



“If you're not making someone else's life better,
then you're wasting your time. Your life will become
better by making other lives better.”
— Will Smith

Under the guidance of Prof. Seema Yadav, Dr. Alberto Boschetti & Packt Group

Advanced Indian Sign Language to text & speech interpreter.

Group Members:

- Abrar Shaikh
- Prajwal Shinde
- Huzefa Ratlamwala



Outline

- Introduction
- Our Research
- Current Scenario
- Our Solution
- Our Strategy
- Data Collection
- Video Pre-processing techniques used
- Deep Learning results

Introduction

It's not what happens to us, but it's how we respond to it that matters the most. That's what has driven us to do this project. We do not see impaired people as impaired. They just have a different mode of communication that we - normal humans cannot understand.

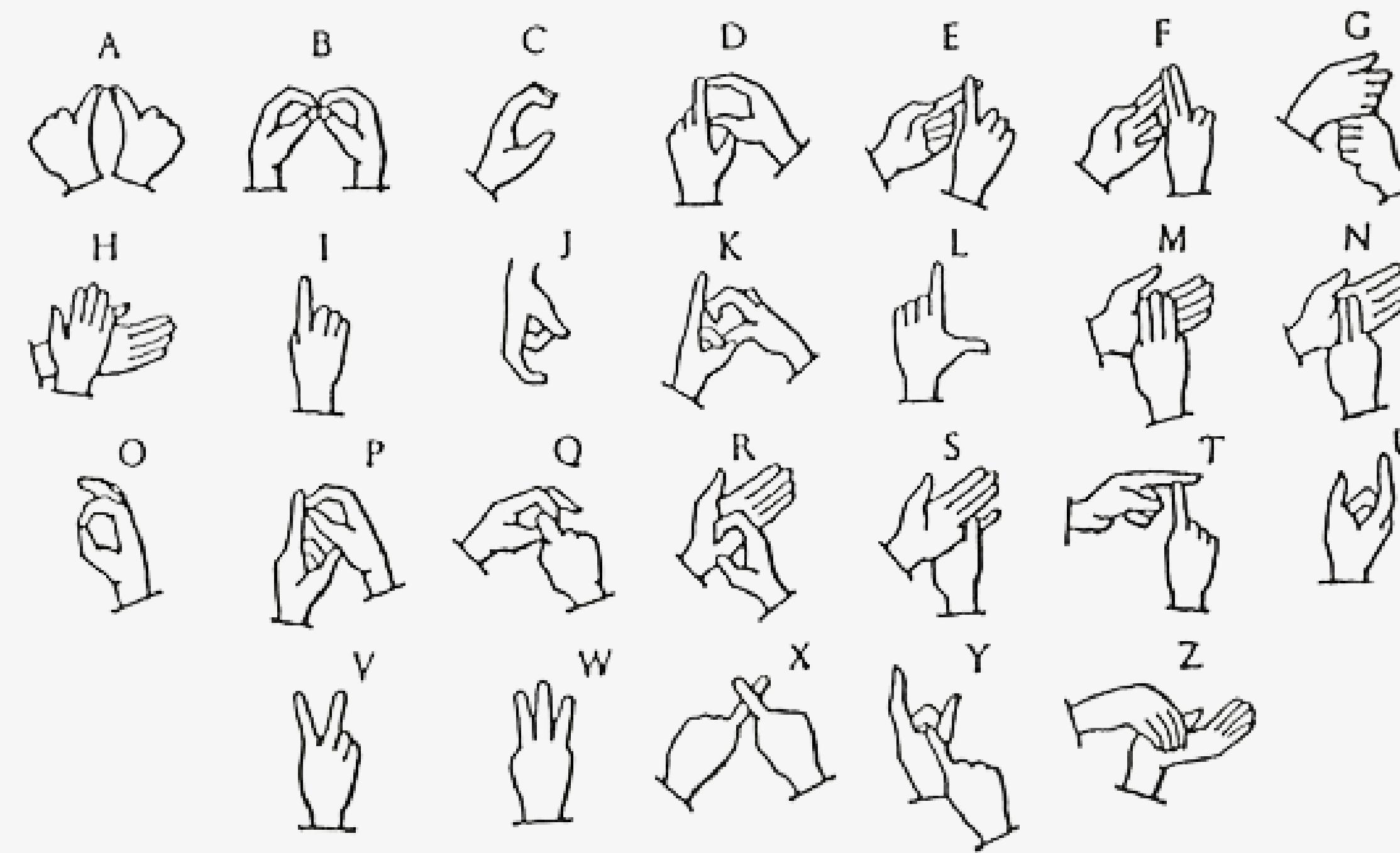


Fig. Alphabets in Indian Sign Language (a mission accomplished)



Our Research



ALI YAVAR JUNG NATIONAL INSTITUTE

To conduct, sponsor, coordinate or subsidize research into all aspects of education and rehabilitation of persons with **hearing impairment**.



There are only about 250 certified sign language interpreters in India, translating for a deaf population of between 1.8 million and 7 million.

[1]



- No ready-made dataset of videos is available
- Very few resources are available on the topic.
- Changing and evolving the mindset of people.

Current Scenario

Hardware Dependency[4]

Kinect based detection.
[2]

Marker Free Solution

Lack of ISL Support [1]



All about the project

Our Vision

**To create a robust platform that can be
used by impaired people all around India.
Which can convert not just alphabets but
words. yess!**

Key points of our system

Marker- Free

Ease of use

Works with
words.

Made by them
for them

How do we get there?

We have a three-step plan
to make this mission happen

Our Strategy

Step 1

Finalize the best possible image processing methods for hand detection

Step 2

Record videos of all the basic gestures with the help of our impaired friends & clean the data

Step 3

Training of a robust RNN model to predict the gestures.

Step 4

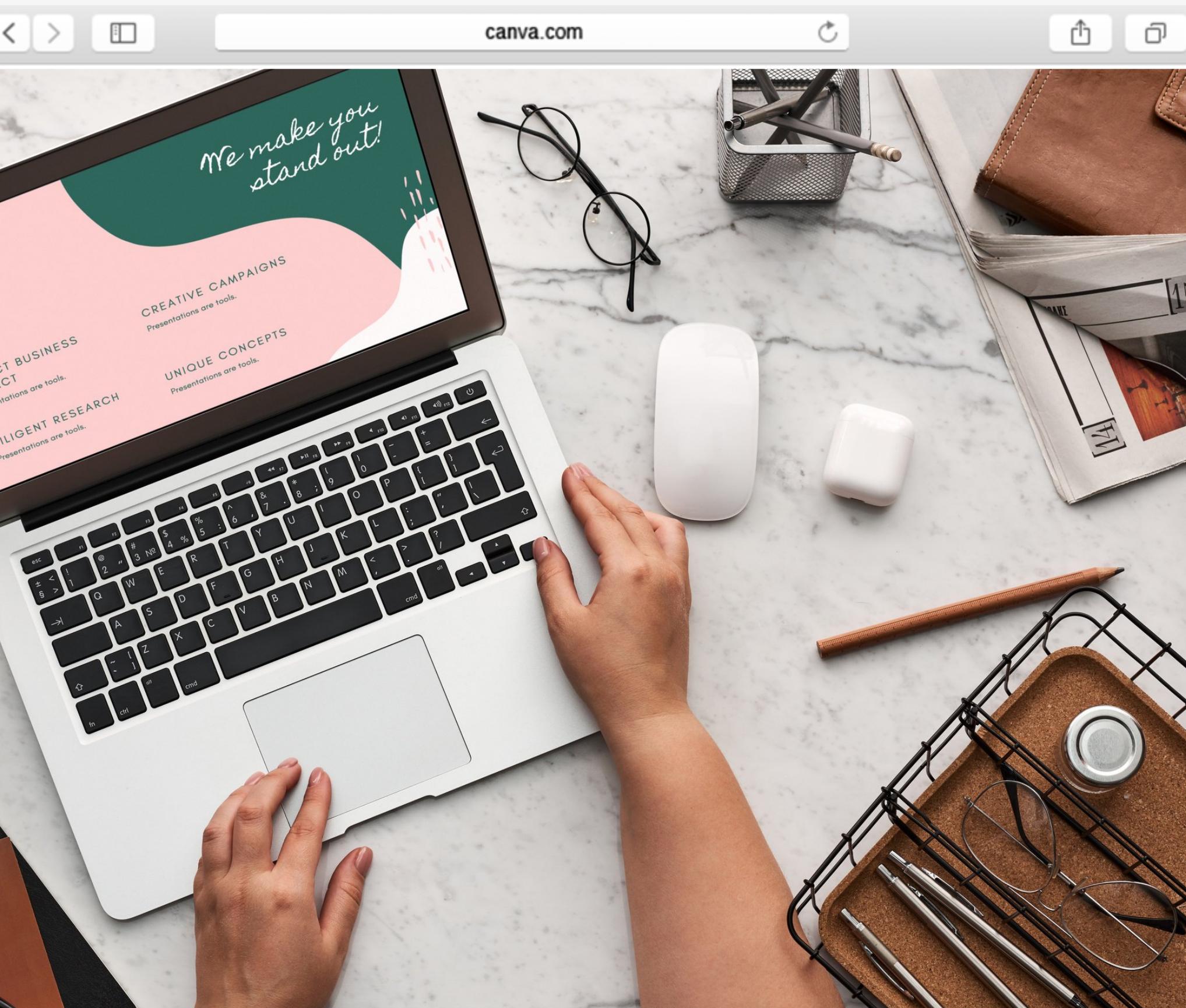
Creation of easy to use UI Interface which can be used to inference the model.

we are here!

step 1

Data Set Collection





Brief Summary

- The entire dataset for the project was created by us.
- For now, we have only worked on 3 gestures "hello", "men" & "women".
- For each gesture, we have collected roughly 10 - 15 videos.

Our Strategy



Step 1

Open the dataset creation script. Once you are ready click on start button



Step 2

It will start recording.
After the gesture is done press stop.



Step 3

The video is then converted into 700 x500 px.



Step 4

The holistic detection is done and converted into a NumPy array and is then saved on the local machine.

Step by Step

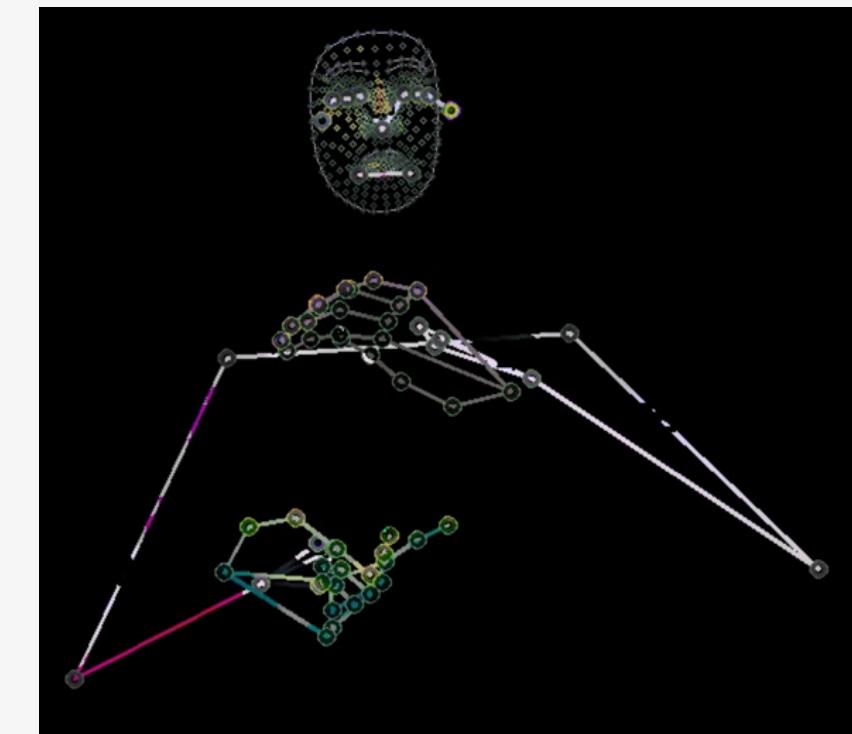
Step 1



Step 2



Step 3



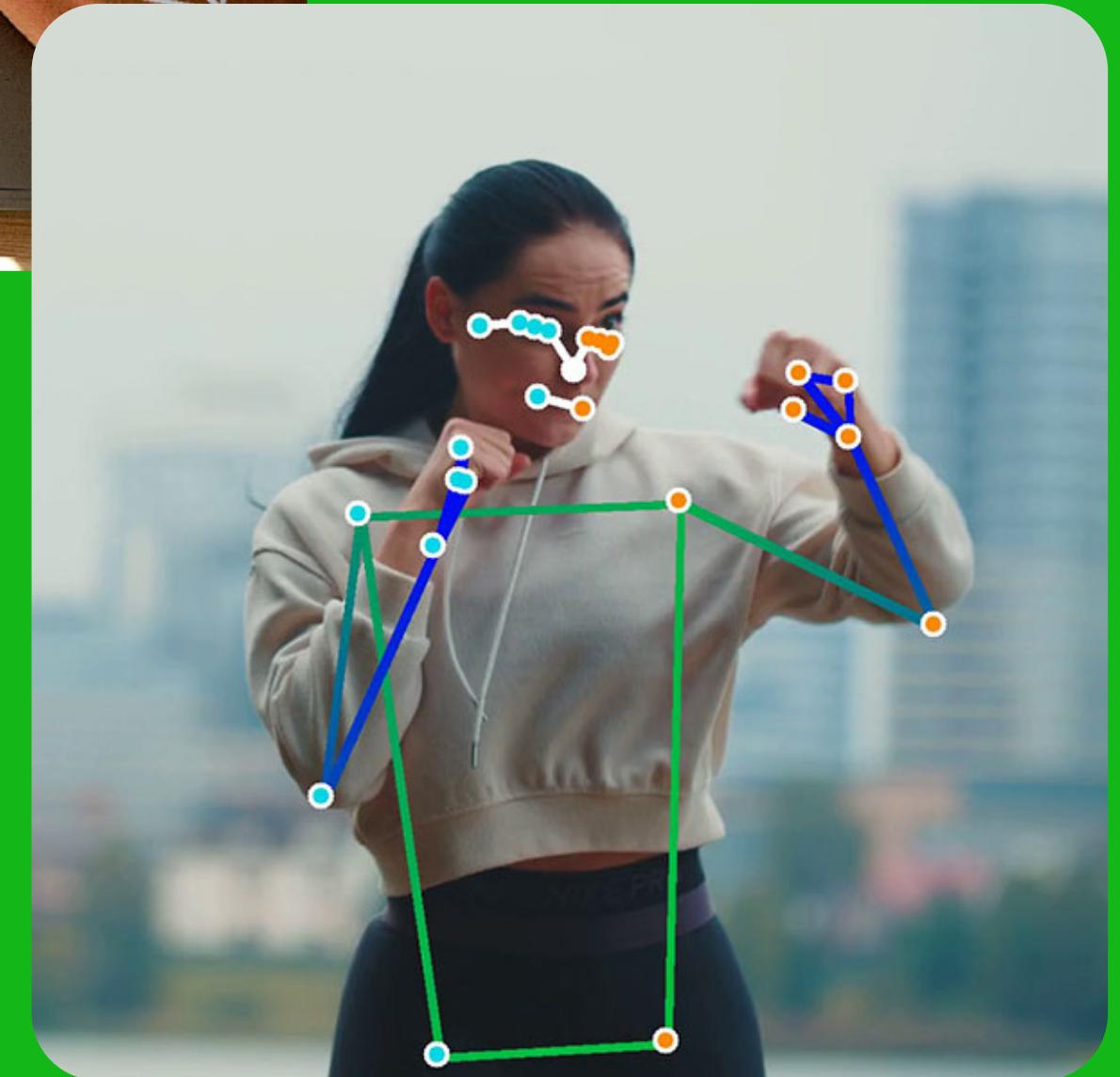
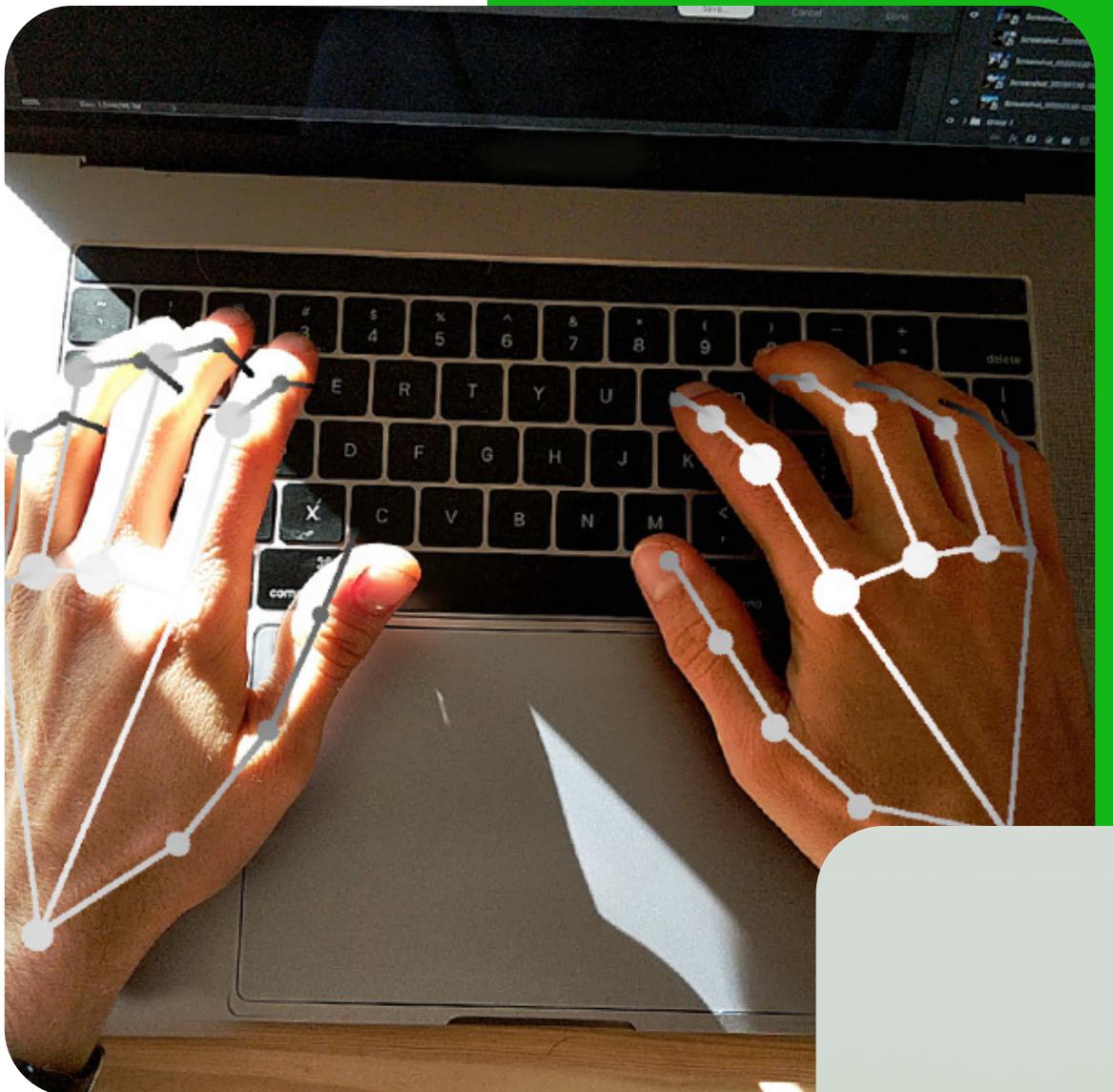
step 2

Pre-Processing



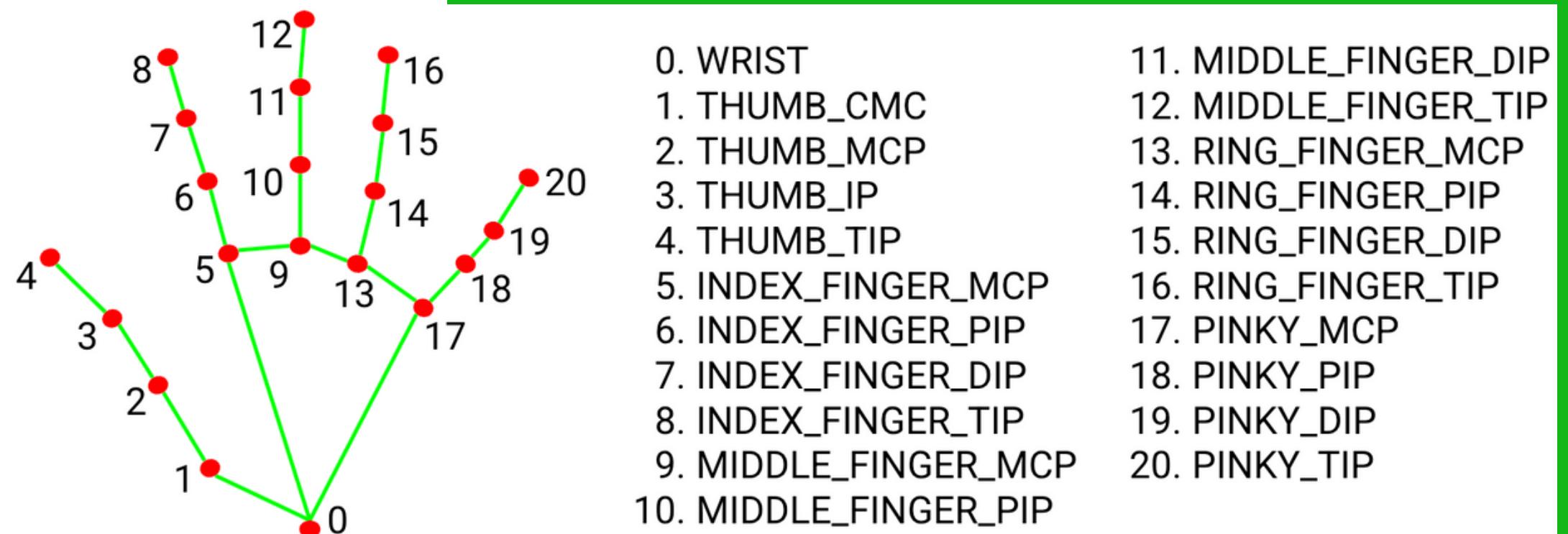
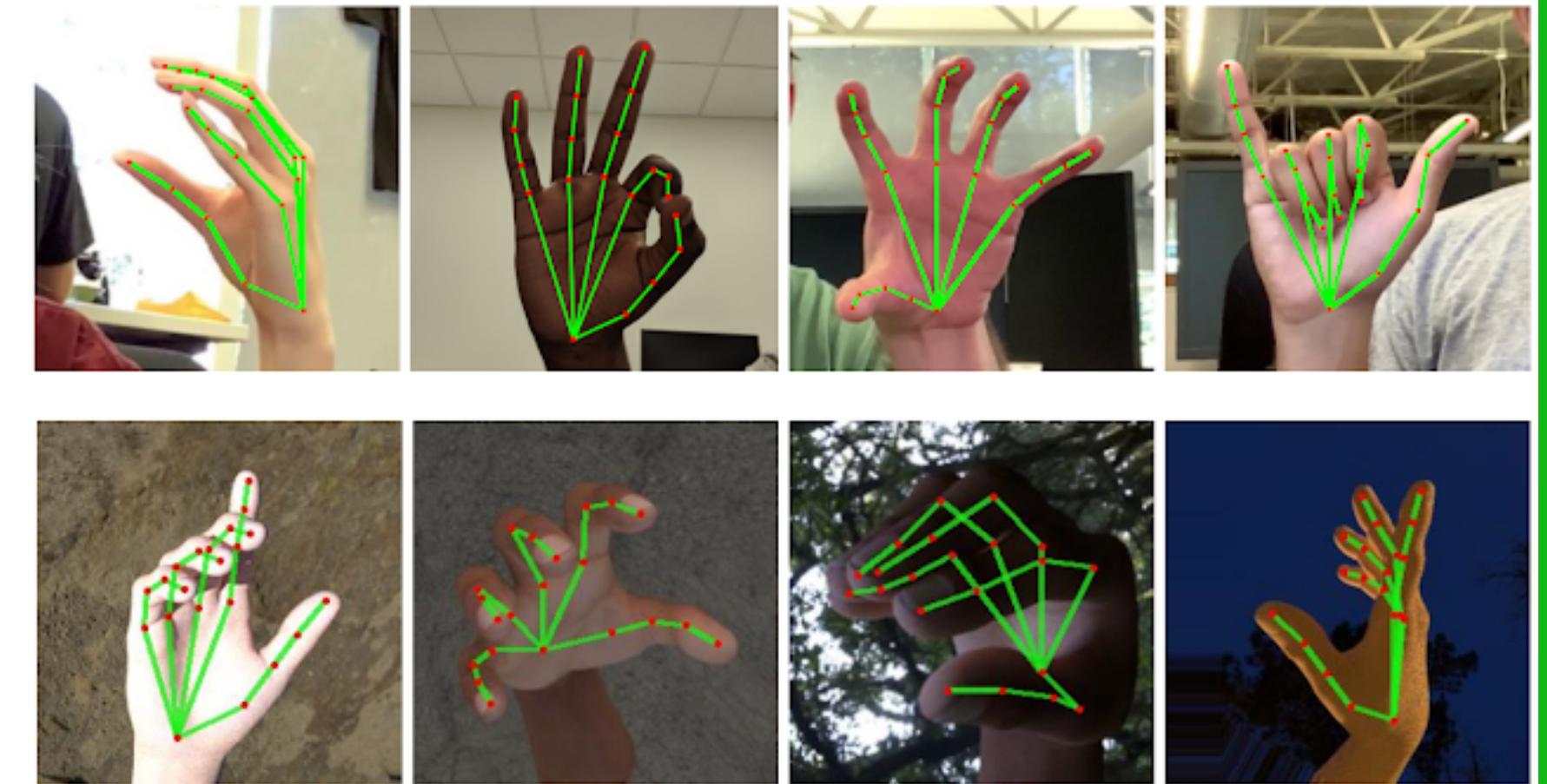
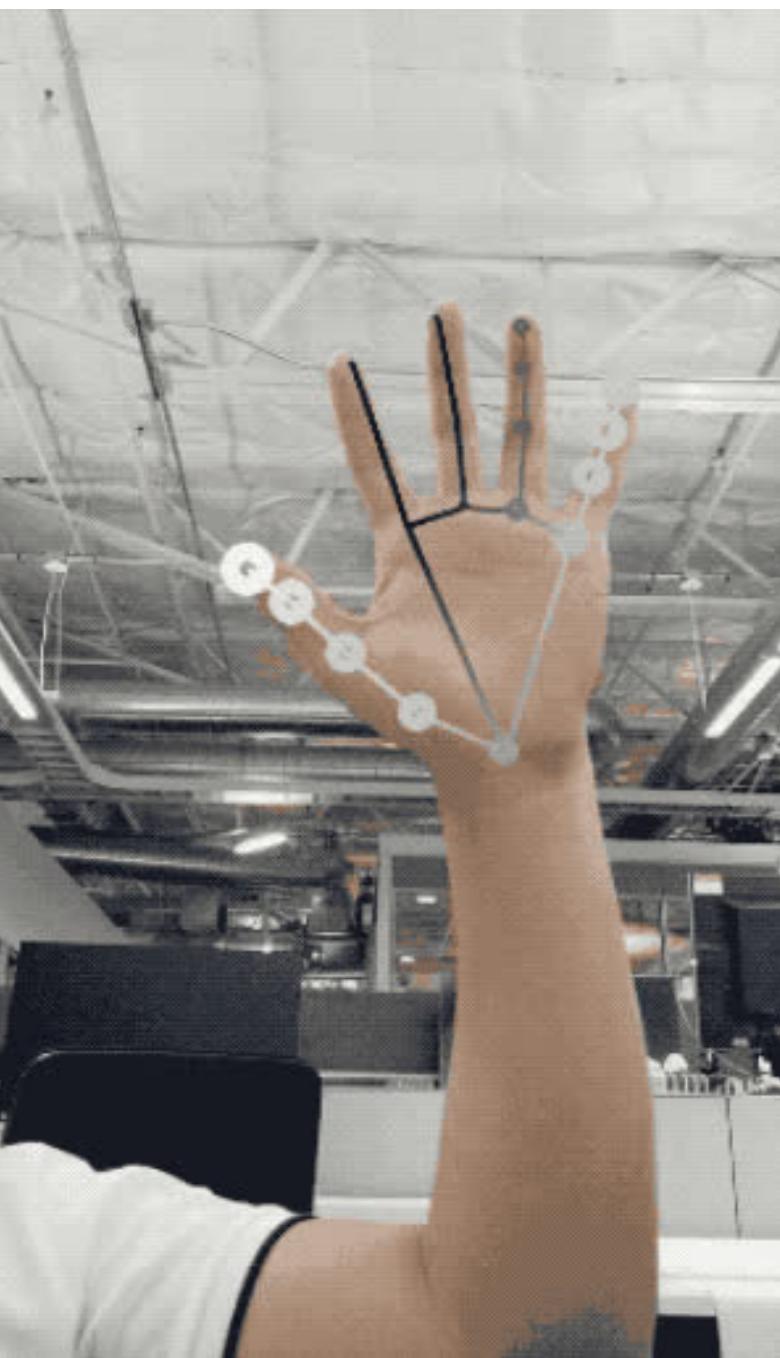
Method 1

Tracking finger movements
and posture



how it works?

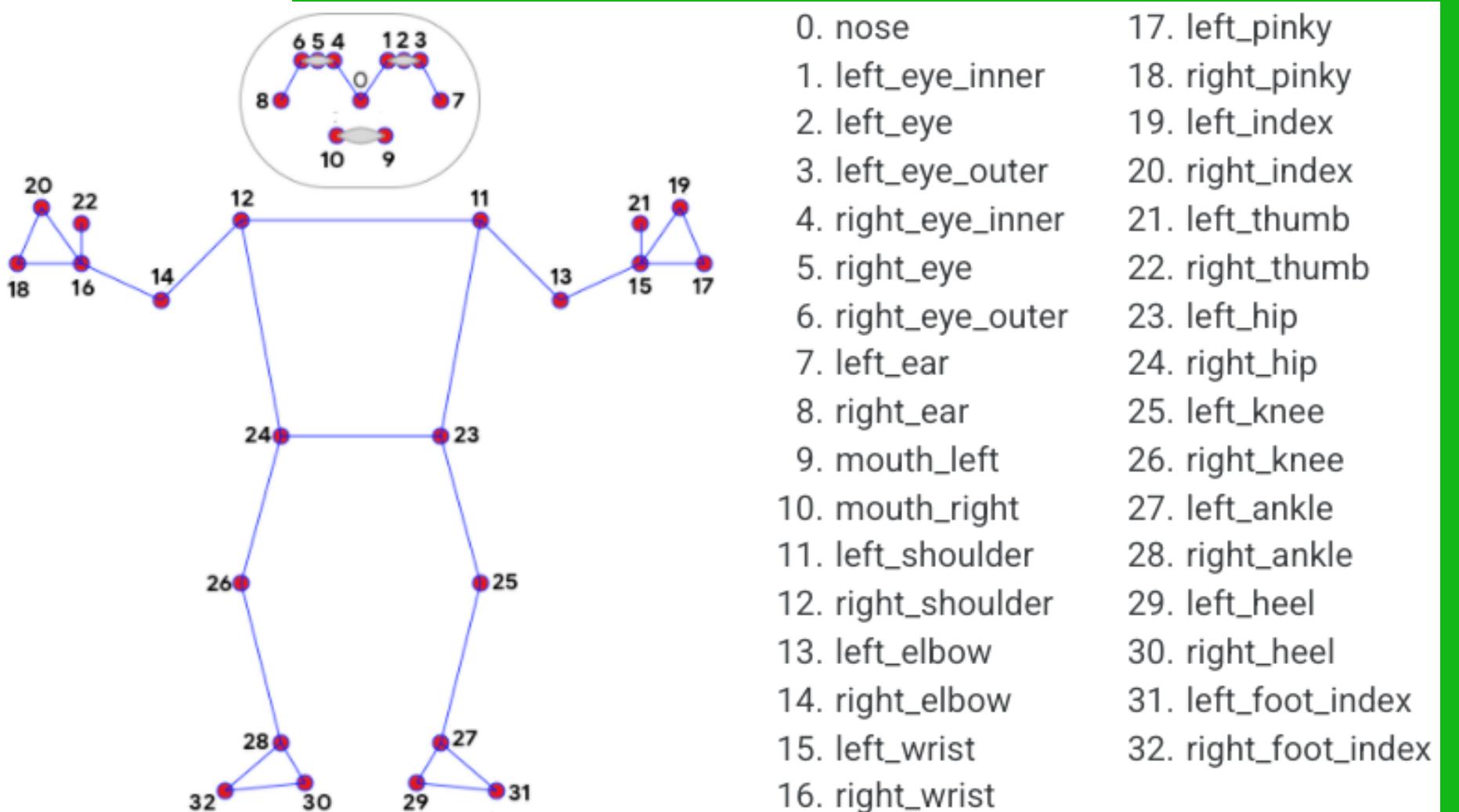
Hand Detection



how it works?

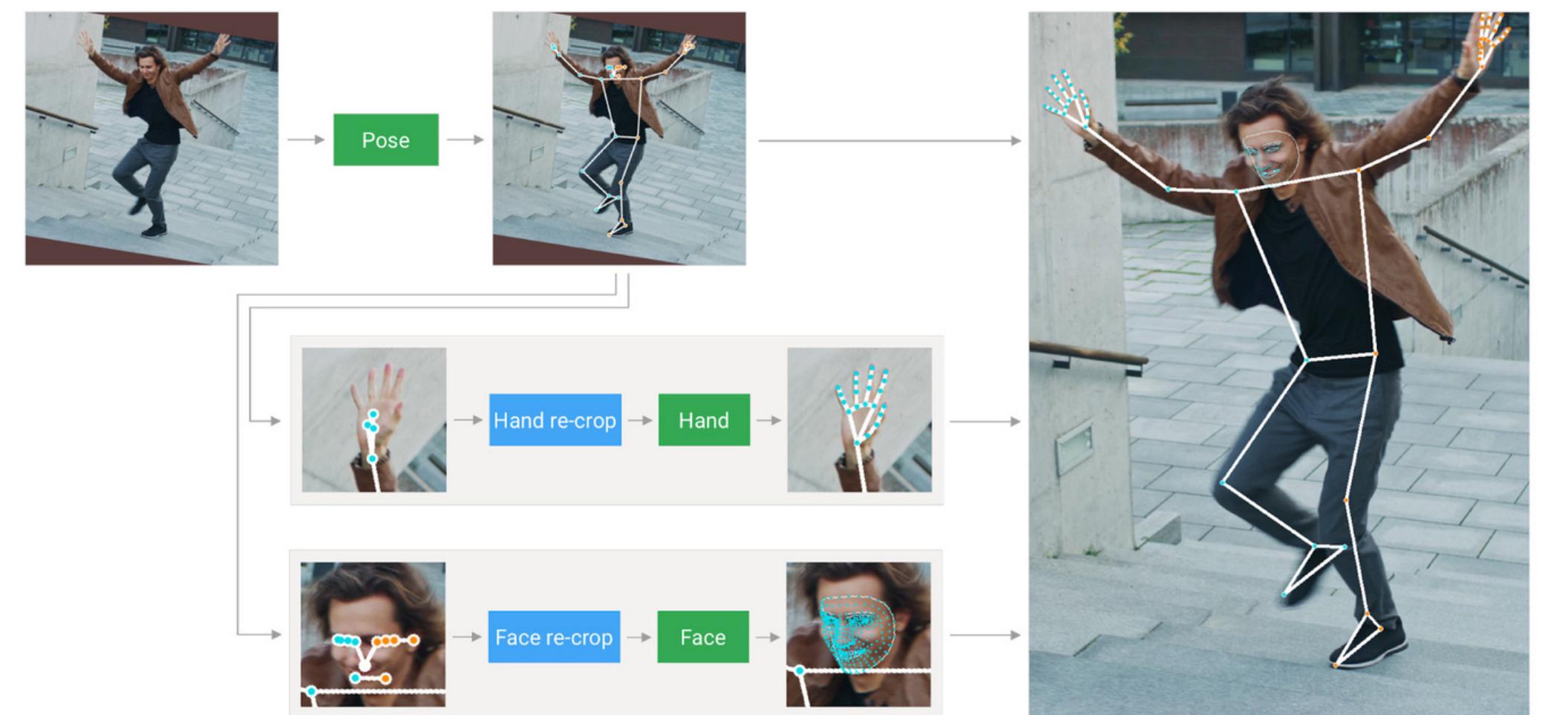
Pose Detection

The detector is inspired by our own lightweight BlazeFace model, used in MediaPipe Face Detection, as a proxy for a person detector. It explicitly predicts two additional virtual keypoints that firmly describe the human body center, rotation and scale as a circle. Inspired by Leonardo's Vitruvian man, we predict the midpoint of a person's hips, the radius of a circle circumscribing the whole person, and the incline angle of the line connecting the shoulder and hip midpoints.



Method 2

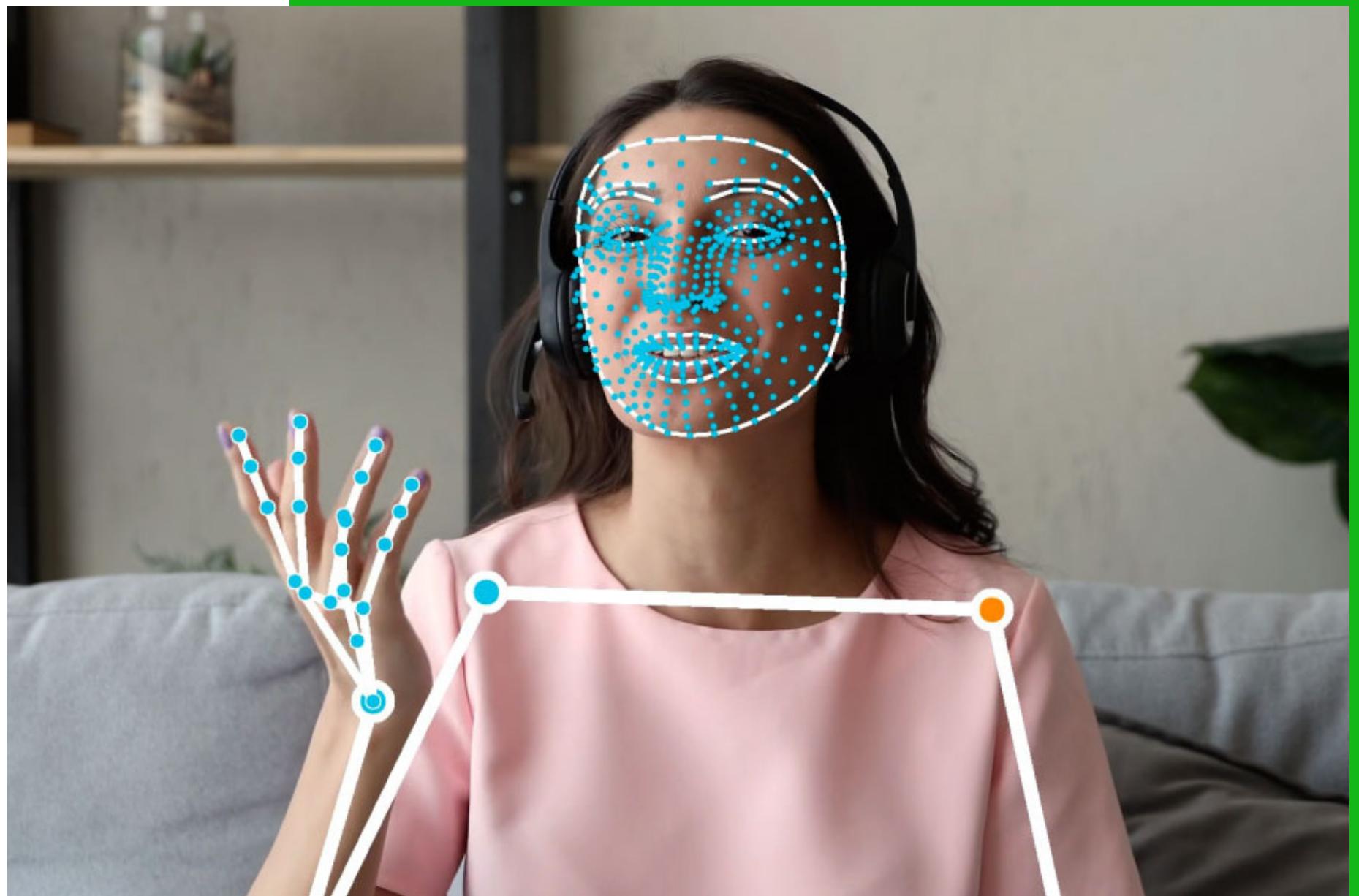
Perform complete holistics tracking
(human pose, face landmarks, and
hand tracking)



how it works?

Holistics Detection

The MediaPipe Holistic pipeline integrates separate models for the **pose**, **face** and **hand** components, each of which are optimized for their particular domain into one.



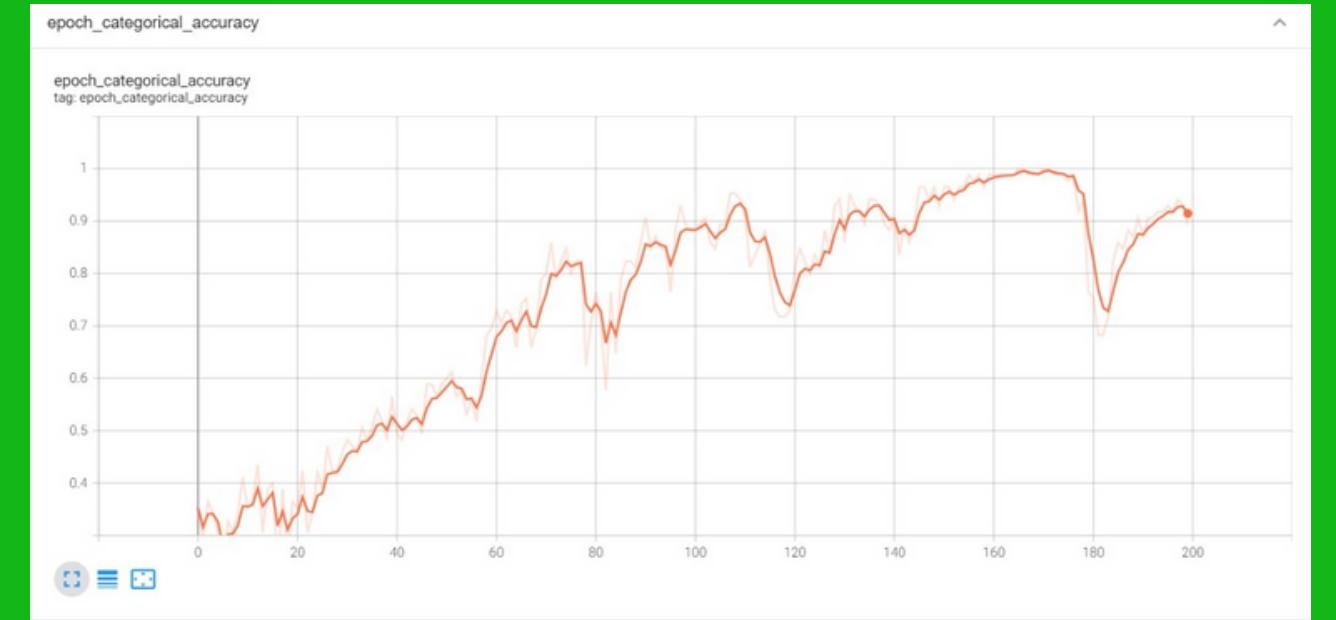
step 2

Deep Learning



LSTM

Long short-term memory is an artificial recurrent neural network architecture used in the field of deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections. It can process not only single data points, but also entire sequences of data.



Validation Accuracy



Validation Loss

Future Scope

1

Creating a multi-platform application.

2

Increase the no of gestures tracked to cover as many gestures as we can.

3

Creating an interface for the impaired person to upload his/her own gesture recordings.

References

- [1] <https://www.pri.org/stories/2017-01-04/deaf-community-millions-hearing-india-only-just-beginning-sign>
- [2] Pratik Likhar; Neel Kamal Bhagat; Rathna G N, Deep Learning Methods for Indian Sign Language Recognition, Berlin, Germany 2021.
- [3] Diksha Hatibaruah; Anjan Kumar Talukdar; Kandarpa Kumar Sarma, A Static Hand Gesture Based Sign Language Recognition System using Convolutional Neural Networks, New Delhi, India 2020.
- [4] Ebey Abraham; Akshatha Nayak; Ashna Iqbal, Real-Time Translation of Indian Sign Language using LSTM, Bangalore, India 2020.
- [5] <https://google.github.io/mediapipe/>

Thank you!