**CHAPTER - 3 Proposed Work**

The system which we are trying to make will be able to closely replicate all the functionalities of an actual mouse. The user should be able to control the mouse just by making a few gestures on their hand. These gestures would be easy to make and memorize. All this would be done without making any physical contact with the mouse pointer devices. The mouse functions which we plan to demonstrate are as follows.

1. Mouse Maneuver
2. Left Click
3. Right Click
4. Double Click
5. Scroll Up
6. Scroll Down

The Maneuver functionality will be done with the fist gesture. The center point of the fist would be calculated, and it would be used as the pointer. Movement of the fist will make the mouse move in the direction of the hand movement.

The Left Click would be done with the help of the two-finger gesture. As soon as a ‘Peace’ symbol is shown, it will be detected as a left click. After a left click function, unless a different gesture is not made, a second left click will not be detected. This will help the system from repeatedly clicking the mouse when the two fingers are shown up.

The Right Click would be done with the palm gesture. If the palm is shown, the system will detect it as a right click. Similar to the left click, unless a change of gesture is not recognized, the system would not register a second right click.

The Double Click function is recognized using the three-finger symbol. Since it is not a very commonly used function, the three-finger gesture has been assigned to it.

The Scroll Down functionality is recognized using the swing gesture. Whenever the user makes a swing gesture, the screen will scroll down 10 clicks at once. If the user does not change his hand gesture, then the screen will continue to keep scrolling until an end is reached or a different gesture is made.

The Scroll Up functionality is done with the swing and index finger combo. This gesture will move the screen up 10 clicks. The scrolling speed is always constant.

**CHAPTER 4 – Implementation**

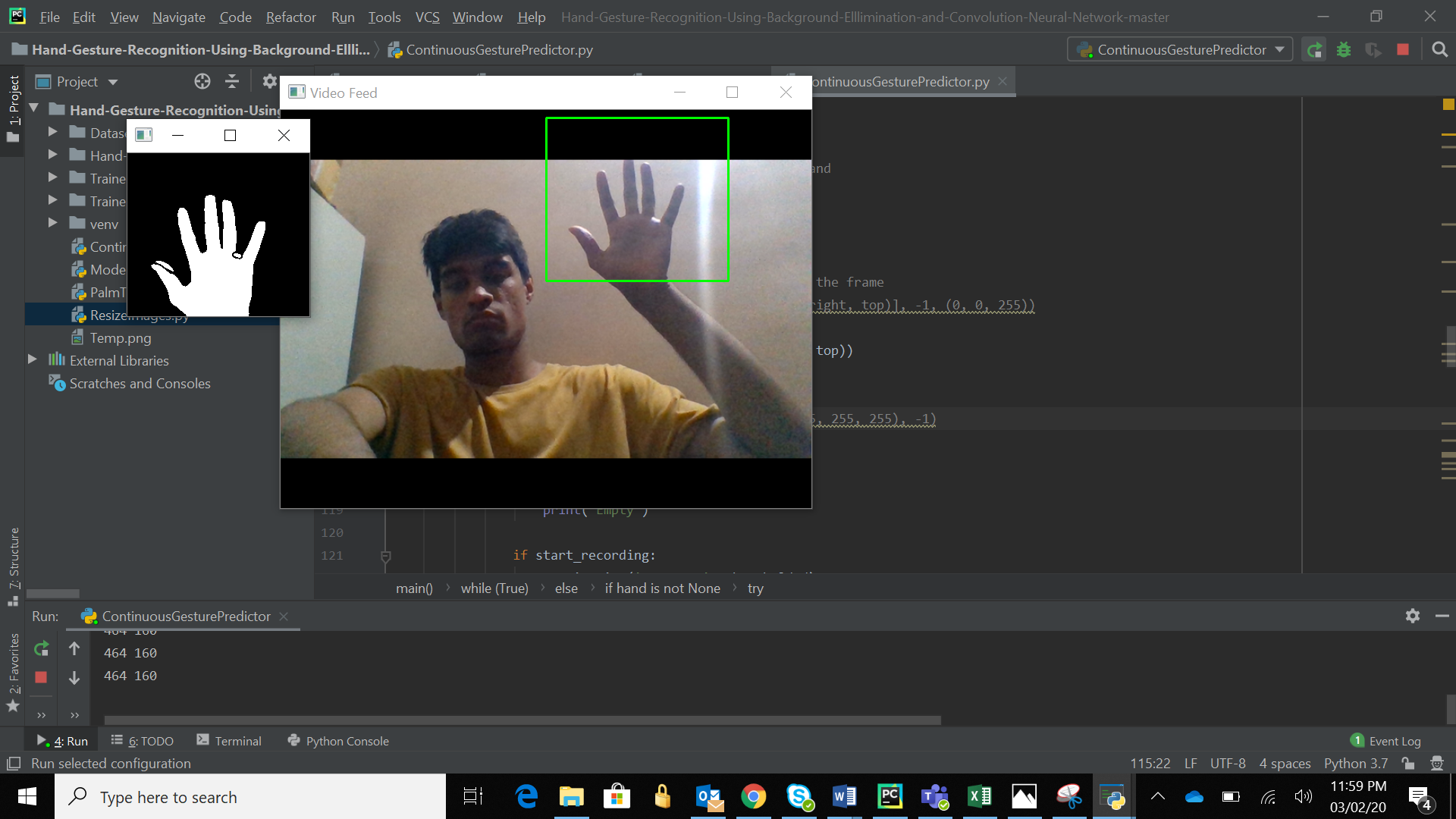
The project is implemented in multiple stages. Those stages are as follows.

* Capture the real time video.
* Extract the individual frames.
* Resize the frame.
* Flip the frames.
* Get the ROI (Region of Interest).
* Convert the ROI to grayscale.
* Blur the ROI.
* Eliminate the background.
* Predict the gesture.
* Draw the counters and calculate the center point.
* Track pointer according to the centroid movement.
* Perform different actions with the different colours extracted.

## **Capturing the real time video**

For the system to sense the movement of hand we need a sensor. We used the web camera in our work for that sensing. The web camera captures the real time video that is it gives us the real time news feed. The captured image is at a particular resolution and at a particular frame rate. We can change the frame rate and resolution according to our needs.

The web camera captures the real time video. This is divided into different frames. These frames are then processed further.

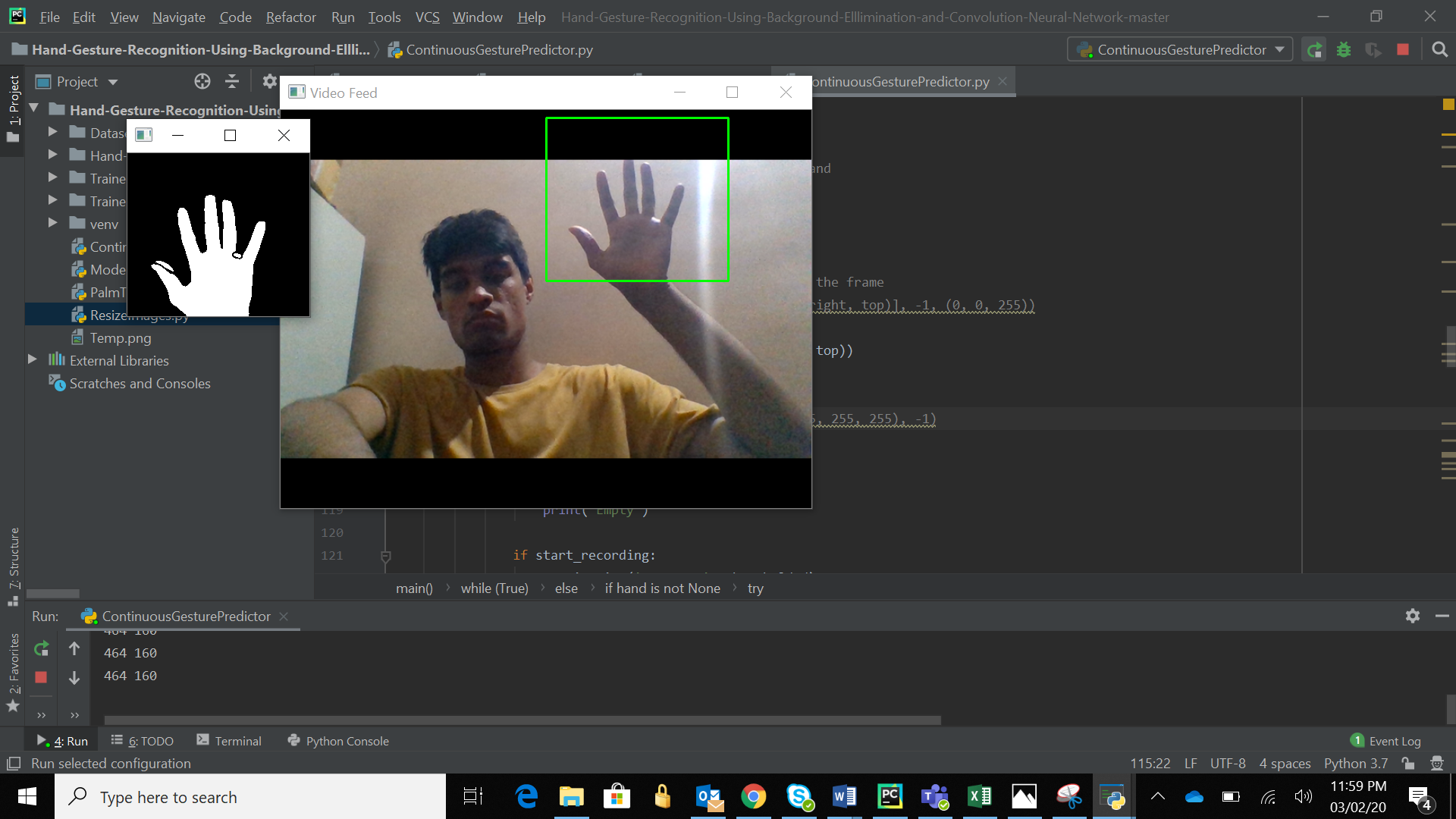


## **Resizing the image**

When the camera captures the real time feed and converts it into the frames then the frame obtained might be of different size. So, we need to resize the image to fit our need. Now this image is inverted. It is like the image we get when we stand in front of the mirror. So, if we move our coloured strip to right it will in turn move to left and so will our mouse pointer. This will create a lot of confusion in operating the mouse pointer. Hence, we need to flip each frame. This is done by vertically inverting the frame.

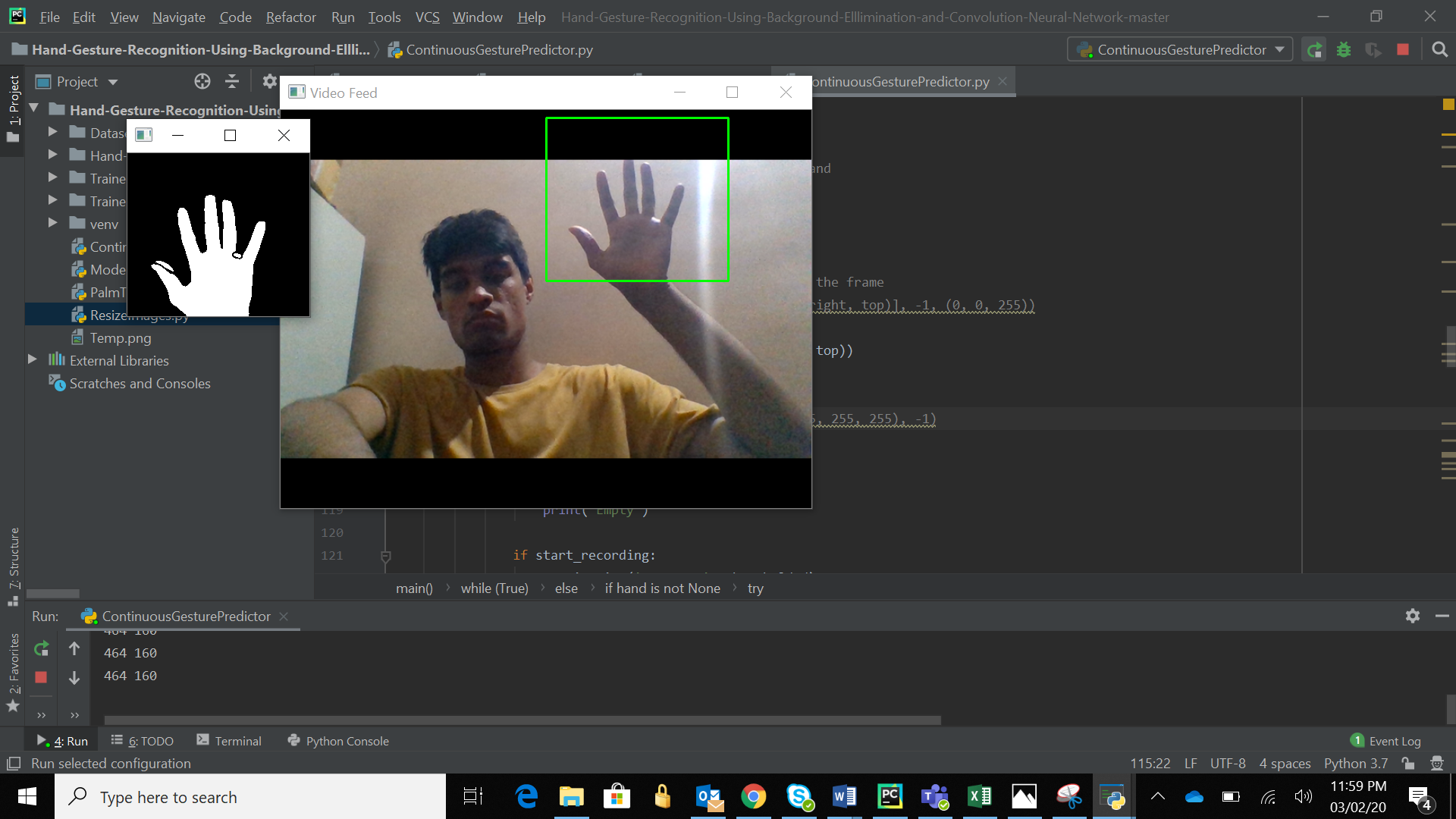
## **Flipping the Image**

Now this image is inverted. It is like the image we get when we stand in front of the mirror. So, if we move our coloured strip to right it will in turn move to left and so will our mouse pointer. This will create a lot of confusion in operating the mouse pointer. Hence, we need to flip each frame. This is done by vertically inverting the frame.



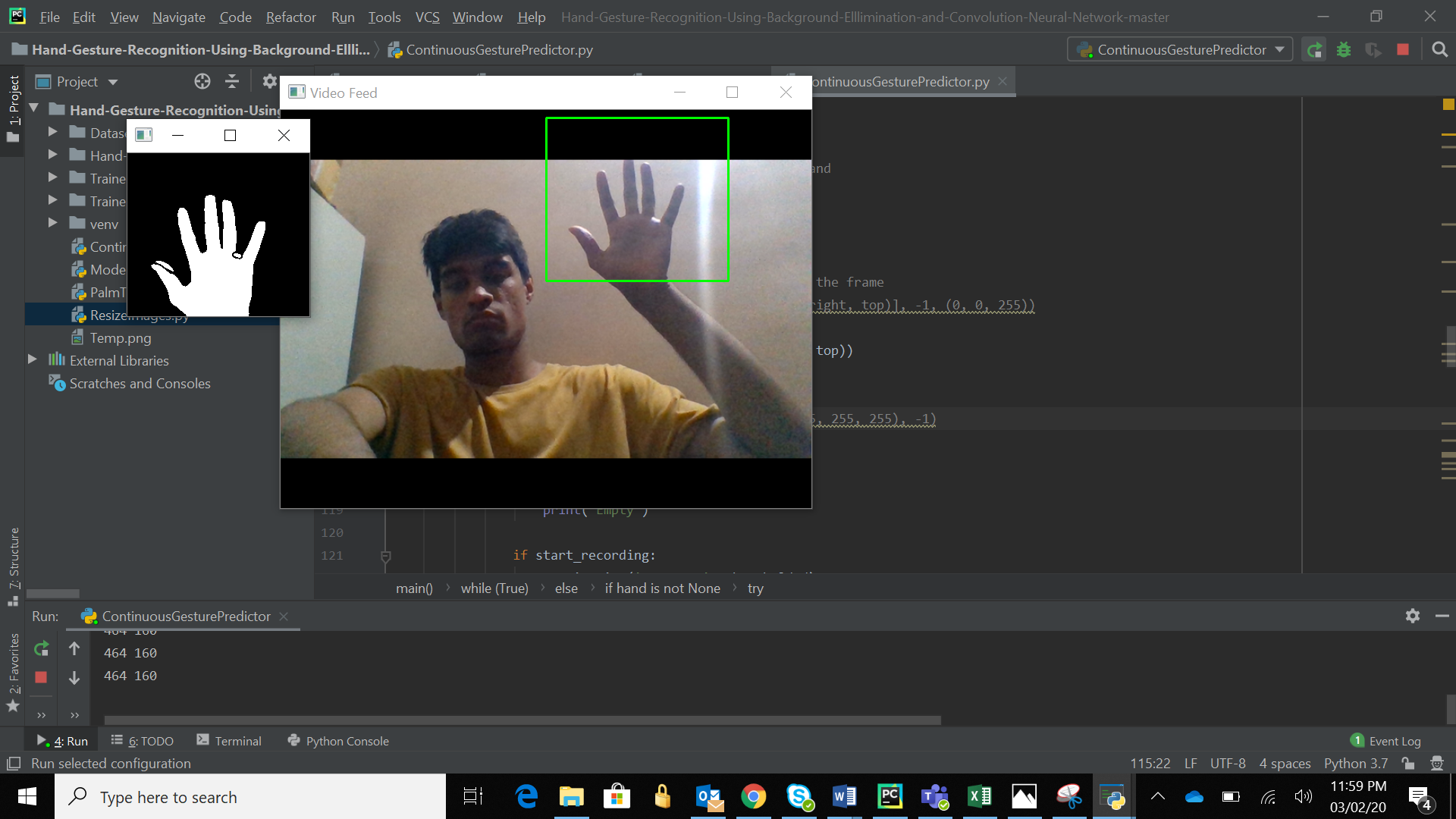
## **Getting the Region of Interest**

The Region of Interest is a small part of the entire frame which is the actual area where all the processing will be done. The remaining part of the frame is ignored. Only the image inside this region is processed. This is done to remove any unnecessary objects in the background which can cause noise.



## **Eliminate the Background**

In this step, all the background images are removed and only the image of the hand is shown. This is important as based on this image, the different gestures made by the hand will be predicted. The hand is shown in white colour and the remaining background is shown as black colour.



## **Predict the gesture**

The image then passes through our convolutional neural network. The gesture is predicted on basis of how the neural network is trained previously. This is an important step as a major part of the accuracy of our system depends on it.

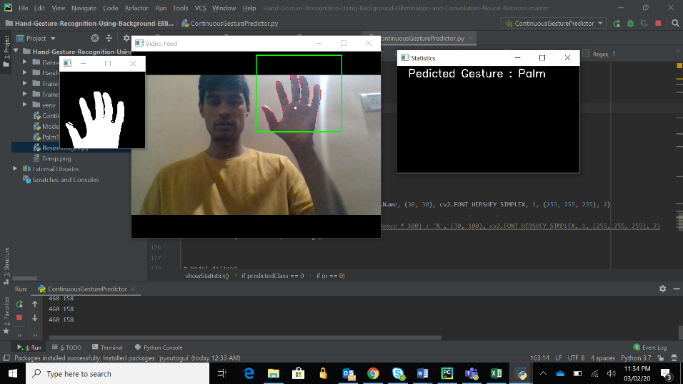


Figure 1 Palm

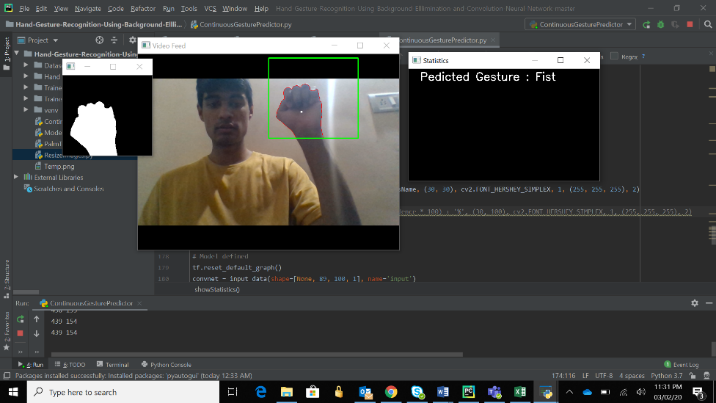


Figure 2 Fist

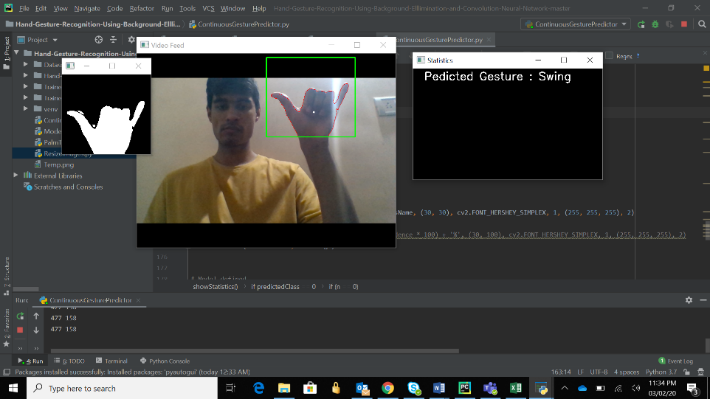


Figure 3 Swing

## **Draw the contours and its centroid**

The contours are drawn around the hand gesture. Contour is the outer line around a certain area. After this we calculate the centroid. It is the center point of the area surrounded by the contour. We control our mouse pointer movement with respect to the calculated centroid.

## **Tracking and performing different actions**

The tracking and different actions performed are done using the inbuilt library in python named PyAutoGUI. This library has many functions for left click, right click or just moving the pointer and many more.