Gesture-Based Control of Multimedia Player Using Python and Open CV

Nikita Kashyap[1*], Pragati Patharia[2], Arun Kumar Kashyap[3]
Department Of Electronics And Communication Engineering, School Of [1,2]Engineering And Technology, Guru Ghasidas Vishwavidyalaya, Bilaspur (C.G.)
[3]Department Of Biotechnology, Govt. E Raghavendra Rao Pg. Science College, Bilaspur (C.G.)
Corresponding Author: Nikita Kashyap nikita2010k@gmail.com

ARTICLEINFO

Keywords: Hand Gesture, Opencv, Python, Multimedia Player

Received : 3 November
Revised : 19 November
Accepted: 22 December

ABSTRACT

The increasing significance of computers in our daily lives, coupled with the rise of ubiquitous computing, has necessitated effective human-computer interaction. Hand gesture recognition systems have emerged as a real-time video-based solution for detecting and interpreting hand gestures, offering intelligent and natural human-computer interaction (HCI) methods. This project focuses on leveraging human hands as input devices for computer operation. Developed using Python and the OpenCV library, the program utilizes a computer webcam to capture and analyze hand shapes and patterns. The program provides real-time feedback by displaying recognized hand gestures on the live video stream. The ultimate outcome of this project is an application that enhances user experiences in contactless systems. The project recognizes and detects human hand motions using the Python computer language through a process flow that includes background subtraction, hand ROI segmentation, contour detection, and finger recognition. Techniques for processing images are used, including hand gesture detection, pattern recognition, thresholding, and contour detection. The processing of incoming photos and the creation of related keystrokes are made possible by OpenCV, a rich set of image processing tools

DOI: https://doi.org/10.59890/ijatss.v1i4.983
E-ISSN: 3026-2410
https://journal.multitechpublisher.com/index.php/ijatss

INTRODUCTION

Speech is the predominant mode of communication among humans, but non-verbal forms of interaction also play a significant role. Although verbal communication offers unambiguous signals, non-verbal indicators like hand gestures can express nuances that are difficult to understand through verbal communication alone. By offering a natural and intuitive form of connection, the creation of a hand gesture recognition system seeks to close the gap between humans and computers [1]. This system facilitates effective communication between humans and machines by recognizing and interpreting hand gestures. The recognized gestures can be employed for various purposes, such as monitoring robots or conveying meaningful information. Hand gestures serve as a rich source of information that may not be easily conveyed through other channels. Therefore, this project focuses on exploring methodologies to detect and interpret the language of hand gestures, enabling more efficient and intuitive communication [2,3].

LITERATURE REVIEW

The ability of computers to comprehend human hand movements or gestures has the potential to streamline tasks and reduce the gap between human and machine interaction [4]. Hand gesture recognition finds applications in numerous fields, including image processing, cyber security, and robotics, among others. Consequently, it has become an active and extensively researched area [4-6]. Numerous experiments have been conducted to validate the effectiveness of gesture detection, examining the results and applications of these systems. These systems provide insights into new techniques for recognizing emotions, thereby expanding our understanding and skills in the area of human-computer interaction

METHODOLOGY

In this research, human hand gesture detection and recognition are achieved using a CNN-based classification approach. The system's overall process flow includes an array of steps, such as hand ROI segmentation using a mask picture, finger segmentation, normalization of segmented finger images, and finger recognition using a CNN classifier [7]. The accompanying picture, which depicts the sequential processes and their relationships, provides a visual representation of the suggested flow of the hand gesture recognition system.

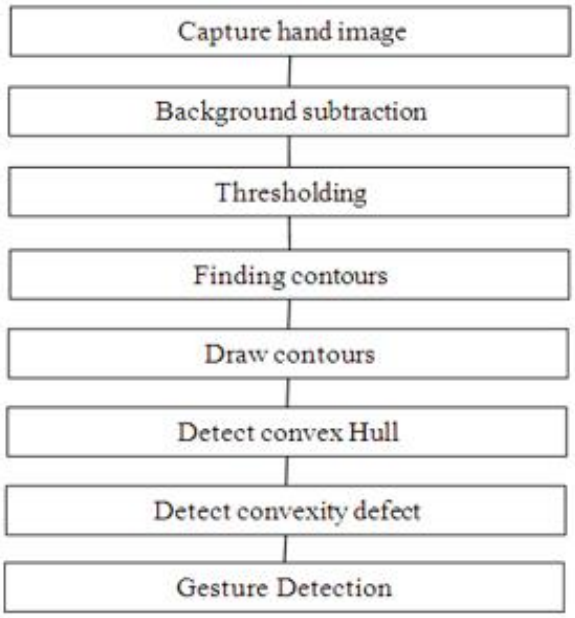


Fig 1. Flow Chart of the Methodology

The system utilizes the Python programming language along with the OpenCV library. For hand gesture identification, it uses a pre-trained transfer learning CNN model. A deep learning technique called transfer learning is using a model created for one task as the basis for another. In this instance, convolutional neural networks (CNNs) created especially for hand identification and recognition are used for testing and training the pre-trained model. The model is created using the Keras and TensorFlow tools, with TensorFlow serving as the backend for effective computations and optimizations and Keras acting as the frontend for Python-based model training. Five steps make up our system, and they are as follows:

1. Image Capturing- Using a Webcam

RGB images are frame by frame acquired during the image-capturing stage. The goal is to photograph the hand without a glove or any other accessories [8].

2. Preprocessing- for Pre-Processing

an important area called the Region of Interest (ROI) is extracted from the video stream instead of processing the entire frame. This selection helps minimize computation time. The ROI is then converted from color to grayscale to enhance processing efficiency. After completing the necessary processing steps, the grayscale image can be restored to its original color space. To further refine the ROI, a Gaussian blur is applied. This blurring technique reduces the impact of high-frequency objects that are not the target of interest. It is important to note that any camera vibrations during this step can affect the algorithm's performance and may lead to inaccuracies [8].

3. Hand Region Segmentation

This process is essential to hand gesture detection systems because it enhances system performance by removing unnecessary information from the video stream. The two main approaches for detecting hands in photos are skin color-based and shape-based. The skin-color-based approach is straightforward yet sensitive to changes in background and illumination [14, 13]. However, the shape-based approach, which relies less on outside influences, uses the convexity principle for hand detection. Edge detection, RGB value analysis, and background subtraction are a few methods that can be utilized to extract the hand region from an image [8, 9]. The background subtraction approach is used in this study to distinguish the hand from the background. Using the running average principle, the system calculates the average backdrop image by focusing on a particular scene for a sufficient number of frames. The background image threshold is calculated by the running average using a weighted algorithm. The hand is positioned in front of the camera after the background has been determined, and the absolute difference between the current frame (which includes the hand as a foreground item) and the backdrop is calculated [10,18]. Background subtraction, a technique that aids in separating the hand region. The hand region is made white and the remainder of the image is black after the image has been thresholded. Applying a threshold is essential for achieving accurate hand segmentation. Mathematically, the thresholding principle can be represented as a function of the pixel intensity

$$f(x) = \begin{cases} 1, \; if \;\; x \geq threshold \\ 0, \; if \;\; x < threshold \end{cases}$$

The motion detection output is the image that results from background removal and thresholding. The results of the hand region segmentation technique are shown in Figure. To remove any remaining minor noise patches, a series of morphological operations like erosion and dilation are carried out.

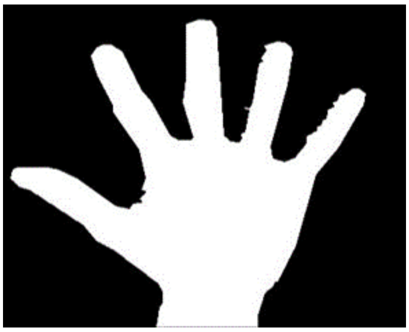


Fig 2. Output of Hand Region Segmentation Process [10]

By following these processes, the hand region is effectively segmented from the image, serving as a crucial step in subsequent hand gesture recognition tasks.

4. Contour-Extraction

Contours are defined as the boundaries or outlines of objects, such as the hand in our case, that are present in an image. These contours are formed by connecting points with similar color values. Contours play a crucial role in shape analysis, object detection, and recognition processes [11].

5. Features Extraction and Recognition

Now we turn to the second part of the research, which is how we determine the number of fingers for recognizing hand gestures. To achieve this, the Convex Hull method is employed [13]. The Convex Hull helps identify the extreme points of the hand, including the top, bottom, left, and right points. These extreme points are crucial for understanding the hand's shape and can be visualized as a group of points surrounding the hand region, as depicted in Figure 3. The convex hull method analyzes the shape's outline and identifies convex and concave points [15]. For a hand, 5 convex points (one per finger) and 4 defects (between adjacent fingers) are expected. By counting these points, we can determine the number of fingers displayed. OpenCV functions like findContours(), convexHull(), and convexityDefects() are used to obtain these points [11]. The output includes centroid coordinates and the number of defects. Additionally, a circle is drawn around the fingers using the palm's center as the center point and a radius of 70% of the maximum distance between the center and extreme points. Deviations from this hull are considered convexity defects [17].

To visualize the hand gesture recognition process, we can refer to an image. As shown in Figure 4, we draw a line linking the start and finish points, and we insert a circle at the farthest point.

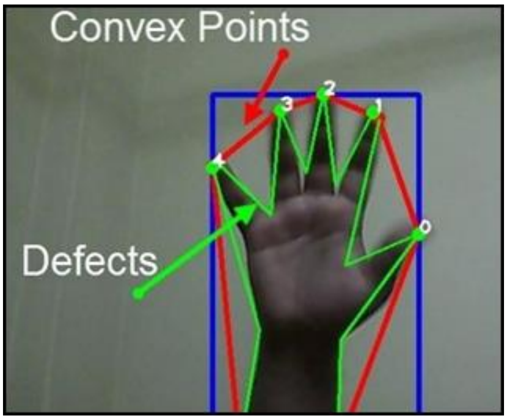


Fig3. Hand Convex Hull [10]

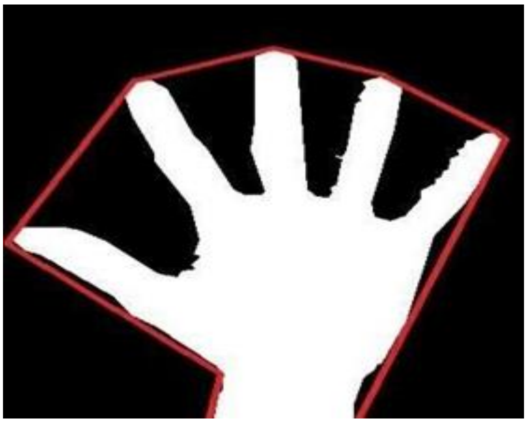


Fig 4. Convexity Defects [11]

In order to detect the angle between fingers, the cosine rule is applied to each defect. The triangle's sides are inversely proportional to one of its angles according to the cosine rule, as depicted in Figure 5. By calculating the angle using the formula [11], we can determine if it is less than 90 degrees, indicating the presence of a finger.

$$C = \sqrt{a^2+b^2 - 2ab\cos\gamma}$$

For Finding Gamma this Formula is Used:

$$\gamma = \cos^{-1}\left(\frac{a^2+b^2 - c^2}{2ab}\right)$$

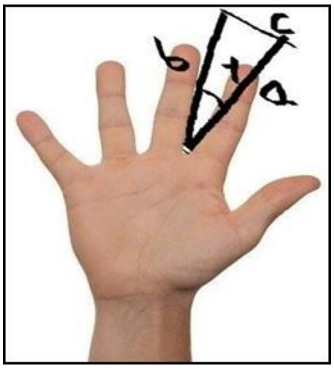


Fig 5. Finding Angle Between Fingers [11]

Once the angle is determined, a circle is drawn around the approximate distance of the farthest point. The finger count is displayed using text in the image, as depicted in Figure 6.

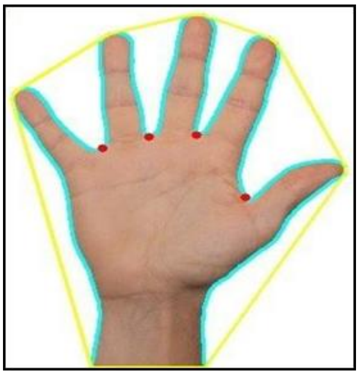


Fig 6. Counting Fingers [11]

The process of distinguishing hand gestures is continuous. After executing the necessary instructions based on the recognized gesture, the system returns to the first step to process the next image and continue the recognition process.

RESULTS AND *DISCUSSION*

Python Version 7 opensource simulation software is used to replicate the suggested hand motion detection and recognition methods. Python scientific distributions have approved this open-source application. The NumPy, OpenCV, and PyAutoGUI library modules are part of the Python software package. These modules are open source and available without a license. Each module is integrated into the Python kernel, and the proposed work is simulated using the Python programming language. This study aims to improve the accuracy of hand gesture recognition. The user doesn't need to wear any sort of hand gloves because the design is so straightforward. Although a regular PC with a web camera can run this recognition software. The simulation portion is highlighted in this result section. After implementing, we obtained the following results:

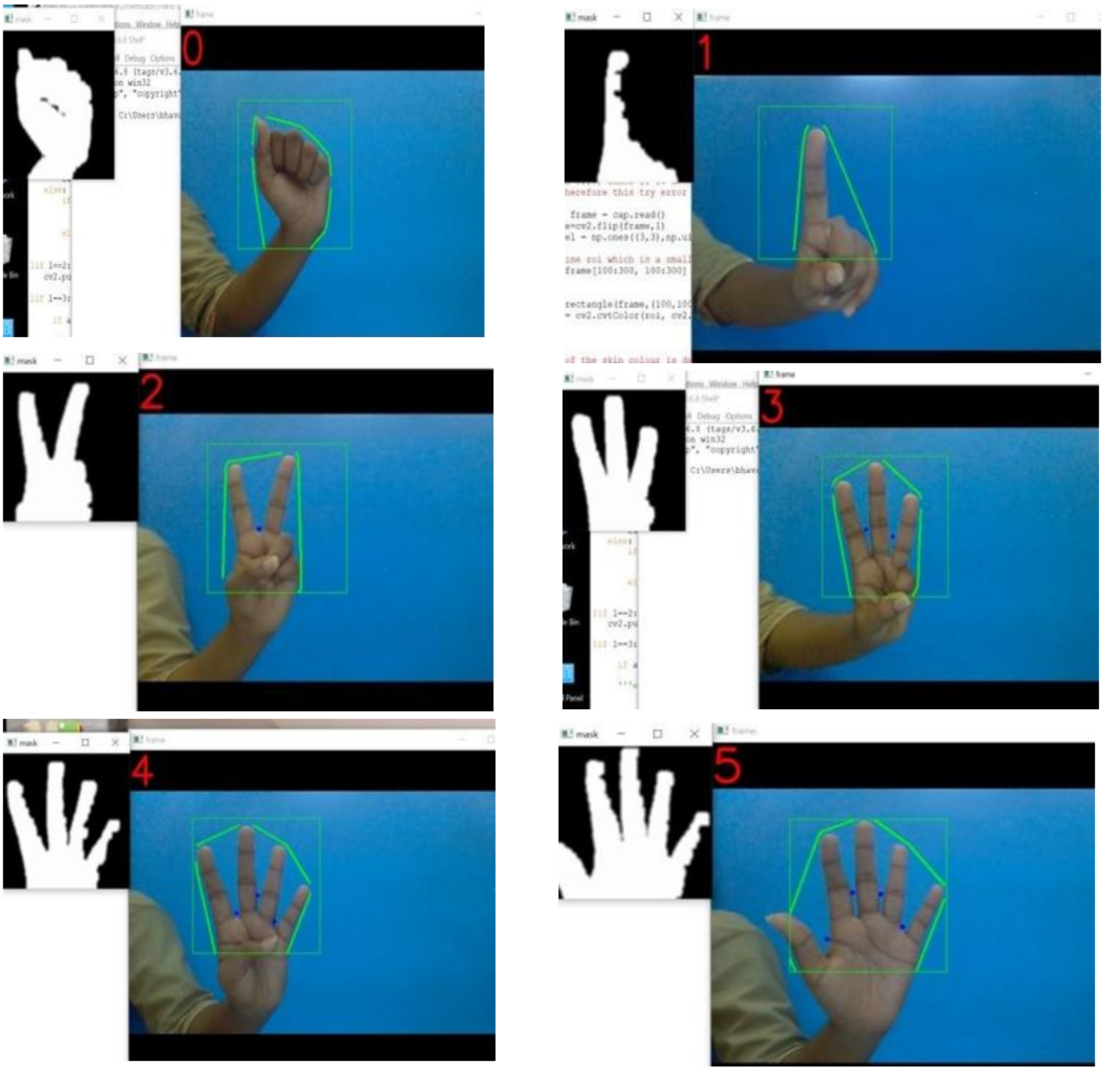


Fig 7 Python Version

During our testing, we evaluated the system in various conditions. We were pleased to find that it achieved an impressive accuracy rate of 98.7% and a recognition rate of 96.6%. It performed exceptionally well when the background was clear and the lighting conditions were moderate, resulting in the highest levels of accuracy.

This study focuses on a hand gesture-based control mechanism for a multimedia player. Users of the system can employ a variety of hand gestures to operate the multimedia player, including those for play, pause, full screen, and stop, increase volume, and reduce volume. The system divides the foreground of the picture using low-cost methods such as skin recognition and an approximate median algorithm [12, 16]. This enables the recognition of hand gestures in a natural and intuitive manner. The process involves capturing the hand image, converting it to the HSV color space, tracking the hand based on color, creating a color-based mask, and filtering the actual color, as shown in Figure 8.

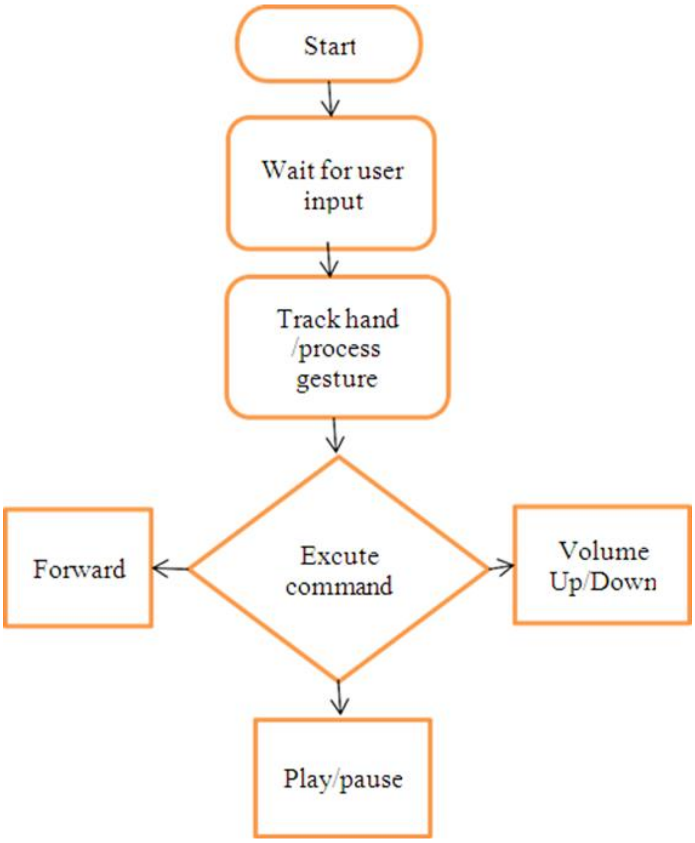
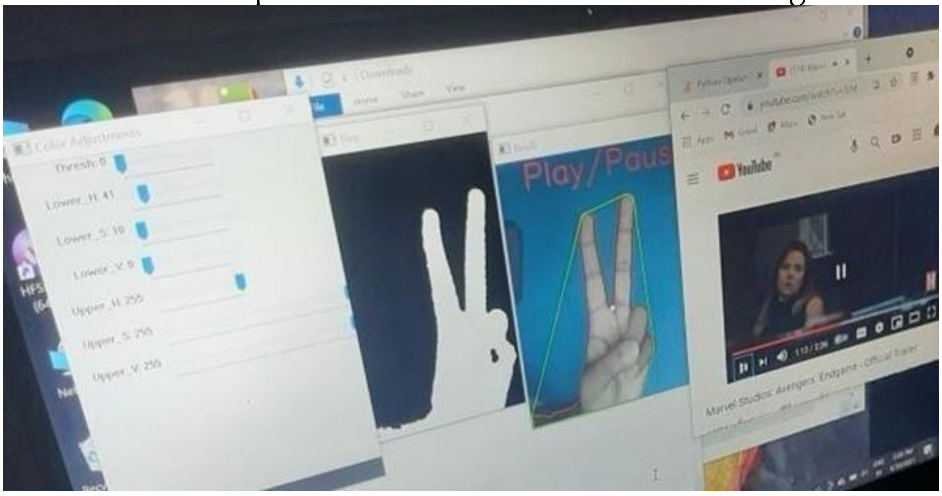


Fig 8. Flow Chart for Controlling the Multimedia Player

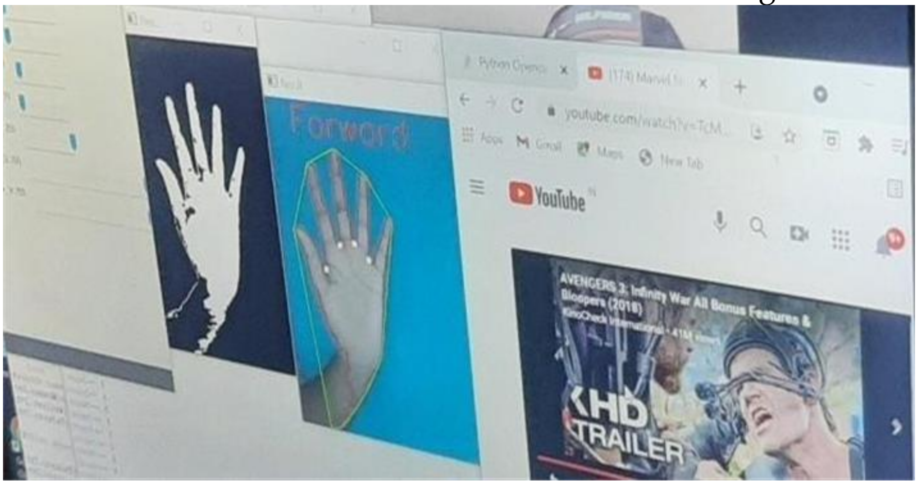
By detecting contours and analyzing convexity defects, the system identifies specific hand gestures and performs corresponding actions. The recognized gestures are then mapped to keyboard keys for controlling the media player. This approach provides a dynamic and interactive way to control media playback using hand gestures.

Upon completing the implementation of the aforementioned steps, we obtained the following results:

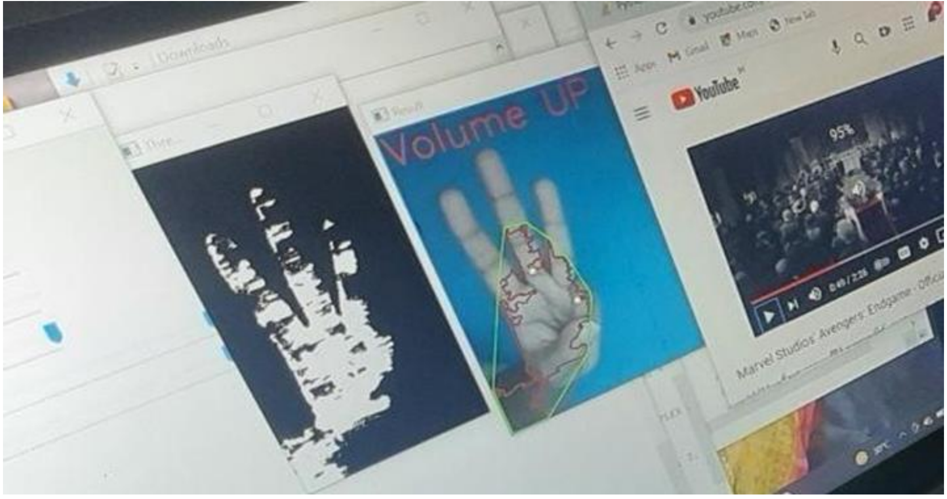
1.The Video Starts or Stops When the User Makes the Two-Finger Motion



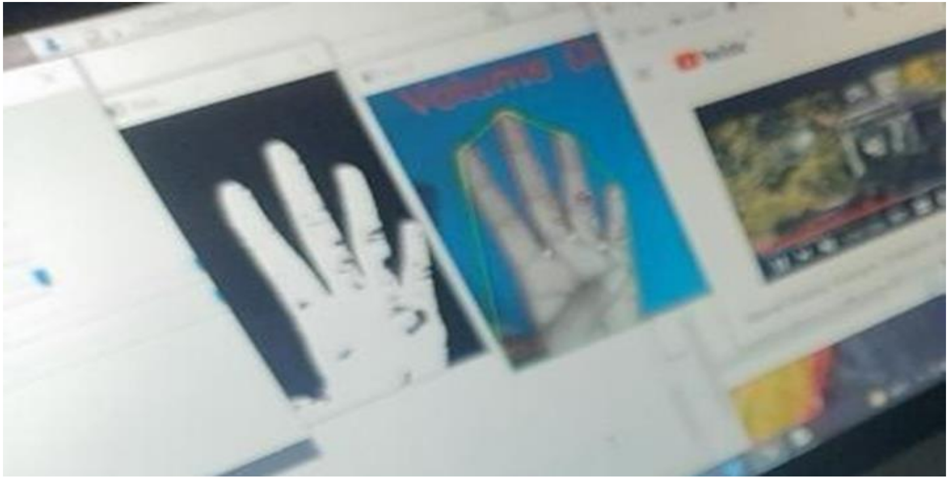
2.The Video is Forwarded When the User Makes the Five-Finger Motion


Forward
youtube.com/watch?v=TcM...
YouTube
AVENGERS 3: Infinity War All Bonus Features & Bloopers (2018)
HD TRAILER

3. The Volume is Up When a User Makes the Three-Finger Motion



4. The Volume is Down When a User Makes the Four-Finger Gesture



These figures demonstrate the successful control of the media player through the utilization of hand gestures

CONCLUSIONS AND RECOMMENDATIONS

In conclusion, this project successfully used Python and OpenCV to build hand gesture recognition by adding hand segmentation and detection methods. The objectives of the project were achieved, which include establishing a comprehensive system to detect, recognize, and interpret hand gestures through computer vision. Additionally, the system was able to generate various numbers and sign languages for effective communication. The increasing popularity of hand gesture recognition is evident in the advancements made by companies such as Microsoft, Samsung, and Sony, with applications spanning across entertainment, artificial intelligence, education, medical, and automation fields. With ongoing research and development, the adoption of gesture recognition technology is expected to become more accessible and cost-effective, combining the power of technology with human interaction.

FURTHER STUDY

This research still has limitations, so it is necessary to carry out further research related to the topic of Gesture-Based Control of Multimedia Player Using Python and OpenCV in order to improve this research and add insight to readers.

REFERENCES

[Ahmad Puad Ismail1, Farah Athirah Abd Aziz1, Nazirah Mohamat Kasim1 and Kamarulazhar Daud, “Hand gesture recognition on python and opencv”, IOP Conf. Series: Materials Science and Engineering, Vol. 1045, 2020.

Aashni Haria, Archanasri Subramanian, Nivedhitha Asokkumar, Shristi Poddar, Jyothi S Nayak, “Hand Gesture Recognition for Human Computer Interaction”, Procedia Computer Science, Vol. 115, pp. 367-374, ISSN 1877-0509, 2017.

Ahmed Kadem Hamed AlSaedi, Abbas H. Hassin AlAsadi, “An efficient hand gestures recognition system,” IOP Conf. Series: Materials Science and Engineering, 745, 2020.

B. Nandwana, S. Tazi, S. Trivedi, D. Kumar and S. K. Vipparthi, "A survey paper on hand gesture recognition," 2017 7th International Conference on Communication Systems and Network Technologies (CSNT), pp. 147-152, 2017, doi: 10.1109/CSNT.2017.8418527.

[Bilvika, K & K, Sneha & M, Sahana & Patil, Tejaswini, “Face and Hand Gesture Recognition System for Controlling VLC Media Player,” International Journal of Scientific Research in Science and Technology. 119-124, 2021.

C. R. P. Dionisio and R. M. Cesar, "A project for hand gesture recognition," Proceedings 13th Brazilian Symposium on Computer Graphics and Image Processing (Cat. No.PR00878), Gramado, Brazil, 2000, doi: 10.1109/SIBGRA.2000.895833.

Daluma, M. S. & Saleh, M. (2017). Foreign Direct Investment and Economic Growth: A Test of Granger Causality in Nigeria. African Journal of Research and Sustainable Development 5(2):2856 – 2011, 2017.

Daluma, M. S. & Saleh, M. (2017). Growth of Non-Oil Sectors: A Key to Diversification and Economic Development in Nigeria. African Journal of Research and Sustainable Development 5(2):2856 – 2011, 2017.

I. Ralev and G. Krastev, "Components and model of implementation of a hand gesture recognition system," 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Ankara, Turkey, pp. 1-4, 2022, doi: 10.1109/HORA55278.2022.9799919.

M. Panwar and P. Singh Mehra, "Hand gesture recognition for human computer interaction," 2011 International Conference on Image Information Processing, pp. 1-7, 2011, doi: 10.1109/ICIIP.2011.6108940.

Madhav Mishra, "Hand Detection and Finger Counting Using OpenCV-Python," Analytics Vidhya, 2020.

P S, Dr-Neethu & Ramadass, Suguna & Sathish, Divya. "An efficient method for human hand gesture detection and recognition using deep learning convolutional neural networks,". Soft Computing, 2020.

S. N. Karishma and V. Lathasree, "Fusion of Skin Color Detection and Background Subtraction for Hand Gesture Segmentation", International Journal of Engineering Research and Technology, vol. 3 no 1 pp 13–18, 2014.

Saleh M. & Daluma M. S. (2017). Impact of Tax Reform on Federal Revenue Generation in Nigeria. African Journal of Innovation in Management Science and Administration 6(1):2856 – 2011, 2017.

Saleh, M. (2017). Effect of Internal Control on Fraud Prevention of Listed Manufacturing Firms in Nigeria. International Journal of Business and Management; Vol. 12, No.3 2017.

Saleh, M., Daluma, M. S. & Mailafia, L. (2021). Effect of Randomized Algorithms on Conservative Structural Equation Model Users. Sub-Sahara Journal of Science+ Business Media 4(11), 222-239, 2021.

Saleh, M., Effect of Board of Directors' Attributes on the Financial Performance of Listed Deposit Money Banks in Nigeria. Developing Country Studies ISSN 2224-6007X (Paper) ISSN 2425-0465(Online). 2(10) - 2016

Sarma, D., Bhuyan, M.K., "Methods, Databases and Recent Advancement of Vision-Based Hand Gesture Recognition for HCI Systems: A Review," SN COMPUT. SCI. 2, 436, 2021.