

Volume Modulation using Hand Gestures

**A Special Assignment Presentation under
FUNDAMENTALS OF IMAGE AND VIDEO PROCESSING**

Submitted by:

Name of Students: Ishan Tewari and Labdhi Sheth

Roll No's: 18BCE080 and 18BCE101

ABSTRACT

- Human imitation for his encompassing surroundings makes him interfere in each details of this great environment. People with hearing impairment are gesturing with one another for delivering a selected message, this technique for correspondence likewise draws in human impersonation regard for cast it on human-computer interaction.
- Over time, analysis in varied fields of computer science have increased the capabilities of sensible homes with refined convolutional machine learning models that constantly analyze sound input for activation phrases and context dependent correction of detected words and phrases in commands.
- Google with its MediaPipe is working on its real time had tracking system



Problem Statement

- The current exploration exertion characterizes a climate where a number of difficulties have been considered for getting the hand motion acknowledgment strategies in the virtual climate.
- Hand tracking system has its own set of barriers as:

The dimension of
the input

Clarity of the
video input

Presence of
multiple hands

Skin texture and
color of the hands

Problem statement continued-

Our project surrounds the topic of volume control using hand gestures and thus the pointers to index finger and thumb is used.

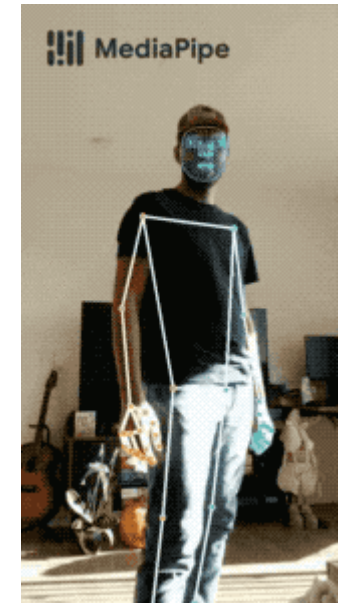
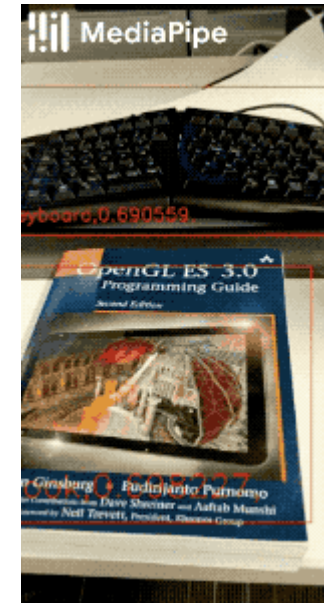
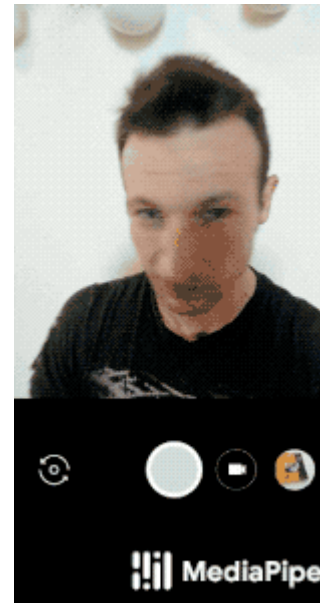
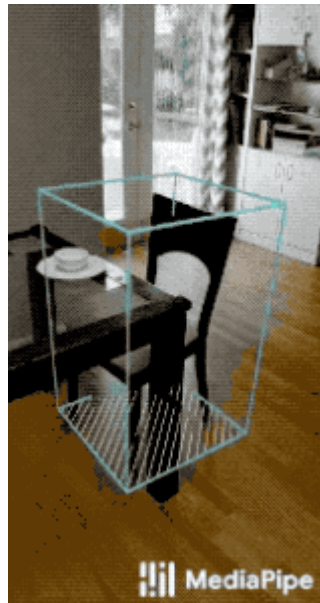
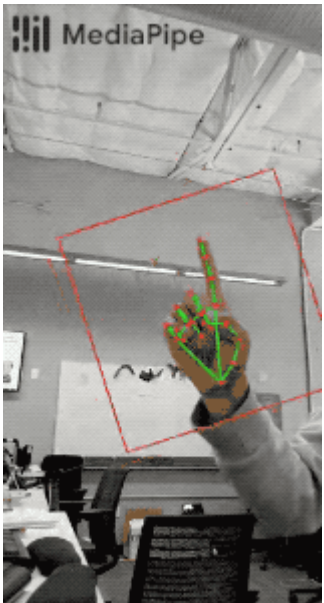
These tasks are performed using Image and video processing by manipulating the image as input and locating the object of importance.

OpenCV library has been used, thus the BGR images are changed to RGB images and to those images we apply MediaPipe algorithm.

Objective

- Hand tracking module which tracks down the palm and locates points on the palm called the landmarks.
- Measure the length of the difference between the position of index finger and the thumb to calculate the volume indication as hand gesture by the user.
- Identify the hand movements and as soon as the user freezes at a point, that volume is set to the device.

What is MediaPipe?

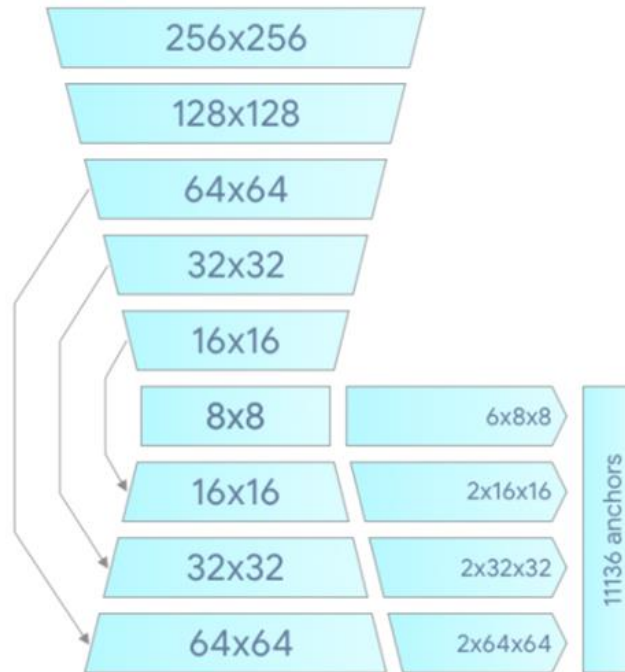


Contribution of MediaPipe

- An efficient two-stage hand tracking pipeline that can track multiple hands in real-time on mobile devices.
- A hand pose estimation model that is capable of predicting 2.5D hand pose with only RGB input.
- An open source hand tracking pipeline as a ready-to-go solution on a variety of platforms, including Android, iOS, Web (Tensorflow.js) and desktop PCs.

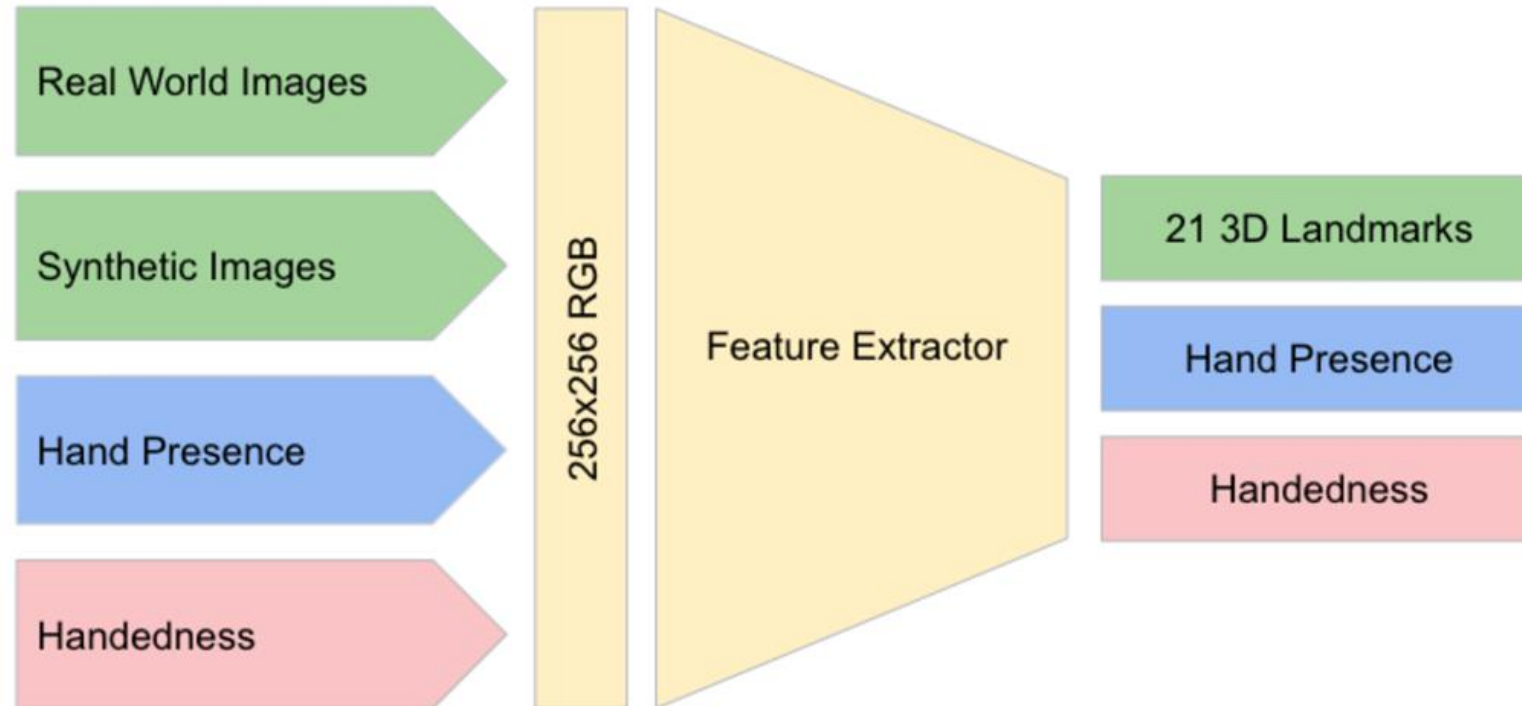
Architecture

Palm Detection Model:



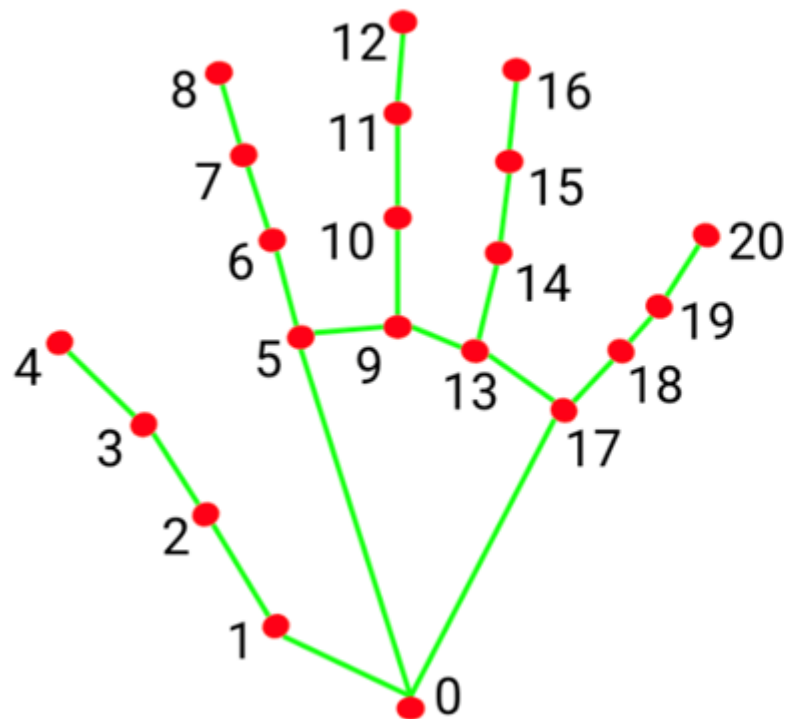
Architecture continued-

Hand Landmark Detection Model



Architecture continued-

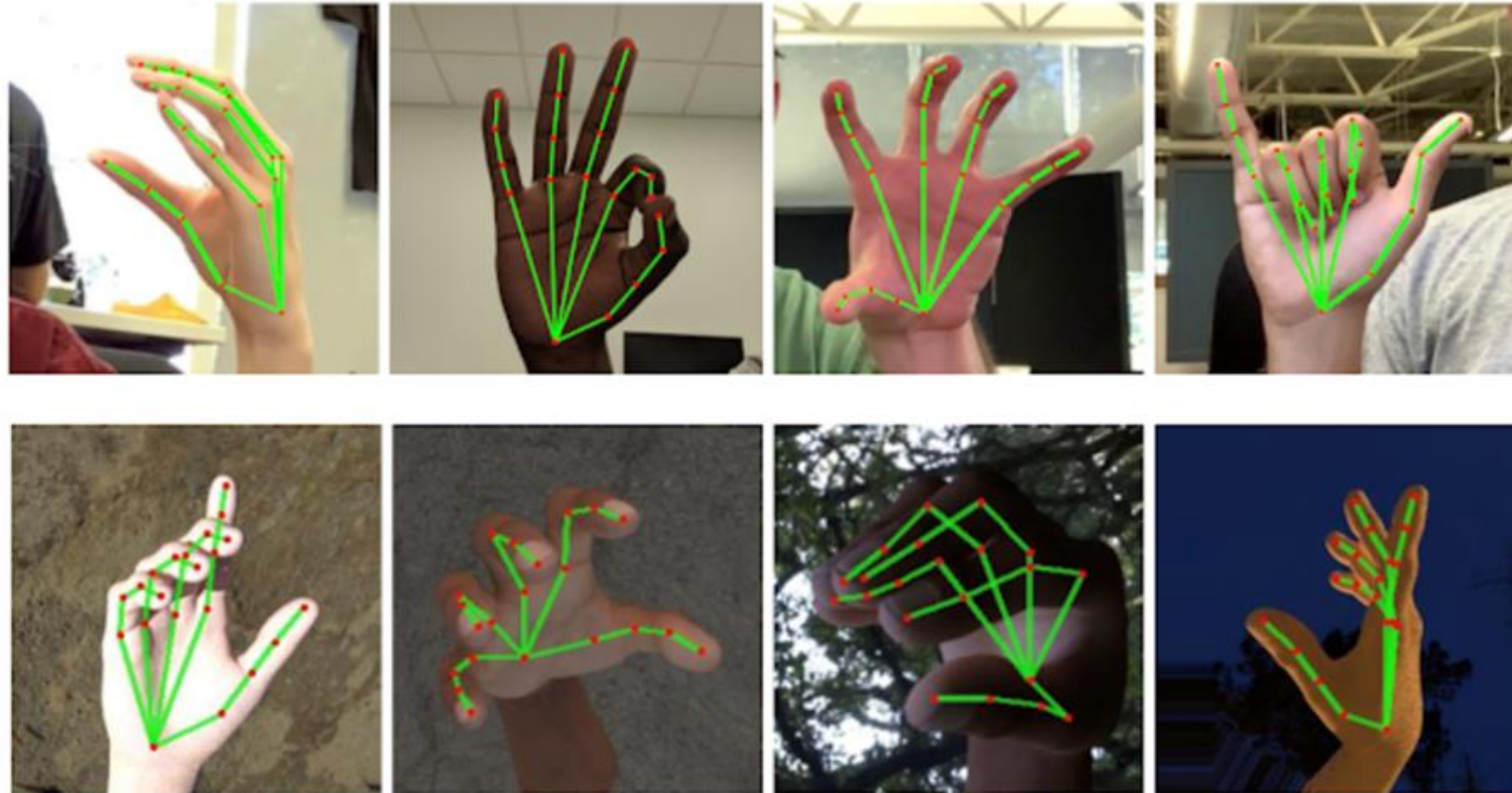
Landmark point



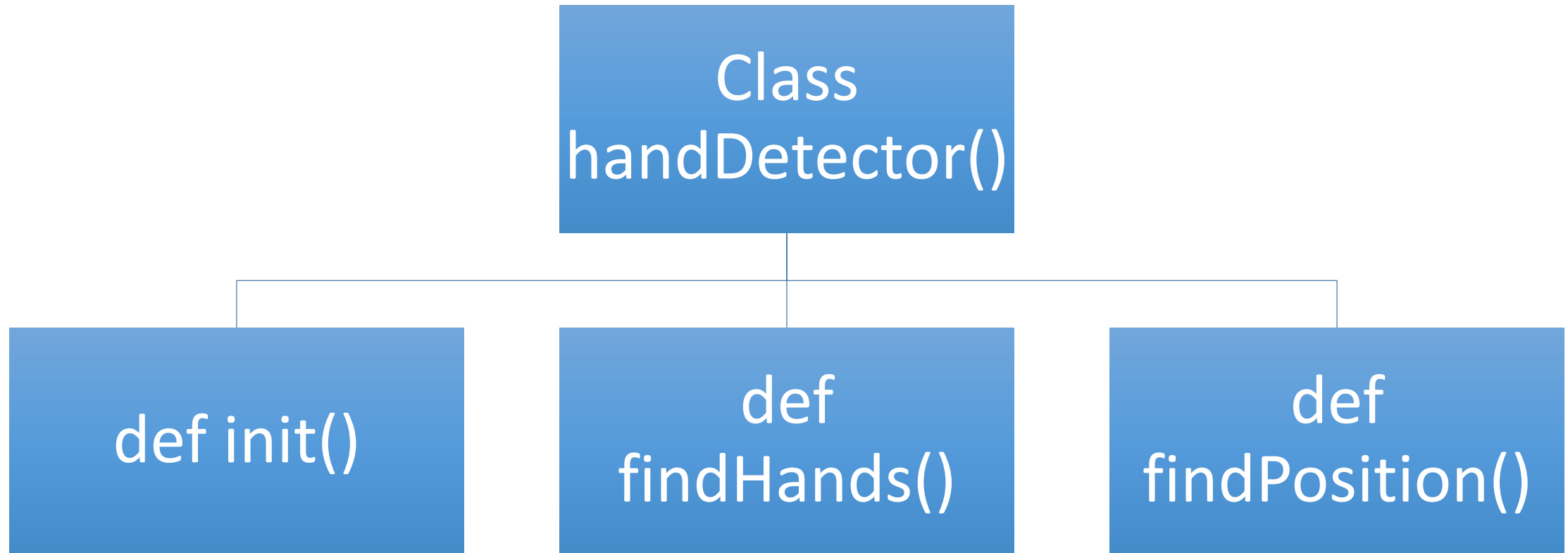
- | | |
|-----------------------|-----------------------|
| 0. WRIST | 11. MIDDLE_FINGER_DIP |
| 1. THUMB_CMC | 12. MIDDLE_FINGER_TIP |
| 2. THUMB_MCP | 13. RING_FINGER_MCP |
| 3. THUMB_IP | 14. RING_FINGER_PIP |
| 4. THUMB_TIP | 15. RING_FINGER_DIP |
| 5. INDEX_FINGER_MCP | 16. RING_FINGER_TIP |
| 6. INDEX_FINGER_PIP | 17. PINKY_MCP |
| 7. INDEX_FINGER_DIP | 18. PINKY_PIP |
| 8. INDEX_FINGER_TIP | 19. PINKY_DIP |
| 9. MIDDLE_FINGER_MCP | 20. PINKY_TIP |
| 10. MIDDLE_FINGER_PIP | |

Architecture continued-

Combined Model



Implementation



Implementation continued-

Use pycaw library to get the details of volume levels of the computer.

Get the landmark location of thumb and index finger.

Calculate distance between landmarks.

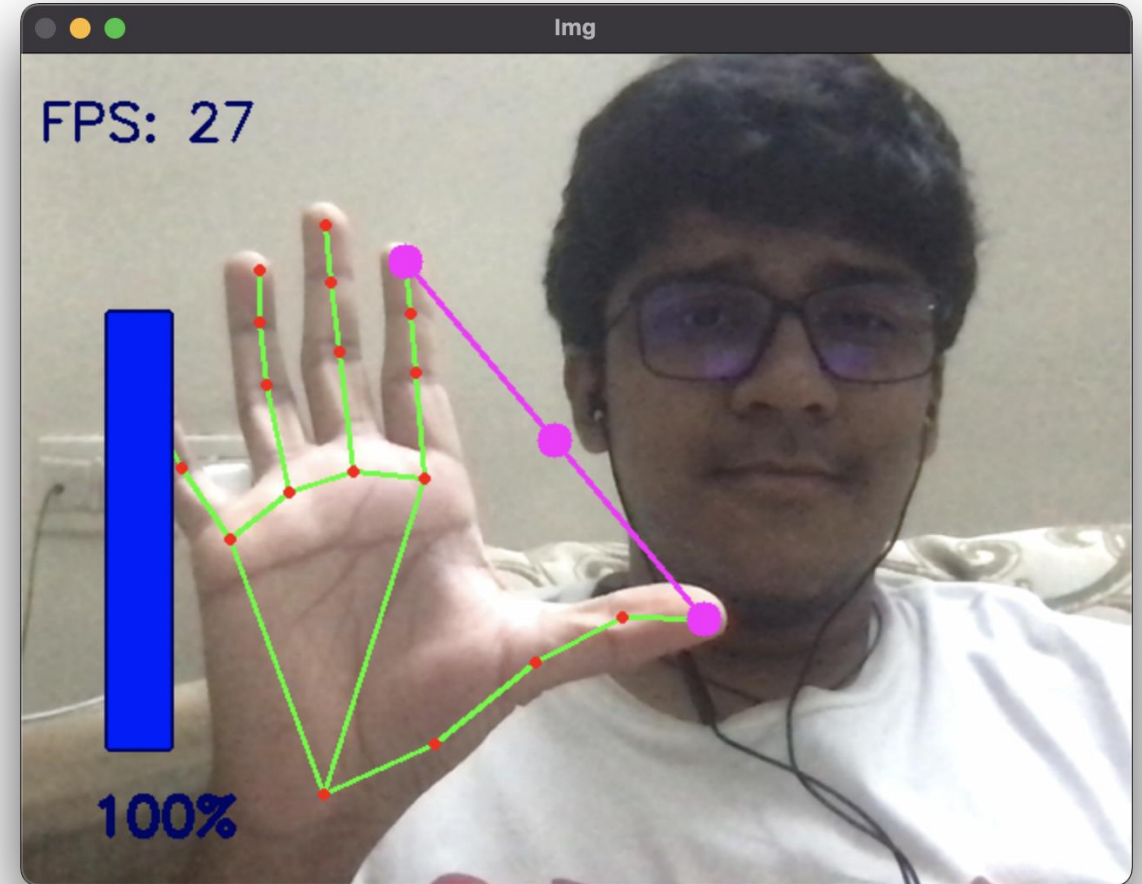
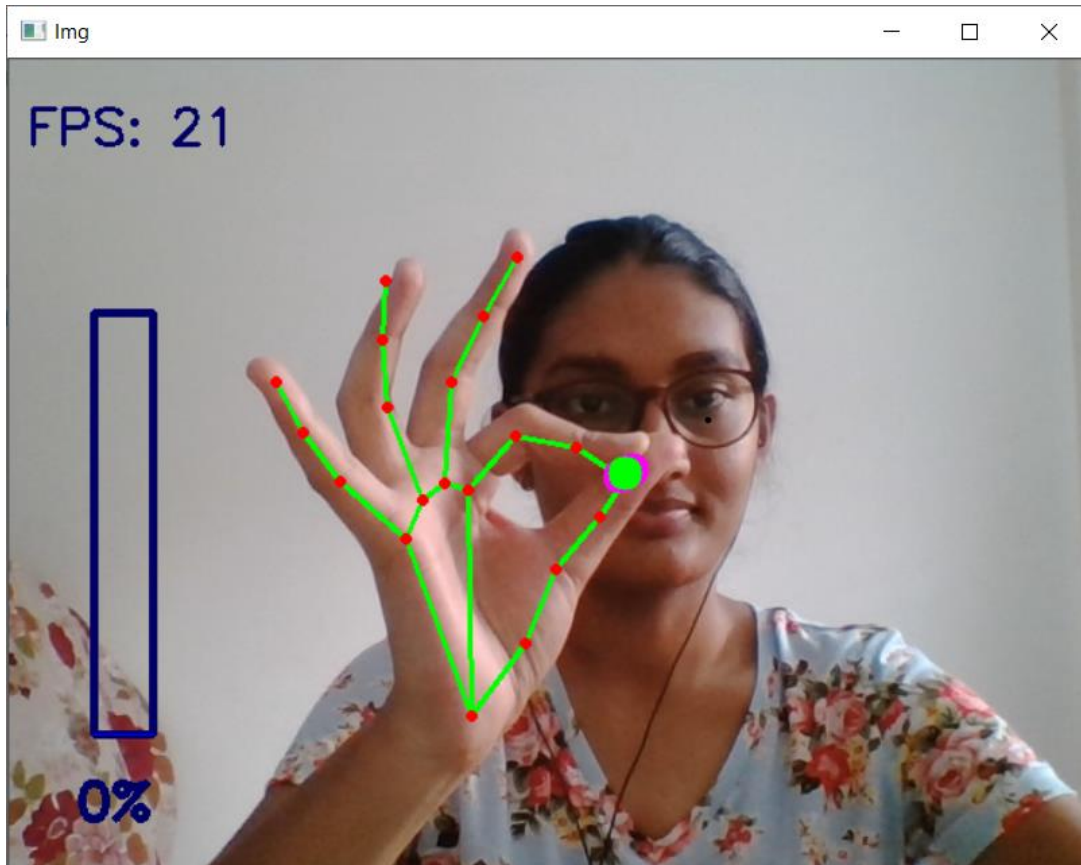
Transform the range of distance to the range of volume calculated using pycaw.

Use this metric to set master volume of the system.

Implementation Results

- The distance of the hand from the camera is inversely proportional to the accuracy of the change in volume i.e. the further away the hand is from the camera, more difficult it is for the model to predict the change in length accurately.
- The naive implementation was a bit too sensitive so we reduced the sensitivity with some fine tuning.

Implementation Results continued-



Conclusion

- Our paper suggests a solution where human computer interaction can be merged with daily use smart home devices where constructing an efficient hand tracking system is an important aspect.
- MediaPipe can be easily used as a tool to accurately determine hand gestures.
- As per the implementation results we have overcome the problem of always terminating the code and always ending on zero volume

Future Scope

Our future scope would be to come up with

1. Some relative distance measure where in the volume perceived remains constant with respect to the hand distance from the camera.
2. Expanding the scope of the project by implementing tracking other gestures
3. Tune the model so as to get finer landmarks for more accurate results.
4. Integrate into smart home devices.
5. Add support for multiple operating systems.

Thank You 😊