Research Proposal: Lip Reading in English Using Machine Learning and Computer Vision

## Title:

Lip Reading in English Using Machine Learning and Computer Vision

## II. Introduction:

This research project aims to develop a lip reading system for the English language by leveraging machine learning and computer vision techniques. The primary objective is to enhance communication accessibility for the hearing impaired and explore applications in transcription, human-computer interaction, and security.g valuable insights for enhancing user experiences and human-avatar interactions.

## III. Research Question:

How can machine learning and computer vision be used to accurately recognize and transcribe spoken English words from lip movements in real-time?

## IV. Literature Review:

A review of existing literature reveals the importance of human pose estimation and its applications in computer vision, animation, and virtual environments. While individual approaches exist for pose estimation, a comprehensive method for predicting and recreating human behavior is an emerging research area.

## V. Objectives:

Investigate existing pose estimation techniques and their limitations.

Develop a novel approach for predicting and recreating human pose gestures in real-time. (key problem)

Evaluate the accuracy and responsiveness of the virtual human in mimicking human gestures.

Explore potential applications in human-computer interaction and entertainment.

## VI. Methodology:

We propose the following methodology for this research:

### A. Pose Estimation:

Review and implement state-of-the-art pose estimation models (e.g., PoseNet, OpenPose,Mediapipe netControl).

Enhance the accuracy of pose estimation for real-time tracking.

### B. Virtual Human Creation:

Develop a virtual human model capable of replicating human gestures.

Integrate the pose estimation model with the virtual human.

### C. Prediction and Real-time Mimicry:(possibility)

Train the model to predict human pose gestures based on input data.

Implement real-time mimicry of predicted gestures by the virtual human.

### D. Evaluation:

Assess the accuracy of pose prediction and mimicry through user studies.

Measure the responsiveness and naturalness of virtual human interactions.

(I find this research proposal quite extensive, so I'm exploring ways to break it down into manageable components. For instance, I'm considering focusing on hand gesture recognition as one achievable aspect of the larger project )