

Survey paper on Hand Gesture Recognition Based Virtual Mouse

Aabha Waichal¹, Mauli Gandhi¹, Amruta Bhanji¹, Srushti Bhagwat¹, Shalaka Deore²

1 Student, Computer Engineering Department

2 Project guide

Modern Education Society's College of Engineering, Pune, India.

Abstract—The invention of computer has been a turning point in the history. It makes our day-to-day life much simpler. The way to communicate with computer i.e., Human Computer Interaction (HCI), is through input and output devices. Mouse is one of the input devices. Mouse is an important part of Human Computer Interaction (HCI). With evolving technology, the cursor can be controlled without physical mouse hardware but using tip of a finger, hand gestures, etc. These hand movements can be recorded through webcam. The movement of the cursor can also be controlled by color detection or fingertip tracking using a webcam. This new era technology is useful in many fields, in many situations. This can be accomplished using extra hardware component or developing a software system. This is achieved through various image processing techniques and gesture recognition techniques without extra hardware. In this paper we have proposed a system which uses hand gesture recognition-based Convolutional Neural Network (CNN) model for controlling the mouse.

Keywords—Virtual mouse, Hand Gesture Recognition, CNN model

I. INTRODUCTION

Mouse is a pointing device which makes human Computer Interaction easy. There has been much development in the types of mice. Beginning from handheld wired mouse to Bluetooth mouse, then touchpad and touch screen to glove-based detector mouse[1] and recently there have been many newly developed image processing techniques[2].

Image processing is a method of performing operations on an image to get useful information. Gesture recognition is a type of image processing process. Many techniques have been explored till now like 3D convolutional neural network[3], contour and convex hull area[4].

Different pre-processing techniques have been invented and employed for accurate and efficient outcome of the previously mentioned gesture recognition techniques like k-cosine and border-tracing [3], background subtraction[4], computing four motion matrices [5].

In this paper we have proposed an interactive way of controlling the movement of mouse by hand gesture. The in-built webcam is used to capture the live feed. Using python libraries hand detection in the image and background subtraction is done. The pre-processed image is then passed to CNN model for recognizing the gesture and accordingly the cursor is controlled. CNN model is trained by the dataset prepared.

II. LITERATURE REVIEW

Ahmed, Muhammad et al., (2021) [6] provide a very good explanation about various techniques of object detection such as Object Detection (OD), Salient Object Detection (SOD) and Category-specific Object Detection (COD). It reviews different existing methods which employ deep learning and compares them based on their performance in challenging environment. It depicts recent research and the advancement in this field comparing their performance using output, time required and comparing them with traditional methods. It also discusses publicly available datasets for this task as well as explains evaluation metrics used for comparison. It also gives a brief idea about the shortcomings and direction in which further advancements can be made.

V. Tiwari et al., (2020) [7] uses VGC16 as pre-trained model to achieve image classification. It compares its results with different models such as Baseline CNN, three block VGG model. The paper also incorporated the data augmentation in the VGC3 model to study its effect on accuracy. The implemented VGC16 model has the accuracy of 98.97% which is quite an improvement over the accuracy of 55.075%, 74.561%, 61.404% of Baseline CNN, VGC3, VGC3+Data augmentation respectively.

T. Palleja et al. (2008) the system controls the mouse by the movements of head and facial features. It uses an image processing algorithm and compute four motion matrix. Ten frame cumulative image is analysed and region of interest is used for detection of the movement. The process takes some time thus reducing the speed at which the mouse responds.

Li Wensheng, Deng Chunjian and Lv Yi (2010) [8] proposed model consists of two parts-server side and client side. Server side: adaptive on-line training is used for detection of finger using colour histogram, tracking of finger using mean shift algorithm, constructing appropriate message and sending such messages to the client and mouse movement is done by BP neural network. Client side: responses to the messages and obtains the coordinates of multiple fingertips by calling the API function provided by the server, then does the appropriate message processing to achieve the control of the application. The results are not consistent using adaptive on-line training because of the varied skin colour.

Kumar, P. & Verma, J. & Prasad, Shitala (2012) proposed an additional device called data glove is used. The data glove captures the current position of the hand and the angles

between the joints and then these features are used to classify the gestures using K-NN classifier.

A. Mhetar et al., (2014) [9] proposed model consisting of an IR camera, projector, laptop system, Teensy 3.1. The model is based on the concept of virtual marker. The IR camera is interfaced with Teensy 3.1 using I2C communication which is connected to the laptop. The virtual marker provides functions of a mouse by using a high-end microcontroller configured as a Human Interface Device (HID) for faster response interfaced with an IR camera using I2C interface. IR Camera tracks IR sources and sends information regarding its position which is converted to X and Y co-ordinates by Teensy and moves mouse pointer to the position pointed by co-ordinates. The working of the model depends on the availability of the IR camera with specific technical requirements.

S. M. S. Shajideen and V. H. Preetha (2018) [10] proposed model which has hand pointing gestures incorporated with other hand gestures in 3D space. Two USB cameras are used which are placed orthogonal to each other to obtain top view and side view of different hand gestures. MATLAB software is used for it. For separate two views, the two detectors are trained and choosing various image samples for various directions at the top & side view. At the training stage, binary patterns are applied for image conversation and feature generation for each sample. Then two cascade detectors were built which depend on AdaBoost feature selection. At the testing stage, the transformation of the input image to the working image, each and every detector monitors & scan the working image in each view separately.

K. S. Varun et al., (2019) [11] developed models which are based on color detection and mouse movement based on highlighted color which is given from the user for the movement of the mouse. A two figure input can be seen where it is forming two rectangles and forming an average point from both the figures. That point will be acting like a mouse pointer. Once the point moves the mouse pointer in the runtime also moves along. So using this the movement of the mouse can be implemented. The updating of the mouse pointers depends on the position of the prespecified colored caps in the mask that is created for understanding the system. The created mask is converted from RGB background to a black and white that will be used for the detection of the prespecified colored objects that will help for the movement of the mouse. If the prespecified colored caps camouflage with the background they will not be detected and thus no movement of the mouse is possible.

S. R. Chowdhury et al., (2020) [12] proposed model consisting of both mouse and keyboard. The Mouse uses a convex hull process for its working, defects are captured or read, using this defects the functions of the mouse are mapped. The process of this image recognition process solely focuses on defects and conditional statements, the convex hull takes the gap of the fingers as defects, so it can be used for multiple gestures and mapping commands.

Sai Mahitha G. et al., (2021) [13] put forth a different type of model. In this model the mouse cursor of the system can be controlled just on appearing our fingers before the computer's web camera. These finger gestures are captured and controlled through a Color Detection technique of webcam. This system

allows us to direct the system pointer by using our finger bearing color caps or tapes and the operations like dragging of files and the left click would be performed by using distinct finger gestures. It also performs the transfer of files among two PC's in a single similar network. This developed system makes use of only a less resolution webcam which acts as a sensor for tracking the user's hands in two dimensions. If the prespecified colored caps camouflage with the background they will not be detected and thus no movement of the mouse is possible.

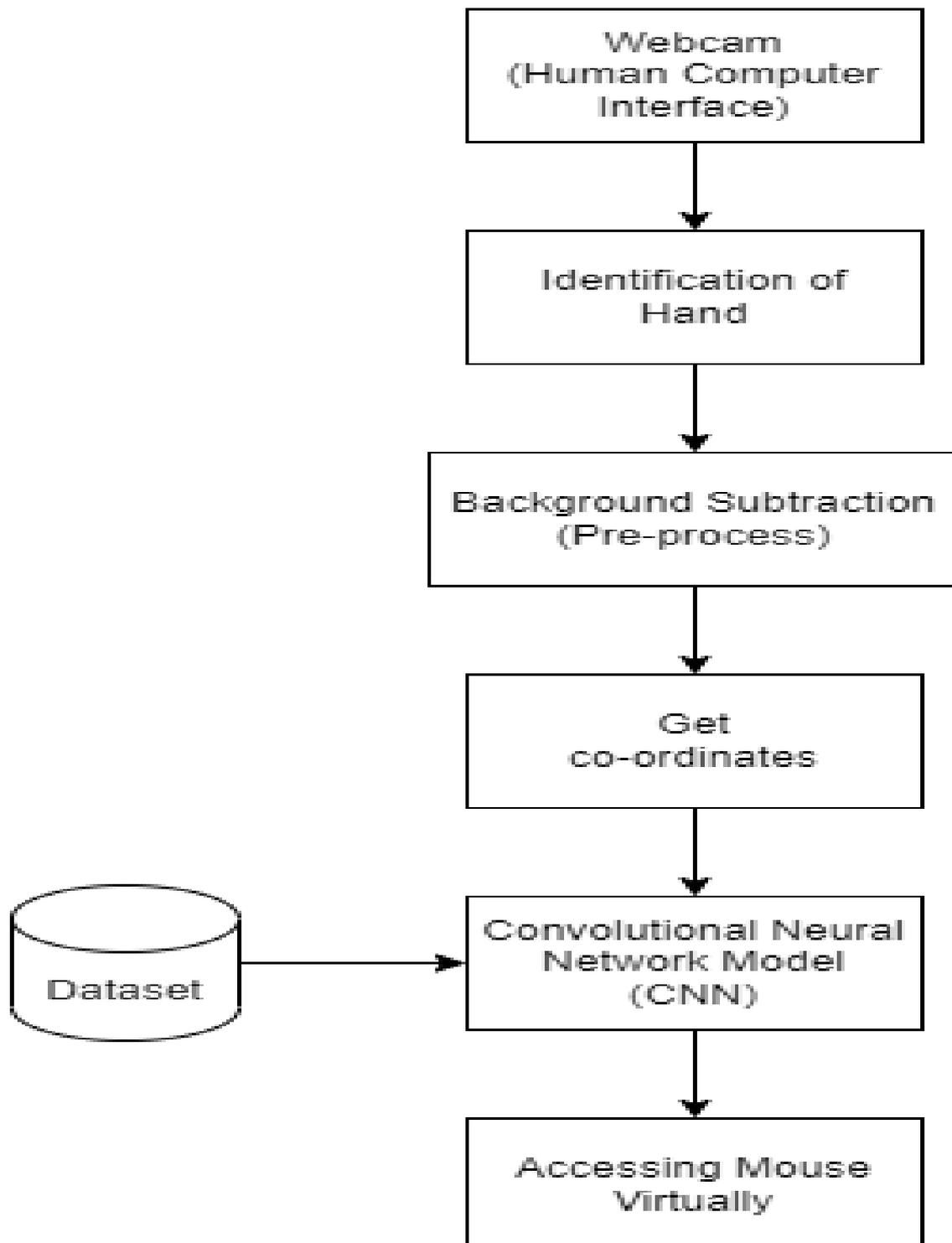
Tran, DS., Ho, NH., Yang, HJ. et al., (2021) [14] proposed virtual mouse method using RGB-D images and fingertip detection. The hand region of interest and the center of the palm are first extracted using in-depth skeleton-joint information images from a Microsoft Kinect Sensor version 2, and then converted into a binary image. Then, the contours of the hands are extracted and described by a border-tracing algorithm. The Kcosine algorithm is used to detect the fingertip location, based on the hand-contour coordinates. Finally, the fingertip location is mapped to RGB images to control the mouse cursor based on a virtual screen. This study still suffers from several limitations that are mainly inherited from Microsoft Kinect.

V. V. Reddy et al., (2020) [15] paper implements algorithm which uses two types of methods for implementation of mouse control. One includes the usage of color caps and other uses the recognition of bare hand gesture recognition. It is divided into two methods; 'fingertip detection' using colored caps and 'gesture recognition'. It involves the interfacing of the video, processing the images by background subtraction. Background subtraction helps in neglecting the steady objects considering only foreground objects. Fingertip detection involves the color identification, circle identification and guessing the fingers. Gesture recognition involves skin color identification, contour detection and convex hull formation and then guessing the gesture. Respective mouse operation can be done accordingly.

We have used this model as our base for the research. We have studied the background subtraction from the model to be incorporated in our project. In this model convex hull is used for the recognition of the gesture, to increase the accuracy of this model we will be replacing convex hull with convolutional neural network.

III. PROPOSED SYSTEM

In this section we have described our proposed system.
Given below is the architecture diagram of our model.



For this project an in-built webcam is needed. Webcam gives live feed to the system. For this the user should be in the webcam coverage range.

The system observes the image frame, if a hand is detected it activates the further processes. For detecting the hand convex hull algorithm is used.

Once the hand is detected in the image frame, the image is pre-processed. Here we have used background subtraction algorithm from OpenCV. Pre-processing is done so that more accurate result can be obtained. Here the hand boundary is identified and then the remaining part is removed.

Using convex hull, the co-ordinates of hand is extracted. We use the co-ordinates for determining the relative position of the mouse pointer. Since it is difficult to move the cursor to all the areas of the desktop screen from a much smaller area of webcam, we have used relative co-ordinate system to cover the whole desktop screen.

The pre-processed image is sent to the Convolutional Neural Network model. CNN Model analysis the image and identifies the gesture present in the image. Accordingly, the movement of the mouse to be done is decided. The model is trained by the dataset prepared. The dataset has pre-processed images of various hand gestures like navigation, click, zooming and scrolling.

To control mouse virtually python library `autopy` is used. As according to the output given by the CNN Model the mouse movement is performed.

IV. CONCLUSION

In this project we have proposed an efficient way of controlling mouse virtually using real-time camera. We have proposed operations like mouse navigation, clicking, scrolling (up and down), zoom in and out. Our approach is capturing the live feed, subtracting the background and passing it to the CNN model. CNN model gives high accurate result. With using background subtraction, we can also provide good results in complex background. This technique can be implemented in various systems like gaming, detection of sign language and communication. This type of system can also be implemented with emphasis over movement of eyes which can be beneficial for patients suffering from paralysis.

REFERENCES

- [1] Kumar, P. & Verma, J. & Prasad, Shitala, "Hand data glove: A wearable real-time device for human-computer interaction" International Journal of Advanced Science and Technology, 2012.
- [2] X. Zhang and S. Xu, "Research on Image Processing Technology of Computer Vision Algorithm" 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL), 2020, pp. 122-124.
- [3] Dinh-Son Tran, Ngoc-Huynh Ho, Hyung-Jeong Yang, Eu-Tteum Baek, Soo-Hyung Kim and Gueesang Lee "Real-Time Hand Gesture Spotting and Recognition Using RGB-D Camera and 3D Convolutional Neural Network" Appl. Sci. 2020, 10, 722.
- [4] L. D. Singh, P. Das and N. Kar, "A pre-processing algorithm for faster convex hull computation" Confluence 2013: The Next Generation Information Technology Summit (4th International Conference), 2013, pp. 413-418
- [5] T. Palleja et al. "Simple and Robust Implementation of a Relative Virtual Mouse Controlled by Head Movements" Conference on Human System Interactions, 2008, pp. 221-224.
- [6] Ahmed, Muhammad et al. "Survey and Performance Analysis of Deep Learning Based Object Detection in Challenging Environments." Sensors (Basel, Switzerland) vol. 21, 15 5116. 28 Jul. 2021, doi:10.3390/s21155116
- [7] V. Tiwari, C. Pandey, A. Dwivedi and V. Yadav, "Image Classification Using Deep Neural Network" 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), 2020, pp. 730-733.
- [8] Li Wensheng, Deng Chunjian and Lv Yi, "Implementation of virtual mouse based on machine vision" The International Conference on Apperceiving Computing and Intelligence Analysis Proceeding, 2010, pp. 367-371.
- [9] A. Mhetar, B. K. Sriroop, A. G. S. Kavya, R. Nayak, R. Javali and K. V. Suma, "Virtual mouse" International Conference on Circuits, Communication, Control and Computing, 2014, pp. 69-72.
- [10] S. M. S. Shajideen and V. H. Preetha, "Hand Gestures - Virtual Mouse for Human Computer Interaction" International Conference on Smart Systems and Inventive Technology (ICSSIT), 2018, pp. 543-546.
- [11] K. S. Varun, I. Puneeth and T. P. Jacob, "Virtual Mouse Implementation using Open CV" 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019, pp. 435-438.
- [12] S. R. Chowdhury, S. Pathak and M. D. A. Praveena, "Gesture Recognition Based Virtual Mouse and Keyboard" 4th International Conference on Trends in Electronics and Informatics (ICOEI), 2020, pp. 585-589.
- [13] Sai Mahitha G., Revanth B., Geetha G., Sirisha R. "Hand Gesture Recognition to Implement Virtual Mouse Using Open Source Computer Vision Library: Python" Proceedings of International Conference on Advances in Computer Engineering and Communication Systems. Learning and Analytics in Intelligent Systems, vol 20. Springer, Singapore, 2021, pp. 435-446.
- [14] Tran, DS., Ho, NH., Yang, HJ. et al. "Real-time virtual mouse system using RGB-D images and fingertip detection" Multimed Tools Appl 80, 10473-10490, 2021.
- [15] V. V. Reddy, T. Dhyanchand, G. V. Krishna and S. Maheshwaram, "Virtual Mouse Control Using Colored Finger Tips and Hand Gesture Recognition" 2020 IEEE-HYDCON, 2020, pp. 1-5.