

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

Инженерная школа информационных технологий и робототехники
 Направление подготовки 09.04.01 «Информатика и вычислительная техника»
 Отделение информационных технологий

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Нейросетевые алгоритмы распознавания эмоционального состояния человека по изображениям его лица

УДК 004.032.26:004.93'1:159.942

Студент

Группа	ФИО	Подпись	Дата
8BM73	Кузнецова Анастасия Сергеевна		

Руководитель ВКР

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

КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

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

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент ООД	Алексеев Николай Архипович			

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
09.04.01 "Информатика и вычислительная техника", профиль "Компьютерный анализ и интерпретация данных"	Спицын Владимир Григорьевич	д.т.н., профессор		

Томск – 2019 г.

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

Код результата	Результат обучения (выпускник должен быть готов)
P1	Применять базовые и специальные естественнонаучные и математические знания в области информатики и вычислительной техники, достаточные для комплексной инженерной деятельности.
P2	Применять базовые и специальные знания в области современных информационных технологий для решения инженерных задач.
P3	Ставить и решать задачи комплексного анализа, связанные с созданием аппаратно-программных средств информационных и автоматизированных систем, с использованием базовых и специальных знаний, современных аналитических методов и моделей.
P4	Разрабатывать программные и аппаратные средства (системы, устройства, блоки, программы, базы данных и т. п.) в соответствии с техническим заданием и с использованием средств автоматизации проектирования.
P5	Проводить теоретические и экспериментальные исследования, включающие поиск и изучение необходимой научно-технической информации, математическое моделирование, проведение эксперимента, анализ и интерпретация полученных данных, в области создания аппаратных и программных средств информационных и автоматизированных систем.
P6	Внедрять, эксплуатировать и обслуживать современные программно-аппаратные комплексы, обеспечивать их высокую эффективность, соблюдать правила охраны здоровья, безопасность труда, выполнять требования по защите окружающей среды.
Универсальные компетенции	
P7	Использовать базовые и специальные знания в области проектного менеджмента для ведения комплексной инженерной деятельности.
P8	Владеть иностранным языком на уровне, позволяющем работать в иноязычной среде, разрабатывать документацию, презентовать и защищать результаты комплексной инженерной деятельности.
P9	Эффективно работать индивидуально и в качестве члена группы, состоящей из специалистов различных направлений и квалификаций, демонстрировать ответственность за результаты работы и готовность следовать корпоративной культуре организации.
P10	Демонстрировать знания правовых, социальных, экономических и культурных аспектов комплексной инженерной деятельности.
P11	Демонстрировать способность к самостоятельному обучению в течение всей жизни и непрерывному самосовершенствованию в инженерной профессии.

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

Инженерная школа информационных технологий и робототехники
Направление подготовки 09.04.01 Информатика и вычислительная техника
Отделение информационных технологий

УТВЕРЖДАЮ:
Руководитель ООП

(Подпись) _____ (Дата) Спицын В.Г.
(Ф.И.О.)

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

Магистерской диссертации

Студенту:

Группа	ФИО
8BM73	Кузнецовой Анастасии Сергеевне

Тема работы:

Нейросетевые алгоритмы распознавания эмоционального состояния человека по изображениям его лица	
Утверждена приказом директора (номер, дата)	№ 1290/с 19.02.2019

Срок сдачи студентом выполненной работы:	
--	--

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	Задание на выполнение работы
Перечень подлежащих исследованию, проектированию и разработке вопросов	– осуществить обзор предметной области, определить степень проработанности задачи; – изучить и исследовать нейронные сети в области детектирования экспрессий и в области распознавания в целом; – на основе изученного материала предложить собственный алгоритм распознавания эмоционального состояния человека; – выполнить реализацию предложенного алгоритма; – выделить проблемы данной научной области, оценить практическую применимость разработанного алгоритма и обозначить перспективы развития предложенного алгоритма – финансовый менеджмент,

	ресурсоэффективность и ресурсосбережение; – социальная ответственность.
Перечень графического материала	– общая структура разработанной модели;
Консультанты по разделам выпускной квалификационной работы	
Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Меньшикова Екатерина Валентиновна
Социальная ответственность	Алексеев Николай Архипович
Названия разделов, которые должны быть написаны на русском и иностранном языках:	
Разработка алгоритма распознавания эмоционального состояния человека, эксперименты и результаты	

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

Задание выдал руководитель:

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

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8BM73	Кузнецова Анастасия Сергеевна		

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

Инженерная школа информационных технологий и робототехники
Направление подготовки 09.04.01 Информатика и вычислительная техника
Уровень образования магистратура
Отделение информационных технологий
Период выполнения осенний / весенний семестр 2018/2019 учебного года

Форма представления работы:

Магистерская диссертация

**КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН
выполнения выпускной квалификационной работы**

Срок сдачи студентом выполненной работы:

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
25.10.18	Аналитический обзор	15
25.01.19	Разработка алгоритма распознавания эмоционального состояния человека	30
25.03.19	Программная реализация разработанного алгоритма	20
30.04.19	Тестирование реализованных программных средств	15
23.05.19	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	10
25.05.19	Социальная ответственность	10

Составил преподаватель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ	Спицын В.Г.	д.т.н.		

СОГЛАСОВАНО:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ	Спицын В.Г.	д.т.н.		

РЕФЕРАТ

Выпускная квалификационная работа содержит 108 страниц, 19 рисунков, 37 таблиц, 44 источника и 2 приложения.

Ключевые слова: нейронные сети, сверточные нейронные сети, классификация, распознавание эмоций, машинное обучение.

Объектом данного исследования является метод распознавания эмоций человека. Предметом исследования являются информационные модели сверточных нейронных сетей, а также описание алгоритма распознавания двигательной активности лица для решения практической задачи.

Научная новизна магистерской диссертации заключается в сокращении числа параметров и времени обучения нейронной сети для определения эмоционального состояния человека с точностью не менее 97%.

Цель работы – разработка нейросетевого алгоритма распознавания эмоционального состояния человека по изображениям его лица.

В процессе исследования были изучены и проанализированы существующие алгоритмы и средства распознавания эмоций человека по изображению его лица, применимые в задаче распознавания эмоционального состояния.

В результате исследования был предложен и программно-реализован собственный алгоритм распознавания эмоций.

Основные конструктивные, технологические и технико-эксплуатационные характеристики: программное приложение, созданное при помощи библиотек OpenCV, TensorFlow, Keras на языке программирования Python в редакторе исходного кода Visual Studio Code.

Область применения: робототехника, системы безопасности, определение эмоционального состояния человека.

Значимость работы заключается в практической разработке алгоритма распознавания эмоций с использованием сверточных нейронных сетей.

В перспективе распознавания эмоциональной составляющей человеческого лица предполагается создание комплекса алгоритмов, которые

бы учитывали все тонкости. Голосовое детектирование вместе с трекингом изменений окружающей среды, распознавание основных психологических телодвижений индивидуума и все это в совокупности с обычным эмоциональным детектором значительно увеличили бы точность любых программных средств, используемых для решения данной проблемы машинного обучения.

ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ, СОКРАЩЕНИЯ И НОРМАТИВНЫЕ ССЫЛКИ

ANN (Artificial neural networks) – искусственные нейронные сети.

CNN (convolutional neural network) – сверточные нейронные сети.

CKP (Cohn-Kanade Plus) – база данных выражений лиц, предназначенная для исследований в области автоматического анализа и синтеза изображений лица и для изучения восприятия.

MMI Facial Expression Database – база данных, предоставляющая большие объемы визуальных данных о выражениях лица.

Кросс-валидация – это метод проверки разработанной модели, чтобы оценить насколько статистический анализ, используемый в этой модели, способен работать на независимом наборе данных. Одна итерация кросс-валидации включает в себя разделение набора данных на части, затем построение модели на одной части (также называемой тренировочным набором), и проверка модели на другой части (также называемой тестовым набором).

AAM (Active Appearance Model) – модель активного внешнего вида. Это алгоритм компьютерного зрения для сопоставления статистической модели формы и внешнего вида объекта с новым изображением.

Candide – это параметризованная маска для лица, специально разработанная для кодирования человеческих лиц на основе моделей. Малое количество полигонов (около 100) позволяет быстро восстанавливать лицо с умеренной вычислительной мощностью. Candide контролируется глобальными и локальными единицами действия (AU). Глобальные соответствуют поворотам вокруг трех осей. Локальные блоки действий управляют мимикой лица, так что можно получить различные выражения.

SVM (Support Vector Machine) – машина опорных векторов.

OpenCV (Open Source Computer Vision Library) – библиотека алгоритмов компьютерного зрения, обработки изображений и численных алгоритмов общего назначения с открытым кодом.

Оглавление

ВВЕДЕНИЕ	12
1 АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ	14
1.1 Описание предметной области.....	14
1.2 Обзор существующих алгоритмов, применяемых для распознавания эмоций	15
1.2.1 GoogLeNet	15
1.2.2 Happy and Routray алгоритм.....	16
1.2.3 Byeon and Kwak алгоритм	16
1.2.4 Song алгоритм.....	16
1.2.5 Luesy алгоритм	16
1.2.6 Anderson алгоритм.....	17
1.2.7 Kotsia и Pitas алгоритм.....	17
1.2.8 Shan алгоритм	17
1.2.9 Zafar алгоритм	18
1.3 Теоретические основы	18
1.3.1 Свертка блоком 1 x 1.....	18
1.3.2 LRN	19
1.3.3 ReLU	20
1.3.4 10-кратная перекрестная проверка	20
1.3.5 SVM	21
1.3.6 Латентное размещение Дирихле (LDA).....	21
1.3.7 Сверточные нейронные сети	22
2 РАЗРАБОТКА АЛГОРИТМА РАСПОЗНАВАНИЯ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ЧЕЛОВЕКА.....	25
2.1 Архитектура разработанной модели	25
2.2 Используемые наборы данных.....	30
2.2.1 Набор данных MMI.....	30
2.2.2 Набор данных СКР.....	31
2.3 Обучение модели	32
3 ПРОГРАММНАЯ РЕАЛИЗАЦИЯ	33

3.1	Использованные технологии	33
3.1.1	Keras.....	33
3.1.2	OpenCV.....	34
3.1.3	TensorFlow.....	34
3.1.4	Dlib.....	34
3.2	Описание функционала реализованного алгоритма	35
3.3	Эксперименты и результаты.....	36
4	РЕЗУЛЬТАТЫ РАБОТЫ	40
5	ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ.....	43
5.1	Потенциальные потребители результатов исследования.....	45
5.2	Сегментирование рынка	46
5.3	Оценка конкурентоспособности технических решений	46
5.4	FAST-анализ.....	47
5.4.1	Выбор объекта FAST-анализа.....	48
5.4.2	Описание главной, основных и вспомогательных функций, выполняемых объектом.....	48
5.4.3	Определение значимости выполняемых функций объектом	48
5.4.4	Анализ стоимости функций выполняемых объектом исследования	50
5.4.5	Построение функционально-стоимостной диаграммы объекта и её анализ	51
5.4.6	Оптимизация функций выполняемых объектом.....	51
5.5	SWOT-анализ	52
5.6	Оценка готовности проекта к коммерциализации	54
5.7	Методы коммерциализации результатов научно-технического исследования	55
5.8	Инициация проекта	56
5.8.1	Цели и результат проекта	56
5.8.2	Организационная структура проекта	58
5.8.3	Ограничения и допущения проекта.....	58
5.9	Планирование управления научно-техническим проектом.....	59
5.9.1	График проведения и бюджет НТИ.....	59

5.10	Бюджет НТИ	62
5.11	Реестр рисков проекта.....	68
5.12	Оценка ресурсной, финансовой и экономической эффективности НТИ	69
6	СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ	74
6.1	Производственная безопасность	76
6.1.1	Анализ выявленных вредных факторов при разработке и эксплуатации проектируемого решения.....	77
6.1.2	Анализ выявленных опасных факторов при разработке и эксплуатации проектируемого решения.....	85
6.2	Экологическая безопасность	87
6.2.1	Анализ влияния процесса исследования на окружающую среду	87
6.2.2	Обоснование мероприятий по защите окружающей среды.....	87
6.3	Защита в чрезвычайных ситуациях	88
6.3.1	Анализ вероятных ЧС, которые могут возникнуть на рабочем месте при проведении исследований.....	88
6.3.2	Обоснование мероприятий по предотвращению ЧС и разработка порядка действия в случае возникновения ЧС	89
6.4	Правовые и организационные вопросы обеспечения безопасности	90
6.4.1	Специальные правовые нормы трудового законодательства.....	90
6.4.2	Организационные мероприятия при компоновке рабочей зоны исследователя	92
	ЗАКЛЮЧЕНИЕ	95
	СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	96
	ПРИЛОЖЕНИЕ А (справочное).....	100
	ПРИЛОЖЕНИЕ Б CD-ДИСК.....	108

ВВЕДЕНИЕ

В последние годы распознавание эмоций стало становиться популярным не только в методах и специфике психологии и социологии.

С учетом всемирной и всепоглощающей идеи об искусственном интеллекте ученые и программисты всего мира пришли к консенсусу необходимости достаточно точного определения эмоций с помощью программных средств.

Распознавание номеров у машин уже вошло в обиход и активно используется правоохранительными органами, детектирование лиц с помощью программ и компьютера уже давно не новинка, так как во многих крупных компаниях стоит система безопасного входа по идентификации сотрудника посредством фотографий.

Но один из самых важных прорывов в сфере безопасности был совершен при введении на особо важные объекты камер с распознаванием отрицательных, враждебно настроенных эмоций, которые помогут вычислить преступника, грабителя или террориста. К тому же, распознавание эмоций может быть применено в робототехнике. Робот может правильно реагировать на состояние человека, при этом сделав взаимодействие человека с роботом более естественным и интеллектуальным.

Особенно актуально правильно определить эмоциональное состояние человека в ситуациях, которые могут быть опасны для жизни. В частности, некоторые современные автомобили содержат систему определения усталости водителя. Такого рода системы позволяют в некоторых случаях избежать аварий, вызванных плохим самочувствием водителя, его невнимательностью или сонливостью. Анализ осуществляется на основе результатов обработки изображений лица человека, полученных с видеокамеры.

Исходя из этого, видно, что данное направление не стоит на месте и успешно развивается. На данный момент значительную нишу в этом научном направлении занимает машинное обучение, в частности нейронные сети. В связи с вышеизложенным, высокую степень актуальности имеет задача

разработки нейросетевого алгоритма распознавания эмоционального состояния человека по изображениям его лица.

Объектом данного исследования являются подходы и методы распознавания мимических выражений лица. Предметом исследования являются модели сверточных нейронных сетей для распознавания эмоций.

Целью работы является разработка нейросетевого алгоритма для определения эмоционального состояния человека по изображению его лица.

Для достижения цели работы необходимо решить следующие задачи:

- осуществить обзор предметной области, определить степень проработанности задачи;
- изучить и исследовать нейронные сети в области детектирования экспрессий и в области распознавания в целом;
- на основе изученного материала предложить собственный алгоритм распознавания эмоционального состояния человека;
- выполнить реализацию предложенного алгоритма;
- выделить проблемы данной научной области, оценить практическую применимость разработанного алгоритма и обозначить перспективы развития предложенного алгоритма.

1 АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ

1.1 Описание предметной области

Люди используют различные формы коммуникации, такие как речь, жесты и эмоции. Распознавание эмоций других людей является важным фактором для правильного и надлежащего понимания друг друга.

С продолжающимися исследованиями в области робототехники, особенно в области человекоподобных роботов, становится интересным интегрировать эти возможности в машины, тем самым делая способ коммуникации более разнообразным и естественным.

Помимо человеко-компьютерного взаимодействия, распознавание эмоций может быть применено и в других областях, таких как видеонаблюдение в системах безопасности или обеспечение безопасности водителя. Возможность определить настроение и эмоциональное состояние водителя может помочь в оценке уровня его внимания, что позволит автоматическим системам адаптироваться.

Ещё одной интересной сферой является реклама, а точнее маркетинговые исследования. Определение реакции покупателя на определенные товары позволит определить его потребности и предлагать товары, наиболее ему интересные.

Задача распознавания эмоций имеет множество подходов. Основное различие между традиционной искусственной нейронной сетью (ANN) и CNN состоит в том, что только последний слой CNN полностью связан, тогда как в ANN каждый нейрон связан с любыми другими нейронами. Сверточные нейронные сети представляют собой особый вид ANN, и было показано, что он хорошо работает в качестве инструмента извлечения признаков при использовании изображений в качестве входных данных [1] и способен работать в режиме реального времени. Это позволяет использовать необработанные входные изображения без какой-либо предобработки или постобработки.

1.2 Обзор существующих алгоритмов, применяемых для распознавания эмоций

Анализ распознавания эмоций, является сложной и комплексной задачей, поэтому существует огромное количество методов и алгоритмов, рассмотренных в работах разных авторов. Некоторые методы очень спорны и имеют существенные недостатки, так как авторы пытаются найти все более изощренный метод, который докажет свою состоятельность и выявит недостатки предыдущих алгоритмов.

Есть некоторые особенности лица человека, из-за которых не существует на данный момент эталонного способа решения задачи. Часто многие методы не учитывают аксессуаров и лицевой растительности, таких, как бороды, усов, неравномерных бровей и объемных причесок, закрывающих или мешающих распознаванию. К тому же, эмоция очень абстрактное выражение, которое очень изменяется от человека к человеку. Часто мы сталкиваемся с проблемой непонимания некоторой эмоции, например, открытый рот не всегда означает удивление. Однако, основной недостаток, как было сказано выше, отсутствие эталонной обучающей выборки, на основе которой создается алгоритм.

Описанные ниже алгоритмы будут использованы для сравнения с разработанным алгоритмом.

1.2.1 GoogLeNet

Christian Szegedy предложил архитектуру под названием GoogLeNet. Это 27-уровневая сеть, состоящая в основном из CNN. Сеть обучается с использованием стохастического градиентного спуска. В Классификационном конкурсе ILSVRC 2014 эта сеть продемонстрировала себя отлично, заняв первое место. Изюминка заключалась в использовании свёрточных блоков 1x1 (NiN) для уменьшения количества свойств перед подачей в «дорогие» параллельные блоки. Подробно преимущество применения таких сверточных блоков будет описано в разделе «Теоретические основы» [2].

1.2.2 Happy and Routray алгоритм

Используя Extended Cohn-Kanade Dataset, Happy and Routray [3] проводят различие между шестью основными эмоциями. Учитывая входное изображение, их решение локализует область лица. Из этого региона, например, пятна на лице, глаза или губы, обнаруживаются, и точки интереса отмечаются. Из патчей, которые имеют наибольшую разницу между двумя изображениями, извлекаются особенности. Размерность элементов уменьшается, а затем передается машине опорных векторов (SVM). Для оценки метода применяется 10-кратная перекрестная проверка. Средняя точность составляет 94,09%.

1.2.3 Byeon and Kwak алгоритм

Распознавание эмоций на основе видео было предложено Byeon и Kwak [4]. Они разработали трехмерную CNN, которая использует группы из пяти последовательных кадров в качестве входных данных. База данных, построенная с использованием эмоций десяти человек, была использована для достижения точности 95%.

1.2.4 Song алгоритм

Song и соавторы использовали глубокую сверточную нейронную сеть для изучения выражений лица. Созданная сеть состоит из пяти слоев с общим количеством нейронов в 65 тысяч. Слои свертки, пулинга, локальные слои фильтра и полносвязный слой используются для достижения точности 99,2% в наборе СКР. Чтобы избежать переобучения, был использован метод прореживания [5].

1.2.5 Luesu алгоритм

Luesu и соавторы разработали набор Extended CohnKanade. Этот набор данных содержит аннотации эмоций. Что касается классификации, они также оценили наборы данных, использующие активные модели внешнего вида (AAM) в сочетании с SVM. Чтобы найти положение и отследить лицо на разных изображениях, они использовали AAM, который создает сетку лица. Из этой сетки они извлекли два вектора признаков. В первом содержались

нормализованные вершины относительно поворота, перемещения и масштаба. Во втором были извлечены серое изображение из данных сетки и входные изображения. Они выбрали стратегию перекрестной проверки, при которой один предмет не учитывается в процессе обучения, и точность достигает 80% [6].

1.2.6 Anderson алгоритм

Anderson и соавторы разработали систему выражения лица, способную распознавать шесть основных эмоций. Их система построена на трех компонентах. Первый – это трекер лица (производная от шаблона отношения) для определения местоположения лица. Второй компонент представляет собой алгоритм оптического потока для отслеживания движения внутри лица. Последний компонент – это сам механизм распознавания. Он основан на машинах опорных векторов и многослойных персептронах. Этот подход был реализован в EmotiChat. Алгоритм достигает точности распознавания 81,82% [7].

1.2.7 Kotsia и Pitas алгоритм

Kotsia и Pitas обнаруживают эмоции, нанося на лицо сетку Candide, маску лица с небольшим количеством полигонов. Сетка изначально размещается на изображении случайным образом, затем ее необходимо вручную разместить на лице человека. На протяжении всей эмоции сетка отслеживается с помощью трекера KanadeLucasTomasi. Информация о геометрическом смещении, предоставляемая сеткой, используется как вектор признаков для мультиклассовых SVM. Распознаваемые эмоции: гнев, отвращение, страх, счастье, грусть и удивление. Они оценивали модель по набору данных Cohn-Kanade, и точность была достигнута на уровне 99,7%[8].

1.2.8 Shan алгоритм

Shan разработал систему, основанную на локальных двоичных шаблонах (LBP). LBP рассчитываются по лицевой области. Из извлеченных LBP получается вектор признаков. Характеристики зависят от положения и

размера субрегионов, в которых рассчитывается LBP. AdaBoost используется для поиска областей изображений, которые содержат наиболее различительную информацию. Были оценены различные алгоритмы классификации, из которых SVM с функциями Boosted-LBP работает лучше всего с точностью распознавания 95,1% в наборе СКР [9].

1.2.9 Zafar алгоритм

В 2013 году Zafar и соавторы [10] предложили систему распознавания эмоций с использованием робастной нормализованной взаимной корреляции (NCC). Используемый NCC – это «корреляция как измененная шкала дисперсии разницы между стандартизированными оценками». Внешние пиксели, которые влияют на соответствие шаблона слишком сильному или слишком слабому, исключаются и не учитываются. Этот подход был оценен в различных базах данных, включая AR FaceDB (85% точность распознавания) и Extended Cohn Kanade Database (90% точность распознавания).

1.3 Теоретические основы

1.3.1 Свертка блоком 1 x 1

Свертка 1×1 была представлена NIN [11] и используется с ReLU. Таким образом, изначально NIN использует её для введения большей нелинейности, чтобы увеличить мощность представления сети, поскольку авторы в NIN считают, что данные представлены в нелинейной форме. В GoogLeNet свертка 1×1 используется как модуль уменьшения размеров для сокращения вычислений. Уменьшая узкое место вычислений, можно увеличить глубину и ширину.

Приведем пример. Предположим, что нужно выполнить свертку 5×5 без использования свертки 1×1 , как показано ниже:

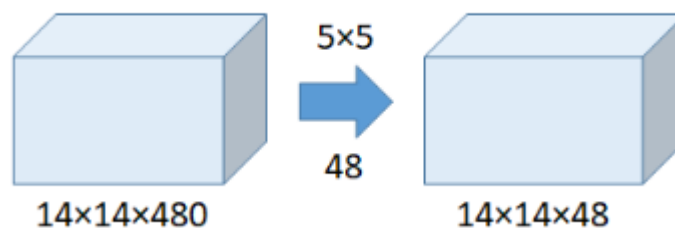


Рисунок 1.1 – Свертка без блока 1 x 1

Тогда число операций будет равно:

$$N = (14 \times 14 \times 48) \times (5 \times 5 \times 480) = 112.9M.$$

Теперь используем сверточный блок 1 x 1:

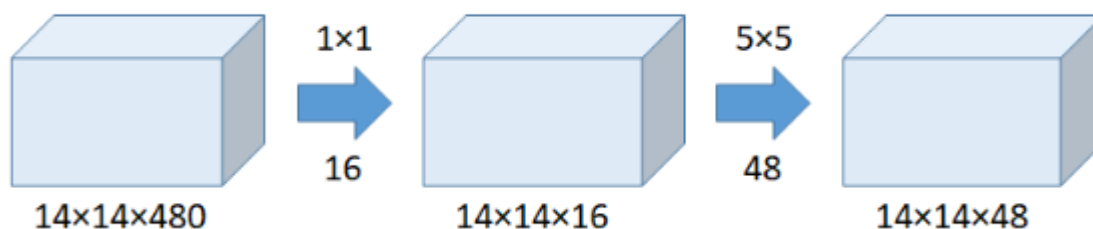


Рисунок 1.2 – Свертка с использованием блока 1 x 1

Число операций для свертки блоком 1x1 = $(14 \times 14 \times 16) \times (1 \times 1 \times 480) = 1.5M$

Число операций для свертки блоком 5x5 = $(14 \times 14 \times 48) \times (5 \times 5 \times 16) = 3.8M$

Общее число операций = $1.5M + 3.8M = 5.3M$

Хорошо видно, что удалось сократить число операций в 22 раза.

1.3.2 LRN

Local Response Normalization (LRN) – слой локальной нормализации.

Данная нормализация используется для предотвращения замедления скорости обучения параметров сети.

Local response normalization является самостоятельным слоем сети. Данная операция нормализует каждое значение входной матрицы по каналу, то есть применяется следующая формула:

$$b_{x,y}^i = \frac{a_{x,y}^i}{\left(k + \alpha \sum_{j=\max(0, i-\frac{n}{2})}^{\min(N-1, i+\frac{n}{2})} (a_{x,y}^j)^2 \right)^\beta} \quad (1)$$

где $a(i, x, y)$ представляет i -й вывод ядра свертки (после ReLU) в позиции (x, y) на карте объектов.

$b(i, x, y)$ представляет собой результат нормализации локального отклика и, конечно, это также вход для следующего слоя.

N – номер свертки ядра.

n – смежный номер ядра свертки, этот номер зависит от автора.

k, α, β - гиперпараметры, обычно $k = 2, \alpha = 10e-4, \beta = 0,75$.

1.3.3 ReLU

ReLU – функция активации под названием «выпрямитель». ReLU имеет следующую формулу $f(x) = \max(0, x)$ и реализует простой пороговый переход в нуле.

Стоит обратить внимание на понятие разреженности активации.

Пусть существует большая нейронная сеть с множеством нейронов. Использование сигмоиды или гиперболического тангенса будет влечь за собой активацию всех нейронов аналоговым способом. Это означает, что почти все активации должны быть обработаны для описания выхода сети. Другими словами, активация плотная, а это затратно. В идеале необходимо, чтобы некоторые нейроны не были активированы, это сделало бы активации разреженными и эффективными.

ReLU позволяет это сделать. Пусть есть сеть со случайно инициализированными весами (или нормализованными), в которой примерно 50% активаций равны 0 из-за характеристик ReLU (возвращает 0 для отрицательных значений x). В такой сети включается меньшее количество нейронов (разреженная активация), а сама сеть становится легче.

1.3.4 10-кратная перекрестная проверка

Перекрестная проверка – это метод оценки прогностических моделей путем разделения исходного образца на обучающий набор для обучения модели и тестовый набор для ее оценки.

В k -кратной перекрестной проверке исходная выборка случайным образом разбивается на k подвыборок одинакового размера. Из k подвыборок одна подвыборка сохраняется в качестве данных проверки для тестирования модели, а оставшиеся $k-1$ подвыборки используются в качестве обучающих данных. Затем процесс перекрестной проверки повторяется k раз (сгибы), причем каждый из k подвыборок используется ровно один раз в качестве данных проверки. Результаты k из сгибов могут затем быть усреднены (или иным образом объединены), чтобы произвести единственную оценку. Преимущество этого метода заключается в том, что все наблюдения используются как для обучения, так и для проверки, а каждое наблюдение используется для проверки только один раз.

В работе упоминается перекрестная проверка, где $k = 10$.

1.3.5 SVM

Метод опорных векторов – это особый класс алгоритмов, который характеризуется использованием ядер, отсутствием локальных минимумов, и используется для решения задач классификации и регрессии.

Основная идея метода заключается в построении гиперплоскости, разделяющей объекты выборки наиболее оптимальным способом. Алгоритм работает в предположении, что чем больше расстояние (зазор) между разделяющей гиперплоскостью и объектами разделяемых классов, тем меньше будет средняя ошибка классификатора.

1.3.6 Латентное размещение Дирихле (LDA)

Латентное размещение Дирихле – порождающая модель, применяемая для машинного обучения и информационного поиска, которая позволяет объяснять результаты наблюдений с помощью неявных групп, что предоставляет возможность выявления причин сходства некоторых частей данных. Примером могут служить слова, собранные в документы. Они являются наблюдениями, и утверждается, что каждый документ состоит из

смеси небольшого количества тем и что появление каждого слова связано с одной из тем документа.

Модель LDA используется для решения классической задачи анализа текстов: создание вероятностной модели большой коллекции текстов [12].

1.3.7 Сверточные нейронные сети

При обработке изображений и, в частности, распознавания объектов на фотографиях возникает проблема большого объема данных. К примеру, самая распространенная база данных с изображениями содержит несколько миллионов картинок с размерами 256×256 и, записанными в RGB формате. Обычная нейронная сеть будет невероятно долго обучаться даже на супермощных компьютерах и, вероятно, придет к проблеме переобучения, которая заключается в отсутствии возможности воспринимать и классифицировать объекты в общем. Для примера с картинками из ImageNet в полностью связанной архитектуре нейронной сети каждый нейрон будет иметь порядка $256 * 256 * 3 = 196608$ весов. Для полноценной классификации необходимо иметь как минимум несколько таких нейронов. Решением данной ситуации является специальная архитектура искусственной нейронной сети, так называемые сверточные нейронные сети.

Сверточные нейронные сети имеют более подходящую архитектуру для работы с изображениями. Слои этой нейронной сети расположены в трех измерениях, которые характерны для картинки: высота, ширина и глубина (в отношении изображений это Red, Green, Blue). Нейроны в слое связаны только с некоторой областью в предшествующем слое, а не со всеми как в полносвязной сети.

Основные слои, которые используются в сверточной сети это слой свертки, слой линейной ректификации, слой субдискретизации и полносвязный слой. У каждого из них своя функция и задача, но в итоге мы получаем сложно структурированную сеть с наилучшими результатами для классификации объектов на изображении. На рисунке 1.3 подробно иллюстрирована архитектура CNN [13]. Рассмотрим подробнее каждый из слоев:

1. Входной

Первый вход у любого типа нейронной сети входной. Этот слой хранит необработанные значения пикселей входного изображения. В случае с изображением это некоторое количество значений по ширине, высоте и глубине.

2. Слой свертки

Сверточный слой – это основной слой в CNN. Данный слой опирается на операции свертки, которая заключается в том, что по всем входным данным пробегается матрица весов ограниченного и небольшого размера, создающая новое значение для некоторого числа пикселей. Такая матрица называется ядром свертки, ее особенностью является то, что весовые коэффициенты такой матрицы определяются во время обучения и заранее неизвестны. К необходимости использования этого слоя можно отнести то, что он значительно уменьшает количество весов, созданных для нейронов, в отличие от полносвязной нейронной сети.

3. Слой с блоком линейной ректификации(ReLU)

Данный слой содержит функцию активацию после операции свертки.

В сверточной нейронной сети используется вместо привычных гиперболической или сигмоидальной функции активации, функции $f(x) = \max(0, x)$. Эта активационная функция показала очень хорошие результаты на задаче классификации изображений. Более того, выбор был сделан на этой функции активации по причине того, что для большого объема данных она позволяет обучаться сети за адекватное время.

4. Слой субдискретизации

На данной этапе происходит пулинг полученных данных, то есть подвыборка. На этой стадии матрица характеристик, она же матрица со значениями пикселей уплотняется до матрицы меньшего размера. На примере изображений пулинг уменьшается размерность по ширине и высоте. Такое преобразование делается на уникальных и непересекающихся областях, обычно из прямоугольника 2×2 выбирается один пиксель, значение которого

максимальное. Тем самым происходит удаление незначительных деталей и разрешение задачи переобучения. Существуют субдискретизации не только с функцией максимизации, но еще и функцией среднего или L2 нормирования.

5. Полносвязный слой (fullyconnected layer)

Данный слой представляет собой обычную нейронную сеть, где каждый нейрон связан со всем предыдущими. Предназначение такого слоя в том, чтобы вычислить уже по имеющимся данным вероятность принадлежности к классу. Значительно сократив объем данных по сравнению с исходным, сеть передает эти данные в полносвязную нейронную сеть, в которой может быть еще несколько слоев внутри.

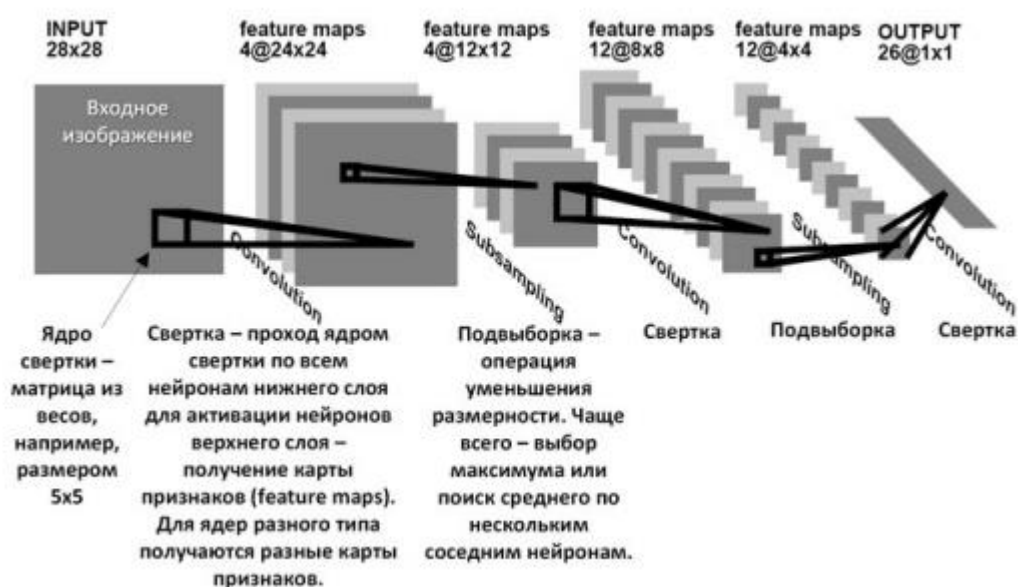


Рисунок 1.3 – Архитектура CNN

У сверточной нейронной сети есть огромные преимущества в распознавании объектов:

- высокая точность при распознавании;
- наилучший выбор при работе с изображениями;
- удобное распараллеливание, и как следствие возможность обучения и реализации на графическом процессоре;
- обучается такая сеть при использовании метода обратного распространения ошибки.

Из недостатков можно смело выделить неопределенность структуры,

многие параметры необходимо настраивать под задачу, то есть эмпирическим путем, методом проб и ошибок, а обучение сети на огромных наборах данных может длиться неделями.

2 РАЗРАБОТКА АЛГОРИТМА РАСПОЗНАВАНИЯ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ЧЕЛОВЕКА

2.1 Архитектура разработанной модели

Предлагаемая архитектура глубоких сверточных нейронных сетей (изображена на рисунке 2.1) состоит из четырех частей.

Первая часть представляет собой автоматическую предобработку данных. Данный этап начинается со сверточного слоя, который применяет к изображению 64 различных фильтра.

Следующий слой – пулинг, который понижает размерность изображений, которые затем нормируются LRN.

Следующие шаги – два блока FeatEx (Parallel Feature Extraction Block), выделенные на рисунке 2.1. Они являются ядром предложенной архитектуры и будут описаны позже в этом разделе. Признаки, выделенные этими блоками перенаправляются на полно-связный слой, который использует их для классификации эмоций.

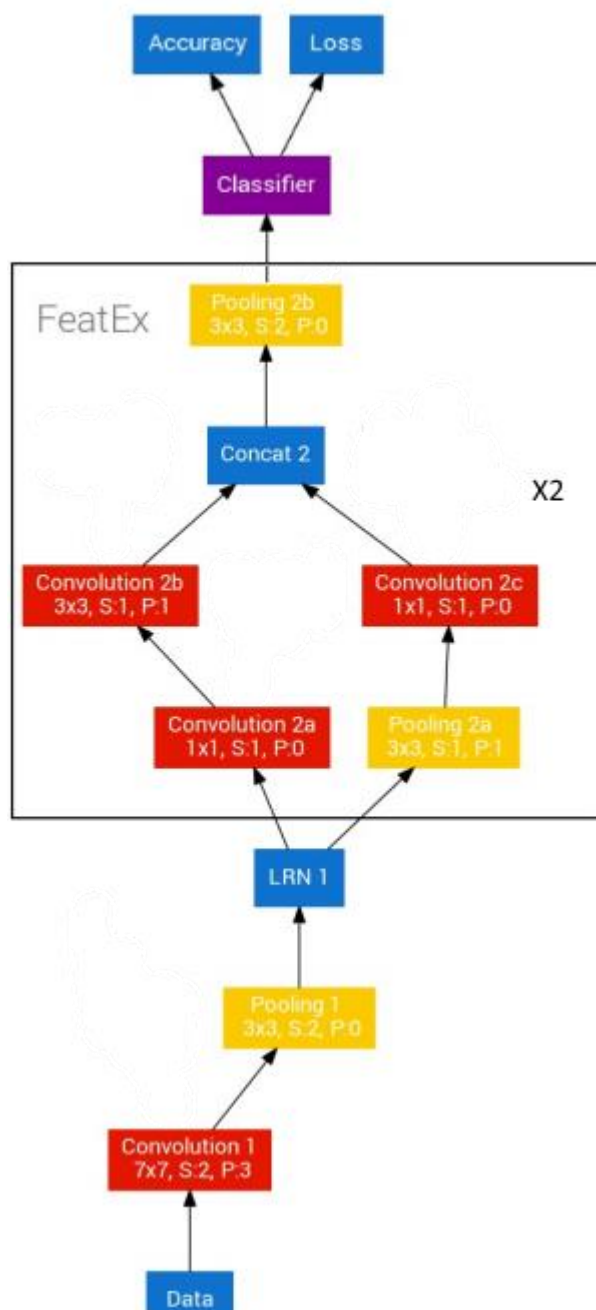


Рисунок 2.1 – Архитектура разработанной модели

Ключевая структура в построенной архитектуре – блок параллельного извлечения признаков (FeatEx). Идея о его использовании связана с успехом GoogleNet. В данной сети использовалась свертка размером 1 x 1 как модуль уменьшения размеров сети для сокращения вычислений. Уменьшая узкое место вычислений можно увеличить ширину и глубину сети. Именно это и послужило вдохновением для создания предлагаемой архитектуры. Оригинальный GoogleNet содержит около 6 миллионов параметров. Предлагаемая модель содержит 417120 параметров.

FeatEx блок состоит из слоев свертки, пулинга и ReLU. Первый сверточный слой в FeatEx уменьшает размерность, поскольку он свертывается с фильтром размером 1×1 . Он усиливается ReLU, который создает желаемую разреженность. Выходной сигнал затем свернут с фильтром размером 3×3 . В параллельном пути слой Max Pooling используется для уменьшения информации перед применением свертки с ядром размера 1×1 . Применение фильтров разного размера отражает различные масштабы, на которых могут появляться лица. Результаты двух параллельных путей объединены для более разнообразного представления входных данных.

Использование блока FeatEx дважды показало хорошие результаты.

Разные слои модели создают свои векторы признаков, которые представлены на рисунке 2.2.

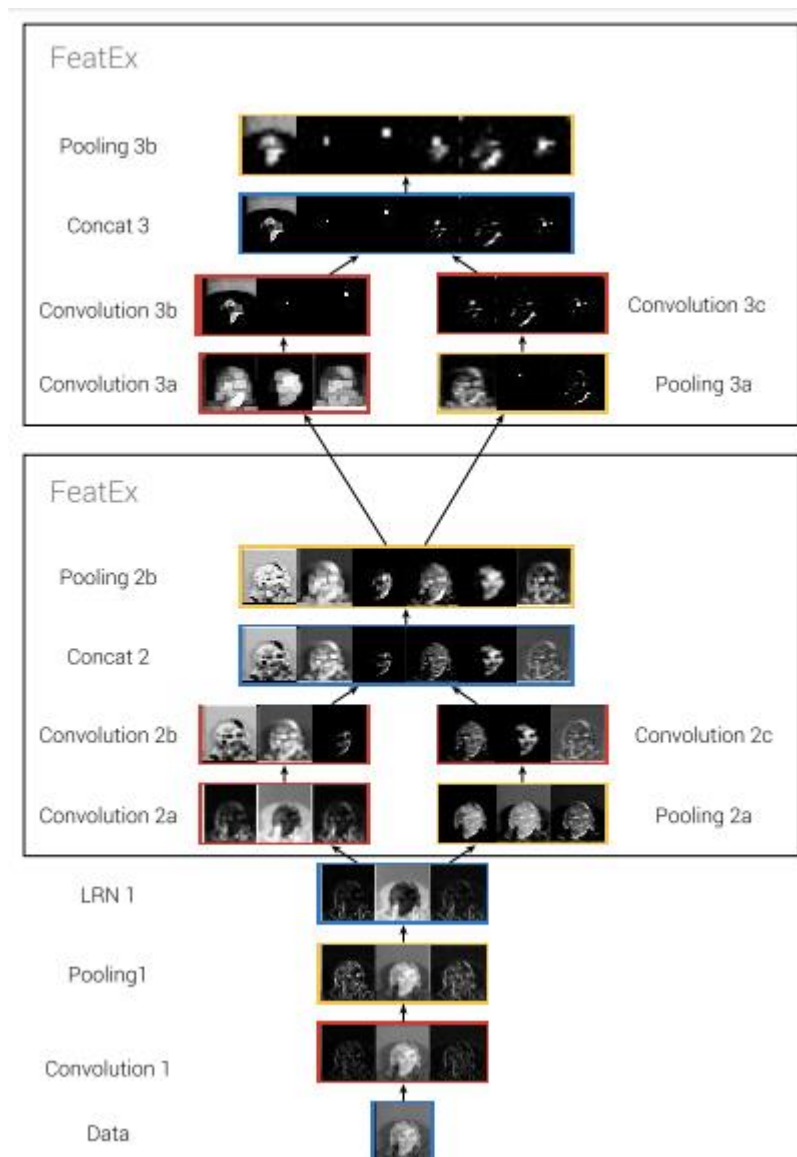


Рисунок 2.2 – Пример визуализации работы различных слоев. Данные взяты из MMI набора данных

Первая часть длится до тех пор, пока LRN 1 не обработает данные и не создаст несколько измененных экземпляров входных данных. На данном этапе будут видны только ребра с низким уровнем абстракции. Первый блок FeatEx создает два параллельных пути извлечения признаков объектов в разном масштабе, которые будут объединены в Concat 2.

Второй блок FeatEx уточняет представление признаков, а также уменьшает размерность.

Эта визуализация доказывает, что объединение двух блоков FeatEx является правильным подходом для создания абстрактного представления признаков.

Выходную размерность каждого слоя можно увидеть в таблице 2.1.

Таблица 2.1 – Размерность на выходе каждого слоя

Слой	Размер на выходе слоя
Data	224×224
Convolution 1	$64 \times 112 \times 112$
Pooling 1	$64 \times 56 \times 56$
LRN 1	$64 \times 56 \times 56$
Convolution 2a	$96 \times 56 \times 56$
Convolution 2b	$208 \times 56 \times 56$
Pooling 2a	$64 \times 56 \times 56$
Convolution 2c	$64 \times 56 \times 56$
Concat 2	$272 \times 56 \times 56$
Pooling 2b	$272 \times 28 \times 28$
Convolution 3a	$96 \times 28 \times 28$
Convolution 3b	$208 \times 28 \times 28$
Pooling 3a	$272 \times 28 \times 28$
Convolution 3c	$64 \times 28 \times 28$
Concat 3	$272 \times 28 \times 28$
Pooling 3b	$282 \times 14 \times 14$
Classifier	$7 \times 1 \times 1$

Описанная архитектура компактна, что позволяет не только обучать модель быстро, но и использовать её для приложений, работающих в режиме реального времени.

Производительность предлагаемой архитектуры обеспечивает эффективность и надежность использования предлагаемой работы в реальных приложениях.

2.2 Используемые наборы данных

В работе были использованы следующие наборы данных: Extended Cohn-Kanade (СКР) и MMI Facial Expression Database.

2.2.1 Набор данных MMI

Набор данных MMI содержит более 2900 видео и 740 изображений 75 человек. Аннотации содержат единицы действия и эмоции. Набор данных содержит веб-интерфейс со встроенным поиском для сканирования базы данных. Видео и изображения являются цветными. Люди разного возраста, разного пола и имеющие разную этническую принадлежность. В наборе данных представлены шесть основных эмоций: гнев, отвращение, страх, счастье, печаль, удивление.

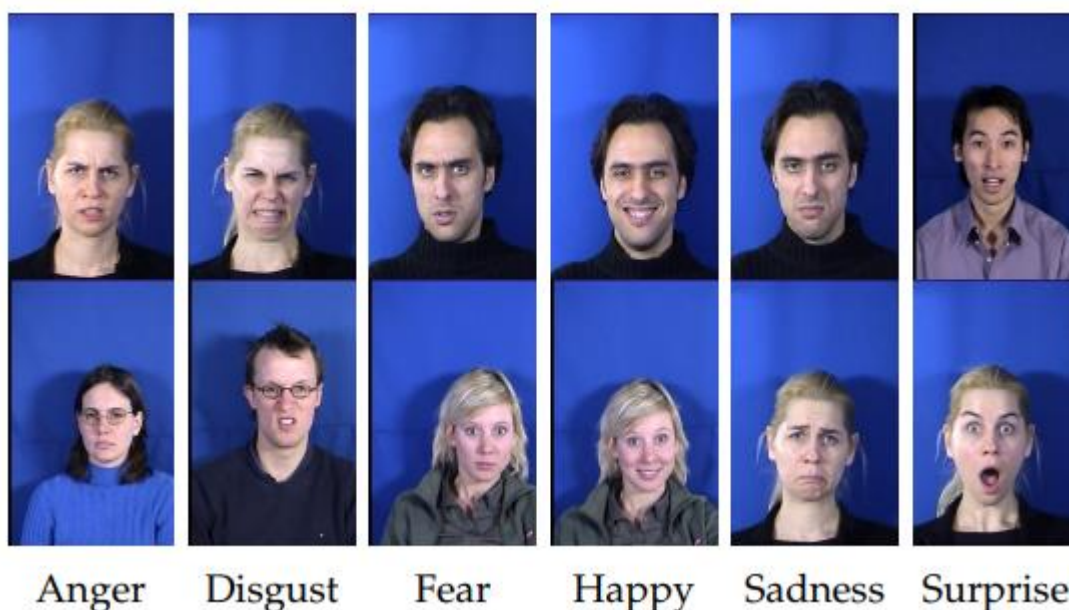


Рисунок 2.3 – Демонстрация различий отображения одних и тех же эмоций в MMI наборе данных

Из каждого видео автоматически было выделено 20 наиболее значимых для отражения сути эмоции кадров. Первые два кадра были отброшены, т.к. они отражают нейтральное выражение. Для выделения этих кадров использовалась разница двух соседних кадров, предварительно переведенных в оттенки серого. Для компенсации шума перед расчетом был применен сглаживающий Гауссов фильтр. После извлечения кадров было получено 3740 изображений.

Для набора данных MMІ в настоящее время наилучшая точность распознавания эмоций составляет 93,33% [14].

2.2.2 Набор данных СКР

Для этого набора данных 210 человек в возрасте от 18 до 50 лет были записаны с изображением эмоции. Присутствуют как женщины, так и мужчины.

Изображения имеют размер 640×490 , а также 640×480 пикселей. Изображения могут быть как в оттенках серого, так и цветные. Всего в этом наборе 593 последовательности, помеченных эмоциями. Представленные эмоции: гнев, отвращение, страх, счастье, печаль, удивление и презрение.

Для лучшей оценки полезно исследовать несколько наборов данных. Данный подход направлен на выяснение того, является ли предлагаемый алгоритм рабочим для эмоций, показанных по-разному.



Рисунок 2.4 – На рисунке отражены различия внутри набора данных Cohn-Kanade Plus (СКР)

В результате сравнения этих наборов данных можно отметить, что эмоция презрение содержится только в наборе СКР (рисунок 2.5).



Рисунок 2.5 –Примеры изображений из MMI (сверху) и СКР (снизу) наборов данных. Эмоции (перечислены слева направо): гнев, грусть, отвращение, счастье, страх, удивление. Эмоция презрение не отображена

Для набора данных СКР в настоящее время лучшая точность для распознавания эмоций составляет точность 99,2% [15].

2.3 Обучение модели

Для обучения модели использовался набор данных MMI Facial Expression Database. Было решено разделить полученные 3740 изображений на обучающую и тестовую выборку в отношении 80% / 20%. Таким образом, для обучения было использовано 2992 изображения.

График зависимости точности модели от количества эпох представлен на рисунке 2.6.

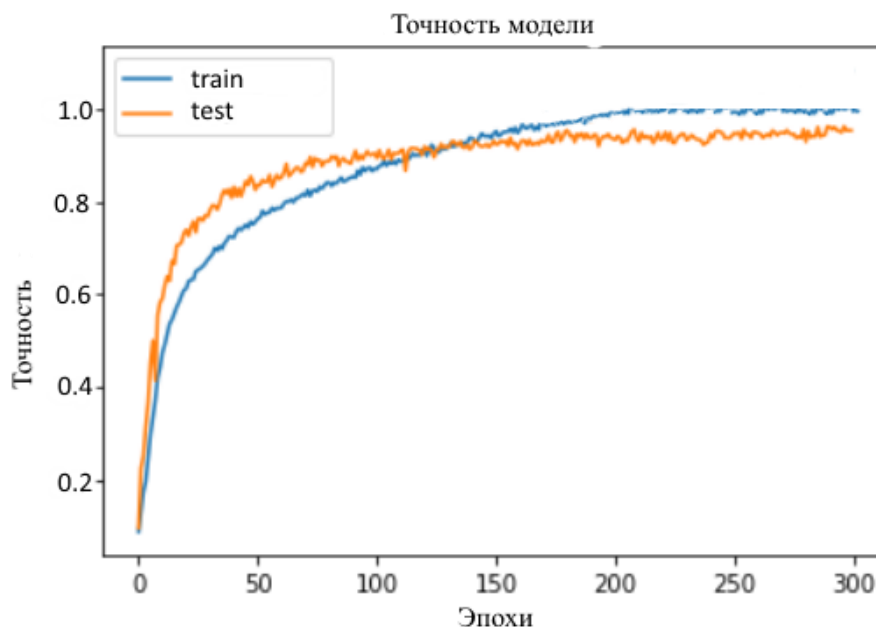


Рисунок 2.6 – График обучения модели

Было замечено, что 130 эпох достаточно для обучения модели. Используя график, можно сделать вывод о переобучении модели после продолжения обучения.

3 ПРОГРАММНАЯ РЕАЛИЗАЦИЯ

3.1 Используемые технологии

3.1.1 Keras

Keras – открытая нейросетевая библиотека, написанная на языке Python. Это библиотека с открытым исходным кодом, которая позволяет легко создавать нейронные сети. Библиотека совместима с TensorFlow, Microsoft Cognitive Toolkit, Theano и MXNet. Tensorflow и Theano являются наиболее часто используемыми численными платформами на Python для разработки алгоритмов глубокого обучения, но они довольно сложны в использовании. Нацелена на оперативную работу с сетями глубинного обучения, при этом спроектирована так, чтобы быть компактной, модульной и расширяемой[16].

Keras занимает второе место в рейтинге используемых фреймворков для глубокого обучения (рисунок 3.1).

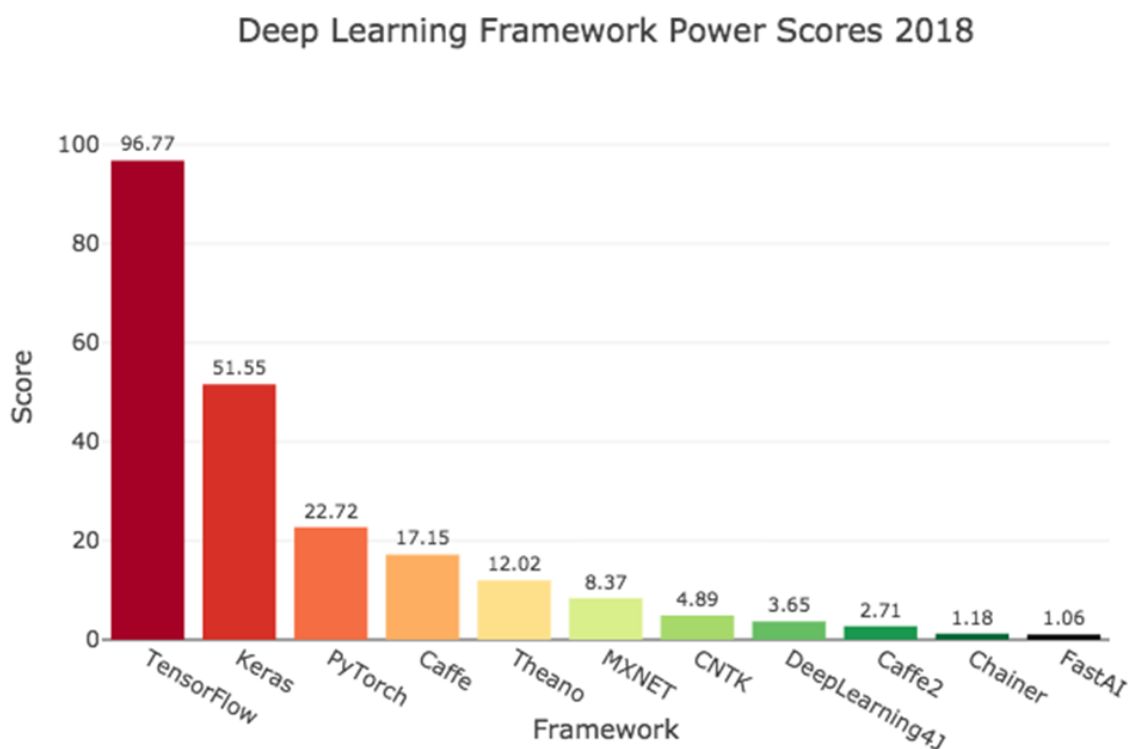


Рисунок 3.1 – Рейтинг используемых фреймворков для глубокого обучения

Исходя из вышеизложенного, было решено использовать популярную связку Keras+TensorFlow для обучения и работы со сверточными нейронными сетями.

3.1.2 OpenCV

OpenCV (англ. *Open Source Computer Vision Library*, библиотека компьютерного зрения с открытым исходным кодом) – библиотека алгоритмов компьютерного зрения, обработки изображений и численных алгоритмов общего назначения с открытым кодом. Реализована на C/C++, также разрабатывается для Python, Java, Ruby, Matlab, Lua и других языков. Может свободно использоваться в академических и коммерческих целях[17].

Также, данная библиотека содержит удобный функционал по работе с изображениями, алгоритмы их обработки, распознавания лиц и пр. Именно поэтому было принято решение об использовании OpenCV.

3.1.3 TensorFlow

TensorFlow – открытая программная библиотека для машинного обучения, разработанная компанией Google для решения задач построения и тренировки нейронной сети с целью автоматического нахождения и классификации образов, достигая качества человеческого восприятия. Основной API был реализован для Python.

Позволяет абстрагироваться от внутренних деталей машинного обучения и сосредоточиться непосредственно на решении своей задачи. Очень мощный фреймворк, позволяет создавать, обучать и использовать нейронные сети любого известного типа[18].

3.1.4 Dlib

Dlib – это универсальная программная библиотека с открытым исходным кодом, написанная на C++. Она известна куда менее, чем OpenCV, но содержит в себе ряд функций: несколько реализаций SVM, несколько

реализаций нейронных сетей, подборку функций матричной геометрии, подборку алгоритмов градиентного спуска, SURF|HOG [19].

Данная библиотека сначала использовалась для распознавания лиц, так как обладает более высокой точностью. Однако, скорость распознавания лиц намного ниже, чем использование алгоритма на основе каскадов Хаара в OpenCV. Применение алгоритма в приложениях, работающих в режиме реального времени было бы невозможным, именно поэтому было решено использовать OpenCV.

3.2 Описание функционала реализованного алгоритма

Реализованный алгоритм был применен в работе в двух режимах: распознавание одиночного изображения и распознавание эмоций в режиме реального времени по видео.

Общий алгоритм работы можно описать следующим образом:

- 1) Захватывается кадр или загружается изображение;
- 2) Происходит предварительная обработка изображения: изменение размера, перевод в градации серого и пр.;
- 3) Предобработанное изображение подается на вход модели, которая возвращает результат в виде вектора, определяющего эмоции;
- 4) Извлекается эмоция с максимальным значением, оно и является наиболее вероятным;
- 5) Рамка с распознанным лицом, эмоция и время её определения демонстрируются пользователю.

3.3 Эксперименты и результаты

Тестирование производилось на ПК следующей конфигурации:

- процессор: Intel Core i7-4500 CPU @ 1.80 Ghz 2.40 Ghz;
- оперативная память: 8 Гб;
- операционная система: Linux Ubuntu.

Для тестирования модели были использованы стандартные наборы данных: Extended Cohn-Kanade (СКР) и часть MMI Facial Expression Database. Одной из причин такого выбора стала их популярность и используемость для тестирования у других авторов и их алгоритмов. Таким образом, кросс-валидация на разных наборах данных позволит объективно оценить качество разработанного алгоритма.

Набор данных СКР

Набор данных СКР является популярным набором данных для проведения оценки работы различных алгоритмов. Чтобы определить, является ли архитектура конкурентоспособной, она была оценена на наборе данных СКР. Для экспериментов все 5870 аннотированных изображений были использованы для 10-кратной перекрестной проверки. Предложенная архитектура оказалась очень эффективной в этом наборе данных со средней точностью 97,6%.

В таблице 3.1 приведены результаты сравнения различных подходов на данном наборе данных.

Таблица 3.1 – Сравнение точности методов на наборе СКР

Автор	Метод	Точность
Aliya Zafar	NCC	90%
Happy et al.	Facial Patches + SVM	94.09%
Lucey et al.	AAM + SVM	≥ 80%
Song et al.	ANN (CNN)	99.2%
	Разработанный метод	97.6%

Точность 90%, представленная автором Zafar, обоснована изображениями, отобранными вручную. Результаты не были валидированы с использованием перекрестной проверки. Оценка классификации представлена на рисунке 3.2. Можно заметить, что некоторые эмоции прекрасно распознаются.

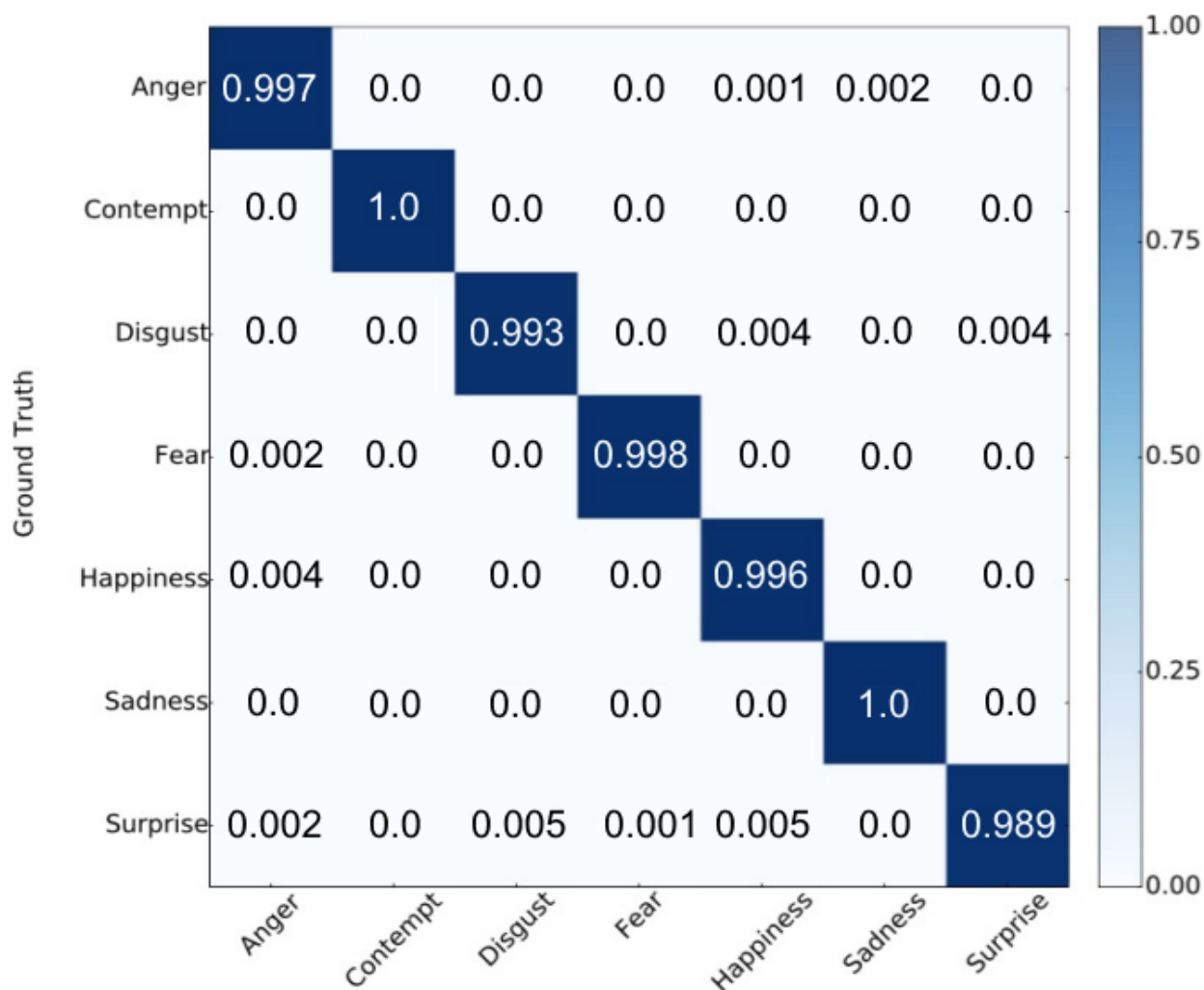


Рисунок 3.2 – Матрица усредненного 10-кратного перекрестного оценивания на наборе данных СКР. Самая низкая точность достигнута на эмоции удивление с 98,79% в то время, как презрение, печаль оба распознаны со 100% точностью

Набор данных ММІ

На тестовой выборке из набора было произведено тестирование. Была применена 10-кратная перекрестная проверка. Её результаты представлены на рисунке 3.3.

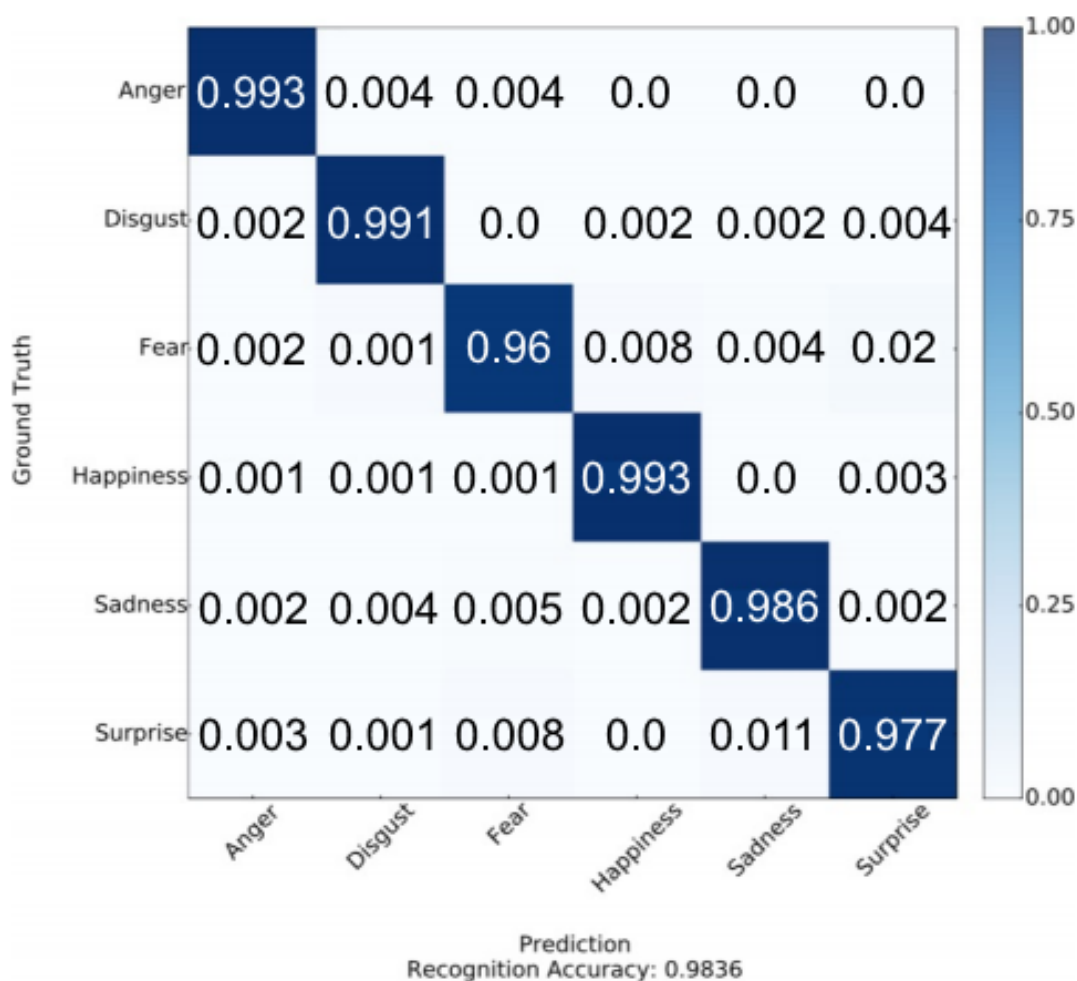


Рисунок 3.3 – Матрица усредненной 10-кратной перекрестной проверки на наборе данных MMI. Самая низкая точность получена на эмоции страха с 93,75%. Счастье распознается с точностью 98,21%.

На приведенном рисунке показано, что точность для страха самая низкая с 93,75%, в то время как счастье почти полностью распознано верно – 98,21%. Страх и удивление – эмоции, которые чаще всех бывают спутаны.

Средняя точность составляет 98,36%. Это лучше чем точность, достигнутая авторами Wang and Yin на данном наборе (таблица 3.2).

Таблица 3.2 – Сравнение точности методов на наборе MMI

Автор	Метод	Точность
Wang and Yin [20]	LDA	93.33%
Wang and Yin	QDC	92.78%
Wang and Yin	NBC	85.56%
	Разработанный метод	98.36%

Таким образом, уменьшив число параметров сети и время обучения модели возможно получить хороший результат, сопоставимый с GoogleNet. Предлагаемая архитектура не зависит от извлечения признаков вручную и работает лучше, чем ранее предложенные сверточные нейронные сетевые подходы.

4 РЕЗУЛЬТАТЫ РАБОТЫ

В данной работе представлен алгоритм, который работает полностью автоматически. Это нейронная сеть, которая требует меньшее количество вычислительных затрат по сравнению со многими существующими архитектурами CNN. Для создания данной архитектуры была введена новая составная структура FeatEx. FeatEx создает широкое представление признаков исходных данных.

По результатам 10-кратной перекрестной валидации результат следующий: в среднем, с точностью распознавания 97,6% на наборе данных СКР и 98,36% в наборе данных MMI. Это показывает, что предлагаемая архитектура способна конкурировать с современными подходами в области эмоций.

Для оценки точности и производительности были использованы следующие метрики:

1. Кросс-энтропийная функция ошибки. Была использована модель глубокого обучения с кросс-энтропийным слоем в конце с семью единицами softmax, поэтому цель состоит в том, чтобы уменьшить потери в мультиклассе / кросс-энтропии.

2. Точность. Это говорит нам о том, насколько точно разработанная модель работает при распознавании эмоций. Определяется как процент верно распознанных эмоций.

3. Метрика путаницы: поскольку проблема разработки заключается в классификации нескольких классов, метрика путаницы поможет нам узнать, какие классы являются более доминирующими по сравнению с другими или к какому классу модель более смещена. Это показало четкую картину результата классификации модели.

Примеры работы алгоритма представлены на рисунках 4.1-4.3.

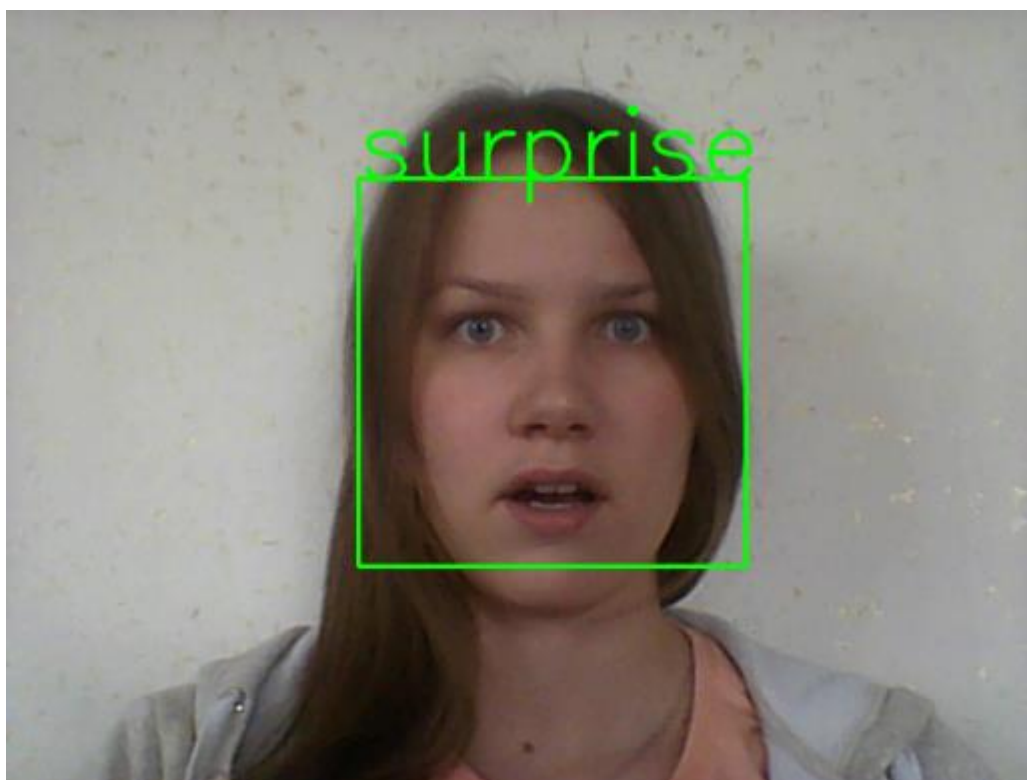


Рисунок 4.1 – Результат распознавания эмоции «удивление»

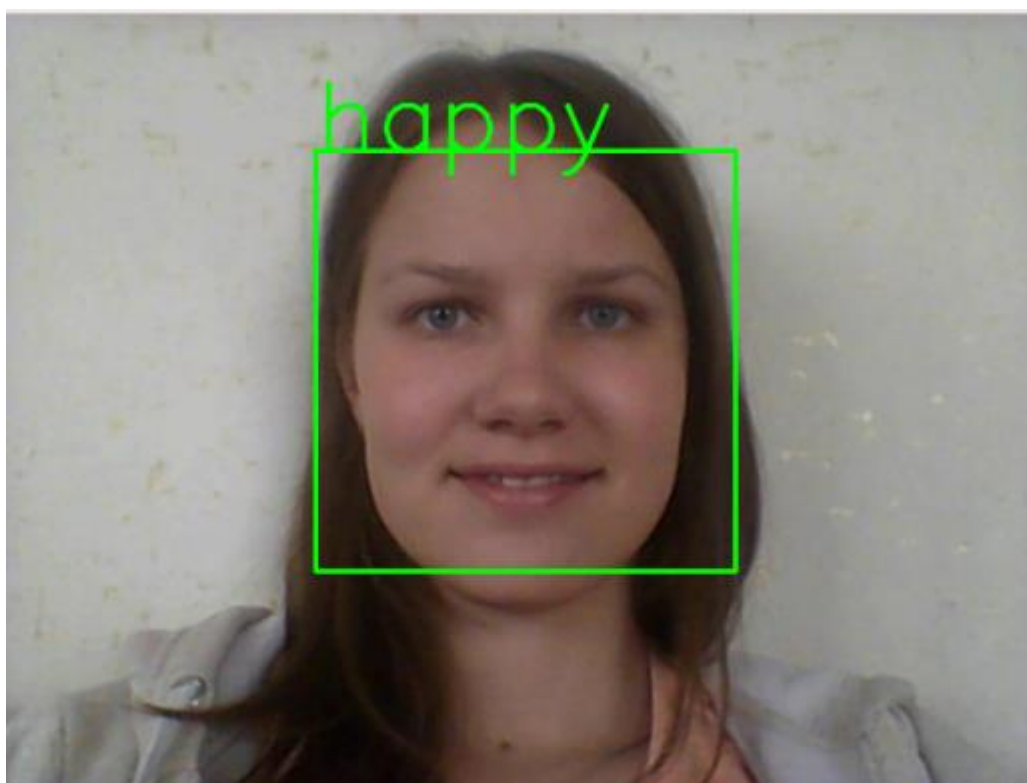


Рисунок 4.2 – Результат распознавания эмоции «счастье»



Рисунок 4.3 – Результат распознавания эмоции «грусть»

Среднее время распознавания при работе в режиме реального времени:

лица – 0.07 сек.

эмоции – 0.02 сек.

РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ

Студенту:

Группа	ФИО
8BM73	Кузнецова А.С.

Инженерная школа	информационных технологий и робототехники	Подразделение	отделение информационных технологий
Уровень образования	магистратура	Направление/специальность	09.04.01 Информатика и вычислительная техника

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

1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	Оклад руководителя – 47104 руб.; Оклад инженера – 22500 руб.; Персональный компьютер: 58000 руб.
2. Нормы и нормативы расходования ресурсов	Норма амортизации – 20% Тариф на электроэнергию – 5,8 руб за кВт.ч
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	Ставка для расчета отчислений во внебюджетные фонды – 30% Ставка для расчета накладных расходов – 30%

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого потенциала, перспективности и альтернатив проведения НИ с позиции ресурсоэффективности и ресурсосбережения	Проведена оценка коммерческого потенциала: 1. Потенциальные потребители результатов исследования. 2. Анализ конкурентных технических решений 3. SWOT-анализ
2. Планирование и формирование бюджета научных исследований	Произведен расчет бюджета научных исследований: 356896 руб
3. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования	Определена ресурсная, финансовая, бюджетная эффективность исследования посредством расчета интегрального финансового показателя, интегрального показателя ресурсоэффективности и эффективности.

Перечень графического материала (с точным указанием обязательных чертежей):

1. Оценка конкурентоспособности технических решений
2. Матрица SWOT

3. <i>Альтернативы проведения НИ</i> 4. <i>График проведения и бюджет НИ</i> 5. <i>Оценка ресурсной, финансовой и экономической эффективности НИ</i>	
Дата выдачи задания для раздела по линейному графику	

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН	Меньшикова Е.В.	канд. филос. наук, доцент		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8BM73	Кузнецова А.С.		

Введение. В работе рассматривается задача разработки нейросетевого алгоритма для определения эмоционального состояния человека по изображению его лица.

Актуальность разработки обусловлена необходимостью получения информации об эмоциональном состоянии человека для обеспечения его безопасности, выявления опасных групп населения по камерам видеонаблюдения, повышения степени интеллектуализации робото-человеческого взаимодействия.

Целью раздела «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение» является проектирование и создание конкурентоспособных разработок, технологий, отвечающих современным требованиям в области ресурсоэффективности и ресурсосбережения.

5.1 Потенциальные потребители результатов исследования

Для анализа потребителей результатов исследования необходимо рассмотреть целевой рынок и провести его сегментирование.

Целевой рынок представляет собой сегменты рынка, на котором в будущем возможна продажа разработки.

Сегментирование – это разделение покупателей на однородные группы, для каждой из которых может потребоваться определенный товар.

Произведем анализ рынка потенциальных потребителей.

Ввиду того, что разработка, которая является конечной целью научного исследования, обладает как научной, так практической ценностью, потенциальными потребителя являются физические и/или юридические лица, занимающиеся научно-исследовательской или практической деятельностью в области определения эмоций человека.

Главными потребителями разработки являются компании, которым необходимо получать данные об эмоциональном состоянии человека. Сферы же их деятельности могут быть разнообразны: робототехника, системы безопасности (камеры видеонаблюдения, определение состояния водителей, впечатления от рекламы и пр.)

5.2 Сегментирование рынка

Так как главными потребителями разработки являются компании, в качестве основных критериев сегментации были выделены отрасль деятельности компании и её размер.

		Реклама	Системы безопасности	Робототехника
Размер компании	Крупные			
	Средние			
	Мелкие			

Как можно заметить из приведенной карты сегментирования рынка услуг по использованию алгоритмов распознавания эмоционального состояния человека, главными потребителями являются крупные компании в сфере рекламы и робототехники и средние в сфере систем безопасности и робототехнике.

5.3 Оценка конкурентоспособности технических решений

В качестве основных конкурентных технических решений были выбраны следующие программные продукты:

- EmoDetect (в таблице обозначен как «1»);
- FaceReader (в таблице обозначен как «2»);
- проект, предлагаемый в текущем исследовании (в таблице обозначен как «3»).

Результаты конкурентного анализа приведены в таблице 5.1.

Таблица 5.1 – Результаты конкурентного анализа

Критерии оценки	Вес	Баллы			Конкурентоспособность		
		Б ₁	Б ₂	Б ₃	К ₁	К ₂	К ₃
Технические критерии оценки ресурсоэффективности							
1. Быстрота распознавания	0,2	5	4	3	1,0	0,8	0,6
2. Универсальность	0,2	4	3	5	0,8	0,6	1,0
3. Простота использования	0,05	5	4	4	0,25	0,2	0,2
4. Качество распознавания	0,25	2	4	4	0,5	1,0	1,0
5. Потенциал разработки	0,1	2	4	4	0,2	0,4	0,4
Экономические критерии оценки эффективности							
6. Доступность	0,1	4	3	4	0,4	0,3	0,4
7. Цена	0,1	4	3	3	0,4	0,3	0,3
Итого:		26	25	27	3,55	3,6	3,9

Пояснение некоторых пунктов:

- под «Универсальностью» подразумевается возможность использования разработки совместно с другими методами;
- под «Доступностью» понимается то, насколько открытой является разработка (находится ли исходный код в открытом доступе);
- критерий «Потенциал разработки» отражает возможности по улучшению алгоритма без внесения значительных изменений в его архитектуру.

По результатам проведённого конкурентного анализа можно сделать вывод, что предложенный метод наиболее конкурентоспособен в сравнении с другими рассмотренными методами. В первую очередь, такая конкурентоспособность обусловлена за счёт универсальности метода, а так же за счёт обеспечения высокого качества распознавания эмоций.

5.4 FAST-анализ

FAST-анализ представляет собой функционально-стоимостной анализ. Далее приведены пять стадий, предполагаемых FAST-анализом.

5.4.1 Выбор объекта FAST-анализа

В рамках магистерской диссертации в качестве объекта FAST-анализа выступает объект исследования – алгоритм распознавания эмоционального состояния человека.

5.4.2 Описание главной, основных и вспомогательных функций, выполняемых объектом

Для описания главной, основных и вспомогательных функций, выполняемых объектом алгоритм распознавания эмоций человека был разделён на пять составных частей: чтение изображения, предобработка изображения, распознавание лица, распознавание эмоции, визуализация результатов.

Таблица 5.2 – Классификация функций, выполняемых объектом исследования

Наименование процесса	Выполняемая функция	Ранг функции		
		Главная	Основная	Вспомогательная
Чтение изображения	Чтение данных в оперативную память			X
Предобработка изображения	Приведение изображения к нужному формату		X	
Распознавание лица	Распознавание лица человека на изображении		X	
Распознавание эмоции	Распознавание эмоции	X		
Визуализация результатов	Отображение результатов распознавания			X

5.4.3 Определение значимости выполняемых функций объектом

Определим значимость функций, выполняемых объектом по методу расставления приоритетов.

Таблица 5.3 – Матрица смежности функций

	Чтение изображе ния	Преобраб отка изображен ия	Распознав ание лица	Распознав ание эмоции	Визуализа ция результат ов
Чтение изображен ия	=	>	>	>	=
Преобраб отка изображен ия	<	=	>	>	<
Распознава ние лица	<	<	=	>	<
Распознава ние эмоции	<	<	<	=	<
Визуализац ия результато в	=	>	>	>	=

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

Преобразуем матрицу смежности в матрицу количественных соотношений функций.

Таблица 5.4 – Матрица количественных соотношений функций

	Чтение изображе ния	Преобраб отка изображен ия	Распознав ание лица	Распознав ание эмоции	Визуализа ция результат ов
Чтение изображен ия	1	1,5	1,5	1,5	1
Преобраб отка изображен ия	0,5	1	1,5	1,5	0,5
Распознава ние лица	0,5	0,5	1	1,5	0,5
Распознава ние эмоции	0,5	0,5	0,5	1	0,5
Визуализац	1	1,5	1,5	1,5	1

ия результато в					
$\sum = 25$	0,14	0,2	0,24	0,28	0,14

Примечание: 0,5 при «<»; 1,5 при «>»; 1 при «=».

5.4.4 Анализ стоимости функций выполняемых объектом исследования

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

Таблица 5.5 – Определение стоимости функций, выполняемых объектом исследования

Наименовани е процесса	Выполняемая функция	Трудоёмкос ть в (ч)	Стоимость часа работы (руб/час)	Себестоимос ть (руб)
Чтение изображения	Чтение данных в оперативную память	10	222	2220
Предобработка изображения	Приведение изображения к нужному формату	40	222	8880
Распознавание лица	Распознавание лица человека на изображении	50	222	11100
Распознавание эмоции	Распознавание эмоции	210	222	46620
Визуализация результатов	Отображение результатов распознавания	30	222	6660

Общая стоимость исследования: 75480 руб.

5.4.5 Построение функционально-стоимостной диаграммы объекта и её анализ

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

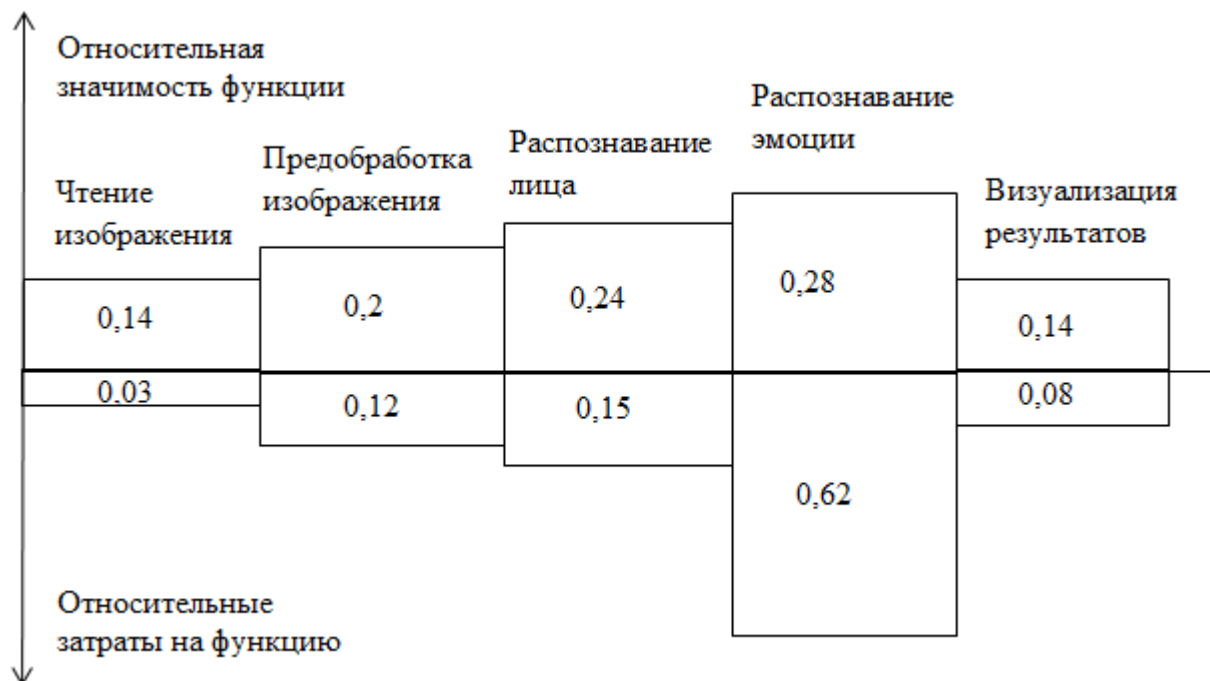


Рисунок 5.1 – Функционально-стоимостная диаграмма

5.4.6 Оптимизация функций выполняемых объектом

Как можно заметить из полученной функционально-стоимостной диаграммы, по функции визуализация результатов наблюдается рассогласование.

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

5.5 SWOT-анализ

В ходе проведения SWOT-анализа была составлена итоговая матрица (таблица 5.6), содержащая описание сильных и слабых сторон проекта, выявленные возможности и угрозы, а также их корреляцию.

Таблица 5.6 – Результаты SWOT-анализа

	Сильные стороны С1. Использование инновационного подхода С2. Большая точность распознавания и доступность, чем у большинства конкурентных разработок С3.Открытость разработки	Слабые стороны Сл1. Проигрыш в точности относительно некоторых конкурентов Сл2. Необходимость наличия большой базы данных для обучения модели
Возможности В1. В виду отсутствия точных методов решения задачи распознавания эмоций человека, отсутствие конкурентных решений В2. Тенденция к увеличению популярности использования алгоритмов нейронных сетей для решения задач по распознаванию В3. Возможность масштабирования и улучшения проекта	В2С2: В связи с пристальным вниманием к алгоритмам на основе нейронных сетей в последнее время, возможно появление новых оптимизаций и эвристик, которые в будущем можно будет легко интегрировать в имеющуюся разработку, тем самым значительно улучшив её конкурентоспособность В3С3: Предложенный метод является менее дорогим, обеспечивая неплохую точность, тем самым обеспечивая значительную универсальность	В1Сл2: Для увеличения точности распознавания требуется большая база данных изображений, значительные временные затраты для обеспечения качественного обучения модели

Угрозы У1. В дальнейшем, используемых ресурсов для обучения может быть недостаточно У2. Недостаток человеческих ресурсов по сравнению с конкурентными разработками.	У1У2С2: Недостаток ресурсов для обеспечения масштабирования проекта, дообучения модели	У1У2Сл1: В случае прекращения исследований по данной теме, разработка не сможет быть реализована в полной мере, не обеспечит улучшения и продвижения в данном направлении среди конкурентов
--	---	--

Интерактивная матрица проекта представлена в таблицах 5.7 – 5.10:

Таблица 5.7 – Сильные стороны и возможности

	Сил1	Сил2	Сил3
В1	-	-	-
В2	-	+	-
В3	+	-	-

Таблица 5.8 – Слабые стороны и возможности

	Слаб1	Слаб2
В1	-	-
В2	-	-
В3	-	+

Таблица 5.9 – Сильные стороны и угрозы

	Сил1	Сил2	Сил3
У1	-	+	-
У2	-	+	-

Таблица 5.10 – Слабые стороны и угрозы

	Слаб1	Слаб2
У1	+	+

У2	+	-
----	---	---

Таким образом, можно сделать вывод, что проект необходимо развивать в направлении наибольшей универсальности и гибкости, с возможностью расширения с помощью дообучения модели на новых данных обучения. При этом необходимо отслеживать и, по возможности, применять новейшие технологии и улучшения в области нейросетевых алгоритмов.

5.6 Оценка готовности проекта к коммерциализации

Для оценки готовности научной разработки к коммерциализации была составлена таблица 5.11, содержащая показатели о степени проработанности проекта с позиции коммерциализации и компетенциям разработчика научного проекта.

Таблица 5.11 – Бланк оценки степени готовности научного проекта к коммерциализации

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

	разработки на рынок		
10.	Разработана стратегия (форма) реализации научной разработки	1	1
11.	Проработаны вопросы международного сотрудничества и выхода на зарубежный рынок	1	1
12.	Проработаны вопросы использования услуг инфраструктуры поддержки, получения льгот	1	1
13.	Проработаны вопросы финансирования коммерциализации научной разработки	1	1
14.	Имеется команда для коммерциализации научной разработки	1	1
15.	Проработан механизм реализации научного проекта	2	3
	ИТОГО БАЛЛОВ	29	32

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

$$B_{\text{сум}} = \sum B_i ,$$

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

B_i – балл по i -му показателю.

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

5.7 Методы коммерциализации результатов научно-технического исследования

Для успешной коммерциализации научно-технических разработок необходимо правильно определить метод коммерциализации. Данный метод определяет успешность продвижения продукта на рынке, коммерческий эффект которого может быть использован для дальнейшего продвижения и развития проекта.

Организация совместного предприятия – является наиболее подходящим методом для разрабатываемого проекта. Один человек не может справиться с

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

5.8 Инициация проекта

Инициация проекта состоит из процессов, которые выполняются для нового проекта или новой стадии проекта. Для этого определяются начальные цели, содержание, фиксируются ресурсы. Также определяются внутренние и внешние заинтересованные стороны проекта.

5.8.1 Цели и результат проекта

В данном разделе необходимо привести информацию о заинтересованных сторонах проекта, иерархии целей проекта и критериях достижения целей.

Под заинтересованными сторонами проекта понимаются лица или организации, которые активно участвуют в проекте или интересы которых могут быть затронуты как положительно, так и отрицательно в ходе исполнения или в результате завершения проекта. Это могут быть заказчики, спонсоры, общественность и т.п. Информация по заинтересованным сторонам проекта представлена в таблице 5.12.

Таблица 5.12 – Заинтересованные стороны проекта

Заинтересованные стороны проекта	Ожидания заинтересованных сторон
НИ ТПУ, школа ИШИТР, направление КАИД	Получение нейросетевого алгоритма для определения эмоционального состояния человека по изображению его лица с целью его дальнейшего

	внедрения и представления на конференциях
Рекламные агентства	Распознавание эмоций позволит проводить анализ покупателя и на основе его эмоционального состояния предлагать соответствующие товары
Службы видеонаблюдения и безопасности	Разработка позволит обнаруживать подозрительные лица, преступников в магазинах и общественных местах
Компании, работающие с искусственным интеллектом	Распознавание эмоций позволит улучшить человеко-машинное взаимодействие, сделать системы, использующие искусственный интеллект более естественными и интеллектуальными

Исходя из ожиданий заинтересованных сторон, была сформирована цель и ожидаемые результаты проведения исследования (таблица 5.13).

Таблица 5.13 – Цели и результат проекта

Цели проекта:	Разработка нейросетевого алгоритма распознавания эмоционального состояния человека
Ожидаемые результаты проекта:	Нейросетевой алгоритм распознавания эмоционального состояния человека с точностью не ниже 96%
Критерии приемки результата проекта:	Точность, выраженная в процентах, и определяемая числом правильно распознанных эмоций
Требования к результату проекта:	Требование:
	Оформление результатов исследования в соответствии с требованиями нормативно-технической документации
	Точность распознавания эмоции у разработанного алгоритма должна быть выше (96%), чем у большинства конкурентных решений

	Стоимость проекта должна быть сопоставима по сравнению с аналогами, а в лучшем случае быть меньшей
--	--

На данном этапе были определены заинтересованные стороны проекта, были определены цели и ожидаемые результаты проекта, а также критерии приемки результатов проекта заинтересованной стороной. В дальнейшем это позволит не допустить перерасход средств и избежать конфликтов между участниками проекта.

5.8.2 Организационная структура проекта

При проведении сложных научно-исследовательских работ необходимо организовать группу способную разработать сам проект, оценить возможные экономические и социальные риски проекта и выполнить требования потенциальных потребителей.

Для выполнения вышенаписанных требований сформирована следующая группа (таблица 5.14).

Таблица 5.14 – Рабочая группа проекта

№ п/п	ФИО, основное место работы, должность	Роль в проекте	Функции	Трудовые затраты, час.
1.	Спицын В.Г., профессор отделения ИТ ТПУ	Руководитель проекта	Составление ТЗ, консультации, проверка полученных результатов	134
2.	Кузнецова А.С., магистрант ТПУ	Исполнитель по проекту	Анализ области разработки, разработка проекта, тестирование, подготовка документации	528
ИТОГО:				662

5.8.3 Ограничения и допущения проекта

Ограничения и допущения сведены в таблице 5.15.

Таблица 5.15 – Ограничения проекта

Фактор	Ограничения/ допущения
Бюджет проекта	356896
Источник финансирования	Частное финансирование
Сроки проекта:	02.02.2019 – 02.06.2019
Дата утверждения плана управления проектом	02.02.2019
Дата завершения проекта	02.06.2019
Прочие ограничения и допущения*	-

5.9 Планирование управления научно-техническим проектом

5.9.1 График проведения и бюджет НТИ

Перечень этапов и работ в рамках проведения научного исследования представлен в таблице 5.16 в виде календарного плана проекта:

Таблица 5.16 – Календарный план

№	Название	Продолжительность, дни	Дата начала работ	Дата окончания работ	Состав участников
1	Составление и утверждение технического задания	8	02.02.19	11.02.19	Руководитель
2	Подбор и изучение материалов по теме	8	12.02.19	20.02.19	Руководитель, инженер
3	Изучение уже существующих решений в данной области	4	21.02.19	26.02.19	Инженер
4	Выбор направления исследований	3	27.02.19	01.03.19	Руководитель, инженер
5	Календарное планирование работ по теме	2	02.03.19	04.03.19	Руководитель, инженер
6	Теоретическая разработка структуры алгоритма, решающего поставленную задачу	6	05.03.19	12.03.19	Инженер

7	Построение пробных вариантов алгоритма	11	13.03.19	25.03.19	Инженер
8	Оценка результатов и итеративное внесение изменений в структуру алгоритма	35	26.03.19	08.05.19	Инженер, руководитель
9	Сравнение результатов работы конечного варианта алгоритма с другими решениями в данной области	8	09.05.19	18.05.19	Инженер
10	Оценка эффективности полученных результатов	4	19.05.19	23.05.19	Инженер, руководитель
11	Оценка целесообразности проведения дальнейших исследований по данной теме	7	24.05.19	31.05.19	Инженер, руководитель

Таблица 5.17 – Календарный план-график работ

№	Вид работ	$T_{k,i}$	Продолжительность выполнения работ											
			Февр			Март			Апр			Май		
			1	2	3	1	2	3	1	2	3	1	2	3
1	Составление и утверждение технического задания	8												
2	Подбор и изучение материалов по теме	8												
3	Изучение уже существующих решений в данной области	5												
4	Выбор направления исследований	3												
5	Календарное планирование работ по теме	2												
6	Теоретическая разработка структуры алгоритма, решающего поставленную задачу	7												
7	Построение пробного варианта алгоритма	12												
8	Оценка результатов и, при необходимости, итеративное внесение изменений в структуру алгоритма	36												
9	Сравнение результатов работы конечного варианта алгоритма с другими решениями в данной области	9												
10	Оценка эффективности полученных результатов	4												
11	Оценка целесообразности проведения дальнейших исследований по данной теме	7												

Примечание к таблице 5.17:

 – Руководитель

 – Инженер

5.10 Бюджет НТИ

В состав затрат на создание проекта включается величина всех расходов, необходимых для реализации комплекса мероприятий, составляющих содержание разработки. Расчёт стоимости выполнения работы производится по следующим статьям затрат:

- заработная плата;
- накладные расходы;
- прочие прямые расходы;
- специальное оборудование для научных работ (амортизация);
- отчисления на социальные нужды.

Амортизация оборудования

В данной статье рассчитывается амортизация используемого оборудования за время выполнения научно-исследовательской работы.

Таблица 5.18 – Затраты на амортизацию оборудования

Наименование оборудования	Балансовая стоимость единицы оборудования, руб.	Срок эксплуатации, лет	Фактическое время работы, дни	Затраты, руб.
Персональный компьютер	58000	5	88	2797

Затраты на амортизационные отчисления составляют:

$$САМ = 58000/5/365 * 88 = 2797 \text{ руб.}$$

Заработная плата

Статья расходов включает заработную плату научного руководителя и студента, а также премии, входящие в фонд заработной платы. Расчет основной заработной платы выполняется на основе трудоёмкости выполнения каждого этапа и величины месячного оклада исполнителя.

Заработная плата рассчитывается по формуле:

$$C_{зп} = Z_{осн} + Z_{доп},$$

где $Z_{осн}$ – основная заработная плата; $Z_{доп}$ – дополнительная заработная плата.

Основная заработная плата рассчитывается по формуле:

$$Z_{осн} = Z_{дн} * T_{раб},$$

где $Z_{дн}$ – среднедневная заработная плата работника, руб;

$T_{раб}$ – продолжительность работ, выполняемых работником, раб.дн.

Среднедневная заработная плата рассчитывается по формуле:

$$Z_{дн} = \frac{Z_m * M}{F_d},$$

где Z_m – месячный должностной оклад работника, руб;

F_d – действительный годовой фонд рабочего времени, раб. дн;

M – количество месяцев работы без отпуска в течение года: при отпуске в 28 раб. дня $M = 11,2$ месяца, 5-дневная неделя; при отпуске в 56 раб. дней $M = 10,4$ месяца, 6-дневная неделя.

Среднедневная заработная плата руководителя:

$$Z_{дн} = \frac{91853 * 10,4}{251} = 3806 \text{ руб}$$

Среднедневная заработная плата инженера (на предприятии):

$$Z_{дн} = \frac{32175 * 10,4}{251} = 1333 \text{ руб}$$

Таблица 5.19 – Баланс рабочего времени

Показатели рабочего времени	Руководитель	Инженер
Календарное число дней	365	365
Количество нерабочих дней:		
- выходные дни;	52	52
- праздничные дни.	14	14
Потери рабочего времени:		
- отпуск;	48	30
- невыходы по болезни.		18
Действительный годовой фонд рабочего времени	251	251

Расчёт месячного должностного оклада:

$$З_{\text{м}} = З_{\text{б}} * (1 + K_{\text{пр}} + K_{\text{д}}) * K_{\text{р}},$$

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

$K_{\text{пр}}$ – премиальный коэффициент, принят за 30% от $З_{\text{б}}$ для преподавателя и 10% от $З_{\text{б}}$ инженера на предприятии;

$K_{\text{д}}$ – коэффициент доплат и надбавок, принят за 20% от $З_{\text{б}}$ для преподавателя;

$K_{\text{р}}$ – районный коэффициент, 1.3 для Томска.

Расчёт месячного должностного оклада руководителя:

$$З_{\text{м}} = 47104 * (1 + 0,3 + 0,2) * 1,3 = 91853 \text{ руб}$$

Расчёт месячного должностного оклада студента на предприятии:

$$З_{\text{м}} = 22500 * (1 + 0,1 + 0) * 1,3 = 32175 \text{ руб}$$

Таблица 5.20 – Расчёт основной заработной платы

Исполнители	З _б	К _{пр}	К _д	К _р	З _м , руб	З _{дн} , руб	Т _р , раб.дн.	З _{осн} , руб
Руководитель	47104	0,3	0,2	1,3	91853	3806	22	83732
Инженер	22500	0,1	0	1,3	32175	1333	88	117304
Итого З_{осн} = 201036 руб								

Дополнительная заработная плата рассчитывается исходя из 10-15% от основной заработной платы, работников.

$$З_{\text{доп}} = З_{\text{осн}} * К_{\text{доп}},$$

где З_{осн} – основная заработная плата;

К_{доп} – коэффициент дополнительной заработной платы принимается равным 0.1.

Дополнительная заработная плата руководителя:

$$З_{\text{доп}} = 83732 * 0.1 = 8373 \text{ руб}$$

Дополнительная заработная плата инженера на предприятии:

$$З_{\text{доп}} = 117304 * 0.1 = 11730 \text{ руб}$$

Таблица 5.21 – Заработная плата исполнителей

Заработная плата, руб	Руководитель	Инженер
Основная	83732	117304
Дополнительная	8373	11730

Итого	92105	129034
--------------	--------------	---------------

Отчисления на социальные нужды

В данной статье расходов отражаются обязательные отчисления по установленным законодательством Российской Федерации нормам:

- государственное социальное страхование (ФСС);
- пенсионный фонд (ПФ);
- медицинское страхование (ФФОМС).

$$C_{\text{внеб}} = K_{\text{внеб}} * (З_{\text{осн}} + З_{\text{доп}}),$$

где $K_{\text{внеб}}$ – коэффициент отчислений на уплату во внебюджетные фонды, принимается равным 30% (в 2019 г.).

Отчисления для руководителя: 27632 руб.

Отчисления для инженера: 38710 руб.

Накладные расходы

Затраты, которые не попали в предыдущие статьи расходов, учитываются в качестве накладных расходов.

$$C_{\text{накл}} = K_{\text{накл}} * (З_{\text{осн}} + З_{\text{доп}}),$$

где $K_{\text{накл}}$ – коэффициент накладных расходов, принимается равным 0,3.

Общие накладные расходы:

$$C_{\text{накл}} = 0,3 * (92105 + 129034) = 66342 \text{ руб}$$

Прочие прямые затраты

В ходе выполнения работы были затрачены энергетические ресурсы. Затраты на электроэнергию можно рассчитать следующим образом:

$$\mathcal{E}_{\text{об}} = P_{\text{об}} \cdot \mathcal{C}_{\text{э}} \cdot t_{\text{об}} = 0.09 * 528 * 5.8 = 276 \text{ руб.}$$

Бюджет разработки

Определим итоговые затраты на научно-технический проект по разработке нейросетевых алгоритмов для определения эмоционального состояния человека по изображениям его лица.

Таблица 5.22 – Бюджет затрат НТИ

№	Статья расходов	Сумма, руб	% от общей суммы
1	Амортизация оборудования	2797	0,78
2	Заработная плата	221139	62
3	Отчисления на социальные нужды	66342	18,6
4	Накладный расходы	66342	18,6
5	Прочие прямые затраты	276	0,02
	Итого	356896	100

5.11 Реестр рисков проекта

Идентифицированные риски проекта включают в себя возможные неопределенные события, которые могут возникнуть в проекте и вызвать последствия, которые повлекут за собой нежелательные эффекты. Собранная информация отражена в таблице 5.23.

Таблица 5.23 – Реестр рисков

№	Риск	Потенциальное воздействие	Вероятность наступления (1-5)	Влияние риска (1-5)	Уровень риска*	Способы смягчения риска	Условия наступления
1.	Нехватка ресурсов для усовершенствования алгоритма	Снижение конкурентоспособности	4	4	Высокий	Постепенное улучшение мощностей оборудования и количества трудовых ресурсов	Появление новых разработок в области нейронных сетей, требующих большего количества ресурсов
2.	Отсутствие финансирования	Прекращение разработки проекта	2	5	Низкий	Поиск компаний, готовых инвестировать в разработку	Недостаточная точность распознавания, по сравнению с аналогами
3.	Используемые технологии станут неактуальными и	Повышение трудозатрат на переработку проекта	3	4	Низкий	Отслеживание новых технологий, постоянное улучшение проекта	Высокие темпы развития области нейронных сетей, появление новых инструментов

Можно сделать вывод, что высоким уровнем риска обладает нехватка ресурсов в дальнейшем. Отметим, что при правильном подходе к разработке и поддержке проекта, таком, как своевременная амортизация оборудования, поддержка актуальности технологий, все риски могут быть устранены.

5.12 Оценка ресурсной, финансовой и экономической эффективности НТИ

Определение бюджета затрат на научно-исследовательский проект по каждому варианту исполнения приведен в таблице 5.24.

Таблица 5.24 – Бюджет вариантов исполнения проекта

№	Статья расходов	Сумма, руб			% от общей суммы
		EmoDetect	FaceReader	Аналог3 (текущий проект)	
1	Амортизация оборудования			2797	0,78
2	Заработная плата			221139	62
3	Отчисления на социальные нужды			66342	18,6
4	Накладные расходы			66342	18,6
5	Прочие прямые затраты			276	0,02
	Итого	420000	670000	356896	100

Определение финансовой эффективности происходит на основе расчета интегрального показателя эффективности научного исследования. Интегральный финансовый показатель разработки определяется как:

$$I_{\text{финр}}^{\text{исп}i} = \frac{\Phi_{p^i}}{\Phi_{\text{max}}},$$

где Φ_{p^i} – стоимость i -го варианта исполнения;

Φ_{max} – максимальная стоимость исполнения.

$$I_{\text{финр}}^{\text{исп}i} = \frac{420000}{670000} = 0,63 \text{ – интегральный финансовый показатель аналога 1}$$

$$I_{\text{финр}}^{\text{исп}i} = \frac{670000}{670000} = 1 \text{ – интегральный финансовый показатель аналога 2}$$

$$I_{\text{финр}}^{\text{исп}i} = \frac{356896}{670000} = 0,53 \text{ – интегральный финансовый показатель текущего проекта}$$

Интегральный показатель ресурсной эффективности вариантов исполнения объекта исследования можно определить следующим образом:

$$I_{pi} = \sum a_i \cdot b_i ,$$

где a_i – весовой коэффициент -го варианта исполнения;

b_i – бальная оценка i -го варианта исполнения разработки, устанавливаемого экспертным путем по выбранной шкале оценивания.

Таблица 5.25 – Сравнительная оценка характеристик вариантов исполнения проекта

Критерии оценки	Вес	Баллы		
		EmoDetect	FaceReader	Аналог3 (текущий)
1. Быстрота распознавания	0,2	5	4	3
2. Универсальность	0,2	4	3	5
3. Простота использования	0,05	5	4	4
4. Качество распознавания	0,25	2	4	4
5. Потенциал разработки	0,1	2	4	4
Итого:	1			

$$I_{pi} = 2,75 \text{ – аналог1}$$

$$I_{pi} = 3 \text{ – аналог2}$$

$I_{pi} = 3,2$ – текущая разработка

Интегральный показатель эффективности вариантов исполнения разработок $I_{испi}$ определяется на основании интегрального показателя ресурсоэффективности и интегрального финансового показателя по формуле:

$$I_{исп.1} = \frac{I_{р-исп1}}{I_{фипр}^{исп1}}, I_{исп.2} = \frac{I_{р-исп2}}{I_{фипр}^{исп2}} \text{ и т. д.}$$

$$I_{исп.1} = \frac{2,75}{0,63} = 4,37, \quad I_{исп.2} = \frac{3}{1} = 3, \quad I_{исп.3} = \frac{3,2}{0,53} = 6,04.$$

Сравнение интегрального показателя эффективности, вариантов исполнения разработки позволяет определить сравнительную эффективность проекта (таблица 5.26) и выбрать наиболее целесообразный вариант из предложенных.

Сравнительная эффективность проекта ($\mathcal{E}_{ср}$):

$$\mathcal{E}_{ср} = \frac{I_{исп1}}{I_{исп2}}$$

Таблица 5.26 – Сравнительная эффективность разработки

№ п/п	Показатели	EmoDetect	FaceReader	Аналог3 (текущий проект)
1	Интегральный финансовый показатель разработки	0,63	1	0,53
2	Интегральный показатель ресурсоэффективности разработки	2,75	3	3,2
3	Интегральный показатель	4,37	3	6,04

	эффективности			
4	Сравнительная Эффективность вариантов исполнения	$A_{н1}/A_{н2}=1,46$ $A_{н1}/A_{н3}=0,72$	$A_{н2}/A_{н1}=0,69$ $A_{н2}/A_{н3}=0,5$	$A_{н3}/A_{н1}=1,38$ $A_{н3}/A_{н2}=2,01$

Из полученных и приведённых в таблице 5.26 результатов, можно сделать вывод, что решение, предлагаемое в магистерской диссертации, превосходит рассмотренные аналоги. При этом преимущество над EmoDetect незначительное.

Заключение. В ходе оценки перспективности и альтернатив проведения научного исследования с позиции ресурсоэффективности и ресурсосбережения коммерческого потенциала для работы «Нейросетевые алгоритмы распознавания эмоционального состояния человека по изображениям его лица» были определены потенциальные потребители разработки – средние и крупные компании в сфере робототехнике, крупные в сфере рекламы и средние в сфере систем безопасности.

Также был произведен анализ конкурентных технических решений, результаты которого показали, что предложенный метод имеет высокую конкурентоспособность за счет своей универсальности. Рассчитанный интегральный показатель конкурентоспособности подтвердил перспективность разработки и проведения исследований в данной области.

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

Разграничение этапов работ позволило структурировать план работы над исследованием, определить ответственных за его этапы. На основе созданного перечня этапов и работ был создан календарный план-график, необходимый для планирования времени выполнения определенного этапа научного исследования.

6 СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ

Студенту:

Группа	ФИО
8BM73	Кузнецовой Анастасии Сергеевне

Школа	информационных технологий и робототехники	Отделение (НОЦ)	информационных технологий
Уровень образования	Магистратура	Направление/специальность	09.04.01 «Информатика и вычислительная техника»

Исходные данные к разделу «Социальная ответственность»:	
Характеристика объекта исследования (вещество, материал, рабочая зона) и области его применения	Нейросетевые алгоритмы распознавания эмоционального состояния человека по изображениям его лица
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
1. Правовые и организационные вопросы обеспечения безопасности	1. Специальные правовые нормы трудового законодательства; 2. Организационные мероприятия при компоновке рабочей зоны.
2. Производственная безопасность 2.1. Анализ выявленных вредных факторов при разработке и эксплуатации проектируемого решения 2.2. Анализ выявленных опасных факторов при разработке и эксплуатации проектируемого решения	Вредные факторы: 1. Отклонения показателей микроклимата; 2. Недостаточная освещенность рабочей зоны; 3. Превышение уровня шума; 4. Повышенный уровень электромагнитных излучений; Опасные факторы: 1. Электрический ток 2. Опасность возникновения пожара
3. Экологическая безопасность	Источники выбросов в атмосферу; Образование сточных вод и отходов. Мероприятия по снижению вредного воздействия на ОС
4. Безопасность в чрезвычайных ситуациях	Вероятные ЧС при разработке и

	эксплуатации проектируемого решения и меры по их предупреждению
--	--

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

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент ООД ШБИП	Алексеев Н.А.	-		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8BM73	Кузнецова А.С.		

Введение. Предметом исследования в рамках дипломной работы являются информационные модели сверточных нейронных сетей, а также описание алгоритма распознавания двигательной активности лица для решения практической задачи, а объектом – нейросетевой алгоритм распознавания эмоций человека.

Разработанный алгоритм может быть применен в сфере рекламы (оценка реакции покупателя на товары), робототехники (более естественное робото-человеческое взаимодействие), системах безопасности (обнаружение подозрительных лиц).

Работы проводятся в учебно-научной лаборатории, далее офис, находящейся в аудитории 401 десятого корпуса Томского Политехнического Университета. Предоставленное рабочее место оборудовано столом, креслом, ЖК монитором и системным блоком компьютера.

6.1 Производственная безопасность

Научно-исследовательская работа выполнялась в рабочем кабинете учебно-научной лаборатории, оснащенном персональными электронно-вычислительными машинами (ПЭВМ). В процессе работы на программиста действуют различные опасные и вредные факторы. Задача охраны труда свести действие этих факторов к минимуму и создать оптимальные условия труда. Работа с ЭВМ регламентируется санитарными правилами и нормами СанПиН 2.2.2/2.4.1340-03 «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы».

Рабочее место находится на четвертом этаже здания и представляет собой комнату длиной – 5 м., шириной – 4 м. и высотой – 3 м. Естественное освещение кабинета осуществляется посредством одного окна размерами 2,2 м. х 1,5 м. Дверь – металлическая, одностворчатая, черного цвета. Высота двери – 2 м., ширина – 1 м. Стены комнаты окрашены водоэмульсионной краской бежевого цвета. Потолок подвесной, плиточный. Пол покрыт линолеумом. Площадь кабинета составляет 20 м², объем – 60 м³.

Помещение оборудовано на четыре рабочих места, где установлены 4 персональных компьютера с жидкокристаллическими экранами. Требования, которые определены к минимальной площади и объему на одно рабочее место (4 м² на человека) – выполняются [21].

Выявленные опасные и вредные факторы согласно ГОСТ 12.0.003-74 (с измен. №1, октябрь 1978г., переиздание 1999 г.) «Классификация вредных и опасных производственных факторов» [22] приведены в таблице 6.1.

Таблица 6.1 – Выявленные опасные и вредные производственные факторы согласно ГОСТ 12.0.003-74

Вредные производственные факторы	Опасные производственные факторы
Отклонение показателей микроклимата в помещении	Электрический ток
Повышенный уровень шума и вибрации на рабочем месте	Повышенный уровень статического электричества
Недостаточная освещенность рабочей зоны	Короткие замыкания
Повышенный уровень электромагнитных излучений	
Психофизиологические факторы: физические нагрузки, умственное перенапряжение, перенапряжение глаз, монотонность труда, эмоциональные перегрузки.	

6.1.1 Анализ выявленных вредных факторов при разработке и эксплуатации проектируемого решения

Отклонение показателей микроклимата в помещении

Микроклимат производственных помещений – это климат внутренней среды этих помещений, который определяется действующими на организм человека сочетаниями температуры, влажности и скорости движения воздуха, а также интенсивности теплового излучения от нагретых поверхностей [23].

Оптимальные значения этих характеристик зависят от сезона (холодный, тёплый), а также от категории физической тяжести работы. Для инженера-программиста она является лёгкой (1а), так как работа проводится сидя, без систематических физических нагрузок.

Согласно требованиям СанПиН 2.2.4.548-96. «Гигиенические требования к микроклимату производственных помещений» [24], оптимальные параметры микроклимата в офисах приведены в таблице 6.2.

Таблица 6.2– Оптимальные значения характеристик микроклимата по СанПиН 2.2.4.548 – 96 [24]

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

Таблица 6.3 – Допустимые величины показателей микроклимата (СанПиН 2.2.4.548 – 96) [24]

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

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

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

После фактических измерений были получены следующие параметры:

- температура воздуха – 24 С (измерено термометром);
- относительная влажность – 54% (измерено гигрометром);

– скорость движения воздуха с закрытыми окнами и дверью – 0,1 м/с (измерено анемометром).

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

Недостаточная освещенность рабочей зоны

Освещение – получение, распределение и использование световой энергии для обеспечения благоприятных условий видения предметов и объектов [23].

В рабочем помещении сочетаются естественное освещение (через окна) и искусственное освещение (использование ламп при недостатке естественного освещения).

Светильники в помещении располагаются равномерно по площади потолка, тем самым обеспечивая равномерное освещение рабочих мест.

Разряд зрительных работ программиста относится к разряду III подразряду Г (высокой точности). Параметры искусственного освещения согласно СНиП 23-05-95 «Естественное и искусственное освещение» указаны в таблице 6.4 [25]. Нормы коэффициента пульсации освещенности для III разряда зрительных работ указаны в таблице 6.5.

Таблица 6.4– Нормативные значения освещённости согласно СНиП 23-05-95 [25]

Характеристика зрительной работы	Наименьший или эквивалентный размер объекта различения, мм	Разряд зрительной работы	Подразряд зрительной работы	Контраст объекта с фоном	Характеристика фона	Искусственное освещение		
						Освещённость, Лк		
						При комб. освещении		При общ.освещении
						всего	В том числе от общего	
Высокой	От 0,3 до	III	Г	Средний	Светлый	400	200	200

точности	0,5			<< Большой	<< Средний			
----------	-----	--	--	---------------	---------------	--	--	--

Таблица 6.5– Нормы коэффициента пульсации освещенности для III зрительных работ [26]

Система освещения		Коэффициент пульсации освещенности для III разряда зрительной работы, %
Общее освещение		15
Комбинированное освещение	а) общее	20
	б) местное	15

Источником света для общего освещения выберем люминесцентные лампы белой цветности (ЛБ). Мощность (ЛБ) 40 Вт. Световой поток 2800 лм.

Для помещения выберем тип светильника для люминесцентных ламп ШОД - 2-80. Количество ламп – 2. Мощность лампы 80 Вт. Длина светильника 1530 мм. Ширина светильника 284 мм.

Для создания благоприятных зрительных условий, наименьшая допустимая высота подвеса светильника ШОД над полом 2,5 м. План рабочего места инженера показан на рисунке 6.1

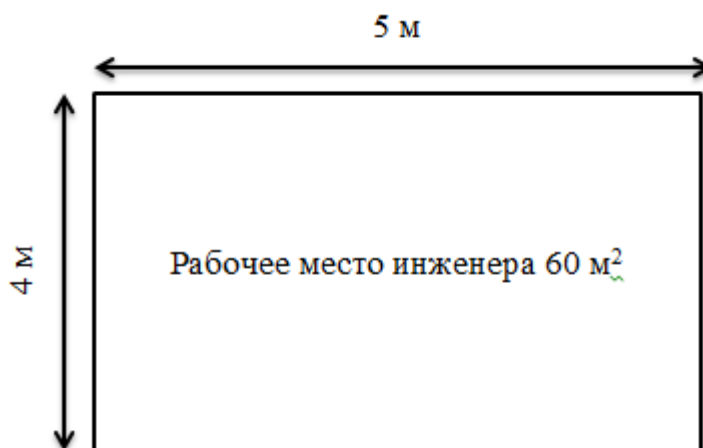


Рисунок 6.1– План рабочего места

Параметры рабочего места инженера: $A=5$ м – длина помещения; $B=4$ м – ширина помещения; $H=3$ м – высота помещения; $S=20$ м² – площадь помещения.

Размещение светильников в помещении определяется следующими параметрами:

$h_c=0,2$ м – расстояние светильников от перекрытия (свес);

$h_n=H-h_c=3-0,2=2,8$ м – высота светильника над полом, высота подвеса;

$h_{pn}=0,7$ м – высота рабочей поверхности над полом;

$h= h_n- h_{pn} = 2,8-0,7=2,1$ м – расчётная высота, высота светильника над рабочей поверхностью.

Уровень необходимого освещения определяется степенью точности зрительных работ.

Индекс помещения находится по формуле:

$$i = \frac{S}{h \cdot (A+B)} = \frac{20}{3 \cdot (4+5)} = 0,74 \quad (6.1)$$

Индекс помещения $i=0,74$

Коэффициент отражения потолка $R_n=70\%$,

Коэффициент отражения стен $R_c=50\%$

По индексу помещения и коэффициентам отражения определяем ближайший коэффициент использования светового потока $\eta = 0,36$.

Размещаем светильники в два ряда. В каждом ряду можно установить 2 светильника ШОД - 2-80. Учитывая, что в каждом светильнике установлено две лампы белой цветности (ЛБ). Мощность (ЛБ) 80 Вт. Световой поток 5200 лм., общее число ламп в помещении $N = 6$.

По формуле 6.2 определяем световой поток лампы:

$$\Phi = \frac{E_n \cdot S \cdot K_z \cdot Z}{N \cdot \eta} \quad (6.2)$$

где Φ – световой поток лампы;

E_n – норма освещенности, 400 Лк;

S – площадь помещения, 20 м²;

K_z – коэффициент запаса, 1,5 – для люминесцентных ламп;

Z – коэффициент неравномерности освещения, 1,1 – для люминесцентных ламп;

N – число ламп в помещении, 6;

η – коэффициент использования светового потока, 0,36.

$$\Phi = \frac{400 * 20 * 1,5 * 1.1}{6 * 0,36} = 6111$$

Делаем проверку выполнения условия:

$$-10\% \leq \frac{\Phi_{\text{расч}} - \Phi_{\text{лампы}}}{\Phi_{\text{расч}}} \cdot 100\% \leq +20\% \quad (6.3)$$

$$\frac{6111 - 5200}{6111} * 100\% = 14,9\%$$

Определяем электрическую мощность осветительной установки:

$$6 * 80 = 480 \text{ Вт} \quad (6.4)$$

Оптимальное расстояние L от крайнего ряда светильников до стены рекомендуется принимать равным L/3.

L – расстояние между соседними светильниками или рядами.

План размещения светильников в рабочем помещении показан на рисунке 6.2.

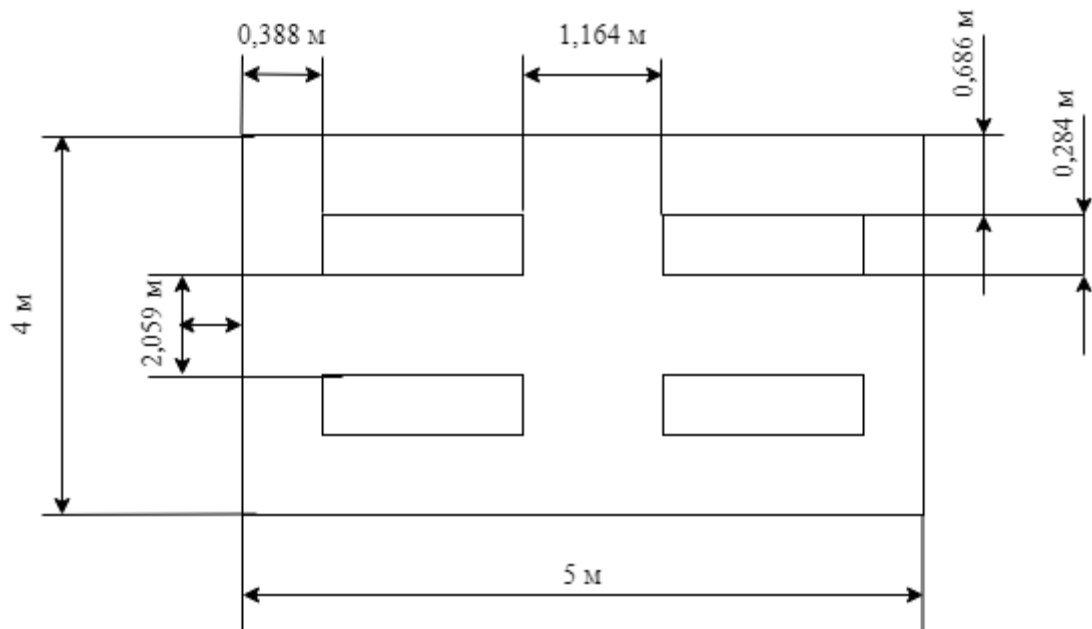


Рисунок 6.2 – План размещения светильников в помещении

Таким образом, при должном обеспечении параметров освещённости, необходимо так же соблюдать правила подбора шрифта и его фона. В ходе

работы, был использован чёрный шрифт и белый фон, размер шрифта – 12-14 пт, что соответствует 0,7 мм для наименьшего объекта различения (точки). Используемые параметры шрифта и фона полностью соответствовали требованиям СНиП 23-05-95 для III разряда зрительных работ.

Повышенный уровень шума и вибрации на рабочем месте

Шум определяется как звук, оцениваемый негативно и наносящий вред здоровью [23]. Основными источниками шума в помещении являются:

- система охлаждения центральных процессоров;
- жесткие диски.

При выполнении основной работы на ПЭВМ уровень шума на рабочем месте не должен превышать 50 дБА. Допустимые уровни звукового давления в помещениях для персонала, осуществляющего эксплуатацию ЭВМ при разных значениях частот согласно СН 2.2.4/2.1.8.562-96 «Шум на рабочих местах, в помещениях жилых, общественных зданий и на территории жилой, застройки», приведены в таблице 6.6 [27].

Таблица 6.6– Допустимые уровни звука на рабочем месте согласно СН 2.2.4/2.1.8.562-96 [27]

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

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

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

Повышенный уровень электромагнитных излучений

Воздействие электромагнитного излучения на человека зависит от напряженностей электрического и магнитного полей, потока энергии, частоты колебаний, размера облучаемого тела [23].

Работа проводилась на современном компьютере, где значения электромагнитного излучения малы и отвечают требованиям СанПиН 2.2.4.1191-03. «Электромагнитные поля в производственных условиях», которые приведены в таблице 6.7 [27].

Таблица 6.7– Допустимые уровни электромагнитных полей согласно СанПиН 2.2.4.1191-03 [27]

Наименование параметров	Допустимые значения
Напряженность электромагнитного поля на расстоянии 50 см. вокруг ВДТ по электрической составляющей должна быть не более: • в диапазоне частот 5 Гц – 2 кГц • в диапазоне частот 2 – 400 кГц	25 В/м 2.5 В/м
Плотность магнитного потока должна быть не более: • в диапазоне частот 5 Гц – 2 кГц • в диапазоне частот 2 – 400 кГц	250 нТл 25 нТл
Напряженность электростатического поля:	20 кВ/м

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

Психофизиологические факторы

Согласно ПОТ РО 29-002-94. «Правила по охране труда для издательств» [28] вредные психофизиологические подразделяются на:

– физические нагрузки (статические и динамические); умственное перенапряжение;

- перенапряжение анализаторов (глаз и др.); монотонность труда;
- эмоциональные перегрузки.

Для снижения влияния вредных и опасных психофизиологических факторов, влияющих на человека, основная деятельность которого предполагает работу за персональным компьютером, было осуществлено [28]:

- размещение персонального компьютера (ПК) на рабочем месте необходимо с учётом границ полей зрения работающего;
- использование рабочего стула с подъемно-поворотным устройством, обеспечивающим регулировку высоты сиденья и спинки, изменение угла наклона спинки;
- использование рабочего стула с подлокотниками. Высота поверхности сиденья регулировалась в пределах 400 – 500 мм. Ширина сиденья составляла – 400 мм, а глубина – 380 мм;
- использование рабочего стула с высотой опорной поверхности спинки – 300 мм, шириной 380 мм, углом наклона спинки в пределах 90 – 100° к плоскости сиденья;
- обеспечение оптимального режима труда и отдыха в течение рабочего дня;
- назначение частых, но коротких перерывов.

6.1.2 Анализ выявленных опасных факторов при разработке и эксплуатации проектируемого решения

Электробезопасность – это система организационных и технических мероприятий и средств, обеспечивающих защиту людей от вредного и опасного для жизни воздействия электрического тока, электрической дуги, электромагнитного поля и статического электричества [29]. Персональный компьютер питается от сети 220В переменного тока с частотой 50Гц. Помещение с ПЭВМ, где проводились описанные выше работы, относится к помещениям без повышенной опасности согласно классификации помещений

по опасности поражения людей электрическим током [30], 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.

Токи статического электричества, наведенные в процессе работы компьютера на корпусах монитора, системного блока и клавиатуры, могут приводить к разрядам при прикосновении к этим элементам. Такие разряды опасности для человека не представляют, но могут привести к выходу из строя компьютера. Для снижения величин токов статического электричества используются нейтрализаторы, местное и общее увлажнение воздуха, использование покрытия полов с антистатической пропиткой.

Короткое замыкание – электрическое соединение двух точек электрической цепи с различными значениями потенциала, не предусмотренное конструкцией устройства и нарушающее его нормальную работу [23]. Короткое замыкание может возникать в результате нарушения изоляции токоведущих элементов или механического соприкосновения неизолированных элементов. Для предотвращения КЗ электрическая цепь содержит предохранитель, который автоматически размыкает цепь, если ток превысил допустимое значение.

В качестве мероприятий по предотвращению возможности поражения электрическим током были соблюдены следующие действия [23]:

- запрет работы на задней панели при включенном сетевом напряжении;
- работы по устранению неисправностей производились квалифицированным персоналом;

- изоляция проводов;
- исключение перегрева приборов;
- слежение за исправностью электропроводки.

6.2 Экологическая безопасность

6.2.1 Анализ влияния процесса исследования на окружающую среду

Помещение с персональным компьютером относится к пятому классу, размер санитарно-защитной зоны которого равен 50 метров, так как работа на персональном компьютере не является экологически опасной.

В данном подразделе рассматривается 3 сферы окружающей среды, на которые оказывается воздействие: загрязнение атмосферного воздуха, образование сточных вод и отходов производства и потребления.

Углекислый газ – один из источников загрязнения, его относят к газам, которые обладают парниковым эффектом. Такие газы, есть ни что другое, как прослойка, которая не дает Земле и космосу осуществлять нормальные теплообменные функции [43].

Человек влияет на водные ресурсы, так как в процессе выполнения он потребляет воду, а продукты жизнедеятельности человека попадают в сточные воды через канализацию [44].

К отходам производства и потребления относятся: утилизация ламп и макулатуры.

6.2.2 Обоснование мероприятий по защите окружающей среды

Отработанная техника утилизируется с уничтожением информации согласно ГОСТ Р 50739-95 «Средства вычислительной техники. Защита от несанкционированного доступа к информации» [33]. Отработанные люминесцентные лампы утилизируются согласно ГОСТ 30772-2001 «Ресурсосбережение. Обращение с отходами. Термины и определения» [34].

В случае выхода из строя они списываются и отправляются на специальный склад, который при необходимости принимает меры по утилизации списанной техники и комплектующих.

В целом, утилизация ламп предполагает передачу использованных ламп предприятиям – переработчикам, которые с помощью специального оборудования перерабатывают вредные лампы в безвредное сырье – сорбент, которое в последующем используют в качестве материала для производства, например тротуарной плитки.

Утилизация макулатуры должна производиться в соответствии с ГОСТ Р 55090-2012 «Ресурсосбережение. Обращение с отходами. Рекомендации по утилизации отходов бумаги» [41] и ГОСТ 10700-97 «Макулатура бумажная и картонная. Технические условия» [42], согласно которым:

- макулатура должна разделяться на три группы: А – высокого качества, Б – среднего качества, В – низкого качества;
- макулатура каждой группы в зависимости от состава, источников поступления, цвета и способности к роспуску должна соответствовать маркам;
- макулатура должна содержать не более определённого количества примесей, определённого для каждой группы.

Все мероприятия по утилизации выполняются посредством сторонних компаний, имеющих соответствующие лицензии на такую деятельность (согласно ст. 17 Федерального закона от 8 августа 2001 г. №128-ФЗ «О лицензировании отдельных видов деятельности» [35]).

6.3 Защита в чрезвычайных ситуациях

6.3.1 Анализ вероятных ЧС, которые могут возникнуть на рабочем месте при проведении исследований

В данном случае на объекте (офис) могут возникать чрезвычайные ситуации (ЧС) следующего характера:

- техногенные;
- экологические;
- природные.

Наиболее типичной ЧС для помещения, котором производится выполнение ВКР, является пожар. Данная ЧС может произойти в случае

замыкания электропроводки оборудования, обрыва проводов, не соблюдения мер пожаробезопасности и т.д.

6.3.2 Обоснование мероприятий по предотвращению ЧС и разработка порядка действия в случае возникновения ЧС

Пожарная безопасность включает в себя комплекс организационных и технических мероприятий, направленных на обеспечение безопасности людей, предотвращения пожара, ограничение его распространения, а также создание условия для успешного тушения пожара.

Пожарная опасность персональных электронно-вычислительных машин, обусловлена наличием в применяемом электрооборудовании горючих изоляционных материалов. Горючими являются изоляция обмоток трансформаторов, различных электромагнитов, проводов и кабелей.

Помещение, где проводились описанные выше работы, по пожарной и взрывной опасности относят к категории Д (пониженная пожароопасность) согласно ППБ-03 [30], характеризующейся отсутствием легковоспламеняющихся веществ и материалов в горячем состоянии. К категории Д относятся помещения, в которых находятся (обращаются) негорючие вещества и материалы в холодном состоянии.

Здание 10-ого корпуса, в котором находится помещение, относится к нестораемым по НПБ 105-95. «Определение категорий помещений и зданий по взрывопожарной и пожарной опасности» [31].

Для того что бы избежать возникновения пожара необходимо проводить следующие профилактические работы, направленные на устранение возможных источников возникновения пожара:

- периодическая проверка проводки;
- отключение оборудования при покидании рабочего места;
- проведение инструктажа работников о пожаробезопасности.

Для предотвращения пожара в аудитории с ПЭВМ имеется:

- углекислотный огнетушитель типа ОУ-2 (данный тип огнетушителя подходит для помещений с электрооборудованием (ГОСТ Р 51057-01 и НПБ 155-02);
- пожарная сигнализация ДИП-3СУ (извещатель пожарный, дымовой оптико-электронный точечный).

В случае возникновения ЧС как пожар, необходимо предпринять меры по эвакуации персонала из офисного помещения в соответствии с планом эвакуации (план эвакуации имеется и размещен в помещении лаборатории). При отсутствии прямых угроз здоровью и жизни произвести попытку тушения возникшего возгорания огнетушителем. В случае потери контроля над пожаром, необходимо эвакуироваться вслед за сотрудниками по плану эвакуации и ждать приезда специалистов, пожарников. При возникновении пожара должна сработать система пожаротушения, издав предупредительные сигналы, и передав на пункт пожарной станции сигнал о ЧС, в случае если система не сработала, по каким либо причинам, необходимо самостоятельно произвести вызов пожарной службы по телефону 112, сообщить место возникновения ЧС и ожидать приезда специалистов.

6.4 Правовые и организационные вопросы обеспечения безопасности

6.4.1 Специальные правовые нормы трудового законодательства

В Российской Федерации вопросы, относящиеся к организации и охране труда при работе за компьютером, регулируются:

- трудовым кодексом;
- «Гигиеническими требованиями к персональным электронно-вычислительным машинам и организации работы» (СанПиН 2.2.2/2.4.1340-03);
- «Типовой инструкцией по охране труда при работе на персональном компьютере» (ТОИ Р-45-084-01).

Ответственность за выполнение данных норм возлагается на юридических лиц, осуществляющих разработку, производство и эксплуатацию ПЭВМ. Например, согласно п. 13.1 Санитарных правил 2.2.2/2.4.1340-03 [21]

лица, работающие с ПЭВМ более 50% рабочего времени (профессионально связанные с эксплуатацией ПЭВМ), должны проходить обязательные предварительные при поступлении на работу и периодические медицинские осмотры в установленном порядке. Такой порядок утвержден приложением 3 к Приказу Минздравсоцразвития РФ от 16.08.2004 N 83 «Об утверждении перечней вредных и (или) опасных производственных факторов и работ, при выполнении которых проводятся предварительные и периодические медицинские осмотры (обследования), и порядка проведения этих осмотров (обследований)» (далее – Приказ N 83) [37].

Кроме вышеупомянутого порядка, Приказом N 83 определены вредные производственные факторы, к которым, в частности, отнесены: – излучения (разд. 3 «Физические факторы» приложения 1); – работы с персональными электронно-вычислительными машинами (ПЭВМ) лиц, профессионально связанных с эксплуатацией ПЭВМ (п. 4.2.3 разд. 4.2 «Зрительно напряженные работы: прецизионные, работы с оптическими приборами и наблюдение за экраном») [37].

Из вышесказанного следует, что работа сотрудников с использованием персональных электронно-вычислительных машин (компьютеров) оказывает негативное воздействие на здоровье, то есть является работой, связанной с вредными и (или) опасными условиями труда.

В соответствии со ст. 147 ТК РФ [38] оплата труда работников, занятых на тяжелых работах, работах с вредными и (или) опасными и иными особыми условиями труда, устанавливается в повышенном размере по сравнению с тарифными ставками, окладами (должностными окладами), установленными для различных видов работ с нормальными условиями труда.

Минимальные размеры таких выплат и условия указанного повышения устанавливаются в порядке, определяемом Правительством РФ, с учетом мнения Российской трехсторонней комиссии по регулированию социально-трудовых отношений. Конкретные размеры повышения оплаты труда определяются работодателем с учетом мнения представительного органа

работников в порядке, установленном ст. 372 ТК РФ [39], либо коллективным договором, трудовым договором.

Организация режимов труда и отдыха при работе с ПЭВМ в соответствии с [21] осуществляется в зависимости от вида и категории трудовой деятельности.

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

При работе с ПЭВМ в ночную смену (с 22 до 6 час), независимо от категории и вида трудовой деятельности, продолжительность регламентированных перерывов следует увеличивать на 30%.

6.4.2 Организационные мероприятия при компоновке рабочей зоны исследователя

Требования к организации рабочих мест пользователей:

- Рабочее место должно быть организовано с учетом эргономических требований согласно ГОСТ 12.2.032-78 «ССБТ. Рабочее место при выполнении работ сидя. Общие эргономические требования» [35] и ГОСТ 12.2.061-81 «ССБТ. Оборудование производственное. Общие требования безопасности к рабочим местам» [36];

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

- На уровне экрана должен быть установлен оригинал-держатель.

На рисунке 6.3. схематично представлены требования к рабочему месту.



Рисунок 6.3 – Организация рабочего места

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

Согласно СанПиН 2.2.2/2.4.1340-03 «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы» [21]:

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

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

ЗАКЛЮЧЕНИЕ

В ходе дипломной работы был выполнен аналитический обзор общедоступных алгоритмов определения эмоционального состояния человека.

На основании проанализированной информации была предложена архитектура модели для определения эмоционального состояния человека по изображениям его лица.

Алгоритм представляет собой нейронную сеть, которая требует меньшее количество вычислительных затрат по сравнению со многими существующими архитектурами CNN. Для создания данной архитектуры была введена новая составная структура FeatEx. FeatEx создает широкое представление признаков исходных данных.

По результатам 10-кратной перекрестной валидации результат следующий: в среднем, с точностью распознавания 97,6% на наборе данных СКР и 98,36% в наборе данных MMI. Это показывает, что предлагаемая архитектура способна конкурировать с современными подходами в области эмоций.

Описанная архитектура компактна, что позволяет не только обучать модель быстро, но и использовать её для приложений, работающих в режиме реального времени.

Производительность предлагаемой архитектуры обеспечивает эффективность и надежность использования предлагаемой работы в реальных приложениях.

В перспективе распознавания эмоциональной составляющей человеческого лица предполагается создание комплекса алгоритмов, которые бы учитывали все тонкости. Голосовое детектирование вместе с трекингом изменений окружающей среды, распознавание основных психологических телодвижений индивидуума и все это в совокупности с обычным эмоциональным детектором значительно увеличили бы точность любых программных средств в этой проблеме машинного обучения.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, "Decaf: A deep convolutional activation feature for generic visual recognition," arXiv preprint arXiv:1310.1531, 2013.
2. GoogLeNet [Электронный ресурс]. – Режим доступа: <https://ai.google/research/pubs/pub43022> (дата обращения: 06.05.2019).
3. S. Happy and A. Routray, "Automatic facial expression recognition using features of salient facial patches," *Affective Computing, IEEE Transactions on*, vol. 6, no. 1, pp. 1–12, Jan 2015.
4. Y.-H. Byeon and K.-C. Kwak, "Facial expression recognition using 3d convolutional neural network."
5. I. Song, H.-J. Kim, and P. B. Jeon, "Deep learning for realtime robust facial expression recognition on a smartphone," in *Consumer Electronics (ICCE), 2014 IEEE International Conference on. IEEE*, 2014, pp. 564–567.
6. P. Lucey, J. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, June 2010, pp. 94–101.
7. K. Anderson and P. W. Mcowan, "A real-time automated system for recognition of human facial expressions," *IEEE Trans. Syst., Man, Cybern. B, Cybern*, pp. 96–105, 2006.
8. I. Kotsia and I. Pitas, "Facial expression recognition in image sequences using geometric deformation features and support vector machines," *Image Processing, IEEE Transactions on*, vol. 16, no. 1, pp. 172–187, Jan 2007.
9. C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image and Vision Computing*, vol. 27, no. 6, pp. 803 – 816, 2009.

10. A. Zafer, R. Nawaz, and J. Iqbal, “Face recognition with expression variation via robust ncc,” in Emerging Technologies (ICET), 2013 IEEE 9th International Conference on, Dec 2013, pp. 1–5.
11. Review: GoogLeNet (Inception v1)— Winner of ILSVRC 2014 (Image Classification), [Электронный ресурс]. – Режим доступа: <https://medium.com/coinmonks/paper-review-of-googlenet-inception-v1-winner-of-ilsvrc-2014-image-classification-c2b3565a64e7> (дата обращения: 06.05.2019).
12. Latent Dirichlet Allocation, [Электронный ресурс]. – Режим доступа: <http://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf> (дата обращения: 08.05.2019).
13. CNN, [Электронный ресурс]. – Режим доступа: <https://www.machinecurve.com/index.php/2018/12/07/convolutional-neural-networks-and-their-components-for-computer-vision/> (дата обращения: 12.05.2019).
14. M. Pantic, M. F. Valstar, R. Rademaker, and L. Maat, “Webbased database for facial expression analysis,” in Proceedings of IEEE Int’l Conf. Multimedia and Expo (ICME’05), Amsterdam, The Netherlands, July 2005, pp. 317–321.
15. Cohn-Kanade AU-Coded Expression Database, [Электронный ресурс]. – Режим доступа: <http://www.pitt.edu/~emotion/ck-spread.htmSdf> (дата обращения: 14.05.2019).
16. Keras: [Электронный ресурс] / Keras documentation. URL: <https://keras.io/> (Дата обращения: 15.05.2019).
17. OpenCV: [Электронный ресурс] – URL: <https://opencv.org/> (Дата обращения: 15.05.2019).
18. Tensorflow: [Электронный ресурс] – URL: <https://www.tensorflow.org/> (Дата обращения: 15.05.2019).
19. Dlib: [Электронный ресурс] – URL: <http://dlib.net/> (Дата обращения: 15.05.2019).

20. J. Wang and L. Yin, "Static topographic modeling for facial expression recognition and analysis," *Comput. Vis. Image Underst.*, vol. 108, no. 1-2, pp. 19–34, Oct. 2007.
21. СанПиН 2.2.2/2.4.1340-03. «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы».
22. ГОСТ 12.0.003-74 (с измен. №1, октябрь 1978г., переиздание 1999 г.) «Классификация вредных и опасных производственных факторов».
23. Безопасность жизнедеятельности: Учебник для вузов /Под ред. К.З. Ушакова. – М.: Изд-во Московского гос. горного университета, 2000.– 430 с.
24. СанПиН 2.2.4.548 – 96. «Гигиенические требования к микроклимату производственных помещений»
25. СНиП 23-05-95. «Естественное и искусственное освещение».
26. СНиП 23-05-2010 «Нормы Коэффициента пульсации помещений административных зданий».
27. СН 2.2.4/2.1.8.562-96 «Шум на рабочих местах, в помещениях жилых, общественных зданий и на территории жилой, застройки».
28. ПОТ РО 29-002-94. «Правила по охране труда для издательств».
29. СанПиН 2.2.4.1191-03. «Электромагнитные поля в производственных условиях».
30. ГОСТ 12.1.009-76 «Электробезопасность. Термины и определения».
31. СНиП 21-01-97. «Пожарная безопасность зданий и сооружений».
32. НПБ 105-95. «Определение категорий помещений и зданий по взрывопожарной и пожарной опасности».
33. ГОСТ Р 50739-95 «Средства вычислительной техники. Защита от несанкционированного доступа к информации».
34. ГОСТ 30772-2001 «Ресурсосбережение. Обращение с отходами. Термины и определения».
35. Федеральный закон от 8 августа 2001 г. №128-ФЗ «О лицензировании отдельных видов деятельности».

36. ГОСТ 12.2.032-78 «ССБТ. Рабочее место при выполнении работ сидя. Общие эргономические требования».

37. ГОСТ 12.2.061-81 «ССБТ. Оборудование производственное. Общие требования безопасности к рабочим местам».

38. Приказ Минздравсоцразвития РФ от 16.05.2005 N 83 («Об утверждении перечней вредных и (или) опасных производственных факторов и работ, при выполнении которых проводятся предварительные и периодические медицинские осмотры (обследования), и Порядка проведения этих осмотров (обследований)»).

39. ТК РФ, Статья 147. Оплата труда работников, занятых на работах с вредными и (или) опасными условиями труда.

40. ТК РФ, Статья 372. Порядок учета мнения выборного органа первичной профсоюзной организации при принятии локальных нормативных актов.

41. ГОСТ Р 55090-2012 «Ресурсосбережение. Обращение с отходами. Рекомендации по утилизации отходов бумаги».

42. ГОСТ 10700-97 «Макулатура бумажная и картонная. Технические условия».

43. ГОСТ Р 56276-2014 «Газы парниковые. Углеродный след продукции. Требования и руководящие указания по количественному определению и предоставлению информации».

44. ГОСТ Р ИСО 14046-2017 «Экологический менеджмент. Водный след. Принципы, требования и руководящие указания».

ПРИЛОЖЕНИЕ А

(справочное)

DEVELOPMENT OF ALGORITHM FOR HUMAN EXPRESSION RECOGNITION EXPERIMENTS AND RESULTS

Студент

Группа	ФИО	Подпись	Дата
8BM73	Кузнецова А.С.		

Консультант Отделения информационных технологий ИШИТР:

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

Консультант-лингвист отделения иностранных языков ШБИП

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

2 DEVELOPMENT OF ALGORITHM FOR HUMAN EXPRESSION RECOGNITION

2.1 Proposed architecture

The proposed deep Convolutional Neural Network architecture (depicted in Figure 2.1) consists of four parts. The first part automatically preprocesses the data. This begins with Convolution 1, which applies 64 different filters. The next layer is Pooling 1, which down-samples the images and then they are normalized by LRN 1. The next steps are the two FeatEx (Parallel Feature Extraction Block) blocks, highlighted in Figure 2.1. They are the core of the proposed architecture and described later in this section. The features extracted by these blocks are forwarded to a fully connected layer, which uses them to classify the input into the different emotions. The described architecture is compact, which makes it not only fast to train, but also suitable for real-time applications. This is also important as the network was built with resource usage in mind.

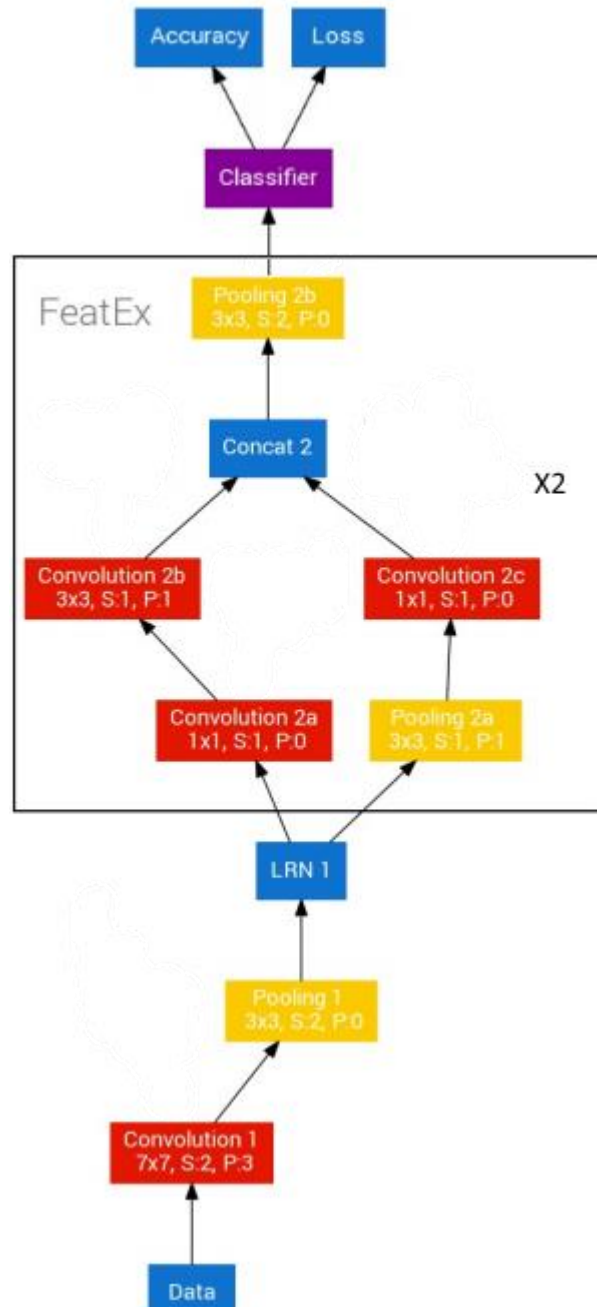


Figure 2.1 – Proposed architecture

The key structure in our architecture is the Parallel Feature Extraction Block (FeatEx). It is inspired by the success of GoogleNet. By reducing the computation bottleneck, the width and depth of the network can be increased. That is what inspired the creation of the proposed architecture. The original GoogleNet contains about 6 million parameters. The proposed model contains 417120 parameters.

The block consists of Convolutional, Pooling, and ReLU Layers. The first Convolutional layer in FeatEx reduces the dimension since it convolves with a filter of size 1×1 . It is enhanced by a ReLU layer, which creates the desired sparseness.

The output is then convolved with a filter of size 3×3 . In the parallel path a Max Pooling layer is used to reduce information before applying a CNN of size 1×1 . This application of differently sized filters reflects the various scales at which faces can appear. The paths are concatenated for a more diverse representation of the input. Using this block twice yields good results.

The different layers of the architecture produce feature vectors as can be seen in Figure 2.2.

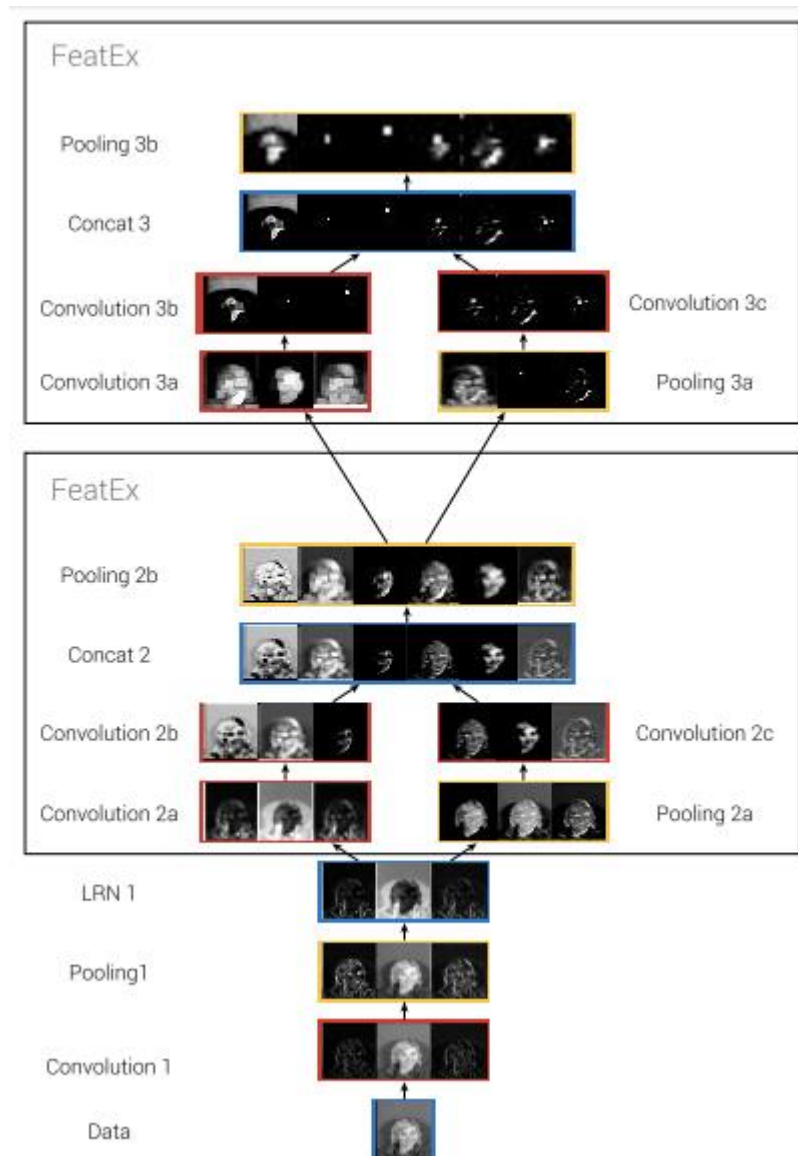


Figure 2.2 – Example visualizations of the different layers. The data is taken from the MMI set

The first part until LRN 1 preprocesses the data and creates multiple modified instances of the input. These show mostly edges with a low level of abstraction. The

first FeatEx block creates two parallel paths of features with different scales, which are combined in Concat 2.

The second FeatEx block refines the representation of the features. It also decreases the dimensionality. This visualization shows that the concatenation of FeatEx blocks is a valid approach to create an abstract feature representation. The output dimensionality of each layer can be seen in Table 2.1.

Table 2.1 –The different output sizes produced by each layer

Layer	Output size
Data	224×224
Convolution 1	$64 \times 112 \times 112$
Pooling 1	$64 \times 56 \times 56$
LRN 1	$64 \times 56 \times 56$
Convolution 2a	$96 \times 56 \times 56$
Convolution 2b	$208 \times 56 \times 56$
Pooling 2a	$64 \times 56 \times 56$
Convolution 2c	$64 \times 56 \times 56$
Concat 2	$272 \times 56 \times 56$
Pooling 2b	$272 \times 28 \times 28$
Convolution 3a	$96 \times 28 \times 28$
Convolution 3b	$208 \times 28 \times 28$
Pooling 3a	$272 \times 28 \times 28$
Convolution 3c	$64 \times 28 \times 28$
Concat 3	$272 \times 28 \times 28$
Pooling 3b	$282 \times 14 \times 14$
Classifier	$7 \times 1 \times 1$

The described architecture is compact, which makes it not only fast to train, but also suitable for real-time applications.

The performance of the proposed architecture endorses the efficacy and reliable usage of the proposed work for real world applications.

2.2 DATASETS

The following datasets were used: The Extended Cohn-Kanade (CKP) and MMI Facial Expression Database.

2.2.1 MMI Dataset

The MMI dataset contains over 2900 videos and 740 images of 75 persons. The annotations contain action units and emotions. The database contains a web-interface with an integrated search to scan the database. The videos/images are colored. The people are of mixed age, different gender and have different ethnical background. The emotions investigated are the six basic emotions: Anger, Disgust, Fear, Happiness, Sadness, Surprise.

From each video the 20 frames, which represent the content of the video the most, have been extracted fully automatically. The first two of these frames have been discarded since they provide neutral expressions.

To determine the frames, the difference between grayscale consecutive frames was calculated. To compensate noise the images have been smoothed using a Gaussian filter before calculation. In total 3740 images have been extracted.

For the MMI dataset, currently the best accuracy for emotion recognition is 93.33%.

2.2.2 CKP Dataset

This dataset presented by contains recordings of emotions of 210 persons at the ages of 18 to 50 years. Both female and male persons are present.

The images are of size 640×490 px as well 640×480 px. They are both grayscale and colored. In total this set has 593 emotion-labeled sequences. The emotions consist of Anger, Disgust, Fear, Happiness, Sadness, Surprise, and Contempt.

Thus for a better evaluation it is helpful to investigate multiple datasets. This aims at investigating whether the proposed approach works on different ways emotions are shown and whether it works on different emotions.

As a result of comparing these datasets, it can be noted that the emotion of contempt is contained only in the CKP dataset.



Figure 2.5 – Example images from the MMI (top) and CKP (bottom). The emotions from left to right are: Anger, Sadness, Disgust, Happiness, Fear, Surprise. The emotion Contempt of the CKP set is not displayed

For the CKP dataset, currently the best accuracy for emotion recognition is 99.2%.

3.3 EXPERIMENTS AND RESULTS

Testing was performed on a PC of the following configuration:

- Processor: Intel Core i7-4500 CPU @ 1.80 Ghz 2.40 Ghz;
- RAM: 8 GB;
- operating system: Linux Ubuntu.

Standard data sets were used: Extended Cohn-Kanade (CKP) and part of the MMI Facial Expression Database for testing the model. The reasons for this choice are their popularity and other authors use these datasets for testing their algorithms. Thus, cross-validation on different data sets will allow an objective assessment of the quality of the developed algorithm.

CKP dataset

The CKP database has been analyzed often and many different approaches have been evaluated in order to "solve" this set. To determine whether the architecture is competitive, it has been evaluated on the CKP dataset. For the experiments all 5870 annotated images have been used to do a 10-fold cross-validation. The proposed architecture has proven to be very effective on this dataset with an average accuracy of 97.6%.

In Table 3.1 different results from state of the art approaches are listed as comparison.

Table 3.1 – Comparison of accuracy of methods on the CKP dataset

Author	Method	Accuracy
Aliya Zafar	NCC	90%
Happy et al.	Facial Patches + SVM	94.09%
Lucey et al.	AAM + SVM	$\geq 80\%$
Song et al.	ANN (CNN)	99.2%
Kuznetsova Anastasia	Developed method	97.6%

The 90% accuracy reported by Zafar is based on hand picked images. The results are not validated using cross-validation.

MMI dataset

The original images were used for testing. A 10-fold cross-validation has been applied. The lowest accuracy is achieved by the emotion Fear with 93.75%. Happiness is recognized with 98.21%. Fear and Surprise are the emotions confused the most. This is better than the accuracies achieved by Wang and Yin (Table 3.2).

Table 3.2 – Summary the current state of the art in emotion recognition on the MMI database

Author	Method	Accuracy
Wang and Yin	LDA	93.33%
Wang and Yin	QDC	92.78%
Wang and Yin	NBC	85.56%
Kuznetsova Anastasia	Developed method	98.36%

Thus, by reducing the number of network parameters and model training time, it is possible to obtain a good result comparable to GoogleNet. The proposed architecture is independent of any hand-crafted feature extraction and performs better than the earlier proposed convolutional neural network based approaches.

ПРИЛОЖЕНИЕ Б
CD-ДИСК