

A PROJECT REPORT ON

CONTROLLING COMPUTER USING HAND GESTURES

SUBMITTED TO MIT SCHOOL OF ENGINEERING, LONI, PUNE IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF
THE DEGREE

BACHELOR OF TECHNOLOGY
(Computer Science & Engineering)

BY

Candidate Name: PRN/ Enrollment No:

Pradnya Rohidas Kedari	MITU18BTCS0121
Shubhangi Jagdish Kadam	MITU18BTCS0102

Under the guidance of

RAJESH PRASAD



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

MIT School of Engineering
MIT Art, Design and Technology University
Rajbaug Campus, Loni-Kalbhor, Pune 412201

2021-22



**MIT SCHOOL OF ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
MIT ART, DESIGN AND TECHNOLOGY UNIVERSITY,
RAJBAUG CAMPUS, LONI-KALBHOR, PUNE 412201**

CERTIFICATE

This is to certify that the dissertation report entitled “**CONTROLLING COMPUTER USING HAND GESTURES**”.

Submitted by:	Name of the Candidate	PRN/ Enrollment No
	Pradnya Rohidas Kedari	MITU18BTCS0121
	Shubhangi Jagdish Kadam	MITU18BTCS0102

is a bonafide work carried out by them under the supervision of Rajesh Prasad and it is submitted towards the partial fulfillment of the requirement of MIT Art, Design, and Technology University, Pune for the award of the degree of Bachelor of Technology (Computer Science and Engineering).

Rajesh Prasad
Guide
Dept. of Computer Sc. & Engg.

Dr. Shraddha Phansalkar
Head of the Department
Dept. of Computer Sc. & Engg.

Dr. Kishore Ravande
Principal
MIT School of Engineering,
MIT Art, Design and Technology,
Rajbaug Campus, Loni Kalbhor, Pune 412201

Seal/Stamp of the College
Place: Loni Kalbhor, Pune
Date: 01/07/2022

DECLARATION

We, the team members, Ms. Pradnya Rohidas Kedari (MITU18BTCS0121) and Ms. Shubhangi Jagdish Kadam (MITU18BTCS0102) hereby declare that the project work incorporated in the present project entitled “**Controlling Computer using Hand Gestures**” is original work. This work (in part or in full) has not been submitted to any University for the award or a Degree or a Diploma. We have properly acknowledged the material collected from secondary sources wherever required. We solely own the responsibility for the originality of the entire content.

Date: 01/07/2022

Pradnya Rohidas Kedari
Shubhangi Jagdish Kadam

Name and Signature of Guide

Rajesh Prasad

Seal/Stamp of the College

Place: Loni Kalbhor, Pune

Date: 01/07/2022



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
MIT SCHOOL OF ENGINEERING,
RAJBAUG CAMPUS, LONI-KALBHOR,
PUNE – 412201

EXAMINER'S APPROVAL CERTIFICATE

The project report entitled “CONTROLLING COMPUTER USING HAND GESTURES” is submitted by Ms. Pradnya Rohidas Kedari (MITU18BTCS0121), and Ms. Shubhangi Jagdish Kadam (MITU18BTCS0102) in partial fulfillment for the award of the degree of Bachelor of Technology (Computer Science & Engineering) during the academic year 2021-22, of MIT-ADT University, MIT School of Engineering, Pune, is hereby approved.

Examiners:

1.

2.

ACKNOWLEDGEMENT

We would like to express our special thanks and gratitude to our guide Rajesh Prasad sir who gave us guidance and support till the completion of our project work by providing all the necessary information to complete this project.

We would also like to extend our gratitude to our HOD Dr. Shraddha Phansalkar ma'am, Dean of Engineering Dr. Rajneeshkaur Sachdeo ma'am, and our Principal Dr. Kishore Ravande Sir for providing us with support and facility that was required.

We are thankful and fortunate enough to get constant encouragement, support and guidance from our guide as well as from project coordinators of the Department of Computer Science & Engineering, which helped us in successfully completing our project work.

By doing this project we have gained a lot of knowledge which can help us in the near future and also this entire project helped us in doing a lot of Research and we came to know about so many new things we are really thankful to them.

Pradnya Rohidas Kedari

Shubhangi Jagdish Kadam

ABSTRACT

The presence on market of the low-cost webcams with, at least, satisfactory qualities open up new directions regarding the implementation of human computer interaction (HCI) interfaces. Gesture is one of the most vivid and dramatic way of communications between human and computer. Hence, there has been a growing interest to create easy-to-use interfaces by directly utilizing the natural communication and management skills of humans.

Hand gesture recognition system contains a decent surveillance present days thanks to simple and straightforward intercommunication between human and machine. The main target of developing hand gesture is to make a much better conveyance between human and computer for delivering information. Hand gestures are a kind of nonverbal way to interact that may be employed in several fields. Research and survey papers supported hand gestures have acquire so many alternative techniques, including those supported on sensor technology and computer vision.

In this project, we decided to design a real-time human-computer interaction system based on hand gesture. Specifically, we will use the convolutional neural network (CNN) in complete process. This application presents a hand gesture interface for controlling a computer that is performing different operations using neural network. The proposed system recognizes a set of specific hand gestures. Our application is based on five phases, Image frame acquisition, Hand tracking, Features extraction, Recognition of gestures and Classification (perform desired operation). A frame from the webcam camera is captured, and then hand detection, hand shape features extraction, and hand gesture recognition will be done.

In this project, we decided to develop a generic Deep Neural Network-Based model for controlling a computer using hand gestures recognition. We will build a system to handle a device using hand gestures.

Our objective is to build and train a model for feature extraction and classification of hand gestures, to use a trained model for classification of hand gestures and recognizing an action and perform an operation accordingly, and to control a computer and perform related activities using hand gestures with the help of computer vision and deep learning algorithm.

CONTENTS

Certificate	2
Declaration	3
Examiner's Approval Certificate	4
Acknowledgement	5
Abstract	6
List of Figures	8
1. INTRODUCTION	9
1.1 Introduction	9
1.2 Market Survey	10
1.3 Research Gap	10
1.4 Problem Statement	10
1.5 Objectives	11
2. CONCEPTS AND METHODS.....	12
2.1 Concepts	12
2.2 Some Basic Terminologies.....	13
2.3 Platform/Technology Used	13
3. LITERATURE SURVEY.....	14
4. PROJECT PLAN.....	18
5. IMPLEMENTATION	20
5.1 Dataset	20
5.2 Implementation	20
5.2.1 Training of Face Emotion Detector Model	20
5.2.2 Recognition of The Hand Gesture	23
6. RESULT	24
7. CONCLUSION AND FUTURE WORK.....	29
BIBLIOGRAPHY	31
ANNEXURE A: List of Publications and Research Paper	
(In its Original formats)	32
ANNEXURE B: Plagiarism Report	33

LIST OF FIGURES

Figure Number: Figure of the table	Page Number
Figure 1.1.1 Methodology of Proposed System	10
Figure 4.1 SDLC Agile Model	18
Figure 5.1.1 Dataset Images	20
Figure 5.2.1 CNN Architecture	21
Figure 5.2.1.3 Workflow of Proposed System	23
Figure 6.1 Train vs Val accuracy	24
Figure 6.2 Train vs Val loss	24
Figure 6.3 Overall Result of Model	24
Figure 6.4 Confusion Matrix	24
Figure 6.5 GUI of System	24
Figure 6.6 Gesture Detection Frame	25
Figure 6.7 Opening the WhatsApp	26
Figure 6.8 Opening the Google Chrome	26
Figure 6.9 Opening the PowerPoint Presentation	27
Figure 6.10 Opening the VLC	27
Figure 6.11 Opening the Microsoft Edge	28

1. INTRODUCTION

1.1 INTRODUCTION

Gesture recognition is a full of life analysis field in Human Computer Interaction technology. It has several employments in virtual environment management, medical applications, sign language translation, robot control, music creation, or home automation. During this project Controlling Computer using Hand Gestures, we are going to make a real-time application using OpenCV and Python. OpenCV is a real-time open-source computer vision and image-processing library. We'll use it via the OpenCV python package.

There has been a special significance recently on Human Computer Interaction (HCI) study to form convenient interfaces by directly using common communication and handling expertise of humans. Among different chassis parts, the hand is the most helpful interaction tool, because of its adroitness. The word gesture is employed for several different cases involving human motion particularly of the hands, arms, and face, only some of these are co-operative or informative.

Gesture recognition is a crucial field that specify human gesture using computer vision techniques and algorithms. There are numerous bodily motion which may develop gesture however the usual type of gesture generation arises from the face and hands. The complete policy of tracing gesture to their representation and changing them to some useful command is called as gesture recognition. Various technologies has been used for the design and implementation of such kind of application.

The CNN or convolutional neural networks are the extremely popular used technique for image classification domain. An image classifier takes an image, or sequence of images that is video as an input and classifies it into one of the possible categories or classes that it was trained to recognize. They have applications in several different fields like driver less cars, defense, education, medical field, fraud detection, etc. There are many algorithms for image classification and also there are some challenges like data overfitting, environment like background color, structure, etc.

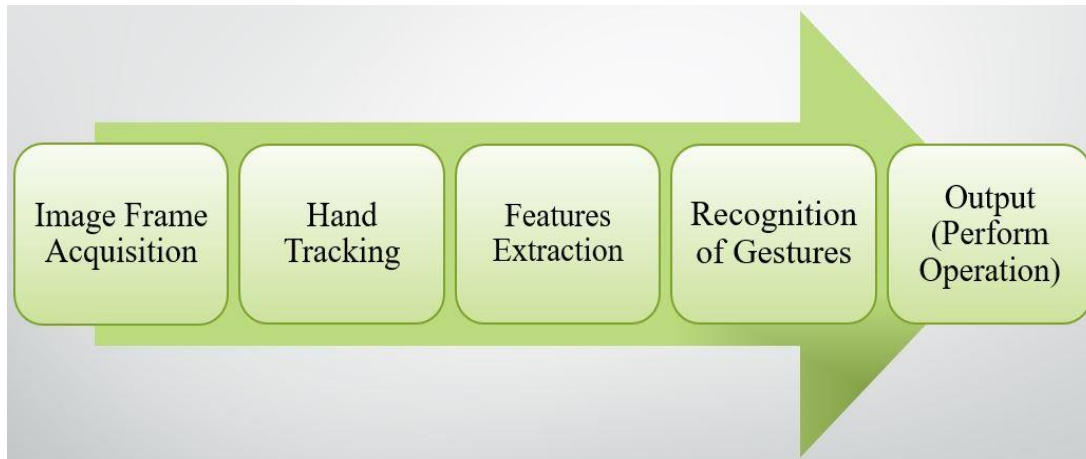


Figure 1.1.1 Methodology of Proposed System

1.2 MARKET SURVEY

The gesture recognition system has emerged as a popular technology over traditional mechanical interaction technologies. The market is segmented on the basis of Type, Application, Technology, Product and geography. Assistive robotics, Sign language detection, Immersive gaming technology, smart TV, virtual controllers, Virtual mouse, etc.

1.3 RESEARCH GAP

- Most of the methods used Arduino and sensors, so it is costly, directly device webcam is used in very few methods.
- There might be miss-recognitions in case the background has elements that resemble the human skin.
- Hand should be within the range limit.
- Dataset overfitting is the main concern.

1.4 PROBLEM STATEMENT

The aim of this project is to develop an interface to enable users to control other their computers by using hand gestures and a webcam. Hence our problem statement becomes, “Controlling Computer using Hand Gestures.”

1.5 OBJECTIVES

- To build and train a model for feature extraction and classification of hand gestures.
- To use a trained model for classification of hand gestures and recognizing an action and perform an operation accordingly.
- To control a computer and perform related activities using hand gestures.

2. CONCEPTS AND METHODS

2.1 CONCEPTS

Python is an interpreted, high-level and general-purpose programming language. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects. Python is often described as a "batteries included" language due to its comprehensive standard library.

Numpy, pandas, Keras, Sklearn, Matplotlib are the open source libraries used in python programming language, used to perform different operations like Numpy is used for handling arrays and matrix operations, Keras is a deep learning API written in Python, running on top of the machine learning library Tensorflow, Keras uses Tensorflow for its backend, Sklearn (Scikit-learn formerly scikits.learn and also known as sklearn) is a free software library features different machine learning algorithms and Matplotlib is a plotting library for python and numerical mathematics extension Numpy.

Imutils are a series of convenience functions to make basic image processing functions such as translation, rotation, resizing, and displaying Matplotlib images easier with OpenCV and Python.

We used OpenCV library (Open-Source Computer Vision Library) which is an open-source computer vision and machine learning software library. OpenCV is a library of programming functions mainly aimed at real-time computer vision. In simple language it is library used for Image Processing. It is mainly used to do all the operation related to Images.

PyAutoGUI is essentially a Python package that works across Windows, MacOS X and Linux which provides the ability to simulate mouse cursor moves and clicks as well as keyboard button presses. Used to programmatically control the mouse and keyboard.

MediaPipe is an open source machine learning platform used for building complex and multi-model applied machine learning pipelines. It can be used to make cutting-edge machine learning models like face detection, multi-hand tracking, object detection, and tracking and many more.

2.2 SOME BASIC TERMINOLOGIES

Artificial Intelligence (AI) - as name suggest can be interpreted to mean incorporating human intelligence to machines. AI is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans and animals.

Machine Learning (ML) - ML is the study of computer algorithms that improve automatically through experience. It is seen as a subset of artificial intelligence. Machine learning algorithms a build mathematical model in order to make predictions or decisions without being explicitly programmed to do so.

Deep Learning (DL) - Deep learning (also known as deep structured learning) is part of a broader family of machine learning methods based on artificial neural networks.

Computer Vision - Computer vision is the field that deals with how computers can gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to understand and automate tasks that the human visual system can do.

Image Processing - Image processing is a method to perform some important operations on an image. In order to get an enhanced high Quality image or to extract the most useful information from that.

Convolutional Neural Network (CNN) - In deep learning, a convolutional neural network is a class of deep neural networks, most commonly applied to analyzing visual imagery. A Convolutional neural network (CNN) is a neural network that has one or more convolutional layers and are used mainly for image processing, classification, etc. CNNs are used for image classification and recognition because of its high accuracy.

2.3 PLATFORM/TECHNOLOGY USED

- ❖ We developed this project in Python using Jupyter Notebook & Google Colab platform.
- ❖ Python (Tensorflow and Keras, Sklearn, Numpy, Pandas)
- ❖ OpenCV
- ❖ Deep Learning Algorithms

3. Literature Survey

In computer science and language technology, gesture recognition is an important topic which interpret human gesture through computer vision algorithms. The entire procedure of tracking gesture to their representation and converting them to some purposeful command is known as gesture recognition. Research and survey papers supported hand gestures have acquire so many alternative techniques, including those supported on sensor technology and computer vision.

The initial approach of communication with computer employing hand gesture was first projected by Myron W. Krueger in 1970 [1]. The purpose of the approach was attained and also the mouse cursor control was accomplished using an external webcam (Genius FaceCam 320), a software package that would paraphrase hand gestures and so turned the acknowledged gestures into OS commands that handled the mouse operations on the display screen of the computer [2]. Selecting hand gesture as an interface in HCI will permit the implementation of a good vary of applications with none physical contact with the computing environments [3]. Nowadays, majority of the HCI relies on devices like keyboard, or mouse, however an enlarging significance in a category of techniques based on computer vision has been came out because of skill to acknowledge human gestures in a habitual manner [4].

The primary aim of gesture recognition is to spot a specific human gesture and carry information to the computer. General objective is to create the computer acknowledged human gestures, to manage remotely with hand poses a good sort of devices [5]. The automated vision-based recognition of hand gesture for management of tools, such as digital TV, play stations and for sign language was take into account as a significant exploration topic lately. However the common issues of those works arise because of several problems, like the complicated and disturbing environments, tone color of skin and also the kind of static and dynamic hand gestures. Hand gestures recognition for TV management is suggested by [6]. During this system, just one gesture is employed to regulate TV by operating user hand. On the display, a hand icon seems that follows the hand of user.

In this paper [7], the actual HCI system that based on gestures and accept gestures uniquely operating one monocular camera and reach out the system to the HRI case has been

evolved. The came out system depends on a Convolution Neural Network classifier to grasp features and to acknowledge gestures. The Hidden Markov Model delivers as a crucial tool for the recognition of dynamic gestures in real time. The method employed HMM, works in actual and is built to operate in static environments. The approach is to make the use of LRB topology of HMM in association with the Baum Welch Algorithm for training and also the Forward and Viterbi Algorithms for testing and checking the input finding sequences and producing the most effective attainable state sequence for pattern recognition [8].

In this paper [9], the system is designed even it appears to be easy to use as compared to latest system or command based system however it is less powerful in spotting and recognition. Require to upgrade the system and attempt to construct further strong algorithm for both detection and recognition despite of the confused background environment and a usual lighting environment. Also require to upgrade the system for several additional categories of gestures as system is built for just six classes. However this system can use to manage applications like power point presentation, games, media player, windows picture manager etc. In this paper [10], hand gesture laptop makes the use of an Arduino Uno, Ultrasonic sensors and a laptop to perform the activities like controlling media, playback and volume. Arduino, Ultrasonic sensors, Python used for serial connection. This type of technology can be employed in the classroom for easier and interactive learning, immersive gaming, interacting with virtual objects on screen.

Arduino UNO and ultrasonic sensors based hand gesture to control a computer where they can play pause videos and scroll up and down pages [11]. This paper [12] suggest an efficient ultrasonic based hand gesture monitoring system which is designed with the help of Arduino microcontroller ATMEGA32. The hand gestures recognized effectively with ultrasonic sensors. It is proved that no extra hardware is needed to identify hand gestures and proved that simple inexpensive ultrasonic sensors can be employed to find several ranges to identify hand gestures. In this paper [13], Arduino UNO ARDUINO and python programming with wired ultrasonic sensor based hand gesture system to control a computer where they can zoom in/out and rotate the image is developed. This is successful trial of working of hand motion sensing system using sensors i.e. ultrasonic sensors and finger contact sensors and using in it to Arduino kits in wireless mode using radio frequency.

HCI for MS office and media player and have their own dataset, used skin colored based technique [14]. Application that switch in a Web browser, Web page scrolling, Task switching, Changing the slides of the presentation, Play/pause the video, Video forward and rewind is implemented. Arduino, PySerial, PyautoGUI, etc. used [15]. This project is built in order to make smart home appliance system. Two deep CNN architectures are evolved in this system which are revised from AlexNet and VGGNet, respectively [16]. Done the implementation of the system using Convolutional Neural Network and Back Propagation methodologies. They built a gesture controlling and hand recognition system for the one who are disables [17].

In this paper they have used the surface electromyography (sEMG) sensors with wearable hand gesture devices and mostly applied classifier is Artificial Neural network for sign language hand gesture. In this authors faced overfitting problem in the dataset [18]. The vision based real-time system with Python programming language and OpenCV libraries and Linux framework was implemented. A realtime human gesture recognition using an automated technology called Computer Vision is demonstrated using Linux operating system and Lenet is a CNN architecture used for training of the gestures [19]. This approach is relied on image comparison and motion recognition mechanism to try to do mouse indicator activities and make choice of the icon.

A Virtual gesture control mouse is an approach is developed to help the cursor of the mouse and perform its operations with the help of real-time camera. This method is also built on the basis of image comparison and motion recognition mechanism to try to do mouse indicator operations and choice of icons [20]. Developed a system that assists user friendly interaction such as full-body game and system generating a three-dimensional environment. It is an exploratory study on the gesture selection for operations and also settled an information recapture system to address several doubt phase of users [21].

Hyperparameter of CNN structures which are relied on Alex Net model, are advanced by heuristic optimization algorithms. This suggested approach is tried out on gestures language digits and at the same time Thomas Moeslund's gesture detection datasets. In this paper they achieved the 94.2% accuracy. Simultaneously, for Thomas Moeslund's gesture recognition dataset, the established method got 98.09% average accuracy performance and 94.33% average

recognition performance [22]. In this paper, they have built an algorithm for real-time hand gesture recognition using convolutional Neural Network (CNN) and they got an average accuracy of 98.76% on the dataset that they have proposed. Dataset consists of total nine hand gestures and 500 image samples for each gesture [23].

A wireless hand gestured controlled fan system is formed. Gyro sensor is employed so as to work out the modification in co-ordinates of the hand. Also used a microcontroller and particularly an Arduino in which coded functions have been run. Hand gesture detection achieved with testing accuracy of 98.61% [24]. To build the gesture detection system and also to minimize the unnecessary information of EMG signals and optimize the dimension of the signal, the Principal Component Analysis (PCA) and GRNN neural network are employed. In this system they have taken total nine static gestures as input observations and made the use of electromyography instrument to examine the characteristics of the signal. The all in all recognition accuracy of the system came to 95.1% [25].

4. PROJECT PLAN

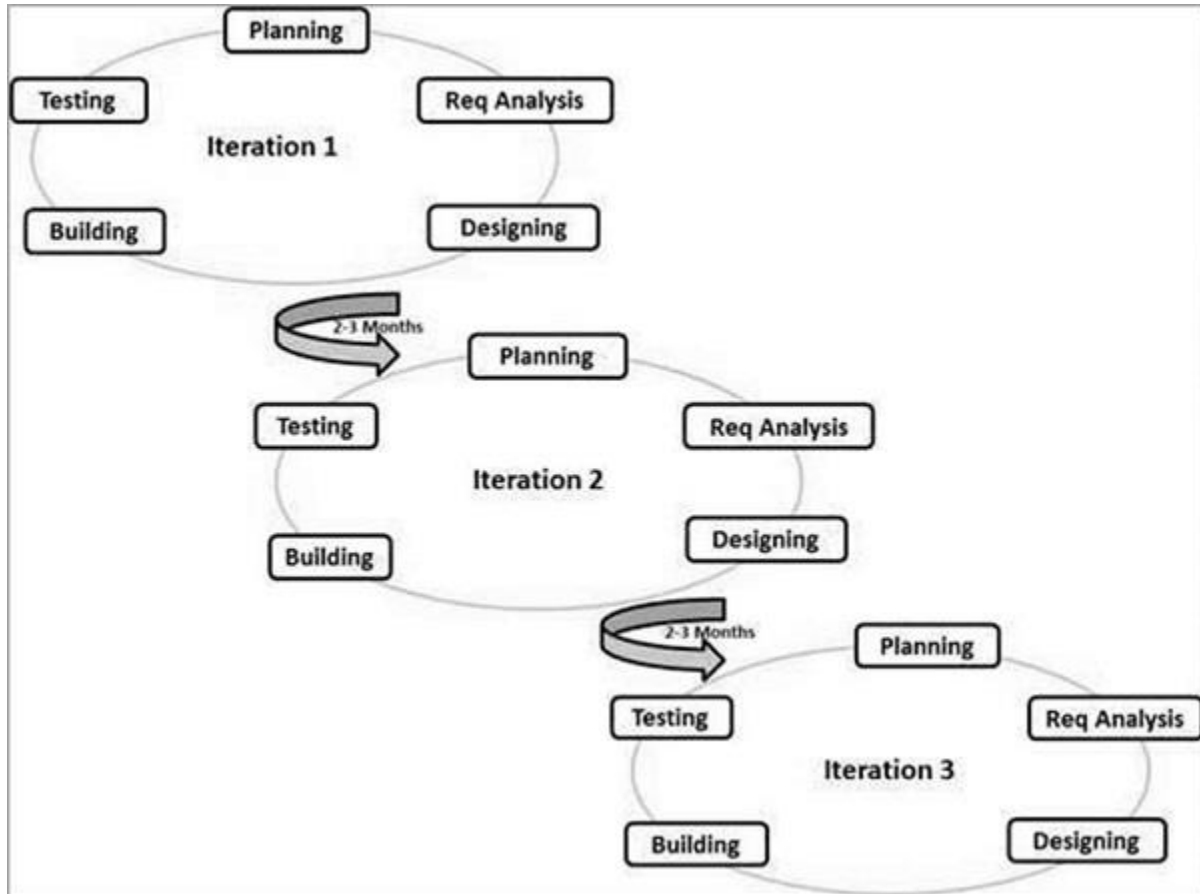


Figure 4.1 SDLC Agile Model

Agile SDLC model is a combination of iterative and incremental process models with focus on process adaptability and customer satisfaction by rapid delivery of working software product. Agile Methods break the product into small incremental builds. These builds are provided in iterations. Each iteration typically lasts from about one to three weeks. Every iteration involves cross functional teams working simultaneously on various areas like –

- Planning
- Requirements Analysis
- Design
- Building (Coding)
- Testing

- Deployment and Maintenance

At the end of the iteration, a working product is displayed to the customer and important stakeholders.

5. IMPLEMENTATION

5.1 DATASET

We tried to use available dataset, but we faced overfitting problem. Thus we create our own dataset for training the model. We took total 10 different hand gestures to perform activities like opening the WhatsApp, PowerPoint presentation, Microsoft Edge, Google Chrome, Video Player, etc. We took total 3000 images for training, 2000 images for testing and 500 images for validation.

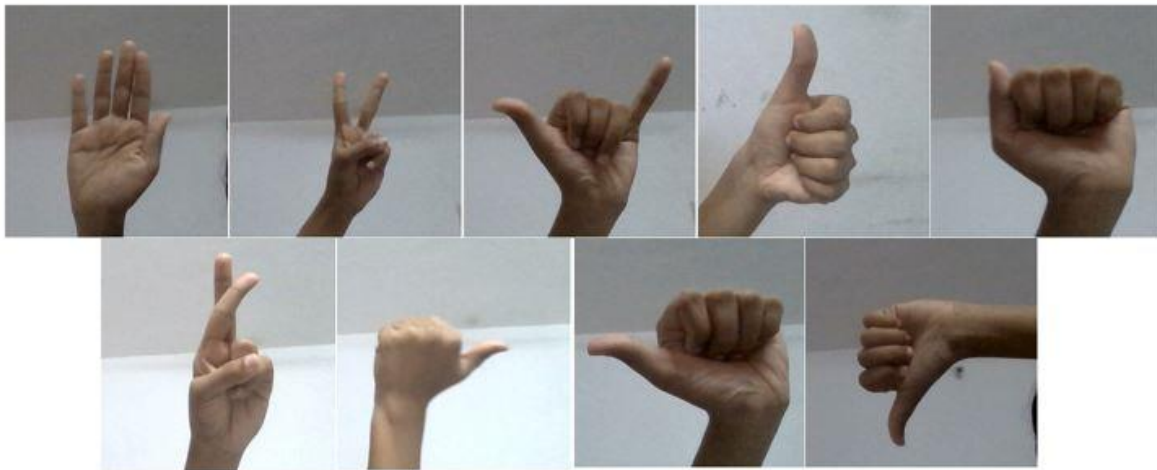


Figure 5.1.1 Dataset Images

5.2 IMPLEMENTATION

5.2.1 Train The Hand Gesture Recognizer Model

In this module first upon we will have a look on the dataset. We have total 3000 images for training, 2000 images for testing and 500 images for validation. We have performed image data augmentation, as someone said “Keras ImageDataGenerator is a gem!” it lets us augment our images in real-time while the model is still training! We can apply any random transformations on each image from dataset as it is passed to the model. This will not only make our model robust but will also save up on the overhead memory!

Now our next task is to train our hand gesture detector model. For that we have used Convolution Neural Network (CNN) architecture revised from VGGNet. CNN are used as black-boxes and their learned features remain hidden.

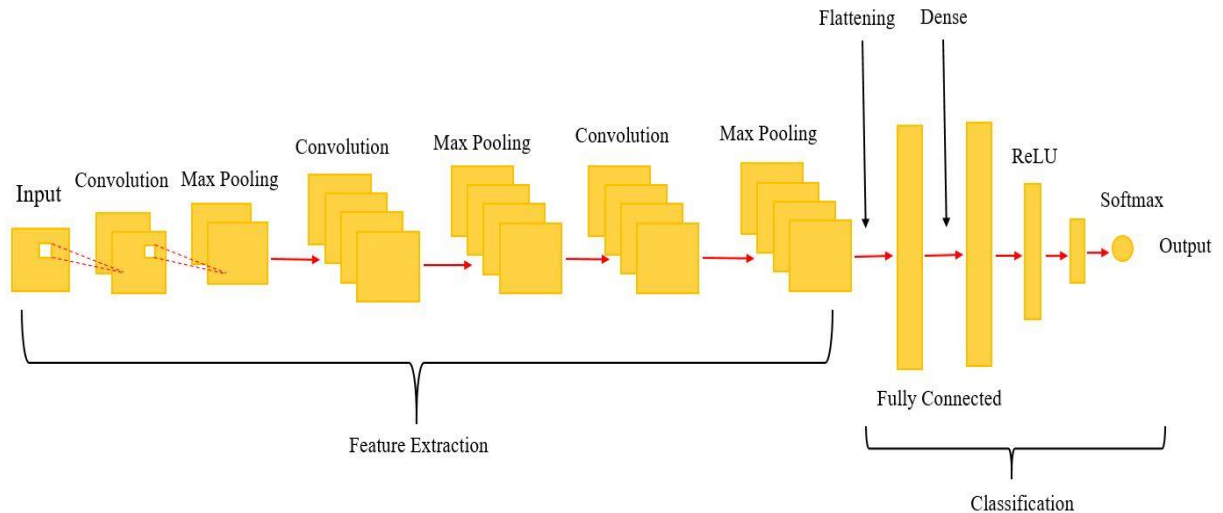


Figure 5.2.1.1 CNN Architecture

In this, for creating the model, we have used VGG19 (inception learning). VGG is a successor of AlexNet. VGG19 is a deep convolution neural network having convolution layer, pooling layer (max pool), fully connected layer, ReLU and softmax in its architecture. The first building block in our plan of attack is convolution operation. In this step, we will touch on feature detectors, which basically serve as the neural network's filters. We will also discuss feature maps, learning the parameters of such maps, how patterns are detected, the layers of detection, and how the findings are mapped out. The second step is pooling, pooling layers are used to reduce the dimensions of the feature maps. Thus, it reduces the number of parameters to learn and the amount of computation performed in the network. The pooling layer summarizes the features present in a region of the feature map generated by a convolution layer, however, will be a specific type of pooling; max pooling. Max pooling is a pooling operation that selects the maximum element from the region of the feature map covered by the filter. Thus, the output after max-pooling layer would be a feature map containing the most prominent features of the previous feature map. The third step will involve flattening, this will be a brief breakdown of the flattening process and how we move from pooled to flattened layers when working with Convolutional Neural Networks. Flatten is the function that converts the pooled feature map to a single column that is passed to the fully connected layer. Fully Connected layers, in this part, everything that we covered throughout the section will be merged together. Dense adds the fully connected layer to the neural network. The dense layer

is a neural network layer that is connected deeply, which means each neuron in the dense layer receives input from all neurons of its previous layer. Thus, they're densely connected. In other words, the dense layer is a fully connected layer, meaning all the neurons in a layer are connected to those in the next layer. Dropout is a technique where randomly selected neurons are ignored during training. They are “dropped-out” randomly. This means that their contribution to the activation of downstream neurons is temporally removed on the forward pass and any weight updates are not applied to the neuron on the backward pass. Rectified Linear Unit or ReLU, the rectified linear activation function or ReLU for short is a piecewise linear function that will output the input directly if it is positive, otherwise, it will output zero. Mathematically, it is defined as $y = \max(0, x)$. The softmax function is a function that turns a vector of K real values into a vector of K real values that sum to 1. The input values can be positive, negative, zero, or greater than one, but the softmax transforms them into values between 0 and 1, so that they can be interpreted as probabilities. If one of the inputs is small or negative, the softmax turns it into a small probability, and if an input is large, then it turns it into a large probability, but it will always remain between 0 and 1. For this reason it is usual to append a softmax function as the final layer of the neural network. Softmax is often used as the final layer in the network, for a classification task.

Now after creation of model, we compiled and trained our model. Then after checking the accuracy, visualization of results by plotting graphs, and saving the model we finished with the training module.

Here, process is taking the image of the gesture as an input using the web-cam and then compress the image by using the CNN algorithm to match the images in the dataset in order to detect the hand moment accurately. The captured image is preprocessed and a hand detector tries to filter out the hand image from the captured image. A CNN classifier is employed to recognize gestures from the processed image after feature extraction.

We used Adam optimizer because accuracy rate of this optimizer is much better than others. The whole workflow of our proposed system is explained through following image,

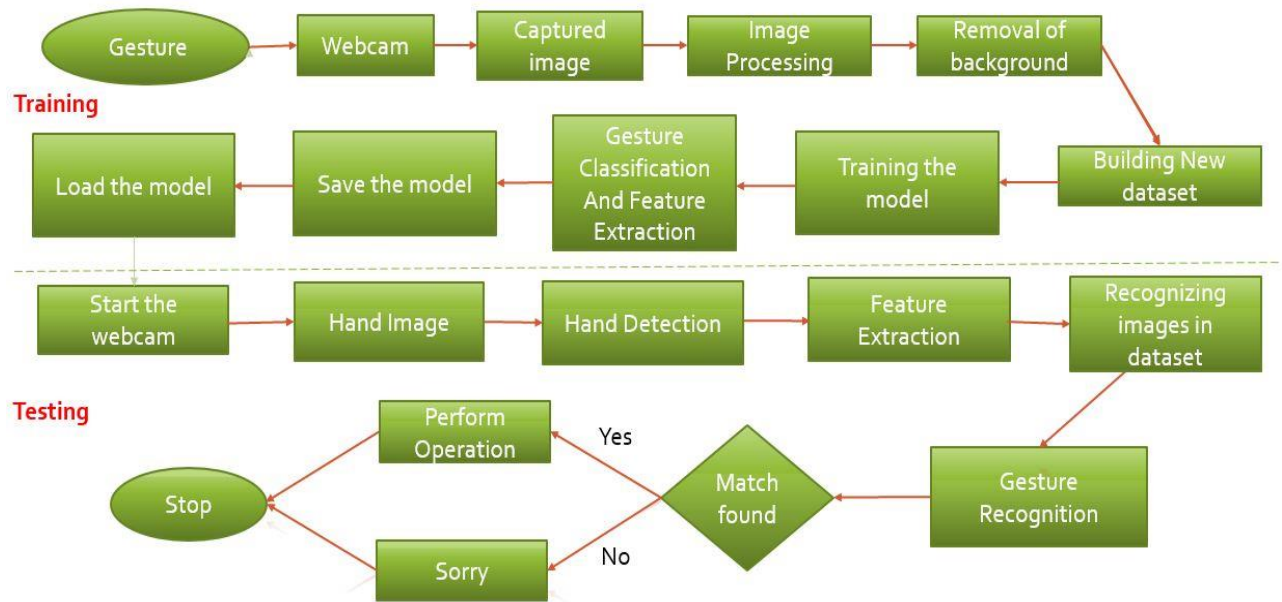


Figure 5.2.1.3 Workflow of Proposed System

5.2.2 Recognize The Hand Gesture And Perform Actions

Now we have trained model. This model will detect the hand, then features will be extracted and if gesture is recognized then particular operation will be perform. We have also added one more functionality in our implementation that tells user which action is performed. We developed GUI for our application.

Now our defined modules are completed and our trained model is detecting the hand and performing specific actions. Thus, we successfully implemented defined problem statement.

6. RESULT

Following are the two plots of train-val accuracy and train-val loss. We got 80.40% validation accuracy.

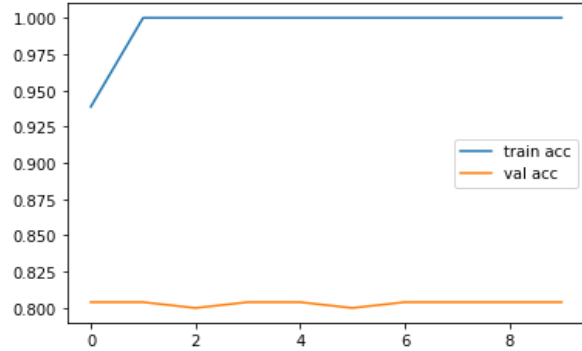


Figure 6.1 Train vs Val accuracy

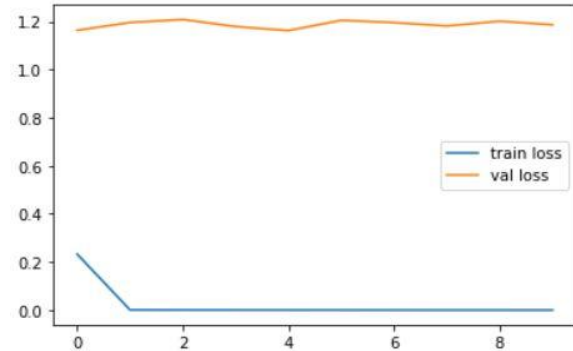


Figure 6.2 Train vs Val loss

We achieved following result of evaluation metrics. Our model is confused in some of the gestures like cross, scissor, up, etc. We achieved 85.90% accuracy at the time of testing our model. Following are the classification reports we got.

	precision	recall	f1-score	support
0	0.30	0.73	0.43	84
1	1.00	1.00	1.00	200
2	1.00	0.87	0.93	229
3	0.99	0.96	0.98	206
4	1.00	1.00	1.00	200
5	0.86	1.00	0.93	173
6	1.00	1.00	1.00	200
7	1.00	0.47	0.64	422
8	0.47	1.00	0.64	94
9	0.96	1.00	0.98	192
accuracy			0.86	2000
macro avg	0.86	0.90	0.85	2000
weighted avg	0.93	0.86	0.87	2000

Figure 6.3 Overall Result of Model

[61	0	0	0	0	0	0	23	0]
[0	200	0	0	0	0	0	0	0]
[0	0	200	2	0	27	0	0	0]
[0	0	0	198	0	0	0	0	8]
[0	0	0	0	200	0	0	0	0]
[0	0	0	0	0	173	0	0	0]
[0	0	0	0	0	0	200	0	0]
[0	0	0	0	0	0	0	0	94]
[139	0	0	0	0	0	0	200	83]
[0	0	0	0	0	0	0	0	192]

Figure 6.4 Confusion Matrix

Below are the images of GUI of our system,



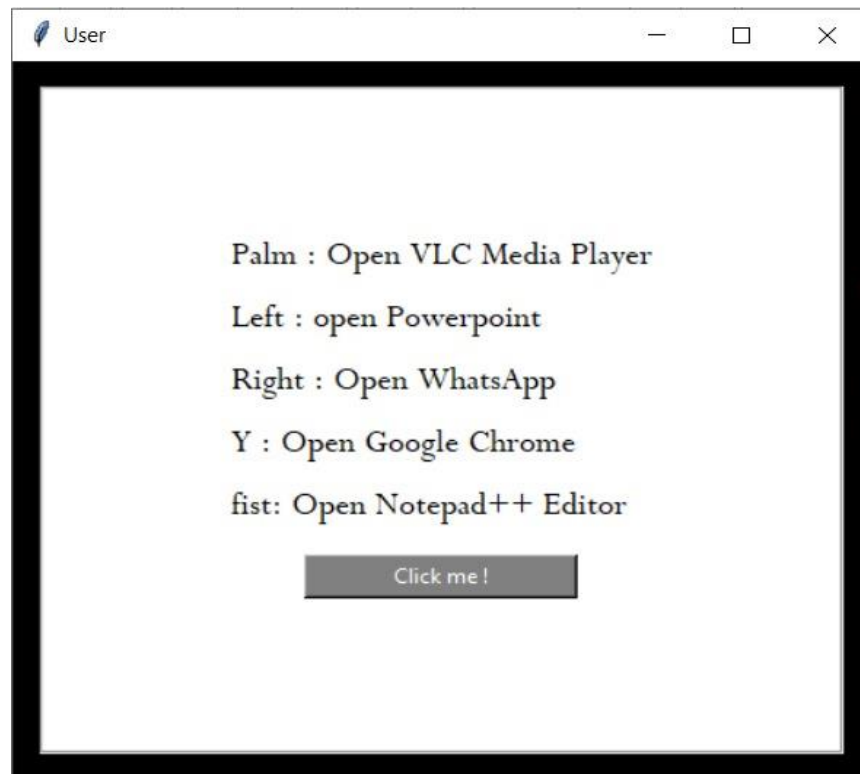


Figure 6.5 GUI of System

Below are the some demons of the developed system. We assigned total 10 hand gestures to perform different operations.



Figure 6.6 Gesture Detecting Frame

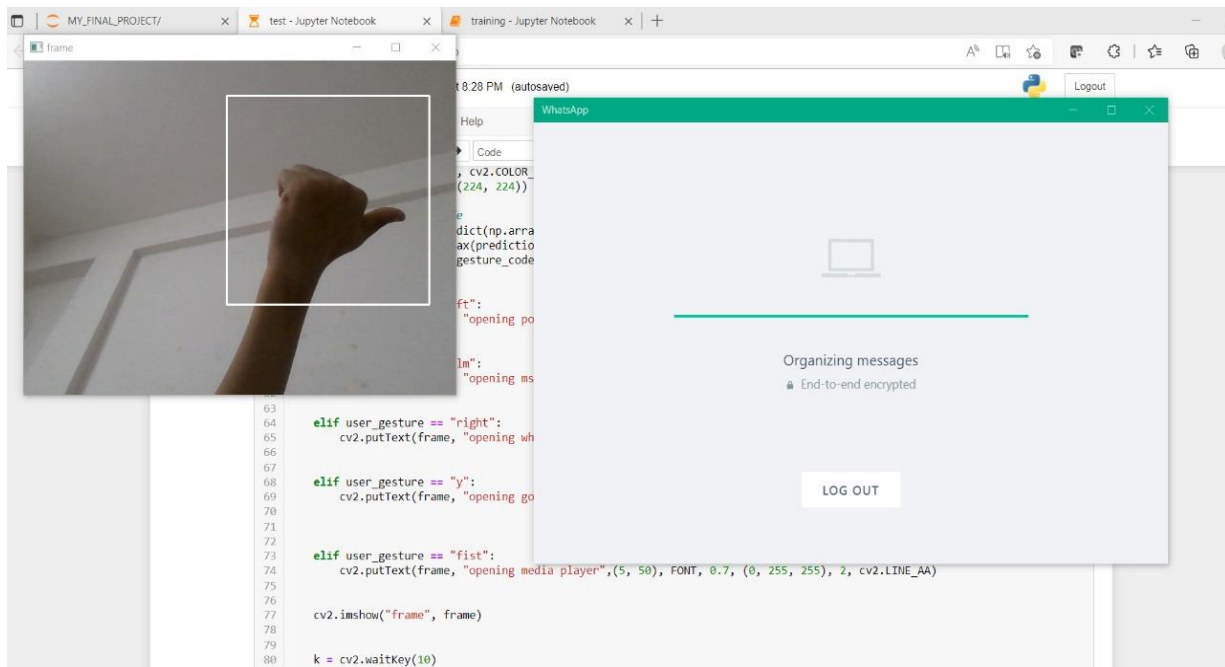


Figure 6.7 Opening the WhatsApp

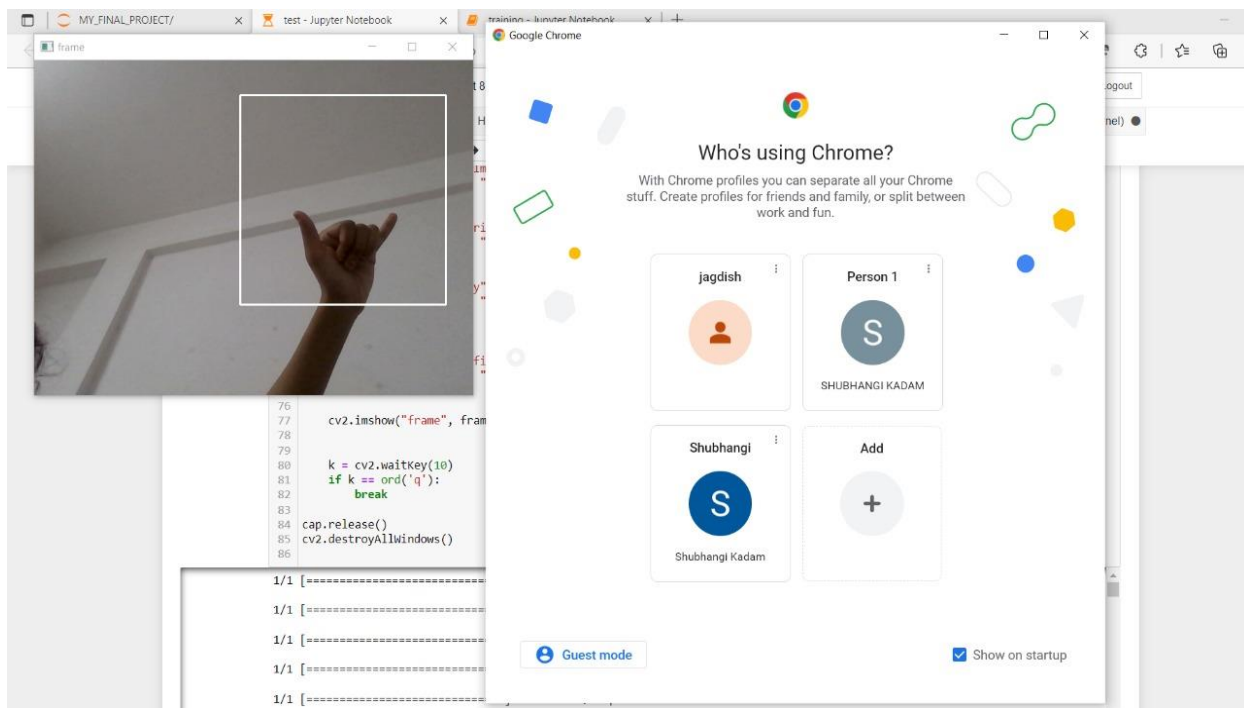


Figure 6.8 Opening the Google Chrome

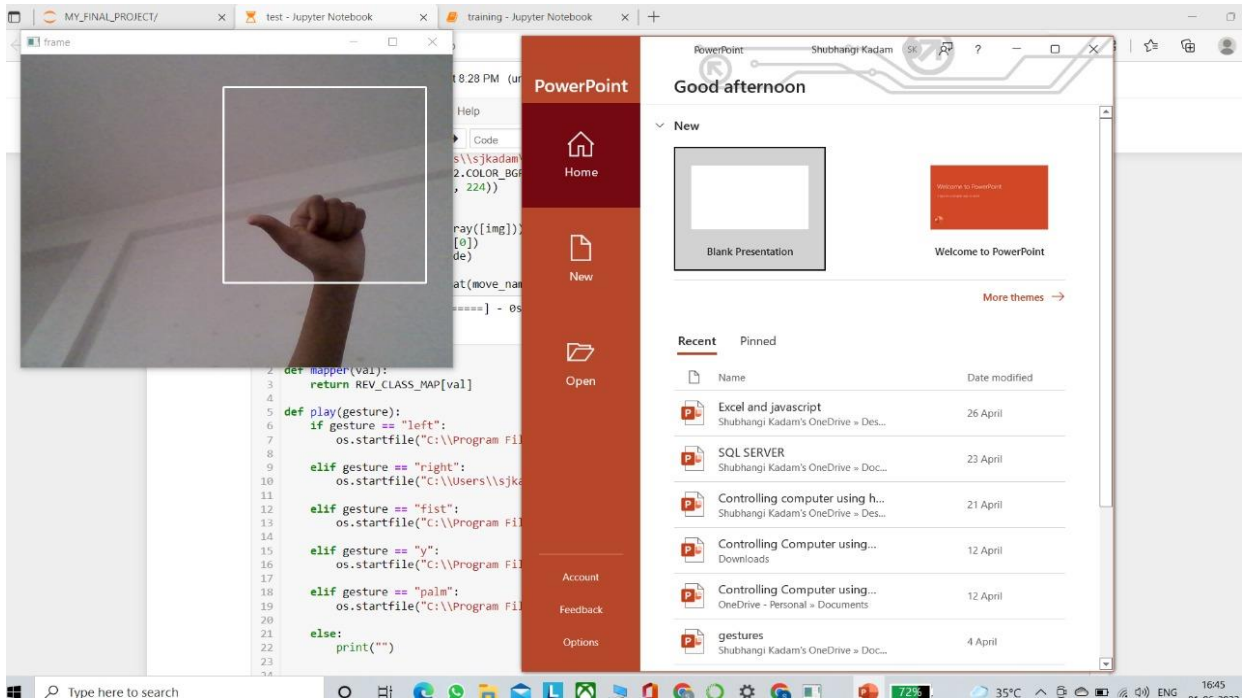


Figure 6.9 Opening the PowerPoint Presentation

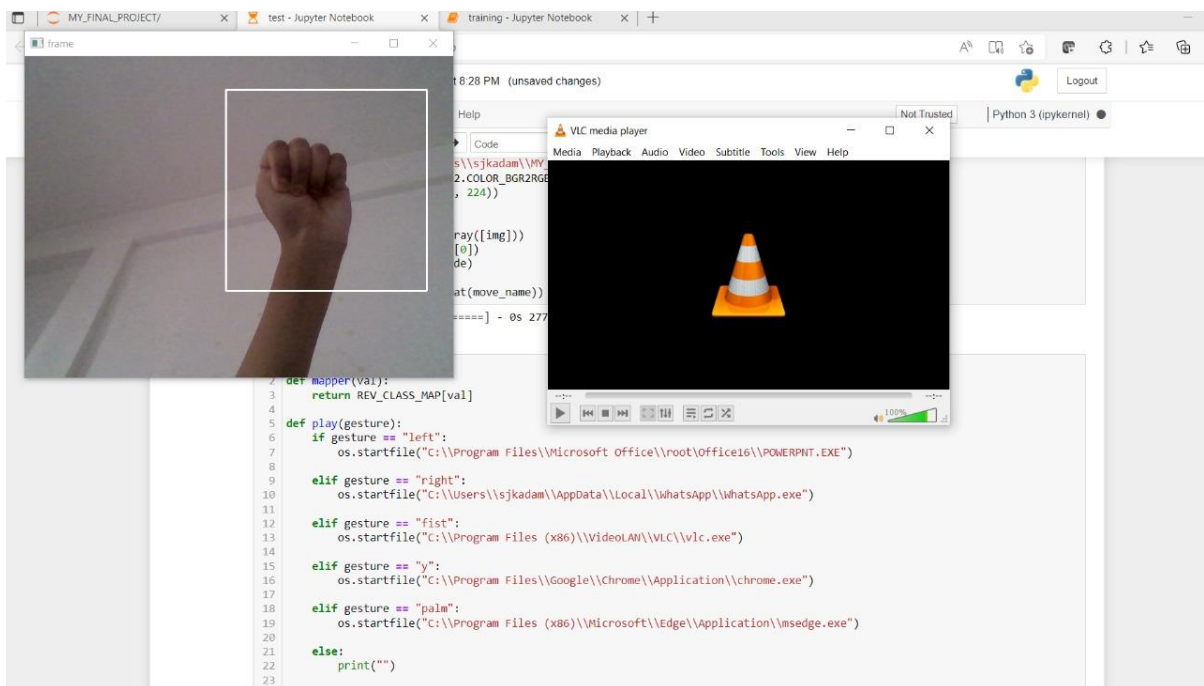


Figure 6.10 Opening the VLC

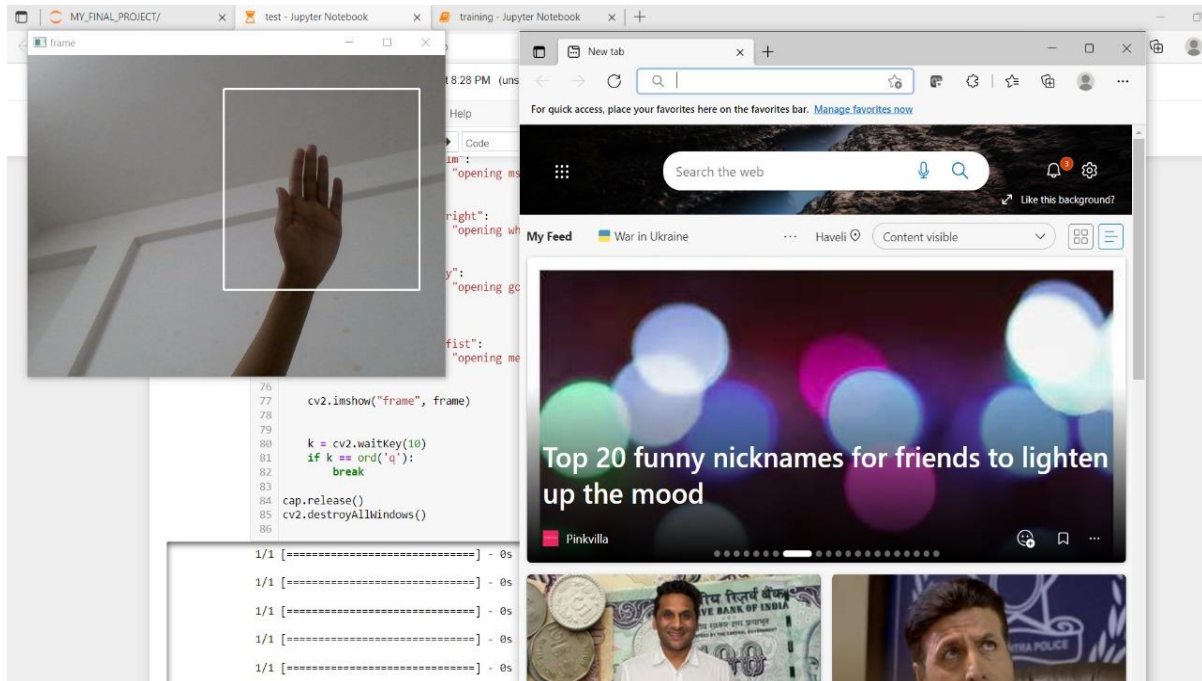


Figure 6.11 Opening the Microsoft Edge

7. CONCLUSION AND FUTURE WORK

With the growth of present technology, and as humans generally makes the use of hand movements that is hand gestures in the daily communication in order to make intentions more clear, hand gesture identification is treated to be a crucial portion of Human Computer Interaction (HCI), which provides devices the capability of detecting and classifying hand gestures, and perform activities subsequently.

Research and analysis in the field of hand gestures has become more popular and exciting. It also allows a way of natural and simple interaction. Standard interactive techniques based on several tools like a mouse, keyboard/touch pad, or touch screen, joystick for gaming and consoles for system management. In this paper we have discussed the overall review of gesture acquisition methods, the feature extraction process, the classification of hand gestures, the challenges that face researchers in the hand gesture recognition process.

In this application, we developed a deep learning model for controlling a computer using hand gestures with the help of Python and OpenCV. It is the cost effective model. We can define a project as creating a dataset, training a model and testing this model in real time. This project have limited scope, but in future we can add more operations like volume up/down, scroll up/down, swipe left/right and many more, and can possible to make completely hand gestures controlling device.

Hand gesture recognition used in many applications like HCI, robotics, sign language, digit and alphanumeric value, home automation, medical applications, gaming etc. Hand gestures recognition provides an interesting interaction field in a several different computer applications. With the development of today's technology, and as humans tend to naturally use hand gestures in their communication process to clarify their intentions, hand gesture recognition is considered to be an important part of Human Computer Interaction (HCI), which gives computers the ability of capturing and interpreting hand gestures, and executing commands afterwards.

It was a wonderful and learning experience for our team while working on this project. This journey of project involved various problems and challenges gave us feel of programmers and gained programming knowledge.

BIBLIOGRAPHY

- [1] M. KRUEGER, Artificial reality II, Addison-Wesley Reading (Ma), 1991.
- [2] Horatiu-Stefan Grif and Cornel Cristian Farcas, "Mouse Cursor Control System Based on Hand Gesture", 9th International Conference Interdisciplinarity in Engineering, INTER-ENG 2015, 8-9 October 2015, Tirgu Mures, Romania.
- [3] H.A JALAB, "Static hand Gesture recognition for human computer interaction", Asian Network for Scientific Information technology journal, pp. 1-7, 2012.
- [4] J C.MANRESA, R VARONA, R. MAS and F. PERALES, "Hand tracking and gesture recognition for humancomputer interaction", Electronic letters on computer vision and image analysis, vol. 5, pp. 96-104, 2005.
- [5] H. HASAN and S. ABDUL-KAREEM, "Static hand gesture recognition using neural networks", Artificial Intelligence Review, vol. 41, pp. 147-181, 2014.
- [6] W. FREEMAN and C. WEISSMAN, "Television control by hand gestures", Proc. of Intl. Workshop on Automatic Face and Gesture Recognition 1995, pp. 179-183, 1995.
- [7] Pei Xu, "A Real-time Hand Gesture Recognition and Human-Computer Interaction System", arXiv:1704.07296v1 [cs.CV] 24 Apr 2017.
- [8] J.R Pansare, Malvika Bansal, Shivin Saxena, Devendra Desale, "Gestuelle: A System to Recognize Dynamic Hand Gestures using Hidden Markov Model to Control Windows Applications", International Journal of Computer Applications (0975 – 8887) Volume 62– No.17, January 2013.
- [9] Ram Pratap Sharma and Gyanendra K. Verma, "Human Computer Interaction using Hand Gesture", Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015).
- [10] Rohit Mukherjee, Pradeep Swethen, Ruman Pasha, Sachin Singh Rawat, "Hand Gestured Controlled laptop using Arduino", International Journal of Management, Technology And Engineering, October 2018.
- [11] Ayushi Bhawal, Debaparna Dasgupta, Arka Ghosh, Koyena Mitra, Surajit Basak, "Arduino Based Hand Gesture Control of Computer", IJESC, Volume 10, Issue No.6, June 2020. [
- [12] Udit Kumar, Sanjana Kintali, Kolla Sai Latha, Asraf Ali, N. Suresh Kumar, "Hand Gesture Controlled Laptop Using Arduino", April 2020.
- [13] Sarita K., Gavale Yogesh, S. Jadhav, "HAND GESTURE DETECTION USING ARDUINO AND PYTHON FOR SCREEN CONTROL", International Journal of Engineering Applied Sciences and Technology, 2020, Vol. 5, Issue 3, ISSN No. 2455-2143, July 2020.
- [14] Ram Pratap sharma, Gyanendra Varma, "Human computer interaction using hand gesture", 2015.
- [15] J. S. Vimali, Senduru Srinivasulu, J. Jabez, S. Gowr, "Hand Gesture Recognition Control for Computers Using Arduino", 2020.
- [16] Chung, Hung-Yuan, Yao-Liang Chung, and Wei-Feng Tsai. "An efficient hand gesture recognition system based on deep CNN." 2019 IEEE International Conference on Industrial Technology (ICIT). IEEE, 2019.
- [17] Varun, Kollipara Sai, I. Puneeth, and T. Prem Jacob. "Hand gesture recognition and implementation for disables using CNN'S." 2019 International Conference on Communication and Signal Processing (ICCSP). IEEE, 2019.

- [18] Yasen, Mais, and Shaidah Jusoh. "A systematic review on hand gesture recognition techniques, challenges and applications." *PeerJ Computer Science* 5 (2019): e218.
- [19] Suresh, Malavika, Avigyan Sinha, and R. P. Aneesh. "RealTime Hand Gesture Recognition Using Deep Learning." *International Journal of Innovations and Implementations in Engineering* 1 (2019).
- [20] Shibly, Kabid Hassan, et al. "Design and development of hand gesture based virtual mouse." 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT). IEEE, 2019.
- [21] Wu, Huiyue, et al. "The gesture disagreement problem in free-hand gesture interaction." *International Journal of Human–Computer Interaction* 35.12 (2019): 1102-1114.
- [22] Ozcan, Tayyip, and Alper Basturk. "Transfer learningbased convolutional neural networks with heuristic optimization for hand gesture recognition." *Neural Computing and Applications* 31.12 (2019): 8955-8970.
- [23] Zhan, Felix. "Hand gesture recognition with convolution neural networks." 2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI). IEEE, 2019.
- [24] Alon, Alvin Sarraga, and Julie Ann B. Susa. "Wireless Hand Gesture Recognition for an Automatic Fan Speed Control System: Rule-Based Approach." 2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA). IEEE, 2020.
- [25] Qi, Jinxian, et al. "Surface EMG hand gesture recognition system based on PCA and GRNN." *Neural Computing and Applications* 32.10 (2020): 6343-6351.

ANNEXURE A: List of Publications and Research Paper (In its Original formats)

1. Publication Name: International Research Journal of Engineering and Technology (IRJET)
Publication site: <https://www.irjet.net/>
Paper Link: <https://www.irjet.net/archives/V9/i2/IRJET-V9I2181.pdf>
Status: **Published**



2. Publication Name: Resbee Publisher Multimedia Research Journal
Publication site: <https://publisher.resbee.org/Journal-publishing-service.php?journal=mr>
Paper Link:
Status: **Submitted**

3. Publication Name: International Research Journal of Engineering and Technology (IRJET)
Publication site: <https://www.irjet.net/>
Paper Link:
Status: **Submitted**

ANNEXURE B: Plagiarism Report



PLAGIARISM SCAN REPORT

	2% Plagiarised		98% Unique	Date	2022-06-08
				Words	834
				Characters	5861

Content Checked For Plagiarism

Controlling Computer using Hand Gestures

Pradnya Kedari¹, Shubhangi Kadam², Rajesh Prasad³

¹Student, Department of Computer Science and Engineering, MIT School of Engineering, MIT Art, Design and Technology, Pune, Maharashtra, India

²Student, Department of Computer Science and Engineering, MIT School of Engineering, MIT Art, Design and Technology, Pune, Maharashtra, India

³Professor, Department of Computer Science and Engineering, MIT School of Engineering, MIT Art, Design and Technology, Pune, Maharashtra, India

-----***-----

Abstract – Human computer interaction platform have many ways to implement as webcams and other devices like sensors are inexpensive and can get easily available in the market. The most powerful way for communication between human and machine is through gesture. For higher conveyance between the human and machine/computer to convey information, hand gesture system is very useful. Hand gestures are a sort of nonverbal type to communicate that may be employed in several fields. Research and survey papers included hand gestures applications have acquire different alternative techniques, including those supported on sensor technology and computer vision.

In this system, we aimed to build a real-time gesture recognition system using hand gestures. Particularly, we will use the convolutional neural network (CNN) in throughout the process. This application presents a hand gesture-based system to control a computer that is performing different operations using neural network. Our application is defined in five phases, Image frame acquisition, Hand tracking, Features extraction, Recognition of gestures and Classification (perform desired operation). An image from the webcam will be captured, and so hand detection, hand shape features extraction, and hand gesture recognition are done.

Key Words: Deep Learning, Computer Vision, Hand Gestures, Convolutional Neural Network, Python, OpenCV

1. INTRODUCTION

Gesture recognition is a popular and in-demand analysis field in Human Computer Interaction technology. It has several employments in virtual environment management, medical applications, sign language translation, robot control, music creation, or home automation. There has been a special significance recently on HCI study. Hand is the one which is most helpful communication tool in several body parts, because of its expertise. The word gesture is employed for several cases involving human motion particularly of the hands, arms, and face, just some of these are informative.

The convolutional neural networks are the most popular employed technique for the image classification task. An image classifier takes an input image, or input sequence of images and categories them into one among the possible classes that it was trained to classify. They have applications in different fields such as medical domain, self-driving cars, educational domain, fraud detection, defense, etc. There are several techniques and algorithms for image classification task and also

there are some challenges like data overfitting. During this project Controlling Computer using Hand Gestures, we are aimed to make a real-time application using OpenCV and Python.

OpenCV is a real-time open-source computer vision and image-processing library. We'll use it with the help of the OpenCV python package.

Fig 1. Methodology of Proposed System

1.1 MARKET SURVEY

Over the traditional mechanical communication technologies, gesture recognition system has become known as a most popular technology. The domain market is divided on the different basis like Technology, Type, Practice, Product, Use and Geography. Assistive robotics, Sign language detection, Immersive gaming technology, smart TV, virtual controllers, Virtual mouse, etc.

1.2 RESEARCH GAP

- Most of the methods used Arduino and sensors, directly device webcam is used in very few methods.
- There might be miss-recognitions of gestures in case the background environment has elements that appears like human skin.
- Hand should be within the range limit.
- Dataset overfitting is the main concern.

2. LITERATURE SURVEY

In computer science and technology era, gesture recognition is a crucial field which can use to translate human gesture using different computer vision techniques and algorithms. There are numerous human body motions which can create gestures but the most common type of gesture generation stand up from the face and hands. The complete policy of tracking gesture to their representation and changing them to some useful command is refer as gesture recognition.

Different techniques and methods has been employed for the design and development of such kind of task.

The starting approach of interaction with computer using hand gesture was first projected by Myron W. Krueger in 1970 [1]. The aim of the perspective was achieved and also the mouse cursor control was performed with the help of an external webcam (Genius FaceCam 320), a software package that would interpret hand gestures and then turned the recognized gestures into OS commands that controlled the mouse actions on the display screen of the computer device [2]. Choosing hand gesture as a communication tool in HCI will allow the development of a good vary of applications with the absence of physical contact with the computing devices [3]. At present, most of the HCI depends on the devices such as keyboard, or mouse, however the growing importance in a category of methods and techniques based on computer vision has been popular because of skill to recognize human gestures in a simple manner [4].

Matched Source



Similarity 4%

Title:[www.irjet.net › archives › V9A Survey Paper on Controlling Computer using Hand Gestures](http://www.irjet.net/archives/V9A/Survey%20Paper%20on%20Controlling%20Computer%20using%20Hand%20Gestures.pdf)

OpenCV is a real-time open-source computer vision and image-processing library. We'll use it via the OpenCV python package. There has been a special significance recently on Human Computer Interaction (HCI) study to form convenient interfaces by directly using common communication and handling expertise of humans.

<https://www.irjet.net/archives/V9/i2/IRJET-V9I2181.pdf/>

PLAGIARISM SCAN REPORT

 <div>2% Plagiarised</div>	 <div>98% Unique</div>	<div>Date</div> <div>2022-06-08</div>
		<div>Words</div> <div>848</div>
		<div>Characters</div> <div>5295</div>

Content Checked For Plagiarism

Detection and recognition of a particular human body gesture and carry information to the computer is the main objective of gesture recognition. Overall objective of this system was to create the human gestures which can be admit by computer device to control a good sort of devices that are at distant using different hand poses [5]. Hand gesture recognition based on the robotic computer vision to handle the devices such as digital TV, play stations, etc. Hand gesture recognition for sign language was considered as a weighty research area lately. But because of different issues, like skin tone color difference, the complex and disturbing environments and also the different static and dynamic hand gestures, the common problem of that system raised. Hand gestures recognition for management of TV is recommended by [6]. In this, only one hand moment is used to control TV. A hand picture looks like icon that follows the hand movements of the user appear on the screen display of TV. In this paper [7], the actual human computer interface that is HCI model which is based on the hand gestures and accept gestures in a unique way to operate using monocular camera and assist the system to the HRI case has been developed. The evolved system relies on a classifier based on Convolution Neural Network to extract features and to recognize particular hand gesture.

The HMM that is Hidden Markov Model considered as a crucial tool for the dynamic gestures recognition in actual. The system employed HMM, operate in present and the general aim to build this system is to operate in static environments. The proposed methodology was for training, to employ the topology named LRB of HMM with the Baum Welch Algorithm and also for testing, the Forward and Viterbi Algorithms and checking the input sequences and building the maximum productive achievable pattern recognition state sequence [8]. In this paper [9], even the developed model seems to be easy to handle as compared to the newest available system or command based system but drawback is developed system is less powerful in spotting the gestures and recognition of the same. So despite of the complex environmental background and a normal lighting environment background, exist system need to improve and require to build further a good network for gesture recognition. This system is built for total six classes. However this exist model can be used to control operations such as power point presentation, windows picture manager, media player, games etc. In this paper [10], using an Arduino Uno and Ultrasonic sensors, operations such as handling media player, volume increase/decrease are performed on laptop. Arduino, Ultrasonic sensors, Python used for serial connection. For interactive and effective learning, such type of system can be used in the teaching classrooms.

Hand Gesture recognition system based on devices like Arduino UNO and several ultrasonic sensors to manage a device where they can control VLC by involving operations like play and pause videos and also for page scroll up and down [11]. This paper [12] suggest a convenient hand gesture monitoring system based on ultrasonic sensors, which is built using Arduino microcontroller ATMEGA32. It is claimed that extra hardware is not require to classify hand gestures and also claimed that simple low cost ultrasonic sensors can be used to notice different range limits to identify hand gestures. In this paper [13], hand gesture system relied on Arduino UNO and python programming with wired ultrasonic sensor is developed to manage a device and they included operations like zoom in/out and image rotation, etc. This trial is successful trial of working of hand motion sensing system using sensors and Arduino kits in wireless mode radio frequency. Hand gesture recognition system for Microsoft Office and media player with their own dataset is developed. [14].

3.1 DATASET

We tried to use available dataset, but we faced overfitting problem. Thus we create our own dataset for training the model. We took total 10 different hand gestures to perform activities like opening the WhatsApp, PowerPoint presentation, Microsoft Word, Microsoft Edge, Google Chrome, Video Player, Xbox, and Paint etc. We took total 3000 images for training, 2000 images for testing and 500 images for validation.

Fig 2. Sample Dataset Images

3.2 IMPLEMENTATION

a. Train The Hand Gesture Recognizer Model

In this module first upon have a glance on the dataset. We have total 3000 images for training, 2000 images for testing and 500 images for validation. We have performed image data augmentation, as somebody said "Keras ImageDataGenerator is a gem!" it lets us augment our images in real-time while the model is still training. We can apply any random transformations on each image from dataset as it is passed to the model. This will not only make our model powerful but will also lay aside on the overhead memory!

Now our next task is to train our hand gesture detector model. For that we have used Convolution Neural Network (CNN) architecture revised from Squeezenet and VGGNet. Learned features of CNN remain hidden and thus it is used as black-box.

Fig 1. CNN Architecture



Matched Source

Similarity 4%

Title: [Freespace_Segmentation-Ford_Otosan_Intern ... - GitHub](#)

https://github.com/recepayddogdu/Freespace_Segmentation-Ford_Otosan_Intern/blob/main/Questions_and_Answers.md

PLAGIARISM SCAN REPORT

	7% Plagiarised		93% Unique	Date	2022-06-08
				Words	876
				Characters	5738

Content Checked For Plagiarism

In this, for creating the model, we have used VGG19 (inception learning). VGG is a successor of AlexNet. VGG19 is a deep convolution neural network having convolution layer, pooling layer (max pool), fully connected layer, ReLU and softmax in its architecture. The first step is convolution operation. In this step, feature detectors are mapped, which basically serve as the neural network's filters. The second step is pooling, pooling layers are used to minimize the dimensions of the feature maps. However, will use a specific type of pooling, max pooling. Max pooling selects maximum features to the next layer. The third step will involve flattening, flatten is the function that converts the pooled feature map to a single column that is passed to the fully connected layer. Fully Connected layers, in this part, everything that we covered throughout the section will be merged together. Dense adds the fully connected layer to the neural network. In the dense layer neurons are supposed to connect deeply. Rectified Linear Unit or ReLU is a linear function that will output the input directly if it is positive, otherwise, it will output zero. Mathematically, it is defined as $y = \max(0, x)$. The softmax function transforms input value into values between 0 and 1, so that they can be interpreted as probabilities. Final output of this layer will always remain between 0 and 1.

For this reason it is usual to append a softmax function as the final layer of the neural network. Softmax is often used as the final layer in the network, for a classification task.

We used Adam optimizer because accuracy rate of this optimizer is much better than others. The whole workflow of our proposed system is explained through following image Fig 2,

Fig 2. Workflow of Proposed System

Now after creation of model, we compiled and trained our model. Then after checking the accuracy, visualization of results by plotting graphs, and saving the model we finished with the training module. Here, complete task is taking the image of the hand gesture as an input using the webcam and then compress the image by using the CNN algorithm to match the images in the dataset in order to detect the hand moment accurately. The captured image is preprocessed, and a hand detector tries to filter out the hand image from the captured image. A CNN classifier is used to recognize gestures from the processed image after feature extraction.

b. Recognize The Hand Gesture and Perform Operations

Now we have trained model. This model will detect the hand, then features will be extracted and if gesture is recognized then particular operation will be perform. We also added one more functionality in our implementation that tells user which action is performed.

4. RESULT AND ANALYSIS

Following are the two plots of train-val accuracy and train-val loss. We got 80.40% validation accuracy.

Fig 3. Train vs Val accuracy Fig 6. Train vs Val accuracy

We achieved following result of evaluation metrics. Our model is confused in some of the gestures like cross, scissor, up, etc. We achieved 85.90% accuracy at the time of testing our model. Following are the classification reports we got.

Fig 7. Overall Result of Model Fig 8. Confusion Matrix

Below are some demons of the developed system. We assigned total 10 hand gestures to perform different operations.

Fig 9. Gesture Detecting Frame Fig 10. Opening the WhatsApp

Fig 11. Opening the Google Chrome Fig 12. Opening the PowerPoint Presentation

Fig 13. Opening the VLC Fig 14. Opening the Microsoft Edge

5. CONCLUSION

With the growth of present technology, and as humans generally makes the use of hand movements that is hand gestures in the daily communication in order to make intentions more clear, hand gesture identification is treated to be a crucial portion of Human Computer Interaction (HCI), which provides devices the capability of detecting and classifying hand gestures, and perform activities subsequently. Research and analysis in the field of hand gestures has become more popular and exciting. It also allows a way of natural and simple interaction. Standard interactive techniques based on several tools like a mouse, keyboard/touch pad, or touch screen, joystick for gaming and consoles for system management. In this paper we also have discussed the overall review of gesture acquisition methods, the feature extraction process, the classification of hand gestures, the challenges that face researchers in the hand gesture recognition process. In this application, we developed a deep learning model for controlling a computer using hand gestures with the help of Python and OpenCV. It is the cost-effective model as we are not using any extra devices and sensors. We can define a project as creating a suitable dataset, training the model and testing this model in real time. This project has limited scope, we assigned total 10 hand gestures to perform different operations, but in future we can add more operations like volume up/down, scroll up/down, swipe left/right and many more, and can be possible to make completely hand gestures controlling device. Hand gesture recognition used in many different applications like robotics, sign language recognition, HCI, digit and alphanumeric value, home automation, medical applications, gaming etc. Hand gestures recognition provides an interesting interaction field in a several different computer science applications.

Matched Source

Similarity 9%

Title: [heartbeat.comet.ml > a-beginners-guide-toConvolutional Neural Networks: An Intro Tutorial | by Derrick ...](#)

May 08, 2018 · Dense adds the fully connected layer to the neural network. from keras.models import Sequential from keras.layers import Convolution2D from keras.layers import MaxPooling2D from keras.layers import Flatten from keras.layers import Dense 2) Initializing the neural network To initialize the neural network we create an object of the Sequential class. <https://heartbeat.comet.ml/a-beginners-guide-to-convolutional-neural-networks-cnn-cf26c5ee17ed/>

Similarity 9%

Title: [General idea about Activation Function. - LinkedIn](#)

<https://www.linkedin.com/pulse/general-idea-activation-function-swagat-panda#:~:text=ReLU%20stands%20for%20Rectified%20Linear,usually%20a%20good%20first%20choice.>

Similarity 7%

Title: [Freespace_Segmentation-Ford_Otosan_Intern ... - GitHub](#)

https://github.com/recepayddogdu/Freespace_Segmentation-Ford_Otosan_Intern/blob/main/Questions_and_Answers.md
