

Hand gesture recognition using machine learning algorithms

Abhishek B¹, Kanya Krishi², Meghana M³, Mohammed Daaniyaal⁴, Anupama H S⁵

^{1,2,3,4}B. E, Computer Science and Engineering, BMS Institute of Technology, Bangalore, India

⁵BMS Institute of Technology, Bangalore, India

Article Info

Article history:

Received Apr 24, 2020

Revised Jun 14, 2020

Accepted Jun 29, 2020

Keywords:

Gesture recognition

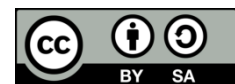
Human–computer interaction

User-friendly interface

ABSTRACT

Gesture recognition is an emerging topic in today's technologies. The main focus of this is to recognize the human gestures using mathematical algorithms for human computer interaction. Only a few modes of human-computer interaction exist, they are: through keyboard, mouse, touch screens etc. Each of these devices has their own limitations when it comes to adapting more versatile hardware in computers. Gesture recognition is one of the essential techniques to build user-friendly interfaces. Usually, gestures can be originated from any bodily motion or state, but commonly originate from the face or hand. Gesture recognition enables users to interact with the devices without physically touching them. This paper describes how hand gestures are trained to perform certain actions like switching pages, scrolling up or down in a page.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Abhishek B,
B.E, Computer Science and Engineering,
BMS Institute of Technology, Bangalore, India.
Email: anupamahs@bmsit.in

1. INTRODUCTION

Gesture recognition is a technique which is used to understand and analyze the human body language and interact with the user accordingly. This in turn helps in building a bridge between the machine and the user to communicate with each other. Gesture recognition is useful in processing the information which cannot be conveyed through speech or text. Gestures are the simplest means of communicating something that is meaningful. This paper involves implementation of the system that aims to design a vision-based hand gesture recognition system with a high correct detection rate along with a high-performance criterion, which can work in a real time human–computer interaction (HCI) system without having any of the limitations (gloves, uniform background etc.) on the user environment. The system can be defined using a flowchart that contains three main steps, they are: learning, detection, recognition as shown in Figure 1.

Learning involves two aspects such as:

- Training dataset: This is the dataset that consists of different types of hand gestures that are used to train the system based on which the system performs the actions.
- Feature extraction: It involves determining the centroid that divides the image into two halves at its geometric centre.

Detection involves three aspects:

- Capture scene: Captures the images through a web camera, which is used as an input to the system.
- Preprocessing: Images that are captured through the webcam are compared with the dataset to recognize the valid hand movements that are needed to perform the required actions.
- Hand detection: The requirements for hand detection involve the input image from the webcam. The image should be fetched with a speed of 20 frames per second. Distance should also be maintained

between the hand and the camera. Approximate distance that should be between hand the camera is around 30 to 100 cm. The video input is stored frame by frame into a matrix after preprocessing.

Recognition consists of:

- Gesture recognition: The number of fingers present in the hand gesture is determined by making use of defect points present in the gesture. The resultant gesture obtained is fed through a 3-dimensional convolutional neural network consecutively to recognize the current gesture.
- Performing action: The recognized gesture is used as an input to perform the actions required by the user.

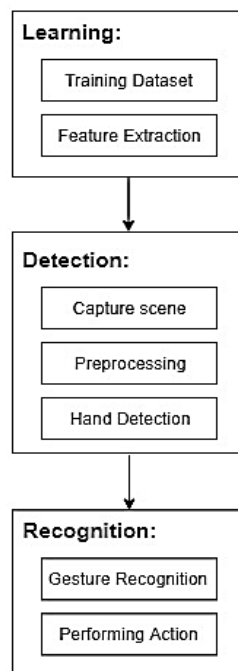


Figure 1. Flowchart of HCI

2. LITERATURE SURVEY

The implementation is divided into four main steps: 1) image enhancement and segmentation, 2) orientation detection, 3) feature extraction, and 4) classification [1]. This work was focused on above four categories but main limitation was change of color was happening very rapidly by the change in the different lighting condition, which may cause error or even failures. For example, due to insufficient light condition, the existence of hand area is not detected but the non-skin regions are mistaken for the hand area because of same color [2]. Involves three main steps for hand gesture recognition system: 1) segmentation, 2) feature representation 3) recognition techniques. The system is based on hand gesture recognition by modeling of the hand in spatial domain. The system uses various 2D and 3D geometric and non-geometric models for modeling. It has used fuzzy c-means clustering algorithm which resulted in an accuracy of 85.83%. The main drawback of the system is it does not consider gesture recognition of temporal space, i.e., motion of gestures and it is unable to classify images with complex background i.e., where there are other objects in the scene with the hand objects [3]. This survey focuses on the hand gesture recognition using different steps like data acquisition, pre-processing, segmentation and so on. Suitable input device should be selected for the data acquisition. There are a number of input devices for data acquisition. Some of them are data gloves, marker, and hand images (from webcam/Kinect 3D Sensor). But the limitation with this work was change in the illumination, rotation and orientation, scaling problem and special hardware which is pretty costlier [4]. The system implementation is divided into three phases: 1) Hand gesture recognition using kinetic camera, 2) algorithms for hand detection recognition, 3) hand gesture recognition. The limitation here is that the edge detection and segmentation algorithms used here are not very efficient when compared to neural networks. The dataset being considered here is very small and can be used to detect very few sign gestures.

The system architecture consists of: 1) image acquisition, 2) segmentation of hand region, 3) distance transforms method for gesture recognition [5]. The limitations of this system involve 1) the numbers of gestures

that are recognized are less and 2) the gestures recognized were not used to control any applications [6]. In this implementation there are three main algorithms that are used: 1) Viola–Jones algorithm. 2) convex hull algorithm, 3) the AdaBoost based learning algorithm. The work was accomplished by training a set of feature set which is local contour sequence. The limitations of this system are that it requires two sets of images for classification. One is the positive set that contains the required images, the other is the negative set that contains contradicting images [7]. The system implementation consists of three components: 1) hand detection 2) gesture recognition, and 3) HCI. It has implemented the following methodology: 1) the input image is preprocessed and the hand detector tries to filter out the hand from the input image, 2) a CNN classifier is employed to recognize gestures from the processed image, while a Kalman Filter is used to estimate the position of the mouse cursor, and 3) the recognition and estimation results are submitted to a control centre which decides the action to be taken. One of the limitations of this system is that it recognizes only the static images [8]. This implementation focuses on detection of hand gestures using java and neural networks. It is divided into two phases: 1) Detection module using java where in the hand is detected using background subtraction and conversion of video feed into HSB video feed thus detecting skin pixels; 2) The second module is the prediction module; a convolutional neural network is used. The input feed image is gained from Java. The input image is fed into the neural network and is analyzed with respect to the dataset images. One of the limitations of this system is that it requires socket programming in order to connect java and python modules.

3. IMPLEMENTATION

A hand gesture recognition system was developed to capture the hand gestures being performed by the user and to control a computer system based on the incoming information. Many of the existing systems in literature have implemented gesture recognition using only spatial modelling, i.e., recognition of a single gesture and not temporal modelling i.e., recognition of motion of gestures. Also, the existing systems have not been implemented in real time, they use a pre captured image as an input for gesture recognition. To overcome these existing problems a new architecture has been developed which aims to design a vision-based hand gesture recognition system with a high correct detection rate along with a high-performance criterion, which can work in a real time HCI system without having any of the mentioned strict limitations (gloves, uniform background, etc.) on the user environment. The design is composed of a HCI system which uses hand gestures as input for communication as show in Figure 2.

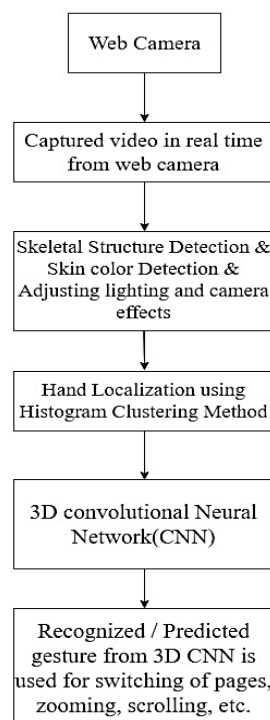


Figure 2. Design of the proposed HCI system

Input to the system is from the web camera or a prerecorded video sequence. Later it detects the skin color by using an adaptive algorithm in the beginning of the frames. For the current user skin color has to be fixed based on the lighting and camera parameter and condition. Once it has been fixed, hand is localized with a histogram clustering method. Then a machine learning algorithm has been used to detect the hand gestures in consecutive frames to distinguish the current gesture. These gestures are used as an input for a computer application as shown in Figure 3. The system is divided into 3 subsystems:

3.1. Hand and motion detection

The Web-camera captures the hand movement and provides it as input to OpenCV and TensorFlow Object detector. Edge detection and skin detection are performed to obtain the boundary of the hand. This is then sent to the 3D CNN.

3.2. Dataset

Dataset is used for training the 3D CNN. Two types of datasets are being used—one for the hand detection and the other for the motion or gesture detection. Hand detection uses EGO dataset, Motion or Gesture Recognition uses Jester dataset.

3.3. 3D CNN

CNN's are a class of deep learning neural networks used for analyzing videos and images. It consists of several layers—input layer, hidden layers, and output layer. It performs back propagation for better accuracy and efficiency. It performs training and verification of the recognized gestures and HCIs take place—turning of the pages, zooming in, and zooming out. The interactions with the computer take place with the help of PyAutoGUI or System Calls.

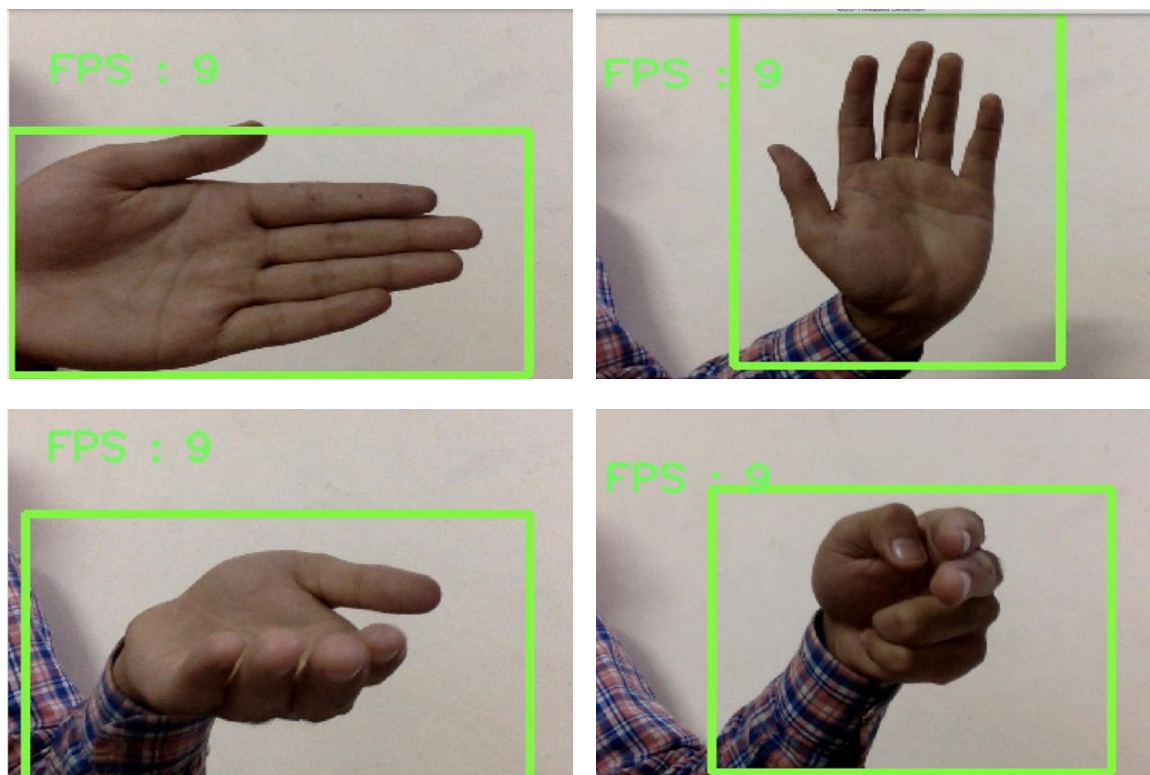


Figure 3. System recognized hand gestures

4. CONCLUSION

The importance of gesture recognition lies in building efficient human-machine interaction. This paper describes how the implementation of the system is done based upon the images captured. Hand detection is done using OpenCV and TensorFlow object detector. And further it is enhanced for interpretation of gestures by the computer to perform actions like switching the pages, scrolling up or down the page.

ACKNOWLEDGEMENT

This work is done, supervised, and supported by the students and faculty members of the Department of Computer Science and Engineering, BMS Institute of Technology, Bangalore.

REFERENCES

- [1] M. Panwar and P. S. Mehra, "Hand gesture recognition for human computer interaction," 2011 *International Conference on Image Information Processing*, Shimla, pp. 1-7, 2011.
- [2] R. Zaman Khan and N. A. Ibraheem, "Comparative Study of Hand Gesture Recognition System," *International Conference of Advanced Computer Science & Information Technology*, 2012.
- [3] A. R. Sarkar, G. Sanyal, and S. Majumder, "Hand Gesture Recognition Systems: A Survey," *International Journal of Computer Applications*, vol. 71, no.15, pp. 25-37, May 2013.
- [4] A. E. Manjunath, B. P. V. Kumar, H. Rajesh, "Comparative Study of Hand Gesture Recognition Algorithms," *International Journal of Research in Computer and Communication Technology*, vol 3, no. 4, April 2014.
- [5] D. R. Jadhav and L. M. R. J. Lobo, "Navigation of PowerPoint Using Hand Gestures," *International Journal of Science and Research (IJSR)*, vol. 4, no. 1, pp. 833-837, 2015.
- [6] R. M. Gurav and P. K. Kadbe, "Real time finger tracking and contour detection for gesture recognition using OpenCV," 2015 *International Conference on Industrial Instrumentation and Control (ICIC)*, pp. 974-977, 2015.
- [7] P. Xu, "A Real-time Hand Gesture Recognition and Human-Computer Interaction System," 2017, [Online]. Available: <https://arxiv.org/abs/1704.07296>
- [8] P. Suganya, R. Sathya, and K. Vijayalakshmi, "Detection and Recognition of Gestures to Control the System Applications by Neural Networks," *International Journal of Pure and Applied Mathematics*, vol. 118, no. 10, pp. 399-405, January 2018.