

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

Институт Кибернетики
Направление подготовки Информатика и вычислительная система
Кафедра Информатики и проектирование систем

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Разработка программных средств для распознавания нечётких дубликатов изображений с помощью кластеризации их пространственных и яркостных характеристик

УДК 004.415:004.93'12

Студент

Группа	ФИО	Подпись	Дата
8ВМ4Б	Петрик Ирина Михайловна		

Руководитель

Должность	ФИО	Учёная степень, звание	Подпись	Дата
Доцент каф. ИПС	Стоянов Александр Кириллович	к.т.н.		

КОНСУЛЬТАНТЫ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Учёная степень, звание	Подпись	Дата

По разделу «Социальная ответственность»

Должность	ФИО	Учёная степень, звание	Подпись	Дата

ДОПУСТИТЬ К ЗАЩИТЕ:

Зав. кафедрой	ФИО	Учёная степень, звание	Подпись	Дата

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

Институт Кибернетики
Направление подготовки Информатика и вычислительная система
Кафедра Информатики и проектирование систем

УТВЕРЖДАЮ:
Зав. кафедрой

(Подпись) (Дата) (Ф.И.О.)

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

магистерской диссертации

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

Студенту:

Группа	ФИО
8ВМ4Б	Петрик Ирина Михайловна

Тема работы:

Разработка программных средств для распознавания нечётких дубликатов изображений с помощью кластеризации их пространственных и яркостных характеристик	
Утверждена приказом директора (дата, номер)	

Срок сдачи студентом выполненной работы:	
--	--

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе <i>(наименование объекта исследования или проектирования; производительность или нагрузка; режим работы (непрерывный, периодический, циклический и т. д.); вид сырья или материал изделия; требования к продукту, изделию или процессу; особые требования к особенностям функционирования (эксплуатации) объекта или изделия в плане безопасности эксплуатации, влияния на окружающую среду, энергозатратам; экономический анализ и т. д.).</i>	Исходными материалами являются коллекции изображений лиц. Объектом исследования и проектирования является разработка программных средств распознавания изображений лиц в коллекции по заданному образцу. Требования к разрабатываемому программному комплексу: 1. Возможность выполнять распознавание с помощью кластеризации только яркостной характеристики. 2. Возможность выполнять распознавание с
--	---

	помощью кластеризации яркостной и координатной характеристик. 3. Возможность выбора распределения: ранговое/естественное 4. Возможность выбора параметров сравнения: косинусное расстояние/эвклидова метрика. 5. Вывод по требованию пользователя графиков. 6. Сохранение необходимой информации в выбранную папку.
Перечень подлежащих исследованию, проектированию и разработке вопросов <i>(аналитический обзор по литературным источникам с целью выяснения достижений мировой науки техники в рассматриваемой области; постановка задачи исследования, проектирования, конструирования; содержание процедуры исследования, проектирования, конструирования; обсуждение результатов выполненной работы; наименование дополнительных разделов, подлежащих разработке; заключение по работе).</i>	1. Обзор литературы 2. Постановка задачи 3. Обоснование и выбор средств программирования 4. Описание алгоритма распознавания изображений лиц на основе кластеризации их пространственно-яркостных характеристик 5. Экспериментальные результаты 6. Заключение и выводы
Перечень графического материала <i>(с точным указанием обязательных чертежей)</i>	
Консультанты по разделам выпускной квалификационной работы <i>(с указанием разделов)</i>	
Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Конотопский В.Ю.
Социальная ответственность	Акулов П.А.
Названия разделов, которые должны быть написаны на русском и иностранном языках:	
Расчёты и аналитика	

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

Задание выдал руководитель:

Должность	ФИО	Учёная степень, звание	Подпись	Дата
Доцент каф. ИПС	Стоянов Александр Кириллович	к.т.н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ВМ4Б	Петрик Ирина Михайловна		

Запланированные результаты обучения по программе

Код результатов	Результат обучения (выпускник должен быть готов)	Требования ФГОС ВО (ФГОС 3+), критерии АИОР, заинтересованных работодателей и студентов
Общепрофессиональные компетенции		
Р1	Воспринимать и самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте.	Требования ФГОС 3+ (ОПК-1; ПК 3-6; ОК-4), критерий 5 АИОР (п. 1.1), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
Р2	Владеть и применять методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе в глобальных компьютерных сетях.	Требования ФГОС 3+ (ОПК-5; ПК-7; ОК-7), критерий 5 АИОР (п. 1.1, 1.2), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и

Код результата тов	Результат обучения (выпускник должен быть готов)	Требования ФГОС ВО (ФГОС 3+), критерии АИОР, заинтересованных работодателей и студентов
		зарубежных работодателей.
Р3	Демонстрировать культуру мышления, способность выстраивать логику рассуждений и высказываний, основанных на интерпретации данных, интегрированных из разных областей науки и техники, выносить суждения на основании неполных данных, анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями.	Требования ФГОС 3+ (ОПК-6; ПК-1,2; ОК-1,2), критерий 5 АИОР (п. 1.2), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
Р4	Анализировать и оценивать уровни своих компетенций в сочетании со способностью и готовностью к саморегулированию дальнейшего образования и профессиональной мобильности. Владеть, по крайней мере, одним из иностранных языков на уровне социального и профессионального общения, применять специальную лексику и профессиональную терминологию языка.	Требования ФГОС 3+ (ОПК-3,4; ПК-11,12; ОК-3), критерий 5 АИОР (п. 1.6, п. 2.2), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов,

Код результатов	Результат обучения (выпускник должен быть готов)	Требования ФГОС ВО (ФГОС 3+), критерии АИОР, заинтересованных работодателей и студентов
		отечественных и зарубежных работодателей.
Профессиональные компетенции		
Р5	Выполнять инновационные инженерные проекты по разработке аппаратных и программных средств автоматизированных систем различного назначения с использованием современных методов проектирования, систем автоматизированного проектирования, передового опыта разработки конкурентно способных изделий.	Требования ФГОС 3+ (ПК-8–12; ОПК-2, ПК-7,6), критерий 5 АИОР (п. 1.3), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
Р6	Планировать и проводить теоретические и экспериментальные исследования в области проектирования аппаратных и программных средств автоматизированных систем с использованием новейших достижений науки и техники, передового	Требования ФГОС 3+ (ПК-1–7; ОПК-6; ОК-4,9), критерий 5 АИОР (п.1.4),

Код результатов	Результат обучения (выпускник должен быть готов)	Требования ФГОС ВО (ФГОС 3+), критерии АИОР, заинтересованных работодателей и студентов
	отечественного и зарубежного опыта. Критически оценивать полученные данные и делать выводы.	соответствующий международным стандартам EUR- ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
P7	Осуществлять авторское сопровождение процессов проектирования, внедрения и эксплуатации аппаратных и программных средств автоматизированных систем различного назначения.	Требования ФГОС 3+ (ПК-13–19; ОПК-5; ОК-8), критерий 5 АИОР (п. 1.5), соответствующий международным стандартам EUR- ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
Общекультурные компетенции		
P8	Использовать на практике умения и навыки в организации исследовательских,	Требования ФГОС

Код результатов	Результат обучения (выпускник должен быть готов)	Требования ФГОС ВО (ФГОС 3+), критерии АИОР, заинтересованных работодателей и студентов
	проектных работ и профессиональной эксплуатации современного оборудования и приборов, в управлении коллективом.	3+ (ОК-5,8; ОПК-1,6; ПК-6,7,11,12), критерий 5 АИОР (п. 2.1, п. 2.3, п. 1.5), соответствующий международным стандартам EUR- ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
Р9	Осуществлять коммуникации в профессиональной среде и в обществе в целом, активно владеть иностранным языком, разрабатывать документацию, презентовать и защищать результаты инновационной инженерной деятельности, в том числе на иностранном языке.	Требования ФГОС 3+ (ОК-2,9; ОПК-4; ПК-1), критерий 5 АИОР (п. 2.2), соответствующий международным стандартам EUR- ACE и FEANI. Запросы студентов, отечественных и

Код результата тов	Результат обучения (выпускник должен быть готов)	Требования ФГОС ВО (ФГОС 3+), критерии АИОР, заинтересованных работодателей и студентов
		зарубежных работодателей.
Р10	Совершенствовать и развивать свой интеллектуальный и общекультурный уровень. Проявлять инициативу, в том числе в ситуациях риска, брать на себя всю полноту ответственности.	Требования ФГОС 3+ (ОК-1,6; ОПК-2; ПК-1,2), критерий 5 АИОР (п. 2.4, п. 2.5) , соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
Р11	Демонстрировать способность к самостоятельному обучению новым методам исследования, к изменению научного и научно-производственного профиля своей профессиональной деятельности, способность самостоятельно приобретать с помощью информационных технологий и использовать в практической деятельности новые знания и умения, в	Требования ФГОС 3+ (ОК-3,4,7; ОПК-3; ПК-7), критерий 5 АИОР (п. 2.6), соответствующий международным стандартам EUR-

Код результатов	Результат обучения (выпускник должен быть готов)	Требования ФГОС ВО (ФГОС 3+), критерии АИОР, заинтересованных работодателей и студентов
	том числе в новых областях знаний, непосредственно не связанных со сферой деятельности, способность к педагогической деятельности.	АСЕ и FEANI. Запросы студентов, отечественных и зарубежных работодателей.

РЕФЕРАТ

Выпускная квалификационная работа 110 с., 10 рис., 19 табл., 22 источника, 3 прил.

Ключевые слова: распознавание нечётких дубликатов, кластеризация, сегментация, нейронная рекуррентная сеть.

Объектом исследования является разработка программных средств распознавания изображений лиц в коллекции по заданному образцу.

Цель работы – исследование возможности распознавания нечётких дубликатов с помощью кластеризации их пространственно-яркостных характеристик.

В процессе исследования проводились эксперименты по распознаванию лиц из коллекции, отобранных Либором Спасеком.

В результате исследования был сделан вывод о том, кластеризация пространственно-яркостных характеристик изображений позволяет распознавать нечёткие дубликаты точнее, чем кластеризация яркости.

Основные конструктивные, технологические и технико-эксплуатационные характеристики: созданное программное обеспечение включает в себя следующие модули – кластеризации, сегментации, загрузки изображений и вычисления их меры близости загруженных.

Степень внедрения: работа была опубликована в Сборнике трудов XIII Международной научно-практической конференции студентов, аспирантов и молодых учёных «Молодёжь и современные информационные технологии».

Область применения: цифровые базы данных, системы компьютерного зрения, поиск изображений в Интернете, военная отрасль.

Экономическая эффективность/значимость работы Оценка экономической эффективности данного проекта невозможна, так как не определена точная область применения.

В будущем планируется использовать полученные результаты для решения задачи распознавания образов для изображений.

Оглавление

Введение	2
1. Обзор литературы	3
2. Объекты и методы исследования	6
2.1. Постановка задачи	6
2.2. Обзор средств, использующихся для реализации проекта	6
3. Расчеты и аналитика	Ошибка! Закладка не определена.
3.1. Теоретический анализ.....	Ошибка! Закладка не определена.
3.2. Структура программы.....	Ошибка! Закладка не определена.
3.3 Описание выполнения программы	Ошибка! Закладка не определена.
4. Результаты проведенного исследования	Ошибка! Закладка не определена.
5. Финансовый менеджмент, ресурсоэффективность	9
5.1 Организация и планирование работ	9
5.1.1 Продолжительность этапов работ.....	10
5.1.2 Расчет накопления готовности проекта.....	33
5.2 Расчет сметы затрат на выполнение проекта.....	34
5.2.1 Расчет затрат на материалы	35
5.2.2 Расчет заработной платы	36
5.2.3 Расчет затрат на социальный налог	37
5.2.4 Расчет затрат на электроэнергию	37
5.2.5 Расчет амортизационных расходов.....	39
5.2.6 Расчет прочих расходов	40
5.2.7 Расчет общей себестоимости разработки.....	40

5.2.8 Расчет прибыли	41
5.2.9 Расчет НДС	42
5.2.10 Цена разработки НИР	42
5.3 Оценка экономической эффективности проекта	42
5.3.1 Оценка научно-технического уровня НИР	42
6. Социальная ответственность	Ошибка! Закладка не определена.
6.1. Производственная безопасность	Ошибка! Закладка не определена.
6.1.1. Вредные факторы и меры защиты от них	Ошибка! Закладка не определена.
6.1.2. Опасные факторы и меры защиты от них	Ошибка! Закладка не определена.
6.1.3. Мероприятия по снижению вредных и опасных факторов	Ошибка! Закладка не определена.
6.2 Экологическая безопасность	Ошибка! Закладка не определена.
6.2.1 Анализ воздействия	Ошибка! Закладка не определена.
6.2.2. Мероприятия по снижению воздействия на окружающую среду	Ошибка! Закладка не определена.
6.3 Безопасность в чрезвычайных ситуациях	Ошибка! Закладка не определена.
6.3.1. Анализ воздействия ЧС	Ошибка! Закладка не определена.
6.3.2 Мероприятия по снижению вероятности возникновения ЧС	Ошибка! Закладка не определена.
6.4. Правовые вопросы обеспечения охраны труда ...	Ошибка! Закладка не определена.
6.4.1. Характерные для проектируемой рабочей зоны правовые нормы госконтроля охраны труда	Ошибка! Закладка не определена.
6.4.2. Организационные мероприятия при компоновке рабочей зоны	Ошибка! Закладка не определена.

Заключение	47
Список публикаций студента	48
Список использованных источников:.....	Ошибка! Закладка не определена.

Введение

С развитием вычислительной техники стало возможным решить ряд задач, возникающих в процессе жизнедеятельности, облегчить, ускорить, повысить качество результата. К одному из методов, позволяющих облегчить работу отдельных видов деятельности, относится распознавание изображений.

Распознавание изображений используется во многих сферах жизнедеятельности: поиск изображений по оригиналу в Интернете, цифровые архивы и библиотеки; анализ содержания космических и аэрофотоснимков, системы технического зрения и т.д. [1]. Одной из проблем, возникающих при решении этой задачи, является распознавание изображений, являющихся несколько искажёнными копиями оригинала – так называемых нечётких дубликатов.

Целью данной работы является создание программного обеспечения для исследования возможности распознавания нечётких дубликатов изображений, основанного на кластеризации их пространственно-яркостных характеристик.

1. Обзор литературы

Ранее в [2] была показана возможность успешного распознавания искажённых копий (нечётких дубликатов) изображения. Для распознавания было применено сравнение ранговых распределений относительных мощностей кластеров их яркости.

Пиксель любого изображения имеет три независимых характеристики – яркость и две пространственные координаты. В [2] для распознавания использована только одна из характеристик – яркость, которая была кластеризована по близким значениям. Но два изображения разных объектов могут иметь очень схожие распределения яркости пикселей, и, как следствие, после кластеризации (сегментации) они будут признаны похожими. Очевидно, это приводит к ошибкам и ограничивает возможности описанного метода. Чтобы избежать таких ошибок, было решено разработать метод, который включает в себя кластеризацию уже трёх характеристик (яркость и координаты x и y) рекуррентной нейронной сетью. Это позволит понять, в какой части изображения находятся пиксели со схожей яркостью.

Существует несколько методов кластеризации, которые основаны на различных представлениях о задаче [3]. Первый подход – вероятностный, к нему относится метод k -средних, k -medians и др. Второй подход основывается на системах искусственного интеллекта: нейронная сеть Кохонена, генетический алгоритм и др. Следующий подход основан на теории графов. Самый популярный алгоритм с использованием этого подхода – алгоритм кратчайшего незамкнутого пути. Большинство этих и других методов имеют один главный недостаток - необходимость задания числа кластеров. Поэтому было решено использовать в данной работе метод

кластеризации рекуррентной нейронной сетью, в котором нет этой необходимости.

Также следует отметить, что многообразие алгоритмов распознавания можно поделить на три группы. В первой группе используется сравнение характерных признаков лица. При этом все алгоритмы, входящие в эту группу, имеют два этапа. Первый этап включает в себя детектирование и локализация лица на изображении. На втором этапе осуществляется выравнивание изображения лица (геометрическое и яркостное), выделение признаков и распознавание, т.е. сравнение вычисленных признаков с заложенными в базу данных эталонами. Недостаток данных алгоритмов – это трудоёмкость; выделение признаков требует существенных затрат времени и вычислительных ресурсов.

Во вторую группу входят нейросетевые алгоритмы [5-7]. Нейронные сети (НС) проходят обучение с помощью учебных примеров. Процесс обучения НС включает в себя выделение ключевых признаков, вычисление их важности и построение взаимосвязей между ними. Свёрточная НС показывает самые лучшие результаты в области распознавания лиц [7], так как она может не зависеть от изменений масштаба, смещений, поворотов, смены ракурса и прочих искажений. Недостатком НС также можно назвать трудоёмкость алгоритма. Если добавить новый эталон в базу данных, потребуется полное переобучение сети на всем имеющемся наборе. Помимо этого, имеются проблемы, связанные с обучением: выбор наилучшего шага оптимизации, попадание в локальный оптимум, переобучение. Также выбор количества нейронов, слоёв и пр. могут вызвать затруднение.

Третья группа алгоритмов не использует признаки, что существенно снижает трудоёмкость распознавания. Данный метод, основанный на таком

принципе, называется беспризнаковым распознаванием. Решение данной задачи опирается на гипотезу компактности, смысл которой в предположении, что объекты со схожими свойствами чаще всего попадают в один класс, чем в разные.

В соответствии с постановкой этой задачи предполагается, что каждый объект распознавания может быть представлен результатами попарных сравнений с оригиналом. Для сравнения используется произвольная действительная функция, называемая расстоянием. Решение о том, схож ли объект (нечёткий дубликат изображения) с оригиналом, в классической постановке обычно строится на основе превышения порога, установленного решающим правилом. Другими словами, решение сводится к сравнению полученного расстояния и определённого числа – порога. Если расстояние превышает порог, считается, что сравниваемое изображение не является нечётким дубликатом оригинала.

2. Объекты и методы исследования

2.1. Постановка задачи

Задачей данной работы является разработка программного средства для распознавания нечётких дубликатов изображений.

Программа должна предоставлять следующие возможности:

1. Выполнение распознавания с помощью кластеризации только яркостной характеристики.
2. Выполнение распознавания с помощью кластеризации яркостной и координатной характеристик.
3. Выбор распределения: ранговое/естественное
4. Выбор параметров сравнения: косинусное расстояние/эвклидова метрика.
5. Вывод по требованию пользователя графиков.
6. Сохранение необходимой информации в выбранную папку.

В ходе работы планируется провести следующие эксперименты:

1. Распознавание изображений с помощью кластеризации пространственно-яркостных характеристик.
2. Распознавание изображений с помощью кластеризации только яркости.

После экспериментов будет произведено сравнение полученных значений.

2.2. Обзор средств, использующихся для реализации проекта

Средой разработки для реализации дипломного проекта была выбрана Microsoft Visual Studio 2015. Microsoft Visual Studio — серия продуктов компании [Майкрософт](#), которая включает в себя интегрированную среду разработки программного обеспечения, а также много других инструментальных средств. С помощью данных продуктов можно

разрабатывать не только [консольные приложения](#), но и приложения с [графическим интерфейсом](#). В Microsoft Visual Studio поддерживаются технологии [Windows Forms](#), [веб-сайты](#), [веб-приложения](#), [веб-службы](#) как в [родном](#), так и в [управляемом](#) кодах для всех платформ, таких как [Microsoft Windows](#), [Windows Mobile](#), [Windows CE](#), [.NET Framework](#).

Основные достоинства: малая зависимость от собственной среды разработки, возможность как компиляции в машинный код, так и интерпретации во время отладки, большое количество библиотек

К недостаткам можно отнести платформенную зависимость конечного кода от компилятора.

Для проведения экспериментов использовались изображения из базы, собранной Либором Спасеком (Libor Spacek). По нашему мнению, они являются наиболее показательными. База включает в себя изображения 375 человек. Каждый человек имеет по 20 изображений на однородном и неоднородном фонах при различных условиях освещения. Также база содержит в себе и изображения лиц, сделанными с разными положениями камеры и наличием эмоций. [8]

5. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

5.1 Организация и планирование работ

Перечень этапов и работ, распределение исполнителей по данным видам работ приведен в таблице 4.

Таблица 4. Перечень работ и продолжительность их выполнения

№ раб	Этапы работы	Исполнители	Загрузка исполнителей
1	Постановка целей и задач, получение исходных данных	НР	НР – 100%
2	Составление и утверждение технического задания	НР	НР – 100%
3	Подбор и изучение материалов по тематике	НР, И	НР – 35% И – 100%
4	Календарное планирование работ по теме	НР, И	НР – 100% И – 10%
5	Разработка алгоритма для решения задачи	НР, И	НР – 40% И – 100%
6	Проектирование общей структуры программы	НР, И	НР – 10% И – 100%
7	Кодирование	И	И – 100%
8	Отладка и тестирование	И	И – 100%

	программы		
9	Оформление расчетно- пояснительной записки	И	И – 100%

5.1.1 Продолжительность этапов работ

Трудовые затраты в большинстве случаев образуют основную часть стоимости разработки, поэтому важным моментом является определение трудоемкости работ каждого из участников научного исследования.

Для определения вероятных (ожидаемых) значений продолжительности работ $t_{ож}$ используется следующая формула:

$$t_{ож} = \frac{3 \cdot t_{min} + 2 \cdot t_{max}}{5},$$

где t_{min} – минимальная продолжительность работы, дн.;

t_{max} – максимальная продолжительность работы, дн.;

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

$$T_{РД} = \frac{t_{ож}}{K_{ВН}} \cdot K_{Д},$$

где $t_{ож}$ – продолжительность работы, дн.;

$K_{ВН}$ – коэффициент выполнения работ, учитывающий влияние внешних факторов на соблюдение предварительно определенных длительностей, в частности ($K_{ВН} = 1$);

K_d – коэффициент, учитывающий дополнительное время на компенсацию непредвиденных задержек и согласование работ ($K_d = 1,15$).

Расчет продолжительности этапа в календарных днях ведется по формуле:

$$T_{KD} = T_{RD} \cdot T_K,$$

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

T_K – коэффициент календарности, позволяющий перейти от длительности работ в рабочих днях к их аналогам в календарных днях, и рассчитываемый по формуле

$$T_K = \frac{T_{КАЛ}}{T_{КАЛ} - T_{ВД} - T_{ПД}},$$

где $T_{КАЛ}$ – календарные дни ($T_{КАЛ} = 366$);

$T_{ВД}$ – выходные дни ($T_{ВД} =$);

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

$$T_K = \frac{366}{366 - 66} = 1,22$$

Таблица 5. Трудозатраты на выполнение проекта

Этап	Исполнители	Продолжительность работ, дни			Трудоемкость работ по исполнителям чел.- дн.			
					$T_{\text{РД}}$		$T_{\text{КД}}$	
		t_{min}	t_{max}	$t_{\text{ож}}$	НР	И	НР	И
1	2	3	4	5	6	7	8	9
Постановка целей и задач, получение исходных данных	НР	2	4	2,8	3,22	–	3,93	–
Составление и утверждение технического задания	НР	2	3	2,4	2,76	–	3,37	–
Подбор и изучение материалов по тематике	НР, И	11	15	12,6	5,07	14,49	6,19	17,68
Календарное планирование работ по теме	НР, И	3	7	4,6	5,29	0,53	6,45	0,65

Разработка алгоритма для решения задачи	НР, И	11	16	13	5,98	14,95	7,30	18,24
Проектирование общей структуры программы	НР, И	6	10	7,6	0,87	8,74	1,06	10,66
Кодирование	И	14	24	18	–	20,7	–	25,25
Отладка и тестирование программы	И	4	7	5,2	–	5,98	–	7,30
Оформление расчетно- пояснительной записки	И	5	9	6,6	–	7,59	–	9,26
Итого:				72,8	23,19	72,98		

Таблица 6- Линейный график работ

Этап	НР	И	Март			Апрель			Май			Июнь	
			10	20	30	40	50	60	70	80	90	100	110
1	3,93	–	■										
2	3,37	–		■									
3	6,19	17,68		■	■								
4	6,45	0,65											
5	7,30	18,24				■	■	■					
6	1,06	10,66						■	■	■			
7	–	25,25							■	■	■	■	
8	–	7,30											
9	–	9,26										■	■

НР – ■ ; И – ■

5.1.2 Расчет накопления готовности проекта

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

Введем обозначения:

- $TP_{\text{общ.}}$ – общая трудоемкость проекта;
- TP_i (TP_k) – трудоемкость i-го (k-го) этапа проекта, $i = \overline{1, I}$;
- TP_i^H – накопленная трудоемкость i-го этапа проекта по его завершении;
- TP_{ij} (TP_{kj}) – трудоемкость работ, выполняемых j-м участником на i-м этапе, здесь $j = \overline{1, m}$ – индекс исполнителя, в нашем примере $m = 2$.

Степень готовности определяется формулой :

$$CG_i = \frac{TP_i^H}{TP_{\text{общ.}}} = \frac{\sum_{k=1}^i TP_k}{TP_{\text{общ.}}} = \frac{\sum_{k=1}^i \sum_{j=1}^m TP_{km}}{\sum_{k=1}^I \sum_{j=1}^m TP_{km}}.$$

Применительно к таблице (2) величины TP_{ij} (TP_{kj}) находятся в столбцах (6, $j = 1$) и (7, $j = 2$). $TP_{\text{общ.}}$ равна сумме чисел из итоговых клеток этих столбцов.

Таблица 7. Нарастание технической готовности работы и удельный вес каждого этапа

Этап	TP_i , %	CG_i , %
Постановка задачи	3,34 8237	3,34 8237
Разработка и утверждение технического задания (ТЗ)	2,86 9917	6,21 8154

Подбор и изучение материалов по тематике	20,3 3898	26,5 57134
Разработка календарного плана	6,05 1783	32,6 08917
Обсуждение литературы	21,7 6354	54,3 72457
Выбор структурной схемы устройства	9,99 2721	64,3 65178
Выбор принципиальной схемы устройства	21,5 2438	85,8 89558
Расчет принципиальной схемы устройства	6,21 8155	92,1 07713
Оформление расчетно-пояснительной записки	7,89 2277	100

5.2 Расчет сметы затрат на выполнение проекта

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

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

- арендная плата за пользование имуществом;
- прочие услуги (сторонних организаций);
- прочие (накладные расходы) расходы.

5.2.1 Расчет затрат на материалы

К данной статье расходов относится стоимость материалов, покупных изделий, полуфабрикатов и других материальных ценностей, расходуемых непосредственно в процессе выполнения работ над объектом проектирования. Сюда же относятся специально приобретенное оборудование, инструменты и прочие объекты, относимые к основным средствам, стоимостью до 40 000 руб. включительно. Цена материальных ресурсов определяется по соответствующим ценникам или договорам поставки. Кроме того статья включает так называемые транспортно-заготовительные расходы, связанные с транспортировкой от поставщика к потребителю, хранением и прочими процессами, обеспечивающими движение (доставку) материальных ресурсов от поставщиков к потребителю. Сюда же включаются расходы на совершение сделки купли-продажи (т.н. транзакции). Приблизительно они оцениваются в процентах к отпускной цене закупаемых материалов, как правило, это $5 \div 20 \%$. Исполнитель работы самостоятельно выбирает их величину в указанных границах.

Таблица 8. Расчет затрат на материалы

Наименование материалов	Цена за ед., руб.	Кол- во	Сумма, руб.
Лицензия Microsoft Visual Studio 2015	26 394	1 экз.	26394
Бумага для принтера формата А4	200	1 уп.	200
Картридж для принтера	1250	1 шт.	1250
Итого:			27844

Допустим, что ТЗР составляют 7 % от отпускной цены материалов, тогда расходы на материалы с учетом ТЗР равны $C_{\text{мат}} = 27844 * 1,07 = 29793,08$ руб.

5.2.2 Расчет заработной платы

Данная статья расходов включает заработную плату научного руководителя и инженера (в его роли выступает исполнитель проекта), а также премии, входящие в фонд заработной платы. Расчет основной заработной платы выполняется на основе трудоемкости выполнения каждого этапа и величины месячного оклада исполнителя. Среднедневная тарифная заработная плата ($ЗП_{\text{дн-т}}$) рассчитывается по формуле:

$$ЗП_{\text{дн-т}} = MO/25$$

учитывающей, что в году 300 рабочих дней и, следовательно, в месяце в среднем 25 рабочих дня (при шестидневной рабочей неделе).

Расчеты затрат на полную заработную плату приведены в таблице 8. Затраты времени по каждому исполнителю в рабочих днях с округлением до целого взяты из таблицы 7. Для учета в ее составе премий, дополнительной

зарплаты и районной надбавки используется следующий ряд коэффициентов: $K_{\text{ПР}} = 1,1$; $K_{\text{доп.ЗП}} = 1,188$; $K_{\text{р}} = 1,3$. Таким образом, для перехода от тарифной (базовой) суммы заработка исполнителя, связанной с участием в проекте, к соответствующему полному заработку (зарплатной части сметы) необходимо первую умножить на интегральный коэффициент $K_{\text{и}} = 1,1 * 1,188 * 1,3 = 1,699$.

Таблица 9. Затраты на заработную плату

Исполнитель	Оклад, руб./мес.	ежедневная ставка, руб./раб.день	Затраты времени, раб.дни	Коэффициент	Фонд з/платы, руб.
НР	23 264,86	930,59	24	1,699	37945,7
И	14 874,45	594,98	73	1,62	70362,3
Итого:					108308

5.2.3 Расчет затрат на социальный налог

Затраты на единый социальный налог (ЕСН), включающий в себя отчисления в пенсионный фонд, на социальное и медицинское страхование, составляют 30 % от полной заработной платы по проекту, т.е. $C_{\text{соц.}} = C_{\text{зп}} * 0,3$. Итак, в нашем случае $C_{\text{соц.}} = 108308 * 0,3 = 32492,4$ руб.

5.2.4 Расчет затрат на электроэнергию

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

$$C_{\text{эл.об.}} = P_{\text{об}} \cdot t_{\text{об}} \cdot Ц_{\text{э}},$$

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

$Ц_{\text{э}}$ – тариф на 1 кВт·час;

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

Для ТПУ ЦЭ = 5,257 руб./кВт·час (с НДС).

Время работы оборудования вычисляется на основе итоговых данных таблицы 2 для инженера ($T_{РД}$) из расчета, что продолжительность рабочего дня равна 8 часов.

$$t_{об} = T_{РД} * K_t,$$

где $K_t \leq 1$ – коэффициент использования оборудования по времени, равный отношению времени его работы в процессе выполнения проекта к $T_{РД}$, определяется исполнителем самостоятельно. В ряде случаев возможно определение $t_{об}$ путем прямого учета, особенно при ограниченном использовании соответствующего оборудования ($K_t = 0,7$).

Мощность, потребляемая оборудованием, определяется по формуле:

$$P_{ОБ} = P_{ном.} * K_C,$$

где $P_{ном.}$ – номинальная мощность оборудования, кВт;

$K_C \leq 1$ – коэффициент загрузки, зависящий от средней степени использования номинальной мощности. Для технологического оборудования малой мощности $K_C = 1$.

Пример расчета затраты на электроэнергию для технологических целей приведен в таблице 9.

Таблица 10. Затраты на электроэнергию технологическую

Наименование оборудования	Время работы оборудования $t_{\text{ОБ}}$, час	Потребляемая мощность $P_{\text{ОБ}}$, кВт	Затраты ОБ, руб.
Персональный компьютер	583,84*0,7	0,4	859,40
Струйный принтер	25	0,1	13,14
Итого:			872,54

5.2.5 Расчет амортизационных расходов

В статье «Амортизационные отчисления» рассчитывается амортизация используемого оборудования за время выполнения проекта.

Используется формула:

$$C_{\text{АМ}} = \frac{N_{\text{А}} * C_{\text{ОБ}} * t_{\text{рф}} * n}{F_{\text{д}}},$$

где $N_{\text{А}}$ – годовая норма амортизации единицы оборудования;

$C_{\text{ОБ}}$ – балансовая стоимость единицы оборудования с учетом ТЗР. При невозможности получить соответствующие данные из бухгалтерии она может быть заменена действующей ценой, содержащейся в ценниках, прейскурантах и т.п.;

$F_{\text{д}}$ – действительный годовой фонд времени работы соответствующего оборудования, берется из специальных справочников или фактического режима его использования в текущем календарном году;

t_{pf} – фактическое время работы оборудования в ходе выполнения проекта, учитывается исполнителем проекта;

n – число задействованных однотипных единиц оборудования.

Для ПК в 2016 г. (300 рабочих дней при шестидневной рабочей неделе) можно принять $F_d = 300 * 8 = 2400$ часа;

Стоимость ПК 29380 руб., время использования 583 часа, тогда $C_{AM}(ПК) = ((1/2,5) * 29380 * 583 * 1) / 2400 = 2854,76$ руб. Стоимость принтера 7000 руб., его $F_d = 500$ час.; $N_A = 0,5$; тогда его $C_{AM}(Пр) = (0,5 * 7000 * 30 * 1) / 500 = 210$ руб.. Итого начислено амортизации 3064,76 руб.

5.2.6 Расчет прочих расходов

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

$$C_{\text{проч.}} = (C_{\text{мат}} + C_{\text{зп}} + C_{\text{соц}} + C_{\text{эл.об.}} + C_{\text{ам}} + C_{\text{нп}}) \cdot 0,1$$

Для нашего примера это:

$$C_{\text{проч.}} = (27844 + 108308 + 32492,4 + 872,54 + 3064,76) \cdot 0,1 = 17258,17 \text{ руб.}$$

5.2.7 Расчет общей себестоимости разработки

Проведя расчет по всем статьям сметы затрат на разработку, можно определить общую себестоимость проекта «Разработка программных средств для задачи сравнения изображений лиц на основе кластеризации меры их близости».

Таблица 11. Смета затрат на разработку проекта

Статья затрат	Условное обозначение	Сумма, руб.
Материалы и покупные изделия	$C_{\text{мат}}$	27844
Основная заработная плата	$C_{\text{зп}}$	108308
Отчисления в социальные фонды	$C_{\text{соц}}$	32492,4
Расходы на электроэнергию	$C_{\text{эл.}}$	872,54
Амортизационные отчисления	$C_{\text{ам}}$	3064,76
Прочие расходы	$C_{\text{проч}}$	17258,17
Итого:		189839,87

Таким образом, затраты на разработку составили $C = 189839,87$ руб.

5.2.8 Расчет прибыли

Прибыль от реализации проекта в зависимости от конкретной ситуации (масштаб и характер получаемого результата, степень его определенности и коммерциализации, специфика целевого сегмента рынка и т.д.) может определяться различными способами. В данном случае она составляет 37967,97 руб. (20 %) от расходов на разработку проекта.

5.2.9 Расчет НДС

НДС составляет 18% от суммы затрат на разработку и прибыли. В нашем случае это $(189839,87 + 37967,97) \cdot 0,18 = 41005,41$ руб.

5.2.10 Цена разработки НИР

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

$$C_{\text{НИР(КР)}} = 189839,87 + 37967,97 + 41005,41 = 268812,25 \text{ руб.}$$

5.3 Оценка экономической эффективности проекта

Оценка экономической эффективности данного проекта невозможна, так как требуется провести сложный и затратный анализ.

5.3.1 Оценка научно-технического уровня НИР

Научно-технический уровень характеризует влияние проекта на уровень и динамику обеспечения научно-технического прогресса в данной области. Для оценки научной ценности, технической значимости и эффективности, планируемых и выполняемых НИР, используется метод балльных оценок. Балльная оценка заключается в том, что каждому фактору по принятой шкале присваивается определенное количество баллов. Обобщенную оценку проводят по сумме баллов по всем показателям. На ее основе делается вывод о целесообразности НИР.

Сущность метода заключается в том, что на основе оценок признаков работы определяется интегральный показатель (индекс) ее научно-технического уровня по формуле:

$$K_{\text{НТУ}} = \sum_{i=1}^3 R_i \cdot n_i,$$

где $I_{\text{НТУ}}$ – интегральный индекс научно-технического уровня;

R_i – весовой коэффициент i -го признака научно-технического эффекта;

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

Таблица 12. Весовые коэффициенты признаков НТУ

Признаки научно-технического эффекта НИР	Характеристика признака НИР	Ri
Уровень новизны	Систематизируются и обобщаются сведения, определяются пути дальнейших исследований	00,4
Теоретический уровень	Разработка способа (алгоритм, программа мероприятий, устройство, вещество и т.п.)	00,1
Возможность реализации	Время реализации в течение первых лет	00,5

Таблица 13. Баллы для оценки уровня новизны

Уровень новизны	Характеристика уровня новизны – n_1	Баллы
Принципиально новая	Новое направление в науке и технике, новые факты и закономерности, новая теория, вещество, способ	8 – 10

Новая	По-новому объясняются те же факты, закономерности, новые понятия дополняют ранее полученные результаты	5 – 7
Относительно новая	Систематизируются, обобщаются имеющиеся сведения, новые связи между известными факторами	2 – 4
Не обладает новизной	Результат, который ранее был известен	0

Таблица 14. Баллы значимости теоретических уровней

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

Таблица 15. Возможность реализации результатов по времени

Время реализации – n_3	Баллы
В течение первых лет	10

От 5 до 10 лет	4
Свыше 10 лет	2

Так как все частные признаки научно-технического уровня оцениваются по 10-балльной шкале, а сумма весов R_i равна единице, то величина интегрального показателя также принадлежит интервалу $[0, 10]$.

В таблице 16 указано соответствие качественных уровней НИР значениям показателя.

Таблица 16. Соответствие уровней НИР

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

Для используемого в пособии примера частные оценки уровня p_i и их краткое обоснование даны в таблице 17.

Таблица 17. Оценки научно-технического уровня НИР

Значимость	Фактор НТУ	Уровень фактора	Выборанный балл	Обоснование выбранного балла
0,4	Уровень новизны	Относительно новая	5	Облегчит поиск схожих изображений

0,1	Теоретический уровень	Разрабо тка способа	6	Уменьшит количество возможных ошибок
0,5	Возможность реализации	В течение первых лет	10	Быстрое внедрение на любой компьютер в любой сфере

Отсюда интегральный показатель научно-технического уровня для нашего проекта составляет:

$$I_{\text{нту}} = 0,4 \cdot 5 + 0,1 \cdot 6 + 0,5 \cdot 10 = 2 + 0,6 + 5 = 7,6$$

Таким образом, исходя из данных таблицы 16, данный проект имеет средний уровень научно-технического эффекта.

Заключение

В ходе проделанной работы было разработано программное средство в среде Microsoft Visual Studio 2015 с поставленными задачами:

Задачей данной работы является разработка программного средства для распознавания нечётких дубликатов изображений.

Программа предоставляет следующие возможности:

1. Выполнение распознавания с помощью кластеризации только яркостной характеристики.
2. Выполнение распознавания с помощью кластеризации яркостной и координатной характеристик.
3. Выбор распределения: ранговое/естественное
4. Выбор параметров сравнения: косинусное расстояние/евклидова метрика.
5. Вывод по требованию пользователя графиков.
6. Сохранение необходимой информации в выбранную папку.

В ходе работы были проведены следующие эксперименты:

1. Распознавание изображений с помощью кластеризации пространственно-яркостных характеристик.
2. Распознавание изображений с помощью кластеризации только яркости.

После проведения всех экспериментов и сравнения полученных значений было сделано несколько выводов:

Во-первых, метод распознавания нечётких дубликатов изображений с использованием яркостно-координатных характеристик является более точным, чем метод с использованием только яркости.

Во-вторых, разработанный метод лучше всего работает с изображениями с естественными искажениями, т.е. изменения освещения и эмоций, чем с искусственно созданными изменениями, выполненными в графическом редакторе.

Список публикаций студента

N п/п	Наименование работы, ее вид	Форма работы	Выходные данные	Объем в п.л. или с.	Соавторы
1	2	3	4	5	6
статьи					
1	Распознавание изображений, основанное на ранговом распределении их пространственно-яркостных характеристик	Электрон.	Сборник трудов XIII Международной научно-практической конференции студентов, аспирантов и молодых учёных «Молодёжь и современные информационные технологии» Том I	2	-