

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ – ПРОЦЕССОВ УПРАВЛЕНИЯ
КАФЕДРА МОДЕЛИРОВАНИЯ ЭКОНОМИЧЕСКИХ СИСТЕМ

Дурандин Даниил Павлович

Магистерская диссертация

**Классификация медицинских изображений с
помощью свёрточных нейронных сетей**

Направление 01.04.02 «Прикладная математика и информатика»

Основная образовательная программа ВМ.5759 «Цифровая экономика»

Заведующий кафедрой,
доктор физ.-мат. наук,
профессор

Смирнов Н.В.

Научный руководитель,
кандидат технических наук,
доцент

Гришкин В.М.

Рецензент,
кандидат технических наук

Тимошенко Д.М.

Санкт-Петербург

2021

Содержание

Введение.....	4
Постановка задачи.....	7
Обзор литературы.....	8
Глава 1. Анализ данных.....	10
1.1 Анализ наборов данных медицинских изображений.....	10
1.1.1 Набор данных OASIS.....	10
1.1.2 Набор данных CBIS-DDSM.....	11
1.1.3 Набор данных BRATS.....	12
1.2 Анализ используемого набора изображений.....	14
1.2.1 Обзор данных.....	14
1.2.2 Описание патологий набора CheXpert.....	17
1.3 Выводы.....	19
Глава 2. Методы классификации медицинских изображений.....	21
2.1 Свёрточные нейронные сети.....	21
2.1.1 Свёрточный слой.....	22
2.1.2 Слой пулинга.....	23
2.1.3 Полносвязные слои.....	23
2.1.4 Residual block.....	24
2.1.5 Dense Block.....	24
2.1.6 Depthwise separable convolution.....	25
2.2 Модель внимания.....	26
2.3 Выводы.....	29
Глава 3. Модель внимания к потенциальным областям интереса на изображении.....	31
3.1 Модель с механизмом внимания.....	31
3.1.1 Карты признаков.....	31
3.1.2 Механизм выделения областей интереса.....	32
3.1.3 Обучение модели.....	34
3.2 Модель на основе анализа частей изображения.....	34

3.2.1 Выделение областей.....	35
3.2.2 Архитектура модели	35
3.3 Ансамблирование моделей.....	37
3.4 Выводы.....	38
Глава 4. Реализация метода.....	40
4.1 Структура программного обеспечения.....	40
4.2 Базовая модель.....	40
4.3 Предложенный подход.....	42
Заключение.....	45
Список литературы.....	46

Введение

Сегодня методы машинного обучения играют всё большую роль в автоматизации медицинских процессов, им находят применение в таких областях как геномная биоинформатика, структурная биоинформатика, анализ медицинских снимков и многих других направлениях исследований.

Так, например, искусственным интеллектом решается задача прогнозирования фолдинга белка, то есть процесс формирования сложной структуры белка. Данные прогнозы способствуют созданию белков, имеющих определённую структуру, что позволяет создавать лекарства на основе белка.

В области геномной биоинформатики искусственный интеллект решает такие задачи как: аннотация геномов и предсказание эффекта мутаций, что позволяет в значительной степени ускорить процесс аннотации, что особенно важно, учитывая стремительно растущее количество данных о последовательностях генома.

Наиболее заметную роль, методы машинного обучения играют в области анализа медицинских изображений, таких как рентген снимки и изображения, полученные с помощью компьютерной томографии. Искусственный интеллект позволяет обнаружить различные патологии на данных изображениях, что позволяет снизить нагрузку на медицинский персонал. Наиболее популярными направлениями в данной области является поиск новообразований на рентген снимках груди и томографии мозга.

Применение машинного обучения в области анализа медицинских изображений сегодня актуально как никогда, учитывая растущую доступность исследований с помощью магнитно-резонансной томографии, сегодня генерируется большое количество данных, которые необходимо анализировать человеку. Внедрение искусственного интеллекта в данный процесс, позволит значительно увеличить скорость обработки данных и снять нагрузку со специалистов в данной области. Высокая степень автоматизации данных процессов позволит проводить более широкую

диспансеризацию населения, что в свою очередь приведёт к увеличению ранней выявляемости заболеваний, а следственно снизит уровень смертности населения, от заболеваний которые имеют слабую симптоматику на раннем этапе.

О необходимости исследований в данной области нам говорит то, что уже сегодня методы машинного обучения внедряются в процесс диагностики заболеваний. Так например, уже сегодня платформа Botkin.AI помогает врачам выявлять рак лёгких на компьютерных томограммах, а модель, разработанная компанией Care Mentor AI позволяет по КТ-снимкам определить процент и степень поражения лёгких COVID-19. Данные сервисы были подключены к единому радиологическому информационному сервису, что говорит о том, что данный подходу уже активно используется для диагностики заболеваний.

Основным источником данных, рассматриваемом в данной работе являются рентген снимки груди с выявленными заболеваниями. Основной задачей данной работы является классификация данных снимков по типу выявленного заболевания. Для решения данной необходимо привлечь методы машинного обучения, в частности свёрточные нейронные сети, которые являются наиболее популярным решением для анализа изображений.

Так же необходимо проанализировать специфику данных снимков, для того чтобы произвести соответствующую предварительную обработку изображений и выбрать наиболее подходящую архитектуру свёрточной сети. Следует так же изучить специфику различных наборов данных медицинских изображений, чтобы выявить общие закономерности в таких данных, для того чтобы результаты данной работы не были слишком специфичны и могли быть применены и для других областей анализа медицинских изображений.

В ходе данной работы был произведён обзор различных методов классификации, выполнен поиск необходимых данных, были предложены

базовый подход и подход, учитывающий специфику данных изображений, данные подходы были реализованы, было произведено сравнение и анализ их результатов работы.

Постановка задачи

В данной работе стоит задача классификации рентген снимков по типам заболеваний. Для её решения необходимо проанализировать различные наборы медицинских изображений, для понимания специфики данных. Разработать архитектуру свёрточной нейронной сети и предложить подход для её обучения. Предложенный подход должен учитывать особенности наборов данных, которые были выведены в ходе предварительного анализа. Реализовать ансамбль, построенный на основе популярных архитектур свёрточных сетей, для сравнения с результатами предложенного метода.

Для этого необходимо:

- провести анализ современных алгоритмов компьютерного зрения, а также способы их применения в рамках поставленной задачи
- проанализировать различные наборы медицинских изображений
- провести анализ подходов к классификации изображений, учитывающие выявленные особенности
- предложить возможные пути улучшения найденных решений
- реализовать предложенный подход
- реализовать базовый метод классификации на основе популярных архитектур свёрточных сетей
- сравнить результаты работы данных подходов

Обзор литературы

Сегодня было исследовано большое количество подходов к задаче классификации медицинских изображений. Было выявлено и исследовано множество специфических проблем, связанных с наборами медицинских изображений и способами их классификации. Одной из таких проблем является сложность аннотирования снимков, так как это требует привлечения большого количества специалистов в данной области, что требует больших затрат.

В статье [1] предлагается подход, который способствует уменьшению необходимого количества аннотированных снимков. Авторы вводят метод, основанный на предобученных моделях, который позволяет использовать небольшую часть набора размеченных данных для аннотирования неразмеченных данных. Особенность подхода заключается в том, что для обучающего набора используется небольшая часть наилучшим образом аннотированных данных, в смысле максимального значения энтропии. В зависимости от значения пересечения аннотированных и размеченных данных, это позволяет сократить обучающую выборку в несколько раз.

Для некоторых областей обработки медицинских изображений характерна работа с изображениями очень высокого разрешения. Основным недостатком одного из популярных подходов, сжатия изображения, является то, что при данном процессе теряется часть важных данных.

В статье [2] предложен подход, решающий данную проблему, авторы статьи предлагают построить ансамбль моделей, каждый на вход из которых подаётся фрагмент исходного изображения. Каждая из моделей кодирует изображение в латентные вектора, определённой размерности. Далее для всех таких полученных векторов осуществляется средний и максимальный пуллинг, результаты которого передаются на полносвязные слои классификатора. Далее результаты классификатора, объединяются с результатами классификации МРТ снимков и финальный результат получается мягким способом голосования по двум моделям. Данный подход

позволяет обрабатывать изображения высокого разрешения без потери части данных и при этом значительно экономить вычислительные ресурсы.

Авторы статей [3-4] предлагают схожие подходы, основанные на ансамбле нескольких свёрточных сетей, которые служат для формирования признаков для дальнейшей классификации. Признаки, извлечённые с последних слоёв свёрточных сетей, затем обрабатываются различными методами классификации машинного обучения.

В следующей работе [5] авторы предлагают архитектуру свёрточной нейронной сети, которая показывает лучшее качество классификации, относительно популярных моделей свёрточных сетей. При этом, за счёт меньшего количества параметров процесс обучения занимает намного меньше времени.

Глава 1. Анализ данных

В последнее время растёт количество и доступность данных связанных с медицинскими изображениями, что увеличивает интерес к анализу снимков в данной области. Наиболее популярными наборами данных являются трёхмерные данные компьютерной томографии и двухмерные срезы слоёв томографии и рентген снимки. В данной главе будет произведён анализ некоторых популярных наборов данных, а также подробный анализ данных, используемых в этой работе.

1.1 Анализ наборов данных медицинских изображений

Далее будет произведён анализ и краткий обзор популярных наборов данных, находящихся в свободном доступе. Обозреваемые данные относятся к области двухмерных изображений: рентгеновских снимков и двухмерных срезов снимков МРТ.

1.1.1 Набор данных OASIS

Серия исследований изображений в открытом доступе (OASIS) - это набор данных [6], посвященный влиянию старения и болезни Альцгеймера, который находится в свободном доступе для научного сообщества. OASIS-3 включает данные 1098 участников, в возрасте от 42 до 95 лет, включая людей с нормальными когнитивными функциями и людей с ранней стадией деменции. OASIS-3 включает структурную и функциональную МРТ (магнитно-резонансную томографию), амилоидную и метаболическую ПЭТ (позитронно-эмиссионную томографию), нейропсихологическое тестирование и клинические данные. Обширные данные, представленные в OASIS-3, могут использоваться как отдельный набор данных или в сочетании с другими открытыми данными, отвечающими на множество тем исследований.

При анализе процессов деменции МРТ снимки могут :

- помочь подтвердить диагноз деменции и тип заболевания, вызывающего деменцию

- предоставить подробную информацию о повреждении кровеносных сосудов при сосудистой деменции
- демонстрируют сокращение в определенных областях мозга - например, лобные и височные доли в основном страдают от сжатия при лобно-височной деменции, в то время как обычно только височные доли поражаются на ранних стадиях болезни Альцгеймера

Как правило, на поздних стадиях болезнь Альцгеймера идентифицируется сокращением различных областей мозга, это можно увидеть на рисунке 1.

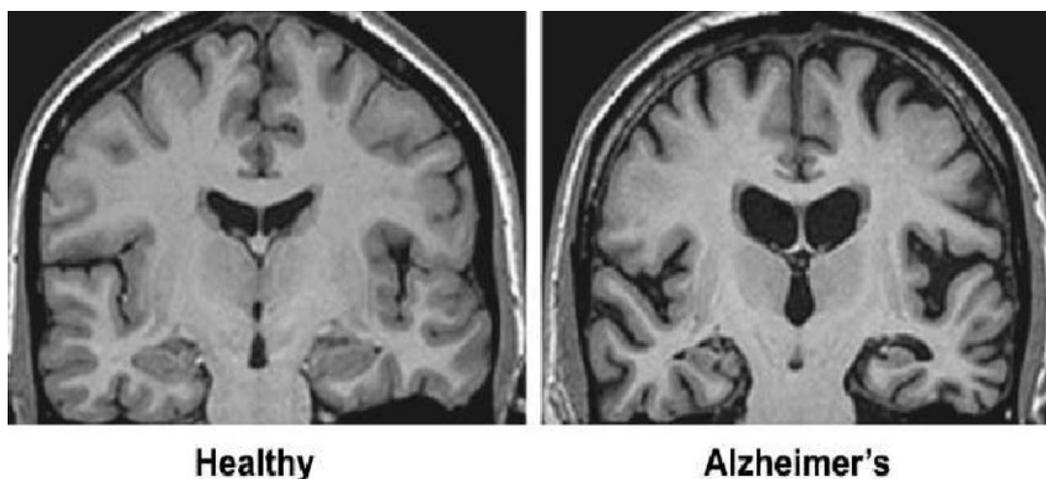


Рисунок 1. МРТ снимки здорового мозга и мозга с признаками болезни Альцгеймера

Таким образом, анализ изображений отдельных участков мозга поможет определить степень поражения. В статье [7] предлагается подход на ансамблировании различных популярных моделей свёрточных сетей. В статье [8] были воспроизведены и проанализированы различные подходы к классификации снимков из этого набора данных.

1.1.2 Набор данных CBIS-DDSM

CBIS-DDSM (Curated Breast Imaging Subset of DDSM) [9] представляет собой обновленную и стандартизированную версию цифровой базы данных для скрининговой маммографии (DDSM). DDSM - это база данных,

содержащая 2620 сканированных маммографических исследований. Он содержит нормальные, доброкачественные и злокачественные случаи с подтвержденной специалистом информацией о патологии. Коллекция CBIS-DDSM включает подмножество данных DDSM, отобранных и обработанных маммографом. Набор данных также включает выделенные области для сегментации и ограничивающие рамки для зон интереса.

На рисунке 2 можно увидеть, выделенные зоны интереса доброкачественных и злокачественных образований, они занимают совсем небольшую область, относительно целого снимка, поэтому в самом наборе данных присутствуют координаты и ограничивающие рамки для областей интереса.

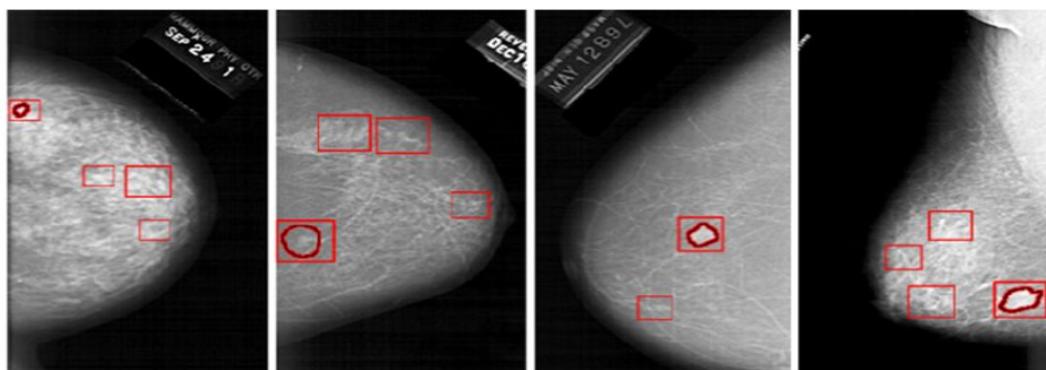


Рисунок 2. Ограничивающие рамки для областей интереса

В большинстве работ, представленных для данного набора, модели нейронных сетей обучаются именно на фрагментах изображения. В следующей работе [10] рассматривается метод на основе механизмов внимания, применительно к этому набору данных. Авторы данной статьи [11] предлагают метод классификации с использованием техники разбиения исходного изображения на части, для выделения зон интереса.

1.1.3 Набор данных BRATS

BRATS[12-14] представляет собой набор предоперационных мультимодальных МРТ-сканирований глиобластомы (GBM / HGG) и глиомы более низкого уровня (LGG) с подтвержденным диагнозом.

Все снимки доступны в виде файлов NIfTI и содержат различные рипы МРТ снимков для каждого случая. Все наборы данных изображений сегментированы вручную от одного до четырех экспертов по одному и тому же протоколу аннотации, и их аннотации были одобрены опытными нейрорадиологами. Аннотации включают GD-усиливающую опухоль (ET - метка 4), перитуморальный отек (ED - метка 2) и некротическое и не усиливающее ядро опухоли (NCR / NET - метка 1), как описано в справочной статье BraTS, опубликованной в IEEE Transactions for Medical Imaging.

На рисунке 3 можно увидеть примеры из данных для обучения BRATS с выделенными областями опухоли, выведенными из аннотаций отдельных экспертов (синие линии) и согласованной сегментацией (пурпурные линии).

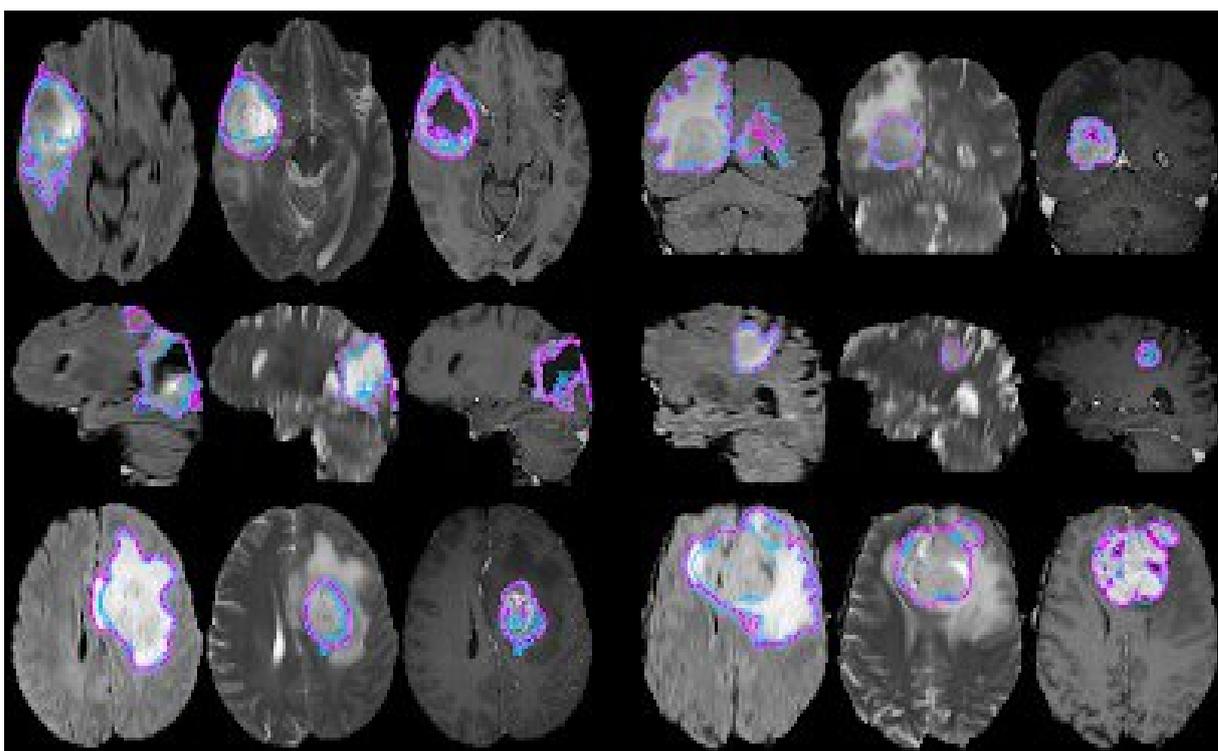


Рисунок 3. Примеры данных из набора BRATS с выделенными областями опухоли

В каждой строке показаны два случая опухоли высокой степени злокачественности (строки 1-4), опухоли низкой степени злокачественности (строки 5-6) или синтетических случаев (последняя строка).

Для данного набора снимков также характерен небольшой размер области интереса, относительно размера всего снимка. В связи с проблемой выделения нескольких перекрывающихся областей, в следующей работе [15] был предложен метод сегментирования, основанный на построении нескольких моделей с механизмов внимания, каждая из которых нацелена на сегментацию области с соответствующей меткой. В качестве базовой модели в данной работе была использована модификация популярной архитектуры свёрточной сети U-Net [16], с добавлением механизма внимания, предложенного в статье [17].

1.2 Анализ используемого набора изображений

В качестве данных для данной работы был выбран набор рентген снимков лёгких CheXpert[18].

Причиной этому послужили такие факторы как:

- возможность использовать набор сжатых изображений в формате jpeg, так как он намного меньший объём памяти
- наибольшая потребность в автоматизации процесса распознавания отклонений в данной области, так как флюорографические исследования проводятся намного чаще прочих

1.2.1 Обзор данных

CheXpert - это большой общедоступный набор данных для интерпретации рентгенограмм грудной клетки, состоящий из 224 316 рентгенограмм 65 240 пациентов. Он был собран из рентгенологических исследований грудной клетки в Стэнфордской больнице, проведенных в период с октября 2002 года по июль 2017 года как в стационарных, так и в амбулаторных центрах, вместе с соответствующими радиологическими отчетами. Пример изображений из набора данных можно увидеть на рисунке 4.

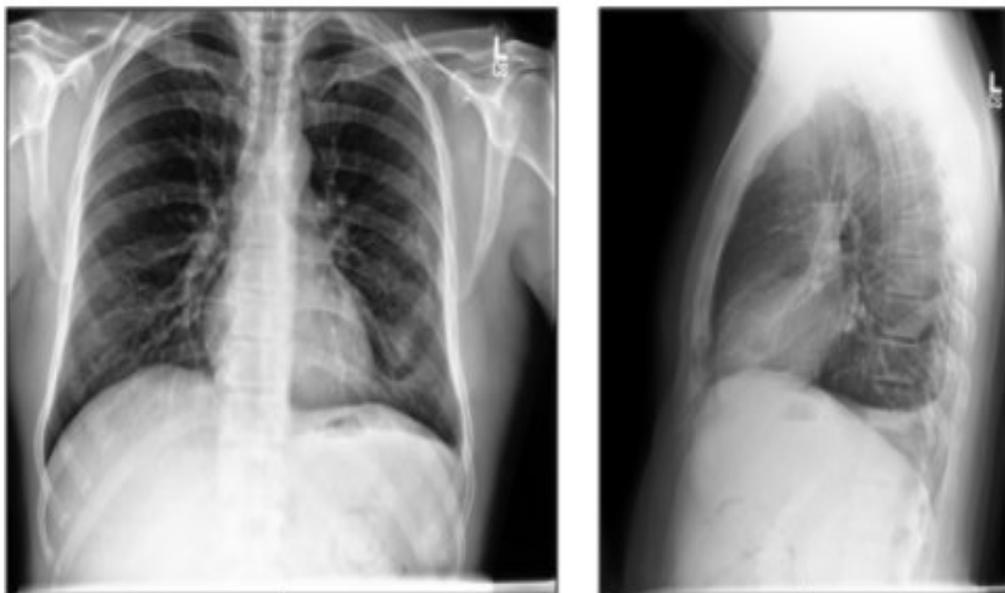


Рисунок 4. Пример снимка в двух проекциях из набора данных CheXpert

Каждый отчет был помечен на наличие 14 наблюдаемых патологий как положительный, отрицательный или неопределенный. Аннотация изображений проводилась с помощью автоматизированной этикетировочной машины на основе правил для извлечения наблюдений из радиологических отчетов с произвольным текстом. Извлечённые метки используются в качестве структурированных меток для изображений.

Аннотация снимков состоит из трех отдельных этапов: извлечение упоминаний, классификация упоминаний и объединение упоминаний. На этапе извлечения упоминаний извлекаются упоминания из списка наблюдений из раздела «Отпечаток» радиологических отчетов, в котором приводятся основные результаты радиографического исследования. На этапе классификации, упоминания наблюдений классифицируются как отрицательные, неопределенные или положительные. На этапе агрегации упоминаний, используется классификация для каждого упоминания наблюдений, чтобы получить окончательную метку для 14 наблюдений (пробел для не упомянутых, 0 для отрицательных, -1 для неопределенных и 1 для положительных). Результаты автоматического аннотирования можно увидеть на рисунке 5.

	Observation	Labeler Output
<p>1. <i>unremarkable</i> <u>cardiomediastinal silhouette</u></p> <p>2. diffuse <u>reticular pattern</u>, which can be seen with an atypical <u>infection</u> or chronic fibrotic change. <i>no focal consolidation</i>.</p> <p>3. <i>no pleural effusion</i> or <u>pneumothorax</u></p> <p>4. mild degenerative changes in the lumbar spine and old right rib <u>fractures</u>.</p>	No Finding	
	Enlarged Cardiom.	0
	Cardiomegaly	
	Lung Opacity	1
	Lung Lesion	
	Edema	
	Consolidation	0
	Pneumonia	u
	Atelectasis	
	Pneumothorax	0
	Pleural Effusion	0
	Pleural Other	
	Fracture	1
	Support Devices	

Рисунок 5. Аннотирование снимков

Стоит отметить, что набор данных является несбалансированным относительно, различных патологий, причём разница между наиболее и наименее представленными патологиями может достигать нескольких десятков тысяч раз. Это можно увидеть на таблице 1.

Таблица 1: Распределение различных патологий в зависимости от их типа.

Патология	Положительный	Неточный	Отрицательный
No Finding	16627	0	171014
Enlarged Cardiomediastinum	9020	10148	168473
Cardiomegaly	23002	6597	158042
Lung Opacity	92669	4341	90631
Lung Lesion	6856	1071	179714
Edema	48905	11571	127165
Consolidation	12730	23976	150935
Pneumonia	4576	15658	167407
Atelectasis	29333	29377	128931
Pneumothorax	17313	2663	167665
Pleural Effusion	75969	9419	102526
Pleural Other	2441	1771	183429
Fracture	7270	484	179887
Support Devices	105831	898	80912

В статье[18], опубликованной авторами набора данных, предлагаются различные подходы к интерпретации наблюдений, помеченных как неопределённые.

В частности, предлагаются следующие подходы:

- игнорирование неопределённых меток во время обучения
- обозначение неопределённых меток как отрицательных
- обозначение неопределённых меток как положительных
- подход основанный на обучении модели, помечая неопределённые метки как отрицательные, и использование данной модель для создания прогнозов, которые сопоставляют каждой метки неопределенности прогноз вероятности, полученный моделью.
- рассматривать метки как отдельный класс

Каждый из данных подходов, позволяет повысить метрики для классификации изображений для одного или нескольких классов, но нет такого метода, который бы давал значительный прирост точности для всех классов.

1.2.2 Описание патологий набора CheXpert

Описания отклонений в данном параграфе были взяты из раздела описания различных патологий публичного ресурса Radiopedia [19].

Ателектаз - это уменьшение вздутия всего или части легкого. Синоним «коллапс» часто используется как синоним ателектаза, особенно когда он тяжелый или сопровождается явным увеличением помутнения легких. На рентгеновских снимках и компьютерной томографии наблюдается уменьшение объема, сопровождающееся повышенной непрозрачностью легкого.

Кардиомегалия - это универсальный термин для обозначения увеличенного сердца. В большинстве случаев для выявления кардиомегалии достаточно рентгеновского снимка грудной клетки (поскольку сердце либо

явно нормальных размеров, либо явно ненормально увеличено). Так же для диагностики можно рассчитать кардиоторакальное соотношение (CTR). CTR измеряет ширину сердечного силуэта и грудной полости; соотношение больше 0,5 является отклонением от нормы.

Консолидация - это термин, используемый для описания консолидации в одной из долей легкого. Это указывает на альвеолярное распространение болезни и чаще всего возникает из-за пневмонии. Консолидация относится к альвеолярным воздушным пространствам, заполненным жидкостью (экссудат / транссудат / кровь), клетками (воспалительными), тканями или другим материалом. Общие особенности консолидации на рентгеновских снимках включают: помутнение воздушного пространства, вызывающее обскурацию легочных сосудов.

Отек легких - это широкий описательный термин, который обычно определяется как аномальное скопление жидкости во внесосудистых отделах легких. Рентген грудной клетки покажет жидкость в стенках альвеол, усиленное затемнение сосудов, и, возможно, плевральный выпот.

Увеличение сердечного силуэта является - общий термин соответствующий нетипичной форме сердечного силуэта, одной из причин которого, в частности, является кардиомегалия.

Перелом - полное или частичное нарушение целостности кости. На рентгеновских снимках чаще всего обнаруживаются переломы ребер, как правило диагностика такого перелома затруднена, если нет смещения.

Поражение легкого - общий термин, который часто связывается, например, с лёгочным узлом. Одиночный легочный узел или «пятно в легком» определяется как четко очерченное, округлое помутнение, не превышающее 3 см в диаметре, которое полностью окружено легочной тканью, не касается корня легочной ткани.

Легочное помутнение представляет собой результат уменьшения соотношения газа и мягких тканей (крови, паренхимы и стромы легких) в

легких. На рентгеновских снимках выделяются области затемнения воздушного пространства в лёгких.

Плевральный выпот - это аномальное скопление жидкости в плевральной полости. Они могут быть результатом различных патологических процессов, которые подавляют способность плевры реабсорбировать жидкость. Небольшие объёмы жидкости плохо различимы на фронтальном рентген снимке, лучше различаются на боковой проекции. Как правило характеризуется затемнением нижней части лёгкого.

Пневмония - это инфекция, которая вызывает воспаление в одном или обоих легких и может быть вызвана вирусом, бактериями, грибами или другими микробами. На снимках представлено в виде участков помутнений, так называемых «белых пятен».

Пневмоторакс, относится к наличию газа (часто воздуха) в плевральной полости. Когда это скопление газа постоянно увеличивается, что приводит к сжатию структур средостения, это может быть опасным для жизни и известно как напряженный пневмоторакс. При пневмотораксе на снимке характерны следующие особенности: видимый висцеральный край плевры виден как очень тонкая резкая белая линия, на периферии этой линии не видно следов легких, периферическое пространство рентгенопрозрачно по сравнению с соседним легким.

Медицинские устройства в грудной клетке регулярно наблюдаются рентгенологами при просмотре рентгенограмм и компьютерных томографов. Среди них могут быть такие как: катетеры, торакостомические трубки и так далее. Как правило их явно видно на снимке.

1.3 Выводы

В данной главе был проведён обзорный анализ нескольких популярных наборов медицинских изображений и подробный анализ, используемого набора данных. Исходя из данного анализа, можно сделать вывод, что большинство медицинских изображений содержат небольшую область

интереса на снимке. Для многих наборов данных предоставляются ограничивающие рамки с выделенными областями интереса, что позволяет обучать нейронные сети не на целом изображении, а только на нужной части. Выбранный для работы набор данных имеет схожую специфику, а так же является несбалансированным относительно представленных патологий.

Глава 2. Методы классификации медицинских изображений

В связи с ростом доступности различных наборов медицинских данных, в последнее время появилось множество предложений по решению специфических проблем, связанных с такими данными. Большинство таких предложений связано с введением новых архитектур нейронных сетей, созданием специфических слоёв для уже существующих архитектур, различными подходами к обучению сетей и способами их ансамблирования.

В данной главе будут проанализированы, наиболее интересные подходы, с точки зрения полученных выводов по анализу медицинских наборов данных. Также в данной главе будут введены общие концепции работы свёрточных нейронных сетей и наиболее популярных сегодня архитектур, которые будут использованы в дальнейшем, в данной работе.

2.1 Свёрточные нейронные сети

Свёрточные нейронные сети - это специальный тип нейронных сетей, направленных на распознавание образов на изображениях. Последнее время, данный тип сетей, является наиболее популярным в решении задач обработки изображений. С помощью данного вида нейронных сетей решаются такие задачи как: классификация и сегментация изображений. В последнее время было создано множество различных типов архитектур и вариаций слоёв для данного типа сетей, для работы с различными задачами в обработке изображений. Как правило свёрточная нейронная сеть состоит из чередования слоёв свёртки и слоёв субдискретизации, которые формируют признаки для выходного полносвязного слоя. На выход нейронной сети подаются метки классов или изображения с размеченными по различным классам пикселями, для осуществления задачи сегментации. Архитектуры свёрточных нейронных сетей крайне разнообразны, но построены они все по схожему принципу чередования слоёв свёртки и пулинга. Типовую архитектуру свёрточной сети, можно увидеть на рисунке 6.

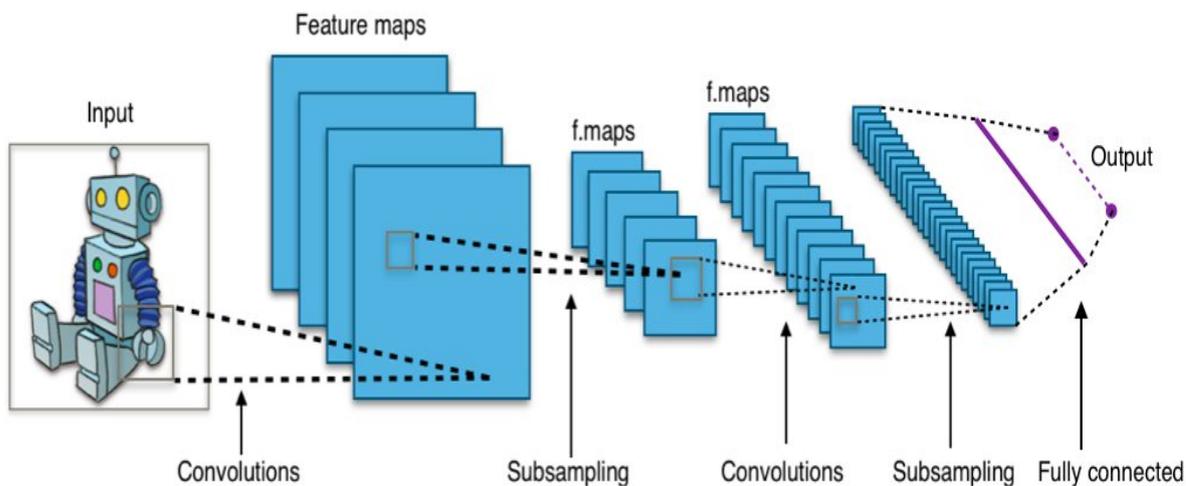


Рисунок 6. Типовая архитектура свёрточной сети

2.1.1 Свёрточный слой

Свёрточный слой является основным блоком сети, который выполняет операцию свёртки. Свёртка осуществляется путём суммирования результата поэлементного умножения двух матриц. Первой матрицей является карта признаков с предыдущего слоя сети. Второй матрицей является так называемый фильтр или ядро свёртки, содержащий веса, параметры формируемые в ходе обучения сети. Фильтр представляет из себя небольшое окно, которое «скользит» по матрице карты признаков и путём поэлементного умножения и суммирования формирует карту признаков следующего слоя. Пример данного процесса представлен на рисунке 7.

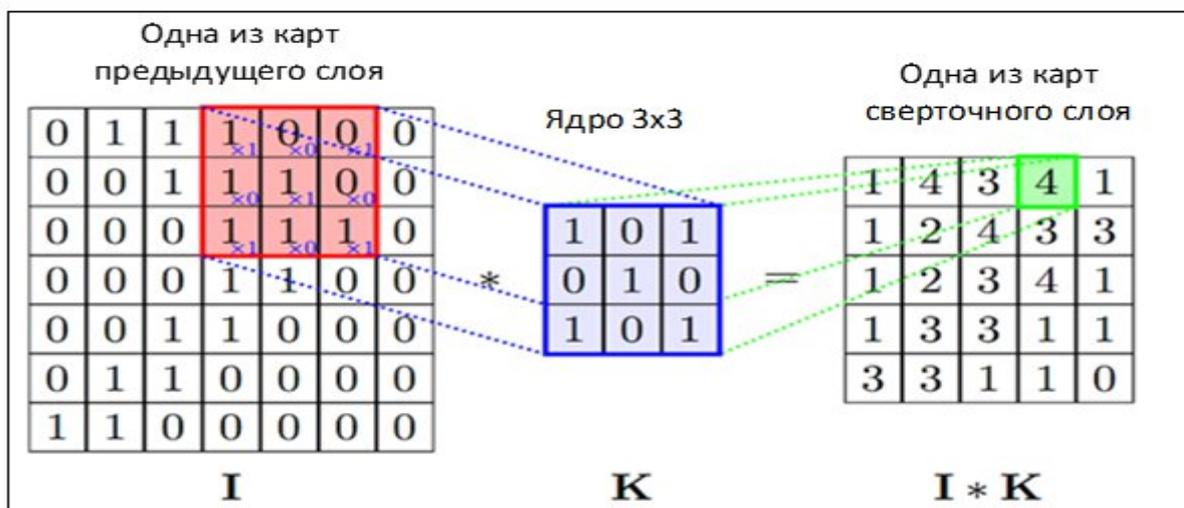


Рисунок 7. Пример операции свёртки

Сформированные в процессе обучения, значения фильтра интерпретируются как объекты определённой формы, а полученные после свёртки значения, как сигнал о наличии или отсутствии данного объекта в определённом месте изображения.

2.1.2 Слой пулинга

Слой пулинга служит для уменьшения размерности карт признаков предыдущего слоя, осуществляется это путём фильтрации матрицы карты признаков, как правило, с помощью функции максимума. Матрица слоя пулинга проходит по всем элементам матрицы карты признаков предыдущего слоя и формирует новую карту размерности в n раз меньшей, чем она была, где n - это размер окна пулинга. Пример такой операции изображён на рисунке 8.

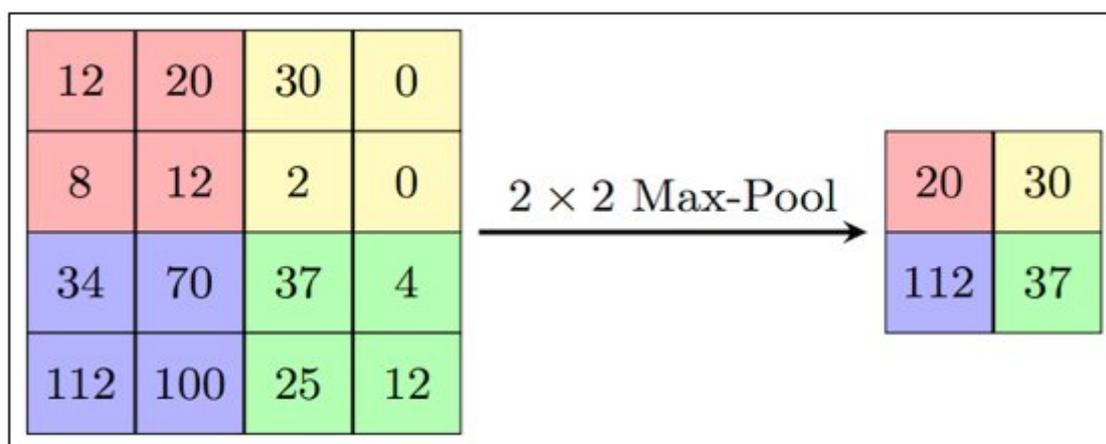


Рисунок 8. Операция пулинга

2.1.3 Полносвязные слои

На вход полносвязным слоям подаётся вектор, полученный после финальной операции пулинга, осуществлённой над всей матрицей карты признаков, таким образом вектор размерности $n \times 1$, где n это количество фильтров финального слоя свёртки в сети подаётся на вход полносвязного слоя. На выход полносвязных слоёв подаётся вектор с метками классов. В случае решения задач сегментации изображения, полносвязные слои

отсутствуют, а на выход последнего свёрточного слоя сети подаётся изображение с размеченными пикселями классов.

2.1.4 Residual block

Одной из серьёзных проблем, возникающих в ходе обучения нейронных сетей это проблема исчезающего и взрывающегося градиента, она возникает из-за того при расчёте ошибки обратного распространения, дифференциации функции потерь производится по правилу цепочки. Таким образом, до глубоких слоёв сети, из-за многократного умножения на небольшие величины, может доходить незначительная величина. Чтобы избежать данной проблемы, была предложена модель Residual block [20]. Основное отличие от стандартных моделей состоит в том, что слои соединяются не только последовательно, но и через несколько последовательно стоящих слоёв. Схематическое изображение блока можно изображено на рисунке 9. На основе данных блоков построены такие популярные модели свёрточных сетей, как ResNet.

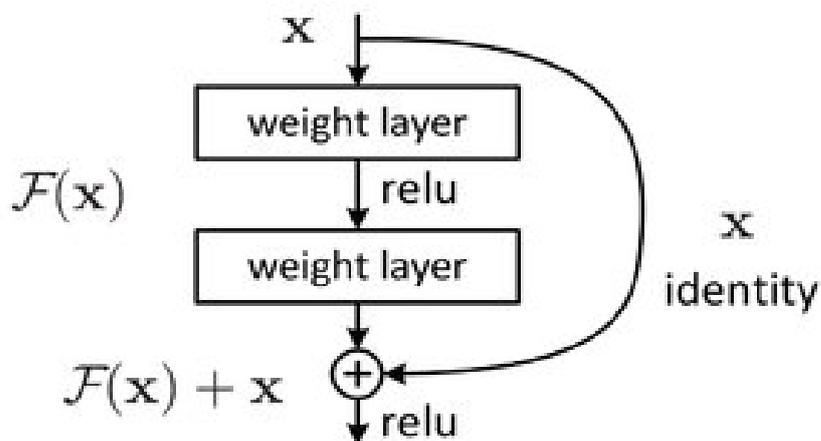


Рисунок 9. Residual block

2.1.5 Dense Block

Идея Dense Block, по своей сути схожа с идеей Residual Block, и решает схожие проблемы, основным отличием является то, что слои в блоках соединяются не с каким-то пропуском, а каждый предыдущий соединяется с

каждым последующим. Так же предыдущие слои не суммируются с последующими, а соединяются или конкатенируются в один тензор. На основе данных блоков построена такая популярная сеть, как DenseNet[21], которая при меньшем числе параметров, позволяет получить схожие с другими сетями результаты. Изображение данного блока можно увидеть на рисунке 10.

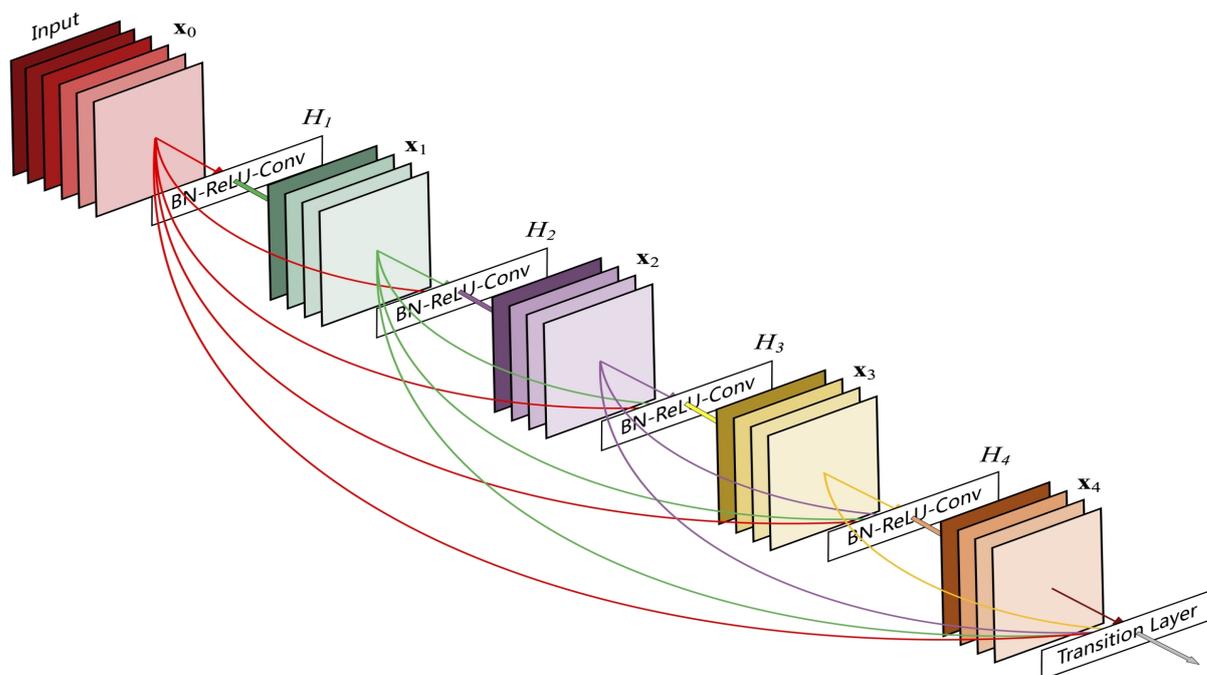


Рисунок 10. Dense Block

2.1.6 Depthwise separable convolution

Данный подход идеологически продолжает архитектуру Inception, которая использовалась, для того чтобы снизить количество гиперпараметров у сети, за счёт одновременной свёртки карты признаков фильтрами разного размера. В основе же данного подхода[22], лежит идея уменьшения количества обучаемых параметров у сети. В данном блоке сначала производится свёртка с фильтром 1×1 , а затем уже к каждому каналу отдельно применяется свёртка 3×3 . Уменьшение количества признаков, получается за счёт того, что первичная свёртка совершается единичным фильтром, что позволяет уменьшить количество параметров в n^2 раз, где n это размер фильтра.

А вторичная свёртка производится, только по каждому полученному каналу отдельно, а не всем вместе, что позволяет сократить количество параметров в k раз, где k это количество каналов в карте признаков, к которой применяется свёртка. Схематичное изображение блока представлено на рисунке 11.

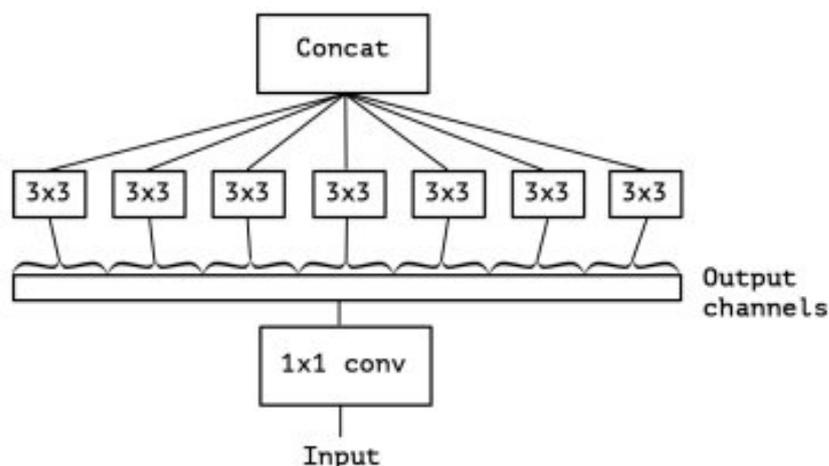


Рисунок 11. Схема блока Depthwise separable convolution

На основе данного подхода работает популярная архитектура MobileNet[23].

2.2 Модель внимания

Модели внимания были предложены для анализа текстовых данных. Они были призваны выделить важность элементов в контексте одного документа. Изначально механизм внимания был представлен в контексте рекуррентных Seq2seq сетей для "обращения внимания" блоков декодеров на скрытые состояния RNN для любой итерации энкодера, а не только последней.

Применение модели внимания для анализа изображений рассмотрим на примере данной статьи [24].

Как уже упоминалось, проблема с использованием всего рентгеновского изображения для классификации заключается в том, что на медицинских изображениях область поражения может быть очень маленькой по сравнению со всем изображением и даже может располагаться где-то вдоль границ, вызывая большое количество шума для классификатора и снижение точности обнаружения.

В статье авторы используют рекурсивное жесткое внимание (то есть жесткое самовнимание) для повышения точности обнаружения путем вырезания отличительных частей изображения и классификации как глобального изображения, так и кадрированной части вместе. Пример выделения областей интереса на снимке можно увидеть на рисунке 12.

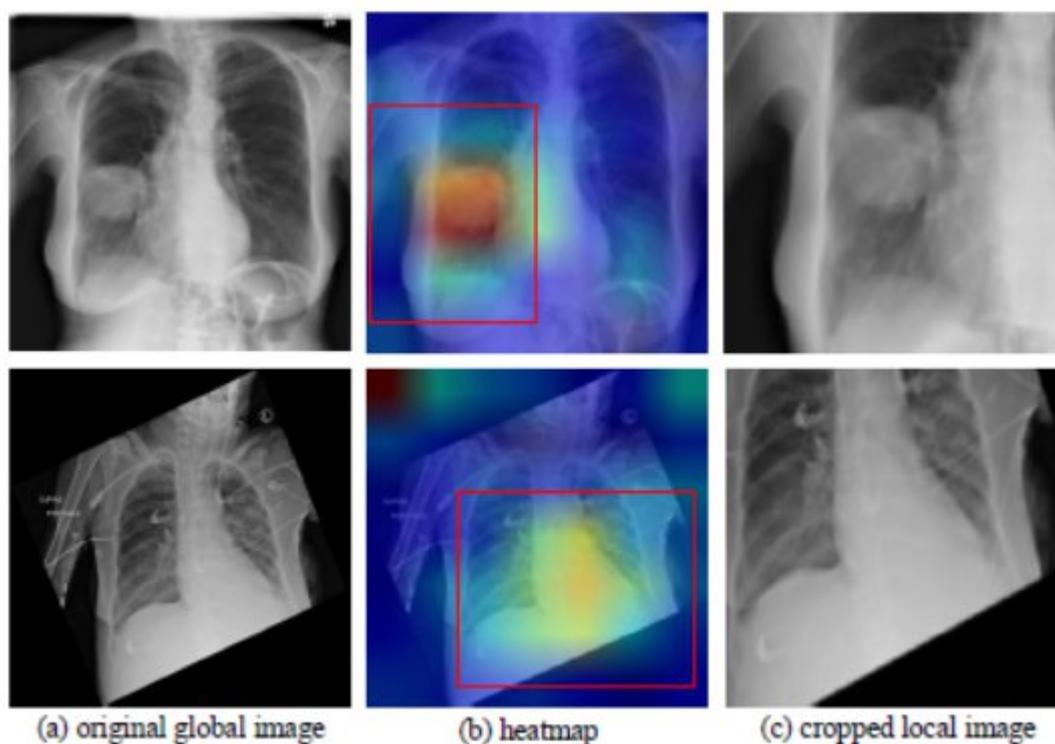


Рисунок 12. Два снимка, их тепловые карты, извлеченные из определенного сверточного слоя в сети, а также полученное обрезанное изображение.

Для классификации используются как исходное, так и обрезанное изображение.

Сеть состоит из трех ветвей :

- Глобальная ветвь , обрабатывающая все изображение и определяющая ROI(область интереса кадрирования)
- Локальная ветвь, использующий механизм внимания для обработки кадрированного сегмента изображения

- Fusion ветвь служит для объединения результатов глобального и локального сегментов сети и выполнение окончательной классификации с использованием полносвязных слоев

Все ветви данной сети выполняют классификацию для прогнозирования наличия патологий. Помимо классификации, глобальная ветвь используется для создания тепловой карты, которая определяет обрезанную область. Тепловая карта создается путем подсчета максимальных значений по каналам на определенном слое. Затем создается маска того же размера, что и тепловая карта. Если значение карты интенсивности пикселей для каждого канала в определенной позиции превышает некоторый порог, пиксель маркируется единицей. В противном случае пиксель будет иметь значение 0. Точки со значением 1 определяются как область интереса. Затем данный сегмент поступает на локальную ветвь, где анализируется уже обрезанный сегмент. После этого, выходы обеих ветвей объединяются для выполнения итоговой классификации.

Сеть обучается в три этапа:

- Сначала обучается глобальная ветвь, предварительно обученная на датасете ImageNet, формируются тепловые карты изображений для поиска области интереса.
- Сегменты полученные, после прохождения глобальной сети подаются в локальную сеть, для более точной классификации. Весовые коэффициенты глобальной ветви при этом замораживаются.
- Выходные данные глобальной и локальной ветвей объединяются и производится итоговая классификация. Веса предыдущих ветвей при этом замораживаются.

Результаты работы данного алгоритма сравнивают с несколькими, ранее предложенными методами [25-27]. В результате сравнения авторы

показывают, что предложенный ими алгоритм улучшил показатели классификации относительно рассматриваемых методов.

Схематично представленную архитектуру предложенной сети, можно увидеть на изображении 13.

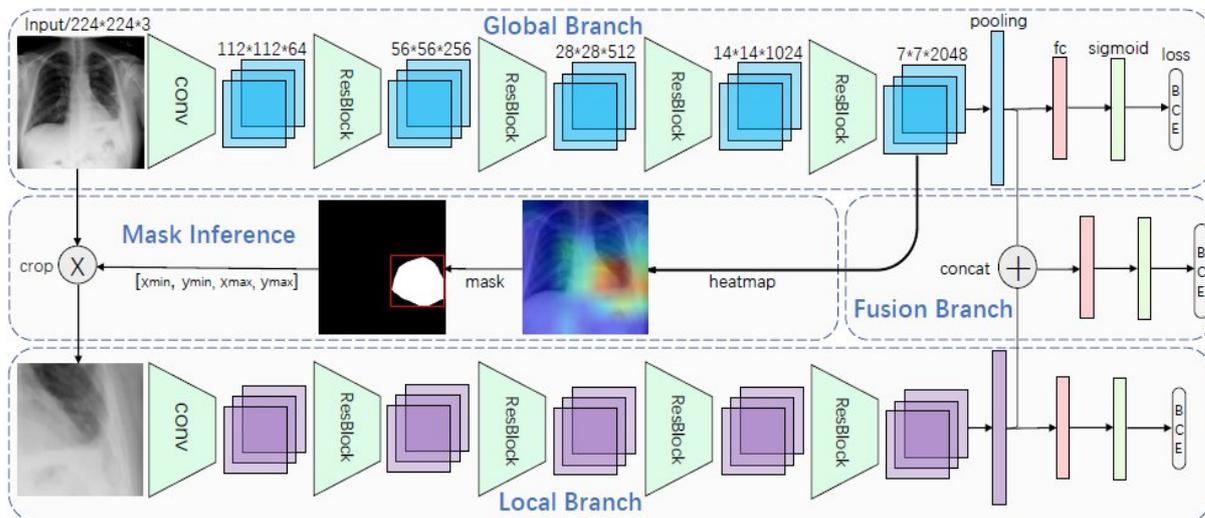


Рисунок 13. Архитектура AG-CNN, основанная на модели ResNet50.

Данный алгоритм действительно улучшает качество классификации но имеет ряд недостатков. Например при выделении зон интереса путём вычисления максимума по всем картам признаков, извлечённых из последнего слоя свёрточной нейронной сети, теряется часть информации. Так как выделение области интереса происходит в той части изображения, где находится только одна определённая патология.

2.3 Выводы

В данной главе были введены общие сведения о работе свёрточных нейронных сетей, которые необходимы для понимания работы обозреваемого далее метода. Также был произведён анализ подхода основанного на механизме внимания. В ходе анализа были выявлены предполагаемые недостатки метода, а именно потеря части информации при выделении областей интереса. Данная информация теряется за счёт того, что из исходного изображения извлекается область с максимальными

значениями пулинга по данному изображению, что может привести к отсечению частей изображения которые содержат другие патологии.

Глава 3. Модель внимания к потенциальным областям интереса на изображении

В данной главе описывается построение базового метода для классификации изображений на основе популярных архитектур свёрточных сетей, приводятся способы их ансамблирования. Также предлагаются методы, которые учитывают специфику медицинских наборов изображений и набора ChexPert в частности. Данные модели и способы их ансамблирования были предложены на основе анализа методов и различных наборов медицинских изображений описанных в предыдущих главах.

3.1 Модель с механизмом внимания

Для решения проблем связанных со спецификой наборов медицинских данных, таких как небольшая зона поражения, относительно всего снимка, была предложена модель с механизмом внимания, которая формирует и извлекает изображения зон интереса из исходного снимка и на их основе совершает предсказания относительно каждой извлечённой части.

3.1.1 Карты признаков

Как упоминалось во второй главе, карты признаков, можно интуитивно интерпретировать как, матрицы по которым можно определить наличие или отсутствие признака закодированного в фильтре. Таким образом дополнительный анализ данных матриц, потенциально может помочь улучшить качество классификации.

Исходя из выводов второй главы, предложенный метод анализа карт признаков, содержит в себе некоторые недостатки, а именно потеря данных за счёт отсутствия концентрации на определённых признаках. Для устранения данного недостатка был предложен метод, позволяющий фокусироваться на каждом из классов.

Для этого необходимо для входного изображения вычислить градиент предсказанного класса по отношению к функции активации последнего

свёрточного слоя. Получив градиент выходного нейрона относительно выходной карты признаков последнего сверточного слоя, мы высчитываем среднюю интенсивность для всех карт признаков. Далее необходимо умножить полученные значения градиента на матрицы карты признаков выходного слоя, а затем просуммировать их. Таким образом мы можем получить предполагаемую область локализации для каждого класса, предсказанного сетью. Пример данного подхода, можно увидеть на следующем рисунке 14.

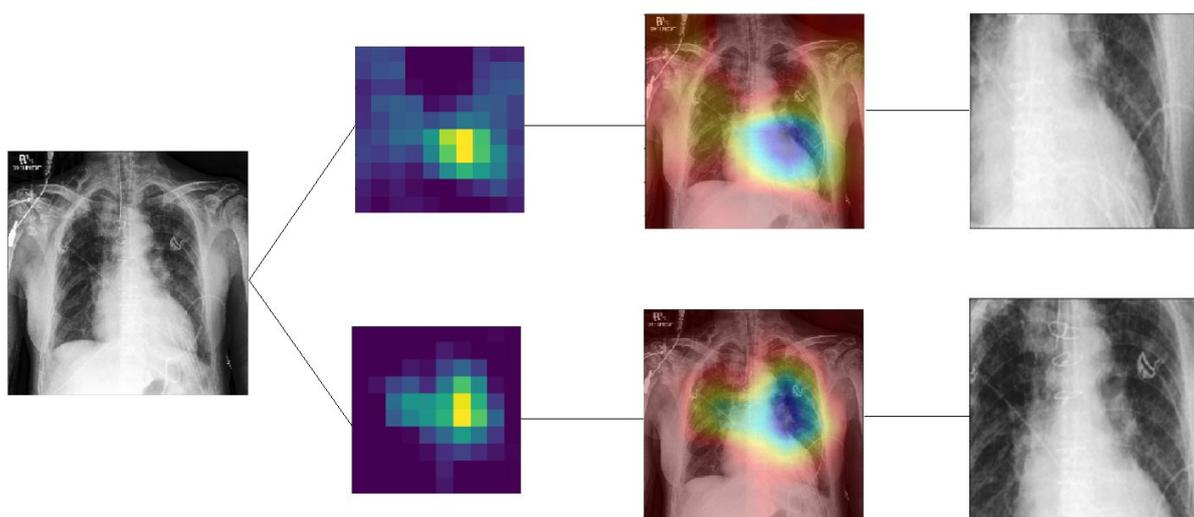


Рисунок 14. Пример получения различных областей интереса в зависимости от предсказанного класса

3.1.2 Механизм выделения областей интереса

Для дальнейшей обработки изображений, необходимо выделить на них, зоны интереса, которые в дальнейшем будут использованы как входные изображения для следующей свёрточной сети. Чтобы выделить такие зоны необходимо сначала сформировать бинарную маску для изображения, здесь возможно два подхода. Первый, это формирование маски на основе порогового значения. Данный метод является самым простым, но имеет недостаток в том, что вводится дополнительный гиперпараметр, величина порога. Лишние гиперпараметры в ходе обучения затрудняют задачу поиска

их оптимальных значений. Для того чтобы этого избежать в данной работе будет использован способ бинаризации с автоматическим выбором порога. Данный метод называется бинаризацией Оцу. Суть данного метода заключается в поиске такого порога, который минимизировал бы дисперсию внутри классов.

После получения бинарного изображения, необходимо выделить на нём интересующую область, данная область является таковой в смысле наибольшего значения связанных между собой пикселей. Пиксели считаются связанными, если они имеют соседей с таким же значением по горизонтали, вертикали и диагонали. Группы связанных пикселей нумеруются различными метками, а затем подсчитывается в какой группе находится наибольшее количество пикселей. Затем из исходного изображения выделяется прямоугольная область, соответствующая максимальным и минимальным значениям по горизонтали и вертикали, для наибольшей группы связанных пикселей. Схематическое изображение данного процесса можно увидеть на рисунке 15.

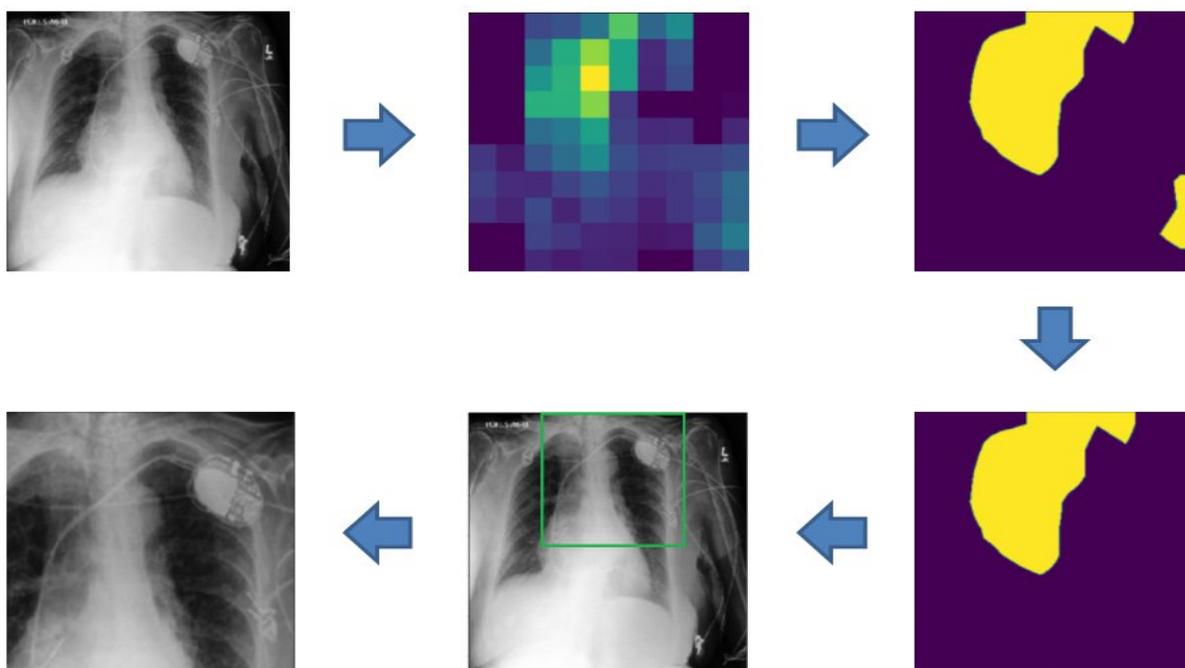


Рисунок 15. Процесс выделения областей интереса на снимке

Важно отметить, что суммирование производится по значениям интенсивности пикселей на карте признаков и результат делится на размер области, выраженный в количестве пикселей в группе. Это делается для того чтобы модель не акцентировала внимание на больших участках с небольшими значениями пикселей.

3.1.3 Обучение модели

На основе метода, описанного выше, для каждого изображения формируются изображения с областями интереса, количество таких изображений равно количеству предсказываемых классов. Изображения для которых не получилось выделить связные области, такой случай может произойти если модель верно предсказала отсутствие патологии на снимке, заменяются на изображения заполненные нулевыми пикселями. Для каждой полученной зоны интереса ставится в соответствие набор меток, где метка соответствующая наличию патологии обозначается единицей, а все остальные нулю.

В случае отсутствия патологии на выделенной зоне интереса все метки будут назначены как нули. Таким образом, мы получаем новый набор данных для обучения который более чем в 10 раз больше оригинального. В соответствии с этим можно рассмотреть два подхода к обучению. Первый это обучение модели на всём полученном наборе данных, а второй заключается в обучении на том наборе данных, который содержит хотя бы одно ненулевое значение.

3.2 Модель на основе анализа частей изображения

Так как, первый подход основан на автоматическом извлечении зон интереса из исходного снимка, это может привести к некоторым проблемам при обучении данной модели. Так как сам процесс извлечения признаков может занимать много времени и кроме того, он сильно зависит от качества работы базового классификатора. Для устранения, потенциальных проблем, был предложен метод лишённый данного недостатка, но сильно зависящий

от истинности исходных предпосылок, лежащих в его основе. Данный подход основан на анализе различных частей исходного снимка, выделенных вручную.

3.2.1 Выделение областей

Для данного подхода области интереса выделялись вручную, исходя из анализа присутствующих в наборе снимков патологий. Для этого на исходном изображении было выделено 5 зон интереса. Левая и правая верхние зоны лёгкого, в этих зонах можно наблюдать различные виды поражения лёгкого. Левая и правая нижние зоны лёгкого, на данных областях часто можно обнаружить различные формы отёка. Центральная зона снимка, на которой можно различить изменения сердечного силуэта. Выделенные сегменты можно увидеть на рисунке 16.

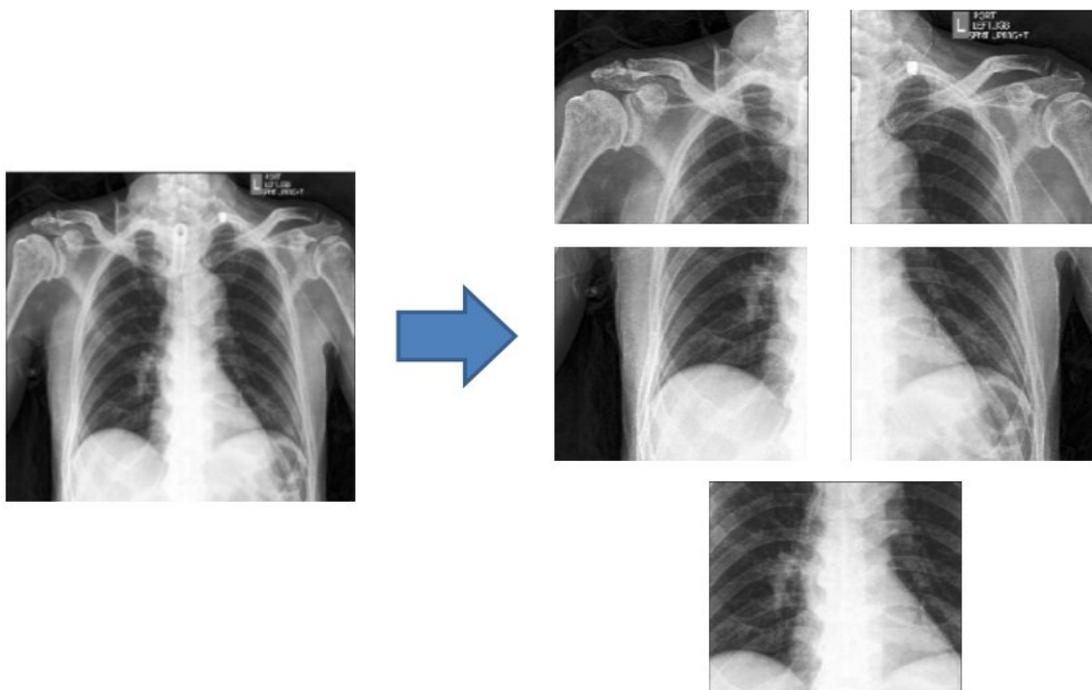


Рисунок 16. Сегментация исходного снимка на области интереса

3.2.2 Архитектура модели

В качестве модели классифицирующей области интереса выделенные вручную, был предложена архитектура свёрточной нейронной базирующаяся

на популярной архитектуре MobileNet. Данная архитектура была выбрана потому что при относительно небольшом количестве обучаемых параметров она показывает результаты сравнимые с более «тяжеловесными моделями». Подход заключается в том, что области интереса, извлечённые из исходного снимка подаются на вход соответствующим свёрточным сетям, у которых исключён финальный слой пулинга. Поэтому на выходе формируют наборы карт признаков, которые объединяются в один вектор, который передаётся слою трёхмерного пулинга, выход из данного слоя передаётся дальше на полносвязные слои. Архитектура сети изображена на рисунке 17.

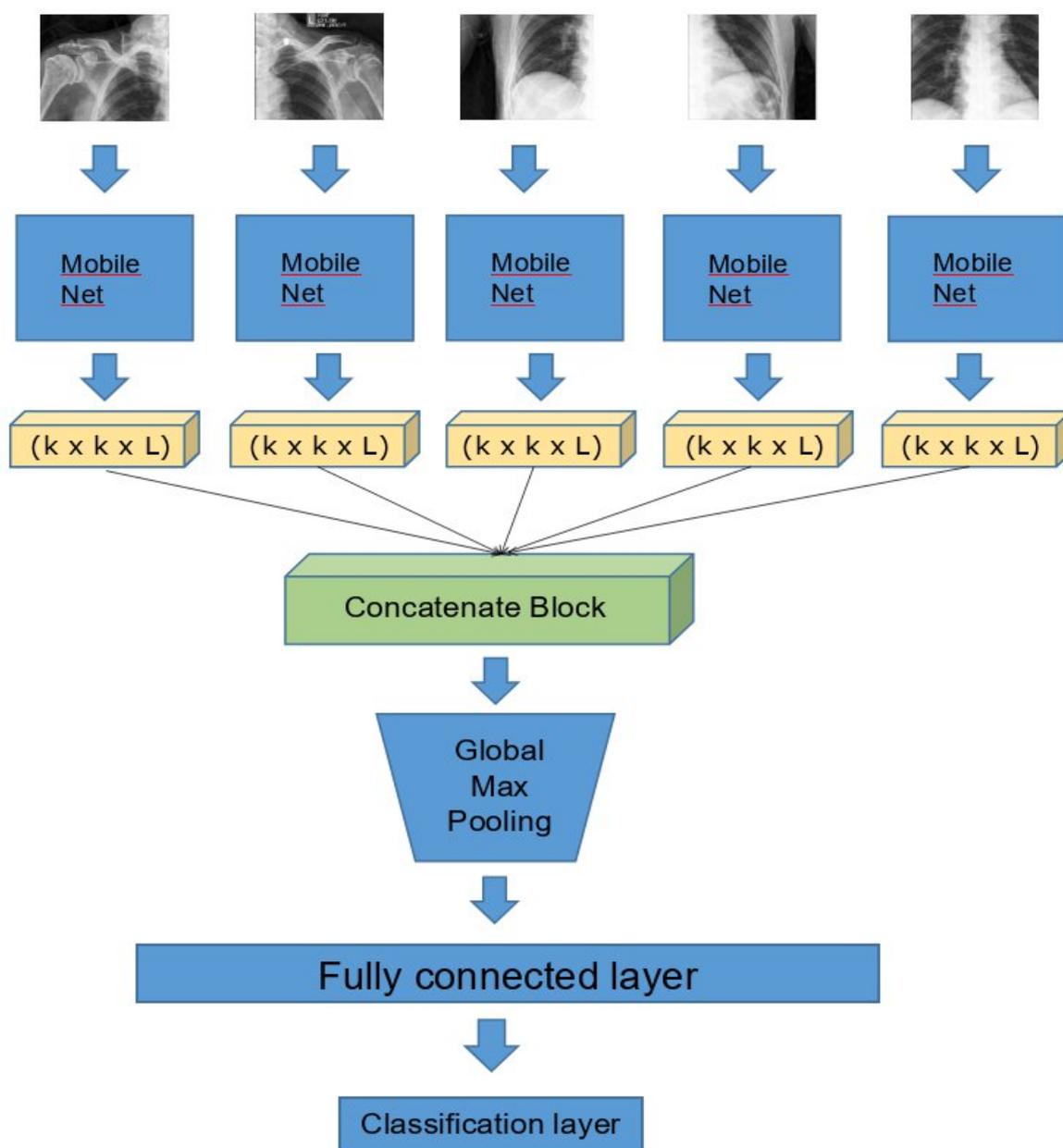


Рисунок 17. Архитектура модели

3.3 Ансамблирование моделей

Для улучшения качества классификации описанные выше модели объединяются в ансамбль. Базовой моделью для данного ансамбля является свёрточная нейронная сеть архитектуры DenseNet. Схему предложенной архитектуры можно увидеть на рисунке 18.

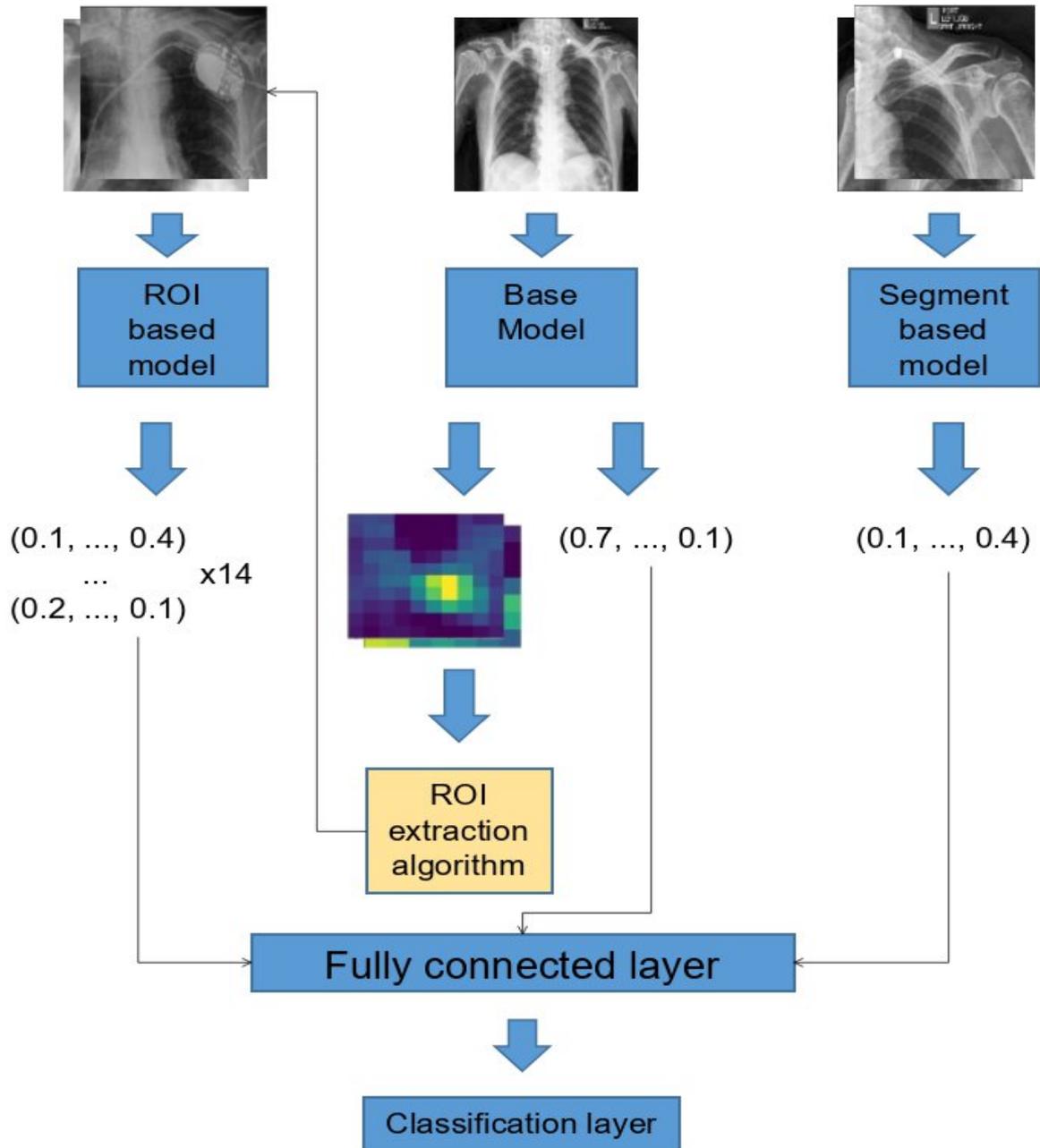


Рисунок 18. Архитектура итогового ансамбля моделей

Базовая модель используется для первичной классификации и для формирования карты признаков для модели внимания. Сформированные карты признаков, подаются на вход модели внимания, которая формирует на

выходе вектора с предсказаниями, количество которых равно количеству классов в представленном наборе. Модель классификации областей выделенных вручную, так же формирует на выходе вектора с предсказаниями относительно классов. Все полученные вектора объединяются в один и передаются на вход полносвязной нейронной сети для осуществления финальной классификации.

Так же стоит отметить, что предложенный подход позволяет значительно сэкономить вычислительные ресурсы. Так, например в рамках используемого набора данных были представлены снимки в двух форматах: DICOM и jpeg. DICOM (Digital Imaging and Communications in Medicine) это медицинский формат создания, хранения, передачи и визуализации цифровых медицинских изображений и документов обследованных пациентов. Как правило, изображения представленные в таком формате, занимают большой объём памяти на жёстком диске. Для сравнения набор данных ChexPert в формате DICOM занимает около 500 гигабайт памяти, а в формате jpeg около 11 гигабайт. Таким образом, в рамках данного подхода, можно обучить базовую модель на данных, которые занимают меньше объёма памяти на диске, а следовательно, при процессе обучения используют меньше оперативной и видео памяти. Таким образом мы получаем информацию о локализации патологий. Затем эта информация может быть использована применительно к набору данных, который занимает большее количество памяти, но теперь уже нет необходимости использовать изображения целиком, а можно использовать только части изображения с выделенными областями интереса. Таким образом данный подход позволяет сэкономить вычислительные ресурсы. Особенно это эффективно для снимков со сверхвысоким разрешением.

3.4 Выводы

Исходя из анализа, приведённого в предыдущих двух главах, были предложены несколько путей решения обозначенных проблем. Был

предложен метод для дополнительного анализа областей интереса, формируемых свёрточной нейронной сетью. Такой подход решает проблему анализа небольших участков исходного снимка и потенциально способен улучшить качество классификации базовой модели. Так же был предложен способ выделения областей интереса без помощи свёрточных сетей и предложен вариант модели, обрабатывающий такие области интереса. В итоге данные подходы были объединены в одном ансамбле вместе с полносвязной нейронной сетью, для осуществления финальной классификации.

Глава 4. Реализация метода

Для реализации были выбраны подходы рассмотренные в третьей главе. Были реализованы указанные архитектуры моделей, также был реализован механизм выделения областей интереса. Было построено базовое решение на основе ансамбля популярных архитектур свёрточных сетей, данные архитектуры кратко описаны во второй главе. Результаты базовой модели использовались для сравнения с показателями полученными предложенными методами.

4.1 Структура программного обеспечения

Данные методы были реализованы на языке Python. Данный язык программирования был выбран из-за своей популярности и доступности множества модулей для работы с данными и реализации многих алгоритмов машинного обучения.

В ходе реализации методов были использованы библиотеки: OpenCV, PIL, skimage - использовались при обработке изображений, а так же для написания алгоритма сегментации зон интереса, tensorflow и keras -для построения и обучения алгоритмов машинного обучения.

Вычисления производились на графической карте Nvidia GeForce GTX 1660 Ti с 6 гигабайтами видеопамяти и на удалённом сервере предоставляемого Google Colaboratory.

4.2 Базовая модель

Данная модель построена для того чтобы сравнить, полученные предложенными методами результаты с результатами работы популярных архитектур свёрточных сетей. Для построения модели использовались такие сети как ResNet50, DenseNet121 и MobileNet. Данные модели были обучены на изображениях с размером 320 на 320 пикселей. Каждая модель обучалась в течение 5 эпох. Затем проводилось мягкое и жёсткое голосование по всем моделям. Результаты работы ансамбля можно увидеть в следующей таблице.

Для оценки качества классификации в данной работе используется метрика AUC ROC (площадь под кривой ошибок). Данная кривая ошибок определяется как зависимость доли верных положительных классификаций от доли ложных положительных классификаций при варьировании порога решающего правила. Она строится на осях TPR (True Positive Rate) и FPR(False Positive Rate) которые определяют количество верно и неверно классифицированных примеров относительно всего объёма меток с положительным и нулевыми значениями соответственно.

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

$$FPR = \frac{FP}{FP + TN}$$

где TP - количество верно предсказанных положительных меток;

TN - количество верно предсказанных отрицательных меток;

FP - количество ложно предсказанных положительных меток;

FN - количество ложно предсказанных отрицательных меток.

ROC кривая строится на графике как отношение TPR к FPR для различных порогов, для которых мы считаем метку положительной. При понижении порога классификации больше элементов классифицируются как положительные, что увеличивает как ложные, так и истинные положительные результаты. AUC ROC оценивается как площадь под полученным графиком. Площадь под данной кривой служит для численного выражения качества классификации. Значения данной метрики получаются определяются для каждого класса в отдельности.

Результаты классификации моделей, выраженные относительно данной метрики, можно увидеть в Таблице 2.

Таблица 2. Результаты предсказаний базовой модели

Класс/Модель	DenseNet121	ResNet50	MobileNet
No Finding	0.84	0.82	0.84
Enlarged Cardiomediastinum	0.64	0.60	0.59
Cardiomegaly	0.82	0.78	0.79
Lung Opacity	0.73	0.69	0.71
Lung Lesion	0.72	0.66	0.69
Edema	0.80	0.77	0.78
Consolidation	0.71	0.68	0.69
Pneumonia	0.72	0.66	0.70
Atelectasis	0.69	0.68	0.70
Pneumothorax	0.77	0.72	0.70
Pleural Effusion	0.85	0.84	0.84
Pleural Other	0.69	0.69	0.66
Fracture	0.74	0.68	0.67
Support Devices	0.81	0.77	0.80

Усреднённые результаты по мягкому голосованию: 0.771

Усреднённые результаты по жёсткому голосованию: 0.753

4.3 Предложенный подход

Для настройки работы ансамбля использовались различные параметры. Так для базовой модели и модели основанной на анализе сегментов в качестве функции потерь использовалась взвешенная функция потерь. В зависимости от соотношения коэффициентов, переданных для данной функции, она будет в большую или меньшую сторону штрафовать ложно положительные результаты. Для нашего подхода, мы уменьшили штрафы за ложно положительные предсказания, так как мы заинтересованы в том, чтобы модель, пусть даже и ложно предположила наибольшее число патологий. Так как в ходе выделения областей интереса данные области могут быть отброшены как неверно помеченные, но получить положительный результат в сегменте изображения, который базовая модель не выделила намного сложнее. В целом такой подход имеет больший смысл для задач

классификации медицинских изображений, так как зачастую выгоднее ложно определить наличие патологии, чем ложно определить её отсутствие.

Используемые параметры и размеры входных изображений для моделей можно увидеть на таблице 3.

Таблица 3. Параметры для моделей

Модель	Архитектура	Функция потерь	Параметры функции потерь	Размер входного слоя
Base model	DenseNet121	weighted binary crossentropy	a = 1 b = 0.8	(320,320,3)
Segment based model	DenseNet121	weighted binary crossentropy	a = 1 b = 0.8	(224 ,224,3)
ROI attention model	DenseNet121	binary crossentropy	-	(224,224,3)
Uniting model	FCN	binary crossentropy	-	(224,)

В таблице 4 представлены метрики для результатов прогноза, отдельно базовой модели и её же в составе различных ансамблей.

Таблица 4. AUC ROC для различных ансамблей моделей

Модель	Avg AUC ROC
Base model	0.771
Base model + Segment based model	0.783
Base model + ROI attention model	0.78
Base model + ROI attention model + Segment based model	0.793

В таблице 5 представлены метрики AUC ROC для итогового ансамбля моделей, для каждого класса в отдельности, также в данной таблице приведены соответствующие значения для базовой модели.

Таблица 5. AUC ROC для каждого класса

Класс/Модель	Предложенный подход	Базовый ансамбль моделей
No Finding	0.85	0.85
Enlarged Cardiomediastinum	0.70	0.67
Cardiomegaly	0.85	0.85
Lung Opacity	0.73	0.73
Lung Lesion	0.78	0.76
Edema	0.81	0.81
Consolidation	0.77	0.74
Pneumonia	0.74	0.72
Atelectasis	0.70	0.67
Pneumothorax	0.86	0.83
Pleural Effusion	0.88	0.87
Pleural Other	0.74	0.71
Fracture	0.80	0.75
Support Devices	0.88	0.85

Исходя из данных, представленных в Таблице 5, можно увидеть, что значения AUC ROC возросли больше остальных, для таких классов как: перелом и вспомогательные устройства. Это может быть связано с тем, что предложенный подход позволяет лучше анализировать небольшие участки на снимке.

4.4 Выводы

Исходя из анализа результатов, можно сделать вывод, что предложенный метод, действительно смог улучшить результаты классификации рентген снимков. Данный подход позволяет повторно анализировать предположения сделанные базовой моделью, и выявлять отсутствие или наличие признаков на выделенных областях интереса. Стоит отметить, что исходя из анализа набора данных, предложенный подход может быть успешно применён и к другим медицинским изображениям, так как стало ясно из анализа данных, выявленная специфика характерна для большинства наборов медицинских данных.

Заключение

В ходе данной работы был проведён анализ популярных наборов медицинских данных. В результате данного анализа была выявлена общая специфика для наборов медицинских изображений. Были проанализированы популярные подходы учитывающие данную специфику. На основе анализа данных подходов были предложены решения, которые потенциально могут исправить недостатки рассмотренных решений. Были разработаны и реализованы архитектуры свёрточных сетей, необходимые для реализации предложенного подхода. Было реализовано ансамблирование моделей. Был предложен и реализован механизм выделения зон интереса на изображениях. В итоге предложенная модель позволила улучшить точность классификации относительно подхода реализованного с помощью ансамбля популярных архитектур свёрточных сетей.

Список литературы

- [1] Hao R., Namdar K., Liu L., Khalvati F. A Transfer Learning Based Active Learning Framework for Brain Tumor Classification [Электронный ресурс] // arXiv preprint arXiv:2011.09265. – 2020 URL: <https://arxiv.org/abs/2011.09265> (дата обращения 15.04.2021).
- [2] Khan M.A., Ashraf I., Alhaisoni M., Damaševičius R., Scherer R., Rehman A. Multimodal Brain Tumor Classification Using Deep Learning and Robust Feature Selection: A Machine Learning Application for Radiologists //Diagnostics (2075-4418). – 2020. – Т. 10. – №. 8.
- [3] Chaunzwa L.T, Hosny A., Xu Y., Shafer A., Diao N., Lanuti M., Christiani D., Mak R. Deep learning classification of lung cancer histology using CT images //Scientific reports. – 2021. – Т. 11. – №. 1. – С. 1-12.
- [4] Sathyakumar K., Munoz M., Singh J., Hussain N., Babu B.A. Automated Lung Cancer Detection Using Artificial Intelligence (AI) Deep Convolutional Neural Networks: A Narrative Literature Review //Cureus. – 2020. – Т. 12. – №. 8.
- [5] Khan H.A., Jue W., Mushtaq M., Mushtaq U.M. Brain tumor classification in MRI image using convolutional neural network //Mathematical Biosciences and Engineering. – 2020. – Т. 17. – №. 5. – С. 6203-6216.
- [6] LaMontagne P.J., Benzinger T., Morris J.C., Keefe S., Hornbeck R., Xiong C., Grant E., Hassenstab J., Moulder K., Vlassenko A.G., Raichle M.E., Cruchaga C., Marcus D. OASIS-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer disease [Электронный ресурс] //MedRxiv. – 2019 URL: <https://www.medrxiv.org/content/10.1101/2019.12.13.19014902v1> (дата обращения 16.04.2021).
- [7] Islam J., Zhang Y. An ensemble of deep convolutional neural networks for Alzheimer's disease detection and classification [Электронный ресурс] //arXiv preprint arXiv:1712.01675. – 2017 URL: <https://arxiv.org/abs/1712.01675> (дата обращения 16.04.2021).
- [8] Wen J., Thibeau-Sutre E., Diaz-Melo M., Samper-González J., Routier A., Bottani S., Dormont D., Durrleman S., Burgos N., Colliot O. Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation //Medical image analysis. – 2020. – Т. 63. – С. 101694.
- [9] Heath M., Bowyer K., Kopans D., Kegelmeyer P., Moore R., Chang K., Munishkumaran S. Current status of the digital database for screening

- mammography //Digital mammography. – Springer, Dordrecht, 1998. – С. 457-460.
- [10] Sun H., Li C., Liu B., Liu Z., Wang M., Zheng H., Feng D.D., Wang S. AUNet: Attention-guided dense-upsampling networks for breast mass segmentation in whole mammograms //Physics in Medicine & Biology. – 2020. – Т. 65. – №. 5. – С. 055005.
- [11] Xi P., Shu C., Goubran R. Abnormality detection in mammography using deep convolutional neural networks //2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA). – IEEE, 2018. – С. 1-6.
- [12] Menze B.H., Jakab A., Bauer S., Kalpathy-Cramer J., Farahani K, Kirby J., et al. The multimodal brain tumor image segmentation benchmark (BRATS) //IEEE transactions on medical imaging. – 2014. – Т. 34. – №. 10. – С. 1993-2024.
- [13] Bakas S., Akbari H., Sotiras A, Bilello M., Rozycki M., Kirby J.S., et al., Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features //Scientific data. – 2017. – Т. 4. – №. 1. – С. 1-13.
- [14] Bakas S., Reyes M., Jakab A., Bauer S., Rempfler M., Crimi A., et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge [Электронный ресурс] //arXiv preprint arXiv:1811.02629. – 2018 URL: <https://arxiv.org/abs/1811.02629> (дата обращения 15.04.2021).
- [15] Awasthi N., Pardasani R., Gupta S. Multi-Threshold Attention U-Net (MTAU) based Model for Multimodal Brain Tumor Segmentation in MRI scans [Электронный ресурс] //arXiv preprint arXiv:2101.12404. – 2021 URL: <https://arxiv.org/abs/2101.12404> (дата обращения 21.04.2021).
- [16] Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation //International Conference on Medical image computing and computer-assisted intervention. – Springer, Cham, 2015. – С. 234-241.
- [17] Oktay O., Schlemper J., Folgoc L.L., Lee M., et al. Attention u-net: Learning where to look for the pancreas [Электронный ресурс] //arXiv preprint arXiv:1804.03999. – 2018 URL: <https://arxiv.org/abs/1804.03999> (дата обращения 16.04.2021).

- [18] Irvin J., Rajpurkar P., Ko M., Yu Y., et al. Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison //Proceedings of the AAAI Conference on Artificial Intelligence. – 2019. – Т. 33. – №. 01. – С. 590-597.
- [19] Электронная энциклопедия радиологии Radiopedia.org [Электронный ресурс] URL: <https://radiopaedia.org/articles/> (дата обращения 12.04.2021)
- [20] He K., Zhang X., Ren S., Sun J., et al. Deep residual learning for image recognition //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – С. 770-778.
- [21] Huang G., Liu Z., Maaten L., Weinberger K.Q. Densely connected convolutional networks //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – С. 4700-4708.
- [22] Chollet F. Xception: Deep learning with depthwise separable convolutions //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – С. 1251-1258.
- [23] Sandler M., Howard A., Zhu M., Zhmoginov A., Chen L. Mobilenetv2: Inverted residuals and linear bottlenecks //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2018. – С. 4510-4520.
- [24] Guan Q., Huang Y., Zhong Z, Zheng Z., Zheng L., Yang Y. Diagnose like a radiologist: Attention guided convolutional neural network for thorax disease classification [Электронный ресурс] //arXiv preprint arXiv:1801.09927. – 2018 URL: <https://arxiv.org/abs/1801.09927> (дата обращения 15.04.2021).
- [25] Kumar P., Grewal M., Srivastava M. M. Boosted cascaded convnets for multilabel classification of thoracic diseases in chest radiographs //International Conference Image Analysis and Recognition. – Springer, Cham, 2018. – С. 546-552.
- [26] Wang X., Peng Y., Lu L., Lu Z., Bagheri M., Summers R.. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – С. 2097-2106.
- [27] Yao L., Poblenz E., Dagunts D., Covington B, Bernard D., Lyman K. Learning to diagnose from scratch by exploiting dependencies among labels

[Электронный ресурс] //arXiv preprint arXiv:1710.10501. – 2017 URL:
<https://arxiv.org/abs/1710.10501> (дата обращения 15.04.2021).