



Indian Institute of Science Education and Research  
Bhopal  
**Computer Vision( DSE/EECS-312)**  
Bonus Assignment-1

Deadline: 14-10-2024, 11:59 PM

Max mark: 10

Please follow the instructions carefully.

---

1. Plagiarism and copying from anywhere (similar submission) can debar you from this course and invite the academic dishonesty policy.
2. The length of the video should be between 3-5 minutes.
3. The video should not be abusive and/or harmful in any sense to anyone.
4. The purpose is to learn the recent advancements of AI in the Computer vision world only. Therefore, no personal interest will be entertained.

**Submission:**

1. Video
  2. A 1-2 page report only in the given overleaf template explaining the topic of your video, the crux, an explanation of the tool used in both questions and its link.
- 
1. Using recent advancements in generative AI and computer vision technologies, create a simulated debate video featuring two prominent historical or contemporary figures. The debate should center around a critical and timely issue, such as climate change, technological ethics, international relations, or social justice. Additionally, consider the ethical implications and technical challenges of creating realistic AI-generated representations of real individuals.  
  
Here is one example: <https://www.linkedin.com/feed/update/urn:li:activity:7118570023987683328?updateEntityUrn=urn%3A7118570023987683328%29>

**Answer:**

I used several tools to create it. First I generated the script using ChatGPT, then I used Verbo ai to generate a video from the face given. Then I used media.io and fakeyou.com to generate the audio for each individuals. Then I finally combined the videos using video editing tool Davinci Relove. The individuals are Leonel messi and Donald trump. The topic of the debate is "Leading world towards AI". The video has been attached.

2. Explore the various open-source toolboxes and applications used in the field of computer vision and come up with some interesting applications or toolboxes related to the Computer Vision field (Inpainting, Pose Estimation, DeepFakes(Image, Audio or Video), Adversarial Example Generation, Style Transfer). Select any toolboxes or applications and use them to test their effectiveness in solving the specific problem you are addressing, and discuss their core functionalities, advantages, and real-world applications.

### Answer:

Recently, I was working on the project of ai gym trainer. So what you do is basically, record your video or turn on your camera for live rep counting. So follow is the idea:

The problem statement is simple, given a video, I have to count the number of reps user has performed and also check if the form of the user is correct or not. I have to do this in real time too, since I want to support video to from live video feed.

### Mediapipe

I discovered this library, which is developed and maintained by Google. I used it to extract. To discuss on some of the core functionalities of mediapipe library, below image describes number of features that this library has.

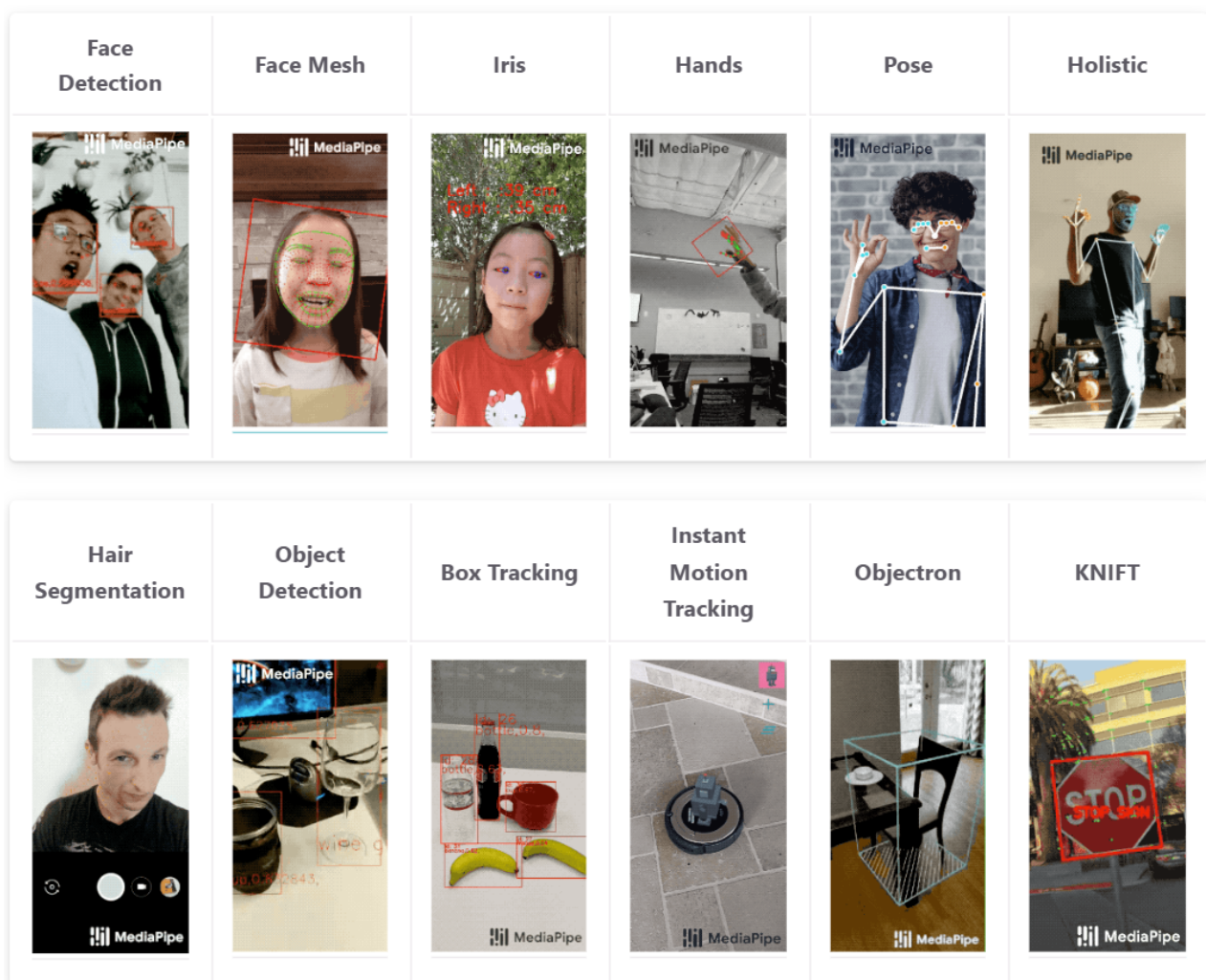


Figure 1: Mediapipe features

## mediapipe

I used it to extract skeleton of the human body currently in frame. This is a pose estimation problem

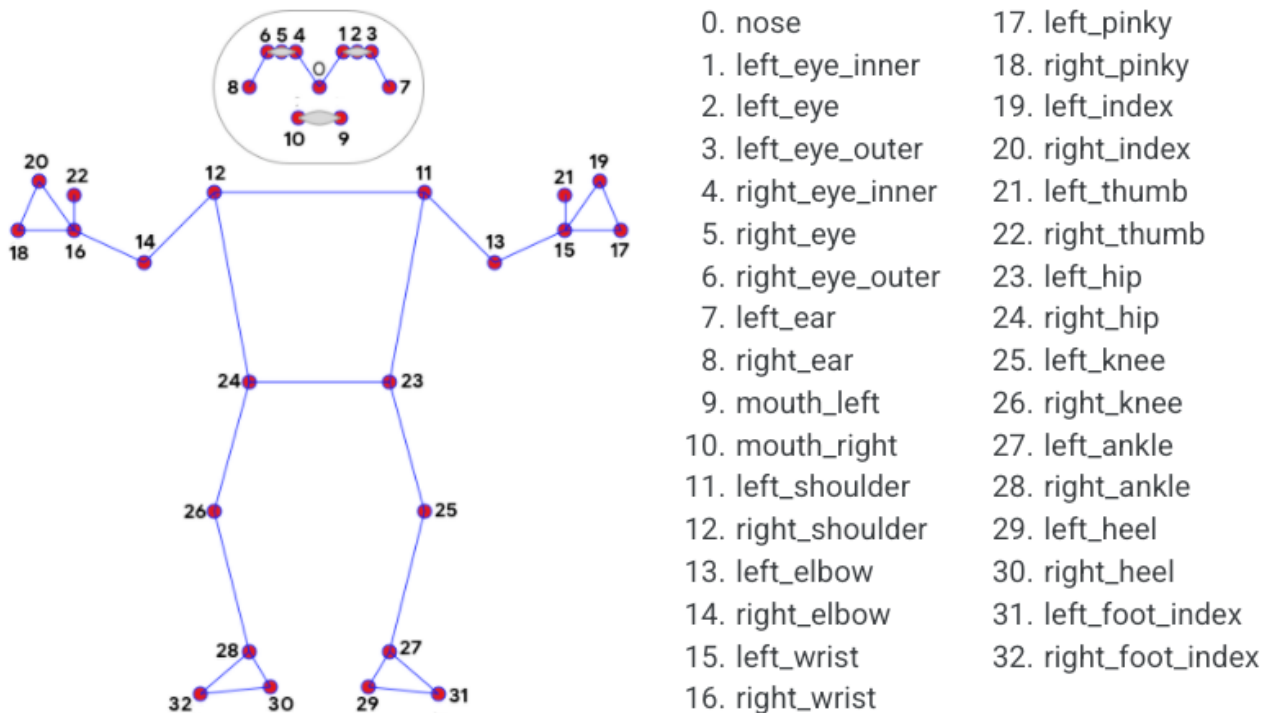


Figure 2: Pose tracking full body landmarks

I then used the extracted skeleton to extract the angles of the joints. I used these angles to figure if the angles cross certain threshold to count the rep. I also used the angles to figure out if the form of the person is correct or not.

I have added a screenshot of working of that project.

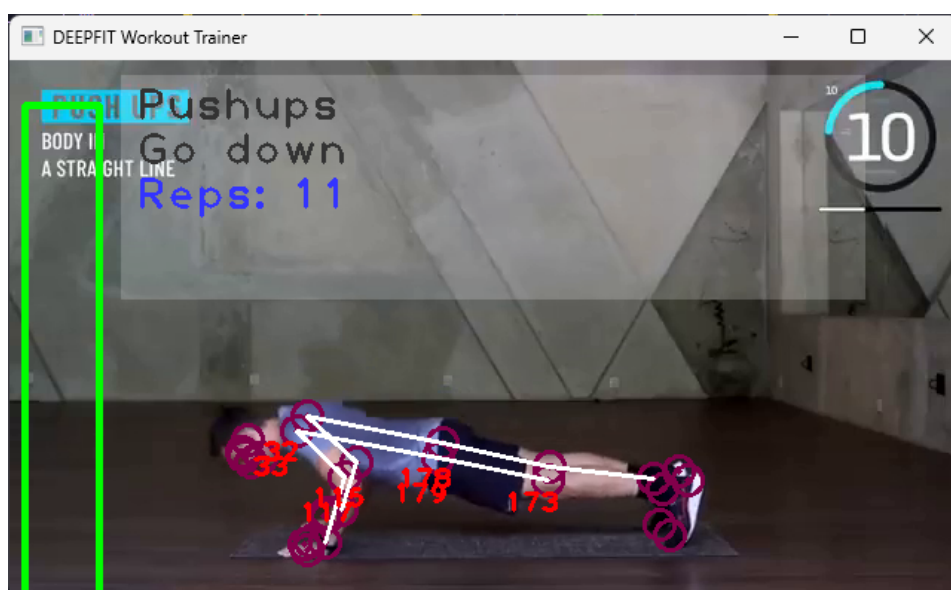


Figure 3: AI gym trainer app