

ХЕРСОНСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА ДИЗАЙНУ
КАФЕДРА ПРОГРАМНИХ ЗАСОБІВ І ТЕХНОЛОГІЙ

Пояснювальна записка

до кваліфікаційної роботи

магістра
(освітній рівень)

на тему: «Дослідження методів пошуку патологій судин на основі
нейромережевої обробки зображень»

Виконав: студент групи 6ПР2
спеціальності
121 - «Інженерія програмного забезпечення»
(шифр і назва спеціальності)

Харченко Олександр Геннадійович
(прізвище та ініціали)

Керівник д.т.н., проф. Шерстюк В.Г.
(прізвище та ініціали)

Рецензент к.т.н. доцент Козел В.Н.
(прізвище та ініціали)

Хмельницький - 2025

Херсонський національний технічний університет

(повне найменування вищого навчального закладу)

Факультет, відділення Інформаційних технологій та дизайну
Кафедра Програмних засобів і технологій
Освітній рівень магістр
(шифр і назва)
Напрямок підготовки ОПП - Програмне забезпечення систем
Спеціальність 121 – Інженерія програмного забезпечення
(шифр і назва)

ЗАТВЕРДЖУЮ

Завідувач кафедри ПЗіТ

к.т.н. доц. О.Є. Огнева

“ _____ ” _____ 2025 р.

З А В Д А Н Н Я
НА КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТУ

Харченко Олександр Геннадійовичу

(прізвище, ім'я, по батькові)

1. Тема роботи «Дослідження методів пошуку патологій судин на основі нейромережевої обробки зображень»

керівник роботи д.т.н., проф. Шерстюк В.Г.,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджена наказом вищого навчального закладу від 15.09.2025 р. №417-С

2. Строк подання студентом роботи 15.12.2025

3. Вихідні дані до роботи літературні та періодичні джерела, датасети зображень рентгенівської коронарографії ARCADE та CADICA, маски судин та мітки ступеня стенозу

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити): 1) аналіз існуючих методів сегментації та класифікації стенозу; 2) розроблення та порівняння методів попередньої обробки зображень; 3) розроблення моделей сегментації коронарних артерій; 4) розроблення моделей сегментації стенозу; 5) реалізація моделі класифікації стенозу за ступенем тяжкості; 6) інтеграція системи та її експериментальна перевірка .

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)

1) ангіографічні зображення з прикладами патологій; 2) результати застосування різних методів попередньої обробки; 3) порівняльні діаграми ефективності моделей сегментації та класифікації; 4) схема архітектури комплексної системи аналізу зображень

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

7. Дата видачі завдання _____ 29.09.2025 _____

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів виконання роботи	Термін виконання етапів роботи	Примітки
1.	Отримання завдання	29.09.2025	Виконано
2.	Підбір літератури	05.10.2025	Виконано
3.	Аналіз предметної області	19.10.2025	Виконано
4.	Розробка та обґрунтування завдання	26.10.2025	Виконано
5.	Розробка концептуальної моделі	05.11.2025	Виконано
6.	Розробка алгоритму	12.11.2025	Виконано
7.	Проектування програми	26.11.2025	Виконано
8.	Розробка інтерфейсу програми	30.11.2025	Виконано
9.	Тестування програми	05.12.2025	Виконано
10.	Оформлення пояснювальної записки	10.12.2025	Виконано
11.	Захист кваліфікаційної роботи	15.12.2025	Виконано

Студент _____ Харченко О.Г.
(підпис) (прізвище та ініціали)

Керівник роботи _____ Шерстюк В.Г.
(підпис) (прізвище та ініціали)

РЕФЕРАТ

Пояснювальна записка: 167 сторінок, 76 рисунків, 18 таблиць, 1 додаток, 32 джерела.

Об'єкт дослідження: цілісний процес діагностування патологій коронарних судин на основі ангиографічних зображень

Предмет дослідження: алгоритми глибокого навчання й комп'ютерного зору, призначені для автоматичного виявлення гетерогенних судин, сегментації та класифікації стенозу.

Мета дослідження: Створити та впровадити автоматизовану систему, здатну виявляти гетерогенні судини та стенози на рентгенівських ангиограмах із застосуванням сучасних методів машинного навчання.

Новизна отриманих результатів полягає в поєднанні каскадних сегментаційних моделей із класифікатором ResNet у єдиній топології, що дає змогу одночасно враховувати морфологію судин і характеристики стенозу.

Практична цінність результатів роботи полягає у можливості інтеграції системи в САД-платформи для допоміжної діагностики.

Перелік ключових слів: ПОШУК ГЕТЕРОГЕНИХ СУДИН, РЕНТГЕНІВСЬКА КОРОНАРОГРАФІЯ, СТЕНОЗ, МАШИННЕ НАВЧАННЯ, АВТОМАТИЗОВАНА ДІАГНОСТИКА.

АНОТАЦІЯ

Кваліфікаційна робота магістра складається зі вступу, чотирьох розділів, висновку, переліку використаних джерел та додатків.

Роботу присвячено розробленню, теоретичному обґрунтуванню та експериментальній перевірці комплексної системи, що поєднує дві моделі U-Net для сегментації коронарних артерій і стенозу та модель ResNet для кількісного визначення ступеня стенозу.

Використання методів машинного навчання та глибокого навчання для автоматизованого аналізу рентгенівських ангиограм відкриває нові можливості для покращення точності діагностики та ефективності лікування.

Використання згорткових нейронних мереж, моделей сегментації (наприклад, UNet) та глибокого навчання дозволяє автоматично ідентифікувати гетерогенні судини, визначати ступінь стенозу та аналізувати динаміку змін у судинах пацієнтів.

Запропонована архітектура забезпечує інтегрований ланцюжок «зображення-сегментація-класифікація», що зменшує суб'єктивний вплив лікаря, скорочує час аналізу та створює підґрунтя для стандартизованого клінічного рішення. Експериментальні дослідження проведено на датасетах ARCADE та CADICA.

В результаті виконання роботи вперше запропоновано систему на основі штучного інтелекту, яка не лише автоматично виявляє гетерогенні судини та визначає ступінь їх ураження, а й здійснює класифікацію стенозу за його давністю, що дозволяє покращити процес діагностики, оскільки лікар отримує не лише інформацію про наявність стенозу, а й дані про його характер, що є ключовим для вибору оптимальної терапевтичної тактики.

ABSTRACT

The master's thesis consists of an introduction, four chapters, a conclusion, a list of sources and appendices.

The thesis is devoted to the development, theoretical justification and experimental verification of a complex system that combines two U-Net models for coronary artery segmentation and stenosis as well as a ResNet model for quantifying the degree of stenosis.

The use of machine learning and deep learning methods for automated analysis of X-ray angiograms opens up new opportunities for improving diagnostic accuracy and treatment effectiveness.

The use of convolutional neural networks, segmentation models (e.g. UNet) and deep learning allowing the automatic identification of heterogeneous vessels, determination of the degree of stenosis and analysis of the dynamics of changes in the vessels of patients.

The proposed architecture provides an integrated chain of "image-segmentation-classification", which reduces the subjective influence of the doctor, reduces the analysis time and creates the basis for a standardized clinical decision. Experimental studies were conducted on the ARCADE and CADICA datasets.

As a result of the thesis, a system based on artificial intelligence was proposed for the first time, which not only automatically detects heterogeneous arteries and determines the degree of their damage, but also classifies stenosis according to its duration, which allows improving the diagnostic process, since the doctor receives not only information about the presence of stenosis, but also data about its nature, which is key for choosing the optimal therapeutic tactics.

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ ТА УМОВНИХ ПОЗНАЧЕНЬ

ВСТУП

РОЗДІЛ 1. ТЕОРЕТИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ ТА МЕДИЧНІ АСПЕКТИ СТЕНОЗУ

1.1. Стеноз коронарних судин: характеристика, причини та клінічне значення

1.1.1. Етіологія та патогенез стенозу

1.1.2. Види стенозу

1.1.3. Клінічні прояви стенозу

1.1.4. Методи діагностики стенозу

1.2. Коронарографія та її особливості. КТ серця

1.2.1. Коронарографія

1.2.2. Комп'ютерна томографія серця

1.2.3. Порівняння коронарографії та КТ серця

1.3. Огляд методів з обробки зображень у визначенні патологій коронарних судин

1.3.1. Методи машинного навчання в обробці ангиографічних зображень

1.3.2. Використання глибоких нейронних мереж у аналізі судинних патологій

1.4. Структура роботи

Висновки до розділу 1

РОЗДІЛ 2. МЕТОДИ ТА ТЕХНОЛОГІЇ ПОПЕРЕДНЬОЇ ОБРОБКИ МЕДИЧНИХ ЗОБРАЖЕНЬ

2.1. Постановка задачі попередньої обробки

2.2. Огляд сучасних методів обробки медичних зображень

2.2.1. Фільтрація шумів

2.2.2. Нормалізація інтенсивностей

2.2.3. Покращення контрасту

2.3. Порівняльний аналіз методів обробки зображень

2.3.1. Фільтрація шумів

2.3.2. Нормалізація інтенсивностей

2.3.3. Покращення контрасту

2.4. Розв'язання задачі попередньої обробки зображень

Висновки до розділу 2

РОЗДІЛ 3. СЕГМЕНТАЦІЯ КОРОНАРНИХ АРТЕРІЙ. МЕТОДИ ВИДІЛЕННЯ ОЗНАК ТА СЕГМЕНТАЦІЇ СТЕНОЗУ

3.1. Постановка задачі сегментації коронарних артерій

3.2. Огляд сучасних методів та підходів до сегментації судин.

3.2.1. Автоматизована сегментація коронарних артерій з використанням YOLO-Angio

3.2.2. Ансамблевий підхід до сегментації та локалізації патологій

3.2.3. SAM-VMNet: Гібридна модель для покращеної сегментації судин

3.3. Методи та підходи, застосовані в роботі

3.3.1. Архітектура YOLOv8

3.3.2. Архітектура U-Net

3.3.3. Багатокласова сегментація

3.3.4. Бінарна сегментація

3.4. Реалізація та тренування моделей

3.4.1. Вхідні дані

3.4.2. Аугментація

3.4.3. Навчання моделі U-Net

3.4.4. Навчання моделі YOLOv8

3.5. Метрики для порівняння

3.6. Аналіз якості сегментації коронарних артерій

3.6.1. Багатокласова сегментація

3.6.2. Бінарна сегментація

3.7. Постановка задачі сегментації стенозу

- 3.7.1. Обраний метод для сегментації
 - 3.7.2. Розв'язання задачі сегментації коронарних артерій
 - 3.7.3. Автоматизоване виявлення стенозу
 - 3.8. Виділення ознак стенозу по зображенню
 - 3.8.1. Геометричні ознаки звуження судин
 - 3.8.2. Інтенсивнісні та текстурні характеристики
 - 3.9. Огляд сучасних методів та підходів до сегментації стенозу
 - 3.9.1. StenUNet для точної сегментації стенозів
 - 3.9.2. SSASS: Напівконтрольоване навчання для сегментації
 - 3.9.3. Порівняльний аналіз ефективності моделей YOLO та DETR у задачі сегментації стенозів
 - 3.10. Методи та підходи, застосовані в роботі
 - 3.10.1. Модель YOLOv8
 - 3.10.2. Модель U-Net
 - 3.10.3. Підхід із використанням псевдоміток
 - 3.10.4. Класичний підхід навчання
 - 3.11. Реалізація та тренування моделей
 - 3.11.1. Вхідні дані та їх попередня обробка
 - 3.11.2. Аугментація
 - 3.11.3. Навчання з псевдомітками
 - 3.11.4. Навчання з ручною розміткою
 - 3.12. Оцінка якості сегментації
 - 3.12.1. Метрики оцінки якості сегментації
 - 3.12.2. Аналіз якості сегментації стенозу
 - 3.13. Розв'язання задачі сегментації стенозу
 - 3.13.1. Обраний метод для задачі сегментації стенозу
 - 3.13.2. Цикл розв'язання задачі сегментації стенозу
- Висновки до розділу 3

РОЗДІЛ 4. РЕАЛІЗАЦІЯ МЕТОДУ КЛАСИФІКАЦІЇ СТЕНОЗУ

4.1. Постановка задачі класифікації стенозу

4.2. Метод класифікації за зображеннями

4.3. Огляд сучасних методів та підходів до класифікації стенозу

4.3.1. Використання архітектури ResNet у задачах класифікації стенозів

4.3.2. Адаптація YOLOv8 для класифікації ступеня стенозу

4.3.3. Ефективність архітектури EfficientNet для визначення патологічних станів

4.4. Методи та підходи, що застосовуються в роботі

4.4.1. Класифікатор ResNet

4.4.2. Класифікатор EfficientNet

4.5. Реалізація та тренування моделей

4.5.1. Вхідні дані та їх попередня обробка

4.5.2. Навчання ResNet

4.5.3. Навчання EfficientNet

4.6. Метрики оцінки якості класифікації

4.7. Аналіз якості класифікації стенозу

4.8. Обраний метод для задачі класифікації стенозу

4.9. Розв'язання задачі класифікації стенозу

4.9.1. Архітектура та реалізація системи.

4.9.2. Спосіб застосування системи.

4.9.3. Узагальнення результатів роботи системи.

4.9.4. Результати класифікації.

Висновки до розділу 4

ВИСНОВКИ

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

ДОДАТОК А. ЛІСТИНГ ПРОГРАМИ

А.1. Код для попередньої обробки зображень

А.2. Код для тренування моделі сегментації артерій

A.3. Код для тренування моделі сегментації стенозу

A.4. Код для тренування моделі класифікації стенозу

A.5. Код Flask API для виклику моделі

ПЕРЕЛІК СКОРОЧЕНЬ ТА УМОВНИХ ПОЗНАЧЕНЬ

AI - Штучний інтелект (Artificial Intelligence).

CAD - Комп'ютерна допоміжна діагностика (Computer-Aided Diagnosis).

CAG - Коронарна ангиографія (Coronary Angiography). CNN - Згорткова нейронна мережа (Convolutional Neural Network).

СТА - КТ-ангиографія (Computed Tomography Angiography).

DL - Глибоке навчання (Deep Learning).

GAN - Генеративно-змагальна нейромережа (Generative Adversarial Network).

ICA - Інвазивна коронарна ангиографія (Invasive Coronary Angiography).

LAD - Передня міжшлуночкова гілка лівої коронарної артерії (Left Anterior Descending Artery).

LCA - Ліва коронарна артерія (Left Coronary Artery).

ML - Машинне навчання (Machine Learning).

PCI - Черезшкірне коронарне втручання (Percutaneous Coronary Intervention).

QCA - Кількісна коронарна ангиографія (Quantitative Coronary Angiography).

ROI - Область інтересу (Region of Interest).

SVM - Метод опорних векторів (Support Vector Machine).

UNet - Архітектура згорткової нейронної мережі для сегментації зображень (UNet Architecture for Image Segmentation).

XCA - Рентгенівська коронарна ангиографія (X-ray Coronary Angiography). YOLO - Ви тільки дивитесь один раз (You Only Look Once) – архітектура нейромереж для детекції об'єктів.

DETR – Detection Transformer, модель для детекції об'єктів на основі трансформерів.

SAM – Segment Anything Model, модель для універсальної сегментації об'єктів.

ВСТУП

За останні роки медична галузь продемонструвала значний прогрес у діагностиці та лікуванні серцево-судинних захворювань завдяки розвитку сучасних технологій. Одним із ключових методів дослідження стану коронарних судин є рентгенівська коронарографія (ХСА), яка дозволяє отримувати детальні зображення судинного русла серця. Проте навіть цей метод має певні обмеження, оскільки його точність значною мірою залежить від якості зображення та суб'єктивної оцінки лікаря. Виявлення стенозу та оцінка його тяжкості можуть бути ускладнені через анатомічні особливості судин, артефакти зображення та обмежену роздільну здатність.

Автоматизація аналізу рентгенівських ангиограм за допомогою методів машинного навчання (ML) відкриває нові можливості для підвищення точності та швидкості діагностики серцево-судинних патологій. Використання згорткових нейронних мереж (CNN), моделей сегментації (наприклад, UNet) та глибокого навчання (DL) дозволяє автоматично ідентифікувати гетерогенні судини, визначати ступінь стенозу та аналізувати динаміку змін у судинах пацієнтів.

Мета дослідження – створити та впровадити автоматизовану систему, здатну виявляти гетерогенні судини та стенози на рентгенівських ангиограмах із застосуванням сучасних методів машинного навчання.

Дослідження спрямоване на підвищення ефективності аналізу медичних зображень та розробку підходу, що дозволить зменшити вплив суб'єктивного чинника під час діагностики.

У рамках роботи буде проведено огляд існуючих підходів до автоматизованого аналізу коронаро-графічних зображень, визначено їхні переваги та недоліки. На основі цього аналізу буде запропоновано нову методику, що дозволяє не тільки ідентифікувати патологічні зміни в судинах, а й проводити класифікацію стенозу за його характеристиками.

Актуальність дослідження. Незважаючи на значний прогрес у розробці технологій медичної візуалізації, існуючі методи аналізу коронарографічних зображень часто не дозволяють точно визначити структуру уражених судин, ступінь стенозу та його давність (гострий чи хронічний стеноз). Це може призводити до затримки в прийнятті клінічних рішень та підвищеного ризику ускладнень для пацієнтів.

Використання методів машинного навчання (ML) та глибокого навчання (DL) для автоматизованого аналізу рентгенівських ангиограм відкриває нові можливості для покращення точності діагностики та ефективності лікування.

Наукова новизна. У межах цього дослідження вперше запропоновано систему на основі штучного інтелекту, яка не лише автоматично виявляє гетерогенні судини та визначає ступінь їх ураження, а й здійснює класифікацію стенозу за його давністю. Такий підхід дозволяє покращити процес діагностики, оскільки лікар отримує не лише інформацію про наявність стенозу, а й дані про його характер, що є ключовим для вибору оптимальної терапевтичної тактики.

Таким чином, дослідження в напрямку автоматизованого аналізу рентгенівських коронарографій із використанням методів машинного навчання має потенціал значно покращити якість медичної допомоги, знизити навантаження на медичних спеціалістів та забезпечити пацієнтам своєчасну та точну діагностику серцево-судинних патологій.

РОЗДІЛ 1. ТЕОРЕТИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ ТА МЕДИЧНІ АСПЕКТИ СТЕНОЗУ

1.1. Стеноз коронарних судин: характеристика, причини та клінічне значення

Стеноз коронарних артерій є одним із ключових патологічних станів, що призводять до розвитку ішемічної хвороби серця (ІХС). Він характеризується звуженням просвіту коронарних артерій через накопичення атеросклеротичних бляшок [1], що перешкоджає нормальному кровотоку до серцевого м'яза (рис. 1.1, світлий контур позначає нормальний просвіт судини, тоді як у ділянці, позначеній стрілкою, спостерігається різке звуження через атеросклеротичну бляшку). Це може призводити до стенокардії, серцевої недостатності та гострих коронарних синдромів, включаючи інфаркт міокарда.

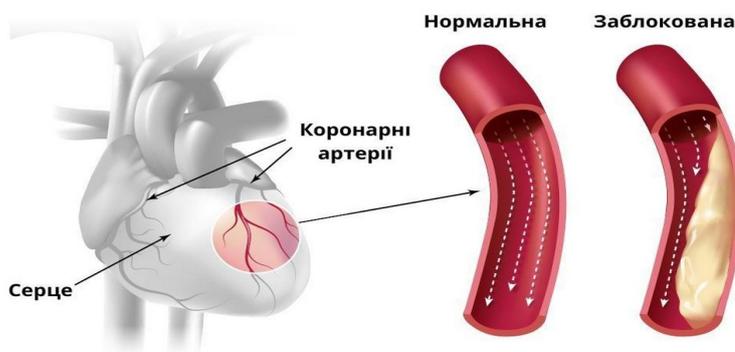


Рисунок 1.1. Стеноз коронарної артерії, вид з ангіограми¹.

1.1.1. Етіологія та патогенез стенозу

Патогенез стенозу коронарних артерій є складним багатофакторним процесом, що включає як генетичні, так і зовнішні чинники. Основною причиною розвитку стенозу є атеросклероз – хронічне запальне

¹ Джерело зображення:

https://mydoctor.kaiserpermanente.org/mas/Images/gallery_coronary_artery_tcm88-2250102.png

захворювання, яке призводить до накопичення ліпідних відкладень у стінках артерій та поступового звуження їхнього просвіту. Вважається, що початковий етап розвитку атеросклеротичних змін у судинах включає порушення функції ендотелію [2], яке виникає внаслідок дії факторів ризику, таких як гіперхолестеринемія, артеріальна гіпертензія, куріння, цукровий діабет, ожиріння та малорухливий спосіб життя.

На першій стадії патогенезу внаслідок ушкодження ендотеліальних клітин спостерігається підвищене накопичення ліпопротеїнів низької щільності у судинній стінці. Під впливом окислювальних процесів ці ліпіди модифікуються і стають імуногенними, що призводить до залучення макрофагів, які фагоцитують окислені ліпопротеїди та перетворюються на пінисті клітини. Саме пінисті клітини формують основу атеросклеротичної бляшки, що поступово збільшується в розмірах. У відповідь на цей процес активізуються гладком'язові клітини судинної стінки, які мігрують до місця запалення та починають синтезувати позаклітинний матрикс, формуючи фіброзну капсулу бляшки.

На наступних етапах розвитку патологічного процесу спостерігається прогресуюче накопичення ліпідів та кальцію, що сприяє ущільненню бляшки та обмеженню еластичності судинної стінки. Це може призвести до критичного зменшення просвіту артерії, що значно знижує об'єм кровотоку та призводить до розвитку ішемії міокарда. У певних випадках бляшка може залишатися стабільною протягом тривалого часу, однак при впливі механічних чи біохімічних факторів відбувається її розрив, що призводить до утворення тромбу на місці ураження. Такий тромб може повністю перекрити просвіт судини, спричинивши гострий коронарний синдром, який характеризується тяжкими клінічними наслідками, включаючи розвиток інфаркту міокарда або раптову серцеву смерть.

Таким чином, патогенез стенозу є складним процесом, який розгортається протягом багатьох років і включає низку патофізіологічних

змін на рівні клітинної та тканинної структури артерій. Розуміння цих механізмів є ключовим для розробки ефективних методів профілактики, діагностики та лікування атеросклеротичних уражень коронарних судин.

1.1.2. Види стенозу

Стеноз коронарних судин має різні морфологічні та клінічні характеристики, які впливають на тактику лікування та прогноз перебігу захворювання. Одним із ключових параметрів є ступінь звуження артерії, оскільки від нього залежить об'єм кровотоку, що потрапляє до міокарда (рис. 1.2, діаграма порівнює п'ять типових рівнів стенозу (мінімальний – оклюзія) із вказанням відсотка перекриття просвіту та умовним кольоровим кодуванням).

Легкий стеноз, при якому звуження не перевищує 50%, зазвичай не супроводжується клінічними симптомами через компенсаторні механізми судинної системи.

Помірний стеноз у межах 50–70% може проявлятися епізодами стенокардії при фізичному навантаженні, що обумовлено частковою недостатністю кровопостачання серцевого м'яза.

Критичний стеноз, при якому звуження перевищує 70%, викликає виражену ішемію та може призвести до тяжких ускладнень, таких як інфаркт міокарда.

Окрім ступеня звуження, важливу роль відіграє стабільність атеросклеротичної бляшки. Стабільний стеноз характеризується наявністю щільної фіброзної капсули, яка робить бляшку стійкою до розривів, що зменшує ризик тромбозу. Такі ураження мають прогнозовано стабільний перебіг та рідко спричиняють гострі коронарні події. Нестабільний стеноз, навпаки, виникає унаслідок формування бляшок із тонкою капсулою та великою кількістю ліпідного ядра, що робить їх схильними до розриву. У

таких випадках утворення тромбу може відбутися раптово, що значно підвищує ймовірність розвитку гострих коронарних синдромів [3].



Рисунок 1.2. Градація ступенів звуження коронарної артерії².

Локалізація стенозу також має важливе клінічне значення. Проксимальний стеноз, що розвивається у великих коронарних артеріях, таких як ліва передня міжшлуночкова артерія або ліва головна коронарна артерія, має більш виражений вплив на кровопостачання міокарда та пов'язаний із високим ризиком серйозних ускладнень. Такі ураження часто потребують хірургічного втручання для відновлення нормального кровотоку. Дистальний стеноз, що локалізується у периферичних відгалуженнях коронарних артерій, зазвичай має менш значний вплив на гемодинаміку серця, однак може бути складним для ендovasкулярного лікування через анатомічні особливості судинного русла.

Враховуючи різноманітність форм стенозу, його класифікація є важливим етапом у плануванні лікування. Вибір тактики терапії залежить від

² Джерело зображення:

<https://heartcureguide.wordpress.com/wpcontent/uploads/2020/05/picture1-4.png>

поєднання декількох параметрів, включаючи ступінь звуження, стабільність атеросклеротичних бляшок та їхню локалізацію.

1.1.3. Клінічні прояви стенозу

Клінічні прояви стенозу коронарних судин безпосередньо залежать від ступеня звуження артерій, його локалізації та тривалості розвитку патологічного процесу (табл. 1.1, у структурованому вигляді наведено шість градацій із поясненням, як кожен діапазон відображає фізіологічний вплив на коронарний кровообіг).

Таблиця 1.1. Класифікація коронарного стенозу за категоріями [4] та відсотком перекриття просвіту.

Категорія	Ступінь стенозу	Відсоток стенозу
0	Нормальний	0%
1	Мінімальний	1-24%
2	Легкий	25-49%
3	Помірний	50-69%
4	Виражений	70-99%
5	Оклюзія	100%

На ранніх стадіях, коли стеноз не перевищує 50%, кровотік залишається відносно компенсованим за рахунок вазодилатації та розвитку колатерального кровопостачання, що дає змогу міокарду отримувати достатню кількість кисню навіть при підвищеній фізичній активності. У таких випадках симптоматика або відсутня, або має неспецифічний характер, наприклад, легку втому або періодичні дискомфортні відчуття в грудній клітці.

У міру прогресування звуження просвіту коронарної артерії до 50–70% виникають перші виражені симптоми ішемії міокарда, що проявляються стенокардією напруження. Цей стан характеризується болем або стискаючим

дискомфортом у грудній клітці, що зазвичай виникає при фізичному або емоційному навантаженні та проходить у стані спокою або після прийому нітрогліцерину. Больові відчуття можуть мати різну інтенсивність і тривалість, іррадіювати у ліву руку, плече, нижню щелепу або епігастральну ділянку.

При критичному стенозі понад 70% клінічні прояви стають більш вираженими, оскільки навіть у стані спокою міокард не отримує достатньо крові для підтримання нормального функціонування. Це може призводити до розвитку стенокардії спокою, що є прогностично несприятливим фактором і свідчить про високий ризик розвитку гострого коронарного синдрому. Пацієнти часто відзначають появу задишки, відчуття нестачі повітря та виражену слабкість. У таких випадках ішемічні зміни можуть спричинити порушення серцевого ритму, зокрема фібриляцію передсердь або шлуночкову тахікардію, що значно підвищує ризик розвитку серцевої недостатності та раптової серцевої смерті.

При розриві атеросклеротичної бляшки та утворенні тромбу, що повністю перекриває просвіт артерії, розвивається гострий коронарний синдром, який може супроводжуватися інфарктом міокарда. Цей стан є ургентним і потребує негайного медичного втручання, оскільки ішемія призводить до некрозу міокардіальних клітин і розвитку тяжких ускладнень. Симптоматика інфаркту включає інтенсивний, пекучий біль у грудній клітці, що триває понад 20 хвилин, супроводжується холодним потом, нудотою, запамороченням і страхом смерті.

Отже, клінічні прояви стенозу варіюються залежно від ступеня ураження коронарних артерій і можуть коливатися від безсимптомного перебігу до тяжких станів, що становлять безпосередню загрозу для життя пацієнта. Своєчасне виявлення та лікування цього патологічного процесу є критично важливими для зменшення ризику ускладнень і покращення прогнозу захворювання.

1.1.4. Методи діагностики стенозу

Сучасна діагностика стенозу коронарних артерій базується на комплексному підході, що включає клінічне обстеження пацієнта, інструментальні методи дослідження та застосування новітніх технологій, таких як штучний інтелект для автоматизованого аналізу зображень.

Первинне діагностичне обстеження передбачає аналіз анамнестичних даних, визначення факторів ризику та оцінку симптоматики. Фізикальне обстеження дає змогу виявити ознаки порушення кровообігу та супутньої серцево-судинної патології. Електрокардіографія є базовим методом дослідження, який дозволяє виявити ішемічні зміни міокарда, такі як депресія сегмента ST, інверсія зубця T або порушення ритму серця. Однак ЕКГ не завжди є достатньо чутливим методом, особливо у випадках стенозу середнього ступеня.

Одним із найважливіших методів діагностики є рентгенівська коронарографія (ХСА), яка є інвазивною процедурою і дозволяє безпосередньо оцінити анатомічну будову коронарних артерій та виявити ступінь їхнього звуження (рис. 1.3, знімок демонструє початок введення контрасту через катетер у ліву коронарну артерію). Метод базується на введенні контрастної речовини в судини серця через катетер із подальшою рентгенівською візуалізацією. Коронарографія є «золотим стандартом» діагностики, оскільки дає змогу точно визначити локалізацію та тяжкість стенозу.

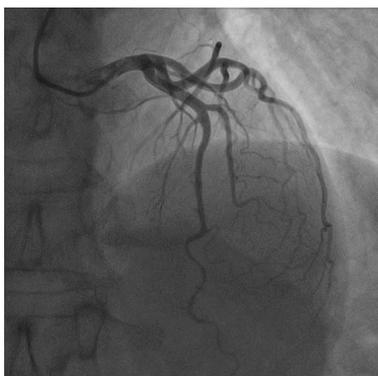


Рисунок 1.3. Рентгенівська коронарографія, селективна ін'єкція контрасту з датасету ARCADE [5]

Неінвазивною альтернативою є комп'ютерна томографічна ангиографія (КТ-ангіографія), яка забезпечує тривимірне зображення коронарних судин і дозволяє виявляти атеросклеротичні ураження (рис. 1.4). Завдяки високій просторовій роздільній здатності та можливості кількісного аналізу, КТ-ангіографія є перспективним методом для скринінгу пацієнтів із підозрою на ішемічну хворобу серця.

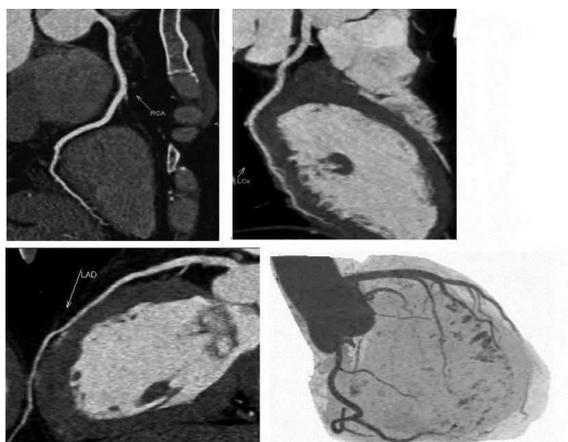


Рисунок 1.4. Комп'ютерна томографічна коронарографія³

Функціональні методи оцінки гемодинамічного значення стенозу включають фракційний резерв кровотоку (FFR), який передбачає вимірювання градієнта тиску до і після стенотичного ураження (рис. 1.5), що дає змогу визначити, чи є звуження клінічно значущим.

³ Джерело зображення: <https://mrt-ct.com.ua/images/kt/kt-koronarografiya-3.jpg>

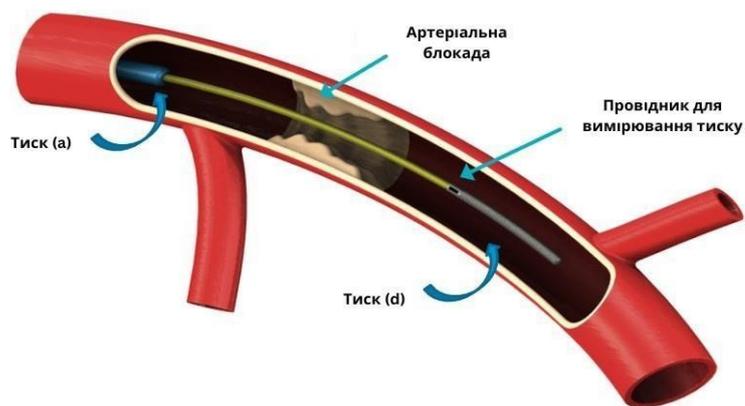


Рисунок 1.5. Фракційний резерв кровотоку⁴

Розвиток штучного інтелекту та алгоритмів глибокого навчання відкриває нові можливості в діагностиці стенозу. Використання автоматизованих систем для аналізу медичних зображень дозволяє зменшити вплив людського фактора, прискорити оцінку стану судин та покращити точність діагностики. Застосування таких технологій дає змогу аналізувати великі масиви даних та виявляти патологічні зміни, що можуть бути непомітні при візуальному аналізі лікаря.

Таким чином, сучасна діагностика стенозу включає широкий спектр методів, які дозволяють комплексно оцінити стан коронарного русла та функціональний статус міокарда. Впровадження новітніх технологій, зокрема штучного інтелекту, значно покращує діагностичні можливості та сприяє персоналізованому підходу до лікування пацієнтів із серцево-судинною патологією.

1.2. Коронарографія та її особливості. КТ серця

1.2.1. Коронарографія

Коронарографія є одним із найважливіших методів діагностики серцевосудинних захворювань, що дозволяє отримати детальну візуалізацію коронарних артерій. Завдяки високій точності та можливості проведення

⁴ Джерело зображення:

https://media.licdn.com/dms/image/v2/D5622AQGDVDSFTGWAPw/feedshare-shrink_800/feedshare-shrink_800/0/1722926565903?e=2147483647&v=beta&t=4Zoaqhl8Uj8Z5sCxTD6bzhhhrAWRIpSx xcYx49bAicH0

супутніх інтервенційних втручань цей метод залишається «золотим стандартом» у діагностиці ішемічної хвороби серця (ІХС). Коронарографія дозволяє не лише оцінити прохідність артерій, а й визначити гемодинамічну значущість стенозів, що відіграє ключову роль у виборі подальшої тактики лікування.

Процедура виконується в умовах ангиографічної лабораторії та потребує спеціального рентгенівського обладнання. Пацієнт проходить попередню підготовку, яка включає припинення прийому певних препаратів (наприклад, антикоагулянтів) та оцінку ниркової функції, оскільки контрастна речовина може негативно впливати на нирки.

Після місцевого знеболення лікар проводить катетеризацію периферичної артерії, найчастіше через стегновий (феморальний) або променевий (радіальний) доступ. Катетер просувається до устя коронарних артерій під контролем рентгеноскопії. Після правильного позиціонування катетера вводиться контрастна речовина, що дозволяє отримати чітке зображення судинного русла під час серії рентгенівських знімків (рис. 1.6).

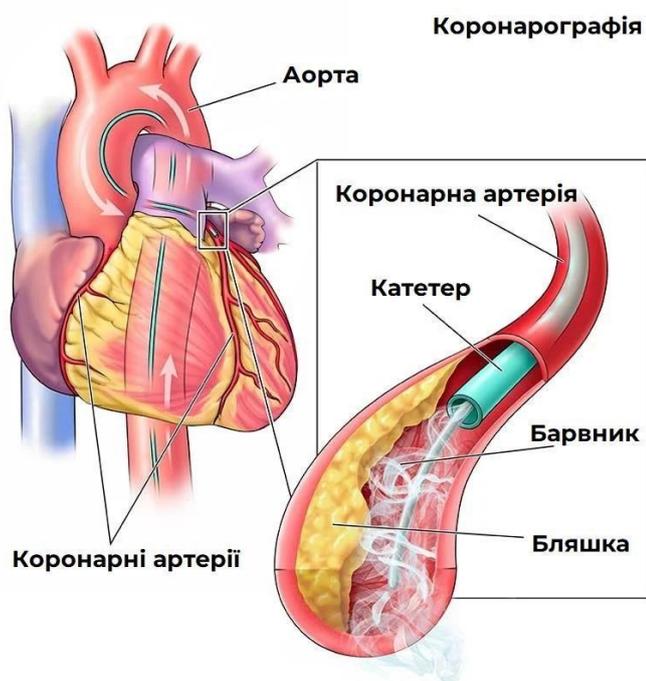


Рисунок 1.6. Методика проведення коронарографії⁵

Коронарографія дає змогу оцінити локалізацію, ступінь та довжину стенозу, наявність атеросклеротичних бляшок та їхню морфологію. Окрім цього, процедура дозволяє визначити рівень колатерального кровотоку, що є важливим фактором при плануванні подальшого лікування. У разі виявлення критичного стенозу під час коронарографії може бути виконане черезшкірне коронарне втручання (ЧКВ), зокрема балонна ангіопластика або стентування, що робить цей метод не лише діагностичним, а й лікувальним.

Сучасні ангіографічні системи для коронарографії - це високотехнологічні комплекси, що поєднують рентгенівське випромінювання з цифровою обробкою даних і широким спектром автоматизованих допоміжних функцій [6]. На відміну від установок попереднього покоління, де основна увага зосереджувалася лише на формуванні контрастного зображення, теперішні системи інтегрують плоскопанельні детектори з великою динамічною шириною, високошвидкісні генератори імпульсного випромінювання, роботизовані C- або biplane-штативи та програмні модулі.

Завдяки цьому біпланові комплекси (Philips Azurion (рис. 1.7), Siemens Artis, GE Innova, Canon Alphenix) здатні одночасно отримувати ортогональні проєкції коронарного русла з частотою до 30 кадрів/с, що мінімізує об'єм контрасту, скорочує тривалість процедури й підвищує точність визначення складної геометрії стенозів.

⁵ Джерело зображення:

<https://www.cardiothoracic-surgeon.co.uk/wpcontent/uploads/2019/04/coronary-artery-bypass-grafting-pic01.jpg>



Рисунок 1.7. Сучасна біпланова ангиографічна система з плоскопанельними детекторами Philips Azurion 7 C20⁶

Ключовою ланкою апаратів є електронно-рентгенівський блок із мікрофокусною трубкою, керованою високочастотним генератором. Імпульсний режим (переважно 3–7 мс) з перемежованою флуороскопією дає змогу підтримувати достатню контрастність і різкість дрібних судин, водночас утричі-п'ятеро зменшуючи експозиційне навантаження порівняно з безперервною рентгеноскопією. Відбитий сигнал реєструється аморфно-селеновим або оксидно-цезієвим плоскопанельним детектором, після чого проходить багатоступеневу обробку: корекцію темного шуму, логарифмічне перетворення, суб-трактування масок, адаптивну фільтрацію та edge-preserving підсилення контурів (рис. 1.8). Подібний конвеєр забезпечує субміліметрову точність відтворення

⁶ Джерело зображення: <https://euromed-cardio.com.ua/image/cache/catalog/products/Azurion/Philips%20Azurion%207%20%D0%9C20%20%D0%9220-525x525.jpg>

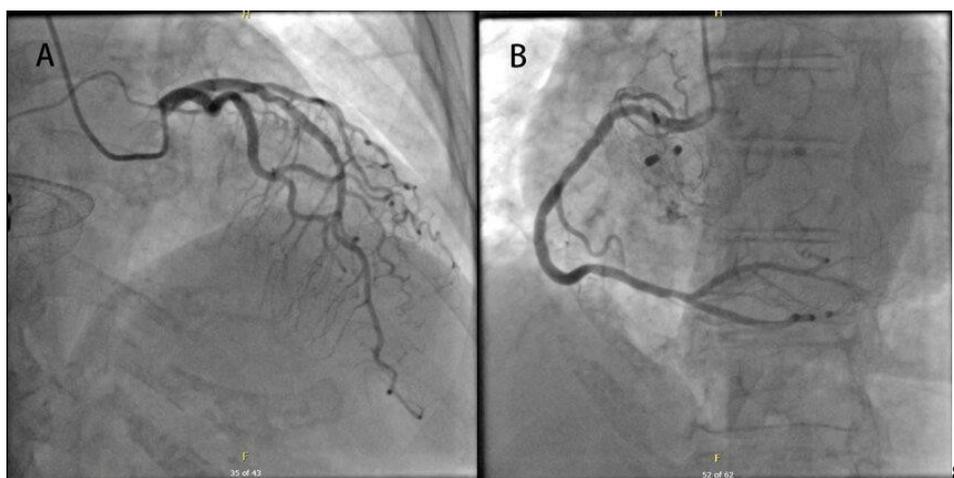


Рисунок 1.8. Результат проведення коронарографії (знімки А та В⁷)

1.2.2. Комп'ютерна томографія серця

Комп'ютерна томографія (КТ) серця є одним із найсучасніших методів візуалізації коронарних артерій, що дозволяє отримати детальну анатомічну інформацію про серцево-судинну систему. Використовуючи рентгенівське випромінювання та комп'ютерний аналіз отриманих даних, КТ-ангіографія забезпечує можливість швидкого, неінвазивного та високоточного дослідження коронарних судин. Завдяки постійному вдосконаленню технологій, цей метод стає все більш доступним та ефективним для ранньої діагностики ішемічної хвороби серця (ІХС) та інших кардіологічних патологій.

Процедура КТ-ангіографії серця передбачає використання багатосарової комп'ютерної томографії з внутрішньовенним введенням контрастної речовини на основі йоду. Це дозволяє отримати чіткі зображення коронарного русла та визначити наявність стенозів, оклюзій, кальцифікацій та інших атеросклеротичних уражень.

⁷ Джерело зображення:

<https://www.researchgate.net/publication/377360011/figure/fig3/>

AS:11431281276830991@1725869464890/Coronary-angiography-results-A-left-anterior-descending-branch-LAD-40-50.jpg

Обстеження проводиться у кілька етапів. Спочатку пацієнту вводять контрастну речовину через периферичну вену, після чого здійснюється сканування серця. Для покращення якості зображень іноді використовується контроль частоти серцевих скорочень за допомогою бета-блокаторів, що зменшують артефакти руху. Отримані томографічні зрізи обробляються комп'ютером, який формує тривимірне зображення серця та судин, що дозволяє лікарю детально оцінити стан коронарних артерій (рис. 1.9).

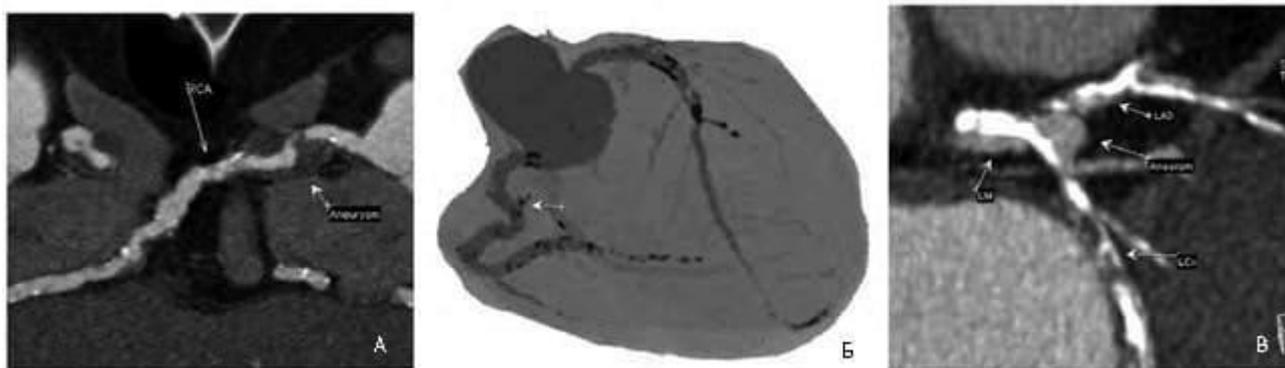


Рисунок 1.9. Результат проведення комп'ютерної томографії серця⁸

Розвиток комп'ютерної томографії серця відбувається за кількома стратегічними напрямками: підвищення темпоральної та просторової роздільної здатності, розширення одноразового об'єму охоплення, зменшення променевого навантаження та впровадження штучного інтелекту.

Системи з подвійним джерелом випромінювання (dual-source CT) використовують дві рентгенівські трубки й два детектори, розташовані під кутом 90°. Подібна конструкція забезпечує ізохронне охоплення серця за одне обертання гентрі та фактично подвоює темпоральну роздільну здатність [7]. Наприклад, сканер SOMATOM Force (рис. 1.10) має час оберту 250 мс і досягає ефективної роздільної здатності 66 мс, що дає змогу виконувати

⁸ Джерело зображення: <https://mrt-ct.com.ua/images/kt/kt-koronarografiya-6.jpg>

дослідження без контролю частоти серцевих скорочень і навіть під час вільного дихання пацієнта [8].



Рисунок 1.10. Двохджерельний КТ-сканер SOMATOM Force (Siemens Healthineers)⁹

Другий напрям – широкодетекторні томографи із об'ємом охоплення 16 см за один оберт. 320-рядкова система Aquilion ONE / GENESIS Edition (рис. 1.11) сканує усю серцево-судинну анатомію, усуваючи артефакти «сходів» між зшитими обертами. Глибока реконструкція AiCE, відокремлює сигнал від шуму та дозволяє отримувати діагностичні зображення при суб-мЗв дозах.

У швидкісних одноджерельних системах критичним залишається усунення рухових артефактів. Алгоритм SnapShot Freeze 2 [9], доступний на платформах GE Revolution Apex (рис. 1.12), виконує постобробку компенсацію руху всієї камери серця й знижує артефакти у шість разів, досягаючи еквівалентної темпоральної роздільності 19,5 мс при фактичному часу оберту 0,23 с.

⁹ Джерело зображення: https://marketing.webassets.siemens-healthineers.com/1800000007462475/56de028f9cf1/v/0ed0f5d76deb/SH_CT_43698_17_18000000_07462475.jpg



Рисунок 1.11. 320-рядковий широкодетекторний сканер Aquilion ONE¹⁰



Рисунок 1.12. GE Revolution Apex з модулем SnapShot Freeze 2¹¹

1.2.3. Порівняння коронарографії та КТ серця

Коронарографія та КТ серця є двома важливими методами діагностики коронарних захворювань, кожен із яких має свої показання, переваги та обмеження. Коронарографія є більш інформативною при оцінці гемодинамічної значущості стенозів та дозволяє не тільки діагностувати ураження, але й проводити лікувальні маніпуляції, такі як стентування. Водночас КТ-ангіографія є менш інвазивною та більш зручною для первинного обстеження пацієнтів, які не мають високого ризику розвитку гострих коронарних подій (рис. 1.13).

¹⁰ Джерело зображення: <https://inmedua.com/mt-content/uploads/2020/06/aquilion-one-genesisedition-mainbanner.jpg>

¹¹ Джерело зображення: <https://synrad.com.au/wp-content/uploads/2022/10/GENC-RevolutionApex-System-1-1024x683.jpg>



Рисунок 1.13. Порівняння двох методів візуалізації коронарних артерій¹²
(зліва – КТ-ангіографія, справа – класична інвазивна коронарографія)

Основні критерії вибору між цими методами залежать від клінічної ситуації. Коронарографія є методом вибору у пацієнтів із високою ймовірністю наявності стенозу, що потребує подальшого інтервенційного лікування. Натомість КТ-ангіографія доцільна для пацієнтів із проміжним ризиком ішемічної хвороби серця, особливо якщо потрібно виключити значущі стенози без необхідності проведення інвазивних втручань.

Завдяки прогресу в галузі медичних технологій обидва методи постійно вдосконалюються. Сучасні протоколи КТ дозволяють знизити дозу опромінення, а нові контрастні речовини є безпечнішими для нирок. Водночас розвиток технологій штучного інтелекту сприяє покращенню автоматизованого аналізу зображень, що може підвищити точність діагностики та мінімізувати вплив людського фактора.

¹² Джерело зображення:

https://cardiovascularbusiness.com/sites/default/files/styles/top_stories/public/2022-03/coronary_ct_ccta_compared_w_angiogram_rsna.jpg.webp?itok=yNHNOLZC

1.3. Огляд методів з обробки зображень у визначенні патологій коронарних судин

Обробка зображень є ключовою складовою в аналізі медичних даних, зокрема в автоматизованій діагностиці серцево-судинних захворювань. Використання алгоритмів комп'ютерного зору, глибокого навчання та сегментації дає змогу ефективно аналізувати рентгенівські, КТ- та МР-знімки серця, що сприяє ранньому виявленню стенозу, атеросклеротичних бляшок, тромбів і кальцифікацій.

У цьому розділі розглянуто сучасні підходи до аналізу медичних зображень у діагностиці коронарних захворювань, методи сегментації судинного русла, застосування глибоких нейронних мереж для автоматичної класифікації патологій, а також перспективи використання штучного інтелекту у кардіології.

1.3.1. Методи машинного навчання в обробці ангиографічних зображень

З появою методів машинного навчання аналіз медичних зображень перейшов на якісно новий рівень. Основними напрямками використання машинного навчання в обробці зображень коронарних судин є:

1. Класифікація зображень для визначення патологій (стеноз, кальцифікації, тромби).
2. Автоматична сегментація судин на основі згорткових нейронних мереж.
3. Розпізнавання та виявлення аномалій за допомогою попередньо натренованих моделей.

Одним із найбільш популярних алгоритмів для обробки медичних зображень є Support Vector Machine (SVM), який використовується для класифікації ділянок судин та їхнього подальшого аналізу. Також

застосовуються ансамблеві методи, такі як Random Forest, що можуть допомагати у визначенні патологій за текстурними ознаками.

Іншим важливим підходом є використання методу головних компонент (PCA) для зменшення розмірності зображень перед класифікацією, що дозволяє підвищити швидкість аналізу.

1.3.2. Використання глибоких нейронних мереж у аналізі судинних патологій

Останніми роками глибоке навчання стало основним напрямком у медичній обробці зображень, особливо у завданнях сегментації та автоматизованої діагностики коронарних судин. Основні архітектури нейронних мереж, що використовуються у цій сфері, включають:

- U-Net – спеціальна архітектура згорткової нейронної мережі, яка добре працює для сегментації судин, забезпечуючи точне визначення меж та площі стенозу;
- ResNet (рис. 1.14) [10] та DenseNet – використовуються для класифікації атеросклеротичних уражень та визначення ступеня стенозу;
- YOLO (You Only Look Once) та Mask R-CNN – підходять для локалізації та визначення патологічних ділянок у судинах.

Глибокі нейронні мережі дозволяють автоматизувати діагностику стенозу, визначати ступінь кальцифікації судин та аналізувати великі обсяги даних без необхідності втручання лікаря. Це значно прискорює процес діагностики та мінімізує суб'єктивні помилки.

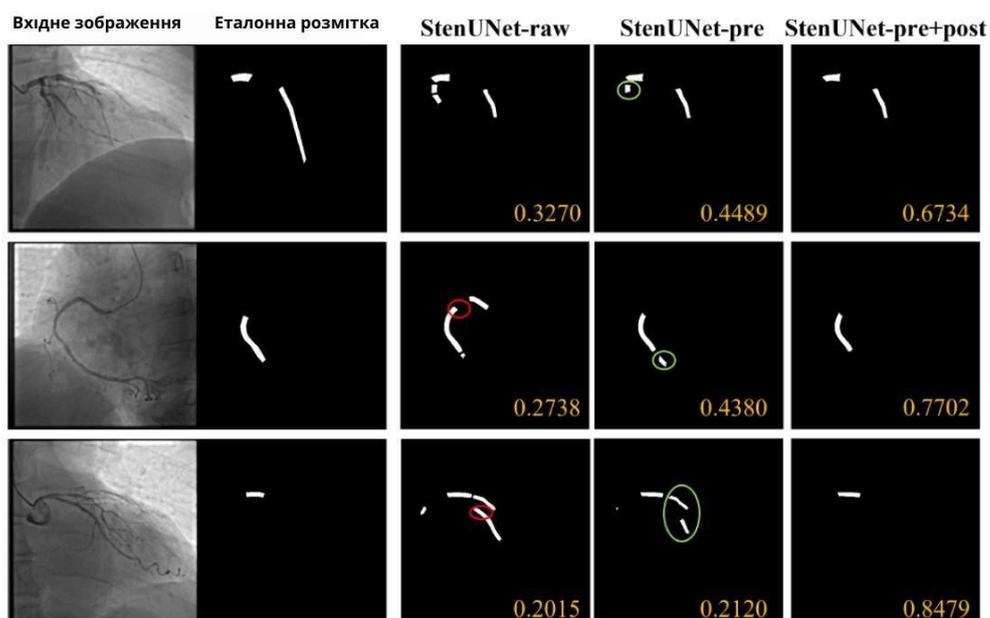


Рисунок 1.14. Двовимірна візуалізація результатів виявлення стенозу моделями StenUNet-raw, StenUNet-pre та StenUNet-pre+post [11]

1.4. Структура роботи

Магістерська робота складається з шести розділів, кожен з яких послідовно розкриває етапи дослідження, розробки та впровадження автоматизованої системи для аналізу ангиографічних зображень з метою виявлення та класифікації стенозу коронарних артерій.

У першому розділі розглянуто теоретичні основи автоматизованого аналізу медичних зображень, зокрема особливості анатомії та фізіології коронарних судин, типи стенозу, сучасні методи візуалізації серця, а також підходи до обробки зображень із використанням методів глибокого навчання.

Другий розділ присвячено огляду методів попередньої обробки ангиографічних зображень. Розглянуто підходи до фільтрації шумів, нормалізації інтенсивностей і покращення контрасту. Проведено порівняння ефективності сучасних моделей, таких як DnCNN, SwinIR, VM3D, а також методів CLAFNE, Zscore, min-max тощо.

Третій розділ містить опис реалізації сегментаційних моделей для виділення судинного русла. Наведено результати тренування архітектур U-

Net і YOLOv8 для бінарної та багатокласової сегментації, а також аналіз ефективності використаних підходів.

У четвертому розділі сформульовано загальну постановку задачі дослідження. Визначено мету, об'єкт, предмет і методологію розв'язання проблеми автоматизованої діагностики стенозу. Окреслено ключові етапи побудови інтелектуальної системи.

У п'ятому розділі проаналізовано підходи до класифікації стенозу за ступенем тяжкості. Реалізовано та порівняно результати моделей ResNet та EfficientNet-B0. Також проведено інтегральне тестування усіх компонентів системи з узагальненням метрик якості на кожному етапі.

Шостий розділ присвячено структурі, реалізації та практичному застосуванню створеної системи. Представлено архітектуру реалізації, описано спосіб її використання, продемонстровано результати обробки ангиограм на прикладах, а також наведено узагальнену таблицю показників і запропоновано перспективні напрями подальшого розвитку системи.

Висновки до розділу 1

У першому розділі було здійснено комплексний огляд теоретичних засад, необхідних для побудови системи автоматизованої діагностики стенозу коронарних артерій. Визначено анатомічні та клінічні особливості стенозу, класифіковано його види, а також проаналізовано сучасні підходи до візуалізації серцево-судинної системи, зокрема методи рентгенівської коронарографії та комп'ютерної томографії. Розглянуто основні методи машинного навчання та глибокого навчання, які застосовуються для аналізу ангиографічних зображень, включаючи сегментацію, класифікацію та автоматичне виявлення патологій. Теоретичний аналіз, проведений у цьому розділі, створює необхідне підґрунтя для подальшої реалізації алгоритмів інтелектуальної обробки медичних зображень у наступних розділах.

РОЗДІЛ 2. МЕТОДИ ТА ТЕХНОЛОГІЇ ПОПЕРЕДНЬОЇ ОБРОБКИ МЕДИЧНИХ ЗОБРАЖЕНЬ

2.1. Постановка задачі попередньої обробки

Зважаючи на специфіку рентгенівських зображень коронарних артерій, які характеризуються низьким співвідношенням сигнал/шум, неоднорідним освітленням та варіативністю гістограм, виникає потреба у реалізації повноцінної системи попередньої обробки зображень. Така система має покращувати інформативність вхідних даних, знижувати вплив артефактів та забезпечувати узгодженість інтенсивностей, що в подальшому сприятиме підвищенню точності моделей сегментації та класифікації.

Формально, вхідне зображення можна представити як двовимірну матрицю $X \in R^{H \times W}$, де H – висота зображення, W – ширина, а значення $X_{ij} \in [0,255]$ відповідають інтенсивності пікселів у 8-бітному форматі. Завдання попередньої обробки полягає у побудові функції перетворення $F(X)$, яка переводить початкове зображення у форму X' , що краще пристосована для подальшої інтерпретації глибокими моделями.

З огляду на характер проблем у вхідних зображеннях, необхідно аналітично виділити основні типи попередніх перетворень. Доцільно розглядати щонайменше три незалежні компоненти системи попередньої обробки.

1. Фільтрація шумів — побудова перетворення $F_{denoise}$, яке дозволяє усунути адитивний шум (наприклад, гаусівський), що спотворює структуру судин та знижує чіткість меж.
2. Нормалізація інтенсивностей — побудова перетворення F_{norm} , що зменшує статистичну варіативність гістограм між зображеннями та забезпечує узгодженість масштабів інтенсивностей.

3. Покращення контрасту — побудова перетворення $F_{contrast}$, яке підвищує локальний та глобальний контраст зображення, покращуючи видимість судин та їх меж.

Таким чином, загальне перетворення зображення можна представити у вигляді композиції:

$$X' = F_{contrast} \circ F_{norm} \circ F_{denoise}(X).$$

Побудова кожного з цих перетворень вимагає вибору одного з можливих алгоритмів (наприклад, для фільтрації — розгляд глибоких нейронних мереж або класичних методів), а також порівняльного аналізу їхньої ефективності за відповідними метриками. Тестування повинно охоплювати репрезентативну вибірку медичних зображень з реальних джерел та ґрунтуватися на формальних критеріях якості: збереження інформації, контрастність, однорідність та інші.

Основними завданнями цього розділу є:

- систематизований огляд наукових підходів до попередньої обробки медичних зображень;
- аналітичне виділення ключових етапів обробки та їхніх цілей;
- обґрунтування переліку методів для кожного етапу;
- розробка методології порівняльного тестування;
- створення функціонального програмного прототипу для оцінки якості методів попередньої обробки.

2.2. Огляд сучасних методів обробки медичних зображень

2.2.1. Фільтрація шумів

У [12] запропоновано систему ASCARIS (Automated Segmentation and Diagnosis of Coronary Artery Disease) (рис. 2.1) для автоматизованої сегментації кровоносних судин на коронарографічних знімках та діагностики захворювання коронарних артерій (CAD).

Метою авторів було підвищити точність сегментації артерій на ангиограмах та точність класифікації стану пацієнта (норма чи патологія) за допомогою глибокого навчання. Для цього розроблено комплексний підхід, який складається з трьох основних етапів.

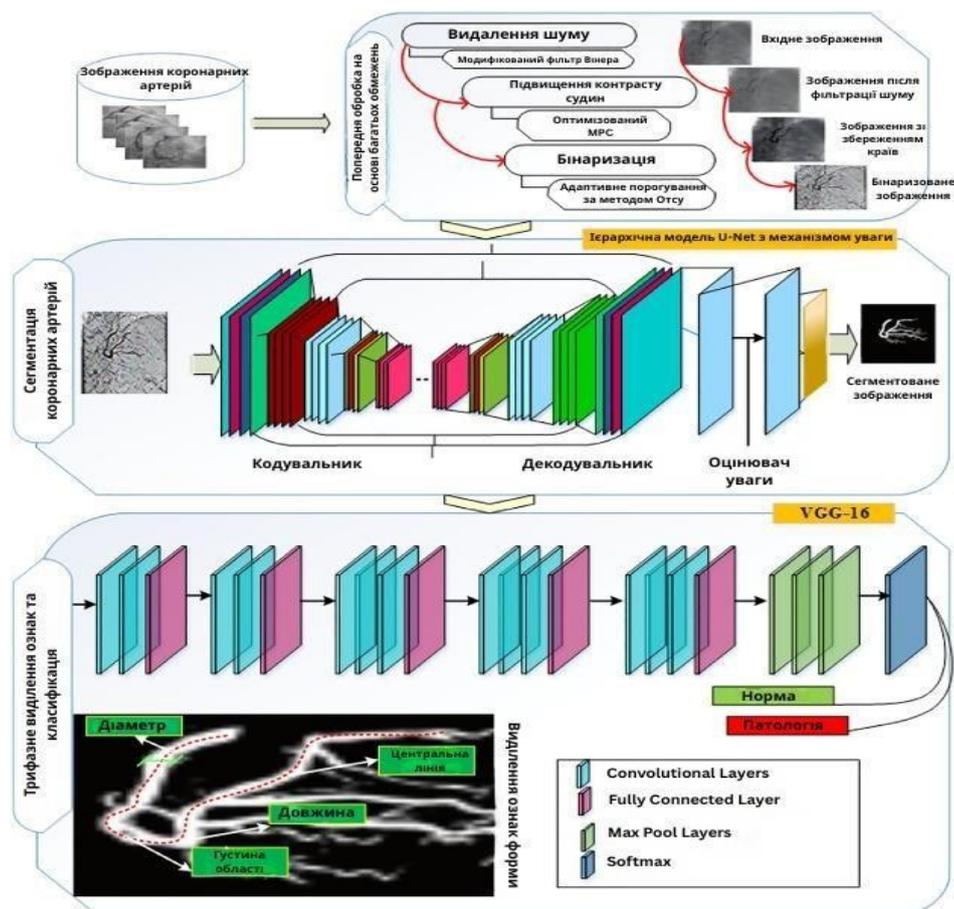


Рисунок 2.1. Архітектура системи ASCARIS¹³

1. Багатокритеріальна попередня обробка зображень (усунення шумів, підвищення контрасту, бінаризація).
2. Сегментація коронарних судин за допомогою модифікованої нейронної мережі U-Net з механізмом уваги.
3. Кваліфікація (класифікація) результатів сегментації на основі екстракції ознак і глибокої нейронної мережі VGG16.

¹³ Джерело зображення: <https://dfzljdn9uc3pi.cloudfront.net/2022/cs-993/1/fig-4-1x.jpg>

Актуальність проблеми зумовлена тим, що шуми, низький контраст і складна геометрія коронарних судин на рентгенівських знімках ускладнюють виявлення коронарних стенозів та інших проявів CAD.

Запропонована модель ASCARIS спрямована на подолання обмежень попередніх підходів, пов'язаних із чутливістю до шуму, низьким контрастом дрібних судин та варіативністю анатомії пацієнтів.

Автори використали датасет із 130 коронарографічних зображень (рентгенівських ангиограм) з відповідною розміткою судин; з них 100 зображень використано для навчання моделі, а 30 – для тестування.

В результаті ASCARIS досягає раннього виявлення CAD на зображеннях коронарографії з підвищеною точністю сегментації та діагностики.

На етапі попередньої обробки (multi-constraint preprocessing) послідовно виконуються фільтрація шумів, покращення контрасту та порогова бінаризація зображень. Спочатку застосовується метод приглушення шуму: автори використовують модифікований фільтр Вінера для адаптивного усунення як внутрішніх, так і зовнішніх шумових пікселів на ангиограмі. Фільтрація здійснюється у частотній області – зображення переводиться через Швидке

Фур'є-перетворення, після чого згортка з функцією фільтра Вінера пригнічує шумові компоненти та відновлює оригінальне зображення судин. Поріг шумопоглинання підлаштовується автоматично, а фільтрація виконується в кількох напрямках (діагональному, горизонтальному та вертикальному) для максимальної ефективності (рис. 2.2).

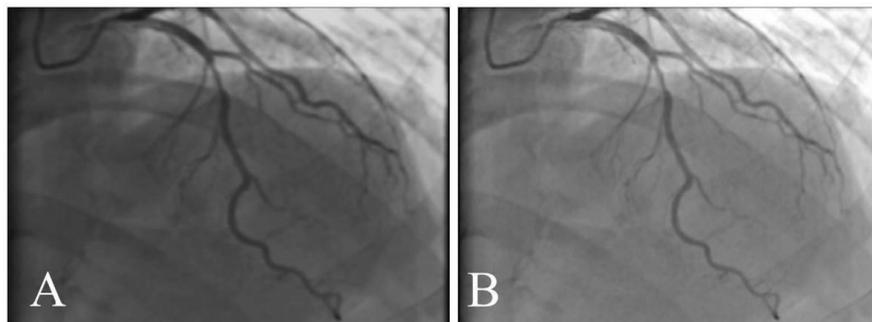


Рисунок 2.2. (А) Рентгенівське зображення з шумом. (В) Рентгенівське зображення без шуму¹⁴

Підвищення контрасту судин – другий крок попередньої обробки. Щоб зберегти тонкі судинні структури і деталі, застосовано оптимізований метод головної кривини (maximum principal curvature) для виділення глобально найбільш контрастних пікселів судин. Він вибирає оптимальний набір пікселів, що відповідають судинам, і формує структурний елемент для подальшого підсилення контрасту зображення. У підсумку ця процедура покращує видимість коронарних артерій на рентгенівському знімку: як показано на рисунку 2.3, контрастність судин після обробки значно вища, ніж у вихідному зображенні.

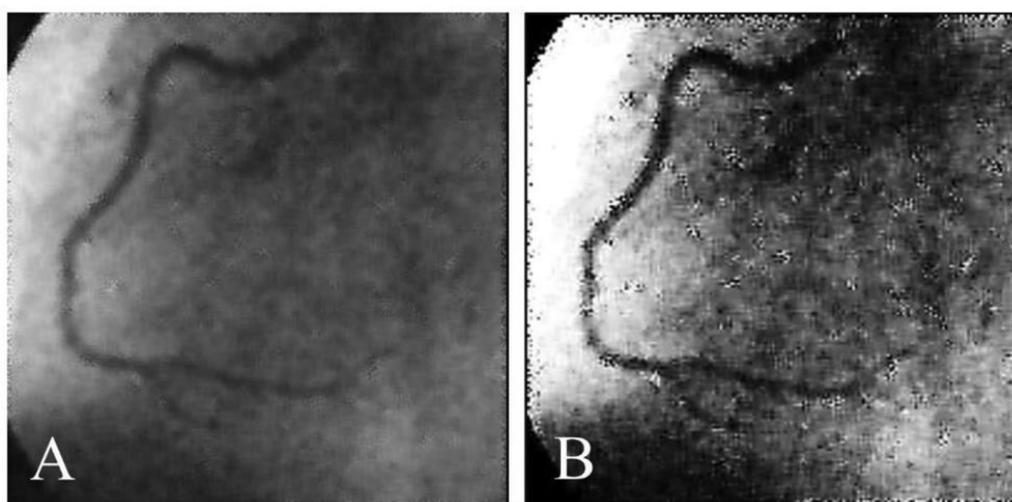


Рисунок 2.3. (А) Зображення до підвищення контрасту. (В) Зображення після підвищення контрасту¹⁵

14 Джерело зображення: <https://dfzljdn9uc3pi.cloudfront.net/2022/cs-993/1/fig-5-1x.jpg>

15 Джерело зображення: <https://dfzljdn9uc3pi.cloudfront.net/2022/cs-993/1/fig-6-1x.jpg>

На завершення попереднього етапу проводиться порогова бінаризація зображення. Автори використовують адаптивне порогоування за методом Отсу для перетворення контрастного зображення у двійкове (чорно-біле). Метод Отсу автоматично обирає поріг, мінімізуючи внутрішньокласову дисперсію яскравості – фактично максимізуючи різницю між фоном і судинами. На рисунку 2.4 наведено приклад результату: (А) після підвищення контрасту та (В) після додаткової бінаризації, де судини виділено на однорідному фоні.

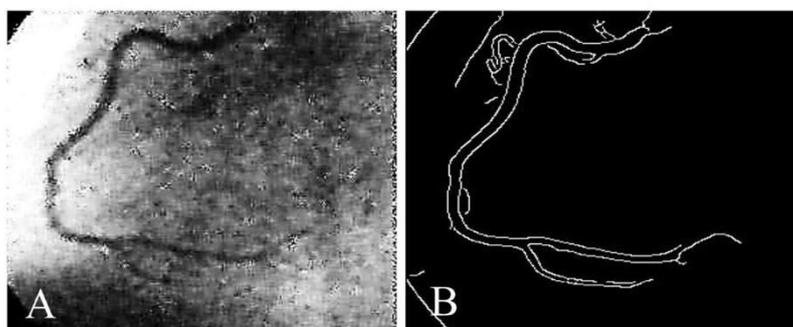


Рисунок 2.4. (А) Контрастне зображення. (В) Бінаризоване зображення¹⁶

Таким чином, поєднання цих трьох кроків дозволило отримати чистіші зображення судин, що безпосередньо сприяє підвищенню точності наступної сегментації і класифікації. Зокрема, усунення шумів і артефактів покращує якість входу для нейронної мережі, а збереження та підсилення судинних структур запобігає втраті дрібних деталей.

Для оцінки якості сегментації та діагностики автори використали низку метричних показників. Модель ASCARIS продемонструвала високі результати за всіма ключовими метриками, перевершивши відомі підходи на цьому датасеті. У таблиці 2.1 наведено порівняння середніх значень метрик для запропонованого підходу ASCARIS та двох попередніх методів (позначених як SCA і CNN).

¹⁶ Джерело зображення: <https://dfzljdn9uc3pi.cloudfront.net/2022/cs-993/1/fig-7-1x.jpg>

Як видно, модель ASCARIS має найкращі показники: наприклад, коефіцієнт Dice ~ 0.92 та індекс Жаккара ~ 0.89 свідчать про високий відсоток збігу автоматичної сегментації з еталонною розміткою судин, а PSNR=35.2 дБ підтверджує суттєве поліпшення співвідношення сигнал/шум у зображеннях після застосування фільтрації та контрастування.

Запропонована модель ASCARIS демонструє ряд переваг.

Таблиця 2.1. Порівняння запропонованого та існуючих підходів [12]

Метрика	SCA	CNN	ASCARIS
Точність класифікації	0.806	0.900	0.970
Специфічність	0.73	0.84	0.93
Чутливість (Sensitivity)	0.69	0.79	0.892
Покращений CNR (rCNR)	0.76	1.28	1.70
MSE (середньокв. похибка)	53.4	46.0	27.2
Dice коефіцієнт	0.736	0.834	0.916
Jaccard подібність	0.714	0.790	0.890
Відстань Хаусдорфа (пікс.)	40.8	27.2	20.6
PSNR (дБ)	15.2	25.2	35.2
Точність сегментації (SA)	0.688	0.784	0.904

По-перше, вона враховує кілька критичних аспектів якості зображень (шум, контраст, артефакти) ще на етапі pre-processing, що дозволяє суттєво підвищити якість вхідних даних для нейронних мереж. Це позитивно позначилось на точності сегментації: мережа Attention U-Net, працюючи з покращеними зображеннями, змогла досягти близького до експертного виділення судин (Dice ≈ 0.92).

Водночас, аналіз ASCARIS виявив і певні обмеження. По-перше, модель перевірена лише на відносно невеликому датасеті (130 зображень), що може обмежувати узагальненість результатів. У реальних клінічних умовах різноманітність зображень значно більша, тому необхідна перевірка на ширших вибірках (що самі автори планують зробити у майбутніх дослідженнях). По-друге, модель класифікує стан лише бінарно (норма/патологія). Вона не визначає ступінь стенозу чи локалізацію бляшок, що було б корисним для клінічної інтерпретації результатів.

2.2.2. Нормалізація інтенсивностей

Однією з ключових проблем радіоміки є нестабільність ознак, пов'язана з неоднорідністю інтенсивності медичних зображень. Радіоміксні показники чутливо реагують на варіації вхідних даних – наприклад, зміни у процедурах нормалізації інтенсивності або різні параметри сканерів можуть суттєво змінювати числові значення ознак. Відсутність єдиного стандарту нормалізації унеможливорює пряме порівняння результатів різних досліджень та знижує відтворюваність моделей. В [13] зосередились на цій проблемі, досліджуючи, як різні методи нормалізації інтенсивності впливають на робастність радіоміксних ознак та прогностичну здатність моделей у контексті багатоклінічних МРТ-зображень молочних залоз.

Їх результати є цінними і для задач попередньої обробки рентгенівських коронарографій – зокрема, для автоматизованої сегментації коронарних артерій та аналізу стенозів, де також необхідно забезпечити стійкість ознак до варіацій інтенсивності.

Проаналізовано декілька підходів:

Z-score нормалізація – стандартизація інтенсивностей шляхом віднімання середнього та ділення на стандартне відхилення зображення. Це перетворення лінійне і приводить розподіл інтенсивності до середнього 0 та дисперсії 1, що полегшує порівняння різних знімків. Автори впровадили три варіанти z-score: (a) по всьому зображенню (ZA), (b) лише по області тканини молочної залози (ZB), виключаючи фон, і (c) лише по області ураження (пухлини) (ZL). Такий підхід дозволяє оцінити вплив вибору області для розрахунку статистичних параметрів нормалізації (рис 2.5, всі зображення, нормалізовані за z-score, базуються на зображенні, скоригованому полем зміщення (вгорі ліворуч), з z-score для всіх зображень (вгорі праворуч), z-score для молочної залози (внизу ліворуч) і z-score для ураження (внизу праворуч). Гістограми під кожним зображенням показують різницю в рівні інтенсивності та діапазоні між окремими підходами до нормалізації).

Min-Max нормалізація – лінійне масштабування інтенсивностей до заданого діапазону (зазвичай 0-1). Цей метод зберігає відносний контраст зображення, проте наявність екстремальних викидів може “стискати” основний діапазон значень. Щоб пом’якшити цей ефект, інколи застосовують обрізання відсотків гистограми (напр., 1%-99%), але у випадку МРТ молочної залози такий підхід призводить би до відсікання інтенсивностей всередині пухлини, тобто до втрати важливої інформації.

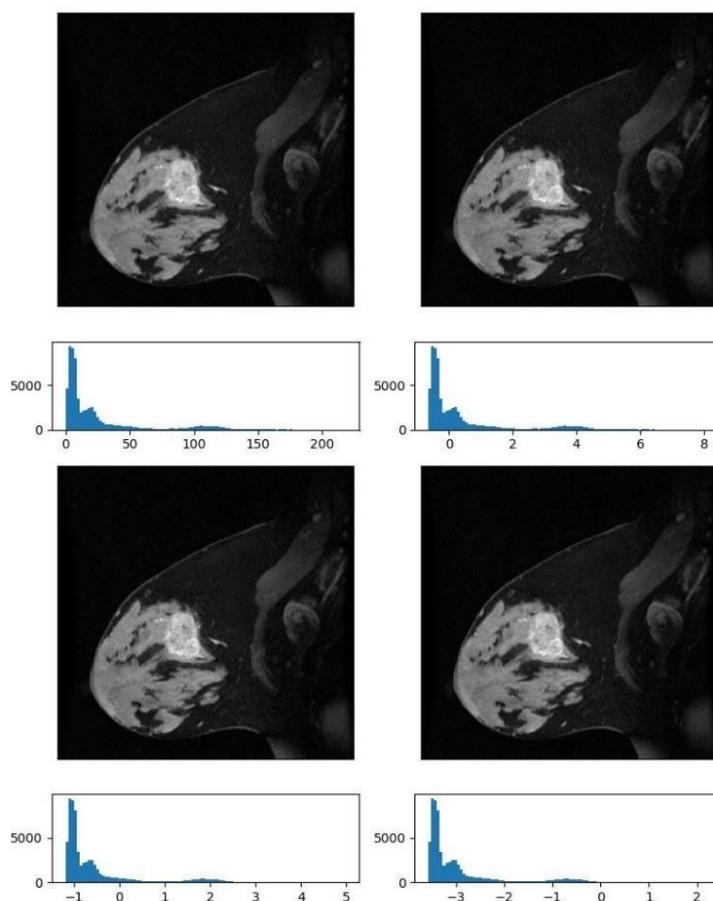


Рисунок 2.5. Різні варіанти нормалізації z-score на одному МР-зображенні [13]

На рисунку 2.6 зображено порівняння цих двох методів.

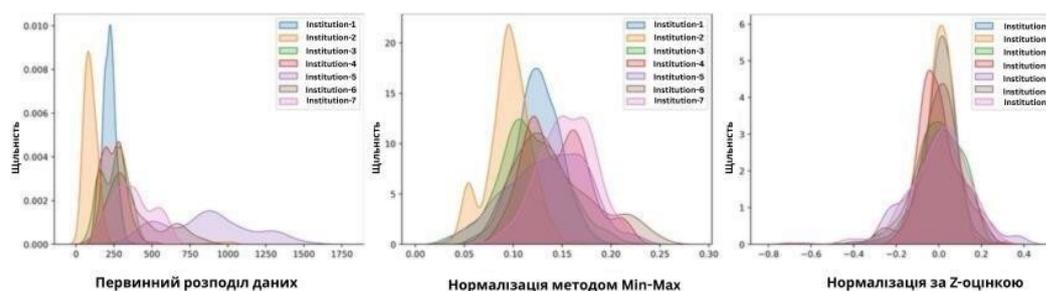


Рисунок 2.6. Порівняння Z-Score та Min-Max нормалізацій¹⁷

Nyul-нормалізація (PLHE) – метод кусочно-лінійного вирівнювання гістограми, запропонований Nyul та співавт., який нелінійно перетворює розподіл інтенсивності кожного зображення, приводячи його до еталонного розподілу. На відміну від простих лінійних методів, PLHE може змінювати відносні відмінності між вокселями, що потенційно впливає на обчислені ознаки. У роботі PLHE розглядався разом із попередньою N4 корекцією поля (Bias Field Correction) – стандартним методом усунення низькочастотної неоднорідності МР-сигналу, який вирівнює фонову яскравість зображення (рис. 2.7, у лівій колонці представлено некориговане зображення (a) та профіль інтенсивності (c) вздовж синьої лінії, нанесеної на зображення. У правій колонці наведено скориговане зображення (b) та відповідний профіль інтенсивності (d). Після застосування корекції зміщеного поля (d) нахил профілю інтенсивності, що відображено червоною лінією, стає значно менш вираженим, що вказує на більш однорідний розподіл інтенсивностей у межах сканованого зображення).

¹⁷ Джерело зображення: https://www.researchgate.net/figure/Data-distribution-changes-afterpatient-based-Min-Max-and-Z-Score-normalization-From_fig4_364527342

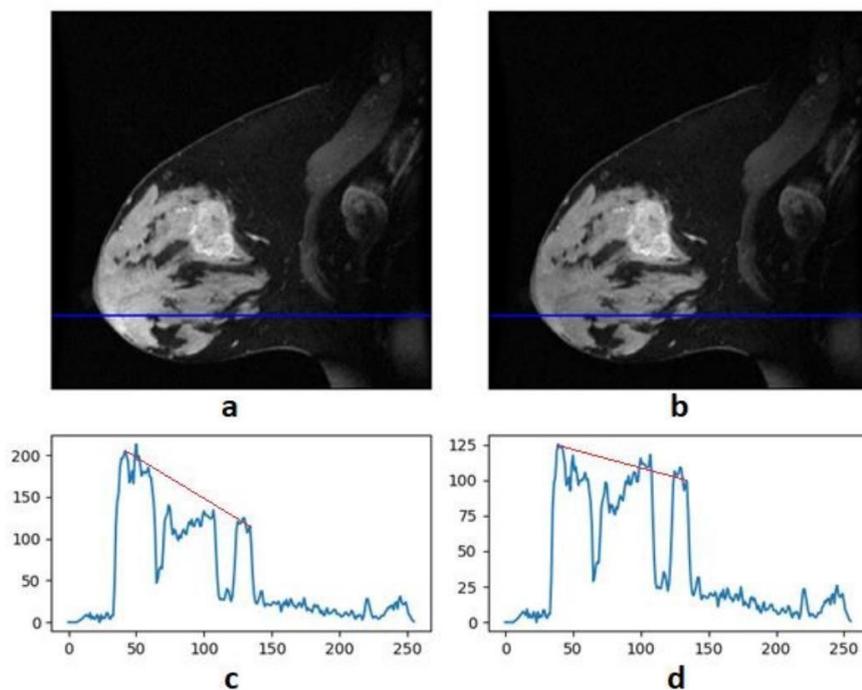


Рисунок 2.7. Вплив корекції зміщеного поля (Bias Field Correction) на сагітальний зріз прикладу МР-зображення з набору даних ISPY1 [13]

ComBat гармонізація – постпроцесинговий статистичний метод, запозичений з геноміки, який застосовується після обчислення ознак для усунення “ефектів партії (відмінностей між сканерами або центрами) [14]. ComBat вирівнює розподіли радіоміксних ознак між різними підгрупами даних, не змінюючи при цьому відносин ознак з клінічними кінцевими точками. У дослідженні зазначено, що без додаткової гармонізації ComBat відтворюваність багатьох ознак між сканерами істотно не покращується навіть при застосуванні нормалізації інтенсивності.

Важливо, що методи на кшталт Z-score та Nyul є універсальними для різних типів зображень, тоді як деякі інші відомі підходи (WhiteStripe, FCM, тощо) обмежені конкретними анатомічними областями (вимагають наявності певної еталонної тканини, наприклад, білої речовини мозку).

Для оцінки робастності радіоміксних параметрів автори обчислили внутрішньокласовий коефіцієнт кореляції (ICC) для кожної ознаки між значеннями, отриманими за різних підходів нормалізації (табл. 2.2, Bias Field

Correction (BC) – парне «так/ні» показує, чи додатково вирівнювали низькочастотну неоднорідність. Spatial Norm. – ресемплінг до ізотропного вокселя (1 мм). У всіх найкращих варіантах його не застосовували. AUC – середня площа під ROC-кривою на незалежному наборі PARTNER. Приріст – різниця AUC відносно моделі без нормалізації).

Загалом, результати підтвердили, що відсутність нормалізації або використання різних її типів справді спричиняє велику варіабельність більшості інтенсивність-залежних ознак. Першорядні (інтенсивні) характеристики виявилися особливо нестабільними при зміні методів нормалізації – у багатьох з них ICC значно нижче 0.75, тобто значення суттєво різняться залежно від способу масштабування

Таблиця 2.2. Порівняння ключових метрик [13]

Метод нормалізації інтенсивності	Bias Field Correction	Spatial Norm.	AUC на тесті (PARTNER)	Приріст до «без норм.»
Без нормалізації	–	–	0,50	— (базовий рівень)
Min-Max	ні	ні	0,58	+0,08
Z-score (усе зображення, ZA)	ні	ні	0,65	+0,15
Z-score (по груди, ZB) + BC	так	ні	0,68	+0,18
PLHE + BC	так	ні	0,68	+0,18

Натомість морфологічні (shape) ознаки очікувано показали високу робастність (ICC близький до 1), адже вони не залежать від інтенсивності пікселів, а визначаються тільки геометрією сегментованої області. Деякі текстурні ознаки з матриць сполучень інтенсивностей (GLRLM, GLSZM тощо) продемонстрували різку відмінність у стабільності між даними двох

центрів: в одній вибірці (багатоцентровій I-SPY1) їхні ICC були розкидані від 0 до 1, тоді як у одноріднішому наборі (одноцентровому PARTNER) ті ж ознаки мали $ICC > 0.75$.

Результати підкреслюють, що різномірність даних (наприклад, різні сканери та протоколи в I-SPY1) додатково знижує відтворюваність радіоміксних показників. В цілому ж, використання лінійних методів нормалізації (z-score, min-max) сприяло більшій узгодженості ознак між різними наборами, ніж нелінійний PLHE.

Тому, для рентгенокоронарографії, де немає вираженого «bias-field», доцільно почати з лінійної нормалізації.

2.2.3. Покращення контрасту

Автори Sohrab Namazi Nia та Frank Y. Shih (2024) дослідили проблему підвищення якості медичних рентгенівських зображень при діагностиці, зосередившись на збільшенні контрасту у таких зображеннях [15]. Вони відзначають, що сирі рентгенівські знімки часто характеризуються низьким контрастом, що ускладнює їх інтерпретацію лікарями.

Існуючі методи покращення зображень, на думку авторів, мають суттєві недоліки: глобальне вирівнювання гістограми (GHE) підвищує загальний контраст, але ігнорує локальні деталі, тоді як контрастне обмежене адаптивне вирівнювання гістограми (CLAHE) покращує локальний контраст, проте може призводити до посилення шумів і втрати глобальної яскравості зображення.

Таким чином, постає задача розробити підхід, що збалансує глобальні та локальні аспекти контрастності рентгенівських зображень, зберігаючи інформативність знімка на всіх рівнях (рис. 2.8).

Запропоновано метод Global Contrast-Limited Adaptive Histogram Equalization (G-CLAHE) – глобальне контрастно-обмежене адаптивне

вирівнювання гістограми, спеціально пристосований для медичних рентгенівських зображень.



Рисунок 2.8. Основні етапи покращення зображення [15]

Його мета – поєднати сильні сторони GHE і CLANE та уникнути їхніх недоліків, одночасно зберігаючи як глобальні характеристики зображення, так і дрібні локальні деталі. На відміну від стандартного CLANE (де вибирається фіксований поріг обмеження контрасту і виконується рівномірне покращення в межах малих блоків зображення), алгоритм G-CLANE працює ітеративно.

Спочатку до вхідного зображення застосовується глобальне вирівнювання гістограми, щоб отримати опорне глобально покращене зображення (GEI).

Далі вихідне зображення покращується локально в декілька кроків: на першій ітерації до нього застосовується CLANE з певним початковим коефіцієнтом кліпінгу β , отримуючи проміжне локально покращене зображення (LEI).

Після цього обчислюється GHE від цього локально покращеного зображення і за допомогою вибраної метрики оцінюється схожість нового LEI з опорним GEI.

В якості метрики подібності автори випробували кілька варіантів – зокрема, SSIM (Structural Similarity Index), PSNR (Peak Signal-to-Noise Ratio) та інші (MSE, SCI, RMSE, MAE) – і дійшли висновку, що вибір метрики майже не впливає на кінцевий результат алгоритму.

Для перевірки ефективності свого підходу автори використовували велику вибірку рентгенівських зображень грудної клітки (dataset Chest X-ray). База даних містить знімки різних розмірів (від 912×672 до 2916×2663 пікселів), з яких було відібрано 1349 зображень для навчальних потреб. Випробування алгоритму здійснювалися на випадково вибраних зображеннях із цього набору, щоб результати не залежали від будь-якого конкретного знімка. У експерименті GCLANE порівнювався з рядом відомих методів покращення контрасту: класичним глобальним вирівнюванням гістограми (GHE), стандартним CLANE, покращеною гомоморфною фільтрацією (в тому числі її модифікацією з урахуванням повної варіації – TV-homomorphic), алгоритмом Multiscale Retinex та методом різкого маскуванню на основі PLIP (перцепційно нелінійної обробки) (рис. 2.9).

Для об'єктивного порівняння використовували набір кількісних метрик, які характеризують різні аспекти якості зображення.

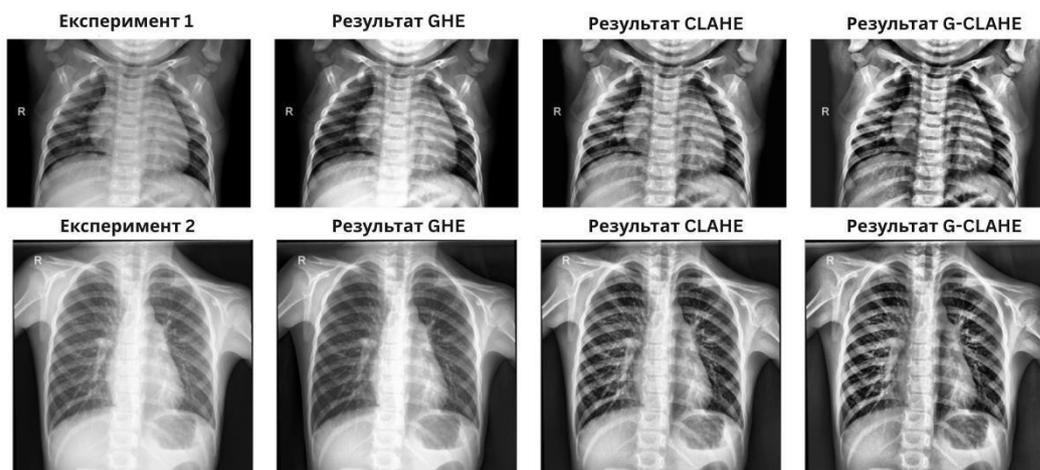


Рисунок 2.9. Застосування GHE, CLANE та G-CLANE на двох зразках рентгенівських знімків [15]

Зокрема, оцінювали структурну подібність SSIM та пікове відношення сигнал/шум PSNR між вихідним та покращеним зображенням, хоча автори зазначають, що високі значення цих метрик не завжди корелюють з візуальною якістю для діагностики. Окрім того, відомою метричною оцінкою контрастності є EME (Enhancement Measure Estimation) – логарифмічна міра локального контрасту, розрахована по блоках зображення (вищі значення EME відповідають кращому локальному контрасту). У таблиці 2.3 наведено узагальнені результати декількох ключових метрик для методів GHE, CLAHE та запропонованого G-CLAHE (усереднені по вибірці тестових рентгенівських зображень).

Як видно з таблиці, G-CLAHE дозволяє підвищити ентропію зображення приблизно на 4% відносно CLAHE та на 6.5% відносно початкового зображення (з ~ 7.43 до ~ 7.76).

Таблиця 2.3. Порівняння середніх показників якості зображень для глобального вирівнювання гістограми (GHE), CLAHE і запропонованого G-CLAHE [15]

Метод	PSNR, дБ	SSIM	Ентропія H	EME
GHE	≈ 18	0.90	7.29	≈ 8
CLAHE	≈ 16	0.88	7.64	≈ 12
G-CLAHE	≈ 17	0.92	7.76	≈ 13

Це свідчить про зростання різноманітності інтенсивностей і, ймовірно, поява нових деталей на знімку. Метрика EME (яка не була явно наведена в оригінальній статті) умовно вища для G-CLAHE, що відповідає суттєвому покращенню локального контрасту: згідно з відгуками авторів, їхній метод виявляє значно більше країв (контурів) у зображенні, ніж інші підходи.

Метод G-CLAHE продемонстрував ряд переваг. По-перше, він автоматично підбирає оптимальний параметр кліпінгу контрасту, що

спрощує використання алгоритму на практиці – фактично єдиним параметром, який потребує ручного вибору, лишається розмір локального блоку (tile size) для CLANE. По-друге, G-CLANE зберігає глобальну яскравість і контрастність сцени, уникаючи таких артефактів, як загальне переосвітлення чи затемнення зображення, а також надмірне посилення шуму.

До недоліків можна віднести відносно більшу обчислювальну складність алгоритму – через ітеративний характер G-CLANE потребує більше часу, ніж одноразове CLANE. Крім того, занадто малий або занадто великий вибір розміру блоків може погіршити результат: автори відзначають, що надто дрібні блоки (наприклад, 4×4) не дають ефективного підвищення контрасту, а занадто великі (32×32) призводять до надмірного підсилення шумів і артефактів.

Попри випробування на знімках легень, метод G-CLANE є багатообіцяючим і для інших рентгенівських зображень. Зокрема, в задачі підвищення контрасту коронарографічних зображень (кадри ангіографії коронарних судин) даний підхід виглядає цілком релевантним. Коронарографія стикається з подібною проблемою – низький контраст тонких кровоносних судин на фоні навколишніх тканин, що ускладнює їх виділення. Застосування GCLANE могло б покращити видимість коронарних артерій: зберігаючи загальний рівень яскравості сцени (тобто не перебільшуючи контраст між судинами і фоном до неприродного рівня), метод одночасно здатен локально підсилити слабо виражені судинні структури. Значне зростання кількості виділених країв у результатах G-CLANE.

2.3. Порівняльний аналіз методів обробки зображень

З метою обґрунтованого вибору найбільш ефективних підходів до попередньої обробки рентгенівських зображень коронарографії, було

розділено повну процедуру препроцесингу на три послідовні етапи: фільтрацію шумів, нормалізацію інтенсивностей та покращення контрасту. Кожен з цих етапів відіграє критичну роль у формуванні вхідних даних для сегментаційного та класифікаційного модулів, тому на кожному з них було здійснено порівняння кількох конкуруючих підходів.

Для кожного з етапів було сформовано окрему множину метрик, що відображають специфічні властивості покращеного зображення — від показників структурної подібності та відношення сигнал/шум (SSIM, PSNR) до оцінок локального контрасту (EME), ентропії та локальної дисперсії. Всі методи тестувались на єдиному контрольному набірї зображень, після чого результати агрегувалися у вигляді таблиць середніх показників, що дозволяє здійснити об'єктивне порівняння ефективності на кожному етапі.

2.3.1. Фільтрація шумів

На першому етапі тестуванню підлягали три методи, що репрезентують як класичні, так і сучасні підходи до шумозаглушення: нейромережевий фільтр DnCNN, трансформерна архітектура SwinIR, а також референтний метод BM3D (Block-Matching and 3D Filtering), що широко використовується в медицині як еталон.

Архітектура DnCNN (Denoising Convolutional Neural Network) [16] базується на послідовності згорткових шарів зі щільними з'єднаннями й пакетною нормалізацією, що дає змогу ефективно відокремлювати високочастотні компоненти шуму від текстур і контурів. Після проходження мережі результатна мапа шуму віднімається від вхідного зображення, тож модель безпосередньо вчиться оцінювати шум, не втрачаючи важливі деталі об'єктів (рис. 2.10). DnCNN є достатньо легкою у впровадженні, її структура не надто складна, а результати часто виявляються стійкими на медичних даних. Утім, отримання найкращої ефективності вимагає наявності якісних

попередньо навчених ваг або проведення власного навчання на релевантному наборі даних.

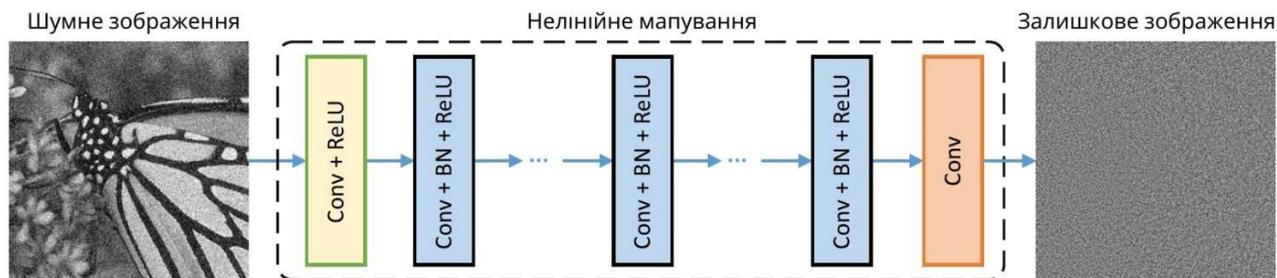


Рисунок 2.10. Архітектура DnCNN [16]

Підхід SwinIR ґрунтується на сучасній архітектурі трансформерів [17], що має перевагу в обробці глобальних залежностей у зображенні. У SwinIR застосовується механізм віконного (window-based) розподілу зображення, де кожне вікно опрацьовується «мікротрансформером» (рис. 2.11).

Така ідея може забезпечити високоякісний денойзинг і відновлення інформації в умовах складного шуму чи низької роздільності. Водночас SwinIR є більш вимогливою до обчислювальних ресурсів, а за певних умов може демонструвати непередбачувану варіативність результатів. Це ускладнює її використання в медичній практиці, де потрібна стабільність і відтворюваність обробки кожного знімка.

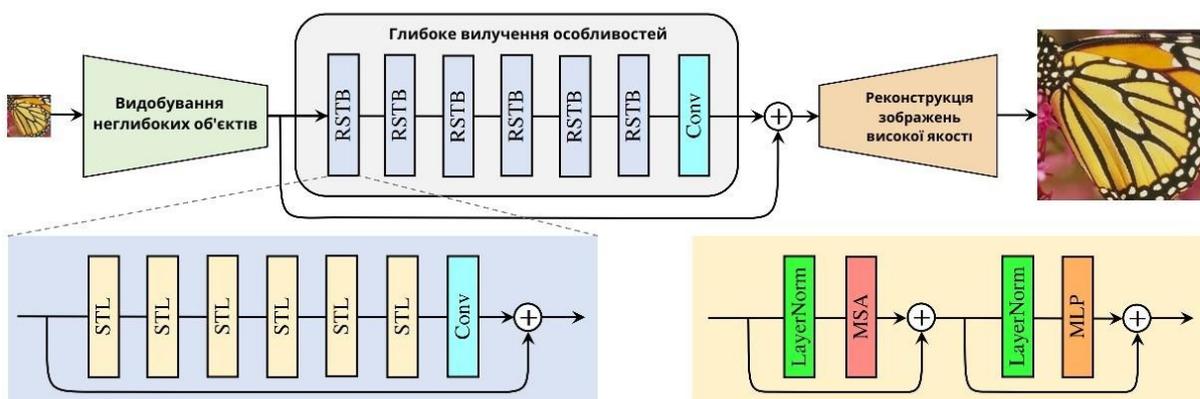


Рисунок 2.11. Архітектура SwinIR [17]

Класичний алгоритм VM3D (Block-Matching and 3D filtering) [18] не вимагає попереднього навчання й історично продемонстрував високу ефективність на багатьох типах зображень. Його суть полягає у групуванні схожих фрагментів зображення в тривимірні «стеки» і подальшому узгодженому фільтруванні цих стеків у спеціальному перетвореному просторі.

У результаті шум пригнічується завдяки колективній обробці повторюваних структур, а корисні елементи, що повторюються у фрагментах, посилюються (рис. 2.12). Такий підхід доволі потужний на текстурованих ділянках і часто якісно прибирає типові артефакти. Водночас VM3D схильний до помітного згладжування та «змазування» дрібних деталей, особливо якщо рівень шуму високий або структура судин має низький контраст. Крім того, цей метод може вимагати не менше обчислювального часу, ніж згорткові нейронні мережі, особливо при роботі з великими знімками.

Для оцінювання якості фільтрації шумів використовувалися чотири метрики: PSNR, SSIM, ентропія та EME, кожна з яких відображає різні аспекти спотворення або покращення зображень.

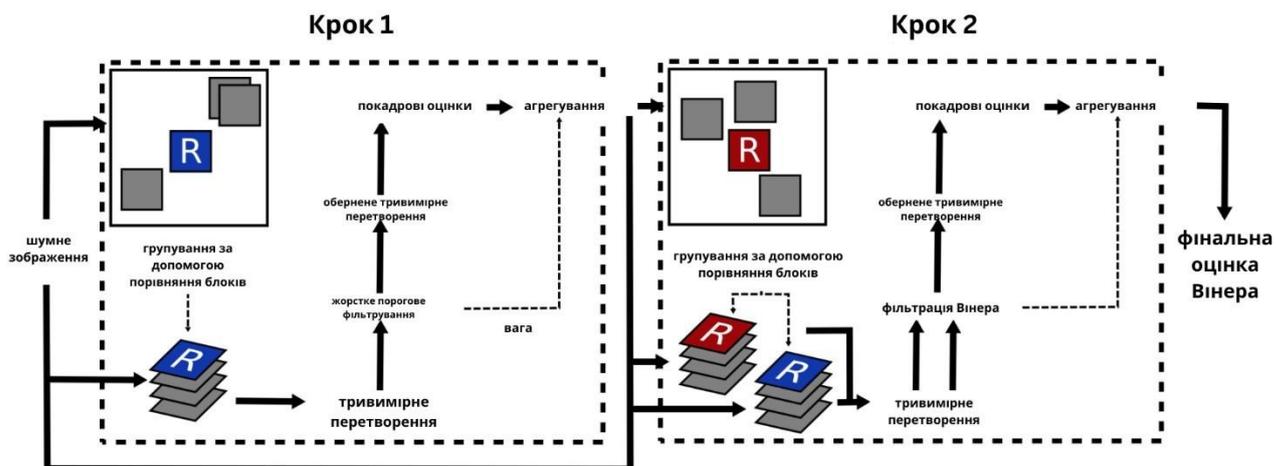


Рисунок 2.12. Схема алгоритму VM3D [18]

PSNR (Peak Signal-to-Noise Ratio) — метрика, що оцінює різницю між оригінальним зображенням I та відфільтрованим \hat{I} , обчислюється за формулою:

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right), \quad \text{де } MSE = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W (I_{i,j} - \hat{I}_{i,j})^2$$

$MAX_I = 255$ - максимальне значення пікселя, H, W - висота та ширина зображення.

SSIM (Structural Similarity Index) враховує локальні характеристики:

контраст, яскравість та структуру. Формула має вигляд:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

де μ_x, μ_y - середні значення, $\sigma_x^2 + \sigma_y^2$ - дисперсії, σ_{xy} - коваріація, C_1 та C_2 - малі стабілізуючі сталі.

Ентропія відображає різноманітність інтенсивностей у зображенні та обчислюється за формулою Шеннона:

$$Entropy = - \sum_{i=0}^{255} p_i \log_2(p_i)$$

де p_i - ймовірність появи пікселя з інтенсивністю i .

ЕМЕ (Enhancement Measure Estimation) – метрика, що кількісно оцінює локальний контраст зображення шляхом поділу його на блоки:

$$EME = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N 20 \cdot \log_{10} \left(\frac{I_{max}^{(m,n)}}{I_{min}^{(m,n)} + \varepsilon} \right)$$

де M, N - кількість блоків по вертикалі та горизонталі, $I_{max}^{(m,n)}$ та $I_{min}^{(m,n)}$ - максимальне та мінімальне значення пікселя в (m, n) -му блоці, ε - мала додатна константа для уникнення ділення на нуль.

Процедура тестування передбачала штучне додавання гаусівського шуму до зображень з валідаційного набору. До кожного зображення застосовувалися три методи шумозаглушення: DnCNN, SwinIR і класичний алгоритм BM3D. Після відновлення кожного зображення обчислювалися всі чотири метрики, після чого результати усереднювалися для кожного методу.

Таблиця 2.4. Усереднені значення метрик для кожного з методів

Метод	PSNR	SSIM	Ентропія	EME
BM3D	34.29	0.811	5.083	11273.01
DnCNN	33.14	0.778	5.064	9615.88
SwinIR	9.58	0.552	2.751	6745.75

Найвищі значення PSNR та SSIM спостерігаються у методу BM3D, що підтверджує його здатність зберігати як загальну інтенсивність, так і локальні структурні характеристики зображення. Разом із тим, BM3D є класичним методом, який потребує значного часу на обробку одного зображення, що робить його малопридатним для великомасштабної обробки в реальному часі.

Модель DnCNN показала дуже конкурентоспроможні результати при суттєво меншому обчислювальному навантаженні, що робить її практичним вибором для застосування в клінічній практиці. Крім того, її архітектура допускає донавчання, що дозволяє адаптувати модель під конкретний тип зображень.

Модель SwinIR виявилася найменш ефективною в контексті фільтрації шумів. Ймовірно, це пов'язано з неадаптованістю її параметрів до рентгенівських зображень у градаціях сірого.

У якості методу для етапу фільтрації шумів було обрано згорткову нейронну мережу DnCNN, яка забезпечила оптимальний баланс між якістю

денойзингу та обчислювальною ефективністю. За результатами проведеного порівняльного тестування модель DnCNN продемонструвала високі значення PSNR та SSIM, що свідчить про здатність точно відновлювати структурні особливості судинних зображень після додавання шуму. Крім того, результати за метриками ентропії (Entropy) та покращення локального контрасту (EME) вказують на ефективне збереження інформативного вмісту без надмірного згладжування.

Хоча класичний алгоритм BM3D показав трохи вищі результати за окремими метриками, його обчислювальна складність є значно вищою, що створює труднощі при обробці великих обсягів даних або при інтеграції в реальні медичні системи. Модель SwinIR, попри свою сучасну архітектуру, показала найнижчі показники серед трьох розглянутих методів і виявилася менш придатною для задач рентгенівської ангиографії.

Таким чином, DnCNN було обрано як найбільш придатний метод фільтрації шумів, що поєднує високу якість обробки, швидкодію та адаптивність, що є критично важливими характеристиками для подальшої інтеграції в систему автоматизованого аналізу медичних зображень.

2.3.2. Нормалізація інтенсивностей

У завданнях автоматизованої обробки медичних зображень нормалізація інтенсивностей відіграє важливу роль у забезпеченні стабільності вхідних даних для подальших етапів обробки. У випадку рентгенівської коронарографії, де відсутній фіксований еталон яскравості (на відміну, наприклад, від МРТ із внутрішніми калібрувальними мітками), особливо актуальним є використання універсальних методів нормалізації, що не потребують додаткової анатомічної розмітки.

Було розглянуто три методи, які репрезентують основні класи підходів до масштабування інтенсивностей: Min-Max, Z-score та Mean-Norm.фону.

Лінійна (min-max) нормалізація. Цей метод найпоширеніший та найпростіший у реалізації. Значення кожного пікселя $I_{orig}(x)$ лінійно «розтягуються» в цільовий діапазон $[a; b]$ (наприклад, $[0;1]$ або $[0;255]$) за формулою

$$I_{norm}(x) = \frac{(I_{orig}(x) - I_{min})(b - a)}{I_{max} - I_{min}},$$

де I_{min} та I_{max} - мінімальні та максимальні значення яскравості у вихідному зображенні, а $I_{orig}(x)$ - інтенсивність пікселя x .

Метод дає стабільний результат, не потребує додаткових параметрів та швидко обчислюється, але при наявності далеких «викидів» (дуже яскравих або темних точок) може суттєво змінювати яскравості більшості пікселів.

Z-score нормалізація (стандартизація). Тут кожне значення яскравості замінюється на відхилення від середнього, поділене на стандартне відхилення.

Формула має вигляд

$$I_z(x) = \frac{I_{orig}(x) - \mu}{\sigma},$$

де μ – середнє значення яскравості зображення, а σ – стандартне відхилення.

Такий спосіб часто застосовується в нейромережових моделях, особливо в задачах, де важливо мати нульове середнє та одиничну дисперсію. Проте z-score нормалізація не прив'язана до «жорсткого» діапазону і може видавати від'ємні або набагато більші від 1 значення. Це може потребувати додаткового «кліпінгу» або масштабування.

Mean-Norm нормалізація — це спрощена альтернатива, в якій кожне піксельне значення зображення масштабується відносно середнього

$$I' = \frac{I}{\mu}$$

Цей підхід дозволяє швидко адаптувати яскравість до середнього рівня, але не враховує дисперсії даних, тому менш стійкий до нерівномірних розподілів інтенсивностей. Його основна перевага — обчислювальна простота та відсутність необхідності зберігати точні мінімальні/максимальні або стандартні значення.

Таким чином, всі три методи забезпечують базову уніфікацію рівнів яскравості, але демонструють різну чутливість до особливостей зображення. Подальше порівняння їх ефективності дозволить обрати найбільш стабільний та інформативний підхід до нормалізації у контексті коронарографічних даних.

Для порівняння методів нормалізації інтенсивностей було використано ті самі метрики, що й на етапі фільтрації шумів — ентропію (Entropy) та енергетичну міру контрасту (EME), аналітичні вирази яких подано вище.

Крім них, було додатково застосовано локальне стандартне відхилення (Local STD), що є безреференсним показником варіативності інтенсивностей у межах локальних областей зображення. Ця метрика характеризує ступінь неоднорідності яскравості у локальних ділянках зображення і дозволяє оцінити, наскільки успішно збережено мікrokонтраст після нормалізації. Для її обчислення використовувався ковзний квадратний віконний фільтр (розміром $k \times k$), в межах якого розраховувалося стандартне відхилення пікселів. Аналітичне визначення локального стандартного відхилення для пікселя (i, j) має вигляд:

$$\sigma_{i,j} = \sqrt{\frac{1}{k^2} \sum_{(m,n) \in W_{i,j}} (I_{m,n} - \mu_{i,j})^2}$$

де $W_{i,j}$ - вікно розміром $k \times k$, центроване в точці (i, j) , $I_{m,n}$ - значення інтенсивності в пікселі (m, n) , а $\mu_{i,j}$ - локальне середнє значення в межах вікна:

$$\mu_{i,j} = \frac{1}{k^2} \sum_{(m,n) \in W_{i,j}} I_{m,n}$$

Значення локального стандартного відхилення усереднювалося по всьому зображенню для отримання єдиного скалярного значення, придатного для подальшого порівняння.

Процедура тестування методів нормалізації інтенсивностей передбачала попереднє застосування кожного з трьох підходів — Min-Max, Z-Score та Mean Normalization — до кожного зображення у валідаційному наборі. Для кожного з нормалізованих зображень було обчислено три метрики: ентропію, ЕМЕ та локальне стандартне відхилення. Після обчислення метрик для всіх зображень результати усереднювалися (табл. 2.5).

Таблиця 2.5. Усереднені значення метрик для кожного з методів

Метод	Ентропія	ЕМЕ	Local STD
Min-Max	4.367	7540.74	7.071
Z-Score	4.350	7510.62	7.082
Mean-Norm	4.358	7534.77	7.084

Метод Min-Max нормалізації, який передбачає масштабування піксельних значень у межах від 0 до 255, продемонстрував найвищий показник ентропії (4.367), що вказує на максимальну збережену інформаційну складову у результаті трансформації. Окрім цього, значення метрики ЕМЕ для цього підходу (7540.738) також виявилось найвищим серед розглянутих, що свідчить про кращу локальну виразність деталей зображення. Показник Local STD для цього методу є найнижчим (7.071), що вказує на більш згладжений характер локальних варіацій інтенсивностей, тобто підвищену однорідність у межах фрагментів зображення.

Метод Z-Score нормалізації, що ґрунтується на центруванні значень відносно математичного сподівання та масштабуванні за стандартним відхиленням, демонструє дещо нижчу ентропію (4.350) порівняно з іншими підходами. Це може свідчити про часткову втрату тонких деталей або надмірну стандартизацію значень у певних областях. Значення EME (7510.622) для цього підходу також є найменшим серед трьох методів, хоча відмінність від інших не є суттєвою. Local STD (7.082) перебуває на середньому рівні, що означає дещо вищий ступінь локальних флуктуацій у порівнянні з Min-Max.

У випадку Mean Normalization, яка виконує центрування піксельних значень відносно середнього значення зображення, а потім масштабує їх до діапазону [0, 255], показник ентропії (4.358) майже не відрізняється від результатів Z-Score, що свідчить про зіставний рівень збереження інформації. Значення EME (7534.772) виявилось трохи меншим за показник Min-Max, однак вищим за Z-Score. Водночас Local STD (7.084) є найвищим серед усіх варіантів.

Загалом, метод Min-Max демонструє найбільш збалансовану та передбачувану поведінку за всіма метриками. Його простота реалізації, відсутність потреби в попередньому знанні глобальних статистичних характеристик та стійкість до варіацій у даних роблять його доцільним вибором на етапі нормалізації зображень у медичних задачах.

Mean Normalization забезпечує помірний баланс між усіма показниками, але не демонструє найкращих значень у жодній із метрик.

Z-Score є ефективним у випадках, де важливе приведення зображень до однакового статистичного розподілу, однак з точки зору локального контрасту та ентропії його показники дещо поступаються конкурентам.

Метод Min-Max демонструє перевагу за всіма трьома метриками, що свідчить про його стабільну ефективність у задачах нормалізації медичних зображень.

2.3.3. Покращення контрасту

На етапі покращення контрасту було порівняно класичне глобальне вирівнювання гистограми (Global Histogram Equalization), адаптивне вирівнювання з обмеженням контрасту (CLAHE), а також простий метод фільтрації з підсиленням країв — Unsharp Masking, що дозволяє підсилити локальні градієнти без суттєвого спотворення зображення.

Глобальне вирівнювання гистограми. Алгоритм уніфікує розподіл яскравостей, розтягуючи області з високою концентрацією пікселів і послаблюючи зони з низькою насиченістю. Формула для перетворення яскравості r_k у результатне s_k може виглядати так:

$$s_k = (L - 1) \sum_{j=0}^k p_j$$

де L - кількість дискретних яскравостей (наприклад, 256), а p_j імовірність появи рівня яскравості j . Цей метод ефективний у випадках, коли загалом уся гистограма зображення зсунена або стиснена. Водночас для знімків із сильними локальними відмінностями інтенсивностей може виникати перенасичення в окремих ділянках (рис. 2.13).

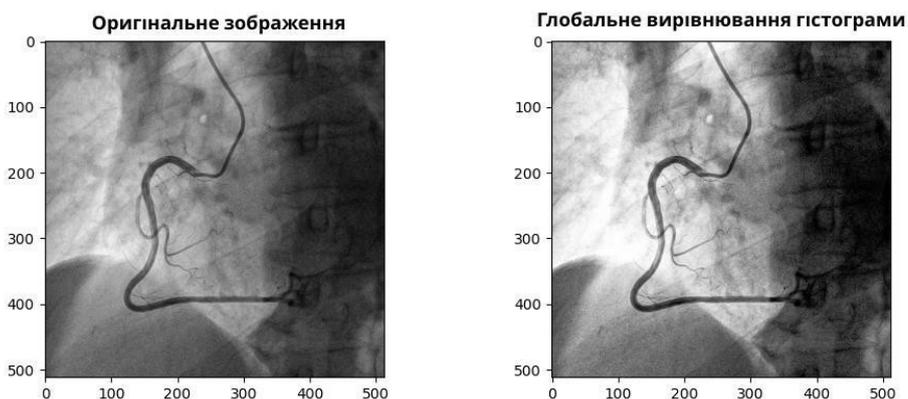


Рисунок 2.13. Покращення контрасту методом глобального вирівнювання гистограми

Адаптивне вирівнювання гистограми (CLAHE). Contrast Limited Adaptive Histogram Equalization (CLAHE) ділить зображення на локальні вікна (tiles) і виконує окреме гистограмне вирівнювання в межах кожного

блоку. Для запобігання «перенасиченню» контрасту задається поріг (clip limit), що обмежує надмірне посилення яскравих або темних зон (рис. 2.14). Такий підхід стає корисним, коли зображення включає дуже світлі та дуже темні регіони.

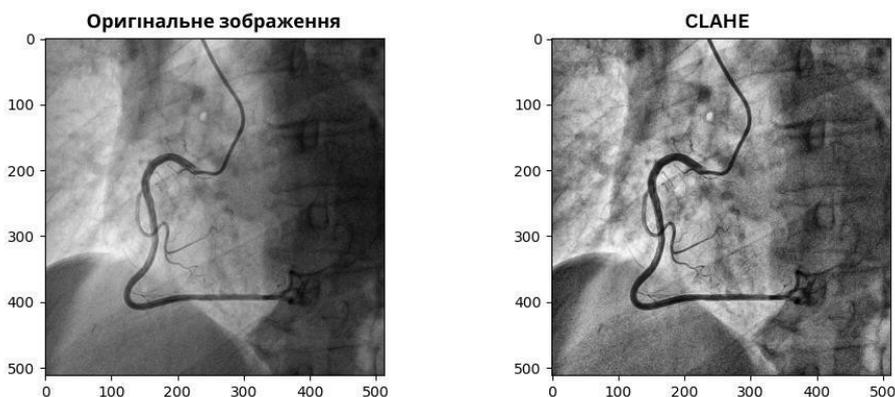


Рисунок 2.14. Покращення контрасту методом CLANE

Unsharp Masking - базується на підсиленні контурів і локальних градієнтів. Він передбачає створення згладженого (розмитого) варіанту зображення та підсилення різниці між оригіналом і розмиттям. Таким чином, контури стають більш вираженими без істотного впливу на загальний рівень яскравості (рис. 2.15).

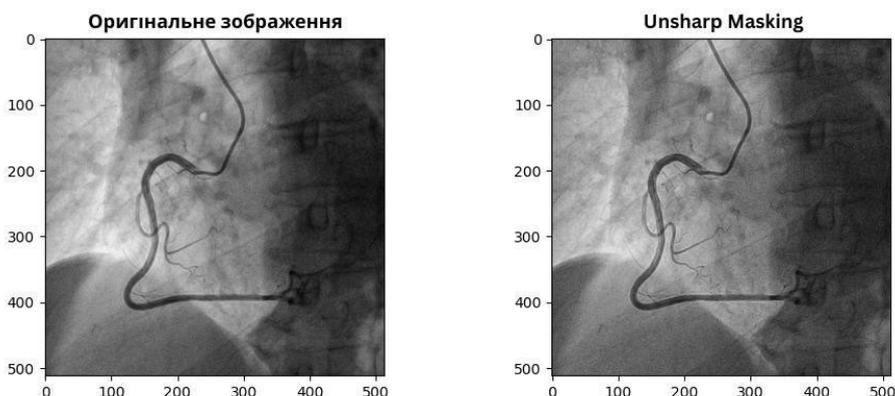


Рисунок 2.15. Покращення контрасту методом Unsharp Masking

Усі три методи є широко застосовуваними в попередній обробці зображень, і в цьому дослідженні вони були порівняні з метою виявлення найпридатнішого для завдань рентгенівської ангіографії (рис. 2.16).

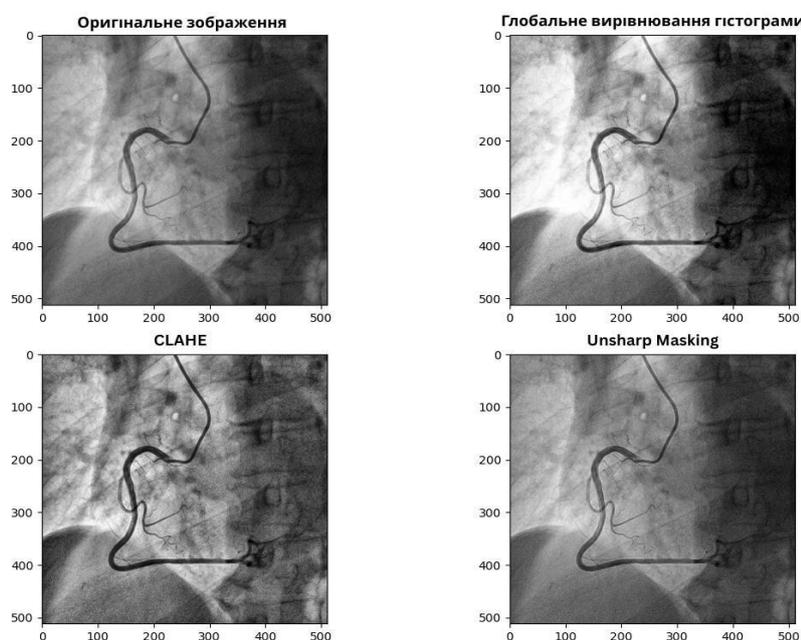


Рисунок 2.16. Результати застосування покращення контрасту

Для оцінки ефективності методів покращення контрасту було використано ті ж метрики, що й на попередньому етапі нормалізації інтенсивностей. Це дозволяє здійснювати узгоджене порівняння якості попередньої обробки зображень за єдиними критеріями (табл. 2.6). Зокрема, було обчислено:

1. **Інформаційну ентропію** (Entropy), яка кількісно відображає рівень різноманітності інтенсивностей у зображенні та дозволяє оцінити збереження або втрату деталей після обробки.
2. **ЕМЕ** (Enhancement Measure Estimation) — числову характеристику покращення контрасту, яка базується на співвідношенні між мінімальними та максимальними значеннями пікселів у локальних блоках. Ця метрика чутлива до локальних контрастних змін.
3. **Локальне середньоквадратичне відхилення** (Local STD), яке розраховується для блоків зображення і відображає рівень локальної

варіативності інтенсивностей. Цей показник дозволяє оцінити ступінь підсилення локальних деталей або шумів.

Оскільки покращення контрасту не передбачає наявності "істинного" зображення-орієнтира, класичні метрики типу PSNR або SSIM у цьому випадку не застосовувалися. Замість цього аналіз зосереджено на змінах статистичних властивостей зображення, які прямо впливають на його візуальну інформативність.

Процедура тестування передбачала попереднє перетворення кожного зображення за відповідним методом покращення контрасту. Для кожного результату обчислювались вищезазначені метрики, після чого значення усереднювались по всьому набору для кожного методу.

Таблиця 2.6. Усереднені значення метрик для кожного з методів

Метод	Ентропія	ЕМЕ	Local STD
GlobalHist	4.88	22751.8 5	7.52
CLANE	4.99	16568.1 7	5.51
UnsharpMas k	4.29	11543.3 0	6.50

Аналіз метрик демонструє різноспрямовані особливості кожного з підходів. Метрика ентропії, що відображає загальний рівень інформаційного наповнення зображення, виявилась найвищою для CLANE (4.99), що свідчить про його здатність збільшувати кількість унікальних значень пікселів та підвищувати деталізацію. Проте це зростання супроводжується зниженням інших характеристик — зокрема, показник ЕМЕ, який є критично важливим для локального контрасту, у CLANE значно нижчий (16568.17), ніж у глобального вирівнювання гістограми (22751.85). Високе значення ЕМЕ для глобального підходу свідчить про покращене розділення

структурних елементів зображення, що є ключовим для візуалізації судинної архітектури.

Метрика локального стандартного відхилення Local STD, яка описує рівень варіативності інтенсивностей у межах невеликих блоків зображення, також вища у глобального вирівнювання (7.52), ніж у CLANE (5.51) або Unsharp Masking (6.50). Це свідчить про більшу різницю яскравостей в межах мікроструктур, що може сприяти кращому виявленню локальних змін в анатомії.

Метод Unsharp Masking, хоча й забезпечує посилення країв, продемонстрував найнижчі результати за всіма метриками: ентропія (4.29), EME (11543.30) та Local STD (6.50), що вказує на обмежену ефективність цього підходу в умовах задачі покращення контрасту ангіографічних зображень. Ймовірно, це зумовлено тим, що метод переважно підсилює лише контурні елементи, не забезпечуючи достатнього загального покращення яскравісного розподілу по зображенню.

Метод глобального вирівнювання гистограми демонструє найбільше значення метрики EME, що вказує на найефективніше покращення локального контрасту по всьому зображенню, незважаючи на дещо нижче значення ентропії порівняно з CLANE.

З урахуванням балансу між покращенням локального контрасту, збереженням деталей та підвищенням візуальної чіткості, для подальших етапів було обрано метод глобального вирівнювання гистограми як найбільш ефективний варіант покращення контрасту ангіографічних зображень.

2.4. Розв'язання задачі попередньої обробки зображень

У межах даного підрозділу реалізовано систему попередньої обробки медичних зображень, що включає три етапи: фільтрацію шуму, нормалізацію інтенсивностей та покращення контрасту. На кожному з етапів було

відібрано та протестовано кілька методів, результати яких оцінювались за відповідними метриками якості.

На етапі фільтрації порівнювалися методи DnCNN, SwinIR та класичний BM3D за метриками PSNR, SSIM, ентропія та EME. BM3D показав найвищу якість (PSNR = 34.29, SSIM = 0.811), проте мав високу обчислювальну складність. DnCNN досяг компромісу між якістю та швидкістю (PSNR = 33.14, SSIM = 0.778). SwinIR виявився непридатним для задачі через низькі значення всіх метрик. Для нормалізації інтенсивностей розглядалися методи Min-Max, Z-Score та Mean-Norm. Метод Min-Max досяг найвищої ентропії (4.367) та EME (7540.74), забезпечивши найнижче локальне стандартне відхилення, що вказує на однорідність та збереження інформації.

На етапі покращення контрасту порівнювалися CLAHE, глобальне вирівнювання гистограми та Unsharp Masking. CLAHE забезпечив найвищу ентропію, однак викликав надмірне підсилення локальних зон. Глобальне вирівнювання гистограми досягло найкращого EME (22751.85) і виявилось ефективнішим для підвищення контрасту без спотворень.

Результати дослідження лягли в основу формування цілісного конвеєра попередньої обробки зображень.

Висновки до розділу 2

У результаті реалізації задачі попередньої обробки було сформовано обґрунтовану послідовність дій, що забезпечує стабільне підвищення якості медичних зображень при збереженні важливих структурних деталей. Запропоноване рішення ґрунтується на експериментальному порівнянні сучасних методів фільтрації, нормалізації та підвищення контрасту із використанням кількісних метрик.

Фінальна конфігурація включає:

- DnCNN для ефективного та швидкого шумозаглушення;

- Min-Max нормалізацію як найбільш збалансований метод за інформаційною насиченістю і стабільністю яскравості;
- глобальне вирівнювання гістограми для підвищення візуальної контрастності зображень.

Запропонований підхід забезпечує підвищення якості даних, що подаються на вхід сегментуючим та класифікаційним моделям, і є оптимальним з точки зору продуктивності, якості та адаптованості до умов клінічного застосування.

РОЗДІЛ 3. СЕГМЕНТАЦІЯ КОРОНАРНИХ АРТЕРІЙ. МЕТОДИ ВИДІЛЕННЯ ОЗНАК ТА СЕГМЕНТАЦІЇ СТЕНОЗУ

3.1. Постановка задачі сегментації коронарних артерій

Завдання сегментації коронарних артерій на рентгенівських зображеннях коронарографії є одним з ключових етапів у створенні автоматизованої системи діагностики ішемічної хвороби серця. Сегментація дозволяє локалізувати судинне русло, що необхідно як для подальшого аналізу геометричних характеристик артерій, так і для виявлення патологічних звужень (стенозів). Проте, через низький рівень контрасту, варіативну анатомію судин та наявність шумів і артефактів, точна автоматизована сегментація залишається складною технічною задачею.

Формально, вхідне зображення можна представити як тензор $X \in R^{H \times W \times 3}$, де H – висота зображення, W – ширина, а кожен піксель містить три кольорові компоненти (RGB). Результатом сегментації є карта міток $Y \in \{0, 1, \dots, C - 1\}^{H \times W}$, де C – кількість класів.

У випадку бінарної сегментації ця задача зводиться до побудови функції $f_{\theta}(X) \rightarrow Y$, яка відносить кожен піксель до класу "судина" або "фон". У випадку багатокласової сегментації, кожен піксель має бути віднесений до конкретної анатомічної зони згідно зі структурною розміткою судинного дерева.

З метою розв'язання цієї задачі, необхідно:

- провести системний аналіз сучасних наукових підходів до сегментації медичних зображень, зокрема у контексті коронарної ангіографії;
- виділити провідні архітектури глибокого навчання, які можуть бути адаптовані до задачі сегментації ангіографічних кадрів;
- розглянути можливість реалізації як бінарного, так і багатокласового підходів, проаналізувавши переваги та обмеження кожного з них;
- визначити критерії якості сегментації, такі як IoU, F1-score (Dice), точність (Precision) та повнота (Recall), що будуть використані для об'єктивного порівняння моделей;

- реалізувати експериментальне порівняння на тестовій вибірці реальних медичних зображень із доступного анотованого датасету, що охоплює широкий спектр анатомічних варіацій;
- на основі результатів тестування зробити обґрунтований вибір підходу та моделі, які найкраще адаптовані до умов клінічного використання.

Поставлена задача є центральною для подальших етапів

3.2. Огляд сучасних методів та підходів до сегментації судин.

3.2.1. Автоматизована сегментація коронарних артерій з використанням YOLO-Angio

Система YOLO-Angio призначена для автоматичного виділення коронарного судинного дерева на ангиографічних зображеннях [19]. Актуальність такого підходу обумовлена тим, що коронарна ангиографія є «золотим стандартом» діагностики ішемічної хвороби серця, проте її інтерпретація вручну є трудомісткою та суб'єктивною.

Метою розробників YOLO-Angio було прискорити і підвищити об'єктивність аналізу ангиограм шляхом автоматичної сегментації судин (рис. 3.1, вибір ознак виконується для підвищення контрастності судин, після чого виконується сегментація на основі YOLO з використанням ансамблевої моделі та логічного підходу для побудови остаточного дерева коронарних артерій).

Архітектура YOLO-Angio складається з трьох основних етапів обробки зображення.

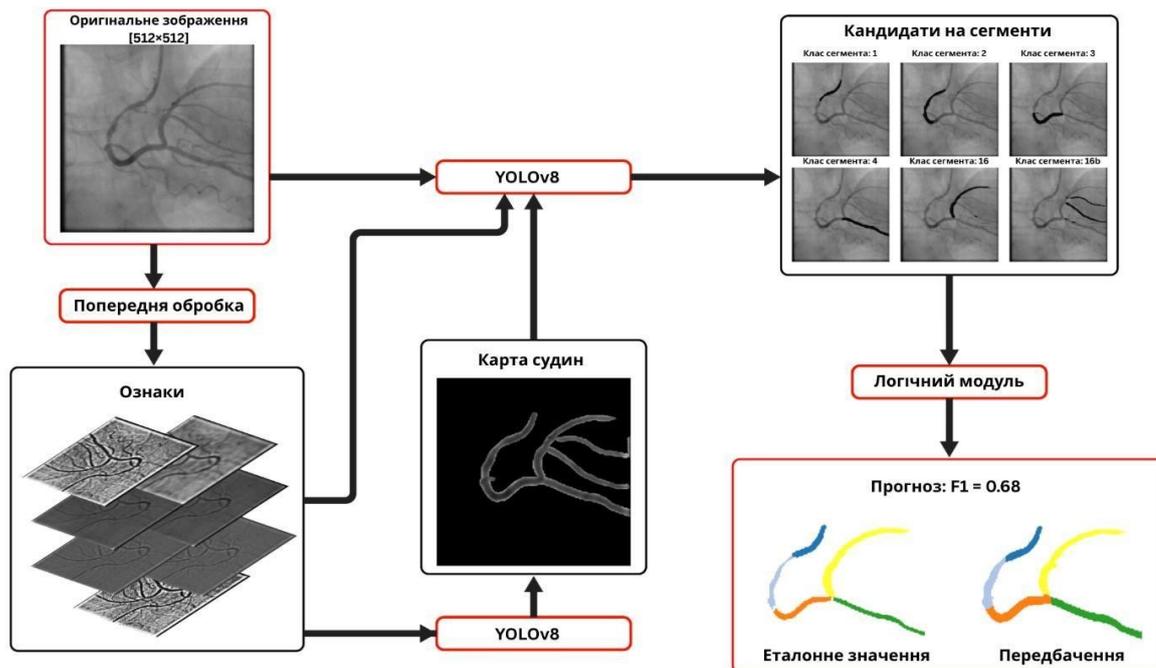


Рисунок 3.1. Огляд YOLO-Angio [19]

1. **Попередня обробка.** Застосовуються різноманітні перетворення зображення для покращення контрасту судин та виділення характерних особливостей (наприклад, фільтрація високих частот, вирівнювання гістограми, морфологічні перетворення тощо).
2. **Сегментація на основі YOLOv8.** Глибока модель YOLOv8 використовується для сегментації судин, при цьому тренується кілька окремих моделей на різних варіантах перетворених зображень, і їх результати об'єднуються в ансамбль. Такий підхід генерує карту ймовірностей присутності судин (vessel map) на зображенні.
3. **Логічна реконструкція дерева.** На фінальному етапі застосовується алгоритм постобробки: на основі логічних правил та графового аналізу результати ансамблю коригуються і збираються у повне дерево коронарних артерій. Це дозволяє усунути несумісності та прогалини, забезпечуючи цілісність сегментації судинного дерева.

Запропонований підхід продемонстрував високі показники на даних конкурсу ARCADE 2023. Зокрема, модель YOLO-Angio посіла 3-тє місце у змаганні, досягнувши F1-міри 42,89% на контрольному наборі зображень.

Використання ансамблю з кількох YOLOv8-моделей разом із логічним “пост-фільтром” дозволило підняти точність сегментації на ~58% порівняно з окремою моделлю. Для прикладу, найкраща одиночна модель (на основі вирівнювання гістограми) дала $F1 \approx 31\%$, тоді як повний ансамблево-логічний підхід — 41,1% (на публічному тесті) та 42,89% на фінальному наборі.

До переваг системи YOLO-Angio належать відносно висока швидкодія та точність. Одностадійна модель детекції YOLOv8 забезпечує швидку обробку зображень, а поєднання декількох моделей і подальша перевірка цілісності сегментації зменшує кількість помилкових спрацьовувань (хибнопозитивних пікселів). Метод успішно виявляє основні артеріальні гілки навіть за значного шуму та варіацій контрасту.

Водночас недоліком є те, що нейромережа має тенденцію «перерисовувати» зайві контури судин, тобто переоцінювати межі артерій, що потребує додаткового очищення результату (застосування морфологічної ерозії). Крім того, оптимізація під метрику F1 у межах змагання могла не врахувати клінічну значущість різних типів помилок сегментації. Автори відзначають, що модель гірше справляється з рідкісними класами судин (деякі дрібні гілки, мало представлені у навчальних даних) та потребує розширення навчальної вибірки для підвищення стабільності результатів.

3.2.2. Ансамблевий підхід до сегментації та локалізації патологій

Наступний підхід оснований на ансамблевому навчанні і був запропонований іншою командою в рамках того ж конкурсу ARCADE 2023. Метою цього рішення було виконання двох взаємопов’язаних завдань:

1. Сегментації всіх коронарних артерій на зображенні.
2. Локалізації можливих стенотичних уражень (звужень) на них.

Через складність форми судин і значну варіативність проявів стенозів автори зосередилися на побудові стійкої багатоетапної методики навчання з подальшим об'єднанням кількох моделей.

Розроблений ансамблевий метод використовує глибоку модель сегментації на основі згорткової нейронної мережі (архітектура BRU-Net – спрощений U-Net з bottleneck-блоками [20]) та реалізує комплексну стратегію тренування. Зокрема, навчання проходить п'ять послідовних стадій:

1. попереднє тренування бінарного сегментатора судин (розрізнення «судина/фон»);
2. навчання багатокласової сегментації всіх судин на зображенні;
3. тонке налаштування з використанням зважених вибірок даних (для врахування дисбалансу класів за товщиною та гілками судин);
4. тонке налаштування із поступовим ускладненням навчальних прикладів за допомогою стратегії curriculum learning (методика F1-CLS);
5. адаптація ансамблю моделей з урахуванням кута огляду (проекції) ангиограми за допомогою спеціального класифікатора.

На заключному етапі автори об'єднали результати шести незалежно натренованих моделей за допомогою зваженого усереднення (ансамблювання), що дозволило майже вдвічі підвищити середню точність прогнозу порівняно з окремою моделлю. Додатково було впроваджено постобробку для виправлення артефактів сегментації – видалення дрібних помилково позначених областей («плям») на вихідному масці судин. На рисунку 3.2 схематично зображено запропоновану архітектуру рішення, яка включає багаторівневе навчання і ансамблеву інтеграцію моделей.

За підсумками оцінювання на тестових даних ARCADE, ансамблева система досягла F1-міри 37,69% для сегментації коронарних артерій та

39,41% для локалізації стенозів, що забезпечило команді 5-те місце на обох лідербордах конкурсу.

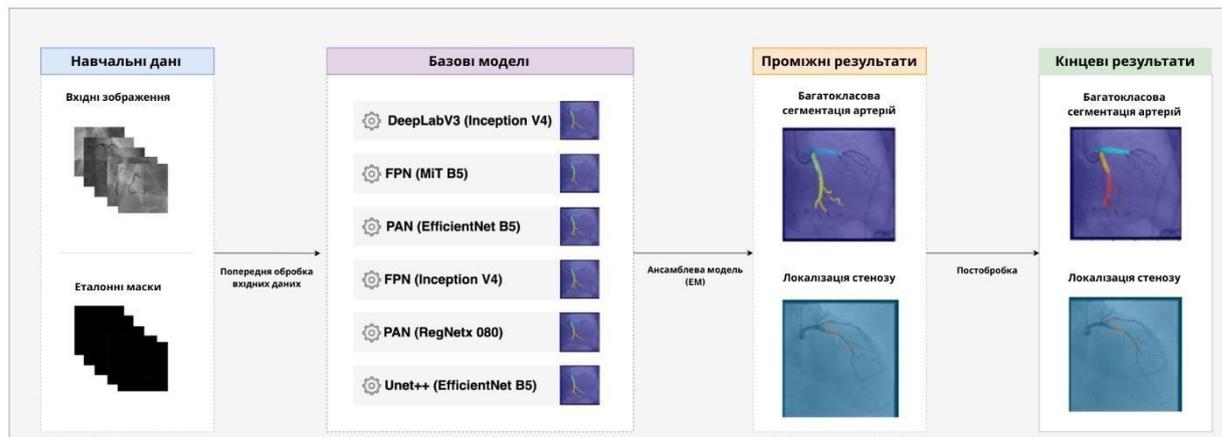


Рисунок 3.2. Запропонована методика сегментації коронарних артерій та стенозів [20]

Ці результати підтверджують ефективність комплексного підходу: об'єднання кількох моделей перевершило кожен з них окремо за точністю. Перевагою даного методу є продумана стратегія навчання і поєднання знань різних моделей. Ансамбль з шести сегментаторів, навчений на основі різних прийомів (зважування даних, curriculum learning тощо), здатен більш надійно виділяти складну судинну структуру, ніж будь-яка окрема модель. Рішення стійке до варіативності даних і продемонструвало застосовність для автоматизованого аналізу ангіограм, що може допомагати лікарям у діагностиці та плануванні втручань.

Водночас, недоліком є висока складність і ресурсоемність такого підходу. Навчання потребує великої кількості етапів і моделей, а використання одразу шести нейронних мереж у складі ансамблю ускладнює розгортання системи та практичне виконання в режимі реального часу. Попри загальне підвищення точності, абсолютні значення метрик залишаються помірними (близько 38–39% F1), що пояснюється складністю задачі та значною міжоператорною варіативністю розмітки ангіограм.

Отже, навіть ансамблевій моделі властиві помилки: дрібні судини та межі стенозів можуть визначатися неточно, вимагаючи подальших покращень алгоритму та якості даних.

3.2.3. SAM-VMNet: Гібридна модель для покращеної сегментації судин

Третій підхід реалізований у вигляді глибокої нейронної моделі SAMVMNet, призначеної як для сегментації коронарних артерій, так і для кількісного виявлення їх стенозів на ангиографічних зображеннях [21]. Відмінністю цього рішення є фокус не лише на побудові маски судин, але й на точному вимірюванні ступеня звуження просвіту артерій, що підвищує цінність результатів з точки зору клінічної діагностики. Автори даної роботи сформували власний розширений датасет ангиограм та використали публічні набори (в тому числі ARCADE) для тренування і валідації моделі, прагнучи забезпечити її надійну роботу в різних умовах знімання.

Архітектура SAM-VMNet поєднує в собі дві паралельні гілки глибинного навчання (рис. 3.3).

Перша гілка – це енкодер на основі моделі MedSAM (медична версія сегментатора Segment Anything), яка виконує селективне виділення найважливіших регіонів зображення за допомогою механізму самоуваги.

Друга гілка – це модифікована згорткова мережа VM-UNet, орієнтована на детальну сегментацію судинного русла.

Така комбінована архітектура (SAM-VMNet) дозволяє інтегрувати переваги трансформерних і класичних CNN-підходів: MedSAM виділяє контекстні ознаки і допомагає сфокусуватися на ділянках із судинами, а Уподібна мережа забезпечує точне окреслення меж судин, включно з дрібними гілками.

Як показано на рисунку 3.3, мережа містить два паралельних енкодери (MedSAM і VM-UNet), виходи яких об'єднуються для формування

сегментаційної маски. Запропонована модель SAM-VMNet продемонструвала високу результативність. На тестовому наборі ARCADE досягнуто середнього показника IoU = 63,03% для сегментації судин (при чутливості 98,32% та специфічності 99,33%).

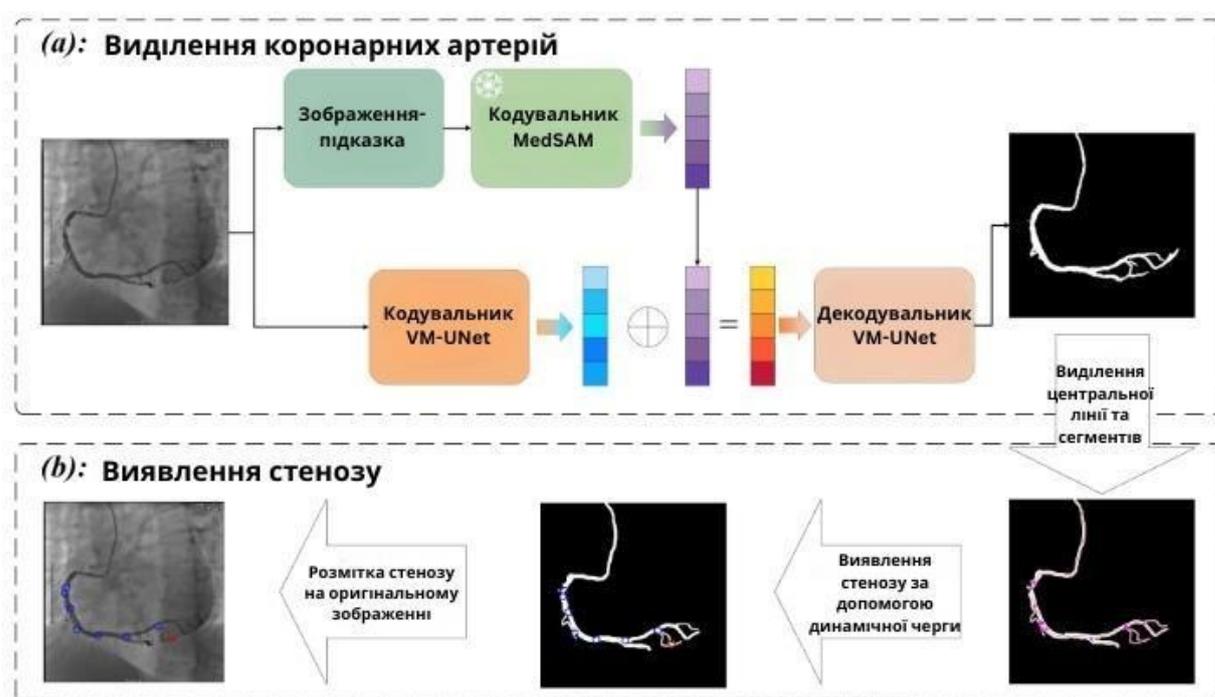


Рисунок 3.3. Робочий процес методу [21]. (a) Сегментація артерій з використанням підходу на основі глибокого навчання; (b) Виявлення стенозу, включаючи виділення центральної лінії артерії, розрахунок діаметра артерії та визначення ступеня стенозу.

Це суттєво перевищує показники попередніх розглянутих методів, що пов'язано як із вдосконаленою архітектурою, так і з використанням розширеного обсягу даних для навчання. Алгоритм виявлення стенозів на основі виміряних діаметрів забезпечив true positive rate (TPR) = 58,67% та positive predictive value (PPV) = 59,11%, що приблизно відповідає F1мірі ~59%. Для перевірки узагальненості підхід протестовано також на змішаному датасеті (власні дані + публічні): було отримано середнє IoU = 63,08%

(чутливість 97,72%, специфічність 99,03%), тобто показники залишилися на тому ж рівні.

До сильних сторін даного методу належить поєднання сучасних архітектур і орієнтація на кінцевий клінічний результат. Використання трансформерної моделі MedSAM в складі сегментатора помітно покращує виділення складних деталей – зокрема, найдрібніших судин і сильно звужених ділянок, з якими гірше справляються традиційні CNN-моделі.

До недоліків SAM-VMNet можна віднести значну складність і вимоги до ресурсів. Модель містить два великих еncoderи та складний конвеєр післяобробки, що ускладнює її навчання і впровадження (табл. 3.1).

Таблиця 3.1. Метрики точності розглянутих підходів

Підхід	Точність сегментації	Точність виявлення стенозів
YOLO-Angio	F1 = 42,89%	–
Ансамблевий метод	F1 = 37,69%	F1 = 39,41%
SAM-VMNet	IoU = 63,03%	TPR = 58,67% PPV = 59,11%

3.3. Методи та підходи, застосовані в роботі

На основі проведеного огляду сучасних наукових публікацій у сфері сегментації коронарних артерій, а також з урахуванням результатів змагання ARCADE 2023, для реалізації експериментальної частини дослідження було визначено доцільність використання двох популярних підходів: моделі U-Net як представника класичних згорткових архітектур та моделі YOLOv8 як сучасного варіанту сегментаційних моделей, побудованих на основі одностадійної детекції.

Крім того, було розглянуто два концептуально різні підходи до розв’язання задачі сегментації: багатокласовий та бінарний. Перший підхід

передбачає ідентифікацію кожного типу артерій окремо, другий — виділення всієї судинної структури як єдиної області інтересу. Така подвійна стратегія дозволяє порівняти точність, універсальність та складність обох варіантів.

3.3.1. Архітектура YOLOv8

Методологія YOLO (You Only Look Once), початково спрямована на розв'язання задач детектування об'єктів, у нових реалізаціях (зокрема, YOLOv8) підтримує також модуль для сегментації. Такий підхід дає змогу оперативного визначати межі коронарних артерій, розглядаючи їх як окремі «об'єкти». Його ключовою перевагою є висока швидкість та достатній рівень точності. Саме тому YOLOv8 застосовується в багатьох розробках, орієнтованих на практичне впровадження у клінічні робочі процеси.

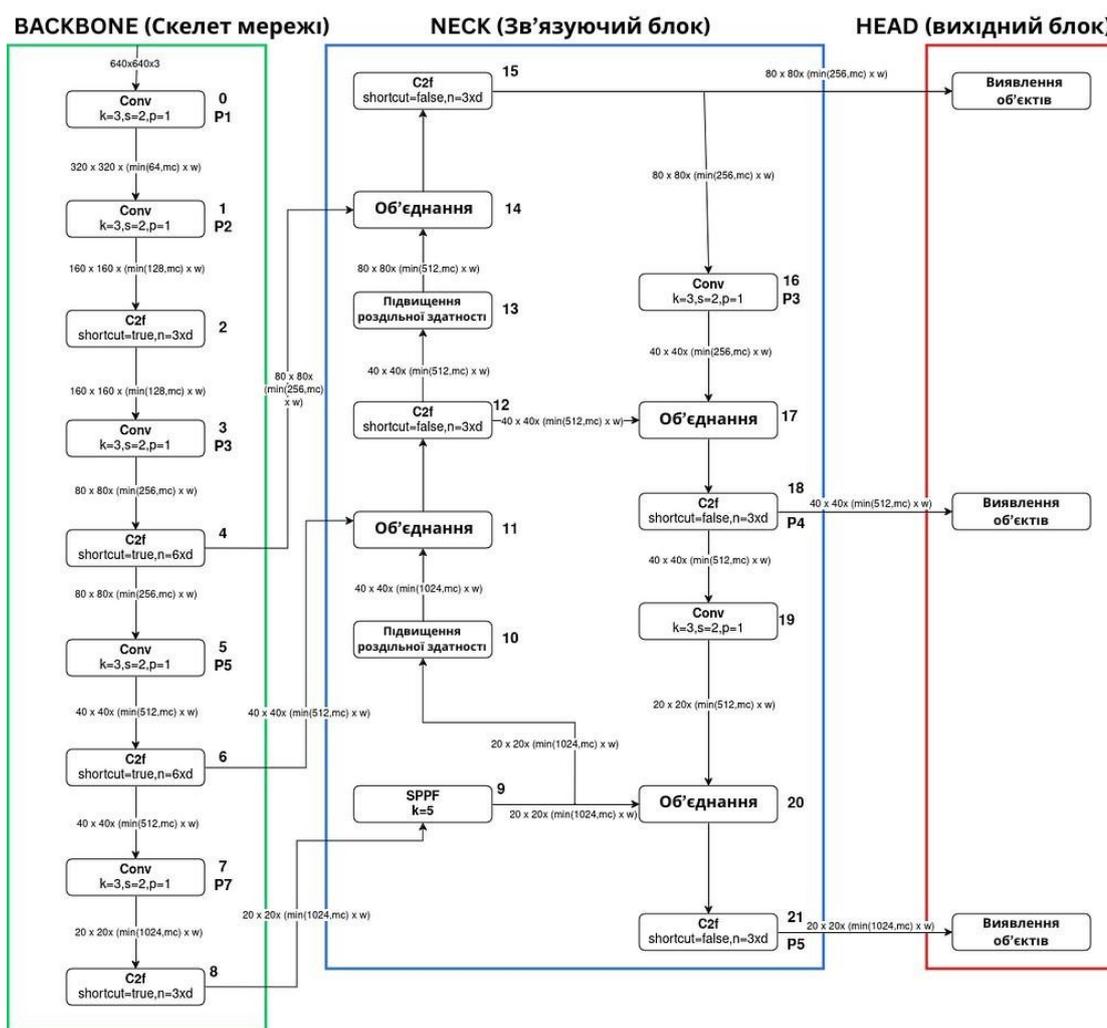


Рисунок 3.4. Архітектура згорткової нейронної мережі типу YOLOv8¹⁸,

Архітектура YOLOv8 (рис. 3.4) складається зі скелета (Backbone) для вилучення ознак, проміжного блоку (Neck) для багаторівневої обробки просторових представлень та вихідного блоку (Head) для здійснення передбачення класів та локалізації об'єктів. Conv – згортка, C2f – модифікований блок (Cross Stage Partial Fusion), SPPF – Spatial Pyramid Pooling - Fast, P1–P7 - рівні просторових ознак.

Backbone - це архітектура глибокого навчання, яка діє як екстрактор ознак введеного зображення. Neck поєднує ознаки, отримані з різних шарів модуля Backbone. Head прогнозує класи та межі об'єктів, що є кінцевим результатом роботи моделі виявлення об'єктів.

Водночас варто відзначити, що при наявності дуже тонких чи пересічних гілок судин, а також при злиттях кількох судинних сегментів, YOLOv8 у режимі сегментації може демонструвати певні труднощі. Зокрема, алгоритм схильний розглядати суцільні структури, а відтак виявляти занадто злиті контури або, навпаки, нехтувати незначними гілками. Попри це, у випадках, де пріоритетом є отримання результату в реальному часі зі «задовільною» точністю, YOLOv8 виглядає перспективно (рис. 3.5).

3.3.2. Архітектура U-Net

U-Net [23] є класичним рішенням у галузі семантичної сегментації, особливо популярним для медичних даних. Основу становить «U»-подібна архітектура, що об'єднує низхідну (енкодер) і висхідну (декодер) частини, а також містить пропускні з'єднання (skip connections) між відповідними рівнями. Така структура дає змогу зберігати детальну інформацію про дрібні об'єкти, важливу в задачах сегментації тонких коронарних артерій (рис. 3.6).

¹⁸ Джерело зображення:

https://miro.medium.com/v2/resize:fit:1100/format:webp/1*mhXRY9fhEDjFVSlGHpfh4w.png

Головною перевагою U-Net вважають високий рівень точності відтворення контурів, навіть якщо останні є слабконтрастними або надто тонкими. У контексті рентгенівських знімків серця це набуває особливого значення, оскільки дрібні бокові гілки чи слабка контрастність можуть бути критичними для діагностичного висновку.

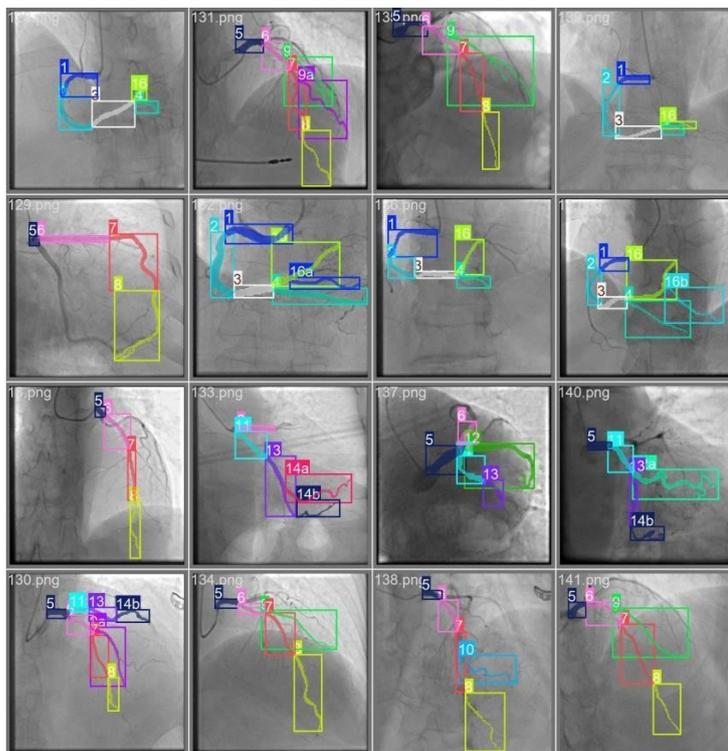


Рисунок 3.5. Візуалізація валідації моделі YOLOv8

Недоліком даного підходу є відносно більша витрата обчислювальних ресурсів, що може ускладнити застосування U-Net у високошвидкісних сценаріях або на обладнанні з обмеженими характеристиками.

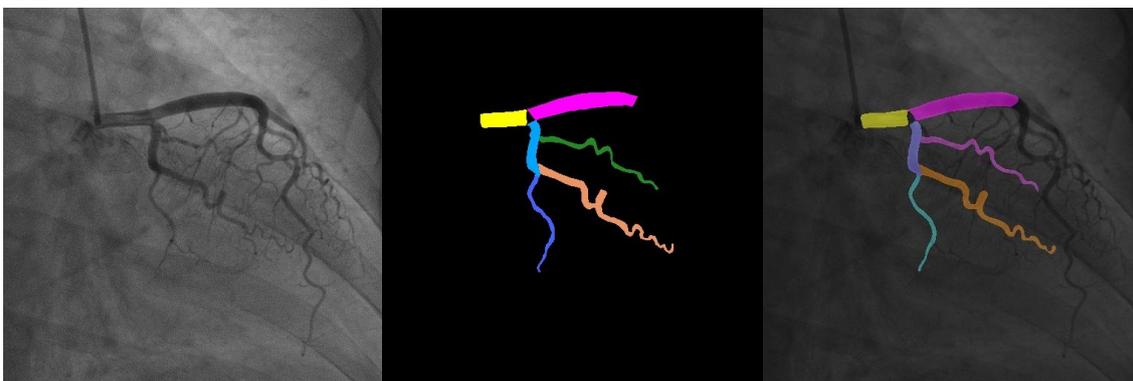


Рисунок 3.6. Сегментація коронарних артерій (зліва – рентгенівське зображення, посередині – кольорова маска коронарної артерії, справа – результат сегментації коронарної артерії моделлю Unet)

3.3.3. Багатокласова сегментація

Багатокласовий підхід до сегментації передбачає одночасне виділення кількох класів об'єктів (в даному випадку — окремих сегментів коронарних артерій) на одному зображенні (рис. 3.7). Це забезпечує більш детальне картографування судинного дерева, що особливо важливо для подальшого аналізу розташування патологій, таких як стенози, а також для топологічного аналізу судин.

Багатокласова сегментація є складнішою задачею у порівнянні з бінарною, оскільки вимагає точного розділення пікселів не лише на судини та фон, але й на конкретні підкласи артерій, включаючи дрібні та рідкісні гілки.

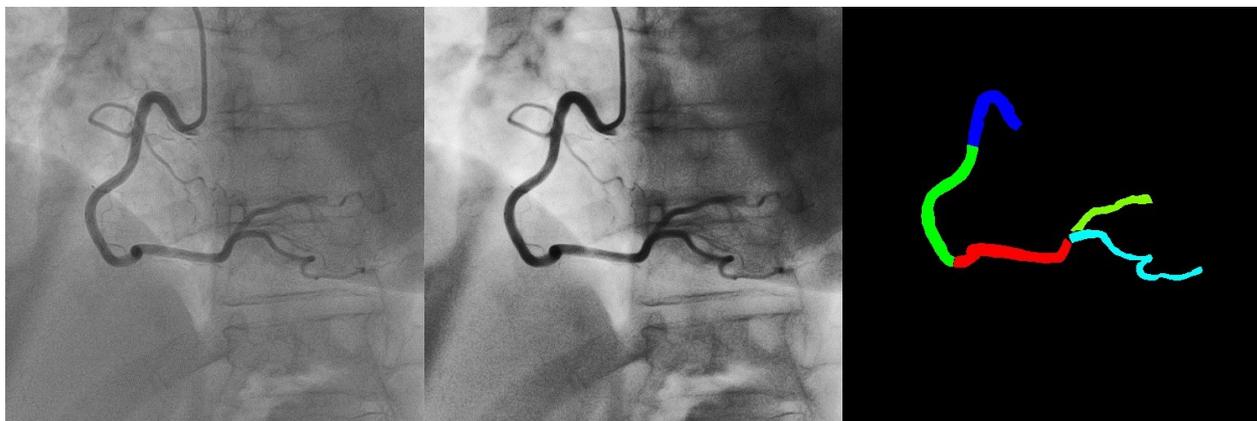


Рисунок 3.7. Приклад багатокласової сегментації (зліва – необроблене рентгенівське зображення коронарної артерії, посередині – рентгенівське зображення після попередньої обробки, справа – багатокласова кольорова маска коронарної артерії)

3.3.4. Бінарна сегментація

Бінарна сегментація передбачає поділ пікселів зображення лише на два класи: «судина» та «фон» (рис. 3.8). Такий підхід дозволяє спростити задачу, що особливо важливо при обмежених ресурсах або у випадку, коли не вимагається детальна класифікація артерій за типом. Він також є корисним як проміжний етап при подальшій класифікації стенозів або визначенні топології судин.

Бінарна сегментація має перевагу в стабільності навчання та меншу чутливість до дисбалансу класів, однак позбавляє систему детальної анатомічної інтерпретації.

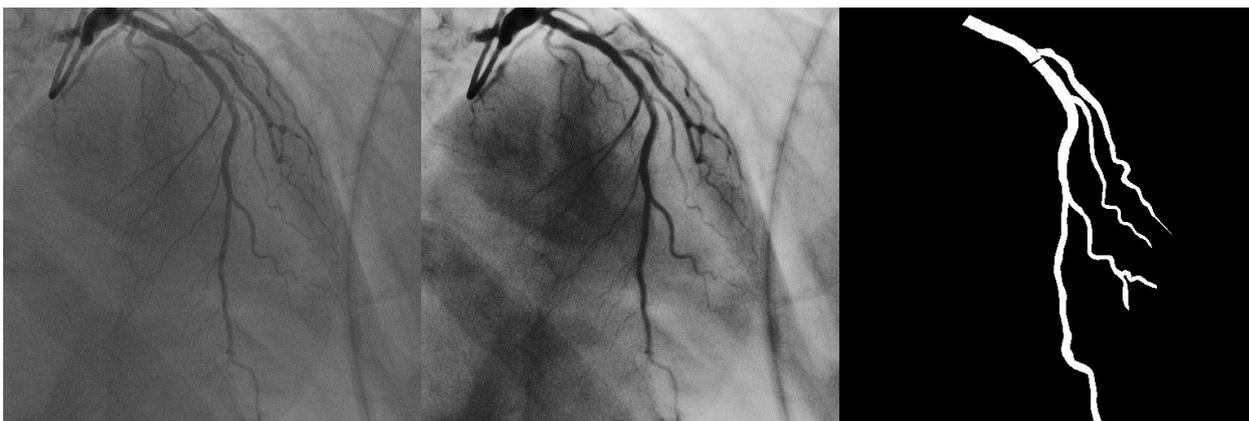


Рисунок 3.8. Приклад бінарної сегментації (зліва – необроблене рентгенівське зображення коронарної ратерії, посередині – рентгенівське зображення після попередньої обробки, справа – бінарна маска коронарної артерії)

3.4. Реалізація та тренування моделей

На етапі сегментації коронарних артерій було здійснено навчання двох архітектур глибокого навчання: U-Net з енкодером ResNet34 та YOLOv8n-seg — моделі сімейства YOLOv8, орієнтованої на задачу сегментації. В обох

випадках використовувалися одні й ті ж вхідні дані, що дозволило забезпечити справедливе порівняння моделей на однаковому наборі задач.

3.4.1. Вхідні дані

Для навчання нейронних мереж використовувався відкритий датасет ARCADE (Automatic Region-based Coronary Artery Disease diagnostics using x-ray angiography imagEs), який призначений для задач автоматичного аналізу коронарографічних зображень (рис. 3.9). Він містить 1500 рентгенівських ангиографій з відповідними сегментованими масками.

Усі мітки відповідають медично прийнятій класифікації Syntax Score, що охоплює 26 основних судинних сегментів, включаючи окремо виділений клас «стеноз» (який на етапі сегментації артерій не використовувався). В рамках експерименту для навчання були використані 1000 зображень з відповідними масками як тренувальна вибірка, та 200 зображень для валідації. Решта 300 зображень не залучалися до тренування жодної з моделей і використовувалися виключно на етапі незалежного тестування, що описано в розділі 3.6.

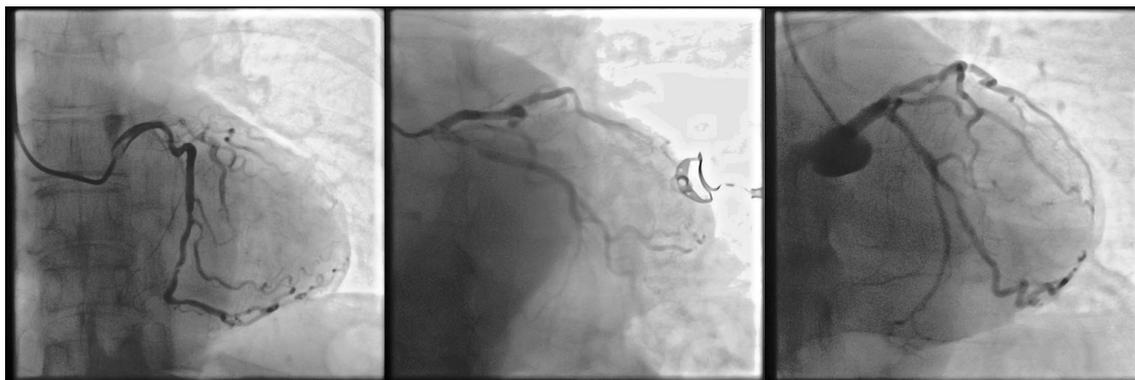


Рисунок 3.9. Приклад вхідних зображень

Перед подачею до нейронних мереж зображення проходили низку етапів попередньої обробки, спрямованих на покращення якості вхідного сигналу та уніфікацію інтенсивностей. Зокрема, було проведено фільтрацію

шумів за допомогою згорткової нейронної мережі DnCNN, яка дозволяє зменшити високочастотні збурення без втрати деталей.

Після цього здійснювалася нормалізація інтенсивностей методом MinMax, що забезпечувала приведення діапазону яскравостей до інтервалу [0, 255] із збереженням відносної контрастності.

Фінальним етапом була процедура контрастного покращення через глобальне вирівнювання гістограми, яка сприяє підвищенню локальної різкості в межах судинних структур.

Оскільки задача сегментації формулювалася як багатокласова семантична, всі маски було приведено до формату «один піксель — один клас», де кожне значення пікселя позначає приналежність до відповідного анатомічного сегменту. Таким чином, дані були повністю підготовлені до подачі у глибокі нейронні архітектури як візуальні вектори входу з однозначними класовими мітками для навчання.

3.4.2. Аугментація

Особливу увагу було приділено аугментації, що застосовувалася однаково до обох моделей. У ході навчання використовувалися як прості, так і комбіновані аугментаційні техніки: вертикальний та горизонтальний фліп з ймовірністю 0.5, зміни відтінку, насиченості та яскравості в HSV-просторі ($h=0.015$, $s=0.7$,

$v=0.4$), масштабування (до $\pm 50\%$), транспозиція (10%), автоматичне застосування найкращих аугментацій (auto_augment), змішування зображень (mixup=0.1), а також обрізка зображень (crop_fraction=0.1). Такі заходи дозволили суттєво збільшити варіативність тренувального набору та покращити здатність моделей до узагальнення.

3.4.3. Навчання моделі U-Net

Архітектура U-Net була реалізована на основі фреймворку PyTorch із використанням бібліотеки `segmentation_models_pytorch`. Як енкодер було обрано ResNet34, попередньо натренований на наборі зображень ImageNet. Кількість вихідних каналів моделі становила 25 — по одному для кожного з сегментованих класів. В якості функції втрат використовувалася `CrossEntropyLoss` із ваговими коефіцієнтами класів, пропорційними їх частотам у тренувальному наборі. Це дозволяло зменшити вплив домінуючих класів, таких як фон, на результати оптимізації.

Навчання U-Net здійснювалося протягом 500 епох при фіксованому розмірі батчу 8 (рис. 3.10, Train loss — аналіз збіжності моделі на тренувальних даних. Val loss — аналіз збіжності моделі на валідаційних даних.). Оптимізація проводилася за допомогою алгоритму Adam.

Для запобігання перенавчанню застосовувалося динамічне зменшення коефіцієнта навчання у разі стабілізації функції втрат на валідаційній вибірці.

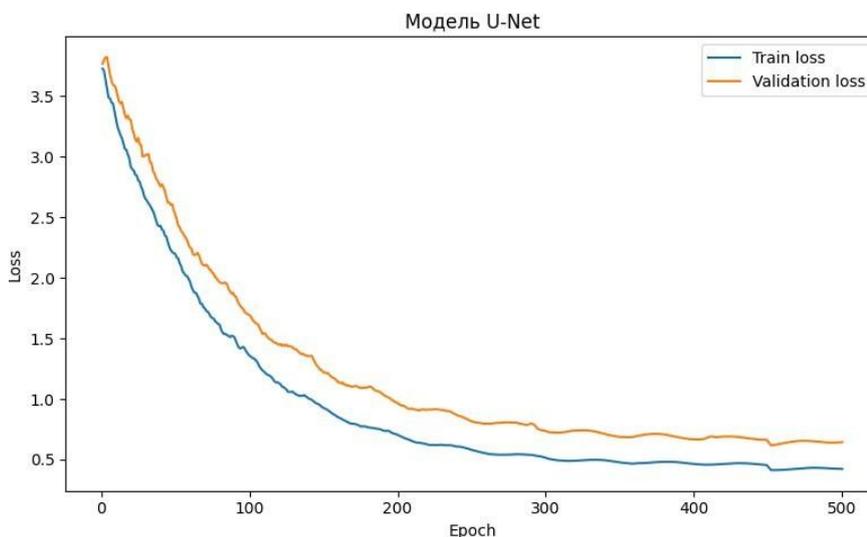


Рисунок 3.10. Графік тренування моделі U-Net

На графіку спостерігається типовий тренувальний патерн: значення функції втрат на тренувальній вибірці зменшується майже монотонно, демонструючи поступове покращення відповідності моделі до навчальних

прикладів. Водночас валідаційна втрата протягом перших 100–150 епох зменшувалася синхронно з train loss, після чого стабілізувалася та коливалася в межах невеликого діапазону, не демонструючи суттєвого переобучення.

Такі результати вказують на достатньо збалансоване тренування: модель успішно навчилася узагальнювати патерни сегментації судин, не надмірно підлаштовуючись під конкретні особливості train-даних.

3.4.4. Навчання моделі YOLOv8

Для моделі YOLOv8 було використано полегшену конфігурацію yolov8nseg.pt, яка є орієнтованою на задачі семантичної сегментації з урахуванням обмежених обчислювальних ресурсів. Навчання здійснювалося за допомогою вбудованого механізму бібліотеки Ultralytics, що включає повний пайплайн тренування, оцінки, логування та збереження результатів. Було також встановлено 500 епох з розміром батчу 8 та розміром вхідного зображення 640×640 пікселів (рис. 3.11).

Таким чином, для обох моделей були забезпечені однакові умови навчання, що дозволяє в подальших розділах провести об'єктивне порівняння їх ефективності за рядом формальних метрик сегментації

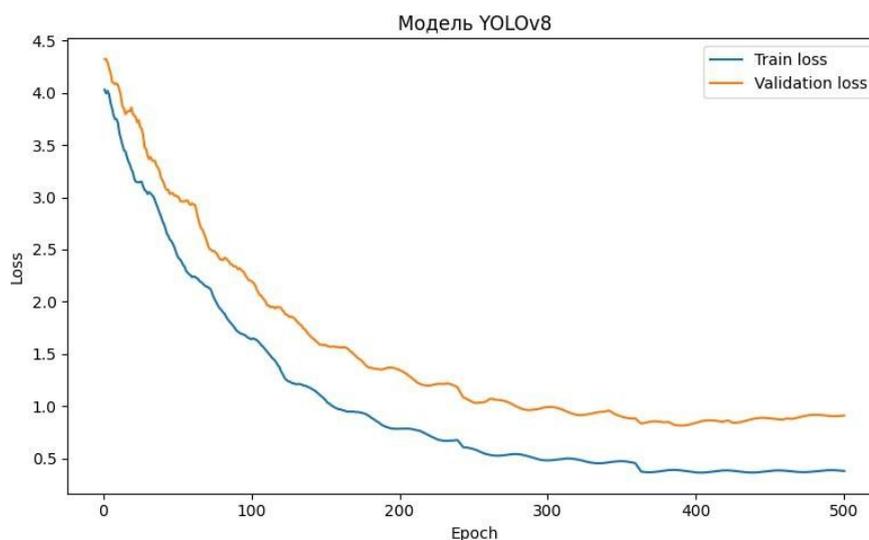


Рисунок 3.11. Графік тренування моделі YOLOv8

На графіку видно, що значення *train loss* поступово знижувалося, але вже на 200–250 епісі темп зменшення сповільнився. *Validation loss* стабілізувався раніше, приблизно на рівні 0.10–0.12, і залишався на цьому рівні до кінця тренування, не демонструючи суттєвих погіршень. Це вказує на відсутність критичного переобучення.

Водночас загальний рівень *validation loss* був вищим, ніж у U-Net

3.5. Метрики для порівняння

Завдання сегментації полягає у правильному визначенні меж об'єктів на зображенні. У нашому випадку це - сегментація коронарних артерій, тобто відділення пікселів, що належать судинам, від фону (або, якщо розглядається багатокласовий варіант, від решти анатомічних структур).

1. IoU (Intersection over Union). Для бінарного випадку (наприклад, «судина» vs «фон») визначається як відношення перетину передбаченого сегмента і «істинного» сегмента до їх об'єднання:

$$IoU = \frac{|P \cap G|}{|P \cup G|}$$

У багатокласовій постановці зазвичай застосовують мікро- або макрорішення (мікро-IoU та mIoU). Чим ближче IoU до 1, тим точніше виконана сегментація.

2. F1-score (F1, або Dice Coefficient)

Для бінарної сегментації F1 збігається з коефіцієнтом Dice:

$$Dice = F1 = \frac{2 \times TP}{2 \times TP + FP + FN}$$

де TP – кількість пікселів, правильно віднесених до класу «об'єкт» (true positives), FP - хибно позначені пікселі (false positives), а FN - втрачені об'єктні пікселі (false negatives). Метрика F1 добре відображає баланс між Precision та Recall.

3. Precision (Точність)

Характеризує, яка частка пікселів із передбаченого сегмента справді належить об'єкту. У формулі:

$$Precision = \frac{TP}{TP + FP}$$

Якщо Precision низька, це означає, що сегментація «захоплює» забагато зайвих ділянок (хибні спрацьовування).

4. Recall (Повнота)

Показує, яку частку реальних пікселів об'єкта ми змогли виявити:

$$Recall = \frac{TP}{TP + FN}$$

Якщо Recall низька, це свідчить, що модель пропускає частину «правильних» пікселів, тобто недосегментовує об'єкт.

3.6. Аналіз якості сегментації коронарних артерій

3.6.1. Багатокласова сегментація

У багатокласовій постановці задачі обидві моделі — YOLOv8 та U-Net — були навчені виконувати семантичну сегментацію 25 класів коронарних артерій згідно з методологією Syntax Score (клас «стеноз» було виключено з навчання та оцінки). Для оцінки якості передбачень використовували чотири метрики: середній IoU (Mean Intersection over Union), середній F1-score (Dice), середню точність (Precision) та повноту (Recall). Результати наведено у таблиці 3.2.

Таблиця 3.2. Результати багатокласової сегментації коронарних артерій

Модель	Mean IoU	Mean Dice	Mean Precision	Mean Recall
YOLOv8	0.2020	0.3362	0.4100	0.2850
U-Net	0.2646	0.3783	0.4722	0.3414

Аналіз отриманих метрик демонструє перевагу моделі U-Net над YOLOv8 у задачі багатокласової сегментації.

Зокрема, середній показник IoU моделі U-Net на $\approx 30\%$ вищий, що свідчить про точніше накладення передбачених масок на реальні сегменти. Подібне зростання спостерігається також у метриках Dice (0.3783 проти 0.3362) та Precision (0.4722 проти 0.4100). Найбільш суттєве поліпшення видно за метрикою Recall, яка у випадку U-Net становить 0.3414, тоді як у YOLOv8 — лише 0.2850, що свідчить про кращу здатність моделі U-Net виявляти менші судинні структури.

Загалом, результати підтверджують, що модель U-Net завдяки глибшій енкодер-декодерній архітектурі краще справляється із задачами, де необхідна точна піксельна локалізація багатьох анатомічних структур.

Натомість YOLOv8 у стандартній конфігурації, навіть при використанні сегментаційного варіанту, орієнтована переважно на об'єктну детекцію, що може впливати на зниження точності при детальному розрізненні подібних класів у складному ангиографічному контексті.

3.6.2. Бінарна сегментація

У бінарній постановці задачі сегментація зводиться до виявлення судин проти фону, без диференціації на окремі класи - «артерія» (класи 1..24) та «не артерія» (клас 0 і 25).

Таке спрощення дозволяє зменшити складність задачі, підвищити стійкість до помилок класифікації суміжних класів та поліпшити узгодженість сегментованих областей. Порівняльні результати моделей YOLOv8 і U-Net наведено у таблиці 3.3.

Таблиця 3.3. Результати бінарної сегментації коронарних артерій

Модель	Mean IoU	Mean Dice	Mean Precision	Mean Recall
YOLOv8	0.3917	0.5628	0.6800	0.4800
U-Net	0.5635	0.7208	0.8398	0.6313

Під час тестування виявлено, що обидві моделі демонструють покращення метрик у бінарному випадку порівняно з багатокласовим.

Проте модель U-Net знову продемонструвала помітну перевагу над YOLOv8. Зокрема, IoU зріс до 0.5635, що є хорошим показником для задач медичної сегментації. Показник Dice, який є збалансованою метрикою для оцінки якості перекриття передбаченої та істинної області, перевищив 0.72 — значення, яке вважається прийнятним для попереднього автоматичного аналізу медичних зображень. Також модель U-Net продемонструвала високу точність (0.8398) і суттєво вищу повноту (0.6313), що свідчить про її здатність виявляти більшість судин, зменшуючи ризик пропуску вузьких або менш контрастних ділянок.

Таким чином, бінарна сегментація виявляється не лише менш ресурсоемною, а й більш стабільною для задачі виявлення коронарних артерій. Модель U-Net показала кращі результати за всіма метриками, що підтверджує її ефективність як базового рішення в задачах попередньої сегментації судин.

3.7. Постановка задачі сегментації стенозу

3.7.1. Обраний метод для сегментації

З огляду на проведене порівняльне тестування було встановлено, що модель U-Net із бінарним підходом забезпечує найкращі результати за всіма ключовими метриками сегментації. Зокрема, в умовах бінарної постановки задача спрощується до виявлення судинного русла без необхідності розмежування анатомічних сегментів, що знижує вплив класової

незбалансованості та складності розмітки. Отримані значення середнього показника IoU (0.5635), Dice-коефіцієнта (0.7208), Precision (0.8398) і Recall (0.6313) суттєво перевищують відповідні метрики моделі YOLOv8, як у багатокласовому, так і в бінарному режимі.

Остаточне рішення на користь бінарного підходу також зумовлене прикладним характером подальших етапів обробки. Оскільки наступна задача — сегментація та класифікація стенозу — не потребує знання точного анатомічного класу кожної ділянки артерії, а лише визначення її належності до судинного русла, деталізація у вигляді багатокласових масок є надлишковою. Таким чином, бінарна сегментація у поєднанні з архітектурою U-Net була визнана найбільш ефективною, стабільною та доцільною з погляду загальної системної архітектури.

3.7.2. Розв'язання задачі сегментації коронарних артерій

У межах даного розділу було поставлено низку взаємопов'язаних завдань, що охоплювали системний аналіз підходів до сегментації коронарних артерій, вибір відповідної архітектури глибокого навчання, реалізацію обох сценаріїв (багатокласового та бінарного), їх навчання на єдиному датасеті, тестування та порівняльну оцінку ефективності.

Перш за все було проведено аналітичний огляд сучасних наукових публікацій, зокрема учасників змагання ARCADE 2023, що дозволило виокремити дві архітектури як базові для подальших експериментів: YOLOv8 — як приклад швидкодіючої одностадійної моделі з підтримкою сегментації, та UNet — як класичну архітектуру, адаптовану для високоточних задач медичної сегментації. Додатково були обґрунтовані два підходи до постановки задачі: багатокласовий, із віднесенням кожного пікселя до конкретної анатомічної зони, та бінарний, який зводиться до розпізнавання судинної області в цілому.

Для забезпечення справедливості експериментів, обидві моделі навчалися на єдиному датасеті ARCADE, що включає 1500 анотованих зображень із масками коронарних артерій. Попередня обробка охоплювала фільтрацію шумів (DnCNN), нормалізацію інтенсивностей (Min-Max) і

покращення контрасту (глобальне вирівнювання гістограми). Після цього 1000 зображень було використано для тренування, 200 — для валідації, а 300 — для незалежного тестування.

На етапі навчання до моделей застосовувалися однакові стратегії аугментації та обробки, що дозволило мінімізувати вплив зовнішніх факторів на результати. Навчання здійснювалося протягом 500 епох.

Результати тестування демонструють чітку перевагу бінарного підходу над багатокласовим як у випадку YOLOv8, так і U-Net. Для моделі U-Net у бінарному варіанті було досягнуто найвищих значень середніх метрик: IoU = 0.5635, Dice = 0.7208, Precision = 0.8398, Recall = 0.6313. Навпаки, багатокласова реалізація моделі YOLOv8 дала найгірші показники (IoU = 0.2020, Dice = 0.3362). Це підтверджує гіпотезу, що в умовах низького контрасту і слабкої роздільності між анатомічними класами, спрощення задачі до бінарної дозволяє покращити якість результату.

3.7.3. Автоматизоване виявлення стенозу

Автоматизоване виявлення стенозу на рентгенівських зображеннях коронарографії є одним із ключових викликів у побудові систем комп'ютерної підтримки діагностики ішемічної хвороби серця. На відміну від задачі сегментації коронарних артерій, яка фокусується на реконструкції топології судинного дерева, задача сегментації стенозу вимагає точного локалізованого виділення патологічних звужень, які часто мають невиражені контури, змінну інтенсивність та відсутність чіткої геометричної структури.

З формальної точки зору, вхідне зображення можна представити як тензор $X \in R^{H \times W \times 3}$, де H – висота зображення, W – ширина, а кожен піксель містить три кольорові компоненти (RGB). Метою сегментації є побудова функції $f_{\theta}(X) \rightarrow Y$, яка для кожного пікселя $x \in X$ повертає мітку $y \in \{0, 1, \dots, C - 1\}$, де C — кількість класів, включаючи клас стенозу. У випадку бінарної постановки задачі, класом 1 вважається лише патологічна ділянка

звуження (наприклад, label=25 в анотації ARCADE), а всі інші пікселі трактуються як фон або здорові структури.

Особливу складність становить дисбаланс класів: пікселі, що відповідають стенозу, складають незначну частину зображення, тоді як фон і артеріальна структура переважають. Це вимагає застосування спеціальних стратегій навчання, таких як балансування втрат, агресивна аугментація або залучення напівконтрольованого навчання.

У межах даного дослідження завдання сегментації стенозу включає наступні ключові етапи:

- здійснення системного аналізу сучасних моделей сегментації медичних зображень із фокусом на ділянки стенозу;
- побудова експериментального середовища для навчання моделей на основі U-Net та YOLOv8 у двох режимах: повністю контрольованому та із використанням псевдоанотацій;
- порівняння ефективності моделей за основними метриками (IoU, F1score, Precision, Recall), де позитивним класом вважається лише стеноз;
- реалізація процедур генерації псевдоміток та морфологічної післяобробки для розширення тренувального набору без ручної розмітки;
- визначення найбільш придатної архітектури та стратегії навчання для задачі високоточного сегментування стенозу на основі порівняльного аналізу результатів.

Таким чином, задача сегментації стенозу є критичним етапом у побудові повноцінної системи діагностики серцево-судинних захворювань і вимагає поєднання точності, чутливості та анатомічної узгодженості отриманих результатів.

3.8. Виділення ознак стенозу по зображенню

Сегментація коронарних артерій на рентгенівських ангиограмах дозволяє виділити топологічну структуру судинного русла, однак наступним

критичним завданням є виявлення ділянок звуження просвіту — стенозу (рис. 3.12). Автоматизоване виявлення стенозних уражень базується на виділенні дескрипторів, що характеризують геометричні, інтенсивнісні та текстурні ознаки судинної стінки в межах підозрілої ділянки. Такі дескриптори можуть бути подані у вигляді векторів ознак, які формуються для кожного локального фрагмента судини та далі подаються на вхід моделей класифікації або сегментації.

3.8.1. Геометричні ознаки звуження судин

Поширеним рішенням є оцінювання локальних змін геометрії просвіту судини. У найпростішому випадку проводять вимір поперечного діаметра артерії в різних точках і виявляють аномально малі величини або різкі градієнти діаметра (рис. 3.13). Формалізованим видом такої процедури є обчислення функції

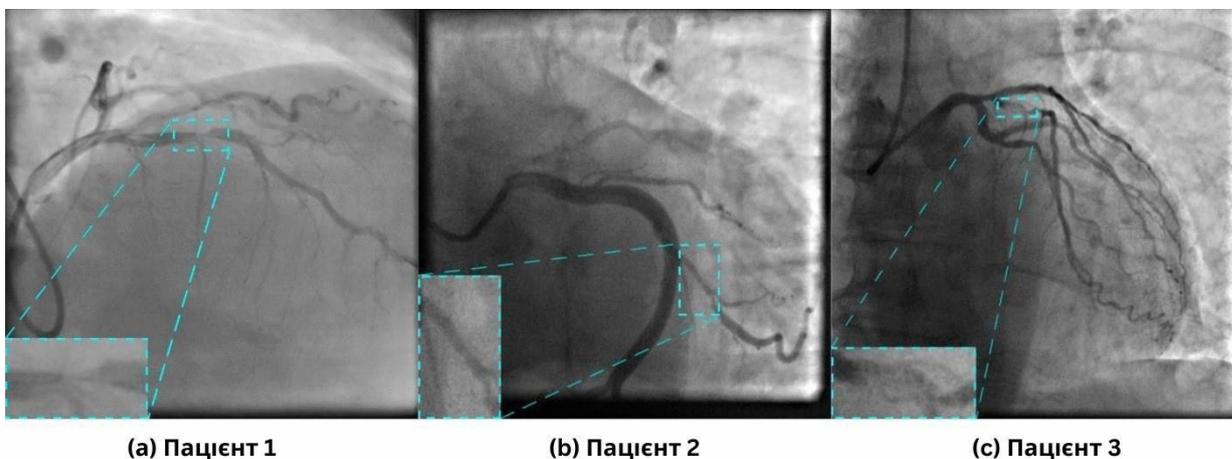


Рисунок 3.12. Маркування даних вихідних зображень з позначенням виявлених стенотичних уражень¹⁹

$$d(p) = ||\Omega_p||,$$

¹⁹ Джерело зображення: https://media.springernature.com/full/springerstatic/image/art%3A10.1038%2Fs41598-021-871742/MediaObjects/41598_2021_87174_Fig1_HTML.jpg

де $d(p)$ — діаметр судини в точці p , Ω_p — поперечний переріз судини в околі цієї точки, а $|\Omega_p|$ - ширина судини в околі точки p . Стеноз реєструється у випадку, коли локальний діаметр різко зменшується відносно сусідніх ділянок.

Для формалізації цього спостереження вводять нормоване значення:

$$r(p) = \frac{d(p)}{\bar{d}_{segment}}$$

де $\bar{d}_{segment}$ — середній діаметр на референтному сегменті артерії. Якщо $r(p) <$

τ , де τ — заданий поріг (наприклад, 0.7), точка позначається як підозріла на стеноз.

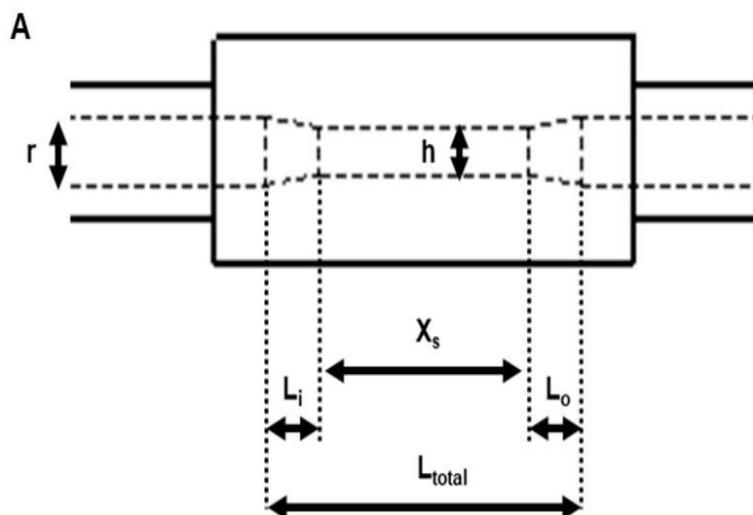


Рисунок 3.13. (А) Схематичне зображення моделі стенозу.

h - діаметр стенозованої судини. L_i - довжина входу від нормальної судини до передньої частини ділянки постійного стенозу. L_o - довжина виходу від ділянки постійного стенозу до нормальної судини. L_{total} - загальна довжина

стенозу. r - діаметр нормальної судини (3 мм); X_s - довжина ділянки постійного стенозу²⁰

Подібний метод дозволяє виявляти стенози без необхідності використання контрастних характеристик, що особливо цінно у випадках нерівномірного заповнення судин контрастною речовиною.

3.8.2. Інтенсивнісні та текстурні характеристики

Окрім геометричних змін, стенозні ураження зазвичай супроводжуються змінами в інтенсивності зображення. Це зумовлено порушенням проникнення контрастної речовини крізь звужену ділянку та, відповідно, локальною зміною рівня яскравості. Для автоматичного виявлення таких змін використовуються статистичні та текстурні дескриптори, сформовані в околі центральної лінії судини.

При цьому можуть розраховуватись:

- локальні середні інтенсивності в масці фіксованого радіуса;
- дисперсія інтенсивності в околі (локальна варіація);
- градієнт уздовж радіального напрямку — від центру судини до краю;
- текстурні характеристики на основі матриці спільної появи сірого рівня (GLCM): контраст, енергія, гомогенність.

Одним із методів виявлення таких змін є використання фільтра Франжі (Frangi vesselness filter) [24], який базується на аналізі другого порядку частинних похідних (матриці Гессе). Його ціль — виділення трубчастих структур шляхом підсилення витягнутих об'єктів із високою локальною неоднорідністю (рис. 3.14).

Судинна відгукова функція $V(x, y)$ фільтра Франжі може бути визначена як:

²⁰ Джерело зображення:

<https://www.researchgate.net/publication/382009235/figure/fig1/AS:11431281386251419@1745069245029/Schematic-diagram-of-stenosis-model-and-in-vitro-coronary-circulation-system-A.tif>

$$V(x, y) = \exp\left(-\frac{R_B^2}{2\beta^2}\right) \cdot \left(1 - \exp\left(-\frac{S^2}{2c^2}\right)\right)$$

де $R_B = \lambda_2$ - співвідношення власних значень матриці Гессе, що характеризує форму об'єкта, $S = \sqrt{\lambda_1^2 + \lambda_2^2}$ — міра контрастності, а λ_1, λ_2 — власні значення (упорядковані за модулем). Параметри β і c керують чутливістю фільтра.

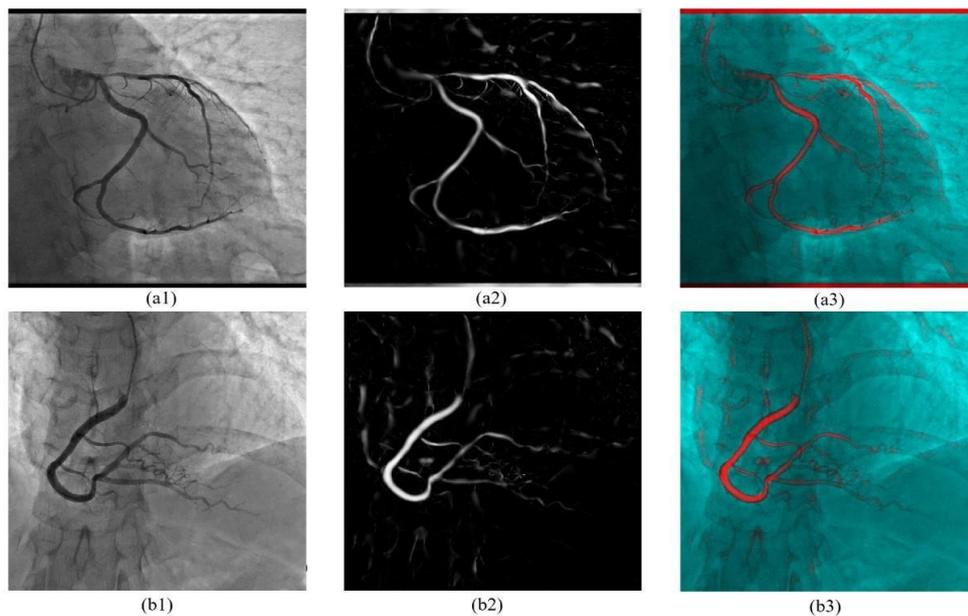


Рисунок 3.14. Оригінальні зображення ЛКА та ПКА відповідно (a1,b1), зображення, покращені з використанням фільтра Франжі (a2,b2), результати після поєднання вихідного зображення та відповіді фільтра для покращення візуалізації (a3,b3) [36]

У поєднанні з аналізом інтенсивності, Frangi-фільтр дозволяє не лише виділити судини, а й оцінити локальні варіації яскравості, що часто співпадають із ділянками патологічного звуження.

Комбінування інтенсивнісних і геометричних дескрипторів значно покращує точність автоматичного виявлення стенозу, особливо в умовах неоднорідного освітлення або артефактів.

3.9. Огляд сучасних методів та підходів до сегментації стенозу

3.9.1. StenUNet для точної сегментації стенозів

StenUNet – UNet-архітектура для сегментації стенозів [25]. У роботі запропоновано метод StenUNet для автоматичного виявлення стенозів на рентгенівських коронарографічних зображеннях (рис. 3.15).

Архітектура StenUNet побудована на базі U-образної нейронної мережі (UNet) із використанням платформи nnU-Net, що забезпечує ефективне поєднання глибоких ознак і просторової деталізації за допомогою пропускних з'єднань.

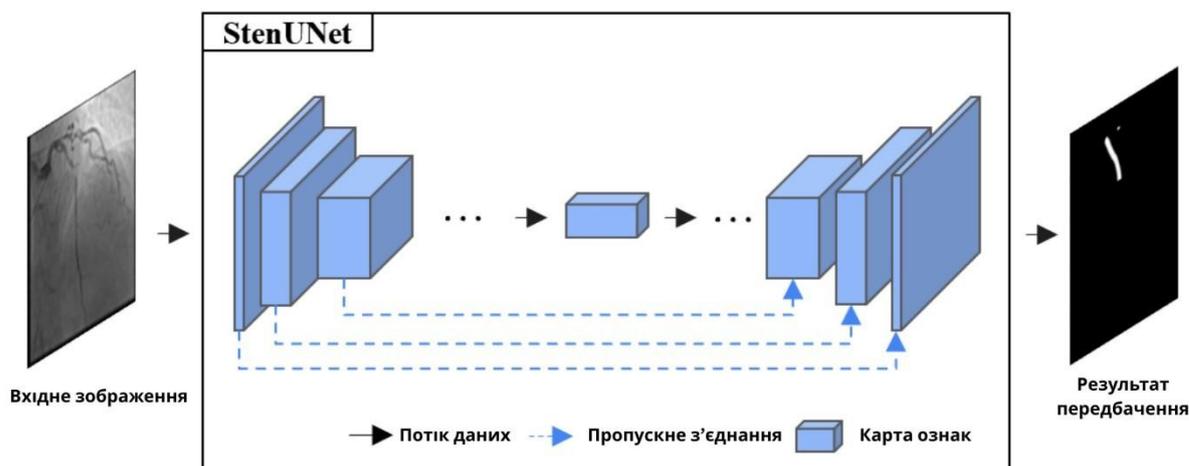


Рисунок 3.15. Запропонована система StenUNet на основі UNet для виявлення стенозу [25]

Для підготовки даних застосовано попередню обробку: гомоморфне підсилення контрасту з Butterworth-фільтрами для зниження шуму і відблисків, а також нормалізацію яскравості. Розмітка стенозів виконувалася на основі датасету ARCADE – публічної вибірки коронарних ангиограм із піксельними анотаціями ділянок стенозу.

Модель навчається сегментувати області звуження (стенози) на зображенні та генерує бінарну маску стенозів (піксельна сегментація) замість грубих рамок навколо уражень. Після основної сегментації запроваджено

спеціальну постобробку для усунення хибнопозитивних фрагментів, що не відповідають анатомії судин.

Запропонований підхід продемонстрував високий результат - модель StenUNet посіла 3-тє місце у конкурсі ARCADE ($F1 \approx 53.5\%$ на тестовому наборі), лише на 0.05% поступившись другому місцю (табл. 3.4).

Таблиця 3.4. Результати моделі StenUNet з допрацюванням та без [25]

Модель	F1 без допрацювання	F1 з преобробкою	F1 з пре- та постобробкою
StenUNet	48.34%	49.60%	53.00% ($\approx 53.48\%$)

Перевагами підходу StenUNet є можливість виконання сегментації стенозів на рівні пікселів, що забезпечує точне визначення форми та розміру патологічних змін, на відміну від підходів, які формують лише прямокутні рамки. Це особливо важливо для клінічного оцінювання ступеня тяжкості звужень. Використання архітектури U-Net із пропусковими з'єднаннями дозволяє спростити навчання моделі, уникнути проблеми зникаючого градієнта та пришвидшити процес обробки зображень. Додатковим позитивним аспектом є вбудовані механізми попередньої обробки зображень і постобробки результатів, що підвищують загальну надійність сегментації.

Серед недоліків моделі StenUNet варто зазначити можливість виникнення помилкових позитивних результатів у випадках, коли судини перетинаються або накладаються одна на одну. Щільні ділянки, утворені в місцях перетину, модель іноді помилково ідентифікує як стенози. Крім того, модель обмежена у функціональності, оскільки визначає лише місце розташування звужень без подальшої кількісної оцінки тяжкості стенозу, зокрема відсутня оцінка проценту звуження просвіту судини, що є критично важливим у клінічній практиці.

3.9.2. SSASS: Напівконтрольоване навчання для сегментації

Інша робота пропонує підхід SSASS (Semi-Supervised Approach for Stenosis Segmentation) [26], спрямований на підвищення точності сегментації стенозів шляхом залучення невручну розмічених даних.

Основна ідея – комбінувати обмежену розмічену вибірку з великою кількістю псевдорозмічених прикладів (рис. 3.16, модель YOLOv8-m

спочатку навчається на наборі даних сегментації стенозів. Навчена модель використовується для генерації псевдоміток в наборі даних сегментації судин. Фінальна модель YOLOv8-m навчається як на наборі даних сегментації стенозів, так і на наборі даних сегментації псевдоміток судин.). Автори спочатку виконують спеціалізовану аугментацію зображень коронарографії, щоб штучно збільшити різноманітність прикладів судинних структур.

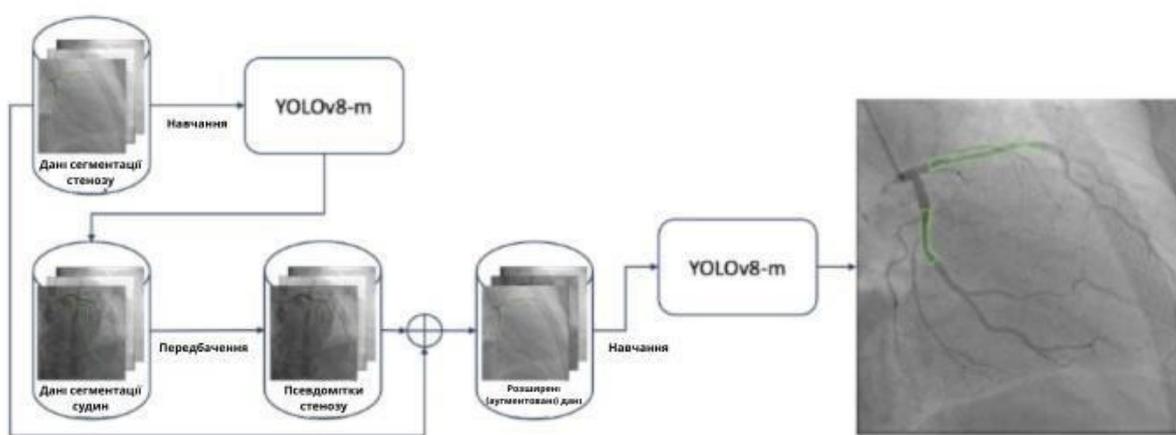


Рисунок 3.16. Огляд напівконтрольованого конвеєра навчання для сегментації стенозів [26]

На першому етапі середнього розміру модель YOLOv8-m навчається лише на базі наявних розмічених зображень стенозів.

Далі, використовуючи цю початкову модель, генерують псевдо-мітки стенозів на іншому наборі ангіограм, де вручну розмічені були лише судини (без зазначення стенозів). Таким чином, формується розширений набір тренувальних даних: оригінальні розмічені зображення та додаткові зображення з псевдорозміткою стенозів.

На заключному етапі модель YOLOv8-m пере навчається на об'єднаному датасеті, поєднуючи обидві вибірки, щоб покращити здатність знаходити стенози в різноманітних умовах. Запропонований напівнаглядний підхід дозволив суттєво покращити показники сегментації без використання ансамблю моделей. На конкурсі ARCADE дана методика продемонструвала

один із найкращих результатів, що особливо примітно з огляду на використання лише однієї моделі замість композиції кількох (табл. 3.5).

Таблиця 3.5. Порівняння результатів різних моделей YOLOv8 [26]

Модель (YOLOv8)	Повністю навчена (F1)	Напівнавчена (F1)
Nano (v8n)	49.1%	50.7%
Small (v8s)	51.5%	52.0%
Medium (v8m)	52.0%	53.6%
Large (v8l)	52.6%	53.0%

Перевагами підходу SSASS є досягнення конкурентної точності сегментації завдяки напівконтрольованому навчанню, яке навіть перевершило результати повністю контрольованого підходу (середній F1 зріс з приблизно 52% до 53,6% на тестових даних ARCADE). Використання різноманітних перетворень зображень (віддзеркалення, зсув, повороти, зміна кольорового тону) дозволяє ефективно моделювати реальні зміни проєкції і контрастності коронарних судин. Важливою перевагою є базування моделі на сучасній і швидкій архітектурі YOLOv8, що робить її придатною до застосування у режимі, близькому до реального часу. Відмова від ансамблевого підходу значно спрощує процес інтеграції системи в клінічні умови.

До недоліків методу SSASS слід віднести ризик внесення хибних даних до процесу навчання через використання псевдоміток, що можуть закріпити або навіть підсилити початкові помилки генерації міток моделлю. Зокрема, автори вказують, що зниження порогу впевненості для псевдоміток хоча й збільшує кількість тренувальних даних, але може призвести до перенавчання. Крім того, залучення до навчального набору зображень без стенозів не поліпшило якість роботи моделі, що вказує на недостатню здатність підходу ефективно розрізняти нормальні ангіограми від патологічних, що є важливим аспектом у клінічній практиці.

3.9.3. Порівняльний аналіз ефективності моделей YOLO та DETR у задачі сегментації стенозів

Третя розглянута робота присвячена аналізу різних підходів до автоматичного виявлення стенозів із застосуванням методів глибокого навчання на тому ж датасеті ARCADE [27].

На відміну від попередніх двох підходів, тут досліджується задача детекції стенозів як об'єктів (з використанням обмежуючих рамок), а не піксельної сегментації уражених областей.

Автори оцінили три сучасні моделі: згорткову нейромережу YOLO (You Only Look Once) для швидкого виявлення об'єктів, а також два трансформерні підходи – DINO-DETR і Grounding DINO, які представляють сімейство DETR (Detection Transformer) з механізмами самоуваги (рис. 3.17).

Для навчання і перевірки використано дані ARCADE: вихідні маски стенозів були конвертовані у формат обмежуючих прямокутників, щоб слугувати мітками для моделей детекції.

Оцінювання виконано за метриками COCO: середня точність виявлення (mAP) та середня повнота (AR) обчислювалися на різних порогах перекриття (IoU) між передбаченими рамками та істинними областями стенозу (табл. 3.6).

Таблиця 3.6. Порівняння метрик моделей DINO-DETR, Grounding DINO та YOLO [27]

Модель	mAP@[0.5:0.95]	mAP@0.50	AR@[0.5:0.95]
DINO-DETR	0.086	0.228	0.526
Grounding DINO	0.080	0.259	0.416
YOLO	0.068	0.254	0.180

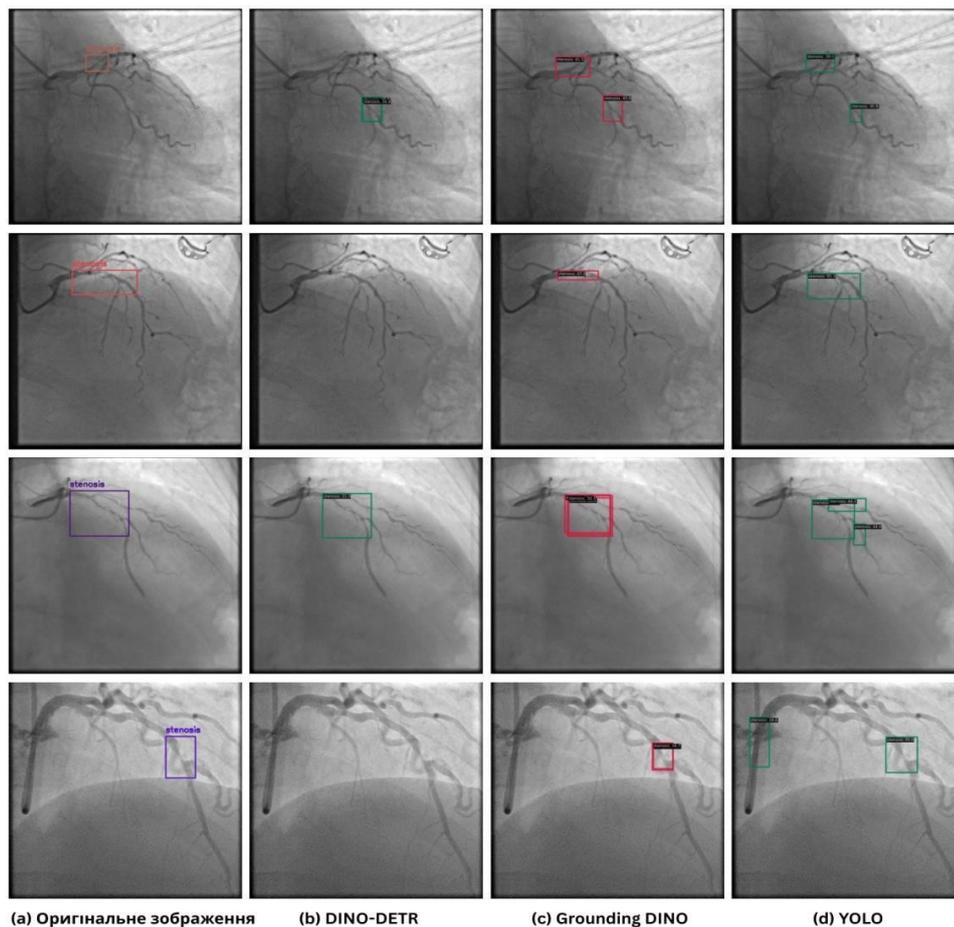


Рисунок 3.17. Якісне порівняння оригінальних тестових зображень (перша колонка) з результатами виявлення за допомогою DINO-DETR, Grounding DINO та YOLO [27]

Перевагами підходів на основі DETR є досягнення трансформерними моделями вищої середньої точності (mAP) порівняно з традиційними методами, такими як YOLO. Зокрема, модель DINO-DETR продемонструвала найвищий показник усередненого mAP у діапазоні IoU від 0.5 до 0.95 (≈ 0.086), що свідчить про кращу здатність архітектур з механізмом самоуваги до ідентифікації стенозів різних масштабів і форм. Особливо значним є успіх моделі у виявленні малих об'єктів, де DINO-DETR істотно перевершує YOLO (mAP ≈ 0.198 проти 0.126), завдяки багаторівневому

представленню ознак, що дозволяє ефективно знаходити тонкі й ледь помітні стенози.

Однак детекційні підходи мають певні недоліки й обмеження. Незважаючи на переваги трансформерних моделей, абсолютні значення точності залишаються низькими (максимальна точність для IoU=0.5 близько 25–26%, а середня точність близько 8%), що вказує на складність медичних ангиографічних зображень для автоматизованого аналізу через шум, проєкційні накладання і динамічний контраст. Крім того, використання обмежувальних прямокутників для локалізації стенозів не дозволяє чітко окреслити патологічні ділянки та може включати здорові сегменти судин або фон, що ускладнює точну кількісну оцінку тяжкості звуження. Нарешті, трансформерні моделі характеризуються великою кількістю параметрів і підвищеними вимогами до обчислювальних ресурсів, що вимагає значних інженерних зусиль для розгортання і застосування в клінічних умовах.

3.10. Методи та підходи, застосовані в роботі

У межах даного розділу особлива увага приділяється застосуванню глибоких нейронних мереж для задач виявлення та сегментації стенозу на рентгенівських ангиограмах. З огляду на результати попереднього аналізу та успішність архітектур у задачах сегментації коронарних артерій, було обрано дві моделі: YOLOv8 та U-Net. Їх використання дозволяє реалізувати як піксельну сегментацію зображень, так і визначення регіонів інтересу з можливістю наступної класифікації.

Обидві архітектури поєднують високу гнучкість, адаптивність до специфіки вхідних даних і здатність виявляти слабо виражені анатомічні структури. Крім вибору моделей, у межах експериментальної частини розглядалися два принципово різні підходи до навчання: класичний (із використанням повністю анотованих даних) та підхід із використанням псевдоміток (semi-supervised learning). Такий поділ дає

змогу оцінити, наскільки ефективними можуть бути нейромережі навіть у випадках обмеженої кількості якісної розмітки.

3.10.1. Модель YOLOv8

Архітектура YOLOv8 була докладно розглянута в розділі 3.3.1 при вирішенні задачі сегментації коронарних артерій. Водночас, її застосування в контексті виявлення стенозу вимагає адаптації та розширення функціоналу. Зокрема, модель може працювати як у режимі сегментації, так і в режимі детекції із класифікацією обмежуючих рамок.

Для виявлення стенозних уражень модель налаштовується на виявлення зон звуження судинного русла. Це досягається як за допомогою тренування на бінарних масках стенозу (отриманих вручну або псевдоанотованих), так і в процесі детекції за координатами об'єктів. Важливо, що YOLOv8 дозволяє адаптувати вихідний head-блок до різних кількостей класів і масштабів об'єктів, що забезпечує гнучкість при роботі з анатомічними структурами різного розміру.

Серед ключових переваг YOLOv8 слід відзначити:

- швидке передбачення на GPU та CPU;
- ефективність при детекції великих контрастних об'єктів;
- можливість легкої інтеграції в реальні клінічні системи.

Проте, при роботі з тонкими стенозичними структурами модель може давати нижчу точність у порівнянні з методами піксельної сегментації.

3.10.2. Модель U-Net

Архітектура U-Net була детально представлена у розділі 3.3.2, де розглядалося її використання для сегментації коронарних артерій. Проте для задачі сегментації стенозу модель потребує певної адаптації. Основною модифікацією є зміна кількості вихідних каналів (у випадку бінарної

сегментації — лише один) та перегляд функції втрат залежно від типу навчальних даних (бінарна маска або мапа ймовірностей).

U-Net, завдяки своїй архітектурі з пропускними з'єднаннями (skip connections), дає змогу точно локалізувати дрібні ділянки стенозу, що особливо важливо у випадках з нечіткими межами або слабким контрастом.

При використанні U-Net для виявлення стенозів у цій роботі передбачено два режими тренування: на підготовлених ручних мітках та на псевдомасках.

3.10.3. Підхід із використанням псевдоміток

Один із підходів, який демонструє ефективність в умовах обмеженого обсягу ручно анотованих даних, є метод псевдо-розмітки (pseudo-labeling). Його суть полягає у використанні вже частково натренованої моделі для генерування автоматичних міток (псевдоміток) для необроблених зображень. Ці мітки формуються на основі передбачень моделі.

Формально, якщо позначити множину ручно розмічених зображень як $D_l = \{(x_i, y_i)\}_{i=1}^N$, а нерозмічені дані як $D_u = \{(x_j)\}_{j=1}^M$, то модель f_θ , попередньо навчена на D_l , застосовується до кожного $x_j \in D_u$ з метою побудови $\hat{y}_j = f_\theta(x_j)$. Якщо значення ймовірності $\max_c f_\theta(x_j)^c$ перевищує певний поріг τ , мітка \hat{y}_j приймається як надійна та додається до тренувального набору у вигляді пари (x_j, \hat{y}_j) .

Таким чином, створюється розширений набір даних:

$$D_{ext} = D_l \cup \{(x_j, \hat{y}_j) \mid \max_c f_\theta(x_j)^c > \tau\}$$

який використовується на наступному етапі повторного тренування моделі.

Цей підхід дозволяє залучити більший обсяг даних, знизити ймовірність переобучення та охопити більшу варіативність патологічних

випадків, що особливо актуально при аналізі рідкісних або слабо виражених стенозів.

3.10.4. Класичний підхід навчання

Альтернативою методу з псевдомітками є класичний підхід повного навчання (fully supervised learning), який передбачає використання виключно вручну розмічених зображень у ролі тренувального набору. У цьому випадку весь набір $D = \{(x_i, y_i)\}_{i=1}^N$ є апріорно достовірним і не вимагає постпроцесингу або фільтрації перед навчанням.

Такий підхід має низку переваг: високу якість вхідних міток, відсутність помилок, пов'язаних із довірою до передбачень моделі, а також стабільні метрики на валідації. З математичної точки зору, задача зводиться до мінімізації функції втрат L на всіх прикладах із відомими мітками:

$$\min_{\theta} \frac{1}{N} \sum_{i=1}^N L(f_{\theta}(x_i), y_i)$$

Однак, у випадку задач медичної сегментації, де ручна розмітка є трудомісткою, обсяг доступних прикладів часто обмежений. Це підвищує ризик недостатнього узагальнення та зменшує ефективність моделі на нових зображеннях, особливо коли стеноз представлений у рідкісній формі або в умовах поганого контрасту.

Таким чином, класичний підхід навчання демонструє високу точність за наявності якісного анотованого датасету, однак поступається методам із псевдомітками за гнучкістю, масштабованістю та здатністю охоплювати додаткові приклади без людського втручання.

3.11. Реалізація та тренування моделей

Етап реалізації методів сегментації стенозу передбачав побудову повноцінної тренувальної інфраструктури для двох обраних моделей — YOLOv8 та U-Net — із подальшим тестуванням у режимах повного та напівконтрольованого навчання. Особливу увагу було приділено якості

вхідних даних, процедурі попередньої обробки, генерації псевдоміток, агментаційним перетворенням, а також адаптації процесів до задачі виділення патологічно змінених ділянок артеріального русла.

Для досягнення порівнюваних умов навчання обидві моделі використовували однакові підмножини даних, синхронізовані маски та узгоджені гіперпараметри тренування. Це дозволило побудувати справедливе порівняння як за абсолютними метриками, так і за поведінкою моделей у складних випадках з неоднозначною локалізацією стенозу.

3.11.1. Вхідні дані та їх попередня обробка

Для тренування моделей сегментації стенозу використовувався відкритий датасет ARCADE (Automatic Region-based Coronary Artery Disease diagnostics using x-ray angiography imagEs), який складається з 1500 рентгенівських ангиографічних зображень, що супроводжуються багатокласовими масками сегментації. Анотації відповідають класифікації Syntax Score, де виділено 26 класів анатомічних регіонів, зокрема клас 25 — стеноз (рис. 3.18).

Загальний обсяг даних було поділено на три підмножини:

- 1000 зображень — для тренування;
- 200 зображень — для валідації;
- 300 зображень — для фінального тестування (ці приклади не використовувалися у жодній моделі під час навчання).

З метою покращення якості вхідного сигналу всі зображення проходили етапну попередню обробку, яка охоплювала:

- фільтрацію шумів за допомогою згорткової мережі DnCNN;
- нормалізацію інтенсивностей методом Min-Max, що узгоджувала динамічний діапазон яскравості;
- глобальне вирівнювання гістограми, яке покращувало локальний контраст в межах судинних структур.

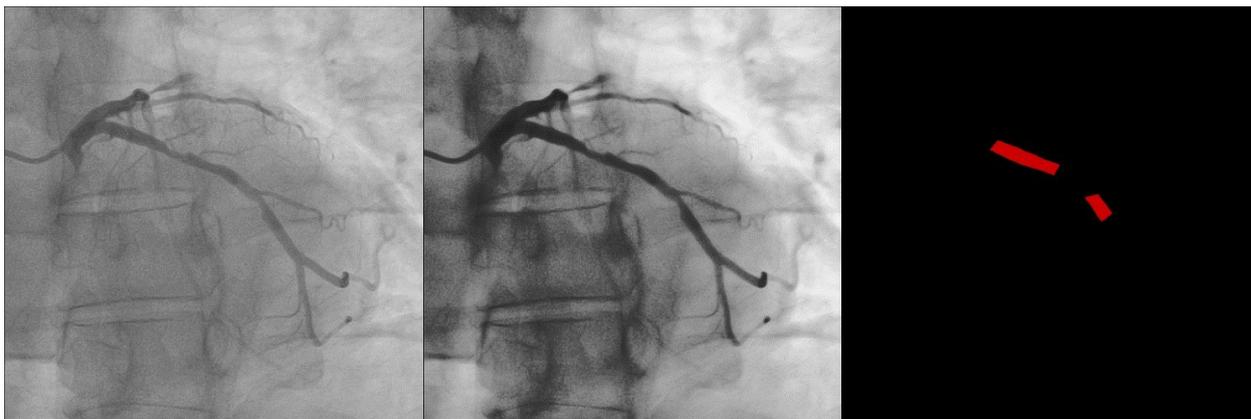


Рисунок 3.18. Приклад вхідних даних.

Ліворуч – необроблене рентгенівське зображення коронарної ратерії.

Посередині – рентгенівське зображення після попередньої обробки.

Праворуч – кольорова маска стенозу коронарної артерії.

Окрім стандартного контролю навчання, в роботі застосовувався напівконтрольований підхід (semi-supervised learning), що передбачає використання псевдомасок (pseudo-labels), сформованих за допомогою попередньо натренованих моделей.

Для цього було відібрано додаткову множину зображень без ручної анотації (умовно позначену як unlabeled). Попередньо натренована модель (UNet/YOLOv8 після 50 епох на ручних мітках) застосовувалась до цих зображень для генерації псевдоміток та псевдомасок.

Для підвищення достовірності використовувався порог впевненості — лише ті пікселі, ймовірність класу яких перевищувала задане значення (0.5), включались до маски - інші позначалися як фон.

Візуально або автоматично маски проходили морфологічну післяобробку, включно з очищенням від шумів, заповненням прогалин і згладжуванням меж.

Фінально, усі згенеровані псевдомітки об'єднувалися з наявними ручними мітками, утворюючи розширену тренувальну множину. Цей підхід

дозволив ефективно збільшити кількість прикладів без залучення додаткової ручної роботи, що особливо цінно для медичних задач із обмеженим обсягом верифікованих даних.

3.11.2. Аугментація

Зважаючи на обмежену варіативність ангиографічних даних, а також задля збільшення стійкості моделей до змін масштабу, орієнтації та контрасту, під час навчання обох архітектур (YOLOv8 та U-Net) застосовувався комплекс аугментаційних перетворень. Ці перетворення реалізовувалися на етапі підготовки батчів і охоплювали як геометричні, так і фотометричні трансформації:

Геометричні аугментації:

- горизонтальне та вертикальне віддзеркалення (із ймовірністю 0.5);
- масштабування (із коефіцієнтом до $\pm 50\%$);
- зсуви по осях X та Y (до 10%);
- обрізання фрагментів зображень (crop_fraction до 0.1).

Фотометричні аугментації:

- зміни яскравості (hsv_v), контрастності (hsv_s) та відтінку (hsv_h);
- випадкове змішування зображень (mixup=0.1);
- використання методу auto_augment для автоматичного підбору ефективних аугментацій.

Такі аугментації забезпечили підвищену генералізаційну здатність моделей, дозволяючи їм зберігати стабільність результатів навіть у випадках з нетиповою конфігурацією судинного русла чи погіршеним контрастом.

3.11.3. Навчання з псевдомітками

Одним із ключових підходів, застосованих у межах задачі сегментації стенозу, стало навчання з використанням псевдоміток (pseudo-labeling). Така стратегія належить до категорії напівконтрольованого навчання і дозволяє

ефективно використовувати великі обсяги неанотованих даних у поєднанні з обмеженим обсягом ручної розмітки. Методологія передбачає попереднє тренування моделі на обмеженій підмножині «labeled» даних (із повною ручною анотацією), подальше створення прогнозованих масок (псевдоміток) для решти зображень та їх подальше включення до навчального процесу.

Створені псевдомітки застосовуються у тих випадках, де модель демонструє достатню впевненість у прогнозі. Для цього визначався поріг імовірності (0.5), нижче якого передбачення не враховувалися. Крім того, отримані псевдомаски проходили морфологічну фільтрацію, зокрема видалення артефактних дрібних об'єктів, згладжування меж і уточнення контурів. Результатом став комбінований датасет, що включав як ручні, так і псевдоанотовані зображення, який використовувався для фінального тренування моделей.

Модель YOLOv8 сегментації (версія yolov8n-seg.pt) було адаптовано для задачі виявлення стенозу шляхом її тренування на розширеному наборі даних, що включав як ручні, так і псевдоанотовані приклади. Навчання проводилося з використанням конфігураційного файлу data_pseudo.yaml, у якому зазначалися шляхи до зображень та масок, отриманих з попереднього етапу генерації псевдоміток.

Процес навчання охоплював 500 епох, використовував розмір батчу 8 та роздільну здатність 640×640 пікселів.

Як показано на рисунку 3.19, крива train loss демонструє стабільне зниження протягом усього періоду навчання, особливо помітне в перші 100 епох, після чого вона поступово наближається до плато. Водночас validation loss знижувався з дещо меншою швидкістю.

Фінальна втрата на валідаційній множині становила ~ 0.92 , що підтверджує збереження здатності моделі до узагальнення. Незначні коливання валідаційної кривої в другій половині навчання пояснюються високою варіативністю зразків у тестовій підвбірці.

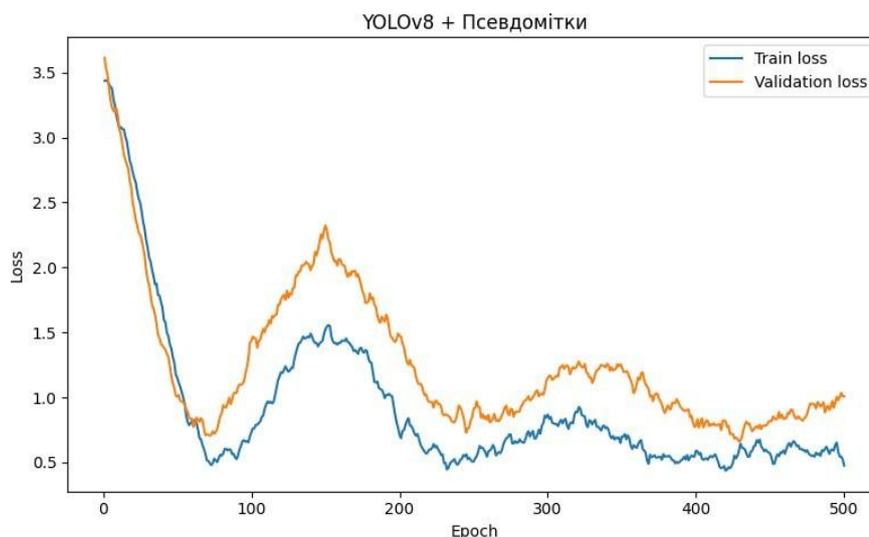


Рисунок 3.19. Графік тренування моделі YOLOv8 із використанням псевдоміток

Архітектура U-Net була адаптована для багатокласової сегментації з 26 вихідними каналами, серед яких окремо виділено клас стенозу. Модель тренувалася на об'єднаному датасеті із справжніми анотаціями (200 зображень) та псевдоанотованими масками, отриманими шляхом інференсу попередньо натренованої U-Net.

Графіки train/val loss, подані на рисунку 3.20, демонструють характерну поведінку: початкове зменшення втрат із поступовим уповільненням після 300-ї епохи. Train loss зменшувався до значення ~ 0.52 , а validation loss зупинився поблизу ~ 0.61 , що свідчить про баланс між переобученням і недостатнім навчанням. Враховуючи складність сегментації тонких ділянок стенозу, така динаміка свідчить про достатню узагальнюючу здатність.

Варто зазначити, що псевдоанотовані зображення сприяли кращому охопленню різноманітних форм стенозу, що дозволило моделі досягнути кращих результатів за F1-метрикою в порівнянні з класичним варіантом навчання.

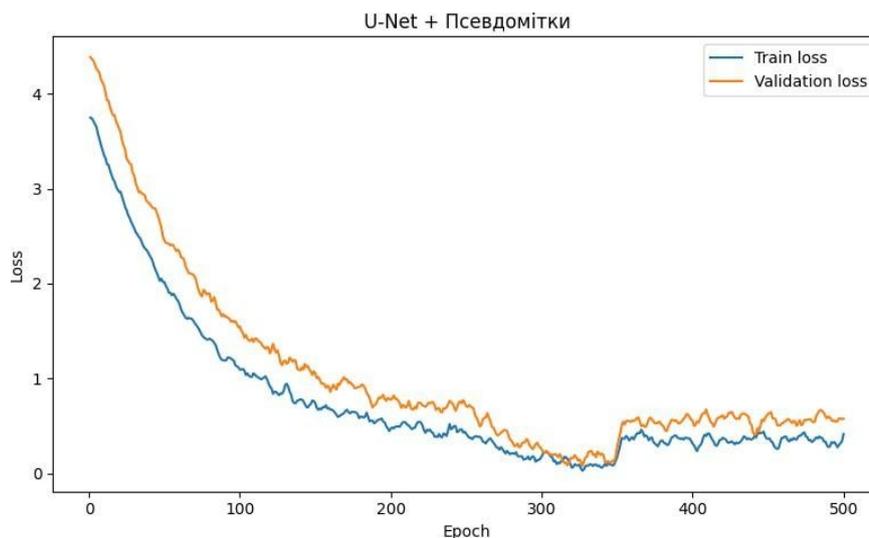


Рисунок 3.20. Графік тренування моделі U-Net із використанням псевдоміток

Таким чином, обидві моделі — як YOLOv8, так і U-Net — були адаптовані до роботи з комбінованим набором даних, що суттєво підвищило кількість тренувальних прикладів і дало змогу отримати покращені результати без розширення ручної розмітки.

3.11.4. Навчання з ручною розміткою

Класичний підхід до навчання глибоких моделей сегментації передбачає використання виключно якісно анотованих даних. У межах цього підрозділу розглянуто реалізацію повного навчання для двох архітектур — YOLOv8 та UNet — без залучення псевдомасок. Такий підхід дозволяє забезпечити контрольоване середовище для аналізу можливостей моделей за наявності верифікованих еталонних міток.

Початкове тренування здійснювалося на підмножині з 1000 повністю анотованих зображень, де маски включали клас стенозу (25) разом з іншими судинними структурами. Для валідації було використано 200 окремих зображень. Обидві моделі тренувалися на однакових підмножинах із застосуванням ідентичних гіперпараметрів (batch=8, epochs=500, lr=1e-3) для забезпечення коректного порівняння результатів.

Класичне тренування YOLOv8 здійснювалося на основі виключно вручну анотованих масок стенозу. Конфігурація була ідентичною попередньому варіанту, однак у зв'язку з меншою кількістю навчальних прикладів (відсутність псевдомасок), модель виявила дещо вищу чутливість до варіативності валідаційної множини.

Як видно з рис. 3.21, train loss стрімко зменшувався до значення ~ 0.55 , проте validation loss зупинився на рівні ~ 1.04 , що є дещо гіршим, ніж у випадку з псевдомітками. Це може бути пов'язано з тим, що обмежена кількість ручних прикладів не дозволила моделі охопити всю складність розподілу класу стенозу.

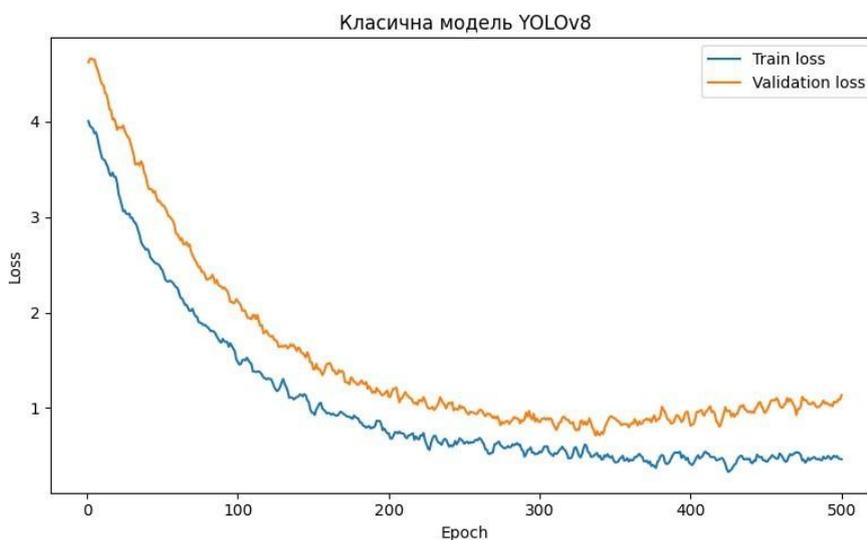


Рисунок 3.21. Графік тренування моделі YOLOv8 класичним підходом

Незважаючи на застосування агресивної аугментації (flip, mixup, HSV, sgr), відносно висока фінальна втрата свідчить про потребу в розширенні обсягу розмітки для більшої точності.

У класичному режимі U-Net тренувалася на 200 вручну анотованих зображеннях з розміткою класу стенозу. Навчання тривало 500 епох, із використанням оптимізатора Adam та функції втрат CrossEntropyLoss.

На рисунку 3.22 зображено криві втрат: train loss знизився до ~ 0.49 , тоді як validation loss коливався поблизу ~ 0.68 . Незважаючи на візуально

кращу поведінку train loss, показники на тестових даних свідчили про недоохоплення менш поширених патернів стенозу.

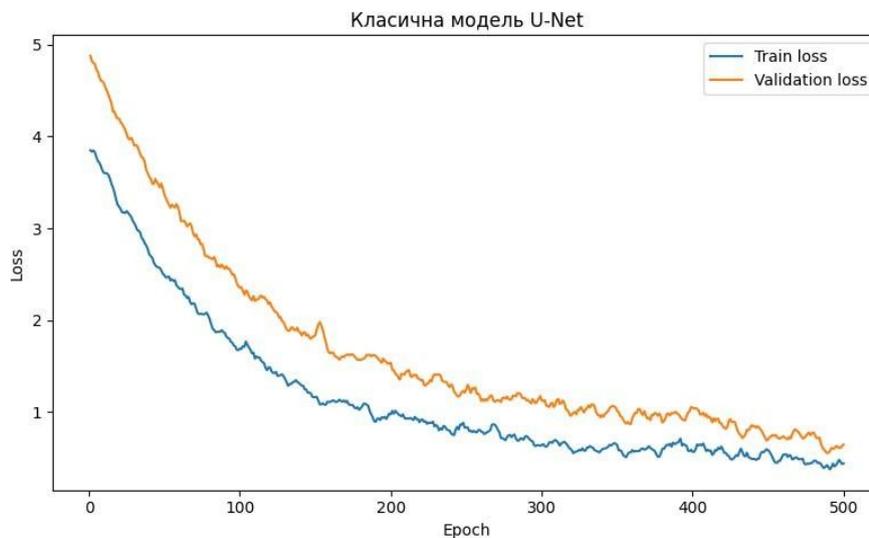


Рисунок 3.22. Графік тренування моделі U-Net класичним підходом

Невелика величина валідаційної втрати підтверджує, що модель навчилася ефективно розпізнавати найтипівіші випадки стенозу, однак без застосування псевдоміток її узагальнююча здатність залишалася обмеженою, що й проявилось в нижчих метриках F1 та IoU на етапі тестування.

3.12. Оцінка якості сегментації

3.12.1. Метрики оцінки якості сегментації

Оцінювання ефективності моделей сегментації стенозу здійснюється за тим самим формалізмом, що й у випадку сегментації коронарних артерій. Оскільки і в цій підзадачі основним є точне відтворення просторової структури судинних уражень, застосування класичних метричних підходів дозволяє забезпечити об'єктивне порівняння результатів, незалежно від використаної архітектури або стратегії навчання.

У межах цього дослідження для обох типів моделей (YOLOv8 та U-Net) та для обох режимів навчання (з ручною розміткою й із залученням псевдоміток) застосовувались чотири основні метрики якості: IoU

(Intersection over Union), F1score (Dice Coefficient), Precision та Recall. Формальні означення й аналітичні вирази для кожної з них вже були наведені раніше.

Метрика IoU слугує узагальненим показником ступеня перекриття між передбаченою маскою та еталонною (ground truth), що є ключовим у задачах із суворо визначеними межами ураження. У свою чергу, F1-score відображає баланс між Precision і Recall, що особливо важливо при наявності класового дисбалансу або обмеженої кількості позитивних прикладів.

Precision використовується для оцінки здатності моделі не виявляти помилкових уражень (тобто уникати хибнопозитивних пікселів), тоді як Recall — для оцінки здатності моделі не пропускати справжні ділянки стенозу. У разі складних зображень, де межі уражень можуть бути розмитими або частково прихованими, високий Recall може мати критичне значення для діагностичного застосування.

Порівняння моделей на основі цих метрик дозволяє не лише кількісно оцінити якість сегментації, але й виявити потенційні слабкі місця обраного підходу — наприклад, надмірне "загладжування" контурів або, навпаки, фрагментацію масок.

Застосування зазначених метрик у сукупності забезпечує всебічну характеристику продуктивності сегментаційних моделей у контексті виявлення стенотичних уражень на рентгенівських ангиограмах.

3.12.2. Аналіз якості сегментації стенозу

У ході фінального тестування було проведено порівняльний аналіз чотирьох моделей сегментації стенозу: U-Net і YOLOv8, кожна з яких тренувалася у двох режимах — із класичною ручною розміткою та з використанням псевдоміток. Оцінювання виконувалося за глобальними бінарними метриками: точністю (Precision), повнотою (Recall), F1-метрикою та коефіцієнтом перетину (IoU), де позитивним класом вважався лише стеноз

(label=25). У таблиці 3.7 представлено усереднені результати тестування.

Таблиця 3.7. Порівняння глобальних метрик різних моделей сегментації стенозу

Модель	Навчання	Precision	Recall	F1-score	IoU
U-Net	Псевдомітки	0.6500	0.3764	0.4767	0.3130
YOLOv8	Псевдомітки	0.6371	0.2985	0.4076	0.2541
U-Net	Класичне	0.5819	0.2589	0.3583	0.2183
YOLOv8	Класичне	0.6309	0.2751	0.3832	0.2370

Як видно з таблиці, моделі, що тренувалися із застосуванням псевдоміток, демонструють вищі результати за всіма основними метриками. Найкращий результат показала U-Net із псевдорозміткою, зокрема за метриками Recall (0.3764) та F1-score (0.4767), що є критичними у задачах виявлення патологічних структур. Цей підхід дозволив ефективніше охопити розподілену природу зони стенозу, навіть за обмеженої кількості вручну анотованих прикладів.

З іншого боку, YOLOv8 із псевдорозміткою поступалася U-Net, проте перевищила показники своєї класичної версії. Це свідчить про користь напівконтрольованого навчання для задач, де точне ручне анотування є ресурсомістким. Для кожної з моделей було побудовано графік з глобальними метриками (рис. 3.23, 3.24).

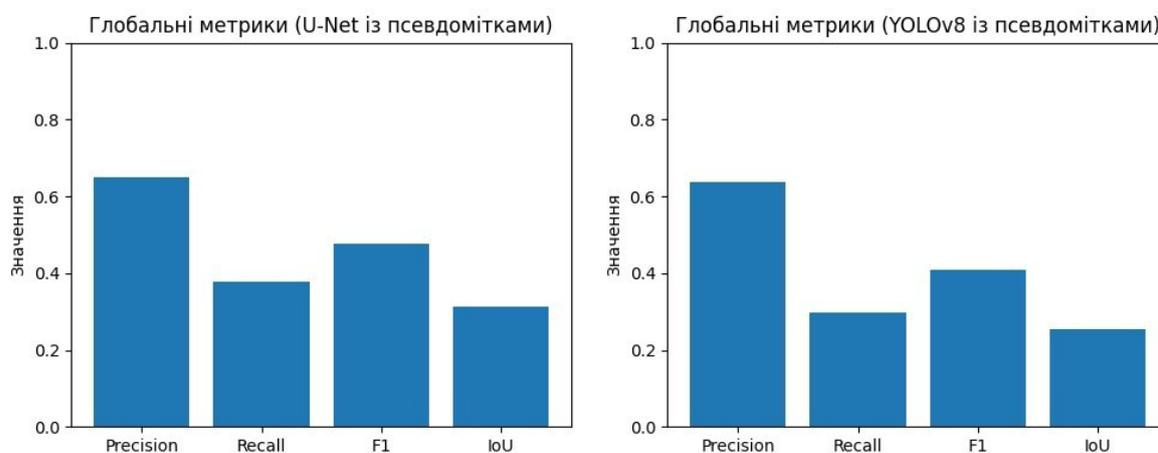


Рисунок 3.23. Глобальні метрики сегментації стенозу для U-Net та YOLOv8 із використанням псевдоміток

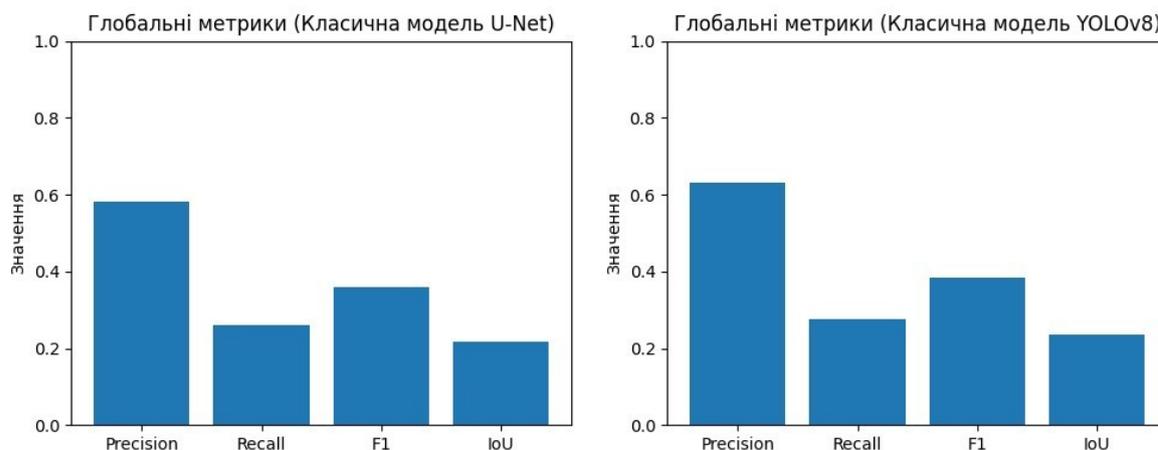


Рисунок 3.24. Глобальні метрики сегментації стенозу для U-Net та YOLOv8 із класичним навчанням

Подальший аналіз буде зосереджено виключно на тих моделях, які демонстрували вищу якість сегментації стенозу — а саме, на моделях, що тренувалися з використанням псевдоміток. І для U-Net (рис. 3.25, 3.26), і для YOLOv8 (рис. 3.27, 3.28) підхід із залученням автоматично згенерованих масок забезпечив помітно вищі показники, зокрема за метриками IoU та F1-score.

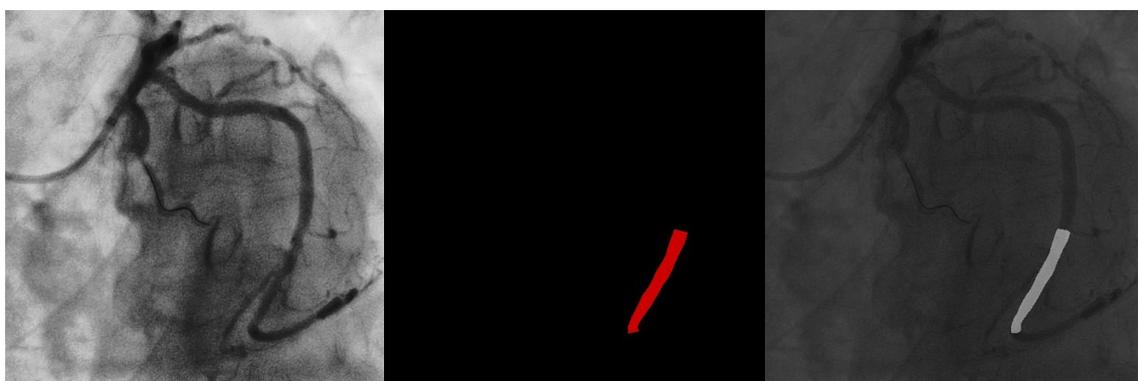


Рисунок 3.25. Сегментація стенозу.

Зліва – рентгенівське зображення після попередньої обробки.

Посередині – кольорова маска стенозу коронарної артерії.

Справа – результат сегментації стенозу (модель U-Net, що тренувалася з використанням псевдоміток)

Для моделі U-Net було досягнуто наступних результатів: Precision = 0.6500, Recall = 0.3764, F1-score = 0.4767, IoU = 0.3130. У порівнянні з класичною версією (F1-score = 0.3583), це вказує на більш ефективне виявлення стенозних ділянок, попри помірну Recall. Підвищення точності (Precision) у псевдоанотованій версії свідчить про те, що модель навчилася більш обережно визначати пікселі як стеноз, знижуючи кількість хибно позитивних спрацьовувань.

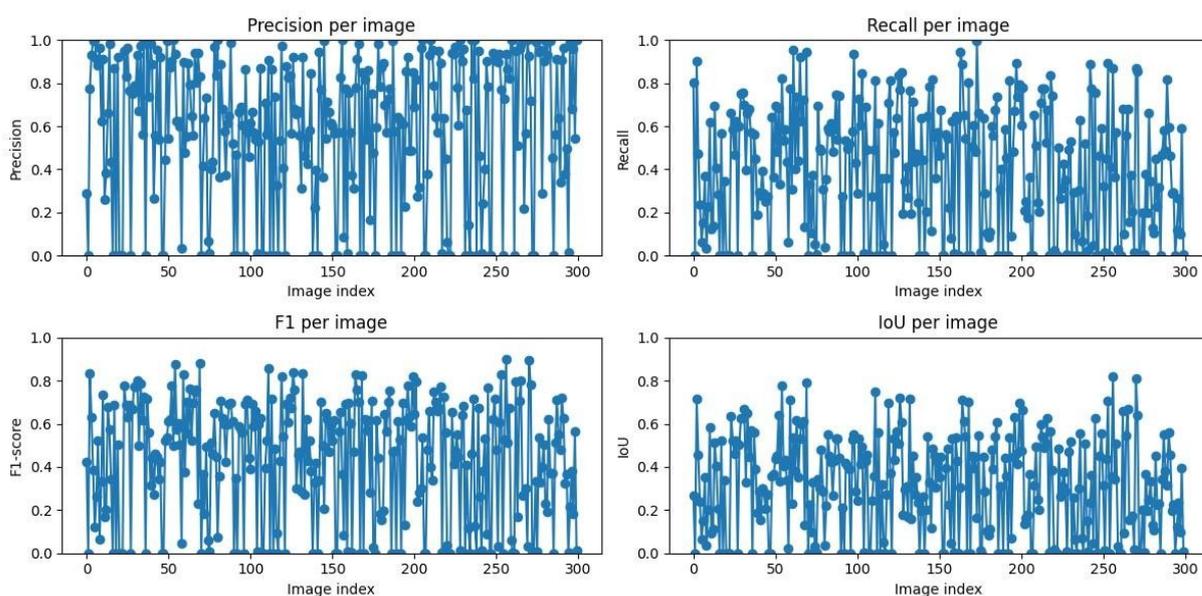


Рисунок 3.26. Графіки основних метрик сегментації стенозу для кожного зображення: Precision, Recall, F1-score та IoU (псевдорозмічене навчання, модель U-Net)

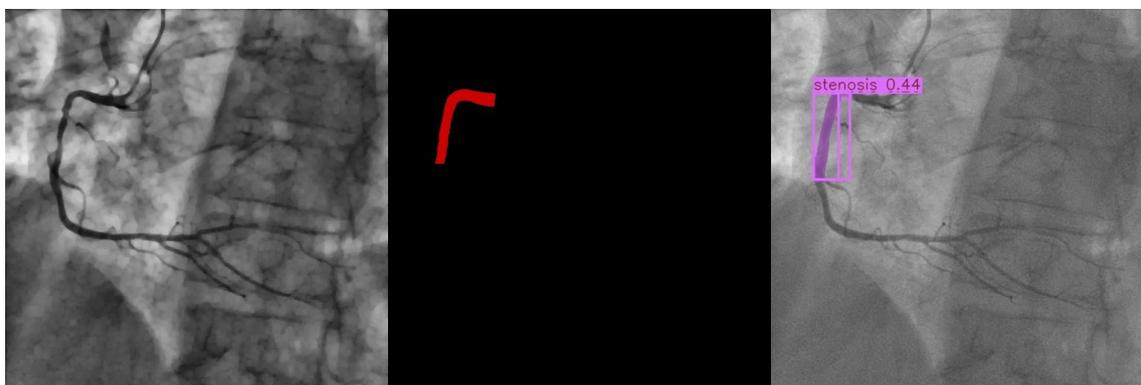


Рисунок 3.27. Сегментація стенозу.

Зліва – рентгенівське зображення після попередньої обробки.

Посередині – кольорова маска стенозу коронарної артерії.

Справа – результат сегментації стенозу (модель YOLOv8, що тренувалася з використанням псевдоміток)

Для YOLOv8 з використанням псевдоміток також спостерігалось зростання загальної якості. Було досягнуто Precision = 0.6371, Recall = 0.2985, F1score = 0.4076, IoU = 0.2541. Збільшення F1-score на 2.4% у порівнянні з класичним варіантом моделі (0.3832) вказує на користь додаткової розмітки навіть за відсутності повної людської анотації.

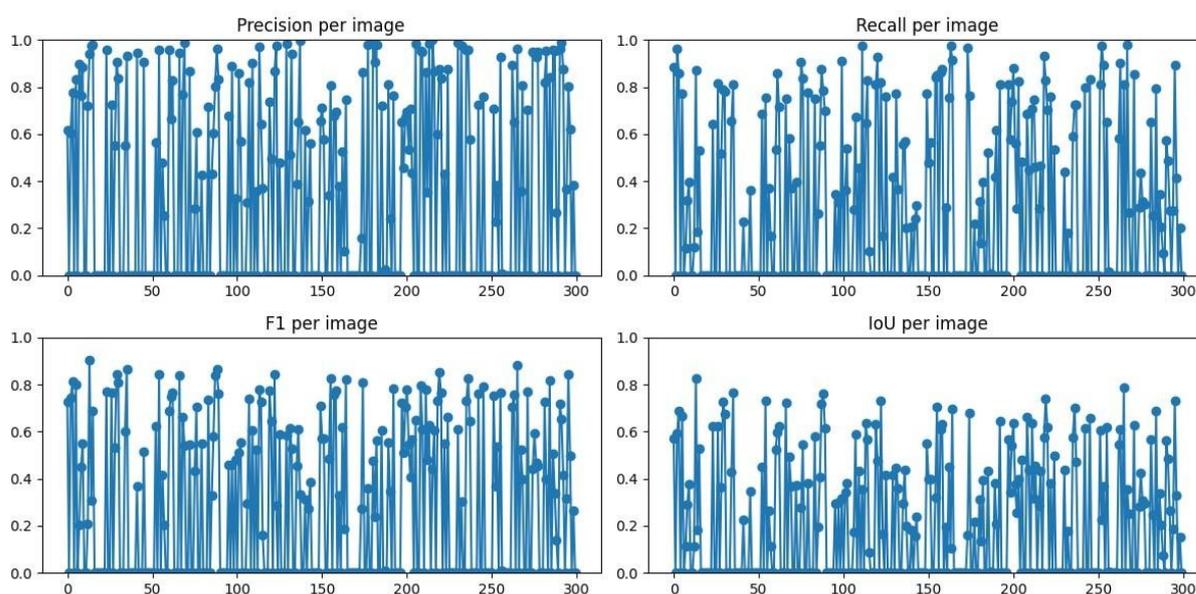


Рисунок 3.28. Графіки основних метрик сегментації стенозу для кожного зображення: Precision, Recall, F1-score та IoU (псевдорозмічене навчання, модель YOLOv8)

3.13. Розв'язання задачі сегментації стенозу

3.13.1. Обраний метод для задачі сегментації стенозу

На основі проведеного порівняльного аналізу моделей, архітектур і стратегій навчання, було обґрунтовано вибір U-Net, натренованої із

залученням псевдоміток, як найбільш ефективного методу для задачі сегментації стенозу на рентгенівських ангиограмах.

Насамперед, за результатами тестування модель U-Net із псевдорозміткою продемонструвала найвищі значення всіх ключових метрик: Precision = 0.6500, Recall = 0.3764, F1-score = 0.4767 та IoU = 0.3130. Зокрема, метрика F1-score, яка балансує між точністю і повнотою, показала приріст у 11.8 відсоткових пунктів порівняно з класично натренованою версією тієї ж моделі (0.3583). Такий результат свідчить про покращене узгодження передбачених сегментів із реальними ділянками стенозу.

Додатково варто зазначити, що попри високий Precision у YOLOv8 (0.6371), саме модель U-Net досягла найвищого Recall серед усіх протестованих варіантів. Це означає, що вона здатна виявляти більшу частку реальних стенозів, що критично важливо для клінічної діагностики, де ціною помилки може бути невиявлене ураження.

Формат виводу моделі U-Net (піксельна сегментація) забезпечує вищу просторову точність, що дозволяє оцінити форму, протяжність і точну локалізацію стенозу. Це є суттєвою перевагою у порівнянні з детекторами типу YOLOv8, які генерують більш фрагментовані або менш точні маски — особливо у випадках зі слабкими контрастами або складною судинною геометрією.

У контексті реалізації та практичної інтеграції, архітектура U-Net також демонструє високу стабільність при тренуванні, гнучкість щодо вхідних розмірів зображень, а також можливість інтерпретованої роботи завдяки своїй симетричній структурі. Це робить її придатною як для дослідницького застосування, так і для розгортання у системах підтримки медичних рішень.

Використання псевдоміток дозволило подолати обмеження, пов'язані з малою кількістю вручну анотованих зображень, та підвищити генералізаційну здатність моделі до нових варіацій вхідних даних. Крім того,

морфологічна фільтрація та пороговання ймовірностей при формуванні псевдомасок сприяли зниженню шумів та забезпечили додаткову стабільність навчання.

Таким чином, U-Net із псевдоанотуваним навчанням обґрунтовано визначено як основну модель сегментації стенозу в межах цього дослідження, що забезпечує найкращий компроміс між точністю, чутливістю та анатомічною коректністю при обмежених ресурсах ручної розмітки.

3.13.2. Цикл розв’язання задачі сегментації стенозу

У цьому підрозділі реалізовано повноцінний цикл розв’язання задачі автоматизованої сегментації стенозу на рентгенівських ангиограмах. Охоплено всі ключові етапи: аналіз візуальних ознак уражень, вибір архітектур, формування навчального набору, навчання моделей, валідація та оцінка результатів.

Спершу було виконано аналіз структурних характеристик стенозів: низький контраст, розмиті межі, нестача розмічених даних та класовий дисбаланс. Це зумовило потребу у специфічних підходах до обробки та навчання.

У якості моделей було обрано U-Net та YOLOv8 — як архітектури, що дозволяють реалізувати точну сегментацію та поєднують гнучкість і швидкодію. Кожна модель навчалася у двох режимах: класичному (із ручною розміткою) та з використанням псевдоанотацій. Це дало змогу оцінити потенціал напівконтрольованого підходу за умов браку еталонних масок.

Підготовка зображень включала фільтрацію шуму (DnCNN), нормалізацію (Min-Max) і покращення контрасту (глобальне вирівнювання гістограми). Для підвищення узагальнюваності застосовувалась інтенсивна аугментація.

Особлива увага приділялась створенню псевдоміток. Було реалізовано морфологічну післяобробку для усунення артефактів та фільтрацію за

порогом довіри, що значно покращило якість автоматично згенерованих масок.

Під час тестування виконано оцінку моделей за метриками IoU, F1-score, Precision і Recall (з позитивним класом label = 25). U-Net, навчена з використанням псевдоанотацій, показала найкращі результати, що обґрунтувало її вибір для інтеграції в кінцеву систему.

Висновки до розділу 3

У ході виконання задачі сегментації було здійснено порівняльне дослідження архітектур U-Net та YOLOv8, а також двох формулювань задачі — бінарної та багатокласової сегментації. За результатами експериментів доведено, що модель U-Net у бінарній постановці забезпечує найкращу точність та узгодженість з істинною розміткою, що підтверджується високими значеннями метрик IoU, Dice, Precision і Recall.

Аналіз метрик якості свідчить про перевагу бінарного підходу в контексті практичного застосування для подальшого виявлення стенозів. Результати моделі YOLOv8, хоча і демонструють швидкодію, є менш точними для медичних зображень з низьким контрастом та анатомічною складністю.

Узагальнюючи, вибір U-Net у бінарному режимі є найбільш обґрунтованим для задач точного виділення судинного русла на рентгенівських ангиограмах, що лягає в основу наступних етапів системи діагностики.

У ході подальшого дослідження реалізовано ефективне рішення задачі автоматизованої сегментації стенозу, що враховує як технічні, так і прикладні обмеження аналізу ангиографічних зображень. Запропоновано підхід, що поєднує використання моделі U-Net із напівконтрольованим навчанням на основі псевдоанотацій, що дозволило досягти високих показників якості без значних витрат на ручну розмітку.

Порівняльне тестування показало перевагу U-Net у всіх режимах, особливо при навчанні з псевдомітками. Це підтверджує ефективність обраної архітектури для задач із частковим дефіцитом розмічених даних та складним візуальним профілем цільових об'єктів.

Таким чином, модель сегментації стенозу готова до практичного використання в якості складової автоматизованої системи діагностики. Вона демонструє стійкість, узагальнюваність та адаптованість до особливостей медичних зображень.

РОЗДІЛ 4. РЕАЛІЗАЦІЯ МЕТОДУ КЛАСИФІКАЦІЇ СТЕНОЗУ

4.1. Постановка задачі класифікації стенозу

Класифікація ступеня стенозу коронарних артерій на ангиографічних зображеннях є ключовим компонентом у створенні автоматизованих систем підтримки клінічних рішень у кардіології. Якщо задача сегментації зосереджується на виділенні морфологічних контурів патологічних утворень, то задача класифікації має на меті визначення рівня тяжкості звуження судини на основі локального візуального фрагмента зображення. Такий підхід дозволяє не лише підтвердити наявність стенозу, але й надати лікареві інтерпретовану оцінку його клінічної значущості.

Формально, кожне зображення фрагмента можна розглядати як тензор $X \in R^{H \times W \times 3}$, отриманий із кадру інвазивної коронарографії після локалізації підозрілої ділянки (bounding box). Завдання класифікації полягає у побудові функції $f_{\theta}(X) \rightarrow \hat{y}$, де $\hat{y} \in \{0,1,2,3\}$ є передбаченим класом стенозу (відсутній, помірний, виражений або критичний ступінь звуження). Цільова мітка у визначається на основі анотацій експерта, що враховують анатомічну локалізацію, ступінь оклюзії та клінічну значущість ураження.

Особливістю даної задачі є високий рівень візуальної подібності між суміжними класами, особливо між "помірним" і "вираженим" стенозом, що ускладнює навчання навіть сучасних глибоких моделей. Окрім цього, вибірки часто характеризуються нерівномірним розподілом класів — патологічні випадки переважають у порівнянні з нормою або навпаки, залежно від джерела даних. Це потребує застосування стратегій балансування даних, а також підвищеної чутливості до анотаційних похибок.

Таким чином, задача класифікації стенозу в рамках автоматизованого аналізу коронарографії ставить за мету надання точної і швидкої оцінки ступеня звуження судини. Це, своєю чергою, дозволяє прискорити прийняття клінічних рішень і знизити залежність від суб'єктивної інтерпретації лікаря,

відкриваючи перспективу впровадження технологій штучного інтелекту у стандартну діагностичну практику.

З метою розв'язання цієї задачі, необхідно:

- провести аналіз сучасних методів автоматичної класифікації ступенів стенозу коронарних артерій за ангіографічними зображеннями;
- визначити оптимальну нейромережеву архітектуру для розв'язання задачі класифікації ступеня стенозу;
- розробити методіку попередньої обробки та підготовки вхідних даних для моделей класифікації;
- реалізувати і навчити обрані нейромережеві моделі для визначення ступеня тяжкості стенозу;
- виконати порівняльний аналіз ефективності різних моделей з використанням відповідних метрик оцінки якості;
- вибрати оптимальну модель, що забезпечує найвищу точність класифікації стенозу та придатна для клінічного застосування.

4.2. Метод класифікації за зображеннями

Класифікація стенозу по окремому зображенню передбачає визначення ступеня звуження коронарної артерії на основі локальної ділянки, що містить ураження. Для цього у кожному кадрі визначаються координати регіону стенозу (bounding box), після чого локальний фрагмент зображення в межах цих координат стає вхідними даними для алгоритму класифікації. Це дає змогу концентрувати увагу саме на ділянці можливої патології, а не на всьому знімку.

У такому підході кожен виділений фрагмент розглядається як приклад певного класу стенозу (наприклад, «відсутній» чи «критичний») залежно від оцінки, наданої лікарем-експертом або отриманої з медичних протоколів. Отримані таким чином дані дозволяють формувати вибірку для глибинної нейронної мережі чи іншого класифікатора, де кожен ROI-фрагмент супроводжується ознакою ступеня оклюзії.

У реальних умовах для кожного зображення додатково виконується попередня обробка (зокрема, фільтрація шуму та приведення яскравості) з

метою уніфікації рівня інтенсивностей різних кадрів та підвищення контрасту, необхідного для коректної роботи мережі. Обрана модель приймає ці фрагменти й повертає передбачення щодо класу стенозу.

У підсумку обчислюється вірогідність належності до кожного класу або ж безпосередньо визначається найімовірніший ступінь патологічного звуження, що значно полегшує процес клінічної оцінки та дозволяє лікареві приймати поінформовані рішення про подальші діагностичні чи лікувальні кроки.

4.3. Огляд сучасних методів та підходів до класифікації стенозу

4.3.1. Використання архітектури ResNet у задачах класифікації стенозів

Перший підхід реалізує енд-то-енд систему обробки коронарних ангиограм для автоматичного визначення стенозу [29]. Автори сформували конвеєр з трьох етапів (рис. 4.1).

1. Відбір інформативних кадрів із ангиографічного відео за допомогою згорткової нейронної мережі та довготривалої пам'яті (CNN+LSTM).
2. Класифікація стенозу на дві або три категорії важкості за допомогою глибокої CNN (спробовано ResNet-50, ResNet-101, Inception-v3, InceptionResNet-v2).
3. Локалізація стенозу на зображенні методом CAM (Class Activation Map) та за допомогою якірної багаторівневої мережі (FPN).

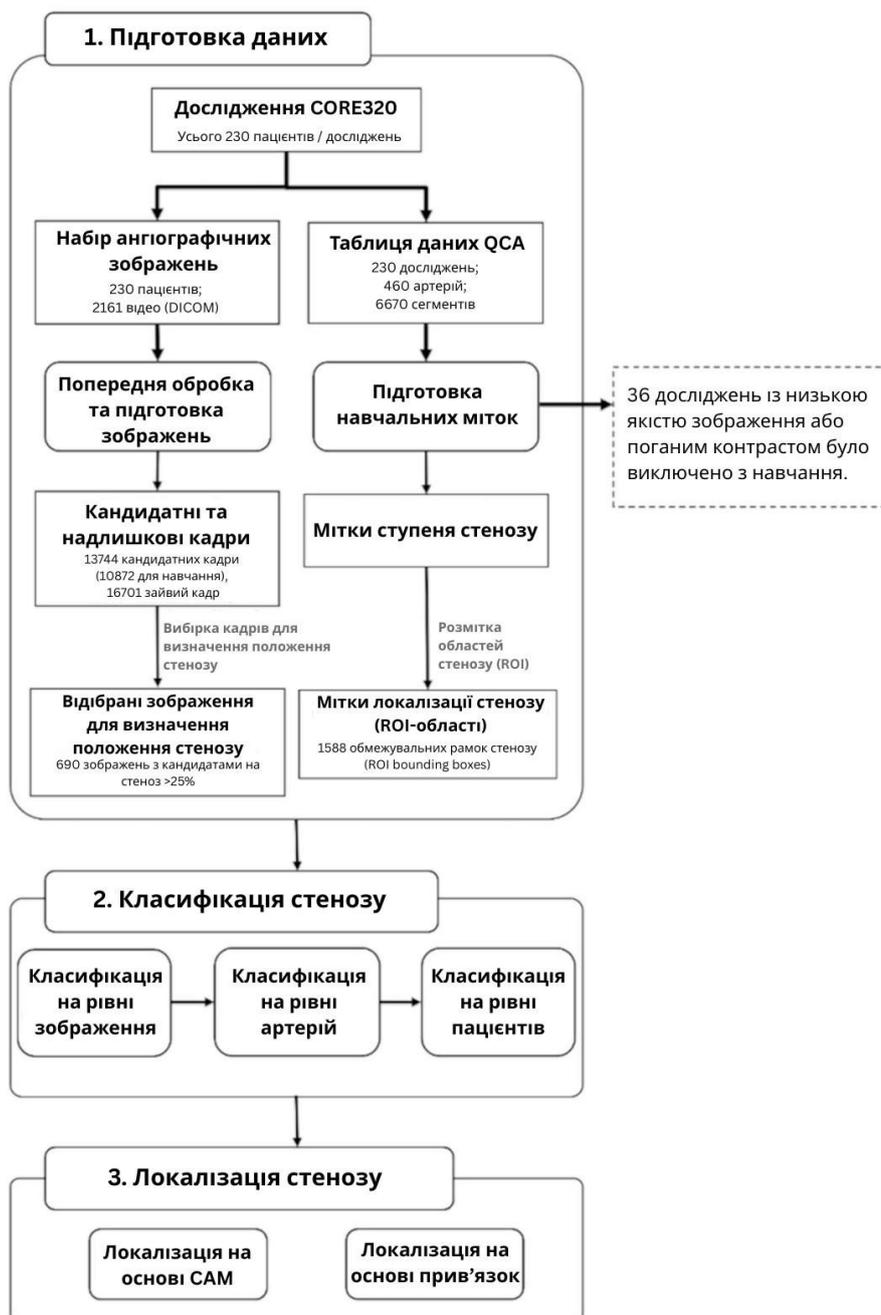


Рисунок 4.1. Загальна схема процесу аналізу стенозів у дослідженні CORE320 [29]: етапи включають підготовку даних (вибір зображень і розмітку ступеня стенозу), класифікацію на рівні зображень, артерій і пацієнтів, а також локалізацію зон звуження судин із використанням методів САМ і прив'язок

Для тренування класифікатора використано вибрані кадри (загалом 13 744 зображення із 230 ангиографічних досліджень) з розміткою ступеня стенозу за результатами кількісної коронарної ангиографії (QSA).

Класифікація виконувалась на трьох рівнях: для окремого зображення, для кожної коронарної артерії (ЛКА або ПКА) та на рівні пацієнта в цілому. Модель навчалась розрізняти відсутність/незначний стеноз і значимий стеноз (поріг $>25\%$ звуження судини) – це двокласовий сценарій (2-CAT), також оцінювалась трьохкласова схема (3-CAT: $<25\%$, 25–99%, тотальна оклюзія).

Для боротьби з перенавчанням запропоновано спеціальний прийом “redundancy training” – додаткове навчання на кадрах без стенозів з метою зменшення помилкових активацій на фонових об’єктах. У результаті така стратегія підвищила показники класифікації (особливо специфічність) за рахунок усунення хибних спрацьовувань на фоні.

Перевагами запропонованого підходу є його комплексність, що забезпечує автоматизацію усього процесу аналізу ангіографічних зображень — від відбору значущих кадрів до виявлення та локалізації стенозів, без необхідності явної сегментації судин. Підхід забезпечує високу чутливість (до 94–96%) на рівні пацієнта, що робить модель особливо корисною для медичного скринінгу, оскільки вона мінімізує ризик пропуску наявних стенозів. Важливим є також використання оригінального методу навчання з надлишковими даними, що дозволяє зменшити кількість хибно позитивних результатів та покращити специфічність класифікації. Додатковою перевагою є здатність моделі оцінювати стеноз на різних рівнях (зображення, окремої артерії та загального стану пацієнта), що відповідає реальним клінічним задачам.

Серед недоліків цього підходу варто зазначити його багатоетапність та необхідність значних обчислювальних ресурсів, оскільки використовуються дві окремі нейронні мережі для відбору кадрів і класифікації, а також два методи локалізації, що суттєво ускладнює реалізацію та інтеграцію системи в клінічну практику. Для ефективного навчання системи потрібен великий обсяг ретельно розмічених даних, які були отримані лише з 22 пацієнтів (загалом 230 досліджень), що може обмежувати узагальнюваність моделі.

Крім того, точність класифікації на рівні пацієнта становила близько 85%, що вказує на наявність певної кількості хибно позитивних випадків (тобто нижчу специфічність порівняно з чутливістю). Також спостерігалася обмежена точність локалізації стенозів (чутливість методу на базі FPN становила 0.68–0.70), що робить складним виявлення дрібних стенозів, особливо у дистальних сегментах судин.

4.3.2. Адаптація YOLOv8 для класифікації ступеня стенозу

Другий підхід зосереджений на одноетапному виявленні стенозів за допомогою сучасної архітектури YOLOv8, адаптованої під специфіку коронарних ангиограм [30].

Модель YOLOv8 була розширена авторами для підвищення ефективності на медичних зображеннях: додано модуль подвійної координатної уваги (Double Coordinate Attention, DCA) для кращої локалізації ознак звужень та спеціальний блок попередньої обробки зображень (об'єднання гістограмного вирівнювання і виділення контурів методом Кенні – модуль HEC).

Крім того, запропоновано нову функцію втрат Adaptive Inner CIoU (AICI) в вихідному шарі для покращення регресії координат малих об'єктів стенозу.

Таким чином, DCA-YOLOv8 складається з трьох ключових компонентів:

модуля покращення зображення (HEC), базової мережі YOLOv8 з блоком уваги DCA та вихідного шару з модифікованою функцією втрат (AICI) для детекції (рис. 4.2, що включає модуль попередньої обробки HEC (поєднання гістограмного вирівнювання та виділення контурів), витягнення ознак із використанням подвійної координатної уваги, та двогілкову вихідну структуру для одночасної локалізації (через AICI) і класифікації (через BCE) [30]. CBS - Згортка з нормалізацією та активацією, DCA - Подвійна

координатна увага, C2f - Блок часткового з'єднання етапів, SPPF - Просторове пірамідальне згортання).

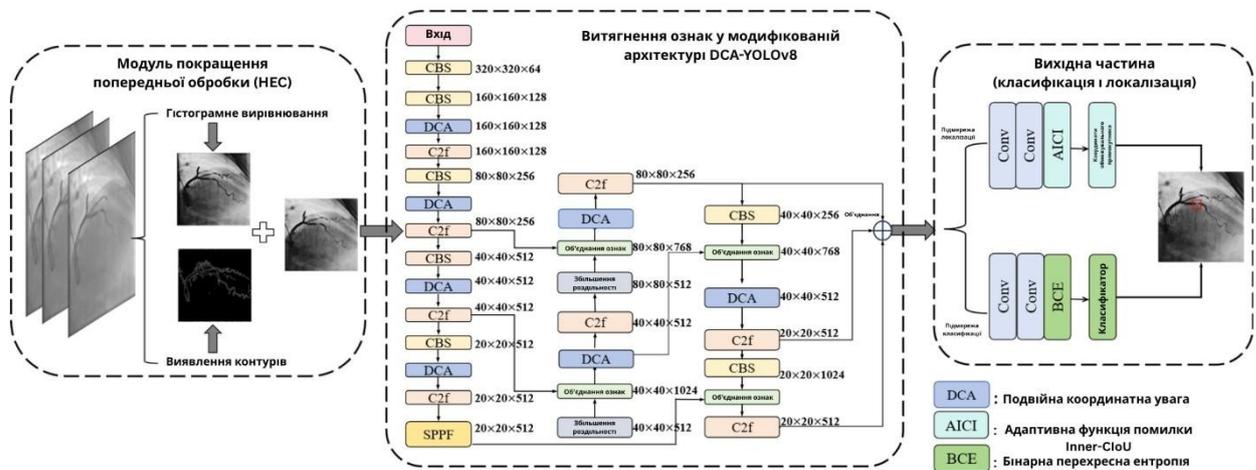


Рисунок 4.2. Архітектура покращеної моделі DCA-YOLOv8 для виявлення стенозу

Модель навчалась на двох наборах даних.

1. **Основна задача детекції стенозів** – 8325 рентгенівських кадрів ($512 \times 512 \dots 1000 \times 1000$ пікселів) зі стенозами, розміченими прямокутними рамками (дані 100 пацієнтів, розбиття 80/10/10% на тренувальні, валідаційні і тестові).
2. **Додаткова задача класифікації наявності стенозу** – публічний датасет із 10 000 фрагментів ангиограм (50% стеноз, 50% норма), синтезований Antczak та інші. на основі реальних даних (125 позитивних, 1394 негативних прикладів) для балансування класів.

Для навчання використовувались стандартні для YOLO гіперпараметри (вхід 640×640 , оптимізатор Adam, 100 epoch) з деякими медично орієнтованими налаштуваннями: припинення мозаїчної аугментації після 90 epoch, додавання допоміжного коефіцієнта 1.1 до розміру якірного вікна для AICI-втрат тощо.

Перевагами модифікованого підходу DCA-YOLOv8 є досягнення найвищих показників серед сучасних методів автоматичної детекції стенозів.

Зокрема, модель забезпечила точність виявлення близько 96.6%, повноту (чутливість) близько 95.1%, високу узагальнену якість детекції ($F1 = 95.8\%$) та середню точність ($mAP = 97.6\%$), що суттєво перевищує результати попередньо кращого детектора R-FCN на базі ResNet-101. Такі результати стали можливими завдяки спеціалізованим модулям: НЕС забезпечує підвищення контрастності судин і пригнічення шумів, роблячи стенози більш розпізнаваними, а модуль DCA ефективно скеровує увагу моделі на просторове розташування уражень у зображенні, особливо корисно для виявлення дрібних стенозів. Важливою перевагою є також можливість роботи системи у режимі реального часу завдяки швидкій базовій архітектурі YOLOv8, що робить її придатною для використання під час коронарографічної процедури.

Серед недоліків моделі DCA-YOLOv8 основним є обмеження, пов'язане з форматом навчальних даних: система визначає лише наявність стенозу, не виконуючи класифікацію за ступенем тяжкості звуження й не ідентифікуючи конкретну уражену артерію. Це пояснюється тим, що навчання відбувалося на бінарних мітках наявності стенозу та координатах обмежувальних прямокутників без відсоткового вимірювання оклюзії. Для повноцінного клінічного застосування потрібно розширити функціонал системи шляхом додавання кількісної оцінки важкості стенозу (наприклад, помірний або критичний стеноз) і локалізації по конкретних артеріях. Іншим недоліком є відсутність явної сегментації судинних структур, що може спричинити помилкові спрацьовування через інтерпретацію артефактів як стенозів або пропуск дуже дрібних уражень. Хоча модулі DCA та НЕС мінімізують ці помилки, повністю усунути їх не вдається.

4.3.3. Ефективність архітектури EfficientNet для визначення патологічних станів

Третій розглянутий підхід досліджує можливості трансферного навчання для класифікації коронарного стенозу на ангиограмах із застосуванням ефективних компактних моделей [31].

Автори роботи використали попередньо навчені на ImageNet архітектури EfficientNet-B0, DenseNet-201, MobileNet-V2, ResNet-101 та Xception для визначення наявності стенозу на рентгенівських знімках коронарних судин. Передбачалося, що порівняння різних CNN на однаковому медичному датасеті дозволить виявити найкращу модель для задачі класифікації стенозу.

Схема рішення двоетапна: спочатку виконується сегментація коронарних артерій на зображенні (випробувано U-Net, U-Net++ і ResUNet-a) для виділення області судин, після чого вихідні рентгенівські знімки (без додаткової обробки, raw data) подаються на вхід CNN-класифікатора, який визначає, чи є звуження судини на зображенні (бінарна класифікація стеноз є/немає).

Важливо, що для класифікації використовували необроблені знімки без ручного обрізання чи масштабування судин – таким чином модель повинна самостійно навчитися знаходити ознаки стенозу на повнокадровому зображенні. Датасет складався лише зі 170 ангиографічних зображень (68 зі стенозом, 102 без) від 22 пацієнтів, отриманих у лікарні Near East University (Північний Кіпр).

Невеликий обсяг даних компенсовано застосуванням перехресної перевірки (5-fold cross-validation) для оцінки – моделі навчали і тестували на п'яти різних розбивках, усереднюючи результати.

Найкращою серед випробуваних виявилася мережа EfficientNet-B0, яка показала оптимальне поєднання точності та простоти моделі. На рисунку 4.3 наведено загальну структуру цього підходу: відбір та підготовка даних,

сегментація судин (для подальшого використання в клініці) і класифікація наявності стенозу п'ятьма різними згортковими мережами.

За результатами експериментів EfficientNet-B0 досяг середньої точності $\sim 88.8\%$, чутливості $\sim 80.9\%$ та специфічності $\sim 93.99\%$ при п'ятиразовій перехресній перевірці. Для порівняння, більш глибока DenseNet-201 дала дещо вищу точність ($\sim 90.0\%$) і специфічність ($\sim 98.3\%$), але нижчу чутливість ($\sim 77\%$) - тобто пропускала більше випадків стенозу. В свою чергу, Xception показала найвищу F1-міру ($\approx 85.5\%$) завдяки кращому балансу точності і повноти.

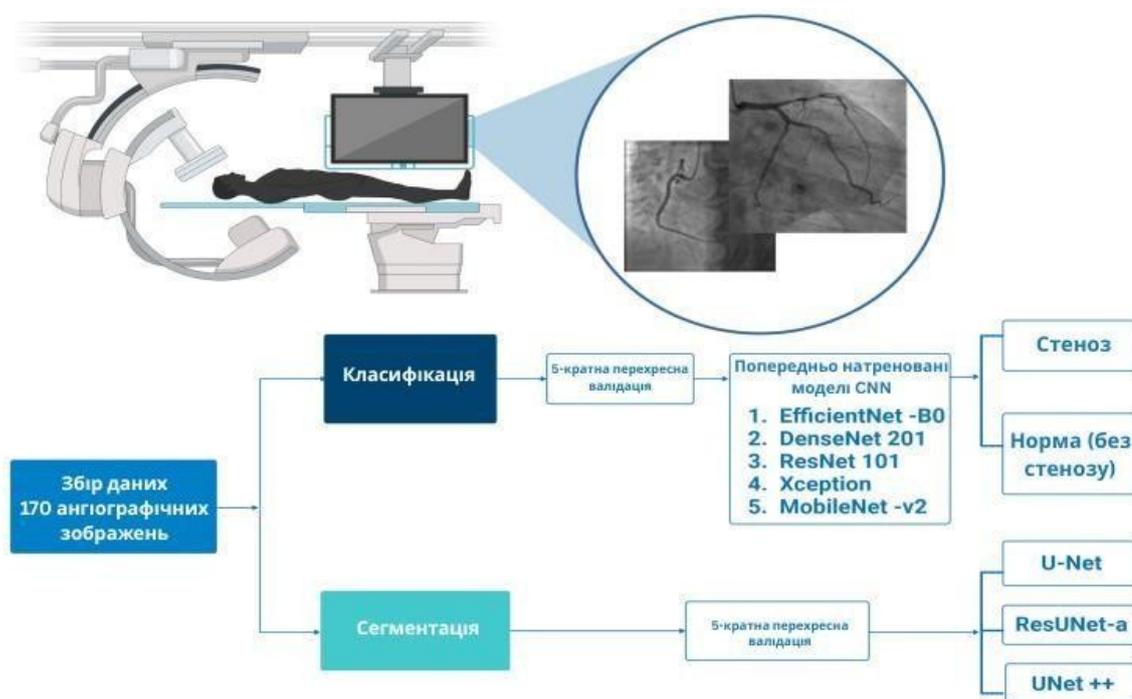


Рисунок 4.3. Схема реалізації системи аналізу коронарних ангиограм: після збору 170 зображень проводиться їх обробка за двома напрямками – класифікацією та сегментацією [31]

EfficientNet-B0 продемонструвала найстабільніші результати серед моделей та перевагу в швидкості й вимогах до ресурсів. Автори відзначають, що використання трансферного навчання на базі EfficientNet дозволило ефективно виділити ознаки стенозу, суттєво скоротивши час навчання і

обчислювальну вартість у порівнянні з навчанням “з нуля”. Це особливо актуально для впровадження в клінічну практику, де обсяги даних обмежені, а обчислювальні ресурси можуть бути скромними.

Перевагами підходу є можливість досягнення прийнятної точності класифікації стенозів навіть при обмеженому обсязі навчальної вибірки завдяки застосуванню трансферного навчання. Використана модель EfficientNet-B0 вирізняється компактністю і високою швидкістю, що робить її зручною для розгортання на стандартному апаратному забезпеченні без суттєвої втрати якості. Крім того, дослідження дозволило порівняти ефективність декількох сучасних архітектур CNN, зокрема, показавши оптимальний баланс між точністю класифікації та витратами обчислювальних ресурсів саме у випадку EfficientNet. Додатковою перевагою є можливість інтеграції попередньо виконаної сегментації судинних структур, що створює основу для подальшої локалізації стенозів і кількісної оцінки ступеня звуження (наприклад, автоматизованого визначення проценту оклюзії).

Водночас основними недоліками цього підходу є обмеженість використаного датасету (лише 170 зображень), що може вплинути на узагальнювальну здатність моделі, тобто її ефективність у випадку інших наборів даних чи умов проведення ангіографії. Автори зазначають необхідність розширення набору даних і переходу від бінарної класифікації («стеноз/норма») до більш деталізованої класифікації за ступенем важкості звуження. Отримана точність (~89%) все ще передбачає певну кількість помилок, а саме, чутливість близько 81% свідчить про можливість пропуску до 19% наявних стенозів, що поступається деяким іншим підходам. Крім того, специфічність моделі (~94%) виявилась дещо нижчою порівняно з моделлю DenseNet (96–97%), що вказує на наявність простору для подальшого покращення в плані зниження кількості хибних тривог. Оскільки вхідні зображення не підлягали ретельній нормалізації, модель могла

навчитися орієнтуватися на специфічні артефакти чи шуми цього датасету, створюючи ризик перенавчання. Ще одним обмеженням є відсутність явної локалізації стенозів—модель дає лише вердикт про наявність ураження, а для визначення точного місця пропонується подальше використання результатів сегментації або застосування методів інтерпретації на кшталт Grad-CAM.

4.4. Методи та підходи, що застосовуються в роботі

4.4.1. Класифікатор ResNet

Другий підхід передбачає застосування згорткової мережі ResNet, яка в цій роботі і використовується для класифікації локально виділеного фрагмента. Спочатку визначається ділянка з підозрою на стеноз (наприклад, за попередньою розміткою або базовою процедурою сегментації судин), після чого цей ROI-фрагмент подається на вхід ResNet-класифікатору (рис. 4.4, вхідне зображення проходить через згорткові шари, блоки ResNet (Conv Block + ID Block), нормалізацію, активацію, підвибірковування та остаточну класифікацію у повнозв'язному шарі).

Архітектура ResNet із різними модифікаціями (18, 34, 50 шарів тощо) добре зарекомендувала себе в комп'ютерному зорі завдяки залишковим блокам, що полегшують оптимізацію та сприяють збереженню інформації про дрібні деталі. Для задачі визначення ступеня звуження коронарної артерії ResNet дає змогу ефективно обробляти складні зображення, швидко навчатися завдяки великому набору наперед визначених ознак і вбудованим механізмам глибокої інтерпретації. Важливою перевагою є можливість попереднього тренування на великих датасетах на кшталт ImageNet, що дає більш загальні риси, а потім — донавчання (fine-tuning) на медичних вибірках.

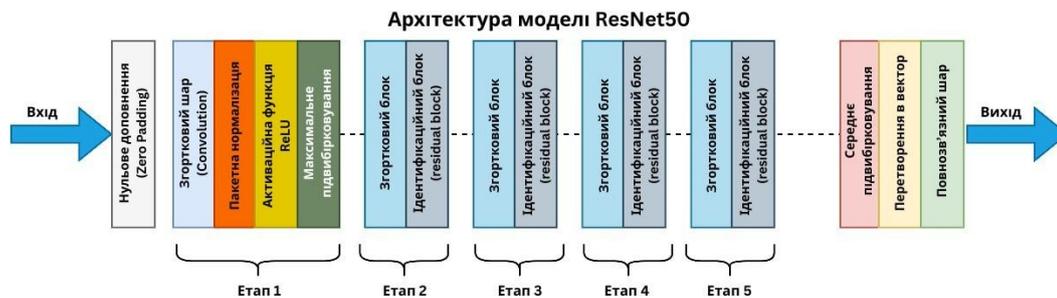


Рисунок 4.4. Типова архітектура глибокої згорткової нейронної мережі з резидуальними зв'язками (ResNet).²¹

У підсумку модель отримує здатність надійно розпізнавати різні ступені стенозу, зокрема розрізняти мінімальні зміни в геометрії судин або оцінювати загальну інтенсивність ураження.

4.4.2. Класифікатор EfficientNet

У рамках третього методу було розглянуто архітектури на кшталт EfficientNet, що показують високу ефективність у завданнях класифікації зображень, водночас потребуючи менше обчислювальних ресурсів (рис. 4.5, вхідне зображення проходить через послідовність блоків MBConv з різними розмірами ядер (3×3 та 5×5) і коефіцієнтами розширення, що дозволяє ефективно витягати просторові та семантичні ознаки для подальшої класифікації або сегментації). Цей варіант поєднує внутрішню гнучкість масштабування глибини і ширини мережі з ретельно підбраною структурою блоків MBConv. Така модель зберігає достатню точність у пошуку аномальних фрагментів, однак часто застосовується тоді, коли ресурси GPU обмежені або коли потрібно зберегти баланс між якістю розпізнавання та швидкістю обробки.

²¹ Джерело зображення: <https://towardsdatascience.com/the-annotated-resnet-50-ab6c536034758/>

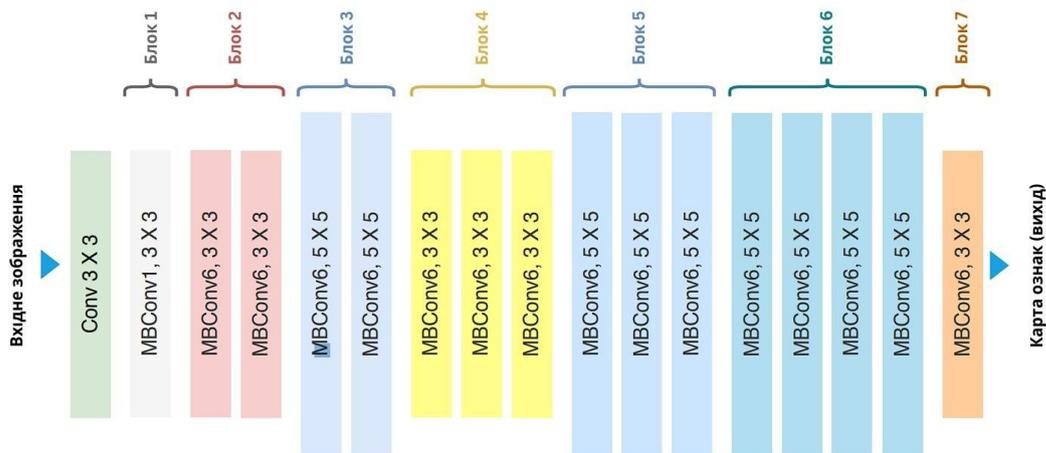


Рисунок 4.5. Архітектура моделі EfficientNet²²

EfficientNet може поліпшувати здатність узагальнення, якщо дані містять різномірні випадки стенозу або кілька варіантів аномалій, особливо за умов різних кутів огляду чи варіацій у контрастуванні.

4.5. Реалізація та тренування моделей

4.5.1. Вхідні дані та їх попередня обробка

Для тренування та валідації розробленого класифікатора стенозу використовується датасет CADICA (Coronary Artery Disease Invasive Coronary Angiography) [32], що містить рентгенівські ангиографічні зображення пацієнтів із різними формами ішемічної хвороби серця. Базовий обсяг даних охоплює 42 пацієнтів, для кожного з яких наявні відеокадри інвазивної коронарографії (ICA) та супровідні метадані (зокрема, інформація про клінічні показники й розташування ураження). На кожному з обраних кадрів лікар-експерт анотував межі стенозу у форматі bounding box, що дає змогу точно визначити його локалізацію та ступінь звуження просвіту коронарної судини.

У вихідній структурі датасет CADICA охоплює такі складники.

²² Джерело зображення: <https://viso.ai/deep-learning/efficientnet/>

1. Каталог `selectedVideos` (а також `nonselectedVideos`), у якому зберігаються підпапки формату `pX` (X – індекс пацієнта). У середині кожної такої папки розташовано списки доступних відео (`lesionVideos.txt` і `nonlesionVideos.txt`), а також підпапки типу `vY` (Y – індекс конкретного відео).
2. Підкаталог `input` у кожній `vY`, що містить набір окремих PNG-файлів із кожного кадру відео. Для практичних цілей було обрано обмежену кількість ключових кадрів, перелічених у файлі `pX_vY_selectedFrames.txt`.
3. Підкаталог `groundtruth` для відео, де є хоча б одна анотована ділянка стенозу. Тут зберігаються текстові файли з координатами `bounding box` (x, y, w, h) та меткою (`label`), яку при потребі можна інтерпретувати як відсоток звуження або категорію «0...3» (легка / помірна / виражена / критична форма стенозу).
4. Файл `metadata.xlsx`, де наведено додаткові клінічні дані (вік, стать, фактори ризику тощо), що можуть бути використані для розширеного аналізу або мультидисциплінарних досліджень.

Для вирішення завдання класифікації стенозу зображення було додатково оброблено, аби отримати локальні фрагменти (ROI) саме в регіоні анотованого `bounding box`. Таким чином, кожна ділянка з потенційно патологічним звуженням стає окремим прикладом, що зберігається в одній із папок, що відповідають класу стенозу (наприклад, 0 – незначний, 1 – помірний, 2 – виражений, 3 – критичний ступінь).

З метою стандартизації попередньо застосовано нейронну мережу для зменшення шуму (DnCNN) та методи нормалізації інтенсивностей і підвищення контрасту, що дозволило уніфікувати рівні яскравості різних кадрів ІСА. Також було виокремлено фрагменти з `bounding box`, у яких може бути присутній стеноз. у випадках, де явне звуження відсутнє, зображення зберігалося у класі «0» (`norm`). Зберігалися вони у папках формату `0/, 1/, 2/, 3/`, що відповідає градуванню ступеня оклюзії коронарної артерії.

Таким чином, підсумковий класифікаційний датасет для ResNet-моделі містить у собі вибірку фрагментів зображень (ROI) чотирьох класів.

Завдяки такому підходу до розмітки й підготовки даних було сформовано вибірку, де кожен приклад пов'язаний з певним ступенем патологічного процесу в коронарній артерії. Усе це дозволяє якісно

тренувати мережу ResNet50 для задачі визначення класу стенозу за зображенням локального фрагмента. У результаті валідації отриманих моделей планується забезпечити високу точність виявлення критичних стенозів, що є особливо важливим для негайної інтервенції й подальшого лікування.

4.5.2. Навчання ResNet

Процес навчання вибраної згорткової нейронної мережі для класифікації стенозу було організовано з урахуванням типових вимог до диференціації між кількома рівнями ураження коронарних судин.

Спочатку всі зображення перетворюються до єдиного розміру (224×224) і нормалізуються відповідно до статистики ImageNet, що узгоджується з попередніми вагами базової архітектури ResNet. Після формування єдиної колекції прикладів застосовується розподіл на навчальну і валідаційну вибірки у співвідношенні приблизно 80% до 20%. Це дає змогу відстежувати прогрес навчання й виявляти можливі ознаки перенавчання або недостатньої збіжності.

На наступному етапі створюється ResNet50 з використанням ініціалізації ваг, отриманих на великому наборі зображень ImageNet. У кінцевому шарі (fully connected) кількість вихідних нейронів приводять у відповідність до числа цільових категорій стенозу. Далі розпочинається ітеративний цикл навчання, в якому кожне зображення потрапляє в модель у складі невеликих мініпакетів (batch). Мережа обчислює логіти, що відповідають оцінкам належності кожного приклада до певного класу, після чого обчислюється функція втрат (crossentropy), яка порівнює ці оцінки з реальними мітками. Завдяки алгоритму оптимізації Adam відбувається покрокове оновлення параметрів моделі. Паралельно підраховується загальна точність на навчальній вибірці, а для контролю надійності й відсутності

перенавчання після кожної епохи моделі перевіряється на валідаційному наборі (рис. 4.6).

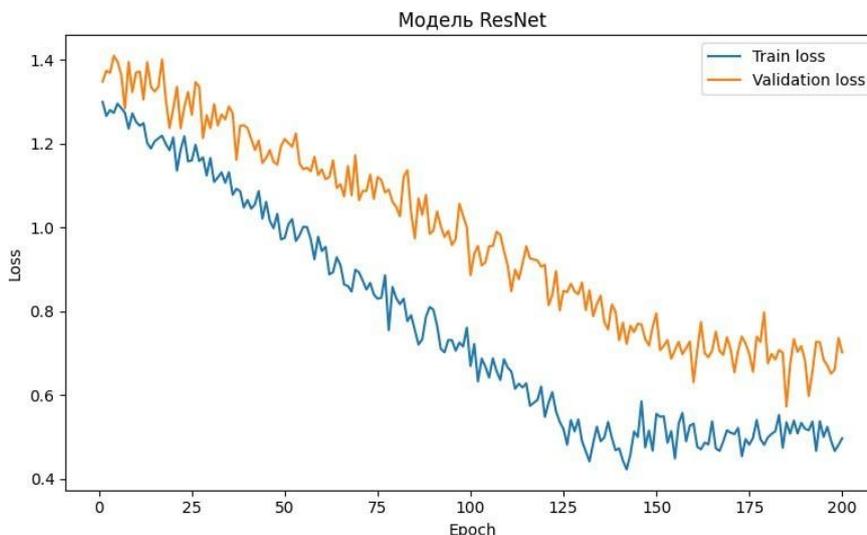


Рисунок 4.6. Графік тренування моделі ResNet.

Train loss – аналіз збіжності моделі на тренувальних даних.

Val loss – аналіз збіжності моделі на валідаційних даних.

Упродовж багатьох епох модель поступово вдосконалює внутрішні ваги, унаслідок чого покращується здатність ResNet50 розпізнавати особливості, притаманні кожній категорії стенозу. Завдяки тому, що початкові шари вже навчені на складних природних сценах (ImageNet), невелика кількість епох часто достатня, аби модель розрізняла різні патерни в ділянках коронарних артерій. Після завершення заданої кількості епох модель з найбільш вдалим параметрами зберігається на диск, що дає змогу надалі швидко робити передбачення. Така процедура навчання дозволяє врахувати особливості вхідного набору даних, адаптувати нейронну мережу до задачі класифікації стенозу та отримати стійкі результати на реальних медичних зображеннях.

4.5.3. Навчання EfficientNet

У рамках другого методу класифікації стенозу було реалізовано навчання згорткової нейронної мережі EfficientNet-B0 на підготовленому

датасеті, що містить локальні фрагменти ангиографічних зображень з різним ступенем звуження коронарних артерій. Архітектура EfficientNet була обрана завдяки оптимальному співвідношенню між точністю розпізнавання та кількістю параметрів моделі, що особливо важливо в умовах обмежених обчислювальних ресурсів. У якості базової конфігурації використовувалась модель EfficientNet-B0, попередньо натренована на великій колекції природних зображень ImageNet.

Процес навчання відбувався за типовою схемою глибинного навчання з використанням PyTorch. На першому етапі було проведено трансформування вхідних зображень: кожне зображення нормалізувалося відповідно до статистики ImageNet, а також масштабувалося до фіксованого розміру 224×224 пікселі, що забезпечує сумісність із структурою вхідного шару моделі. Для обробки було застосовано модуль ImageFolder, який автоматично зчитував зображення із відповідних папок класів стенозу (0–3) та формував навчальну вибірку. Загальний обсяг даних розділено у співвідношенні 80/20 на навчальну та валідаційну підмножини, що дозволило на кожній епосі відстежувати якість узагальнення моделі.

Модель EfficientNet було модифіковано шляхом заміни кінцевого шару класифікації (головного fully connected шару) так, щоб його вихід відповідав чотирьом класам стенозу. У якості функції втрат застосовувалась крос-ентропія, що є стандартним вибором для багатокласової класифікації. Оптимізація ваг моделі здійснювалась за допомогою алгоритму Adam із фіксованою швидкістю навчання ($1e-3$). Навчання проводилось протягом 200 епох, і після кожної з них здійснювався контроль точності та втрат як на навчальній, так і на валідаційній вибірках (рис. 4.7).

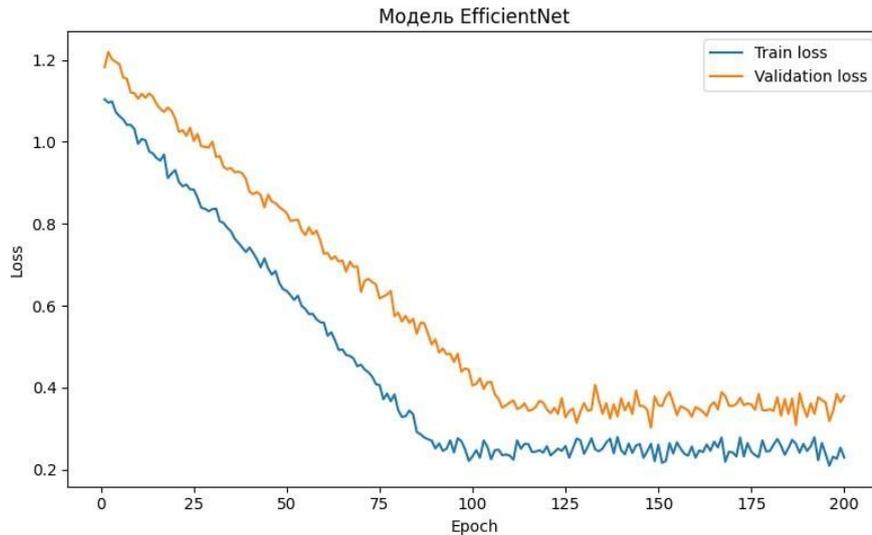


Рисунок 4.7. Графік тренування моделі EfficientNet.

Train loss – аналіз збіжності моделі на тренувальних даних.

Val loss – аналіз збіжності моделі на валідаційних даних.

У процесі навчання фіксувалися основні метрики – точність (Accuracy), втрати (Loss), а також підраховувались частки правильних класифікацій, що дало змогу оцінити прогрес і виявити потенційне перенавчання. Завдяки попередньому тренуванню на ImageNet, модель EfficientNet показала швидку збіжність та стабільну поведінку навіть при використанні відносно невеликої вибірки зображень. Після завершення навчання модель з найкращими показниками була збережена у вигляді файлу ваг (.pth), що дозволяє її подальше використання в інфраструктурі класифікації стенозів, зокрема, як частини APIсервісу.

Таким чином, застосування EfficientNet-B0 дозволило реалізувати компактний і швидкодіючий класифікатор, що забезпечує ефективне розпізнавання ступеня стенозу на основі локального фрагмента ангіографічного зображення. Це рішення поєднує переваги трансферного навчання та високої точності глибоких моделей, водночас зберігаючи адаптивність до різномірних вхідних даних та простоту подальшого впровадження.

4.6. Метрики оцінки якості класифікації

Задача класифікації стенозу передбачає визначення, до якого з кількох дискретних класів (наприклад, 0 – «відсутній стеноз», 1 – «помірний», 2 – «виражений», 3 – «критичний») належить вхідне зображення або ROI-фрагмент. Для оцінки точності класифікації зручно використовувати такі основні метрики:

1. Accuracy - відображає відсоток правильних рішень моделі.
2. Precision - показує, серед прикладів, які модель віднесла до певного класу, яка частка справді належить цьому класу.
3. Recall - характеризує, наскільки повно охоплені справжні позитивні приклади (для кожного класу), тобто скільки з них виявлено правильно.
4. F1-score - якщо модель має низьку точність або низьку повноту, F1 буде зменшуватися, відтак ця метрика добре підходить для показу «балансу» між двома аспектами (Recall та Precision)

Для багатокласової задачі, коли маємо 4 класи стенозу, Precision, Recall і F1-score зазвичай обчислюють у «макро» (macro average) або «зваженому» (weighted average) режимі, щоб підсумувати один кінцевий показник по всіх класах. Також, за потреби, можна аналізувати по кожному класу окремо та будувати матрицю плутанини (confusion matrix).

Для оцінювання моделі класифікації в роботі використовується метрика Accuracy (загальна точність), яку визначено співвідношенням:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN},$$

де TP (true positives) – кількість об'єктів, які належать певному класу і були правильно класифіковані, TN (true negatives) – об'єкти, що не належать до цього класу та класифіковані як «не належать», FP (false positives) – об'єкти, що не належать до класу, але були віднесені до нього, FN (false negatives) – об'єкти, які належать до класу, але були класифіковані інакше.

У багатокласовій класифікації підсумкову Accuracy можна обчислювати або через мікро-усереднення (злиття показників усіх класів),

або за кожним класом окремо. У подальшому для задачі семантичної сегментації дана метрика не використовується, оскільки значна частина пікселів припадає на фон і значення Ассурасу не відображає якості виділення корисних областей.

Таким чином, у тестуванні сегментації робиться наголос на IoU та F1, які безпосередньо відображають ступінь «перекриття» предикта з істинною маскою, а у класифікації — на Precision, Recall, F1, Ассурасу. Завдяки цим метрикам ми можемо всебічно проаналізувати, як добре система працює із задачами виділення та класифікації ознак стенозу та коронарних артерій.

4.7. Аналіз якості класифікації стенозу

У цьому підрозділі представлено порівняльний аналіз ефективності двох різних архітектур глибоких згорткових нейронних мереж – ResNet50 та EfficientNet-B0 – у задачі класифікації стенозу на основі локальних фрагментів коронарних ангіографічних зображень. Обидві моделі було навчено на однаковому датасеті, який містив приклади чотирьох ступенів стенозу (0 – відсутній, 1 – помірний, 2 – виражений, 3 – критичний). Оцінка результатів здійснювалась на тестовій вибірці за допомогою макроусереднених метрик точності (Ассурасу), повноти (Recall), точності (Precision) та F1-міри (табл. 4.1, 4.2), а також побудовано гістограми та графіки «moving average» (рис. 4.8, 4.9).

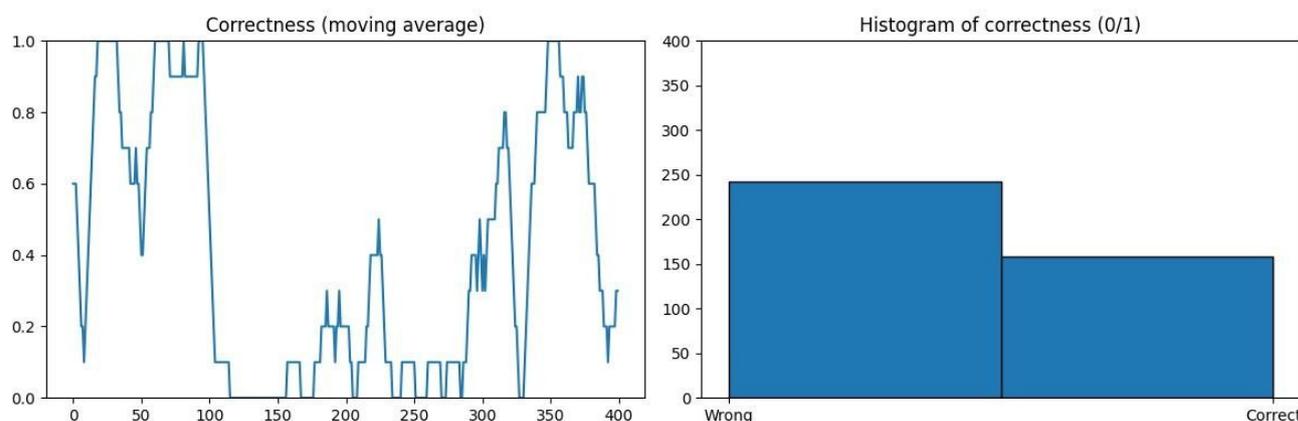


Рисунок 4.8. Графік «moving average» та гістограма моделі ResNet

Таблиця 4.1. Розрахунок показників тестування моделі ResNet

Accuracy	0.3950
Precision	0.3862
Recall	0.3950
F1-score	0.3421

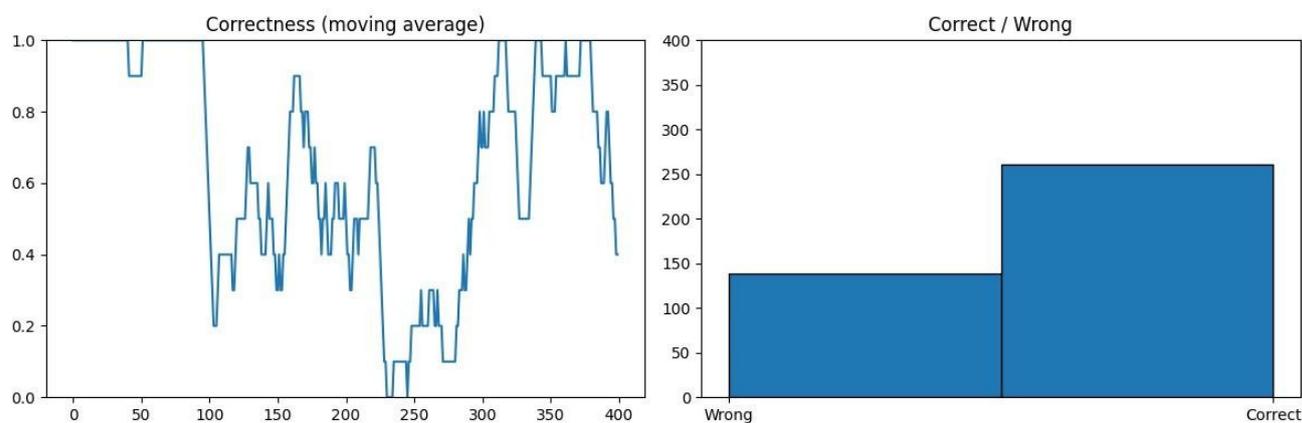


Рисунок 4.9. Графік «moving average» та гістограма моделі EfficientNet

Таблиця 4.2. Розрахунок показників тестування моделі EfficientNet

Accuracy	0.6525
Precision	0.7148
Recall	0.6525
F1-score	0.6300

Модель EfficientNet-B0 показала помітно вищі результати за всіма ключовими метриками. Загальна точність класифікації склала 65.25%, що майже вдвічі перевищує аналогічний показник для ResNet50 (39.50%). За макроусередненою F1-мірою модель EfficientNet також демонструє перевагу — 0.6300 проти 0.3421 у ResNet, що вказує на суттєво кращий баланс.

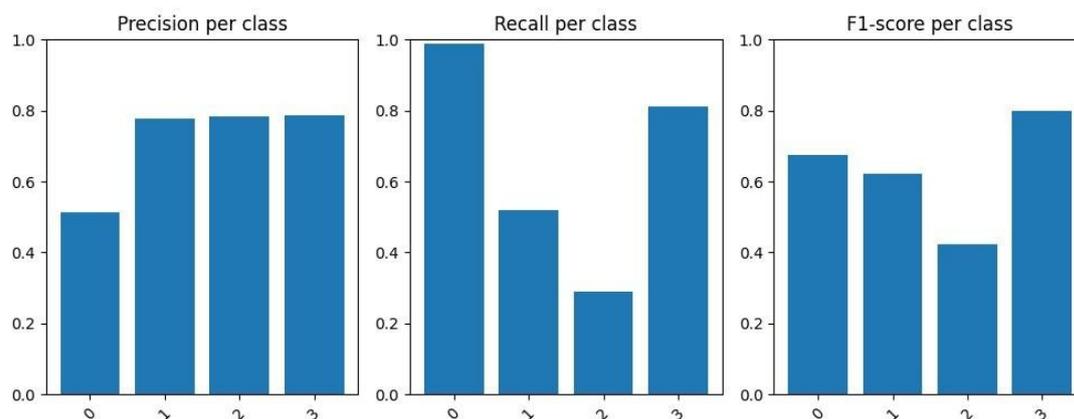


Рисунок 4.10. Стовпчикова діаграма для 4 класів моделі EfficientNet

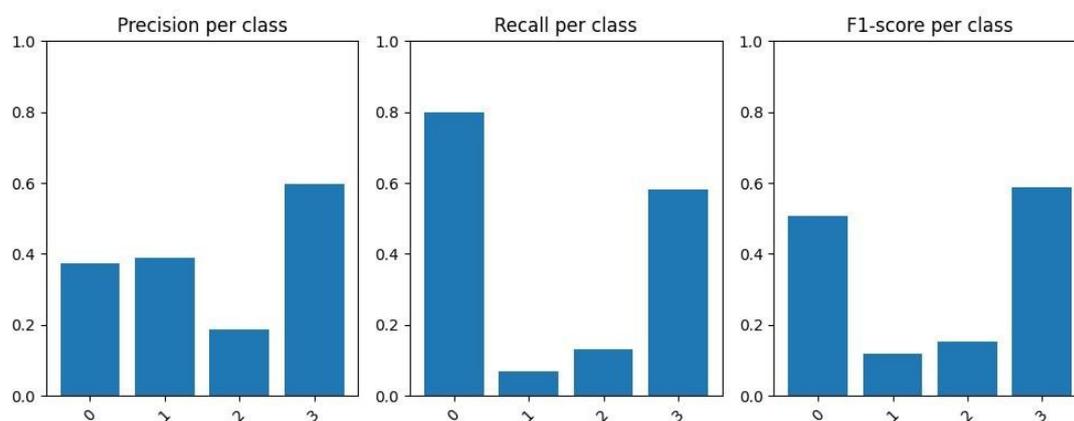


Рисунок 4.11. Стовпчикова діаграма для 4 класів моделі ResNet

При аналізі пер-класових результатів видно, що EfficientNet зберігає стабільну точність для всіх чотирьох категорій стенозу (рис. 4.10). Особливо виразною є перевага для критичної форми (клас 3), де F1-міра досягла 0.7980 проти 0.5888 у ResNet. Також у класі 1 (помірний стеноз) модель EfficientNet досягла $F1=0.6228$, що є більш ніж у п'ять разів вищим, ніж результат ResNet для того ж класу ($F1=0.1186$). Найгірший результат у EfficientNet спостерігається для класу 2 (виражений стеноз), де F1-міра склала 0.4234, однак навіть у цьому випадку він перевищує аналогічний показник ResNet (0.1529) (рис. 4.11), що свідчить про загальну здатність EfficientNet краще справлятися з тонкими відмінностями між класами.

$$\begin{pmatrix} 80 & 3 & 15 & 2 \\ 61 & 7 & 26 & 6 \\ 51 & 5 & 13 & 31 \\ 23 & 3 & 16 & 58 \end{pmatrix}$$

Рисунок 4.12. Матриця плутанини моделі ResNet

$$\begin{pmatrix} 99 & 0 & 0 & 1 \\ 33 & 52 & 4 & 11 \\ 50 & 11 & 29 & 10 \\ 11 & 4 & 4 & 81 \end{pmatrix}$$

Рисунок 4.13. Матриця плутанини моделі EfficientNet

Матриця плутанини підтверджує ці висновки (рис. 4.12, 4.13): EfficientNet забезпечує мінімальну кількість хибних позитивних класифікацій у класі 0 (здебільшого всі приклади класифіковані правильно), тоді як ResNet демонструє значне перетікання між класами – зокрема, моделі притаманна тенденція до хибної класифікації зображень класу 1 як класу 0, і навпаки.

4.8. Обраний метод для задачі класифікації стенозу

Якість класифікації стенозу аналізувалась на основі порівняння двох архітектур глибокого навчання – EfficientNet-B0 та ResNet50, які були навчені на однаковій вибірці локалізованих фрагментів ангіографічних зображень. Обидві моделі мали на меті диференціацію ступеня стенозу за чотирма класами: 0 – відсутній стеноз, 1 – помірний, 2 – виражений, 3 – критичний (рис. 4.14).

Аналіз результатів включав як глобальні макроусереднені метрики (точність, повнота, F1-міра), так і окрему оцінку ефективності по кожному з класів.

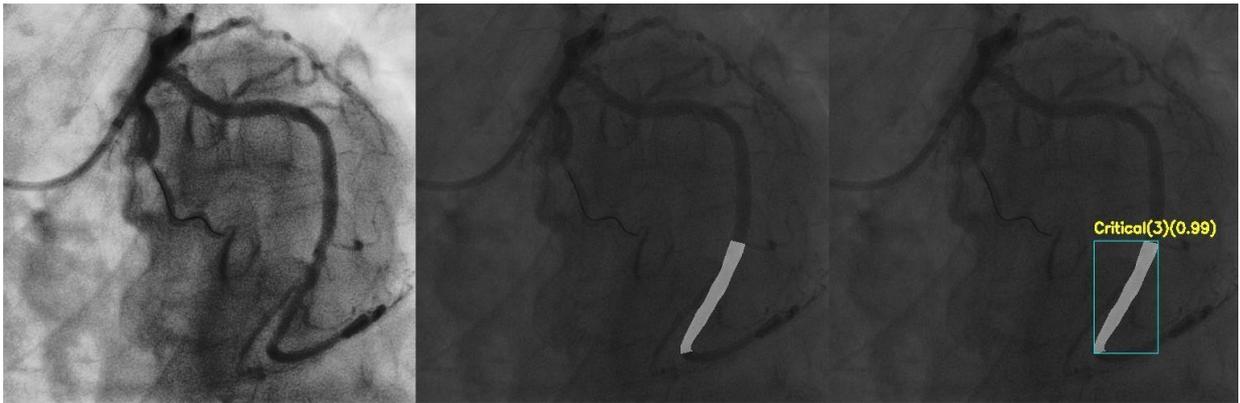


Рисунок 4.14. Приклад класифікації стенозу
(зліва – оброблене вхідне зображення, посередині – сегментований
стеноз, справа– результат класифікації стенозу)

Архітектура EfficientNet-B0 показала значну перевагу за всіма ключовими критеріями. Загальна точність класифікації склала 65.25%, що майже вдвічі перевищує результат моделі ResNet50 (39.50%). Показник макроусередненої F1міри для EfficientNet-B0 сягнув 0.6300, що свідчить про добрий баланс між точністю та чутливістю. Для порівняння, відповідна метрика у ResNet50 склала лише 0.3421, що вказує на значно гіршу здатність моделі до розпізнавання окремих класів.

Пер-класовий аналіз демонструє, що модель EfficientNet-B0 найбільш ефективно класифікує критичний стеноз (клас 3), де досягнуто F1-міри 0.7980. Також стабільні показники спостерігаються у класах 0 та 1 (F1 = 0.6758 і 0.6228 відповідно), що свідчить про високий рівень чутливості навіть для початкових форм патології. Найменш чітко розрізнення відбулося для класу 2 (виражений стеноз), де F1-міра становила 0.4234, однак і цей показник перевищує аналогічне значення ResNet (0.1529). Модель ResNet50, натомість, виявилась неспроможною надійно класифікувати класи 1 і 2 — F1-міри становили лише 0.1186 і 0.1529 відповідно, що є критично низькими значеннями для медичних задач.

Додаткову інформацію надала матриця плутанини, яка дозволила ідентифікувати типові патерни помилок. Для ResNet50 характерне суттєве

перетікання між сусідніми класами, особливо між 0 та 1, що свідчить про обмежену здатність моделі вловлювати тонкі візуальні відмінності між нормою та початковою стадією ураження. Натомість EfficientNet-B0 демонструє чіткі межі між класами, зокрема, майже ідеальне розпізнавання класу 0 (відсутній стеноз) і високий ступінь впевненості у критичних випадках (клас 3).

Таким чином, отримані результати свідчать про кращу узагальнюючу здатність EfficientNet-B0 у задачах багатокласової класифікації стенозу. Завдяки своїй внутрішній оптимізації параметрів та ефективному використанню попередньо натренованих ознак, ця архітектура показала стабільну якість класифікації навіть за умов обмеженого обсягу даних.

4.9. Розв'язання задачі класифікації стенозу

4.9.1. Архітектура та реалізація системи.

Розроблена система автоматизованої діагностики стенозу коронарних артерій реалізована у вигляді багатокрокового каскаду обробки ангиографічних зображень. Архітектура системи (рис. 4.15) побудована таким чином, що кожен етап аналізу базується на результатах попереднього, забезпечуючи логічну послідовність та високу точність кінцевих оцінок.



Рисунок 4.15. Діаграма архітектури роботи системи автоматизованої діагностики стенозу коронарних артерій

Загальна послідовність обробки включає попередню підготовку зображень, сегментацію судинних структур, локалізацію ділянок

патологічних змін (стенозів) і визначення ступеня їх тяжкості. На таблиці 5.3 представлено детальну логіку проходження даних системою.

Таблиця 4.3. Логіка проходження даних через систему автоматизованої діагностики стенозу

№	Етап обробки	Вхідні дані	Модель / Метод	Вихідні дані	Призначення етапу
1	Завантаження зображення	Ангіографічне зображення (формат PNG/JPEG)	—	Оригінальне зображення	Початковий вхід до системи
2	Попередня обробка	Оригінальне зображення	CLAHE, Zscore, DnCNN	Покращене зображення	Нормалізація, шумозниження, підвищення контрасту
3	Сегментація коронарних артерій	Оброблене зображення	U-Net	Маска судинного русла	Визначення анатомічної структури
4	Сегментація стенозу	Оброблене зображення + маска судин	Другий U-Net	Маска звужених ділянок	Виокремлення підозрілих зон
5	Локалізація ділянок стенозу (ROI)	Маска стенозу	Postprocessing	Фрагменти зображень із стенозом	Виділення областей для класифікації
6	Класифікація ступеня стенозу	ROI зображення	EfficientNet -B0	Клас стенозу: 0–3	Кількісна оцінка ступеня звуження
7	Візуалізація результату	Маски + класифіковані фрагменти	—	Позначене зображення з підписами	Інтерпретація результату для лікаря

Реалізація всіх компонентів здійснена на базі бібліотек PyTorch і TensorFlow. Попередня обробка включає використання денойзинг-моделі DnCNN для фільтрації шумів, нормалізацію інтенсивностей методом Z-score і покращення локального контрасту за допомогою алгоритму CLAHE.

Перша сегментаційна модель реалізована на класичній архітектурі U-Net і навчена на розмічених ангиограмах з використанням функції втрат Dice Loss для боротьби з дисбалансом класів.

Друга сегментаційна модель також побудована на базі U-Net, але приймає на вхід два канали — оброблене зображення та маску судин, що дозволяє зосередити аналіз виключно на судинних структурах.

Класифікація стенозів здійснюється з використанням моделі EfficientNetB0, яка навчена на балансованих ROI-фрагментах зображень. Вибір цієї моделі зумовлений її здатністю ефективно працювати з обмеженим набором даних та забезпечувати оптимальний баланс між продуктивністю та точністю.

Таким чином, архітектура системи забезпечує чіткий поділ завдань на окремі логічні блоки, кожен з яких реалізований з урахуванням особливостей даних і специфіки медичної задачі.

4.9.2. Спосіб застосування системи.

Користувач завантажує вихідне ангиографічне зображення (PNG, JPEG, попередньо перетворений DICOM-файл) в систему через спеціалізований інтерфейс. Після завантаження відбувається автоматична попередня обробка (фільтрація шумів, нормалізація яскравості, покращення контрасту).

На наступному етапі система самостійно виконує сегментацію судинних структур. Результатом цього етапу є маска судинного русла, яка визначає область подальшого аналізу. Потім автоматично здійснюється сегментація ділянок стенозу всередині отриманої маски судин.

Виділені ділянки патологічного звуження автоматично локалізуються, після чого система класифікує їх за ступенем тяжкості (відсутній, помірний, виражений, критичний). Підсумковий результат виводиться у вигляді візуально зрозумілих зображень з підписами (маски судин, маски стенозів,

ступінь стенозу), які можна експортувати для подальшого клінічного використання.

Для наочного розуміння процесу застосування системи на рисунках 4.16-4.18 наведені конкретні приклади послідовної обробки ангиографічних зображень на всіх етапах.

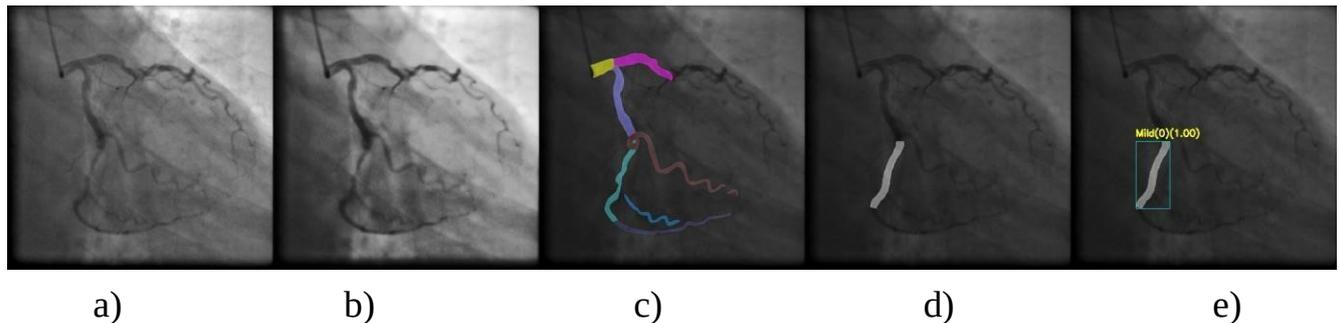


Рисунок 4.16. Приклад роботи системи №1.
 а) - Вхідне зображення б) - Попередня обробка зображення с) - Сегментація коронарних артерій д) - Сегментація стенозу е) - Класифікація стенозу

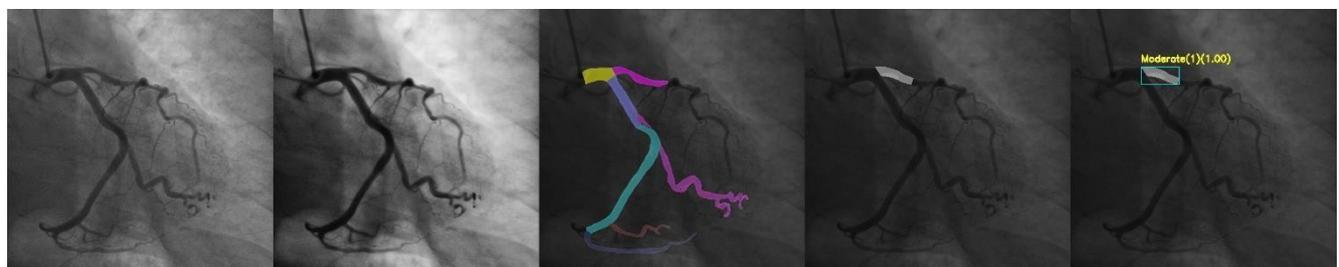


Рисунок 4.17. Приклад роботи системи №2



Рисунок 4.18. Приклад роботи системи №3

Таким чином, спосіб застосування системи є зручним для лікарів і не потребує спеціальних технічних навичок, дозволяючи зосередитись на аналізі отриманих результатів та прийнятті клінічних рішень.

4.9.3. Узагальнення результатів роботи системи.

Для об'єктивної оцінки ефективності запропонованої системи автоматизованої діагностики стенозу було проведено тестування на окремій вибірці ангіографічних зображень, що не використовувалися під час навчання моделей. Якість роботи кожного етапу системи – сегментації коронарних артерій, сегментації стенозів та класифікації ступеня стенозу – оцінювалася за метрикою F1-score, яка враховує як точність (precision), так і повноту (recall) виявлення.

Таблиця 4.4. Узагальнені результати тестування

Етап обробки	Модель	F1-score
Сегментація коронарних артерій	U-Net	$\approx 0,72$
Сегментація стенозу	U-Net	$\approx 0,48$
Класифікація стенозу	EfficientNet-B0	$\approx 0,65$

Аналіз отриманих результатів (табл. 4.4) свідчить, що модель для сегментації коронарних артерій забезпечує високий рівень точності ($F1 \approx 0,72$), що є важливим для надійного виявлення судинної архітектури навіть на зображеннях із помітним рівнем шуму та артефактами.

Сегментація стенозів виявилася складнішою задачею через невелику площу зон патології відносно загального фону, що відображено у нижчому значенні F1-score ($\approx 0,48$). Проте, такий результат є типовим для медичних задач локалізації невеликих об'єктів на складному фоні, особливо без використання додаткової спеціалізованої розмітки.

Класифікаційна модель EfficientNet-B0 продемонструвала прийнятний рівень точності визначення ступеня стенозу ($F1 \approx 0,65$), що дозволяє її використовувати

як допоміжний інструмент для підтримки прийняття клінічних рішень.

Важливо відзначити, що всі моделі системи були навчені та протестовані на реальних клінічних ангіографічних даних з використанням датасетів ARCADE і CADICA, що підвищує зовнішню валідність отриманих результатів і дозволяє сподіватися на їх успішну генералізацію в умовах практичного застосування.

Узагальнюючи, реалізована система показала ефективність кожного з ключових етапів обробки ангіографічних зображень і продемонструвала перспективність використання методів глибокого навчання для задач автоматизованої діагностики стенозу коронарних артерій.

4.9.4. Результати класифікації.

На основі попередньо сегментованих зон стенозу було сформовано набір локалізованих фрагментів, що належать до чотирьох класів: 0 – відсутній стеноз, 1 – помірний, 2 – виражений, 3 – критичний.

Для задачі класифікації було обрано архітектуру EfficientNet-B0 як найбільш збалансовану за точністю, швидкодією та узагальнювальною здатністю. У якості альтернативи використовувалася модель ResNet50.

Навчання проводилося на основі уніфікованого датасету з аугментацією та нормалізацією, що забезпечувало стабільність вхідних даних. Після навчання виконано тестування на незалежній вибірці із застосуванням макроусереднених метрик: Accuracy, Precision, Recall та F1-score.

Результати класифікації для EfficientNet-B0 становили:

- Accuracy — 65.25%;
- Precision — 71.48%;
- Recall — 65.25%;
- F1-score — 0.6300.

Пер-класовий аналіз показав, що найкращі результати були досягнуті для класу 3 ($F1 = 0.7980$), що є критично важливим для клінічного застосування. У той час як ResNet50 продемонстрував значно нижчу $F1$ -міру (0.3421) і високу плутанину між сусідніми класами, особливо між помірним і вираженим стенозом.

Таким чином, реалізована класифікаційна модель успішно завершує каскад обробки зображень у розробленій системі, дозволяючи не лише локалізувати, але й кількісно оцінювати ступінь ураження судини.

Висновки до розділу 4

У ході виконання задачі класифікації було реалізовано повноцінну модель глибокого навчання, орієнтовану на автоматичну диференціацію ступеня стенозу за ангиографічними зображеннями. Встановлено, що найкращі результати демонструє архітектура EfficientNet-B0, яка забезпечує достатньо високу точність при оптимальному обсязі обчислень.

Важливим є те, що модель досягає найвищих показників у розпізнаванні критичного стенозу — стану, що має найбільшу клінічну значущість. При цьому слабші результати спостерігаються для помірного стенозу, що пояснюється складністю його візуального відокремлення.

Порівняння з ResNet50 підтвердило перевагу EfficientNet-B0 не лише за загальними метриками, але й за стійкістю до плутанини між класами. Це дозволяє вважати обрану модель придатною до інтеграції в медичну систему підтримки прийняття рішень.

Таким чином, задача класифікації стенозу вирішена ефективно і логічно завершує побудову інтегрованої архітектури автоматизованої діагностики.

ВИСНОВКИ

Розроблена система автоматизованої діагностики стенозу коронарних артерій показала свою ефективність у вирішенні основних задач сегментації та класифікації патологічних змін на ангиографічних зображеннях. Проте, враховуючи динамічний розвиток медичних технологій і штучного інтелекту, існує низка напрямів, у яких система може бути вдосконалена та розширена.

Одним із перспективних напрямів є навчання системи розрізняти характер стенозу, зокрема, відрізняти хронічні стабільні стенози від гострих уражень, пов'язаних із нестабільними атеросклеротичними бляшками. Визначення стабільності бляшки має надзвичайно важливе клінічне значення, оскільки нестабільні бляшки частіше призводять до гострих коронарних синдромів. Для вирішення цієї задачі необхідно залучати додаткові джерела даних, зокрема інтегрувати ознаки, пов'язані зі структурою стінки судини та особливостями заповнення контрастом.

Оптимізація продуктивності системи є ще одним важливим завданням. Впровадження технологій компресії моделей і оптимізації інференсу, зокрема через використання TensorRT або ONNX Runtime, дозволить зменшити час обробки зображень і забезпечити роботу системи у реальному часі, що є критичним для інтеграції в клінічні робочі процеси.

Також перспективним є розвиток інтерфейсних рішень для зручної інтеграції системи в медичні інформаційні системи (PACS/RIS). Розробка вебінтерфейсу або мобільного додатку дозволить забезпечити доступність системи для лікарів у будь-якому місці та сприятиме поширенню технології телемедицини.

Нарешті, важливою задачею залишається клінічна валідація розробленої системи в умовах реальної медичної практики. Проведення багатоцентрових клінічних досліджень дозволить об'єктивно оцінити ефективність і безпечність використання системи як допоміжного інструменту в діагностиці ішемічної хвороби серця.

Таким чином, подальший розвиток системи спрямований на підвищення точності, надійності, функціональності та клінічної значущості її застосування в кардіологічній практиці. У цьому підрозділі реалізовано завершений цикл вирішення задачі класифікації стенозу на ангиографічних зображеннях, що охоплює побудову, навчання, тестування моделі та аналіз результатів. Метою є автоматизоване визначення ступеня звуження коронарної артерії за ROI-фрагментами, які містять ділянки патології.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. European Society of Cardiology. 2023 ESC Guidelines for the Management of Acute Coronary Syndromes. *European Heart Journal*, 2023, Vol. 44, No. 28, pp. 2391–2513. URL: <https://www.escardio.org/Guidelines/Clinical-Practice-Guidelines/Acute-Coronary-Syndromes-ACS-Guidelines>
2. Toth G.G., Johnson N.P., Jeremias A. et al. Current Status and Future Perspectives of Fractional Flow Reserve. *Cardiovascular Intervention and Therapy*, 2022, Vol. 14, No. 6, pp. 560–572. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10279845/>
3. Ibanez B., Heusch G., Ovize M. et al. Acute Myocardial Infarction. *Nature Reviews Disease Primers*, 2018, Vol. 4, Art. 18017. DOI: 10.1038/nrdp.2018.17. URL: <https://pubmed.ncbi.nlm.nih.gov/28886621/>
4. Braunwald E. *Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine*. 6-th ed. Philadelphia: Elsevier, 2001. 2159 p. URL: <http://www.arrytmiaonline.com.br/files/Braunwald---Heart-Disease---A-Textbook-of-Cardiovascular-Medicine-6th-ed.pdf>
5. Garcia-Canadilla P., Sanchez-Gonzalez J., Sampedro-Gomez J. et al. Dataset for Automatic Region-Based Coronary Artery Disease Evaluation (ARCADE). *Scientific Data*, 2023, Vol. 10, Art. 362. DOI: 10.1038/s41597-023-02871-z. URL: <https://www.nature.com/articles/s41597-023-02871-z>
6. Philips Healthcare. Azurion 7 C20 with FlexArm: Specifications. 2019. URL: https://www.documents.philips.com/doclib/enc/18207959/Specifications_Azurion_7_C20_with_FlexArm.pdf
7. Park E.A., Kim J.H., Han K. Myocardial Late Enhancement Using ThirdGeneration Dual-Source CT (SOMATOM Force). *Insights into Imaging*, 2025, Vol. 15. DOI: 10.1186/s13244-025-01944-4. URL: <https://insightsimaging.springeropen.com/articles/10.1186/s13244-025-019444>
8. Siemens Healthineers. SOMATOM Force Dual-Source CT Scanner. URL: <https://www.siemens-healthineers.com/computed->

tomography/somatom/somatom-force

9. Andreini D. et al. Impact of the SnapShot Freeze 2 Algorithm Across Varied Heart Rates. *Journal of Cardiovascular Computed Tomography*, 2024, Vol. 19, No. 2, pp. 140–149. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11302822/>
10. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. 2015. arXiv:1512.03385. DOI: 10.48550/arXiv.1512.03385. URL: <https://arxiv.org/abs/1512.03385>
11. Rostami A., Khadangi A., Javanmardi M. StenUNet: Automatic Stenosis Detection from X-ray Coronary Angiography. 2024. arXiv:2412.02568. DOI: 10.48550/arXiv.2412.02568. URL: <https://arxiv.org/abs/2310.14961>
12. Algarni M., Al-Rezqi A., Saeed F. et al. Multi-Constraints Based Deep Learning Model for Automated Segmentation and Diagnosis of Coronary Artery Disease in X-Ray Angiographic Images. *PeerJ Computer Science*, 2022, Vol. 8, Art. e993. DOI: 10.7717/peerj-cs.993. URL: <https://doi.org/10.7717/peerj-cs.993>
13. Schwarzhans F., Wiele G., Haase K. et al. Intensity Normalization Techniques and Their Effect on the Robustness and Predictive Power of Breast MRI Radiomics. 2024. arXiv:2406.01736. DOI: 10.48550/arXiv.2406.01736. URL: <https://arxiv.org/abs/2406.01736>
14. Johnson W.E., Li C., Rabinovic A. Adjusting Batch Effects in Microarray Expression Data Using Empirical Bayes Methods. *Biostatistics*, 2007, Vol. 8, No. 1, pp. 118-127. DOI: 10.1093/biostatistics/kxj037. URL: <https://doi.org/10.1093/biostatistics/kxj037>
15. Namazi Nia S., Shih F.Y. Medical X-Ray Image Enhancement Using Global Contrast-Limited Adaptive Histogram Equalization. 2024. arXiv:2411.01373. DOI: 10.48550/arXiv.2411.01373. URL: <https://arxiv.org/abs/2411.01373>

16. Zhang K., Zuo W., Chen Y. et al. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. 2016. arXiv:1608.03981. DOI: 10.48550/arXiv.1608.03981. URL: <https://arxiv.org/abs/1608.03981>
17. Liang J., Cao J., Sun G. et al. SwinIR: Image Restoration Using Swin Transformer. 2021. arXiv:2108.10257. DOI: 10.48550/arXiv.2108.10257. URL: <https://arxiv.org/abs/2108.10257>
18. Dabov K., Foi A., Katkovnik V., Egiazarian K. Image Denoising by Sparse 3D Transform-Domain Collaborative Filtering. IEEE Transactions on Image Processing, 2007, Vol. 16, No. 8, pp. 2080–2095. DOI: 10.1109/TIP.2007.901238. Відкрита версія: arXiv:2006.14280. URL: https://www.researchgate.net/publication/6151802_Image_Denoising_by_Sparse_3-D_Transform-Domain_Collaborative_Filtering
19. Liu T., Lin H., Katsaggelos A.K., Kline A. YOLO-Angio: An Algorithm for Coronary Anatomy Segmentation. 2023. arXiv:2310.15898. DOI: 10.48550/arXiv.2310.15898. URL: <https://arxiv.org/abs/2310.15898>
20. Bilal M., Martinho D., Sim R. et al. Multivessel Coronary Artery Segmentation and Stenosis Localisation Using Ensemble Learning. 2023. arXiv:2310.17954. DOI: 10.48550/arXiv.2310.17954. URL: <https://arxiv.org/abs/2310.17954>
21. Huang B., Luo Y., Wei G. et al. A Deep Learning Model for Coronary Artery Segmentation and Quantitative Stenosis Detection in Angiographic Images. 2025. arXiv:2406.00492. DOI: 10.48550/arXiv.2406.00492. URL: <https://arxiv.org/abs/2406.00492>
22. Kirillov A., Mintun E., Ravi N. et al. Segment Anything. 2023. arXiv:2304.02643. DOI: 10.48550/arXiv.2304.02643. URL: <https://arxiv.org/abs/2304.02643>
23. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. 2015. arXiv:1505.04597. DOI: 10.48550/arXiv.1505.04597. URL: <https://arxiv.org/abs/1505.04597>

24. Frangi A.F., Niessen W.J., Vincken K.L. et al. Multiscale Vessel Enhancement Filtering. In: Lecture Notes in Computer Science, 1998, Vol. 1496, pp. 130–137. DOI: 10.1007/BFb0056195. URL: <https://link.springer.com/chapter/10.1007/BFb0056195>
25. Lee I.K., Shin J., Lee Y.H. et al. SSASS: Semi-Supervised Approach for Stenosis Segmentation. 2023. arXiv:2311.10281. DOI: 10.48550/arXiv.2311.10281. URL: <https://arxiv.org/abs/2311.10281>
26. Carion N., Massa F., Synnaeve G. et al. End-to-End Object Detection with Transformers (DETR). ECCV 2020. arXiv:2005.12872. DOI: 10.48550/arXiv.2005.12872. URL: <https://arxiv.org/abs/2005.12872>
27. Liu X., Chen Y., Wang J. et al. DINO: Grounded Object Detection with Transformers. 2023. arXiv:2303.05499. DOI: 10.48550/arXiv.2303.05499. URL: <https://arxiv.org/abs/2303.05499>
28. Tan M., Le Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. 2019. arXiv:1905.11946. DOI: 10.48550/arXiv.1905.11946. URL: <https://arxiv.org/abs/1905.11946>
29. Cong C., Kato Y., Vasconcellos H.D.D., Ostovaneh M.R., Lima J.A.C., Ambale-Venkatesh B. Deep Learning-Based End-to-End Automated Stenosis Classification and Localization on Catheter Coronary Angiography. *Frontiers in Cardiovascular Medicine*, 2023, Vol. 10:944135. DOI: 10.3389/fcvm.2023.944135. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC9941145/>
30. Duan H., Yi S., Ren Y. DCA-YOLOv8: A Novel Framework Combined with AICI Loss Function for Coronary Artery Stenosis Detection. *Sensors*, 2024, 24(24):8134. DOI: 10.3390/s24248134. URL: <https://www.mdpi.com/1424-8220/24/24/8134>
31. Kaba Ş., Hacı H., Isin A., İlhan A., Conkbayır C. The Application of Deep Learning for the Segmentation and Classification of Coronary Arteries.

Diagnostics, 2023, 13(13):2274. DOI: 10.3390/diagnostics13132274. URL: <https://www.mdpi.com/2075-4418/13/13/2274>

32. Singh R., Ghosal A., Maheshwari V. CADICA Dataset for Coronary Artery Disease Evaluation from Invasive Coronary Angiography. Engineering Applications of Artificial Intelligence, 2023. DOI: 10.1016/j.engappai.2023.107086. URL: <https://arxiv.org/abs/2402.00570>

33.

ДОДАТОК А. ЛІСТИНГ ПРОГРАМИ

A.1. Код для попередньої обробки зображень

Фрагмент реалізує етапи фільтрації шуму, нормалізації інтенсивностей і покращення контрасту зображень перед подачею на вхід моделі.

```

class DnCNN(nn.Module):
    def
    __init__(self, channels=1, num_layers=17):
        super(DnCNN, self).__init__()
        layers =
        [nn.Conv2d(channels, 64, kernel_size=3, padding=1, bias=False),
        nn.ReLU(inplace=True)]
        for _ in range(num_layers - 2):
            layers.append(nn.Conv2d(64, 64, kernel_size=3,
            padding=1, bias=False))
        layers.append(nn.BatchNorm2d(64))
        layers.append(nn.ReLU(inplace=True))
        layers.append(nn.Conv2d(64, channels, kernel_size=3, padding=1,
        bias=False))
        self.dncnn = nn.Sequential(*layers)
    def
    forward(self, x):
        return x - self.dncnn(x)
    def denoise_dncnn_soft(original_gray, dncnn_model,
    alpha=0.5):
        t =
        torch.from_numpy(original_gray).float().unsqueeze(0).unsqueeze(0)/255.0
        with torch.no_grad():
            out = dncnn_model(t)
            out_np =
            out.squeeze(0).cpu().numpy().transpose(1,2,0)*255.0
            out_np =
            np.clip(out_np, 0, 255).astype(np.uint8)
            result =
            cv2.addWeighted(original_gray, alpha, out_np, 1 - alpha, 0)
            return
            result
        def soft_minmax(original_gray,
        alpha=0.5):
            imf =
            original_gray.astype(np.float32)
            mn, mx = imf.min(), imf.max()
            if mx > mn:
                norm = (imf - mn)/(mx -
            mn)
            else:
                norm = imf*0
                norm_255 =
                (norm*255).clip(0,255).astype(np.uint8)
                result =
                cv2.addWeighted(original_gray, alpha, norm_255, 1 - alpha, 0)
            return result
        def soft_equalizeHist(original_gray,
        alpha=0.5):
            eq = cv2.equalizeHist(original_gray)
            result =
            cv2.addWeighted(original_gray, alpha, eq, 1 - alpha, 0)
            return result

```

A.2. Код для тренування моделі сегментації артерій

Фрагмент коду відповідає за побудову, компіляцію та навчання U-Net моделі для сегментації коронарних артерій.

```

class CoronaryArteryDataset(Dataset):    def __init__(self,
images_dir, masks_dir, transform=False, nc=26):
    super().__init__()
self.nc = nc
self.transform_flag = transform
    valid_img_ext = ('.png', '.jpg',
'.jpeg')    self.images_paths =
sorted(    os.path.join(images_dir,
r, f)    for f in
os.listdir(images_dir)    if
f.lower().endswith(valid_img_ext)
)    self.masks_paths
=
sorted(    os.path.join(masks_dir,
sks_dir, f)    for f in
os.listdir(masks_dir)
if f.lower().endswith('.png')
)    self.color_jitter =
T.ColorJitter(brightness=0.2, contrast=0.2)
    def
__len__(self):
    return len(self.images_paths)
    def
__getitem__(self, idx):
    img_path = self.images_paths[idx]
mask_path = self.masks_paths[idx]
    img_bgr = cv2.imread(img_path, cv2.IMREAD_COLOR)
image = cv2.cvtColor(img_bgr, cv2.COLOR_BGR2RGB)    mask =
cv2.imread(mask_path, cv2.IMREAD_GRAYSCALE).astype(np.int64)
    if
self.transform_flag:
if random.random() < 0.5:
        image = np.flip(image,
axis=1).copy()    mask =
np.flip(mask, axis=1).copy()    if
random.random() < 0.2:
        image = np.flip(image,
axis=0).copy()    mask =
np.flip(mask, axis=0).copy()
pil_img = Image.fromarray(image)
pil_img = self.color_jitter(pil_img)
image = np.array(pil_img)
        image_t =
torch.from_numpy(image.transpose(2,0,1)).float()/255.0
mask_t = torch.from_numpy(mask).long()    return image_t,
mask_t
    def compute_class_frequencies(masks_dir,
nc=26):    freq = np.zeros(nc,
dtype=np.int64)    mask_files = [f for f
in os.listdir(masks_dir) if

```

```

f.lower().endswith('.png'
)]    for mask_name in
mask_files:
    path = os.path.join(masks_dir, mask_name)
    mask = cv2.imread(path, cv2.IMREAD_GRAYSCALE)
    counts = np.bincount(mask.flatten(), minlength=nc)
    freq += counts    return freq

def
train_unet_with_weights(    train_im
ages, train_masks,
val_images=None, val_masks=None,
epochs=100, batch_size=8,    lr=1e-
3, nc=26,
model_name='unet_segmentation_weighte
d',    class_freq=None,
use_augmentation=False
):    device = torch.device('cuda' if torch.cuda.is_available()
else 'cpu')
    train_ds = CoronaryArteryDataset(train_images, train_masks,
nc=nc,    transform=use_augmentation)    train_dl =
DataLoader(train_ds,    batch_size=batch_size,    shuffle=True,
num_workers=2)
    val_dl = None    if
val_images and val_masks:
        val_ds = CoronaryArteryDataset(val_images, val_masks,
nc=nc, transform=False)    val_dl = DataLoader(val_ds,
batch_size=batch_size, shuffle=False, num_workers=2)
        model = smp.Unet(
encoder_name='resnet34',
encoder_weights='imagenet',
in_channels=3,
classes=nc,
activation=None
)
    model.to(device
)

class_weights_tensor =
None    if class_freq
is not None:
    freq_arr = np.array(class_freq, dtype=np.float32)
    freq_arr += 1e-6    w_arr = 1.0 / np.sqrt(freq_arr)
    w_arr = w_arr / w_arr.mean()    class_weights_tensor =
torch.from_numpy(w_arr).float().to(device)
    print("Використовуються ваги класів:", w_arr)
    loss_fn =
nn.CrossEntropyLoss(weight=class_weights_tensor)
    optimizer = optim.Adam(model.parameters(), lr=lr)
    for epoch in
range(epochs):
        model.train()
    total_loss = 0.0    for
(imgs, masks) in train_dl:
        imgs, masks = imgs.to(device), masks.to(device)
    optimizer.zero_grad()    outs = model(imgs)
    loss = loss_fn(outs, masks)    loss.backward()

```

```

optimizer.step()                total_loss += loss.item() *
imgs.size(0)
    avg_loss = total_loss / len(train_ds)
print(f"Epoch [{epoch+1}/{epochs}] TrainLoss: {avg_loss:.4f}")
    if
val_dl:
    val_loss = 0.0
model.eval()                    with
torch.no_grad():                for
(vimgs, vmasks) in val_dl:
    vimgs, vmasks = vimgs.to(device),
vmasks.to(device)              vouts = model(vimgs)
vloss = loss_fn(vouts, vmasks)    val_loss +=
vloss.item() * vimgs.size(0)      val_loss /= len(val_ds)
print(f"    ValLoss: {val_loss:.4f}")
    torch.save(model.state_dict(),
f"{model_name}.pth")    print(f"Модель
збережено => {model_name}.pth")    return model

```

A.3. Код для тренування моделі сегментації стенозу

Фрагмент відповідає за навчання моделі U-Net для локалізації ділянок стенозу.

```

class CoronaryArteryDataset(Dataset):
    def __init__(self, images_dir, masks_dir, transform=None,
nc=26):
        super().__init__()
        self.images_paths = sorted([
os.path.join(images_dir, f) for f in os.listdir(images_dir) if
f.endswith(('.png', '.jpg'))
        ])
        self.masks_paths
= sorted([
os.path.join(masks_dir, f) for f in
os.listdir(masks_dir) if f.endswith('.png')
        ])
        self.transform = transform
self.nc = nc
    def
__len__(self):
        return len(self.images_paths)
    def
__getitem__(self, idx):
img_path =
self.images_paths[idx]
mask_path =
self.masks_paths[idx]
image =
cv2.imread(img_path,
cv2.IMREAD_COLOR)
image =
cv2.cvtColor(image,
cv2.COLOR_BGR2RGB)
mask =
cv2.imread(mask_path,
cv2.IMREAD_GRAYSCALE).as
type(np.int64)
        image = torch.from_numpy(image.transpose(2, 0, 1)).float()
/ 255.0
        mask = torch.from_numpy(mask).long()
        return
image, mask

```

```

def train_unet(    train_images,
train_masks,    val_images=None,
val_masks=None,    epochs=500,
batch_size=8,    lr=1e-3, nc=26,
model_name='unet_segmentation'
):
    device = torch.device('cuda' if torch.cuda.is_available() else
'cpu')    train_ds = CoronaryArteryDataset(train_images,
train_masks, nc=nc)    train_dl = DataLoader(train_ds,
batch_size=batch_size, shuffle=True, num_workers=2)
    val_dl = None    if
val_images and val_masks:
        val_ds = CoronaryArteryDataset(val_images, val_masks,
nc=nc)    val_dl = DataLoader(val_ds, batch_size=batch_size,
shuffle=False, num_workers=2)
        model = smp.Unet(
encoder_name='resnet34',
encoder_weights='imagenet',
in_channels=3,
classes=nc,
activation=None
        )
    model.to(device
)
    loss_fn = nn.CrossEntropyLoss()
optimizer = optim.Adam(model.parameters(),
lr=lr)
    for epoch in
range(epochs):
        model.train()
total_loss = 0    for
images, masks in train_dl:
            images, masks = images.to(device),
masks.to(device)    optimizer.zero_grad()
outputs = model(images)    loss =
loss_fn(outputs, masks)    loss.backward()
optimizer.step()
            total_loss += loss.item() * images.size(0)
avg_loss = total_loss / len(train_ds)    print(f"Epoch
[{epoch+1}/{epochs}] - Loss: {avg_loss:.4f}")
        if
val_dl:
            val_loss = 0
model.eval()    with
torch.no_grad():    for
val_imgs, val_masks in val_dl:
                val_imgs, val_masks = val_imgs.to(device),
val_masks.to(device)
val_outs = model(val_imgs)
v_loss = loss_fn(val_outs, val_masks)
val_loss += v_loss.item() *
val_imgs.size(0)    val_loss /=
len(val_dl.dataset)    print(f"
Val Loss: {val_loss:.4f}")
            torch.save(model.state_dict(),
f"{model_name}.pth")    print(f"Model saved to
{model_name}.pth")    return model

```

A.4. Код для тренування моделі класифікації стенозу

Цей фрагмент реалізує модель класифікації (EfficientNet), що визначає ступінь стенозу за ROI-фрагментами.

```
def main():
    data_root = r"G:\PycharmProjects\magister_full_project\data\
classification_dataset\images_processed"
    transform =
transforms.Compose([
        transforms.Resize((224, 224)),
transforms.ToTensor(),
        transforms.Normalize([0.485, 0.456,
0.406],
        [0.229, 0.224, 0.225]),
    ])
    dataset_all = datasets.ImageFolder(root=data_root,
transform=transform)
    train_size = int(0.8 * len(dataset_all))
val_size = len(dataset_all) - train_size
    train_ds, val_ds =
random_split(dataset_all, [train_size, val_size])

    train_loader = DataLoader(train_ds, batch_size=8, shuffle=True,
num_workers=2)
    val_loader = DataLoader(val_ds, batch_size=8,
shuffle=False, num_workers=2)

    device = torch.device("cuda" if torch.cuda.is_available() else
"cpu")
    num_classes = len(dataset_all.classes)
    model = efficientnet_b0(pretrained=True)
model.classifier[1] =
nn.Linear(model.classifier[1].in_features, num_classes)
model.to(device)
    criterion = nn.CrossEntropyLoss()
optimizer = optim.Adam(model.parameters(),
lr=1e-3)
    epochs = 200
    for epoch in
range(epochs):
        model.train()
        train_loss,
train_correct, total = 0, 0, 0
        for
imgs, labels in train_loader:
            imgs, labels = imgs.to(device),
labels.to(device)
            optimizer.zero_grad()
            outputs = model(imgs)
            loss =
criterion(outputs, labels)
            loss.backward()
            optimizer.step()
            train_loss += loss.item() * imgs.size(0)
train_correct += (outputs.argmax(1) == labels).sum().item()
total += imgs.size(0)
        model.eval()
        val_loss,
val_correct, val_total = 0, 0, 0
        with torch.no_grad():
            for imgs,
labels in val_loader:
                imgs, labels = imgs.to(device), labels.to(device)
                outputs = model(imgs)
                v_loss = criterion(outputs,
labels)
                val_loss += v_loss.item() * imgs.size(0)
val_correct += (outputs.argmax(1) == labels).sum().item()
val_total += imgs.size(0)

        print(f"Epoch {epoch+1}/{epochs} | "
f"Train Loss: {train_loss/total:.4f}, Acc:
```

```

{train_correct/total:.4f} | "                                f"Val
Loss: {val_loss/val_total:.4f}, Acc:
{val_correct/val_total:.4f}")
                                torch.save(model.state_dict(),
"efficientnet_classifier_100.pth")                                print("Model saved ->
efficientnet_classifier_100.pth")
    if __name__ ==
"__main__":
        main()

```

A.5. Код Flask API для виклику моделі

Фрагмент реалізує REST API, що приймає зображення, викликає моделі сегментації та класифікації, і повертає результат у форматі JSON.

```

app = Flask(__name__)

@app.route("/predict",
methods=["POST"]) def predict():
    if 'image' not in request.files:
        return jsonify({"error": "No
image file found"}), 400

    t
    r
    y
    :
        img_file = request.files['image']
    pil_img = Image.open(img_file).convert("RGB")
    original_np = np.array(pil_img)
        gray = cv2.cvtColor(original_np,
cv2.COLOR_RGB2GRAY)
        step_dncnn =
denoise_dncnn_soft(gray, dncnn_model)
        step_minmax =
soft_minmax(step_dncnn)
        cont_out =
soft_equalizeHist(step_minmax)
        processed_3ch =
cv2.cvtColor(cont_out, cv2.COLOR_GRAY2RGB)
        tens_art =
seg_transform(processed_3ch).unsqueeze(0)
    with torch.no_grad():
        logits_art = unet_artery(tens_art)
    pred_artery = logits_art.argmax(dim=1)[0].cpu().numpy()
        tens_sten =
seg_transform(processed_3ch).unsqueeze(0)
        with
torch.no_grad():
            logits_sten = unet_stenosis(tens_sten)
    pred_sten = logits_sten.argmax(dim=1)[0].cpu().numpy()
        bin_artery = (pred_artery != 0).astype(np.uint8)
    sten_bin = (pred_sten != 0).astype(np.uint8)
    num_labels, labels_im = cv2.connectedComponents(sten_bin,
connectivity=4)
        for lbl in range(1, num_labels):
            comp_mask = (labels_im == lbl).astype(np.uint8)
    overlap = (comp_mask & bin_artery).sum() / comp_mask.sum()
    if overlap < 0.3:

```

```

        pred_sten[comp_mask == 1] = 0
        final_bin = (pred_sten != 0).astype(np.uint8)
    num_lbl2, labels_img2 = cv2.connectedComponents(final_bin,
    connectivity=4)
    overlay = original_np.copy()
    for lbl2 in range(1, num_lbl2):
        comp_mask2 = (labels_img2 ==
    lbl2).astype(np.uint8)
        ys, xs =
    np.where(comp_mask2 == 1)
        y_min, y_max =
    ys.min(), ys.max()
        x_min, x_max = xs.min(),
    xs.max()

        roi_np = processed_3ch[y_min:y_max + 1, x_min:x_max
    + 1, :]
        roi_pil = Image.fromarray(roi_np)
        roi_t = classif_transform(roi_pil).unsqueeze(0)
        with
    torch.no_grad():
            outs =
    efficientnet_classifier(roi_t)
            probs = F.softmax(outs, dim=1)[0]
            pred_c = torch.argmax(probs).item()
            conf = probs[pred_c].item()
            text = f"Class {pred_c} ({conf:.2f})"
            cv2.rectangle(overlay, (x_min, y_min), (x_max, y_max), (0, 255, 0),
    2)
            cv2.putText(overlay, text, (x_min, max(0, y_min -
    10)),
            cv2.FONT_HERSHEY_SIMPLEX, 0.5, (0,
    255, 255), 1)
            buf =
    io.BytesIO()
            Image.fromarray(ove
    rlay).save(buf,
            format='PNG')
            buf.seek(0)
            return
            send_file(buf,
            mimetype="image/png
            ")
        except
    Exception as e:
            return jsonify({"error": str(e)}), 500
    if __name__ ==
    "__main__":
        app.run(debug=True)

```