

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ
Н. Г. ЧЕРНЫШЕВСКОГО»**
Кафедра информатики и программирования

**РАЗРАБОТКА МОБИЛЬНОГО ПРИЛОЖЕНИЯ
ДЛЯ ДЕТЕКТИРОВАНИЯ И РАСПОЗНАВАНИЯ НОМЕРА
МАРШРУТА ПРИБЛИЖАЮЩЕГОСЯ ОБЩЕСТВЕННОГО
ТРАНСПОРТА**
АВТОРЕФЕРАТ БАКАЛАВРСКАЯ РАБОТА

студента 4 курса, 441 группы
направления 02.03.03 – Математическое обеспечение и администрирование
информационных систем
факультета компьютерных наук и информационных технологий
Орлова Даниила Андреевича

Научный руководитель

ст. преп. кафедры ИиП

А.А Казачкова

Заведующий кафедрой

к.ф.-м.н., доцент

М. В. Огнева

Саратов 2023

ВВЕДЕНИЕ

Актуальность темы.

В настоящее время, в нашу жизнь активно внедряется множество новых технологий, как для улучшения уровня жизни, так и для работы. Но некоторые люди могут быть ограничены в своих возможностях, подвергаясь разным факторам. Так технологии созданы и для того, чтобы компенсировать ограничения людей.

В данной работе будет рассмотрено применение технологий машинного обучения и обработки данных для создания приложения, которое поможет людям со слабым зрением пользоваться общественным транспортом – распознавать номера приближающегося общественного транспорта. В интернете существует множество схожих работ [1-5], которые созданы для помощи людям с ограниченными возможностями, но ни одна из них не реализует конкретную задачу детектирования номера маршрута общественного транспорта.

Целью данной дипломной работы является создание мобильного приложения для распознавания номера маршрута приближающегося общественного транспорта.

Задачи:

- обзор средств и методов машинного обучения;
- обзор особенностей нейронных сетей и их архитектуры;
- обзор средств разработки мобильных приложений;
- сбор и разметка данных;
- разработка нейронной сети для сегментации нужной части изображения;
- определение номера маршрута приближающегося общественного транспорта;
- разработка мобильного приложения для работы с нейронной сетью;
- оценка результата.

Методологические основы.

Проектирование и разработка мобильных приложений, а также разработка проектов с использованием машинного обучения представлена в работах: Бурков А., Чернигов В., Колымыцева Е. П., Крутько Д. А. Володина Л. А., Мартышкин А. И., Николенко С.

Практическая значимость работы.

Практическая значимость работы заключается в использовании приложения людьми с ограниченными возможностями, а также для повышения их уровня жизни.

Структура и объём работы.

Бакалаврская работа состоит из введения, 7 разделов, заключения, списка источников и 5 приложений. Общий объём работы – 59 страниц, из них 43 страниц – основное содержание, включая 16 рисунков, цифровой носитель в качестве приложения, список источников информации – 26 наименований.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первый раздел «Обзор аналогов» посвящен рассмотрению существующих аналогов, которые предназначены для помощи и улучшения жизни слабовидящих людей. Однако данные приложения не имеют тех же специализаций, что и рассматриваемая работа.

TapTapSee - приложение, которое позволяет идентифицировать объекты, фотографируя их и получая звуковое описание.

NaviLens - приложение, разработанное для общественного транспорта в Барселоне.

Google Lens - приложение, использующее искусственный интеллект и компьютерное зрение для распознавания объектов и предоставления информации о них, включая текст, изображения и штрих-коды.

PlantNet - приложение, помогающее идентифицировать различные виды растений и предоставляющее дополнительные сведения о них.

Также рассматриваются подходы компьютерного зрения, включая выделение контуров и готовые OCR модели.

Все перечисленные аналоги могут облегчить жизнь слабовидящим людям, но не обладают необходимым функционалом, который планируется разработать в данной работе.

Второй раздел «Основные определения» отражает в себе основные понятия связанные с работой, такие как:

1. Машинное обучение – подраздел искусственного интеллекта (ИИ) и науки о данных, специализирующийся на использовании данных и алгоритмов для имитации процесса наработки опыта человеком с постепенным повышением точности.

2. Машинное обучение с учителем решает задачу, где имеется множество объектов и возможных ответов. Зависимость между объектами и ответами неизвестна, но известна обучающая выборка, состоящая из пар "объект, ответ". Целью является создание алгоритма, который способен предоставлять точные ответы для любого объекта. Функционал качества или функционал

ошибки используется для измерения точности ответов, а учитель может быть либо сама обучающая выборка, либо человек, предоставляющий правильные ответы для заданных объектов.

3. Типы данных, которые используются в задачах машинного обучения с учителем, это:

- матрица входных объектов – каждый объект описывается набором своих характеристик – признаков;
- матрица расстояний между объектами – каждый объект описывается расстояниями до всех остальных объектов обучающей выборки;
- изображения или видеоряд;
- более сложные случаи, специфичные под каждую задачу, например графы, тексты и результат запросов к базам данных.

Третий раздел «Нейронная сеть». В данном разделе представлен обзор структуры и функционирования нейронных сетей, а также введение в глубокое обучение и его значимость. В конце раздела представлен обзор сети архитектуры YoloV7, которая известна своей способностью обнаруживать объекты в реальном времени.

Четвертый раздел «Мобильное приложение на платформе .NET» включает в себя рассмотрение создания мобильного приложения на платформе .NET с использованием Xamarin. Описывается архитектура Xamarin, включая взаимодействие общего кода и кода базовой платформы, а также компиляцию для Android и Xamarin.iOS. Рассматривается роль языка XAML в разработке пользовательского интерфейса и его отделение от логики времени выполнения. Также рассматривается платформа Xamarin.Forms, позволяющая создавать пользовательские интерфейсы с использованием XAML и кода программной части на C#, а также возможность использования платформенных функций без создания пользовательских «отрисовщиков» или эффектов.

Пятый раздел «Используемые технологии» описывает инструменты и технологии, которые использовались в данной работе, это:

- Jupyter Notebook

- Платформа Roboflow
- Платформа Wandb.ai
- Веб-фреймворк Flask
- Библиотека Tensorflow
- Библиотека OpenCV
- OCR модели
 - Библиотека Python-tesseract
 - Библиотека EasyOCR

Шестой раздел «Реализация веб-интерфейса для работы с нейронной сетью» описывает практическую часть данной работы. Здесь представлено описание создания датасета включающего 688 фотографий различных транспортных средств. Изображения были собраны из открытых источников, таких как поисковая система Google и специализированный сайт Urban Electric Transit. Фотографии были преобразованы в формат JPG с разрешением 640x640 пикселей и глубиной цвета 24. Затем данные были размечены на платформе Roboflow, где для каждого изображения была определена нужная область с помощью Bounding box, которая была преобразована в текстовые данные. Датасет и его аннотации хранятся на платформе в виде версий, доступных для обучения моделей. Для загрузки датасета использовалась библиотека Roboflow. Используемый датасет был разделён на тренировочную, валидационную и тестовую выборки. Для запуска обучения модели была использована команда с указанием конфигурации и параметров, включая количество эпох обучения, файл конфигурации и веса модели. Обучение проводилось на графическом процессоре Tesla K80 с 13 Гб памяти. Было обучено три модели с разным количеством эпох: tiny-модель, вторая модель с 100 эпохами и третья модель с 130 эпохами для проверки переобучения. Tiny-модель имела наименьшую точность, но использовала наименьшее количество ресурсов системы, а также обучилась значительно быстрее остальных. Метрики Tiny-модели представлены на рисунке 1. Её итоговая точность составила – 0.48 из 1. Что является ожидаемым результатом её работы, ввиду легковесности. Обучение

двух других моделей дало лучшие результаты. Из рисунка 2 видно, что модели до 80 эпохи идут вровень, но влияние количества эпох очевидно. Исходя из рисунка 2 видно, что у «большой» модели чаще случаются ложные срабатывания. Это



Рисунок 1 — Метрики Tiny-модели

означает, что большая модель способна выдавать неправильные координаты целевого участка изображения, среди которых возможен правильный. Дальнейший рост числа эпох привёл бы к переобучению. Tiny-модели

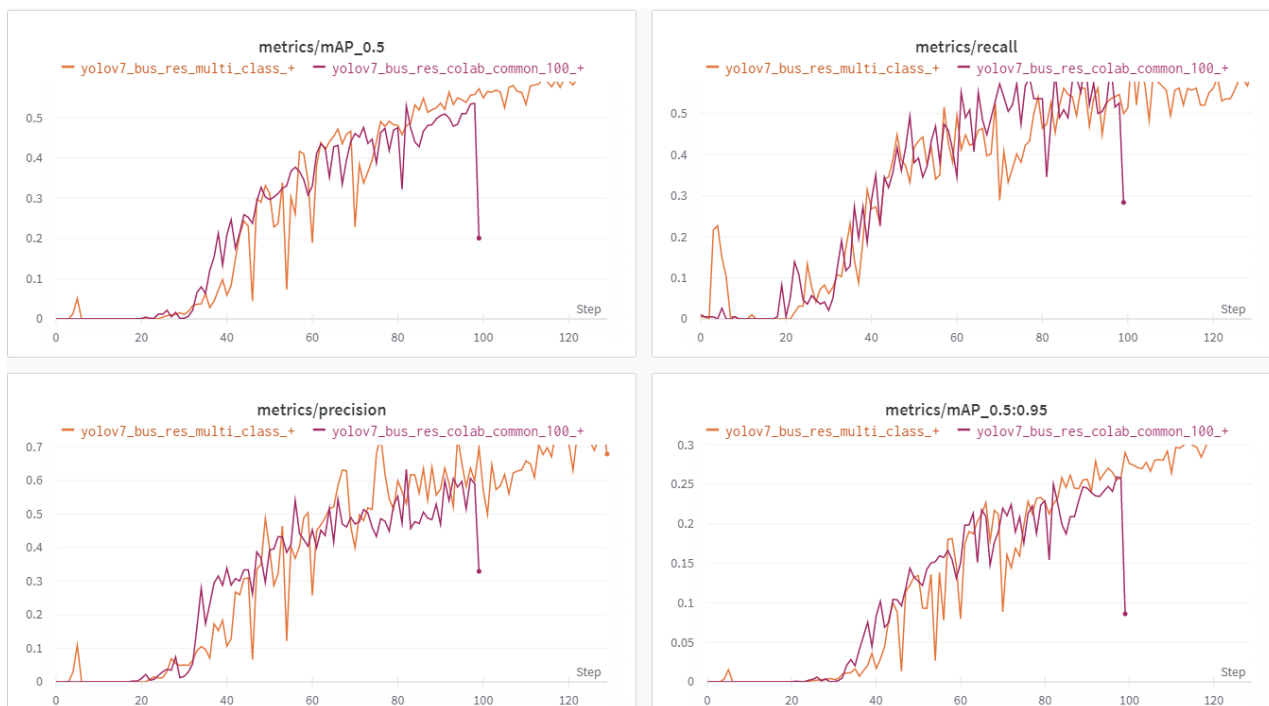


Рисунок 2 — Результат обучения 100 и 130 эпох

Также здесь был произведён анализ проведённых экспериментов, представленных в таблице 1. В частности, планировалось использовать модели EasyOcr и py-tesseract для распознавания символов на изображении. Они предварительно обучены и хорошо зарекомендовали себя среди разработчиков.

Таблица 1 – Сравнение экспериментов по детектированию изображения

	easyOcr	py-tesseract	Совокупность
Без предобработки \ С увеличенным размером рамки	0.75	0.13	0.68
Без предобработки \ С базовым размером рамки	0.53	0.1	0.38
С предобработкой \ С увеличенным размером рамки	0.46	0.005	0.43
С предобработкой \ С базовым размером рамки	0.47	0.15	0.37

Было решено использовать EasyOcr, так как по результатам проведенных, в ходе данной работы, экспериментов она дает наилучшую точность.

Полученная точность для тестового образца с учетом сегментированного изображения составляет 0,757. Это отличный результат работы программы. Совокупность моделей менее удовлетворительна, а также использует больше системных ресурсов и занимает больше времени на выполнение. Py-tesseract работает довольно плохо, полученная точность близка к 0, эта модель не подходит для данного проекта. В Таблице 1 представлена точность каждого метода при решении задачи классификации.

Седьмой раздел «Реализация мобильного приложения» включает в себя описание разработки мобильного приложения для получения изображения маршрутного транспортного средства либо с камеры телефона, либо из памяти устройства, после чего следует отправка изображения на сервер, с последующей его обработкой и отправкой результата обратно на устройство. Результат работы всего приложения представлен на рисунке 3.

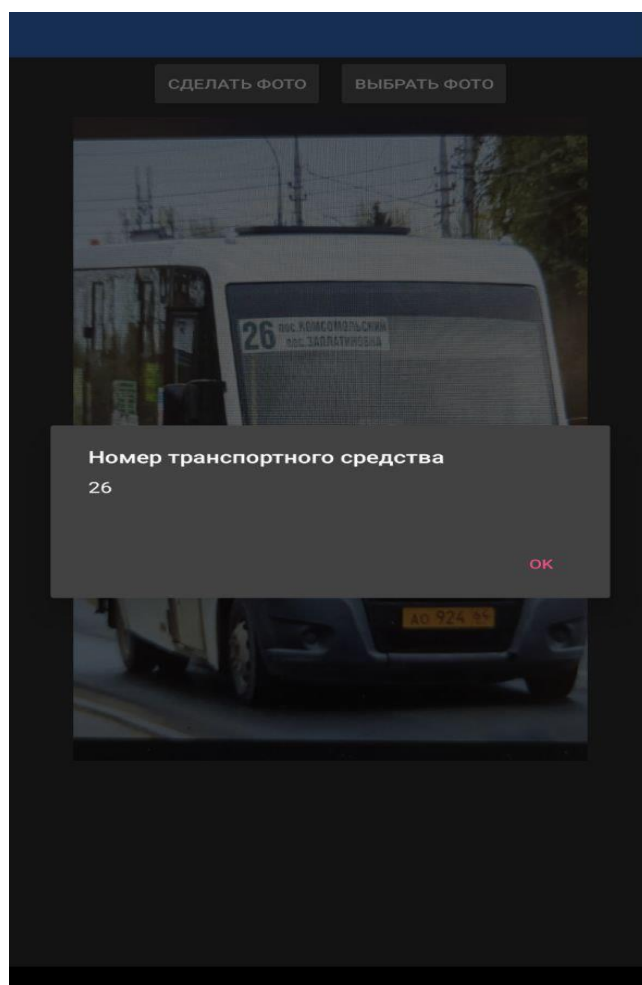


Рисунок 3 – Результат работы мобильного приложения

ЗАКЛЮЧЕНИЕ

В результате выполнения дипломной работы было разработано мобильное приложение для детектирования номера приближающегося общественного транспорта. Были проведены исследования и анализ существующих решений в данной области, определены требования к приложению, разработана архитектура и реализованы методы машинного обучения для поставленной задачи.

Были выполнены следующие задачи:

- обзор средств и методов машинного обучения в разделе 3;
- обзор особенностей нейронных сетей и их архитектуры в разделах 3.1 – 3.2;
- обзор средств разработки мобильных приложений в разделе 4;
- сбор и разметка данных в разделе 6.1;
- разработка нейронной сети для сегментации нужной части изображения в разделе 6.2;
- определение номера маршрута приближающегося общественного транспорта в разделе 6.3;
- разработка мобильного приложения для работы с нейронной сетью в разделе 7;
- оценка результата.

Приложение получает изображение либо с камеры телефона, либо загружает его из памяти устройства, после чего отправляется на сервер, где происходит его обработка: сначала при помощи обученной модели для сегментации нужного участка фотографии, а после языковой моделью. Полученный результат отправляется обратно на устройство, где происходит вывод результат, либо же вывод сообщения об ошибке.

Приложение работает стабильно, потребляет низкое количество ресурсов системы, а также является расширяемым, поскольку основная вычислительная часть находится на удалённом сервере.

Основные источники информации.

1. Крутько Д. А. / Применение сверточных нейронных сетей в задаче распознавания цифр / Д. А. Крутько // Журнал Молодежь. Общество. Современная наука, техника и инновации – 2022 г. № 21. – С 51-52. Яз. рус.
2. Колымыцева Е. П. / Машинное обучение для классификации изображений / Е. П. Колымыцева // Международная научно-техническая конференция молодых ученых БГТУ им. В.Г Шухова / Статья в сборнике трудов конференции – 2022г. 103-108 с.
3. Володина Л. А. / Применение нейросетевых технологий в области распознавания образов / Л. А Володина // Сборник статей IV Международного научно-исследовательского конкурса – 2022 г. 7-9 с.
4. Мартышкин А. И. / Особенности работы сверточных нейронных сетей: архитектура и применение / А. И. Мартышкин, А. А. Зоткина / Журнал Современные информационные технологии – 2022 г. № 36. – С. 11-13.
5. Людоговская М. А / Применение сверточных нейронных сетей при обработке видеоинформации в транспортных системах / М. А Людоговская, М. А. Кабанов // Конференция: Системы управления, сложные системы: моделирование, устойчивость, стабилизация, интеллектуальные технологии / Материалы VIII Международной научно-практической конференции – 2022 г. 96-101 с.
6. Чернигов В. Разработка мобильных приложений на C# для iOS и Android. [Книга] Чернигов В. / ДМК издательство – 2020. – 188с.
7. Распознавание текста с помощью OCR [Электронный ресурс]: интернет-статья. – URL: <https://habr.com/ru/articles/471542/> (дата обращения 21.03.2023)
8. Бурков А. Машинное обучение без лишних слов / Бурков А. 2020. – 192 с.