

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 школы (НОЦ) – Отделение информационных технологий

### МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
<b>Система анализа геологических карт с помощью инструментов компьютерного зрения и машинного обучения</b>

УДК: 004.932.2:004.853:550.8:528

Студент

Группа	ФИО	Подпись	Дата
8ПМ8И	Чугунов Роман Анварович		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент	Аксенов С.В.	к.т.н.		

### КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент	Меньшикова Е.В.	к.филос.н.		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент	Горбенко М.В.	к.т.н.		

По разделу на английском языке

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель	Пичугова И.Л.	—		

### ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
доцент	Губин Е.И.	к.ф.-м.н.		

## ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ООП

Код результата	Результат обучения
<b><i>Общие по направлению подготовки 09.04.04 «Программная инженерия»</i></b>	
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
<b><i>Профиль «Технологии больших данных»/ «Big data solutions»</i></b>	
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 школы (НОЦ) – Отделение информационных технологий

УТВЕРЖДАЮ:  
 Руководитель ООП

\_\_\_\_\_  
 (Подпись)    (Дата)    (Ф.И.О.)

**ЗАДАНИЕ**  
**на выполнение выпускной квалификационной работы**

В форме:

Магистерской диссертации
--------------------------

Студенту:

Группа	ФИО
8П8МИ	Чугунов Роман Анварович

Тема работы:

<b>Система анализа геологических карт</b>	
Утверждена приказом директора	№59-62/с от 28.02.2020 г.

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

**ТЕХНИЧЕСКОЕ ЗАДАНИЕ:**

<p><b>Исходные данные к работе</b></p>	<p style="text-align: center;">Изображения геологических карт (1453 шт.)</p> <p style="text-align: right;">структурных</p>
<p><b>Перечень подлежащих исследованию, проектированию и разработке вопросов</b></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</i></p>	<ol style="list-style-type: none"> <li>1. Аналитический обзор подходов к решению задач:             <ul style="list-style-type: none"> <li>– Сегментация изображений;</li> <li>– Классификация изображений;</li> <li>– Детекция объектов на изображениях;</li> </ul> </li> <li>2. Реализация методов решения данных задач</li> </ol>

<i>работы; наименование дополнительных разделов, подлежащих разработке; заключение по работе).</i>	<p>3. Реализация следующих модулей системы анализа изображений геологических карт:</p> <ul style="list-style-type: none"> <li>– Модуль анализа изображений с применением методов машинного обучения;</li> <li>– Модуль доработки результатов анализа;</li> </ul> <p>4. Финансовый менеджмент;</p> <p>5. Социальная ответственность;</p> <p>6. Заключение.</p>
<b>Перечень графического материала</b>	Рисунки и таблицы для описания исследования предметной области, реализованных модулей системы, составления финансовой стороны и уточнения социальной ответственности.

**Консультанты по разделам выпускной квалификационной работы**

*(с указанием разделов)*

Раздел	Консультант
Социальная ответственность	Горбенко Михаил Владимирович, доцент ООД ШБИП, к.т.н.
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Меньшикова Е.В., доцент ОСГН ШБИП, к.ф.н.
Раздел на английском языке	Пичугова Инна Леонидовна, старший преподаватель ОИЯ ШБИП

**Названия разделов, которые должны быть написаны на русском и иностранном языках:**

1 Модуль анализа изображений геологических карт
2 Модуль доработки результатов анализа
3 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение
4 Социальная ответственность

<b>Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику</b>	01.02.2020
---	------------

**Задание выдал руководитель:**

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОИТ ИШИТР	Аксенов С.В.	к.т.н		

**Задание принял к исполнению студент:**

Группа	ФИО	Подпись	Дата
8ПМ8И	Чугунов Роман Анварович		

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  
 Уровень образования магистратура  
 Отделение школы (НОЦ) – Отделение информационных технологий  
 Период выполнения: весенний семестр 2019 /2020 учебного года

Форма представления работы:

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)
09.02.2019	1 Модуль анализа изображений геологических карт / Описание используемых подходов	30
18.04.2019	2 Модуль доработки результатов анализа / Описание модуля редактирования результатов работы алгоритмов машинного обучения	30
11.05.2019	3 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	20
27.05.2019	4 Социальная ответственность	20

**СОСТАВИЛ:**

**Руководитель ВКР**

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОИТ ИШИТР	Аксенов С.В.	К.Т.Н.		

**СОГЛАСОВАНО:**

**Руководитель ООП**

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОИТ ИШИТР	Губин Е.И.	К.Ф.-М.Н.		

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

Студенту:

<b>Группа</b>	<b>ФИО</b>
8ПМ8И	Чугунов Роман Анварович

<b>Школа</b>	<b>ИШИТР</b>	<b>Отделение школы (НОЦ)</b>	<b>ОИТ</b>
Уровень образования	Магистратура	Направление/специальность	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 инженера – 21760; - Оклад руководителя от организации – 35000; - Оклад научного руководителя – 35120;
2. Нормы и нормативы расходования ресурсов	- Годовая норма амортизации составляет 33.3 %.
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	- Размер страховых взносов равный 30%; - Районный коэффициент г. Томск.

**Перечень вопросов, подлежащих исследованию, проектированию и разработке:**

1. Оценка коммерческого и инновационного потенциала НТИ	- Анализ потенциальных потребителей; - Диаграмма Исикавы; - Анализ конкурентных технических решений.
2. Разработка устава научно-технического проекта	Постановка целей проекта, определение ожидаемого результата и критериев успешности проекта.
3. Планирование процесса управления НТИ: структура и график проведения, бюджет, риски и организация закупок	- Построение диаграммы Ганта; - Планирование работ; - Формирование бюджета затрат; - Анализ рисков проекта.
4. Определение ресурсной, финансовой, экономической эффективности	Расчет финансового показателя эффективности. Описание общих выводов.

**Перечень графического материала (с точным указанием обязательных чертежей):**

1. Сегментация рынка
2. Результаты оценки конкурентных систем анализа картографической информации
3. Диаграмма Исикавы
4. Календарный график проекта
5. Бюджет затрат
6. Реестр рисков

<b>Дата выдачи задания для раздела по линейному графику</b>	
---	--

**Задание выдал консультант:**

<b>Должность</b>	<b>ФИО</b>	<b>Ученая степень, звание</b>	<b>Подпись</b>	<b>Дата</b>
Доцент	Меньшикова Екатерина Валентиновна	Кандидат философских наук		

**Задание принял к исполнению студент:**

<b>Группа</b>	<b>ФИО</b>	<b>Подпись</b>	<b>Дата</b>
8ПМ8И	Чугунов Роман Анварович		

## ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа	ФИО
8ПМ8И	Чугунов Роман Анварович

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

Тема ВКР:

Система анализа геологических карт	
<b>Исходные данные к разделу «Социальная ответственность»:</b>	
1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения	Разработка программных модулей системы анализа геологических карт, целью которой является автоматизация процесса создания трехмерной модели структуры нефтяного пласта на основе изображения геологической карты.
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
<b>1. Правовые и организационные вопросы обеспечения безопасности:</b> – специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства; – организационные мероприятия при компоновке рабочей зоны.	- ГОСТ 12.2.032-78 ССБТ - СанПиН 2.2.4.548-96 - СанПиН 2.2.4/2.1.8.562-96 - СанПиН 2.2.2/2.4.1340-03 - ГОСТ 12.1.009-2009 - ГОСТ 12.1.038-82 ССБТ - ГОСТ Р 22.3.03-94
<b>2. Производственная безопасность:</b> 2.1. Анализ выявленных вредных и опасных факторов 2.2. Обоснование мероприятий по снижению воздействия	Выявленные опасные факторы: - Электрический ток (источник – ПК) - Короткое замыкание - Статическое заземление (источник – ПК) Выявленные вредные факторы: - Освещение - Микроклимат - Шум - Психофизиологические факторы: нервно-психические перегрузки
<b>3. Экологическая безопасность:</b>	Отсутствует воздействие объекта на атмосферу, гидросферу. Воздействие на литосферу происходит при утилизации ПК, используемого для разработки, а также утилизации люминесцентных ламп освещения.
<b>4. Безопасность в чрезвычайных ситуациях:</b>	В ходе разработки программы возможно возникновение пожара на рабочем месте
Дата выдачи задания для раздела по линейному графику	
01.02.2020	

**Задание выдал консультант:**

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Горбенко Михаил Владимирович	Кандидат технических наук		

**Задание принял к исполнению студент:**

Группа	ФИО	Подпись	Дата
8ПМ8И	Чугунов Роман Анварович		

## **РЕФЕРАТ**

Выпускная квалификационная работа содержит 107 с., 43 рис., 21 табл., 24 источников, 1 приложение.

Ключевые слова: анализ данных, компьютерное зрение, сегментация изображений, детекция объектов, классификация изображений, машинное обучение, глубокое обучение.

Объект исследования – процесс работы с геологическими картами, в частности, в геологии и нефтегазовой сфере.

Предмет исследования – использование методов машинного обучения для анализа изображений в геологии.

Цель работы – исследование методов машинного обучения и их использование для реализации модулей системы анализа изображений структурных геологических карт.

В процессе исследования проводилось изучение, анализ, тестирование и сравнение различных существующих методов анализа изображений для решения задач классификации, сегментации изображений и детекции объектов.

В результате исследования были реализованы следующие модули системы: модуль анализа изображений геологических карт и модуль доработки результатов анализа.

Область применения: разработанная система может быть использована на предприятиях нефтегазовой отрасли, а также в научно-исследовательских организациях данной отрасли.

## **ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ, СОКРАЩЕНИЯ**

Машинное обучение – семейство алгоритмов, способных обучаться. Выделяют три основных области машинного обучения: обучение без учителя, обучение с частичным привлечением учителя, обучение с учителем и обучение с подкреплением.

Компьютерное зрение (иначе техническое зрение) — теория и технология создания машин, которые могут производить обнаружение, отслеживание и классификацию объектов.

Глубокое обучение – совокупность методов машинного обучения, основанных на обучении представлениям, а не специализированным алгоритмам под конкретные задачи.

Структурные карты – карты, отображающие какую-либо опорную геологическую поверхность (кровля или подошва стратиграфических подразделений, маркирующие слои и горизонты, поверхности несогласий, разрывных разрушений, залежей полезных ископаемых, водоносных горизонтов и т.п.), скрытую на глубине. При построении структурных карт используются данные, полученные при геологической съёмке, бурении скважин, проведении горных выработок или при геофизических исследованиях, по которым устанавливаются высотные отметки опорной поверхности в разных точках площади исследования. Изображение формы и глубины залегания даётся с помощью стратоизогипс, методика построения которых аналогична изогипсам с учётом всех данных геологического строения территории.

## ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	12
МОДУЛЬ АНАЛИЗА ИЗОБРАЖЕНИЙ ГЕОЛОГИЧЕСКИХ КАРТ .	14
Введение в глубокое обучение .....	14
Дизайн решения.....	20
Детекция скважин (символов) .....	22
Распознавание значений.....	25
Распознавание изолиний .....	28
Описание алгоритма .....	28
Исходный подход к извлечению изолиний.....	28
Извлечение изолиний .....	29
Векторизация линий .....	33
Объединение частей линий .....	34
Кластеризация по цвету и толщине линий.....	38
Сегментация изображений для определения областей карт .....	40
Процесс обучения .....	43
Процесс предсказания на новых данных.....	45
МОДУЛЬ ДОРАБОТКИ РЕЗУЛЬТАТОВ АНАЛИЗА.....	47
Интерфейс системы.....	47
Разработка модуля доработки результатов анализа .....	49
Интерфейс модуля доработки.....	49
Обоснование выбора фреймворка .....	54
Реализация интерфейса .....	57
«ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ» .....	59
«СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ» .....	78
ЗАКЛЮЧЕНИЕ .....	89

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ .....	91
Приложение II.....	94

## **ВВЕДЕНИЕ**

Основным направлением данной работы является исследование современных методов анализа изображений.

Целью дипломной работы является разработка двух модулей системы анализа геологических карт. Данная система предназначена для автоматизации процесса создания трехмерных структурных моделей на основе геологических карт, а также представления разнородных геологических карт в единой системе, доступной пользователям через веб-интерфейс. Значительное количество картографической информации нефтегазовых компаний хранится в виде чертежей на бумаге, что несет за собой ряд проблем:

- Невозможность быстрого поиска карт по различным критериям: по региону, по дате создания, по схожести между собой, по создателю карты и т.д.;
- Различия в представлении различных карт;
- Проблемы в отслеживании актуальности карт;
- Сложность сравнения объектов, представленных на двух разных картах;

Система включает в себя следующие модули:

1. Серверная часть (backend) – взаимодействие с базой данных, аналитической частью, а также ответы за запросы клиентской части приложения.

2. Клиентская часть (frontend) – веб-интерфейс пользователя, который включает функции просмотра и редактирования имеющихся карт, добавления новых, вызова методов их автоматического анализа на основе алгоритмов компьютерного зрения, а также экспорта результатов в формате GeoJSON.

3. Аналитическая часть – модуль, целью которых является извлечение полезной информации из карт, такой как: наличие и координаты нефтяных скважин, различные виды изолиний и глубина их залегания.

Серверная и клиентская часть взаимодействуют посредством REST API. Клиентская часть реализована на JavaScript (Vue.js), серверная часть реализована на Python (Flask).

Конкретной целью данной дипломной работы является разработка аналитической части системы, а также одного из модулей клиентской части – модуля редактирования результатов аналитической части.

Аналитический модуль системы извлекает изолинии на структурных картах, распознает их глубину, а также детектирует имеющиеся скважины.

Модуль редактирования результатов предназначен для исправления ошибок в работе аналитического модуля, т.е., является частью пользовательского интерфейса.

## МОДУЛЬ АНАЛИЗА ИЗОБРАЖЕНИЙ ГЕОЛОГИЧЕСКИХ КАРТ

### Введение в глубокое обучение

Глубокое обучение широко используется в различных задачах компьютерного зрения. С ростом объемов данных, доступных для обучения, а также с увеличением вычислительных мощностей область компьютерного зрения все больше переходит от статистических методов к глубоким нейронным сетям. На рисунке 3 приведен сравнительный график производительности видеокарт P100, V100 и A100, используемых в компьютерах Nvidia DGX различных версий (2016 – 2020 гг.), в задаче обучения контекстуальной языковой модели BERT.

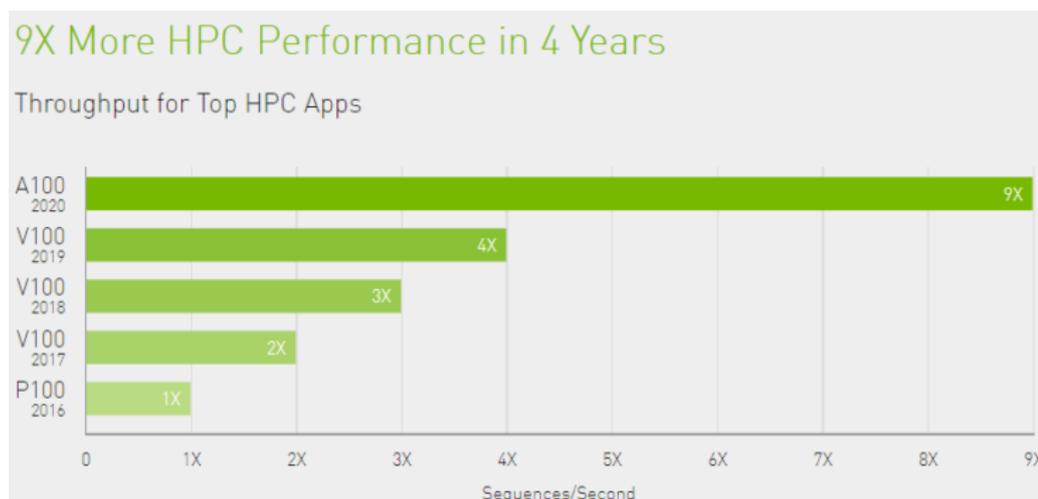


Рисунок 1 – Сравнительная производительность видеокарт P100, V100 и A100, последовательностей/сек

Наиболее частыми задачами в области компьютерного зрения, решаемых с помощью глубокого обучения, являются:

1. Классификация изображений.
2. Детекция объектов.
3. Сегментация изображений.
4. Перенос стиля.
5. Окрашивание изображений.
6. Реконструкция изображений.

7. Повышение разрешения изображений (Image Super-Resolution).
8. Генерация изображений.

Классификация изображений подразумевает определения класса изображений. Данную задачу можно разделить на несколько подзадач:

- Бинарная классификация – классификация на два класса.
- Мультиклассификация – классификация на более, чем 2 класса.
- Многозначная классификация – классификация объектов, которые могут одновременно относиться к нескольким классам.

Известными простыми примерами классификации (мультиклассификации) являются датасет MNIST (рисунок 2) и CIFAR-10 (рисунок 3).



Рисунок 2 – MNIST

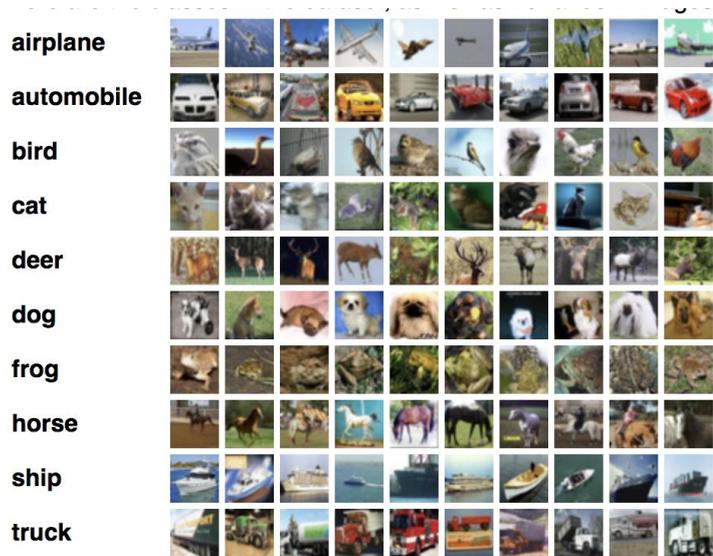


Рисунок 3 – CIFAR-10

В глубоком обучении часто применяется подход Transfer Learning – он основан на обучении нейронной сети решать задачу (как правило, задачу классификации) на большой датасете с дальнейшим переобучением для решения необходимой задачи. Например, если требуется создать сеть для сегментации изображений, часть слоев (кодировщик, предназначенный для извлечения признаков из изображения) можно заимствовать у сети, обученной классифицировать изображения на 1000 классов на датасете ImageNet. Множество архитектур для сегментации и детекции могут использовать различные подсети для извлечения признаков из изображений. Таким образом, такие архитектуры могут иметь различный backbone – сеть, на которой они основаны.

Детекция объектов кроме задачи классификации, включает в себя задачу локализации. Так как на изображении часто имеются несколько объектов разных классов, данная задача считается более сложной, чем обычная классификация. Результатом работы алгоритмов детекции являются ограничивающие рамки вокруг детектированных объектов. Каждая рамка относится к некоторому классу. На рисунке 4 представлена визуализация результатов работы сети Faster R-CNN на датасете MS COCO. Другим известным датасетом для детекции является PASCAL VOC 2012.

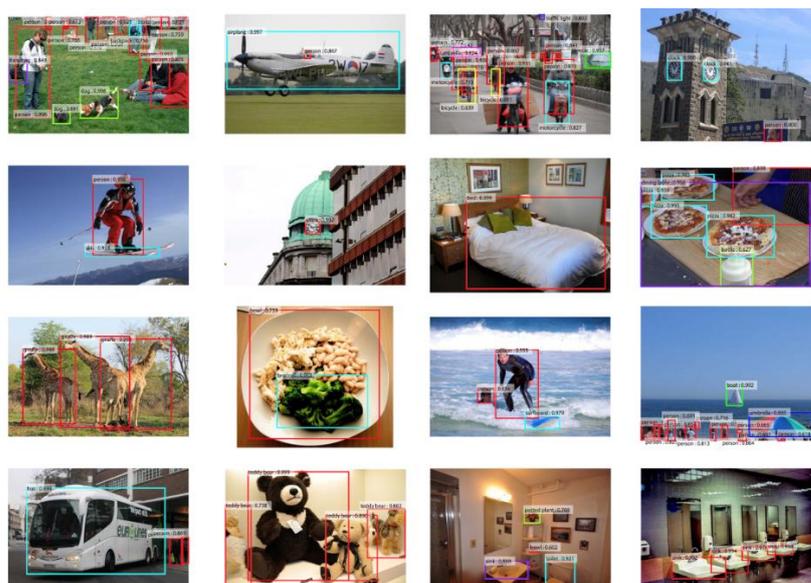


Рисунок 4 – результаты работы сети Faster R-CNN датасета MS COCO

Сегментация изображений – задача распознавания объектов, в которой для каждого объекта должна быть определена его маска – область, которую он занимает. Данную задачу можно рассматривать как задачу разбиения изображения на сегменты. В отличие от задачи детекции, сегментация изображений подразумевает определение пикселей, которые относятся к нужным объектам. Датасеты PASCAL VOC 2012 и MS COCO также могут быть использованы для сегментации изображений (рисунок 5).

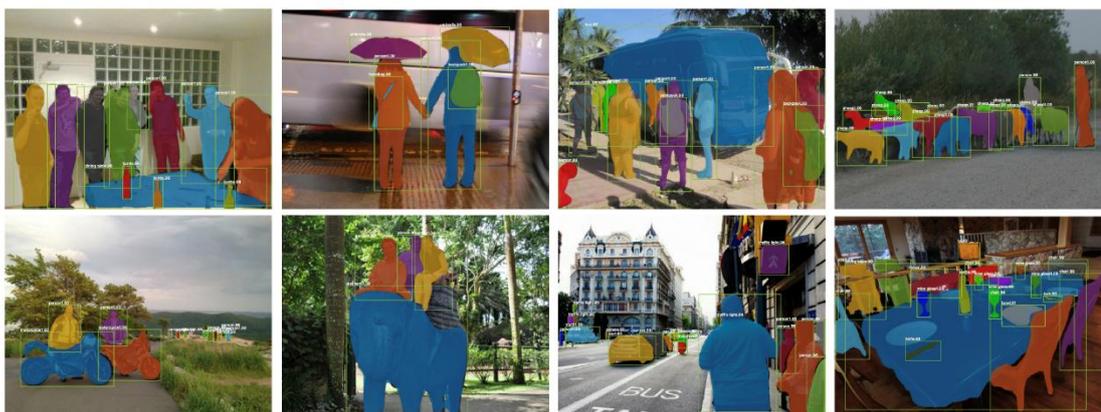


Рисунок 5 – Сегментация изображений датасета MS COCO

Задача переноса стиля – это задача извлечения «стиля» из одного или нескольких изображений и применения этого стиля к другому изображению. Решение данной задачи может потребоваться при создании различных фильтров для изображений (рисунок 6).

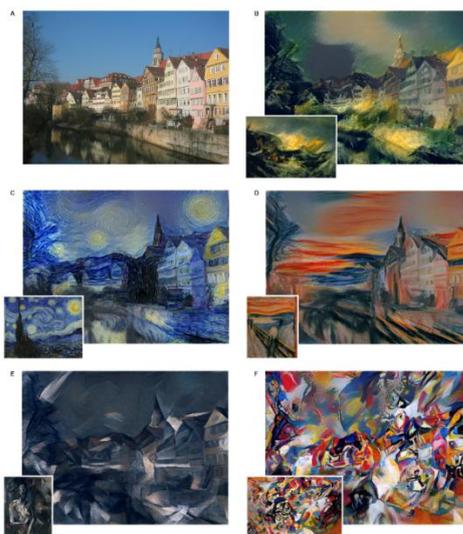


Рисунок 6 – Пример переноса стиля

Окрашивание изображений заключается в переводе черно-белых изображений в цветное. Решение данной задачи также имеет мало применений, кроме как использование в качестве фото-фильтра (рисунок 7) или для обработки черно-белых кинофильмов.

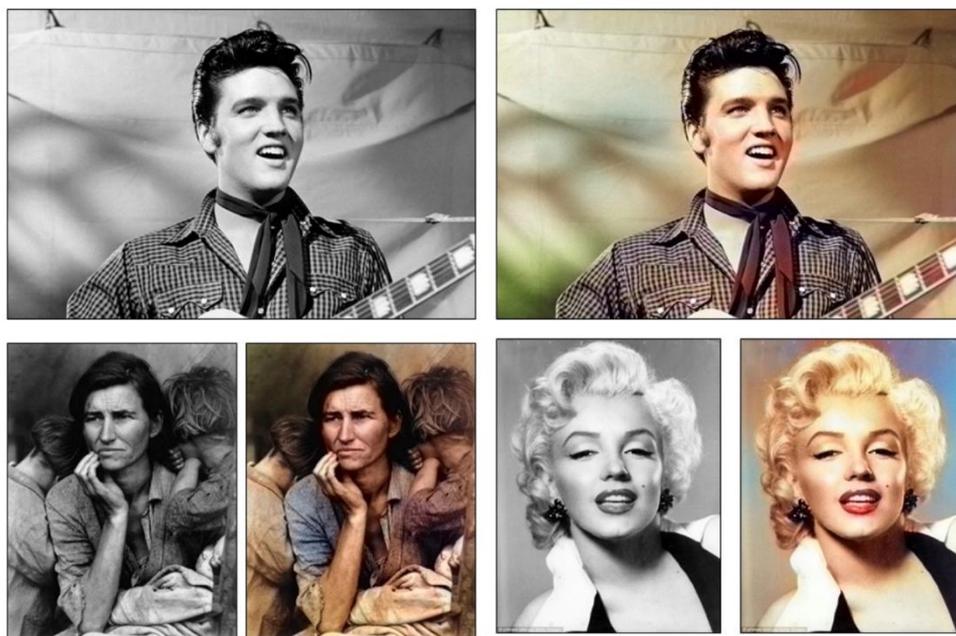


Рисунок 7 – Окрашивание изображений

Восстановление изображений – задача восстановления шумных или поврежденных изображений (рисунок 8).

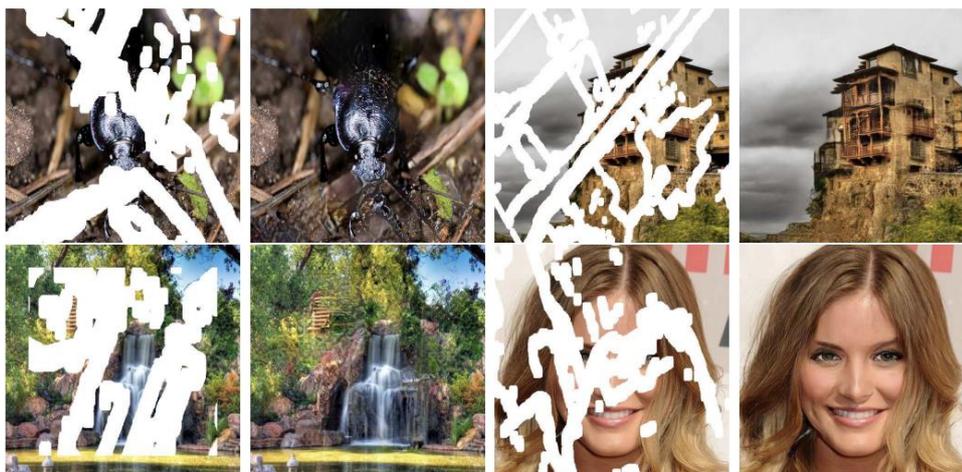


Рисунок 8 – Восстановление изображений

Image Super-Resolution – задача создания изображения более высокого качества, чем исходное изображение. Часто модели, которые используются

для Super-Resolution, могут использоваться в задаче восстановления изображений. Сравнение результатов работы алгоритмов Super-Resolution представлено на рисунке 9.



Рисунок 9 – Сравнение результатов работы различных алгоритмов для повышения разрешения с исходным (до сжатия) изображением (bicubic – алгоритм на основе интерполяции, не подразумевает обучения)

Генерация изображений – задача создания совершенно новых изображений или изображений, измененных определенным образом (добавление объекта на изображение, изменение стиля объекта на изображении). Пример генерации изображений представлен на рисунке 10.

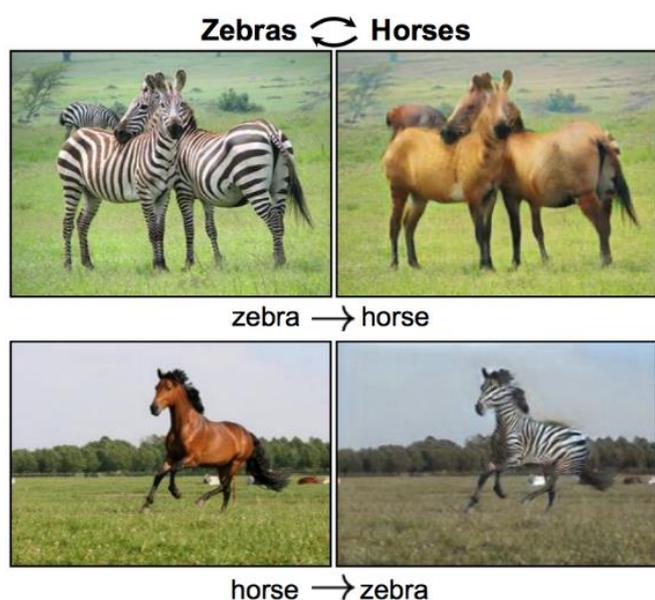


Рисунок 10 – Результат работы алгоритма генерации изображений

В аналитическом модуле системы используются алгоритмы детекции объектов, бинарной сегментации, а также ранее использовался алгоритм бинарной классификации.

### Дизайн решения

Система интеллектуальной оцифровки карт реализована в виде web-приложения.

Схема приложения представлена на рис. 11. Основные компоненты системы включают в себя:

- Пользовательский интерфейс в браузере (клиентская часть приложения) на фреймворке Vue.js.
- Web-сервис на Python (фреймворк Flask), реализующий необходимые интерфейсу функции (серверная часть приложения).
- Фоновый вычислительный модуль, выполняющий обработку карт и подготовку данных для web-сервиса (модуль анализа изображений).

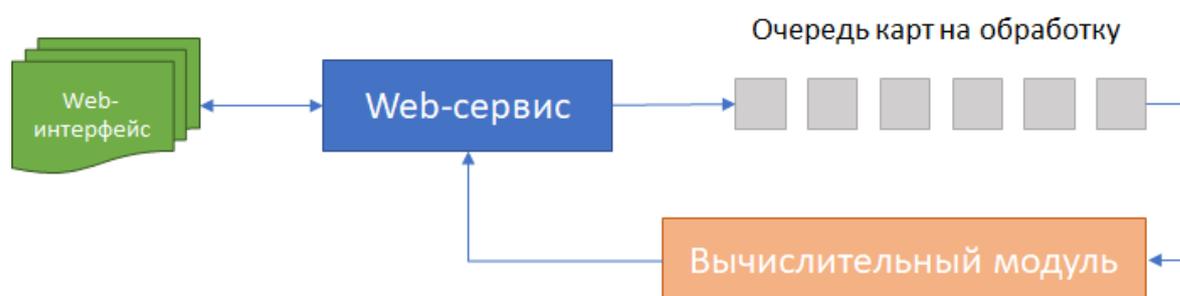


Рисунок 11 – Схема взаимодействия модулей

Модуль анализа изображений, реализованный внутри вычислительного модуля, состоит из следующих этапов (рисунок 12):

- Извлечение символов (символы – это объекты на изображении структурной геологической карты, которые извлекаются с помощью методов детекции, на данный момент к ним относятся только скважины различных типов).
- Извлечение цифр и чисел (цифры объединяются в числа на основе их взаимного расположения).
- Извлечение изолиний.

- Сопоставление чисел и изолиний.
- Этап, не являющийся частью общего процесса обработки изображения: бинарная сегментация изображений для определения области карт.

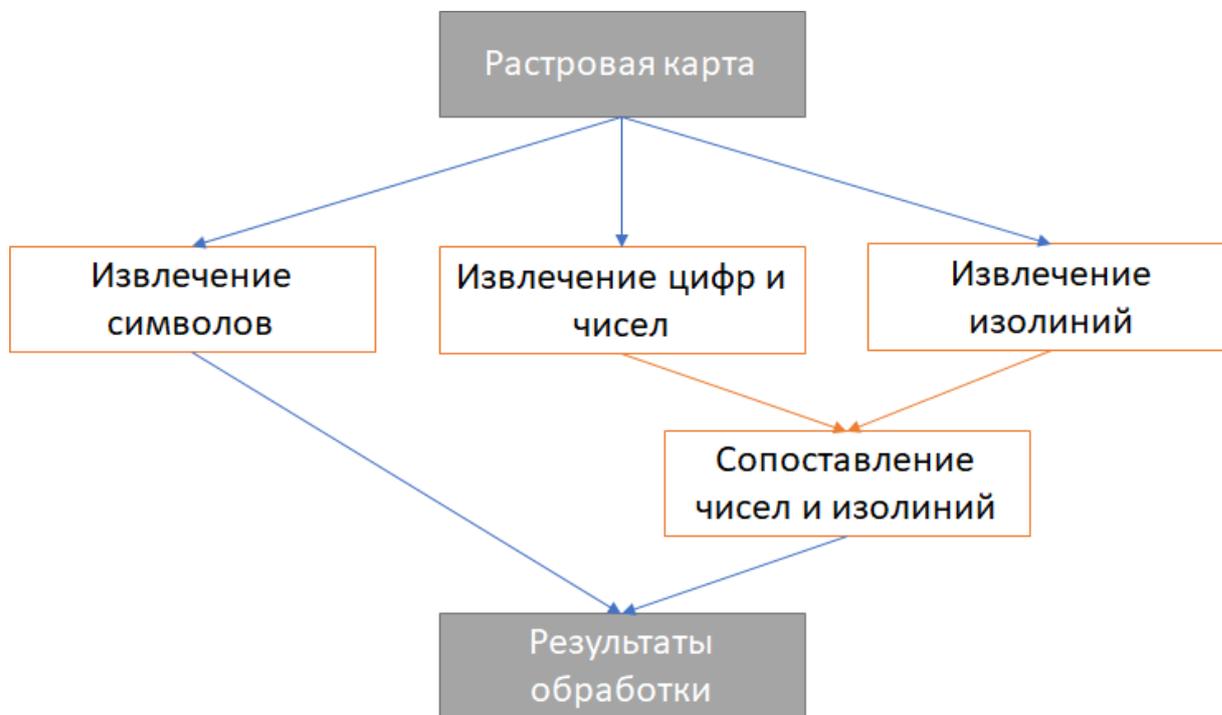


Рисунок 12 – Этапы обработки растровой карты

Алгоритм бинарной сегментации карт на изображении не включен в общий процесс обработки изображения на рисунке 12, так как вызывается независимо в тот момент, когда пользователь выбирает область для анализа – изначально ему предлагается область, определенная нейронной сетью (потенциальная область, содержащая карту).

Таким образом, целью модуля анализа изображений геологических карт является извлечение изолиний на карте, детекция чисел, определяющих глубину залегания изолиний, объединение линий с их глубинами, детекция символов (скважин), а также определение областей, содержащих карты на изображении. Работа модуля основана на алгоритмах компьютерного зрения, в частности, алгоритмов машинного обучения. Использование машинного обучения обусловлено тем, что карты могут сильно отличаться друг от друга,

также, невозможно предугадать, на каких типах карт в будущем будет использоваться система Заказчиком. Этим обусловлена необходимость создания алгоритмов, робастных на большом количестве возможных типов структурных карт. Создание подобной робастной экспертной системы на основе эвристических правил и алгоритмов не представляется возможным. Алгоритмы без обучения наиболее часто требуют тонкой настройки параметров под конкретный набор данных.

### **Детекция скважин (символов)**

Для распознавания скважин на изображении требуется использовать алгоритмы детекции. Современные алгоритмы детекции как правило основаны на использовании сверточных нейронных сетей. Нейронная сеть обучается определять координаты объектов и классы данных объектов на основе обучающей выборки.

Свёрточная сеть для классификации и обнаружения символов на картах была построена с использованием архитектур Inception [1] и Single Shot Detector (SSD) [2]. На вход сети подаётся фрагмент цветного (RGB) изображения размером  $64 \times 64$  пикселя, затем следует группа свёрточных слоёв и три Inception-блока, разделённые слоями субдискретизации с ядрами размера  $2 \times 2$ . Все свёрточные слои имеют функцию активации ReLU. Выходы последнего Inception-блока подаются независимо на три свёрточных слоя – выходы двух из них, с ядрами размера  $3 \times 3$  и тождественными функциями активации используются для классификации и локализации символов, соответственно. Данные, поданные на третий слой, затем также используются для классификации и локализации символов. Структура сети представлена на рисунке 13. Логиты – это векторы, полученные на выходе некоторых блоков сети.

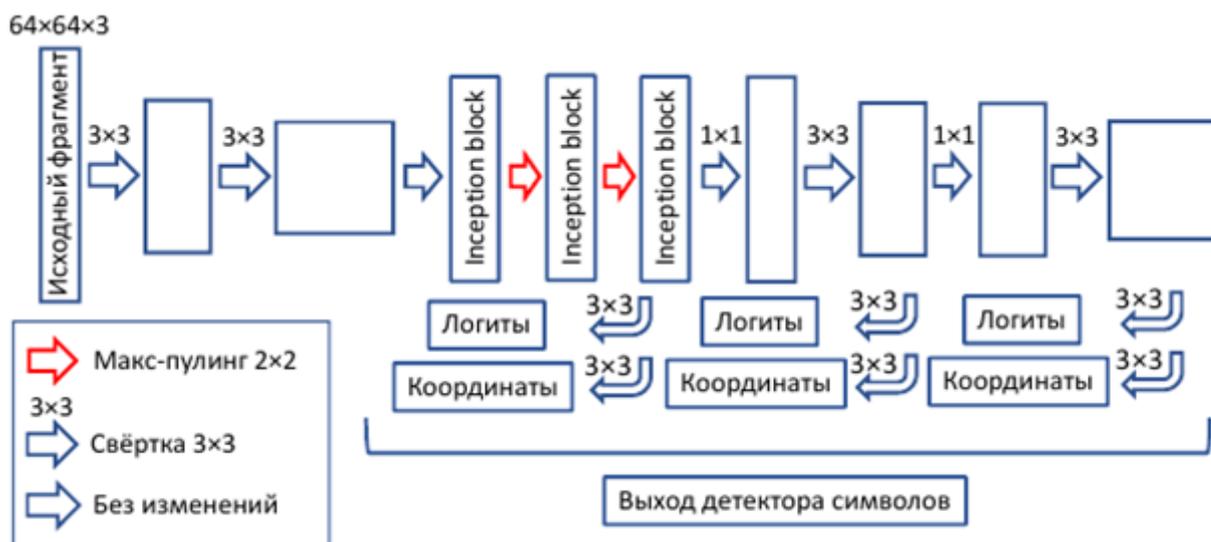


Рисунок 13 – Структура свёрточной сети, использованной для классификации и локализации символов на карте

В связи с тем, что использованная сеть принимает на вход изображение небольшого размера, реальные карты обрабатываются перекрывающимися фрагментами. Для повышения стабильности распознавания в обучающую выборку были добавлены представительные положительные и отрицательные примеры. Так, использовались различные начертания одних и тех же символов, а в качестве фона применялись синтетические данные, моделирующие структуру карт и содержащие линии различного начертания, и текстовые фрагменты. Примеры из обучающей выборки приведены на рисунке 4.



Рисунок 14 – Примеры из обучающей выборки

После обработки исходного изображения один символ часто обнаруживается множество раз (рисунок 5 (а)). Интерпретация результатов обнаружения символов значительно упрощается, если каждому обнаруженному символу соответствует один ограничивающий прямоугольник (рисунок 15 (б)).

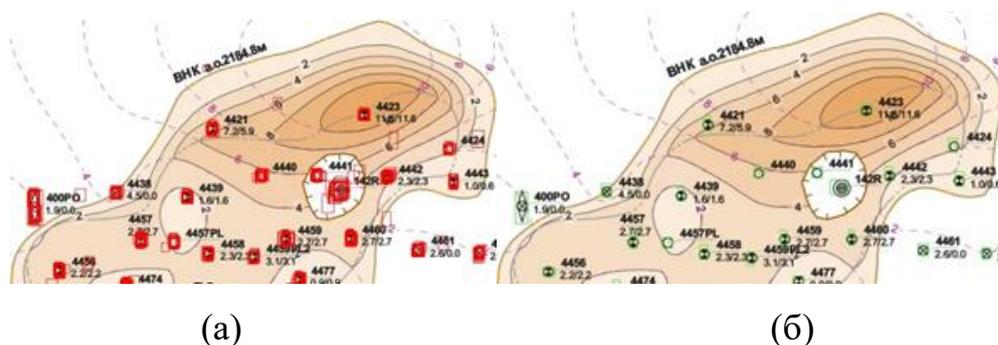


Рисунок 15 - Результат работы детектора до (а) и после (б) подавления не-максимумов

Для пост-обработки первичных результатов работы детектора был использован метод подавления не-максимумов (non-maximum suppression) [3, 4]. Данный метод имеет две стадии: фильтрация ограничивающих прямоугольников и отбрасывание ограничивающих прямоугольников. На первой стадии происходит отбор тех ограничивающих прямоугольников, для которых максимальная вероятность класса выше некоторого заранее заданного порога. Вторая стадия осуществляется итеративно: на каждой итерации из необработанных ограничивающих прямоугольников выбирается тот, для которого вес предсказанного класса максимален (опорный ограничивающий прямоугольник). Затем отыскиваются все ограничивающие прямоугольники, для которых мера Жаккара сходства с опорным не ниже некоторого заранее заданного порога, и отбрасываются. При этом метка класса, присваиваемая опорному ограничивающему прямоугольнику, выбирается как наиболее часто встречающаяся среди всех отброшенных ограничивающих прямоугольников, включая опорный. Итерации выполняются до тех пор, пока все ограничивающие прямоугольники не будут обработаны.

Ввиду отсутствия размеченной выборки настоящих (не созданных алгоритмом генерации синтетических данных) структурных карт, оценка детекции скважин проводилась на тестовой подвыборке синтетической

выборки. Оценка производилась по метрике mAP (mean average precision). Оценка алгоритма на тестовой выборке равнялась 84,5%.

Алгоритм показал высокую точность детекции на синтетической выборке, а также на настоящих структурных картах (по экспертной оценке). Ошибки детекции как правило обусловлены появлением новых классов объектов – скважин, которые обозначены другим символом, отличающимся от всех виденных ранее. Обозначение скважин зависит не только от их типа, но также меняется в зависимости от создателя карты. Карты, созданные различными организациями, наиболее часто имеют множество отличий по стилю, обозначениям, цветовой гамме.

### Распознавание значений

Извлечение значений также выполняется с использованием сети на основе подхода SSD. При этом входной сверточный каскад заменен на Xception [5], а в выход детектора помимо координат добавлено определение ориентации (рисунок 16).

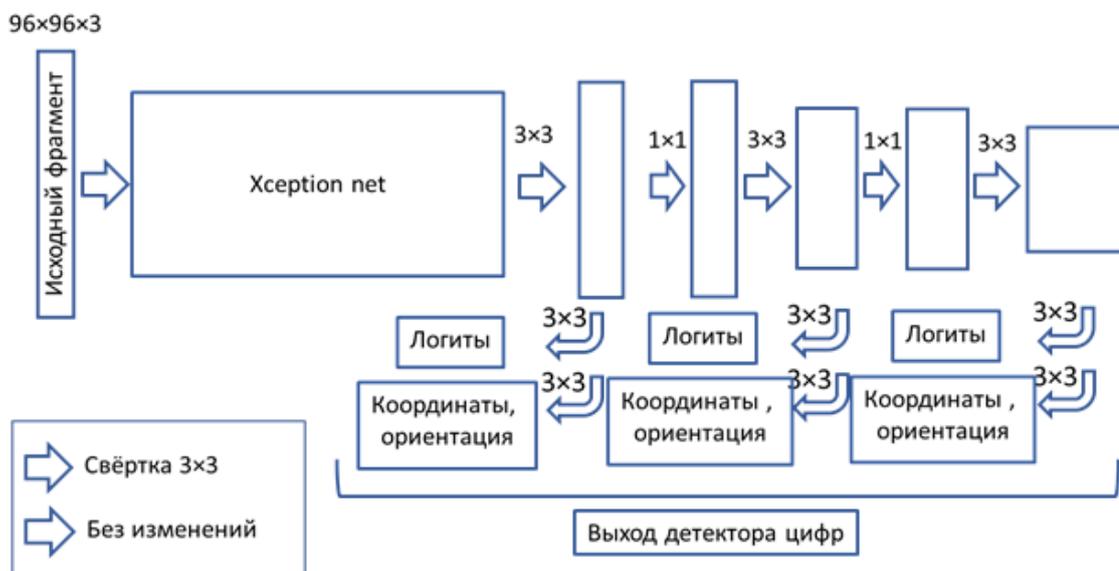


Рисунок 16 - Структура свёрточной сети, использованной для классификации и локализации цифр на карте

Модификация архитектуры обусловлена в том числе тем, что исходная архитектура была достаточна для достижения высокой точности детекции скважин, но распознавание чисел оказалось более сложной задачей. На

структурных картах часто встречаются трех-четырёхзначные числа (иногда с десятичной частью), а также более длинные обозначения и названия. Ошибка хотя бы в одном символе приводит к неправильному определению всего числа или обозначения.

Также после подавления не-максимумов добавлен еще один этап обработки, собирающий отдельные цифр в числа с использование обнаруженной информации о положении, размере и ориентации цифр.

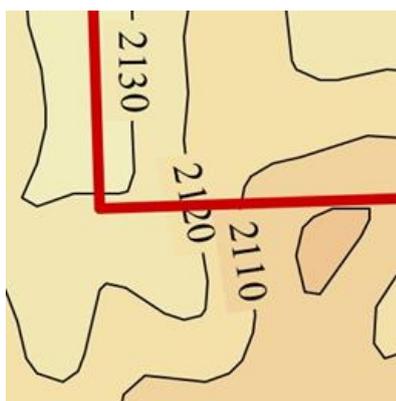
Обучение детектора цифр выполнено на полностью синтетических данных, аналогично подходу использованному в [6]. Примеры из обучающей выборки приведены на рисунке 17.



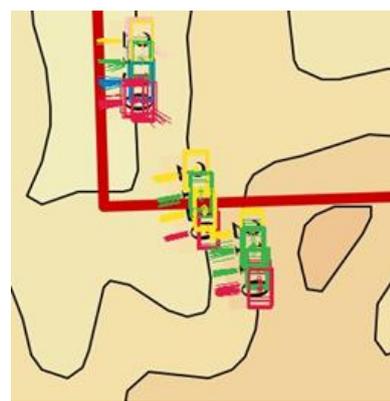
Рисунок 17 - Примеры из обучающей выборки

Алгоритм распознавания чисел и обозначений состоит из следующих шагов:

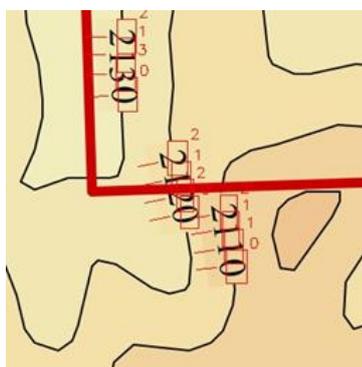
1. Исходное изображение карты разбивается на перекрывающиеся области 64x64 пикселя. Данные участки отправляются в нейронную сеть (рисунок 8 (а)).
2. Результатом работы сети является множество ограничивающих прямоугольников с их направлением (рисунок 8 (б)).
3. Подавление не-максимумов оставляет по одному ограничивающему прямоугольнику на цифру (рисунок 8 (в)).
4. Цифры объединяются в числа на основе их взаимного расположения (рисунок 18 (г)).



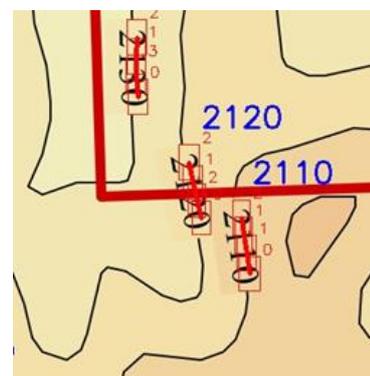
(а)



(б)



(в)



(г)

Рисунок 18 - Исходное изображение (а), результат работы нейросети(б), результат подавления не-максимумов(в), собранные числа(г)

Алгоритм распознавания чисел и обозначений также оценивался на синтетической тестовой выборке. Результат оценки оказался несколько хуже, чем детекция скважин – 79,8% mAP. Такой результат может быть обусловлен тремя факторами:

1. Числа имеют меньший размер, чем часть возможных обозначений скважин, в связи с чем их чаще перекрывают другие объекты.
2. Классов возможных чисел и букв больше, чем классов скважин.
3. Цифры как правило ориентированы также, как изолиния, которой они принадлежат, в то время как скважины всегда имеют одну и ту же ориентацию, небольшие отклонения возможны только в случае, если изображение карты получены путем сканирования карты.

## Распознавание изолиний

### Описание алгоритма

Наиболее сложной проблемой в ходе создания аналитического модуля стало распознавание изолиний. Под распознаванием изолиний подразумевается:

1. Определение изолиний на изображении карты.
2. Векторизация каждой определенной линии – определение ее как последовательность координат точек  $[(x_0, y_0), (x_1, y_1), \dots, (x_N, y_N)]$ . Векторизация линии позволяет экспорт полученных линий в общепринятых форматах данных геоинформационных систем, в частности, в GeoJSON.

3. Определение глубины залегания каждой линии. Часть линий подписана числами на картах, однако, часть не имеет обозначений.

Алгоритм извлечения изолиний состоит из следующих шагов:

1. Выделение линий с помощью алгоритма сегментации на основе нейронной сети.
2. Векторизация линий с целью получить векторное представление линий выделенного слоя.
3. Объединение частей линий в изолинии, а также связь линий и чисел (значений глубин).
4. Оценка свойств отдельных линий и кластеризация по цвету и толщине линий.

### Исходный подход к извлечению изолиний

Изначально, алгоритм сегментации изолиний не был завязан на использование нейронной сети, а основывался на использовании доработанного алгоритма кластеризации KMeans для выделения линий на изображении по цвету. Кластеризация проводилась итеративно, на каждом шаге выбранная область изображения разделялась на 2 кластера по цвету. Кластеры, имеющие низкую дисперсию значений, удовлетворяли условию и удалялись из очереди на кластеризацию. Кластеры, имеющие высокую дисперсию значений пикселей изображения, кластеризовались снова (на 2

кластера). Такой подход позволял разбивать черно-белые изображения на 2 кластера – фон и линии/значения, а более-современные структурные карты, в том числе с градиентной заливкой областей, на десятки кластеров, среди которых были кластеры линий.

Для удаления неподходящих кластеров – т.е., кластеров, не имеющих линий, был обучен бинарный классификатор. На вход классификатора (на основе сверточной нейронной сети) поступало изображение кластера, результатом работы являлся класс кластера – полезный кластер с линией или кластер без линий. Поскольку каждое изображение карты разделялось на 2-64 (случаи, когда кластеров было больше, редки), а ответом фактически является 0 или 1, выборка для обучения классификатора была собрана за многократно меньшее время, чем время, необходимое для разметки реальных данных или создания генератора правдоподобных карт для обучения. Полученная модель глубокого обучения позволяла удалять значительную часть «мусорных» кластеров, не добавлять их в БД и не показывать в пользовательском интерфейсе.

Использование данной кластеризации в совокупности с классификатором кластеров позволяло обойтись без обучающей выборки для сегментации линий и давало приемлемый уровень качества на имеющихся картах, однако, качество кластеризации (т.е., сегментации изображений без учителя) снижалось на сканированных бумажных картах в связи с тем, что области изображения больше не были однородными или градиентными, как на изначально цифровых картах.

Появление сканированных карт потребовало переработки имеющегося алгоритма извлечения изолиний в сторону более робастных к изменениям входных данных алгоритмов.

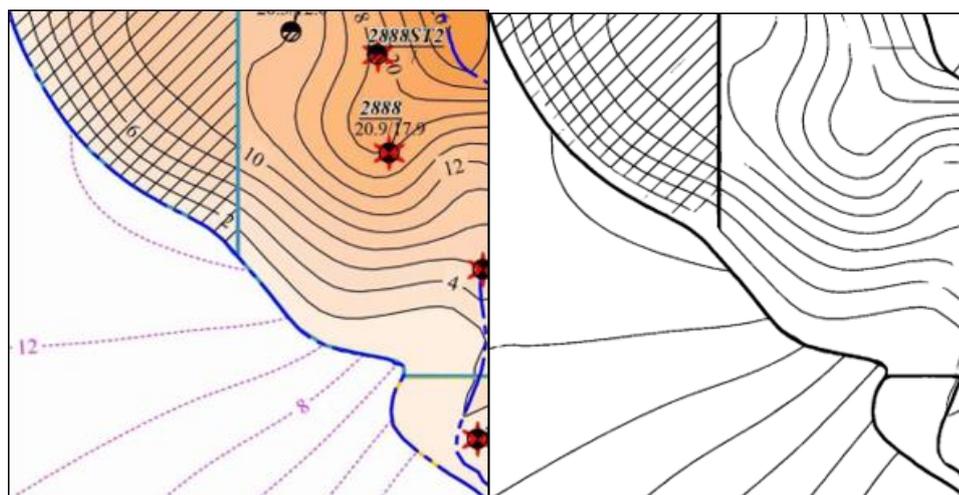
#### Извлечение изолиний

Для определения линий на изображении, а также отделения их от шума используется нейронная сеть для бинарной сегментации изображений. Несмотря на то, что существуют алгоритмы, позволяющие напрямую получать

векторные линии по растровым изображениям, такие алгоритмы не получили широкого распространения и развития, редко используются и сравниваются в релевантных задачах в литературе. Схожей задачей является определение дорог на спутниковых снимках – в такой задаче стандартом на данный момент является использование нейронной сети для бинарной сегментации с последующей пост-обработкой полученных масок дорог. Поэтому и в задаче определения изолиний был использован схожий подход на основе бинарной сегментации.

Несмотря на то, что на каждом изображении могут присутствовать несколько классов изолиний и других линий, один и тот же тип линий может выглядеть совершенно по-разному на различных картах. В связи с этим от разработки генератора карт с различными классами изолиний было решено отказаться, так как одно и то же обозначение может указывать на различные классы изолиний в зависимости от карты.

В будущем, если будет указан конкретный стиль карт, которые требуется анализировать, и если генератор карт будет доработан, то алгоритм бинарной сегментации может быть заменен на сегментацию на несколько классов или даже на инстанс-сегментацию линий. Инстанс-сегментация не только определяет изображения-маски областей линий, но и может разделять линии одного класса между собой. Инстанс-сегментация не использовалась в данной системе, поскольку ее значительно сложнее обучить с высокой точностью распознавания, чем бинарную сегментацию изображений. Пример применения такой сети к фрагменту изображения карты приведен на рисунке 19.



(a)

(б)

Рисунок 19 – выделение линий на изображении:

(a) – изображение до обработки,

(б) – выявленные линии

Выделение изолиний производится сетью, имеющей структуру U-Net [8], однако в качестве базовых блоков используются блоки Xception [9]. Структура сети изображена на рисунке 20.

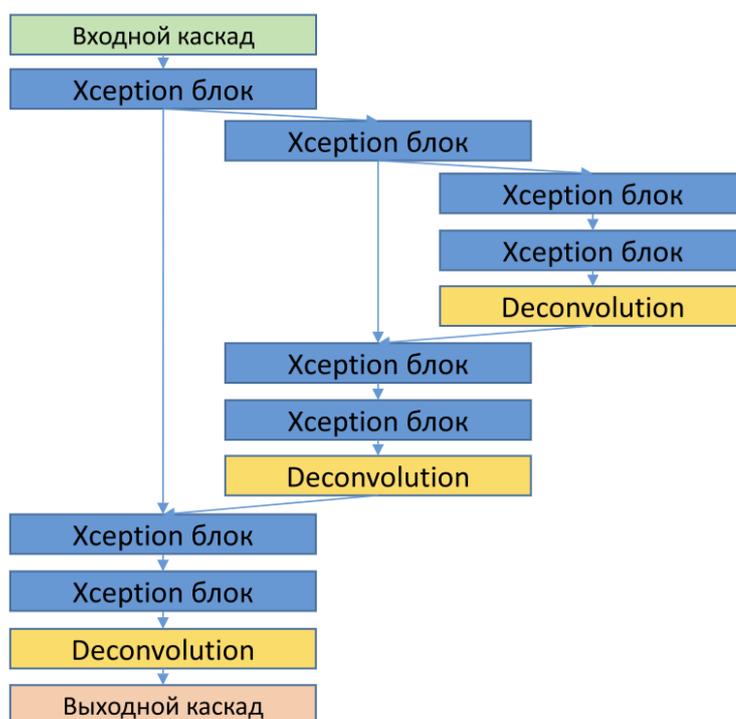
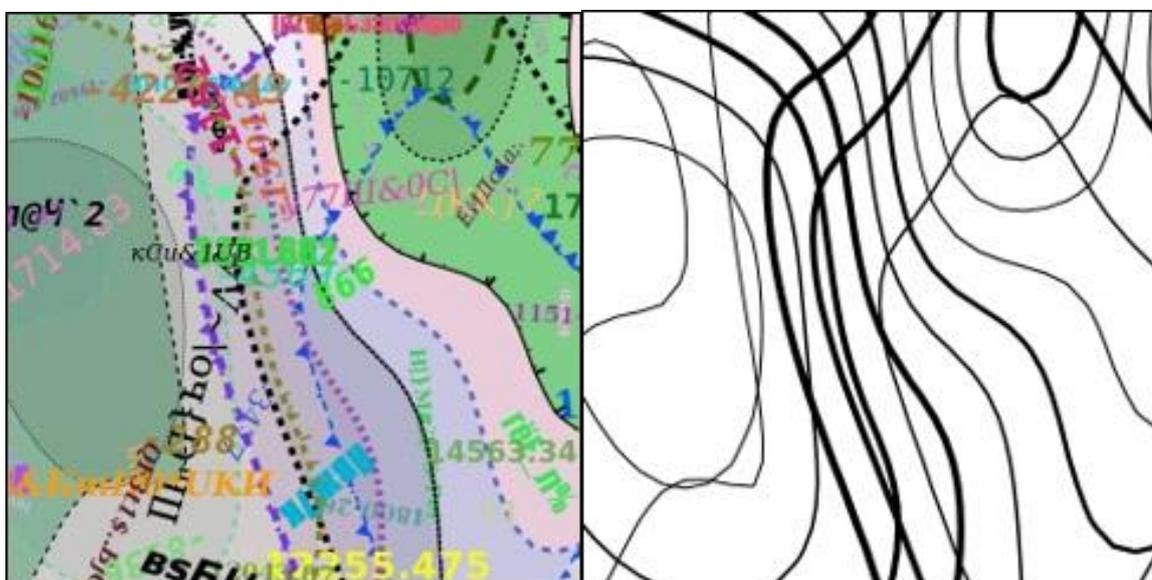


Рисунок 20 – структура сети для выделения линий.

Сеть применяется последовательно к частям изображения размером 256x256, затем результат объединяется.

Обучение сети выполнялось на полностью синтетической выборке. Такой подход использован в связи с большим разнообразием и одновременно маленьким размером доступной выборки. Чем более разнообразные данные требуется анализировать, тем больше данных требуется для обучения. Однако, реальных данных было недостаточно, даже если разметить их полностью. В то же время, ручная разметка изолиний – долгий и трудоемкий процесс, разметка даже одной карты может потребовать несколько часов работы. Пример элемента обучающей выборки на рисунке 21.



(a)

(б)

Рисунок 21 – выделение линий на изображении:

(a) – синтетическое изображение,

(б) – ожидаемый результат работы детектора.

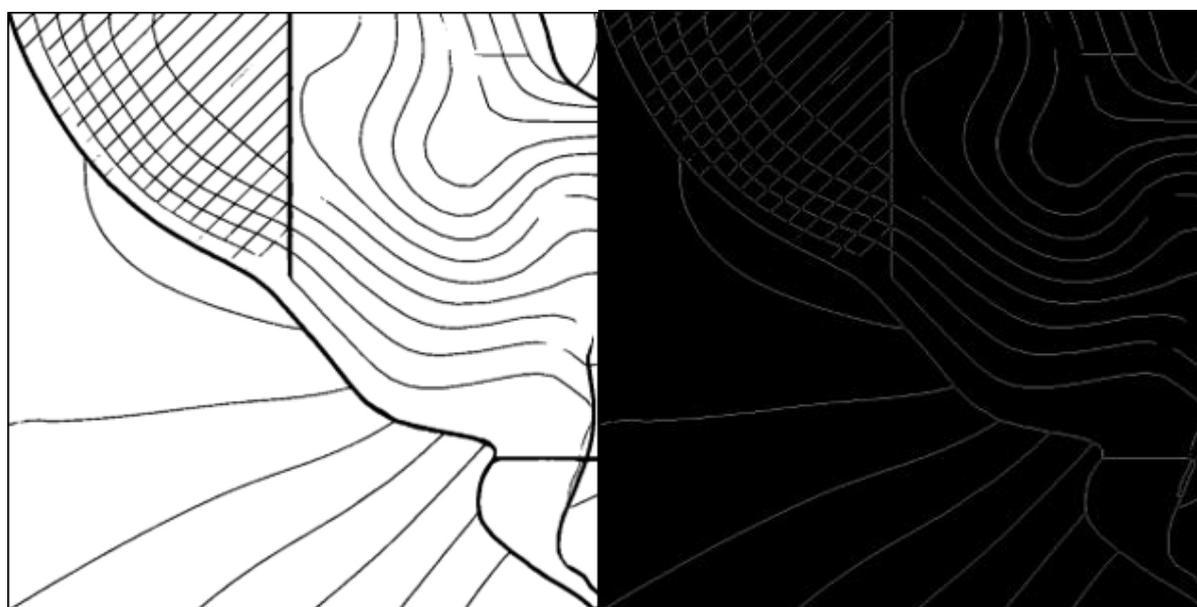
Оценка сегментации производилась на синтетической тестовой выборке (в связи со сложностью сбора ручной разметки). Метрикой качества является IOU (intersection over union, отношение пересечения предсказанной области и истинной к объединению данных областей). Была получена оценка

качества 0.88 IOU. Истинные линии имеют различную толщину, как правило, несколько пикселей.

Использование бинарной сегментации позволило получить схожий с кластеризацией уровень качества на цифровых картах и значительно улучшить точность на сканированных картах.

#### Векторизация линий

В результате выделения изолиний получают бинарные изображения. Пример изображения, подходящего на векторизацию, представлен на рисунке 22 (а). Результат работы алгоритма векторизации представлен на рисунке 22 (б).



(а)

(б)

Рисунок 22 – выделение фрагментов линий на изображении:

(а) – бинаризованное изображение до обработки,

(б) – выявленные фрагменты линий

На рисунке 10 (а) представлено бинаризованное изображение изолиний одного цвета (при этом они могут быть разного типа). Необходимо извлечь все точки, относящиеся к одной линии. В силу того, что линии могут быть разорваны на предыдущем этапе и могут пересекаться линии разных типов,

извлекаться будут фрагменты линий. Их соединение происходит на следующем этапе. Кроме того, стоит отметить, что линии имеют толщину, обусловленную их отрисовкой на исходном растровом изображении.

Этап извлечения векторизованных фрагментов линий имеет следующий алгоритм:

1. Скелетизация бинаризованного изображения. Алгоритм скелетизации [7] подразумевает извлечение срединной линии. Линии после скелетизации имеют толщину в 1 пиксель.

2. Разрыв линий в местах точек пересечения. При бинарной сегментации несколько различных изолиний могли быть объединены, в связи с этим их нужно разорвать и в дальнейшем объединить на основе их направления.

3. Уменьшение шума изображения. Удаление слишком коротких линий – линии длиной в несколько пикселей не могут быть точно объединены в изолинии на следующем этапе алгоритма.

4. Извлекаются непрерывные упорядоченные последовательности точек. Каждая непрерывная последовательность точек, представляет собой набор точек. Таким образом, был совершен переход от бинарной маски, позволяющей определить области линий на изображении, к массивам координат отдельных линий (но не изолиний, поскольку изолинии могут состоять из нескольких линий).

Результатом работы является массив фрагментов линий, извлечённых из изображения. Каждый фрагмент представлен упорядоченным массивом точек, относящихся к этому фрагменту.

#### Объединение частей линий

Полученные на предыдущем этапе линии требуется объединить в изолинии. Также требуется связать полученные изолинии со значениями их глубины.

Первым шагом перед объединением частей линий является их упрощение – уменьшение количества точек, приходящихся на каждую линию.

На предыдущем этапе массив точек линий содержит избыточное количество точек – по точке на пиксель. Такая высокая точность не требуется конечному пользователю, а также может привести к снижению производительности работы модуля редактирования результатов работы аналитического модуля в интерфейсе пользователя (в веб-браузере). Также, большое количество точек увеличивает время работы алгоритма.

Для упрощения линий используется алгоритм Рамера – Дугласа – Пекера. Это алгоритм, позволяющий уменьшить число точек кривой, аппроксимированной большей серией точек. Алгоритм был независимо открыт Урсом Рамером в 1972 и Давидом Дугласом и Томасом Пекером в 1973. Также алгоритм известен под следующими именами: алгоритм Рамера – Дугласа – Пекера, алгоритм итеративной ближайшей точки и алгоритм разбиения и слияния.

Суть алгоритма состоит в том, чтобы по данной ломаной, аппроксимирующей кривую, построить ломаную с меньшим числом точек. Алгоритм определяет расхождение, которое вычисляется по максимальному расстоянию между исходной и упрощённой кривыми. Упрощенная кривая состоит из подмножества точек, которые определяются из исходной кривой (рисунок 23).

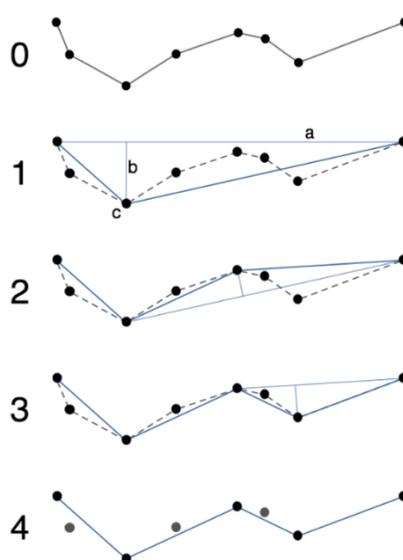


Рисунок 23 – Пример работы алгоритма Рамера – Дугласа – Пекера

Упрощение линий позволяет многократно уменьшить число точек линий без заметных для конечного пользователя искажений кривых.

Следующим шагом является объединение линий в изолинии. Линии объединяются на основе использования эвристических методов. Для объединения линий рассчитывается матрица функции расстояния между каждым концом каждой линии. Таким образом, если на изображении присутствует 5000 линий, полученная матрица имеет размер 10000x10000 значений. Функция расстояния между концами одной линии не рассчитывается, поскольку не имеет смысла объединять линию с ней самой же. Вычисление матрицы векторизовано и основано на использовании библиотеки NumPy для векторных и матричных вычислений. Это позволило ускорить расчет для изображения среднего размера с нескольких минут до нескольких секунд. Чтобы снизить потребность в RAM, данная матрица рассчитывается по частям таким образом, чтобы число элементов, рассчитываемых одновременно, не превышало заданный порог, обусловленный ограничением на использование RAM.

Функция расстояния заключается в расчете нескольких величин:

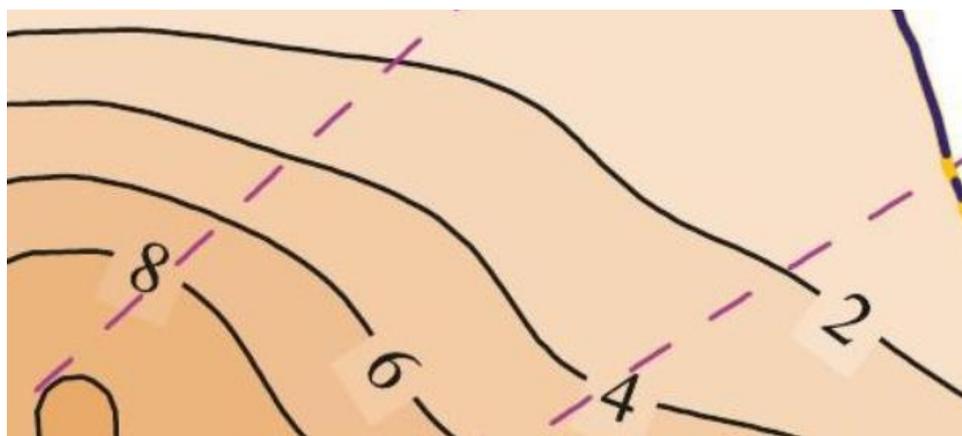
- Расстояние между концами линий (расстояние в пикселях делится на квадратный корень из площади изображения, таким образом, расстояние учитывает размеры изображения).
- Разность углов направления данных концов линий. Угол направления конца линии определяется по взвешенному среднему направлений последних нескольких отрезков между точками данного конца линии. Такой подход обусловлен тем, что определения направления только по последней точке и предпоследней или одной из последних на практике привело к недостаточно высокой точности определения угла направления линии. Концы линий более вероятно будут объединены, если разность углов направления данных концов близка к 180 градусам – т.е., концы линий направлены друг на друга.

- Средняя разность направлений концов линий и направления вектора, соединяющего данный конец линий (рассчитываемый в данный момент) с концом другой линии. Разница между направлением конца линии и направлением такого соединительного вектора должна быть небольшой.

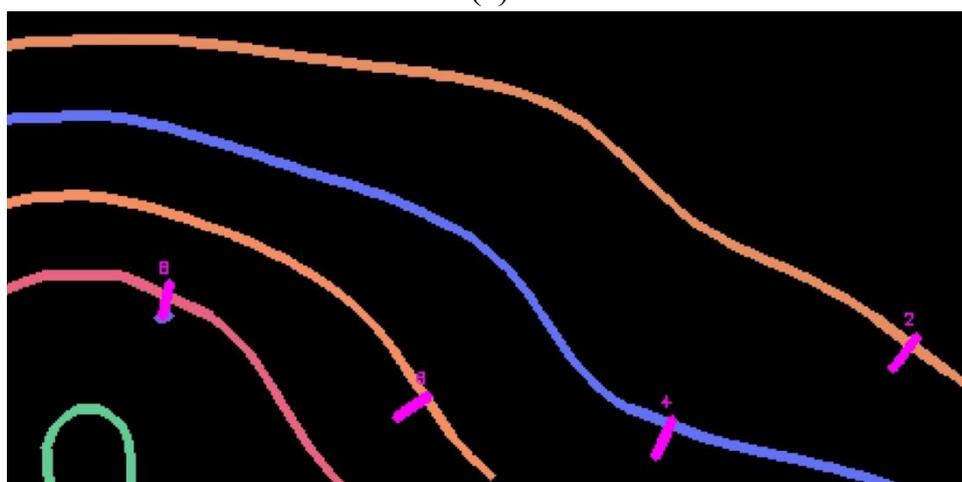
Взвешенная сумма полученных оценок является общим расстоянием между двумя концами двух линий.

Важным шагом после расчета матрицы расстояния является последовательное объединение линий. Концы линий с небольшим значением функции расстояния между ними должны быть объединены, однако, каждый конец линии должен быть объединен только с одним концом другой линии – поскольку изолиния является группой последовательно соединенных линий. Поэтому объединение линий происходит последовательно, начиная с линий с минимальным значением функции расстояния. Если два конца линии уже были объединены, то они не могут соединиться с какими-либо еще концами линии.

Объединение чисел с изолиниями также происходит на этапе объединения линий в изолинии. На представленной Заказчиком выборке, изолинии, если они подписаны, то всегда подписаны цифрами поверх изолиний. Поэтому, для чисел была рассчитана средняя линия – отрезок, соединяющий центральную точку на левой стороне ограничивающей рамки с центральной точкой правой стороны ограничивающей рамки (с учетом направления цифр, которое также определяется нейронной сетью в алгоритме распознавания цифр). Данный отрезок добавляется в массив линий для объединения в изолинии и числа становятся частью изолинии также, как другие линии. Так как несколько разных чисел может объединиться с одной линией, конечным значением глубины изолинии является мода распознанных значений (за исключением некорректно распознанных значений глубины, которые не могут быть преобразованы в число). Результат связывания чисел и линий представлен на рисунке 24.



(a)



(б)

Рисунок 24 – Пример выделенных изолиний с соответствующими значениями глубин: (а) – исходное изображение; (б) – обработанное изображение

Кластеризация по цвету и толщине линий.

На последнем этапе распознавания изолиний они разделяются на разные типы на основе их толщины и цвета. В связи с различием обозначений линией от карты к карте, конкретный класс линий назначить не представляется возможным, если аналитический модуль должен быть универсален к типу структурных карт.

Оценка свойств линий производится путем анализа содержания областей, обнаруженных на этапе выделения изолиний. Пример таких областей приведен на рисунке 25.

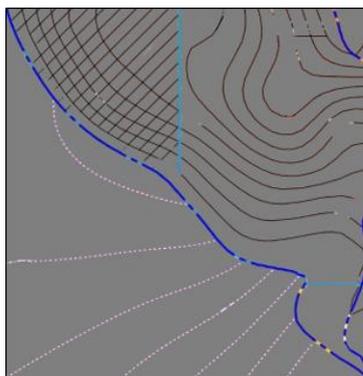


Рисунок 25 – области, содержащие изолинии

Каждой линии, выделенной на предыдущем этапе, определяется соответствующая ей область на исходной карте. Для этой области вычисляются следующие характеристики: средний цвет и средняя ширина линии. Далее выполняется кластеризация линий методом KMeans [10] в 4х мерном пространстве свойств - (R, G, B, Ширина). Поскольку ширина, как правило, изменяется в меньших пределах, чем значения красного, зеленого и желтого каналов цвета, для ширины был подобран коэффициент, позволяющий добиться наилучшего качества разделения типов линий.

Результат работы алгоритма представлен на рисунке 26. Различные типы линий выделены разным цветом. Разделение линий по типу позволяет отобразить разные типы линий на разных слоях для пользователя.

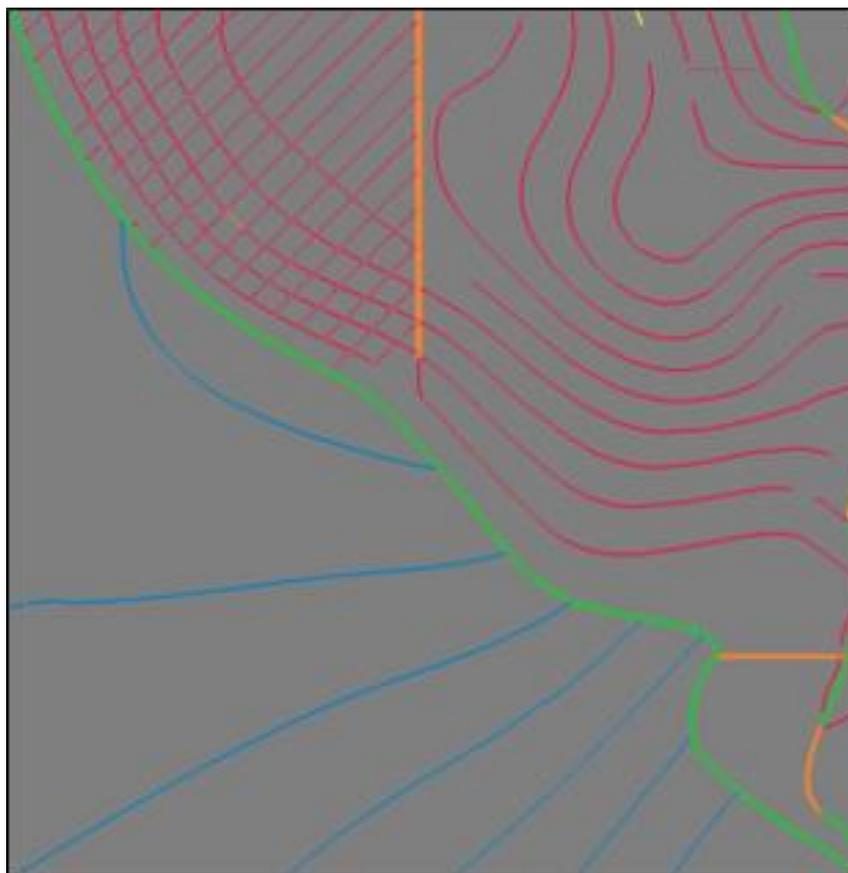


Рисунок 26 - результат кластеризации линий, каждый кластер линий  
отображен своим цветом

### **Сегментация изображений для определения областей карт**

Отдельным алгоритмом, не связанным с другими методами анализа в аналитическом модуле, является алгоритм бинарной сегментации изображений, обученным определять области карт.

Определение областей карт позволяет автоматически предлагать пользователю зоны на изображении, которые имеет смысл анализировать. Кроме карт изображения также могут содержать таблицы, подписи чертежа, заметки и пр. Подобная информация на изображении приводит к увеличению времени обработки, а также к повышению количества ошибок в ходе работы алгоритмов.

На рисунках 27-29 представлены несколько примеры различных структурных карт.



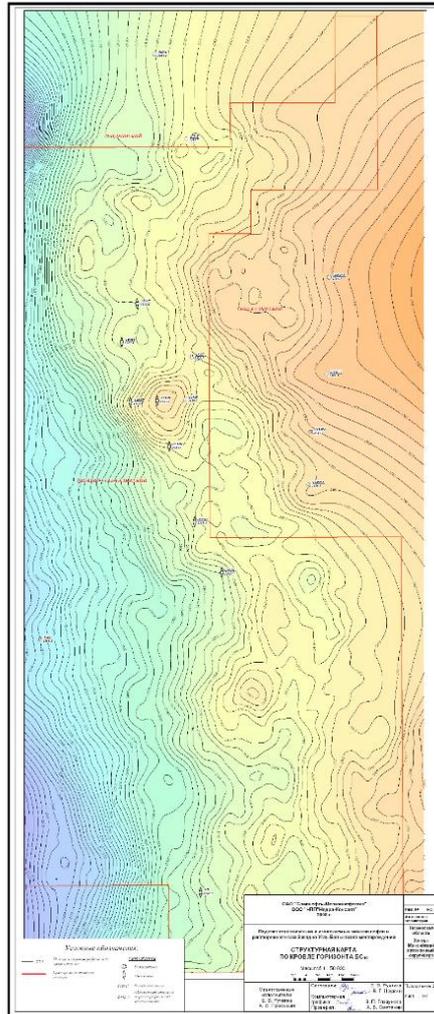


Рисунок 29 – пример карты

В связи с тем, что карты в выборке сильно отличаются друг от друга, использование нейронной сети для решения задачи бинарной сегментации является потенциально более надежным и качественным подходом, чем использование эвристик.

Фреймворком разработки был выбран Catalyst – библиотека, основанная на Pytorch. Это open-source библиотека для ускорения разработки алгоритмов машинного обучения в области глубокого обучения и обучения с подкреплением.

Для быстрой разметки данных для задачи сегментации был использован сервис Supervisely – один из крупнейших сервисов для разметки данных.

## Процесс обучения

Алгоритм бинарной сегментации предназначен для получения из изображений геологических карт (с 3 каналами RGB) бинарной маски, разделяющей два класса областей изображения – фон, не содержащий полезной информации, и область карты.

В ходе разметки данных на площадке Supervisely (рис. 30) было размечено 1429 геологических карт заказчика (приведенных к разрешению 1024x1024), однако, многие из изображений повторялись не один раз с небольшими изменениями – в связи с этим при обучении применялся ряд аугментаций.

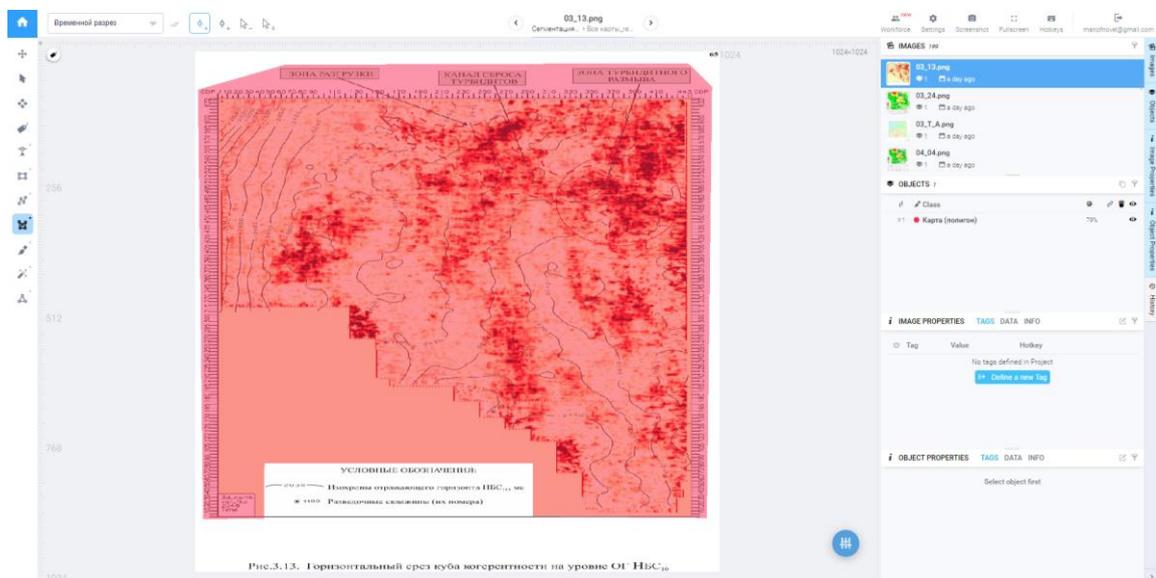


Рисунок 30 – пример размеченного изображения в сервисе Supervisely

В выборке присутствовали как современные карты, так и более старые карты, созданные в СССР.

Определение областей карт решается как задача сегментации, а не детекции, поскольку карты зачастую имеют сложную форму, далекую от прямоугольной. Несмотря на то, что на данный момент пользователю в системе предлагаются прямоугольные области для анализа, а следовательно, можно использовать алгоритм для детекции, выбор был сделан в сторону сегментации, так как в будущем возможно добавление в систему функции выбора области для анализа произвольной формы.

Обучающая выборка подвергалась аугментациям – случайным изменениям в ходе обучения, позволяющим разнообразить данные. Использовался достаточно «сильный» набор аугментаций для каждого изображения:

- С вероятностью 50%: случайный поворот до 90 градусов.
- С вероятностью 50%: вырез случайных небольших прямоугольных областей.
- С вероятностью 30%: небольшие изменения контрастности и яркости.
- С вероятностью 30%: grid distortion – нарезка изображения на клетки и случайные искажения полученной сетки.
- С вероятностью 30%: добавление шума к HSV значениям пикселей.

Использовалась сеть архитектуры ResnetFPUNet – сеть, объединяющую идеи Feature Pyramid Network, U-Net и ResNet. Слои энкодера инициализированы на основе слоев предобученной на выборке ImageNet сети ResNet-34.

Для обучения использовался оптимизатор RAdam с Lookahead [3]. Коэффициент скорости обучения уменьшался в 4 раза, если в ходе обучения значение функции ошибок на проверочной выборке не уменьшался две эпохи подряд.

Функция ошибок составляется как сумма трех функций: Dice, IoU и binary crossentropy (последняя функция с коэффициентом 0.8). Критерием остановки обучения являлся выход IoU на плато (5 эпох без улучшения результата) на проверочной выборке.

В ходе проверки на проверочной подвыборке, которая состояла из 20%, которые не использовались в обучении (а также не имели практически идентичных «двойников», нередко попадавших в обучающей выборке), метрика IoU достигла значений в 0.98. Данная метрика рассчитывается как среднее отношение площади пересечения истинной и предсказанной области к их объединению. Высокое значение метрики может быть обусловлено тем,

что большая часть изображений имеет достаточно легко определяемые 1-3 большие области карты.

#### Процесс предсказания на новых данных

После обучения нейронной сети на фреймворке Catalyst, обученные веса сети были извлечены из контрольной точки Catalyst, чтобы снизить зависимость от данной библиотеки (которая находится на стадии активной разработки). Новые изображения перед сегментацией преобразуются к нужному разрешению – 1024x1024 (обучающая выборка была размечена в таком разрешении). Использование большего разрешения в ходе обучения могло привести к снижению качества предсказаний и необходимости в разметке большего числа данных, использование меньшего разрешения могло заметно снизить качество сегментации, поскольку многие изображения имеют относительно небольшие, но важные детали. Такие детали было бы невозможно определить на изображении маленького разрешения не только нейронной сети, но и человеку.

Обобщающую способность алгоритма можно оценить по результатам на случайно подобранных геологических изображениях из сети Интернет (рис. 31, 32, 33).

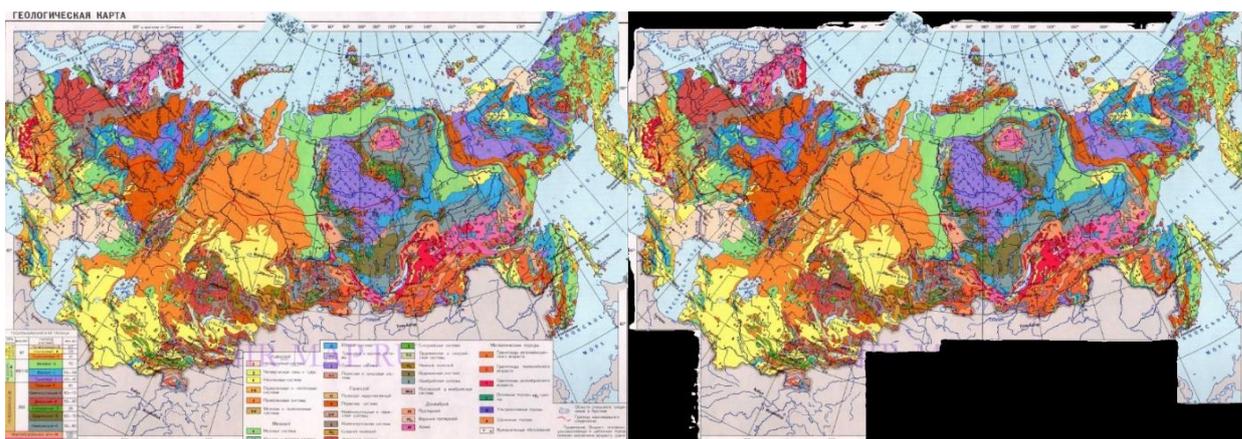


Рисунок 31 – Результат работы алгоритма

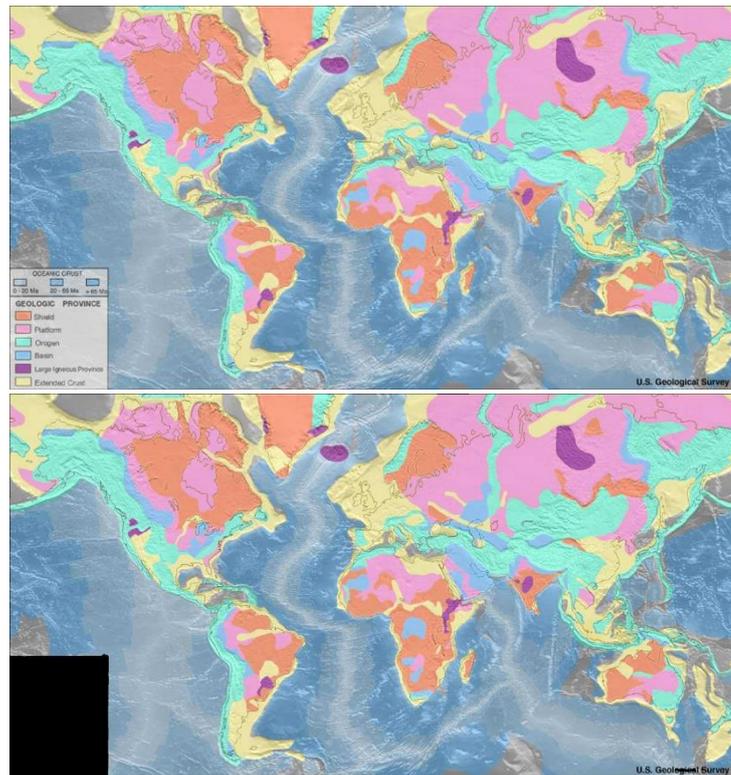


Рисунок 32 – результат работы алгоритма



Рисунок 33 – результат работы алгоритма (изображение осталось практически неизменным, поскольку карта занимает все изображение целиком)

## МОДУЛЬ ДОРАБОТКИ РЕЗУЛЬТАТОВ АНАЛИЗА

### Интерфейс системы

Пользовательский интерфейс (frontend) веб-приложения реализован на фреймворке Vue.js. Интерфейс имеет следующие пользовательские функции:

1. Аутентификация пользователей.
2. Загрузка карты для анализа.
3. Поиск имеющихся карт в каталоге, а также отображение их на карте мира.
4. Страница просмотра результатов анализа и их доработка.
5. Выгрузка результатов в формате GeoJSON.

Необходимость в возможности доработки результатов распознавания обусловлена тем, что целью системы не является полная замена человека в процессе распознавания скважин и изолиний, а только автоматизация данного процесса. Другой причиной является неизбежность ошибок в ходе работе алгоритмов распознавания на большом количестве различных изображений карт.

Главная страница с картой приложения представлена на рисунках 34 и 35.

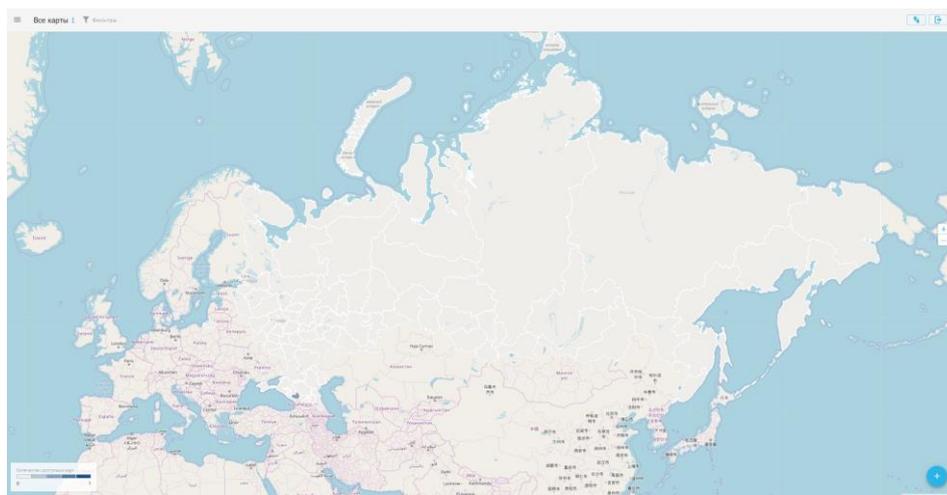


Рисунок 34 – Главная страница приложения с закрытым каталогом

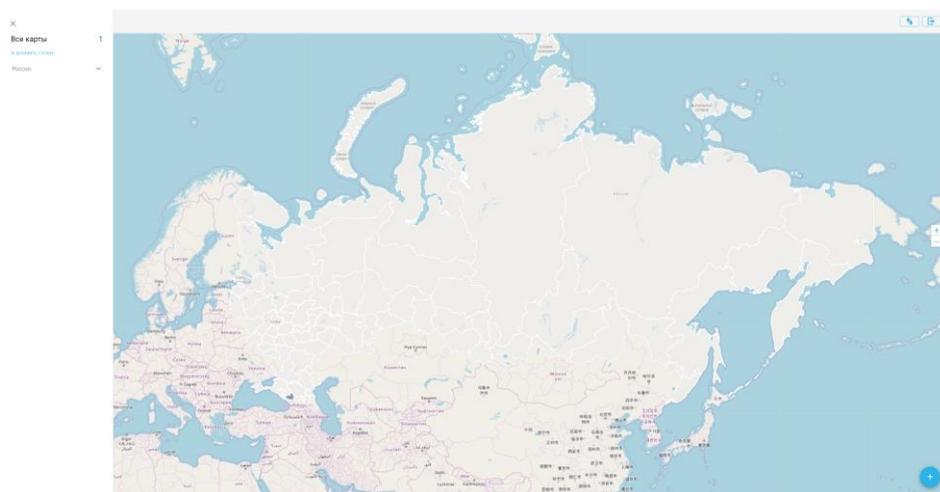


Рисунок 35 – Главная страница приложения с открытым каталогом карт

В левой части интерфейса находится каталог загруженных карт с разделением по странам и регионам, также там находится кнопка фильтрации каталога. В правом верхнем углу пользователю доступны кнопки для экспорта карт и аутентификации. Только аутентифицированный пользователь может пользоваться системой. В правом нижнем углу находится кнопка для загрузки карт. В левом нижнем углу – легенда расцветки регионов РФ по количеству загруженных карт.

На рисунке 36 представлена страница загрузки карт. Загрузка карт производится в 36 этапа, сначала пользователь должен загрузить файл карты, затем указать информацию о карте, необходимую, в том числе, для помещения ее в правильный раздел каталога, после же требуется выбрать прямоугольную область для анализа на карте.

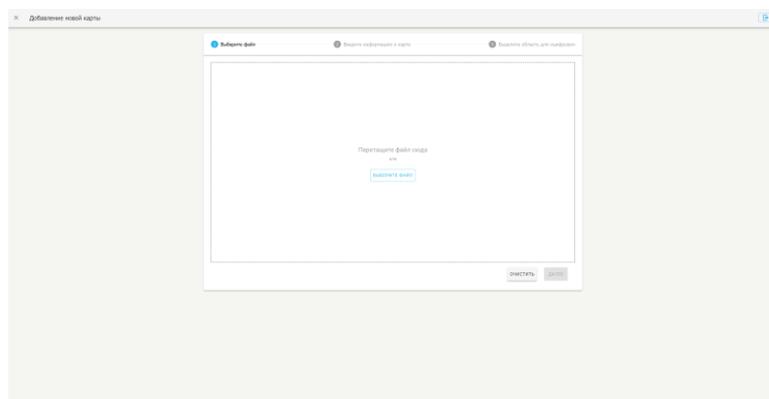


Рисунок 36 – Страница загрузки карты

## Разработка модуля доработки результатов анализа

### Интерфейс модуля доработки

Каталог карт позволяет перейти к просмотру отдельной карты. Страница просмотра карты позволяет редактировать результаты анализа изображения. После исправления ошибок карту можно экспортировать в формате GeoJSON. На рисунке 37 представлено изображение интерфейса страницы редактирования результатов.



Рисунок 37 – Страница редактирования результатов анализа

В левой части страницы редактирования представлены слои изолиний и скважин, доступные для просмотра на карте. Слоям можно изменить название или удалить их целиком. Изолинии на карте могут быть отфильтрованы по длине и по глубине изолинии, а также по наличию значения глубины.

В правом нижнем углу находится легенда значений изолиний. Цвет изолинии определяется значением ее глубины (от фиолетового к зеленому). Красная часть спектра не используются, поскольку красный цвет имеют выбранные пользователем линии. На рисунке 38 представлена страница редактирования карты с отображением распознанных изолиний и скважин. На рисунке 39 более подробно представлена распознанная карта. На данном

изображении черные изолинии – это изолинии без распознанных значений глубины. Таким изолиниям требуется назначить глубину вручную. Корректность значений других изолиний можно установить по сравнению подписи со значением на карте, а также по цвету – ошибочные значения наиболее часто отличаются от корректных на порядки, благодаря чему их цвет резким образом выделяется на фоне корректных изолиний.



Рисунок 38 – Страница редактирования результатов анализа с отображением полученных результатов



Рисунок 39 – Результаты распознавания

В данном случае тестирование проводилось на небольшой области исходной карты, поэтому разметка такого изображения вручную не займет много времени, однако, даже на изображениях такого масштаба работа аналитического модуля позволила значительно уменьшить трудоемкость разметки.

На рисунке 40 можно увидеть верхнюю панель инструментов, доступных при редактировании результатов анализа. Функции верхней панели инструментов:

1. Функция сохранения внесенных изменений – в этом случае изменения, сделанные пользователем в интерфейсе, синхронизируются с базой данных приложения.

2. Экспорт результатов анализа.

3. Перезапуск анализа – пользователь может выбрать другую область для анализа.

4. Кнопка помощи. При нажатии на нее, пользователю показывается модальное окно (рисунок 41) с пользовательской инструкцией. Инструкция знакомит пользователей с набором доступных инструментов, а также горячими клавишами для их вызова.

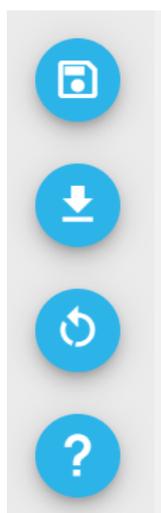


Рисунок 40 – верхняя панель инструментов карты

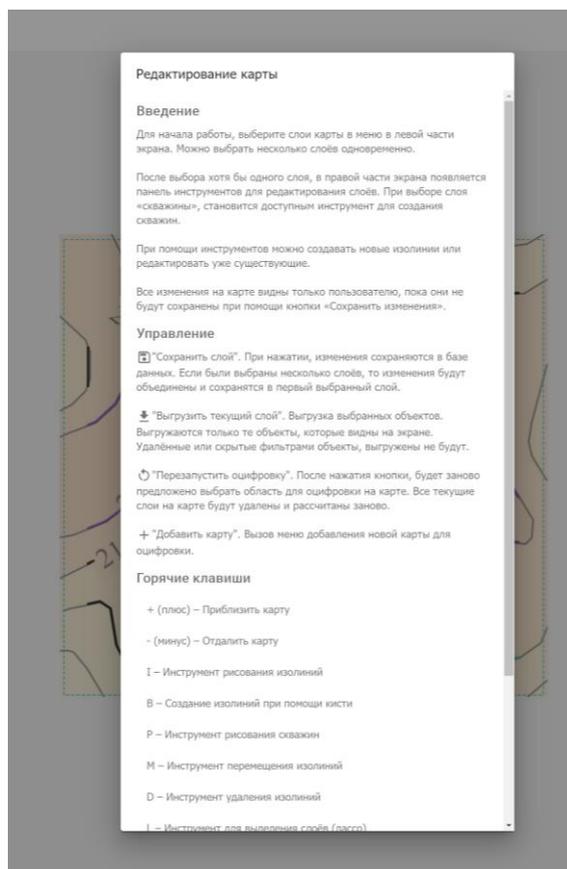


Рисунок 41 – Модальное окно с пользовательской инструкцией

На рисунке 42 изображена нижняя панель инструментов редактора результатов анализа. Нижняя панель инструментов имеет следующие функции:

1. Приближение карты.
2. Отдаление карты.
3. Инструмент создания изолиний – данный инструмент позволяет создавать изолинии путем последовательных кликов, каждый из которых добавляет точку к линии.
4. Инструмент переноса линий. Данный инструмент позволяет переносить линию или несколько выбранных линий.
5. Инструмент удаления линий. Данный инструмент предназначен для удаления линий по клику.
6. Кисть. Данный инструмент позволяет создавать изолинии движением мыши с зажатой левой кнопкой. Работа инструмента основана

вызове функции добавления точки к линии при сдвиге мыши, при этом вызов функции находится в «тормозящем» декораторе – функция не может быть вызвана чаще, чем в 150 мс. Такое ограничение задано для того, чтобы не создавать линии из огромного количества точек, которые будет неудобно редактировать и которые будут уменьшать производительность редактора. Значение ограничения в 150 мс было подобрано так, чтобы задержка была достаточно большой, чтобы выполнять свою функцию, но не настолько большой, чтобы создавать неудобства при использовании.

7. Лассо. Данный инструмент позволяет выбирать несколько линий сразу для передвижения или удаления.

8. Удаление. Данная кнопка позволяет удалить выбранные линии. Эту же функцию можно вызвать клавишей Delete.

9. Отмена. Снимает выделение с выбранных линий.

10. Назад. Позволяет отменить последние действия.

11. Вперед. Позволяет повторить действия, отмененные инструментом «Назад». Работа инструментов «Назад» и «Вперед» основана на хранении истории состояний редактора. История включает в себя до 30 состояний редактора. Данные инструменты также можно было реализовать путем создания множества функций, которые будут повторять или отменять действия тех или иных инструментов, а также созданием истории примененных инструментов. Однако, подход на основе хранения истории состояний значительно проще в реализации, а также не требует изменений при добавлении или удалении каких-либо инструментов в интерфейс. Кроме того, работу данного подхода значительно проще протестировать.

12. Ластик. Данный инструмент позволяет использовать ластик для удаления отдельных точек или линий. Размер (радиус области) ластика можно менять с помощью дополнительного слайдера, появляющегося в интерфейсе при включении ластика.



На основе исследования различных источников в сети Интернет [14, 15, 16].

Преимущества React:

1. Изолированные компоненты позволяют оптимизировать процесс обновления элементов на странице, что повышает скорость работы интерфейса.

2. Компоненты React – повторное использование компонентов позволяет снизить трудозатраты на разработку.

3. Стабильный код, основанный на передачи данных от родительских компонентов к дочерним. Передача данных из дочерних в родительские в основном выполняется с помощью функций, переданных из родительских компонентов.

4. Библиотека с открытым исходным кодом с большим количеством готовых инструментов.

Недостатки React:

1. Сложность изучения.

2. Данный фреймворк значительно популярнее других, что приводит к большому количеству обновлений и инноваций. Таким образом, не всегда можно найти актуальную документацию.

Преимущества Vue.js:

1. Проще в изучении, чем React.

2. Подробная документация.

3. Простая интеграция компонентов и повторное использование компонентов.

4. Удобное расширение разработчика для браузера, позволяющее контролировать состояние приложения. В отличие от аналогичного расширения для React, просмотр Vuex store (переменных состояния приложения, аналог Redux для React) возможно без дополнительных настроек в коде. Также, при просмотре компонентов Vue.js в данном расширении нет

огромного количества компонентов-обертки, которые затрудняют анализ структуры и поиск необходимого компонента, как в расширении для React.

Недостатки Vue.js:

1. По мнению некоторых разработчиков, Vue.js является «слишком гибким» фреймворком.
2. Сообщество разработчиков на данном фреймворке меньше, так как фреймворк относительно новый. Разработчики могут столкнуться с проблемами, на которые будет трудно найти ответ.

Angular является более старым фреймворком по сравнению с Vue и React. Он был выпущен в 2010 году.

Преимущества Angular:

1. Структура приложения основана на компонентах, как у React и Vue.js.
2. Основным языком является TypeScript, который во многих случаях упрощает разработку.
3. Высокая производительность.
4. Крупная экосистема.

Недостатки Angular:

1. Между Angular и AngularJS большая разница, поэтому перенос старых приложений занимает много времени.
2. Сложность фреймворка – код для каждого компонента и управления его жизненным циклом может быть разбит по нескольким файлам.
3. Сложность изучения.
4. Неполная документация команд консоли.

Для определения наиболее подходящего фреймворка требуется составить сравнительную таблицу. Сравнительная таблица (таблица 1) была составлена на основе информации из исследованных источников [14, 15, 16].

Таблица 1. Сравнительная таблица веб-фреймворков

Критерий	Вес	Vue.js	React	Angular
Производительность	0,10	4	5	5
Гибкость	0,15	5	5	4
Количество готовых инструментов	0,15	3	4	5
Простота изучения	0,25	5	4	3
Опыт использования	0,35	5	4	1

На основании таблицы 1 можно рассчитать взвешенную сумму критериев (таблица 2).

Таблица 2. Взвешенная сумма критериев

Vue.js	4,6
React	4,25
Angular	2,95

На основе составленных таблиц сравнения, был сделан выбор в пользу Vue.js.

#### Реализация интерфейса

Для ускорения разработки небольшого приложения использовалась также библиотека Vuetify. Данная библиотека содержит значительное количество UI компонентов, которые можно стилизовать для своего приложения.

Модуль редактирования результатов анализа изображений основан на библиотеке Leaflet.js и плагине Leaflet-Geoman к ней.

Библиотека Leaflet.js – ведущая JavaScript библиотека с открытым исходным кодом для создания интерактивных карт. Данная библиотека используется в приложении для создания глобальной карты на главной странице приложения, а также для отображения проанализированных карт.

Leaflet-Geoman – библиотека, добавляющая в Leaflet возможность создания и редактирования слоев с полигонами, ломаными линиями, окружностями и т.д. Данная библиотека лучше всего подходила для реализации требуемых инструментов, однако, в ней из необходимых функций были реализованы только создание ломаных линий путем последовательных кликов на карте, перенос линий, а также ластик.

Таким образом, кроме интеграции данной библиотеки с существующим приложением и настройки ее для работы с изображением, полученным с сервера приложения, требовалось значительно переработать ее функции, добавив следующие инструменты: инструмент удаления линий по клику, кисть, лассо, удаление выделенных линий, отмена выбора линий, функции «Назад» и «Вперед». Кроме реализации этих инструментов на странице редактора также были добавлена поддержка горячих клавиш для них, а также автоматическое отключение выбранного инструмента при выборе другого – по какой-то причине такой функции не было в исходной библиотеке.

Также была добавлена возможность объединять линии между собой на случай разрывов между ними путем выбора одного конца линии и клике на конец другой линии. Такая функция позволяет не создавать заново разорванную в каком-либо месте линию – разрывы в линиях наиболее часто возникают из-за слишком большого промежутка в линии, занятого значением глубины.

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

Студенту:

Группа	ФИО
8ПМ8И	Чугунов Роман Анварович

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

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

4. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	- Оклад инженера – 21760; - Оклад руководителя от организации – 35000; - Оклад научного руководителя – 35120;
5. Нормы и нормативы расходования ресурсов	- Годовая норма амортизации составляет 33.3 %.
6. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	- Размер страховых взносов равный 30%; - Районный коэффициент г. Томск.

### Перечень вопросов, подлежащих исследованию, проектированию и разработке:

5. Оценка коммерческого и инновационного потенциала НТИ	- Анализ потенциальных потребителей; - Диаграмма Исикавы; - Анализ конкурентных технических решений.
6. Разработка устава научно-технического проекта	Постановка целей проекта, определение ожидаемого результата и критериев успешности проекта.
7. Планирование процесса управления НТИ: структура и график проведения, бюджет, риски и организация закупок	- Построение диаграммы Ганта; - Планирование работ; - Формирование бюджета затрат; - Анализ рисков проекта.
8. Определение ресурсной, финансовой, экономической эффективности	Расчет финансового показателя эффективности. Описание общих выводов.

### Перечень графического материала (с точным указанием обязательных чертежей):

1. Сегментация рынка
2. Результаты оценки конкурентных систем анализа картографической информации
3. Диаграмма Исикавы
4. Календарный график проекта
5. Бюджет затрат
6. Реестр рисков

Дата выдачи задания для раздела по линейному графику

### Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Меньшикова Екатерина Валентиновна	Кандидат философских наук		

### Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ПМ8И	Чугунов Роман Анварович		

## Введение

Целью дипломной работы является разработка некоторых модулей системы анализа геологических карт. Данная система предназначена для автоматизации процесса создания трехмерных структурных моделей на основе геологических карт, а также представлении разнородных геологических карт в единой системе, доступной пользователям через веб-интерфейс. Значительное количество картографической информации нефтегазовых компаний хранится в виде чертежей на бумаге, что несет за собой ряд проблем:

- Невозможность быстрого поиска карт по различным критериям: по региону, по дате создания, по схожести между собой, по создателю карты и т.д.;
- Различия в представлении различных карт;
- Проблемы в отслеживании актуальности карт;
- Сложность сравнения объектов, представленных на двух разных картах;

Система включает в себя следующие модули:

1. Серверная часть (backend) – взаимодействие с базой данных, аналитической частью, а также ответы за запросы клиентской части приложения.
2. Клиентская часть (frontend) – веб-интерфейс пользователя, который включает функции просмотра и редактирования имеющихся карт, добавления новых, их автоматического анализа на основе алгоритмов компьютерного зрения, а также экспорта результатов в формате GeoJSON.
3. Аналитическая часть – модули, целью которых является извлечение полезной информации из карт, такой как: наличие и координаты нефтяных скважин, различные виды изолиний и их глубина.

Конкретной целью данной дипломной работы является разработка аналитической части системы, а также одного из модулей клиентской части – модуля редактирования результатов аналитической части.

Таким образом, выполняемая работа соответствует работе разработчика программного обеспечения.

Целью данного раздела является определение перспективности и успешности разрабатываемой системы анализа геологических карт. Для достижения этой цели были поставлены следующие задачи:

1. Определить потенциальных потребителей системы
2. Провести анализ конкурентных решений
3. Проанализировать факторы и определить связанные с ними проблемы, используя диаграмму Исикавы

4. Определить цель и результат проекта, составить рабочую группу проекта, а также зафиксировать финансовые ресурсы
5. Спланировать работу, распределить задачи между участниками проекта и определить трудоемкость
6. Сформировать бюджет проекта
7. Провести анализ рисков

## 1. Предпроектный анализ

### 1.1 Потенциальные потребители результатов исследования

Для определения потенциальных потребителей требуется определить целевой рынок и произвести его сегментирование. Для того, чтобы определить организации, которым необходима данная разработка, было проведено сегментирование целевого рынка. Сегментация проводилась по следующим критериям: размер организации и потребность в разрабатываемой системе. Карта сегментирования представлена в таблице 1.

Таблица 1 – Сегментирование рынка

		Потребность в хранении, поиске и анализе большого количества геологических карт	
		Есть потребность	Потребность отсутствует
Размер организации	Крупные	Крупные нефтегазовые компании	
	Средние	Компании среднего размера нефтегазовой отрасли, связанные с использованием картографической информации	
	Мелкие		Небольшие организации, не обладающие значительным количеством картографической информации

Таким образом, конечными потребителями являются крупные и средние предприятия нефтегазовой отрасли, производственная деятельность которых связана с использованием геологических карт.

Примерами могут быть такие геологические компании, как ПАО Роснефть, ПАО Газпром и т.п., а также множество компаний меньшего размера.

## 1.2 Анализ конкурентных технических решений

Наиболее близким по набору функций программным обеспечением к разрабатываемой системе, при этом, широко используемым, является ПО Isoline GIS (Б<sub>к1</sub>). Таким образом, Isoline GIS является прямым конкурентом разрабатываемой системы.

Программное обеспечение Isoline обладает всеми возможностями обычной ГИС и отвечает специфическим требованиям нефтяной геологии:

- Отображение специальных картографических объектов (скважин, сейсмопрофилей, карт в изолиниях и т. п.) в соответствии с принятыми стандартами.
- Поддержка геологических объектов (наклонных скважин, сейсмических профилей и разрезов, структурных карт с разломами, 3D моделей и других).
- Функции работы с главным для геологов и геофизиков типом данных – картами в изолиниях.
- Одновременное использование ГИС как интерпретационного пакета.
- Минимальные временные затраты в ходе обмена данными с другими геологическими пакетами (импорт/экспорт).

Данный пакет обладает большим набором функций, нежели разрабатываемая система, однако, его недостатком является отсутствие функции автоматического извлечения изолиний, их значений, а также скважин, из картографических изображений. Также к преимуществам системы можно отнести то, что система имеет веб-интерфейс, а значит, ее можно использовать на любом ПК, на котором установлен современный браузер.

В таблице 2 представлены результаты сравнения разрабатываемой системы и Isoline GIS.

Таблица 2 – Результаты оценки конкурентных систем анализа картографической информации

Критерии оценки	Вес критерия	Баллы		Конкурентоспособность		
		Б <sub>ф</sub>	Б <sub>к1</sub>	К <sub>ф</sub>	К <sub>к1</sub>	
<b>Технические критерии оценки ресурсоэффективности</b>						
1. Удобство	в	0,15	3	5	0,45	0,75

использовании					
2. Практическая польза алгоритмов для конечного пользователя	0,15	5	3	0,75	0,45
3. Потребность в ресурсах памяти	0,03	5	3	0,15	0,09
4. Потенциальные возможности функционального расширения системы	0,18	5	3	0,9	0,54
5. Функциональная мощность	0,09	2	5	0,18	0,45
6. Удобство графического интерфейса	0,05	4	5	0,2	0,25
<b>Экономические критерии оценки ресурсоэффективности</b>					
1. Актуальность на рынке	0,04	5	4	0,2	0,16
2. Цена	0,2	4	3	0,8	0,6
3. Послепродажное обслуживание	0,04	3	4	0,12	0,16
4. Открытость на рынке	0,02	3	3	0,06	0,06
5. Требование сертификации разработки	0,05	5	5	0,25	0,25
<b>Итого</b>	<b>1</b>	<b>44</b>	<b>43</b>	<b>4.06</b>	<b>3.76</b>

Из таблицы сравнения собственного сервиса анализа геологических карт с конкурентным Isoline GIS, можно сделать вывод, что уязвимость конкурентных решений связана отсутствием функций автоматизации процесса анализа, а также отсутствием веб-интерфейса, позволяющим использовать систему практически с любого устройства.

Таким образом, реализация собственного сервиса является более предпочтительнее, чем использование стороннего существующего решения.

При ее проектировании и создании системы применялись современные методы для анализа изображений, с применением технологии машинного

обучения. Это является одним из преимуществ перед системами, основанными на использовании ручного труда.

### 1.3 Диаграмма Исикавы

Диаграмма причины-следствия Исикавы (Cause-and-Effect-Diagram) – это графический метод анализа, формирования причинно-следственных связей между факторами и проблемами в исследуемой проблеме или ситуации. Диаграмма представлена на рисунке 1.

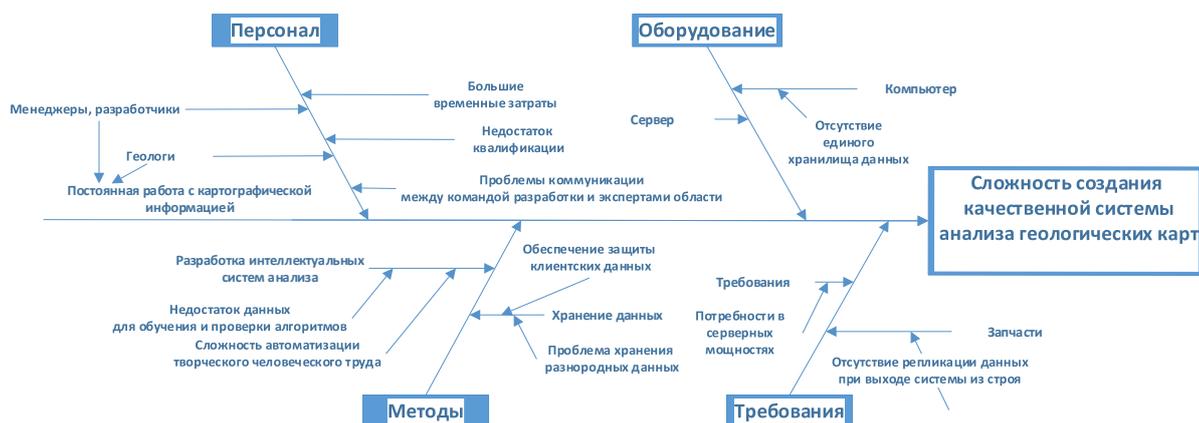


Рисунок 1 – Диаграмма Исикавы

В диаграмме Исикавы, представленной выше представлены основные проблемы, касающиеся четырех основных факторов: персонал, оборудование, методы, материалы. Стоит отметить, что сервис анализа и обработки геологических карт на сегодняшний день не является чем-то популярным в силу того, что технологии в области компьютерного зрения пока не получили широкого распространения в нефтегазовой отрасли, поэтому обучение персонала и объяснение принципов систем подобного рода является нетривиальной задачей.

## 2. Инициация проекта

В процессе инициации проекта определяются начальные цели и содержание, а также фиксируются финансовые ресурсы. Определяются внутренние и внешние заинтересованные стороны проекта, которые будут взаимодействовать и влиять на общий результат научного проекта.

### 2.1 Цели и результат проекта

В данном разделе приведена информация о заинтересованных сторонах проекта, целях проекта и критериях их достижения. Информация по заинтересованным сторонам представлена в таблице 3.

Таблица 3. Заинтересованные стороны проекта

Заинтересованные стороны	Ожидание сторон
Организация-заказчик	Готовое решение по анализу, хранению и поиску картографической информации
Пользователь	Повышение удобства и снижение времени работы с геологическими картами

В таблице 4 представлена иерархия целей проекта и критерии достижения данных целей.

Таблица 4 – Цели и результаты проекта

<b>Цели проекта:</b>	Разработка системы анализа картографической информации
<b>Ожидаемые результаты:</b>	Внедрение системы внутри организации-Заказчика
<b>Критерии приемки результата проекта:</b>	Система реализована и работает Алгоритмы работают на геологических картах Заказчика, а также имеют удовлетворительные оценки точности извлечения изолиний: – По метрикам точности на тестовой выборке данных; – По оценке пользователей
<b>Требования к результату проекта:</b>	1. Выполнены все пункты технического задания 2. Проект успешно принят Заказчиком

## 2.2 Организационная структура проекта

В таблице 5 представлена информация о рабочей группе проекта.

Таблица 5 – Рабочая группа проекта

	ФИО, место работы, должность	Роль в проекте	Функции
	Чугунов Роман Анварович, ведущий инженер по анализу данных и машинному обучению ООО Эко-Томск	Инженер	1. Проектирование 2. Реализация 3. Внедрение
	Кульневич Алексей Дмитриевич, руководитель отдела исследований и разработок ООО Эко-Томск	Руководитель проекта от организации	1. Проверка разработки 2. Помощь во внедрении
	Аксенов Сергей Владимирович, ТПУ, кандидат технических наук	Научный руководитель	1. Составление научных целей и задач 2. Проверка документации

### 2.3 Ограничения и допущения проекта

Ограничения проекта – это все факторы, которые могут послужить ограничением степени свободы участников команды проекта, а также «границы проекта» - параметры проекта или его продукта, которые не будут реализованы в рамках данного проекта. Ограничения проекта представлены в таблице 6.

Таблица 6 – Ограничения проекта

Фактор	Ограничения
Бюджет проекта	300000 рублей
Источник финансирования	ООО Эко-Томск
Сроки проекта	29.01.2020-01.06.2020
Дата утверждения плана управления проектом	29.01.2020
Дата завершения проекта	01.06.2020

В данном разделе были рассмотрены цели и ожидаемый результат, организационная структура, а также ограничения и допущения проекта. К ограничениям относятся бюджет и сроки.

## 3. Планирование проекта

### 3.1 Структура работ в рамках проекта

В данном разделе производится декомпозиция задач и разграничение зоны ответственности между участниками проекта. В таблице 7 представлено распределение исполнителей по видам работ.

Таблица 7 – Распределение задач

Основные этапы	№	Содержание работ	Должность исполнителя
Постановка задач	1	Описание требований	Руководитель от организации, научный руководитель
	2	Анализ предметной области	Руководитель от организации, научный руководитель, инженер
	3	Разработка технического задания	Руководитель от организации, инженер

Проектирование	4	Разработка архитектуры сервиса	Руководитель от организации, инженер
Разработка	5	Разработка аналитического модуля системы	Инженер
	6	Разработка веб-сервиса	Инженер
	7	Разработка модуля обработки входных данных	Инженер
Отладка решения и внедрение	8	Развертывание системы	Инженер
	9	Внедрение разработки	Руководитель от организации, инженер
Оформление документации	10	Написание отчетной документации	Инженер
	11	Проверка работы	Научный руководитель

### 3.2 Определение трудоемкости выполнения работ

Для определения ожидаемых сроков выполнения проекта необходимо оценить его трудоемкость. Воспользуемся формулой:

$$t_{ожі} = \frac{3 * t_{mini} + 2 * t_{maxi}}{5} \quad (1)$$

где  $t_{ожі}$  – ожидаемая трудоемкость выполнения  $i$ -ой работы чел.-дн.;

$t_{mini}$  – минимально возможная трудоемкость выполнения заданной  $i$ -ой работы (оптимистическая оценка: в предположении наиболее благоприятного стечения обстоятельств), чел.-дн.;

$t_{maxi}$  – максимально возможная трудоемкость выполнения заданной  $i$ -ой работы (пессимистическая оценка: в предположении наиболее неблагоприятного стечения обстоятельств), чел.-дн.

Исходя из ожидаемой трудоемкости работ, определяется продолжительность каждой работы в рабочих днях  $T_{pi}$ , учитывающая параллельность выполнения работ несколькими исполнителями:

$$t_{pi} = \frac{t_{ожі}}{Ч_i} \quad (2)$$

где  $t_{pi}$  – продолжительность одной работы, раб.дн.;

$t_{ожі}$  – ожидаемая трудоемкость выполнения одной работы, чел.-дн.;

$Ч_i$  – численность исполнителей, выполняющих одновременно одну и ту же работу на данном этапе, чел.

Для удобства составления календарного плана и графика работ необходимо перевести длительность каждого из этапов из рабочих дней в календарные дни. Для этого воспользуемся следующей формулой:

$$T_{ki} = T_{pi} * k_{\text{кал}} \quad (3)$$

где  $t_{ki}$  – продолжительность выполнения  $i$ -й работы в календарных днях;

$t_{pi}$  – продолжительность выполнения  $i$ -й работы в рабочих днях;

$k_{\text{кал}}$  – коэффициент календарности.

Коэффициент календарности определяется по следующей формуле:

$$T_{\text{кал}} = \frac{T_{\text{кал}}}{T_{\text{кал}} - T_{\text{вых}} - T_{\text{пр}}} = \frac{365}{365 - 66} = 1,22 \quad (4)$$

где  $t_{\text{кал}}$  – количество календарных дней в году;

$t_{\text{вых}}$  – количество выходных дней в году;

$t_{\text{пр}}$  – количество праздничных дней в году.

В соответствии с производственным календарем (для 6-дневной рабочей недели) в 2019 году 365 календарных дней, 299 рабочих дней, 66 выходных/праздничных дней. В таблице 8 представлены подробные временные расчеты этапов отдельных видов работ.

Таблица 8 – Временные показатели проведения научного исследования

Наименование работы	Исполнители работы	Трудоемкость работ, чел-дни			Длительность работ, дни	
		$t_{\text{min}}$	$t_{\text{max}}$	$t_{\text{ож}}$	$T_p$	$T_k$
Описание требований	Руководитель организации от	1	2	1,4	2	3
	Научный руководитель	1	2	1,4	2	3
Анализ предметной области	Руководитель организации от	2	3	2,4	3	4
	Научный руководитель	1	2	1,4	2	3
	Инженер	5	8	6,2	7	9
Разработка технического задания	Руководитель организации от	1	2	1,4	2	3
	Инженер	3	5	3,8	4	5
Разработка архитектуры сервиса	Руководитель организации от	2	4	2,8	3	4
	Инженер	12	15	13,2	14	18

Разработка аналитического модуля системы	Инженер	25	28	26,2	27	33
Разработка веб-сервиса	Инженер	7	9	7,8	8	10
Разработка модуля обработки входных данных	Инженер	7	9	7,8	8	10
Развертывание системы	Инженер	2	3	2,4	3	4
Внедрение разработки	Руководитель организации от	1	2	1,4	2	3
	Инженер	2	3	2,4	3	4
Написание отчетной документации	Инженер	45	55	49	49	60
Проверка работы	Научный руководитель	1	2	1,4	2	3

Для иллюстрирования календарного плана-графика используется диаграмма Ганта, где представлено наглядное отображение графика и распределения работ между участниками проекта. Диаграмма представлена в таблице 9.

Таблица 9 – Календарный график проекта

	Содержание работ	Должность исполнителя	Продолжительность выполнения работ													
			Февр.		Март			Апрель			Май			Июнь		
1	Описание требований	РО, НР	■	■												
2	Анализ предметной области	РО, НР, И	■	■	■											
3	Разработка технического задания	РО, И		■	■											
4	Разработка архитектуры сервиса	РО, И		■	■											
5	Разработка аналитического модуля системы	И (50%)				■	■	■	■	■						
6	Разработка веб-сервиса	И (50%)								■	■					
7	Разработка модуля обработки входных данных	И (50%)										■	■			
8	Развертывание системы	И (50%)											■			
9	Внедрение разработки	РО, И (50%)											■	■		
10	Написание пояснительной документации	И (50%)				■	■	■	■	■	■	■	■	■		
11	Проверка работы	НР													■	

#### 4. Бюджет проекта

Бюджет научно-технического исследования включает в себя стоимость всех расходов, необходимых для выполнения работ по магистерской диссертации. По окончании формирования бюджета, были выявлены следующие группы затрат:

- основная заработная плата исполнителей;
- дополнительная заработная плата исполнителей;
- отчисления во внебюджетные фонды (страховые отчисления);
- амортизация;
- накладные расходы.

##### 4.1 Расчет амортизации

При расчете материальных затрат учитывались вся стоимость материалов, которая понадобилась на выполнение данной магистерской работы. В ходе выполнения работы был использован ПК организации. Срок полезного пользования для офисных машин составляет 2-3 года, компьютер взят на 3 года. Стоимость ПК составляет 40000 рублей. Написание работы составляет 4 месяца. Тогда норма амортизации рассчитывается следующим образом:

$$A_n = \frac{1}{3} \times 100\% = 33.33\% \quad (5)$$

Годовые амортизационные отчисления составляют:

$$A_r = 40000 \times 0.33 = 13200 \text{ рублей} \quad (6)$$

Ежемесячные амортизационные отчисления составляют:

$$A_m = \frac{13200}{12} = 1100 \text{ рублей} \quad (7)$$

Итоговая сумма амортизации основных средств составляет:

$$A = 1100 \times 4 = 4400 \text{ рублей} \quad (8)$$

##### 4.2 Расчет основной заработной платы исполнителей

Данный раздел включает в себя заработную плату научного руководителя, руководителя от организации и инженера. Расчет выполняется на основе трудоемкости выполнения каждого этапа и величины месячного оклада исполнителя.

Трудоемкость исполнителей на разных стадиях выполнения магистерской диссертации была просуммирована и представлена в виде количества дней. Таким образом, если согласно плану, работа над проектом должна вестись в течение 127 дней, то реально затраченное время каждого исполнителя в днях отличается от данной цифры, а также от того, что можно увидеть из диаграммы Ганта (так как на ней декомпозиция происходит с точностью до дней, а не часов).

Для расчета основной заработной платы необходимо рассчитать среднедневную заработную плату:

$$Z_{\text{дн}} = \frac{Z_{\text{м}} \times M}{F_{\text{д}}} \quad (9)$$

где:  $Z_{\text{м}}$  – оклад работника за месяц, руб.;

$M$  – количество месяцев работы без отпуска в течение года, 10,4;

$F_{\text{д}}$  – действительный годовой фонд рабочего времени персонала,  $365 - 66 - 48 = 251$  день.

Основная заработная плата рассчитывается по следующей формуле:

$$Z_{\text{осн}} = Z_{\text{дн}} \times T_{\text{р}} \quad (10)$$

где:  $T_{\text{р}}$  – продолжительность работ, выполняемых научно-техническим работником, раб. дн.;  $Z_{\text{дн}}$  – среднедневная заработная плата работника, руб.

Расчет основной заработной платы представлен в таблице 10.

Месячный должностной оклад работника:

$$Z_{\text{м}} = Z_{\text{б}} \cdot (1 + k_{\text{пр}} + k_{\text{д}}) \cdot k_{\text{р}}, \quad (11)$$

где  $Z_{\text{б}}$  – базовый оклад, руб.;

$k_{\text{пр}}$  – премиальный коэффициент, (определяется Положением об оплате труда);

$k_{\text{д}}$  – коэффициент доплат и надбавок (в НИИ и на промышленных предприятиях – за расширение сфер обслуживания, за профессиональное мастерство, за вредные условия: определяется Положением об оплате труда);

$k_{\text{р}}$  – районный коэффициент, равный 1,3 (для Томска).

Таблица 10 – Расчет основной заработной платы

Исполнители	$Z_{\text{б}}$ , руб.	$k_{\text{пр}}$	$k_{\text{д}}$	$k_{\text{р}}$	$Z_{\text{м}}$ , руб	$Z_{\text{дн}}$ , руб.	$T_{\text{р}}$ , раб. дн.	$Z_{\text{осн}}$ , руб.

Научный руководитель	35120	0	0	1,3	45656	1891,72	6	11350,34
Руководитель от организации	35000	0	0	1,3	45500	1885,26	12	22623,11
Инженер	21760	0	0	1,3	28288	1172,09	74	86734,84

#### 4.3 Расчет дополнительной заработной платы исполнителей

В данную статью включается сумма выплат, предусмотренных законодательством о труде. Например, оплата очередных и дополнительных отпусков; оплата времени, связанного с выполнением государственных и общественных обязанностей; выплата вознаграждения за выслугу лет и т. п. (в среднем — 12% от суммы основной заработной платы).

Расчеты дополнительной заработной платы представлены в таблице 11.

Таблица 11 – Расчет дополнительной заработной платы

Исполнитель	Основная з/п, руб.	К.	Дополнительная з/п, руб.
Научный руководитель	11350,34	0,12	1362,04
Руководитель от организации	22623,11		2714,77
Инженер	86734,84		10408,18
<b>Итого С<sub>доп</sub></b>			14484,99

#### 4.4 Расчет итоговой заработной платы

Исходя из расчетов, обозначенных в таблицах 10 и 11, была рассчитана итоговая заработная плата исполнителей, которая представлена в таблице 12.

Таблица 12 – Итоговая заработная плата

Исполнитель	Основная з/п, руб	Дополнительная з/п, руб	Итоговая з/п, руб
Научный руководитель	11350,34	1362,04	12712,38
Руководитель от организации	22623,11	2714,77	25337,88
Инженер	86734,84	10408,18	97143,02
<b>Итого С<sub>зп</sub></b>			135193,28

#### 4.5 Расчет отчислений во внебюджетные фонды

При составлении расходов на проект, необходимо учитывать обязательные отчисления по установленным законодательством Российской Федерации нормам органам государственного социального страхования (ФСС), пенсионного фонда (ПФ) и медицинского страхования (ФФОМС) от затрат на оплату труда работников, что в сумме составляет 30,2%.

Отчисления представлены в таблице 13.

Таблица 13 – Отчисления во внебюджетные фонды

Исполнитель	Зарплата, руб	Сумма отчислений, руб
Научный руководитель	12712,38	3839,14
Руководитель от организации	25337,88	7652,04
Инженер	97143,02	29337,19
<b>Итого С<sub>внеб</sub></b>		40828,37

#### 4.6 Расчет накладных расходов

В эту статью включаются затраты на управление и хозяйственное обслуживание, которые могут быть отнесены непосредственно на конкретную тему. Кроме того, сюда относятся расходы по содержанию, эксплуатации и ремонту оборудования, производственного инструмента и инвентаря, зданий, сооружений и др. В расчетах эти расходы принимаются в размере 70 - 90 % от суммы основной заработной платы научно-производственного персонала данной научно-технической организации.

Накладные расходы составляют 80-100 % от суммы основной и дополнительной заработной платы, работников, непосредственно участвующих в выполнении темы. Возьмем  $k_{\text{накл}} = 80\%$ .

Расчет накладных расходов производится по формуле:

$$C_{\text{накл}} = k_{\text{накл}} \times (Z_{\text{осн}} + Z_{\text{доп}}) = 0,8 \times 135193,28 = 108\,154,62 \quad (12)$$

где  $k_{\text{накл}}$  – величина коэффициента накладных расходов;

$Z_{\text{осн}}$  – основная заработная плата;

$Z_{\text{доп}}$  – дополнительная заработная плата.

#### 4.7 Формирование бюджета проекта

Исходя из произведенных расчетов, сумма всех расходов была рассчитана и представлена в таблице 14.

Таблица 14 – Бюджет затрат

Расходы	Сумма, руб	%
Основная заработная плата исполнителей	86734,84	34,07
Дополнительная заработная плата исполнителей	14484,99	5,69
Отчисления во внебюджетные фонды	40828,37	16,04
Амортизация	4400	1,73
Накладные расходы	108154,62	42,48
<b>Итого</b>	<b>254602,82</b>	<b>100</b>

Рассчитанный бюджет не превышает бюджета в 300000 рублей, указанного в ограничениях проекта (таблица 7, раздел 4.1).

#### 5. Реестр рисков проекта

Идентифицированные риски проекта включают в себя возможные неопределенные события, которые могут возникнуть в проекте и вызвать последствия, которые повлекут за собой нежелательные эффекты. Информация о потенциальных рисках приведена в таблице 15.

Таблица 15 – Реестр рисков

№	Риск	Потенциальное воздействие	Вероятность наступления	Влияние риска	Уровень риска	Способы смягчения	Условия наступления
1	Технологическое отставание	Снижение качества работы алгоритмов по сравнению с конкурентами	1	3	высокий	Правильная архитектура системы для возможности итеративного изменения	Недостаточный анализ предметной области
2	Недостаточная производительность	Снижение конкурентоспособности	1	2	средний	Использование	Технологическое

	ость системы	особности по сравнению с аналогами				алгоритмов машинного обучения сводит риск к минимуму	отставание алгоритмов системы
3	Узкий сегмент потребителей	Высокая конкуренция	2	3	средний	Более детальный анализ конкурентов	Реализация широко потребляемой системы
4	Увеличение требований со стороны заказчика	Появление дополнительных трудозатрат	2	1	низкий	Обсуждение проекта с заказчиком на каждом этапе сотрудничества	Ошибки в составлении технического задания

В результате идентификации рисков были рассмотрены потенциальные риски, часть которых могут привести к неконкурентоспособности разработанной системы. Однако, воздействие данных рисков возможно свести к минимуму, если следовать представленным способам смягчения.

## **6. Определение экономической эффективности исследования**

По описанным критериям, разобранных в данном разделе, можно сформировать итоговый результат эффективности исследования.

Потенциальными потребителями данной работы являются организации, имеющие в своих бизнес-процессах временные затраты на анализ геологических карт: НИИ геологии и нефтегазовые компании.

Анализ конкурентоспособности показал, что разработанный сервис имеет несколько преимуществ перед конкурентами, в частности, автоматизация труда геолога за счет использования алгоритмов машинного обучения.

Для повышения качества управления проектом работ была произведена декомпозиция задач и разграничение зоны ответственности между участниками проекта. Общая длительность исследования составила около четырех месяцев.

В результате идентификации рисков были рассмотрены потенциальные риски, часть которых могут привести к неконкурентоспособности

разработанной системы. Для их устранения предложены методы смягчения рисков.

Рассчитываемая величина затрат составила 254 602,82 рублей. Исходя из ограничений, накладываемых на проект, максимальный бюджет не должен превышать 300000 рублей.

Для определения экономической эффективности разработки был рассчитан интегральный финансовый показатель, который определяется по следующей формуле:

$$I_{\text{фин}} = \frac{\Phi_p}{\Phi_{\text{max}}}, \quad (13)$$

где  $I_{\text{фин}}$  – интегральный финансовый показатель разработки,  $\Phi_p$  – стоимость исполнения работ,  $\Phi_{\text{max}}$  – максимально допустимая стоимость исполнения проекта.

$$I_{\text{фин}} = \frac{254602,82}{300000} = 0,85 \quad (14)$$

Таким образом значения финансового показателя составляет 0,85, что свидетельствует об эффективном использовании финансовых ресурсов.

Одним из основных положительных эффектов исследования является систематический анализ, проработка методов анализа картографической информации, а также реализация веб-сервиса, что позволит в перспективе получить пользовательскую обратную связь и доработать сервис до промышленных масштабов.

## ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

<b>Группа</b>	<b>ФИО</b>
8ПМ8И	Чугунов Роман Анварович

<b>Школа</b>	<b>ИШИТР</b>	<b>Отделение (НОЦ)</b>	<b>ОИТ</b>
Уровень образования	Магистратура	Направление/специальность	Программная инженерия

Тема ВКР:

Система анализа геологических карт	
<b>Исходные данные к разделу «Социальная ответственность»:</b>	
1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения	Разработка программных модулей системы анализа геологических карт, целью которой является автоматизация процесса создания трехмерной модели структуры нефтяного пласта на основе изображения геологической карты.
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
<b>1. Правовые и организационные вопросы обеспечения безопасности:</b> <ul style="list-style-type: none"> <li>– специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства;</li> <li>– организационные мероприятия при компоновке рабочей зоны.</li> </ul>	- ГОСТ 12.2.032-78 ССБТ. - СанПиН 2.2.4.548-96. - СанПиН 2.2.4/2.1.8.562-96. - СанПиН 2.2.2/2.4.1340-03. - ГОСТ 12.1.009-2009. - ГОСТ 12.1.038-82 ССБТ. - ГОСТ Р 22.3.03-94.
<b>2. Производственная безопасность:</b> 2.1. Анализ выявленных вредных и опасных факторов 2.2. Обоснование мероприятий по снижению воздействия	Выявленные опасные факторы: - Электрический ток (источник – ПК). - Короткое замыкание. - Статическое заземление (источник – ПК). Выявленные вредные факторы: - Освещение. - Микроклимат. - Шум. - Психофизиологические факторы: нервно-психические перегрузки. - Электромагнитное излучение.
<b>3. Экологическая безопасность:</b>	Отсутствует воздействие объекта на атмосферу, гидросферу. Воздействие на литосферу происходит при утилизации ПК, используемого для разработки, а также утилизации люминесцентных ламп освещения.
<b>4. Безопасность в чрезвычайных ситуациях:</b>	В ходе разработки программы возможно возникновение пожара на рабочем месте

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

**Задание выдал консультант:**

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Горбенко Михаил Владимирович	Кандидат технических наук		

Задание принял к исполнению студент:

<b>Группа</b>	<b>ФИО</b>	<b>Подпись</b>	<b>Дата</b>
8ПМ8И	Чугунов Роман Анварович		

## **Введение**

Целью дипломной работы является разработка некоторых модулей системы анализа геологических карт. Данная система предназначена для автоматизации процесса создания трехмерных структурных моделей на основе геологических карт, а также представления разнородных геологических карт в единой системе, доступной пользователям через веб-интерфейс. Значительное количество картографической информации нефтегазовых компаний хранится в виде чертежей на бумаге, что несет за собой ряд проблем:

- Невозможность быстрого поиска карт по различным критериям: по региону, по дате создания, по схожести между собой, по создателю карты и т.д.
- Различия в представлении различных карт.
- Проблемы в отслеживании актуальности карт.
- Сложность сравнения объектов, представленных на двух разных картах.

Система включает в себя следующие модули:

4. Серверная часть (backend) – взаимодействие с базой данных, аналитической частью, а также ответы за запросы клиентской части приложения.
5. Клиентская часть (frontend) – веб-интерфейс пользователя, который включает функции просмотра и редактирования имеющихся карт, добавления новых, их автоматического анализа на основе алгоритмов компьютерного зрения, а также экспорта результатов в формате GeoJSON.
6. Аналитическая часть – модули, целью которых является извлечение полезной информации из карт, такой как: наличие и координаты нефтяных скважин, различные виды изолиний и их глубина.

Конкретной целью данной дипломной работы является разработка аналитической части системы, а также одного из модулей клиентской части – модуля редактирования результатов аналитической части.

Таким образом, выполняемая работа соответствует работе разработчика программного обеспечения. Местом выполнения работы является офис предприятия ООО «Эко-Томск».

### **1. Правовые и организационные вопросы обеспечения безопасности**

Для обеспечения охраны безопасности здоровья сотрудников, организации необходимо соблюдать ряд правил для офисного помещения и персональных компьютеров, изложенных в трудовом праве российской федерации.

Работодатели при установке ПК обязаны выполнить следующий перечень требований:

1. К помещению.

2. К освещению.
3. К организации медицинского обследования пользователей.
4. И т.п.

Также немаловажным фактором при выборе компьютеров для сотрудников является возможность конструкции компьютера изменять положение ПК в различных плоскостях (горизонтальные или вертикальные), с возможной устойчивой фиксацией в положении, которая удобна пользователю. Цвет корпуса ПК должен быть спокойным без блестящих деталей, которые бы вызывали повышенную утомляемость глаз. Экран монитора должен содержать регулировку яркости и контрастности, что каждый работник мог установить нужный режим, которые будут соответствовать чувствительности глаз.

Устройство рабочего стола должна быть использовано для оптимального размещения используемого оборудования. Кроме того, форма рабочего стола должна быть удобна для поддержания рациональной позы пользователя, так, чтобы он мог менять положения своего тела для предупреждения утомления. В соответствии с ГОСТ 12.2.032-78 ССБТ «Рабочее место при выполнении работ сидя. Общие эргономические требования» [17] к рабочему месту предъявляются следующие основные требования:

1. При организации рабочего места следует учитывать антропометрические показатели женщин (если работают только женщины) и мужчин (если работают только мужчины); если работают и женщины, и мужчины — общие средние показатели женщин и мужчин.
2. Конструкцией рабочего места должно быть обеспечено оптимальное положение работающего, которое достигается регулированием высоты рабочей поверхности, сиденья и пространства для ног.
3. Продолжительность рабочего дня не должна превышать 40 часов в неделю. Вид трудовой деятельности за компьютерным устройством (компьютер, мобильное устройство), в рамках выполнения выпускной квалификационной работы, соответствует группе «В» — творческая работа в режиме диалога с компьютерным устройством. Категория данной трудовой деятельности соответствует III (до 6 часов непосредственной работы за компьютером).

## **2. Производственная безопасность**

### **2.1 Анализ выявленных вредных и опасных факторов**

В процессе рассмотрения безопасности в рабочей зоне необходимо было выявить вредные и опасные факторы, которые могут возникнуть на рабочем месте. Факторы считаются вредными, если его воздействие на человека может привести его к заболеванию. Опасный фактор может привести человека к травме.

Также необходимо было описать мероприятия по защите разработчика и пользователей конечного продукта от действия данных факторов. В таблице 16 приведены негативные факторы, влияющие на работу с компьютером.

Таблица 16 – Негативные факторы при работе с компьютером

Наименование видов работ и параметров производственного процесса	Факторы (ГОСТ 12.0.003-74 ССБТ)	Нормативные документы
<b>Вредные факторы</b>		
Работа с компьютером	Повышенная или пониженная температура воздуха рабочей зоны	СанПиН 2.2.4.548-96
	Повышенная или пониженная влажность воздуха рабочей зоны	СанПиН 2.2.4.548-96
	Повышенный уровень шума на рабочем месте	СанПиН 2.2.4/2.1.8.562-96
	Недостаточная освещенность рабочей зоны	СанПиН 2.2.1/2.1.1.1278-03
	Повышенный уровень электромагнитных излучений	СанПиН 2.2.2/2.4.1340-03
<b>Опасные факторы</b>		
Работа с компьютером	Опасность поражения электрическим током	ГОСТ 12.1.038-82
	Пожаровзрывоопасность	ГОСТ 12.1.041-83

## 2.2 Обоснование мероприятий по снижению воздействия

### *Микроклимат*

Одним из необходимых условий труда является обеспечение нормального микроклимата в рабочей зоне, оказывающие значительное влияние на самочувствие человека.

Так как работа с использованием персонального компьютера является основной, (диспетчерские, операторские, расчетные, кабины и посты управления, залы вычислительной техники и др.) и связана с нервно-эмоциональным напряжением, должны быть обеспечены оптимальные параметры микроклимата для категории работ 1а и 1б в соответствии с действующими санитарно-эпидемиологическими нормативами микроклимата производственных помещений. На других рабочих местах следует

поддерживать параметры микроклимата на допустимом уровне, соответствующем требованиям указанных выше нормативов.

Содержание вредных химических веществ в таких рабочих зонах не должно превышать предельно допустимых концентраций загрязняющих веществ в атмосферном воздухе населенных мест в соответствии с действующими гигиеническими нормативами. Также, в помещениях с персональными компьютерами должна проводиться ежедневная влажная уборка.

Нормативным документом, отвечающим за гигиенические требования к микроклимату производственных помещений, является СанПиН 2.2.4.548-96 [18]. В документе указаны все нормативные требования к микроклимату на рабочих местах, у всех видов производственных помещений.

В таблице № 17 представлены фактические, оптимальные и допустимые показатели микроклимата рабочей зоны.

Таблица 17 – Параметры микроклимата на рабочем месте

Период года	Кат. раб.	Температура воздуха, 0С			Относительная влажность воздуха, %			Скорость движения воздуха, м/с		
		Факт.	Опт.	Доп.	Факт.	Опт.	Доп.	Факт.	Опт.	Доп.
Холодный	Ia	23	22-24	20-25	55	40-60	15-75	0,1	0,1	0,1
Теплый	Ia	23	23-25	21-28	50	40-60	15-75	0,1	0,1	0,1-0,2

Оптимальный микроклимат является необходимым на рабочих местах, так как создает комфортное нахождение человека в рабочей зоне, а также обеспечивает его высокий уровень работоспособности. Такие микроклиматические условия обеспечивают благоприятное состояние организму человека и не вызывают отклонений в состоянии его здоровья.

### *Шум*

Превышение уровня шума является распространенным вредным фактором на рабочем месте. Его нарушение влечет за собой негативное воздействие не только на органы слуха, но и на весь организм человека в целом, через центральную нервную систему.

Среди источников шума выделяют принтеры, систему охлаждения, множительная техника, осветительные приборы дневного света, а также шумы, проникающие извне.

Уровень шума не должен превышать значений, установленных СанПиН 2.2.4/2.1.8.562-96 [19] и составлять не более 50 дБА. Допустимые значения уровней звукового давления представлены в таблице № 18.

Таблица 18 – Допустимые значения уровней звукового давления компьютера

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

Снижению уровня шума способствует установка звукопоглощающих материалов (плиты, панели), подвесных акустических потолков, а также установка малошумного оборудования.

#### *Уровень освещенности*

Недостаточная освещенность рабочей зоны оказывает негативное влияние на зрительную систему человека. Происходит снижение концентрации внимания, усталость центральной нервной системы, что приводит к снижению производительности труда.

Уровень освещения на поверхности рабочего стола в зоне размещения документа, согласно СанПиН 2.2.2/2.4.1340-03 [21], должен быть в диапазоне от 300 до 500 лк. Уровень освещенности экрана не должен превышать 300 лк. Яркость осветительных приборов, находящихся в поле зрения, не должна превышать 200 кд/м<sup>2</sup>.

Офис, в котором выполняется работа, представляет собой квадратное помещение со стороной  $A = 12$  м, высотой  $H = 3,5$  м, количество светильников 24 шт, однако, каждый состоит из 4 люминесцентных ламп, таким образом, количество ламп  $n = 96$  шт.

Определим расчетную высоту подвеса светильников над рабочей поверхностью ( $h$ ) по формуле:

$$h = H - h_p - h_c, \quad (1)$$

где  $H$  – высота потолка в помещении, м;  $h_p$  – расстояние от пола до рабочей поверхности стола, м;  $h_c$  – расстояние от потолка до светильника, м.

Вычислим расчетную высоту подвеса светильников над рабочей поверхностью по формуле 1 для данного офисного помещения:

$$h = 3,5 - 0,8 - 0,05 = 2,65 \text{ м}$$

Индекс помещения определяется по формуле (2):

$$i = \frac{S}{h(A+B)}, \quad (2)$$

где  $S$  – площадь помещения, м<sup>2</sup>;  $A$  – длина комнаты, м;  $B$  – ширина комнаты, м;  $h$  – высота подвеса светильников, м.

Индекс данного офисного помещения:

$$i = \frac{144}{2,65 \cdot (24)} = 2,26.$$

Исходя из того, что в помещении свежепобеленный потолок, а также свежепобеленные стены с окнами без штор, согласно методическим указаниям [20], примем коэффициенты отражения от стен  $\rho_c = 70\%$  и потолка  $\rho_n = 50\%$ . По

таблице коэффициентов использования светового потока (с.14, таблица 11 в методических указаниях [20]) для соответствующих значений  $i$ ,  $\rho_c$ ,  $\rho_n$ , примем  $\eta=0,62$ .

Освещенность помещения рассчитывается по формуле:

$$E_{\phi} = \frac{n \cdot \eta \cdot \Phi}{S \cdot k_z \cdot z}; \quad (3)$$

где  $\Phi$  – световой поток светильника, лм;  $S$  – площадь помещения, м<sup>2</sup>;  $k_z$  – коэффициент запаса;  $n$  – число светильников;  $\eta$  – коэффициент использования светового потока,  $z$  – коэффициент неравномерности освещения.

Коэффициент запаса  $k_z$  учитывает запыленность светильников и их износ. Для помещений с малым выделением пыли  $k_z = 1,5$ . Для люминесцентных ламп  $z = 1,1$ . В помещении находятся светильники ЛВО 4×18 CSVT, с люминесцентными лампами типа L 18W/640 с потоком  $F = 1200$  лм. Учитывая все параметры, рассмотренные выше, найдем освещенность (формула 5.3):

$$E_{\phi} = \frac{96 \cdot 0,62 \cdot 1200}{144 \cdot 1,5 \cdot 1,1} = 300,6 \text{ лк.}$$

В рассматриваемом помещении освещенность должна составлять 300 лк согласно СНиП 23-05-95. В данном помещении освещенность находится в пределах нормы, однако, на самой границе допустимого диапазона значений освещенности. Дополнительные источники освещения не нужны, если лампы вовремя подлежат замене.

#### *Электромагнитное излучение*

Вся электротехника, и компьютеры в том числе, производят электромагнитное излучение. В таблице 19 представлены временные допустимые уровни электромагнитных полей, создаваемые компьютерами на рабочих местах, согласно СанПиНу 2.2.2/2.4.1340-03 [21].

Таблица 19 – Временные допустимые уровни электромагнитных полей

Наименование параметров		Временные допустимые уровни электромагнитных полей
Напряженность электрического поля	в диапазоне частот 5 Гц-2 кГц	25 В/м
	в диапазоне частот 2 кГц-400 кГц	2,5 В/м
Плотность магнитного потока	в диапазоне частот 5 Гц-2 кГц	250 нТл
	в диапазоне частот 2 кГц-400 кГц	25 нТл
Поверхностный видеомонитора	электростатический потенциал экрана	500В

### *Психофизиологические факторы*

Продолжительность непрерывной работы за компьютерным устройством, без регламентированного перерыва, не должна превышать 2 часа. Длительность регламентированных перерывов составляет 20 минут (после 1,5-2,0 часа от начала рабочей смены и обеденного перерыва). Также, необходимо уделять время нерегламентированным перерывам (микропаузы), длительность которых составляет 1-3 минуты.

Для снижения воздействия вредных факторов, устанавливаются перерывы в работе для отдыха сотрудников. Суммарное время регламентированных перерывов при работе с персональным компьютером зависит от категории трудовой деятельности и уровня нагрузки за рабочую смену. В таблице 20 приведено суммарное время отдыха для каждой категории работ.

Таблица 20 – Суммарное время перерывов

Категория работы с ПК	Уровень нагрузки за рабочую смену при видах работ с ПК			Суммарное время регламентированных перерывов при 8-часовой смене, мин.
	Группа А, количество знаков	Группа Б, количество знаков	Группа В, количество знаков	
I	до 20000	до 15000	до 2	50
II	до 40000	до 30000	до 4	70
III	до 60000	до 40000	до 6	90

Как уже было указано выше, данная работа относится к III категории.

### *Статическое электричество*

В помещениях, оборудованных ПЭВМ, токи статического электричества чаще всего возникают при прикосновении персонала к любому из элементов ПЭВМ. Такие разряды опасности для человека не представляют, однако кроме неприятных ощущений могут привести к выходу оборудования из строя.

Для предотвращения образования и защиты от статического электричества в помещении используются нейтрализаторы и увлажнители, а полы имеют антистатическое покрытие в виде поливинилхлоридного антистатического линолеума.

Также в СанПиН 2.2.2/2.4.1340-03 установлен максимальный допустимый электростатический потенциал экрана видеомонитора – 500 В.

В качестве мер уменьшения влияния вредных факторов на пользователя используются защитные фильтры для мониторов, увлажнители воздуха. Должны использоваться розетки с заземлением. Требуется проводить регулярную влажную уборку.

### *Электрический ток*

Среди распространенных опасностей в рабочей зоне находится и поражение электрическим током. Опасность поражения определяется величиной тока проходящего через тело человека  $I$  или напряжением прикосновения  $U$ . Напряжение считается безопасным при напряжении прикосновения  $U < 42$  В.

При получении человеком разряда электрического тока могут быть получены электротравмы, электрические удары и даже летальный исход (согласно ГОСТ 12.1.009-2009 [22]).

Для защиты от поражения электрическим током следует выполнить следующие пункты:

1. Обеспечить недоступность токоведущих частей от случайных прикосновений.
2. Электрическое разделение цепей.
3. Устранить опасность поражения при появлении напряжения на разных частях.

В таблице 21 представлены предельно допустимые значения напряжения прикосновения и тока на рабочем месте (согласно ГОСТ 12.1.038-82 [23]).

Таблица 21 – Допустимые значения напряжения прикосновения и тока

Род тока	Напряжения прикосновения, В	Ток, мА
	Не более	
Переменный, 50 Гц	2,0	0,3
Постоянный	8,0	1,0

Согласно электробезопасности, рабочее место относится к помещениям без повышенной опасности поражения электрическим током. Данный фактор характеризуется отсутствием условий, создающих повышенную или особую безопасность от разряда электрическим током.

### **3. Экологическая безопасность**

Экологическая безопасность и охрана окружающей среды являются одними из важнейших факторов при выполнении работ любого характера. При работе в офисном помещении за персональным компьютером отсутствуют выбросы в окружающую среду и нет влияния на жилищную зону.

Так как при разработке данной магистерской диссертации использовался персональный компьютер, необходимо описать правильную утилизацию компьютерного лома после его выхода из строя. В соответствии с постановлением правительства юридическим лицам запрещено

самостоятельно утилизировать компьютерную технику. Для этого необходимо найти специальную компанию, которая занимается утилизацией в частном порядке.

В нормативном документе СанПиН 2.2.2/2.4.1340-03 [21], даются следующие общие рекомендации по снижению опасности для окружающей среды, исходящей от компьютерной техники:

1. Применять оборудование, соответствующее санитарным нормам и стандартам экологической безопасности.
2. Применять расходные материалы с высоким коэффициентом использования и возможностью их полной или частичной регенерации.
3. Отходы в виде компьютерного лома утилизировать.
4. Использовать экономичные режимы работы оборудования.

#### **4. Безопасность в чрезвычайных ситуациях**

Для устранения возможных причин возгорания следует проводить следующие мероприятия:

1. Организация мероприятий:
  - Противопожарный инструктаж обслуживающего персонала.
  - Обучение персонала техники безопасности.
  - Разработка инструкций, планов эвакуаций и пр.
2. Эксплуатационные мероприятия:
  - Соблюдение эксплуатационных норм оборудования.
  - Выбор и использование современных автоматических средств пожаротушения.
3. Технические мероприятия:
  - Профилактический осмотр и ремонт оборудования.
  - Соблюдение противопожарных мероприятий при устройстве электропроводок, оборудования, систем отопления и пр. [24]

#### **5. Выводы**

В результате анализа были определены опасные и вредные производственные факторы для офисного помещения, в котором ведется работа над проектом.

Были определены значения нормативных показателей вредных и опасных факторов на основе документов, регламентирующих меры безопасности на рабочих местах. Также, были составлены рекомендации по улучшению условий труда с целью минимизации влияния вредных и опасных факторов на работника. Кроме того, было оценено воздействие на окружающую среду и установлены правила утилизации отходов. Проведен расчет уровня освещенности рабочего места в помещении, в котором ведется разработка.

Полученное значение освещенности соответствует нижней допустимой границе уровня освещенности.

Были установлены мероприятия, которые необходимо проводить с целью устранения возможных причин возгорания.

На основе правовых норм трудового законодательства были определены основные требования к организации рабочего места.

На основе проведенного анализа может быть сделан вывод о том, что рабочее место соответствует нормативным требованиям производственной безопасности и охраны труда.

## ЗАКЛЮЧЕНИЕ

Несколько областей компьютерного зрения было исследовано в данной работе для того, чтобы разработать два модуля системы анализа структурных геологических карт – модуль анализа изображений и модуль исправления результатов анализа.

Несколько современных методов машинного и глубокого обучения были применены на практике, такие как:

- Бинарная сегментация на основе сверточных нейронных сетей (две сети с разными задачами)
- Бинарная классификация
- Кластеризация
- Детекция объектов

Различные подходы были исследованы в процессе разработки. Все они были оценены на синтетической выборке. На синтетической тестовой выборке были рассчитаны метрики mAP, IOU, F1 score. Однако, наиболее важной метрикой качества оставалась экспертная оценка из-за недостатка размеченных данных.

В будущем запланировано несколько модификаций в разработанных модулях:

1. Нейронная сеть для бинарной сегментации будет заменена на сеть для сегментации на несколько классов. Это позволит извлекать изолинии, структурные разломы, границы лицензионных участков и другую информацию на геологических картах.
2. Генеративно-состязательные сети будут оценены как возможный источник более разнообразных синтетических данных.
3. Разработанная система будет использована для аннотации реальных данных с меньшими трудозатратами. Собранные данные позволят улучшить существующие алгоритмы.

4. Добавление распознавания тектонических нарушений на картах.  
Такую функцию можно реализовать с помощью замены бинарной сегментации изолиний на сегментацию на несколько классов.

Другим важным направлением развития является улучшение UI/UX. Пользователи системы часто сравнивают ее с несколькими другими геоинформационными системами. Некоторые из подобных систем были в разработке многие годы. Следовательно, улучшение клиентской части приложения необходимо для того, чтобы догнать данные ГИС системы по удобству и набору функций. Для достижения этой цели необходимо сделать значительно количество модификаций в текущем модели редактирования результатов и клиентской части приложения в целом.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2015, pp. 1–9.
2. W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A. C. Berg, arXiv:1512.02325 [cs.CV], 2016.
3. Dalal, N., Triggs, B.: Histograms of Oriented Gradients for Human Detection. CVPR. (2005)
4. Felzenszwalb, P., Girshick, R., McAllester, D., Ramanan, D.: Object detection with discriminatively trained part based models. TPAMI 32 (9) (2010) 1627–1645
5. François Chollet, Xception: Deep Learning with Depthwise Separable Convolutions, arXiv:1610.02357v3 [cs.CV]
6. Ankush Gupta, Andrea Vedaldi, Andrew Zisserman, Synthetic Data for Text Localisation in Natural Images, arXiv:1604.06646 [cs.CV].
7. A fast parallel algorithm for thinning digital patterns, T. Y. Zhang and C. Y. Suen, Communications of the ACM, March 1984, Volume 27, Number 3.
8. O. Ronneberger, P. Fischer and T. Brox, "U-net: Convolutional networks for biomedical image segmentation", Proc. Med. Image Comput. Comput.-Assisted Intervention, pp. 234-241, 2015.
9. F. Chollet. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.
10. Dan Pelleg and Andrew Moore. 1999. Accelerating exact k-means algorithms with geometric reasoning. In Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '99). Association for Computing Machinery, New York, NY, USA, 277–281.
11. React: a JavaScript library for building user interfaces [Электронный ресурс]. URL: <https://reactjs.org/>

12. Vue.js: the progressive JavaScript framework [Электронный ресурс]. URL: <https://vuejs.org/>
13. Angular: one framework. Mobile & desktop [Электронный ресурс]. URL: <https://angular.io/>
14. Angular vs. Vue vs. React. [Электронный ресурс]. URL: <https://www.codeinwp.com/blog/angular-vs-vue-vs-react/>
15. Vue.js: comparison with other frameworks. [Электронный ресурс]. URL: <https://vuejs.org/v2/guide/comparison.html>
16. Angular, React, Vue: a performance comparison. [Электронный ресурс]. URL: <https://blog.logrocket.com/angular-vs-react-vs-vue-a-performance-comparison/>
17. ГОСТ 12.2.032-78 ССБТ. Рабочее место при выполнении работ сидя. Общие эргономические требования.
18. СанПиН 2.2.4.548–96. Гигиенические требования к микроклимату производственных помещений.
19. СанПиН 2.2.4/2.1.8.562–96. Шум на рабочих местах, в помещениях жилых, общественных зданий и на территории застройки.
20. Назаренко О. Б. Расчёт искусственного освещения. Методические указания к выполнению индивидуальных заданий для студентов дневного и заочного обучения всех направлений и специальностей ТПУ. – Томск: Изд. ТПУ, 2008. – Режим доступа: [http://window.edu.ru/resource/149/75149/files/raschet\\_isk\\_osv.pdf](http://window.edu.ru/resource/149/75149/files/raschet_isk_osv.pdf), свободный. – (дата обращения: 29.05.2020).
21. СанПиН 2.2.2/2.4.1340–03. Санитарно-эпидемиологические правила и нормативы «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы».
22. ГОСТ 12.1.009-2009. Система стандартов безопасности труда. Электробезопасность.

23. ГОСТ 12.1.038-82 ССБТ. Электробезопасность. Предельно допустимые уровни напряжений прикосновения и токов.

24. ГОСТ Р 22.3.03-94. Безопасность в ЧС. Защита населения.  
Основные положения

## Приложение II

(справочное)

### Review of Key Scientific Papers in Image Classification, Object Detection and Image Segmentation

Студент

Группа	ФИО	Подпись	Дата
8ПМ8И	Чугунов Р.А.		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Аксенов С.В.	к.т.н.		

Консультант-лингвист отделения иностранных языков ШБИП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель	Пичугова И.Л.	—		

Algorithms from multiple areas of Deep Learning are used in this master thesis: object detection, image classification, image segmentation. This section contains a study of several important scientific papers in these fields.

The most often used architecture for image segmentation was introduced in 2015 in paper U-Net: Convolutional Networks for Biomedical Image Segmentation [1].

The authors of the paper presented a network and training strategy that relies on data augmentation to use the available annotated samples more efficiently. The architecture they proposed consists of a contracting path to capture context and a symmetric expanding path that enables precise localization. They showed that such a network outperforms the prior best methods (a sliding-window convolutional network) on the ISBI cell tracking challenge 2015 in these categories by a large margin. Moreover, their network architecture is fast. Segmentation of a 512x512 image takes less than a second on a common GPU in 2015. It takes even less on modern hardware and software. The authors published their implementation online. The implementation is based on Caffe. Popularity of this framework declined during past 5 years.

While convolutional networks were already existed for a long time in 2015, their success was limited due to the size of the available training sets and the size of the considered networks. The typical use of convolutional networks is on classification tasks, where the output to an image is a single class label. However, in many visual tasks, especially in biomedical image processing, the desired output should include localization, i.e., a class label is supposed to be assigned to each pixel. Furthermore, thousands of training images are usually beyond reach in biomedical tasks (as well as many other tasks). A common approach before U-Net was to train a network in a sliding-window setup to predict the class label of each pixel by providing a local region (patch) around that pixel as input. This network can localize and the training data in terms of patches is much larger than the number of training images. Such network won the EM segmentation challenge at ISBI 2012 by a large margin.

However, this approach has two drawbacks. Firstly, it is quite slow because the network must be run separately for each patch and there is a lot of redundancy due to overlapping patches. Secondly, there is a trade-off between localization accuracy and the use of context. Large patches require more max-pooling layers that reduce the localization accuracy, while small patches allow the network to see only little context.

The authors of U-Net (Olaf Ronneberger, Philipp Fischer, Thomas Brox) built a more elegant architecture. This architecture is often used even in 2020. They proposed the so-called “fully convolutional network”. They made use of an idea to supplement a usual contracting network by successive layers, where pooling operators are replaced by upsampling operators. Upsampling operators increase the resolution of the output. In order to localize, high resolution features from the contracting path are combined with the upsampled output. A successive convolution layer can then be trained to assemble a more precise output based on this information.

Another important modification in their architecture is that in upsampling part it has also a large number of feature channels. These channels allow the network to propagate context information to higher resolution layers. As a consequence, the expansive path is more or less symmetric to the contracting path, and yields a u-shaped architecture. The network does not have any fully connected layers.

The authors also proposed to use elastic deformations to the available training images. It is called data augmentation. Data augmentation allows the network to learn invariance to such deformations and data transforms, without the need to see these transformations in the annotated image corpus.

A common challenge in cell segmentation tasks is the separation of touching objects of the same class. In this paper the use of a weighted loss function was proposed. The separating background labels between touching cells obtain a large weight in this loss function.

The authors evaluated the U-Net architecture on very different biomedical segmentation applications. Thanks to data augmentation with elastic deformations,

it only needed very few annotated images and had a very reasonable training time of only 10 hours on Nvidia Titan GPU (6 GB). It is an outdated version of a professional GPU. Nowadays, even some gaming GPUs are more powerful.

The resulting network and training strategy are applicable not only to various biomedical segmentation problems. They can be applied to segmentation tasks in different scientific areas.

SSD (Single Shot MultiBox Detector) was introduced in the same year as U-net (2015) [2]. SSD was used in this master thesis to detect oil wells on images. This architecture is designed for object detection in real-time. Faster R-CNN (another architecture for object detection) uses a region proposal network to create boundary boxes and utilizes those boxes to classify objects. While it was considered the state-of-the-art in terms of accuracy, the whole process runs at 7 frames per second. It is far below what is required for real-time processing. SSD speeds up the process by eliminating the need of the region proposal network. To recover the decline in accuracy, SSD applies a few improvements including multi-scale features and default boxes. These improvements allow SSD to match the Faster R-CNN's accuracy using lower resolution images, which further pushes the speed higher.

Experimental results on the PASCAL VOC, COCO, ILSVRC datasets confirm that SSD has competitive accuracy to methods that utilize an additional object proposal step and is much faster, while providing a unified framework for both training and inference (the process of predicting outputs for new data samples after training). The authors of the paper made several important contributions:

- They introduced SSD, a single-shot detector for multiple categories that is faster than the previous state-of-the-art for single shot detectors (YOLO – You Only Look Once), and significantly more accurate. In fact, it is equally accurate as slower techniques that perform explicit region proposals and pooling (including Faster R-CNN).

- The core of SSD is predicting category scores and box offsets for a fixed set of default bounding boxes using small convolutional filters applied to feature map.
- To achieve high detection accuracy they produce predictions of different scales from feature maps of different scales, and explicitly separate predictions by aspect ratio.
- These design features lead to simple end-to-end training and high accuracy, even on low resolution input images, further improving the speed vs. accuracy trade-off.
- The authors conducted experiments including timing and accuracy analysis on models with varying input size evaluated on PASCAL VOC, COCO, ILSVRC datasets and are compared to a range of state-of-the-art (in the 2015) approaches.

The SSD approach is based on a feed-forward convolutional network that produces a fixed-size collection of bounding boxes and scores for the presence of object class instances in those boxes, followed by a non-maximum suppression step to produce final detections. Non-maximum suppression removes redundant predicted bounding boxes. Given a list of proposal boxes  $B$ , corresponding confidence scores  $S$  and overlap threshold  $N$ , the algorithm outputs a list of filtered proposals  $D$  in five steps:

1. Select the proposal with highest confidence score, remove it from  $B$  and add it to the final proposal list  $D$ .
2. Compare this proposal with all the proposals – calculate the intersection over union metric of this proposal with every other proposal. If the IOU is greater than threshold  $N$ , remove that proposal from  $B$ .
3. Take the proposal with the highest confidence from the remaining proposals in  $B$  and remove it from  $B$  and add it to  $D$ .
4. Once again calculate the IOU of this proposal with all the proposals in  $B$  and eliminate the boxes which have higher IOU than the threshold.

5. This process is repeated until there no more proposals left in B.

Non-maximum suppression is used in this master thesis too as it is a common post-processing step in object detection.

The authors evaluated SSD in on multiple datasets. They achieved results similar to Faster R-CNN in terms of quality, but won by a huge margin in terms of speed. Their methods SSD300 and SSD512 even outperformed Faster R-CNN in both speed and accuracy. SSD300 and SSD512 have the same settings except that they have different input sizes (300x300 vs. 512x512) Although, another network architecture, Fast YOLO can run at 155 frames per second, it has lower accuracy by almost 22% (mean average precision, mAP). SSD300 was the first real-time method to achieve above 70% mAP. About 80% of the forward time is spent on the base network (VGG16 in the classic architecture case). Therefore, using a faster base network (for instance, Inception, Xception) could event further improve the speed, which can make the SSD512 model real-time as well.

Another groundbreaking work in the previous 5 years is the introduction of Residual Networks [3].

ResNets made it possible to train up to hundreds or even thousands of neural layers and still achieves compelling performance. Taking advantage of its powerful representational ability, the performance of many computer vision applications other than image classification has been improved, such as object detection and face recognition.

Deeper neural networks are more difficult to train. The paper's authors reformulated the layers as learning residual functions with reference to the layer inputs, instead of learning unreferenced functions. They provided comprehensive empirical evidence showing that these residual networks are easier to optimize. On the ImageNet dataset they evaluated residual nets with a depth of up to 152 layers which is 8 times deeper than VGG nets (state-of-the-art architecture before ResNet) but still having lower complexity. This architecture won the 1<sup>st</sup> places on the ILSVRC 2015 classification task and COCO 2015 competitions.

The depth of representations is of central importance for many visual recognition tasks. Solely due to ResNet's extremely deep representations a relative improvement of 28% was achieved on the COCO object detection dataset.

Learning better networks is not as easy as stacking more layers. One of the obstacles is the problem of vanishing/exploding gradients. This problem hampers convergence from the beginning of training. This problem has been largely addressed by normalized initialization of model's weights and intermediate normalization layers, which enable networks with tens of layers to start converging for stochastic gradient descent (SGD) with back-propagation method. When deeper networks became able to start converging, a degradation problem was exposed. With the network depth increasing, accuracy gets saturated and then degrades rapidly. Unexpectedly, such degradation is not caused by overfitting, and adding more layers to a suitable deep model leads to higher training error.

The degradation of training accuracy indicates that not all systems are similarly easy to optimize. This problem was addressed in the paper by introducing a deep residual learning network. The idea of this architecture comes from a hypothesis that stacking layers should not degrade the network performance, because we could simply stack identity mappings (layers that do not do anything) upon the current network. The resulting architecture would perform the same. This indicates that the deeper model should not produce a training error higher than its shallower counterparts.

Instead of hoping each few stacked layers directly fit a desired underlying mapping, authors explicitly let these layers fit a residual mapping. They hypothesise that it is easier to optimize the residual mapping than to optimize the original, unreferenced mapping. Basically, the residual block structure is very simple: the block consists of two convolutional layers and input signal sums up with the output signal of the block.

As a matter of fact, ResNet was not the first to make use of shortcut (identity) connections. Highway Network introduces gated shortcut connections. These parameterized gates control how much information is allowed to flow across the

shortcut. Similar idea can be found in the Long Short-Term Memory cell, a type of neuron, in which there is a parameterized forget gate that controls how much information will flow to the next time step. Therefore, ResNet can be thought of as a special case of Highway Network.

Although, experiments show that Highway Network performs no better than ResNet, which is kind of strange because the solution space of Highway Network contains ResNet, therefore it should perform at least as well as ResNet. This suggests that it is more important to keep these “gradient highways” clear than to go for larger solution space.

There are multiple recent variants and interpretations of ResNet. It was inevitable given its popularity. Its architecture is getting studied heavily. Densely connected CNN (DenseNet) further exploits the effects of shortcut connections – it connects all layers directly with each other. In this novel architecture, the input of each layer consists of the feature maps of all earlier layers, and its output is passed to each subsequent layer. The feature maps are aggregated with depth-concatenation.

Another ResNet variant called ResNeXt is very similar to the Inception module of [4]. They both follow the split-transform-merge paradigm, except in this variant, the outputs of different paths are merged by adding them together, while in [4] they are depth-concatenated. Other difference is that in [4] each path inside blocks is different (1x1, 3x3 and 5x5 convolution) from each other, while in this architecture, all paths share the same topology.

Inception modules are in use in this master thesis. They are the indispensable part inside oil wells detection algorithm.

Inception was responsible for setting the new state of the art for classification and detection in the ImageNet Large-Scale Visual Recognition Challenge 2014 (ILSVRC14). The main hallmark of this architecture is the improved utilization of the computing resources inside the network. This was achieved by a carefully crafted design that allows for increasing the depth and width of the network while keeping the computational budget constant. To optimize quality, the architectural decisions

were based on the Hebbian principle and the intuition of multi-scale processing. The authors of this paper also introduced GoogLeNet, a 22 layers deep network.

The most straightforward way of improving the performance of deep neural networks is by increasing their size. This includes both increasing the depth and width. Some drawbacks of this method were already highlighted above, while another drawback is that bigger size typically means a larger number of parameters. This makes the enlarged network more prone to overfitting, especially if the number of labelled examples in the training set is limited. It can become a major bottleneck, since the creation of high quality training sets can be tricky and expensive. Furthermore, larger networks usually require more powerful computational resources. For example, in a deep vision network, if two convolutional layers are chained, any uniform increase in the number of their filters results in a quadratic increase of computation. If the added capacity is used inefficiently (for example, most weights end up to be close to zero), then a lot of computation is wasted. Since in practice the computational budget is always finite, an efficient distribution of computing resources is preferred to an indiscriminate increase of size, even when the main objective is to increase the quality of results.

The fundamental way of solving these issues would be by ultimately moving from fully connected to sparsely connected architectures, even inside the convolutions.

On the downside, today's computing infrastructures are very inefficient when it comes to numerical calculation on non-uniform sparse data structures. Even if the number of arithmetic operations is reduced 100 times, the overhead of lookups and cache misses is so dominant that switching to sparse matrices would not pay off. The gap is widened even further by the use of steadily improving, highly tuned, numerical libraries that allow for extremely fast dense matrix multiplication, exploiting the minute details of the underlying CPU or GPU hardware. Also, non-uniform sparse models require more sophisticated engineering and computing infrastructure. Most current vision oriented machine learning systems utilize sparsity in the spatial domain just by the virtue of employing convolutions. However,

convolutions are implemented as collections of dense connections to the patches in the earlier layer.

The main idea of the Inception module architecture is based on finding out how an optimal local sparse structure in a convolutional vision network can be approximated and covered by readily available dense components.

The inception network was an important milestone in the development of CNN classifiers. Prior to Inception, most popular convolutional neural networks were just stacked convolutional layers deeper and deeper.

The Inception network on the other hand was complex (heavily engineered). It used a lot of tricks to improve performance. There are several versions of Inception. Most popular of them are the following: Inception v1, Inception v2, Inception v3, Inception v4 and Inception-ResNet. Each version is an iterative improvement over the previous one.

There are some reasons that lead to Inception v1 introduction:

1. Salient parts in the image can have extremely large variation in size.
2. Because of this huge variation in the location of the information, choosing the right kernel size for the convolution operation becomes tough. A larger kernel is preferred for information that is distributed more globally, and a smaller kernel is preferred for information that is distributed more locally.
3. Deep networks are prone to overfitting. It also hard to pass gradient updates through the entire network.
4. Naively stacking large convolution operations is computationally expensive.

As a solution to these problems Inception v1 block consists of filters with multiple sizes. These filters operate on the same level the network becomes wider. A naïve implementation of Inception module performs convolution on an input with 3 different sizes of filters (1x1, 3x3, 5x5). Additionally, max pooling is also performed. The outputs are concatenated and sent to the next inception module.

As it was stated before, deep neural networks are computationally expensive. To make it cheaper, the authors limit the number of input channels by adding an

extra 1x1 convolution before 3x3 and 5x5 convolutions. Though adding an extra operation may seem counterintuitive, 1x1 convolutions are far cheaper than 5x5 convolutions. The 1x1 convolution is introduced after the max pooling layer. Using such modules, GoogLeNet was built. It is a pretty deep classifier. As with any very deep network, it is a subject to the vanishing gradient problem. The authors introduced two auxiliary classifiers to prevent this problem. The total loss function is a weighted sum of the auxiliary loss and the real loss.

Inception v2 and Inception v3 had a number of upgrades which increased the accuracy and reduced the computational complexity. Neural networks perform better when convolutions do not alter the dimensions of the input drastically. Reducing the dimensions too much may cause loss of information, known as a “representational bottleneck”. Additionally, using factorization methods, convolutions can be made more efficient in terms of computational complexity.

Inception v2 block was different to v1 in a way that 5x5 convolution was replaced by two 3x3 convolutions. A 5x5 convolution is almost 3 times as expensive as two stacked 3x3 convolutions. Convolutions of filter size  $N \times N$  were replaced by  $1 \times N$  and  $N \times 1$  convolutions. This results in the same output but is 33% cheaper.

Inception v3 incorporated the above upgrades and added some additional improvements:

- RMSProp optimizer.
- Factorized 7x7 convolutions.
- Label smoothing.
- Batch normalization in the auxiliary classifiers.

Inception v4 has different initial set of operations performed before introducing the Inception blocks. Inception v3 introduced specialized “Reduction blocks” which are used to change the width and height of the feature map.

Inception-ResNet is a hybrid architecture inspired by the performance of the ResNet. The Inception blocks were changed in order to make use of residual connections.

Xception [5] is a more recent version of Inception architecture (introduced in 2016). In authors' point of view, the Inception module is an intermediate step in-between regular convolution and the depthwise separable convolution operation (a depthwise convolution followed by a pointwise convolution). In this light, a depthwise separable convolution can be understood as an Inception module with a maximally large number of paths.

Xception stands for Extreme version of Inception. Xception performed even better than Inception v3. This model architecture makes use of the modified depthwise convolution:

1. The order of operations: the original depthwise separable convolutions as usually implemented perform first channel-wise spatial convolution and then perform 1x1 convolution whereas the modified depthwise separable convolution perform 1x1 convolution first then channel-wise spatial convolution. This is claimed to be unimportant because when it is used in stacked setting, there are only small differences appeared at the beginning and at the end of all the chained Xception modules.
2. In the original Inception module, there is a non-linearity (activation function) after first operation. In Xception there is no intermediate non-linear activation function. ReLU activation functions are used in Inception module. Yet it has a problem called "dying ReLU". "Dead" neurons always output the same value for any input. That means that it takes no role in discriminating between inputs. Once a neuron with ReLU activation ends up in this state, it is unlikely to recover, because the function gradient at 0 is also 0, so gradient descent learning will not alter neuron's weights. Many other activation functions were introduced later to counter this problem (Leaky ReLU, ELU, SELU, etc.).

Xception was evaluated on ILSVRC dataset and it outperformed VGGNet, ResNet and Inception v3.

In this master thesis Xception modules are in use inside digit and character recognition network. While wells detection was implemented using Inception

modules, detection of randomly oriented digits required a more sophisticated architecture.

In this section a few deep learning scientific papers were reviewed. Each of these papers can be called a major milestone in machine learning and computer vision. They greatly influenced the work described in the thesis.

## References

1. U-Net: Convolutional Networks for Biomedical Image Segmentation. Olaf Ronneberger, Philipp Fischer, Thomas Brox. [Электронный ресурс]. URL: <https://arxiv.org/abs/1505.04597>
2. SSD: Single Shot MultiBox Detector. Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg. [Электронный ресурс]. URL: <https://arxiv.org/abs/1512.02325>
3. Deep Residual Learning for Image Recognition. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. [Электронный ресурс]. URL: <https://arxiv.org/abs/1512.03385>
4. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, p. 1–9, 2015.
5. Xception: Deep Learning with Depthwise Separable Convolutions. François Chollet. [Электронный ресурс]. URL: <https://arxiv.org/abs/1610.02357>