International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 SJIF 2023: 8.176 ISSN: 2582-3930

Emblem: Sign Language Recognition Based on Convolutional Neural Network

Roshani A.Ahire
Department of computer Engineering
Brahma valley college of Engineering,
Nashik. Maharashtra, India.
Email : roshuahire@gmail.com

Poonam R. Gaikwad
Department of computer Engineering
Brahma valley college of Engineering,
Nashik. Maharashtra, India.
Mail: gaikwadpoonam1234@gmail.com

Aishwarya A. Bhamre
Department of computer Engineering
Brahma valley college of Engineering,
Nashik. Maharashtra, India.
Email: bhamre.aish123@gmail.com

Amol K. Kharule
Department of computer Engineering
Brahma valley college of Engineering, Nashik.
Maharashtra, India.
Email : amolkharule1@gmail.com

Harshal V. Kadam
Department of computer Engineering
Brahma valley college of Engineering, Nashik.
Maharashtra, India.
Email : Harshkdn666@gmailcom

ABSTRACT- **Voice and Language is the main thing for human to communicate with each other. Due to hearing ability we can understand thoughts of each other. Even nowadays we can give commands using voice recognition. But what if one absolutely cannot hear anything and eventually cannot speak. So The Sign Language is the main communicating tool for hearing impaired and mute people, and also to ensure an independent life for them, the automatic interpretation of sign language is an extensive research area. With the use of image processing and artificial intelligence, many techniques and algorithms have been developed in this area. Every sign language recognition system is trained for recognizing the signs and converting them into required pattern. The proposed system aim to provide speech to speechless, in this paper the double handed Indian Sign Language is captured as a series of images and it's processed with the help of Python and then it's converted to speech and text.**

***Index Terms*- Convolutional neural network, Gesture recognition, Image processing, conversion, American sign language, image processing, convolutional neural network, data glass.**

I. INTRODUCTION

Sign languages are vivid on wide and world level. There are multiple sign languages in world which are regular in use that are ASL (American Sign Language) ISL (Indian Sign Language), BSL (Bangladesh Sign Language), MSL (Malaysian Sign Language). These languages are Built and Developed with lots of work and practical testing with intention of feasibility to the deaf and dumb persons. Any language is created with its word and its meaning. Sign Language is created as "Sign" and "Action of That Sign". Because here we are not able to make them understand meaning of sign by writing word. As they are deaf and cannot listen from birth so we cannot teach them words.

Sign language is a system of non-verbal communication using visual gestures, symbols, facial expressions, body language etc. without the need of vocal communication. Any sign language has three main level of communication. The first one is the finger spelling which means that for every alphabet in the language, there are corresponding signs allocated. This is the most basic form of sign

language. Every word can be spelled by the proper use of the signs in sequence. The second level of communication is that there are some special signs which correspond to a whole word to speed up thewhole communication process. The third one is the communication through the body language, eye movement, facial expression etc.But this form of communication is usually not preferred due to its misinterpretation and tardiness. So, among these three, the first two level of communication are the most used by the hearing disabled community for its speed and reliability. So, in this paper we presented a system of interpreting American Sign Language. People with hearing disabilities will use sign language in front of a raspberry-pi camera and then the images extracted through the camera will be processed by the raspberry-pi and matched against the existing database and then the interpreted result will be shown by a data glass.

Given that sign language could be a visual language, it's no surprise that you simply also can learn to higher perceive visual communication. Signing is that the communication secretly. Whether or not its clubs, film theatres or libraries, you do not have to whisper or yell to speak to your friends in sticky state of affairs. The remainder of paper deals with the prevailing system, planned models and implementations.

II. MOTIVATION OF THE PROJECT

We are motivated with aim to use new technologies for better humanity. We found Machine learning like technologies can be used for conquering the backwardness occurred because of this physical disability.

III. PROBLEM DEFINATION & OBJECTIVES

A random person if visited to deaf person and if deaf person is in problem and trying to explain it then it is very difficult to understand what exactly he is trying to say. Delay in detecting his Sign Language can turn into big critical problem for that deaf person. These kind of people cannot spend normal life. They face communication issues at every point. Also they get boundaries and limitations to their dreams and professional aims. Hence they get demotivated and Inferiority Complex

IV. RESEARCH CHALLENGES.

Objective is to give them ability to be expressive in ideas and thoughts. They can get helped in increasing their motivation and confidence and it will help them to think positively and to conquer that physical disability. To develop system with using latest technologies and tools we are keeping objective to overcome from this global level problem. This system will definitely can be- come step into innovation of this global level problem solution. Our system can be Prototype and Proof of Concept for global level solution. This system can be used by Deaf and Deaf persons and also normal person can have this system with them and deaf person can perform sign in from of camera and sign can be converted to text or speech. The mute community all over the globe facing many problems while communicating. The normal and dumb people can communicate only in one way i.e. sign language, but many times communicating with normal persons they noticed difficulty

V. METHODOLOGY

MACHINE LEARNING ALGORITHMS

CNN :-Artificial Intelligence has been witnessing a monumental growth in bridging the gap between the capabilities of humans and machines. Re- searchers and enthusiasts alike, work on numerous aspects of the field to make amazing things happen. One of many such areas is the domain of Computer Vision. The agenda for this field is to enable machines to view the world as humans do, perceive it in a similar manner and even use the knowledge for a multitude of tasks such as Image Video recognition, Image Analysis Classification, Media Recreation, Recommendation Systems, Natural Language Processing, etc. The advancements in Computer Vision with Deep Learning has been constructed and perfected with time, primarily over one particular algorithm — a Convolutional Neural Network.

A Convolutional Neural Network (Convent/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics. The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area.

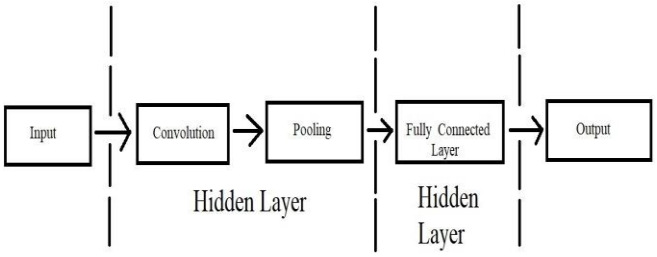

Fig. 5.1 Block diagram.

Convolutional Neural Network is state of the art neural network model for recognizing what's in an image or what an image is. Basic operations of a convolutional neural network are shown in the block diagram. A. Convolution Convolution is a form of mapping or creating a feature map from a primary dataset. Basically, there is going to be a moving region over the entire image either looking for something or trying to create a feature map with a region of pixels (n*n) in order to classify the image.

Our convolutional neural network model consists of one input layer, 7 hidden layers, 1 fully connected layer and 1 output layer. For optimization, we used Adam optimizer and we trained the model for 10 epochs.

The figure 5.1, gives a simplistic view of how a neural network works. Again, the inputs are X1, X2 and X3. Consider the little circles as artificial neurons. The inputs are connected to the each of the neurons and each of the connections has a unique weight associated with it. This is the 1st hidden layer of neuron and each of the neurons are connected to next layer of neurons (2nd hidden layer) with another unique set of weights and then they are connected to output layer and again with each unique weight . This is an example of a deep neural network. Because if a model has one hidden layer, it is termed as a regular neural network but if it has more than one, it is called a deep neural network.

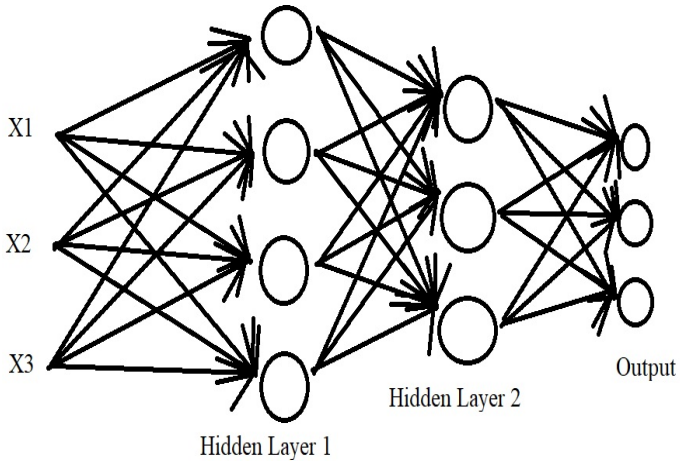

Fig. 5.2 Convolutional Neural Network.

VI. SYSTEM DESIGN

Feature Extraction: Feature extraction involves reducing the number of resources required to describe a large set of data. When performing analysis of complex data one ofthe major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power, also it may cause a classification algorithm to over fit to training samples and generalize poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy. Many machine learning practitioners believe that properly optimized feature extraction is the key to effective model construction. Results can be improved using constructed sets of application-dependent features, typically built by an expert. One such process is called feature engineering.

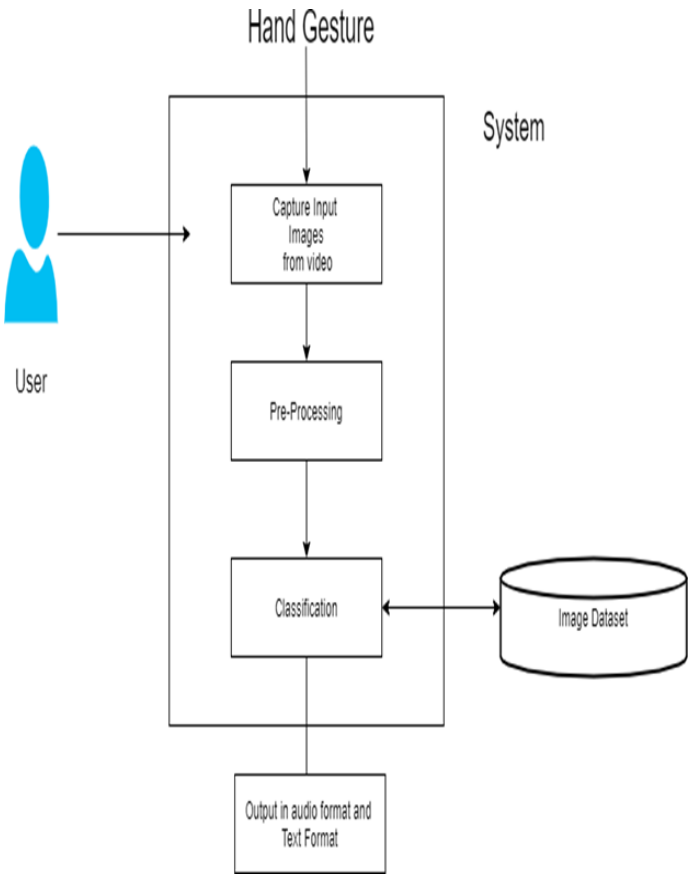

Fig. 6.1 System Architecture.

Sign language brings many benefits to all children regardless of whether they are deaf or struggling with their hearing. As well as helping children to communicate and fully express themselves, it also improves their social skills by increasing their confidence and self-esteem. This app was designed by bilingual (English and ASL) Deaf people and is meant to teach conversational ASL. Using more than 1000 videos, it's packed with features to make learning ASL fun and easy.

VII. EXISTING SYSTEM AND NEED FOR SYSTEM

In the existing systems, BSL uses a two-handed fingerspelling system, com- pared to the one-handed system used in ASL (and FSL). Many American Deaf believe that one handed finger-spelling makes for faster finger-spelling than two-handed systems. However, anecdotal evidence has it that in a challenge between proficient ASL and BSL speakers, neither finger-spelling system proved to be faster; both finished reciting the alphabet at the same time. So that supposed "disadvantage" is rendered invalid. According to many Europeans, American signers tend to fingerspell "too much" compared to the rate of finger-spelling in many European sign languages, including BSL. This may be true; several examples of BSL signs for concepts that do not have a sign in ASL and are often finger-spelled for lack of a formal sign. This is one of the advantages of BSL, but that is not intrinsic to the language itself and it reveals a cultural value. On the other hand, that many BSL signs are often derived from their initialized (English) base, while many ASL signs have been developed without initialization (Including the influence of signed English systems) so one might see that as a "disadvantage".

Detecting Sign Language Characters in Real Time Using Media Pipe and Keras. Sign Language is a form of communication used primarily by people hard of hearing or deaf. This type of gesture-based language allows people to convey ideas and thoughts easily overcoming the barriers caused by difficulties from hearing issues.

VIII. EXPERIMENTAL RESULT

The figure 9 and 10 show accuracy and loss of the trained model. From the graph we can see that we were able to attain an accuracyof 91.25% and loss was reduced to 0.40669 which is far better than that can be achieved through any other algorithm.

Letter ”P” interpreted

Letter ”R” interpreted.

This model is able to recognize 43 letters and gestures (26 alphabets,7 gestures and 10 Numbers). At first, we had trained the model with 5 hidden layers and for 5 epochs and with a dataset of 103,200 samples (2400 samples each). But we were able to achieve an accuracy level of only 74.37%. So, we increased the no. of samples, hidden layers and epochs. Then we were able to achieve an accuracy of 90.25%.

IX. CONCLUSION

Sign Language is a tool to reduce the communication gap between deaf-mute people and normal person. This system which is proposed above gives the methodology which aims to do the same as the two-way communication is possible. This method proposed here facilitates the conversion on the sign into speech. This overcomes the requirement of a translator since real time conversion is used. The system acts a voice of the person who is deaf-mute. This project is a step towards helping a specially challenged people.

This can be further enhanced by making it more user friendly, efficient, portable, compatible for more signs and as well as dynamic signs. This can be further improvised so as to making it compatible for the mobile phones using the built-in camera of the phone. We can increase the distance at which it can be used by using a longer trans-receiver module or over Wi-Fi.

X. FUTURE WORK

In future work, proposed system can be developed and implemented using Raspberry Pi. Image processing part should be improved so that In future work, proposed system can be developed and implemented using Raspberry Pi.

Image Processing part should be improved so that System would be able to communicate in both directions i.e.it should be capable of converting normal language to sign language and vice versa. We will try to recognize signs which include motion.

Moreover we will focus on converting the sequence of gestures into text i.e. word and sentences and then converting it into the speech which can be heard. We’ll also focus on multiple languages in one system like American, Indian, Malesian, Rassian Sign Language. We can also create Mobile Application In future work.

XI. REFERENCES

1) Galib Ibne Haidar, Hasin Ishraq Reefat, 2020 ,Glove Based American Sign Language Interpretation Using Convolutional Neural Network and Data Glass.

2) Neha Poddar, Shrushti Rao, Shruti Sawant, Vrushali Somavanshi, Prof. Sumita Chandak (Feb 2015) "Study of Sign Language Translation us- ing Gesture Recognition", International Journal of Advanced Research in Computer and Communication Engineering.

3) Jayshree R. Pansare, Maya Ingle (2016)," Vision-Based Approach for American Sign Language Recognition Using Edge Orientation Histogram", at 2016 International Conference on Image, Vision and Computing.

4) Arslan Arif, Syed Tahir Hussain Rizvi, Iqra Jawaid, Muhammad Adam Waleed, Techno-Talk: An American Sign Language (ASL) Translator, at CoDIT'16 - April 6-8, 2016, Malta.

5) Justin K. Chen, Debabrata Sengupta, Rukmani Ravi Sundaram," Sign Language Gesture Recognition with Unsupervised Feature Learning"

6) Matheesha Fernando, Janaka Wijayanayaka,"Low cost approach for Real Time Sign Language Recognition" at 2013 IEEE 8th International Conference on Industrial and Information Systems, ICIIS 2013, Aug. 18-20, 2013, Sri.

7) Aarthi M Vijayalakshmi, "SIGN LANGUAGE TO SPEECH CON-VERSION", at 2016 FIFTH INTERNATIONAL CONFERENCE ON RECENT TRENDS IN INFORMATION.

8) Caixia Wu, Chong Pan,YufengJin, Shengli Sun, and Guangyi Shi Shaox- ing, "Improvement of Chinese Sign Language Translation System based on Collaboration of Arm and Finger Sensing Nodes", At The 6th An- nual IEEE International Conference on Cyber Technology in Automa- tion, Control and Intelligent Systems June 19-22, 2016.

9) Poonam Chavan1, Prof. Tushar Ghorpade2, Prof. Puja Padiya, "In- dian Sign Language to Forecast Text using Leap Motion Sensor and RF Classifier" at 2016 Symposium on Colossal Data Analysis and Net- working (CDAN).

10) Ms.Manisha D. Raut, Ms. Pallavi Dhok, Mr. Ketan Machhale, Ms. Jaspreet Manjeet Hora, "A System for Recognition of Indian Sign Lan- guage for Deaf People using Otsu's Algorithm", International Research Journal of Engineering and Technology (IRJET)

11) C. Swapna, Shabnam Shaikh "Literature survey on hand gesture recog- nition for video processing" International Journal of Advanced Re- search in Computer Science and Software Engineering (IJARCSSE).