iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VI **Month of publication:** June 2022

DOI: https://doi.org/10.22214/ijraset.2022.44951

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

Hand Gesture Recognizer with Alert Message Display and Vocalizer for People with Disabilities Using Deep Learning

Dr G. C. Bhanuprakash[1], Meghana Dinesh[2], Ishaka Fatnani[3], Nevethitha Prasad[4], Mansi M[5]
[1]*Professor And Head, Department Of Computer Science & Engineering, Sir M Visvesvaraya Institute Of Technology, Bengaluru, Karnataka, India.*
[2, 3, 4, 5]*Student, Department Of Computer Science & Engineering, Sir M Visvesvaraya Institute Of Technology, Bengaluru, Karnataka, India.*

Abstract: ***Communication is the basic medium by which we can interact and understand each other but unfortunately this process can prove to be difficult for the people with disabilities. This project is a small contribution to address this issue and help them communicate better. It uses the concept of neural networks in order to help the people with disabilities communicate better with just their hand gestures. The hand gesture recognizer and actuator system acts as an interface between computer and human using hand gestures. This project is a combination of gesture identification, classification and it displays the message pertaining to the gesture while simultaneously reading the message out loud. The user has to perform a particular gesture, the webcam captures this and identifies and recognizes it against a set of known gestures and displays and vocalizes the gesture. This is a very useful hands-free approach. There could be an emergency situation for them and in order to inform people around them, this system is an effective tool. The system has a set of samples to be trained against and then the trained model is tested against a set of test data.***

Keywords: Hand Gesture Detection, Neural Networks, Deep Learning, Alert, People with Disabilities.

I. INTRODUCTION

People with disabilities are an integral part of our society. With the advent of science and technology, efforts are being made to develop certain systems that allow accessibility. Hand gestures or sign languages have always been used by people with hearing and speech difficulties to communicate but it is not always easy to find someone who understands sign language and can translate it. Human-computer interaction system can be installed and used anywhere possible. Gestures are basically the physical action performed by a person to convey a message. Hand gesture is a technique of non-verbal communication for human beings.

Object detection is a computer technology related to computer vision and image processing that deals with detecting instances of semantic objects of a certain class (such as humans, buildings, or cars) in digital images and videos. Well-researched domains of object detection include face detection and Pedestrian Detection. This technology has applications in many areas of computer vision, including image retrieval and video surveillance. In recent times a lot of papers have been published that make use of CNN to build a hand gesture recognition system but there are a very few of them connecting the recognized hand gesture to perform useful tasks. Our motive through this project was not only to build a hand gesture recognizer but also an actuator to perform required tasks such as displaying of an alert message dedicated to that particular hand gesture and also vocalizing the message in order to grasp the attention of the people around them.

This project is a small contribution to addressing this issue and assisting people with disabilities in communicating more effectively. It employs the concept of neural networks to assist them in communicating short messages using only hand gestures. An interface between the computer and the user is provided by the hand gesture recognizer and actuator system. This project combines gesture recognition with message classification; based on the gesture, it displays and vocalises the appropriate message. The user must make a specific gesture that the webcam records, recognises against a database of recognised gestures, and then displays the message that goes with it. This hands-free technique is incredibly practical. A set of training samples are sent to the system, and the trained model is then tested using a set of test data. Once a suitable accuracy is achieved, the final step is making use of TensorFlow to load the model. Media Pipe is then used to recognize the gesture and feed it to the model. The webcam is then activated wherein the real time video is captured and the message corresponding to the gesture is displayed as well as vocalized.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

II. LITERATURE SURVEY

Debashish Bal et al. in his paper includes a hand glove system for the deaf and mute patients which uses their hand gesture pattern for any communication. The hardware used by the authors to design the gloves includes flex sensor, to measure the quantity of bending or deflection. The MPU 6050 records acceleration, orientation, and velocity using a 3-axis gyrometer and 3-axis accelerometer. Flex sensors are connected to MPU 6050 with Raspberry Pi 3D. Several dynamic hand movement recordings are obtained 20 times from 88 individuals to provide a total of 1760 samples for the purpose of identifying gestures. Finally, the model is evaluated using 23 physically challenged individuals, each of whom performs the hand gesture 20 times, with an overall accuracy rate of 92.7 percent. Although this glove system has three flex sensors, five flex sensors can be employed to get superior performance. [1]

Zehao Wang et al. proposes in their paper a block sparse based time-frequency (TF) feature extraction method for recognizing dynamic hand gestures using millimetre-wave radar sensors. The author uses the idea that the micro-Doppler features in hand-gestures are more clustered than independently scattered points. This radar-based hand gesture recognition consists of two steps, namely data representation and feature extraction, and classification. There are three methods used, they are, Block Sparse Model is applied to model the TF distribution of gestures, then extraction of TF features is done by block orthogonal matching pursuit (BOMP) algorithm, and finally kNN classifier is used to classify the extracted block features. The dataset consists of 400 hand gestures, each subject repeating a specific gesture for 25 times. The highest accuracy is 89.5% in the classification stage. [2]

Bhargav Jethwa et al. in his paper include the communication between hearing and speech impaired people and non-disabled people. Communication is carried out through the mobile phone of the specially abled person in the form of audio and text notifications. The authors have made use of IMU sensors to record temporal signals. These signals are transmitted to the mobile phone of the disables using Bluetooth 2.0. Here, the signals are pre-processed to remove the noise using Kalman filter. The pre-processed data is used for training in the 1D CNN model. The output of this proposed system is in the form of audio and text of the identified gesture in the di-abled mobile.

The training accuracy achieved is 98.06% and the testing accuracy is 97.96%. The time taken by the model to predict the pattern and display result once the data is reached into the phone is 0.56sec. [3]

Mohd. Aquib Ansari et al. in his paper use a robust approach of dynamic hand gesture recognition for human machine interaction (HMI) systems. Web cam is used to capture images in real time for the working of the system. In RGB and HSV colour space, skin colour modelling is done for segmenting geometrical approximation of the hand using contours. The hand may be found in the image using ROI.

Tracking fingers is done using contour defects, while tracking hands is done using centroid tracking. This method employs eight different types of dynamic hand movements, and the system has effectively recognised each one of them. There are two interface technique types, glove based, and vision based. The methodology proposed was by using Median filter, Skin Colour Segmentation, Morphological Operation which then sends pre-processed images for detection of hand. The accuracy is up to 95% for dynamic hand gesture recognition. [4]

Dr. P. Muralidhar et al. proposes a real time system in the paper [5] can be customized to user-specific dynamic hand movements and can carry out certain tasks, especially for the motor impaired people. The paper makes use of Cascade detection algorithm in order to track the movements of the hand and to trace the path and it also uses Siamese neural network for the purpose of customization and recognition of the gestures. When operating the system for the first time, the user has to record the desired number of gestures for performing various operations respectively. Once the gestures are recorded successfully, the user can perform any of those gestures and the matching of the images takes place between the already stored images and the image of the traced trajectory of the current gesture. The gesture which gives the highest probability on matching is then performed.

M. I. N. P. Munasinghe designed and developed a system in his paper [6] that makes use of Motion History Images (MHI) and Feedforward Neural Networks for recognizing gestures in front of a web camera real time. Firstly, Gaussian mixture-based background/forehead segmentation algorithm and then median filtering is applied, then binary thresholding with Otsu's binarization and then the processed frames are merged and using a developed algorithm based on structural similarity to generate cumulative motion history image. To classify the gestures the author has used a stochastic gradient-based optimizer. A Neural Network is used to determine the probabilities for each category, if the category having the highest probability surpasses 0.8, it can be identified as the correct gesture. The experiment results in real time using motion history images with the neural networks is shown to give considerable satisfactory results.

©IJRASET: All Rights are Reserved

III. METHODOLOGY

Our first step is to capture the real-time video though the web camera. Tis real-time video is fed to the model after being put through a few pre-processing steps to ensure the model has a clear image of the gesture for classification or detection.

A. Pre-Processing Steps: Some of the pre-processing steps we are performing are:

1) Resizing: We are capturing images and later we convert it to the standard dimensions so that all the images are of the same dimension making the processing easier to carry forward.

2) Filtering: Filtering is a process we commonly use for the purpose of image enhancement. By enhancing the image, we are actually giving more significance to the parts of the image that are of our concern. The unwanted noise introduced to the image due to poor illumination and other such factors is removed in this process.

3) Segmentation: It is the process in which we divide an entire image into smaller parts or partitions so that the overall complexity of processing the image reduces and the various insights to be drawn out becomes easier.

4) Edge Detection: It is the main process among other pre-processing steps. Here we will be able to detect the edges in the image with the help of some filters like Sobel, Laplacian, and many other sharpening filters.

5) Feature Extraction: The pre-processed image is then used for the process of feature extraction. In this step, the dimensionality of the image is being reduced. Features are like properties of an image that help us to identify it. So here, the various required features are obtained from the pre-processed image and extracted to form a labelled data which can be used for classification.

6) Classification: It is a process in which the labels will be assigned to the image based on the features extracted in the previous step. The number of classes present in the classification purely depends upon the application of the project.

7) Gesture Recognition: In this process the images are classified into respective classes defined in the previous step by taking into account the features extracted from the image. The gesture recognised by this step can further be used to perform the predefined set of actions associated with respect to the gesture performed. All these processes are used for training a set of training examples and later it is used to work against a set of test examples to get the desired output. It is trained in a convolution neural network (CNN) which is a deep learning method.

8) Actuator: Once the gesture is recognized by the system, it gives an alert signal in the form of audio which can alert the care takers to take appropriate action.

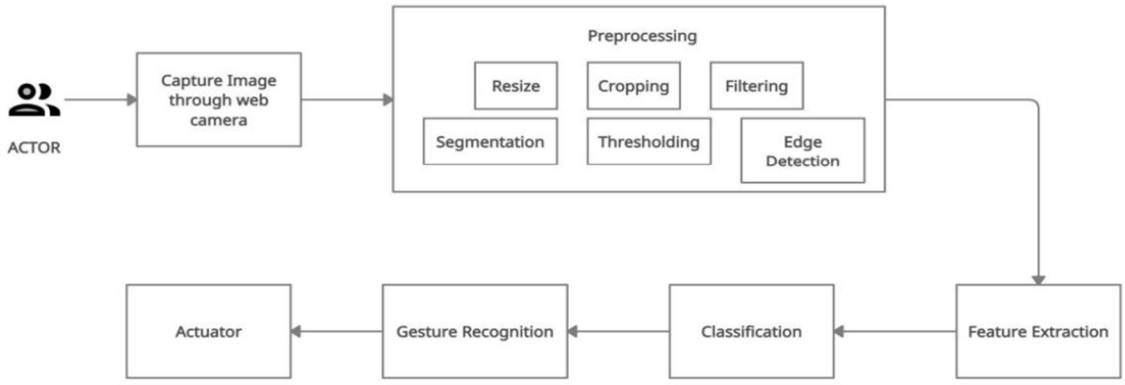


Fig. 1 System Architecture

A high-fidelity hand and finger tracking solution is MediaPipe Hands. It uses machine learning (ML) to extrapolate 21 3D hand landmarks from a single frame. Our solution delivers real-time performance on a cell phone, and even scales to several hands, unlike existing state-of-the-art systems, which mostly rely on powerful desktop environments for inference. MediaPipe Hands uses an ML pipeline made up of several interconnected models: an orientated hand bounding box is produced by a palm detection model that uses the entire image as input. a hand landmark model that runs on the palm detector-cropped picture region and outputs highly accurate 3D hand keypoints.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

IV. RESULTS

The image is being detected with a 99% accuracy. The lag caused by the load on the system is the only drawback. The model detects 10 gestures and can be trained to add more custom gestures to the model. The self-training ability makes it more user friendly.

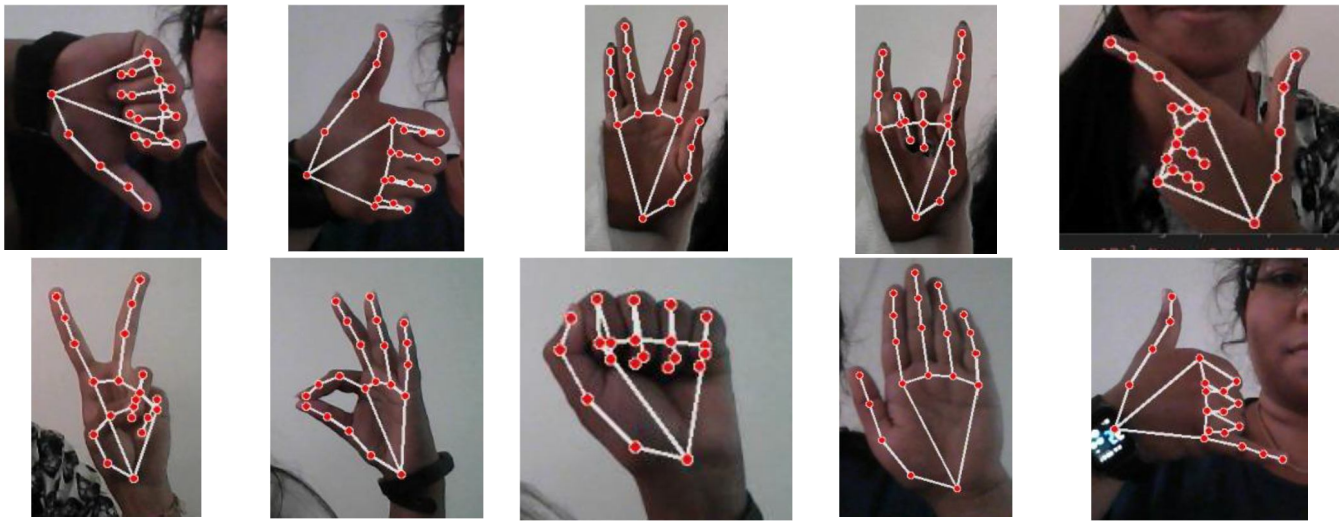


Fig. 2 Gestures Recognized by the System

V. CONCLUSION

Gesture Recognition provides the most important means for non-verbal interaction among the people especially for people with hearing and speech impairments. Different applications of hand gesture recognition have been implemented in different domains. This project is outlines a hand gesture recognition system that is designed to be capable of recognizing hand gestures in real time. The aim is to compare different algorithms in training and testing approaches to discover the best algorithm to extract and classify hand gesture recognition. The ability of neural nets to generalize makes them natural for gesture recognition. Therefore, the video captured in real time is used to recognize the customized gesture for the alert message. The audio feature of the system allows people around to be alerted as to what message is being conveyed by the gesture. The future of such a software is to reduce the lag caused due to the addition of the voiceover. This would make the system more pleasing to the user as well as increase the efficiency. A bigger development would be for the software to send alert messages to emergency services if necessary. This feature, if added, could be lifesaving in many circumstances.

REFERENCES

[1] Debasish Bal et al. “Dynamic Hand Gesture Pattern Recognition Using Probabilistic Neural Network”, 021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS) | 978-1-6654-4067-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/IEMTRONICS52119.2021.9422496

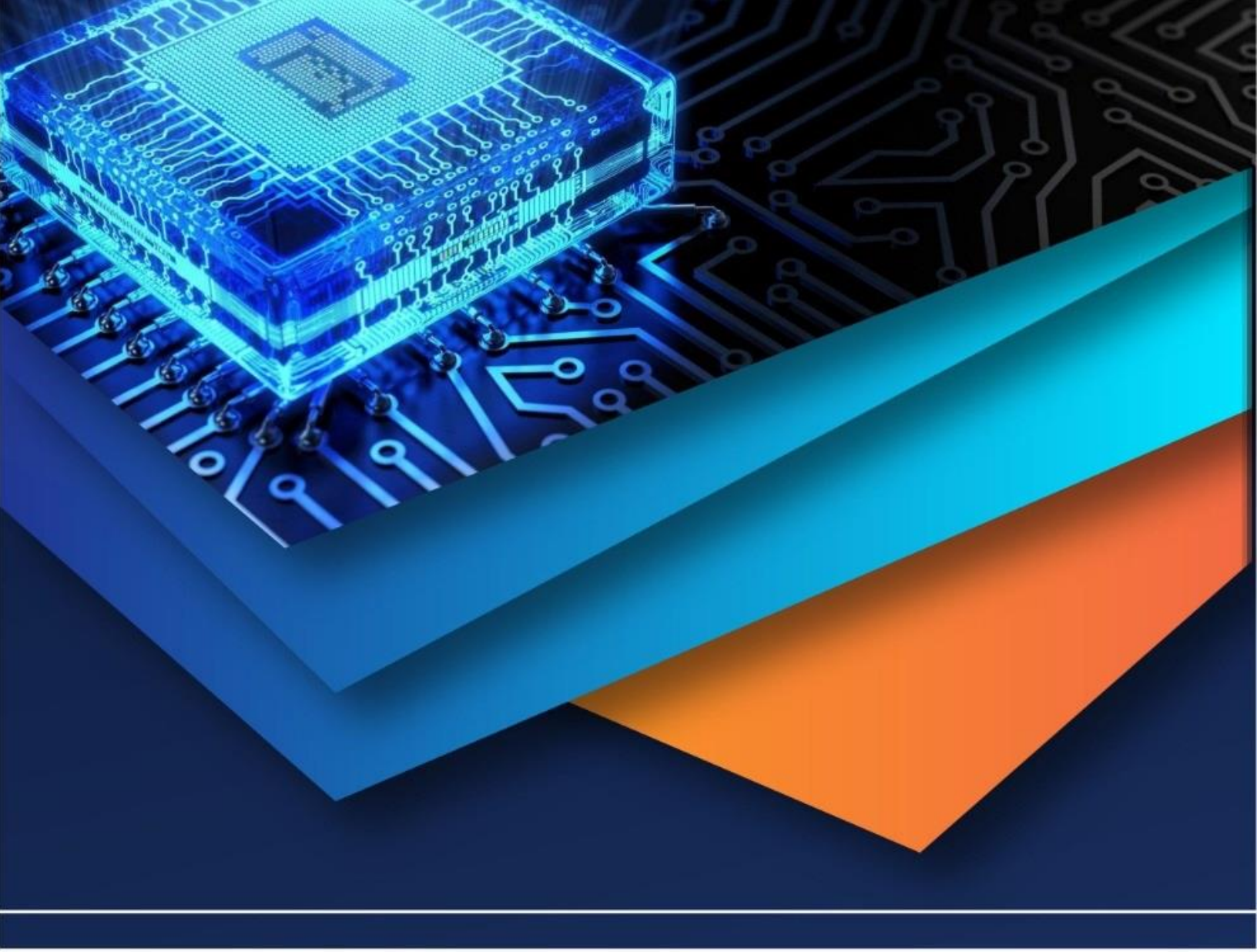
[2] Zehao Wang et al. “Feature Extraction for Dynamic Hand Gesture Recognition Using Block Sparsity Model”; 2021 IEEE/MTT-S International Microwave Symposium - IMS 2021 | 978-1-6654-0307-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/IMS19712.2021.9574796.

[3] Dr. P. Muralidhar et al. “Customizable Dynamic Hand Gesture Recognition System for Motor Impaired people using Siamese neural network”; ISBN: 978-1-5386-8448-1/19/$31.00 ©2019 IEEE.

[4] Abhi Zanzarukiya et al. “Assistive Hand Gesture Glove for Hearing and Speech Impaired”, Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020) IEEE Xplore Part Number: CFP20J32-ART; ISBN: 978-1-7281-5518-0.

[5] Mohd. Aquib Ansari et al” An Approach for Human Machine Interaction using Dynamic Hand Gesture Recognition”; 2019 IEEE Conference on Information and Communication Technology (CICT); 978-1-7281-5398-8/19/$31.00 ©2019 IEEE.

[6] M. I. N. P. Munasinghe “Dynamic Hand Gesture Recognition Using Computer Vision and Neural Networks”; 2018 3rd International Conference for Convergence in Technology (I2CT); 978-1-5386-4273-3/18/$31.00 ©2018 IEEE.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)