ISSN 2063-5346 Section A-Research paper

SMART FRAMEWORK FOR VISUAL COMMUNICATION TO THE PEOPLE WITH DISABILITY

Prof *Mallikarjuna* Rao G[a]*, Nandoori Vasavi[b], Salunkhe Preethi[c], Kaviti Lavanya[d], Jampani Divya[e]

[a] Computer Science Engineering,
Gokaraju Rangaraju Institute of Engineering and Technology,
Hyderabad, India
gmallikarjuna628@grietcollege.com

[b] Computer Science Engineering,
Gokaraju Rangaraju Institute of Engineering and Technology
Hyderabad, India
vasavi.nandoori10@gmail.com

[c] Computer Science Engineering,
Gokaraju Rangaraju Institute of Engineering and Technology
Hyderabad, India
preethisalunkhe@gmail.com

[d] Computer Science Engineering,
Gokaraju Rangaraju Institute of Engineering and Technology
Hyderabad, India
lavanyakaviti93900@gmail.com

[e] Computer Science Engineering,
Gokaraju Rangaraju Institute of Engineering and Technology
Hyderabad, India
jampanidivya7@gmail.com

Abstract:

Effective communication facilitates the sharing of emotions, expressions, and thoughts among individuals, thereby fostering stronger interpersonal connections. There are around 5% of the population of the world who suffer Deaf or hearing-impaired. Complex arm movements with articulated hand forms and facial expressions are used for this purpose. However, the scope for them to communicate with normal people is limited. Our objective is to leverage technology to address this communication gap. To achieve this goal, we present a sophisticated framework in this paper that integrates sign language with cutting-edge machine learning and deep learning techniques. By utilizing the capabilities of Media pipe and TensorFlow with Python, our framework's robustness is significantly improved.Custom made data set is hybridized with American sign language data set during the experimentation. Experimentation results are encouraging with an accuracy of 93%. Our approach holds significant promise in facilitating communication for those with hearing impairment and can lead to a more inclusive society.

Keywords *:* Machine learning, Convolution Neural Networks, Hand Gestures, Arm movements.

INTRODUCTION

Individuals with hearing-impaired and voiceless mostly communicate with others by using hand and body motions in Sign Language (SL), both within their own community and with other people.

Their syntax,vocabulary, and meaning are all unique from both spoken and written language. In spoken language, certain words and grammatical arrangements are mapped to articulating sounds in order to express sense. Visual hand and body gestures are employed in sign language to communicate important ideas. These hand motions can be categorized, and associated text can be produced, using Deep Learning algorithms and image processing.Lingchen Chen et.al [1], in their survey paper discussed various approaches for hand gestures in sign language along with human computer interface.

Sign language makes extensive use of gestures to mimic movement language, which is composed of a series of hand and arm motions. There are various hand gestures and sign languages used in various nations. According to some sources, certain words that have not been translated can be understood by employing the corresponding hand gestures for each alphabet. Moreover, sign language has unique motions for each letter of the English alphabet and every number from 0 to 9.

Based on these, two categories of gesture—static gesture and dynamic gesture—are used to classify sign languages.

The alphabet and numbers are represented using a static gesture, however some concepts are expressed using a dynamic gesture. Words, sentences, etc. are all dynamic. While the latter encompasses the motion of the hands, the head, or both, the former is limited to hand motions. Three main elements make up sign language, which is a visual language, including finger-spelling, a vocabulary of signs at

ISSN 2063-5346

the word level, and non-manual aspects. The latter, on the other hand, transmits information by using keywords and finger spelling, which spells words letter by letter.

Despite several research efforts over the past few decades, designing a sign language translator is still fairly difficult. Even identical signs can appear very differently depending on the signer and the angle. Using the help of a convolutional neural network, this work focuses on developing a static sign language translator. In order to deal with embedded devices, standalone applications, and online apps that have fewer resources, we developed a lightweight network.

LITERATURE REVIEW

Initially, the hand gesture recognition system used a sensor glove to identify sign language gestures by detecting the bending of the fingers through various sensors, such as curvature sensors, flex sensors, and angular displacement sensors. However, this method had limitations as individuals with sensitive skin could experience harmful effects, and wired connections caused discomfort with extended use. Later, a color-based recognition system was developed that used a glove with color blemishes to locate palm movements and make it easier for the camera sensor to trace the geometric shape of the hand. However, the use of gloves still restricted the level of spontaneous interaction with the system. Another system that used motion-based hand recognition with the AdaBoost algorithm was also developed. This system used picture data frames to identify hand gestures, but efficiency and precision were reduced when multiple gestures were active, and a dynamic background had an adverse effect. Additionally, a depth-based recognition system was proposed, which used a depth camera to produce 3D geometric data of the hand. While this system proved to be immune to the effects of skin color, lighting, and background cast, its limited accessibility, cost, and spatial requirements hindered its practicality.

N.Mohamed and colleagues conducted a comprehensive analysis of the hand gestures that are linked to sign language [2]. Based on which many sign level recognition systems were made.

In their study, Davinder Kumar et al. conducted a thorough assessment of the anatomical and physiological foundation that serves as the basis for hand gesture recognition [3]. I. Dhall and his co-authors explored the automation of hand gestures[4], while S. Nitnaware and his team utilized the multiprocessing library to create an entertainment framework that facilitates visual communication for individuals who are both deaf and mute[5]. S. Shanmugam et. al[6] employed convolution neural networks for developing hand gesture recognition, and J. -H. Sun and his team[7] proposed the use of the CamShift algorithm with an AdaBoost classifier for real-time hand gesture recognition. Al Farid F[8] suggested sub-categories for gesture-based recognition and summarized the associated methodologies. Additionally, G.M. Rao and his team[9][10] investigated computer vision-based techniques for detecting partial facial features and determining liveliness, particularly through the analysis of eye blinking. Pramod Kumar Pisharady et al. [11] presented a comprehensive assessment of gesture recognition algorithms, incorporating both qualitative and quantitative approaches, while Ming Jin Cheok et al. [12] offered a comprehensive framework for the gesture recognition process. Similar methods using deep learning methods have been employed by M.Rao et. al [13][14] for black fungus detection and handwritten digit analysis.

PROPOSED METHOD

The machine learning framework is developed to predict hand movements from a video using Convolution Neural Network in the suggested method. The model implements a pattern classification method to recognize the gestures provided from a web camera and perform actions corresponding to the output.

The following figure illustrates the architectural framework
.

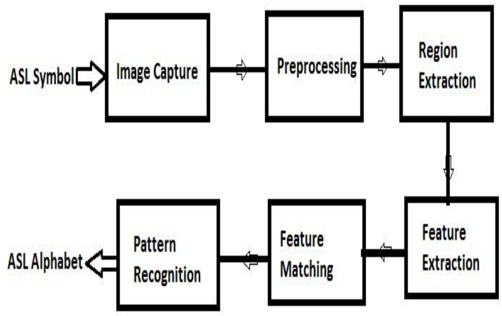

Figure 1: Architectural framework

Import required packages

For the model to be implemented, the following packages must be loaded in the Python platform (Framework is developed using Python 3.8.8).

- OpenCV – 4.5

Open-Source Computer Vision Library is referred to as OpenCV. It falls within the category of machine learning software.. It is used for the acceleration of computer vision streaming.

- Media Pipe – 0.8.5

Google created Media Pipe, a flexible platform for the machine learning solutions. The framework is incredibly lightweight, cross-platform, and open-source. Media Pipe includes a number of machine learning (ML) per-trained solutions, including face and object detection, pose estimation, hand and object recognition, and more. To identify the hand and its major features, we'll first use Media Pipe. For every detected hand, Media Pipe returns a total of 21 crucial points.

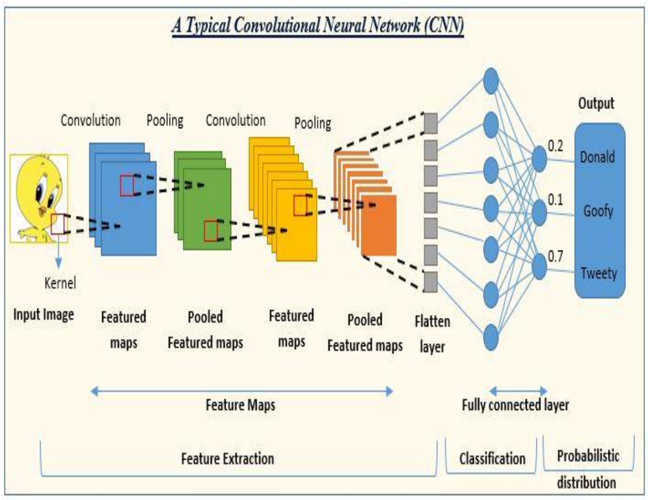

Figure 2. Model of Convolution Neural Networks

● Tensorflow – 2.5.0

The Google Brains team created TensorFlow, an open-source library for machine learning and deep learning. Deep neural networks are given special consideration, despite the fact that they are suitable to several applications. The term "artificial neural networks" also applies to neural networks. It is a part of artificial intelligence and the core of deep learning algorithms. The human brain served as an inspiration for the development of neural networks. It resembles the communication between neurons in living things. Neural networks are made up of node layers, which are composed of layers such as an input layer, one or more hidden layers, and an output layer.

● Numpy – 1.19.3

It is a library under the Python language, that allows users to access data types like lists, dictionaries, tuples, and many other mathematical functions employed in the construction of this system.

Reading frames from Webcam

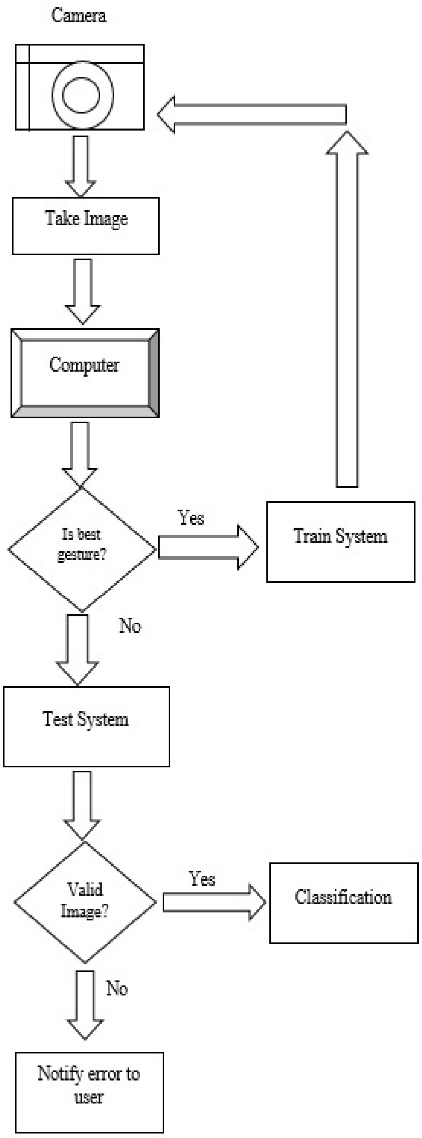

Figure 3. The flowchart of the methodology used.

Mathematically, this procedure is described as :$Z = X * f$
Where X represents an input image and f represents the filter used.Ifthe image's dimensions are (n, n)

and the filter's dimensions are (f, f), then output's dimensions will be ((n-f+1), (n-f+1)).

Image Acquisition:

The first step involved capturing the camera's frame and determining the relevant area within the image. Due to the need for accurate real-time translation, a powerful computer vision technique capable of detecting even minor variations between similar data was required. OpenCV offers over 150 different color-space conversion methods, but only two were considered in this instance: BGR to Grey and BGR to HSV. The HSV format was chosen for this study as it facilitates easier extraction of the gesture compared to other formats, following the conversion of the colored RGB format.

Segmentation:

Segmentation is another crucial element of the suggested process. In the simple implementation, the process exhibits a binary object that illustrates the segmentation. The background is comprised of black pixels, while the foreground is represented by white pixels.. In simple systems, a single parameter known as the intensity threshold controls segmentation.

Classification:

The mask images obtained are trained in Keras, a Python-based high-level API that can be implemented over TensorFlow, through a convolution neural network named Lenet. An example of a convolution neural network (CNN, or ConvNet) are deep neural networks used to analyse the visual representation process. CNN's use a multilayer perceptual variation that is intended to need the least amount of prepossessing.

The Tensor Flow-based Keras system is used to build the neural network after optimization procedures, with the following hyper parameter values: This could seem intimidating to someone who is unfamiliar with deep learning. The most difficult information is concealed behind the scenes using KERAS' simple and adaptable Network Training API.

Identify key-points/gestures of the hand-gesture

Using the pre collected data stored in the database from the system, the live video stream classifies the given hand gesture and identifies with the gesture which is already stored in the database based on the similarities among them. The pre collected data has numerous data pictures for each gesture to detect the live gesture more accurately under every circumstance. To be briefer, it uses Supervised Machine learning technique for the identification of the gesture.

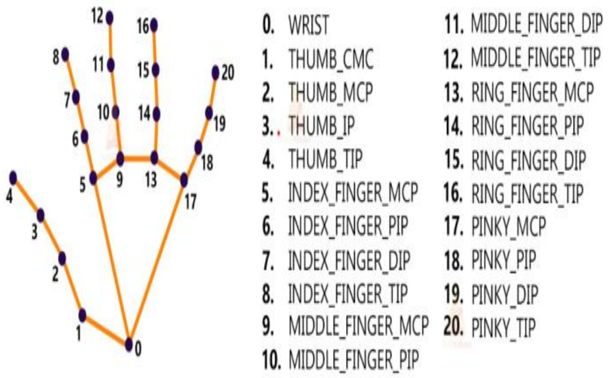

Figure 4. The model of hand used in sign recognition

We used the standard American sign-level language as the base for our sign level recognition.

Figure 5. The Standard American sign-level language

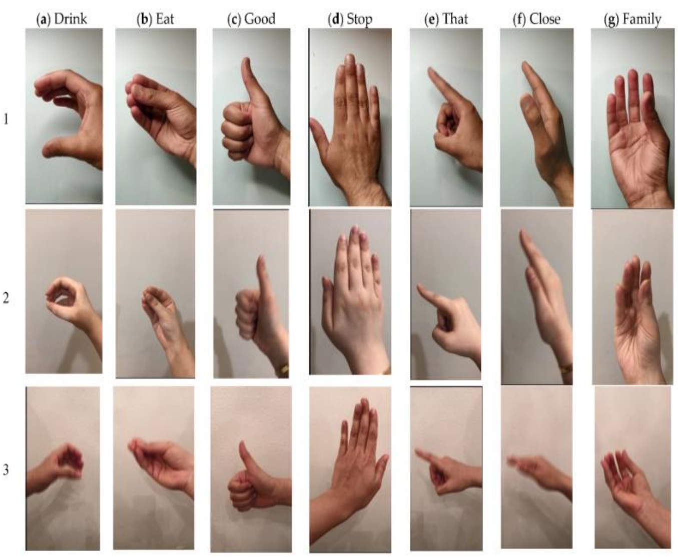

Figure 6. Actions in Standard American sign-level language

RESULTS

The images below show a live demonstration of our model's ability to recognise English alphabets.

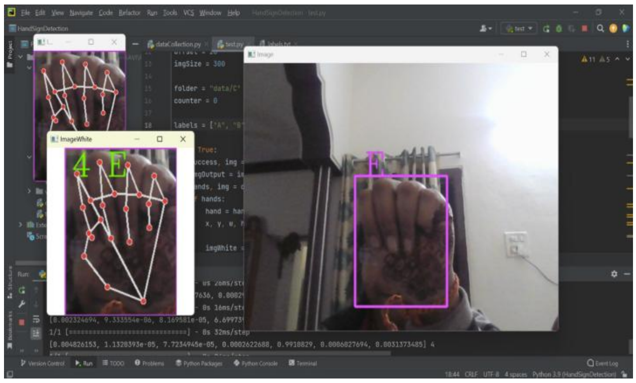

Figure 7. The Model Recognizing the hand gesture as "F"

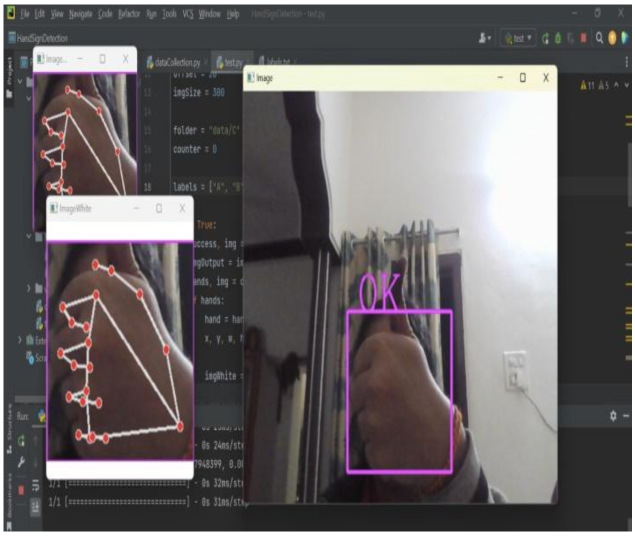

Figure 8. The Model recognizing the hand gesture as "Ok".

CONCLUSION

In a nutshell, this approach has been successful in a variety of project objectives, including:

1.Designing a sophisticated system that leverages computer vision, Python, and OpenCV to identify, categorize, and interpret hand gestures with high accuracy and efficiency.

2.Creating a comprehensive library of hand gestures, including both numerical and sign language variations, that aligns with the project's overarching objectives.

SCOPE

This method for recognizing hand gestures effectively addresses the challenge of efficiently processing and extracting video frames.This technology enables the identification and potential utilization of diverse hand gestures as computer inputs. Additionally, it has the capacity to convert

ISSN 2063-5346

hand gestures that represent numbers into prompt commands to efficiently execute required actions.Future studies may focus on more reliable algorithms that can accurately handle dynamic images while also being used for hand gesture identification.

REFERENCES

[1] B. Nandwana, S. Tazi, S. Trivedi, D. Kumar and S. K. Vipparthi, "A survey paper on hand gesture recognition," 2017 7th International Conference on Communication Systems and Network Technologies (CSNT), Nagpur, India, 2017, pp. 147-152, doi: 10.1109/CSNT.2017.8418527. (2017)

[2] N. Mohamed, M. B. Mustafa and N. Jomhari, "A Review of the Hand Gesture Recognition System: Current Progress and Future Directions", IEEE Access, **vol.9**,pp.157422-157436,2021,,doi: 10.1109/ACCESS.2021.3129650. (2021)

[3] Davinder Kumar and Aman Ganesh, a critical review on hand gesture recognition using sEMG:Challenges, application, processes and techniques, Journal of Physics:Conference Series, **Volume 2327**, 4th International Conference on Intelligent Circuits and Systems, DOI 10.1088/1742-6596/2327/1/012075.

[4] I. Dhall, S. Vashisth and G. Aggarwal, "Automated Hand Gesture Recognition using a Deep Convolutional Neural Network model," 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2020, pp. 811-816, doi: 10.1109/Confluence47617.2020.9057853.(2020)

[5] S. Nitnaware and A. Bagde, "Hand Gesture Recognition to Facilitate Tasks for the Disabled," 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 2022, pp. 949-953, doi: 10.1109/ICAIS53314.2022.9743056. (2022)

[6] S. Shanmugam, L. S A, P. Dhanasekaran, P. Mahalakshmi and A. Sharmila, "Hand Gesture Recognition using Convolutional Neural Network," 2021 Innovations in Power and Advanced Computing Technologies (i-PACT), Kuala Lumpur, Malaysia, 2021, pp. 1-5, doi: 10.1109/i-PACT52855.2021.9696463 (2021)

[7] J. -H. Sun, T. -T. Ji, S. -B. Zhang, J. -K. Yang and G. -R. Ji, "Research on the Hand Gesture Recognition Based on Deep Learning," 2018 12th International Symposium on Antennas, Propagation and EM Theory (ISAPE), Hangzhou, China, 2018, pp. 1-4, doi: 10.1109/ISAPE.2018.8634348(2018).

[8] Al Farid F, Hashim N, Abdullah J, Bhuiyan MR, Shahida Mohd Isa WN, Uddin J, Haque MA, Husen MN. A Structured and Methodological Review on Vision-Based Hand Gesture Recognition System. J Imaging. 2022 May 26;8(6):153. doi: 10.3390/jimaging8060153. PMID: 35735952; PMCID: PMC9224857.

[9] Rao, G.M., Kumar, P., Kumari, G.V., Pande, A., Babu, G.R. (2011). An Investigation into the Use of Partial Face in the Mobile Environment. Advances in Visual Computing. ISVC 2011. Lecture Notes in Computer Science, vol 6939. Springer, Berlin,Heidelberg. DOI:10.1007/978-3-642-24031-7_53(2011).

[10] Dhadigi Naga Nishanth, G. MallikarjunaRao,Liveness Detection Based on Human eye Blinking for Photo Attacks,International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249-8958 (Online), **Volume-9** Issue-1,October, 2019(2019)

[11] P. K. Pisharady and M. Saerbeck, "Recent methods and databases in vision-based hand gesture recognition: A review", Comput. Vis. Image Understand., **vol. 141**, pp. 152- 165, Dec. 2015.

[12] M. J. Cheok, Z. Omar and M. H. Jaward, "A review of hand gesture and sign language recognition techniques", Int. J. Mach. Learn. Cybern., **vol. 10**, no. 1, pp. 131-153, Jan. 2017.

[13] D. Dusa and M. R. Gundavarapu, "Smart Framework for Black Fungus Detection using VGG 19 Deep Learning Approach," 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2022, pp. 1023-1028, doi: 10.1109/ICACCS54159.2022.9785123.

[14] M. R. Gundavarapu, V. V. R. Yannam, A. Velagala, S. R. Lankela, S. K. G and S. C. Regonda, "Smart Bot for Handwritten Digit String Recognition," 2022 International Conference for Advancement in Technology (ICONAT), Goa,India,2022,pp.1-5, doi:10.1109/ICONAT53423.2022.9726081.