GURUSHIYA SP

Project Plan

SIH - NR1167 (DRDO)

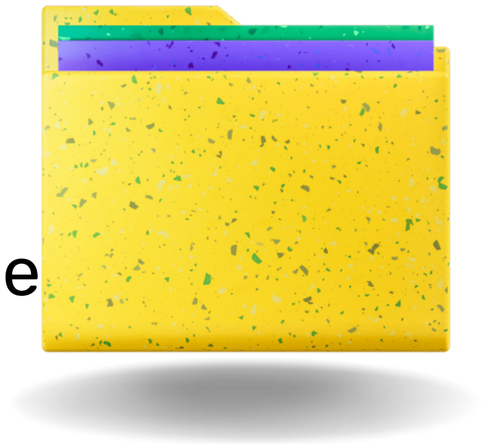
Gesture Enabled Commands for Operating Laptop or PC

Problem Statement

"Developing a solution for Gesture enabled commands for operating Laptops/PCs for frequently used operations on daily basis" has been listed in the official Smart India Hackathon (SIH) portal by DRDO.

Few gestures could be defined for frequently used tasks- save, exit, print, screen-lock, screen unlock, system shut down, system restart. Save, print and exit operations are context sensitive meaning that it is applicable for current application. For example if word document is open and the gesture for save is done then the document will saved, if print gesture is done then printer dialog will open etc. Similarly a gesture could be defined for close/exit which will close the current application. If no application is opened then it will work as system shut down. It is similar to Alt+F4 key press functionality on windows PC.

Overview



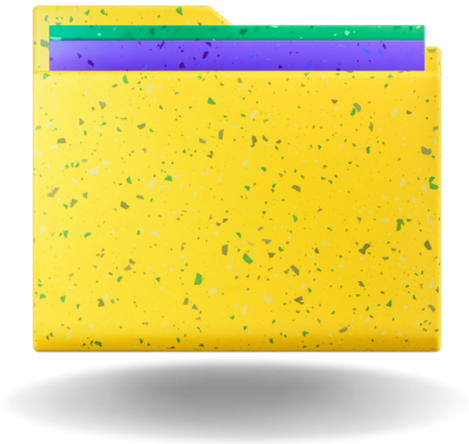
Humans operate computers basically with the help of mouse and keyboard.

Apart from these two methods, computers can also be operated using physical hand.

This project aims to develop a solution for Gesture enabled commands to operate Laptops/PCs for frequently used operations.

Back to Agenda Page

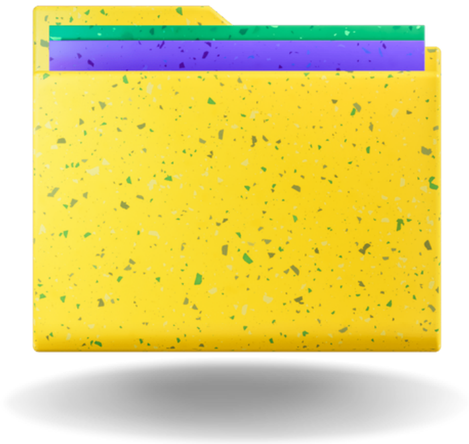
Media Pipe



MediaPipe is a customizable machine learning solutions framework developed by Google. It is an open-source and cross-platform framework, and it is very lightweight. MediaPipe comes with some pre-trained ML solutions such as face detection, pose estimation, hand recognition, object detection, etc.

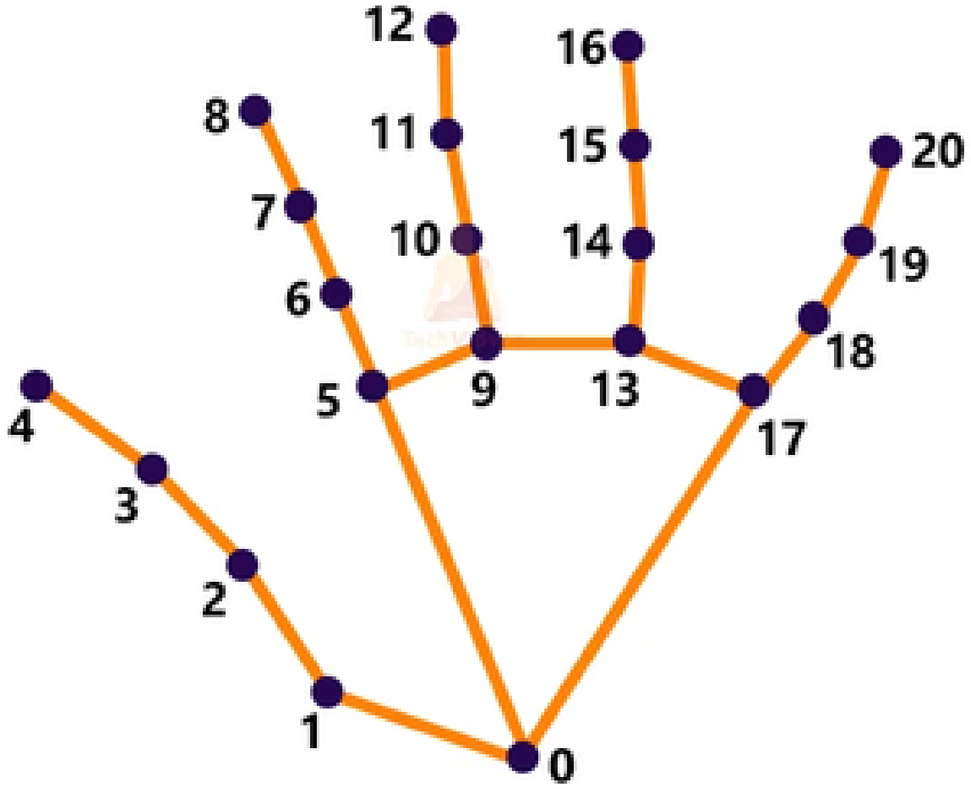
Back to Agenda Page

Media Pipe Hands



MediaPipe Hands is a machine learning solution developed by Google that allows for real-time hand tracking and gesture recognition in videos. It uses a deep neural network to detect 21 3D landmarks of the hand, including the fingertips, knuckles, and wrist, in each frame of a video. This technology can be used for a variety of applications, such as augmented reality, sign language recognition, and hand gesture-based control systems. MediaPipe Hands is available as a pre-built TensorFlow Lite model, making it easy to integrate into mobile and web applications.

21 Key Points



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

10 different gestures that the model recognises

Okay

Peace

Thumbs Up

Thumbs Down

Call me

Stop

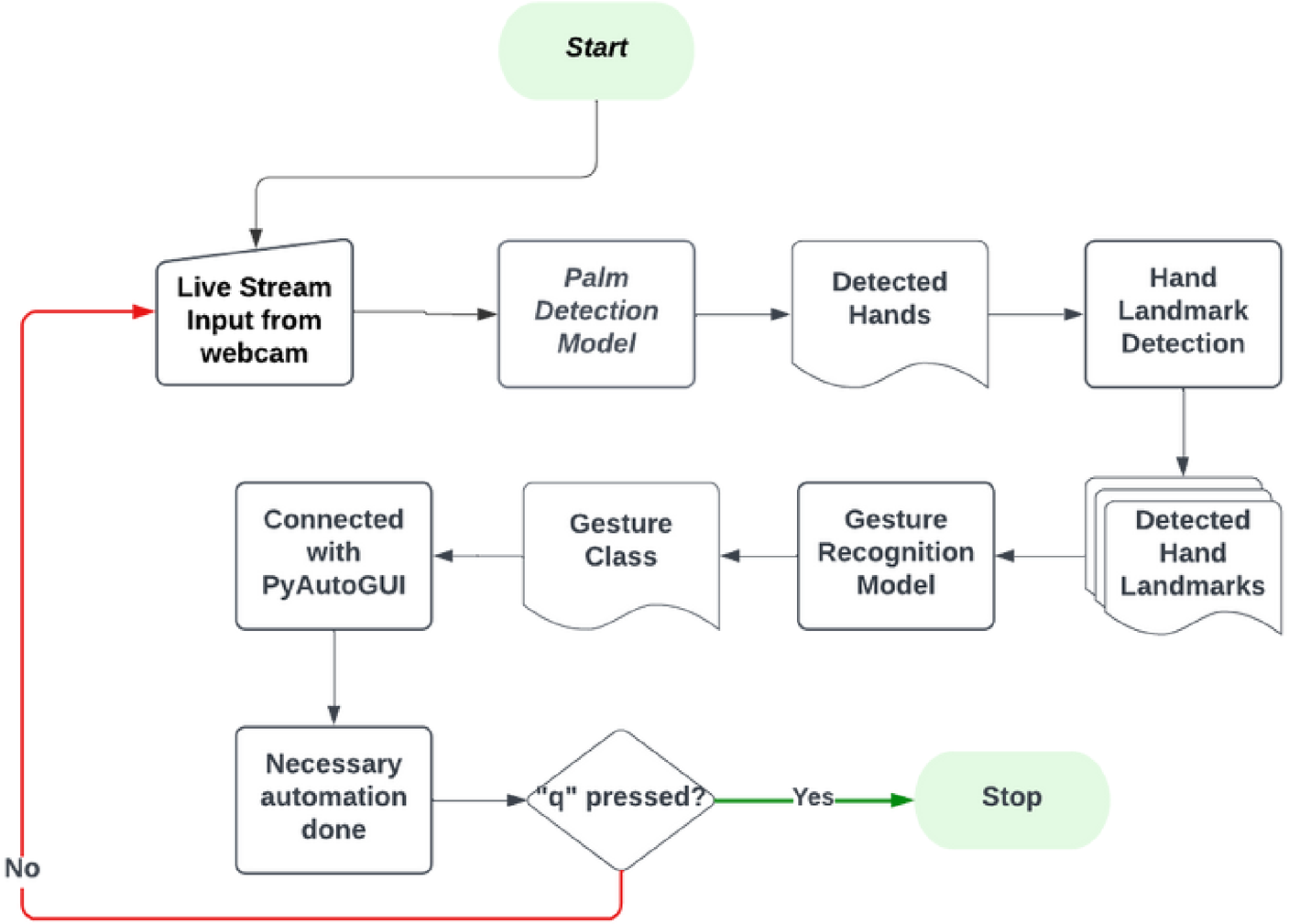
Rock

Live Long

Fist

Smile

Process Flow



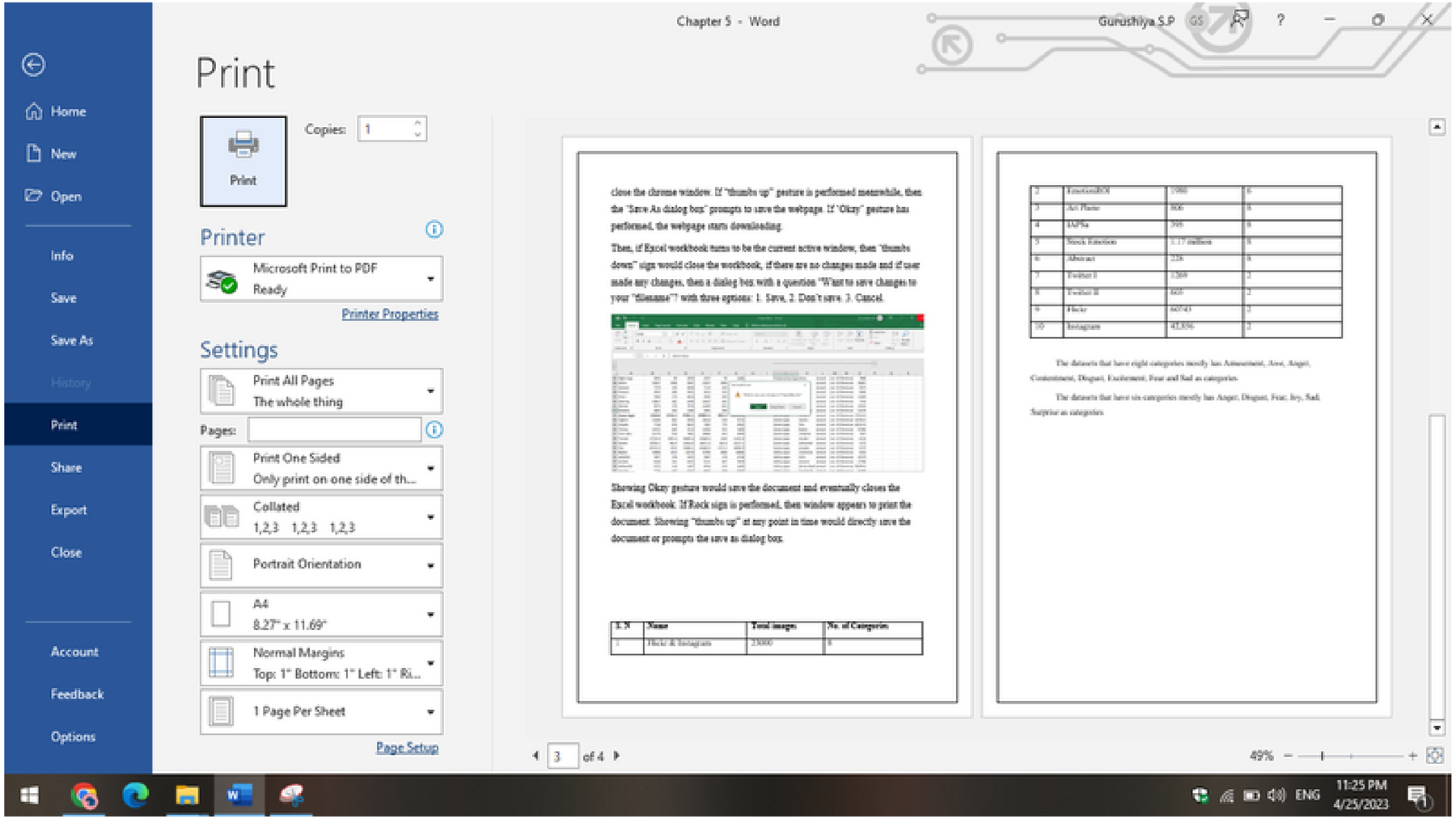
Operations and respective gestures

1. Okay - Enter  Class 0
2. Rock - Print (ctrl + P)  Class 6
3. Thumbs Up - Save (Ctrl + s)  Class 2
4. Thumbs Down - Close (Alt + F4)  Class 3
5. Stop - Close current tab (ctrl + w)  Class 5

Connect class_ID with PyAutoGUI

```
classID = np.argmax(prediction)
if (classID == 2):
    pyautogui.hotkey('ctrl', 's')
elif (classID == 3):
    pyautogui.hotkey('Alt', 'f4')
elif (classID == 0):
    pyautogui.press('enter')
elif (classID == 6):
    pyautogui.hotkey('ctrl', 'p')
elif (classID == 5):
    pyautogui.hotkey('ctrl', 'w')
```

Rock sign in a word doc



Operations and respective gestures specific to media player

1. One - Right Arrow (Frwd 5 seconds)

2. Two - Left Arrow (Bkwd 5 seconds)

3. Three - Up Arrow (Raise volume)

4. Four - Down Arrow (Low volume)

5. Five - Space bar (Pause or Play)

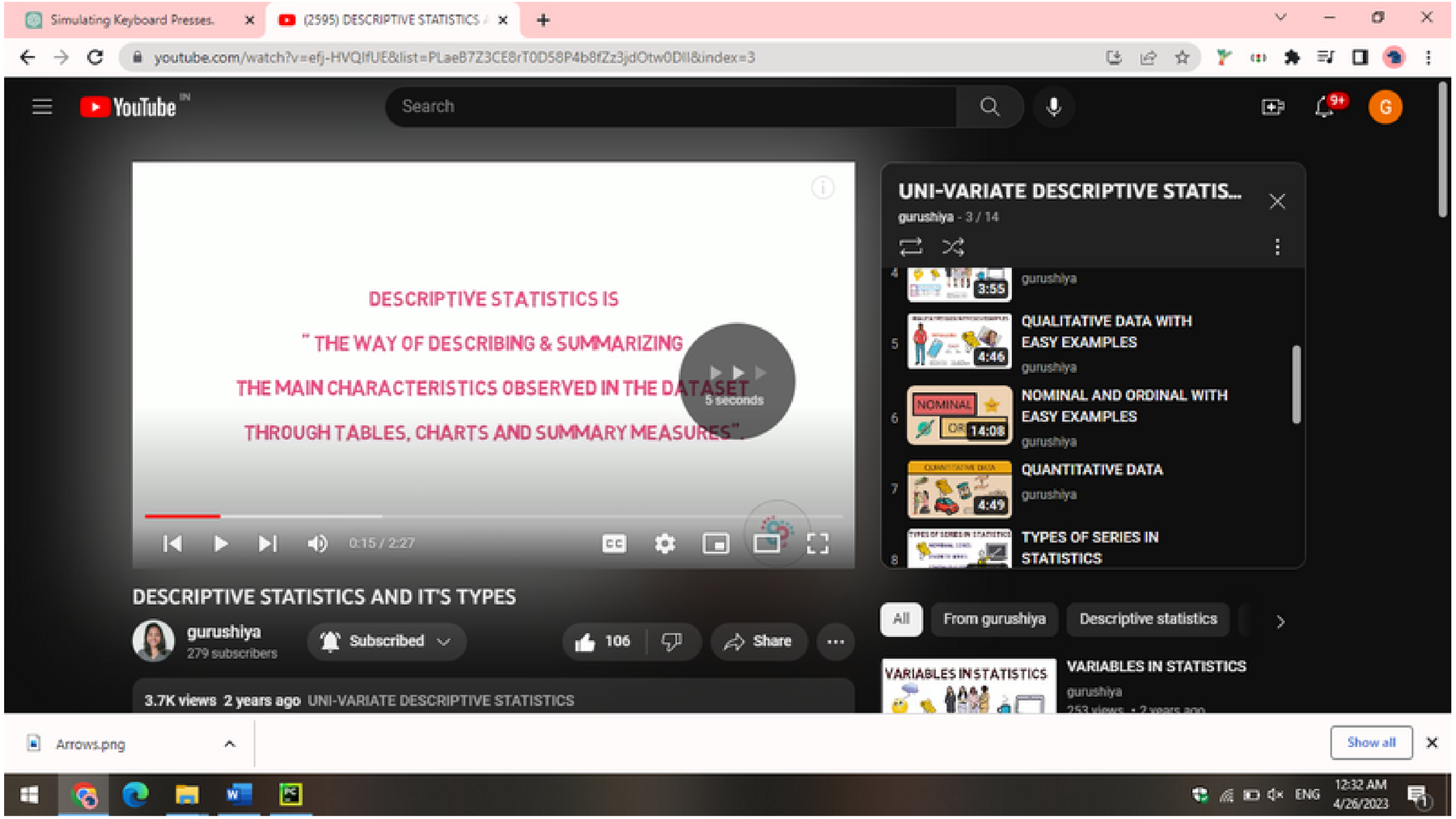
Connect finger counts with PyAutoGUI

```
if (cnt == 1):
    pyautogui.press("right")
elif (cnt == 2):
    pyautogui.press("left")
elif (cnt == 3):
    pyautogui.press("up")
elif (cnt == 4):
    pyautogui.press("down")
elif (cnt == 5):
    pyautogui.press("space")
```

Calculate number of fingers

```
def count_fingers(lst):
    cnt = 0
    thresh = (lst.landmark[0].y * 100 - lst.landmark[9].y * 100) / 2
    if (lst.landmark[5].y * 100 - lst.landmark[8].y * 100) > thresh:
        cnt += 1
    if (lst.landmark[9].y * 100 - lst.landmark[12].y * 100) > thresh:
        cnt += 1
    if (lst.landmark[13].y * 100 - lst.landmark[16].y * 100) > thresh:
        cnt += 1
    if (lst.landmark[17].y * 100 - lst.landmark[20].y * 100) > thresh:
        cnt += 1
    if (lst.landmark[5].x * 100 - lst.landmark[4].x * 100) > 6:
        cnt += 1
    return cnt
```

Showing one – right arrow



GURUSHIYA SP

Thank You

Back to Agenda Page