

Conversion of Sign Language into Text Using Machine Learning Technique

Sayali Gore^{1*}, Namrata Salvi², Swati Singh³

^{1,2,3}Department of Information Technology, Bharati Vidyapeeth College of Engineering, Navi Mumbai, India

Abstract: Communication is giving, receiving or exchanging ideas, information, signals or messages through appropriate media, to give information or to express emotions. It is very important and basic need for human beings. But there are some people in our society who born dumb and deaf or deaf due to some medical issues. These people faced many challenges while communicating with each other and also to the normal people. Sign language is one of the commonly used method by these people for communication. For this translator is required for communication between the person who knows the sign language and to whom they want convey their message. But for many instances translator is not available which creates a communication gap. This can be overcome with the help of using Machine learning algorithm. The main aim of this work is to provide such system which will convert the hand gestures into text using Convolutional Neural Network (CNN) algorithm. This will help deaf and dumb people to communicate very efficiently.

Keywords: Convolution Neural Network, Machine Learning, OpenCV, Sign conversion, Sign Language.

1. Introduction

The communication brings people together, closer to each other. It bridges the gap between individuals and groups through flow of information. But the people with hearing and speaking inability find it very difficult to communicate. Sign language serves a key to overcome this barrier. Impaired people find it very difficult to understand the non-verbal communication as a result sign language interpreter is required to overcome this challenge.

The proposed work aims to build such system which will convert hand gestures into text using CNN algorithm. It is an algorithm which takes image as an input, assign importance to various objects in the image and able to classify various features from each other. The trained dataset will pass through the several layers of CNN such as convolutional layer, pooling layer and fully connected layer. The output layer uses Softmax function to predict the result.

2. Motivation

Communication is one of the fundamental requirement to the existence and survival for the human Beings in the society. We got motivated to build such system by observing our society in day-to-day life. So many people in our country facing the communication problem such as deaf and dumb people i.e.

person who cannot speak or hear properly. These people have difficulties such as hearing and speaking while communication. So, for the better communication this system will provide a bridge for communication. Machine learning is getting popular in all industries with the main purpose of improving revenue and decreasing costs; by using Machine learning technique they automate and optimize their process to solve challenging tasks very efficiently [6], [7].

3. Problem Statement

Sign language is the basic communication method used by hearing impaired people. Hand gestures are used for communication purpose and these people face problems in communicating with other people without a translator. The proposed system aims to fill the communication gap using machine learning by creating a system in which hand gesture can be converted into text using CNN algorithm.

4. Literature Review

Hand gesture recognition interface for visually impaired and blind people. Issue-1, May 2019. The different technology used in this paper was Naïve Bayes, Neural Networks, Decision Tree, (SVM) Support Vector Machine with the accuracy of 88%, 90%. It is introduced due to the need of developing an electronic device that can translate sign language into speech in order to make the communication take place between the mute communities with the general public possible [1].

A System for Accelerometer-Based Gesture Classification Using Artificial Neural Network. The technology used in this paper Hidden Neuron a Bluetooth-enabled IMU mounted on the wrist provides hand motion trajectory information to a local terminal. A multi-layer feed forward artificial neural network (ML-FFNN). Gesture recognition is a technique which is used to understand and analyze the human body language and interact with the user accordingly. This in turn helps in building a bridge between the machine and the user to communicate with each other [2].

Hand Gesture to Speech Conversion using Matlab. The technology used in this paper were Skin colour segmentation hand gesture recognition speech playback. As computer innovation keeps on developing, the requirement for characteristic correspondence amongst people and machines

*Corresponding author: sayaligore2110@gmail.com

additionally increments [3].

Hand Gesture Recognition and Voice Conversion System for Dumb People. The technology used in this paper was Smart speaking system for mute people Raspberry Pi, Flex Sensors, Accelerometer Sensor. Gesture recognition is the technique which is used for the interaction between human and computer or it can be for any automation implementation purpose [4].

Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine Techniques. The technology used in this paper K-means, object detection, object recognition, scale invariant feature transform (SIFT), support vector machine (SVM). The communication between a dumb and hearing person poses to be an important disadvantage compared to communication between blind and ancient visual people [5].

5. Methodology

The proposed recognition technique relies on a convolutional neural network model (CNN) with a feature mapped output layer. The input is captured through the camera using OpenCV and passed through the CNN classifier. In the classifier, the captured input is initially taken and passed through the convolution layer. This layer extracts features from the images. The next layer is the Pooling layer, which reduces the number of parameters when the images are too large. The number of convolution layer and pooling layer can be increased for more accuracy. After passing through several convolution and pooling layers the output is flattened into a vector and sent into a fully connected layer. The last layer is the Softmax layer. This layer is used to show the output.

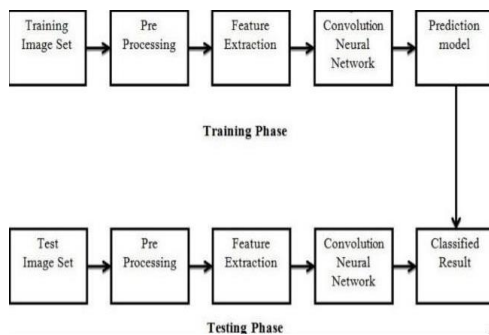


Fig. 1. Block diagram

6. Requirements

A. Dataset

A large dataset of various images of sign languages of special gestures and alphabets and numerical used for more accuracy. Multiple images of a single posture are taken in different conditions and lighting which gives high accuracy rate.

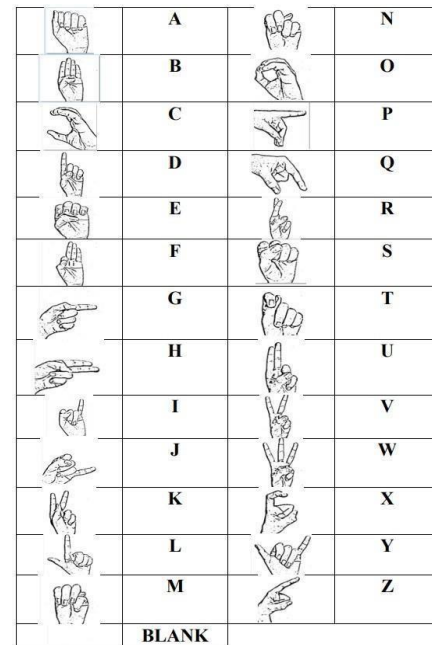


Fig. 2. Dataset diagrams

B. CNN Model

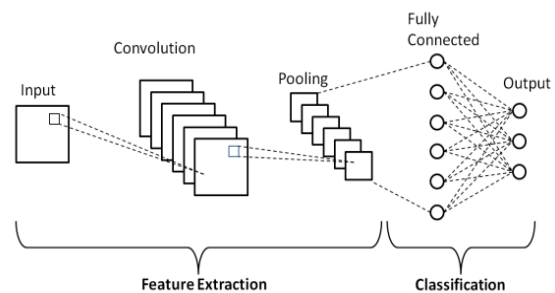


Fig. 3. CNN model diagram

CNN model takes image as an input, process it and classify it under certain categories. In CNN the input images are passed through several layers like the convolution layer, pooling layer, fully connected layers which consists of activation functions.

C. Python Libraries

1. **OpenCV:** OpenCv stands for Open Computer Vision which is a library in Python used for Image Processing. This will be required to perform various actions and processing of the captured images.
2. **Matplot library:** Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy.
3. **Numpy:** Numpy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.

7. Result

- The application translates sign language into the English alphabet.
- The text generated as the output is then converted into

sentence format.

- By using this application dumb and deaf people can communicate with others easily without facing any barriers. This system will decrease the communication gap between the deaf and dumb people

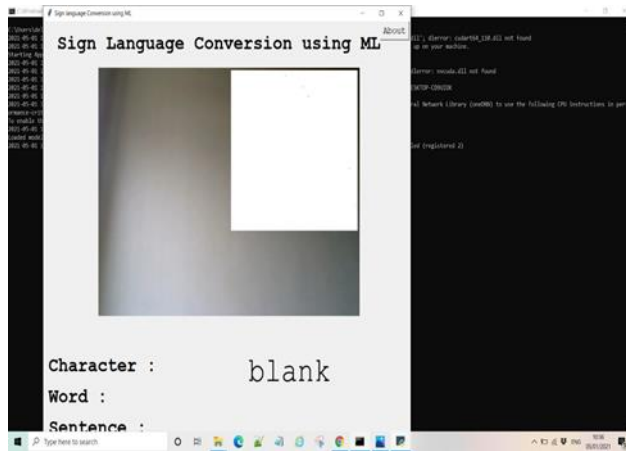


Fig. 4. User Interface

8. Conclusion

In this paper, we have proposed a conversion of sign language into text using CNN. This algorithm is particularly

used to improve the recognition accuracy under challenging conditions such as a change in scale, rotation and translation. Using large dataset increases the accuracy of the result. This proposed work will be beneficial to decrease the communication gap between impaired people.

References

- [1] Ashwini M. Deshpande, Snehal R. Kalbhor, "Video-Based Marathi Sign Language Recognition and text Conversion using Convolutional Neural Network", Conference Paper 2019.
- [2] Hitesh Kumar Sharma, Punit Kumar, Prashant Ahlawat, Nishank Parashar, Rakesh Soni, Yash Manchanda, "Deep Learning Based Accurate Hand Gesture Recognition Using Enhanced CNN Model, June 2020.
- [3] Nasser H. Dardas and Nicolas D. Georganas, Fellow, "Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine Techniques," IEEE 2011.
- [4] Robert M. Stephenson, Member, IEEE, Ganesh R. Naik, Senior Member, IEEE and Rifai Chai, "A System for Accelerometer-Based Gesture Classification Using Artificial Neural Networks".
- [5] "Recognition of Urdu Handwritten Characters Using Convolutional Neural Network" L3i Lab, Université de La Rochelle Av. Michel Crépeau, 17000 La Rochelle, France
- [6] Patil, V., Ingle, D.R. An association between fingerprint patterns with blood group and lifestyle based diseases: a review. *Artif. Intell. Rev* 54, 1803–1839 (2021).
- [7] P. N. Vijaykumar and D. R. Ingle, "A Novel Approach to Predict Blood Group using Fingerprint Map Reading," 2021 6th International Conference for Convergence in Technology (I2CT), 2021, pp. 1-7.