

Разработка приложения-собеседника с возможностью идентификации пользователей

Самохвалова Полина Сергеевна ИУ9-82Б
Руководитель: Посевин Данила Павлович

Актуальность

В настоящее время приложения-собеседники применяются во многих областях. Они используются при решении следующих задач:

- 1) при разработке бесконтактных диалоговых систем и интерфейсов
- 2) в роботизированных системах для взаимодействия с пользователями
- 3) в исследовательских целях для разработки систем искусственного интеллекта
- 4) для сбора датасетов для диалоговых систем и приложений-собеседников
- 5) в образовательных целях

Цель работы

Целью выпускной квалификационной работы является реализация приложения-собеседника с возможностью идентификации пользователей при помощи распознавания лиц.

Постановка задачи

Для достижения поставленной цели необходимо решить следующие задачи:

- рассмотреть способ идентификации пользователей на основе распознавания лиц
- рассмотреть различные виды нейронных сетей, используемых для создания приложений-собеседников
- реализовать пользовательский интерфейс приложения – отображение видеопотока с веб-камеры, интерфейс для распознавания лиц и ввода имен, а также интерфейс для общения с ботом
- реализовать базу данных для хранения информации, необходимой для работы приложения
- реализовать распознавание лиц и возможность ввода имен
- реализовать нейронную сеть для работы чат-бота
- реализовать возможность общения с ботом в пользовательском интерфейсе
- провести тестирование приложения

Выбор стека технологий

Для написания программы был выбран язык программирования Python и его библиотеки и модули:

- PyTorch
- OpenCV
- face_recognition
- speech_recognition
- psycopg2
- os
- time
- tkinter
- PIL
- shutil
- pickle
- sys
- re

Нейронные сети, используемые для создания приложений-собеседников

1. Рекуррентные нейронные сети (RNN) представляют собой класс нейронных сетей, разработанных для обработки и анализа последовательностей данных.
2. LSTM (Long Short-Term Memory) является одним из типов RNN. LSTM используется для решения проблемы затухания градиентов, позволяет сохранять информацию на протяжении длительных временных интервалов.
3. Модель GRU (Gated Recurrent Unit) основана на тех же принципах, что и LSTM, но использует меньше фильтров и операций для вычисления нового скрытого состояния.
4. Механизм внимания позволяет нейронной сети фокусироваться на определенных частях входных данных, что приводит к улучшению качества ответов в приложениях-собеседниках.

Нейронные сети, используемые для создания приложений-собеседников

5. Transformer – архитектура глубоких нейронных сетей, основанная на механизме внимания. Основное преимущество трансформеров по сравнению с рекуррентными нейронными сетями заключается в их высокой эффективности в условиях параллелизма.
6. Модель sequence to sequence применяется для преобразования последовательности входных данных в последовательность выходных данных.
7. BERT построена на архитектуре трансформер, но с увеличенным числом и размером слоев, отсутствующей декодирующей частью и глубокой двунаправленностью, то есть рассмотрением контекста с двух сторон.
8. GPT – это еще одна модель Transformer, предобученная на больших корпусах текста.

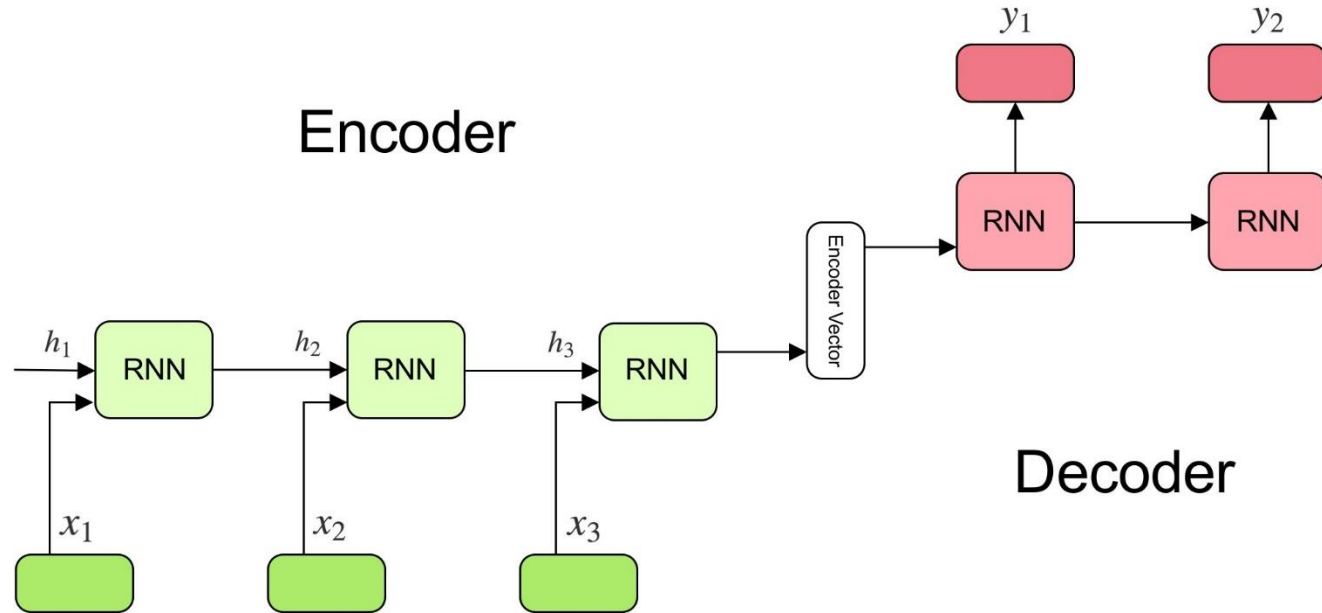
Выбор модели нейронной сети для разработки приложения

Для реализации приложения-собеседника была выбрана модель нейронной сети sequence to sequence с энкодером RNN GRU, декодером RNN GRU и механизмом внимания. Рассмотрим преимущества выбранной модели.

1. Гибкость последовательной генерации.
2. Улучшенная способность перевода входной последовательности в выходную, благодаря использованию механизма внимания.
3. Использование блоков GRU для эффективной обработки последовательностей.
4. Возможность обучения на парных последовательностях.

Модель sequence to sequence

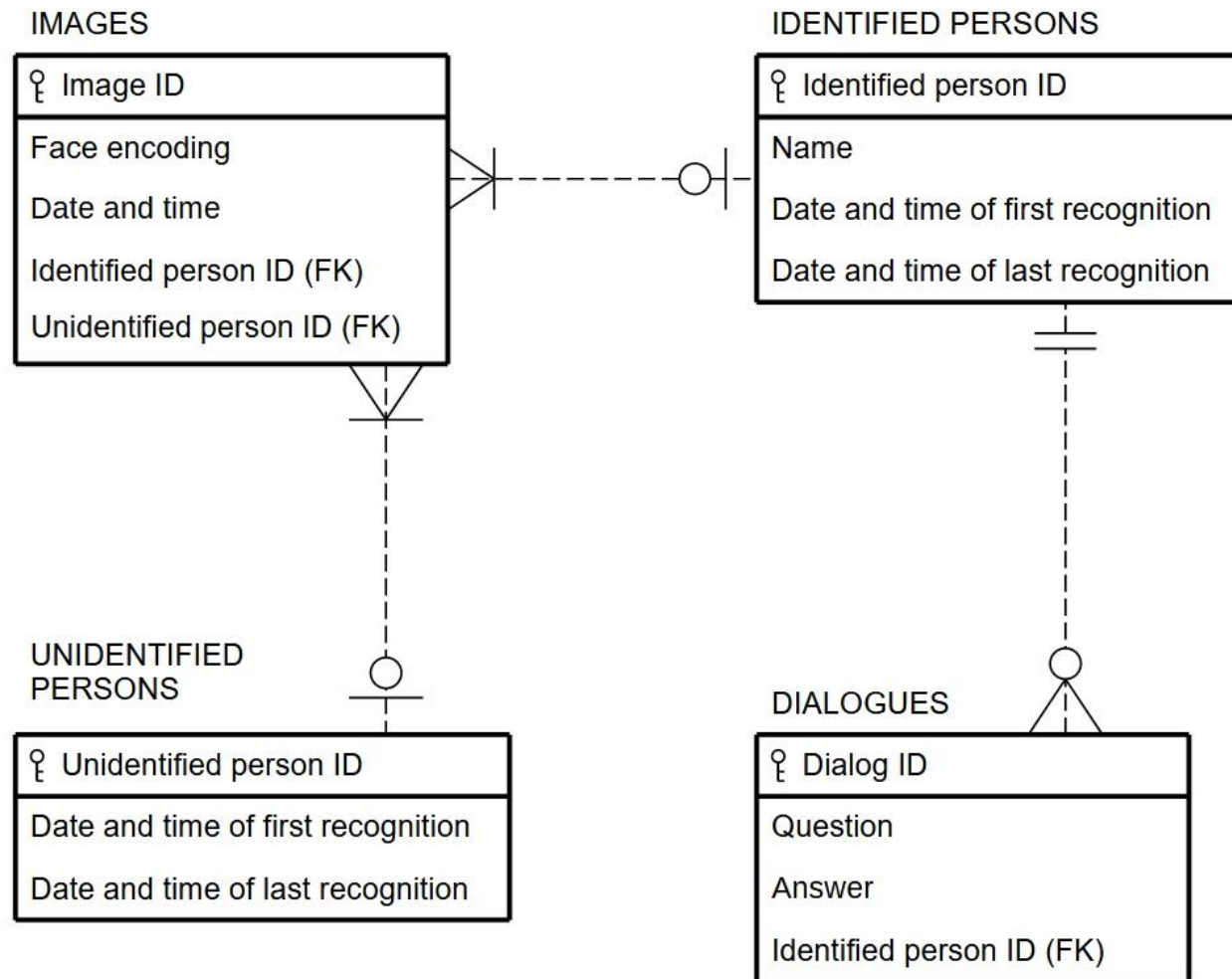
Модель sequence to sequence применяется для преобразования последовательности входных данных в последовательность выходных данных.



Реализация энкодера и декодера, использование механизма внимания

- Энкодер RNN выполняет итерацию входного предложения по одному токену за раз, на каждом временном шаге получая вектор «вывода» и вектор «скрытого состояния».
- Декодер RNN использует контекстные векторы энкодера и внутренние скрытые состояния для генерации следующего слова в последовательности.
- Механизм внимания позволяет декодеру обращать внимание на определенные части входной последовательности, а не использовать весь фиксированный контекст на каждом шаге. Внимание вычисляется с использованием текущего скрытого состояния декодера и выходных данных энкодера.

Создание базы данных для хранения информации, необходимой для работы приложения, реляционная модель



Этапы разработки приложения

1) Реализация пользовательского интерфейса приложения.

1. Реализация интерфейса распознавания лиц и ввода имен.
2. Реализация интерфейса общения с ботом.

2) Реализация базы данных.

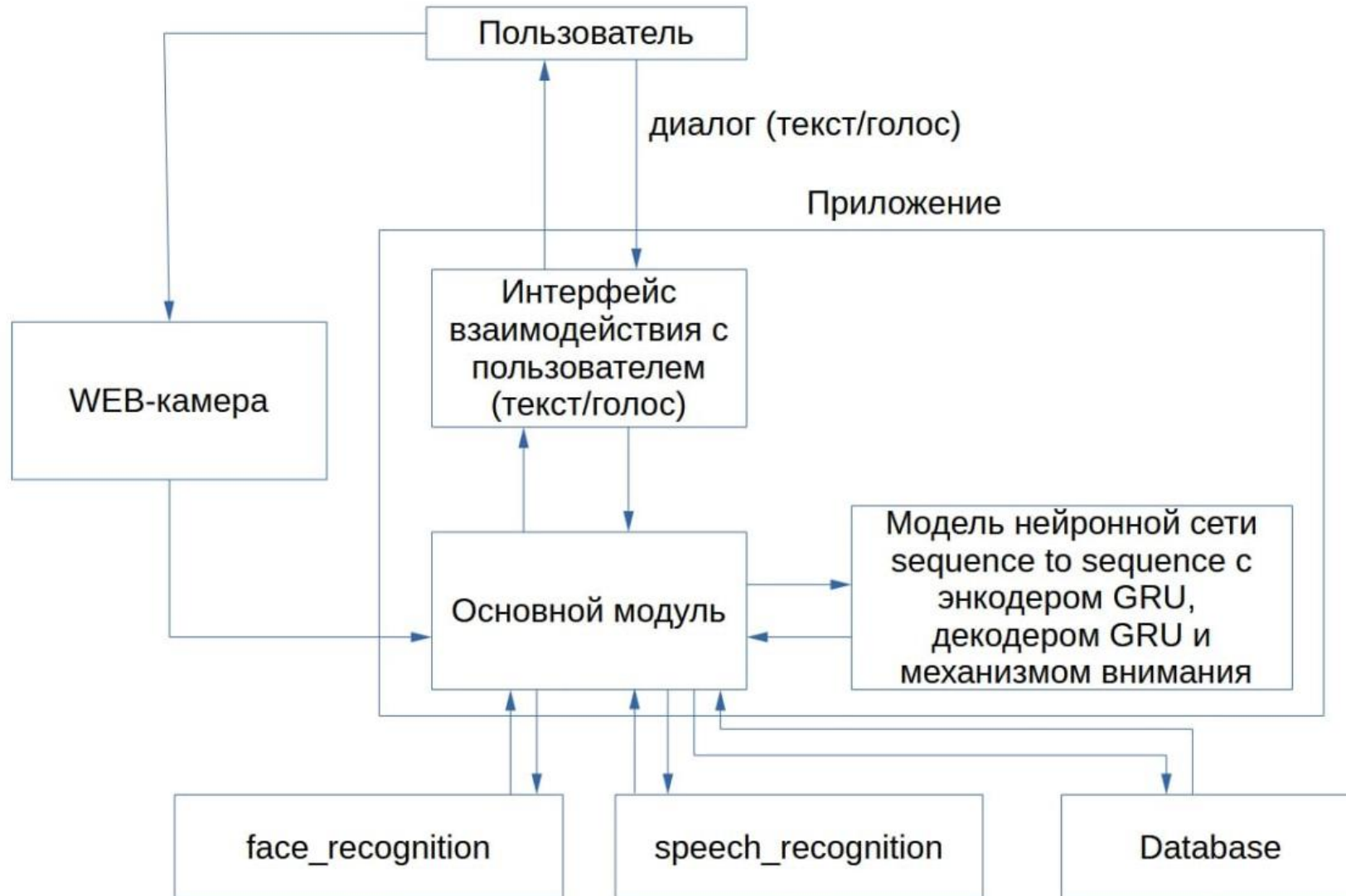
1. Реализация модели «сущность-связь», описание сущностей.
2. Преобразование модели «сущность-связь» в реляционную модель.
3. Описание таблиц.

3) Реализация распознавания лиц.

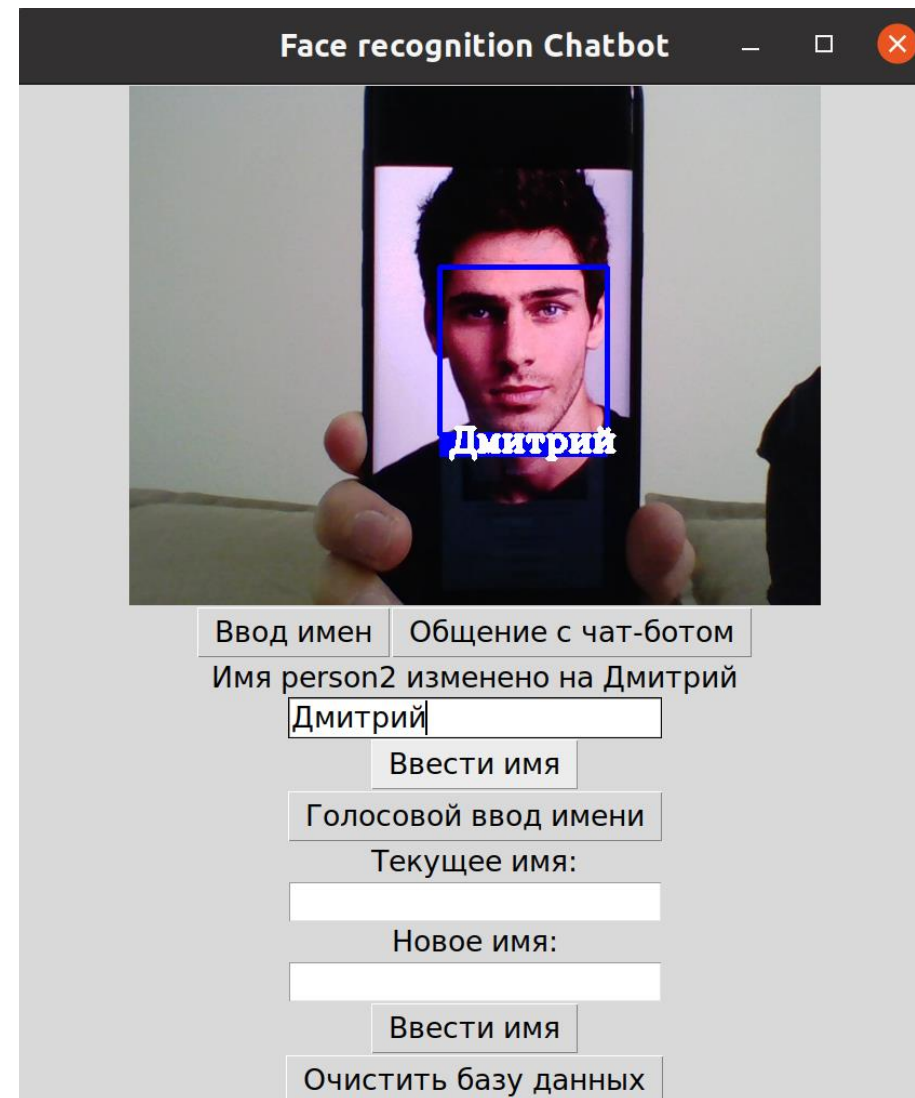
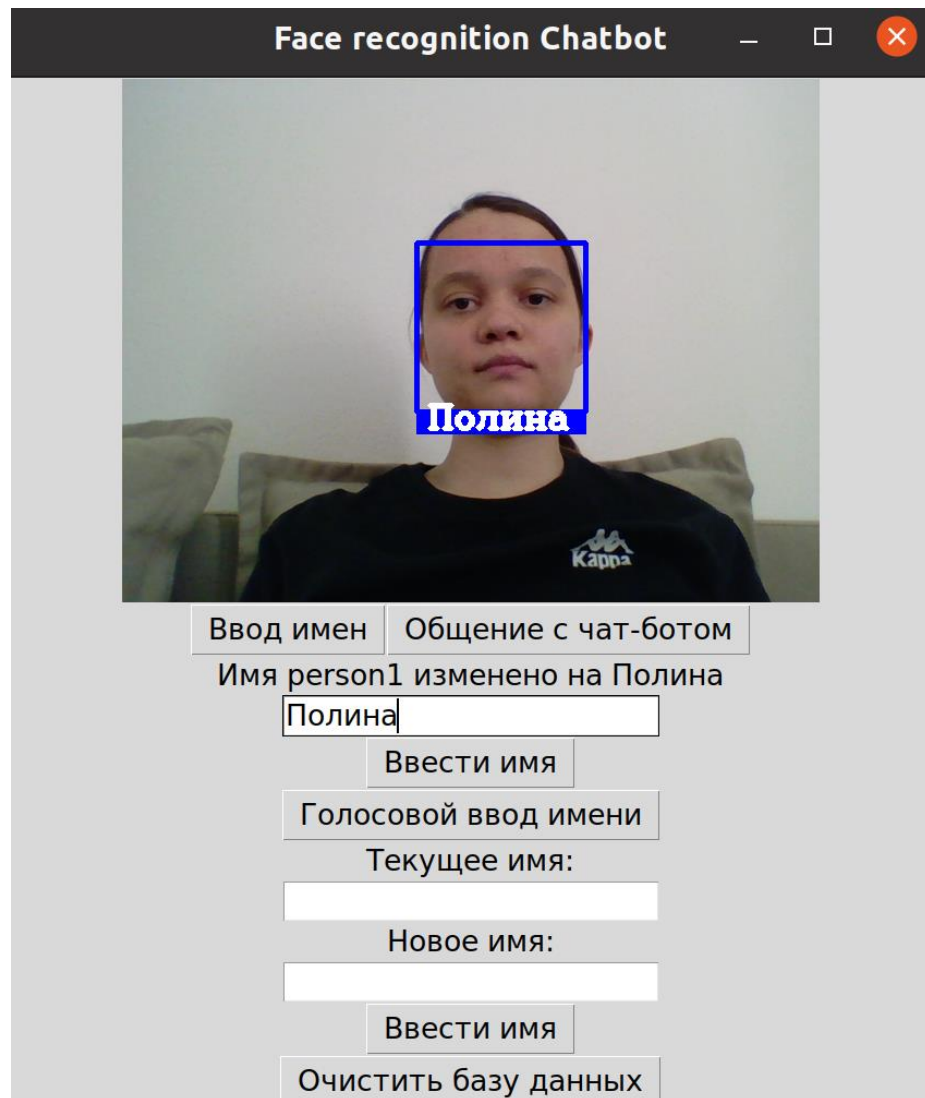
4) Реализация чат-бота.

1. Выбор данных для обучения.
2. Предварительная обработка данных.
3. Реализация модели нейронной сети sequence to sequence с энкодером RNN GRU, декодером RNN GRU и механизмом внимания.

Схема работы приложения

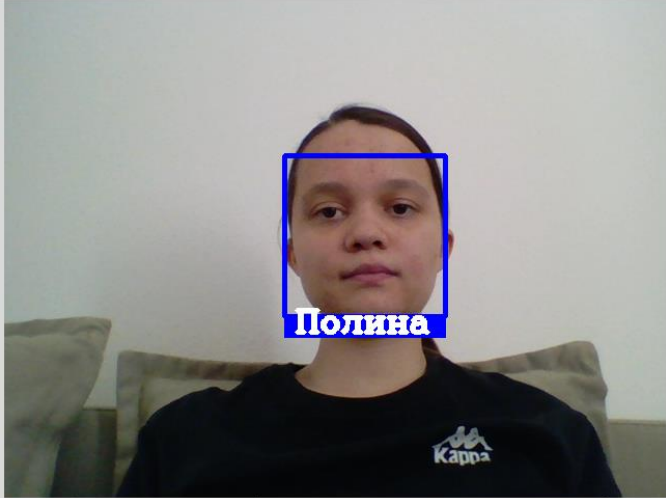


Распознавание лиц и ввод имен



Общение с ботом

Face recognition Chatbot



Ввод имен

Общение с чат-ботом

Полина: Привет!

Бот: привет !

Полина: Как дела?

Бот: нормально .

Полина: Что делаешь?

Бот: смотрю .

Полина: Я иду гулять

Бот: я тоже

Полина: Пойдешь со мной гулять?

Бот: конечно .

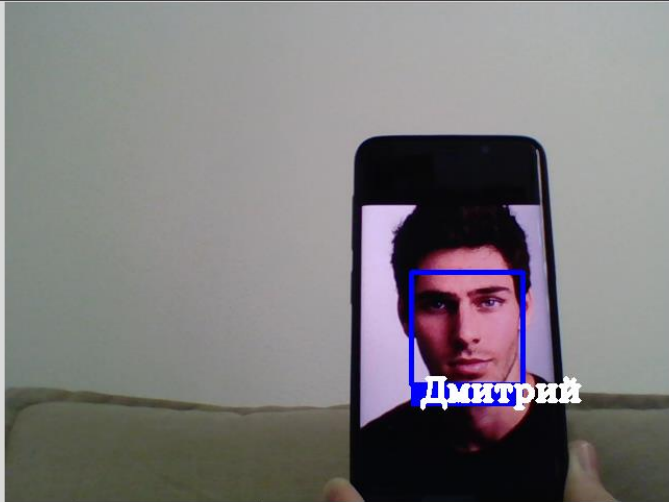
Ввести сообщение

Ввести сообщение голосом

Очистить диалог

Общение с ботом

Face recognition Chatbot



Дмитрий

Ввод имен

Общение с чат-ботом

Дмитрий: Здравствуй!

Бот: здравствуй !

Дмитрий: Ты где?

Бот: в швейцарии .

Дмитрий: Как прошел твой день?

Бот: прекрасно .

Дмитрий: Я пойду гулять в парк

Бот: я тоже

Дмитрий: Когда ты пойдешь гулять в парк?

Бот: завтра .

Ввести сообщение

Ввести сообщение голосом

Очистить диалог

Заключение

В результате выполнения выпускной квалификационной работы были выполнены следующие задачи:

- рассмотрен способ идентификации пользователей на основе распознавания лиц
- рассмотрены различные виды нейронных сетей, используемых для создания приложений-собеседников
- реализован пользовательский интерфейс приложения
- реализована база данных для хранения информации, необходимой для работы приложения
- реализовано распознавание лиц и возможность ввода имен
- реализована нейронная сеть для работы чат-бота – модель sequence to sequence с энкодером RNN GRU, декодером RNN GRU и механизмом внимания
- реализована возможность общения с ботом в пользовательском интерфейсе
- проведено тестирование приложения