



Ссылка на статью:

// Ученые записки УлГУ. Сер. Математика и информационные технологии. УлГУ. Электрон. журн. 2020, № 2, с. 30-34.

Поступила: 27.05.2020

Окончательный вариант: 12.07.2020

© УлГУ

УДК 004.94

## Мобильное приложение для перевода с языка жестов в режиме реального времени

Копылов А.В. \*, Волков М.А.

\* [alexkopylov123@gmail.com](mailto:alexkopylov123@gmail.com)

УлГУ, Ульяновск, Россия

---

В работе дано описание программного продукта, позволяющего осуществлять перевод тактильной речи американского жестового языка на буквы английского языка. Многопоточное мобильное приложение на базе ОС Android для распознавания жестов использует натренированную сверточную нейронную сеть с архитектурой ResNet18 с подобранными гиперпараметрами модели сети и выбранными параметрами аугментации. Одним из важных преимуществ данного продукта является возможность работы в режиме реального времени на мобильных устройствах, без использования сети Интернет. Разработанный прототип мобильного приложения является универсальным, что позволяет при смене языка заменить только файл с моделью нейронной сети (при условии схожести их архитектур) и файл с буквами - классами.

**Ключевые слова:** мобильное приложение, язык жестов, тактильная речь, сверточные нейронные сети, Android, Python, Kotlin.

---

### Введение

В настоящее время активно растет применение нейронных сетей в различных областях нашей жизни, в том числе в распознавании образов. Распознавание языка жестов глухих людей как раз является такой задачей. Для ее решения необходимо выбрать подходящую нейронную сеть, собрать и обработать для нее данные, затем ее натренировать, разработать Android-приложение и загрузить в него нейронную сеть.

Общение - это бесконечно важный навык, который позволяет людям обмениваться информацией. Мы общаемся с помощью речи, жестов, языка тела, письма или с помощью визуальных средств, причем речь является одним из наиболее часто используемых способов общения.

По данным ВОЗ, более 5% населения мира – 466 миллионов человек, в том числе 34 миллиона детей – страдают от инвалидизирующей потери слуха: потеря слуха в слыша-

щем лучше ухе, превышающая 40 дБ у взрослых людей и 30 дБ у детей [1]. В России более 13 миллионов человек имеют проблемы со слухом.

Люди с нарушением слуха общаются с помощью жестовых языков. Система жестового общения глухих имеет сложную структуру и включает две разновидности жестовой речи: русскую и калькирующую. Жестовая речь – это общение при помощи средств жестового языка – самобытной лингвистической системы, обладающей своеобразной лексикой, грамматикой и т.д.

Дактильная речь калькирует лингвистическую структуру словесного языка. Дактильная речь – вторичная знаковая система. Жесты здесь являются эквивалентами слов, а порядок их следования такой же, как в обычном предложении [2].

Если многие глухие могут общаться между собой или читать по губам, то для человека из мира слышащих понять речь глухого бывает очень трудно ввиду объективных причин. Более того, многие люди имеют, помимо глухоты, еще и немоту.

Для решения этой проблемы разработано Android - приложение, которое использует сверточную нейронную сеть, для перевода дактильной речи.

## 1. Функционал программного продукта

В качестве основы для создания программного продукта выбрано мобильное приложение на базе ОС Android, не использующее сеть Интернет. Основными преимуществами такого подхода являются:

- мобильность – использование данного приложения из любой точки мира при условии наличия устройства на базе ОС Android;
- офлайн-режим – не потребуются подключение к Интернету, все необходимые вычисления происходят на мобильном устройстве.

Разработанное многопоточное мобильное приложение для перевода с языка жестов отслеживает движение только кисти руки (см. рис. 1).

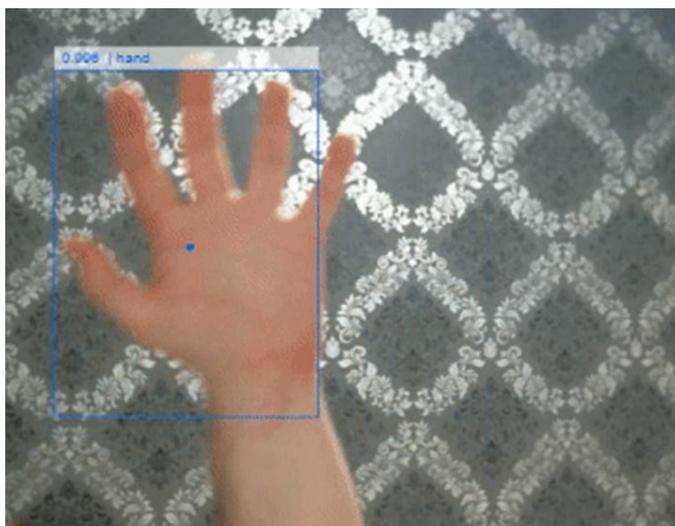
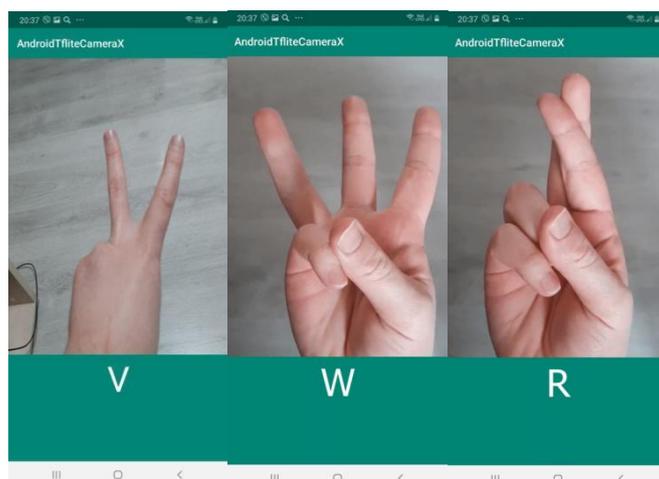


Рис. 1. Захват приложением движений кисти руки.

Приложение работает следующим образом: при наведении камеры смартфона на руку говорящего внизу приложения отображается перевод жеста в букву (см. рис. 2).



**Рис. 2.** Распознавание в приложении букв по жестам.

## 2. Используемые технологии

Сбор и обработка данных, а также тренировка нейронной сети выполнены в среде Jupiter Notebook на языке Python. Машины для сбора, хранения и обработки данных предоставлены сервисом Google Cloud в рамках пробной версии для академических целей. Для тренировки нейронной сети загружены с сайта Kaggle.com следующие данные:

1. [grassknotted/asl-alphabet](#)
2. [kuzivakwashe/significant-asl-sign-language-alphabet-dataset](#)
3. [ahmedkhanak1995/sign-language-gesture-images-dataset](#)
4. [rajarshighoshal/asltestimages](#)
5. [danrasband/asl-alphabet-test](#)
6. [augustohaja/asl-lucas](#)

Для расширения набора данных использованы следующие параметры аугментации:

- `do_flip=True`
- `max_rotate=10.0`
- `max_zoom=1.1`
- `max_lighting=0.15`
- `max_warp=0.2`
- `p_lighting=0.75`

Архитектурой для нейронной сети стала ResNet18[3], представленная на рис. 3.

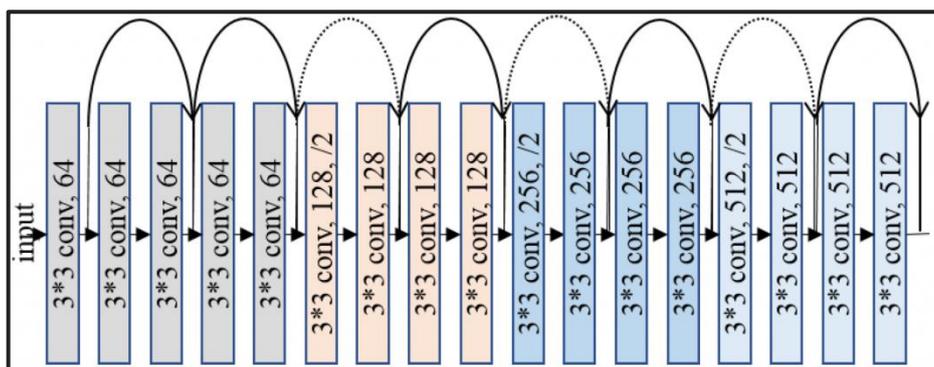


Рис. 3. Архитектура ResNet18.

Гиперпараметры модели для нейронной сети:

- Optimizer = Adam;
- LR = slice(1.10E-06,1.10E-01);
- Epoch = 6;
- Input\_size = 224x224;
- BS = 128.

Для работы над приложением использованы следующие библиотеки: pandas, seaborn, matplotlib, PIL, sklearn, PyTorch, fast, а также для обучения технологиям – курсы Джереми Говарда [4]. Создание Android приложения на языке Kotlin осуществлено посредством среды IDE Android Studio.

Архитектура для приложения основана на работе Виктора Мёнье [5] и представлена на рисунке 4.

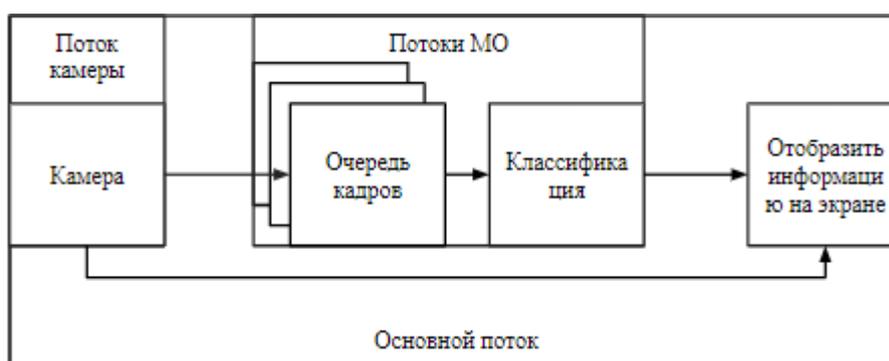


Рис. 4. Архитектура Android приложения

## Заключение

Разработанное многопоточное мобильное приложение для перевода с языка жестов поможет наладить коммуникацию между миром слышащих и миром глухих.

Созданная нейронная сеть может быть перетренирована для перевода с других языков. На данный момент в приложении поддерживается только перевод с американской дактильной речи на английский.

Архитектура приложения универсальна. При замене языка для перевода достаточно заменить только файл с моделью нейронной сети (при условии схожести архитектур этих сетей) и файл с буквами - классами.

Дальнейшие пути развития системы: улучшение распознавания путем тренировки на новых данных и подбора различных гиперпараметров нейросети; добавление новых языков; добавление алгоритма, соединяющего буквы в слова.

## Список литературы

1. Глобальный веб-сайт Всемирной организации здравоохранения. Режим доступа: <https://www.who.int/ru/news-room/fact-sheets/detail/deafness-and-hearing-loss> (дата обращения 27.05.2020).
2. Зайцева Г.Л. *Жестовая речь. Дактилология*: Учеб. для студ. высш. учеб. заведений Гуманит. изд. центр ВЛАДОС, 2000. - 192 с.
3. He K., Zhang X., Ren S, Sun J. *Deep Residual Learning for Image Recognition*. Режим доступа: <https://arxiv.org/pdf/1512.03385.pdf> (дата обращения 27.05.2020).
4. Howard J. *fast.ai Practical Deep Learning for Coders, v3*. Режим доступа: <https://fast.ai> (дата обращения 27.05.2020).
5. Meunier V. *HandPose* Режим доступа: <https://github.com/MrEliptik/HandPose>. [arxiv.org/pdf/1512.03385.pdf](https://arxiv.org/pdf/1512.03385.pdf). (дата обращения 27.05.2020).