iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VI **Month of publication:** June 2022

DOI: https://doi.org/10.22214/ijraset.2022.43645

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

HCI Based Virtual Controlling System

Sakshi Thorat[1], Sakshi Wakade[2], Sushant Jogdand[3], Yashraj Chavan[4], Prof. Neha Chaube[5]
[1,2,3,4] *Students, Dept of CSE, MIT School of Engineering, Pune, Maharashtra, India*
[5] *Professor, Dept of CSE, MIT School of Engineering, Pune, Maharashtra, India*

Abstract: ***Researchers around the globe are working on making our devices more interactive and making them function with minimal physical contact in this research project. The proposed system is an interactive computer system that can operate without a physical keyboard or mouse. This system will benefit everyone, particularly immobilized people with special needs operating a physical keyboard and mouse. So in the system, they have developed an interface that uses visual hand-gesture analysis. These gestures are used to assist those who are having trouble controlling or operating computers or gadgets. The model is being developed in such a way that it can assist us in recognizing and implementing it. Regarding hand gestures, our interface uses OpenCV, Python, and computer vision algorithms that can detect different finger orientations, distinguish the user's hand from the background, and distinguish significant hand movements from unwanted hand movements.***
Keywords: Machine Learning, Computer Vision, Image Processing, OpenCV, Human-Computer Interaction (HCI).

I. INTRODUCTION

We live in a technologically advanced world, and technology is at the heart of our daily existence. Technology has progressed to the point where you can do a variety of tasks with a single touch. Technology enables us to fly around the world. Today, however, we don't even need to touch the devices; we can simply use our voices to manage everything. The world has become a better place to live because of technologies like Siri and Alexa. Human-to-human communication has progressed to human-to-machine communication. In the world of technology, text-to-speech has become a fad.

Machines have been trained by humans for a long time. However, better human-machine interaction is required to help machines better understand humans. The hand movements are about to begin. In the world of technology, hand gestures have grown rather popular. One of the most important technological breakthroughs has been the ability to use hand gestures without touching the device.

Hand gestures can be used to control the entire gadget. The most efficient technique for manipulating technology from now on is to train the computer by simply making simple hand movements. The purpose of gesture interpretation The goal is to push the boundaries of advanced human-computer communication, bringing HCI performance closer to that of human-human interaction.

Human-Computer Interaction focuses on the study of human-computer interfaces. HCI is a field that straddles computer science and behavioral science. From desktops to mobile screens and handheld computers, human-computer interaction has progressed over time. HCI has paved the way for technologies like eye-trackers, wearables, and virtual assistants.

Hand gestures can be used to control computer functions, with the system recognizing the hand gesture and performing the function accordingly. It has the ability to act like a computer's cursor.

II. LITERATURE REVIEW

The author focuses on the development of hand gesture models using convolution neural networks and backpropagation techniques to make these models easier to use. The author describes the findings of hand gesture recognition using deep learning models on the Kaggle dataset. Upon successful implementation, it was observed that accuracy of 1.00 was attained. Furthermore, despite having additional resources, it was discovered that CNN outperformed all other models in terms of training iterations. CNN was shown to be the best model for hand gestures, according to the authors. [1]

The authors wanted to create a virtual game that recognized hand gestures. They used image processing techniques for virtual game interactions, which were implemented in C++ using the OpenCV library. To interact with the virtual game, they used four-handed motions. After implementation, the outcome is represented graphically [2].

The authors study proposes an effective representation of a blurred picture detection system based on edge type and sharpness analysis using the Laplacian operator, which can determine whether the image is blurred or not, as well as the extent of a blur, using the Variance of Laplacian. This project presents a simple, reliable, and fast picture noise estimation approach. The results show that the proposed algorithm performs well for different types of images over a large range of noise variances [3].

©IJRASET: All Rights are Reserved

The author has discussed extraction methodologies and image preprocessing approaches, as well as application areas for hand gestures. The authors used a multivariate Gaussian distribution to discern the non-geometric aspects of hand movements. Various methods for gesture detection are explored in this paper, including neural networks, HMMs, fuzzy c-means clustering, and the use of an orientation histogram for feature representation. There is a detailed overview of modern recognition systems as well as an explanation of gesture recognition challenges [4].

The author created a hand gesture-based real-time human-computer interaction system. They developed a simple method during the HCI stage to avoid erroneous recognition induced by noises (primarily brief, false motions) and thereby increase interaction reliability. The proposed system is extremely expandable, allowing it to be used in human-robotic or other human-machine interaction scenarios that require more complicated command formats than mouse and keyboard events [5].

III. PROPOSED METHODOLOGY

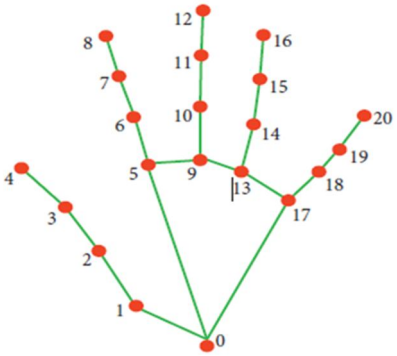
A. System Development

The flowchart of the real-time virtual mouse system explains the various functions and conditions used in the system.

The Virtual Mouse System employs a camera. The proposed AI virtual mouse system is based on the frames captured by a laptop or PC's webcam. The video capture object is created using the Python computer vision library OpenCV, and the web camera begins capturing video, as shown in the figure. The web camera captures images and sends them to the virtual system.

Capturing and processing the video The virtual mouse system uses a webcam to capture each frame until the program is terminated. Video frames are converted from BGR to RGB colour space to find the hands in the video frame by frame, as shown in the following code:

```
def findHands(self, img, draw = True):
imgRGB=cv2.cvtColor(img,cv2.COLOR_BGR2RG)
self.results = self.hands.process(imgRGB)
```



0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP

Fig - 1: Fingertip Ids.

Detecting which finger is up and performing the appropriate mouse function At this stage, Ate detects which finger is up by using the tip ID of the respective finger discovered via the MediaPipe and the respective coordinates of the fingers that are up, and the appropriate mouse function is executed as a result.

Computer Vision is used to move the mouse cursor around the computer window based on hand gestures and hand tip detection. The mouse pointer is made to travel around the computer window using the AutoPy program if theIf the index finger with tip Id = 1 is up or both the index finger with tip Id = 1 and the middle finger with tip Id = 2 is up, the mouse cursor is made to roam around the computer window using the AutoPy application.

To perform a left-button click with the mouse. The computer is directed to utilise the pynput to conduct a left button click if both the index finger with tip Id = 1 and the thumb finger with tip Id = 0 are up and the distance between the two fingers is less than 30px.

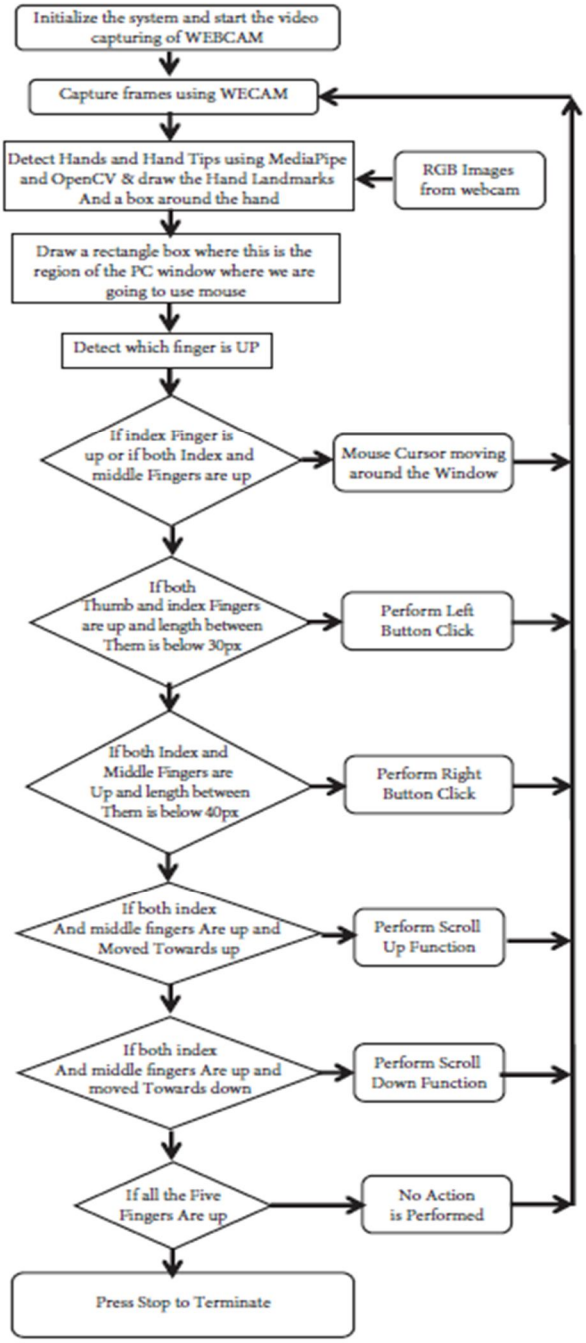


Fig -2: Flow chart

The MediaPipe framework is used for hand gesture detection and tracking, and the OpenCV library is used for computer vision. The algorithm employs machine learning concepts to track and recognize hand gestures and hand tips. MediaPipe is a Google open-source framework that is used for applying a machine learning pipeline. Because it is constructed with time-series data, the MediaPipe framework is useful for cross-platform development. The MediaPipe framework is multimodal, which means it may be used with a wide range of audio and video formats. The MediaPipe framework is used by developers to build and analyze systems using graphs, and it has also been used to develop systems for application purposes.

The steps involved in a MediaPipe-based system are carried out in the pipeline configuration. The pipeline being developed can run on a variety of platforms, allowing for scalability in mobile and desktop environments. The MediaPipe framework is composed of three major components: performance evaluation, a framework for retrieving sensor data, and a collection of reusable components known as calculators. A pipeline is a graph made up of components called "calculators," each of which is connected by streams through which data packets flow. Developers can replace or define custom calculators anywhere in the graph, allowing them to create their own applications. The calculators and streams combined form a data-flow diagram; the graph is created with MediaPipe, with each node representing a calculator and connected by streams. A single-shot detector model is used in real-time to detect and recognize a hand or palm. The MediaPipe employs a single-shot detector model. First, the hand detection module trains for a palm detection model because palms are easier to train. Furthermore, non-maximum suppression works much better on small objects like palms or fists. A hand landmark model entails locating joint or knuckle co-ordinates in the hand region,

OpenCV is a computer vision library that includes image processing algorithms for object detection. OpenCV is a Python programming language library that allows for the development of real-time computer vision applications. Image and video processing makes use of the OpenCV library.

B. Methodology

In this section, we will discuss the requirements of this project, its system architecture, and the required technologies to create it. Firstly, a web camera is being used to capture the hand gesture. It captures the hand and the fingertips. To capture these details, it has made use of MediaPipe since MediaPipe is used for applying a machine learning pipeline. Along with the media pipe, OpenCV is also used for image processing of hand gestures. In this system, a tip_id is assigned to each fingertip. Whichever tip is up, the task assigned to that tipid will be performed. Different models have been used to detect different parts of the image. models such as the hand detection model, hand landmark model, etc. All functions are performed on the basis of the fingertips.

At first, the image is captured as normal and then converted to RGB format in order to get a better understanding of the actual hand image. The color detection part starts from the RGB layer.

C. Working

When it executes the command, the system will begin capturing video through the webcam, and then it will capture frames, i.e. RGB images, based on the user's movement or gesture. Hands and hand tips will be detected using media pipe and OpenCV, and the hand landmarks and a box around the hand will be drawn. This system will now use the box around the hand, including the hand, like the mouse. It will now be determined which finger is upright. The mouse cursor will move around the window if the index finger is up or both the index and middle fingers are up. It will perform a left button click if the index finger is bent and the middle finger is upright. It will perform a right-button click if the index finger is upright and the middle finger is bent. If both the index and middle fingers are up and moved down, it will perform the scroll down function; if all five fingers are up, no action will be taken.

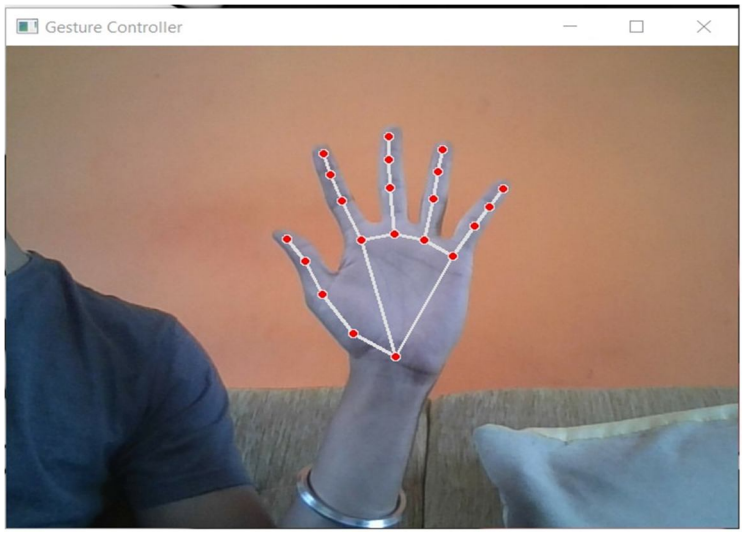


Fig - 3:Neutral GestureIt is used to stop or halt the execution of a gesture.

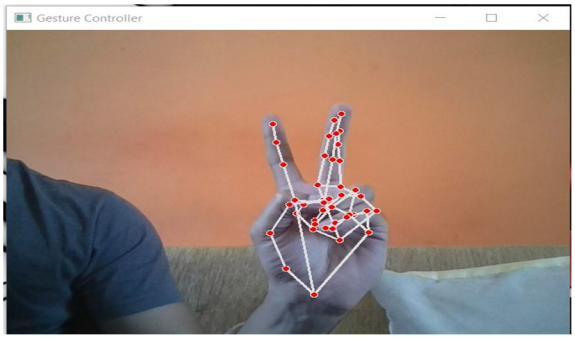


Fig - 4: The index and middle finger are used to move the cursor along with the screen.

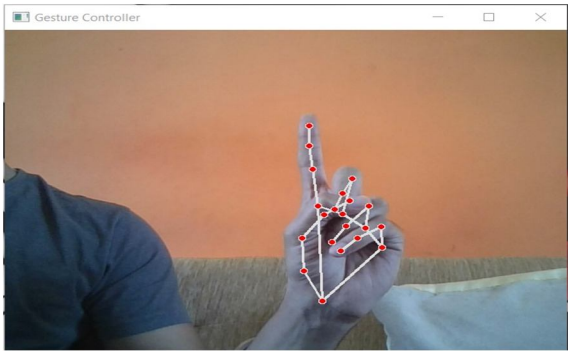


Fig - 5: Gesture for right-click.

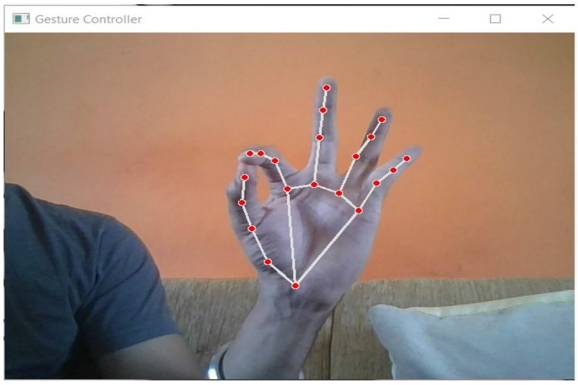


Fig - 6: The Pinch gesture is for vertical and horizontal scrolling. It is also used for volume and brightness control if those settings are open on the system.

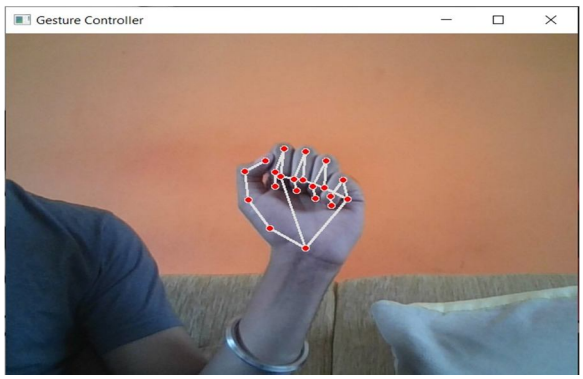


Fig - 7: Gesture for drag and drop functionality and for selecting multiple items at once.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

IV. CONCLUSIONS

A few techniques had to be implemented because accuracy and efficiency play an important role in making the program as useful as an actual physical mouse. After implementing such an application, there is a significant replacement of the physical mouse, i.e., there is no need for any physical mouse. This motion tracking mouse performs every physical mouse movement (virtual mouse). With the help of this system, tasks like delivering presentations and minimising workspace will be easier to do. The goal of this application will be met, and mouse cursor control will be implemented utilising a webcam.

A. Future Scope

With gesture recognition technology on computers, there are many excellent features that can be achieved.

With this gesture recognition system, we can use computers without any help from physical devices such as mouse and keyboards. What will be required is just a screen and high-resolution cameras.

This technology can make work easy as just one big screen can be used by researchers to do all the work.

So gesture recognition can be like this: there is a screen that shows many file details along with a scrollbar, so using your hand gesture you can scroll the screen up and down, or if you want to select the file, then point and zoom in so that the file can be opened. You can go back and forward by using your hand in the right or left direction.

Make your system more secure. The user can add a gesture that will be known only by them, and every time he needs to lock or unlock the system, he can just use the gesture. This will save so much time. If we expand this technology, we can also add gesture recognition in other sectors, such as:

Consumer electronics: Smartphones or TVs with embedded cameras can be used to interact with media applications—control the playback of songs or movies. We can use gestures to play or pause content, mute or unmute the sound, increase or decrease volume, or skip to the next song.

Automotive: The automotive industry will use hand gestures for infotainment systems to control in-car media players and phone systems. But gestures can also be used to control lights and for GPS navigation.

Healthcare: Gesture recognition can help doctors keep surgical wards sterile. By controlling the camera without touching the screen or reviewing medical documentation, the surgeon can reduce the risk of infection.

Thinking of the future scope of this system, we can see that a user can use his system without any use of physical devices like a keyboard or mouse so, with the help of a good camera, the user can operate his system from a distance after getting its range to a few meters.

REFERENCES

[1] Kollipara Sai Varun, I. Puneeth, T. Prem Jacob, “Hand Gesture Recognition and Implementation for Disables using CNN’S”, 2019 International Conference on Communication and Signal Processing (ICCSP), 10.1109/ICCSP.2019.8697980,25 April 2019

[2] Siddharth S. Rautaray, Anupam Agrawal, "Interaction with the virtual game through hand gesture recognition," 2011 International Conference on Multimedia, Signal Processing, and Communication Technologies, 10.1109/MSPCT.2011.6150485, 13 February 2012

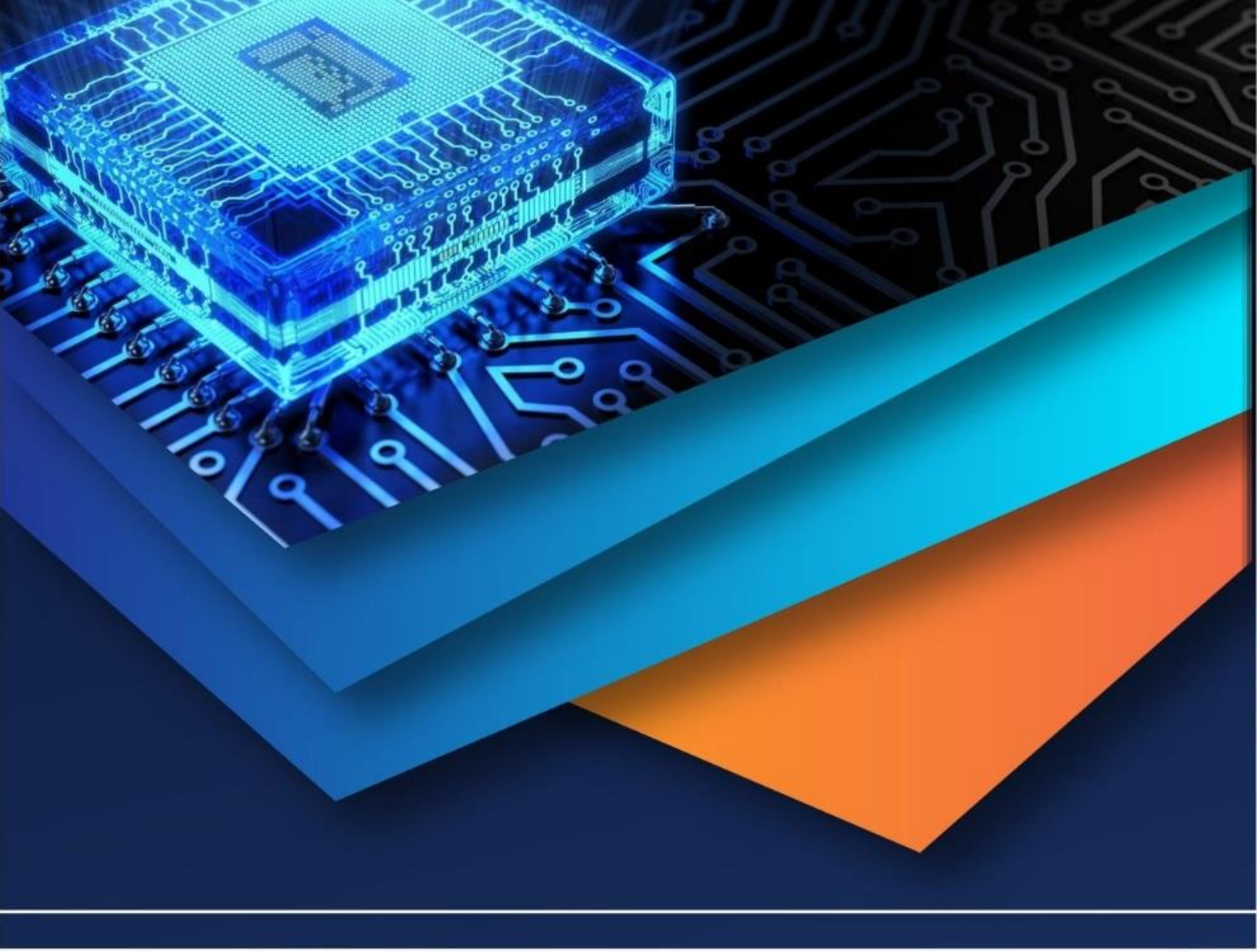
[3] Gaurav Raj, et. al. (2016), "Blur image detection using Laplacian operator and Open-CV", International Conference on System Modeling & Advancement in Research Trends, pp. 63-67.

[4] Emile A. Hendriks, (2008),’’ Sign Language Recognition by Combining Statistical DTW and Independent Classification’’, IEEE Transactions on pattern analysis and machine intelligence. Issue: 11. pp 2040- 2046.

[5] Pei Xu, “A Real-time Hand Gesture Recognition and"Human-Computer Interaction System”, April 2017.

[6] Rafiqul Zaman Khan, Noor Ibraheem, “Hand Gesture Recognition: A Literature Review”, International Journal of Artificial Intelligence & Applications,10.5121/ijaia.2012.3412 , August 2012.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)