права на об'єкти всередині сховища, а й генерувати загальнодоступні посилання на них за допомогою технології Pre-signed URLs. Цей підхід проілюстровано на рисунку 4.2.

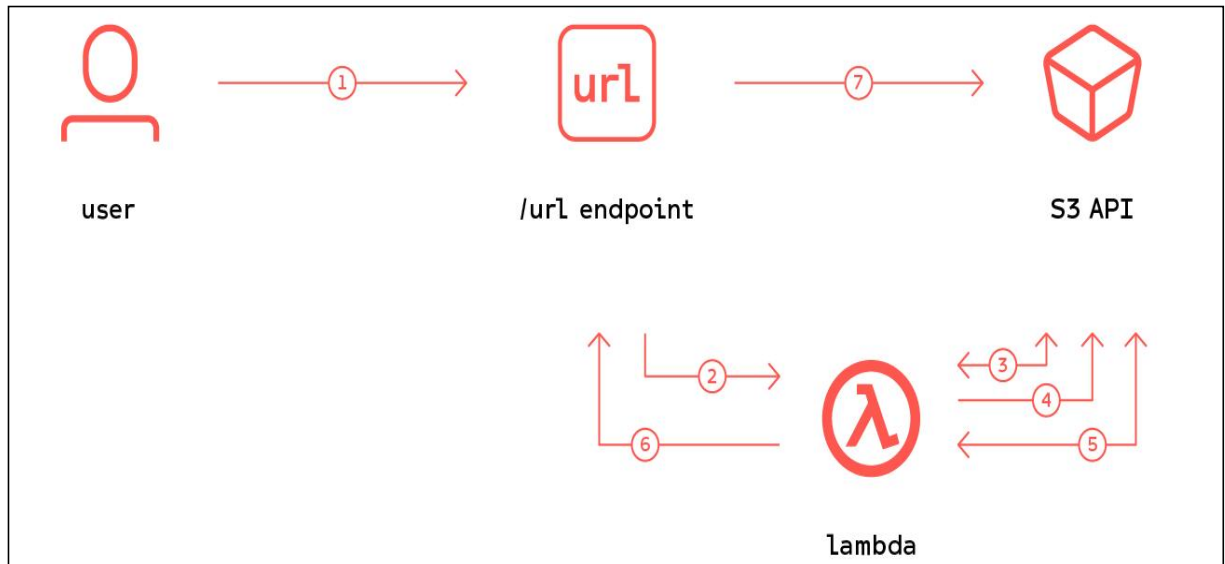


Рисунок 4.2. Використання технології Pre-signed URLs для взаємодії користувача та S3. У даному випадку, під позначанням Lambda буде сервіс ImageRepository, котрий для кожного зображення в серії буде генерувати свою попередньо підписану URL-адресу [<https://www.serverless.com/blog/s3-one-time-signed-url/>].

По-друге, потрібно подбати про зберігання супутніх даних сканувань. Сюди входить інформація про саме сканування, зображення, які вдалося витягнути (попри те, що зображення зберігатимуться у сховищі S3, потрібно подбати про те, аби тримати зв'язок між об'єктами S3 та фактичним скануванням). S3 виступає зовнішнім сховищем даних. Окрім цього, оскільки система є розподіленою, передбачається декілька різнорідних кластерів баз даних, і не можна цілком покладатися на вбудовані механізми цілісності у базі даних, такі як зовнішні ключі.

Під час планування архітектури на стадії моделювання даних та зв'язків у системі, було прийнято рішення розділити бази даних для сервісів обстежень

PatientScans, сервісу ImageRepository (рисунок 4.3), а також сервісу SystemManager. Це рішення було прийнято з огляду на декомпоновність цих компонентів, адже в майбутньому, це позбавить клієнтів від необхідності використовувати систему виключно в цілому, давши змогу зосередитися на функціональності, котра дійсно необхідна для кожного випадку.

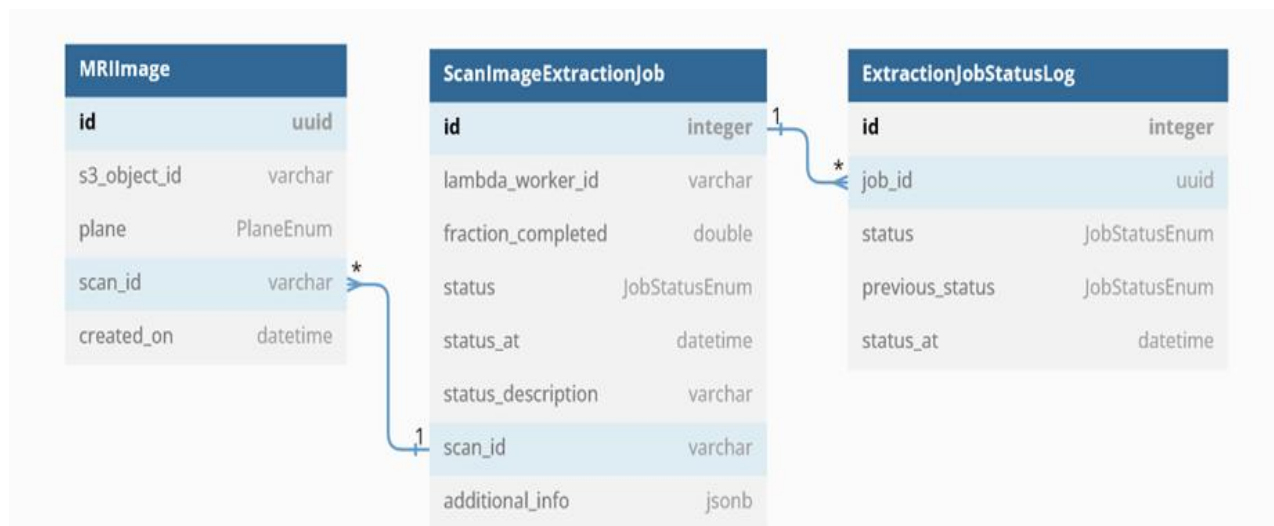


Рисунок 4.3. Діаграма зв'язків між сутностями у БД сервісу ImageRepository

Отож, в рамках розробки було створено три окремі бази даних, зі схемами зв'язків між сутностями, що зображені на рисунках 4.3 та 4.4.

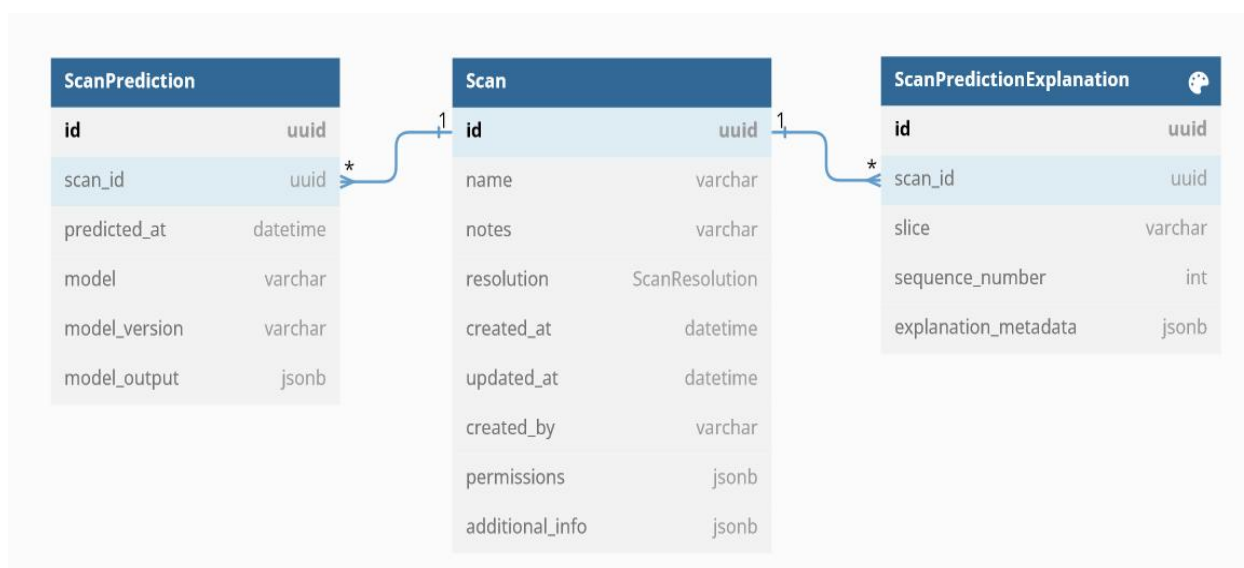


Рисунок 4.4. Діаграма зв'язків між сутностями у базі даних сервісу PatientScans

Як можна бачити, на рисунках наведено опис таблиць та зв'язків, що використовуються у базах даних кожного сервісу. Як було зазначено раніше, перехресні зв'язки між базами даних забезпечуватимуться за допомогою програмного коду на рівні самих сервісів. У якості СУБД для системи використовується PostgreSQL, на двигуні Amazon RDS. Діаграма зв'язків подана на рисунку 4.5.

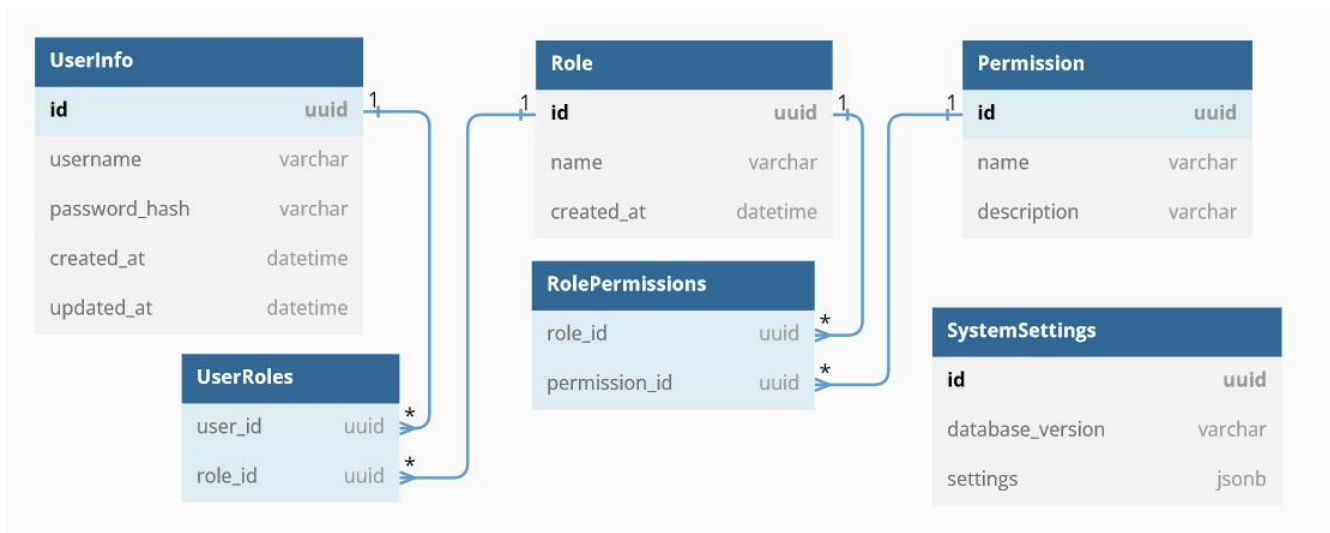


Рисунок 4.5. Діаграма зв'язків між сутностями у базі даних сервісу SystemManager

4.3. Опис бізнес-процесів всередині інформаційної системи

Для кращої деталізації процесів, що відбуваються всередині інформаційної системи, нижче наведено декілька UML-діаграм послідовностей, що описують взаємодію користувача та системи для здійснення типових завдань.

Завантаження файлів DICOM у систему.

Процес завантаження файлів DICOM в систему починається із авторизованого користувача, котрий на головній сторінці сервісу Dashboard натискає на кнопку “Імпортувати обстеження”. Користувач потрапляє у діалогове вікно, де йому пропонується обрати тип файлу для завантаження (за замовчуванням, на даний момент доступна лише єдина опція DICOM), де користувач вибирає файл, та вводить назву для сканування, і натискає на кнопку

«Завантажити».

Після цього, в системі відбуваються наступні події:

1. Сервіс PatientScans створює запис у базі даних обстежень, куди записує ім'я обстеження та супутню інформацію – назву завантаженого файлу, час створення обстеження та інформацію від сервісу SystemManager про користувача та права на об'єкт обстеження. Дана інформація передається у вигляді заголовків до HTTP-запитів, котрі проходять крізь проксі-сервіс APIGateway, котрий в свою чергу відповідає за автентифікацію користувачів через сервіс SystemManger.
2. Оскільки сервіс PatientScans не є відповідальним за екстракцію зображень DICOM, він робить запит до сервісу ImageRepository, на екстракцію зображень для даного обстеження, передаючи ідентифікатор поточного запису у БД PatientScans для підтримки зв'язності даних між сервісами, а також вихідний файл DICOM.
3. ImageRepository створює новий об'єкт ScanImageExtractionJob – Вихідний файл DICOM завантажується на Amazon S3 для можливості подальшого зневадження, а також в якості резервного джерела даних у випадку неочікуваних проблем. Сервіс зберігання зображень також здійснює кодування вхідного файлу Base64 [89] та надсилає його разом із ідентифікатором створеного об'єкту ScanImageExtractionJob до черги на сервісі Amazon SQS.
4. При додаванні елемента у чергу Amazon SQS активується подія, котра запускає під-сервіс ImageRepository. Converter, який розкодує даний файл із черги, та перетворює його на послідовності зображень із відповідних розрізів MPT. На виході процесу екстракції зображень матимемо набір зображень, кожне із яких буде прив'язано до вихідного обстеження через записаний ідентифікатор обстеження у моделі ImageExtraction, а також до конкретного розрізу MPT.

Схематика даного процесу зображена у двох стадіях (початкове завантаження та подальша екстракція з файлу), у вигляді UML діаграм

послідовностей, які можна бачити на рисунках 4.6 та 4.7.

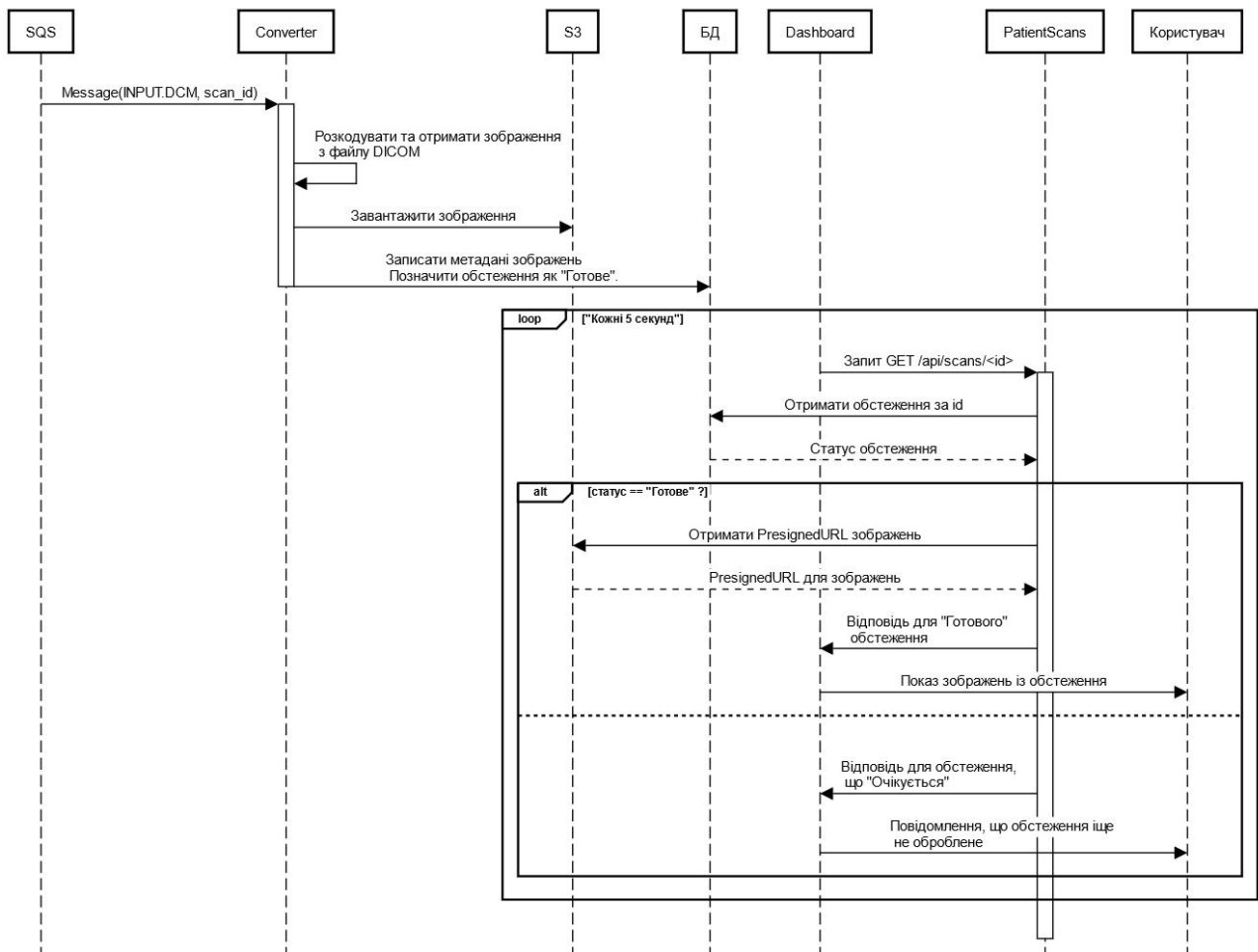


Рисунок 4.6. Діаграма послідовностей завантаження зображень DICOM в систему. Стадія перша - завантаження файлу

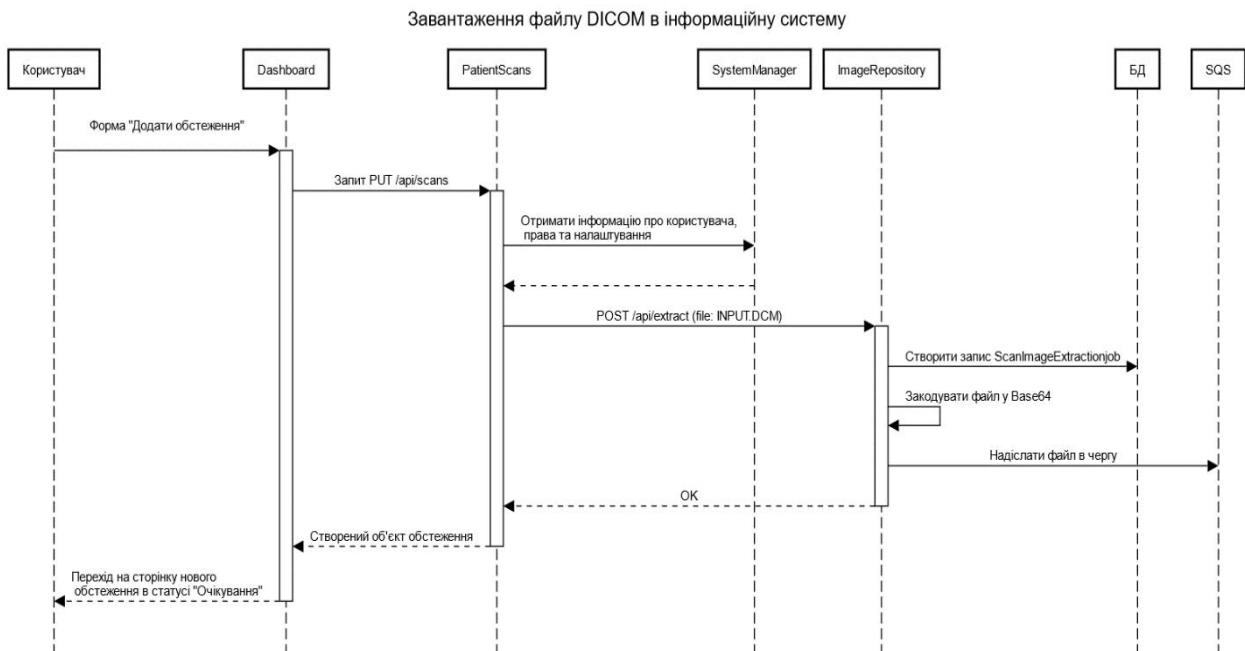


Рисунок 4.7. Діаграма послідовностей завантаження зображень DICOM в систему. Стадія друга – обгортка та очікування на зображення

Щодо реалізації процесу отримання зображень для заданого обстеження, цей варіант використання також дуже природньо поєднується із завантаженням, і є частково описаним в UML-діаграмі послідовностей для другої стадії. По суті, як тільки процес Converter закінчить екстракцію зображень, та промаркує їх за розрізами МРТ, він запише метадані для кожного зображення у БД (час створення, розріз МРТ, ідентифікатор обстеження, тощо), а також завантажить дані зображення на S3, також записавши у метадані унікальний ідентифікатор зображення на S3.

Згодом для отримання та показу зображень користувачеві сервіс PatientScans буде виконувати запит на отримання метаданих зображень до сервісу ImageRepository, здійснюючи фільтрування по заданому ідентифікатору обстеження, і, можливо, додатково за розрізом та номером зображення у послідовності розрізу.

Виконання моделі автоматизованого діагностування на обстеженні.

Після того, як файли обстеження завантажені, а саме обстеження

переходить у стан «Готове» (точніше, коли під-сервіс Converter закінчує роботу із вхідним файлом DICOM, та завантажує промарковані знімки з кожного розрізу МРТ разом із метаданими), на сторінці обстеження з'являється кнопка «Автоматизоване діагностування». При натиску на цю кнопку користувач запускає процес автоматизованого діагностування на всі три підтримувані діагнози: розрив ACL, розрив меніска, та аномальну форму коліна. Цей процес також відбувається у декілька етапів:

1. Сервіс Dashboard робить запит до сервісу PatientScans про підготовку автоматизованого діагностування для конкретного обстеження.
2. Сервіс PatientScans робить API-запит до сервісу MRNet-Combined, передаючи виключно ідентифікатор конкретного обстеження.
3. Сервіс MRNet-Combined, маючи ідентифікатор даного обстеження, робить запит на сервіс ImageRepository з метою витягнення бінарних даних зображення із Amazon S3.
4. Маючи посилання на зображення S3, згенеровані ImageRepository, сервіс MRNet-Combined робить обробку та обрізання даних зображень до потрібних розмірів, групування зображень за розрізами МРТ, та, власне, виклик блоків MRNet для кожного розрізу та завдання діагностування.
5. На виході матимемо ймовірності передбачення для кожного завдання діагностування для даного обстеження.

Схематично цей процес зображено на рисунку 4.7.

Останнім важливим процесом у даній інформаційній системі, котрій потрібно описати, є саме виконання пояснення для кожного знімка МРТ. Для кінцевого користувача системи, лікаря-рентгенолога, це дозволить здійснити візуальний огляд регіонів обраного зображення, які вплинули на результат передбачення найбільше.

Аналогічно, як із етапом передбачення, для цього використовується окремий сервіс ExrAIn, який за заданим ідентифікатором обстеження, а також розрізом МРТ та номером зображення у розрізі, може здійснити візуалізацію

сегментів зображення, котрі позитивно або негативно вплинули на наявність чи відсутність певного діагнозу. Важливо зазначити, що для пояснення, користувач повинен мати заздалегідь обране зображення на певному розрізі, що є зумовлено труднощами із побудовою інтерпретаційної моделі, описаними у розділі 3. Процес виконання моделі ExplAIIn полягає у генерації контрафактного зображення на основі заздалегідь відібраного елемента із середини серії, і його параметризованої вставці у оригінальну серію зображень МРТ, на вхід до моделі класифікатора, з метою отримати результат передбачення із заданим відхиленням λ .

Прогресивно генеруються контрафактні зображення, враховуючи початкову ймовірність класифікації для даного розрізу зображень, з кінцевою метою отримати зображення, серії з якими будуть прокласифіковані найближче до значень 0.05 та 0.95.

Далі, будується попіксельна різниця інтенсивності між даними зображеннями, що і буде пояснювальним зображенням. Схематика процесу пояснення зображена на рисунку 4.8.

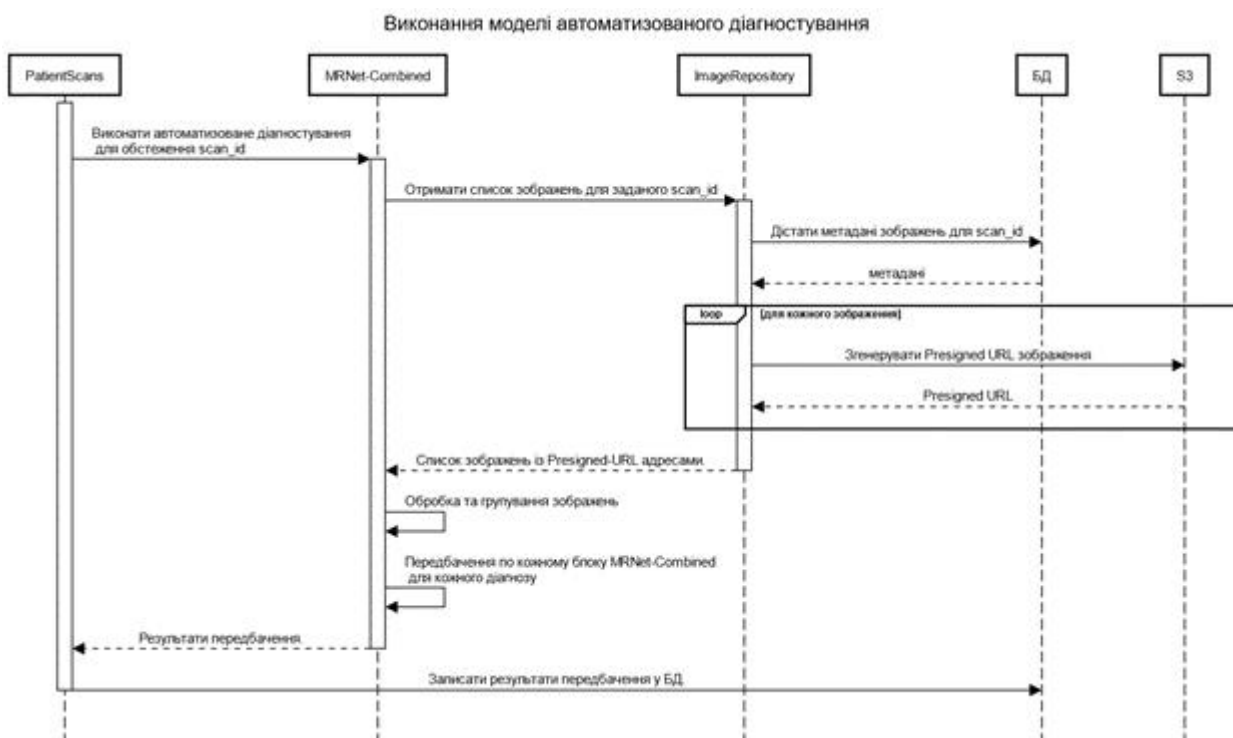


Рисунок 4.8. Процес виконання моделі автоматизованого діагностування для обстеження в системі

Діаграма станів для виконання інтерпретаційної моделі ExplAIIn подана на рисунку 4.9.

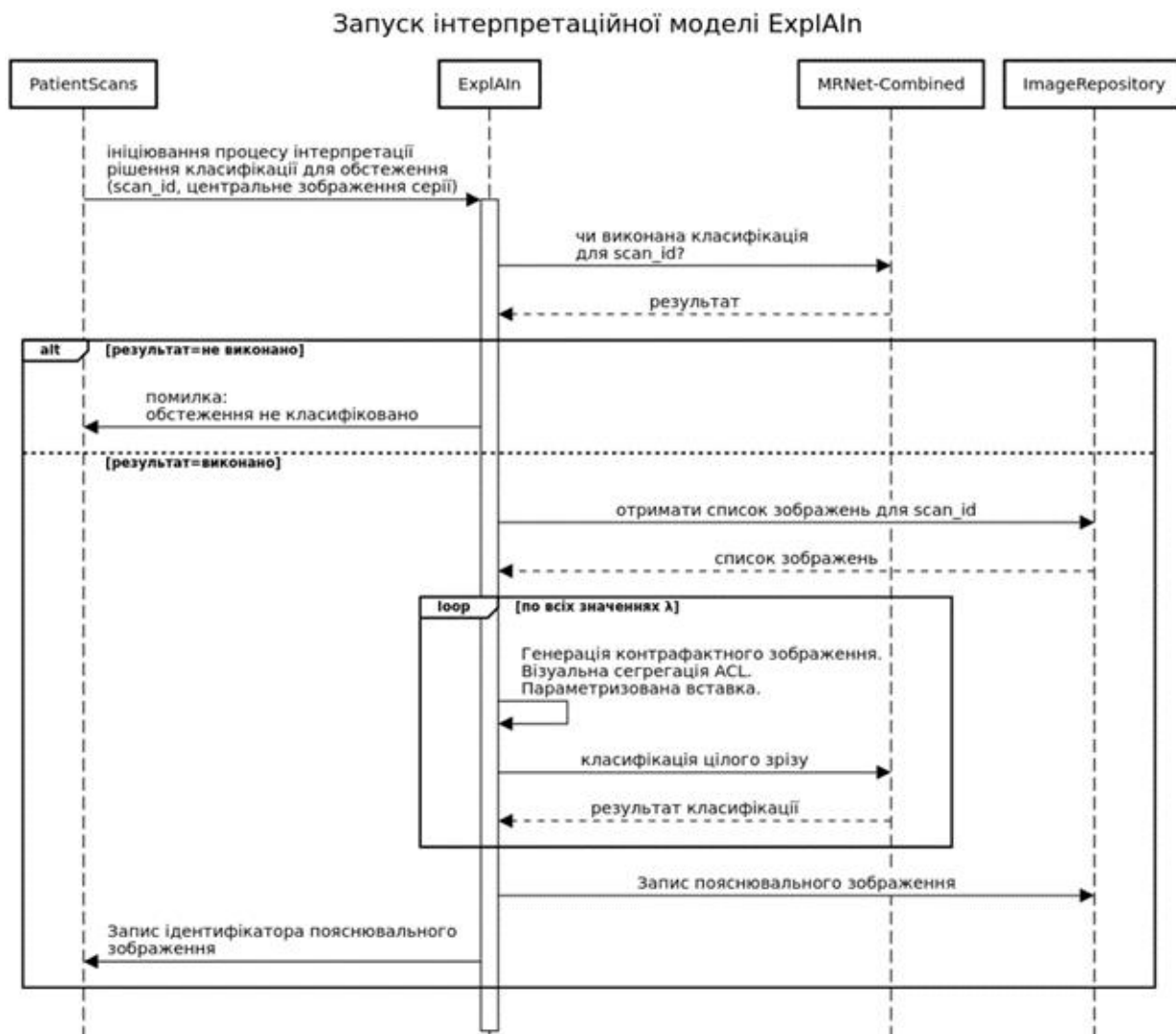


Рисунок 4.9. Діаграма станів для виконання інтерпретаційної моделі ExplAIIn

4.4. Зменшення вартості розгортання системи. Підрахунок вартості

При побудові архітектури інформаційної системи, у цій роботі здагується використання сервісу безсерверних обчислень AWS Lambda. Основною перевагою використання даного сервісу є те, що оплата стягується виключно за виклики лямбда-функцій, тобто час “простою” інформаційної системи буде оптимізованим в плані використання коштів.

Змоделюємо типове навантаження на кластер інформаційної системи для закладів охорони здоров’я в межах міста Львів. Використовуючи дані з

медичної інформаційної системи Helsi [90], у Львові на даний момент можна здійснити запис до 30 лікарів-рентгенологів.

Нехай кожен лікар-рентгенолог за день може прийняти в середньому 32 пацієнти. Це припущення робиться на основі того, що маємо 8-годинний робочий день, помножений на 4 візити за годину при 15-хвилинній тривалості візиту кожного пацієнта.

Виходячи із обчислень, для одного міста Львова матимемо навантаження у розмірі 960 пацієнтів на день. Припускаємо, що для кожного пацієнта потрібно буде завантажити дані сканування МРТ коліна (звісно, не всі звернення пацієнтів будуть пов'язані із травмами коліна та необхідністю обстеження МРТ, проте припускаючи, що це потрібно для усіх пацієнтів, можна розрахувати навантаження на інформаційну систему із достатнім запасом міцності в плані масштабування), а також виконати систему автоматизованого діагностування, і також викликати інтерпретаційну модель хоча б для одного зображення із одного розрізу МРТ.

На основі цих даних здійснимо обчислення кількості викликів сервісів.

На момент написання даного дисертаційного дослідження, преїскурант сервісу AWS Lambda є наступним [91]:

- \$0.20 за кожен мільйон запитів до сервісу;
- додатково \$0.00001667 за кожен Гігабайт-секунду обчислювальних ресурсів (під однією Гігабайт-секундою розуміється одна секунда виконання лямбда-функції, що вимагає один гігабайт оперативної пам'яті).

Розіб'ємо вимоги щодо кількості пацієнтів на власне виклики до системи для кожного пацієнта.

1. Завантаження та обробка файлів DICOM. Здійснення даної операції потребує до 1 гігабайту оперативної пам'яті, враховуючи, що окрім самого DICOM для МРТ, розмір якого в середньому становить 30МБ, також потрібні ресурси для обробки та зберігання у базі даних метаданих

для зображень. Середнє значення тривалості екстракції файлу DICOM на даній системі становило 5.7113 секунди. На виході маємо 5.7713 Гб*с.

2. Виконання моделі автоматизованого діагностування для кожного розрізу МРТ. Тут ми здійснювали тестування із 10Гб доступного обсягу пам'яті (це - максимально доступний об'єм на сервісі AWS Lambda). При цьому завантаження та виконання моделі для кожного розрізу МРТ, а також їх зважування за допомогою попередньо натренованої моделі логістичної регресії тривало в середньому 23.3506 секунд. На виході маємо 233.506 Гб*с обчислень для кожного пацієнта.
3. Підрахуймо кількість обчислень для виконання інтерпретаційної моделі. В рамках її виконання вона здійснює запити до моделі автоматизованого діагностування для одного розрізу, щонайменше 4 рази (у разі успішної генерації контрафактних зображень з першого разу), проте цей показник може варіюватися і до 15 викликів. В середньому модель виконує генерацію потрібних зображень за 6 викликів моделі автоматизованого діагностування. Можемо поєднати це із середнім значення гігабайт-секунд, обчислених у попередньому пункті, поділеному на 3 (оскільки працюємо лише в рамках одного розрізу МРТ). Сама інтерпретаційна модель використовує 4Гб оперативної пам'яті для генерації контрафактних зображень через автокодувальник, і, виключаючи час, необхідний на виклик моделі автоматизованого діагностування, триває в середньому 28.9597 секунд. Підсумувавши усі значення, маємо 816,3567 гігабайт-секунд.

Підсумувавши значення для кожного прогнозованого типу викликів лямбда-функції, отримаємо наступне:

$$P_{load} = 816.3567 + 233.506 + 5.7713 = 1055.6341(Gb \cdot s)$$

Також, окрім вказаних трьох варіантів використання, варто зарезервувати деяку кількість обчислюваної потужності на рутинні завдання, на кшталт, власне, перегляду зображень із DICOM. Для простоти обчислень, заокруглимо значення до 1200 гігабайт-секунд обчислень для кожного пацієнта.

Маючи 960 пацієнтів щодня, отримаємо:

$$1200Gb \cdot s \times 960 \times 0.000001667 USD / (Gb \cdot s) = 1.920384USD$$

Отже, при заданих параметрах обчислювального навантаження, використовуючи сервіс AWS Lambda, прогнозована вартість використання основних варіантів використання системи складає менше двох доларів США на день.

Звісно сюди не враховано інші сервіси на кшталт хмарного сховища S3, або сервісів RDS чи черг Amazon SQS, але дані витрати будуть сталими також і для варіанту розгортання системи, використовуючи традиційну архітектуру із хмарними серверами. У даному ж випадку, нас цікавить суто вартість для основних варіантів використання системи.

Тепер здійснимо обчислення вартості основних варіантів використання системи, використовуючи альтернативу від того самого хмарного провайдера Amazon Web Services — EC2. Для простоти, зосередимося лише на моделі автоматизованого діагностування, та інтерпретаційній моделі.

Для виконання першої з метою підтримки запитів для 960 пацієнтів на день, а також враховуючи запити від інтерпретаційної моделі з метою пояснення серії зображень МРТ із контрафактними вставками, найближчим типом серверу, який може задовільнити ці потреби є t3a.2xlarge, із ціною використання у 0.1504 доларів США за годину.

Для інтерпретаційної моделі потрібно використати більший тип інстансу - r5a.xlarge, який має 32 Гб оперативної пам'яті, що зумовлено необхідністю генерації контрафактних зображень. На цьому типі інстансу якраз можна запустити інтерпретаційну модель як сервіс. Ціна за цей тип інстансу сягає 0.2260 доларів США за годину.

Підсумувавши, отримаємо

$$0.2260 USD/h + 0.1504 USD/h = 0.3764 USD/h .$$

Помноживши отримане значення на 24 години, отримаємо вартість основних варіантів використання системи у 9.03 доларів на день, що є майже у п'ять разів більше, ніж при використанні сервісу AWS Lambda. Зважаючи на те,

що для виконання сервісів екстракції зображень також потрібні будуть інстанси на кшталт t3.medium або трішки дешевші t3.small, можна з впевненістю припустити, що AWS Lambda виправдовує себе для цього застосування в плані ефективності використання ресурсів.

Окрім цього, у разі меншого навантаження на систему, аніж прогнозоване, вартість розгортання при використанні традиційного підходу із хмарними серверами буде сталою, в той час як із безсерверним підходом, вартість буде пропорційно залежати від навантаження.

Висновки до розділу 4

У даному розділі дисертаційного дослідження розроблено архітектуру інформаційної системи для здійснення автоматизованого діагностування серій зображень МРТ колінного суглоба, а також інтерпретацій рішень класифікації щодо заданих діагнозів.

При побудові такої архітектури використовується безсерверний підхід, з метою подальшого зменшення вартості розгортання такої системи, на хмарному провайдері Amazon Web Services.

Попри складнощі у підлаштуванні бізнес-процесів даної інформаційної системи до архітектури без серверів, кінцевий результат прорахунку вартості експлуатації системи із такою архітектурою є у 4.7 раз дешевшим, аніж при використанні відповідного аналогу на традиційній архітектурі із хмарними серверами.

Висновки

У дисертаційній роботі розв'язано актуальне наукове завдання покращенні точності та інтерпретованості методів автоматизованого діагностування колінного суглоба.

1. Здійснено аналіз літератури, включаючи порівняльний аналіз існуючих методів та засобів для автоматизації процесу діагностування ушкоджень колінного суглоба.. Це дало змогу виділити невирішені задачі існуючими методами та здійснити постановку задачі.
2. Розроблено модель порівняння ефективності каркасних мереж виділення ознак для задач автоматизованого діагностування МРТ з метою виділення найбільш продуктивних каркасних мереж шляхом використання шару виділення ознак, як варіативного, та різних вхідних шарів класифікаторів.
3. Розроблено метод діагностування для МРТ колінного суглоба на основі розроблення моделі на базі ансамблю різнотипових каркасних мереж, що дало змогу підвищити точність прийнятих рішень.
4. Отримав подальший розвиток метод запобігання перенавчання моделі на основі аугментації наборів даних шляхом застосування послідовності поворотів та випадкових афінних перетворень, що дає змогу застосовувати цей метод до серійних зображень. Показано, що точність класифікації на аугментованих даних зросла на 11.8%.
5. Проведено порівняльний аналіз для набору моделей класифікації, заснованих на оригінальній архітектурі MRNet, з різними каркасними мережами виділення ознак - Alexnet, VGG11, VGG16, Resnet і Efficientnet. MRNet-VGG16 показує найкращі результати для виявлення розривів ACL. Для виявлення аномалій форми коліна, MRNet-VGG16 домінує MRNet-Resnet для достовірності від 0,5 до 0,75 та MRNet-VGG11 для достовірності понад 0,8.
6. Здійснено розробку методу запобігання перенавчання моделі на основі аугментації наборів даних шляхом застосування послідовності

поворотів та випадкових афінних перетворень, що дає змогу застосовувати цей метод до серійних зображень.

7. Розроблено інтерпретаційну модель діагностування пошкоджень колінного суглоба для серії знімків магнітно-резонансної томографії шляхом побудови карти помірності на основі латентної змінної із заданим значенням зсуву, що дає змогу забезпечити цілісність зображень в розрізі, тобто, згенероване зображення лише мінімально відрізняється від попереднього та наступного у розрізах, а також визначити відхилення точності передбачення при процесах класифікації та пояснення. У цьому ж розділі розроблено методологію інтерпретування для задач автоматизованого діагностування розриву ACL, використовуючи підхід зсуву латентної змінної, шляхом параметризованої вставки контрафактних зображень у середину серії. Підібрано оптимальний параметр $F=3$ для такої вставки, що забезпечує найменшу усереднену різницю ймовірностей класифікації у 0.096.
8. Розроблено архітектури інформаційної системи діагностування на основі знімків МРТ, використовуючи підхід безсерверних обчислень з метою зменшення витрат на розгортання системи. Показано, що такий підхід дозволяє скоротити у 4,7 разів кінцеві витрати на запуск та налагодження подібних ІС у комерційному середовищі для медичних даних.

Список використаних джерел

- [1] Shakhovska N., Pukach P. Comparative Analysis of Backbone Networks for Deep Knee MRI Classification Models. *Big Data and Cognitive Computing*. 2022. Vol. 6(3):69. <https://doi.org/10.3390/bdcc6030069>
- [2] Пукач П. Аналіз каркасних мереж виявлення ознак в моделях глибинного навчання. *Вісник Національного університету «Львівська політехніка» «Інформаційні системи та мережі»*. 2022. Вип. 12. С. 169-176. <https://doi.org/10.23939/sisn2022.12.169>
- [3] Пукач П. Огляд та аналіз основних каркасних мереж виявлення ознак для класифікації зображень мрт в моделях глибинного навчання. *Вісник Хмельницького національного університету. Серія: Технічні науки*. 2022. 6, част. 1. С. 183-187.
- [4] Pukach P. Y., Vovk M., Pukach P. P. On one nonlinear mathematical model of blood circulation with the vessel walls reaction within the hereditary theory. *Прикарпатський вісник НТШ. Серія: Число*. 2022. 17, № 64. С. 31-43. [https://doi.org/10.31471/2304-7399-2022-17\(64\)-31-43](https://doi.org/10.31471/2304-7399-2022-17(64)-31-43)
- [5] Shakhovska N., Darmoriz I., Vyklyuk Y., Kryvenchuk Y., Pukach P. Visualization of the epidemics forecasting results. *CEUR Workshop Proceedings*. 2021. Vol. 3038: Proceedings of the 4th International conference on informatics & data-driven medicine IDDM 2021, Valencia, Spain, November 19 - 21, 2021. P. 283-292.
- [6] Pukach P.Y., Vovk M., Mylyan Y., Bilushchak H., Pukach P.P. On one nonlinear mathematical model of blood circulation with the vessel walls reaction within the hereditary theory. *CEUR Workshop Proceedings*. 2020. Vol. 2753: Proceedings of the 3rd International conference on informatics & data-driven medicine IDDM 2020, Vaxjo, Sweden, November 19-21, 2020. P. 136-141.
- [7] IHS Markit LTD (Prepared for the AAMC), The Complexities of Physician

- Supply and Demand: Projections from 2019 to 2034 AAMC, 2021. URL: <https://www.aamc.org/data-reports/workforce/data/complexities-physician-supply-and-demand-projections-2019-2034> (дата звернення: 15.11.2022).
- [8] Triantaphyllou E., Yanase Y. A systematic survey of computer-aided diagnosis in medicine: Past and present developments. *Expert Systems with Applications*. 2019. 138, 112821. <https://doi.org/10.1016/j.eswa.2019.112821>
- [9] Echegaray S., Gevaert O., Shah R., Kamaya A., Louie J., Kothary N., Napel S. Core samples for radiomics features that are insensitive to tumor segmentation: method and pilot study using CT images of hepatocellular carcinoma. *J. Med. Imaging*. 2015. 2(4):041011. <https://doi.org/10.1117/1.JMI.2.4.041011>.
- [10] Mougiakakou S.G., Golemati S., Gousias I., Nicolaides A.N., Nikita K.S. Computer-aided diagnosis of carotid atherosclerosis based on ultrasound image statistics, laws' texture and neural networks. *Ultrasound Med. Biol.* 2007. 33(1):26-36. <https://doi.org/10.1016/j.ultrasmedbio.2006.07.032>
- [11] Chen S., Suzuki K., MacMahon H.. Development and evaluation of a computer-aided diagnostic scheme for lung nodule detection in chest radiographs by means of two-stage nodule enhancement with support vector classification. *Med. Phys.* 2011 38(4):1844-58. <https://doi.org/10.1118/1.3561504>
- [12] 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. *J. Magn. Reson. Imaging*. 2017. 45(5):1257-1275. <https://doi.org/10.1002/jmri.25620>
- [13] Gore J.C. Artificial intelligence in medical imaging. *Magn. Reson. Imaging*. 2020. 68:A1-A4. <https://doi.org/10.1016/j.mri.2019.12.006>
- [14] Субботін С. О. Нейронні мережі: теорія та практика. Навч. посіб. Житомир: Вид. О. О. Євенок, 2020. 184 с.
- [15] K. Hinkelmann. Neural Networks, p. 7. University of Applied Sciences, Northwestern Switzerland, 2018. URL:<https://web.archive.org/web/>

20181006235506/http://didattica.cs.unicam.it/lib/exe/fetch.php?media=didattica:magistrale:kebi:ay_1718:ke-11_neural_networks.pdf. (дата звернення: 10.01.2023).

- [16] Bollobás B. *Modern Graph Theory*. 5th Edition. New York: Springer, 2010. 394 c. <https://doi.org/10.1007/978-1-4612-0619-4>
- [17] Hammoud Z., Kramer F. Multilayer networks: aspects, implementations, and application in biomedicine. *Big Data Anal.* 2020. 5(2). <https://doi.org/10.1186/s41044-020-00046-0>
- [18] Deng L., Yu. D. Deep Learning: Methods and Applications. *Foundations and Trends in Signal Processing*. 2014. Vol. 7, Issue 3-4. P. 197-387. <https://doi.org/10.1561/20000000039>
- [19] Schmidhuber J. Deep learning in neural networks: An overview. *Neural Networks*. 2015. Vol. 61. P. 85-117. <https://doi.org/10.1016/j.neunet.2014.09.003>
- [20] Koivu A., Korpimäki T., Kivelä P., Pahikkala T., Sairanen M. Evaluation of machine learning algorithms for improved risk assessment for Down's syndrome. *Computers in Biology and Medicine*. 2018. 98. P. 1-7. <https://doi.org/10.1016/j.combiomed.2018.05.004>
- [21] Zhang W., Hasegawa A., Itoh K., Ichioka Y. Image processing of human corneal endothelium based on a learning network. *Appl. Opt.* 1991. 10. 30 (29):4211-7. <https://doi.org/10.1364/AO.30.004211>
- [22] Pagad N. S., Pradeep N., Almuzaini K. K., Maheshwari M., Gangodkar D., Shukla P., Alhassan M. Clinical Text Data Categorization and Feature Extraction Using Medical-Fissure Algorithm and Neg-Seq Algorithm. *Computational Intelligence and Neuroscience*. 2022. Volume 2022. Article ID 5759521. <https://doi.org/10.1155/2022/5759521>
- [23] Jouffroy J., Feldman S.F., Lerner I., Rance B., Burgun A., Neuraz A. Hybrid Deep Learning for Medication-Related Information Extraction From Clinical

- Texts in French: MedExt Algorithm Development Study. *JMIR Med. Inform.* 2021. 9(3):e17934. <https://doi.org/10.2196/17934>
- [24] Valueva M. V., Nagornov N. N., Lyakhov P. A., Valuev G. V., Chervyakov N. I. Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. *Mathematics and Computers in Simulation.* 2020. 177. P. 232-243.
<https://doi.org/10.1016/j.matcom.2020.04.031>
- [25] Géron A. Ahead of AI #5: ReAIval of Ideas. 2019. URL: magazine.sebastianraschka.com. (дата звернення: 07. 03. 2023)
- [26] Štajduhar I., Mamula M., Miletić D., Ünal G. Semi-automated detection of anterior cruciate ligament injury from MRI. *Comput. Methods Programs Biomed.* 2017. 140. P.151-164. <https://doi.org/10.1016/j.cmpb.2016.12.006>
- [27] Breiman L. Random forests. *Machine Learning.* 2001. 45. P. 5–32.
<https://doi.org/10.1023/A:1010933404324>
- [28] Witten I., Frank E., Hall M. Data Mining: Practical Machine Learning. Morgan Kaufmann, 2011. 629 с.
- [29] Bien N., Rajpurkar P., Ball R.L., Irvin J., Park A., Jones E., Bereket M., Patel B.N., Yeom K.W., Shpanskaya K., Halabi S., Zucker E., Fanton G., Amanatullah D.F., Beaulieu C.F., Riley G.M., Stewart R.J., Blankenberg F.G., Larson D.B., Jones R.H., Langlotz C.P., Ng A.Y., Lungren M.P. Deep-learning-assisted diagnosis for knee magnetic resonance imaging: Development and retrospective validation of MRNet. *PLoS Med.* 2018. 27. 15(11):e1002699.
<https://doi.org/10.1371/journal.pmed.1002699>
- [30] Welcome to Python.org! URL: <https://www.python.org/>. (дата звернення: 20. 11. 2022).
- [31] Pydicom. URL: <https://pydicom.github.io/>. (дата звернення: 20. 11. 2022).
- [32] Nyúl L.G., Udupa J.K. On standardizing the MR image intensity scale. *Magn. Reson. Med.* 1999. 42(6). P. 1072-1081. [https://doi.org/10.1002/\(sici\)1522-](https://doi.org/10.1002/(sici)1522-)

2594(199912)42:6<1072::aid-mrm11>3.0.co;2-m

- [33] Krizhevsky A., Sutskever I., Hinton G. E. ImageNet Classification with Deep Convolutional Neural Networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, Red Hook, NY, USA, 2012. P. 1-8.
- [34] Hoai M. Regularized max pooling for image categorization. *Proceedings British Machine Vision Conference*, Nottingham, UK, September 1 - 5 2014. <http://dx.doi.org/10.5244/C.28.32>
- [35] Bishop C. Pattern recognition and machine learning. Berlin: Springer, 2006. 738 c.
- [36] Azcona D., McGuinness K., Smeaton A. A Comparative Study of Existing and New Deep Learning Methods for Detecting Knee Injuries using the MRNet Dataset. *International Conference on Intelligent Data Science Technologies and Applications (IDSTA)*, Valencia, Spain, October 19 - 22 2020. P. 149-155. <http://dx.doi.org/10.1109/IDSTA50958.2020.9264030>
- [37] Singlaa S., Eslamib M., Pollackc B., Wallaced S., Batmanghelichc K. Explaining the Black-box Smoothly - A Counterfactual Approach. *Medical Image Analysis*. 2023. Vol. 84. <https://doi.org/10.1016/j.media.2022.102721>
- [38] Deng J., Dong W., Socher R., Li L.-J., Li K., Fei-Fei L. ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, USA, 2009. P. 248-255. <https://doi.org/10.1109/CVPR.2009.5206848>
- [39] Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv*. 2014. <https://doi.org/10.48550/arXiv.1409.1556>
- [40] Russakovsky O., Deng J., Su H., Krause J., Satheesh S., Ma S., Huang Z., Karpathy A., Khosla A., Bernstein M., Berg A., Fei-Fei L. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*.

2015. 115. P. 211–252. <https://doi.org/10.1007/s11263-015-0816-y>
- [41] He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. *arXiv*. 2015. <https://doi.org/10.48550/arXiv.1512.03385>
- [42] Srivastava R. K., Greff K., Schmidhuber J. Highway Networks. *arXiv*. 2015. <https://doi.org/10.48550/arXiv.1505.00387>
- [43] Huang G., Liu Z., van der Maaten L., Weinberger K. Q. Densely Connected Convolutional Networks. *arXiv*. 2016. <https://doi.org/10.48550/arXiv.1608.06993>
- [44] Nvidia CUDA Home Page. URL: <https://developer.nvidia.com/cuda-zone> (дата звернення: 03.12.2022).
- [45] Zagoruyko S., Komodakis N. Wide Residual Networks. *arXiv*. 2016. <https://doi.org/10.48550/arXiv.1605.07146>
- [46] Zoph B., Vasudevan V., Shlens J., Le Q. V. Learning Transferable Architectures for Scalable Image Recognition. *arXiv*. 2017. <https://doi.org/10.48550/arXiv.1707.07012>
- [47] Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z. Rethinking the Inception Architecture for Computer Vision. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016. P. 2818-2826. <https://doi.org/10.1109/CVPR.2016.308>
- [48] Tan M., Le Q. V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *arXiv*. 2019. <https://doi.org/10.48550/arXiv.1905.11946>
- [49] Tsai C., Kiryati N., Konen E., Eshed I., Mayer A. Knee Injury Detection using MRI with Efficiently-Layered Network (ELNet). *Third Conference on Medical Imaging with Deep Learning*, Montreal, QC, Canada, 2020. P. 784-794.
- [50] Yerushalmy J. Statistical problems in assessing methods of medical diagnosis with special reference to x-ray techniques. *Public Health Reports*. 1947. 62 (40). P. 1432-1449. <https://doi.org/10.2307/4586294>
- [51] Fawcett T. An introduction to ROC analysis. *Pattern Recognition Letters*. 2006.

- 27(8). P. 861-874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- [52] Swets J. A. Signal detection theory and ROC analysis in psychology and diagnostics: collected papers. Mahwah, NJ: Lawrence Erlbaum Associates. 1996.
- [53] Y. Sasaki. The Truth of the F-measure. 2007. URL: <https://www.cs.odu.edu/mukka/cs795sum09dm/Lecturenotes/Day3/F-measure-YS-26Oct07.pdf> (дата звернення: 26.05.2021)
- [54] Aziz Taha A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Medical Imaging*. 2015. 15 (29). <https://doi.org/10.1186/s12880-015-0068-x>
- [55] Cohen J. A Coefficient of Agreement for Nominal Scales. *Educ. Psychol. Meas.* 1960. 20(1). P. 37- 46. <https://doi.org/10.1177/001316446002000104>
- [56] McHugh M. Interrater reliability: The kappa statistic. *Biochem. Med.* 2012. 22(3). P. 276-282.
- [57] Weston A. A (very) brief overview of Google Cloud Platform. URL: <https://towardsdatascience.com/a-very-brief-overview-of-google-cloud-platform-9dcce7bc1b1c>. (дата звернення: 04.12.2022).
- [58] ImageNet Competition Leaderboard. URL: <https://paperswithcode.com/sota/image-classification-on-imagenet>. (дата звернення: 04.12.2022).
- [59] Shorten C., Khoshgoftaar T. A survey on Image Data Augmentation for Deep Learning. *J. Big Data*. 2019. 6. 60. <https://doi.org/10.1186/s40537-019-0197-0>
- [60] Goodfellow I. J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative Adversarial Networks. *arXiv*. 2014. <https://doi.org/10.48550/arXiv.1406.2661>
- [61] Wang J., Perez L. The Effectiveness of Data Augmentation in Image Classification using Deep Learning. *arXiv*. 2017. <https://doi.org/10.48550/arXiv.1712.04621>
- [62] Masitah Abdul Lateh et al. Handling a Small Dataset Problem in Prediction

Model by employ Artificial Data Generation Approach: A Review. *J. Phys. Conf. Ser.* 2017. 892. 012016

- [63] Bach S., Binder A., Montavon G., Klauschen F., Müller K.-R., Samek W. On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation. 2015. *PLoS ONE*. 10(7). e0130140. <https://doi.org/10.1371/journal.pone.0130140>
- [64] Samek W., Wiegand T., Müller K.-R. Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models. *arXiv*. 2017. <https://doi.org/10.48550/arXiv.1708.08296>
- [65] Caruana R., Lou Y., Gehrke J., Koch P., Sturm M. Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August 2015. P. 1721–1730. <https://doi.org/10.1145/2783258.2788613>
- [66] Flaxman B., Goodman S. European union regulations on algorithmic decision-making and a "right to explanation". *AI MAGAZINE*. 2017. 38(3). P. 51-57.
- [67] Maity A. Improvised Salient Object Detection and Manipulation. *Int. Journ. of Image, Graphics and Signal Processing*. 2016. 8(2). P. 53-60. <https://doi.org/10.5815/ijigsp.2016.02.07>
- [68] Arun N., Gaw N., Singh P., Chang K., Aggarwal M., Chen B., Hoebel K., Gupta S., Patel J., Gidwani M., Adebayo J., Li D. M., Kalpathy-Cramer J. Assessing the (Un)Trustworthiness of Saliency Maps for Localizing Abnormalities in Medical Imaging. *Radiol. Artif. Intell.* 2021. 6. 3(6). P. 200-267. <https://doi.org/10.1148/ryai.2021200267>
- [69] Arun N., Gaw N., Singh P., Chang K., Hoebel K., Patel J., Gidwani M., Kalpathy-Cramer J. Assessing the validity of saliency maps for abnormality localization in medical imaging. *Medical Imaging with Deep Learning*, Montreal, Canada, 2020.

- [70] RSNA Pneumonia Detection Challenge (2018), Radiological Society of North America. URL: <https://www.rsna.org/education/ai-resources-and-training/ai-image-challenge/rsna-pneumonia-detection-challenge-2018>. (дата звернення: 15.12.2022).
- [71] Selvaraju R. R., Das A., Vedantam R., Cogswell M., Parikh D., Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2017. P. 618-626. <https://doi.org/10.1109/ICCV.2017.74>
- [72] Singla S., Pollack B., Chen J., Batmanghelich K. Explanation by Progressive Exaggeration. *arXiv*. 2019. <https://doi.org/10.48550/arXiv.1911.00483>
- [73] Cohen J.P., Luck M., Honari S. Distribution Matching Losses Can Hallucinate Features in Medical Image Translation. *Lecture Notes in Computer Science*. 2018. Vol. 11070. https://doi.org/10.1007/978-3-030-00928-1_60
- [74] Mirza M., Osindero S. Conditional Generative Adversarial Nets. *arXiv*. 2014. <https://doi.org/10.48550/arXiv.1411.1784>
- [75] Ross A. S., Hughes M. C., Doshi-Velez F. Right for the Right Reasons: Training Differentiable Models by Constraining their Explanations. *Proc. of the 26 Intern. Joint Conf. on Artificial Intelligence*, Melbourne, Australia, August 19-25, 2017. P. 2662-2670. <https://doi.org/10.24963/ijcai.2017/371>
- [76] Cohen J. P., Brooks R., En S., Zucker E., Pareek A., Lungren M. P., Chaudhari A. Gifsplanation via Latent Shift: A Simple Autoencoder Approach to Counterfactual Generation for Chest X-rays. *Proceedings of Machine Learning Research*. 2021. 143. P. 74–104.
- [77] Wickramasinghe C. S., Marino D. L., Manic M. ResNet Autoencoders for Unsupervised Feature Learning From High-Dimensional Data: Deep Models Resistant to Performance Degradation. *IEEE Access*. 2021. 9. P. 40511-40520. <https://doi.org/10.1109/ACCESS.2021.3064819>
- [78] Van Dyck P., Froeling M., Heusdens C. H. W., Sijbers J., Ribbens A., Billiet T.

- Diffusion tensor imaging of the anterior cruciate ligament following primary repair with internal bracing: A longitudinal study. *Journal of Orthopaedic Research*. 2021. 39. P. 1318-1330. <https://doi.org/10.1002/jor.24684>
- [79] Joyce J. M. Kullback-Leibler Divergence. *International Encyclopedia of Statistical Science* / за ред. М. Lovric. Berlin, 2011. P. 720–722.
- [80] Fowler M. Microservices. URL: <https://martinfowler.com/articles/microservices.html> (дата звернення: 22. 12. 2022).
- [81] Amazon API Gateway. Amazon. URL: <https://aws.amazon.com/api-gateway/>. (дата звернення: 12. 12. 2022).
- [82] Amazon Simple Queue Service. URL: <https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDeveloperGuide/welcome.html> (дата звернення: 02. 12. 2022).
- [83] Freeman E., Hupfer S., Arnold K. *JavaSpaces Principles, Patterns, and Practice*. London: Addison-Wesley Professional, 1999. 368 с.
- [84] Getting started with Amazon RDS. URL: https://docs.aws.amazon.com/ru/rds/latest/UserGuide/CHAP_GettingStarted.html (дата звернення: 30. 11. 2022).
- [85] Cloud Object Storage - Amazon S3. URL: <https://aws.amazon.com/s3/> (дата звернення: 29. 11. 2022).
- [86] Amazon introduces Lambda, Containers at AWS re:Invent. URL: <https://web.archive.org/web/20160710231219/http://sdtimes.com/amazon-introduces-lambda-containers/> (дата звернення: 30. 11. 2022).
- [87] Що таке Медична інформаційна система. URL: <https://web.archive.org/web/20211108183124/https://blog.h24.ua/shho-take-mis/> (дата звернення: 30. 11. 2022).
- [88] Amazon S3 Intelligent-Tiering. URL: <https://aws.amazon.com/s3/storage-classes/intelligent-tiering/> (дата звернення: 30. 12. 2022).
- [89] Base64 encoding and decoding - Web APIs. MDN Web Docs. URL:

https://developer.mozilla.org/enUS/docs/Web/API/WindowBase64/Base64_encoding_and_decoding (дата звернення: 25. 12. 2022).

[90] HELSI - інформаційна система для пацієнтів. URL: <https://helsi.me> (дата звернення: 25. 02. 2023).

[91] A. W. Services, Amazon Lambda - Pricing. URL: <https://aws.amazon.com/lambda/pricing/> (дата звернення: 29. 01. 2023).

[92] URL:<https://www.sciencedirect.com/science/article/abs/pii/S0957417419305238?via%3Dihub>.

[93] Tesla P100 Data Center Accelerator, NVIDIA. URL: <https://www.nvidia.com/en-us/data-center/tesla-p100/> (дата звернення: 04. 12. 2022).

[94] Configure compute resources for prediction, Google. URL: <https://cloud.google.com/vertex-ai/docs/predictions/configure-compute> (дата звернення: 04. 12. 2022).

ДОДАТОК А. АКТИ ВПРОВАДЖЕННЯ

ЗАТВЕРДЖУЮ

Генеральний директор

Львівського онкологічного регіонального
лікувально-діагностичного центру»

Ковальчук І.В.

» _____ 2023 р.



АКТ

про впровадження результатів дисертаційної роботи аспіранта кафедри «Системи штучного інтелекту» Національного університету «Львівська політехніка»

Пукача Павла Петровича

Цей акт підтверджує, що результати дисертаційної роботи Пукача Павла Петровича були використані для розроблення системи анімування зображень у Львівському онкологічному регіональному лікувально-діагностичному центрі (м. Львів) у 2022-2023 рр.

Впровадження дисертаційних досліджень П.П. Пукача полягає у наступному:

- Розроблено модель та метод аналізу ефективності мереж глибинного навчання для задач комп'ютерного бачення. Це стало основою для розроблення моделі опрацювання зображення.
- Розроблено інтерпретаційну модель для глибинних мереж у різних задачах класифікації. Це дало змогу зберігати проаналізовані об'єкти в базі даних без необхідності їх повторного опрацювання.
- Розроблено інформаційну систему для автоматизованого інтерпретованого діагностування зображень МРТ колінного суглоба.

Даний акт не є підставою для взаємних фінансових розрахунків.

Лікар кабінету МРТ

Тертична Г.П.

Лікар-рентгенолог, завідувач кабінетом КТ

Милян Ю.П.



ЗАТВЕРДЖУЮ

Проректор з наукової роботи
Національного університету
«Львівська політехніка»

Іван ДЕМИДОВ

2023 р.

АКТ

використання наукових результатів дисертаційної роботи Пукача Павла Петровича, представленої на здобуття наукового ступеня доктора філософії

Комісія в складі: голови комісії - начальника науково-дослідної частини д.т.н., с.н.с. Небесного Р.В. та членів комісії - завідувача кафедри СШ Шаховської Н.Б., професора кафедри СШ Яковини В.С., доцента кафедри СШ Хавалка В.М., доцента кафедри СШ Кривенчука Ю.П. цим актом підтверджують, що результати дисертаційної роботи Пукача П.П., зокрема,

- модель аналізу ефективності мереж глибинного навчання для задач комп'ютерного бачення;
- модель автоматизованого діагностування зображень МРТ колінного суглоба на основі ансамблю сучасних мереж комп'ютерного бачення;
- методика аугментації медичних зображень на основі афінних перетворень

використано у науково-дослідних роботах фінансованих Міністерством освіти і науки України, що виконувались на кафедрі систем штучного інтелекту і включено до звіту: «Інформаційна технологія формування психофізичного портрету в умовах стресових ситуацій» (№ держ. реєстру 0119U002257).

Отримані автором результати використано:

- при розробленні систем комп'ютерного бачення;
- при розробленні засобів автоматизованого діагностування зображень МРТ;
- при розробленні засобів аугментації медичних зображень.

Голова комісії:

начальник
науково-дослідної частини
д.т.н., с.н.с.

Роман НЕБЕСНИЙ

Члени комісії:

завідувач кафедри СШ

Наталія ШАХОВСЬКА

професор кафедри СШ

Віталій ЯКОВИНА

доцент кафедри СШ

Віктор ХАВАЛКО

доцент кафедри СШ

Юрій КРИВЕНЧУК

“ЗАТВЕРДЖУЮ”

Проректор з науково-педагогічної роботи
Національного університету
"Львівська політехніка"



Олег ДАВИДЧАК

" 15 " 03 2023 р.

А К Т

про впровадження в навчальний процес результатів
дисертаційної роботи

Пукача Павла Петровича

Цей акт складено про те, що результати дисертаційної роботи Пукача Павла Петровича впроваджено у навчальний процес кафедри "Системи штучного інтелекту" Національного університету "Львівська політехніка".

Впровадження результатів дисертаційної роботи полягає в їхньому використанні при викладанні навчальних дисциплін як окремих розділів лекційних курсів, так і в циклах лабораторних робіт.

Зокрема, для викладання дисципліни 6.122.04.Е.188 «Аналіз великих даних» для студентів освітньо-кваліфікаційного рівня «бакалавр», що навчаються за напрямом 122 "Комп'ютерні науки", використано такі результати:

- інтерпретаційна модель для глибинних мереж у різних задачах класифікації на основі методу латентного зсуву;
- методи побудови архітектур інформаційних системи на базі безсерверних обчислень.

Директор ІКНІ,
д.т.н., професор

Микола МЕДИКОВСЬКИЙ

Завідувач кафедри СШ,
д.т.н., професор

Наталія ШАХОВСЬКА

Доцент кафедри СШ,
к.е.н., доцент

Наталія БОЙКО