Volume 3, Issue 3, 2025

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

GESTURE-DETECT: A DEEP LEARNING FRAMEWORK USING CONVOLUTION NEURAL NETWORK FOR HAND GESTURE

Qamar-un-Nisa Abrejo[*1], Laraib Kehar[2], Fahad Shamim[3], Sassue Khatoon[4]

[*1,2,3,4]*Institute of Biomedical Engineering and Technology, LUMHS, Jamshoro, Pakistan*

[*1]qamarunnisaabrejo@gmail.com, [2]laraib.kehar@lumhs.edu.pk, [3]fahad.shamim@lumhs.edu.pk, [4]sassue.khatoon@lumhs.edu.pk

DOI: https://doi.org/10.5281/zenodo.15461715

Keywords
Deep Learning (DL), Neural Network (NN), Convolution Neural Network (CNN), Hand gesture recognition (HGR), MATLAB, real time, human computer interaction (HCI)

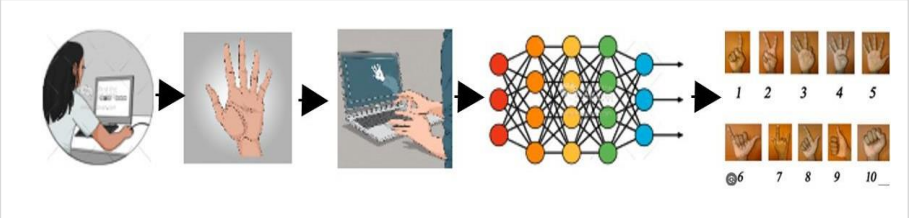
Article History
Received on 07 April 2025
Accepted on 07 May 2025
Published on 16 May 2025

Corresponding Author: *
Qamar-un-Nisa Abrejo

Abstract

Hand gesture recognition refers to non-verbal mode of communication used by deaf and mute community to communicate with normal individuals. Human Computer Interaction (HCI) covers somewhat an interface, comprising hand motions also known as Gestures. A framework possibly used to detect hand motions and express data to operate device. This is an important domain in HCI that comprises interface of device and users. Hand motion detection involves recording specific gestures and detecting them using a camera. Hand motions serve multiple communication purposes. It can help those with disabilities, including as hearing and speech difficulties, as well as stroke patients, communicate and meet their fundamental needs. The technique of Human computer interaction is crucial in health sectors as for as the people with hearing loss and have speech disorders. Regardless of substantial advances for hand motion detection in computer vision-based system for language interpretation, in this sector a significant challenge remains which is limited in scope and only considers common motions. As a result, an innovative hand-vision based Deep learning model which is Convolution Neural Network provides various advantages, including increased accuracy, robustness to variations, and real-time performance. This framework can amplify for synchronous (Real-time) functioning, trained from larger datasets, and can be effortlessly scaled to handle complicated recognition tasks ultimately enabling more effective forms of HCI.

Graphical Abstract:



.

I. INTRODUCTION

Across 350 million or the 5% individuals around the worlds are suffering from acute or by birth hearing losses. This statistical data demonstrates the wide number of hearing loss affected people who depend upon alternative modes of communication to communicate with people who does not have any hearing losses to be normal, focusing the significance of the accessible communications solutions. According to the World Health Organization (2015), these issues affect a significant

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

number of the population, stressing the need of meeting their communication needs [1].

Around 1 in 12 children or 5% to 8% of preschool-aged children experience speech-language disorders (Disability info: speech and language disorders Factsheet (FS11), 2008), which are the most prevalent childhood disabilities. A speech and language disorder can be defined as a speech or sound production, fluency, voice, or language disability that has significant effects on a student's academic career in addition to their social, emotional, or professional development. Youngsters may stutter or have trouble pronouncing words correctly [2]. The impressions that are often being used for social interactions are "speech" and "gestures". Gestures in the form of sign language is utilized by people who have acute hearing loss or are dumb completely across the world for communicating with others who does not have any hearing problems. A gesture is a mixture of facial expressions and different hand movements to express the situation to the fellow people. The sign language serves as a complete meaning of sentences or a particular word [3].

The use of Sign language can only be possible by specialized trainings. The use of gesture language is also difficult to normal people to understand [3]. Hand gestures are an essential part of in-person or face-to-face communication. Thus, all hand gestures and an individual's body language are important in face-to-face communication [4]. In today's times, human-machine interaction is crucial to operate machines virtually through commands provided by humans [5]. Human-computer interactions goals to upgrade the usability and improve the user experiences by developing logical, effective, and easy-to-use frameworks. The fundamental objective of human computer interface is to utilize User-interface interactions, guaranteeing technology meets user need productively. The interface of hand signals in this human-computer interface has captured consideration among individuals due to its efficient way to interact with computer systems.

A hand motion is an approach to communicate a person using movements of fingers with palm motions. In human-computer interaction, Gestures can be integrated with devices in order to eliminate the need of an extra input device cause for making the human-computer interaction successful, there is a need of a device must be able to recognize hand gestures. [6]. By the use of this kind of human-computer interaction, a user could run a diverse range of gadgets virtually using hand gestures [7]. Integrating sign language into human-computer interaction is significant for facilitating communication individuals with acute or usual hearing losses. Designing architecture that can recognize and understand the facial expressions, hand motions and body language is a part of this whole integration. By introducing such technologies like video analysis, deep learning, and motion capture, these frameworks enable the computer system to recognize sign language in an authentic and easy-to-use pattern. Sign language implementation enhances the access thus user interface as human-computer interaction improves to showcasing more uniqueness, participating in bridging the communication barriers and making technology more accessible for everyone [8].

Moreover, a hand motion detection is to be capable to correctly detect every finger and palm orientations in real-time so it can efficiently replace a need of mouse and a keyboard [9]. Modern discoveries in deep learning have wide-ranging outcomes, broadening beyond established domains like activity recognition, health diagnosis and energy identifications by executing computer knowledge, hand motion can be used to assist as a crucial connection between the two normal and impaired individual and communities, eliminating the communication difference more effectively. This can have a notable impact as it would encourage the people who have hearing losses to connect more fruitfully to the individuals with no hearing problems in their environments, achieving equal access to education, employments, and other more opportunities that were beyond grasp. In the following study, we introduce a computer hand-vision based interface for recognizing hand motions in sign language translation systems [10].

This research emphasizes a crucial algorithm to detect a variety of static hand motions implementing HCI using both image extraction and supervised feed-forward NN with back propagation training algorithm for recognition based on hand segmentation [11]. This proposed work approaches

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

the gesture recognition system implementing image blocking and brightness matching factor, achieving 96% accuracy. It focuses on appearance-based gestures for efficient recognitions [12]. This approach proposes a reliable marker-less hand gesture recognition system, detecting both static and dynamic motions. This system converts hand gestures into commands like opening websites, controlling presentations and launching applications. It Demonstrates a simple HCI with minimal requirements of hardware [13].

This paper adds temporal gesture recognition to enhance crucial, context sensitive interactions, with SFAM enabling user defined gesture training [14]. This paper demonstrates the integration of hand gesture with an electronic device to control its parameters [15]. In this paper, a gesture recognition system for HCI is provided. It examines thoroughly knowledge of body regarding software development frameworks, representation and recognition techniques, and taxonomies for gestures. This mainly focuses the three phases i.e., detection, traction and recognition. It also analyzes the parameters and classifies application into core and advanced regions. It also highlights formally given shortcomings and suggestions to improves gesture related HCI frameworks to increase the efficiency for future goals [16]. As technology speeds up its pace, it seems like people are getting more comfortable in communication and other chores via computer on a daily basis. Human- computer interaction will advance to a point where users will need to give more dynamic instructions to the computer system than is currently possible with a mouse and keyboard. With the aid of deep neural networks, computer tools such as Gesture Recognition may significantly promote this type of dynamic control [17].

In this research, two area technique for hand gesture is being presented. The region of interest is isolated in the first phase by removing background pixels with the help of color space segmentation. In the second stage, the classified images are compiled into a CNN model for image classification [18]. The goal of this proposed work is to examine three distinct methods for hand gesture recognition (HGR) that use fingertip detection. A new technique known as "Curvature of Perimeter" is introduced, and it is used as a virtual mouse. The only tools utilized by the system are a webcam and algorithms created with MATLAB's computer vision, image, and video processing toolboxes [19].

In this research, an artificial neural network was recently used as a gesture categorizer. For training, 120 hand gesture photos were employed. 95% is the average classification rate that was ultimately attained. Developing a distinct input method for controlling the media player, approaches to provide several gesture commands that may be helpful in virtual applications. Our method outperforms other hand motion detection systems in terms of accuracy, according to comparisons [20].

II. METHODOLOGY

In this research, Deep Learning Technique is being used to recognize hand gestures. The network is built implementing a CNN. The dataset for single gesture is taken using a webcam and used as input to recognize. Each gesture consists 300 images and the whole dataset is of 6000 images for 20 gestures.

Collect a wide set of hand motion in various directions in the form of images from the different sources or capture directly on your own using webcam of MATLAB and make sub-folders for each gesture individually. Ensure that the dataset includes a variety of hand motions, hand positions, lightened and clear backgrounds. Divide the whole dataset into training and testing subsets, maintaining a good distribution of samples across all categories.

For the data collection, make a Folder of all gestures input images and Load hand gesture images into MATLAB using function like imageDataStore. In preprocessing phase, preprocess the photos to standardize the format and improve model performance. To maintain consistency in input dimensions, resize the photos uniformly using imresize. To increase data variability and improve model generalization, use transformations such as rotation, scaling, and translation. Break the dataset into the training and testing sets using methods such as splitEachLabel to ensure that the models are trained and evaluated separately.

In initialization of model training, construct a convolution neural network architecture for hand gesture identification using deep learning Toolbox in MATLAB. Demonstrate the architecture by

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

defining the number and configuration of convolution, pooling and fully connected layers. Initialize the CNN model with pre-training architectures, Specify the training option (such as optimizing technique), learning rate, mini-batch size and epoch count. The CNN model is compiled by the training dataset implementing trainNetwork function. To confirm the convergence and avoid overfitting, monitor training progress with validation data and visualize training curves (for example accuracy loss) with MATLAB charting methods in the command window.

Lastly initiate Testing after training, exercise the trained dataset to examine the CNN model' proficiency to extend the earlier unseen data. Implement the Load function to obtain the trained model's parameters and architecture. Label hand motion images from the testing dataset with the classify function (or simply make separate folders for each gesture and set current directory to that folder), which applies the trained model to input images and predicts their classifications and show outputs. Rest the block diagram for each step is shown above in Figure 1.

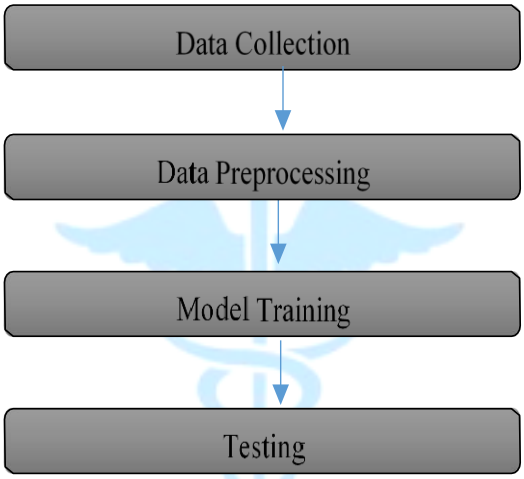


Figure 1: Block diagram for steps implemented in gesture recognition training model

III. RESULTS

In this model, 20 different hand movements were trained into the hand gesture recognition system. Integration of a diverse range of hand movements enabled a comprehensive examination of the model's capabilities across multidimensional settings. Providing a more detailed understanding of its performance. Results shows that the model is effective in recognizing a variety of hand movements, focusing its potential applications in virtual reality, sign language translation and human-machine interactions. Moreover, the model's flexibility and customizability were enhanced by implementing a wide range of gestures, making it more compatible for launching in different environments and user interactions. All encompassing, the results emphasize the effectiveness and extensibility of the developed hand motion recognition system in accommodating a variety of gesture inputs, hence enhancing its utility and applicability across different domains. After execution of Dataset of trained movements, the real-time output can be seen on a different Tab of MATLAB. When webcam detects no gesture, it shows the name of the user as shown in Figure 2.

Similarly, rest 20 gestures are individually assigned a sentence with a gesture to be shown while performing that specific hand motion in front of the webcam. In this collage referred as Figure 3.

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

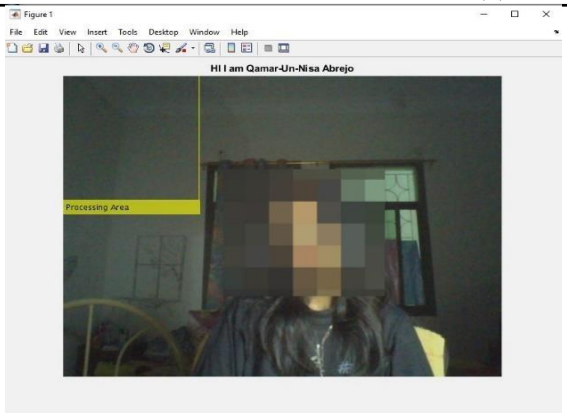


Figure 2: The first output with no sign language is showing name

Represents the nine images, each frame performs a different hand motion accompanied by an assigned sentence that exemplifies its meaning. These gestures and their assigned sentences work together to convey a message or emotion to the viewer, enhancing understanding and engagement to the normal people and the affected

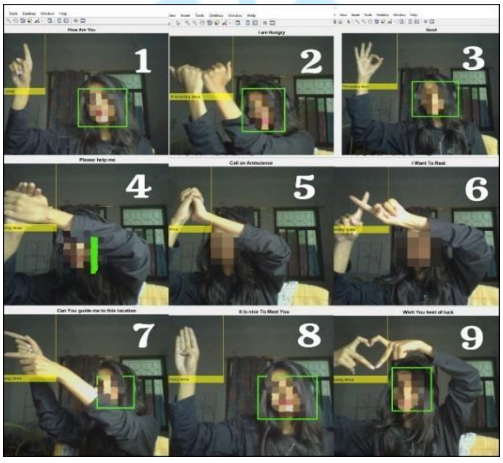


Figure 3: A collage for 9 gestures with assigned sentences

The model was carefully validated to make sure that it performs its intended functions correctly. The gesture recognition system that has been developed and examined has gone through and tested several cycles of testing that are meant to produce valid and accurate outputs. The procedure was done to establish the capacity of the system to identify and distinguish the learnt gestures in a real-life environment and in real time interactions. It has been discovered through the graphs that the total amount of successes that were made at particular time intervals is in exact detail to test that how many times a gesture can be shown accurately, and failures in terms of error or delays that a sentence takes to process the assigned sentence on the window as shown in Figure 4.

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

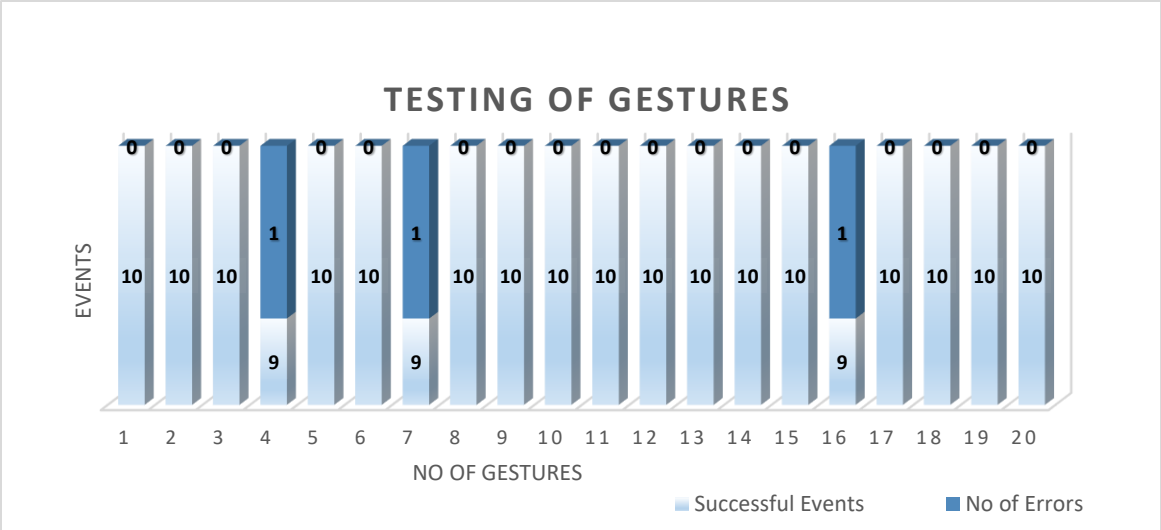


Figure 4: A Graphical representation of No. of successful events and No. of errors

IV. CONCLUSION

Within this research, we developed and tested "GESTURE-DETECT" which is an effective hand gesture recognition system developed via Convolution neural network using Deep learning toolbox in MATLAB. The system's excellent performance in detecting various kinds of hand movements demonstrated CNN's ability for gesture categorization tasks. Rapid Experiments were facilitated by the processing of MATLAB environment, which made data collection and model creation easier.

As a rule, deep learning more particularly convolution neural network (CNN) offers significant benefits in terms of both recognition accuracy and flexibility in handling intricate gesture variations. This approach has the potential to be used in real-time systems including dynamic gaming, virtual reality user interfaces, and sign language interpretation. In order to enhance generalization across various hand sizes, orientations, and lighting conditions, Future work will focus on improving the technology for real-time use and including more diverse datasets to improve generalization across different hand scales, orientations and lighting conditions. Implementing this technique in integrated systems might also make wider application in portable and mobile devices possible.

V. REFERENCES

Aurangzeb, "Deep Learning Approach for Hand Gesture Recognition: Applications in Deaf Communication and Healthcare," *Comput. Mater. Contin.*, vol. 78, no. 1, pp. 127–144, 2024, doi: 10.32604/cmc.2023.042886.

Adebayo, "THE IMPACT OF SPEECH AND LANGUAGE DISORDERS ON PUPILS ', LEARNING AND SOCIALIZATION IN SCHOOLS OF KALOMO DISTRICT-ZAMBIA," vol. 1, no. 2, pp. 141–152, 2014.

Husain, "G Esture R Ecognition S Ystem Using M Atlab : a L Iterature R Eview," vol. 2, no. 11, pp. 425–432.

Bhushan, "An Experimental Analysis of Various Machine Learning Algorithms for Hand Gesture Recognition," *Electron.*, vol. 11, no. 6, 2022, doi: 10.3390/electronics11060968.

Neethu, "An efficient method for human hand gesture detection and recognition using deep learning convolutional neural networks," *Soft Comput.*, vol. 24, no. 20, pp. 15239–15248, 2020, doi: 10.1007/s00500-020-04860-5.

Alnuaim, "Human-Computer Interaction with Hand Gesture Recognition Using ResNet and MobileNet," *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/8777355.

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

Maung, "Real-time hand tracking and gesture recognition system using neural networks," *World Acad. Sci. Eng. Technol.*, vol. 38, pp. 470–474, 2009.

Siby, "Hand Gesture Recognition," no. 3, pp. 1946–1949, 2015.

Agrawal, "Real Time Hand Gesture Recognition for Human Computer Interaction," *Proc. - 6th Int. Adv. Comput. Conf. IACC 2016*, no. February 2016, pp. 470–475, 2016, doi: 10.1109/IACC.2016.93.

Alnaim, "Hand Gesture Recognition using Deep Learning Hand Gesture Recognition using Deep Learning Neural Networks," no. December, pp. 1–161, 2019.

Ummah, "No 主観的健康感を中心とした在宅高齢者における　健康関連指標に関する共分散構造分析Title," *Sustain.*, vol. 11, no. 1, pp. 1–14, 2019, [Online]. Available: http://scioteca.caf.com/bitstream/handle/12 3456789/1091/RED2017-Eng-8ene.pdf?sequence=12&isAllowed=y%0A http://dx.doi.org/10.1016/j.regsciurbeco.2008.06.005%0Ahttps://www.researchgate.net/publication/305320484_SISTEM_PEMBETUNGAN_TERPUSAT_STRATEGI_MELESTARI

Kim, "Brightness Factor Matching for Gesture Recognition System Using Scaled," vol. 75, no. 3, pp. 1134–1139, 2009.

Haria, "Hand Gesture Recognition for Human Computer Interaction," *Procedia Comput. Sci.*, vol. 115, pp. 367–374, 2017, doi: 10.1016/j.procs.2017.09.092.

J. R, "A Method for Temporal Hand Gesture Recognition Keywords 2. The Previous Gesture Recognition System," pp. 1–5.

Freeman, "Television Control by Hand Gestures," *Proc. Int. Work. Autom. Face Gesture Recognit.*, pp. 179–183, 1995, [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.152.8786&rep=rep1&type=pdf

Rautaray, "Vision based hand gesture recognition for human computer interaction: a survey," *Artif. Intell. Rev.*, vol. 43, no. 1, pp. 1–54, 2015, doi: 10.1007/s10462-012-9356-9.

Chakradhar, "International Journal of Research Publication and Reviews Next Word Prediction Using Deep Learning," vol. 3, no. 11, pp. 1182–1189, 2022.

Chang, "An Exploration into Human– Computer Interaction: Hand Gesture Recognition Management in a Challenging Environment," *SN Comput. Sci.*, vol. 4, no. 5, 2023, doi: 10.1007/s42979-023-01751-y.

Farooq, "Real time hand gesture recognition for computer interaction," *2014 Int. Conf. Robot. Emerg. Allied Technol. Eng. iCREATE 2014 - Proc.*, no. April, pp. 73–77, 2014, doi: 10.1109/iCREATE.2014.6828342.

Jalab, "Human computer interface using hand gesture recognition based on neural network," *2015 5th Natl. Symp. Inf. Technol. Towar. New Smart World, NSITNSW 2015*, 2015, doi: 10.1109/NSITNSW.2015.7176405