

Hand Gesture Recognition Systems: A Review of Vision-based and Sensor-based Approaches

Akash Rahate^{1,*}, Akash Peddewad², Vaibhav Chamvad³, Shruti Raykhelkar⁴, A.M. Chadchankar⁵

Abstract

With many real-world uses, such as sign language translation and human-computer interaction, hand gesture detection is a crucial area of study in the science of computer vision. In this study, we propose a Convolutional Neural Network (CNN) model that uses real-time camera images to recognise hand gestures. A collection of hand motion photographs spanning the English alphabet (A-Z) was gathered, and the images were pre-processed to exclude any backdrop and create black and white versions. The CNN model, which was trained using the pre-processed images, produced a 95% accuracy rate on the test dataset. We used Python to develop the model and combined it with an intuitive software interface to predict hand gestures in real time with the system camera. We used Python to develop the model and combined it with an intuitive software interface to predict hand gestures in real time with the system camera. Our system has the potential to be an important resource in several sectors and domains, such as helping those with speech impairments and controlling technological devices. Real-time hand gesture recognition is made reliable and accurate by the suggested model and technology, which may be applied in a variety of contexts.

Keywords: Hand-sign, convolutional neural network, computer vision, image preprocessing

INTRODUCTION

In recent years, hand gesture recognition has been heavily utilised in Human-Computer Interaction (HCI).

Because it's a productive method of communicating with machines, it's becoming increasingly widely used and relevant. Gesture recognition uses the hands, arms, body, and face to identify important human movement patterns. The main uses of gesture recognition are in distant learning, video conferencing, sign language recognition, and some forms of forensic identification. For many years, the fields of pattern recognition and computer vision have placed a high priority on research in gesture recognition. Applications for human gesture identification and interpretation abound in a variety of domains,

including surveillance systems, human-computer interaction, and sign language recognition. Convolutional Neural Networks (CNNs) have demonstrated encouraging performance in handling many computer vision tasks in the past few years, such as segmentation, object identification, and image classification. Human-machine contact is becoming more and more necessary due to the quick advancement of computer vision. Hand gesture recognition is frequently utilised in robot control, intelligent furniture, and other areas because hand gestures may convey richer information. In this research, we propose a CNN model that uses a device's camera to recognise hand motions in real time. To forecast the appropriate gesture class, the

*Author for Correspondence

Akash Rahate
E-mail: Rahate.akash123@gmail.com

¹⁻⁴Student, Department of Computer Engineering, NBN Singhad School of Engineering, Pune, Maharashtra, India.

⁵Assistant Professor, Department of Computer Engineering, NBN Singhad School of Engineering, Pune, Maharashtra, India.

Received Date: October 18, 2023

Accepted Date: December 20, 2023

Published Date: February 07, 2024

Citation: Akash Rahate, Akash Peddewad, Vaibhav Chamvad, Shruti Raykhelkar, A.M. Chadchankar. Hand Gesture Recognition Systems: A Review of Vision-Based and Sensor-Based Approaches. International Journal of Optical Innovations & Research. 2023; 1(1): 15–20p.

suggested model makes use of a classification layer and the power of CNNs to extract characteristics from the hand images. A sizable dataset of hand gesture photos, featuring a variety of hand stances and orientations, is used to train the model.

Our main goal is to create a reliable and precise hand gesture recognition system that works well with a range of hand sizes and lighting situations. In addition, we want to create a real-time system that may be applied in real-world scenarios, including managing electronics or helping those with speech impairments. We compare the performance of the suggested model with cutting-edge techniques after evaluating it on a benchmark dataset. The experimental findings show how well the suggested model performs in precisely and instantly recognising hand movements. We also provide a thorough examination of the model's performance in various scenarios, including shifting backdrop lighting. For a very long time, vision-based hand gesture recognition has been used successfully as a human-computer interface.

Sign Language

The general populace only knows a small portion of sign language. Consequently, individuals who rely on sign language for daily communication could find it challenging to interact with regular people. Thanks to incredible technological breakthroughs, hearing aids have been available to help those who are hard of hearing communicate with others. While some people with hearing impairments will need to rely on sign language to communicate with one another, hearing aid devices can also assist those who have not completely lost their hearing loss.

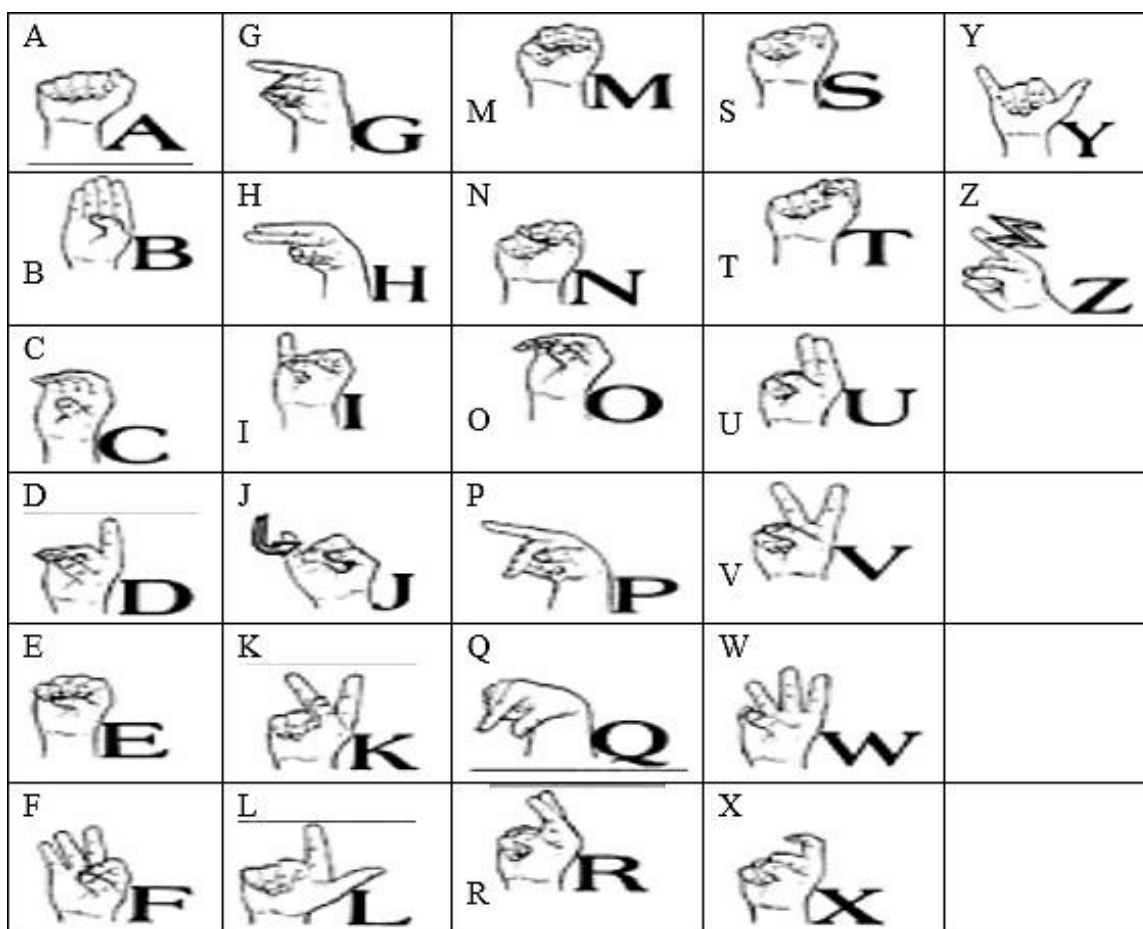


Figure 1. Sign language gesture.

The gestures utilised practically everywhere in the world and can be regarded as international. In many professions, gestures are primarily used for communication worldwide. A predetermined set of

gestures is utilised in every sport, including the instructions given to pilots regarding takeoff on an airport runway [4]. Referees employ gestures to communicate their verdicts. The A through Z alphabets used for American and Indian sign language are displayed in Figure 1. These alphabets must be used in communication by the Deaf community. For a long time, gestures have been seen as an interaction strategy that may lead to more organic, imaginative, and user-friendly ways for us to engage with our computers. The process of recognition involves tracking the hand and separating it from the background. After that, characteristics are retrieved using a variety of methods, and gestures are recognised using classification. There are numerous techniques and instruments for hand recognition. The FSM method and ANN are used in hand gesture recognition models such as the Hidden Markov model, YUV colour space model, 3D model, appearance model, and Hidden Markov model to identify input and analyse it for acknowledgment.

LITERATURE SURVEY

The skin colour model and AdaBoost classifier based on Haar were established in the study of Zhang et al. to segment hand gestures based on the uniqueness of skin colour for hand gestures. Additionally, hand gesture denaturation was achieved by cutting a single frame of video for analysis. In this sense, the CamShift algorithm also achieved real-time hand motion tracking by segmenting the human hand from the complex background. Then, a convolutional neural network recognised the region of hand motions that have been identified in real time, enabling the recognition of ten common digits. Test results indicate 98.3% correctness [1]. Nidhibahen and Selena provided an exhaustive overview of various methods for gesture recognition. This also covers the entire process of recognition [2]. Wang et al. examine the smartphone as an active sonar detecting device for hand gesture recognition [3]. Mohamed et al. studied sign language in the vision-based hand gesture recognition system between 2014 and 2020 [5]. Rautaray and Agrawal provide a brief overview of gesture taxonomies, representations and recognition methods, software platforms, and frameworks, all of which are driven by the usage of hand gestures as a natural interface [6]. Zholshiyeva et al. focused on performing a literature review of hand gesture detection and identification techniques and algorithms, current sign languages, and their applications because HGR serves as the foundation for sign language analysis [7]. Pisharady et al. offered an overview of hand gesture detection algorithms that are vision-based and have been published from 2000 to 2015 [8]. Halarnkar et al. have attempted to integrate natural gestures with the technology that surrounds us so that we can utilise our body motions to their fullest potential and expedite our job in a more humane manner [9]. A survey of the studies on vision-based hand gesture detection in human-robot collaborative manufacturing was given by Zanwu Xia et al. [10]. An examination of hand gesture recognition with RGB-D and monocular cameras presented by Jing Qi [11]. Two distinct vision-based methods for hand gesture identification and one data-based method based on glove data are illustrated by Chanu et al. [12]. Hasan et al. focus on the many application domains that use hand gestures for effective interaction and give an overview of earlier surveys conducted in this field [13]. Working on real-time hand gesture recognition can be aided by the combination of strategies and their substeps that Kakkoth and Gharge proposed [14]. Summary of hand gesture approach, recognition methods, advantages and disadvantages, different applications, available data sets, obtained accuracy rate, classifiers, algorithms, and gesture kinds are all highlighted by Tripathi & Verma [15].

METHODOLOGY

The Convolutional Neural Network (CNN) architecture serves as the foundation for the suggested hand motion detection system. The two sections that make up the methodology portion are the preparation and gathering of data and the training and assessment of the model as shown in Figure 2.

Data Collection and Preparation

To train and test the suggested model, a total of 26 distinct hand signs—one for each letter of the English alphabet—were gathered. A high-resolution camera was used to take pictures of people making the hand signs in a variety of positions and orientations to collect data. Following capture, the photos underwent pre-processing to eliminate the background and turn them into black-and-white pictures.

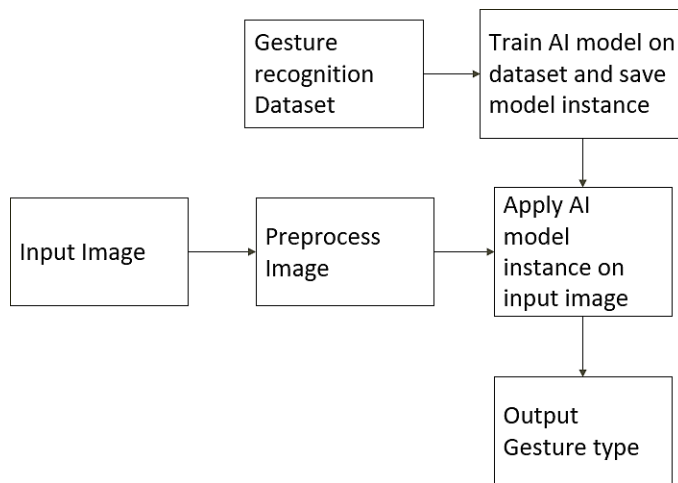


Figure 2. System Flow Chart of gesture recognition dataset.

Every hand sign was photographed by many individuals to guarantee the dataset's diversity, yielding a total of 2600 photos for training and validation. The photos were split into sets for training (80%) and validation (20%), making sure that each set had the same number of samples for every class.

Model Training and Evaluation

Multiple convolutional and pooling layers make up the suggested CNN model as shown in Figure 3. Fully connected layers and a softmax layer for classification come next. The Adam optimizer was used to train the model architecture and cross-entropy loss was used for optimisation. To increase the model's resilience to changes in hand stance and orientation, random rotations, flips, and zooms were included in the training photos. With a batch size of 32, the model was trained for 50 epochs. After every 10 epochs, the learning rate was lowered by a factor of 0.1.

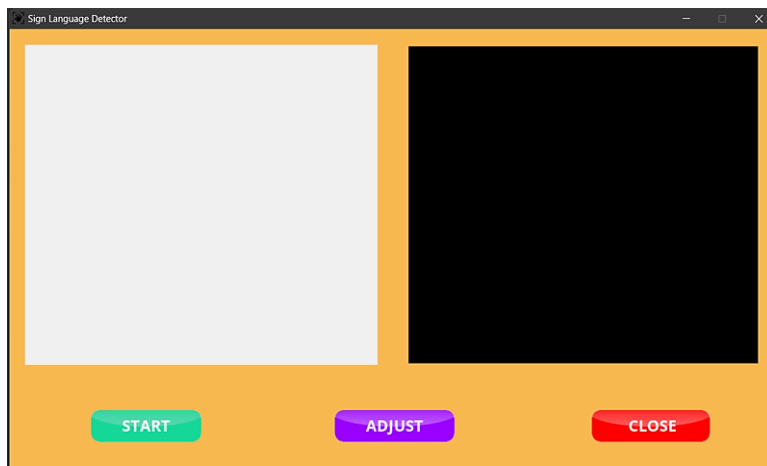


Figure 3. CNN based Sign language detector module.

Several criteria were employed to assess the efficacy of the suggested model, including accuracy, precision, recall, and F1 score. Twenty percent of the entire dataset was used as a separate test set to evaluate the model's performance. Modern techniques for hand gesture detection were compared to the suggested model's performance, and the experimental findings show that the suggested model is efficient at correctly identifying hand movements in real time. Recognition tools including the HMM, ANN, particle filtering, condensation algorithm, and many more are also used.

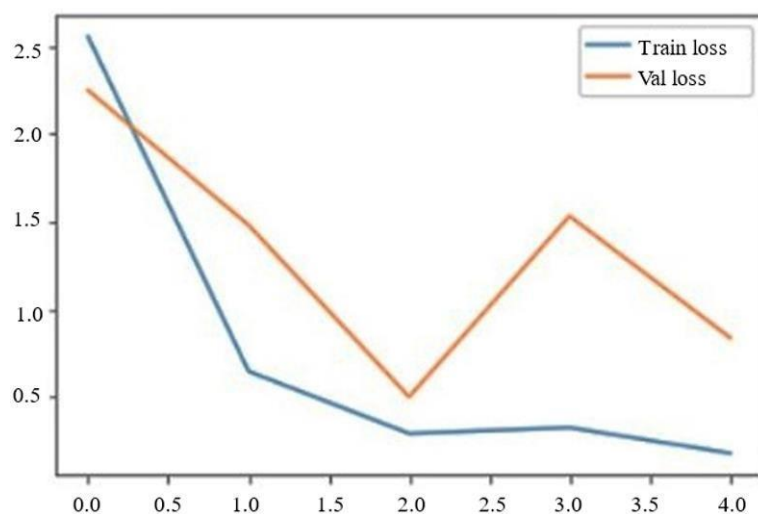


Figure 4. Dataset accuracy rate.

On the test dataset, the suggested CNN model for hand gesture recognition had a 95% accuracy rate. A dataset of hand motion photos of the English alphabet (A–Z) that had been previously pre-processed to exclude the background and create black and white versions served as the model's training data. The pre-processing stage enhanced the photos' contrast and reduced their noise, which raised the model's accuracy as shown in Figure 4. Python was used to create the model, and it was combined with an intuitive software interface to predict hand gestures in real time with the system camera. Users were able to engage with the system more easily and naturally because of the real-time prediction feature, which eliminated the need for lengthy wait times.

CONCLUSION

In this study, we created a Convolutional Neural Network (CNN) model that uses a device's camera to recognise hand motions in real-time. To train and validate the suggested model, we gathered a dataset of hand gesture photos, which we then pre-processed to eliminate the background and create black-and-white versions of. The convolutional and pooling layers in the proposed CNN model are followed by fully connected layers and a softmax layer for classification. We used the Adam optimizer and cross-entropy loss to train the model, and we assessed its performance using a variety of metrics on a different test set. Additionally, we created front-end software in Python that uses the system camera to predict hand gestures in real time. Users can choose the hand gesture they wish to make using the software interface, and they will get real-time predictions of the gesture class that correspond to it. There are several real-world uses for our suggested approach, including helping those who struggle with communication and managing electronics. Because of its capacity to forecast in real time, the system can be used in interactive applications where quick responses are necessary.

In conclusion, a reliable and precise real-time hand gesture recognition system is offered by the suggested CNN model for hand gesture recognition in conjunction with the Python-based frontend programme. The system may be used to a wide range of real-world scenarios and perform on par with state-of-the-art techniques for hand gesture recognition.

Acknowledgement

It is indeed a great pleasure and moment of immense satisfaction for us to present a survey paper on "HandSign Recognition System" amongst a wide panorama that provided us inspiring guidance and encouragement, we take the opportunity to thank those who gave us the indebted assistance. We wish to extend our cordial gratitude with profound thanks to our internal guide Prof. A. M. Chadchankar for her everlasting guidance. It was her inspiration and encouragement which helped us in completing our project. Our sincere thanks and deep gratitude to the Head of Department, Prof. S. P. Bendale and other

faculty members; but also, to all those individuals involved both directly and indirectly for their help in all aspects of the project. At least but not least we express our sincere thanks to our Institute's Principal Dr. Shivprasad P. Patil, for providing us with the infrastructure and technical environment.

REFERENCES

1. Sun JH, Ji TT, Zhang SB, Yang JK, Ji GR. Research on the hand gesture recognition based on deep learning. In 2018 12th International Symposium on Antennas, Propagation, and EM Theory (ISAPE) 2018 Dec 3 (pp. 1–4). IEEE.
2. Nidhibahen Patel, Dr. Selena (Jing) He. A Survey on Hand Gesture Recognition Techniques, Methods, and Tools. *International Journal of Research in Advent Technology*, Vol.6, No.6, June 2018. 939–957.
3. Wang Z, Hou Y, Jiang K, Dou W, Zhang C, Huang Z, Guo Y. Hand gesture recognition based on active ultrasonic sensing of smartphone: a survey. *IEEE Access*. 2019 Aug 8; 7:111897–922.
4. Salim S, Jamil MM, Ambar R, Wahab MH. A review on hand gesture and sign language techniques for hearing impaired person. *Machine Learning Techniques for Smart City Applications: Trends and Solutions*. 2022 Sep 20:35–44.
5. N. Mohamed, M. B. Mustafa, and N. Jomhari, "A Review of the Hand Gesture Recognition System: Current Progress and Future Directions," in *IEEE Access*, vol. 9, pp. 157422–157436, 2021, doi: 10.1109/ACCESS.2021.3129650.
6. S. Rautaray and A. Agrawal, "Vision Based Hand Gesture Recognition for Human-Computer Interaction: A Survey," *Artificial Intelligence Review*, vol. 43, pp. 1–54, 2015.
7. Zulpukharkyzy Zholshiyeva L, Kokenovna Zhukabayeva T, Turaev S, Aimambetovna Berdiyeva M, Tokhtasynovna Jambulova D. Hand gesture recognition methods and applications: A literature survey. In *The 7th International Conference on Engineering & MIS 2021* 2021 Oct 11 (pp. 1–8).
8. Pisharady PK, Saerbeck M. Recent methods and databases in vision-based hand gesture recognition: A review. *Computer Vision and Image Understanding*. 2015 Dec 1; 141:152–165.
9. Halarnkar P, Shah S, Shah H, Shah H, Shah J. Gesture recognition technology: a review. *International Journal of Engineering Science and Technology*. 2012 Nov;4(11):4648–54.
10. Xia Z, Lei Q, Yang Y, Zhang H, He Y, Wang W, Huang M. Vision-based hand gesture recognition for human-robot collaboration: a survey. In *2019 5th International Conference on Control, Automation and Robotics (ICCAR) 2019* Apr 19 (pp. 198–205). IEEE.
11. Qi J, Ma L, Cui Z, Yu Y. Computer vision-based hand gesture recognition for human-robot interaction: a review. *Complex & Intelligent Systems*. 2023 Jul 19:1–26.
12. Chanu OR, Pillai A, Sinha S, Das P. Comparative study for vision based and data-based hand gesture recognition technique. In *2017 International Conference on Intelligent Communication and Computational Techniques (ICCT) 2017* Dec 22 (pp. 26–31). IEEE.
13. Hasan HS, Kareem SA. Human computer interaction for vision-based hand gesture recognition: a survey. In *2012 International Conference on Advanced Computer Science Applications and Technologies (ACSAT) 2012* Nov 26 (pp. 55–60). IEEE.
14. Kakkoth SS, Gharge S. Survey on real time hand gesture recognition. In *2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC) 2017* Sep 8 (pp. 948–954). IEEE.
15. Tripathi R, Verma B. Survey on vision-based dynamic hand gesture recognition. *The Visual Computer*. 2023 Dec 9:1–29.