

Министерство образования и науки Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Институт Кибернетики

Направление подготовки 09.04.01 Информатика и вычислительная техника

Кафедра Информационных систем и технологий

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы

Метод обнаружения расфокусированных и смазанных изображений лиц на основе применения свёрточных нейронных сетей

УДК 004.932.4:004.032.26

Студент

Группа	ФИО	Подпись	Дата
8ВМ5А	Казиев Александр Борисович		

Руководитель

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

КОНСУЛЬТАНТЫ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент каф. МЕН	Конотопский Владимир Юрьевич	к.э.н., доцент		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент каф. ЭБЖ	Извеков Владимир Николаевич	к.т.н., доцент		

ДОПУСТИТЬ К ЗАЩИТЕ:

Зав. кафедрой	ФИО	Ученая степень, звание	Подпись	Дата
Зав. кафедрой ИСТ	Мальчуков Андрей Николаевич	к.т.н., доцент		

Томск – 2017 г.

**ЗАПЛАНИРОВАННЫЕ РЕЗУЛЬТАТЫ ПО ОСНОВНОЙ
ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЕ ПОДГОТОВКИ МАГИСТРОВ
09.04.01 «ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА», ИК
ТПУ, ПРОФИЛЬ «КОМПЬЮТЕРНЫЙ АНАЛИЗ И ИНТЕРПРЕТАЦИЯ
ДАННЫХ»**

Код результатов	Результат обучения (выпускник должен быть готов)	Требования ФГОС ВО (ФГОС 3+), критерии АИОР, заинтересованных работодателей и студентов
Общепрофессиональные компетенции		
Р1	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.	Требования ФГОС 3+ (ОПК-1; ПК 3-6; ОК-4), критерий 5 АИОР (п. 1.1), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
Р2	Владеть и применять методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе в глобальных компьютерных сетях.	Требования ФГОС 3+ (ОПК-5; ПК-7; ОК-7), критерий 5 АИОР (п. 1.1, 1.2), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
Р3	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.	Требования ФГОС 3+ (ОПК-6; ПК-1,2; ОК-1,2), критерий 5 АИОР (п. 1.2), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
Р4	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	Требования ФГОС 3+ (ОПК-3,4; ПК-11,12; ОК-3), критерий 5 АИОР (п. 1.6, п. 2.2), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов,

Код результатов	Результат обучения (выпускник должен быть готов)	Требования ФГОС ВО (ФГОС 3+), критерии АИОР, заинтересованных работодателей и студентов
	профессионального общения, применять специальную лексику и профессиональную терминологию языка.	отечественных и зарубежных работодателей.
Профессиональные компетенции		
P5	Выполнять инновационные инженерные проекты по разработке аппаратных и программных средств автоматизированных систем различного назначения с использованием современных методов проектирования, систем автоматизированного проектирования, передового опыта разработки конкурентно способных изделий.	Требования ФГОС 3+ (ПК-8–12; ОПК-2, ПК-7,6), критерий 5 АИОР (п. 1.3), соответствующий международным стандартам EUR- ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
P6	Планировать и проводить теоретические и экспериментальные исследования в области проектирования аппаратных и программных средств автоматизированных систем с использованием новейших достижений науки и техники, передового отечественного и зарубежного опыта. Критически оценивать полученные данные и делать выводы.	Требования ФГОС 3+ (ПК-1–7; ОПК-6; ОК-4,9), критерий 5 АИОР (п.1.4), соответствующий международным стандартам EUR- ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
P7	Осуществлять авторское сопровождение процессов проектирования, внедрения и эксплуатации аппаратных и программных средств автоматизированных систем различного назначения.	Требования ФГОС 3+ (ПК-13–19; ОПК-5; ОК-8), критерий 5 АИОР (п. 1.5), соответствующий международным стандартам EUR- ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
Общекультурные компетенции		
P8	Использовать на практике умения и навыки в организации исследовательских, проектных работ и профессиональной эксплуатации современного оборудования и приборов, в управлении коллективом.	Требования ФГОС 3+ (ОК-5,8; ОПК-1,6; ПК-6,7,11,12), критерий 5 АИОР (п. 2.1, п. 2.3, п. 1.5), соответствующий международным стандартам EUR- ACE и FEANI. Запросы студентов,

Код результатов	Результат обучения (выпускник должен быть готов)	Требования ФГОС ВО (ФГОС 3+), критерии заинтересованных работодателей и студентов
		отечественных и зарубежных работодателей.
P9	Осуществлять коммуникации в профессиональной среде и в обществе в целом, активно владеть иностранным языком, разрабатывать документацию, презентовать и защищать результаты инновационной инженерной деятельности, в том числе на иностранном языке.	Требования ФГОС 3+ (ОК-2,9; ОПК-4; ПК-1), критерий 5 АИОР (п. 2.2), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
P10	Совершенствовать и развивать свой интеллектуальный и общекультурный уровень. Проявлять инициативу, в том числе в ситуациях риска, брать на себя всю полноту ответственности.	Требования ФГОС 3+ (ОК-1,6; ОПК-2; ПК-1,2), критерий 5 АИОР (п. 2.4, п. 2.5), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
P11	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.	Требования ФГОС 3+ (ОК-3,4,7; ОПК-3; ПК-7), критерий 5 АИОР (п. 2.6), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.

Министерство образования и науки Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Институт Кибернетики

Направление подготовки 09.04.01 Информатика и вычислительная техника

Кафедра Информационных систем и технологий

УТВЕРЖДАЮ:

Зав. кафедрой

_____ Мальчуков А.Н.
(Подпись) (Дата) (Ф.И.О.)

ЗАДАНИЕ

на выполнение выпускной квалификационной работы

В форме:

Магистерской диссертации

Студенту:

Группа	ФИО
8ВМ5А	Казиеву Александру Борисовичу

Тема работы:

Метод обнаружения расфокусированных и смазанных изображений лиц при помощи свёрточных нейронных сетей

Утверждена приказом директора (дата, номер)

От 20.02.2017 №898/с

Срок сдачи студентом выполненной работы:

13.06.2017

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	Разработка метода обнаружения расфокусированных или смазанных изображений лиц при помощи аппарата свёрточных нейронных сетей. На метод налагаются требования точности и быстродействия, соответствующие условиям практического применения.
Перечень подлежащих исследованию, проектированию и разработке вопросов	<ol style="list-style-type: none">1. Обзор методов оценки/обнаружения размытых изображений, моделей классификаторов, подбор алгоритма обучения и библиотеки глубокого обучения.2. Подготовка обучающей выборки.3. Подбор параметров модели классификатора (подбор параметров обучения, проектирование и выбор архитектуры свёрточных нейронных сетей).

	4. Сравнительный анализ работы обученных моделей классификатора, формулировка метода обнаружения размытых изображений лиц. 5. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение. 6. Социальная ответственность.
Перечень графического материала	1. Примеры изображений обучающих выборок для различных векторов признаков. 2. Схемы модулей архитектур свёрточных нейронных сетей. 3. Диаграммы результатов тестирования на каждом этапе подбора параметров обучения. 4. Диаграммы результатов тестирования на каждом этапе подбора архитектуры.

Консультанты по разделам выпускной квалификационной работы	
Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Конотопский Владимир Юрьевич
Социальная ответственность	Извеков Владимир Николаевич

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	
---	--

Задание выдал руководитель:

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

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ВМ5А	Казиев Александр Борисович		

Министерство образования и науки Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Институт Кибернетики

Направление подготовки 09.04.01 Информатика и вычислительная техника

Уровень образования Магистратура

Кафедра Информационных систем и технологий

Период выполнения весенний семестр 2016/2017 учебного года

Форма представления работы:

Магистерская диссертация

КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН
выполнения выпускной квалификационной работы

Срок сдачи студентом выполненной работы:	13.06.2017
--	------------

Дата контроля	Название раздела (модуля) /вид работы (исследования)	Максимальный балл раздела (модуля)
08.02.2017	Обзор методов оценки/обнаружения размытых изображений, моделей классификаторов, подбор алгоритма обучения и библиотеки глубокого обучения	15
06.04.2017	Подготовка обучающей выборки	25
03.05.2017	Подбор параметров модели классификатора (подбор параметров обучения, проектирование и выбор архитектуры свёрточных нейронных сетей)	20
13.05.2017	Сравнительный анализ работы обученных моделей классификатора, формулировка метода обнаружения размытых изображений лиц	20
06.06.2017	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	10
08.06.2017	Социальная ответственность	10

Составил преподаватель:

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

СОГЛАСОВАНО:

Зав. кафедрой	ФИО	Ученая степень, звание	Подпись	Дата
ИСТ	Мальчуков Андрей Николаевич	к.т.н., доцент		

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

Студенту:

Группа	ФИО
8ВМ5А	Казиев Александр Борисович

Институт	Кибернетики	Кафедра	Информационных систем и технологий
Уровень образования	Магистрант	Направление/специальность	09.04.01 Информатика и вычислительная техника

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

1. <i>Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих</i>	– <i>стоимость материально-технических, энергетических и человеческих ресурсов НИ.</i>
2. <i>Нормы и нормативы расходования ресурсов</i>	
3. <i>Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования</i>	– <i>ставки социального налога и НДС.</i>

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. <i>Оценка коммерческого и инновационного потенциала НТИ</i>	– <i>оценка коммерческого потенциала НТИ.</i>
2. <i>Разработка устава научно-технического проекта</i>	
3. <i>Планирование процесса управления НТИ: структура и график проведения, бюджет, риски и организация закупок</i>	– <i>график выполнения ВКР</i>
4. <i>Определение ресурсной, финансовой, экономической эффективности</i>	

Перечень графического материала (с точным указанием обязательных чертежей):

1. *График проведения и бюджет НТИ;*
2. *Оценка ресурсной, финансовой и экономической эффективности НТИ.*

Дата выдачи задания для раздела по линейному графику

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент каф. МЕН	Конотопский Владимир Юрьевич	к.э.н., доцент		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ВМ5А	Казиев Александр Борисович		

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»**

Студенту:

Группа	ФИО
8ВМ5А	Казиёву Александрю Борисовичу

Институт	Кибернетики	Кафедра	Информационных систем и технологий
Уровень образования	Магистратура	Направление/специальность	09.04.01

Исходные данные к разделу «Социальная ответственность»:

1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения	Разработка метода обнаружения размытых изображений лиц, предназначенного для применения в системах распознавания лиц и системах оценки качества изображений лиц.
2. Знакомство и отбор законодательных и нормативных документов по теме	Выбор подходящих нормативов и документов (ГОСТ 12.0.003-74, ГОСТ 12.1.005-88, ГОСТ 12.1.004-91, ГОСТ 12.1.019, СНиП 21-01-97 СанПиН 2.2.4.548-96, СанПиН 2.2.1/2.1.1.1278-03, СанПиН 2.2.2/2.4.1340-03, Технологические регламенты), для обеспечения соответствия условий труда Трудовому кодексу РФ.

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

<p>1. Производственная безопасность</p> <p>1.1. Анализ выявленных вредных факторов проектируемой производственной среды в следующей последовательности:</p> <ul style="list-style-type: none"> • физико-химическая природа вредности, её связь с разрабатываемой темой; • действие фактора на организм человека; – приведение допустимых норм с необходимой размерностью (со ссылкой на соответствующий нормативно-технический документ); – предлагаемые средства защиты (сначала коллективной защиты, затем – индивидуальные защитные средства) <p>1.2. Анализ выявленных опасных факторов проектируемой производственной среды в следующей последовательности</p> <ul style="list-style-type: none"> • механические опасности (источники, средства защиты); • термические опасности (источники, средства защиты); • электробезопасность (в т.ч. статическое электричество, молниезащита – источники, средства защиты); • пожаровзрывобезопасность (причины, профилактические мероприятия, первичные средства пожаротушения) 	<ul style="list-style-type: none"> – требования к помещениям для работы с ПЭВМ – освещённость на рабочем месте; – уровень шума на рабочем месте; – микроклимат (пониженная или повышенная температура, пониженная или повышенная влажность воздуха); – электромагнитное излучение; – повышенные нагрузки на органы зрения; – монотонность труда. <ul style="list-style-type: none"> – электрический ток.
--	---

<p>2. Экологическая безопасность:</p> <ul style="list-style-type: none"> • защита селитебной зоны • анализ воздействия объекта на атмосферу (выбросы); • анализ воздействия объекта на гидросферу (сбросы); • анализ воздействия объекта на литосферу (отходы); • разработать решения по обеспечению экологической безопасности со ссылками на НТД по охране окружающей среды. 	<ul style="list-style-type: none"> - анализ воздействия объекта ВКР и области его использования на ОС; - разработка решений по обеспечению экологической безопасности
<p>3. Безопасность в чрезвычайных ситуациях:</p> <ul style="list-style-type: none"> • перечень возможных ЧС на объекте; • выбор наиболее типичной ЧС; • разработка превентивных мер по предупреждению ЧС; • разработка мер по повышению устойчивости объекта к данной ЧС; • разработка действий в результате возникшей ЧС и мер по ликвидации её последствий 	<ul style="list-style-type: none"> - выбор и описание возможных ЧС: типичная ЧС – пожар; - разработка превентивных мер по предупреждению ЧС; - разработка действий в результате возникшей ЧС и мер по ликвидации её последствий.
<p>4. Правовые и организационные вопросы обеспечения безопасности:</p> <ul style="list-style-type: none"> • специальные (характерные для проектируемой рабочей зоны) правовые нормы трудового законодательства; • организационные мероприятия при компоновке рабочей зоны 	<ul style="list-style-type: none"> - специальные правовые нормы трудового законодательства; - организационные мероприятия при компоновке рабочей зоны программиста.
<p>Перечень графического материала:</p>	
<p>При необходимости представить эскизные графические материалы к расчётному заданию (обязательно для специалистов и магистров)</p>	<ul style="list-style-type: none"> - план размещения светильников; - план эвакуации при пожаре.

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

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Извеков В.Н.	к.т.н., доцент		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ВМ5А	Казиев Александр Борисович		

РЕФЕРАТ

Выпускная квалификационная работа 118 страниц (без приложений), 25 рисунков, 22 таблицы, 59 источников.

Ключевые слова: оценка качества изображений, классификация изображений, расфокусированные, смазанные, размытые изображения, свёрточные нейронные сети, частотный спектр.

Объектом исследования является входной контроль качества изображений лиц для систем распознавания и поисковых баз данных.

Целью работы является разработка метода обнаружения расфокусированных или смазанных изображений лиц при помощи аппарата свёрточных нейронных сетей (СНС).

В процессе исследования производились получение и подготовка обучающей выборки, подбор параметров обучения и архитектуры СНС, сравнительный анализ эффективности использования различных входных векторов признаков для обучения СНС.

Основные конструктивные, технологические и технико-эксплуатационные характеристики: реализация СНС и алгоритмов обучения на языке C++ при помощи библиотеки Dlib-ml.

Степень внедрения: готовится к внедрению.

Область применения: входной контроль качества изображений лиц для систем распознавания и поисковых баз данных.

Значимость работы заключается в том, что был впервые предложен метод обнаружения размытых изображений лиц на основе СНС, приведены данные об эффективности использования различных входных векторов признаков и оптимальные параметры обучения СНС.

В будущем планируется дальнейшее развитие метода, тестирование на дополнительных тестовых выборках, оптимизация и усовершенствование архитектуры СНС, проведение обширного сравнения с другими методами с публикацией результатов.

ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ, СОКРАЩЕНИЯ И НОРМАТИВНЫЕ ССЫЛКИ

Определения:

Обучение модели классификатора – процесс оптимизации параметров модели (в случае ИНС весовых коэффициентов) таким образом, чтобы классификатор успешно справлялся с задачей классификации.

Градиентный спуск – метод нахождения экстремумов функции путём движения по вектору градиента функции в пространстве параметров.

Фурье-образ – функция, описывающая частотный спектр входного сигнала, результат применения преобразования Фурье или дискретного преобразования Фурье.

Сокращения:

ИНС – искусственная нейронная сеть.

СНС – свёрточная нейронная сеть.

ДПФ – дискретное преобразование Фурье.

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	16
1. ТЕОРЕТИЧЕСКАЯ ЧАСТЬ	19
1.1. Классификация методов оценки качества изображений.....	19
1.2. Неэталонные методы оценки размытия неспецифичных изображений 21	
1.3. Модели классификаторов	26
1.3.1. Байесовский классификатор	26
1.3.2. Метод опорных векторов	28
1.3.3. Свёрточные нейронные сети	30
1.4. Модификации метода градиентного спуска	33
1.4.1. Пакетный и стохастический градиентный спуск	33
1.4.2. Ускоренный градиентный спуск Нестерова	35
1.4.3. Adagrad.....	36
1.4.4. RMSProp и AdaDelta	37
1.4.5. Adam.....	38
1.4.6. Выбор модификации метода градиентного спуска.....	38
1.4.7. Батч-нормализация	39
1.4.8. Регуляризация	40
1.5. Библиотеки глубокого обучения для реализации СНС	41
1.5.1. Torch.....	42
1.5.2. Theano.....	42
1.5.3. Pylearn2	43
1.5.4. Caffe.....	43
1.5.5. Dlib-ml.....	44
1.5.6. Выбор библиотеки глубокого обучения.....	44
1.6. Модификации архитектуры СНС	45
1.6.1. Архитектура LeNet	45
1.6.2. Архитектура ResNet.....	47
1.6.3. Архитектура Inception	47
2. ПРОЕКТИРОВАНИЕ И ОБУЧЕНИЕ СНС	50

2.1.1.	Выборка исходных изображений	50
2.1.2.	Варианты входного вектора признаков	52
2.1.3.	Тестовая выборка	54
2.2.	Подбор параметров обучения СНС	54
2.2.1.	Процедура подбора параметров обучения	54
2.2.2.	Подбор параметров обучения для выборки исходных изображений	56
2.2.3.	Подбор параметров обучения для выборки модульных составляющих частотного спектра	60
2.2.4.	Подбор параметров обучения для выборки фазовых составляющих частотного спектра	64
2.2.5.	Оптимальные параметры обучения СНС	68
3.	ПОДБОР АРХИТЕКТУРЫ СНС	69
3.1.	Критерии подбора вариантов архитектуры СНС	69
3.2.	Варианты архитектуры СНС	69
3.3.	Выбор архитектур по критерию быстродействия	70
3.4.	Выбор архитектур по критерию точности	71
4.	ФОРМУЛИРОВКА МЕТОДА ОБНАРУЖЕНИЯ РАСФОКУСИРОВАННЫХ И СМАЗАННЫХ ИЗОБРАЖЕНИЙ ЛИЦ	77
5.	ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ	79
5.1.	Организация и планирование работ	79
5.2.	Продолжительность этапов работ	80
5.3.	Расчёт накопления готовности проекта	84
5.4.	Расчёт сметы затрат на выполнение проекта	85
5.4.1.	Расчёт затрат на материалы	86
5.4.2.	Расчёт заработной платы	86
1.1.1.	Расчет затрат на социальный налог	86
5.4.3.	Расчет затрат на электроэнергию	87
5.4.4.	Расчет амортизационных расходов	87
5.4.5.	Расчет расходов, учитываемых непосредственно на основе платежных (расчетных) документов (кроме суточных)	88

5.4.6. Расчет прочих расходов	88
5.4.7. Расчет общей себестоимости разработки.....	88
5.4.8. Расчет прибыли	89
5.4.9. Расчет НДС	89
5.4.10. Цена разработки НИР	89
5.5. Оценка экономической эффективности проекта.....	90
5.6. Оценка научно-технического уровня НИР	91
6. СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ.....	93
Аннотация	93
Введение.....	93
6.1. Производственная безопасность	94
6.1.1. Анализ вредных и опасных факторов, которые может создать объект исследования	94
6.1.2. Обоснование мероприятий по защите персонала предприятия от действия опасных и вредных факторов (техника безопасности и производственная санитария)	96
6.2. Экологическая безопасность	106
6.3. Безопасность в чрезвычайных ситуациях	106
6.4. Правовые и организационные вопросы обеспечения безопасности. 109	
6.4.1. Специальные правовые нормы трудового законодательства 109	
6.4.2. Организационные мероприятия при компоновке рабочей зоны программиста	109
ЗАКЛЮЧЕНИЕ	112
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	114
ПРИЛОЖЕНИЕ А. ЧАСТЬ ВКР НА АНГЛИЙСКОМ ЯЗЫКЕ.....	120
ПРИЛОЖЕНИЕ Б. ПРОГРАММНЫЙ КОД РЕАЛИЗАЦИИ АРХИТЕКТУР СНС.....	135
ПРИЛОЖЕНИЕ В. CD-ДИСК	139

ВВЕДЕНИЕ

Задачи, которые ставятся перед современными системами видеонаблюдения в составе систем распознавания лиц, предполагают работу в сложных условиях, из-за чего входные изображения подвержены искажениям различного рода [1]. Одним из наиболее частых искажений является размытие в результате смазывания в движении или расфокусировки. Изображения, подверженные размытию, теряют значительную часть высокочастотной информации и становятся непригодными для дальнейшей обработки распознавателем, так как падает вероятность правильной классификации лица [2]. Имеет смысл отбраковывать такие изображения на этапе между детектированием и распознаванием лиц, с целью снижения нагрузки на вычислительную систему и уменьшения количества ложноположительных срабатываний распознавателя. Данная процедура также может быть актуальной на этапе входного контроля эталонных изображений, входящих, к примеру, в поисковые базы данных.

Объектом исследования, таким образом, является входной контроль качества изображений лиц для систем распознавания и поисковых баз данных. Предметом исследования и целью магистерской диссертации является разработка метода обнаружения расфокусированных или смазанных (далее размытых) изображений лиц при помощи аппарата свёрточных нейронных сетей (СНС), который удовлетворял бы требованиям точности и быстродействия, соответствующим условиям практического применения. В ходе выполнения работы также ставится задача исследовать сравнительную эффективность различных вариантов вектора признаков, среди которых рассматриваются исходные изображения и частотный спектр изображений, полученный при помощи дискретного преобразования Фурье.

Для достижения поставленной цели были сформулированы следующие задачи:

1. Провести обзор нестандартных методов оценки/обнаружения размытых изображений, используемых моделей классификаторов, подбор алгоритма обучения и библиотеки глубокого обучения с обоснованием выбора.
2. Подготовить обучающую выборку изображений лиц с низкой и высокой степенью размытия для обучения СНС и тестовую выборку для тестирования. Сформулировать варианты входного вектора признаков и подготовить обучающие выборки для каждого из них путём обработки изображений исходной обучающей выборки.
3. Произвести подбор параметров обучения СНС, проектирование и подбор архитектуры СНС, для достижения наилучшего результата на каждом из предложенных вариантов входного вектора признаков.
4. Произвести сравнительный анализ эффективности использования каждого входного вектора признаков в задаче обнаружения размытых изображений лиц при помощи СНС.
5. Сформулировать метод обнаружения размытых изображений лиц, удовлетворяющий требованиям к быстродействию и точности классификации и пригодный для применения в практических условиях.

В рамках выполнения данной магистерской диссертации были достигнуты результаты, обладающие практической значимостью. Была проделана работа по получению и подготовке обучающей выборки изображений лиц, которая в дальнейшем может быть использована для обучения и тестирования моделей классификаторов. Были произведены эмпирический подбор оптимальных параметров обучающей модели, проектирование и подбор архитектуры СНС в рамках поставленных требований и выполнен анализ результатов, позволяющий оценить сравнительную эффективность использования различных видов входных векторов признаков, полученных из обучающей выборки. Конечным результатом работы является метод обнаружения расфокусированных и смазанных изображений лиц на основе СНС, включающий рекомендации по

проектированию и обучению СНС на подготовленной обучающей выборке, а также реализация метода в виде обученной СНС. Важнейшей составляющей апробации результатов является тот факт, что метод принят к использованию предприятием-заказчиком, помимо этого подготовлены доклады научно-технических конференций.

1. ТЕОРЕТИЧЕСКАЯ ЧАСТЬ

1.1. Классификация методов оценки качества изображений

Предпосылками для разработки метода оценки качества изображений является наличие или отсутствие определённых данных, а именно: информации об исходном изображении идеального качества, информации о содержащихся в изображении искажениях и релевантных знаний о функционировании зрительной системы человека. Данные виды информации также являются критериями для классификации методов оценки качества изображений.

Первым критерием для классификации является наличие информации об исходном изображении, которое изначально не содержит искажений и может быть использовано в качестве эталона при оценке качества изображения, содержащего искажения. Согласно этому критерию методы оценки качества изображений разделяются на три класса:

1. Эталонные методы (full-reference image quality assessment (IQA), FR-IQA), которые предполагают наличие исходного изображения. В данном случае оценка качества изображения коррелирует со степенью схожести с эталонным изображением, в связи с чем речь идёт в первую очередь о сравнении изображений между собой.
2. Неэталонные методы (no-reference IQA, NR-IQA), которые предполагают отсутствие исходного изображения. Данные методы, как правило, наиболее сложны в реализации.
3. Псевдоэталонные методы (reduced-reference IQA, RR-IQA), для применения которых вместе с изображением, содержащим искажения, должна быть предоставлена ограниченная информация об исходном изображении, объём которой, как правило, значительно меньше самого исходного изображения.

Вторым критерием классификации являются рамки применения метода. По этому критерию методы оценки качества изображений разделяются на два больших класса:

1. Общего назначения (general-purpose IQA), не привязанные к ограниченному набору видов искажений, которым может подвергаться изображение. Такие методы должны быть достаточно гибкими, чтобы быть применимыми в широком спектре задач.
2. Специализированные методы оценки, применимые в узком спектре задач и предназначенные, как правило, для оценки уровня ограниченного набора видов искажений или даже одного определённого вида.

Наконец, третий критерий основывается на использовании нисходящего или восходящего подхода к реализации модели для оценки качества изображений:

1. Восходящий подход является наиболее распространённым. Он заключается в изучении функционирования каждого компонента зрительной системы человека и последующем моделировании их значимых составляющих и психофизиологических особенностей с целью получения необходимой функциональности.
2. Нисходящий подход основан на моделировании общей функциональности зрительной системы человека без использования информации о её внутреннем строении. Этот подход рассматривает зрительную систему человека в качестве «чёрного ящика» и не привязан к конкретной реализации тех или иных функций.

В рамках данной работы рассматривается неэталонный способ оценки качества изображений лиц в контексте оценки уровня размытия этих изображений. Результаты оценки используются для бинарной классификации изображений, то есть их распределения по двум категориям – с высокой или

низкой степенью размытия, что позволяет говорить о задаче обнаружения размытых изображений лиц. Использование свёрточной нейронной сети, аппарат которой основан на изучении и моделировании работы зрительной системы человека, в именно визуальной коры головного мозга, относится к способам, разработанным в рамках восходящего подхода [3].

1.2. Неэталонные методы оценки размытия неспецифичных изображений

Методы оценки размытия изображения в общем случае основаны на двух подходах [4]:

- анализе частотного спектра изображений, полученного при помощи дискретного преобразования Фурье;
- анализе структурной информации изображения, такой как наличие резких границ или перепадов в пространстве изображения.

Преобразование Фурье – один из наиболее распространённых инструментов спектрального анализа сигналов, применяемых в различных отраслях науки и техники. С его помощью функция $y(x)$, описывающая некоторый сигнал, может быть представлена в частотном пространстве путём её разложения на гармонические составляющие, то есть синусоиды с определёнными частотами и амплитудами. Представление функции в частотном пространстве называется частотным спектром функции, а функция, описывающая частотный спектр называется фурье-образом. В общем случае для получения фурье-образа используется следующая формула:

$$f(\omega) = \int_{-\infty}^{+\infty} y(x) * e^{-\omega x i} dx. \quad (1.1)$$

Для восстановления функции по её фурье-образу можно использовать формулу обратного преобразования:

$$y(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\omega) * e^{\omega xi} d\omega. \quad (1.2)$$

Так как сигналы, кодирующие изображения, задаются в виде набора дискретных значений, для получения частотного спектра изображения нужно использовать дискретное преобразование Фурье, осуществляемое при помощи следующей формулы:

$$f(k) = \sum_{n=0}^{N-1} y(n) * e^{-\frac{2\pi nki}{N}}, \quad (1.3)$$

где $n = 0 \dots (N - 1)$ – номер дискретного значения исходного сигнала, а $k = 0 \dots (N - 1)$ – номер значения фурье-образа.

Как и в случае непрерывных значений, существует формула обратного преобразования для получения исходного сигнала по фурье-образу:

$$y(n) = \frac{1}{N} \sum_{k=0}^{N-1} f(k) * e^{\frac{2\pi nki}{N}}. \quad (1.4)$$

Значения фурье-образа $f(k)$ принимают комплексные значения, из которых могут быть вычислены модульные и фазовые составляющие, характеризующие соответственно амплитуду и фазу одной гармоники частотного спектра [5].

Преобразуя двумерные данные, такие как изображения, мы получаем двумерный частотный спектр, который можно визуализировать, вычислив его модульные или фазовые составляющие. Примеры представлены на рисунке 1.1.

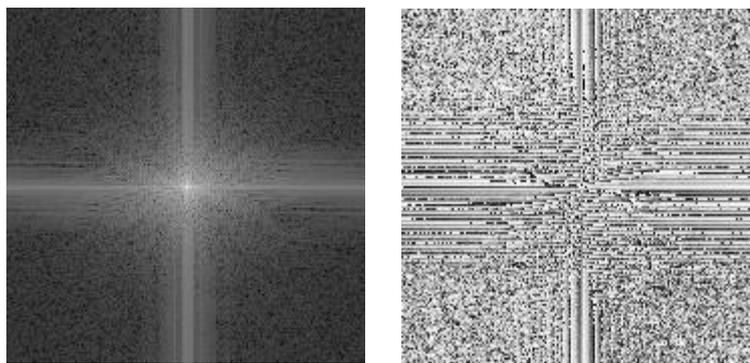


Рисунок 1.1. Примеры изображений модульных (слева) и фазовых (справа) составляющих частотного спектра изображения.

Известно, что искажение размытия, применённое к изображению, снижает как резкость (качество воспроизведения крупных деталей изображения), так и чёткость (различимость мелких деталей). Потеря информации о резких переходах и мелких деталях изображения говорит о том, что происходит потеря высокочастотной информации в изображении. Анализ частотного спектра как раз и позволяет судить о наличии и качестве информации определённых частот в исходном изображении.

В работе [6] предлагается следующий алгоритм анализа частотного спектра с целью оценки уровня размытия в изображении:

1. Получить частотный спектр изображения при помощи ДПФ.
2. Центрировать частотный спектр и вычислить модульные составляющие каждой гармоники.
3. Рассчитать T_H – общее количество модульных составляющих, чья величина превышает пороговое значение, равное $M_{Mod} / 1000$, где M_{Mod} – максимальная величина модульных составляющих.
4. Получить оценку качества изображения по отношению к искажению размытия по формуле:

$$FM = \frac{T_H}{M * N'} \quad (1.5)$$

где M и N – размеры изображения.

Представленный алгоритм был реализован и использован в данной работе в качестве основы для получения минимальной точности классификации на подготовленной в рамках работы тестовой выборке изображений лиц. Результат тестирования алгоритма приводится в подразделе 3.4. Сами авторы алгоритма не приводят данных о точности классификации при помощи данного алгоритма.

В работе [7] предлагается ещё один метод, разработанный в рамках данного подхода, в котором характеристики частотного спектра изображения в целом или его области на вход байесовского классификатора для получения многоклассовой классификации по уровню размытия изображения. По результатам тестирования на собственной тестовой базе изображений метод демонстрирует точность классификации, равную 76.98%.

Существенным недостатком перечисленных методов является игнорирование структурной информации в изображении, так как изображение модульных составляющих частотного спектра не содержит информацию о пространственном распределении высокочастотных и низкочастотных элементов изображения в явном виде. Таким образом, изображение, подверженное, к примеру, искажению импульсного шума, которое может быть результатом сбоя в настройке устройств видеонаблюдения, будет считаться резким вне зависимости от различимости информативной части изображения, то есть частей лица. Наличие дополнительных искажений или неравномерности областей размытия в пространстве изображения также снижает эффективность этих алгоритмов.

Другой подход к оценке размытости изображений заключается в анализе структурной информации в изображении. В работе [8] предлагается метрика общего размытия изображения на основе анализа границ изображения. В рамках алгоритма в изображении обнаруживаются все вертикальные границы, после чего вычисляется ширина каждой границы. На основе средней ширины границ выводится оценка размытости изображения. Авторы работы приводят

результаты в 96% точности классификации на собственной выборке резких изображений и искусственно полученных расфокусированных изображений.

В работе [9] используется вейвлет-преобразование Хаара для классификации типов границ в изображении и последующей классификации изображения на предмет размытости при помощи простых решающих правил. В этой же работе предлагается алгоритм для получения оценки степени размытия. Согласно результатам тестирования, приведённым авторами работы, предложенный метод даёт общую точность классификации, равную 98.60%, на тестовой выборке из резких изображений базы Corel и искусственно полученных размытых изображений.

Существенным недостатком алгоритмов, основанных на анализе границ в изображениях, является сильная зависимость результата от количества и качества границ в изображении, и, как следствие, узкий диапазон пригодных для анализа естественных изображений. Данные алгоритмы также становятся неэффективными при наличии дополнительных искажений в изображении или неравномерности областей размытия.

Среди прочих подходов к неэталонной оценке уровня размытия в изображении можно привести использование алгоритмов деконволюции, направленных на выделение ядра искажения размытия и его применение в попытке устранить искажение. Преимуществом такого подхода является возможность получения ядра искажения высокой сложности в таких задачах, как устранение размытия, вызванного движением рук оператора, однако этот подход избыточен в случае, если устранение искажения не требуется. Существенным недостатком этого подхода также является крайне низкая устойчивость к случаям, когда в одном изображении в разных областях наблюдаются различные искажения размытия или их отсутствие [10, 11].

Следует отметить, что перечисленные методы были разработаны в рамках задачи оценки размытости изображений в общем случае и данные об

их эффективности по отношению к специфической задаче оценки размытости изображений лиц в открытом доступе отсутствуют.

1.3. Модели классификаторов

В общем случае, в рамках решения задачи классификации производятся следующие шаги:

1. Получение входного вектора признаков на основе исходной информации, описывающей классифицируемый объект.
2. Классификация, в ходе которой полученный вектор признаков передаётся определённому решающему правилу.

В данном подразделе перечислены наиболее распространённые виды моделей классификаторов и кратко приводятся их преимущества и недостатки.

1.3.1. Байесовский классификатор

Байесовский подход является одним из самых распространённых методов решения задач классификации. Используется следующее решающее правило:

$$h(x) = \arg \max_{y \in Y} p(y | x), \quad (1.6)$$

где $p(y | x)$ – апостериорная вероятность класса.

Эта вероятность, как правило, неизвестна, однако возможно вычислить приближение условной плотности $p(x | y)$. Используя теорему Байеса, можно переписать решающее правило в следующем виде:

$$h(x) = \arg \max_{y \in Y} p(y | x) = \arg \max_{y \in Y} \frac{p(x | y)p(y)}{p(x)} = \arg \max_{y \in Y} p(x | y) p(y). \quad (1.7)$$

Наконец, из решающего правила в общем виде можно вывести решающее правило наивного байесовского классификатора, для обучения

которого необходимо лишь оценить априорные вероятности классов $p(y)$ и условные плотности $p(x_i | y)$:

$$h(x) = \arg \max_{y \in Y} \prod_{i=1}^n p(f_i(x) = x_i | y) p(y), \quad (1.8)$$

где $f_1(x), \dots, f_n(x)$ – признаковые функции, описывающие объект x и принимающие значения x_1, \dots, x_n (предполагается, что значения функций – независимые случайные величины).

Пусть обучающая выборка состоит из L объектов x^1, \dots, x^L , каждый из которых имеет вид $x^j = [x_1^j, \dots, x_n^j]$. Обозначим через y^j значение целевой функции на объекте x^j . Поскольку пространство $Y = \{y_1, \dots, y_k\}$ в задачах классификации является конечным, априорная вероятность класса y_k можно оценить его частотой:

$$p(y_k) = \frac{|\{j | y^j = y_k\}|}{L}. \quad (1.9)$$

Пусть каждый из признаков f_i принимает значения на конечном множестве X_i . В этом случае условные вероятности $p(f_i(x) = x_i | y_k)$ можно оценить на основе частот $c(x_i, y_k) = |\{j | y^j = y_k, f_i(x^j) = x_i\}|$. Наиболее простой является формула

$$p(x_i | y_k) = \frac{c(x_i, y_k) + \alpha}{\sum_{z \in X_i} (c(z, y_k) + \alpha)}, \quad (1.10)$$

при которой условная вероятность $p(f_i(x) = x_i | y_k)$ пропорциональна величине $c(x_i, y_k) + \alpha$, где α – некоторое положительное число, добавляемое, чтобы избежать нулевых значений вероятности.

Основные преимущества байесовского подхода:

1. Решающее правило оптимально и может быть выписано в явном аналитическом виде, а также достаточно просто реализуется программно.

2. При классификации оцениваются априорные вероятности принадлежности объекта каждому классу. Эта информация может быть в дальнейшем использована, например, в задаче оценки рисков.

Недостатки подхода:

1. Условные плотности восстанавливаются на основе конечных выборок данных, из-за чего классификатор перестает быть оптимальным после их подстановки.
2. Нельзя выделить безусловно лучший метод восстановления плотности, из-за чего приходится подбирать метод эмпирически [12].

1.3.2. Метод опорных векторов

В рамках метода опорных векторов предлагается рассматривать каждое входное изображение в качестве точки в n -мерном пространстве, при этом n соответствует размерности входных данных (ими может быть общее число пикселей изображения или определённые признаки). Задача распознавания сводится к нахождению гиперплоскости в n -мерном пространстве, которая отделяла бы точки, соответствующие изображениям определённого класса, от всех остальных. Метод опорных векторов предполагает множество таких гиперплоскостей, поэтому целью ставится отыскание оптимальной гиперплоскости, расстояние от которой до ближайшей точки максимально.

Пусть входные данные представлены в следующем виде:

$$\{(x_0, y_0), (x_1, y_1), \dots, (x_m, y_m)\},$$

где x_i – входное изображение, а y_i – класс, к которому оно относится, представленный целым числом. Разделяющая гиперплоскость определяется параметром w – нормалью от точки к плоскости и задаётся уравнением $wx - b = 0$. Задача, таким образом, сводится к минимизации $\|w\|$. Согласно теореме Куна-Таккера [13], можно представить гиперплоскость в виде линейной комбинации векторов признаков обучающей выборки:

$$w = \sum_{i=1}^n a_i x_i y_i, \quad (1.11)$$

где a_i – множители Лагранжа. Таким образом, метод относится к разделу линейных опорных векторов. В случае, когда входные данные могут быть линейно разделены, можно использовать метод ядер, предложенный Б. Босером, И. Гийон и В. Вапником, который заключается в том, что объекты обучающей выборки отображаются в пространство более высокой размерности, при этом в новом пространстве выборка линейно разделима. Ядром классификатора называют выражение, определяющее отображение выборки из одного пространства в другое, на практике могут быть использованы следующие ядра: радиальная и гауссова базисные функции, полиномиальная функции, сигмоида и т.д. [14]

Модель на основе метода опорных векторов можно обучить различными методами квадратичного программирования, например, методом последовательной минимальной оптимизации [13].

В случае с обнаружением размытия изображения обучение модели происходит на основе выборки из изображений идеального качества и изображений с размытием. На основе каждого изображения строится гистограмма величин градиента, выраженная в виде вектора размерности n_1 , и гистограмма направлений градиента, выраженная в виде вектора размерности n_2 . Таким образом каждый образец обучающей выборки состоит из вектора входных данных размерности $n_1 + n_2$ и метки класса, к которому относится образец.

1.3.3. Свёрточные нейронные сети

Свёрточные нейронные сети были разработаны с целью получения инструмента, способного использовать локальные признаки изображений. Необходимость такой особенности заключается в свойстве стационарности изображений, заключающемся в том, что объекты могут по-разному располагаться в пространстве изображения, формируя множество инвариантных репрезентаций объекта. Каждая из этих репрезентаций по-прежнему должна была быть отнесена к одному и тому же классу.

Сама концепция использования локальных признаков подкрепляется свидетельствами в области нейробиологии. В работе Д. Хьюбела и Т. Визела [15] утверждается, что каждая клетка визуальной коры головного мозга проявляет чувствительность к определённому ограниченному участку поля зрения, называемому рецептивным полем. Рецептивные поля примыкают друг к другу, таким образом, перекрывая всё поле зрения, при этом соответствующие клетки работают как локальные фильтры, дающие отклик на присутствие в рецептивном поле определённых примитивных структур. Клетки с более широкими рецептивными полями также проявляют инвариантность к незначительным сдвигам объекта в поле зрения.

Изучение визуальной коры головного мозга привело к разработке моделей [16, 17, 18], которые эмулируют её поведение, а одной из самых успешных моделей в области распознавания изображений [19] стали свёрточные нейронные сети.

СНС являются модификацией многослойного перцептрона, включающей свёрточные слои, слои подвыборки или субдискретизации и полносвязные слои. Каждый нейрон какого-либо скрытого слоя СНС связан только с ограниченным локальным участком входного изображения, реализуя принцип локальной связности, что позволяет СНС использовать

преимущества входных данных двумерной структуры. На рисунке 1.2 представлена стандартная архитектура СНС (архитектура LeNet).

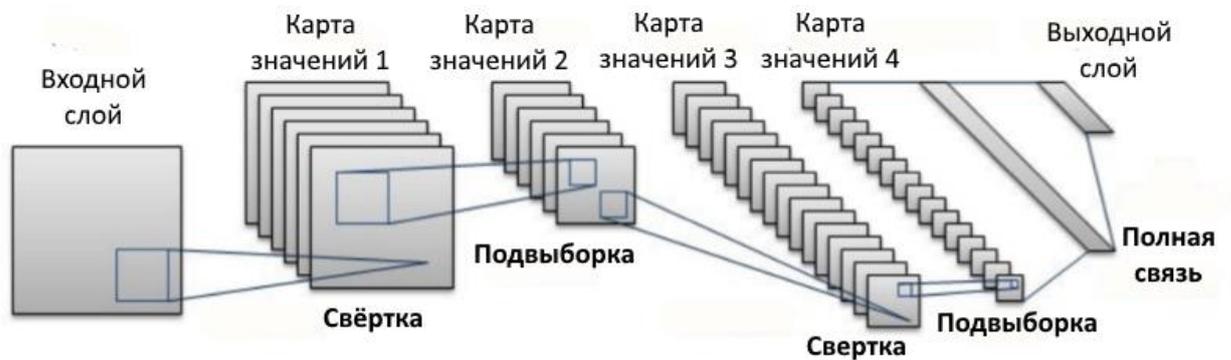


Рисунок 1.2. Архитектура сверточной нейронной сети

Основной особенностью СНС является использование разделяемых весов, что означает, что все нейроны определённого скрытого слоя обладают общим набором входных весов, который составляет один фильтр. При прямом проходе СНС производит операцию свёртки входного изображения набором таких фильтров, генерируя карты признаков – двумерные матрицы, содержащие результат свёртки конкретным фильтром [20].

В свою очередь, слои субдискретизации выполняют агрегацию значений, полученных в результате свёртки, группируя элементы карт признаков определённого размера. Субдискретизация предназначена для снижения вариативности данных, делая СНС более устойчивой к сдвигам локального признака в пределах определённого региона. Таким образом, сдвиг признака на некоторое значения в пределах окна субдискретизации не приведёт к изменению поведения соответствующего нейрона скрытого слоя, который инкапсулирует данный локальный признак. Для агрегации обычно используется функция вычисления среднего или максимального значения в пределах окна субдискретизации.

Таким образом, последовательное применение слоёв свёртки и субдискретизации позволяет СНС извлекать представления, независимые от расположения конкретных локальных признаков в пространстве изображения.

Объекты и наборы признаков, существенные для классификации, могут находиться в любой его части.

Благодаря операциям свёртки СНС обладают свойством инвариантности по отношению к таким аффинным преобразованиям, как сдвиг элементов изображения по вертикальной или горизонтальной оси, однако СНС не проявляют подобной устойчивости к прочим аффинным преобразованиям - вращению, масштабированию или отражению. Чтобы нивелировать этот недостаток, используются различные эвристические методы, такие как использование отражённых и повернутых копий исходных изображений или пирамид изображений в различном масштабе [21]. Для обработки объектов, расположенных в разных ориентациях в трёхмерном пространстве, модель должна быть обучена различающимся наборам признаков [22].

Частично опираясь на описанные принципы работы СНС, можно сформулировать основные преимущества и недостатки данной модели. Основными преимуществами СНС являются:

1. Использование механизма разделяемых весов, что значительно уменьшает количество параметров по сравнению с другими видами ИНС и ускоряет работу за счёт широких возможностей для распараллеливания процесса вычислений.
2. Относительная устойчивость к инвариантным преобразованиям сдвига наборов признаков в пространстве изображений за счёт использования слоёв субдискретизации.
3. СНС сочетают функцию классификатора с функцией выделения признаков путём получения фильтров свёртки в процессе обучения СНС. При помощи данных фильтров могут быть выделены оптимальные наборы признаков, сформулировать которые иным способом было бы невозможно или затруднительно. Данные наборы признаков также эффективно аппроксимируют структурную информацию, содержащуюся в изображении (то есть любую информацию,

касающуюся взаимного расположения признаков в пространстве изображения).

К недостаткам СНС в общем случае относятся:

1. Высокая продолжительность обучения методами обратного распространения ошибки.
2. Малая изученность математического аппарата СНС, что затрудняет разработку методических рекомендаций по подбору параметров СНС и обучающей модели. На сегодняшний день данные параметры приходится подбирать эмпирическим путём.

На основании описанных выше качественных преимуществ СНС было выдвинуто предположение, что аппарат СНС покажет высокую эффективность в решении задачи обнаружения размытых изображений лиц. С целью проверки данного предположения в рамках работы и была поставлена задача разработки метода обнаружения размытых изображений лиц на основе аппарата СНС, обладающего максимально возможной точностью в рамках ограничений, установленных условиями практического применения. Сравнительный анализ характеристик и общей эффективности каждой из представленных в данном разделе моделей классификаторов является предметом дальнейшей научно-исследовательской работы.

1.4. Модификации метода градиентного спуска

1.4.1. Пакетный и стохастический градиентный спуск

Для обучения ИНС, в том числе СНС, наиболее широко применяется метод градиентного спуска и его различные модификации. В теории ИНС метод градиентного спуска более известен как метод обратного распространения ошибки. Оптимизируемыми параметрами в данном случае служат весовые коэффициенты сети, а целью оптимизации является минимизация функции ошибки выходного слоя ИНС.

Первоначальный вариант, а именно метод пакетного градиентного спуска, предполагает подачу на вход сети каждого объекта обучающей выборки и расчёт функции ошибки и величины изменения (градиента) каждого весового коэффициента сети для каждого объекта. После этого рассчитанные градиенты суммируются, и только тогда происходит непосредственно модификация весовых коэффициентов. На практике большие размеры обучающей выборки приводят к тому, что обучение происходит крайне медленно, поэтому метод модифицируется таким образом, что изменение весовых коэффициентов производится после каждого расчёта градиентов на одном отдельном элементе обучающей выборки. Такой метод называется стохастическим градиентным спуском.

Стохастический градиентный спуск модифицирует веса сети W величиной градиента функции ошибки $\nabla J(W)$ со знаком минус. Коэффициент скорости обучения α служит весовым коэффициентом градиента. Таким образом, вектор весов ИНС модифицируется согласно формуле:

$$W_{t+1} = W_t + \Delta W_{t+1} = W_t - \alpha \nabla J(W_t). \quad (1.12)$$

Модификация весов ИНС в методе стохастического градиентного спуска необязательно должна происходить после подачи каждого конкретного объекта обучающей выборки. Она может также производиться после суммирования градиентов, рассчитанных на относительно небольшом ограниченном наборе объектов обучающей выборки (мини-батче), что приводит к снижению влияния шума обучающей выборки на результат обучения [23].

В ходе применения метода градиентного спуска в контексте ИНС функция ошибки каждого нейрона зависит от величины ошибки нейронов всех последующих слоёв сети. Такое количество степеней свободы приводит к возможным проблемам в ходе минимизации функции ошибки:

1. Сложная поверхность отклика целевой функции, в котором области сильной нелинейности могут чередоваться с широкими плато, что приводит к большим колебаниям градиента в зависимости от текущего положения значения оптимизируемой функции.
2. Неоднородное обновление параметров, так как входные данные могут содержать редко встречающиеся, но информативные признаки.
3. Необходимость подбора скорости обучения. Слишком малая скорость обучения может привести к остановке в локальных минимумах и в целом долгой сходимости, а слишком большая – к выходу из узкого глобального минимума.

Экспериментируя с формулами модификации весов сети, можно получить модификации метода градиентного спуска, обладающие определёнными преимуществами для преодоления перечисленных проблем.

1.4.2. Ускоренный градиентный спуск Нестерова

Ускоренный градиентный спуск Нестерова реализует две основные идеи для модификации метода градиентного спуска:

1. Использование импульса движения значения целевой функции в пространстве параметров. Продолжительное движение в определённом направлении позволяет предположить, что двигаться в данном направлении следует и дальше. Величина изменения каждого параметра, то есть весового коэффициента, таким образом, должна учитывать величину изменения на предыдущем шаге с некоторым весовым коэффициентом μ .
2. Заглядывание вперёд по вектору движения. Расчёт градиента происходит не в точке W в пространстве параметров, а в точке $W + \mu\Delta W_{t-1}$, в которую мы собираемся сместиться.

Формула изменения веса в методе ускоренного градиентного спуска Нестерова имеет следующий вид:

$$W_{t+1} = W_t + \mu \Delta W_{t-1} - \alpha \nabla J(W_t + \mu \Delta W_{t-1}). \quad (1.13)$$

Данный метод позволяет добиться значительного ускорения обучения, а также уменьшения влияния шума в обучающей выборке за счёт влияния предыдущих модификаций весов на последующие [24].

1.4.3. Adagrad

Adagrad (от англ. adaptive gradient) является модификацией метода градиентного спуска, направленной на обработку одновременно редких и информативных параметров.

Обновление i -того весового коэффициента в данном случае производится по следующей формуле:

$$(W_{t+1})_i = (W_t)_i - \alpha * \frac{(\nabla J(W_t))_i}{\sqrt{\sum_{t'=1}^t (\nabla J(W_{t'}))_i^2}}. \quad (1.14)$$

В данной формуле учитываются изменения веса на всех предыдущих итерациях обучения. Таким образом, весовые коэффициенты модифицируются с учётом того, насколько типичный признак они фиксируют, при этом наиболее часто обновляющиеся элементы модифицируются меньше. Другим достоинством метода является отсутствие необходимости тщательного подбора скорости обучения, так как скорость обучения будет затухать в ходе выполнения алгоритма [25].

1.4.4. RMSProp и AdaDelta

Алгоритмы RMSProp и AdaDelta являются модификациями алгоритма Adagrad, направленными на предотвращение паралича алгоритма в результате чрезмерного увеличения полной суммы обновлений веса на предыдущих итерациях.

В алгоритме RMSProp вместо полной суммы обновлений используется бегущее среднее предыдущих обновлений, рассчитываемой по формуле:

$$E((W_t)_i) = \delta E((W_{t-1})_i) + (1 - \delta)(\nabla J(W_t))_i^2. \quad (1.15)$$

Значение коэффициента δ обычно принимается равным 0.99. Формула для расчёта модификации веса приобретает вид:

$$(W_{t+1})_i = (W_t)_i - \alpha * \frac{(\nabla J(W_t))_i}{\sqrt{E((W_t)_i)}} [26]. \quad (1.16)$$

Алгоритм AdaDelta, в свою очередь, отличается тем, что использует стабилизирующий член RMS , пропорциональный ΔW_{t+1} . Это сделано для того, чтобы совпадали размерности W и ΔW . Обновление весов в данном случае происходит в три этапа: накопление квадрата градиента по формуле (1.15), обновление весового коэффициента и обновление величины $RMS(\Delta W)$, согласно следующим формулам:

$$(W_{t+1})_i = (W_t)_i - \alpha * \frac{RMS(\Delta W_{t-1})_i}{RMS(\nabla J(W_t))_i} (\nabla J(W_t))_i, \quad (1.17)$$

$$RMS(\nabla J(W_t))_i = \sqrt{E((W_t)_i) + \varepsilon}, \quad (1.18)$$

где ε – константа для обеспечения ненулевого значения RMS на первом шаге [27].

1.4.5. Adam

Алгоритм Adam является обобщением алгоритма AdaGrad, использующим одновременно идею инерционного движения значения целевой функции, как в методе ускоренного градиентного спуска Нестерова, так и идею более слабого обновления весовых коэффициентов для более типичных признаков. В алгоритме Adam обновление весовых коэффициентов происходит согласно следующим формулам:

$$(m_t)_i = \beta_1(m_{t-1})_i + (1 - \beta_1)(\nabla J(W_t))_i, \quad (1.19)$$

$$(\Delta W_t)_i = \beta_2(\Delta W_{t-1})_i + (1 - \beta_2)(\nabla J(W_t))_i, \quad (1.20)$$

$$(W_{t+1})_i = (W_t)_i - \alpha * \frac{\sqrt{1 - (\beta_2)_i^t}}{1 - (\beta_1)_i^t} * \frac{(m_t)_i}{\sqrt{(\Delta W_t)_i + \varepsilon}}. \quad (1.21)$$

Авторы работы рекомендуют использовать следующие значения коэффициентов: $\beta_1 = 0.9$, $\beta_2 = 0.999$ и $\varepsilon = 10^{-8}$ [28].

1.4.6. Выбор модификации метода градиентного спуска

Составим общую таблицу особенностей модификаций метода градиентного спуска из следующих: стохастический градиентный спуск (1), ускоренный градиентный спуск Нестерова (2), Adagrad (3), RMSProp/AdaDelta (4), Adam(5).

Таблица 1.1. Особенности модификаций метода градиентного спуска

Особенность	1	2	3	4	5
Идея инерционного движения значения целевой функции	-	+	-	-	+
Идея заглядывания вперёд по вектору движения	-	+	-	-	-
Уменьшение обновления типичных признаков	-	-	+	+	+
Устойчивость к параличу алгоритма	+	+	-	+	+

Для обучения СНС в данной работе был выбран алгоритм Adam. Алгоритм Adam является самой современной из представленных модификацией метода градиентного спуска, реализующей большую часть полезных концепций своих предшественников.

1.4.7. Батч-нормализация

Независимо от выбранного алгоритма оптимизации весовых коэффициентов СНС в процессе обучения целесообразно использовать процедуру батч-нормализации. Основное назначение батч-нормализации состоит в уменьшении ковариационного сдвига. Так называется эффект, при котором обучаемая модель замедляется в обучении в результате большого разнообразия обучающих данных.

В контексте ИНС этот эффект возникает во внутренних слоях сети. Каждый слой получает входные данные с выхода предыдущего слоя и может расцениваться как отдельная компонента обучающей системы. Так как весовые коэффициенты сети меняются в ходе обучения, то меняется и распределение на входные данные компонент [29]. Обучение будет сходиться быстрее, если произвести предварительную нормализацию и декорреляцию входных данных каждой компоненты обучающей системы. В случае применения батч-нормализации данные нормализуются перед входом большей части внутренних слоёв ИНС.

Приведём алгоритм батч-нормализации:

- Для нормализуемого слоя размерности D нормализуем k -ое измерение x по формуле:

$$\hat{x}^k = \frac{x^k - \mu_B}{\sigma_B^2}, \quad (1.22)$$

где μ_B и σ_B^2 – математическое ожидание и дисперсия для мини-батча $B = \{x_i^k, \dots, m\}$, рассчитываемые по следующим формулам:

$$\mu_B = \frac{1}{m} \sum_{i=1}^m x_i^k, \quad (1.23)$$

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^m (x_i^k - \mu_B)^2. \quad (1.24)$$

- Получаем новую нормализованную величину:

$$y^k = \gamma^k \hat{x}^k + \beta^k, \quad (1.25)$$

где параметры γ , β настраиваются в процессе обучения для каждой размерности.

- Добавляем итоговое преобразование в ИНС перед нелинейной функцией активации.

Таким образом, применение батч-нормализации:

- уменьшает ковариационный сдвиг внутренних слоёв ИНС, за счёт чего ускоряется обучение;
- позволяет менее тщательно подбирать параметр скорости обучения;
- не влияет на применимость метода обратного распространения ошибки, так как является дифференцируемым преобразованием [30].

1.4.8. Регуляризация

В области машинного обучения типичной трудностью при оптимизации параметров обучаемой модели является переобучение. Переобучение – это явление, когда в ходе обучения результаты классификации тестовой выборки, не пересекающейся с обучающей, начинают ухудшаться из-за чрезмерной адаптации модели к обучающей выборке. Таким образом, модель в процессе обучения со временем всё лучше классифицирует обучающую выборку, но хуже классифицирует объекты за её пределами, что понижает её эффективность в рамках практического применения.

Типичной мерой противодействия переобучению является применение регуляризации, то есть модификации оптимизируемой функции, в данном случае путём наложения штрафа за высокую сложность модели. Существуют два основных вида регуляризации:

1. L1-регуляризация или регуляризация Лассо, модифицирующая функцию ошибки по формуле:

$$L1 = J(W) - \frac{\lambda}{2} \sum_i (W)_i, \quad (1.26)$$

где λ – коэффициент редукции весов.

2. L2-регуляризация или регуляризация Тихонова, модифицирующая функцию ошибки по формуле:

$$L1 = J(W) - \frac{\lambda}{2} \sum_i (W)_i^2. \quad (1.27)$$

L1-регуляризация имеет свои преимущества, например, способность неявной селекции признаков за счёт возможности зануления признаком с малыми значения, однако на практике чаще применяется L2-регуляризация, так как она обеспечивает более гладкую регуляризацию и более универсальна в применении (L1-регуляризация как правило, хуже выполняет свои функции, когда речь идёт о градиентных методах оптимизации) [31].

1.5. Библиотеки глубокого обучения для реализации СНС

Для получения наилучших результатов при реализации СНС имеет смысл использовать одну из программных библиотек, реализующих алгоритмы глубокого обучения. Как правило, такие библиотеки свободно распространяемы для применения в практических задачах, хорошо оптимизированы, поддерживают вычисления на графических видеокартах (graphics processing unit, GPU), что особенно актуально для реализации СНС, обладающих широким потенциалом для параллельных вычислений.

В настоящее время существует большое количество программных библиотек в данной области, приведём наиболее широко используемые из них и их основные характеристики.

1.5.1. Torch

Torch – библиотека для научных вычислений с широкой поддержкой алгоритмов машинного обучения. Разрабатывается Idiap Research Institute, New York University и NEC Laboratories America. Библиотека Torch реализована на языке Lua, для ускорения вычислений при помощи GPU используется библиотека CUDA. На данный момент поддерживаются операционные системы Linux, FreeBSD, Mac OS X.

Torch позволяет создавать сложные нейросети с помощью механизма контейнеров. Контейнер — это класс, объединяющий объявленные компоненты нейросети в одну общую конфигурацию, которая в дальнейшем может быть передана в процедуру обучения. Компонентом нейросети могут быть не только полносвязные или сверточные слои, но и функции активации или ошибки, а также готовые контейнеры [32].

1.5.2. Theano

Theano — это расширение языка Python, позволяющее эффективно вычислять математические выражения, содержащие многомерные массивы. Theano разработана в лаборатории LISA для поддержки быстрой разработки алгоритмов машинного обучения.

Библиотека реализована на языке Python, поддерживается на операционных системах Windows, Linux и Mac OS. В состав Theano входит компилятор, который переводит математические выражения, написанные на языке Python в эффективный код на C или CUDA.

Theano предоставляет базовый набор инструментов для конфигурации нейросетей и их обучения. Возможна реализация многослойных полностью

связанных ИНС, СНС, рекуррентных нейронных сетей, автокодировщиков и ограниченных машин Больцмана. Также предусмотрены различные функции активации, в частности, сигмоидальная, softmax-функция, кросс-энтропия. В ходе обучения используется пакетный градиентный спуск [33].

1.5.3. Pylearn2

Pylearn2 – библиотека глубокого обучения, реализованная на языке Python и являющаяся дополнением библиотеки Theano. Библиотека разрабатывается с 2011 года в университете Монреаля и распространяется под лицензией BSD 3-Clause. Поддерживается операционная система Linux. Для ускорения вычислений используется библиотека Cuda-convnet, реализованная на CUDA/C++, что обеспечивает эффективное использование GPU.

Pylearn2 позволяет создавать полностью связанные ИНС, СНС, автокодировщики и ограниченные машины Больцмана, поддерживает различные виды функций ошибки и несколько методов обучения, среди которых пакетный и стохастический градиентный спуск, а также нелинейный метод сопряжённых градиентов. Для хранения структур ИНС в Pylearn2 используются файлы конфигурации в формате YAML, являющимся быстрым и удобным инструментом для сериализации объектов [34].

1.5.4. Caffe

Библиотека Caffe разработана на языке C++, поддерживает операционные системы Linux и OS X, а также Windows. Библиотека была разработана в калифорнийском университете в Беркли в 2013 году, поддерживается Центром Зрения и Обучения Беркли и распространяется под лицензией BSD 2-Clause. Для векторных и матричных вычислений Caffe использует библиотеку BLAS, а также использует библиотеку cuDNN на основе технологии CUDA для ускорения вычислений на GPU.

Библиотека позволяет создавать полностью связанные ИНС и СНС, поддерживает широкий набор функций ошибки и позволяет использовать для обучения стохастический градиентный спуск, алгоритм AdaGrad и ускоренный градиентный спуск Нестерова. Для хранения структур ИНС, исходных данных и параметров обучения в библиотеке Caffe используются файлы конфигурации в формате prototxt [35].

1.5.5. Dlib-ml

Библиотека глубокого обучения Dlib-ml разрабатывается с 2009 г. и с тех пор активно развивается (последнее обновление в марте 2017 г). Она представляет из себя свободно распространяемую кроссплатформенную библиотеку, написанную на языке C++. Dlib-ml предоставляет те же возможности применения алгоритмов глубокого обучения, в том числе проектирования и применения СНС, что и описанные выше библиотеки, но помимо этого имеет ряд преимуществ:

- Удобный и доступный API, ориентированный на использование в коммерческой разработке на языке C++.
- Использование шаблонных классов для определения слоёв нейронных сетей, что делает процесс проектирования архитектуры нейронной сети более простым, а также ускоряет выполнение программы.

Библиотека Dlib-ml поддерживает вычисления на GPU при помощи библиотеки CUDA. Для обучения СНС применяется алгоритм Adam или метод стохастического градиентного спуска [36].

1.5.6. Выбор библиотеки глубокого обучения

В таблице 1.2 представлены критерии выбора библиотеки глубокого обучения и соответствие им каждой из рассмотренных библиотек.

Таблица 1.2. Таблица соответствия библиотек критериям выбора

Критерий	Torch	Theano	Pylearn2	Caffe	Dlib-ml
Поддержка ОС Windows	-	+	+	+	+
Реализация на языке C++	-	-	-	+	+
Поддержка вычислений на GPU	+	+	+	+	+
Наличие актуальной документации	+	+	-	+	+

Библиотека Dlib-ml имеет следующие преимущества по сравнению с библиотекой Caffe:

- работает в 1.6 раз быстрее при использовании вычислений на GPU [24];
- имеет меньшее количество зависимостей, что ускоряет работу и компиляцию программ;
- не хранит структуру нейросетей во внешних файлах, слои сети библиотеки Dlib-ml определены в виде шаблонных классов и задаются внутри программы.

В связи с этим была выбрана библиотека Dlib-ml.

1.6. Модификации архитектуры СНС

1.6.1. Архитектура LeNet

Архитектура LeNet была предложена Я. ЛеКуном в 1998 году в одной из первых работ, посвящённых практической разработке и применению СНС. В данной работе была продемонстрирована высокая эффективность СНС данной архитектуры в задаче распознавания рукописных символов, обучение и тестирование производилось при помощи базы изображений MNIST [37].

Работой, которая принесла свёрточным сетям настоящую популярность, стала работа А. Крижевского, И. Сатскевера и Д. Хинтона, опубликованная в 2012 году. В ней авторами была предложена модификация СНС архитектуры LeNet, содержащая 5 слоёв свёртки, 3 слоя субдискретизации и 3 полносвязных слоя. Назначением СНС было участие в состязании ILSVRC12 (ImageNet Large-Scale Visual Recognition Challenge 2012), участники которого использовали различные алгоритмы машинного зрения для классификации объектов на изображениях среди 1000 категорий. СНС, получившая название AlexNet, выиграла состязание, показав ошибку классификации по методу топ-5, равную 15.4% (последующий участник показал результат в 26.2%), что стало настоящим прорывом в области классификации изображений и обусловило высокий интерес к СНС в последующие годы [19].

Архитектура LeNet предполагает прямолинейное и последовательное расположение слоёв свёртки и субдискретизации с завершением из нескольких полносвязных слоёв, интерпретирующих результаты извлечения признаков. В своём первоначальном виде, предложенном ЛеКуном, СНС данной архитектуры состоит из семи слоёв: входного слоя, двух пар из слоёв свёртки и субдискретизации и двух полносвязных слоёв [37]. Именно эта архитектура представлена на рисунке 1.2 пункта 1.3.3 данной работы. Модификации архитектуры LeNet производятся путём:

- изменения количества слоёв свёртки и субдискретизации;
- изменения количества и размера фильтров слоёв свёртки;
- изменения размера и шага областей субдискретизации;
- изменения количества и размера полносвязных слоёв;
- добавления слоёв нормализации;
- изменения нелинейной функции активации нейронов.

Модификации СНС данной архитектуры можно использовать в любых задачах оценки и классификации, связанных с обработкой локальных признаков изображений.

1.6.2. Архитектура ResNet

СНС архитектуры ResNet была разработана отделением Microsoft Research Asia в 2015 году. Сверхглубокая сеть из 152 последовательных слоёв установила новые рекорды в задачах классификации, обнаружения и локализации объектов на изображениях. Модуль архитектуры ResNet представлен на рисунке 1.3.

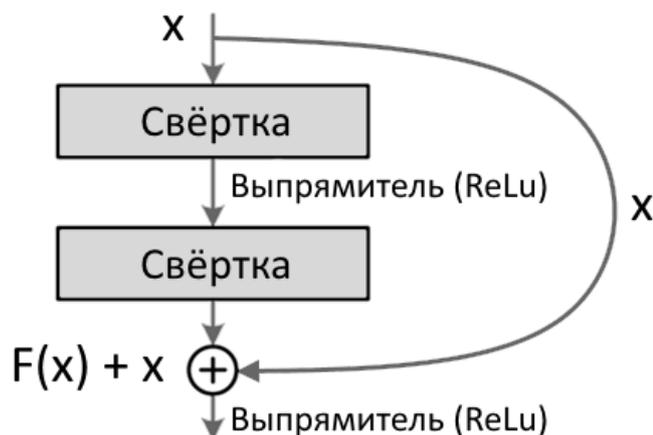


Рисунок 1.3. Модуль архитектуры ResNet

Дополнительная связь, передающая входной вектор на выход модуля из нескольких свёрточных слоёв, где происходит конкатенация входных данных и результата обработки слоями свёртки, обеспечивает получение на выходе модификации входных данных взамен нового представления, практически не содержащего информации о данных на предыдущих этапах, как это происходит в традиционных СНС. Благодаря этому достигается ускорение и качественное улучшение процесса обучения СНС за счёт снижения затухания градиента, что особенно актуально для глубоких СНС. Microsoft ResNet стала победителем ILSVRC15, показав ошибку классификации по методу топ-5, равную 3.6% [38].

1.6.3. Архитектура Inception

Наиболее прямолинейным методом увеличения точности распознавания при помощи СНС является увеличение размера их структуры, что

подразумевает как увеличение в глубину (количества слоёв), так и в ширину (количество вычислительных единиц в каждом слое). Однако увеличивающийся размер сети подразумевает также увеличение количества параметров и вероятности переобучения, особенно, если обучающая выборка невелика. Другим недостатком этого подхода является значительное увеличение затрачиваемых вычислительных ресурсов. К примеру, если два свёрточных слоя глубокой СНС для визуального распознавания связаны между собой, любое равномерное увеличение количества фильтров приведёт к квадратичному увеличению затрат ресурсов.

Фундаментальным решением проблемы может быть переход от полносвязных архитектур к распределённо связанным архитектурам, в том числе в отношении свёрточных слоёв. Основным результатом работы [39] говорит о том, что если вероятностное распределение данных можно представить глубокой распределённой нейронной сетью, то оптимальная архитектура может быть спроектирована слой за слоем путём анализа корреляционной статистики между активациями нейронов последнего слоя и связующих нейронов, между выходами которых существует устойчивая связь.

Применяя данный подход, команда исследователей из корпорации Google, совместно с сотрудниками Университета Северной Каролины и Мичиганского университета, создали своё воплощение глубокой свёрточной нейронной сети с архитектурой под названием Inception. Конечная реализация сети, принимавшая участие в соревновании по детектированию и классификации объектов ILSVRC14, получила название GoogLeNet и показала лучший результат по показателю ошибки классификации по методу топ-5, равный 6.66% (вариант сети, оптимизированный для классификации и обученный на обучающих наборах ILSVRC12), опередив каскад СНС под названием VGG с результатом, равным 7.33%. Основная идея, на которой базируется архитектура Inception, заключается в нахождении оптимальной локально распределённой структуры для свёрточной нейронной сети, которая

могла бы быть аппроксимирована и реализована при помощи современных методов для плотных вычислений [40].

Модуль архитектуры Inception представлен на рисунке 1.4.

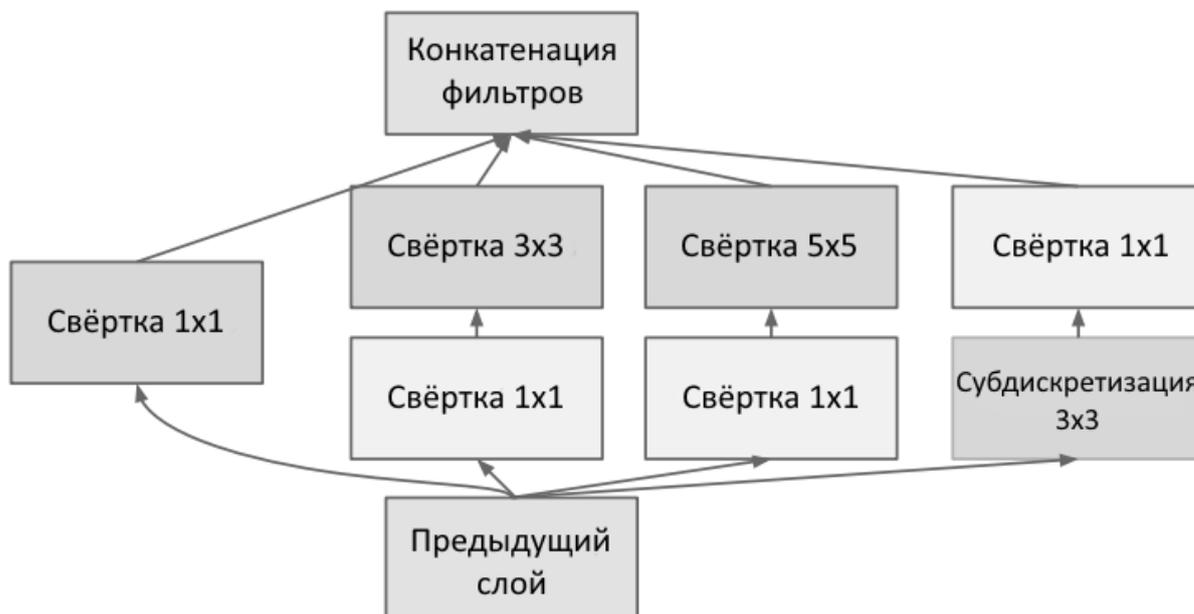


Рисунок 1.4. Модуль архитектуры Inception с использованием 1x1 свёрточных слоёв для сокращения размерности

2. ПРОЕКТИРОВАНИЕ И ОБУЧЕНИЕ СНС

2.1.1. Выборка исходных изображений

Формирование множества обучающих данных достаточного количества и качества имеет важное значение в задачах машинного обучения. Неправильно сформированная или недостаточного размера обучающая выборка может свести на нет всю эффективность алгоритмов обучения, в связи с чем среди специалистов в области машинного обучения принято считать наличие хороших обучающих данных одним из самых важных условий [41].

Для успешного обучения СНС необходимо было сформировать максимально большую обучающую выборку, разделённую на части по количеству классов. В рамках данной работы решалась задача бинарной классификации, поэтому заполнению подлежали два класса изображений лиц – с высокой и низкой степенью размытия. Для формирования обучающей выборки были использованы инструменты библиотек OpenCV и Dlib-ml для предобработки изображений лиц и метод экспертной оценки для их сортировки.

На первом этапе работы были получены видеофайлы, на которых было заснято большое количество лиц в сложных условиях (для этого использовались камеры наблюдения, установленные в общественных местах с большим потоком людей). Была написана программа для обнаружения и кадрирования лиц при помощи библиотеки OpenCV. После обработки полученных видеофайлов было получено более 200000 изображений лиц различного качества.

В данной работе предложен и реализован следующий алгоритм предобработки изображений лиц:

1. масштабирование и центрирование относительно расположения глаз, обнаружение ключевых точек лица производилось средствами библиотеки Dlib-ml;
2. преобразование в оттенки серого;
3. эквализация гистограммы.

Каждое изображение было вручную рассортировано на четыре класса:

1. изображения лиц с высокой степенью размытия;
2. изображения лиц без размытия;
3. изображения лиц с промежуточной степенью размытия;
4. бракованные изображения (засвеченные или затемнённые лица с низкой величиной градиента яркости, лица, перекрытые другими объектами и т.д.).

В результате сортировки была сформирована обучающая выборка из 35000 изображений лиц с высокой степенью размытия и 75000 изображений лиц без размытия. Примеры изображений каждого из четырёх классов представлены на рисунке 2.1.

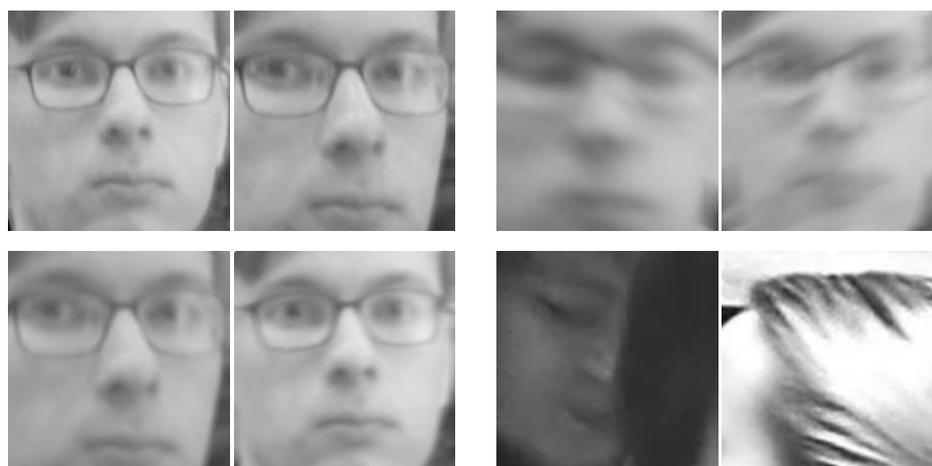


Рисунок 2.1. Примеры изображений лиц без размытия (вверху слева), с высокой степенью размытия (вверху справа), с промежуточной степенью размытия (внизу слева) и бракованных изображений (внизу справа)

2.1.2. Варианты входного вектора признаков

В качестве первого варианта вектора признаков для обучения и практического применения СНС были рассмотрены интенсивности пикселей исходного изображения в оттенках серого. Изображения лиц являются специфичными, то есть содержат определённые структурные закономерности в определённых областях, которые различаются в изображениях с различной степенью размытия. СНС ориентированы на извлечение подобных структурных зависимостей в двумерных входных данных, что позволяет предположить эффективность использования исходных изображений в качестве векторов признаков.

В качестве других вариантов входного вектора признаков были рассмотрены изображения, полученные на основе частотного спектра исходных изображений. Поводом к этому послужило предположение, что СНС будет способна извлечь сложные нелинейные зависимости, в том числе пространственные, из изображений частотного спектра таким же образом, как и из исходных изображений. В отличие от других методов оценки качества на основе анализа частотного спектра, в которых способ получения векторов признаков изначально задан авторами метода, в данном случае большую часть работы по извлечению признаков берёт на себя СНС, формирующая карты признаков полученными в процессе обучения фильтрами.

Способность СНС извлекать нелинейные зависимости между пикселями изображения порождает дополнительное требование к выбору вариантов входного вектора признаков. Вектора признаков должны подбираться таким образом, чтобы содержащаяся в одних векторах информация не могла бы быть с малыми затратами извлечена СНС в ходе обучения с использованием других векторов признаков. Использование исходных изображений и частотного спектра изображений удовлетворяет этому правилу.

Таким образом, было рассмотрено три варианта входного вектора признаков:

1. Вектор признаков на основе исходного изображения.
2. Вектор признаков на основе модульных составляющих частотного спектра.
3. Вектор признаков на основе фазовых составляющих частотного спектра.

Вторая обучающая выборка была составлена на основе модульных составляющих частотного спектра исходных изображений. Массив модульных составляющих может быть легко визуализирован в виде изображения. На рисунке 2.2 приведены примеры входных данных второй выборки.

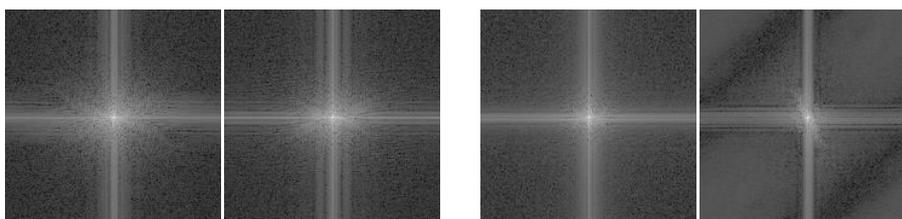


Рисунок 2.2. Примеры изображений модульных составляющих частотного спектра изображений без размытия (слева) и с высокой степенью размытия (справа)

Третья обучающая выборка была составлена на основе фазовых составляющих частотного спектра исходных изображений. На рисунке 2.3 приведены примеры входных данных третьей выборки.

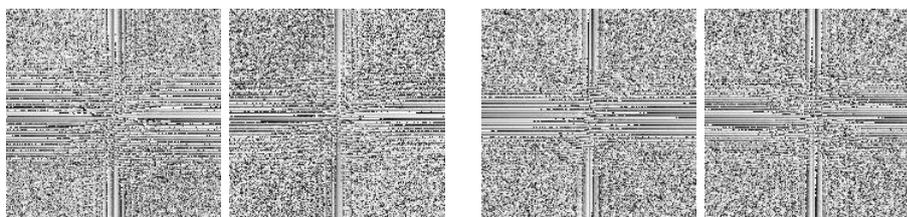


Рисунок 2.3. Примеры изображений фазовых составляющих частотного спектра изображений без размытия (слева) и с высокой степенью размытия (справа)

2.1.3. Тестовая выборка

Была сформирована тестовая выборка из изображений четырёх различных источников с разными условиями получения входных данных. Изображения тестовой выборки должны быть максимально разнообразными и не входить в обучающую выборку, чтобы можно было судить о степени переобучения сети, опираясь на результаты тестирования.

В общей сложности тестовая выборка для каждого варианта вектора признаков состоит из 3300 изображений. Тестирование производится путём классификации каждого входного вектора тестовой выборки, после чего количество правильных классификаций делится на их общее количество для получения точности обученной модели.

2.2. Подбор параметров обучения СНС

2.2.1. Процедура подбора параметров обучения

Процедура подбора параметров обучения СНС состояла из следующих этапов:

1. Подбор начальной скорости обучения. Начальная скорость обучения является самым значимым параметром обучения, так как определяет скорость движения значения целевой функции по поверхности отклика. Слишком низкое значение скорости обучения может привести к невозможности выхода решения из области, сходящейся к одному локальному минимуму, значительно снижая пространство поиска. Слишком высокая скорость обучения может привести к случайному переходу решения из области, сходящейся к более оптимальному локальному минимуму, к области, сходящейся к менее оптимальному.
2. Подбор параметра регуляризации. Модификация изменения весов уменьшает влияние переобучения СНС и, помимо этого, замедляет

обучение. На предыдущем этапе данный параметр устанавливался равным нулю.

3. Уточняющий подбор начальной скорости обучения. Изменение параметра регуляризации может привести к изменению поведения обучающей модели таким образом, что подобранное значение начальной скорости обучения перестанет быть оптимальным, поэтому после подбора параметра регуляризации производится проверка ближайших значений скорости обучения.
4. Подбор предельного количества итераций без прогресса. Данный параметр используется для определения момента уменьшения скорости обучения. Увеличение предельного количества итераций без прогресса, таким образом, удлиняет обучение, давая обучающей модели время на более тщательный поиск оптимального решения. Однако чрезмерно высокое значение данного параметра может привести к неоправданно высокой продолжительности обучения или вовсе к потере свойства сходимости алгоритмом обучения. На предыдущих этапах данный параметр устанавливался равным 300.

На каждом этапе были определены границы и значения параметров для эмпирической оценки их эффективности. Для каждого значения параметра при прочих параметрах, установленных по умолчанию или в значения, подобранные на предыдущих этапах, было проведено 10 процедур обучения сети, после чего были вычислены средние значения и стандартное отклонение результатов тестирования сетей на тестовой выборке. Получение этих данных необходимо в связи с тем, что каждая процедура обучения СНС начинается с инициализации параметров СНС (весовых коэффициентов) случайными значениями, и, таким образом, в начале обучения положение значения целевой функции в пространстве параметров СНС оказывается различным. Значение стандартного отклонения для определённого набора параметров обучения позволяет судить об относительной стабильности получения результатов обучения независимо от начального положения значения целевой функции.

В общей сложности на данном этапе было проведено 690 полных процедур обучения СНС.

2.2.2. Подбор параметров обучения для выборки исходных изображений

На рисунке 2.4 приведены диаграммы результатов подбора начальной скорости обучения СНС на основе выборки исходных изображений. В этом и последующих пунктах более тёмным цветом выделены столбцы, соответствующие СНС с наибольшей точностью.

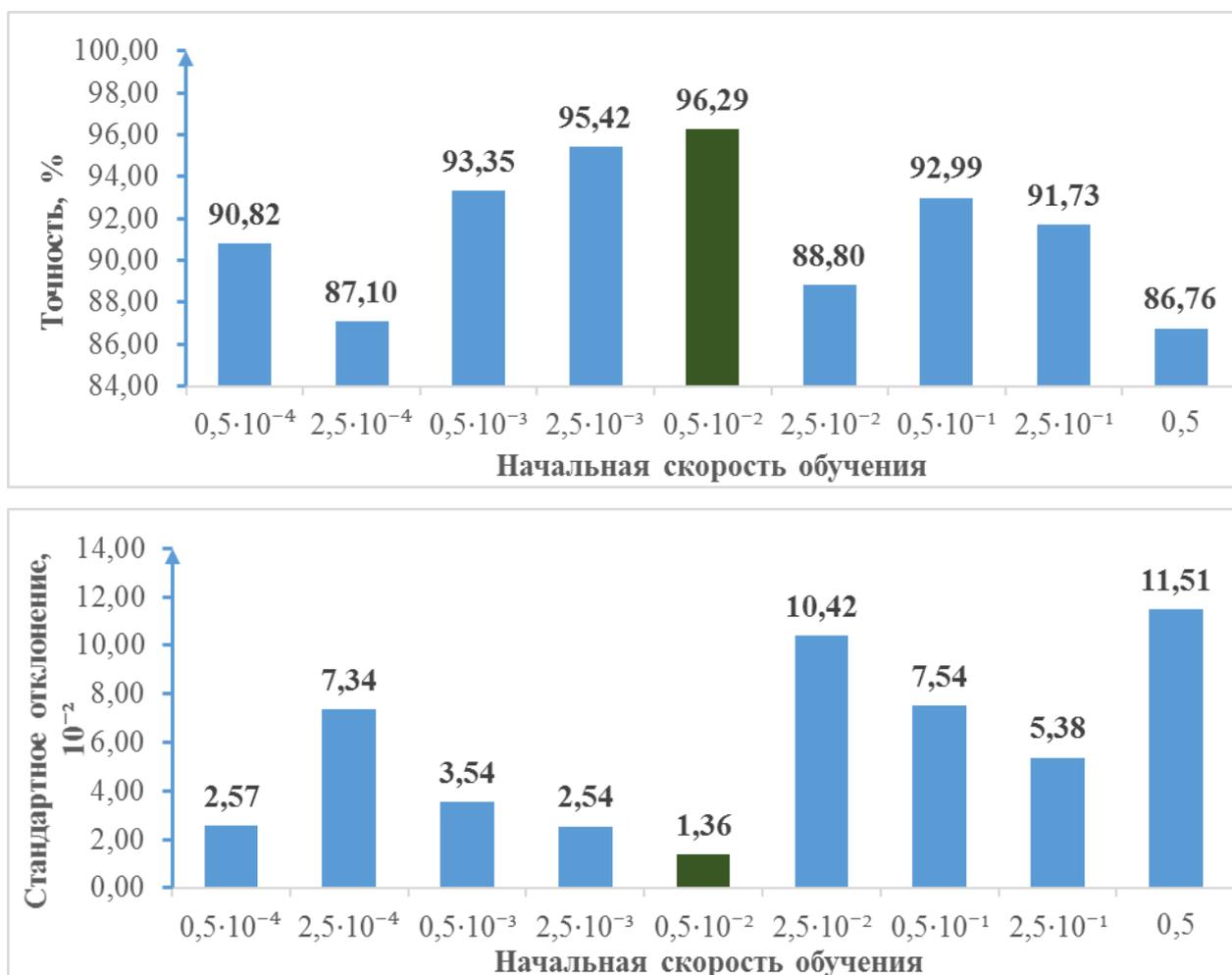


Рисунок 2.4. Среднее значение и стандартное отклонение точности классификации при различных значениях начальной скорости обучения для выборки исходных изображений

Таким образом, оптимальным значением начальной скорости обучения для выборки исходных изображений является значение, равное $0.5 \cdot 10^{-2}$. При

более низком или высоком значении точность классификации начинает падать, что говорит о попадании в менее оптимальные локальные минимумы поверхности отклика целевой функции в результате невозможности выхода из локального минимума или выхода из глобального минимума в результате слишком большой скорости обучения. Резкие уменьшения средней точности при некоторых значениях параметра коррелируют с высокими значениями стандартного отклонения, что говорит о наличии особенно низких значений точности, полученных в рамках обучений с данным значением параметра.

На рисунке 2.5 приведены диаграммы результатов подбора параметра регуляризации.

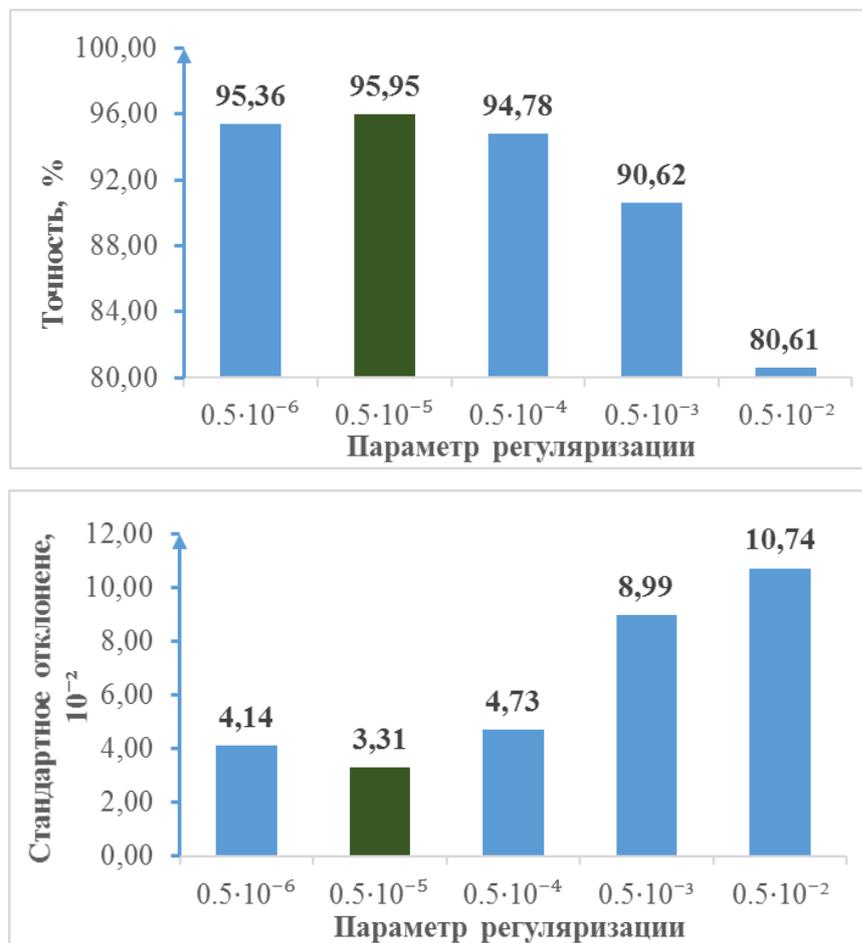


Рисунок 2.5. Среднее значение и стандартное отклонение точности классификации при различных значениях параметра регуляризации для выборки исходных изображений

На основе результатов подбора параметра регуляризации можно сделать вывод, что влияние переобучения для данной модели классификатора и данного размера обучающей выборки в целом невелико, поскольку точность падает незначительно при очень малом значении параметра ($0.5 \cdot 10^{-6}$). Оптимальным является значение, равное $0.5 \cdot 10^{-5}$. При увеличении параметра до $0.5 \cdot 10^{-4}$ и выше точность классификации начинает стремительно уменьшаться, так как параметр регуляризации излишне замедляет движение значения целевой функции в сторону уменьшения, не позволяя двигаться к более оптимальным локальным минимумам или глобальному минимуму.

На рисунке 2.6 приведены диаграммы результатов уточнения подбора начальной скорости обучения.

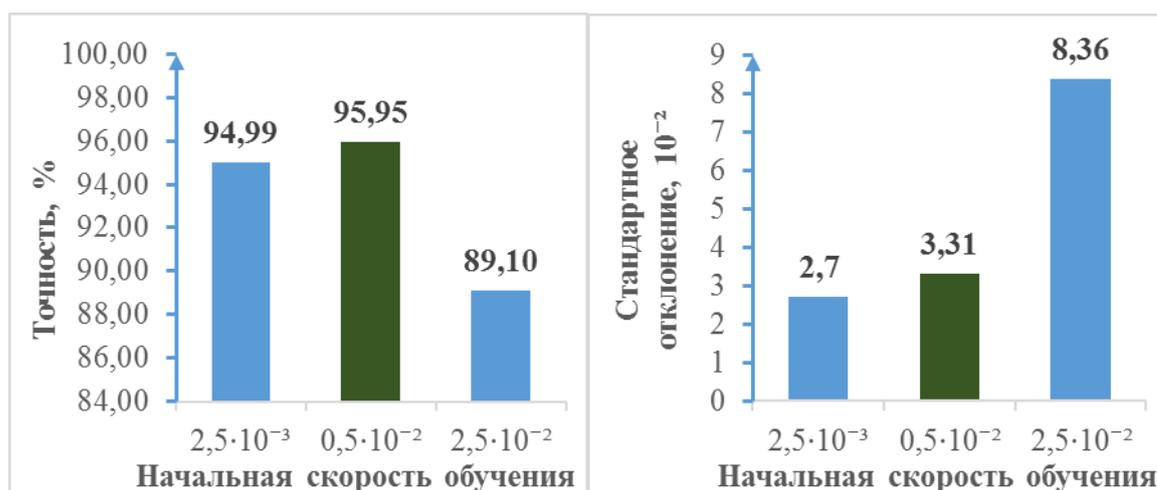


Рисунок 2.6. Среднее значение и стандартное отклонение точности классификации при различных значениях начальной скорости обучения для выборки исходных изображений (уточняющий этап)

Исходя из полученных данных, можно сделать вывод, что оптимальное значение параметра начальной скорости обучения не изменилось.

На рисунке 2.7 приведены диаграммы результатов подбора предельного количества итераций без прогресса.

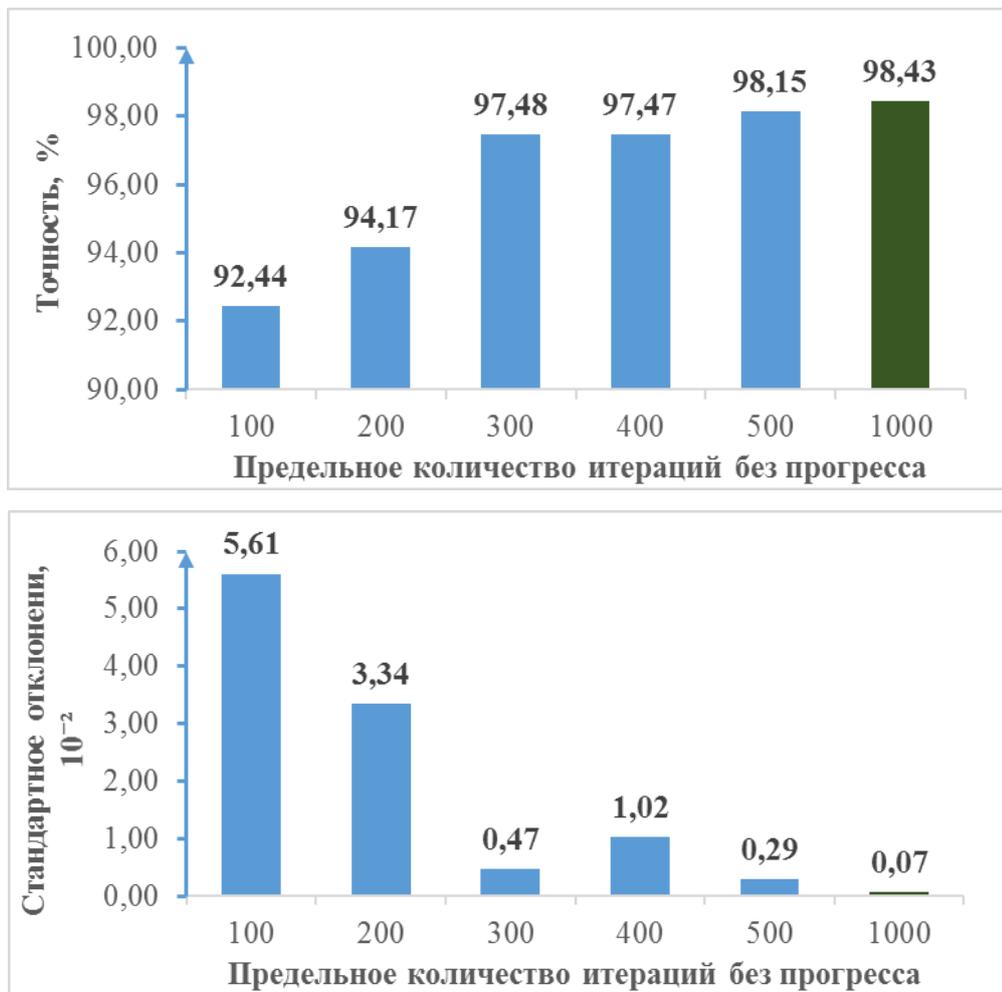


Рисунок 2.7. Среднее значение и стандартное отклонение точности классификации при различных значениях предельного количества итераций без прогресса

Наиболее оптимальным значением предельного количества итераций без прогресса из диапазона вариантов оказалось значение, равное 1000. Это значение является максимальным среди выбранных вариантов, и, возможно, дальнейшее увеличение параметра может привести к лучшим результатам, однако это приведёт к многократно более продолжительному времени обучения СНС, что делает невозможным получение указанных результатов в рамках временных ограничений данной работы.

Относительно низкое значение стандартного отклонения для оптимальных значений каждого параметра говорит о стабильности получения указанных результатов.

2.2.3. Подбор параметров обучения для выборки модульных составляющих частотного спектра

На рисунке 2.8 приведены диаграммы результатов подбора начальной скорости обучения СНС на основе выборки модульных составляющих частотного спектра.

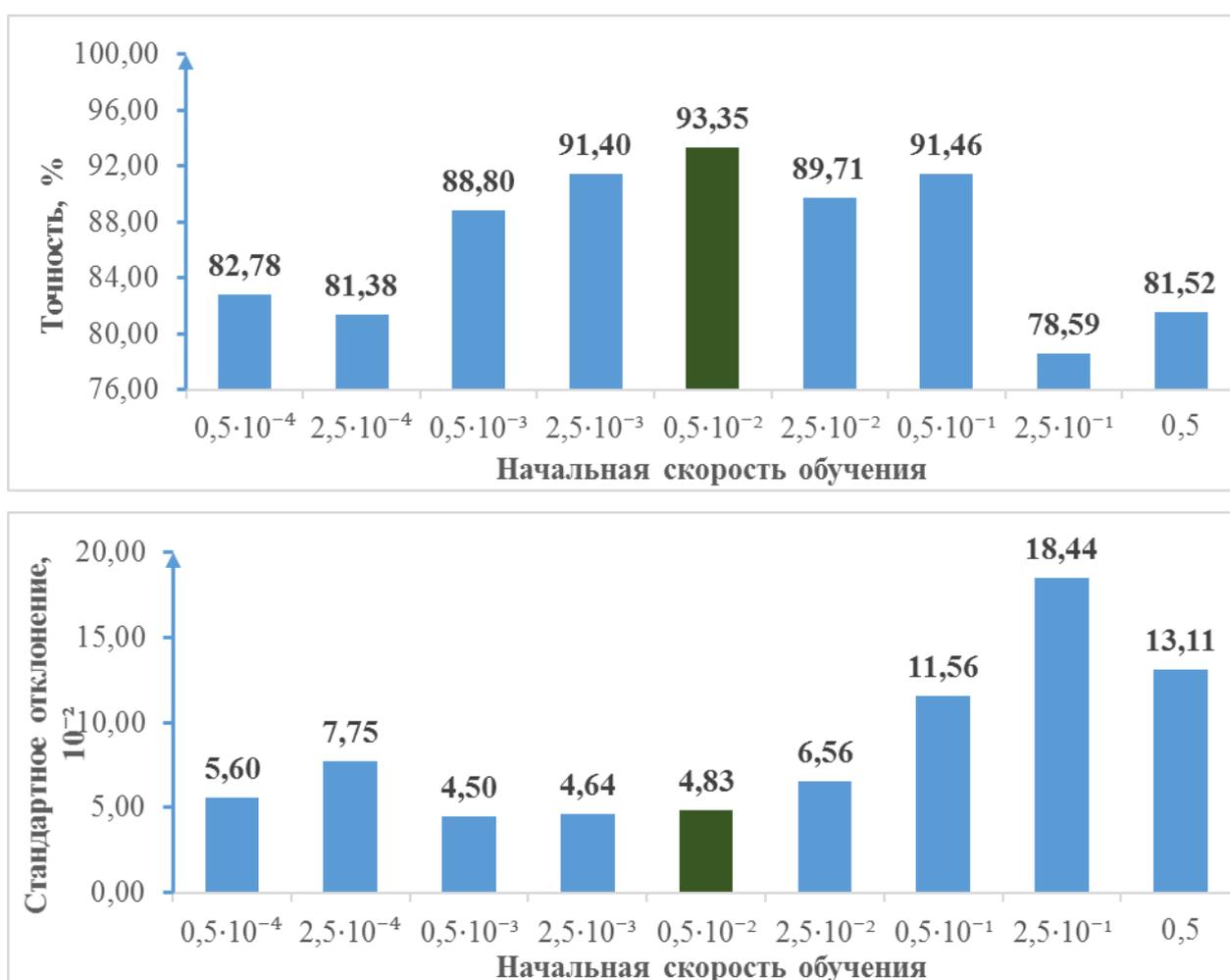


Рисунок 2.8. Среднее значение и стандартное отклонение точности классификации при различных значениях начальной скорости обучения для выборки модульных составляющих частотного спектра

Оптимальным значением начальной скорости обучения для выборки модульных составляющих частотного спектра является значение, равное $0,5 \cdot 10^{-2}$, так же, как и для выборки исходных изображений.

На рисунке 2.9 приведены диаграммы результатов подбора параметра регуляризации.

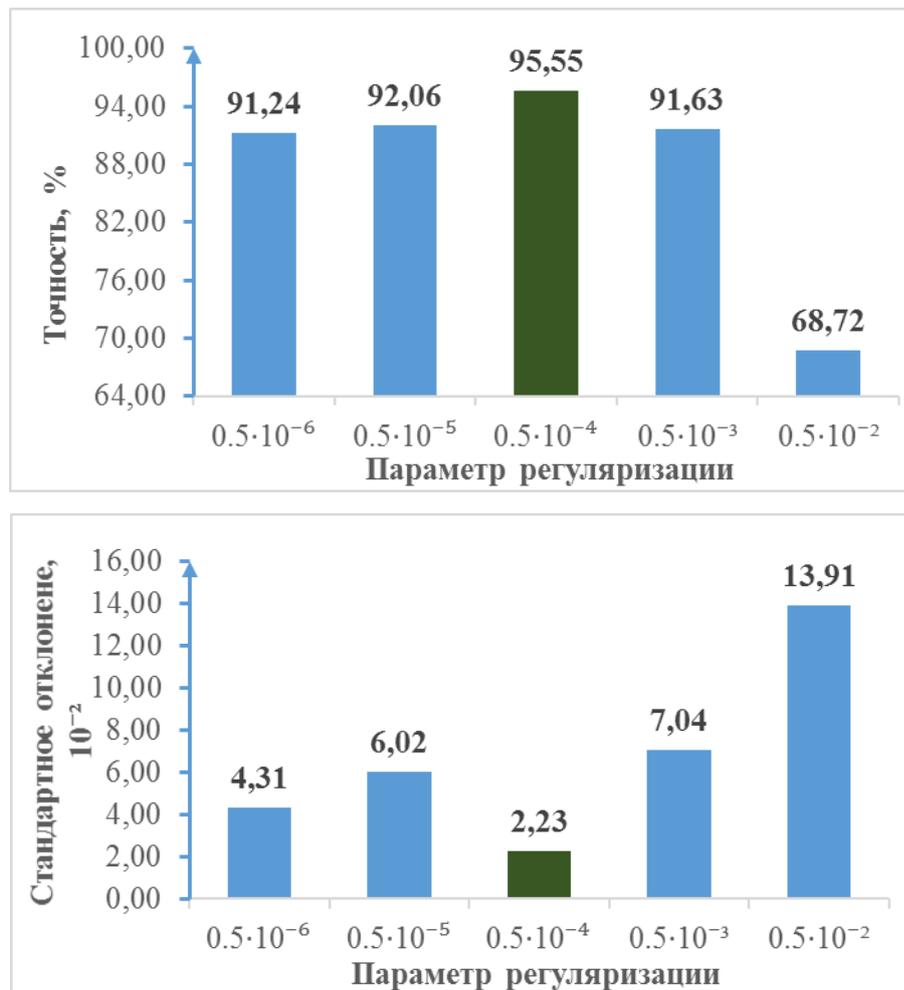


Рисунок 2.9. Среднее значение и стандартное отклонение точности классификации при различных значениях параметра регуляризации для выборки модульных составляющих частотного спектра

Из приведённых диаграмм видно, что оптимальное значение параметра регуляризации для данной выборки составило $0.5 \cdot 10^{-4}$, что говорит о большей подверженности модели переобучению при использовании данного входного вектора признаков.

На рисунке 2.10 приведены диаграммы результатов уточнения подбора начальной скорости обучения.

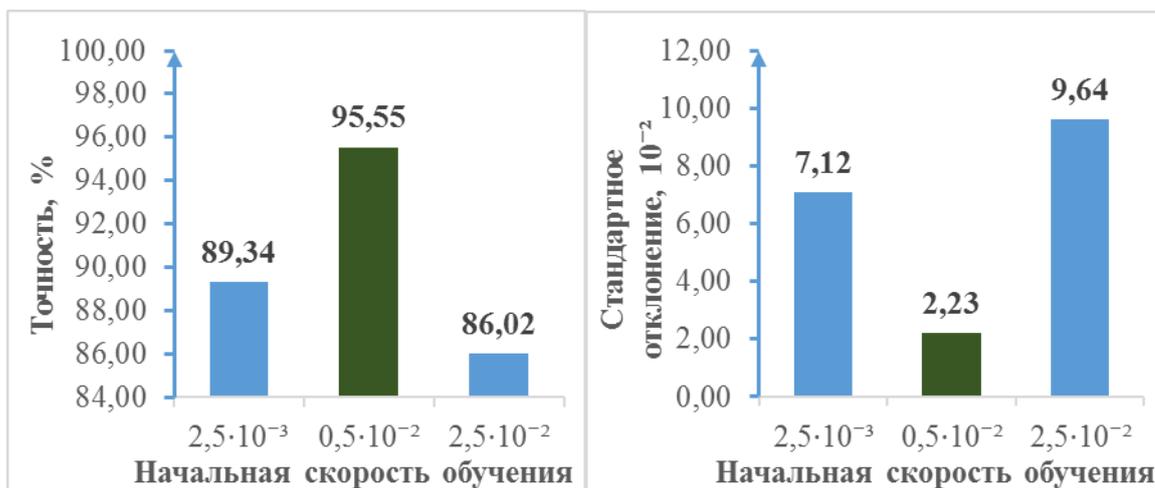


Рисунок 2.10. Среднее значение и стандартное отклонение точности классификации при различных значениях начальной скорости обучения для выборки модульных составляющих частотного спектра (уточняющий этап)

Исходя из полученных данных, можно сделать вывод, что оптимальное значение параметра начальной скорости обучения не изменилось.

На рисунке 2.11 приведены диаграммы результатов подбора предельного количества итераций без прогресса.

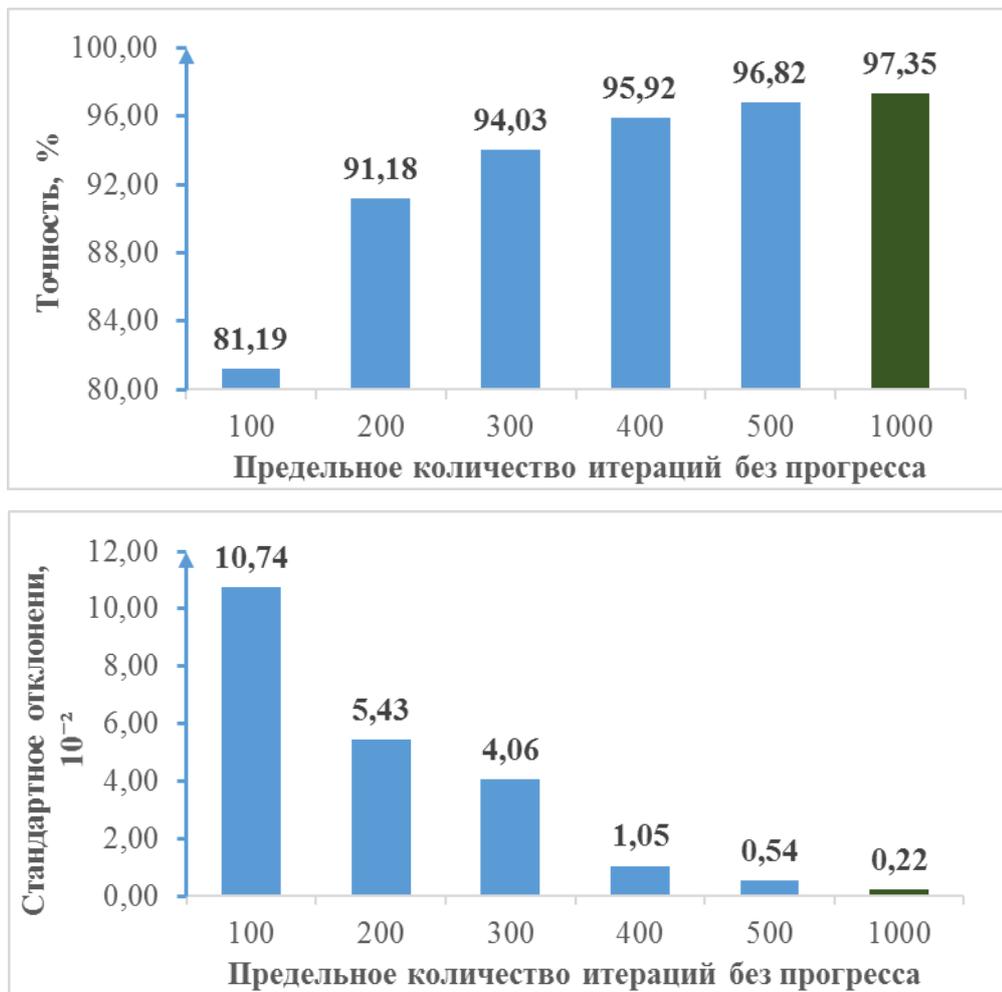


Рисунок 2.11. Среднее значение и стандартное отклонение точности классификации при различных значениях предельного количества итераций без прогресса для выборки модульных составляющих частотного спектра

Как и в случае выборки исходных изображений, лучшим оказалось максимальное значение предельного количества итераций из диапазона вариантов. По указанным в предыдущем разделе причинам получение результатов для более высоких значений параметра невозможно, поэтому в качестве конечного результата подбора данного параметра было принято значение, равное 1000.

2.2.4. Подбор параметров обучения для выборки фазовых составляющих частотного спектра

На рисунке 2.12 приведены диаграммы результатов подбора начальной скорости обучения СНС на основе выборки фазовых составляющих частотного спектра.

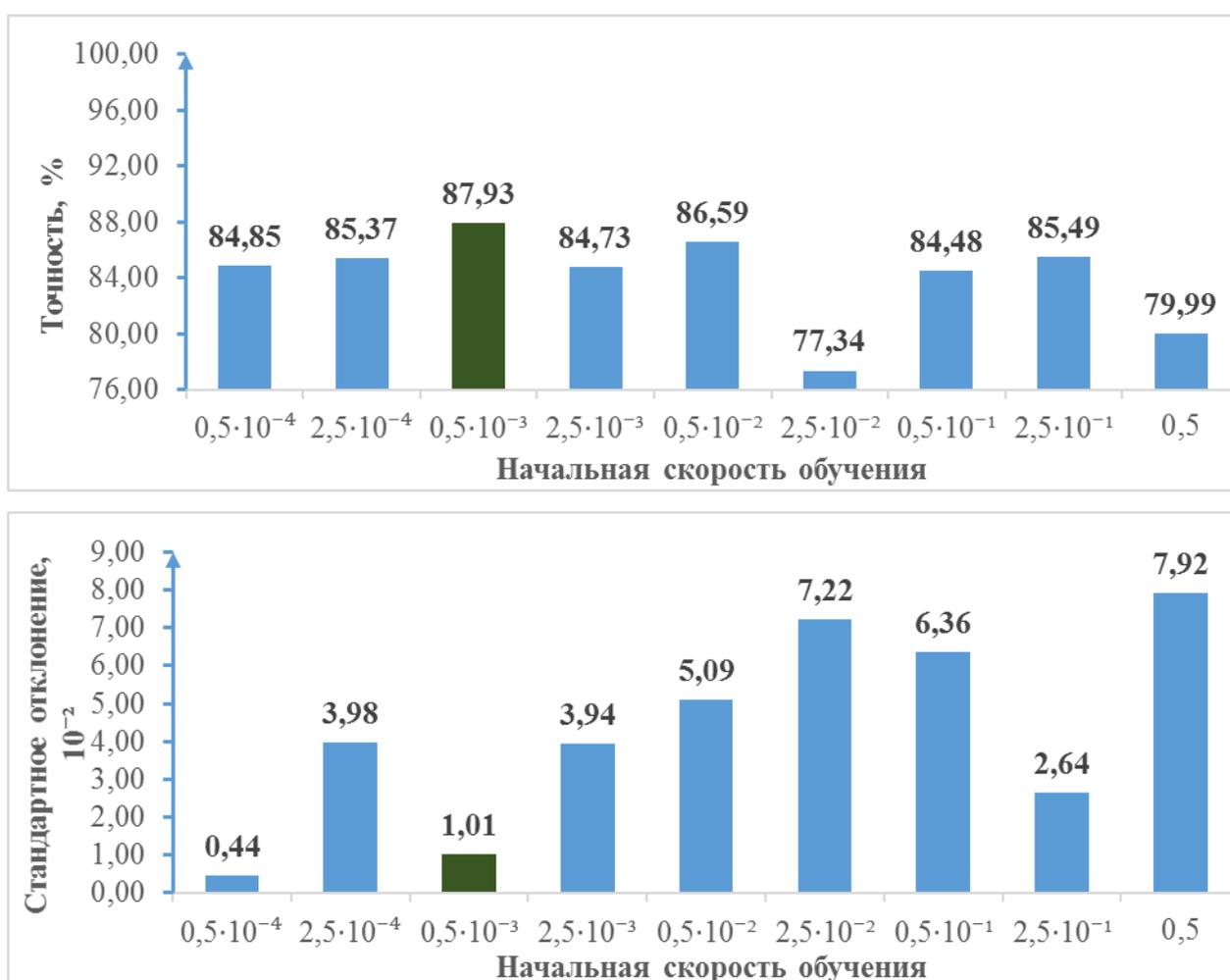


Рисунок 2.12. Среднее значение и стандартное отклонение точности классификации при различных значениях начальной скорости обучения для выборки фазовых составляющих частотного спектра

Для выборки фазовых составляющих частотного спектра оптимальным оказалось значение начальной скорости обучения, равное $0.5 \cdot 10^{-3}$. Данное значение параметра также соответствует одному из самых низких значений стандартного отклонения по сравнению с другими значениями из диапазона вариантов.

На рисунке 2.13 приведены диаграммы результатов подбора параметра регуляризации.

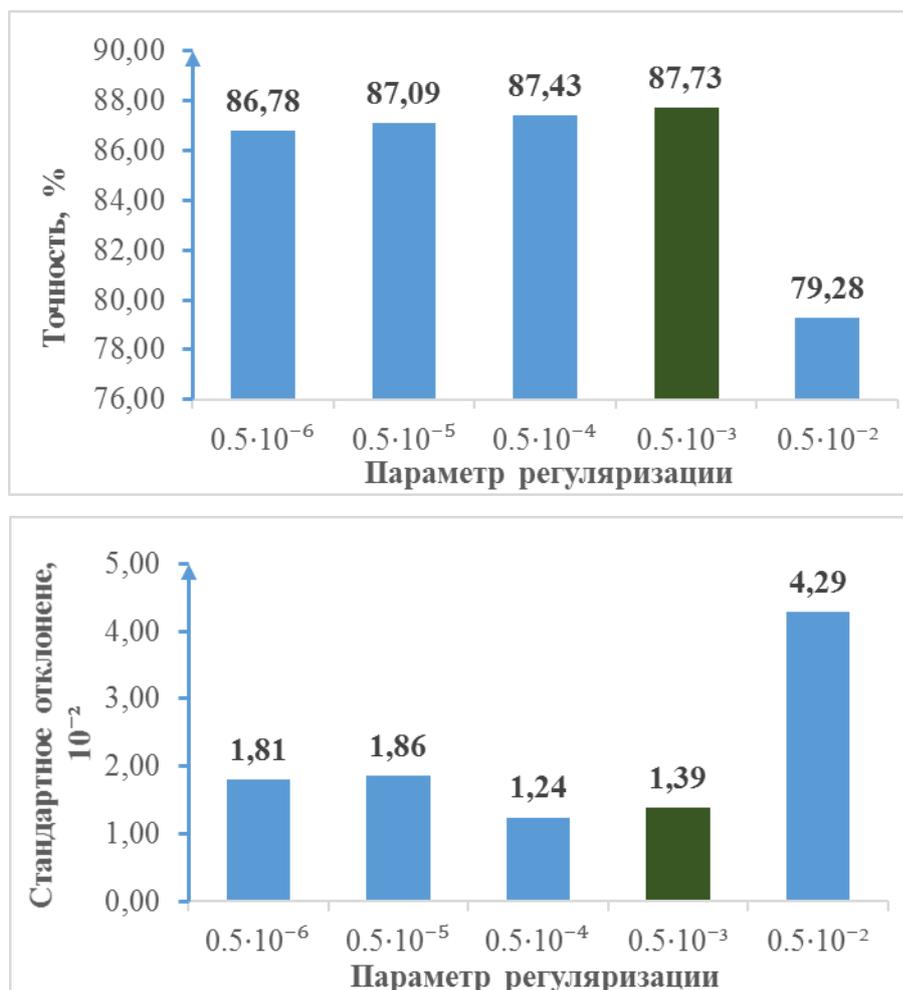


Рисунок 2.13. Среднее значение и стандартное отклонение точности классификации при различных значениях параметра регуляризации для выборки фазовых составляющих частотного спектра

Оптимальное значение параметра регуляризации для данной выборки составило $0.5 \cdot 10^{-3}$, что соответствует максимальному значению среди всех выборок и говорит о наибольшей подверженности модели переобучению при использовании данного входного вектора признаков.

На рисунке 2.14 приведены диаграммы результатов уточнения подбора начальной скорости обучения.

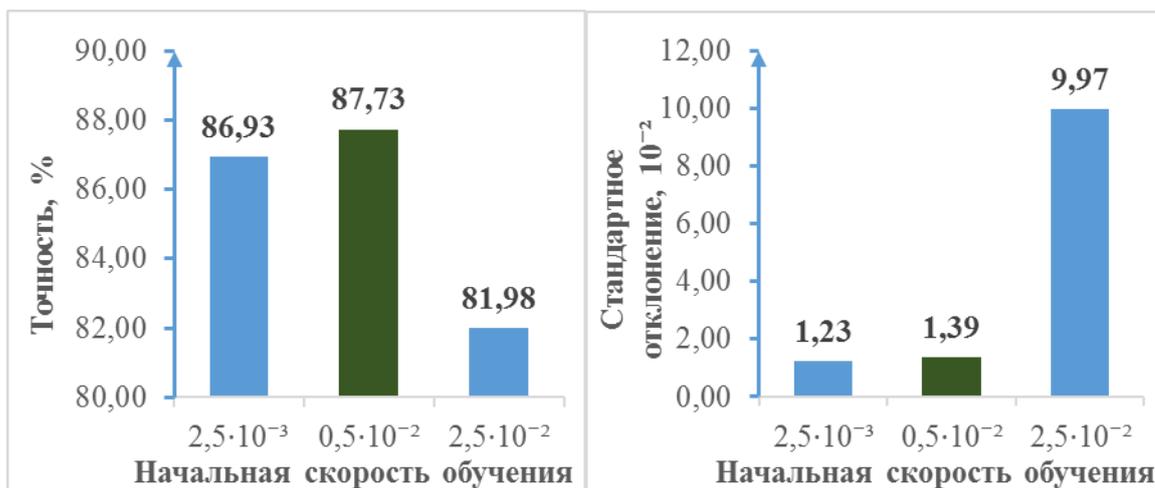


Рисунок 2.14. Среднее значение и стандартное отклонение точности классификации при различных значениях начальной скорости обучения для выборки фазовых составляющих частотного спектра (уточняющий этап)

Исходя из полученных данных можно сделать вывод, что оптимальное значение параметра начальной скорости обучения не изменилось.

На рисунке 2.15 приведены диаграммы результатов подбора предельного количества итераций без прогресса.

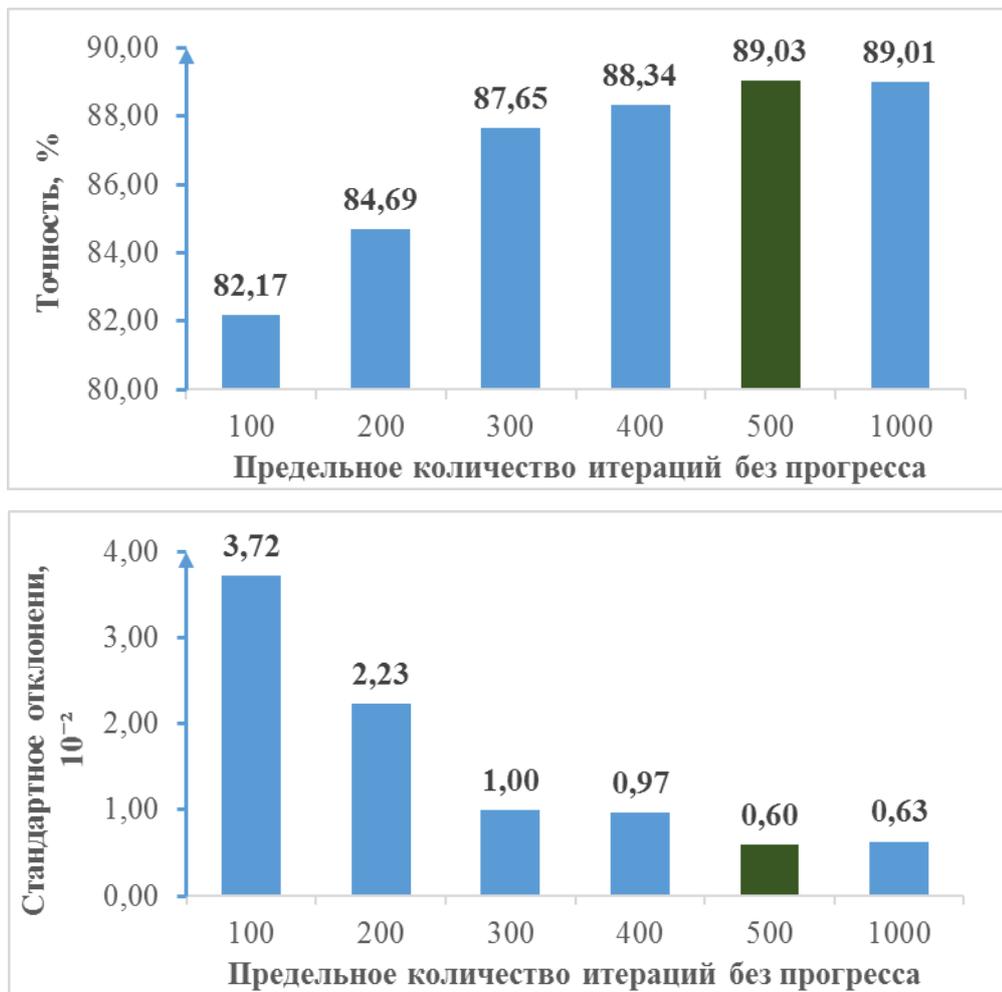


Рисунок 2.15. Среднее значение и стандартное отклонение точности классификации при различных значениях предельного количества итераций без прогресса для выборки фазовых составляющих частотного спектра

В случае данной выборки наилучший результат был достигнут при значении предельного количества итераций без прогресса, равном 500. Дальнейшее увеличение данного параметра не привело к улучшению результатов тестирования СНС, что позволяет предположить, что выбранное значение действительно является оптимальным для данного сочетания модели классификатора и варианта входного вектора признаков

2.2.5. Оптимальные параметры обучения СНС

В таблице 2.1 приведены оптимальные параметры обучения СНС, полученные в результате процедуры подбора, описанной в пункте 2.2.1.

Таблица 2.1. Оптимальные параметры обучения СНС

Параметр	Выборка исходных изображений	Выборка модульных составляющих	Выборка фазовых составляющих
Начальная скорость обучения	$0,5 \cdot 10^{-2}$	$0,5 \cdot 10^{-2}$	$0,5 \cdot 10^{-3}$
Параметр регуляризации	$0,5 \cdot 10^{-5}$	$0,5 \cdot 10^{-4}$	$0,5 \cdot 10^{-3}$
Количество итераций без прогресса	1000	1000	500

При использовании данных параметров обучения параметры классифицирующей модели будут наилучшим образом оптимизированы для классификации изображений лиц на основе соответствующих входных векторов признаков. Из различия параметров обучения видно, как выбор входного вектора влияет на продолжительность обучения и подверженность классифицирующей модели переобучению.

3. ПОДБОР АРХИТЕКТУРЫ СНС

3.1. Критерии подбора вариантов архитектуры СНС

В рамках поставленной задачи, обусловленных условиями практического применения, было выдвинуто требование к быстродействию СНС, сформулированное следующим образом: время однократного прямого прохода СНС не должно превышать 20 мс при запуске на ЭВМ, обладающей следующими характеристиками:

- CPU с частотой 3.5 ГГц, 8 ядер;
- 8 Гб ОЗУ;
- GPU с частотой 1.7 ГГц, 8 Гб видеопамяти.

Быстродействие сети напрямую зависит от количества параметров, то есть весовых коэффициентов сети. Таким образом, соответствие требованию к быстродействию на практике можно реализовать, делая архитектуру СНС более компактной, то есть выбирая небольшое количество слоёв и малые размеры фильтров и полносвязных слоёв сети.

Критерий быстродействия является наиболее приоритетным. Вторым критерием выбора архитектуры СНС является точность классификации, которая должна быть максимальной по сравнению с остальными и не должна быть ниже точности, демонстрируемой простым алгоритмом оценки размытия на основе анализа частотного спектра изображения без использования СНС.

3.2. Варианты архитектуры СНС

В ходе проектирования вариантов архитектуры СНС было выбрано 8 вариантов архитектуры с максимально возможной вариацией в пределах ограничения на компактность архитектуры. Выбранные варианты относятся к следующим категориям:

- 5 вариантов архитектуры LeNet с различным количеством пар свёрточных слоёв и слоёв подвыборки, а также различным размером полносвязного слоя (названия архитектур BlurNet01, BlurNet02, BlurNet03, BlurNet04 и BlurNet05);
- 2 варианта архитектуры ResNet минимально возможного размера в рамках особенностей архитектуры с различным количеством слоёв (архитектуры BlurNetResNet01 и BlurNetResNet02);
- 1 вариант архитектуры Inception минимально возможного размера в рамках особенностей архитектуры (архитектуры BlurNetInception01).

Программный код реализации каждого варианта сети при помощи библиотеки глубокого обучения Dlib-ml приводится в Приложении Б.

3.3. Выбор архитектур по критерию быстродействия

На обучающей выборке исходных изображений было обучено по одному экземпляру СНС каждой архитектуры из представленных выше. После этого было запущено несколько прямых проходов сети, получено время каждого прохода и вычислено среднее значение. Результаты тестирования СНС представлены в таблице 3.1.

Таблица 3.1. Результаты тестирования СНС на время прямого прохода

Архитектура СНС	Краткое обозначение	Время прямого прохода, мс
BlurNet01	BN1	15,0
BlurNet02	BN2	9,3
BlurNet03	BN3	8,5
BlurNet04	BN4	18,4
BlurNet05	BN5	23,5
BlurNetResNet01	BNRS1	23,8
BlurNetResNet02	BNRS2	21,5
BlurNetInception01	BNI1	46,0

Как можно увидеть из данных таблицы 3.1, требованию к быстродействию соответствуют архитектуры BlurNet01, BlurNet02, BlurNet03, BlurNet04. К тестированию на точность классификации также была допущена архитектура BlurNetResNet02 как наиболее быстрая из оставшихся архитектур.

3.4. Выбор архитектур по критерию точности

Таким же образом, как и на этапе подбора параметров обучающей модели, было получено значение средней точности и стандартного отклонения для СНС каждой из выбранных архитектур на основе 10 процедур обучения. Процедура тестирования была проведена для каждой из обучающих выборок, так как точность классификации определённой архитектуры зависит от особенностей входного вектора признаков. В общей сложности на данном этапе было проведено 150 полных процедур обучения СНС.

На рисунках 3.1-3.3 приведены результаты тестирования СНС различных архитектур для каждого варианта входного вектора признаков.

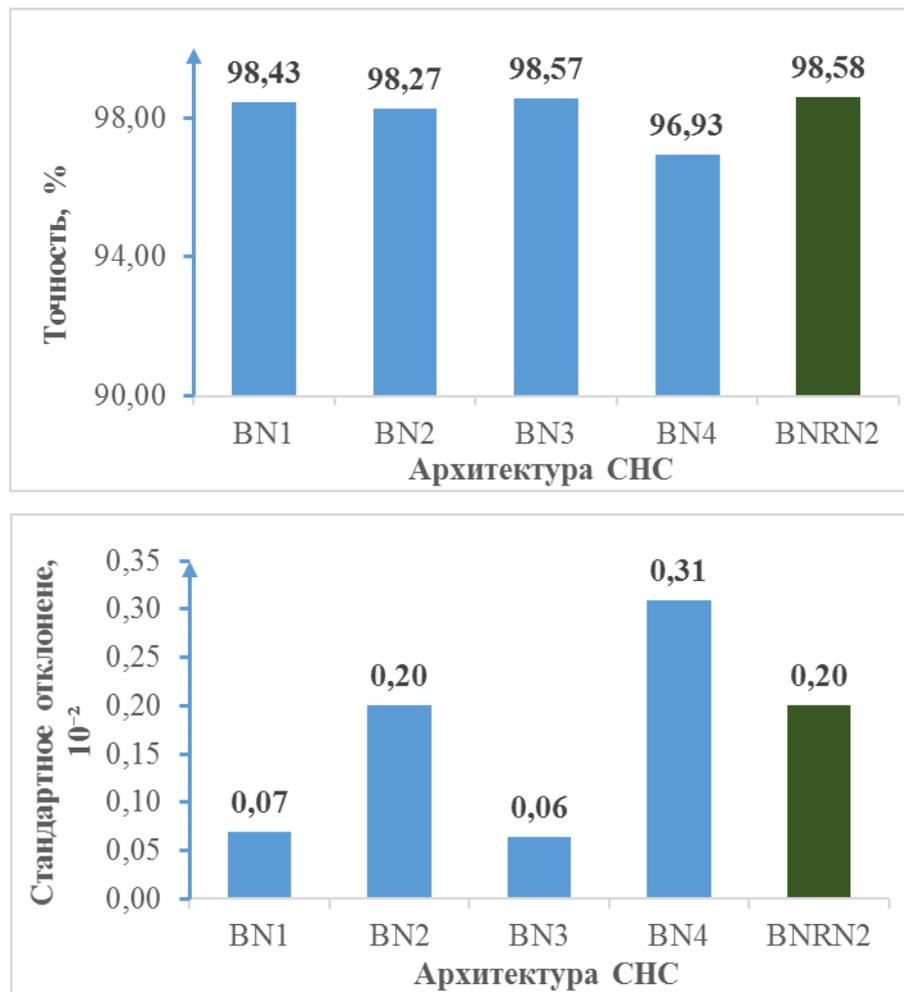


Рисунок 3.1. Среднее значение и стандартное отклонение точности классификации СНС различных архитектур для выборки исходных изображений

Исходя из полученных результатов, для выборки исходных изображений наибольшую точность показала архитектура BlurNetResNet02, второй результат показала архитектура BlurNet03. Однако первая архитектура по сравнению со второй даёт прирост в точности всего лишь в 0.01% одновременно со значительным увеличением времени прямого прохода. Исходя из этого, лучшей следует признать архитектуру BlurNet03.

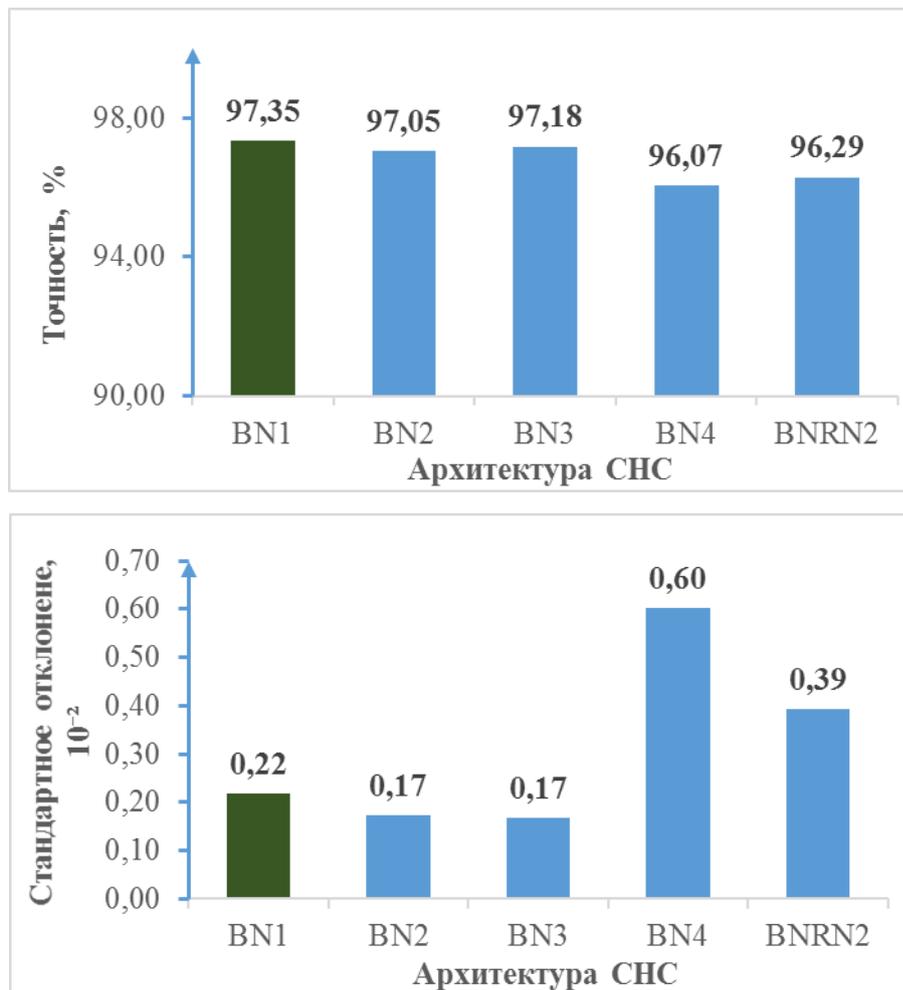


Рисунок 3.2. Среднее значение и стандартное отклонение точности классификации СНС различных архитектур для выборки модульных составляющих частотного спектра

Исходя из полученных результатов, для выборки модульных составляющих частотного спектра лучшей оказалась архитектура BlurNet01, лучший после неё результат с отставанием в 0.17% продемонстрировала архитектура BlurNet03.

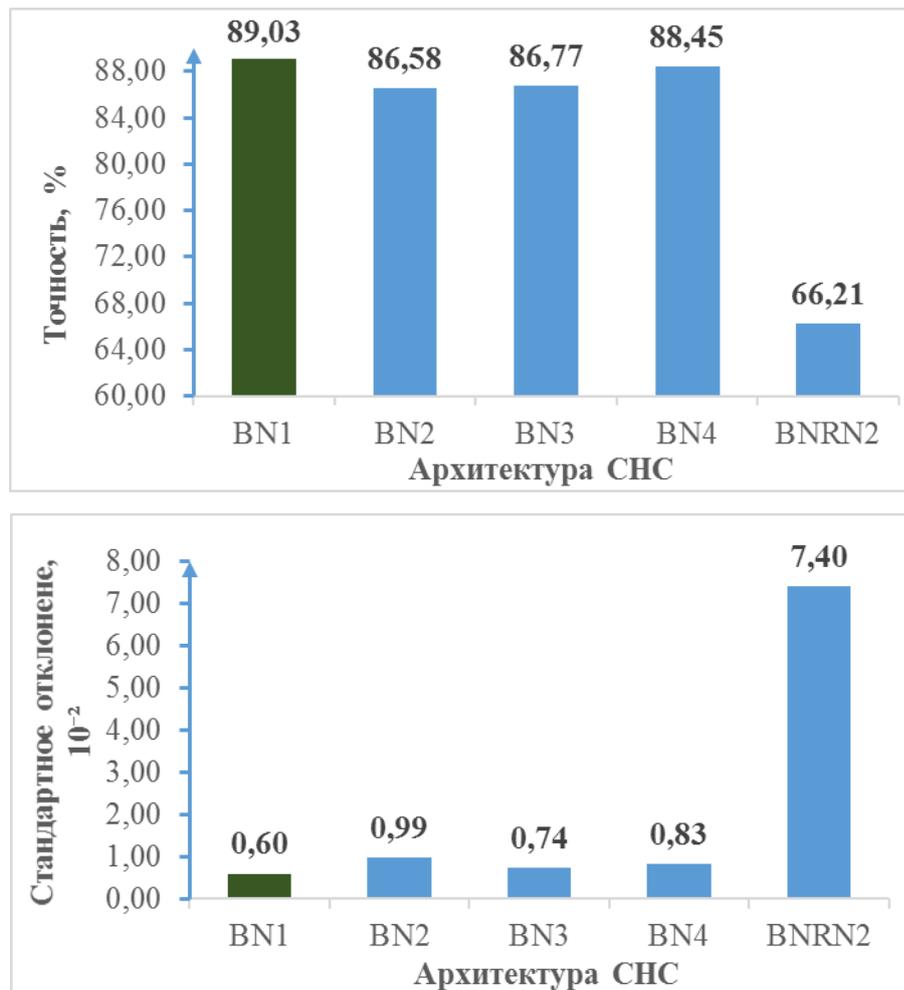


Рисунок 3.3. Среднее значение и стандартное отклонение точности классификации СНС различных архитектур для выборки фазовых составляющих частотного спектра

Для выборки фазовых составляющих частотного спектра лучшей оказалась архитектура BlurNet01.

В таблице 3.2 приведены результаты тестирования лучших вариантов СНС с оптимальными параметрами для каждого варианта входного вектора признаков.

Таблица 3.2. Результаты тестирования лучших вариантов СНС.

Параметр	Выборка исходных изображений	Выборка модульных составляющих	Выборка фазовых составляющих
Архитектура	BlurNet03	BlurNet01	BlurNet01
Средняя точность, %	98.57	97.35	89.03
Максимальная точность, %	98.66	97.70	89.85
Среднее отклонение точности	0.0006	0.0022	0.006

Наиболее эффективным оказался подход, при котором в качестве вектора входных признаков использовалось исходное изображение лица, что говорит о высокой информативности признаков, содержащихся в данных специфичных изображениях, для СНС с точки зрения оценки размытости и последующей бинарной классификации. Данный подход показывает как наибольшую точность, так и наибольшую стабильность, то есть вероятность высоких результатов обучения при использовании данного вектора признаков и подобранных параметров обучения.

Подход, при котором в качестве входного вектора признаков использовались модульные составляющие гармоник частотного спектра показал меньшую эффективность относительно первого подхода. Наименее эффективным оказался третий подход, при котором использовались фазовые составляющие гармоник частотного спектра.

С целью задания минимальной планки точности классификации был реализован алгоритм оценки размытия изображений на основе анализа частотного спектра, предложенный в работе [6] (описание алгоритма приводится в подразделе 1.2 данной работы). Реализованный алгоритм был протестирован на полученной тестовой выборке на предмет точности классификации после дополнительного подбора порогового значения оценки

размытия. Результаты тестирования, в том числе данные о значениях оценки размытия, которые использовались при подборе порогового значения, приведены в таблице 3.3.

Таблица 3.3. Результаты тестирования алгоритма [6] на основе анализа частотного спектра

Минимальное значение оценки размытия среди изображений с низкой степенью размытия	0.0087
Максимальное значение оценки размытия среди изображений с низкой степенью размытия	0.1054
Среднее значение оценки размытия среди изображений с низкой степенью размытия	0.0365
Минимальное значение оценки размытия среди изображений с высокой степенью размытия	0.0045
Максимальное значение оценки размытия среди изображений с высокой степенью размытия	0.0563
Среднее значение оценки размытия среди изображений с высокой степенью размытия	0,0193
Пороговое значение оценки размытия	0.024
Точность классификации, %	77.12

Из полученных результатов можно увидеть, что все СНС, продемонстрировавшие наилучший результат среди СНС для своего входного вектора признаков, превосходят минимальную планку точности, установленную простым алгоритмом на основе анализа частотного спектра изображения [6].

4. ФОРМУЛИРОВКА МЕТОДА ОБНАРУЖЕНИЯ РАСФОКУСИРОВАННЫХ И СМАЗАННЫХ ИЗОБРАЖЕНИЙ ЛИЦ

На основании результатов, представленных в предыдущих разделах данной работы, можно сформулировать метод обнаружения расфокусированных и смазанных (размытых) изображений лиц на основе СНС, состоящий из следующих основных положений:

1. Изображение лица подвергается процедуре предобработки:
 - a. обнаружение положения глаз в кадре, поворот и масштабирование изображения для расположения глаз в определённой позиции;
 - b. преобразование в оттенки серого;
 - c. эквализация гистограммы.
2. В качестве входного вектора признаков используется массив интенсивностей пикселей исходного изображения.
3. Для классификации изображения лиц используется СНС архитектуры BlurNet03 (варианты архитектур СНС приводятся в Приложении Б данной работы).
4. Для обучения СНС используется модификация алгоритма градиентного спуска Adam со следующими параметрами обучения:
 - a. начальная скорость обучения, равная $0.5 \cdot 10^{-2}$;
 - b. параметр регуляризации, равный $0.5 \cdot 10^{-5}$;
 - c. предельное количество итераций без прогресса, равное 1000;
 - d. прочие параметры устанавливаются согласно рекомендациям разработчиков алгоритма [28].
5. Для обучения СНС используется обучающая выборка изображений лиц, разбитая на два класса: изображения лиц с низкой степенью размытия (не менее 35000 изображений) и высокой степенью размытия (не менее 75000 изображений).

6. Для реализации СНС используется библиотека глубокого обучения Dlib-ml.
7. В результате выполнения указанных положений можно получить результат до 98.66% точности классификации изображений лиц на два указанных класса.
8. Продолжительность классификации одного изображения на оборудовании, конфигурация которого указана в пункте 3.1.1 данной работы, в среднем составит 8.5 мс.

5. ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ

5.1. Организация и планирование работ

В число исполнителей научно-исследовательской работы входят научный руководитель (НР) и студент магистратуры (Исполнитель, И). В таблице 5.1 приведены перечень работ и загрузка каждого из исполнителей. В период выполнения научно-исследовательской работы и научный руководитель, и Исполнитель работали в рамках шестидневной рабочей недели.

Таблица 5.1. Перечень работ и загрузка исполнителей

Этапы работы	Исполнители	Загрузка исполнителей
Постановка целей и задач, получение исходных данных	НР	НР – 100%
Составление и утверждение ТЗ	НР, И	НР – 100% И – 10%
Подбор и изучение материалов по тематике, обсуждение литературы	НР, И	НР – 30% И – 100%
Разработка календарного плана	НР, И	НР – 100% И – 10%
Подготовка обучающей выборки	НР, И	И – 100%
Подбор параметров модели классификатора	И	И – 100%
Сравнительный анализ работы обученных моделей классификатора, формулировка метода обнаружения размытых изображений лиц	И	И – 100%
Оформление расчетно-пояснительной записки	И	И – 100%
Подведение итогов	НР, И	НР – 60% И – 100%

5.2. Продолжительность этапов работ

Произведём опытно-статистический метод расчёта продолжительности этапов работ экспертным способом. Для определения вероятных (ожидаемых) значений продолжительности работ $t_{ож}$ применим формулу (5.1).

$$t_{ож} = \frac{3t_{min} + 2t_{max}}{5}. \quad (5.1)$$

Для построения линейного графика необходимо рассчитать длительность этапов в рабочих днях, а затем перевести ее в календарные дни. Расчет продолжительности выполнения каждого этапа в рабочих днях ($T_{РД}$) произведём по формуле:

$$T_{РД} = \frac{t_{ож}}{K_{ВН}} \cdot K_{Д}, \quad (5.2)$$

где $K_{ВН}$ – коэффициент выполнения работ, учитывающий влияние внешних факторов (на каждом из представленных этапов $K_{ВН} = 1$);

$K_{Д}$ – коэффициент, учитывающий дополнительное время на компенсацию непредвиденных задержек и согласование работ (на каждом из представленных этапов возьмём $K_{Д} = 1.2$).

Расчет продолжительности этапа в календарных днях ведется по формуле:

$$T_{КД} = T_{РД} \cdot T_{К}, \quad (5.3)$$

где $T_{КД}$ – продолжительность выполнения этапа в календарных днях;

$T_{К}$ – коэффициент календарности, позволяющий перейти от длительности работ в рабочих днях к их аналогам в календарных днях и рассчитываемый по формуле (5.4).

$$T_{К} = \frac{T_{КАЛ}}{T_{КАЛ} - T_{ВД} - T_{ПД}}, \quad (5.4)$$

где $T_{КАЛ}$ – календарные дни ($T_{КАЛ} = 365$);

$T_{ВД}$ – выходные дни ($T_{ВД} = 52$);

$T_{ПД}$ – праздничные дни ($T_{ПД} = 10$).

Таким образом, $T_K = \frac{365}{365-52-10} = 1,205$.

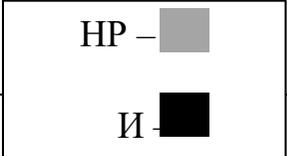
Результаты расчёта трудозатрат на выполнение проекта приведены в таблице 5.2. Линейный график работ представлен в таблице 5.3.

Таблица 5.2. Трудозатраты на выполнение проекта

Этап	Исполнители	Продолжительность работ, дни			Трудоемкость работ по исполнителям чел.-дн.			
				Трд		Ткд		
		t _{min}	t _{max}	t _{ож}	НР	И	НР	И
1	2	3	4	5	6	7	8	9
Постановка целей и задач, получение исходных данных	НР	2.00	4.00	2.80	3.36	–	4.05	–
Составление и утверждение ТЗ	НР, И	2.00	3.00	2.40	2.88	0.29	3.47	0.35
Подбор и изучение материалов по тематике, обсуждение литературы	НР, И	10.00	14.00	11.60	4.18	13.92	5.03	16.77
Разработка календарного плана	НР, И	2.00	4.00	2.80	3.36	0.34	4.05	0.40
Подготовка обучающей выборки	И	36.00	48.00	40.80	–	48.96	–	59.00
Подбор параметров модели классификатора	И	14.00	22.00	17.20	–	20.64	–	24.87
Сравнительный анализ работы обученных моделей классификатора, формулировка метода обнаружения размытых изображений лиц	И	5.00	9.00	6.60	0.00	7.92	0.00	9.54
Оформление расчетно-пояснительной записки	И	8.00	14.00	10.40	–	12.48	–	15.04
Подведение итогов	НР, И	5.00	8.00	6.20	4.46	7.44	5.38	8.97
Итого:				100,80	18,24	111,98	21,98	134,94

Таблица 5.3. Линейный график работ

Этап	НР	И	01.17		02.17			03.17			04.17			05.17			06.17	
			10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160
1	4.05	–	■															
2	3.47	0.35	■	■														
3	5.03	16.77		■	■													
4	4.05	0.40			■													
5	–	59.00				■	■	■	■	■	■	■	■	■	■	■	■	■
6	–	24.87										■	■	■	■	■	■	■
7	–	20.82													■	■	■	■
8	–	15.04														■	■	■
9	5.38	8.97															■	■



5.3. Расчёт накопления готовности проекта

Цель данного пункта – оценка текущих состояний (результатов) работы над проектом. Величина накопления готовности работы показывает, на сколько процентов по окончании текущего (i -го) этапа выполнен общий объем работ по проекту в целом.

Степень готовности определяется формулой

$$СГ_i = \frac{TP_i^H}{TP_{\text{общ.}}} = \frac{\sum_{k=1}^i TP_k}{TP_{\text{общ.}}} = \frac{\sum_{k=1}^i \sum_{j=1}^m TP_{km}}{\sum_{k=1}^I \sum_{j=1}^m TP_{km}}, \quad (5.5)$$

где $TP_{\text{общ.}}$ – общая трудоемкость проекта;

TP_i (TP_k) – трудоёмкость i -го (k -го) этапа проекта, $i = \overline{1, I}$;

TP_i^H – накопленная трудоёмкость i -го этапа проекта по его завершении;

TP_{ij} (TP_{kj}) – трудоёмкость работ, выполняемых j -м участником на i -м этапе, здесь $j = \overline{1, m}$ – индекс исполнителя.

В таблице 5.4 приведены нарастание технической готовности работы и удельный вес каждого этапа.

Таблица 5.4. Нарастание технической готовности и удельный вес этапов работы

Этап	ТР _i , %	СГ _i , %
Постановка целей и задач, получение исходных данных	2.58	2.58
Составление и утверждение ТЗ	2.43	5.01
Подбор и изучение материалов по тематике, обсуждение литературы	13.90	18.91
Разработка календарного плана	2.84	21.75
Подготовка обучающей выборки	37.60	59.34
Подбор параметров модели классификатора	15.85	75.19
Сравнительный анализ работы обученных моделей классификатора, формулировка метода обнаружения размытых изображений лиц	6.08	81.28
Оформление расчетно-пояснительной записки	9.58	90.86
Подведение итогов	9.14	100.00

5.4. Расчёт сметы затрат на выполнение проекта

В состав затрат на создание проекта включается величина всех расходов, необходимых для реализации комплекса работ, составляющих содержание данной разработки. Расчет сметной стоимости ее выполнения производится по следующим статьям затрат:

- материалы и покупные изделия;
- заработная плата;
- социальный налог;
- расходы на электроэнергию (без освещения);
- амортизационные отчисления;
- командировочные расходы;
- оплата услуг связи;
- арендная плата за пользование имуществом;

- прочие услуги (сторонних организаций);
- прочие (накладные расходы) расходы.

5.4.1. Расчёт затрат на материалы

К данной статье расходов относится стоимость материалов, покупных изделий, полуфабрикатов и других материальных ценностей, расходуемых непосредственно в процессе выполнения работ над объектом проектирования.

В рамках выполнения данного проекта затраты на материалы отсутствуют.

5.4.2. Расчёт заработной платы

Данная статья расходов включает заработную плату научного руководителя и инженера (в его роли выступает исполнитель проекта), а также премии, входящие в фонд заработной платы. Расчет основной заработной платы выполняется на основе трудоемкости выполнения каждого этапа и величины месячного оклада исполнителя и приведен в таблице 5.5.

Таблица 5.5. Затраты на заработную плату

Исполнитель	Оклад, руб./мес.	Среднедневная ставка, руб./раб. день	Затраты времени, раб. дни	Коэффициент	Фонд з/платы, руб.
НР	33162.87	1335.60	19	1.699	43114.40
И	25000.00	1006.85	112	1.699	190575.92
Итого:					233690.32

1.1.1. Расчет затрат на социальный налог

Затраты на единый социальный налог (ЕСН), включающий в себя отчисления в пенсионный фонд, на социальное и медицинское страхование,

составляют 30 % от полной заработной платы по проекту. В данном случае $C_{\text{соц}} = 233690.32 * 0.3 = 70107.10$ руб.

5.4.3. Расчет затрат на электроэнергию

Данный вид расходов включает в себя затраты на электроэнергию, потраченную в ходе выполнения проекта на работу используемого оборудования, рассчитываемые по формуле

$$C_{\text{эл.об.}} = P_{\text{об}} * t_{\text{об}} * Ц_{\text{э}}, \quad (5.6)$$

где $P_{\text{об}}$ – мощность, потребляемая оборудованием, кВт;

$Ц_{\text{э}}$ – тариф на 1 кВт*час, $Ц_{\text{э}} = 4.28$ р./(кВт*час);

$t_{\text{об}}$ – время работы оборудования, час.

Таблица 5.6. Затраты на электроэнергию технологическую

Наименование оборудования	Время работы оборудования $t_{\text{об}}$, час	Потребляемая мощность $P_{\text{об}}$, кВт	Затраты $Э_{\text{об}}$, руб.
Персональный компьютер	896*0.8	0.6	1840.74
Итого:			1840.74

5.4.4. Расчет амортизационных расходов

Амортизационные расходы составляют

$$C_{\text{ам}} = C_{\text{ам}}(\text{ПК}) = \frac{0.4 * 42000 * 896 * 1}{2384} = 6314.09 \text{ руб.}$$

5.4.5. Расчет расходов, учитываемых непосредственно на основе платежных (расчетных) документов (кроме суточных)

В рамках выполнения проекта не было произведено затрат, учитываемых на основе платёжных (расчётных) документов.

5.4.6. Расчет прочих расходов

В статье «Прочие расходы» отражены расходы на выполнение проекта, которые не учтены в предыдущих статьях, их следует принять равными 10% от суммы всех предыдущих расходов.

$$\begin{aligned} C_{\text{проч.}} &= (C_{\text{мат}} + C_{\text{зп}} + C_{\text{соц}} + C_{\text{эл.об.}} + C_{\text{ам}} + C_{\text{нп}}) * 0,1 = \\ &= (0 + 233690.32 + 70107.10 + 1840.74 + 6314.09 + 0) * 0.1 = \\ &= 31195.23 \text{ руб.} \end{aligned}$$

5.4.7. Расчет общей себестоимости разработки

Проведя расчет по всем статьям сметы затрат на разработку, можно определить общую себестоимость проекта. Смета затрат на разработку проекта приведена в таблице 5.7.

Таблица 5.7. Смета затрат на разработку проекта

Статья затрат	Условное обозначение	Сумма, руб.
Материалы и покупные изделия	$C_{\text{мат}}$	0
Основная заработная плата	$C_{\text{зп}}$	233690.32
Отчисления в социальные фонды	$C_{\text{соц}}$	70107.10
Расходы на электроэнергию	$C_{\text{эл.}}$	1840.74
Амортизационные отчисления	$C_{\text{ам}}$	6314.09
Непосредственно учитываемые расходы	$C_{\text{нр}}$	0
Прочие расходы	$C_{\text{проч}}$	31195.23
Итого:		343147.48

Таким образом, затраты на разработку проекта составили $C = 343147.48$ руб.

5.4.8. Расчет прибыли

Прибыль в размере 20% от полной себестоимости проекта составляет 68629.50 руб.

5.4.9. Расчет НДС

НДС составляет 18% от суммы затрат на разработку и прибыли. В данном случае НДС равняется $(343147.48 + 68629.50) * 0.18 = 74119.85$ руб.

5.4.10. Цена разработки НИР

Цена равна сумме полной себестоимости, прибыли и НДС, в данном случае

$$C_{\text{НИР(КР)}} = 343147.48 + 68629.50 + 74119.85 = 485896.83 \text{ руб.}$$

5.5. Оценка экономической эффективности проекта

Областью применения разработанного метода служат распределённые системы обнаружения и распознавания лиц. В случае обнаружения размытых изображений лиц, данные изображения не пересылаются сервером обнаружения лиц серверу распознавания лиц. Таким образом, отфильтровывается от 35 до 48% изображений (в зависимости от установленного порога бинарной классификации по степени размытия), что приводит к существенному уменьшению требований к пропускной способности каналов связи между серверами и снижению нагрузки на сервер распознавания лиц, что выражается в меньших затратах на оборудование и электроэнергию.

Другой областью использования разработанного метода является автоматизированный входной контроль качества эталонных материалов для поисковых баз данных систем видеонаблюдения. Эффект проекта выражается в снижении количества ошибок в работе систем, использующих эти базы данных, и, соответственно, повышении их надёжности.

Ожидаемый экономический эффект проекта носит непрямой коммерческий характер, так как разработанный метод является составной частью продуктов предприятия, повышает эффективность и надёжность их работы, и, следовательно, влияет на стоимость и объёмы продаж. Это выражается в дополнительной прибыли предприятия, при этом внедрение разработки не требует дополнительных транзакционных затрат. Определение конкретных значений дополнительной прибыли предприятия и срока окупаемости инвестиций выходят за рамки данной работы, так как требует сбора данных о составе, количестве и особенностях условий работы программно-аппаратных решений на предприятиях, внедряющих данную систему.

5.6. Оценка научно-технического уровня НИР

В таблице 5.8 приведены факторы и коэффициенты, необходимые для оценки научно-технического уровня НИР.

Таблица 5.8. Оценка научно-технического уровня НИР

Значимость	Фактор НТУ	Уровень фактора	Выбранный балл	Обоснование выбранного балла
0,4	Уровень новизны	Новая	5	Получены экспериментальные данные об эффективности применения аппарата СНС в специфической задаче
0,1	Теоретический уровень	Разработка способа	1	Разработка метода обнаружения размытых изображений лиц и его реализации
0,5	Возможность реализации	В течение первых лет	10	Предприятие готово немедленно применить разработку на практике

Отсюда интегральный показатель научно-технического уровня проекта составляет:

$$I_{\text{НТУ}} = 0.4 * 5 + 0.1 * 1 + 0.5 * 10 = 2 + 0.1 + 5 = 7.1$$

В таблице 5.9 указано соответствие качественных уровней НИР значениям рассчитанного показателя $I_{\text{НТУ}}$.

Таблица 5.9. Соответствие качественных уровней НИР значению $I_{НТУ}$

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

Таким образом, исходя из данных таблицы 9, данный проект имеет высокий уровень научно-технического эффекта.

6. СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ

Аннотация

Представление понятия «Социальная ответственность» сформулировано в международном стандарте (МС) IC CSR-08260008000: 2011 «Социальная ответственность организации».

В соответствии с МС социальная ответственность - ответственность организации за воздействие ее решений и деятельности на общество и окружающую среду через прозрачное и этическое поведение, которое:

- содействует устойчивому развитию, включая здоровье и благосостояние общества;
- учитывает ожидания заинтересованных сторон;
- соответствует применяемому законодательству и согласуется с международными нормами поведения (включая промышленную безопасность и условия труда, экологическую безопасность);
- интегрировано в деятельность всей организации и применяется во всех ее взаимоотношениях (включая промышленную безопасность и условия труда, экологическую безопасность) [42].

Раздел включает в себя следующие составляющие: техногенная безопасность, региональная безопасность, организационные мероприятия обеспечения безопасности, особенности законодательного регулирования проектных решений и безопасность в чрезвычайных ситуациях.

Введение

Объект исследования - разработка метода обнаружения размытых изображений лиц, предназначенного для применения в системах распознавания лиц и системах оценки качества изображений лиц.

Работа в рамках выполнения магистерской диссертации заключается в разработке программного кода и обработке данных в формате изображений и на всех своих этапах предполагает большой объем работы с ЭВМ, поэтому важным критерием безопасности является организация рабочего места и режима трудовой деятельности.

6.1. Производственная безопасность

6.1.1. Анализ вредных и опасных факторов, которые может создать объект исследования

Согласно номенклатуре, опасные и вредные факторы по ГОСТ 12.0.003-74 [43] делятся на следующие группы:

- физические;
- химические;
- психофизиологические;
- биологические;

Перечень опасных и вредных факторов, влияющих на персонал в заданных условиях деятельности, представлен в таблице 6.1.

Таблица 6.1. Перечень опасных и вредных факторов при разработке ПО

Источник фактора, наименование видов работ	Факторы		Нормативные документы
	Вредные	Опасные	
<ul style="list-style-type: none"> • работа с ПЭВМ; 	<ul style="list-style-type: none"> • Повышенная или пониженная температура; • Повышенная или пониженная влажность воздуха; • Зрительное напряжение; • Отсутствие или недостаток естественного света; • Повышенный уровень шума; • Электромагнитные излучения. 	<ul style="list-style-type: none"> • Электрический ток. 	<ul style="list-style-type: none"> • Гигиенические требования к микроклимату производственных помещений СанПиН 2.2.4-548-96; • Нормы естественного и искусственного освещения предприятий, СНиП 23-05-95; • Допустимые уровни шумов в производственных помещениях. ГОСТ 12.1.003-83. ССБТ; • Гигиенические требования к персональным электронно-вычислительным машинам и организации работы, СанПиН 2.2.2/2.4.1340-03;

6.1.2. Обоснование мероприятий по защите персонала предприятия от действия опасных и вредных факторов (техника безопасности и производственная санитария)

Требования к помещениям для работы с ПЭВМ

Основным документом, определяющим условия труда на персональных ЭВМ, являются «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы». Санитарные нормы и правила СанПиН 2.2.2/2.4.1340-03, которые были введены 30 июня 2003 года [42].

Магистерская диссертация выполнялась в рабочем помещении отдела разработки ПО ООО «НПК «Техника дела». Параметры рабочего помещения: длина – 7,12 м, ширина – 8,06 м, высота – 3,4 м. Для естественного освещения помещение оснащено двумя окнами, имеющими одинаковые размеры: ширина – 1,35 м, высота – 2,28 м. Дверь в помещение изготовлена из дерева и имеет коричневый цвет. На подвесном потолке установлена прессованная плитка из полистирола. На стенах поклеены обои белого цвета. Напольным покрытием является линолеум. Площадь помещения равна 57,39 м², объём – 195,12 м³. Рабочее помещение относится к классу помещений без повышенной опасности, так как отсутствуют условия, создающие повышенную или особо повышенную опасность.

Согласно СанПиН 2.2.2/2.4.1340-03 [44], норма площади рабочего места с персональным компьютером составляет 4,5 м². В рассматриваемом помещении установлено десять рабочих мест с персональными компьютерами и жидкокристаллическими экранами. Соответственно, на одного человека приходится 5,74 м², что соответствует вышеуказанным требованиям.

Освещённость

Рабочее (общее) освещение – это основное освещение, обеспечивающее нормальные условия для нахождения человека в помещении. Освещение в недостаточной степени может привести к напряжению зрения, ослаблению внимания и наступлению преждевременной утомленности. Слепление, резь в глазах и раздражение могут быть вызваны чрезмерно ярким освещением. Свет на месте труда может создать сильные тени или отблески, а также дезориентировать работающего. Перечисленные причины могут привести к профзаболеваниям [45].

По характеристике зрительной работы труд программиста относится к 3 разряду зрительной работы, т.е. наименьший размер объекта различения от 0,3 до 0,5 мм (точка). Это значит, что нормативное значение освещенности рабочего места должно быть минимум 200 лк (СанПиН 2.2.2/2.4.1340–03, СНиП 23-05-95) [44, 46].

Рассчитаем фактическую освещенность рассматриваемого рабочего помещения. Длина и ширина помещения равны соответственно 7,12 и 8,06 м, высота – 3,4 м. Рассчитаем индекс помещения:

$$i = \frac{S}{h * (A + B)}, \quad (6.1)$$

где i – индекс помещения;

S – площадь помещения, m^2 ;

h – высота помещения, м;

A – длина помещения, м;

B – ширина помещения.

$$i = \frac{57,39}{3,4 * (7 + 8)} = 1,1.$$

Исходя из значения индекса помещения можно определить, что коэффициент использования рассматриваемого светового светильника с

люминесцентными лампами равен 46% [47]. Рассчитаем освещенность по формуле, учитывая, что в помещении 14 светильников по 4 лампы в каждом:

$$E_{\text{факт}} = \frac{N * n * \Phi_{\text{ст}} * \eta}{S * K_3 * Z}, \quad (6.2)$$

где $E_{\text{н}}$ – фактическая освещенность;

N – число светильников в помещении;

n – число ламп в светильнике;

$\Phi_{\text{ст}}$ – величина стандартного светового потока, лм;

η – коэффициент использования светового потока;

S – площадь помещения;

K_3 – коэффициент запаса;

Z – коэффициент неравномерности освещения.

Известно, что $\Phi_{\text{ст}}$ равняется 720 лм, а мощность равняется 20 Вт для люминесцентных встроенных ламп для освещения общественных и жилых помещений (ЛВО). [47], K для помещений с малым выделением пыли равен 1,5, а Z для люминесцентных ламп равен 1.

Рассчитаем значение фактической освещенности:

$$E_{\text{факт}} = \frac{14 * 4 * 720 * 0,46}{57,39 * 1,5 * 1} = 215,4 \text{ лк.}$$

Данное значение близко к нормативному. Рассчитаем численную оценку разности между фактическим значением освещенности и нормативным:

$$\Delta E = \frac{(E_{\text{факт}} - E_{\text{н}})}{E_{\text{н}}} * 100\%, \quad (6.3)$$

где ΔE – показатель разности между фактической освещенностью и нормативной;

$E_{\text{факт}}$ – фактическое значение освещенности;

$E_{\text{н}}$ – нормативное значение освещенности.

$$\Delta E = \frac{(215,4 - 200)}{200} * 100\% = 7.7\%.$$

Зная, что допустимое отклонение освещенности может быть $\pm 20\%$, можно сделать вывод, что система освещения в помещении соответствует норме.

На рисунке 6.1 приведён план размещения светильников в рабочем помещении.

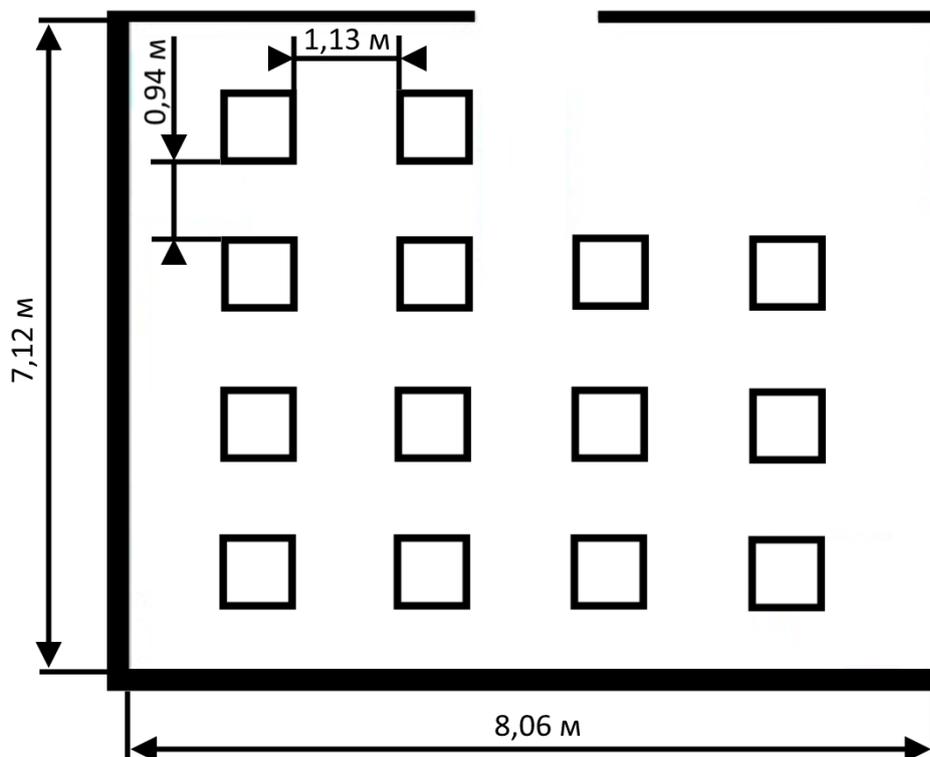


Рисунок 6.1. План размещения светильников на рабочем месте

Уровень шума

Люди, которым приходится работать в условиях длительного шума, обычно имеют головные боли, раздражительность, сталкиваются со снижением памяти, повышенной утомляемостью, также у многих понижен аппетит, есть боли в ушах и т. д. Перечисленные факты снижают производительность, работоспособность человека, а также качество труда [48].

Основными источниками шума в помещении являются системы охлаждения центральных процессоров и работающие жёсткие диски

размещённых в помещении компьютеров. Также источником шума является система вентиляции и шумы поступающий извне помещения.

Во избежание негативных последствий от производственного шума, его необходимо регулировать в соответствии с нормами, которые указаны в ГОСТ 12.1.003-83 «ССБТ. Общие требования безопасности» [49] для соответствия СанПиН 2.2.2/2.4.1340–03 [44].

Допустимые уровни звука и звукового давления для рабочего места разработчика-программиста согласно вышеуказанному ГОСТу 12.1.003-83 [49] представлены в таблице 6.2.

Таблица 6.2. Предельно допустимые уровни звука (ГОСТ 12.1.003-83)

Вид трудовой деятельности/ Частоты	Уровни звука и звукового давления, дБ, в октавных полосах со среднегеометрическими частотами, Гц								
	31,5	63	125	250	500	1000	2000	4000	8000
Научная деятельность, проектирование, программирование, Рабочие места проектно-конструкторских бюро, программистов вычислительных машин и т.д.	86	71	61	54	49	45	42	40	38

Допустимый уровень звукового давления колеблется от 38 дБ до 86 дБ при частоте от 8000 Гц до 31,5 Гц, соответственно. В рассматриваемом помещении уровень шума в течение суток находится в пределах диапазона от 34 до 45 дБ при частоте около 2000 Гц, из чего можно сделать вывод, что в помещении наблюдается повышенный уровень шума, выходящий за допустимые пределы и не соответствует требованиям СанПиН 2.2.2/2.4.1340–

03 [44]. Для уменьшения воздействий шума можно использовать следующие методы, согласно СНиП 23-03-2003 [50]:

1. применение звукоизоляции;
2. экранирование рабочих мест, то есть установка перегородок между рабочими местами;
3. установка оборудования, производящего минимальный шум.

Микроклимат помещения

Компьютеры могут привести к увеличению температуры и снижению относительной влажности в помещении. Поэтому необходимо соблюдать параметры микроклимата. В СанПиН 2.2.4.548 – 96 установлены величины параметров микроклимата, создающие комфортные условия [51].

Требования к микроклимату зависят от категории тяжести работ. Работа программиста относится к легкой категории, то есть к категории 1б (СанПиН 2.2.4.548 – 96) [51]. Для каждой категории тяжести работ есть значения показателей микроклимата.

В таблицах 6.3, 6.4 представлены данные показатели для теплого периода года (плюс 10 °С и выше) и для холодного периода года.

Таблица 6.3. Оптимальные величины показателей микроклимата
(СанПиН 2.2.4.548 – 96) [51]

Период года	Температура воздуха, °С	Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	21-23	20-24	40-60	0,1
Теплый	22-24	21-25		0,1

Таблица 6.4. Допустимые величины показателей микроклимата
(СанПиН 2.2.4.548 – 96) [51]

Период года	Температура воздуха, °С	Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	19-24	18-25	15-75	0,1-0,2
Теплый	20-28	19-29		0,1-0,3

В рассматриваемом помещении вентиляция осуществляется естественным и механическим путём. Тем не менее, температура в помещении в холодное время года может опускаться до 18-19 °С, а в теплое время года подниматься до 29-30 °С. Из этого можно сделать вывод, что температура воздуха в помещении выходит за рамки допустимых показателей.

Если температура воздуха отличается от нормальной, то время пребывания в таком помещении должно быть ограничено в зависимости от категории тяжести работ. В таблицах 6.5, 6.6 приведены показатели рекомендуемого времени работы в помещении при температуре ниже и выше допустимых величин.

Таблица 6.5. Рекомендуемое время работы при температуре воздуха ниже допустимых величин (СанПиН 2.2.4.548 – 96) [51]

Температура воздуха, °С	Время пребывания, не более, ч
17	6
18	7

Таблица 6.6. Рекомендуемое время работы при температуре воздуха выше допустимых величин (СанПиН 2.2.4.548 – 96) [51]

Температура воздуха, °С	Время пребывания, не более, ч
30,0	5
29,5	5,5
29,0	6

Непрерывное время работы людей в помещении варьируется от 1,5 до 6,5 часов, из чего можно сделать вывод, что время работы в помещении может превышать рекомендуемые параметры.

К мероприятиям по оздоровлению воздушной среды в производственном помещении относятся правильная организация вентиляции и кондиционирования воздуха, отопление помещений. В зимнее время в помещении предусматривается система отопления. В помещении установлен кондиционер Mitsubishi Electric MS-GF80VA.

Электромагнитное излучение

Электромагнитное излучение - распространяющееся в пространстве возмущение электрических и магнитных полей [52]. Источниками электромагнитного излучения в данном исследовании являются мониторы и системный блок.

На человека, работающего за компьютером, действует электромагнитное излучение. Опасность этого излучения заключается в том, что его влияние негативно сказывается на здоровье работника, а именно: на нервной, эндокринной и сердечно-сосудистой системах.

Оценка величины уровней ЭМП, проведенная по паспортным данным компьютера и монитора, показала их соответствие нормам ТСО–03 и СанПиН 2.2.2/2.4.1340–03 [44]. В таблице 6.7 приведены нормы уровня ЭМП, которым соответствует техника в кабинете.

Таблица 6.7. Допустимые уровни ЭМП, создаваемых ПК (СанПиН 2.2.2/2.4.1340–03)[3]

Наименование параметров		ВДУ ЭМП
Напряженность электрического поля	в диапазоне частот 5 Гц - 2 кГц	25 В/м
	в диапазоне частот 2 кГц - 400 кГц	2,5 В/м
Плотность магнитного потока	в диапазоне частот 5 Гц - 2 кГц	250 нТл
	в диапазоне частот 2 кГц - 400 кГц	25 нТл
Электростатический потенциал экрана видеомонитора		500 В

Источниками электромагнитного излучения в кабинете выступают системные блоки и мониторы включённых компьютеров. Для того, чтобы снизить воздействие таких видов излучения, рекомендуют применять такие мониторы, у которых уровень излучения понижен (MPR-II, TCO-92, TCO-99), а также установить защитные экраны и соблюдать режимы труда и отдыха.

Электрический ток

К опасным факторам относят поражение электрическим током согласно ГОСТ 12.0.003-74 [43]. Компьютеры, расположенные в помещении, питаются от сети 220В переменного тока с частотой 50Гц. Помещение с ПЭВМ, где проводились работы, относится к помещениям без повышенной опасности, согласно классификации помещений по опасности поражения людей электрическим током [53], так как отсутствуют следующие факторы:

- сырость;
- токопроводящая пыль;
- токопроводящие полы;
- высокая температура;
- возможность одновременного прикосновения человека к имеющим соединение с землёй металлоконструкциям зданий, технологическим аппаратам и механизмам и металлическим корпусам электрооборудования.

К мероприятиям по предотвращению возможности поражения электрическим током относятся [44]:

1. обеспечение недоступности токоведущих частей путем использования изоляции в корпусах оборудования;
2. применение средств коллективной защиты от поражения электрическим током;
3. применение защитного заземления;

4. применение защитного зануления;
5. применение защитного отключения;
6. использование устройств бесперебойного питания.
7. запрет на работы на задней панели при включенном сетевом напряжении;
8. проведение работ по устранению неисправностей только квалифицированным персоналом;
9. слежение за исправностью электропроводки.

В соответствии с ГОСТ 12.0.004-2015 [54] обучение и инструктаж по безопасности труда носит непрерывный многоуровневый характер.

Профессиональная подготовка персонала, повышение его квалификации, проверка знаний и инструктажи проводятся в соответствии с требованиями государственных и отраслевых нормативных правовых актов по организации охраны труда и безопасной работы персонала.

Электротехнический персонал до допуска к самостоятельной работе должен быть обучен приемам освобождения пострадавшего от действия электрического тока, оказания первой помощи при несчастных случаях.

Персонал, обслуживающий электроустановки, должен пройти проверку знаний Правил безопасности и других нормативно-технических документов (правил и инструкций по технической эксплуатации, пожарной безопасности, пользованию защитными средствами, устройства электроустановок) в пределах требований, предъявляемых к соответствующей должности или профессии, и иметь соответствующую группу по электробезопасности.

Персонал обязан соблюдать требования Правил безопасности, инструкций по охране труда, указания, полученные при инструктаже.

Инструктаж по характеру и времени проведения подразделяют на следующие виды [54]:

- вводный инструктаж;
- первичный и повторный инструктажи на рабочем месте;
- внеплановый инструктаж;
- целевой инструктаж.

6.2. Экологическая безопасность

При рассмотрении влияния персонального компьютера на атмосферу, гидросферу и литосферу выявлены особо вредные отходы согласно ГОСТ Р 51768-2001 [58], а именно электронная техника, требующая утилизации: компьютеры, оргтехника и т.п. Утилизация такого оборудования является достаточно сложной, так как они имеют сложную структуру. Непосредственная переработка большей части компонентов включает в себя их сортировку, последующую гомогенизацию и отправку для повторного использования, т.е. с предварительным помолом или переплавкой.

В случае выхода из строя компьютеров, они списываются и отправляются на специальный склад, который при необходимости принимает меры по утилизации списанной техники и комплектующих. В настоящее время в Томской области утилизацией занимаются две компании: городской полигон и ООО НПП «Экотом». Утилизацией опасных бытовых отходов занимаются компании: ООО «Торем», ООО «СибМеталлГрупп».

6.3. Безопасность в чрезвычайных ситуациях

К возможным ЧС, которые могут произойти в рабочем помещении, относятся:

- пожары и взрывы.

Одними из наиболее вероятных и разрушительных видов чрезвычайных ситуаций являются пожар или взрыв на рабочем месте, в связи с чем важное значение имеет пожарная безопасность (ГОСТ 12.0.003-74 [43]). Пожарная безопасность осуществляется системой пожарной защиты и системой

предотвращения пожара. Должны быть приняты следующие меры противопожарной безопасности:

1. обеспечение эффективного удаления дыма, т.к. в помещениях, имеющих оргтехнику, содержится большое количество пластиковых веществ, выделяющих при горении летучие ядовитые вещества и едкий дым;
2. обеспечение правильных путей эвакуации;
3. наличие огнетушителей и пожарной сигнализации;
4. соблюдение всех противопожарных требований к системам отопления и кондиционирования воздуха.

Основные причины возникновения пожаров:

1. нарушение правил пожарной безопасности;
2. перегрузка электросети;
3. неисправность прибора;
4. разряд молнии и неисправность молниеотвода.

По взрыво- и пожароопасности все помещения, согласно техническому регламенту НПБ 105-03 [55], делятся на 5 категорий, в зависимости от применяемых на производстве веществ и их количества. Рассматриваемое рабочее помещение относится к пожароопасной категории В, в качестве горючего материала выступает бумага [56]. Рабочее помещение оснащено ручным углекислотным огнетушителем ОУ-2, а также аптечкой первой помощи согласно требованиям ГОСТ Р 51057-01 [57].

Всякий работник при обнаружении пожара должен (ППБ 01-03 [59]):

1. незамедлительно сообщить об этом в пожарную охрану;
2. принять меры по эвакуации людей, каких-либо материальных ценностей согласно плану эвакуации;
3. отключить электроэнергию, приступить к тушению пожара первичными средствами пожаротушения.

Рабочее помещение оснащено ручным углекислотным огнетушителем ОУ-2, а также аптечкой первой помощи согласно требованиям ГОСТ Р 51057-01 [57]. При возникновении пожара должна сработать система пожаротушения, передав на пункт пожарной станции сигнал о ЧС. В случае если система не сработала, то необходимо самостоятельно произвести вызов пожарной службы по телефону 101, сообщить место возникновения ЧС и ожидать приезда специалистов. На рисунке 6.2 представлен план эвакуации при чрезвычайных ситуациях.



Рисунок 6.2. План эвакуации этажа при пожаре и других чрезвычайных ситуациях

6.4. Правовые и организационные вопросы обеспечения безопасности

6.4.1. Специальные правовые нормы трудового законодательства

В соответствии с государственными стандартами и правовыми нормами обеспечения безопасности предусмотрена рациональная организация труда в течение смены, которая предусматривает:

1. Длительность рабочей смены не более 8 часов;
2. Установление двух регламентируемых перерывов (не менее 20 минут после 1-2 часов работы, не менее 30 минут после 2 часов работы);
3. Обеденный перерыв не менее 40 минут.

Обязательно предусмотрен предварительный медосмотр при приеме на работу и периодические медосмотры.

Каждый сотрудник должен пройти инструктаж по технике безопасности перед приемом на работу и в дальнейшем пройти инструктаж по электробезопасности и охране труда.

Предприятие обеспечивает рабочий персонал всеми необходимыми средствами индивидуальной защиты.

Оплата труда, социальные пособия, дополнительные выплаты устанавливаются в соответствии со степенью вредности и опасности выполняемых обязанностей.

6.4.2. Организационные мероприятия при компоновке рабочей зоны программиста

Требования к рабочему месту при работе с ПЭВМ

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. На рисунке 6.3 показана правильная организация рабочего места [44].

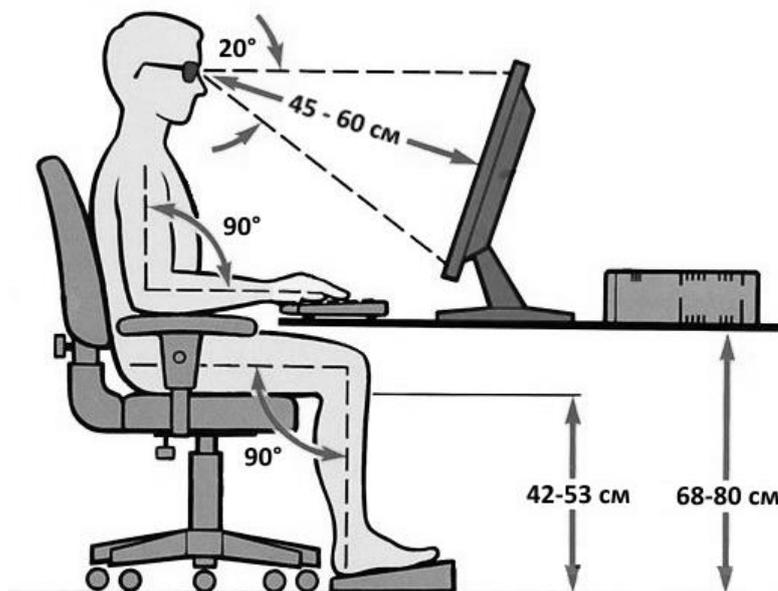


Рисунок 6.3. Организация рабочего места.

Согласно СанПиН 2.2.2/2.4.1340-03 [44]:

- 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° ниже уровня глаза;
- дисплей должен находиться на расстоянии 45-60 см от глаз;
- локтевой сустав при работе с клавиатурой нужно держать под углом 90° ;
- 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.

ЗАКЛЮЧЕНИЕ

В рамках данной работы были решены следующие задачи:

1. Получена и подготовлена обучающая выборка изображений лиц с низкой и высокой степенью размытия, которая может быть в дальнейшем использована в задачах глубокого обучения для оптимизации параметров классифицирующих моделей. Размер выборки составляет 110000 изображений.
2. Произведён аналитический обзор неэталонных методов оценки/обнаружения размытых изображений, используемых моделей классификаторов, подбор алгоритма обучения и библиотеки глубокого обучения с обоснованием выбора.
3. Произведены подбор оптимальных параметров обучения СНС для каждого из предложенных вариантов входного вектора признаков. Спроектированы несколько вариантов архитектуры СНС и произведён выбор архитектур, дающих наилучший результат на каждом из предложенных вариантов входного вектора признаков.
4. Произведён сравнительный анализ эффективности использования каждого входного вектора признаков в задаче обнаружения размытых изображений лиц при помощи СНС. Наиболее эффективным оказалось использование вектора признаков, соответствующего исходному изображению лица (98.66% точности классификации), что говорит о высокой информативности признаков, содержащихся в специфичных изображениях лиц, для СНС с точки зрения классификации по степени размытости. Менее эффективным оказалось использование модульных составляющих гармоник частотного спектра (97.70% точности классификации), а наименее эффективным оказалось использование фазовых составляющих (89.85% точности классификации).
5. Сформулирован метод обнаружения размытых изображений лиц, предложенный впервые и удовлетворяющий требованиям к

быстродействию и точности классификации, наложенным условиями практического применения. Метод позволяет добиться точности классификации в 98.66% и демонстрирует продолжительность классификации одного изображения на ЭВМ приведённой конфигурации, равную 8.5 мс. Результатом реализации метода является обученный экземпляр СНС, готовый к применению в решении задачи обнаружения размытых изображений лиц.

Таким образом, в ходе работы были выполнены все сформулированные задачи, и цель работы была достигнута. Метод принят к использованию предприятием-заказчиком, помимо этого подготовлены доклады научно-технических конференций.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Are Existing Procedures Enough? Image and Video Quality Assessment: Review of Subjective and Objective Metrics / S. Ouni, M. Chamnah, M. Herbin, E. Zagrouba // Farnand, S. P. Image Quality and System Performance V / S.P. Farnand, F. Gaykema // Proceedings of SPIE. – 2008. – Vol. 6808. – P. 68080Q-1–68080Q-11.
2. Igor, V. Examining the Impact of Blur on Recognition by Convolutional Networks / V. Igor, Ch. Ayan, Sh. Gregory // arXiv. – 2016. – №1611.05760.
3. Wang, Z. Modern image quality assessment: Synthesis lectures on image, video & multimedia processing / Z. Wang, A. C. Bovik. – San Rafael: Morgan & Claypool Publishers, 2006. – 156 p.
4. Ramakrishnan, H. N. Detection and estimation of image blur: master's thesis. – Missouri University of Science and Technology, 2010. – 72 p.
5. Оппенгейм А. В. Цифровая обработка сигналов: пер. с англ. / А. В. Оппенгейм, Р. В. Шафер., С. Ф. Боев (ред., пер. с англ.) – М.: Техносфера, 2012. – 1046 с.
6. Kanjar De. Image Sharpness Measure for Blurred Images in Frequency Domain / Kanjar De, V. Masilamani // Procedia Engineering. – 2013. – Vol. 64. – P. 149-158.
7. Image Partial Blur Detection and Classification [статья]
8. Marzialono, P. A no-reference perceptual blur metric / P. Marzialono, F. Dufaux, S. Winkler, T. Ebrahimi // Proceedings of IEEE international conference on Image Processing. – 2002. – Vol. 3. – P. 57-60.
9. Tong, H. Blur detection for digital images using wavelet transform / H. Tong, M. Li, H. Zhang, C. Zhang // Proceedings of IEEE international conference on Multimedia and Expo. – 2004. – Vol. 1. – P. 17-20.
10. Fergus, R. Removing camera shake from a single photograph / R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, W. T. Freeman. – Association for

- Computing Machinery, ACM Trans. Graph. – 2006. – Vol. 25. – No. 3. – P. 787-794.
11. Jia, J. Single image motion de-blurring using transparency // Computer Vision and Pattern Recognition, CVPR. – 2007. – P. 1-8.
 12. Воронцов К.В. Лекции по статистическим (байесовским) алгоритмам классификации [Электронный ресурс]. – 2008. – URL: <http://www.ccas.ru/voron/download/Bayes.pdf>, свободный. – Загл. с экрана. – Яз. рус. Дата обращения: 10.03.2017 г.
 13. Duin, R. P. W. Open issues in pattern recognition / R. P. W. Duin, E. Pekalska // Computer Recognition Systems. Advances in Soft Computing. – 2005. – Vol. 30. – P. 27–42.
 14. Hearst, M.A. Support vector machines / M.A. Hearst, S.T. Dumais, E. Osuna, J. Platt, B. Scholkopf // IEEE Intelligent Systems. – 1998. – Vol. 13. – №4. – P. 18–28.
 15. Hubel, D. H. Receptive fields and functional architecture of monkey striate cortex / D. H. Hubel, T. N. Wiesel // The Journal of physiology. – 1968. – №195. – P. 215–243.
 16. LeCun, Y. Backpropagation applied to handwritten zip code recognition // Neural computation. – 1989. – Vol. 1. – №4. – P. 541–551.
 17. Fukushima, K. Neocognitron: A hierarchical neural network capable of visual pattern recognition // Neural networks. – 1988. – Vol. 1. – №2. – P. 119–130.
 18. Serre, T. Object recognition with features inspired by visual cortex / T. Serre, L. Wolf, T. Poggio // Computer Vision and Pattern Recognition. – 2005. – №2. – P. 994–1000.
 19. Krizhevsky, A. Imagenet classification with deep convolutional neural networks / A. Krizhevsky, I. Sutskever, G.E. Hinton // Advances in neural information processing systems. – 2012. – P. 1097–1105.

20. Bundy, A. Difference of Gaussians / A. Bundy, L. Wallen // Catalogue of Artificial Intelligence Tools. – 1984. – P. 30–35
21. Deng, L. Recent advances in deep learning for speech research at Microsoft // Acoustics, Speech and Signal Processing (ICASSP). – 2013. – P. 8604–8608
22. Lee, H. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations // Proceedings of the 26th Annual International Conference on Machine Learning. – 2009. – P. 609–616.
23. Bottou, L. Stochastic Gradient Descent Tricks // Neural Networks: Tricks of the Trade: Lecture Notes in Computer Science / G. Montavon, G. B. Orr, K.-R. Muller (Eds.) – Springer, 2012. – P. 421-436.
24. Sutskever, I. On the Importance of Initialization and Momentum in Deep Learning / I. Sutskever, J. Martens, G. Dahl, G. Hinton // Proceedings of the 30th International Conference on Machine Learning. – 2013. – Vol. 28. – P. III-1139–III-1147
25. Duchi, J. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization / J. Duchi, E. Hazan, Y. Singer // The Journal of Machine Learning Research. – 2011. – Vol. 12. – P. 2121-2159.
26. Tieleman, T. RMSProp: Divide the gradient by a running average of its recent magnitude / T. Tieleman, G. Hinton // Coursera: Neural Networks for Machine Learning: Technical report. – 2012. – P. 26-31.
27. Zeiler M. Adadelta: An adaptive learning rate method // arXiv. – 2012. – №1212.5701.
28. Kingma, D. Adam: A Method for Stochastic Optimization / D. Kingma, J. Ba // arXiv. – 2015. – №1412.6980.
29. LeCun, Y. Efficient backprop // Neural Networks: Tricks of the Trade: Lecture Notes in Computer Science / G. Montavon, G. B. Orr, K.-R. Muller (Eds.) – Springer, 2012. – P. 9-48.

30. Ioffe, S. Batch normalization: Accelerating deep network training by reducing internal covariate shift / S. Ioffe, C. Szegedy // arXiv. – 2015. – №1502.03167.
31. Feature selection, L1 vs. L2 regularization, and rotational invariance
32. Torch: A scientific computing framework for LUAИТ [Электронный ресурс]. – URL: <http://torch.ch>, свободный. – Загл. с экрана. – Яз. англ. Дата обращения: 10.03.2017 г.
33. Theano at a Glance [Электронный ресурс]. – URL: <http://deeplearning.net/software/theano/introduction.html>, свободный. – Загл. с экрана. – Яз. англ. Дата обращения: 10.03.2017 г.
34. Pylearn2 dev documentation [Электронный ресурс]. – URL: <http://deeplearning.net/software/pylearn2>, свободный. – Загл. с экрана. – Яз. англ. Дата обращения: 10.03.2017 г.
35. Caffe – Deep learning framework by BAIR [Электронный ресурс] // berkeleyvision.org – URL: <http://caffe.berkeleyvision.org>, свободный. – Загл. с экрана. – Яз. англ. Дата обращения: 10.03.2017 г.
36. King, D. E. Dlib-ml: A Machine Learning Toolkit // Journal of Machine Learning Research. – 2009. – Vol. 10. – P. 1755-1758.
37. LeCun, Y. Gradient-based learning applied to document recognition / Y. LeCun, L. Bottou, Y. Bengio, P. Haffner // Proceedings of the IEEE. – 1998. – Vol. 86. – №11. – P. 2278-2324.
38. He, K. Deep Residual Learning for Image Recognition / K. He, X. Zhang, S. Ren, J. Sun // arXiv. – 2015. - №1512.03385.
39. Arora S. Provable Bounds for Learning Some Deep Representations / A. Bhaskara, R. Ge, T. Ma // arXiv. – 2013. – №1310.6343v1.
40. Szegedy, C. Going deeper with convolutions / C. Szegedy et al. // arXiv. – 2014. – №1409.4842.
41. Кафтанников И. Л. Проблемы формирования обучающей выборки в задачах машинного обучения / И. Л. Кафтанников, А. В. Парасич //

- Вестник Южно-Уральского государственного университета. – 2016. – Т. 16. – № 3.
42. Международный стандарт ИС CSR-08260008000: 2011 «Социальная ответственность организации»;
 43. ГОСТ 12.0.003-74 ССБТ. Опасные и вредные производственные факторы. Классификация. - М.: Издательство стандартов, 2001. – 4 с.
 44. СанПиН 2.2.2/2.4.1340-03. Гигиенические требования к персональным электронно-вычислительным машинам и организации работы.– М.: Информационно-издательский центр Минздрава России, 2003. – 54 с.;
 45. Искусственное освещение производственных помещений и его расчет [Электронный ресурс] // Open Manager. – URL: <http://www.openmanager.ru/opmans-289-1.html>, свободный. – Загл. с экрана. – Яз. рус., Дата обращения: 23.05.2017.
 46. СНиП 23-05-95. Естественное и искусственное освещение. – М.: Центр проектной продукции в строительстве, 2011. – 70 с.;
 47. ГОСТ 6825-91. Лампы люминесцентные трубчатые для общего освещения. – М.: Издательство стандартов, 1992. – 242 с.;
 48. Борьба с шумом на производстве: Справочник / Е.Я. Юдин, Л.А. Борисов; Под общ. ред. Е.Я. Юдина – М.: Машиностроение, 1985. – 400с., ил.;
 49. ГОСТ 12.1.003-83. ССБТ. Общие требования безопасности. – М.: Издательство стандартов, 2002. – 13 с.;
 50. СНиП 23-03-2003. Защита от шума. – М.: Госстрой России, 2004. – 34 с.;
 51. СанПиН 2.2.4.548 – 96. Гигиенические требования к микроклимату производственных помещений. – М.: Информационно-издательский центр Минздрава России, 1997. – 20 с.;

52. Безопасность жизнедеятельности. /Под ред. Н.А. Белова - М.: Знание, 2000 - 364с.;
53. Классификация помещений по опасности поражения людей электрическим током – Правила устройства электроустановок, 1999 г., извлечение.
54. ГОСТ 12.0.004-2015 «ССБТ. Организация обучения безопасности труда. Общие положения» – М.: Стандартинформ, 2016. – 40 с.;
55. НПБ 105-03. Определение категорий помещений и зданий по взрывопожарной и пожарной опасности. – М.: ВНИИПО, 2003. – 31 с.;
56. СП 12.13130.2009. Определение категорий помещений, зданий и наружных установок по взрывопожарной и пожарной опасности. – М.: Проспект, 2010. – 32 с.;
57. ГОСТ Р 51057-01. Огнетушители переносные. Общие технические требования. Методы испытаний. – М.: Издательство стандартов, 2001. – 48 с.
58. ГОСТ Р 51768-2001. Ресурсосбережение. Обращение с отходами. Методика определения ртути в ртутьсодержащих отходах. Общие требования. – М: Издательство стандартов, 2001. - 13 с.;
59. ППБ 01-03. Правила пожарной безопасности в Российской Федерации. – М.: ФГУ ВНИИПО МЧС России, 2003. – 111 с.;

ПРИЛОЖЕНИЕ А. ЧАСТЬ ВКР НА АНГЛИЙСКОМ ЯЗЫКЕ

Подраздел 1.4. Modifications of gradient descent method

Подраздел 1.5. Deep learning software libraries for CNN implementation

Подраздел 1.6. CNN architecture modifications

Студент:

Группа	ФИО	Подпись	Дата
8ВМ5А	Казиев Александр Борисович		

Консультант кафедры ИСТ:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент	Мирошниченко Евгений Александрович	к.т.н.		

Консультант – лингвист кафедры ИЯИК:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
старший преподаватель	Рыбушкина Светлана Владимировна	-		

1.4. Modifications of gradient descent method

1.4.1. Batch and stochastic gradient descent

The gradient descent method and its various modifications are the most widely applied methods for artificial neural networks (ANN) training including convolutional neural networks (CNN). The gradient descent method is also known as back-propagation algorithm in case of ANN training. The weights of an ANN are the parameters that require optimization and the goal of the optimization is minimization of the error function of the last layer.

The original variant of the gradient descent method also called batch gradient descent assumes feeding every object of a training set to the ANN, calculation of error function and updating each weight for each of the objects and only then summing up and applying the calculated modifications. In practice, it causes the training process to be extremely long assuming we have a big training set. That is why the method is modified so as the weights are modified upon every gradient calculation based on a single object of the training set. Such a modification is called stochastic gradient descent (SGD).

SGD updates the weights W by a negative gradient $\nabla J(W)$ of the error function. The learning rate α is a weight coefficient to the gradient. Thus, the weights are updated by the formula:

$$W_{t+1} = W_t + \Delta W_{t+1} = W_t - \alpha \nabla J(W_t). \quad (1.12)$$

Modification of the weights in the SGD does not necessarily take place after the submission of each object of the training sample. It can be done after summing the gradients calculated on a relatively small set of objects of the training set (mini-batch), which reduces the effect of the noise present in the training set [23].

In context of ANN training, the error of each neuron depends on errors of neurons of all subsequent layers of the network. Such a big number of free

parameters leads to possible complications in the course of minimizing the target error function:

1. A complex landscape of the target function, in which regions of strong nonlinearity neighbor with wide plateaus, which leads to large gradient oscillations depending on the current position of the target function value.

2. Non-uniform updating of parameters, as the input data can contain rare, yet informative features.

3. Necessity of empirical selection of a learning rate. A too small value of the learning rate can lead to a halt in a local minima and long convergence in general and a too big one can lead to exiting a narrow global minimum.

Experimenting with formulas of weights update leads to modifications of the gradient descent method, which has certain advantages that allow overcoming the listed complications.

1.4.2. Nesterov's accelerated gradient

Nesterov's accelerated gradient method implements two basic ideas to modify the gradient descent method:

3. Using the momentum of the motion of the target function value in the parameters space. Continuous movement in a certain direction suggests that it is would be proficient to move further in this direction. The magnitude of the update of each parameter, i.e., the weighting coefficient, must therefore take into account the magnitude of its update on the previous step with a certain weighting factor μ .
4. Looking ahead in direction of the motion vector. The gradient is not calculated at point W in the parameters space but in point $W + \mu\Delta W_{t-1}$, which the target function is going to move into.

In Nesterov's accelerated gradient descent method the weights are updated by the formula:

$$W_{t+1} = W_t + \Delta W_{t+1} = W_t - \alpha \nabla J(W_t). \quad (1.12)$$

This method results in a significantly faster training process as well as reduced effect of the noise present in the training set because of the previous weight updates affecting subsequent ones [24].

1.4.3. Adagrad

Adagrad (derived from *adaptive gradient*) is a modification of gradient descent method addressing the problem of rare, yet informative features.

In this method each i -th weight is updated by the following formula:

$$(W_{t+1})_i = (W_t)_i - \alpha * \frac{(\nabla J(W_t))_i}{\sqrt{\sum_{t'=1}^t (\nabla J(W_{t'}))_i^2}}. \quad (1.14)$$

This formula takes into account the previous weight updates. Thus, the weight is updated with consideration of the typicality of the feature it represents, the more frequently updated weights being modified less. Another advantage of the method is innecessity of thorough empirical selection of the learning rate because of the fact that the learning rate fades out during the training process [25].

1.4.4. RMSProp и AdaDelta

The RMSProp and AdaDelta algorithms are modifications of the Adagrad algorithm. These modifications address the problem of algorithm paralysis due to excessive increase of the sum of the previous weight updates.

In the RMSProp algorithm the running average of previous updates is used instead of their full sum. The running average is calculated by the following formula:

$$E((W_t)_i) = \delta E((W_{t-1})_i) + (1 - \delta)(\nabla J(W_t))_i^2. \quad (1.15)$$

The value of δ is usually set to 0.99. In this method the weights are updated by the formula:

$$(W_{t+1})_i = (W_t)_i - \alpha * \frac{(\nabla J(W_t))_i}{\sqrt{E((W_t)_i)}} [26]. \quad (1.16)$$

The AdaDelta algorithm uses a stabilizing term by the name of *RMS*, which is proportional to ΔW_{t+1} . It is being done to equalize dimensionalities of W and ΔW . The weight updates are taking place in three stages: accumulating the squared gradient by the formula (1.15), weight update and *RMS*(ΔW) update by the following formulas:

$$(W_{t+1})_i = (W_t)_i - \alpha * \frac{RMS(\Delta W_{t-1})_i}{RMS(\nabla J(W_t))_i} (\nabla J(W_t))_i, \quad (1.17)$$

$$RMS(\nabla J(W_t))_i = \sqrt{E((W_t)_i) + \varepsilon}, \quad (1.18)$$

where ε is a constant designed to prevent *RMS* from equaling zero on the first step of the training [27].

1.4.5. Adam

The Adam algorithm is a generalization of the AdaGrad algorithm that utilizes both the concept of the motion momentum of the target function value, as in the Nesterov's accelerated gradient method, and the concept of decreasing updates of more frequently updated parameters. In the Adam algorithm weight updates are calculated according to following formulas:

$$(m_t)_i = \beta_1(m_{t-1})_i + (1 - \beta_1)(\nabla J(W_t))_i, \quad (1.19)$$

$$(\Delta W_t)_i = \beta_2(\Delta W_{t-1})_i + (1 - \beta_2)(\nabla J(W_t))_i, \quad (1.20)$$

$$(W_{t+1})_i = (W_t)_i - \alpha * \frac{\sqrt{1 - (\beta_2)_i^t}}{1 - (\beta_1)_i^t} * \frac{(m_t)_i}{\sqrt{(\Delta W_t)_i + \varepsilon}}. \quad (1.21)$$

The developers of the method recommend using the following values of the constants: $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\varepsilon = 10^{-8}$ [28].

1.4.6. Selection of modification of gradient descent method

Table 1.1 shows the compilation of features of the listed modifications of the gradient descent method. The designations are as follows: SGD (1), Nesterov's accelerated gradient (2), Adagrad (3), RMSProp/AdaDelta (4), Adam(5).

Table 1.1. Features of the gradient descent method modifications

Feature	1	2	3	4	5
Utilizing motion momentum of the target function value	-	+	-	-	+
Looking ahead in the direction of the motion vector	-	+	-	-	-
Decreasing updates of more frequently updated parameters	-	-	+	+	+
Robustness against algorithm paralysis	+	+	-	+	+

In this master's dissertation the Adam algorithm was selected for CNN training considering the following: the algorithm is one of the most recent and advanced modification of gradient descent method that includes the most of its predecessors' valuable features.

1.4.7. Batch-normalization

Regardless of the selected algorithm of CNN weights optimization it is expedient to use the process of batch-normalization. Its main purpose is to decrease covariant shift, which is an effect of slowing down the training process due to high level of variety within the training set.

In the context of ANN training this effect takes place in the inner layers of the net. Every layer receives data from the output of the previous layer and thus can be regarded as an individual component of the training system. The updating of the weights of the ANN leads to modification of input data distribution of each component [29]. The training will converge faster if a preliminary normalization and decorrelation of the input data of each component occurs. In case of batch-

normalization the data is normalized on the input of the most part of the ANN's inner layers.

The algorithm of the batch-normalization is as follows:

- Normalize k -th dimension of data x for a layer of dimensionality D :

$$\hat{x}^k = \frac{x^k - \mu_B}{\sigma_B^2}, \quad (1.22)$$

where μ_B and σ_B^2 are the expected value and dispersion for the mini-batch $B = \{x_i^k, \dots, m\}$, calculated by formulas:

$$\mu_B = \frac{1}{m} \sum_{i=1}^m x_i^k, \quad (1.23)$$

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^m (x_i^k - \mu_B)^2. \quad (1.24)$$

- Acquire new normalized value:

$$y^k = \gamma^k \hat{x}^k + \beta^k, \quad (1.25)$$

where parameters γ, β are set during the training for each of the dimensions.

- Add the final representation into the ANN before the non-linear activation function.

Thus, the batch-normalization:

- decreases the covariant shift of the inner layers of the ANN increasing the speed of the training process;
- allows a less thorough process of the learning rate selection;
- does not prevent usage of the back-propagation method as the batch-normalization is a differentiable transformation [30].

1.4.8. Regularization

Overfitting of the classifier model is a typical complication in the area of machine learning. It is a situation when the results of testing the model on a testing set (which does not overlap with the training set) start becoming worse due to overly tight adaptation of the model to the training set. Thus, the model starts classifying the objects of the training set better over time, showing progressively worse results on the objects not included into the training set, which dramatically reduces its practical effectiveness.

A typical measure to counteract overfitting is using regularization, i.e. modifying the target function in the way that penalizes the model for excessive complexity. There are two major kinds of regularization:

3. L1-regularization or lasso regularization, which proposes the following modifications:

$$L1 = J(W) - \frac{\lambda}{2} \sum_i (W)_i, \quad (1.26)$$

where λ is the weight reduction coefficient.

4. L2- regularization or Tikhonov regularization, which proposes the following modification:

$$L1 = J(W) - \frac{\lambda}{2} \sum_i (W)_i^2. \quad (1.27)$$

L1-regularization has certain advantages such as ability of implicit feature selection by mean of possibility for features with low value to become zero. Yet in practice the L2-normalization is used more widely as it provides a more smooth regularization and is generally more universal (L1-regularization usually shows worse results in case of gradient-based optimization methods) [31].

1.5. Deep learning software libraries for CNN implementation

To achieve the best possible results implementing the CNN it is expedient to use one of the software libraries, which implement the deep learning algorithms. Such libraries are usually freely distributed, well optimized and support calculations on graphics processing units (GPU), which is especially relevant for CNNs, that provide wide possibilities for parallel processing.

There is a plenty of such libraries available nowadays, so the review of the most widely used ones will be given.

1.5.1. Torch

Torch is a software library used for scientific research with a wide support of machine learning algorithms. It was developed by Idiap Research Institute, New York University and NEC Laboratories America. The Torch library is implemented in the Lua language, the CUDA library is used to implement GPU calculations. The library supports Linux, FreeBSD and Mac OS X operational systems.

Torch allows creating sophisticated ANNs using the mechanism of containers. A container is a class, which integrates ANN's components into a single configuration that is later passed to the training process. The components of the ANN may be fully-connected or convolutional layers as well as non-linear activation functions, loss layers or other containers [32].

1.5.2. Theano

Theano is an extension of Python programming language that allows effective calculation of mathematical expressions upon multidimensional arrays. Theano was developed in LISA laboratory in order to streamline the machine learning algorithm development.

The library is implemented in Python language and supports Windows, Linux and Mac OS X operational systems. Theano contains a compiler to transform the mathematical expressions in Python into effective code in C or CUDA.

The Theano library provides a basic set of tools for configuring and training of the ANNs. It allows implementation of multilayer fully-connected ANNs, CNNs, recurrent ANNs, autoencoders and restricted Boltzmann machines. Several activation functions are available such as sigmoid, softmax or cross-entropy functions. The batch gradient descent method is used for the training process [33].

1.5.3. Pylearn2

Pylearn2 is a deep learning software library implemented in Python language. It is essentially an extension of the Theano library. Pylearn2 is being developed since 2011 in the Monreal University and is distributed under the BSD 3-clause license. The library supports Linux operational system and uses the Cuda-convnet library based upon CUDA/C++ to boost GPU calculations.

Pylearn2 allows implementing fully-connected ANNs, CNNs, autoencoders and restricted Boltzmann machines, it supports several kinds of activation functions and a variety of training methods such as the batch and stochastic gradient descent methods as well as nonlinear conjugate gradient method. The library uses YAML-formatted configuration files to store structures of the ANNs that is a quick and convenient approach to object serialization [34].

1.5.4. Caffe

The Caffe library was implemented in C++ language, it supports Linux, Mac OS X and Windows operational systems. The library was developed in University of California at Berkeley in 2013, currently it is being supported by The Berkeley Vision and Learning Center and distributed under the BSD 2-Clause license. Caffe uses the BLAS library for vector and matrix calculations as well as the cuDNN library based upon CUDA/C++ to boost GPU calculations.

The library allows implementing fully-connected ANNs and CNNs, supports a wide range of loss functions and a variety of training methods such as SGD, AdaGrad or Nesterov's accelerated gradient. Caffe uses prototxt-formatted

configuration files to store structures of the ANN's, training and implementation parameters [35].

1.5.5. Dlib-ml

The Dlib-ml library was developed in 2009 and is being actively extended since then (its last update took place in March of 2017). It is a freely distributed cross-platform library implemented in C++ language. As well as the other listed libraries, Dlib-ml provides wide capabilities of implementing deep learning algorithms including designing and application of CNNs, but also has several advantages:

- A convenient API designed for usage in commercial C++ software development.
- Using template classes to define and store ANN's architectures that streamlines the process of ANN architecture design.

The Dlib-ml library uses CUDA library to support GPU calculations and implements SGD and Adam training algorithms [36].

1.5.6. Deep learning software library selection

The criteria of deep learning software library selection as well as compliance with the criteria by each of the listed libraries are given in table 1.2.

Table 1.2. Compliance with the criterion by each library

Criterion	Torch	Theano	Pylearn2	Caffe	Dlib-ml
Windows support	-	+	+	+	+
C++ implementation	-	-	-	+	+
GPU support	+	+	+	+	+
Availability of up-to-date documentation	+	+	-	+	+

The Dlib-ml library has several advantages compared to Caffe:

- x1.6 faster functioning using calculations on GPU [24];
- less dependencies that streamlines the execution and compilation of programs;
- storing ANN's structure within programs as template classes instead of keeping it on disk.

Based on this Dlib-ml was chosen for CNN's implementation and training in this master's dissertation.

1.6. CNN architecture modifications

1.6.1. LeNet architecture

The LeNet architecture was proposed by Y. LeCunn in 1998 in one of the first papers dedicated to implementation and practical application of CNNs. In this paper it was demonstrated that CNNs of the proposed architecture perform efficiently in the task hand-written symbols recognition, illustrating it with training and testing on the MNIST dataset [37].

The first work that provided the vast popularity for the CNNs was a paper by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton published in 2012. In this paper a modification of LeNet architecture was proposed that consisted of 5 convolutional layers, 3 pooling layers and 3 fully-connected layers. The CNN was intended to participate in ILSVRC12 (ImageNet Large-Scale Visual Recognition Challenge 2012) competition, in which the competitors used different algorithms of computer vision to solve the tasks of classification of objects from 1000 categories. The proposed CNN by the name AlexNet has won the competition demonstrating result of 15.4% top-5 classification error (leaving behind the second competitor with result of 26.2%), that appeared to be a breakthrough in the area of object classification and conditioned the high level of interest to the CNNs for the following years [19].

The LeNet architecture proposes straightforward and consequent arrangement of convolutional and pooling layers with addition of several fully-connected layers to interpret the results of feature extraction. In the original form it consists of 7

layers: an input layer, two pairs of convolutional and pooling layers and two fully-connected layers [37]. This architecture is given in figure 1.2 in paragraph 1.3.3 of this dissertation. Modification to the LeNet architecture may be applied by the means of:

- altering amount of convolutional and pooling layers;
- altering amount and size of filters in convolutional layers;
- altering size and stride in pooling layers;
- altering amount and size of fully-connected layers;
- adding batch-normalization;
- altering the type of non-linear activation functions of the neurons.

Modifications of the CNNs of this architecture may be used in any task of evaluating and classification connected with processing local features of the images.

1.6.2. ResNet architecture

The ResNet architecture was developed by Microsoft Research Asia division in 2015. The very deep CNN consisting of 152 consequent layers set new records in the tasks of classification, detection and localization of objects within images. A module of ResNet architecture is given in figure 1.3.

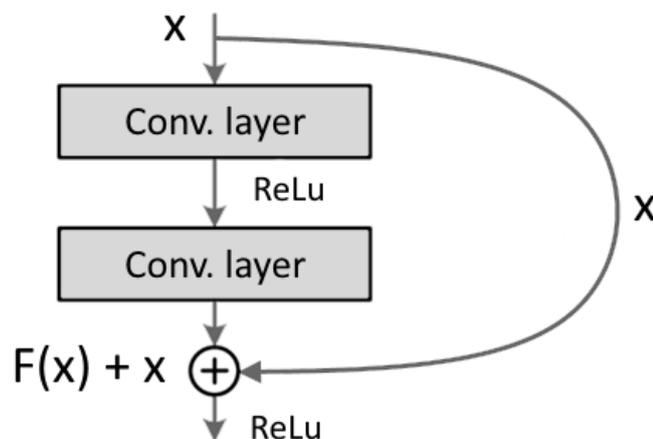


Figure 1.3. A ResNet module

An additional connection is presented which passes a copy of the input data to the output of the stack of convolutional layers, where both handfuls of data are being concatenated. This way the output becomes a slightly modified representation of the input data instead of a completely new representation, which contains no information about the input data, as in traditional CNNs. This measure accelerates the training process and make it more efficient by the mean of addressing the vanishing gradient problem, which is especially relevant for deep CNNs. Microsoft ResNet became the winner of ILSVRC15 competition by demonstrating result of 3.6% top-5 error [38].

1.6.3. Inception architecture

The most straightforward way of increasing CNN's effectiveness is by expanding their structure, which may infer increasing its depth (the amount of layers) or increasing its width (the amount of neurons in each layer). However, expanding the CNN architecture also leads to the increased amount of parameters which makes it prone to overfitting, especially it the training set is not relatively small. Another drawback of this approach is a dramatically increased computational cost. For example, a linear addition of filters to one of the two directly connected convolutional layers of a deep CNN would lead to a quadratic growth of computational cost.

A fundamental solution to this problem may be a transition from fully-connected architecture to sparsely-connected ones and this can apply to convolutional layers as well. The main result of [39] is the observation, that in case if the input data probability distribution can be estimated by a deep sparsely-connected ANN, an optimal architecture can be derived layer by layer by means of analyzing the correlation between activations of neurons of the last layer and stably related connecting neurons.

A team of researchers from Google in collaboration with the University of North Carolina and the University of Michigan used this approach to create their

own implementation of a deep CNN with the architecture by the name Inception. The final variant of the CNN took part in ILSVRC14 competition of detection and classification under the name of GoogLeNet and demonstrated the best result of 6.66% top-5 error (this variant was optimized for classification tasks and trained on ILCVRC12 training set). The second result was demonstrated by a cascade of CNNs by the name of VGG with the result of 7.33% top-5 error. The main idea of the Inception architecture is finding an optimal sparsely-connected CNN structure that could be approximated by modern dense computing methods [40].

A single module of Inception architecture is given in figure 1.4.

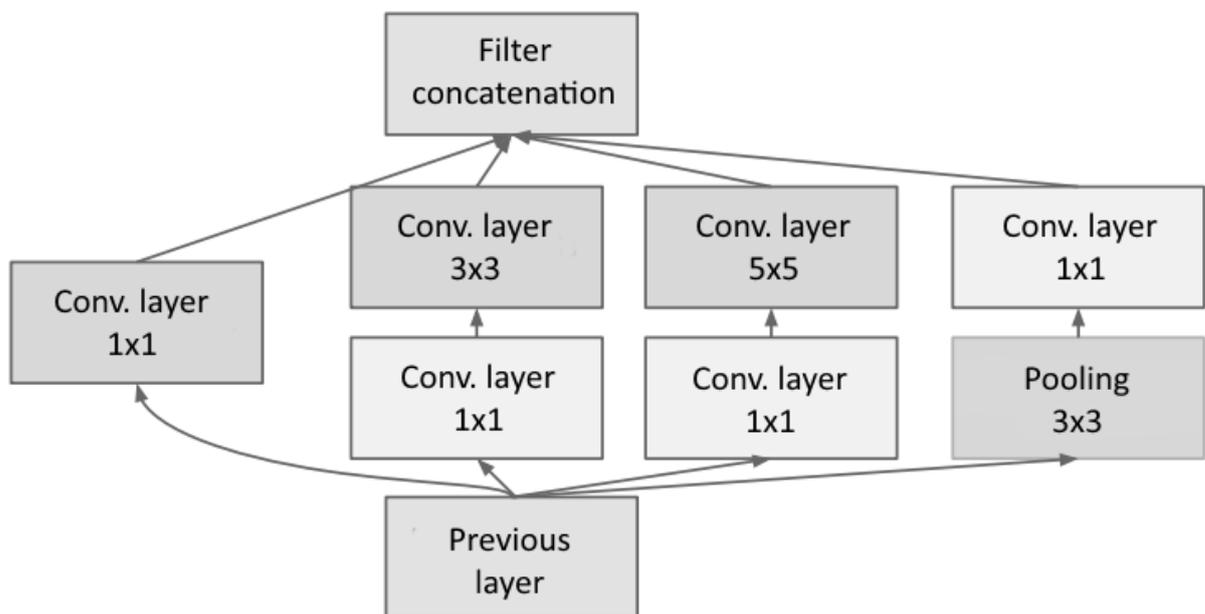


Figure 1.4. An Inception module that uses 1x1 convolutional layers for dimensionality reduction


```

template <typename SUBNET> using block_a4 = relu<bn_con<con<10,
1, 1, 1, 1, max_pool<3, 3, 1, 1, SUBNET>>>>;
template <typename SUBNET> using incept_a =
inception4<block_a1, block_a2, block_a3, block_a4, SUBNET>;

template <typename SUBNET> using block_b1 = relu<bn_con<con<4,
1, 1, 1, 1, SUBNET>>>>;
template <typename SUBNET> using block_b2 = relu<bn_con<con<4,
3, 3, 1, 1, SUBNET>>>>;
template <typename SUBNET> using block_b3 = relu<bn_con<con<4,
1, 1, 1, 1, max_pool<3, 3, 1, 1, SUBNET>>>>>>;
template <typename SUBNET> using incept_b =
inception3<block_b1, block_b2, block_b3, SUBNET>;

using BlurNetInception01 = loss_multiclass_log<
    fc<2,
    relu<
    bn_fc<fc<100,
    max_pool<2, 2, 2, 2, incept_b<
    max_pool<2, 2, 2, 2, incept_a<
    input_rgb_image_sized<183>
    >>>>>>>>>;

```

ПРИЛОЖЕНИЕ В. CD-ДИСК