

# Sign Language Recognition System

Akriti Goyal, Deepanshu Dhar, Paras A Nair, Chirag Saini, Supreetha S M, Chetana Prakash  
Department of Computer Science and Engineering,  
BIET Davangere

**Abstract**— Communication plays one of the important role in survival within the society. People with ability to speak and hear are privileged one. By being one of a privileged it is one's responsibility to help such people. This is a paper regarding a process or way that will help deaf and dumb person to communicate with normal people and gives an opportunity to investigate a new area of technology.

**Keywords**— Convolutional neural network, skin segmentation, Indian sign language.

## I. INTRODUCTION

Humans are living in a society as a community because they are very similar when comes to the way of living, way of communicating, way of expressing etc. But after a glance on people, it was found that some are not as privileged as all of the others. Those are deaf and dumb people. These people have a large gap with normal speaking people under some circumstances such as communication and expression.

Various works have been done using the technologies that are available but still there has not been any work which may revive the communication gap among people. There are several models made in different part of regions according to the requirement of the people such as Arabic sign language, British sign language (BSL), American sign language (ASL), Turkish sign language (TSL), etc. Indian sign language also sometimes called as Indo- Pak sign language have also been developed but there is much limitations to it.

### Problem statement:

Not a proper system has been developed in order to resolve the problem of people with speaking disorder and deaf. Also multiple videos are available on websites that are not too helpful and moreover there is very less dual system (sign to speech and vice versa) developed.

### Objective:

- To aim towards developing a proper Indian sign language with all sets of alphabets and numbers.
- To carry out two-way communication. The spoken words will be taken as input and signs will be given as output.
- To investigate and implement different machine learning classifiers like Logistic Regression, Naïve Bayes classifier, K-nearest neighbours (KNN), support vector machine (SVM) and a neural network technique Convolution Neural Networks (CNN).

### Proposed system:

A perfectly working model idea for Indian sign language, that includes all the alphabets and digits. It also includes dual mode of output that is, sign language converted into text/speech and speech converted in sign.

## II. RELATED WORK

Sawant Pramada, Deshpande Saylee, Nale Pranita [1]: This paper gives different idea about pre-processing of image. It introduces coordinate mapping for skin segmentation. This paper is based on a fast and efficient method of coordinate mapping. After converting RGB image to grey scale the number of fingers will be taken in consideration. We don't need a perfect angle of camera but instead we need a clear image which can show the number of fingers present in it. The image of open hand will be taken and the tip of fingers will be converted to white colour. Then tips of all the fingers will be given a coordinate point. The position and the number of points in image will be used to get the similarity among the image and will also help during the training.

Er. Aditi Kalsh, Dr. N.S. Garewal [2]: This paper uses edge detection technique. This technique is old but one of most accurate technique. Error rate is found very low in this technique. Canny edge detection is a widely used method due to its ease of use. The jump and changes in edges is very easy to detect and recognise. By finding edge in any image it is easy to use remove all extra background and get a definite shape of image. This gives a benefit of fast learning.

Christopher Lee and Yangsheng Xu [3]: A glove based research that was developed for first time, recognised 14 letters from alphabets. Later there were many gloves developed that could recognise all the letters from alphabets. There were many limitations to this method as their new users faced recognition problem, also the movement of hand was a problem as there were many wires connected to it.

Sanil Jain, Kadi Vinay Sameer Raja [4]: This paper introduced use of videos of hand signs instead of instant camera capturing hand images that to be taken as input. From the video the frame will be extracted and then there will be feature extraction of the obtained frame. Then extracted images were given for training.

## III. SYSTEM DIAGRAM AND WORKFLOW

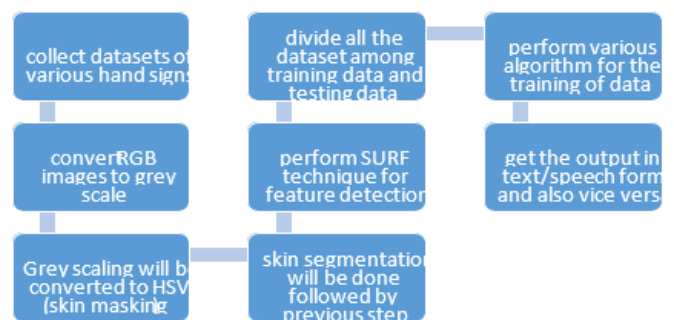


Fig 1: working of system

The above fig 1 refers to the working of the system. At beginning the dataset will be collected. Those datasets will go under various processes before getting into CNN layers. The processes which are required to go through before feeding the image to CNN layer are as follows.

#### a) PRE-PROCESSING OF IMAGE

##### Grey scaling:

Each image is formed by pixels. These pixels are always present with some values (0 to 255) regardless of colours. So it is needed to convert these values to binary numbers. For conversion into binary number (0 and 1), the image should be on scale of black to white i.e. greyish shade.

##### Segmentation:

The main aim of segmentation is to remove the unrequired background. This is done using skin masking. The first step towards skin masking is conversion of image to grey scale followed by conversion of RGB image to HSV. Any image that include colour space is defined by HSV (hue, saturation, value) as shown in fig 2. Hue represents colour, saturation represents shade of grey and value represents no. of shade of the specific colour. As result of this process, the output of the coloured image is given by lower and upper bound of image.

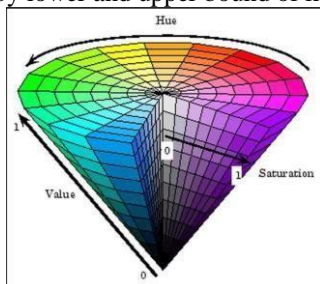


Fig 2: HSV explanation

##### Feature Detection and Extraction:

The SURF (Speeded Up Robust Feature) technique is used to extract descriptors from the segmented hand images as shown in fig 3. It is composed of three steps called feature extraction, feature description and feature matching. It is a fast and accurate algorithm comparison of the image and finding similarities between them.

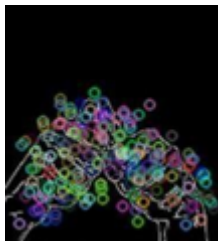


Fig 3: SURF image

#### b) FEEDING THE IMAGE TO CNN LAYER

The image will be feed to CNN layers. These layers extract all the features and use them to train the model. CNN has mainly 3 layers: convolutional layer, pooling layer and fully connected as shown below in fig 4.

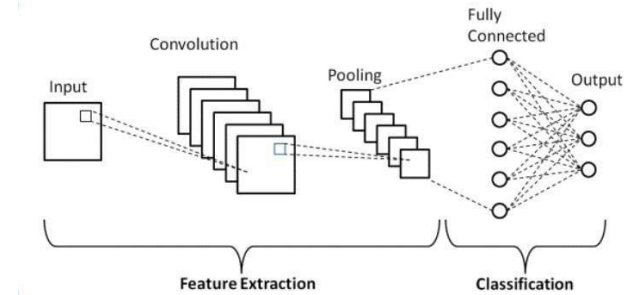


Fig 4: CNN basic architecture

#### c) CLASSIFICATION

For training the data set the required algorithms and classifiers as mentioned below. The classifiers that has to be used are as follows:

**CNN:** It is used for image recognition and classification that means to learn the features and scale of the image in different position. The convolutional layer captures feature like edges, color, gradient orientation etc. The pooling layer is used to reduce the number of parameters(feature) in order to learn. Then locally connected layer accepts input from preceding layer, computes the class scores, and outputs a 1-D array with the same size as the number of classes.

**KNN:** KNN is used for high level of prediction. This is used for feature similarity that means new point is assigned based on how closely it resembles from the points in training set. It is a famous method used for image classification. It uses  $k=1$  and Euclidean-distance to get smallest data process which is used for ranking.

**Naïve Bayes classifier:** Naïve Byes classifier helps in building the fast machine learning model that can make quick predictions. It makes use of a language model to assign class labels to some instances, based on a set of features which can be numerically represented using statistical techniques.

**Support vector machine:** SVM is a set of supervised learning methods that are used for classification, regression and outlier's detection. Here all data items are plotted in n dimensional plotting as a point where n is the quantity of highlights with the value of every component being the value of that specific coordinator. At that point, classification is performed by finding hyperplane that isolates and separates the classes.

**Logistic regression:** logistic regression estimates the parameters of logistic model. It is used to understand the relationship dependent and independent variable.

##### Pattern matching and getting text

Training of the dataset will be done in the same way as explained above. When it is the turn of testing, an image of some sign using hand will be given to camera. If the pattern gets matched, it will soon give the output as text/speech and also vice versa.

#### IV. RESULT AND DISCUSSION

The use of above methods and implementation of several classifiers gives us the following results as mentioned in below images. Signs are given as input through camera and text is given as output. As mentioned above in the paper, dual mode of communication implemented. Spoken words are given as input and the sign are taken as output using google speech API.

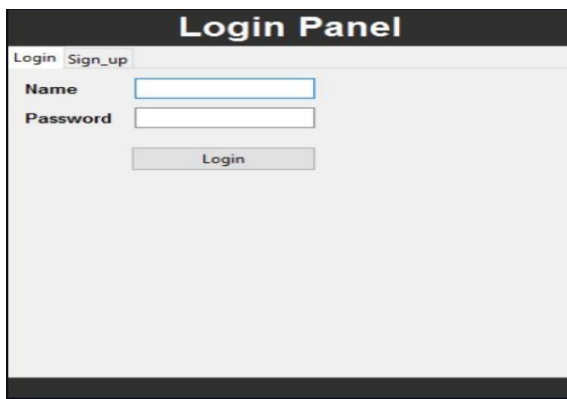


Fig 5: Login box

In the beginning, a login box will be opened as shown above in fig 5. The login panel consist of login option with name and password area. For the new user sign up option is also provided.

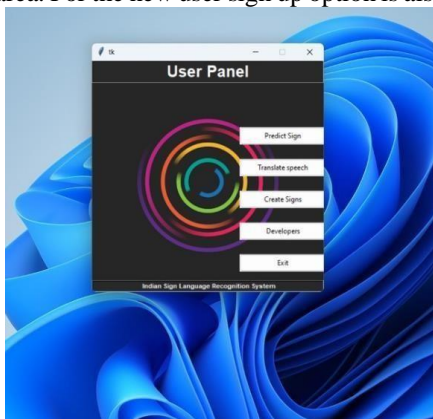


Fig 6: User box

Followed by login panel, there is a user panel that opens with multiple options as referred in fig no 6.

The panel consists of 5 options namely predict sign, translate speech, create sign, developers and exit.

By selecting translate speech option from user panel (shown above), spoken word will be translated to sign. For testing, a word "open" is taken. When the word open is spoken, the sign of each letter of that word appears one by one on the screen. Here is the output of word shown below, consisting of fig no. 7,8, 9 and 10.

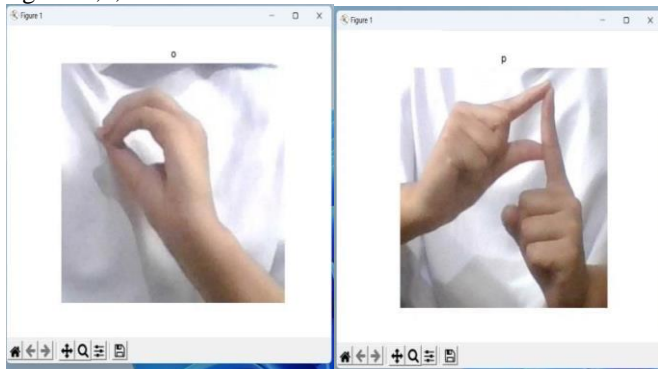


Fig 7: Alphabet O

Fig 8: Alphabet P

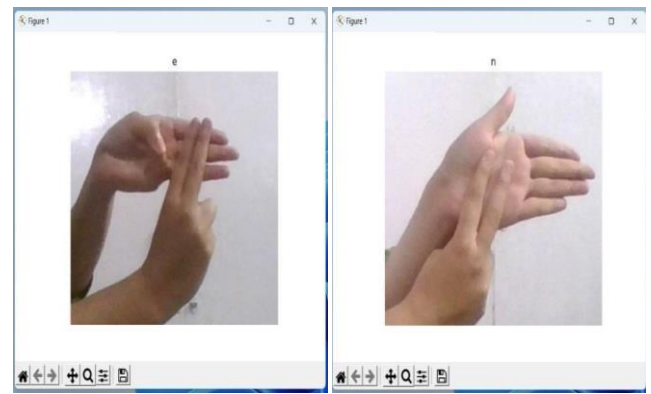


Fig 9: Alphabet E

Fig 10: Alphabe

## V. CONCLUSION

This paper bridges the gap between normal people and deaf/dumb people by providing a new means of communication. It gives an opportunity to work and implement one's knowledge in order to help less fortunate using advanced technology. An attempt is made to work on a model for Indian sign language, that includes all the alphabets and digits. An incredible use of all the techniques is also done in order to provide the best and suitable information and implementation.

## REFERENCES

- [1] [Sanjay Meena. A Study of Hand Gesture Recognition Technique. Master Thesis, Department of Electronics and Communication Engineering, National Institute of Technology, India.
- [2] Olena Lomakina. Development of Effective Gesture Recognition System. TCSET'2012, Lviv-Slavske, Ukraine.
- [3] P. Subha Rajam and Dr. G. Balakrishnan. Real Time Indian Sign Language Recognition System to aid Deaf-dumb People. ICCT, IEEE.
- [4] Wilson A.D, Bobick A.F. Learning visual behavior for gesture analysis. In Proc. IEEE Symposium on Computer Vision.
- [5] <http://www.happinesspages.com/baby-sign-language-FAQ.html>
- [6] <http://mi.eng.cam.ac.uk/~cipolla/lectures/PartIB/old/IB-visualcodebook.pdf>
- [7] <https://github.com/shackenberg/Minimal-Bag-of-Visual-Words-Image-Classifer/blob/master/sift3>.
- [8] <http://en.wikipedia.org/wiki/YIQ>
- [9] <http://en.wikipedia.org/wiki/YUV>
- [10] Das, A., Gawde, S., Surattwala, K., Kalbande, D. "Sign Language Recognition Using Deep Learning on Custom Processed Static Gesture Images". 2018 International Conference on Smart City and Emerging Technology (ICSCET).
- [11] Rao, G. A., Syamala, K., Kishore, P. V. V., Sastry, A. S. C. S. "Deep convolutional neural networks for sign language recognition".
- [12] Mahesh Kumar NB. "Conversion of Sign Language into Text". International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 9.
- [13] Bantupalli, K, Xie, Y. (2018). "American Sign Language Recognition using Deep Learning and Computer Vision". IEEE International Conference on Big Data (Big Data)
- [14] Kumar, A., Thankachan, K., & Dominic, M. M. "Sign language recognition". 3rd International Conference on Recent Advances in Information Technology (RAIT).
- [15] Pankajakshan, P. C., & Thilagavathi B.. Sign language recognition system. International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)