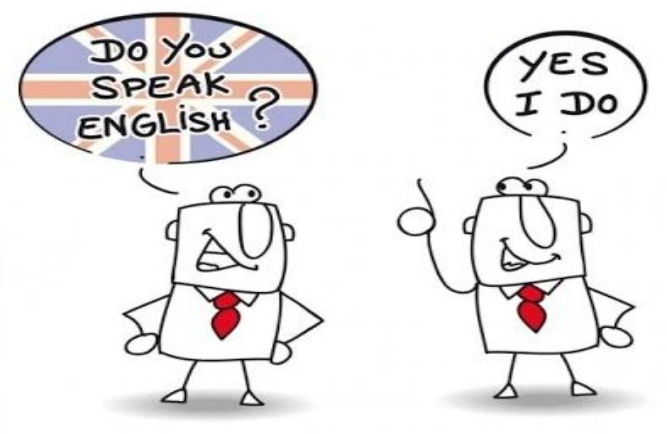
Two-way Communication between Deaf(Dumb) and Normal People using Deep Learning

Introduction



- Humans need a language to communicate with each other so, what's the issue here?
- Regular(English) vs Sign Language

- Regular and Sign Languages are completely different one.
- Can we create a bridge between both languages? Yes

Translators are bridges

- Human SL Translators
- Hand Gloves

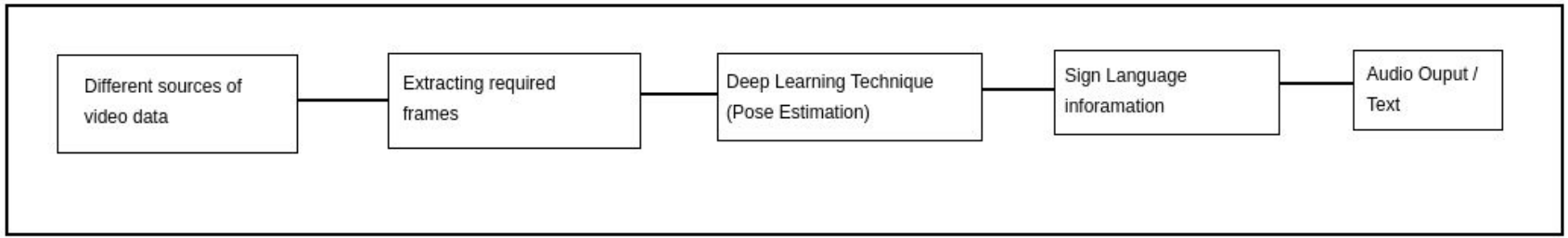
Motivation

In order to provide the right to communicate effectively, where a person with hearing impairment can communicate with the one who has no prior knowledge of sign language. This bridge enables normal individuals to understand the sign language and communicates with them.

Two-way Communication

Work Flow

Step 1. When a Deaf-Dumb person(DDP), wants to communicate with the Normal Person (NP), we can capture the hand gestures and facial expressions via a device camera, analyzing them using the deep learning techniques and the appropriate meanings are output as voice/text.



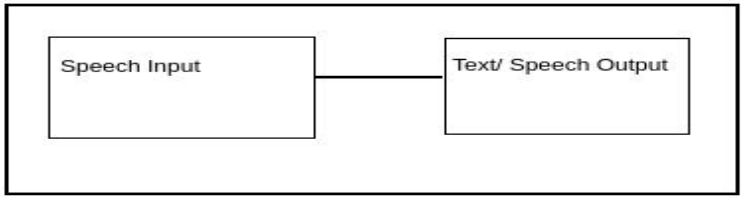
Step-by-step approach

1. Capturing videos of different sign words based on the ISL
2. Cleaning the videosets according to individual signs
3. Extracting the frames out of each videoset
4. Normalising the images
5. Estimating the pose and finding the key points the images using Pose Net
6. Training the neural network, based on the poses
7. The analysed information, is outputed.

Challenges

1. Latency : Processing can be slow
2. Lack of Gesture Language : Different people, different gestures
3. Robustness: The hand gesture motions can’t be read optimally because different environment conditions
4. Performance: Gesture recognition require intensive work do be done, it’s difficult to run on resource constrained devices.

Step 2. In response to it, NP can record the information, which wants to be conveyed, this information is converted using the voice to text API as a text, which can be read by the DDP.

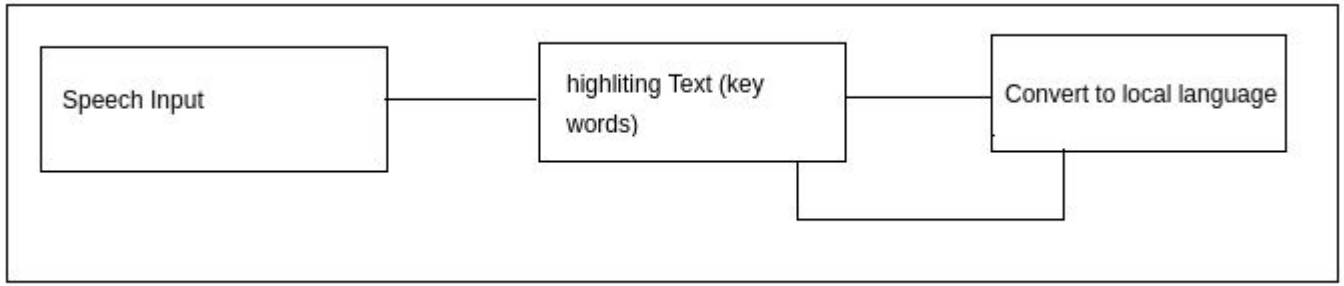


I Would you like to have lunch with me?

👉 You-me lunch go?

I The yellow ball is in the cupboard

👉 Cupboard ball yellow inside



Step-by-step approach

1. Speech input is recorded using the microphone
2. Text is generated from the input-ed speech
3. Keywords of the text is highlighted.
4. Language translation APIs to translate into desired language.

Conclusion and Future Work

- Using Pose Estimation of an individual gather precise information, rather taking time to train to the same pose in different angles(slight deviations)
- These two-way approach, provides a chance to increase the communication between two different language community people.
- In future, more efficient ways can be developed using 3D avatars for better communications.
- Based on similar idea, applications can be developed by integrating the code into devices which enables hassle less mode for two-way communication and better understandings.