Gesture-Buddy: Your Virtual Friend

A PROJECT REPORT

Submitted by

Project Contributors	
20BCS1682	Saumyamani Bhardwaz
20BCS1762	Rohan Godha
20BCS2621	Niraj Kumar Yadav
20BCS5455	Pranam Manav Upagade

in partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

Chandigarh University

DEC 2022

BONAFIDE CERTIFICATE

Certified that this project report **"Gesture-Buddy: Your Virtual Friend"** is the bonafide work of the mentioned students, who carried out the project work under my/our supervision.

1. Saumyamani Bhardwaz
2. Rohan Godha
3. Niraj Kumar Yadav
4. Pranam Manav Upagade

Dr. Sandeep Singh Kang

HEAD OF DEPARTMENT

Tushar Semwal(E13016)

SUPERVISOR

Submitted for the project viva-voce examination held on__/ /2022.

INTERNAL EXAMINER

EXTERNAL EXAMINER

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENT

We would like to express our deep and sincere gratitude to Tushar Semwal Sir (Supervisor), Charnpreet Kaur Ma'am (Co-supervisor), panelist as well as our respected Dr. Sandeep Singh Kang Sir (H.O.D.) for giving us the opportunity to do the project and providing valuable guidance throughout this project.

Their dynamism, vision and exquisite efforts have deeply inspired us. They taught us the methodology to carry out the project and to present the project work as clearly as possible. It was a great privilege for us to study and work under their guidance. We owe the completion of my project to our project Mentors for their continuous support and guidance.

Also, we would like to thank our friends and family who helped us a lot in gathering information and guiding us from time to time while making this project. Despite their busy schedules during this pandemic, they helped us in finalizing the project in the limited time frame.

Regards,

Project Team (Gesture-Buddy: Your Virtual Friend)

ABSTRACT

Hand gesture recognition systems have received a lot of attention in recent years due to the variety of applications of and its ability to efficiently interact with machines through human-computer interaction. Our project outlines a new hand gesture recognition system. A major problem for hand gesture recognition systems is facing the challenge of the gesture system. Also check out the recent postural methods and the presented gesture recognition system. A summary of research results on hand gesture methods, databases, and comparisons between key stages of gesture recognition is also presented. Finally, we discussed future application areas, using Open CV, Computer Vision, and Python.

Gesture recognition has been recognized as a natural way for the communication especially for elder or impaired people. Hand gesture recognition is an important research issue in the field of human-computer interaction (HCI), because of its extensive applications in virtual reality, sign language recognition, and computer games. There is a number of algorithms addressing different aspects of the Gesture recognition problem have been proposed. While image-based techniques have been widely studied, it may be affected by lighting conditions, large variations of the hand gesture and textures. Recently, with the developments of new technologies and the large availability of inexpensive depth sensors, real time gesture recognition has been faced by using depth information and avoiding the limitations due to complex background and lighting situations.

This project introduces an enhanced automated model for hand gesture recognition using OpenCV,cvzone, pyautogui and hand-dot-finger points. The proposed model uses both the depth and color information from Hand Tracking Module to detect the hand shape, which ensures the robustness in cluttered environments. The proposed model consists of two major modules, namely, hand detection and gesture recognition. Experiments have been conducted on large dataset to demonstrate the efficiency of the proposed model. The experimental results show an outstanding performance in the terms of accuracy, recall and precision.

CHAPTER 1.
INTRODUCTION

1.1. Client Identification/Need Identification/Identification of relevant Contemporary issue

The essential aim of building hand gesture recognition system is to create a natural interaction between human and computer where the recognized gestures can be used for controlling a robot or conveying meaningful information. How to form the resulted hand gestures to be understood and well interpreted by the computer considered as the problem of gesture interaction. Human computer interaction (HCI) also named Man-Machine Interaction (MMI) refers to the relation between the human and the computer or more precisely the machine, and since the machine is insignificant without suitable utilize by the human.

There are two main characteristics should be deemed when designing a HCI system as mentioned in functionality and usability. System functionality referred to the set of functions or services that the system equips to the users, while system usability referred to the level and scope that the system can operate and perform specific user purposes efficiently. The system that attains a suitable balance between these concepts considered as influential performance and powerful system. Gestures used for communicating between human and machines as well as between people using sign language.

1.2. Identification of Problem

Gestures can be static (posture or certain pose) which require less computational complexity or dynamic (sequence of postures) which are more complex but suitable for real time environments. Different methods have been proposed for acquiring information necessary for recognition gestures system. Some methods used additional hardware devices such as data glove devices and color markers to easily extract comprehensive description of gesture features. Other methods based on the appearance of the hand using the skin color to segment the hand and extract necessary features, these methods considered easy, natural and less cost comparing with methods mentioned before.

1.3. Identification of Tasks

The tasks that we went under included learning about open CV, hand tracker, hand tracking module, opening system, Python, 21 data points of our hands, comparators & operators, different points of our face, different parameters of screen such as height, width, camera distance and frame, masking etc. We also had do a lot of research about this project features and how can we optimize them to create the best set of modules.

1.4. Timeline

In this project, we started as a group of individuals who had a keen interest in learning about open CV series, Python and machine learning and the applications that they can provide us with opportunity to learn the various applications in the real-world scenario. At the initial step we learned about how can we detect an object through the use of computer vision through our laptops and then we started to do some coding and successfully we directed our hands in this through the use of open CV. Then we prepare for better application and studied more. We concluded with creating rock paper scissors game based on AI and comparators in python.

Meanwhile when we were making this project, we thought about how can we make the lives of individuals easier and hance, we tried to create and interact with the mouse and keyboard through the use of only eyes and fingertips.

1.5. Organization of the Report

The further chapters include literature review, analysis and features description which can help us learn more about OpenCV and Hand Detection and Gesture Identification.

CHAPTER 2.
LITERATURE REVIEW/BACKGROUND STUDY

2.6. Timeline of the reported problem

Hand gestures can be used as a natural communication with the computer; with today's technologies, the number of applications that could apply hand gesture recognition is rapidly increasing.
In this section, we will be discussing the most recent applications that were presented in the years 2016 to 2018:

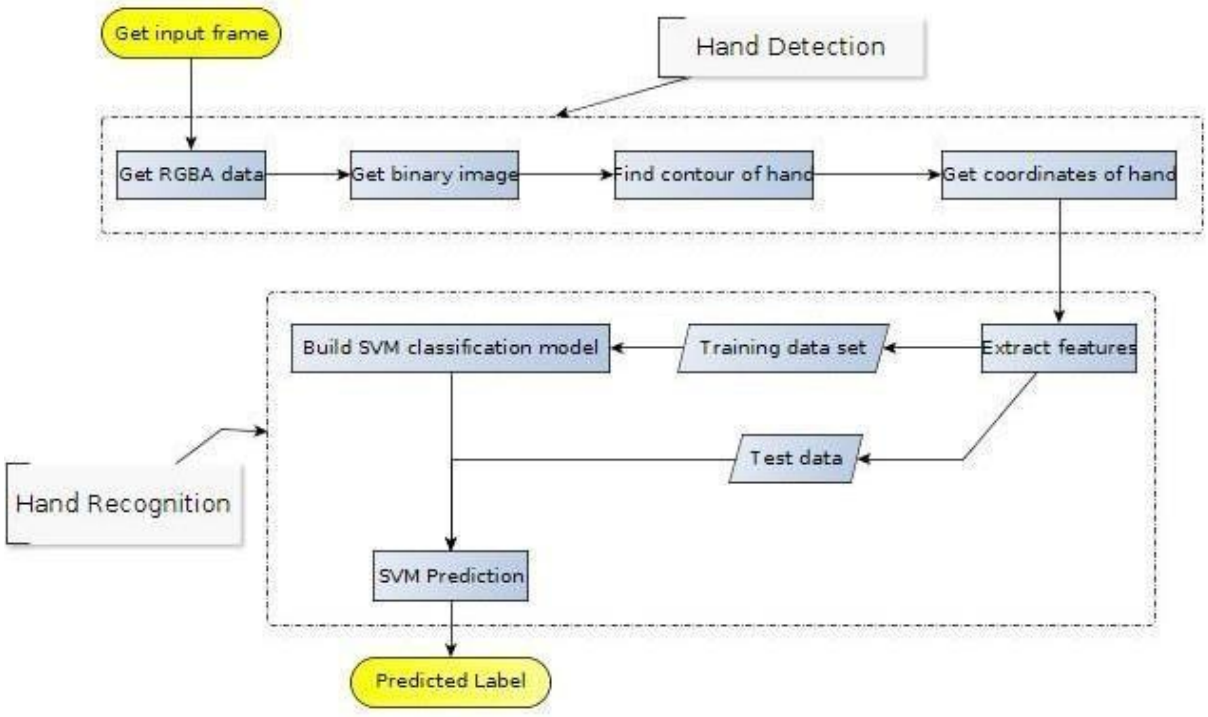
The work of Weiguo et al. (2017), Chenyang, Yingli & Matt (2016), Rasel et al. (2017), Shaun et al. (2017), Deepali & Milind (2016), Nabeel, Rosa & Chan (2017); and Anshal, Heidy & Emmanuel (2017) proposed using hand gesture recognition for American Sign Language (ASL) for people who are hearing impaired, whereas in Shaun et al. (2017) they did not evaluate the letters that are dynamic (like: j and z).

In Nabeel, Rosa & Chan (2017) 36 gestures were studied including 26 ASL alphabets and 10 ASL numbers. In Anshal, Heidy & Emmanuel (2017) two different translation paradigms; English characters (alphabet) and complete words or phrases were proposed.

2.7. Proposed solutions

Gesture recognition system was proposed in Soumya & Muzameel (2017), Mudra is an expressive form of gesture that is mainly used in Indian classical dance form where the gesture is in visual form to connect with the audience. The authors in Zhiwen et al. (2017) presented a real-time hand gesture recognition by using Kinect sensor, to control mouse by user hands for operations such as 'clicking', 'dragging' and 'dropping', and engaged/disengaged gestures. Gesture recognition in Karthik et al. (2017) was used for Bio Robotics, the paper focused on presenting a sensor based human gesture recognition for the Hand Cricket game.

2.8. Bibliometric analysis



1.1 Color-Based Hand Gesture Recognition

2.9. Problem Definition

For hand gesture recognition through SVM, first hand motion is acquired using a web camera, during processing. A bounding box is built to identify the hand during the pre-processing step in order to separate the foreground from the background. This is accomplished by using differences between successive frames.

To extract the characteristics and locate the hand area, KNN requires entering parameters. In order to achieve this, the KNN system adapts four parameters: the ratio of the vertical and horizontal lines, the rectangle situation that contains the median point and covers the hand shape, and a parameter providing details about the horizontal distribution surrounding the rows' mean within the entire hand area, and the standard deviation.

2.10. Goals/Objectives

The list of objectives that will need to be achieve for this project:

(1) to establish a complete system for detecting, recognizing and interpreting hand gesture recognition through computer vision using Python and OpenCV.

(2) to create the numbers and sign languages of hand gesture shown in the system that will meets the name of the project.

- Narrow, specific statements about what is to be learned and performed
- Precise intentions
- Tangible
- Concrete
- Can be validated or measure

CHAPTER 3.

DESIGN FLOW/PROCESS

3.1. Evaluation & Selection of Specifications/Features

The features that we are mainly focusing upon are:

1. Multi Hand Identifier
2. Hand Sign Detection
3. Rock Paper Scissors
4. AI Mouse
5. Eye Mouse
6. Rectangle Screen Scroller

3.2. Design Constraints

Designing of the hand gesture recognition is one of the complicated job that involves two major problem.

i. Firstly is the detection of hand. User hand is detected by using webcam in real-time video. The problem would be the unstable brightness, noise, poor resolution and contrast. With various information of image like color, hand posture and shape based (shape of hand) in a real time would affect the recognition of gestures.

ii. Second problem is to create the sign that is suitable to be used for one hand in a time. The extraction of hand need to be followed to determine each number and sign used. The extraction of hand involves the contour and convexity defects. Convexity defect gives an issue on how to calculate the depth of the defects. Some of the defects have far greater depth than others so to identify the depth would include some equations.

3.3. Analysis and Feature finalization subject to constraints

The detected hand in the video are recognized to identify the gestures. At this stage, the process involves are the segmentation and edge detection. For this project, the prototype itself has its own limitation in terms of hardware implementation. First of all, the video

capturing device used for this project can only configure up to 640 by 480 pixel resolutions. This might be the disadvantage of getting the hardware to properly work for detecting the hand. To solve this, the web camera must be in a fixed position up to 30 inches (75 cm) to help capture the video more efficiently. Another limitation to this project is the variation in plane and pose. Variations can be classified as rotation, translation and scaling of camera pose. The direction of the web camera is important to obtain a good video without any shadowing effects and the intensity of light in a room is enough. By selecting the correct light intensity and direction, the shadows can be reduced to a minimum.

3.4. Design Flow

1. A system based on KL Transforms can be proposed to identify different hand gestures. Five steps are to be adopted: filtering the skin, cropping the palm, detecting the edges, extracting the features, and classification. Primarily the hand can be detected by filtering the skin, and cropping the palm was executed to extract out only the palm portion. In order to extract the outline images of palm, the extracted image are processed by performing the Canny Edge Detection technique. After palm extraction, the hand features were extracted employing K-L Transform technique, eventually the input gesture can be recognized using proper classifier. The system was tested for 10 different hand gestures, and the obtained recognizing rate was 96%.

2. Technique or design 2 is about a design process in which researchers measured the distance from the hand edges to the centroid of the hand, along a number of radials spaced equally around a circle. This informed that the general “shape” of the gesture.

3. Moreover, another seven different algorithms for hand-feature extraction: These algorithms aimed to recognize and furthermore classify static hand gestures. Gesture vocabulary was defined, with 10 gestures, and videos from 20 persons performing the gestures for hand feature extraction were recorded. The study aims to test the robustness of all algorithms individually. Each algorithm is applied to scale, translation then rotation invariance. The analysis of both data and obtained results indicated further pre-processing

on the video frames is a must in order to reduce the number of variant feature values acquired for the same hand posture. The depth video images acquired with the Kinect had low resolution with some noise, so it was inferred that some impreciseness on data recordings results from those difficulties, leading to more complicated class learning. It was found that the radial signature and the centroid distance were the best shape descriptors discussed in this paper in terms of robustness and computation complexity.

4. Gesture recognition is an active research field in Human-Computer Interaction technology. It has many applications in virtual environment control and sign language translation, robot control, or music creation. In this machine learning project on Hand Gesture Recognition, we are going to make a real-time Hand Gesture Recognizer using the MediaPipe framework and Tensorflow in OpenCV and Python. OpenCV is a real-time Computer vision and image-processing framework built on C/C++. But we'll use it on python via the OpenCV-python package. MediaPipe comes with some pre-trained ML solutions such as face detection, pose estimation, hand recognition, object detection, etc. We'll first use MediaPipe to recognize the hand and the hand key points. MediaPipe returns a total of 21 key points for each detected hand. These key points will be fed into a pre-trained gesture recognizer network to recognize the hand pose.

   It works on a very basic algorithm:

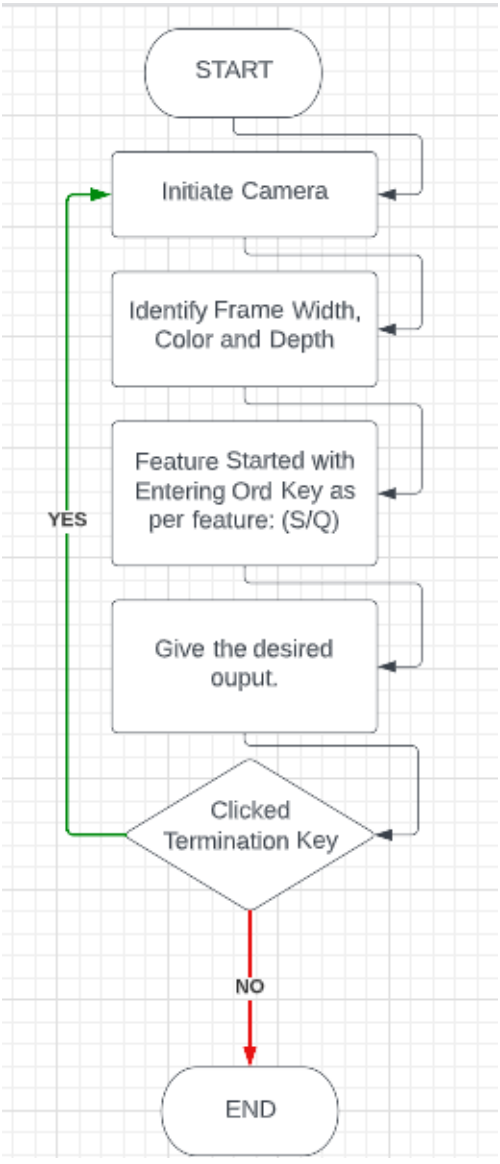
   a) Import necessary packages.
   b) Initialize models.
   c) Read frames from a webcam.
   d) Detect hand keypoints.
   e) Recognize hand gestures

3.5. Design selection

The first step in any hand processing system is to detect and locate the hand in the real-time video from the webcam. The detection of hand is challenging because of variation in pose, orientation, location and scale. Also, different intensity of light in the room adds to the variability. In the process of detection of hand, according to Mohamed [1], hand gesture recognition generally involves multiple levels such as image acquisition,

pre-processing, feature extraction and gesture recognition. Image acquisition involve capturing image in the video frame by frame using a webcam. The captured images go through the image pre-processing process which involves color filtering, smoothing and thresholding. Feature extraction is a method that extracting features of the hand image such as hand contours while gesture recognition is a method to recognize hand gesture by extracting the features.

3.6. Implementation plan/methodology



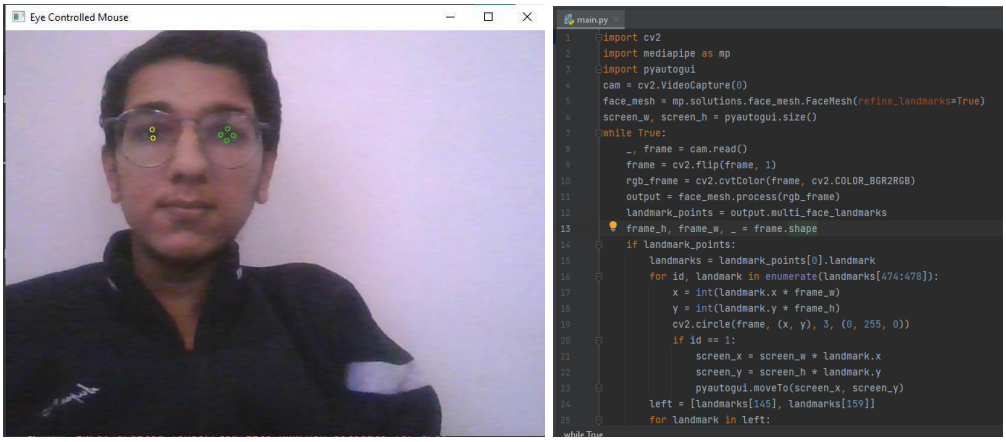
3.1 Flowchart of the Project

CHAPTER 4.

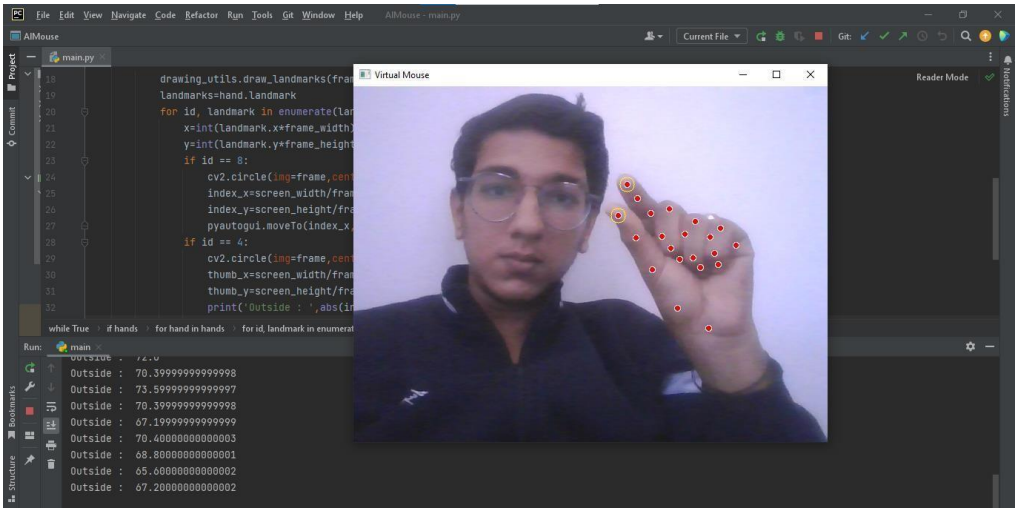
RESULTS ANALYSIS AND VALIDATION

4.1. Implementation of solution

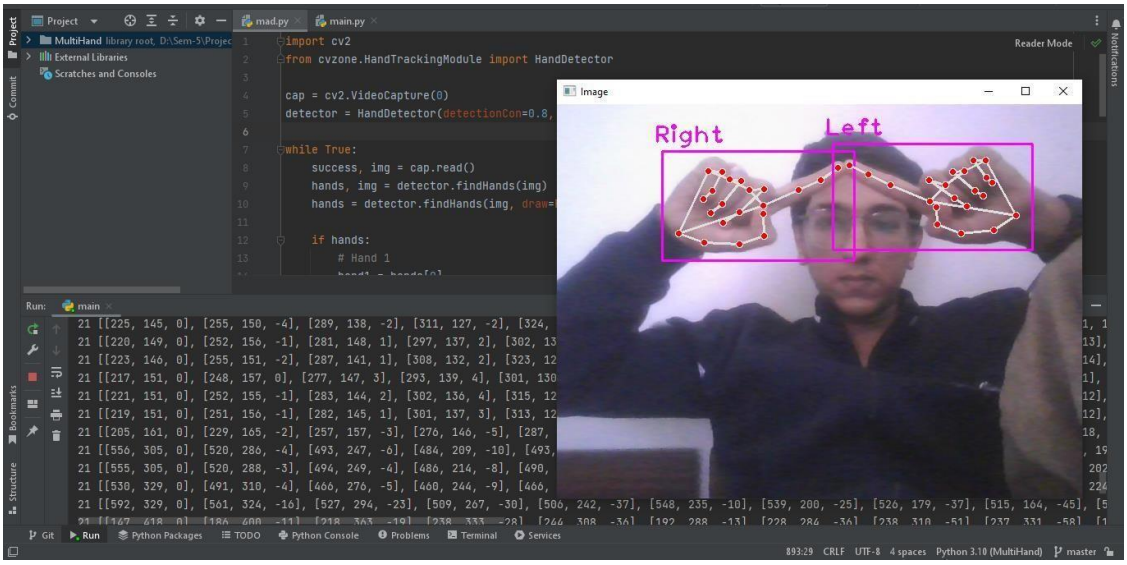
1. Eye Controlled Mouse



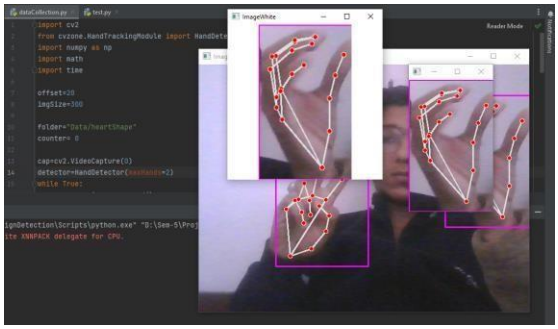
2. AI Finger Controlled Mouse



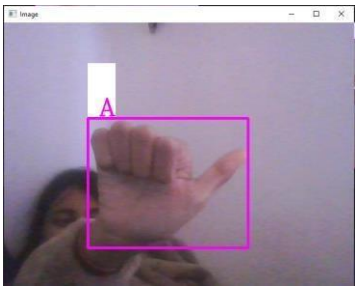
3. MultiHand



4. Hand Sign Detection

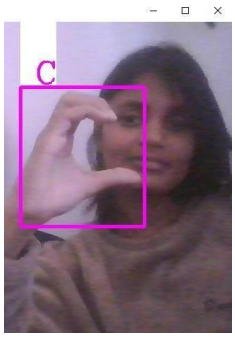


A:

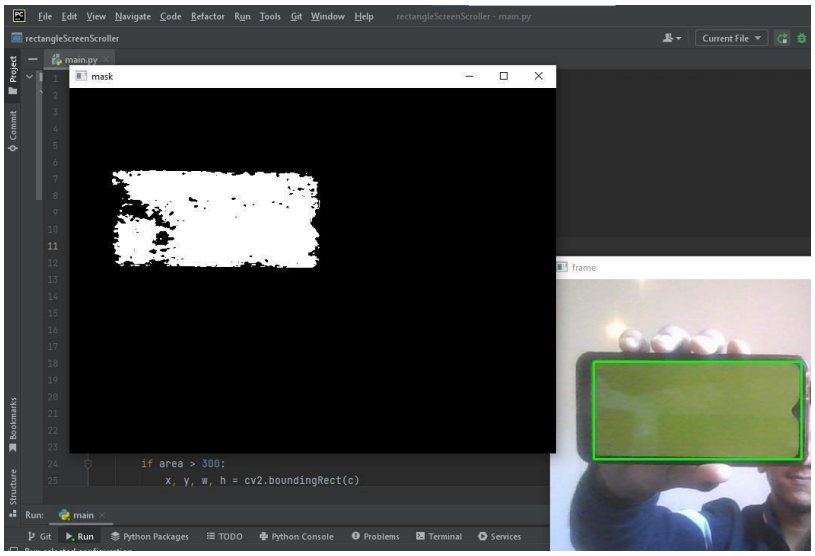


B:

C:



5. Rectangle Screen Scroller



6. Rock Paper Scissors

CHAPTER 5.
CONCLUSION AND FUTURE WORK

5.1. Conclusion

Gesture recognition has several advantages, including increased safety and the convenience of using deliberate gestures to control any device functions rather than a potentially confusing menu system. Gesture recognition is a promising technology, but it only represents a small portion of the huge capacity that ML algorithms have in regard to voice commands, eye movements, and body movements. In the coming time as technological advancement happens, we might see more exposure in the field of gesture recognition.

5.2. Future work

Despite the high performance of some of the current methods discussed in this research, hand gesture recognition is still an evolving topic that needs more experiments. The hand gesture recognition method also needs to be extended to cover all areas of information technology and artificial intelligence, such as tablets, smartphones, gaming consoles, smart televisions, laptops, and desktops (Hexa, 2017).

It is no doubt that hand gesture recognition is capable of enabling natural communication and intelligence into applications that humans use every day. Hand gesture recognition is employing the principle of perceptual computing and changing the methods of human computer interaction (HCI) making them less complex and more enjoyable.

Applications such as home control systems, healthcare systems, gaming technologies, automobiles, televisions, home automations, and robotics are expected to be able to use hand gesture recognition to represent the communication between the user and the devices (Hexa, 2017). Furthermore, some of the applications are very sensitive and in need of having a high recognition accuracy almost close to 100% to be able to use them without causing any damage or danger to human lives; such as applications of the health field, the transportation field, and the flight operation field.

REFERENCES

[1] S. Mitra and A. Tinku, "Gesture recognition: A survey", IEEE Transactions on Systems Man and Cybernetics Part C (Applications and Reviews), vol. 37, no. 3, pp. 311-324, 2007.

[2] A. Erol, G. Bebis, M. Nicolescu, R. D. Boyle and X. Twombly, "Vision-based hand pose estimation: A review", Comput. Vision Image Understand, vol. 108, no. 1, pp. 52-73, Oct. 2007.

[3] J. P. Wachs, M. Kolsch, H. Stern and Y. Edan, "Vision-based hand-gesture applications", Commun. ACM, vol. 54, no. 2, pp. 60-71, Feb. 2011.

[4] I. Dejmal and M. Zacksenhouse, "Coordinative Structure of Manipulative Hand-Movements Facilitates Their Recognition", IEEE. Trans. Biomed. Eng., vol. 53, no. 12, pp. 2455-2463, Nov. 2006.

[5] P. G. Kry and D. K. Pai, "Interaction capture and synthesis", ACM Trans. Graph., vol. 25, no. 3, pp. 872-880, Jul. 2006.

[6] M. H. Yang, N. Ahuja and M. Tabb, "Extraction of 2D Motion Trajectories and Its Application to Hand Gesture Recognition", IEEE Trans. Pattern Anal. Mach. Intell.,

S. Mitra and A. Tinku, "Gesture recognition: A survey", IEEE Transactions on Systems Man and Cybernetics Part C (Applications and Reviews), vol. 37, no. 3, pp. 311-324, 2007.

[7] A. Erol, G. Bebis, M. Nicolescu, R. D. Boyle and X. Twombly, "Vision-based hand pose estimation: A review", Comput. Vision Image Understand, vol. 108, no. 1, pp. 52-73, Oct. 2007.

[8] J. P. Wachs, M. Kolsch, H. Stern and Y. Edan, "Vision-based hand-gesture applications", Commun. ACM, vol. 54, no. 2, pp. 60-71, Feb. 2011.

[9] I. Dejmal and M. Zacksenhouse, "Coordinative Structure of Manipulative Hand-Movements Facilitates Their Recognition", IEEE. Trans. Biomed. Eng., vol. 53, no. 12, pp. 2455-2463, Nov. 2006.

[10] P. G. Kry and D. K. Pai, "Interaction capture and synthesis", ACM Trans. Graph., vol. 25, no. 3, pp. 872-880, Jul. 2006.
[11] M. H. Yang, N. Ahuja and M. Tabb, "Extraction of 2D Motion Trajectories and Its Application to Hand Gesture Recognition", IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 8, pp. 1061-1074, Nov. 2002.
[12] Y. Wu, J. Lin and T. S. Huang, "Analyzing and Capturing Articulated Hand Motion in Image Sequences", IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 12, pp. 1910-1922, Dec. 2005.
[13] C. W. Ng and S. Ranganath, Real-time gesture recognition system and application. Image Vis. Comput, vol. 20, no. 13, pp. 993-1007, 2012.
[14] K. Abe, H. Saito and S. Ozawa, "Virtual 3-D interface system via hand motion.recognition from two cameras", IEEE Trans. Syst. Man Cybern. A, vol. 32, no. 4, pp. 536-540, 2012.
[15] J. Suarez and R. R. Murphy, "Hand Gesture Recognition with Depth Images: A Review", Proc. RO-MAN, pp. 411-417, 2012.
[16] Y. R. Wang, W. H. Lin and L. Yang, "A novel real time hand detection based on skin-color", Proc. ISCE, pp. 141-142, 2013.
[17] M. Kolsch and M. Turk, "Robust Hand Detection", Proc. FG, pp. 614-619, 2004.
[18] B. Stenger, P. R. S. Mendonca and R. Cipolla, "Model-Based 3D Tracking of an Articulated Hand", Proc. CVPR, pp. 310-315, 2001.
[19] Z. Ren, J. Yuan, J. Meng and Z. Zhang, "Robust Part-Based Hand Gesture Recognition Using Kinect Sensor", IEEE Trans. on Multimedia, vol. 15, no. 5, pp. 1110-1120, Aug. 2013.
[20] R. P. Mihail, N. Jacobs and J. Goldsmith, "Real Time Gesture Recognition With 2 Kinect Sensors", Proc. IPCV, pp. 1-7, 2012.