Group 8 Iteration 2

For developing model which recognizes the human body as we specified in the earlier Iterations we are using Media pipe Library and for deploying in Web RTC we are planning to use Stream lit, a python library for easy of use and effective visual communication.

Mediapipe is a framework for machine learning that offers a broad range of solutions for handling multimedia data. It has the ability to estimate 3D human poses, which can be utilized to create a model of a person in a three-dimensional space.

Streamlit is a free Python library that allows developers to construct customized web applications for data science and machine learning. Its user-friendly interface makes it simple to create interactive interfaces that can be deployed on the web.

WebRTC is a collection of protocols and APIs used to facilitate real-time communication through the internet. It is often utilized to create audio and video communication applications that allow for peer-to-peer connections.

Steps:

1. Begin by installing the necessary libraries, which include Mediapipe, Streamlit, and WebRTC.
2. Create a 3D human pose estimation model using Mediapipe that can be used to model a person in a three-dimensional space.
3. Develop a Streamlit application that accepts a video stream as input and leverages the Mediapipe model to estimate the pose of the human body in the video.
4. Incorporate WebRTC into the Streamlit application to enable real-time communication between users.
5. Finally, deploy the application on a cloud platform or server to make it accessible to users over the web.

Mediapipe offers several machine learning models for different applications, including 3D human modeling. Mediapipe's 3D human pose estimation model, in particular, can be utilized to assess a person's 3D pose in real-time from a 2D video stream, like a feed from a webcam.

The model's architecture is composed of a convolutional neural network (CNN) and a temporal model that can gauge the 3D joint positions of a person's body, encompassing the head, torso, arms, and legs. When given a 2D image or video frame, the model returns the 3D coordinates of each joint in the image.

Once the 3D joint positions are estimated, they can be employed to develop a 3D model of the person in real-time, which can be visualized in a variety of ways, such as through a 3D graphics engine or projected onto a virtual environment. This opens up various potential applications, including virtual try-on, gaming, and augmented reality experiences.