**Face detection and pose estimation:**

**Objective:** To implement a script that can detect faces and its landmarks. Using those landmarks, determine the head pose of the face under focus.

**Face detection:**

For this, I used a SSD model. The face under focus will be highlighted with a green box while the other faces in the video will be highlighted with red boxes. The face that has the biggest bounding box is considered the face under focus.

**Landmark detection:**

In this project, I have used eyes, nose, mouth and chin as landmarks. Although, it would be more useful to use ears as one of the landmark as it could help with determining the head pose. Unfortunately, I was not able to implement that. That is one of the limitation of the script.

**Head pose estimation:**

I detected the head angles using the OpenCV functions to derive the translation and rotation vector. However, it is not extremely accurate. While it is capable of detecting right and left turns, it struggles a bit to determine the up and down movement of the head.

The angles are indicated at the nose tip. The theta value which indicates whether the head Is looking up or down is shown in green and phi value which indicates whether the head Is looking right, or left is shown in red. Their values are also stored in a list and can be plotted when the video ends or when the user presses the enter button.

If theta value is above 50, then the head is looking down and if theta value is below -50, then the head is looking up.

If the phi value is above 50, then the head is turned right and if the phi value is below -50, then the head is turned left.

**Note: I have tested the code only in recorded video and not on live webcam due to technical difficulties on my side. I have included the code that automatically uses webcam and the line that takes in recorded video in line 19 and line 20. Please feel free to switch between both.**

**Additional problems:**

1. If the head goes out of focus and comes back in, will your algorithm start tracking again?

Yes, the algorithm will start tracking when the face gets back in to the frame

1. How are you prioritizing tracking for objects in focus?

I am only predicting the landmarks and detecting the head pose for the face in focus. While the other faces will be marked by a bounding box, I don’t predict the head pose for those faces.

1. Are you able to track with foreign objects like headphones, spectacles, etc?

Yes, I am able track with headphones and spectacles. However, the tracking is not good for faces with mask.

1. If there is occlusion (example: if someone is drinking from a bottle) will the tracking stop?

Yes, when the person lifts their head when drinking from a bottle, there is threshold above which it stops tracking. More rigorous training and landmark points such as ears and neck are required to avoid this.

1. Anything else that you think will be important for a real-world application but might not have been mentioned above. Explain why it is important

I have tried using the script for faces with masks. Further improvement is needed to make It track faces with mask.