

Инженерная школа информационных технологий и робототехники  
Направление подготовки 09.04.02 «Информационные системы и технологии»  
Отделение информационных технологий

### МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
<b>Разработка алгоритмов распознавания символов на изображениях табличек домов</b> УДК 004.021-047.84:004.932.75'1

Студент

Группа	ФИО	Подпись	Дата
8ИМ71	Белков Сергей Геннадьевич		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ	Спицын В. Г.	Д.Т.Н., профессор		

### КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель ОСГН	Потехина Н. В.	—		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель ООД	Атепаева Н.А.	—		

### ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ	Савельев А.О.	К.Т.Н.		

**ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ООП ПРОГРАММЫ  
МАГИСТРОВ 09.04.02 «ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ»**

Код	Результат обучения (выпускник должен быть готов)
<b>Общепрофессиональные и профессиональные компетенции</b>	
P1	Воспринимать и самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте.
P2	Владеть и применять методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе в глобальных компьютерных сетях.
P3	Демонстрировать культуру мышления, способность выстраивать логику рассуждений и высказываний, основанных на интерпретации данных, интегрированных из разных областей науки и техники, выносить суждения на основании неполных данных, анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями.
P4	Анализировать и оценивать уровни своих компетенций в сочетании со способностью и готовностью к саморегулированию дальнейшего образования и профессиональной мобильности. Владеть, по крайней мере, одним из иностранных языков на уровне социального и профессионального общения, применять специальную лексику и профессиональную терминологию языка.
P5	Разрабатывать стратегии и цели проектирования, критерии эффективности и ограничения применимости, новые методы, средства и технологии проектирования геоинформационных систем (ГИС) или промышленного программного обеспечения.
P6	Планировать и проводить теоретические и экспериментальные исследования в области создания интеллектуальных ГИС и ГИС технологии или промышленного программного обеспечения с использованием методов системной инженерии.
P7	Осуществлять авторское сопровождение процессов проектирования, внедрения и сопровождения ГИС и ГИС технологий или промышленного программного обеспечения с использованием методов и средств системной инженерии, осуществлять подготовку и обучение персонала.
P8	Формировать новые конкурентоспособные идеи в области теории и практики ГИС и ГИС технологий или системной инженерии программного обеспечения. Разрабатывать методы решения нестандартных задач и новые методы решения традиционных задач. Организовывать взаимодействие коллективов, принимать управленческие решения, находить компромисс между различными требованиями как при долгосрочном, так и при краткосрочном планировании.
<b>Общекультурные компетенции</b>	
P9	Использовать на практике умения и навыки в организации исследовательских, проектных работ и профессиональной эксплуатации современного оборудования и приборов, в управлении коллективом.
P10	Свободно пользоваться русским и иностранным языками как средством делового общения.
P11	Совершенствовать и развивать свой интеллектуальный и общекультурный уровень. Проявлять инициативу, в том числе в ситуациях риска, брать на себя всю полноту ответственности.
P12	Демонстрировать способность к самостоятельному обучению новым методам исследования, к изменению научного и научно-производственного профиля своей профессиональной деятельности, способность самостоятельно приобретать с помощью информационных технологий и использовать в практической деятельности новые знания и умения, в том числе в новых областях знаний, непосредственно не связанных со сферой деятельности, способность к педагогической деятельности.

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

Школа Информационных технологий и робототехники  
 Направление подготовки 09.04.02 «Информационные системы и технологии»  
 Отделение школы (НОЦ) Информационных технологий

УТВЕРЖДАЮ:  
 Руководитель ООП  
 \_\_\_\_\_ Савельев А.О.  
 (Подпись) (Дата) (Ф.И.О.)

### ЗАДАНИЕ на выполнение выпускной квалификационной работы

В форме:

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

Студенту:

Группа	ФИО
8ИМ71	Белкову Сергею Геннадьевичу

Тема работы:

<b>Разработка алгоритмов распознавания символов на изображениях табличек домов</b>
--

Утверждена приказом директора (дата, номер)	№ 1734/с от 05.03.2019 г
---	--------------------------

Срок сдачи студентом выполненной работы:	01.06.2019
--	------------

### ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

<b>Исходные данные к работе</b>	<ul style="list-style-type: none"> <li>• Объекты исследования: алгоритмы распознавания текста на изображениях адресных табличек домов</li> <li>• Предмет исследования: системы распознавания изображений</li> <li>• Документ, описывающий требования: задание на выполнение работы</li> </ul>
<b>Перечень вопросов, подлежащих исследованию, проектированию и разработке</b>	<ul style="list-style-type: none"> <li>• Особенности внешнего вида табличек адресов домов</li> <li>• Перевод изображений в градиенты серого</li> <li>• Бинаризация изображений</li> <li>• Выделение сегментов с буквенными или числовыми символами</li> <li>• Распознавание фотографий с числовыми и буквенными символами</li> <li>• Особенности проектирования нейронных сетей</li> <li>• Способы обучения нейронных сетей</li> </ul>

<b>Перечень графического материала</b>	<ul style="list-style-type: none"> <li>• Примеры адресных табличек домов</li> <li>• Иллюстрации работы алгоритма обработки изображений</li> <li>• Структура нейронных сетей и их составляющих</li> <li>• Графики функций активации</li> </ul>
<b>Консультанты по разделам выпускной квалификационной работы</b>	
<b>Раздел</b>	<b>Консультант</b>
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Старший преподаватель Отделения социально-гуманитарных технологий Потехина Нина Васильевна
Социальная ответственность	Старший преподаватель Отделения общетехнических дисциплин Атепаева Наталья Александровна
Раздел на английском языке	Доцент Отделения иностранных языков Сидоренко Татьяна Валерьевна Доцент Отделения информационных технологий Мирошниченко Евгений Александрович
<b>Названия разделов, которые должны быть написаны на русском и иностранном языках:</b>	
Анализ предметной области (Analysis of the objective area)	

<b>Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику</b>	01.02.2019
---	------------

**Задание выдал руководитель:**

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ	Спицын Владимир Григорьевич	д.т.н., профессор		01.02.2019

**Задание принял к исполнению студент:**

Группа	ФИО	Подпись	Дата
8ИМ71	Белков Сергей Геннадьевич		01.02.2019

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

Школа Информационных технологий и робототехники  
 Направление подготовки 09.04.02 «Информационные системы и технологии»  
 Уровень образования Магистратура  
 Отделение школы (НОЦ) Информационных технологий  
 Период выполнения осенний / весенний семестр 2018/2019 учебного года

Форма представления работы:

Магистерская диссертация
--------------------------

### КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН выполнения выпускной квалификационной работы

Срок сдачи студентом выполненной работы:	05.06.2019
--	------------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
20.02.2019	Анализ предметной области	20
15.03.2019	Разработка алгоритма	25
15.04.2019	Программная реализация	25
20.05.2019	Обучение и тестирование	15
25.05.2019	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	5
29.05.2019	Социальная ответственность	5
29.05.2019	Раздел на английском языке	5

**СОСТАВИЛ:**

**Руководитель ВКР**

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ	Спицын Владимир Григорьевич	д.т.н., профессор		01.02.2019

**СОГЛАСОВАНО:**

**Руководитель ООП**

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент отделения ИТ	Савельев Алексей Олегович	к.т.н.		01.02.2019

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

Студенту:

Группа	ФИО
8ИМ71	Белкову Сергею Геннадьевичу

Школа	ИШИТР	Отделение (НОЦ)	Отделение информационных технологий
Уровень образования	Магистратура	Направление/специальность	09.04.02 «Информационные системы и технологии»

Тема ВКР:

Разработка алгоритмов распознавания символов на изображениях табличек домов

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

1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	Стоимость материальных ресурсов определялась согласно прейскурантам компаний Оклад руководителя – 33664 р. Оклад инженера – 21760 р.
2. Нормы и нормативы расходования ресурсов	Накладные расходы 16 %; Районный коэффициент 30%; Норма амортизации ПЭВМ 33,33%;
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	Коэффициент отчислений на уплату во внебюджетные фонды 30,2%

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

1. Оценка коммерческого и инновационного потенциала НТИ	Анализ конкурентных технических решений, FAST-анализ, SWOT-анализ, оценка готовности проекта к коммерциализации.
2. Разработка устава научно-технического проекта	Определение заинтересованных сторон проекта, целей и результатов проекта. Разработка иерархической структуры работ.
3. Планирование процесса управления НТИ: структура и график проведения, бюджет, риски и организация закупок	Планирование этапов работы, исполнителей, формирование трудоемкости и графика разработки НИ. Расчет сметы затрат.
4. Определение ресурсной, финансовой, экономической эффективности исследования	Расчет научно-технического эффекта

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

- Оценочная карта сравнения конкурентности технических решений
- FAST-анализ
- Функционально-стоимостная диаграмма
- Результаты SWOT-анализа
- Бланк оценки степени готовности научного проекта к коммерциализации
- Иерархическая структура работ
- Линейный график работ
- Календарный план научно-исследовательского проекта

Дата выдачи задания для раздела по линейному графику	1.03.2019
--	-----------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель ОСГН	Потехина Нина Васильевна	—		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ИМ71	Белков Сергей Геннадьевич		

## ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

<b>Группа</b>	<b>ФИО</b>		
8ИМ71	Белкову Сергею Геннадьевичу		
<b>Школа</b>	<b>ИШИТР</b>	<b>Отделение (НОЦ)</b>	<b>Отделение информационных технологий</b>
<b>Уровень образования</b>	Магистратура	<b>Направление/специальность</b>	09.04.02 «Информационные системы и технологии»

Тема ВКР:

Разработка алгоритмов распознавания символов на изображениях табличек домов	
<b>Исходные данные к разделу «Социальная ответственность»:</b>	
1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения	Объект исследования – алгоритмы распознавания текста на изображениях адресных табличек домов Область применения – обработка изображений уличных панорам Рабочая зона – кабинет с рабочими местами: стол, стул, компьютер, сетевое оборудование для подключения к интернету.
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.	– Трудовой кодекс Российской Федерации от 30.12.2001 197-ФЗ (ред. от 01.04.2019). – СанПиН 2.2.4.548-96 – СанПиН 2.2.2/2.4.1340-03 – ГОСТ 12.2.032-78
<b>2. Производственная безопасность:</b> 2.1. Анализ выявленных вредных и опасных факторов 2.2. Обоснование мероприятий по снижению воздействия	Анализ выявленных вредных и опасных факторов: – отклонение показателей микроклимата; – повышенный уровень шума на рабочем месте; – недостаточная освещенность рабочей зоны; – повышенный уровень электромагнитных полей. – повышенное значение напряжения в электрической цепи, замыкание которой может произойти через тело человека.
<b>3. Экологическая безопасность:</b>	Негативно влияющие на экологию факторы могут быть связаны с эксплуатацией компьютера, на котором разрабатывается программное обеспечение.
<b>4. Безопасность в чрезвычайных ситуациях:</b>	Вероятные ЧС – пожар. Установка общих правил поведения и рекомендаций во время пожара, план эвакуации.

<b>Дата выдачи задания для раздела по линейному графику</b>	01.03.2019
---	------------

**Задание выдал консультант:**

<b>Должность</b>	<b>ФИО</b>	<b>Ученая степень, звание</b>	<b>Подпись</b>	<b>Дата</b>
Старший преподаватель ООД	Атепаева Наталья Александровна			

**Задание принял к исполнению студент:**

<b>Группа</b>	<b>ФИО</b>	<b>Подпись</b>	<b>Дата</b>
8ИМ71	Белков Сергей Геннадьевич		

## РЕФЕРАТ

Выпускная квалификационная работа 100 страниц, 38 рисунков, 34 источника, 30 таблиц и 2 приложения.

Ключевые слова: обработка изображений, распознавание символов, сверточные нейронные сети.

Объектом исследования являются алгоритмы распознавания текста на изображения, содержащих адресные таблички домов.

Цель работы – разработать алгоритм по распознаванию названия улицы и номера дома на адресной табличке дома.

В результате исследования были проанализированы виды адресных табличек домов, возможности их распознавания и применения распознавания адресов, разработан алгоритм распознавания текста на изображения, содержащих адресные таблички домов, реализован этап предобработки изображения для дальнейшего распознавания символов.

Область применения: картографические сервисы, обход системы защиты от спама reCaptcha.

Значимость работы заключается в определении местоположения здания по снимку для географической привязки адреса здания к карте.



## **ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ, СОКРАЩЕНИЯ И НОРМАТИВНЫЕ ССЫЛКИ**

**ИНС** – искусственная нейронная сеть

**ARGB** – Alpha Red Green Blue, обозначения каналов изображения

**YUV** – цветовая модель, в которой цвет состоит из трёх компонентов: яркость (Y) и два цветоразностных компонента (U и V).

**CNN** – сверточная нейронная сеть

**KNN** – Метод k-ближайших соседей

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ .....	14
1 АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ .....	15
1.1 Описание проблемы .....	15
1.2 Описание адресных табличек .....	15
1.3 Описание машинного обучения .....	17
1.4 Искусственные нейронные сети .....	18
1.4.1 Общее описание .....	18
1.4.2 Обобщенная структура ИНС .....	19
1.4.3 Нейроны смещения .....	22
1.4.4 Метод градиентного спуска .....	25
1.4.5 Метод обратного распространения .....	26
1.5 Описание сверточной нейронной сети .....	27
1.5.1 Свертка .....	28
1.5.2 Субдискретизация .....	29
1.5.3 Softmax .....	30
1.6 Описание алгоритма ближайших соседей .....	30
1.7 Анализ существующий решений .....	31
1.8 Вывод по разделу 1 .....	32
2 РАЗРАБОТКА АЛГОРИТМА .....	33
2.1 Описание алгоритма .....	33
2.2 Предобработка изображения .....	34
2.2.1 Преобразование в градации серого .....	34
2.2.2 Пороговая классификация изображения .....	35
2.2.3 Поиск сегментов изображения .....	36

2.3	Выбор структуры ИНС .....	37
2.3.1	Структура для распознавания таблички .....	37
2.3.2	Структура для распознавания символа .....	39
2.4	Формирование обучающей выборки .....	40
2.5	Обучение ИНС.....	41
3	ПРОГРАММНАЯ РЕАЛИЗАЦИЯ .....	42
3.1	Используемые технологии .....	42
3.2	Программные модули .....	42
3.2.1	Augment.....	43
3.2.2	KNN .....	43
3.2.3	CNN .....	43
3.2.4	Dataset.....	43
4	ОБУЧЕНИЕ И ТЕСТИРОВАНИЕ.....	44
4.1	Входные данные.....	44
4.2	Статистика обучения и распознавания .....	45
4.3	Демонстрация распознавания .....	46
5	ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ.....	49
5.1	Предпроектный анализ .....	49
5.1.1	Анализ конкурентных технических решений.....	49
5.2	FAST-анализ .....	51
5.3	SWOT-анализ.....	54
5.4	Оценка готовности проекта к коммерциализации.....	56
5.5	Инициация разработки .....	57
5.6	Организация и планирование работ .....	59

5.6.1	Иерархическая структура работ .....	59
5.6.2	План разработки.....	59
5.6.3	Продолжительность этапов работ .....	60
5.6.4	Расчет нарастания технической готовности работ.....	63
5.7	Расчет сметы затрат .....	65
5.7.1	Расчет заработной платы.....	65
5.7.2	Расчет отчислений от заработной платы .....	66
5.7.3	Расчет амортизации .....	66
5.7.4	Расчет накладных расходов .....	67
5.7.5	Расчет общей сметы.....	67
5.8	Оценка научно-технического уровня НИР.....	67
5.9	Вывод по разделу 5 .....	70
6	СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ.....	71
6.1	Правовые и организационные вопросы обеспечения безопасности .....	71
6.2	Правовые нормы трудового законодательства .....	71
6.2.1	Организационные мероприятия при компоновке рабочей зоны ...	72
6.3	Производственная безопасность .....	74
6.3.1	Отклонение показателей микроклимата.....	75
6.3.2	Превышение уровня шума .....	76
6.3.3	Недостаточность освещенности рабочей зоны.....	77
6.3.4	Повышенный уровень электромагнитных полей .....	78
6.3.5	Повышенное значение напряжения в электрической цепи, замыкание которой может произойти через тело человека. ....	79
6.4	Экологическая безопасность.....	79
6.5	Безопасность в чрезвычайных ситуациях.....	80

ЗАКЛЮЧЕНИЕ .....	83
СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ .....	84
ПРИЛОЖЕНИЕ А .....	87
ПРИЛОЖЕНИЕ Б.....	100

## **ВВЕДЕНИЕ**

Способность автоматически распознать адрес на фотографии с геопривязкой и связывать обработанный номер с известным адресом улицы, помогает с высокой степенью точности определить местоположение здания. Классическим примером таких фотографий являются изображения с «Google Street View» и «Яндекс.Панорамы», состоящие из большого количества географически привязанных панорамных изображений.

Для решения задачи распознавания адресных табличек могут быть применены искусственные нейронные сети (ИНС) с предварительной обработкой изображения. Разработка алгоритма распознавания на основе ИНС позволит достичь высокой скорости идентификации адреса для различных сфер применения.

**Целью** данной работы является разработка алгоритма распознавания адресных табличек домов с использованием нейронных сетей.

### **Задачи:**

- Изучение форматов табличек домов
- Изучение средств и методов, необходимых для обработки и распознавания изображений
- Разработка алгоритма

# **1 АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ**

## **1.1 Описание проблемы**

Распознавание номера улицы – это особый вид распознавания последовательности. Задача состоит в том, чтобы по заданному изображению идентифицировать номер дома и название улицы на изображении.

Хотя распознавание текста в ограниченных областях, таких как обработка текстовых документов, хорошо изучена, распознавание произвольного многосимвольного текста на фотографиях все еще остается сложной задачей. Эта трудность возникает из-за широкой изменчивости внешнего вида текста: из-за большого разнообразия шрифтов, цветов, стилей, ориентаций и расположения символов. Проблема распознавания дополнительно усложняется факторами окружающей среды, такими как: освещение, тени, зеркальность и окклюзия, а также факторами получения изображения, такими как разрешение, движение и размытие фокуса.

## **1.2 Описание адресных табличек**

Большинство домовых знаков и адресных табличек в Российской Федерации выполнены в соответствии стандарту ГОСТ Р 52290-2004 «Знаки дорожные. Общие технические требования» [1]. Однако, применение стандарта относится к выбору шрифтов и отступов. Взаимное расположение буквенных и цифровых обозначений может варьироваться от знака к знаку, что осложняет задачу распознавания. В ходе исследования были выявлены следующие виды адресных табличек:

1. Прямоугольные – стандартные таблички прямоугольной формы. Горизонтальное взаимное расположение цифровой и буквенной частей. Номер дома может находиться как слева, так и справа относительно названия улицы (рисунок 1.1).



Рисунок 1.1– Прямоугольные таблички

2. Закругленные – форма подложки может быть закруглена сверху или снизу. Возможно горизонтальное и вертикальное взаимное расположение цифровой и буквенной частей (рисунок 1.2).



Рисунок 1.2 – Закругленные таблички

3. Вертикальные – форма подложки в основном прямоугольная, взаимное расположение цифровой и буквенной частей вертикальное (рисунок 1.3).





Рисунок 1.3 – Вертикальные таблички

4. Номерные – адресная табличка содержит только номер дома. Форма подложки варьируется от прямоугольной до круглой (рисунок 1.4).



Рисунок 1.4 – Номерные таблички

В данной работе рассматривается распознавание знаков 1 типа.

### 1.3 Описание машинного обучения

Машинным обучением называется процесс, обработки большого числа примеров, выявления закономерностей для последующего прогнозирования характеристик новых данных [5].

Существуют два типа машинного обучения.

1. **Обучение по прецедентам**, или индуктивное обучение. Данный тип построен на использовании частных эмпирических данных для выявления общих закономерностей.

2. **Дедуктивное обучение** использует знания экспертов, составляет на их основе базу данных. Дедуктивное обучение обычно относят к сфере

экспертных систем. В связи с этим, тип обучения по прецедентам воспринимают как синоним термину машинное обучение.

Существуют разногласия в понимании разности между искусственным интеллектом и машинным обучением. В целом, программы, решающие простые для человека задачи, можно назвать искусственным интеллектом. В свою очередь, машинное обучение позволяет обучаться машине на примерах точно так же, как человеку легче понять что-нибудь на примере.

Помимо того, что машинное обучение находится на стыке таких дисциплин, как методы оптимизации и математическая статистика, оно также имеет свои особенности. Основные из них связаны с проблемами переобучения и вычислительной эффективности.

Большинство методов индуктивного обучения были разработаны в качестве альтернативы классическим статистическим подходам. Также они откосятся к категории извлечения информации (Data Mining).

## **1.4 Искусственные нейронные сети**

### **1.4.1 Общее описание**

Нейронная сеть – это последовательность нейронов, соединенных между собой синапсами. Структура нейронной сети пришла в мир программирования прямоком из биологии. Благодаря такой структуре, машина обретает способность анализировать и даже запоминать различную информацию. Нейронные сети также способны не только анализировать входящую информацию, но и воспроизводить ее из своей памяти [2].

Нейронные сети используются для решения сложных задач, которые требуют аналитических вычислений подобных тем, что делает человеческий мозг. Самыми распространенными применениями нейронных сетей является:

Классификация – распределение данных по параметрам. Например, на вход дается набор людей и нужно решить, кому из них давать кредит, а кому нет. Эту работу может сделать нейронная сеть, анализируя такую информацию как: возраст, платежеспособность, кредитная история и тд.

Предсказание – возможность предсказывать следующий шаг. Например, рост или падение акций, основываясь на ситуации на фондовом рынке.

Распознавание – в настоящее время, самое широкое применение нейронных сетей. Используется в Google, когда вы ищете фото или в камерах телефонов, когда оно определяет положение вашего лица и выделяет его и многое другое.

### **1.4.2 Обобщенная структура ИНС**

Главные составляющие нейронной сети: нейроны и связи между ними (синапсы).

**Нейрон** – это вычислительная единица, которая получает информацию, производит над ней простые вычисления и передает ее дальше. Они делятся на три основных типа: входной, скрытый и выходной.

У каждого из нейронов есть 2 основных параметра: входные данные (input data) и выходные данные (output data). В случае входного нейрона: Вход=Выход. В остальных, в поле input попадает суммарная информация всех нейронов с предыдущего слоя, после чего, она нормализуется, с помощью функции активации и попадает в поле output.

**Синапс** – это связь между двумя нейронами. У синапсов есть 1 параметр – вес. Благодаря ему, входная информация изменяется, когда передается от одного нейрона к другому.

Полная модель искусственного нейрона с синапсами изображена на рисунке 1.5.

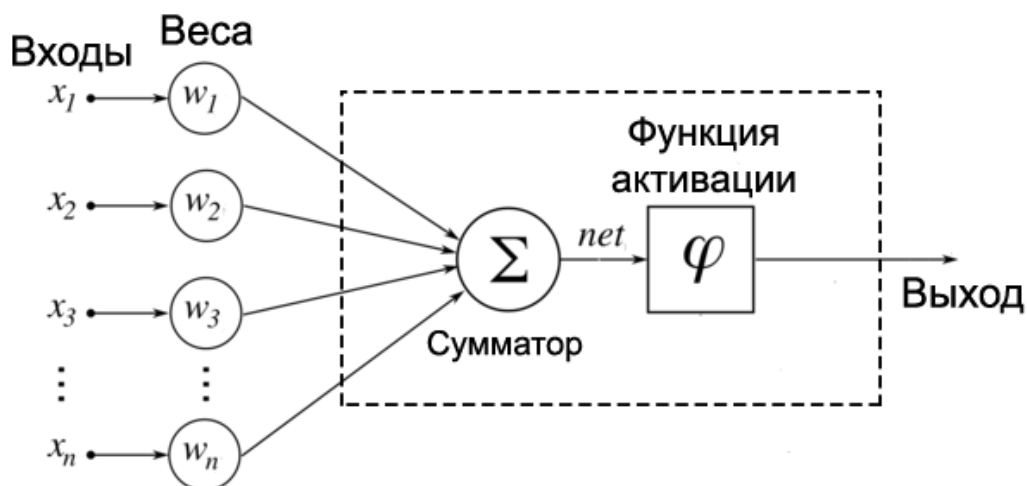


Рисунок 1.5 – Модель нейрона

В том случае, когда нейросеть состоит из большого количества нейронов, вводят термин **слоя**. Соответственно, есть входной слой, который получает информацию,  $n$  скрытых слоев (обычно их не больше 3), которые ее обрабатывают и выходной слой, который выводит результат.

На рисунке 1.6 приведен пример структуры ИНС с тремя слоями.

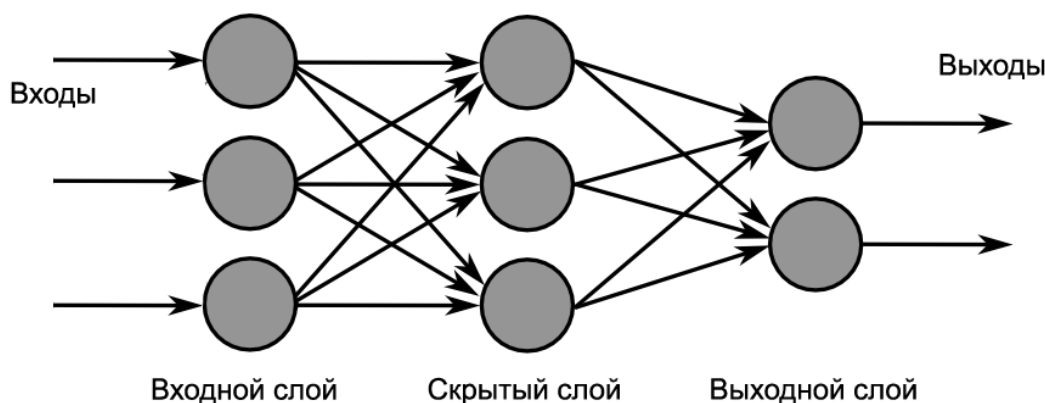


Рисунок 1.6 – Структура ИНС с тремя слоями.

**Функция активации** – это функция, принимающая взвешенную сумму как аргумент. Значение этой функции и является выходом нейрона [3].

Основные виды функций активации:

- **Функция единичного скачка.** Самый простой вид функции активации. Выход нейрона может быть равен только 0 или 1. Если взвешенная сумма больше определенного порога  $b$ , то выход нейрона равен 1. Если ниже, то 0 [4].

Графическое представление функции изображено на рисунке 1.7. На горизонтальной оси расположены величины взвешенной суммы. На вертикальной оси — значения выходного сигнала.

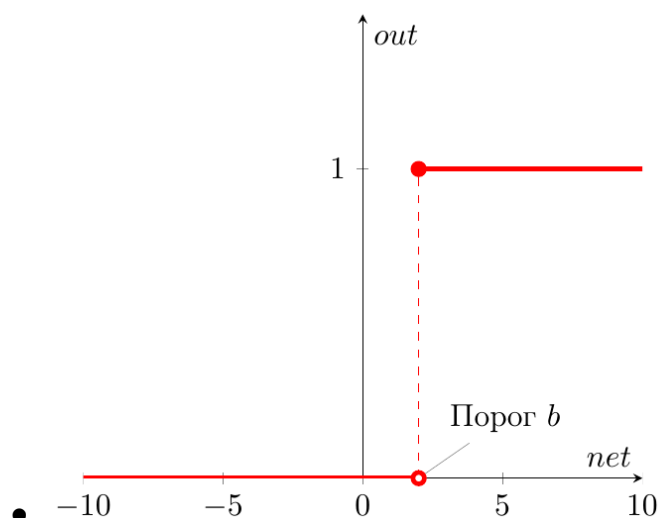


Рисунок 1.7 – Функция единичного скачка

- **Сигмоидальная функция.** Существует целое семейство сигмоидальных функций, некоторые из которых применяют в качестве функции активации в искусственных нейронах. Самая часто используемая в нейронных сетях сигмоида – логистическая функция (рисунок 1.8).

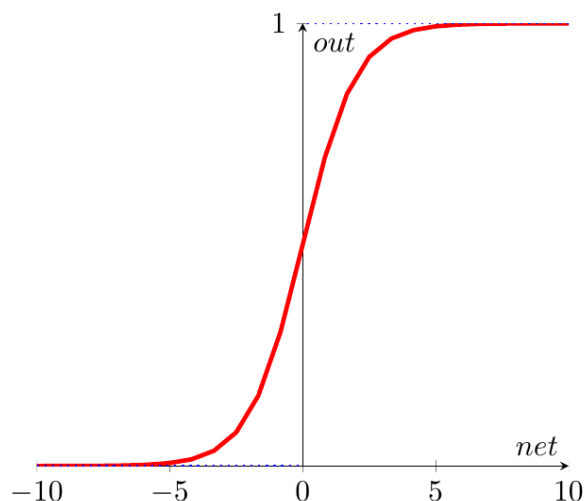


Рисунок 1.8 – Логистическая функция

Аналитическое представление функции выглядит следующим образом (формула 1.1):

$$out(net) = \frac{1}{1 + \exp(-a \cdot net)} \quad (1.1)$$

где  $a$  – коэффициент, характеризующий степень крутизны функции  
 $net$  – взвешенная сумма весов

Преимущества:

1. вне зависимости от аргумента (взвешенной суммы), выходной сигнал всегда будет в пределах от 0 до 1
2. более гибкая, чем функция единичного скачка, результатом может быть не только 0 и 1, но и любое число между ними
3. во всех точках она имеет производную, и эта производная может быть выражена через эту же функцию

- **Функция активации ReLU.** Выпрямленная линейная функция активации (rectified linear unit, ReLU) (рисунок 1.9). Вычисление сигмоиды и гиперболического тангенса требует выполнения ресурсоемких операций, таких как возведение в степень. Применение ReLU существенно повышает скорость сходимости градиентного спуска (в некоторых случаях до 6 раз по сравнению с сигмодой и гиперболическим тангенсом [15]).

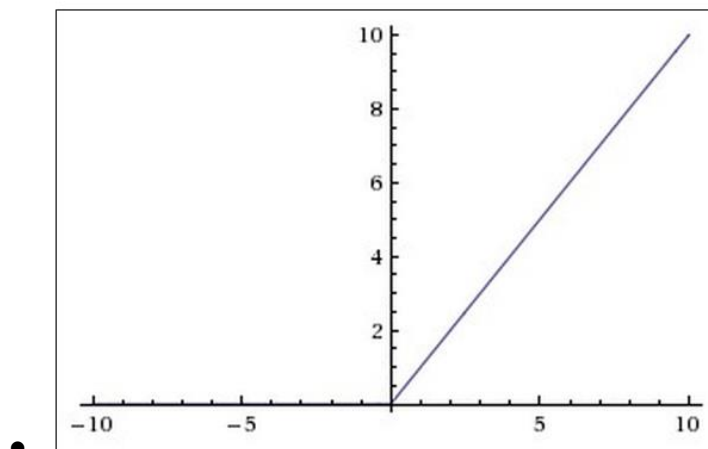


Рисунок 1.9 – Функция ReLU

Уравнение функции выражается формулой 1.2:

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \quad (1.2)$$

### 1.4.3 Нейроны смещения

Нейрон смещения или bias-нейрон – это вид нейронов, используемый в большинстве нейросетей. Особенность этого типа нейронов заключается в том,

что его вход и выход всегда равняются 1 и они никогда не имеют входных синапсов. Нейроны смещения могут, либо присутствовать в нейронной сети по одному на слое, либо полностью отсутствовать. Соединения у нейронов смещения такие же, как у обычных нейронов – со всеми нейронами следующего уровня, за исключением того, что синапсов между двумя bias нейронами быть не может. Следовательно, их можно размещать на входном слое и всех скрытых слоях, но никак не на выходном слое, так как им попросту не с чем будет формировать связь [17]. На рисунке 1.10 приведен пример нейрона смещения  $B_1$ .

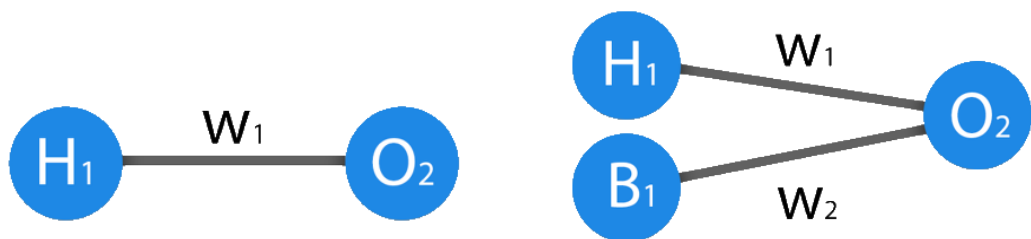


Рисунок 1.10 – Нейрон смещения

Нейрон смещения нужен для того, чтобы иметь возможность получать выходной результат, путем сдвига графика функции активации вправо или влево. Тогда можно установить, что выход  $O_2$  будет равен входу  $H_1$ , умноженному на его вес, и пропущенному через функцию активации (формулы 1.3, 1.4):

$$\text{input} = H_1 * w_1 + H_2 * w_2 + b_3 \quad (1.3)$$

$$b_3 = \text{bias} * w_3 \quad (1.4)$$

Когда в ходе обучения регулируются веса скрытых и выходных нейронов, меняется наклон функции активации (рисунок 1.11).

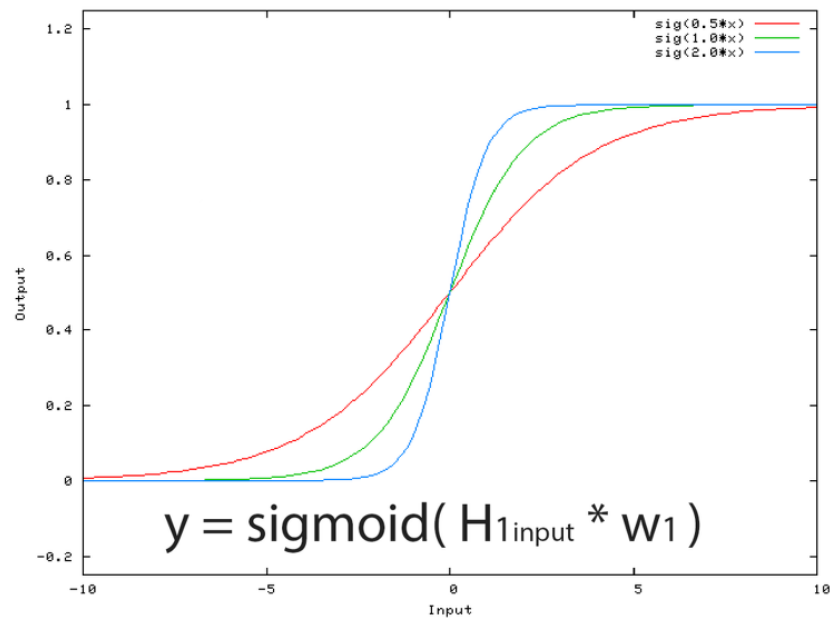


Рисунок 1.11 – Сигмоидальная функция без нейрона смещения

Регулирование веса нейронов смещения может дать нам возможность сдвинуть функцию активации по оси X и захватить новые участки (рисунок 1.12). Иными словами, если точка, отвечающая за решение, будет находиться на графике левее или правее, то нейронная сеть никогда не сможет решить задачу без использования нейронов смещения.

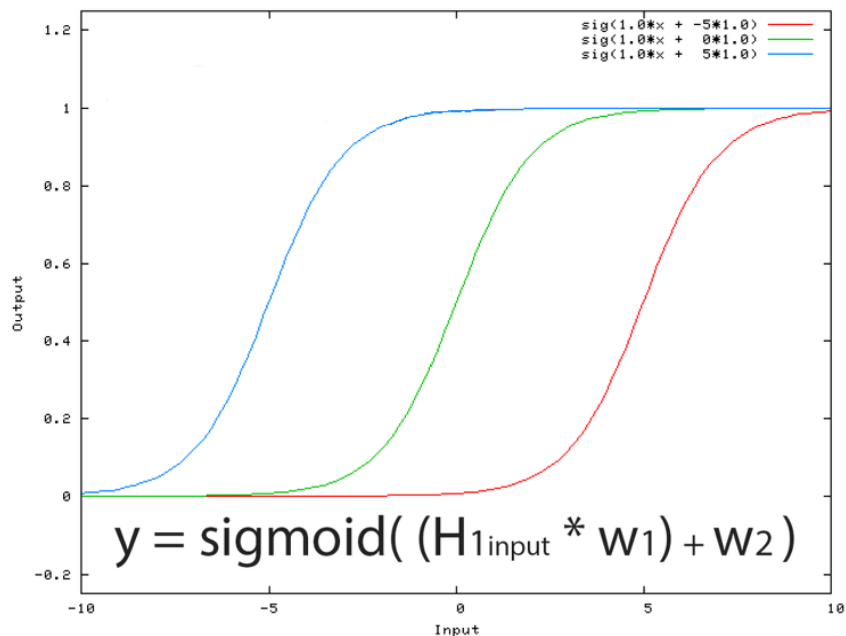


Рисунок 1.12 – Сигмоидальная функция с нейроном смещения



#### 1.4.4 Метод градиентного спуска

Градиентный спуск – метод нахождения локального экстремума (минимума или максимума) функции с помощью движения вдоль градиента [12]. Градиент представляет собой вектор, который определяет крутизну склона функции и указывает его направление относительно точки на поверхности или графике. Чтобы найти градиент нужно взять производную от функции по данной точке –  $f'(w_1)$  на рисунке 1.13.

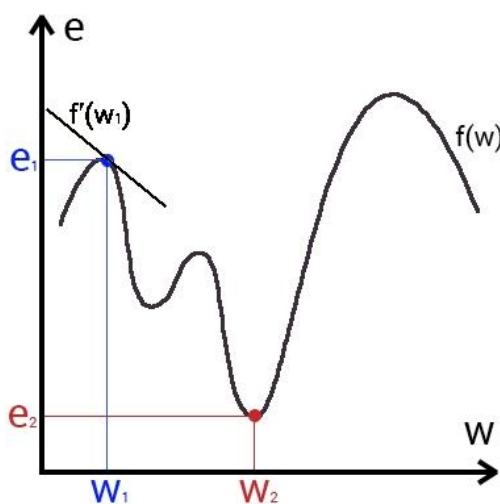


Рисунок 1.13 – Метод градиентного спуска

Метод градиентного спуска в чистом виде может «застевать» в локальных минимумах функции. Для преодоления проблемы метода градиентного спуска используется с методом моментов. Идея метода в сохранении части скорости спуска в каждой итерации. Фактически, происходит добавление к приращению веса приращения веса из предыдущей итерации, умноженной на коэффициент момента (формула 1.5) [13].

$$\Delta W_t = \eta \cdot \nabla E + \mu * \Delta W_{t-1} \quad (1.5)$$

где  $\eta$  – коэффициент скорости обучения,

$\nabla E$  – градиент функции потерь,

$\mu$  – коэффициент момента,

$\Delta W_{t-1}$  – изменение весов на предыдущей итерации.

Вместе с моментом в методе обратного распространения также используется такой параметр как скорость обучения. Он позволяет управлять величиной коррекции весов на каждой итерации.

#### **1.4.5 Метод обратного распространения**

Выполнение алгоритма начинается с создания произвольно сгенерированных весов для многослойной сети. Затем процесс, описанный ниже, повторяется до тех пор, пока средняя ошибка на входе не будет признана достаточно малой. Продвижение вперед по сети рассчитывает активации нейронов и выход, продвижение назад – градиент.

Шаг 1. Берется пример входного сигнала  $E$  с соответствующим правильным значением выхода  $C$ .

Шаг 2. Рассчитывается прямое распространение  $E$  через сеть – определяются весовые суммы  $S_i$  и функция активации  $f(in)$ .

Шаг 3. Начиная с выходов, выполняется обратное движение через нейроны выходного и промежуточного слоя, при этом рассчитывается значения ошибок (формулы 1.6 и 1.7):

для выходного нейрона:

$$\delta_0 = (OUT_{ideal} - OUT_{actual}) * f'(in) \quad (1.6)$$

для всех скрытых нейронов:

$$\delta_h = \sum (w_i * \delta_i) * f'(in) \quad (1.7)$$

где  $w$  – значения весов входящих синапсов для данного нейрона,

$\delta$  – ошибки для соответствующих весов,

$f(in)$  – функция активации.

Шаг 4. На данном шаге используется метод градиентного спуска, описанный в пункте 0. Веса в сети обновляются следующим образом (формула 1.8):

$$\Delta w_i = \varepsilon * grad_w + \alpha * \Delta w_{i-1} \quad (1.8)$$

где  $\varepsilon$  – скорость обучения,

$\alpha$  – момент

## 1.5 Описание сверточной нейронной сети

Свёрточные нейронные сети (convolutional neural network, CNN) сильно отличаются от других видов сетей. Обычно они используются для обработки изображений, реже для аудио. Типичным способом применения CNN является классификация изображений. Такие сети обычно используют «сканер», который не парсит все данные за один раз. Например, если имеется изображение  $200 \times 200$ , не будут обрабатываться все 40 тысяч пикселей. Вместо это сеть считает квадрат размера  $20 \times 20$  (обычно из левого верхнего угла), затем сдвинется на 1 пиксель и считает новый квадрат, и т.д. Эти входные данные затем передаются через свёрточные слои. Эти слои имеют свойство сжиматься с глубиной, причём часто используются степени двойки: 32, 16, 8, 4.

После каждого этапа свёртки следует этап подвыборки формируются несколько карт признаков, каждая карта признаков проходит этап подвыборки для уменьшения размерности и фильтрации незначимых деталей. После нескольких циклов свёртки и подвыборки, конечные карты признаков разворачиваются в вектор признаков. Схематическое изображение работы свёрточных слоёв изображено на рисунке 1.14.

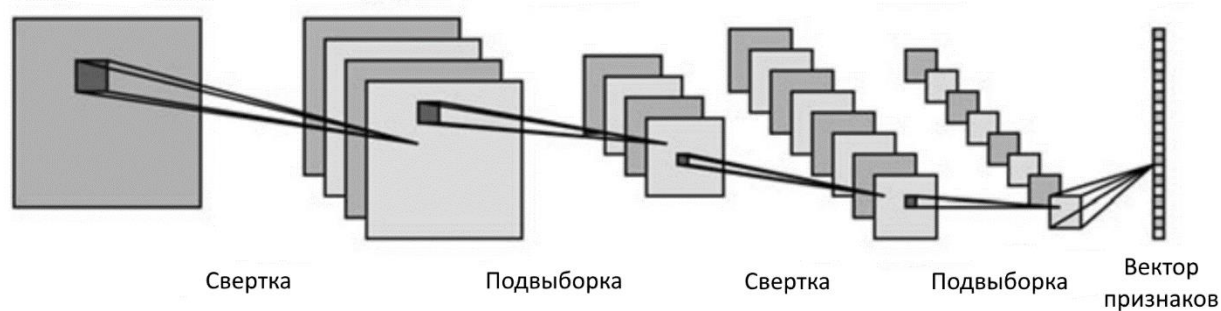


Рисунок 1.14 – Свёртка изображения

На практике к концу сверточной нейронной сети прикрепляют нейронные сети прямого распространения для дальнейшей обработки данных. Такие сети называются глубинным. На рисунке 1.15 представлено схематическое изображение свёрточной ИНС.

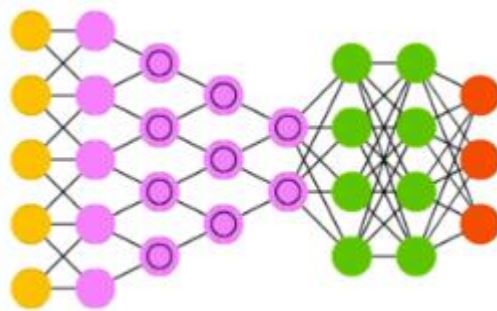


Рисунок 1.15 – Схема CNN

Сверточные нейронные сети обеспечивают частичную устойчивость к изменениям масштаба, смещениям, поворотам, смене ракурса и прочим искажениям. Сверточные нейронные сети объединяют три архитектурных идеи, для обеспечения инвариантности к изменению масштаба, повороту сдвигу и пространственным искажениям:

- локальные рецепторные поля (обеспечивают локальную двумерную связность нейронов);
- общие синаптические коэффициенты (обеспечивают детектирование некоторых черт в любом месте изображения и уменьшают общее число весовых коэффициентов);
- иерархическая организация с пространственными подвыборками.

На данный момент сверточная нейронная сеть и ее модификации считаются лучшими по точности и скорости алгоритмами нахождения объектов на сцене. Начиная с 2012 года, нейросети занимают первые места на известном международном конкурсе по распознаванию образов ImageNet.

Далее описаны основные операции свёрточной нейронной сети

### 1.5.1 Свертка

На этапе свёртки происходит поэлементное умножение матрицы изображения на фильтр, сложение всех элементов и присвоение его центральному пикселю фильтра. Затем фильтр сдвигается на 1 пиксель вправо и повторяется умножение матриц. Графическое представление операции свёртки представлено на рисунке 1.16.

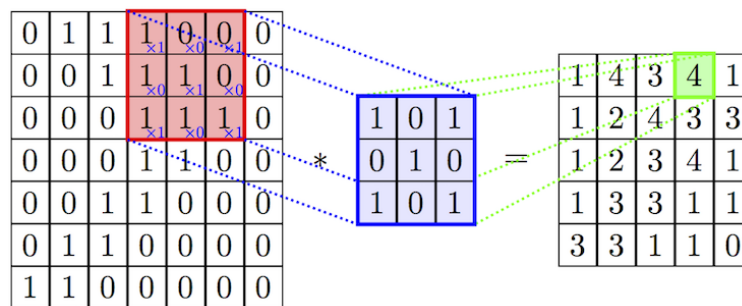


Рисунок 1.16 – Операция свёртки

Задача построения свёрточных слоёв состоит в настройке параметров фильтров таким образом, чтобы наилучшим образом извлечь уникальные характеристики изображения.

Для сохранения размерности после операции свёртки используется нулевой отступ (zero padding) – добавление к границам карты нулевых пикселей. Это приводит к заполнению матрицы таким образом, что размерность матрицы вывода будет такой же, как у исходного ввода (рисунок 1.17).

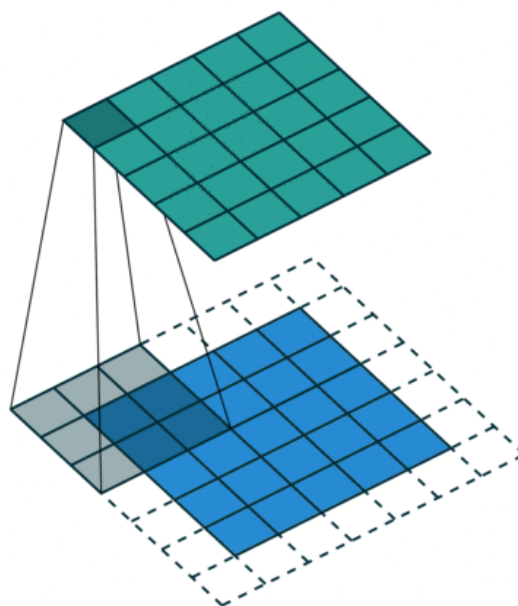


Рисунок 1.17 – Свёртка с нулевым отступом

### 1.5.2 Субдискретизация

Цель этапа – уменьшение размерности карт предыдущего слоя. Если на предыдущей операции свертки уже были выявлены некоторые признаки, то для дальнейшей обработки настолько подробное изображение уже не нужно, и оно

уплотняется до менее подробного. К тому же фильтрация уже ненужных деталей помогает не переобучаться.

В процессе сканирования ядром подвыборочного слоя (фильтром) карты предыдущего слоя, сканирующее ядро не пересекается в отличие от сверточного слоя. Обычно, каждая карта имеет ядро размером 2x2, что позволяет уменьшить предыдущие карты сверточного слоя в 2 раза. Вся карта признаков разделяется на ячейки 2x2 элемента, из которых выбираются максимальные по значению. Операция подвыборки (или MaxPooling – выбор максимального) изображена на рисунке 1.18.

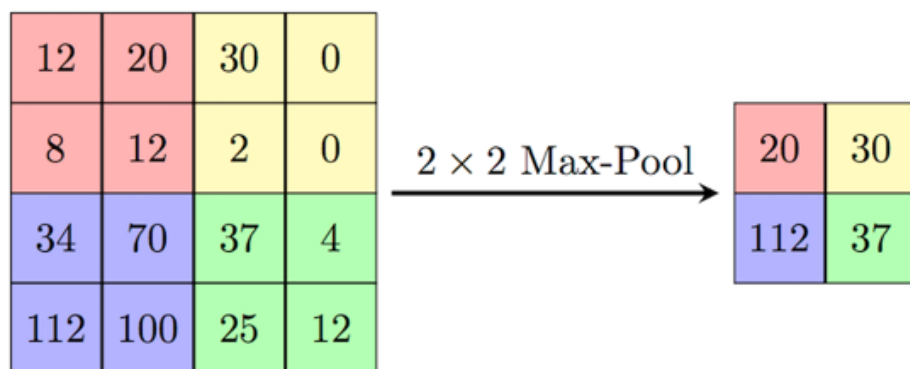


Рисунок 1.18 – Операция подвыборки

### 1.5.3 Softmax

Функция Softmax применяется в машинном обучении для задач классификации, когда количество возможных классов больше двух (для двух классов используется логистическая функция). Сумма всех выходных сигналов равна 1. Координаты  $\sigma_i$  полученного вектора при этом трактуются как вероятности того, что объект принадлежит к классу  $i$  (формула 1.9).

$$\sigma(x_j) = \frac{e^{x_j}}{\sum_i e^{x_i}} \quad (1.9)$$

## 1.6 Описание алгоритма ближайших соседей

Алгоритм k ближайших соседей (k Nearest Neighbor) – это один из достаточно простых алгоритмов классификации. Алгоритм является классическим в области Machine Learning. Задача классификации в машинном обучении – это задача отнесения объекта к одному из заранее определенных

классов на основании его формализованных признаков. Каждый из объектов в этой задаче представляется в виде вектора в N-мерном пространстве, каждое измерение в котором представляет собой описание одного из признаков объекта.

Для обучения классификатора необходимо иметь набор объектов, для которых заранее определены классы. Это множество называется обучающей выборкой, её разметка производится вручную, с привлечением специалистов в исследуемой области.

Когда появляются неразмеченные данные, алгоритм проходит два основных шага:

1. Поиск k ближайших размеченных точек данных (k ближайших соседей)
2. На основе классов соседей, kNN определяет к какому классу отнести новые данные (классификация)

Для упорядоченных значений атрибутов данных находится Евклидово расстояние (формула 1.10):

$$D_E = \sqrt{\sum_i^n (x_i - y_i)^2} \quad (1.10)$$

### **1.7 Анализ существующий решений**

На данный момент единственным найденным решением в сфере распознавания номеров домов является технология, используемая в проекте Google Maps [14]. Технология разработана в университете Карнеги-Меллона, которую в последствии приобрела корпорация Google. Эта технология находит и считывает номера улиц в режиме просмотра улиц и сопоставляет их с существующими адресами, чтобы точно определить точное местоположение дома на Картах Google. В статье [16] указано, что эта система способна точно определять и считывать сложные числа в режиме просмотра улиц с точностью до 96%. Также результаты системы используются Google reCAPTCHA для демонстрации пользователям. «reCAPTCHA» - один из самых безопасных

обратных тестов тьюринга, который использует искаженный текст и изображения, чтобы отличить людей от ботов.

### **1.8 Вывод по разделу 1**

На момент проведения аналитического анализа в свободном доступе не найдено готовых решений в сфере распознавания адресных табличек домов, что делает невозможным сравнение с аналогами.

Для разработки алгоритма выбрано использование детектора границ Канни для этапа предобработки изображений. Среди найденных границ необходимо обнаружить замкнутые контуры. Обработка найденных контуров производится свёрточной ИНС, так как она обеспечивает устойчивость к смещениям, поворотам, смене ракурса и прочим искажениям.

В ходе обучения нейронной сети будут использоваться нескольких тысяч изображений, что сильно скажется на скорости обучения, поэтому в качестве функции активации выбрана функция ReLU как обеспечивающая высокую скорость сходимости.



## 2 РАЗРАБОТКА АЛГОРИТМА

### 2.1 Описание алгоритма

Алгоритм состоит из трёх главных этапов: поиск таблички, распознавание текста самой таблички и вывод распознанного текста (рисунок 2.1).



Рисунок 2.1 – Общий алгоритм

Последовательность работы алгоритма на этапах «Поиск таблички» и «Распознавание текста» имеет следующий вид:

Шаг 1. Преобразование цветного изображения в градации серого.

Шаг 2. Пороговая классификация изображения.

Шаг 3. Поиск сегментов изображения с символами таблички.

Шаг 4. Фильтрация сегментов.

Шаг 5. Распознавание сегментов.

Первые 4 шага представляют собой предобработку изображения. Выделение отдельных сегментов с символами упрощает задачу для его распознавания. В случае распознавания текста, вместо поиска на всём изображении, задача сводится к распознаванию каждого символа отдельно.

На 5 шаге для классификации сегментов используются отдельно друг от друга два подхода:

- алгоритм ближайшего соседа
- свёрточная ИНС

Описанный выше алгоритм предназначен как для поиска адресной таблички на изображении, так и для распознавания символов на табличке. Для

этого на 4 шаге используются две разные нейросети, первая обучена на распознавание таблички, вторая – на распознавание символов.

Алгоритм ближайшего соседа был выбран в качестве классического алгоритма классификации в машинном обучении. На основе его результатов будет приведена целесообразность применения нейронных сетей для задачи классификации символов адресной таблички.

Идентифицируемое число представляет собой последовательность символов,  $s = s_1, s_2, \dots, s_n$ . При определении точности цифрового преобразователя вычисляется отношение входных изображений, для которых каждый элемент  $s_i$  последовательности длины  $n$  предсказаны правильно.

## 2.2 Предобработка изображения

Далее приведено описание каждого шага алгоритма на примере нижеприведенной адресной таблички (рисунок 2.2).



Рисунок 2.2 – Входное изображение

### 2.2.1 Преобразование в градации серого

Каждый пиксель изображения формируется при помощи сочетания четырех каналов ARGB: альфа-канала, красного, зеленого и синего. Альфа-канал отвечает за прозрачность пикселя (100% – пиксель полностью непрозрачный, 0% – полностью прозрачный). Сочетание значений RGB каналов определяет цвет пикселя. Каждый канал несёт в себе 8 бит информации (1 байт), соответственно интенсивность канала может меняться в диапазоне от 0 до 255 [6].

Для того, чтобы преобразовать цветное изображение в черно-белое, нужно воспользоваться формулой перевода из RGB в YUV (формула 2.1):

$$Y = 0.299 * R + 0.587 * G + B * 0.114 \quad (2.1)$$

Полученное значение присвоить RGB каналам этого же пикселя (то есть оно будет одинаковое). Альфа-канал оставляем без изменений. Такую операцию нужно проделать с каждым пикселем изображения. В OpenCV для преобразования изображения в градации серого используется функция `cvtColor()` [7]. Пример перевода цветного изображения в градации серого представлен на рисунке 2.3.



Рисунок 2.3 – Изображение в градациях серого

## 2.2.2 Пороговая классификация изображения

Пороговая классификация представляет собой определение каждого пикселя изображения к определённому классу, используя пороговое значение пикселя.

Самый простой способ для перевода изображения в двухцветное состояние – задать порог. Например, если у изображения 256 цветов, то можно задать порог 128. Основные методы бинаризации изображения [8]:

1. Отсечение по порогу яркости.
2. Метод Отса.
3. Метод Берисена.
4. Метод Эйквеля.
5. Метод Ниблэка.

Учитывая частое использование в табличках всего двух цветов, в данной работе была достаточной бинаризация по порогу яркости. В качестве порога задавалось среднее значение пикселей. Пример перевода бинаризованного изображения представлен на рисунке 2.4.



Рисунок 2.4 – Бинаризованное изображение

### 2.2.3 Поиск сегментов изображения

В данной работе для поиска границ используется метод Канни как самый популярный метод выделения границ. Хотя работа Канни была проведена на заре компьютерного зрения (1986 г.), детектор границ Канни до сих пор является одним из лучших детекторов [9].

Шаги детектора:

Шаг 1. Убрать шум и лишние детали из изображения.

Шаг 2. Рассчитать градиент изображения.

Шаг 3. Сделать края тонкими (edge thinning).

Шаг 4. Связать края в контуры (edge linking).

Детектор использует фильтр на основе первой производной от гауссианы. Так как он восприимчив к шумам, необходимо применять данный метод на предварительно обработанных изображениях. Сначала, исходные изображения нужно свернуть с гауссовым фильтром.

В OpenCV, детектор границ Канни реализуется функцией `cvCanny()`, которая обрабатывает только одноканальные изображения [7].

Из найденных контуров находятся замкнутые и исключаются те, которые не содержат цифры или буквы. На рисунке 2.5 представлен пример выделения сегментов.



Рисунок 2.5 – Сегментированное изображение

## 2.3 Выбор структуры ИНС

Структура нейронной сети в общем случае представляет собой входной, выходной и скрытый слои. Количество нейронов во входном слое определяется длиной вектора признаков. Количество нейронов в скрытом слое определяется эмпирическим путём. Количество нейронов в выходном слое определяется количеством распознаваемых классов. В случае распознавания домовых знаков, количество классов составляет 45 (10 цифр, 33 буквы, точка, знак дроби).

Для определения количества нейронов для скрытого слоя существует два подхода: усложнение и упрощение структуры ИНС. В первом случае берётся количество нейронов, которого заведомо не хватает для получения хорошего результата. После чего добавляется новый нейрон или слой, если после его добавления ошибка уменьшается. Во втором случае наоборот берётся ИНС с избыточной структурой. Пока ошибка падает, удаляется нейрон или слой.

### 2.3.1 Структура для распознавания таблички

Схематичное изображение слоёв ИНС для обнаружения таблички представлено на рисунке 2.6.

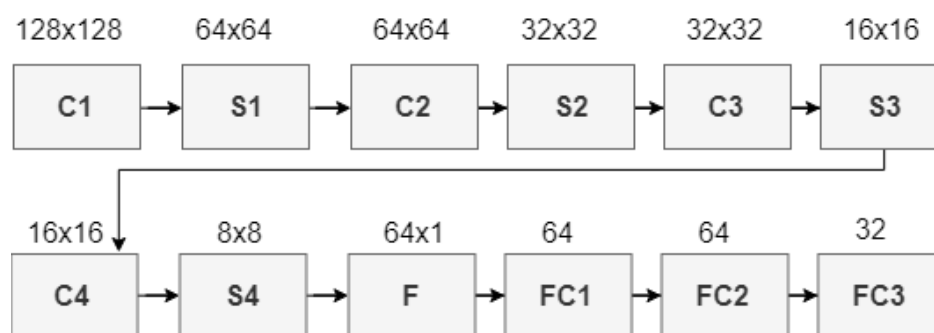


Рисунок 2.6 – Слои первой ИНС

Первые 4 пары слоёв представляют чередование операций свёртки и подвыборки:

- Свёрточный слой (C1)

Размер карты 128x128. Размер фильтра 3x3. Количество фильтров 4.

На выходе 4 карты признаков.

- Слой подвыборки (S1)

Размер карты 64x64. Размер фильтра 2x2. На выходе 4 карты признаков.

- Свёрточный слой (C2)

Размер карты 64x64. Размер фильтра 3x3. Количество фильтров 4.

На выходе 16 карт признаков.

- Слой подвыборки (S2)

Размер карты 32x32. Размер фильтра 2x2. На выходе 16 карт признаков.

- Свёрточный слой (C3)

Размер карты 32x32. Размер фильтра 3x3. Количество фильтров 2. На выходе 32 карты признаков.

- Слой подвыборки (S3)

Размер карты 16x16. Размер фильтра 2x2. На выходе 16 карт признаков

- Свёрточный слой (C4)

Размер карты 16x16. Размер фильтра 3x3. Количество фильтров 2. На выходе 64 карты признаков.

- Слой подвыборки (S4)

Размер карты 8x8. Размер фильтра 2x2. На выходе вектор признаков длиной 64 (F)

Последними в нейронной сети расположены слои FC1, FC2 и FC3. Они представляют собой классический трехслойный персептрон и состоят из обычных нейронов. Данные слои предназначены для осуществления классификации признаков, которые были извлечены на предыдущих слоях нейронной сети.

- Полносвязный слой (FC1)

Состоит из 64 нейронов, каждый принимает на вход один признак из вектора признаков.

- Полносвязный слой (FC2)

Состоит из 32 нейронов.

- Полносвязный слой (FC3)

Состоит из 1 нейрона. Выход нейрона показывает присутствие адресной таблички.

### 2.3.2 Структура для распознавания символа

Схематичное изображение слоёв ИНС для распознавания символов представлено на рисунке 2.7.

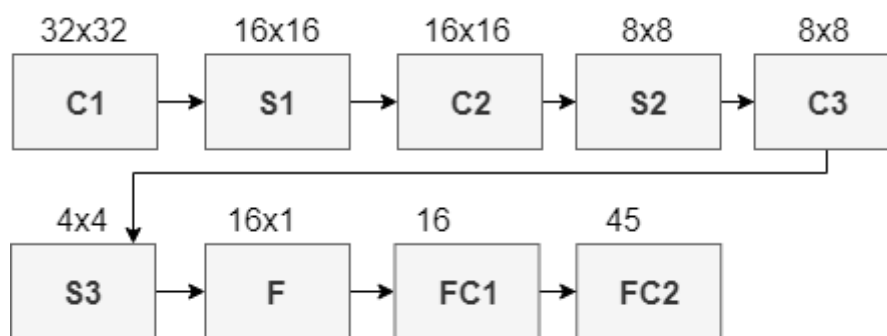


Рисунок 2.7 – Слои второй ИНС

- Свёрточный слой (C1)

Размер карты 32x32. Размер фильтра 3x3. Количество фильтров 8.

На выходе 8 карт признаков.

- Слой подвыборки (S1)

Размер карты 16x16. Размер фильтра 2x2. На выходе 8 карт признаков.

- Свёрточный слой (C2)

Размер карты 16x16. Размер фильтра 3x3. Количество фильтров 8.

На выходе 64 карты признаков.

- Слой подвыборки (S2)

Размер карты 8x8. Размер фильтра 2x2. На выходе 64 карты признаков.

- Свёрточный слой (C3)

Размер карты 8x8. Размер фильтра 3x3. Количество фильтров 2.

На выходе 128 карт признаков.

- Слой подвыборки (S3)

Размер карты 4x4. Размер фильтра 2x2. На выходе вектор признаков длиной 128.

- Полносвязный слой (FC1)

Состоит из 128 нейронов, каждый принимает на вход один признак из вектора признаков.

- Полносвязный слой (FC1)

Состоит из 128 нейронов, каждый принимает на вход один признак из вектора признаков.

## **2.4 Формирование обучающей выборки**

Для обучения нейронной сети необходимо предварительно сформировать датасет – набор данных с заранее определенной принадлежностью к классам. В данной работе датасетом являются изображения символов, встречающихся в адресных табличках домов, с обозначением изображенного символа. Непосредственно из датасета формируется обучающая выборка.

Решение задачи формирования обучающей выборки может производиться с помощью стандартного статистического подхода – перекрёстной проверки. Идея подхода состоит в проверке модели на данных отличных от используемых во время обучения.

В данном методе происходит разбиение имеющегося в наличии набора данных на обучающее множество и тестовое множество. Обучающее множество разбивается на два несвязанных подмножества:

– Обучающее подмножество, используемое для формирования модели.



– Тестовое подмножество, используемое для проверки точности распознавания модели.

Так как в свободном доступе не был найден подходящий готовый датасет, данные были собраны вручную с помощью сервисов «Google Maps» и «Яндекс Панорамы». Процесс сбора достаточного для обучения количества изображений занимал слишком много времени, поэтому для расширения уже собранного датасета был составлен и применен алгоритм, вносящий искажения и шумы.

## **2.5 Обучение ИНС**

Обучение представляет собой настройку весов нейронов с помощью моделирования среды, в которую встроена ИНС [10]. Данное определение процесса обучения нейронной сети подразумевает следующие этапы:

1. Нейронная сеть принимает внешние стимулы.
2. Изменение свободных параметров сети в сторону уменьшения ошибки.
3. После обновления параметров нейронная сеть отвечает на возбуждения другим способом.

На втором этапе для изменения параметров сети применяется алгоритм обратного распространения ошибки. Обратное распространение является одним из самых часто используемых алгоритмов для изменения весов связей при обучении. Ошибка распространяется от выходного к входному слою: в противоположном направлении от прохождения сигнала сети [11].

### 3 ПРОГРАММНАЯ РЕАЛИЗАЦИЯ

#### 3.1 Используемые технологии

Алгоритм реализован с помощью языка программирования Python [18], библиотеки для работы с изображениями OpenCV [19], библиотеки для машинного обучения Tensorflow [20].

#### 3.2 Программные модули

Для структурирования и разделения логики программная реализация алгоритма разнесена на несколько модулей. На рисунке 3.1 изображены модули и связь между ними.

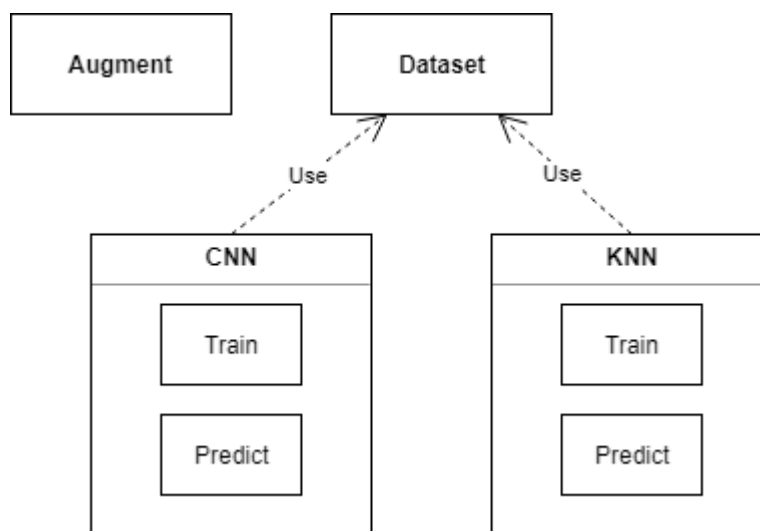


Рисунок 3.1 – Диаграмма модулей

Каждый из элементов выполняет следующие функции:

Augment – модуль для расширения исходной выборки естественных изображений с помощью добавления искажений и шумов.

KNN – модуль для обучения и распознавания алгоритма ближайшего соседа. Содержит подмодули для обучения и распознавания, Train и Predict соответственно.

CNN – модуль для обучения свёрточной нейронной сети и распознавания с помощью неё. Также содержит свои подмодули Train и Predict.

Dataset – чтение и подготовка изображений из датасета.

Далее описаны основные методы каждого модуля

### 3.2.1 Augment

rotate – поворот изображения на случайный градус

addBlur – добавление случайного размытия

addNoise – добавление случайного шума

### 3.2.2 KNN

flattenImage – перевод изображения в одномерный массив для сохранения в файл. На рисунке 3.2 представлен файл, каждая строка которого является одномерным изображением.

flattenClassifications – перевод набора классификаторов в одномерный массив для передачи в метод train.

train – запуск тренировки KNN-модели.

1	5.1000000000000000e+01	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02
2	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02
3	8.0000000000000000e+01	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02
4	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02
5	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02
6	5.0000000000000000e+01	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02
7	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02
8	7.1000000000000000e+01	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02
9	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02
10	0.0000000000000000e+00	0.0000000000000000e+00	0.0000000000000000e+00	0.0000000000000000e+00	0.0000000000000000e+00	0.0000000000000000e+00
11	8.3000000000000000e+01	2.5500000000000000e+02	1.9800000000000000e+02	2.6000000000000000e+01	0.0000000000000000e+00	0.0000000000000000e+00
12	0.0000000000000000e+00	2.1000000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.5500000000000000e+02	2.6000000000000000e+01
13	1.7300000000000000e+02	2.5500000000000000e+02	1.6000000000000000e+01	0.0000000000000000e+00	0.0000000000000000e+00	0.0000000000000000e+00
14	8.5000000000000000e+01	2.5500000000000000e+02	1.6700000000000000e+02	2.1000000000000000e+01	0.0000000000000000e+00	0.0000000000000000e+00

Рисунок 3.2 – Файл с одномерными изображениями

### 3.2.3 CNN

createConvolutionalLayer – создание сверточного слоя с этапом подвыборки после свертки

createFlattenLayer – преобразование карт признаков в одномерный массив для передачи в полносвязный слой

createFcLayer – создание полносвязного с предыдущим слоя

showProgress – вывод статистики по каждой эпохе

train – запуск построенного графа нейронной сети

### 3.2.4 Dataset

loadTrain – чтение всех изображений из датасета, приведение изображений к одному размеру, предобработка.

readTrainSets – вызов метода load\_train, разделение датасета на тренировочную и проверочную выборки. На проверочную выборку отводится 20% от всего датасета.

## 4 ОБУЧЕНИЕ И ТЕСТИРОВАНИЕ

### 4.1 Конфигурация оборудования

Обучение и тестирование проводилось на персональном компьютере со следующей конфигурацией:

- процессор: Intel Core i5-3210M 2.5 ГГц;
- ОЗУ: 8 Гб;
- ПЗУ: твердотельный накопитель Kingston SA400S37;
- операционная система: Linux Ubuntu;

### 4.2 Входные данные

Пример входных данных для обучения свёрточной ИНС для поиска адресных табличек представлен на рисунке 4.1.



Рисунок 4.1 – Выборка исходных изображений

Пример естественных изображений для распознавания символов, собранных с помощью открытых источников представлены на рисунке 4.2.



Рисунок 4.2 – Естественные изображения символов

Изображения, полученные с помощью модуля Augment. Пример изображений расширенной выборки для буквы «Е» представлен на рисунке 4.3.

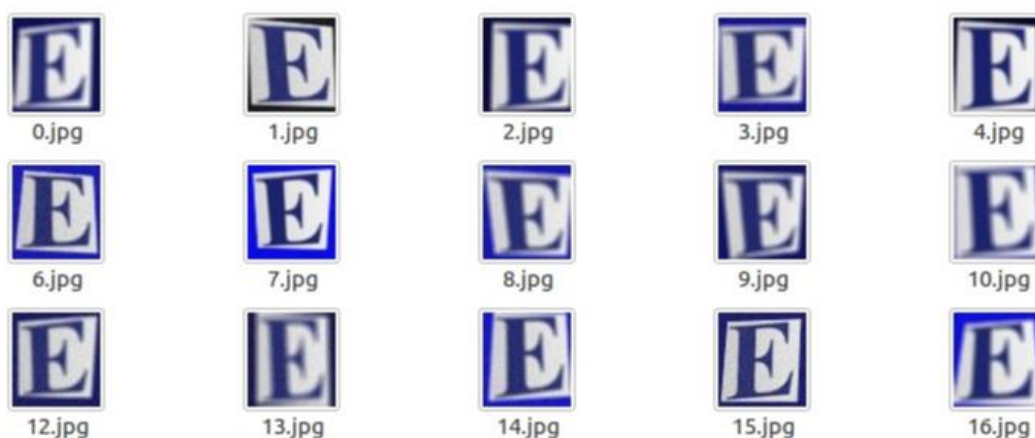


Рисунок 4.3 – Расширенная выборка символов

### 4.3 Статистика обучения и распознавания

Результаты точности для поиска адресных табличек представлены в таблице 4.1.

Таблица 4.1 – Результаты поиска адресных табличек

Слои	Размер выборки на каждый класс	Точность	Эпохи	Время обучения
3 сверточных	500	72,8 %	55	1.2ч
4 сверточных	500	84,3 %	71	1.5ч

Следующая проверка обучения ИНС производилась с целью подтвердить целесообразность расширения обучающей выборки путём внесения искажений. Сравнение результатов приведено в таблице 4.2.

Критерием прекращения работы алгоритма обучения является малое изменение ошибки в течение 10 итераций.

Таблица 4.2 – Результаты распознавания ИНС

<b>Тип данных для обучения</b>	<b>Размер выборки на каждый класс</b>	<b>Точность</b>	<b>Эпохи</b>	<b>Время обучения</b>
Естественные изображения	100	75,7 %	129	8,5ч
Расширенная выборка	500	89,2 %	320	17,8ч

Исходя из проведенного эксперимента, видно значительное возрастание точности распознавания. Для дальнейших экспериментов применяется расширенная выборка.

Для проверки целесообразности применения для распознавания свёрточной ИНС был проведён сравнительный тест с алгоритмом ближайшего соседа (KNN). Результаты приведены в таблице 4.3.

Таблица 4.3 – Результаты распознавания KNN

<b>Тип данных для обучения</b>	<b>Размер выборки на каждый класс</b>	<b>Точность</b>	<b>Время</b>
Естественные изображения	100	36,3%	23с
Расширенная выборка	500	38,9 %	120с

По результатам в таблицах 4.2 и 4.3 заметно значительное возрастание точности распознавания при применении ИНС по сравнению с алгоритмом ближайшего соседа.



#### 4.4 Демонстрация распознавания

Изображение для демонстрации алгоритма представлено на рисунке 4.4.



Рисунок 4.4 – Исходное изображение

Для выделения контуров использовался поиск с помощью библиотеки OpenCV с модификатором поиска самых внешних контуров RETR\_EXTERNAL. Из найденных контуров отбираются контуры с минимальной площадью в 1000 пикселей. Результат выделения контуров представлен на рисунке 4.5.



Рисунок 4.5 – Выделение контуров

Выделенные контуры последовательно обрабатываются на предмет присутствия адресной таблички. Контур, содержащий адресную табличку подаётся на следующий цикл сегментации с поиском всех контуров на изображении. Каждый контур передаётся на вход второй свёрточной ИНС.

Графическое отображение сегментации изображения представлено на рисунке 4.6.

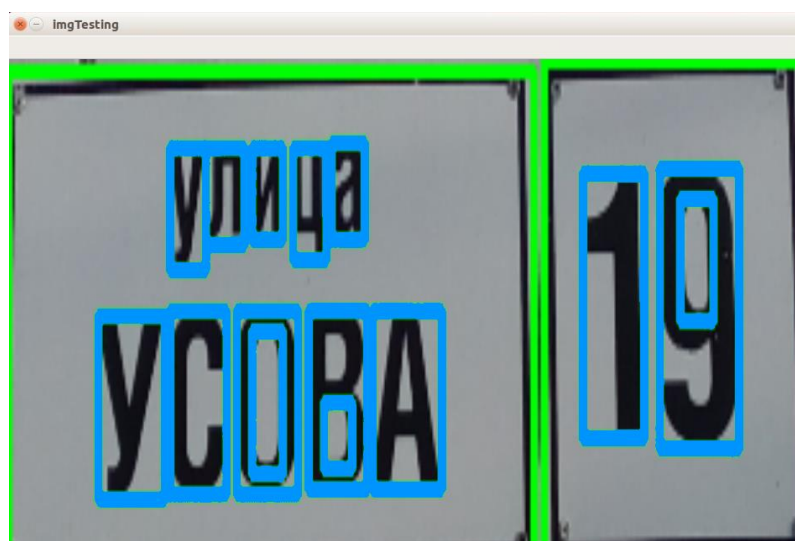


Рисунок 4.6 – Результат сегментации

Пример вывода нейронной сети для распознавания одного символа изображен на рисунке 4.7. Каждое значение представляет собой вероятность, из символов, имеющих максимальную вероятность, на основе максимальных значений вероятностей составляется выходная строка.

```
(image-classifier) user@...:~/PythonEnvironments/image-classifier$ python predict.py  
[0.02451666, 0.02839908, 0.61756054, 0.01664993, 0.00105205, 0.00291787, 0.01800255, 0.0285  
112212, 0.01970273, 0.00133887, 0.00784733, 0.00464257, 0.00082445, 0.00347422, 0.01382702,  
9, 0.00469606, 0.00679531, 0.00334315, 0.01868503, 0.01260803, 0.01303068, 0.00915168, 0.00  
1548089, 0.01327979, 0.02477706, 0.00308764, 0.02440661, 0.0005547, 0.01848016, 0.01242948,  
7, 0.00511223]
```

Рисунок 4.7 – Вывод нейронной сети

#### 4.5 Вывод по разделу 4

Реализованный алгоритм с применением сверточной нейронной сети показывает хорошую точность распознавания. Повышение точности распознавания возможно за счёт создания более объёмной обучающей выборки, содержащей большое разнообразие табличек и символов. Также использование более производительного аппаратного обеспечения позволит сократить время обучения, что увеличит количество проводимых экспериментов со структурой ИНС.



## **5 ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ**

### **5.1 Предпроектный анализ**

Разрабатываемое решение в рамках магистерской диссертации, заключается в разработке и реализации алгоритма для распознавания адресных табличек домов. Разработка алгоритма распознавания на основе ИНС позволит достичь высокой скорости идентификации адреса для различных сфер применения.

Одним из наиболее перспективных применений является разработка на основе описываемого в данной работе алгоритма программной системы для обработки изображений уличных панорам. Система позволит находить и распознавать адреса домов без участия человека. Распознанные адреса возможно привязать к карте с помощью GPS.

Потенциальным потребителем данной разработки являются организации, нуждающиеся в автоматизации распознавания адресных табличек домов. Сферы применения разработки: от нужд обычных пользователей до специализированных задач.

Целью данного раздела является анализ совокупности факторов, которые определяют коммерческую привлекательность разработки, ее перспективность и успешность.

Основными задачами является оценка перспективности разработки с помощью анализа конкурентных технических решений, FAST-анализа, SWOT-анализа, оценки готовности проекта коммерциализации, а также расчет стоимости и составление графика проведения работ.

#### **5.1.1 Анализ конкурентных технических решений**

Далее проведем анализ конкурентных разработок для того чтобы оценить возможности продукта выстоять против конкуренции со стороны других программных решений.

На данный момент, в сфере распознавания адресных табличек отсутствуют прямые конкуренты. В связи с этим, проводится сравнение с

программными продуктами, имеющими функции распознавания простого текста. Для анализа выбраны следующие решения:

Б1 – Разрабатываемое решение;

Б2 – Yandex Translator. Переводчик с модулем для детектирования и распознавания текста;

Б3 – ABBYY FineReader. Программа для оптического распознавания текста.

Результаты анализа приведены в таблице 5.1.

Таблица 5.1 – Оценочная карта сравнения конкурентности технических решений

Критерии оценки	Вес	Баллы			Конкурентоспособность		
		Б1	Б2	Б3	К1	К2	К3
1	2	3	4	5	6	7	8
Технические критерии оценки ресурсоэффективности							
1. Удобство в эксплуатации	0,1	9	9	9	0,9	0,9	0,9
2. Эргономичность	0,08	4	6	7	0,32	0,48	0,56
3. Надежность	0,06	5	5	5	0,3	0,3	0,3
4. Потребность в ресурсах памяти	0,09	7	8	8	0,63	0,72	0,72
5. Простота эксплуатации	0,08	8	8	8	0,64	0,64	0,64
6. Качество интеллектуального интерфейса	0,07	7	6	6	0,49	0,42	0,42
7. Точность распознавания	0,3	9	5	4	2,7	1,5	1,2
Экономические критерии оценки эффективности							
1. Цена	0,1	9	9	2	0,9	0,9	0,2
2. Послепродажное обслуживание	0,08	8	8	8	0,64	0,64	0,64
3. Уровень проникновения на рынок	0,04	8	6	5	0,32	0,24	0,2
<b>Итого:</b>	<b>1</b>	<b>74</b>	<b>70</b>	<b>62</b>	<b>7,84</b>	<b>6,74</b>	<b>5,78</b>

Анализ конкурентных технических решений определяется по формуле 5.1:

$$K = \sum B_i \cdot B_j, \quad (5.1)$$

где K – конкурентоспособность научной разработки или конкурента;

B<sub>i</sub> – вес показателя (в долях единицы);

B<sub>i</sub> – балл i-го показателя

Конкурентоспособность проекта 1 исполнения относительно проекта 2 исполнения:  $K_{1,2} = 7,84 / 6,74 = 1,16$ .

Конкурентоспособность проекта 1 исполнения относительно проекта 3 исполнения:  $K_{1,3} = 7,84 / 5,78 = 1,36$ .

По результатам проведённого анализа можно сделать вывод, что предложенное решение наиболее конкурентоспособно в сравнении с другими рассмотренными решениям. В первую очередь, конкурентоспособность обусловлена высокой точностью распознавания и уровнем проникновения на рынок.

## 5.2 FAST-анализ

В рамках магистерской диссертации в качестве объекта FAST-анализа выступает алгоритм распознавания адресных табличек домов.

Для описания главной, основных и вспомогательных функций, выполняемых объектом алгоритм был разделён на пять составных частей: чтение изображения, предобработка изображения, детектирование таблички, распознавание символа, визуализация результатов (таблица 5.2).

Таблица 5.2 – Классификация функций, выполняемых объектом исследования

Наименование процесса	Выполняемая функция	Ранг функции		
		Главная	Основная	Вспомогательная
Чтение изображения	Чтение данных в оперативную память			X
Предобработка изображения	Приведение изображения к нужному формату		X	
Детектирование таблички	Поиск таблички на изображении		X	
Распознавание символа	Определение к какому символу относится каждый сегмент	X		
Визуализация результатов	Отображение результатов распознавания			X

Определим значимость функций, выполняемых объектом по методу расставления приоритетов (таблица 5.3).

Таблица 5.3 – Матрица смежности функций

	Чтение изображения	Предобработка изображения	Детектирование таблички	Распознавание символа	Визуализация результатов
Чтение изображения	=	>	>	>	=
Предобработка изображения	<	=	>	>	<
Детектирование таблички	<	<	=	>	<
Распознавание символа	<	<	<	=	<
Визуализация результатов	=	>	>	>	=

Примечание: «<» – менее значимая; «=» – одинаковые функции по значимости; «>» – более значимая.

Преобразуем матрицу смежности в матрицу количественных соотношений функций (таблица 5.4).

Таблица 5.4 – Матрица количественных соотношений функций

	Чтение изображения	Предобработка изображения	Детектирование таблички	Распознавание символа	Визуализация результатов
Чтение изображения	1	1,5	1,5	1,5	1
Предобработка изображения	0,5	1	1,5	1,5	0,5
Детектирование таблички	0,5	0,5	1	1,5	0,5
Распознавание символа	0,5	0,5	0,5	1	0,5
Визуализация результатов	1	1,5	1,5	1,5	1
$\Sigma=25$	0,14	0,2	0,24	0,28	0,14

Оценим уровень затрат на выполнение каждой функции с использованием нормативного метода (таблица 5.5). Стоимость функций определим исходя из количества часов, затраченных на их выполнение, а также стоимости часа работы дипломника на предприятии, где выполнялась дипломная работа, так как фактических материалов при разработке программного кода не было затрачено.

Таблица 5.5 – Определение стоимости функций

Наименование процесса	Выполняемая функция	Трудоёмкость в (ч)	Стоимость часа работы (руб/час)	Себестоимость (руб)
Чтение изображения	Чтение данных в оперативную память	2	165	330
Предобработка изображения	Приведение изображения к нужному формату	20		3 300
Детектирование таблички	Поиск таблички на изображении	120		19 800
Распознавание символа	Определение к какому символу относится каждый сегмент	90		14 850
Визуализация результатов	Отображение результатов распознавания	8		1 320

Общая стоимость исследования: 39 600 руб.

На рисунке 5.1 представлена информация об объекте исследования, собранная в рамках предыдущих стадий, обобщённая в виде функционально-стоимостной диаграммы.

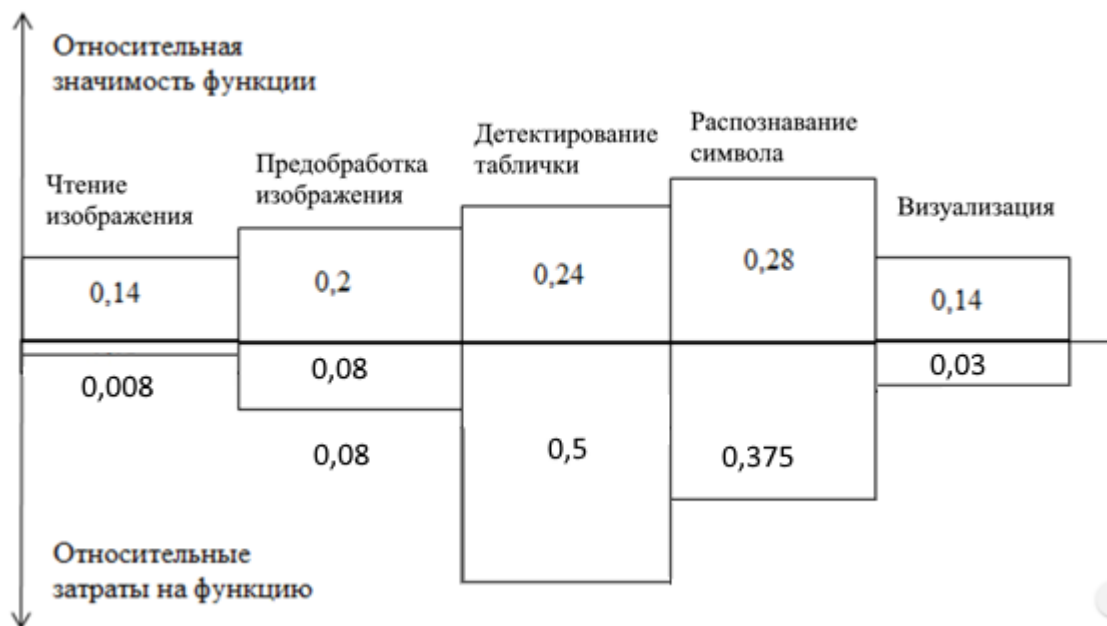


Рисунок 5.1 – Функционально-стоимостная диаграмма

Исходя из полученной функционально-стоимостной диаграммы, по функциям чтения изображения и визуализации результатов наблюдается

рассогласование. Необходимо оптимизировать данные функции. Достигнуть этого можно путём введения ограничений на набор методов чтения и визуализации (ввести некоторый базовый функционал), тем самым затратив меньшее число часов на разработку этих методов.

### **5.3 SWOT-анализ**

Результаты SWOT-анализа учитываются при разработке структуры работ, выполняемых в рамках разработки проекта. Учитывая отраслевую специфику объекта исследования магистерской диссертации необходимо провести комплексный анализ проекта.

В ходе проведения SWOT-анализа была составлена итоговая матрица (таблица 5.6).

Таблица 5.6 – Результаты SWOT-анализа

	<b>Сильные стороны</b> С1. Использование нейронных сетей С2. Большая точность распознавания С3. Открытость разработки	<b>Слабые стороны</b> Сл1. Падение точности распознавания для некоторых видов табличек Сл2. Необходимость наличия большой базы данных для обучения модели
<b>Возможности</b> В1. Рост вычислительной мощности процессоров В2. Тенденция к увеличению популярности использования алгоритмов нейронных сетей для решения задач по распознаванию	В1С1С3: Увеличение скорости обучения нейронной сети и качества распознавания В2С2: В связи с пристальным вниманием к алгоритмам на основе нейронных сетей, возможно появление новых оптимизаций и эвристик, которые в будущем можно будет интегрировать в имеющуюся разработку, значительно улучшив её конкурентоспособность	В1Сл1: Более производительное аппаратное обеспечение позволит качественнее обучать модели, что увеличит точность распознавания В2Сл2: Увеличение базы данных изображений, за счёт исследовательского сообщества
<b>Угрозы</b> У1. Недостаток производительности аппаратного обеспечения для обучения У2. Недостаток человеческих ресурсов по сравнению с конкурентными разработками	У1С1С2: Вариативность структуры нейронной сети и улучшение алгоритмов обучения позволит обойти аппаратные ограничения У2С3: Засчёт открытости разработки привлечение новых специалистов в области нейронных сетей, возможность найма сотрудников работать удаленно	У1У2Сл1Сл2: Повышение квалификации сотрудников для разработки более производительных алгоритмов У2Сл2: Автоматизация сбора изображений

Таким образом, можно сделать вывод, что проект необходимо развивать в направлении наибольшей универсальности и гибкости, с возможностью расширения с помощью дообучения модели на новых данных обучения. Необходимо отслеживать и применять новейшие технологии и улучшения в области нейросетевых алгоритмов, при этом повышать квалификацию сотрудников для успешного применения появляющихся технологий и разработки своих.

## 5.4 Оценка готовности проекта к коммерциализации

Для оценки готовности научной разработки к коммерциализации была составлена таблица 5.7, содержащая показатели о степени проработанности проекта с позиции коммерциализации и компетенциям разработчика научного проекта

Таблица 5.7 – Бланк оценки степени готовности научного проекта к коммерциализации

п/п	Наименование	Степень проработанности научного проекта	Уровень имеющихся знаний у разработчика
1.	Определен имеющийся научно-технический задел	4	4
2.	Определены перспективные направления коммерциализации научно-технического задела	4	4
3.	Определены отрасли и технологии (товары, услуги) для предложения на рынке	4	5
4.	Определена товарная форма научно-технического задела для представления на рынок	2	3
5.	Определены авторы и осуществлена охрана их прав	1	2
6.	Проведена оценка стоимости интеллектуальной собственности	2	1
7.	Проведены маркетинговые исследования рынков сбыта	2	2
8.	Разработан бизнес-план коммерциализации научной разработки	1	1
9.	Определены пути продвижения научной разработки на рынок	2	2
10.	Разработана стратегия (форма) реализации научной разработки	1	1
11.	Проработаны вопросы международного сотрудничества и выхода на зарубежный рынок	1	1
12.	Проработаны вопросы использования услуг инфраструктуры поддержки, получения льгот	1	1
13.	Проработаны вопросы финансирования коммерциализации научной разработки	1	1
14.	Имеется команда для коммерциализации научной разработки	1	1
15.	Проработан механизм реализации научного проекта	2	3
	<b>ИТОГО БАЛЛОВ</b>	<b>29</b>	<b>32</b>



Оценка готовности научного проекта к коммерциализации (или уровень имеющихся знаний у разработчика) определяется по формуле 5.2:

$$B_{\text{сум}} = \sum B_i \quad (5.2)$$

где  $B_{\text{сум}}$  – суммарное количество баллов по каждому направлению;

$B_i$  – балл по  $i$ -му показателю

После проведения анализа, можно сделать вывод, что перспективность проведения коммерциализации средняя. По вопросам маркетинговых исследований, финансирования коммерциализации, необходимо привлечение в команду проекта специалистов из данных областей.

Для успешной коммерциализации научно-технических разработок необходимо правильно определить метод коммерциализации. Данный метод определяет успешность продвижения продукта на рынке, коммерческий эффект которого может быть использован для дальнейшего продвижения и развития проекта. Организация совместного предприятия – является наиболее подходящим методом для разрабатываемого проекта.

Один человек не может справиться с разработкой проекта, исследованием актуальных технологий и разработок, продумыванием направления дальнейшего движения проекта, поиском источников финансирования и пр. К тому же, с расширением проекта в будущем, будет необходимо задействовать большее число трудовых ресурсов. Создание совместного предприятия позволит выделить отделы по разработке программного обеспечения, отдел по перспективному развитию, отдел маркетинга и другие, что позволит структурировать работу и облегчить управление и поддержку проекта.

### **5.5 Инициация разработки**

В рамках инициации разработки формулируются цели и ожидаемые результаты работы. Также определяются заинтересованные стороны разработки и возможные ограничения. Заинтересованные в данной разработке стороны представлены в таблице 5.8.

Таблица 5.8 – Заинтересованные стороны разработки

<b>Заинтересованные стороны</b>	<b>Ожидания заинтересованных сторон</b>
НИ ТПУ, школа ИШИТР	Получение нейросетевого алгоритма для распознавания адресных табличек домов на изображении с целью его дальнейшего внедрения и представления на конференциях
Студент	Защита магистерской диссертации

Цели и результат проекта отображены в таблице 5.9

Таблица 5.9 – Цели и результат проекта

<b>Цели разработки:</b>	Разработка нейросетевого алгоритма распознавания адресных табличек домов
<b>Ожидаемые результаты разработки:</b>	Точность, выраженная в процентах, и определяемая числом правильно распознанных табличек
<b>Критерии приемки результата разработки</b>	Точность распознавания не ниже 90%
<b>Требования к результату разработки:</b>	Оформление результатов исследования в соответствии с требованиями нормативнотехнической документации
	Точность распознавания у разработанного алгоритма должна быть выше 90%
	Стоимость проекта должна быть сопоставима по сравнению с аналогами, а в лучшем случае быть меньшей

В таблице 5.10 представлена рабочая группа разработки, определена роль и основные функции каждого участника в разработке.

Таблица 5.10 – Рабочая группа разработки

<b>№ п/п</b>	<b>ФИО, основное место работы, должность</b>	<b>Роль в разработке</b>	<b>Функции</b>	<b>Трудо-затраты, рабочие дни.</b>
1	Спицын В.Г., НИ ТПУ, профессор	Научный руководитель	Утверждение основных разделов, выдача заданий к исполнению, координирование деятельности исполнителя	10
2	Белков С.Г., НИ ТПУ, студент	Исполнитель	Исполнение поставленных задач	107
<b>ИТОГО:</b>				<b>117</b>

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

## **5.6 Организация и планирование работ**

### **5.6.1 Иерархическая структура работ**

Иерархическая структура работ для данной разработки представляет собой детализацию укрупненной структуры работ (рисунок 5.2).

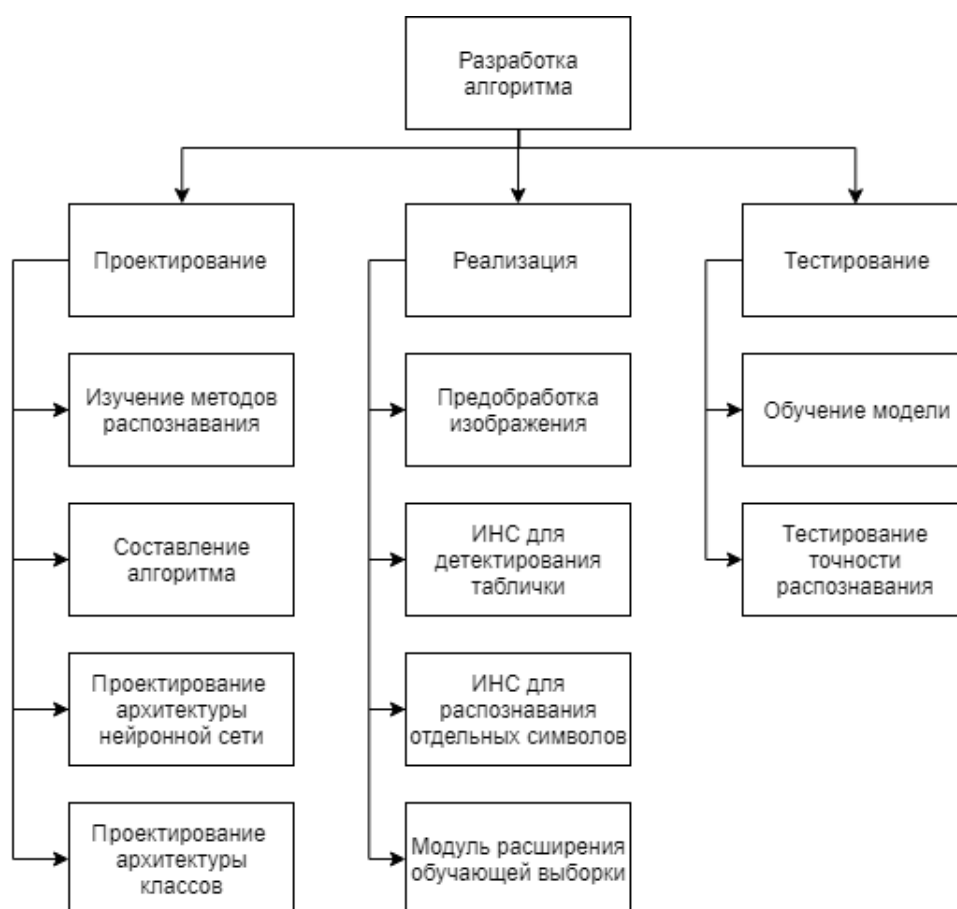


Рисунок 5.2 – Иерархическая структура работ

Задачи по созданию данной разработки разделены на три основных блока: проектирование, реализация и тестирование. Основная часть работ фокусируется на реализации нейронных сетей и обучении оптимальной модели данных.

### **5.6.2 План разработки**

Для организации оптимального процесса реализации проекта необходимо спланировать занятость каждого участника разработки, а также

сроки проведения отдельных работ. Этот этап служит для составления полного перечня проводимых работ, определения исполнителей для данных работ, а также формируется время продолжительности работы.

Наиболее удобным, простым и наглядным способом для этих целей является использование линейного графика. Для его построения составим перечень работ и соответствие работ своим исполнителям, продолжительность выполнения этих работ и сведем их в таблицу 5.11.

Таблица 5.11 – Перечень работ и продолжительность их выполнения

Этапы работы	Исполнители	Загрузка исполнителей
Постановка целей и задач, получение исходных данных	НР, И	НР – 60% И – 40%
Подбор и изучение материалов по тематике	НР, И	НР – 30% И – 70%
Разработка календарного плана	НР, И	НР – 40% И – 60%
Обсуждение литературы	НР, И	НР – 20% И – 80%
Анализ алгоритмов для решения поставленной задачи	НР, И	НР – 20% И – 80%
Планирование архитектуры программного обеспечения	НР, И	НР – 10% И – 90
Разработка программного обеспечения	И	И – 100%
Тестирование и отладка программного обеспечения	И	И – 100%
Оформление расчетно-пояснительной записки	И	И – 100%
Подведение итогов	НР, И	НР – 70% И – 30%

НР – научный руководитель

И – исполнитель

### 5.6.3 Продолжительность этапов работ

Для определения ожидаемого значения продолжительности работ тоже применяется вероятностный метод – метод двух оценок  $t_{\min}$  и  $t_{\max}$  (формула 5.3).

$$t_{\text{ож}} = \frac{(3t_{\min} + 2t_{\max})}{5} \quad (5.3)$$

где  $t_{\min}$  – минимальная трудоемкость работ, чел/дн.,

$t_{\max}$  – максимальная трудоемкость работ, чел/дн.

Для выполнения перечисленных в таблице 4.11 работ требуются специалисты:

- инженер – в его роли действует исполнитель НИР (ВКР);
- научный руководитель.

Для построения линейного графика необходимо перевести длительность этапов в рабочих днях в календарные дни. Расчет продолжительности этапа в календарных днях ведется по формуле 5.4:

$$T_{\text{КД}} = T_{\text{РД}} \cdot T_{\text{К}} \quad (5.4)$$

где  $T_{\text{РД}}$  – продолжительность выполнения этапа в рабочих днях,  
 $T_{\text{КД}}$  – продолжительность выполнения этапа в календарных днях,  
 $T_{\text{К}}$  – коэффициент календарности.

Коэффициент календарности рассчитывается по формуле:

$$T_{\text{КД}} = \frac{T_{\text{кал}}}{T_{\text{кал}} - T_{\text{вых}} - T_{\text{пр}}} = \frac{365}{365 - 66} = 1,22$$

где  $T_{\text{КАЛ}}$  – календарные дни,  
 $T_{\text{ВЫХ}}$  – выходные дни,  
 $T_{\text{ПР}}$  – праздничные дни.

Таблица 5.12 – Трудозатраты на выполнение проекта

Этап	Исполнители	Продолжительность работ,			Длительность работ, чел/дн.			
		дни			Т <sub>рд</sub>		Т <sub>кд</sub>	
		t <sub>min</sub>	t <sub>max</sub>	t <sub>ож</sub>	НР	И	НР	И
Постановка целей и задач, получение данных	НР, И	1	3	1,8	1,0	0,72	1	1
Подбор и изучение материалов по тематике	НР, И	10	15	12	3,6	8,4	4	10
Разработка календарного плана	НР, И	2	4	2,8	1,12	1,68	1	2
Обсуждение литературы	НР, И	2	5	3,2	0,64	2,56	1	3
Анализ алгоритмов для решения поставленных задач	НР, И	10	17	12,8	2,56	10,24	3	13
Планирование архитектуры программного обеспечения	НР, И	15	22	17,8	1,78	16,02	2	20
Разработка программного обеспечения	И	30	50	38	0	38	0	46
Тестирование и отладка программного обеспечения	И	20	25	22	0	22	0	27
Оформление расчетно-пояснительной записки	И	5	7	5,8	0	5,8	0	7
Подведение итогов	НР, И	2	4	3	2,1	0,9	2	1
<b>Итого:</b>				<b>117</b>	<b>10</b>	<b>107</b>	<b>14</b>	<b>130</b>

#### 5.6.4 Расчет нарастания технической готовности работ

Величина нарастания технической готовности работы показывает процент выполнения работы на каждом этапе (формула 5.5):

$$H_i = \frac{t_{Hi}}{t_0} \cdot 100\% \quad (5.5)$$

где  $t_{Hi}$  – нарастающая трудоемкость с момента начала работы  $i$ -го этапа,  
 $t_0$  – общая трудоемкость.

Общая трудоемкость вычисляется по формуле 5.6:

$$t_0 = \sum_{i=1}^n t_{OЖi} \quad (5.6)$$

где  $t_{OЖi}$  – ожидаемая продолжительность  $i$ -го этапа.

Удельный вес каждого этапа  $Y_i$  определяется по формуле 5.7:






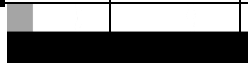





$$Y_i = \frac{t_{OЖi}}{t_0} \cdot 100\% \quad (5.7)$$



Результаты вычислений  $H_i$  и  $Y_i$  отражены в таблице 13.

Таблица 5.13 – Нарастание технической готовности и удельный вес этапов

Этап	$H_i, \%$	$Y_i, \%$
Постановка целей и задач, получение данных	1,51	1,51
Подбор и изучение материалов по тематике	11,59	10,08
Разработка календарного плана	13,949	2,35
Обсуждение литературы	16,63	2,689
Анализ алгоритмов для решения поставленных задач	27,39	10,756
Планирование архитектуры программного обеспечения	42,35	14,958
Разработка программного обеспечения	74,28	31,93
Тестирование и отладка программного обеспечения	92,77	18,48
Оформление расчетно-пояснительной записки	97,64	4,87
Подведение итогов	100	2,35

Таблица 5.14 – Линейный график работ

Этап	Т <sub>кд</sub>		Февраль			Март			Апрель			Май		
	НР	И	10	20	30	10	20	30	10	20	30	10	20	30
1	1	1												
2	4	10												
3	1	2												
4	1	3												
5	3	13												
6	2	20												
7	0	46												
8	0	27												
9	0	7												
10	2	1												

НР  И 



## 5.7 Расчет сметы затрат

В состав затрат на разработку ПО и проведение эксперимента включается стоимость всех расходов, необходимых для реализации комплекса работ, составляющих содержание данной разработки. Расчет сметной стоимости на выполнение данной разработки производится по следующим статьям затрат:

- материалы и покупные изделия;
- заработная плата;
- отчисления в социальные фонды;
- амортизационные отчисления;
- накладные расходы.

### 5.7.1 Расчет заработной платы

Данная статья расходов включает заработную плату научного руководителя и инженера, а также премии, входящие в фонд заработной платы. Расчет заработной платы выполняется на основе трудоемкости выполнения каждого этапа и величины месячного оклада исполнителя. Среднедневная заработная плата рассчитывается по формуле 5.8:

$$\text{Дневная з/плата} = \frac{\text{Месячный оклад}}{24,9 \text{ дня}} \quad (5.8)$$

Расчеты затрат на заработную плату приведены в таблице 5.15. При расчете учитывалось, что в году 299 рабочих дня и, следовательно, в месяце 24,9 рабочих дня. Также был принят во внимание коэффициент, учитывающий коэффициент по премиям  $K_{\text{ПР}} = 1,1$  и районный коэффициент  $K_{\text{РК}} = 1,3$  ( $K = K_{\text{ПР}} * K_{\text{РК}} = 1,1 * 1,3 = 1,43$ ).

Таблица 5.15 – Затраты на заработную плату

Исполнитель	Оклад руб./мес.	Среднедневная ставка руб./день	Затраты времени, дни	Коэффициент	Фонд з/платы, руб.
НР	33 664	1346,56	10	1,43	13 304
И	21 760	870,4	107	1,43	132 333,53
<b>Итого:</b>					<b>145 637,54</b>

### 5.7.2 Расчет отчислений от заработной платы

Затраты по этой статье составляют отчисления в социальные фонды. Отчисления по заработной плате определяются по следующей формуле 5.9:

$$C_{\text{соц}} = K_{\text{соц}} * C_{\text{зп}}, \quad (5.9)$$

где  $K_{\text{соц}}$  – коэффициент, учитывающий размер отчислений из заработной платы. Данный коэффициент составляет 30,2% от затрат на заработную плату и включает в себя:

- отчисления в пенсионный фонд;
- на социальное страхование;
- на медицинское страхование.

Отчисления из заработной платы составили:

$$C_{\text{соц}} = 0,302 * 145\,637,54 = 43\,982,54 \text{ руб.}$$

### 5.7.3 Расчет амортизации

В статье «Амортизационные отчисления» от используемого оборудования рассчитывается амортизация за время выполнения работы для оборудования, которое имеется в наличии. Амортизационные отчисления рассчитываются на время использования ЭВМ по формуле 5.10:

$$C_{\text{ам}} = \frac{H_A \cdot C_{\text{об}}}{F_d} \cdot t_{\text{вт}} \cdot n \quad (5.10)$$

где  $H_A$  – годовая норма амортизации,  $H_A = 33,3\%$ ,

$C_{\text{об}}$  – цена оборудования,  $C_{\text{об}} = 50000$  руб.,

$F_d$  – действительный годовой фонд рабочего времени,  $F_d = 2\,392$  часов,

$t_{\text{вт}}$  – время работы вычислительной техники при создании программного продукта,  $t_{\text{вт}} = 930$  часа,  $n$  – число задействованных ПЭВМ,  $n = 1$ .

Затраты на амортизационные отчисления составили:

$$C_{\text{ам}} = \frac{0,333 * 50000}{2392} * 930 * 1 = 6\,473,45$$

#### 5.7.4 Расчет накладных расходов

В статье «Накладные расходы» отражены расходы на разработку проекта, которые не учтены в предыдущих статьях. Накладные расходы составляют 16 % от единовременных затрат на выполнение технического продукта и проводятся по формуле:

$$C_{\text{нак}} = (C_{\text{зп}} + C_{\text{соц}} + C_{\text{ам}}) * 0,16$$

$$C_{\text{нак}} = (145\,637,54 + 43\,982,54 + 6\,473,45) * 0,16 = 31\,375$$

#### 5.7.5 Расчет общей сметы

Проведя расчет сметы затрат на разработку, можно определить общую стоимость разработки проекта (таблица 5.16).

Таблица 5.16 – Смета затрат

Статья затрат	Сумма, руб.
Заработная плата	145 637,54
Отчисления в социальные фонды	43 982,54
Амортизационные отчисления	6 473,45
Накладные расходы	31 375
Итого	227 468,53

Таким образом, расходы на разработку составили  $C = 227\,468,53$  руб.

### 5.8 Оценка научно-технического уровня НИР

Научно-технический уровень характеризует, в какой мере выполнены работы и обеспечивается научно-технический прогресс в данной области. Для оценки научной ценности, технической значимости и эффективности, планируемых и выполняемых НИР, используется метод бальных оценок. Бальная оценка заключается в том, что каждому фактору по принятой шкале присваивается определенное количество баллов. Обобщенную оценку проводят по сумме баллов по всем показателям или рассчитывают по формуле. На этой основе делается вывод о целесообразности НИР.

Сущность метода заключается в том, что на основе оценок признаков работы определяется коэффициент ее научно-технического уровня по формуле 5.11:

$$K_{\text{НТУ}} = \sum_{i=1}^3 R_i \cdot n_i \quad (5.11)$$

где  $K_{\text{НТУ}}$  – коэффициент научно-технического уровня,

$R_i$  – весовой коэффициент  $i$ -го признака научно-технического эффекта,

$n_i$  – количественная оценка  $i$ -го признака научно-технического эффекта, в баллах.

Таблица 5.17 – Весовые коэффициенты признаков НТУ

Признак НТУ	Характеристика признака	$R_i$
Уровень новизны	Систематизируются и обобщаются сведения, определяются пути дальнейших исследований	0.4
Теоретический уровень	Разработка способа (алгоритм, программа мероприятий, устройство, вещество и т.п.)	0.1
Возможность реализации	Время реализации в течение первых лет	0.5

Таблица 5.18 – Баллы для оценки уровня новизны

Уровень новизны	Характеристика уровня новизны	Баллы
Принципиально новая	Новое направление в науке и технике, новые факты и закономерности, новая теория, вещество, способ	8-10
Новая	По-новому объясняются те же факты, закономерности, новые понятия дополняют ранее полученные результаты	5-7
Относительно новая	Систематизируются, обобщаются имеющиеся сведения, новые связи между известными факторами	2-4
Не обладает новизной	Результат, который ранее был известен	0

Таблица 5.19 – Баллы значимости теоретических уровней

Теоретический уровень полученных результатов	Баллы
Установка закона, разработка новой теории	10
Глубокая разработка проблемы, многоспектральный анализ, взаимодействия между факторами с наличием объяснений	8
Разработка способа (алгоритм, программа и т. д.)	6
Элементарный анализ связей между фактами (наличие гипотезы, объяснения версии, практических рекомендаций)	2
Описание отдельных элементарных факторов, изложение наблюдений, опыта, результатов измерений	0,5

Таблица 5.20 – Возможность реализации по времени

Время реализации	Баллы
В течение первых лет	10
От 5 до 10 лет	4
Свыше 10 лет	2

Результаты оценок признаков научно-технического уровня приведены в таблице 5.21.

Таблица 5.21 – Оценки научно-технического уровня НИР

Признак НТУ	Выбранный балл	Обоснование выбранного балла
Уровень новизны	4	Облегчит обработку адресных табличек домов
Теоретический уровень	6	Разработана и реализован алгоритм, решающий поставленные в рамках задания задачи
Возможность реализации	7	Полученный продукт находится на стадии тестирования

Исходя из оценки признаков, показатель научно-технического уровня для данного проекта составил:

$$K_{\text{нту}} = 0,4 \cdot 4 + 0,1 \cdot 6 + 0,5 \cdot 7 = 5,7$$

В таблице 5.22 приведена градация показателя уровня научно-технического эффекта.

Таблица 5.22 – Оценка уровня научно-технического эффекта

Уровень НТЭ	Показатель НТЭ
Низкий	1-4
Средний 4-7	4-7
Высокий	8-10

Таким образом, исходя из таблицы 5.22, данный проект имеет средний уровень научно-технического эффекта.

### 5.9 Вывод по разделу 5

Потенциальными потребителями разработки являются организации, нуждающиеся в автоматизации распознавания адресных табличек домов.

Анализ конкурентных технических решений показал, что предложенный метод имеет высокую конкурентоспособность за счет высокой точности распознавания, дешевизны и уровня проникновения на рынок.

Сильные и слабые стороны решения, его возможности и угрозы, а также корреляция этих показателей были определены в ходе SWOT-анализа. Полученные показатели позволили определить направление дальнейшего развития предложенного алгоритма для достижения наибольшей конкурентоспособности. При этом было установлена необходимость применения новейших технологий в сфере нейронных сетей и повышении квалификации персонала для поддержания гибкости, высокой точности распознавания и уменьшения времени обучения модели.

Разграничение этапов работ позволило структурировать план работы над исследованием, определить ответственных за его этапы. На основе созданного перечня этапов и работ был создан календарный план-график. Согласно которому общая длительность работ составляет 117 дней. Бюджет затрат на исследование составляет 227 468 руб.

## **6 СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ**

### **Введение**

В рамках магистерской диссертации разработан и реализован алгоритм, предназначенный для сегментации и распознавания адресных табличек домов. Разработка велась исключительно при помощи компьютера. Сферы применения разработки: от нужд обычных пользователей до специализированных задач. Одним из наиболее перспективных применений является разработка на основе описываемого в данной работе алгоритма программной системы для обработки изображений уличных панорам. Независимо от конкретного применения, взаимодействие пользователя с разработанной программой в любом случае производится с помощью программных и аппаратных средств ПЭВМ, а также с помощью периферийных устройств, подключенных к ПЭВМ.

Данный раздел посвящен анализу вредных и опасных факторов производственной среды для операторов ПЭВМ, а также разработке программ по минимизации воздействия вредоносного и опасного влияния выявленных факторов, снижению вредных воздействий на окружающую среду и защите в чрезвычайных ситуациях.

### **6.1 Правовые и организационные вопросы обеспечения безопасности**

#### **6.2 Правовые нормы трудового законодательства**

Регулирование отношений между работником и работодателем, касающихся оплаты труда, трудового распорядка, особенности регулирования труда женщин, детей, людей с ограниченными способностями и проч., осуществляется законодательством РФ, а именно трудовым кодексом РФ [21]. Продолжительность рабочего дня не должна быть меньше указанного времени в договоре, но не больше 40 часов в неделю. Для работников до 16 лет – не более 24 часов в неделю, от 16 до 18 лет и инвалидов I и II группы – не более 35 часов.

Возможно установление неполного рабочего дня для беременной женщины; одного из родителей (опекуна, попечителя), имеющего ребенка в

возрасте до четырнадцати лет (ребенка-инвалида в возрасте до восемнадцати лет). Оплата труда при этом производится пропорционально отработанному времени, без ограничений оплачиваемого отпуска, исчисления трудового стажа и других прав.

В течение рабочего дня работнику должен быть предоставлен перерыв для отдыха и питания продолжительностью не более двух часов и не менее 30 минут, который в рабочее время не включается. Всем работникам предоставляются выходные дни, работа в выходные дни осуществляется только с письменного согласия работника

Выполняемая работа по виду трудовой деятельности относится к группе Б – работа по вводу информации и по категории тяжести к категории В – суммарное время непосредственной работы с ПЭВМ за рабочую смену не более 6 часов за смену.

#### **6.2.1 Организационные мероприятия при компоновке рабочей зоны**

Требования к организации рабочих мест пользователей:

– Рабочее место должно быть организовано с учетом эргономических требований согласно ГОСТ 12.2.032-78 «ССБТ. Рабочее место при выполнении работ сидя. Общие эргономические требования» [22] и ГОСТ 12.2.061-81 «ССБТ. Оборудование производственное. Общие требования безопасности к рабочим местам» [23].

Конструкция рабочей мебели (рабочий стол, кресло, подставка для ног) должна обеспечивать возможность индивидуальной регулировки соответственно росту пользователя и создавать удобную позу для работы. Вокруг ПК должно быть обеспечено свободное пространство не менее 60-120см;

На уровне экрана должен быть установлен оригинал-держатель. На рисунке 6.1. схематично представлены требования к рабочему месту.





Рисунок 6.1 – Организация рабочего места

Работа программиста связана с постоянной работой за компьютером, следовательно, могут возникать проблемы, связанные со зрением. Также неправильная рабочая поза может оказывать негативное влияние на здоровье. Таким образом, неправильная организация рабочего места может послужить причиной нарушения здоровья и появлением психологических расстройств.

Согласно СанПиН 2.2.2/2.4.1340-03 «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы» [24]:

- яркость дисплея не должна быть слишком низкой или слишком высокой;
- размеры монитора и символов на дисплее должны быть оптимальными;
- цветовые параметры должны быть отрегулированы таким образом, чтобы не возникало утомления глаз и головной боли.
- опоры для рук не должны мешать работе на клавиатуре;
- верхний край монитора должен находиться на одном уровне с глазом, нижний – примерно на  $20^\circ$  ниже уровня глаза;
- дисплей должен находиться на расстоянии 45-60 см от глаз;
- локтевой сустав при работе с клавиатурой нужно держать под углом  $90^\circ$ ;

– каждые 10 минут нужно отводить взгляд от дисплея примерно на 5- 10 секунд;

– монитор должен иметь антибликовое покрытие; – работа за компьютером не должна длиться более 6 часов, при этом необходимо каждые 2 часа делать перерывы по 15-20 минут;

– высота стола и рабочего кресла должны быть комфортными.

### **6.3 Производственная безопасность**

Для обеспечения производственной безопасности необходимо проанализировать воздействия на человека вредных и опасных производственных факторов, которые могут возникать при разработке или эксплуатации проекта. Производственный фактор считается вредным, если воздействие этого фактора на работника может привести к его заболеванию. Производственный фактор считается опасным, если его воздействие на работника может привести к его травме. Возможные опасные и вредные факторы согласно ГОСТ 12.0.003-2015 [25] приведены в таблице 6.1.

Таблица 6.1 – Возможные опасные и вредные факторы

<b>Факторы (ГОСТ 12.0.003-2015)</b>	<b>Этапы работ</b>			<b>Нормативные документы</b>
	<b>Разработка</b>	<b>Изготовление</b>	<b>Эксплуатация</b>	
1. Отклонение показателей микроклимата	+	+	+	СанПиН 2.2.4.548-96 [27]
2. Превышение уровня шума	+	+	+	ГОСТ 12.1.003-2014 [28] СанПиН 2.2.4.3359-16 [29]
3. Недостаточная освещенность рабочей зоны	+	+	+	СанПиН 2.2.1/2.1.1.1278-03 [30]
4. Повышенный уровень электромагнитных полей	+	+	+	ГОСТ 12.1.006–84 [31] СанПиН 2.2.4.3359-16 [29]
5. Повышенное значение напряжения в электрической цепи, замыкание которой может произойти через тело человека.	+	+	+	СанПиН 2.2.2/2.4.1340–03 [32]

Далее более подробно рассмотрены опасные и вредные факторы и мероприятия по снижению их воздействия.

### **6.3.1 Отклонение показателей микроклимата**

Микроклимат производственных помещений – это климат внутренней среды этих помещений, который определяется действующими на организм человека сочетаниями температуры, влажности и скорости движения воздуха, а также интенсивности теплового излучения от нагретых поверхностей.

Оптимальные значения этих характеристик зависят от сезона (холодный, тёплый), а также от категории физической тяжести работы. Для инженера-программиста она является лёгкой (Ia), так как работа проводится сидя, без систематических физических нагрузок.

Согласно требованиям СанПиН 2.2.4.548-96. «Гигиенические требования к микроклимату производственных помещений», оптимальные и допустимые параметры микроклимата в офисах приведены в таблицах 6.2 и 6.3.

Таблица 6.2 – Оптимальные значения характеристик микроклимата

<b>Период года</b>	<b>Кат. раб.</b>	<b>Температура воздуха, °С</b>	<b>Относительная влажность воздуха, %</b>	<b>Скорость движения воздуха, м/с</b>
Холодный	Ia	22-24	40-60	0,1
Теплый	Ia	23-25	40-60	0,1

Таблица 6.3 – Допустимые величины показателей микроклимата

<b>Период года</b>	<b>Кат. раб.</b>	<b>Температура воздуха, °С</b>	<b>Относительная влажность воздуха, %</b>	<b>Скорость движения воздуха, м/с</b>
Холодный	Ia	19-24	15-75	0,1-0,2
Теплый	Ia	20-28	15-75	0,1-0,3

Если температура воздуха отличается от нормальной, то время пребывания в таком помещении должно быть ограничено в зависимости от категории тяжести работ.

Для создания благоприятных условий труда и повышения производительности, необходимо поддерживать оптимальные параметры микроклимата производственных помещений. Для этого предусмотрены следующие средства: центральное отопление, вентиляция (искусственная и естественная), искусственное кондиционирование. Все из вышеперечисленных средств используются на рабочем месте.

Можно отметить, что фактические значения температуры, влажности и скорости движения воздуха соответствуют нормам, установленным в СанПиН 2.2.4.548-96. Для соответствия этим нормам производится проветривание помещения каждый час и ежедневная влажная уборка.

### **6.3.2 Превышение уровня шума**

Шум определяется как звук, оцениваемый негативно и наносящий вред здоровью. Основными источниками шума в помещении являются:

- система охлаждения центральных процессоров;
- жесткие диски.

При выполнении основной работы на ПЭВМ уровень шума на рабочем месте не должен превышать 50 дБА. Допустимые уровни звукового давления в помещениях для персонала, осуществляющего эксплуатацию ЭВМ при разных значениях частот согласно СН 2.2.4/2.1.8.562–96 «Шум на рабочих местах, в помещениях жилых, общественных зданий и на территории жилой, застройки», приведены в таблице 6.4.

Таблица 6.4 – Допустимые уровни звука на рабочем месте

Вид трудовой деятельности, рабочее место	Уровни звукового давления, дБ, в октавных полосах со среднегеометрическими частотами, Гц									Уровни звука и эквивалентного звука, дБА
	31,5	63	125	250	500	1000	2000	4000	8000	
Научная деятельность, проектирование, программирование,	86	71	61	54	49	45	42	40	38	50

Для снижения уровня шума, производимого персональными компьютерами, регулярно проводилось техническое обслуживание:

- чистка от пыли;
- замена смазывающих веществ;
- применение звукопоглощающих материалов.

### **6.3.3 Недостаточность освещенности рабочей зоны**

Недостаточная освещенность рабочей зоны является вредным производственным фактором, возникающим при работе с ПЭВМ, уровни 71 которого регламентируются СП 52.13330.2016. Причиной являются недостаточность естественного освещения, искусственного освещения, пониженная контрастность. Недостаточность освещения снижает производительность труда, увеличивает утомляемость, количество допускаемых ошибок, ухудшению зрения.

Установлено, что свет, помимо обеспечения зрительного восприятия, воздействует на нервную оптико-вегетативную систему, систему формирования иммунной защиты, рост и развитие организма и влияет на многие основные процессы жизнедеятельности, регулируя обмен веществ и устойчивость к воздействию неблагоприятных факторов окружающей среды. Освещение должно включать в себя как естественное, так и искусственное. Для источников искусственного освещения применяют люминесцентные лампы типа ЛБ.

Минимальный размер объекта различия входит в диапазон 0,5 до 1,0, следовательно, работа относится к разряду IV. Подразряд Г, т.к. контраст объектов различия с фоном большой, сам фон светлый. В соответствии с СП 52.13330.2016 уровень искусственного освещения должен быть не менее 300 лк. Пульсация при работе с компьютером не должна превышать 5%, рассчитывается по формуле:

$$K_{\Pi} = \frac{E_{\max} - E_{\min}}{2E_{\text{ср}}} 100 \quad (6.1)$$

где  $E_{\max}$  и  $E_{\min}$  – максимальное и минимальное значения освещенности соответственно за период ее колебания, лк;

$E_{\text{ср}}$  – среднее значение освещенности за этот же период, лк.

Увлечение коэффициента пульсации освещенности снижает зрительную работоспособность, повышает утомляемость, воздействует на нервные элементы коры головного мозга и фоторецепторные элементы сетчатки глаз. Для снижения пульсации необходимо использовать светильники, в которых лампы работают от переменного тока частотой 300 Гц и выше, так как при такой частоте пульсация не оказывает влияния на общую и зрительную работоспособность.

#### **6.3.4 Повышенный уровень электромагнитных полей**

Воздействие электромагнитного излучения на человека зависит от напряженностей электрического и магнитного полей, потока энергии, частоты колебаний, размера облучаемого тела.

Работа проводилась на современном компьютере, где значения электромагнитного излучения малы и отвечают требованиям СанПиН 2.2.1/2.1.1.1278-12. «Электромагнитные поля в производственных условиях», которые приведены в таблице 6.5.

Таблица 6.5 – 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 5 кГц-400 кГц	25 нТл
Поверхностный видеомонитора	электростатический потенциал экрана	500В

Основной способ снижения вредного воздействия – это увеличение расстояния от источника (не менее 50 см от пользователя). При работе за компьютером специальные экраны и другие средства индивидуальной защиты применены не были.

### **6.3.5 Повышенное значение напряжения в электрической цепи, замыкание которой может произойти через тело человека.**

Поражение электрическим током является опасным производственным фактором, поскольку оператор ПЭВМ имеет дело с электрооборудованием. Нормы электробезопасности на рабочем месте регламентируются СанПиН 2.2.2/2.4.1340-03, вопросы требований к защите от поражения электрическим током освещены в ГОСТ 12.1.019-2017 ССБТ.

Электробезопасность – система организационных и технических мероприятий и средств, обеспечивающих защиту людей от вредного и опасного воздействия электрического тока, электрической дуги, электромагнитного поля и статического электричества.

Помещение, где расположено рабочее место оператора ПЭВМ, относится к помещениям без повышенной опасности ввиду отсутствия следующих факторов: сырость, токопроводящая пыль, токопроводящие полы, высокая температура, возможность одновременного прикосновения человека к имеющим соединение с землей металлоконструкциям зданий, технологическим аппаратам, механизмам и металлическим корпусам электрооборудования.

Для оператора ПЭВМ при работе с электрическим оборудованием обязательны следующие меры предосторожности: перед началом работы нужно убедиться, что выключатели и розетка закреплены и не имеют оголённых токоведущих частей, при обнаружении неисправности оборудования и приборов необходимо, вызвать мастера ответственного за оборудование.

### **6.4 Экологическая безопасность**

В данном разделе рассматривается воздействие на окружающую среду деятельности по разработке проекта, а также самого продукта в результате его реализации на производстве.

Разработка и работа за ПЭВМ не являются экологически опасными работами, потому объект, на котором производилась разработка продукта и объекты, на которых он будет использован операторами ПЭВМ относятся к предприятиям пятого класса, размер селитебной зоны для которых равен 50 м.

Разработанный проект магистерской диссертации, не наносит вреда окружающей среде ни на стадиях его разработки, ни на стадиях эксплуатации. Однако, средства, необходимые для его разработки и эксплуатации могут наносить вред окружающей среде. Например, аккумуляторы. Для их утилизации необходимо обращаться в специальные организации, занимающиеся приемом, утилизацией и переработкой.

В нормативном документе СанПиН 2.2.1/2.1.1.1278-12, даются следующие общие рекомендации по снижению опасности для окружающей среды, исходящей от компьютерной техники:

- применять оборудование, соответствующее санитарным нормам и стандартам экологической безопасности;
- применять расходные материалы с высоким коэффициентом использования и возможностью их полной или частичной регенерации;
- отходы в виде компьютерного лома утилизировать;
- использовать экономные режимы работы оборудования.

## **6.5 Безопасность в чрезвычайных ситуациях**

Наиболее вероятной чрезвычайной ситуацией является возникновение пожара, так как на рабочем месте располагается большое количество ЭВМ. В соответствии с нормами пожарной безопасности, помещения с ЭВМ относятся к категории В (пожароопасные).

Основные причины возникновения возгораний: нарушение правил эксплуатации электрического оборудования, эксплуатация его в неисправном состоянии, перегрузка электрических сетей, применение неисправных осветительных приборов, электропроводки и устройств.

Для предупреждения возгораний в помещении необходимо соблюдать: установленный режим эксплуатации электрических устройств, противопожарные нормы и правила при установке оборудования, проводить технические осмотры и планово-предупредительные ремонты оборудования и



технических средств противопожарной защиты и пожаротушения (огнетушители) согласно утвержденному графику.

В помещении должен быть установлен углекислотный огнетушитель типа ОУ-5 для тушения пожаров. В случае угрозы возникновения ЧС необходимо отключить электропитание, вызвать по телефону пожарную команду, эвакуировать людей из помещения согласно плану эвакуации (рисунок 6.2). При наличии небольшого очага пламени можно воспользоваться подручными средствами с целью прекращения доступа воздуха к объекту возгорания

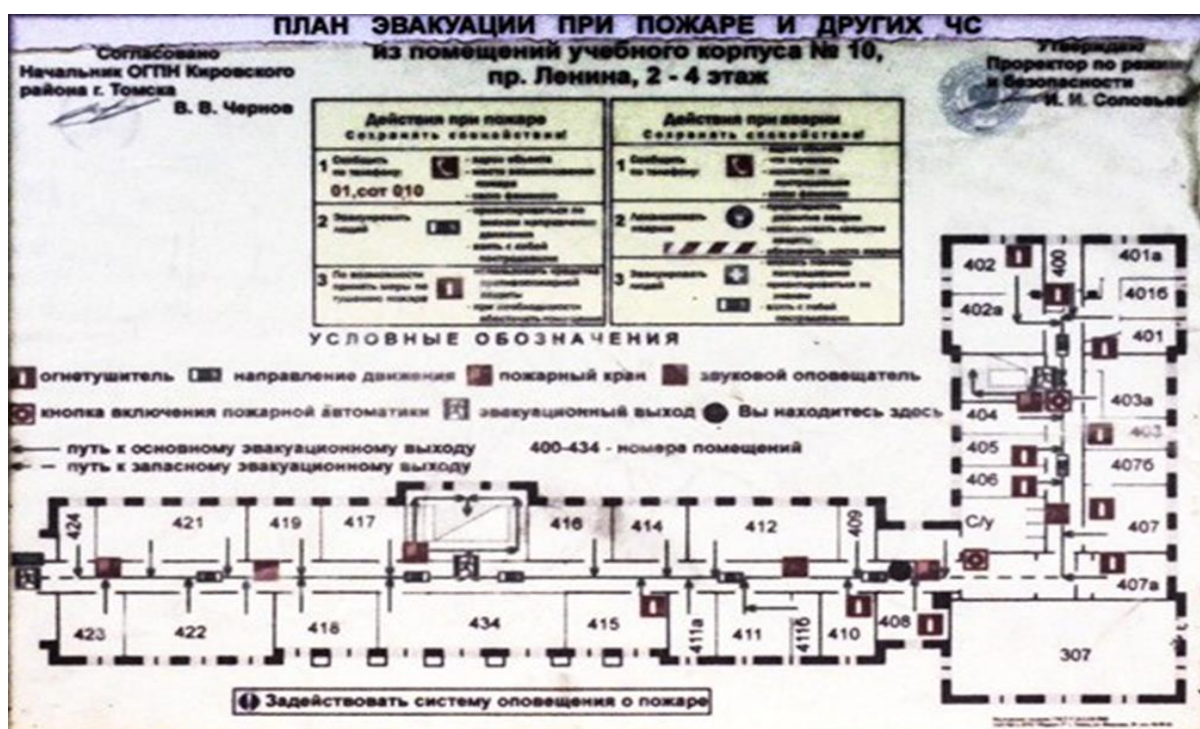


Рисунок 6.2 – План эвакуации

### **Выводы по разделу**

В работе рассмотрены основные аспекты производственной, экологической безопасности, безопасности в чрезвычайных ситуациях, правовых и организационных вопросов обеспечения безопасности, имеющие отношение к деятельности программиста в учебной лаборатории. Подобные аспекты регламентируются законодательством РФ и носят, как рекомендательный, так и обязательный характер.

Рассмотренные нормы и правила позволили получить базовые знания о правилах поведения в определённых ситуациях в рабочем помещении, о правах работника и обязанностях работодателя в течение рабочего процесса. Исполнение изученных норм и правил призвано обеспечить безопасность и ресурсоэффективность трудовой деятельности.

## **ЗАКЛЮЧЕНИЕ**

В работе представлен анализ существующих видов адресных табличек домов. Составлен алгоритм для распознавания адресных табличек домов с использованием сверточной нейронной сети. Для составления алгоритма изучены методы обработки и классификации изображений. Алгоритм включает в себя предобработку изображения, сегментирование и последующую классификацию. В качестве алгоритма классификации было проведено сравнение точностей распознавания между сверточной нейронной сетью и алгоритмом ближайшего соседа, что показало целесообразность применения сверточной нейронной сети. Составлена архитектура для двух сверточных ИНС: для детектирования адресной таблички и для распознавания символов.

В качестве демонстрации алгоритмов разработано консольное приложение на языке Python. В приложении реализованы как этап предобработки изображения, так и классификации. Классификация сегментов выполнена с помощью сверточной ИНС с обратным распространением ошибки и функцией активации ReLU. Максимальная точность распознавания в ходе экспериментов достигала 89%

## СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. ГОСТ Р 52290-2004 // Федеральное агентство по техническому регулированию и метрологии. – Москва, 2006. – 125 с
2. Нейронные сети для начинающих [Электронный ресурс]. – URL: <https://habr.com/ru/post/312450/>. (Дата обращения 15.03.2019).
3. Глава 3. Основы ИНС // Нейронные сети [Электронный ресурс]. – URL: <https://neuralnet.info/chapter/основы-инс> (Дата обращения 15.03.2019).
4. Вишник М.И. Обобщенные функции // Соровосовский образовательный журнал. – 1997. – №12. С. 112-117.
5. Знакомство с машинным обучением // Google stories [Электронный ресурс]. – URL: <https://www.google.com/intl/ru/about/stories/machine-learning-qa/>. (Дата обращения 15.03.2019).
6. Преобразование цветного изображения в черно-белое // Программирование на C, C# и Java [Электронный ресурс]. – URL: <https://vscode.ru/prog-lessons/preobrazovanie-tsvetnogo-izobrazheniya-v-chno-beloe.html>. (Дата обращения 15.03.2019).
7. OpenCV [Электронный ресурс]. – URL: <https://opencv.org/>. (Дата обращения 15.03.2019).
8. Янковский А.А., Бугрий А.Н.. Критерии выбора метода бинаризации при обработке изображений лабораторных анализов // Научно-технический журнал «АСУ и приборы автоматизации». – 2010. – №153. С. 53-56.
9. Пелевин Е.Е., Балясный С.В.. Оптимальные алгоритмы выделения контуров изображения в системе технического зрения // Juvenis scientia. – 2016. – №6. С. 6-8.
10. Хайкин С. Нейронные сети. – 2-е изд. – М.: Вильямс, 2006. – 1103с
11. Джонс М. Программирование искусственного интеллекта в приложениях / М.Джонс: Пер. с англ. Осипов А.И. – М.: ДМК Пресс, 2013. – 312 с
12. Градиентный спуск [Электронный ресурс]. – URL: [https://ru.wikipedia.org/wiki/Градиентный\\_спуск](https://ru.wikipedia.org/wiki/Градиентный_спуск). (Дата обращения 15.03.2019).

13. Николенко С., Кадури́н А., Архангельская Е. Глубокое обучение – 1-е изд. – СПб.: Питер, 2018. – 480 с.
14. Street View and reCAPTCHA technology just got smarter [Электронный ресурс]. – URL: <https://security.googleblog.com/2014/04/street-view-and-recaptcha-technology.html>. (Дата обращения 20.04.2019).
15. Multi-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks [Электронный ресурс]. – URL: <https://arxiv.org/abs/1312.6082>. (Дата обращения 20.04.2019).
16. ImageNet Classification with Deep Convolutional Neural Networks [Электронный ресурс]. – URL: <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>. (Дата обращения 22.04.2019).
17. Нейронные сети для начинающих. [Электронный ресурс]. – URL: <https://habr.com/ru/post/313216/>. (Дата обращения 22.04.2019).
18. Python. [Электронный ресурс]. – URL: <https://www.python.org/>. (Дата обращения 10.03.2019).
19. OpenCV. [Электронный ресурс]. – URL: <https://www.opencv.org/>. (Дата обращения 10.03.2019).
20. TensorFlow. [Электронный ресурс]. – URL: <https://www.tensorflow.org/>. (Дата обращения 10.03.2019).
21. Трудовой кодекс Российской Федерации. [Электронный ресурс]. – URL: [http://www.consultant.ru/document/cons\\_doc\\_LAW\\_34683/](http://www.consultant.ru/document/cons_doc_LAW_34683/). (Дата обращения 10.05.2019).
22. ГОСТ 12.2.032-78. [Электронный ресурс]. – URL: <http://docs.cntd.ru/document/1200003913/>. (Дата обращения 10.05.2019).
23. ГОСТ 12.2.061-81 ССБТ. [Электронный ресурс]. – URL: <http://docs.cntd.ru/document/gost-12-2-061-81-ssbt>. (Дата обращения 10.05.2019).
24. СанПиН 2.2.2/2.4.1340-03. [Электронный ресурс]. – URL: <http://docs.cntd.ru/document/901865498>. (Дата обращения 10.05.2019).

25. ГОСТ 12.0.003-2015. [Электронный ресурс]. – URL: <http://docs.cntd.ru/document/1200136071>. (Дата обращения 10.05.2019).
26. ГОСТ 12.0.003-2015. [Электронный ресурс]. – URL: <http://docs.cntd.ru/document/1200136071>. (Дата обращения 10.05.2019).
27. СанПиН 2.2.4.548-96. [Электронный ресурс]. – URL: <http://docs.cntd.ru/document/901704046>. (Дата обращения 10.05.2019).
28. ГОСТ 12.1.003-2014 [Электронный ресурс]. – URL: <http://docs.cntd.ru/document/1200118606>. (Дата обращения 10.05.2019).
29. СанПиН 2.2.4.3359-16 [Электронный ресурс]. – URL: <http://docs.cntd.ru/document/420362948>. (Дата обращения 10.05.2019).
30. СанПиН 2.2.1/2.1.1.1278-03 [Электронный ресурс]. – URL: <http://docs.cntd.ru/document/901859404>. (Дата обращения 10.05.2019).
31. ГОСТ 12.1.006–84 [Электронный ресурс]. – URL: <http://docs.cntd.ru/document/5200272>. (Дата обращения 10.05.2019).
32. СанПиН 2.2.2/2.4.1340–03 [Электронный ресурс]. – URL: <http://docs.cntd.ru/document/901865498>. (Дата обращения 10.05.2019).
33. Николенко С., Кадури́н А., Архангельская Е. Глубокое обучение – 1-е изд. – СПб.: Питер, 2018. – 480 с.
34. Криницына З. В., Видяев И. Г. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение / З.В. Криницына, И.Г. Видяев; Томский политехнический университет. – Томск: Изд-во Томского политехнического университета, 2014. – 73 с

## ПРИЛОЖЕНИЕ А

### Раздел 1

#### АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ ANALYSIS OF THE OBJECTIVE AREA

Студент:

Группа	ФИО	Подпись	Дата
8ИМ71	Белков Сергей Геннадьевич		

Консультант ОИТ ИШИТР:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент	Мирошниченко Е. А.	К.Т.Н.		

Консультант – лингвист ОИЯ ШБИП:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент	Сидоренко Т. В.	К.П.Н.		

# **1 ANALYSIS OF THE OBJECTIVE AREA**

## **1.1 Problem description**

Street number recognition is a specific kind of sequence recognition. The problem is to identify the house number and the street name in the image for a given image.

Although text recognition in limited areas, such as word processing, is well researched, recognizing arbitrary multi-character text in photographs is still a challenge. This difficulty arises because of the wide variation in the appearance of the text: due to the wide variety of fonts, colors, styles, orientations and position of characters. The problem of recognition is further complicated by environmental factors, such as lighting, shadows, specularity and occlusion, as well as image acquisition factors, such as resolution, motion and blurring of the focus.

## **1.2 Description of address plates**

Most house signs and address plates in the Russian Federation are made in accordance with the state standard 52290-2004 «Road signs. General technical requirements». However, the application of the standard refers to the choice of fonts and indents. The mutual arrangement of alphabetic and numerical designations may vary from sign to sign, which complicates the recognition issue. During the research revealed the following types of address plates:

1. Rectangular - standard rectangular plates. Horizontal mutual arrangement of the numeric and alphabetic parts. The house number can be either on the left or on the right relative to the street name (Figure A.1).





Figure A.1 – Rectangular plates

2. Rounded - the shape of the substrate can be rounded at the top or bottom. Possibly the horizontal and vertical mutual arrangement of the numeric and alphabetic parts (Figure A.2).



Figure A.2 – Rounded plates

3. Vertical - the shape of the substrate is mainly rectangular, the relative position of the numeric and alphabetic parts is vertical (Figure A.3).



Figure A.3 – Vertical plates

4. Numeric - the address label contains only the house number. The shape of the substrate varies from rectangular to round (Figure A.4).



Figure A.4 – Numeric plates

In this paper discusses character recognition of type 1.

### 1.3 Description of machine learning

Machine learning is a process in which the system processes a large number of examples, identifies patterns and uses them to predict the characteristics of new data.

There are two types of training. Precedent learning, or inductive learning, is based on identifying common patterns from partial empirical data. Deductive training involves the formalization of expert knowledge and their transfer to a computer as a knowledge base. Deductive learning is usually referred to the field of expert systems,

therefore the terms machine learning and learning from precedents can be considered synonymous.

Artificial intelligence and machine learning can be understood differently. In general, programs that solve simple problems for a person can be called artificial intelligence. It is known that it is easier for people to understand something, if you explain it with an example. Machine learning also allows computers to learn from examples.

Machine learning is at the junction of mathematical statistics, optimization methods and classical mathematical disciplines, but it also has own specifics related to the problems of computational efficiency and retraining. Many methods of inductive learning were developed as an alternative to classical statistical approaches. Many methods are closely related to data mining.

## **1.4 Artificial neural networks**

### **1.4.1 General Description**

A neural network is a sequence of neurons connected by synapses. The structure of the neural network came into the programming world straight from biology. Thanks to this structure, the machine gets the ability to analyze and even memorize various information. Neural networks are also able not only to analyze incoming information, but also to reproduce it from its memory.

Neural networks are used to solve complex problems that require analytical calculations similar to the human brain. The most common applications of neural networks are:

Classification – the distribution of data by parameters. For example, a set of people is given at the entrance and you need to decide which of them to give a credit, and who does not. This work can be done by a neural network, analyzing such information as: age, solvency, credit history, and so on.

Prediction – the ability to predict the next step. For example, rising or falling stocks, based on the situation on the stock market.

Recognition – currently, the widest application of neural networks. Used by Google when you are looking for a photo or in the camera phones, when it determines the position of your face and highlights it, and more.

### 1.4.2 Generalized structure of the ANN

The main components of the neural network: neurons and connections between them (synapses).

**A neuron** is a computational unit that receives information, performs simple calculations on it, and passes it on. They are divided into three main types: input, hidden and output.

Each of the neurons has 2 basic parameters: input data and output data. In the case of an input neuron: Input = Output. In the others, the total information of all neurons from the previous layer gets into the input field, after which it is normalized using the activation function and enters the output field.

**A synapse** is a connection between two neurons. Synapses have 1 parameter - weight. Thanks to him, the input information changes when it is transmitted from one neuron to another.

The complete model of an artificial neuron with synapses is shown in Figure A.5.

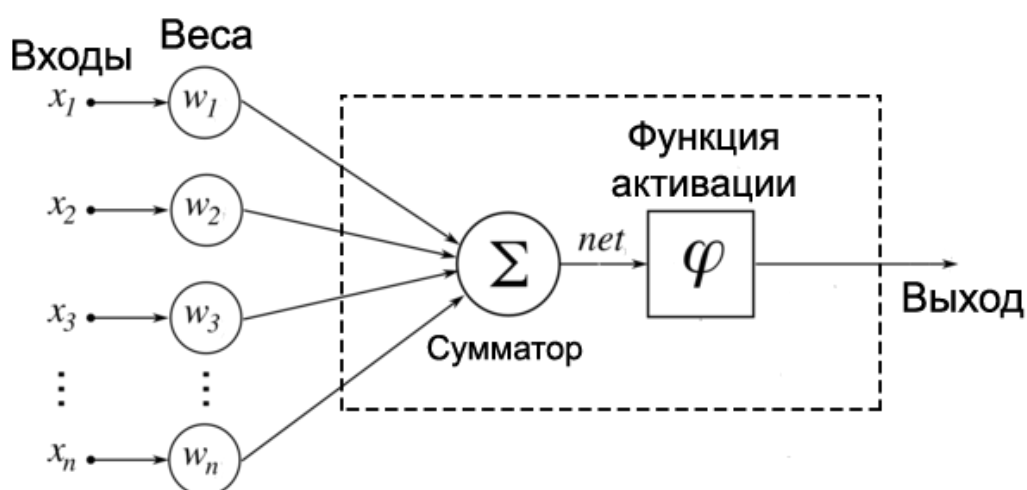


Figure A.5 – Neuron model

In the case when the neural network consists of a large number of neurons, the term layer is introduced. There is an input layer that receives information, N

hidden layers (usually no more than 3 of them) that process it and an output layer that displays the result.

Figure A.6 shows an example of the structure of an ANN with three layers.

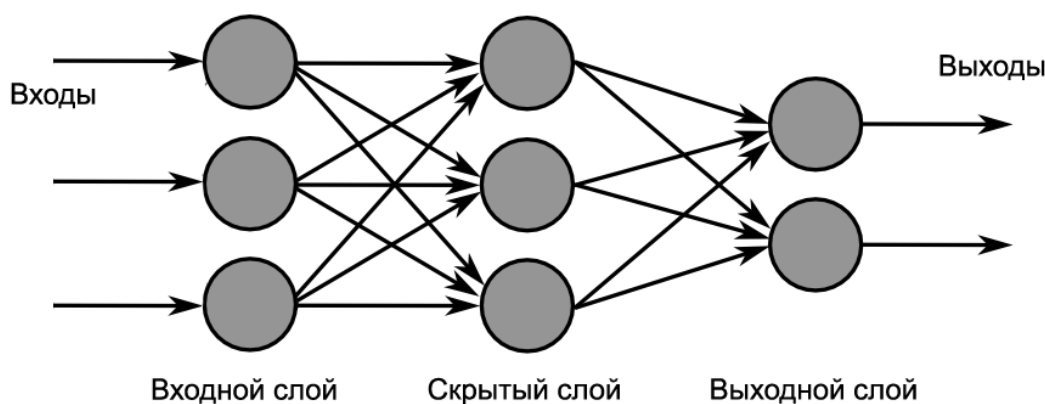


Figure A.6 – Neural network structure

An activation function is a function that takes a weighted sum as an argument. The value of this function is the output of the neuron.

The main types of activation functions:

- **Single jump function.** The easiest kind of activation function. The output of a neuron can only be 0 or 1. If the weighted sum is greater than a certain threshold  $b$ , then the output of the neuron is 1. If lower, then 0.

A graphical representation of the function is shown in Figure A.7. The weighted sum values are located on the horizontal axis. On the vertical axis - the values of the output signal.

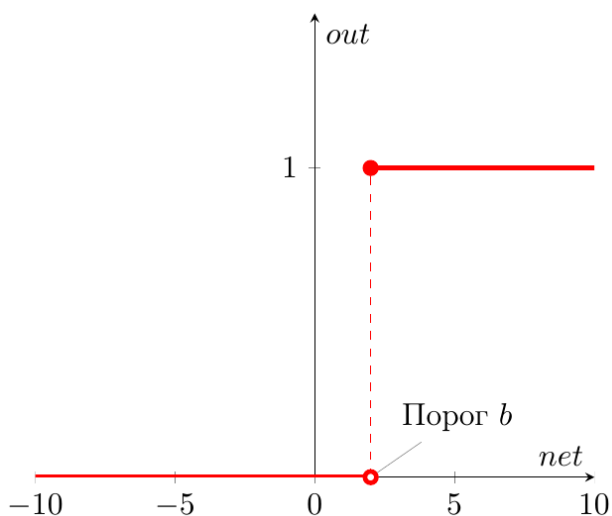


Figure A.7 – Single jump function

- **Sigmoidal function.** There is a whole family of sigmoid functions, some of which are used as an activation function in artificial neurons. The most commonly used sigmoid in neural networks is the logistic function (Figure A.8).

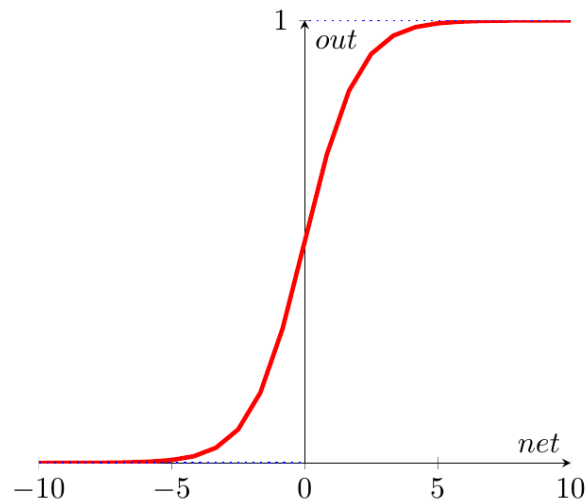


Figure A.8 – Logistic function

The analytical representation of the function is below:

$$out(net) = \frac{1}{1 + \exp(-a \cdot net)}$$

where  $a$  – the coefficient characterizing the degree of steepness of the function

$net$  – weighted sum of weights

Benefits:

1. regardless of the argument (weighted sum), the output signal will always be in the range from 0 to 1
2. more flexible than the single hop function, the result can be not only 0 and 1, but any number between them
3. it has a derivative at all points, and this derivative can be expressed through the same function

- **Activation function ReLU.** Rectified linear activation function (rectified linear unit, ReLU) (Figure A.9). Calculating sigmoids and hyperbolic tangent requires demanding operations such as exponentiation. The use of ReLU significantly increases the rate of convergence of the gradient descent (in some cases up to 6 times compared with sigmoid and hyperbolic tangent).

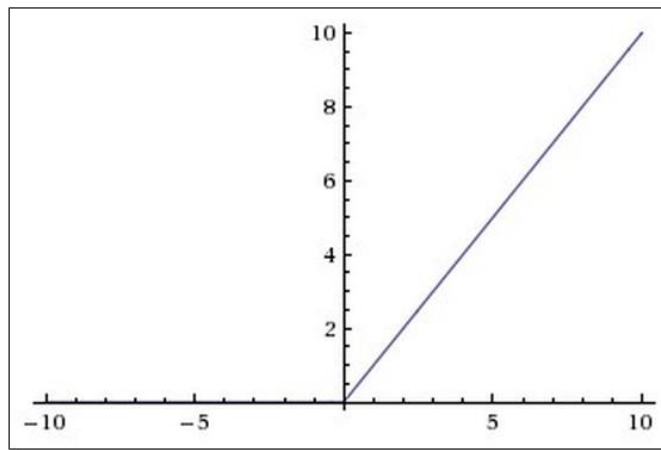


Figure A.9 – ReLU function

The equation of the function is expressed by the formula:

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$$

### 1.4.3 Bias neuron

The bias neuron is a type of neuron used in the most of neural networks. The feature of this type of neuron is that its input and output always equal 1 and they never have input synapses. Displacement neurons can either be present in the neural network one by one on a layer, or be completely absent. Connections in neurons of displacement are the same as in normal neurons – with all neurons of the next level, except that there can be no synapses between two bias neurons. Consequently, they can be placed on the input layer and all hidden layers, but not on the output layer, since they simply will not have anything to form a connection with. Figure A.10 shows an example of a B1 displacement neuron.

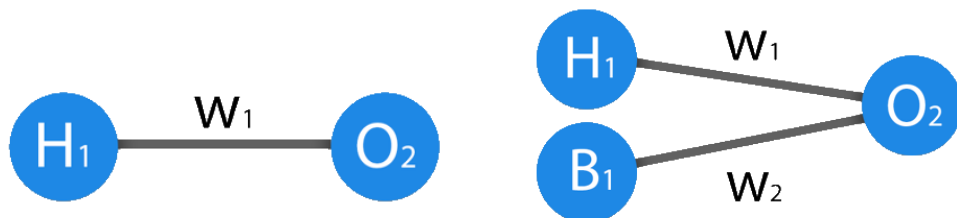


Figure A.10 – Bias neuron

Neuron bias is needed in order to be able to get the output by shifting the graph of the activation function to the right or left. Then you can establish that the

output O2 will be equal to the input H1 multiplied by its weight, and passed through the activation function:

$$\text{input} = H_1 * w_1 + H_2 * w_2 + b_3$$

$$b_3 = \text{bias} * w_3$$

When the weights of hidden and output neurons are regulated during training, the slope of the activation function changes (Figure A.11).

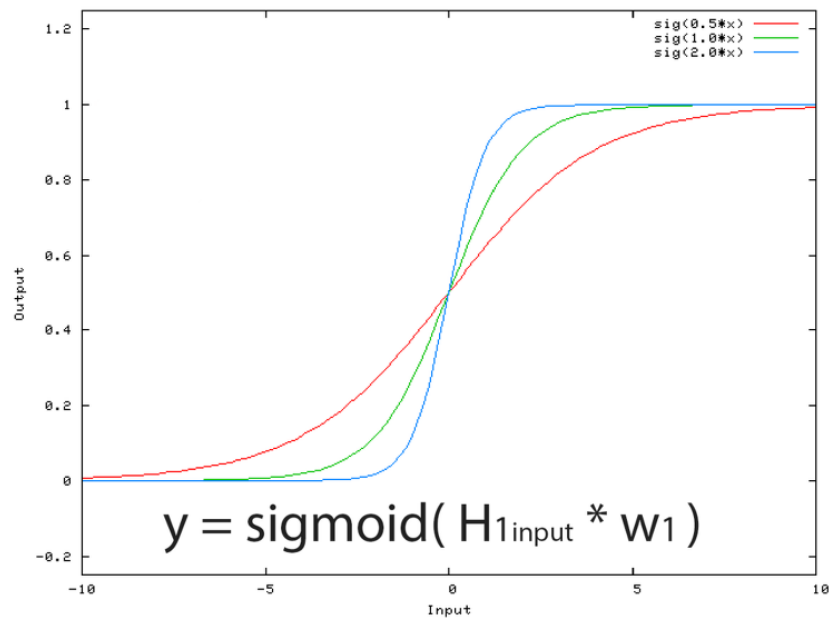


Figure A.11 – Sigmoidal function without bias neuron

Adjustment of the weight of the neuron bias can give the opportunity to shift the activation function along the X axis and capture new areas (Figure A.12). In other words, if the point responsible for the solution is located on the graph to the left or to the right, then the neural network will never be able to solve the problem without the use of displacement neurons.



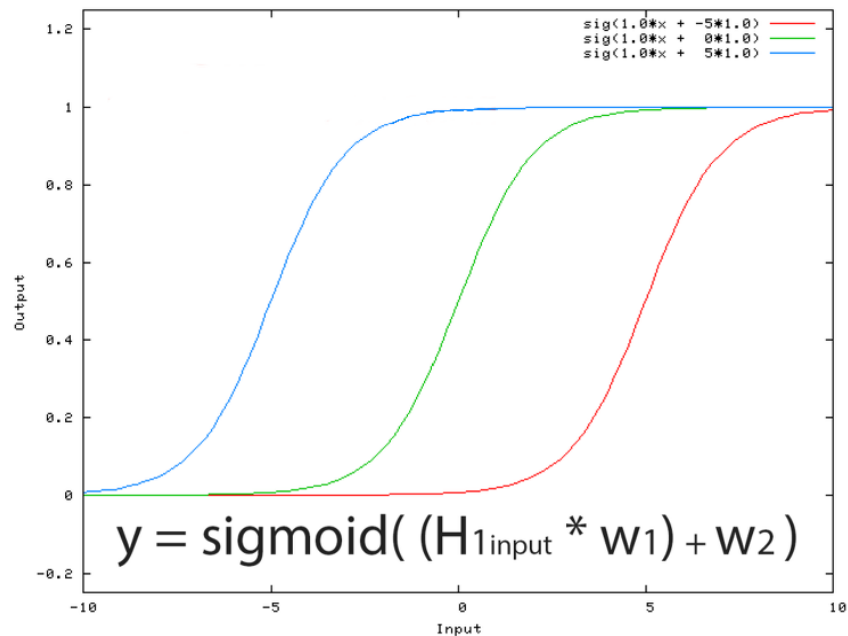


Figure A.12 – Sigmoidal function with bias neuron

#### 1.4.4 Gradient Descent Method

Gradient descent is a method of finding a local extremum (minimum or maximum) of a function by moving along the gradient. A gradient is a vector that determines the slope of a function and indicates its direction relative to a point on a surface or a graph. To find the gradient, you need to take the derivative of the function at a given point -  $f'(w_1)$  in Figure A.13.

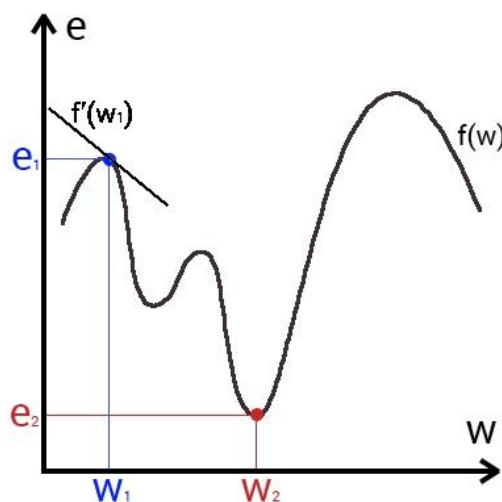


Figure A.13 – Gradient Descent Method

The pure gradient descent method can be stuck in local minima of the function. To overcome the problem of the gradient descent method is used with the

method of moments. The idea of the method is to save part of the descent speed in each iteration. In fact, the weight increment from the previous iteration multiplied by the moment coefficient is added to the weight increment.

$$\Delta W_t = \eta \cdot \nabla E + \mu * \Delta W_{t-1}$$

where  $\eta$  – the coefficient of learning speed,

$\nabla E$  – the gradient of the loss function,

$\mu$  – the moment coefficient,

$\Delta W_{t-1}$  – change of weights on the previous iteration.

Together with the moment in the method of back propagation, such parameter as the speed of learning is also used. It allows you to control the amount of correction scales at each iteration.

#### **1.4.5 Back Propagation Method**

The execution of the algorithm begins with the creation of arbitrarily generated weights for a multilayer network. Then the process described below is repeated until the average input error is found to be sufficiently small. Advance along the network counts the activation of neurons and output, the advance back - gradient.

Step 1. Take an example of the input signal  $E$  with the corresponding correct value of the output  $C$ .

Step 2. Calculate the direct distribution of  $E$  through the network - weights  $S_i$  and the activation function  $f(in)$  are determined.

Step 3. Starting with the outputs, the reverse movement through the neurons of the output and intermediate layer is performed, and the error values are calculated (equations 4 and 5):

for the output neuron

$$\delta_0 = (OUT_{ideal} - OUT_{actual}) * f'(in)$$

for all hidden neurons

$$\delta_h = \sum (w_i * \delta_i) * f'(in)$$

where  $w$  – weights of incoming synapses for a given neuron,

$\delta$  – errors for the appropriate weights,

$f(in)$  – activation function.

Step 4. In this step, the gradient descent method described in step 0 is used.  
The weights in the network are updated as follows:

$$\Delta w_i = \varepsilon * \text{grad}_w + \alpha * \Delta w_{i-1}$$

where  $\varepsilon$  – the learning rate

$\alpha$  – moment

## **ПРИЛОЖЕНИЕ Б**

CD-Диск