WIRELESS APPLICATION CONTROL USING HAND GESTURES

A Project Report

Submitted By

Vankamamidi Phani Saketh

Pamarthi Venkata Sandeep

Palle Chandra Hasan Reddy

Kataru Nehemya

200303125069

200303125061

200303125060

200303126119

in Partial Fulfilment For the Award of

the Degree of

BACHELOR OF TECHNOLOGY

COMPUTER SCIENCE AND ENGINEERING

Under the Guidance of

Asst. Prof. Zeenal Patel

Assistant Professor

PARUL UNIVERSITY

VADODARA

October - 2023

PARUL UNIVERSITY

CERTIFICATE

This is to Certify that Project - 2 -Subject code 203105400 of 7th Semester entitled "Wireless Application Control Using Hand Gestures" of Group No. PUCSE_136 has been successfully completed by

- VANKAMAMIDI PHANI SAKETH - 200303125069
- PAMARTHI VENKATA SANDEEP - 200303125061
- PALLE CHANDRA HASAN REDDY - 200303125060
- KATARU NEHEMYA - 200303126119

under my guidance in partial fulfillment of the Bachelor of Technology (B.TECH) in Computer Science and Engineering of Parul University in Academic Year 2023- 2024.

Date of Submission :-

Head of Department,

Project Guide:-Asst. Prof. Zeenal Patel

Dr. Amit Barve

CSE, PIET

Project Coordinator:- Prof. Yatin Shukla

Parul University

External Examiner

Acknowledgements

"The single greatest cause of happiness is gratitude."

-Auliq-Ice

We are happy to present this project. 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 every one who has helped us make this project a success.

We heartily thank our Project mentor Prof. Yatin Shukla Dept of Computer science and Engineering, for his constant follow up and advice throughout the course of the Project work.

We gracefully thank our Project guide Prof. Zeenal Patel , 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 Information Science Department for their help and kind co-operation. Lastly we thank our parents and friends for their encouragement and support given to us in order to finish this precious work.

Parul University,

Vadodara

Abstract

Hand gestures provide a natural way for humans to interact with computers to perform a variety of different applications. The human-computer interaction became an importance in bringing the idea that the link between a user and a computer should look more and more like one between two human beings. Thus, with increasing improvement in technology, response time and ease of operations are the concerns. Here is where human-computer interaction comes into show. This interaction is open and challenges the used devices such as the keyboard and mouse for input. Gestures are natural and are frequently used in day-to-day communications. Therefore, communicating using gestures with computers creates a whole new standard of interaction. In this project, with the help of computer vision and deep learning techniques, user hand movements (gestures) are used to control the media player, PPT slides, minimizing and closing the applications etc. The proposed web application enables the user to use their local device camera to identify their gesture and execute the control over the media player and similar applications (without any additional hardware). It increases effectiveness and makes communication easy in allowing the user control his/her laptop/desktop from a particular distance. We are thinking that in future people will move to wireless technologies and gestures are part of that so, we have made our choice to do this project.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Project problem statement

In our increasingly interconnected world, the demand for user-friendly and efficient methods of interacting with digital devices and applications has never been higher. The reliance on traditional input devices like keyboards and touchscreens can be limiting, especially in scenarios where users need to control applications from a distance or when their hands are occupied. This project undertaken to solve this challenge by creating a system that allows users to control applications wireless using hand gestures. The primary issue this project aims to address is the need for a more accessible and natural way to interact with technology. This involves overcoming the constraints of current input methods, developing intuitive and responsive gesture recognition technology, and ensuring that the solution is flexible and adaptable to various applications and user preferences. By focusing on these fundamental challenges, this project aims to improve the overall user experience, making technology more accessible and natural for a wide range of users in different contexts, from smart home control to virtual reality gaming.

1.2 Scope

The scope of the project encompasses the design, development, and implementation of a wireless application control system that utilizes hand gestures as the primary input method. The project will target a wide range of applications, including but not limited to smart home automation, multimedia playback, virtual reality (VR) environments, and gaming. Some key components and objectives are

- **Gesture Recognition System:** Develop a real-time gesture recognition system capable of detecting and interpreting a predefined set of hand gestures accurately. This system should be adaptable to different environments and lighting conditions to ensure consistent performance.

- **Wireless Communication:** Implement wireless communication protocols (e.g., Bluetooth, Wi-Fi) to establish a seamless connection between the gesture recognition module and the target applications or devices. Ensure low latency and reliable data transmission.
- **Application Integration:** Integrate the gesture control system with a variety of applications, demonstrating its versatility and adaptability. Showcase practical use cases, such as controlling lights, appliances, media players, and immersive VR experiences.
- **User Interface (UI):** Design an intuitive user interface that allows users to configure and customize gesture commands for specific applications. Ensure ease of use and accessibility for users with varying levels of technical expertise.

1.3 Aim

The aim of this project is to design, develop, and implement an innovative and human interactive system for wireless application control using hand gestures as the primary input method. Our primary objectives are to create a user hand movements (gestures) to control the media player, PPT slides, minimizing and closing the applications .

1.4 Objective

1. **Media Player Control:**
    - Develop a module that allows users to control media playback (play, pause, stop, volume adjustment) using hand gestures.
    - Ensure compatibility with popular media player applications (e.g., VLC, Windows Media Player) to offer users a seamless experience.
2. **PPT Slide Navigation:**
    - Implement a feature for advancing and reversing PowerPoint (PPT) slides through intuitive hand gestures.
    - Ensure smooth integration with Microsoft PowerPoint and support for common presentation functionalities.
3. **Application Window Management:**
    - Create functionality for users to minimize and close open application windows through specific hand gestures.

- Ensure compatibility with a variety of applications, including standard desktop software and web-based applications.

4. **Gesture Recognition with CNN:**

   - Integrate a Convolutional Neural Network (CNN) model to enhance the accuracy and robustness of hand gesture recognition.
   - Train the CNN model with a diverse dataset of hand gestures to improve recognition performance.

Chapter 2

Literature Review

2.1 Literature Survey

1. **PAPER 1:- A Deep Convolutional Neural Network Approach for Static Hand Gesture Recognition[1]**

   - **Author Name :** Aditya V. and Rajesh R.
   - **Publication Year :** 2020
   - This paper proposes a method for the automatic recognition of hand postures using convolutional neural networkswith deep parallel architectures. The proposed model avoids the need for hand segmentation, which is a very difficult task in images with cluttered backgrounds. Eventhough many different segmentation techniques are possible based on skin color, hand shape etc., they all fail in giving proper results when applying to images with other background objects. This approach also avoids the hectic job of deriving potential feature descriptor capable of recognizing the various gesture classes. The method has been tested on the datasets with very large numbers of gesture classes with images having both uniform as well as complex backgrounds and the results are promising. The performance of the proposed method in this paper has been evaluated on two publicly available datasets through five fold cross validation. The performance analysis using the statistical metrics such as accuracy, precision, recall and F1-score shows greater recognition power of the proposed CNN model.

2. **PAPER 2:- A Real Time Hand Gesture Recognition System Using Motion History Image[2]**

   - **Author Name :** Hsieh Chiung, Liou, Dunga-huna, & Lee David

- **Publication Year :** 2010
- In this paper, a face based adaptive skin color model and a motion history image based hand moving direction detection method were proposed. There are four dynamic hand gestures hand moving up, moving down, moving left, and moving right and two static hand gestures fist and waving hand defined in this paper. These hand gestures are natural and simple. Harr-like features were designed to detect the four directional dynamic hand gestures. Static hand gestures were extracted by the face based adaptive skin color model and detected by checking a face based ROI. Five persons were invited to test our proposed system. Experimental results showed the accuracy is 94.1% in average and demonstrated the feasibility of proposed system.

3. **PAPER 3:- Controlling Media Player with Hand Gestures using Convolutional Neural Network[3]**
   - **Author Name :** Nagalapuram , Roopashree, Varshashree, Dheeraj, & Nazareth
   - **Publication Year :** 2021
   - This paper presents the work to control the media player using the hand gesture recognition system. The OpenCV techniques are used to capture the images, 2-Dimensional Convolutional Neural Network is used to extract features and predict the gestures, PyAutoGUI is used to control the keyboard keys whenever a gesture is integrated with it is predicted. A custom dataset of 7 gestures is collected to test the proposed model. This model is also tested with these gestures in real-time to examine the accuracy of the proposed system. The proposed CNN model achieved a high accuracy of 98%, providing a user-friendly, cost-effective approach to interaction with computer systems. Thus, the proposed system is a true real-time model with low to negligible latency.

4. **PAPER 4:- Controlling Power Point Using Hand Gestures in Python[4]**
   - **Author Name :** Idrees.M, Ahmad.A, Butt.M, Danish.
   - **Publication Year :** 2021
   - This research paper focused on removing one such distraction by allowing the presenter to manage slides solely by gesturing in front of the camera.They managed to map specific gestures for one action on the slides, including the next slide, previous slide, zoom in

and out,opening a highlighter/pen, and play/pause videos. This project uses the python language, and the library used for machine learning is Py Torch. They used a dataset named 20BN-jester Dataset V1 for the project, which had 148,092 gestures. The results achieved were excellent, which can make hand gestures a righteous mode of presenting.

5. **PAPER 5:-Development of an Automated Hand Gesture Software to Control Volume for Computer[5]**

    - **Author Name :** Sakthimohan.M, Elizabeth Rani.G, Navaneethakrishnan.M, Revathi.S, Naik, Reddy.P.V .
    - **Publication Year :** 2023
    - This paper concludes that an application that allowed user to perform hand gestures for clean software control. An imaginative and prescient-based hand Gesture system that does not require any unique markers or gloves and may operate in real-time on a commodity laptop with low-cost cameras. Specially, the system can music the top positions of the counters and index finger for every hand. An automated software provided a robust hand gesture recognition in the computer and it do not require any markers. The software is more user friendly. For any type of finger distance of both index and thumb fingers the accuracy predicted 99.3 % and the time run out only 0.5 seconds for any angular position of finger. In future we planned that for other facial or body expressions the volume to be control for all applications.

6. **PAPER 6:- Hand Gesture Recognition for Human Computer Interaction [6]**

    - **Author Name :** Haria.A, Subramaniana.A, Asokkumara.N, Poddar.S, Nayaka.
    - **Publication Year :**2017
    - In this paper a robust gesture recognition system that did not utilize any markers, hence making it more user friendly and low cost. In this gesture recognition system, They have aimed to provide gestures, covering almost all aspects of HCI such as system functionalities, launching of applications and opening some popular websites. In future we would like to improve the accuracy further and add more gestures to implement more functions. Finally, they target to extend our domain scenarios and apply our tracking mechanism into a variety of hardware including digital TV and mobile devices. They also aim to extend this mechanism to a range of users including disabled users

7. **PAPER 7:- Hand Gesture Recognition Using 3D-CNN Model[7]**
    - **Author Name :** Al-Hammadi Muhammad, Abdul, Alsulaiman.
    - **Publication Year :** 2020
    - This paper proposed system used linear mapping for frame selection. To over come the lack of huge annotated dataset of hand gestures, transfer learning had been investigated. The pre-trained C3D structure was fine tuned in the proposed system for that purpose.The proposed system achieved excellent accuracy in most of the cases and still need more enhancement for the signer independent mode. The performance of the proposed system has been compared with some other state of the art methods in the literature. As for the future work, we will generalize the evaluation of the proposed system on more hand gesture classes. We will also investigate the effect of different parameters on the system performance such as the frames' selection technique, the optimization method, and the loss function.
8. **PAPER 8:- Vision-based hand-gesture applications[8]**
    - **Author Name :** Wachs, Kölsch, Stern,Edan.
    - **Publication Year :** 2011
    - This paper provides a comprehensive exploration of the applications and advancements in vision-based hand-gesture recognition technology. It delves into the state-of-the-art methods and techniques used in computer vision and machine learning to recognize and interpret hand gestures accurately. The paper discusses how these technologies are harnessed in various domains, including human-computer interaction, gaming, virtual reality, and medical applications. It also highlights the role of gesture control in enhancing user experiences, improving accessibility, and enabling touch less interactions with electronic devices and software interfaces. The paper addresses the challenges, opportunities, and future directions in the field of vision-based hand-gesture applications, making it a valuable resource for researchers, developers, and professionals interested in this rapidly evolving area of technology.
9. **PAPER 9:- Dynamic Hand Gesture Recognition Using 3DCNN and LSTM with FSM Context-Aware Model[9]**

- **Author Name :** Hakim, Timothy K.Shih, Kasthuri Arachchi.S.P, Aditya Wisnu, Yi-Cheng, Chih-Yang.
- **Publication Year :** 2019
- This paper presents the work to solve the gesture recognition problem on real-time application situation by combining RGB and depth modalities as the input for the deep learning model. The combination of 3DCNN and LSTM model could extract the spatio-temporal features of the gesture sequence, especially with the dynamic gesture recognition. In the term of using it in a real-time application, adding the FSM controller model could narrow the gesture classification search on the model into a smaller one that could ease the model work and enhance the accuracy result. Dataset collection of 24 gestures were designed associated with a smart TV-like environment to test the proposed model. As for the application's real-time testing, eight gestures were to examine the robustness of our work. The result suggests if the FSM controller may enhance the accuracy result in real-time applications.

10. **PAPER 10:-A Sliding Window Approach to Natural Hand Gesture Recognition using a Custom Data Glove[10]**

    - **Author Name :** Granit Luzhnica, Elizabeth Lex, Viktoria Pammer.
    - **Publication Year :** 2016
    - In this paper, we presented a gesture recognition system built for recognising 31 natural and interaction-oriented hand gestures. Our feature extraction is based on statistics and spectral properties of a sliding window over the data stream. We show that our features are highly discriminative for natural hand gestures and we achieve an accuracy of 98.5% with our gesture recognition system, which relies on linear discriminant analysis for dimensionality reduction and logistic regression for classification. Moreover, accuracy does not significantly suffer (98.2%) when the computationally expensive FFT features are removed. The main contribution of this paper lies in showing that all selected gestures can be recognised very well, given the sensors on the custom data glove and selected features extracted using sliding window approach.

11. **PAPER 11:- Static Hand Gesture Recognition Based on Convolutional Neural Networks[11]**

    - **Author Name :** Raimundo.F,Pinto, Carlos Borges, Antonio Almeida , and Ialis Paula.

- **Publication Year :** 2019
- This research paper is dedicated to the study and application of Convolutional Neural Networks (CNNs) in the context of static hand gesture recognition. It presents a comprehensive examination of how CNNs are employed to effectively identify and interpret static hand gestures. The paper explores the development of CNN architectures tailored for this specific task, encompassing data collection, model training, and evaluation metrics. It likely showcases experimental results demonstrating the CNN's accuracy in recognizing various static hand gestures, with potential comparisons to other recognition methods. With a focus on enhancing accuracy and robustness, the paper discusses the implications of its findings for practical applications, particularly in human-computer interaction, sign language translation, and accessibility tools. As a valuable contribution to the field of computer vision and gesture recognition, this research paper is a valuable resource for researchers, engineers, and practitioners interested in leveraging CNNs for static hand gesture recognition tasks.

12. **PAPER 12:- Hand gesture recognition on python and Open-cv[12]**

    - **Author Name :** Ahmad Puad Ismail , Farah Athirah Abd Aziz , Nazirah Mohamat Kasim and Kamarulazhar Daud.
    - **Publication Year :** 2021
    - This research paper focuses on the development of a hand gesture recognition system using the Python programming language and the Open-CV (Open Source Computer Vision) library. It provides an in-depth exploration of the methodology and techniques used to create an effective hand gesture recognition solution. The paper likely discusses various aspects of the system, including data collection, preprocessing, and the design of gesture recognition algorithms. It may showcase how the OpenCV library is leveraged for image processing and feature extraction. Additionally, the paper is expected to present experimental results and performance evaluations to demonstrate the accuracy and robustness of the hand gesture recognition system. This research paper serves as a valuable resource for those interested in the practical implementation of hand gesture recognition systems using widely available and accessible tools like Python and OpenCV.

13. **PAPER 13:- Hand gesture classification using a novel CNN-crow search algorithm[13]**

- **Author Name :** Thippa Reddy Gadekallu, Mamoun Alazab, Rajesh Kaluri ,Praveen Kumar Reddy Maddikunta , Sweta Bhattacharya , Kuruva Lakshmanna , Parimala
- **Publication Year :** 2021
- This research paper introduces a novel approach to hand gesture classification by combining Convolutional Neural Networks (CNNs) with a specialized optimization technique called the Crow Search Algorithm (CSA). The paper provides a detailed exploration of this innovative fusion of deep learning and optimization. It is likely to discuss the methodology of training CNN models for hand gesture recognition, focusing on data preparation, model architecture, and training parameters. Additionally, the paper delves into the integration of CSA, an optimization algorithm inspired by the foraging behavior of crows, to enhance the CNN's performance in gesture classification. It is expected to present experimental results showcasing the effectiveness of this hybrid approach, demonstrating high accuracy in recognizing and classifying a variety of hand gestures. This research paper contributes a unique perspective to the field of hand gesture classification, offering insights into the synergy between CNNs and nature-inspired optimization techniques like CSA, with potential applications in human-computer interaction, sign language translation, and more.

14. **PAPER 14:- Controlling Multiple Applications with Hand Gesture Using Convolution Neutral Network[14]**

    - **Author Name :** Raj Kumar , Smitha , Adiga, Md Altaf Raja, Anuhya, Nandini
    - **Publication Year :** 2022
    - This research paper presents an innovative approach to human-computer interaction by harnessing Convolutional Neural Networks (CNNs) for the recognition of hand gestures to control multiple applications simultaneously. The paper offers a comprehensive exploration of the methodology and techniques employed in this unique system. It is expected to cover the entire process, from data collection and pre processing to the design and training of CNN models specifically tailored for hand gesture recognition. Additionally, the paper likely discusses the seamless integration of these gesture controls into various applications across different domains, exemplifying their versatility and adaptability. Experimental results would likely be presented, demonstrating the system's capability to accurately and efficiently control multiple applications through hand

gestures. This research paper represents a valuable contribution to the field of human-computer interaction, offering insights into the practical implementation of CNN-based hand gesture recognition for enhanced multi-application control, with potential applications in smart homes, multimedia, and beyond.

15. **PAPER 15:- A Smart Vision Based Single Handed Gesture Recognition system using deep neural networks[15]**

    - **Author Name :** Suguna.R , Rupavathy.N , Asmetha Jeyarani.R
    - **Publication Year :** 2021
    - This paper is about a vision based noninvasive and cost effective hand gesture recognition system has been proposed in this work. Several research works have been carried out on different datasets. The proposed CNN architecture performs well with minimum filters and layers. The simplicity in design reduces the response time during testing. Analyzing the misclassified samples and improving the training size may enhance the recognition rate. By tuning the hyper parameter the recognition rate has improved yielding an accuracy score of 98.1% in test dataset.

16. **PAPER 16:- Video Controlling Using Hand Gestures for Disabled People[16]**

    - **Author Name :** Stella Nadar, Simran Nazareth, Kevin Paulson, NilambriNarkar.
    - **Publication Year :** 2021
    - This research paper addresses an important aspect of assistive technology by presenting a system that enables video control through hand gestures, specifically designed to cater to the needs of disabled individuals. The paper explores the development and implementation of a gesture recognition system, likely utilizing computer vision techniques and machine learning algorithms, to interpret hand movements and translate them into video control commands. It is expected to discuss the process of data collection, including the use of diverse datasets and image preprocessing techniques to accommodate various users. The research likely emphasizes the importance of accessibility and user-friendliness, focusing on how individuals with physical disabilities can navigate and interact with video content independently. Practical applications of this technology, such as controlling video playback or selecting content through gestures, would be detailed. The paper is likely to underscore the positive impact of this

innovation on disabled individuals' quality of life and independence, making it a valuable contribution to the field of assistive technology

17. **PAPER 17:- Gesture Recognition using CNN and RNN[17]**

    - **Author Name :** Rajalakshmi.J, Kumar.P
    - **Publication Year :** 2020
    - This paper contains the Gesture Recognition System for controlling electronic appliances is implemented using the Deep learning model by combining the Convolution Neural Network and Recurrent Neural Network. The CNN model will extract the efficient features from the hand gestures and the RNN will identify the temporal patters to recognize the dynamic gestures. In future, more devices can be controlled by including various gestures and will improves the quality of gesture image.

18. **PAPER 18:- Sign Language Recognition Using Neural Network[18]**

    - **Author Name :** Kaustubh Jadhav , Abhishek Jaiswal , Abbas Munshi , Mayuresh Yerendekar
    - **Publication Year :** 2017
    - This paper has been developed from classifying only static signs and alphabets, to the system that can successfully recognize dynamic movements that comes in continuous sequences of images. Researcher nowadays are paying more attention to make a large vocabulary for sign language recognition systems. Many researchers are developing their Sign Language Recognition System by using small vocabulary and self-made database. Large database build for Sign Language Recognition System is still not available for some of the country that involved in developing Sign Language Recognition System.The neural networks are one of the more powerful tools in the identification system and pattern recognition. The system presents a performance pretty good to identify the static images of the sign alphabetic language. The system shows that the first stage can be useful for deaf persons or with speech disability for communicating with the rest of the people who do not know the language. In this work, the developed hardware architecture is used as image recognizing system but it is not only limited to this application, it means, the design can be employed to process other type of signs.

19. **PAPER 19:- Arduino based Hand Gesture Control of Computer Application[19]**

- **Author Name :** Gaurav Sawardekar, Parthil Thaker, Rishiraj Singh, Vaishali Gaikwad (Mohite)
- **Publication Year :** 2018
- This research paper presents an innovative system that leverages Arduino-based technology to enable hand gesture control of computer applications. The paper delves into the development and implementation of a practical and cost-effective solution for enhancing human-computer interaction. It is expected to discuss the architecture and components of the Arduino-based system, detailing how it captures and processes hand gestures through sensors or cameras. The paper is likely to address the software integration necessary to interpret these gestures as commands for various computer applications, making it accessible to a wide range of users. Practical applications of this technology, such as controlling multimedia playback, navigating slideshows, or managing desktop applications, would likely be explored. The research paper is likely to highlight the benefits of this approach in terms of accessibility and user convenience, as well as its potential in assisting individuals with physical disabilities or those seeking hands-free control options. This work represents a valuable contribution to the field of human-computer interaction and demonstrates the practicality of using Arduino technology for enhancing the usability of computer applications.

20. **PAPER 20:- Hand Gesture Control Car[20]**

    - **Author Name :** Rutwik Shah, Vinay Deshmukh, Viraj Kulkarni, Shatakshi Mulay, Madhuri Pote
    - **Publication Year :** 2020
    - This paper contains the introduction of IoT and combination of IoT and physical devices makes life easier. The tasks which are dangerous and hazardous can be done very easily. The introduction of IoT also completes the tasks in a very short duration of time. Moreover the human errors are reduced on a great scale and results are achieved with great accuracy.The limitations such as great power consumptions can be overcomed by use of strong batterys and low power consumption sensors. Human hand can move in various directions however only five of them are recognized by the car. So further addition to the project is detecting other hand gestures and getting the output from the car accordingly.

Chapter 3

Methodology

3.1 Work flow

- The working flow-chart of gesture recognition system is below.The methodology that we used for our project consists of different phases.

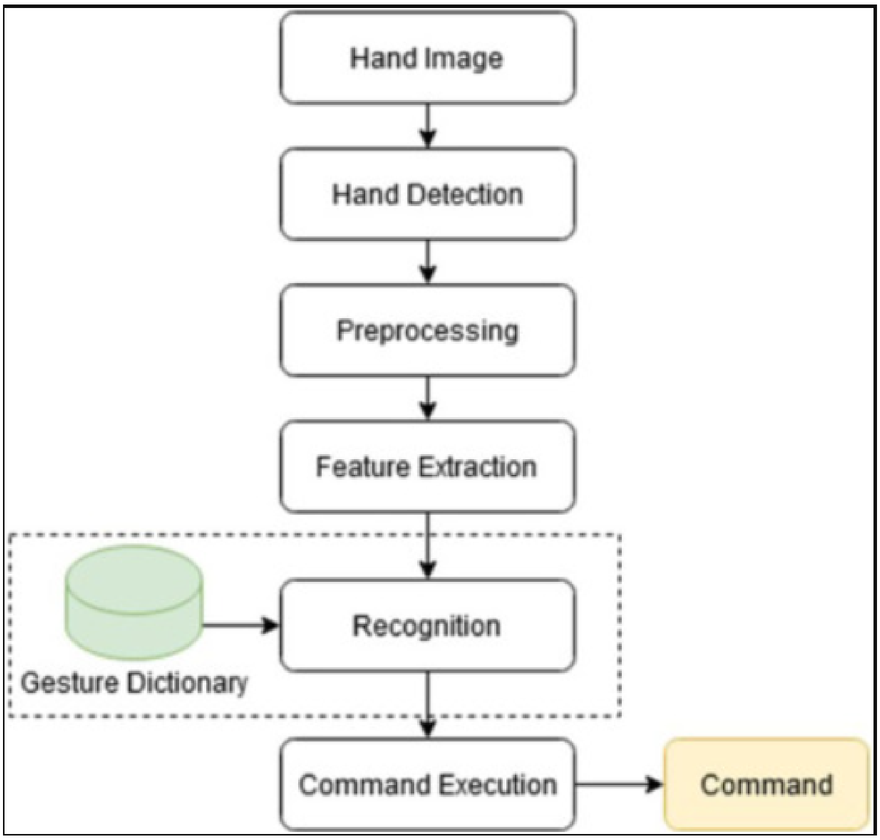

Figure 3.1: Flow chart of hand gesture recognition system

- **Hand Image:** These phase contains the images of hand symbols that will control the applications of your project.We used a dataset containing hand gesture images, which were categorized into multiple classes representing different gestures. The dataset was divided into training and testing sets for model evaluation.The dataset images in this project we used are color images only. “01_one”, “02_two”, “03_three”, “04_four”, “05_five”, “little & index fingers”, “thumb & index fingers”, “thumb & little fingers”, “thumb, index & little fingers”, “none” are the mentioned folders names which contains 200 images each. 2000 images in training data and 2000 images in testing data totally 4000 images. The dataset we taken can be consider as a medium dataset for deep learning techniques.

- **Hand Detection:** These phase detect the symbols of hands with the help of webcam.
- **Preprocessing:**Preprocessing is a crucial step in many computer vision and machine learning projects, including those involving hand detection or gesture recognition. The preprocessing steps we need to perform may vary depending on the specific requirements of your project and the quality of your data. here are some common preprocessing steps.
  - **Data Collection and Labeling:** Collect a diverse and representative dataset of hand images or videos that covers various hand poses, lighting conditions, and backgrounds.Label the data by annotating hand regions or gestures in each frame or image. Proper labeling is essential for supervised learning.
  - **Data Augmentation:**Augment your dataset by applying random transformations to the images or frames. Common augmentations include rotation, scaling, translation, and flipping. Data augmentation helps improve model robustness and generalization.
  - **Normalization:**Normalize pixel values in images or frames to a standard range (e.g., [0, 1] or [-1, 1]). Normalization helps improve model convergence during training.
  - **Data Splitting:**Split your dataset into training, validation, and testing subsets. This allows you to train your model on one subset, tune hyper parameters on another, and evaluate performance on a separate subset.
  - **Data Serialization:**Serialize your preprocessed data into a suitable format for efficient training. Common formats include TFRecords or HDF5.
  - **Data Loader:**Implement a data loader or generator that loads batches of preprocessed data during training. This helps optimize memory usage.

  - **Data Normalization:**Apply normalization techniques specific to your project. For gesture recognition, you may need to calculate features like hand key points or use other methods to represent hand positions.

- **Feature Extraction:**Feature extraction is a critical step in many computer vision and machine learning projects, including those involving hand detection and gesture recognition. Feature extraction involves transforming raw data (such as images or video frames) into a more compact and representative representation that can be used for model training and analysis.

- **Recognition:** It is used for Recognizing the symbols and control the applications of the project.

- Input dataset as the symbols to control hand gestures applications. The symbols are used for controlling hand gestures applications are shown in fig3.2.

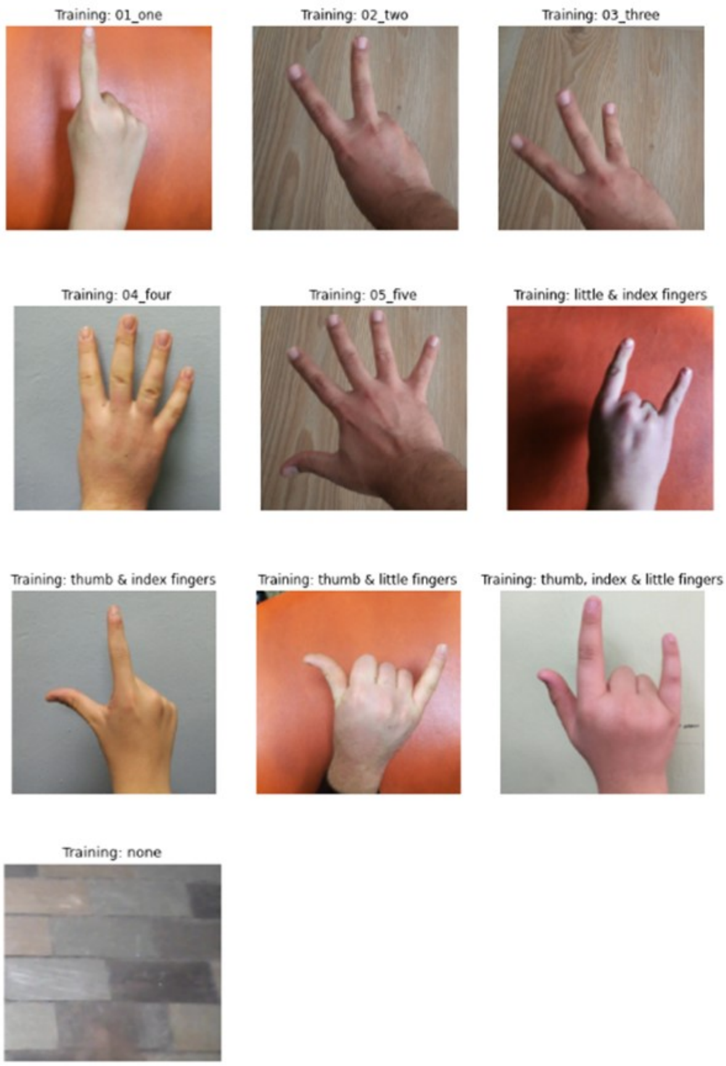

Figure 3.2: Symbols for controlling hand gestures applications.

3.2 CNN model and VGG-16 CNN model

3.2.1 Convolutional Neural Network(CNN) :

- A Convolutional Neural Network (CNN) is a type of deep learning algorithm that is particularly well-suited for image recognition and processing tasks. It is made up of multiple layers, including convolutional layers, pooling layers, and fully connected layers. The convolutional layers are the key component of a CNN, where filters are applied to the input image to extract features such as edges, textures, and shapes. The output of the convolutional layers is then passed through pooling layers, which are used to down-sample the feature maps, reducing the spatial dimensions while retaining the most important information. The output of the pooling layers is then passed through one or more fully connected layers, which are used to make a prediction or classify the image. CNNs are trained using a large dataset of labeled images, where the network learns to recognize patterns and features that are associated with specific objects or classes. Once trained, a CNN can be used to classify new images, or extract features for use in other applications such as object detection or image segmentation. CNNs have achieved state-of-the-art performance on a wide range of image recognition tasks, including object classification, object detection, and image segmentation. They are widely used in computer vision, image processing, and other related fields, and have been applied to a wide range of applications, including self-driving cars, medical imaging, and security systems.

- **Convolutional Neural Network Design:**

  - The construction of a convolutional neural network is a multi-layered feed-forward neural network, made by assembling many unseen layers on top of each other in a particular order. It is the sequential design that give permission to CNN to learn hierarchical attributes. In CNN, some of them followed by grouping layers and hidden layers are typically convolutional layers followed by activation layers. The pre-processing needed in a ConvNet is kindred to that of the related pattern of neurons in the human brain and was motivated by the organization of the Visual Cortex.The convolution layers contains units called feature maps and each of them is connected to the local patches in the previous layer through filter bank. Same filter bank is used in all the units of a feature map, and different filter banks are used in different feature maps in a layer. This architecture enables to easily identify the distinctive local patterns from images, even it is located at different parts of the image. The local weighted sum obtained through filtering operation is passed through a non-linear function called ReLu(Rectified Linear Unit) to stabilize the convolved results. The pooling operation is incorporated in the CNN structure to group the semantically similar features from the convolution layer. Thus the architecture of a CNN contains two or three convolution layers with the non linear activation and pooling layers, followed by more convolutional layers with pooling and activation, and a final fully connected layer that performs the classification.

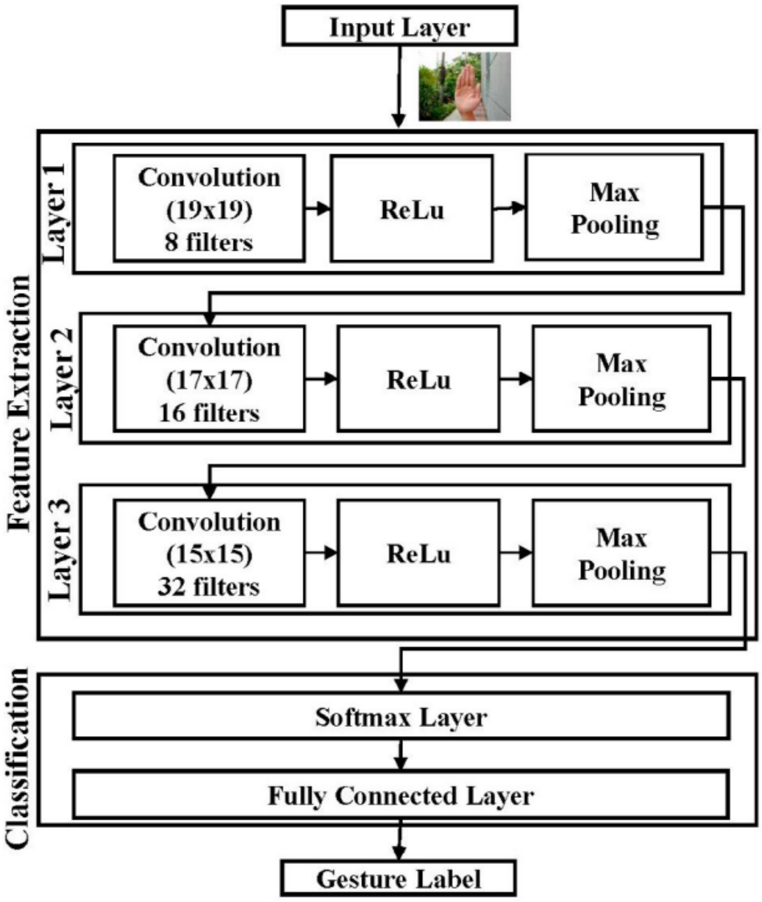

Figure 3.3: CNN Model Design

Hyperparameter	Value/Description
Number of Convolutional Layers	3
Number of Filters (per layer)	[32, 64, 128]
Filter Size (per layer)	[3x3, 3x3, 3x3]
Pooling Type (per layer)	Max Pooling
Pooling Size (per layer)	[2x2, 2x2, 2x2]
Activation Function	ReLU
Dropout Rate	0.25 (after each Conv layer)
Fully Connected Layers	2
Number of Neurons (FC layers)	[512, 10]
Activation Function (FC layers)	ReLU (except the output layer)
Output Layer Activation Function	Softmax
Learning Rate	0.001
Optimizer	Adam
Batch Size	32
Number of Epochs	10

Table 3.1: Hyperparameters for CNN

3.2.2 VGG-16 CNN MMODEL:

- VGG-16, short for Visual Geometry Group 16, is a widely recognized Convolutional Neural Network (CNN) architecture that was developed by the Visual Geometry Group at the University of Oxford. It's known for its simplicity and effectiveness in image classification tasks.
- **VGG-16 CNN model architecture overview:** VGG-16 consists of 16 layers, including 13

convolutional layers, followed by three fully connected layers. The convolutional layers are divided into five groups, with each group containing multiple convolutional layers followed by max-pooling layers. The last three fully connected layers are responsible for classifying the input image.

- **Convolutional and Pooling Layers:** The convolutional layers use small filters (typically 3x3) to extract various features from the input image. The deeper the network, the more abstract and complex features it can capture. After each convolutional layer, a max-pooling layer reduces the spatial dimensions of the feature maps, helping to decrease the computational load and increase the receptive field.
- **Filter Sizes and Depth:** VGG-16 utilizes 3x3 filters throughout the network. It maintains a consistent depth of 64 filters for the first two convolutional groups and then doubles the depth for the subsequent groups (128, 256, 512, and 512). This depth increase enables the model to capture more complex patterns.
- **Fully Connected Layers:** After the convolutional layers, the architecture appends three fully connected layers, each with 4096 units. These layers combine the high-level features learned by the convolutional layers and perform the final classification.
- **Activation Function:** Rectified Linear Unit (ReLU) activation functions are used after each convolutional and fully connected layer. ReLU helps introduce non-linearity into the network, allowing it to learn more complex mappings between inputs and outputs.
- **Dropout:** VGG-16 employs dropout regularization in the fully connected layers to prevent overfitting. Dropout randomly ”drops out” a fraction of neurons during training, forcing the network to become more robust and less reliant on specific neurons.
- **Softmax Activation:** The final layer of the network uses the softmax activation function to produce class probabilities. This makes VGG-16 suitable for multi-class classification tasks.
- **Output Layer:** The output layer has a number of units equal to the number of classes in the classification problem. The predicted class is the one with the highest probability in the softmax output.
- **Image Preprocessing:** Input images are typically resized to a fixed size (e.g., 224x224) and normalized by subtracting the mean pixel values. VGG-16’s primary contribution is in demonstrating the effectiveness of deep convolutional networks in image recognition

tasks. However, its architecture's depth and number of parameters make it computationally intensive and prone to overfitting on smaller datasets. Despite these limitations, it has paved the way for even deeper and more efficient architectures, such as ResNet and Inception, which build upon VGG's principles while introducing additional innovations.

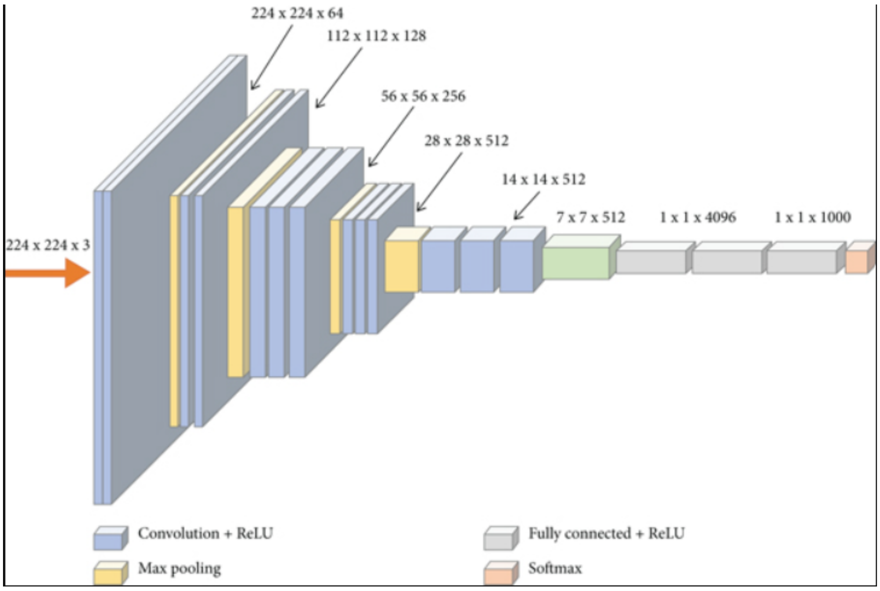

Figure 3.4: VGG-16 CNN Model Design

Hyperparameter	Value/Description
Base Pretrained Model	VGG-16 (pre-trained on ImageNet)
Input Image Size	224x224 pixels
Number of Convolutional Layers	13 (5 Conv blocks + 3 Fully connected)
Number of Filters (per layer)	[64, 128, 256, 512, 512]
Filter Size (per layer)	[3x3, 3x3, 3x3, 3x3, 3x3]
Max Pooling (per block)	After 2 Conv layers
Activation Function	ReLU
Fully Connected Layers	3 (4096, 4096, and output classes)
Activation Function (FC layers)	ReLU (except the output layer)
Output Layer Activation Function	Softmax
Learning Rate	Typically in the range of 0.0001 to 0.001
Optimizer	Adam or SGD
Weight Initialization	He Normal (or Glorot Uniform)
Batch Normalization	Applied after each Conv and FC layer
Dropout Rate	0.5
Weight Decay (L2 Regularization)	1e-4 (or other suitable value)
Batch Size	Typically 32, 64, or 128
Number of Epochs	10

Table 3.2: Hyperparameters for VGG-16 CNN

3.3 Results

- Implementation Outputs for the project.

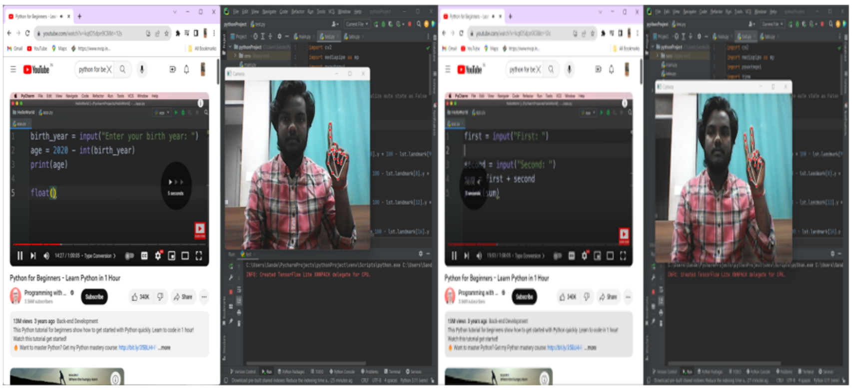

Figure 3.5: Forwarding & Backward gesture

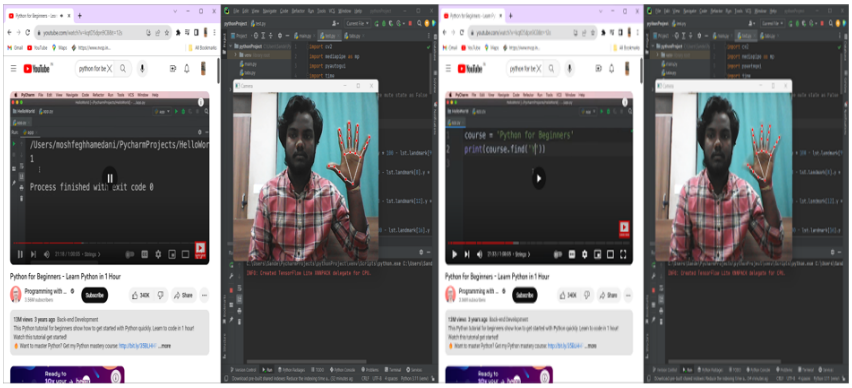

Figure 3.6: Play/Pause gesture

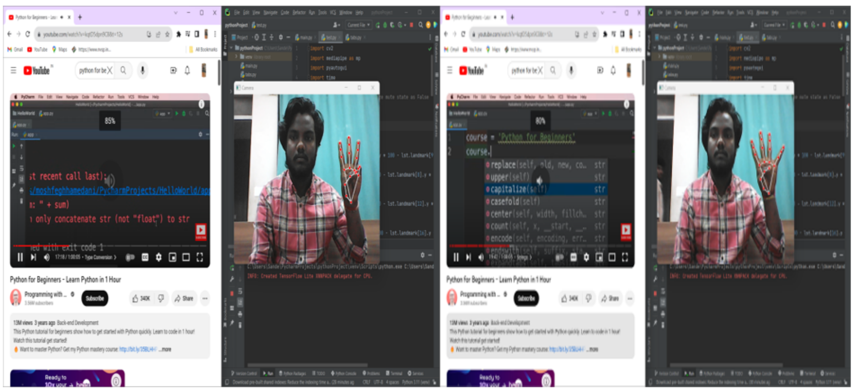

Figure 3.7: Volume Up and Volume Down gesture

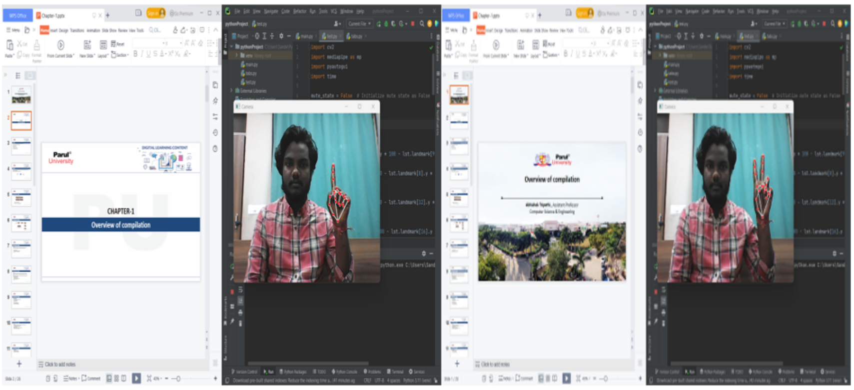

Figure 3.8: PPT Slide Forwarding & Backward gesture

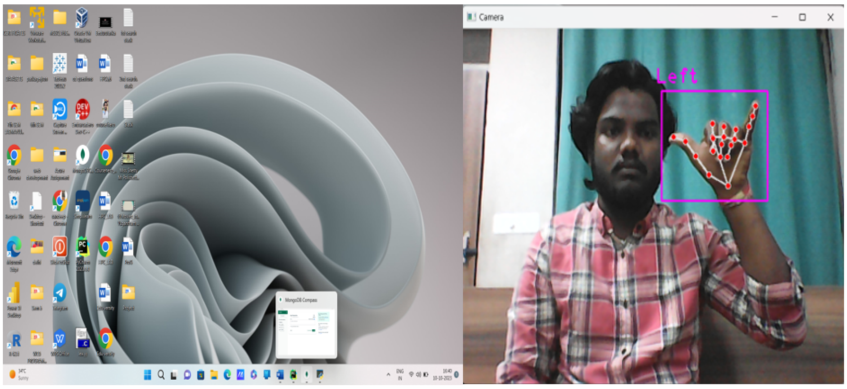

Figure 3.9: Minimize window gesture

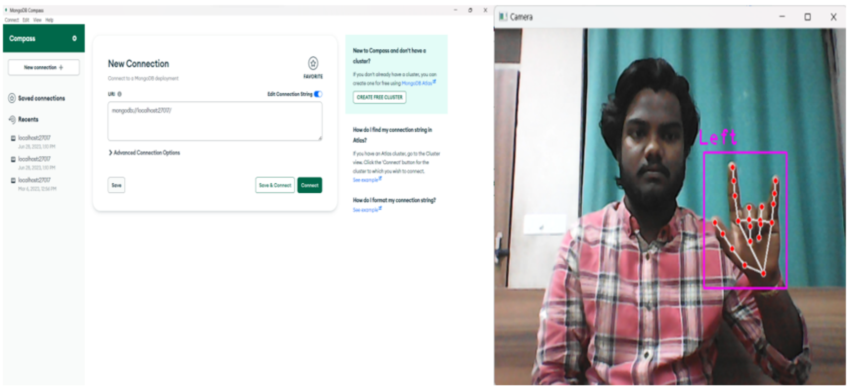

Figure 3.10: Maximize window gesture

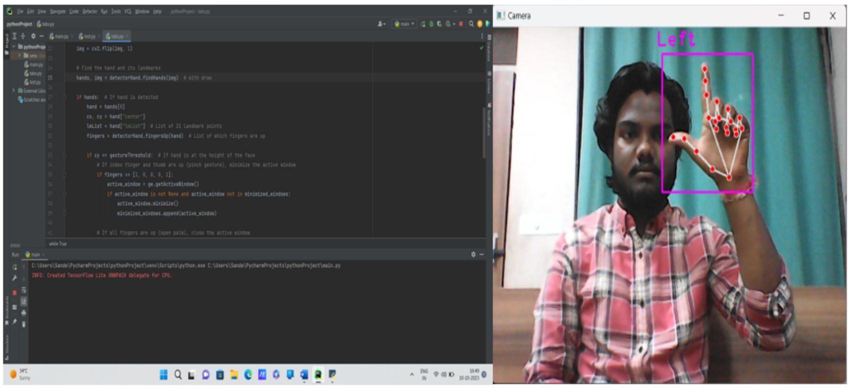

Figure 3.11: Restore window gesture

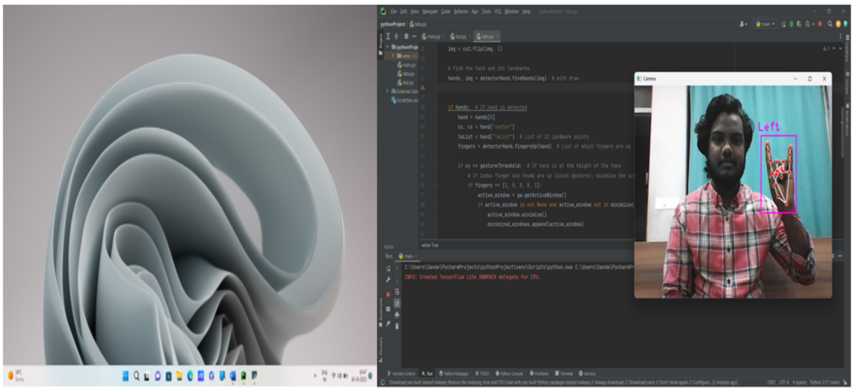

Figure 3.12: Close window gesture

3.3.1 CNN Model Performance :

The Convolutional Neural Network (CNN) model we implemented demonstrated an accuracy of approximately 81.85% on the test set. The classification report provided insights into the precision, recall, and F1-score for each gesture class. Notably, some gestures, such as the ”thumb and index fingers” sign, achieved relatively better accuracy. The model to capture intricate patterns in the images.

Figure 3.13: CNN Model Accuracy Graph

	Precision	Recall	F1-Score	Support
01_one	0.71	0.91	0.80	200
02_two	0.62	0.84	0.71	200
3_three	0.73	0.55	0.62	200
04_four	0.78	0.86	0.82	200
05_five	0.80	0.66	0.72	200
little and index fingers	0.97	0.88	0.92	200
none	1.00	1.00	1.00	200
thumb and index fingers	0.91	0.84	0.88	200
thumb and little fingers	0.91	0.86	0.89	200
thumb, index and little fingers	0.90	0.81	0.85	200
accuracy			0.82	2000
macro avg	0.83	0.82	0.82	2000
weighted avg	0.83	0.82	0.82	2000

Table 3.3: CNN Model Classification Report

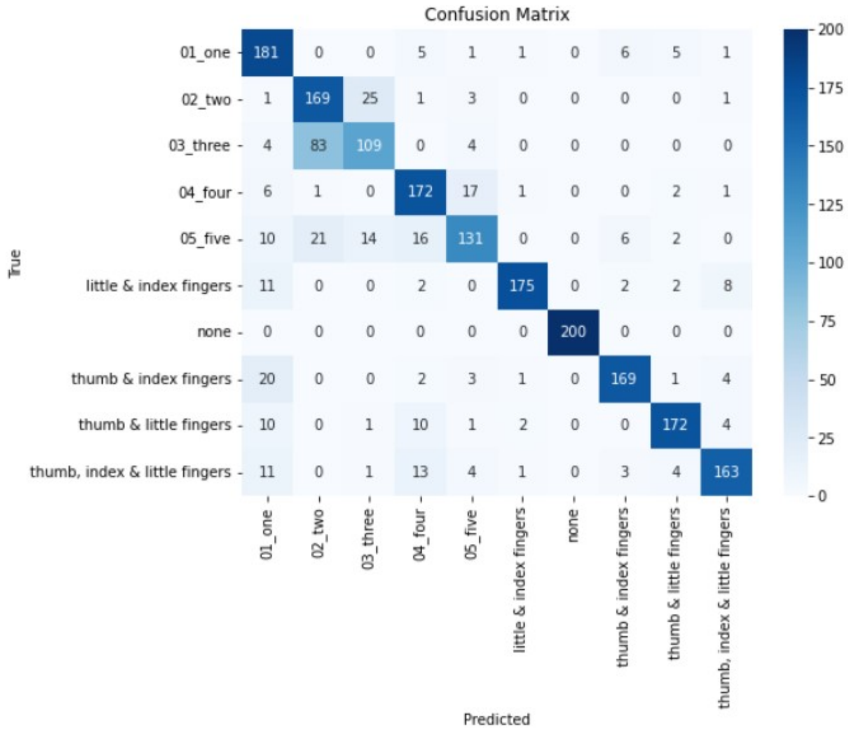

Figure 3.14: CNN Model Confusion Matrix

3.3.2 VGG-16 Based Model Performance:

In contrast, the VGG16-based model exhibited a significantly improved accuracy of 91.00%. Leveraging transfer learning from the pre-trained VGG16 architecture allowed the model to extract high-level features from the images, resulting in good generalization and accuracy. This model outperformed the CNN model across all evaluation metrics, indicating the benefits of using a well-established architecture for our task.

Figure 3.15: VGG-16 CNN Model Accuracy Graph

	Precision	Recall	F1-Score	Support
01_one	0.92	0.92	0.92	200
02_two	0.89	0.87	0.88	200
3_three	0.90	0.73	0.81	200
04_four	0.93	0.91	0.92	200
05_five	0.84	0.93	0.88	200
little and index fingers	0.92	0.95	0.94	200
none	1.00	1.00	1.00	200
thumb and index fingers	0.95	0.92	0.93	200
thumb and little fingers	0.94	0.95	0.95	200
thumb, index and little fingers	0.90	0.99	0.94	200
accuracy			0.92	2000
macro avg	0.92	0.92	0.92	2000
weighted avg	0.92	0.92	0.92	2000

Table 3.4: VGG-16 CNN Model Classification Report

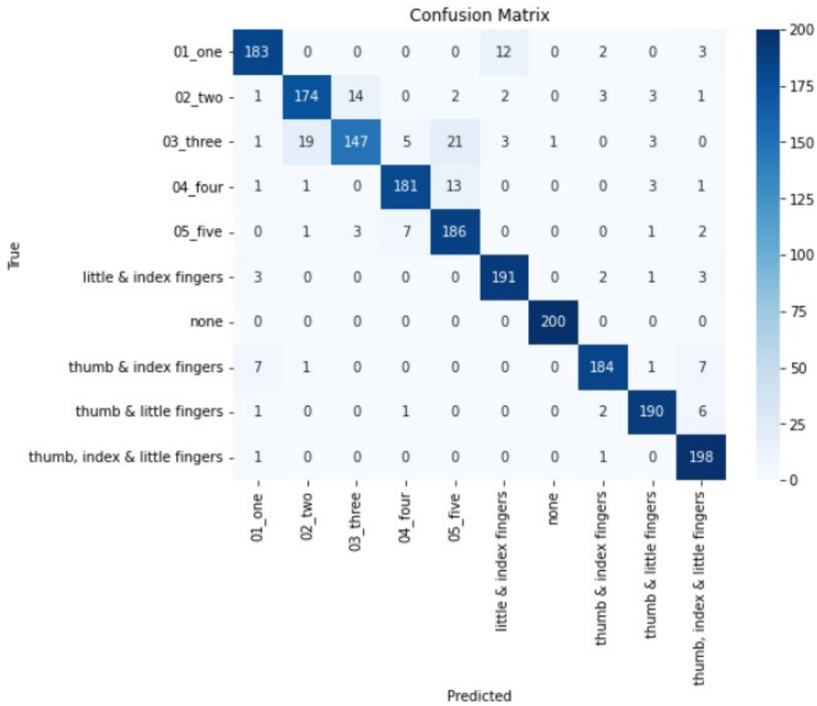

Figure 3.16: VGG-16 CNN Model Confusion Matrix

	VGG-16 Model	CNN Model
Accuracy	91.00%	81.85%
Precision (macro avg)	0.92	0.83
Recall (macro avg)	0.92	0.82
F1-score (macro avg)	0.92	0.82
Precision (weighted avg)	0.92	0.83
Recall (weighted avg)	0.92	0.82
F1-score (weighted avg)	0.92	0.82

Table 3.5: Comparison of Classification Reports

- **The gesture recognition detect hand symbols from a considerable distance, typically ranging from 5 to 6 feet.**

Chapter 4

Conclusion and Future Work

4.1 Conclusion:

This paper presents the work to control the media player controller, PPT slide changing, minimizing, maximizing, closing and restore the previous minimize window (or) application using the hand gesture recognition system. The Open-CV techniques are used to capture the images, 2- Dimensional Convolutional Neural Network is used to extract features and predict the gestures, PyAutoGUI is used to control the keyboard keys whenever a gesture is integrated with it is predicted. A custom dataset of 10 gestures is collected to test the suggested model. This model is also tested with these gestures in real-time to examine the accuracy of the recommended system. The VGG-16 CNN model achieved a high accuracy of 88.94%, providing a user-friendly, cost-effective approach to interaction with computer systems. The future scope is to work on improving the gesture recognition capabilities in different environments such as in medical fields, AI.

4.2 Future Work:

- In future we can also add so many applications control of any device using gesture recognition.
- Here we have done with the Cnn model and Vgg-16 Cnn model and we got a good accuracy. So, now we can try to improve the accuracy of gesture recognition by applying various algorithms in deep learning .
- In future we can also improve the ranging distance of the hand symbols detection from web cam as compared to now.

References

1 V.Aditya, & R.Rajesh (2020). A Deep Convolutional Neural Network Approach for Static Hand Gesture Recognition. Committee of the Third International Conference on Computing and Network Communications (CoCoNet'19), Volume(171), Pages 2353-2361.

2 Hsieh Chiung, Liou, Dunga-huna, & Lee David (2010). A Real-Time Hand Gesture Recognition System Using Motion History Image. 2nd International Conference on Signal Processing Systems (ICSPS), Volume(II), Page 394-398.

3 Nagalapuram, G. D., Roopashree, S., Varshashree, D., Dheeraj, D., & Nazareth, D. J. (2021). Controlling Media Player with Hand Gestures Using Convolutional Neural Network. IEEE Mysore Sub Section International Conference (MysuruCon), Volume(1), Pages 1-8.

4 Idrees, M., Ahmad, A., Butt, M. A., & Danish, H. M. (2021). Controlling PowerPoint Using Hand Gestures in Python. Webology (ISSN: 1735-188X), Volume(18), Page 1372-1388.

5 Sakthimohan, M., Elizabeth Rani, G., Navaneethakrishnan, M., Revathi, S., Naik, B. T.,& Reddy, P. V. (2023). Development of an Automated Hand Gesture Software to Control Volume for Computer. International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS), Volume(1), Pages 382-386.

6 Haria, A., Subramaniana, A., Asokkumara, N., Poddar, S., & Nayaka, J. S. (2017). Hand Gesture Recognition for Human-Computer Interaction. 7th International Conference on Advances in Computing & Communications, ICACC, Volume(115), Pages 367-374.

7 Al-Hammadi, M., Muhammad, G., Abdul, W., & Alsulaiman, M. (2020). Hand Gesture Recognition Using 3D-CNN Model. IEEE Consumer Electronics Society, Volume(1), Pages 2162-2248.

8 Wachs, J. P., Kölsch, M., Stern, H., & Edan, Y. (2011). Vision-based hand-gesture applications. Communications of the ACM, 54(2), 60-71.

9 Hakim, N. L, Timothy K.Shih, Kasthuri Arachchi.S. P, Aditya Wisnu, Yi-Cheng, Chih-Yang. (2019). Dynamic Hand Gesture Recognition Using 3DCNN and LSTM with FSM Context-Aware Model. Volume(1), Pages 1-19.

10 Granit Luzhnica, Elizabeth Lex, Viktoria Pammer. A Sliding Window Approach to Natural Hand Gesture Recognition using a Custom Data Glove. In: 3D User Interfaces (3DUI); 2016 IEEE Symposium on ; 2016 Mar 19 ; New York : IEEE; 2016 ; p.81-90.

11 Raimundo F. Pinto Jr. , Carlos D. B. Borges, Antoˆnio M. A. Almeida , and Ia´lis C. Paula Jr. Static Hand Gesture Recognition Based on Convolutional Neural Networks. Volume 2019, Article ID 4167890, 12 pages.

12 Ahmad Puad Ismail , Farah Athirah Abd Aziz , Nazirah Mohamat Kasim and Kamarulazhar Daud. Hand gesture recognition on python and opencv. Ahmad Puad Ismail et al 2021 IOP Conf. Ser.: Mater. Sci. Eng. 1045 012043. 11 pages.

13 Thippa Reddy Gadekallu,Mamoun Alazab ,Rajesh Kaluri , Praveen Kumar Reddy Maddikunta , Sweta Bhattacharya,Kuruva Lakshmanna,Parimala. Hand gesture classification using a novel CNN-crow search algorithm. Complex & Intelligent Systems (2021) 7:1855–1868.

14 Raj Kumar E M, Smitha G Adiga, Md Altaf Raja, T V Anuhya, Dr. C Nandini . Controlling multiple applications with hand gesture using CNN. Publised by International Journal of Creative Research Thoughts (IJCRT) . Volume 10, Issue 4 April 2022 — ISSN: 2320-2882.

15 Suguna R , Rupavathy N , Asmetha Jeyarani R . A Smart Vision Based Single Handed Gesture Recognition system using deep neural networks. I3CAC 2021, June 07-08 ,11 Pages.

16 Stella Nadar, Simran Nazareth, Kevin Paulson, NilambriNarkar. Video controlling using hand gestures for disable people. International Journal of Trendy Research in Engineering and Technology. Volume 5 Issue 2 April 2021,ISSN No 2582-0958.

17 Rajalakshmi J, Kumar P . Gesture Recognition using CNN and RNN.International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878 (Online), Volume-9 Issue-2, July 2020.

18 Kaustubh Jadhav , Abhishek Jaiswal , Abbas Munshi , Mayuresh Yerendekar. Sign language recognition using neural network. Volume 1 Issue 1,I SSN: 2581-4419.

19 Gaurav Sawardekar, Parthil Thaker, Rishiraj Singh, Vaishali Gaikwad (Mohite). Arduino based Hand Gesture Control of Computer Application. October 29 ,2018 . 6 Pages.

20 Rutwik Shah, Vinay Deshmukh, Viraj Kulkarni, Shatakshi Mulay, Madhuri Pote. Hand Gesture Control Car. International Journal of Engineering Research & Technology (IJERT) ,Published by ICSITS - 2020 Conference Proceedings, Volume 8, Issue 05, ISSN: 2278-0181.