Gestures To Control Computer Applications

Mohd Talha Zeeshan[1], Mohd Ishthyaq Hussain [2], B Karthik Raj[3]

[1]Student, Dept. of Electronics and Communication Engineering, SNIST, Ghatkesar, Hyderabad, Telangana, India
[2]Student, Dept. of Electronics and Communication Engineering, SNIST, Ghatkesar, Hyderabad, Telangana, India
[32]Student, Dept. of Electronics and Communication Engineering, SNIST, Ghatkesar, Hyderabad, Telangana, India

---***---

Abstract - Gesture to control computer applications means we use gestures to control our computer system here in this project we have used hand gesture to control the two main application of computer that is controlling a VLC media player and presentation and we have also tried to use our eyes as cursor. This type of application is very much helpful during the Covid era as we know that many people got affected to covid by simply touch other objects so there instead of using joystick, keyboard and mouse without any direct contact we can interact with the computer device. As every application it also has pros and cons. It also helpful for many people who are suffering with physical disabilities. As we know that the world is very large many people loose their fingers and hands in accidents so such people can also access the computer easily. By using hand gestures we a can play/pause a video, increase and decrease the volume and can forward and rewind a video. In a presentation we can using our finger as a pointer, we can draw or write on a presentation slides, we can erase what ever we have written or drawn on the slide and we can move the slide to forward and backward. Here we are using only python and its libraries and pycharm IDE we can also do the same project by using arduino and also other similar devices.

***Key Words*:** **Gestures, Python, VLC Media Player, Libraries, Pycharm IDE.**

1.INTRODUCTION

A computing method called gesture recognition makes an effort to identify gestures and analyse them using mathematical algorithms. Gesture recognition is not just restricted to hand motions; it can also be applied to every part of the body. There is a global conference devoted to gesture and facial recognition, which is a burgeoning area of computer science. The possibilities for using the field will expand as well as the field itself. Gesture recognition software can be used with touch screens, cameras, or other peripheral devices, among other methods, to improve human-computer interaction.

For many individuals today, recognising touch screen gestures has become second nature. Although certain computers and operating systems provide individualised gesture recognition, Nowadays, the majority of people are aware of the pinch-to-zoom feature on touch screens, which they may use to get a closer look at anything. Nearly all user interfaces, including those on smartphones and desktop computers, are compatible with this particular gesture. Touchscreens make it relatively simple for people to interact with computers.

A camera and motion sensor are used in vision-based gesture recognition technology to monitor and translate human motions in real time. The monitoring of depth data is also possible with more recent cameras and software, which helps enhance gesture recognition. Users may quickly engage with the application to get the required results thanks to real-time picture processing. As an illustration, the Xbox Kinect used a camera to interpret player gestures throughout various games.

Deep learning algorithms have also been tested in tests where a camera was used to watch a person's walk in order to estimate their risk of falling and provide advice on how to reduce that risk.

Devices that employ specialised cameras and programmes focused on hand tracking have been developed, such as those from Leap Motion, to enhance the results of motion-tracking. Such applications can improve accuracy by concentrating solely on hand gesture detection, enabling users to interface with their systems conveniently and totally hands-free.

2. LITERATURE SURVEY

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.

3. PROPOSED METHODOLOGY

This project made use of a number of computer vision-related methods. These consist of the ones utilised in colour segmentation, Morphological filtering, Extraction of Features, Contours, Convex Hull and controlling media player and presentation using pyautogui.

3.1.1 Data Flow Digaram

The data flow diagram essentially shows how the data control flows from one module to another. Unless the input filenames are correctly given the program cannot proceed to the next module. Once the user gives, the correct input filenames parsing is done individually for each file.

The data flow diagram essentially shows how the data control flows from one module to another. Unless the input filenames are correctly given the program cannot proceed to the next module. Once the user gives, the correct input filenames parsing is done individually for each file. The required information is taken in parsing and an adjacency matrix is generated for that. From the adjacency matrix, a lookup table is generated giving paths for blocks.

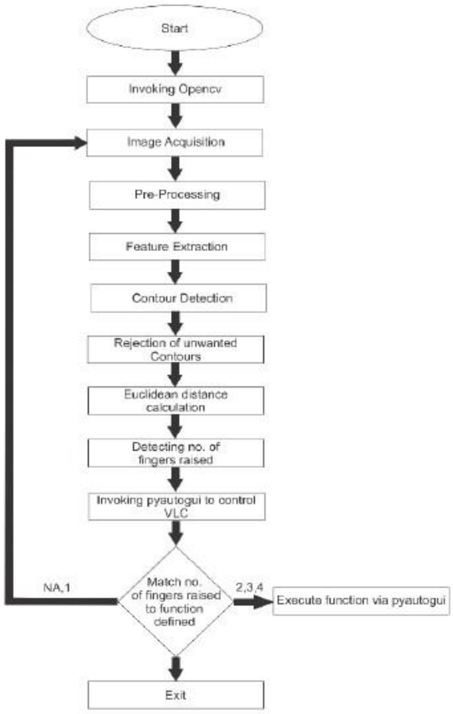

Fig-1: Flow Chart

In addition, the final sequence is computed with the lookup table and the final required code is generated in an output file. In case of multiple file inputs, the code for each is generated and combined together.

3.1.2 Component Diagram

The component diagram for the gesture detection System include the various unit for input and output operation. For our designed we have mainly two process in which one is to capture the image through camera which is done by invoking openCV and other is pre-processing done by the system. The processing includes two Unit which is used to process the image capture by the camera, Firstly the preprocessing unit detects the metadata ig the image and its trajectory that is the orientation in which the hand fingers was raised ,then its sent for further pre-processing.

In Further pre-processing system used the algorithm for extracting the feature to recognize the fingers raised The feature are extracted using the metadata and information in previous pre-processing steps.It is interesting to note that all the sequence of activities that are taking place are via this module itself, i.e. the parsing and the process of computing the final sequence. The parsing redirects across the other modules until the final code is generated.

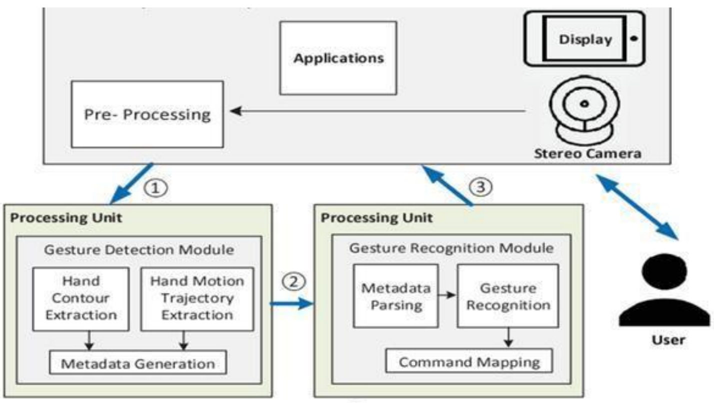

Fig-2: Component Diagram

3.1.3 Activity Schematic

An activity diagram displays the order in which a complicated process's many phases must occur. An activity is represented by a circle holding the operation's name. A transition caused by the completion is denoted by an outgoing solid arrow connected to the end of the activity sign.

Activity diagrams are visual depictions of processes with choice, iteration, and concurrency supported by activities and actions. Activity diagrams in the Unified Modelling Language are meant to represent both organisational and computational operations (i.e. workflows). Activity diagrams display the entire control flow. arrow-connected activity diagrams are up of a small number of shapes. The most significant shapes include:

Round rectangles stand in for actions, diamonds for choices, bars for the beginning or end of concurrent activities, a black circle for the workflow's initial state, and an encircling black circle for its conclusion (final state).

Similar to the other four diagrams, activity diagrams serve similar fundamental goals. It captures the system's dynamic behaviour. The message flow from one item to another is depicted using the other four diagrams, whereas the message flow from one activity to another is depicted using the activity diagram.

An activity is a specific system function. Activity diagrams are used to build the executable system utilising forward and reverse engineering methodologies in addition to helping to see a system's dynamic nature.

Recognition of hand Gesture includes various activities to be performed. As shown in the figure 5.5 the first activity is to start the camera to capture the image. This activity automatically invoke the camera using openCV library in python as the execution of the program starts. Then on the basis of the capture image , the gesture information is extracted. This information is used to extract the features such contours , convex hull and the defects point. On the basis of which the number of fingers in front of the camera is Recognized .As the Finger information is extracted ,the application automatically perform the action such as play, pause, seek forward, seek backward etc on the basis of number of fingers raise, on the VLC Media Player .

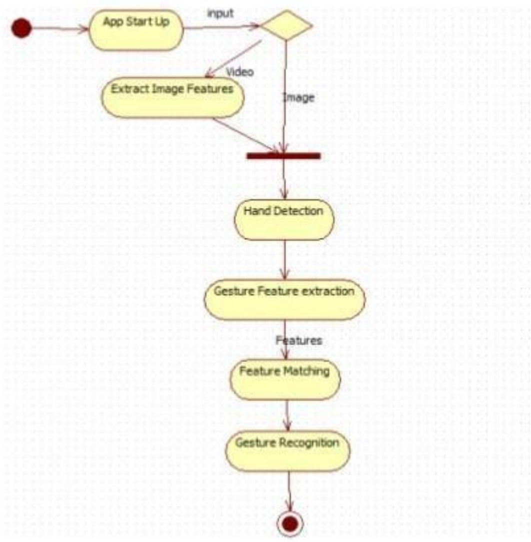

Fig-3: Activity digarm.

3.1.4 Sequence Diagram

Sequence diagram are an easy and intuitive way of describing the behaviour of a system by viewing the interaction between the system and the environment. A sequence diagram shows an interaction arranged in a time sequence. A sequence diagram has two dimensions: vertical dimension represents time; the horizontal dimension represents the objects existence during the interaction.

The shown Sequence diagram explain the flow of the program. As this System is based on Human Computer Interaction so it basically include the user, computer and the medium to connect both digitally that is web camera. As the execution of the program starts, it firstly invoke the web camera to take RGB image of the hand. Then the image is segmented and filtered to reduce the noise in the image .After the removal of the noise the hand gestures are detected i.e the number of fingers raised are preprocessed, on the basis of feature are extracted. After Feature extraction is done, By using conditionals statements in the program the feature are matched .As the features matches ,it automatically controls the media player and gives us the required results.

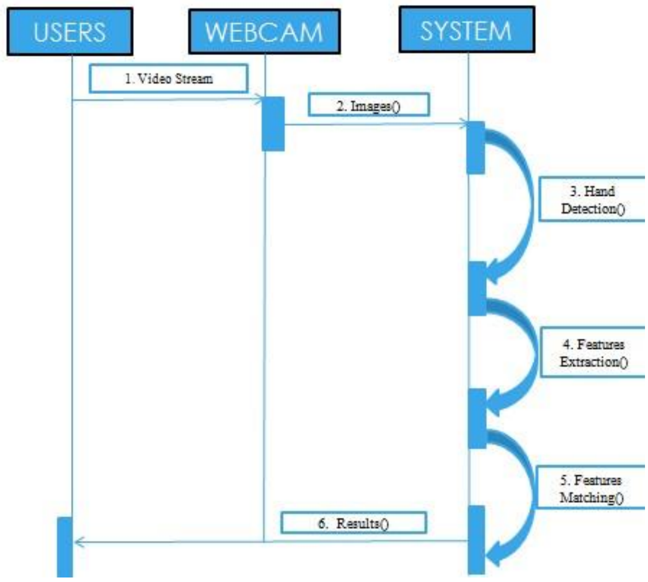

Fig-4 : sequence diagram.

3.2 Obtaining Data

The first step is to collect the picture from the camera and designate a region of interest in the frame. This is vital because the image may include many variables, and these factors may have unintended effects, which greatly reduces the amount of data that has to be processed. A webcam is utilised to take the picture, which continually records frames and provides the raw data needed for processing. The input image in this case is an uint8. The RGB picture that was procured needs to be treated first, or pre-processed, before the components are divided and an acknowledgement is given.

3.2.1 Threshold-Based Colour Segmentation

Identification of certain areas within a picture is known as segmentation.

The following diagram illustrates the algorithm used for thresholding-based colour segmentation:

• Use the camera to record a picture of the gesture.

• Establish the range of HSV values for skin tone that will be used as thresholds.

• Change the image's colour space from RGB to HSV.

• Change all of the pixels to white if they fall inside the threshold values.

• Change all remaining pixels to black.

The segmented image should be saved as an image file.

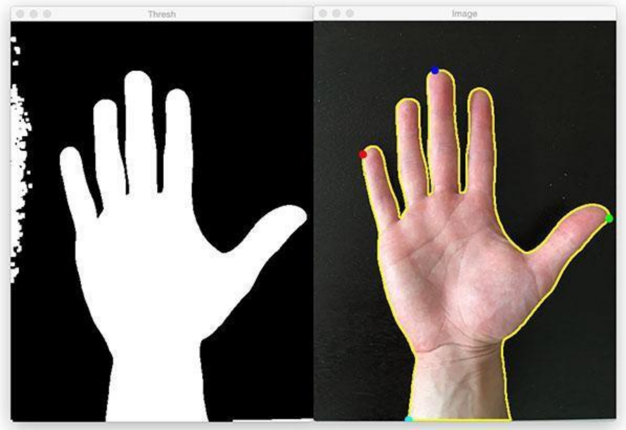

Fig-5 : Segmentation

3.2.2 Morphological Filtering

Morphological image processing is a collection of nonlinear operations related to the shape or morphology of features in an image. According to Wikipedia, morphological operations rely only on the relative ordering of pixel values, not on their numerical values, and therefore are especially suited to the processing of binary images. Morphological operations can also be applied to greyscale images such that their light transfer functions are unknown and therefore their absolute pixel values are of no or minor interest.

3.2.3 Dilation

The binary picture is extended from its initial form during the dilation process. The structural element determines how the binary picture is enlarged. In comparison to the picture itself, this structural element is tiny; its typical size is 3 by 3.

Let's define B as the structural element and X as the reference image. Equation defines the dilation operation.

XB = z|[(B)ZX]XX

X⊕B={z|[(B^)Z∩X]∈X}

Then, this structural component will be moved to the right. One of the black squares of B is found to overlap or intersect the black square of X at location. Hence, position. Black will be inserted inside the square. The structural element B is also a transition between left to right and from top to bottom on the image X to yield the dilated image.

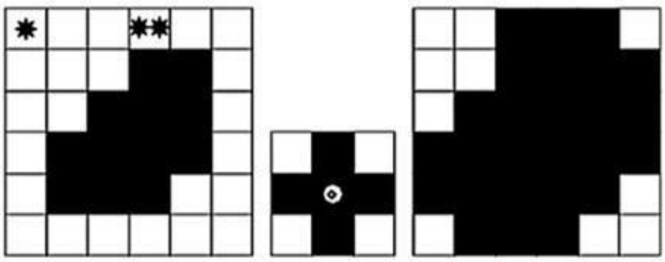

Fig-6: Dilation.

3.2.4 Erosion

Dilation's opposite process is erosion. When an image is enlarged via dilation, it is reduced in size by erosion. The structural element determines how the image is reduced. With a 3 × 3 size, the structural element is often smaller than the picture.

When compared to greater structuring-element sizes, this will guarantee faster calculation times. The erosion process will shift the structural element from left to right and from top to bottom, almost identical to the dilatation process.

X⊖B= {z|(B^)Z∈X}X⊖B={z|(B^)Z∈X}

According to the equation, the structuring element is only taken into account when it is a subset of or equal to the binary image X. Fig. 6.3 illustrates this procedure. Again, the white square denotes the number 0, while the black square denotes the number 1. Starting at point •, the erosion process begins.

Because there isn't a complete overlap in this instance, the pixel at location • will stay white.

The identical condition is then seen when the structural element is moved to the right. Because there isn't total overlapping at position u, the white colour will be applied to the black square shown by the asterisk (*).

The structuring element is then moved once more until its centre is in the location indicated by the asterisk (••).

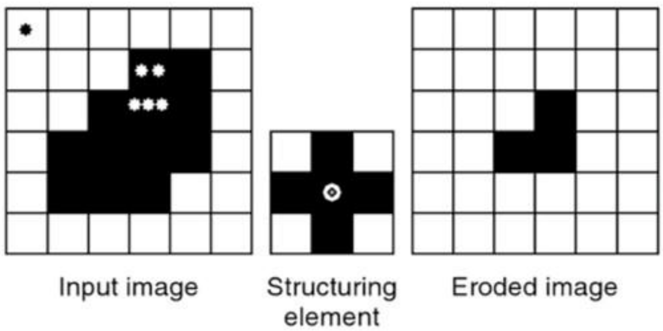

Fig-7: Erosion

There may be some "1s" in the background if the division is not continuous; this is known as background noise. Additionally, there is a chance that the system recognised a motion incorrectly; this is known as gesture noise. The aforementioned mistakes should be eliminated if we want faultless gesture contour detection. To provide a smooth, closed, and finished hand action, a morphological separating (filtering) strategy is used to group dilation (enlargement) and erosion (disintegration).

3.2.5 Extraction of Features

It is possible to use a pre-made or pre-processed photograph, but the final image loses its unique highlights. The following are the attributes that may be retrieved: Locating Contours ii.

Locating and fixing convex hull iii. Operations in

Mathematics

Fig-8: Feature Extraction.

1. Contours: This refers to the hand's orientation, regardless of whether the hand is placed vertically or horizontally. Assuming that the hand is vertical and the box bounding is ten lengths long, we first attempt to determine the orientation by length to breadth ratio.

They will be wider than the same-sized box, and if the hand is horizontally positioned so that the bounding box's width is more than the bounding box's width, they will be longer than the box's

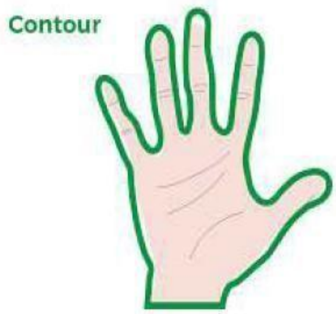

Fig-9: Detection of contour.

Identifying and correcting convex hulls: A hand posture may be identified by its orientation and the number of fingers that are visible. We need to process just a portion of the finger on the hand that we have previously processed by determining and evaluating the centroid in order to acquire the total number of fingers that are exhibited in hand motions.

3. This may be determined using the following formula: angle = math.acos (((b**2 + c**2 - a**2)/ (2*b*c)) * 57 In order to distinguish between the various fingers and to identify each one individually, this formula calculates the angles between the two fingers. In order to extract the right number of raised fingers into the image, we may additionally calculate the length of each raised or compressed finger's coordinate points using the centroid as a source of perspective.

The input video and turn it into frames. The test results are displayed below:

1. The procured RGB picture has to be pre-processed, as illustrated in figure 7.1, before the components are separated and an acknowledgement is created.

2. We utilised Otsu's thresholding method in this research.

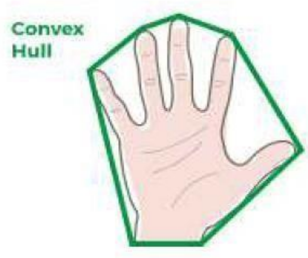

Fig-10: Detected Convex Hull.

Cluster-based thresholding is carried out automatically using Otsu's thresholding. One threshold approach is used to divide the image's pixel histogram into its component parts.

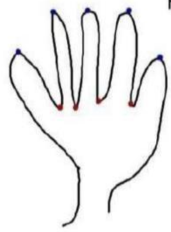

Fig-11: convex and defects points.

Contours are the continuous points on the border that change in integrity and have the same hue or intensity. The contours are an excellent tool for item detection and recognition as well

4. RESULTS

we can get the output by running the python code in any integrated development environment here we use pycharm which is an IDE. Here we will be able to control a video player

The suggested system performs different operations using hand gestures, namely the number of raised fingers in the region of interest. A hand gesture recognition system, as seen in figure , identifies shapes and/or orientations in order to assign the system a task to

4.1 Results for controlling a video player

Now the below figure shows hand landmarks detection based on mediapipe.

1.When we give our hand as an input the video is captued and land marks are detected

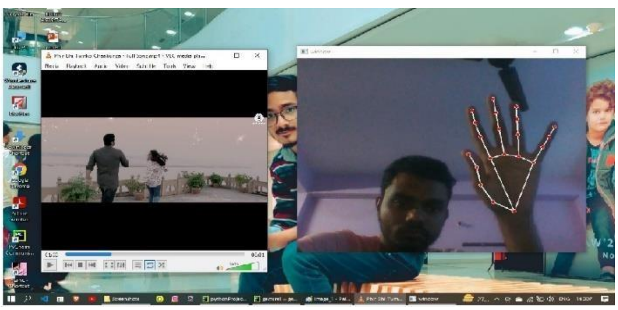

Fig-12 : Detection of Hand.

2.When we fold our hand the video player is paused

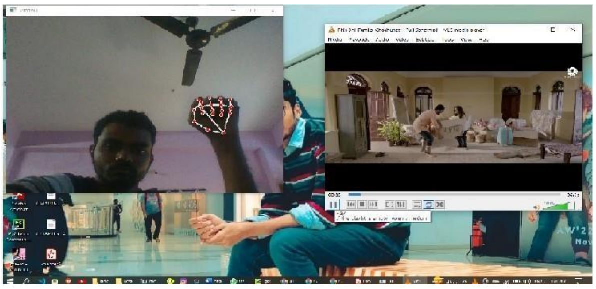

Fig-13: Pausing of a Video Player.

3. when we open or stretch our five fingers video starts playing.

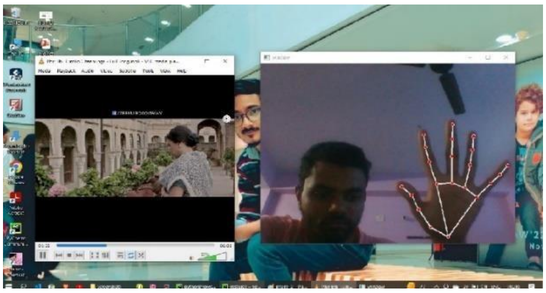

Fig-14: Video Starts Playing.

4. When we raise our middle finger,ring finger and little finger the volume is raised or increases.

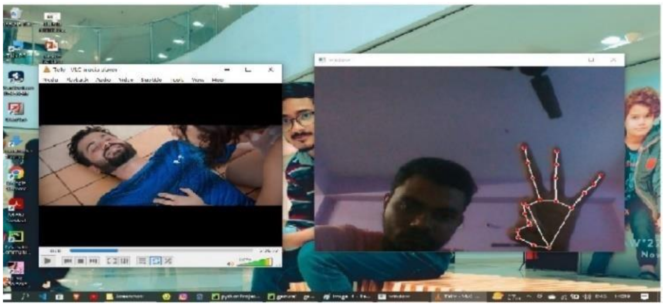

Fig-15: Volume Increases.

5. When we lower our four fingers leaving thumb finger volume decrease.

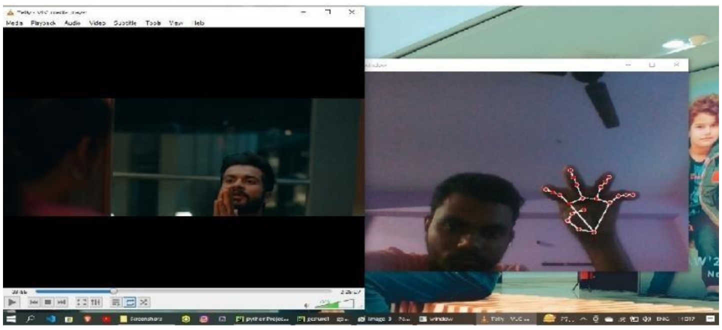

Fig-16: Volume Decreases.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 01 | Jan 2023 www.irjet.net p-ISSN: 2395-0072

6. Use the small finger to to move forward the video by ten seconds

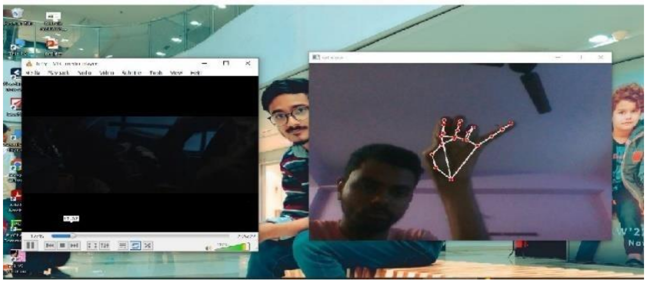

Fig-17: Forward

7.Raise two fingers to rewind the video

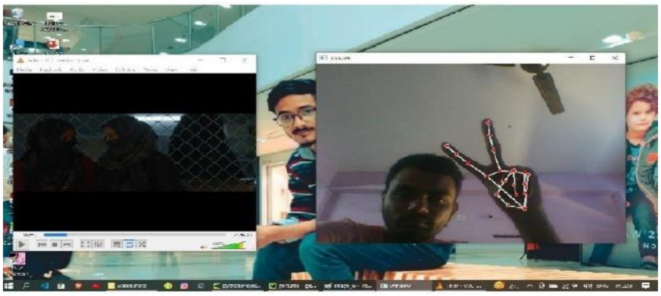

Fig-18: Rewind

4.2 Results for controlling a presentation

8. To get a pointer on slide use index and middle finger join two fingers then we get a pointer on screen.

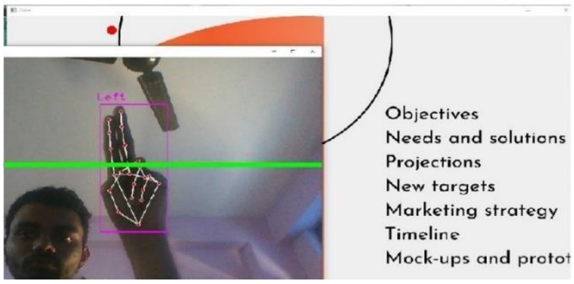

Fig-19: Pointer.

9. To draw or write on slide you can use single index finger

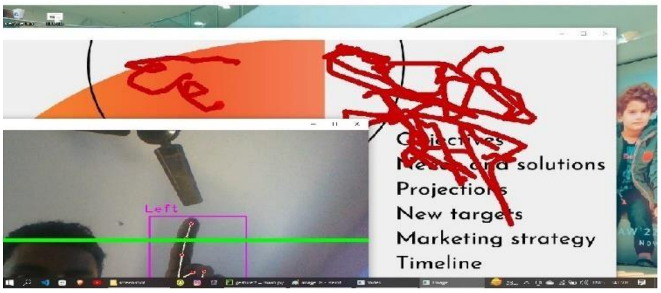

Fig-20: Draw/Write

10.To erase the drawing or the written text raise index, middle and ring fingers.

Fig-21: Erase

11. To change to next slide use the smallest finger.

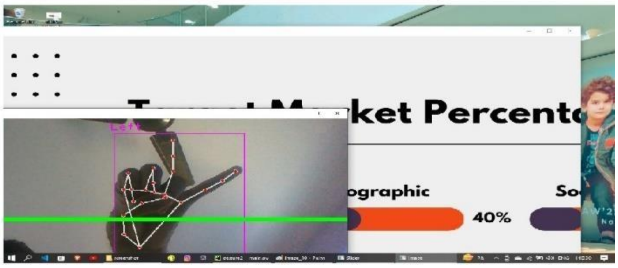

Fig-22: Next slide

11.To go to previous slide move your thumb finger upwards.

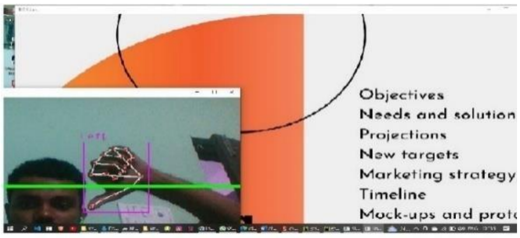

Fig-23: Previous slide.

12. When we move our eyes it acts as a cursor.

Fig-24: Cursor.

5. CONCLUSION

The Hand gesture detection system for controlling UI, a standalone application for manipulating the different user interface controls and/or applications like VLC Media Player, was conceived, created, and built in this project. During the analysis phase, we gathered data on the numerous gesture recognition systems now in use, their methodologies, algorithms, and success/failure rates. As a result, we carefully compared various solutions and evaluated their effectiveness. We created the system architecture diagrams and the data flow diagram of the

When converting a grayscale image to a binary image, the notion of thresholding states is used to locate outlines. For instance, despite the form around the hand, the lighting over the photograph of the hand may be uneven as a result of the contours that were drawn around the dark areas. That should be prevented by altering the limit. For an accurate examination of gesture recognition, the background of the images should be clear. Additional sporadic monitoring minutes are useful for ensuring that the shapes of the layout picture and the individual's picture are the same.

6. REFERENCES

[1] Oudah M, Al-Naji A, Chahl J. Hand Gesture Recognition Based on Computer Vision: A Review of Techniques. J Imaging. 2020 Jul 23;6(8):73. doi: 10.3390/jimaging6080073. PMID: 34460688; PMCID: PMC8321080.

[2] Karapinar Senturk, Z., & Bakay, M. S. (2021). Machine Learning Based Hand Gesture Recognition via EMG Data. ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, 10(2). https://doi.org/10.14201/ADCAIJ2021102123136

[3] Bakheet, S., Al-Hamadi, A. Robust hand gesture recognition using multiple shape-oriented visual cues. J Image Video Proc. 2021, 26 (2021). https://doi.org/10.1186/s13640-021-00567-1

[4] Abhilash Dayanandan, Akshay Chakkungal, Anooj Kommeri, Deepak Koppuliparam bil, Dr. Prashant Nitnaware (May 2020). Gesture Controlled Media Player using TinyYoloV3. IRJET: International Research Journal of Engineering and J.M. (2020). A wearable biosensing system with in-sensor adaptive machine learning for hand gesture recognition. Nature Electronics, 4, 54-63.

[7] Rubin Bose, S. and Sathiesh Kumar, V. 'In-situ Identification and Recognition of Multi-hand Gestures Using Optimized Deep Residual Network'. 1 Jan. 2021 : 6983 – 6997. kkkkkk

[8] Verdadero, M. S., Martinez-Ojeda, C. O., & Cruz, J. C. D. (2018). Hand Gesture Recognition System as an Alternative Interface for Remote Controlled Home Appliances. In 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM) (pp. 1-5). IEEE.

[9] https://www.geeksforgeeks.org/find-and-draw-contoursusing-opencv-python/

[10] https://learnopencv.com/contour-detection-using-opencvpython-c/

[11] https://learnopencv.com/convex-hull-using-opencv-inpython-and-c/

[12] https://theailearner.com/2020/11/09/convexity-defectsopencv/

[13] Research Gate, Google.

[14] Cevikcan, Ustunug A, Industry 4.0: Managing The Digital Transformation, Springer Series in Advanced Manufacturing, Switzerland. 2018. the following DOI: 10.1007/978-3-319-57870-5.

[15] Pantic M, Nijholt A, Pentland A, Huanag TS, HumanCentered Intelligent Human-Computer Interaction (HCI2): How Far We From Attaining It?,International Jounal of Autonomous and Adaptive Communications Systems (IJAACS), vol.1 no.2, 2008. pp 168-187. DOI: 10.1504/IJAACS.2008.019799.

[16] Hamed Al-Saedi A.K, Hassin Al-Asadi A, Survey of Hand Gesture Recognition System. IOP Conferences Series: Journal of Physics: Conferences Series 1294 042003. 2019.