

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)

Кафедра Штучного інтелекту
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

Автоматичний розрахунок фракції викиду лівого
шлуночка серця за допомогою моделей глибинного навчання
(тема)

Виконав:
студент 2 курсу, групи СШМ-19-2
Даниленко К. О.
(прізвище, ініціали)

Спеціальність 122 Комп'ютерні науки
(код і повна назва спеціальності)

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

Освітня програма Системи штучного інтелекту
(повна назва спеціалізації)

Керівник проф. Терзіян В. Я.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____
(підпис)

В.О. Філатов
(прізвище, ініціали)

2021 р.

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)
Кафедра Штучного інтелекту
(повна назва)
Рівень вищої освіти другий (магістерський)
Спеціальність 122 Комп'ютерні науки
(код і повна назва)
Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)
Освітня програма Системи штучного інтелекту (СШІ)
(повна назва)

ЗАТВЕРДЖУЮ:
Зав. кафедри _____
(підпис)
«_____» _____ 20__ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Даниленко Катерині Олександрівні
(прізвище, ім'я, по батькові)

1. Тема роботи Автоматичний розрахунок фракції викиду лівого шлуночка серця за допомогою моделей глибинного навчання

затверджена наказом університету від 29 березня 2021 р. № 390Ст

2. Термін подання студентом роботи до екзаменаційної комісії 19 травня 2021 р.

3. Вихідні дані до роботи Науково-технічні публікації, медичні публікації, дані Інтернет-джерел та наукових робіт щодо ручного та автоматичного підрахунку фракції викиду лівого шлуночка серця, документація мови програмування Python та її бібліотек

4. Перелік питань, що потрібно опрацювати в роботі _____

- 1) Аналіз предметної галузі _____
- 2) Методи та алгоритми сегментації лівого шлуночка _____
- 3) Алгоритм для автоматичного підрахунку фракції викиду _____
- 4) Вибір програмних засобів для реалізації алгоритму автоматичного розрахунку фракції викиду _____
- 5) Проведення експериментальних досліджень _____

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) _____

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

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

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Отримання завдання на дипломну роботу	29.03.2021	виконано
2	Аналіз предметної галузі та постановка завдання	01.04.2021 – 04.04.2021	виконано
3	Аналіз методів та алгоритмів для сегментації лівого шлуночка	04.04.2021 – 09.04.2021	виконано
4	Аналіз алгоритму для підрахунку фракції викиду, аналіз архітектури MFP-UNet	09.04.2021 – 12.04.2021	виконано
5	Розробка алгоритму для автоматичного підрахунку фракції викиду	12.04.2021–15.04.2021	виконано
6	Тестування	15.04.2021 – 16.04.2021	виконано
7	Оформлення пояснювальної записки	29.03.2021 – 19.04.2021	виконано
8	Оформлення графічних матеріалів	10.05.2021	виконано
9	Попередній захист	12.05.2021	виконано
10	Захист перед ЕК	19.05.2021	

Дата видачі завдання 29 березня 2021 р.

Студент _____
(підпис)

Керівник роботи _____
(підпис) _____
(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка: 94 с., 2 табл., 25 рис., 2 дод., 33 джерела.

АВТОМАТИЗАЦІЯ, АНАЛІТИЧНИЙ МЕТОД ПІДРАХУНКУ, ЛІВИЙ ШЛУНОЧОК СЕРЦЯ, МОДЕЛІ ГЛИБИННОГО НАВЧАННЯ, СЕГМЕНТАЦІЯ, ФРАКЦІЯ ВИКИДУ.

Об'єкт дослідження – процес автоматичного підрахунку фракції викиду лівого шлуночка серця за допомогою моделей глибинного навчання та аналітичного методу.

Предмет дослідження – алгоритми та методи глибинного навчання, які використовуються для автоматичного підрахунку фракції викиду та сегментації лівого шлуночка, аналітичні методи підрахунку фракції викиду на медичних двовимірних зображеннях.

Мета роботи – дослідження методів для автоматичного підрахунку фракції викиду та розробка алгоритму, який виконує задачу автоматичного підрахунку за допомогою моделі глибинного навчання та аналітичного методу.

Методи дослідження – методи підрахунку фракції викиду вручну на двовимірних зображеннях, алгоритми та методи обробки медичних зображень, моделі глибинного навчання для сегментації медичних зображень, аналітичні методи підрахунку.

РЕФЕРАТ

Пояснительная записка: 94 с., 2 табл., 25 рис., 2 доп., 33 источника.

АВТОМАТИЗАЦИЯ, АНАЛИТИЧЕСКИЙ МЕТОД ПОДСЧЕТА, ЛЕВЫЙ ЖЕЛУДОЧЕК СЕРДЦА, МОДЕЛИ ГЛУБИННОГО ОБУЧЕНИЯ, СЕГМЕНТАЦИЯ, ФРАКЦИЯ ВЫБРОСА.

Объект исследования – процесс автоматического подсчета фракции выброса левого желудочка сердца с помощью моделей глубинного обучения и аналитического метода.

Предмет исследования – алгоритмы и методы глубинного обучения, которые используются для автоматического подсчета фракции выброса и сегментации левого желудочка, аналитические методы подсчета фракции выброса на медицинских двумерных изображениях.

Цель работы – исследование методов для автоматического подсчета фракции выброса и разработка алгоритма, который выполняет задачу автоматического подсчета с помощью модели глубинного обучения и аналитического метода.

Методы исследования – методы подсчета фракции выброса вручную на двумерных изображениях, алгоритмы и методы обработки медицинских изображений, модели глубинного обучения для сегментации медицинских изображений, аналитические методы подсчета.

ABSTRACT

Explanatory note: 94 p., 2 tabl., 25 fig., 2 app., 33 sources.

AUTOMATION, ANALYTICAL CALCULATION METHOD, DEEP LEARNING MODELS, EJECTION FRACTION, LEFT VENTRICLE OF THE HEART, SEGMENTATION

The object of research is the process of automatic calculation of the ejection fraction of the left ventricle of the heart using models of in-depth training and analytical method.

The subject of research is algorithms and methods of deep learning, which are used for automatic calculation of the ejection fraction and segmentation of the left ventricle, analytical methods for calculating the ejection fraction on medical two-dimensional images.

The purpose of the work is to study the methods for automatic calculation of the emission fraction and to develop an algorithm that performs the task of automatic calculation using the model of deep learning and analytical method.

Research methods are methods of calculating the emission fraction manually on two-dimensional images, algorithms and methods of medical image processing, models of in-depth training for segmentation of medical images, analytical methods of calculation.

ЗМІСТ

Перелік умовних позначень.....	7
Вступ.....	8
1 Аналіз предметної галузі.....	10
1.1 Виявлення проблем та актуалізація рішень	10
1.2 Поняття фракції викиду лівого шлуночка	13
1.3 Постановка задачі	22
2 Методи та алгоритми сегментації лівого шлуночка.....	23
2.1 Методи сегментації лівого шлуночку. Архітектура U-Net.....	23
2.2 Навчання U-Net моделі.....	30
2.3 Аугментація даних.....	34
2.4 Модифікації U-Net архітектури.....	38
2.5 CoordConv для покращення сегментації медичних зображень	51
3 Алгоритм для автоматичного підрахунку фракції викиду.....	53
3.1 Загальний опис алгоритму для підрахунку фракції викиду.....	53
3.2 Метод сегментації лівого шлуночка.....	55
3.3 Аналітичний метод підрахунку довжини, об'єму та фракції викиду..	61
4 Програмні засоби для реалізації алгоритму автоматичного розрахунку фракції викиду	66
5 Експериментальні дослідження.....	70
Висновки	75
Перелік джерел посилання	77
Додаток А Реалізація алгоритму автоматичного підрахунку фракції викиду лівого шлуночка серця.....	81
Додаток Б Відомість кваліфікаційної роботи.....	94

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

КТ – комп'ютерна томографія;

ЛШ – лівий шлуночок;

МРТ – магнітно-резонансна томографія;

СН – серцева недостатність;

ФВ – фракція викиду;

ШІ – штучний інтелект;

BCE – Binary Cross Entropy – функція витрат бінарна хрестова ентропія;

EDV – End-Diastolic Volume – кінцево-діастолічний об'єм лівого шлуночка;

EF – Ejection Fraction – фракція викиду;

ESV – End-Systolic Volume – кінцево-систолічний об'єм лівого шлуночка;

GAN – Generative adversarial network – генеративні змагальні мережі;

IoU – Intersection over Union – перетин-над-союзом, функція помилок;

RCNN – Recurrent Convolutional Neural Network – рекурентна згорткова нейронна мережа;

ReLU – Rectified Linear Unit – функція активації в згортковій нейронній мережі;

SV – Stroke Volume – ударний об'єм крові.

ВСТУП

Штучний інтелект (ШІ) повільно пронизує безліч галузей протягом останніх декількох років. ШІ в медицині став особливо актуальною темою в останні роки, допомагаючи лікарям точніше діагностувати пацієнтів, робити прогнози щодо майбутнього здоров'я пацієнтів та рекомендувати кращі методи лікування.

Щоб правильно діагностувати захворювання, потрібні роки медичної підготовки. Навіть тоді діагностика часто є важким, трудомістким процесом. У багатьох сферах попит на експертів значно перевищує наявну пропозицію. Це створює напруження для лікарів і часто затримує рятувальну діагностику пацієнта.

Машинне навчання – особливо алгоритми глибокого навчання – нещодавно досягло величезних успіхів у автоматичному діагностуванні захворювань, зробивши діагностику дешевшою та доступнішою. Той факт, що багато завдань можна автоматизувати і є основним аргументом на користь широкого використання штучного інтелекту в медицині. Автоматизація часто означає, що завдання виконуються швидше, а також звільняє час для медичних працівників, коли вони можуть виконувати інші обов'язки, які не можна автоматизувати і, отже, розглядаються як більш ефективне використання людських ресурсів [1]. Зменшується час на діагностування одного пацієнта шляхом використання автоматизованих методів підрахунків певних вимірювань для виявлення того чи іншого захворювання, що дає можливість збільшити кількість обстежень в день без втрати якості діагностування.

Алгоритми глибинного навчання можуть навчитися бачити закономірності подібно до того, як їх бачать лікарі. Ключова відмінність полягає в тому, що для навчання алгоритмам потрібно велика кількість конкретних прикладів – багато тисяч, які буває важко отримати, особливо, якщо захворювання рідкісне. Також важливо зазначити, що машинне

навчання буде корисне в тих областях, де діагностична інформація, яку досліджує лікар, є оцифрованою – зрозумілою для машини. Такими областями можуть бути:

- виявлення раку легенів або інсультів на основі КТ;
- оцінка ризику раптової серцевої смерті або інших серцевих захворювань на основі електрокардіограм та зображень МРТ серця;
- класифікація уражень шкіри на знімках шкіри;
- пошук показників діабетичної ретинопатії на зображеннях очей [2].

Оскільки в вище вказаних областях існує достатня кількість відповідних даних, алгоритми стають такими ж точними в діагностиці, як і експерти. Різниця полягає в тому, що алгоритм може робити висновки за частки секунди, і його можна відтворити по всьому світі, і в кожному куточку планети люди будуть мати можливість отримати доступ до однаково високої якості радіологічної діагностики найкращого фахівця за низькою ціною.

Застосування машинного навчання в діагностиці тільки починається – більш амбітні системи включають поєднання безлічі джерел даних (КТ, МРТ, геноміка та протеоміка, дані пацієнтів і навіть рукописні файли) для оцінки захворювання або його прогресування. Та слід зазначити, що навряд чи ШІ повністю зможе замінити лікарів. Натомість системи ШІ будуть використовуватися для висвітлення потенційно злоякісних уражень або небезпечних серцевих закономірностей для експерта – дозволяючи лікареві зосередитись на інтерпретації цих сигналів.

1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ

1.1 Виявлення проблем та актуалізація рішень

Протягом останнього десятиліття серцево-судинні захворювання залишаються основною основою смерті у всьому світі. За підрахунками Всесвітньої організації охорони здоров'я, щороку у всьому світі через серцево-судинні захворювання помирає понад 17,9 мільйона смертей, і 80% з них – від ішемічної хвороби та мозкового інсульту [2]. Величезна кількість смертей поширена серед країн із низьким та середнім рівнем доходу [3]. Багато схильних факторів, таких як особисті та професійні звички та генетична схильність, пояснюють серцеві захворювання. Різні звичні фактори ризику, такі як куріння, надмірне вживання алкоголю та кофеїну, стрес та фізична бездіяльність, а також інші фізіологічні фактори, такі як ожиріння, гіпертонія, високий рівень холестерину в крові та наявні захворювання серця є схильними факторами для серцевих захворювань.

Серцева недостатність (СН) – складний клінічний синдром, який виникає внаслідок функціонального або структурного порушення шлуночків, що призводить до симптоматичної дисфункції лівого шлуночка (ЛШ) [4]. Симптоми виникають внаслідок недостатнього серцевого викиду, що не встигає за метаболічними потребами організму. Це провідна причина серцево-судинної захворюваності та смертності у всьому світі, незважаючи на досягнення в галузі терапії та профілактики. Це може бути наслідком розладів перикарда, міокарда, ендокарда, клапанів серця, великих судин або деяких метаболічних відхилень. Ефективна, точна та рання медична діагностика серцево-судинних захворювань відіграє вирішальну роль у вживанні профілактичних заходів для запобігання смерті.

Якщо у пацієнта серцева недостатність, медичному працівнику, який проводить діагностику, важливо визначити фракцію викиду лівого

шлуночка (ФВЛШ) серця. Дане вимірювання допомагає визначити міри, які необхідно вжити для лікування та перевірити наскільки ефективно працює лікування.

Фракція викиду (ФВ) показує, наскільки добре лівий шлуночок (або правий шлуночок) перекачує кров з кожним серцевим ритмом. У більшості випадків ФВ визначає кількість крові, яка відкачується з лівого шлуночка кожного разу, коли він скорочується [5]. Лівий шлуночок є основною насосною камерою серця. Дане вимірювання виражається у відсотках. ФВ, який нижче норми, може бути ознакою серцевої недостатності. Якщо у пацієнта серцева недостатність і рівень ФВ нижче за норму (знижений), даний показник допомагає дізнатися, наскільки важким є стан пацієнта.

Точне вимірювання ФВЛШ є дуже важливим для ведення пацієнтів із серцево-судинними захворюваннями. ФВЛШ також має прогностичне значення при прогнозуванні несприятливих наслідків у пацієнтів із застійною серцевою недостатністю, після інфаркту міокарда та після реваскуляризації [4]. ФВЛШ відіграє важливу роль у оцінці ступеня тяжкості зниження систолічної функції серця, і, отже, є корисним для керівництва лікуванням різних серцево-судинних захворювань. Показання для вимірювання ФВЛШ включають:

- визначення анатомії та функції лівого шлуночка;
- оцінка глобальної функції лівого шлуночка – якісна та кількісна;
- оцінка пацієнтів з ознаками та симптомами, що свідчать про серцево-судинні захворювання;
- визначення категорії серцевої недостатності;
- оцінка перед процедурою, для якої систолічна функція лівого шлуночка може бути фактором ризику або протипоказанням;
- оцінка вроджених вад серця;
- оцінка клапанних розладів.

Для вимірювання ФВ можуть бути використані кілька методів візуалізації, включаючи ехокардіографію, магнітно-резонансну

томографію (МРТ), комп'ютерну томографію (КТ), радіонуклідну ангіографію та комп'ютерну томографію. В даний час МРТ широко вважається золотим стандартом для вимірювання ФВ, значною мірою завдяки корисності томографічних методів для подолання потреби в радіаційному опроміненні та вирішення геометричної складності серцевих камер, особливо правого шлуночка [6].

ФВ на ехокардіографії, як правило, вимірюється лише в лівому шлуночку, оскільки лівий шлуночок повністю доступний за допомогою стандартних ехокардіографічних оглядів, а геометрія камери дозволяє надійно оцінювати ФВ як параметр з ортогональних 2D зображень, використовуючи витягнутий еліпс модель. Незважаючи на те, що ці оцінки виявились надійними, у більшості ехокардіографічних оцінок ФВ залишається ступінь операторської залежності. Тим не менше, ехокардіографічно оцінена ФВ залишається найбільш широко застосовуваним підходом через низьку вартість, відсутність випромінювання та все більшу доступність. Але варто зазначити, що підрахунок ФВ на 2D зображеннях ехокардіограм може зайняти багато часу.

Ехокардіографія є ефективним методом контролю стану серця і дозволяє проводити ранню діагностику до появи симптомів. Це неінвазивна діагностична методика, яка надає інформацію щодо серцевої функції та гемодинаміки. Ехокардіографію найчастіше використовують для діагностування серцево-судинних захворювань після електрокардіографії та рентгенографії грудної клітки. Ехокардіограма проводиться за допомогою ультразвукових хвиль для створення рухомої картини серця. Перевагами даного методу діагностики є те, що звукові хвилі нешкідливі для людського організму та безболісні, а портативні системи забезпечують зручність з миттєвими відтворюваними результатами.

Однак у клініках первинної медичної допомоги та бідніших регіонах є недостатньо досвідчені рентгенологи для інтерпретації ехокардіографії. Тому існує потреба у розробці системи, яка могла б допомагати медичному

працівнику швидко та якісно інтерпретувати ехокардіографічне дослідження серця, виводити рекомендації щодо подальшого лікування та виконувати автоматичні розрахунки певних вимірювань, необхідних для діагностування.

1.2 Поняття фракції викиду лівого шлуночка

Перед тим, як дати поняття фракції викиду крові необхідно розглянути з яких частин складається наше серце та як саме воно працює. Розташоване між легеньми посередині грудної клітки, серце перекачує кров через мережу артерій і вен, відому як серцево-судинна система. Він підштовхує кров до органів, тканин і клітин організму. Кров доставляє кисень та поживні речовини до кожної клітини та виводить вуглекислий газ та інші відходи, що утворюються цими клітинами. Кров переноситься від серця до решти тіла через складну мережу артерій, артеріол і капілярів. Кров повертається до серця через вени і венули.

Серце складається з чотирьох камер: двох верхніх камер, відомих як ліве передсердя і праве передсердя, і двох нижніх камер, званих лівим і правим шлуночками. Воно також складається з чотирьох клапанів: трикуспідального, легеневого, мітрального та аортального клапанів [5] рисунок 1.1.

Робота серця складається з наступних етапів:

– праве передсердя отримує кров, що не має кисню, з найбільших вен тіла – верхньої порожнистої вени та нижньої порожнистої вени – і перекачує її через трикуспідальний клапан до правого шлуночка;

– правий шлуночок перекачує кров через легеневий клапан до легенів, де вона насичується киснем;

– ліве передсердя отримує кисневу кров з легенів і перекачує її через мітральний клапан до лівого шлуночка;

– лівий шлуночок перекачує багату киснем кров через аортальний клапан в аорту та решту тіла.

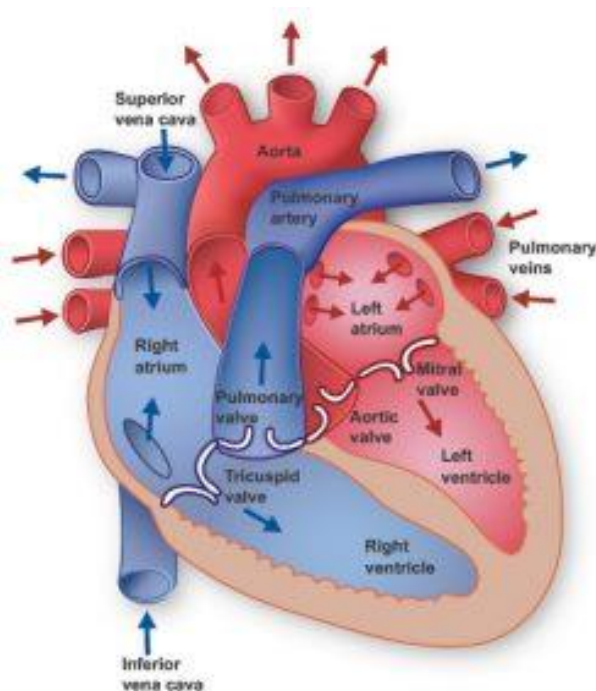


Рисунок 1.1 – Будова серця

Фракція викиду крові (англ. Ejection Fraction – EF) – це показник, що визначає ступінь ефективності роботи серцевого м'яза під час удару. Це вимірювання відсотка крові, яка залишає серце кожного разу, коли воно стискається. Коли серце стискається, воно викачує (викидає) кров з двох нижніх камер (шлуночків). Коли серце розслабляється, шлуночки наповнюються кров'ю. Яким би сильним не було скорочення, серце ніколи не може викачати всю кров із шлуночка. Термін «фракція викиду» стосується відсотка крові, яка відкачується із наповненого шлуночка при кожному серцебитті. Фракцію викиду зазвичай вимірюють лише в лівому шлуночку. Лівий шлуночок є основною насосною камерою серця. Він перекачує багату киснем кров у головну артерію вашого тіла (аорту) до решти частини тіла. Фракція викиду лівого шлуночка є центральним показником систолічної функції лівого шлуночка [8].

Ударний об'єм (англ. Stroke volume – SV) – це об'єм крові в мілілітрах, що викидається зі шлуночка через скорочення серцевого м'яза, який здавлює ці шлуночки. SV – різниця між кінцевим діастолічним об'ємом (англ. End-Diastolic Volume – EDV) та кінцевим систолічним об'ємом (англ. End-Systolic Volume – ESV) (формула 1.1).

$$SV = EDV - ESV. \quad (1.1)$$

Кінцевий діастолічний об'єм – це кількість крові, яка знаходиться в шлуночках до скорочення серця. Кінцевий діастолічний об'єм відноситься до кількості крові в лівому шлуночку в кінці діастоли, безпосередньо перед початком систоли. Діастола виникає, коли серцевий м'яз розслабляється, а камери наповнюються кров'ю.

Кінцевий систолічний об'єм – це кількість крові, що залишається в шлуночку в кінці систоли, після скорочення серця. Систола виникає, коли шлуночки скорочуються, витісняючи кров з лівого шлуночка в іншу частину тіла.

Розраховується фракція викиду крові наступним чином (формула 1.2):

$$EF = \frac{SV}{EDV} = \frac{EDV - ESV}{EDV} \cdot 100\%. \quad (1.2)$$

Фракцію викиду лівого шлуночка можна розрахувати, використовуючи різні способи, або суб'єктивно, використовуючи візуальну оцінку, або об'єктивно кількісними методами. Перевага полягає у застосуванні кількісних показників для оцінки ФВЛШ, щоб мінімізувати мінливість та сприяти більшій точності та точності вимірювання [5].

Неінвазивні методи оцінки включають:

– ехокардіограму. Це найпоширеніший тест, що використовується для вимірювання частки викиду. Під час ехокардіограми звукові хвилі

використовуються для створення зображень серця та крові, яка прокачується через серце;

– катетеризацію серця. Під час катетеризації серця тонку пластикову трубку (катетер) вставляють у кровоносну судину на руці або нозі, а потім обережно ведуть до серця. Зображення, зроблені під час катетеризації, можуть виміряти частку викиду серця;

– магнітно-резонансну томографію (МРТ). МРТ використовує магнітне поле та радіохвилі для створення зображень поперечного перерізу певних частин тіла. Коли МРТ використовується для дослідження серця, це називається серцево-судинною МРТ;

– комп'ютерну томографію (КТ). КТ використовує рентгенівські промені для створення зображень поперечного перерізу певних частин вашого тіла. Коли КТ використовується для дослідження серця, це відоме як серцева КТ;

– скануванню ядерної медицини (Nuclear medicine scan). Під час ядерного сканування в кров потрапляє незначна кількість радіоактивного матеріалу. Потім спеціальні камери виявляють радіоактивний матеріал у вашій крові, коли він протікає через ваше серце та легені.

В даний час для вимірювання ЛШН можна використовувати різні методи за допомогою ехокардіографії. Використовувані методи відрізняються один від одного залежно від типу використовуваного ехокардіографічного зображення та рівнянь, що використовуються для визначення об'ємів лівого шлуночка. У свою чергу виділяють такі типи ехокардіографічних зображень:

- м-режим;
- двовимірний;
- тривимірний.

Отримані вимірювання можуть бути лінійними (одновимірними), площинними (двовимірними) або об'ємними (тривимірними) вимірами. Відсутність іонізуючого випромінювання діє на користь ехокардіографії.

На рисунках 1.2, 1.3, 1.4 представлені приклади вимірювання об'єму лівого шлуночка на різних ехокардіографічних зображеннях.

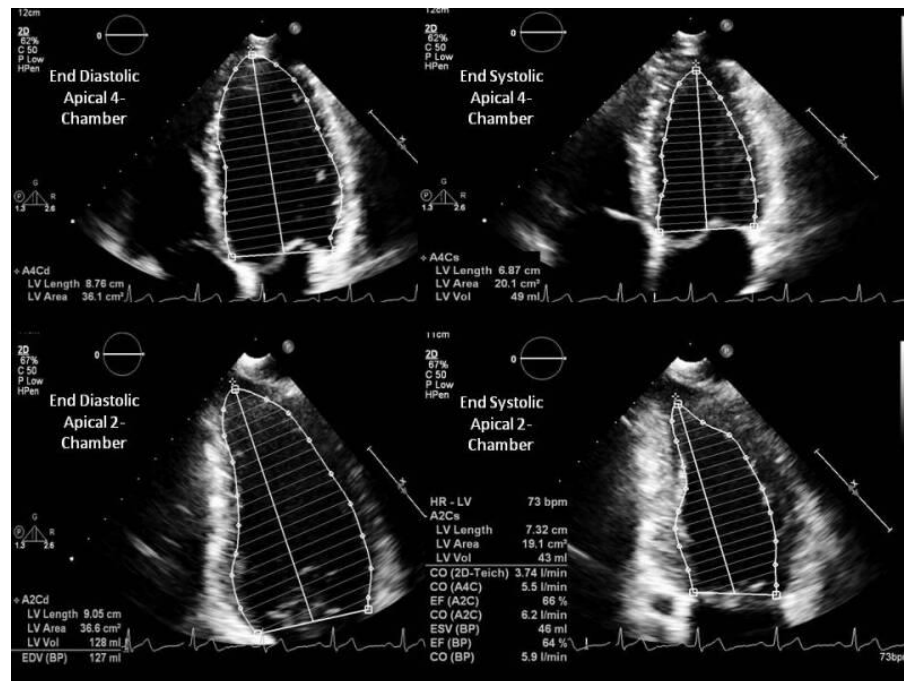


Рисунок 1.2 – Вимірювання об'єму лівого шлуночка за допомогою двовимірного ехокардіографічного зображення

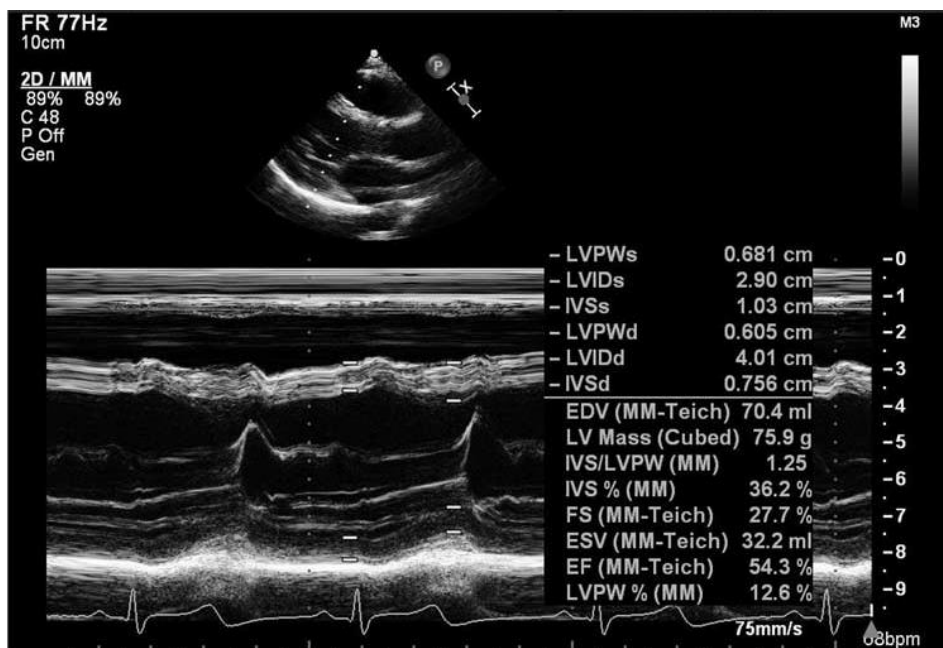


Рисунок 1.3 – Вимірювання об'єму лівого шлуночка в М-режимі

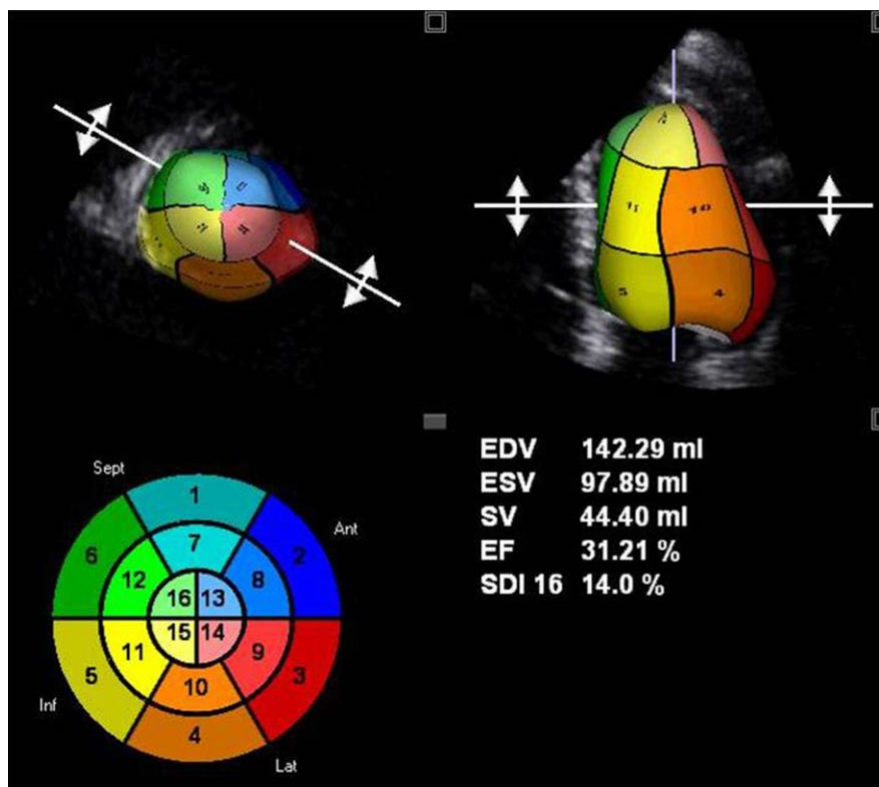


Рисунок 1.4 – Вимірювання об'єму лівого шлуночка за допомогою тривимірного ехокардіографічного зображення

Алгоритми, за допомогою яких можна розрахувати загальний коефіцієнт ФВЛШ, кінцевим діастолічний та кінцевим систолічний об'єм:

– модифіковане правило Сімпсона. Лівий шлуночок вважається сумою циліндра (від основи серця до мітрального клапана), усіченого конуса (від рівня мітрального клапана до рівня сосочкових м'язів) та іншого конуса віднесена до серцевої верхівки. Передбачається, що ці три секції мають однакову висоту;

– еліпсоїдна модель із використанням біпланних даних. Дві перпендикулярні площини ехо (мітральний клапан та апікальний вид) замінили дві ангіографічні проекції. Апікальна площина малої (перегородково-задньо-латеральної) осі виводиться з області зображення та його найбільшої довжини. Для цієї моделі мітральний план передбачається посередині між основою і вершиною. Мала вісь мітральної

площини виводиться із площі та перегородково-задньобічного розміру зображення мітрального листа;

– еліпсоїдна модель, що використовує дані з однією площиною: Площа та довжина апікального ехокардіографічного зображення були замінені на стандартне рівняння площі-довжини [5].

Біпланний метод дисків (модифіковане правило Сімпсона) є рекомендованим на сьогодні двовимірним методом для оцінки ФВЛШ. Модифікований метод Сімпсона (біпланний метод дисків) – це спосіб, що вимагає відстеження площі порожнини ЛШ. Американське товариство ехокардіографії рекомендує цей метод для вимірювання ФВЛШ. Цей метод вимагає вимірювання ФВЛШ шляхом простеження ендокардіальної межі як у верхівковому чотирикамерному, так і в двокамерному вигляді в кінці систоли та кінці діастоли. Ці трасування в решті-решт поділяють порожнину ЛШ на заздалегідь визначену кількість дисків (зазвичай 20). Обсяги дисків базуються на трасуваннях, отриманих в результаті дослідження. Деякі переваги цієї методики включають швидку оцінку візуалізації без впливу іонізуючого випромінювання, переносимість, неінвазивність та недорогість. Точність результатів збільшується за рахунок використання декількох оглядів та внутрішньовенного контрастування, що допомагає у кращому визначенні кордону ендокарда, що покращує визначення ФВЛШ.

Для того, щоб розрахувати фракцію викиду лівого шлуночка на двовимірному ехокардіографічному зображенні за допомогою біпланного методу дисків, необхідно на усьому серцевому циклі визначити зображення, які будуть відповідати діастолі та систолі лівого шлуночка. Коли дані зображення знайдені, медичний працівник вручну описує площу лівого шлуночка на кожному із зображень. Далі отримана площа шлуночка розділяється вздовж його довгою осі на серії дисків з однаковою висотою. Варто зазначити, що для підрахунків ФВЛШ за допомогою біпланного методу дисків, використовуються зображення апікального 2-камерного

виду (англ. Apical 2-chamber view – A2C) та апікального 4-камерного виду (англ. Apical 4-chamber view – A4C) (рисунок 1.5).

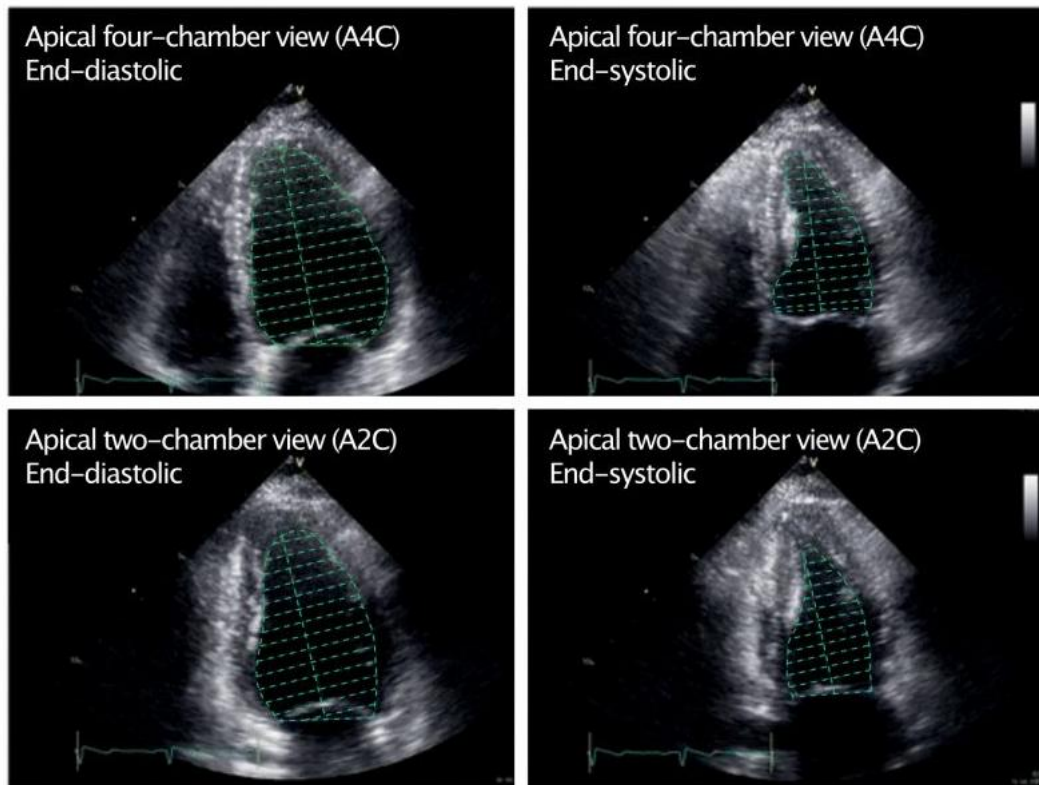


Рисунок 1.5 – Визначення площі ЛШ та поділ на n -ну кількість дисків на ехокардіографічних зображеннях A2C та A4C

На рисунку 1.6 зображений біпланний метод дисків, шлуночок на апікального 2-камерного виду та апікального 4-камерного виду розділений на рівну кількість дисків. a_i та b_i – це довжини i -го диску. L – довжина лівого шлуночка по довгій осі. Висота шлуночка приймається від середньої точки прямої лінії через кільце мітрального клапана до самої дистальної точки верхівки ЛШ.

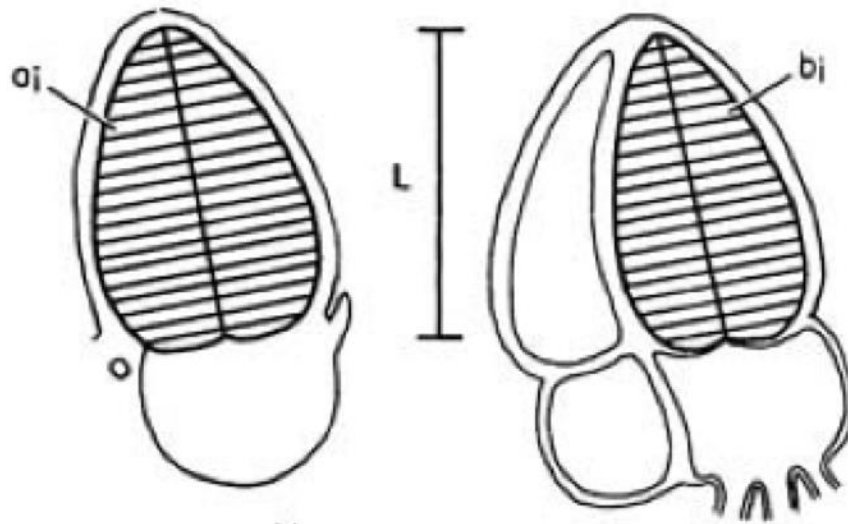


Рисунок 1.6 – Біпланний метод дисків

Об'єм кожного диска отримується добутком площі та висоти диска (формула 1.3):

$$Disk V_i = \frac{\pi \cdot (a_i \cdot b_i) \cdot L}{4n}, \quad (1.3)$$

де a_i та b_i – це довжини i -го диску,

L – довжина лівого шлуночка по довгій осі,

n – обрана кількість дисків.

Остаточна формула для розрахунку об'єму лівого шлуночка виглядає так (формула 1.4):

$$LV V = \frac{\pi}{4} \sum_{i=1}^n \frac{(a_i \cdot b_i) \cdot L}{n}. \quad (1.4)$$

Таким чином рахується об'єм лівого шлуночка на двовимірних зображеннях. Після знаходження об'єму на зображеннях діастолі та систолі застосовується формула 1.2 для підрахунку фракції викиду.

1.3 Постановка задачі

Завданням кваліфікаційної роботи є дослідження існуючих методів для автоматичного підрахунку фракції викиду лівого шлуночка та реалізація алгоритму автоматичного підрахунку за допомогою моделей глибинного навчання та аналітичних методів обробки зображень.

Для досягнення поставленої мети необхідно розглянути наступні питання:

- проаналізувати існуючі проблеми в даній області та визначити шляхи їх вирішення;

- проаналізувати існуючі підходи щодо ручного та автоматичного підрахунку фракції викиду;

- розглянути існуючі архітектури моделей сегментації, їх модифікації, визначити переваги та недоліки;

- обрати архітектуру моделі сегментації, яка задовольняє для вирішення задачі сегментації лівого шлуночка та стійка до складностей, які виникають при сегментації медичних зображень;

- реалізувати обрану архітектуру моделі сегментації;

- провести порівняльний аналіз обраної архітектури з класичними методами;

- реалізувати алгоритм для автоматичного підрахунку на основі обраної архітектури моделі сегментації з аналітичним підходом для підрахунку фракції викиду лівого шлуночка з отриманих даних з моделі сегментації.

2 МЕТОДИ ТА АЛГОРИТМИ СЕГМЕНТАЦІЇ ЛІВОГО ШЛУНОЧКА

2.1 Методи сегментації лівого шлуночку. Архітектура U-Net

Підрахунок фракції викиду (EF) лівого шлуночка (LV) вимагає точного окреслення ендокарда лівого шлуночка як в кінцевій діастолі (ED), так і в кінцевій систолі (ES). Для того, щоб автоматично розраховувати фракцію викиду на двовимірних зображеннях, необхідно спочатку сегментувати лівий шлуночок на усіх фреймах у ехокардіографічному відео для 2- та 4-х камерних видів. Після цього потрібно підрахувати площини отриманих сегментів на кожному фреймі серцевого циклу та обрати фрейми з мінімальною та максимальною площиною ЛШ окремо для 2- та 4-х камерних видів. Отримані фрейми будуть відповідати діастолі та систолі ЛШ.

Сегментація – це процес, який розподіляє зображення на регіони [9]. Це підхід до обробки зображень, який дозволяє розділяти об'єкти та текстури на зображеннях. Сегментація лівого шлуночка (ЛШ) дозволяє обчислити безліч серцевих параметрів, а також створити високоякісні анатомічні моделі для хірургічного планування, керівництва та регіонального аналізу серця. Незважаючи на те, що ручне розмежування шлуночка вважається золотим стандартом, воно забирає багато часу і дуже сприйнятливим до мінливих та внутрішньо-спостережних змін. Отже, необхідні швидкі, надійні та точні напівавтоматичні або повністю автоматичні алгоритми сегментації.

Існує багато традиційних способів виконати сегментацію зображення. Серед них є: точкові, лінійні та крайові методи виявлення, порогове значення, кластеризація на основі регіонів, піксельна кластеризація, морфологічні підходи тощо. На даний час розроблено безліч методів сегментації із згортковими нейронними мережами, які стали необхідними для вирішення більш складних завдань з сегментацією зображень.

Для глибинного навчання зазвичай потрібні великі масиви даних для навчання моделей. Але часто стає питання де дістати достатню кількість даних для навчання якісних моделей глибинного навчання, особливо, коли мова йде про вузькоспеціалізовану галузь, де збір даних можливий тільки під керівництвом висококваліфікованого спеціаліста або самим спеціалістом. Однією з таких галузей є медицина. Але окрім цього, в потрібно пам'ятати про закон про захист персональних даних пацієнта. Тому всі дані пацієнтів, які підготовлюються спеціалістом для подальшого використання іншими спеціалістами у глибинному навчанні повинні бути анонімізованими. Виникає необхідність в такій архітектурі нейронної мережі, яка дозволить навчити достатньо точну на якісну модель на відносно невеликому наборі даних.

U-Net – це архітектура сегментації зображень, розроблена головним чином для медичного аналізу зображень, яка дозволяє точно сегментувати зображення, використовуючи обмежену кількість навчальних даних [10]. U-Net є більш успішним, ніж звичайні моделі, з точки зору архітектури та з точки зору сегментації зображень на основі пікселів, сформованих із згорткових нейронних мережевих шарів. Ці риси забезпечують U-Net дуже високу корисність для медичного візуалізаційного співтовариства і призвели до широкого прийняття U-Net як основного інструменту для завдань сегментації в медичній візуалізації. Успіх U-net виявляється в його широкому використанні на всіх основних медичних зображеннях, від КТ та МРТ до рентгенівських променів та мікроскопії. Крім того, хоча U-Net в основному є інструментом сегментації, існували випадки використання U-Net в інших програмах.

U-Net бере назву від архітектури, яка при візуалізації здається подібною до літери U, як показано на рисунку 1.7. Вхідні зображення отримуються як сегментована вихідна карта. Найбільш особливий аспект архітектури у другій половині. Мережа не має повністю підключеного

рівня (fully connected layer). Використовуються лише згорткові шари. Кожен стандартний процес згортки активується функцією активації ReLU.

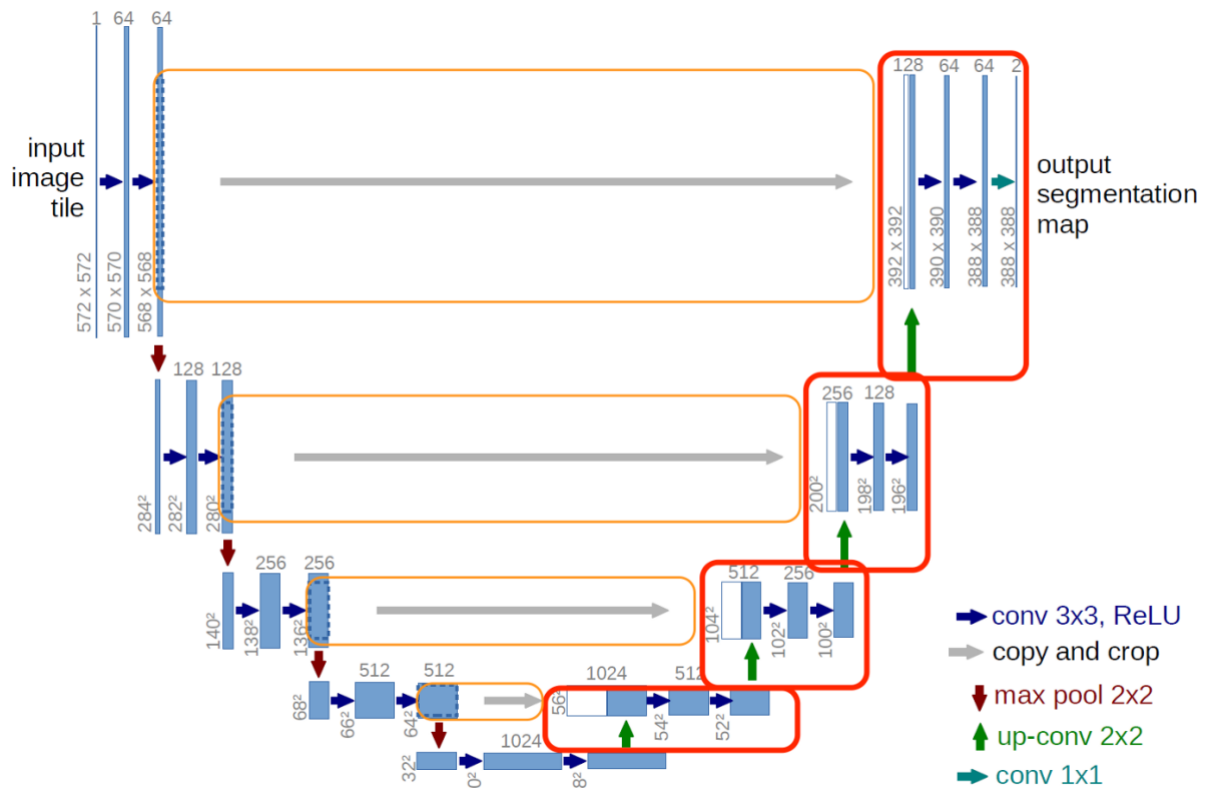


Рисунок 2.1 – Візуалізація класичної моделі U-Net

Процес зменшення розмірності по висоті та ширині, який застосовується по всій згортковій нейронній мережі, тобто шару об'єднання, застосовується у вигляді збільшення розміру у другій половині моделі. Шар об'єднання зменшує інформацію про висоту і ширину, підтримуючи постійну кількість каналів вхідної матриці. Шар об'єднання – це крок, який використовується для зменшення складності (кожен елемент матриці зображення називається пікселем). Таким чином, шар об'єднання відноситься до пікселя, який представляє групи пікселів.

Шар згортки є фундаментальним компонентом архітектури конволюційних нейронних мереж, який виконує виділення ознак та

зазвичай складається з комбінації лінійних та нелінійних операцій, тобто операції згортки та функції активації [11]. Згортка – це спеціалізований тип лінійних операцій, що використовується для вилучення ознак, де на вхідних даних застосовується невеликий масив чисел, який називається ядром. Поелементно добуток між кожним елементом ядра та вхідним масивом даних обчислюється в кожному розташуванні масиву і підсумовується для отримання вихідного значення у відповідному положенні вихідного масиву, що називається картою ознак. Ця процедура повторюється із застосуванням кількох ядер для формування довільної кількості карт функцій, які представляють різні характеристики вхідних масивів. Отже, різні ядра можна розглядати як різні екстрактори особливостей. Два ключових гіперпараметри, що визначають операцію згортки, – це розмір ядра та кількість. Перший, як правило, становить 3×3 , але іноді 5×5 або 7×7 . Останній є довільним і визначає глибину вихідних характеристичних карт.

Звичайна операція згортки без додаткових гіперпараметрів не дозволяє центру кожного ядра перекривати крайній елемент вхідного масиву і зменшує висоту та ширину вихідної карти функцій порівняно з вхідним масивом. Заповнення (padding), як правило, нульове заповнення, – це техніка вирішення цієї проблеми, коли рядки та стовпці нулів додаються з кожного боку вхідного масиву, щоб відповідати центру ядра на самому зовнішньому елементі і зберігати однакову в площині розмірність за допомогою операції згортки. Сучасні архітектури конволюційних нейронних мереж зазвичай використовують нульове заповнення, щоб зберегти площинні розміри. Без нульового заповнення кожна наступна карта об'єктів буде зменшуватися після операції згортки.

Відстань між двома послідовними положеннями ядра називається кроком (stride), який також визначає операцію згортки. Загальним вибором кроку є 1, однак іноді використовується крок, більший за 1, для того, щоб досягти дискретизації функціональних карт. Альтернативним методом для проведення дескрипції є операція об'єднання.

Ключовою особливістю операції згортки є розподіл ваги: ядра розподіляються по всіх позиціях зображення. Розподілення ваги створює наступні характеристики операцій згортки [12]:

- забезпечення інваріантності локальних шаблонів ознак, отриманих за допомогою трансляції ядер, коли ядра переміщуються по всіх позиціях зображення і виявляють вивчені локальні шаблони;
- вивчення просторових ієрархій шаблонів ознак за допомогою понижувальної дискретизації в поєднанні з операцією об'єднання, що призводить до захоплення все більшого поля зору;
- підвищення ефективності моделі за рахунок зменшення кількості параметрів для вивчення в порівнянні з повністю підключеними (fully connected) нейронними мережами.

У звичайній згортковій нейронній мережі U-Net (ліва частина моделі) кожен нейрон пов'язаний з усіма нейронами попереднього шару, і кожний зв'язок має свій ваговий коефіцієнт. У згорткових нейронних мережах в операціях згортки використовується матриця ваг невеликого розміру, яка «рухається» по всьому вхідному зображенню. Шар згортки виконує додавання результатів поелементного добутку кожного фрагмента зображення на матрицю – ядро згортки. Вагові коефіцієнти ядра згортки невідомі і встановлюються в процесі навчання [13].

Оператор згортки обчислюється за формулою 2.1:

$$(I \cdot K)_{xy} = \sum_{i=1}^h \sum_{j=1}^w K_{i,j} \cdot I_{x+i-1,y+j-1}, \quad (2.1)$$

де I – двовимірне зображення розміром $x \times y$,

K – матриця розмірності $h \times w$ (ядро згортки).

Операція активації являє собою нелінійну функцію і визначає вихідний сигнал нейрона. Нейрони з цією функцією активації називаються ReLU (rectified linear unit) і мають наступну формулу 2.2 [14]:

$$f(x) = \max(0, x). \quad (2.2)$$

Дана функція активації істотно прискорює процес навчання за рахунок простоти обчислень. Шар об'єднання (pooling) здійснює зменшення розміру зображення в n разів по кожній осі шляхом об'єднання сусідніх пікселів в областях розміром $n \times n$ в один піксель. У якості отриманого пікселя може використовуватися середнє значення об'єднаних пікселів або їх максимум.

Шар об'єднання забезпечує типову операцію зниження дискретизації, яка зменшує розмірність площини карт функцій, щоб ввести незмінність перекладу до невеликих зсувів та спотворень, а також зменшити кількість наступних параметрів, що піддаються навчанню. Слід зазначити, що в жодному з шарів об'єднання немає параметра, що піддається навчанню, тоді як розмір фільтра, крок та відступ – це гіперпараметри в операціях об'єднання, подібно до операцій згортки.

Найпопулярнішою формою операції об'єднання є максимальне об'єднання, яке вилучає елемент з вхідних карт функцій, виводить максимальне значення в кожному елементі та відкидає всі інші значення. На практиці зазвичай використовується максимальне об'єднання з фільтром розміром 2×2 та кроком 2. Це зменшує розмір площини карт функцій у два рази. На відміну від висоти та ширини, розмір глибини карт об'єктів залишається незмінним.

В U-Net при виконанні пулінгу застосовується максимізація по області 2×2 пікселя. Після проходження декількох серій згорткових і субдискретизуючих шарів вхідне зображення перетворюється від конкретної сітки пікселів з високою роздільною здатністю до більш абстрактних карток ознак.

Далі в декодуючій частині мережі (права частина) використовуються серії шарів, аналогічних із шарами кодуєчої (лівої) частини, тільки замість субдискретизуючого шару, що зменшує розмір зображення, застосовується

шар зворотної згортки, який збільшує розмір зображення в n разів по обох осях. Крім того, в декодуєчій частині виконується конкатенація карт ознак з картами ознак кодуєчої частини. На останньому рівні згортка 1×1 використовується для формування вихідного сегментованого зображення. Для того, щоб забезпечити безперешкодну операції згортки вихідної сегментації, важливо обрати розмір вхідного зображення таким чином, щоб усі операції максимального об'єднання 2×2 застосовувались до шару з рівним розміром x та y . Пікселі в прикордонній області симетрично додаються навколо зображення, щоб можна було безперервно сегментувати зображення. За допомогою цієї стратегії зображення сегментується повністю. Метод заповнення (додавання пікселів) важливий для застосування моделі U-Net до великих зображень, інакше роздільна здатність буде обмежена ємністю пам'яті графічного процесора.

Підсумовуючи характеристики моделі U-Net, можна виділи особливо корисною – це створення дуже детальних карт сегментації з використанням дуже обмежених зразків. Дана риса має велике значення для медичної спільноти, оскільки правильно марковані зображення часто обмежені [15]. Це досягається використанням випадкової еластичної деформації на навчальних даних, що дозволяє мережі вивчати ці варіації, не вимагаючи нових маркованих даних. Іншою проблемою є відокремлення об'єктів одного класу, що вирішується шляхом застосування зваженої функції втрат, яка карає модель, якщо не вдається розділити два об'єкти [16]. U-Net також набагато швидше навчається, ніж більшість інших моделей сегментації, завдяки навчанню на основі контексту.

За допомогою U-Net моделей можливо отримати зображення-маски для кожного фрейму відео послідовності серцевого циклу, віднайти за розміром площі фрейми діастолі та систолі. Наступним кроком після отримання фреймів з масками буде застосування аналітичних методів для роботи із зображеннями для виділення контуру, знаходження довжини ЛШ

та розбиття усієї площини ЛШ на рівні за висотою диски та застосувати Біпланний метод для підрахунку фракції викиду [14].

Аналітичні методи є достатньо точними у підрахунках, тому похибка у вимірах може бути через помилки сегментації у моделі глибинного навчання. Тому модель для отримання сегментів ЛШ повинна бути точною та стійною до складностей, які можуть статися при сегментації медичних зображень[16],[17].

2.2 Навчання U-Net моделі

Етикетки даних та основних істин є найважливішими компонентами досліджень із застосуванням методів глибокого навчання чи інших методів машинного навчання. Ретельний збір даних та створення етикеток для них, за допомогою яких можна навчити та протестувати модель, є обов'язковим для успішного проекту глибокого навчання, однак отримання високоякісних маркованих даних може бути дорогим та трудомістким. Незважаючи на те, що може існувати безліч наборів даних медичних зображень, відкритих для громадськості, у цих випадках особливу увагу слід приділяти якості основних етикеток.

Набір даних, як правило, поділяються на три набори: навчальний, валідаційний та тестовий набір, хоча існують деякі варіанти, такі як перехресне підтвердження. Навчальний набір використовується для навчання мережі, де значення втрат обчислюються шляхом прямого розповсюдження, а параметри, що навчаються, оновлюються за допомогою зворотного розповсюдження. Валідаційний набір використовується для оцінки моделі під час тренувального процесу, тонкої настройки гіперпараметрів та виконання вибору моделі. Тестовий набір в ідеалі використовується лише один раз після навчання моделі, щоб оцінити ефективність остаточної моделі, яка була доопрацьована та обрана в процесі навчання за допомогою навчальних та валідаційних наборів.

Навчання мережі – це процес пошуку ядер у шарах згортки та ваг у повністю зв’язаних шарах (fully connected layers), що мінімізують різницю між передбаченнями вихідних даних та заданими мітками істинності в наборі навчальних даних. Алгоритм зворотного розповсюдження – це метод, який зазвичай використовується для навчання нейронних мереж, де функція втрат та алгоритм оптимізації градієнтного спуску відіграють важливу роль. Ефективність моделі для конкретних ядер та ваг обчислюється функцією втрат шляхом прямого розповсюдження на навчальному наборі даних, а параметри, що підлягають навчанню, а саме ядра та ваги, оновлюються відповідно до величини втрат за допомогою алгоритму оптимізації, який називається зворотним розповсюдженням або градієнтним спуском.

Функція збитків, яку також називають функцією витрат, вимірює сумісність між вихідними прогнозами мережі через пряме розповсюдження та заданими основними мітками істинності. Тип функції збитків є одним із гіперпараметрів і його потрібно визначати відповідно до заданої задачі.

Функція втрат є однією з важливих складових навчання моделей медичної сегментації зображень. Функції втрат визначають, як моделі нейронних мереж обчислюють загальну похибку з їх залишків для кожної навчальної партії. Це, в свою чергу, впливає на те, як вони коригують свої коефіцієнти при зворотному розповсюдженні, тому вибір функції втрат має прямий вплив на продуктивність моделі.

Перехресна ентропія [18] визначається як міра різниці між двома розподілами ймовірностей для даної випадкової величини або набору подій. Вона широко використовується для цілей класифікації, і оскільки сегментація є класифікацією на рівні пікселів, дана функція також застосовується. Найчастіше використовувана функція втрат для завдання сегментації зображень – це бінарна перехресна ентропія для пікселів. Ця втрата вивчає кожен піксель окремо. Дана функція витрат визначається за формулою 2.3:

$$L_{BCE(y,y')} = -(y \log(y') + (1 - y) \log(1 - y')), \quad (2.3)$$

де y' – передбачене значення моделі.

Збалансована поперечна ентропія (BCE) являється різновидом перехресна ентропія та подібна до зваженої поперечної ентропії. Єдина відмінність полягає в тому, що в цьому, крім позитивних прикладів, також зважуються і негативні приклади. Збалансовану перехресну ентропію можна визначити наступним чином (формула 2.4):

$$L_{BCE(y,y')} = -(\beta \cdot y \log(y') + (1 - \beta) \cdot (1 - y) \log(1 - y')), \quad (2.4)$$

де β – коефіцієнт ваги.

Інша функція втрат, яка широко використовується для завдань сегментації зображень, базується на коефіцієнті *Dice*, який є мірою перекриття двох зразків. Цей показник коливається від 0 до 1, де коефіцієнт *Dice* 1 позначає ідеальне та повне перекриття. Коефіцієнт *Dice* був спочатку розроблений для бінарних даних і може бути розрахований як (формула 2.5):

$$L_{Dice} = \frac{2|A \cap B|}{|A| + |B|}, \quad (2.5)$$

де $|A \cap B|$ – загальні елементи між множинами A і B ,

$|A|$ – кількість елементів у множині A (і також для множини B).

Для випадку оцінки коефіцієнта *Dice* на передбачуваних масках сегментації ми можемо наблизити $|A \cap B|$ як елементальне множення між передбаченням і цільовою маскою, а потім підрахувати отриману матрицю.

Також існує комбінація між вище представленими функціями витрат, яка має назву *BCE-Dice Loss* [19]. Дана втрата поєднує втрати *Dice* зі стандартною бінарною перехресною ентропією (*BCE*), яка зазвичай є

типовими для моделей сегментації. Поєднання двох методів забезпечує певне розмаїття втрат, водночас отримуючи вигоду від стабільності BCE . Дана функція витрат визначається за формулою 2.6:

$$L_{BCEDice} = \lambda(L_{BCE}) + \gamma(1 - L_{Dice}), \quad (2.6)$$

де λ та γ – вага для індивідуальної функції витрат.

Показник перетину через союз (Intersection over Union, IoU), який також називають індексом Жакарда, є методом кількісної оцінки процентного перекриття між цільовою маскою та результатами прогнозування. Дана метрика тісно пов'язана з коефіцієнтом Dice, який часто використовується як функція втрат під час тренування. Метрика IoU вимірює кількість загальних пікселів між цільовою маскою та масками передбачення, поділену на загальну кількість пікселів, присутніх в обох масках. Показник перетину через союз визначається формулою 2.7:

$$IoU = \frac{A \cap B}{A \cup B}, \quad (2.7)$$

де A – цільова маска,

B – результуюча маска.

Перетин $A \cap B$ складається з пікселів, знайдених як у масці передбачення, так і в масці основної істини, тоді як об'єднання $A \cup B$ просто складається з усіх пікселів, знайдених або в масці передбачення, або в цільовій масці. Оцінка IoU обчислюється для кожного класу окремо, а потім усереднюється по всіх класах, щоб отримати загальний, середній бал IoU нашого семантичного прогнозування сегментації.

Альтернативна метрика для оцінки семантичної сегментації полягає у простому повідомленні про відсоток пікселів на зображенні, які були правильно класифіковані. Точність пікселів зазвичай повідомляється для

кожного класу окремо, а також в цілому для всіх класів. Розглядаючи точність пікселів для кожного класу, оцінюється бінарна маска; *True Positive (TP)* представляє піксель, який правильно передбачає належність до даного класу (відповідно до цільової маски), тоді як *True Negative (TN)* представляє піксель, який правильно визначено як такий, що не належить до даного класу, *False Positive (FP)* – це кількість пікселів, які були класифіковані як належні до даного класу, хоча фактично вони не належать, а *False Negative (FN)* – кількість пікселів, помилково класифікованих як належних до класу. Формула 2.8 описує підрахунок точності:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2.8)$$

Ця метрика може іноді давати оманливі результати, коли представлення класу невелике на зображенні, оскільки міра буде необ'єктивною, в основному повідомляючи про те, наскільки добре було ідентифіковано негативні випадки (тобто ті регіони, де класу немає).

Подібно до коефіцієнта *Dice*, *Sensitivity Specificity Loss* широко використовуються для оцінки прогнозів сегментації. У цій функції втрат вирішується проблема дисбалансу класу, використовуючи параметр w . Функція втрат визначається за формулою 2.9:

$$SSL = w \cdot \frac{TP}{TP+FN} + (1 - w) \cdot \frac{TN}{TN+FP}, \quad (2.9)$$

де w – коефіцієнт ваги.

2.3 Аугментація даних

Також важливим інструментом для навчання моделей глибокого навчання є аугментація даних, особливо коли основна вибірка даних

обмежена, що присутнє медичним наборам даних. Збільшення даних – це процес, який дозволяє збільшити обсяг навчальних даних шляхом внесення деяких обґрунтованих модифікацій або перетворень у існуючі дані [20]. У медичній візуалізації це зазвичай робиться з трансформаціями, які однаково застосовуються як до зображень, так і до етикеток, створюючи викривлені версії навчальних даних. Методи збільшення часто використовують трансформації, такі як обертання, відбиття та еластичні деформації, які створюють навчальні зображення, які дуже нагадують один конкретний приклад навчання. Незважаючи на те, що зображення, сформовані таким чином, помітно відрізняються від навчальних зображень, ця техніка збільшення допоміжна, що покращує ефективність виконання різноманітних завдань машинного навчання.

Для того, щоб успішно побудувати добре узагальнюючі глибинні моделі, потрібен величезний обсяг фактичних даних, щоб уникнути перенавчання (overfitting) моделі та «запам'ятовування» навчальних даних і, як результат, неможливість генералізувати дані з тренувального набору.

Алгоритми збільшення даних для сегментації можна розділити на наступні основні категорії [21]:

- алгоритми, що використовують різні трансформації вихідних даних, включаючи афінні трансформації зображень;
- еластичні перетворення;
- перетворення на рівні пікселів.

У афінних підходах реальні дані зображень піддаються різним операціям (обертання, масштабування, обрізання, гортання або переклади), щоб збільшити кількість прикладів тренувань. Такі традиційні методи збільшення даних у цілому дають дуже корельовані зображення, тому можуть запропонувати мало покращень для процесу навчання глибокої мережі та подальшого узагальнення за тестовими даними. Крім того, вони також можуть генерувати анатомічно неправильні приклади, наприклад,

використовуючи обертання. Тим не менше, афінні перетворення зображень є тривіальними для реалізації як у 2D, так і у 3D зображеннях.

Випадкове перевертання створює дзеркальне відображення вихідного зображення вздовж однієї (або більше) вибраної осі. Зазвичай реальні зображення можна перевертати вздовж горизонтальної осі, що не стосується вертикальної, оскільки частини зображення вгорі і внизу не завжди є «взаємозамінними». Подібним чином у цьому контексті можна використовувати поворот зображення на кут α навколо центрального пікселя. Ця операція супроводжується відповідною інтерполяцією відповідно до вихідного розміру зображення.

Операція перекладу зміщує ціле зображення на задану кількість пікселів у вибраному напрямку, одночасно застосовуючи відступ. Це дозволяє мережі не концентруватися на особливостях, присутніх переважно в одній просторовій області, але змушує модель навчитися просторово-інваріантних ознак. Як і у випадку з обертанням переклад зображення на задану кількість пікселів вздовж обраної осі (або осей) може створити корисні зображення. Однак ця процедура може бути не «корисною» для всіх архітектури глибоких мереж – згорткові нейронні мережі використовують згортки та операції об'єднання, які є внутрішньо просторово-інваріантними.

Введення масштабованих версій оригінальних зображень до навчального набору може допомогти глибокій мережі засвоїти цінні глибокі особливості незалежно від їх початкового масштабу. Оскільки об'єкти на зображенні можуть різнитися за розміром, масштабування дійсно може привести до корисних доповнень зображень до навчального набору. Оскільки різні глибокі архітектури вимагають зображень постійного розміру, масштабування зазвичай поєднується з обрізанням, щоб зберегти початкові розміри зображення. Крім того, обрізання може обмежити поле зору лише тими частинами зображення, які є важливими.

Існують методи збільшення, які не змінюють геометричну форму зображення (отже, всі геометричні особливості залишаються незмінними

під час процесу збільшення), але впливають на значення інтенсивності пікселів (або локально, або по всьому зображенню). Такі операції можуть бути особливо корисними при аналізі медичних зображень, коли різні навчальні зображення отримуються в різних місцях та за допомогою різних сканерів, отже, можуть бути внутрішньо неоднорідними за інтенсивністю пікселів, градієнтами інтенсивності або «насиченістю». Під час збільшення рівня пікселів інтенсивність пікселів зазвичай змінюється за допомогою випадкового або нульового середнього гауссового шуму (зі стандартним відхиленням, що відповідає відповідному виміру даних), із заданою ймовірністю (попередня операція називається випадковою інтенсивністю варіація). Інші операції на рівні пікселів включають зсув та масштабування значень інтенсивності пікселів (і модифікацію яскравості зображення), застосування гамма-корекції та її декількох варіантів, різкість, розмиття тощо. Цей тип збільшення даних часто використовується для даних з великими розмірами, оскільки його можна зручно застосувати до вибраних вимірів.

Для полегшення проблем, пов'язаних з основними підходами до збільшення даних (включаючи проблему створення корельованих вибірок даних), були запропоновані різні підходи до створення штучних даних (GAD). Генеративні змагальні мережі (GAN) широко використовуються для розширення наборів медичних даних. Основною метою GAN є створення нового прикладу даних (за допомогою генератора), який неможливо відрізнити від реальних даних дискримінатором (генератор конкурує з дискримінатором). Існує така архітектура GAN, яка використовує генератор грубого до тонкого, який має на меті зафіксувати різноманітність навчальних даних та створити доповнені приклади. Незважаючи на те, що GAN дозволяють ввести незмінність та стійкість глибинних моделей щодо не лише афінних перетворень (наприклад, обертання, масштабування чи перевертання), але й деяких варіацій форми та зовнішнього вигляду, збіжність змагального тренування та точки його

рівноваги залишаються відкриті питання. Нарешті, існують сценарії, в яких генератор подає безліч дуже подібних прикладів, які не можуть поліпшити узагальнення системи – це відоме як проблема колапсу режиму.

2.4 Модифікації U-Net архітектури

Оскільки потенціал U-Net все ще зростає, створюється все більше нових модифікацій класичної архітектури, які влучно використовувати для конкретної задачі, враховуючи її особливості [22].

Attention U-Net. Часто бажаною рисою в мережі обробки зображень є здатність зосереджуватися на конкретних об'єктах, які мають важливе значення, ігноруючи непотрібні ділянки. Мережа Attention досягає цього, використовуючи ворота уваги. Механізм уваги використовується для збільшення можливостей CNN у різних завданнях комп'ютерного зору, включаючи класифікацію, виявлення, сегментацію, субтитри до зображень та візуальну відповідь на питання. Механізми уваги спрямовують модель на підкреслення найбільш помітних рис, корисних для конкретних завдань, уникаючи марних особливостей. Ворота уваги – це блок, який обробляє функції, які не мають відношення до поточного завдання. Ворота уваги бувають у двох формах: жорсткій та м'якій. Жорстка увага працює на основі виділення відповідних регіонів шляхом обрізання зображення або пропозиції ітеративного регіону. Оскільки жорстка увага може вибрати лише одну область зображення за раз, це має два наслідки, вона недиференційована і вимагає посилення навчання, і це означає, що для даного регіону на зображенні мережа може або звертати «увагу», або ні, без проміжків між ними. Як результат, стандартне зворотне поширення не може бути здійснено, і для вибірки точності на різних етапах зворотного поширення необхідна вибірка методом Монте-Карло. Враховуючи, що точність залежить від того, наскільки якісно зроблено відбір проб, існує

потреба в інших техніках, таких як навчання підкріплення, щоб зробити модель ефективною.

М'яка увага працює шляхом зважування різних частин зображення. Області високої релевантності помножуються на більшу вагу, а області низької релевантності позначаються меншими вагами. У міру навчання моделі більша увага приділяється регіонам з більшими вагами. На відміну від жорсткої уваги, ці ваги можна застосовувати до багатьох ділянок зображення. Завдяки детермінованій природі м'якої уваги, вона залишається диференційованою і її можна навчити за допомогою стандартного зворотного поширення. Оскільки модель тренується, зважування також тренується так, що модель стає кращою при вирішенні, на які деталі звернути увагу.

Для поліпшення ефективності сегментації, Khened та Roth [23] поклалися на додаткові попередні моделі локалізації об'єктів для розділення локалізації та подальших етапів сегментації. Цього можна досягти, інтегрувавши ворота уваги поверх архітектури U-Net, не навчаючи додаткових моделей. На рисунку 2.2 представлена архітектура Attention U-Net.

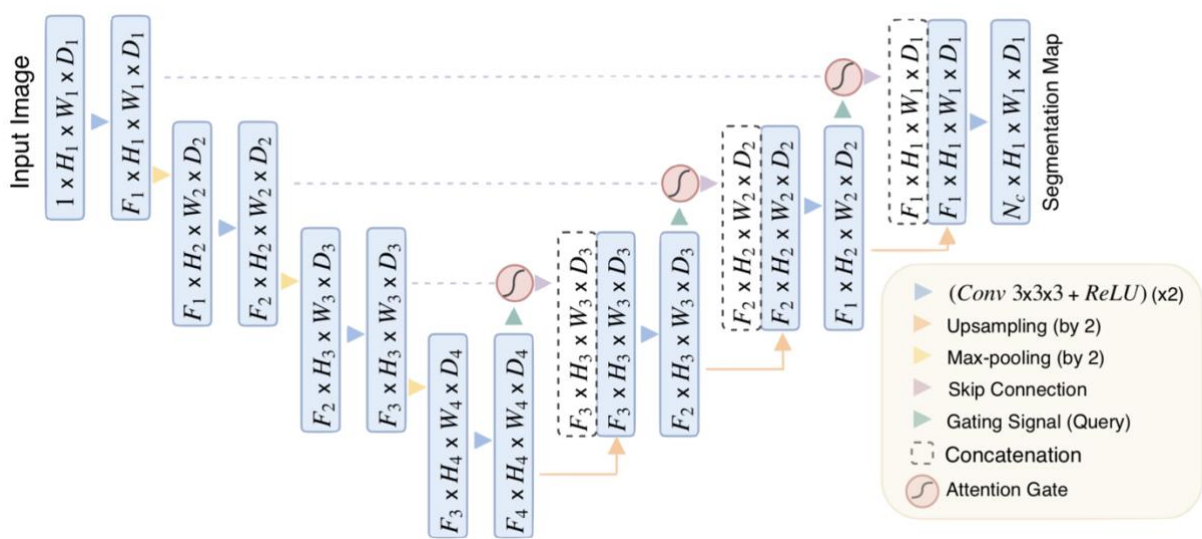


Рисунок 2.2 – Архітектура мережі Attention U-Net

Як результат, ворота уваги, вбудовані в U-Net, можуть покращити чутливість та точність моделі до пікселів переднього плану, не вимагаючи значних накладних витрат. Ворота уваги можуть поступово приглушати відповіді функцій у нерелевантних фонових регіонах.

Кожен шар в експансивному шляху має заслінку уваги, через яку повинні пройти відповідні ознаки із контракуючого шляху, перш ніж об'єднати об'єкти з пробними елементами в експансивному шляху. Повторне використання заслінки уваги після кожного шару значно покращує ефективність сегментації, не додаючи моделі занадто великої обчислювальної складності.

Блок уваги корисний у моделях кодерів-декодерів, таких як мережа U, оскільки він може надавати локалізовану інформацію про класифікацію на відміну від загальної класифікації. В U-Net це дозволяє різним частинам мережі зосередитися на сегментуванні різних об'єктів.

Inception U-Net. Більшість алгоритмів обробки зображень, як правило, використовують фільтри фіксованого розміру. Налаштування моделі для пошуку правильного розміру фільтра часто може бути громіздким. Більше того, фільтри фіксованого розміру підходять лише для зображень із помітними деталями подібного розміру. У багатьох додатках аналіз розглядає зображення з великими варіаціями форм і розмірів у видному регіоні. Одним із варіантів вирішення цієї проблеми було б використання більш глибоких мереж, які можуть читати деталі високого рівня в різних розмірах і формах. Однак такі глибокі мережі досить обчислювально витратні. Альтернативне рішення, яке називається Inception мережею, використовує фільтри різних розмірів на одному шарі в мережі. Кожен звичайний згортковий шар в початковій мережі U-Net замінюється модулем Inception, який включає безліч наборів згорток 3×3 , згортка 1×1 , максимальне об'єднання 3×3 і каскадних згорток 3×3 [24]. На рисунку 2.3 зображена архітектура мережі Inception U-Net. Набір цифр, показаний під

кожним початковим модулем, вказує загальну кількість використаних фільтрів, а також висоту, ширину та глибину вхідної карти.

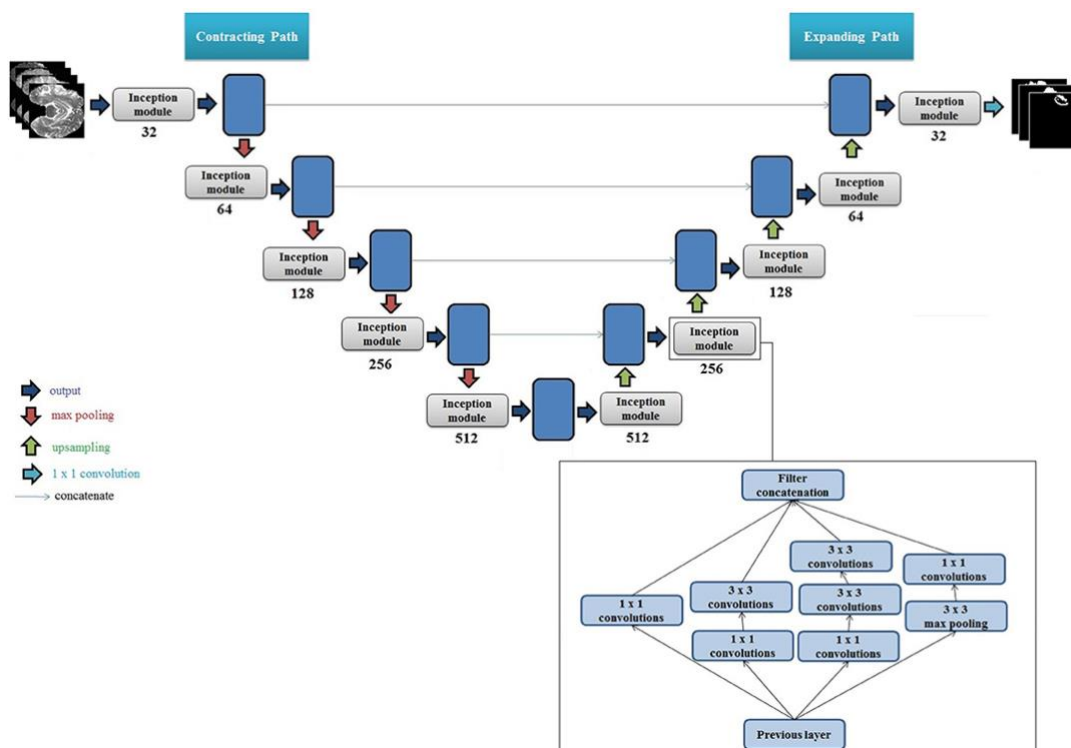


Рисунок 2.3 – Архітектура мережі Inception U-Net

Кількість фільтрів на кожному шарі подвоюється на стороні кодера, а розмір вихідної карти функцій (висота та ширина) зменшується вдвічі. Помноження на 4 для кожного значення глибини відбувається завдяки 4 варіаціям фільтра в модулі Inception, який генерує 4 набори однакових за розміром карт функцій. Потім карти функцій дескриптуються за допомогою максимального об'єднання, що зменшує їхню висоту та ширину вдвічі. Цей процес повторюється до досягнення вузького місця, після чого виконується операції U-Net, які знаходяться у центрі моделі. І навпаки, на шляху, що розширюється, висота і ширина карт об'єктів подвоюється, а глибина зменшується вдвічі на кожному шарі до досягнення результату – маски сегментації для даного вхідного зображення. Крім того, кожен набір карт функцій, сформований на шляху контракування, об'єднується у відповідні

карти об'єктів на шляху, що розширюється. В даній мережі також використовується функція активації ReLU для кожного шару та виконується нормалізацію партії (batch normalization) у кожному модулі Inception.

На останньому шарі на шляху, що розширюється, вихідні висота та ширина еквівалентні висоті та ширині вихідних вхідних зображень. Потім застосовується згортковий шар з розміром ядра 1×1 , щоб зменшити глибину останньої карти об'єктів, щоб дорівнювати кількості класів. Потім застосовується піксельна функція активації, яка потім перетворює зменшену карту функцій у бінарні зображення сегментації. На рисунку 2.3 знизу праворуч представлена внутрішня архітектура одного початкового модуля з декількома згортковими фільтрами та фільтрами максимального об'єднання. Числа в кожному блоці представляють розмір фільтра згортки.

Виходи різних фільтрів об'єднуються та передаються на наступний шар. Початкова мережа здатна досить ефективно аналізувати зображення з різними виділеними регіонами завдяки різним розмірам фільтрів. Щоб зменшити обчислювальну складність, початкова мережа додає згортку 1×1 перед кожним фільтром 3×3 або більшим для зменшення розмірності. Крім того, об'єднання шарів також може бути додано паралельно в кожному початковому модулі.

Residual U-Net. Цей варіант U-Net заснований на архітектурі ResNet. Мотивацією ResNet було подолання труднощів у навчанні дуже глибоких нейронних мереж. Відомо, що нейронні мережі здатні швидше сходиться до рішення, чим більше шарів присутній. Однак експериментальні результати показали, що збільшення кількості шарів призводить до насичення, а подальше збільшення може спричинити погіршення характеристик. Ця деградація виникає внаслідок втрати ідентичності ознак у глибших нейронних мережах, спричинених зменшенням градієнтів у ваговому векторі. ResNet зменшує цю проблему, використовуючи пропускання з'єднань, які беруть карту об'єктів з одного шару і додають її на інший шар глибше в мережі [25]. Така поведінка дозволяє мережі краще зберігати карти

функцій у глибших нейронних мережах та забезпечувати покращену продуктивність таких глибших мереж. У залишковій U-мережі на кожному блоці в мережі вхід до першого згорткового рівня додається до виходу з другого згорткового шару за допомогою пропускання з'єднання. Це пропускане з'єднання застосовується перед вибіркою вниз або підбіркою у відповідних шляхах в U-мережі. Використання залишкових з'єднань пропуску допомагає полегшити проблему зникаючого градієнта дозволяючи проектувати моделі U-Net з більш глибокими нейронними мережами. На рисунку 2.4 представлена архітектура Residual U-Net.

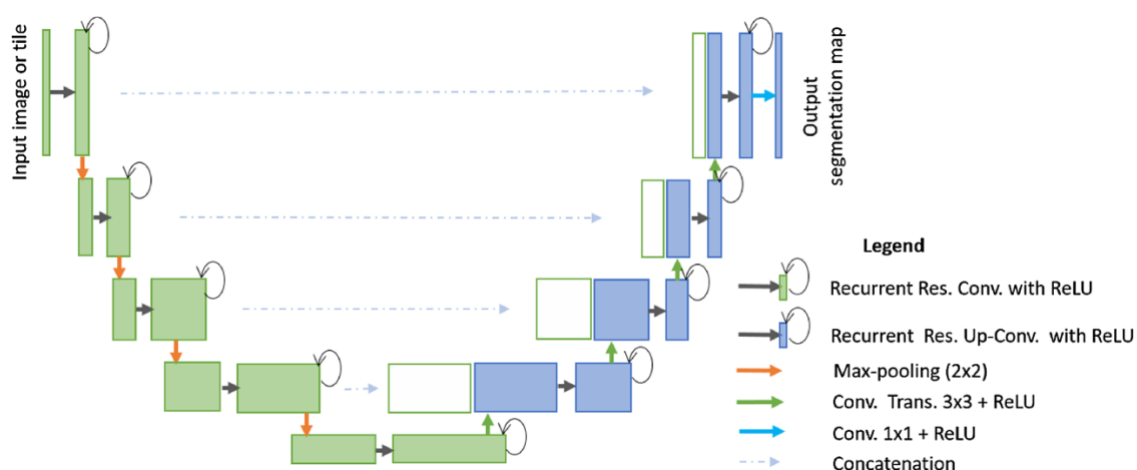


Рисунок 2.4 – Архітектура мережі Residual U-Net

Рекурентна згорткова нейронна мережа (RCNN) включає повторювані петлі зворотного зв'язку в згортковий шар. Зворотній зв'язок застосовується як після згортки, так і після активації, і подає карту об'єктів, створену фільтром, назад у відповідний шар. Властивість зворотного зв'язку дозволяє одиницям оновлювати свої карти функцій на основі контексту суміжних одиниць, забезпечуючи кращу точність та продуктивність.

Dense U-Net. Щільні U-мережі використовують блоки DenseNet [19] замість регулярних шарів. Хоча модель ResNet дозволяє більш глибокі нейронні мережі, вона не усуває проблему зникнення градієнтів. Архітектура ResNet також з часом погіршує продуктивність із збільшенням

шарів. DenseNet – це глибока архітектура навчання, побудована на базі ResNet із двома ключовими змінами. По-перше, кожен шар у блоці отримує об'єкт або карту ідентичності від усіх попередніх шарів. І друга важлива зміна полягає в тому, що карти ідентичності поєднуються шляхом конкатенації за каналами в тензори, на відміну від ResNet, де карти ідентичності підсумовуються шляхом складання по елементах. На рисунку 2.5 зображена архітектура мережі Dense U-Net.

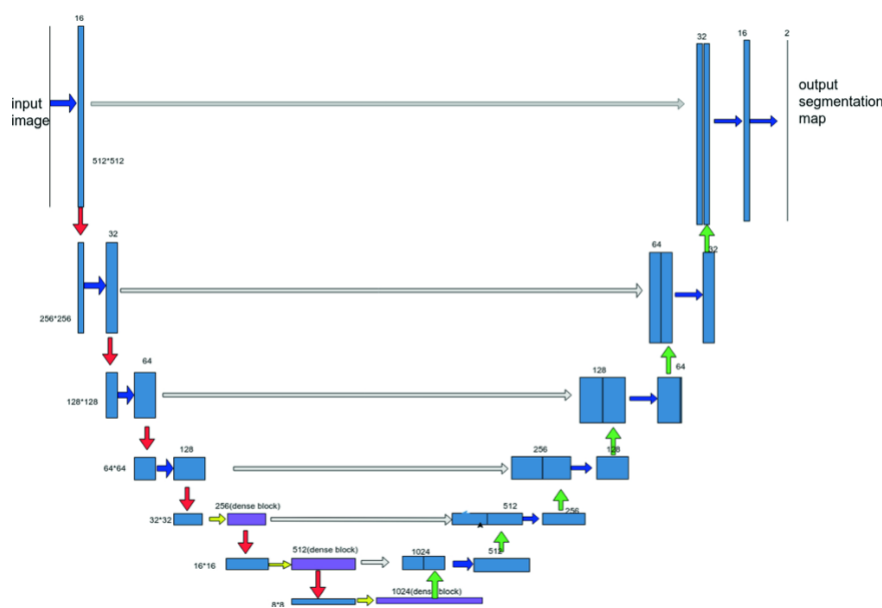


Рисунок 2.5 – Архітектура мережі Dense U-Net

Отже, відображення ідентичності кожного шару залежить не тільки від попереднього шару, але й від усіх шарів перед ним у блоці. Це дозволяє DenseNet зберігати всі карти ідентичності з попередніх шарів і суттєво сприяти розповсюдженню градієнта. Наслідком є те, що кожен шар може мати менше каналів, оскільки інформація легше зберігається між шарами, що призводить до вищої точності з меншою кількістю обчислень, і, як наслідок, може дозволити глибоке вивчення моделей з більшою кількістю шарів. При реалізації U-мережі кожен традиційний блок U-Net замінюється щільним блоком із двох або більше згорткових шарів. Прийняття щільних блоків дозволяє отримати більш глибокі моделі U-Net, які можуть

сегментувати об'єкти на зображенні з більшою відмінністю. Цей атрибут щільних U-Net є дуже популярним в аналізі медичних зображень через те, що предмети на таких зображеннях надзвичайно близькі і часто перекриваються. Застосування щільної U-мережі було виявлено при аналізі пухлин головного мозку, сегментації кровоносних судин сітківки, сегментації судин головного мозку, меланоми, раку легенів, рак печінки та багатоорганна сегментація.

U-Net ++ – ще одна потужна форма архітектури U-Net, натхненна DenseNet [26]. Вона використовує густу мережу з'єднань пропуску як проміжну сітку між контракуючим та експансивним шляхами. Це допомагає мережі, поширюючи більше семантичної інформації між двома шляхами, дозволяючи їй більш точно сегментувати зображення. У традиційній U-мережі карти функцій контракуючого шляху безпосередньо об'єднуються у відповідні шари експансивного шляху. Однак U-Net ++ має ряд вузлів пропуску з'єднання між кожним відповідним шаром. Архітектура U-Net ++ представлена на рисунку 2.6.

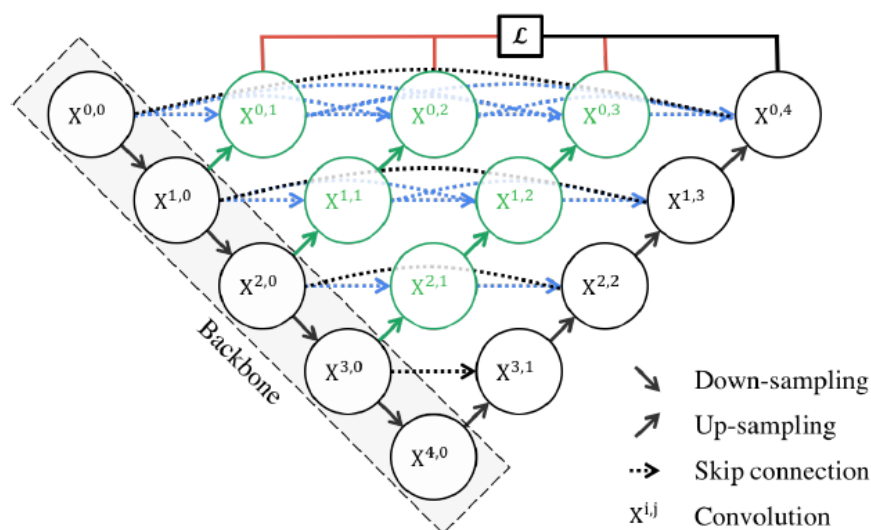


Рисунок 2.6 – Архітектура мережі U-Net ++

Кожен блок підключення пропуску отримує всі карти об'єктів від усіх попередніх блоків на одному рівні, а також пробну карту об'єктів від

безпосереднього нижнього блоку. Тому кожен рівень еквівалентний щільному блоку. Це розташування мінімізує втрату семантичної інформації між двома шляхами. Кількість проміжних блоків підключення пропусків залежить від кількості шару і лінійно зменшується, спускаючись по шляху стиснення. На відміну від базової U-мережі, яка має єдину пряму конкатенацію від скорочувального шляху до експансивного шляху, U-net ++ має ряд проміжних згорткових блоків між двома шляхами [27]. Кожен посередницький та експансивний блок отримує об'єднані карти функцій з усіх попередніх блоків на тому самому рівні, а також пробну карту функцій із блоку, що знаходиться безпосередньо під ним. U-Net ++ широко застосовуються в сегментації медичних зображень: сегментацію ядер клітин, ракової тканини, серцевих структур та судин та органів малого тазу.

MFP-Unet. Існує безліч модифікацій даної архітектури і усі вони створенні для вирішення певних складностей для тих чи інших зображень. Таким чином була створена архітектура для сегментації саме лівого шлуночку на ехокардіографічних зображеннях під назвою MFP-Unet [28]. Ця архітектура є багатофункціональною пірамідою U-Net (MFP-Unet) і являє собою поєднання розширених мереж U-Net та Feature Pyramid. Як згадувалося раніше, шлях стиснення архітектури U-Net створює додаткову кількість карт функцій, але в той же час роздільна здатність зменшується. Це є причиною застосування процесу об'єднання на кожному рівні шляху розширення. Поєднання інформації глибоких функцій та низькорівневих функцій з високою роздільною здатністю дозволяють алгоритму зберегти як точність, так і роздільну здатність. Незважаючи на таку архітектуру, U-Net використовує лише останні функції шару в шляху розширення для етапу сегментації (останній шар). Архітектура MFP-Unet пропонує змінити архітектуру U-Net, щоб карти об'єктів на всіх рівнях шляху розширення були включені окремо.

Розглядаючи архітектуру U-Net, MFP-Unet додає додаткові рівні згортки для вилучення карт функцій з усіх рівнів шляху розширення, щоб

бути включеними в процес сегментації останнього шару. Це включення обіцяє функціональна пірамідальна мережа. Шлях кодера залишається незмінним, але для розширення U-мережі використовувались розширені фільтри згортки. На шляху декодера шари архітектури U-Net залишилися незмінними, але додаються додаткові шари для підвищення точності та роздільної здатності кожного рівня на окремих картах функцій (рисунок 2.7).

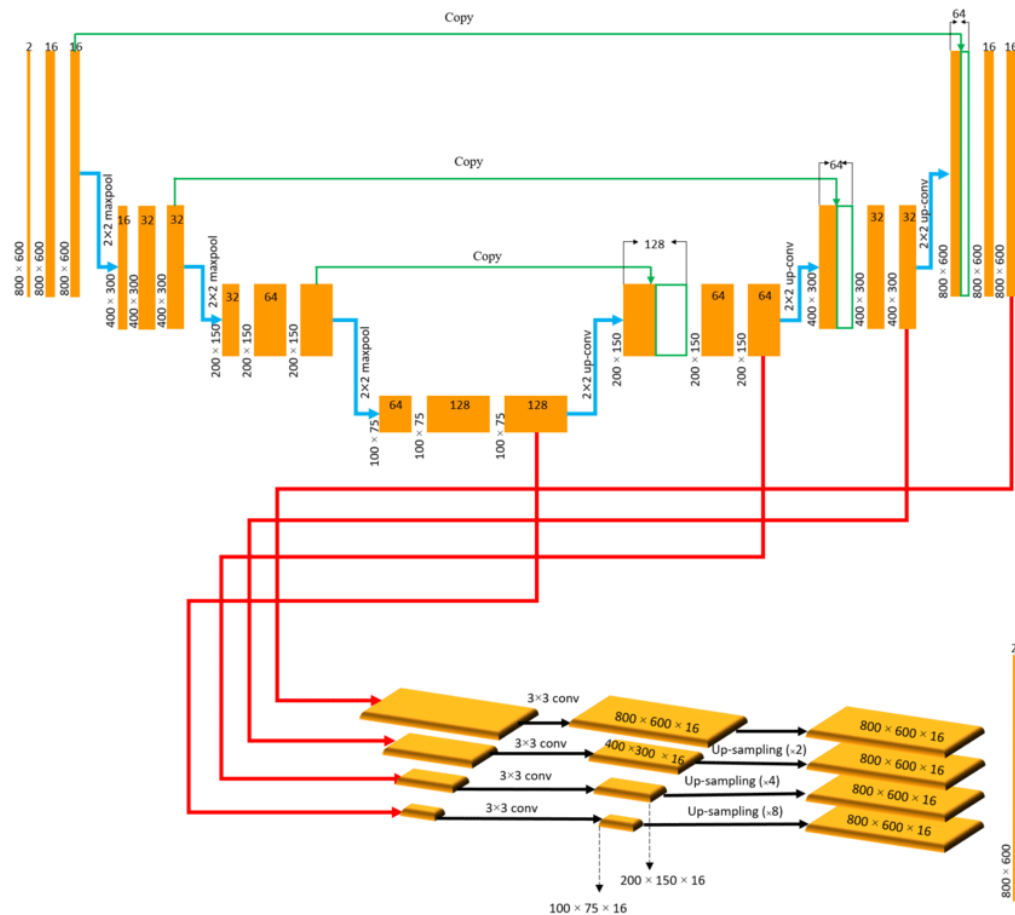


Рисунок 2.7 – Архітектура MFP-Unet

Нарешті, замість того, щоб вирішити проблему сегментації, маючи лише останню карту об'єктів, збережені карти об'єднуються та проходять через мережевий класифікатор. Таким чином архітектура MFP-Unet долає проблему роздільної здатності та точності в шляхах кодера-декодера, і

через, що дана архітектура є лише модифікацією U-Net, вона зберігає в собі усі її переваги.

Adversarial U-Net – це установка, при якій дві мережі конкурують одна з одною з метою покращення їх продуктивності. Генеративні змагальні мережі (Generative Adversarial Networks, GAN) – це новий тип змагальних процесів, що використовуються для отримання нових даних. Структура складається з двох мереж: дискримінатора та генератора. Механізм дискримінації D – це класифікатор, який навчений визначати, чи отримано дане вхідне зображення із набору даних, чи створене генератором G. На рисунку 2.8 зображена загальна архітектура GAN.

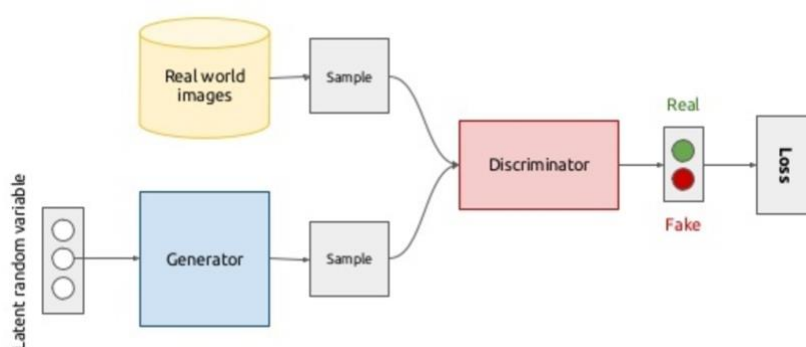


Рисунок 2.8 – Архітектура Generative Adversarial Networks

Дискримінатор D проходить стандартну підготовку під контролем CNN, і для кожного введення зображення він видає ймовірність зображення, яке створюється G з метою мінімізації рівня помилок при класифікації «підробок» як реальних зображень набору даних. Генератор G створює зображення, які періодично подаються на дискримінатор. Щоб допомогти генератору створювати переконливі зображення, його функція градієнта є функцією функції градієнта дискримінатора під час кроку, коли дискримінатор подається на фальшиве зображення. Це дозволяє генератору регулювати свою вагу у відповідь на вихід дискримінатора. Крім того, для створення варіацій на зображеннях, що генеруються генератором, до нього

передається випадковий шум. Мета генератора – перевершити дискримінатор, тобто максимізувати частоту помилок дискримінатора. Цей мінімаксий взаємозв'язок призводить до виникнення змагальної мережі, де дві мережі конкурують між собою.

Враховуючи достатній час, змагальна мережа повинна досягти оптимального стану, коли дискримінатор завжди виводить ймовірність $\frac{1}{2}$, незалежно від того, чи є зображення із набору даних чи генератора, що означає, що він більше не може відрізнити реальні зображення від синтетичних зображень, виробляється генератором. Потім отриманий генератор можна використовувати для штучного створення зображень певного предмета.

Цю структуру можна додатково розширити, щоб обмежити GAN у створенні обмеженої смуги синтетичних зображень, керуючи її мітками та вхідними зображеннями. Ця зміна відома як умовна GAN. Adversarial U-Net – це тип умовних GAN. Архітектура мережі представлена на рисунку 2.9.

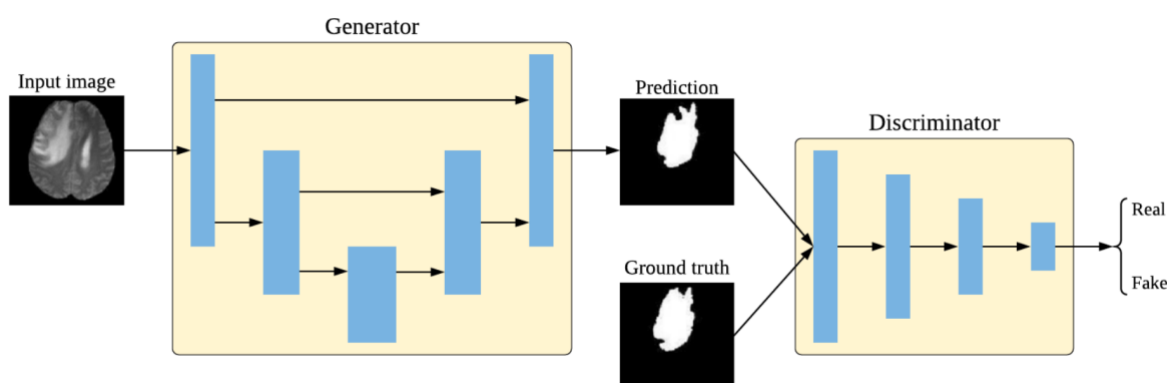


Рисунок 2.9 – Архітектура Adversarial U-Net

Мережа генераторів побудована на основі архітектури U-Net, тоді як дискримінатор залишається незмінним. Конструкція U-Net дозволяє генератору приймати зображення як вхідний сигнал замість випадкових

шумів. Ключова відмінність у змагальних U-мережах полягає в тому, що метою генератора є не створення нових зображень, а скоріше трансформовані зображення. Цей вихід генератора G оцінюється за дискримінатором D , який тренується на вручну перетворених зображеннях. В ідеалі, після належної підготовки, генератор зможе досягти тієї самої трансформаційної здатності, що і ручна трансформація людини. Потім отриманий генератор можна використовувати для застосування його функції перетворення на нових зображеннях, що було б значно швидше, ніж лікар, який вручну перетворив зображення. Adversarial U-Net знайшли широкий спектр застосувань, включаючи кількісне картографування чутливості мозку, виявлення пухлин головного мозку та раку молочної залози, сегментацію судин сітківки, сегментацію серцевих структур та реєстрацію зображень структур мозку.

2.5 CoordConv для покращення сегментації медичних зображень

Згорткові нейронні мережі (CNN), клас моделей глибокого навчання, вже давно стали потужними інструментами з видатною продуктивністю завдяки різноманітним програмам, таким як виявлення об'єктів та семантична сегментація. Незважаючи на їхній прорив, продуктивність CNN все ще схильна до деградації через відсутність просторових особливостей, що було б особливо корисно для сегментації зображень, де потрібно приймати рішення відносно даних піксель до класу чи ні. Нещодавні досягнення в цій галузі були зосереджені на інтеграції місцезнаходження об'єкту в моделі CNN, щоб подолати цю проблему.

Шар CoordConv, нещодавно представлений у [29], дозволяє фільтру згортки враховувати просторові координати пікселів. Метою CoordConv є вивчення відображення між координатами в декартовому просторі та координатами в просторі з пікселями інтересу. CoordConv показав свій

потенціал для локалізації об'єктів та викликав інтерес до сегментації зображень, де CoordConv також показав покращення результату [30].

Шар CoordConv – це просте розширення до стандартного згорткового шару. Як все говорилося раніше, згорткові шари використовуються в безлічі додатків, оскільки вони мають ряд переваг: мають порівняно мало параметрів навчання, їх швидко обчислювати на сучасних графічних процесорах і вони вивчають функцію, яка є інваріантною.

Шар CoordConv зберігає перші дві з цих властивостей – кілька параметрів та ефективні обчислення, але дозволяє мережі навчитися зберігати або відкидати третю – незмінність трансляції – як це потрібно для задачі, що вивчається. Може здатися, що позбавлення інваріантності перекладу перешкоджатиме здатності мереж засвоювати узагальнюючі функції. Однак розподіл невеликої кількості пропускнуої спроможності мережі для моделювання інваріантних аспектів проблеми, що не перекладаються, може забезпечити побудовання моделі, яка набагато більше піддається навчанню, які також узагальнюють набагато краще. Рівень CoordConv може бути реалізований як просте розширення стандартної згортки, при якому додаткові канали створюються за допомогою екземпляра та заповнюються (постійною, нетренованою) інформацією про координати, після чого вони об'єднуються по каналу до вхідного подання і застосовується стандартний згортковий рівень.

Конкретно, i -координатний канал являє собою матрицю $h \times w$ rank-1 із першим рядком, заповненим 0, другим рядком 1, третім 2-ма і т. д. Канал координат j подібний, але зі стовпцями, заповненими значення замість рядків.

На рисунку 2.10 зображено операцію згортки з додаванням двох додаткових координат, i та j .

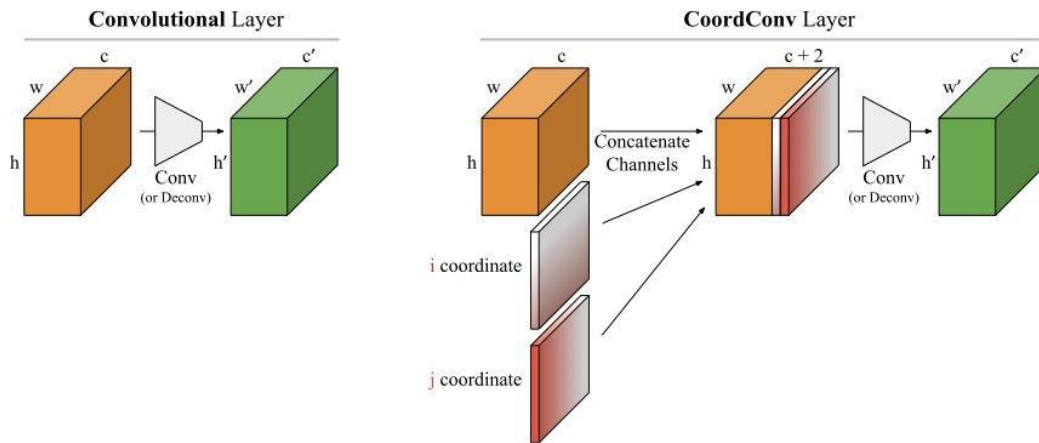


Рисунок 2.10 – Операція згортки разом з додатковими шарами координат

Також до даних шарів застосовується остаточне лінійне масштабування значень координат i та j , щоб дані значення потрапили в діапазон $[-1, 1]$. Для згортки у двох вимірах достатньо двох координат (i, j) , щоб повністю вказати вхідний піксель, але при бажанні можуть бути додані додаткові канали. Для інших шарів у мережі кількість рядків і стовпців буде відповідати висоті та ширині карти об'єктів на цьому конкретному шарі. Отже, CoordConv можна застосовувати до будь-якого шару, а не лише до вхідного шару.

3 АЛГОРИТМ ДЛЯ АВТОМАТИЧНОГО ПІДРАХУНКУ ФРАКЦІЇ ВИКИДУ

3.1 Загальний опис алгоритму для підрахунку фракції викиду

Ехокардіограма – це тест, який використовує ультразвук, щоб показати, як працює серцевий м'яз і клапани. Звукові хвилі роблять рухомі зображення серця, щоб медичний працівник міг добре розглянути його розміри та форму. Під час дослідження серце розглядаються з різних планів, на яких оцінюються різноманітні характеристики та величини. В алгоритмі автоматичного розрахунку фракції викиду, який пропонується в даній роботі, фокусуватися будемо на двокамерному- та чотирьох-камерному виду зображення серця.

Як вже було розглянуто раніше, для того, щоб підрахувати об'єм лівого шлуночка та фракцію викиду, необхідно на ехокардіографічній відео послідовності на повному серцевому циклі (зазвичай це 30-40 кадрів) знайти фрейми з найбільшою на найменшою площиною лівого шлуночка. Площа ЛШ на кожному фреймі знаходитиметься за допомогою глибинного навчання, а саме моделі сегментації. U-Net архітектура, яка показала багатообіцяючі результати у сегментації медичних зображень, побудована та модифікована на повністю згортковій нейронній мережі. Для її навчання не потрібні великі масиви даних, що є великою перевагою. Вона пропонує дві основні модифікації: симетрія мережі та пропуск з'єднань між шляхами кодера-декодера. Ці аспекти об'єднують карти функцій кодування з відповідними картами функцій декодування. Рисунок 3.1 демонструє алгоритм автоматичного розрахунку, який використовується в даній роботі.

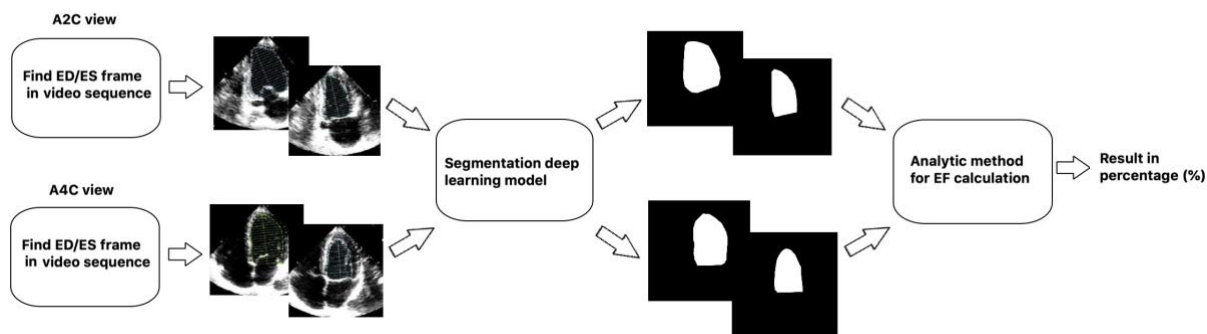


Рисунок 3.1 – Алгоритм автоматичного розрахунку фракції викиду ЛШ

Для ефективної діагностики існує надзвичайно важлива необхідність сегментації лівого шлуночка (ЛШ) при розрахунках клінічних показників, таких як кінцеві діастолічні та кінцеві систолічні об'єми, фракція викиду, маса лівого шлуночка тощо. Звичайно фахівець-кардіолог вручну визначає межу ендокарда ЛШ на кінцевій діастолі та систолі серцевого циклу. Ця процедура є не лише виснажливим та трудомістким завданням, але також страждає від складної відтворюваності та мінливості внутрішньо та між спостерігачами. Для вирішення цієї проблеми потрібен автоматичний метод сегментації, щоб прискорити процес, одночасно зменшуючи мінливість фахівців. Однак проблема сегментації ЛШ стикається з кількома проблемами. По-перше, деякі характеристики ультразвукових зображень, наприклад, низьке відношення сигналу до шуму, слабе відлуння тощо, генерують зображення з незрозумілими межами, що в деяких випадках робить розмежування ЛШ надзвичайно складним. По-друге, визначення області, що представляє інтерес (тобто області ЛШ), є складним через наявність більше ніж однієї анатомічної структури на ехокардіографічних зображеннях. Нарешті, немає простого зв'язку між фізичними властивостями тканини та інтенсивністю пікселів на зображеннях. Для вирішення проблем описаних вище в даній роботі буде використовуватися модель сегментації MFP-UNet, яка являє собою модифікацію класичної архітектури U-Net у поєднанні з функціональною пірамідною мережею та

розроблена безпосередньо для сегментації лівого шлуночка серця. Немає потреби навчати дві моделі сегментації для двокамерного- та чотирьох-камерного зображення окремо, тому що основним об'єктом як на одному так і на іншому зображенні є лівий шлуночок, для моделі сегментації буде спільного для двох видів. На вхід моделі будуть подаватися зображення, попередньо оброблені, чорно-білого кольору одноканальні, бо колір зображення не впливає на якісний результат даної моделі, але використання одноканального зображення зменшить кількість параметрів. На виході моделі будуть очікуватися чорно-білі маски лівого шлуночка.

Після того, як будуть отримані маски, необхідно застосувати аналітичні методи обробки зображень. Такі методи існують у Python бібліотеці OpenCV. З її допомогою буде проведено морфологічний розбір зображення маски для виділення контуру ЛШ, а також знайдені основні точки для підрахунку довжини лівого шлуночку по довгій осі – від основи серця до середини лінії мітрального клапана. Для застосування формули Біпланного методу підрахунку ФВ залишилося знайти об'єм ЛШ, шляхом його розділення на n -ну кількість дисків, як для 2-камерного виду так и для 4-х камерного. Дана маніпуляція також буде проводитися шляхом застосування функціоналу OpenCV. Використовуючи формулу Біпланного підрахунку ФВ, буде отримане значення, яке виявляє, який відсоток насиченої киснем крові буде викачано ЛШ в організм.

Даний алгоритм буде працювати в режимі реального часу з відображенням результуючої маски, яку повертає модель сегментації, а також з виводом підрахованого значення ФВ.

3.2 Метод сегментації лівого шлуночка

Серед методів, заснованих на глибокому навчанні, згортована нейронна мережа здатна в значній мірі вирішувати проблеми сегментації

зображень. Це пояснюється її характеристиками незмінності зсуву та потужною здатністю до вилучення особливостей зображення.

Незважаючи на винахідливість, U-Net безпосередньо ігнорує вплив карт функцій у різних масштабах. Завдяки цьому, коли карти функцій збираються до вибірки в шляху декодера, лише останній шар підключається до вихідної карти сегментації, як і будь-яка інша мережа прямого пересилання. Крім того, блоки в різних масштабах шляху декодера не можуть спільно використовувати функції, що може перешкоджати потоку інформації та спричиняти зайві параметри. Хоча можливо забезпечити введення останнього шару, використовуючи всі кроки шляху декодера.

Для вирішення цієї проблеми пропонується використовувати функціональні пірамідні мережі вбудовані в архітектуру U-Net. Піраміди об'єктів є основним компонентом в системах розпізнавання для виявлення об'єктів у різному масштабі. Ключовою ідеєю є вилучення карт функцій з усіх рівнів шляху декодера окремо та використання їх для остаточної класифікації пікселів.

Функціональні пірамідні мережі (Feature Pyramid Networks, FPN) – це ефективна архітектура для виявлення об'єктів у різних масштабах [28]. Існує три підходи до розпізнавання об'єктів у різних масштабах:

- піраміди одного і того ж зображення в різних масштабах;
- піраміди особливостей: особливості зображень виділяються в них ієрархічно. Це ключовий момент, який змушує дослідників створювати піраміду об'єктів, а не піраміду зображень;
- піраміда об'єктів з однаковою смисловою силою в усіх масштабах: цю мету легко досягти, створивши додатковий шлях, який додає функції високого рівня від верхніх шарів до нижніх шарів, що мають вищу роздільну здатність, але семантично слабкі особливості.

Є два шляхи, якими поширюється вхідне зображення. Перший – це «шлях знизу вгору», а другий – «шлях зверху вниз». Для визначення піраміди функціональних карт просторові розміри слід регулярно

зменшувати знизу вгору, щоб створити різні рівні піраміди. Це включає кілька модулів згортки з іменами {C1, C2, C3, C4 та C5}. Коли вхідне зображення проходить через мережу, просторова розмірність зменшується в 0,5 разів, що еквівалентно подвоєнню кроку. Зокрема, є кроки {2, 4, 8, 16 та 32} для відповідних модулів {C1, ..., C5}. Архітектура функціональної пірамідної мережі представлена на рисунку 3.2.

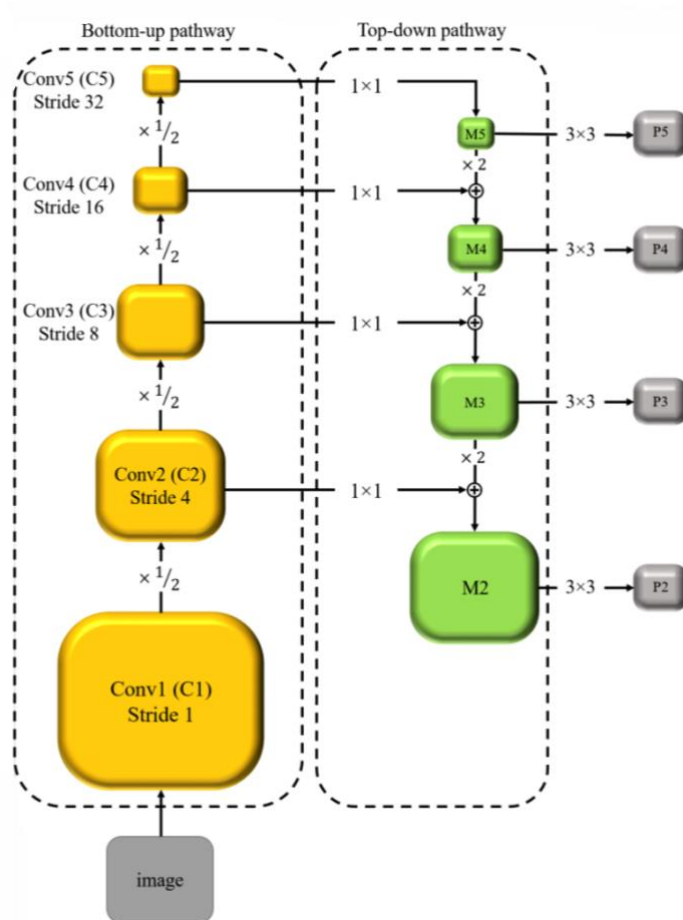


Рисунок 3.2 – Архітектура функціональної пірамідної мережі

Шлях зверху вниз забезпечує можливість перенесення семантично більш сильних функцій у вищу роздільну здатність. Потім ці функції доповнюються відповідними функціями шляху знизу вгору через пропускання з'єднань. Існує кілька конкретних процесів, які дозволяють зробити карти об'єктів знизу вгору придатними для об'єднання з картами

об'єктів зверху вниз. Для зменшення глибини каналу C5 застосовується згортковий фільтр 1×1 . Потім застосовується згортка 3×3 для створення P5, яка є першою картою об'єктів, що використовується для подальшого прогнозування. На кожному рівні, щойно створена карта об'єктів вибирається та підсумовується в результаті застосування 1×1 згортки до відповідної карти об'єктів у шляху знизу вгору. Ця процедура повторюється доти, доки остання карта роздільної здатності не буде створена з найвищою якістю.

Така архітектура U-Net разом з функціональною пірамідною мережею являє собою багатофункціональною пірамідою U-Net (Multi-feature pyramid U-Net, MFP-Unet). Як згадувалося раніше, шлях стиснення архітектури U-Net створює додаткову кількість карт функцій, отже, роздільна здатність зменшується. Це є причиною застосування процесу об'єднання на кожному рівні шляху розширення. Поєднання інформації глибоких функцій та низькорівневих функцій з високою роздільною здатністю дозволяють алгоритму зберегти як точність, так і роздільну здатність. Незважаючи на таку архітектуру, класична архітектура U-Net використовує лише останні функції шару в шляху розширення для етапу сегментації (останній шар моделі). Архітектура MFP-Unet змінює U-Net таким чином, щоб карти об'єктів на всіх рівнях шляху розширення були включені окремо.

Враховуючи архітектуру U-Net, MFP-Unet додає додаткові шари згортки для вилучення карт функцій з усіх рівнів шляху розширення, щоб бути включеними в процес сегментації останнього шару. Це включення обіцяє функціональна пірамідальна мережа. Таким чином, до кожного рівня декодера додається простий окремий шлях, який є згортковим шаром 3×3 з 16 картами об'єктів і шаром для вибірки, щоб зробити розмір карт об'єктів такими ж, як і останній шар. Нарешті, усі отримані карти об'єднань об'єднуються, щоб скласти остаточну карту об'єктів із 64 каналами, яка відображається у фільтрі згортки 1×1 для процесу класифікації.

Останній шар на кожному рівні шляху декодера називається up_i ($i = 1, 2, 3, 4$). У класичній мережі U-Net повністю зв'язаним шаром є шар згортки з ядром 1×1 , його вихід буде розрахований за формулами 3.1 та 3.2 з $m = 1$ та $l = L$ (останній шар).

$$x_{ij}^l = \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} w_{ab} y_{(i+a)(j+b)}^{l-1}, \quad (3.1)$$

де $m \times m$ – розмір фільтра становить,

w_{ab} – ваги,

y^{l-1} – вихід попереднього шару.

$$y_{ij}^l = ReLU(x_{ij}^l), \quad (3.2)$$

де $ReLU$ – функція активації.

Для архітектури MFP-Unet формули для підрахунку виходу будуть виглядати наступним чином (формула 3.3, 3.4):

$$y_{ij}^{L-1} = concat \left[\frac{conv(up^4, 16), conv(up^3, 16)}{conv(up^2, 16), conv(up^1, 16)} \right], \quad (3.3)$$

де $concat[...]$ – це операція конкатенації,

$conv(up^i, 16)$ – одиниця згортання (шари згортки та ReLU), що виконується на up^i з вихідними 16 картами функцій.

$$y_{ij}^L = ReLU(w_{00} y_{ij}^{L-1}). \quad (3.4)$$

Шлях кодера залишається незмінним, але використовуються розширені фільтри згортки, щоб бути більш надійними у розширеній U-Net. У шляху декодера шари архітектури U-Net залишались незмінними, але додавались додаткові шари для підвищення точності та роздільної здатності

кожного рівня на окремих картах функцій. Нарешті, замість того, щоб вирішити проблему сегментації, маючи лише останню карту об'єктів, збережені карти об'єднуються в одну кінцеву та передаються мережевому класифікатору.

Попередня обробка зображень є важливим кроком у процесі прийняття рішень та комп'ютерного зору. Ця процедура може включати операції з придушення шуму, посилення контрасту тощо. Усі ці операції мають різноманітні методики, які використовуються відповідно до типу зображення та мети попередньої обробки. В даній роботі використовуються сірі медичні зображеннями ехокардіографії у 2-х та 4-камерному виді, і метою є окреслення меж ЛШ. Межі ЛШ – це бічна стінка та перегородка, які виглядають яскравішими в техніці візуалізації ультразвуком. Тому в даній роботі використовується метод попередньої обробки для посилення контрасту, щоб зробити гарне поле між межами та іншими яскравими пікселями на зображенні.

Одним із способів підвищення контрастності зображень є порогове значення. Це ефективний і простий інструмент для багатьох заявників, де рівні сірого в об'єктах на зображенні відрізняються від тих, що належать до фонових пікселів. Порогове обмеження виконується у двох основних процедурах: глобальній та локальній. У глобальному пороговому значенні існує одне порогове значення для всього зображення, тоді як при локальному пороговому рівні всі пікселі порівнюються з пороговим значенням, яке обчислюється на основі деяких статистичних даних сусідніх пікселів у вікні [28].

Однією з популярних методів локального порогового значення є техніка Ніблака. У цьому методі значення порогу визначається із застосуванням середнього та стандартного відхилення пікселів у вікні. Тут ми використали цю техніку для глобального порогового значення, щоб зробити межі яскравішими на ехокардіографічних зображеннях [31]. Значення середнього та стандартного відхилення обчислюються з

урахуванням цілого зображення. Нове зображення, яке отримується при проходженні через пороговий фільтр, може використовуватися як другий канал на вхід до нейронної мережі.

Для покращення результатів сегментації використовувався метод додавання двох додаткових шарів координат до вхідного зображення *CoordConv*, який являє собою простим розширенням класичного згорткового шару.

Медична сегментація зображення також вимагає загальної інформації про зображення, щоб побачити, як розташовані органи відносно один одного. U-Net моделі складно охоплювати все зображення в шарах стискання та розширення, та через це мережа не здатна виявити, що на ультразвуковому зображенні є лише один лівий шлуночок. Підходом для вирішення цієї слабкості було додавання ще двох шарів зменшення вибірки. Однак даний метод збільшує параметри мережі. Замість цього в даній роботі будуть використані розширені ядра згортки [32] для збільшення сприйнятливих полів мережі.

Для навчання моделі сегментації вирішено використовувати функцію витрат *BCE-Dice Loss*, формула якої визначалася у попередньому розділі, а в ролі функцій для оцінки точності результатів моделі виступатимуть *Dice*, *Intersection over Union* та *Accuracy*, описані формулами 2.5, 2.7, 2.8 відповідно.

3.3 Аналітичний метод підрахунку довжини, об'єму та фракції викиду

Після того, як модель сегментації закінчить навчання, вона буде використовуватися для отримання масок ЛШ на фреймах кінцевої діастоли та систоли на двокамерному та чотирьох камерному зображеннях серця.

Слід зазначити, що у деяких випадках ЛШ на двокамерному та чотирьох камерному зображеннях може мати різний зум. Тому для фреймів діастоли обирається зображення серед двокамерного та чотирьох

камерного, де довжина ЛШ найбільша. Якщо на двокамерному, то зображення маски для чотирьох камерного виду збільшуються пропорційно, щоб довжина стала однаковою, якщо довжина більша на двокамерного, то збільшуються маски чотирьох камерного виду.

Дана процедура необхідна для точних розрахунків. У іншому випадку можлива помилка в підрахунках. На рисунку 3.3 продемонстрований приклад вирівнювання зображень по довжині:

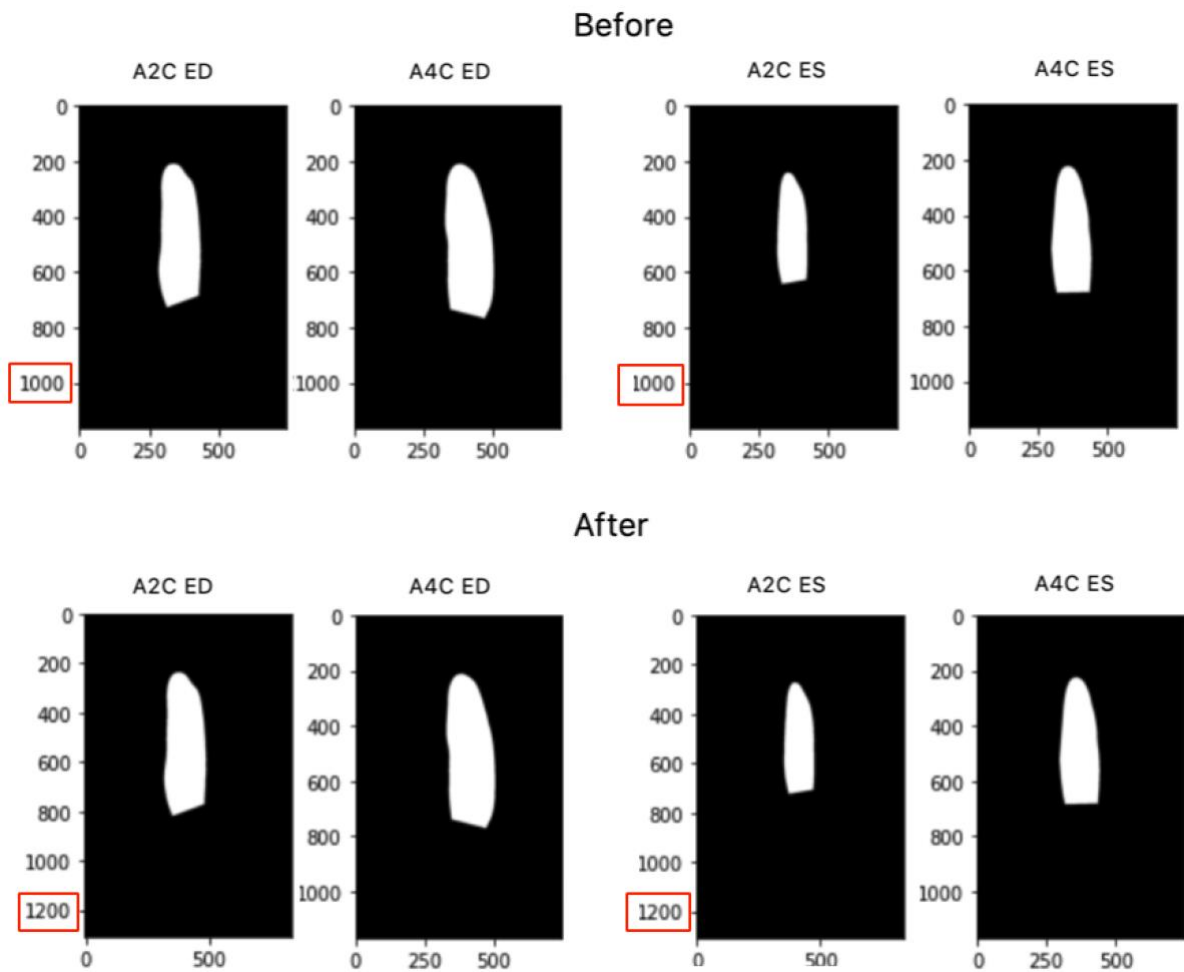


Рисунок 3.3 – Вирівнювання зображень двокамерного та чотирьох камерного виду по довжині ЛШ

На рисунку 3.3 видно, особливо на двокамерних зображеннях, що довжини не співпадають, і на дослідження двокамерного виду не було приближено, як на чотирьох камерному. Тому, зображення на двокамерного

виду були збільшені по висоті на 200 пікселів, а також по ширині на відповідну кількість пікселів, щоб залишити пропорції.

Коли зображення двокамерного та чотирьох камерного виду будуть вирівняні по довжині ЛШ, застосовуються засоби бібліотеки Python OpenCV, яка має багато корисних функцій та методів для обробки зображень. А саме на отриманих масках необхідно знайти три точки, які являються верхівкою серця (apex of heart), точки початку та кінця мітрального клапану, за допомогою який буде знаходитися точка середини мітрального клапану.

Спочатку необхідно знайти контур ЛШ з отриманої маски використовуючи метод для знаходження країв Canny Edge Detection. Результатом даного методу буде контур ЛШ. В даній роботі буде застосовуватися геометричний метод, де лівий шлуночок буде описуватися колом, таким чином будуть знайдені принаймні 2 точки, які будуть лежати на даному колі. На рисунку 3.4 показано приклад даного методу:

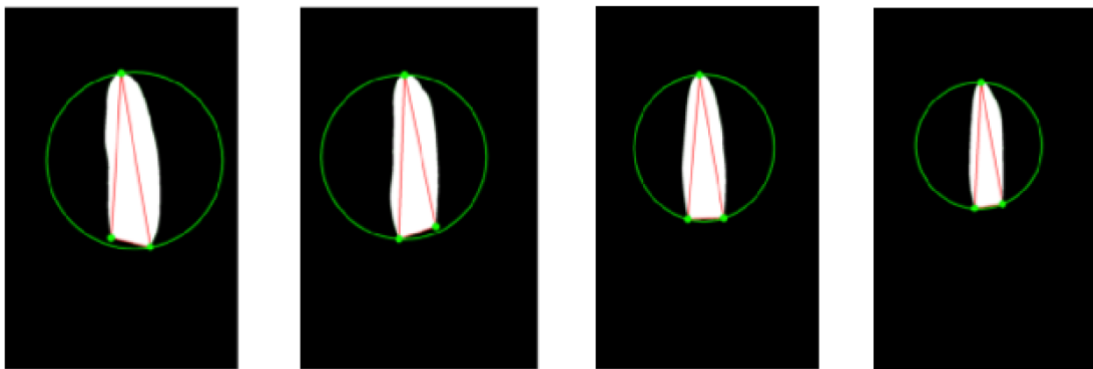


Рисунок 3.4 – Знаходження точок методом описаного кола

Іноді одна з точок основи мітрального клапану може не лежати на описаному колі, як на перших двох зображеннях на рисунку 3.4. У даному випадку вона знаходиться програмно в циклі, шляхом пересування на один

піксель вгору з протилежної сторони від знайденої точки мітрального клапану, доки не буде знайдений білий піксель.

Коли усі три точки знайдені потрібно підрахувати координати точки середини основи мітрального клапану. Дана операція виконується також за допомогою геометричної операції.

Для підрахунку довжини лівого шлуночка знадобляться точки верхівки серця та середини мітрального клапану. Отримана довжина буде вимірюватися в пікселях, і для того щоб трансформувати її в сантиметри необхідно мати інформації про те, скільки на один сантиметр приходить пікселів. Таким чином це дозволить отримати величини об'єму ЛШ. Якщо дана інформація у наборі даних відсутня, об'єм лівого шлуночку не можливо підрахувати у величині см^3 , бо вона буде відповідати пікселям. Але для підрахунків фракції викиду це не має значення, бо дана величина вимірюється у відсотках (%). Фракція викиду буде розрахована за формулою 1.2.

На рисунку 3.5 зеленим кольором виділені точки верхівки серця та середини основи мітрального клапану та результат застосування Біпланного методу дисків.

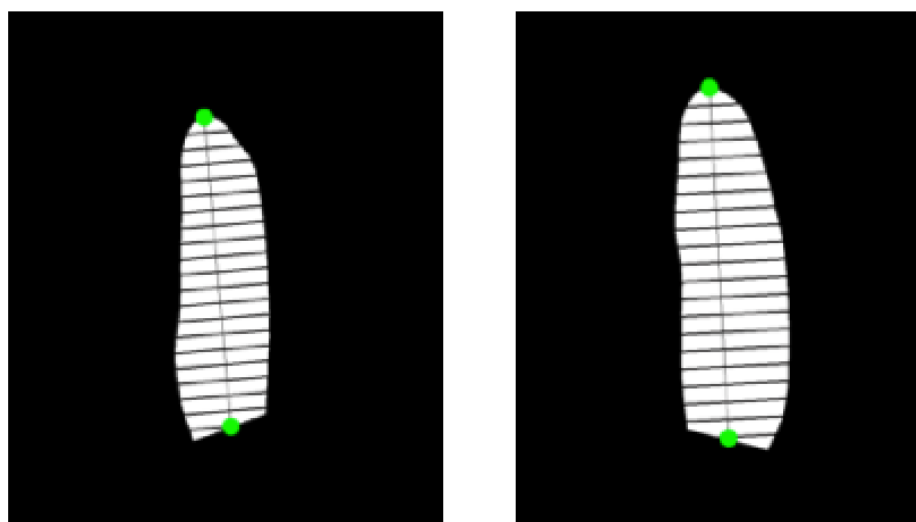


Рисунок 3.5 – Застосування Біпланного методу дисків до масок ЛШ

Щоб провести розрахунки об'єма ЛШ за формулою Біпланного методу дисків Сіпмона (формула 1.4), потрібно розбити маску ЛШ на 20 дисків однакової висоти, і довжини діаметрів дисків та їх висоту використати у формулі. Головним правилом є те, що лінії дисків повинні бути перпендикулярні лінії довжини ЛШ, тобто мати кут 90° . За допомогою методів OpenCV можливо знайти кут, на який потрібно відхилити пряму лінію, щоб вона була під кутом 90° до іншої лінії.

Формула Біпланного методу застосовується для об'єму ЛШ у кінцевій діастолічній та систолічній фазі та отримується фракція викиду крові ЛШ.

Таким чином отримаємо точний аналітичний метод, який дозволяє швидко та якісно підрахувати фракцію викиду на отриманих масках ЛШ із моделі сегментації без втручання людини.

4 ПРОГРАМНІ ЗАСОБИ ДЛЯ РЕАЛІЗАЦІЇ АЛГОРИТМУ АВТОМАТИЧНОГО РОЗРАХУНКУ ФРАКЦІЇ ВИКИДУ

Для реалізації алгоритму автоматичного розрахунку фракції викиду було обрано мову програмування Python версії 3.8. Python – високорівнева мова програмування загального призначення, це потужний інструмент для створення програм найрізноманітнішого призначення. З його допомогою можна вирішувати завдання різних типів. Python має велику кількість стандартних бібліотек, які мають багато корисних функцій. Python підтримує структурне, об'єктно-орієнтоване, функціональне, імперативне і аспектно-орієнтоване програмування. Python – інтерпретована мова програмування. це дозволяє значно спростити налагодження програм

До переваг, які має мова програмування Python, можна віднести простоту, яка досягається за рахунок того, початковий код виконується методом інтерпретації, можливість роботи програми на Python з будь-якою операційною системою, велика кількість стандартних та додаткових модулів, які можна легко інсталювати.

Також Python має багато бібліотек для роботи з відео та зображеннями. Серед таких бібліотек можна виділити OpenCV, scikit-image, PIL (Python Imaging Library). Дані бібліотеки використовувалися для обробки зображень та аналітичного підрахунку фракції викиду.

OpenCV – бібліотека комп'ютерного зору і машинного навчання з відкритим вихідним кодом. У неї входять понад 2500 алгоритмів, в яких є як класичні, так і сучасні алгоритми для комп'ютерного зору і машинного навчання. Ця бібліотека має інтерфейси на різних мовах,. OpenCV використовується для роботи із зображеннями і має наступні функції:

- розпізнавання об'єктів та тексту;
- усунення спотворень;
- виявлення подібності та форми об'єктів;
- стеження за переміщенням об'єкта;

– розпізнавання рухів, жестів і багато іншого.

Для побудови глибоких нейронних мереж використовувався Python Framework PyTorch. PyTorch – це бібліотека машинного навчання з відкритим кодом, заснована на бібліотеці Torch, яка використовується для таких програм, як комп'ютерний зір та обробка природних мов, в основному розроблена дослідницькою лабораторією AI Facebook (FAIR). Це безкоштовне програмне забезпечення з відкритим кодом, випущене під модифікованою ліцензією BSD.

PyTorch надає дві функції високого рівня:

- тензорні обчислення (як NumPy) з сильним прискоренням за допомогою графічних процесорів (GPU);
- глибокі нейронні мережі, побудовані на основі автоматичної системи диференціації на основі типу.

Для розробки програмного коду та тестування використовувалося середовище Jupyter Notebook. Jupyter Notebook – це потужний інструмент для розробки та презентацій проектів Data Science в інтерактивному режимі. Він об'єднує код і виводить усі дані одного документа, що містять текст, математичні порівняння та візуалізацію. Такий кроковий підхід забезпечує швидкий, послідовний процес розробки, завдяки виводу для кожного блоку, який виконується відразу.

Для тренування моделі сегментації був використаний набір даних SAMUS, що містить двовимірні апікальні чотирикамерні та двокамерні послідовності огляду, отримані від 500 пацієнтів, та який доступний для завантаження у вільному доступі [33].

Загальний набір даних SAMUS складається з клінічних обстежень 500 пацієнтів, придбаних в університетській лікарні Сент-Етьєн (Франція) та включених у дослідження в рамках норм, встановлених місцевим етичним комітетом лікарні. Придбання було оптимізовано для вимірювання фракції викиду лівого шлуночка. Повний набір даних був отриманий ультразвуковими сканерами GE Vivid E95 (GE Vingmed Ultrasound, Horten

Norway) із зондом GE M5S (GE Healthcare, США). Не було запроваджено жодного додаткового протоколу, ніж той, що застосовується у клінічній рутині. Для кожного пацієнта 2D апікальні чотирикамерні та двокамерні оглядові послідовності експортувались із програмного забезпечення для аналізу EchoPAC (УЗД GE Vingmed, Horten, Норвегія). Ці стандартні серцеві види були обрані для цього дослідження, щоб дозволити оцінити значення фракції викиду лівого шлуночка на основі Біпланного методу дисків Сімпсона.

Повний набір даних складається з набору даних для навчання та містить 450 пацієнтів та тренувальний набір даних – 50 пацієнтів.

Набір навчальних даних складається наступним чином:

- 366 пацієнтів з хорошою / середньою якістю зображення;
- 84 пацієнти з поганою якістю зображення;
- 222 пацієнти з фракцією викиду нижче 45%;
- 87 пацієнтів з фракцією викиду вище 55%.

Набір даних для тестування складається наступним чином:

- 40 пацієнтів з хорошою / середньою якістю зображення;
- 10 пацієнтів із поганою якістю зображення;
- 24 пацієнти з фракцією викиду нижче 45%;
- 10 пацієнтів з фракцією викиду вище 55%.

Через те, що початковий набір даних не має окремо створеного набору для валідації моделі після її навчання, набір для тренування був розбитий на тренувальний, 400 пацієнтів, та валідаційний 50 пацієнтів, серед яких:

- 38 пацієнтів з хорошою / середньою якістю зображення;
- 12 пацієнти з поганою якістю зображення;
- 30 пацієнти з фракцією викиду нижче 45%;
- 20 пацієнтів з фракцією викиду вище 55%.

Принаймні один повний серцевий цикл був отриманий для кожного пацієнта на кожному огляді, що дозволяє вручну анотувати серцеві структури при кінцевій діастолі та систолі, та додані в набір даних. Окрім

відео послідовності серцевого циклу додаються окремо зображення кінцевої діастолі та систолі, а також зображення маски для кожного виду. Маски складаються із трьох слоїв, де на кожному слої сегментована певна частина лівого шлуночка: ендокард лівого шлуночка (внутрішній шар ЛШ), епікарда (зовнішній шар, що становить стінку ЛШ) та межі лівого передсердя. В даній роботі були використані лише маски ендокарда ЛШ для навчання моделі. Вхідні зображення у наборі даних надаються у форматі файлу raw / mhd. Для роботи із зображеннями даного формату використовувалась бібліотека SimpleITK мови Python. SimpleITK – це інструментарій для аналізу зображень з великою кількістю компонентів, що підтримують загальні операції фільтрації, сегментації та реєстрації зображень. Він побудований на основі інструментарію сегментації та реєстрації Insight ІТК з метою забезпечити спрощений інтерфейс для ІТК.

Також для кожного пацієнта у наборі даних надається інформація про кількість фреймів у відео послідовності та номери фреймів, що відповідають кінцевій діастолі та систолі, вік та стать, якість записаної відео послідовності та розрахунки об'єму ЛШ і фракція викиду, розрахована Біпланним методом Сімпсона.

Для більш швидкого відкриття файлів, які знаходяться у форматах raw / mhd, файли були збережені у форматі NumPy при першому їх відкритті і надалі використовувалися для тренування моделі.

5 ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ

В даній кваліфікаційній роботі був розроблений алгоритм автоматичного підрахунку фракції викиду крові лівого шлуночка, реалізовано модель сегментації MFP-UNet (додаток А), яку було створено саме для даної задачі, та аналітичний алгоритм підрахунку по отриманим маскам з даної моделі, а також проведено ряд експериментів залежності відповідний параметрів до якості сегментації та підрахунків.

Даний алгоритм повинен працювати з відеопослідовністю, роблячи сегментацію ЛШ на кожному фреймі, щоб обрати фрейми, де ЛШ мінімальний та максимальний за площею, як правило на одному кардіоциклі (30-40 фреймів). Вимоги до моделі в даній задачі формуються таким чином, щоб результат виходу моделі отримувався швидко. На це впливає кількість параметрів моделі, глибина моделі, розмір вхідного зображення. На додаток до моделі MFP-UNet було реалізовано класичну архітектуру U-Net і її модифікацію UNet++ та проведено порівняльну характеристику по кількості параметрів даних моделей. В таблиці 5.1 представлені дані характеристики.

Таблиця 5.1 – Порівняння U-Net архітектур за кількістю параметрів

Назва архітектури	Кількість параметрів
MFP-UNet	813,505
Класична U-Net	839,489
UNet++	3,293,617

З таблиці видно, що найбільшу кількість параметрів має модель UNet++, тому для даної задачі краще обрати архітектури MFP-UNet та класичну U-Net, так як вони мають майже однакову кількість параметрів і це буде впливати на отримання результату.

Для навчання моделей використовувалися такі параметри: $learning\ rate = 0.001$, $weight\ decay = 0.00005$, $batch\ size = 32$. Adam optimizer був обраний в якості оптимайзера для навчання моделей. Кількість епох для тренування моделі дорівнює 50. Розмір вхідного зображення дорівнює 160×160 та має один канал. Перед тренуванням зображення було нормовано у діапазоні $[-1, 1]$. Також була додана аугментація даних, щоб збільшити кількість варіацій масок, щоб модель мала змогу генералізувати дані та не досягти перетренування; зсув на випадкову кількість пікселів, повороти, зміна яскравості зображення додаються з вірогідністю в 50% кожен під час тренування моделі. Слід зазначити, що ЛШ при цьому завжди залишається у полі зору.

На рисунку 5.1 проілюстровано результати сегментації двох моделей.

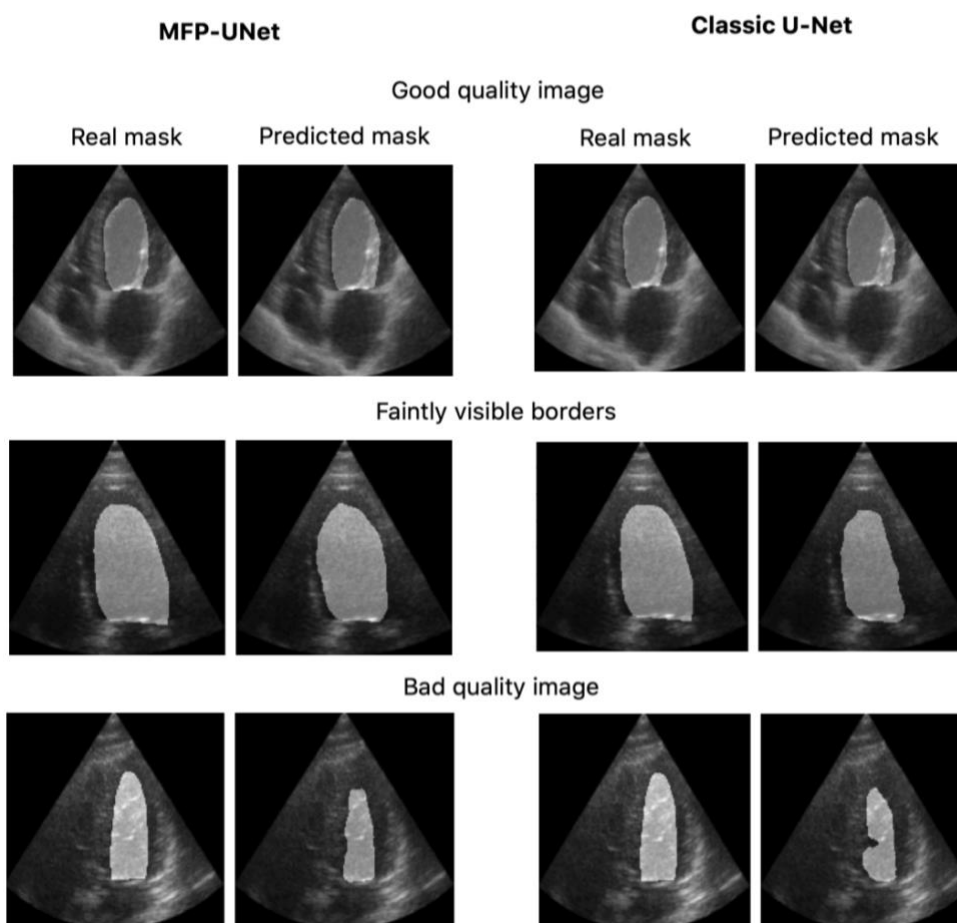


Рисунок 5.1 – Результати сегментації моделей MFP-UNet та U-Net

Як видно с рисунку 5.1 на зображенні хорошої якості дві моделі показали гарний результат, але у разі поганої якості зображення або розмитих країв ЛШ MFP-UNet має кращі результати.

Також в таблиці 5.2 представлені усереднені результати оцінок, отриманих на валідаційному наборі даних.

Таблиця 5.2 – Порівняння результатів U-Net архітектур за оцінками IoU, Dice та Accuracy

Назва архітектури	IoU	Dice	Accuracy
MFP-UNet	0.8653	0.9205	0.9867
Класична U-Net	0.8718	0.8992	0.9715

З рисунка 5.1 та таблиці 5.2 можна зробити висновки, що модифікації, які були здійснені до U-Net архітектури, щоб отримати MFP-UNet, покращили якість отриманих масок, хоча не збільшили кількість параметрів моделі.

На рисунку 5.2 представлені результати наступного експерименту.

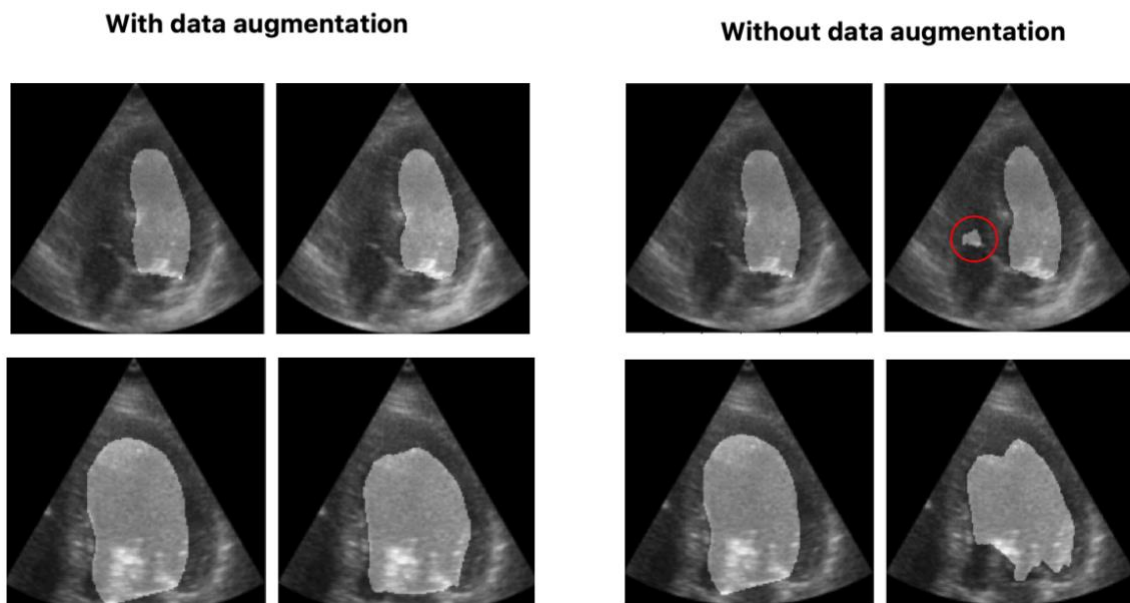


Рисунок 5.2 – Порівняння результатів масок без та з аугментацією даних

Після отриманих результатів для моделі MFP-UNet було проведено наступні декілька досліджень, а саме вплив розміру *batch size* на точність моделі та наявності аугментації даних. Збільшення *batch size* у два рази не вплинуло на якість моделі, але зменшило кількість часу на навчання. На відміну від *batch size*, відсутність аугментації даних вплинула на результат моделі: дещо знизилась оцінка по Dice – 0.9161 та IoU – 0.8544, з'явилися додаткові артефакти та погіршилась якість результуючої маски.

У першому експерименті модель MFP-UNet навчалась без додавання додаткових шарів координат для CoordConv шару. В наступному експерименті були додані додаткові шари та отримані результати: оцінка по Dice збільшилась з 0.9205 до 0.9289, а IoU – з 0.8653 до 0.8844.

Для алгоритму автоматичного підрахунку фракції викиду було вирішено використовувати MFP-UNet з CoordConv шаром на вході та з навчальними параметрами, описаними вище. Результати, які були отримані з використанням моделі сегментації та аналітичного алгоритму представлені на рисунках 5.3, 5.4.

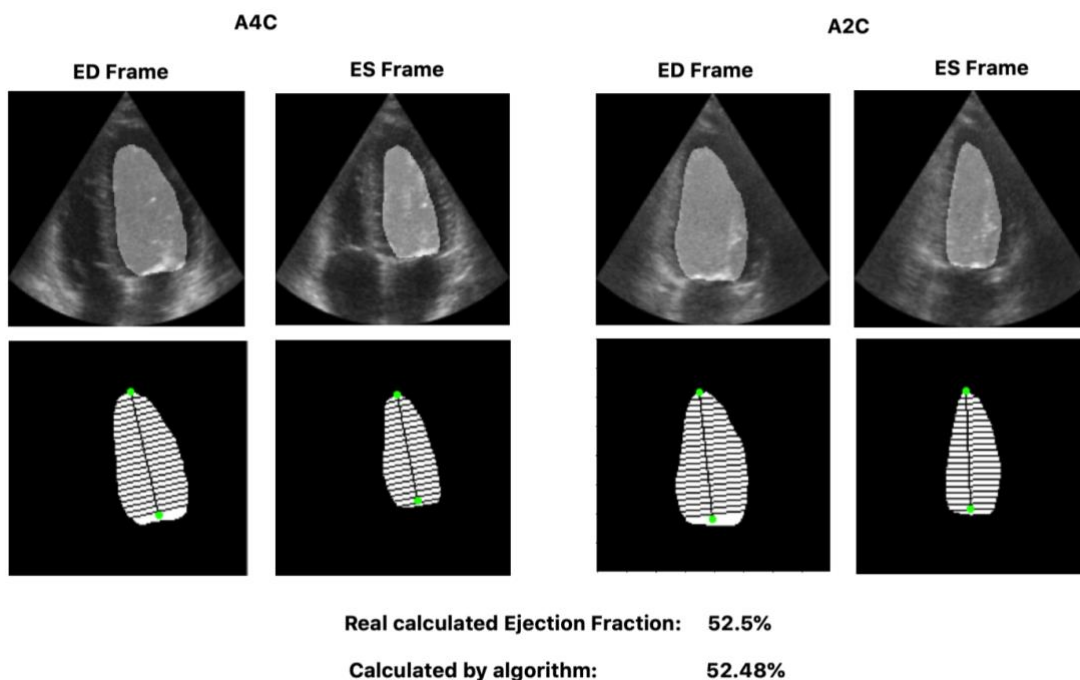


Рисунок 5.3 – Результат 1 роботи алгоритму

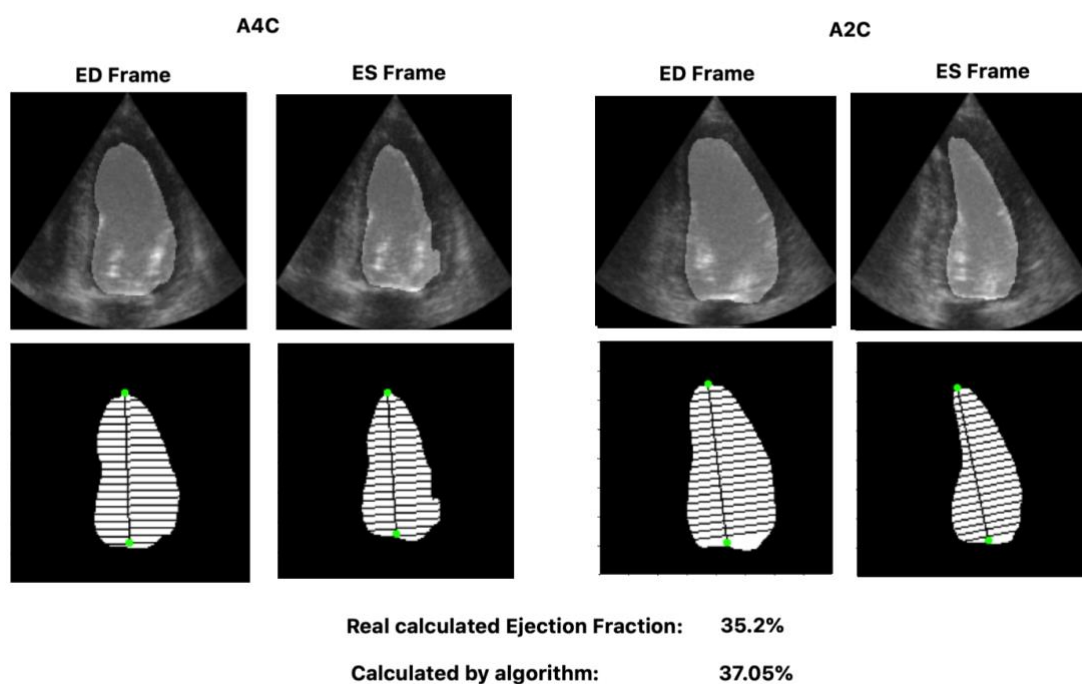


Рисунок 5.4 – Результат 2 роботи алгоритму

Для аналітичного підходу використовувалися методи OpenCV для підготовки маски до підрахунків на основі Біпланного методу дисків. Точність реалізованого підходу перевірялася на 450 пацієнтах на реальних масках та усереднений результат похибки становив $\pm 1.67\%$.

У цілому похибка повного алгоритму склала $\pm 5.55\%$ – це усереднений результат по валідаційному наборі даних. З рисунків 5.3 та 5.4 видно, що алгоритм правильно підрахував фракцію викиду ЛШ у пацієнта з нормальною роботою серця (52.48%) та з вадами серця (37.05%). При фракції викиду нижче 50% слід робити додаткову діагностику.

ВИСНОВКИ

Методи та алгоритми глибинного навчання широко впроваджуються в медичну галузь за останні роки. Вони допомагають лікарям точніше діагностувати пацієнтів, робити прогнози щодо майбутнього здоров'я пацієнтів та рекомендувати кращі методи лікування. На даний момент серцево-судинні захворювання знаходяться на першому місці, як причина високої смертності у всьому світі. Своєчасне діагностування та лікування захворювання відіграють напевно найважливішу роль у скороченні смертності від даного захворювання. Діагностика пацієнта займає певний час, тому кількість пацієнтів обстежених за день може бути достатньо невеликою. Дану проблему може вирішити автоматизування задач та підрахунків під час діагностики пацієнта, які можуть бути автоматизовані, що допоможе медичним працівникам швидше і навіть точніше діагностувати захворювання, вимикаючи суб'єктивний фактор. Серед таких задач є підрахунок фракції викиду лівого шлуночка серця.

Фракція викиду лівого шлуночка (ФВЛШ) є важливим прогностичним показником серцево-судинних захворювання. Це вимірювання, виражене у відсотках, кількості крові, яку відкачує лівий шлуночок при кожному скороченні. Аналіз 2D ехокардіографічних зображень відіграє вирішальну роль у клінічній рутині для вимірювання морфології та функції серця та для встановлення діагнозу. Такий аналіз базується на інтерпретації клінічних показників, які отримують при низькому рівні обробки зображень, таких як сегментація та відстеження. Наприклад, вилучення фракції викиду лівого шлуночка вимагає точного розмежування лівого ендокард шлуночків як у кінцевій діастолі, так і в кінці систоли. Підрахунок фракції викиду лівого шлуночка вимагає точного окреслення ендокарда лівого шлуночка як в кінцевій діастолі так і в кінцевій систолі. Для того, щоб автоматично розраховувати фракцію викиду на двовимірних зображеннях, необхідно спочатку сегментувати лівий

шлуночок на усіх фреймах у ехокардіографічному відео для 2- та 4-х камерних видів. Після цього потрібно підрахувати площини отриманих сегментів на кожному фреймі серцевого циклу та обрати фрейми з мінімальною та максимальною площиною ЛШ окремо для 2- та 4-х камерних видів. Отримані фрейми будуть відповідати діастолі та систолі ЛШ. В задачі сегментації використовується модель глибинного навчання U-Net та її модифікації. U-Net – це техніка сегментації зображень, розроблена головним чином для медичного аналізу зображень, яка дозволяє точно сегментувати зображення, використовуючи обмежену кількість навчальних даних.

Наступним кроком після отримання фреймів з масками було застосування аналітичних методів для роботи із зображеннями для виділення контуру, знаходження довжини ЛШ та розбиття усієї площини ЛШ на рівні за висотою диски та застосування Біпланний метод для підрахунку фракції викиду.

Також було проведено ряд досліджень, які показали, що модифікація класичної архітектури U-Net MFP-UNet показує кращі характеристики по Dice та Ассигасу оцінкам, а використання новітнього підходу – додавання CoordConv шару – та використання аугментації даних збільшує точність результату моделі по віще вказаним оцінкам та якість отриманих масок. Обрана модель для вирішення задачі сегментації лівого шлуночку є стійкою по проблем, які виникають під час сегментації медичних зображень, такі як низька якість зображення, розмиті границі ЛШ, а також мати високий рівень відгуку.

Будо розроблено алгоритм на основі MFP-UNet та аналітичного підходу с використання Біпланного методу дисків для підрахунку фракції викиду та отамані результати. Точність даного підходу становить – 92% з похибкою $\pm 5.55\%$.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Terziyan, V., Tkachuk, A., Tsymbal, A., & Puuronen, S. Intelligent Medical Diagnostics System Based on Integration of Statistical Methods // Informatica Medica Slovenica. 1996. Vol. 3, pp. 109-114.
2. Seckeler M., Hoke T. The worldwide epidemiology of acute rheumatic fever and rheumatic heart disease // Clin Epidemiol. 2011.
3. Gaziano T., Bitton A., Anand S., Abrahams-Gessel S., Murphy A. Growing epidemic of coronary heart disease in low-and middle-income countries // Curr Probl Cardiol. 2010. №35, pp. 72–115.
4. Hajouli S., Ludhwani D. Heart Failure and Ejection Fraction // Treasure Island (FL): StatPearls Publishing. 2021.
5. Kosaraju A., Goyal A., Grigorova Y., Makaryus A. Left Ventricular Ejection Fraction // Treasure Island (FL): StatPearls Publishing. 2021.
6. Long O., Ye X., Zhao O. Artificial intelligence and automation in valvular heart diseases // Cardiology Journal, 2020. Vol. 27, no.4, pp. 404-420.
7. H. Feigenbaum. Echocardiography: An overview // Journal of the American College of Cardiology, 2010. Vol. 1, pp. 216-224.
8. Waghilokar K.B, Fisher C.M, Goodson A., Herrick C.D, Rees M., Toscano E., MacRae C.A, Scirica B.M., Desai A.S., Murphy S.N. Extraction of Ejection Fraction from Echocardiography Notes for Constructing a Cohort of Patients having Heart Failure with reduced Ejection Fraction (HFrEF) // J Med Syst. 2018. №42, p. 209.
9. Nikolenko S., Kadurin A., Arkhangel'skaya E. Glubokoe obuchenie. Pogruzhenie v mir neyronnykh setey [Deep learning. Immersion in the world of neural networks] // St. Petersburg: Izdatel'skiy dom «Piter» Publ. 2018. p.482
10. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation // International Conference on Medical Image

Computing and Computer-Assisted Intervention - MICCAI 2015. Lecture Notes in Computer Science. 2015. Vol. 9351, pp. 234-241.

11. Guide to convolution arithmetic for deep learning. URL: <https://arxiv.org/abs/1603.07285> (дата звернення: 07.03.2021).

12. Yamashita R., Nishio M., Do R., Togashi K. Convolutional neural networks: an overview and application in radiology // Insights into Imaging. 2018. Vol. 9, pp. 611–629.

13. Nair V. and Hinton G.E. Rectified linear units improve restricted boltzmann machines // Proceedings of the 27th International Conference on International Conference on Machine Learning (ICML'10). 2010. pp. 807-814.

14. Bulwer B., Solomon S., Janardhanan R. Echocardiographic Assessment of Ventricular Systolic Function // Essential Echocardiography. 2007. pp. 89-117.

15. Dangi S., Yaniv Z., Linte C.A. Left Ventricle Segmentation and Quantification from Cardiac Cine MR Images via Multi-task Learning // Stat Atlases Comput Models Heart. 2019. pp. 21–31.

16. Leclerc S., Smistad E., Pedrosa J., Ostvik A., Cervenansky F., Espinosa F., Espeland T., Berg E., Jodoin P., Grenier T., Lartizien C., Dhooge J., Lovstakken L., Bernard O. Deep Learning for Segmentation Using an Open Large-Scale Dataset in 2D Echocardiography // IEEE Trans Med Imaging. 2019. Vol. 38(9), pp. 2198-2210.

17. Smistad E., Ostvik A., Salte I., Melichova D., Nguyen T., Haugaa K., Brunvand H., Edvardsen T., Leclerc S., Bernard O., Grenne B., Lovstakken L. Real-Time Automatic Ejection Fraction and Foreshortening Detection Using Deep Learning // IEEE Trans Ultrason Ferroelectr Freq Control. 2020. Vol. 67(12), pp. 2595-2604.

18. J. Shruti. A survey of loss functions for semantic segmentation // IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), 2020.

19. Kheneda M., Kollerathua V., Krishnamurthia G. Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers // *Medical Image Analysis*. 2019. Vol. 51, pp. 21-45.

20. Data Augmentation for Brain-Tumor Segmentation: A Review. URL: <https://www.frontiersin.org/articles/10.3389/fncom.2019.00083/full> (дата звернення: 01.04.2021).

21. Semi-supervised Task-driven Data Augmentation for Medical Image Segmentation. URL: <https://arxiv.org/pdf/2007.05363.pdf> (дата звернення: 01.04.2021).

22. Siddique N., Sidike P., Elkin C., Devabhaktuni V. U-Net and its variants for medical image segmentation: theory and applications // *Computer Science, Engineering*, 2020.

23. Attention U-Net: Learning Where to Look for the Pancreas. URL: https://smcdonagh.github.io/papers/attention_u_net_learning_where_to_look_for_the_pancreas.pdf (дата звернення: 18.03.2021).

24. Inception Modules Enhance Brain Tumor Segmentation. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6640210/> (дата звернення: 18.03.2021).

25. Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation. URL: <https://arxiv.org/pdf/1802.06955.pdf> (дата звернення: 20.03.2021).

26. Review: UNet++ — A Nested U-Net Architecture (Biomedical Image Segmentation). URL: <https://sh-tsang.medium.com/review-unet-a-nested-u-net-architecture-biomedical-image-segmentation-57be56859b20> (дата звернення: 18.03.2021).

27. UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. URL: <https://arxiv.org/pdf/1912.05074.pdf> (дата звернення: 18.03.2021).

28. Moradi S., Oghli M., Alizadehasl A., Shiri I., Oveisi N., Oveisi, Majid Maleki M., Dhooge J. MFP-UNet: A novel deep learning based approach for left ventricle segmentation in echocardiography // *Physica Medica*, 2019.
29. An intriguing failing of convolutional neural networks and the CoordConv solution. URL: <https://arxiv.org/pdf/1807.03247.pdf> (дата звернення: 10.04.2021).
30. Investigating CoordConv for Fully and Weakly Supervised Medical Image Segmentation. URL: <https://hal-normandie-univ.archives-ouvertes.fr/hal-03088385/document> (дата звернення: 11.04.2021).
31. Automated Segmentation of Pulmonary Lobes using Coordination-Guided Deep Neural Networks. URL: <https://arxiv.org/abs/1904.09106> (дата звернення: 05.04.2021).
32. Mozaffari M., Sankoff D., Lee W. BowNet: Dilated convolutional neural network for ultrasound tongue contour extraction // *The Journal of the Acoustical Society of America*. 2019. Vol. 146(4), pp. 2940-2941.
33. CAMUS Dataset. URL: <https://www.creatis.insa-lyon.fr/Challenge/camus/>. (дата звернення: 12.04.2021).