

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра информатики и программирования

**РАЗРАБОТКА ПРИЛОЖЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ ОБЪЕКТОВ В
УСЛОВИЯХ ОГРАНИЧЕННОСТИ ДАННЫХ
АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ**

студента 4 курса 441 группы

направления 02.03.03 Математическое обеспечение и администрирование
информационных систем

факультета компьютерных наук и информационных технологий

Клочкова Сергея Михайловича

Научный руководитель:

доцент, к. ф.-м. н.

Огнева М.В.

подпись, дата

Зав. кафедрой:

доцент, к. ф.-м. н.

Огнева М.В.

подпись, дата

Саратов 2025

ВВЕДЕНИЕ

Актуальность темы.

Развитие мощностей вычислительной техники и технологий искусственного интеллекта позволило компьютерному зрению сделать качественный скачок вперед. В настоящее время CV-технологии применяются во многих жизненно важных отраслях, например, в медицине, ритейле и производстве.

Машинное обучение, а в частности глубинное обучение и нейронные сети, стало основополагающими методами в задачах компьютерного зрения. Но, несмотря на кажущуюся универсальность, искусственный интеллект имеет большое количество слабостей и ограничений, таких как требования к объему данных (к примеру, "проклятие размерности"). В таких ситуациях необходимо использовать так называемое "классическое" компьютерное зрение, то есть компьютерное зрение, построенное на алгоритмах. Основной сутью CV-алгоритмов являются выявление характерных признаков изображений и их дальнейшая обработка.

Использование алгоритмов компьютерного зрения вместо методов машинного обучения обусловлено жесткими ограничениями в размерах датасета (не всегда есть возможность сделать очень много фотографий одного и того же экспоната), возможностью быстрого обновления базы данных (добавление нового экспоната потребовало бы очень ресурсозатратного переобучения модели). Для распознавания сложных объектов на изображениях в условиях ограниченного по размерам датасета наилучшую эффективность показывают алгоритмы выявления ключевых точек.

Цель бакалаврской работы – Разработать приложение, позволяющее выполнять распознавание объектов в условиях ограниченности данных на примере музейных экспонатов.

Поставленная цель определила **следующие задачи**:

1. выполнить обзор аналогов разрабатываемого продукта;
2. изучить различные алгоритмы вычисления ключевых точек изображения;

3. установить необходимый инструментарий для работы и изучить его;
4. провести тестирование различных алгоритмов вычисления ключевых точек изображения и способов хранения их дескрипторов;
5. на основе проведённого тестирования выбрать алгоритм вычисления ключевых точек изображения и способ хранения дескрипторов;
6. спроектировать архитектуру серверного ПО и связь с клиентским приложением;
7. разработать приложение с использованием изученных инструментов.

Методологические основы разработки приложения для распознавания объектов в условиях ограниченности данных представлены в работах Уэйнмена, Хоуса, Кэлера, Ильященко и Горячкина.

Практическая значимость бакалаврской работы.

Практическая значимость работы определяется следующими факторами:

- разработано программное приложение, способное распознавать музейные экспонаты по изображениям даже при ограниченном наборе обучающих данных, что соответствует реальным условиям работы многих музеев;
- предложенные и реализованные подходы позволяют использовать приложение для создания цифровых гидов, интерактивных каталогов и систем дополненной реальности для музеев, не обладающих большими цифровыми архивами;
- приложение может быть интегрировано в образовательные, туристические и просветительские проекты, расширяя доступ посетителей к информации об экспонатах;
- представленное решение может быть адаптировано и использовано в других сферах, связанных с распознаванием визуальных объектов при нехватке данных (например, в библиотеках, архивах, галереях);
- результаты могут быть использованы для создания прототипов интеллектуальных систем помощи музейным сотрудникам в идентификации и каталогизации экспонатов.

Структура и объём работы. Бакалаврская работа состоит из введения, 10 разделов, заключения, списка использованных источников и 21 приложений. Общий объем работы – 140 страниц, из них 59 страниц – основное содержание, включая 11 рисунков и 1 таблицы, цифровой носитель в качестве приложения, список использованных источников информации – 25 наименований.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первый раздел «Обзор аналогов» посвящен анализу существующих решений в области распознавания объектов.

В разделе описаны современные системы, основанные на нейросетях, такие как Google Cloud Vision API, Amazon Rekognition и другие. Приведены их достоинства (высокая точность, гибкость, автоматизация) и недостатки (требуют больших объемов данных и вычислительных ресурсов). Сделан вывод о том, что в условиях ограниченности данных классические методы (например, по ключевым точкам) предпочтительнее.

Выводы: для ограниченных датасетов классические алгоритмы на основе ключевых точек являются более подходящим выбором, чем нейросетевые подходы.

Второй раздел «Обзор алгоритмов» посвящен теоретическому анализу алгоритмов детектирования ключевых точек.

Рассмотрены три алгоритма: BRISK, ORB и AKAZE. Приведены их принципы работы, математические формулы, схемы и сравнительные характеристики.

Выводы: каждый алгоритм имеет свои особенности.

Третий раздел «Используемые технологии» посвящен описанию программных средств, примененных в работе.

Описан выбор языка программирования (C++17), библиотек (OpenCV, spdlog, wfrest, GoogleTest, Doxygen, Swagger, Qt и др.), а также инструменты контейнеризации (Docker и docker-compose).

Выводы: рациональный выбор инструментов обеспечил производительность и масштабируемость приложения.

Четвертый раздел «Проектирование архитектуры приложения» посвящен созданию структуры программной системы.

Разработана двухуровневая клиент-серверная архитектура. Серверная часть построена по модульному принципу: модуль ядра, работы с сетью и базой данных.

Выводы: модульная архитектура обеспечивает гибкость и расширяемость. Обработка изображений происходит на сервере, клиент отправляет запросы.

Пятый раздел «Тестирование алгоритмов поиска ключевых точек» посвящен сравнению алгоритмов по метрикам скорости и точности. Проведено тестирование на датасете из фотографий музейных экспонатов. Рассмотрены различные подходы к хранению дескрипторов (массив, конкатенация и т.д.).

Выводы: оптимальным оказался алгоритм ORB с хранением объединенных дескрипторов. AKAZE показал хорошую точность, но был слишком медленным.

Шестой раздел «Разработка модуля базы данных» посвящен реализации хранения и поиска объектов.

Были реализованы:

- выбор и интеграция NoSQL СУБД Apache Cassandra;
- асинхронная работа с сопоставителями дескрипторов;
- эффективная локальная база в оперативной памяти;
- алгоритмы поиска, добавления и удаления объектов;
- пул сопоставителей с автоматическим обновлением.

Выводы: модуль оптимизирован под многопоточную работу, что обеспечивает высокую производительность и отказоустойчивость.

Седьмой раздел «Разработка модуля ядра» посвящен реализации основного функционала обработки изображений.

Были реализованы:

- логирование и конфигурация;
- вычисление дескрипторов;
- добавление и получение данных об экспонатах.

Выводы: модуль ядра успешно выполняет задачи подготовки и обработки изображений.

Восьмой раздел «Разработка модуля работы с сетью» посвящен организации взаимодействия сервера и клиента.

Реализован REST API, получающий изображения, возвращающий данные, выгружающий базу.

Выводы: модуль стабильно работает с HTTP-запросами, интегрирован с другими частями сервера.

Девятый раздел «Тестирование» посвящен проверке модулей системы.

Выполнены модульные тесты с использованием GoogleTest для базы данных, ядра и сетевого взаимодействия.

Выводы: все модули успешно прошли тестирование, обеспечена корректность работы.

Десятый раздел «Клиентские приложения» посвящен разработке интерфейсов взаимодействия с пользователем.

Были разработаны:

- администрирующее desktop-приложение на Qt;
- Telegram-бота на C++ с асинхронной отправкой изображений и ответами от сервера.

Выводы: обеспечена удобная работа для конечных пользователей, реализована отправка и получение данных в реальном времени.

ЗАКЛЮЧЕНИЕ

В ходе данной работы был проведен обзор аналогов планируемого ПО, было проведено сравнительное тестирование алгоритмов выявления ключевых точек, на основе которого был выбран наиболее эффективный алгоритм, подходящий для задачи поиска музейных экспонатов по изображениям. Также была спроектирована архитектура серверного ПО, и было разработано клиент-серверное приложение для распознавания музейных экспонатов в условиях ограниченности данных.

В дальнейшем планируется разработка Android-клиента взамен текущему телеграм боту, добавление интерактивных карт музейных выставок в само клиентское приложение и их удобного редактора в администрирующее приложение. Также планируется переход с обработки отдельных изображений на обработку видеопотока, поступающего с камеры мобильного устройства.

Основные источники информации:

1. Weinmann, M.. From Early Concepts to Modern Computer Vision / M.. Weinmann, G.. Farinella, S.. Battiato, R.. Cipolla. Advanced Topics in Computer Vision. Advances in Computer Vision and Pattern Recognition. — London: Springer, 2013.
2. Howse, J.. Learning OpenCV 4 Computer Vision with Python 3 Third Edition / J.. Howse, J.. Minichico. — Birmingham, UK: Packt Publishing Ltd, 2020.
3. Кэлер, А.. Изучаем OpenCV 3. Разработка программ компьютерного зрения на C++ с применением библиотеки OpenCV / А.. Кэлер, Г.. Брэдски. — Москва: Лань, 2017.
4. Iliashenko, O. Yu.. Possibilities of using computer vision for data analytics in medicine / O. Yu.. Iliashenko, E. L.. Lukyanchenko // IZVESTIYA OF SARATOV UNIVERSITY. MATHEMATICS. MECHANICS. INFORMATICS. — 2022. — T. 22. — С. 224–232.

5. Горячкин, Б. С.. Компьютерное зрение / Б. С.. Горячкин, М.А.. Китов // E-Scio. — 2020.