

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ
Кафедра дискретной математики и алгоритмики

ШЕПЕЛЕВА Виктория Дмитриевна

**АВТОМАТИЧЕСКОЕ ОБНАРУЖЕНИЕ АНОМАЛИЙ ГРУДНОЙ
КЛЕТКИ НА ОСНОВЕ АНАЛИЗА ЦИФРОВЫХ РЕНТГЕНОВСКИХ
ИЗОБРАЖЕНИЙ**

Магистерская диссертация
специальность 1-31 81 09 «Алгоритмы и системы обработки больших объемов
информации»

Научный руководитель
Ковалев Василий Алексеевич
кандидат технических наук

Допущена к защите

«__» _____ 2020 г.

Зав. кафедрой дискретной математики и алгоритмики

_____ В. М. Котов

доктор физ.-мат. наук, профессор

Минск, 2020

ОГЛАВЛЕНИЕ

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ.....	4
АГУЛЬНАЯ ХАРАКТАРЫСТЫКА РАБОТЫ	5
ABSTRACT	6
ВВЕДЕНИЕ	7
ГЛАВА 1.....	8
ОСНОВНЫЕ ПОНЯТИЯ И ОБЗОР ЛИТЕРАТУРЫ.....	8
1.1 Общая информация о нейронных сетях	9
1.2 Сверточные нейронные сети.....	10
1.3 Архитектуры моделей сверточных нейронных сетей	11
1.4 Обзор существующих подходов.....	14
1.5 Выводы.....	16
ГЛАВА 2.....	17
ПРИМЕНЕНИЕ МЕТОДОВ КЛАССИФИКАЦИИ ПАТОЛОГИЙ ГРУДНОЙ КЛЕТКИ	17
2.1 Исходные данные.....	17
2.2 Предобработка изображений	22
2.2.1 Эквиализация.....	22
2.2.1 Сегментация	24
2.2.3 Добавление шума	25
2.4 Эксперименты	26
2.4.1 Multi-label классификация	26
2.4.2 ROC-кривая, AUC	27
2.4.3 Результаты	28
2.4.2 Взвешенная функция потерь	31
2.5 Выводы.....	32
ГЛАВА 3.....	33
ТЕПЛОВЫЕ КАРТЫ	33
3.1 Алгоритмы построения тепловых карт.....	33
3.1.1 Коэффициент корреляции Пирсона	33
3.1.2 Логистическая регрессия	34

3.1.3 Grad-CAM	35
3.1.4 Альтернативный подход к построению тепловых карт.....	36
3.2 HSV	36
3.3 Предобработка изображений	38
3.3.1 Алгоритм Inpainting	38
3.4 Эксперименты	39
3.5 Выводы.....	44
ЗАКЛЮЧЕНИЕ	45
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	46

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Магистерская диссертация, 47 стр., 25 рис., 7 табл., 16 источников

РЕНТГЕНОВСКИЕ ИЗОБРАЖЕНИЯ, ПАТОЛОГИИ ГРУДНОЙ КЛЕТКИ, ЗАДАЧА КЛАССИФИКАЦИИ, НЕЙРОННАЯ СЕТЬ, ТЕПЛОВЫЕ КАРТЫ

Объект исследования – методы классификации патологий грудной клетки, построения тепловых карт рентгеновских изображений.

Цель работы – изучение методов классификации патологий грудной клетки, построения тепловых карт, реализация системы автоматизированной диагностики, исследование подходов для повышения качества классификации.

Методы исследования – анализ, эксперимент, тестирование, сравнение.

Результаты исследования:

1. Изучены подходы к решению задачи автоматического обнаружения патологий грудной клетки;
2. Подготовлены исходные данные, рассмотрены и реализованы методы предобработки изображений;
3. Реализованы методы классификации патологий грудной клетки, проведены вычислительные эксперименты;
4. Исследованы и реализованы методы построения тепловых карт, проведены эксперименты по улучшению локализации очагов патологий;
5. Получена нейросетевая модель для использования в системе автоматизированной диагностики.

Область применения – медицина, программное обеспечение рентгенодиагностики.

АГУЛЬНАЯ ХАРАКТЕРЫСТЫКА РАБОТЫ

Магістарская дысертацыя, 47 с., 25 мал., 7 табл., 16 крыніц

РЭНТГЕНАЎСКІЯ ЗДЫМКІ, ПАТАЛОГІІ ГРУДНОЙ КЛЕТКІ, ЗАДАЧА КЛАСІФІКАЦЫІ, НЕЙРОННАЯ СЕТКА, ЦЕПЛАВЫЯ КАРТЫ

Аб'ект даследавання - метады класіфікацыі паталогій грудной клеткі, пабудовы цеплавых карт рэнтгенаўскіх здымкаў.

Мэта работы - вывучэнне метадаў класіфікацыі паталогій грудной клеткі, пабудовы цеплавых карт, рэалізацыя сістэмы аўтаматызаванай дыягностыкі, даследаванне падыходаў для павышэння якасці класіфікацыі.

Метады даследавання - аналіз, эксперымент, тэставанне, параўнанне.

Вынікі даследавання:

1. Вывучаны падыходы да вырашэння задачы аўтаматычнага выяўлення паталогій грудной клеткі;
2. Падрыхтаваны зыходныя дадзеныя, разгледжаны і рэалізаваны метады перадапрацоўкі малюнкаў;
3. Рэалізаваны метады класіфікацыі паталогій грудной клеткі, праведзены вылічальныя эксперыменты;
4. Даследаваны і рэалізаваны метады пабудовы цеплавых карт, праведзены эксперыменты па паляпшэнню лакалізацыі ачагоў паталогій;;
5. Атрымана нейросетевая мадэль для выкарыстання ў сістэме аўтаматызаванай дыягностыкі.

Вобласць прымянення - медыцына, праграмнае забеспячэнне рэнтгенадыягностыкі.

ABSTRACT

Master thesis, 47 p., 25 figures, 7 tables, 16 sources

**X-RAY IMAGES, CHEST PATHOLOGIES, CLASSIFICATION PROBLEM,
NEURAL NETWORK, HEATMAPS**

Object of research - methods of chest pathologies classification, generation heatmaps of x-ray images.

The aim of the work is to study methods of chest pathologies classification, to generate heatmaps, to develop an automated diagnostic system, to research approaches of classification quality improvement.

Research methods - analysis, experiment, testing, comparison.

The results of the study:

1. The approaches to solving the problem of chest pathologies automatic detection were studied;
2. Initial data was prepared, methods of image preprocessing were reviewed and implemented;
3. Methods of chest pathologies classification were implemented. Research of them was conducted;
4. Methods for generating heatmaps were implemented. Experiments for improving localization pathologic niduses were conducted;
5. A model has been obtained for use in an automated diagnostic system.

Field of application – medicine, X-ray diagnostic software.

ВВЕДЕНИЕ

В современном мире скрининг населения является одним из способов своевременного обнаружения болезней легких, а также различных патологий грудной клетки. Соответственно, чем раньше будет поставлен диагноз, тем эффективнее будет процесс лечения. Важность построения системы автоматизированной диагностики можно аргументировать, например, тем, что в современном мире зафиксировано большое число смертей от туберкулеза и пневмонии. В некоторых случаях постановка диагноза радиологом является затруднительной задачей и требует повторного создания рентгеновского снимка, а также длительного наблюдения за возможными изменениями. Инструменты автоматизированной диагностики могут дополнять процесс принятия решений о диагнозе, одновременно устраняя пробелы в экспертной радиологической интерпретации при проведении скрининга.

В последнее время зачастую используются глубокие нейронные сети, обученные на больших базах данных, для разработки систем компьютерной диагностики. В частности, для анализа рентгеновских цифровых изображений используются сверточные нейронные сети. Они достигают высокой точности при решении задачи классификации патологий грудной клетки, а также активно используются в других областях исследования изображений.

В рамках магистерской диссертации были изучены и реализованы методы классификации патологий грудной клетки, проведены исследования по влиянию предобработки изображений на качество классификации, изучены и исследованы методы построения тепловых карт, проведены эксперименты по улучшению локализации очагов патологий. В частности, в первой главе проведен обзор и анализ существующих подходов, даны основные определения и понятия нейронных сетей, приведено описание основных архитектур сверточных нейронных сетей. Во второй главе дана полная характеристика исходных данных для задачи, описаны методы предобработки изображений, проведенные эксперименты для классификации, а также приведен сравнительный анализ полученных результатов. В третьей главе проведен обзор методов построения тепловых карт, описаны эксперименты и детали реализации, направленные на построение и улучшение локализации тепловых карт, проиллюстрированы примеры полученных изображений, проведен сравнительный анализ.

ГЛАВА 1

ОСНОВНЫЕ ПОНЯТИЯ И ОБЗОР ЛИТЕРАТУРЫ

Основной задачей системы автоматизированной диагностики является эффективное и быстрое предоставление надежного дополнительного “мнения”, сравнимого с мнением радиолога, о наличии патологии, включая сложные подозрительные случаи. Работа такой системы заключается в том, чтобы для исходного рентгеновского изображения предсказать вероятность наличия патологии, в частности, идентифицировать болезнь, а также предоставить изображение с выделенными очагами злокачественных образований грудной клетки. В данной диссертации рассматривается отдельная область грудной клетки - лёгкие.

Современным примером такой системы является CAD4TB, [1], задача которой состоит в распознавании туберкулеза на рентгенах. В шестой версии программного обеспечения были достигнуты следующие величины метрик: чувствительность (англ. *sensitivity*) 90%, специфичность (англ. *specificity*) 76%. Наряду с этим данная система осуществляет построение тепловых карт локализации очагов туберкулеза.

Для создания системы необходима большая база данных. В частности, в рамках проведения экспериментов использовалась база данных, которая содержит информацию о результатах скрининга населения двухмиллионного города, проводимых в период 2001-2014 г. Типичной особенностью скрининга населения является то, что итоговая база данных содержит большую часть рентгеновских изображений без патологий, в то время как различные виды отклонений представлены пропорционально естественному уровню заболеваний. Трудность построения системы автоматизированной диагностики заключается в том, что в данных присутствуют случаи, когда радиолог не может поставить точный диагноз, а также, когда диагноз поставлен неверно. Соответственно, заранее присутствует определенный процент ошибки. Также в базе данных имеются рентгеновские изображения, содержащие дефекты.

Важной особенностью рентгеновских изображений является то, что они являются полутонными, а также их пониженная контрастность. Таким образом, основной аспект в разработке системы принадлежит не только созданию блока классификации наличия патологии, но и предобработке изображений.

Основными современными методами распознавания изображений являются сверточные нейронные сети. В следующих разделах приведено описание математической модели нейронных сетей и примеров архитектур. Также приведен обзор и анализ существующих подходов по распознаванию патологий грудной клетки.

1.1 Общая информация о нейронных сетях

Нейронная сеть представляет собой математическую модель, которая имитирует поведение человеческого мозга, в частности, синаптические связи между нейронами. В биологии нейрон является клеткой, состоящей из тела, ядра и отростков. Множество ответвлений от тела нейрона называется дендритами. Длинные дендриты называются аксонами. Дендрит принимает сигнал и, соответственно, выступает в роли входа, аксон передает сигнал другому нейрону по синаптической связи и выступает в роли выхода. Тело нейрона играет роль сумматора. Таким образом, математическая модель нейрона показана на рисунке 1, где веса регулируют импульс сигнала, т.е. его важность в последующей передаче.

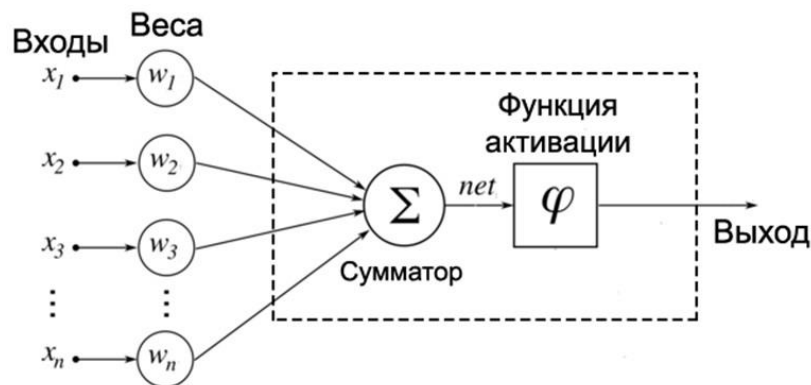


Рисунок 1 – Математическая модель нейрона, [2]

Соответственно, выход проиллюстрированной выше модели вычисляется по формуле:

$$f(x) = \varphi\left(\sum_{i=1}^n \omega_i x_i + b\right) \quad (1)$$

Глубокие нейронные сети представляют собой множество чередующихся слоев линейности и нелинейности, где сумматор является линейным преобразованием, а функция активации - нелинейным. Функция активации значительно влияет на способность к обучению глубоких нейронных сетей, так как, исключив слои нелинейности, сеть сводится к простой математической модели нейрона, которая, в свою очередь, не обладает хорошей способностью решать задачу нелинейно разделимых в пространстве объектов разных классов.

Примеры функций активации:

1. Сигмоидная функция

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

2. Гиперболический тангенс

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3)$$

3. ReLU

$$f(x) = \max(0, x) \quad (4)$$

4. LeakyReLU

$$f(x) = \max(0.1x, x) \quad (5)$$

Одним из широко применяемых методов обучения является обучение с учителем, при котором используются данные с известным для них верным значением выхода. В процессе обучения участвует некоторая заранее заданная функция потерь, которая сравнивает предсказание нейронной сети с уже известным желаемым выходом. Таким образом, задача сводится к тому, чтобы минимизировать значение функции потерь, корректируя определенным образом величины весов слоев нейронной сети. Самым популярным методом обновления весов при обучении является метод обратного распространения ошибки. Он состоит из двух этапов:

1. Прямое распространение ошибки. На вход сети подаются исходные данные, для которых заранее известно верное соответствие. Далее данные проходят этапы вычисления на каждом слое сети, и в течение этого процесса сохраняются значения функции активации, включая последний слой выхода;

2. Обратное распространение ошибки. На этом этапе происходит вычисление значения функции потерь. Далее вычисляются значения частных производных функции потерь по каждому весу на последнем слое. Затем происходит умножение полученных величин на параметр, называемый скоростью обучения, и данные значения используются для обновления весов. После этого происходит передача вычисленных значений частных производных на предыдущий слой, и действия повторяются. Данная процедура изменения весов называется градиентным спуском. Также существует его модификация, а именно стохастический градиентный спуск. Среди других алгоритмов оптимизации можно выделить метод среднеквадратичного распространения ошибки (RMSProp) и метод адаптивной оценки момента (Adam).

1.2 Сверточные нейронные сети

Сверточная нейронная сеть - специальная архитектура нейронной сети, которая в большинстве случаев используется для распознавания образов на изображениях. Такая архитектура учитывает двумерную топологию изображения, что повышает точность распознавания. Также этот тип сети устойчив к небольшим аффинным преобразованиям. Большим преимуществом сверточных сетей перед полносвязными сетями в распознавании изображений

является значительно меньшее число параметров. Основными отличиями сверточных нейронных сетей является наличие сверточных слоев, а также слоев дискретизации (пулинга).

Сверточный слой состоит из набора карт. Каждая карта обладает своим ядром (фильтром). Ядро свертки фильтра обрабатывает выход предыдущего слоя путем прохождения по фрагментам, для каждого фрагмента вычисляется сумма поэлементного произведения. Ядро свертки состоит из весов, которые являются обучаемыми параметрами сети. Таким образом, оно находит определенные признаки на изображении.

Слой субдискретизации (пулинга) представляет собой алгоритм прохода скользящего окна заданного размера по изображению, в процессе которого вычисляется некоторая функция. Примером такой функции может служить вычисление максимума, среднего значения, суммы. Одним из основных параметров данного преобразования является шаг, с которым осуществляется проход. В большинстве случаев шаг выбирается таким образом, чтобы прямоугольники, покрытые скользящим окном, не пересекались. Соответственно, такое преобразование приводит к уменьшению пространства признаков, выделяя наиболее значимые из них. Также можно отметить, что пулинг инвариантен к небольшим сдвигам и поворотам.

1.3 Архитектуры моделей сверточных нейронных сетей

Для VGG16 на вход подаются изображения 224x224. Далее изображение проходит последовательно через сверточные слои с ядром фильтра размера 3x3. Также между ними находятся слои пулинга с окном размера 2x2 и шагом 2. После всех сверточных слоев идет три полносвязных слоя: первые два имеют 4096 каналов, последний 1000. Последний слой является слоем предсказания, поэтому он зачастую подвергается изменениям в связи с поставленной задачей, в частности, изменяется размерность выхода и применяемая после него функция активации. Архитектура сети изображена на рисунке 2.

Для DenseNet121 на вход подаются изображения 224x224, [4]. Архитектура сети состоит из последовательно чередующихся Dense блоков и Transition блоков. Dense блок представляет собой последовательно применения попарно сверток 1x1 и 3x3. Каждый слой в этом блоке принимает на вход все предыдущие карты признаков этого же слоя, при этом признаки конкатенируются. Transition блок состоит из свертки 1x1 и усредненного пулинга с окном 2x2 и шагом 2. Последние два слоя: глобальный усредненный пулинг с окном 7x7 и полносвязный слой на 1000 каналов. Архитектура сети изображена на рисунке 3.

Сеть GoogleNet состоит из девяти Inception модулей, рис. 4, [5]. На вход сети подаются изображения 224x224. Inception модуль представляет собой

параллельную комбинацию сверточных фильтров 1×1 , 3×3 , 5×5 . Отличительная особенность заключается в применении сверток 1×1 . Они значительно уменьшают количество признаков, что, в свою очередь, сильно сокращает объем вычислений. Такой способ избавления от избыточности аргументирован тем, что входные признаки коррелированы. Также GoogleNet имеет три training head, в каждой из которых есть полносвязные слои для предсказания классов. Это было сделано для обеспечения предотвращения дискриминации признаков на ранних слоях обучения и регуляризации. При обучении сети функции потерь каждого блока классификации суммируются с весами. Архитектура сети представлена на рисунке 5.

Для InceptionV3 на вход подаются изображения размера 299×299 . Сеть представляет собой модификацию GoogleNet. Изменения для Inception модуля заключаются в другом порядке применения сверточных фильтров, а также в замене MaxPooling на AveragePooling. Архитектура сети изображена на рисунке 6.

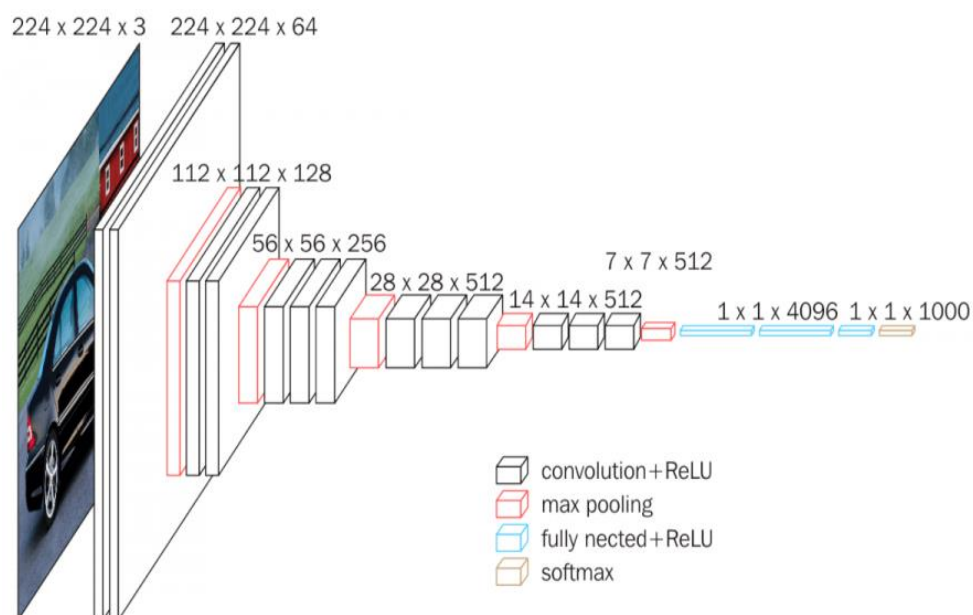


Рисунок 2 – Архитектура сети VGG16, [3]

Архитектура Xception характеризуется наличием сепарабельных сверток, исходя из гипотезы о том, что межканальные и пространственные корреляции можно полностью разделить, [7]. Сеть состоит из 36 сверточных слоев, которые формируют основу извлечения признаков. Эти сверточные слои разделены на 14 модулей, т.е. архитектура представляет собой стек из сепарабельных сверток и остаточных соединений, см. рис. 7. Сепарабельные свертки раскладывают обычную свертку на две, одна из которых анализирует межканальную корреляцию, а другая – пространственную корреляцию в пределах одного канала. Это обеспечивает снижение параметров сети, а также ресурсов, необходимых для обучения. На вход сети подаются изображения 299×299 .

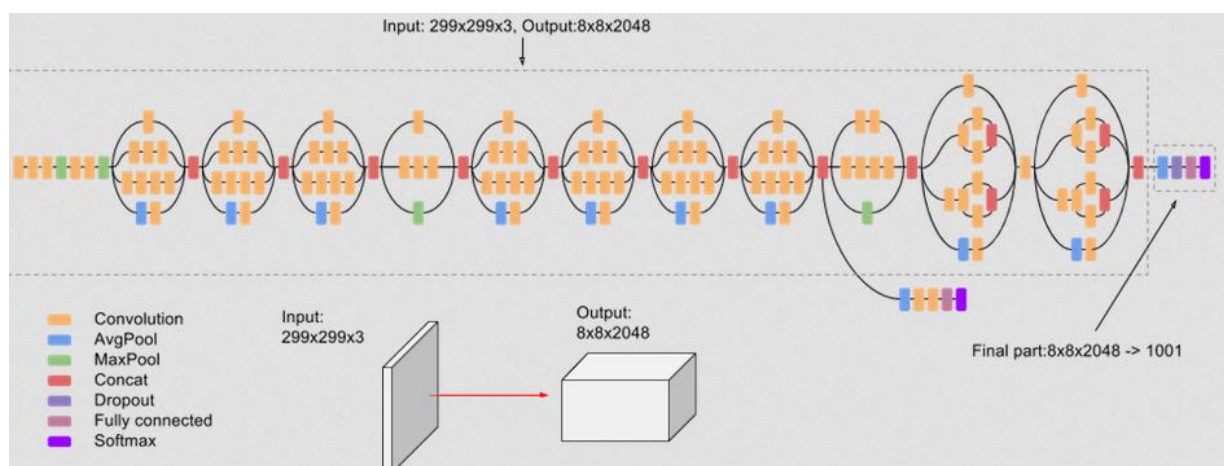


Рисунок 6 – Архитектура сети InceptionV3, [6]

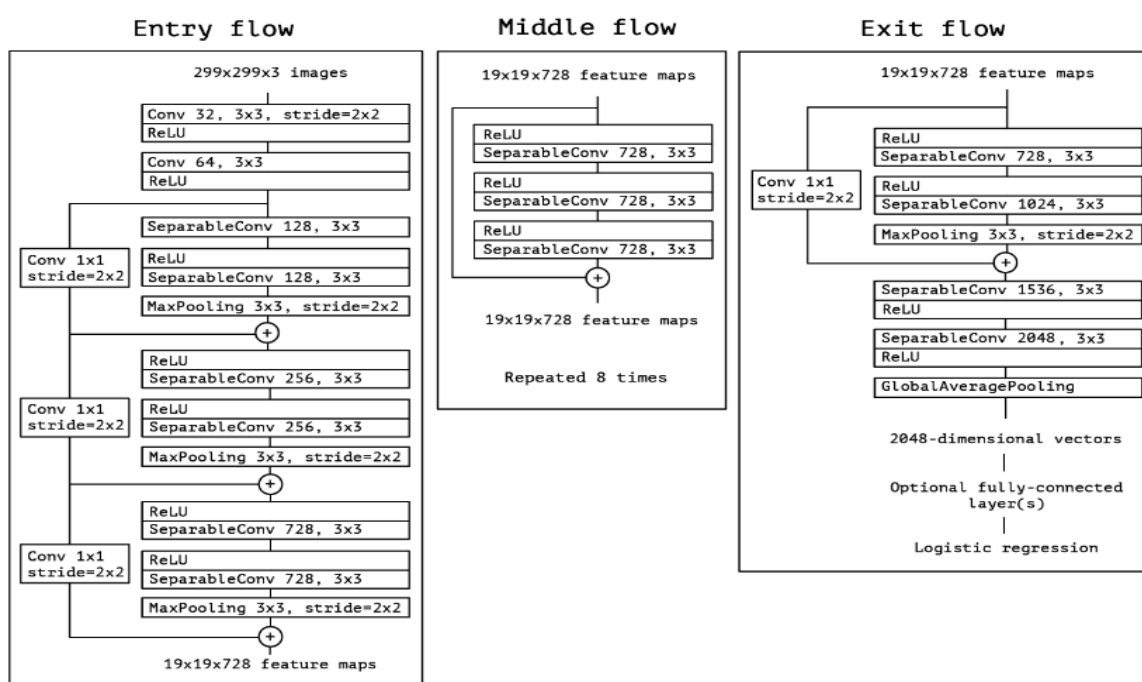


Рисунок 7 – Архитектура сети Xception, [7]

Особенность архитектуры ResNet заключается в использовании блоков, состоящих из двух обучаемых слоев, к выходу которых добавляется значение входа для этого блока. Такие блоки позволяют обучать сеть, состоящую из нескольких сотен, тысячи слоев. Также в данной модели применяются bottleneck блоки, состоящие из применения свертки 1x1 с меньшим числом карт (обычно, четверть от входа), затем свертка 3x3 с таким же количеством карт, в конце свертка 1x1 для увеличения числа признаков (обычно, равное числу признаков входа блока).

1.4 Обзор существующих подходов

В работе [8] были проведены эксперименты с использованием сверточных нейронных сетей для классификации наличия туберкулеза на рентгеновских

снимках. Препроцессинг состоял из получения маски легких после сегментации, затем исходное изображение обрезалось по границам полученной маски. Для классификации использовались следующие модели сверточных нейронных сетей, предобученные на imagenet: AlexNet, VGG-16, VGG-19, ResNet-50, Xception. Процесс обучения происходил только для небольшого блока полносвязных слоев, расположенного за последним сверточным слоем. В статье лучшие результаты показала сеть AlexNet. Наряду с перечисленными моделями рассматривалась небольшая сверточная нейронная сеть, состоящая из трех сверточных слоев с фильтром 3x3, после каждого из которых применялся MaxPooling с окном 2x2 и шагом 2. Затем расположен полносвязный слой и слой Dropout. В конце расположен полносвязный слой предсказания с двумя выходами и функцией активации softmax. Результаты такой модели оказались хуже в сравнении с AlexNet по метрикам точности и AUC.

В статье [4] были проведены эксперименты для задачи бинарной классификации наличия пневмонии на рентгеновских изображениях и multi-label классификации 14 болезней, а также были построены тепловые карты расположения очагов инфекции. Для классификации использовалась модель сверточной нейронной сети DenseNet-121, инициализированная весами модели, предобученной на imagenet. Для бинарной классификации последние полносвязный слой были заменены на один полносвязный слой с одним выходом и сигмоидной функцией активации. В качестве функции потерь использовалась взвешенная бинарная перекрёстная энтропия. В процессе обучения участвовали все слои сети. Аугментация исходных данных заключалась в применении случайным образом операции горизонтального отражения. В качестве оценки качества обучения была вычислена F-метрика, значение которой составляло 0,435. Для оценки качества multi-label классификации была посчитана метрика AUC, в частности для пневмонии значение составило 0,768. Построение тепловых карт выполнялось посредством использования метода САМ, который заключается в следующем:

Рассматривается выход последнего сверточного слоя для поданного на вход сети изображения. Обозначим через f_k k-ю карту признаков этого слоя и пусть $w_{c,k}$ веса последнего слоя для k-ой карты признаков и класса c. Тепловая карта M_c вычисляется по формуле:

$$M_c = \sum_k w_{c,k} f_k \quad (6)$$

За основу магистерской диссертации были приняты исследования, описанные в статье [9]. В ней были рассмотрены и реализованы два метода классификации рентгеновских изображений. Первый из них направлен на

бинарную классификацию наличия патологии. Для этого были обучены модели сверточных нейронных сетей VGG16, VGG19, InceptionV3. Далее выход предпоследнего полносвязного слоя перед слоем классификации использовался в качестве признаков, используемых для обучения логистической регрессии. Лучшее значение метрики AUC 0,851 было получено при обучении VGG16. Второй метод был направлен на решение задачи multi-label классификации, в частности семи патологий. Для обучения были использованы также модели VGG16, VGG19, InceptionV3. В данном случае для конкретных патологий лучшие значения метрики AUC показала модель InceptionV3. В свою очередь для каждой обученной модели были построены тепловые карты с использованием коэффициентов корреляции Пирсона, вычисленных между бинарными метками класса наличия патологии и усредненным выходом активации последнего сверточного слоя для каждого канала.

Соответственно, одной из основных задач магистерской диссертации заключается в повышении значений метрик качества классификации, полученных ранее, в частности, в статье [9].

1.5 Выводы

В данной главе были изложены следующие пункты:

1. Основные понятия нейронных сетей, в частности, сверточных нейронных сетей;
2. Приведено описание моделей сверточных нейронных сетей, а также перечислены особенности их архитектур;
3. Произведен обзор существующих подходов к решению поставленной задачи, дана краткая характеристика результатов исследований.

ГЛАВА 2

ПРИМЕНЕНИЕ МЕТОДОВ КЛАССИФИКАЦИИ ПАТОЛОГИЙ ГРУДНОЙ КЛЕТКИ

2.1 Исходные данные

Исходные данные представляют собой рентгеновские изображения грудной клетки, находящиеся в системе PACS. Все изображения имеют формат DICOM. Данный формат наиболее часто используется для хранения медицинских изображений, а также их передачи между медицинскими учреждениями. Помимо самого изображения формат DICOM позволяет также хранить текстовую информацию о пациенте, например, возраст, дата снимка, наличие болезни и т.д.

Наряду с изображениями имеется база данных SQL, которая содержит 1908926 записей скрининга населения в период 2001-2014 годах. Каждая запись соответствует одному рентгеновскому изображению и включает следующую информацию о пациенте: номер карты, пол, возраст, описание радиолога относительно того, что изображено на рентгене, параметры изображения (высота, ширина, размер). Примеры записей:

10, 1, 67, Аорта развернута , уплотнена.Спондилез., 1625, 1536, 3055381.

10, 1, 69, Легкие без видимых очаговых и инфильтративных теней. Аорта развернута, уплотнена.Левый купол диафрагмы уплощен., 1625, 1536, 2997021

10, 1, 71, Эмфизема. Пневмосклероз. Аорта уплотнена, развернута. Спондилез., 1500, 1536, 2570157

10, 1, 72, Пневмосклероз. Аорта уплотнена, развернута. Спондилез., 1625, 1536, 3179511

10, 1, 73, Пневмосклероз. Аорта уплотнена, развернута. Спондилез., 1625, 1536, 3194332

11, 1, 72, Аорта развернута , уплотнена., 1625, 1536, 2989613.

Описание радиологов содержит информацию о возможных болезнях сердечно-сосудистой системы (сердце, кровеносные сосуды и т.д.), аномалий скелета (сколиоз, деформация ребер и т.д.), а также заболеваний легких. Описание написано в произвольной форме, иногда с опечатками. Также в нем может содержаться информация о ранее перенесенных болезнях или описан анамнез. Это сделано для того, чтобы сформировать представление о течении болезни пациента, возможных рецидивах или выздоровлении. Таким образом, в базе данных может содержаться несколько записей, относящихся к одному и тому же пациенту. В свою очередь, стоит отметить, что записи могут включать

информацию о предположительных диагнозах, которые требуют дополнительного подтверждения.

В диссертации выделяются шесть видов патологий легких: бронхит, эмфизема, очаговые тени, пневмония, пневмосклероз, туберкулез. Каждая из патологий характеризуется своими признаками и имеет определенный вид на рентгене. Примечательно то, что, например, пневмония имеет ярко выраженный очаг, а также в сложном случае может затенять все легкое.

Приведем краткое описание каждой патологии и ее особенностей.

Бронхит. Является заболеванием дыхательной системы, при котором в большей степени поражаются бронхи. Повреждение и воспаление бронхов может возникать как самостоятельный процесс, а также может быть осложнением на фоне хронических заболеваний легких. Бронхит характеризуется отеком бронхов, а также большим отделением слизи. В результате этого нарушается циркуляция воздуха к легким и из них. Выделяют острую и хроническую форму. Хроническая форма характеризуется наличием кашля с выделением мокроты, который продолжается более трех месяцев в году. На рентгенах могут выделены четыре вида симптомов: изменения легочного рисунка (уплотнение стенок бронхов), эмфизема в сочетании с участками лобулярного и ацинозного уплотнения, морфологические изменения бронхов (изменение контуров бронхов, их калибра), изменения функции бронхов, [10].

Эмфизема. Проявляется наличием лишнего (избыточного) не полностью выдыхаемого воздуха в легких. Вследствие этого на легкие оказывается давление, в результате которого воздухоносные каналы расширяются и утрачивают эластичность. Легкие становятся более уязвимы к бактериальной агрессии, наблюдается дополнительная нагрузка на сердечно-сосудистую систему. Различают везикулярную (или собственно эмфизему), интерстициальную и ряд других форм эмфиземы легких. Основными факторами возникновения эмфиземы считаются курение и производственное загрязнение воздуха. Зачастую встречается сочетание эмфиземы и хронического бронхита. В основном от эмфиземы страдают пожилые люди. На рентгенах эмфизема проявляется в виде следующих признаков: увеличение воздухонаполнения альвеол, обеднение сосудистого рисунка преимущественно в периферических отделах легочных полей, наличие булл.

Очаговые тени. Под очаговой тенью понимают пятно менее 1 см в диаметре. В большей степени очаги в легких встречаются при воспалительных заболеваниях, например, очаговый туберкулез. Очаговая тень характеризуется следующими признаками: локализация, распространенность, контуры, интенсивность. Выделяют следующую интенсивность очага: малоинтенсивная

тень, среднеинтенсивная тень, плотный очаг. Также тени могут образовывать группы.

Пневмосклероз. Патологический процесс, при котором происходит замещение легочной ткани соединительной и сопровождается нарушением газообмена, а также эластичности. Вследствие разрастания соединительной ткани происходит деформация бронхов. Процесс носит необратимый характер. В основном встречается у людей пожилого возраста. Наибольшее значение в развитии пневмосклероза имеет инфекция. Бактерии (стрептококки, стафилококки), некоторые грибки, возбудители паразитарных инфекций, вызывая нагноение и некроз легочной ткани, ведут к развитию бронхоэктазов и пневмоциррозу. На рентгенах отмечаются усилением и деформацией линейных элементов легочного рисунка по ходу бронхиальных разветвлений за счет уплотнения стенок бронхов.

Пневмония. Является вирусным заболеванием, при котором происходят очаговые, сегментарные или тотальные поражения легочной ткани. Пневмония чаще всего протекает в раннем или пожилом возрасте. Симптомами пневмонии являются появление сильного озноба и лихорадки, головные боли, боли при дыхании, мокрый кашель. Тяжелая форма пневмонии без проведения надлежащего лечения может привести к летальному исходу. К возникновению острых, особенно очаговых, пневмоний предрасполагают различные хронические заболевания легких и врожденные дефекты бронхов и легких. На рентгенах наличие пневмонии характеризуется понижением прозрачности участка легкого, при тотальной пневмонии возможно полное затемнение легочного поля.

Туберкулез. Инфекционное заболевание, основными симптомами которого являются: кашель с мокротой (возможно с примесью крови), лихорадка, слабость и др. Важно на начальных стадиях обнаружить туберкулез, чтобы вовремя назначить лечение. Особенностью является то, что при заражении инфекция может продолжительное время находится в инкубационном периоде, при этом человек никаких симптомов не испытывает. При излечении туберкулеза в дальнейшем возможны рецидивы болезни.

В исходной базе данных содержатся записи, одним из полей которых является описание радиолога. В соответствии с описанием, по ключевым словам производилась разметка записей относительно перечисленных ранее классов болезней.

Для бронхита ключевыми словами являлись бронхит, а также сокращение бронх. Для эмфиземы ключевые слова эмфизема, эмф-ма. Для фиброза: фиброз, фибр. Для очаговых теней: очаговые тени, очаговая тень, очаг. тени, очаг. тень, очаговоподобная тень, очаговоподобные тени. Для

пневмосклероза: пневмосклероз, пн-з, пн/скл. Для туберкулеза: туберкулез, туб-з, tbs, tbc.

Сложность разметки данных, используя ключевые слова, состояла в наличии опечаток в описании, а также описаний о перенесенных заболеваниях или анамнеза. Описания с анамнезом исключались в большей степени вручную. Также помечались описания, где диагноз ставился под вопросом, т.е. не был подтвержден. Стоит отметить, что в процессе разметки записям с неподтвержденным диагнозом приписывалась метка наличия соответствующей патологии. Это сделано для того, чтобы система классификации имитировала следующую логику: если есть подозрение, что у пациента может быть некоторое заболевание, то естественным образом нужно провести дополнительные исследования, а не утверждать, что пациент здоровый. Еще одна особенность разметки заключалась в том, что записи, которые содержали пневмонию и туберкулез под вопросом, отмечались наличием обеих патологий. Аргументом таких действий является то, что они имеют похожий вид на рентгеновских снимках.

В результате разметки число записей, соответствующее каждой болезни, указано в таблице 1.

Таблица 1 - Количество размеченных записей

Наименование болезни	Число записей
Бронхит	1193
Эмфизема	55180
Очаговые тени	7449
Пневмосклероз	103130
Пневмония	8606
Туберкулез	1314

Кроме перечисленных классов также вводится бинарный класс наличия патологий, который принимает значение 1 при наличии признака хотя бы одной из перечисленных выше болезней, 0 в противном случае. Соответственно записей с наличием патологий 121099.

Также производится разметка записей с описанием, соответствующим здоровым пациентам. Ключевыми фразами являются: Лёгочные поля чистые. Корни структурны, Легочные поля чистые. Корни структурны, Легочные поля чистые, корни структурны, Лёгочные поля чистые, корни структурны, Легочные поля без видимой патологии, Лёгкие и сердце в пределах возрастных изменений,

Очаговые и инфильтративные тени не определяются. Корни легких структурные, При компьютерной рентгенографии патологических изменений не обнаружено, При компьютерной рентгенографии патологических изменений не выявлено, При компьютерной рентгенографии органов грудной клетки патологических изменений не выявлено, Видимые лёгочные поля без очаговых и инфильтративных теней, Видимые лёгочные поля без очагово - инфильтративных теней, Легкие без видимых очаговых и инфильтративных теней. Корни легких структурные, Легкие без видимых очаговых и инфильтративных теней. Корни структурные, Легкие без видимых очаговых и инфильтративных теней, Органы грудной полости без паталогических изменений, Свежих видимых очаговых либо инфильтративных изменений не выявлено, Свежих видимых очаговых либо инфильтрат. изменений не выявлено, Свежих видимых очаговых либо инфильтрат. теней не выявлено. Сердце и д-ма без особенностей. Синусы свободны, Органы грудной полости без паталогических изменений, Очаговые и инфильтративные тени не определяются. Корни легких структурные. Границы сердца, средостения не расширены. Контур диафрагмы ровные. Синусы свободные, Очаговые и инфильтративные тени не определяются. Корни легких структурные. Границы сердца, средостения не расширены. Синусы свободные, Очаговые и инфильтративные тени не определяются. Корни легких структурные. Границы сердца, средостения не расширены. Контур диафрагмы ровные, Очаговые и инфильтративные тени не определяются. Корни легких структурные. Границы сердца, средостения не расширены, Очаговые и инфильтративные тени не определяются, Очаговые и инфильтративные тени не определяются. Корни легких структурные. На выдохе, Очаговых и инфильтративных теней в легких не выявлено. Корни структурны.

Число записей, соответствующее здоровым пациентам, составило 1221945.

Как отмечалось ранее каждой записи в базе данных соответствует одно рентгеновское изображение. Каждое изображение представлено одним каналом яркости размера 16 бит. Разрешение изображений изменяется от 520x576 пикселей до 2800x2531. Примеры исходных изображений показаны на рисунке 8.

Обучающая выборка формируется следующим образом: рассматриваются только пациенты возрастом от 18 до 90, каждому изображению с патологией ставится в соответствие изображение здорового пациента того же возраста или в пределах ± 10 от данного значения, если число пациентов с патологией больше здоровых при одинаковом возрасте, а также того же пола. Таким образом, обеспечивается сбалансированность выборки по бинарной метке наличия патологии, а также по возрасту, так как в легкие в течение жизни

подвергаются изменениям, в частности, деформации легочных полей (увеличение), снижению эластичности тканей.

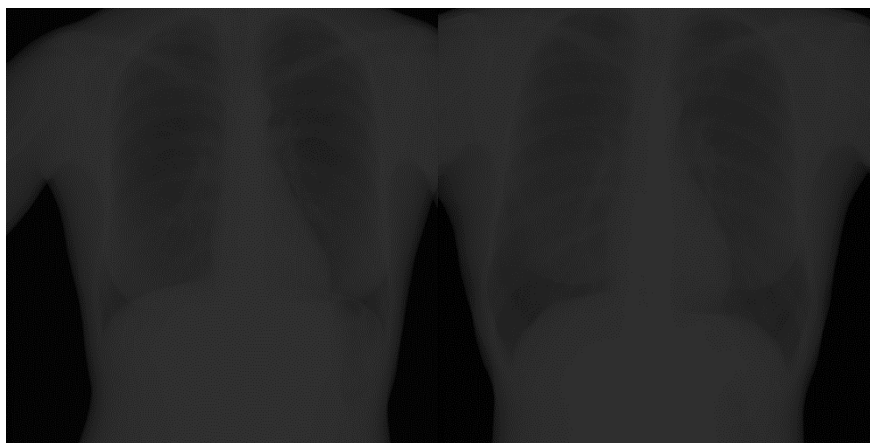


Рисунок 8 - Примеры исходных изображений

Статистика по выборке представлена в таблице 2.

Таблица 2 – Статистика выборки

	Женский пол без патологии	Женский пол с патологией	Мужской пол без патологии	Мужской пол с патологией
Общее количество	68555	68555	49815	49815
Среднее значение	60.81	71.67	59.26	67.92
Стандартное отклонение	9.20	11.60	10.69	14.09

Далее выборка разбивается на 3 группы train, validation, test в соотношении 70%, 20%, 10% соответственно. Стоит отметить, что разбиение происходит в отношении пациентов, идентифицированных номером амбулаторной карты, а не записей, так как одному пациенту может соответствовать несколько рентгеновских изображений. В итоге в train 163743 изображения (81879 здоровых, 81864 с патологией), в validation 46509 (23339 здоровых, 11736 с патологией), в test 23351 (11736 здоровых, 11579 с патологией).

2.2 Предобработка изображений

Предобработка изображений включает эквализацию, сегментацию, добавление шума, [9]. Перед этапом эквализации все изображения уменьшаются до размера 256x256.

2.2.1 Эквализация

Процесс эквализации, [11], начинается с построения гистограммы распределения яркости. Гистограмма определяется на множестве значений яркости $[0, 2^p]$, p - число бит для кодирования изображения, и отображает

количество пикселей в изображении для каждого значения яркости. Пример изображен на рисунке 9. Далее гистограмма нормируется в диапазон значений от 0 до 1, т.е. выполняется деление для каждого столбца гистограммы, значения числа пикселей, на общее число пикселей изображения. Таким образом, по гистограммам можно определить преобладание темных или светлых оттенков. Эквиализация гистограммы заключается в том, чтобы преобразовать ее к более равномерному распределению. Т.е. для улучшения визуального восприятия изображения необходимо, чтобы гистограмма содержала все значения яркости примерно в одинаковом количестве. Для этого вычисляется кумулятивная сумма для гистограммы для каждого значения яркости. Т.о., определена функция, аргументом которой является значение яркости, а значением - величина кумулятивной суммы в нем. Далее производится преобразование над изображением следующим образом: анализируем каждый пиксель и значение в нем, по ранее определенной функции ставим для этого пикселя значение кумулятивной суммы от 0 до 1. На выходе получается изображение со значениями пикселей от 0 до 1. Производим его преобразование в полутоновое домножением всех пикселей на 2^p и отбрасыванием остатка. Результатом процесса эквиализации для изображения на рисунке 9 является рисунок 10.

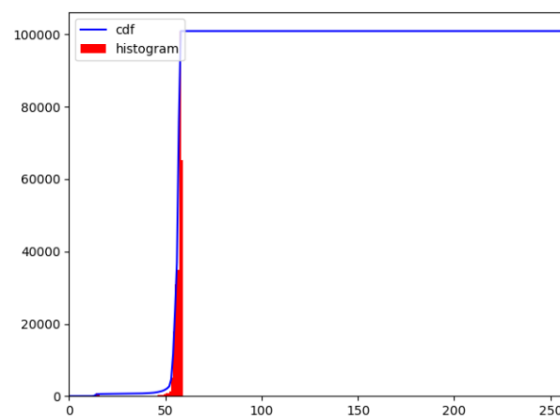
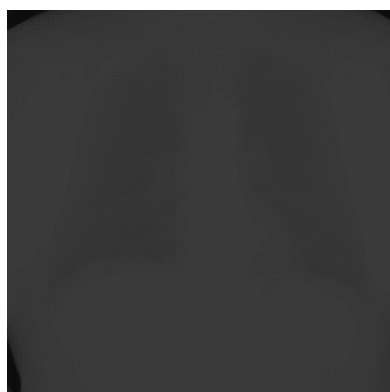


Рисунок 9 – Пример гистограммы яркости для полутонового изображения

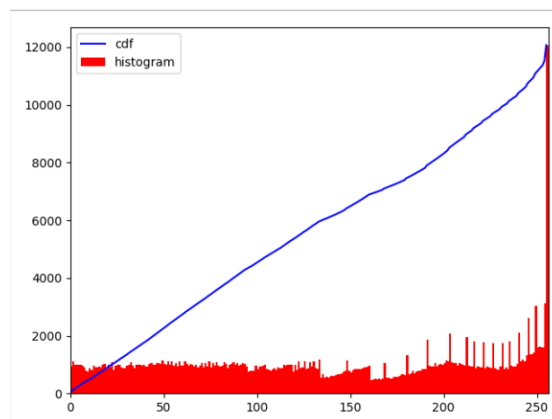
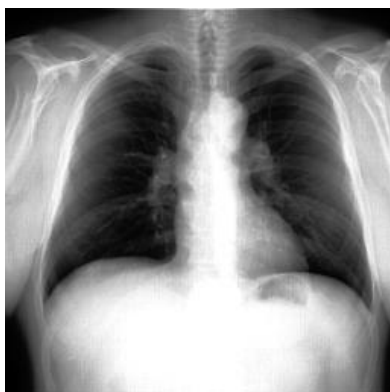


Рисунок 10 – Результат эквиализации гистограммы

2.2.1 Сегментация

Сегментация рентгеновских изображений заключается в вычислении соответствующих им масок, которые выделяют границы легких. Процесс сегментации выполнялся с использованием нейронной сети Unet [12], предобученной на публично доступной базе изображений JSRT. Архитектура Unet показана на рисунке 11.

Архитектура состоит из двух путей: сужающегося и расширяющегося. Сужающийся путь включает в себя повторное применение сверток 3×3 с последующим нелинейным преобразованием ReLU и max pooling 2×2 для понижения разрешения. Расширяющийся путь состоит из блоков, которые повышают дискретизацию карт свойств. Каждый блок состоит из свертки 2×2 , которая уменьшает количество каналов свойств, объединения с соответствующим образом обрезанной картой свойств из сужающегося пути, и двух сверток 3×3 , за которыми следует ReLU. На последнем слое используется свертка 1×1 для того, чтобы сопоставить каждый 64-компонентный вектор признаков с необходимым количеством классов.

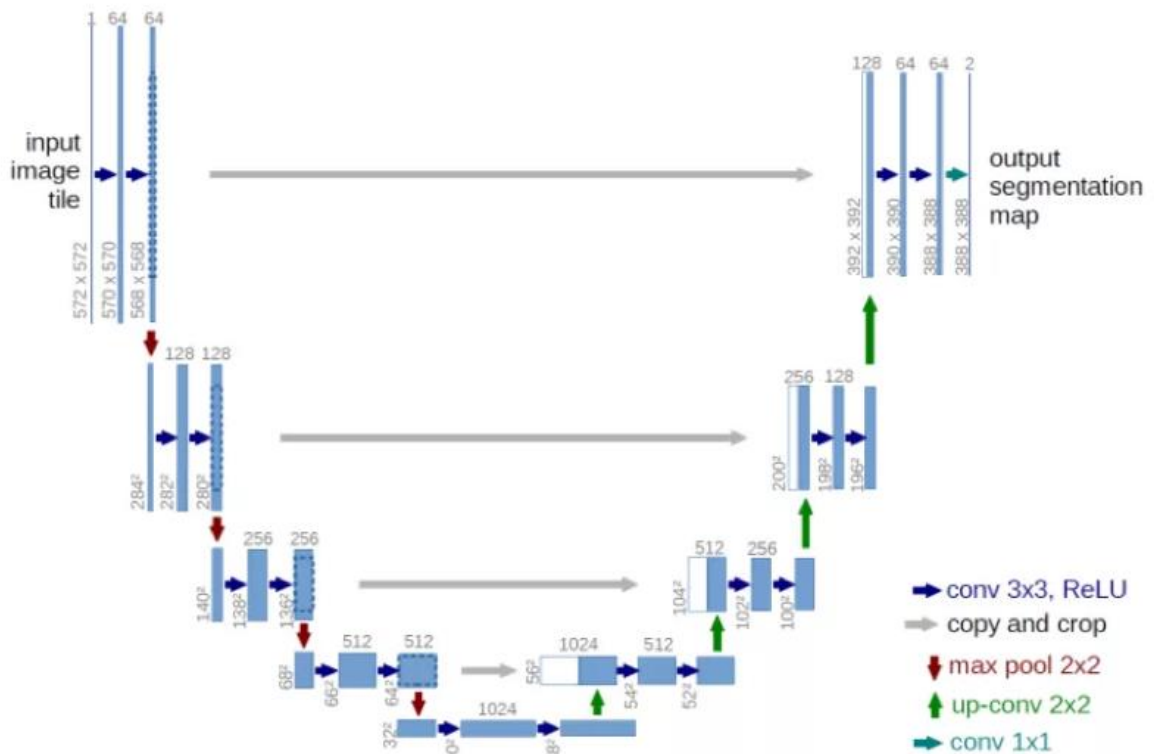


Рисунок 11 – Архитектура сети Unet, [13]

На вход предобученной сети Unet подавались рентгеновские изображения, предварительно подвергнутые процессу эквализации. Выход представлял собой маску, которая приводилась в бинарный вид путем отсечения по порогу 0,5. Затем маска проходила процесс валидации, который заключался в следующем:

1. Применение морфологических операций для вычисления и удаления областей небольшой площади, а именно меньше 2% от площади изображения. А также применение морфологических операций заполнения дыр площади меньшей 2% от площади изображения. Сделано это по причине того, что выход сети может содержать некоторые шумы (дефекты);

2. Вычисляется число связных компонент маски, состоящих из единиц. Если число компонент оказывается меньше двух, то такая маска считается невалидной, и изображение, которому она соответствует, не включается в множество для обучения;

3. Для каждой компоненты связности производится подсчет площади занимаемой ею области. Далее выделяются две компоненты с наибольшим значением площади для вычисления их пропорционального отношения, причем деление осуществляется меньшего на большее. Если соотношение оказывается меньше, чем 0.55, то, аналогично пункту 2, маска считается невалидной;

4. В том случае, когда маска удовлетворяет всем перечисленным выше условиям, остальные компоненты связности кроме двух наибольшей площади обнуляются. Соответственно, полученная маска участвует в дальнейших преобразованиях.

2.2.3 Добавление шума

Следующий этап заключается в нормализации исходного изображения по следующему алгоритму: для маски, состоящей из двух несвязных областей легких, вычисляется выпуклая оболочка, далее анализируются значения яркости пикселей, принадлежащих внутренней части выпуклой оболочки и находятся первый и 99-й перцентиль. Нормализованное изображение получается по формуле 7:

$$img = 0.2 + 0.6 * \frac{img - q_1}{q_{99} - q_1}, \quad (7)$$

q_{99} – 99 – й перцентиль, q_1 – 1 – й перцентиль

Далее нормализованное изображение умножается на бинарную маску и обрезается по границам легких с константными отступами. После умножения фон приобретает черный цвет, что соответствует сильному перепаду яркости на границах легких. Для того чтобы этого избежать, фон заполняется шумом, имеющим нормальное распределение, с математическим ожиданием и стандартным отклонением, соответствующим математическому ожиданию и стандартному отклонению распределения пикселей изображения, для которых значение маски равно 1, т.е. пикселей, принадлежащих зоне легких. На рисунке 12 приведен процесс предобработки для изображения.

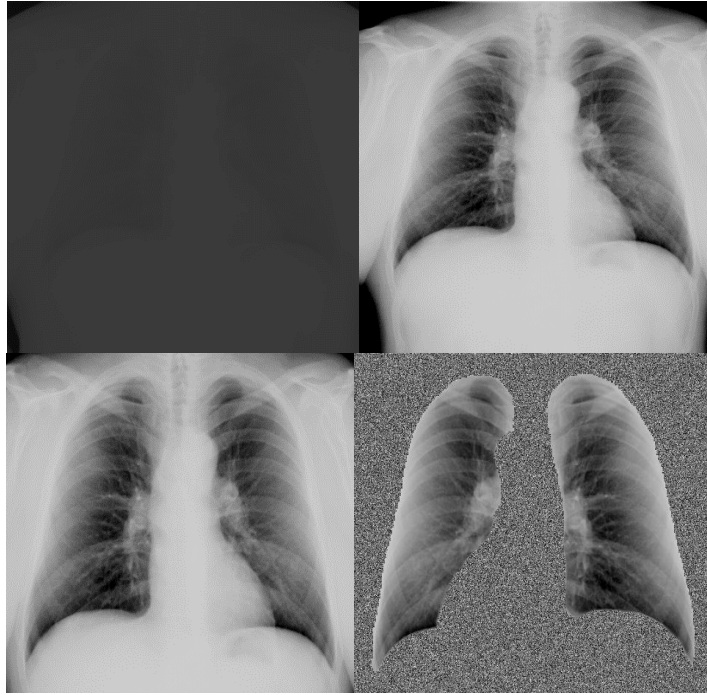


Рисунок 12 - Процесс предобработки изображения (1 - исходное изображение, 2 - нормализованное изображение, 3 - обрезанное изображение, 4 – изображение с шумом)

2.4 Эксперименты

2.4.1 Multi-label классификация

Для решения задачи классификации были обучены следующие архитектуры сверточных нейронных сетей: VGG16, DenseNet121, InceptionV3, GoogleNet, ResNet50, Xception. Тип классификации был выбран multi-label, при котором для одного объекта могут предсказываться одновременно несколько меток различных классов. Число классов для предсказания семь: один класс указывает на наличие (отсутствие) патологии, остальные определяют каждый из классов болезней, перечисленных в главе 2.1. Для multi-label классификации на последнем полносвязном слое каждой сети применялась сигмоидная функция активации. Выбор такой функции обусловлен тем, что ее применение к выходам слоя осуществляется независимо. Также в качестве функции потерь использовалась бинарная перекрестная энтропия:

$$H_p(q) = -\frac{1}{N} \sum_{i=1}^N y_i \log(p(y_i)) + (1 - y_i) \log(1 - p(y_i)) \quad (8)$$

Помимо перечисленных выше сетей проводились эксперименты по обучению VGG16 с уменьшенным числом параметров, т.е. она имеет следующую последовательность слоев: 2 сверточных слоя 3x3 с числом каналов 56 (в исходной 64), слой max пулинга 2x2, 2 сверточных слоя 3x3 с числом каналов 112 (в исходной 128), слой max пулинга 2x2, 3 сверточных слоя 3x3 с

числом каналов 224 (в исходной 256), слой max пулинга 2x2, 3 сверточных слоя 3x3 с числом каналов 448 (в исходной 512), слой max пулинга 2x2, 3 сверточных слоя 3x3 с числом каналов 448 (в исходной 512), слой max пулинга 2x2, 2 полносвязных слоя с числом каналов 3136 (в исходной 4096), 1 полносвязный слой с числом каналов 7. Число параметров уменьшалось, так как архитектура VGG16 разрабатывалась для предсказания 1000 классов изображений imagenet. Данная задача имеет всего 7. Обозначим данную сеть VGG_v2.

Для обучения было выбрано 30 эпох и размер батча 32. В качестве инструмента обучения был выбран пакет keras с tensorflow. Для тренировочной выборки производилась аугментация с целью уменьшить переобучение. Она состояла в следующем: диапазон вращения 10, диапазон сдвига по ширине 0.1, диапазон сдвига по высоте 0.1, диапазон увеличения 0.1. Также значения пикселей всех изображений были приведены в диапазон от -0.5 до 0.5. В качестве алгоритма оптимизации была выбрана оптимизация методом Адама со скоростью обучения 10^{-5} . В процессе обучения если в течение трех эпох значение функции потерь на валидационном множестве не улучшает ее минимального значения, величина скорости обучения уменьшается в 10 раз. Обучение останавливается, если в течение восьми эпох значение минимума функции потерь на валидационном множестве не улучшается.

2.4.2 ROC-кривая, AUC

ROC-кривая для бинарного классификатора отображает соотношение долей истинно положительных объектов и ложноположительных. Истинно положительными объектами являются те, для которых метка класса 1 была предсказана верно. Ложноположительными объектами являются те, для которых была предсказана метка 1, однако истинное значение класса отличное.

Истинно положительная пропорция также называется чувствительностью (англ. sensitivity) и вычисляется по формуле:

$$TPR(a, X^m) = \frac{\sum_{i=1}^m [a(x_i) = +1][y_i = +1]}{\sum_{i=1}^m [y_i = +1]}, \quad (9)$$

$a(x)$ – классификатор, X^m – выборка, состоящая из m объектов,
 y_i – истинные метки класса для объектов выборки

Ложноположительная пропорция также называется специфичностью (англ. specificity) и вычисляется по формуле:

$$FPR(a, X^m) = \frac{\sum_{i=1}^m [a(x_i) = +1][y_i = -1]}{\sum_{i=1}^m [y_i = -1]} \quad (10)$$

Соответственно, график представляет собой набор точек (FPR, TPR) при различном выборе порога и показывает соотношение между увеличением чувствительности и уменьшением специфичности. Метрика AUC вычисляется как величина площади под ROC-кривой. При значении этой величины 1 достигается 100% точность классификации на выборке. Таким образом, при использовании данной метрики необходимо стремиться к ее увеличению, однако при AUC=1 существует большая вероятность того, что модель переобучена.

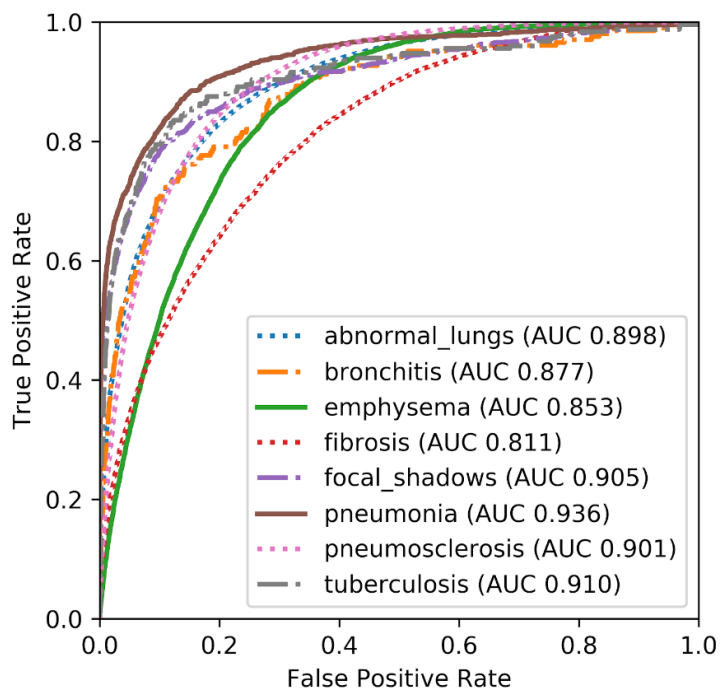
2.4.3 Результаты

Значения AUC для каждой из моделей пункта 2.4.1 описаны в таблице 3. На рисунках 13 - 14 приведены графики ROC-кривых для каждой из сетей VGG16_v2 и Xception, которые показали наилучшие результаты.

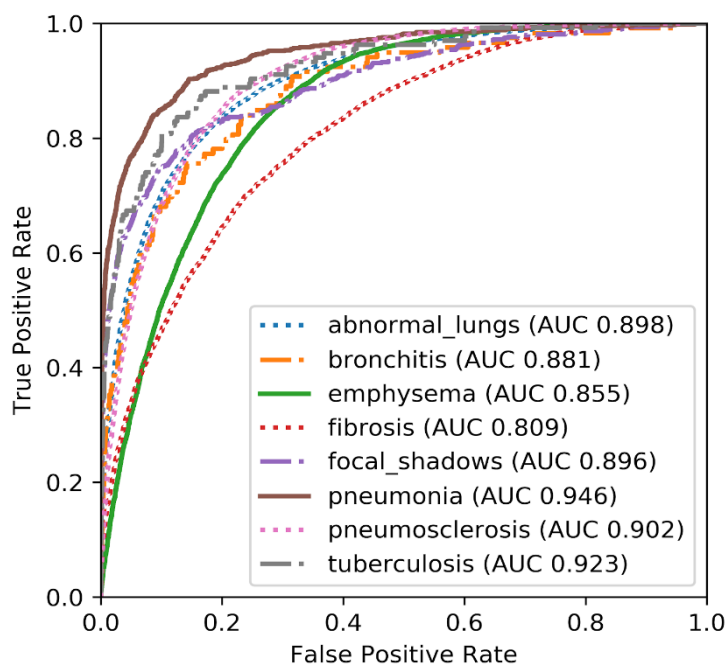
Таблица 3 - Значения AUC

	VGG16	VGG16_v2	DenseNet121	InceptionV3	GoogleNet	ResNet50	Xception
Наличие патологии	Val: 0.892 Test: 0.893	Val: 0.898 Test: 0.898	Val: 0.878 Test: 0.882	Val: 0.869 Test: 0.872	Val: 0.892 Test: 0.891	Val: 0.891 Test: 0.891	Val: 0.901 Test: 0.9
Бронхит	Val: 0.872 Test: 0.898	Val: 0.877 Test: 0.881	Val: 0.852 Test: 0.885	Val: 0.843 Test: 0.874	Val: 0.863 Test: 0.902	Val: 0.833 Test: 0.831	Val: 0.882 Test: 0.896
Эмфизема	Val: 0.852 Test: 0.85	Val: 0.853 Test: 0.855	Val: 0.847 Test: 0.846	Val: 0.836 Test: 0.834	Val: 0.839 Test: 0.841	Val: 0.837 Test: 0.839	Val: 0.842 Test: 0.841
Очаговые тени	Val: 0.896 Test: 0.886	Val: 0.905 Test: 0.896	Val: 0.877 Test: 0.862	Val: 0.885 Test: 0.869	Val: 0.894 Test: 0.891	Val: 0.898 Test: 0.891	Val: 0.913 Test: 0.907
Пневмония	Val: 0.931 Test: 0.941	Val: 0.936 Test: 0.946	Val: 0.905 Test: 0.923	Val: 0.898 Test: 0.909	Val: 0.925 Test: 0.93	Val: 0.933 Test: 0.943	Val: 0.938 Test: 0.947
Пневмо-склероз	Val: 0.902 Test: 0.902	Val: 0.901 Test: 0.902	Val: 0.894 Test: 0.894	Val: 0.887 Test: 0.886	Val: 0.897 Test: 0.896	Val: 0.891 Test: 0.892	Val: 0.903 Test: 0.902
Туберкулез	Val: 0.888 Test: 0.894	Val: 0.91 Test: 0.923	Val: 0.838 Test: 0.841	Val: 0.877 Test: 0.859	Val: 0.875 Test: 0.881	Val: 0.897 Test: 0.908	Val: 0.922 Test: 0.937

Как видно из таблицы лучшие значения AUC имеет Xception. Стоит также отметить, что значения AUC вычислялись для тех весов сети, на которых достигается минимум функции ошибки на валидационном множестве.

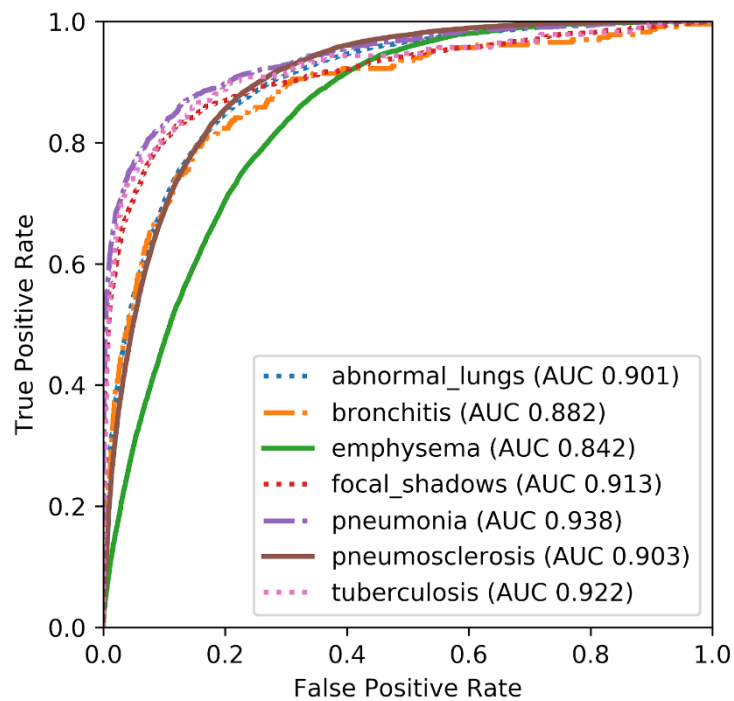


a

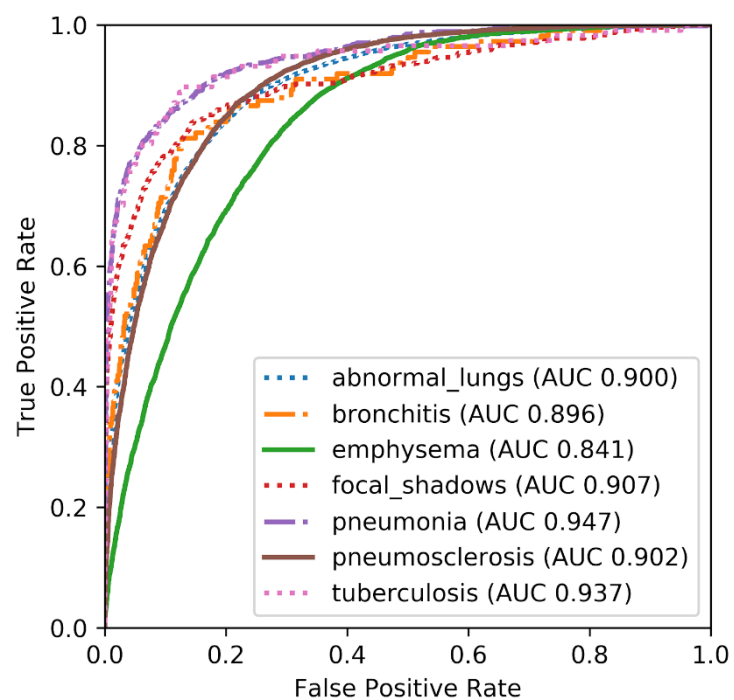


б

Рисунок 13 - ROC-кривые для VGG16_v2(а - для валидации, б - для теста)



a



б

Рисунок 14 - ROC-кривые для Xception (а - для валидации, б - для теста)

В таблице 4 приведены значения порогов отсечения для VGG16_v2 и Xception, на которых достигается максимальное качество классификации.

Таблица 4 - Оптимальные значения порогов отсечения

	VGG16_v2 Val	VGG16_v2 Test	Xception Val	Xception Test
Наличие патологии	0.4149	0.4062	0.7301	0.6488
Бронхит	0.0036	0.0036	0.0028	0.0066
Эмфизема	0.1076	0.1080	0.4397	0.4280
Очаговые тени	0.0349	0.0233	0.0423	0.033
Пневмония	0.0175	0.0158	0.0216	0.0201
Пневмосклероз	0.2493	0.2298	0.61	0.5722
Туберкулез	0.0047	0.0043	0.0025	0.0061

Стоит отметить, что пороговые значения для разных классов сильно отличаются, а именно для некоторых из них в 10 раз. Это объясняется тем, что количество примеров в классах значительно различается.

2.4.2 Взвешенная функция потерь

Вследствие того, что для разных классов значения порогов отсечения сильно отличаются в порядке, была применена техника обучения для несбалансированной выборки по классам, а именно использование взвешенной функции потерь:

$$H_p(q) = -\frac{1}{N} \sum_{i=1}^N w * y_i \log(p(y_i)) + (1 - w)(1 - y_i) \log(1 - p(y_i)) \quad (11)$$

Выбор значения веса w для каждого класса производился двумя способами:

$$1. \quad w = \frac{N}{N+P}, \quad (12)$$

N – число негативных примеров, P – число позитивных

$$2. \quad w = \frac{n_{samples}}{n_{clones}}, \quad (13)$$

$n_{samples}$ – число всех примеров,

n_{clones} – число примеров класса, несущие признак

При обучении модели VGG16 с такой функцией потерь были получены противоречивые результаты, а именно для случаев с предсказанием туберкулеза 96% одновременно предсказание наличия патологии было 2%. В данном случае наблюдался перевес в сторону классов с небольшим количеством примеров. Также при обучении график изменений функции потерь на валидационном множестве представлял собой ломаную, см. рис. 15, т.е. не наблюдался обычным образом спад значения функции потерь, а затем увеличение, свойственное

моменту переобучения. Это явление можно объяснить слишком сильной разбалансировкой выборки по классам.

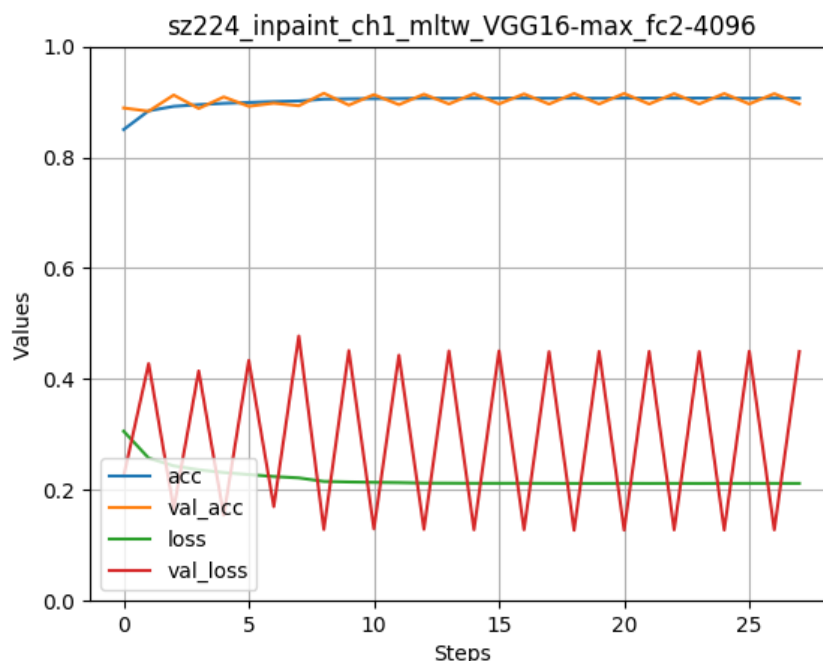


Рисунок 15 – График обучения VGG16 со взвешенной функцией потерь

2.5 Выводы

В данной главе было рассмотрено:

1. Исходные данные, их состав, формирование выборки, статистика по классам;
2. Предобработка изображений, особенности реализации;
3. Описаны и проанализированы эксперименты с обучением моделей сверточных нейронных сетей;
4. Произведено сравнение моделей с использованием метрики AUC и ROC-кривой;
5. Проведены эксперименты со взвешенной функцией потерь.

ГЛАВА 3

ТЕПЛОВЫЕ КАРТЫ

Тепловые карты применяются для локализации участков предсказанной болезни на рентгеновских снимках. В данной главе перечислены методы их построения и проведенные эксперименты.

3.1 Алгоритмы построения тепловых карт

3.1.1 Коэффициент корреляции Пирсона

Для построения тепловой карты используется модель сверточной нейронной сети, предварительно обученная для решения задачи классификации. Для каждого примера из тренировочной выборки вычисляется выход на последнем сверточном слое после применения функции активации. Затем для каждого канала полученного выхода вычисляется среднее значение по матрице. Таким образом, каждый элемент тренировочной выборки описывается вектором размерности, равной числу каналов последнего сверточного слоя. Далее для каждого канала вычисляется коэффициент корреляции Пирсона по формуле (14), где первый вектор в качестве i -й компоненты имеет значение соответствующей каналу компоненты вектора, описывающего i -й объект выборки, а второй вектор состоит из меток класса тренировочной выборки.

$$r = \frac{\sum (x - m_x)(y - m_y)}{\sqrt{\sum (x - m_x)^2 (y - m_y)^2}}, \quad (14)$$

m_x – среднее значение вектора x , m_y – среднее значение вектора y

Каждый полученный коэффициент возводится в квадрат и умножается на сигнум от изначального значения. Формула расчета коэффициента корреляции Пирсона устанавливает тесноту связи признаков в том случае, если связь между ними имеет линейный характер. Его значение изменяется в пределах от -1 до 1. Соответственно, 1 означает полную линейную положительную взаимосвязь, -1 означает полную линейную обратную взаимосвязь, 0 – отсутствие линейной корреляции.

Тепловая карта для изображения вычисляется путем поканального умножения выхода последнего сверточного слоя после активации на коэффициенты корреляции Пирсона и суммирования. Далее результат нормируется путем деления на свое максимальное значение и умножается на предсказанную вероятность наличия болезни.

3.1.2 Логистическая регрессия

Логистическая регрессия – метод машинного обучения, который используется для предсказания вероятности того, что объект принадлежит определенному классу. Простейший случай представляет собой разделение объектов на два класса. Идея метода заключается в том, чтобы разделить все пространство признаков, описывающих объекты тренировочной выборки, на два полупространства линейной границей (плоскостью, в двумерном случае – прямой). Соответственно, уравнение n -мерной плоскости имеет вид:

$$t = w_0 + \sum_{i=1}^n w_i x_i \quad (15)$$

Для объекта тренировочной выборки, описываемого n -мерным вектором, при подстановке в уравнение плоскости (15) возможны три варианта:

1. Значение находится в пределах $(0, \infty)$. В этом случае считается, что объект принадлежит классу «+» с вероятностью P_+ в пределах $(0,5; 1]$. Чем выше полученное значение, тем дальше от разделяющей плоскости расположен объект, соответственно, большая вероятность того, что он принадлежит конкретному классу;
2. Значение находится в пределах $(-\infty, 0)$. В этом случае считается, что объект принадлежит классу «-», а вероятность P_+ находится в пределах $[0; 0,5)$;
3. Значение равно 0. В этом случае модель не может определить, какому классу принадлежит объект, и $P_+ = 0,5$.

Так как величина, полученная при подстановке вектора признаков в уравнение плоскости, является действительной, а вероятность изменяется от -1 до 1, то необходимо произвести дополнительное преобразование по формуле (16):

$$P_+ = \sigma(t) = \frac{1}{1 + e^{-t}} \quad (16)$$

График функции изображен на рисунке 16. Модель логистической регрессии имеет вид (17). В качестве меток класса используется 0 и 1.

$$L_{log}(X, y, w) = \sum_{i=1}^N \left(-y_i \log_e \sigma(w^T x_i) - (1 - y_i) \log_e (1 - \sigma(w^T x_i)) \right), \quad (17)$$

где $\sigma(w^T x_i) = \frac{1}{1 + e^{-w^T x_i}}$, X - множество объектов выборки, w - вектор весов, y - вектор меток класса

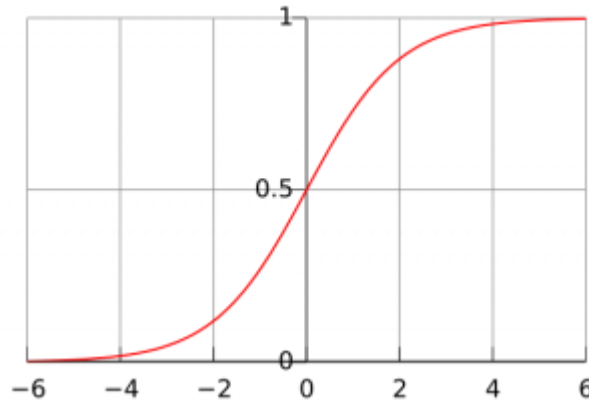


Рисунок 16 – График логистической функции

Построение тепловой карты осуществляется следующим образом:

Для каждого экземпляра тренировочной выборки вычисляется вектор таким же способом, как и в пункте 3.1.1. Однако вместо коэффициентов корреляции Пирсона в данном случае обучается модель логистической регрессии. Меткой класса является признак наличия болезни 0 или 1. Вычисление тепловой карты происходит путем умножения выхода последнего сверточного слоя после активации на коэффициенты логистической регрессии. Затем, как и в пункте 3.1.1, результат нормируется и умножается на предсказанную вероятность болезни.

3.1.3 Grad-CAM

Grad-CAM, [14], использует информацию о градиенте, которая поступает на последний сверточный слой, для того чтобы назначить степень важности каждому нейрону в зависимости от решаемой задачи. Нейроны на этом слое пытаются найти семантическую информацию о классе изображения.

Вначале вычисляются градиенты оценки для каждого класса s , y^s , относительно карт активаций признаков последнего сверточного слоя A^k . Затем по высоте и ширине применяется глобальный усредненный пулинг и вычисляются веса важности для нейронов a_k^s , (18).

$$a_k^s = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^s}{\partial A_{ij}^k} \quad (18)$$

Веса a_k^s представляют собой частичную линейризацию глубокой нейронной сети в направлении от A и отражают важность карты признаков k для класса s . Далее вычисляется линейная взвешенная комбинация карт активации на выходе последнего сверточного слоя и коэффициентов a_k^s . К результату применяется функция ReLU, (19).

$$L^c = \text{ReLU}\left(\sum_k a_k^c A^k\right) \quad (19)$$

ReLU применяется по причине того, что нас интересуют только те признаки, которые положительно влияют на класс, т.е. пиксели, интенсивность которых увеличивают предсказание. Отрицательные признаки, вероятно, не относятся к предсказываемому классу. Однако, если их не занулять, они могут привести к тому, что тепловая карта будет выделять большую площадь объекта предсказываемого класса. Это приводит к плохой локализации.

Данный алгоритм можно обобщать на выход любого сверточного слоя.

3.1.4 Альтернативный подход к построению тепловых карт

Идея альтернативного подхода заключается в том, чтобы изменить архитектуру сверточной нейронной сети таким образом, чтобы выход сверточного слоя представлял собой тепловую карту без применения дополнительных преобразований. Соответственно, архитектура сверточной нейронной сети изменяется следующим образом: после последнего сверточного слоя модели VGG16 добавляются сверточные слои с фильтрами 1x1 и функцией активации ReLU, на последнем из них используется сигмоидная функция активации и число выходов составляет число предсказываемых классов. В конце сети используется глобальный максимальный пулинг. Применение сверток 1x1 объясняется тем, что они анализируют межканальную корреляцию и производят суммирование значений пикселей по каналам с некоторыми весами. Т. о., эти преобразования подобны тем, которые производятся при вычислении тепловых карт алгоритмами, описанными в предыдущих пунктах, только с выходом последнего сверточного слоя.

3.2 HSV

Цветовая модель HSV, [15], состоит из трех основных компонент:

1. Параметр H определяет цветовой тон, значение которого изменяется от 0 до 1 либо от 0° до 360°. Минимальное значение соответствует красному цвету. Шкала оттенков представлена на рисунке 17;
2. Параметр S определяет насыщенность цвета и изменяется от 0 до 1. Наименьшее значение параметра соответствует минимальной насыщенности и приближению к более серому оттенку;
3. Параметр V определяет яркость цвета и изменяется от 0 до 1. Наименьшее значение параметра приближает цвет к черному оттенку.



Рисунок 17 – Шкала оттенков Hue, [15]

Перевод модели HSV в RGB осуществляется по следующим формулам:

$$H_i = \left\lfloor \frac{H}{60} \right\rfloor \bmod 6, H \in [0, 360] \quad (20)$$

$$V_{min} = (1 - S) * V \quad (21)$$

$$a = (V - V_{min}) * \frac{H \bmod 60}{60} \quad (22)$$

$$V_{inc} = V_{min} + a \quad (23)$$

$$V_{dec} = V - a \quad (24)$$

Таблица 5 - Соответствие HSV и RGB

H_i	R	G	B
0	V	V_{inc}	V_{min}
1	V_{dec}	V	V_{min}
2	V_{min}	V	V_{inc}
3	V_{min}	V_{dec}	V
4	V_{inc}	V_{min}	V
5	V	V_{min}	V_{dec}

Исходные рентгеновские изображения являются полутоновыми. Соответственно, тепловые карты являются также полутоновыми. Для преобразования в цветное изображение использовалась модель HSV. Алгоритм заключается в следующем:

1. В качестве параметра H принимается полученная тепловая карта. Перед этим все её значения преобразуются в диапазон от 0 до 1. Далее карта преобразуется по формуле:

$$heatmap = 0.7 * (1 - heatmap) \quad (25)$$

В данной формуле разность применяется для того, чтобы максимальному значению на тепловой карте соответствовал красный цвет. Умножение на 0,7 необходимо для того, чтобы уменьшить диапазон изменения тона до синего;

2. В качестве параметра S принимается бинарная маска легких (значения 0 и 1), которая умножается на 0,5. Таким образом, фон остается серым;

3. В качестве параметра V принимается нормализованное исходное изображение.

3.3 Предобработка изображений

Предобработка изображений состоит из тех же этапов, что и в главе 2.2, однако без операции добавления шума. Это сделано по причине того, что нейронная сеть на последнем сверточном слое имеет свои максимальные значения, которые реагируют на шум, т.е. происходит локализация, в частности, фона. Соответственно, в дальнейшем использовались изображения без заполнения фона, а также были проведены эксперименты с его заполнением с использованием алгоритма Inpainting. Заполнение фона можно аргументировать тем, что нейронная сеть не имеет необходимости учить признаки, относящиеся к другим органам, за исключением легких.

3.3.1 Алгоритм Inpainting

Алгоритм inpainting применяется для восстановления изображений, [16]. Исходными данными для алгоритма являются изображение с дефектами и маска, которая показывает, какие области необходимо восстановить. В статье [16] описывается алгоритм, основанный на применении уравнений Навье-Стокса для гидродинамики, (26).

$$v_t + v * \nabla v = -\nabla p + \nu \Delta v; \nabla * v = 0, \quad (26)$$

v_t - производная v по t , v - вектор скорости, p - давление, ν - коэффициент кинематической вязкости, ∇ - оператор Набла, Δ - векторный оператор Лапласа.

Также вводятся такие параметры, как функция потока ψ , которая удовлетворяет формуле (27), и ω - завихренность, (28).

$$\nabla^\perp \psi = v \quad (27)$$

$$\omega = \nabla \times v \quad (28)$$

Для применения выше перечисленных уравнений к изображениям проводится следующая аналогия терминов:

1. функция потока ψ - интенсивность изображения I ;
2. скорость потока $v = \nabla^\perp \psi$ - направление изофоты $\nabla^\perp I$;
3. завихренность ω - сглаживание $\omega = \Delta I$;
4. коэффициент кинематической вязкости ν - анизотропная диффузия ν .

Наряду с уравнениями определяются граничные условия. Пусть Ω - область, которую необходимо восстановить, а $\partial\Omega$ - ее граница. Полагая, что интенсивность изображения на границе I_0 , а также, что мы знаем ΔI_0 , получаем условие (29).

$$I|_{\partial\Omega} = I_0, \quad (29)$$

Уравнения Навье-Стокса гарантирует непрерывность функции интенсивности изображения I и ее направления изофоты через границу восстанавливаемой области.

Таким образом, алгоритм заключается в перемещении вдоль границы от известных значений пикселей к неизвестным. Он продолжает изофоты (линии, которые соединяют точки с одинаковой интенсивностью) и сопоставляет векторы градиента на границе области. Далее значение пикселя заполняется цветом так, чтобы уменьшить значение дисперсии.

Для рентгеновских изображений применялись следующие манипуляции:

1. Сглаживание изображения гауссовым фильтром;
2. Изображение умножалось на маску. Предварительно к маске применялась морфологическая операция эрозия;
3. Выполнялся алгоритм inpaint.
4. Исходное изображение умножалось на маску после операции эрозии и прибавлялось к полученному изображению из предыдущего пункта.

Пример предобработки с использованием inpaint показаны на рисунке 18.

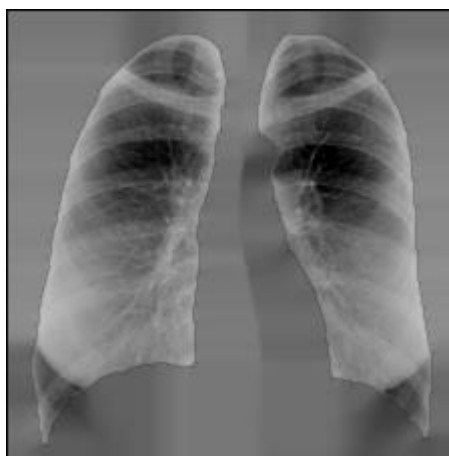


Рисунок 18 - Изображение после алгоритма inpainting

3.4 Эксперименты

Изначально эксперименты проводились на моделях, обученных в главе 2. При построении тепловых карт для изображений с зашумленным фоном были замечены дефекты. Например, для здоровых людей тепловая карта показывала наличие болезни на границах легких. Объяснением этого может служить то, что распределение шума совпадает с распределением пикселей в выделенных маской участках легких. Т.е. сеть в данном случае извлекала признаки из фона.

Также проблемой было то, что, например, сеть VGG16 предсказывала наличие болезни у пациента в то время, как тепловая карта не показывала очагов патологии. VGG16 имеет слой Flatten перед полносвязными слоями блока классификации, что приводит к затуханию признаков. Таким образом, были проведены эксперименты с заменой данного слоя на различные виды пулингов. Например, GlobalMaxPooling лучше подходит к построению тепловой карты, так как оставляет признаки, влияющие на предсказание, не изменяя их локальности. Также были проведены эксперименты с GlobalAveragePooling.

Так как изображения с зашумленным фоном привели к неудовлетворительным результатам, то последующие исследования проводились с исходными нормализованными изображениями. Примеры тепловых карт изображены на рисунке 19. Однако точность предсказания для таких изображений оказалась ниже в сравнение с моделями, обученными на зашумленных изображениях.

Далее были проведены эксперименты с другими способами заполнения фона, например, алгоритм inpainting из пункта 3.3.1. Примеры тепловых карт представлены на рисунках 20-22.

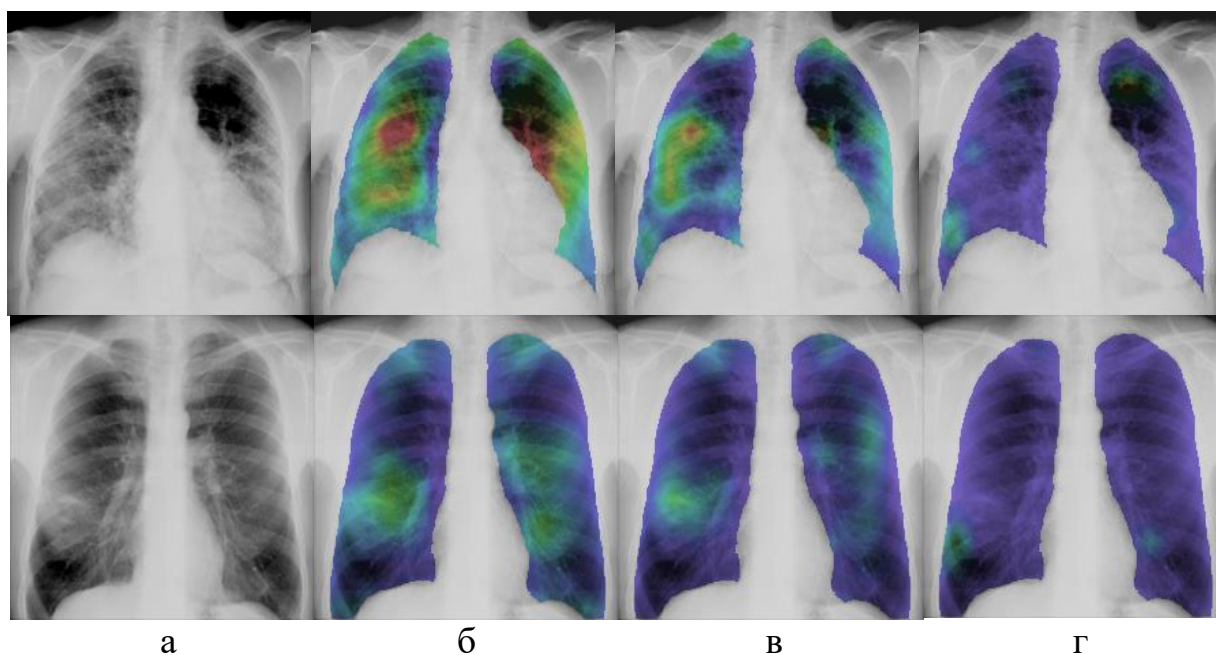
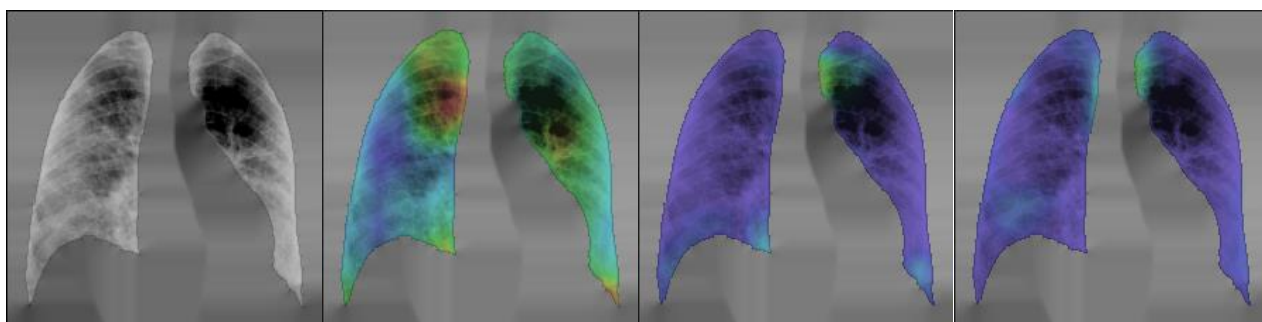


Рисунок 19 - Тепловые карты для нормализованных изображений, VGG16 с GlobalMaxPooling а) исходное изображение, б) Grad-CAM, в) логистическая регрессия, г) корреляция Пирсона (очаговые тени, пневмония)



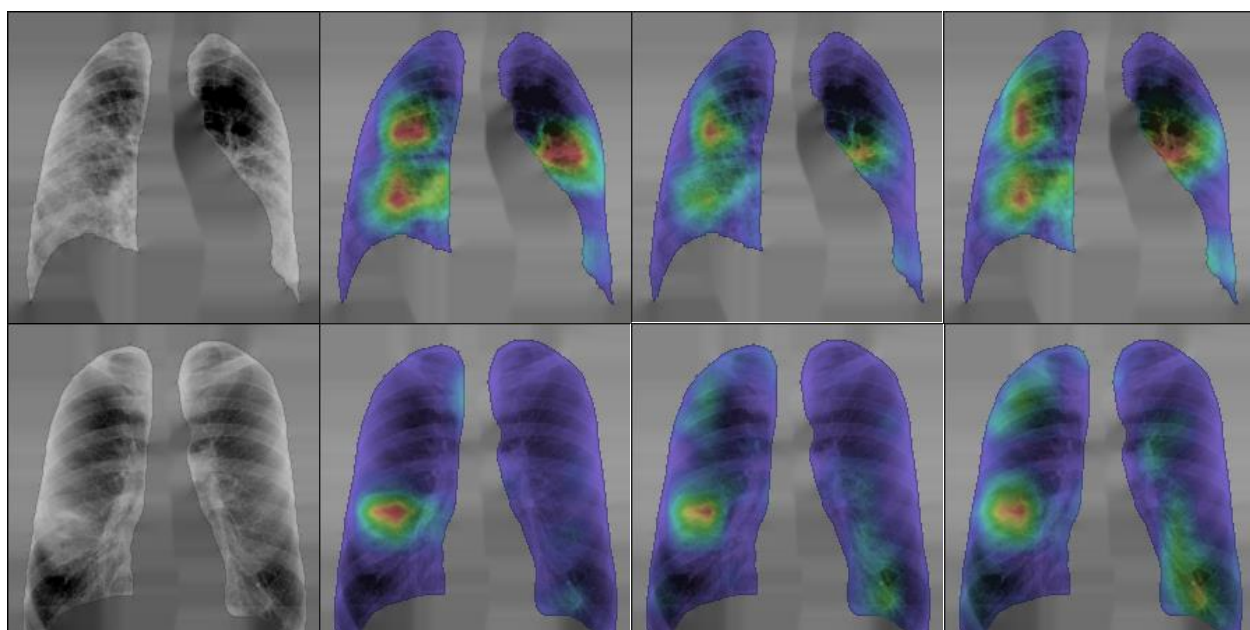
а

б

в

г

Рисунок 20 - Тепловые карты для изображений с inpaint, VGG16 с Flatten а) исходное изображение, б) Grad-CAM, в) логистическая регрессия, г) корреляция Пирсона (очаговые тени)



а

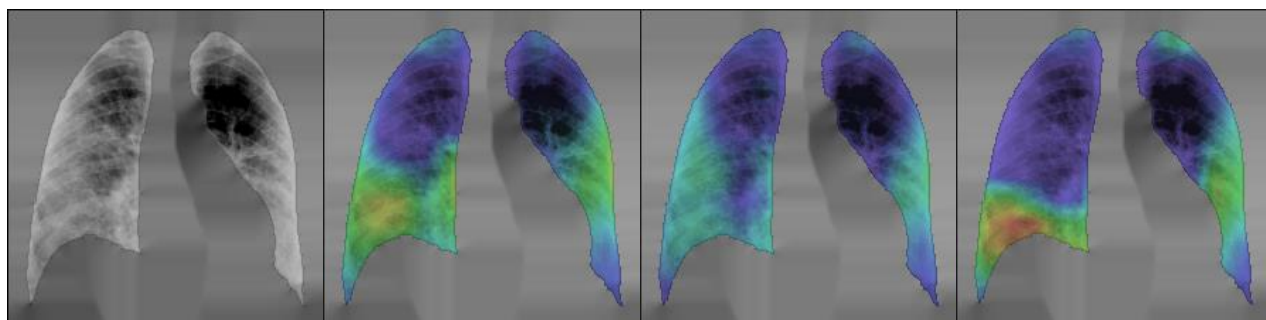
б

в

г

Рисунок 21 - Тепловые карты для изображений с inpaint, VGG16 с GlobalMaxPooling а) исходное изображение, б) Grad-CAM, в) логистическая регрессия, г) корреляция Пирсона (очаговые тени, пневмония)

Можно отметить, что слои Flatten и GlobalAveragePooling показывают некорректные тепловые карты для пациентов с патологией, так как в большей степени выделяют участки, находящиеся на границах легких.



а

б

в

г

Рисунок 22 - Тепловые карты для изображений с inpaint, VGG16 с GlobalAveragePooling а) исходное изображение, б) Grad-CAM, в) логистическая регрессия, г) корреляция Пирсона (очаговые тени)

В таблице 6 приведены значения метрики AUC для обученной VGG16 с различными видами предобработки изображений.

Таблица 6 - Значения AUC для VGG16

	GlobalMaxPooling, нормализованные изображения	Flatten, inpainting	GlobalMaxPooling, inpainting	GlobalAveragePooling, inpainting
Наличие патологии	Val: 0.881 Test: 0.88	Val: 0.871 Test: 0.872	Val: 0.882 Test: 0.883	Val: 0.871 Test: 0.871
Бронхит	Val: 0.859 Test: 0.887	Val: 0.85 Test: 0.874	Val: 0.868 Test: 0.896	Val: 0.843 Test: 0.869
Эмфизема	Val: 0.836 Test: 0.838	Val: 0.834 Test: 0.833	Val: 0.839 Test: 0.837	Val: 0.829 Test: 0.83
Очаговые тени	Val: 0.876 Test: 0.883	Val: 0.881 Test: 0.885	Val: 0.882 Test: 0.893	Val: 0.871 Test: 0.879
Пневмония	Val: 0.874 Test: 0.881	Val: 0.859 Test: 0.868	Val: 0.912 Test: 0.921	Val: 0.824 Test: 0.819
Пневмо- склероз	Val: 0.897 Test: 0.896	Val: 0.889 Test: 0.889	Val: 0.893 Test: 0.893	Val: 0.892 Test: 0.891
Туберкулез	Val: 0.833 Test: 0.814	Val: 0.836 Test: 0.824	Val: 0.846 Test: 0.846	Val: 0.805 Test: 0.765

Тепловые карты, получаемые с использованием методов корреляции Пирсона и логистической регрессии, имеют явно выделенные очаги на изображениях с предобработкой inpainting в сравнение с нормализованными изображениями. Также значение величины AUC выше для предобработки inpainting и применения GlobalMaxPooling.

Дальнейшие эксперименты были направлены на бинарную классификацию наличия патологии, а также с использованием максимального пулинга, так как он дает наилучшие результаты в локализации тепловых карт. Также для более усовершенствованной локализации очагов проводились эксперименты с подаваемыми на вход сети изображениями большего размера. Например, для VGG19 размеры изображения выбирались пропорционально 32. Наряду с этим, для VGG19 изменялось число выходов полносвязных слоев на 512, равное числу

выходов последнего сверточного слоя. Примеры построенных тепловых карт можно наблюдать на рисунках 23-24.

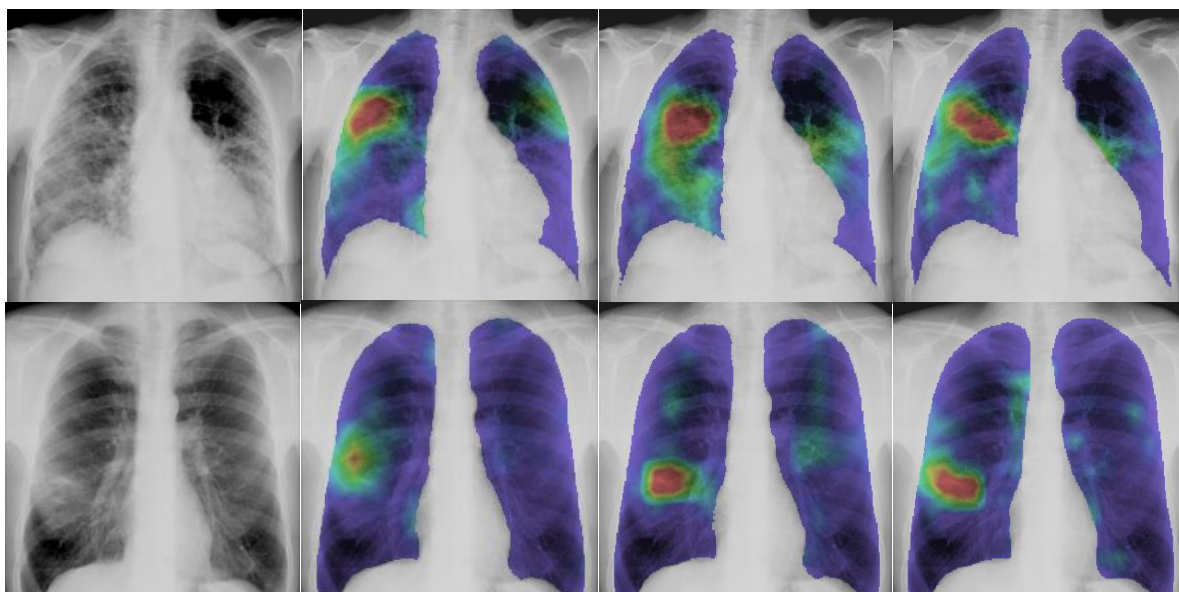


Рисунок 23 – Тепловые карты, построенные методом Grad-CAM, на модели VGG19, слева направо: исходное изображение, 224x224, 288x288, 320x320 (очаговые тени, пневмония)

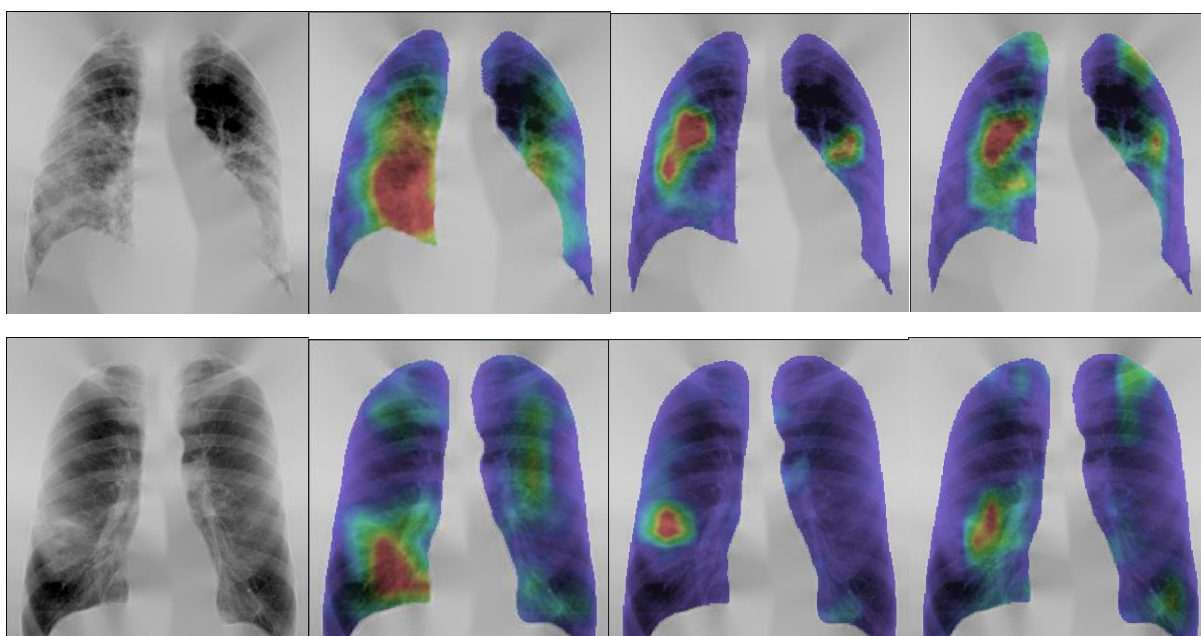


Рисунок 24 – Тепловые карты, построенные методом Grad-CAM, на модели VGG19, предобработка inpaint, слева направо: исходное изображение, 224x224, 288x288, 320x320 (очаговые тени, пневмония)

В таблице 7 приведены значения AUC для метки наличия патологии в соответствие с проведенными экспериментами.

Можно отметить, что при увеличении размеров для нормализованных изображений площадь выделенных очагов увеличивается, для изображений с

предобработкой inpainting наблюдается, наоборот, уменьшение площади. Также для нормализованных изображений увеличивается значение AUC.

Таблица 7 - Значения AUC для бинарной метки наличия патологии на модели VGG19

	AUC
224x224	0.898
288x288	0.901
320x320	0.906
Inpaint, 224x224	0.902
Inpaint, 288x288	0.9
Inpaint, 320x320	0.897

На рисунке 25 представлены тепловые карты, построенные методом из пункта 3.1.4, с использованием модели сверточной нейронной сети VGG16. Для обучения сети для бинарной классификации использовались нормализованные изображения размера 320x320. Значение AUC составило 0,893.

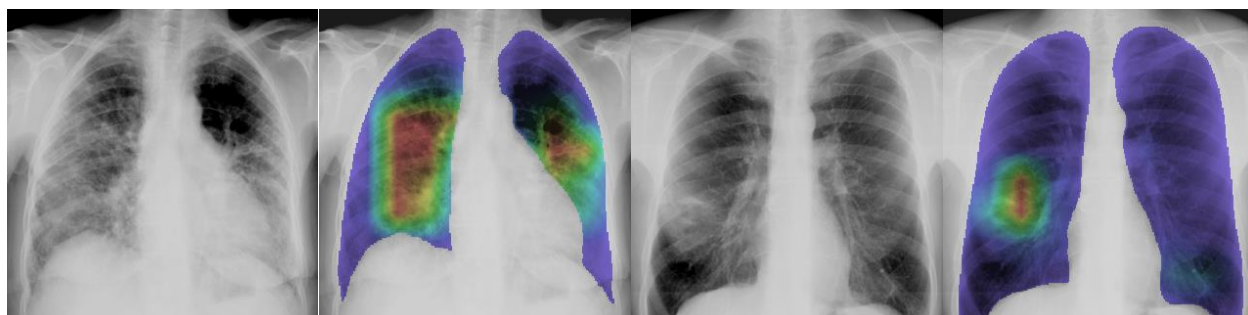


Рисунок 25 – Тепловые карты для изображений 320x320, VGG16

Тепловые карты, построенные данным методом, выделяют участки очагов, подобно методу Grad-CAM, однако существенным отличием является то, что на примерах не наблюдается ложных очагов на границах легких.

3.5 Выводы

В данной главе рассмотрено:

1. Алгоритмы построения тепловых карт. Перечислены особенности их реализации;
2. Произведено описание модели HSV;
3. Определен процесс предобработки изображений;
4. Проведены эксперименты с различной предобработкой изображений и различными размерами изображений.

ЗАКЛЮЧЕНИЕ

В магистерской диссертации были выполнены следующие задачи:

1. В первой главе рассмотрены основные понятия системы автоматизированной диагностики; нейронных сетей; приведен обзор исследований;
2. Размечены исходные данные, подготовлена обучающая выборка;
3. Проведена предобработка исходных изображений с использованием сегментации и нормализации;
4. Проведены эксперименты для классификации multi-label с различными моделями нейронных сетей;
5. Проведено сравнение результатов классификации с использованием метрики AUC;
6. Реализовано несколько алгоритмов построения тепловых карт;
7. Проведены эксперименты для построения тепловых карт с различной предобработкой и применением различных слоев сверточной сети.
8. Проведены эксперименты для построения тепловых карт с использованием моделей сверточных нейронных сетей, обученных на изображениях, имеющих различные размеры.

Направления дальнейших исследований:

1. Рассмотреть обучающую выборку, где здоровых людей больше, чем с патологией, которая будет отражать распределение патологии в современном мире;
2. Обучение других моделей сверточных нейронных сетей;
3. Применение методов борьбы с переобучением для возможного улучшения качества классификации;
4. Исследовать другие методы предобработки изображений.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Computer aided detection of tuberculosis on chest radiographs: An evaluation of the CAD4TB v6 system / Keelin Murphy [et al.]. – 2020. – 11 p.
2. Нейронные сети. 2019. <https://future2day.ru/nejronnye-seti/>
3. VGG16 – сверточная сеть для выделения признаков изображений. 2018. <https://neurohive.io/ru/vidy-nejrosetej/vgg16-model/>
4. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning / P.Rajpurkar [et al.]. – 2017. – 7p.
5. Going deeper with convolutions / C. Szegedy [et al.]. – Google Inc., 2014. – 12p.
6. A convolutional neural network for automatic identification and classification of fall army worm moth / Francis Chulu [et al.]. – University of Zambia, IJACSA, 2019. – No. 7. Vol. 10. – P. 112-118.
7. Chollet F., Xception: Deep Learning with Depthwise Separable Convolutions / F. Chollet. – Google Inc., 2017. – 8p.
8. Comparing deep learning models for population screening using chest radiography / R. Sivaramakrishnan [et al.]. – 2018.
9. Liauchuk V., Detection of lung pathologies using deep convolutional networks trained on large x-ray chest screening database / V. Liauchuk, V. Kovalev. – 2019.
10. Большая медицинская энциклопедия. <https://бмэ.орг>
11. Histogram equalization. https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_imgproc/py_histograms/py_histogram_equalization/py_histogram_equalization.html
12. O. Ronneberger, U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention - (MICCAI), Munich, Germany 5-9 October;
13. U-Net: нейросеть для сегментации изображений. 2018. <https://neurohive.io/ru/vidy-nejrosetej/u-net-image-segmentation/>

14. Selvaraju R. R., Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization / R. R. Selvaraju, M. Cogswell, A. Das. – 2019.
15. HSV (Цветовая модель).
[https://ru.wikipedia.org/wiki/HSV_\(цветовая_модель\)](https://ru.wikipedia.org/wiki/HSV_(цветовая_модель))
16. Bertalmio M., Navier-Stokes, Fluid Dynamics, and Image and Video Inpainting / M. Bertalmio, A. L. Bertozzi, G. Sapiro.