



HAND GESTURE RECOGNITION

By Shitij Agrawal

INDEX

- INTRODUCTION
- COMPONENTS USED
- IMPLEMENTATION AND WORKING
- OUTPUT SCREEN-SHORTS
- CONCLUSION

INTRODUCTION

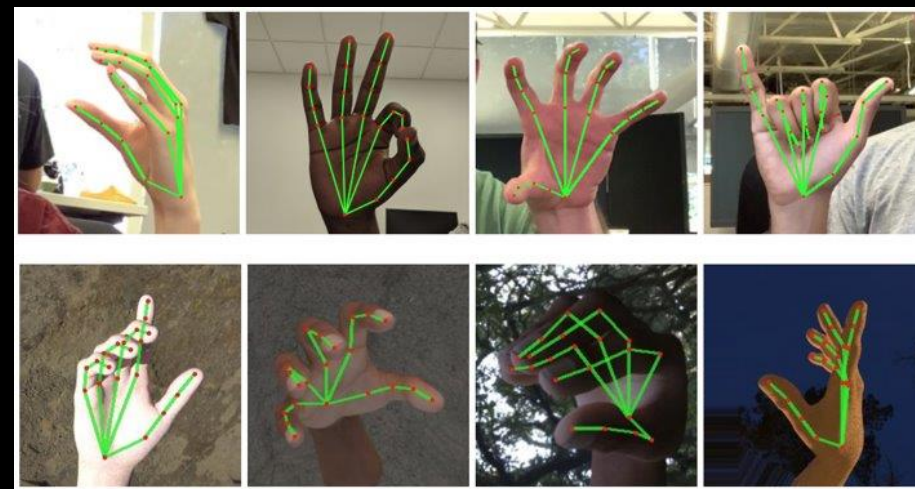
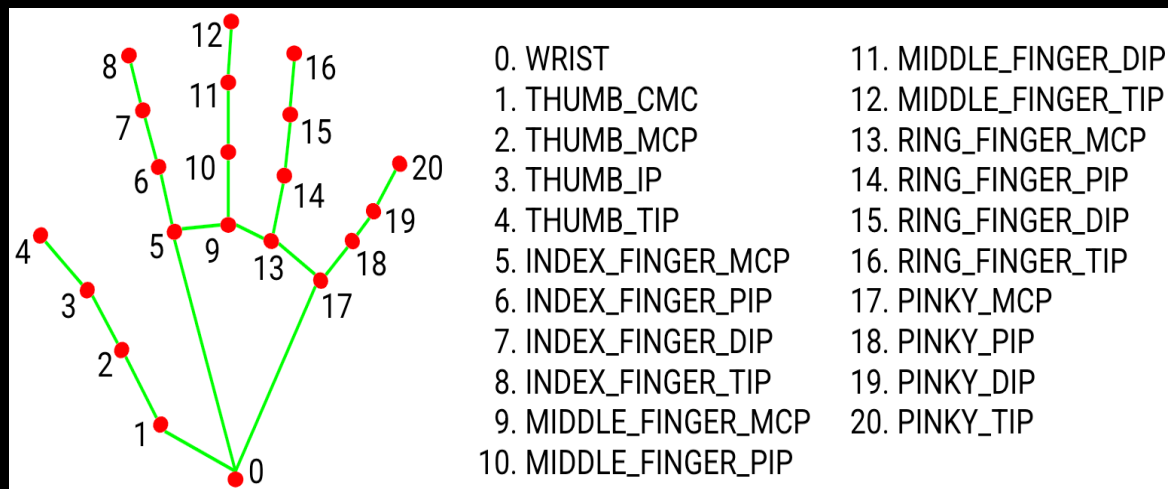
- IN TODAY'S WORLD, WE HAVE CREATED LOTS OF SHORTCUT TECHNIQUES FOR OUR DAY-TO-DAY WORK LIKE DIGITAL ASSISTANT TO MANAGE OUR BASIC WORKS, CONTROLLING HOME BY AUTOMATED WAY, ETC.
- THE HAND GESTURE RECOGNITION CAN PROVIDE ONE OF SUCH WAY TO MAKE OUR DAY-TO-DAY LIFE EASY
- HAND GESTURE RECOGNITION CAN ALSO BE VERY USEFUL FOR THE PEOPLES WHO HAS ANY KIND OF DISABILITY

COMPONENTS USED

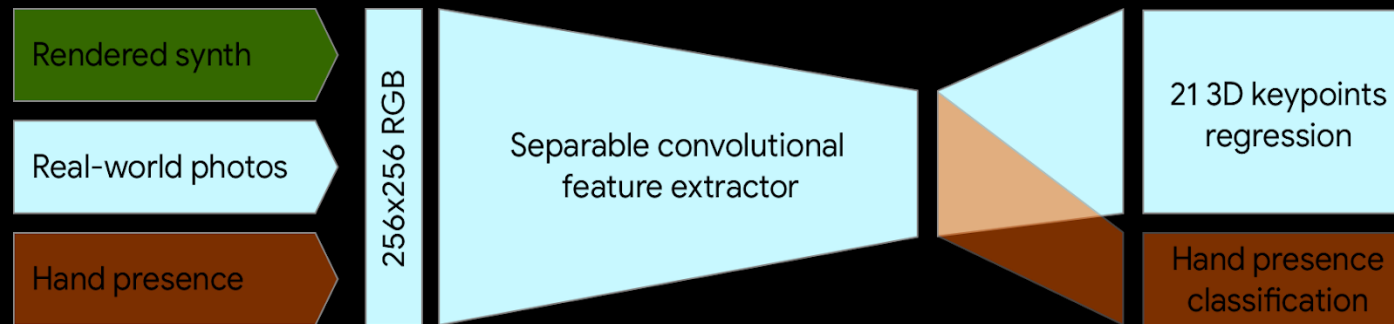
- PYTHON
- MEDIAPIPE
- OPENCV2
- NUMPY
- TENSORFLOW
- SKLEARN
- CSV

IMPLEMENTATION AND WORKING

To obtain ground truth data, ~30K real-world images with 21 3D coordinates are manually annotated, as shown below (taken Z-value from image depth map, if it exists per corresponding coordinate). To better cover the possible hand poses and provide additional supervision on the nature of hand geometry, we also render a high-quality synthetic hand model over various backgrounds and map it to the corresponding 3D coordinates.



IMPLEMENTATION AND WORKING

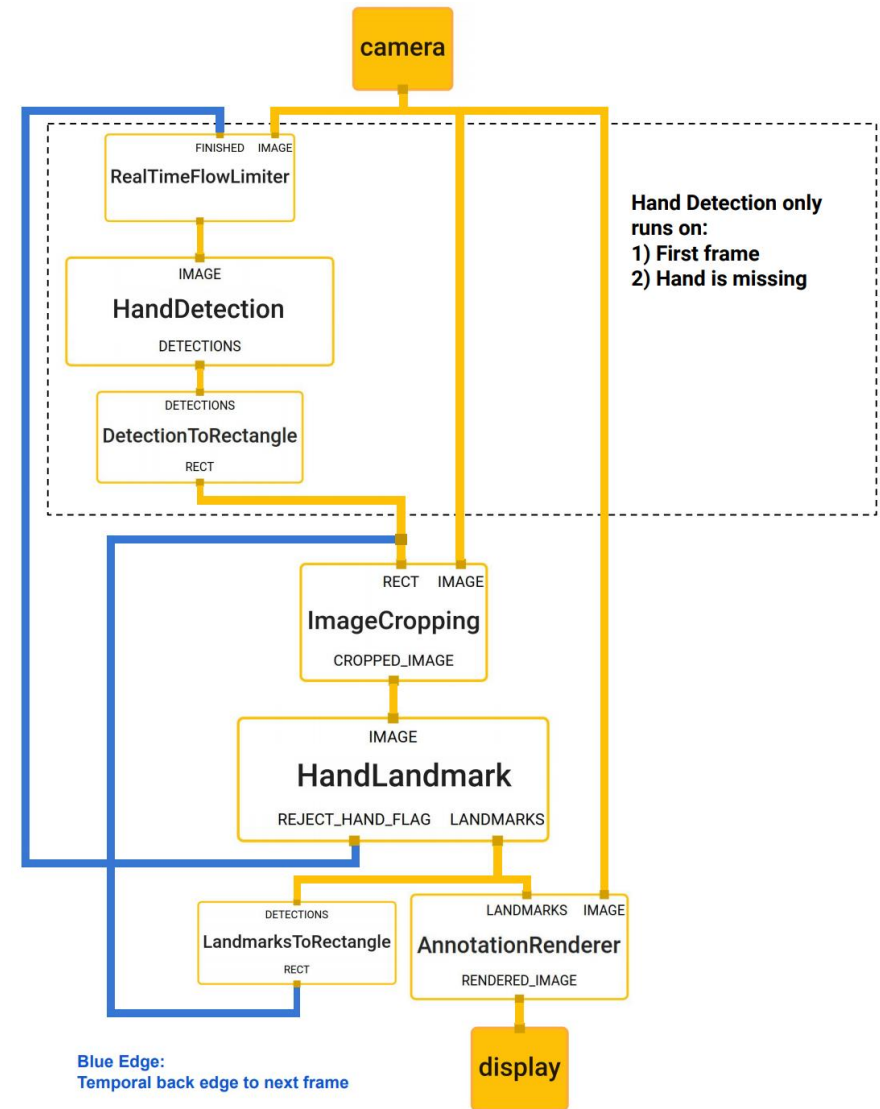


The table summarizes regression accuracy depending on the nature of the training data. Using both synthetic and real world data results in a significant performance boost.

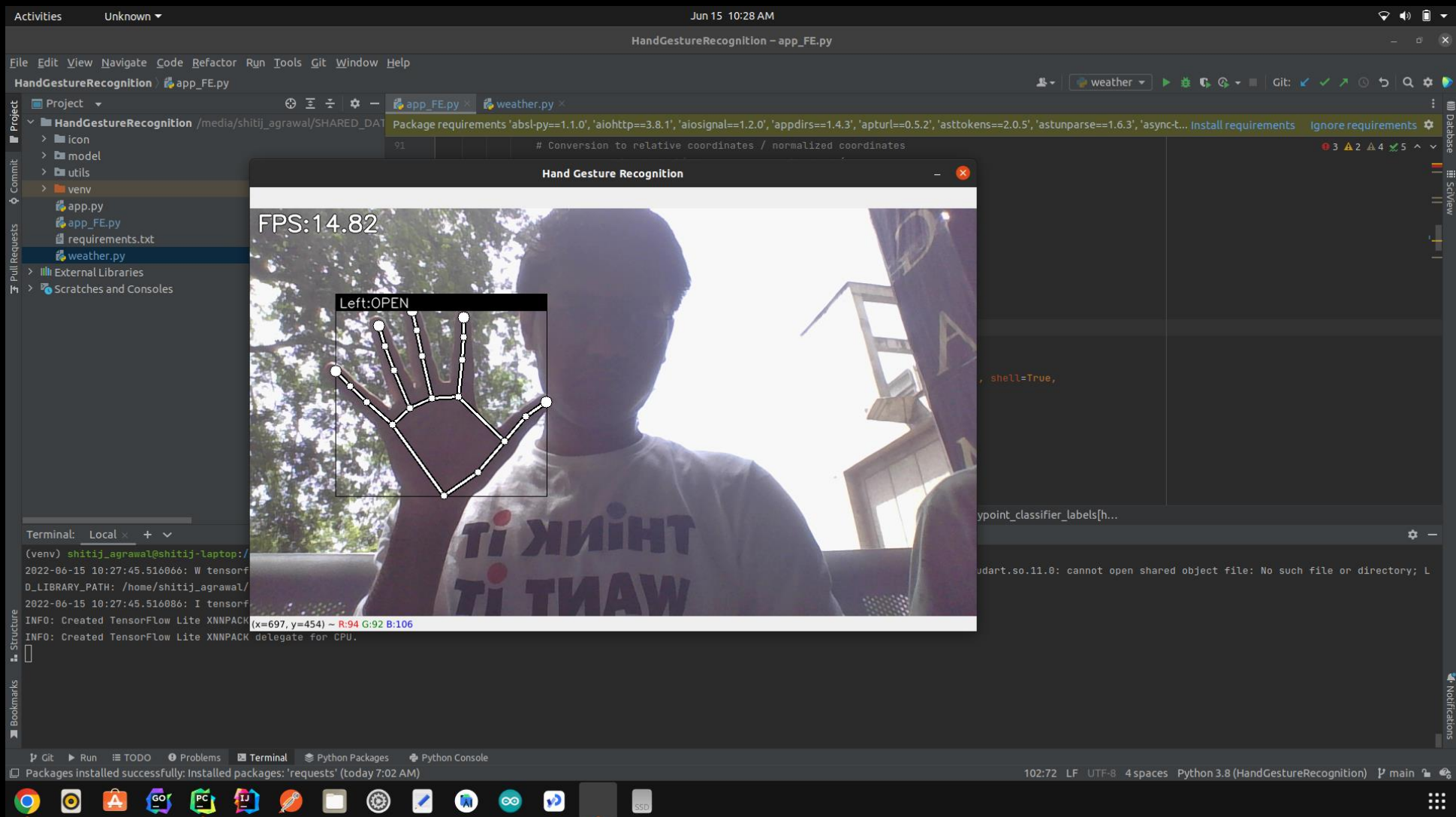
Dataset	Mean regression error normalized by palm size
Only real-world	16.1 %
Only rendered synthetic	25.7 %
Mixed real-world + synthetic	13.4 %

IMPLEMENTATION AND WORKING

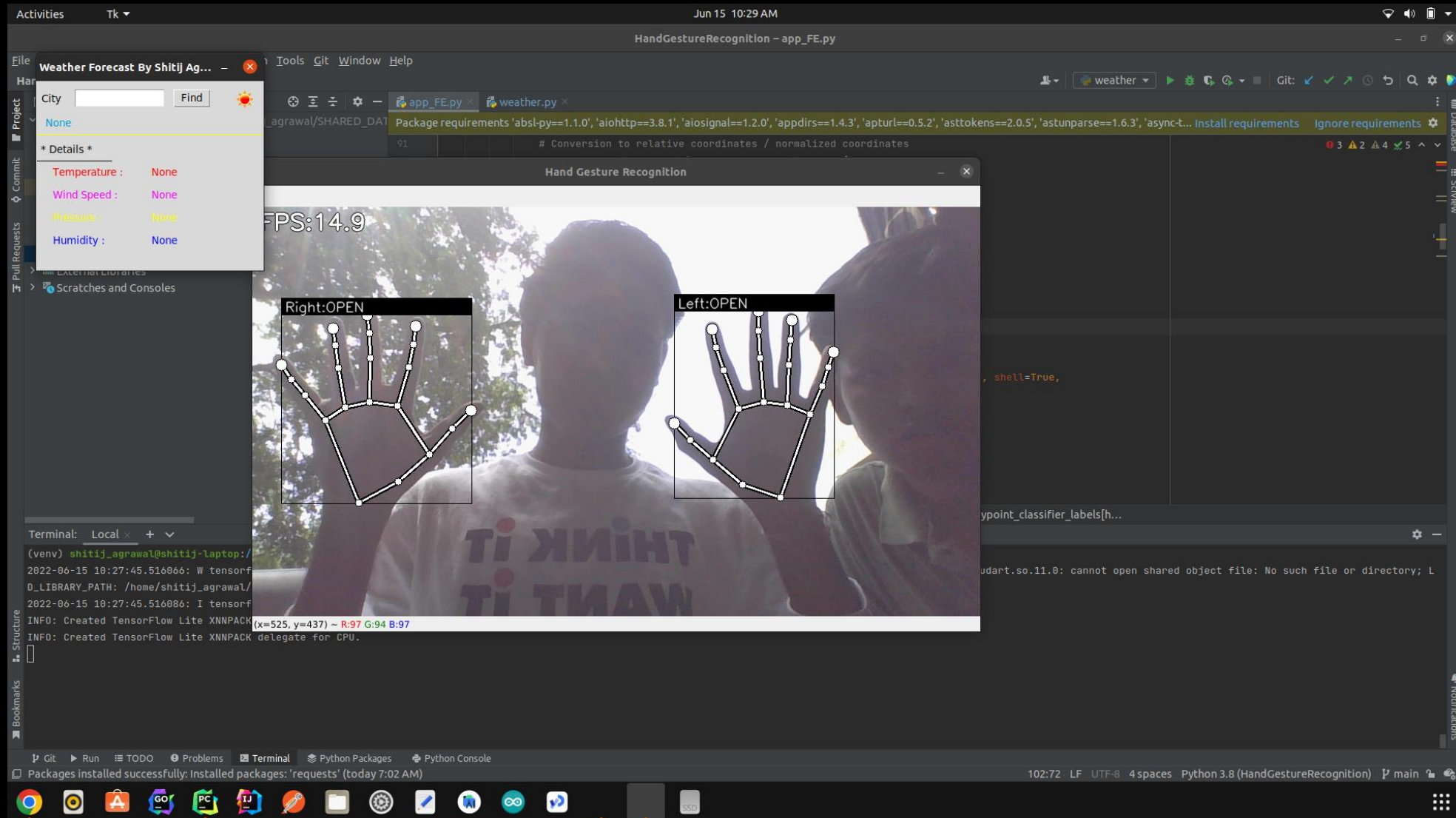
The graph consists of two subgraphs—one for hand detection and one for hand keypoints (i.e., landmark) computation. One key optimization MediaPipe provides is that the palm detector is only run as necessary (fairly infrequently), saving significant computation time. We achieve this by inferring the hand location in the subsequent video frames from the computed hand key points in the current frame, eliminating the need to run the palm detector over each frame. For robustness, the hand tracker model outputs an additional scalar capturing the confidence that a hand is present and reasonably aligned in the input crop. Only when the confidence falls below a certain threshold is the hand detection model reapplied to the whole frame.



OUTPUT SCREEN SHORTS



OUTPUT SCREEN SHORTS



CONCLUSION

- THIS PROJECT CAN HELP TO MAKE LOTS OF APPLICATIONS RELATED TO THE HAND GESTURE
- THIS IS A PORTABLE AND REUSABLE MODULE WHICH CAN BE EASILY REUSED IN MANY DIFFERENT PROJECTS