

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»
ФАКУЛЬТЕТ БІОМЕДИЧНОЇ ІНЖЕНЕРІЇ

(повна назва інституту/факультету)

КАФЕДРА БІОМЕДИЧНОЇ КІБЕРНЕТИКИ

(повна назва кафедри)

«На правах рукопису»
УДК 004.891.3 + 616.1

«До захисту допущено»:
Завідувач кафедри БМК

_____ Євген НАСТЕНКО
(підпис) (ім'я, ПРІЗВИЩЕ)
« ___ » грудня 2022р.

Магістерська дисертація

на здобуття ступеня магістра
за освітньо-професійною програмою
«Комп'ютерні технології в біології та медицині»
зі спеціальності 122 «Комп'ютерні науки»

на тему: Нейронна мережа для виявлення відхилень грудної клітини на рентгенівських знімках

Виконав: студент II курсу, групи БС-11мп

НЕРУС ВЛАДИСЛАВ МИКОЛАЙОВИЧ

(прізвище, ім'я, по батькові)

(підпис)

Науковий керівник:

зав. каф. БМК ФБМІ, проф. Настенко Євген Арнольдович

(посада, науковий ступінь, вчене звання, прізвище, ім'я, по-батькові)

(підпис)

Рецензент:

(посада, науковий ступінь, вчене звання, прізвище, ім'я, по-батькові)

(підпис)

Засвідчую, що у цій магістерській дисертації немає запозичень з праць інших авторів без відповідних посилань.

Студент _____
(підпис)

Київ – 2022 року

**Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»**

Інститут (факультет) _____

БІОМЕДИЧНОЇ ІНЖЕНЕРІЇ

(повна назва)

Кафедра _____

БІОМЕДИЧНОЇ КІБЕРНЕТИКИ

(повна назва)

Рівень вищої освіти – другий (магістерський)

спеціальність _____

122 «Комп'ютерні науки»

Освітньо-професійна програма «Комп'ютерні технології в біології та
медицині»

(код і назва)

ЗАТВЕРДЖУЮ

Завідувач кафедри БМК

Свген НАСТЕНКО

(підпис)

(ім'я, ПРІЗВИЩЕ)

«31» серпня 2022р.

ЗАВДАННЯ

на магістерську дисертацію студенту

Нерусу Владиславу Миколайовичу

(прізвище, ім'я, по батькові)

1. Тема дисертації **Нейронна мережа для виявлення відхилень грудної клітини на рентгенівських знімках**

науковий керівник дисертації

Настенко Євген Арнольдович, професор

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від «09» листопада 2022 р. № 4133-с

2. Термін подання студентом дисертації: 01-05 грудня 2022 року

3. Об'єкт дослідження: **зображення грудної клітини, які були отримані за допомогою рентгену**

4. Вихідні дані: **публічний датасет Chest Xray-14, що містить рентгенівські зображення з 14 типами різних патологій**

5. Перелік завдань, які потрібно розробити: **1.Провести аналіз наукових джерел, які стосуються проблематики класифікації, детекції патологій на рентгенівських знімках грудної клітини 2. Дослідити сучасні архітектури нейронних мереж, які застосовуються в задачах мульти-класифікації. 3. Реалізувати сучасну архітектуру нейронної мережі та додати класифікуючий шар вкінці. 4. Провести попередню обробку даних, зробити розширення даних**

та просторову і кольорову аугментацію. 5.Провести тренування нейронної мережі за допомогою сучасних оптимізаторів. 6. Розрахувати стартап-проект для виведення результатів роботи на ринок. 7. Розробити концептуальні висновки за результатами магістерської дисертації.

6. Орієнтовний перелік графічного (ілюстративного) матеріалу: **35 рисунків, 15 таблиць, презентація на 16 слайдів.**

7. Орієнтовний перелік публікацій: **1 публікація у фаховому виданні**

8. Консультанти розділів дисертації: -

9. Дата видачі завдання: **31 серпня 2022 р.**

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації	Примітка
1	Отримання завдання на МД	31 серпня 2022р.	виконано
2	Практика	з 01.09.2022р. по 28.10.2022р.	виконано
3	Вставлення результатів практики в розділи МД	5 листопада 2022р.	виконано
4	Дописування скриптів для візуалізації	6-7 листопада 2022р.	виконано
5	Перетреновунвання нейронної мережі	7-12 листопада 2022р.	виконано
6	Оформлення додаткових візуальних даних	12-18 листопада 2022р.	виконано
7	Підготовка та оформлення розділів МД	18-26 листопада 2022р.	виконано
8	Складання висновків до розділів	27-28 листопада 2022р.	виконано
9	Написання загальних висновків до магістерської роботи	29 листопада 2022р.	виконано
10	Подання в електронному вигляді МД та анотації до неї на перевірку нормоконтролера та UNICHECK .	1-34 грудня 2022р	виконано
11	Передзахист МД та допуск до захисту дисертації	4-6 грудня 2022р.	виконано
12	Подання МД рецензенту. Отримання рецензії.	7-15 грудня 2022р.	виконано
13	Подання пакету документів по МД до захисту в ЕК	15-18 грудня 2022р.	виконано
14	Захист МД в ЕК	19 грудня 2022р.	виконано

Студент

_____ (підпис)

Владислав НЕРУС

_____ (ім'я, ПРІЗВИЩЕ)

Науковий керівник

_____ (підпис)

Євген НАСТЕНКО

_____ (ім'я, ПРІЗВИЩЕ)

Нормоконтролер

_____ (підпис)

Галина КОРНІЄНКО

_____ (ім'я, ПРІЗВИЩЕ)

РЕФЕРАТ

Магістерська дисертація за темою «Система діагностики тромбоемболії легеневої артерії за даними комп'ютерної томографії органів грудної клітини» виконана студентом кафедри біомедичної кібернетики ФБМІ Нерусом Владиславом Миколаєвичом зі спеціальності 122 «Комп'ютерні науки» за освітньо-професійною програмою «Комп'ютерні технології в біології та медицині», та складається зі: вступу; 4 розділів (літературний огляд, теоретична частина, практична частина, аналіз стартап-проекту), висновків до кожного з цих розділів; загальних висновків; списку використаних джерел, який налічує 68 джерел. Загальний обсяг роботи 105 сторінок.

Актуальність теми. Респіраторні захворювання, такі як пневмонія, є поширеним станом легеневої інфекції, а хвороба COVID-19 стала небезпечною для життя хворобою, яка з'явилася наприкінці 2019 року та вразила весь світ. Пневмонія — це смертельна інфекція нижніх дихальних шляхів, яка відноситься до категорії гострих захворювань і, як повідомляється, є основною причиною смерті в усьому світі. У 2017 році на нього припадало 15% дитячих смертей, які трапилися протягом року [34]. Крім того, літні люди мають високий ризик захворіти на пневмонію, що призводить до критичних станів. Однак, якщо діагностувати та лікувати на ранній стадії, пов'язаний ризик можна мінімізувати [34]. Тому, система для автоматичного діагностування хвороб грудної клітини може бути використана як рекомендувальний інструмент рентгенолога задля зменшення помилок та людських помилок.

Мета і завдання дослідження. Метою роботи є розробка та навчання нейронної мережі для класифікації хвороб грудної клітини за рентгеном. Її досягнення передбачає вирішення наступних завдань:

1. Огляд літератури із обраної тематики.
2. Аналіз існуючих нейронних мереж.
3. Аугментація даних.
4. Нормалізація вибірки даних.

5. Розробка архітектури нейронної мережі.

6. Реалізація чи імпорт нейронної мережі та її тренування.

Об'єкт дослідження. Рентгенівський знімок грудної клітини.

Предмет дослідження. Згорткові нейронні мережі в задачах мульти-класифікації рентгенівських зображень, аугментація зображень

Методи дослідження. Методи розширення даних, методи боротьби з дисбалансом класів, операції згортки, пулінгу, згорткові нейронні мережі, оптимізація нейронних мереж.

Публікації. За результатами виконаної роботи було опубліковано 1 наукову статтю:

1. Нерус В. М. Використання нейронних мереж для діагностики захворювань грудної клітини на рентгенограмах // Міжнародний науковий журнал "Інтернаука". — 2022. — №14. <https://doi.org/10.25313/2520-2057-2022-14-8414>

Ключові слова. нейронна мережа, CNN, рентген, КТ зображення, комп'ютерний зір, аугментація.

Бібліографічний опис МД

Нерус В. М..Нейронна мережа для виявлення відхилень грудної клітини на рентгенівських знімках: магістерська дис. : 122 Комп'ютерні науки / Нерус Владислав Миколайович. – Київ, 2022. – 105 с.

ABSTRACT

Master's thesis on " A neural network for detecting the chest abnormalities on X-ray images" is executed by the student of the department of biomedical cybernetics (Faculty of Biomedical Engineering) Nerus Vladyslav Nikovaevich in the specialty 122 "Computer science" on the educational and professional program "Computer and technology" consists of: introduction; 4 sections (literature review, theoretical part, practical part, analysis of a startup project), conclusions to each of these sections; general conclusions; a list of used sources, which includes 69 sources and applications. The total volume of the work is 105 pages

Relevance of the topic. Respiratory diseases such as pneumonia is a common lung infection condition and COVID-19 disease has become a life-threatening disease that emerged in later 2019 and has been impacted the entire world. Pneumonia is a fatal lower respiratory infection under the acute diseases category and has been reported to be a major cause of deaths around the world. In 2017, it was accountable for 15% of child deaths that happened during the year[34]. In addition, older people have a high risk of getting pneumonia that leads to critical conditions. However, if diagnosed and treated early, the associated risk can be minimized [34]. Therefore, the system for automatic diagnosis of chest diseases can be used as a recommended tool of the radiologist to reduce mistakes and human errors.

Objective of the study. The purpose of the work is the development and training of a neural network for the classification of chest diseases based on X-rays. Its achievement involves solving the following tasks:

1. Review of the literature on the selected topic.
2. Analysis of existing neural networks.
3. Augmentation of data.
4. Normalization of ata sampling.
5. Development of neural network architecture.
6. Implementation or import of a neural network and its training.

Object of study. X-ray images

Subject of study. Convolutional neural networks in multi-classification of X-ray

images.

Research methods. Methods of data expansion, methods of combating class imbalance, convolution, pooling operations, convolutional neural networks, optimization of neural networks.

Publications. As a result of the work done 1 scientific articles were published:

1. *Nerus V. M.* A neural network for detecting the chest abnormalities on X-ray images // *International scientific journal "Internauka"*. — 2022. — No. 14. <https://doi.org/10.25313/2520-2057-2022-14-8414>

Keywords. neural network, CNN, X-ray, CT imaging, computer vision, augmentation.

ЗМІСТ

ВСТУП.....	7
РОЗДІЛ 1 ОГЛЯД ЛІТЕРАТУРИ.....	10
1.1. Огляд існуючих рішень проблем	10
Висновки до розділу 1	28
РОЗДІЛ 2 ТЕОРЕТИЧНА ЧАСТИНА.....	30
2.1. Опис проблем мануальної діагностики торакальної патології клітини	30
2.2. Методи діагностики патології органів грудної клітки.	31
2.3. Проблеми задачі класифікації рентгенівських зображень	34
2.4. Штучні нейронні мережі	36
2.5. Згорткові нейронні мережі	39
2.6. Додаткові операції в згорткових нейронних мережах	46
2.6.2. Функції активації	49
2.6.3. Оптимізатори нейронних мереж	54
2.6.4. Механізм уваги в сучасних нейронних мережах	57
Висновки до розділу 2	60
РОЗДІЛ 3 ПРАКТИЧНА ЧАСТИНА.....	61
3.1. Опис використовуваних даних	61
3.2. Вибір мови програмування	64
3.3. Порівняння Pytorch та Tensorflow	65
3.4. Аугментація даних	69
3.5. EfficientNet як основна моделі	72
3.6. Боротьба з дисбалансом класів	75
3.7. Тренування моделі.	77
3.8. Тестування моделі.	79
Висновки до розділу 3	80
РОЗДІЛ 4 РОЗРАХУНОК СТАРТАП ПРОЕКТУ.....	82
4.1. Опис ідеї проекту	82
4.2. Організація проекту	83
4.3. Канва бізнес-проекту	83
4.4. Ключові види діяльності	84

4.5. Ціннісні пропозиції та споживачі	86
4.6. Взаємовідносини зі споживачами та канали збуту	88
4.7. Обґрунтування ресурсів та витрат проекту	89
4.8. Грошовий потік та оцінка вартості проекту	92
ЗАГАЛЬНІ ВИСНОВКИ.....	95
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	96

ВСТУП

В даний час розвиток технологій штучного інтелекту та машинного навчання, а також накопичення великих обсягів медичних зображень відкривають можливість створення систем автоматичного аналізу медичних зображень.

Рентген грудної клітки – це одне з найпростіших і найпоширеніших обстежень, що застосовуються зараз. Його автоматизація значно знизить навантаження на рентгенологів, дозволить оцінювати якість їх роботи та знизить ймовірність лікарських помилок.

При медичному дослідженні легень за допомогою рентгенограм грудної клітини можна виділити достатньо велику кількість таких патологій як рак, пневмоторакс, пневмонія, фіброз, емфізема та туберкульоз. Головним захворюванням, через яке люди проходять регулярні флюорографічні дослідження, є туберкульоз. Він являється дуже заразним і небезпечним захворюванням, тому проведення рентгенологічного дослідження грудної клітки, є обов'язковим і щорічною процедурою для кожної людини у багатьох країнах [33].

За останні роки обсяг і складність інформації, яка надається радіологам для звітування, надзвичайно зросла. Враховуючи складність новітніх методів візуалізації, зокрема КТ та МРТ, зараз інтерпретація клінічних зображень займає більше часу, ніж процес їх отримання[43].

Робоче навантаження може бути фактором збільшення ймовірності помилок у радіологічних звітах [18]. Різноманітні дослідження показали, що більшість аномальних знахідок на звичайних рентгенограмах виявляються протягом перших кількох секунд пошуку зображення, при цьому кількість істинно позитивних знахідок різко зменшується через короткий час. Крім того, нещодавнє дослідження зорової акомодатії та продуктивності рентгенологів показало, що здатність зосереджуватися та виявляти переломи зменшується наприкінці робочого дня [28]. Довший робочий день може лише посилити це

зниження продуктивності, а отже, і безпеки.

Нещодавнє національне дослідження робочого навантаження консультантів-рентгенологів в Ірландії показало, що в 2009 році середній ірландський радіолог виконував 128% робочого навантаження, яке вважалося відповідним як контрольний показник, виміряний в Австралії [17, 8] . Зростаюча кількість і складність візуалізаційних досліджень вимагає відповідного збільшення робочої сили радіологів.

Тому існує зацікавленість у розробці комп'ютерних алгоритмів для пошуку аномалій на рентгенограмі, які згодом стануть допоміжним інструментом на допомогу радіологам. Одним із методів машинного навчання є нейронні мережі.

Нейронні мережі є універсальними системами інтерполяції функцій. Нейронна мережа з оптимальною архітектурою здатна апроксимувати функцію виділення патологій та перекласти завдання формалізації ознак патологій з людини на нейронну мережу, що спрощує класифікацію знімків грудної клітки.

Метою даної роботи є застосування методів машинного навчання в області пошуку патологій на рентгенограмах легень і проведення відповідних досліджень. Для досягнення цієї мети були поставлені такі *завдання*:

1. Провести підготовку даних для машинного навчання, розробити методи попередньої обробки знімків, що полегшують процес навчання.
2. Розробити моделі машинного навчання: побудувати найбільш оптимальну архітектуру моделей та реалізувати їх.
3. Провести експерименти з реалізованими моделями машинного навчання: навчити моделі, вибрати метрику для оцінки ефективності моделей, протестувати моделі та порівняти результати роботи моделей, використовуючи вибрану метрику.
4. Розрахувати стартап-проект для виведення стартапу на ринок.
5. Розробити висновки за результатами магістерського дослідження

Об'єктом дослідження в даній роботі є рентгенограми з публічного датасету.

Предметом даного дослідження виступають сучасні методи аугментації даних та архітектури нейронних мереж .

Методами дослідження в даній роботі являються: методи розширення даних, методи боротьби з дисбалансом класів, операції згортки, пулінгу, згорткові нейронні мережі, оптимізація нейронних мереж.

Наукова новизна

Полягає в запропонованій комбінації сучасних технік таких як аугментація та що ще більш важливо збалансування класів в поєднанні з сучасними архітектурами нейронних мереж

Практична значимість

Пов'язана з автоматизацією оцінок рентгенологічних зображень.

Публікації. За результатами виконаної роботи було опубліковано 1 наукову статтю:

1. Нерус В. М. Використання нейронних мереж для діагностики захворювань грудної клітини на рентгенограмах // Міжнародний науковий журнал "Інтернаука". — 2022. — №14.
<https://doi.org/10.25313/2520-2057-2022-14-841> 4

Структура дисертації

Магістерська дисертація за темою «Нейронна мережа для виявлення відхилень грудної клітини на рентгенівських знімках» виконана студентком Нерусом Владиславом Миколайовичом зі спеціальності 122 «Комп'ютерні науки» за освітньо-професійною програмою «Комп'ютерні технології в біології та медицині» побудована за класичним типом та викладена на 105 сторінках. Вона складається : вступу; 4 розділів (літературний огляд, теоретична частина, практична частина, аналіз стартап-проекту), висновків до кожного з цих розділів; загальних висновків; списку використаних джерел, який налічує 6 джерела та додатків (9 – на кирилиці, 59 – на латиниці). В дисертації представлено 35 рисунків і 15 таблиць.

РОЗДІЛ 1

ОГЛЯД ЛІТЕРАТУРИ

1.1. Огляд існуючих рішень проблем

Перші діагностичні системи для легеневих захворювань почали з'являтися в кінці 20 століття, але вони не були дуже успішними. Головною проблемою того часу була нестача комп'ютерних потужностей для запуску алгоритмів, необхідних для швидкої обробки зображень. Загалом захворювання легенів можна виявити за допомогою основних методів обробки зображень, але це було обчислювально тяжко в минулому. Однак за останні 20-30 років галузь графічних та звичайних процесорів швидко розвивалася, що стало однією з причин для розвитку нейронних мереж, тому автоматична діагностика стала більш поширеною і простою в обчислювальному і технологічному плані. Згорткові нейронні мережі краще справляються з такими завданнями, як аналіз рентгенівських зображень, ніж попередні нейронні мережі, не потребуючи допомоги з боку людей.

Оскільки COVID-19 став проблемою, багато компаній, які змайються інтелектуальним аналізом даних взялися за розроблення рішення на основі ШІ для діагностики COVID-19 на рентгенівських знімках грудної клітки, як наприклад Google [29].

Звичайно, такі інструменти використовують методи глибинного навчання штучних нейронних мереж, що досить успішно використовується для класифікації та розподілу за типом захворювання, пов'язані зі специфічною інфекцією або відсотком ураження легень, або спостереження за прогресом захворювання.

На сьогоднішній день існують різні типи нейронних мереж, за допомогою яких можна досліджувати та вирішувати різні проблеми. Типові проблеми які вирішують нейронні мережі: класифікація захворювань за знімком (можуть вказати тип захворювання), вказати на тяжкість

захворювання, зробити детекцію враженої частини органа, та проводити моніторинг хвороби в реальному часі.

Штучний інтелект може допомогти зменшити навантаження на радіологів, визначаючи ознаки та симптоми коронавірусної інфекції іншими методами. В деяких випадках, алгоритмічна оцінка зображення знижує необхідність застосування нуклеїнових тестів. Згідно із цією статтею «Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning», глибоке навчання було використано для аналізу рентгенівських знімків з метою діагностики туберкульозу та інших захворювань. Дана модель досягла досить високої точності класифікації під 93%[27].

У схожому дослідженні де були використані нейронні мережі на основі згорток, якщо говорити точніше, то розробники використовували чотири згортокові шари після чого два повнозв'язних шара в поєднанні із сучасними підходами в датааугментації та методів оптимізації досягли точності в більше ніж 93% в задачі класифікації пневмонії на основі нейронних мереж[51].

Глибока згорточна нейронна мережа (CNN) була створена групою авторів з Корнельського університету, яка використовується для автоматичної діагностики COVID-19 за допомогою попередньо навчених моделей. У дослідженні використовувалася сучасна модель глибокого навчання під назвою ResNet50, а потім використала техніку під назвою «трансферне навчання», щоб ще більше покращити її продуктивність. Це призвело до дивовижної точності 96%. [35].

Таким чином, можна зробити висновок, що сучасні глибокі нейронні мережі здатні вивчати та запам'ятовувати багато інформації. Мережі можуть точно класифікувати захворювання легенів за допомогою архітектур зі згортовими шарами та досягати точності у більше ніж 90% .

Для наочності, лікарі, які відвідували дослідження Лондонського коледжу радіології, виявили, що відсоток випадків респіраторних захворювань, які вони правильно ідентифікували, класифікували на рентгенівських знімках, у середньому становив близько 82,6%[47].

Сукупність технологій, які використовуються в глибоких нейронних мережах, уже є досить точними, хоча вона все ще знаходиться на стадії вивчення. Незважаючи на те, що моделі штучного інтелекту в деяких випадках дають більш точні результати, ніж радіологи, коли справа доходить до класифікації зображень, ще занадто рано сподіватися, що всі радіологи будуть замінені моделями ШІ. Основною перешкодою для цього є велика кількість необхідних етапів регулювання та сертифікації, а також стандартизований формат впровадження таких систем і якість зображення для широкого спектру продукції рентгенівського обладнання.

Щоб прийняти правильне рішення, важливо взяти до уваги ймовірність того, що група експертів запропонує краще рішення, ніж будь-хто з них самотійно. Тому точність загальної класифікації лікаря і моделі вища, ніж точність аналогічної класифікації їх окремо. Враховуючи всі ці фактори, можна вважати, що зараз краще не використовувати повністю автоматизовані системи, а використовувати їх як частину системи підтримки прийняття рішень, щоб допомогти лікарям приймати рішення. Лікар може використовувати модель, щоб зменшити кількість помилок. Класифікація захворювань у спосіб, який не враховує різні типи помилок, і аналізу причин які їх викликають, може призвести до упущених захворювань. Однією з причин, чому лікарі роблять менше помилок на практиці, є те, що вони намагаються уникати помилок першого роду[12].

Помилки 1 го типу, які можуть мати неприємні наслідки, наприклад пацієнти можуть лікуватися від впневмонії, коли вони її насправді не мають, і це може означати, що лікарні в кінцевому підсумку будуть переповнені людьми, яким не потрібно бути там, і це може коштувати багато грошей . Згідно з дослідженнями[13], більшість пацієнтів, які звертаються до лікарень, насправді не мають жодних патологій. Автоматизована система класифікації, заснована на використанні штучного інтелекту, допоможе відсіювати цих пацієнтів і точніше і швидше ідентифікувати тих пацієнтів, які потребують лікування або навіть госпіталізації.

На даний момент існує багато різних підходів до проблеми автоматичного аналізу рентгенологічних зображень. Ван запропонував покращити структуру шляхом додаткової інтеграції перехідні шари в CNN. Його мережа показує хороші результати великий набір даних ChestX-ray 14 і є однію із фундаментальних перших робіт[6]. П. Раджпуркар пропонує використовувати 121-шаровий згортковий нейрон мережа DenseNet-121 як основа, поєднана з нелінійною функцією активація нейронів. Ця мережа отримала назву CheXNet. В діагностиці пневмонії, CheXNet успішно перевищив середній показник лікарів рентгенологів)[58]. (див. рис. 1.1).

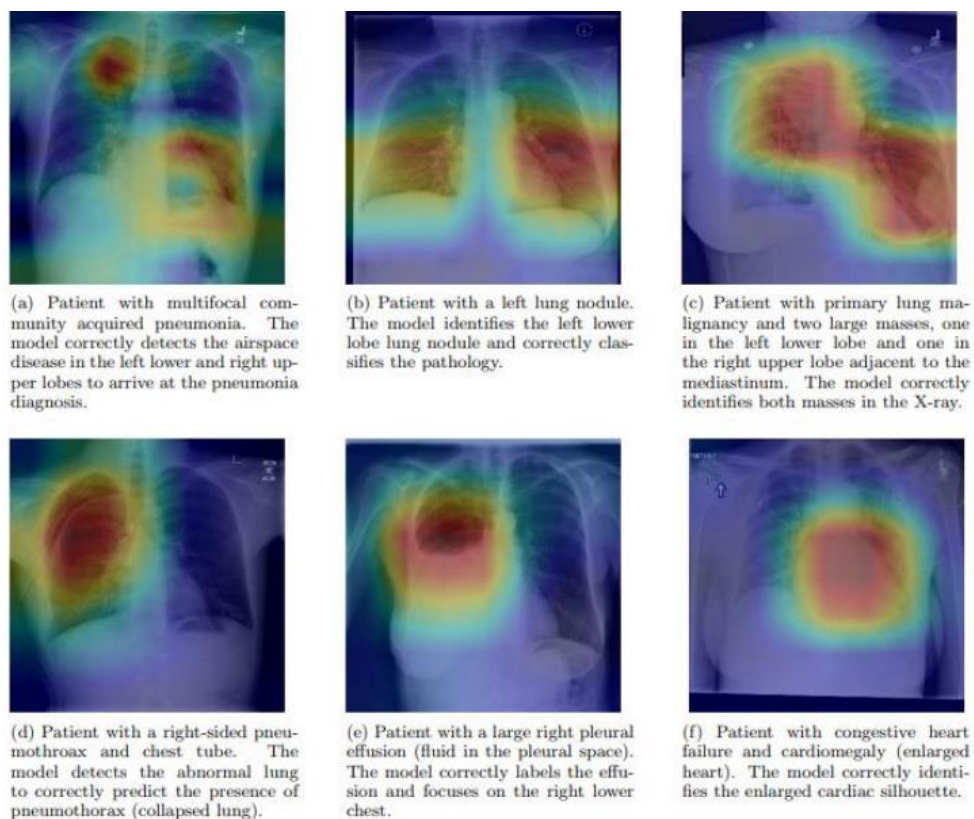


Рисунок 1.1 Результат роботи CheXNet [58]

CheXNet локалізує патології, за допомогою яких він ідентифікує карти активації класу, які підсвічують області рентгенівського зображення, які найсильніше впливають на вибір нейронної мережі для конкретного патології. Йо пропонує об'єднати CNN з рекурентними нейронними мережами (RNN) для діагностики органів грудної клітки з урахуванням залежностей між

різними патологіями [26]. Його підхід проілюстровано на рис 1.2.

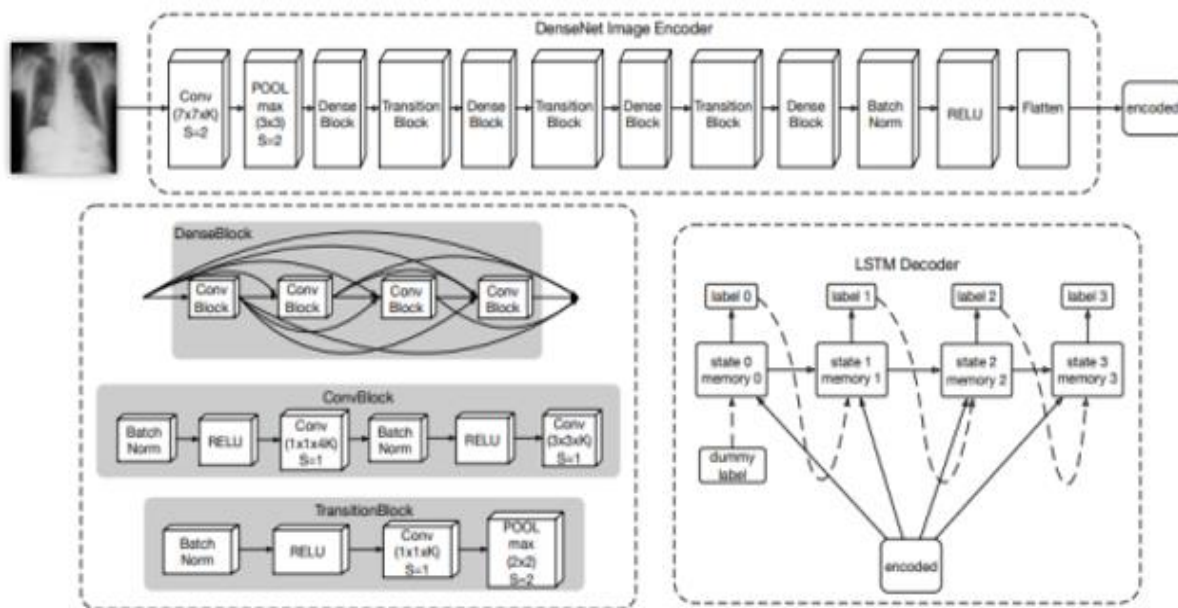


Рисунок 1.2 Архітектура запропонована Яо [26]

Ліет ін. використовував попередньо навчений ResNet(див. рис. 1.3) для вилучення окремих частин зображення, які потім подаються в згорткову мережу, з метою отримати карту ймовірностей для кожної патології. Вхідне зображення в його мережі спочатку обробляється CNN. Далі створюється «Шар нарізки» – базові пакети. Пакети передаються до згорткової мережі розпізнавання для остаточної класифікації (див. рис. 1.4) [9].

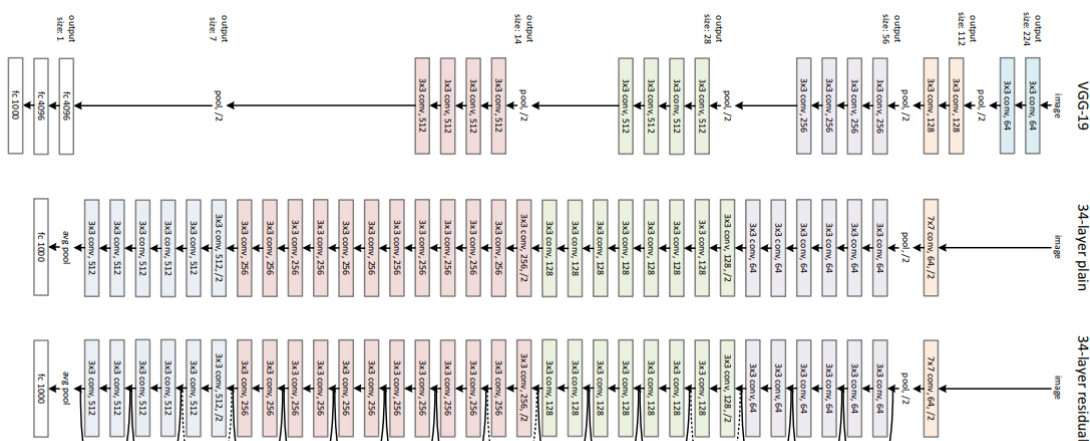


Рисунок 1.3 Базова архітектура ResNet

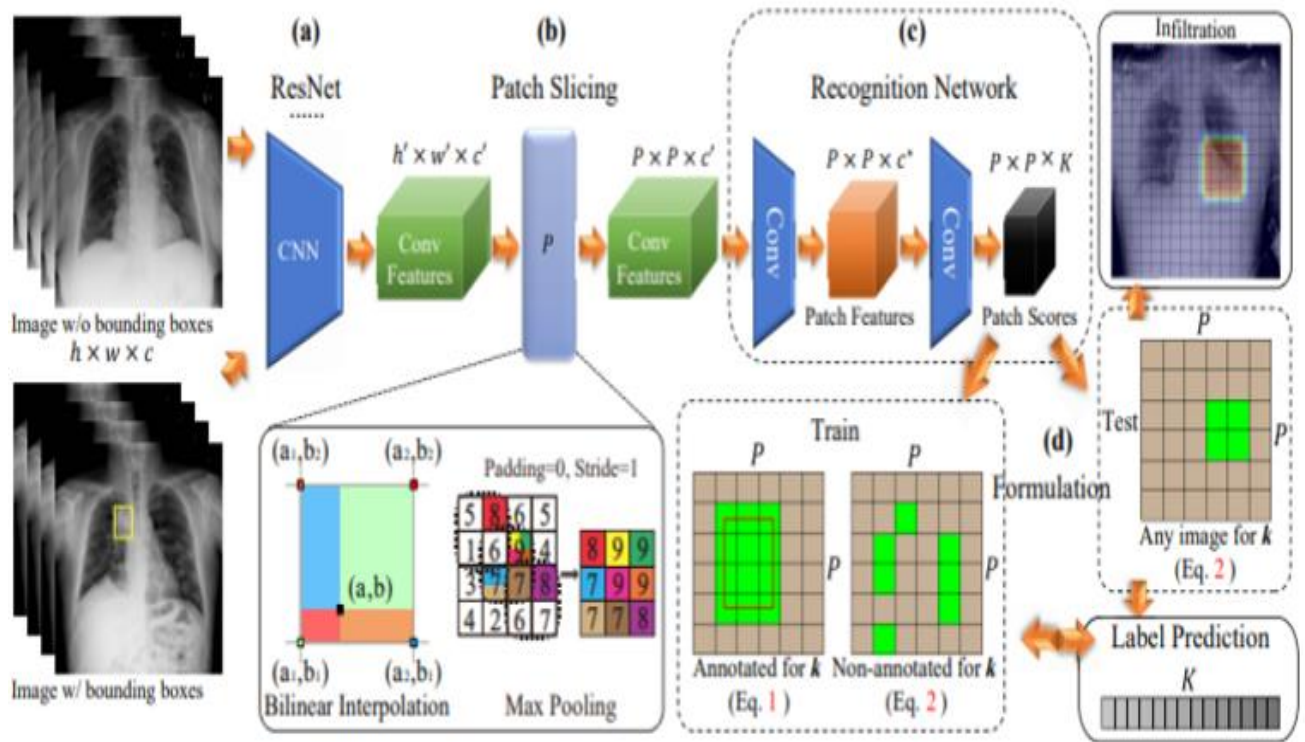


Рисунок 1.4 Архітектура мережі запропонована Ліетом[9]

Чжічен Хуан і Донгмей Фу пропонують згорткову нейронну мережу з множинною увагою для виявлення патологій грудної клітки на рентгенівських знімках. Архітектура запропонована Хуаном і Донгмей Фу зображена трохи нижче, на якому видно (див. рис. 1.5) [10].

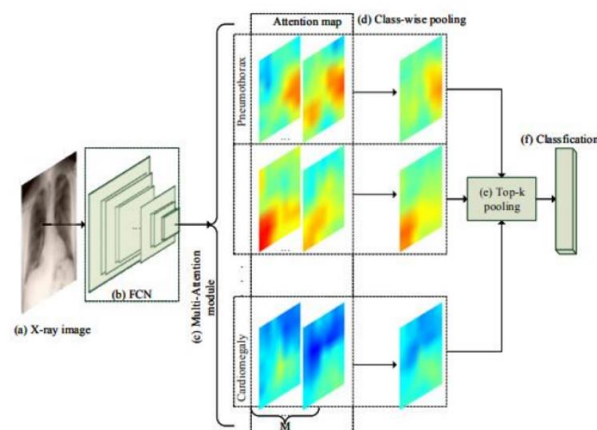


Рисунок 1.5 Архітектура мережі Хуана Чжічена[10]

Нарешті, Ян та інші запропонували комбіновані компресійні блоки паралельно з використанням DenseNet для покращення продуктивності моделі.

(Рис. 1.6). Крім того, вони також використовували передачу карт функцій для локалізації і виділення підозрілих ділянок [19].

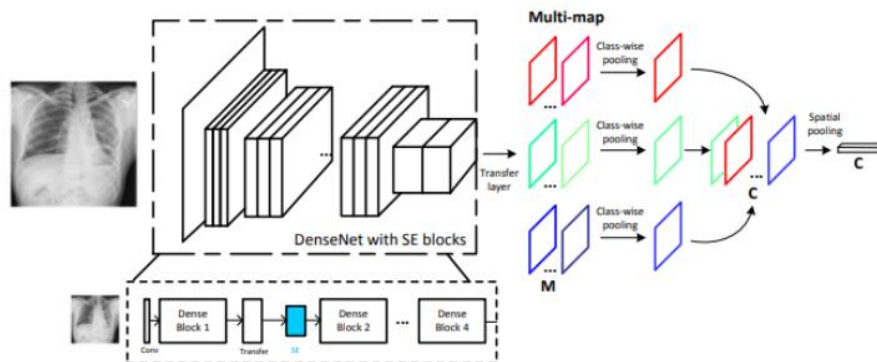


Рисунок 1.6 Архітектура мережі запропонована Яном[19]

1.2. Мережі з механізмом уваги до патологій

Для вирішення задач класифікації рентгенограм органів грудної клітки, були запропоновані нейронні мережі CONAF і RAMAF розроблені в Королівському коледжі Лондона. Як правило, в таких задачах використовується класифікація на два класи об'єктів: зображення, що містять ознаки патологій і зображення без ознак патології. Наразі для збереження рентгенограм у цифровому вигляді використовують DICOM[36].

Переваги формату DICOM:

1. висока якість зображення та підтримка 65536 відтінків сірого;
2. є загальновизнаним промисловим стандартом.;
3. формат забезпечує можливість запису не тільки інформації про зображення, а також великий обсяг пов'язаної інформації, від параметрів пацієнта до характеристик обладнання [36].

У цій частині розділа описано подання класифікатора зображень на основі згорткових нейронних мереж (CONAF (рис. 1.7), Convolution networks зі зворотним зв'язком уваги). Класифікатор був розроблений в Лондонський коледжі [39]. В ході роботи була поставлена проблема розпізнавання

рентгенограм, які, швидше за все, містять одне або більше уражень патології. Хоча локалізація уражень на знімку не була основною задачею, цю інформацію можна отримати від навченої нейронної мережі для створення карти значущості, наприклад карти температури, показуючи, де на початковому етапі найімовірніше знаходяться ураження. Розглянута архітектура використовує ці карти для виведення механізму програмної уваги [39]. Для рентгенограм, що містять анотації, карти довіри можна порівнювати з апіорними даними (поле, виділене радіологом) для обчислення похибки локалізації.

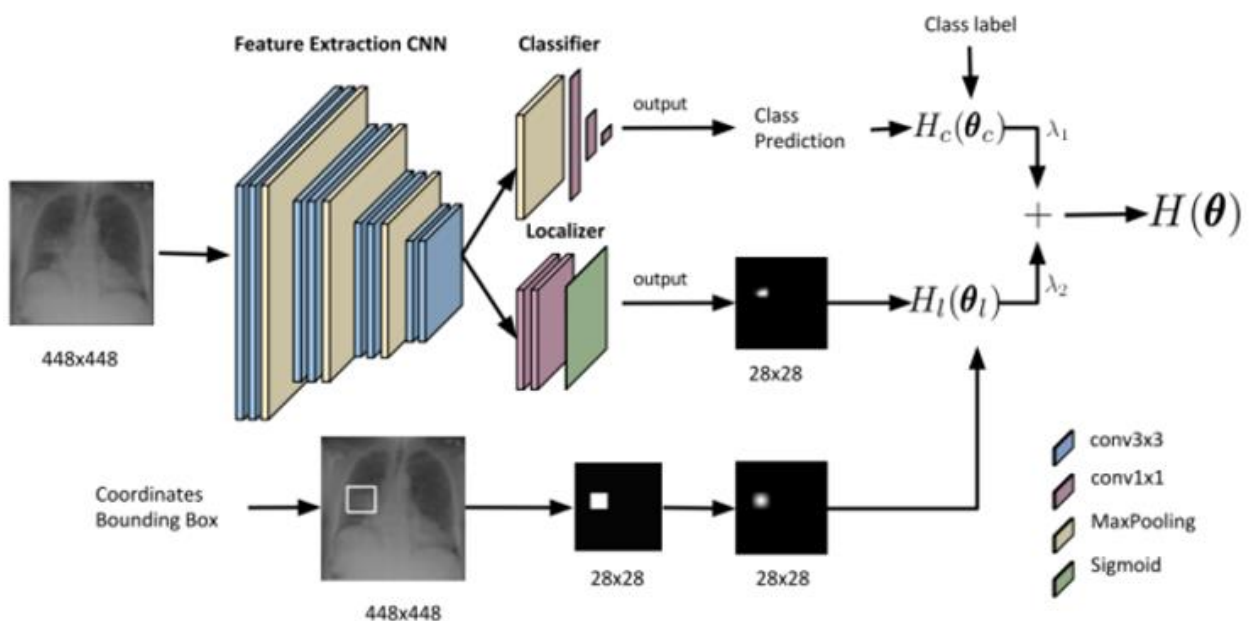


Рисунок 1.7 Архітектура мережі CONAF [39]

Ідея уваги в глибоких нейронних мережах натхненна системою зорової уваги людини. Просторова увага дозволяє людям вибірково обробляти візуальну інформацію через визначення пріоритетів області в полі зору і значно покращити як розпізнавання, так і виявлення, особливо на зображеннях із додатковим шумом. Дотримуючись того самого принципу, нейронні мережі можна навчити зосереджуватися на певних частинах вхідного сигналу, які, як видається, тісно пов'язані з поставленим завданням. Конволюційна мережа

використовується для вивчення особливостей зображення, які мінімізують помилку класифікації, а карти помітності виводяться у слабо контрольований спосіб як частина процесу навчання. Карта помітності виділяє частини зображення, які, ймовірно, будуть пов'язані з класом, передбаченим мережею. Автори пропонують гібридний підхід до навчання з використанням функції втрат, яка, окрім штрафування за помилку класифікації, також штрафує за показник невідповідності між неявним розташуванням патології в мережі, представленим картою помітності, отриманою під час навчання, та реальним положенням патології, коли та відома. Велика помилка вказує на те, що поточне представлення мережі не точно описує візуальні моделі патології та забезпечує додатковий механізм для самовдосконалення через зворотне поширення помилки. Отримана архітектура, згорточна нейронна мережа зі зворотним зв'язком уваги (CONAF), гарантує, що представлення зображення отриманий мережею, оптимізує ефективність як класифікації, так і локалізації.

Рекурентна модель уваги зі зворотним зв'язком (RAMAF). Другий підхід реалізує механізм жорсткої уваги, у якому кожна рентгенограма обробляється за кінцеву кількість послідовних кроків. На відміну від попереднього методу, на кожному кроці як вхідні дані використовується лише частина зображення. На кожному кроці часу алгоритм вибирає наступне місце для відвідування з розподілу ймовірностей, який обчислюється на основі інформації, отриманої моделлю на всіх попередніх кроках. Уся інформація, накопичена випадковим шляхом по зображенню, зрештою використовується для класифікації зображення. Оцінка класифікації використовується як сигнал винагороди для оновлення розподілу ймовірностей, що контролює послідовність місць зображення, які слід відвідати, щоб більше уваги врешті приділялося відповідним частинам зображень, таким як легені. Запропонована авторами архітектура RAMAF (модель повторюваної уваги зі зворотним зв'язком уваги) призначає вищі винагороди щоразу, коли шлях, відвіданий алгоритмом під час навчання, збігається з правильним розташуванням вузлика, якщо це доступно. Ця стратегія винагороди змушує модель переглядати області зображення, які,

ймовірно, містять ділянки вузликів, з більшою швидкістю, що призводить до вищої швидкості конвергенції та підвищення ефективності класифікації. Удосконалена модель отримала назву RAMAF (рис. 1.8.) (повторна увага модель з анотаційним зворотним зв'язком) [39]. Модель працює методом обробки зображення на частини, одну за одну, і вчиться рухатися візуалізації шляхом виконання обстежень стратегічно обраних територій. Після кожного кроку алгоритм перевіряє більшу частину зображення, а оптимальна політика, яка диктує «де шукати», що мінімізує помилку класифікації.

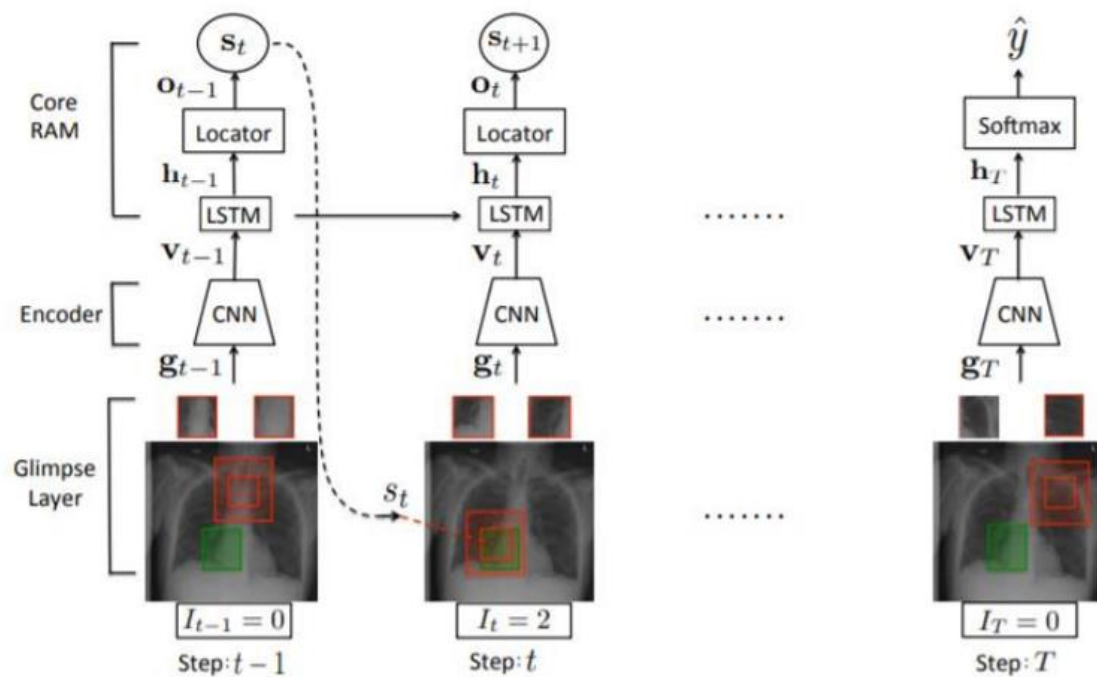


Рисунок 1.8 Архітектура мережі RAMAF [39]

Як видно з малюнка, після первинного відбору монітор фрагмент постачається з шаром Encoder, який витягує інформацію. Функція цього шару є його стійкість до шуму. У цьому додатку виникає складне візуальне середовище, що характеризується високою варіативністю як яскравість, так і складність об'єкта. Це пов'язано з мінливістю анатомічні особливості пацієнтів. Далі кожен контрольований фрагмент пропускається через два згорткові нейронні мережі з наступними операціями підвибірки.

1.3. Фільтрація кісткової тканини в мережах

Проблему автоматичного виявлення та подальшої фільтрації кісткової тканини ребер на рентгенівських знімках розглядається вже досить давно. Незважаючи на це, фільтрації не приділено багато уваги, хоча застосування цього методу має велике значення в клінічній практиці і являється досить корисним при діагностиці легеневих захворювань. Крім того, використання цієї техніки в кінцевому підсумку може підвищити точність класифікації легень на 25-30% [5].

Всі існуючі алгоритми фільтрації кісткової тканини всередині груднини складаються з кількох послідовних ступенів. Вивчивши кілька таких алгоритмів [5, 1, 15], був обраний алгоритм, описаний у [5]. Першим критерієм відбору було той факт, що за відсутності будь-яких попередньо позначених даних, було необхідно, щоб алгоритм використовував принцип навчання без учителя. Цьому критерію відповідали два зі знайдених алгоритми, описані в [15] і [1].

Найбільшу складність фільтрації кісткової тканини викликає стадія визначення різниці між інтенсивністю текстури кісткової та легеневої тканин. Цей крок необхідний для безпосереднього видалення силуетів кісток шляхом мінімізації різниці між інтенсивностями. Для обчислення інтенсивності текстури використовується лінійна модель: інтенсивність краю в точці (x, y) визначається як $ax + by$. Фактично задача визначення інтенсивностей зводиться до задачі оптимізації вибору найкращих параметрів для моделі (a, b) . Для вирішення цієї проблеми порівнювані алгоритми використовують різні підходи.

Отже, в [1] автори пропонують використовувати спеціальний генетичний алгоритм (RCGA), заснований на тому, що в генетичні алгоритми є вбудованими принципи, запозичені з біології та генетики, що, у свою чергу, дозволить досягти хороших результатів. Проте запропонований алгоритм авторів має кілька недоліків. По-перше, при розрахунку відмінностей

інтенсивності однорідність текстури кістки може бути спотворена. Крім того, додається висока обчислювальна складність генетичних алгоритмів. На це вказують автори [1] і запропонували алгоритм, заснований на регресійній моделі, яка враховує однорідність кісток, при цьому витрачаючи менше часу на розрахунки. Алгоритм, описаний у [1], працює в припущенні, що величина різниці між інтенсивністю кісткової та легеневої тканин може змінюватися від центру кістки до її меж, однак на сусідніх ділянках уздовж параболічної кривої ця величина залишається практично незмінною. Це припущення дозволяє методу спочатку передбачити різницю інтенсивності в центрі кістки, а потім розширити це значення у вертикальному напрямку, щоб мінімізувати різницю між краями ребер і легеневою тканиною вище або нижче кордонів. Ось короткий опис основних кроків цього алгоритму:

1. Вилучення шаблону

Першим кроком у пошуку ребра є вирізання шаблону ребра із вмісту знімка. Спочатку виконується першочергове виділення країв ребра за допомогою певного оператора детектування країв, після чого до зображення застосовується проста схема вибірки для фіксації шаблону, що знаходиться в центрі області легені. За допомогою певної оцінки, вибирається кілька відповідних шаблонів-кандидатів для кожної легені окремо.

2. Знаходження крайніх точок меж ребра

Для пошуку меж країв використовується логарифмічний фільтр Габора (фільтр Лога-Габора) [57]. Цей фільтр часто використовується розпізнавати межі об'єктів на зображенні. На цьому кроку виконується первинне виділення ребер, починаючи з «початкових точок». «Початкові точки» розташовані там, де краї перетинаються із зовнішньою межею легень. Варто зазначити, що на цьому етапі структура краю частково виділена.

3. Вибір найкращого шаблону

Цей крок зменшує потенційну кількість відповідних шаблонів для правої та лівої легені, обчисливши відстань між межами ребер на шаблоні та структурами, знайденими на попередньому кроці. Шаблон з найменшою

відстанню вибирається для подальшої роботи системи

4. Оцінка ширини ребра

Щоб уможливити подальший вибір ребра, обчислюється середня висота (або ширина) краю. Основна мета цього кроку - знайти найдовшу можливу вертикальну лінію, які можна вписати в крайову структуру. Встановивши праву та ліву межі допустимої ширини краю, обчислення середнього значення на основі всіх дійсних вертикальних ліній, знайдених у структурі ребер.

5. Регулювання параболи

На цьому кроці виправляється вертикальне положення шаблону. Зрушення виконується шляхом оптимізації коефіцієнта при вільному члені полінома $p(x)$, заданого на кроці 3. Залишаючи всі інші коефіцієнти без змін, p_{015} обчислюються значення $p(n+1)$ для кожного ребра. Таким чином, знайдені значення дозволяють визначити кінцеве положення шаблону.

6. Розширення крайових меж

Кожне ребро всередині грудної клітини має певні особливості в своїй геометрії, а отже, і глобальний шаблон у формі параболи, що виходить із знайденої «початкової точки», може не вписатися в структуру реберної кривизни. Цю проблему можна вирішити кількома способами. Автори пропонують власний підхід на основі уточнення положення центру ребра шляхом перевизначення коефіцієнтів того самого полінома $p(x)$, характерного для ребра. Це розширить край лінійно, досягнувши передбачуваної висоти ребра. Отже, цей крок дозволяє більш повно виділити поверхню краю.

7. Фільтрація кісткової тканини

Завершальним етапом методу є усунення з рентгенівського зображення шумів, викликаних наявністю кісткової тканини ребер. Ця операція виконується шляхом придушення силуетів ребра зі збереженням інформації про розташовану під ними легенева тканину. Придушення починається з пікселів від вертикального центру ребра та адаптивно розширюється до країв ребра. По суті, необхідно визначити інтенсивність ураження легеневої тканини під ребром [1].

1.4. Попередня обробка зображень

Щоб перетворити зображення в більш просту та ефективну форму необхідна попередня обробка. Це зменшує шуми на зображенні та допомагає визначити фокусні зони. Адаптивний підхід вирівнювання гистограми з обмеженим контрастом використовується для покращення місцевого контрасту. Встановлення порогів використовується для зменшення шуму на зображенні. Пікселі, які знаходяться нижче порогу на гистограмі яскравості зображення вважаються шумом. Для покращення використовується фонові корекція та фільтрація порогових результатів. Корекція фону мінімізує вплив різних умов освітлення, а фільтр зменшує випадковий шум, присутній на зображенні. Згадані раніше підходи зазвичай використовується для підвищення якості цифрових зображень застосування в машинному навчанні[4].

Для детального вивчення систем САПР, найбільш часто використовуваних стратегій попередня обробка - це розширення даних і нормалізація кольору. Із зростанням обсягу даних навчальний набір даних покращується на за допомогою різних методів трансформації зображення, таких як масштабування, перевертання, дзеркальне відображення, розмивання та додавання шуму. Ці перетворення змінюють морфологію зображення. Крім того, гістопатологічні зображення відрізняються за кольором і освітленням через неузгодженість у підготовці гістопатологічних слайдів (ряба розкраска) і різні умови освітлення під час цифрового захоплення зображення. Тому важливий етап попередньої обробки патогістологічне зображення - нормалізація кольору плям. У цьому методі вибирається еталонне гістопатологічн зображення з ідеальним замалюванням пятна.

Значення кольорів регулюються таким чином, щоб вони відповідали значенням кольорів цільового зображення. Деякі дослідження показують, що точкова нормалізація покращує продуктивність класифікатора на основі

глибокого навчання. Дослідники використовуються нормалізацію кольору для розробки систем САПР на основі ретельного дослідження гістопатологічних зображень раку молочної залози. Порівняння нещодавно методи, що використовуються для нормалізації кольору плями для гістопатології зображення представлені в дослідженні Roy et al. Деякі дослідження свідчать про дуже ефективні методи нормалізації кольору в цифрових зображеннях. Наприклад, дослідники запропонували метод нормалізації кольору, який перетворює зображення RGB до колірному простору на основі сприйняття. Метод передає колір між вихідним і цільовим зображеннями шляхом обчислення середнього значення та стандартного відхилення для кожної осі зображення[4].

1.5. Покращення зображень за допомогою фільтрів

Основне призначення цих фільтрів — зменшення шуму, а також фільтра можна використовувати для підкреслення певних рис на зображенні. Для видалення шуму можна використовувати адаптивний фільтр Адаптивний медіанний фільтр виконує просторову обробку, щоб визначити, які пікселі зображення постраждали імпульсні перешкоди[22] (рис. 1.9.).

Медичні образи часто забруднені імпульсивними, адитивними або мультиплікативними шумами через низку неідеальностей у процесі формування зображення. Шум зазвичай спотворює медичне зображення шляхом заміни деяких пікселів вихідного зображення з новими пікселями зі значеннями яскравості близько або дорівнює мінімуму чи максимуму допустимого динамічного діапазон яскравості. Ідентифікація типу шуму в медичному зображенні здійснюється в два етапи. На першому етапі використовується критерій для визначення наявності імпульсних шумів. Якщо результат цього критерію негативний, зображення потім передається на другий етап, щоб ідентифікувати адитивний або мультиплікативний характер шуму.

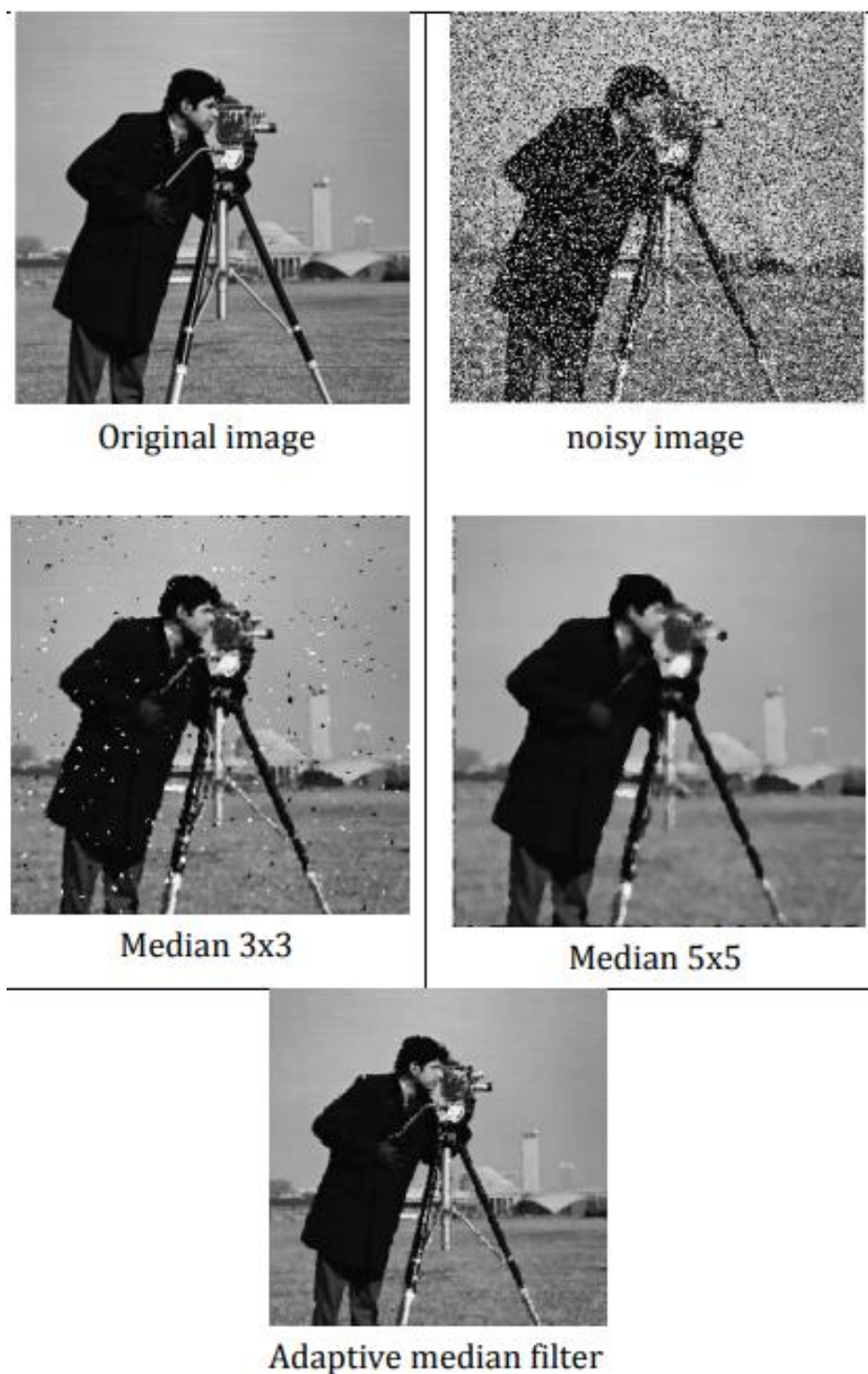


Рисунок 1.9 Застосування медіанного фільтра [22]

Адаптивний медіанний фільтр (AMF) класифікує пікселів як шум, порівнюючи кожен піксель на зображенні з його оточенням сусідні пікселі. Розмір підлягає коригуванню, як і поріг для порівняння. Піксель, який відрізняється від більшості своїх сусідів, а також структурно не відповідає пікселям, як виглядає, позначено як імпульсний шум. Потім ці шумові пікселі

замінюються середнім значенням піксель для пікселів у середовищі, які пройшли перевірку шумове маркування. Спочатку йде перетворення зображення в градації сірого за допомогою спеціальної функції, потім застосовуючи адаптивну середню фільтрацію до отриманого зображення. Щоб покращити контраст для виділення області, в обробці зображень використовують метод покращення балансового контрасту (BCET) див. рис. 1.10 [24].



Рисунок 1.10 Застосування фільтра BCET[24]

Зазвичай при обробці медичного зображення, необхідне посилення контрастності для зони інтересу. Наприклад, Contrast Limited Adaptive, вирівнювання гистограми (CLAHE) дозволяє “покращити” зображення та отримати найкращі характеристики медичних зображень для правильної діагностики (рис. 1.11) [55].

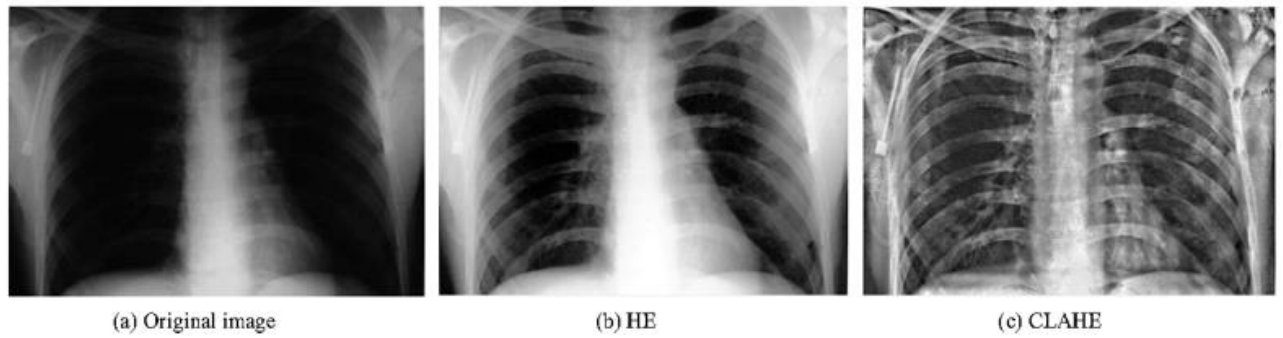


Рисунок 1.11 Застосування фільтра CLAHE[55]

1.6. Методи витягування ознак на основі глибокого навчання та їх обмеженість

Вилучення ознак на основі глибокого навчання викликало значний інтерес серед дослідників до класифікації гістопатологічних зображення. Про це свідчить дослідження [11]. Автори в [37] виконали декілька експерименти з набором гістопатологічних даних раку молочної залози ("BrecaKHis") за допомогою функцій LBP, CLBP, LPQ, ORB, GLCM і PFTAS. Пізніше ці автори реалізували модель, засновану на глибокому навчанні для класифікація набору даних BrecaKHis. Модель досягла вищою точність у порівнянні з моделями класифікації на основі особливості LBP, CLBP, LPQ, ORB і GLCM. Нещодавно Convolutional Neural Network мережа (CNN), заснована на багатошаровому навчанні, досягла найвищого значення P_{rr} для набору даних BrecaKHis. Згадані раніше дослідження базуються на вивченні моделі CNN з нуля. Однак для цього типу дослідження потрібен великий набір даних для підвищення ефективності навчання. Отримання такого великого набору даних з місцевих лікарень важко через такі фактори, як асоціація з патологічними лабораторіями, дозволи на доступ до даних пацієнтів та авторизація результатів діагностики. Незважаючи на ці труднощі, в дослідження [13] представило набір гістопатологічних даних раку молочної залози під назвою BrecaKHis, яка стала загальнодоступною для подальшого дослідження. Але цей набір даних відносно невеликий і надає обмежені зразки для вивчення

порівняно з великими наборами даних, такі як ImageNet. Причому навчання з нуля вимагає багато зусиль для точного налаштування параметрів CNN і ще більш тривалого часу для навчання в порівнянні з раніше вивченими моделями CNN.

Гістопатологічні зображення мають спільні риси, такі як лінії, краї та кольори із зображеннями з набору даних ImageNet. Тому таким чином, CNN, раніше вивчені на ImageNet, можуть бути використані для визначення особливостей гістопатологічних зображень. Таким чином, дослідники зацікавлені у використанні попередньо навчених CNN замість CNN, навчених з нуля. Дослідження [2] демонструє, що попередньо навчені CNN є хорошою заміною для CNN, навчених з нуля для діагностики раку молочної залози за допомогою гістопатології. У цьому творі використовується переказ дослідження для розробки CAD-моделей раку молочної залози.

У деяких дослідженнях [13] CNN використовувався для класифікації набору даних BreastHist. У цій статті використовувалися різні комбінації тренувальних і тестових вибірок, тому їх продуктивність не можна порівняти з [42]. У дослідженні Spanhol et al [42], раніше вивченому для видалення використовувався BVLC CaffeNet (модифікований AlexNet). Особливості (DeCAF). У цій статті також змінено розмір вхідного зображення для використання раніше навченої CNN, що, у свою чергу, може призвести до втрати інформації.

Висновки до розділу 1

В даному розділі було показано, що сучасні моделі глибокого навчання вже широко досліджуються як науковцями, інститутами так і великими компаніями для класифікації рентгенівських знімків та локалізації на них уражень. В більшості статей автори велику увагу звертають на попередню обробку даних. Тому в розділі було розглянуто додатково матеріали і статті, що описують такі теми як фільтрація кісткової тканини на рентгені грудної

клітини, покращення зображень за допомогою фільтрів таких як медіанний адаптивний фільтр і тд.

Слід зауважити, що в останні роки, багато розробників починають додавати додатковий функціональний модуль з механізмом уваги. Цей модуль може додатково видати хітмапу картинки, на якій темними колами покаже де на даній картинці вірогідніше всього знаходиться ураження.

РОЗДІЛ 2

ТЕОРЕТИЧНА ЧАСТИНА

2.1. Опис проблем мануальної діагностики торакальної патології клітини

До групи патологій органів грудної клітки входять захворювання, які становлять серйозну загрозу життю та здоров'ю людей [56]. Більшість часто з них зустрічаються: пневмонія, туберкульоз легень, ателектаз, легеневі абсцеси, гідроторакс, пневмоторакс та ін. Відмінною рисою більшості легеневих захворювань є складний діагноз. Характерними ознаками патологій є індивідуальний характер і може змінюватися від людини до людини [62]. Але існує ряд загальних факторів, які провокують виникнення і розвиток патології легенів. Вони поділяються на зовнішні та внутрішні. Внутрішні фактори включають:

- генетична схильність;
- ожиріння;
- образ життя.

Зовнішні фактори включають:

- несприятлива екологія в місці проживання;
- професійні сенсibiliзуючі речовини;
- активне і пасивне куріння;
- деякі особливості харчування;
- наявність хронічних захворювань легень.

Розглянемо докладніше деякі патології, згадані вище. Інфільтрація - ущільнення легеневої тканини за рахунок коміркових елементів. Рентген показує нерегулярні ущільнення. Виникає в результаті інфекції або травми.

Пневмонія - це запальний процес, що виникає в легенях випадок зараження.

Рентген показує:

- затемнення легеневої тканини;
- контрастність легеневого малюнка;
- збільшення кореня легені на боці запалення.

Легеневий фіброз - заміщення легеневої тканини сполучною.

Виникає з кількох причин:

- проникнення інфекції;
- наявність алергії;
- тривале вдихання пилу.

На рентгенограмі характеризується змінами нижньої відділи легенів.

Ателектаз - це патологія, яка характеризується відсутністю повітря в легеневої тканини. Виникає в результаті зміни прохідності бронхів.

Рентген показує:

- нечіткі межі серцевих відділів;
- тригранне затемнення;
- підняття коренів легенів.

Гідроторакс - скупчення рідини в плеврі. Зазвичай виникає в поєднанні з іншими патологіями. Стискає і зменшує розмір легеневої тканини. Затінений на рентгені середньої інтенсивності. Пневмоторакс – це патологія, що характеризується накопиченням повітря в плеврі. Виникає в результаті різних механічних пошкодження. На рентгенівському знімку характеризується зниженим купол діафрагми, просвітлення без легеневого малюнка. Всі патології мають свої характерні ознаки, за якими існує можливість класифікувати. Існують також різні способи виявлення патологій грудної клітини.

2.2. Методи діагностики патології органів грудної клітки.

Використовується для діагностики патологій органів грудної клітки різноманітні неінвазивні методи дослідження внутрішніх органів. Серед них магнітно-резонансна томографія, флюорографія, рентгенографія,

флюорографія та ін. В усіх них завдання експерта – оглянути зображення і виявити ознаки патологій, серед яких можуть бути такі особливості:

1. Зони просвітлення і затемнення.
2. Відсутність вогнищевих та інфільтративних тіней.
3. Порушення в системі кровообігу та кровопостачання.
4. Відсутність кістозних утворень, застійних явищ.
5. Структури коренів легкі, їх розміри та важкість, лімфатичні вузлів.
6. Посилення малюнка, фіброз.
7. Кальцинати.
8. Зміни діафрагми.

Розглянемо деякі способи ідентифікації вище особливості легень, а також їх переваги та недоліки. Флюорографія - це рентгенологічний метод дослідження легень і органів грудної клітки, результати якого фіксуються на спец чутлива плівка [15]. За допомогою флюорографії можна виявити запалення тканин, пухлини, ущільнення і рубцеві освіти, а також сторонні предмети. До переваг методу можна віднести:

- мінімальна витрата часу;
- низьке радіаційне навантаження порівняно з іншими методами;

До недоліків методу можна віднести:

- великий ступінь похибок при використанні плівки апарату;
- невеликий розмір плівки, через що важко помітити дрібні зміни або вогнища патології;
- відсутність можливості розглянути детально підозрілу ділянку.

Рентгенографія є одним з основних методів рентгенологічного дослідження, що полягає в отриманні зображення за допомогою рентгенівські промені. Цей метод є дуже точним та інформативним [13].

До переваг методу можна віднести:

- відносна простота дослідження;
- низька вартість дослідження;

- відносно низьке радіаційне навантаження на об'єкт;

До недоліків методу можна віднести:

- великі витрати часу на обробку зображення;
- труднощі з проведенням досліджень в різних проєкціях.

Рентгеноскопія – метод отримання зображення легенів за допомогою рентгена на спеціальному екрані. Він дозволяє побачити різні патології легеневих тканин, наявність рідини або повітря в плеврі.

До переваг методу можна віднести:

- здатність проводити дослідження в реальному масштабі
- вміння проводити дослідження в різних проєкціях.

До недоліків методу можна віднести:

- неможливість побачити дрібні вогнища патологій, які можуть обернутися метастазами або згодом ускладнити лікування;
- відносно висока доза радіації.

Комп'ютерна томографія – метод, який дозволяє детально оглянути досліджуваний об'єкт (в даному випадку органи грудної клітки) по шарах. У комп'ютерній томографії зображення отримують за допомогою рентгенівські промені. Отримані зображення передаються на комп'ютер для подальше лікування у спеціаліста та діагностика [21].

Метод дозволяє виявити:

- запалення тканин;
- вроджені патології;
- пухлини;
- онкологічні захворювання;
- ущільнення і рубцеві трансформації;
- Травма грудної клітки.

Переваги:

- більш детальне дослідження;
- більш чітка візуалізація структур органу.

Недоліки методу:

- при нанесенні контрасту потрібна підготовка;
- наявність обмежень для проведення досліджень с
- висока вартість дослідження.

Позитронно-емісійна томографія - медичний погляд дослідження, що дозволяють добре вивчити стан органів. для метод отримання зображення використовує радіоактивні хімікати речовини та спеціальний сканер.

Метод застосовується до:

- виявлення вогнищ;
- визначення наявності метастазів;
- виявлення змін легневих структур.

Переваги методу:

- дозволяє отримати інформацію про якість і характер патологічні зміни.
- дозволяє дуже рано виявити утворення.

Недоліки методу:

- необхідність введення радіаційної речовини для якісного дослідження;
- зображення містять менше інформації, ніж інші томографічні дослідження;
- висока вартість дослідження.

Підсумовуючи, можна сказати, що способів обстеження органів грудної клітки багато, що застосовуються на перевірку наявностей або розвиненості патологій.

2.3. Проблеми задачі класифікації рентгенівських зображень

Завдання діагностики захворювань легенів за допомогою рентгенівського зображення може бути складнішим, ніж стандартне завдання розпізнавання або класифікації об'єктів на зображеннях, оскільки різниця між об'єктами різних класів може бути незначною, а в деяких випадках її не видно навіть для людини, що займається цим професійно. Тому часто людина не може розрізнити випадки на рентгенівських знімках. Навіть лікарі, які дуже

добре вміють робити свою справу, час від часу можуть помилятися. Згідно з деякими згаданими дослідженнями, в середньому точність ручної класифікації становить близько 80%[47]. Більшість людей, у тому числі непрофесійні радіологи, не зможуть правильно визначити пухлину без значної кількості інформації.

Архітектури нейронних мереж стараються розроблятися таким чином, щоб мати можливість ідентифікувати дрібні невидимі елементи на зображенні. Перш ніж будувати таку мережу, важливо провести ретельні дослідження та тестування. Однією з найбільших проблем використання ШІ в медицині є отримання правильних медичних даних для навчання алгоритмів. Це потребує багато часу та досвіду людей, які дуже спеціалізуються в цій галузі.

У багатьох країнах діють закони, які захищають конфіденційність людей і зберігають конфіденційність їхніх медичних даних. Це ускладнює доступ і використання інформації про своє здоров'я. У відкритому доступі є багато рентгенівських знімків, і більшість із них не дуже якісні. Низька якість стане причиною, що навчання глибокої нейронної мережі високої якості може зайняти багато часу, і для цього потрібно багато схожих зображень, але не дасть високих показників метри на виході. Рентген може допомогти діагностувати захворювання легенів, але ми не маємо достатньо навчальних даних для ефективного використання цього методу. Нам потрібно більше даних, щоб підвищити точність.

Існує можливість створення централізованої платформи для зберігання медичних даних, яка дозволить різним фахівцям збирати дані для власного використання у сфері ШІ, а також забезпечить належну оптимізацію відповідного законодавства. Це дозволить нам підвищити ефективність машинного навчання в медицині, використовуючи різноманітні дані з різних джерел. Щоб діагностувати захворювання легенів за допомогою аналізу рентгенівських зображень, нам потрібно спочатку використати техніку, яка називається попередньою обробкою, а потім додати вхідні дані. Далі в цій статті ми поговоримо про деякі способи, які допоможуть вам.

Рентгенографія — це найпростіший спосіб визначити, чи є у когось проблеми з легенями, і це особливо корисно для людей, які не можуть дозволити собі часто відвідувати лікаря. За допомогою рентгенографії можна побачити захворювання на ранніх стадіях. Даний вид діагностики широко застосовується у всьому світі, і залишається широковикористовуваним і пріоритетним у широкій медичній практиці. Існують інші, більш детальніші способи діагностики проблеми, але це найпоширеніший спосіб.

Багатьом лікарям для встановлення діагнозу потрібна серія рентгенівських знімків. Одна з проблем полягає в тому, що зображення зазвичай мають низьку роздільну здатність, і це може ускладнити перегляд деталей. Крім того, якщо використовуються дані з різних джерел, можуть виникнути проблеми з нерівномірністю та шумом. Ось чому багатьом лікарям потрібна серія з кількох зображень, щоб отримати більш точний діагноз.

Нейронна мережа може мати проблеми з якістю даних, що надсилаються, або може бути недостатньою роздільною здатністю в даних. Дивлячись на рентгенівські знімки пацієнта, важливо пам'ятати, що його тіло може перебувати в різних положеннях. Наприклад, іноді пацієнт лежить, а іноді він може рухатися. Машинне навчання може бути складним у використанні, якщо ви не враховуєте певні речі.

Отже, опираючись на все те що було описано вище можна зробити висновок, що кожен тип проблеми комп'ютерного зору складний і багатогранний в предметній області медицини, ніяка із задач не являється дефолтною та тривіальною, і потребує багато часу, даних, ресурсів та досліджень для ретельного вивчення.

2.4. Штучні нейронні мережі

Сьогодні штучні нейронні мережі використовуються по-різному. Деякі з них використовуються в комп'ютерних програмах і веб-сайтах. Можна виділити класи нейронних мереж які, використовуються в таких речах, як:

- Розпізнавання зображень (або класифікація зображень) — це завдання ідентифікації зображень і класифікації їх в одному з кількох попередньо визначених окремих класів. Отже, програмне забезпечення та програми для розпізнавання зображень можуть визначити, що зображено на зображенні, і відрізнити один об'єкт від іншого. Галузь дослідження, спрямована на забезпечення машин цієї здатністю, називається комп'ютерним зором. Будучи одним із завдань комп'ютерного зору (CV), класифікація зображень служить основою для вирішення різних проблем CV, включаючи класифікацію зображень із локалізацією – розміщення зображення в заданому класі та малювання рамки навколо об'єкта, щоб показати, де він розташований на зображенні[61];
- Прогнозування. Ідея рекурентних нейронних мереж (RNN) полягає у використанні послідовної інформації. У традиційній нейронній мережі ми припускаємо, що всі входи незалежні один від одного. Але для багатьох завдань це не є оптимальною ідеєю. RNN називають рекурентними, тому що вони виконують одне й те саме завдання для кожного елемента послідовності, при цьому результат залежить від попередніх обчислень.

Інший спосіб думати про RNN полягає в тому, що вони мають «пам'ять», яка фіксує інформацію про те, що було обчислено до цього часу. Теоретично RNN можуть використовувати інформацію в довільно довгих послідовностях, але на практиці вони обмежені оглядом лише на кілька кроків (докладніше про це пізніше). Мережі LSTM сьогодні досить популярні. LSTM не мають принципово відмінної архітектури від RNN, але вони використовують іншу функцію для обчислення прихованого стану. Комірка LSTM має 5 важливих компонентів, які дозволяють моделювати як довгострокові, так і короткострокові дані: стан комірки, прихований стан, входні ворота, забуді ворота та вихідні ворота. Однією з найважливіших переваг LSTM є їхня здатність запам'ятовувати довгострокові послідовності[12];

- Кластеризація. CentNN — це алгоритм неконтрольованого змагального навчання, заснований на класичному алгоритмі кластеризації k -середніх, який оцінює центроїди пов'язаних груп кластерів у даті навчання. CentNN не вимагає ні попередньо визначеного розкладу для коефіцієнта навчання, ні загальної кількості ітерацій для кластеризації. У статті результати моделювання проблем кластеризації та проблем стиснення зображень показують, що CentNN сходиться набагато швидше, ніж звичайні алгоритми із сумісною якістю кластеризації, тоді як інші алгоритми можуть давати нестабільні результати залежно від початкових значень коефіцієнта навчання та загальної кількості ітерацій.;
- Апроксимація. Нейронні мережі можуть виконувати досить хорошу роботу з апроксимації безперервних функцій. Використовуючи лінійні операції та каскадні з'єднання, можна створити програму, яка зможе обчислювати будь-яку безперервну функцію з певною точністю. Нейрони можуть мати будь-яку нелінійну поведінку, від простої сигмоподібної кривої до вейвлета, синусоїди або полінома. Складність мережі залежить від типу нелінійної функції, яку вона використовує, але будь-яка нелінійність може допомогти мережі наблизити роботу безперервного автомата[63].
- Асоціативна пам'ять. Стиснення даних допомагає нейронним мережам запам'ятовувати зв'язки між різними параметрами, що дає змогу виразити більше інформації про дані в компактному вигляді. Зворотний процес — відновлення вихідного набору даних про частину інформації — називається (авто)асоціативною пам'яттю. Пов'язана пам'ять допомагає відновити вихідний сигнал, зображення або пам'ять, уникнувши перешкод або пошкоджень.

Існує багато різних способів використання штучних нейронних мереж, і це лише приблизний перелік.

2.5. Згорткові нейронні мережі

Є кілька недоліків класичних ML алгоритмів, особливо коли мова йде про обробку зображень. ML використовує один перцептрон для кожного входу (наприклад, піксель у зображенні, помножений на 3 у випадку RGB). Кількість ваг швидко стає неконтрольованою для великих зображень. Для зображення розміром 224 x 224 пікселя з 3 кольоровими каналами потрібно навчитися близько 150 000 ваг тільки на першому слої! У результаті виникають труднощі під час тренувань.

Ще одна поширена проблема полягає в тому, що класичні ML алгоритми по-різному реагують на введення (зображення) і його зміщену версію — вони не інваріантні до перекладу. Наприклад, якщо зображення kota з'являється у верхньому лівому куті зображення на одному зображенні та в нижньому правому куті іншого зображення, ML спробує виправити себе та припустить, що кіт завжди з'являтиметься в цій частині зображення.

Очевидно, що класичний ML – не найкраща ідея для обробки зображень. Одна з головних проблем полягає в тому, що просторова інформація втрачається, коли зображення вирівнюється в MLP. Вузли, розташовані близько один до одного, важливі, оскільки вони допомагають визначити особливості зображення. Таким чином, нам потрібен спосіб використовувати просторову кореляцію характеристик зображення (пікселів) таким чином, щоб ми могли бачити kota на нашому зображенні незалежно від того, де він може з'явитися. Сподіваюся, зрозуміло, чому MLP — це жахлива ідея для обробки зображень. Тепер давайте підемо далі та обговоримо, як CNN можна використовувати для вирішення більшості наших проблем [69].

Згорткові нейронні мережі – це один з видів моделей глибинного навчання, який використовується частіше всього для обробки зображень. За деякими структурними елементами CNN схожі із організацією зорової кори людей і використовуються для освоєння просторових ієрархій і патернів в зображеннях. Згортковий шар, пули об'єднання та повністю пов'язаний шар

як правило є основними структурними елементами зготнової нейронної мережі[48] .

Згортка – це спеціалізований тип лінійної операції, що використовується для витягування ознак, коли невеликий масив чисел, що називається ядром, застосовується до входу, що представляє собою масив чисел, що називається тензором. Поелементне (рух фільтра див. рис. 2.2) множення між кожним елементом ядра та вхідним тензором обчислюється в кожному місці тензору та підсумовується для отримання вихідних значень у відповідному положенні вихідного тензору, називаючи картою активацій (рис. 2.1).

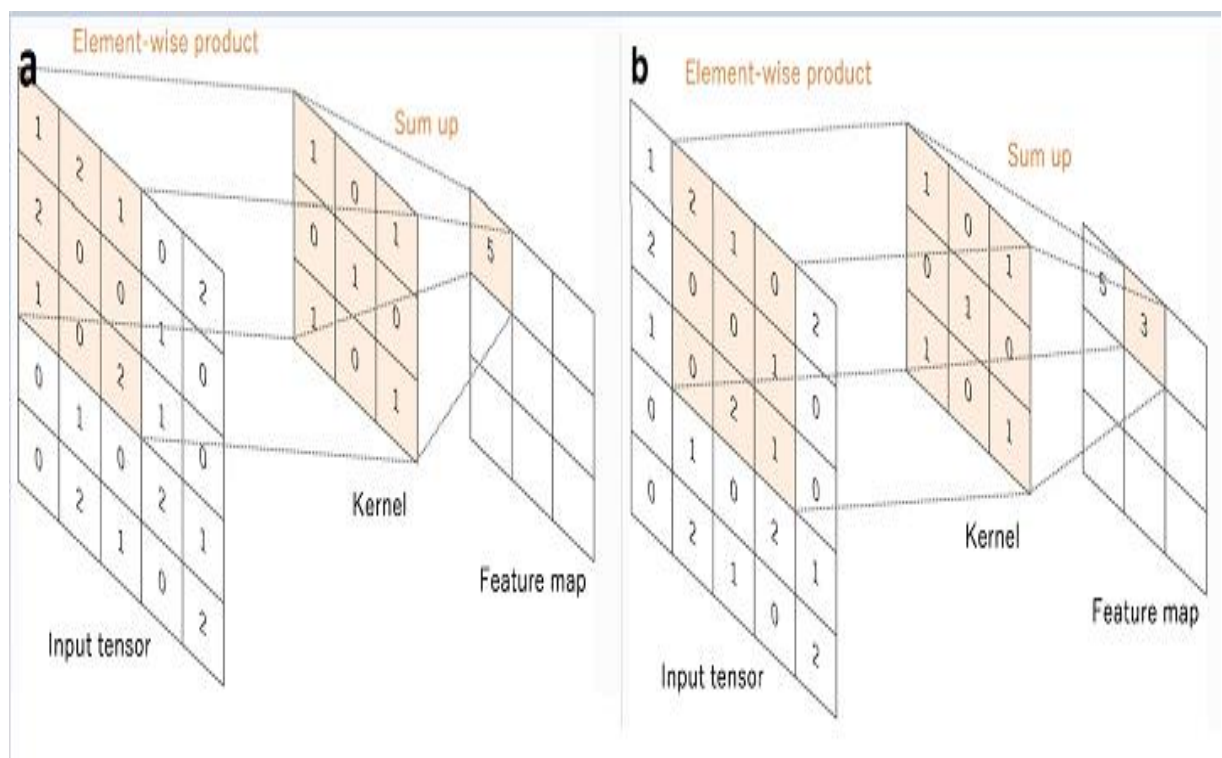


Рисунок 2.1 – Принцип згортки зображення [53]

Згортковий шар містить набір фільтрів, параметри яких необхідно вивчити. Висота і вага фільтрів менші, ніж у вхідного об'єму. Кожен фільтр згортається разом із вхідним об'ємом для обчислення карти активації, що складається з нейронів. Іншими словами, фільтр пересувається по ширині та висоті входу, а скалярний добуток між входом і фільтром обчислюється в кожній просторовій позиції.

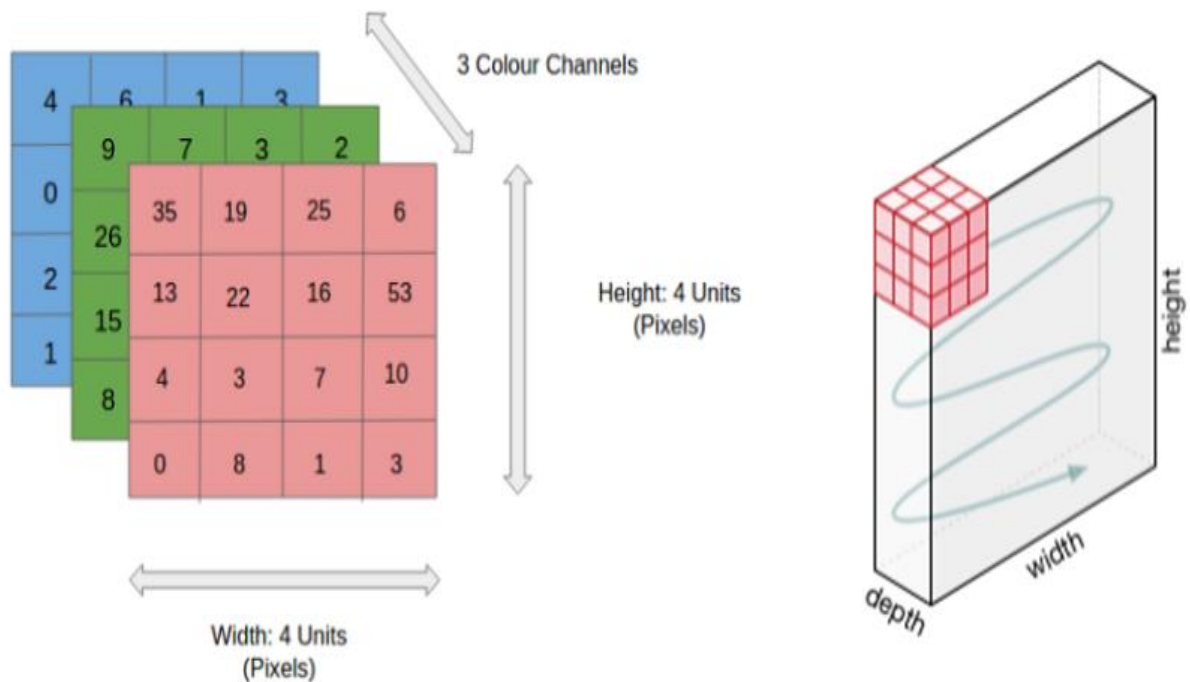


Рисунок 2.2 – Згортка у 3-ьох каналному зображенні [52]

Вихідний об'єм згорткового шару виходить укладанням карт активації всіх фільтрів уздовж вимірювання глибини. Оскільки ширина та висота кожного фільтра розроблені так, щоб бути меншими за вхід, кожен нейрон у карті активації пов'язаний лише з невеликою локальною областю вхідного об'єму. Іншими словами, розмір рецептивного поля кожного нейрона невеликий і дорівнює розміру фільтра. Локальний зв'язок мотивується архітектурою зорової кори тварин [46], де рецептивні поля клітин малі. Локальна зв'язність згорткового рівня дозволяє мережі вивчати фільтри, які максимально реагують на локальну область вхідних даних, таким чином використовуючи просторову локальну кореляцію вхідних даних (для вхідного зображення піксель більше корельований до найближчих пікселів, ніж до віддалені пікселі). Крім того, оскільки карта активації отримується шляхом виконання згортки між фільтром і входом, параметри фільтра спільно використовуються для всіх локальних позицій. Розподіл ваги зменшує кількість параметрів ефективності вираження,

ефективності навчання та хорошого узагальнення

Існує два типи результатів операції - один, у якого згорнутий елемент зменшується розмірністю порівняно з вхідним, і другий, у якого розмірність або збільшується, або залишається незмінною, це досягається за допомогою додавання симетричних або нульових пікселів на краї зображень [69].

CNN для класифікації мають два компоненти (рис 2.3):

1. Бекбон – це частина що відповідає за вилучення признаков з об'єкту, на виході з бекбону ми отримуємо, карти активацій. Що далі будуть подаватися на класифікаційну частину.

2. Хед (в нашому випадку класифікаційна частина). Група повнозв'язних шарів приймають рішення на основі вхідної інформації. Вони призначають імовірність об'єкту на зображенні таким, яким його передбачає алгоритм[41].

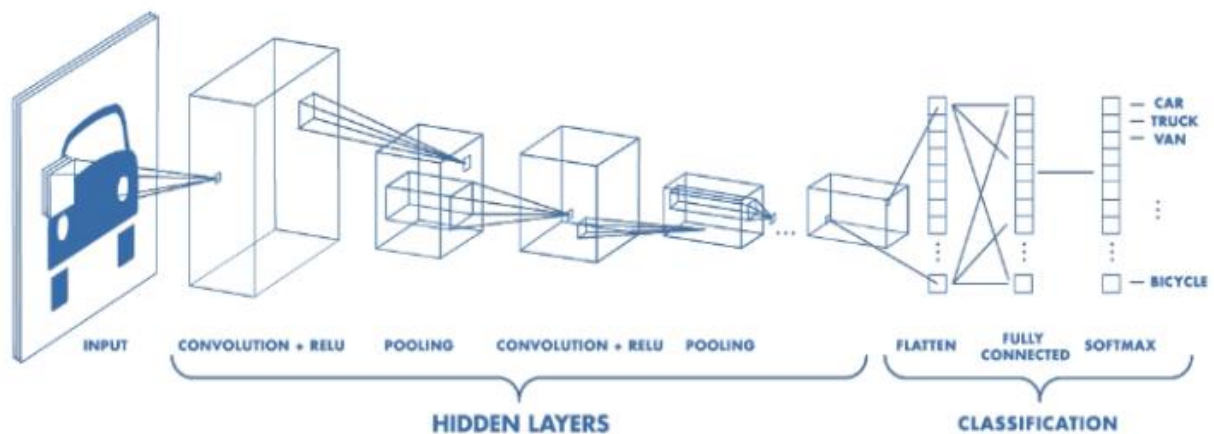


Рисунок 2.3 – Класична структура згорткової нейронної мережі з класифікатором [52]

Класифікувати зображення означає визначити, який клас найкраще описує зображення. Комп'ютер отримує на вхід купу зображень.. Комп'ютер приймає на вхід масив пікселів (рис.2.4).

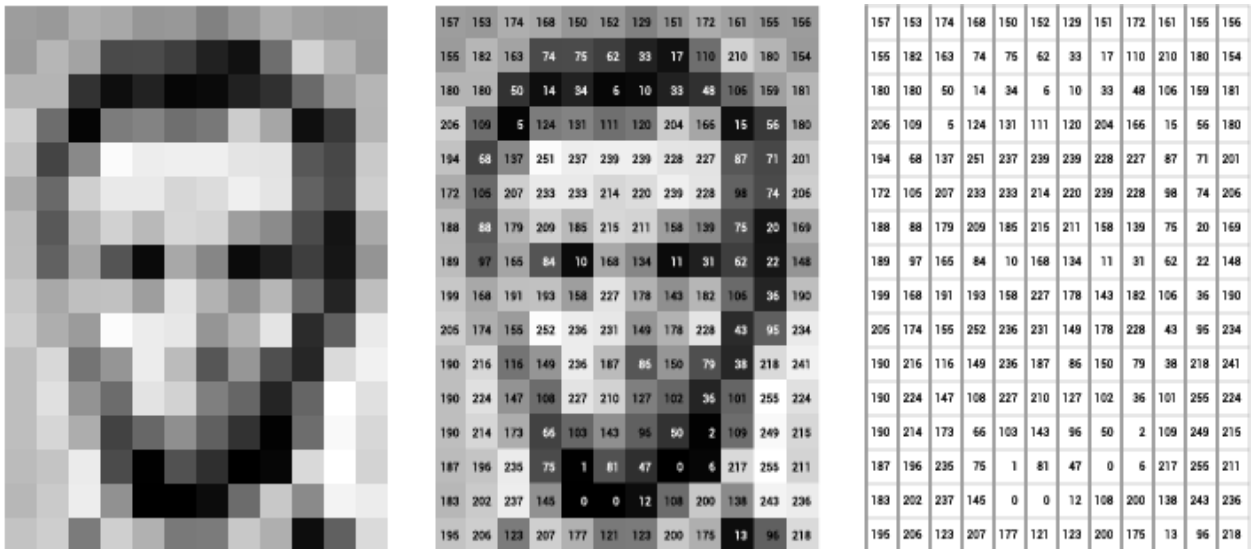


Рисунок 2.4 – Представлення зображення у пам'яті комп'ютера [41]

Як приклад, уявімо, що ми підвантажуюємо картинку розміром 224×224 , отже в пам'яті комп'ютера це буде масив з цілими числами від 0 до 255 розміром $224 \times 224 \times 3$ (в даному випадку «3» - це показник що вказує на канали RGB, а «224» - висота і ширина в пікселях).

Кольорове зображення складається з маленьких квадратів, кожен з яких складається з трьох кольорів: червоного, зеленого та синього. Масив $224 \times 224 \times 3$ відповідає загальній кількості інформації про картинку. Кожен піксель на екрані має певний колір на основі числа від 0 до 255. Це число описує, наскільки яскравим є цей колір у цій точці екрана.

Мережа може виконувати класифікацію зображень через пошук характеристик базового рівня, наприклад, границь і викривлень, а потім з допомогою побудови більш абстрактних концепцій через групи згорткових шарів [69].

Додавання шару Fully Connected дозволяє досліджувати різні комбінації функцій, представлених на виході згорткового шару. FC розглядає функцію, яка може бути нелінійною в цьому просторі. [52].

Після згорткових шарів ми отримуємо тензор-матрицю з числами, ця матриця, щоб побудувати FC шар переводиться в вектор-стовпець, шляхом послідовного вилучення елемента з матриці та вставляння у вектор

стовпець[52]. Умовно процедура отримання вектор-стовпця показана на рис. 2.5.

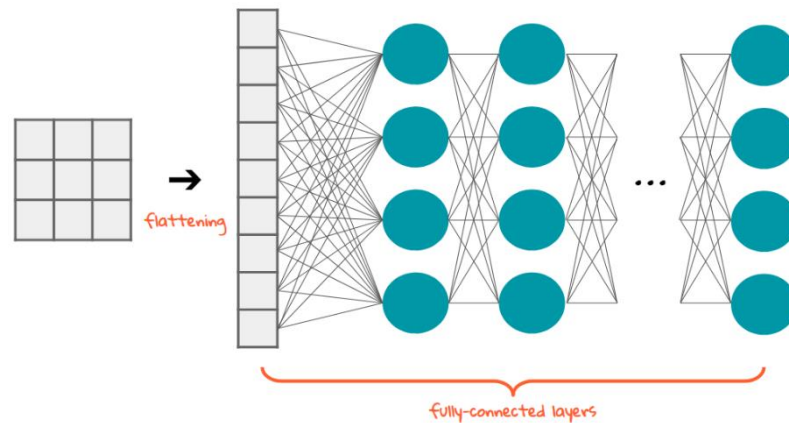


Рисунок 2.5 – Флатеннінг зображення [25]

Нейронна мережа має на своїх виходах деякі числа, які вона використовує для класифікації речей. Кожне число говорить вам щось про введене зображення. Модель зрозуміти, наскільки ймовірно, що вхідний екземпляр належить до певного класу. Функція Softmax використовується для нормалізації виходу та підвищення інтерпретабельності останнього шару нейронної мережі. Щоб ця логіка працювала, вам потрібно вставити кожен вихідний перцептрон у функцію Softmax, яка перетворить отриманий вектор m розмірності l у вектор $s(z)$ розмірності l дійсних чисел в інтервалі від 0 до 1, сума вихідних елементів вихідного вектора завжди дорівнює 1. Функція softmax() функція приймає число як вхідні дані та передбачає значення даного елемента виводу, яке є сумою всіх елементів введення, помножених на відповідну вагу.

$$p(m)_j = \frac{e^{m_j}}{\sum_{l=1}^L e^{m_l}} \quad (2.1)$$

де z – вхідний вектор, e^{m_j} – стандартна експоненціальна функція для вхідного j -того елемента вектора, e^{m_l} – стандартна експоненціальна функція для вхідного вектора, L – кількість класів.

Персептрон — це мінімальний структурний елемент нейронної мережі, який допомагає створювати повноцінну зв'язану нейронну мережу. Нейромережа використовує зв'язки між нейронами, щоб надати кожному нейрону особливу вагу. У згортковій мережі кожен шар використовує обмежену вагову матрицю, яка «переміщується» по всьому шару з самого початку, безпосередньо вздовж вхідного зображення. Після кожного зсуву ця вагова матриця формує сигнал активації для нейрона наступного шару. Наступний шар показує, які патерни присутні на попередньому шарі та їхні координати. Цю карту називають «карту ознак», оскільки вона показує, які об'єкти присутні на зображенні [3]. Схематично рух фільтра зображено на рис. (2.6).

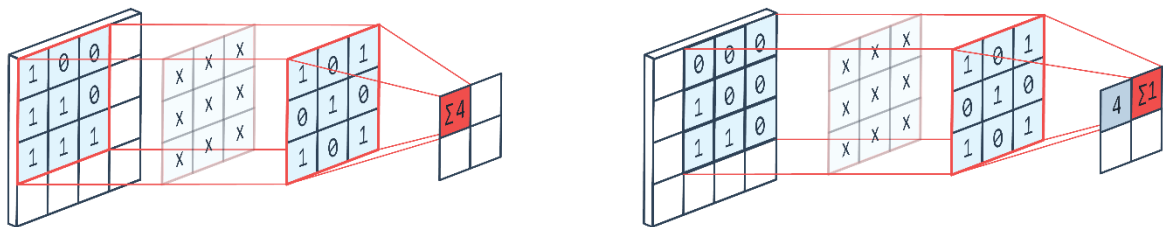


Рисунок 2.6 – Рух фільтра вздовж зображення

В згортковій нейронній мережі набір ваг не один, а тисячі, а іноді і мільйони, що кодує елементи зображення. Про це такі ядра згортки не закладаються дослідником заздалегідь, а формуються самостійно шляхом навчання мережі класичним методом зворотного поширення помилки, прохід кожним набором ваг формує свій власний примірник карти ознак, роблячи нейронну мережу багатоканальною (багато незалежних карт признаков на одному шарі) [69], візуалізацію фільтрів на різних рівнях можна побачити на рис. 2.7

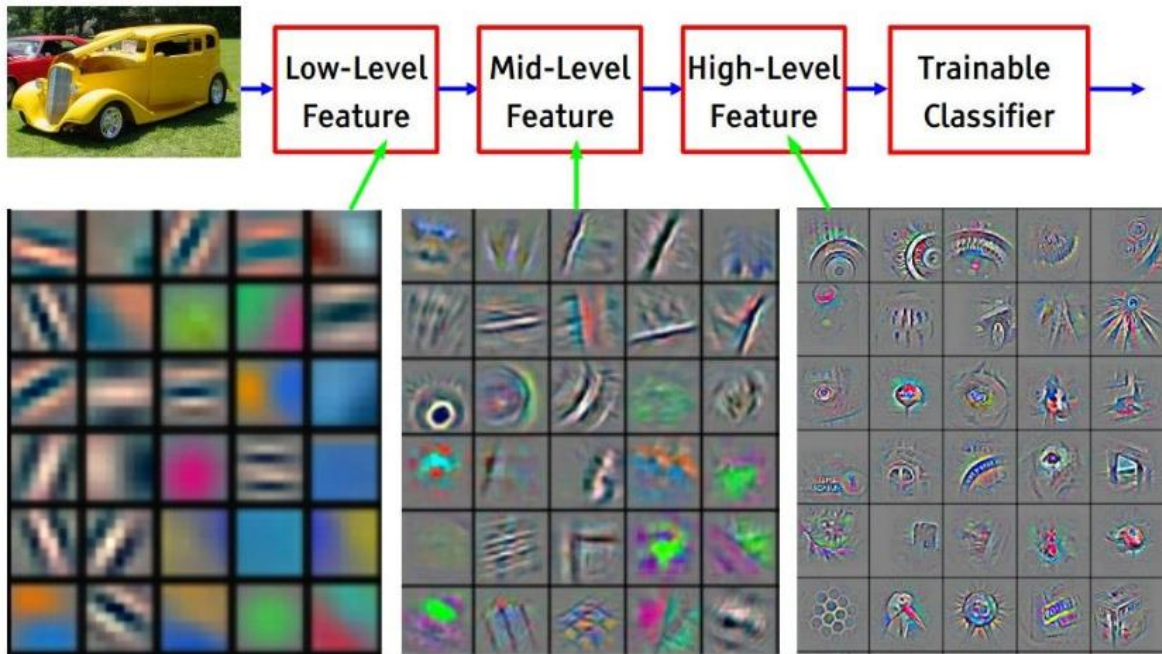


Рисунок 2.7 – Візуалізація фільтрів [23]

Згорткові мережі добре розпізнають шаблони, але їх може бути важко налаштувати, оскільки існує багато різних налаштувань, які потрібно вибрати для свого комп'ютера. Існує багато різних параметрів, які можуть впливати на згорткову нейронну мережу (CNN). Деякі з цих параметрів включають кількість шарів, розмір ядер, що використовуються для кожного шару, кількість ядер, що використовуються для кожного шару, крок зсуву ядра під час обробки шару, необхідність шарів підвибірки та тип зменшення розмірності (наприклад максимальний, середній тощо). [67].

2.6. Додаткові операції в згорткових нейронних мережах

2.6.1. Pooling (Об'єднання)

Шар об'єднання – це новий шар, який додається після шару згортки. Зокрема, після застосування нелінійності (наприклад, ReLU) до карт функцій, виведених згортковим шаром; наприклад, шари в моделі можуть виглядати наступним чином:

Додавання шару об'єднання після згорткового шару є загальним шаблоном, який використовується для впорядкування шарів у згортковій нейронній мережі, який може повторюватися один або кілька разів у даній моделі.

Рівень об'єднання працює з кожною картою об'єктів окремо, щоб створити новий набір такої ж кількості об'єднаних карт об'єктів.

Об'єднання передбачає вибір операції об'єднання, подібно до фільтра, який буде застосовано до карт об'єктів. Розмір операції об'єднання або фільтра менший за розмір карти функцій; зокрема, це майже завжди 2×2 пікселі, що застосовуються з кроком у 2 пікселі (рис. 2.8):

Це означає, що рівень об'єднання завжди зменшуватиме розмір кожної карти об'єктів у 2 рази, напр. кожен розмір зменшується вдвічі, зменшуючи кількість пікселів або значень у кожній карті функцій до однієї чверті розміру.

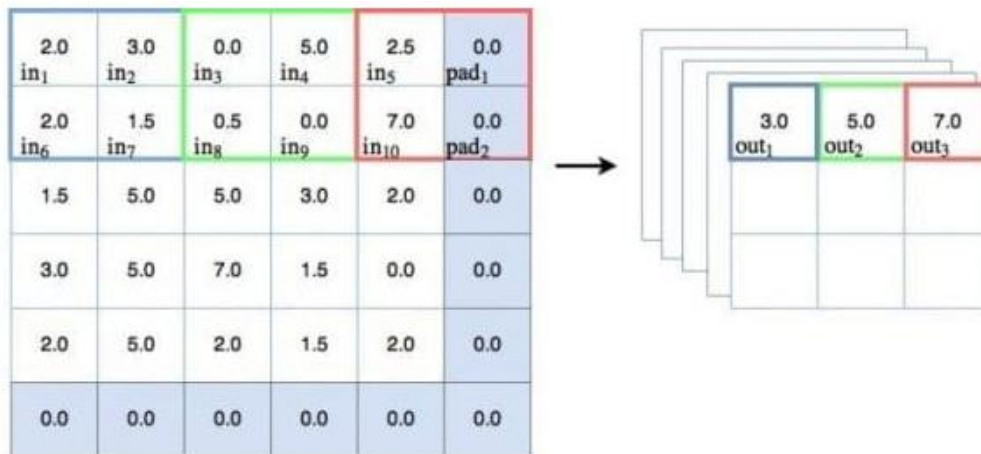


Рисунок 2.8 – Операція max pooling

Наприклад, шар об'єднання, застосований до карти об'єктів розміром 6×6 (36 пікселів), призведе до вихідної об'єднаної карти об'єктів розміром 3×3 (9 пікселів).

На рис. 2.8 видно, що пулінгове вікно рухається вправо з кроком 2 одиниці. На даному рисунку проілюстровано кроки тільки по осі x, але даний рух відбувається і по осі y з кроком 2. Другими словами, крок пулінгу буде

позначатися як $[2,2]$. Чим більший крок пулінгу тим більш зменшений буде вихід після операції над картою чи зображенням. Як можна побачити на попередньому рисунку, вхідний тензор, що має розмір 6×6 буде зменшено до 3×3 на виході. В загальних термінах пулінг можна віднести до категорії даунсемплінгу так як він зменшує кількість тренувальних паматерів та полугшує модель.

Концепція медіанного пулінгу вперше була використана в [32] та попередньо представлена в [40]. На рис. 2.9 показано як медіанний пулінг об'єднання виконує зменшення дискретизації шляхом поділу вхідних даних на прямокутні області об'єднання та обчислення середнього значення значення кожної області [31].

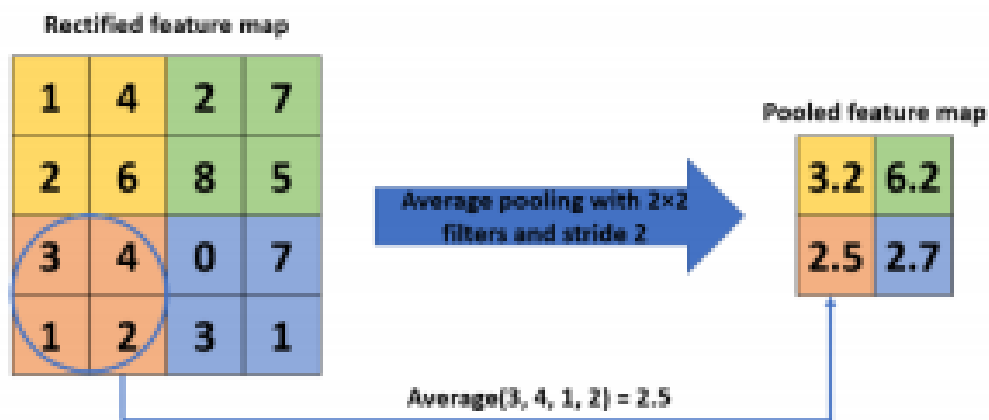


Рисунок 2.9 – Операція mean pooling [40]

Шар об'єднання допомагає зменшити розмір згорнутого елемента в просторі. Зменшення кількості вимірів у даних полегшить їх обробку комп'ютером. Вилучення домінуючих функцій, інваріантів обертання та позиціонування з даних допомагає вам ефективніше вивчати моделі. Максимальне об'єднання усуває всі шумові сигнали, а також поєднує придушення шумів із зменшенням розмірності. Усереднене об'єднання — це спосіб зробити речі більш упорядкованими шляхом придушення шуму. Максимальне об'єднання є кращим способом (рис 2.10), ніж усереднення[40].

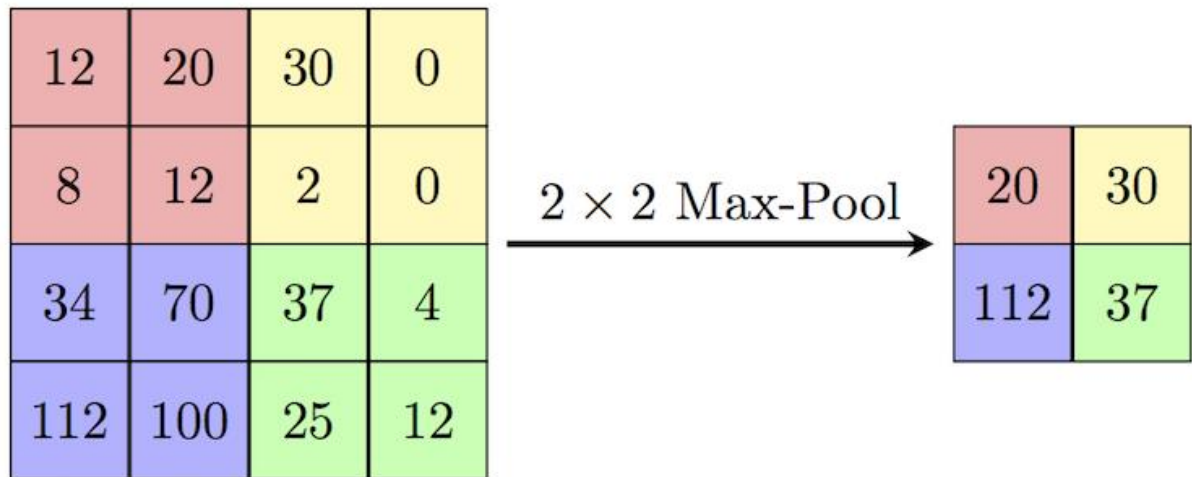


Рисунок 2.10 – Операція max pooling[40]

2.6.2. Функції активації

Однією з речей, яка допомагає будувати архітектуру нейронної мережі, є вибір правильної функції активації нейронів (розташування функції активації див. рис 2.11). Функції активації визначає, наскільки нейронна мережа придатна для навчання.. Але коли мережа стає глибшою, у вас може виникнути більше проблем із її навчанням. Однією з причин є так зване ослаблення градієнтів (*gradient vanishing problem*). Помилка множиться на похідну функції активації на кожному рівні, коли вона поширюється мережею. Похідна традиційної сигмоїдальної функції активації менше одиниці по всій області, тому після кількох шарів помилка буде близькою до нуля. Якщо функція активації має дуже круту похідну, це може призвести до швидкого зростання помилки та, зрештою, до нестабільності в навчанні мережі. [68].

Функцію активації неможливо пояснити у відриві від нейрона. Головні вимоги до неї - безперервність і монотонність по всій області визначення. Як таку функцію беруть або логістичний сигмоїд або гіперболічний тангенс. Вони насичуються на кордонах. Тобто нескінченне зростання аргументу не призводить до зростання функції.

Функція активації тісно пов'язана з теоремою Колмогорова Арнольда.

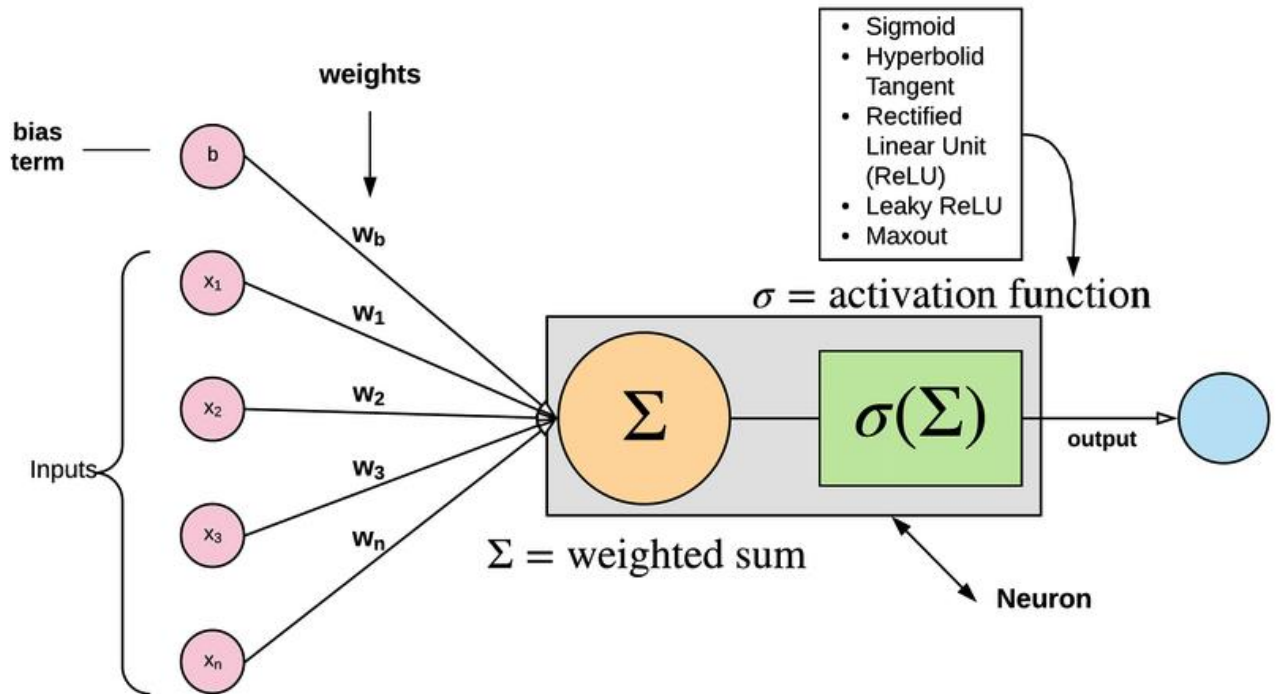


Рисунок 2.11 – Розташування функції активації

Сигмоїда

Сигмоїдальна функція є однією з безперервних активуючих функцій, що на вхід отримує будь-яке число, а як результат повертає число з проміжку від 0 до 1. В процесі активації негативні числа перетворюються в 0, а додатні в 1. Функція сигмоїди одна з перших застосовувалася в сітках, так як вона проста в обчисленні та інтерпретабельності. Сигмоїда виражається за формулою.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2.2)$$

де x – вхідне дійсне число, e – експонента

Одним з недоліком сигмоїдальної функції є те що при великих вхідних значеннях її градієнт близький до 0, що може погано впливати на навчання нейронної мережі

Графік сигмоїдальної функції наведений на малюнку нижче (рис. 2.12):

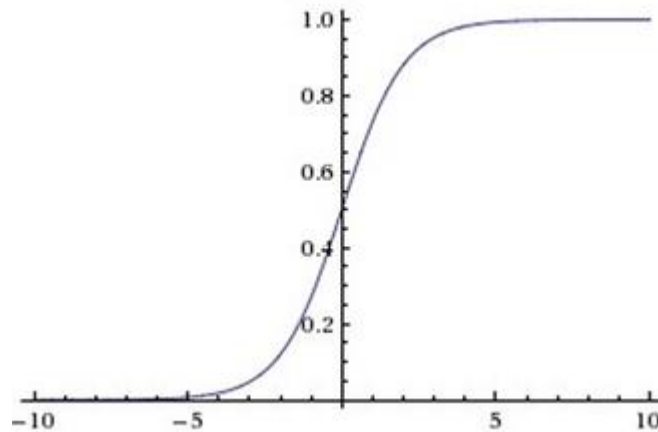


Рисунок 2.12 – Сигмоїдальна функція активації [68]

Зворотне поширення помилки працює шляхом множення локального градієнта на глобальний градієнт. Оскільки градієнти перемножаються, то якщо градієнт у точці буде малим, це обнулить загальний градієнт нейрона, і може викинути його з навчання. Важливо бути обережним, встановлюючи ваги сигмовидних нейронів, щоб вони не були надто великими, щоб нейрон міг давати найточніші результати. Якщо вихідні ваги занадто великі, мережа навчатиметься повільно, оскільки більшість нейронів перебуватимуть у стані насичення. [68].

ReLU

Rectified Linear Unit є найпоширенішою функцією активації в моделях глибокого навчання. Функція повертає 0, якщо отримує негативне значення, але для будь-якого позитивного значення x вона повертає це значення назад [69]. Отже, це можна записати як:

$$f(x) = \max(0, x) \quad (2.3)$$

де x – вхідне дійсне число, $\max()$ – функція максимуму

Графік функції ReLU наведений на малюнку нижче (рис. 2.13.):

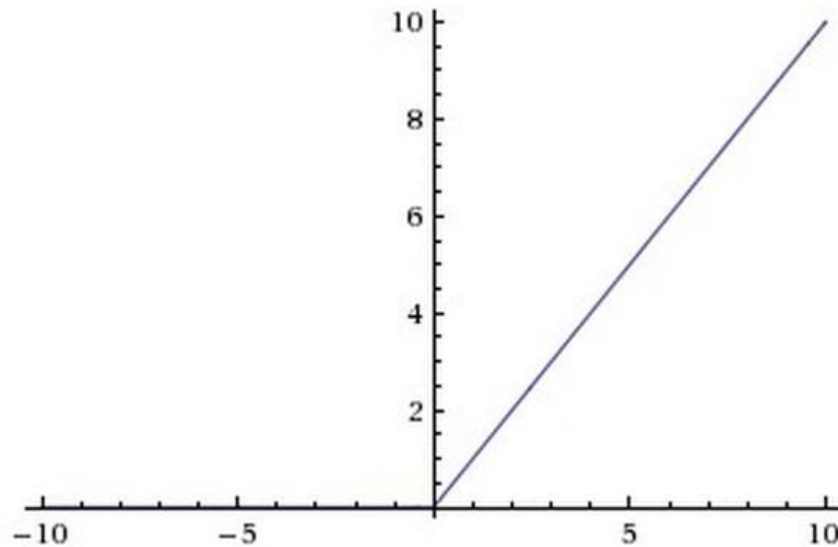


Рисунок 2.13 – функція активації ReLU [68]

Як бачите, ReLU наполовину випрямлений (знизу). $f(z)$ дорівнює нулю, коли z менше нуля, і $f(z)$ дорівнює z , коли z більше або дорівнює нулю. Діапазон: [від 0 до нескінченності). І функція, і її похідна монотонні. Але проблема полягає в тому, що всі від'ємні значення одразу стають нульовими, що зменшує здатність моделі належним чином підганяти або навчатися на основі даних. Це означає, що будь-який негативний вхід, наданий функції активації ReLU, одразу перетворює значення на графік на нуль, що, у свою чергу, впливає на кінцевий графік, не відображаючи від'ємні значення належним чином [69].

Існує достатньо багато різних варіантів функцій активацій (див. рис. 2.14). На рис зображено функції активацій, їх формули, схематичні графіки, та похідні.










Name	Plot	Equation	Derivative
Identity		$f(x) = x$	$f'(x) = 1$
Binary step		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$
Logistic (a.k.a. Soft step)		$f(x) = \frac{1}{1 + e^{-x}}$	$f'(x) = f(x)(1 - f(x))$
Tanh		$f(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
ArcTan		$f(x) = \tan^{-1}(x)$	$f'(x) = \frac{1}{x^2 + 1}$
Rectified Linear Unit (ReLU)		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Parameteric Rectified Linear Unit (PReLU) [2]		$f(x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Exponential Linear Unit (ELU) [3]		$f(x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
SoftPlus		$f(x) = \log_e(1 + e^x)$	$f'(x) = \frac{1}{1 + e^{-x}}$

Рисунок 2.14 – різновиди функцій активацій [45]

Також нижче (рис. 2.15) наведено поведінку градієнта функцій активацій на числовому проміжку.

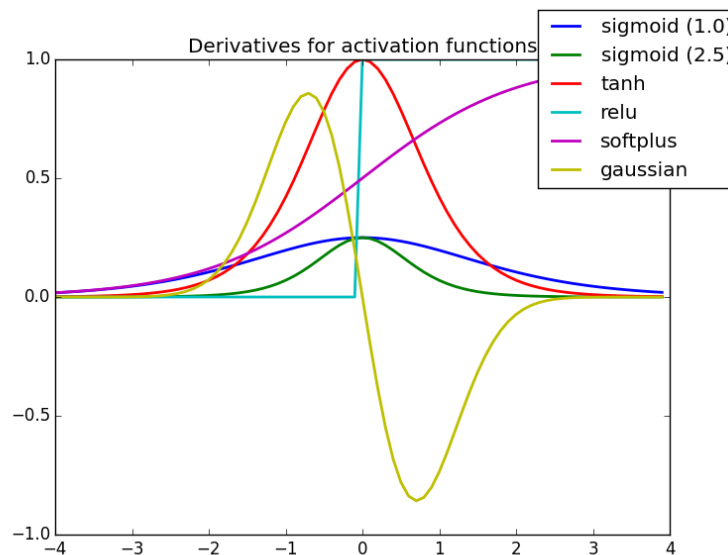


Рисунок 2.15 – похідні функцій активацій [45]

2.6.3. Оптимізатори нейронних мереж

Дивлячись на проблема оптимізації з точки зору геометріх в нейронних мережах наявна поверхня втрат. При відповідному наборі вагів – ми отримуємо відповідний вихідний результат мережі. Оптимізація нейронної мережі – пошук мінімум цієї поверхні [65].

Проблематика оптимізацій нейронних мереж описується наступним:

1. Вибір learning rate. При малому коефіцієнту швидкості навчання потрібно багато часу, щоб знайти мінімум, а велика швидкість навчання частіше всього вказує на те, що мережа не здатна знайти мінімум функції втрат.
2. Неможливість виходу з локального мінімуму
3. Зміна морфологія поверхні втрат. Оскільки набір даних не являється репрезентативним, модель з часом буде можливість покращити і додати нових даних для навчання. Нові дані можуть суттєво вплинути на морфологію функції помилки, і попередньо знайдений мінімум, вже не буде глобальним мінімум всієї поверхні.

Задля пошуку мінімального значення функції втрат використовують методи з класу оптимізаторів. Ці методи відповідають за зниження помилки та надання максимально точних вихідних даних.

Гرادієнтний спуск

Градiєнтний спуск (GD) — це ітеративний алгоритм оптимізації першого порядку, який використовується для знаходження локального мінімуму/максимуму заданої функції. Цей метод зазвичай використовується в машинному навчанні (ML) і глибокому навчанні (DL), щоб мінімізувати функцію витрати/втрати (наприклад, у лінійній регресії). Через його важливість і простоту впровадження цей алгоритм зазвичай викладають на початку майже всіх курсів машинного навчання. Алгоритм записується таким чином [46]:

$$\theta_i = \theta_{i-1} - \alpha * \nabla J(\theta_{i-1}) \quad (2.4)$$

де θ_i – значення вагів на i -тій ітерації, θ_{i-1} – значення вагів на попередній ітерації, $\nabla J(\theta_{i-1})$ – значення градієнту вагів на попередній ітерації, α – швидкість навчання [69].

Переваги звичайного градієнтного спуску:

1. Невисокі обчислювальні навантаження.
2. Відносно проста реалізація.
3. Простий в розумінні.

Недоліки звичайного градієнтного спуску:

1. Не стійкий до локальних мінімумів.
2. Ваги змінюються на всьому наборі даних. Тому якщо набір дуже великий навчання відбувається дуже повільно.
3. Потрібно виділяти багато пам'яті в оперативній пам'яті.

Стохастичний градієнтний спуск

Одним із підвидів градієнтного спуску є стохастичний градієнтний спуск, його відмінність в тому, що він частіше оновлює параметри сітки. Оптимізатор, змінює ваги на кожному прикладі навчання. Обраховується за формулою[46]:

$$\theta_i = \theta - \alpha * \nabla J(\theta_{i-1}; x(i); y(i)) \quad (2.5)$$

Де θ_i – значення вагів на i -тій ітерації, θ_{i-1} – значення вагів на попередній ітерації, $\nabla J(\theta_{i-1}; x(i); y(i))$ – значення градієнту вагів на попередній ітерації з навчальними прикладами $\{x(i); y(i)\}$, α – швидкість навчання [69].

Параметри моделі обновлюються часто і швидко, і після кожної поданої картинки, ваги міняються це призводить до дуже великої дисперсії в параметрах.

Переваги стохастичного градієнтного спуску:

1. Часта зміна вагів сітки.
2. Відсутня необхідність у резеруванні великого обсягу оперативної пам'яті.

Недоліки стохастичного градієнтного спуску:

1. Висока дисперсія параметрів сітки.
2. Потрібно додатково налаштовувати і робити гнучкою швидкість навчання, щоби після деякого часу тренування швидкість навчання зменшувалась[46].

Міні-пакетний градієнтний спуск

Мініградієнт є одним з найкращих серед варіантів градієнтного спуску. Він являє собою золоту середину між 2ма попередніми варіантами. В міні-пакетний градієнтний спуск картинка подаються партіями, пакетами. Модель отримує пакет картинок пропускає через себе кожну картинку і рахує загальну помилку на всіх картинках пакету, після цього оптимізатор оптимізує ваги відповідно до загальної помилки.

$$\theta = \theta - \alpha * \nabla J(\theta; B(i)) \quad (2.6)$$

де θ_i – значення вагів на i -тій ітерації, θ_{i-1} – значення вагів на попередній ітерації, $\nabla J(\theta_{i-1}; x(i); y(i))$ – значення градієнту вагів на попередній ітерації з навчальними прикладами які подаються пакетом $\{B(i)\}$, α – швидкість навчання [69].

Переваги пакетного градієнтного спуску:

1. Має не таку велику дисперсію як SGD, при цьому часто оновлює параметри

2. Пам'ять потрібно виділяти тільки на картинки з пакету

У всіх типів градієнтного спуску є свої проблеми:

1. Підбір learning rate, так як при низькому коефіцієнті можель буде довго сходитися.
2. Сталий learning rate для всіх параметрів
3. Схильність опинитися в пастках локальних мінімумів [46].

Адам

[Адам](#) (Адаптивна оцінка моменту) працює з імпульсами першого та другого порядку. Інтуїція, що стоїть за Адамом, полягає в тому, що не параметри нейронної мережі не можуть змінюватися так швидко лише тому, що можемо перестрибнути мінімум, ми хочемо трохи зменшити швидкість для ретельного пошуку [69].

Замість того щоб адаптувати швидкість навчання параметрів на основі середнього першого моменту (середнього значення), Adam як в RMSProp, також використовує середнє значення других моментів градієнтів. Зокрема, алгоритм обчислює експоненціальне ковзне середнє градієнта і квадратичний градієнт, а параметри beta1 і beta2 керують швидкістю загасання цих ковзних середніх [69].

2.6.4. Механізм уваги в сучасних нейронних мережах

Останнім часом багато різних мереж SOTA використовували ці механізми уваги, які значно покращили та уточнили результати в реальному часі.

Легка мережа та прості реалізації спростили безпосереднє включення в частину вилучення функцій згорткових нейронних мереж[50].

Модулі уваги використовуються, щоб змусити CNN навчатися та зосереджуватися на важливій інформації, а не вивчати некорисну довідкову інформацію. У випадку виявлення об'єктів корисною інформацією є об'єкти або кадрування цільового класу, які ми хочемо класифікувати та локалізувати на зображенні.

Модуль уваги складається з простого 2D-згорткового шару, MLP (у випадку уваги каналу) і сигмоїдної функції (рис. 2.16) в кінці для створення маски вхідної карти функцій.

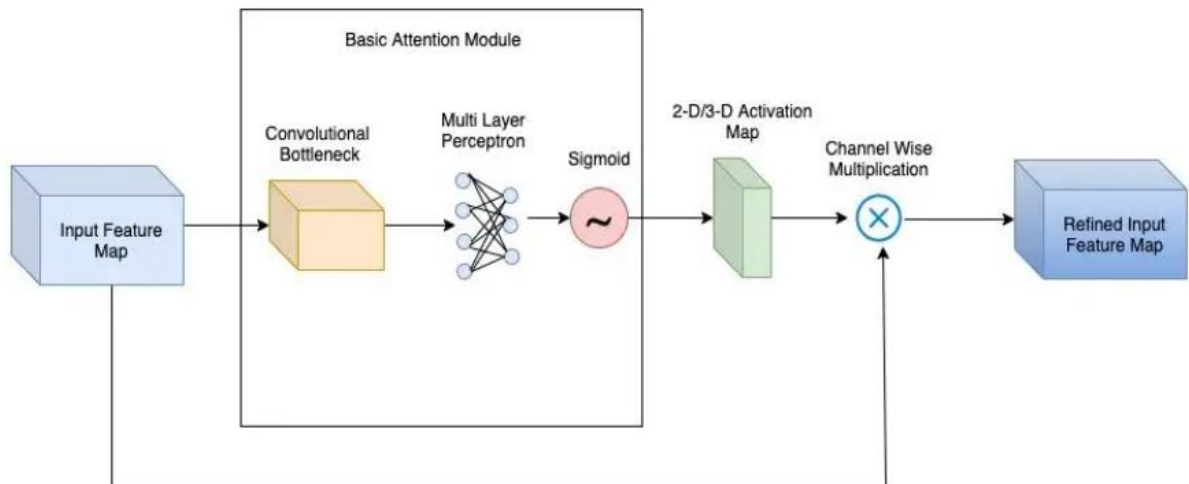


Рисунок 2.16 – Структура механізму уваги [50]

Він приймає карту ознак як вхідну інформацію та дає $1 \times H \times W$ як вихідну карту уваги. Ця карта уваги потім множиться поелементно на карту вхідних функцій, щоб отримати більш уточнений і виділений результат.

Як правило, механізми уваги застосовуються до просторових і каналних розмірів. Ці два механізми уваги, а саме генерація просторової та каналної карти уваги виконується послідовно, як показано на рис. 2.17, або паралельно.

Ці механізми уваги вперше експериментували в залишкових архітектурах. На наведеному нижче малюнку (рис. 2.17) зображено загальну структуру рівня уваги.

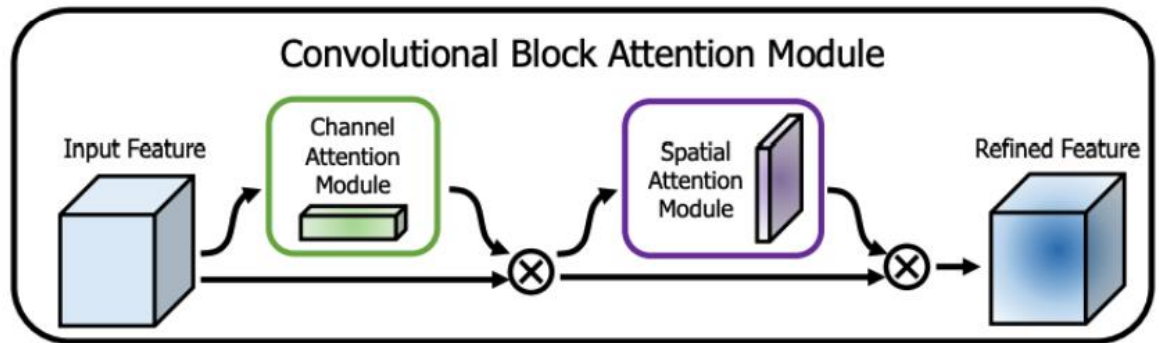


Рисунок 2.17 – Канали уваги [50]

Кроки для створення карти просторової уваги:

- Виконання глобального пулінгу, об'єднання карти ознак, після чого буде отриманий о каналний вектор $F_c \in C \times 1 \times 1$.
- Передання повнозв'язного шару одного прихованого шару розміру C/r . Тут r — коефіцієнт зменшення для прихованого каналу (Наприклад, якщо довжина вектора каналу становить 1024, а коефіцієнт зменшення r — 16, то кількість нейронів у прихованому шарі становитиме 64).
- Додавання рівня нормалізації партії перед цим MLP.

Також інсує механізм уваги з так званою технологією бутлнека (рис. 2.18)

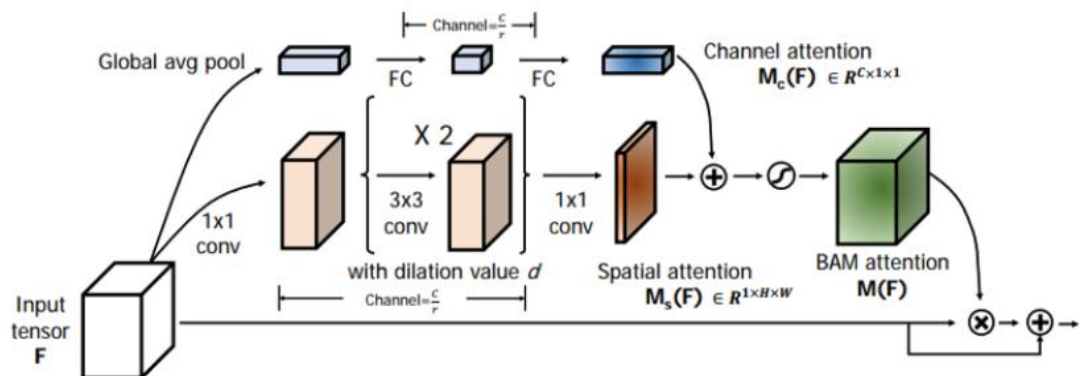


Рисунок 2.18 – бутлнек модуль механізму уваги [50]

Вхідна карта функцій F передається через ланцюжки згорткових шарів 1×1 і 3×3 . Тут згорткові шари 3×3 містять значення розширення $d=4$. Значення розширення використовується для збільшення ефективного рецептивного поля мережі. На рис. 2.19 ми бачимо ефект збільшення значення розширення [50].

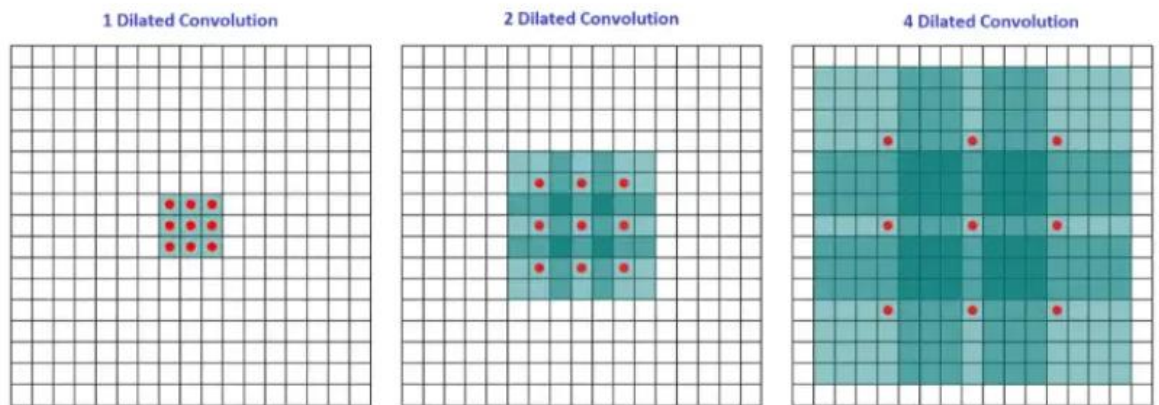


Рисунок 2.19 – рецептивне поле [50]

Висновки до розділу 2

В даному розділі було розглянуто проблематика рентгенології, зашальні методи діагностики хвороб грудної клітини. Після чого було детально проаналізовано та описано теорію нейронних мереж, їх сучасні переваги та недоліки в порівнянні з класичним машинним навчанням, сфери застосування глибокого навчання.

В передостанніх підрозділах значна увага була виділена на теорію згорткових нейронних мереж, наведено багато ілюстрацій з механіки нейронних мереж. Додатково було описано такі операції як пулінг, оптимізація, механізм уваги та функції активації в нейронних мережах.

РОЗДІЛ 3

ПРАКТИЧНА ЧАСТИНА

3.1. Опис використовуваних даних

Існує маса датасетів, що містять рентгенівські знімки легень як здорового, так і патологічного характеру. Після вивчення кількох прикладів відповідних наборів даних, перевагу було віддано датасету, підготовленому Клінічним Центром у складі Національних Інститутів Здоров'я США (NIH Clinical Center) [21]. Набір даних називається NIH Chest X-rays Dataset. Ключовим критерієм вибору стала кількість знімків, що міститься в наборі, оскільки однією з особливостей методів глибокого навчання є потреба якомога більшої кількості вихідних даних. Повну інформацію про кожен з розглянутих наборів наведено в табл. 3.1.

Таблиця 3.1.

Характеристики датасетів з рентген зображеннями

Назва датасету	Кількість	Розширення	Формат
NIH Chest X-rays dataset	112120	1024x1024	PNG
Shenzhen Hospital X-ray Set	662		JPEG
Montgomery Country chest Xray dataset	138	4020x4892	DICOM
JSRT database	247	2048x2048	PNG
Belarus tuberculosis database	464	2248x2248	DICOM

Вибраний датасет надається у вільному доступі та містить трохи більше 112 тисяч знімків. Зображення представлені у форматі PNG із глибиною кольору 8 біт. Кожен знімок має роздільну здатність 1024x1024 пікселя. Після аналізу вмісту датасету було встановлено співвідношення здорових знімків близько 60% датасету, та відповідно, зображень, що містять будь-яку патологію близько 40% набору даних.

Ось приклад здорового зображення та зображення з патологією (рис. 3.1.):

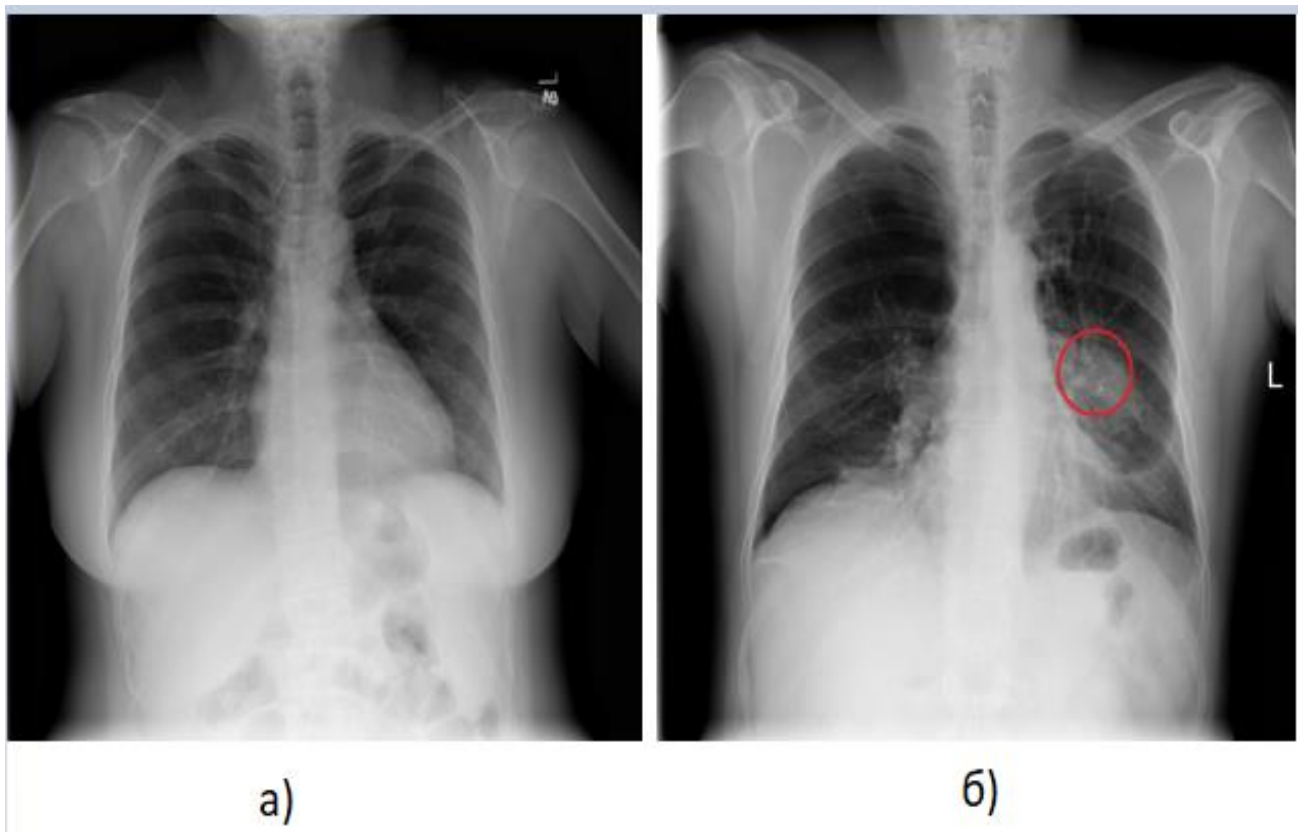


Рисунок 3.1 – приклади знімків з датасету

Далі буде проілюстровано графіки, що відображають розподіл міток та розподіл за віком в розрізі гендерного признака (рис.3.2. – рис.3.3.)

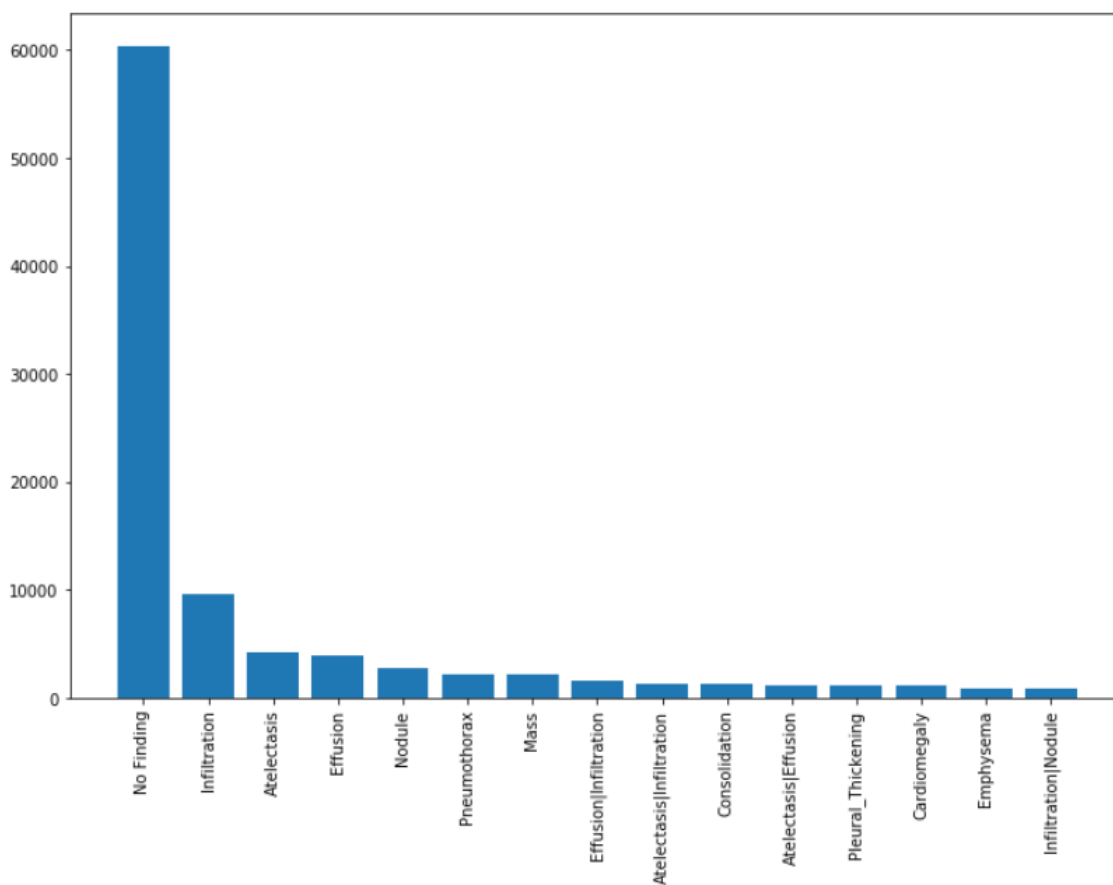


Рисунок 3.2 – розподіл міток

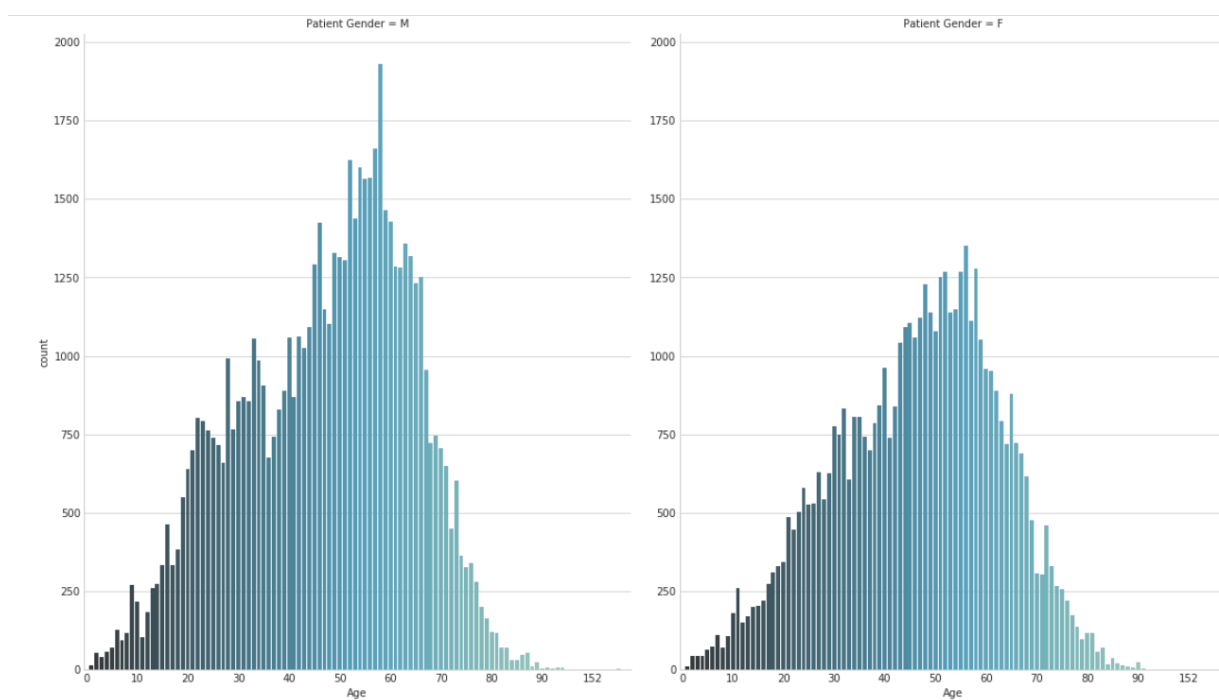


Рисунок 3.3 – розподіл за віком та гендером

3.2. Вибір мови програмування

Сьогодні існує багато мов, які використовуються для створення нейронної мережі. Далі ми розглянемо основні мови, для яких бібліотеки глибокого навчання. Мова MATLAB є інструментом, який забезпечує взаємодія оператора з різними можливостями для аналізу, збору та представлення даних. Ця мова має характерні переваги та недоліки мова, що живе в закритій екосистемі.

Переваги:

- мову легко вивчати;
- наявність готових модулів для роботи з нейронними мережами;
- часті оновлення.

недоліки:

- перевантаженість операторами та функціями;
- висока вартість програмного забезпечення;
- вузька спрямованість.

Java — це мова високого рівня, синтаксис якої базується на C++, тому Java схожа на C. Однак синтаксис Java простіший, тобто дозволяє новачкам швидше вчитися та ефективніше використовувати код досягнення конкретних результатів.

Переваги:

- безпечність;
- незалежність від платформи;
- автоматичне керування пам'яттю.

недоліки:

- багатослівний і складний код;
- немає вбудованого дизайну.

Python — мова програмування високого рівня для різного роду завдань [7].

Переваги:

- легко навчатися;
- велика ком'юніті;
- наявність різноманітних сервісів і фреймворків.

недоліки:

- складний у використанні API;
- погана продуктивність.

Для вирішення проблеми було вирішено використовувати Python. Так як ця мова має ряд істотних переваг і чудово підходить для створення згорткової нейронної мережі. Далі ми розглянемо основні фреймворки, для яких призначені працювати з нейронними мережами.

3.3. Порівняння Pytorch та Tensorflow

TensorFlow — це бібліотека з відкритим кодом, створена Google на Python і C++. Цей фреймворк є хорошим рішенням для створення комплексного проекту. Крім того, він підтримується великою спільнотою та має потужний потенціал засоби моніторингу процесу навчання моделей та візуалізації. Torch — це науково-обчислювальне середовище, яке широко підтримується алгоритми машинного навчання. Це забезпечує максимальну гнучкість і швидкість побудови наукових алгоритмів, що робить цей процес надзвичайно простий. Крім того, цей каркас поставляється з великою кількістю бібліотек, орієнтованих на машинне навчання.

Keras — це бібліотека на основі Python із можливістю запуску над TensorFlow і Theano. Keras підтримує широкий діапазон шарів нейронної мережі, такі як згорткові шари, рекурентні або щільні. Для вирішення проблеми було вирішено використовувати Keras і TensorFlow, оскільки ці фреймворки чудово підходять для впровадження нейронної мережі.

І TensorFlow, і PyTorch пропонують корисні абстракції, які полегшують розробку моделей за рахунок скорочення шаблонного коду. Вони

відрізняються тим, що PyTorch має більш «пітонячий» підхід і є об'єктно-орієнтованим, тоді як TensorFlow пропонує різноманітні варіанти.

Сьогодні PyTorch використовується для багатьох проектів глибокого навчання, і його популярність зростає серед дослідників ШІ, хоча з трьох основних фреймворків він найменш популярний. Тенденції показують, що незабаром це може змінитися.

Коли дослідники хочуть гнучкості, можливостей налагодження та короткої тривалості навчання, вони обирають PyTorch. Він працює на Linux, macOS і Windows.

Завдяки добре задокументованому фреймворку та великій кількості навчених моделей і посібників TensorFlow є улюбленим інструментом багатьох професіоналів галузі та дослідників. TensorFlow пропонує кращу візуалізацію, що дозволяє розробникам краще налагоджувати та відстежувати процес навчання. Однак PyTorch забезпечує лише обмежену візуалізацію.

TensorFlow також перемагає PyTorch у розгортанні навчених моделей у виробництві завдяки фреймворку TensorFlow Serving. PyTorch не пропонує такої структури, тому розробникам потрібно використовувати Django або Flask як внутрішній сервер.

Якщо ми почнемо говорити про розгортання, TensorFlow наразі є явним переможцем: він має TensorFlow Serving, який є платформою для розгортання ваших моделей на спеціалізованому сервері gRPC. Мобільний також підтримується.

Коли ми повернемося до PyTorch, ми можемо використовувати Flask або іншу альтернативу для кодування REST API поверх моделі. Це також можна зробити за допомогою моделей TensorFlow, якщо gRPC не підходить для вашого випадку використання. Однак TensorFlow Serving може бути кращим варіантом, якщо продуктивність викликає занепокоєння.

Tensorflow також підтримує розподілене навчання, якого PyTorch наразі не має. Критеріями відбору бібліотеки для даного дослідження були обрані:

1. Швидкість навчання (на GPU) Для кожної комбінації нейронної мережі, розміру вхідного зображення та розміру пакету було виконано 100 кроків навчання та виміряно середній час кроку навчання. Повні результати вимірювань представлені в табл. 3.1. Прочерки в таблицях відповідають такій комбінації параметрів, при якій процесу не вистачало відеопам'яті. На рис. 3.4 показані криві, що відповідають відносному часу кроку навчання фреймворків для кожної обраної комбінації вхідних параметрів (за одиницю береться найменший час кроку в цій комбінації).

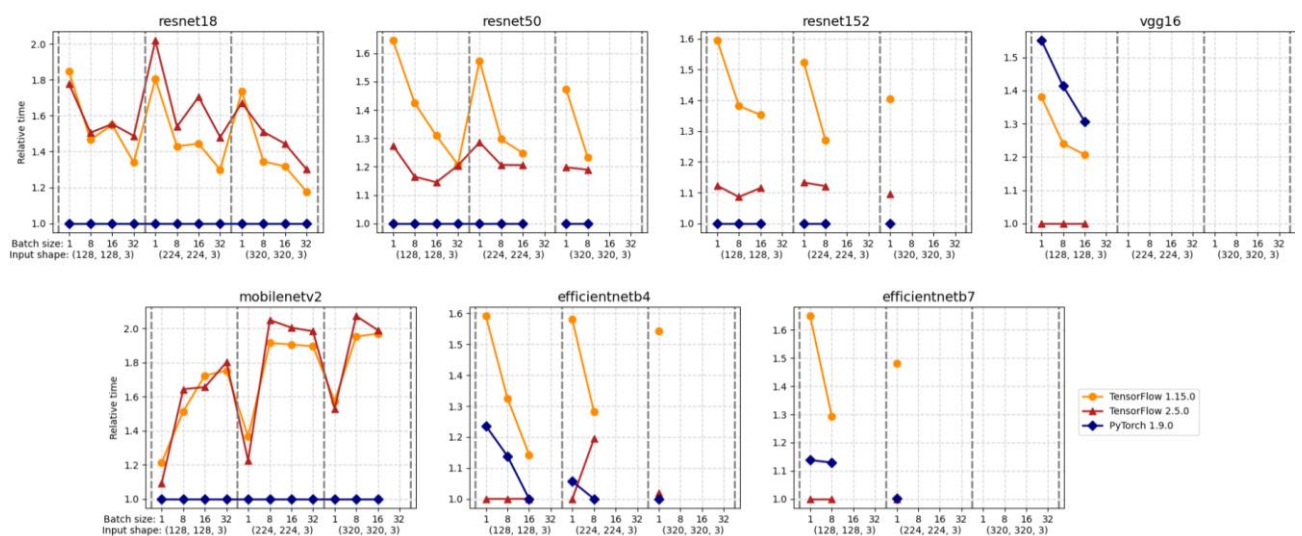


Рисунок 3.4 – Швидкість навчання на GPU

У більшості випадків PyTorch працює краще, ніж TensorFlow. Приблизно підраховано, що швидкість навчання в PyTorch 1.9.0 на 30% швидша, ніж у TensorFlow 2.5.0, і на 45% швидша, ніж у TensorFlow 1.15.0. Однак помітно, що при навчанні vgg16 і effectnet у TensorFlow 2.5.0 він показує кращу продуктивність, хоча в цілому близьку до продуктивності в PyTorch 1.9.0. Оптимізацію відеопам'яті в TensorFlow 2.5.0 покращено порівняно з TensorFlow 1.15.0, що дозволяє навчати більше нейронних мереж. Оптимізація відеопам'яті в PyTorch 1.9.0 здається трохи кращою, ніж у TensorFlow 2.5.0.

2. Швидкість висновку (на GPU) Середній час кроку виведення було виміряно подібним чином, повні результати вимірювань представлени на

рисунку кривих, що відповідають відносному часу кроку виведення фреймворків, показані на рис. 3.5.

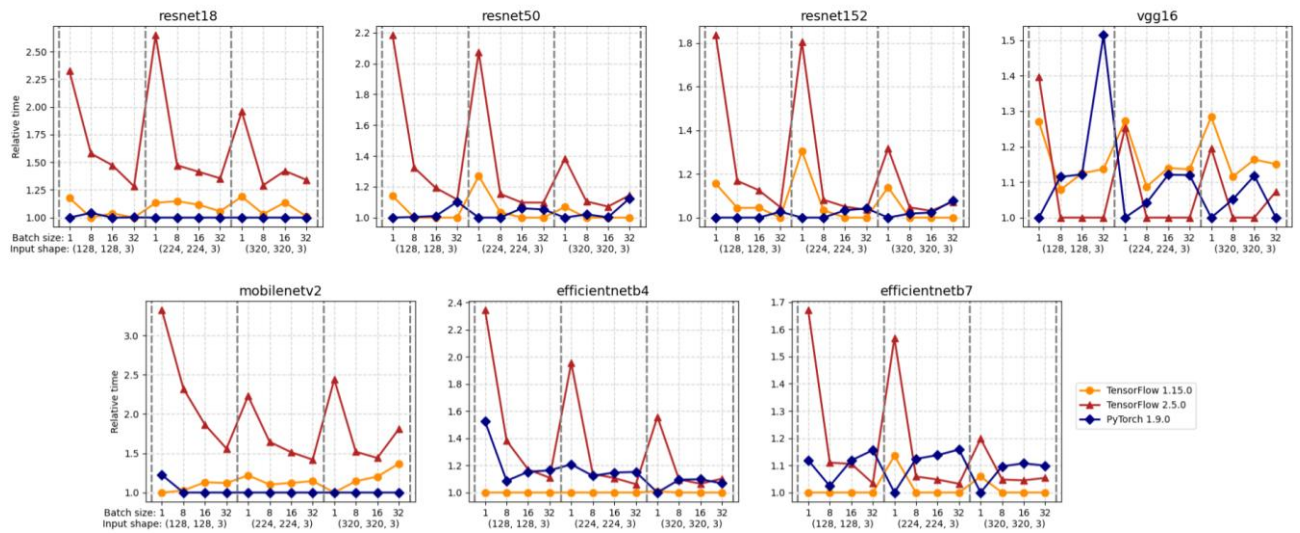


Рисунок 3.5 – швидкість інференса

Під час виведення на GPU TensorFlow 1.15.0 і PyTorch 1.9.0 демонструють приблизно однакову продуктивність: з грубою. Згідно з оцінками, PyTorch 1.9.0 в середньому на 5% швидше за використання архітектур ResNet і VGG, а TensorFlow 1.15.0 в середньому на 5% швидше за використання MobileNet та архітектури EfficientNet. У той же час TensorFlow 2.5.0 в середньому на 30% повільніше у всіх випадках. Основною причиною цього, на мою думку, є використання динамічного графа обчислень: при переході на статичний швидкість виведення в TensorFlow 2.5.0 стає майже такою ж, як і в TensorFlow 1.15.0.

3. Швидкість висновку (на ЦП) На рис. 3.6 показано криві, що відповідають відносному часовому кроку висновку рамки на CPU3.

Швидкість виведення (на ЦП) Тут результати виявилися досить несподіваними: у більшості випадків PyTorch 1.9.0 демонструє низьку продуктивність, і чим складніша нейронна мережа, тим нижча продуктивність. Приблизна оцінка швидкості виведення ЦП: TensorFlow 2.5.0 у середньому на 5% швидший, ніж TensorFlow 1.15.0, і на 40% швидший, ніж PyTorch 1.9.0.

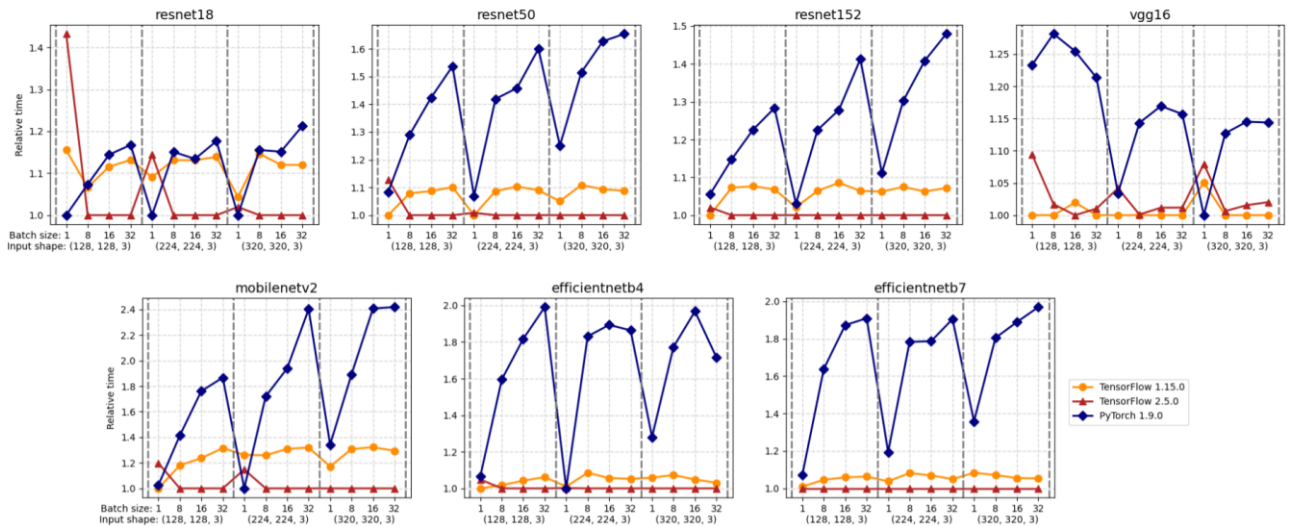
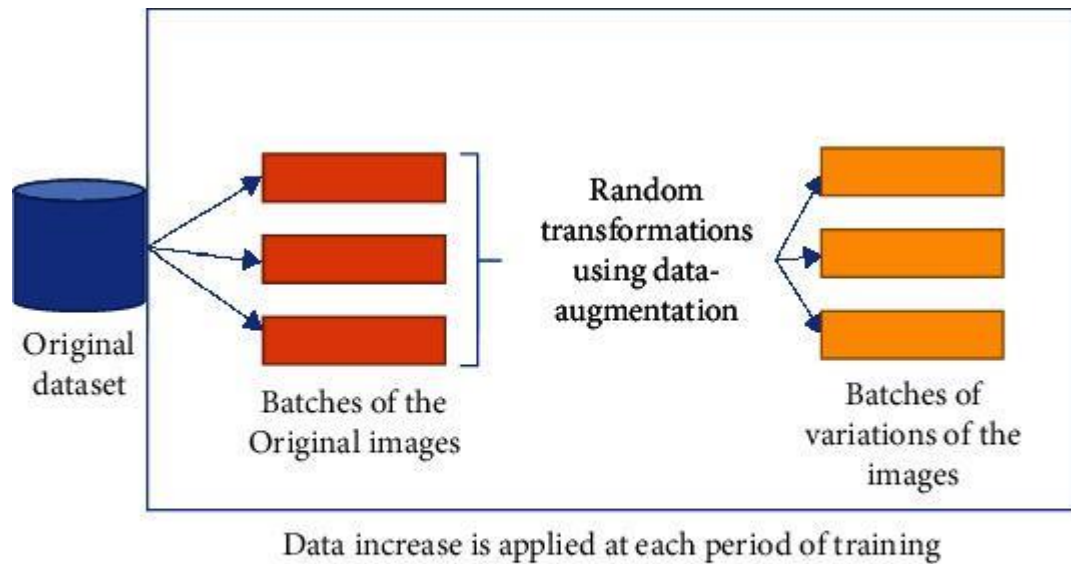


Рисунок 3.6 – швидкість інференсу на CPU

Цікаво, що під час висновку на ЦП перехід від динамічного графіка до статичного в TensorFlow 2.5.0 не впливає на швидкість. В результаті цих даних було вирішено обрати бібліотеку tensorflow.

3.4. Аугментація даних

Збільшення даних Збільшення даних полягає у створенні більшої кількості навчальних даних (у кількості та/або різноманітності) з існуючих навчальних зразків, збільшення зразків за допомогою ряду випадкових перетворень, які, як правило, включають обертання, пов'язані з перетвореннями, переклади та масштабування тощо. Під час навчання мета полягає в тому, щоб модель ніколи не бачила точно те саме зображення двічі і, таким чином, краще узагальнювала, відоме як динамічне збільшення даних (рис. 3.7.)



The model is always trained on different data

Рисунок 3.7 – Пайплайн аугментації

Доповнення даних за допомогою класу `ImageDataGenerator`. Клас `ImageDataGenerator` також дозволяє застосовувати доповнення даних. Доповнення Keras зосереджено на створенні різноманітності даних, що відображаються в моделі від епохи до епохи, зберігаючи вихідну кількість даних. Створюючи екземпляр класу `ImageDataGenerator`, визначаються перетворення, а потім, під час навчання, збираються пакети із зображеннями, перетвореними в реальному часі. Він прагне максимально використати навчальні приклади, щоб модель ніколи не бачила точно те саме зображення двічі; це допомагає уникнути переобладнання та краще узагальнює модель.

Щоб реалізувати розширення даних, була визначена функція таким чином, щоб вона повертала примірник до класу `ImageDataGenerator`, налаштованого на виконання перетворень даних. Враховуючи попередні статті, в яких аналізувалися рентгенограми органів грудної клітки, були визначені такі трансформації:

- Довільне обертання в діапазоні $-5/+5$ градусів
- Випадковий горизонтальний і вертикальний зсув на 5%
- Випадковий діапазон збільшення 15%

- Випадкове спотворення вздовж осі на 0,1 градуса створює або виправляє кути сприйняття
- Адитивний гаусівський шум, випадкове розмиття.

Відступ для збереження постійного розміру вхідних зображень, дзеркальне відображення сусідніх пікселів, щоб заповнити проміжки відсутніми пікселями

Запрограмована функція використовує метод `flow_from_dataframe` для застосування перетворень, визначених генератором, і збирання пакетів зображень. Динамічне збільшення даних було застосовано під час навчання моделі шляхом виклику функції в кожній епосі.

Щоб правильно оцінити здатність моделі до узагальнення, до набору даних перевірки не було застосовано розширення даних, тому довелося запрограмувати іншу функцію за допомогою іншого генератора без включення параметрів перетворення, згаданих вище. У цьому випадку реалізація використовує метод `flow_from_dataframe` перед навчанням класифікатора для створення пакетів зображень.

У результаті ми можемо виділити дві основні категорії природного упередження. Перший — це афінні перетворення, які можуть генерувати штучні приклади різного положення пацієнта, природного нахилу пацієнта та відповідних зрушень легеневого поля. Горизонтальний переверот зображення призводить до дзеркального відображення серця й аорти, що відповідає рідкісному стану *situs inversus*, і.г. синдром дзеркальних органів. Другий — це перетворення кольорів, які можуть генерувати штучні приклади, пов'язані з різними властивостями рентгенівських сканерів, різною відстанню між тілом пацієнта та рентгенівським сканером через положення AP – PA (рис.3.8.).

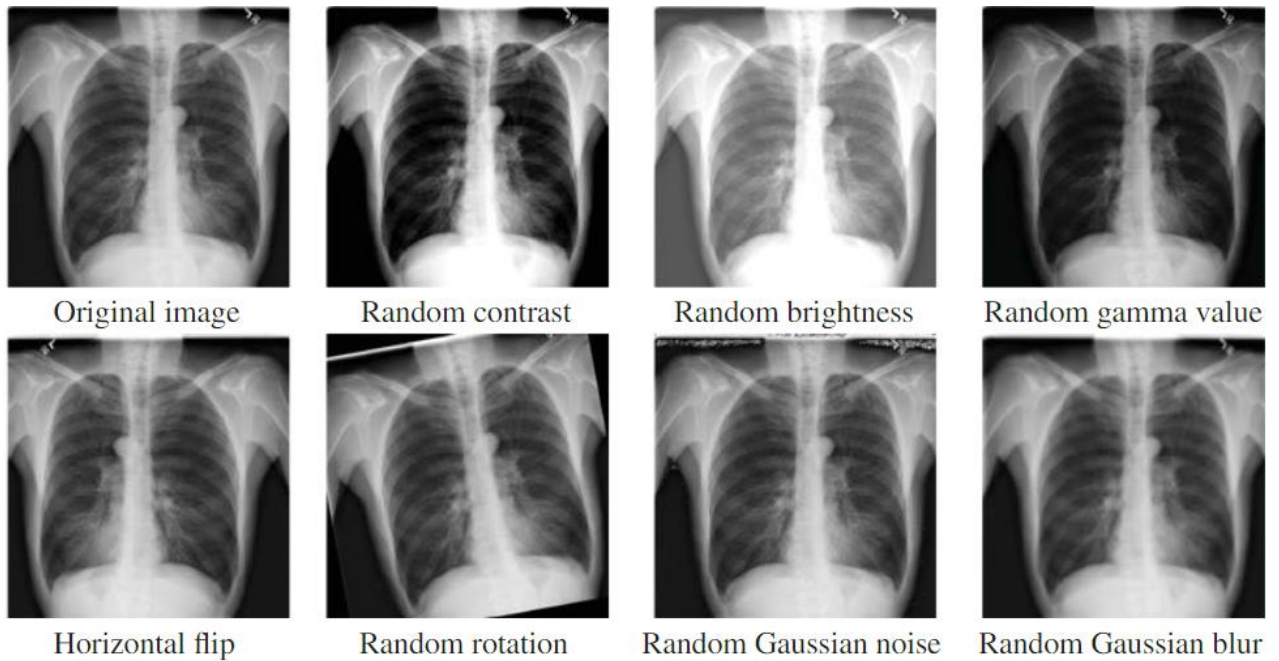


Рисунок 3.8 – Приклад уагментації

В деяких сучасних статтях наведено порівняння AUC значень, та було встановлено що контраст, гаусівський шум і розмиття продемонстрували гіршу середню продуктивність. Однак випадкові повороти, горизонтальне переверот і яскравість показали середню ефективність більше ніж на 1 % приріст, а їх комбінована модель показала найкращий середній показник ефективності.

3.5. EfficientNet як основна моделі

Нещодавно Google AI Research опублікував статтю під назвою «EfficientNet: переосмислення масштабування моделі для згорткових нейронних мереж». У цій статті автори пропонують нову архітектуру, яка забезпечує найсучаснішу точність класифікації в ImageNet, водночас у 8,4 раза менша та в 6,1 раза швидша за висновками, ніж найкраща існуюча CNN. Він досягає високого рівня точності на багатьох інших наборах даних, таких як CIFAR-100, квіти та автомобілі. Хороші результати на кількох наборах даних показують, що архітектуру можна використовувати для перенесення навчання.

Наївний спосіб підвищити продуктивність нейронних мереж – це поглибити CNN. Чудовим прикладом може бути resnet which, який має кілька варіацій від 18 до 202. Зроблення CNN глибшим або ширшим може підвищити продуктивність, але це вимагає великих обчислювальних витрат. Отже, нам потрібен якийсь спосіб збалансувати наші постійно зростаючі пошуки продуктивності та витрати на обчислення. У статті автори пропонують новий метод масштабування моделі, який використовує простий складений коефіцієнт для масштабування CNN у більш структурований спосіб. Цей метод допомагає їм вирішити, коли збільшити глибину чи ширину мережі.

Автори хотіли оптимізувати для точності та ефективності. Отже, вони виконали пошук нейронної архітектури. Цей пошук дав архітектуру Efficient-V0, яка виглядає досить простою та зрозумілою для реалізації (рис. 3.9).

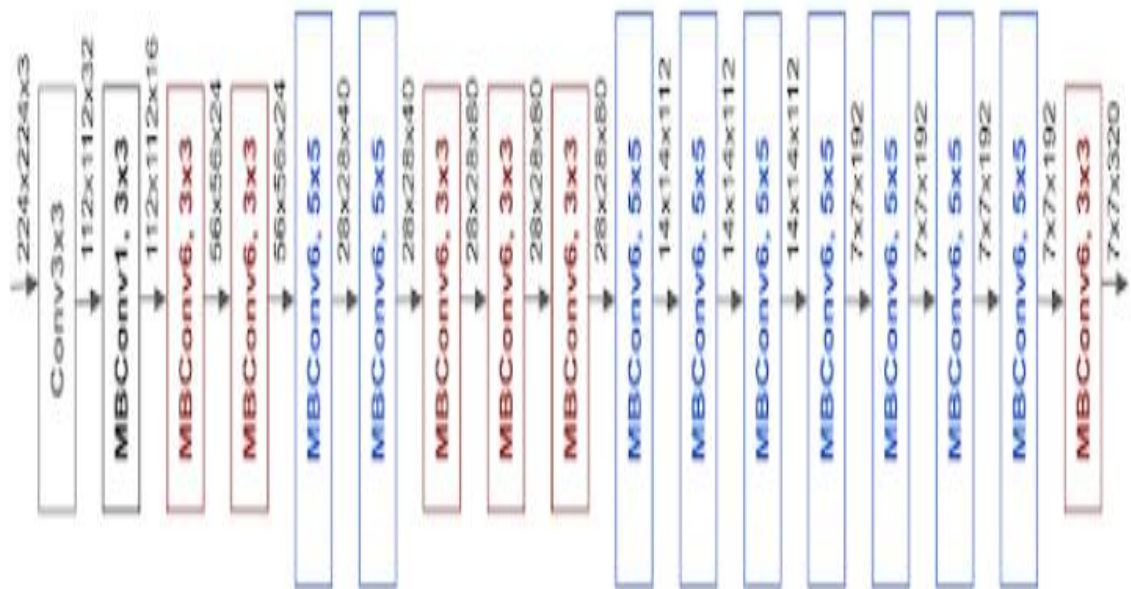


Рисунок 3.9 – Приклад архітектури ЕфіншНет

Як видно з графіка продуктивності, EfficientNet використовує менше параметрів і досягає дуже високої точності. Для отримання додаткової інформації зверніться (рис.3.10).

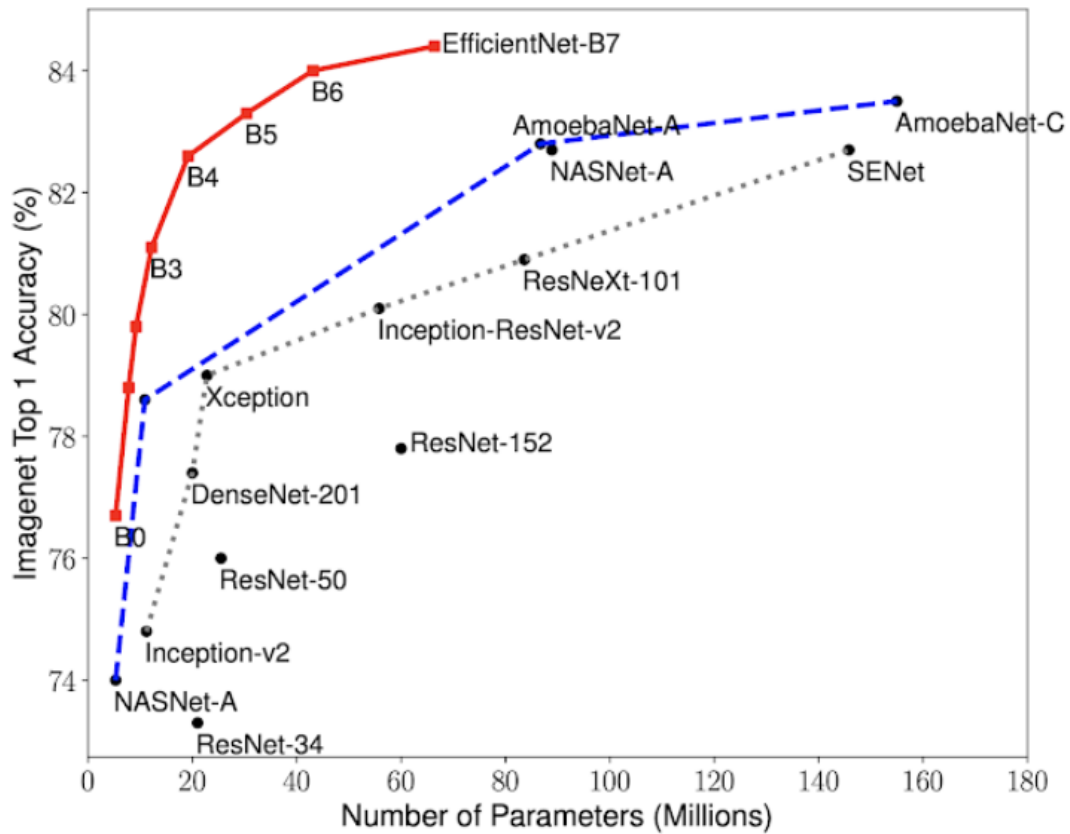


Рисунок 3.10 – Порівняння моделі з іншими

Проведемо імпорт моделі в нашій бібліотеці, та глянемо на кількість тренувальних та нетренувальних параметрів, також до моделі були додані останні повнозв'язні шари, які саме і роблять класифікацію (рис. 3.11).

```

Downloading data from https://github.com/Callidior/keras-applications/1
oaugment_notop.h5
44113920/44107200 [=====] - 2s 0us/step
Model: "sequential"

```

Layer (type)	Output Shape	Param #
efficientnet-b3 (Model)	(None, 8, 8, 1536)	10783528
global_average_pooling2d (G1)	(None, 1536)	0
dense (Dense)	(None, 1024)	1573888
dense_1 (Dense)	(None, 14)	14350

```

Total params: 12,371,766
Trainable params: 12,284,470
Non-trainable params: 87,296

```

Рисунок 3.11 – Імпорт моделі

3.6. Боротьба з дисбалансом класів

Однією з проблем роботи з медичними діагностичними наборами даних є великий дисбаланс класів, присутній у таких наборах даних. Давайте побудуємо графік частоти кожної з міток у нашому наборі даних:

На цьому графіку ми бачимо, що поширеність позитивних випадків суттєво різниться в залежності від різних патологій. (Ці тенденції також відображають тенденції в повному наборі даних.)

- Патологія грижі має найбільший дисбаланс з часткою випадків позитивного тренування близько 0,2%.

- Але навіть інфільтраційна патологія, яка має найменшу кількість дисбалансу, має лише 17,5% випадків навчання, позначених як позитивні.

В ідеалі ми б навчили нашу модель, використовуючи рівномірно збалансований набір даних, щоб позитивні та негативні випадки навчання однаково сприяли втратам.

Вплив дисбалансу класів на функцію втрат. Давайте розглянемо це ближче. Припустимо, що ми використали б нормальну втрату крос-ентропії для кожної патології.

Вплив дисбалансу класів на функцію втрат. Давайте розглянемо це ближче. Припустимо, що ми використали б нормальну втрату крос-ентропії для кожної патології. Використовуючи формули крос-ентропії, ми можемо побачити, що якщо існує великий дисбаланс із дуже малою кількістю позитивних випадків навчання, наприклад, тоді у втраті домінуватиме негативний клас (рис 3.12).

Проблема дисбалансів класів гостро стоїть в таких сферах як медицина, телекомунікації, банківська сфера, астронавтика. Для того, щоб зрозуміти проблему глибше, уявімо, що в нашому датасеті, 99 відсотків хворі на туберкульоз. Тому, наша модель може постійно видавати результат, що в хворого туберкульоз і вона буде права в 99 відсотків випадків. Але інші захворювання вона детектити не зможе. Основною проблемою такої поведінки

є дисбаланс класів в тестовому наборі, який в свою чергу може не відобразити генеральної сукупності

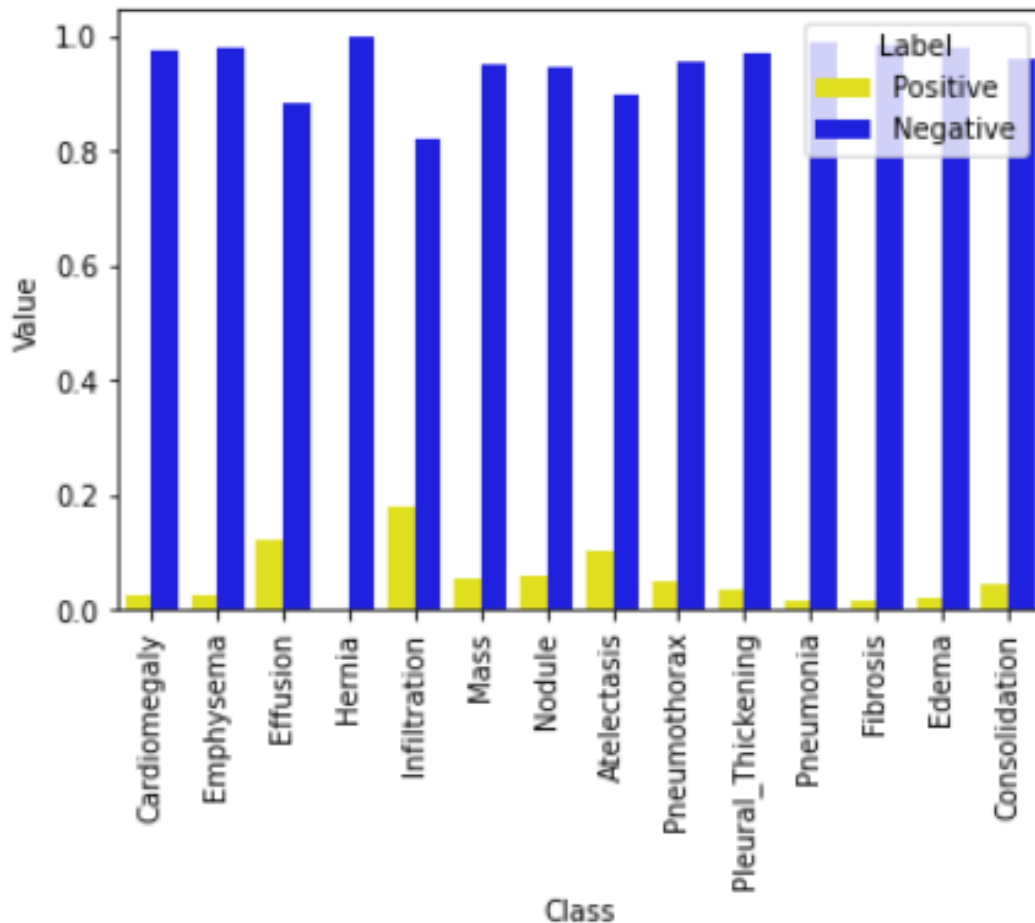


Рисунок 3.12 – Баланс класів без обробки

Як ми бачимо на наведеному вище графіку, внесок позитивних випадків значно нижчий, ніж негативних. Але ми хочемо, щоб внески були рівними. Одним із способів зробити це є множення кожного прикладу з кожного класу на специфічний для класу ваговий коефіцієнт, w_{pos} і w_{neg} , щоб загальний внесок кожного класу був однаковим.

Якщо ми використовуємо звичайну функцію перехресних ентропійних втрат із дуже незбалансованим набором даних, як ми бачимо тут, тоді алгоритм матиме стимул надавати пріоритет мажоритарному класу (тобто негативному в нашому випадку), оскільки він більше сприяє втратам.

Тепер давайте збалансуємо класи, а потім використаємо їх для візуалізації (рис. 3.13)

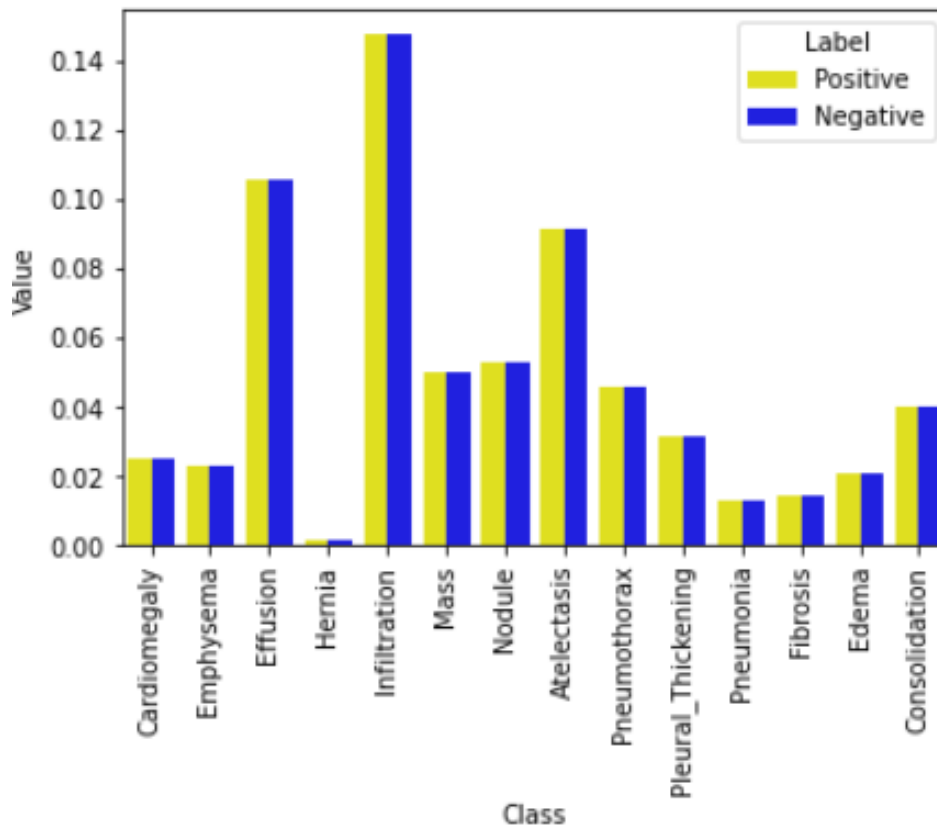


Рисунок 3.13 – Баланс класів після обробки

3.7. Тренування моделі.

Коли створили нейронну мережу, ми випадково ініціалізуємо наші ваги. Очевидно, що це не дасть вам дуже хороших результатів. У процесі навчання ми хочемо почати з нейронної мережі, що працює погано, і закінчити мережею з високою точністю. З точки зору функції втрат, ми хочемо, щоб наша функція втрат була значно нижчою в кінці навчання. Покращення мережі можливо, тому що ми можемо змінити її функцію, налаштувавши ваги. Ми хочемо знайти іншу функцію, яка працює краще, ніж початкова.

Задача навчання еквівалентна задачі мінімізації функції втрат. Навіщо мінімізувати втрати замість максимізації? Виявляється, втрату оптимізувати набагато легше.

Існує багато алгоритмів оптимізації функцій. Ці алгоритми можуть базуватися на градієнті чи ні, у тому сенсі, що вони використовують не лише інформацію, надану функцією, а й її градієнт. Один із найпростіших алгоритмів на основі градієнта — називається. Стохастичний градієнтний спуск. В даній роботі мережа оптимізувалася за допомогою пакетного градієнтного спуску, графік навчання можна побачити на рис. 3.14.

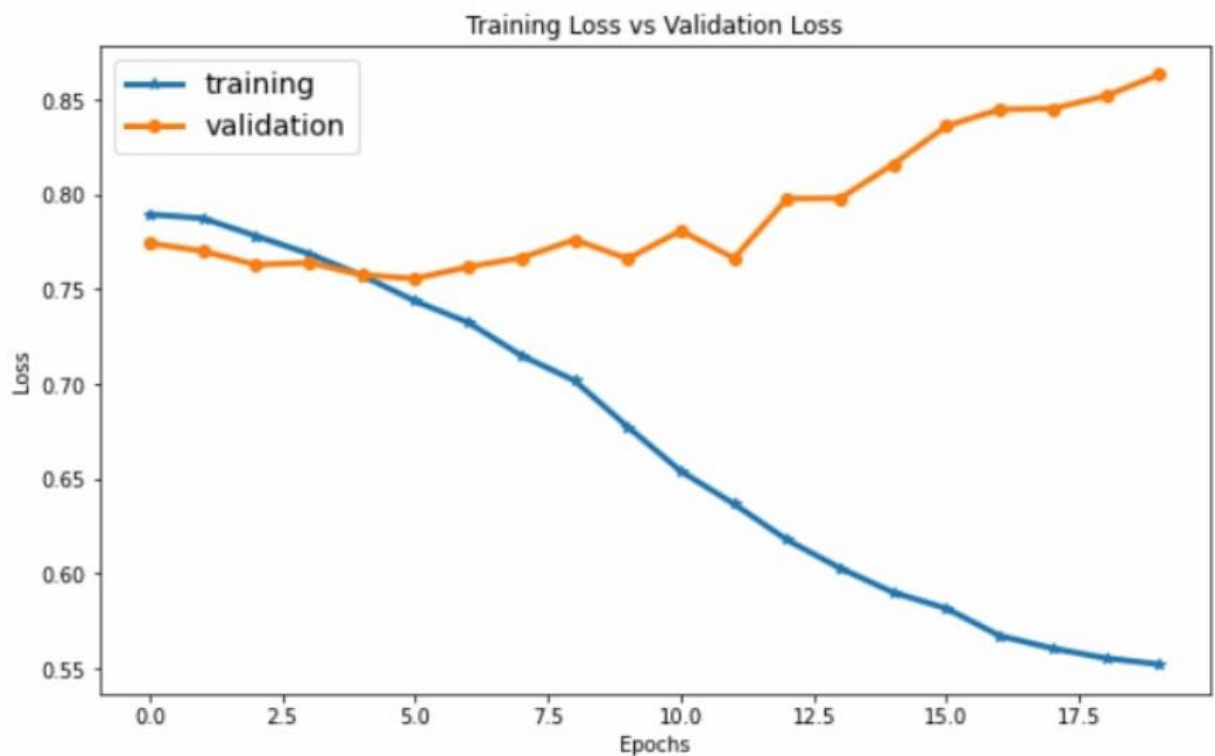


Рисунок 3.14 – Тренування нейронної сітки

Як видно на рисунку помилка на тренувальних даних постійно зменшується, це означає що наша мережа вивчає наші дані, але найбільш цікавою для нас на даному графіку є точка мінімальної похибки на валідаційних даних, як ми бачимо з графіку, мінімальна похибка на валідаційних даних була на 6 епосі, тому потрібно зберегти ваги саме цієї моделі і використовувати її реальних задачах з новими даними. Ваги нейронної мережі в бібліотеці тензофлуо можна зберегти у спеціальному форматі “h5”, задля цього варто скористатися функцією `tf.save`.

3.8. Тестування моделі.

Після збалансування класів було проведено навчання моделі EfficientNet-V4 після чого було проведено тестування (рис 3.15).

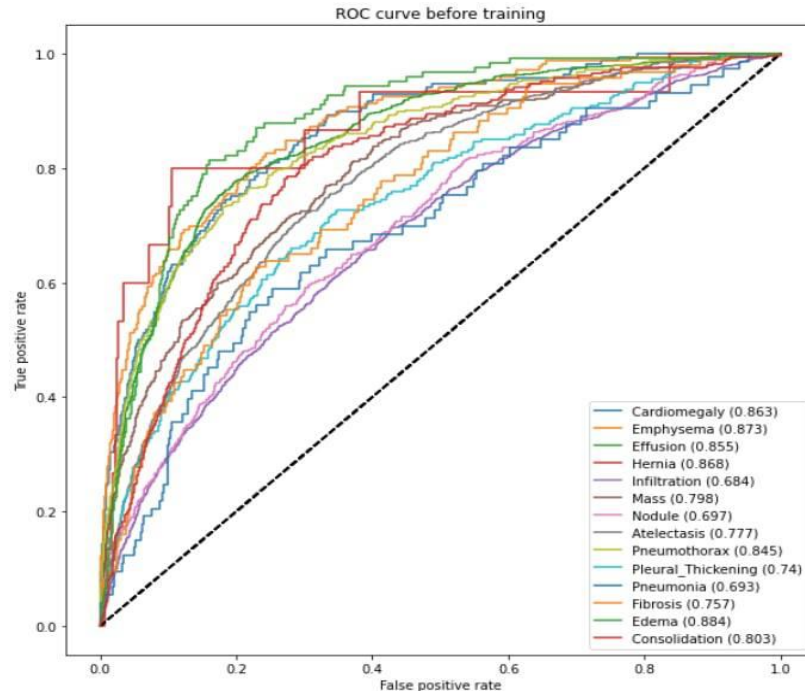


Рисунок 3.15 – РОК-АУК крива

Після цього ваги моделі з найкращими показниками були збережені у форматі h5 та протестовані на тестовій вибірці. Головною метрикою для оцінки мультикласифікації стала AUC метрика, на цю метрику орієнтуються більшість досліджень даної проблеми мультикласифікації. Результати ROC-AUC кривої наведено на (рис. 3.15). З рисунка видно що середнє значення АУК метрики більше ніж 80 відсотків

Задля візуальної перевірки нашої сітки напишемо окремий скрипт, який буде візуалізувати (рис. 3.16) такі параметри: картинку грудної клітки пацієнта, його реальну хворобу “Dx”, та той результат що предиктить нейронка “PDx”.



Рисунок 3.16 – Візуалізація предікту

Як видно з даного рисунку, на 4 прикладах, те що предіктила нейронка і та хвороба яка справді була в пацієнта – сходилися. Слід зауважити, що попередньо ми вихід нейронки де останній слой був софтмакс форматували у вірогідність за допомогою спеціального модуля SciPy.

Висновки до розділу 3

В даному розділі було описано попередню обробку і аугментацію даних , імпорт моделі та її тестування після тренування. Також було проведено аналіз датасету з яким працюю. Велика увага зосередилася на виборі інструментів.

В результаті даного дослідження можна зробити висновок, що основною

проблемою з якою треба боротися в рамках даного дослідження – дисбаланс класів. Для боротьби була написана спеціальна функція помилки, що враховувала вагу кожного класу з датасету, та надавала завчати лише ті хвороби, кількість картинок яких найбільше в датасеті.

В ході роботи була використана одна нейронних мереж з передового класу EfficientNet.

РОЗДІЛ 4

РОЗРАХУНОК СТАРТАП ПРОЕКТУ

4.1. Опис ідеї проекту

Останнім часом все частіше обговорюється застосування AI у медицині. І раніше можна було застосовувати експертні системи та алгоритми класифікації до завдань постановки діагнозу. Однак, є одна область AI, яка досягла найбільших успіхів в останні роки, а саме область розпізнавання зображень та згорткові нейронні мережі. На деяких тестах алгоритми AI у розпізнаванні картинок перевершили людину. Ось два приклади: Large Scale Visual Recognition Challenge та German Traffic Sign Recognition Benchmark.

Відповідно, виникла ідея застосувати AI до області розпізнавання зображень там, де лікарі займаються розпізнаванням зображень, а саме до аналізу знімків і, для початку, рентгенівських знімків.

Рентгенографія грудної клітки – це класичне дослідження, частіше проведене в процесі скринінгу населення для діагностики патологічних змін у грудній клітці. За допомогою цього методу можуть бути діагностовано або виявлено підозри на такі часто діагностовані захворювання:

- Пневмонія,
- Туберкульоз,
- Пневмоторакс,
- Інтерстиціальні захворювання легень,
- Серцева недостатність (набряк легень),
- Грижа стравохідного отвору діафрагми.

Виходячи з усього вищесказанного виникла ідея розробити програмне забезпечення, яке зможе діагностувати захворювання по рентгену грудної клітини. На західному ринку та в Індії є схожі продукти, які за рентгенограмою вміють детектувати рак на ранній стадії, та інші відхилення в грудній клітині.

Даний продукт може включати перевагу, підвантаження ні тільки одного знімку, а цілої папки з картинками та отримання результату для рядку знімків.

4.2. Організація проекту

Назва проекту: Нейронна мережа для виявлення відхилень на рентгенівських знімках грудної клітини.

Організація проекту створена на базі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського». Можливий пошук клінік, медичних закладів які виступатимуть партнерами.

4.3 Канва бізнес-моделі проекту

Канва бізнес-моделі дозволяє сформулювати стратегічне управління проекту (табл.4.1).

Таблиця 4.1

Канва бізнес-моделі проекту

(8) КЛЮЧОВІ ПАРТНЕРИ Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського, рентген відділення	(4) КЛЮЧОВІ ВИДИ ДІЯЛЬНОСТІ Дослідно-технологічна робота, обробка рентгенологічного зображення	(5) ЦІННІСНІ ПРОПОЗИЦІЇ Можливість діагностувати одні з найпоширеніших хвороб грудної клітини на рентгенівському знімку. Також лікарі зможуть бачити вірогідність тієї чи іншої хвороби, враження.	(6) ВЗАЄМОВІДНОСИНИ ЗІ СПОЖИВАЧАМИ B2B лікарі можуть обробляти дані своїх пацієнтів навіть віддаленно	(5) СПОЖИВЧІ СЕГМЕНТИ Люди з підозрою на туберкульоз, пневмонію та інші хвороби грудної клітини Лікарі рентгенологічного відділення
	(6) КЛЮЧОВІ РЕСУРСИ Технології розробки програмних продуктів, та команда розробників.		(6) КАНАЛИ ЗБУТУ Презентація на медичних конференціях, прямі емейли на пошту рентгенологічних відділень, клінік	

Продовження таблиці 4.1

<p align="center">(7) СТРУКТУРА ВИТРАТ</p> <p>Витрати: підтримка серверу, заробітна плата робітникам, оренда офіса</p>	<p align="center">(8) ПОТОКИ НАДХОДЖЕННЯ ДОХОДІВ</p> <p>Грошовий потік буде спрямований від медичних установ, приватних клінік</p>
---	---

4.4. Ключові види діяльності

4.4.1. Вид проекту за характером інновації

За характером інновацій даний проект можна ідентифікувати як запровадження нової технології та дослідно-технологічна робота.

Вид проекту за характером інновації Вид проекту за характером інновації – дослідно-технологічна робота. Ключовим видом діяльності стартап-проекту буде розробка прикладного програмного застосунку для діагностування поширених хвороб грудної клітини за рентгенвським знімком.

В даний час пневмонія - одне з найсерйозніших легеневих захворювань у всьому світі. За даними Всесвітньої організації охорони здоров'я щорічно повідомляється майже про два мільйони смертей від пневмонії.

Тяжка пандемія коронавірусу COVID-19 з початку 2020 року ще більше посилила летальний кінець легеневих захворювань. Численні клінічні дослідження підтвердили, що інфекція COVID-19 викликає важку форму пневмонії у великої кількості людей. У той же час бактеріальні та вірусні патогени викликають різні форми пневмонії, що потребує різних підходів до лікування. І хоча бактеріальна пневмонія негайно лікується антибіотиками, вірусний тип цього захворювання вимагає екстреної допоміжної допомоги. робить критично важливою своєчасну та точну діагностику пневмонії. Окрім пневмонії програмне забезпечення зможе виділяти ще 13 типів відхилень грудної клітини таких як: ефузії, кардіомегалія, ателектаз і тд.

4.4.2. Спрямованість проекту

- випуск ПЗ, конкурентоспроможного на світовому ринку
- збільшення продуктивності та поліпшення умов праці
- поліпшення ефективності діагностики та лікування хворих

Сфера медицини перебуває у постійному пошуку нових способів діагностики, лікування, моніторингу і зниження показників смертності пацієнтів, тому технологія зменшить час та збільшить точність діагностування.

4.4.3. Висновок щодо науково-технічного рівня ідеї

Проаналізувавши ринок подібного програмного забезпечення, було визначено, що даний проєкт немає аналогів в Україні та має аналоги на Заході.

4.4.4. Основні бізнес процеси проєкту

В ході роботи було встановлено, що основними бізнес процесами буде робота з даними, інженерна складова продукту, розробка нейронних мереж, аугментації, тренування та тестування даних моделей на тестових вибірках. Також окремо буде потрібно розділ з деплоєм продукту. В таблиці 4.2 наведено опис основних бізнес-процесів проєкту.

Таблиця 4.2

Бізнес-процеси проєкту

Група процесів	Бізнес-процес	Ступінь опрацювання бізнес-процесу	
		є реалізованим	буде реалізованим
Препроцесінг даних	Пошук даних	+	
	Імпорт даних	+	
	Попередня обробка даних, аугментація даних	+	
Написання нейронної мережі	Підбір архітектур	+	
	Реалізація архітектур		+
	Навчання та тестування нейронної мережі на нових даних, що призведе до зменшення похибки на тестових даних		+

Продовження таблиці 4.2

Розробка додатку	Реалізація бекенду ПЗ		+
	Реалізація інтуїтивного інтерфейсу		+
	Можливість підгружати декілька зображень одночасно і отримувати результати		+
Обслуговування споживача	Підтримка користувачів		+
	Вдосконалення продукту		+

4.5. Ціннісні пропозиції та споживачі

Ціннісна пропозиція – сукупність переваг, які проект може запропонувати споживачу.

4.5.1. Характер формування споживчої цінності проекту

Основна цінність проекту заключається в підтримці лікарів-рентгенологів. Система буде допомагати поставити діагноз. Це в свою чергу може зменшити вплив людського фактору та пришвидшити діагностику.

4.5.2. Зміст ідеї проекту

Основна ідеологія стартапу, сфера застосування, сегменти споживачів, та основна цінність даного проекту відображена в табл. 4.3.

Таблиця 4.3

Зміст ідеї проекту

<i>Зміст ідеї</i>	<i>Напрямки застосування</i>	<i>Сегменти споживачів</i>	<i>Цінність для споживачів</i>
Автоматична діагностика відхилень на рентгнівському знімку грудної клітини	1. Медицина	Лікарі-рентгенологи пацієнти	Підвищення точності діагностики, пришвидшення встановлення діагностики, зменшення людського фактору

Продовження таблиці 4.2

за допомогою нейронних мереж	2. Наука	Дослідники , науковці	Внесення додаткових знань та результатів дослідження в наукову базу
------------------------------	----------	-----------------------	---

4.5.3. Аналіз ідеї проекту

В цьому підпункті було проаналізовано схожі стартапи, виявилось що є лише декілька аналогів зі схожими задачами, один із них «ChestLink».

Слід зауважити, що в зв'язку з епідемією попит на схожі продукти зріз і багато передвоих компаній зайнялися ресьорсчем в даній темі і на ринок за 1-2 останніх роки вийшло багато гравців, але все ж залишаються вільні ніші, та функціонал який ще нереалізований у цих продуктах.

Характеристика сторін проудкту наведена в табл. 4.4.

Таблиця 4.4

Визначення сильних, слабких та нейтральних характеристик проекту

<i>№</i>	<i>Техніко-економічні характеристики ідеї</i>	<i>W (слабка сторона)</i>	<i>N (нейтральна сторона)</i>	<i>S (сильна сторона)</i>
1.	Використання методів штучного інтелекту і машинного навчання	–	–	+
2.	Наявність даних	–	+	–
3.	Створення програмного забезпечення без фінансування	+	–	–
4.	Можливість похибки моделей	+	–	–

4.5.4. Технологічний аналіз

В процесі реалізації проекту було застосовано фреймворки машинного навчання, деталі можна бачити в таблиці 4.5.

Таблиця 4.5

Технічна складова проекту

<i>№</i>	<i>Ідея проекту</i>	<i>Технології її реалізації</i>	<i>Наявність технологій</i>	<i>Доступність технологій</i>
1.	Побудова моделей	Python	Наявна	Доступна
2.	Розробка інтерфейсу додатку, щоб скинути данні і отримати результат	Бібліотек а Django	Наявна	Доступна
<i>Обрана технологія реалізації ідеї проекту: Python, Django, Pytorch, Pandas, sklearn</i>				

4.5.5. SWOT-аналіз проекту.

За допомогою методики SWOT-аналіз було визначено його сильні сторони проекту, його слабкі сторони можливості розвитку та різні загрози (таблиця 4.6)

Таблиця 4.6

SWOT-аналіз проекту

<i>Сильні сторони (S): Використання сучасних методів штучного інтелекту</i>	<i>Слабкі сторони (W): Відсутність українських даних</i>
<i>Можливості (O): Отримання ПЗ, яке зможе пришвидшити роботу та підвищити точність</i>	<i>Загрози (T): Недонавченість чи перенавченість моделей, неправильна класифікація хвороби</i>

4.6. Взаємовідносини зі споживачами та канали збуту

Було проведено аналіз майбутнього споживчого сегменту, їх поведінкові характеристики, вимоги до продукту та канали збуту, результати наведено в (таблиці 4.7)

Таблиця 4.7

Взаємовідносини між споживачами та канали збуту

<i>№</i>	<i>Сегмент споживачів</i>	<i>Особливості поведінки</i>	<i>Вимоги споживачів</i>	<i>Канали збуту</i>
1.	Медичні заклади, лікарі-рентгенологи	Для адекватної поведінки користувачів необхідно: - просто логіка інтерфейсу - відсутність перебоїв - наявність туторіала	- до ПЗ: 1) Інтуїтивність додатка 2) Наявність документації 3) Наявність підтримки	Прямий: Листи до медичних закладів, продаж ПЗ на вебсайті Непрямий: Реклама на медичних конференціях

4.7. Обґрунтування ресурсів та витрат проекту

4.7.1 Визначення ціни

Було проведено аналіз на товари зі схожим функціоналом та продукти аналоги, на основі чого було розраховано орієнтовну ціну за продукт, що зазначено в табл 4.8

Таблиця 4.8

Визначення ціни

<i>№</i>	<i>Рівень цін на товари-замінники</i>	<i>Рівень цін на товари-аналоги</i>	<i>Рівень доходів цільової групи споживачів</i>	<i>Верхня та нижня межі встановлення ціни на продукцію</i>
1.	7000 - 10000 грн	12000 - 30000 грн	10000-20000 грн	Нижня межа: 10000грн Верхня межа: 15000грн

Розрахункова ціна продукції орієнтовно становить 10 тис грн.

4.7.2. Визначення обсягу виробництва продукції

Результати розрахунку обсягів виробництва на найближчі 5 років наведено в табл. 4.9. За розрахункам загальна потреба в даному продукті буде високою на наступні 5 років. Обсяги виробництва можуть зростати з кожним роком.

Таблиця 4.9

Визначення обсягу виробництва

<i>Показник</i>	<i>Значення по роках</i>				
	<i>2022</i>	<i>2023</i>	<i>2024</i>	<i>2025</i>	<i>2026</i>
<i>Загальна потреба в продукції</i>	висока	висока	висока	висока	висока
<i>Можливі річні обсяги випуску в натуральних показниках</i>	100	400	800	1300	1800
<i>Ціна одиниці продукції (тис. грн.)</i>	10	10	12	12	15
<i>Річні обсяги випуску в вартісних показниках (млн. грн.)</i>	1	4	9.6	15.6	27

4.7.3. Розрахунок загальних початкових інвестиційних витрат

Результати розрахунку початкових інвестицій наведено в табл. 4.10

Таблиця 4.10

Визначення обсягу фінансування

<i>№</i>	<i>Назва етапу</i>	<i>Терміни виконання</i>	<i>Обсяги фінансування, тис. грн.</i>
1.	Проведення досліджень	3 місяці	50
2.	Написання моделей та перевірка їх	2 місяці	150
3.	Написання інтерфейсу програми	1 місяць	100
4.	Організація проекту	6 місяців	80
5.	Інші витрати	-	-
<i>Разом</i>		1 рік	380

4.7.4. Розрахунок виробничих витрат

Структура витрат необхідних для реалізації поточної діяльності наведено в табл. 4.11. Для обрахунку структури витрат необхідно врахувати витрати на орену приміщення, загальногосподарські витрати, витрати на збут,

оплату праці, витрати на матеріальні ресурси. Витрати були розраховані на 3 роки (22, 23, 24).

За розрахунками сукупні витрати за 22 рік становитимуть – 994 тисячі, за 23 – близько півтора мільйонів і за 24 – до двох мільйонів.

Таблиця 4.11

Визначення виробничих витрат

4№ з/п	Стаття витрат	Сукупні витрати за період, тис. грн.		
		2022	2023	2024
1.	Загальногосподарські витрати	-	-	-
1.1.	витрати на оренду та утримання приміщень, обладнання	100	150	200
1.2.	комунальні витрати	14	20	25
1.3.	витрати на збут, просування та рекламу	80	100	150
2.	Витрати на матеріальні ресурси (комплектуючі, сировина)	-	-	-
3.	Витрати на оплату праці	800	1200	1600
4.	Інші витрати (якщо є)	-	-	-
Разом:		994	1 470	1 975

4.7.5. Розрахунок загальних витрат на реалізацію проекту по роках

Результат розрахунку загальних витрат наведений в табл. 4.12

Таблиця 4.12

Визначення загальних витрат

Показник	Значення по роках, грн.			Разом
	2022	2023	2024	
Інвестиційні витрати (п. 7.2)	380 000	1 000 000	1 000 000	2 380 000
Виробничі витрати (п. 7.3)	994 000	1 470 000	1 975 000	4 439 000

Продовження таблиці 4.2

Обсяг загальних витрат, в тому числі за рахунок	1 374 000	2 470 000	2 975 000	6 819 000
– власних коштів	100 000	500 000	500 000	1 100 000
– коштів інвестора	1 274 000	1 970 000	2 475 000	5 719 000

4.8. Грошовий потік та оцінка вартості проекту

4.8.1. Формування грошового потоку від реалізації проекту. Чистий дисконтований дохід (NPV, Net Present Value) – це різниця між надходженнями та інвестиціями, отриманими за весь період реалізації проекту.

Результати розрахунку формування грошового потоку наведені в табл.

4.13

Таблиця 4.13

Формування грошового потоку

№	Показник	Значення по роках, тис. грн.					Разом
		2022	2023	2024	2025	2026	
1.	Надходження від проекту	1 000 000	4 000 000	9 600 000	15 600 000	27 000 000	57.2 млн грн
2.	Загальні витрати	1 374 000	2 470 000	2 975 000	2 975 000	2 975 000	12.769 млн грн
3.	Грошовий потік (CF)	-374 000	1 530 000	6 625 000	12 625 000	24 025 000	44.431 млн
4.	Акумуляований грошовий потік (ACF)	-374	1 156	7 781	20 406	44 431	-

4.8.2. Розрахунок вартості проекту

Розрахунок вартості проекту було проведено за декількома способами і результати розрахунку можна побачити в 4.14. Даний розрахунок проводиться для уточнення вартості проекту. Це одна з ключових метрик перед інвесторами

та задля можливості отримання інвесторського фінансування треба детально дослідити цю метрику, та сформувані різні репорти з наведеною метрикою, розрахованою декількома способами, це дозволить побачити, це дозволить зробити фінансовий хелікоптер вью проекту.

Таблиця 4.14

Розрахунок вартості проекту

<i>Найменування методу</i>	<i>Умови застосування</i>	<i>Отримана оцінка</i>	<i>Частка венчурного інвестора</i>	<i>Рішення щодо інвестування в проект</i>
Метод мультиплікаторів	Стартап перебуває на завершальній стадії розробки	3.2 млн. грн	25%	Інвестувати
Метод «хокейної ключки»	Стартап перебуває на завершальній стадії розробки	3.1 млн. грн	25%	Інвестувати
Метод скорингу	Стартап перебуває на завершальній стадії розробки	3.2 млн. грн	25%	Інвестувати

Висновки до розділу 4

В даному розділі було детально розглянено привабливість проекту з точки зору інвестицій в нього, як в стартап.

Можна затвердити, що західний ринок на сьогоднішній день, розвиває ринок автоматичного діагностування ретгенівських зображень, а на українському поки немає схожих стартапів, тому це може стати гарною відправною точкою для софт ланчу проекту.

ЗАГАЛЬНІ ВИСНОВКИ

Розглянуто наукову літературу за темою роботи, в результаті чого виявлено основні підходи до класифікації та розпізнавання відхилень грудної клітини. Проаналізовано найпоширеніші способи обробки зображень з метою використання їх для виділення корисної інформації з сигналів. У результаті виконання практики було обрано та реалізовано нейронну мережу для класифікації 14 поширених захворювань за рентгенівськими зображеннями пацієнтів.

Під час виконання практики було описано всі основні процеси взаємодії з нейронною мережею: імпортування архітектури, підготовка даних, збалансування класів та тестування. Попередньо було розглянено вхідні дані та їх обробка. В ході написання коду було використано мову програмування Python та одну з найпоширеніших бібліотек для глибокого навчання tensorflow.

Великою проблемою в ході роботи стала проблема дисбалансу класів в датасеті задля чого було написано спеціально функцію, яка враховує вагу кожного класу, і спеціально заставляє нейронну мережу фокусуватися на тих класах які не є широко представленими в датасеті.

В результаті роботи було отримано модель Ефішн Нет, яка показала найкращі результати за метрикою АУК. Дана модель була збережена, її ваги були збережені у форматі "h5" та при потребі може бути заімпортована та при певному налаштуванні сервера, може бути застосована в деплойменті.

Як результат, було написано модуль який візуалізує зображення та вірогідність хвороби.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. A nonparametric-based rib suppression method for chest radiographs / L. Jiann-Shu, W. Jing-Wein, W. Hsing-Hsien, Y. Ming-Zheng // *Computers & Mathematics with Applications*. — 2012. — Vol. 64. — P. 1390–1399
2. A. Padma and Dr.R. Sukanesh ,“Automatic Diagnosis of Abnormal Tumor Region From Brain Computed Tomography Images Using Wavelet Based Statistical Texture Features” , *International Journal of Computer Science, Engineering and Information Technology (IJCSEIT)*, Vol.1, No.3, August 20119
3. Adit D. A Beginner's Guide To Understanding Convolutional Neural Networks [Електронний ресурс] / Deshpande Adit – Режим доступу до ресурсу: <https://adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks>
4. Assistant, Sindhu & H K, Shashikala. (2021). IMAGE PRE-PROCESSING TECHNIQUES FOR X-RAY MEDICAL IMAGES: A SURVEY. 2320-2882.
5. Atlas-based rib-bone detection in chest X-rays / S. Candemir, S. Jaeger, S. Antani et al. // *Computerized Medical Imaging and Graphics*. — 2016. — Vol. 51. — P. 32–39
6. B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba. Learning Deep Features for Discriminative Localization. CVPR'16
7. Berlin L. Liability of interpreting too many radiographs. *AJR Am J Roentgenol*. 2000;175(1):17–22.
8. Brady AP. Measuring Consultant Radiologist workload: method and results from a national survey. *Brady AP Insights into Imaging*. 2011;2(3):247–60
9. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.
10. C. Yan, J. Yao, R. Li, Z. Xu, and J. Huang, “Weakly supervised deep learning for thoracic disease classification and localization on chest xrays,” in *Proceedings of*

the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics. ACM,2018, pp. 103–110

11. C.-H. Wei, S. Y. Chen, and X. Liu, “Mammogram retrieval on similar mass lesions,” *Computer Methods and Programs in Biomedicine*, vol. 106, no. 3, pp. 234–248, 2012. M. Mustra and M. Grgic, “Robust automatic breast and pectoral muscle segmentation from scanned mammograms,” *Signal Processing*, vol. 93, no. 10, pp. 2817–2827, 2013

12. Casper Solheim Bojer, Understanding machine learning-based forecasting methods: A decomposition framework and research opportunities, *International Journal of Forecasting*, Volume 38, Issue 4, 2022

13. D. C. Pereira, R. P. Ramos, and M. Z. do Nascimento, “Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm,” *Computer Methods and Programs in Biomedicine*, vol. 114, no. 1, pp. 88–101, 2014.

14. Deep Learning with Lung Segmentation and Bone Shadow Exclusion Techniques for Chest X-Ray Analysis of Lung Cancer / Yu. Gordienko, P. Gang, J. Hui et al. // ArXiv. — 2017. — Access mode: <https://arxiv.org/abs/1712.07632>

15. Eliminating rib shadows in chest radiographic images providing diagnostic assistance / H. Ogul, B.B. Ogul, A.M. Agildere et al. // *Computer Methods and Programs in Biomedicine*. — 2016. — Vol. 127. — P. 174–184.

16. Emanuele Pesce, Samuel Withey, Petros-Pavlos Ypsilantis, Robert Bakewell, Vicky Goh “Learning to detect chest radiographs containing pulmonary lesions using visual attention networks”, Department of Biomedical Engineering, King’s College London, London, UK.

17. Faculty of Radiologists, RCSI Measuring Consultant Radiologist Workload in Ireland: Rationale, Methodology and results from a national survey. Dublin: Faculty of Radiologists. March 2011. Available online from: <http://www.radiology.ie/about/docs/MeasuringConsultantRadiologistWorkloadIreland.pdf>

18. Fitzgerald R. Error in Radiology. *Clin Radiol*. 2001;56(12):938–46

19. H.-C. Shin, K. Roberts, L. Lu, D. Demner-Fushman, J. Yao, and R. M. Summers, “Learning to read chest x-rays: Recurrent neural cascade model for automated image annotation,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2497–2506.

20. Health National Institutes Of. NIH Chest X-rays. — 2017. — Access mode: <https://www.kaggle.com/nih-chest-xrays/data> (online; accessed: 25.04.2018)

21. Health National Institutes Of. NIH Clinical Center provides one of the largest publicly available chest xray datasets to scientific community. — 2017. — Access mode: <https://www.nih.gov/news-events/news-releases/nih-clinical-center-provides-one-largest-publicly-available-ches>(online; accessed: 25.04.2018)

22. Hetvi Son , Darshana Sankhe, “Image Restoration using Adaptive Median Filtering” Of Electronics Engineering, D.J Sanghvi College of Engineering, Mumbai, India. Oct 2019

23. How to visualize filters after the 1st layer trained by CNNs? [Электронный ресурс]. — 2016. — Режим доступа до ресурсу: <https://stackoverflow.com/questions/38450104/how-to-visualize-filters-after-the-1st-layer-trained-by-cnns>

24. Image processing. [Электронный ресурс] – Режим доступа до ресурсу: <https://www.imageprocessing.com/2017/11/balance-contrast-enhancement-technique.html>

25. Jiwon J. The Most Intuitive and Easiest Guide for Convolutional Neural Network [Электронный ресурс] / Jeyong Jiwon. – 2019. – Режим доступа до ресурсу: <https://towardsdatascience.com/the-most-intuitive-and-easiest-guide-for-convolutional-neural-network-3607be47480>

26. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778

27. Kermany D., Goldbaum M., Cai W., Valentim C. Identifying medical diagnoses and treatable diseases by deep learning. Cell. 2018. P. 1122–1131.

28. Krupinski EA, Berbaum KS, Caldwell RT, Schartz KM, Kim J. Long radiology workdays reduce detection and accommodation accuracy. *J Am Coll Radiol*. 2010;7(9):698–704

29. Kyle W. Google details AI that classifies chest X-rays with human-level accuracy [Электронный ресурс] / Wiggers Kyle. – 2019. – Режим доступа до ресурсу: <https://venturebeat.com/ai/google-details-ai-that-classifies-chest-x-rays-with-human-level-accuracy/>

30. LeCun Y. Deep learning / Yann LeCun. // *Nature*. – 2015. – №521
 Russakovsky O. ImageNet Large Scale Visual Recognition Challenge / O. Russakovsky, J. Deng. // *International Journal of Computer Vision*. – 2015. – №115.

31. LeCun, Y., et al. Handwritten digit recognition with a back-propagation network. in *Advances in neural information processing systems*

32. LeCun, Y., et al., Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998. 86(11): p. 2278-2324

33. Lung Diseases [Электронный ресурс] – Режим доступа до ресурсу: <https://medlineplus.gov/lungdiseases.html>

34. Meedeniya D, Kumarasinghe H, Kolonne S, Fernando C, Díez IT, Marques G. Chest X-ray analysis empowered with deep learning: A systematic review. *Appl Soft Comput*. 2022 Sep;126:109319. doi: 10.1016/j.asoc.2022.109319. Epub 2022 Jul 18. PMID: 36034154; PMCID: PMC9393235.

35. Narin A., Kaya C., Pamuk Z. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. New York: Cornell University Press, 2020. 31

36. Noumeir R. Benefits of the DICOM structured report. *J Digit Imaging*. 2006 Dec;19(4):295-306. doi: 10.1007/s10278-006-0631-7. PMID: 16752206; PMCID: PMC3045165.

37. Nuryanti Mohd Salleh, Harsa Amylia Mat Sakim and Nor Hayati Othman, “Neural Networks to Evaluate Morphological Features for Breast Cells

Classification” IJCSNS International Journal of Computer Science and Network Security, vol. 8 No. 9, pp. 51-58, September 2008.

38. P. Agrawal, M. Vatsa, and R. Singh, “Saliency based mass detection from screening mammograms,” *Signal Processing*, vol. 99, pp. 29–47, 2014. [14] Y. Zhang, N. Tomuro, J. Furst, and D. S. Raicu, “Identifying the optimal segmentors for mass classification in mammograms,” in *Proceedings of the Medical Imaging 2015: Image Processing*, S. Ourselin and M. A. Styner, Eds., February 2015.

39. Pesce, Emanuele & Ypsilantis, Petros & Withey, Samuel & Bakewell, Robert & Goh, Vicky & Montana, Giovanni. (2017). Learning to detect chest radiographs containing lung nodules using visual attention networks.

40. Pooling Methods in Deep Neural Networks, a Review [Электронный ресурс] – Режим доступа до ресурсу: <https://arxiv.org/ftp/arxiv/papers/2009/2009.07485.pdf>

41. Pranav M. Computer Vision: How Can Our Computers See and Make Sense Out of It? [Электронный ресурс] / Mendiratta Pranav. – 2020. – Режим доступа до ресурсу: <https://medium.com/swlh/computer-vision-how-can-our-computers-see-and-make-sense-out-of-what-they-see-f6dd777aed07>

42. Renske de Boer, Fedde van der Lijn, Henri A. Vrooman, Meike W. Vernooij, M. Arfan Ikram, Monique M.B. Breteler, Wiro J. Niessen, “Automatic segmentation of brain tissue and white matter lesions in MRI”, *Biomedical Imaging: From Nano to Macro*, 2007. ISBI 2007. 4th IEEE International Symposium, pp. 652 – 655

43. Robinson PJ. Radiology’s Achilles’ heel: error and variation in the interpretation of the Röntgen image. *Br J Radiol.* 1997;70(839):1085–98

44. Ruihua Guo. Tuberculosis Diagnostics and Localization in Chest X-Rays via Deep Learning Models // *Front. Artif. Intell.*, 05 October 2020.

45. SAGAR S. Activation Functions in Neural Networks [Электронный ресурс] / SHARMA SAGAR. – 2017. – Режим доступа до ресурсу: <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6>

46. Sanket D. Various Optimization Algorithms For Training Neural Network [Электронный ресурс] / Doshi Sanket – Режим доступа до ресурсу: <https://towardsdatascience.com/optimizers-for-training-neural-network-59450d71caf6>

47. Satia I., Bashagha S., Bibi A., Ahmed R. Assessing the accuracy and certainty in interpreting chest X-rays in the medical division. Clin Med. London

48. Seelwan Sathitratanaheewin. Deep learning for automated classification of tuberculosis-related chestX-Ray: dataset distribution shift limits diagnostic performance generalizability // ScienceDirect, Volume 6, Issue 8, August 2020.

49. Seelwan Sathitratanaheewin. Deep learning for automated classification of tuberculosis-related chestX-Ray: dataset distribution shift limits diagnostic performance generalizability // ScienceDirect, Volume 6, Issue 8, August 2020.

50. Shreedjav T. Understanding Attention Modules: CBAM and BAM [Электронный ресурс] / Trivedi Shreedjav. – 2020. – Режим доступа до ресурсу: <https://medium.com/visionwizard/understanding-attention-modules-cbam-and-bam-a-quick-read-ca8678d1c671#:~:text=Attention%20Module%3A%20What%20is%3F,and%20localize%20in%20an%20image>.

51. Stephen O., Sain M., Maduh U., Jeong D. An efficient deep learning approach to pneumonia classification in healthcare. J Healthcare Eng. 2019. P. 471–480

52. Sumit S. A Comprehensive Guide to Convolutional Neural Networks [Электронный ресурс] / Saha Sumit. – 2018. – Режим доступа до ресурсу: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>

53. VGG16 – Convolutional Network for Classification and Detection [Электронный ресурс] // VGG16 – Convolutional Network for Classification and Detection. – 2018. – Режим доступа до ресурсу: neurohive.io/en/popular-networks/vgg16/

54. Wen, H., Qi, W., & Shuang, L. (2016). Medical X-Ray Image Enhancement Based on Wavelet Domain Homomorphic Filtering and CLAHE. *2016 International Conference on Robots & Intelligent System (ICRIS)*, 249-254. Режим доступа до ресурсу: <https://www.semanticscholar.org/paper/Medical-X-Ray-Image-Enhancement-Based-on-Wavelet-Wen-Qi/8e6bf00a2fd9e3344b09f86f04e5c540bfd5aac9>

55. Wen, He et al. “Medical X-Ray Image Enhancement Based on Wavelet Domain Homomorphic Filtering and CLAHE.” *2016 International Conference on Robots & Intelligent System (ICRIS)* (2016): 249-254.

56. WHO | WHO Mortality Database [электронный ресурс] URL: https://www.who.int/healthinfo/statistics/mortality_rawdata/en/

57. Wikipedia. Log Gabor filter. — 2012. — Access mode: https://en.wikipedia.org/wiki/Log_Gabor_filter (online; accessed: 25.04.2018).

58. Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen, and H. Greenspan, “Chest pathology identification using deep feature selection with nonmedical training,” *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 6, no. 3, pp. 259–263, 2018.

59. Yamashita R. Convolutional neural networks: an overview and application in radiology / Rikiya Yamashita. // SpringerOpen. – 2018.

60. Z. Li, C. Wang, M. Han, Y. Xue, W. Wei, and L.-J. Li, “Thoracic disease identification and localization with limited supervision,” in *Computer Vision and Pattern Recognition (CVPR), 2018 IEEE Conference on*. IEEE, 2018, pp. 8290–8299

61. Zhong-Qiu Zhao, Member, IEEE, Peng Zheng, Shou-tao Xu, and Xindong Wu. “Object Detection with Deep Learning: A Review”.

62. Авшаров Е. М., Абгарян М. Г., Сафарянц С. А. Обработка медицинских изображений, как необходимый инструментарий медицинского диагностического процесса [Электронный ресурс]. – Электрон, дан. – [2017]. – Режим доступа: www.courseas.ru/download/pdf/Processing_of_medical_images.pdf

63. Горбань А.Н. Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей. Сибирский журнал вычислительной математики, 1998, т. 1, № 1. С. 12–24.

64. Машинное обучение. Классификация [Электронный ресурс]. – Электрон. издан. – [2017]. – Режим доступа: <http://www.machinelearning.ru/>.

65. Оптимизация нейронной сети [Электронный ресурс]. – 2019. – Режим доступа до ресурсу: <https://www.machinelearningmastery.ru/neural-network-optimization-7ca72d4db3e0>

66. Реализуем и сравниваем оптимизаторы моделей в глубоком обучении [Электронный ресурс]. – 2020. – Режим доступа до ресурсу: <https://habr.com/ru/company/skillfactory/blog/525214/>

67. Свёрточная нейронная сеть [Электронный ресурс] – Режим доступа до ресурсу: <https://u.to/ScVbGw>

68. Сверточная нейронная сеть, часть 1: структура, топология, функции активации и обучающее множество [Электронный ресурс] – Режим доступа до ресурсу: <https://habr.com/ru/post/348000/>

69. Нерус В. М.. Згортькова нейронна мережа для виявлення ознак туберкульозу: бакалаврська робота. : Біомедична інженерія / Нерус Владислав Миколайович. – Київ, 2021. – 60 с.

