

Направление конкурса: Большие данные, искусственный интеллект, финансовые технологии и машинное обучение

Переводчик с русского языка жестов

Автор:

Демидович Ольга Юрьевна,

9 класс МБОУ «Инженерный лицей НГТУ»

Ленинского района

города Новосибирска

конт.тел: +7 909-531-31-45

Наставники:

Гончаров Иван Олегович, инженер машинного обучения

Weights&Biases

Шрайнер Борис Александрович, руководитель
лаборатории искусственного интеллекта НГПУ

Новосибирск 2022

Оглавление

Введение.....	3
Теоретическая часть.....	5
Как работает программа	7
Заключение	9
Список используемой литературы:	10

Введение

В мире существует около 130 жестовых языков, в каждом — около четырёх тысяч жестов, которые производят руками. Люди с ОВЗ ежедневно используют русский язык жестов, однако их общение в магазинах, в государственных органах, медицинских и юридических учреждениях сопряжено с проблемами непонимания жестового языка остальными людьми.

В России, по словам Президента Ассоциации переводчиков жестового языка Лилии Ионичевской, более 2 млн глухих и слабослышащих людей и всего около 1 тыс. переводчиков с русского жестового языка (РЖЯ). Руководитель отдела переводчиков жестового языка Российской специализированной академии искусств Варвара Ромашкина предполагает, что это число еще меньше. Причем большинство из них работают в Москве. Президент Всероссийского общества глухих Станислав Иванов, говорит о нехватке около четырёх тысяч переводчиков с РЖЯ. Однако даже если проблема нехватки переводчиков будет решена, время, отведенное на услуги такого специалиста для одного человека с проблемами речи и слуха, ничтожно мало. Тем не менее, со стороны государства для данной категории лиц по постановлению Премьера Михаила Мишустина будет в два раза увеличен объем услуг сурдоперевода для людей с нарушением слуха. С 2022 года они будут получать 84 оплачиваемых государством часа перевода, вместо положенных ранее 40. Это всего лишь 13-14 минут в день.

Указанная проблематика создает преграду для социализации данной категории лиц. Особо остро это проблема ощутилась во время пандемии, когда нужно оперативно общаться с медицинскими и социальными работниками. Для решения данной проблемы было разработано приложение- переводчик с РЖЯ, которое с помощью камеры и нейронных сетей узнаёт жесты глухих и переводит их в текст.

Проект направлен облегчить общение с людьми с ограниченными возможностями здоровья с людьми, которые не знают жестовый язык, при

этом ограничена возможность использования клавиатуры. Используются такие технологии, как OpenCV, NumPy, TensorFlow, Mediapipe, русский жестовый язык. Целью проекта является снятие барьера в общении, облегчение социализации людей с ОВЗ по слуху. Проект создан, чтобы решить проблему общения и понимая этой категории лиц, ведь для них жестовый язык - родной и единственный.

Теоретическая часть

Для работы программы используется следящий стек технологий (стек технологий — это набор языков программирования, фреймворков и ПО, необходимых для разработки приложения):

Python — это скриптовый язык программирования. Он универсален, поэтому подходит для решения разнообразных задач и многих платформ, начиная с iOS и Android и заканчивая серверными ОС.

Русский жестовый язык (РЖЯ) — национальная лингвистическая система, обладающая собственной лексикой и грамматикой, используемая для общения глухих и слабослышащих, живущих в России[2], а также на территории СНГ (Белоруссия, Казахстан), а также Украины.

OpenCV — это open source библиотека компьютерного зрения, которая предназначена для анализа, классификации и обработки изображений. OpenCV написана на языке высокого уровня (C/C++) и содержит алгоритмы для: интерпретации изображений, калибровки камеры по эталону, устранение оптических искажений, определение сходства, анализ перемещения объекта, определение формы объекта и слежение за объектом, 3D-реконструкция, сегментация объекта, распознавание жестов и т.д.

Пакет NumPy является незаменимым помощником Python. Он тянет на себе анализ данных, машинное обучение и научные вычисления, а также существенно облегчает обработку векторов и матриц.

Tensorflow (далее — TF) — довольно молодой фреймворк для глубокого машинного обучения, разрабатываемый в Google Brain, в 2015 году система была переведена в свободный доступ с открытой лицензией Apache 2.0. для решения задач построения и тренировки нейронной сети с целью автоматического нахождения и классификации образов, достигая качества человеческого восприятия

MediaPipe — это фреймворк с открытым исходным кодом, представленный Google, который помогает создавать мультимодальные

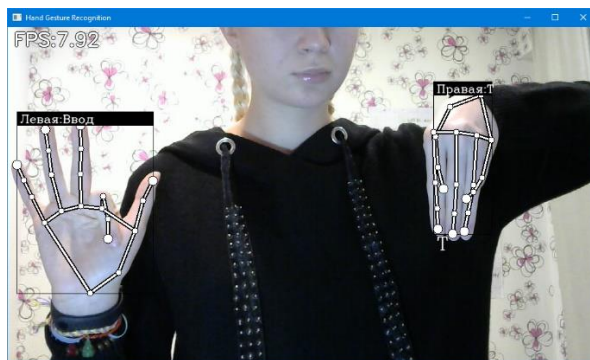
конвейеры машинного обучения. Разработчик может создать прототип, не углубляясь в написание алгоритмов и моделей машинного обучения, используя существующие компоненты. Эта структура может использоваться для различных приложений для обработки изображений и мультимедиа (особенно в виртуальной реальности), таких как обнаружение объектов, распознавание лиц (MediaPipe BlazeFace), отслеживание рук, пальцев (MediaPipe Hands) и прочего. MediaPipe поддерживает различные аппаратные и операционные платформы, такие как Android, iOS и Linux, предлагая API на C++, Java, Objective-C и т.д.

Как работает программа

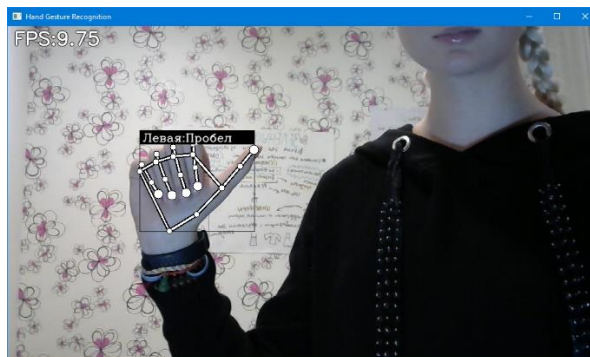
Для обучения нейросети использовалось 10 000 наборов относительных координат по каждому жесту РЖЯ.

Программа с помощью веб-камеры видит жест правой руки, и определяет его принадлежность к букве по точкам на суставах ладони, которые добавляет к изображению сама программа. В левой руке заложен ввод и удаление знаков.

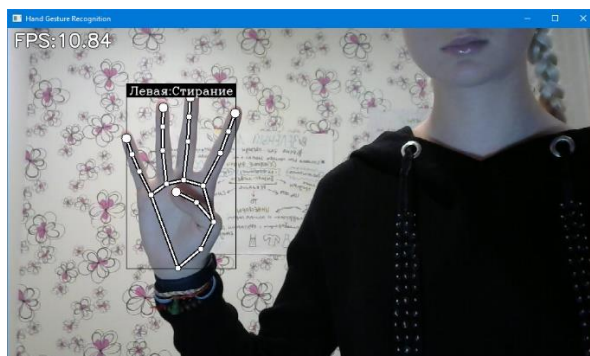
Жесты левой руки:



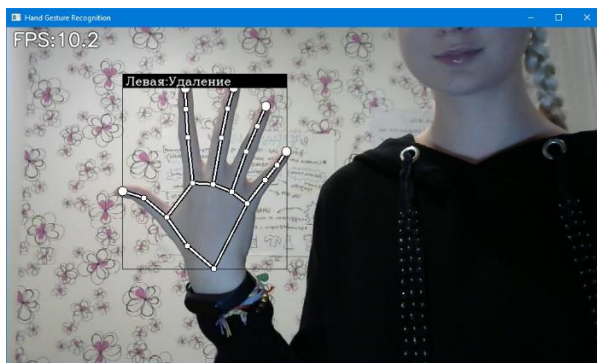
жест “Ввод” - программа добавляет к сообщению, находящемуся под правой рукой, введенную букву;



жест “Пробел” - программа добавляет к сообщению разделяющий знак - пробел;



жест “Стирание” - происходит по-знаковое удаление введенных букв;



жест “Удаление” - всё
сообщение удаляется

Жесты правой руки: буквы русского языка жестов

Для работы программы бралась модель от TensorFlow, модель от Google Mediapipe, которая определяла ключевые точки на руках. Эти точки кодировались в вектор и на их основе собирались данные - примеры того, как выглядят ключевые точки на руках для разных букв. На основе собранных данных тренируется простейшая полно-сцепленная нейросеть. На входные данные выдается этот вектор, в котором закодированы 21 точка. На выходе определяется, какой нейрон горит, и выводится соответствующая буква. Вектор сделан так, что точки берутся относительно друг друга и нормализуются в прямоугольнике. Эти точки - это не просто значения пикселей, а значения от нуля до единицы и означают то, как далеко располагаются точки друг от друга.

Программа использует и во многом основывается на проекте определения положения пальцев рук.

Заключение

Все жестовые языки уникальны и отличаются в разных странах и очень важно, чтобы моя программа развивалась именно в России, ведь это позволяет людям, владеющим русским жестовым языком, свободнее общаться. Использование программы в публичных местах, таких как магазины, государственные и медицинские учреждения - где нет возможности использовать клавиатуру.

Разработанное приложение эффективно и надежно распознает жесты русского жестового языка вне зависимости от особенностей человека, фона и освещения. Оно позволяет буквально “набирать” текст, не прикасаясь к клавиатуре и сенсорному экрану.

На примере этого проекта можно понять, как использовать Mediapipe для других проектов на основе компьютерного зрения.

Для дальнейшего развития проекта предполагается:

- Добавление жестов, которые обозначают сразу целое слово.
- Разработка учебного приложения, которое бы можно использовать для обучения русскому языку жестов.
- Выпуск перчаток, которые могли бы также считывать жесты, так как программа использует точки на руках
- Голосовая интерпретация введенных слов на языке жестов

Список используемой литературы:

1. Hand gesture recognition using mediapipe
<https://github.com/kinivi/hand-gesture-recognition-mediapipe>
2. MediaPipe Hands <https://google.github.io/mediapipe/solutions/hands>
3. Глава Всероссийского общества глухих сообщил о нехватке около 4 тыс. сурдопереводчиков, 4.10.2021 <https://tass.ru/obschestvo/12569661>
4. Наталья Костарнова «Правительство просят не экономить на жестах», 16.03.2021 <https://www.kommersant.ru/doc/4730061>
5. Русский жестовый язык https://www.youtube.com/watch?v=-Q0VQZbLlHo&list=PL_VzOPUOoGKQQ7hyco7uz5apagoF99raf