THINK INDIA JOURNAL

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

Indian Sign Language Recognition System

Kanha Gupta[1], Beesetty Yogendra[2] , Shalini[3], *Awadhesh Kumar Shukla[4], Sripada Manasa Lakshmi[5]

Student[(1,2,3)], Assistant Professor[(4,5)]

Lovely Professional University, Punjab, India[(1,2,3,4,5)]

Kanhagupta66@gmail.com[1],123balaanand@gmail.com[2], shalinikssingh4@gmail.com[3], awadhesh.23866@lpu.co.in[4], manasa.23777@lpu.co.in[5]

Corresponding Author :- Awadhesh Kumar Shukla

Abstract-Sign Language is a language that is used by deaf and dumb people to convey their meaning to others. Sign language can be considered a very versatile language because it consists of a combination of various hand shapes, hand movements, including arms and body. Our system is trained to recognize Indian sign alphabets which are the building blocks of words that form a sentence. Our paper proposes a system that will help deaf and dumb people to communicate with people as these people consist of 5% of the population of the world which is 466 million people of which 34 million are children. Our system uses a camera to access your live image and gives the trained machine your images that check and gives the sign you have shown. It also uses text to speech to tell us.

Keywords- Sign language, PCA, Machine Learning

1. Introduction

Society of deaf and dumb having a medium of communication by sign language. This will help them in the connection between humans and computers. Traditional humans will simply act and communicate with each other, however, the person with hearing and speaking disabilities face issues in the act with alternative hearing folks while not a translator. Typically, the deaf got to look for the assistance of the language interpreter to translate their thoughts to traditional folks and the other way around. However, this method seems to be terribly pricey and doesn't work throughout the life amount of a handicapped person. So, a system which will mechanically acknowledge language becomes a necessity. Introducing such a system would cause minimizing the gap between deaf and traditional folks in society. The language in use at an area depends on the culture and speak the language at that place. ISL could be a commonplace and well-developed manner of communication for deaf folks in Indian nation and speaking in English. Completely different symbols are concerned in numerous alphabets for Indian language. There's no standardized language for all deaf folks across the globe.

In language recognition, there are two methods first one is vision-based and another one is glove based. Within the glove-based methodology, the user should wear a tool that carries a load of

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

cables to attach the device to a pc. Such devices are overpriced and scale back the naturalness of language communication.

The need for the only camera is there in vision based which further directly deal with the hand signs. It's a 2-Step Process: sign capturing and sign analysis. Vision-based ways offer a natural setting to the user and scale back the complications as within the glove-based methodology.

The objective of the planned work is to expeditiously acknowledge the signs of Indian language (ISL) exploitation the vision-based methodology. The signs captured underneath uniform background are thought-about for coaching set construction to avoid alternative skin-colored objects within the background and to form the system additional correct.

2. Literature review

Now all the sign language in Indian sign language is double-handed that's why it is more complexed than American sign language. The model that is chosen for recognizing the Indian sign language is Artificial Neural Networks (ANN), Support Vector Machines (SVM), Hidden Markov Model (HMM). All the research work done by using this model is a static sign of Indian sign language controlled under some conditions. Some of the systems use an only dark background. Some systems use only a bare hand and only one-color background.
A facial sign is not considered in any system yet. However only hand sign is included for better results. The researcher should focus on mainly different backgrounds, different signs based on hand size, some differences in the distance while capturing the signs, lightning condition and color features.[3]

Indian Sign Language has been an emerging field in artificial Intelligence. The main difficulty faced by the researcher is hand signs vary from motion to appearance. The Model defines three steps: Preprocessing, Feature Extraction, and Classification. Preprocessing includes skin filtering, histogram checking then Feature Extraction includes eigenvalue and eigenvector lastly eigenvalue weighted includes Euclidian Distance for recognizing the hand signs. This proposed model only works with a bare hand for good accuracy. There are only 24 alphabets considered as recognizing the sign. Fortunately, 96.25% is accuracy.[6]

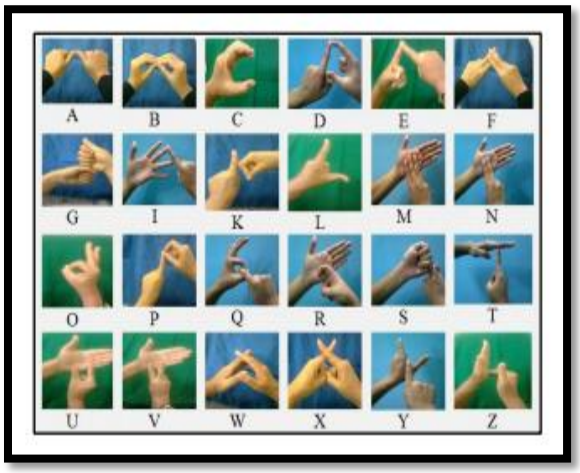

Figure 1: Sign Language *[6]*

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

W Ahmed focused on the property of covariant to scale, rotate and translate Fourier Descriptor was chosen for feature extraction. The classifier was used in this model was Nearest Neighborhood with Euclidian Distance. Datasets for this model was considered along with different skin color, varying distance from the camera and different illumination ranges. The success rate was achieved at 94.15%.

They used the methodology of feature extraction using Fourier Descriptors for using the hand sign in real-time, the problem occurs like a little bit variation in angle(rotation), scaling (hand size) and translation for feature extraction task.
For calculating feature, they found the centroid of the hand blob.

Centroid distance is calculated by the distance of the boundary point from the centroid(xc, etc.). $r(i) = \sqrt{([x(i) - x_c]^2) + ([y(i) - y_c]^2)}$

In Fourier Descriptor formulae the given value was put.

$$[Z_n] = \text{DFT}[\text{r(i)}] = \frac{1}{\text{N}} \sum_{i=0}^{N-1} r(i) e^{-\frac{2\pi i n}{N}}$$

$$\text{here, n} = 1 \ldots 256, \text{N} = 256$$

The true success rate with nonparametric Nearest Neighbors was achieved at 94.15%.[10]

Karishma Dixit et al. proposed vision-based approach which efficiently recognizes the hand signs and translates its meaning in the text. There is not any standard sign language all over the world. It differs from region to region because of their spoken languages. The feature vector is a prominent part of any sign language recognition model. Feature vector Hu invariant moment and the structural shape descriptor is used for the feature extraction process. For classification, a multiclass SVM is used for training the model and recognizing and classifying the sign. Support Vector machine examines its closest pair in 30 secs called support vector. After searching the closest pair, it draws the line for connecting them. So, they discovered a similar group in one class and dissimilar in another. 720 images are simulated and trained by this algorithm. However, there are many challenges like background color, illumination variation, hand movement, and position that occurs in building the system. The accuracy of this proposed algorithm of SVM was able to recognize the hand signs by 96 %.[11]

Vandana Soni et al discussed about PCA(Principal Component Analysis) along with the neural network. Principal Component Analysis is an algorithm for the high dimensional data analysis where the dimensionality reduction is done for the reduction of the input variable. It simplifies the problems by finding the projections that will maximize the variance of the datasets. They discussed that they captured the image and convert the image from RGB to HSV format and use city block distance method, use in Eigenvalue of a particular image and convert into text format and text word converted invoice format. In this method, the input image was transferred to skin

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

filtering then to hand cropping then to a binary image than to the grey image which shows the eigenvalue then value is converted to text match in a database and translate a voice.
PCA is calculated as finding the eigenvector of the covariance matrix of the current signs. Eigenvectors are stored as corresponding to their eigenvalue in decreasing order called a feature vector. Finally, the result for the PCA model is achieved at 94% which means the model recognizes the sign 94% correctly.
Further, this model was combined with a neural network for predicting more robust and recognizing the sign at higher efficiency [4]

Kanika et al. achieved the accuracy rate at a maximum of 96%. In the earlier proposed model, the hand signs scaling, rotation and translation were a huge challenge. But in OpenCV, the SIFT image feature gives a set of features which is not affected by these problems. 26 alphabets and numbers. were taken as input to the system and stored in a one-dimensional array. The output for communicating the deaf and dumb people is provided as an interface in OpenCV.[7]

Enrique Orduna et al. proposed mainly three steps for any recognition system that is preprocessing, feature extraction and classification. They captured image is converted into HSV color for preprocessing. for extracting the feature Hu moments method is used. After that Support Vector Machine is used for classification of the hand signs for the ISL. The building model is tested on webcam. In SVM all the similar groups occur in the same class and dissimilar in a different class. The accuracy rate for this proposed model is 96%.[2]
M. Suresh et al. created a two-way system. The first translation system is –sign to speech – if a deaf-dumb person can communicate with a normal person. Image from camera acquired is binarized, noise removed, the boundary of finger detected, finding exact text match and finally text with audio is sent to the receiver. The second translation system is –speech to sign –a normal person speaks, it is converted to text and then appropriate text meant sign will be displayed.[1]
Kumud Tripathi et al. performed experiments on 10 different sentences. The experiments are performed on continuous gestures of both static and dynamic. They recorded sentences 10 times out of which 6 were used for training and rest for testing. The results they obtained from correlation and Euclidean distance have higher Accuracy. The orientation histogram with 36 bins gives higher accuracy than 18 bins because it measures the angle of change of hand much appropriately.[5]

3. Methodology

This project aims to create an Indian language recognition System - i.e. given an input of somebody sign language an alphabet or numbers, the model ought to be able to predict that alphabet or range is being signed. Such a model has many use-cases and may be deployed either through an app or web site for folks to use. it's high connexon within the lives of not solely the deaf-dumb however additionally anyone who interacts with them.

There are 1.1 million deaf-mute folks in the Indian nation. 98% of those folks are illiterate, creating language their solely methodology of communication. This trend isn't expected to vary anytime shortly, as solely a pair of deaf youngsters attend college. The number of individuals within the country with a handicap was calculable at thirteen lakhs in the 2011 Indian census.

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

However, this might be a case of great under-reporting: The National Association of the Deaf estimates that there are 18 million people are deaf. the important numbers can be even larger as most countries have regarding 3% of their population as experiencing deafness, that interprets to regarding four crore folks. With such an oversized population of the state looking on language, it important to support them.

Though the signs employed in the country vary from region to region, many tries are being created to order a convention that's said because of the Indian language (ISL). However, this development continues to be in its aborning stage and isn't reaching most deaf colleges. Even otherwise, a calculable solely a pair of deaf youngsters attend formal education. These twin issues of a developing commonplace and therefore the restricted reach of formal education.

Proposed Methodology

1) A Graphic User Interface which allows the user to capture the image. This phase is called image acquisition.
2) In GUI a designated section for showing our hand sign.
3) After the image is captured. It goes through preprocessing which are:
    a. Set minimum and maximum skin color range for YCrCb format.
    b. Convert RGB to YCrCb.
    c. Noise removal and smoothing.
    d. Convert RGB to Grayscale.
4) Feature Extraction
5) Classification using CNN algorithm, and build a model.
6) Predict from the model.
7) Give text output. Convert text to speech, Now the processes in with images

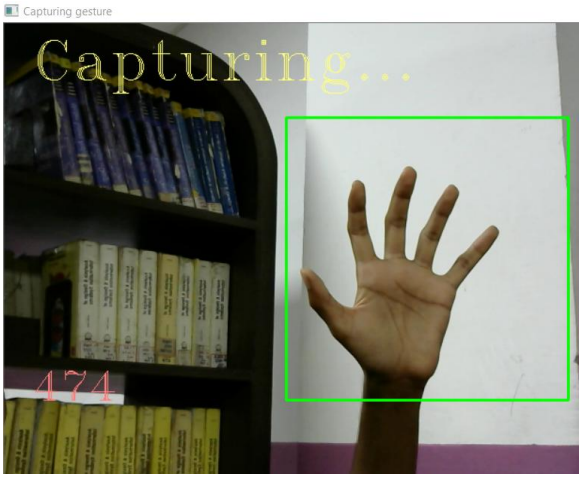

Figure 2: Input Image

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

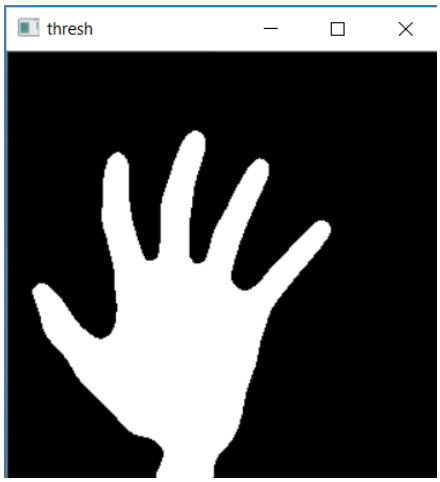

Figure 3: Output Image after thresh-holding

RGB to YCrCb

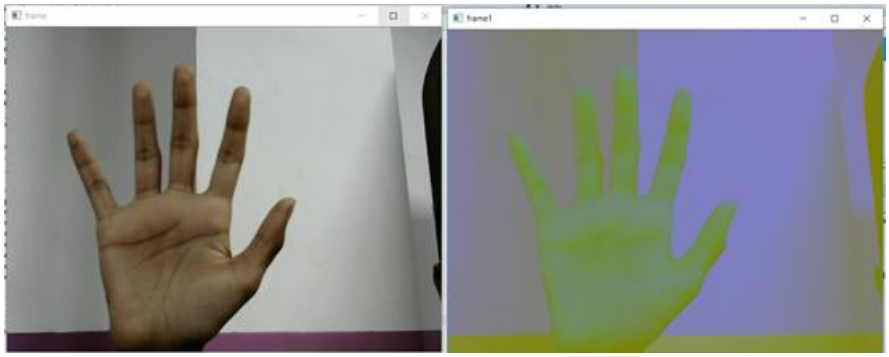

Figure 4: Converting RGB to YCrCb

YCrCb refers to

i. **Y→** Luminance component obtained from RGB after gamma correction
ii. **Cr = R – Y →** It takes the distance from the red component to the luminance component
iii. **Cb = B – Y→** It takes the distance from blue component to luminance component

Figure 5: Skin Range YCrCb

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

There is a set range in YCrCb that is for skin color that is for minimum [0,133,77] and for maximum [235,173,127]. In this range the image is transformed.

Figure 6: Gaussian Blur

Gaussian blur takes the image and transforms the original image into a softly blurred version of the given image.

Figure 7: Median Filter

The median blur technique works in a similar way comparing to other averaging techniques.
The median blur processes each element of the image and replaces it by the median of all the pixels of the surrounding pixels.

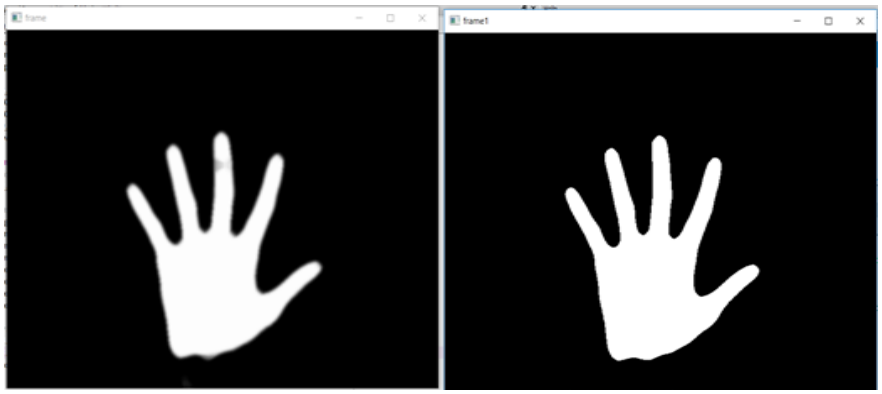

Figure 8: Thresholding the images

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

The threshold function checks whether the pixel value is greater than or less than the threshold value if it is greater than it will be assigned a value if not another value. cv2.threshold() is used to invoke the thresholding function.
There are several requirements to apply thresholding which are:

1. The image is in grayscale format.
2. The threshold value to classify pixel.
3. Maximal value if the pixel value exceeds the threshold value

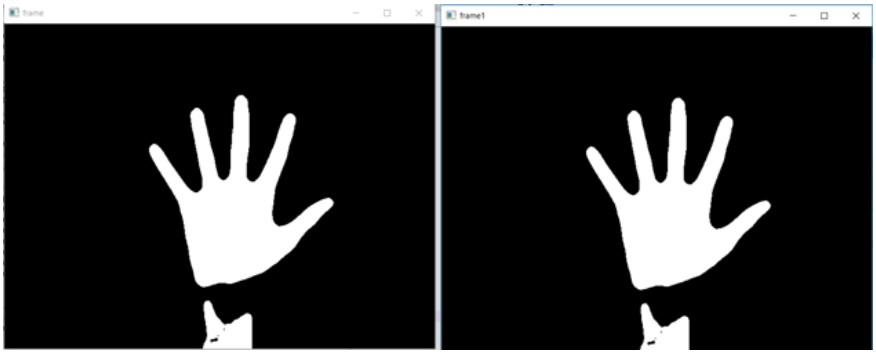

Figure 9: Merging

As the name suggests it merges the images that are passed to its function.
Cropping the section, we get our desired output.

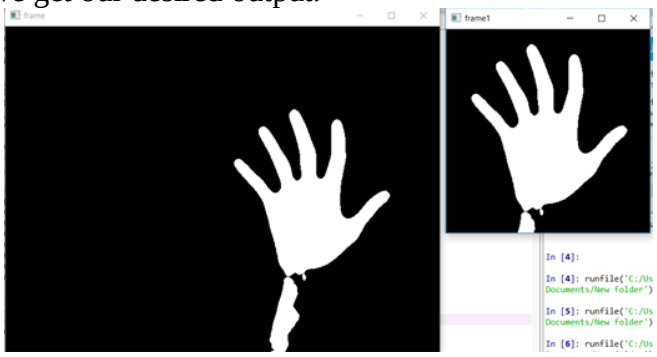

Figure 10: Cropping images

These are the steps we have followed in preprocessing the image before training or testing.
Then we have used CNN classification to train our system. We have chosen CNN because of its compatibility with training images and have better accuracy.

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

We have given the Graphs of Cross-Entropy Loss and Classification Accuracy.

```
Epoch 00020: val_acc did not improve from 0.99980
CNN Error: 0.02%
```

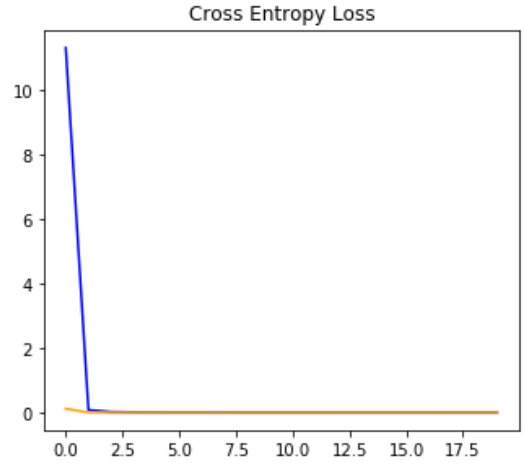

Figure 11: Cross Entropy Loss

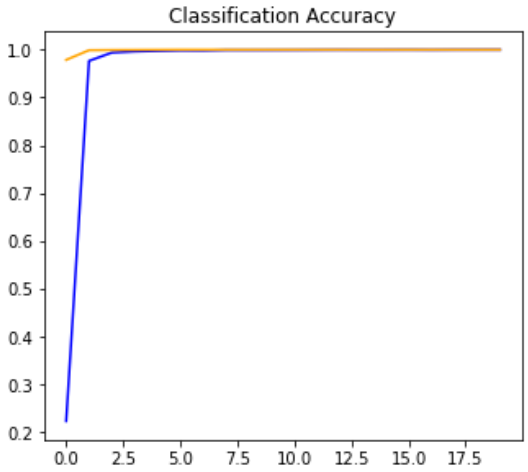

Figure 12: Accuracy Graph

Conclusion

We have made a system to help the common people to understand what the deaf and dumb people are saying. To do that we have a need for a webcam that by capturing the image we convert the image in YCrCb image and do preprocess to the image. After that the preprocessing we check from the database to match the image and predict the image. If the predicted is text same for 20 frames of images than the prediction is confirmed and prints on the screen. After the letters, the person as to say has been finished the word or set of letters is converted into speech by the system.

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

References

[1] Anand, M. Suresh, A. Kumaresan, and N. Mohan Kumar. "An Integrated Two Way ISL (Indian Sign Language) Translation System--A New Approach." *International Journal of Advanced Research in Computer Science* 4.1 (2013).
[2] Nair, Anuja V., and V. Bindu. "A review on Indian sign language recognition." *International Journal of Computer Applications* 73.22 (2013).
[3] Tripathi, Kumud, and Neha Baranwal GC Nandi. "Continuous Indian sign language gesture recognition and sentence formation." *Procedia Computer Science* 54 (2015): 523-531.
[4] Singha, Joyeeta, and Karen Das. "Recognition of Indian sign language in live video." *arXiv preprint arXiv:1306.1301* (2013).
[5] Rojasara, Dipali, and Nehal G. Chitaliya. "Indian Sign Language Recognition-A Survey." *International Journal of Engineering Research and Technology (IJERT)* 2.10 (2013).
[6] Deora, Divya, and Nikesh Bajaj. "Indian sign language recognition." *2012 1st International Conference on Emerging Technology Trends in Electronics, Communication & Networking*. IEEE, 2012.
[7] Ahmed, Washef, Kunal Chanda, and Soma Mitra. "Vision based hand gesture recognition using dynamic time warping for Indian sign language." *2016 International Conference on Information Science (ICIS)*. IEEE, 2016.
[8] Dixit, Karishma, and Anand Singh Jalal. "A Vision-Based Approach for Indian Sign Language Recognition." *International Journal of Computer Vision and Image Processing (IJCVIP)* 2.4 (2012): 25-36.