Hand Gesture Based System Control

A Project Report

Submitted to the APJ Abdul Kalam Technological University

in partial fulfillment of requirements for the award of degree

Bachelor of Technology

in

Computer Science and Engineering

(Artificial Intelligence and Machine Learning)

by

GOVIND A (SCT20AM033)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SREE CHITRA THIRUNAL COLLEGE OF ENGINEERING

PAPPANAMCODE, THIRUVANANTHAPURAM, -695018

KERALA

JUNE 2023

Declaration

We undersigned hereby declare that the project report **“Gesture Based System Control”**, submitted for partial fulfilment of the requirements for the award of the degree of Bachelor of Technology of the APJ Abdul Kalam Technological University, Kerala is a bonafide work done by us under the supervision of **Smt. Binu Rajan M R**. This submission represents our ideas in our own words and where ideas or words of others have been included, we have adequately and accurately cited and referenced the original sources. We also declare that we have adhered to the ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in my submission. We understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed as the basis for the award of any degree, diploma or similar title of any other University.

Thiruvananthapuram

03-07-2023

GOVIND A

DEPT. OF COMPUTER SCIENCE AND ENGINEERING

SREE CHITRA THIRUNAL COLLEGE OF ENGINEERING

TRIVANDRUM

2022 - 23

CERTIFICATE

This is to certify that the report entitled **"Hand Gesture Based System Control"** submitted by **Govind A (SCT20AM033)** to the APJ Abdul Kalam Technological University in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineering is a bonafide record of the project work carried out by them under my/our guidance and supervision. This report in any form has not been submitted to any other University or Institute for any purpose.

Project Guide **Project Coordinator**

Head of Dept.

Acknowledgement

We are grateful to present this project after completing it successfully. This project would not have been possible without the guidance, assistance and suggestions of many individuals. We would like to express our deep sense of gratitude and indebtedness to each and everyone who has helped us make this project a success. We heartily thank our **Head of Department, Dr Soniya B, Dept. of Computer Science and Engineering, Sree Chitra Thirunal College of Engineering** for her constant encouragement and inspiration in taking up this project. We heartily thank our Project coordinator **Smt. Soja Salim, Assistant Professor, Dept. of Computer Science and Engineering**, for his constant follow-up and advice throughout the course of the Project work. We gratefully thank our Project guide, **Smt. Binu Rajan M R, Assistant Professor, Dept. of Computer Science and Engineering**, for her encouragement and advice throughout the course of the Project work. Special thanks to all the staff members of the Computer Science and Engineering Department for their help and kind cooperation. Lastly, we thank our parents and friends for their encouragement and support given to us in order to finish this precious work.

Abstract

In this swiftly advancing world, the need for physical contact with a traditional keyboard or touchscreen is becoming rapidly irrelevant. To overcome this limitation, a new system can be introduced that will minimize the reliance on these conventional devices by employing human-computer interface. This project thereby aims to develop a gesture-based system control using Mediapipe, a powerful machine learning framework that provides various modules for processing multimedia data, including image and video data. This enhances the overall user experience as it enables a broader range of users to access and control technology thereby reducing the need for physical contact with a keyboard or other systems. This system will permit users to control various applications and devices by showing hand gestures in front of a camera, which will then be captured and processed by the Mediapipe framework. The system will recognize specific gestures and translate them into commands which will be sent to the corresponding applications or devices. The project will involve the development of a user interface for selecting the applications and devices to be controlled, as well as the design of the gesture recognition algorithm using the Mediapipe framework. The final product will provide a hands-free and intuitive interface for controlling various systems, making it useful in various scenarios, such as home automation, multimedia playback, and gaming.

Contents

List of Figures

Chapter 1

INTRODUCTION

Hand gestures have always played a significant role in human communication, ranging from basic non-verbal cues to intricate sign languages and cultural gestures. They serve as a powerful means of expression and connection between individuals.

In recent years, hand gesture technology has emerged as a groundbreaking form of interaction, allowing users to control devices through the natural movement of their hands and fingers. This technology holds immense potential to transform the way we engage with our smartphones, tablets, smartwatches, and other electronic devices, providing a more intuitive and seamless user experience.

By harnessing the power of hand gestures, users can navigate through menus, scroll through content, zoom in and out, and perform various commands without the need for physical touch or traditional input methods. This advancement opens up new possibilities for individuals with physical limitations, making technology more accessible and inclusive for all.

Moreover, hand gesture technology is not limited to personal devices; it can be applied in diverse fields such as gaming, virtual reality, healthcare, and industrial automation. Its versatility and adaptability make it a promising tool for enhancing user interfaces and revolutionizing the way we interact with technology.

As hand gesture technology continues to advance, researchers and developers are exploring innovative ways to refine its accuracy, responsiveness, and versatility. With ongoing advancements, we can expect even more sophisticated hand gesture recognition systems that can interpret complex gestures and further expand the boundaries of human-computer interaction.

Chapter 2

LITERATURE REVIEW

A Literature Review is a comprehensive summary of previous research on a topic. Here the existing systems are reviewed and a problem statement is proposed.

2.1 Existing Systems

Paper 1: Recognition of Hand Gestures using Mediapipe Hands

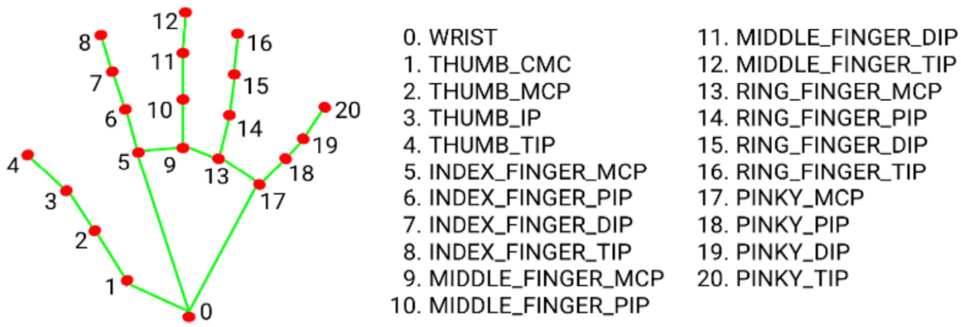

Figure 2.1: Hand gesture features

Paper 2: Virtual Mouse Control Using Hand Gestures

- Hand gestures are used to control the pointer's movement. The center of the palm should be discovered initially in order to locate the hand.
- Determine the hand center form has the benefit of being simple and straightforward to perform. The shortest distance between each point inside the inscribed circle and the contour was measured, with the point with the greatest distance being recorded as the center.
- The radius of the hand was calculated as the distance between the center of the hand and the hand contour.The hand center was determined for each successive frame, and the tip of the finger was known and used for hand tracking utilizing the hand center.

Paper 3: Human Computer Interaction using Virtual User Computer Interaction System

- Development of a virtual user computer interaction system keyboard layout is displayed on the screen and characters can be selected by moving the fingertip.
- A system that can teach a computer to understand the physical world from vision.
- A framework was proposed which detects typing motion with two-hand fingers in the air
- An eye blink-controlled keyboard system was designed where the detection of blinking of the eye was used to select the desired key
- A real-time hand gesture recognition interface called VUCIS (Virtual User Computer Interaction System) is used to operate a mouse, keyboard, and stylus.

Paper 4: Eye-Blink Detection System for Virtual Keyboard

- Real-time data is gathered for tracking and estimation of eye gaze

- Translates eye movement into a message for severely impaired individuals to interact with others using an interface adapted to their needs
- Facial landmark detection process is used to assign a specific index for the eyes
- Identifies eye blinking by calculating the ratio between normal eye blinking and targeted blinking behavior

Paper 5: Finger Recognition and Gesture-Based Virtual Keyboard

- Track the user's hand and perform cursor functions using different hand gestures.
- The system has the potential of being a viable replacement for the computer mouse

One	Forward
Two	Volume up
Three	Volume down
Four	Backward
Five	Close

2.2 Problem Statement

To develop a system that allows users to interact with digital devices, such as computers or smartphones, through human gestures rather than traditional input devices such as a mouse or keyboard.

Additionally, the system should be designed to be user-friendly and intuitive, ensuring that users can easily learn and remember the different hand gestures and their associated actions.

2.3 Proposed Solution

The system incorporates cameras and sensors to detect and track hand and finger movements. Through the utilization of machine learning algorithms and AI, these gestures are recognized and translated into commands that control the device. The system's hardware and software components are designed and developed, followed by a calibration process to ensure accurate detection and tracking of hand movements.

This gesture-based system offers a convenient and intuitive method of device control, enhancing user experience and interaction.

2.4 Objectives

The objectives of this project are:

- To Replace keyboard and mouse with human gesture induced control.
- To create a user friendly interface between humans and the device.
- To provide a lightweight platform to modify system controls in real-time

Chapter 3

METHODOLOGY

The system involves the use of cameras and sensors to detect the movement and position of hands and fingers. The system uses machine learning algorithms and AI to recognize hand gestures and translate them into commands that control the device.

The implementation of the system involves the following steps:

- Design and development of the hardware and software components required for the system.
- Calibration of the cameras and sensors to detect and track hand movements.
- Collection of training data to train the machine learning models.
- Development of the gesture recognition algorithm to translate hand gestures into device commands.
- Integration of the gesture recognition algorithm with the device's software to control the device based on the user's hand gestures.
- The system is tested and optimized to ensure that it operates efficiently and accurately.
- User testing and feedback are also essential to improve the system's performance and usability.

3.1 Functional Requirements

This step involves identifying the specific requirements and objectives of this system. This includes hand detection, gesture recognition, gesture mapping and gesture feedback.

3.1.1 Volume Module

1. **Adjustment and control over volume**

   The system provides users with the ability to fine-tune and customize the audio output to their preference thereby ensuring an optimal listening experience in various environments and for different types of content.

2. **Continuous hand tracking**

   The system employs continuous hand movement detection to facilitate volume control. By constantly monitoring the movements of the user's hand, the system can accurately interpret gestures and translate them into volume adjustments.

3. **Real-time feedback**

   The system provides visual feedback to indicate successful gesture recognition, ensuring that users are aware of their actions being registered by the system.

3.1.2 Screen Brightness Control Module

1. **Screen brightness management**

   Users can easily tailor the screen brightness to match their lighting conditions, individual comfort levels, and aesthetic preferences.

2. **Real-time hand tracking**

   Users can modify the screen brightness using intuitive gestures. The system accurately captures and analyzes the position, orientation, and motion of the user's hands enabling precise recognition and mapping of gestures to specific actions.

3. **Constant feedback**

   The system incorporates visual cues as indicators to provide feedback, thereby assuring users of the successful recognition of their gestures by the system

3.1.3 Virtual Keyboard Module

1. **On-screen keyboard**

   The virtual keyboard serves as a graphical user interface, enabling users to input characters and commands visually through an on-screen representation of a keyboard in a more convenient and user-efficient manner.

2. **Continuous Hand tracking**

   The system monitors user's hand movements for browsing, selection, and character input through gestures. This streamlined interaction enhances the input process and user experience.

3. **Error elimination**

   The system renders effortless error elimination provision with a backspace key, allowing users to easily remove entered characters and correct mistakes for improved accuracy.

4. **Google search results**

   Upon inputting the desired text and initiating the search by clicking the designated button, Google furnishes a selection of links that align with the searched phrase.

3.2 System Design

A design plan that outlines the architecture, modules, and components of the system is created. This is to determine the technologies and tools required for implementation, such as a database for storing student and bus data, a mobile application for user interfaces, and communication systems for real-time tracking.

3.2.1 Application Architecture Design

The proposed system utilizes a camera to detect specific hand gestures, particularly the signs for letters V, A, O, and H. When the V sign is detected, the user is granted access to the volume module, allowing them to control the system's audio output by adjusting the volume through system settings. Similarly, a comparable process is employed for controlling the brightness level and activating the virtual keyboard. In the case of the H sign being recognized, the system exits from the program, providing a straightforward and efficient way to terminate the application. This gesture-based control system offers a convenient and intuitive means of interacting with the system's functionalities.

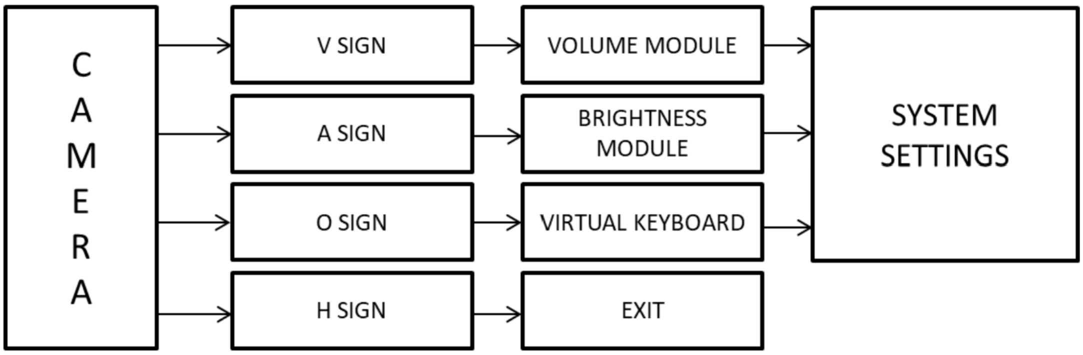

Figure 3.1: Application Architecture Diagram

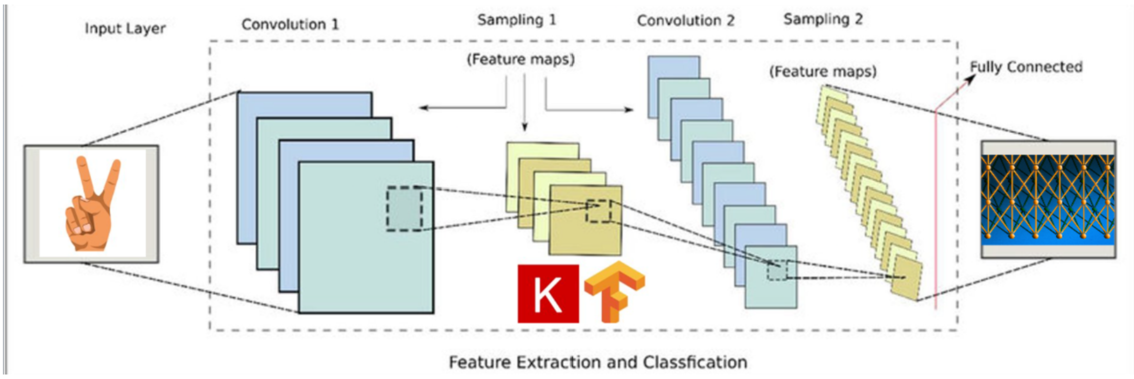

Figure 3.2: CNN Architecture

3.2.2 Use Case View

Actors:

1. **User:** The individual interacting with the system using hand gestures.
2. **System Control Window**: User interface for recognizing gestures and controlling the system.

Use Cases:

1. **Home Screen:**

   Description: This use case represents the window that opens up when the system is initiated. Main Functionality: Provides access to other modules and functionalities within the system. Associated Actors: User

2. **Brightness Control:**

   Description: This use case enables users to control the brightness of the PC using hand gestures. Main Functionality: Allows users to adjust the screen brightness by performing specific hand gestures. Associated Actors: User

3. **Volume Control:**

   Description: This use case allows users to control the volume of the PC using hand gestures. Main Functionality: Enables users to adjust the audio volume by utilizing hand gestures. Associated Actors: User

4. **Virtual Keyboard:**

   Description: This use case represents the window displaying a virtual keyboard, enabling users to type letters using hand gestures. Main Functionality: Displays a virtual keyboard on the screen and recognizes finger gestures to input letters. Associated Actors: User

5. **Exit:**

   Description: This use case allows the user to exit the system. Main Functionality: Terminates the system and closes all associated windows. Associated Actors: User

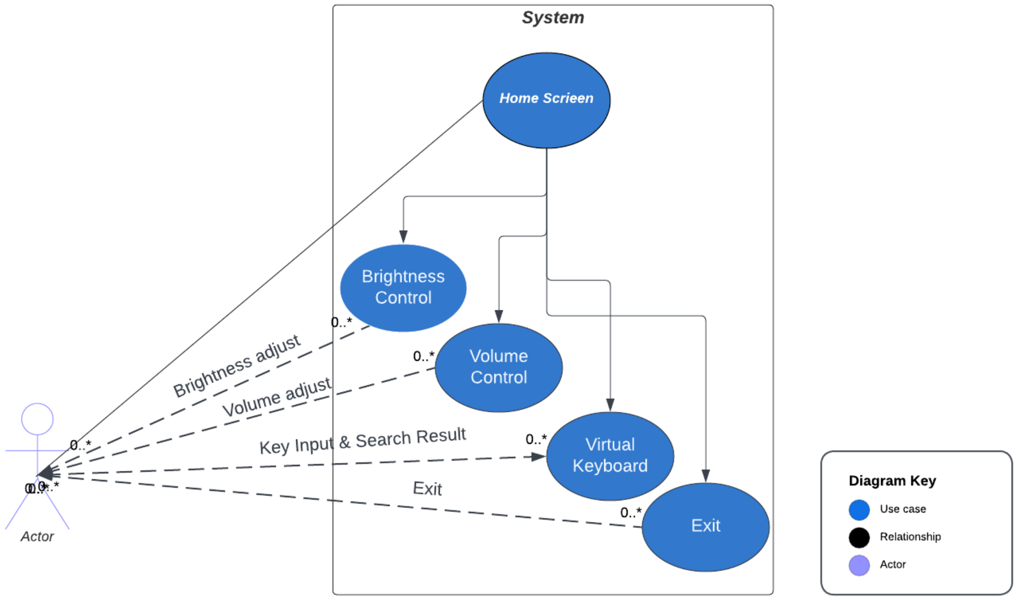

Figure 3.3: Use Case Diagram

3.3 Technology Stack

We have used VSCode to emulate and runvarious built in library function in python and have used Google Colab with tesla t4 GPU to train CNN model using TensorFlow-Keras Framework

3.3.1 OpenCV

Open Source Computer Vision Library is used to capture the video input from the camera, analyze hand gestures, and track hand movements in real-time. It is used here to input image dataset for cnn and real time camera input for the recognition module.

3.3.2 Mediapipe

Mediapipe offers a wide range of pre-built models and tools for tasks like hand tracking, pose estimation, and gesture recognition. MediaPipe can be used to process each frame in the video data, detect and track hand gestures, and extract relevant information for controlling the system.Here we have used hand landmark 1 and 5 to calculate the Euclidean distance between the fingers.

3.3.3 Tensorflow(CNN)

TensorFlow provides a robust and efficient framework for developing CNN models, while Keras offers a user-friendly model construction. Epoch of 65 along with a learning rate of 0.001 with a testsize of 0.2.The output of the model is saved in .h5 format along with class labels.

3.3.4 Tkinter

By integrating Tkinter with gesture recognition algorithms and sensors, developers can create interactive graphical interfaces that respond to gestures and enable system control. Tkinter provides a versatile and intuitive framework for building user interfaces, making it suitable for implementing gesture-based control systems with ease and flexibility. With Tkinter, a homepage GUI was constructed to switch to corresponding functionalities.

3.3.5 Subprocess

This library is used to integrate and run multiple python scripts in a single program. Here we can avoid the over complexity of code by constructing a stream of process.

3.3.6 Googlesearch

Google search library in python helps us to toggle google search directly from the computer vision window. This helps user to experience a handfree search experience.

3.3.7 Pynput

Pynput is a controller library interface in python that helps user to emulate keyboard functionalities without using the hard keyboard. The function we use hear includes Alphabets , Numbers, Backspace and Search Button

Chapter 4

IMPLEMENTATION

Implementation is the process of putting a project plan into action to produce the deliverables. The implementation plan defines how this project is executed. The implementation plan provides a roadmap for the effective execution of the project in order to achieve desired outcomes.

4.1 Modules used

Camera

Using Computer Vision Library

Volume

Using AudioUtil Library

Brightness

Using sbc Library

Virtual Keyboard

Using pynput Library

Google Search

Using googlesearch Library

GUI

Using Tkinter GUI Library

CNN

Using Tensorflow Library

4.2 Model Building

4.2.1 Model Features

MODEL : CNN

EPOCHS : 75

LEARNING RATE : 0.001

CLASSES : 5

TEST : 20

DATASET SIZE : 4000 images

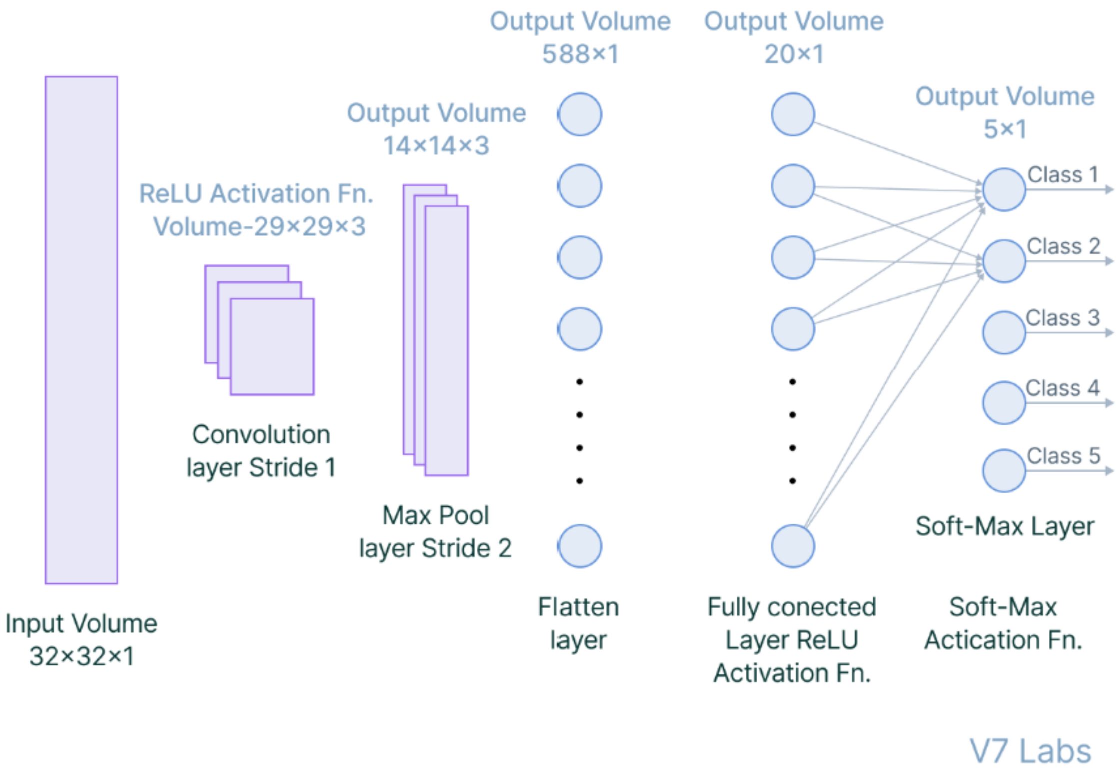

Figure 4.1: Model Features

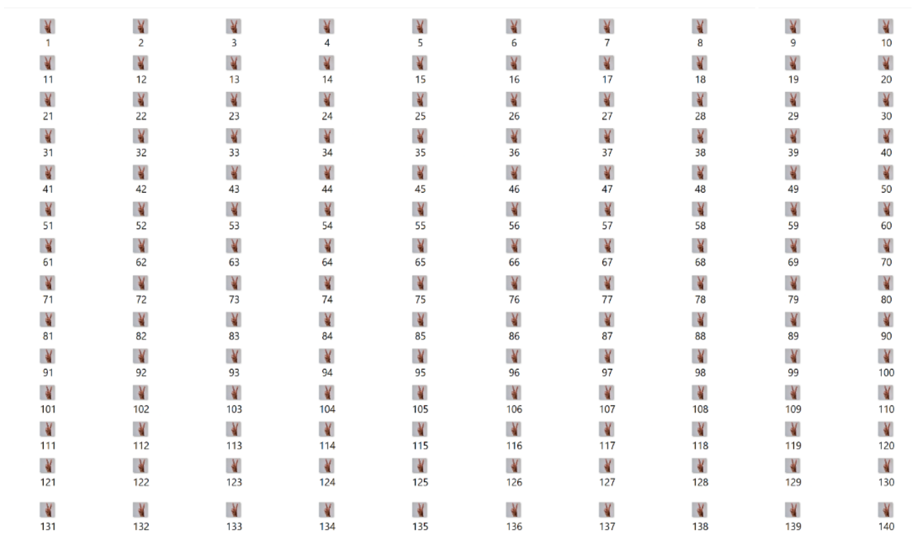

Figure 4.2: Raw Data

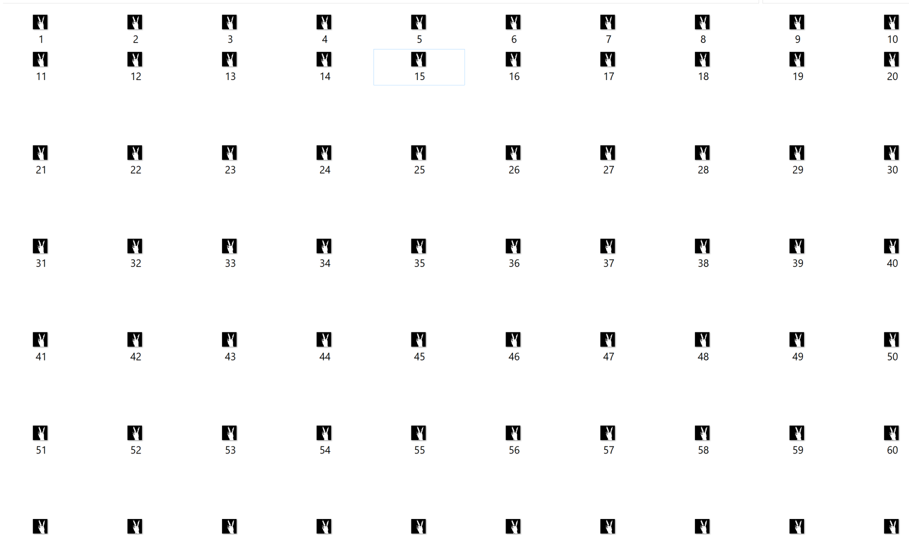

Figure 4.3: Preprocessed Data

Figure 4.4: Noisy Data

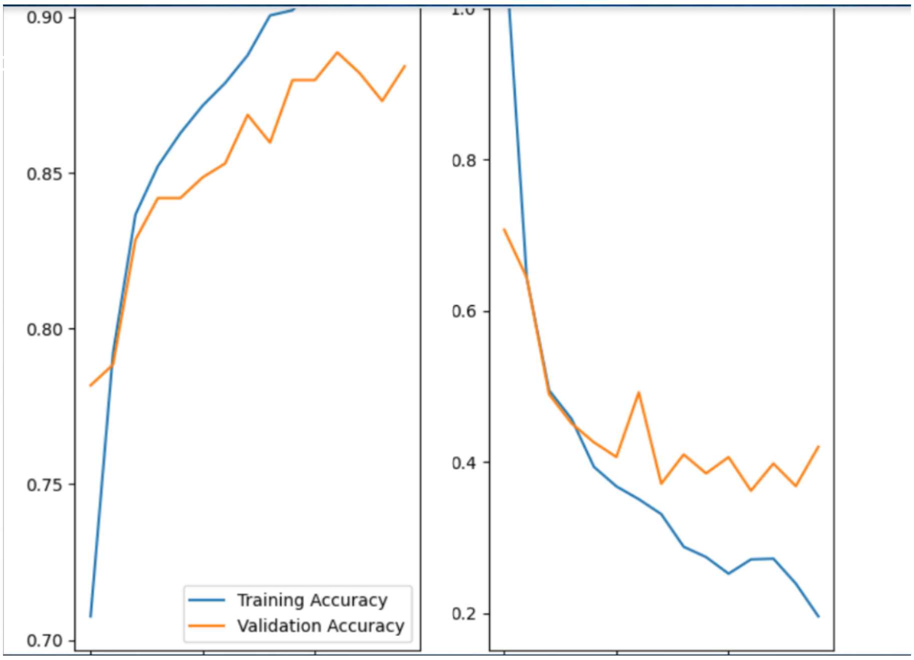

Figure 4.5: Model Accuracy

Chapter 5

TESTING

Testing is a crucial process aimed at identifying any deviations from the expected functionality of a system. The success or failure of a project is determined by whether or not there are any deviations from the intended behavior of the system.

5.1 Unit Testing

In unit testing, individual units are determined if they are fit for use. A unit is the smallest testable part of an application. In procedural programming, a unit may be an individual function or procedure. Unit tests are created by programmers or occasionally by white box testers.

5.2 Integration Testing

Integration testing is the activity of finding faults when testing the individual components together. Structural testing is the culmination of integration testing involving all the components of the system. This testing strategy ensures that software and subsystems work together as a whole. It tests the interface of all modules to make sure that the modules behave properly when integrated together.

5.3 Functional Testing

Functional testing is a type of testing that seeks to establish whether each application feature works as per the software requirements. Each function is compared to the corresponding requirement to ascertain whether its output is consistent with the end user's expectations. Functional testing is a quality assurance (QA) process and a type of black-box testing that bases its test cases on the specifications of the software component under test.

5.4 Load Testing

Load testing generally refers to the practice of modelling the expected usage of a software program by simulating multiple users accessing the program concurrently. It is the process of putting demand on a system and measuring its response. Load testing examines how the system behaves during normal and high loads and determines if a system, piece of software, or computing device can handle high loads given a high demand of end-users. This tool is typically applied when a software development project nears completion.

Chapter 6

RESULT

The result provides a comprehensive overview of the system developed for controlling various devices using hand gestures. The system has successfully implemented features like screen brightness control, volume control and usage of virtual keyboard.

Home Page

The home page of the system comprises of four gestures namely signs of V,A,O and H for volume control, screen brightness control, virtual keyboard and exit respectively.

Figure 6.1

Modules

The section presents the output obtained when selecting one of the three available options, as shown in the provided figures. The figures visually depict the outcomes associated with each option, providing a clear representation of the obtained results. This analysis helps in understanding and interpreting the consequences of selecting a specific option within the given context.

Volume Control

In the provided figure, it is evident that both the volume in the taskbar and the display box on the screen undergo changes simultaneously. The figure visually demonstrates the simultaneous observation of these two elements, highlighting their synchronized behavior. This observation allows for a clear understanding of the relationship and interaction between the volume control and the on-screen display.

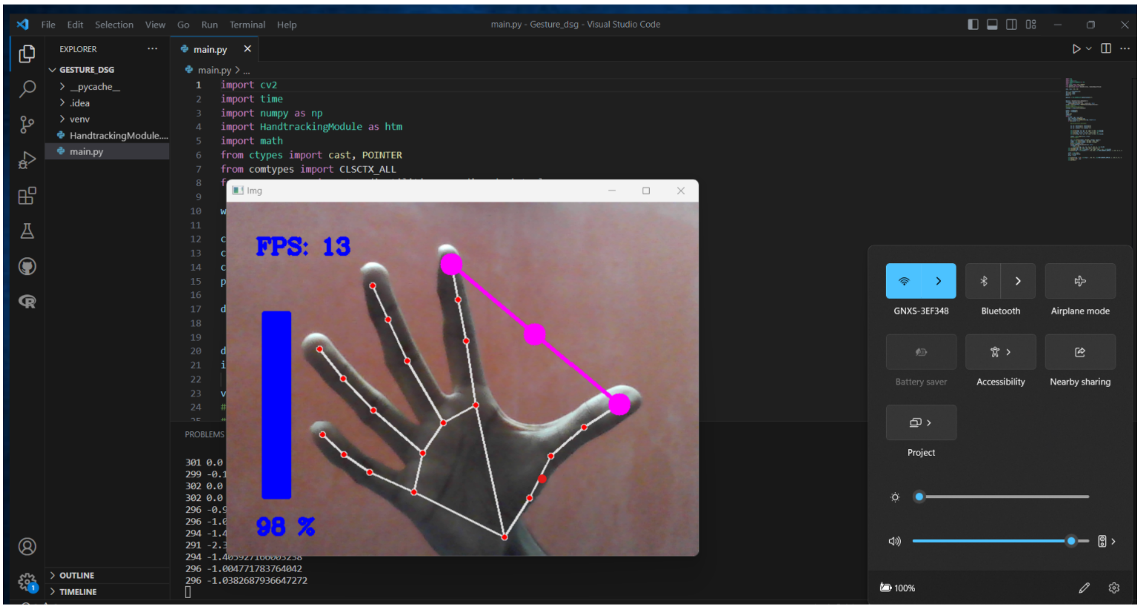

Figure 6.2: Volume Control

Screen Brightness Control

By measuring the relative distance between the index and thumb finger, it is possible to adjust the brightness of the screen. The figure illustrates this relationship, showcasing the direct impact of the measured distance on the screen brightness. This correlation is also reflected in the corresponding change observed in the taskbar, providing a visual confirmation of the brightness adjustment.

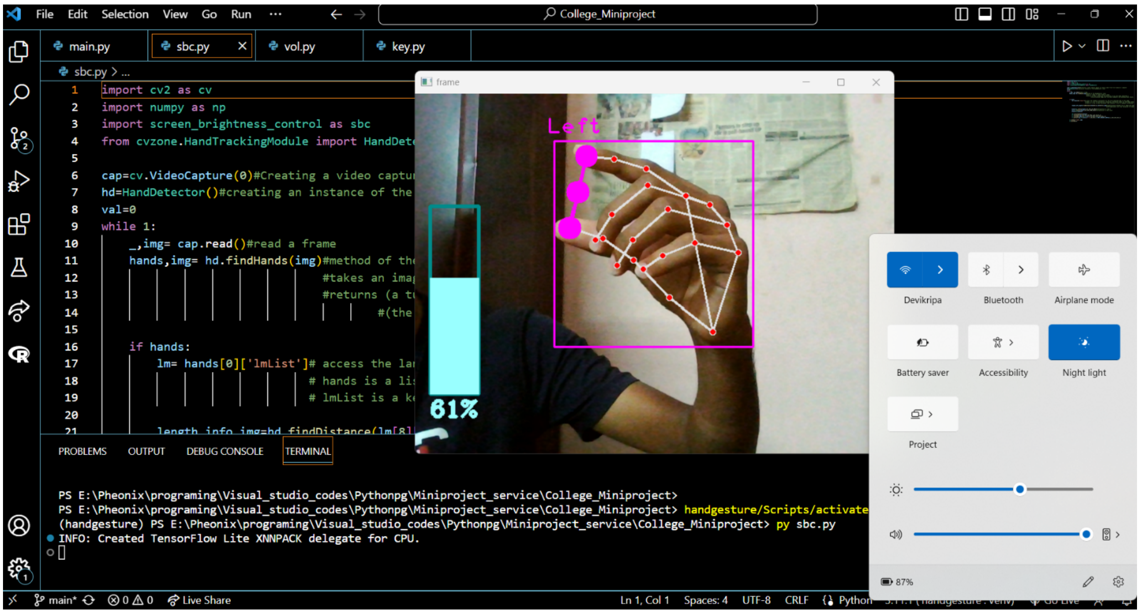

Figure 6.3: Screen Brightness Control

Virtual Keyboard

The screen displays a virtual keyboard that can be navigated by moving a finger. To select a letter, two fingers are used. Mouse control is simulated by using a finger. Upon entering the desired phrase and pressing enter, the Google Search module generates URL links related to the topic.

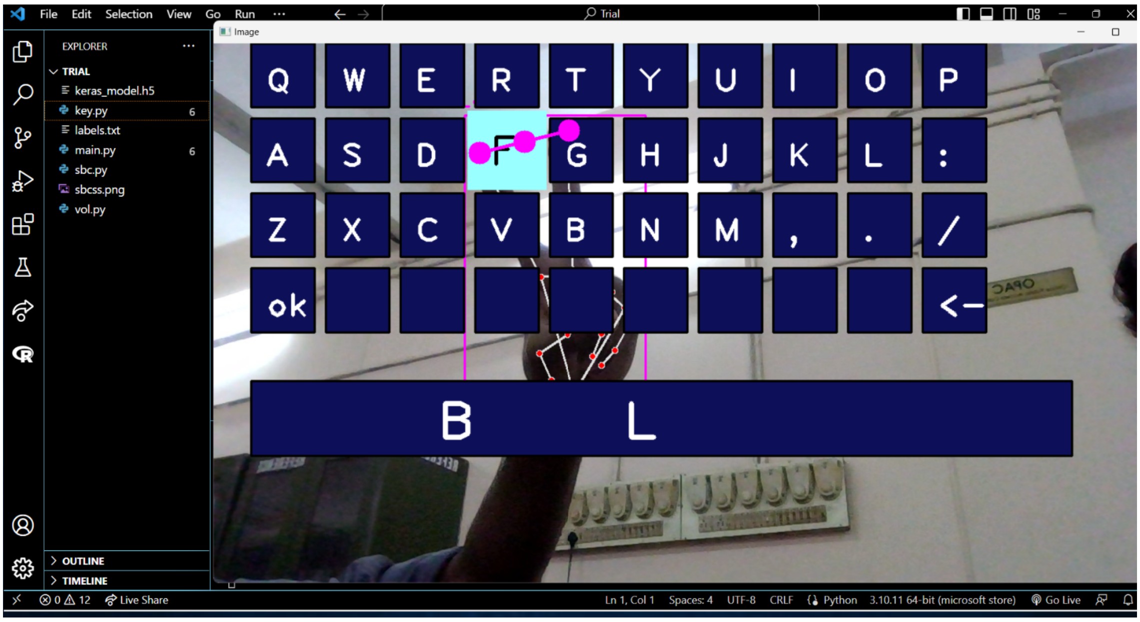

Figure 6.4: Virtual Keyboard

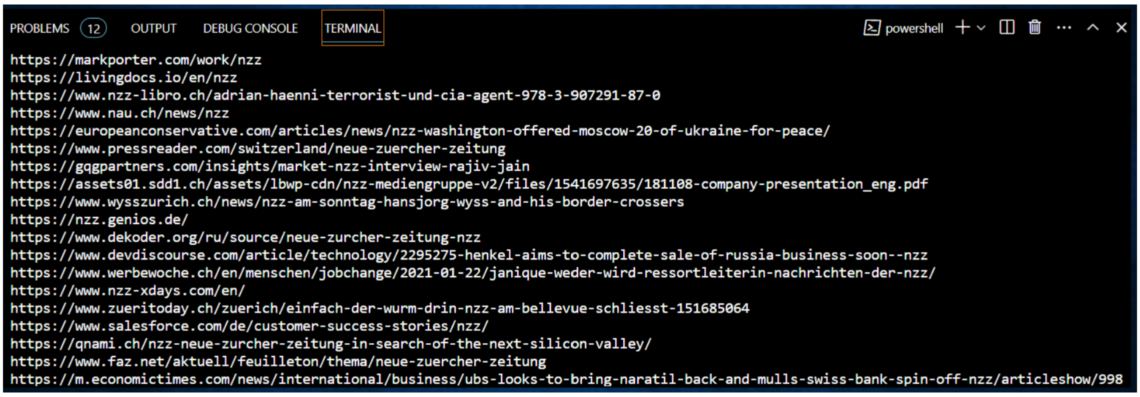

Figure 6.5: Google Search

Chapter 7

CONCLUSION

The gesture-based system enables users to interact with digital devices using human gestures instead of conventional input devices like keyboards or mice. User-friendliness and intuitiveness are key considerations in its design, ensuring that users can quickly grasp the hand gestures and their corresponding actions. The system aims to provide a seamless and natural way of controlling and navigating digital interfaces, enhancing the overall user experience. By eliminating the need for physical peripherals, this technology offers a more immersive and hands-free interaction paradigm, empowering users with a new level of control and convenience.

References

[1] International Research Journal of Modernization in Engineering Technology and Science.

[2] International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056.

[3] Shaurya Raj, Tuhin Kalia, Shivansh Aggarwal, Sahil Jaglan, Nikita Nijhawan and Mugdha Sharma.

[4] Hashmi, N., Kawar, J., Kawar, S., Patil, K. T. (2019). Eye-Blink Detection System for Virtual Keyboard. Smt. Indira Gandhi College of Engineering, Navi Mumbai, Maharashtra, India.

[5] Research Gate, www.researchgate.net.

[6] Wikipedia, https://en.wikipedia.org/wiki/Main$_{P}age$.