

Инженерная школа информационных технологий и робототехники
 Направление подготовки 09.04.04 Программная инженерия
 Отделение информационных технологий

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Информационно-аналитическая система для автоматизированной диагностики патологий органов грудной клетки

УДК 004.652.032.26:621.386.8

Студент

Группа	ФИО	Подпись	Дата
8ПМ8И	Кулебякин Алексей Сергеевич		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Аксёнов Сергей Владимирович	К.Т.Н.		

КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Меньшикова Екатерина Валентиновна	к. филос. н.		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ООД ШБИП	Горбенко Михаил Владимирович	К.Т.Н.		

По разделу на английском языке

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель	Пичугова Инна Леонидовна	—		

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Губин Евгений Иванович	к.ф.-м.н.		

ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ООП

Код результата	Результат обучения
<i>Общие по направлению подготовки 09.04.04 «Программная инженерия»</i>	
P1	Способность проводить научные исследования, связанные с объектами профессиональной деятельности
P2	Способность разрабатывать новые и улучшать существующие методы и алгоритмы обработки данных в информационно-вычислительных системах
P3	Способность составлять отчеты о проведенной научно-исследовательской работе и публиковать научные результаты
P4	Способность проектировать системы с параллельной обработкой данных и высокопроизводительные системы
P5	Способность осуществлять программную реализацию информационно-вычислительных систем, в том числе распределенных
P6	Способность осуществлять программную реализацию систем с параллельной обработкой данных и высокопроизводительных систем
P7	Способность организовывать промышленное тестирование создаваемого программного обеспечения
<i>Профиль «Технологии больших данных»/ «Big data solutions»</i>	
P8	Способность исследовать и анализировать большие данные, создавать их модели и интерпретировать структуры данных в таких моделях
P9	Способность понимать принципы создания, хранения, управления, передачи и анализа больших данных с использованием новейших технологий, инструментов и систем обработки данных в высокопроизводительных сетях
P10	Способность применять теорию распределенной системы управления базами данных к традиционным распределенным системам реляционных баз данных, облачным базам данных, крупномасштабным системам машинного обучения и хранилищам данных

Министерство науки и высшего образования Российской Федерации
 федеральное государственное автономное
 образовательное учреждение высшего образования
 «Национальный исследовательский Томский политехнический университет» (ТПУ)

Инженерная школа информационных технологий и робототехники
 Направление подготовки 09.04.04 Программная инженерия
 Отделение информационных технологий

УТВЕРЖДАЮ:
 Руководитель ООП
 _____ Губин Е.И.
 (Подпись) _____ (Дата) (Ф.И.О.)

ЗАДАНИЕ на выполнение выпускной квалификационной работы

В форме:

Магистерской диссертации

(бакалаврской работы, дипломного проекта/работы, магистерской диссертации)

Студенту:

Группа	ФИО
8ПМ8И	Кулебякин Алексей Сергеевич

Тема работы:

Информационно-аналитическая система для автоматизированной диагностики патологий органов грудной клетки	
Утверждена приказом директора (дата, номер)	№59-62/с от 28.02.2020 г.

Срок сдачи студентом выполненной работы:	15.06.2020
--	------------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе <i>(наименование объекта исследования или проектирования; производительность или нагрузка; режим работы (непрерывный, периодический, циклический и т. д.); вид сырья или материал изделия; требования к продукту, изделию или процессу; особые требования к особенностям функционирования (эксплуатации) объекта или изделия в плане безопасности эксплуатации, влияния на окружающую среду, энергозатратам; экономический анализ и т. д.).</i>	<ol style="list-style-type: none"> 1. Набор данных, состоящий из 112120 рентгеновских снимков во фронтальной плоскости грудной клетки 2. Требование разработать систему для их анализа
--	--

<p>Перечень подлежащих исследованию, проектированию и разработке вопросов</p> <p><i>(аналитический обзор по литературным источникам с целью выяснения достижений мировой науки техники в рассматриваемой области; постановка задачи исследования, проектирования, конструирования; содержание процедуры исследования, проектирования, конструирования; обсуждение результатов выполненной работы; наименование дополнительных разделов, подлежащих разработке; заключение по работе).</i></p>	<ol style="list-style-type: none"> 1. Проведение обзора конкурентных решений 2. Описание метода, позволяющего определять патологии на рентгенограммах 3. Реализация алгоритма обнаружения пневмонии на рентгенограмме 4. Разработка приложения для работы с алгоритмом 5. Финансовый менеджмент; 6. Социальная ответственность; 7. Заключение.
<p>Перечень графического материала</p> <p><i>(с точным указанием обязательных чертежей)</i></p>	<p>Рисунки и таблицы для описания исследования предметной области, реализованных модулей системы, составления финансовой стороны и уточнения социальной ответственности.</p>
<p>Консультанты по разделам выпускной квалификационной работы</p>	
<p>Раздел</p>	<p>Консультант</p>
<p>Финансовый менеджмент, ресурсоэффективность и ресурсосбережение</p>	<p>Меньшикова Екатерина Валентиновна, доцент ОСГН ШБИП, к. филос. н.</p>
<p>Социальная ответственность</p>	<p>Горбенко Михаил Владимирович, доцент ООД ШБИП, к.т.н.</p>
<p>Раздел на английском языке</p>	<p>Пичугова Инна Леонидовна, старший преподаватель ОИЯ ШБИП</p>

<p>Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику</p>	<p>01.03.2020</p>
--	-------------------

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
<p>Доцент ОИТ ИШИТР</p>	<p>Аксёнов Сергей Владимирович</p>	<p>к.т.н.</p>		<p>01.03.2020</p>

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
<p>8ПМ8И</p>	<p>Кулебякин Алексей Сергеевич</p>		<p>01.03.2020</p>

Инженерная школа информационных технологий и робототехники
 Направление подготовки 09.04.04 Программная инженерия
 Уровень образования – магистратура
 Отделение информационных технологий
 Период выполнения: весенний семестр 2019/2020 учебного года

Форма выполнения работы:

Магистерская диссертация

(бакалаврская работа, дипломный проект/работа, магистерская диссертация)

КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН выполнения выпускной квалификационной работы

Срок сдачи студентом выполненной работы:	15.06.2020
--	------------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
01.04.2020	Модуль автоматизированной диагностики патологий органов грудной клетки	30
01.05.2020	Программная оболочка для работы с модулем диагностики патологий	30
24.05.2020	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	20
31.05.2020	Социальная ответственность	20

Составил преподаватель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Аксёнов Сергей Владимирович	к.т.н.		

СОГЛАСОВАНО:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Губин Евгений Иванович	к.т.н., доцент		

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И
РЕСУРСОСБЕРЕЖЕНИЕ»**

Студенту:

Группа	ФИО
8ПМ8И	Кулебякин Алексей Сергеевич

Школа	ИШИТР	Отделение	ОИТ
Уровень образования	Магистратура	Направление/специальность	09.09.04 Программная инженерия

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:

1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	Оклад руководителя - 35120 руб. Оклад инженера - 21760 руб. Стоимость материальных ресурсов определялась согласно прейскурантам компаний
2. Нормы и нормативы расходования ресурсов	Накладные расходы 10%; Районный коэффициент 30%. Норма амортизации ПЭВМ 33,3%
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	Коэффициент отчислений на уплату во внебюджетные фонды 30 %

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого и инновационного потенциала НИИ	Анализ потенциальных потребителей результатов исследования, оценка качества и перспективности проекта по технологии QuaD, SWOT-анализ
2. Разработка устава научно-технического проекта	Инициация проекта: определение заинтересованных сторон проекта, целей и результатов проекта
3. Планирование процесса управления НИИ: структура и график проведения, бюджет, риски и организация закупок	План проекта, определение трудоемкости выполнения работ, разработка графика проведения научного исследования, расчет бюджета разработки
4. Определение ресурсной, финансовой, экономической эффективности	Описание потенциального эффекта

Перечень графического материала (с точным указанием обязательных чертежей):

1. Оценочная карта для QuaD-анализа разработки
2. Матрица SWOT разработки
3. Диаграмма Ганта
4. График проведения разработки
5. Бюджет затрат
6. Основные мероприятия по снижению рисков

Дата выдачи задания для раздела по линейному графику	01.03.2020
---	------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Меньшикова Екатерина Валентиновна	к. филос. н.		01.03.2020

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ПМ8И	Кулебякин Алексей Сергеевич		

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»**

Студенту:

Группа	ФИО
8ПМ8И	Кулебякин Алексей Сергеевич

Школа	ИШИТР	Отделение (НОЦ)	ОИТ
Уровень образования	Магистратура	Направление/специальность	09.04.04 Программная инженерия

Тема ВКР:

Информационно-аналитическая система для автоматизированной диагностики патологий органов грудной клетки	
Исходные данные к разделу «Социальная ответственность»:	
1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения	Разработка представляет собой информационно-аналитическую систему для автоматизированной диагностики патологий органов грудной клетки. Система может использоваться медицинским персоналом везде, где есть доступ к сети интернет.
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
1. Правовые и организационные вопросы обеспечения безопасности: – специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства; – организационные мероприятия при компоновке рабочей зоны.	– Трудовой кодекс Российской Федерации от 30.12.2001 N 197ФЗ – Федеральный Закон от 27.07.2006 N 152-ФЗ (ред. от 25.07.2011) «О Персональных Данных» – ГОСТ 12.2.032-78 ССБТ
2. Производственная безопасность: 2.1. Анализ выявленных вредных и опасных факторов 2.2. Обоснование мероприятий по снижению воздействия	– Отклонение показателей микроклимата – Превышение уровня шума – Отсутствие или недостаток естественного и искусственного света – Поражение электрическим током: – Умственное перенапряжение:
3. Экологическая безопасность.	Несущими угрозу окружающей среде являются следующие объекты: – Компьютерное оборудование.
4. Безопасность в чрезвычайных ситуациях.	Наиболее вероятной ЧС является пожар в здании, в т.ч. из-за короткого замыкания электропроводки. Возможными ЧС могут являться наводнения, удары молний, взрывы, эпидемии инфекционных заболеваний

Дата выдачи задания для раздела по линейному графику	01.03.2020
--	------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ООД ШБИП	Горбенко Михаил Владимирович	К.Т.Н.		01.03.2020

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ПМ8И	Кулебякин Алексей Сергеевич		

Реферат

Выпускная квалификационная работа состоит из 76 страниц, включает в себя 16 рисунков, 19 таблиц и 1 приложение. При работе были использованы 23 источников литературы.

Ключевые слова: история болезни, патологии органов грудной клетки, машинное обучение, веб-приложение, автоматизированная диагностика.

Цель работы – разработка инструмента анализа рентгенологических снимков для локализации патологий в области лёгкого.

Объектом исследования является практическая часть разработки алгоритма обнаружения торакальных заболеваний и патологий на рентгенограммах, а также приложения для работы с ним с целью постановки предварительного диагноза.

Область внедрения данной работы – здравоохранение.

В работе выполнен анализ имеющихся на рынке решений со схожим функционалом, рассмотрены алгоритмы обнаружения патологий на рентгенограммах, обучена и протестирована модель машинного обучения, разработано веб-приложение для простой и удобной работы с моделью.

В качестве дальнейших шагов планируется расширение функционала для диагностики других патологий и заболеваний, тестирование системы в медицинских учреждениях г. Томска

Определения, обозначения, сокращения

В пояснительной записке использовались следующие термины с соответствующими определениями, обозначениями и сокращениями:

рентгенограмма – зарегистрированное на светочувствительном материале (фотоплёнке, фотопластинке) изображение объекта, возникающее в результате взаимодействия рентгеновских лучей с веществом.

локализация – процесс определения местоположения.

DenseNet – Densely Connected Convolutional Network

AUC-ROC – Area Under the Curve - Receiver Operating Characteristics

AI – Artificial Intelligence (Искусственный интеллект)

ML – Machine Learning (Машинное обучение)

DL – Deep Learning (Глубокое обучение)

Оглавление

Введение.....	13
1 Аналитический обзор предметной области.....	15
1.1 Обзор конкурентных решений.....	15
1.1.1 ChestEye.....	15
1.1.2 qXR.....	15
1.2 Алгоритмы глубокого обучения для работы с рентген-снимками.....	17
1.3 Набор данных ChestX-ray14.....	19
2. Разработка модуля для обнаружения пневмонии на рентгенограммах.....	20
2.1 Выбор программного обеспечения.....	20
2.2 Используемые библиотеки.....	20
2.3 Построение и обучение DenseNet-121.....	21
2.4 Выводы по разделу.....	23
3. Проектирование и разработка приложения для обнаружения патологий органов грудной клетки по рентген-снимкам.....	25
Результаты.....	29
4. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение.....	30
4.1 Предпроектный анализ.....	30
4.1.1 Технология QuaD.....	30
4.1.2 SWOT-анализ.....	32
4.1.3 Оценка готовности научно-исследовательского проекта к коммерциализации.....	33
4.2 Инициация разработки.....	34
4.3 Планирование управления разработкой.....	35
4.3.1 План разработки.....	35
4.3.2 Разработка графика проведения разработки.....	36
4.4 Бюджет разработки.....	37
4.4.1 Расчет материальных затрат.....	37
4.4.2 Амортизационные затраты.....	37
4.4.3 Зарботная плата исполнителей.....	38
4.4.4 Отчисления во внебюджетные фонды.....	40
4.4.5 Накладные расходы.....	40
4.4.6 Формирование бюджета.....	40
4.5 Определение потенциального эффекта разработки.....	42
5. Социальная ответственность.....	43
5.1 Правовые и организационные вопросы обеспечения безопасности.....	43

5.2	Производственная безопасность.....	45
5.2.1	Анализ выявленных вредных и опасных факторов	45
5.2.2	Обоснование мероприятий по снижению воздействия	46
5.2.2.1	Отклонение показателей микроклимата.....	46
5.2.2.2	Превышение уровня шума	47
5.2.2.3	Отсутствие или недостаток света	49
5.2.2.4	Повышенное значение напряжения в электрической цепи	50
5.2.2.5	Умственное перенапряжение.....	51
5.3	Экологическая безопасность.....	52
5.4	Безопасность в чрезвычайных ситуациях	54
5.4.1	Мероприятия по предотвращению ЧС и разработка порядка действий в случае возникновения ЧС	54
5.5	Выводы по разделу.....	56
	Заключение.....	58
	Список достижений	59
	Список используемых источников.....	61
	Приложение А.....	64

Введение

Современный мир постоянно меняется, автоматизация проникает во всё большее количество отраслей и отрасль здравоохранения не является исключением. На приёме пациента доктору больше не нужно заполнять такие данные, как имя, дата рождения, пол и т.д. При сборе анамнеза доктору могут предлагаться подходящие ключевые слова и возможные варианты диагнозов на основании симптомов и предыдущих записей из истории болезни.

Однако процессы, связанные с более сложными данными, например, рентген-снимками всё ещё нуждаются в автоматизации, поскольку число собираемых данных неуклонно растёт, особенно в 2020 году из-за вспышки COVID-19.

С развитием моделей машинного обучения, а также постоянно пополняющимся объёмом снимков становится возможным диагностировать патологии с точностью высококвалифицированного врача-радиолога и даже лучше.

Целью работы является разработка инструмента анализа рентгенологических снимков для локализации патологий в области лёгкого. В качестве инструментов обучения модели используются глубокие нейронные сети.

Основные задачи исследования:

1. провести анализ существующих решений для автоматизированной диагностики патологий органов грудной клетки;
2. разработать алгоритм обнаружения патологий органов грудной клетки (пневмонии), используя размеченные рентген-снимки;
3. провести тестирование алгоритма для оценки точности обнаружения;
4. разработать программную оболочку для работы с алгоритмом.

Под объектом исследования понимается разработка алгоритма обнаружения торакальных заболеваний и патологий на рентгенограммах, а также приложения для работы с ним с целью постановки предварительного диагноза.

Предметом исследования является обзор существующих решений со схожим функционалом, архитектуры глубокой нейронной сети, подходящей для решения поставленных задач.

Научная новизна работы заключается в улучшении качества обнаружения патологий на рентгенограммах грудной клетки, а также разработке программного решения с перспективой его внедрения в медицинские учреждения г. Томска.

Практическая значимость результатов магистерской диссертации заключается в приложения для диагностики и локализации патологий по загружаемым рентгенограммам грудной клетки.

Результаты работы планируются к использованию в отделении лучевой диагностики СибГМУ г. Томска.

1 Аналитический обзор предметной области

1.1 Обзор конкурентных решений

На сегодняшний день на рынке представлено множество решений для телемедицины, а также для автоматизации обработки медицинских данных. Большинство из них позволяют провести консультацию со специалистом путём видеоконференции. Однако можно отметить два решения, обладающих функционалом для диагностики патологий по рентгенограммам.

В данном разделе будут представлены описания каждого из решений. В конце раздела будут выбраны основные параметры разрабатываемой системы диагностики.

1.1.1 ChestEye

Одним из решений для автоматизированной обработки и диагностики рентгенограмм является платформа ChestEye компании Ohipit, базирующейся в Польше [1]. Решение компании позволяет диагностировать 75 различных патологий органов грудной клетки. Платформа идентифицирует рентгенограммы без аномалий и создает предварительные отчеты. Предварительные отчеты о здоровых пациентах создаются только в тех случаях, когда система максимально уверена в результатах, примерно у 20-40% от всего числа здоровых пациентов. Минимизируя работу рентгенологов, необходимую для получения диагностики по рентгенограммам здоровых пациентов, ChestEye повышает производительность радиологов и позволяет медицинским специалистам сосредоточить свое внимание на случаях с отклонениями. К сожалению, компания не сообщает о качестве и точности работы платформы.

1.1.2 qXR

Другим, более перспективным решением можно считать разработки компании Qure.ai, базирующейся в Индии [2]. Система qXR v2.0 позволяет обнаружить аномальные результаты на рентгенограммах грудной клетки. Основной сценарий использования — отделение рентгенограмм здоровых пациентов от рентгенограмм пациентов с патологиями органов грудной клетки.

qXR также включает в себя собственный алгоритм, который проверяет рентгенограммы на наличие признаков туберкулеза.

Преимущество qXR по сравнению с ChestEye заключается в возможности сгенерировать описание найденной патологии, включающее в себя название, размер и расположение патологии. Полученные данные используются для предварительного заполнения отчёта о проведении рентгенографии.

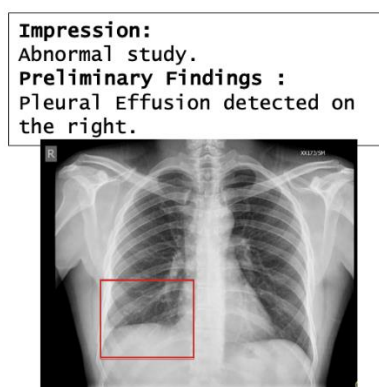


Рисунок 1 – Текстовый отчёт о работе системы qXR

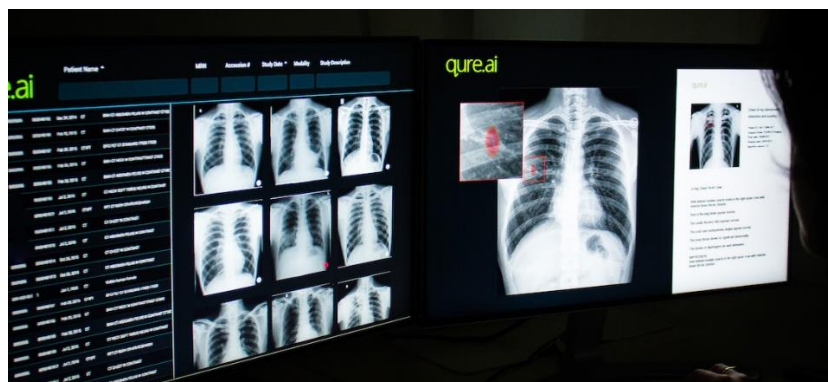


Рисунок 2 – Пользовательский интерфейс системы qXR

Таким образом, следует разработать решение, которое не только не будет уступать рассмотренным выше конкурентным решениям, но и превосходить их. В частности, планируется значительно улучшить интерфейс по работе с медицинскими данными, а также заложить в систему возможность для расширения функционала в будущем и масштабирования на большое количество медицинских учреждений.

1.2 Алгоритмы глубокого обучения для работы с рентген-снимками

Алгоритмы глубокого обучения, такие как глубокие нейронные сети, известны своей высокой производительностью и качеством работы, когда дело доходит до обработки изображений [3; 4; 5].

Существует множество архитектур нейронных сетей, которые подходят для поиска паттернов на изображениях. Одна из них – DenseNet (Densely Connected Convolutional Network) [6]. Изначально она была разработана для обучения на наборе данных ImageNet, содержащей более 14 миллионов изображений, разбитых на 21 841 категорию. Сеть можно легко переобучить для работы с различными наборами данных, в том числе и с рентген-снимками.

Более ранние нейронные сети обладали меньшим количеством слоёв (LeNet – 5 слоёв, VGG – 19 слоёв). DenseNet же может содержать от 121 до 264 скрытых слоёв, что значительно усложняет её архитектуру.

DenseNet соединяет каждый слой с каждым последующим слоем последовательно, в то время как традиционные сверточные сети с L слоями имеют L соединений, по одному между каждым слоем – DenseNet имеет $L(L + 1) / 2$ прямых соединений. Карты признаков всех предыдущих слоев используются в качестве входных данных для каждого слоя, а его собственные карты признаков используются в качестве входных данных для всех последующих слоев. Сети DenseNet имеют ряд неоспоримых преимуществ: они уменьшают проблему исчезающего градиента, улучшают распространение признаков, стимулируют повторное использование признаков и существенно сокращают количество параметров.

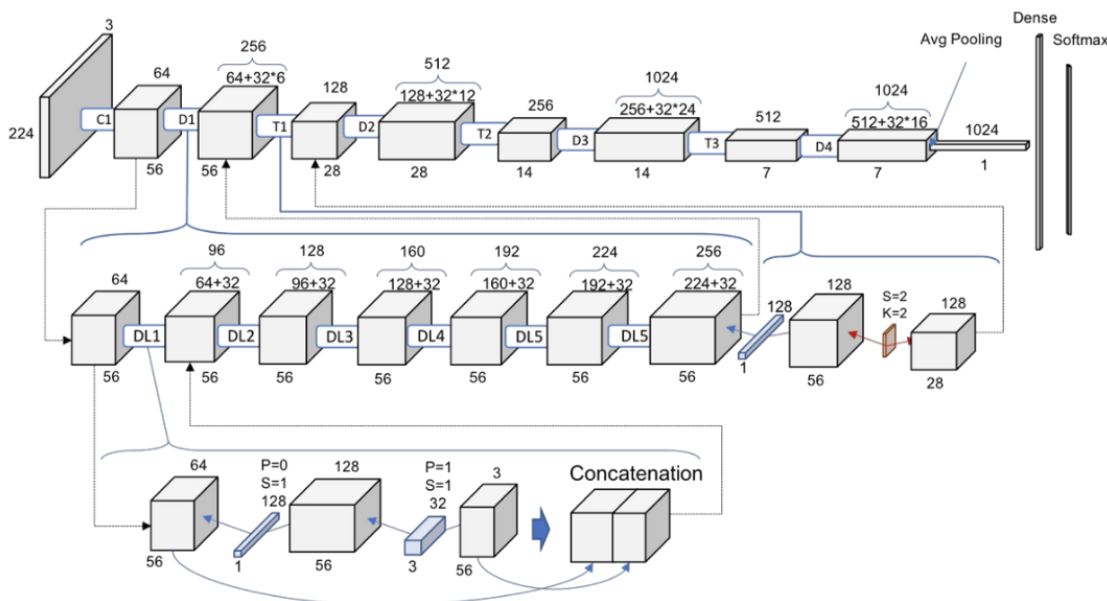


Рисунок 3 – Архитектура свёрточной нейронной сети DenseNet-121

Здесь C_x – это свёрточный слой x , D_x – Dense-блок x , T_x – Transition-блок x , DL_x – Dense-слой x .

Если вы посмотрите внимательно, каждый Dense-слой добавляет к предыдущему 32 новые карты признаков. Поэтому количество карт признаков увеличивается с 64 до 256 каждые 6 слоев. Кроме того, Transition-блок выполняет свертку 1×1 со 128 фильтрами. Затем следует pooling-слой 2×2 с шагом 2, в результате чего количество карт признаков сокращается вдвое.

Однако необходимо также понять, что происходит внутри каждого Dense слоя в каждом Dense-блоке, чтобы понять, как работает вся сеть. На самом глубоком уровне можно увидеть, как на самом деле достигается такое поведение добавления 32-кратного количества слоев. Каждый Dense-слой имеет свертку 1×1 со 128 фильтрами для уменьшения размера карт признаков и выполнения вычислительно более сложной свертки 3×3 с выбранным коэффициентом скорости роста равным 32 картам признаков.

Затем входной блок и результат этих двух операций (которые одинаковы для каждого Dense-слоя в каждом Dense-блоке) объединяются в процессе добавления новой информации к общему знанию сети.

1.3 Набор данных ChestX-ray14

Набор данных ChestX-ray14 является популярным общедоступным набором, содержащим 112120 рентгеновских снимков во фронтальной плоскости грудной клетки, полученных у 30805 уникальных пациентов [7]. Каждое изображение в наборе данных снабжено аннотацией о 14 различных патологиях органов грудной клетки. Данные аннотации были получены с использованием методов автоматического извлечения данных из радиологических отчетов.

Для наглядности и в качестве демонстрации работы сети DenseNet-121 в данной работе из всех 14 патологий для обнаружения была выбрана только пневмония, как наиболее распространённая. Изображения, у которых в аннотации была указана пневмония, были помечены как положительные примеры, а все остальные изображения были помечены как отрицательные. Для задачи обнаружения пневмонии набор данных был случайным образом разделен на обучающую (28744 пациента, 98637 изображений), валидационную (1672 пациента, 6351 изображение) и тестовую (389 пациентов, 420 изображений) выборки. В выборках нет случаев, когда снимки одного и того же пациента могут оказаться в разных выборках. Перед подачей изображений в сеть изображения были уменьшены до 224×224 пикселей и нормализованы на основе среднего значения и стандартного отклонения изображений в наборе данных ImageNet, использованном для предобучения сети. Также обучающая выборка была расширена за счёт горизонтального отражения некоторых изображений в случайном порядке.



Рисунок 4 – Пример изображения из набора данных ChestX-ray14

2. Разработка модуля для обнаружения пневмонии на рентгенограммах

2.1 Выбор программного обеспечения

В качестве языка программирования для решения поставленной задачи был выбран язык Python. Python – высокоуровневый интерпретируемый язык программирования, построенный на идеях императивного, объектно-ориентированного и функционального программирования. Язык создан Гвидо ван Россумом в 1989 году и с тех пор непрерывно совершенствуется.

К преимуществам Python можно отнести:

- открытость его кода;
- стандарт PEP-8 стимулирует разработчиков на написание простого и хорошо читаемого кода;
- возможность простого разделения кода на отдельные модули;
- большое количество библиотек для машинного обучения;
- возможность разработки серверной части веб-приложения.

Исходя из данных преимуществ, язык программирования Python идеально подходит для решения поставленной задачи, т.е. разработки исполняемого модуля для обнаружения пневмонии.

2.2 Используемые библиотеки

В ходе реализации проекта были использованы некоторые Python библиотеки для работы с определенными типами данных и использования специальных функций, расширяющих возможности языка.

Библиотеки NumPy и Pandas для работы с многомерными массивами, матрицами, проведения над ними математических операций, а также для использования структуры DataFrame.

Библиотеки PIL и PyTorch были использованы для работы с изображениями и для построения и обучения сети DenseNet соответственно.

2.3 Построение и обучение DenseNet-121

Модель машинного обучения, использующую в качестве основы глубокую свёрточную сеть DenseNet-121 можно либо написать вручную, используя заранее написанные функции для слоёв сети из PyTorch, либо использовать уже предобученную на наборе данных ImageNet модель сети.

Поскольку в основную архитектуру сети изменения вносятся не будут, логичнее воспользоваться вторым вариантом. В таком случае, необходимо воспользоваться средним и стандартным отклонением полученным в процессе предобучения на ImageNet для нормализации.

Чтобы адаптировать DenseNet-121 под задачу необходимо заменить последний полносвязный слой в модели на слой с одним выходом и применить сигмоидальную нелинейность.

Обучение сети производилось с использованием оптимизатора Adam. На вход за один раз подаётся мини-батч из 16 изображений для ускорения процесса обучения. Стартовый коэффициент скорости обучения был выбран равным 0.01 и уменьшается в 10 раз каждый раз, когда ошибка на валидационной выборке перестаёт изменяться после очередной эпохи обучения. Задача состоит в том, чтобы выбрать модель с наименьшей ошибкой на валидационной выборке.

Для того, чтобы оценить производительность и качество обучения модели был использован метод подсчёта площади под кривой AUC (Area Under the Curve) ROC (Receiver Operating Characteristics). Кривая AUC-ROC – это измерение производительности для задачи классификации при различных настройках пороговых значений. ROC – это кривая вероятности, а AUC – степень или мера отделимости, она показывает, насколько хорошо модель способна отделить один класс от другого. Чем выше AUC, тем лучше модель способна отличить пациентов с патологией, например, пневмонией и без неё.

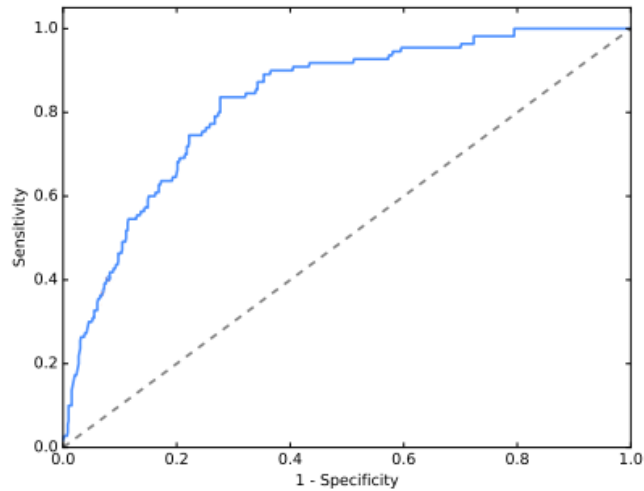


Рисунок 5 – Кривая AUC-ROC (среднее для 14 классов)

Для того, чтобы убедиться, что сеть была хорошо обучена, был проведён следующий эксперимент: были взяты ещё три модели – DenseNet-169 с большим числом слоёв по сравнению с DenseNet-121, а также ResNet-50 и SEResNet-50 и обучены на тех же на том же наборе данных.

В результате была получена сравнительная таблица 1, в которой указаны значения AUC-ROC для каждой модели по каждому классу, а также среднее значение по всем классам.

Таблица 1 – Значения AUC-ROC для DenseNet-121 в сравнении с другими сетями

Патология	DenseNet-121	DenseNet-169	ResNet-50	SEResNet-50
Atelectasis	0.8248	0.8156	0.8063	0.8144
Cardiomegaly	0.8982	0.8993	0.9072	0.8981
Effusion	0.8757	0.8792	0.8779	0.8743
Infiltration	0.6738	0.6737	0.6638	0.6756
Mass	0.8469	0.8313	0.8261	0.8342
Nodule	0.7643	0.7491	0.728	0.7583
Pneumonia	0.7781	0.7735	0.7551	0.7608
Pneumothorax	0.8541	0.8602	0.8382	0.8638
Consolidation	0.8107	0.8096	0.8069	0.8092
Edema	0.9033	0.9103	0.9073	0.9081
Emphysema	0.9209	0.91	0.8724	0.9194
Fibrosis	0.8284	0.818	0.7981	0.8291
Pleural Thickening	0.7735	0.7658	0.7431	0.766
Hernia	0.9194	0.9355	0.863	0.9322
AUC-ROC Mean	0.8337	0.8308	0.8138	0.8317

Из таблицы 1 можно видеть, что DenseNet-121 способна отделять некоторые патологии хуже, чем другие модели, взятые для сравнения, но среднее значение AUC-ROC, а также значение для пневмонии является лучшим из четырёх моделей.

2.4 Выводы по разделу

Таким образом, имея полученные данные, можно сделать вывод, что модель была выбрана верно и её можно использовать для обнаружения пневмонии на рентген-снимках.

Изменение гиперпараметров таких, как стартовое значение коэффициента скорости обучения не дало серьёзных улучшений. Снижая его до 0.001 и до 0.0001, увеличило время обучения, но поскольку в процессе обучения его значение всё равно опускается до этих значений, неудивительно, что не наблюдается улучшение производительности. Однако его увеличение до 0.1 снизило среднее значение AUC-ROC до 0.7956.

Дальнейших улучшений можно добиться если:

- использовать набор данных с большим количеством рентгенограмм, включающий в себя также боковую проекцию грудной клетки, а также базовую информацию о пациенте – его возраст, пол и ранее обнаруженные патологии;
- разработать модель с более совершенной архитектурой нейронной сети.

Оба варианта потребуют увеличения вычислительных мощностей для обучения модели, а также большое количество хорошо обученных специалистов-радиологов для разметки данных.

Полученные результаты показывают высокий потенциал разработанного решения и могут быть использованы, как минимум, для отделения рентгенограмм здоровых пациентов от снимков с найденными патологиями. Более того, если будет выявлена какая-либо патология, медицинский персонал будет сразу же уведомлен об этом, чтобы проверить результаты и либо утвердить предложенный диагноз, либо назначить некоторые дополнительные тесты,

чтобы убедиться в правильности / неправильности постановки предварительного диагноза, если специалист не доверяет предложенным рекомендациям.

3. Проектирование и разработка приложения для обнаружения патологий органов грудной клетки по рентген-снимкам

С целью проверки и демонстрации работоспособности предлагаемого классификатора, а также для упрощения работы с ним, необходимо создать приложение, основной целью которого является помощь врачу в постановке диагноза.

Приложение должно выполнять следующие задачи: загрузка, обработка и анализ рентген-снимков грудной клетки пациентов, сохранение информации о проведённом обследовании.

На основании поставленных цели и задач приложения можно перейти к его проектированию.

3.1 Проектирование приложения и выбор средств разработки

На рисунке 6 представлена диаграмма последовательности работы приложения по загрузке, анализу и сохранению данных пациента в систему.

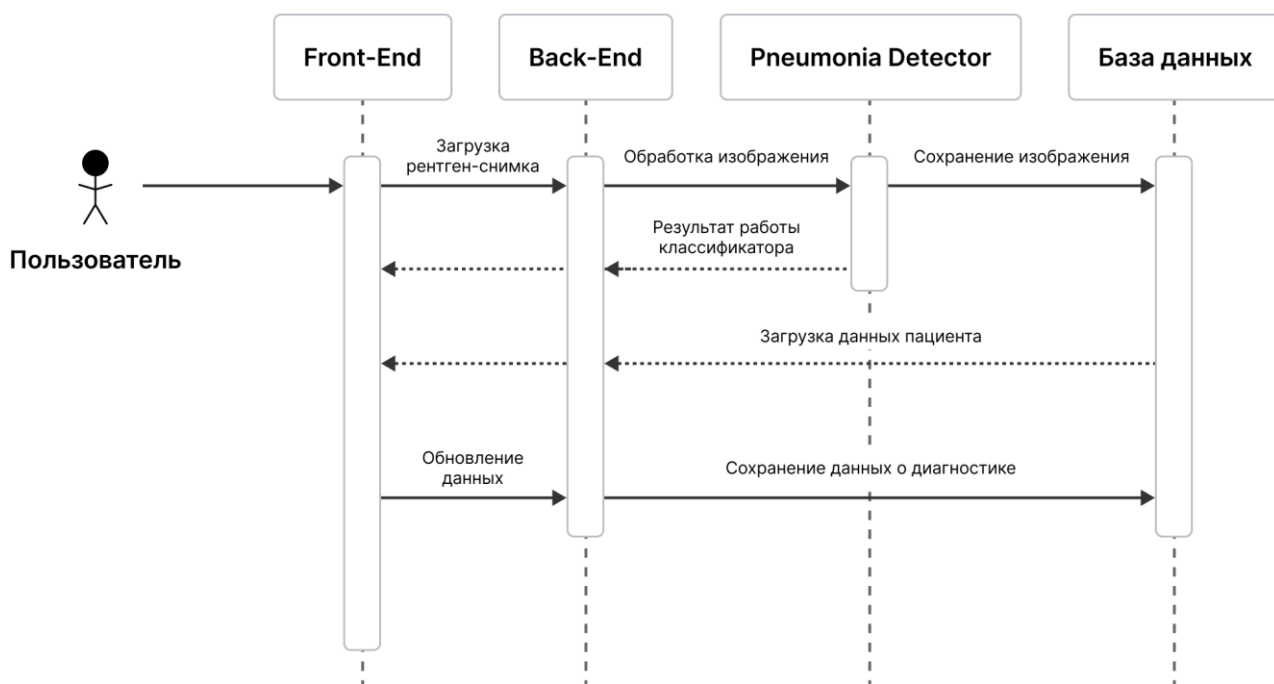


Рисунок 6 – Диаграмма последовательности работы приложения для обнаружения патологий по рентген-снимкам

Как видно из представленной диаграммы, пользователь взаимодействует с приложением через Front-End часть. В свою очередь, Front-End часть реализует

основную логику и операции с базой данных, вызовы модуля для диагностики пневмонии и отдает конечный ответ пользователю средствами REST API. Модуль для обработки изображений, классификации и, как следствие, диагностики представляет из себя отдельный процесс, который запускается при обращении к нему со стороны сервера. После обработки изображения и работы классификатора, полученный результат отправляется обратно на клиентскую часть для отображения вместе с данными о пациенте, взятыми из базы данных. При необходимости внести какие-либо изменения в данные о посещении пациентом специалиста это легко можно сделать на стороне клиента, а после сохранить данные о приёме и диагностике в базу данных.

Основываясь на поставленных задачах, а также особенностях работы классификатора, для разработки приложения был выбран следующий стек технологий:

- React.js для разработки Front-End части;
- язык Python совместно с библиотекой Flask, преимуществом которой можно отметить удобство использования в проектах с небольшим функционалом;
- поскольку данные о посещении пациентом врача-радиолога представляют собой записи в «медицинскую карту» будет логично использовать документно-ориентированную систему управления базами данных MongoDB.



```
1 {
2   first_name: "Jake",
3   last_name: "Brown",
4   sex: "Male",
5   birthday: "04.07.1991",
6   records: [
7     { date: "24.05.2020",
8       anamnesis: "Cough, sputum production, formation of rales, dyspnea",
9       diagnosis: "Bronchitis",
10      recommendations: "Chest X-Ray" },
11     { date: "04.06.2020",
12       scan: "/images/04062020_113442.png",
13       diagnosis: "Pneumonia",
14       diagnosis_likelihood: 0.84 }
15   ]
16 }
```

Рисунок 7 – Пример медицинской карты пациента из базы данных

Как можно заметить приложение является не очень сложным по своей архитектуре, поскольку фокусируется на конкретной задаче – диагностике пневмонии. За счёт модульности инструментов диагностики возможно дальнейшее увеличение функционала приложения без ущерба усложнения его архитектуры. К тому же, если добавить возможность входа в систему, используя учётные записи, принадлежащие различным медицинским учреждениям, то систему можно будет легко масштабировать как в масштабах города, так и страны и всего мира.

3.2 Проектирование пользовательского интерфейса приложения

Простой и понятный пользовательский интерфейс является ключом к повышению скорости выполнения любой задачи, и задача диагностики патологий не является исключением. Поскольку пользоваться приложением врачи будут каждый день, необходимо максимально направить фокус на работу с данными пациента и исключить любую нерелевантную информацию.

Базовый сценарий использования приложения представлен на рисунке 8. Руководствуясь им можно спроектировать интерфейс, который позволит выполнять эти действия.

Главная рабочая зона приложения представляет собой список медицинских записей пациента. Непосредственно самого пациента можно выбрать, перейдя в «Меню навигации» → «Пациенты» → выбрать пациента, после этого будет возможно добавить новые записи в медицинскую карты пациента, прикрепляя рентгенограмму или же просто внося необходимые данные о диагнозе, если рентгенограмма не проводилась или не требовалась для конкретного случая.

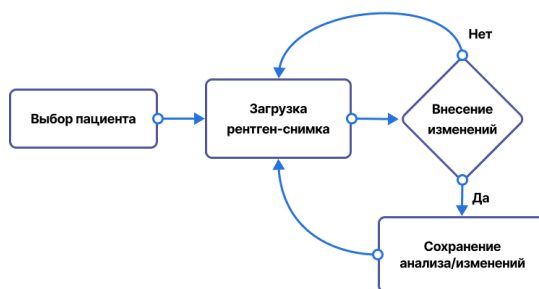


Рисунок 8 – Базовый сценарий работы с приложением

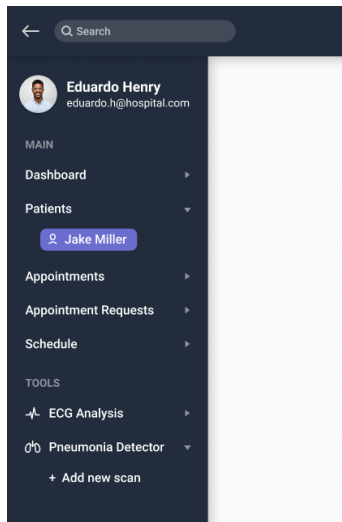


Рисунок 9 – Меню навигации

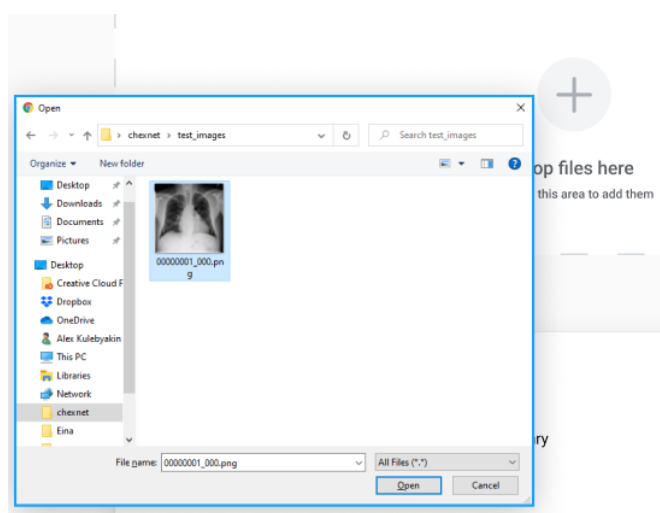
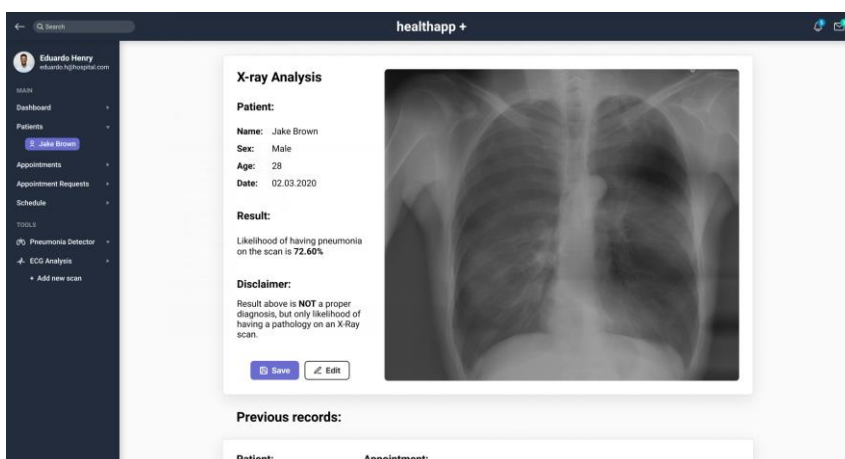


Рисунок 10 – Процесс добавления рентгенограммы в систему



(a)



(б)

Рисунок 11 – Интерфейс приложения: (а) при просмотре результата, (б) при наведении на рентгенограмму

На рисунке 11 показан пример работы с загруженным в систему снимком – специалисту предлагается вероятный диагноз и возможность либо согласиться с ним, сохранив результаты обследования, либо же отредактировать и после – сохранить. При наведении на рентгенограмму отобразится тепловая карта с вероятным расположением патологии на снимке.

Как можно заметить из рисунков 9-11 приложение обладает простым интерфейсом и позволяет легко работать с рентген-снимками пациентов, поэтому можно считать, что свои задачи приложение выполняет максимально эффективно.

Результаты

В разработке конечного приложения применялись следующие инструменты и технологии: React.js, Python, Flask, MongoDB, PyTorch.

С помощью созданного приложения возможно быстро и точно диагностировать патологии органов грудной клетки, а при необходимости внести изменения в диагностику или провести дополнительные тесты.

Несмотря на то, что в рамках текущего законодательства невозможно использовать систему без проверки результата квалифицированным специалистом, она всё же может быть использована для помощи врачам при постановке диагноза.

В будущем планируется протестировать систему в отделении лучевой диагностики СибГМУ и других медицинских учреждениях г. Томска.

4. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

4.1 Предпроектный анализ

В рамках данной работы разрабатывается информационно-аналитическая система для автоматизированной диагностики патологий органов грудной клетки. Данная система может быть использована в медицинских учреждениях для поддержки принятия решения врача.

Потенциальными потребителями подобных систем являются предприятия здравоохранения: поликлиники, больницы. Такие системы помогут врачу быстрее и, иногда, более корректно принять решение о способе лечения данного пациента. Другими потребителями можно считать представителей научного сообщества, заинтересованных в улучшении качества и точности диагностики, а также расширении функционала системы.

Целью данного раздела является анализ совокупности факторов, которые определяют коммерческую привлекательность разработки, ее перспективность и успешность. Основными задачами является оценка перспективности разработки, ее готовности к коммерциализации, выявление потенциальных угроз, а также расчет стоимости и составление графика проведения работ.

4.1.1 Технология QuaD

Технология QuaD представляет собой гибкий инструмент измерения характеристик, описывающих качество новой разработки и её перспективность на рынке и позволяющие принимать решение целесообразности вложения денежных средств в научно-исследовательский проект. В основе технологии лежит нахождение средневзвешенной величины показателей. Оценочная карта представлена в таблице 2.

Таблица 2 – Оценочная карта технологии QuaD

Критерии оценки	Вес критерия	Баллы	Максимальный балл	Относительное значение (3/4)	Средневзвешенное значение (5x2)
1	2	3	4	5	6
Производительность	0,07	75	100	0,75	0,0525
Отказоустойчивость	0,08	90	100	0,90	0,072
Унифицированность	0,05	60	100	0,60	0,03
Безопасность	0,1	70	100	0,70	0,07
Пользовательский интерфейс	0,12	95	100	0,95	0,114
Простота пользования	0,11	90	100	0,90	0,099
Функциональная мощность	0,1	75	100	0,75	0,075
Масштабируемость	0,1	80	100	0,80	0,08
Конкурентоспособность продукта	0,08	70	100	0,70	0,056
Перспективность рынка	0,07	85	100	0,85	0,0595
Цена	0,05	70	100	0,70	0,035
Финансовая эффективность научной разработки	0,07	85	100	0,85	0,0595
Итого	1				0,8025

Оценка качества и перспективности по технологии QuaD определяется по формуле:

$$P_{cp} = \sum P_i \cdot 100 = 0,8025 \cdot 100 = 80,25, \quad (1)$$

где P_{cp} – средневзвешенное значение показателей качества и перспективности научной разработки; P_i – средневзвешенное значение показателя.

Значение P_{cp} позволяет говорить о перспективах разработки и качестве проведенного исследования. Если значение показателя P_{cp} получилось от 100 до 80, то такая разработка считается перспективной. Если от 79 до 60 – то перспективность выше среднего. Из таблицы можно сделать вывод, что разработку можно считать перспективной.

4.1.2 SWOT-анализ

SWOT-анализ разработанного протокола представляет собой двухэтапный комплексный анализ разработки. Результаты анализа представлены в таблице 3.

Таблица 3 – Матрица SWOT разработки

	Сильные стороны разработки: S1. Невысокая стоимость. S2. Простота использования. S3. Скорость. S4. Возможность расширения функционала. S5. Перспективы для улучшения качества диагностики.	Слабые стороны разработки: W1. Техническая сложность. W2. Отсутствие финансирования для разработки. W3. Слабый функционал в момент запуска на рынок. W4. Отсутствие репутации на рынке. W5. Недостаточно высокая точность диагностики.
Возможности: O1. Сокращение финансирования системы здравоохранения, поэтому есть потребность в автоматизации. O2. Востребованность инструментов автоматизации работы с медицинскими данными. O3. Привлечение специалистов со всего мира для расширения функционала и качества диагностики.	Данная разработка является востребованным, гибким и масштабируемым средством автоматизации диагностики патологий по рентген-снимкам грудной клетки. Модульность инструментов для диагностики позволит легко расширить функционал системы.	Несмотря на спрос на систему, её функционал на начальном этапе может не удовлетворить потенциальных покупателей, а снижения финансирования может также негативно сказаться на процессе разработки.
Угрозы: T1. Отсутствие спроса на продукт. T2. Запрет на использования системы в медицинских учреждениях T3. Несвоевременное финансовое обеспечение со стороны государства и медицинских учреждений.	Наступление сценариев большинства угроз маловероятно. Требования крупных компаний могут превышать функционал системы.	Из-за высоких требований к качеству и точности диагностики, может случиться ситуация, когда улучшение этих параметров будет стоить крайне дорого.

Необходимо отметить, что особенностью данного решения является его скорость, простота использования и модульность. Расширение функционала со

временем позволит ещё больше автоматизировать диагностику пациентов. Поддержание стабильной ценовой политики и оказание услуг по настройке под конкретные медицинские учреждения будет напрямую способствовать конкурентоспособности разрабатываемого решения.

4.1.3 Оценка готовности научно-исследовательского проекта к коммерциализации

Одной из важных задач в ходе выполнения данного раздела является оценка готовности разработки к коммерциализации. Оцениваемыми параметрами являются как научная, так и коммерческая составляющая. Таблица 4 представляет собой бланк оценки степени готовности разработки к коммерциализации.

Таблица 4 – Бланк оценки степени готовности разработки к коммерциализации

№ п/п	Наименование	Степень проработанности разработки	Уровень имеющихся знаний у разработчика
1.	Определен имеющийся научно-технический задел	4	4
2.	Определены перспективные направления коммерциализации научно-технического задела	4	5
3.	Определены отрасли и технологии (товары, услуги) для предложения на рынке	4	3
4.	Определена товарная форма научно-технического задела для представления на рынок	3	4
5.	Определены авторы и осуществлена охрана их прав	3	3
6.	Проведена оценка стоимости интеллектуальной собственности	3	3
7.	Проведены маркетинговые исследования рынков сбыта	2	3
8.	Разработан бизнес-план коммерциализации научной разработки	1	1
9.	Определены пути продвижения научной разработки на рынок	3	3
10.	Разработана стратегия (форма) реализации научной разработки	5	4
11.	Проработаны вопросы международного сотрудничества и выхода на зарубежный рынок	3	4

12.	Проработаны вопросы использования услуг инфраструктуры поддержки, получения льгот	1	1
13.	Проработаны вопросы финансирования коммерциализации научной разработки	1	2
14.	Имеется команда для коммерциализации научной разработки	2	3
15.	Проработан механизм реализации разработки	5	5
	Итого:	44 / 75	48 / 75

Из таблицы можно сделать вывод о том, что перспективность разработки является средней и для дальнейшей коммерциализации желательно вовлечение сторонних специалистов в этой области.

Успех продвижения товара на рынок во многом зависит от правильности выбора метода коммерциализации. Для данного разработки наиболее приемлемым методом коммерциализации является предоставление заказчику на основе договора доступа к сервису. К тому же, в процессе использования программного продукта потребуется его доработка и адаптация к конкретной среде использования, а данный метод коммерциализации это позволяет.

4.2 Инициация разработки

В рамках инициации разработки формулируются цели и ожидаемые результаты работы. Также определяются заинтересованные стороны разработки и возможные ограничения. Заинтересованные в данной разработке стороны представлены в таблице 5.

Таблица 5 – Заинтересованные стороны разработки

Заинтересованные стороны	Ожидания заинтересованных сторон
Научный руководитель, инженер	Разработать и внедрить в медицинских учреждениях г. Томска систему автоматизированной диагностики патологий органов грудной клетки
Медицинский персонал	Уменьшение времени работы с пациентами, благодаря автоматизации процесса работы с медицинскими данными
Пациенты	Сокращение ожидания результатов диагностики за счёт использования системы
Научное сообщество	Возможность внедрить свои исследования и разработки

Цели и результат проекта отображены в таблице 6.

Таблица 6 – Цели и результат разработки

Цель проекта:	Разработать алгоритмическое и программное обеспечение для диагностики патологий органов дыхания по рентген-снимкам пациента
Ожидаемые результаты:	1. Формализованное описание системы 2. Программная реализация системы
Критерии приемки результата проекта:	Прохождение функционального тестирования решения на тестовых данных
Требования к результату проекта:	1. Выполнены все пункты технического задания 2. Разработанный функционал полностью соответствует проектным решениям

В таблице 7 представлена рабочая группа разработки, определена роль и основные функции каждого участника в разработке.

Таблица 7 – Рабочая группа разработки

№ п/п	ФИО, основное место работы, должность	Роль в разработке	Функции	Трудозатраты, час.
1.	Аксёнов Сергей Владимирович, ТПУ ОИТ ИШИТР, доцент	Научный руководитель	Утверждение основных разделов, выдача заданий к исполнению, координирование деятельности исполнителя	29
2.	Кулебякин Алексей Сергеевич, ТПУ ОИТ ИШИТР, магистрант гр. 8ПМ8И	Исполнитель	Исполнение поставленных задач	228
Итого:				257

Данный раздел отражает тот факт, что выполняемая работа имеет довольно большой объем. Заинтересованные стороны проекта ожидают достаточно высококачественные результаты, которые необходимо достичь исполнителю.

4.3 Планирование управления разработкой

4.3.1 План разработки

Чтобы отразить ключевые события по ведению разработки, необходимо составить календарный план. План представлен в таблице 8.

Таблица 8 – Календарный план разработки

Код работы	Название	Длительность, дни	Дата начала работ	Дата окончания работ	Состав участников
1	Выбор научного руководителя	6	15.02.20	21.02.20	Кулебякин А.С.
2	Составление и утверждение темы	2	22.02.20	25.02.20	Аксёнов С.В.
3	Составление календарного плана-графика выполнения работы	4	26.02.20	29.02.20	Кулебякин А.С. Аксёнов С.В.
4	Подбор и изучение литературы по теме магистерской работы	11	01.03.20	14.03.20	Кулебякин А.С.
5	Анализ предметной области	6	15.03.20	21.03.20	Кулебякин А.С.
6	Поиск исходных данных	3	22.03.20	25.03.20	Кулебякин А.С.
7	Построение классификатора	13	26.03.20	09.04.20	Кулебякин А.С.
8	Проектирование пользовательского интерфейса	18	10.04.20	30.04.20	Кулебякин А.С.
9	Согласование выполненной работы с научным руководителем	5	01.05.20	08.05.20	Кулебякин А.С. Аксёнов С.В.
10	Выполнение других частей работы (фин. менеджмент, соц. ответственность)	14	09.05.20	26.05.20	Кулебякин А.С.
11	Подведение итогов, оформление работы	4	27.05.20	31.05.20	Кулебякин А.С. Аксёнов С.В.
Итого:		86	15.02.20	31.05.20	

4.3.2 Разработка графика проведения разработки

В рамках планирования научно-исследовательского проекта построим календарный и линейный графики. Календарный план проекта представлен на рисунке 12.

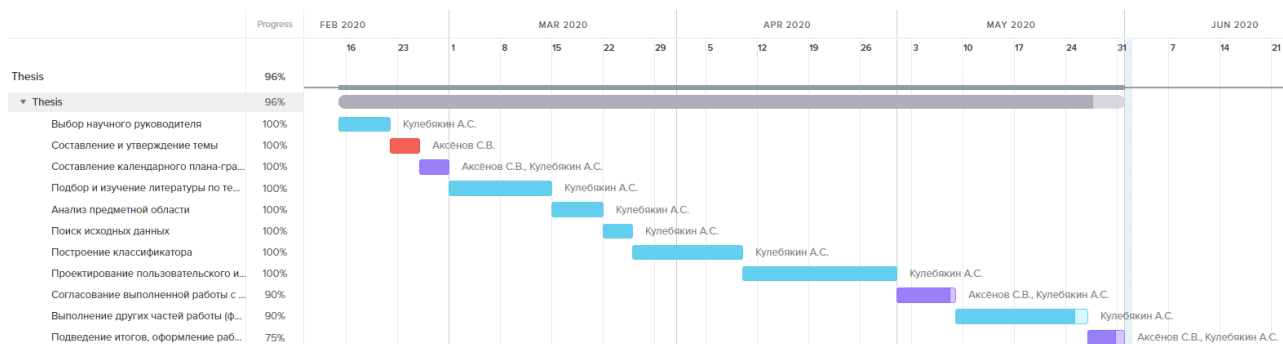


Рисунок 12 – Диаграмма Ганта

Из диаграммы Ганта наглядно видны границы этапов научно-исследовательского проекта. Длительность его выполнения 3,5 месяца.

4.4 Бюджет разработки

Бюджет научно-исследовательского проекта должен быть основан на достоверном отображении всех видов расходов, связанных выполнением проекта. В процессе формирования бюджета разработки используется следующая группировка затрат по статьям:

1. Материальные затраты;
2. Амортизационные отчисления;
3. Основная заработная плата исполнителей;
4. Дополнительная заработная плата исполнителей;
5. Отчисления во внебюджетные фонды (страховые отчисления);
6. Накладные расходы.

4.4.1 Расчет материальных затрат

Для проведения исследования какие-либо специальные материалы и комплектующие не приобретались. Единая сумма на канцелярские принадлежности составляет 2000 рублей.

4.4.2 Амортизационные затраты

Для проведения исследования специальное оборудование не покупалось, поэтому при расчете затрат учитывается только амортизация. Первоначальная стоимость ПК магистранта, используемого для проведения исследований, составляет 50000 рублей. Срок полезного использования данной машины – 3

года, из которых 4 месяцев (15.02.2020 – 15.06.2020) машина использовалась для написания и защиты ВКР.

Норма амортизации:

$$A_n = \frac{1}{n} \cdot 100\% = \frac{1}{3} \cdot 100\% = 33,33\% \quad (2)$$

Годовые амортизационные отчисления:

$$A_г = 50000 \cdot 0,33 = 16500 \text{ рублей} \quad (3)$$

Ежемесячные амортизационные отчисления:

$$A_m = \frac{16500}{12} = 1375 \text{ рублей} \quad (4)$$

Итоговая сумма амортизации составит $4 * 1375 = 5500$ рублей

4.4.3 Заработная плата исполнителей

Статья включает основную заработную плату работников, непосредственно занятых выполнением НИИ, (включая премии, доплаты) и дополнительную заработную плату. Она рассчитывается по формуле:

$$Z_{зп} = Z_{осн.} + Z_{доп.}, \quad (5)$$

где $Z_{осн.}$ – основная заработная плата; $Z_{доп.}$ – дополнительная заработная плата (12-20 % от $Z_{осн.}$).

Среднедневная заработная плата рассчитывается по формуле:

$$Z_{дн.} = \frac{Z_M \cdot M}{F_D}, \quad (6)$$

где Z_M – месячный должностной оклад работника, руб.; М – количество месяцев работы без отпуска в течение года; F_D – действительный годовой фонд рабочего времени научно- технического персонала, раб. дн.

Таблица 9 - Баланс рабочего времени (для 6-дневной недели)

Показатели рабочего времени	Дни
Календарное число дней	365
Количество нерабочих дней (выходные, праздничные дни)	66
Потери рабочего времени (отпуск, невыходы по болезни)	56
Действительный годовой фонд рабочего времени	243

Для расчета основной заработной платы инженера берем оклад, равный окладу 21760 руб. Для расчета основной заработной платы руководителя в расчет возьмем оклад, равный 35120 руб.

Месячный оклад работника рассчитывается по формуле:

$$Z_M = Z_{TC} \cdot (1 + k_{np} + k_d) \cdot k_p, \quad (7)$$

где: Z_{TC} – заработная плата по тарифной ставке, руб.; k_{np} – премиальный коэффициент, равный 0,3 (т.е. 30% от Z_{TC}); k_d – коэффициент доплат и надбавок составляет примерно 0,2 – 0,5 (в НИИ и на промышленных предприятиях – за расширение сфер обслуживания, за профессиональное мастерство, за вредные условия: 15-20% от Z_{TC}); k_p – районный коэффициент, равный 1,3 (для Томска).

Расчёт основной заработной платы приведён в таблице 10.

Таблица 10 – Расчет основной заработной платы

Исполнители	k_{np}	k_d	k_p	$Z_{дон.}$ руб.	T_p , раб. дн.	$Z_{осн.}$ руб.
Инженер	0	0	1,3	1 210,7	84	101 698,8
Научный руководитель	0,3	0,2	1,3	2 931	15	43 965

В дополнительную заработную плату входят суммы выплат, предусмотренные трудовым кодексом, например, оплата ежегодных и дополнительных отпусков. Запланируем дополнительную заработную плату в размере 15 % от основной заработной платы исполнителей. В таблице 11 представлен расчет затрат на заработную плату исполнителей.

Таблица 11 – Затраты на заработную плату без отчислений

Исполнители	$Z_{осн.}$ руб.	$Z_{дон.}$ руб.	Z_n руб.
Инженер	101 698,8	15 254,82	116 953,62
Научный руководитель	43 965	6 594,75	50 559,75
Итого:	145 663,8	21 849,57	167 513,37

4.4.4 Отчисления во внебюджетные фонды

Общие тарифы страховых взносов в 2020 году в ИФНС:

- 22 % – на пенсионное страхование;
- 2,9 % – страхование по временной нетрудоспособности;
- 5,1 % – медицинское страхование;

Величина отчислений во внебюджетные фонды определяется исходя из формулы:

$$Z_{внеб.} = k_{внеб.} \cdot (Z_{осн.} + Z_{доп.}), \quad (8)$$

где $k_{внеб.}$ – коэффициент отчислений на уплату во внебюджетные фонды (пенсионный фонд, фонд обязательного медицинского страхования и пр.).

Таким образом, с учетом результатов расчета зарплат на заработную плату величина отчислений во внебюджетные фонды составляет:

$$Z_{внеб.} = 0,3 \cdot 167\,513,37 = 50\,254,01 \text{ руб.} \quad (9)$$

4.4.5 Накладные расходы

При выполнении проекта могут возникнуть косвенные издержки – накладные расходы, возникающие дополнительно к основным затратам, например, на консультационные услуги, оплату коммунальных услуг, расходы на услуги связи (телефон, интернет) и так далее. Расчет накладных расходов определяется по формуле:

$$Z_{нак.} = \sum St \cdot k_{нр}, \quad (10)$$

где $k_{нр}$ – коэффициент, учитывающий накладные расходы; St – затраты по статьям накладных расходов.

Величину коэффициента накладных расходов можно взять в размере 10%.

$$Z_{нак.} = \frac{(2000+5500+167\,513,37+50\,254,01)}{10} = 22\,526,74 \text{ руб.} \quad (11)$$

4.4.6 Формирование бюджета

Рассчитанная величина затрат научно-исследовательской работы является основой для формирования бюджета затрат проекта. Определение

бюджета затрат на научно-исследовательский проект по каждому варианту исполнения приведен в таблице 12.

Таблица 12 – Бюджет разработки научно-исследовательского проекта

Наименование	Сумма, руб.	Удельный вес, %
Материальные затраты	2 000	0,81
Затраты на специальное оборудование	5 500	2,22
Затраты на основную заработную плату	145 663,8	58,78
Затраты на дополнительную заработную плату	21 849,57	8,82
Страховые взносы	50 254,01	20,28
Накладные расходы	22 526,74	9,09
Общий бюджет	247 794,12	100

Общий бюджет разработки составляет 247794,12 рублей. Данный бюджет является приемлемым для перспективной разработки.

4.4.7 Риски при работе над проектом

Проведение любого научно-исследовательского проекта сопряжено с возникновением различных рисков. Предварительное определение рисков помогает своевременному принятию мер по предотвращению возникновения угроз или минимизации их последствий.

Таблица 13 – Реестр рисков

Риск	Потенциальное воздействие	Вероятность (1-5)	Влияние риска (1-5)	Уровень риска	Способы смягчения риска	Условия наступления
Кадровый риск	Отсутствие исполнителей проекта	4	5	Высокий	Повышение мотивации исполнителей проекта	Потеря интереса к деятельности проекта
Технический риск	Потеря файлов проекта; Нехватка мощностей	3	5	Высокий	Резервное копирование; Покупка доп. мощности	Отказ оборудования
Доступ к данным	Отсутствие данных для работы системы	1	5	Высокий	Заключение договора на поставку данных	Закрытие доступа к рентген-снимкам

Из анализа реестра рисков можно заключить, что первым и вторым типами рисков обладает практически каждый проект. Риск же потери доступа к данным во время выполнения данного научно-исследовательского проекта существенен для реализации, однако маловероятен.

4.5 Определение потенциального эффекта разработки

Результаты оценки востребованности разработки можно считать положительными, поскольку, во-первых, были выявлены потенциальные потребители настоящего решения. Во-вторых, в РФ пока не запущен подобный сервис в коммерческое пользование, а значит не придётся бороться с конкурентами за потребителей. В-третьих, проведенный SWOT анализ показал перспективность разработки. Расширение функционала, оказание услуг по настройке или консультированию под каждое конкретное медицинское учреждение совместно с поддержанием стабильной ценовой политики позволит сохранять свою конкурентоспособность при выходе других игроков на рынок.

Кроме того, в данной главе был разработан план и сформирован бюджет технического решения. Продолжительность проекта составила 86 рабочих дня, а общий бюджет затрат составил 247794,12 рублей – это абсолютно приемлемые показатели для создания перспективного продукта. Разработанный реестр рисков отражает потенциальные пути преодоления внешних и внутренних рисков и способствует успешной реализации проекта, а также его дальнейшее существование.

5. Социальная ответственность

Целью дипломной работы является разработка информационно-аналитической системы для автоматизированной диагностики патологий органов грудной клетки. Система позволит врачам получить предварительный диагноз по рентген-снимку грудной клетки пациента и существенно сократить время обследования. Создание Системы было инициировано Центром трансляции медицинских технологий СибГМУ совместно с ТПУ. [1]

В данном разделе будут рассмотрены опасные и вредные факторы, оказывающие влияние на производственную деятельность инженера-разработчика ПО, воздействия объекта исследования на окружающую среду, правовые и организационные вопросы, а также мероприятия в чрезвычайных ситуациях. Проектирование, разработка и тестирование системы производилось в аудитории Томского политехнического университета. Основным средством работы является персональный компьютер с выходом в сеть Интернет.

5.1 Правовые и организационные вопросы обеспечения безопасности

Нормативное регулирование охраны труда при осуществлении трудовой деятельности за компьютерами осуществляется посредством документов, перечисленных в списке литературы.

Однако, непосредственно при выполнении работы сотрудником-инженером нормальные условия для выполнения работниками норм выработки обязан обеспечить работодатель согласно "Трудовому кодексу Российской Федерации" от 30.12.2001 N 197-ФЗ. [2] К таким условиям, в частности, относятся:

- исправное состояние помещений, сооружений, машин, технологической оснастки и оборудования;
- своевременное обеспечение технической и иной необходимой для работы документацией;
- надлежащее качество материалов, инструментов, иных средств и предметов, необходимых для выполнения работы, их своевременное предоставление работнику;

- условия труда, соответствующие требованиям охраны труда и безопасности производства;
- создание и функционирование системы управления охраной труда;
- соответствующие требованиям охраны труда условия труда на каждом рабочем месте;
- режим труда и отдыха работников в соответствии с трудовым законодательством и иными нормативными правовыми актами, содержащими нормы трудового права;
- обучение безопасным методам и приемам выполнения работ и оказанию первой помощи пострадавшим на производстве, проведение инструктажа по охране труда, стажировки на рабочем месте и проверки знания требований охраны труда;
- недопущение к работе лиц, не прошедших в установленном порядке обучение и инструктаж по охране труда, стажировку и проверку знаний требований охраны труда;
- организацию контроля за состоянием условий труда на рабочих местах, а также за правильностью применения работниками средств индивидуальной и коллективной защиты;
- информирование работников об условиях и охране труда на рабочих местах, о риске повреждения здоровья, предоставляемых им гарантиях, полагающихся им компенсациях и средствах индивидуальной защиты;
- принятие мер по предотвращению аварийных ситуаций, сохранению жизни и здоровья работников при возникновении таких ситуаций, в том числе по оказанию пострадавшим первой помощи;
- санитарно-бытовое обслуживание и медицинское обеспечение работников в соответствии с требованиями охраны труда, а также доставку работников, заболевших на рабочем месте, в медицинскую организацию в случае необходимости оказания им неотложной медицинской помощи.

Нормативные положения СанПиНа 2.2.2/2.4.1340-03 предъявляют определенные требования к оснащению рабочего места, предусматривающего длительную работу за ПК [3]:

Таблица 14 – Нормы оборудования рабочих мест

Высота перегородок, разделяющих рабочие места	Не менее 1,5 метров
Ширина рабочего стола	От 80 до 140 см
Глубина рабочего стола	От 80 до 100 см
Высота рабочего стола	7,25 см
Расстояние от глаз до монитора	От 60 до 70 см
Расстояние клавиатуры от края стола	От 10 до 30 см
Сидение	Должно позволять регулировку по высоте, повороту и углу наклона спинки (регулировки должны быть независимыми друг от друга)
Подставка для ног	Ширина – от 30 см, глубина – от 40 см, с углом наклона до 20 градусов

5.2 Производственная безопасность

5.2.1 Анализ выявленных вредных и опасных факторов

В данном подразделе анализируются вредные и опасные факторы, которые могут возникать при проведении исследований в лаборатории, при разработке или эксплуатации проектируемого решения.

Согласно ГОСТ 12.0.003-2015 [4] в таблице 15 представлены возможные вредные и опасные факторы. Работа по разработке программного обеспечения делится на три основных этапа: проектирование, разработка и эксплуатация.

Таблица 15 – Вредные и опасные факторы при работе за компьютером

Факторы (ГОСТ 12.0.003-2015)	Этапы работ			Нормативные документы
	Проектирование	Разработка	Эксплуатация	
1. Отклонение показателей микроклимата	+	+	+	СанПиН 2.2.4.548–96 [12] СанПиН 2.2.2/2.4.1340-03 [10]
2. Превышение уровня шума	+	+	+	ГОСТ 12.1.003-2014 [13]
3. Отсутствие или недостаток естественного света	+	+	+	СП 52.13330.2011 [14]
4. Недостаточная освещенность рабочей зоны	+	+	+	СП 52.13330.2011 [14]
5. Повышенное значение напряжения в электрической цепи, замыкание которой может произойти через тело человека	+	+	+	ГОСТ 12.1.019-2017 [16]
6. Умственное перенапряжение	+	+	+	ТОИ Р-45-084-01 [17]

5.2.2 Обоснование мероприятий по снижению воздействия

Теоретически, объект исследования (система диагностики патологий) при определенных условиях, не предусмотренных разработчиками, способен привести к перегрузкам в ПЭВМ и вызвать определенные последствия. Поскольку разработка не является осязаемым объектом и неотделима от ПЭВМ, вредные и опасные факторы, которые могут быть прямо или косвенно отнесены к разработке, относятся и к рабочему месту.

5.2.2.1 Отклонение показателей микроклимата

Показатели микроклимата должны обеспечивать сохранение теплового баланса человека с окружающей средой и поддержание оптимального или допустимого теплового состояния организма. Отклонение показателей от нормы в пределах допустимых величин могут вызвать локальные ощущения теплового дискомфорта и напряжение механизмов терморегуляции.

Согласно СанПиН 2.2.4.548-96 [12], санитарные правила устанавливают гигиенические требования к показателям микроклимата рабочих мест

производственных помещений с учетом интенсивности энергозатрат работников, времени выполнения работы, периодов года и содержат требования к методам измерения и контроля микроклиматических условий.

Оптимальные микроклиматические условия установлены по критериям оптимального теплового и функционального состояния человека. Они обеспечивают общее и локальное ощущение теплового комфорта в течение 8-часовой рабочей смены при минимальном напряжении механизмов терморегуляции, не вызывают отклонений в состоянии здоровья, создают предпосылки для высокого уровня работоспособности и являются предпочтительными на рабочих местах.

Категория работ определена, как наименее энергозатратная, т.е. не сопровождающаяся какой-либо физической нагрузкой (сидячий вид деятельности, умственный труд) – 1а. Оптимальные величины показателей микроклимата представлены в таблице 16.

Таблица 16 – Оптимальные величины показателей микроклимата на рабочих местах

Период года	Категория работ	Температура воздуха, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Теплый	1а	22-24	60-40	0,1
Холодный	1а	23-25	60-40	0,1

Для определения температуры воздуха в помещении использовался настенный термометр. Показания температуры воздуха составляют около 23 °С. Отсюда можно сделать вывод, что микроклимат на рабочем месте является оптимальным для работы в данное время года.

5.2.2.2 Превышение уровня шума

Превышения уровня шума является вредным фактором на рабочем месте. Постоянный шум, превышающий допустимые значения, не только воздействует на органы слуха, но и влияет на общее самочувствие работника, способствует ослаблению организма, а также снижает работоспособность.

Согласно ГОСТ 12.1.003-2014 [13], машины, которые в процессе работы могут производить шум, неблагоприятно воздействующий на работников, следует конструировать и изготавливать с учетом последних достижений технологии и принципов проектирования, позволяющих снизить излучаемый шум.

Наибольшим шумовым событием на рабочем месте для исследователя является ПЭВМ.

СанПиН 2.2.2/2.4.1340-03 регулирует допустимые значения уровней звукового давления в октавных полосах частот и уровня звука, создаваемого ПЭВМ. Данные представлены в таблице 17.

Таблица 17 – Допустимые значения уровней звукового давления в октавных полосах частот и уровня звука, создаваемого ПЭВМ

Уровни звукового давления в октавных полосах со среднегеометрическими частотами									Уровни звука в дБА
31,5 Гц	63 Гц	125 Гц	250 Гц	500 Гц	1000 Гц	2000 Гц	4000 Гц	8000 Гц	
86 дБ	71 дБ	61 дБ	54 дБ	49 дБ	45 дБ	42 дБ	40 дБ	38 дБ	50

В помещении был произведён замер уровня звука, он составил в среднем 47,2 дБА, что удовлетворяет данным условиям.

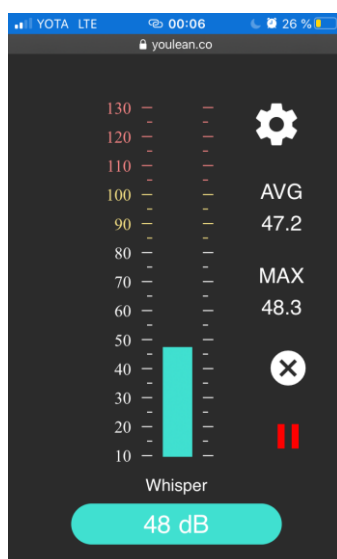


Рисунок 13 – Измерение уровня шума в помещении

5.2.2.3 Отсутствие или недостаток света

Отсутствие естественного света в темное время суток является естественным явлением, актуальным для г. Томска в течение рабочего дня в зимний период в силу короткого светового дня. Поэтому основным источником освещения на рабочем месте является искусственное освещение. Недостаточная освещенность рабочей поверхности может сказаться на здоровье работника и его производительности труда. Согласно таблице 18, освещенность на рабочем месте должна составлять 300 – 500 лк. [14]

Таблица 18 – Требования к освещению на рабочих местах, оборудованных ПК

Освещенность на рабочем месте	300 – 500 лк
Освещенность на экране ПК	Не выше 300 лк
Блики на экране	Не выше 40 кд/м ²
Прямая блеклость источника света	200 кд/м ²
Показатель ослепленности	Не более 20
Показатель дискомфорта	Не более 15
Отношение яркости:	
– между рабочими поверхностями	3:1 – 5:1
– между поверхностями стен и оборудования	10:1
Коэффициент пульсации:	Не более 5%

В помещении (рисунок 15) используется светодиодная лампа с параметрами, представленными на рисунке 14. Для того, чтобы рассчитать освещённость на рабочем месте необходимо узнать световой поток, испускаемый лампой, он составляет 4200 лм. [15]



Рисунок 14 – Параметры светодиодной лампы в помещении

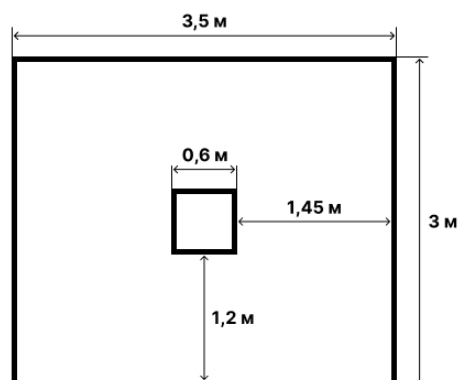


Рисунок 15 – План помещения и размещения светильников со светодиодными лампами

Рассчитать освещённость можно по формуле $E = \Phi / S$, где E – освещённость рабочего пространства, Φ – световой поток, S – площадь помещения. Получаем $E = 4200 \text{ лм} / (3,5 \text{ м} \cdot 3 \text{ м}) = 400 \text{ лк}$, что удовлетворяет условиям, приведенным в таблице 18.

5.2.2.4 Повышенное значение напряжения в электрической цепи

К опасным факторам на рабочем месте относится поражение электрическим током, поскольку исследователь взаимодействует с электрооборудованием.

Согласно ГОСТ 12.1.019-2017 ССБТ [16], во избежание поражения работника электрическим током следует соблюдать следующие базовые принципы:

1) Проводящие части, находящиеся под опасным рабочим, наведенным, остаточным напряжением, не должны быть доступными, а доступные проводящие части не должны находиться под опасным напряжением при нормальных условиях (при отсутствии повреждения), а также в случае единичного повреждения.

2) Защиту при нормальных условиях (защиту от прямого прикосновения) обеспечивают посредством основной защиты, а защиту при условиях единичного повреждения (защиту при косвенном прикосновении) обеспечивают посредством защиты при повреждении.

Степень опасного воздействия на человека электрического тока электрической дуги зависит от:

- 1) величины напряжения прикосновения, электрического сопротивления тела человека, силы тока, протекающей через него, а также величины падающей энергии электрической дуги;
- 2) рода (постоянный, переменный, выпрямленный) тока и частоты переменного электрического тока;
- 3) пути протекания тока через тело человека и площади контакта электрической дуги с поверхностью тела человека;
- 4) продолжительности воздействия электрического тока и электрической дуги на организм человека;
- 5) индивидуальных особенностей организма человека;
- 6) условий внешней среды.

5.2.2.5 Умственное перенапряжение

Работая за ПЭВМ, работник также находится под влиянием еще одного вредного производственного фактора, нервно-психической перегрузки в виде умственного перенапряжения.

Согласно ТООИ Р-45-084-01 [17], виды трудовой деятельности разделяются на 3 группы: группа А - работа по считыванию информации с экрана компьютера с предварительным запросом; группа Б - работа по вводу информации; группа В - творческая работа в режиме диалога с компьютером. При выполнении в течение рабочей смены работ, относящихся к различным видам трудовой деятельности, за основную работу с компьютером следует принимать такую, которая занимает не менее 50% времени в течение рабочей смены или рабочего дня. Уровень нагрузки представлен в таблице 6.

Таблица 19 – Уровень нагрузки за рабочую смену при работе с компьютером

Категория работ	Уровень нагрузки за рабочую смену при работе с компьютером		
	группа А, количество знаков	группа Б, количество знаков	группа В, час.
III	До 60000	До 40000	До 6,0

5.3 Экологическая безопасность

Деятельность по разработке ПО не связана с производством, поэтому влияние на окружающую среду минимально. При работе над проектом применяются рекомендации по минимизации влияния на окружающую среду.

С учетом того, что наиболее значительную часть твердых бытовых отходов (до 40 % в развитых странах) составляет бумага для печати — актуальным является осуществление её утилизации для повторного использования в промышленном производстве, что является наиболее экономически эффективным способом обращения с отходами согласно ГОСТ Р 55090-2012 «Ресурсосбережение. Обращение с отходами. Рекомендации по утилизации отходов бумаги» [18].

При завершении срока службы ПК их можно отнести к отходам электронной промышленности. Переработка таких отходов осуществляется разделением на однородные компоненты, химическим выделением пригодных для дальнейшего использования компонентов и направлением их для дальнейшего использования согласно ГОСТ Р 55102-2012 «Ресурсосбережение. Обращение с отходами. Руководство по безопасному сбору, хранению, транспортированию и разборке отработавшего электротехнического и электронного оборудования, за исключением ртульсодержащих устройств и приборов» [19].

Перечень элементов и содержащее их отработанное электротехническое и электронное оборудование, которые должны быть отдельно собраны при выводе отработавшего электротехнического и электронного оборудования из эксплуатации:

- конденсаторы, содержащие ПХБ;
- печатные платы с площадью поверхности больше 10 см²;
- картриджи;
- пластик;
- электронно-лучевые трубки;

- элементы отработавшего электротехнического и электронного оборудования;
- жидкокристаллические экраны (если необходимо, вместе с корпусом) с поверхностью более 100 см² и все экраны с подсветкой газоразрядными лампами;
- внешние электрические кабели; - элементы, содержащие огнеупорные керамические слои;
- конденсаторы, содержащие электролит (размер хотя бы одной из сторон конденсатора должен быть 25 мм или более).

Люминесцентные лампы относят к ртутьсодержащим отходам, и для их утилизации действует Постановление Правительства РФ от 03.09.2010 № 681 (ред. от 01.10.2013) «Об утверждении Правил обращения с отходами производства и потребления в части осветительных устройств, электрических ламп, ненадлежащие сбор, накопление, использование, обезвреживание, транспортирование и размещение которых может повлечь причинение вреда жизни, здоровью граждан, вреда животным, растениям и окружающей среде» [20]. Согласно постановлению, устанавливается порядок обращения с отходами производства и потребления в части осветительных устройств, электрических ламп, ненадлежащие сбор, накопление, использование, обезвреживание, транспортирование и размещение которых может повлечь причинение вреда жизни, здоровью граждан, вреда животным, растениям и окружающей среде.

Не допускается самостоятельное обезвреживание, использование, транспортирование и размещение отработанных ртутьсодержащих ламп потребителями отработанных ртутьсодержащих ламп, а также их накопление в местах, являющихся общим имуществом собственников помещений многоквартирного дома, за исключением размещения в местах первичного сбора и размещения и транспортирования до них.

Сбор отработанных ртутьсодержащих ламп у потребителей осуществляют специализированные организации.

Отходы, не подлежащие переработке и вторичному использованию, подлежат захоронению на полигонах.

5.4 Безопасность в чрезвычайных ситуациях

Возможными ЧС могут являться наводнения, удары молний, взрывы, эпидемии инфекционных заболеваний.

Наиболее вероятным ЧС является пожар. Источниками возгорания может стать электропроводка, внутренние работающие устройства ПК и другие электроприборы, находящиеся в помещении.

Поражающими факторами пожаров в помещении являются токсическое воздействие горючих материалов (в т.ч. отравление угарным газом), экстремальный нагрев среды, а также обломки и осколки при нарушении целостности конструкций здания [21].

Для минимизации возможности возникновения фактора пожара необходимо проводить пожарную профилактику. Пожарная профилактика представляет собой комплекс организационных и технических мероприятий, направленных на обеспечение безопасности людей, на предотвращении пожара, ограничение его распространения, а также создание условий для успешного тушения пожара. Для профилактики пожара чрезвычайно важна правильная оценка пожароопасности, определение опасных факторов и обоснование способов и средств пожар предупреждения и защиты.

Одно из условий обеспечения пожаробезопасности - ликвидация возможных источников воспламенения.

Обогревание помещения открытыми электронагревательными приборами могут привести к пожару, т.к. в помещении находятся бумажные документы и справочная литература. Следовательно, использование открытого нагревательного прибора неприемлемо.

5.4.1 Мероприятия по предотвращению ЧС и разработка порядка действий в случае возникновения ЧС

Согласно ГОСТ Р 22.3.03-94 [22], обеспечение безопасности людей в ЧС, обусловленных природными стихийными бедствиями, техногенными авариями

и катастрофами, а также применением современного оружия (военные ЧС) является общегосударственной задачей, обязательной для решения всеми территориальными, ведомственными и функциональными органами управления и регулирования, службами и формированиями, а также подсистемами, входящими в Российскую систему предупреждения и действий в чрезвычайных ситуациях (РСС).

Мероприятия по защите людей от источников ЧС должны планироваться в объемах, гарантирующих непревышение нормативного воздействия на них возможных поражающих факторов для расчетной ЧС.

Для защиты жизни и здоровья населения в ЧС следует применять следующие основные мероприятия гражданской обороны, являющиеся составной частью мероприятий РСЧС:

- 1) укрытие людей в приспособленных под нужды защиты населения помещениях производственных, общественных и жилых зданий, а также в специальных защитных сооружениях;
- 2) эвакуацию населения из зон ЧС;
- 3) использование средств индивидуальной защиты органов дыхания и кожных покровов;
- 4) проведение мероприятий медицинской защиты;
- 5) проведение аварийно-спасательных и других неотложных работ в зонах ЧС.

Мерами по предупреждению ЧС являются:

- 1) Соблюдение техники безопасности при работе с ПЭВМ. Использование только исправного оборудования.
- 2) Своевременное проведение ТО и ППР электроустановок согласно утвержденного графика и технических средств противопожарной защиты и пожаротушения.
- 3) Установка противопожарной сигнализации.
- 4) Своевременное проведение инструктажа рабочего персонала. В случае угрозы возникновения ЧС (пожара) необходимо вызвать противопожарную

службу, отключить электроэнергию и, следуя плану эвакуации, эвакуировать находящихся в помещении людей и покинуть помещение. В случае, если очаг возгорания является небольшим, и нет угрозы поражения электрическим током, можно использовать углекислотные огнетушители ОУ-5 высокого давления с зарядом жидкой двуокиси углерода, согласно ГОСТ 8050-85 [23]. Расположение огнетушителей отмечено на плане эвакуации людей при пожаре и других ЧС из помещений общежития №12, отмеченном на рисунке 16.



Рисунок 16 – План эвакуации при пожаре и других ЧС

5.5 Выводы по разделу

В результате проведенного анализа были выявлены вредные и опасные производственные факторы для компьютерной аудитории кафедры, и работы, связанной с проектом.

Для вредных и опасных факторов, из нормативных документов, были определены значения нормативных показателей, которые сравнивались со значениями, действующими при разработке проекта. Приведены рекомендации по улучшения условий труда и минимизации влияния на работника вредных и опасных факторов. Проведен анализ воздействия на окружающую среду и обозначены проблемы утилизации отходов. Определена категория помещения по пожароопасности, а также действия по минимизации риска возникновения пожара.

Также были рассмотрены правовые нормы трудового законодательства, применимые к условиям настоящего проекта, определены основные требования к организации рабочего места.

После проведенного анализа можно сделать вывод, что рабочее место соответствует всем нормативным требованиям производственной безопасности и охраны труда.

Заключение

В ходе данной работы был запрограммирован и протестирован алгоритм обнаружения патологий органов грудной клетки на рентгенограмме. Для взаимодействия с алгоритмом было спроектировано и разработано веб-приложение, позволяющее за несколько секунд получить рекомендацию при постановке диагноза врачом-радиологом.

Качество обнаружения на данном этапе пока не позволяет использовать систему без проверки результатов квалифицированным специалистом, но помогает сократить время и стоимость диагностики.

При последующей доработке как алгоритма, так и приложения, продукт можно будет легко вывести на рынок, а при увеличении спроса – масштабировать его на огромное количество медицинских учреждений и дополнять требуемым функционалом за счёт модульности как алгоритма, так и приложения.

В будущем планируется протестировать систему в отделении лучевой диагностики СибГМУ и других медицинских учреждениях г. Томска – это позволит дополнительно удостовериться в коммерческом потенциале разработанной системы.

Список достижений

1. Диплом I степени на конкурсе по решению бизнес-кейсов в сфере медицины «МедХакатон», г. Томск, сентябрь 2018 г.
2. Диплом I степени на хакатоне «Цифровые решения для нефтегазового сектора», г. Томск, октябрь 2018 г.
3. Повышенная государственная академическая стипендия II степени в номинации "За достижения в научно-исследовательской деятельности" по итогам осеннего семестра 2018/19 учебного года, приказ № 9752/с от 2 ноября 2018 г.
4. Стипендия программы FIRST+ для обучения в университете Аалто (Финляндия, г. Эспоо) в весеннем семестре 2018-2019 уч. года, ноябрь 2018
5. Стипендия Правительства Российской Федерации по приоритетным направлениям модернизации и технологического развития российской экономики, весенний семестр 2018-2019 уч. г., приказ №1612/с от 01.03.2019 г.
6. Повышенная государственная академическая стипендия III степени в номинации "За достижения в научно-исследовательской деятельности" по итогам весеннего семестра 2018/19 учебного года, приказ № 2194/с от 21 марта 2019 г.
7. Диплом за лучшее решение на международном хакатоне "Nordic Health Hackathon", Финляндия, г. Хельсинки, март 2019 г.
8. Повышенная государственная стипендия по научно-исследовательской деятельности, ОСЕНЬ 2019/2020 уч. года. Приказ от 03.10.2019 № 9236/с.
9. Диплом за 1 место в Конкурсе инновационных и предпринимательских проектов ТПУ 10К "Идеи для жизни" в номинации "Инновационные проекты в IT-сфере", г. Томск, декабрь 2019 г.
10. Стипендия Правительства РФ студентам ТПУ, обучающимся по специальностям или направлениям подготовки, соответствующим

приоритетным направлениям модернизации и технологического развития российской экономики на весенний семестр 2019-2020 учебного года. Приказ № 58-22/с от 27.02.2020, номер в приказе 4

11. Повышенная государственная стипендия по научно-исследовательской деятельности, ВЕСНА 2019/2020 уч. года. Приказ от 20.03.2020 № 80-5/с

Список используемых источников

1. ChestEye CAD. Automate healthy patient diagnosis and reporting. Oxipit.ai. [Электронный ресурс] URL: <https://oxipit.ai/products/chesteye/>
2. Automated Chest X-ray Interpretation – qXR v2.0. Qure.ai. [Электронный ресурс] URL: <http://qure.ai/qxr.html>
3. The Future of Artificial Intelligence in Healthcare. HealthTech. [Электронный ресурс] URL: <https://healthtechmagazine.net/article/2020/02/future-artificial-intelligence-healthcare>
4. The potential for artificial intelligence in healthcare. Future Healthcare Journal. NCBI. [Электронный ресурс] URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6616181/>
5. Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. (2018). Artificial intelligence in radiology. Nature reviews. Cancer, 18(8), 500–510. (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6268174/>)
6. G. Huang, Z. Liu and L. van der Maaten, “Densely Connected Convolutional Networks,” 2018.
7. Wang, Xiaosong, Peng, Yifan, Lu, Le, Lu, Zhiyong, Bagheri, Mohammadhadi, and Summers, Ronald M. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. arXiv preprint arXiv:1705.02315, 2017.
8. Центр трансляции медицинских технологий. – СибГМУ: [Электронный ресурс]. URL: https://ssmu.ru/ru/nauka/structura_nauka/centr_transl_jacii_medicinskikh_tekhnologij/ (Дата обращения: 20.05.2020)
9. Трудовой кодекс Российской Федерации от 30.12.2001 N 197-ФЗ (ред. от 01.04.2019)
10. СанПиН 2.2.2/2.4.1340-03 Гигиенические требования к персональным электронно-вычислительным машинам и организации работы (с изменениями на 21 июня 2016 года)

- 11.ГОСТ 12.0.003-2015 Система стандартов безопасности труда (ССБТ).
Опасные и вредные производственные факторы. Классификация
- 12.СанПиН 2.2.4.548-96 Гигиенические требования к микроклимату
производственных помещений
- 13.ГОСТ 12.1.003-2014 Система стандартов безопасности труда (ССБТ).
Шум. Общие требования безопасности
- 14.СП 52.13330.2011 Свод правил. Естественное и искусственное
освещение.
- 15.Светодиодный комплект для изготовления и переделки светильников.
ЦентрБиоСвет: [Электронный ресурс]. URL: https://www.xn--1001-t4dx6d.xn--p1ai/goods/39339975-svetodiodny_komplekt_dlya_izgotovleniya_i_peredelki_svetilnikov_affina_new (Дата обращения: 20.05.2020)
16. ГОСТ 12.1.019-2017 Система стандартов безопасности труда (ССБТ).
Электробезопасность. Общие требования и номенклатура видов
защиты
17. ТОИ Р-45-084-01 Типовая инструкция по охране труда при работе на
персональном компьютере
- 18.ГОСТ 55090-2012. Ресурсосбережение. Обращение с отходами.
Рекомендации по утилизации отходов бумаги // Электронный фонд
правовой и нормативно-технической документации. [Электронный
ресурс]. URL: <http://docs.cntd.ru/document/1200103182> (дата
обращения: 07.06.2020).
- 19.ГОСТ 55102-2012. Ресурсосбережение. Обращение с отходами.
Руководство по безопасному сбору, хранению, транспортированию и
разборке отработавшего электротехнического и электронного
оборудования, за исключением ртутьсодержащих устройств и
приборов // Электронный фонд правовой и нормативно-технической
документации. [Электронный ресурс]. URL:
<http://docs.cntd.ru/document/1200104723> (дата обращения: 07.06.2020).

20. Постановление Правительства РФ от 03.09.2010 N 681 (ред. от 01.10.2013) "Об утверждении Правил обращения с отходами производства и потребления в части осветительных устройств, электрических ламп, ненадлежащие сбор, накопление, использование, обезвреживание, транспортирование и размещение которых может повлечь причинение вреда жизни, здоровью граждан, вреда животным, растениям и окружающей среде // Государственная система правовой информации [Электронный ресурс]. URL: <http://pravo.gov.ru/proxy/ips/?docbody=&nd=102141053> (дата обращения: 07.06.2020).
21. ГОСТ Р 22.0.07-95 Безопасность в чрезвычайных ситуациях. Источники техногенных чрезвычайных ситуаций. Классификация и номенклатура поражающих факторов и их параметров
22. ГОСТ Р 22.3.03-94. Безопасность в чрезвычайных ситуациях. Защита населения. Основные положения
23. ГОСТ 8050-85 Двуокись углерода газообразная и жидкая. Технические условия (с Изменениями N 1, 2, с Поправкой)

Приложение А

(справочное)

Application of Deep Neural Networks for Pneumonia Detection

Студент

Группа	ФИО	Подпись	Дата
8ПМ8И	Кулебякин Алексей Сергеевич		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Аксёнов Сергей Владимирович	к.т.н., доцент		

Консультант-лингвист отделения иностранных языков ШБИП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ст. преподаватель	Пичугова Инна Леонидовна	-		

Introduction

Technology in the fields of life sciences and healthcare is constantly evolving. New manufacturing techniques have made it possible to make more effective therapies, and advances in research have led to more cost-effective solutions to conditions that were once considered too expensive to treat.

One of the most interesting aspects of the evolution of medical technology is the constantly increasing involvement of automation in various medical operations, particularly in diagnostics.

Before the involvement of automation, essential diagnostic procedures relied fully on manual procedures. While conventional methods were successful to a certain degree, they were hampered by certain limitations. For instance, manual testing of specimens is slow and prone to mistakes caused by poor human judgment.

Automation helps to get rid of these limitations to achieve more accurate results in less time. A delay in the diagnostic process puts the patient at risk of the illness getting worse as it remains undiagnosed longer. However, trying to accelerate the process also makes it more likely for medical professionals to come up with inaccurate results. To ensure proper diagnosis, the technology used to automate diagnostic processes must be able to deliver both speed and precision.

Overall, the involvement of automation leads to 3 main advantages:

- **more affordable treatment:** diagnosis is faster and more accurate with automation and doctors will be able to recommend the right medicine to patients before illnesses require more expensive treatment options;
- **safer solutions:** more accurate diagnosis means there is a lower risk of complications associated with patients receiving ineffective or incorrect treatment. Doctors will not have to second-guess their recommendations, and patients can rest easily knowing that they will always receive the treatment they need;
- **more patients treated:** by reducing the time it takes to complete a diagnosis, the laboratory can perform more tests with the right automated tools.

The goal of this master thesis is to develop an application for detecting thoracic diseases on chest X-ray scans. The application should also have an ability to be scaled

to a large number of medical institutions and have modular structure for future functionality expansion.

Competitor analysis

Currently there are lots of solutions for telemedicine, as well as for automated medical data processing, but only two of them have similar functionality to the one that should be designed and developed during the master's program.

The first solution is the ChestEye developed by Oxipit, a Poland-based company [1]. ChestEye support 75 most common radiological findings. It identifies chest X-ray images with no abnormality and produces preliminary reports. Preliminary healthy patient reports are produced only in cases where the system is highly confident of the results (20-40% of all healthy patient cases). By minimizing radiologist input required for healthy patient chest X-ray reporting, ChestEye improves radiologist productivity and allows medical specialists to focus their attention on cases with abnormalities. Unfortunately, Oxipit does not provide any information about their system's quality and accuracy.

The second and more promising solution is the tool called qXR 2.0 developed by the company Qure.ai based in India [2]. qXR detects abnormal chest X-ray findings. It can be used to separate normal from abnormal X-rays, for pre-read assistance, or as a radiology audit tool. qXR also includes a proprietary algorithm that screens X-rays for the signs of tuberculosis. The key feature of qXR is that it generates a description of the X-ray findings, including name, size and location of the abnormality, which is used to pre-fill radiology reports. qXR was built with deep learning and have been trained and tested using a growing database (over 2.5 Million) of X-rays from diverse sources.

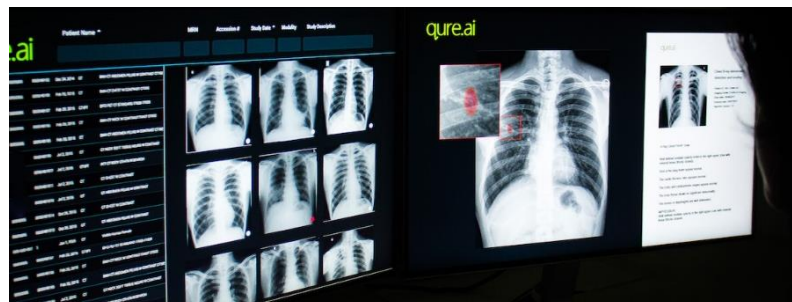


Figure 1. User interface of the qXR system

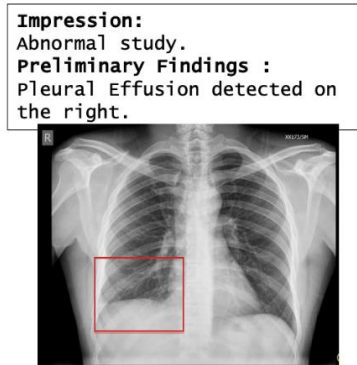


Figure 2. Full text report that pre-fills the radiology template

As it was mentioned, a proposed solution should not only have the functionality similar to the two described above, but also surpass them. In particular, it is planned to significantly improve the user interface for working with medical data, as well as to make the system scalable to an infinite number of medical institutions with an opportunity to expand its functionality in the future.

Artificial Intelligence

Most of the automated solutions including those mentioned above are, in fact, based on machine learning (ML) algorithms and to be more specific – deep learning (DL) models. Both DL and ML are parts of artificial intelligence (Figure 3). Many people are surprised to learn that AI is nothing new. It dates back centuries, but first became viable in the 1900s – and really took off in the 1950s. When computer scientist Alan Turing developed the Turing Test, he created the first test of machine intelligence against that of a human counterpart [3].

After that, each decade saw advancements in AI ranging from industrial robots in the automotive industry to interactive computer programs that could communicate with humans (i.e. the first inception of chatbots), sci-fi film depictions of humanized bots, and voice assistants in smartphones that are programmed with natural language processing (NLP) capabilities.

All of these advancements led to AI becoming commonplace in our lives so much that we often overlook that many of our day-to-day processes are fueled by artificial intelligence.

Machine learning is a subset of AI. That is, all machine learning counts as AI, but not all AI counts as machine learning. For example, symbolic logic – rules engines, expert systems and knowledge graphs – could all be described as AI, and none of them are machine learning.

One aspect that separates machine learning from the knowledge graphs and expert systems is its ability to modify itself when exposed to more data; i.e. machine learning is dynamic and does not require human intervention to make certain changes. That makes it less brittle, and less reliant on human experts.

The “learning” part of machine learning means that ML algorithms attempt to optimize along a certain dimension; i.e. they usually try to minimize error or maximize the likelihood of their predictions being true. This has three names: an error function, a loss function, or an objective function.

Deep learning is a subset of machine learning. Usually, when people use the term deep learning, they are referring to deep artificial neural networks, and somewhat less frequently to deep reinforcement learning.

Deep artificial neural networks are a set of algorithms that have set new records in accuracy for many important problems, such as image recognition, sound recognition, recommender systems, natural language processing etc. For example, deep learning is part of DeepMind’s well-known AlphaGo algorithm, which beat the former world champion Lee Sedol at Go in early 2016, and the current world champion Ke Jie in early 2017.

Deep is a technical term. It refers to the number of layers in a neural network. A shallow network has one so-called hidden layer, and a deep network has more than one layer. Multiple hidden layers allow deep neural networks to learn features of the data in a so-called feature hierarchy, because simple features recombine from one layer to the next to form more complex features. Neural networks with many layers pass input data (features) through more mathematical operations than nets with few layers, and, therefore, they are more computationally intensive to train. Computational intensity is one of the hallmarks of deep learning, and it is the reason why a new kind

of chips called GPUs (Graphics Processing Units) is in demand to train deep-learning models.

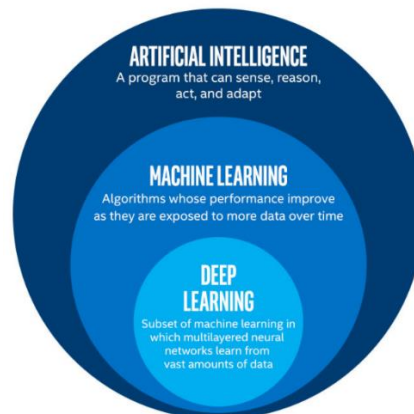


Figure 3. Venn diagram of AI, machine learning and deep learning

The truth is that AI, as its core, is merely programming. The use and application of AI and machine learning in healthcare enterprise systems are still relatively new. In the current ecosystem, with increasing healthcare costs, AI and machine learning algorithms can help to save lots of money by using it for patients' data analysis.

Both patients and healthcare professionals generate a huge amount of data. There are several medical data types: text, numeric, time-series and images. Image data is by far the most complex data to work with.

The ethics of AI are currently without guidelines and regulations. Many assume that AI has an objectivity that puts it above questions of morality, but in healthcare it is not the case. AI algorithms are only as fair and unbiased as the learnings, which come from the environmental data.

Applications of AI and machine learning for healthcare

Machine learning and AI are transforming the healthcare industry and significantly improving outcomes. However, it is unlikely that AI agents will ever completely replace doctors. Machine learning algorithms can improve diagnostics, predict certain outcomes, and will be a basis of personalized care.

The potential applications of machine learning in healthcare are vast. Predicting disease, detecting the risk of cancers or other pathologies that can be found on X-ray scans, suggesting meditation courses – are potential applications. Doctors make

challenging decisions every day. It is vital that these decisions are as informed as possible. Since it is illegal to fully rely on machine learning algorithms when it comes to human lives, the best use of AI and ML right now is to help doctors rationalize decisions and suggest better treatment options [4].

As for AI-assisted radiology, there is a huge amount of electronically-stored medical imaging data and deep learning algorithms which are the best suited for working with imaging data can be fed with this kind of dataset to detect and discover patterns and anomalies. Algorithms can interpret the imaging data like a highly trained radiologist can – identifying suspicious spots on the skin, lesions, tumors, and brain bleeds. The usage of AI/ML tools for assisting radiologists is, therefore, going to expand exponentially.

A critical problem in the healthcare right now is a lack of well-trained radiologists. In most circumstances, such skilled workers are under enormous strain due to the deluge of digital medical data. An average radiologist needs to produce interpretation results for one image every 3–4 seconds to meet the demand. [5]

Deep Learning for X-ray scan analysis

Deep learning algorithms like deep neural networks are known for their high performance when it comes to image processing.

There are several neural network architectures that could work for finding patterns on X-ray scans. One of them is called Densely Connected Convolutional Network [6], DenseNet. It was originally designed to be trained on the ImageNet dataset but could be easily retrained to work with different datasets due to its adaptive architecture.

Previously developed neural networks had from 5 to 20 hidden layers inside them. DenseNet is the next step, since it is increasing the depth of deep convolutional networks to 121, 169, 201 and 264 hidden layers.

DenseNet connects each layer to every other layer in a feed-forward fashion. Whereas traditional convolutional networks with L layers have L connections – one between each layer and its subsequent layer – DenseNet has $L(L + 1) / 2$ direct

connections. For each layer, the feature-maps of all preceding layers are used as inputs, and its own feature-maps are used as inputs into all subsequent layers. DenseNets have several compelling advantages: they alleviate the vanishing-gradient problem, strengthen feature propagation, encourage feature reuse, and substantially reduce the number of parameters.

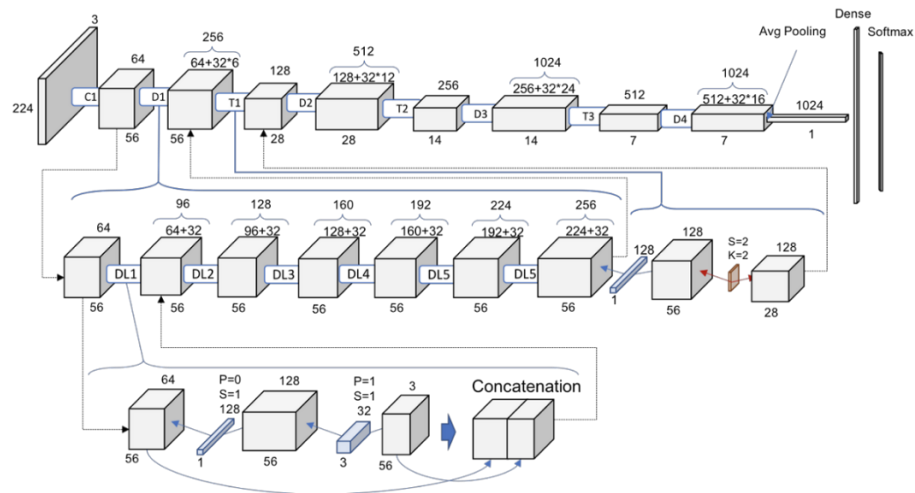


Figure 4. Full schematic representation of DenseNet-121

Here C_x is a Convolution Layer x , D_x – Dense Block x , T_x – Transition Block x , DL_x – Dense Layer x .

If you look closely, every Dense Layer is adding to the previous volume these 32 new feature maps. This is why it goes from 64 to 256 after 6 layers. In addition, Transition Block performs as 1×1 convolution with 128 filters followed by a 2×2 pooling with a stride of 2, resulting on dividing the size of the volume and the number of feature maps on half. A couple of statements from this pattern observation:

- The volume within a Dense Block remains constant;
- The volume and the feature maps are halved after every Transition Block.

However, it is necessary to understand what is going on inside every Dense Layer within each Dense Block to get the idea of how the whole network works. On the deepest level it can be seen how actually this behavior of adding 32 times the number of layers is achieved. Each Dense Layer has a 1×1 convolution with 128 filters to reduce the feature maps size and perform a more expensive 3×3 convolution with this chosen 32 number of feature maps of growth rate.

Then, the input volume and the result of the two operations (which are the same for every Dense Layer within every Dense Block) are concatenated, in the action of adding new information to the common knowledge of the network.

The Dataset and Training

The ChestX-ray14 dataset is a popular publicly available dataset released by Wang et al. which contains 112,120 frontal-view X-ray images of 30,805 unique patients [7]. Each image in the dataset is annotated with up to 14 different thoracic pathology labels using automatic extraction methods on radiology reports.

For the sake of simplicity and as a proof of concept, only pneumonia was diagnosed in this master thesis. Images that have pneumonia as one of the annotated pathologies were labeled as positive examples and all other images were labeled as negative examples. For the pneumonia detection task, the dataset was randomly split into training (28744 patients, 98637 images), validation (1672 patients, 6351 images), and test (389 patients, 420 images) sets. There is no patient overlap between the sets. Before inputting the images into the network, the images were downscaled to 224×224 pixels and normalized based on the mean and standard deviation of images in the ImageNet training set. It is also reasonable to augment the training data with random horizontal flipping to make the model more flexible.



Figure 5. Sample image from the ChestX-ray14 dataset

The training process is fairly simple – the final fully connected layer of the DenseNet-121 is replaced with one that has a single output, after which a sigmoid nonlinearity is applied. The weights of the network are initialized with weights from a model pretrained on ImageNet. The network is trained end-to-end using Adam

optimizer. The network's input on each training epoch is a minibatch of 16 images. Initial learning rate is set to be 0.01 and it decays by a factor of 10 each time the validation loss plateaus after an epoch, and picks the model with the lowest validation loss.

To test the model's performance an AUC (Area Under the Curve) ROC (Receiver Operating Characteristics) method was used. AUC – ROC curve is a performance measurement for classification problem at various thresholds settings. ROC is a probability curve and AUC represents degree or measure of separability. It tells how much model is capable of distinguishing between classes. The higher the AUC is, the better the model is at predicting 0s as 0s and 1s as 1s. By analogy, the higher the AUC is, the better the model is at distinguishing between patients with disease and no disease.

An excellent model has AUC near to the 1 which means it has good measure of separability. A poor model has AUC near to the 0 which means it has worse measure of separability. In fact, it means it is reciprocating the result. It is predicting 0s as 1s and 1s as 0s. In addition, when AUC is 0.5, it means the model has no class separation capacity whatsoever.

The model's performance is shown in Table 1. However, to ensure that the chosen model has the best performance it is necessary to compare it to other popular deep neural networks.

Table 1. DenseNet-121 performance compared to other popular deep neural networks

Pathology	DenseNet-121	DenseNet-169	ResNet-50	SEResNet-50
AUC - ROC Mean	0.8337	0.8308	0.8138	0.8317
Atelectasis	0.8248	0.8156	0.8063	0.8144
Cardiomegaly	0.8982	0.8993	0.9072	0.8981
Effusion	0.8757	0.8792	0.8779	0.8743
Infiltration	0.6738	0.6737	0.6638	0.6756
Mass	0.8469	0.8313	0.8261	0.8342
Nodule	0.7643	0.7491	0.728	0.7583
Pneumonia	0.7781	0.7735	0.7551	0.7608
Pneumothorax	0.8541	0.8602	0.8382	0.8638

Consolidation	0.8107	0.8096	0.8069	0.8092
Edema	0.9033	0.9103	0.9073	0.9081
Emphysema	0.9209	0.91	0.8724	0.9194
Fibrosis	0.8284	0.818	0.7981	0.8291
Pleural Thickening	0.7735	0.7658	0.7431	0.766
Hernia	0.9194	0.9355	0.863	0.9322

As it can be seen in Table 1, in some cases DenseNet-121 performs worse than other networks. However, the mean AUC-ROC and especially the one for detecting pneumonia is the best of all 4 models. It means that DenseNet-121 is indeed well-suited for the task of detecting pneumonia on X-ray scans. Increasing the number of hidden layers will not give better results as seen in the comparison between DenseNet-121 and DenseNet-169. Additionally, the training time is increasing exponentially with the increasing number of layers. With the current configuration it took about 3 hours to train the DenseNet-121.

Changing hyperparameters like learning rate will not significantly improve the performance either – learning rate equal 0.001 or even 0.0001 will just make the training process slower and 0.1 and higher will make the model perform worse.

The further improvements could be made by switching to:

- a different dataset with a larger number of X-ray scans that involves not only frontal-view scans but also side-views and basic patient’s data like age, sex and previously found pathologies;
- a neural network with a different, more advanced architecture.

These two options will require much more computing power to train the model and even more qualified radiologists to label the data.

However, the current results show the high potential of the developed solution and could be used to at least separate the patients without pathologies from those who have them. Moreover, if any pathology is identified, a healthcare specialist will be notified as soon as possible to review the finding and either approve the proposed diagnosis or set some further tests to make sure that the diagnosis is correct/wrong if the specialist does not trust the recommendation.

Conclusion

The current state of AI and Deep Learning algorithms allow us to automate not only routine tasks like filtering e-mail spam but even more complex ones like detecting a variety of pathologies or developing new drugs, for instance.

A constantly increasing demand for the qualified diagnostics force humanity to rely more on automated solutions like the one described above. It is possible that it will take years to integrate intelligent tools into the existing healthcare systems, but it is absolutely necessary. As an example, the COVID-19 pandemic happening right now that has already taken over 400,000 lives could have been suppressed much earlier if humanity had developed and integrated solutions to automatically decode virus' genome and detect the atypical pneumonia on earlier stages.

However, since much more laboratories and researchers are constantly working on developing preventive solutions, there is a high chance that the future pandemics will not be so fatal and will be defeated before it does a lot of harm to the humanity.

References

1. ChestEye CAD. Automate healthy patient diagnosis and reporting. Oxipit.ai. [Электронный ресурс] URL: <https://oxipit.ai/products/chesteye/>
2. Automated Chest X-ray Interpretation – qXR v2.0. Qure.ai. [Электронный ресурс] URL: <http://qure.ai/qxr.html>
3. The Future of Artificial Intelligence in Healthcare. HealthTech. [Электронный ресурс] URL: <https://healthtechmagazine.net/article/2020/02/future-artificial-intelligence-healthcare>
4. The potential for artificial intelligence in healthcare. Future Healthcare Journal. NCBI. [Электронный ресурс] URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6616181/>
5. Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. (2018). Artificial intelligence in radiology. Nature reviews. Cancer, 18(8), 500–510. (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6268174/>)
6. G. Huang, Z. Liu and L. van der Maaten, “Densely Connected Convolutional Networks,” 2018.
7. Wang, Xiaosong, Peng, Yifan, Lu, Le, Lu, Zhiyong, Bagheri, Mohammadhadi, and Summers, Ronald M. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. arXiv preprint arXiv:1705.02315, 2017.