

Министерство науки и высшего образования Российской Федерации

Федеральное государственное автономное образовательное учреждение
высшего образования «Уральский федеральный университет имени первого
Президента России Б. Н. Ельцина» (УрФУ)

Институт радиоэлектроники и информационных технологий – РТФ

ОТЧЕТ
о проектной работе
по теме: Искусственный интеллект
по дисциплине: Проектный практикум

Команда: Carleones

Тимлид: Бабаджанов Фуокатджон Фарходович РИ-140914

Разработчик: Гафурянов Артур Маратович РИ-14

Разработчик: Андреев Александр Сергеевич РИ-140914

Екатеринбург
2025

СОДЕРЖАНИЕ

1. Целевая аудитория
 2. Определение проблемы
 3. Подходы к решению проблемы
 4. Анализ аналогов
 5. Сценарии использования
 6. Требование к продукту и MVP
- Заключение

ВВЕДЕНИЕ

В современном мире технологии машинного обучения и искусственного интеллекта стремительно развиваются и находят всё большее применение в повседневной жизни. Одним из востребованных направлений в данной области является оптическое распознавание символов (OCR — Optical Character Recognition), в том числе и рукописного текста. Несмотря на значительный прогресс в распознавании печатных символов, обработка рукописных данных до сих пор остаётся сложной задачей из-за большого количества индивидуальных особенностей почерка, различий в стилях написания, наклоне и форме букв.

Целью нашего проекта стало создание удобного и доступного инструмента, способного распознавать рукописный текст с изображений, отправляемых пользователями. В качестве основной платформы взаимодействия с пользователями мы выбрали мессенджер Telegram, благодаря его популярности, простоте интеграции ботов и широким возможностям API. В результате был разработан Telegram-бот, способный принимать изображения с рукописным текстом, обрабатывать их с использованием алгоритмов компьютерного зрения и нейросетевых моделей, и возвращать пользователю результат в виде распознанного текста.

В рамках данного проекта была реализована полная цепочка обработки изображения: от получения фотографии через Telegram до выдачи пользователю текста. Особое внимание мы уделили выбору и обучению модели для распознавания рукописного текста, предобработке изображений, обеспечению устойчивости системы к "шумным" данным, а также разработке логики взаимодействия с пользователем. Наш бот способен обрабатывать как печатные, так и более сложные рукописные тексты с различной степенью

разборчивости, что расширяет сферу его применения — от автоматизации ввода данных до помощи в оцифровке архивов, заметок или анкет.

В ходе разработки мы столкнулись с рядом интересных технических и проектных задач, таких как определение оптимального формата входных изображений, обработка различных типов освещённости, уменьшение числа ошибок при распознавании, а также обеспечение стабильной работы Telegram-бота на серверной стороне. Кроме того, проект стал для нас отличной возможностью применить на практике знания в области машинного обучения, нейронных сетей, Python-программирования и API-интеграций.

Таким образом, наш проект не только решает практическую задачу распознавания рукописного текста, но и демонстрирует потенциал современных технологий в упрощении и автоматизации взаимодействия между человеком и компьютером. Мы уверены, что такие решения будут становиться всё более актуальными в ближайшем будущем, особенно в образовательной, юридической и архивной сферах.

1. ЦЕЛЕВАЯ АУДИТОРИЯ

Telegram-бот для распознавания рукописного текста ориентирован на широкий круг пользователей, которым необходимо быстро и удобно переводить рукописную информацию в цифровой формат. Наша целевая аудитория включает как частных пользователей, так и представителей профессиональных сообществ. Ниже представлены ключевые категории пользователей, на которых ориентирован наш проект:

1. Студенты и школьники

Студенты и учащиеся часто сталкиваются с необходимостью расшифровки и перевода в текстовый формат своих конспектов, домашних заданий, лабораторных работ и заметок. Использование Telegram-бота позволяет упростить этот процесс, особенно при подготовке рефератов и письменных работ.

2. Преподаватели и учителя

Для преподавателей бот может стать инструментом автоматизации проверки домашних заданий, анализа рукописных тестов и составления учебных материалов. Также он может быть полезен при переводе архивных заметок в электронный вид.

3. Работники архивов и библиотек

Сотрудники архивных служб, историки и библиотекари часто работают с большим объёмом рукописной информации. Автоматизация распознавания таких данных поможет ускорить процессы оцифровки и облегчить доступ к историческим материалам.

4. Специалисты в области документооборота

Юристы, нотариусы, бухгалтеры и другие специалисты, работающие с письменными заявлениями и анкетами, могут использовать бота для оперативного перевода рукописных данных в цифровую форму.

5. Широкая аудитория пользователей Telegram

Так как бот реализован в виде Telegram-бота, им могут воспользоваться любые пользователи мессенджера, которым нужно быстро распознать рукописный текст — например, чтобы отправить его по почте, вставить в документ или сохранить в заметках.



Рисунок 1 – Анализ целевой аудитории

2. ОПРЕДЕЛЕНИЕ ПРОБЛЕМЫ

В повседневной жизни и профессиональной деятельности люди по-прежнему широко используют рукописный текст: от конспектов и заметок до анкет, заявлений и архивных документов. Несмотря на активное развитие цифровых технологий, значительное количество информации по-прежнему создаётся вручную — особенно в образовательной, юридической, административной и научной сферах. Однако ручной ввод такой информации в цифровой формат остаётся трудоёмким, неэффективным и подверженным ошибкам.

Оцифровка рукописных данных традиционно требует значительных временных затрат, внимательности и аккуратности. При больших объёмах информации этот процесс становится утомительным и малопродуктивным. Кроме того, не каждый пользователь обладает навыками работы с профессиональными инструментами для оптического распознавания текста (OCR), а многие из этих решений имеют платный доступ или требуют установки дополнительного ПО. Всё это создаёт барьер для быстрой и доступной цифровизации рукописного контента.

Существующие решения по распознаванию текста, как правило, ориентированы на печатные документы и имеют ограниченную точность при работе с рукописными изображениями. Рукописный текст отличается высокой вариативностью — каждая надпись может быть уникальной в зависимости от почерка, толщины линий, угла наклона, фона, освещённости, качества съёмки и других факторов. Это делает задачу распознавания сложной и требует использования более гибких и адаптивных алгоритмов, в частности — нейросетевых моделей.

Также важно отметить, что большинство существующих OCR-систем работают через десктопные приложения или веб-интерфейсы, что требует от

пользователя определённого уровня технической подготовки. Между тем, Telegram — один из самых популярных мессенджеров, и именно его простота и универсальность делают его идеальной платформой для быстрого взаимодействия с инструментами на базе ИИ.

Таким образом, основная проблема, которую решает наш проект, заключается в отсутствии простого, доступного и мобильного способа распознавания рукописного текста с изображений. Наш Telegram-бот предоставляет пользователю возможность моментально отправить фотографию с рукописным текстом и получить результат без необходимости скачивать отдельные приложения или разбираться в сложных интерфейсах. Это делает процесс цифровизации заметок, анкет, форм и других рукописных документов простым, быстрым и эффективным.

3. ПОДХОДЫ К РЕШЕНИЮ ПРОБЛЕМЫ

Для эффективного решения задачи распознавания рукописного текста с изображений было рассмотрено несколько технических и архитектурных подходов, каждый из которых имеет свои особенности, преимущества и ограничения. В нашем проекте мы выбрали комбинацию современных технологий машинного обучения, компьютерного зрения и удобной платформы для взаимодействия с пользователем — Telegram. Ниже представлены ключевые подходы, использованные при разработке Telegram-бота.

1. Использование нейросетевых моделей для распознавания рукописного текста

Одним из наиболее перспективных подходов в современной практике является использование сверточных нейронных сетей (CNN) и рекуррентных нейронных сетей (RNN), а также трансформеров (Transformers) для обработки изображений и последовательностей. Для нашей задачи мы использовали одну из предобученных моделей, обученную на базе рукописных датасетов (например, EMNIST, IAM или собственный датасет).

Модель сначала извлекает признаки из изображения с помощью сверточных слоёв, а затем интерпретирует последовательность символов. В некоторых вариантах используется архитектура CRNN (Convolutional Recurrent Neural Network), которая хорошо справляется с задачей распознавания текста в строке. Также может применяться CTC-loss (Connectionist Temporal Classification), позволяющий обрабатывать входы с переменной длиной и без необходимости точной разметки.

2. Предобработка изображений

Качество входного изображения играет ключевую роль в точности распознавания. Поэтому нами реализованы методы предобработки изображений, включающие:

Приведение к стандартному размеру и цветовой гамме (чёрно-белое изображение).

Бинаризация (например, методом Оtsu) для повышения контрастности.

Фильтрация шума и сглаживание фона.

Коррекция перспективы и выравнивание текста при перекосах.

Такая обработка позволяет модели лучше фокусироваться на символах и уменьшает влияние внешних факторов, таких как тени, блики или неровности бумаги.

3. Интеграция с Telegram API

Telegram предоставляет простой и мощный API для создания ботов, включая возможность принимать изображения, текстовые команды и отправлять ответы в чате. Благодаря этому наш бот работает в интуитивно понятной среде, привычной для большинства пользователей.

Алгоритм работы бота:

1. Пользователь отправляет фотографию рукописного текста.
2. Бот получает изображение и запускает процедуру обработки.
3. Предобработанное изображение передаётся в модель для распознавания.

4. Результат в виде текста возвращается пользователю в том же чате.

Такой подход обеспечивает минимальное время отклика и высокую доступность сервиса.

4. Гибкая архитектура и масштабируемость

Бот построен с учётом возможности масштабирования и обновления. В случае необходимости можно легко заменить модель на более точную или дообученную, расширить поддержку языков или добавить функциональность (например, сохранение истории распознаваний, экспорт в PDF/Word и т.д.).

Эти подходы позволили нам создать надёжную, быструю и доступную систему распознавания рукописного текста, ориентированную на широкую аудиторию пользователей. В следующем разделе будет подробно рассмотрена техническая реализация проекта, включая выбор библиотек, фреймворков и инструменты для развертывания.

4. АНАЛИЗ АНАЛОГОВ

На сегодняшний день существует несколько популярных решений для распознавания текста с изображений, включая и рукописный текст. Однако большинство из них ориентированы на работу с печатными документами или требуют установки дополнительных программ. Ниже кратко рассмотрены некоторые из них.

1. Google Lens / Google Keep

Эти инструменты позволяют распознавать текст с фотографий прямо на смартфоне. Однако их функциональность ограничена рамками приложений, и пользователю необходимо вручную делать снимок, открывать нужное приложение и копировать результат. Поддержка рукописного текста нестабильна и зависит от качества почерка и изображения.

2. Microsoft OneNote OCR

OneNote предлагает встроенную функцию распознавания текста с изображений, включая рукописный. Однако эта функция больше ориентирована на работу в рамках экосистемы Microsoft и предназначена для пользователей, активно использующих офисные решения.

3. Онлайн-сервисы (например, i2OCR, NewOCR)

Это веб-платформы, которые позволяют загрузить изображение и получить текст. Большинство из них работают с печатным текстом и слабо адаптированы под рукописный. Также их использование не всегда удобно с мобильных устройств, а некоторые сервисы ограничены по количеству бесплатных распознаваний.

В отличие от вышеперечисленных решений, наш Telegram-бот:

не требует установки приложений;

работает напрямую в мессенджере;

ориентирован именно на рукописный текст;
обеспечивает быструю и интуитивно понятную обработку.

Таким образом, наш проект закрывает важную нишу — распознавание рукописного текста в мобильной и доступной форме.

5. СЦЕНАРИИ ИСПОЛЬЗОВАНИЯ

Созданный Telegram-бот ориентирован на широкий круг задач, связанных с переводом рукописной информации в цифровой формат. Ниже представлены основные сценарии использования, демонстрирующие его универсальность и практическую ценность.

1. Оцифровка учебных конспектов

Студенты и школьники могут использовать бота для перевода собственных или чужих конспектов в электронный формат. Это позволяет быстро собрать материалы для подготовки к экзаменам, отправить заметки одноклассникам или структурировать записи по темам.

2. Распознавание домашних заданий и упражнений

Бот может быть полезен для преподавателей, которым нужно быстро просмотреть или проверить выполненные задания, присланные в виде фотографий. Также учащиеся могут использовать его для хранения своих работ в электронном виде.

3. Оцифровка архивных документов

Историки, архивисты и библиотекари часто сталкиваются с необходимостью перевода старых рукописных записей в цифровую форму. С помощью Telegram-бота они могут быстро обрабатывать страницы рукописных книг, писем и заметок.

4. Автоматизация анкет и опросов

Организации, работающие с опросными листами или анкетами, заполненными вручную, могут использовать бота для быстрого извлечения текстовой информации и переноса её в базы данных или электронные таблицы.

5. Быстрая передача рукописных заметок

В повседневной жизни пользователи часто пишут заметки от руки — например, список покупок, номер телефона, идея проекта. Вместо перепечатывания вручную, достаточно отправить фото в бота и скопировать результат.

6. Поддержка людей с ограничениями по зрению или моторике

Бот может использоваться как вспомогательное средство для пользователей, которым сложно печатать текст с клавиатуры — например, при физических ограничениях. Они могут писать текст вручную, делать фото и использовать бота для преобразования в цифровую форму.

№	Сценарий использования	Описание	Оценочная доля %
1	Оцифровка учебных конспектов	Быстрый перевод записей с лекций и уроков в цифровой формат	30%
2	Распознавание для домашних заданий	Помощь преподавателям и ученикам в обработке рукописных работ	20%
3	Оцифровка архивных документов	Работа с историческими записями, письмами, дневниками и книгами	15%
4	Автоматизация анкет и вопросов	Быстрая оцифровка форм, опросов и Бланков	10%
5	Передача рукописных заметок	Упрощённый способ передать список, мысль или контакт написав от руки	15%
6	Поддержка пользователей с ограничениями	Альтернатива вводу с клавиатуры для людей с двигательными или зрительными трудностями	

6. ТРЕБОВАНИЯ К ПРОДУКТУ И К MVP

Общие требования к продукту

Чтобы Telegram-бот для распознавания рукописного текста был полезным, надёжным и удобным, мы выделили ряд требований, которым должна соответствовать финальная версия продукта:

Функциональные требования:

Возможность приёма изображений от пользователей (JPG, PNG).

Предобработка изображений (чёрно-белый фильтр, увеличение контраста, устранение фона).

Распознавание рукописного текста с использованием модели машинного обучения.

Возврат результата пользователю в виде текста в чате Telegram.

Поддержка кириллицы и латиницы.

Устойчивость к ошибкам: бот должен давать понятные сообщения при сбоях или нераспознанном тексте.

Нефункциональные требования:

Быстрое время отклика (желательно ≤ 5 секунд).

Простота использования (интерфейс — обычный Telegram-чат).

Масштабируемость — возможность подключать новые языки и модели.

Защита пользовательских данных (не хранить изображения дольше, чем нужно для распознавания).

Поддержка мобильных устройств (Telegram работает на всех популярных платформах).

Минимально жизнеспособный продукт (MVP)

На начальном этапе разработки мы сосредоточились на создании MVP-версии, которая реализует основные функции без второстепенного функционала. Это позволило протестировать ключевую идею проекта с минимальными затратами времени и ресурсов.

Возможности MVP:

Приём фото от пользователя в Telegram.

Автоматическая обработка изображения и его анализ.

Распознавание текста с помощью заранее обученной модели.

Ответ пользователю в том же чате: распознанный текст.

Ограничения MVP:

Только один язык (например, русский).

Отсутствие пользовательской настройки модели или формата вывода.

Минимальный интерфейс (без кнопок, только сообщения).

Нет сохранения истории и файлов.

Нет обучения на новых данных.

Цель MVP

Целью MVP было проверить:

Насколько точно работает выбранная модель на реальных фотографиях.

Удобно ли пользователям взаимодействовать с ботом в Telegram.

Сколько времени уходит на обработку и распознавание.

Какие трудности чаще всего возникают у пользователей.

По результатам тестирования MVP можно будет принять решение о том, какие функции нужно доработать, какие добавить, а какие — исключить. Также это поможет обосновать потребность в дальнейшем развитии проекта и привлечении дополнительных ресурсов (например, серверов или расширенной модели).

Заключение

В рамках данного проекта была разработана и протестирована рабочая версия Telegram-бота, способного распознавать рукописный текст с изображений с использованием современных методов компьютерного зрения и машинного обучения. Основное внимание было уделено простоте использования, доступности и скорости взаимодействия пользователя с системой.

В ходе работы над проектом мы:

провели анализ существующих решений на рынке и выделили их слабые стороны;

определили целевую аудиторию и сценарии использования бота;

разработали архитектуру системы и реализовали MVP-продукт;

провели предварительное тестирование и убедились в жизнеспособности идеи.

Telegram-бот показал, что даже простое и понятное решение может эффективно решать повседневные задачи — от оцифровки учебных конспектов до автоматизации ввода рукописных данных. Благодаря использованию готовых технологий и модульному подходу, проект имеет широкие перспективы масштабирования и улучшения.

Созданный MVP может стать основой для более сложных и интеллектуальных систем с возможностью обучения на пользовательских данных, интеграции с внешними сервисами (Google Docs, PDF-экспорт, базы данных) и поддержки распознавания нескольких языков.

Разработка такого бота продемонстрировала важность связи между техническими знаниями и реальными потребностями пользователей. Мы считаем, что проект имеет потенциал для практического применения как в образовательной, так и в профессиональной среде, а дальнейшее развитие — вопрос времени и ресурсов.

