

Національний університет "Львівська політехніка"

Інститут комп'ютерних наук та інформаційних технологій



Кафедра систем штучного інтелекту
Спеціальність 122 «Комп'ютерні науки»



ПОЯСНЮВАЛЬНА ЗАПИСКА

до магістерської кваліфікаційної роботи на тему:

«Розроблення моделі глибокого навчання для класифікації зображень і виявлення об'єктів на медичних зображеннях»

Студента(ки) групи

КНСШ-23

Воробель А.О.

Керівник роботи

(підпис)

(Косаревич Р.Я.)

Консультанти

(підпис)

(_____)

(підпис)

(_____)

Завідувач

кафедри СШ

(Шаховська Н. Б.)

Національний університет “Львівська політехніка”
Інститут комп’ютерних наук та інформаційних технологій
Кафедра систем штучного інтелекту
Спеціальність 122 «Комп’ютерні науки»

ЗАТВЕРДЖУЮ:

Завідувач кафедри СШІ _____

“ ____ ” _____ 2023р.

**ЗАВДАННЯ
НА МАГІСТЕРСЬКУ КВАЛІФІКАЦІЙНУ РОБОТУ
СТУДЕНТОВІ**

Вороблю Адріану Олександровичу

1. Тема роботи: «Розроблення моделі глибокого навчання для класифікації зображень і виявлення об’єктів на медичних зображеннях » затверджена наказом університету № 4420-4-06 від 13.10.2023

2. Термін подання закінченої роботи: _ 01.12.2023

3. Вихідні дані до (проекту) роботи: модель нейронної мережі для виявлення та класифікації переломів на рентгенограмах.

4. Зміст розрахунково-пояснювальної записки: (перелік питань, що належить розробити): аналітичний огляд літературних та інших джерел, системний аналіз та обґрунтування проблеми, методи та засоби вирішення проблеми, практична реалізація.

5. Перелік графічного матеріалу (з точним зазначенням обов’язкових креслень): діаграма PRISMA, хронологія розвитку моделі, архітектора моделі, візуалізація метрик, рентгенівські знімки, візуалізація роботи методів обробки зображень, оцінкові метрики моделі.

6. Консультанти по роботі, із зазначенням розділів проекту, що стосується їх

Розділ	Консультант	Підпис, дата	
		Завдання видав	Завдання прийняв
Консультант з іноземної мови			

7. Дата видачі завдання 10.03.2023

Керівник роботи _____
(підпис)

Завдання прийняв до виконання _____
(підпис)

КАЛЕНДАРНИЙ ПЛАН

№ п/п	Назва етапів дипломної роботи	Термін виконання етапів роботи	Примітка
1	Огляд літератури	10.03.2023 – 03.04.2023	
2	Створення архітектури	03.04.2023 – 15.05.2023	
3	Розробка алгоритмічного та програмного забезпечення	15.05.2023 – 11.09.2023	
4	Експериментальні дослідження	11.09.2023 – 06.11.2023	
5	Оформлення записки	06.11.2023 – 30.11.2023	

Керівник роботи _____
(підпис)

Студент _____
(підпис)

АНОТАЦІЯ

Магістерська кваліфікаційна робота виконана студентом групи КНСШ-23 Вороблем Адріаном Олександровичем. Тема “Розроблення моделі глибокого навчання для класифікації зображень і виявлення об’єктів на медичних зображеннях”. Робота направлена на здобуття ступеня магістр за спеціальністю 122 «Комп’ютерні науки».

Об’єктом дослідження є дослідження процес класифікації зображень та виявлення об’єктів на медичних зображеннях.

Предметом дослідження є розробка рішення з використанням нейромережевої архітектури та методів глибокого навчання для досягнення найвищої точності та швидкості виявлення об’єктів на медичних зображеннях.

Досягнення мети відбувається за рахунок використання моделі YOLOv8 та її тренування. Апробацію роботи моделі здійснено на основі медичних наборів даних у відкритому доступі, що містять рентгенівські знімки, для розв’язання задачі виявлення переломів та їх класифікації.

У результаті виконання дипломної роботи розроблено дві моделі глибокого навчання, одна з яких дозволяє виявляти переломи на рентгенівських знімках та класифікувати їх на декілька класів, а інша — лише виявляти.

Загальний обсяг роботи: 56 сторінок, 40 рисунків, 21 посилання.

Ключові слова: комп’ютерний зір, класифікація, машинне навчання, глибокі неронні мережі, YOLOv8, рентгенівські знімки, перелом.

ABSTRACT

The master's qualification work was completed by the student of the CSAI-23 group Vorobel Adrian Oleksandrovych. Topic is "Development of a deep learning model for image classification and object detection in medical images". The work is aimed at obtaining a master's degree in the specialty 122 "Computer Science".

The object of research is the process of image classification and object detection on medical images.

The subject of the research is the development of a solution using neural network architecture and deep learning methods to achieve the highest accuracy and speed of object detection in medical images.

The goal is achieved by using the YOLOv8 model and its training. The approbation of the model's work was carried out on the basis of medical data sets in open access, containing X-ray images, to solve the problem of fracture detection and their classification.

As a result of the thesis, two deep learning models were developed, one of which allows you to identify fractures on X-ray images and classify them into several classes, and the other only detects them.

Total volume of work: 56 pages, 40 figures, 21 references.

Keywords: computer vision, classification, machine learning, deep neural networks, YOLOv8, X-rays, fracture.

ЗМІСТ

АНОТАЦІЯ	4
ABSTRACT	5
ВСТУП.....	7
1. АНАЛІТИЧНИЙ РОЗДІЛ	11
1.1. Методологія пошуку наукових джерел	11
1.2. Огляд літературних джерел	13
1.3. Критичний аналіз літературних джерел	18
1.4. Висновки до розділу 1	28
2. ДОСЛІДНИЦЬКИЙ РОЗДІЛ	29
2.1. Вибір моделі нейронної мережі	29
2.2. Аналіз архітектури вибраної моделі	31
2.3. YOLOv8	34
2.4. Вибір на аналіз наборів даних	37
2.5. Аналіз методів обробки навчального набору даних	39
2.6. Висновки до розділу 2	42
3. РОЗДІЛ АПРОБАЦІЙ ТА РЕЗУЛЬТАТІВ	44
3.1. Підготовка набору даних	44
3.2. Тренування моделі	46
3.3. Обговорення результатів	48
3.4. Висновки до розділу 3	59
ВИСНОВОК	60
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	62
ДОДАТОК А	65

ВСТУП

Використовуючи моделі глибокого навчання, можна автоматизувати процес аналізу медичних зображень і рентгенівських знімків. Використання таких моделей може скоротити час, необхідний для аналізу великої кількості зображень, і підвищити точність діагностики та лікування.

Сьогодні розробка моделей глибокого навчання для медичних зображень є одним із основних напрямків досліджень штучного інтелекту та машинного навчання. За допомогою цих моделей можна класифікувати різні типи зображень, наприклад рентгенівські, КТ та МРТ, а також виявляти патологію та інші зміни, які не завжди помітні людському оку.

Важливим завданням у цій галузі є точне виявлення об'єктів чи класифікація медичних зображень для діагностики захворювань і виявлення патології, що дає змогу лікарям ставити точніші діагнози та підбирати ефективні методи лікування. Крім того, використання моделей глибокого навчання може виявляти невеликі зміни в зображеннях, які можуть вказувати на прогресування захворювання на ранній стадії, що дозволяє швидше та ефективніше лікувати пацієнтів.

Однак, щоб досягти високої точності моделей глибокого навчання для медичної візуалізації, необхідно мати велику кількість високоякісних даних. Важливо також пам'ятати, що ймовірність неправильного діагнозу та його наслідків може бути дуже високою, тому досягнення високої точності є надзвичайно важливим завданням при розробці моделей глибокого навчання для такої задачі. Це дозволяє підвищити та пришвидшити рівень діагностики та лікування пацієнтів, знизити ризик помилкового діагнозу.

Крім того, застосування моделей глибокого навчання до медичної візуалізації може революціонізувати спосіб діагностики та лікування захворювань. За допомогою цих моделей можна аналізувати тисячі зображень, що дає можливість отримати більш точну та об'єктивну

інформацію про стан пацієнта, а також із більшою швидкістю продукувати результати аналізів. Крім того, це може допомогти розробити індивідуальне лікування на основі аналізу унікальних характеристик кожного пацієнта.

Отже, розробка моделей глибокого навчання для класифікації медичних зображень і виявлення патологій є важливим кроком на шляху до покращення діагностики та лікування захворювань. Використання цих моделей може допомогти покращити лікування та підвищити його ефективність, що може мати позитивний вплив на здоров'я та якість життя пацієнтів.

Одним із частих звернень пацієнтів до лікаря є підозра на перелом кістки. Проблема полягає в тому, що виявлення переломів на рентгенівських знімках залишається складним завданням для медичних фахівців і вимагає значних зусиль та часу. Іноді перелом може бути настільки непомітний, що пацієнт може навіть не підозрювати, що він є. Оптимізація процесу виявлення переломів є важливою задачею, оскільки швидка та точна діагностика є критично важливою для пацієнтів, які потребують невідкладної медичної допомоги.

Вирішення цієї проблеми є доволі важливим, оскільки:

- Навіть досвідчені лікарі можуть допускати помилки при виявленні переломів на рентгенівських знімках. Це може призвести до надлишкових обстежень, затримок у лікуванні та інших негативних наслідків для пацієнтів.
- У медичних установах щодня створюються тисячі рентгенівських знімків, і обробка кожного з них вручну вимагає великих зусиль і часу.
- При різних критичних ситуаціях та нещасних випадках, коли пацієнти потребують негайної діагностики, важливо швидко та надійно виявляти переломи на рентгенівських знімках.
- Використання глибокого навчання для виявлення переломів на рентгенівських знімках може посприяти покращенню точності та швидкості діагностики.

Тому, проблема виявлення переломів на рентгенівських знімках є актуальною і вимагає розвитку та впровадження сучасних технологій на основі глибокого навчання для покращення діагностики та надання якісної медичної допомоги пацієнтам.

Метою цієї роботи є розробка рішення із використанням моделей глибокого навчання для виявлення перелому на рентгенівських знімках та їх подальшої класифікації.

Для досягнення поставленої мети будуть виконані наступні завдання:

1. Дослідження існуючих технологій глибокого навчання для виявлення об'єктів та класифікації медичних зображень.
2. Вибір та тренування нейромережевої архітектури для виявлення та класифікації медичних зображень та виявлення на них патологій.
3. Збір та передобробка медичних зображень для створення навчальної та тестової вибірок.
4. Оптимізація (за потреби) вибраної архітектури за допомогою налаштування гіперпараметрів та методів регуляризації.
5. Оцінка точності та швидкості розробленої моделі на тестовій вибірці медичних зображень.

Об'єктом дослідження є дослідження процес класифікації зображень та виявлення об'єктів на медичних зображеннях.

Предметом дослідження є розробка рішення з використанням нейромережевої архітектури та методів глибокого навчання для досягнення найвищої точності та швидкості виявлення об'єктів на медичних зображеннях.

Методами дослідження є методи обробки зображення, модель нейронної мережі, що використовуються для роботи з зображеннями.

Новизною дослідження є саме рішення для виявлення об'єктів на їх класифікації на медичних зображеннях, за допомогою використання нейромережевої архітектури та методів глибокого навчання.

Практична цінність полягає у тому, що при великому потоці пацієнтів важливо швидко надати допомогу кожному, що потребує допоміжного автоматизованого інструменту.

Особистий внесок магістранта полягає в розробці моделі нейронної мережі для допомоги у виявленні об'єктів на медичних зображеннях.

Апробація результатів роботи. На основі проведених наукових досліджень було подано для публікації наукову статтю по темі магістерської роботи на наукову конференцію “Нейромережні технології та їх застосування «НМТІЗ-2023»”, що проходить 13-14 грудня 2023 року.

1. АНАЛІТИЧНИЙ РОЗДІЛ

1.1. Методологія пошуку наукових джерел

У цій роботі проведено аналітичний огляд наукових джерел згідно з стандартизованою методологією PRISMA, що містить загальні рекомендації щодо огляду наукових та базових особливостей мета-аналізу.

В якості одного з наукових джерел я буду використовувати наукометричну базу Scopus. Scopus є одним з найбільших міжнародних наукових баз даних, який охоплює різні наукові галузі, зокрема науку про здоров'я. Він забезпечує доступ до більш ніж 75 мільйонів досліджень, статей, конференційних матеріалів та книг з усього світу, що відображає величезний обсяг наукових даних, які були опубліковані протягом багатьох років. Окрім цього він має доступний та простий інтерфейс із можливістю фільтрувати результати на вищому рівні, а також можливість прописати запит на більш низькому рівні, за допомогою ключових слів, що дає змогу глибше профільтрувати знайдені джерела.

```
( TITLE-ABS-KEY ( "bone" AND ( "fracture" OR "crack" OR "broken" ) AND "detection" AND ( "x-ray" OR "xray" OR "x ray" ) AND "deep learning" ) AND SUBJAREA ( comp OR medi ) AND PUBYEAR > 2020 )
```

Рис.1.1. Початковий запит пошуку релевантних до теми статей по базі Scopus

Запит на рис.1.1 означає, що буде йти пошук по статтях, анотаціях та наукових документах з вільним доступом, у розділі тематики комп'ютерних наук (comp — computer science) та медичної тематики (medi — medical), починаючи з 2020 року, які мають містити слова “bone”, “fracture” або “crack” або “broken”, “detection”, “x-ray” чи “xray” чи “x ray”, та “deep learning” у назві, анотації або в ключових словах. За даним запитом було знайдено 126 документів.

Іншою базою наукових джерел я обрав ресурс PubMed, оскільки тема пов'язана з медициною. За наступним запитом, — “deep learning” + “bone” +

("fracture" або "crack" або "broken") + "detection", 2020-2023, — було знайдено 53 документи (рис.1.2).

("deep learning"[Title/Abstract] AND "bone"[Title/Abstract] AND ("fracture"[Title/Abstract] OR "crack"[Title/Abstract] OR "broken"[Title/Abstract]) AND "detection"[Title/Abstract] AND 2020/01/01:2023/12/31[Date - Publication]) AND ((frrft[Filter]) AND (fha[Filter]))

Рис.1.2. Запит пошуку релевантних до теми статей по базі PubMed

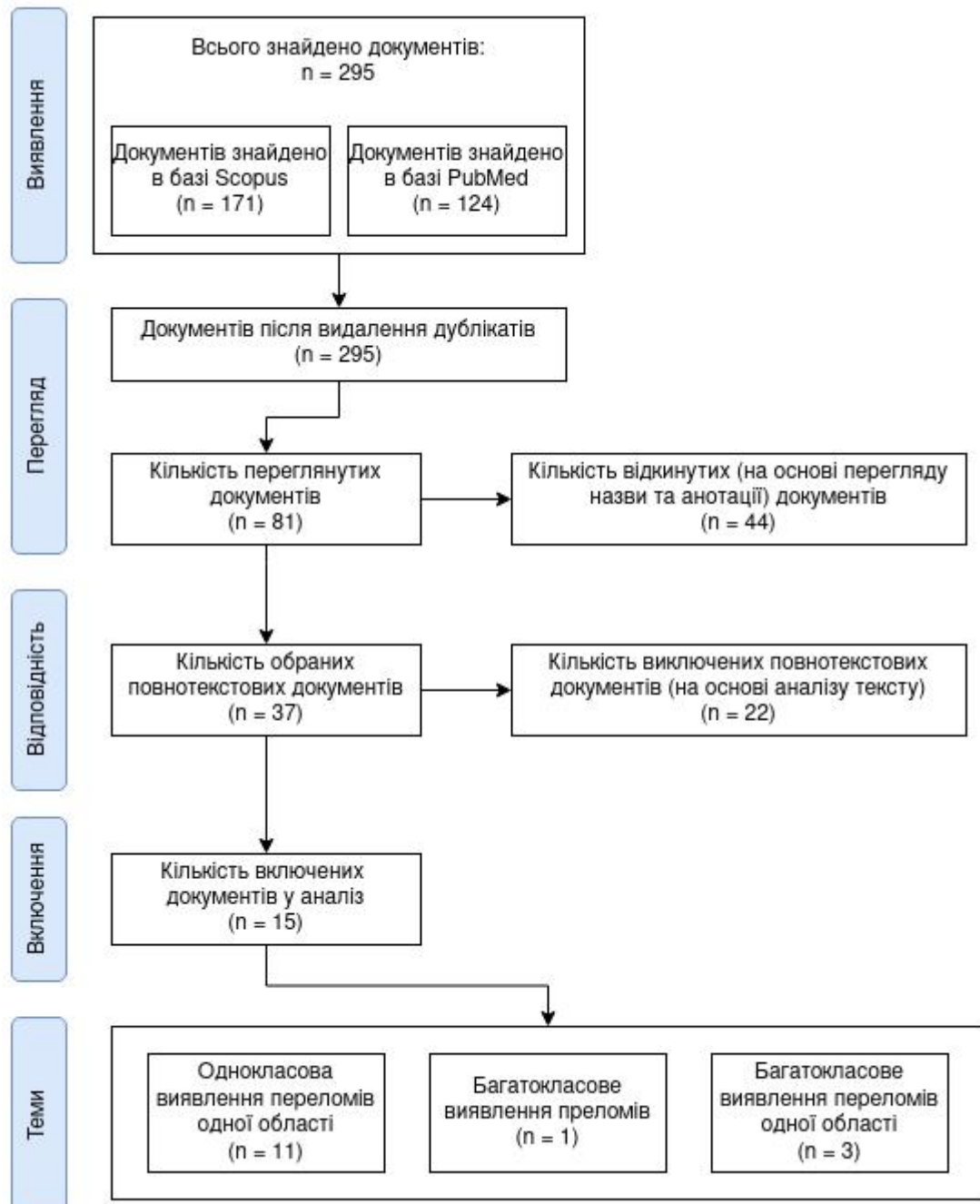


Рис.1.3. Методологія виконання аналітичного огляду

1.2. Огляд літературних джерел

На основі проведеного пошуку літературних джерел було включено N статей у подальший огляд та аналіз. Їх зведено у табл. 1.1.

Таблиця 1.1. Огляд існуючих робіт

№	Назва	Інструментарій	Задача	Рік	Посилання
1	AC-Faster R-CNN: an improved detection architecture with high precision and sensitivity for abnormality in spine x-ray images	AC-Faster R-CNN, Deformable Convolution Feature Pyramid Network, Abnormality Capture Head, Contrast Limited Adaptive Histogram Equalization	Розробка моделі глибокого навчання для точного виявлення та класифікації аномалій на рентгенограмах хребта	2022	[1]
2	Application of convolutional neural networks for distal radio-ulnar fracture detection on plain radiographs in the emergency room	DenseNet-161, ResNet-152, Gradient-weighted Class Activation Mapping	Оцінка ефективності моделей глибокого навчання у виявленні переломів зап'ястя на рентгенограмах	2020	[2]
3	Automated detection of	DCNN, Grad-CAM	Розробка високоефективної	2023	[3]

	scaphoid fractures using deep neural networks in radiographs		моделі глибокого навчання для виявлення переломів ліктьової кістки		
4	Automated fracture detection in the ulna and radius using deep learning on upper extremity radiographs	RegNetX006, EfficientNet B0, InceptionResNet V2	Розробка та оцінка моделі глибокого навчання для виявлення одного або декількох переломів в ліктьовій або променевій кістках на рентгенограмах	2023	[4]
5	Automated universal fractures detection in X-ray images based on deep learning approach	Cascade R-CNN, Mask R-CNN, SSD, NAS-FPN, AUGFPN, Ada-ResNeSt, AC-BiFPN	Розробка універсальної системи комп'ютерної діагностики переломів на рентгенограмах на основі глибокого навчання	2022	[5]
6	Automatic detection and localization of thighbone fractures	ResNeXt, RPN, RoI, R-CNN, PASCAL VOC	Розробка та оцінка нової методики для автоматизованого	2022	[6]

	in X-ray based on improved deep learning method		виявлення переломів стегнової кістки на рентгенограмах за допомогою глибокого навчання		
7	Bone Fracture Detection Using Deep Supervised Learning from Radiological Images: A Paradigm Shift	-	Огляд різних підходів до виявлення та класифікації переломів за допомогою глибокого навчання	2022	[8]
8	Cascading AB-YOLOv5 and PB-YOLOv5 for rib fracture detection in frontal and oblique chest X-ray images	AB-YOLOv5, PB-YOLOv5	Розробка каскадної архітектури моделей глибокого навчання для ідентифікації переломів ребер	2023	[10]

9	Evaluation of ensemble strategy on the development of multiple view ankle fracture detection algorithm	DCNN, ImageNet, AUROC, grad-CAM, Adam optimizer	Дослідження можливості та ефективності використання глибоких згорткових нейронних мереж (DCNNs) для виявлення переломів кісточки ноги на рентгенограмах її різних проекцій	2023	[11]
10	Fracture Detection in Wrist X-ray Images Using Deep Learning-Based Object Detection Models	SABL, RegNet, RetinaNet, PAA, Libra R-CNN, FSAF, Faster R-CNN, Dynamic R-CNN, DCN, ResNet50	Порівняння різних моделей глибокого навчання та розробка ансамблю моделей для вирішення задачі ідентифікації перелому зап'ястя	2022	[13]
11	Hybrid SFNet Model for Bone Fracture Detection and Classification Using ML/DL	LRN, DCNN, Canny, SFNet	Розробка мультимодальної гібридної моделі SFNet для виявлення переломів на рентгенограмах	2022	[15]

12	Lower-extremity fatigue fracture detection and grading based on deep learning models of radiographs	ResNet-50, triplet branch network, AUC	Створення моделі для виявлення переломів нижніх кінцівок з використанням глибокого навчання	2023	[17]
13	Multiclass datasets expand neural network utility: an example on ankle radiographs	Inceptionv3-network, ImageNet, ROC AUC, AUC	Створення моделі для виявлення наявності перелому та оцінки його важкості на рентгенограмах стопи	2023	[19]
14	Transfer learning-based ensemble convolutional neural network for accelerated diagnosis of foot fractures	InceptionResNet V2, MobilenetV1, ResNet152V2, ROC, AUC, F1	Розробка системи штучного інтелекту для удосконалення діагностики переломів стопи на рентгенограмах	2023	[20]
15	Using Artificial Intelligence to Diagnose Osteoporotic Vertebral Fractures on Plain Radiographs	YOLOV5, YOLOV5-XRAY, Encoder, Decoder		2023	[21]

Далі, проведемо детальний аналіз отриманих робіт з табл. 1.1.

1.3. Критичний аналіз літературних джерел

В даному розділі проведемо детальний аналіз кожної з робіт, отриманих у розділі 1.2.

У статті [1] розглядається проблема локалізації та ідентифікації захворювань на рентгенограмах хребта. Метою дослідження є розробка мережі для об'єктного виявлення, яка точно локалізує та класифікує аномальні ділянки на рентгенограмах хребта. Запропонована модель, AC-Faster R-CNN, використовує глибоке навчання та має структуру злиття ознак Deformable Convolution Feature Pyramid Network та структуру захоплення аномалій Abnormality Capture Head. Вона також використовує поєднання розширених і деформованих згорток для кращого збору інформації на різних масштабах уражень. Для покращення точності виявлення використовується алгоритм підвищення контрастності - Contrast Limited Adaptive Histogram Equalization - для попередньої обробки зображень. Результати експериментів на тестовому наборі з 1007 рентгенограм показують, що модель AC-Faster R-CNN перевершує базову модель і інші архітектури для виявлення аномалій. Зокрема, середній середній показник IoU 50% складає 39.8%, а точність та чутливість при оптимальній точці кривої Precision-Recall становлять 48.6% та 46.3%, відповідно, що досягає поточного рівня передових технологій в області виявлення.

Дана робота показує, що модель AC-Faster R-CNN демонструє високу точність та чутливість в завданнях виявлення аномалій на рентгенограмах хребта та ефективно локалізує та ідентифікує аномальні ділянки, а також виявляє приховані переломи. Проте набір даних був недостатньо великим, а також робота була спрямована лише на один вид переломів.

У дослідженні [2] автори оцінювали ефективність двох моделей глибокого навчання, а саме DenseNet-161 і ResNet-152, в детекції переломів зап'ястя. Для цього вони зібрали зображення пацієнтів, які зверталися з травмами зап'ястя в екстрених відділеннях лікарень. Датасет, який був

створений з січня 2018 по травень 2020 року, був розділений на навчальну і тестову вибірки у співвідношенні 90% до 10%. Для навчання використовувались дві моделі глибокого навчання, і для оцінки їхньої ефективності використовувалася ROC-AUC. Для виділення областей на рентгенограмах, які вплинули на рішення моделей, використовувався метод градієнтно-зваженої активаційної карти. Під час навчання моделей використовувалося значну кількість рентгенограм - 4,551 зображень з переломами та 4,443 зображень без переломів. Результати показали, що моделі DenseNet-161 і ResNet-152 мали високі показники чутливості, специфічності та точності в детекції переломів зап'ястя.

Дане дослідження показало, що обидві моделі можуть успішно допомогти виявити переломи зап'ястя з задовільною точністю в екстрених ситуаціях в лікарнях. Проте робота була спрямована лише на одну область переломів, що робить її обмежено ефективною.

У роботі [3] основна мета полягала в розробці високоефективної моделі глибокого навчання для виявлення переломів ліктьової кістки, які можуть бути видимими на рентгенограмах та тих, які є прихованими і не видимими на звичайних рентгенограмах. Робота використовувала 525 зображень рентгенограм зап'ястя: 250 з них представляли нормальний стан ліктьової кістки, 219 - ліктьові кістки з переломами, і 56 - ліктьові кістки з прихованими переломами. Ці зображення були отримані від Департаменту ортопедії Медичного коледжу Кастурби. Для досягнення поставленої мети була розроблена модель глибокого навчання на основі згорткових нейронних мереж (CNN) для розділення зображень на два класи (норма і перелом) та трьох класів (норма, перелом і прихований перелом). Для локалізації перелому використовувався метод градієнтно-зваженої активаційної карти (Grad-CAM). В результаті експериментів були отримані такі результати: для двокласової класифікації модель досягла чутливості 92%, специфічності 88%, точності 90% та площі під кривою характеристики робочої характеристики

(AUC) 0.95. Для трьох класів модель показала загальну чутливість 85%, специфічність 91%, точність 90% і AUC 0.88.

Плюс цієї роботи, що розроблена модель є ефективною у виявленні переломів ліктьової кістки на рентгенограмах, включаючи і приховані переломи, і може бути корисною у клінічних умовах, оскільки не вимагає препроцесингу (сегментації) рентгенограм. Проте використаний датасет був доволі обмеженим у області перелому та й самому об'ємі даних.

У дослідницькій роботі [4] зусилля були спрямовані на виявлення одного або кількох переломів в ліктьовій або променевої кістках за допомогою глибокого навчання, які були навчені на рентгенограмах верхніх кінцівок. В рамках дослідження використовувався набір даних, який складався з різних типів рентгенограм кінцівок, включаючи 4,480 зображень із переломами і 4,383 зображень без переломів, всі переломи стосувались, як попередньо згадувалось, ліктьової або променевої кісток. Запропонований метод мав два основних етапи: препроцесінг, який включав видалення радіографічного фону та прибирання некісткової тканини, та використання глибоких моделей навчання, таких як RegNetX006, EfficientNet B0 та InceptionResNetV2, для обробки зображень, які містять тільки кісткову тканину, з метою виявлення одного або декількох переломів ліктьової або променевої кісток. Ефективність запропонованого методу була оцінена за допомогою різних метрик, таких як точність, F1-показник, коефіцієнт кореляції Меттью, ROC-AUC, чутливість, специфічність та точність, використовуючи метод хрестової перевірки (10-fold cross-validation). Результати показали, що найкращу класифікаційну ефективність було досягнуто за допомогою запропонованого передпроцесінгу та архітектури RegNetX006, з такими показниками, як точність (0.9921), F1-показник (0.9918), коефіцієнт кореляції Меттью (0.9842), AUC (0.9918), чутливість (0.9974), специфічність (0.9863) та точність (0.9923).

Дана стаття показала, що запропонований метод препроцесінгу здатний виявляти переломи ліктьової та променевої кісток за допомогою штучного

інтелекту. Проте це є і мінусом, оскільки не завжди є час чи можливість зробити препроцесинг рентгенограм.

У статті [5] було розроблено систему комп'ютерної діагностики переломів на рентгенограмах на основі глибокого навчання з метою підвищення точності та універсальності виявлення переломів у різних анатомічних областях. Раніше існуючі системи зосереджувалися на виявленні переломів в певних частинах тіла. У роботі використано попередню обробку зображень для поліпшення якості рентгенограм, а також методи збільшення обсягу даних. Була розроблена система для автоматичного виявлення переломів на рентгенограмах з використанням модифікованої нейромережі Ada-ResNeSt та методу виявлення AC-BiFPN. Для оцінки результатів було створено власний набір даних для універсального виявлення переломів на основі відкритого набору даних MURA. За результатами експериментів, система показала високу точність виявлення переломів (68.4% за метрикою AP) при прийнятній швидкості обробки зображень (122 мс на зображення) на тестовому наборі MURA-D, і це, за словами авторів, є обіцяючими результатами порівняно з іншими сучасними системами виявлення.

Дана робота має схожу ціль з цією, та може вважатись аналогом чи конкурентом. Що стосується точності, то дана робота має менше 70%, що є непоганим показником, але й не достатнім.

Наступна робота [6] розглядає застосування глибокого навчання для виявлення переломів стегнової кістки на рентгенограмах. Автори пропонують нову двоетапну структуру мережі на основі згорткових нейронних мереж для виявлення таких переломів. В новій мережі враховано баланс інформації на кожному рівні функційної піраміди ResNeXt. У дослідженні використовувався набір даних з 3842 рентгенограмами стегнової кістки. Для порівняння запропонованої методики було використано інші сучасні методи виявлення об'єктів за допомогою наперед навчених моделей та transfer learning на тому ж наборі даних. Крім того, було проведено кілька експериментів для визначення впливу окремих компонентів запропонованої

системи та різних налаштувань гіперпараметрів на виявлення переломів. Результати експериментів показали, що середня точність запропонованої системи досягла 88.9% в виявленні переломів стегнової кістки.

Результати даного дослідження підтверджують ефективність та перевагу запропонованої методики порівняно з іншими сучасними методами виявлення. Проте як недолік можна виділити вузьку специфікацію моделі, а саме переломів стегнової кістки.

У статті [8] були розглянуті різні підходи до виявлення та класифікації переломів. Діагностика та прогнозування рентгенограм шляхом експертів-радіологів є витратною за робочим часом та працею процедурою, яку можна автоматизувати за допомогою технік виявлення переломів. Згідно з висновками багатьох дослідників, основною перешкодою для створення високоефективного алгоритму класифікації є недостатність розмічених даних для навчання моделей. Поки не існує стандартної моделі, яку можна використовувати з наявними даними. Автори намагалися продемонструвати використання глибокого навчання (DL) в медичних зображеннях та його потенціал для точної діагностики хвороби. Однак існує сумнів у повній інтеграції DL у рентгенологію, оскільки є страх серед радіологів втратити роботу. Замість цього, основна мета полягає в допомозі радіологам у їх роботі, а не в їх заміщенні. Висновок полягає в тому, що моделі на основі згорткових нейронних мереж (CNN), зокрема InceptionNet і XceptionNet, продемонстрували високу ефективність у виявленні переломів на рентгенограмах.

Дана стаття хорошим оглядом існуючих технологій, проте не містить конкретних заходів, щодо розробки моделі.

У роботі [10] досліджують використання глибокого навчання для виявлення та класифікації переломів ребер на рентгенограмах грудної клітки. На відміну від інших аномалій грудної клітки, дослідження щодо переломів ребер є обмеженим. Більшість існуючих глибоких моделей зосереджені на використанні фронтальних рентгенограм грудної клітки (CXR). Дослідники

використали набір даних EDARib-CXR, який містить 369 фронтальних та 829 косих рентгенограм грудної клітки. У роботі автори представляють дві моделі для виявлення переломів, AB-YOLOv5 і PB-YOLOv5, які були навчені та оцінені на наборі даних EDARib-CXR. AB-YOLOv5 — це модифікована мережа YOLOv5, яка включає додаткову гілку для покращення роздільної здатності ознак на останньому рівні згорткової мережі. PB-YOLOv5 зберігає ту ж структуру, що і в оригінальній YOLOv5, але використовує фрагменти зображень під час навчання для збереження ознак малих об'єктів на зменшених зображеннях. Крім того, автори пропонують нову дворівневу каскадну архітектуру, яка поєднує обидві моделі виявлення (AB-YOLOv5 і PB-YOLOv5). Ця структура демонструє покращені метрики на тестовому наборі даних, досягаючи значення AP30 на рівні 0.785.

Дане дослідження показало розробку нової дворівневої архітектури моделей глибокого навчання, які здатні ідентифікувати та локалізувати переломи ребер на фронтальних та косих рентгенограмах грудної клітки. Попри це робота була зосереджена лише на одному класі переломів.

У статті [11] розглядається можливість та ефективність використання глибоких згорткових нейронних мереж (DCNNs) для виявлення переломів кісточки ноги на рентгенограмах її різних проекцій. Рентгенограми кісточки є основним інструментом для діагностики переломів цієї області, і застосування алгоритмів DCNN на цих зображеннях може покращити точність та ефективність діагностики. Дослідники натренували DCNN за допомогою набору зображень, який включав 3102 рентгенограми лоджки. Окремо були навчені моделі DCNN для антеропостеріорних (AP) та латеральних (Lat) проекцій рентгенограм. Різні методи ансамблю, такі як "sum-up," "severance-OR," та "severance-Both," були використані для об'єднання результатів моделей, які використовують різні проекції. Результати показали, що окремо моделі AP і Lat мали високі показники чутливості, специфічності та точності, а саме 79%/84%, 90%/86%, 88%/86%, 83%/85%, та 0.816/0.850, відповідно. Використання різних методів ансамблю

дозволило оптимізувати результати в залежності від клінічної ситуації: “sum-up” метод згенерував результати використавши дві моделі та отримав оцінку AUROC в 0.917 (95% CI: 0.863–0.972) із 87% точності, “severance-OR” мав результати кращої чутливості в 90%, і “severance-Both” метод показав вищий показник специфічності, 94%.

Дане дослідження підтвердило, що глибокі нейронні мережі можуть бути ефективно використані для виявлення переломів кісточки ноги, а вибір методу ансамблю може залежати від клінічних потреб. З недоліків даного дослідження, можна виділити вузьку специфіку зображень переломів.

У роботі [13] дослідники вивчали використання глибокого навчання для виявлення переломів зап'ястя на рентгенограмах з метою підтримки лікарів, особливо в екстрених медичних службах. Для досягнення цієї мети, вони використовували різні моделі глибокого навчання, такі як SABL, RegNet, RetinaNet, PAA, Libra R-CNN, FSAF, Faster R-CNN, Dynamic R-CNN і DCN, з різними структурами для виявлення переломів на рентгенограмах. Всього було розроблено 20 різних процедур для виявлення переломів. Для покращення результатів, було розроблено п'ять різних ансамблевих моделей, які були подалі використані для створення унікальної моделі виявлення переломів, яка отримала найкращий результат серед 26 різних моделей з середнім показником точності (AP50) на рівні 0,8639.

У дослідженні [15] автори розробили новий метод для діагностики переломів кісток на рентгенограмах, який використовує глибоке навчання та вдосконалений алгоритм Canny для виявлення переломів і здорових ділянок кістки. Запропонована модель, відома як Hybrid Scale Fracture Network (SFNet), є двомасштабною послідовною глибокою моделлю. Ця модель виявляє переломи кісток ефективно та вимагає менше обчислювального часу порівняно з іншими сучасними моделями глибокого навчання. Інновація цього дослідження полягає в тому, що воно використовує вдосконалений алгоритм Canny для виявлення контурів на рентгенограмах, що локалізують область перелому. Далі сірі зображення разом з відповідними зображеннями

країв, отриманими за допомогою алгоритму Canny, подаються на вхід до запропонованої глибокої мережі SFNet для навчання та оцінки. Результати дослідження показали, що модель SFNet з використанням алгоритму Canny досягла найвищої точності, F1-показника та чутливості на рівні 99,12%, 99% та 100% в діагностиці переломів кісток. Це свідчить про те, що використання алгоритму Canny покращує продуктивність глибокого навчання в цій задачі.

У статті [17] дослідники розглядали можливість використання глибокого навчання для діагностики та оцінки важкості переломів нижніх кінцівок на звичайних рентгенограмах. Для цього було використано 1151 зображення (682 для tibiofibula, нижня пряма кістка ноги, та 469 для стопи) з переломами та 2842 зображення (2000 для tibiofibula та 842 для стопи) без аномалій з двох клінічних центрів. Після позначення уражень, зображення з одного центру були розподілені у співвідношенні 7:1:2 для побудови моделей, а решта зображень з іншого центру використовувалися для зовнішньої перевірки. Для розробки діагностичних моделей для виявлення та оцінки були використані мережа ResNet-50 та триплет-гілка. Впродовж внутрішнього тестування та зовнішньої перевірки моделей були оцінені чутливість, специфічність та площа під кривою характеристики робочої характеристики (AUC) для моделей виявлення та точність за матрицею нерозуміння для моделей оцінки важкості перелому. Результати показали, що моделі виявлення мають високу чутливість та добру специфічність для переломів тибіофібули та стопи та перевершують візуальну оцінку навіть від старших рентгенологів. Загальна точність діагностичної моделі склала 78.5%/62.9% для tibiofibula та 74.7%/61.1% для стопи на внутрішньому тестувальному/зовнішньому валідаційному наборах.

Моделі для виявлення та оцінки показали обіцяні результати проте можуть бути використані лише для допомоги в діагностиці переломів нижніх кінцівок.

У роботі [19] вивчалася можливість використання штучного інтелекту в комп'ютерному баченні для покращення діагностики на рентгенограмах

стопи. Методика дослідження включала багатовимірну класифікацію рентгенограм стопи, кількість яких становила 1493 зображення, оцінюючи різні характеристики, включаючи впевненість у наявності перелому. На основі цих зображень була навчена нейронна мережа для виявлення переломів за допомогою спеціального протоколу підготовки даних та навчання, а саме залучення двох справжніх спеціалістів для розмітки зображень та поділ на класи. В якості моделі була використана Inceptionv3-network пре-тренувана на ImageNet. Результати показали, що моделі, навчені на цьому датасеті, перевершили більшість подібних моделей з літератури, досягнувши значення площі під кривою характеристики робочої характеристики (ROC AUC) на рівні 0,943. Виключення рентгенограм стопи з ознаками хірургічного втручання покращило класифікацію перелому (AUC 0,955), в той час як обмеження тренувального набору даних лише здоровими стопами з та без перелому не мало системного впливу на результати. Наприкінці, дослідники виявили, що наявність ознак хірургічного втручання може впливати на модель, і вилучення таких зображень покращило результати. Однак вилучення інших патологій, окрім перелому, не вплинуло на ефективність моделі, що свідчить про корисність врахування різноманітності ознак для навчання надійних моделей. Таким чином, використання багатокласових даних дозволяє оцінювати різні характеристики зображень і поглиблює наше розуміння образів патологій.

Дане дослідження показує, що при класифікації перелому варто враховувати багат різних ознак, а не лише обмежений набір.

У дослідженні [20] розглядалася проблема діагностики переломів стопи, яка може бути ускладнена через складну структуру стопи з численними кістками, зв'язками, сухожиллями та м'язами. Дослідники розробили систему, яка допомагає удосконалювати діагностику переломів стопи та порівнювали ефективність III на різних групах з різним рівнем кваліфікації. Для поліпшення видимості рентгенограм використовувалася техніка контрастного адаптивного гістограмного вирівнювання. Дані також

піддавалися аугментації для запобігання перенавчанню. Оброблені рентгенограми подавалися на вхід у ансамбль моделі, яка базується на згорткових нейронних мережах. Використовувалися три моделі: InceptionResNetV2, MobilenetV1 і ResNet152V2. Для оцінки результатів використовувалися різні метрики, включаючи криву характеристики робочої характеристики (ROC), площу під кривою (AUC) та F1. Результати показали, що ансамбль моделі має кращу здатність до класифікації (F1: 0.837, AUC: 0.95, Точність: 86.1%) порівняно з окремими моделями, які показали точність на рівні 82.4%.

Дане дослідження демонструє використання ансамлю моделей, для ідентифікації перелому стопи. Проте використання ансамлю може збільшувати час роботи системи, що може бути недоліком.

У праці [21] було розроблено та продемонстровано систему діагностики переломів хребця, використовуючи глибоке навчання, що допомагає мережі виконувати виявлення та сегментацію скелетних структур, а також ідентифікацію та градацію переломів хребця. В якості навчальної та внутрішньої валідаційної вибірки було використано 11 397 рентгенограм з шести медичних центрів у Шанхаї. Для зовнішньої валідації було залучено 1276 учасників зі служби невідкладної допомоги Шанхайської шостої лікарні (1276 рентгенограм). Радіологи провели оцінку всіх рентгенограм та використали метод Генанта для діагностики та градації переломів хребця як основу для порівняння результатів. В якості моделі була використана система на базі YOLOv5, а саме YOLOv5-XRAY, енкодер для витягування ознак, та два декодера для ідентифікації та сегментації. Система AI_OVF_SH продемонструвала високу точність та швидкість обчислень у виявленні та сегментації скелетних структур. У внутрішній валідаційній вибірці модель AI_OVF_SH показала точність 97,41%, чутливість 84,08% і специфічність 97,25% для всіх переломів. Чутливість і специфічність для помірних переломів становили відповідно 88,55% і 99,74%, а для серйозних переломів - 92,30% і 99,92%. У зовнішній валідаційній вибірці точність, чутливість і

специфічність для всіх переломів становили відповідно 96,85%, 83,35% і 94,70%. Для помірних переломів чутливість та специфічність складала відповідно 85,61% і 99,85%, а для серйозних переломів - 93,46% і 99,92%.

1.4. Висновки до розділу 1

В даному розділі був проведений аналіз літературних джерел за темою детекції та класифікації переломів кісток. Було використано дві бібліометричні бази — Scopus та PubMed. Використовуючи розширені механізми запитів було знайдено близько трьохсот релевантних статей, з яких 15 було обрано для подальшого глибшого аналізу.

Більшість робіт включали в себе розробку моделей, чи ансамблів моделей, глибокого навчання для вирішення задачі однокласового (перелом однієї області) виявлення переломів. Декілька робіт досліджували мультикласове виявлення переломів в межах однієї області. Одна робота представляла універсальну систему виявлення переломів в різних областях.

З інструментарію використаного у працях можна виділити ResNet, та різні її варіації та модифікації, Faster R-CNN та сімейство YOLO. Для оцінки здебільшого використовувались метрики AUC, ROC-AUC та F1. Також часто зустрічались моделі пре-треновані на ImageNet. Для обробки зображень використовувались аугментації із зміненням контрастності та яскравості, в окремому випадку був використаний метод Canny для виявлення контурів на зображенні.

Підсумовуючи всі проаналізовані роботи, можна сказати, що більшість робіт зосереджуються на досягненні максимального результату навчаючи модель на наборі зображень якоїсь однієї області тіла (стопа, ребра, зап'ястя, тощо), і лише одна робота була присвячена розробки моделі для виявлення переломів у різних областях тіла. Саме цій задачі є присвячена дана робота.

2. ДОСЛІДНИЦЬКИЙ РОЗДІЛ

2.1. Вибір моделі нейронної мережі

Аналізуючи літературні джерела, описані у п.1.2, мою увагу привернула абревіатура YOLO. Доволі велика кількість розробників обирають цю архітектуру нейронної мережі для вирішення задачі виявлення та класифікації об'єктів на зображеннях чи відео, а отже є сенс розглядати даний алгоритм як засіб реалізації поставленої задачі.

YOLO, або “You Only Look Once”, — це популярна модель для виявлення об'єктів, відома своєю швидкістю та точністю. Ця модель є одноразовим детектором, який використовує повністю згортову нейронну мережу (CNN) для обробки зображення.

Дана модель може бути використана для багатьох задач, таких як: класифікація, виявлення, сегментація, відслідковування чи навіть виявлення пози.

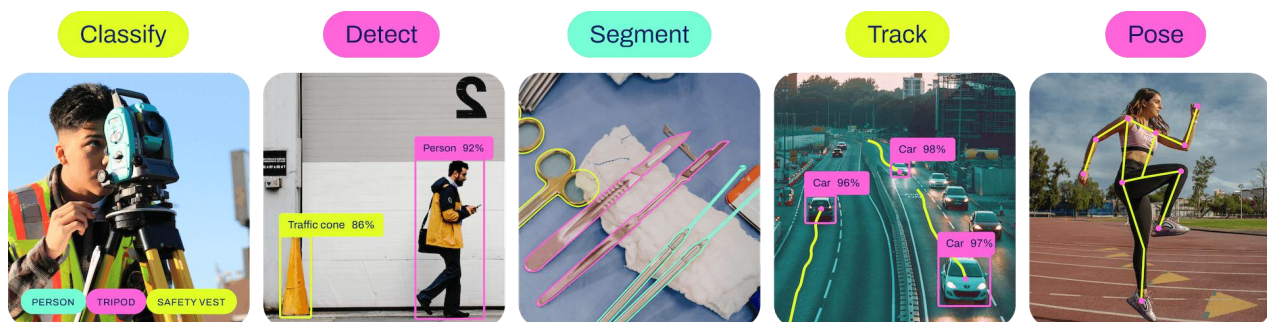


Рис.2.1.Набір задач на які розраховані моделі YOLO

В межах даної роботи буде використана комбінація із можливих використаннях моделі — виявлення та класифікація об'єктів.

Одноразове виявлення об'єктів використовує один прохід вхідного зображення для прогнозування наявності та розташування об'єктів на ньому. Обробка зображення за один прохід робить такі моделі обчислювально ефективними.

Однак одноразове виявлення об'єктів зазвичай менш точне, ніж інші методи, і менш ефективно для виявлення невеликих об'єктів. Такі

алгоритми можна використовувати для виявлення об'єктів у реальному часі в середовищах з обмеженими ресурсами.

Дворазове виявлення об'єктів використовує два проходи вхідного зображення, щоб зробити передбачення щодо присутності та розташування об'єктів. Перший прохід використовується щоб створити набір пропозицій або потенційних координатних місць розташування об'єктів, а другий прохід використовується для перевірки та уточнення цих пропозицій і створення остаточних прогнозів. Цей підхід більш точний, ніж одноразове виявлення об'єкта, але також дорожчий з точки зору обчислень.

Модель YOLO пропонує використовувати наскрізну (end-to-end) нейронну мережу, яка робить передбачення рамок об'єктів (bounding boxes) і ймовірностей класів одночасно. Дотримуючись такого підходу до виявлення об'єктів, YOLO досягла дуже високих результатів, значно перевершивши інші алгоритми виявлення об'єктів у реальному часі.

У той час як такі алгоритми, як Faster RCNN, працюють шляхом виявлення можливих цікавих регіонів за допомогою Region Proposal Network, а потім виконують розпізнавання цих регіонів окремо, YOLO виконує всі свої передбачення за допомогою єдиного повністю підключеного шару. Методи, які використовують мережі Region Proposal Network, виконують кілька ітерацій для одного зображення, тоді як YOLO обходиться лише однією ітерацією.

З моменту першого випуску YOLO, що відбулось у 2015 році, було розроблено чимало нових версій цієї моделі, кожна з яких базується на попередній та вдосконалює батьківську версію моделі. Остання її варіація була випущена у 2023 році із номером версії 8. На рис.2.2 можна оглянути хронологію, яка демонструє розвиток сімейства алгоритмів YOLO за останні роки.

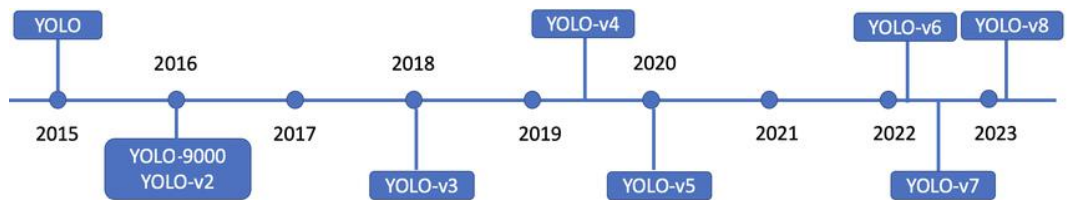


Рис.2.2. Хронологія розвитку YOLO

Окрім цього, з рис.2.2 можна побачити, що дана модель не просто популярна й користується успіхом, а й має великий потенціал для покращення, про що говорить частота оновлення версій. Саме тому для даної роботи була вибрана дане сімейство алгоритмів.

2.2. Аналіз архітектури вибраної моделі

Для того щоб почати ефективно реалізовувати поставлену задачу за допомогою вибраної моделі, потрібно спочатку проаналізувати, як вона працює, аби краще зрозуміти її принцип роботи. Ця частина важлива тим, що допоможе краще зрозуміти як підібрати гіперпараметри.

Отже, архітектура YOLO складається із 3 основних частин:

- Backbone (хребет) - згорткова нейронна мережа, яка об'єднує та формує характеристики зображення з різною деталізацією.
- Neck (шия) - серія шарів для змішування та комбінування функцій зображення, щоб передати їх у передбачення.
- Head (голова) - споживає функції з шиї та виконує кроки передбачення рамок та класу.

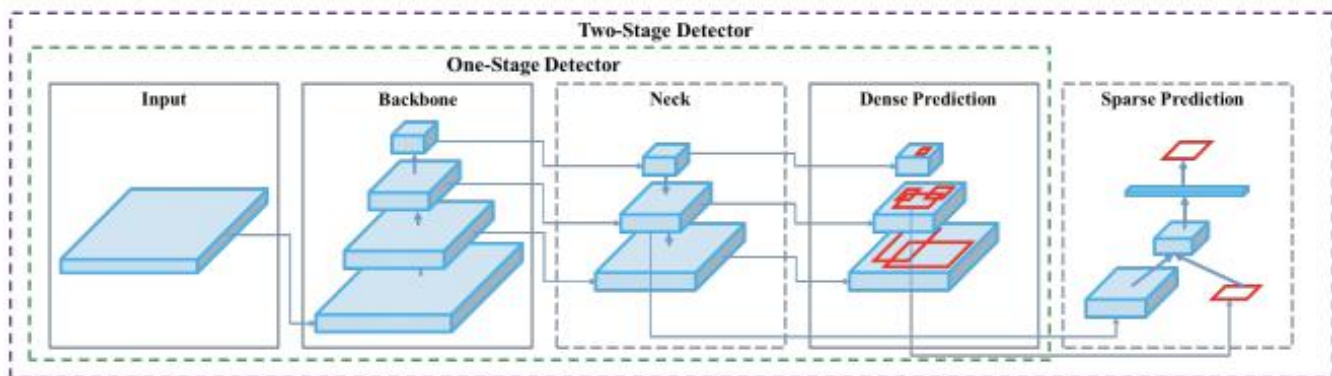


Рис.2.3. Архітектура YOLO

Алгоритм приймає зображення на вхід а потім використовує просту глибоку згорткову нейронну мережу для виявлення об'єктів на зображенні. Нижче представлена архітектура згорткової нейронної мережі, яка формує хребтову частину YOLO.

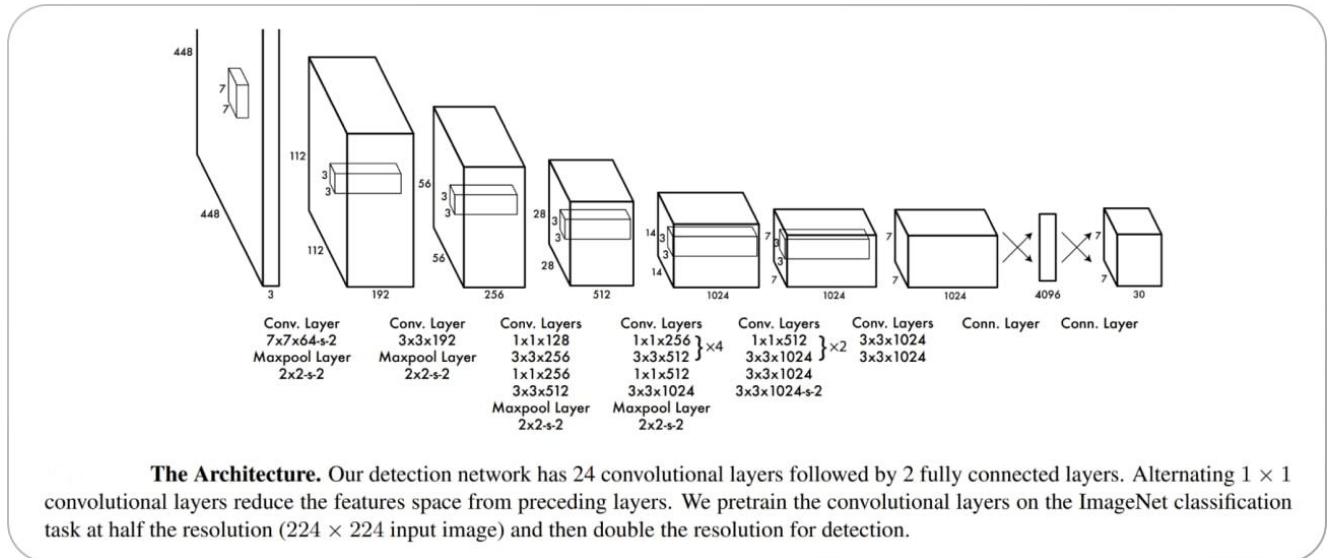


Рис.2.4. Згорткова нейронна мережа хребтової частини YOLO

Перші 20 шарів згортки моделі є попередньо навчені за допомогою ImageNet шляхом підключення тимчасового Avg Pooling шару та Fully Connected шару. Потім ця попередньо навчена модель перетворюється для виконання виявлення об'єктів, оскільки попередні дослідження показали, що додавання згортки (convolution) та підключених (connected) шарів до попередньо навченої мережі підвищує продуктивність. Останній повністю пов'язаний шар YOLO передбачає як імовірності класу, так і координати рамки об'єкту.

Алгоритм YOLO ділить вхідне зображення на квадратну сітку $N \times N$. Якщо центр об'єкта потрапляє в клітинку сітки, ця клітинка відповідає за виявлення цього об'єкта. Кожна клітинка сітки передбачає прямокутники-рамки, там де вважає що знаходить об'єкт, і оцінки достовірності для цих прямокутників. Ці оцінки достовірності відображають, наскільки модель впевнена в тому, що рамка містить об'єкт, і наскільки точно передбачена сама рамка.

YOLO може передбачати кілька рамок на одну клітинку сітки, що призводить до того, що в нас на зображенні з'являється декілька передбачень на один об'єкт. Під час навчання ми звісно ж хочемо, щоб за кожен об'єкт відповідала лише одна рамка (передбачення, яке найбільш точно обрамлює об'єкт та має найбільшу оцінку достовірності). YOLO призначає один предиктор, який буде «відповідальним» за прогнозування об'єкта на основі прогнозу, який має найвищий поточний IOU. IOU, або Intersection Over Union, — це метрика оцінки точності виявлення об'єкту, яка обраховує перекриття передбаченого регіону об'єкта із справжнім (вручну додана рамка, анотація для навчання). Чим більше перекриття, тим краща оцінка.

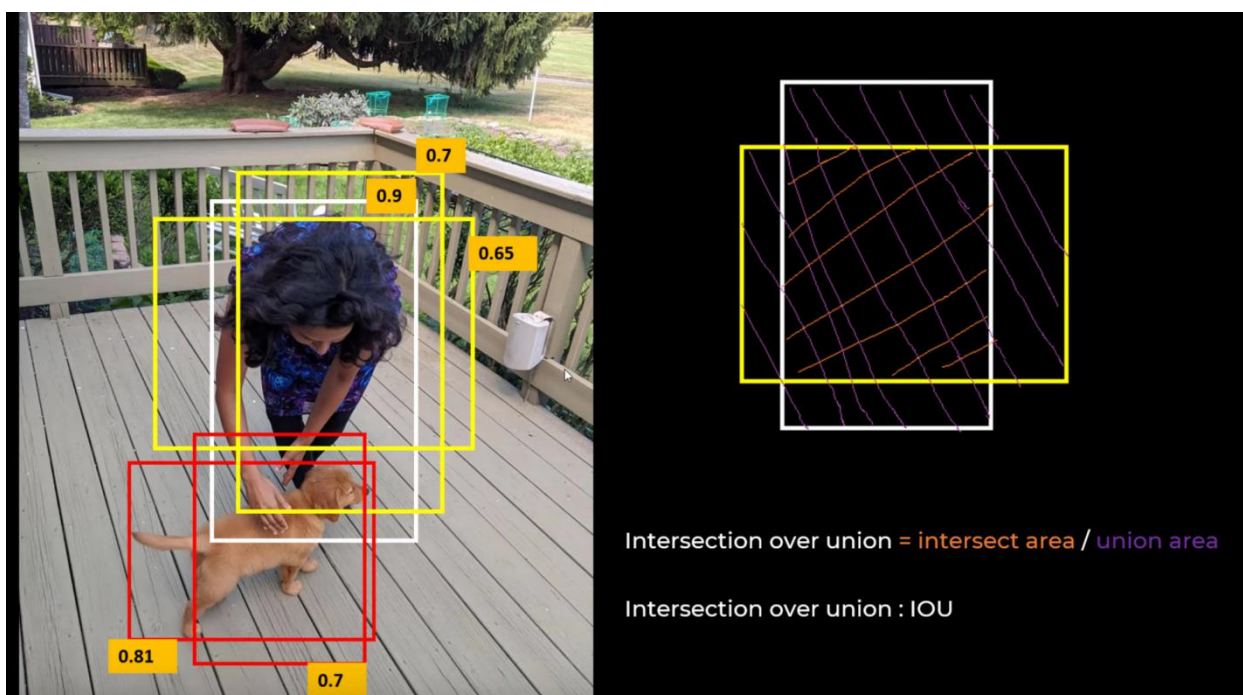


Рис.2.5. Візуалізація метрики IOU

Така робота алгоритму призводить до спеціалізації між предикторами рамки. Кожен предиктор стає кращим у прогнозуванні певних розмірів, пропорцій або класів об'єктів, покращуючи загальну оцінку запам'ятовування.

Одним із ключових методів, що використовується в моделях YOLO, є немаксимальне придушення (non-maximum suppression, NMS). NMS — це етап постобробки, який використовується для підвищення точності та

ефективності виявлення об'єктів. Під час виявлення об'єктів зазвичай для одного об'єкта на зображенні генеруються кілька рамок (див. рис.2.5). Ці обмежувальні рамки можуть перекриватися або розташовуватися в різних положеннях, але всі вони представляють той самий об'єкт. NMS використовується для ідентифікації та видалення зайвих або неправильних обмежувальних рамок і для виведення єдиної обмежувальної рамки для кожного об'єкта на зображенні.

2.3. YOLOv8

YOLOv8 — це на даний момент найновіша версія зі всього сімейства, випущена 10 січня 2023 року. Як раніше згадувалось, кожна версія моделі покращує та вирішує проблеми попередньої, а отже можна припустити, що восьма версія є найкращою. Але чи дійсно це так?

Перші 4 версії даної моделі були написані на C, та підтримувались через глибоку оболонку фреймворку Darknet. Після чого компанія Ultralytics переписала третю версію моделі на PyTorch, відтестували та випустили 5 версію.

YOLOv5 швидко стала популярною завдяки своїй гнучкій структурі Pythonic. Ця структура дозволила спільноті винаходити нові вдосконалення моделювання та швидко ділитися ними в репозиторії за допомогою подібних методів PyTorch.

За останні два роки від репозиторію YOLOv5 PyTorch відгалужуються різні моделі, зокрема Scaled-YOLOv4, YOLOR і YOLOv7. Інші моделі з'явилися по всьому світу на основі власних реалізацій на основі PyTorch, наприклад YOLOX і YOLOv6. Паралельно кожна модель YOLO принесла нові методи, які продовжують підвищувати точність і ефективність моделі.

Наприклад починаючи із 7 версії, модель отримала дві нові техніки тренування. Перша з них називається “мульти-масштабна стратегія” і її суть полягає в тому, що модель навчається на зображеннях різного масштабу, а потім комбінує передбачення. Це допомагає покращити виявлення об'єктів різних розмірів і форм. Друга техніка має назву “Focal Loss”, яка має на меті

вирішення проблеми незбалансованості класів, що часто виникає в задачах виявлення об'єктів. Дана техніка надає більше ваги для передбачень складних випадків, і зменшує вагомість передбачень простих.

Із випуском восьмої моделі сімейства YOLO інтерфейс використання став ще зручнішим, точність моделі покращилась (про це згодом) та збільшилось число ком'юніті навколо цілого сімейства YOLO. Модель тепер має особистий модуль, який можна вставити через PIP (менеджер модулів Python), також різні задачі розбиті по різних файлах, що спрощує роботу з моделлю.

Порівняно з попередніми моделями дана версія має дві основні зміни:

- Anchor-Free Detection
- Мозаїчна аугментація

Anchor boxes, або рамки об'єкта, — відігравали важливу роль в навчанні моделей комп'ютерного зору. Вони чітко вказували де знаходиться об'єкт і де його центр. Також вони вирішували проблему, коли центри двох об'єктів попадали в одну точку. У восьмій моделі розробники відійшли від цього методу, і це сталося з декількох причин: 1) відсутність генералізації (тренування моделі на заданих рамках робить модель жорсткою та нееластичною, після чого в неї важко запихати нові дані), 2) відсутність правильних рамок при нерівностях об'єктів (нерівності неможливо відобразити належним чином за допомогою багатокутних блоків).

Під час тренування YOLOv8 застосовує багато аугментацій до зображень, і одна з нових — мозаїчна аугментація. Цей тип аугментації “склеює” декілька зображень разом, що вчить модель бачити об'єкти в інших позиціях чи часткових появах. Такий процес доволі затратний по продуктивності, тому він включається лише для останніх 10 епох.

По показних ефективності восьма модель показує кращі результати серед своїх попередників. YOLOv8 має більше параметрів, ніж його попередники, такі як YOLOv5, але менше параметрів, ніж YOLOv6. Він пропонує приблизно на 33% більше mAP (mean average precision) для

моделей розміру n і загалом більший mAP у всіх напрямках (рис.2.6). А на рис.2.7 ми можемо спостерігати швидший час передбачення серед усіх інших моделей YOLO. Слід зауважити, що mAP визначався на датасеті COCO, що містить 330 тисяч зображень з 80 категоріями. Він часто використовується як стандарт оцінки моделей.

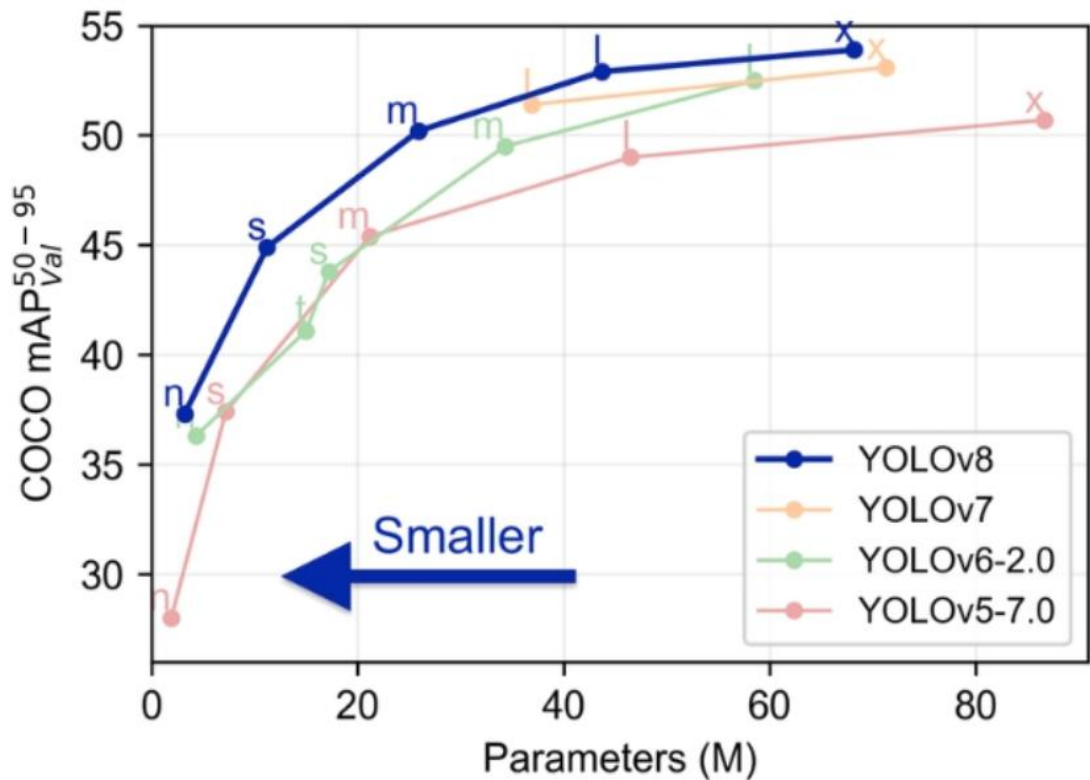


Рис.2.6. Кількість параметрів моделей відносно точності

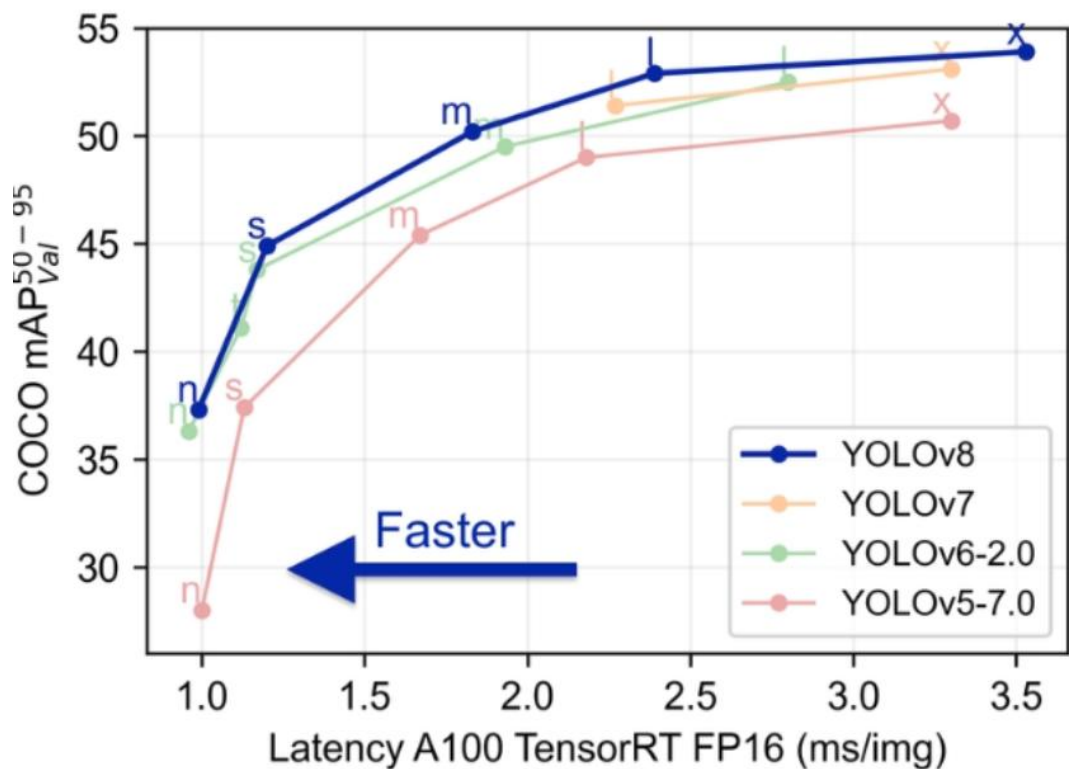


Рис.2.7. Швидкість моделей відносно точності

2.4. Вибір на аналіз наборів даних

Модель без якісних даних ніколи не зможе досягнути високої точності, тому вибір датасету є наступним важливим кроком. Для цієї роботи були проаналізовані багато різних датасетів, та відібрано декілька з них для глибшого аналізу та, можливої, майбутньої комбінації цих наборів разом.

Набір “FracAtlas” [12] містить 4083 зображення із різними випадками переломів у різних місцях тіла. Великою перевагою даного датасету є те, що він вже містить теги (анотації, або англійською “labels”) до усіх зображень, а отже для нього не потрібно підготовки, що збереже багато часу. Також хороша якість зображень є безумовною перевагою для поставленої задачі. Проте, після аналізу виявилось, що з чотирьох тисяч зображень, десь трохи більше трьох тисяч є зображеннями без переломів. Тому даний датасет годиться лише малою часткою своїх зображень. Більше того, він має лише один клас — “перелом”, решта

фотографій без переломів не мають класифікації, що в термінах моделі YOLO вважається як “фон”.

Іншим набором даних є “bone fracture detection using x-rays” [9] з публічної бази датасетів Kaggle, який містить 9467 зображень. На відміну від попереднього набору, цей не є протегований, проте зразу поділений на тренувальну та тестову вибірки. Відсутність анотацій є недоліком, оскільки ручне тегування зображень займає чимало часу (що буде продемонстровано в наступному пункті). Перевагою цього набору є те, що в тренувальній вибірці застосовані трансформації зображень, як от поворот на певну кількість градусів, що дасть змогу моделі вивчити переломи з різних кутів. Ще один недолік — наявність лише двох класів, “перелом” та “без перелому”, що фактично відповідає попередньому проаналізованому набору даних.

Наступний набір даних “fracture types” [14] містить 1250 зображень різних ділянок тіла. Класи поділяються на чотири види: A, B, C та normal. Клас A означає простий перелом (одна точка перелому), B означає перелом середньої складності (дві точки перелому), C означає складний перелом (три та більше точки перелому, розробленість кістки) і normal означає кістку без перелому. Перевагою є наявність чотирьох класів, середній об’єм, поділ на тренувальний, тестовий та валідаційний сеті, а також він є протегованим. З недоліків можна вважати наявність класу normal, який є переважаючим у вибірці та робить цей набір даних незбалансованим. Проте, застосувавши певні методи, цей клас можна легко прибрати і залишити зображення, де є лише цей клас, як “фон”.

Ще один знайдений набір “long bones x rays” [16] містить 1175, які поділені на 9 класів, схожих до попереднього датасету, — A1, A2, A3, B1, B2, B3, C1, C2, C3. Перевагами є схожість із датасетом [14] та відсутність зображень без переломів. З недоліків є те, що нема анотацій (тегів) до зображень.

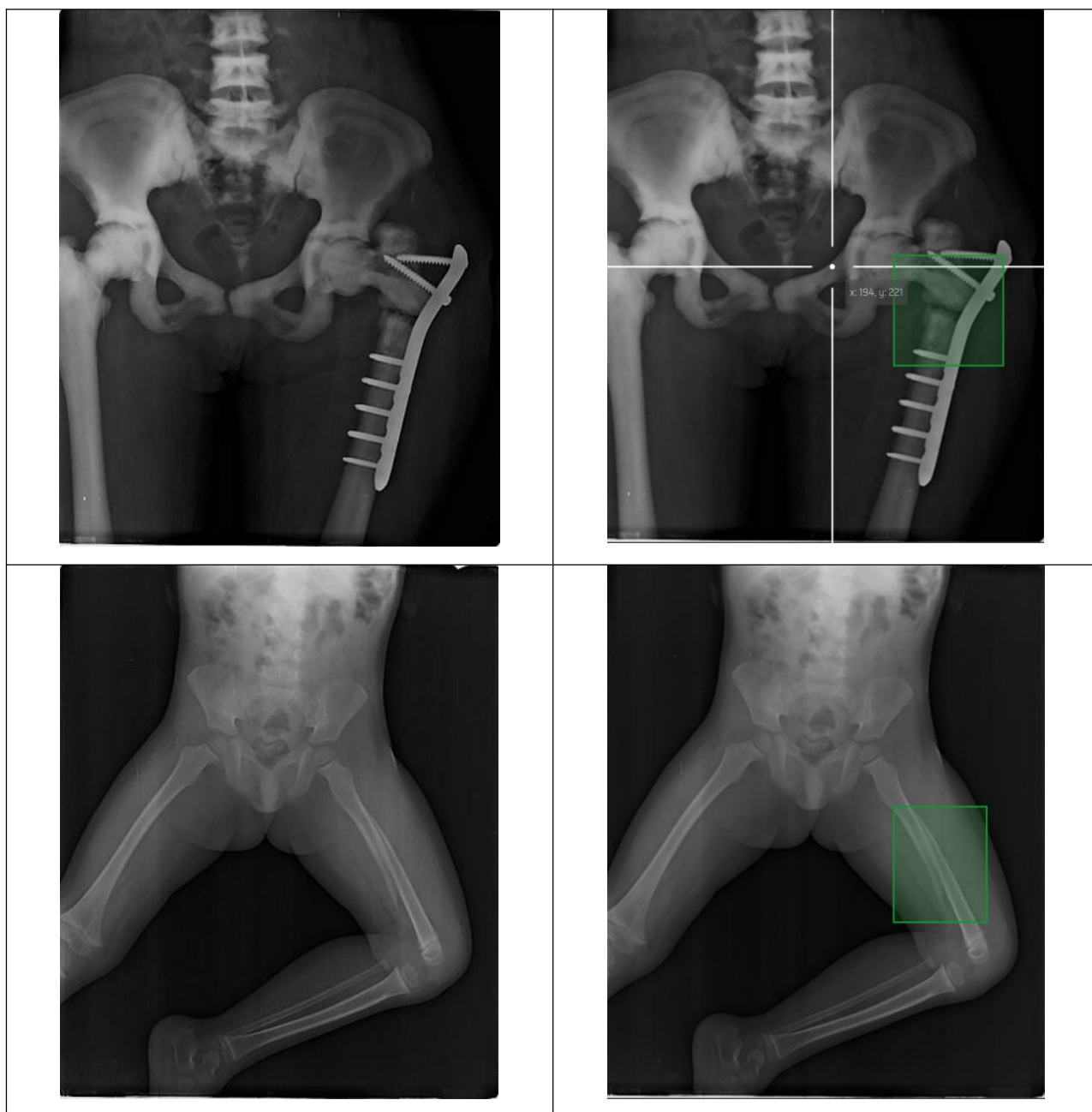
Інший знайдений набір даних “Bone fracture detection” [7] містить близько 1000 зображень в непоганій якості. Набір містить один клас, який позначає перелом, що є недоліком. З переваг — наявність анотацій.

2.5. Аналіз методів обробки навчального набору даних

Проте наявність зображень не є достатнім фактором для початку роботи із моделлю YOLO. Як вже було згадано, для навчання моделі потрібно вказати рамки, щоб модель “розуміла”, де знаходиться потрібний нам об’єкт. Для цього можна використати готові веб-застосунки [18], які набагато полегшують роботу з анотаціями (labeling) до зображень. Проте все одно ручної роботи не уникнути, оскільки хто як не людина зможе найточніше визначити цільовий об’єкт. Нижче продемонстровано роботу із тегування зображень датасету.

Таблиця 2.1. Приклад тегування зображень для навчання





Після завершення тегування зображень дані про встановлені рамки експортуються в спеціальному форматі, який очікує YOLO. Це є текстові (.txt) файли, що названі за ім'ям відповідного зображення, які містять 5 значень: число класу та 4 координати рамки. Файли, що відповідають зображенням без перелому кісток — пусті.

Location: /			
Name	Size	Type	Modified
IMG0000777.txt	37 bytes	plain text ...	30 вересня 2023, 11:18
IMG0000822.txt	37 bytes	plain text ...	30 вересня 2023, 11:18
IMG0001774.txt	37 bytes	plain text ...	30 вересня 2023, 11:18

Рис.2.8. Приклад експортованого архіву з тегами та рамками

```
1 0 0.783610 0.568770 0.248158 0.205502
```

Рис.2.9. 1-ий приклад вмісту файлів-лейблів

```
1 0 0.762930 0.565534 0.210738 0.215210
```

Рис.2.10. 2-ий приклад вмісту файлів-лейблів

Маючи в наявності протегований набір даних, можна розбити його на тренувальний, тестовий та валідаційний набори. Структура цих наборів також має бути не аби-яка, оскільки YOLOv8 очікує спеціально продуманий формат: кожна з трьох директорій має містити дві під-директорії — labels та images, де друга має містити самі зображення, а перша — текстові файли-анотації, названі один в один, як відповідне зображення. Також має бути конфігураційний файл формату YAML, який вказує де знаходяться директорії із навчальними даними, кількість класів та їх назви (сам конфігураційний файл може містити набагато більше різних параметрів, тут було наведено лише необхідний мінімум для роботи моделі).

```

Open  data.yaml  Save
~/Documents/University/1_Marictp/DATASET_2

1 train: ../train
2 val: ../valid
3 test: ../test
4
5 nc: 3
6 names: ['A', 'B', 'C']
7

```

Рис.2.11. Приклад конфігураційного файлу

2.6. Висновки до розділу 2

В даному розділі було проведено аналіз інструментарію використаного в досліджених джерелах та обґрунтовано вибір моделі для подальшої реалізації поставленої задачі, якою стала модель YOLO (п.2.1).

Модель YOLO є одноразовим детектором, що дозволяє алгоритму бути швидким та обчислювально ефективним. Цим він може забезпечувати швидкість надання передбачень для поставленої задачі. Проте такого виде детектори не мають великої точності для малих об'єктів, що спричинить складнощі при виявленні малих переломів.

Після цього в пункті 2.2 було наведено детальний опис архітектури моделі, та основних принципів її роботи. Архітектуру YOLO можна поділити на три частини — хребет, шию та голову, де перша частина складається із згорткової нейронної мережі, яка виявляє ознаки, друга — комбінує та змішує їх, і третя — робить передбачення розташування об'єкта та його класу. Окрім цього в даному пункті були описані важливі деталі моделі, такі як non-maximum suppression та IOU.

В п.2.3 було проведено аналіз останньої моделі сімейства — YOLOv8 — та її відмінності від попередніх. Тут описані зміни до роботи моделі, такі як Anchor-Free Detection та мозаїчна аугментація, які покращують та удосконалюють тренувальний процес. Також були згадані деякі техніки, які були унаслідковані із v7 моделі та мають важливе значення в процесі навчання моделі. В кінці пункту були наведені графіки порівняння точності та швидкості нової моделі відносно декількох старіших моделей.

В наступному пункті, 2.4, відбувся аналіз 4 відібраних наборів даних, де були наведені їх переваги та недоліки. Загалом можна підмітити, що більшість наборів даних (ті, що не були вказані в пункті) містять зображення зап'ясть та рук. Цим можна також пояснити наявність великої кількості

досліджень на дану тему. Для даної ж роботи були проведені пошуки наборів даних із зображеннями переломів різних ділянок тіла.

В останньому пункті були описані метод підготовки набору даних для роботи із моделлю YOLOv8. Для того щоб модель запрацювала, потрібно протегувати зображення, які не мають анотацій, розбити загальний набір на спеціальну структуру директорій та заповнити конфігураційний файл.

Підсумовуючи даний розділ, YOLOv8 є хорошим претендентом на вирішення поставленої задачі. Окрім цього, були знайдені декілька наборів даних із зображеннями та розбиттям на класи. А також було досліджено спосіб підготовки даних для навчання моделі.

3. РОЗДІЛ АПРОБАЦІЙ ТА РЕЗУЛЬТАТІВ

3.1. Підготовка набору даних

Навчальний набір даних був сформований із двох датасетів, проаналізованих в п.2.4., а саме “fracture types” [14] та “long bones x rays” [16]. Оскільки вони мають схожість в класах, було вирішено звести їх до одної класифікації. Оскільки один набір має три класи, а інший 9, рішення була на користь узагальнення більшої кількості класів до меншої. Таким чином: A1, A2 та A3 стали A, B1, B2 та B3 стали B і C1, C2, C3 — C. Наступним кроком, за допомогою онлайн-інструменту [18] для всіх зображень набору [16] були створені анотації шляхом виділення місця чи кількох місць перелому у прямокутник, який формує чотири координати для моделі. Після цього, наступним кроком треба розбити нашу вибірку на три частини — навчальну, валідаційну та тестову. Для цього був написаний python-скрипт, який за допомогою модуля random вибирає картинку та відповідну їй анотацію, та копіює у три різні директорії, формуючи необхідну для алгоритму YOLO структуру.

Наступний кроком було застосування методу адаптивного вирівнювання гістограми з обмеженим контрастом, або CLAHE (Contrast-Limited Adaptive Histogram Equalization). Цей алгоритм є варіантом вирівнювання гістограми, який широко використовується в програмах обробки зображень для покращення контрастності зображення, уникаючи надмірного посилення шуму та зберігаючи деталі зображення. Основна ідея CLAHE полягає в тому, щоб виконувати вирівнювання гістограми локально, у менших областях зображення, а не глобально.

У традиційному вирівнюванні гістограми зображення ділиться на невеликі ділянки або частини зображення, які не перекриваються, і кожне зображення вирівнюється окремо. Цей процес може призвести до надмірного посилення шуму в підзображеннях із низьким контрастом, що призведе до шумного та штучного вигляду зображення. На противагу, CLAHE ділить

зображення на ділянки, що перекриваються, і застосовує вирівнювання гістограми до кожної області окремо. Це дозволяє покращити контрастність, зберігаючи локальні деталі та уникаючи надмірного посилення шуму.

CLANE складається з двох основних етапів: підвищення контрастності та обмеження контрастності. На першому кроці локальне вирівнювання гістограми застосовується до кожної невеликої області зображення, що призводить до підвищення контрастності. На другому кроці контрастність зображення обмежується шляхом застосування нелінійної функції до гістограми, так що кількість пікселів з дуже високою або дуже низькою інтенсивністю зменшується. Нелінійна функція визначається параметром, який визначає величину обмеження контрастності, яке буде застосовано.

CLANE має декілька переваг перед традиційним вирівнюванням гістограми. По-перше, він може зберегти локальні деталі та ефективніше підвищити контраст. По-друге, це може уникнути надмірного посилення шуму, що особливо важливо для медичних зображень, де зображення можуть бути шумними. Нарешті, це простий і обчислювально ефективний алгоритм, що робить його придатним для програм реального часу.

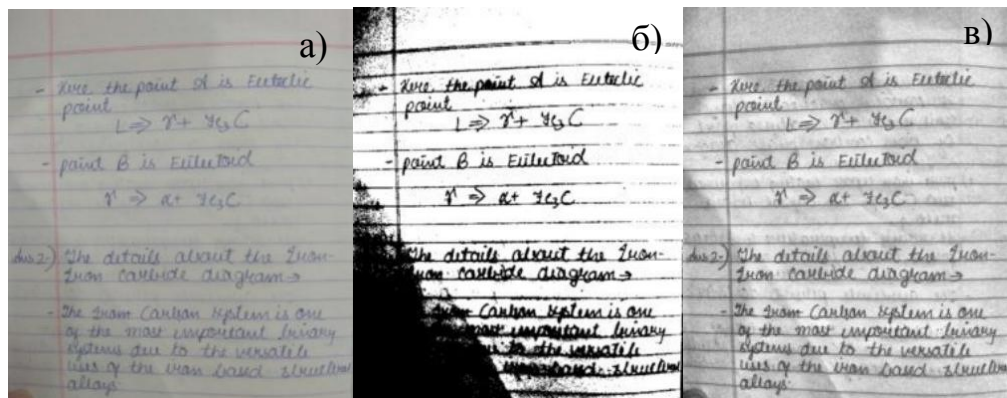


Рис.3.1. а) оригінальне зображення; б) застосування звичайного вирівнювання гістограми; в) застосування CLANE

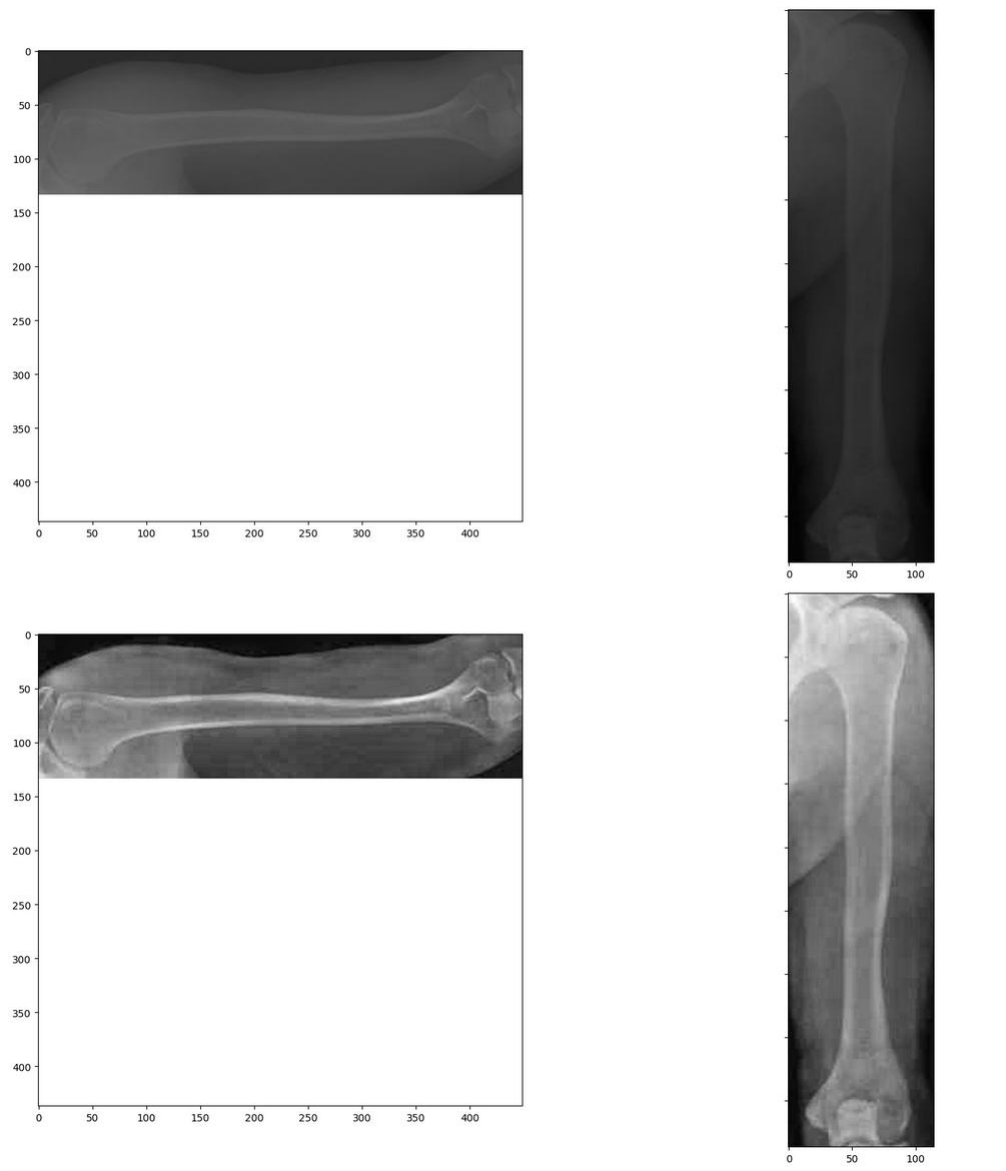


Рис.3.2. Приклад застосування CLAHE на рентгенівських знімках
(зверху оригінал, знизу оброблена)

3.2. Тренування моделі

Використання моделі YOLOv8 не складає труднощів, оскільки вона має простий Python інтерфейс та широку документацію, в якій наведені приклади. Модуль з моделлю надає можливості запускати процеси (тренування, валідація, передбачення, тощо) як за допомогою консольних команд, так і за допомогою коду.

Вибір середовища для тренування впав на два варіанти — Google Colab або Kaggle Notebook. Другий варіант виглядав більш привабливо, оскільки даний веб ресурс надає кращі обчислювальні ресурси. Проте критичним

недоліком виявилось те, що директорія із набором даних є захищеною від запису, а моделі YOLOv8 необхідно створити кеш безпосередньо в директорії з датасетом (і варіанту створити в іншій директорії не було). Тому було обрано середовище Google Colab. Дане середовище надає можливості апаратного прискорення за допомогою графічного процесора. У безплатному доступі є лише одна така машина, а саме Tesla T4 із 16 ГБ пам'яті GDDR6 та 2,560 ядрами CUDA, а також 13 ГБ оперативної пам'яті та 78 ГБ простору на диску.

Модель YOLOv8 має декілька варіацій — n (nano), s (small), m (medium), l (large) та x (x-large). Кожна з них відрізняється кількістю ваг та швидкістю.

Model	size (pixels)	mAP ^{val} 50-95	Speed CPU ONNX (ms)	Speed A100 TensorRT (ms)	params (M)	FLOPs (B)
YOLOv8n	640	37.3	80.4	0.99	3.2	8.7
YOLOv8s	640	44.9	128.4	1.20	11.2	28.6
YOLOv8m	640	50.2	234.7	1.83	25.9	78.9
YOLOv8l	640	52.9	375.2	2.39	43.7	165.2
YOLOv8x	640	53.9	479.1	3.53	68.2	257.8

Рис.3.3. Порівняння різних варіантів моделі YOLOv8

Для цієї роботи була вибрана YOLOv8s, оскільки це є максимально можливий варіант в межах ресурсів машини Tesla T4.

Останнім кроком є підбір гіперпараметрів, яких в нашої моделі є близько 50. Для цієї роботи були змінені лише декілька з них, а саме:

- Кількість епох: epochs=150,
- Кількість зображень в батчі: batch=64,
- Кількість епох без покращення метрик: patience=30,
- Рзмір зображення: imgsiz=640,
- Спосіб кешування: cache='disk',
- Одно- чи мультикласовість: single_cls=False,

- Претренована модель: pretrained=True,
- Оптимізатор: optimizer='Adam',
- Швидкість навчання: lr0=0.001,
- Кількість градусів аугментації повороту зображень: degrees=15,
- Кількість епох з мозаїчною аугментацією: close_mosaic=75.

3.3. Обговорення результатів

Оцінка результатів буде відбуватися за кількома метриками: Precision, Recall, mAP50, mAP50-95.

Precision — метрика, яка розраховується як відношення кількості “true positive” передбачень до загальної кількості “positive” передбачень. Тобто, якщо модель передбачила 100 переломів, але лише 90 з передбачень були правильними, тоді метрика precision буде дорівнювати 90%. Іншими словами, ця метрика позначає точність виявлення об’єкту.

Recall — метрика, яка розраховується як відношення кількості “true positive” передбачень до загальної кількості справжніх (“true”) об’єктів. Тобто, якщо модель передбачила 75 переломів, а їх насправді було 100, тоді метрика recall буде дорівнювати 75%. Іншими словами, ця метрика позначає чутливість до виявлення об’єктів.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

Рис.3.4. Формули обрахування precision та recall

Mean average precision, або mAP — метрика, яка розраховується як середнє значення між AP по всіх класах. AP, або Average Precision, розраховується як середнє зважене Precision для якогось порогового значення чи проміжку. Для прикладу mAP50 означає, що порогове значення 0.5, або 50%, а mAP50-95 означає пороговий проміжок 0.5-0.95, або 50-95%.

Порогове значення в даній метриці береться для IoU (цю метрику ми вже згадували в п.2.2).

Тож перейдемо до результатів тренування моделі YOLOv8 на комбінованому датасеті з трьома класами.

Модель показала наступний результат по метриках: Precision — 64.8%, Recall — 64.8%, mAP50 — 63.3% та mAP50-95 — 26.0%. Ці значення можна інтерпретувати наступним чином: дана модель має посередню точність та може точно класифікувати більшу частину випадків, також mAP50 показує, що більшість передбачених позицій об'єкту мають перекриття із справжніми позиціями мінімум на 50%.

З табл.3.1 можна побачити ці метрики для кожного з класів: А, В та С. Можна помітити, що модель найточніше класифікує третій клас, найбільше правильних класифікацій надає першому класу, та найточніше визначає позицій також першого класу.

Таблиця 3.1. Значення метрик для класів

Клас	Precision	Recall	mAP50	mAP50-95
всі	0.648	0.648	0.633	0.26
А	0.662	0.752	0.704	0.264
В	0.576	0.559	0.563	0.223
С	0.704	0.632	0.632	0.293

Нормалізована confusion-матриця (рис.3.5) також показує, що модель краще всього справляється із випадками класу А, яких скорше всього є більшість в наборі даних (оскільки це простий, найбільш поширений тип перелому). Після того клас С, та В. З проблемних моментів можна помітити, що доволі багато картинок без переломів (background) було передбачені як класи А та В, що може свідчити про ймовірність про некоректні анотації класів.

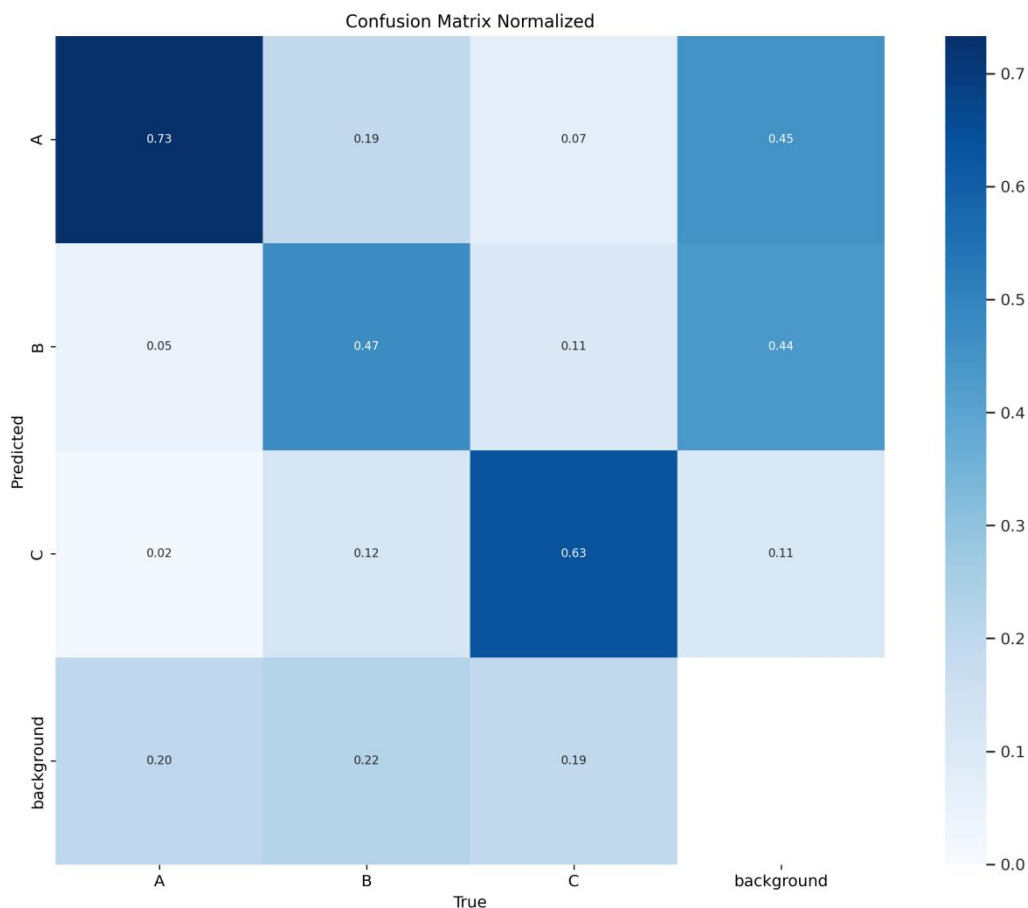


Рис.3.5. Нормалізована confusion-матриця

На рис.3.6 можна чітко побачити причини даної поведінки моделі та підтвердження попередніх гіпотез, а саме домінуючу поведінку класу А. Основна причина це тотальна незбалансованість класів В та С відносно класу А.

Проте, окрім цього, можна спостерігати й згадану в п.2.3 техніку “Focal Loss”. Попри те, що клас А значно переважає у вибірці, решта класів все одно мають близькі до нього результати значень Precision та Recall. Ба більше, не зважаючи на те, що класу С найменше, він має більше значення ніж клас В у метриці Recall, та більше значення ніж інші два класи у метриці Precision. Ця техніка може бути дуже корисна у випадку обмеженої кількості даних.

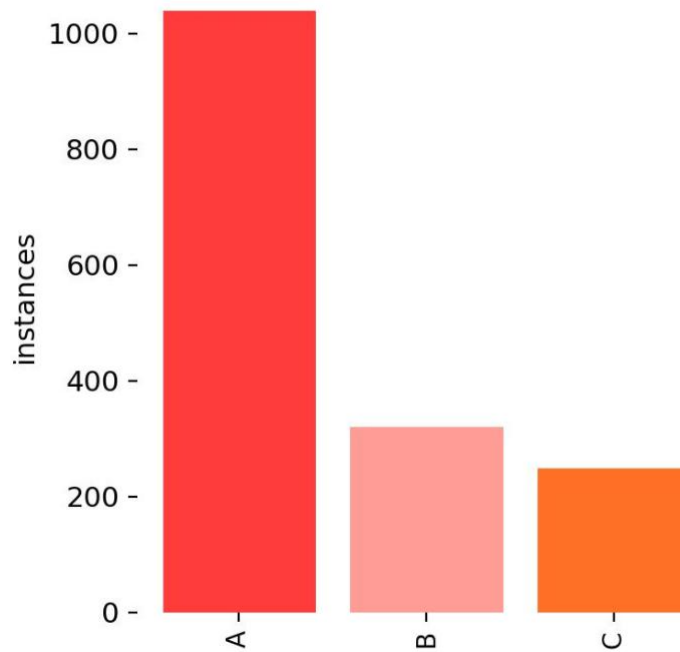


Рис.3.6. Стовпчикова діаграма розподілу класів

З графіку кривої Precision-Recall (рис.3.7), можна побачити що для класу A можна спостерігати доволі хороше значення метрики Precision, приблизно 0.75, разом із таким ж значенням Recall, приблизно 0.75 також. Клас B має більш стрімке підіння Precision ніж клас C, проте трішки кращі показники при Recall більше ~ 0.7 .

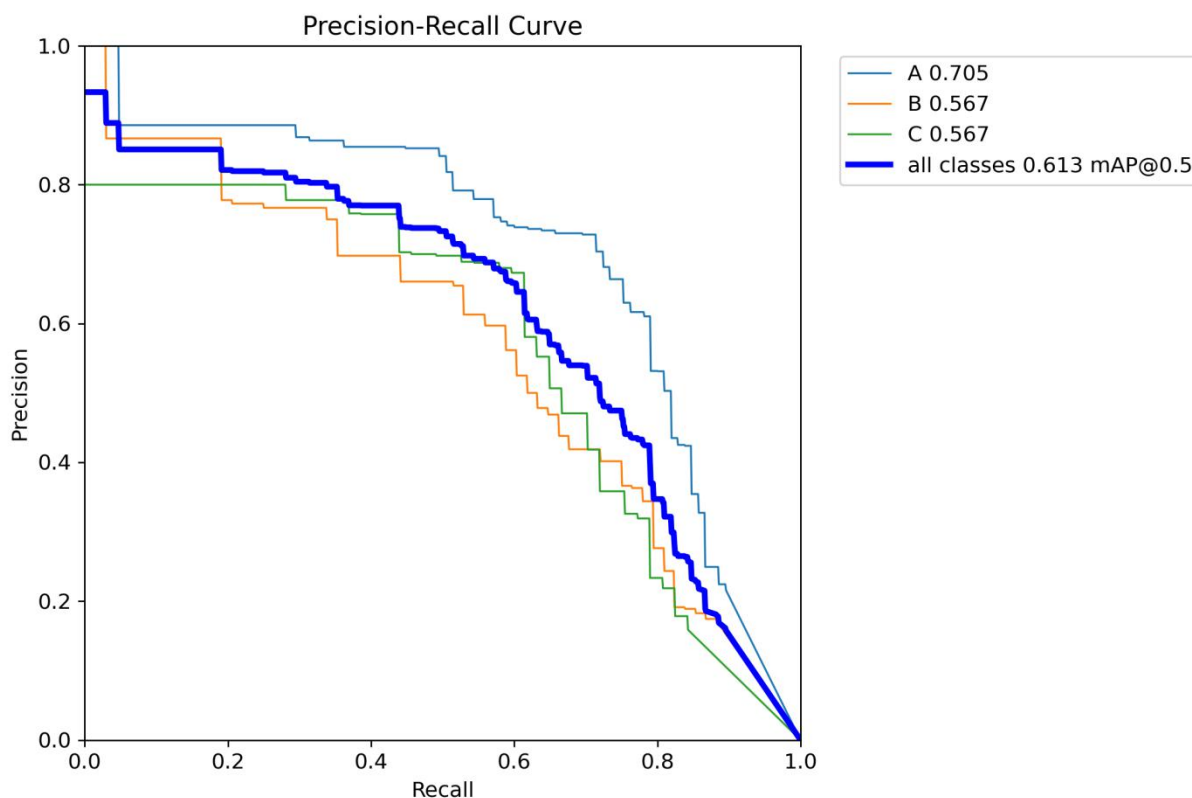


Рис.3.7. Графік кривої Precision-Recall

На рис.3.8 та рис.3.9 можна спостерігати графіки функції втрат на тренувальній та валідаційній вибірках. По ліву сторону зображений графік втрат на передбаченні рамки, по праву сторону — графік втрат передбачення класу. На тренувальній вибірці (рис.3.8) ми можемо спостерігати стабільне спадання функції для передбачення рамки, для передбачення класу можна помітити різке зниження на 75 епохах. Це можна пояснити тим, що на 75 епосі включається мозаїчна аугментація, що дозволяє краще навчитись моделі розрізняти класи.

На валідаційній вибірці (рис.3.9) можна спостерігати доволі нестабільну функцію втрат на передбаченні рамки, яка з часом починає зростати.

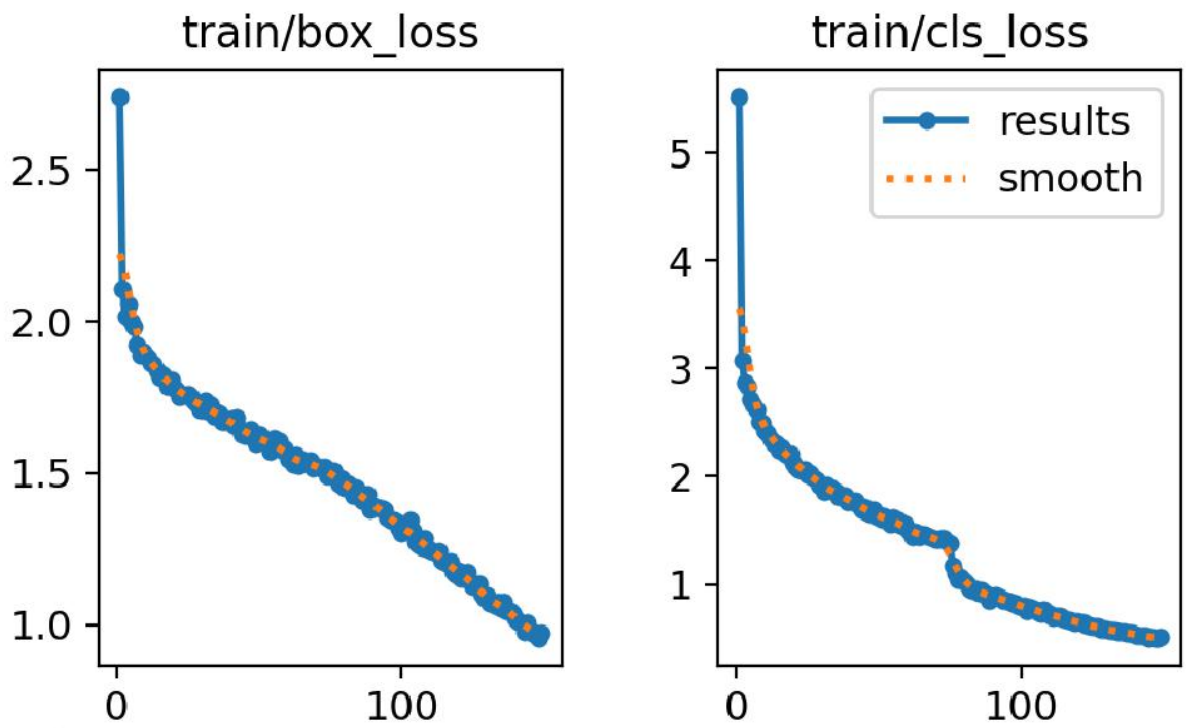


Рис.3.8. Графік функції втрат на тренувальній вибірці

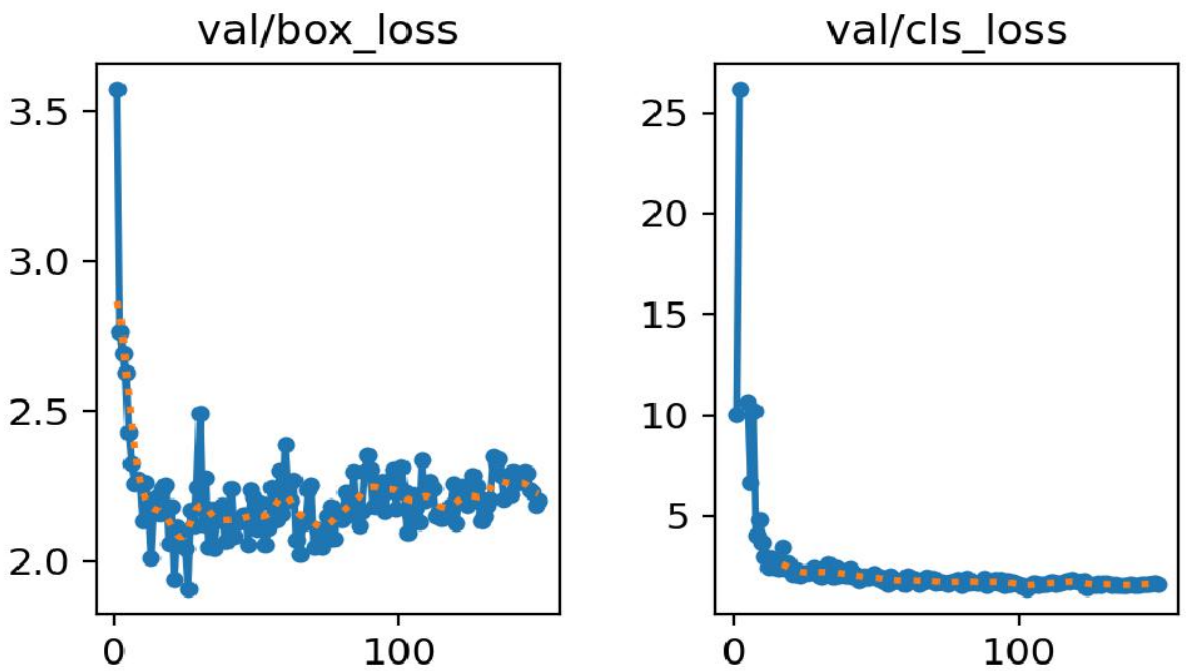





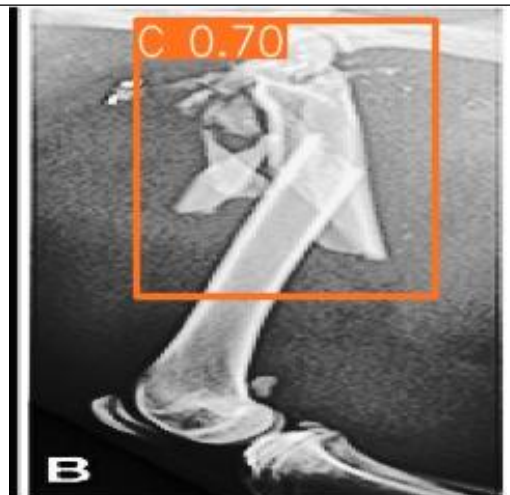


Рис.3.9. Графік функції втрат на валідаційній вибірці

Також нижче в таблиці 3.2 буде наведена демонстрація роботи моделі з передбачення та класифікації переломів.

Табл.3.2. Приклади роботи моделі

Оригінал	Передбачення
	
 В	 В
 В	 В



Також окрім мультикласового тренування було проведено експеримент з одним класом — “перелом”, який мав на меті перевірити точність оприділення місця перелому, без класифікації тяжкості. Тренування відбувалося на датасеті “Bone fracture detection” [7]. В результаті модель показала наступний результат по метриках: Precision — 85.2%, Recall — 73.5%, mAP50 — 79.9% та mAP50-95 — 34.2%.

На кривій Precision-Recall можна спостерігати доволі довгий проміжок з високим Precision, що означає високий шанс правильного передбачення класу. Проте із збільшенням значення Recall є суттєве зменшення точності Precision.

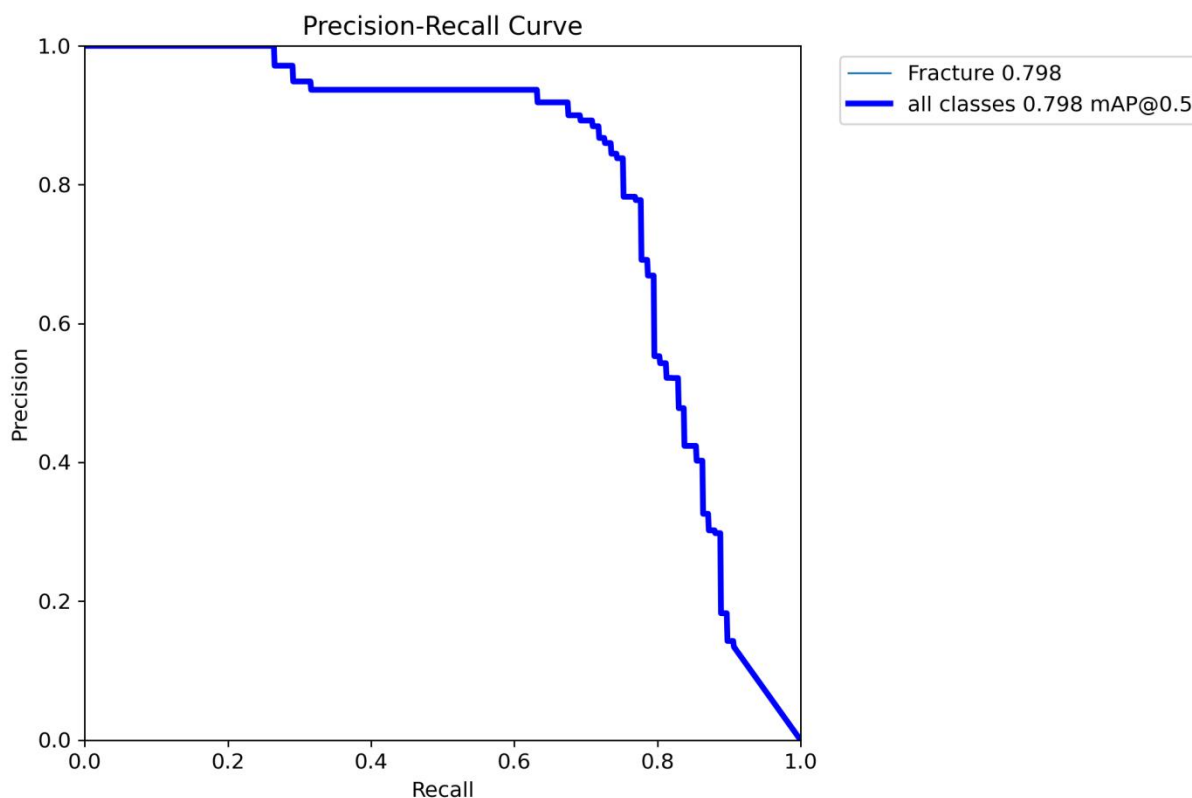


Рис.3.10. Графік кривої Precision-Recall

На рис.3.11 можна спостерігати графіки функції втрат на тренувальній вибірці однокласової моделі. На відміну від графіків мультикласової моделі тут не видно різкого зниження функції втрат для класу на моменті мозаїчної аугментації, що є цілком логічним явищем, оскільки клас один і немає потреби для моделі вивчати його варіації та відмінності. Також можна помітити стабільне зменшення функції втрати для передбачення позиції об'єкту.

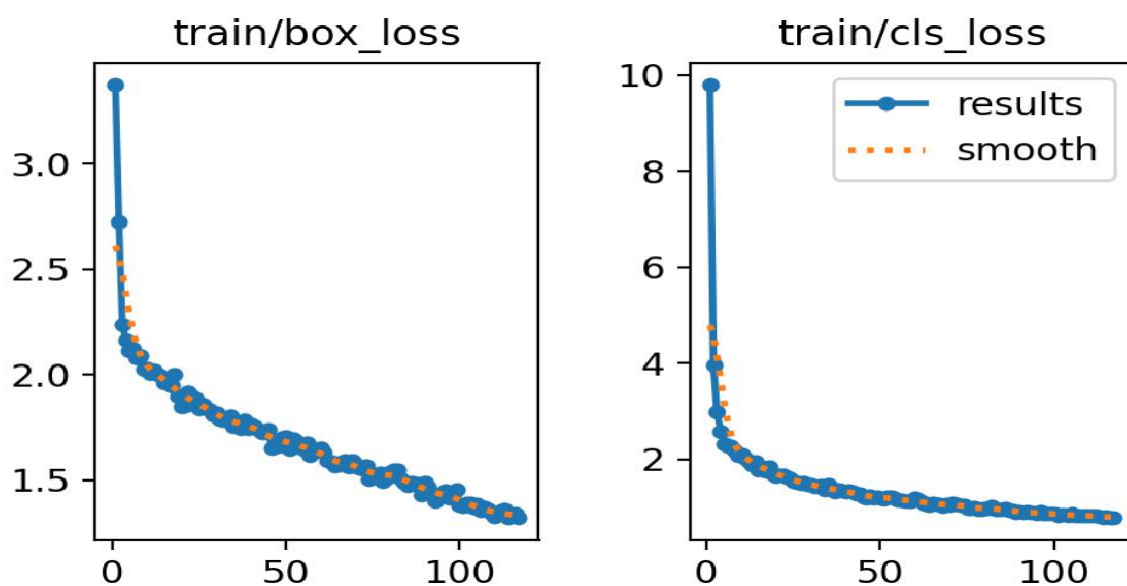


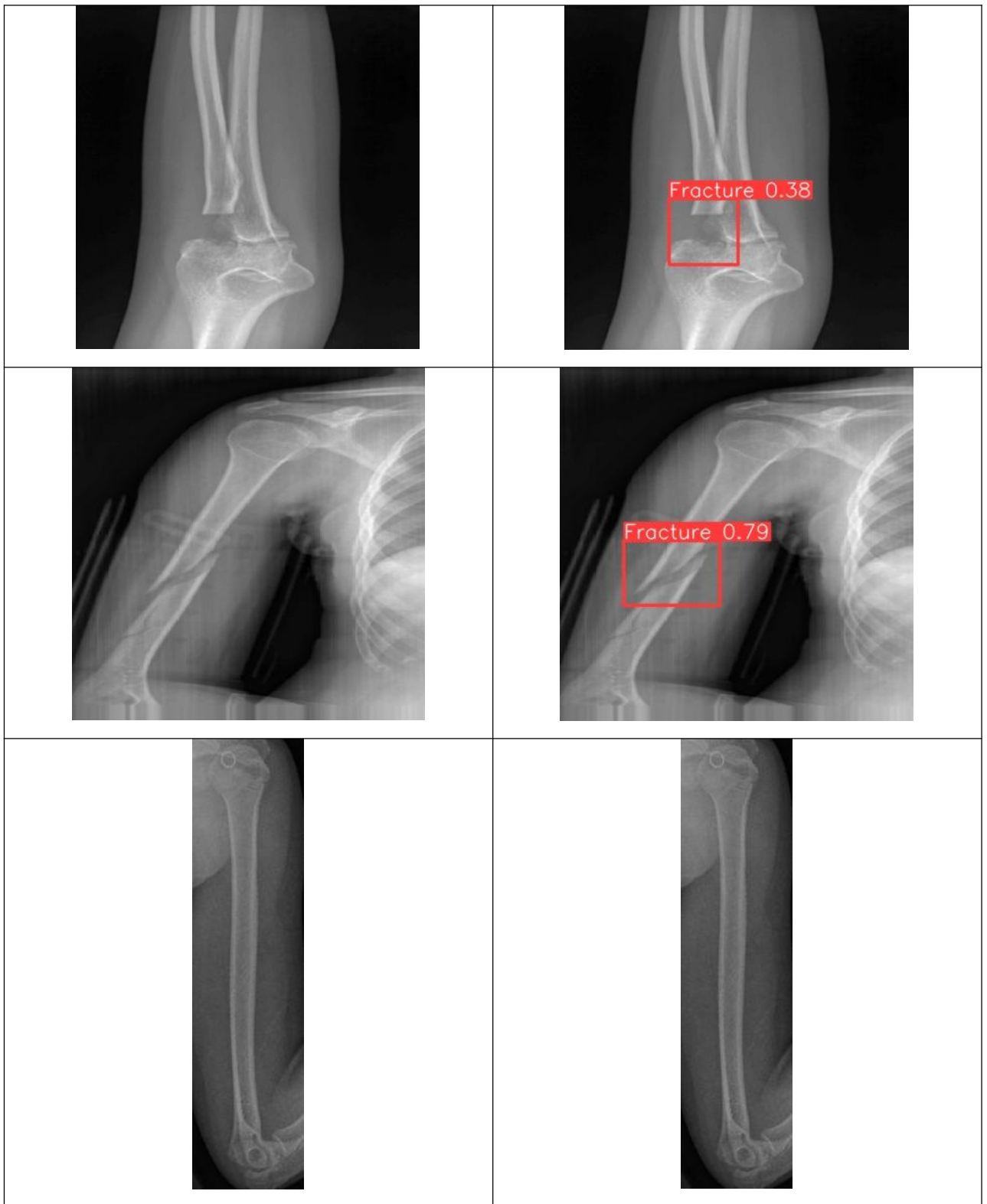


Рис.3.11. Графік функції втрат на тренувальній вибірці

Далі в таблиці 3.3 будуть продемонстровані приклади роботи цієї моделі.

Табл.3.3. Приклади передбачень однокласової моделі

Оригінал	Передбачення
	



3.4. Висновки до розділу 3

В даному розділі були проведені експерименти з навчанням моделі та оцінка її результатів. В ході експериментів були застосовано дві стратегії: мультикласова та однокласова. Оцінка моделей включала в себе декілька метрик, а саме Precision, Recall, mAP50 та mAP50-95.

Мультикласова стратегія включала в себе тренування моделі YOLOv8 на датасеті із трьома класами А, В та С, які репрезентують складність перелому. Результати оцінки даної моделі виявились доволі посередніми: загальні Precision із значенням 0.648 вказує на переважно правильні предбачення класу, Recall із значенням 0.648 вказує на переважно точні передбачення місця перелому, та mAP50 із значенням 0.633 вказує на те, що передбачене місце перелому співпало із аноцією на 63% в середньому.

Однокласова стратегія, як можна зрозуміти із назви, включала в себе тренування лише на одному класі, що фактично дорівнює простій детекції перелому. Результати оцінки даної моделі виявились кращими за попередні: Precision — 0.852, Recall — 0.735, та mAP50 — 0.799. Проте варто мати на увазі, що даний експеримент проводився на іншому датасеті.

Отже, що могло повпливати на такий результат? Скорше всього, мультикласовий набір даних мав гіршу якість зображень, хоч і більшу їх кількість. Також фактором могли бути і некоректно вказані класи в анотаціях, оскільки набір даних є у відкритому доступі, тому достеменно невідомо наскільки професійно автор протегував зображення. Окрім цього мультикласовий датасет був скомбінований із двох, тому це теж могло зіграти свою роль. Однокласовий датасет не був комбінованим, та мав візуально кращу якість зображень. Також відсутність різноманіття класів суттєво полегшило задачу моделі.

Підсумовуючи, можна сказати, що модель YOLOv8 явно має перспективи у застосуванні для вирішення задачі виявлення переломів на рентгенівських знімках, проте без якісного набору даних з коректними анотаціями важко досягти високих результатів.

ВИСНОВОК

Кожна людина хоча б раз у житті зверталась до лікаря, адже медична сфера є невід'ємною частиною людства. Саме тому якість медичних послуг є дуже важливим аспектом, оскільки чим точніше буде визначений діагноз, тим краще пройде лікування, а чим швидше буде визначений діагноз, тим більше шансів вчасно надати допомогу в лікуванні, чи навіть запобігти хворобі.

Проте навіть найдосвідченіші лікарі можуть робити помилки. Є випадки, коли помилку можна допустити, але бувають ситуації коли немає права на помилку. Тому досить важливо мати джерело, чи інструмент, за допомогою якого можна перевірити свою гіпотезу, чи пришвидшити пошук проблеми. На цьому кроці вступає в гру штучний інтелект, який давно вже себе зарекомендував як інструмент покращення ефективності роботи людини не в одній сфері.

У даній праці була проведена робота з розробки моделі глибокого навчання для класифікації та виявлення переломів кісток на рентгенівських знімках. В ході аналізу теоретичних матеріалів було обрано архітектуру моделі сімейства YOLO, а саме восьму версію, як претендента на вирішення поставленої задачі. Далі було проведено два експеримента: виявлення та класифікація на три класи, та лише виявлення переломів. Результати оцінки моделей виконувались за допомогою метрик, таких як Precision, Recall, mAP50 та mAP50-95. Їх розуміння та аналіз надали уявлення про ефективність розроблених моделей.

Мультикласова стратегія, яка передбачала тренування моделі YOLOv8 на датасеті із трьома класами, продемонструвала обмежені результати. Ця модель виявилася менш точною, з Precision та Recall на рівні 0.648, та mAP50 у 0.633. Помітно, що складність датасету та, ймовірно недостатня якість, вплинула на загальну ефективність моделі.

Натомість, однокласова стратегія, орієнтована на простіше завдання детекції переломів, продемонструвала кращі результати з Precision 0.852, Recall 0.735, та mAP50 0.799. Важливо відзначити, що цей експеримент проводився на іншому датасеті, що може пояснити поліпшення в ефективності.

Однак, попри досягнуті покращення в однокласовому підході, слід врахувати фактори, які могли вплинути на менш успішні результати в мультикласовому варіанті, такі як низька якість зображень та можливі неточності в анотаціях.

У підсумку, модель YOLOv8 проявила свій потенціал у виявленні переломів на рентгенівських знімках, але вагомий вплив якості та репрезентації даних на результати підкреслює необхідність подальших досліджень та вдосконалення як самої моделі, так і якості наборів даних.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Zhong B. AC-faster r-cnn: an improved detection architecture with high precision and sensitivity for abnormality in spine x-ray images / B. Zhong, J. Yi, Z. Jin // Physics in medicine and biology. — 2023. — Vol. 68, No. 19.
2. Kim M. W. Application of convolutional neural networks for distal radio-ulnar fracture detection on plain radiographs in the emergency room / M. W. Kim, J. Jung, S. J. Park, [et al.] // Clinical and Experimental Emergency Medicine. — 2021. — Vol. 8, No. 2. — P. 120–127.
3. Singh A. Automated detection of scaphoid fractures using deep neural networks in radiographs / A. Singh, A. A. Ardakani, H. W. Loh, [et al.] // Engineering Applications of Artificial Intelligence. — 2023. — Vol. 122.
4. Erdaş Ç. B. Automated fracture detection in the ulna and radius using deep learning on upper extremity radiographs / Ç. B. Erdaş // Joint Diseases and Related Surgery. — 2023. — Vol. 34, No. 3. — P. 598–604.
5. Lu S. Automated universal fractures detection in x-ray images based on deep learning approach / S. Lu, S. Wang, G. Wang // Multimedia Tools and Applications. — 2022. — Vol. 81, No. 30. — P. 44487–44503.
6. Guan B. Automatic detection and localization of thighbone fractures in x-ray based on improved deep learning method / B. Guan, J. Yao, S. Wang, [et al.] // Computer Vision and Image Understanding. — 2022. — Vol. 216.
7. Bone fracture detection / URL: <https://universe.roboflow.com/fracture-uofxm/bone-fracture-detection-ivsy6>.
8. Meena T. Bone fracture detection using deep supervised learning from radiological images: a paradigm shift / T. Meena, S. Roy // Diagnostics (Basel, Switzerland). — 2022. — Vol. 12, No. 10. — P. 2420.
9. Bone fracture detection using x-rays / URL: <https://www.kaggle.com/datasets/vuppalaadithyasairam/bone-fracture-detection-using-xrays/data>.

10. Tsai H.-C. Cascading ab-yolov5 and pb-yolov5 for rib fracture detection in frontal and oblique chest x-ray images / H.-C. Tsai, N.-H. Lu, K.-Y. Liu, [et al.] // IET Computer Vision. — 2023. — Vol. 17, No. 7. — P. 750–762.
11. Cheng C.-T. Evaluation of ensemble strategy on the development of multiple view ankle fracture detection algorithm / C.-T. Cheng, C.-P. Hsu, C.-H. Ooyang, [et al.] // The British Journal of Radiology. — 2023. — Vol. 96, No. 1145. — P. 20220924.
12. FracAtlas: a dataset for fracture classification, localization and segmentation of musculoskeletal radiographs / figshare, 2023 / URL: https://figshare.com/articles/dataset/The_dataset/22363012.
13. Hardalaç F. Fracture detection in wrist x-ray images using deep learning-based object detection models / F. Hardalaç, F. Uysal, O. Peker, [et al.] // Sensors. — 2022. — Vol. 22, No. 3.
14. Fracture types object detection dataset and pre-trained model by modern academy for engineering and technology / URL: <https://universe.roboflow.com/modern-academy-for-engineering-and-technology/fracture-types>.
15. Yadav D. P. Hybrid sfnet model for bone fracture detection and classification using ml/dl / D. P. Yadav, A. Sharma, S. Athithan, [et al.] // Sensors (Basel, Switzerland). — 2022. — Vol. 22, No. 15. — P. 5823.
16. Long bones x rays classification dataset and pre-trained model by modern academy for engineering and technology / URL: <https://universe.roboflow.com/modern-academy-for-engineering-and-technology/long-bones-x-rays>.
17. Wang Y. Lower-extremity fatigue fracture detection and grading based on deep learning models of radiographs / Y. Wang, Y. Li, G. Lin, [et al.] // European Radiology. — 2023. — Vol. 33, No. 1. — P. 555–565.
18. Make sense / URL: <https://www.makesense.ai/>.
19. Kim S. Multiclass datasets expand neural network utility: an example on ankle radiographs / S. Kim, P. Rebmman, P. H. Tran, [et al.] // International Journal

of Computer Assisted Radiology and Surgery. — 2023. — Vol. 18, No. 5. — P. 819–826.

20. Kim T. Transfer learning-based ensemble convolutional neural network for accelerated diagnosis of foot fractures / T. Kim, T. S. Goh, J. S. Lee, [et al.] // Physical and Engineering Sciences in Medicine. — 2023. — Vol. 46, No. 1. — P. 265–277.

21. Shen L. Using artificial intelligence to diagnose osteoporotic vertebral fractures on plain radiographs / L. Shen, C. Gao, S. Hu, [et al.] // Journal of Bone and Mineral Research. — 2023. — Vol. 38, No. 9. — P. 1278–1287.

ДОДАТОК А

<https://github.com/AVorobel/masters-work>