International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
Impact Factor (2020): 6.733

Indian Sign Language Recognition System for Deaf and Dumb Using CNN

Srushti Raut[1], Panav Patel[2], Sohan Vichare[3], Gayatri Hegde[4], Rahul Durvas[5]

[1, 2, 3, 5]Department of Information Technology, Pillai College of Engineering, Panvel, India

[4]Professor, Department of Information Technology, Pillai College of Engineering, Panvel, India

Abstract: ***One of the nonverbal communication methods utilized in sign language is the hand sign. It is mostly used by deaf and dumb people to communicate with other people or among themselves having difficulty in speaking or hearing. Many people throughout the world have created numerous sign language project which consists of alphabets and numbers, but the end users find them neither versatile nor cost-effective. This project shows a conceptual framework capable of automatically recognizing sign language in order to assist deaf and dumb individuals in communicating more efficiently with everyone. The people in society frequently overlook or reject those who are considered to be stupid or illiterate. Additionally, it is exceedingly challenging for regular people to understand and converse with them. Typically, these folks rely on some other person or a translator in order to communicate with the society. A translator or a person won't be available every time to communicate. The community of the deaf and dumb uses sign language exclusively as a means of communication.***

Keywords: Sign Language, CNN, Open-CV, Hand Gesture.

1. Introduction

To understand the world around us and make sense of life, each of us has a certain way to comprehend and communicate. sizable population is cut off from the mainstream hearing-dominated society and at risk of losing credibility because they are unable to communicate with those who can only utilize speech. Among these is the use of sign language. Instead of using words to communicate, it uses body gestures to create a visual language. To communicate, the language uses visible indicators like as hands, eyes, facial emotions, and gestures. Although sign language is mostly used by those who are deaf, many hearing people also use it. Sign language, like any spoken language, has its own grammar and principles that has evolved over time. There is no "universal" sign language, just like there isn't one for spoken languages. The sign language used in various country varies depending on the region and culture. For example, Indian Sign Language (ISL) is different from American Sign Language (ASL) which is also different from British Sign Language (BSL).

Objectives

a) To build an open-source software to provide robust and accurate hand gesture recognition and tracking.
b) To build a system that can help normal people interact with speech and hearing-impaired people with ease.
c) Making dumb individuals feel comfortable expressing their emotions to normal people is the goal of our endeavor.
d) To build a system to be able to interpret the few hand signs made by the deaf people.

2. Literature Survey

Hand gestures are frequently used in sign language as a method of nonverbal communication. In addition, people with hearing disabilities or speech impairments use it to communicate with one another and with non-disabled people. It is possible to distinguish between letters and digits with today's technology. This project is an addition to the current system that will enable the system to recognize text in the form of words and phrases. Real-time hand tracking, hand segmentation, feature extraction, and gesture detection are the four modules. The technology that is being developed would be a real-time system that would use image processing to interpret live motions of signs. Among the methods used for processing data, such as images, and for image analysis and classification are Open CV (an image processing package) and CNN (a deep learning model). Accurately identifying single-handed and double-handed movements is made simple and affordable by this method. Millions of deaf people may undoubtedly benefit from using this technology to interact with normal people. One of the most important factors in a society's ability to survive is communication. Deaf and dumb people interact with one another through sign language, but it might be difficult for non-deaf people to understand them. Thus, communication is accomplished using sign language.

Despite intensive study of American sign language, Indian sign language is very different from it. The ISL language uses two hands (20 of the 26 available), while the ASL language uses only one hand. Features are usually hidden when using both hands because the hands overlap. Additionally, efforts to detect ISL gestures have been hampered by a paucity of datasets and the fact that regional variations in sign language exist. With the help of this technology, hearing people and the deaf will be able to communicate with one another, which is the first step toward using Indian sign language. This research's expansion to words and common phrases will not only make it easier for dumb and deaf people to communicate with others, but it may also help develop autonomous systems that can understand and support them. The system's objective is to identify Indian Sign Language alphabets using the associated gesture. In American Sign Language, this topic has received extensive research, whereas in Indian Sign

Language it has gotten minimal attention. Instead of using high-end tools like gloves or the Kinect to address this problem, our approach is to identify gestures in photos taken with a webcam, extract certain features from those features, and then classify the features using computer vision and machine learning methods.

A number of earlier efforts in this subject were considered. Various researchers produced these various works. Different researchers employed a range of techniques, such as vision-based, data-glove-based, artificial neural network, fuzzy logic, genetic algorithm, hidden Markov model, support vector machine, etc. The works that have come before are listed below. Numerous studies have been conducted on various facets of this strategy. The suggested framework's examination of the literature reveals that numerous attempts to solve sign identification in videos and images utilizing different techniques and algorithms have been made.

Review Paper on Sign Language Recognition for the Deaf and Dumb by R Rumana, Reddygari Sandhya Rani, Mrs. R. Prema, in year 2021 [1]:

For the deaf and dumb, an efficient real-time vision-based system for recognizing American sign language has been developed for asl alphabets. Throughout the endeavor, there were numerous difficulties. The dataset was a very first problem. It intended to handle raw photographs as well, just like CNN in Keras, because it was much more convenient to work with square photos alone. They opted to create their own dataset because it was impossible to discover an existing dataset for that. The second problem was choosing which filter to apply to the images in order to extract the proper characteristics and use that image as input for the CNN model. They experimented with a variety of filters, including binary threshold, canny edge detection, and gaussian blur, but in the end decided to use the gaussian blur filter. The model's accuracy continued to have issues after it was trained in earlier stages, but these issues were eventually fixed by extending the input image size and also by upgrading the dataset.

Indian Sign Language Recognition using Convolutional Neural Network by Rachana Patil, Abhishek Bahuguna, and Mr. Gaurav Datkhile, in the year 2021 [2]: In this study, The CamShift approach was used to separate the human hand from the complex background and detect real-time hand gestures. The recognition of the area of hand movements that were observed in real time using a convolutional neural network is followed by the identification of 10 frequently used digits. The proposed system comprises a dataset with 400 images for each category, 4000 hand gestures, and a total of 1600 images for training. This experiment yields a 98.3 percent accuracy rate.

Machine Learning Techniques for Indian Sign Language Recognition by **Kusumika Krori Dutta, in the year 2017 [3]:** In this, the dataset was divided into two groups in, with one serving as training and the other as testing. 70% of the total data make up the training set, and the other 30% are utilized for testing. We also did experiments on the same (30% or 70%) dataset that is used for CNN classifier training and testing. The outcomes of these experiments are completely accurate.

Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique by **Joyeeta Singha and Karen Daas in the year 2013 [4]:** In this study many techniques including Skin Filtering, Hand cropping, classifier, and feature extraction were applied. Using MATLAB 7.6 (R2008a) and related hardware, the suggested system was put into practice. For this system, 24 alphabets of Indian sign language were taken into consideration, each with 10 samples, for a total of 240 images that were photographed. Different photos were analyzed, and it was discovered that the new categorization method had a 97% accuracy rate. Using the Eigen value weighted Euclidean distance between Eigen vectors as a classification method, the success rate increased from 87% to 97%. With a success rate of 97%, it attempted to increase the recognition rate in comparison to earlier works. Additionally, we took into account both hands in our paper. Since the visuals were so dynamic, the letters "H" and "J" from the 26 alphabets were not taken into consideration. In the future, we hope to cope with fluid gestures. Furthermore, only 240 photos were taken into account in this paper; therefore, we intend to expand it in the future.

Artificial Neural Network Based Method for Indian Sign Language Recognitionby **Adithya V., Vinod P. R., Usha Gopalakrishnan, in the year 2013 [5]:**In this, a method based on artificial neural networks was suggested for recognizing Indian sign language. It proposed a technique to find a 32-set of combinations—10 for each up and down movement of the fingers—in order to get matching Tamil letters. The method required converting decimal numbers from the up/down position of the fingers into categories that could recognize Tamil alphabets. A collection of static data was recorded as 640 x 480-pixel pictures. Images were converted from RGB to greyscale using palm image extraction. The experimental finding showed that the dynamic approach had a reported accuracy rate of 98.75% and a recognition rate of 96.87% for the static method.

3. Proposed System

The proposed system is a real-time system where the live sign gestures will be recognized in words or sentences. OpenCV (an image processing library) is used for database creation. CNN (a deep learning model) is used for processing data such as images and is used for image recognition and classification. After predicting the gesture, the text is converted to speech.

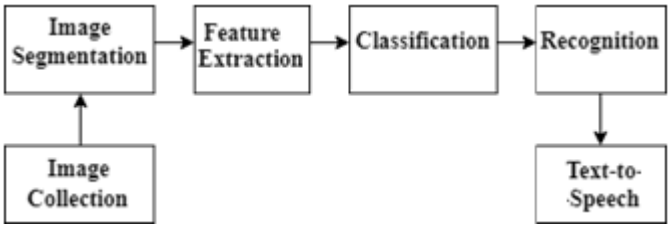


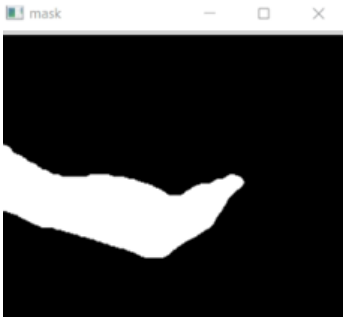
Figure 1: Proposed System Methodology

The proposed System consists of the following steps:

Image Collection: In this, images were captured in real time using Open-CV to create a dataset as there is no dataset available for words of Indian Sign Language.

Image Segmentation: The process of image segmenting involves taking objects or signs out of the surrounding environment of a captured image. The segmentation procedure makes use ofmasking, and canny edge detection. To identify gestures, the hand's position is be detected and segmented. As real-time image captured by camera contains a collection of frames, the RGB image frame is converted and so the input images are converted to grayscale and after converting, the medium blur is applied to remove unnecessary noise.

Masking: By using masking, one can inform sequence-processing layers that some time steps in an input are absent and should be omitted during data processing.



Canny Edge Detection: Canny Edge Detection: A multi-stage method is used by the Canny edge detector, an edge detection operator, to identify a variety of edges in images. With the use of the Canny edge detection technology, the amount of data that needs to be processed can be drastically reduced while still extracting meaningful structural information from various vision objects. It is frequently used in many computer vision systems. Canny has discovered that the prerequisites for using edge detection on various vision systems are rather simple.

Features Extraction: Feature extraction is the creation of new and smaller set of features that captures most of the features of the raw data. Predefined features such as geometrical feature (position, angle, distance, etc.), color feature, and others are extracted that are later used as signs for classification or recognition. The machine store images in the form of numbers and cannot recognize images that are in their original form. So, the raw image is first converted into gray scale image and then is transformed into a matrix of numbers which ranges from 0 to 255 these numbers indicate the intensity of brightness of the pixel. Smaller number denotes black color (closer to 0) and larger number denotes white color (closer to 255).

Training Dataset: A proper dataset is created of the gestures so that the images captured while communicating can be compared. The dataset is split into two sets i.e. training set and test set. This consists of random sampling without replacement about 80 percent of the rows and putting them into the training set. The remaining 20 percent is put into the test set.

Classification: CNN is the classifier that is employed in this project. CNN is utilized for picture characterization and identification because of their high precision. 32 filters of size 3x3 is applied to the input image in the first layer of CNN along with the ReLu activation function so that returns a positive value and zero even if the input number is negative. A MaxPooling layer to downscale the output of CNN is applied next adding another convolution layer with 32 filters of 3x3 followed by another MaxPooling layer. For the 3rd CNN layer, a layer of 64 filters of

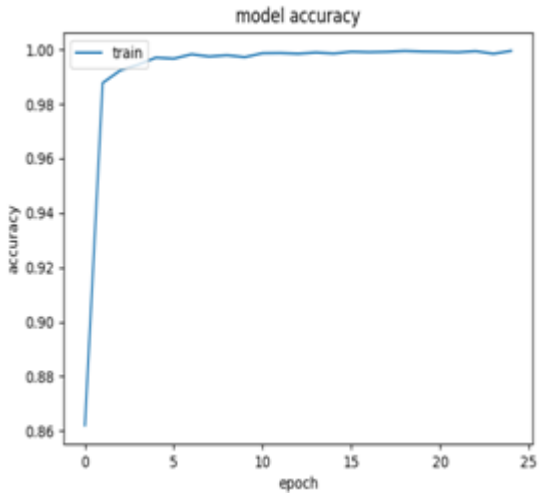
3x3 size is added followed by another MaxPooling layer. Then the output of the 3rd MaxPooling layer is flattened into a 1D array which is used as a input to the fully connected layer. The ReLu activation function and 256 neurons followed by a Dropout layer is added with fully connected layer that drops 50 percent of neurons to prevent overfitting.

Text-to-speech: The hand sign recognized by CNN is displayed and then is converted to speech using the python text to speech (pyttsx3) library.

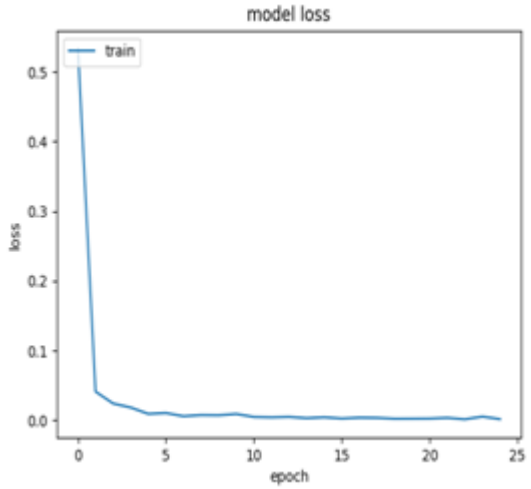
4. Results

An accuracy of 99% is achieved over 25 epoch which recognizes all 26 Alphabets, Numbers ranging from 1-9 and Words, some of which are Age, Address, Bank, Country, Dance, Football, Give me, Hello, How are you, Home, Language, Learn, Love, Name, Namaste, Play, Sign, Study, Thank You, Time, Water bottle, What, You, Your.

a) Graph Plot for Accuracy



b) Graph Plot for Loss



5. Conclusion

The primary form of communication for Deaf and Dumb persons in India is Indian Sign Language. This study provides a comprehensive CNN Model implementation for the recognition of Indian sign language. In section IV, Step wise Implementation has been discussed which includes Image Collection, Image Segmentation, Feature Extraction, Classification, and Text-to-Speech. In section V, the results are presented. Real Time Recognition of hand sign gestures is shown in the results and after displaying the text output, it can be converted into speech.

References

[1] A Review Paper on Sign Language Recognition for The Deaf and Dumb published by R Rumana, Reddygari Sandhya Rani, and Mrs. R. Prema, in the year 2021.
[2] Indian Sign Language Recognition using Convolutional Neural Network published by Rachana Patil, Abhishek Bahuguna, and Mr. Gaurav Datkhile, in the year 2021.
[3] Machine Learning Techniques for Indian Sign Language Recognition published by Kusumika Krori Dutta, in the year 2017.
[4] Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique published by Joyeeta Singha, Karen Daas, in the year 2013.
[5] Indian Sign Artificial Neural Network Based Method for Indian Sign Language Recognition published by Adithya V., Vinod P. R., and Usha Gopalakrishnan, in the year 2013.
[6] Translation of Sign Language for Deaf and Dumb People published by Suthagar S., K. S. Tamilselvan, P. Balakumar, B. Rajalakshmi, C. Roshini, in the year 2020.
[7] Sign Language Recognition: State of The Art. ARPN Journal of Engineering and Applied Sciences Published by Sahoo, A. K., G. S. Mishra, and K. K. Ravulakollu in the year 2014.
[8] Sign-Language Recognition-System Published by S. Shirbhate1, Mr. Vedant D. Shinde2, Ms. Sanam A. Metkari3, Ms. Pooja U. Borkar4, Ms. Mayuri A. Khandge in the year 2020.
[9] Hand Gesture Recognition for Sign Language Recognition: A Review", International Journal of Science, Engineering and Technology Research (IJSETR), Published by Pratibha Pandey, Vinay Jain, in the year March 2015.
[10] A Review on Hand Gesture Recognition System Published By J.SSonkusare, N.B Chopade, R.Sor, and S.L Tade, in the year 2017.
[11] Continuous Indian Sign Language Gesture Recognition and Sentence Formation, Kumud Tripathi, Neha Baranwal and G. C. Nandi, in year 2015.
[12] Sign Language Recognition Using Convolutional Neural Networks, Pigou, L., Dieleman, S., Kindermans, P.-J., Schrauwen, B, in the year 2014.
[13] Deep Convolutional Neural Networks for Sign Language Recognition, G.A. Rao, K. Syamala, P.V.V Kishore, A.S.C.S. Sastry, in the year 2018.