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

Sign Language Recognition based on hand symbol using Machine Learning

Kamlesh Patil[1] , Kshitija Patankar[2], Shraddha Lokhande[3], Vaishnavi Pathade[4], Mrunmayee Khaire[5]

[1] Information Technology Department, Bharati Vidyapeeth's College of Engineering for Women, Pune, Maharashtra, India. [2] Information Technology Department, Bharati Vidyapeeth's College of Engineering for Women, Pune, Maharashtra, India. [3] Information Technology Department, Bharati Vidyapeeth's College of Engineering for Women, Pune, Maharashtra, India.

[4] Information Technology Department, Bharati Vidyapeeth's College of Engineering for Women, Pune, Maharashtra, India.

[5] Information Technology Department, Bharati Vidyapeeth's College of Engineering for Women, Pune, Maharashtra, India.

Abstract - The communication among a deaf and listening person poses an excessive hassle. Gestural communication is called Sign language. Sign language uses gestures instead of sound to convey meaning, simultaneously combining hand shapes, orientations and movement of the hands and facial expressions to express fluidly a speaker's thoughts. Signs are used to communicate words and sentences to the audience. The purpose of the Sign Language Recognition system is to develop an algorithm for recognition of hand gestures with reasonable accuracy, where the input to the pattern recognition system will be given from the hand. It recognizes the pattern and displays the pattern in the form of the text. The information given by camera is collected and stored in the database. The exact meaning of the information is matched with the samples stored previously in the database and is printed. The image is processed considering the parameters like the number of fingers used, the angles between them and then the information is displayed in the form of text. The proposed system is user friendly, as it is easy to use and capable of building efficient and effective human computer interaction.

***Key Words*: Sign Language Recognition, Convolutional Neural Network, Application Programming Interface, Tensor-flow, OpenCV**

1. INTRODUCTION

It is quite difficult for the mute people to coordinate with the normal common people of society. Therefore, it is very necessary to develop such a smart system that can be helpful to such people so that they can communicate easily with the normal/common people of society. Sign language recognition is a project implementation for designing a model in which a web camera is used for capturing images of hand gestures which is done by open cv. After capturing images, labeling of images is required and then pre trained model SSD Mobile net v2 is used for sign recognition. Therefore, an effective path of communication is developed between deaf and normal audiences. Because of the advancements in machine learning, it is possible to detect sign language in real time. We have used the Tensorflow Object Detection pipeline, OpenCv python library and transfer learning to train the deep learning model that detects sign languages in Real time. Sign language is learned by mute people and usually it is not known to normal people, so it becomes a challenge for communication between a normal and mute person. There it striked our mind to bridge the gap between mute and normal people to make communication easier. The purpose of Sign Language Recognition (SLR) systems is to provide an efficient and accurate way to convert sign language into text for the hearing impaired or enable very young children to interact with computers(recognizing sign language), among others.

2. LITERATURE SURVEY

Deaf and mute people can interact with other people in the world by means of Sign language by using hand gestures. Due to advancement in the Machine Learning field, it is possible to detect Sign language. For designing a model in Sign language detection images are captured using a web camera which is done by open cv. Then after capturing images, labeling of images is done and then pre-trained model SSD is used for sign recognition. Three steps which must be followed in real time for implementing projects are namely: 1. Input can be obtaining footage of the user signing. 2. Next step is classifying each frame in the video to a sign. 3.Output will be reconstructing and displaying the most likely Sign from classification scores. The OpenCv python library, Tensorflow Object Detection pipeline and Machine learning model that detects sign languages in Real time are utilized for implementation of this project. A user defined dataset is used in this project. It is a collection of over 50 images. This dataset contains a total of 5 symbols i. e, Hello, Yes, No, I Love You and Thank You, which is useful while dealing with the real time application. SSD stands for single-shot multi box detector. From image processing images are pre-processed and hand is extracted from the background and a dataset is formed. The dataset contained 24 different signs of the

English alphabet which follows one-handed Indian Sign Language rules.The images captured by the camera were converted into a dataset containing 240 images consisting of 10 images of each sign. This dataset is used to train and test the Convolutional Neural Network model formed in the machine learning segment. All the steps are performed using Python programming language in the Visual Studio Code 2019 programming platform. Several libraries used in this project are OpenCV2, Pandas, Keras, Numpy etc. The Convolutional Neural Network proposed here is tested on both a custom-made dataset and also with real-time hand gestures performed by people of different skin tones. More than 500 people worldwide suffer from physical, sensory or mental abilities. Their lives are often hampered by physical or social barriers that make it difficult for them to participate fully in society and enjoy equal rights and opportunities. Many mute people use sign language in which they use hand gestures instead of sound to convey the meaning. Sign helps to communicate words and sentences to the audience. The input to the sign language recognition system will be given from the hand.

The system will recognize the pattern and display the pattern in the form of text or label. When mute people show their hand in front of the camera ,it is matched with the sample stored previously in the database and is printed. In this project, images in real time are captured from a USB camera and the detected frames are saved. When a hand gesture is detected, a bounding box is drawn around the gesture. Depending on the recognized gesture, a corresponding text is displayed in the result window. Communication between deaf-mute people becomes more difficult in everyday life. It is very difficult for mute people to coordinate with normal people around. Gestural type of communication called sign language which is widely used. They proposed a mute people communication exchange device that translates the hand gestures to text messages. Fingertip detection and the detections of centers of palms. The purpose of this project is to facilitate humans with the help of a deaf mute communication and to bridge the gap between mute and normal people.

3. SYSTEM MODEL AND ARCHITECTURE

The sign language recognition system is based on collected frames on a web camera of laptop or PC. OpenCV is used for image processing. Taking Photographs: Under order to gain improved accuracy through a huge dataset, several photographs of distinct sign language signals were collected from various angles and in variable light conditions.

Segmentation: After capturing the whole image, a specific section from the full image is picked that contains the important features and sign language symbols to be predicted. For the sign to be detected, bounding boxes are used. These boxes should be tightly packed around the picture region to be detected. The hand movements that were labeled were given specific names. The labeling was done with the Label Img tool. Image selection for training and testing purposes.

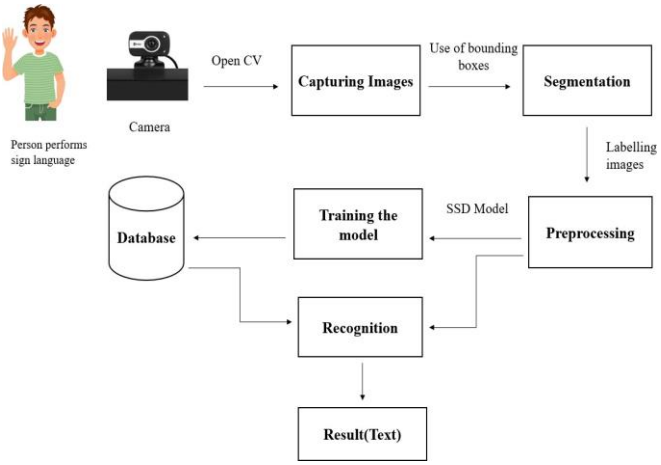


Fig. 1 System Architecture

4. ALGORITHM

Convolutional Neural Network(CNN):-

A Neural Network is a computational learning system that uses a network of functions to understand and translate a data input of one form into a desired output. The concept of the artificial neural network was inspired by human biology and the way neurons of the human brain function together to understand inputs from human senses. A convolutional neural network is a feed-forward neural network that is generally used to analyze visual images by processing data with grid-like topology. It's also known as a ConvNet. A convolutional neural network is used to detect and classify objects in an image. It uses a special technique called Convolution. This system accepts input in the form of a Live dataset. We know that we're doing data processing and training the dataset on the system, therefore we're employing modules: Pre-processing, Feature extraction, and classification, all of which use our CNN algorithm. So First Input as a Images Live dataset, then pre-processed the dataset (pre-processing step is clean the Image and Remove Blur part) After that, then the system extracts the parameters or features of the in the extraction section. Then, in classification, where we utilize our CNN algorithm to recognize Sign Language.

The network is made up of multiple layers, including convolutional layers, pooling layers, and fully connected layers. The convolutional layers apply filters to the input image to extract useful features, while the pooling layers downsample the image to reduce its size and computational complexity. Finally, the fully connected layers combine the features learned from the previous layers and use them to make predictions about the image.

The flow of CNN is illustrated in the Fig. 2

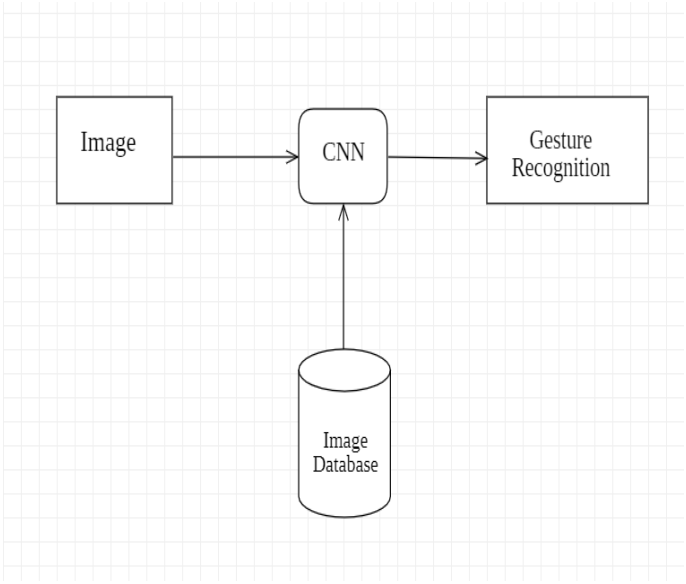


Fig. 2 Convolution Neural Network

A typical CNN consists of several types of layers :

1. Convolutional layer: This layer applies a set of learnable filters to the input image, performing convolution operations to extract features from the image. Each filter extracts a specific feature, such as edges or shapes, and produces a feature map.
2. Activation layer: This layer applies a non-linear activation function, such as ReLU (Rectified Linear Unit), to the output of the convolutional layer. This helps to introduce non-linearity and make the network more expressive.
3. Pooling layer: This layer reduces the spatial dimensions of the feature maps by performing pooling operations, such as max pooling or average pooling. This helps to reduce the computation cost and prevent overfitting.
4. Dropout layer: This layer randomly drops out a certain percentage of the neurons in the network during training, which helps to prevent overfitting.
5. Fully connected layer: This layer connects every neuron in the previous layer to every neuron in the current layer, and performs a linear transformation followed by an activation function.
6. Softmax layer: This layer is typically used for classification tasks, and applies the softmax function to the output of the last fully connected layer to obtain a probability distribution over the output classes.

These layers are typically stacked together in a feedforward manner to form the CNN architecture. The output of one layer serves as the input to the next layer, and the weights of the layers are learned through backpropagation during training.

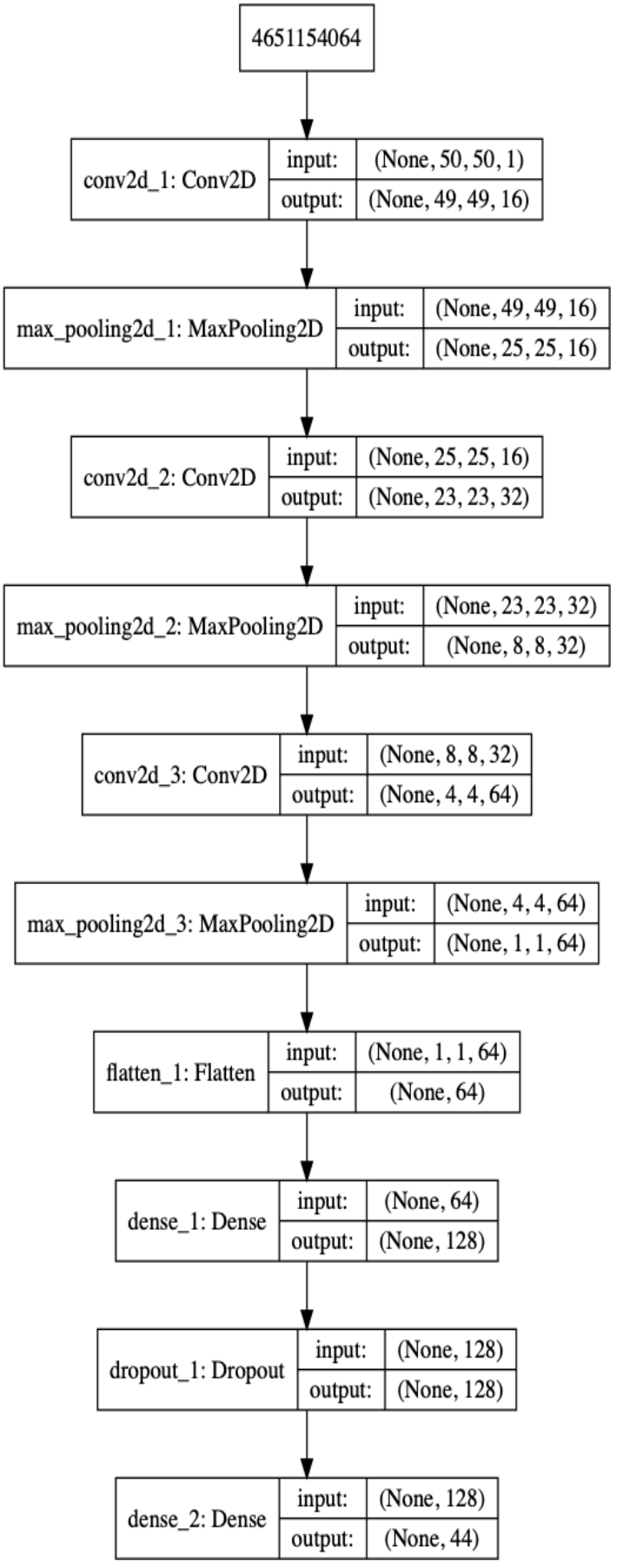


Fig. 3 CNN Model

Following figure illustrates the Confusion matrix for accuracyof the predicted labels.

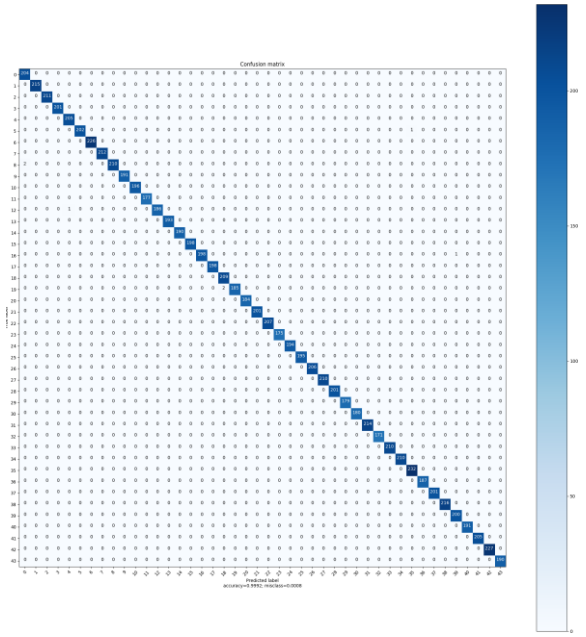


Fig. 4 Confusion Matrix

4. DATASET

For this project, a user defined dataset is used. It is a collection of over 500 images. This dataset contains a totalof 5 symbols i.e.Best of Luck,I love you, Like, You,
Remember. Also it contains letters from A to Z and also numbers from 0 to 9 which is quite useful while dealing withthe real time application.

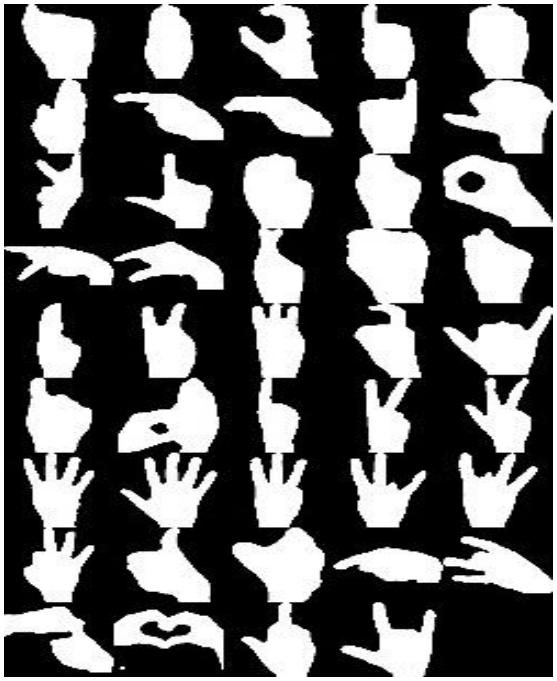
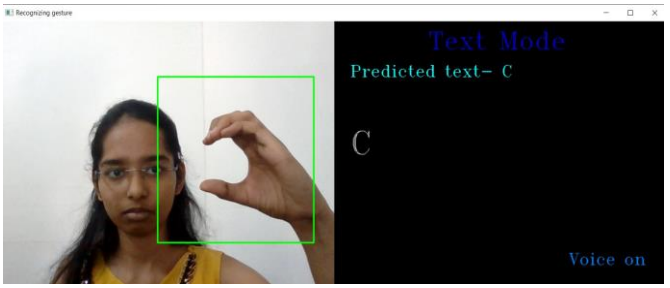
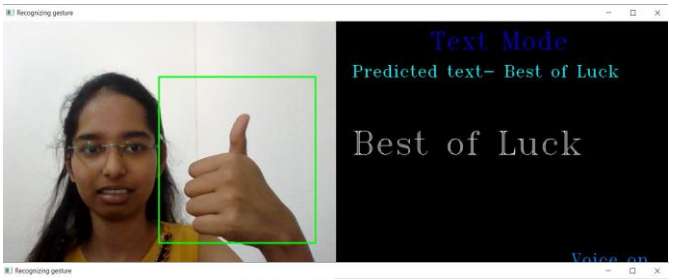
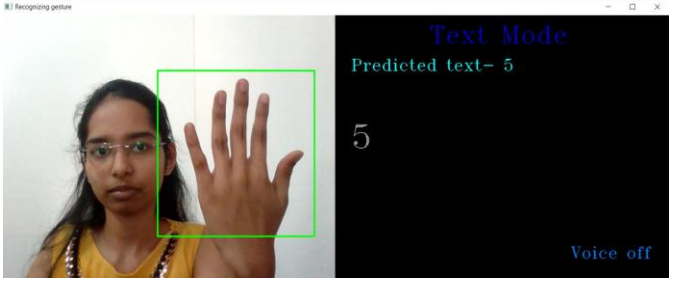


Fig. 5 Hand Gestures

4. RESULT







5. APPLICATIONS AND FUTURE SCOPE

Provides an efficient and accurate way to convert sign language into text.Helps the mute people to communicate with the others as well as amongst themselves.Helps in the process of social inclusion.Enhances the level of confidence among the disabled Makes life easier for the mute people.

FUTURE SCOPE :

This technique might be implemented in elementary schools to teach children sign language at a young age.

5. CONCLUSION

Sign Language Recognition System has been developed from classifying only static signs and alphabets, to the system that can successfully recognize dynamic movements that come in continuous sequences of images. Researchers nowadays are paying more attention to making a large vocabulary for sign language recognition systems. Many researchers are developing their Sign Language Recognition System by using small vocabulary and self made databases. Large database built for Sign Language Recognition System is still not available for some of the countries that are involved in developing Sign Language Recognition System. The classification method of identifying sign language is also varied among researchers. Using their own ideas and limitations for the Sign Language Recognition System, the comparison of one method to another method is still subjective. Fair and direct comparisons between approaches are limited because of the variation of sign language in different countries and the difference in limitations set by each researcher. Variation of sign language in most of the country is based on their grammar and their way to present each word, such as presenting the language by word or by sentence.

6. REFERENCES

[1] Journal article – Sahil Rawal, Dhara Sampat, Priyank Sanghvi. "Real Time Sign Language Detection", "International Research Journal of Engineering and Technology (IRJET)",Volume: 09 Issue: 04 | Apr 2022

[2] Journal article – Iram Haider, Mohammad Ammar Mehdi, Asjad Amin, Kashif Nisar, "A Hand Gesture Recognition based Communication System for Mute people", "IEEE 23rd International Multitopic Conference (INMIC)", June 2021

[3] Journal article –Jinalee Jayeshkumar Raval,Ruchi Gajjar,"Real-time Sign Language Recognition using Computer Vision", "International Conference on Signal Processing and Communication (ICPSC)", July 2021

[4] Journal article -Anupama H S1,Dr.Cauvery N K2,Lingaragu G M3,"Hand Gesture Recognition for Mute People", "International Journal of Advanced Research in Computer and Communication Engineering",Volume:3,Issue 10,October 2020

[5] Garcia, B., & Viesca, S. A. (2016). Real-time American sign language recognition with convolutional neural networks. Convolutional Neural Networks for Visual Recognition, 2, 225-232

[6] Van den Bergh, M., & Van Gool, L. (2011, January). Combining RGB and ToF cameras for real-time 3D hand gesture interaction. In 2011 IEEE workshop on applications of computer vision (WACV) (pp. 66-72). IEEE

[7] Website or Webpage Author's name. Sign Language Recognition Using Python and OpenCV https://dataflair.training/blogs/sign-language recognition-python-opencv/