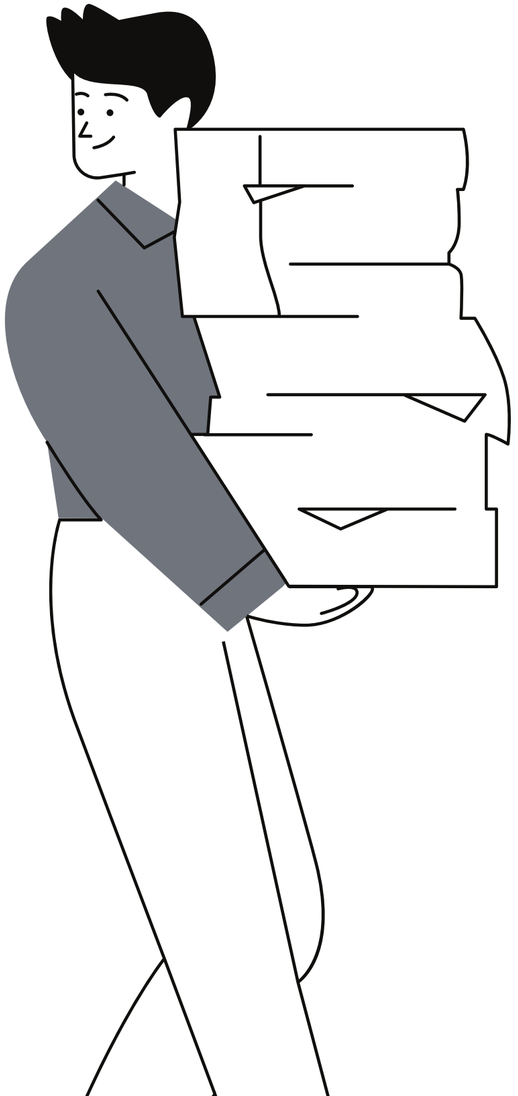
Department of Science & Technology

PR324

Team VIZION

Making Video-Conferencing more realistic, one frame at a time!

Team Leader : Parth Sarthi Prasad
College Code: 1-3512549572.



Use Cases

The gaze correction tool we have developed can be easily added to many platforms like zoom, google meets, skype, slack etc. Here are few of the significant use cases that our solution covers.

1

Realistic and natural communication.

2

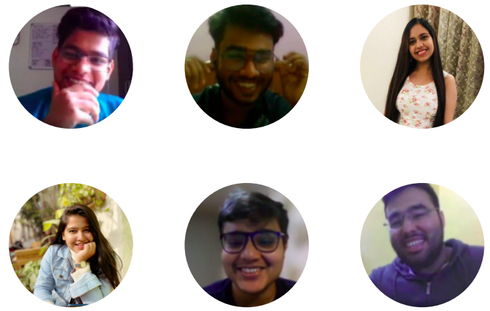
Teachers to increase the impact of their lessons

3

Can be used by the vulnerable elderly population to contact their loved ones.

4

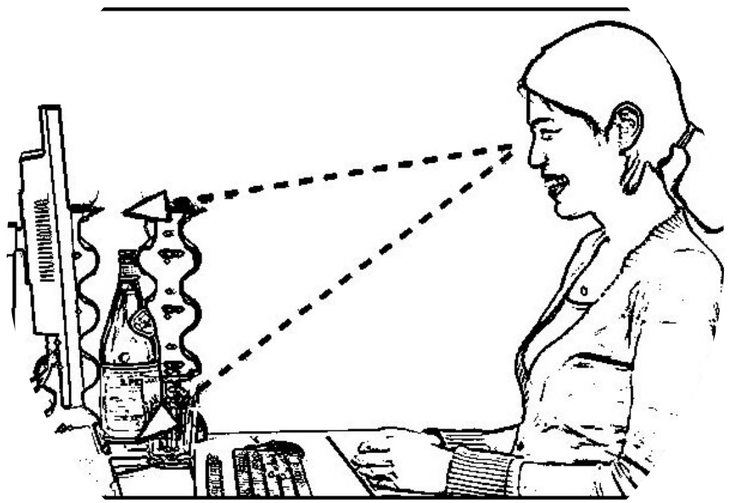
Can be used by managers to create a more realistic workspaces online.



Introduction

In current scenario, when all physical communication is replaced with video chatting mechanists. Over a hundred million video calls take place everyday and this number is bound to increase as more institutions adapt to the "New Normal".

However, most to these tools lack ability to recreate the atmosphere of real conversation because the eye contact is missing as camera and screen are misaligned, thus it seems unnatural.



1. Portalibilty
2. Computational Efficiency
3. Platform Independence
4. Cost Efficiency

Our Solution

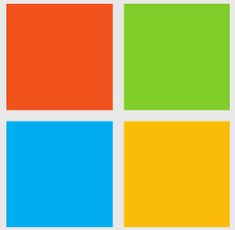
Team Vizion

1 App integration

2 Plugin based solution

3 Machine Learning model

Linux



Android

For the Android App we are using :-

- Dynamo Db for Storing Data
- JWT Library for Video Token Generation
- Twillio API for Video Calling
- SparkAr android SDK

1. Open the Android App.
2. Vote for your favorite ideas using stars.
3. Circle or mark up any promising ideas.

Plugin

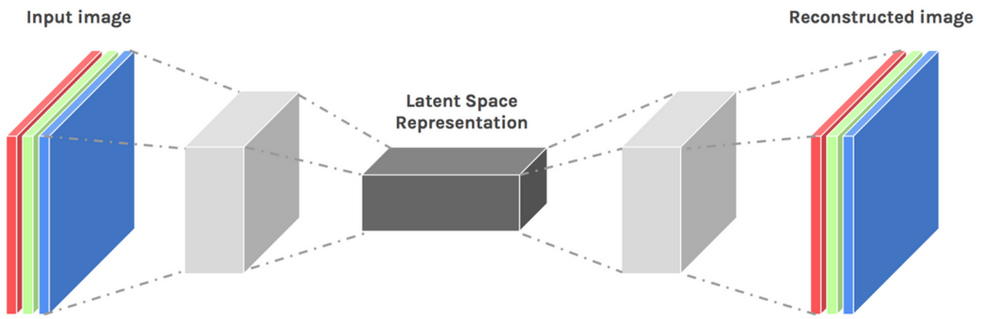
The plugin based solution, can run on edge device, this helps to reducing load on central server and making solutions viable for modern systems as well as older models using a remote server chosen if model faces a computational bottleneck, thus using a microservice on cloud rather than preferred edge device.

1. Create a new webcam sink using pyfakewebcam/ OBS.
2. Run ML model and select Input camera feed.
3. Choose output camera from meet link such as slack/meet/zoom/skype, etc

ML Model

What we have here is a generative convolutional autoencoder model trained for an Inpainting based task. While some gaze redirection models require the face and eye pose to be provided as inputs to generate the corrected image, our model has no such requirements. Our model is an inpainting based solution wherein, we simply remove the eyes from the input images and the model tries to fill the removed space in the image.
As our model was trained only on a dataset of forward gazing images, the induced selection bias causes it to generate forward gazing images.The input and output size of model is same at 128x128. Due, to limited computational resources, the size of the images and hence image clarity could not be greatly increased.

1. Input camera feed frame by frame and generate input mask to pass on only Eye specific region.
2. Pass on the mask to gaze correction model to return corrected eye gaze.
3. fill the removed spaced from the returned result in original frame and pass it to video sink.

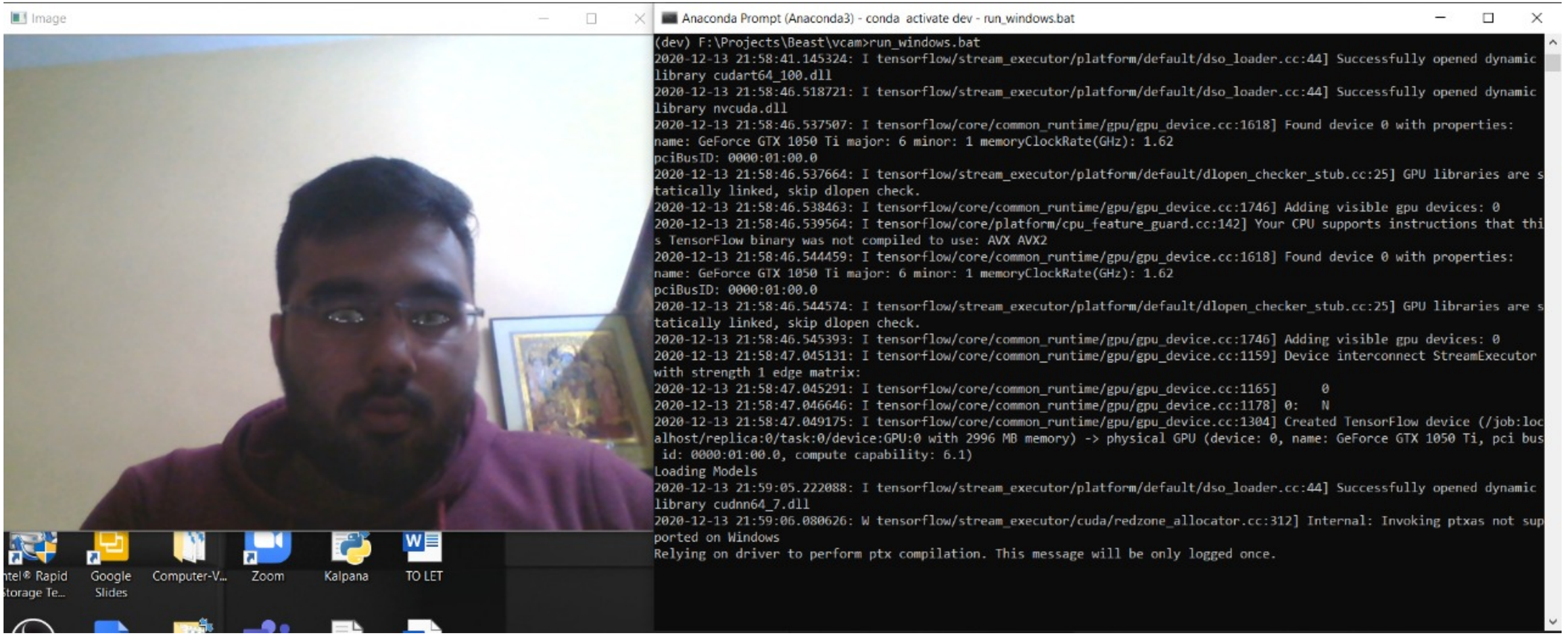


Check it out, Yourself!

https://youtu.be/sjX9b8svnaY

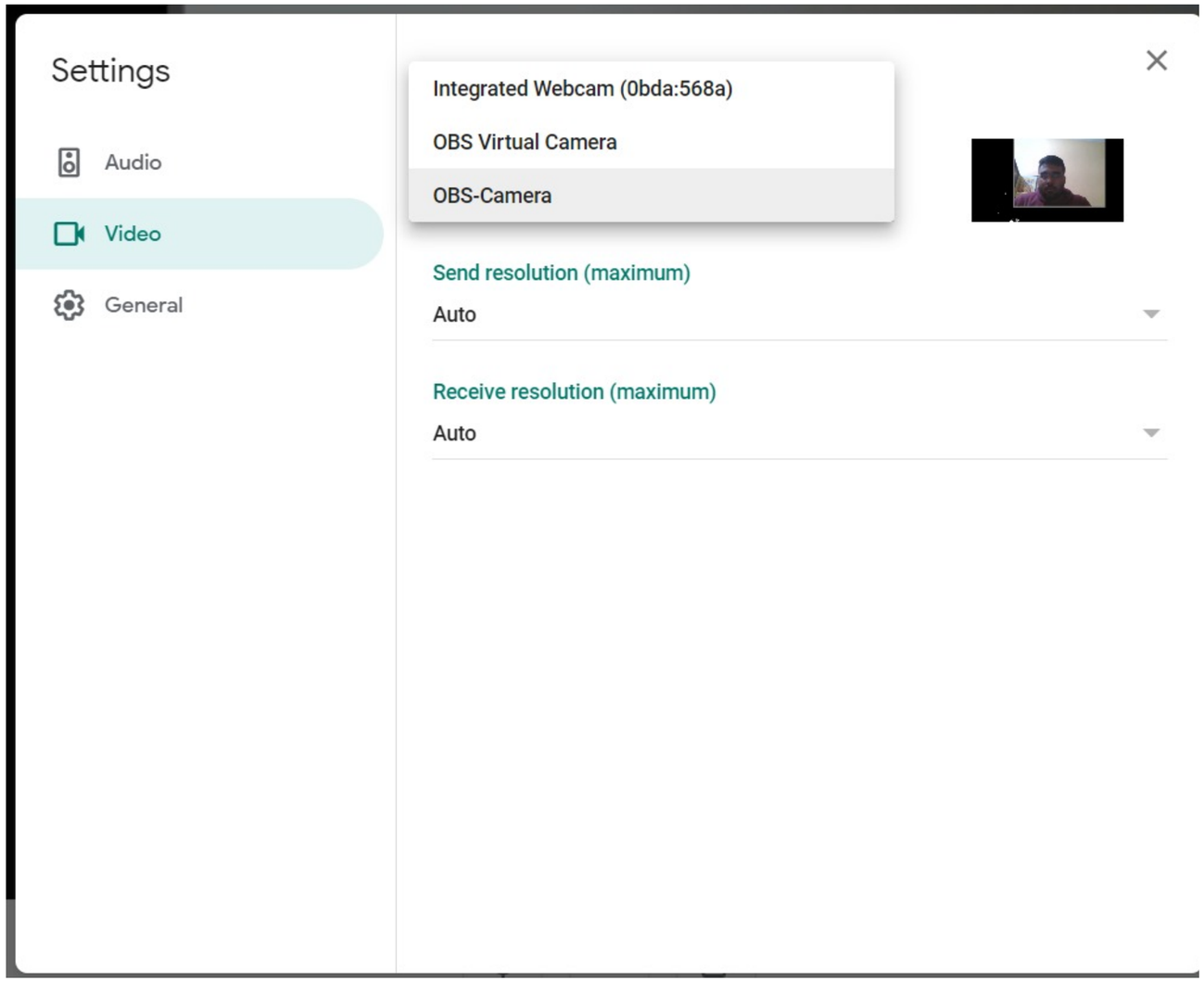
Team Vizion

1 Running the script



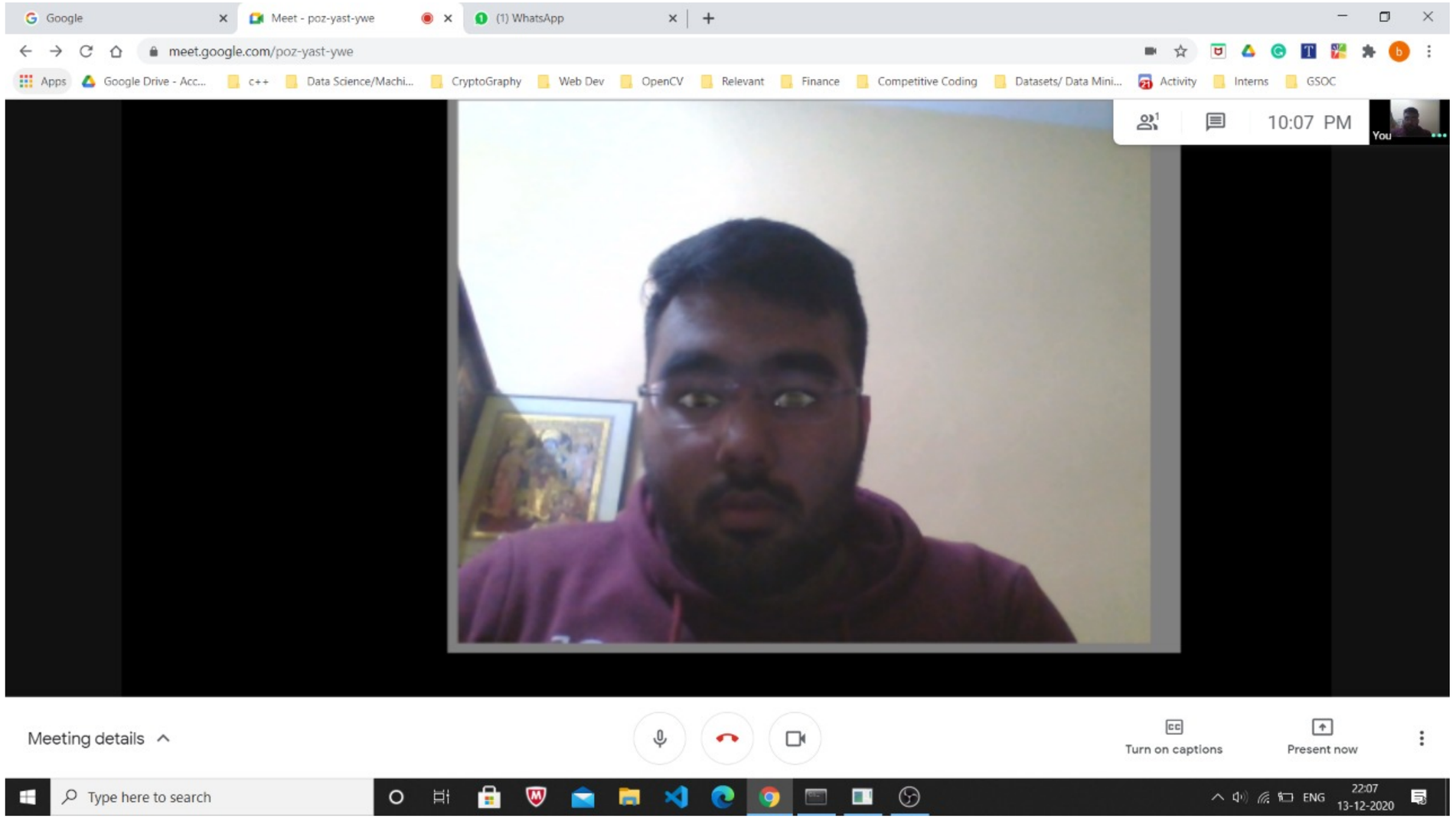
Team Vizion

2 Select OBS camera



Team Vizion

3 Use Google Meet.



Future Vizion

1. Android Application
2. Web Plugin
3. ML Model

As this is a basic tool and need of the hour , we have plans to expand the reach of this tool and make it available on every platform in an easily accessible and widely affordable manner for mass usage .Also we plan to improve our machine learning model to provide much more realistic view.

We hope to bring this tool to everyone as soon as possible so that even in tough times like these we can stay connected and strong together

Team Vizion

We're done!

Thank You