

# Sign Language To Text Converter System

Athira Jayakumar<sup>1</sup>, Aparna Nambiar<sup>2</sup>, Arya Anilkumar<sup>3</sup>, Sreelekshmi Menon<sup>4</sup>, Prof. Shubhangi Chavan<sup>5</sup>

<sup>1,2,3,4,5</sup> Dept of Computer Engineering  
<sup>1,2,3,4,5</sup> Pillai College Of Engineering, Panvel, Navi Mumbai

**Abstract-** *The Sign Language is the primary medium used by speech or hearing impaired people so as to communicate with rest of their peers. Sign language is how specially abled persons express their emotions. In this project we will create an interface where the video input will be taken and hand gestures will be recognized to find the hand region by eliminating all other unwanted portion in video with the help of OpenCV. In the first part we make a classifier model using the Keras implementation of convolutional neural network using python [1]. Based on this approach the results are predicted and the word is obtained. The Word Segmentation module that includes techniques like dividing the text, parsing, scoring and segmenting will be used in the next phase to obtain the final textual output which corresponds to the sign provided as input. There are almost 138 to 300 sign languages used across the globe. We would be using the American Sign Language(ASL) that uses single handed gestures as the input and generate the corresponding letters or text in English.*

**Keywords-** OpenCV, CNN, Word Segmentation, Parsing

## I. INTRODUCTION

The project will create an interface where the video input is taken and hand gestures will be recognized . OpenCV is used to find the hand region by eliminating all other unwanted portions in the video and it will be further processed and stored. The second phase of the project will focus upon the NLP techniques. The word segmentation module provided by natural language processing which includes techniques like parsing, scoring and segmenting is used to obtain the final output which is the text that corresponds to the language. American Sign Language (ASL) is used and that uses single hand gestures as the input and generate corresponding text or letters in English language. A deep learning approach will be used which can classify the sign using the convolutional neural network. In the first part we make a classifier model using the Keras implementation of convolutional neural network (cnn) using python . In the next phase another real-time system which used skin segmentation to find the Region of Interest in the frame which shows the bounding box. The segmented region is fed to the classifier model to predict the sign [1].

## II. LITERATURE SURVEY

### A. Feature extraction and Selection

It is observed that the authors have followed the classical machine-learning workflow. Each of the 12 datasets used was randomly shuffled 10 times. The differences between the results were quantified using statistical methods on the 12 databases, 10 shuffles, 4 folds, 12 classifiers which gives a total of 5760 measurements. In 8 cases of classifiers out 12, the greedy search leads to the worst classification results. The largest PCC values were obtained by evolutionary algorithms, reranking search, linear forward selection, the tabular search[7].

### B. Digital image processing

The research team observed that large number of image processing applications, tools and techniques helps to extract complex features of an image. Image processing works on single dimensional image to multidimensional and see what actually in the image. Image processing is the really important for developing technologies in the real time aspect. This paper discusses the overview of an image processing applications, tools and techniques. But OpenCv is the optimal tool used as it provides all the basic algorithm[9].

### C. Parsers and Parsing Approaches

It is observed that the dependency parsers constructed using the dependency parsing technique can be quite helpful as can serve as resources for various research purposes. In this method the sentence is given as input to the system and an output is produced which is basically a dependency tree. The Paninian framework is the basis that can be used for parsing English language. It can also be understood that for the Indian languages the same framework can be used however with much better accuracy results. From the research of the authors it is also observed that Minipar Parser performs worse as far as the quality of the results are concerned while the MST parser is very complex machine learning approach. The major advantage of this framework is that it offers good performance and produce all the grammatical information of the sentences. The major advantage of the

dependency parser is that it resolves ambiguity to a great extent[4].

#### D. Sign Language using Convolutional Neural Network

CNNs takes its inspiration from the visual cortex of the human brain. The artificial neurons connects to a region called a receptive field. This is obtained by performing discrete convolutions on the image with filter values as weights that are trainable. More than one filters are applied for each channel, and together with the activation functions of the neurons, they form what is called a feature maps. This is followed by a scheme of pooling, where only the interesting information of the feature maps are clubbed together. These techniques are performed in multiple layers[1].

#### E. Extraction Of Features from an image

Feature extraction is finished once the preprocessing in character recognition system. the first task of pattern recognition is to require Associate in Nursing input pattern and properly assign it as one of the potential output categories. This method may be categorized into 2 general stages: Feature selection and Classification. It is vital to the entire process since the classifier won't be able to acknowledge from poorly hand-picked options. As patterns have totally different orientation, styles etc., numerous image preprocessing techniques are applied first off. In this paper, a quick overview of feature extraction techniques could also be taken and it may be determined that that feature extraction technique can be higher for the work to be done supported complexity, type of image (e.g. grey, color image)[6].

### III. PROPOSED WORK

The process of taking in the hand gestures as inputs requires algorithms that produces an accurate output that is the corresponding text to the hand gesture. The proposed architecture is a six step process that starts from taking in input followed by the preprocessing state and the NLP techniques to obtain the final textual output.

#### 3.1 System Architecture

The system architecture is given in Figure 1. Each block is described in this Section.

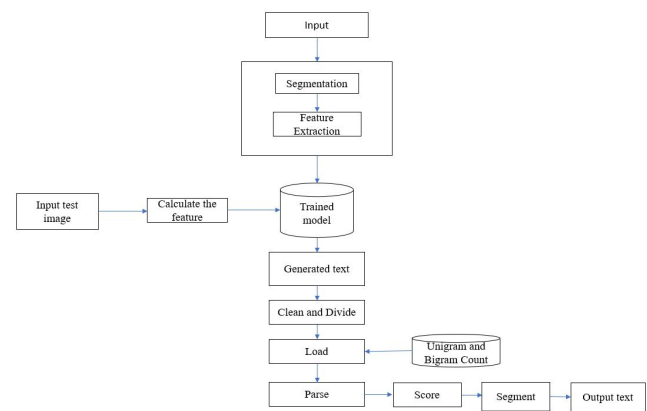


Fig. 1 Proposed system architecture

**A. Input :** The first step is to provide the input sign on which all the further processes need to be carried out to obtain the desired result.

**B. Preprocessing:** Preprocessing is concerned with the task of removing unwanted noises from the input provided. The preprocessing consists of Segmentation, Feature Extraction and Image Detection. Segmentation refers to the process of partitioning the images into different segments. Here segmentation includes separation of the background and other such undesired parts from the captured images. Preprocessing is concerned with the task of removing unwanted noises from the input provided. The preprocessing consists of Segmentation and Feature Extraction. Segmentation refers to the process of partitioning the images into different segments. Here segmentation includes separation of the background and other such undesired parts from the captured images.

**C. Image Classification:** After the preprocessing stage, the image that is obtained is compared with the images present in the sample gestures dataset. The dataset contains multiple images for the signs. The image generated is compared with these images. If it matches with the image already stored, it can be carried ahead for further processing, otherwise the generated image would be discarded. The CNN classifier is trained and used in this procedure.

**D. Clean and Divide:** Now to obtain proper final output, NLP techniques are used. Word segmentation is done to obtain the proper output. The word segment library in python is used here. It has the function clean in it that eliminates all non-alphanumeric characters. The divide function yields the prefix and suffix pairs from the text with length of prefix not exceeding the limit.

**E. Load:** The load function loads all the unigram and bigram counts from the disk. The unigram data includes the most commonly used 333,000 words. Similarly, bigram data

includes only the most commonly used 250,000 phrases. Every word and phrase is lowercased and the punctuations are removed.

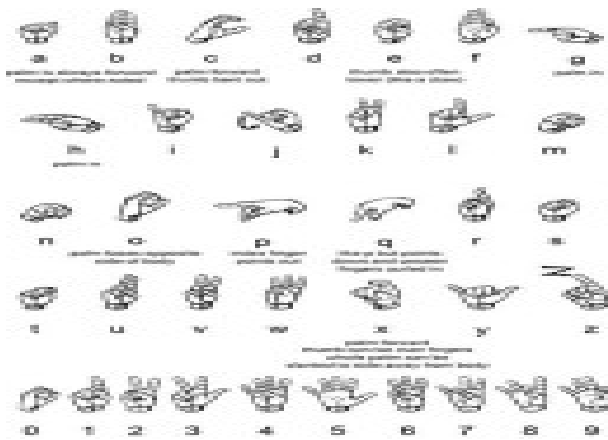
**F. Parse:** In this stage parsing is done. This function reads and parses the tab separated file word and count pairs.

**G. Score:** Scoring the word is done in the context of the previous word.

**H. Segment:** Finally a list of words is returned that is the best segmentation of the text that is generated.

**I. Output:** The segmented words are taken as the meaningful output text.

#### IV. DATASET



The dataset for the image classification basically consists of all the alphabets from A to Z.

#### V. ACKNOWLEDGMENT

It is our privilege to precise our sincerest regards to our guide Prof. Shubhangi Chavan for the important inputs, able guidance, encouragement, whole-hearted cooperation and constructive criticism throughout the period of this work. We deeply categorical our sincere due to our Head of the Department Dr. Sharvari Govilkar and our Principal Dr. Sandeep M. Joshi for encouraging and permitting us to present this work.

#### REFERENCES

- [1] Sajanraj T D, Beena M V, "Indian Sign Language Numeral Recognition Using Region of Interest Convolutional Neural Network", (2018).
- [2] Sampada.S.Wazalwar,UrmilaShrawankar,"Interpretation of sign language into English using NLP techniques", (2017)

- [3] Sayli Dixit and Urmila Shrawankar,"Framing Sentences from Sign Language Symbols using NLP ISBN-13:978-93-5260-126-4", (2016)
- [4] Amita Rani, Komal Mehla, Ajay Jangra, "Parsers and Parsing Approaches: Classification and State of the Art", (2015).
- [5] Timothy G. Wade, James D. Wickham, "Mappingspatial patterns with morphological image processing", (2015).
- [6] Gaurav Kumar, Pradeep Kumar Bhatia, "A Detailed Review of Feature Extraction in Image Processing Systems", (2014)
- [7] Madalina Cosmina Popescu, Lucian Mircea Sasu, "Feature Extraction, Feature selection and Machine Learning for Image Classification: A case study", (2014)
- [8] ElisamargarethSibarani, MhD. Nadial, EvyPanggabean, "A Study of Parsing Process on Natural Language processing", (2013)
- [9] Dr V. Arulmozhi, "Digital Image Processing-A Quick Review", (2013)
- [10] <https://textprocessing.org/getting-started-with-python-word-segmentation>