



Gesture Based Desktop Automation by Adopting Artificial Intelligence and Computer Vision Techniques

Kishor Kumar Reddy C¹ and Anisha P R²

^{1,2}Stanley College of Engineering & Technology for Women, Hyderabad, India

Abstract:

Latest advancements in technology paved its way to handle complicated tasks in a much simpler manner. Good communication skills are essential to allow others and yourself to understand information more accurately and quickly. For a convincing communication to happen, physical gestures can be measured as a powerful mean. Considering the gesture-based communication as a criterion, we can even facilitate our work while using computers. This sort of improvement can be done by Automation. It plays a vital role in order to make this happen. The main priority of this system is to be simple, easy and user friendly without making any special hardware. Gesture based desktop automation system provides us with an innovative, natural, user friendly way of interaction with the computer which is more familiar to the human beings. In this, first, the model is pre trained on the images of different hand gestures, like, showing numbers with fingers – 0,1,2,3,4,5. This model uses the integrated webcam to capture the video frame. The image of the gesture captured in the video frame is compared with the pre-trained model and the gesture is identified. Desktop automation functionalities like opening browser, file explorer, calculator, notepad etc. will be executed if the gesture is recognized. This system will work as one of the futuristic of Artificial Intelligence and Computer Vision with user interface.

Keywords: Automation, Imagery, Webcam, Computer vision, Artificial Intelligence

1. Introduction

Modern developments in computer software and related technologies have provided a valuable service to the users. In everyday life, physical gestures are a powerful means of communication. Use of the full potential of physical gesture is also something that most human computer dialogues lack [12]. Automation is the key to such development. The modern IT environment needs to be dynamic, and should be able to scale quickly—automation is necessary. As computers have become effectively a part of our lives, people began

to think deliberately if there was a more natural way to interact with these devices, other than making use of the traditional mouse and keyboard. This has paved its way to the advent of a new arena: Human Computer Interaction (HCI).

The prominence of HCI is rising day by day as the industry requires a lot more of Virtual and Augmented Reality as it excites people to make use of an advance level of interaction that suits the term ‘virtual reality’, which mimics reality, where we use our voice, senses and gestures to communicate with the world. When we



think about the term 'basic interaction', the first thought is to touch things by moving our hands. This further leads to highlight the usage of hand gesture recognition as it is the base of HCI [11].

In recent years, human computer interface research has been limited to approaches focused on the use of a keyboard, joystick, mouse, graphics panel and so on. This development has rapidly progressed with the advent of voice, live displays, sound, vision, and other techniques that offer a much higher quality and seamless interactions between a human and machine [14].

Humans are able to recognize body and sign language easily. This is possible due to the combination of vision and synaptic interactions that were formed along brain development. In order to replicate this skill in computers, some problems need to be solved: how to separate objects of interest in images and which image capture technology and classification technique are more appropriate, among others. This paper intends to discover and identify the prevailing options for hand gesture recognition in a common framework. Every individual today, possesses a laptop with a front-facing camera. If we concentrate on this, then we can probably build an easier way to communicate with the masses [13].

The chore of hand gesture recognition is one of the vital and key problem in Computer Vision. The primary step in any hand processing system is to spot the hand and to localize it in the image. The hand detection is a challenging task due to

the variations in the posture, angle, site and scale. And also, diverse lighting conditions increase additional inconsistency.

This builds a real time gesture classification system that can automatically detect gestures in natural lighting condition. In order to accomplish this objective, a real time gesture-based system is developed to identify gestures. Gesture Based Desktop Automation is to enhance desktop automation based on gesture recognition. This system will work as one of futuristic of Artificial Intelligence and computer vision with user interface. The main priority of this system is to be simple, easy and user friendly without making any special hardware. All computation will occur on single PC or workstation. Only hardware used is Digital Camera to detect the gestures and automate the functionalities of desktop to simplify our everyday tasks.

Recently, the range of gestures available for communicating with handheld devices and consumer electronic devices has expanded expeditiously. This creates a system that can control the desktop applications using Gesture-based detection which enables us to access the desktop functionalities effortlessly and proficiently [15].

The Gesture recognition with various hand motions is intensifying at incredible speed for the ultramodern products and services. Many hi-tech companies are emerging technology on the basis of hand gesture recognition system which comprises of corporations like Samsung, Microsoft, Sony and including the devices like Laptop,



Professional and LED lights, Hand held devices. etc. The verticals are including Gesture technology in several arenas like Artificial Intelligence, Entertainment, Education, Automation and Medical fields.

The future aim in the field of Gesture Recognition is to make it cost effective and economical which is possible when proper amount of investment is applied on research and development. Converting data into features with the fusion of technology and human intelligence is a commendable feature which has paved a way to many innovations. This gesture recognition technology has been very effective in the advancement of Smart phones. They have come across massive expanse by enhancing the viewing and reading features with touchless gestures using this gesture recognition system. Google Glass is also an application under similar cadre.

Nowadays, this technology has also been entrenched into smart televisions and alike applications, which can be effortlessly controlled and managed by Hand gestures and Voice options. Hand Gesture may also be practised in terms of Medical Assistance and Robotic Nurse in the medical department. With the rapid growth in technology, revolutionary changes regarding the future makes it quite unpredictable but the point to be noted is that, the forthcoming of gesture recognition will be evergreen and is going to stay for a long time with more beneficial, exciting and life touching experiences [16].

Gesture control technology shows great potential in education. The future work

of the use of gesture control technology in higher education is to increase experimentation in the teaching environment. For the best results and user satisfaction, it is also expected to perform more calibrations on the product and improve the accuracy of the software. Another direction of future work should include testing gesture control technology, as well as augmented reality (AR) and virtual reality (VR). This will be the next step in the virtual laboratory, which can be used in many scientific disciplines and engineering laboratories.

2. Relevant Work

Udit Kumar et al. proposed a paper: “Hand Gesture Controlled Laptop Using Arduino”: In this paper, the operations of the computer or laptop’s desktop functionalities are controlled through various hand gestures of human. The design mainly involved with HMI (Human Machine Interface) system that is able to control the system applications such as volume offsetting, scroll vertical and horizontal, tab shifting etc., without using any mouse, keyboard, or joystick. The system control with simple unique gestures of hands reduces the space between user and machine. In the present paper basic Arduino Uno is used to support the hand gesture-based system control [1].

Dr. Jabbar Raheem Rashed et al. proposed a paper: “New method for hand gesture recognition using wavelet neural network”. The paper presents a technique where an image of hand gesture will be passed through four stages, preprocessing,



normalization, feature extraction, and classification. The wavelet neural network is used to build information models because it can handle very complex interactions. The practical system is simulated using MATLAB. The performance of suggested method is excellent compared with traditional approaches and the final results of selected patterns recognition are powerful [2].

Smit Desai et al. proposed a paper: “Human Computer Interaction Through Hand Gestures for Home Automation Using Microsoft Kinect”: In this paper we have used Microsoft Kinect for image capturing along with some important computer vision (CV) and digital image processing techniques (DIP) for hand gesture recognition. Arduino Uno microcontroller and relay circuits are used for controlling electrical devices. The algorithm presented gives here is an assistive application useful for physically challenged and senior citizens with an accuracy of 88% [3].

Sushmita Nooka et al. proposed a paper: “Gesture Controlled Home Automation Using CNN”: This paper method discusses the latest approach. The previous methods have seen the use of accelerometers, which are fixed on the hand as they are considered to be the best tool for carrying out such practices. These models are accurate but not flexible and portable. Our system is based on remote control home automation where-in our remote is an android application. In our system, the smartphone camera is used to give gestures as input to the gesture recognition model to

filter and predict the gesture. Web-based automation communicates directly with raspberry pi to control home appliances [4].

Udit Kumar et al. proposed a paper: “Access lens: a gesture-based screen reader for real-world documents”: This paper describes the development of Access Lens hardware and software, the iterative design of Access Lens in collaboration with blind computer users, and opportunities for future development. Access Lens is a new interaction method that uses computer vision-based gesture tracking to enable blind people to use accessible gestures on paper documents and other physical objects, such as product packages, device screens, and home appliances by converting text to speech [5].

Tomi Heimonen et al. proposed a paper: “Designing Gesture-Based Control for Factory Automation”: This reports the development and evaluation of a gesture-based interaction prototype for controlling the loading station of a factory automation system. In this context, gesture-based interaction has the potential to free users from the tedious physical controls but it must also account for safety considerations and users’ perceptions. An optimal interaction is challenging but practical in order to make the interaction an emotionally pleasurable and novel experience [6].

Haitham Badi et al. proposed a paper: “Recent methods in vision-based hand gesture recognition”. The proposed system presents a recognition algorithm to recognize a set of six specific static hand gestures, namely: Open, Close, Cut, Paste,



Maximize, and Minimize. The hand gesture image is passed through three stages, namely, pre-processing, feature extraction, and classification. The back-propagation learning algorithm is employed in the multi-layer neural network classifier. The results show that hand contour method has a performance of (71.30 %) recognition, while complex moments have a better performance of (86.90 %) recognition rate [7].

Dinh-Son-Tran et al. proposed a paper: “Real-Time Hand Gesture Spotting and Recognition Using RGB-D Camera and 3D Convolutional Neural Network”. In this paper, we propose a novel method for fingertip detection and hand gesture recognition in real-time using an RGB-D camera and a 3D convolution neural network (3DCNN). This system can accurately and robustly extract fingertip locations and recognize gestures in real-time. The experimental results showed that our system has a high level of accuracy of hand gesture recognition and is considered to be a good approach to a gesture-based interface [8].

Priyanka D. Hatwar et al. proposed a paper: “Home Automation System Based on Gesture Recognition System”: This paper presents an effective method to overcome the problems faced by old age or disable people using speech recognition. This system is both software and hardware designed using MATLAB R2009a. This system is divided into three main parts namely voice train process, voice recognition process and integration of hardware with MATLAB.

This system used speaker dependent method. This proposed design is novel in the way that it is controlling loads by speech recognition using MATLAB to turn on/off loads via parallel port of a computer [9].

Yanan Xu et al. proposed a paper: “Review of Hand Gesture Recognition Study”. This paper reviewed the current study status and application of gesture recognition aiming to summarize the commonly used hand gesture recognition methods, analysis their strength and weak points, and list the challenging problems in current research of hand gesture recognition system. It also tells that Studies on hand gesture recognition is very important for the development of new human-centered human-computer interaction [10].

3. Proposed Architecture

In this, first, the model is pre trained with the images of different hand gestures, such as, showing numbers with fingers as 0,1,2,3,4,5. This model uses the integrated webcam to capture the video frame. The image of the gesture captured in the video frame is compared with the pre-trained model and the gesture is identified. If the gesture predicted is 0, then it opens Notepad; if 1 is detected, then it opens File explorer; if 2 is detected, then it opens Calculator. Likewise, if 3,4,5 is detected then it opens Google, YouTube, Facebook respectively. The figure below describes the architecture.

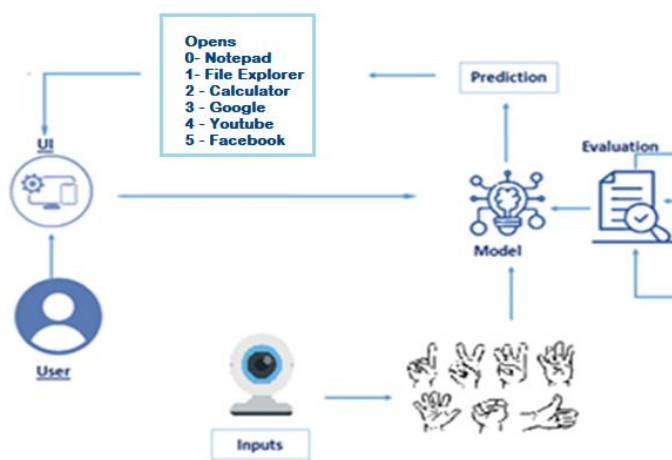


Fig. 1: Proposed
Architecture

4. Discussion & Results

After completion of model building, flask app is built and executed. Then the following results are showcased.

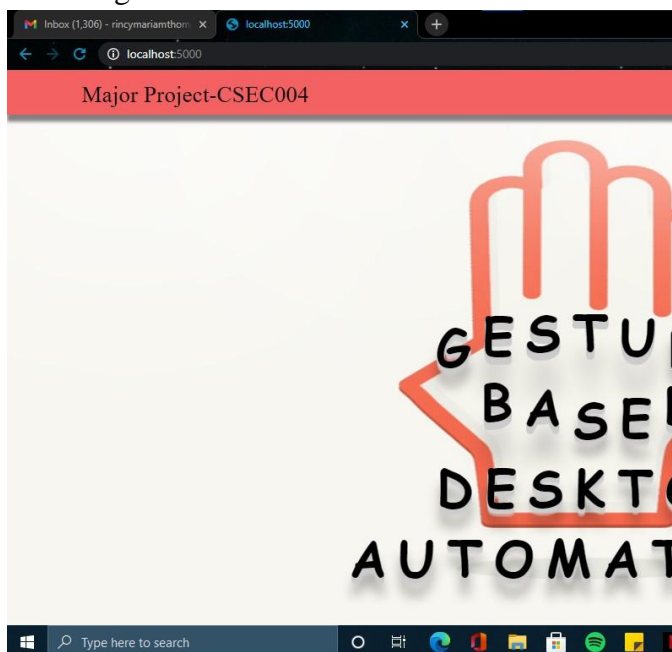
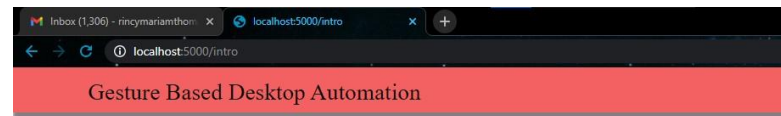


Fig 2: Home page

When the application runs, the localhost displays the above web page i.e., home page.



Gesture based desktop automation system provides us with an in way of interaction with the computer which is more familiar to th first, the model is pre trained on the images of different hand gest fingers - 1,2,3,4. This model uses the integrated webcam to capt of the gesture captured in the video frame is compared with the pr identified. Desktop automation functionalities like open file explorer, youtube etc. will be executed if the gesture is recogni of the futuristic of Artificial Intelligence and Computer V



Fig 3: Introduction

When "Introduction" button is clicked, the above webpage is displayed where the introduction is defined.

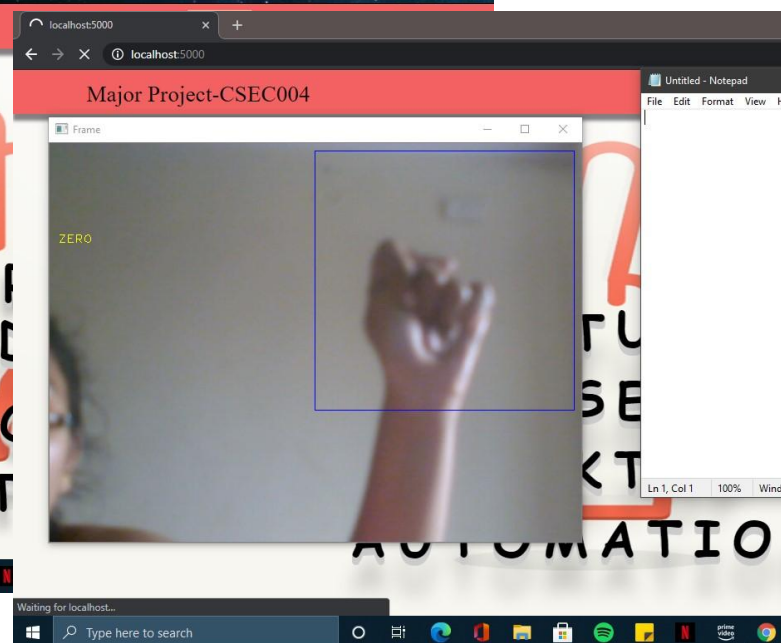


Fig 4: Gesture-0

The "Launch" button launches the camera in order to detect the hand gesture. When the

gesture is identified, the respective action for the gesture shown is performed. In the above figure, it is noticed that “Notepad” is opened when 0 is detected.

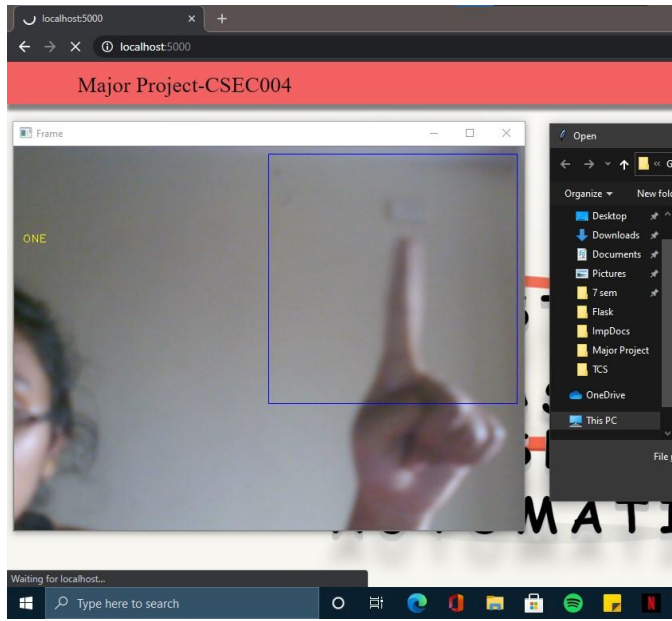
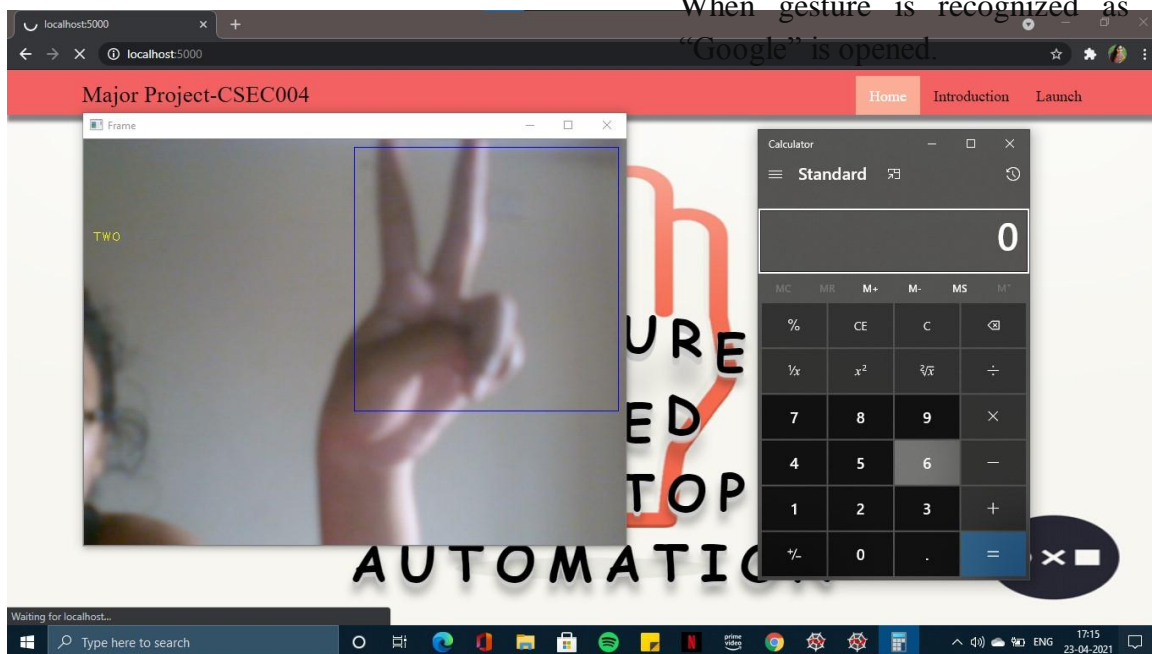


Fig 5: Gesture – 1

Likewise, “File Explorer” is opened when 1 is detected.

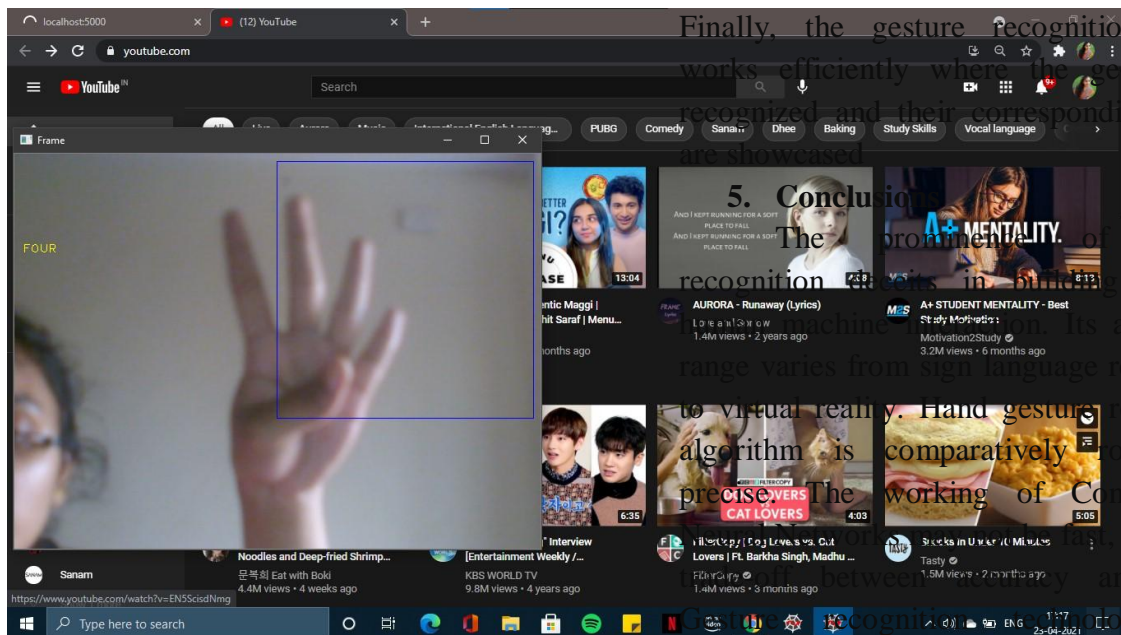


**Fig 6:
Gesture – 2**

“Calculator” is opened when gesture is recognized as 2.

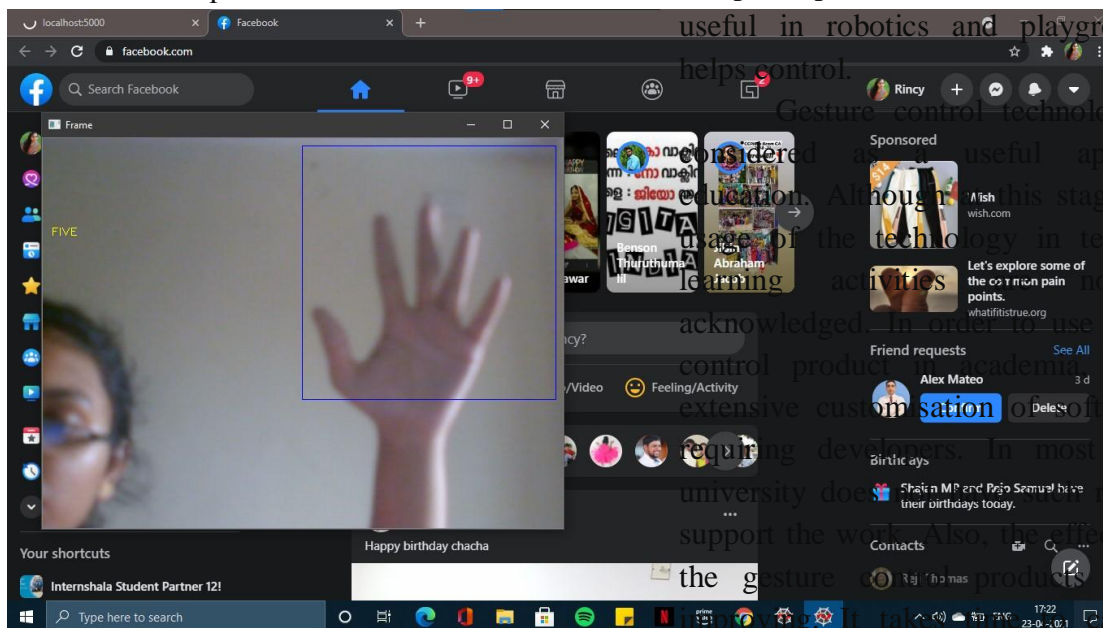
**Fig
7:Gesture – 3**

When gesture is recognized as 3, then “Google” is opened.



**Fig
8:Gesture – 4**

When gesture is recognized as 4, then “YouTube” is opened.



**Fig
9:Gesture – 5**

“Facebook” is opened when gesture detected is 5.

Finally, the gesture recognition system works efficiently where the gestures are recognized and their corresponding results are showcased.

The prominent feature of gesture recognition is in building effective machine learning. Its application range varies from sign language recognition to virtual reality. Hand gesture recognition algorithm is comparatively robust and precise. The working of Convolutional Neural Network (CNN) is also there is between accuracy and speed.

Gesture recognition technology uses mathematical algorithms to interpret human gestures and actions. The camera detects the gesture, which is then interpreted by the computer process. This technology is very useful in robotics and playgrounds, and helps control.

Gesture control technology can be considered as a useful approach in education. Although at this stage, practical usage of the technology in teaching and learning activities is not widely acknowledged. In order to use the gesture control product in academia, it requires extensive customization of software, often requiring developers. In most cases, the university does not have resources to support the work. Also, the effectiveness of the gesture control product still needs to be calibrated.

In order to control the product, users will need to spend time training and practicing. Gesture control products are not



yet as intuitive as they claim to be or have the potential to be.

The performance of the proposed method is highly dependent on the results of manual detection. If there is a moving object with a skin colour, the object exists in the hand detection result, thereby reducing the performance of gesture recognition. However, machine learning algorithms can distinguish hands from the bottom. Time-of-flight (ToF) cameras provide depth information, which can improve hand detection performance. Therefore, in future work, machine learning methods and ToF cameras can be used to solve complex potential problems and improve the robustness of manual detection.

By using hand gestures, the user can operate the virtual environment. The user's hand gesture is mimicked by the virtual hand. During the initialisation segment, the user provides the sample gestures to the model. In the recognition segment, the system matches up the input gestures against the sample gestures. This requires no special hardware, apart from the camera. The recognition algorithm built is relatively fast, simple and easy to use, which can be run on a workstation in real-time. It is robust against varying lighting environments. It also upholds the accuracy, even in a situation where different hands are used. Hence, gesture based desktop automation promises extensive range of applications in almost every field

References

1. Udit Kumar, Sanjana Kintali, Kolla Sai Latha, Asraf Ali, N. Suresh

- Kumar. (2020, April). Hand Gesture Controlled Laptop Using Arduino. Test Engineering & Management, 83, 20083-20086.
2. Dr. Jabbar Raheem Rashed, Mr. Hasanain Abbas Hasan. (2017, January). New method for hand gesture recognition using wavelet neural network. Journal of Engineering and Sustainable Development, 21(01), 65-73.
3. Smit Desai, Apurva Desai. (2017, April 8). Human Computer Interaction Through Hand Gestures for Home Automation Using Microsoft Kinect. International Conference on Communication and Networks, 508, 19-29.
4. Sushmita Nooka, Ninad Kheratkar, Aboli Pathak, Shreyash Kumbhar, Ashwini Jarali. (2020, March). Gesture Controlled Home Automation Using CNN. International Research Journal of Engineering and Technology, 07(03), 5391-5395.
5. Shaun K. Kane, Brian Frey, Jacob O. Wobbrock. (2013, April). Access lens: a gesture-based screen reader for real-world documents. SIGCHI Conference on Human Factors in Computing Systems, 347-350.
6. Omi Heimonen, Jaakko Hakulinen, Markku Turunen, Jussi P. P. Jokinen, Tuuli Keskinen, Roope Raisamo. (2014, September). Designing Gesture-Based Control for Factory



- Automation. IFIP Conference on Human-Computer Interaction, 1-8.
7. Haitham Badi. (2016, April 20). Recent methods in vision-based hand gesture recognition. International Journal of Data Science and Analytics, 1, 77-87.
 8. Dinh-Son Tran, Ngoc-Huynh Ho, Hyung-Jeong Yang, Eu-Tteum Baek, Soo-Hyung Kim, Gueesang Lee. (2020, January 20). Real-Time Hand Gesture Spotting and Recognition Using RGB-D Camera and 3D Convolutional Neural Network. Applied Sciences, 10(722), 1-15.
 9. Priyanka D Hatwar, Neha A. Wahile, Isha M. Padiya. (2017, March). Home Automation System Based on Gesture Recognition System. International Journal of Emerging Technologies in Engineering Research, 5(3), 48-53.
 10. Yanan Xu, Yunhai Dai, (2017). Review of Hand Gesture Recognition Study and Application. Contemporary Engineering Sciences, 10(8), 375-384.
 11. Anisha P R, Kishor Kumar Reddy C and Nguyen Gia Nhu, "Blockchain Technology: A Boon at the Pandemic Times – A Solution for Global Economy Upliftment with AI and IoT", EAI/Springer Innovations in Communication and Computing, 2022.
 12. Kishor Kumar Reddy C, Anisha P R, Shastry R, Ramana Murthy B V, "Comparative Study on Internet of Things: Enablers and Constraints", Advances in Intelligent Systems and Computing, 2021
 13. Kishor Kumar Reddy C, Anisha P R, Apoorva K, "Early Prediction of Pneumonia using Convolutional Neural Network and X-Ray Images", Smart Innovation, Systems and Technologies, 2021
 14. Kishor Kumar Reddy C and Vijaya Babu B, "ISPM: Improved Snow Prediction Model to Nowcast the Presence of Snow/No-Snow", International Review on Computers and Software, 2015
 15. Kishor Kumar Reddy C, Rupa C H and Vijaya Babu B, "SLGAS: Supervised Learning using Gain Ratio as Attribute Selection Measure to Nowcast Snow/No-Snow", International Review on Computers and Software, 2015
 16. Kishor Kumar Reddy C, Rupa C H and Vijaya Babu B, "A Pragmatic Methodology to Predict the Presence of Snow/No-Snow using Supervised Learning Methodologies", International Journal of Applied Engineering Research, 2014.