OPEN ACCESS

International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-4, November 2019

Interpretation and Translation of American Sign Language for Hearing Impaired Individuals using Image Processing

Shreyas Rajan, Rahul Nagarajan, Akash Kumar Sahoo, M. Gowtham Sethupati

***Abstract:** The growth of technology has influenced development in various fields. Technology has helped people achieve their dreams over the past years. One such field that technology involves is aiding the hearing and speech impaired people. The obstruction between common individuals and individuals with hearing and language incapacities can be resolved by using the current technology to develop an environment such that the aforementioned easily communicate among one and other. ASL Interpreter aims to facilitate communication among the hearing and speech impaired individuals. This project mainly focuses on the development of software that can convert American Sign Language to Communicative English Language and vice-versa. This is accomplished via Image-Processing. The latter is a system that does a few activities on a picture, to acquire an improved picture or to extricate some valuable data from it. Image processing in this project is done by using MATLAB, software by MathWorks. The latter is programmed in a way that it captures the live image of the hand gesture. The captured gestures are put under the spotlight by being distinctively colored in contrast with the black background. The contrasted hand gesture will be delivered in the database as a binary equivalent of the location of each pixel and the interpreter would now link the binary value to its equivalent translation delivered in the database. This database shall be integrated into the mainframe image processing interface. The Image Processing toolbox, which is an inbuilt toolkit provided by MATLAB is used in the development of the software and Histogramic equivalents of the images are brought to the database and the extracted image will be converted to a histogram using the 'imhist()' function and would be compared with the same. The concluding phase of the project i.e. translation of speech to sign language is designed by matching the letter equivalent to the hand gesture in the database and displaying the result as images. The software will use a webcam to capture the hand gesture made by the user. This venture plans to facilitate the way toward learning gesture-based communication and supports hearing-impaired people to converse without trouble.*

***Keywords:** Sign Language, Image Processing, database, hand gesture.*

Manuscript published on November 30, 2019.

* Correspondence Author

G.Shreyas Rajan*, pre-final year student Dept of Computer Science & Engineering at SRM Institute of Science & Technology, Ramapuram.

Rahul Nagarajan**,** pre-final year student Dept of Computer Science & Engineering at SRM Institute of Science & Technology, Ramapuram.

Akash Kumar **Sahoo,** pre-final year student Dept of Computer Science & Engineering at SRM Institute of Science & Technology, Ramapuram.

M.Gowtham Sethupathi, Assistant Professor Dept of Computer Science & Engineering at SRM Institute of Science & Technology, Ramapura

I. INTRODUCTION

For an unfortunate individual, the inability to speak and/or hear poses a great challenge as sign language is their sole medium of communication.

This is because sign language is predominantly fathomable among people sharing the same disability; which essentially hinders or rather neutralizes layman interaction. This is the place innovation tolls in,

with plenty of answers for decode communication through signing. We aim to devise a method which can effectively destroy the communicative obstruction between deaf-mutes and a commoner. The project serves to decipher American Sign Language (ASL) visuals and convert it to textual (English) format such that words or entire sentences can be obtained. We plan to improvise a doable methodology for communication through sign language that would be reasonably easy to execute. The project shall provide a convenient interface for the deaf-mutes, integrated with the webcam. Using the webcam, ASL shall be recognized, processed and converted to alphabetic or numerical values such that entire sentences are eventually obtained and thus the dream of communication among an enabled and disabled shall come true to its fullest extent.

II. LITERATURE SURVEY

Various papers deal with the conversion of American Sign Language to English. The conversion is either done to English text to speech [7] or simply converted to English text. There are two basic methods of approach to converting ASL to speech [4] which is the Sensor-based sign recognition system [1] and recognition of sign language using Image Processing.

Most of the Sensor-based method [1] [10] uses the Microsoft Kinect, which is a mobility detector equipped with a camera, used to sense the movement. [1] Microsoft Kinect is a movement sensor that was created by Microsoft for the Xbox 360 (gaming console) which enables the gamer to manipulate the game without having contact with the controller. Microsoft Kinect can also be used for the Windows Operating system. Some of the papers which were using the sensor-based recognition used a hand glove connected to sensors [5]. Sensors connected to the glove read the gesture and sends accurate information to the front-end system. The framework [5] proposes Artificial Neural Networks and the use of sensors for the recognition of American Sign Language. This framework likewise utilizes a tangible Cyber-Glove and a Flock of Birds 3-D movement tracker to isolate the hand motions.

The Sensor-based recognition system works accurately for about 96% and the systems are capable of translation of sign language to English [1][5][10]. However, there are downsides to this method of approach. [5] Albeit the greater part of the ASL-letters relies upon finger-bending, some of them likewise depend upon hand orientation and two of them are dynamic.

There are similitudes among 'g' and 'q', 'h' and 'u', 'k' and 'p'. These couples have a similar handshape; however, their hand orientations distinct from the others. Extra (costly) equipment – Microsoft's Kinect Sensor is a fundamental part of the framework. Hence to resolve these downsides of the sensor-based hand gesture recognition system, the second method is used in many systems. This is the 'Recognition of hand gestures using Image Processing Technology'. Most of the system for hand gesture recognition today uses image processing technology.

There is a plethora of mythology/algorithms utilized in the image processing technique. Image processing is combined along with Artificial Intelligence [2] to extract the image, [9]The conventional Convex Hull algorithm is combined with Image Processing for the recognition of strong points on hand [4], Image processing with Euclidian Distance to recognize the letters [6], and Image Processing is combined with the Hidden Markov Models (HMM) [8].

[2] The proposed framework intends to perceive some exceptionally fundamental components of gesture-based communication and interpretation of them into meaningful content. To initiate, the video will be captured frame-by-frame, the captured video will be developed, and the suitable picture will be removed, this acquired picture will be additionally handled utilizing BLOB analysis and will be sent to the database; here the captured picture will be contrasted and the one present in the database and the corresponding picture will be utilized to detect sign language.

[7] The framework proposes a method for moment Hand Gesture Recognition and highlight extraction utilizing a web camera - The picture is pre-processed and the limit is utilize99d to expel commotion from the picture. Furthermore, region covering is applied to fill an opening in the fundamental closer view. At the stage of pre-processing, the picture is passed on to include the extraction stage. The test picture is characterized in the closest neighbor's class in preparing the set. The characterization results are shown to the user and through the windows text-to-speech API gesture.

[8] An extensible framework that uses a colored-camera to track hands continuously and deciphers American Sign Language (ASL) utilizing Hidden Markov Models (HMM's). The following phase of the framework procedure delivers a coarse depiction of hand shape, orientation, and trajectory. The hands are followed by their color: in the first trail through decidedly hued gloves and the second, through their regular skin tone. In the two cases, the result data is contributed to HMM for acknowledgment of the marked words. This framework can convey an exactness rate of 91.9% whilst adequate lighting conditions.

[9] The framework proposes a perfect procedure; whose genuine objective is to accomplish the transliteration of 24 static gesture-based communication letters in order and quantities of Sign Language into humanoid or machine-justifiable English manuscript-processing operations of the marked info signal are done in the principal stage. In the following stage, the different locale properties of pre-prepared signal pictures are determined. In the last stage, in light of the properties determined in the previous stage, the transliteration of the marked signal into content has been completed.

[3] This paper aims to implement hand gesture recognition using current technologies which include Image Processing and Neural Networks. The paper proposes to use Neural Networks to train the system for the database and use Image Processing techniques to identify the hand gestures and to display the corresponding output. The Algorithms used to implement the system are COHST- Combined Orientation Histogram and Statistical and Wavelet features for the static recognition of hand gestures.

[4] This system proposes the use of Image processing for the recognition of the hand gestures. The paper intends to use a simpler approach of image-based recognition rather than sensor-based recognition. The system uses tools such as Microsoft Visual Studio as an IDE and the OpenCV library. This system also uses a webcam to capture real-time images for sign detection. The tracking of a hand gesture from a captured image is done using the Convexity Hull algorithm.

[6] This framework helps standard speakers in perceiving signals utilized in American Sign Language. The framework illustrated in this paper is finished with the assistance of MATLAB. In this methodology, the signs are caught utilizing a webcam. The pictures caught are then handled further and the highlights are extricated from the caught pictures utilizing PCA. Correlation of the highlights is finished utilizing Euclidean Distance with the preparation sets. Least Euclidean separation perceives the character.The Image Processing method produces accurate results almost in any condition. Hence the Technology used in this project will be the Image Processing techniques.

III. ARCHITECTURE DIAGRAM

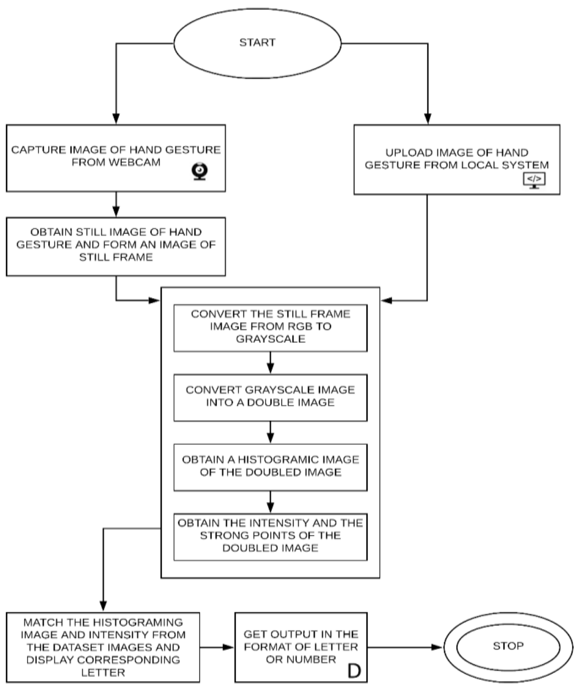

Fig 1 Architecture Diagram of the system

The architecture diagram denotes the workflow of the proposed system which will be covered in the forthcoming section. The basic idea is to get the input image via webcam or local system and feed the system. The system then performs the image processing techniques on the input image and matches it with the image value in the dataset and gives the output in text format.

IV. PROPOSED SYSTEM

This system aims to facilitate acquisition and decyphering of sign-language for all individuals. The system will be capable of capturing the hand gesture of an individual through the front end application and the webcam and display the output in the format of English Alphabets or Numerals. This system can be divided into different modules based on its work, the first being Capturing and retrieving the live image. The next module is the conversion of the raw image to a processed one by applying the image processing technique. The following module is the equalization of the handled image with the image in the dataset. The algorithm and the work process of every module are explained in the subsequent section

V. CAPTURING AND RETRIEVAL OF LIVE IMAGE

This is the first and foremost module of the system. In the module, the user show's hand gesture of the specific alphabet or number to be retrieved to the live webcam. However, this process can also be done using a static method. The required hand gesture can be captured before inputting the image. This image can then be uploaded to the system from the stored path of the image. This is referred to as the static method and the retrieval of an image via the webcam is known as the dynamic method. This system then takes the hand gesture's image as the input and sends it to the processing part of the system where this image is further processed to match it with the value in the Dataset.

VI. PROCESSING THE IMAGE USING EDGE DETECTION ALGORITHMS

The processing of the image in this framework is implemented using MATLAB. In this module, the input image is manipulated in a way such that its intensity and pixel values match the values of the corresponding image in the dataset. This is achieved through the application of Edge detection on the image.

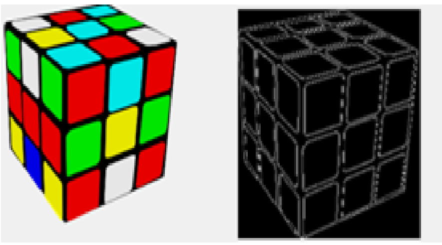

Fig 2 Edge detection using 'Sobel'

A. Algorithm for edge detection using Sobel

1. Read in the input image.
   I = read ('hadgesture.tif'); Display(I)

Fig 3 Hand Gesture

2. The image to be fragmented varies enormously in contrast to the foundation picture. Changes conversely can be recognized by administrators that ascertains the slope of a picture. To generate a binary mask, quantify the gradient image and apply a limit.
   Threshold = edge (I,' Sobel')
   Fudge Factor = 0.5
   BW = edge (I,'Sobel', Threshold * Fudge Factor)
3. Dilate the Image
   Add pixels to the boundaries of objects to the binary mask by the vertical structuring followed by horizontal structuring entity. The 'imdilate' method is used for this purpose
   BW = dilate (BWs, [se90 se0])
   Display BW

4. Fill Interior Gaps: The Hand gesture still contains dilated gaps.
   BW1 = fill (BW,'holes')
   Display BW1

5. Remove linked Objects on Border
   BW2 = clear border (BW1,4)
   Display BW2
6. Smooth the Image
   SeD = strel ('diamond',1)
   BW3 = erode (BW2,SeD)
   BW4 = erode (BW3,SeD)
   Display BW4

This gives us the edge detected image which would look as follows:

Fig 4 Edge Detected Hand Gesture

VII. COMPARISON OF THE PROCESSED IMAGE WITH THE IMAGE IN THE DATASET

This module is implemented using the Histogram matching algorithm

- Choose the range of the histograms, and for every range, account for the number of pixels with a color in the particular range

- For every individual pixel, look at the colour intensity's value and add a tally to the bucket. We split up each bucket total by the pixels in the whole image to obtain a normalized Histogram for the particular color channel.
- We perform Edge Detection technique on the image for the texture histogram. All the points have normal vectors pointing in the direction at right angle to the edge. We quantize the normal vector's angle from 0 – PI (0 to 180 degree)
- After adding up the number of edge points in all the directions, an un-normalized histogram is obtained which is normalized by dividing each bucket by total no. of edges in the image
- To quantify the texture scale histogram, for each edge, we gauged the separation to the following nearest edge with a similar orientation. After this, we put those qualities on a histogram and standardize it by separating by the no. of edge points.
- To distinguish two pictures, we take the total estimation of the distinction between every histogram and add the values. The following is the formula to compare two histogram images

$$d(H_1, H_2) = \frac{\sum_I (H_1(I) - \bar{H_1})(H_2(I) - \bar{H_2})}{\sqrt{\sum_I (H_1(I) - \bar{H_1})^2 \sum_I (H_2(I) - \bar{H_2})^2}}$$

Where

$$\bar{H_k} = \frac{1}{N} \sum_J H_k(J)$$

` and N is the total number of histogram bins, I and J are the initial images to be compared.

Here H_1 is the first Histogram and H_2 is the second Histogram. The above mentioned is the formula to compare two histogram images

- The smallest value matches with the image in the Database.

$$Match\ \% = \frac{Total\ number\ of\ Pixels\ matched}{Total\ number\ of\ pixels}\ X\ 100$$

VIII. RESULTS

The outcome of the system encompasses AMERICAN SIGN LANGUAGE-letters and numbers. The input image is expected to be processed and be converted into its corresponding value such that the hearing-impaired individual can benefit from it. The following contains a dataset of certain letters and numbers.

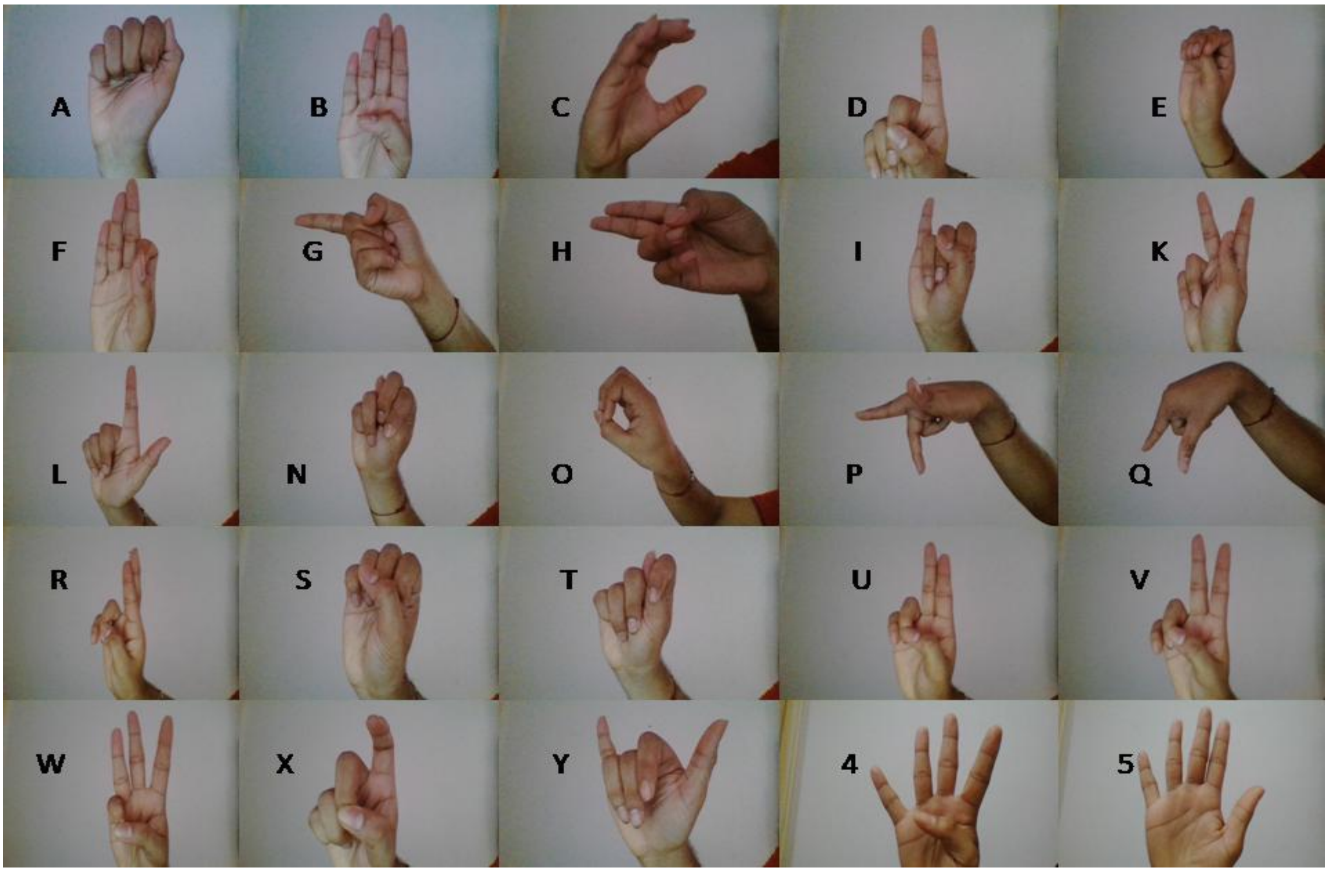

Fig 5 Input Hand Gesture Images

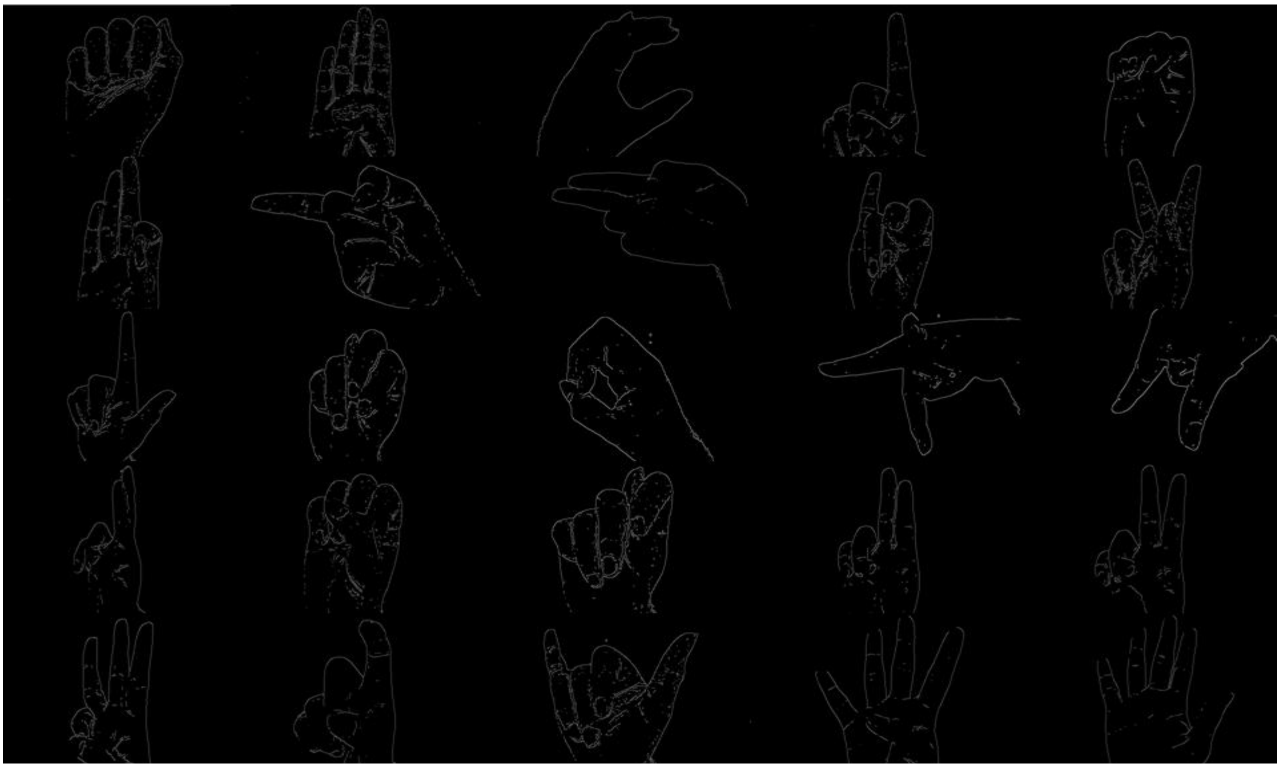

Fig 6 Edge detected Output image

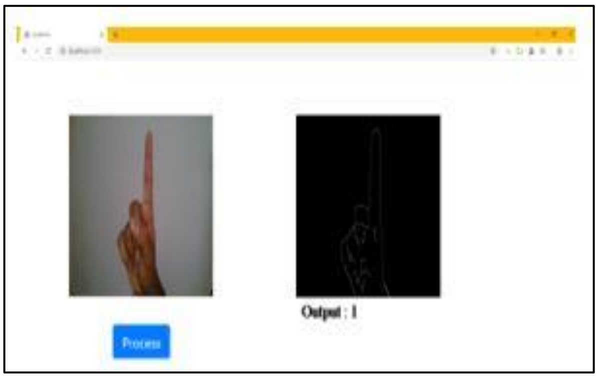

Fig 7 Output for the number '1'

The values of the processed image match with the values of the image in the dataset with an accuracy of 96.54%

.

Table- I Accuracy Table for some Alphabets

ALPHABET	NO. OF IMAGES IN DATASET	NO OF IMAGES MATCHED	ACCURACY
A	50	48	96%
C	50	48	96.2%
E	50	47	94.2%
G	50	45	90%
L	50	46	92.3%
S	50	45	91%
Y	50	49	99%

Upon observing the table explicitly, we can notice that there is a total of 50 images in the dataset from which 48 images have been matched for the alphabet 'A' with an exactness yield of 96%. Likewise, for the alphabet 'C', 48/50 images have been matched, with an exactness rate of 96.2%. 'E' on a downside has only 47/50 matches with a rate of 94.2%. 'G' has a further decrement of 45/50 matches, with the lowest rate of 90%. 'L' shows a minute improvement with a 46/50 match and a rate of 92.3%. 'S' turns out to be on a lower spectrum with only 45 matches and a rate of 91%. Y is on the apex with 49/50 matches and a mind-boggling rate of 99%.

To visually portray the data, we are going to resort to the Line chart and Line-Clustered column chart as given below:

Retrieval Number: D6966118419/2019©BEIESP
DOI:10.35940/ijrte.D6966.118419
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

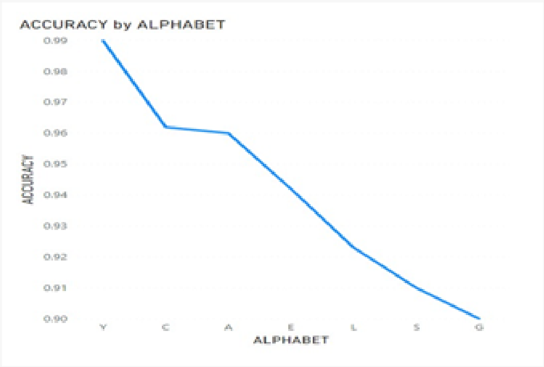

Fig 8 Accuracy vs Alphabet Line Chart

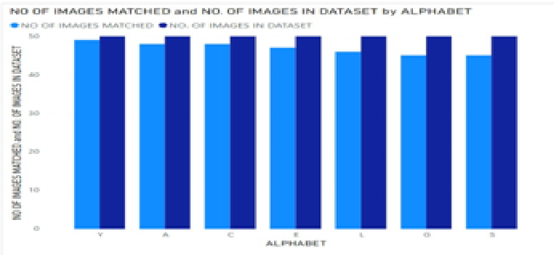

Fig 9 Images matched vs Images in dataset

IX. CONCLUSION

This paper deals with the day-to-day problems of communication amongst the Hearing-impaired individuals and thus a solution has been proposed regarding the same. There are two major drawbacks to this system. The letter 'J' and 'Z' has some motion as a part of its gesture and hence the system will not be able to input these letters. The system's dataset is fixed and limited to alphabets and numbers only. Furthermore, Artificial Neural Network algorithms can be applied in the future for the image comparison module. Future works on this system can be implemented by adding the hand gestures of simple meaningful words that can be easily recognized.

REFERENCES

1. Conversation Of Sign Language To Speech With Human Gestures Rajaganapathy. S1, Aravind. B2, Keerthana. B3, Sivagami. M4 1, 2, 3MS Software Engineering, 4Assistant Professor VIT University – Chennai INDIA.
2. Translation Of Sign Language Finger-Spelling To Text Using Image Processing Krishna Modi B. Tech. (Comp.), Amrita More B. Tech. (Comp.); Mukesh Patel School Of Technology And Management Engineering JVPD Scheme Bhaktivedanta Marg, Vile Parle (W), Mumbai-400 056
3. COHST And Wavelet Features Based Static ASL Numbers Recognition Asha Thalangea, Dr. S. K. Dixit, Assistant Professor, E&TC Dept., Walchand Institute Of Technology, Solapur – 413006, Maharashtra, India
4. Sign Language Recognition Using Image-Based Hand Gesture Recognition Techniques. Ashish S. Nikam P.G. Student, Dept. Of Elec. &Telecom, Aarti G. Ambekar Asst. Professor, Dept. Of Elect. &Telecom Dwarkadas J. Sanghvi College Of Engg. Mumbai, India
5. Recognition Of Finger Spelling Of American Sign Language With Artificial Neural Network Using Position/Orientation Sensors And Data Glove; Cemil Oz And Ming C. Leu, Department Of Mechanical And Aerospace Engineering University Of Missouri-Rolla, Rolla, MO 65409, USA
6. Online Hand Gesture Recognition &Classification For Deaf &Dumb Nitesh S. Dr. M. S. Nagmode2, Mr. R. D. Komati3 Student1, Professor2, Ass. Professor3 Department Of Electronics And Telecommunication MIT College Of Engineering, Pune
7. Sign Language Recognition Using Image Processing; Vijay Moore, Sanket Sangamnerkar, Vaibhav Thakare Dnyaneshwari Mane, Rahul Dolas Dept. Of Information Technology, AISSMS Institute Of Information Technology
8. Real-Time American Sign Language Recognition From Video Using Hidden Markov Models Thad Starner And Alex Pentland Room E15-383, The Media Laboratory Massachusetts Institute Of Technology 20 Ames Street, Cambridge MA
9. American Sign Language Recognition System: An Optimal Approach; Shivashankara S Department Of Computer Science &Engineering, Srinath S Department Of Computer Science &Engineering, Sri Jayachamarajendra College Of Engineering, Mysuru, India
10. Sign Language Recognition With The Kinect Sensor Based On Conditional Random Fields Hee-Deok Yang Department Of Computer Engineering, Chosun University, Seosuk-Dong, Dong-Ku, Gwangju 501-759

AUTHORS PROFILE

G.Shreyas Rajan is a pre-final year student in Dept of Computer Science & Engineering at SRM Institute of Science & Technology, Ramapuram. He has experience in Mobile Application development and Web designing and development

Rahul Nagarajan is a pre-final year student in Dept of Computer Science & Engineering at SRM Institute of Science & Technology, Ramapuram. He has experience in Business Intelligence and Database Management Systems

Akash Kumar Sahoo is a pre-final year student in Dept of Computer Science & Engineering at SRM Institute of Science & Technology, Ramapuram. He has experience in Cloud Computing technologies, Web hosting, and Web Designing.

M.Gowtham Sethupathi is an Assistant Professor in the Dept of Computer Science & Engineering at SRM Institute of Science & Technology, Ramapuram. He has presented a paper "Application Access Control On Android-Based On Context and Security Profiles" in "National Conference on Computing and Communication", published in the proceedings at Anna University Tirunelveli.

Retrieval Number: D6966118419/2019©BEIESP
DOI:10.35940/ijrte.D6966.118419
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)