ISSN No:-2456-2165

# Sign Language Recognition using Machine Learning: A Survey

Anagha.G[1], Sinchana.S.Bharadwaj[1], Sumana.M.R[1], Varshini.Krishna.Mohan[1], Nagaraj.A[2*] ,
[1]Under Graduate Student,[2] Associate Professor,
Dept. Of Computer Science & Engineering Jyothy Institue of Technology,
Visvesvaraya Technological University Thataguni Post,
Bengaluru-560082, India

**Abstract:- It is said that more than one billion people in the world are disabled. One of the only ways they can communicate among themselves or with people who don't have this disability is sign language. Sign language is a creative way to communicate within the deaf community by using gestures done by hand and other means that do not require talking. It involves combining hand movements, their shapes, etc. Body movements and facial expressions are also taken into consideration. All these factors help convey the person's thoughts in a fluent manner. Most of the general public who aren't disabled have no knowledge about sign language. Even out of the few who are aware, the majority don't know how to use it for communication, this stops them from interacting with the deaf and mute people. Through this software, we want to raise awareness towards sign language and help bridge the gap by creating a sign language interpreter that recognizes hand gestures.**

***Keywords:-*** *Sign Language Recognition, Mobile Application, Convolutional Neural Network, Machine Learning.*

## I. INTRODUCTION

Sign language is a way of communication for the deaf and sign language recognition is the method through which the hand gestures in sign language are recognized by using images or videos.

Each country has its own sign language. There are more than 300 sign languages in the world, each differing from another based on what part of the world they belong to. Sign language may even be different in the same country based on the accent. region,etc. One of the most prominent sign language systems known is ASL which is the American standard language. ASL is a natural language with its own structure and unlike spoken English. It is a visual language that uses the movement and placement of hands to convey the meaning of words. Facial expressions and body movements also play an integral part in the language.

In India ISL, Indian Sign Language is used. This is different from ASL as it uses its own vocabulary and grammar. Even though ISL is used in India, a part of the deaf community still use ASL so for this proposed system we have decided to use ASL.

There are two different types of approaches to recognize sign language, they are contact-based systems and vision-based systems. Contact-based systems as the name suggests are based on contact. Gloves are used in these systems, which have sensors on them, and thus the movements are captured. Even though this system provides high accuracy it is not cost-effective. Vision-based systems are the ones in which the signs are detected using visions, for this purpose either static images are used or real-time images are captured. Sign language is all about perspective hence the vision-based is an effective method.

The proposed algorithm in this solution is CNN, convolutional neural network that is a part of deep learning. It is mostly used for analyzing visual representations. Tensor flow is also used which is a tool used to develop and train the model.

## II. LITERATURE SURVERY(ABSTRACT)

[1]In this publication, an effort has been made to highlight the work of American Sign Language researchers and compare their work. Asl - American sign language is used , as it is the most widely used sign language. This system uses Microsoft Kinect for image acquisition or extracting the dataset. Feature extraction is done using PCANet (Principal component analysis network) and the classification of of ASL letter is done using COnvolution neural network(CNN). The majority of SL and gesture identification problems have been solved using statistical modeling techniques such as PCA and support vector machines. The system proves to be 98% effective due to the fact that the application requires more features or distinction to be accurate. Less number of signs were used to train the model, leading to poor user experience.

[2]This system is a real-time sign language translator that translates sign language to text. It uses GoogLeNet architecture to train the data .It uses Convolution neural network to classify each frame in the video to a letter, and it reconstructs to display the most likely word from the classification.This is a a robust model that correctly classifies letters a-e in the majority of cases when used by first-time users, and another that correctly classifies letters a-k in the majority of situations Given the dataset restrictions and the promising findings obtained, we are hopeful that with further research and data, we will be able to build a fully generalizable translator for all ASL letters. Lack of variety in the datasets, the validation accuracies were not directly replicable when tested on the web application. Better letter segmentation is required, as well as a more smooth method for retrieving photographs from users at a faster rate.

[3]Sign Language Recognition is one of today's fastest-growing disciplines of study. In these disciplines, many innovative techniques have lately been developed. For the

ISSN No:-2456-2165

recognition of diverse Indian Sign Languages, the suggested method employs Eigen value weighted Euclidean distance as a classification technique. Skin Filtering, Hand Cropping, Feature Extraction, and Classification are the four aspects of the system. In this paper, 24 signs were investigated, each with 10 samples, for a total of 240 photos, with a recognition rate of 97 percent. Key highlight of this system is that signs can be done using both the hands. One drawback is that the alphabets 'H'and 'J'are not added as they are considered dynamic gestures.

[4]A system that uses hand detection as a learning tool for beginners in sign language has been developed. This system is based on explicit skin-color space thresholding, which is a skin-color modeling technique. The skin-color range has been predetermined to separate pixels (hand) from non-pixels (background). For image categorization, the photos were loaded into a model called the Convolutional Neural Network (CNN). Keras was utilized for image training.The system achieved an average testing accuracy of 93.67 percent when given suitable illumination and a uniform background, with 90.04 percent credited to ASL alphabet recognition, 93.44 percent for number recognition, and 97.52 percent for static word identification, outperforming prior related studies. The method is used to perform quick computations.

[5]SLR (sign language recognition) is a difficult yet crucial study subject for numerous computer vision systems.Attempt to improve communication between the deaf and the hearing.Hearing-impaired individuals. An accurate method is presented in this paper .and a comprehensive deep learning-based sign language technique. Video sequences are also recognised. The method is based on skeletal aspects of the hand and body retrieved from RGB videos and, As a result, it obtains highly discriminative skeletal data for gesture detection without the use of any additional equipment. Experimentation with a big publicly available dataset of sign language demonstrates the superiority of our methods over others.The proposed SLR system is the first attempt to combine a vision-based technique (i.e., video sequence processing) with reliable skeletal data extraction without the use of data gloves or other sensors that constrain a signer's mobility.This work introduces a novel SLR system that proposes the extraction and processing of hand and body skeletal data from video sequences, attempting to overcome the limitations of earlier approaches.

[6]This system proposes a system for the recognition of Mexican Sign Language. They are using a computer vision system. This system uses a digital camera and four LED reflectors to capture the sign language. They are using artificial neural networks as their pattern recognition model. The MSL has 27 signs in which 21 and static and in this system those 21 signs are being recognized.Multi Layer Perceptron is used to recognize signs and normalized central moments are used to represent the frames properly.

[7]This paper uses an open-source framework of Media Pipe and Support Vector Machine(SVM) to recognize the sign languages. They use four different types of datasets for this model, they are Indian, American, Italian, and Turkey .Media Pipe is a framework that helps people in developing audios and video data cross-platform in machine learning pipelines. This is done in three stages, the first stage is detecting and extraction of hand features, stage two is cleaning the data and the third stage is analyzing sign languages using machine learning algorithm. The average accuracy of this system is 99%.

[8]This paper is for recognizing the Indian Sign Language. The ISL is not as common as ASL because of the lack of a datasets, ISL also uses two hands to make gestures,which makes it hard to detect. Hence, considering all this paper proposes to bridge the communication gap. Machine learning algorithms like SVM and random forest are used for training. The proposed method in this paper is to capture the image,process it, extract the features and then classify them.

This system recognizes the sign languages by using Convolutional neural networks (CNNs), Microsoft Kinect and GPU acceleration. Italian gestures are recognized in this system.The dataset used is ChaLearn Looking at People 2014, which consists of 20 gestures by different users in different environments. So the video data is processed using 2D convolutions, one CNN is used for extracting hand gestures and the other one for upper body features. Data augmentation is used during training to reduce overfitting.The accuracy is up to 91.70%

This system uses smart gloves to recognize the gesture that the person wants to portray. It uses artificial language to achieve this. The normal glove solutions only detect motions of hands or can only recognize single gestures. So this system proposes a system which can recognize sentences as well and with more accuracy. The method of segmentation and non segmentation is used for this feat which is assisted in deep learning model. A total of 20 sentences and 50 words are recognized. The gestures that are recognized by the sensors in the glove are projected into a virtual place and they are then translated to text and audio. The average rate of correctness is 86.67%.

ISSN No:-2456-2165

## III. LITERATURE SURVEY

The below list (**TABLE I)** outlines survey of papers related to the topic in brief with possible gaps/limitations within the proposed system.

| Papers | Title | Authors | Year Of Publication | Proposed system | Gaps |
|---|---|---|---|---|---|
| [1] | Static sign language recognition using deep learning | Lean Karlo S. Tolentino, Ronnie O. Serfa Juan, August C. Thio-ac, Maria Abigail B. Pamahoy, Joni Rose R. Forteza, and Xavier Jet O. Garcia | 2019 | A sign language learning system which is based on the modelling technique of skin color. A specific range is agreed upon that will extract pixels from non pixels. | This system can recognize only static sign languages. |
| [2] | Sign language recognition based on hand and body skeletal data | Dimitrios Konstantinidis, Kosmas Dimitropoulos, and Petros Daras | | A methodology for sign language recognition that uses hand and skeletal features that are extracted from RGB videos. | Hand joints are considered for recognition and other hand parts may cover the said joints. |
| [3] | Real-time vernacular sign language recognition using mediapipe and machine learning | Arpita Halder and Akshit Tayade | 2021 | A sign language recognition system which uses MediaPipe's open framework to detect hand gestures. | This system only recognizes alphabets and numbers. |
| [4] | Real-time American Sign Language Recognition with Convolutional Neural Networks | Brandon Garcia and Sigberto Alarcon Viesca | | Real-time sign language translator that uses GoogleNet architecture to train the data to classify each frame in the video to a letter, and reconstructs characters to display the most likely word from the classification. | The validation accuracies are not directly replicable when tested on the web application. Better letter segmentation is required. |
| [5] | Review on sign language detection using machine learning | Anamika Srivastava1 and Vikrant Malik | 2020 | The proposed system uses Microsoft Kinect for image acquisition and PCA net architecture for feature extraction. SVM technique is used to classify the data and get desired output. | Application requires more features or distinction to be accurate. Less number of signs were used to train the model, leading to poor user experience. |
| [6] | Sign language recognition using machine learning algorithm | Radha S.Shirbhate,Vedant D. Shinde, Sanam A.Metkari, Pooja U. Borkar, and Mayuri A | 2020 | A sign language recognition system that uses a variety of algorithms such a SVM, Random forest and Hierarchical classification. | The system can be useful for static ISL numeral signs only. |
| [7] | Sign language recognition using convolutional neural networks | Lionel Pigou(B), Sander Dieleman, Pieter-Jan Kindermans, and Benjamin Schrauwen | | Sign language recognition system using 2D Convolutions and pooling method | The system only recognises 20 Italian gestures. |
| | Indian sign language recognition using eigen value weighted | Joyeeta Singha and Karen Das | | An Indian sign language recognition system using Eigen value weighted | The system cannot recognize the letters 'H' and 'J' as they are |

| | | | | | |
|---|---|---|---|---|---|
| [8] | euclidean distance based classification technique | | 2013 | Euclidean distance as a classification technique. | dynamic gestures. |
| [9] | Automatic mexican sign language recognition using normalized moments and artificial neural networks | Francisco Solís, David Martínez, and Oscar Espinoza | 2016 | The system presents a computer vision system for automatic recognition of Mexican Sign Language(MSL). It uses a digital camera and 4 LED reflectors to reduce shadows and for improving segmentation of hand with background. | This system is trained only to recognize static images. LED reflectors are used which increases the cost. |
| [10] | AI enabled sign language recognition and VR space bidirectional communication using triboelectric smart glove | Feng Wen, Zixuan Zhang, Tianyiyi He, and Chengkuo Lee | 2021 | The system presents an artificial intelligence enabled sign language recognition system. It uses sensing gloves which are configured with 15 triboelectric sensors, using Convolutional Neural Networks(CNN) to process the input data. | It uses sensors which is not cost effective. It uses text from able bodies for communication process which is not very practical. |

Table 1

## IV. EXISTING SOLUTION

ASL recognition isn't a new problem in computer vision. Researchers have utilized classifiers from a range of categories over the last two decades, which we may generally divide into linear classifiers, neural networks, and Bayesian networks. The research was also based on a variety of input sensors, gesture segmentation, feature extraction, and classification approaches.

The proposed system analyzes the American Sign Language gestures and then converts them into human readable text.The majority of the works designed to address this problem have used one of two approaches: contact-based systems, such as system gloves, or vision-based systems, which rely solely on cameras. Contact based method is the one where the signer has to wear a hardware glove and do the signs , and the hand movements get captured . This system is uncomfortable for practical and daily use, despite having an accuracy of 90%.Static and dynamic recognition are two types of vision-based methods. Static is concerned with the identification of static gestures (two-dimensional images), whereas dynamic is concerned with the capturing of motions in real time. This entails the employment of a camera to record motion.Vision is a key factor in sign language, and every sign language is intended to be understood by one person located in front of another, from this perspective, a gesture can be completely observable.

Sensor-based devices, such as SignSpeak, were employed in many research. This device used a variety of sensors, including flex and contact sensors for finger and palm movements, as well as accelerometers and gyros for hand movement; the gloves were then trained to recognise different gestures using Principal Component Analysis, and each gesture was then classified into alphabets in real time.In visual-based SLR, several strategies have been established.Many people experimented with image and video processing since sign language comprises both static and dynamic gestures.

Another existing solution for communication for hearing impaired people is a chat application. Chat programmes have evolved into a great tool for individuals to connect with one another in a variety of languages. There are many different chat software that are utilized by different individuals in different languages, but there isn't one that allows you to interact with sign languages. The recent release of low-cost depth sensors, such as the widely used Microsoft Kinect sensor, has facilitated the development of new gesture detection algorithms. Depth photographs generate a three-dimensional model of the scene, which can be used to ease tasks like people segmentation and tracking, body component detection, and motion prediction.

## V. PROPOSED SOLUTION

Developing an android application aided by machine learning techniques to recognize hand gestures can be done with ease. Hand gesture recognition shows how fast the algorithms detect the gestures in a single shot. The faster and stable it can be, the user experience will be smoother and better. The proposed SLR system constitutes the first attempt to merge a vision-based approach (i.e., processing of images) with the accurate extraction of skeletal data without employing data gloves or other sensors that limit the movements of a signer. The application opens to a camera, which detects the hand movements of the signer. Hand gesture features in the uploaded images are extracted and used to recognize the type of gesture.

- **How different does our sign language interpreter work using Convolutional Neural Network as compared to previous efforts?**
  - We are developing a real-time sign language interpreter that can seamlessly recognize and phrase American Sign Language gestures.
  - This technology is a first of its kind to be implemented in an android application which will help a user understand sign language anywhere anytime with just the use of a phone.
  - In terms of convenience, a desktop or web application would always require a desktop or laptop and it would be impractical to make the disabled person sit in front of a desktop to communicate with an able body.
  - This project intends to educate users about sign language while helping them to translate sign language to English by using the sign language scanner feature.

## VI. CONCLUSION

Sign language recognition is a difficult task in today's world due to the gap between able-bodied and hearing impaired people. The proposed system is to make this task easier with the help of an android application. Sign language is recognized by taking the input from camera and then recognizing which hand gesture is made by the user. We have deduced few algorithms that can be used for sign language recognition. This system can be used to educate the able-bodied and bridge the gap between able-bodied and hearing impaired. It takes real-time hand gestures and gives the output. The resources being used are well suited with today's technologies making it robust and accurate.

## REFERENCES

[1.] Lean Karlo S. Tolentino, Ronnie O. Serfa Juan, August C. Thio-ac, Maria Abigail B. Pamahoy, Joni Rose R. Forteza, and Xavier Jet O. Garcia, "Static sign language recognition using deep learning," International Journal of Machine Learning and Computing, Vol. 9, No. 6, December 2019.

[2.] Dimitrios Konstantinidis, Kosmas Dimitropoulos, and Petros Daras, "Sign language recognition based on hand and body skeletal data," ITI-CERTH, 6th km Harilaou-Thermi, 57001, Thessaloniki, Greece.

[3.] Arpita Halder and Akshit Tayade, "Real-time vernacular sign language recognition using mediapipe and machine learning," International Journal of Research Publication and Reviews Vol (2) Issue (5) (2021).

[4.] Brandon Garcia and Sigberto Alarcon Viesca, "Real-time American Sign Language Recognition with Convolutional Neural Networks".

[5.] Anamika Srivastava1 and Vikrant Malik, "Review on sign language detection using machine learning," Journal of Critical Reviews ISSN- 2394-5125 Vol 7, Issue 10, 2020.

[6.] Radha S. Shirbhate, Vedant D. Shinde, Sanam A.Metkari, Pooja U. Borkar, and Mayuri A. Khandge, "Sign language recognition using machine learning algorithm," International Journal of Engineering and Technology, Volume: 07 Issue: 03, Mar 2020.

[7.] Lionel Pigou(B), Sander Dieleman, Pieter-Jan Kindermans, and Benjamin Schrauwen, "Sign language recognition using convolutional neural networks," ELIS, Ghent University, Ghent, Belgium.

[8.] Joyeeta Singha and Karen Das, "Indian sign language recognition using eigen value weighted euclidean distance based classification technique," International Journal of Advanced Computer Science and Applications, Vol. 4, No. 2, 2013.

[9.] Francisco Solís, David Martínez, and Oscar Espinoza, "Automatic mexican sign language recognition using normalized moments and artificial neural networks," Scientific Research Publishing, 2016.

[10.] Feng Wen, Zixuan Zhang, Tianyiyi He, and Chengkuo Lee, "AI enabled sign language recognition and VR space bidirectional communication using triboelectric smart glove," Nature Communications, 2021.