International Conference on Artificial Intelligence and Machine Learning
In association with International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com) | Volume 8, Issue 5, July-August-2021

Deaf Sign Interpreter Service

Shivanand Manjaragi*[1], Revati Patil[1], Chaitra Shende[1], Aruna Gadakeri[1], Sourab Nagare[1]
*[1]Department of Computer Science and Engineering, Hirasugar Institute of Technology, Nidasoshi, Belgaum, Karnataka, India

ABSTRACT

Sign language is the way of communication for hearing impaired people. There is a challenge for common people to communicate with deaf people which makes this system helpful in assisting them. This paper aims at implementing computer vision which can take the sign from the users and convert them into text in real time. The proposed methodology contains four modules such as: image capturing, pre-processing classification and prediction. This system uses a camera, which captures various gestures of the hand. Next, the captured image is pre-processed, the edges are determined an edge detection algorithm. Finally, a template-matching algorithm identifies the sign and display the text. As the output is text, one can easily interpret the meaning of a particular sign. This also curtails the difficulty to communicate with the deaf. The system is implemented by using OpenCV-Python. The system uses various libraries and finally translates the sign gestures to text with accuracy.
Keywords: Deaf Sign Interpreter, Deep learning, template-matching algorithm, RCNN

I. INTRODUCTION

The deaf and dumb people find it difficult to communicate their needs in their day to day life. The Sign language is developed so that they can communicate with other people using hand movements. Sign languages are developed primarily to aid deaf and dumb people. They use a concurrent and specific combination of hand movements, hand shapes and orientation in order to convey particular information. One such set of language is the Indian Sign Language (ISL) system which is predominantly used in south Asian countries. Certain aspect that distinguishes ISL from other sign languages is that ISL devoid of any temporal inflections in its finger spelling chart and also the usage of both the hands. Many people in India are speech and/or hearing impaired, and they thus use hand gestures to communicate with other people. However, apart from a handful number of people, not everyone is aware of this sign language and they may require an interpreter which can be inconvenient and expensive. This paper aims to narrow this communication gap by developing software which can predict the ISL alphanumeric hand gestures in real time.

The objective of this system is to elevate people with hearing disability and help them socialize with common people. It is a form of non-verbal communication.

- Reorganization of sign language automatically from video stream to help deaf and dumb people, this system helps them to communicate more effectively with each other or normal people.

- Reorganization of hand gesture from the video stream and converting gesture to text form which helps deaf and dumb people to commutate effectively.
- This reorganization system provides us an innovative, natural, user friendly way of communication with the computer which is more familiar to the human beings.

The organization of this paper is as follows. The Section 2 is literature review, Section 3 presents the details of the proposed methodology, Section 4 consists of result analysis and section 5 is conclusion and future work.

II. LITERATURE REVIEW

[1]. Deaf Mute Communication Interpreter- A Review: This paper aims to cover the various prevailing methods of deaf-mute communication interpreter system. The two broad classifications of the communication methodologies used by the deaf – mute people are - Wearable Communication Device and Online Learning System. Under Wearable communication method, there are Glove based systems, Keypad method and Handicom Touch- screen. All the above mentioned three sub-divided methods make use of various sensors, accelerometer, a suitable micro-controller, a text to speech conversion module, a keypad and a touch- screen. The need for an external device to interpret the message between a deaf –mute and non-deaf- mute people can be overcome by the second method i.e. online learning system. The Online Learning System has different methods. The five subdivided methods are- SLIM module, TESSA, Wi-See Technology, SWI_PELE System and Web- Sign Technology.

[2]. An Efficient Framework for Indian Sign Language Recognition Using Wavelet Transform: The proposed ISLR system is considered as a pattern recognition technique that has two important modules: feature extraction and classification. The joint use of Discrete Wavelet Transform (DWT) based feature extraction and nearest neighbor classifier is used to recognize the sign language. The experimental results show that the proposed hand gesture recognition system achieves maximum 99.23% classification accuracy while using cosine distance classifier.

[3]. Hand Gesture Recognition Using PCA in: In this paper authors presented a scheme using a database driven hand gesture recognition based upon skin color model approach and thresholding approach along with an effective template matching with can be effectively used for human robotics applications and similar other applications.. Initially, hand region is segmented by applying skin color model in YCbCr color space. In the next stage thresholding is applied to separate foreground and background. Finally, template based matching technique is developed using Principal Component Analysis (PCA) for recognition.

[4]. Hand Gesture Recognition System for Dumb People: Authors presented the static hand gesture recognition system using digital image processing. For hand gesture feature vector SIFT algorithm is used. The SIFT features have been computed at the edges which are invariant to scaling, rotation, addition of noise.

[5]. An Automated System for Indian Sign Language Recognition in: In this paper a method for automatic recognition of signs on the basis of shape based features is presented. For segmentation of hand region from the images, Otsu's thresholding algorithm is used, that chooses an optimal threshold to minimize the within-class variance of threshold black and white pixels. Features of segmented hand region are calculated using Hu's invariant moments that are fed to Artificial Neural Network for classification. Performance of the system is evaluated on the basis of Accuracy, Sensitivity and Specificity.

[6]. Hand Gesture Recognition for Sign Language Recognition: A Review in: Authors presented various methods of hand gesture and sign language recognition proposed in the past by various researchers. For deaf and dumb people, Sign language is the only way of communication. With the help of sign language, these physical impaired people express their emotions and thoughts to other person.

[7]. Design Issue and Proposed Implementation of Communication Aid for Deaf & Dumb People in: In this paper author proposed a system to aid communication of deaf and dumb people communication using Indian sign language (ISL) with normal people where hand gestures will be converted into appropriate text message. Main objective is to design an algorithm to convert dynamic gesture to text at real time. Finally after testing is done the system will be implemented on android platform and will be available as an application for smart phone and tablet pc.

[8]. Real Time Detection and Recognition Of Indian And American Sign Language Using Sift In: Author proposed a real time vision based system for hand gesture recognition for human computer interaction in many applications. The system can recognize 35 different hand gestures given by Indian and American Sign Language or ISL and ASL at faster rate with virtuous accuracy. RGB-to- GRAY segmentation technique was used to minimize the chances of false detection. Authors proposed a method of improvised Scale Invariant Feature Transform (SIFT) and same was used to extract features. The system is model using MATLAB. To design and efficient user friendly hand gesture recognition system, a GUI model has been implemented

[9]. A Review on Feature Extraction for Indian and American Sign Language in: Paper presented the recent research and development of sign language based on manual communication and body language. Sign language recognition systems typically elaborate three steps pre processing, feature extraction and classification. Classification methods used for recognition are Neural Network(NN), Support Vector Machine(SVM), Hidden Markov Models(HMM), Scale Invariant Feature Transform(SIFT),etc.

[10]. SignPro- An Application Suite for Deaf and Dumb. In: Author presented application that helps the deaf and dumb person to communicate with the rest of the world using sign language. The key feature in this system is the real time gesture to text conversion. The processing steps include: gesture extraction, gesture matching and conversion to speech. Gesture extraction involves use of various image processing techniques such as histogram matching, bounding box computation, skin color segmentation and region growing. Techniques applicable for Gesture matching include feature point matching and correlation based matching.

The other features in the application include voicing out of text and text to gesture conversion.

From the literature review, we have come across multiple solutions and study on detection and interpretation of sign language for deaf and dumb. From the literature review studied, it can be noted that the majority of solutions proposed by the research scholars is hardware based solution which consists of sensory gloves which can track the gestures to determine the sign language using MEMS sensors mounted on the system. The comparative study can be performed concluded that the sensor based solutions require extra hardware to be attached to the deaf and dumb person's body to interpret the sign language. This is not only hectic to wear but also has accuracy issues. Other Computer Vision based algorithms use SURF and SIFT for Detection of hand gestures. This required more optimization to get accurate results. Thus our proposed solution using deep learning RCNN neural network is more accurate and by using this we train the model. This model is user friendly to operate using the developed python GUI application and can be scaled anytime by altering the data.

III. PROPOSED METHODOLOGY

The proposed system consists of development of Deaf sign Interpreter service using Python and Deep learning. The system consists of training a machine learning model capable of detection and interpreter of sign languages for deaf and dumb using deep learning and Python. The proposed system consists of a python GUI application which can accept the image data feed form the Camera. The system then uses deep learning to detect the sign language in the image and then interprets it using trained deep learning RCNN neural network to detect the sign language. The sign interpreted is showed as text in GUI application. Thus proposed system not only can help the deaf and dumb to communicate easily but also avoid the necessity of person with the knowledge of sign language who works as an interpreter between the deaf and dumb people and common people.

Sign language is evolving as an inevitable communication method for the hearing impaired persons. The basic element of the sign language is the sign language alphabets. By using these alphabets the system is built by using the Deep learning and CNN technologies. This system is basically built to help impaired people for communication purpose.

The proposed work consists of development of sign language interpreter for deaf and dumb using deep learning. As shown in block diagram the system consists of python GUI application, Python backend with OpenCV and a deep learning model capable of detecting or recognizing the deaf Sign language. The GUI application developed will have the option of loading the image for the interpretation of the sign language or starting a live video stream to detect the sign language in the live video. The Camera captures the video and feeds the frame by frame data to the python backend where OpenCV performs image preprocessing to feed it to the deep learning model. The Deep learning neural network is trained on sign language data using transfer learning approach to detect and interpret sign language using the provided dataset and to predict the sign message which is conveyed by the deaf and dumb people. The preprocessed images are fed to the trained neural network which will display the text as output regarding the detected sign. The system thus forms and better and accurate solution of sign language interpretation which is not only accurate but also can be scaled in future since it is based on deep learning.

A. System Architecture

The System architecture is shown in Fig. 1. In our system we are using a pre-trained model. The first process is Data acquisition. Here we are collecting data in terms of images. Lots of data is collected, it is further used for preprocessing. In preprocessing, the LabelImg software is used for generating XML files. Which were further converted into TF-Record files. By using these TF-Record files the RCNN deep learning model is trained. This trained model further used for creating GUI application. Here the images will be feed on the top of this model. Then the inference session collects the features of the image, where capabilities are learned from the deep learning trained model for the sign detection. After detecting the gesture, python Tkinter GUI is used to provide fast and easy way to create GUI application. And finally detected gesture will be displayed as an output in text format.

RCNN Algorithm:

Faster RCNN is the modified version of Fast RCNN. The major difference between them is that Fast RCNN uses selective search for generating Regions of Interest, while Faster RCNN uses “Region Proposal Network”, aka RPN. RPN takes image feature maps as an input and generates a set of object proposals, each with an objectness score as output.

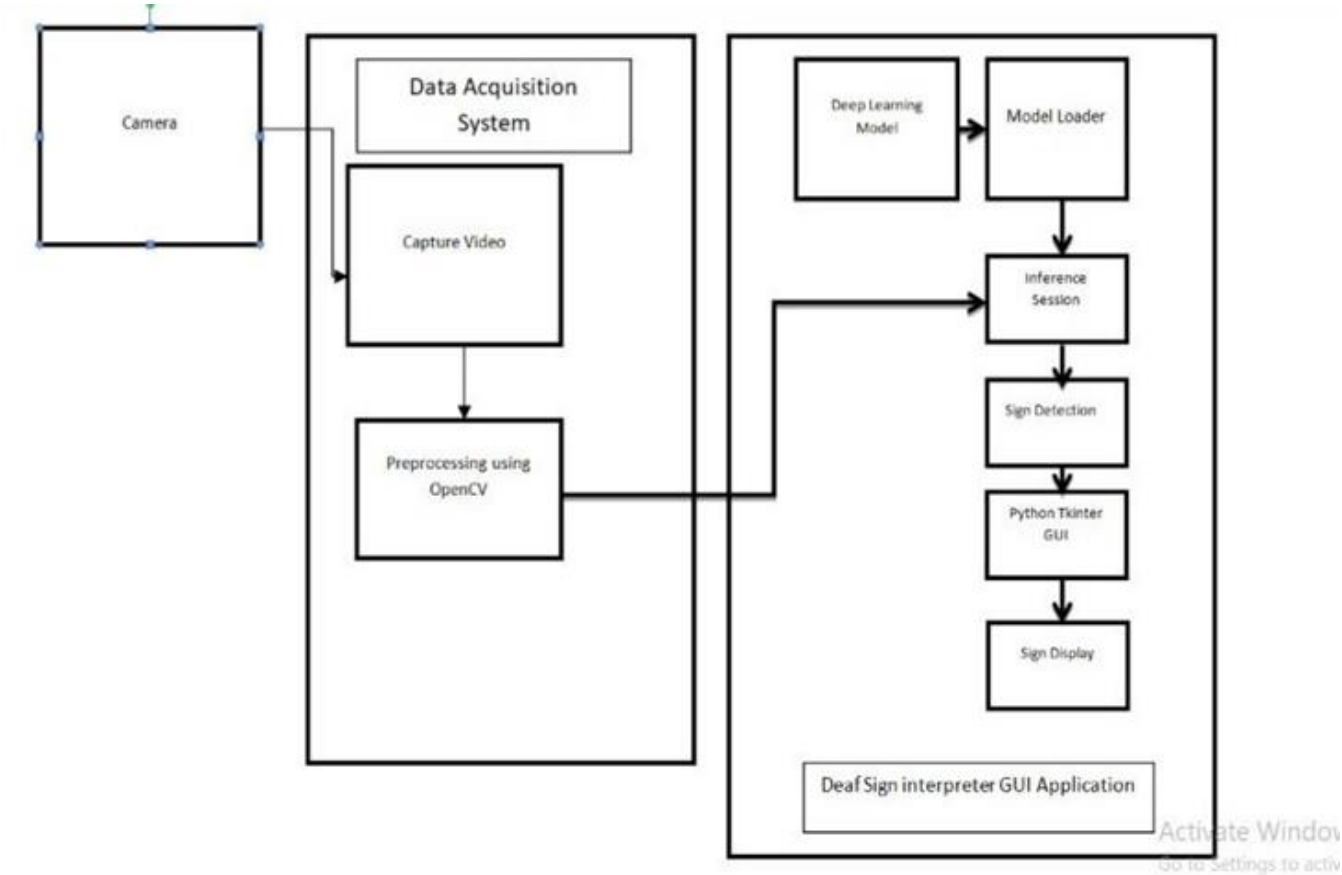


Fig.1 System Architecture

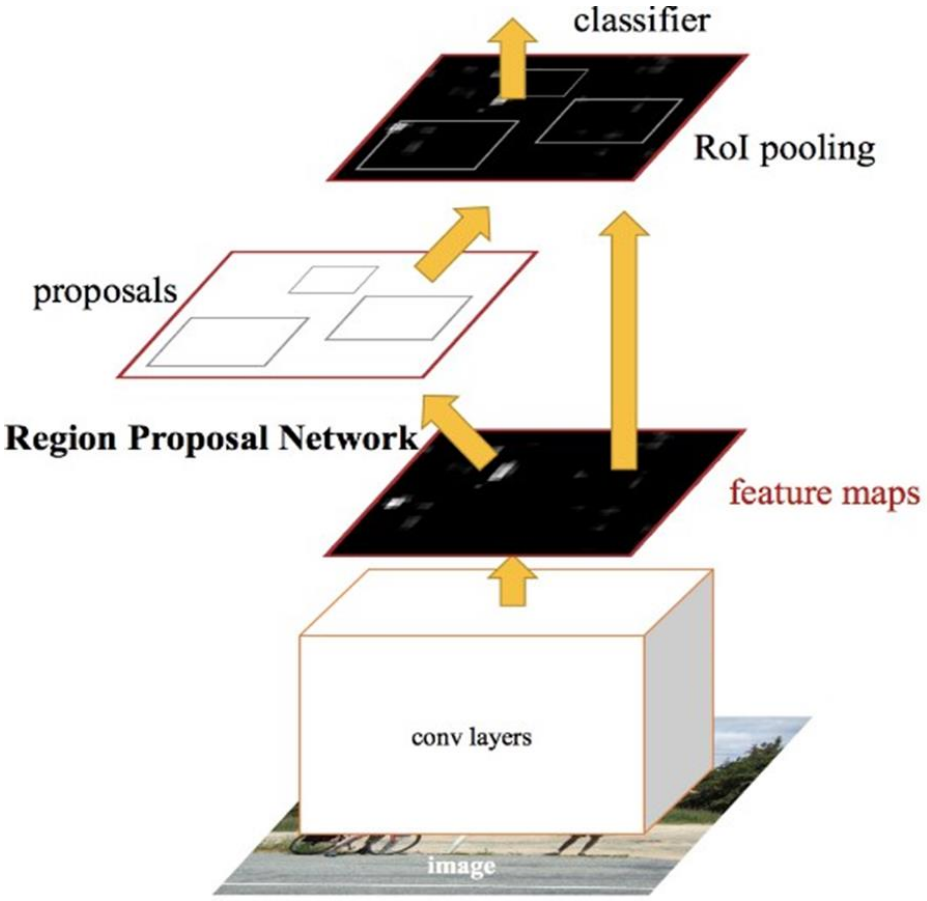


Fig.2 Faster RCNN Working

The below steps are typically followed in a Faster RCNN approach:

- We take an image as input and pass it to the ConvNet which returns the feature map for that image.
- Region proposal network is applied on these feature maps. This returns the object proposals along with their objectness score.
- A ROI pooling layer is applied on these proposals to bring down all the proposals to the same size.
- Finally, the proposals are passed to a fully connected layer which has a softmax layer and a linear regression layer at its top, to classify and output the bounding boxes for objects.

B. Implementation

The below diagram shows the detailed components of the system model.

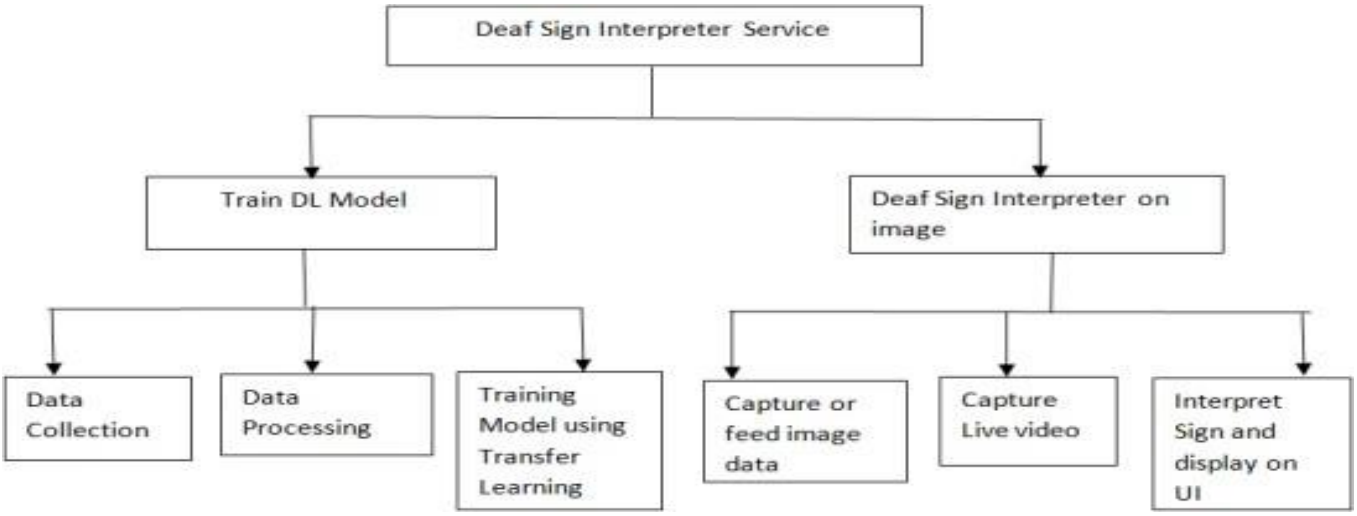


Fig.3 Data Decomposition Diagram

The different components of the system are:

Train DL model

The first step in sign language detection is the training of DL model. The pre-trained model is trained and then once the model is trained the successful inference on the model can be performed by interacting with the deep learning model to display the results. The training of DL model involves following functions:

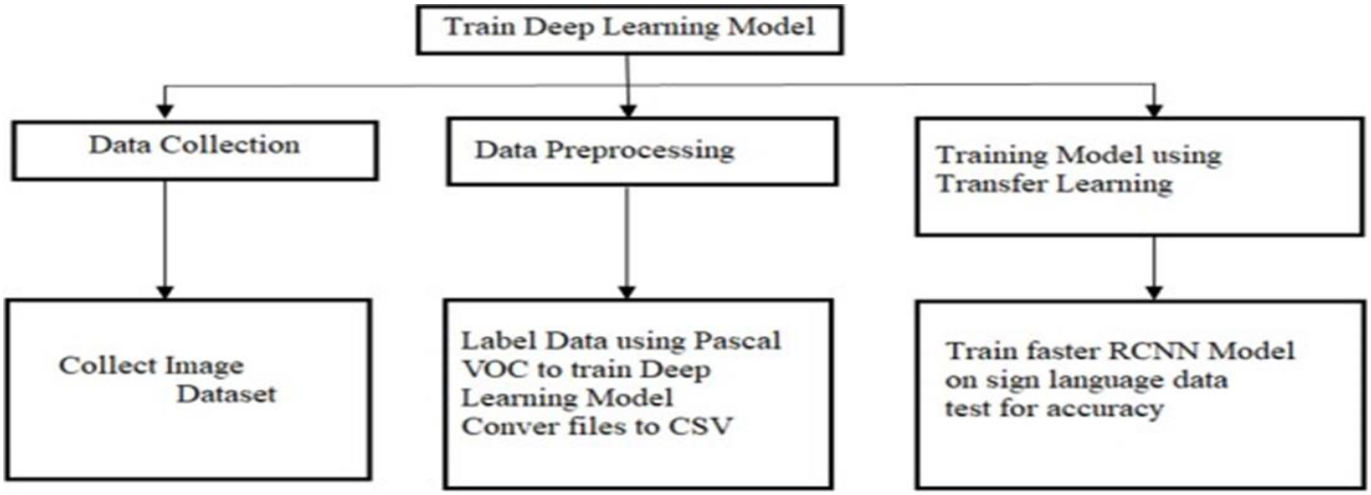


Fig.4 Training Deep Learning Model

- Data collection: Since we are using the transfer learning approach to train the existing model, we require data. The data collection involves image data of the number of sign language samples to train the model. We collect the data samples of each sign language instance and then used the same data for training the model. Once the data is collected the data is split into train and test prior to preparing the data for training.
- Data Pre-processing: The data preprocessing involves resizing and labeling the image data based on the deep learning network chosen for transfer learning. All the images in the test and train are resized and the sign are labeled properly using the Labeling Software. The labeled files are then converted into CSV which will be used generate the train record and test record files used for training the DL model
- Training the model using transfer learning approach: Once the dataset required for training is ready we need to train the model using transfer learning approach. The RCNN network architecture is chosen and the model is retrained on the deaf sign data for detecting of inference the deaf signs. Once the model training is complete the results are tested for accuracy. If accuracy is missing the parameters for training are fine tuned and the model training is repeated till we get accurate results or till the loss function remains constant.
- Deaf Sign interpretation APP: Once the trained model is ready the python app for deaf sign interpretation will be designed. The python app will load the image of video data and perform deaf sign detection and plot the results on image or video feed. The functions of the deaf sign interpreter app are.

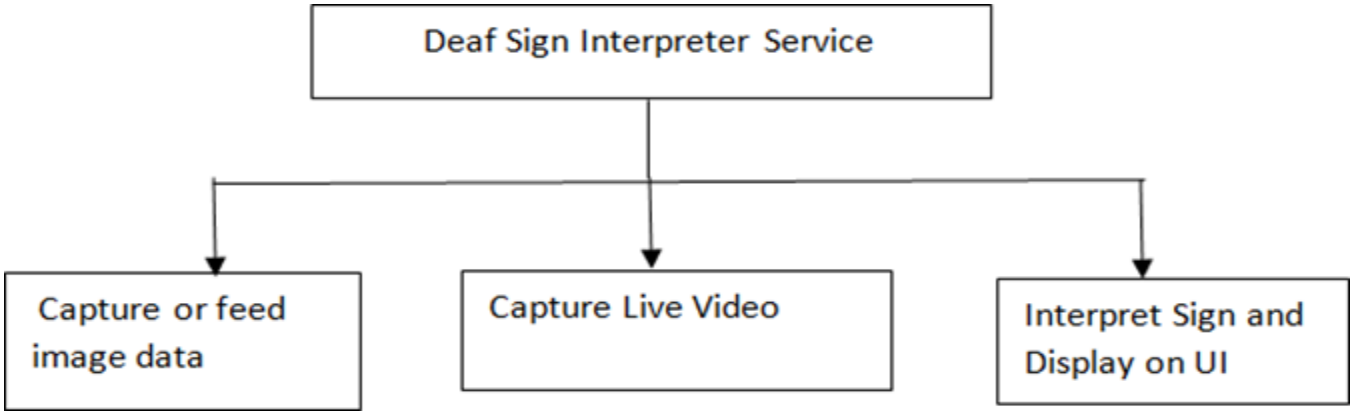


Fig.5 Deaf Sign interpretation APP

- Capture Image or feed image: The image to be interpreted can be fed using this module.
- Capture Live video: The video can be fed or live video can be directly processed using deaf sign interpreter service.
- Interpret the Sign and display on UI: The input image or video split into frame by frame will be fed to the trained model. If sign is recognized the model returns detected sign, occurrence, confidence and dictionary of coordinates of occurrence. This will be plotted on the Output UI image or video feed using Python OpenCV.

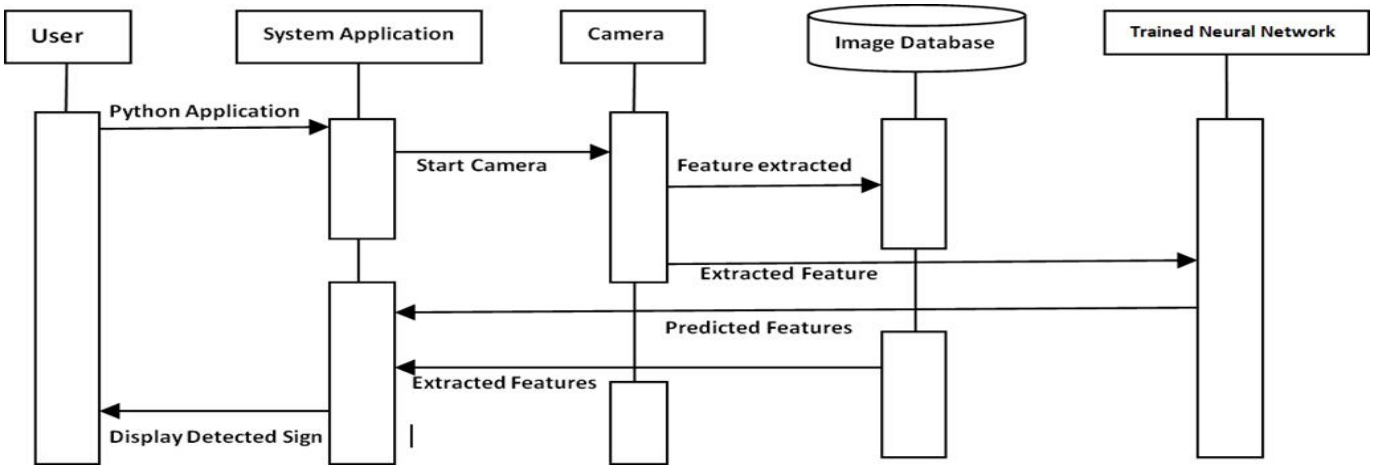


Fig. 6 Sequence Diagram

User first opens the python application, starts the camera and captures the image of hand gestures made. Then the hand gesture features are scaled and extracted by the algorithm. The extracted features are taken by the trained model to predict the output. After predicting the output the sign text will be displayed on the screen.
User first opens the python application, starts the camera and captures the image of hand gestures made. Then the hand gesture features are scaled and extracted by the algorithm. The extracted features are taken by the trained model to predict the output. After predicting the output the sign text will be displayed on the screen.
The algorithms of the each module are given below:

Data Collection:

Step 1 : Start
Step 2 : Collect real-time images as Dataset.
Step 3 : Split the Dataset into 80% training set and 20% testing set.
Step 4 : End

Image labeling:

Step 1 : Start
Step 2 : Label data using Pascal VOC to Train Deep Learning Model.
Step 3 : Convert files to CSV.
Step 4 : End

Symbol Detection:

Step 1 : Start
Step 2 : We take an image as input and pass it to the ConvNet which returns the feature map for that image.
Step 3 : Region proposal network is applied on these feature maps. This returns the object proposals along with their objectness score
Step 4 : A ROI pooling layer is applied on these proposals to bring down all the proposals to the same size.
Step 5 : Finally, the proposals are passed to a fully connected layer which has a softmax layer and a linear regression layer at its top, to classify and output the bounding boxes for objects.

User Interaction:

Step 1 : Start

Step 2 : Capture or feed image data.

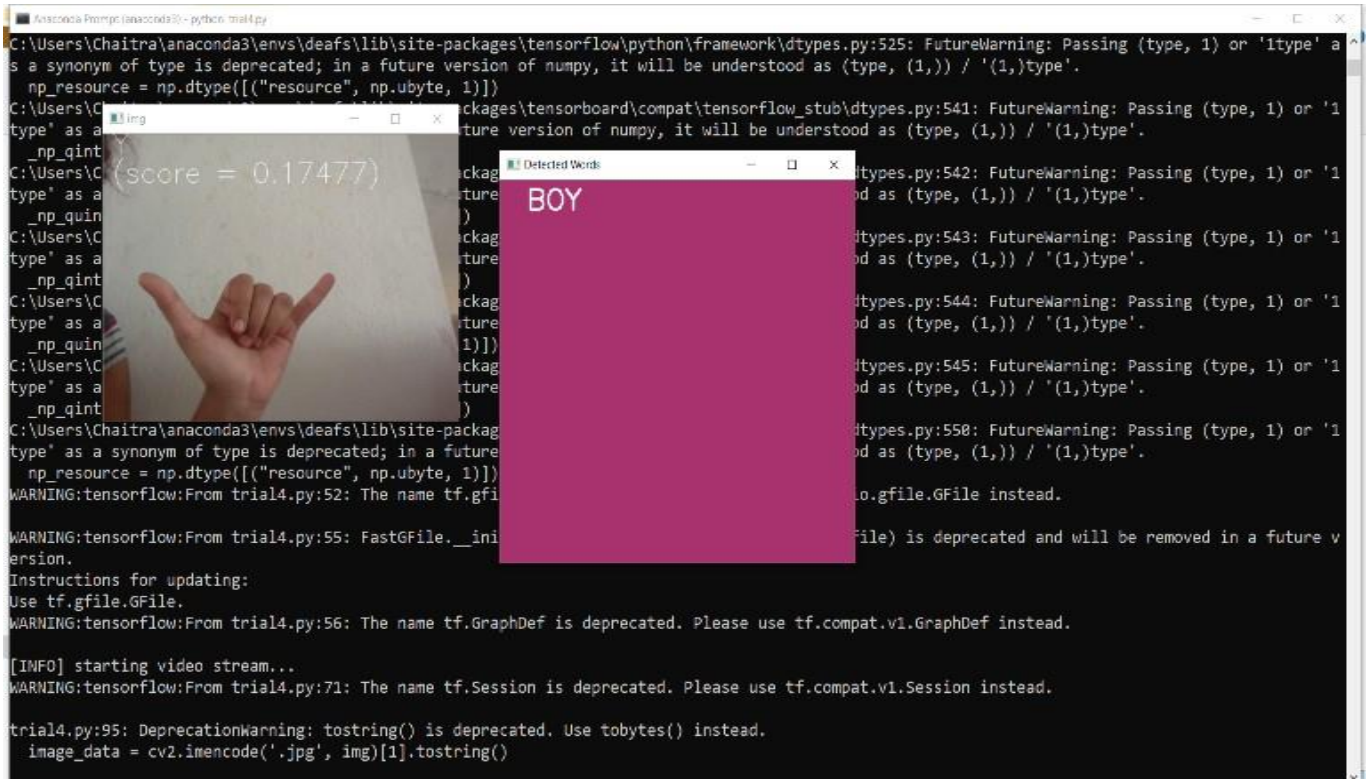
Step 3 : Capture Live Video.

Step 4 : Interpret Sign and Display on UI.

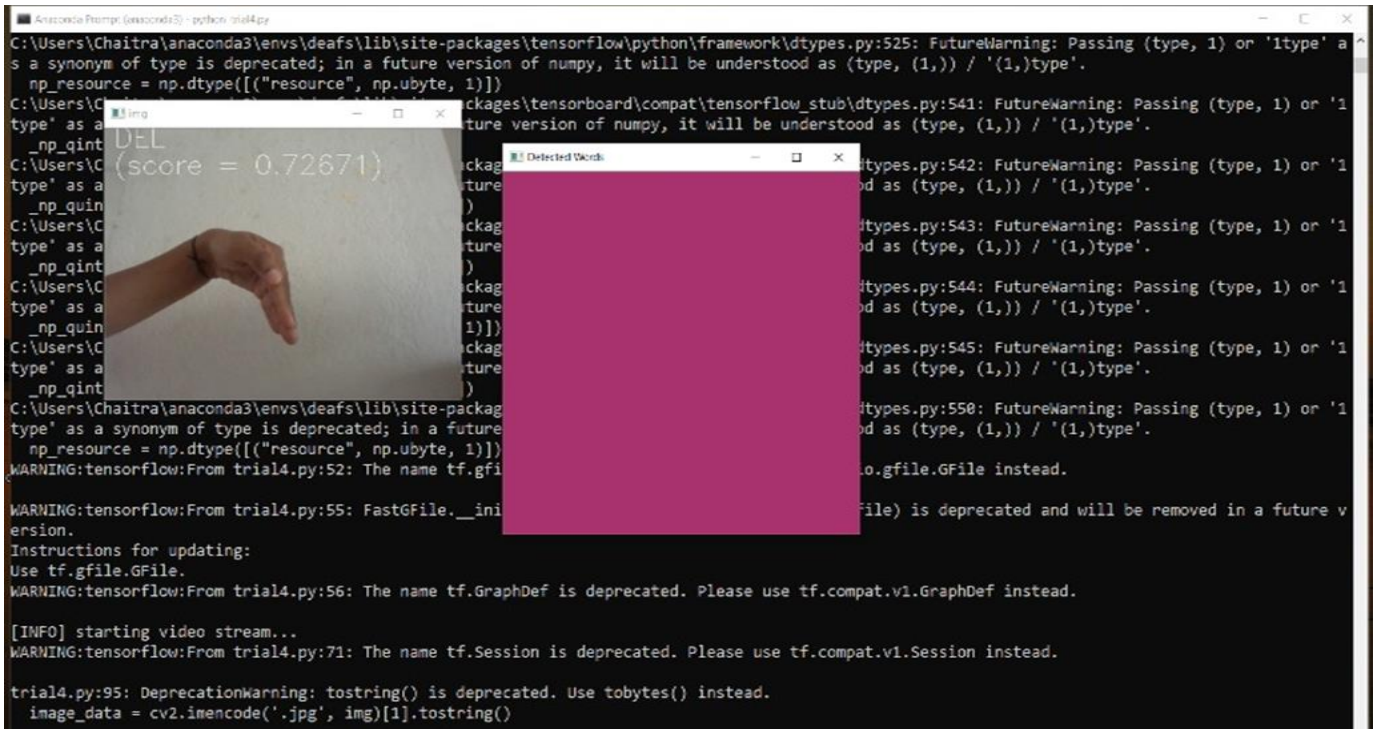
Step 5 : End

IV. RESULTS

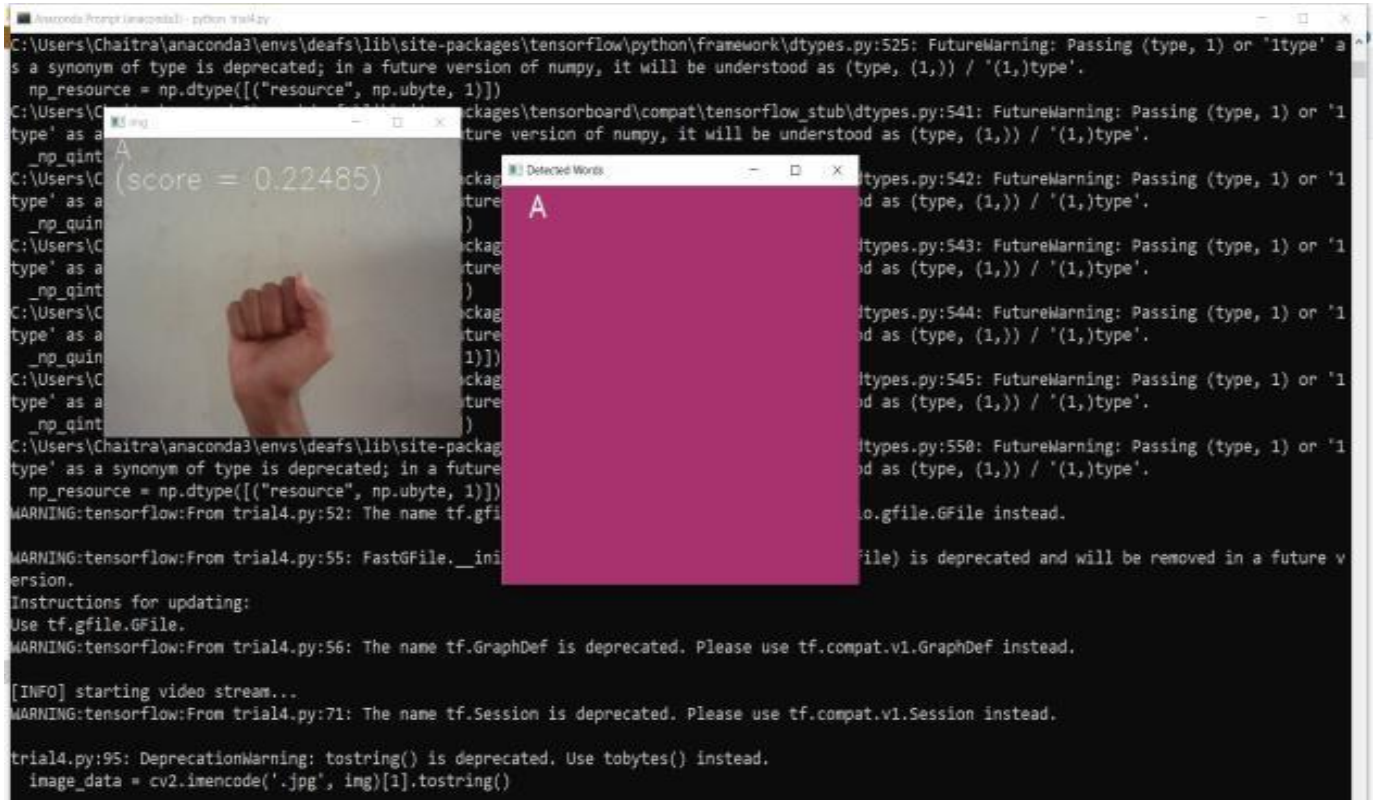
Below are some snapshots which include the human interface design with the computer.



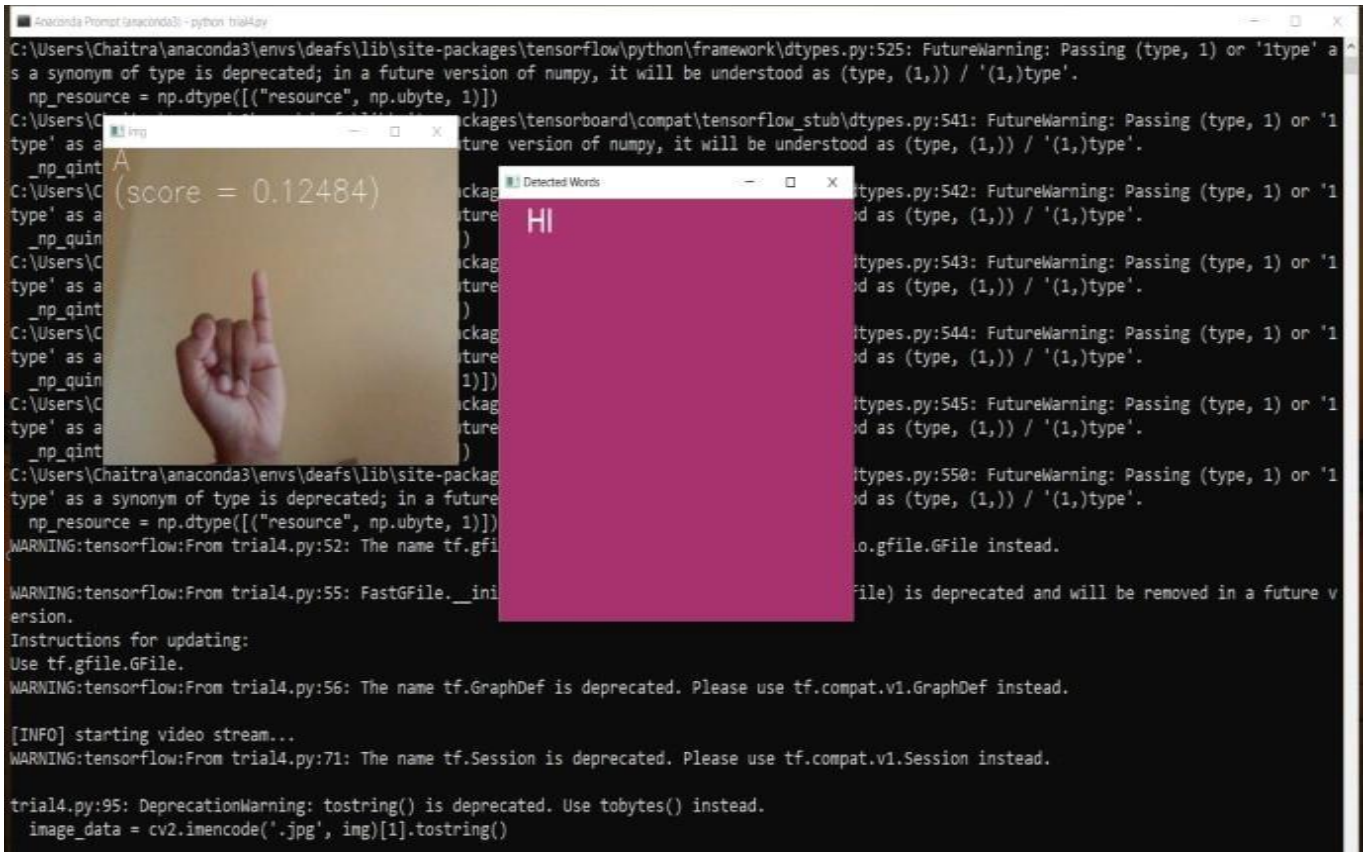
In above snapshot we want to display the word BOY so we used the signs and after some trail we displayed the word BOY successfully.



In the above snapshot the Delete sign will be detected.



In the above snapshot we can see that the letter A will be displayed by using the sign language.



In the above snapshot the word HI will be displayed by using the sign Language.

The accuracy of recognizing the symbol is 85% and in future this can be improved by training the model with larger dataset.

V. CONCLUSION AND FUTURE WORK

The proposed system not only can help the deaf and dumb to communicate easily but also avoid the necessity of person with the knowledge of sign language who works as an interpreter between the deaf and dumb people and common people. Thus our proposed solution using deep learning is user friendly to operate using the developed python application and can be scaled anytime by altering the training data.

This work has wide scope for future modification. In Future this work can be modified for accuracy by training on huge datasets as machine learning is data hungry. More the data, more the accuracy. At present we have worked on only basic sign letters. It can be further used to train on different sign languages and can be modified.

VI. REFERENCES

[1]. Sunitha K. A, Anitha Saraswathi.P, Aarthi.M, Jayapriya. K, Lingam Sunny, "Deaf Mute Communication Interpreter- A Review", International Journal of Applied Engineering Research ,Volume 11, pp 290-296 , 2016.

[2]. Mathavan Suresh Anand, Nagarajan Mohan Kumar, Angappan Kumaresan, " An Efficient Framework for Indian SignLanguage Recognition Using Wavelet Transform" Circuits and Systems, Volume 7, pp 1874-1883, 2016.

[3]. Mandeep Kaur Ahuja, Amardeep Singh, "Hand Gesture Recognition Using PCA", International Journal of Computer Science Engineering and Technology (IJCSET), Volume 5, Issue 7, pp. 267-27, July 2015.
[4]. Sagar P.More, Prof. Abdul Sattar, "Hand gesture recognition system for dumb people",
[5]. International Journal of Science and Research (IJSR)
[6]. Chandandeep Kaur, Nivit Gill, "An Automated System for Indian Sign Language Recognition", International Journal of Advanced Research in Computer Science and Software Engineering.
[7]. Pratibha Pandey, Vinay Jain, "Hand Gesture Recognition for Sign Language Recognition: A Review", International Journal of Science, Engineering and Technology Research (IJSETR), Volume 4, Issue 3, March 2015 .
[8]. Nakul Nagpal,Dr. Arun Mitra.,Dr. Pankaj Agrawal, "Design Issue and Proposed Implementation of Communication Aid for Deaf & Dumb People", International Journal on Recent and Innovation Trends in Computing and Communication ,Volume: 3 Issue: 5,pp- 147 – 149.
[9]. Neelam K. Gilorkar, Manisha M. Ingle, "Real Time Detection And Recognition Of Indian And American Sign Language Using Sift", International Journal of Electronics and Communication Engineering & Technology (IJECET), Volume 5, Issue 5, pp. 11-18 , May 2014
[10]. Neelam K. Gilorkar, Manisha M. Ingle, "A Review on Feature Extraction for Indian and American Sign Language", International Journal of Computer Science and Information Technologies, Volume 5 (1) , pp314-318, 2014.
[11]. Ashish Sethi, Hemanth , Kuldeep Kumar,Bhaskara Rao ,Krishnan R, "Sign Pro-An Application Suite for Deaf and Dumb", IJCSET , Volume 2, Issue 5, pp-1203-1206, May 2012.
[12]. Priyanka Sharma,"Offline Signature Verification Using Surf Feature Extraction and Neural Networks Approach", International Journal of Computer Science and Information Technologies, Volume 5 (3) , pp 3539-3541, 2014.