SIGN LANGUAGE RECOGNITION: A SURVEY

Mokshak Ketan Dagli
MCA Department
Jain University, Bengaluru, Karnataka, India

Dr. Preeti Savant
MCA Department
Jain University, Bengaluru, Karnataka, India

Abstract— **Sign Language Recognition is a growing field in research where different people are trying to develop or propose a system that can help the specially-able people of the society in the best possible way so that they can interact and communicate with the people of the society easily and without the problem of them learning the sign language. This language is basically used by the deaf and dumb people of society. This paper describes various ways a sign language recognition system has been built or has been proposed by different researchers. It also aims to have a better understanding of various recognition systems and various methods of recognizing and predicting the hand signs/hand gestures of the specially able people of the society. In return to predict a value which is in a human-readable format that is text. This paper describes in detail about the various ways and how one method of sign language recognition is different from other based on performance, results, difficulty in developing the model, etc.**

Keywords— **K-Nearest, LDA, IOT Device, Static Sign, Dynamic Sign, TTS.**

I.   INTRODUCTION

According to the latest survey done by the World Health Organization as of Nov 2021, 15% of the total population are specially able out of which 5% to 7.5% are those who are deaf and dumb. They are the ones who cannot speak or listen to others. The mode of communication to these deaf and dumb people can only be done using Sign Language. Sign Language is a way to interact or communicate with people using hand gestures. Sign language uses hand gestures and other ways of non-verbal behaviors to convey their intended meaning [1]. So, it is very difficult for people to communicate or interact with these specially able people using sign language as it is difficult for everyone to know and understand sign language.

Therefore, to make things easier for these specially able people to interact with people of the society a sign language recognition application is to be developed which can be used to predict or recognize the hand signs or the hand gestures demonstrated by the specially able person. As people face a lot of difficulties in understanding or figuring the hand signs so to remove the hindrance and to remove the barrier between them a system needs to be developed to translate or predict the hand signs of people into a human-readable format such as text. A lot of technologies have already been developed as this is a growing field around research where technologies like machine learning, deep learning, etc. have been used to develop a system to solve and predict hand signs.

II.   APPROACHES

As per the growth in the field of technology in terms of machine learning, deep learning, neural networks, IoT, etc., there are various ways the Sign language recognition system can be developed based on various platforms. Sign language recognition can be developed in two forms either using Static Signs or using Dynamic Signs. There are multiple ways and approaches to how you can achieve Sign Language recognition using these forms.

A few ways to which you can achieve this is by developing a machine learning model that predicts the hand sign, an IoT device that can detect the hand sign, etc. These machine learning models are developed in various ways for example using the K-nearest Neighbor algorithm, LDA (Linear Discriminant Analysis) algorithm, etc. which is used to predict the hand sign depicted by the user. A few models have been explained in the latter part of this section.

A.   Static Sign Based Recognition –

Most of the research done under sign language recognition are performed using static signs. These research and method use various machine learning algorithms, python libraries, third party libraries, etc. to achieve it sign language recognition.

Omar Rashid & et. al [2] used support vector machines to recognize one hand finger of American Sign Language. It is only used to recognize static number and alphabets of ASL. Ali Karaim and et. al [3] has developed a system that can detect 32 static alphabets of Persian sign language. Neural networks are used for classification process and manually cropping is used for segmentation. The rate of accuracy using this model is 94.06%. Ravi Kiran and et. al [4] used fingertip

detection to recognize hand sign. No use of special gloves is used in this system. Various other researchers have used support vector machine and IOT gloves to detect various sign in different languages.

B. Dynamic Sign Based Recognition –

Dynamic Sign Based recognition format uses a lot of technologies combined to achieve it. It has a very high scope in the future as it can help people get more realistic and positive response from these systems. These systems development requires high concentration and high programming skills to get the most accurate value. These systems also include emotion detection which can help in a more effective manner with advanced words recognition.

Hong Li and et. al [5] used contour-based feature to recognize dynamic single hand sign. Dynamic programming is used for this system. The accuracy achieved using this system is 88% to 96%. The same method is used by Yang and et.al [6] with an additional feature in terms of predicting value, where the end results are compared with CRF and latent dynamic CRF. Due to this additional feature accuracy rate has increased by 40%. Hand signs formed using two hands can be detected using a system developed by Anup nanny and et.al [7]. They have used direction histograms, Euclidean distance, and K-nearest neighbor for classification. The recognition part is done used the hidden Markov model. Accuracy obtained using this system is 93.2%.

III. RELATED WORK

A. Conversion Of Sign Language To Text & Speech Using Machine Learning Techniques.

Victoria A Adewale & Adegoke O Olamiti used a machine-learning algorithm to develop the Sign Language Recognition system. They used the K-Nearest Neighbor algorithm of machine learning to predict the most accurate value based on the input provided by the user. The most difficult part of the system was to perform image segmentation and object recognition which was achieved with the help of SURF and FAST algorithm in the field of machine learning. These algorithms were used to perform various functionalities like feature extraction, feature selection, etc. In this process, an image is captured and extracted from the system to perform prediction. The extracted image is then processed where the non-required objects are removed from the image to perform better and accurate predictions.

After cleaning the image, it is passed to the K-nearest algorithm to perform the prediction was based on the data set the object in the image has been compared based on this a predicted value is passed to the output screen. The output screen includes two formats of output where one includes text which is easy for any human to read and understand and the format is speech for those who cannot read text data. This text

to speech feature is achieved using a text to speech library in python. The pictorial representation of the system is depicted in the Fig 1.0. Therefore, this is a system which can be used only through laptops and computers enabled with camera and speakers, so it makes it difficult for people to carry it everywhere. This system has scored an accuracy of 78%. This system requires a vast data set, a better feature extraction, and selecting algorithm so that it can provide a cleaner image for predicting the value.
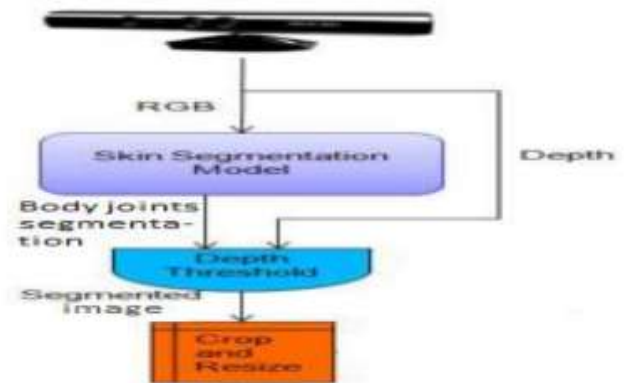

Fig 1.0: Architecture of the system.

B. Conversion Of Sign Language Into Text.

Mahesh Kumar used a combination of algorithms and machine learning techniques to develop this system. The main algorithm used to predict and recognize the hand sign from the provided input was achieved through LDA - Linear Discriminant Analysis method. LDA explains the model the difference between the class of data. This is a dimensionality reduction technique that was used for the classification of data. In this model, the first input needs to be pre-processed where the image/input is cleaned using various algorithm where image acquisition and segmentation is performed. Feature extraction and selection are also performed before running the model for predicting the data.
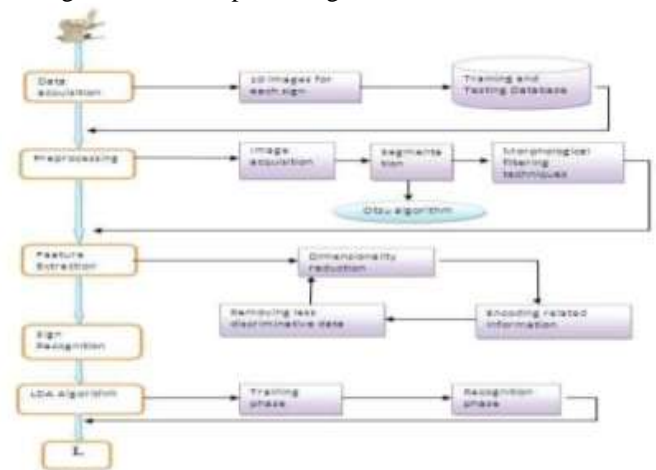

Fig 2.0: Steps in LDA for this system.

As shown in the Fig 2.0 this system the model is developed in the following manner where:

I. We compute the d-dimensional means vectors for the different classes from the data set.
II. Compute the scatter matrices.
III. Compute the eigenvectors and respective eigenvalues for the scatter matrices.
IV. Sort the eigenvectors by decreasing eigenvalues and choose the K eigenvectors with the largest eigenvalues to form a d * k dimensional matrix.
V. Use the matrix to transform the sample into the new subspace.

This system is developed using the LDA algorithm for recognition with the help of various other algorithms for data pre-processing. The system can only be deployed on a laptop which eventually solves the problem but cannot be accessible by people in a handy manner.

C. Real time Indian Sign Language Recognition System to aid deaf-dumb people.

P. S. Rajam & G. Balakrishnan used a unique way to predict the hand gesture or the hand signs. In this model, the hand sign is predicted based on the fingers captured in the image. For instance, when an image is captured by the camera of the device, that image is then passed for pre-processing which is used to clean and extract the important and necessary features from the input image which includes the hand sign right from the palm to the tip of every finger. In this system, every fingertip must be visible. In the proposed model 32 combinations of binary number signs are developed which are loaded at run time. It is developed in such a fashion that it can identify which finger is raised and which is not. This means that the unique binary number formation is done based on the fingers raised by the user for the hand sign. Every finger right from the middle, ring, index, little, and thumb finger are necessary to be visible in the image. Even if the finger is pointed downwards, it still needs to be visible. This unique method helps the machine to more accurately predict the hand sign shown by the user. This algorithm identifies which finger is raised by the user and then maps it to the data set which is a combination of the binary set.
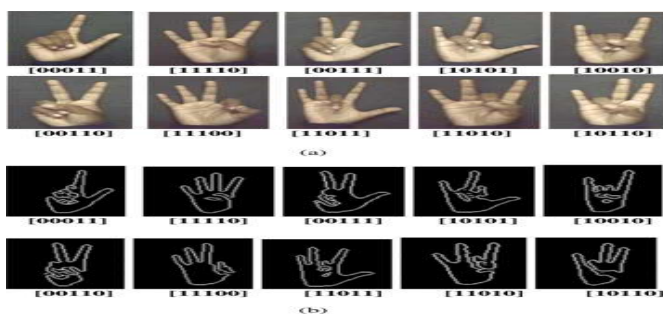

Fig 3.0: Sign Language Recognition using Binary Codes.

As shown in the above Fig 3.0 where when a user wants to depict a letter "L" he symbolizes it with 2 fingers pointing upwards which is the index and thumb finger whereas all the other fingers are pointing downwards. This device then captures an image of this hand sign in such a manner where the whole palm is visible. Then the image of the palm along with the fingers is passed for pre-processing where the feature extraction, selection, cleaning of the image, etc. is done. The image is then converted into a grayscale image which is later passed to the machine learning algorithm which recognizes which fingers are pointed upwards and generated a binary code which is then compared with the one from the data set to check if it is available. The one closest to it will be displayed as the output or will be considered as the result.

The pros and cons of this model are that it has a fixed number of binary conversions of the hand signs. If the model fails to predict or recognize the find tips then it cannot generate the corresponding binary version. It can only be used from a device like a laptop, or a computer enabled with a camera which leads it to a non-handy model. After performing many tests, the accuracy of this model is up to 96.87% which is very high when compared to the other models discussed above.

D. Sign Language To Speech Conversion.

P. Vijayalakshmi & M. Aarthi used a combination of algorithms and the latest technology which is IoT (Internet of things) to develop a Sign Language Recognition System. In this proposed idea a huge number of sensors and their modules are used to perform or predict hand spelling which is also known as the hand sign. Here an IOT device in the shape of Glows is developed with a lot of sensors attached to it which produces an enormous number of values from the movement of the hand and its fingers.

Sensors used in the system with its usage are as follows:

I. Flex Sensor - It is used to measure the degree to which the fingers are bent.
II. Accelerometer Sensor - It is used to measure the degree to which the finger is tilted.
III. Tactile Sensor - It is used to measure the physical interaction between the fingers.
IV. Arduino Microcontroller - It is used as an intermediate between the sensor and the model which recognizes the hand sign.

As depicted in the Fig 4.0 where for instance, when a hand sign is depicted by a user wearing these glows huge amount of data which is generated by the sensors is passed to the Arduino board which collects these values and passes it to the calculating phase of the board which performs various calculations based on the collected data. This microcontroller used here is a multi-purpose unit as it is used to collect values from the sensors and perform a calculation based on the

collected data. The collected data from the sensors which get continuously streamed to the microcontroller compares these readings with the predefined threshold values which are fed to the microcontroller in advance and then a corresponding gesture is recognized and respectively a text is displayed. This system also has TTS (Text to speech) module integrated into it which is used to convert the recognized text into a synthesized speech at run-time. An HMM (Hidden Markov Model) based Speech Synthesis system is used in this system to perform the text to speech conversion. The main parameters of the HMM-based speech synthesis system are the Training phase and Synthesis phase which is used to synthesize the speech in the English Language.
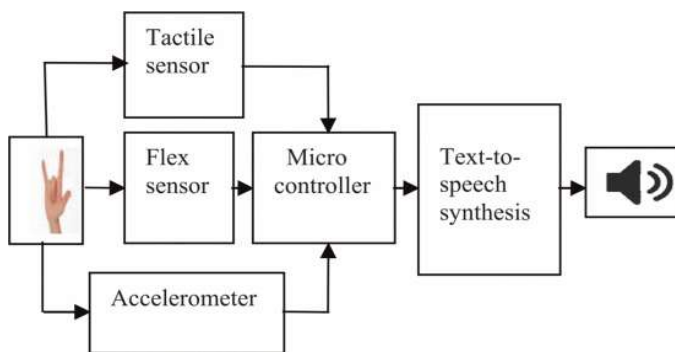

Fig 4.0: Flow of the sensors in the device.

Therefore, this system is more advanced and has an integration of an IoT device which is a glow fitted with a lot of sensors that retrieve huge amounts of data on the movement of the hand. The pros and cons of this idea are that this system is not handy it always requires the user to wear a glow that is connected to an Arduino board which is used to predict the value and so on. If more signs are to be recognizable then more predefined values are to be incorporated within the board in advance. The average accuracy using the proposed system is at a rate of 87.5% across all gestures.

IV. CONCLUSION

This paper analyzes various methods and technologies used to develop Sign Language Recognition Systems. It explains in detail about different systems in terms of execution, result, accuracy obtained, pros & cons. This paper will give you a short summary about different research papers related to Sign Language Recognition Systems.

V. REFERENCES

[1]. P. Lokhande, R. Prajapati, and S. Pansare, "Data gloves for sign language recognition system," IJCA Proceedings on National Conference on Emerging Trends in Advanced Communication Technologies, vol. NCETACT 2015, no. 1, pp. 11–14, 2015.

[2]. O. Rashid, A. Al-Hamadi, and B. Michaelis, "Utilizing invariant descriptors for finger spelling American sign language using SVM," in Proceedings of the 6th international conference on Advances in visual computing - Volume Part I, 2010, pp. 253–263.

[3]. A. Karami, B. Zanj, and A. K. Sarkaleh, "Persian sign language (PSL) recognition using wavelet transform and neural networks," Expert Syst. Appl., vol. 38, no. 3, pp. 2661–2667, 2011.

[4]. J. Ravikiran, K. Mahesh, S. Mahishi, R. Dheeraj, S. Sudheender, and N. V. Pujari, "Automatic recognition of sign language images," in Lecture Notes in Electrical Engineering, Dordrecht: Springer Netherlands, 2009, pp. 321–332.

[5]. H. Li and M. Greenspan, "Model-based segmentation and recognition of dynamic gestures in continuous video streams," Pattern Recognition., vol. 44, no. 8, pp. 1614–1628, 2011.

[6]. R. Yang, S. Sarkar, and B. Loeding, "Handling movement epenthesis and hand segmentation ambiguities in continuous sign language recognition using nested dynamic programming," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 3, pp. 462–477, 2010.

[7]. A. Nandy, J. S. Prasad, S. Mondal, P. Chakraborty, and G. C. Nandi, "Recognition of isolated Indian sign language gesture in real time," in Communications in Computer and Information Science, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 102–107.

[8]. V. Adewale and A. Olamiti, "Conversion of Sign Language to text and speech using machine learning techniques," JOURNAL OF RESEARCH AND REVIEW IN SCIENCE, vol. 5, no. 1, 2018.

[9]. M. Kumar, "Conversion of Sign Language into text," Ripublication.com. [Online]. Available: https://www.ripublication.com/ijaer18/ijaerv13n9_90.pdf. [Accessed: 19-Jan-2022].

[10]. P. S. Rajam and G. Balakrishnan, "Real time Indian Sign Language Recognition System to aid deaf-dumb people," in 2011 IEEE 13th International Conference on Communication Technology, 2011, pp. 737–742.

[11]. P. Vijayalakshmi and M. Aarthi, "Sign language to speech conversion," in 2016 International Conference on Recent Trends in Information Technology (ICRTIT), 2016, pp. 1–6.