HAND GESTURE DETECTION AND RECOGNITION

A PROJECT REPORT

Submitted by

AASHISH YADAV (21BAI10116)
UJJWAL SAGAR (21BAI10047)
NISHANT SETH (21BAI10168)
DIVYA AGARWAL (21BAI10159)
SHUBHAM SINGH (21BAI10247)

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE ENGINEERING WITH SPECIALIZATION IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING

VIT BHOPAL UNIVERSITY

KOTHRIKALAN, SEHORE

MADHYA PRADESH - 466114

FEB 2023

VIT BHOPAL UNIVERSITY, KOTHRIKALAN, SEHORE
MADHYA PRADESH – 466114

BONAFIDE CERTIFICATE

Certified that this project report titled **"HAND GESTURE RECOGNITION AND DETECTION USING MACHINE LEARNING"** is the bonafide work of "**AASHISH YADAV (Register No :21BAI10116), UJJWAL SAGAR (Register No :21BAI10047), NISHANT SETH (Register No :21BAI10168), DIVYA AGARWAL (Register No :21BAI10159), SHUBHAM SINGH (Register No :21BAI10247)"** who carried out the project work under my supervision. Certified further that to the best of my knowledge the work reported at this time does not form part of any other project/research work based on which a degree or award was conferred on an earlier occasion on this or any other candidate.

PROGRAM CHAIR
Suthir S, Assistant Professor (Senior)
School of Computer Science and Engineering
VIT BHOPAL UNIVERSITY

PROJECT GUIDE
Dr. Komarasamy G, Assistant Professor
School of Computer Science and Engineering
VIT BHOPAL UNIVERSITY

The Project Exhibition II Examination is held on 4 MAR 2023

ACKNOWLEDGEMENT

First and foremost, I would like to thank the Lord Almighty for His presence and immense blessings throughout the project work.

I wish to express my heartfelt gratitude to Dr. Suthir, Head of the Department, School of Aeronautical Science for much of his valuable support encouragement in carrying out this work.

I would like to thank my internal guide Mr. Dr. Komarasamy G, for continually guiding and actively participating in my project, giving valuable suggestions to complete the project work.

I would like to thank all the technical and teaching staff of the School of Aeronautical Science, who extended directly or indirectly all support.

Last, but not least, I am deeply indebted to my parents who have been the greatest support while I worked day and night for the project to make it a success.

LIST OF TABLES

ABSTRACT

Hand gesture recognition is a technology that uses computer vision techniques to detect and interpret hand gestures made by a person.

These gestures can be used as a form of non-verbal communication and input for various types of systems, such as human-computer interaction, sign language recognition, and gaming.

A typical hand gesture recognition system includes a camera or depth sensor to capture the image or depth data of the hand.

The system can be trained using machine learning algorithms with a dataset of labeled hand gestures.

LITERATURE REVIEW

Hand gestures are an interesting interaction insight in a variety of computer applications. Two principal questions must be answered when using them: The first question is what technology to use for collecting raw data from the hand. Generally, two types of technologies are available for collecting this raw data. The first one is a glove input device and the second way of collecting raw data is to use computer vision. Accuracy of a glove input device depends on the type of sensor technology used; usually, the more accurate the glove is, the more expensive it is. In a vision-based solution, one or more cameras placed in the environment record hand movement. Both types of solutions have their own advantages and disadvantages, and the question that which solution to use is a difficult one. However, when using a hand posture or gesture-based interface, the user does not want to wear the device and be physically attached to the computer. If vision-based solutions can overcome some of their difficulties and disadvantages, they appear to be the best choice for raw data collection.

PROJECT DESCRIPTION AND OUTLINE

1.1 Introduction

The essential aim of building hand gesture recognition system is to create a natural interaction between human and computer where the recognition gestures can be used for controlling a robot or conveying meaningful information. The aim of this technique is the proposal of a real time vision system for its application within visual interaction environments through hand gesture recognition, using general purpose hardware and low-cost sensors, like a simple personal computer and an USB Webcam, so any user could make use of it in his/her office or home.

The development in computer vision makes it possible to approach the interface problems from a human perspective, making the communication between computer and humans more natural . When we as humans communicate, we use our voice and parts of our body such as our face and arms in making gestures. Recently, many attempts have been made to create system that through computer vision could be able to understand gestures . There are two main characteristics that should be deemed when designing a HCI system as mentioned in : Functionality and Usability. System functionality referred to the set of functions or services that the system equips to the users , while system usability referred to the level and scope that the system can operate and perform specific user purposes efficiently . The system that attains a suitable balance between these concepts considered as influential performance and powerful system .

1.2 Motivation for the Work

The motivation for gesture recognition researchers is to develop a system which can detect the gestures and these gestures are widely used for conveying the information or to control the devices. Camera based solutions for gesture recognition has been widely used in numerous applications and capability to communicate through Human Computer Interaction. A new device called Leap Motion Controller is discussed which provides the complete information of hands helps to track hand movements and gestures through API (Application Programming Interface). An idea of real time hand gesture recognition process through this device is explained along with insight of existing machine learning models. Finally, an attempt is made to explain the complexities with the device and the models along with its features.

1.3 Problem Statement

We faced problems in,

- At its first version it was only able to identify few gestures only and failed to recognize a few times.
- To Make the code execute faster i.e., improving the loading time of the Hand recognition system.

1.4 Objective of the Work

The primary objective of this project is to develop a software application that recognize words or sentences using Hand Gesture recognition. The following are the specific objectives of the project:

- Human-computer interaction: The system should allow users to interact with computers and other electronic devices using hand gestures, making the interaction more natural and intuitive.

- Sign language recognition: The system should be able to recognize and interpret sign language gestures, allowing for more effective communication with people who are deaf or hard of hearing.
- Gaming: The system should be able to recognize and respond to hand gestures made by players, allowing for more immersive and interactive gaming experiences.
- Virtual reality: The system should be able to recognize and respond to hand gestures made by users, allowing for more realistic and intuitive interaction in virtual environments.

1.5 Process Flow

- Data acquisition: The system captures images or video of the user's hands using a camera or other input device.
- Preprocessing: The images or video are preprocessed to remove noise and enhance the features of the hand gestures. This may include image cropping, normalization, and background subtraction.
- Feature extraction: The system extracts feature from the preprocessed images or video that are relevant for hand gesture recognition. This may include extracting the shape, size, and movement of the hand.
- Classification: The system uses the extracted features to classify the hand gestures. This may include training a machine learning model on a dataset of labeled hand gestures.
- Output: The system generates an output based on the recognized hand gesture. This may include controlling a cursor on a screen or sending commands to another device.
- Feedback & Evaluation: Depending on the specific application, the system may be able to provide feedback to the user, or be evaluated against a set of ground truth samples, to improve its recognition accuracy over time.

RELATED WORK INVESTIGATION

2.1 Work Investigation on the Topic

Gesture recognition is a type of perceptual computing user interface that allows computers to capture and interpret human gestures as commands. The general definition of gesture recognition is the ability of a computer to understand gestures and execute commands based on those gestures.

Background:

- There has been significant work on hand gesture recognition using deep learning in recent years.
- Researchers have used various architectures such as convolutional neural networks (CNNs), long short-term memory (LSTM) networks, and combination of CNNs and LSTM networks for this task.
- Many of the recent works have used depth data from depth sensors like Kinect and Leap Motion to train these models.
- Additionally, some works have used data from RGB cameras and other modalities like infrared and thermal cameras to improve the robustness of the models.

Existing Approaches:

The primary goal in studying gesture recognition is to introduce a system that can detect specific human gestures and use them to convey information or for command and control purposes. Therefore, it includes not only tracking of human movement, but also the interpretation of that movement as significant commands. Two approaches are generally used to interpret gestures for HCI applications. The first approach is based on data gloves (wearable or direct contact) and the second approach is based on computer vision without the need to wear any sensors.

Hand Gestures Based on Instrumented Glove Approach

The wearable glove-based sensors can be used to capture hand motion and position. In addition, they can easily provide the exact coordinates of palm and finger locations, orientation, and configurations by using sensors attached to the gloves. However, this approach requires the user to be connected to the computer physically, which blocks the ease of interaction between user and computer. In addition, the price of these devices is quite high. However, the modern glove based approach uses the technology of touch, which more promising technology and it is considered Industrial-grade haptic technology.

Where the glove gives haptic feedback that makes user sense the shape, texture, movement and weight of a virtual object by using microfluidic technology.

Hand Gestures Based on Computer Vision Approach

The camera vision-based sensor is a common, suitable and applicable technique because it provides contactless communication between humans and computers. Different configurations of cameras can be utilized, such as monocular, fisheye, TOF and IR. However, this technique involves several challenges, including lighting variation, background issues, the effect of occlusions, complex background, processing time traded against resolution and frame rate and foreground or background objects presenting the same skin color tone or otherwise appearing as hands.

REQUIREMENT ARTIFACTS

3.1 Hardware and Software requirements

Hardware requirements:

- Modern Operating System.
- x86 64-bit CPU (Intel / AMD architecture)
- 4 GB RAM.
- 5 GB free disk space.

Software requirements:

We can either use **VS Code** or **PyCharm** to compile and run the code and for this project we'll be using VS Code.

VS Code: Visual Studio Code is a streamlined code editor that supports development operations such as debugging, task execution, and version control. It aims to give developers only the tools they need to speed up their code, build and debug cycle, leaving more complex workflows to her IDE with rich features such as Visual Studio IDE.

VS Code also shares same features as PyCharm.

3.2 Modules and Library requirements

3.2.1 TensorFlow

TensorFlow is an end-to-end open-source platform for machine learning. TensorFlow is a rich system for managing all aspects of a machine learning system; however, this class focuses on using a particular TensorFlow API to develop and train machine learning models.

3.2.2 NumPy

NumPy is a numerical computing library in Python that provides fast and efficient operations for arrays and matrices. It is widely used in scientific computing and data analysis. NumPy provides tools for linear algebra, random number generation, and statistical operations. It also integrates well with other libraries in the scientific computing stack, such as SciPy and Matplotlib.

3.2.3 OpenCV

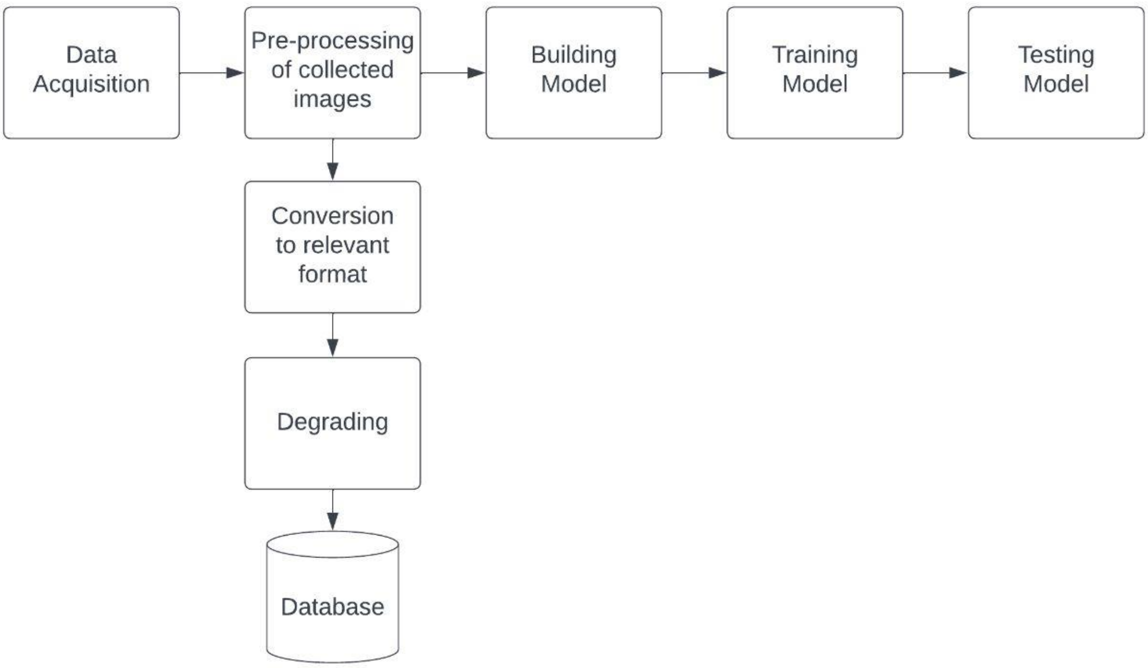
OpenCV (Open-Source Computer Vision Library) is an open-source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products. Being an Apache 2 licensed product, OpenCV makes it easy for businesses to utilize and modify the code.

3.2.4 Matplotlib.pyplot

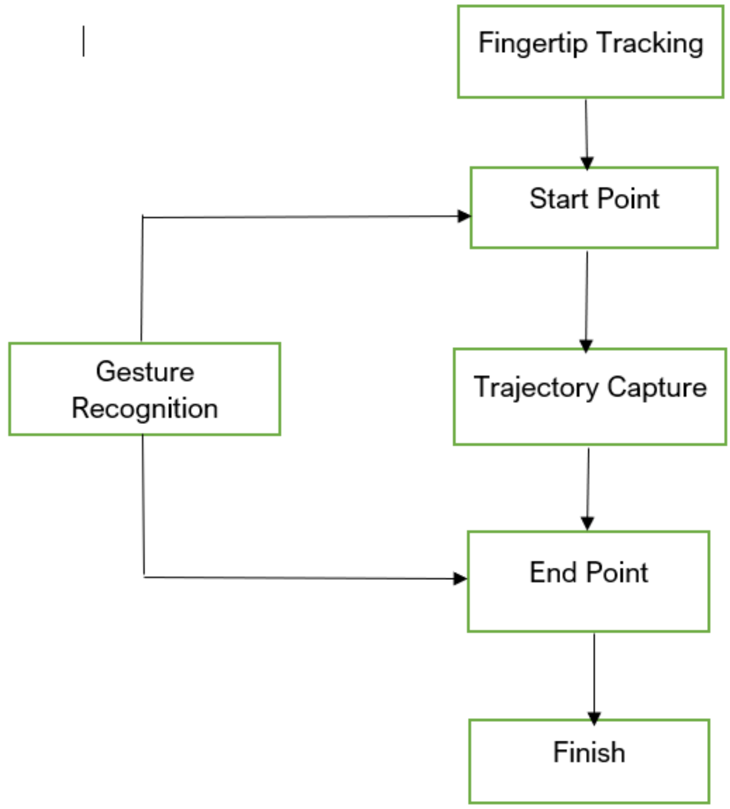
Matplotlib is a plotting library for Python that allows users to create a wide range of static, animated, and interactive visualizations. Matplotlib provides functions for creating 2D and 3D plots, histograms, bar charts, scatter plots, and more. It integrates well with other libraries in the scientific computing stack, such as NumPy and Pandas, and is a popular tool for data visualization in Python.

DESIGN METHODOLOGY AND ITS NOVELTY

4.1 Work flow



4.2 Flowchart

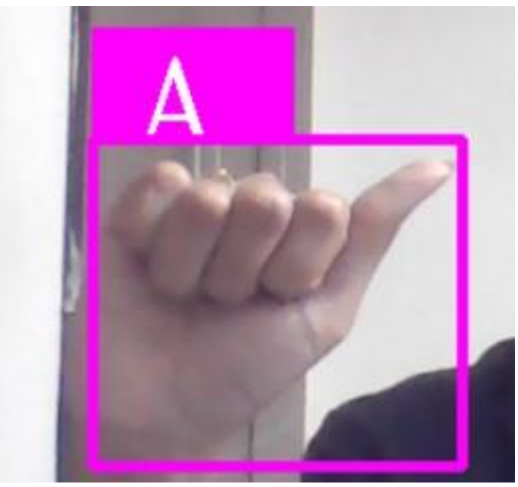

Fingertip Tracking
Start Point
Gesture Recognition
Trajectory Capture
End Point
Finish

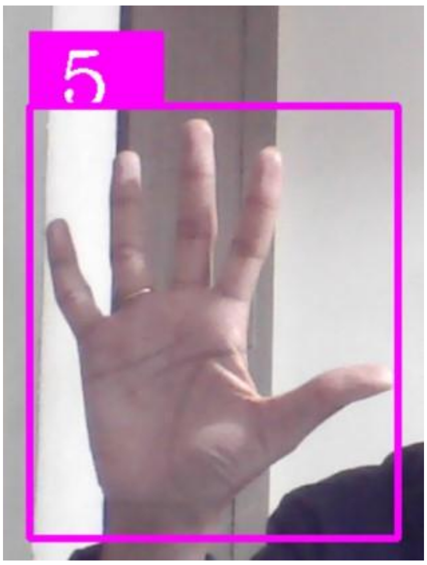
TECHNICAL IMPLEMENTATION & ANALYSIS

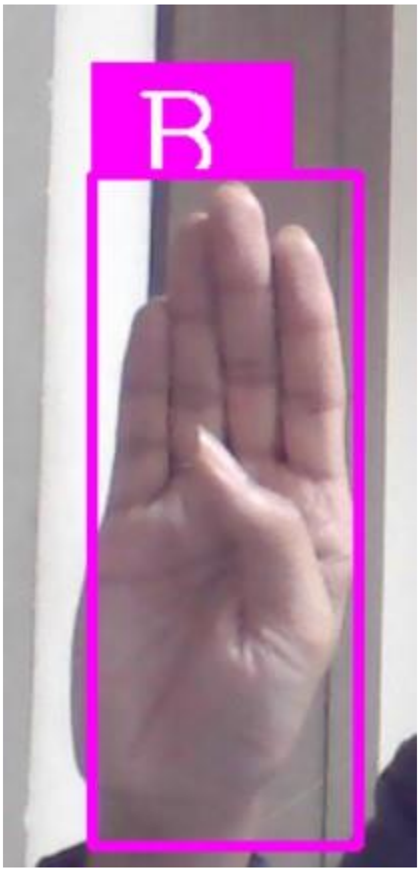
5.1 Explanation of Code

```python
import cv2
from cvzone.HandTrackingModule import HandDetector
from cvzone.ClassificationModule import Classifier
import numpy as np
import math
import time

cap = cv2.VideoCapture(0)
detector = HandDetector(maxHands=2)
classifier = Classifier("Model/keras_model.h5", "Model/labels.txt")

offset = 20
imgSize = 300

folder = "Data/5"
counter = 0

labels = ["A", "B", "C", "D", "E", "5"]

while True:
    success, img = cap.read()
    imgOutput = img.copy()
    hands, img = detector.findHands(img)
    if hands:
        hand = hands[0]
        x, y, w, h = hand['bbox']

        imgWhite = np.ones((imgSize, imgSize, 3), np.uint8)*255
        imgCrop = img[y-offset:y + h + offset, x - offset:x + w + offset]
```

```
    if aspectRatio > 1:
        k = imgSize/h
        wCal = math.ceil(k*w)
        imgResize = cv2.resize(imgCrop, (wCal, imgSize))
        imgResizeShape = imgResize.shape
        wGap = math.ceil((imgSize-wCal)/2)
        imgWhite[:, wGap:wCal + wGap] = imgResize
        prediction, index = classifier.getPrediction(imgWhite,draw=False)
        print(prediction, index)


    else:
        k = imgSize / w
        hCal = math.ceil(k * h)
        imgResize = cv2.resize(imgCrop, (imgSize, hCal))
        imgResizeShape = imgResize.shape
        hGap = math.ceil((imgSize - hCal) / 2)
        imgWhite[hGap:hCal + hGap, :] = imgResize
        prediction, index = classifier.getPrediction(imgWhite, draw=False)

    cv2.rectangle(imgOutput, (x - offset, y - offset-50), (x - offset + 90, y - offset-50+50), (255, 0, 255), cv2.FILLED)
    cv2.putText(imgOutput, labels[index], (x,y-20), cv2.FONT_HERSHEY_COMPLEX, 1.7, (255,255, 255),2)
    cv2.rectangle(imgOutput, (x-offset, y-offset), (x+w+offset, y+h+offset), (255, 0, 255), 4)

    cv2.imshow("ImageCrop", imgCrop)
    cv2.imshow("ImageWhite", imgWhite)

cv2.imshow("Image", imgOutput)
cv2.waitKey(1)
```






B


E

PROJECT OUTCOME AND APPLICABILITY

6.1 Project applicability on real world applications

The Hand Gesture Detection and Recognition system has several real-world applications and can be used in a variety of industries and fields. Some of the most notable applications include:

- **Medical Operations:**
  Gestures can be used to control the distribution of resources in hospitals, interact with medical instrumentation, control visualization displays, and help handicapped users as part of their rehabilitation therapy. Some of these concepts have been exploited to improve medical procedures and systems; for example, a technology which satisfied the “come as you are” requirement, where surgeons control the motion of a laparoscope by making appropriate facial gestures without hand or foot switches or voice input. Simply hand gestures into doctor-computer interfaces, describing a computer-vision system that enables surgeons to perform standard mouse functions, including pointer movement and button presses, with hand gestures that satisfy the “intuitiveness” requirement.
- **Gesture Based Gaming Control:**
  Computer games are a particularly technologically promising and commercially rewarding arena for innovative interfaces due to the entertaining nature of the interaction. Users are eager to try new interface paradigms since they are likely immersed in a challenging game-like environment. In a multi-touch device, control is delivered through the user’s fingertips. Which finger touches the screen is irrelevant; most important is

where the touch is made and the number of fingers used. In computer-vision-based, hand-gesture-controlled games, the system must respond quickly to user gestures, the “fast-response” requirement. In games, computer-vision algorithms must be robust and efficient, as opposed to applications (such as inspection systems) with no real-time requirement, and where recognition performance is the highest priority. Research efforts should thus focus on tracking and gesture/posture recognition with high-frame-rate image processing.

CONCLUSIONS AND RECOMMENDATION

- There are deaf who uses hand sign to convey their message but normal people some time might know sign language.
- Individuals who are using hand signs to communicate covertly or to trigger certain actions.
- Sometime it is difficult for audience and players to interpret hand signals used by coaches and players to communicate during games.
- In conclusion, hand recognition systems have become increasingly important in recent years due to their wide range of applications in various fields such as robotics, gaming, and security systems. These systems use various techniques such as computer vision, machine learning, and deep learning algorithms to detect, track, and recognize hand gestures and movements.
- Hand recognition systems have been shown to be highly accurate and efficient in identifying hand gestures, which has made them a popular choice for applications such as sign language recognition, human-computer interaction, and virtual reality systems. These systems have also been used in security systems to detect unauthorized access, and in medical fields to assist with diagnosis and treatment.

- However, hand recognition systems still face some challenges, including variations in lighting conditions, occlusion, and noise, which can affect their accuracy. Despite these challenges, ongoing research in this field is continually improving the performance of hand recognition systems, making them more reliable and robust.
- Overall, hand recognition systems have great potential for a wide range of applications, and as technology continues to evolve, we can expect even more advanced systems to be developed in the future.

REFERENCES

- **Paper 1 , Abdullah Mujahid , Real-Time Hand Gesture Recognition Based on Deep Learning YOLO v3 Model , (MDPI , DOI:10.3390/app11094164, Year of publication 2021)**
- **Paper 2 , Ali Moin , A wearable bio-sensing system with in-sensor adaptive machine learning for hand gesture , (MDPI , DOI-**

https://doi.org/10.1038/s41928-020-00510-8 , Year of publication 2021)

- Paper 3 , Ankita Wadhawan , Deep learning-based sign language recognition system for static sign , (DOI- https://doi.org/10.1007/s00521-019-04691-y , Year of publication 2021)

- Paper 4, Munir Oudah , Hand Gesture Recognition Based on Computer Vision , (MDPI , 10.3390/jimaging6080073 , Year of publication 2020)
- Paper 5 , Muneer Al-Hammadi , Deep Learning-Based Approach for Sign Language Gesture Recognition With Efficient Hand Gesture Representation , (IEEE, 10.1109/ACCESS.2020.3032140 , Year of publication 2020)

- **Paper 6 , Bin Hu , Deep Learning Based Hand Gesture Recognition and UAV Flight Controls , (Researchgate , DOI-doi.org/10.1007/s11633-019-1194-7, Year of publication 2020**
- **Paper 7 , Ali Raza Asif , Performance Evaluation of Convolutional Neural Network for Hand Gesture Recognition Using EMG ,(MDPI , DOI - doi.org/10.3390/s20061642 , Year of publication 2020)**