#AndroidDevChallenge

Tell us what your idea is

I intend to build an application for helping people with hearing and speech disabilities. This application will enable them to interact with people around them effortlessly without the need for a Sign Language Translator. They won't feel embarrassed or panic of being misinterpreted while communicating with people who do not understand Sign Language.

By removing this communication barrier, we can give them an equal opportunity to express their opinion and move shoulder to shoulder with other normal people, where speech or language doesn't hold them back anymore from realizing their potential.

The application will be able to convert ASL into Speech & Text and vice-versa. The application will be using the Google Camera Application in its core and utilizing On-Device Machine Learning capabilities from the ML Kit using modules like Object Detection & Tracking, AutoML Vision Edge and developing & training Custom Models, including Federated learning capability to the devices, would be an added advantage and providing the scope for customization based on the geographic variations of the sign language. This can also serve as a plugin, which can be integrated with all on-demand video chat services.

Tell us how you plan on bringing it to life

The project is still in the conceptual design stage and I am working on the basic architecture for the application, moving forward I prefer to follow the Iterative SLDC model as we have clearly defined the final product requirement which is the main task moving forward, we need to finalize the details related to the selection of machine learning modules and creating custom codes in a sequential manner, being able to refine our understanding of the system and improvise on the performance aspects with time. Just to list down a few of the important tasks moving forward:

1. Designing the combined model to recognize pipeline facial expressions and Hand Tracking Technology which can precisely detect the shape and motion of hands, as it is very important we combine both of these to accurately identify the signs.   **(Mid Feb)**

2. The next step will be preparation of machine learning based Segmentation mask for each frame to in-focus the foreground subject, considering the semantic cues to be able to track the change in the relative positioning of the prime features to track their relative motion with reference to the proceeding and succeeding frame.   **(End of March)**

3. Then being able to compare the frames and grouping them as activities, with the help of TensorFlow's Deep Learning models which is already trained with all the ASL signs to derive their meanings.   **(End of April)**

This is the plan till 1st of May, 2019. In addition to the above features after launch we would like to include some additional features over time:

4. Considering the sign language has variations over geographies, and it is next to impossible to accommodate all these variations, the model can be improved by adding on Federated learning

capability to the trained model to enhance the accuracy further and also being able to customize the model based on their geography.

Help & Inspiration from Google:

We are always inspired by Google's commitment to simplify the things and taking up complex challenges and coming up with innovative solutions.

It is very challenging task for implementing the proper models and obtaining the required functionality, we require the support of strong expertise in ML Kit and TensorFlow Lite.

Particularly the technologies which are of interest and look forward for guidance from the teams who have worked on developing these latest features are very much the inputs we require to move forward in building our application.

- Single camera Software Portrait Mode

- Hand Tracking Technology which can precisely detect the shape and motion of hands
     By Google Researchers Valentin Bazarevsky & Fan Zhang

- Real-Time AR Self-Expression with Machine Learning
     By Google Researchers Artsiom Ablavatski & Ivan Grishchenko

- Depth estimation using Semantic cues in addition to the parallax

- Synthetic Shallow Depth of Field

- Federated Learning: complementing the traditional centralized machine learning.

Tell us about you

My name is Sanket Rout, I have close to 3 years of experience in software development; have worked on several projects related to New Product Development. Developed consumption based utility billing system, Also part of the team that worked on smart factory concept trying to automate the operations & monitoring of complete factory via an intelligent dashboard, worked in developing the dashboard to monitor the operational status of critical Healthcare instrument involving training the ML modules with past failure data to enhance the predictive maintenance module. Have hands on experience in implementing deep learning neural network from scratch.

My Teammate Sankalp Rout, is final year of Undergrad student from IIIT, Bhubaneswar. He has worked on several academic projects on AI / ML technologies, like making a an application for controlling the music player from the web cam using hand gestures, another application for sentiment analyser for music player theme selection. He has completed several certification courses on AI, Ml & Deep Learning from the Coursera & Deeplearning AI by Andrew NG. He is an active member of the GDG Bhubaneswar.