

Правительство Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Санкт-Петербургский государственный университет»

Кафедра информационно-аналитических систем

Ефремов Александр Александрович

Поиск аномалий в рентгеновских снимках
при помощи глубокого машинного
обучения с использованием
предварительной обработки снимков для
сегментации легких и удаления костей

Бакалаврская работа

Научный руководитель:
к.ф.-м.н., старший преподаватель кафедры информатики СПбГУ
Салищев С. И.

Рецензент:
к.ф.-м.н., доцент кафедры параллельных алгоритмов СПбГУ
Евдокимова Т. О.

Санкт-Петербург
2018

SAINT-PETERSBURG STATE UNIVERSITY

Chair of Analytical Information Systems

Alexander Efremov

Detecting Anomalies in Chest X-Ray Images
by Means of Deep Learning with Image
Preprocessing by Lung Segmentation and
Bone Exclusion Techniques

Bachelor's Thesis

Scientific supervisor:
senior lecturer Sergei Salishev

Reviewer:
associate professor Tatjana Evdokimova

Saint-Petersburg
2018

Оглавление

Введение	4
1. Постановка задачи	6
2. Обзор	7
2.1. Основные понятия	7
2.2. Описание используемых данных	9
2.3. Существующие решения	10
2.3.1. Усиление локального контраста	11
2.3.2. Сегментация легких	11
2.3.3. Фильтрация костной ткани	12
2.3.4. Метрики качества классификации	19
3. Особенности реализации	21
3.1. Подготовка данных для обучения	21
3.2. Особенности методов предобработки	22
3.3. Выбор модели для классификации	22
4. Эксперименты	24
5. Результаты	25
Заключение	27
Список литературы	29

Введение

Диагностика легочных заболеваний сохраняет свою значимость для многих жителей нашей планеты. Особенно она актуальна для стран с высоко развитой промышленностью в совокупности с недостаточным контролем над очистительными сооружениями, что, в свою очередь, повышает риск подверженности таким заболеваниям, как пневмония, туберкулез, рак легких и другим. В связи с этим возникает потребность в регулярном мониторинге состояния легких для предупреждения или обнаружения заболевания, прежде чем оно нанесет серьезный вред здоровью.

Рентгенография — технически простой и наиболее доступный для населения метод предварительной диагностики заболевания, в том числе отслеживания заболевания на ранней стадии. Данный тип исследования широко распространен по всему миру и вопреки наличию других, более сложных и эффективных способов диагностики, он остается приоритетным в медицинской практике даже в хорошо оснащенных клиниках и медицинских центрах. Несмотря на все преимущества данного метода, анализ и выявление признаков заболевания по флюорографическим снимкам является весьма сложной задачей, требующей участия нескольких высоко квалифицированных специалистов. Сложность задачи заключается, как правило, в низком разрешении изображений и наличии различного рода “шумов” (засвеченное изображение, наличие посторонних предметов и т.д.).

Помимо шумов, дополнительные трудности вносит наличие на снимках естественных анатомических структур (например, костей, других перекрывающих органов), которые, зачастую, скрывают аномалии, содержащиеся в легочной ткани. Выделение и последующая фильтрация подобного рода структур, с сохранением информации о легочной ткани под ними, является важной задачей для систем автоматизированного проектирования (САПР), используемых для рентгенографического исследования грудной клетки и может значительно повысить точность обнаружения аномалий.

Так, авторы [2] провели исследование эффективности применения методов сегментации и фильтрации костной ткани ребер на рентгеновских снимках и выяснили, что использование данных методов позволяет повысить итоговую точность классификации. К примеру, использование метода фильтрации костной ткани позволило авторам повысить точность классификации примерно на 30% против случая, когда снимки классифицировались без какой-либо предварительной обработки. Таким образом, итоговая точность оказалась около 70%. Конечно, даже этот результат нельзя назвать высоким, но здесь стоит отметить, что авторы использовали для классификации самую простейшую модель сверточной нейронной сети, без конфигурирования её параметров.

В данной работе будут рассмотрены предложенные авторами методы предварительной обработки рентгеновских снимков и проведено сравнение качества классификации на наборах данных с разной степенью предобработки. Отличительной чертой данной работы можно считать использование набора снимков, содержащего значительно больший объем изображений, нежели тот, что использовали авторы указанной статьи. Это позволит объективно сравнить результаты, полученные в данной работе, с результатами авторов.

1. Постановка задачи

Пользуясь методами глубокого машинного обучения, проанализировать эффективность применения методов выделения контура легких и исключения тени костей из грудной клетки при анализе рентгеновских снимков и выявлении подозрительных областей.

Для достижения данной цели были поставлены следующие задачи:

- Изучить существующие методы глубокого машинного обучения в контексте применения к медицинским задачам
- Найти и подготовить набор данных, необходимый для обучения алгоритмов
- Проанализировать и выбрать наиболее подходящие алгоритмы для предобработки и классификации изображений
- Провести необходимые эксперименты
- Проанализировать полученные результаты
- Предоставить сравнительный анализ работы различных сочетаний алгоритмов

2. Обзор

В данной главе приводятся необходимые понятия из теории машинного обучения, описываются методы предобработки рентгеновских снимков, а также данные, используемые в данной работе.

2.1. Основные понятия

Глубокое обучение (ГО) представляет собой набор методов машинного обучения (МО), использующих многослойную систему нелинейных преобразователей для извлечения и дальнейшего преобразования признаков [12]. В контексте нейронных сетей это буквально означает, что нейросеть содержит больше одного скрытого слоя. Методы глубокого обучения основываются на изучении представления данных, а не на алгоритмах для решения специфических задач, которые лежат в основе “неглубоких” методов. Это позволяет им строить более высокие абстракции, всё больше приближая эти методы к принципам работы человеческого мозга. В отличие от обычных методов МО, которые делят задачу несколько небольших частей, а затем собирают всё в один результат, методы ГО используют для решения поставленных задач комплексный подход. Более того, если обычные методы МО нуждаются в заранее определенных для работы функциях, то методы ГО создают эти функции самостоятельно, основываясь на результатах, получаемых в процессе обучения. Другими словами, концепция ГО подразумевает, что машина не просто обучается, а постоянно совершенствует сам способ обучения. Методы ГО, так же как и методы МО, подразделяются на методы обучения с учителем, с частичным привлечением учителя и обучения без учителя.

Сверточная нейронная сеть (Convolutional neural network) или CNN — это класс глубоких нейронных сетей прямого распространения, успешно применяющихся при распознавании визуальных образов, в том числе и при распознавании изображений [11]. Работа сети устроена таким образом, что принимая изображение, сеть выделяет некоторые конкретные особенности этого изображения, после чего переходит к получе-

нию каких-то более абстрактных деталей и так вплоть до выделения абстракций высокого уровня. При этом, в процессе обучения сеть самостоятельно перенастраивается и выстраивает наиболее подходящую иерархию абстрактных признаков, пренебрегая маловажными (на её взгляд) деталями и выделяя наиболее значимые. Как правило, сверточная нейронная сеть состоит из достаточно большого количества различных слоев. Приведем основные типы слоев, которые обычно входят в состав CNN: слой свертки (convolutional layer), слой субдискретизации или подвыборки (pooling layer), слой активации (ReLU layer) и полносвязный слой (fully connected layer). Пример типичной сверточной нейронной сети изображен на рисунке 1. Подробно с теорией глубокого обучения, а также с устройством сверточных нейронных сетей можно ознакомиться в [4].

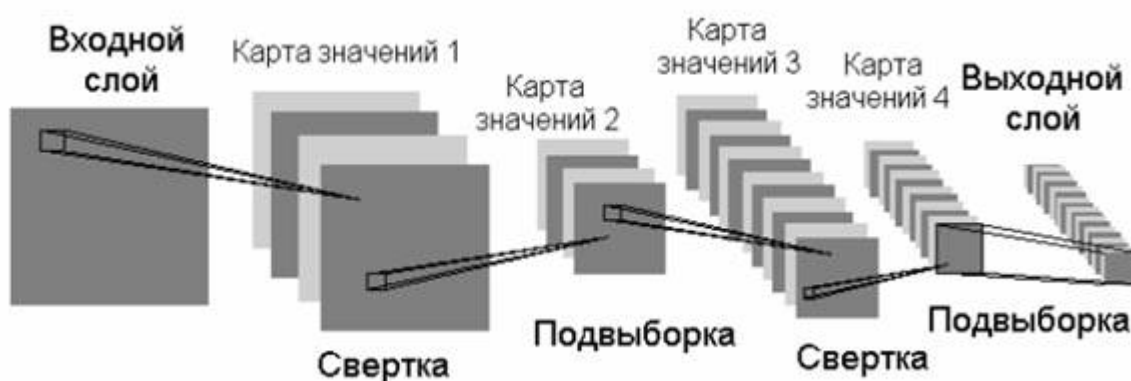


Рис. 1: Обобщенная схема сверточной нейронной сети

Одним из этапов данной работы является классификация рентгеновских снимков на два класса: здоровые легкие и легкие, содержащие какую-либо патологию. Любая задача классификации ставится следующим образом: имеется множество объектов (ситуаций), разделённых некоторым образом на классы. Задано конечное множество объектов, для которых известно, к каким классам они относятся. Это множество называется обучающей выборкой (training set). Классовая принадлежность остальных объектов не известна. Требуется построить алгоритм, способный указать наименование класса, к которому относится произвольный объект из исходного множества [7]. Для проверки рабо-

ты обученного алгоритма следует также подготовить тестовую выборку (validation set). Именно на этом множестве и будет оцениваться точность классификации.

Весь процесс работы алгоритмов машинного обучения, как правило, разбивается на несколько логически следующих друг за другом частей. На рисунке 2 приведена наиболее часто встречающаяся схема, содержащая основные этапы работы таких алгоритмов:

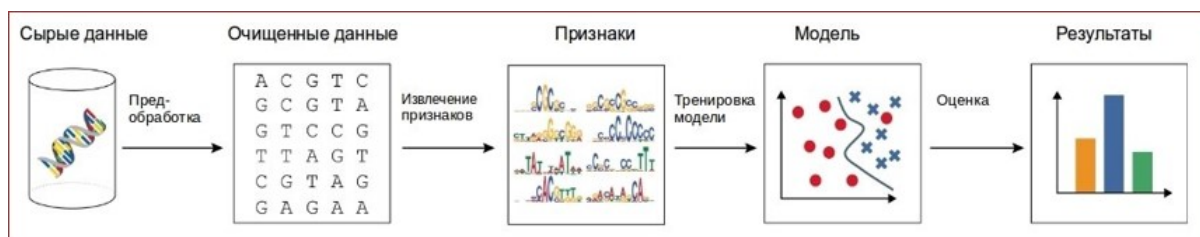


Рис. 2: Классическая схема этапов машинного обучения

2.2. Описание используемых данных

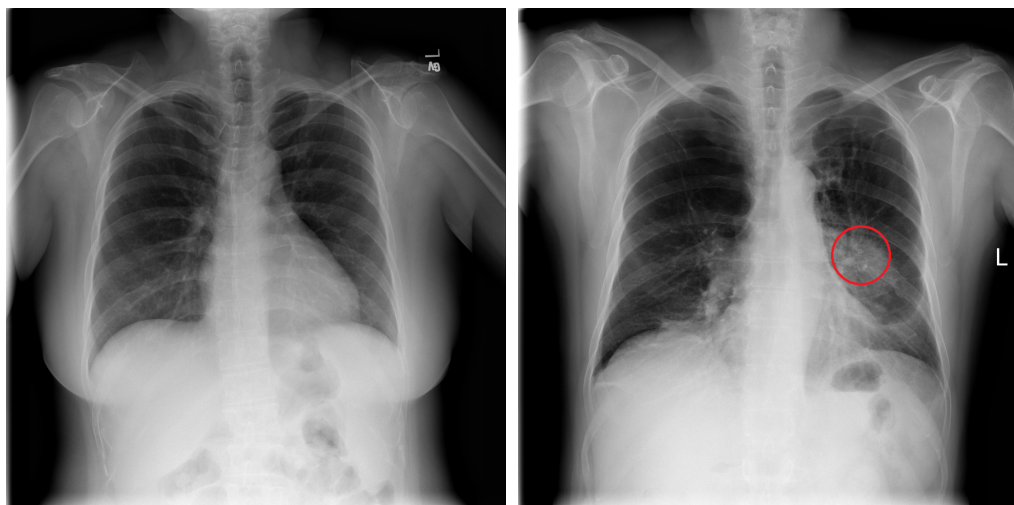
Существует масса датасетов, содержащих рентгеновские снимки легких, как здорового, так и патологического характера. После изучения нескольких примеров подходящих наборов данных, предпочтение было отдано датасету, который был подготовлен Клиническим Центром в составе Национальных Институтов Здоровья США (NIH Clinical Center) [6]. Набор данных именуется как NIH Chest X-rays Dataset. Ключевым критерием выбора стало количество снимков, содержащееся в наборе, поскольку одной из особенностей методов глубокого обучения является потребность в как можно большем количестве исходных данных. Полная информация о каждом из рассмотренных наборов приведена в таблице 1.

Выбранный датасет предоставляется в свободном доступе и содержит чуть больше 112 тысяч снимков [5]. Изображения представлены в формате PNG с глубиной цвета 8 бит. Каждый снимок имеет разрешение 1024x1024 пикселя. После анализа содержимого датасета было установлено соотношение здоровых снимков (около 60% датасета)

Название датасета	Без патологии	С патологией	Общее количество	Формат	Разрешение
Montgomery Country chest X-ray dataset	80	58	138	DICOM	4020x4892
Belarus tuberculosis database	-	464	464	DICOM	2248x2248
NIH Chest X-rays dataset	63170	48950	112120	PNG	1024x1024
JSRT database	93	154	247	PNG	2048x2048
Shenzhen Hospital X-ray Set	326	336	662	JPEG	N/A

Таблица 1: Сравнение наборов рентгеновских снимков легких

к снимкам, содержащим какую-либо патологию (около 40% датасета). Приведем пример здорового снимка и снимка с патологией:



(a) Без патологии

(b) С патологией

Рис. 3: Примеры снимков из датасета NIH Chest X-rays Dataset

2.3. Существующие решения

Одним из ключевых этапов машинного обучения является этап предварительной обработки данных. В зависимости от специфики изучаемой области и преследуемых целей, выбираются те или иные способы

предобработки. В данной работе будут использованы следующие методы предобработки рентгеновских снимков: усиление локального контраста, сегментация легких и фильтрация костной ткани ребер.

2.3.1. Усиление локального контраста

Одним из значимых факторов, влияющих на качество изображения, является уровень контраста изображения. Если контраст слишком низкий или слишком высокий, то изображение вряд ли будет пригодным для обучения. Для повышения контрастности изображения используют методы усиления контраста. В работе было использовано общее решение для повышения локального контраста рентгеновских изображений:

$$L_{LCE} = \frac{L - L'}{\sqrt{L^2 - (L')^2}} \quad (1)$$

где L — исходное изображение, а L' — изображение, к которому применено размытие Гаусса.

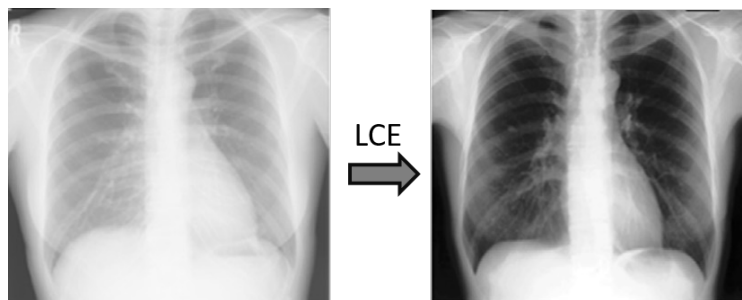


Рис. 4: Пример работы метода усиления локального контраста

2.3.2. Сегментация легких

Приступая к анализу рентгеновских снимков легких, как правило, начинают с выделения контура легких. Так, под задачей сегментации легких подразумевается распознавание и последующее отделение области легких от сердца и других органов, присутствующих на снимках. Для решения данной задачи была выбрана сверточная нейронная сеть на базе архитектуры UNet. Данная архитектура наилучшим образом зарекомендовала себя в применении к задачам сегментации

биомедицинских изображений [9]. К недостаткам данной архитектуры можно отнести требование высоких вычислительных мощностей, а также относительно слабое распределение снимков по видам патологий. Первый недостаток устраняется наличием соответствующего оборудования, позволяющего производить обучение с использованием графического процессора (GPU). Второй — с большой долей вероятности не повлияет на результаты обучения, поскольку в данной работе рассматривается задача бинарной классификации, а именно задача определения наличия, либо отсутствия патологии на снимке.

Задействованный в данной работе метод сегментации легких применяется к исходным изображениям и включает в себя следующие шаги:

- Поиск границ легких посредством обучения сверточной нейронной сети на базе UNet для сегментации легких на исходном наборе данных с подготовленными вручную “масками” (границами легких)
- Определение границ легких в виде черно-белых “масок” с помощью обученной сверточной нейронной сети для каждого исходного изображения из набора данных
- Выделение области интересов (правое и левое легкие) из исходных снимков



Рис. 5: Пример работы метода сегментации легких

2.3.3. Фильтрация костной ткани

Проблема автоматического обнаружения с последующей фильтрацией костной ткани ребер на рентгеновских снимках рассматривает-

ся уже довольно долгое время. Несмотря на это, именно фильтрации уделялось не очень много внимания, хотя применение данного метода имеет большое значение в клинической практике и является весьма полезным при диагностике легочных заболеваний. Более того, использование данного приема, в конечном счете, позволяет повысить точность классификации легких на 25-30% [2].

Все существующие алгоритмы фильтрации костной ткани внутри грудной клетки состоят из нескольких последовательно выполняющихся шагов. После изучения нескольких таких алгоритмов [1, 3, 13], был выбран алгоритм, описанный в [3]. Первым критерием отбора стал тот факт, что в отсутствие каких-либо предварительно размеченных данных, необходимо было, чтобы алгоритм использовал принцип обучения без учителя. Этому критерию удовлетворяли два из найденных алгоритмов, описанных в [3] и [13]. Необходимо было сравнить эти два алгоритма и использовать тот, что имеет больше преимуществ. Наибольшую трудность при фильтрации костной ткани вызывает этап определения разности интенсивности текстуры кости и легочной ткани. Данный этап нужен для непосредственного удаления силуэтов костей за счет сведения разницы между интенсивностями к минимуму. Для вычисления интенсивности текстур используется линейная модель: интенсивность ребра в точке (x,y) определяется как $ax+b$. По сути, задача определения интенсивностей сводится к задаче оптимизации подбора наилучших для модели параметров (a,b) . Для решения этой задачи сравниваемые алгоритмы используют разные подходы. Так, в [13] авторы предлагают использовать специальный генетический алгоритм (RCGA), основываясь на том, что в генетические алгоритмы заложены принципы, заимствованные из биологии и генетики, что, в свою очередь, позволит достичь хороших результатов. Однако, предложенный авторами алгоритм имеет несколько недостатков. Во-первых, при вычислении разностей интенсивности однородность текстуры кости может быть искажена. Ко всему прочему, добавляется высокая вычислительная сложность генетических алгоритмов. На это указывают авторы [3] и предлагают алгоритм, основанный на модели регрессии, который учи-

тывает однородность кости, затрачивая при этом меньше времени на вычисления.

Алгоритм, описанный в [3], работает в предположении, что значение разности интенсивностей кости и легочной ткани может меняться от центра кости к её границам, однако, это значение остается практически неизменным на соседних участках вдоль параболической кривой. Это допущение позволяет методу сначала предсказать разницу интенсивности в центре кости, а затем расширить это значение в вертикальном направлении для сведения разницы между границами ребер и легочной тканью, находящейся сверху или снизу от границ, к минимуму. Таким образом, в данной работе было решено использовать алгоритм, предложенный в [3]. Приведем краткое описание основных шагов этого алгоритма:

1. Извлечение шаблона

Первым шагом при поиске ребра является извлечение шаблона ребра из содержимого снимка. Этот шаг позволяет системе работать по принципу обучения без учителя. Сначала производится предварительное выделение границ ребра с помощью специального оператора обнаружения границ, после этого к изображению применяется простая схема выборки для фиксации шаблона, обнаруженного в центре области легкого. С помощью определенной оценки выбираются несколько подходящих шаблонов-кандидатов для каждого из легких по отдельности.

2. Поиск крайних точек границ ребра

Для поиска границ ребер применяется логарифмический фильтр Габора (Log-Gabor filter) [10]. Данный фильтр часто используют для распознавания границ объектов на изображении. На данном шаге производится первичное выделение ребер, начиная от “стартовых точек”. “Стартовые точки” располагаются в месте, где ребра пересекаются с внешней границей легких. Стоит отметить, что на данном шаге структура ребра выделяется частично.

3. Выбор лучшего шаблона

На данном шаге производится сокращение числа потенциально подходящих шаблонов для правого и левого легких путем вычисления расстояния между границами ребер на шаблоне и структурами, найденными на предыдущем шаге. Шаблон с наименьшим таким расстоянием выбирается для дальнейшей работы системы. После того, как произведен выбор лучшего шаблона для одного из легких, шаблон для другого легкого создается путем зеркального отражения первого. Определение симметрии для создаваемого шаблона происходит в три этапа: поскольку форма шаблона должна представлять собой эллиптическую кривую, необходимо определить коэффициенты p_1, p_2, \dots, p_{n+1} полинома $p(x)$ степени n ($n \geq 2$), которые лучше всего соответствуют данным по y . Затем производится подсчет симметрии и, наконец, определяется конечное положение шаблона по отношению к центру масс области легкого.

$$p(x) = p_1x^n + p_2x^{n-1} + \dots + p_nx + p_{n+1} \quad (2)$$

4. Оценка ширины ребра

Для возможности дальнейшего выделения ребра производится вычисление средней высоты (или ширины) ребра. Главная цель данного шага — найти максимально длинную вертикальную линию, которую можно вписать в структуру ребра. Устанавливая правую и левую границу допустимой ширины ребра, производится вычисление среднего значения на основе всех допустимых вертикальных линий, найденных в структуре ребра.

5. Настройка параболы

На этом шаге производится корректировка вертикального положения шаблона. Сдвиг производится за счет оптимизации коэффициента при свободном члене полинома $p(x)$, приведенного на шаге 3. Оставляя все остальные коэффициенты неизменными, про-

изводится подсчет значений p_{n+1} для каждого ребра. Таким образом, найденные значения позволяют определить конечное положение шаблона.

6. Расширение границ ребер

Каждое ребро внутри грудной клетки имеет некоторые особенности в своей геометрии, в связи с чем, глобальный шаблон в виде параболы, исходящий из найденной “стартовой точки”, может не вписаться в структуру кривизны ребра. Эту проблему можно решить несколькими способами. Авторы предлагают свой подход, основанный на уточнении положения центра ребра путем переопределения коэффициентов все того же полинома $p(x)$, специфичных для ребра. Это позволит расширить ребро линейно, достигнув предсказанной высоты ребра. Таким образом, данный шаг позволяет выделить поверхность ребра более полно.

7. Фильтрация костной ткани

Заключительным этапом метода является удаление с рентгеновского снимка шумов, вызванных присутствием костной ткани ребер. Данная операция производится путем подавления силуэтов ребер с сохранением информации о легочной ткани, расположенной под ними. Подавление начинается с пикселей от вертикального центра ребра и адаптивно расширяется до границ ребра. По сути, необходимо определить интенсивность легочной ткани под ребром по следующей формуле:

$$S' = S - D \quad (3)$$

где $S' = (s_1', s_2', \dots, s_{k'})$ — интенсивность легочной ткани под ребром, $S = (s_1, s_2, \dots, s_l)$ — видимая интенсивность и $D = (d_1, d_2, \dots, d_k)$ — интенсивность кости; k — число пикселей, расположенных вдоль вертикального направления от центра ребра.

Неизвестными в данном случае являются значения интенсивно-

стей кости (d_1, d_2, \dots, d_k) , которые подлежат определению. Для подсчета этих значений авторы используют геометрию анатомической модели поперечного сечения ребра, показанной на рисунке 6а.

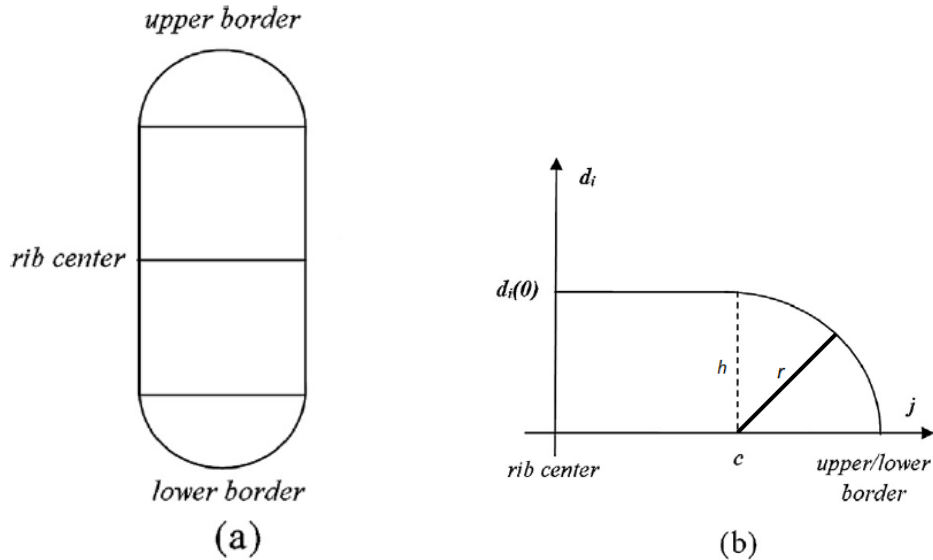


Рис. 6: (а) Анатомическая модель поперечного сечения ребра и (б) соответствующая ей геометрическая модель силуэта ребра

Исходя из этой модели, значение i -го пикселя d_i (i отвечает за перемещение вдоль параболической кривой) считается неизменным до достижения вертикального положения c , где ребро начинает сужаться. Приняв вертикальную позицию центра ребра за 0, интенсивность ребра $d_i(j)$ для вертикального положения j находится из следующей системы:

$$d_i(j) = \begin{cases} \frac{d_i(0)}{rh - c} \sqrt{(rh - c)^2 - (j - c)^2}, & j > c \\ d_i(0), & j \leq c \end{cases} \quad (4)$$

Здесь нужно задать два параметра: c (точка, где ребро становится тоньше) и $d_i(0)$ (разность интенсивностей для центра ребра). Авторы задают $c = rh/4$, полагая, что это сгладит переходы между границей ребра и легочной тканью. Вторым параметром авторы

предлагают определять, используя следующую модель регрессии:

$$d_i(j) = \frac{\sum_{-\frac{rh}{2} \leq y \leq \frac{rh}{2}} i - w < x < i + w^{S_x(y)} - \left(\sum_{\frac{rh}{2} < y < rh} i - w < x < i + w^{S_x(y)} + \sum_{-y < -\frac{rh}{2}} i - w < x < i + w^{S_x(y)} \right)}{2wrh} \quad (5)$$

где w — параметр, отвечающий за ширину окна, в котором производится сравнение интенсивностей между соответствующими областями. Этот параметр задается вручную. После нескольких экспериментов авторы задали значение $w = 20$.

Формально, разность интенсивностей, которую нужно вычесть из центра ребра, вычисляется путем вычитания среднего значения интенсивности пикселей чуть над верхней границей ребра и тех, которые находятся чуть под нижней границей ребра из наблюдаемой средней интенсивности в центре ребра. Такой подход имеет смысл, поскольку совокупная интенсивность рентгенографического изображения грудной клетки линейно изменяется в вертикальном направлении из-за расположения пациента при проведении обследования.

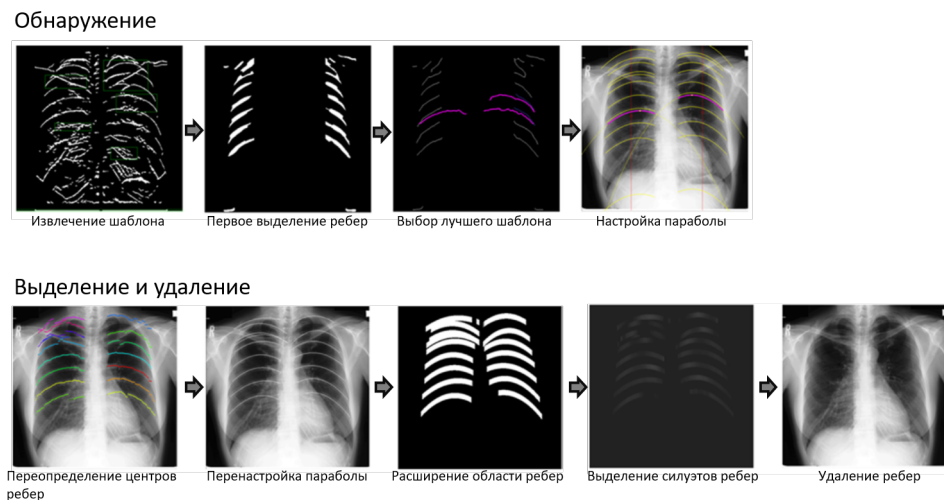


Рис. 7: Пример работы метода фильтрации костной ткани

2.3.4. Метрики качества классификации

Для понимания того, насколько хорошо работает тот или иной алгоритм, необходимо ввести численную метрику оценки его качества. Введение подобных метрик также позволит легко сравнить качество работы разных алгоритмов между собой. В данной работе были использованы следующие метрики:

1. Accuracy

Простейшим методом оценивания качества работы алгоритма является оценка общего числа верно предсказанных элементов внутри тестовой выборки. Данная метрика представляет собой отношение количества элементов, по которым алгоритм вынес верное решение к размеру всей обучающей выборки:

$$Accuracy = \frac{P}{N}, \quad (6)$$

где P — число верно предсказанных элементов, N — размер обучающей выборки.

Важно отметить тот факт, что данная оценка плохо учитывает точность предсказания для каждого класса по отдельности, поэтому результат может оказаться далек от истины в случае распределения элементов обучающей выборки в сторону какого-то одного или нескольких классов. Проще говоря, можно получить высокую общую точность, но при этом для какого-то отдельного класса она окажется очень низкой. Поэтому данная оценка была использована после проведения основных экспериментов для большей наглядности конечных результатов.

2. F_1 -мера

Данный подход основан на двух оценках точности, называемых Precision (точность) и Recall (полнота), которые определяются сле-

дующим образом:

$$P = \frac{TP}{TP + FP} \quad (7)$$

$$R = \frac{TP}{TP + FN}, \quad (8)$$

где TP — количество верно предсказанных элементов, $TP + FP$ — количество всех предсказанных элементов, $TP + FN$ — количество ожидаемых элементов.

Precision показывает, сколько полученных от классификатора положительных ответов являются верными. Чем больше точность, тем меньше промахов. Recall характеризует способность классификатора определять как можно больше положительных ответов из ожидаемых. Очевидно, чем выше обе эти оценки, тем лучше точность предсказания. Однако, на практике одновременное достижение максимума обеих оценок невозможно. Поэтому для качественной оценки классификатора применяется так называемая F_1 -мера, которая связывает Precision и Recall и определяется как их среднее гармоническое:

$$F_1 = 2 \cdot \frac{PR}{P + R} \quad (9)$$

F_1 -мера дает наиболее полную картину при оценке точности работы классификатора, учитывая распределение данных по классам, что является её неоспоримым преимуществом перед другими существующими оценками. Также данная метрика позволяет избежать таких случаев, как, например, отнесение всех данных к одному классу, когда значение Precision будет очень низким, а значение Recall, напротив, слишком высоким. Таким образом, данная оценка использовалась в качестве основной при проведении экспериментов.

3. Особенности реализации

В данной главе приводятся некоторые особенности отдельных этапов работы. В качестве основного языка программирования использовался Python 2.7. Также использовались вспомогательные библиотеки машинного обучения Keras и Tensorflow.

3.1. Подготовка данных для обучения

Перед тем, как подавать рентгеновские снимки на сегментацию или фильтрацию костной ткани ребер, необходимо тщательно подготовить изображения для достижения лучшего результата. Самым первым шагом стал анализ содержимого снимков. Оказалось, что многие снимки непригодны для обучения на них нейросетей. В связи с чем, было принято решение вручную удалить “плохие” снимки. Снимок признавался “плохим” в следующих случаях (примеры на рисунке 8):

- Снимок сделан не в прямой проекции
- Положение грудной клетки на снимке слишком искажено
- Снимок содержит много посторонних предметов в области грудной клетки, таких как: провода от медицинских приборов, булавки, пуговицы, кардиостимуляторы и так далее
- Снимок слишком размыт, либо засвечен

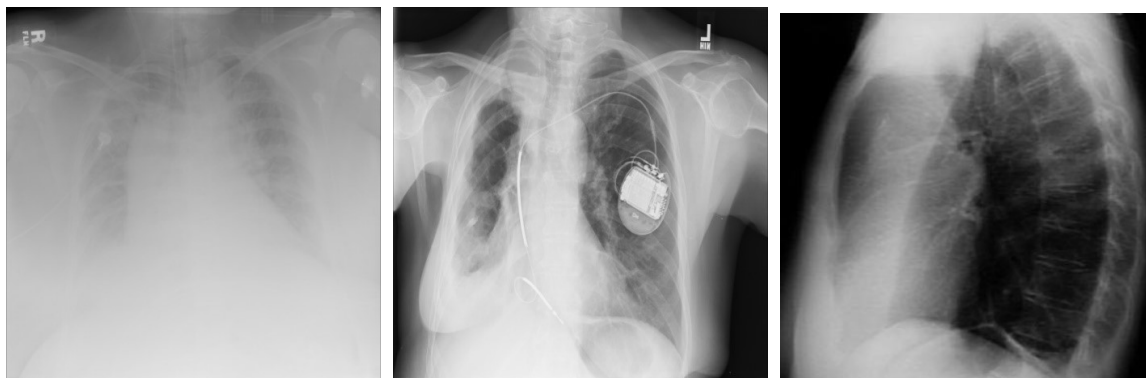


Рис. 8: Примеры “плохих” снимков

В результате значительная часть снимков была отсеяна. Несмотря на это, количество оставшихся снимков вполне удовлетворяет запросам нейросети. Результаты первого этапа обработки представлены в таблице 2.

	До	После
Без патологии	63337	27320
С патологией	48783	13863
Всего	112120	41169

Таблица 2: Результаты ручной фильтрации снимков

Далее необходимо произвести сжатие изображений, поскольку чем больше разрешение снимка, тем медленнее происходит обучение. Все изображения были приведены к разрешению 224x224 пикселей. Помимо сжатия, очень полезным приемом является повышение контраста отдельных частей изображения (LCE). Данный прием позволяет повысить качество изображений, сделав очертания соответствующих структур (легких, ребер, сердца и т.д.) более четкими и различимыми, что, в свою очередь, улучшает точность работы других алгоритмов. Описание данного метода можно найти в предыдущей главе.

3.2. Особенности методов предобработки

В условиях отсутствия “масок” для рассматриваемого набора данных, а также для упрощения и ускорения процесса выполнения работы для сегментации данных была использована уже обученная модель. Обучение проводилось на двух датасетах: Montgomery и JSRT [8].

3.3. Выбор модели для классификации

В связи с тем, что целью данной работы является, в первую очередь, анализ эффективности методов сегментации и фильтрации костной ткани на снимках, для самой классификации была выбрана мо-

дель, удовлетворяющая минимальным требованиям: простая структура и небольшое время обучения. Также не производилось специальное конфигурирование модели, поскольку задачи получения максимально возможной точности и минимальных потерь при классификации в данной работе поставлено не было. Основываясь на этих требованиях, была выбрана стандартная модель сверточной нейронной сети на базе архитектуры AlexNet. Сеть содержит восемь слоев: 5 слоев свертки и 3 полносвязных слоя. Структура этой нейронной сети показана на рисунке 9. В данной нейросети используется функция активации ReLU. ReLU слой расположен после каждого сверточного и полносвязного слоя.

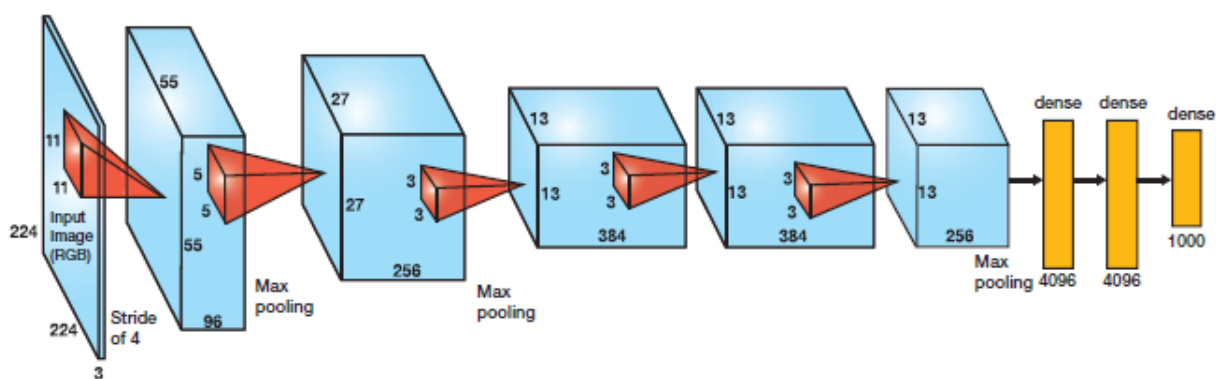


Рис. 9: Архитектура нейросети AlexNet

4. Эксперименты

Для проведения экспериментов был использован предварительно отфильтрованный вручную набор NIH Chest X-rays Dataset. Было подготовлено четыре набора снимков с разной степенью предобработки: исходный набор без предобработки, набор с сегментированными снимками, набор со снимками без костей, а также набор, содержащий снимки, обработанные обоими методами. Примеры снимков каждого из наборов представлены на рисунке 10. Количество и состав снимков во всех использованных наборах одинаков. 90% наборов использовалось для обучения и 10% — для валидации. Тренировка сети производилась на изображениях с разрешением 224x224 пикселя, на каждую итерацию подавалось по 256 снимков. Итого, для тренировки было использовано 24600 здоровых снимков и 11800 снимков с патологиями. Оставшиеся снимки были использованы для валидации.



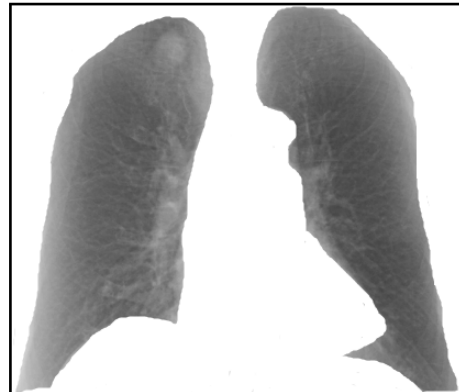
(a) Без обработки



(b) Сегментированное



(c) Без костей

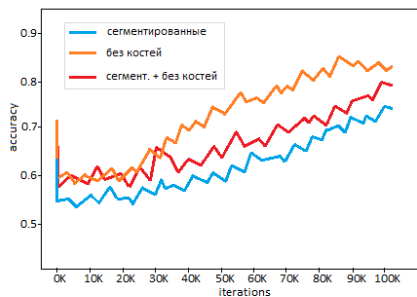


(d) Сегментированное, без костей

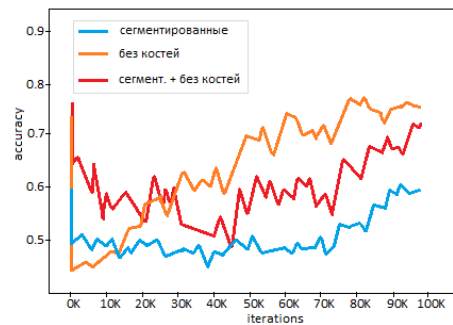
Рис. 10: Примеры снимков для обучения

5. Результаты

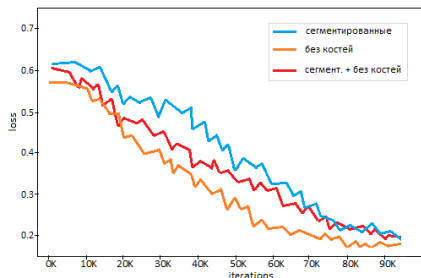
Для обучения предварительно было подготовлено четыре набора данных с разной степенью предобработки. Обучение проводилось на сверточной нейронной сети архитектуры AlexNet. После того, как обучение было завершено, последовала валидация обученных алгоритмов. Валидационный набор состоит из 4000 снимков: 2500 снимков со здоровыми легкими и 1500 — с какой-либо аномалией. Хуже всего результаты оказались у сети, обученной на данных без какой-либо предварительной обработки (за исключением усиления контраста), что, впрочем, было вполне ожидаемо. Точность классификации составила около 46%. На фоне данного результата, та же сеть, использующая данные, предобработанные хотя бы одним из методов сегментации или фильтрации костной ткани, показали в сравнении гораздо лучшие результаты. На рисунке 11b можно увидеть, что точность одного из вариантов обучения составила порядка 75%.



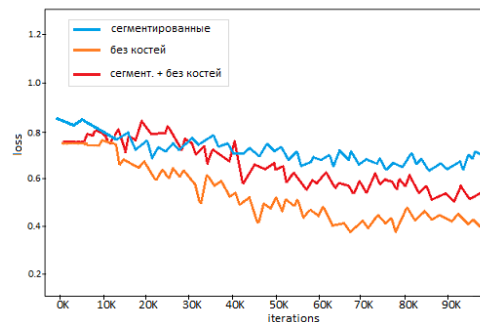
(a) Training accuracy



(b) Validation accuracy



(c) Training loss



(d) Validation loss

Рис. 11: Результаты классификации

В целом, при тренировке, все три оставшихся варианта с предоб-

работкой снимков продемонстрировали повышение качества классификации (рисунки 11a и 11c), однако на валидационном наборе результаты оказались несколько хуже. Также наблюдается увеличение разницы между значениями функции потерь для тренировочного и тестового наборов данных, это можно объяснить вероятным переобучением сети (рисунок 11d).

Также после проведения всех экспериментов были подсчитаны основные метрики оценки качества классификации. Результаты приведены в таблице 3. Наибольшее значение F_1 -меры показала сеть, обученная на снимках без костей и составило 0.70. Исходя из результатов для варианта обучения на исходных снимках, общее улучшение составило почти 30%, что подтверждает пользу применения методов предобработки рентгеновских снимков при распознавании аномалий в легких.

Использованные снимки	Precision	Recall	F_1 -мера	Общая точность
Без обработки	0.46	0.38	0.42	0.452
Сегментированные	0.63	0.51	0.56	0.587
Без костей	0.75	0.66	0.70	0.753
Сегментированные + без костей	0.72	0.64	0.68	0.730

Таблица 3: Оценка качества работы классификаторов

Для наглядности результатов также представлена валидационная матрица ошибок (confusion matrix) для метода фильтрации костной ткани (таблица 4). Из таблицы видно, что число снимков (как здоровых,

Без костей		Expected	
		Normal	Abnormal
Actual	Normal	1647	542
	Abnormal	853	958

Таблица 4: Матрица ошибок для тестовой выборки

так и с аномалиями), неверно классифицированных сетью, достаточно велико (853 и 542 снимка соответственно). Сведение этих значений к минимуму позволит ещё больше повысить качество работы системы, а следовательно, и уровень доверия к её результатам врачей-рентгенологов.

Заключение

В рамках данной выпускной квалификационной работы были достигнуты следующие результаты:

- Произведен обзор существующих методов глубокого обучения и алгоритмов предварительной обработки рентгеновских снимков.
- Выбран и подготовлен подходящий набор данных для обучения.
- Для каждого этапа обучения был подобран соответствующий алгоритм, выбрана наиболее подходящая архитектура нейронной сети.
- Проведены эксперименты посредством обучения нейросетей на данных с предобработкой и без неё.
- Представлен сравнительный анализ работы алгоритмов, была подтверждена эффективность применения методов сегментации легких и фильтрации костной ткани внутри грудной клетки.

По результатам данной работы сделаны следующие выводы:

1. Использование различных методов предварительной обработки рентгеновских снимков легких позволяет повысить итоговое качество распознавания аномалий в грудной клетке. Этап предобработки является одним из значимых при автоматизированном анализе рентгеновских снимков.
2. Наиболее эффективным оказался метод фильтрации костной ткани ребер. После его применения точность классификации возросла почти на 30% и составила 71%. При этом использование данного метода также показало лучшее значение F_1 -меры (0.71). Также стоит отметить достигнутое в результате применения данного метода значение метрики Recall, а именно 0.67. Полученное значение, хоть и является лучшим в данной работе, но, тем не менее,

этого мало для практического использования данного метода, когда значение Recall должно быть близким к 0.99. Представляется возможным повысить данное значение в дальнейшем за счет рассмотрения более сложных архитектур сверточных нейросетей для классификации рентгеновских снимков.

Список литературы

- [1] Atlas-based rib-bone detection in chest X-rays / S. Candemir, S. Jaeger, S. Antani et al. // Computerized Medical Imaging and Graphics. — 2016. — Vol. 51. — P. 32–39.
- [2] Deep Learning with Lung Segmentation and Bone Shadow Exclusion Techniques for Chest X-Ray Analysis of Lung Cancer / Yu. Gordienko, P. Gang, J. Hui et al. // ArXiv. — 2017. — Access mode: <https://arxiv.org/abs/1712.07632>.
- [3] Eliminating rib shadows in chest radiographic images providing diagnostic assistance / H. Ogul, B.B. Ogul, A.M. Agildere et al. // Computer Methods and Programs in Biomedicine. — 2016. — Vol. 127. — P. 174–184.
- [4] Goodfellow I., Bengio Y., Courville A. Deep Learning. — MIT Press, 2016. — Access mode: <http://www.deeplearningbook.org>.
- [5] Health National Institutes Of. NIH Chest X-rays. — 2017. — Access mode: <https://www.kaggle.com/nih-chest-xrays/data> (online; accessed: 25.04.2018).
- [6] Health National Institutes Of. NIH Clinical Center provides one of the largest publicly available chest x-ray datasets to scientific community. — 2017. — Access mode: <https://www.nih.gov/news-events/news-releases/nih-clinical-center-provides-one-largest-publicly-available-chest-x-ray-datasets-to-scientific-community> (online; accessed: 25.04.2018).
- [7] MachineLearning.ru. Классификация. — 2011. — Access mode: <http://www.machinelearning.ru/wiki/index.php?title=Классификация> (online; accessed: 25.04.2018).
- [8] Petsiuk V. Lung Segmentation (2D). — 2017. — Access mode: <https://github.com/imlab-uiip/lung-segmentation-2d> (online; accessed: 25.04.2018).

- [9] Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, LNCS. — 2015. — Vol. 9351. — P. 234–241.
- [10] Wikipedia. Log Gabor filter. — 2012. — Access mode: https://en.wikipedia.org/wiki/Log_Gabor_filter (online; accessed: 25.04.2018).
- [11] Wikipedia. Convolutional neural network. — 2018. — Access mode: https://en.wikipedia.org/wiki/Convolutional_neural_network (online; accessed: 25.04.2018).
- [12] Wikipedia. Deep learning. — 2018. — Access mode: https://en.wikipedia.org/wiki/Deep_learning (online; accessed: 25.04.2018).
- [13] A nonparametric-based rib suppression method for chest radiographs / L. Jiann-Shu, W. Jing-Wein, W. Hsing-Hsien, Y. Ming-Zheng // Computers & Mathematics with Applications. — 2012. — Vol. 64. — P. 1390–1399.