JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Mouse Cursor Control Using Hand Gesture Recognition

[1] **Mohammad Naveed Khan,** [2] **Sheikh Auqid Irfan,** [3] **Shah Faisal Ahmad,** [4] **Mr. Khalid Makhdoomi**
[1,2,3]B. E CSE, SSM College of Engineering
[4] Assistant Professor, Department of CSE, SSM College of Engineering

Abstract

In this project, we design a human computer interaction system using hand gestures. Hand gesture recognition (HGR) is a natural way of Human Machine Interaction and has been applied on different areas. Here we explore the use of hand gestures as a means of human-computer interactions for computer applications. Our goal is to make a system that will recognise a few sets of hand gestures and correspondingly realize the robust control of mouse cursor and mouse clicks with a higher accuracy of gesture recognition. More specifically it is the use of a convolution neural network (CNN) to recognize the hand gestures in even the harshest lightning/background conditions using only a simple webcam. We will try to adjust the cursor control to the relative movement of our specific gesture set for mouse movement and different mouse events like clicks to different gestures. The final system will be highly extendable and can be further implemented in advanced technologies like augmented reality (AR), virtual reality (VR) for easy and hassle-free navigation and control without the need for keyboard/mouse for computers and controllers for VR/AR headsets.

Keywords: Human Computer Interaction, Hand Gestures, CNN, Mouse Control.

INTRODUCTION

Computer technology has tremendously grown over the past decade and has become a necessary part of everyday live. The primary computer accessory for human computer interaction (HCI) is the mouse. The mouse is not suitable for HCI in some real-life situations, such as with human robot interaction (HRI). There have been many researches on alternative methods to the computer mouse for HCI (Quam and David 1990; Zhu et al 2005). The most natural and intuitive technique for HCI, that is a viable replacement for the computer mouse is with the use of hand gestures (Yeo et al 2013). This project is therefore aimed at investigating and developing a CC system using hand gestures.

Project Justification

Most laptops today are equipped with webcams, which have recently been used in security applications utilizing face recognition. In order to harness the full potential of a webcam, it can be used for vision-based CC, which would effectively eliminate the need for a computer mouse or mouse pad. The usefulness of a webcam can also be greatly extended to other HCI application such as a sign language database (Starner et al 1998) or motion controller (Rehg and Kanade 1994). Over the past decades there have been significant advancements in HCI technologies for gaming purposes, such as the Microsoft Kinect and Nintendo Wii. These gaming technologies provide a more natural and interactive means of playing videogames. According to Benedetti (2009), motion controls is the future of gaming and it have tremendously boosted the sales of video games, such as the Nintendo Wii which sold over 50 million consoles within a year of its release.

HCI using hand gestures is very intuitive and effective for one-to-one interaction with computers and it provides a Natural User Interface (NUI). There has been extensive research towards novel devices and techniques for cursor control using hand gestures (Yeo et al 2013; Jophin et al 2012). Besides HCI, hand gesture recognition is also used in sign language recognition (Guan et al 2008), which makes hand gesture recognition even more significant.

Project Objectives

- Research existing methods and accessories for cursor control and the suitability of visually based methods for cursor control.
- Investigate tracking algorithms utilized for cursor control.
- Develop and implement a computer application that utilizes alternate methods for cursor control.
- Compare the accuracy and precision of the application with alternative accessories (such as the Microsoft Kinect/Computer Mouse).

System Overview

Our main goal will be to train an image classifier (CNN) that is able to recognise a set of hand gestures and use those predictions to control our mouse cursor. We will set a specific gesture for the 3 basic mouse functions hover, left click and right click. We will feed our model test samples in real-time to make predictions and use those predictions for our actions. Our main goal is to produce a scalable model where the users can generate their custom training data and further increase the number of classifications.

Figure 1 : System Overview

Data Generation

Here we use OpenCV and MediaPipe to generate a boundary box around our hand and generate gesture specific data images for training our model. For our model a total of 3 gestures were trained with training data for each gesture having around 1500 images and the validation data having 450 images.

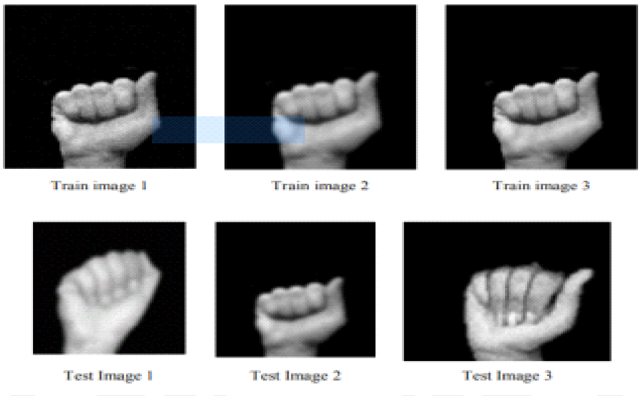

Figure 2: Test Data Sample

Data Augmentation

Since our goal was make a scalable model where we can easily increase the number of gestures our model can classify, we used physical image augmentation on each class of our training data by either flipping, rotating and shifting the images. No augmentation was done on the validation data.

Figure 3: Augmentation Data

Creating a Convolutional Neural Network

We used EfficientNetB0 as our base model with ImageNet weights and max pooling, removing its output layer we defined a functional neural network connecting it to the second last layer of EfficientNetB0 model. We trained the custom layers for classification of our images and saved the weights of the custom layers in Keras file.

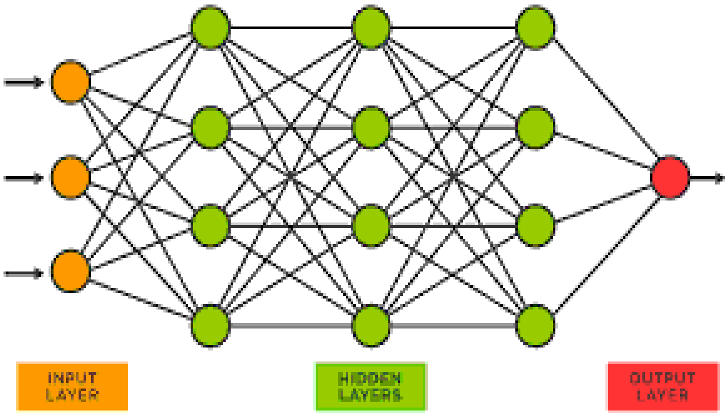

Figure 4: Convolution Neural Network

Gesture Recognition

We will use MediaPipe for hand detection and feed the detected hand after processing it to our trained model for making predictions in real-time. Using our current Hardware, we were able to achieve a detection rate of 17-20 detections per seconds.

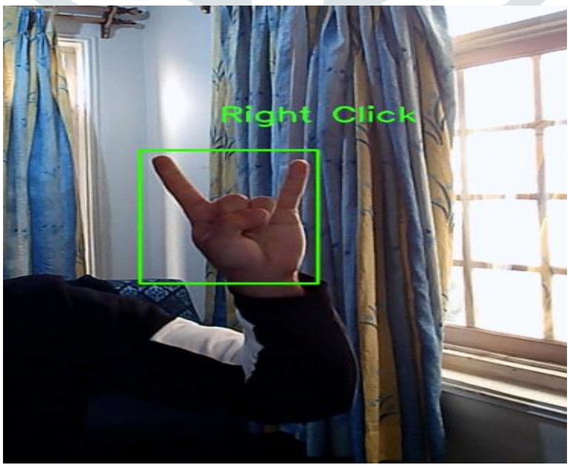

Figure 5: Model Testing on Live Feed

Cursor Control

We define a function for scaling the points of our detected hand on input video frame to that of the resolution of our monitor for controlling the cursor. Also, we will assign specific predictions for the left and right click operations of the mouse.

```
def resolution_scaling(x,y):
    new_x=1920                  #input("Enter Current Monitor Horizontal Resoultion: ")
    new_y=1080                  #input("Enter Current Monitor Vertical Resoultion: ")

    var_x=int((x/1280)*new_x)
    var_y=int((y/720)*new_y)

    return var_x,var_y
```


CONCLUSIONS

We trained a neural network from scratch using self-generated data and were able to achieve a test accuracy of **98.18%**. Our method was able to surpass the usual limits faced by other models that are based on identifying specific color on your hands or using various other methods. Our model was able to extract the useful features from the training data and use that data to make accurate predictions. Using those predictions, we were able to program a simple script for controlling the mouse cursor based on the predictions from our model.

FUTURE SCOPE

There are generally two approaches for hand gesture recognition, which are hardware based (Quam 1990; Zhu et al 2006), where the user must wear a device, and the other is vision based (Shrivastava 2013; Wang and Popović 2009), which uses image processing techniques with inputs from a camera. The proposed system is vision based, which uses image processing techniques and inputs from a computer webcam. Vision based gesture recognition systems are generally broken down into four stages, skin detection, hand contour extraction, hand tracking and gesture recognition. The input frame would be captured from the webcam and the skin region would be detected using skin detection. The hand contour would then be found and used for hand tracking and gesture recognition. Hand tracking would be used to navigate the computer cursor and hand gestures would be used to perform mouse functions such as right click, left click, scroll up and scroll down. The scope of the project would therefore be to design a vision-based CC system, which can perform the mouse function previously stated.

RESULTS

At the end of the training our CNN was able to achieve a **Validation Accuracy** of **99.10%** and a **Test Accuracy** of **98.18%.** Since our data set was very small with the normal 80:20 training to validation data split we were able to achieve a high Validation Accuracy but our models Test Accuracy was not sufficient to make accurate predictions. So, after a lot of trail and error we found the perfect Training to Validation Data split in out use case to be 60:40 respectively. Here are the results obtained with our 60-40 Training-Validation Data split when the model was trained for 30 Epochs:

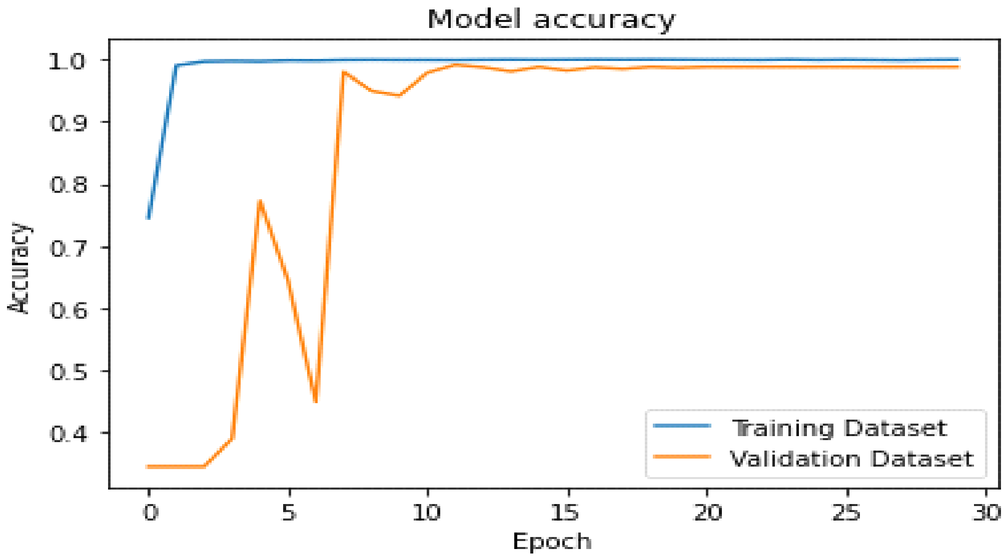

Figure 6: Training to Validation Accuracy

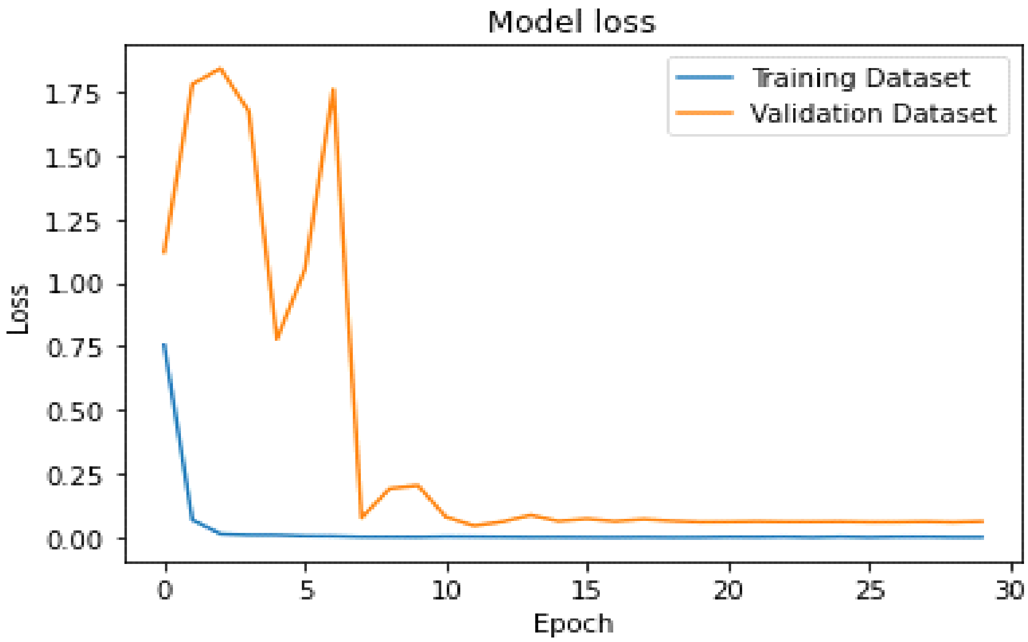

Figure 7: Training to Validation Loss

6.3 Test Data Accuracy of Model

6/6 [============================] – 0s 33ms/step – loss: 0.0610 – accuracy: 0.9818

Test Accuracy: 98.18181991577148%

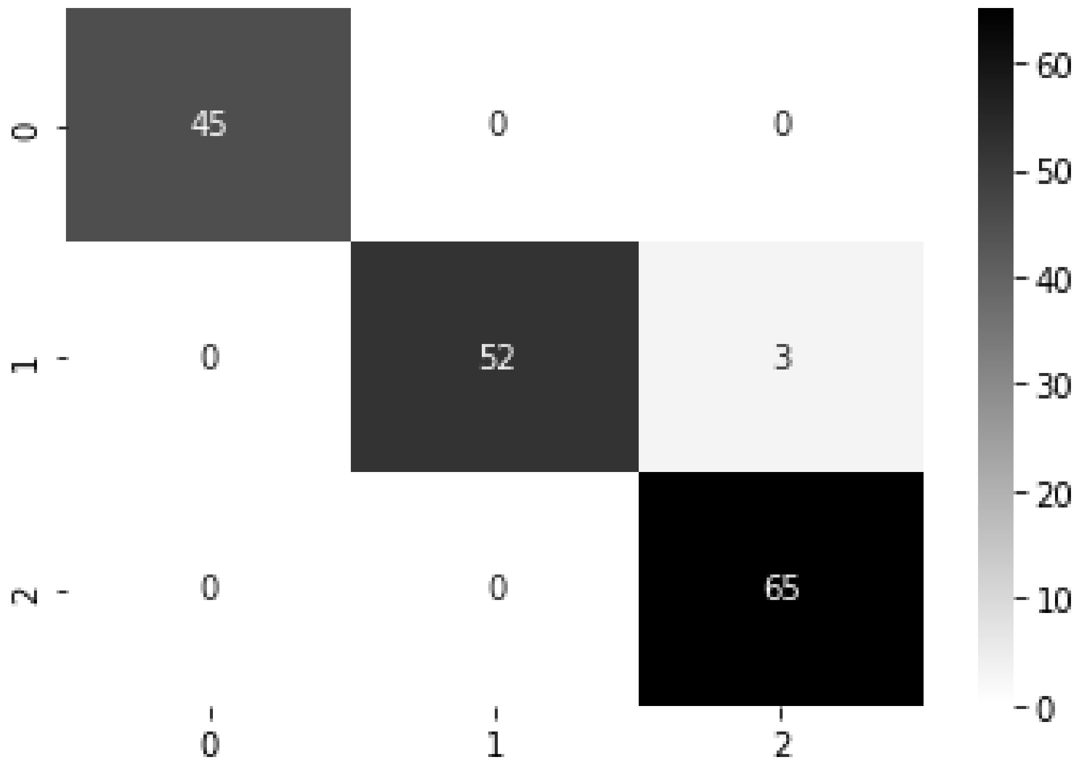

Figure 8: Confusion Matrix Heatmap for our Test Data

REFERENCES

[1] Quam, D. L. (1990, May). Gesture recognition with a dataglove. In *IEEE Conference on Aerospace and Electronics* (pp. 755-760). IEEE.

[2] Starner, T., Weaver, J., & Pentland, A. (1998). Real-time american sign language recognition using desk and wearable computer based video. *IEEE Transactions on pattern analysis and machine intelligence*, *20*(12), 1371-1375.

[3] Rehg, J. M., & Kanade, T. (1994, May). Visual tracking of high dof articulated structures: an application to human hand tracking. In *European conference on computer vision* (pp. 35-46). Springer, Berlin, Heidelberg.

[4] Jophin, S., Sheethal, M. S., Philip, P., & Bhruguram, T. M. (2012). Gesture based interface using motion and image comparison. *International Journal of Advanced Information Technology*, *2*(3), 37.

[5] Zhang, F., Bazarevsky, V., Vakunov, A., Tkachenka, A., Sung, G., Chang, C. L., & Grundmann, M. (2020). Mediapipe hands: On-device real-time hand tracking. arXiv preprint arXiv:2006.10214.

[6] Mahitha, G. S., Revanth, B., Geetha, G., & Sirisha, R. (2021). Hand Gesture Recognition to Implement Virtual Mouse Using Open Source Computer Vision Library: Python. In Proceedings of International Conference on Advances in Computer Engineering and Communication Systems (pp. 435-446). Springer, Singapore.

[7] Huang, Z., Pan, Z., & Lei, B. (2017). Transfer learning with deep convolutional neural network for SAR target classification with limited labelled data. Remote Sensing, 9(9), 907.

[8] Oudah, M., Al-Naji, A., & Chahl, J. (2020). Hand gesture recognition based on computer vision: a review of techniques. journal of Imaging, 6(8), 73.