ACTION VOCALIZER

Narendra Kumar S [1],Vishwas S S[2], Shreeoum S Hebse[3], Samarth N G[4], Jason W George[5]

[1]Professor,[2]Final year Student,[3]Final year Student, [4]Final year Student,[5]Final year Student
Department of Computer Science and Engineering, Jawaharlal Nehru New College of Engineering, Shimoga

---

Abstract - Individuals primarily communicate with one another. Blind and deaf people use signlanguagetocommunicatewithothers.Signlanguagewascreatedasalanguagetohelp the Deaf communicate with one another. Hearing impairment (HI) according to WHO (2018) data 63,000,000 people with significant auditory loss adult-onset deafness is predicted to affect 7.6% of Indians, whereas childhood-onset deafness affects 2%.

As previously said, the number of persons afflicted by hearing impairment and adult-onset deafness rises every year, making Indian sign language a crucial part of initial part of our community. Our study on this issue will assist this demographic in communicating with the broader public and not falling behind in our social growth. The primary goal of this project is to raise awareness of sign language and help deaf and dumb community to communicate with world.

This system translates hand gesture into speech and text. Hand gesture is analysed as part of the identification. The technology provides text output,which helps deaf people, and speech output, which helps blind people and humans communicate more effectively

***Key Words*:**Action Vocalizer, Tenserflow, Mediapipe, image processing, CNN,Keras.

1.INTRODUCTION

The ability to communicate and convey an opinion or a message via any languageby speaking is the biggest boon for a human being. However, some unfortunate people don't have this opportunity. Instead of a language, they use sign language for conveying their message by making certain hand gestures. But the problem is, not everyone can understand the sign language via hand gestures. As a result, communication with such people has always been difficult. In such cases, hand gestures can be recognised by using deep learning and corresponding output can be provided for the identified sign, so that the communication happen smoothly. Hand gesture recognition is one of the most sophisticated areas where computer vision and artificial intelligence have helped improve communication with physically challenged individuals in emergency circumstances, such as pain, calling for help, and soon. Sign language provides a definite way of communication while preserving their uniquegrammatical patterns. Its emphasis the conceptually predetermined movements of the hands, arms, head, and body to considerably build a gesture language.

A sign language recognition system aims to convert sign language into text or voice accurately, efficiently, and in a straightforward manner. This is achieved through digital image processing and various classification algorithms that detect the flow of alphabets and translate sign language words and phrases. Hand gestures are vital for nonverbal communication and have significant relevance in our daily lives. While there is no universal sign language, a limited number of individuals are familiar with it, making it less effective as a means of widespread communication. Sign language recognition helps bridge the communication gap for people with disabilities, enabling them to understand and interact with others without misunderstanding.

2. RELATED WORKS

In the study [1], data acquisition is made using appropriate sensors, cameras, or microphones. The mean filter method is used for noise reduction, image segmentation is done using Threshold method and the image is identified and classified using the CNN and ANN algorithms. The main advantage of this implementation is, it does not require any pricy technology and CNNS can capture very complex details in the hand gestures, as they are expert at extracting hierarchical features from images. But it is difficult to find threshold value and thresholding cannot be applied to a multiple channel image.

The Binarized Neural Network (BNN) boosts the performance by improving the speed of training and reduces the memory size in case of [2]. First the image is converted from RGB to HSV colour space then gaussian blurring is applied to remove noise and for segmentation, Otsu thresholding is used, morphological operations are also applied. Unlike the other studies, here BNN is used instead of CNN, which is faster, more efficient and it reduces computational complexity. But it misclassifies some signs because of similar kind of shapes and limited to small number of classes.

3. METHODOLLOGY

There are mainly 6 steps:

- Capturing of images using camera .
- Pre-processing of images.
- The pre-processed image to CNN.
- After pre-processing with help of Mediapipe,Tenserflow certain operations will be performed.
- Comparing the image with the trained data.
- Finaly then displaying the suitable word.

4. PROPOSED SYTEM

It is important to design an architecture which is not only good at learning features but also is scalable to massive datasets. The Convolutional layers take advantage of inherent properties of images. They use Convolution of image and filters to generate invariant features which are passed on to the next layer. At first the training of the model is done by sending segmented image and the modelgets trained, then

image that has to be classified is taken from web camera which is furtherprocessed. Then the pre-processing is done for the real-time input on whichthe model performs according to the training methods. The output of corresponding sign interms of text is obtained.

Figure 4.1 shows the Architecture of the proposed system, which consists of pre- processing, training, classification of the image and conversion.

Pre-Processing of Images**:** The image taken from the camera is subject to pre-processing This pre – processing step involves Mediapipe, Tenerflow, Cvzone followed by operations.

CNN model: The segmented image is sent to the CNN model for classification. Here the image gets classified based on the trained CNN model, and it outputs a text based on classification of Hand Gesture image. The CNN model that is considered here is MobileNetV2 which is light weight and suitable for low end devices.

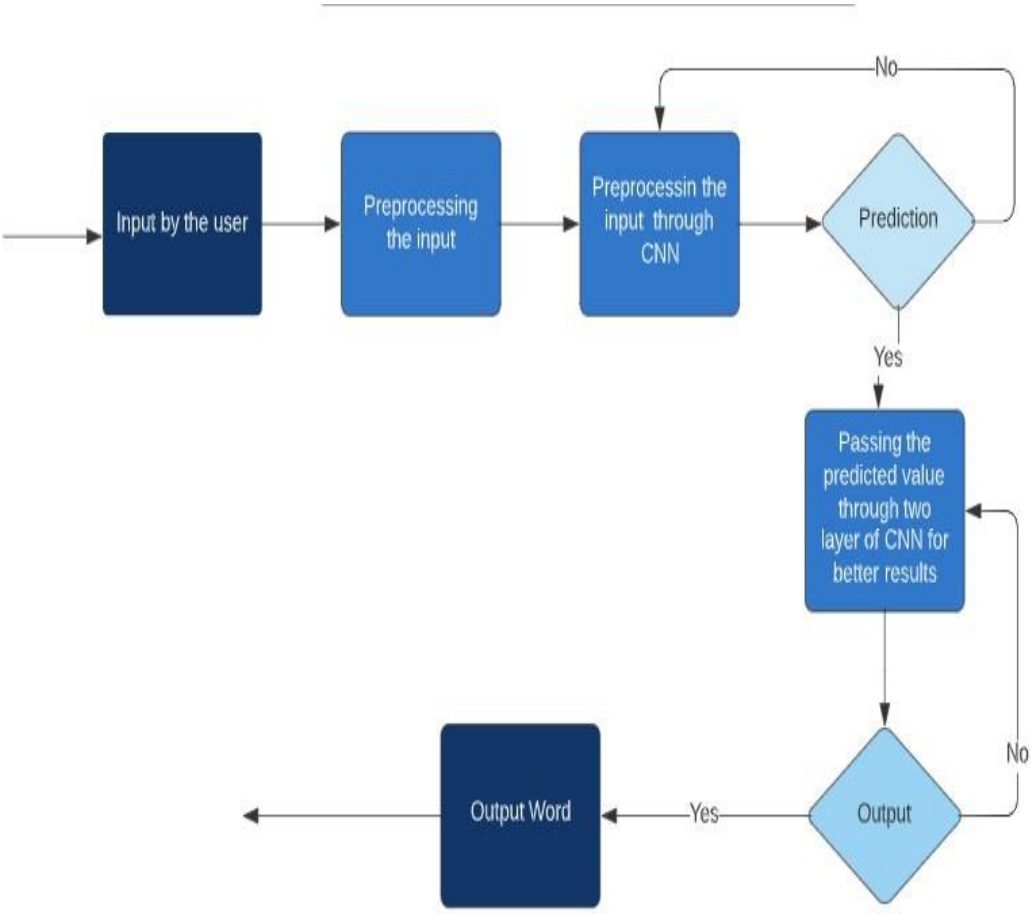

FIG 4.1 SYSTEM ARCHITECTURE

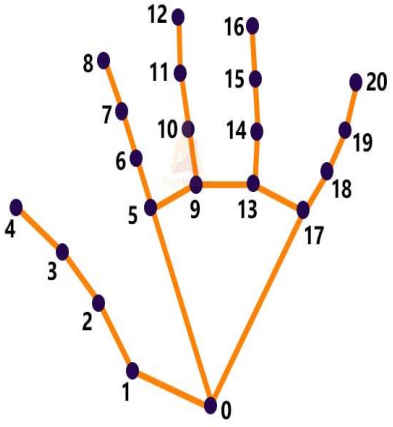

FIG 4.2 key points

These key points will be fed into a pre-trained gesture recognizer network to recognize the hand pose

6. TECHNOLOGY USED

Visual Studio Code is a coding editor that helps to start coding quickly and it is a completely free platform. It supports to code in multiple programming languages, withoutswitching editors. Visual Studio Code has support for many languages, like Python, Java, JavaScript, C++ and many more.

Media pipe is a customizable machine learning solutions framework developed by Google. It is an open-source and cross-platform framework, and it is very lightweight. MediaPipe comes with some pre-trained ML solutions such as face detection, pose estimation, hand recognition, object detection, etc.

TensorFlow is an open-source library for machine learning and deep learning developed by the Google brains team. It can be used across a range of tasks but has a particular focus on deep neural networks.

7. OUTPUT

- **Detection of Sign Language**: This application canrecognize any sign language. Whenever user add the respective dataset in preprocessing step then after data training the respective application will be capable of recognizing it and then it will predict the label accordingly.

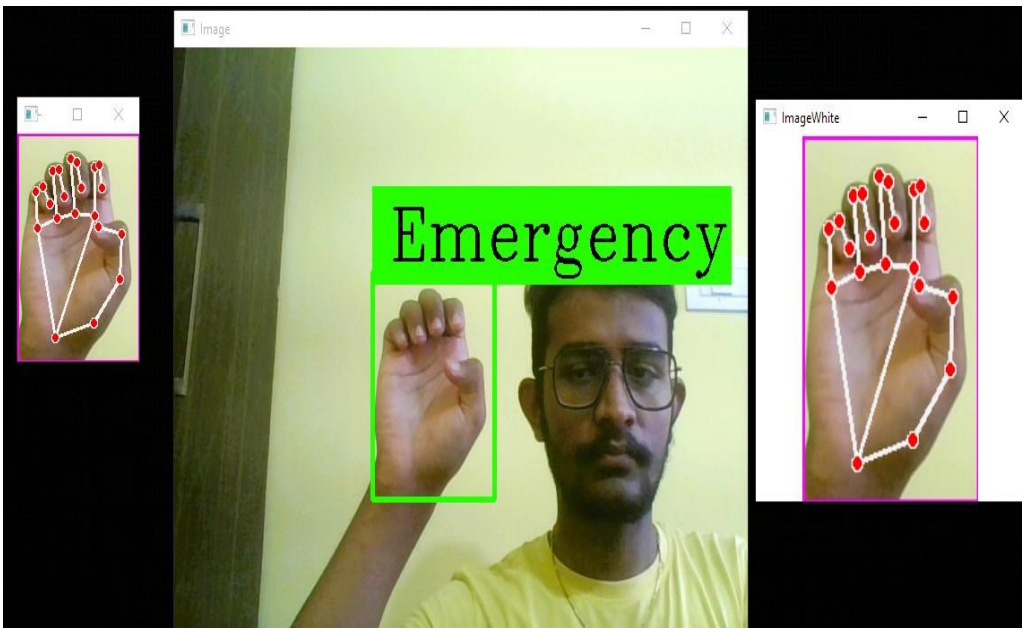

Fig. 7.1 Application detecting Indian Sign language sign 'Emergency'.

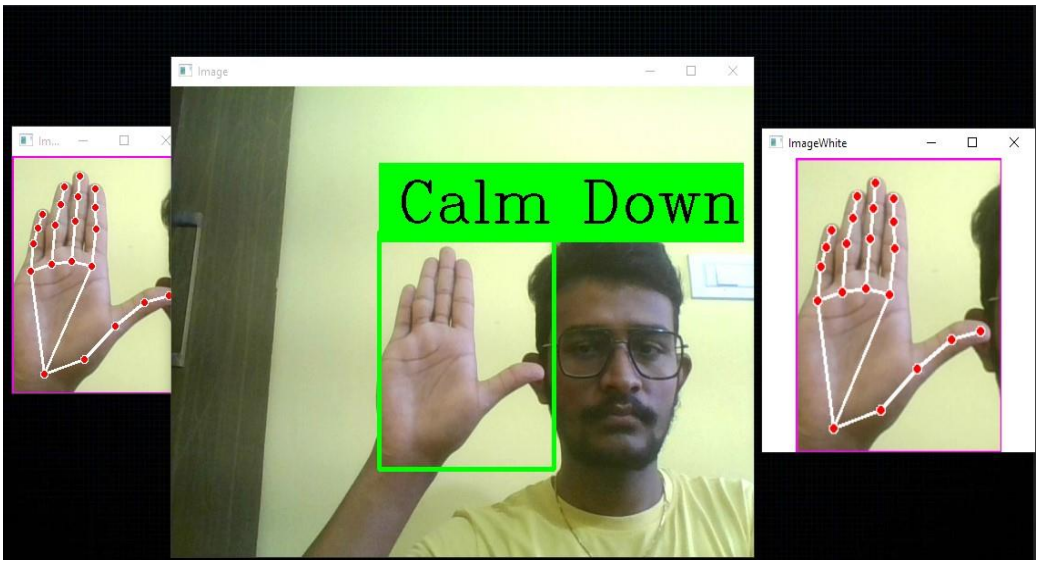

Fig. 7.2 Application detecting Indian Sign language sign 'Calm Down'

To summarize, the results covered in the project are with respect to trained dataset, which is the process of finding optimal values for a learning algorithm and a Standalone application to recognize and translate hand gestures. Special features of application such asdetection of any Sign Language, automatic detection and to increase the processing accuracy a separate white board is created and main palm is captured which shows the advantages of the proposed application over other

8. CONCLUSION

Hand gesture recognition technology is powerful tools that have the potential to bridge the communication gap between the deaf and mute community and the hearing world. The sole purpose of this project is to bridge the communication gap between deaf, dumb and hearing world. The proposed system improvesaccess to education and engages specially abled people in social interactions.

9. FUTURE SCOPE

- The proposed system can be further improved to efficiently work with any kind of background.
- Further advancement can be done by including videos in data set.
- The Proposed system can be capable of having many sign languages and a option to select according to requirement.
- Further advancement can be done by including emergency, greeting words or sentences in dataset.

10. REFRENCES

[1] Surekha P, Niharika Vitta, Pranavi Duggirala, Teja Sree Desani, Venkata Surya Saranya, "Hand Gesture Recognition and Voice, Text Conversion Using CNN andANN", Second International Conference on Artificial Intelligence and Smart Energy, 2022.

[2] Dr. K. Venkatachalam, Dr. V. Purandhar Reddy, M.Amudhan, A.Raguraman, E.Mohan, "An Implementation of K-Means Clustering for Efficient Image Segmentation", 10th IEEE International Conference on Communication Systems and Network Technologies, 2021.

[3] V. J. Dhanawansa, T. P. Rajakaruna, "Sinhala Sign Language InterpreterOptimized for Real-Time Implementation on a Mobile Device", 10th International Conference on Information and Automation for Sustainability ICIAFS, 2021.

[4] Qazi Mohammad Areeb, Mohammad Nadeem, "Deep Learning Based Hand Gesture Recognition for Emergency Situation: A Study on Indian Sign Language", International Conference on Data Analytics for Business Industry (ICDABI),2021.

[5] Pratik Likhar, Neel Kamal Bhagat, Dr. Rathna G N, "Deep Learning Methods For Indian Sign Language Recognition",10th IEEE International Conference on Consumer Electronics (ICCE-Berlin), 2020.

[6] S. Vanaja, R.Preetha, S.Sudhain, "Hand Gesture Recognition For Deaf And Dumb Using Cnn Technique", International Conference on Communication and Electronics Systems(ICCES), 2021.

[7] Mohita Jaiswal, Vaidehi Sharma, Abhishek Sharma, Sandeep Saini, Raghuvir Tomar, "An Efficient Binarized Neural Network for Recognizing Two Hands Indian Sign Language Gestures in Real-time Environment", 17th India Council Conference (INDICON), 2020.

[8] Sadia Sultana, Umme Subrina Jannat Rounok Afza Doha, Mohammad MahadiHassan, Muhammed J.A. Patwary, "Bangla Sign Language Recognition From Hand Gesture Using Convolutional Neural Network", International Conference on Innovation in Science, Engineering and Technology (ICISET), 2022.

[9] Diksha Hatibaruah, Anjan Kumar Talukdar, Kandrapa Kumar Sarma, "A Static Hand Gesture Based Sign Language Recognition System Using Convolutional Neural Networks", IEEE India Council International Conference (INDICON),2020.