Indian Sign Language Recognition System for differently abled (Hard of hearing)

By - R.Srikar(AM.EN.U4AIE19051)

Project Guide - Dr. Geetha M

Motivation

- Deaf and Hard of hearing people cannot interact with the computer since they cannot see the screen nor hear what it says via speakers.
- We aim at creating a chatbot which interacts with these people using ISL(Indian Sign Language) and understanding what they say via ISL hand signs in the webcam feed.
- This assists them in basic interaction with computer like inputting text and filling forms.

Indian Sign Language

- Indian Sign Language(ISL) is a method of representing alphabets and numbers using hand signs.
- This can be used as a medium between deaf and dumb people for communicating with each other.
- We use this as a communication medium in our approach with out chatbot and the user.

Research Problem

- There is a need for a chatbot that can effectively communicate with hard of hearing individuals using Indian Sign Language (ISL) as the common medium which assists them in filling forms and providing input to the computer.
- The lack of communication channels for the deaf and dumb community makes it challenging for them to access technology and fill out online forms, which can be a major hindrance in their daily lives.
- The chatbot will act as a bridge between the user and the computer, enabling them to interact with technology effortlessly.
- By incorporating ISL into the chatbot, we can ensure that the communication between the user and the computer is seamless and effective, providing them with the opportunity to access technology like never before.

Problem statement

- By utilizing ISL as a medium of communication, implementing an AI-based chatbot that can effectively assist hard of hearing individuals in filling out online forms and interacting with technology with an efficient sign recognition module that addresses the challenges of movement epenthesis, signer-dependent variations, and complex backgrounds,.
- Furthermore, optimizing the sign recognition module for real-time streaming video input and fast response will provide a seamless and effective interaction between the user and the computer, improving their quality of life.

Literature Review

- Sign Language Recognition for Static and Dynamic Gestures by Jay Suthar
- Zero-Shot Sign Language Recognition: Can Textual Data Uncover Sign Languages? by Yunus Can Bilge
- A Review Paper on Sign Language Recognition of Engineering System For Deaf And Research & Dumb People using Image Technology Processing by Manisha U. Kakde
- Attention-Based Sign Language Recognition Network Utilizing Key frame Sampling and Skeletal Features by Wei Pan

Challenges

- Detect ISL hand signs for all 36 characters(26 alphabets + 10 numbers).
- Provide methods to process raw model output into meaningful sentences or text.
- Handle different backgrounds in the webcam feed.
- The model must detect the hand sign in the image really fast.
- The whole architecture must be scalable to serve multiple users.
- Tackle the problem of movement epenthesis.
- Tackle the problem of different signers showing signs.

Research Gap/ the need for your work

- Model used in the first paper of literature review has an accuracy of around 73 percentile which is not the percentile of accuracy we want in our model.
- In the second and third and fourth papers of literature review, it becomes really tough for the models proposed to be used in a chatbot since the response needs to be as fast as possible and those model architectures are really big.
- The third concern is the scalability part wherein we have to make the chatbot to work for multiple users at a time and this issue is not addressed in any research papers we reviewed.
- Our work addresses all these problems and provides a robust solution for detecting ISL hand signs using a chatbot.

Proposed Methodology

- User shows ISL hand signs which are captured using a webcam.
- We get the coordinates of hands from the image using mediapipe framework.
- Transformation functions were applied to the hand coordinates to make all hands similar and in same position.
- The transformed coordinates are used to train a neural network which will later predict hand signs according to the coordinates.
- Basic NLP is applied to the model output for removing unwanted characters, applying spellchecking and word segmentation if required.
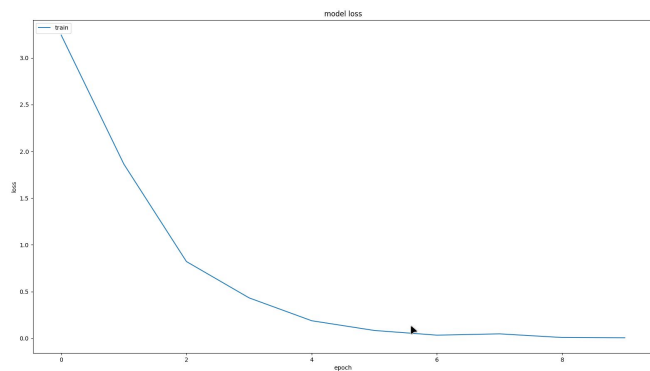
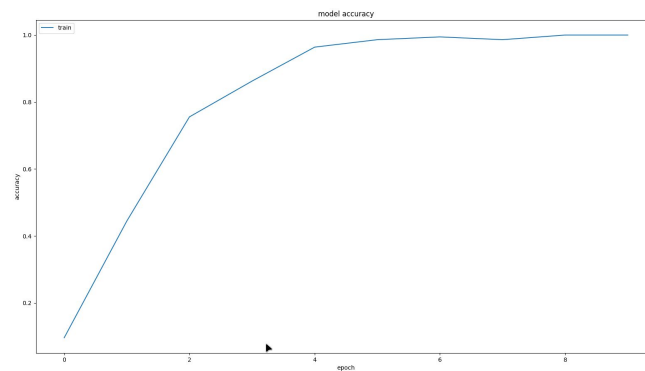# Final Architecture and Block Diagram
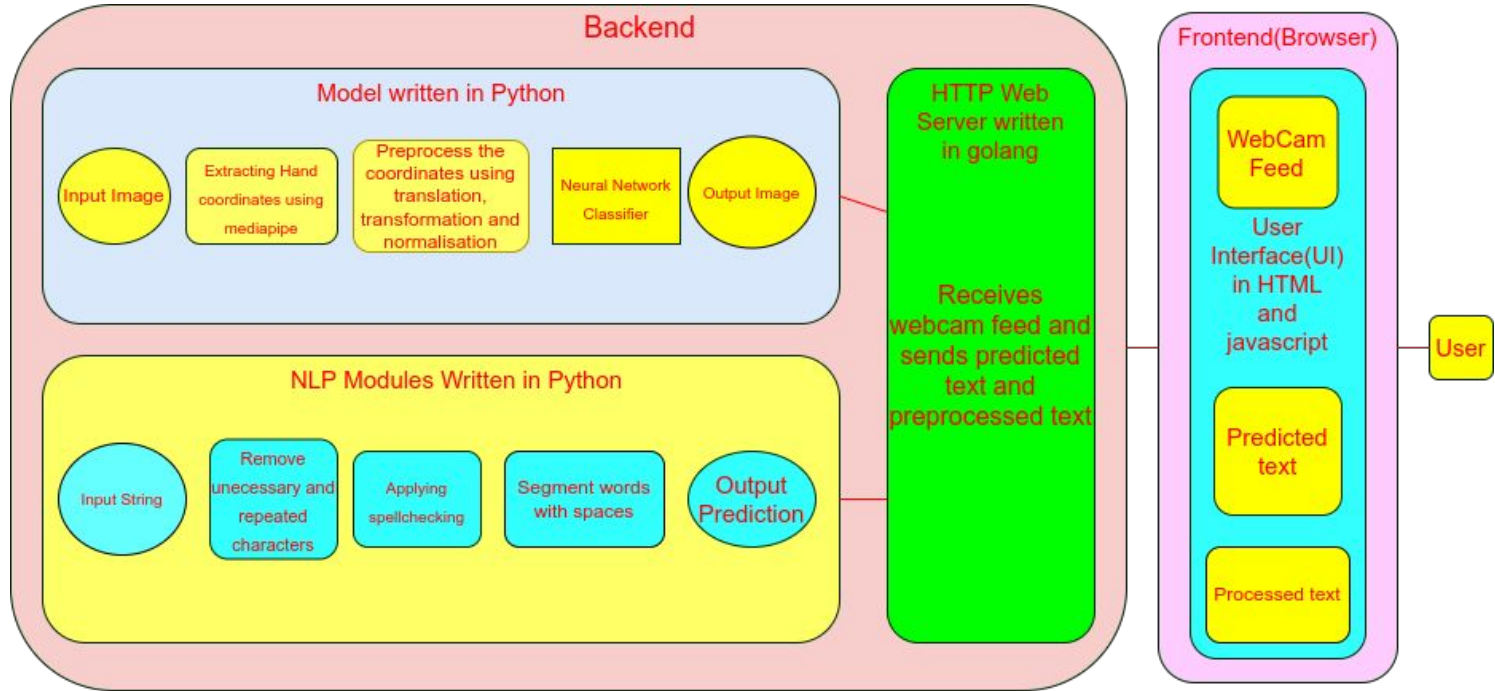
Results and Discussion

- The model with the new dataset gave an incredible boosted the training accuracy.
- The transformation function gave a great boost to the accuracy since it made all the hands look a bit similar.
- The model is now able to detect all the 36 characters very easily from the live feed.

Results

Loss per epoch graph

Accuracy per Epoch graph

Chatbot UI Results(Continued)

NLP Results

Word Segmentation for "HI HOW ARE YOU I AM DOING FINE IS SRIKAR THERE IN HIS ROOM"	HIHOWAREYOUIAMDOIN GFINEISSRIKARTHEREIN HISHOME	HI HOW ARE YOU I AM DOING FINE IS SRI KART HERE IN HIS HOME
Spell Checking for "CUT"	CUV	CUT

```
Temp sentence: HIHOWAREYOUIAMDOINGFINEISSRIKARTHEREINHISH
Temp sentence cleared
Temp sentence: HIHOWAREYOUIAMDOINGFINEISSRIKARTHEREINHISHO
Temp sentence cleared
Temp sentence: HIHOWAREYOUIAMDOINGFINEISSRIKARTHEREINHISHOM
Temp sentence cleared
Temp sentence: HIHOWAREYOUIAMDOINGFINEISSRIKARTHEREINHISHOME
Correct word: ['hi', 'how', 'are', 'you', 'i', 'am', 'doing', 'fine', 'is', 'sri', 'kart', 'here', 'in', 'his', 'home']

[N]:[xeltron@xeltron]
[/media/xeltron/main/Coding/web_development/fyp/ml]
[$]:
```

```
Temp word cleared
Temp word cleared
Temp word cleared
Temp Word: C
Temp word cleared
Temp Word: CU
Temp Word: CUV
Correct word: cut
[I]:[xeltron@xeltron]
```

Conclusion and Future Scope

- Model detects all the ISL hand signs really good but it should work on hands of people whose hands are really big.
- Since our dataset is not very large and only consists of 20 images per character totaling to 720 images for 36 characters, increasing the size of the dataset might give even better accuracy for predictions for all kinds of hands.
- The dataset could also consist of left handed coordinate samples for left handed predictions. Currently there are no left handed coordinate samples.
- Working on the dataset can increase the model prediction accuracy by a great extent in realtime.
- Also our NLP modules cannot handle nouns well. Some research can be done to find a better way to handle nouns.

Target Publication & Publication status

- Springer

References

- https://indiansignlanguage.org

The End
Thank You