International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 05 | May - 2022 Impact Factor: 7.185 ISSN: 2582-3930

ISL Interpreter

Yash Varia[1], Aditya Chaurasiya[1], Jiten Patel[1], Ms. Theres Bemila[2]

[1]UG Student, Department of Information Technology, Shah and Anchor Kutchhi Engineering College, Mumbai
[2]Assistant Professor, Department of Information Technology, Shah and Anchor Kutchhi Engineering College, Mumbai

---

Abstract - It is observed that we Human Beings communicate with each other through different mediums such as verbally, digitally through text or via gestures, and so on. Understanding what the other person is trying to convey is important in today's world where miscommunication can cause severe consequences. Hearing and Speech impaired people have trouble filling this communication gap which creates a huge negative impact on their day-to-day life. They use Sign language to overcome this communication barrier which is not enough as the majority of people do not know Sign language. For this reason, the implementation of a Digital real-time system will be very helpful for the Speech impaired community. As a result, we have proposed a real-time, easy-to-use, free Indian Sign Language Interpreter System that is compatible with varying lighting and environmental conditions and has low computational complexity using Image processing and Artificial Intelligence.

Key Words: Artificial Intelligence, Free, Image Processing, Indian Sign Language, Varying lighting conditions.

1.INTRODUCTION

Gestures is a movement of part of the body, especially a hand or the head, to express an idea or meaning. Gestures are of different types and in them hand gestures can be dissected into two groups which are Global motion where the whole hand produces a movement, and the latter is Local motion where only the fingers are taken in consideration.

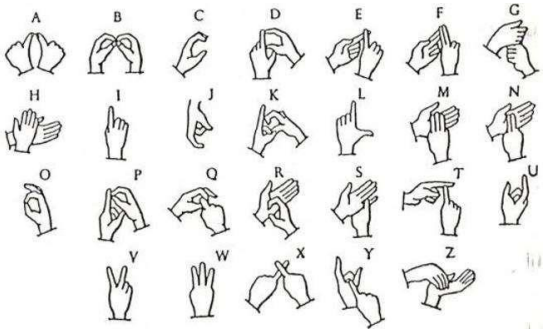


Fig. 1 – Representation of ISL Alphabets

Here we are focusing on Indian Sign Language which is used to communicate by the Hearing-Speech impaired people. In a recent study regarding Sign Language Awareness conducted in our country, it was found that there are only some hundred Certified Sign Language Interpreters whereas there are approximately 2.8 million people who suffer from Speech or Hearing disability. Hence, to fill this disproportionate gap, we have produced an Al/Ml based ISL Interpreter to convert Sign Language signs to text.

After conducting lot of research online and studying various published papers we have found that most of the research focuses on One-Hand signs or detection using multicolored hand gloves. In this paper we have proposed a system that can recognize Sign Language that are static two-hand or one-hand gestures in nature. This system will also be able to form words and convert them into different languages.

2. REVIEW OF LITERATURE

Miss Juhi Ekbote and Mrs. Mahesweta Joshi proposed utilizing ANN and SVM to recognize ISL gestures.[2]ISL alphabets and numerals are converted to text using the provided methods. The camera's input is transformed to YCbCr shading space, and then feature extraction is performed using HOG, shape descriptors, SIFT, PCA, and Fourier Descriptors. The recognition is then done using a support vector machine and a neural network.

Kandarpa Sarma, Ananya Choudhury, and Anjan Talukdar Under a complex background, I proposed a Conditional Random Field based Indian Sign Language Recognition System. [3] The methods proposed transform ISL gestures into words. Face detection is performed using Haar cascade, and the face is then removed from the frame, followed by morphological operations to filter noise, and motion detection is performed by frame differencing of the input frames. Then, between the outputs, a logical ANI operation is conducted to obtain the final ham segmented output. After that, the output is displayed, which includes contour processing, hand tracking, feature extraction, and recognition using CRF.

To transform ISL fingerspelling into speech, Paromita Das, Riya De, Sudipta Paul, Mainak Chowdhury, and Biswarup Neogi suggested ANALYTICAL STUDY AND OVERVIEW ON GLOVE BASED INDIAN SIGN LANGUAGE INTERPRETATION TECHNIQUE [4]. Tennis gloves with five flex sensors and a gyroscope sensor are used to collect data, which is then turned into a digital signal using an ADC. The input is then encoded in parallel to serial and transferred to the mobile device via a WIFI module. The motion is then classified by the app on the mobile device and forwarded to TTS Engine.

To convert fingerspellings and numerical into text, Adithya V, Vinod P.R., and Usha Gopalakrishnan proposed an Artificial Neural Network Based Method for Indian Sign Language Recognition [5]. The following steps are used to capture images of the hand for hand segmentation using a webcam. The image is color-space converted to YCbCr. Skin-colored pixels in input image images are identified using thresholding, after which various morphological operations are performed to reduce noise. Following hand segmentation, feature extraction is performed, for which a new shape feature based on the image's distance transform is used. Following that, the extracted features are used to train the feed-forward neural network to recognize the sign.

3.PROPOSED METHOD

A. DATA ACQUISITIONS

The Database which we have used for implementing this system was built by us as the readily available online dataset where not up to the mark, incomplete or inaccurate. Dataset was generated by capturing images through the built-in laptop webcam. Around 1200 images for each alphabet were captured which were then mirrored to obtain flip images that helped us to cover both the hands, so in total 2400 images for each alphabetic letter were produced for our dataset. These images were then reduced to 50x50 px to reduce the size of the dataset. The dataset covers all A-Z letters and we have split this dataset into two parts i.e., Training data and Validation data.

B. CREATING HISTOGRAM FOR THE HAND

A histogram is a graphical representation of the tonal values of your image. In other words, it shows the amount of tones of particular brightness found in your photograph ranging from black (0% brightness) to white (100% brightness). [6]

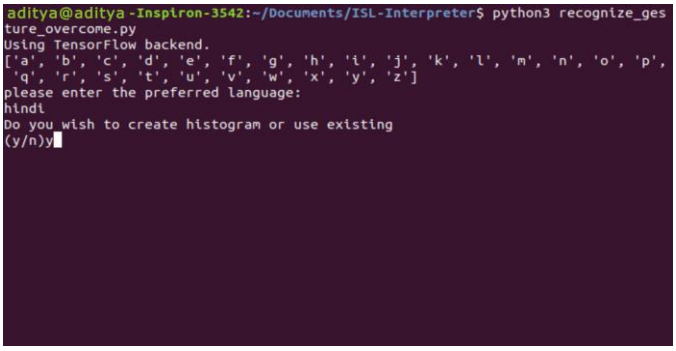


Fig. 2 User Screen when using Script

In order to determine the skin pixels in the image we have utilized Histogram. The idea behind using Histogram is that it makes the application adaptable to the changes in the lighting condition. E.g., If the user is using the application in a well-lit room, then the image will easily recognizable as compared to the usage in a dark room or at night.

By generating a Histogram each time, the system adapts to the lighting changes and can also easily take a dynamic approach to skin color segregation.

Fig. 3 Creating Histogram

In Fig. - 3 User Histogram creation starts when user places his hand inside the ROI.

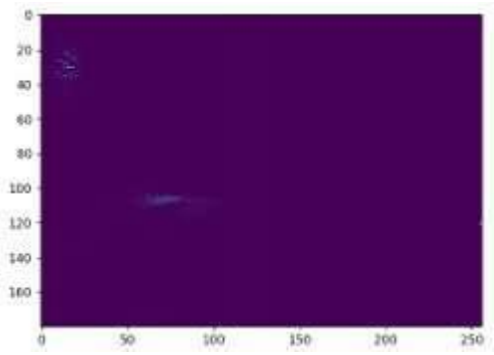


Fig. 4 Visual Representation of Histogram

In Fig. - 4 Histogram in Visualized by taking data from the past step in the form of HSV color model.

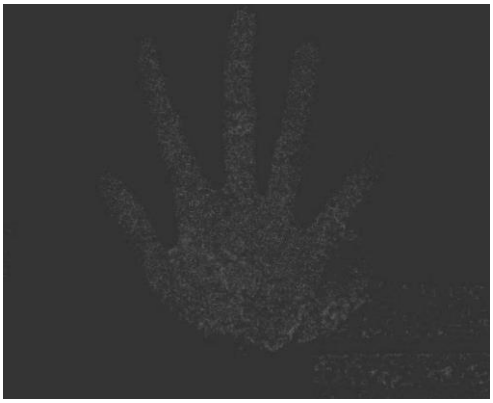
C. SKIN COLOR SEGMENTATION

The Histogram obtained is now used to Back project skin pixels.

1) Back Projection

Back Projection is a way of recording how well the pixels of a given image fit the distribution of pixels in a Histogram model. To state it simply, in Back Projection we calculate the histogram model of a feature and then use it to further find the feature in that image.

Back Projection is a way of recording how well the pixels of a given image fit the distribution of pixels in a Histogram model. To state it simply, in Back Projection we calculate the histogram model of a feature and then use it to further find the feature in that image.



Brightness of **Fig. - 5** has been increased for better vision

2) Filtering and Smoothing

The image obtained in the Back Projection is not very accurate contains lot of noise. In order to smoothen the image and reduce noise from it, various filtering and smoothing operations are performed.

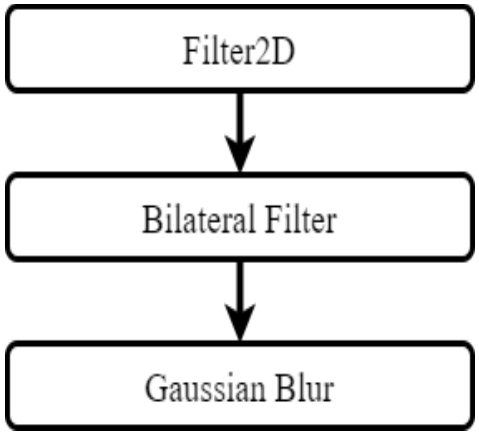


Fig. 6 Flow of the Entire process of Filtering and Smoothing

The function Filter2D (Fig.– 6) is used to convolve a kernel with an image. Operation is like this: keep this kernel above a pixel, add all the 225 pixels below this kernel, take its average and replace the central pixel with the new average value. It continues this operation for all the pixels in the image. [8]

Fig. 7 Filter2D

The Bilateral Filter (Fig.-7) is used to effectively remove noise while keeping edges sharp. [9]

Fig. 8 Bilateral Filter

Gaussian Blur (Fig.- 8) is used to further reduce image noise and reduce details where a Kernel value of 5x5 is used.

Fig. 9 Gaussian Blur

3) Thresholding

Thresholding is a simple method of Image Segmentation. Through Thresholding Binary Images can be produced from Grayscale image. [12]

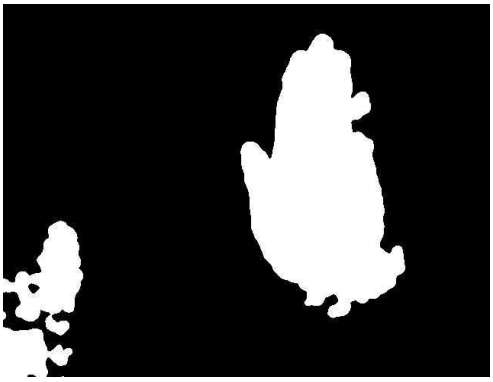


Fig. 10 Image obtained after Thresholding

D. CLASSIFICATION

The image obtained after filtering and smoothing process is used as the input in the classification step. A variety of classification techniques are available which are useful for recognizing the gestures. Classification consists of two phases training and testing phase. Image classification is the most important part of digital image analysis. [10]

In the implementation of our system, we have used Convolutional Neural Network for recognizing the ISL letters.

In deep learning, a convolutional neural network (CNN or ConvNet) is a class of deep neural networks, most commonly applied to analyzing visual imagery. CNNs are regularized versions of multilayer perceptrons. Multilayer perceptrons usually refer to fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer. [11] In the neural network that we have created we have 4 main layers. Input layer followed by two hidden layers and lastly the output layer. High accuracy is proficient by using CNN.

3.RESULTS

This segment provides the execution aftereffects of the gesture-based communication acknowledgment framework. Below pictures represent the testing outcome. The framework was actualized utilizing Python 3.6.7.

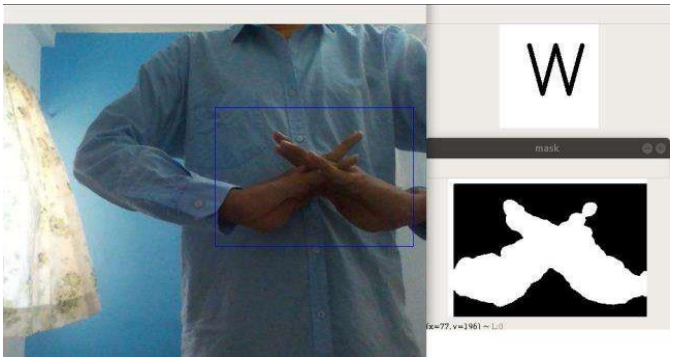


Fig. 11 Gesture and output for Letter W

Fig. 12 Gesture and output for Letter E

The word "Cat" can also be converted into different languages which helps the Impaired to communicate in multiple languages seamlessly.

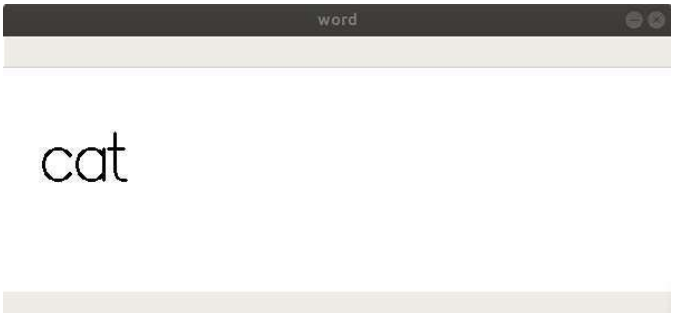


Fig. 13 Formation of Word "CAT"

Outcome from the testing of the Progream

Number of Test Images	Correctly Classified	Wrongly Classified	Accuracy (%)
26*5	115	15	88.46%

4. CONCLUSIONS

This research paper entitled as "**ISL Interpreter**" which is an AI based project gives us a brief idea about the problems faced by the Deaf-Speech impaired community in India and how this system can mitigate their problems by implementing it on a large scale without spending exorbitant amount.

The proposed system makes use of Image processing and Convolutional Neural Network at its core to function, and the trained model has a validation accuracy of 99%.

REFERENCES

[1] https://www.pri.org/stories/2017-01-04/deaf-community-milIons-hearing-india-only-just-beginning-sign
[2] Hardik Rewari, Vishal Dixit, N. Hema, "Automated Sign Language Interpreter", 2018 Eleventh International Conference on Contemporary Computing.
[3] Ananya Choudhury, Anjan Kumar Talukdar, Kandarpa Kumar Sarma, "A Conditional Random Field Based Indian Sign Language Recognition System under Complex Background", 2014 Fourth International Conference on Communication Systems and Network Technologies.
[4] Paromita Das, Riva dc, Sudipta Paul, Mainak Choudhury, Biswarup Ncogi, "Analytical study and overview on glove based Indian Sign Language interpretation technique", Michael Faraday IET International Summit 2015.
[5] V. Adithya, P.R. Vinod, Usha Gopalkrishnan, "Artificial neural network-based method for Indian Sign Language recognition", 2013 IEEE Conference on Information & Communication Technologies.
[6] https://photographylife.com/understanding-histograms-in-photography
[7] https://docs.opencv.org/2.4.13.7/doc/tutorials/imgproc/histograms/back_projection/back.projection.html
[8] https://docs.opencv.org/3.1.0/d4/d13/tutorial_py_filtering.html
[9] Madhuri Sharma, Ranjna Pal and Ashok Kumar Sahoo, "Indian Sign language Recognition using Neural Networks and KNN Classifiers", ARPN Journal of engineering and Applied Sciences, Vol. 9, No. 8, August 2014.
[10] https://en.wikipedia.org/wiki/Convolutional_neural_network