National Conference on
"Role of Engineers in Nation Building" organized by VIVA Institute of Technology, Mumbai
(6th and 7th March 2020)
International Journal of Engineering Research & Science (IJOER)
ISBN:[978-93-5391-287-1] ISSN: [2395-6992] [Vol-6, Issue-3, March- 2020]

An Analysis of Sign Language Analyzer and Translator

Saurabh Patil[1], Jaydeep Rathod[2], Jay Patil[3], Pallavi Vartak[4]
[1]Department of Computer Engineering Mumbai University, Mumbai
patil.saurabh60@gmail.com
[2]Department of Computer Engineering Mumbai University, Mumbai
jaydeeprathod48@gmail.com
[3]Department of Computer Engineering Mumbai University, Mumbai
j716449@gmail.com
[4]Department of Computer Engineering Mumbai University, Mumbasi
pallavivartak@viva-technology.org

Abstract**—***Human Beings interact verbally with each other to convey their thoughts, ideas and experiences with the people around them. But, this is not in the case for the Dumb and Deaf people. They have to communicate using Sign Languages and sometimes it is not possible for normal people to understand it. Dumb and Deaf people face many problems because of these disabilities as they cannot apply for Jobs. Also, if they are working it becomes quite difficult for the co-workers to coordinate. Through Sign Languages, it is possible to convey messages without acoustic sounds. Normal people are not able to understand the Sign language which gives rise to a major problem of miscommunication. To overcome this problem, we have proposed a system which aims to make dumb and deaf people communicate with normal people with less difficulty, reducing the communication gap between them. In the proposed hand gestures are captured, processed and converted into speech or text. The proposed system would use a camera to capture the hand gestures of the people with disabilities and it would display message on the screen This system will help to ease communication between normal and hearing-impaired people.*

Keywords— Analyzer , dumb and deaf , sign language , speech , translator

I. INTRODUCTION

Human Beings interact verbally with each other to convey their thoughts, ideas and experiences with people around them. This is not in the case for Dumb and Deaf people. They use Sign languages to communicate with people around them. But, sometimes it is not possible to convey their thoughts and this leads to miscommunication. However, people around these Dumb and Deaf people are not available to help them because of this barrier of Signs between them. Some normal humans are aware of the signs for conveying some things while some are not and this stops normal humans from helping those Dumb and Deaf People. Thus, to convey those messages properly and communicate with these Dumb and Deaf people there is a serious requirement of a system that could make these communication very much easier and smooth. A system wherein you could held a phone in front of these Dumb and Deaf people and using the video recorder you will get the meanings/words of the signs they are making. So, these would help both for proper communication. There are some systems in the market that predicts the alphabets of the signs done. But, this system takes more time to build a word and a sentence. So, prediction of a word instead of an alphabet could make it easier and faster for communicating with these Dumb and Deaf people.

II. LITERATURE SURVEY

V. Troung, et. al. [1] have proposed a system that can automatically detect static hand signs of alphabets in American Sign Language(ASL). The proposed system adopted the two combined concepts AdaBoost and Haar-like classifiers. In this work to increase accuracy of the system, a huge database for training process was used, and it generated great results. A data set of 28000 samples of hand sign images, 1000 images for each hand sign of positive training images in different scales, illumination, and the data set of 11100 samples of negative images was used for implementing and training the translator.

International Journal of Engineering Research & Science (IJOER)
ISBN:[978-93-5391-287-1] ISSN: [2395-6992] [Vol-6, Issue-3, March- 2020]

Experiments show that the system can recognize all signs with a precision of 98.7%. Input of the system is a live video and output is text and speech.

A. Jayprakash, et. al. [2] have proposed a solution to the communication barrier of deaf and dumb community with the society. An android application was developed that translates the gestures using less memory and less CPU processing time. The application can be divided into two parts, adding a new gesture into the dataset and recognizing existing gestures from the dataset. The main method to identify gestures is using descriptors while the histogram matching is done in order to reduce the dataset for comparing using descriptors. The initial phase of comparison using histogram matching is done to identify the gestures. Gestures that are close to test sample. Further, only those samples that are subjected to Oriented Fast and Rotated BRIEF (Binary Robust Independent Element Features) based comparison reducing the CPU time.

A. Khaleghi, et. al. [3] have proposed a machine learning based system. The system deals with the classification of single and double handed Indian sign language recognition using machine learning algorithm with the help of MATLAB. The system was implemented on MATLAB and PC with Intel i5, 2.2GHz processor and 4GB RAM. Features were extracted from the set of images and used for training the conjugate gradient back propagation neural network (supervised learning). The images were also trained separately using K-NN technique. The single and double handed Indian Sign language is acquired and classified into English alphabets and numbers using K-NN and back propagation techniques, where PCA is used for dimensionality reduction. For, K-NN techniques with K=1, the system achieved 100% recognition rate whereas using back propagation technique the system achieved 94-96% recognition rate.

C. Zhao, et. al. [4] have proposed a system to create a vision-based application which offers sign language translation to text. The existing system aids communication between signers and non-signers. The existing system extracts temporal and spatial features from the video sequences. Inception, a CNN (Convolution Neural Network) for recognizing spatial features was used. The system then uses RNN (Recurrent Neural Network) to train on temporal features. The dataset used is a American Sign Language Dataset that consists of video sequences. In the existing system CNN and RNN are trained independently. In this system recordings are independent of each other, there is no continuation of gestures between two recordings. Problems faced by this system is with facial features and skin tones. While testing with different skin tones the model dropped accuracy if it hadn't been trained on a certain skin tone and was made to predict on it.

P. Singh, et. al. [5] have proposed two classification techniques PNN and KNN. In this system hand gestures are captured and processed with the help of MATLAB and then converted into speech and text. Following parameters are taken into consideration while capturing, extracting and pre-processing are, lighting changes, different sizes and shapes of user hand, background, skin colour, distance from the camera and the angle of position of the hand. The system uses 7Hu techniques for feature extraction and KNN classifier. To be precise, 82% accuracy is achieved by the existing system. The quality of English speech is good because of the inbuilt MATLAB function. Hindi speech is a database of records on recorded sound file. In this ".wav" audio file format is used for Hindi speech. In future work, ISL (Indian Sign language) could be converted into ASL (American Sign Language).

B. Gupta, et. al. [6] have proposed a system for Indian sign languages that recognizes static images of the signed alphabets. Indian Sign Language has single handed as well as double handed gestures. First, the signs are classified as single-handed or double-handed then for both categories two kinds of features, namely Histograms of Oriented Gradient (HOG) and Scale Invariant Feature Transform (SIFT) are extracted for a set of training images. Then, they are combined in a single matrix. Correlation is computed for these matrices and is fed to a K- Nearest Neighbour Classifier to obtain the resultant classification of the test. The dataset has 520 images for the training segment and 260 images for testing. The existing system could be further improved by including other forms of gestures such as mimics, expression and finger spelling as well.

International Journal of Engineering Research & Science (IJOER)
ISBN:[978-93-5391-287-1] ISSN: [2395-6992] [Vol-6, Issue-3, March- 2020]

K. Dutta, et. al. [7] have proposed a system that captures a series of images and it is processed using MATLAB and then converted to speech and text. The system was developed to translate the double handed Indian Sign Language to both text and speech. The existing system is based on eigenvalue algorithm. It employs Shi-Tomasi corner detector. Feature detection method used in this paper is Tomasi's good feature to track. In this paper, two methodologies have been implemented. One of the methodology used is based on statistical method and other is based on centroid algorithm. In statistical method, single gesture with 10 different images were processed with proper threshold fixed. In centroid algorithm number of active fingers were calculated after forming two centric windows. Main drawback in the system is that they are position and background dependant which restricts its usage. Output text which is obtained after image and word processing is further converted to speech using text to speech synthesis.

N. Thiracitta, et. al. [8] have proposed a methodology to solve the barrier between dumb and deaf and the ordinary people using Markov Models. The system uses Argentina Sign Language that provides data as video. Pre-processing is done using edge detection and skin detection using the help of Contrast Adaptive Histogram Equalization (CLAHE) for image enhancement. Then the, features are extracted by its movement. At last, data is trained and classified using Modified Hidden Markov Model. 83% accuracy is achieved by classifying 10 signs using Gaussian Hidden Markov Model and 72% accuracy is achieved using Multinomial Hidden Markov Model. Drawbacks of this system are that the signers are using gloves, but this research is using skin recognition, making of the features aren't accurate. Future work for this system is improvement in noise reduction, dataset collection, better feature extraction, or better model to be used.

M. Taskiran, et. al. [9] presented a real-time sign language recognition system for people who aren't aware of sign languages. Sign Language used for the system is American Sign Language. In this paper, Convolution neural network model was trained by using dataset collected in 2011 by Massey University. After network training is completed, the network model and network weights are recorded for the real-time system. In the real-time system the skin colour is determined for a certain frame for hand use, and the hand gesture is determined using the convex hull algorithm, the registered neural network and network weights are used in real time to define hand gestures. Accuracy of this system is 98.05%.

D. Naglot, et. al. [10] have proposed a system using new digital sensor called "Leap Motion Controller" in which the signs are captured. LMC is a 3D non-contact motion sensor which can track and detect hands, fingers, bones and finger like objects. The existing system used Multilayer Perceptron (MLP) neural network with Back Propagation (BP) algorithm to build a classification model which takes feature set as input. Different signs are recognized using Multi-Layer Perceptron neural network. Dataset of total 520 samples (20 samples of each alphabet) is used to execute MLP. Features provided to ANN to train the system are distance between fingers tip position to palm center and distance between consecutive fingers tip position. Recognition rate of the proposed system is 96.15%.

S. Patel, et. al. [11] proposed a system in which gesture recognition and acquisition is done using camera. Initially it captures images in the plain RGB format, after that image is processed in the Skin Detection section where the algorithm of HSV-Skin Detection plays the role of detection of skin with the help of some threshold values of which are denoted by YCrCb. The next step is the blob detection algorithm where the skin detected image is processed to the grayscale image where the pixel are in the format of plain black and white pixels. Thus it further guides the algorithm into the contour detection where it accepts the gray image and if the contour area is found to be greater than 200 then it filled with white pixels. Finally the output is generated through text and speech form. The output of this system is very efficient, consistent and of high approximation of gesture processing and speech analysis.

M. Aarthi, et. al. [12] proposed a system with gesture recognition module and a Text-to-Speech synthesizer . Compared to other gestures, hand gesture plays an important role, as it expresses the user's views in less time. Flex sensor-based gesture recognition module is developed to recognize English alphabets and few words. Text-to-Speech synthesizer based on HMM is built to convert the corresponding text. The flex sensor is interfaced with the digital ports of Atmega328 microcontroller.

International Journal of Engineering Research & Science (IJOER)
ISBN:[978-93-5391-287-1] ISSN: [2395-6992] [Vol-6, Issue-3, March- 2020]

The output from the microcontroller is the recognized text which is fed as input to the speech synthesizer. Arduino microcontroller processes the data for each particular gesture made. The system is trained for different voltage values for each letter. Gestures performed by multiple users have been tested for all the letters in ASL.

Pei Xu, [13]have proposed a real-time human-computer interaction system is designed based on hand gestures, who recognizes gestures only using one monocular camera and extends the system to the HRI(Human Robot Interaction) case. Hand detection, gesture recognition and human-computer interaction (HCI) are the three major components of the system based on recognition and comprehends the robust control of mouse and key-board events with a higher accuracy of gesture recognition. The developed system only supports static gestures. The future work will be to investigate robust classifiers for dynamic gestures and develop a gesture-based HCI or HRI system with the support of complex motion recognition.

JoyeetaSingha, et.al. [14] have proposed a vision-based approach is used to build a dynamic hand gesture recognition system. Various challenges such as complicated background, change in illumination and occlusion make the detection and tracking of hand difficult in any vision-based approaches. To overcome such challenges, a hand detection technique is developed by combining three-frame differencing and skin filtering. Thus, a system with optimal features was selected using analysis of variance combined with incremental feature selection. These selected features were then fed as an input to the ANN, SVM and kNN model. These individual classifiers were combined to produce classifier fusion model. Fivefold cross-validation has been used to evaluate the performance of the proposed model. One-way analysis of variance test, Friedman's test and Kruskal–Wallis test have also been conducted to validate the statistical significance of the results.

AashniHariaa, et. al.[15] have proposed a robust marker- less hand gesture recognition system which can efficiently track both static and dynamic hand gestures. Static gestures are those gestures which do not change for a certain duration of time and dynamic hand gestures are gestures that change within a certain time. Detection module applies different image processing techniques such as colour conversion, thresholding, finding defects etc. The image is classified using Haar cascade to detect gestures if no defects are found. When a moving closed palm gestures are recognized for 5 continuous frames, it is considered to be a dynamic swipe motion. It is used when Microsoft PowerPoint is running in the foreground, to swipe to the next slide within the presentation. The accuracy for proposed system is estimated by the number of times it correctly recognises a gesture when it is shown 10 times in succession

III. ANALYSIS TABLE

The Table 2.1 is a summary of research papers on Sign Language Analyzer and Translator. The Table 2.1 states the different techniques used for recognition of signs . The accuracy varies as per the system used.

Table 2.1:Analysis Table

Sr. No.	Paper Name	Advantages	Accuracy
1.	A Translator for American Sign language to Text and Speech [1]	AdaBoost and Haar-like classifiers	98.7%.
2	Sign language Translator for mobile platforms [2]	Oriented Fast and Rotated Brief Algorithm	70%
3	Machine Learning techniques for Indian Sign languages[3]	PCA and ANN Algorithm in MATLAB.	Doublehanded - 95.84% Single handed - 94.88%
4	American Sign language recognition using Deep Learning and Computer vision [4]	Deep Learning - CNN and RNN	-

International Journal of Engineering Research & Science (IJOER)
ISBN:[978-93-5391-287-1] ISSN: [2395-6992] [Vol-6, Issue-3, March- 2020]

5.	Moment based Sign language recognition for Indian Languages [5]	PNN and KNN classification techniques.	82%
6	K- Nearest Correlated Neighbour Classification for Indian Sign language Gesture Recognition using feature[6]	KNN Classification technique.	Double handed -97.50% Single handed - 91.11%
7	Double Handed Indian Sign Language to Speech and text [7]	Eigenvalue algorithm andMATLAB	-
8	The Comparison of Some Hidden Markov Models for Sign Language Recognition[8]	Gaussian Hidden MarkovModel Multinomial Hidden Markov Model	83% 72%
9	A Real-Time System For Recognition Of American Sign Language By Using Deep Learning[9]	CNN, Skin Detection and Convex Hull	98.05%
10	Real Time Sign Language Recognition using the Leap Motion Controller[10]	Leap Motion Controller and Multi-Layer Perceptron	96.15%
11	Hand-Gesture Recognition for Automated Speech Generation[11]	HSV model LargeBlob Detection, Flood Fill and Contour Extraction.	-
12	3D-CNN-Based Fused Feature Maps with LSTM applied to Action Recognition[12]	Hidden Markov Model,flex sensor.	99%
13	A Real-time Hand Gesture Recognition and Human-Computer Interaction System[13]	Robust control of mouse and key-board events with a higher accuracy of gesture recognition	-
14	Dynamic hand gesture recognition using vision-based approach for human–computer interaction[14]	Various challenges such as complicated background, change in illumination and occlusion are removed in this system to make the hand detection more accurate.	-
15	Hand Gesture Recognition for Human Computer Interaction[15]	A marker free gesture recognition system. Detects both static and dynamic gestures.	84%-94% in plain background

International Journal of Engineering Research & Science (IJOER)
ISBN:[978-93-5391-287-1] ISSN: [2395-6992] [Vol-6, Issue-3, March- 2020]

IV. CONCLUSION

The proposed system is an approach to translate sign language to text or speech. Due to the complex signs, using only normal data alone may not be optimal for translation for signs. In particular, this system reduces the communication gap between dumb and deaf people and the normal people. The steady growth of technology all over the world, it demands to have a system which will be able to translate sign languages to a spoken language. The output of the proposed system will be able to translate words or phrases which are commonly used while having a conversation. Considering the use of CNN to translate sign language into text will yield better results and accuracy.

ACKNOWLEDGEMENTS

We would like to express a deep sense of gratitude towards our mentor Prof. DnyaneshwarBhabad, Department of Computer Engineering for his constant encouragement and valuable suggestions. The work that we have been able to present is possible because of timely guidance and support.

REFERENCES

[1] Vi N.T. Truong, Chuan-Kai Yang, Quoc-Viet Tran, "A Translator for American Sign Language to Text and Speech", IEEE, Japan, 2016.

[2] Mahesh M, ArvindJayaprakash, Geetha M, "Sign Language Translator for mobile platforms", IEEE, India, 2017, pp 1176-118.

[3] KusumikaKrori Dutta, Sunny ArokiaSwamy Bellary, "Machine Learning Techniques for Indian Sign Language Recognition", ICCTCEEC, India, 2017, pp 4896-4899.

[4] KshitijBantupalli, Ying Xie, "American Sign Language Recognition using Deep Learning and Computer Vision", IEEE, USA, 2018.

[5] Umang Patel, Aarti G. Ambekar, "Moment Based Sign Language Recognition For Indian Languages", IEEE, India, 2017.

[6] Bhumika Gupta, Pushkar Shukla, Ankush Mittal, "K-Nearest Correlated Neighbor Classification for Indian Sign Language Gesture Recognition using Feature Fusion", ICCCI, India, 2016.

[7] KusurnikaKrori Dutta, Satheesh Kumar Raju, Anil Kumar G, Sunny ArokiaSwarny, "Double Handed Indian Sign Language to Speech and Text", ICIIP, India, 2015, pp 374-377.

[8] Suharjito, Herman Gunawan, NaradaThiracitta, GunawanWitjaksono, "The Comparison of Some Hidden Markov Models for Sign Language Recognition", IEEE, Indonesia, 2018, pp 6-10.

[9] Murat Taskiran, Mehmet Killioglu, and NihanKahraman, "A Real-Time System For Recognition Of American Sign Language By Using Deep Learning", IEEE, Turkey, 2018, pp 258-261.

[10] DeepaliNaglot, Milind Kulkarni, "Real Time Sign Language Recognition using the Leap Motion Controller", IEEE, India, 2017.

[11] Sunny Patel, UjjayanDhar, SurajGangwani, Rohit Lad, PallaviAhire, "Hand-Gesture Recognition for Automated Speech Generation", IEEE, India, 2016, pp 226-231.

[12] Aarthi M, Vijayalakshmi P, "Sign Language to Speech Conversion", Fifth International Conference on Recent Trends in Information Technology, India, 2016, pp 1-6.

[13] Pei Xu, "A Real-time Hand Gesture Recognition and Human-Computer Interaction System", Department of Electrical and Computer Engineering, University of Minnesota, Twin Cities.

[14] Joyeeta Singha1, Amarjit Roy1, RabulHussainLaskar, "Dynamic hand gesture recognition using vision-based approach for human–computer interaction", Department of ECE, NIT Silchar, Silchar, India.

[15] AashniHaria, Archanasri Subramanian, NivedhithaAsokkumar, ShristiPoddar, JyothiNayak, "Hand Gesture Recognition for Human Computer Interaction", 7th International Conference on Advances in Computing & Communications, ICACC-2017, Elsevier, 22-24 August 2017, Cochin, India, p. 367-374