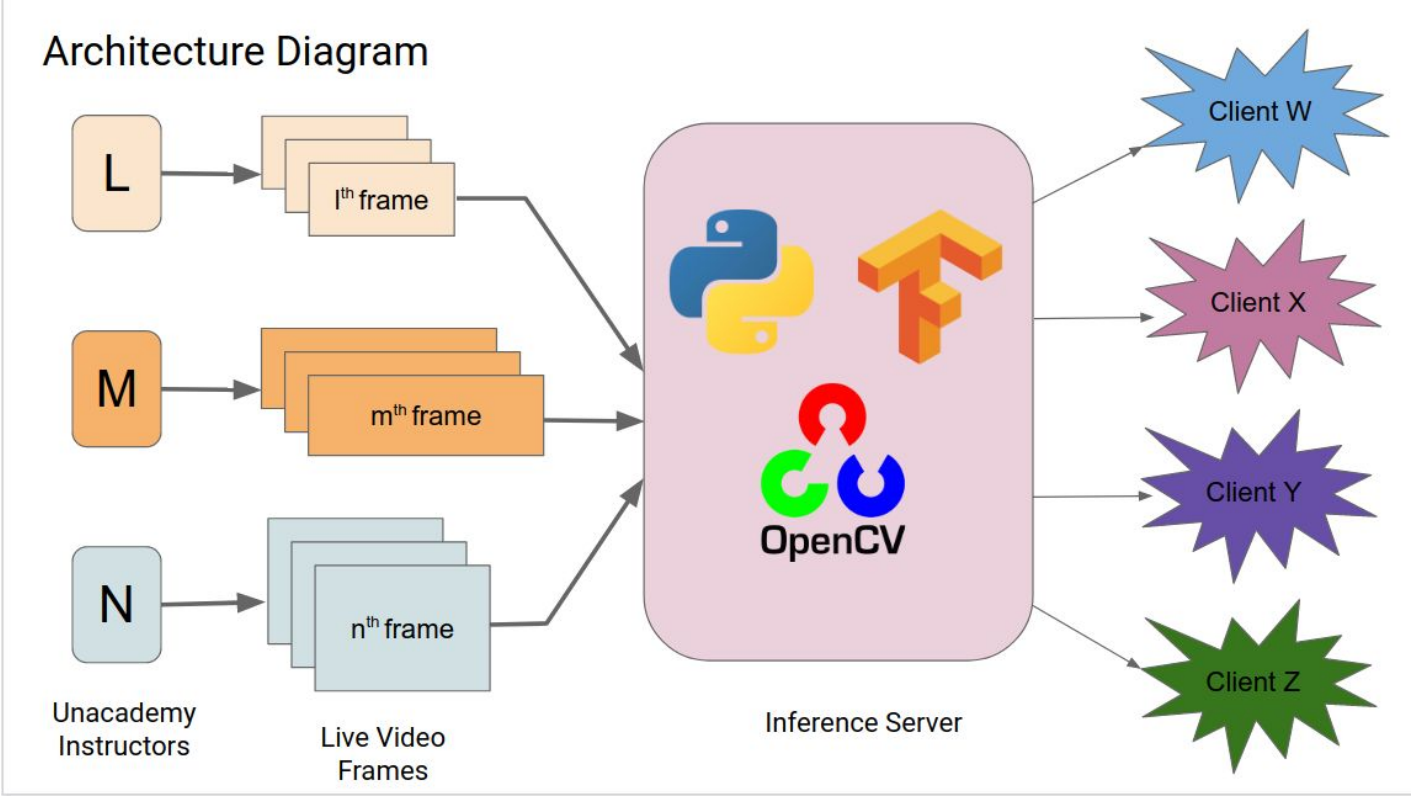
Project Design Document

Project Goals:

1. To remove the monocolor background from instructor videos.
2. To develop a robust algorithm that works in real-time
3. The algorithm will work even when the background consists of real world scenes (curtains, windows, doors, etc)
4. Display the output in a web based interface

Architecture



L, M and N are the Unacademy Instructors as shown in the above figure. They are sending the live video stream to the server. The server takes the frame from each of the instructor and starts to process the frames. The MobileNet V2 model will be used to do the inference once on the

video file. The instructor is mandatory to be present in the first frame sent by the camera to the server. The traditional computer vision technique such as Gaussian Blurring is applied to model so that the background can be blurred. The instructor will be segmented out and will be pasted on the next frame so that the computation is saved. The MobileNet V2 achieves 25 FPS on the live video stream.

Inputs

1. Video stream of the instructor to the inference server.

User Interface

1. Video stream displayed on the client browser.
2. Video is streamed in real-time.
3. The user interacts with the web interface to play the live video

Configuration

Server Config

1. GPU based server
2. Python packages: Tensorflow, Numpy, OpenCV

Client Config

1. Web browser

Data Model

1. Deep Neural Net Models stored on the Inference Server

Conclusion:

1- 25FPS is easily achieved by the model without sacrificing the accuracy.