

HAND GESTURE RECOGNITION AND TEXT VOICE CONVERSION FOR DEAF AND DUMB

JAWAHARLAL NEHRU NEW COLLEGE OF ENGINEERING

COMPUTER SCIENCE AND ENGINEERING

Students:	Prajwal M Dixit	prajwalmdixit@gmail.com
	Prashanth Patel G P	patelgp.56@gmail.com
	Priya G E	priyagowda302001@gmail.com
	Shubhashree T S	shubhashreets09@gmail.com

Guide:	Narendra Kumar S	narendra@jnnce.ac.in
---------------	------------------	----------------------

Keywords: CNN model, Image, Hand Gesture, Pre-processing.

Introduction

Sign 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 accidents, calling for help, and so on. Hand gesture detection is used for deaf communication as well as to allow gesture-based signaling systems. Hand gesture recognition is a very complicated field that has been utilized to improve deaf communication as well as to allow gesture-based signaling systems. In the subject of sign recognition, deep learning is recently gaining the general implementation of sign language recognition. Pre-processing includes choosing several frames from the video, adding various filters if necessary, and resizing the frames according to the model's input.

Surekha P et. al., in [1] proposed a hand gesture recognition model, where the recognized gesture is further converted to text and voice using CNN and ANN. The methodology followed in [1] can be summarized as follows: 1. Data acquisition. 2. The pre-processing steps are applied to the raw data, such as noise removal or normalization. 3. The training process. 4. The data augmentation techniques are used to improve the model's robustness. 5. Noise reduction is done using mean filter and image segmentation is done using threshold method. The proposal [1] has a limitation, to boost user engagement and make the system more robust, the app may be linked with other mobile and IoT devices.

Objectives

- To convert hand motion into voice and text: This project will be able to provide a separate device for effective segmentation of hand picture from the image captured and transform the hand gesture that have been identified into text and speech using effective convolution neural network.
- To provide freedom for hand motion: To provide freedom for orientation of hand movement (Hand can be placed anywhere within the frame).
- To provide efficient segmentation: To segment hand orientation efficiently.

- To display specific text: To display the text on the screen and fed into Google Text To Speech to provide an audio output.

Methodology

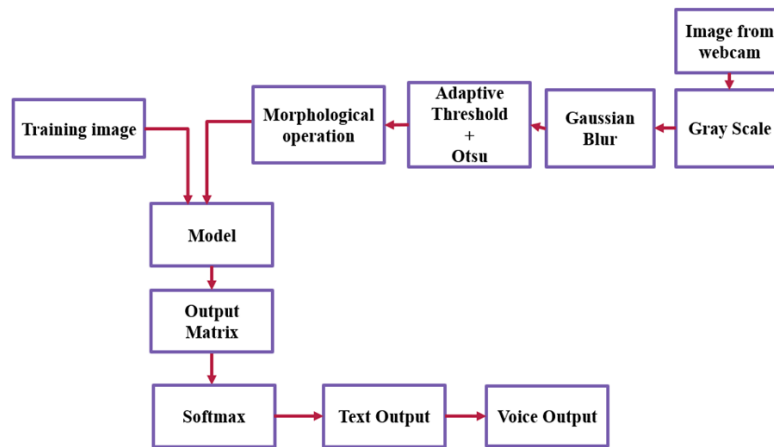


Fig.1 System Architecture

The methodology of the proposed hand gesture recognition system is shown in Fig. 1. The proposed technique of hand gesture recognition scheme follows different phases, namely pre-processing of image, transfer of Pre-processed image to CNN model and text to voice conversion. At first the training of the model is done by sending segmented image and the model gets trained, then image that has to be classified is taken from web camera which is further connected to Raspberry Pi.

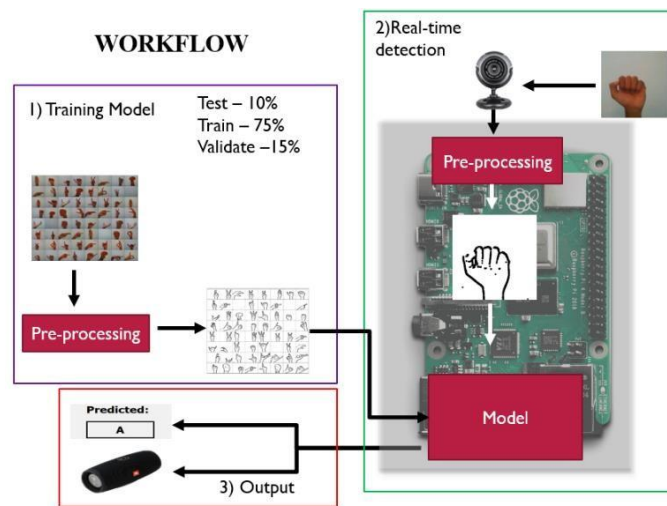


Fig.2 Workflow

Pre-Processing of Images: The image taken from the camera is subject to pre-processing. This pre – processing step involves Grayscale, Gaussian blur, Adaptive Threshold with Otsu threshold followed by Morphological 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.

Text to Voice Conversion: The classified text is converted to voice.

Results and Conclusions

A Standalone application that can recognize hand gestures in real-time and translate them into text and speech is developed to make proposed hand gesture recognition system more accessible. Figure 3 shows a snapshot of the applications user interface.

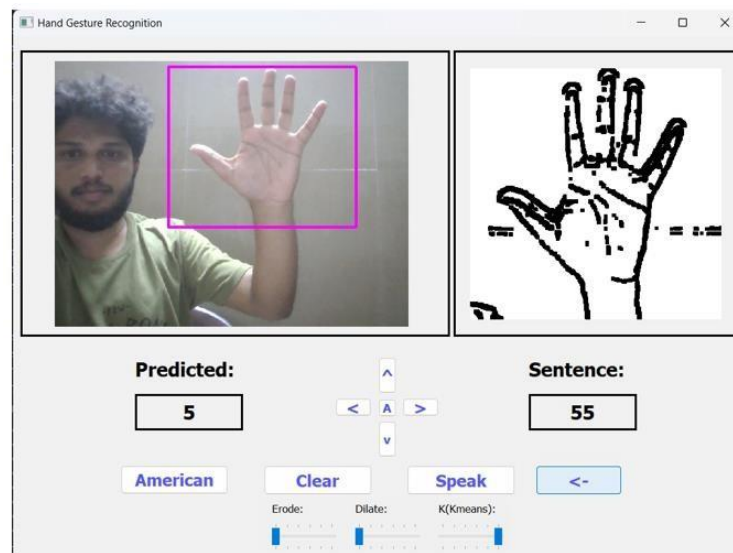


Fig.3 Standalone Desktop Hand Gesture Recognition Application.

Hand gesture recognition and text-to-voice conversion technologies are 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 improves access to education and engages specially abled people in social interactions. Through this project both American and Indian sign language can be translated. As MobileNetV2 is a light weight model specifically designed for mobile devices, it can be deployed in low end devices.

Future Scope

- The proposed system can be further improved to efficiently work with any kind of background.
- Further advancement can be done by including emergency, greeting words or sentences in dataset.

Reference

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