Rajiv Gandhi Institute of Technology
Chola Nagar, Bengaluru, Karnataka - 560032

Sign Language Interpreter using Deep Learning

Under the guidance of : Asst. Prof. Pushplata Dubey

- Bahaduri Prachiti Jagdish - 1RG17CS009
- Mohammed Azim - 1RG17CS030
- Mohammed Touseef - 1RG17CS031
- Prajwal B Mani - 1RG17CS037

ABSTRACT

Speech Impairment is a disability, which affects an individual's ability to communicate using speech and hearing. This brings about the difficulty for both the sign and non - sign language speakers to communicate with each other. With recent advances in deep learning and computer vision, the focus of our project is to create a vision of an end to end Convolutional Neural Network that will be trained on the ASL(American Sign Language) dataset then modeled on robust architectures like GoogLeNet/MobileNet architecture and deploy it on an android application so that it will have more accessibility and provides an ease of use, thus aiding communication between signers and non-signers. It is a challenging and interesting problem that if solved will bring a leap in social and technological aspects alike.

MOTIVATION

- Communication is one of the basic requirements for survival in society.
- The lack of ISL dataset and often leads to obscurity of features due to overlapping of hands this leads us to use ASL.
- Our project aims at taking the basic step in bridging the communication gap between normal people and deaf and dumb people.
- Effective extension of this project to words and common expressions may not only make the deaf and dumb people communicate faster and easier with outer world, also providing a boost in developing autonomous systems for understanding and aiding them.

PROBLEM IDENTIFICATION

According to the projection of data it is expected to rise to 900 million by 2050 that is double the current stats.

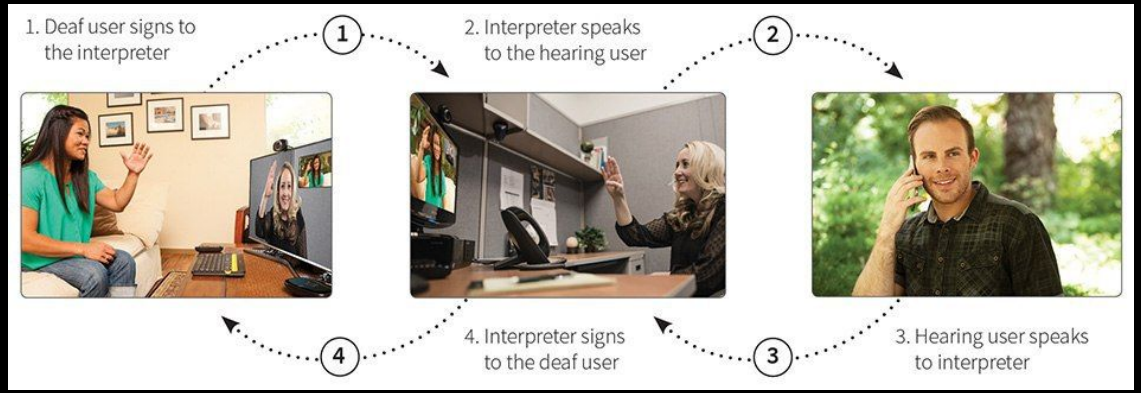
Hearing loss can affect a person in three main ways:

- Education
- Job
- Social Withdrawal

OBJECTIVES 

1. Create a robust android application that aids in communication for deaf people.
2. Use deep learning with transfer learning techniques to build a neural network model that as to ability to learn pattern in video and classify the images.
3. An error correction model for identifying the pattern mismatch and correcting it for the audio or text input format.
4. To build a model that has higher accuracy with less bias and overfitting problems.
5. To enhance the model to classify at least 20 classes or more .

EXISTING SYSTEM



- **<u>Traditional - based approach</u>**

A sign language interpreter is a person trained in translating between a spoken and a signed language.

EXISTING SYSTEM

- **Glove - based approach**
  - The signers are required to wear a sensor glove.
  - The task will be simplified during segmentation process by wearing glove.

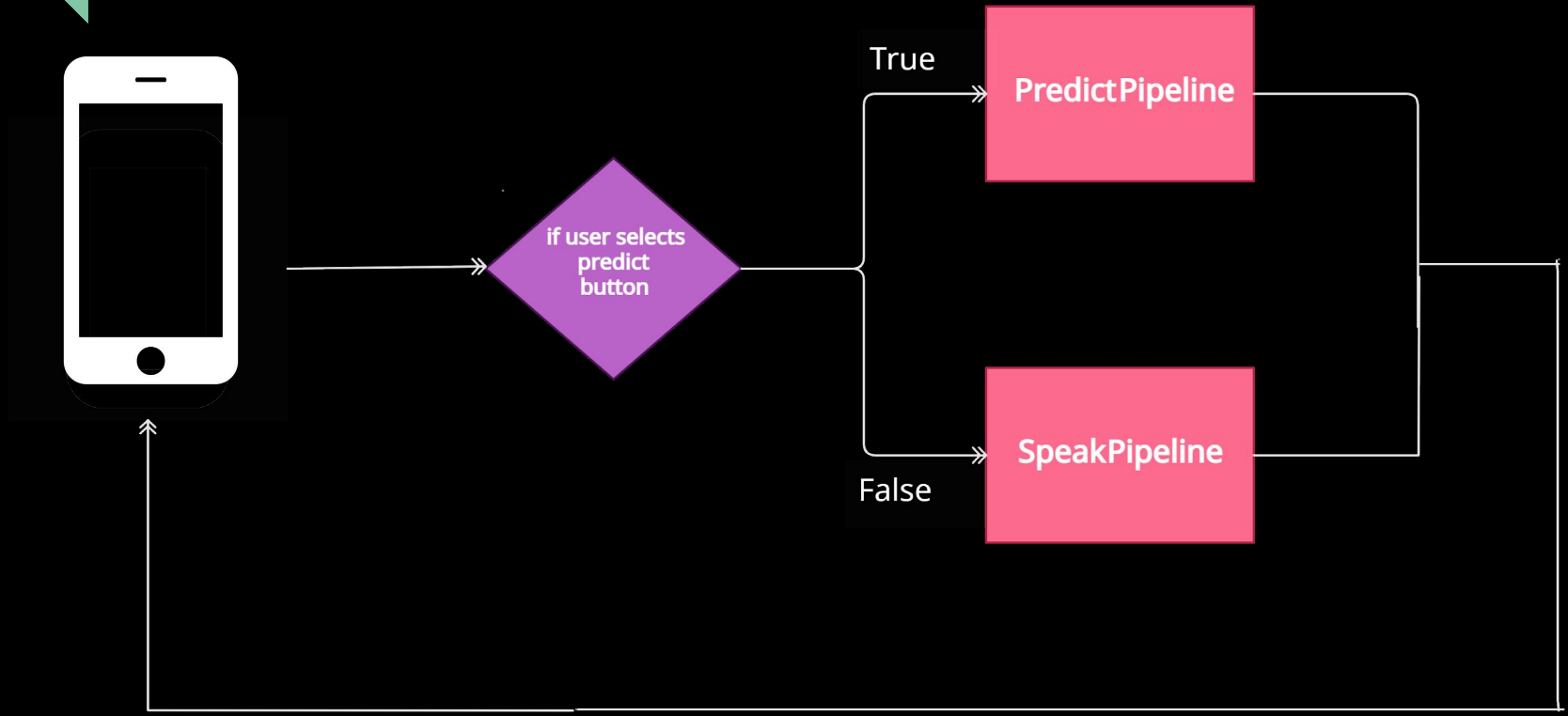
DRAWBACKS

- Lack of availability of the interpreter for that moment.
- The gloves based approach is that the signer has to wear the sensor hardware along with the glove during the operation of the system.
- The cost of traditional based approach will vary from 100 to 200 dollars/hrs and for gloves based approach its costs around 40,000 dollars / pair

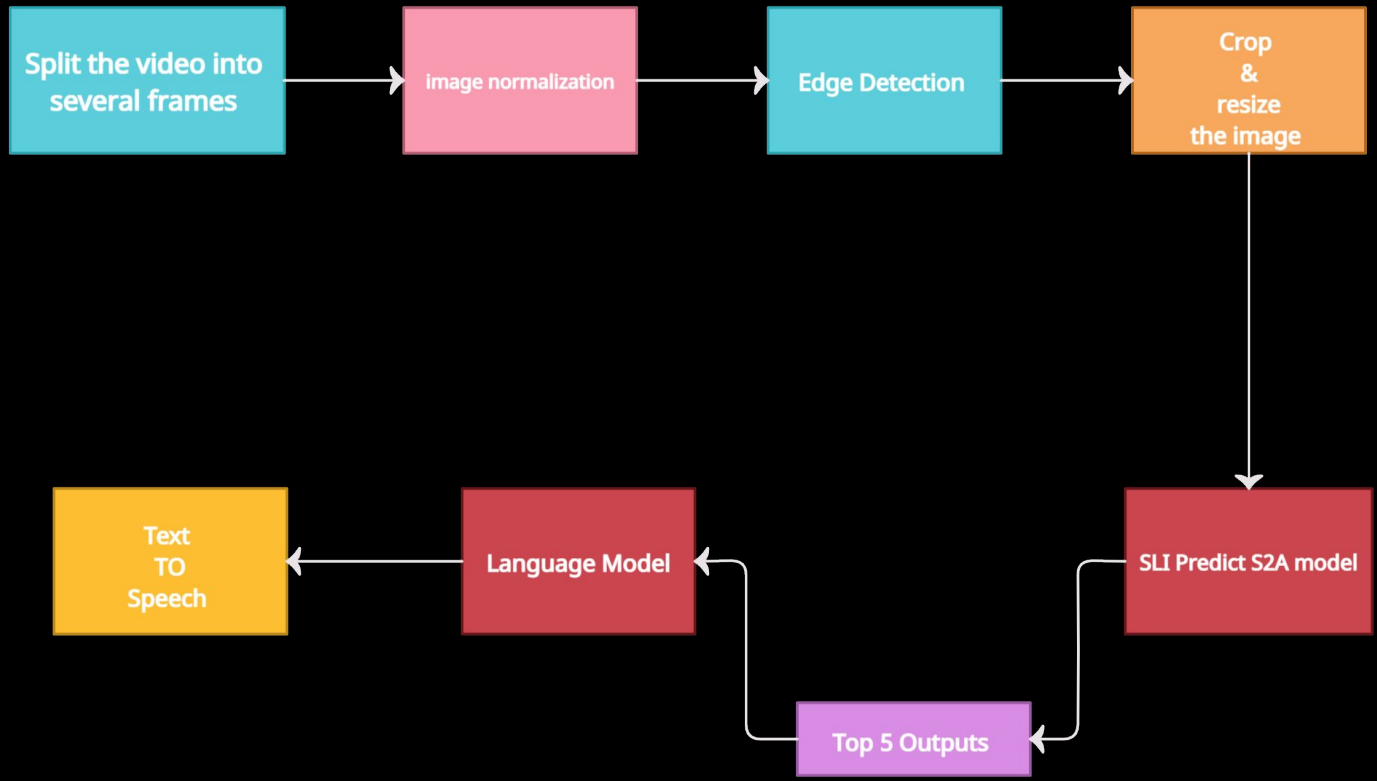
PROPOSED SYSTEM

We have introduced a refined approach using advanced deep learning techniques that is flexible to use by any mobile OS. In this presentation, we propose an architecture to detect signs. Image pre-processing makes the existing data and input data normalized. The pipeline allows us to automate machine learning workflow. Transfer learnings make it easy to make the models to learn even small details that are hard to capture in a small network. Kivy gives us an advantages overview by giving us a docile structure to compile to any version of the mobile app.

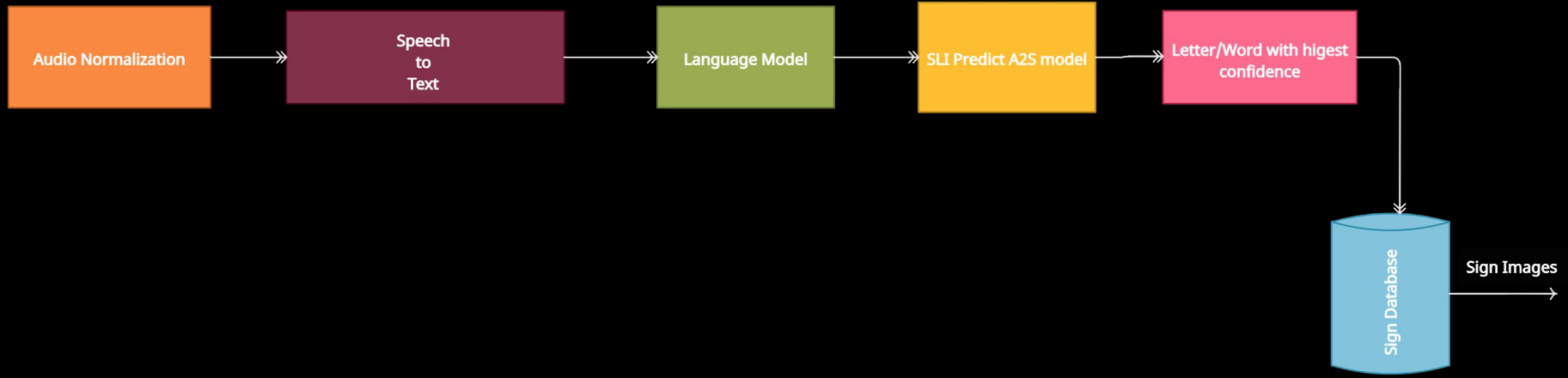
ARCHITECTURE OF PROPOSED SYSTEM



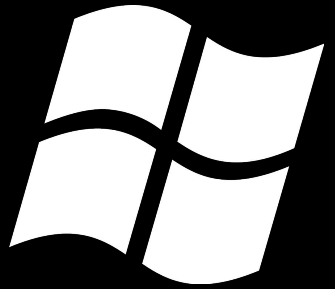
PREDICT PIPELINE



SPEAK PIPELINE

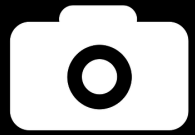


HARDWARE AND SOFTWARE REQUIREMENT SPECIFICATION



- HARDWARE REQUIREMENTS:

For Typical Operating System (Windows 10)

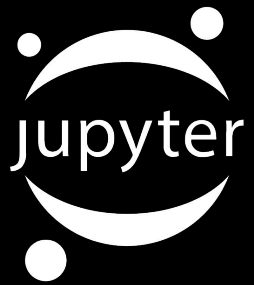


- Intel Core i3-3rd gen processor or later.
- 8GB RAM & 120GB disk space
- A good GPU support
- I/O devices
- Any external or inbuilt camera with minimum pixel resolution 200 x 200 (300pi or 1501pi) 4-megapixel cameras and up.

HARDWARE AND SOFTWARE REQUIREMENT SPECIFICATION

❑ SOFTWARE REQUIREMENTS:

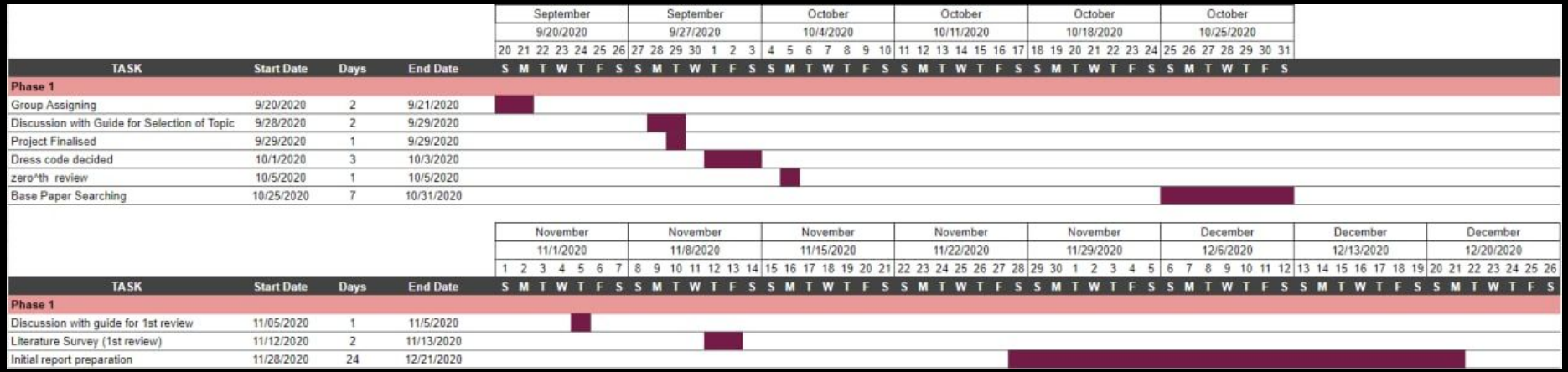
- Front End : Android (using Kivy)
- Video & image processing : Open CV, scikit-image
- Neural Network Modeling : Tensorflow(tf.lite), Keras, Pytorch
- Architectures : Google LeNet, MobileNet, VGG 19
- Ubuntu : deploying .apk using buildozer kivy
- Editor : Visual code, jupyter notebook



POSSIBLE OUTCOME OF THE PROJECT

- UI Interface to access model functionalities
- Improved accuracy
- Dynamic Tracing
- Continuous integration pipelines that enable model redeployment

GANTT CHART



TASK	Start Date	Days	End Date
Phase 1			
Group Assigning	9/20/2020	2	9/21/2020
Discussion with Guide for Selection of Topic	9/28/2020	2	9/29/2020
Project Finalised	9/29/2020	1	9/29/2020
Dress code decided	10/1/2020	3	10/3/2020
zero^th review	10/5/2020	1	10/5/2020
Base Paper Searching	10/25/2020	7	10/31/2020

TASK	Start Date	Days	End Date
Phase 1			
Discussion with guide for 1st review	11/05/2020	1	11/5/2020
Literature Survey (1st review)	11/12/2020	2	11/13/2020
Initial report preparation	11/28/2020	24	12/21/2020

CONCLUSION

A sign language is a natural mode of communication used by the deaf community. India has large population of speech and hearing impaired but a very small number of certified sign language interpreters are available. Research in hand gesture recognition has gained attention with advancement in the field of computer vision. A Sign Language Interpreter (SLI) decodes and understands the information conveyed by signs. SLI can be a major breakthrough in helping a common people to communicate with the deaf and can help in bridging this communication gap. A SLI can be designed based on video/image processing and deep learning techniques which requires a standard dataset, determination of an optimal feature set, and an appropriate classification technique.