SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

Kattankulathur, Chengalpattu District - 603203

18CSC305J – ARTIFICIAL INTELLIGENCE

MINI PROJECT REPORT

VOLUME CONTROL USING HAND DETECTION

Gudied by:

RENUKA DEVI .P

Submitted By:

KEERTHIKA KANCHIPATI	RA2011003010766
MANEESHA BHUPALEM	RA2011003010763
THOTA NAGA NANDINI	RA2011003010787

COLLEGE OF ENGINEERING & TECHNOLOGY
SRM INSTITUTE OF SCIENCE & TECHNOLOGY
S.R.M. NAGAR, KATTANKULATHUR – 603 203

BONAFIDE CERTIFICATE

Certified that this project report **" Volume Control using Hand Detection"** is the bonafide work of **"Maneesha Bhupalem,Keerthika Kanchipati,Thota Naga Nandini"** of III Year/VI Sem B.tech(CSE) who carried out the mini project work under my supervision for the course 18CSC303J-Database Management systems in SRM Institute of Science and Technology during the academic year 2022-2023(Even sem).

SIGNATURE

Faculty name
Faculty Designation
Department name and seal

ABSTRACT

We are developing a volume controller in which we are using hand gestures as the input to control the system, Opencv module is basically used in this implementation to control the gestures. This system basically uses the web camera to record or capture the images /videos and accordingly on the basis of the input, the volume of the system is controlled by this application. The main function is to increase and decrease the volume of the system.
We can use our hand gestures to control the basic operation of a computer like increasing and decreasing volume. Therefore ,people will not have to learn machine-like skills which are a burden most of the time.This type of hand gesture systems provides a natural and innovative modern way of non verbal communication. These systems has a wide area of application in human computer interaction. *A systems that can recognize a hand motion in real time video is hand gesture recognition. Hand gestures are categorized according to their subject matter. The design of hand gesture recognition is one of the more difficult jobs, as it combines two significant issues. The detection of the hand is the first step and creating a sign that can only be utilized by one hand at a time. It can be used in a variety of settings, including human-computer interaction and sign language. The basic concept of hand segmentation and the hand detection system, which use the Haar-cascade classifier, may be used to construct hand gesture recognition using The configuration comprises a single camera that captures the user's gesture and feeds it into the system. A fundamental goal of gesture recognition is to develop a system that can recognize specific human gestures and utilize them to send information for device control. With real-time gesture recognition, a user can operate a computer by making a specific gesture in front of a computer's camera.* Computer interaction is one of the significant challenges that elderly and disabled people face in today's computerized world. In recent years, hand gesture recognition has emerged as one of the most natural human-machine interactions in software development, particularly for facilitating friendly and flexible human-computer interaction. This article proposed a system architecture and software configuration for developing a real-time volume control hand gesture on a Raspberry Pi equipped with a camera programmed in Python using the Open-Source Computer Vision (OpenCV) library. The hand gesture recognition system's primary objective is to establish communication between humans and computerized systems in order to control volume. According to the experimental

results, the hand gesture recognition system performs well when controlling volume. The system is capable of operating in real-time for any individual.

INTRODUCTION

Hand gestures is the powerful communication medium for Human Computer Interaction (HCI).Several input devices are available for interaction with computer, such as keyboard, mouse, joystick and touch screen, but these devices does not provide easier way to communicate. In this, the system which is proposed will consists of desktop and laptop interface, hand gesture can be used by the users need to wear data gloves also can use the web camera or separate cameras for recording the hand gestures. The first and most important step toward any hand gesture recognition system is to implement hand tracking system. Some Sensor devices are generally used in DataGlove based methods for digitizing hand and finger motions into multi parametric data. Other sensors used in this system will collect hand configuration and hand movements.

Hand gesture recognition is a system that can recognize a hand motion in a real-time video in everyday life. Hand gestures are classified according to their subject of interest. One of the goal in this implementation is to design a system for hand gesture recognition. One of the most critical issues in computer vision is the task of recognizing hand motions. Human computer interaction (HCI) systems that entail hand processing tasks such as hand detection and hand gesture recognition have become more advanced as information and media technology has progressed. Detecting and locating the hand in real-time footage from the webcam is the initial stage in any hand processing system. Because of the variety in position, orientation, placement and scale, detecting a hand might be difficult. Variability is also aided by the varying levels of light in the room. Hand gesture recognition often requires numerous levels of processing, including image acquisition, pre-processing, feature extraction, and gesture identification. Using a webcam, image acquisition entails recording an image in a video frame by frame. The collected images are subjected to colour filtering, smoothing as part of the image pre-processing. Feature extraction is a technique for extracting features from a hand image, such as hand outlines, whereas gesture recognition is a technique for extracting features and recognizing hand gestures. Designing a hand gesture recognition system is a difficult task that comprises two major issues. The first is the detection of a person's hand. A webcam is used to detect the user's hand in real-time video in this acquisition some issue arises such as inconsistencies in brightness, noise,resolution,

and contrast. To identify the gestures, the technique involves segmentation and edge detection and with help of the openCV module we obtain the hand gesture and able to control the volume.

LITERATURE SURVEY:

Hand gesture detection and utilizing them to control certain set of devices operations and allowing interaction with computer system without the aid of mouse and keyboard. In this paper we draws along the same line but we attributed the use of Haar-cascade classifier to identify hand gesture. Some of the related works in this fifield are described brieflfly as follows [1]A non-local algorithm for hand gestures was proposed by A. Buades, B. Coll, and J. Morel. At the moment, finding finger movement algorithms remains a valid task. Functional analysis and statistics collide. Despite the fact that most recently presented approaches have a high level of sophistication, Algorithms have not yet reached a satisfactory degree of performance applicability. All work admirably when the model matches the algorithm assumptions, but they all fail in general, producing defects in analysing the pixels through the camera. The primary goal of this study is to define a generic mathematical and experimental technique for comparing and classifying conventional hand movement recognition algorithms. [2] For the no required elements in the video frame, Golam Moktader Daiyan et al. (2012) suggested a high performance decision based median filter. This technique detects noise pixels iteratively over numerous phases before replacing them with the median value. Noise detection is accomplished by enlarging the field of view. Mask till 77% to keep the extraction of local data going. Furthermore, if the algorithm fails to find a noise-free pixel at 7 7, the processing pixel is replaced by the last processed pixel. If the noise-free median value isn't available in the 7th processing window, the last processed pixel is used to determine if it is noise-free. The method chooses a window size if the last processed pixel is noisy. Calculate the number of 0s and 255s in the processing window using the 1515 dimension. Then, in the selected window, replace the processed pixel with 0 or 255, whichever is higher in number. [3]Rajeshwari Kumar Dewangan et.al accurate object information and obtain a location using a deep learning object recognition technique.Object recognition algorithms are designed based on theSingle Shot MultiBox Detector (SSD) structure, an object recognition deep learning model, to detect objects using a camera. [4] H. Jabnoun et, al suggested the system that restores a central function of the visual system which is the identification of surrounding objects which is

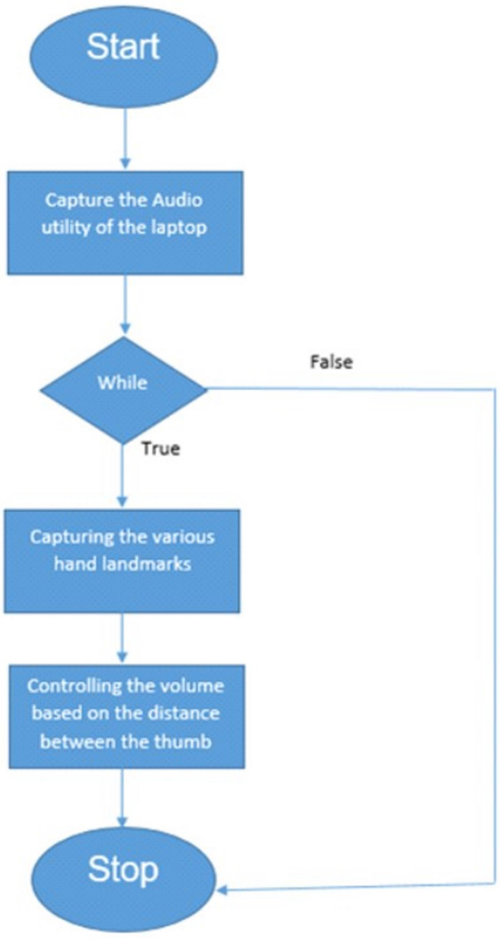
based on the local features extraction concept. Using SFIT algorithm and keypoints matching showed good accuracy for detecting objects. [5] Košale U, Žnidaršic P, Stopar K suggested that Detection of obstacles is performed by Time of Flight (ToF) sensors, whose ranging data is then processed with an on-board microcontroller and send via Bluetooth connection to the belt. The belt is equipped with a
second microcontroller, which interprets the data and presents it to the wearer with 15 vibration motors arranged in a square grid. The glasses are worn on the head, whereas the belt is attached around the stomach area. But the number of sensors detecting the obstacle decreased with the distance. Circle and square were detected better than triangle. This suggests that different shapes trigger different responses of sensors on glasses. A. Jaiswal et al. [6] proposed an approach that used user generated picture denoising. The remaining task is broken down into four steps. The first image is denoised using a filtering process, and the second image is denoised using a different method. Wavelet-based approaches are used to denoise pictures, filtering, third hard thresholding, and thresholding. Finally, the approach was applied to a noisy image concurrently. PSNR output results are calculated by comparing all cases, the MSE (mean square error) is obtained. On the basis of PSNR, MSE, and image visual quality, experiments are conducted on 512 X 512 noisy images with noise variance of 0.04, output of median filter, Wiener filter, hard thresholding, and hard thresholding with median filtering. When the filtering and wavelet thresholding techniques are combined, they produce excellent results.

PROPOSED SYSTEM:

As shown in figure 3.1, which indicates the proposed architecture of the system used in the volume controller. Here our input is hand gesture which is captured using web camera. Then the GUI (graphical user interface) helps to display the

hand that conveys information and it processes the actions of the hand gestures of the user. By recognizing the gestures, the user is able to control the volume of the system which is final output.

METHODOLOGY:

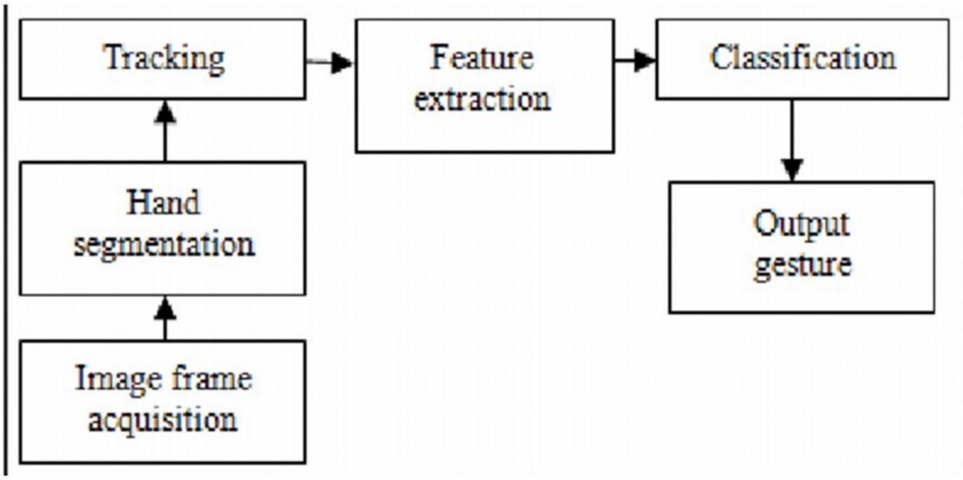


The fig 4.1 indicates the flow chart of operation. Firstly, initiate start the program and then import various modules used for the AI recognition and various audio utilities which are of main concern. Next, it capture the area of interest by detecting various contours. Later it execute the loop to detect various hand landmarks. After getting the hand land marks, it verify the distance between the thumb and index figure tip. The frame is displayed giving the final values of the reading with complete decrease and increase in the volume using AI. The program is executed till the loop is iterated. Once it completes the iterations it comes out of the loop and the program
stops.

ALGORITHM:

Step 1: Start the Program
Step 2: Importing the Various Modules-Open Vision used for the AI recognition and various audio utilities which are of main concern.
Step 3: Capturing the Area of Interest by detecting the various contours and differentiating the white and balck region of the interest
Step 4: Executing the Loop to detect the various hand landmarks
Step 5: Getting the Hand Landmarks and verifying the distance between the index and thumb finger based on the algorithm given
Step 6: Display the frame giving the final values of the reading with complete decrease and increase in the volume using Artificial
Intelligence. The Program is executed till the loop is iterated once it completes the iterations it comes out of the loop and program stops
Step 7: Stop

ARCHITECTURE DIAGRAM:



CODING AND WORKING:

Installed libraries-

```
import cv2

import mediapipe as mp

from math import hypot

from ctypes import cast, POINTER

from comtypes import CLSCTX_ALL

from pycaw.pycaw import AudioUtilities, IAudioEndpointVolume

import numpy as np
```

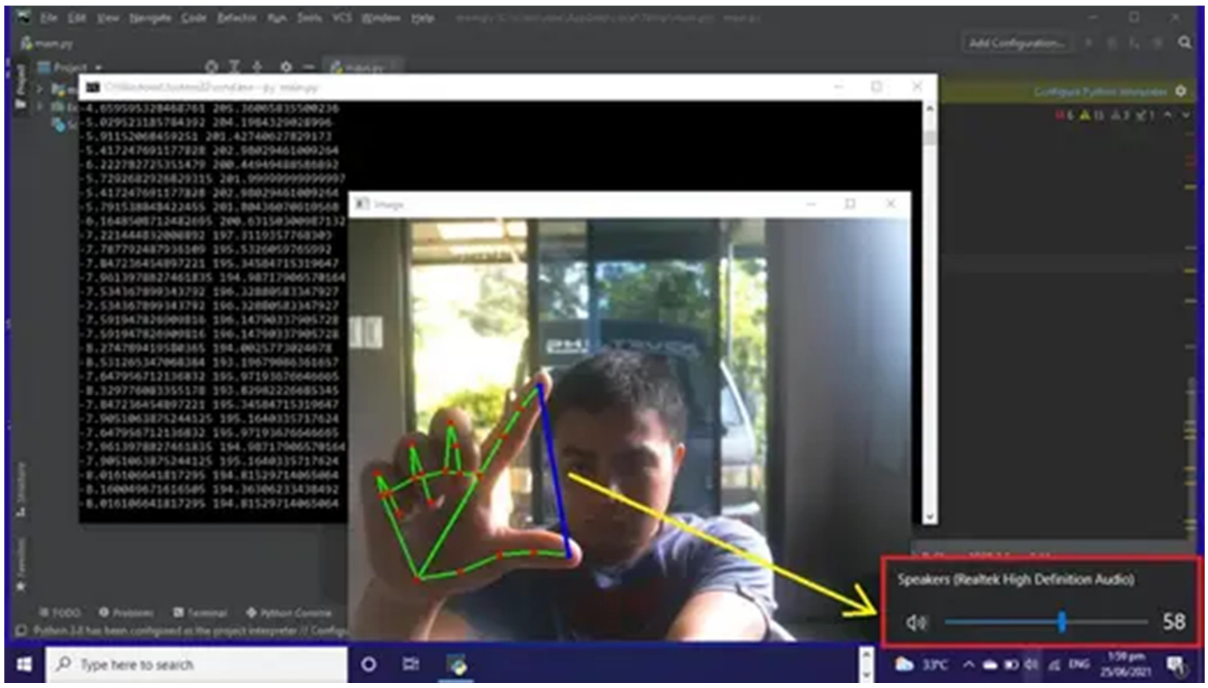
SOUCE CODE-

```
cap = cv2.VideoCapture(0)

mpHands = mp.solutions.hands
```

```
hands = mpHands.Hands()

mpDraw = mp.solutions.drawing_utils

devices = AudioUtilities.GetSpeakers()interface =
devices.Activate(IAudioEndpointVolume._iid_, CLSCTX_ALL, None)

volume = cast(interface, POINTER(IAudioEndpointVolume))


volMin,volMax = volume.GetVolumeRange()[:2]

while True:

    success,img = cap.read()

    imgRGB = cv2.cvtColor(img,cv2.COLOR_BGR2RGB)

    results = hands.process(imgRGB)


    lmList = []

    if results.multi_hand_landmarks:

        for handlandmark in results.multi_hand_landmarks:

            for id,lm in enumerate(handlandmark.landmark):

                h,w,_ = img.shape

                cx,cy = int(lm.x*w),int(lm.y*h)

                    lmList.append([id,cx,cy])


mpDraw.draw_landmarks(img,handlandmark,mpHands.HAND_CONNEC
TIONS)
```

```
if lmList != []:
    x1,y1 = lmList[4][1],lmList[4][2]
    x2,y2 = lmList[8][1],lmList[8][2]

    cv2.circle(img,(x1,y1),4,(255,0,0),cv2.FILLED)
      cv2.circle(img,(x2,y2),4,(255,0,0),cv2.FILLED)
    cv2.line(img,(x1,y1),(x2,y2),(255,0,0),3)

    length = hypot(x2-x1,y2-y1)
    vol = np.interp(length,[15,220],[volMin,volMax])
    print(vol,length)
    volume.SetMasterVolumeLevel(vol, None)

    # Hand range 15 - 220
    # Volume range -63.5 - 0.0

cv2.imshow('Image',img)
if cv2.waitKey(1) & 0xff==ord('q'):
    break
```

OUTPUT:



CONCLUSION:

In conclusion, volume control with hand detection using OpenCV is a novel and innovative way to control the volume of devices. It provides a handsfree and intuitive method of adjusting the volume, which can be particularly useful in situations where manual controls are not accessible or practical. The module uses computer vision techniques to detect and recognize hand gestures, and machine learning algorithms to classify them and map them to corresponding volume levels. The system is highly customizable and can be adjusted to recognize different hand gestures and map them to different volume levels. Overall, volume control with hand detection using OpenCV is a promising technology that has the potential to revolutionize the way we interact
with devices and technology.

REFERENCES

[1] Coll, Bartomeu, Antoni Buades & Morel, Jean-Michel. (2005). A Review of Image Denoising Algorithms, with a New One. SIAM Journal on Multiscale Modeling and Simulation. 4. 10.1137/040616024.

[2] Daiyan, Golam & Mottalib, M.A.. (2012). High performance decision based median filter for salt and pepper noise removal in images. 107-112. 10.1109/ICCITechn.2012.6509707.

[3]Rajeshwari Kumar Dewangan, Siddharth Chaubey, "Object Detection System with Voice Output using Python," *2021 International Journal of Science & Engineering Development Research,* Vol. 6, Issue 3, pp. 15-20, 2021.

[4] H. Jabnoun, F. Benzarti, H. Amiri, "Object detection and identification for blind people in video scene," *2015 15th International Conference on Intelligent Systems De- sign and Applications (ISDA)*, pp. 363-367, 2015.

[5] Košale U, Žnidaršic P, Stopar K, "Detection of different shapes and materials by glasses for blind and visually impaired," *2019 6th Student Computer Science Research Conference*, Vol. 57, 2019.

[6]Ayushi Jaiswal Ajay Somkuwar Jayprakash Upadhyay '*Image denoising and quality measurements by using filtering and wavelet based techniques'* AEU - International Journal of Electronics and Communications 68(8):699-705

[7]E. Barri, A. Gkamas, E. Michos , C. Bouras, C. Koulouri, S.A.K. Salgado, "Text to Speech through Bluetooth for People with Special Needs Navigation," *International Conference on Ne8tworking and Services,* 2010.