

Voice Guided, Gesture Controlled Virtual Mouse

Rajat Dudhapachare
Department of Information Technology
G H Raisoni College of Engineering
Nagpur, India
dudhapachare_rajat.it@ghrce.raisoni.net

Mayur Awatade
Department of Information Technology
G H Raisoni College of Engineering
Nagpur, India
awatade_mayur.it@ghrce.raisoni.net

Pushpak Kakde
Department of Information Technology
G H Raisoni College of Engineering
Nagpur, India
kakde_pushpak.it@ghrce.raisoni.net

Nihal Vaidya
Department of Information Technology
G H Raisoni College of Engineering
Nagpur, India
vaidya_nihal.it@ghrce.raisoni.net

Mayur Kapgate
Department of Information Technology
G H Raisoni College of Engineering
Nagpur, India
kapgate_mayur.it@ghrce.raisoni.net

Rajesh Nakhate
Department of Information Technology
G H Raisoni College of Engineering
Nagpur, India
rajesh.nakhate@raisoni.net

Abstract— The mouse is a prime example of HCI advancement. Although the modern wireless mouse or Bluetooth mouse still requires peripherals like energy cells and cards to connect to the computer, it is not entirely peripheral-free. The suggested AI virtual mouse system addresses the aforementioned issues by capturing hand motions with an external or digital camera, then improving the system's accuracy using voice assistants and hand point detection through object recognition. The system's foundation consists of machine learning techniques. To operate a computer, navigate, and execute actions like left- and right-clicking, digital hand motion can be used instead of a physical mouse. The program employs machine learning for hand identification and Python modules, for voice assistance. The suggested approach eliminates the need for human involvement and computer-controlled equipment in the battle against the spread of COVID-19 by implementing fundamental mouse operations together with brightness, volume control, and the ability to manage noise variations.

Keywords—*Gesture, HCI, Voice assistant, Python, Mediapipe, OpenCV, Image Processing, Virtual mouse*

I. INTRODUCTION

Throughout the past several years, there has been an increase in interest in the development of human-computer interaction systems. As software packages that give capabilities for certain processing and facilitate connection between pcs and computing devices are widely accessible, students, engineers, and scientists are urged to concentrate on creating natural, intuitive user interfaces. This research proposes an AI based totally digital mouse device that uses voice assistant and hand gestures to permit mouse-like actions on a computer. The recommended system's foremost reason is to enable the person to perform mouse pointer capabilities and scroll operations by using the use of digital camera of pc i.e. built in camera. The AI virtual mouse system detects the movement of a user's fingertip to conduct mouse cursor actions like clicking, scrolling, pointing, and perform brightness and volume control, either via a web camera, lens or a in-built camera [7].

Voice assistant can respond to user inquiries with pertinent information by listening for particular keywords and removing background noise. A voice assistant is a sort of helper that employs digital technology and human voice synthesis and algorithms to pay attention to specific spoken commands and deliver pertinent information or carry out certain tasks.

The movement of a visual object is one mouse function that has been replaced by hand movements. The system is designed to be inexpensive, and it takes use of inexpensive input devices like a webcam to record hand movements. To manipulate materials, preset command-based movements are modelled.

II. PROBLEM STATEMENT

The suggested virtual mouse idea can be used to clear up troubles that occur within the real world, inclusive of the ones where in there is not enough room to use a bodily mouse and those that affect individuals or people who aren't able to cope with a physical mouse because of hand problems. The COVID-19 situation also makes it dangerous to use gadgets by touching them because doing so ought to result in a situation wherein the virus is spread by way of touching the gadgets. The suggested AI digital mouse can be used to clear up those troubles due to the fact movement and hand points spotting is utilized to manipulate the numerous computer mouse features by using the usage of a web-cam or an innate camera.

When utilizing a Bluetooth mouse or a remote, a few specific equipment, such as the mouse, the computer connection device, and aside from the battery to power the mouse to operate a utilized, the client manages the PC mouse activity by using his or her natural camera or visual camera throughout the process.

III. RELATED WORK

Although some relevant studies have been done on virtual mouse employing hand gesture identification while trying gloves and using colored points in the fingers for movement identification, but they are not perfect in mouse operations. Because of gloves, some users may not be able to use the gloves properly, and in other situations, colour tip detection may not be successful, the identification is not as accurate. A few attempts are made to recognize the hand gesture input using a webcamera.

- Human-computer interactions rely heavily on Hand Gesture Recognition, as demonstrated by Varun et al. The authors demonstrated the numerous technological advancements that are currently taking place, such as biometric verification, which is often seen in phones or smart phones. Hand gesture identification is another contemporary method of HCI, allowing us to operate our systems by waving our hands in front of a webcam. This study is offered

based on this concept. The ideas and approaches for colour identification and the virtual mouse are completely clarified in this study [7].

- An mutable skin colour system and a moving history photo palm movement control identification technique were employed by ChenChiung Hsieh et al. The paper's main shortcoming is its inability to recognize hand movements with more complexity [3].
- Reddy et al. suggested a virtual mouse interface that recognized hand gestures and identified finger tips. The two finger tracking techniques used in this study are hand gesture detection and employing colored caps. There are three essential phases in this process: monitoring hand gestures, finger detection utilizing color identification, and cursor implementation. In this work, a convex hull is formed around the contour that is detected to produce hand gesture tracking. Using the area ratio of the produced hull and contour, hand features are retrieved. To evaluate this method in real-world circumstances, thorough experiments are run [2].
- "Cursor Control via Hand Gestures" was a proposal made by Vinay Pasi, Saurabh Singh, and Pooja Kumari for the IJCA Journal in 2016 . The approach recommended using several bands to carry out varying mouse activities. The con is that multiple colours are needed to perform mouse operations [5].
- "Virtual Mouse Using Hand Gesture" was an idea put out by Chaithanya C, Lisho Thomas, Naveen Wilson, and Abhilash SS in 2018, where model identification is made through colors. But very few mouse operations are carried out[6].
- Hand gesture recognition method provides a straightforward and natural method to communicate in the modern world, according to research by Boruah et al. A significant component of modern e-learning is the creation of training methods that make use of technology-based practical things to boost association and collaboration among the tutor and the learner. In this research, we suggest an collaborative learning assistance device based on a hand motion identification system that uses vision. The structure recognizes hand gestures using MediaPipe [1].
- The research on "Cursor Control System Using Hand Gesture Recognition". was made but the restriction in this study is that saved frames must be processed in order to partition hands and identify skin pixels[4].

IV. METHODOLOGY

The proposed artificial intelligence virtual mouse comprises of data from camera pictures on a computer or laptop. As you can see, the Python computer vision software OpenCV is used to create a video capture object, which causes the web camera to start recording. A camera captures the frames, which are then transferred to a computerized artificial intelligence.

A. Taking the input

Handles camera using digicam, an object obtained from opencv that catches the video stream and assigns the height and width in pixels for each collected frame; also, utilized flip image and BGR to RGB picture conversion.

B. Hand Detection

Establishing major or minor hands utilizing mediapipe framework points, achieving a maximum of two hands, a minimum detection confidence of 50%, a minimum tracking confidence of 50%, and obtaining finger status.

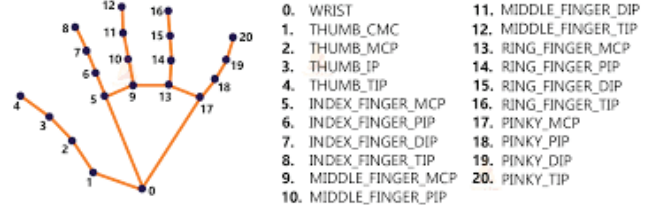


Fig. 1. Hand Landmark points

C. Hand Gesture

First of all, the condition of your fingers today If the finger is extended, 1; otherwise, 0. Measure the distance between the fingertip, middle knuckle, and base knuckle to establish the "finger" position.

From index to pinky for the thumb, there are zero points for each finger [[8,5,0], [12,9,0], [16,13,0], and [20,17,0]]. calculating the Euclidean distance between the "point" to determine the ratio and finger state.

$$D = \sqrt{[(X2 - X1)^2 + (Y2 - Y1)^2]}$$

D. To control position of pointer

We must first determine the center of every red item that is identified, and we will do this by finding the middle value of the highest and the lowest points of the neighboring tips. Currently we having two co-ordinates from the intersection of the two items, which allows us to see their commonality and obtain the red object in the image. The detected point is frequently switched from the particular screen resolution to the camera resolution. Currently, the position is often set based on the mouse pointer. To control mouse function with Gestures

E. To control function with gestures

To navigate the computer display with the mouse. If the fore finger is raise up with point Id 1 or if both the fore finger and middle finger are up, the Python AutoPy package may be used to have the mouse pointer move around the PC display.

Click the mouse to select the appropriate button. So when middle finger with point Id = 2 and the index finger with point Id = 1 are both up and their spacing is smaller than 40 pixels, the PC is programmed to press the right bouton of mouse.

For Nothing to Happen on the Screen. The pc is programmed to stop performing any mouse activities over the display if all the fingers are raised with point Id values of 0, 1, 2, 3, and 4.

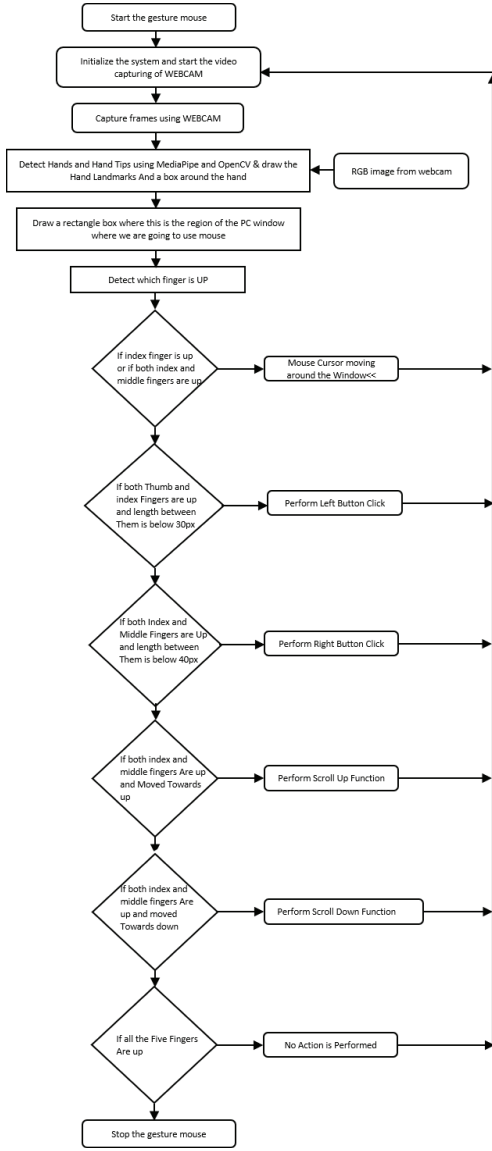


Fig. 2. Workflow for Hand Gesture

F. Methodology used by voice assistant

Headings, we use sapi5 and pyttsx3 right away to give our program the capacity to communicate with the system voice. Text-to-speech software is implemented in Python via the pyttsx3 module.

In contrast to many other libraries, it works with Python 2 and 3 and may even be used offline. Thanks to Microsoft's Speech Application Programming Interface (SAPI), Windows program may make advantage of voice recognition and synthesis. The main function then defines the program's capabilities. The following is what the recommended system is anticipated to be able to perform.

- The assistant continually asks the user for feedback and watches for new commands. Depending on the user's requirements, the receiving or listening time can be changed.
- If the helper doesn't comprehend the instruction the first time, it will keep inquiring the user to pardon it.
- This assistant may be set to talk in either a male or female voice, depending on the user's preferences.

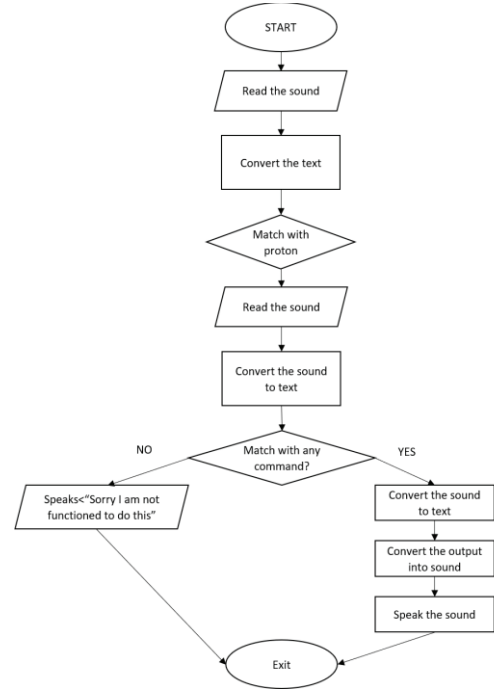


Fig. 3. Workflow for Voice Assistant

V. IMPLEMENTATION

A. Language

Python3 is a general-purpose, high-level programming language. It is designed to prioritize indentation and code readability.

Python has dynamic typing and garbage series. It helps some of computing paradigms, which includes structured programming, object-orientated programming, procedural programming, and practical programming (specially this). additionally, it is understand as a "batteries blanketed" language because of its massive popular library.

Guido van Rossum developed Python language in the late 1980s to take place of the ABC programming language. Python 0.9.0 was released for use for the first time in 1991[36]. In 2000, Python 2.0 became a reality. Python 3.0 was a big upgrade that was only partly backwards compatible with earlier versions when it was released in 2008.

In projects that contain artificial intelligence and machine learning, Python is broadly utilised with the aid of libraries like TensorFlow, Keras, Pytorch, and scikit-learn. when you consider that it's miles a programming language with a modular device, easy syntax, and effective text processing facilities, Python is commonly used for processing natural languages.

B. Library

A collection of interconnected modules makes up the Python Library. provides content packages that can be used many times across different applications. It simplify and makes programming useful for the editor.

Packages like OpenCV, mediapipe, pyttsx3, and speech recognition have been utilised by us.

cv2: OpenCV is an excellent tool for processing pictures and may be used to handle tasks like tracking an item and detecting faces in digital photos.

Mediapipe: A pass-platform library known as Mediapipe developed by using Google provides outstanding in-constructed machine learning (ML) approaches for computers imaginative and prescient troubles. Python users might also utilise the OpenCV laptop imaginative and prescient library, which is extensively used for photograph enhancement, evaluation, popularity, and identification.

pyttsx3: Pyttsx3 is a Python text-to-speech translation module. Unlike rival libraries, it works offline and is suitable with Python 2 and 3. A programme invokes the factory function pyttsx3.init to obtain a reference to a Pyttsx3 ().

speech recognition: Voice identification refers to a machine's capacity to hear spoken terms and identify them. The spoken words may also then transformed into text, a query may be posed, or a reaction can be given by using using Python's speech recognition package. Even a few machinery may be set up to reply to those spoken words.

C. Flowchart

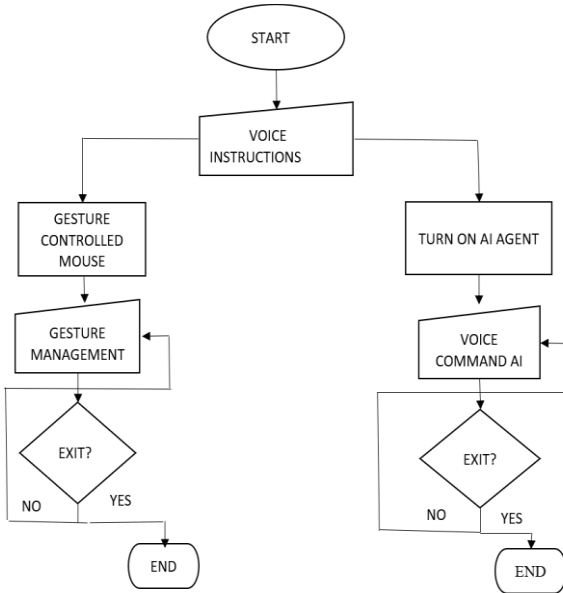


Fig. 4. Flowchart for overall system

VI. RESULT

The concept of using computer vision to enhance HCI is presented in the suggested AI virtual mouse. Due to the scarcity of usable datasets, comparing evaluation of the virtual mouse is challenging. Webcam has been positioned at various distances from the user to follow hand motions and identify hand tips in an effort to assess their abilities under varied illumination conditions. The AI virtual mouse system is tested numerous times under various lighting conditions, including strong light, low light, near proximity to the webcam, and four feet or more away from the screen.

The AI virtual mouse technology has a 99 percent accuracy rate. Based on its accuracy of 99 percent, we may draw the understanding that the suggested Virtual mouse system has operated satisfactorily. Because the right click is the most difficult motion for computers to understand, its

accuracy is poor. The right click's accuracy is low since the motion needed to perform the precise mouse operation is more challenging.

The other gestures are also quite accurate. Our virtual mouse model performed remarkably well, with accuracy of 99 percent, when compared to earlier methods. When in comparison with existing virtual mouse concepts, it is clear that the proposed AI virtual mouse has outperformed all others in terms of usability and accuracy.

The suggested model's unique capability to simulate real-world mouse actions on a computer, including left- and right-click, up-down scrolling, and mouse pointer movement, utilising fingertip sensing, is its key innovation.

A. Gesture Controller

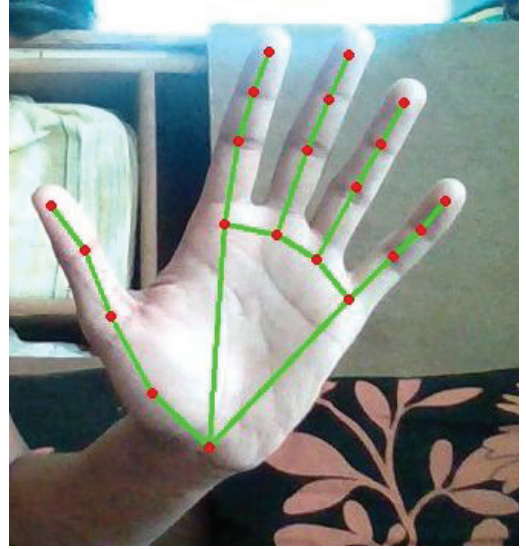


Fig. 5. Hand gesture recognition

B. Voice Assistant

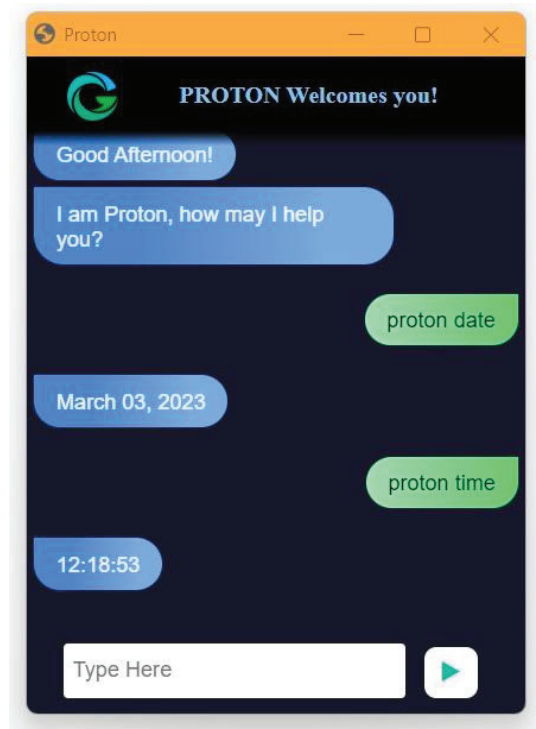


Fig. 6. Voice assistant

VII. APPLICATION

The virtual mouse system may be used in a number of situations, including those where it is impossible or difficult to use a genuine mouse. The technology improves human-computer interaction while reducing the need for devices.

1. The advised AI virtual mouse may be used to perform the laptop mouse operations with out the usage of a bodily mouse during the COVID-19 situation because it isn't always safe to use the devices by means of contacting them as it can bring about a state of affairs where the virus is propagated by means of touching the gadgets.
2. It may be used to operate virtual augmented and reality games instead of wired or wireless mouse.
3. Without the need of gadgets, the system advised can be utilized to operate robots and automated systems.
4. These controls can be used by those with hand issues to use the computer's mouse.
5. The suggested method may be utilized for virtual prototyping in designing and architecture.

VIII. CONCLUSION

The primary aim of the virtual mouse concept is to do away with the requirement for a hardware mouse in favour of using hand movements to move mouse cursor across screen. By processing the frames to execute the required mouse operations based on the detected fingers and hand tip, the suggested system may be realised via a webcam or an embedded camera.

Those with impairments will find the voice-enabled personal assistant more practical and time-saving in today's busy lives. When it comes to performing some of the tasks the user requests, this helper performs an excellent job. Moreover, this assistant can do a broad range of operations in response to a particular voice command, such as sending text messages to the user device, automating YouTube, and retrieving information from Google and Wikipedia. We have been able to automate multiple services with a unified command thanks to the voice assistant. This tool makes it easier for users to do the bulk of their tasks, such as internet searches. Our aim is for this programme to be completely capable of replacing the need for human server admins. As this system is made on Anaconda community, so any adjustments may be made immediately. The project's modular configuration allows additional modification and the addition of new features without affecting how the current system functions.

REFERENCES

- [1] B. J. Boruah, A. K. Talukdar and K. K. Sarma, "Development of a Learning-aid tool using Hand Gesture Based Human Computer Interaction System," 2021 Advanced Communication Technologies and Signal Processing (ACTS), 2021.
- [2] V. V. Reddy, T. Dhyanchand, G. V. Krishna and S. Maheshwaram, "Virtual Mouse Control Using Colored Finger Tips and Hand Gesture Recognition," 2020 IEEE- HYDCON, 2020, pp. 1-5, doi:10.1109/HYDCON48903.2020.9242677.
- [3] C.-Chiung Hsieh, D.-Hua Liou, & D. Lee, , "A real time hand gesture recognition system using motion history image," Proc. IEEE Int'l Conf. Signal Processing Systems (ICSPS), 2.10.1109/ICSPS.2010.5555462, 2010.
- [4] S. U. Dudhane, "Cursor control system using hand gesture recognition," *IJARCCCE*, vol. 2, no. 5, 2013.
- [5] K. P. Vinay, "Cursor control using hand gestures," *International Journal of Critical Accounting*, vol. 0975–8887, 2016.
- [6] L. Thomas, "Virtual mouse using hand gesture," *International Research Journal of Engineering and Technology (IRJET)*, vol. 5, no. 4, 2018
- [7] J. Katona, —A review of human-computer interaction and virtual reality research fields in cognitive Info Communications, || *Applied Sciences*, vol. 11, no. 6, p. 2646, 2021
- [8] Balamurugan. C1, Arumuga Kumar. M2, Arun N3 and Deepak proposed a study on "HCI SYSTEM WITH HAND GESTURE" in *International Research Journal of Engineering and Technology (IRJET)*.
- [9] D.-S. Tran, N.-H. Ho, H.-J. Yang, S.-H. Kim, and G. S. Lee, —Real-time virtual mouse system using RGB-D images and fingertip detection, || *Multimedia Tools and Applications Multimedia Tools and Applications*, vol. 80, no. 7, pp. 10473–10490, 2021.
- [10] K. H. Shibly, S. Kumar Dey, M. A. Islam, and S. Iftekhar Showrav, —Design and development of hand Gesture based virtual mouse, || in *Proceedings of the 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, pp. 1–5, Dhaka, Bangladesh, May 2019
- [11] P. W. Jaronde, A. Vyas and M. Gaikwad, "A Survey on Energy Aware Cognitive Radio Network," 2022 2nd International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC), Mathura, India, 2022, pp. 1-6, doi: 10.1109/PARC52418.2022.9726654.
- [12] D. Manikkule and P. Jaronde, "Encapsulation of Full Adder Using 180nm CNTFET," 2019 9th International Conference on Emerging Trends in Engineering and Technology - Signal and Information Processing (ICETET-SIP-19), Nagpur, India, 2019, pp. 1-6, doi: 10.1109/ICETET-SIP-1946815.2019.9092017.
- [13] D. Manikkule and P. Jaronde, "Design of Decoder Using Ternary Inverter," 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), Bombay, India, 2019, pp. 1-3, doi: 10.1109/I2CT45611.2019.9033760.
- [14] Wazalwar, S.S., Shrawankar, U. (2021). Distributed Education System for Deaf and Dumb Children and Educator: A Today's Need. In: Singh Mer, K.K., Semwal, V.B., Bijalwan, V., Crespo, R.G. (eds) *Proceedings of Integrated Intelligence Enable Networks and Computing. Algorithms for Intelligent Systems*. Springer, Singapore.
- [15] S Wazalwar, U Shrawankar , *Community Cloud Service Model for People with Special Needs, Cloud Computing Technologies for Smart Agriculture and Healthcare*, 47-55, Chapman and Hall/CRC, 2021
- [16] N. A. Karotiya, N. P. Wyawahare and S. L. Haridas, "Review paper on point to point communication with the use of power over Ethernet based on VOIP system on asterisk," 2016 2nd International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), Chennai, India, 2016, pp. 729-732, doi: 10.1109/AEEICB.2016.7538390.
- [17] A. P. Tembhurne and N. P. Wyawahare, "Alamouti based on spatial modulation for data transmission," *International Conference for Convergence for Technology-2014*, Pune, India, 2014, pp. 1-4, doi: 10.1109/I2CT.2014.7092179.
- [18] S. Patil, S. Vairagade and D. Theng, "Machine Learning Techniques for the Classification of Fake News," 2021 International Conference on Computational Intelligence and Computing Applications (ICCICA), Nagpur, India, 2021, pp. 1-5, doi: 10.1109/ICCICA52458.2021.9697267.
- [19] A. Khan and A. V. Sakhare, "A Gestalt of Data Exploration and Conception," 2015 Fifth International Conference on Communication Systems and Network Technologies, Gwalior, India, 2015, pp. 908-912, doi: 10.1109/CSNT.2015.240.
- [20] A. K. Dudhpachare, T. V. Kuthe, C. V. Lalke, N. P. Wyawahare, R. Agrawal and P. Daigavane, "Process of RO's Waste Water Reuse & Water Management in Society by Using IOT Automation," 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2022, pp. 1546-1550, doi: 10.1109/ICAAIC53929.2022.9792805.
- [21] Potbhare, P.D., Rahul Agrawal, Sorte, S., Lokulwar, P., "IoT based Smart Air Pollution Monitoring System", "International Conference on Applied Artificial Intelligence and Computing, ICAAIC 2022, 2022, pp. 1829–1834.
- [22] Malewar, P., Kadu, R., Kakde, R., Wyawahare, N.P., Agrawal, R., "A Review Paper on Data Sensing and Acquisition Complexity and

Accuracy in Green House Monitoring Systems”, “International Conference on Emerging Trends in Engineering and Technology, ICETET, 2022, 2022-April.

- [23] Prajval Kadam, Mayur Sontakke, Aniket Narale, Rahul Agrawal, Swati Sorte, “Design of Mobile Solar Powered Public Station Using RFID” IEEE International Conference on Electronics and Renewable Systems (ICEARS), 16-18, March 2022. (Covered by Scopus).
- [24] Kshipra Rajguru, Pramod Tarpe, Vedanti Aswar, Kartik Bawane, Swati Sorte, Rahul Agrawal, “Design and Implementation of IoT Based Sleep Monitoring System for Insomniac People”, IEEE International Conference on Artificial Intelligence and Smart Systems (ICAIS 2022), 23-25, February 2022. (Covered by Scopus).
- [25] Prakash Yadav, Dr.Sanjay Dorle, Rahul Agrawal, “Arrhythmia Detection on ECG Signals Using Neural Network Approach” IEEE International Conference on Computational Intelligence & Computing Applications 2021, 26-27 Nov.2021. (Covered by Scopus).