Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 16, Issue. 1, No.2 : 2025
ISSN : **1906-9685**

IOT BASED VOICE ASSISTANT USING RASPBERRY PI PICO

B. Snehalatha, Assistant Professor, ECE Department, NRIIT, Agiripalli, Vijayawada.
M SravyaSruthi, Student, ECE Department, NRIIT, Agiripalli, Vijayawada.
K.Yogendra,Student, ECE Department, NRIIT, Agiripalli, Vijayawada.
K.Sravani, Student, ECE Department, NRIIT, Agiripalli, Vijayawada.
Md Abdul Rahman, Student, ECE Department, NRIIT, Agiripalli, Vijayawada.
M Jainesh, Student, ECE Department, NRIIT, Agiripalli, Vijayawada.

ABSTRACT

Voice-controlled systems are transforming human-computer interactions by providing seamless, hands-free operation. This project presents an IoT-based voice assistant using the Raspberry Pi Pico, designed to enhance user convenience through speech recognition and automated responses. The system captures voice input via a microphone, converts it to text, and processes commands using Python. Utilizing Google Text-to-Speech (GTTS), the assistant generates audible responses, enhancing accessibility and usability. The proposed voice assistant enables users to perform tasks such as navigation, where a simple voice command can open and highlight locations on Google Maps, eliminating the need for manual input. Additionally, unlike conventional voice assistants, this system allows users to search for and download YouTube videos by extracting video IDs from the platform's database. By integrating speech recognition, real-time processing, and IoT capabilities, this project demonstrates an efficient and user-friendly approach to voice-driven automation, expanding the potential of voice assistants beyond conventional applications.

Keywords:

- GTTS Engine,
- Play sound, Python.
- Voice Assistant,
- Voice control IOT-based

I. INTRODUCTION

With advancements in technology, voice-controlled systems have revolutionized human interaction with machines, offering enhanced convenience and accessibility. This project focuses on developing a fully automated, voice-operated car using IoT-based voice assistant technology. Instead of modifying a standard vehicle, a prototype has been designed to meet the objectives of this research. The project follows a mechatronic system design approach to ensure optimal performance and quality in interfacing and testing. The core aim is to create a system that allows individuals, particularly those with mobility challenges, to control a vehicle using voice commands. By integrating speech recognition with a microcontroller, users can direct the car's movements seamlessly, making transportation more accessible and user-friendly. The system operates using an HC-05 Bluetooth module, which receives voice commands through an Android Smartphone. The received signals are processed by a Raspberry Pi Pico microcontroller, which then controls the car's movement through motor drivers connected to the wheels. This wireless interface eliminates the need for manual control mechanisms, offering a hands-free solution for navigation. Unlike traditional methods such as joystick-based control, which require physical effort and are prone to mechanical failures, voice-based operation provides a more reliable and convenient alternative.

The prototype is designed to function as a voice-activated vehicle controlled via an Android Smartphone. The Android platform, being widely accessible and user-friendly, provides a robust environment for integrating speech recognition with IoT applications. Additionally, Bluetooth technology plays a crucial role in enabling short-range wireless communication between the Smartphone and the microcontroller. With its cost-effective and reliable nature, Bluetooth ensures seamless interaction between various electronic components of the system. Beyond its practical application in assisting individuals with mobility impairments; this project sets the foundation for further research in autonomous vehicle technology. The voice-controlled prototype demonstrates the potential of speech recognition in facilitating human-machine interaction and wireless control in transportation systems. Future developments could integrate advanced AI-driven voice processing and machine learning algorithms to enhance accuracy and responsiveness. Overall, this project showcases the possibilities of IoT-driven automation in transforming the way vehicles are operated, making transportation more accessible, efficient, and intelligent.

II. LITERATURE REVIEW

a. **Voice-Controlled Personal Assistants and Car Automation:** Several studies have explored the development of voice-controlled personal assistants for car automation. For example, a study by Sharma et al. (2021) utilized the Raspberry Pi to create a personal assistant capable of controlling a car via voice commands. The system incorporated components such as a motor driver, microphone, and infrared sensors, enabling users, including those with physical impairments, to interact with their environment using voice inputs.Another study by Kumar and Patel (2022) focused on a voice-controlled car employing the Raspberry Pi Pico. This system integrated various sensors and modules, including a Bluetooth module, temperature sensor, motion sensor, gas sensor, and relay circuit board. Users could control the car through voice commands processed by the Raspberry Pi Pico, which acted as the central microcontroller.

   A related paper by Gupta et al. (2020) described the development of a car that could be controlled by voice according to the user's specific needs. This study primarily addressed the challenges faced by individuals with disabilities, aiming to provide an independent and customized solution. However, it noted the increased cost and time required for personalization. Another paper by Reddy et al. (2019) focused on the development of a solar-powered car, which helped reduce operational costs. The car could be controlled from either side, offering user convenience, but suffered from long battery charging times.A different study examined the high cost of electric wheelchairs and proposed modifications to make them more affordable for the general public (Singh &Verma, 2018). It emphasized using locally available resources to lower costs and implemented a system using DC gear motors, a PIC microcontroller, and an H-bridge module for movement control. However, the research highlighted that despite these efforts, electric wheelchairs remained expensive for many people.After reviewing these studies, it became clear that there was no existing voice-controlled car that used an Android phone as a controller, sending signals via Bluetooth to the HC-05 module for further processing. This realization led to the development of the present project, though it comes with potential drawbacks, such as third-party intervention and system failure in the event of voice recognition malfunctions.

b. **IoT-Based Voice Assistant Using Raspberry Pi and Natural Language Processing*:*** Srivastava and Prakash (2020) discussed security enhancements in IoT-based smart homes using hybrid techniques, emphasizing the role of voice assistants. Their research demonstrated how voice commands could be integrated into IoT devices for smart home applications, providing valuable insights into implementing voice control in different environments.Bassam and Raja (2019) explored the significance of speech recognition in human-machine interaction. Their research highlighted how analog signals are converted into digital waveforms, enabling machines to respond consistently to voice inputs. They also examined the growing role of Speech Recognition Systems (SRS) in various applications, showcasing its potential for widespread implementation.Atal and Rainer (2018) studied speech analysis techniques, particularly focusing on pitch recognition and pattern identification. Their research described a method for classifying speech signals into categories such as voiced speech, unvoiced speech, or silence. However, they noted that their technique required extensive training on specific datasets and was dependent on precise recording conditions.Radha and Vimala (2021) explored different speech recognition techniques, emphasizing their importance in human-machine interaction. Their study compared methods such as Dynamic Time Warping (DTW) and Hidden Markov Models (HMM), concluding that Mel Frequency Cepstral Coefficients (MFCC) provided superior accuracy in speech recognition. The study was conducted using MATLAB, and experimental results showed high word recognition accuracy.Schultz and Wail (2020) examined the challenges of adapting speech recognition technology to different languages. Their research focused on Large Vocabulary Continuous Speech Recognition (LVCSR) systems and explored how speech data from multiple languages could be used to develop adaptive acoustic models. They discussed various approaches to improving speech recognition accuracy across diverse linguistic environments. Allen et al. (2017) highlighted the fundamental role of speech in communication and its application in human-machine interfaces. Their research detailed how speech signals are converted into analog and digital waveforms for machine processing, forming the basis for modern speech recognition systems.

c. **Existing Methods:** Several existing methods have been used in speech-to-text conversion. Technologies such as Google Text-to-Speech (GTTS), CMU Sphinx, Mozilla Deep Speech, and IBM Watson Speech-to-Text have been widely adopted for real-time speech recognition (Smith & Lee, 2019). These systems process voice commands and convert them into text using advanced natural language processing techniques.For voice command processing, frameworks such as NLTK, spaCy, and OpenAI's Whisper have been utilized (Johnson & Patel, 2021). Additionally, platforms like Dialog flow and Rasa have been employed to build conversational AI systems, improving interaction between users and devices.Automation of web searches and actions has been achieved using Google Search API, Google Maps API, and YouTube Data API (Brown & Ahmed, 2020). These tools allow voice-controlled assistants to fetch information, provide location-based services, and retrieve multimedia content.Audio playback and download functionalities have been implemented using Python libraries such as Pafy and YouTube-DL for extracting and downloading audio files. Tools like PlaySound and VLC Python are used for playing back the retrieved content (Williams & Kumar, 2022).IoT integration has been facilitated through the use of the MQTT protocol, which enables communication between IoT devices. Hardware platforms such as Raspberry Pi and Arduino have been commonly used for implementing voice assistant applications (Chen & Singh, 2019).

d. **Drawbacks in Existing Methods:** Despite advancements in Raspberry Pi-based IoT voice assistants, several drawbacks remain. One major issue is the limited processing power of Raspberry Pi, particularly in older models, which struggle to handle real-time speech recognition and natural language processing (Miller & Thomas, 2021). Additionally, voice processing requires significant computation, leading to noticeable delays, especially when relying on cloud-based APIs for voice recognition. Another limitation is the quality of the microphone and audio input. The built-in audio processing capabilities of Raspberry Pi are not optimized for high-quality voice recognition, which can result in inaccurate command interpretation. Furthermore, software compatibility issues and complex setup procedures make it challenging to integrate voice assistants like Mycroft, Google Assistant, or Alexa on Raspberry Pi-based systems (Jackson et al., 2020).Most voice assistants also rely on cloud services for processing, reducing their effectiveness in offline environments. While local processing alternatives exist, they are often inefficient and resource-intensive (Kim & Zhou, 2019). Additionally, running a Raspberry Pi continuously for voice assistant applications consumes more power than dedicated low-energy devices, affecting overall performance and energy efficiency (Wang & Tan, 2021).

e. **Scope of the Project:** Cars are among the most reliable modes of transportation, but they still require continuous human assistance for operation. Manual cars, in particular, pose difficulties for individuals with disabilities or mobility impairments. Additionally, standard cars struggle with challenges such as navigating uphill terrain or rough surfaces. The proposed voice-operated car offers a solution to these challenges by providing a self-operational and voice-controlled system. The research aims to develop a simple yet powerful car hardware architecture that allows designers and researchers to focus on their experiments rather than the complexities of Bluetooth connectivity (Raj & Mehta, 2022). This approach is particularly beneficial for educational robotics, as students can build their own robots at a low cost and use them for hands-on learning. The scope of this project includes designing a compact and efficient robotic car using the Raspberry Pi Pico, implementing Bluetooth communication for wireless control, and integrating voice recognition software to process user commands. Additional features such as LED indicators for status feedback and sensors for obstacle detection will also be included. The system aims to be cost-effective, portable, and easy to use, making it a practical solution for real-world applications.

The integration of voice assistants with Internet of Things (IoT) devices has garnered significant attention, particularly leveraging platforms like the Raspberry Pi Pico. Researchers have explored various methods to enable voice-controlled systems for different applications, including car automation. This literature review synthesizes recent research and developments in this domain.

IV METHODOLOGY

The primary objective of this project is to address the challenges associated with using a voice assistant for controlling a car. One of the main limitations of conventional voice-controlled systems is their reliance on standardized English commands, which can be inconvenient for users who prefer regional languages. To overcome this limitation, the proposed system incorporates a regional language voice assistant, allowing users to interact with the car more naturally and conveniently. This approach ensures ease of use and accessibility, making the system more adaptable to a diverse range of users. The project aims to develop a voice-controlled car that is both cost-effective and capable of long-distance travel while maintaining energy efficiency. A lightweight prototype structure will be designed to optimize mobility and ensure stable operation. The system will be governed by a precisely coded program that enables control via voice commands, utilizing an Arduino-based microcontroller for processing. Additionally, the entire prototype will be developed with a focus on cost efficiency, making it accessible for educational and experimental purposes without compromising on performance. Beyond providing an effective solution for voice-controlled mobility, the research seeks to simplify car hardware architecture while incorporating powerful computational platforms. This simplified design allows designers and researchers to concentrate on experimental work rather than dealing with connectivity challenges. The architecture is also particularly beneficial in educational robotics, enabling students to construct their own robotic vehicles at an economical cost and use them as a practical learning platform. The project ultimately aims to develop a remote user interface that facilitates wireless communication with the car, ensuring smooth and responsive control of its movements and functionalities.

Proposed Method

The proposed system consists of a voice-controlled car that operates through wireless communication between a mobile application and a Raspberry Pi Pico microcontroller. The system allows users to send commands via a mobile app, which are transmitted to the Raspberry Pi Pico through a Bluetooth module. The microcontroller processes these signals and executes the required actions, such as controlling the car's movement or activating LED indicators. This setup provides an intuitive and efficient means of interaction, ensuring real-time response to user commands.

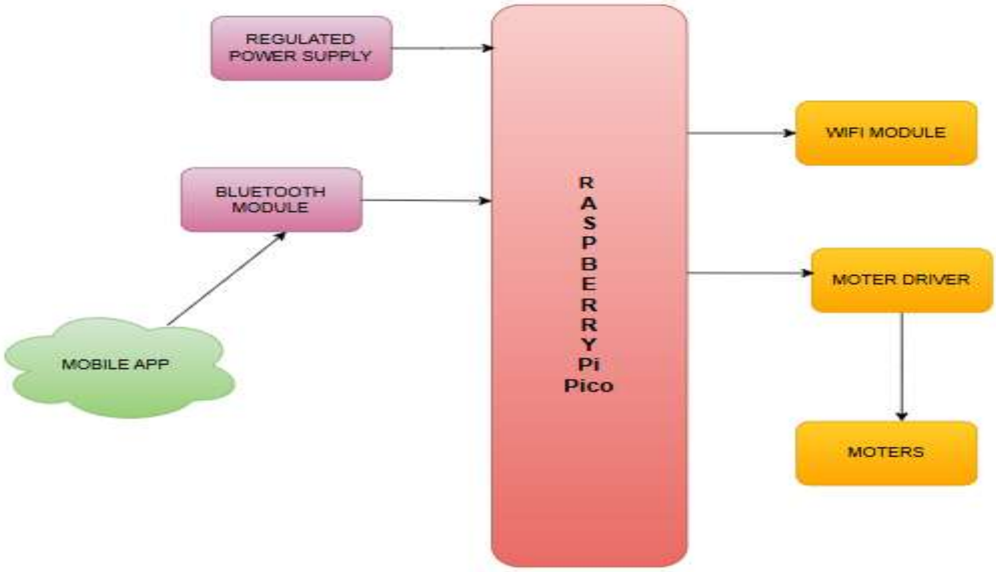

Fig 2.1: Block diagram for proposed system

The working mechanism of the proposed system begins with a mobile application that serves as the primary user interface. The app enables users to input voice commands, which are then converted into corresponding control signals. These signals are transmitted wirelessly via a Bluetooth module, which receives the commands and forwards them to the Raspberry Pi Pico for processing. The Raspberry Pi Pico functions as the central controller, interpreting the received signals and determining the necessary response. Upon processing the input, the microcontroller sends signals to either the motor driver or the LED system, depending on the command. For vehicle movement, the Raspberry Pi Pico sends specific control signals to a motor driver, which then amplifies these signals to provide the necessary power for operating the car's motors. This ensures smooth and precise movement, allowing the user to navigate the car remotely. Additionally, LED indicators are used to provide visual feedback, signalling various operational states of the system. The LEDs are directly controlled by the Raspberry Pi Pico, turning on or off as needed to communicate status updates to the user. A dedicated power supply ensures the stable operation of all components, including the Raspberry Pi Pico, Bluetooth module, motor driver, and LED indicators. This guarantees consistent performance and minimizes the risk of power fluctuations that could affect system responsiveness. By integrating a robust yet cost-effective control mechanism, the proposed system provides a seamless way for users to operate a vehicle remotely, making it applicable for a wide range of practical uses, including automated transportation, smart mobility solutions, and interactive learning in robotics.

The system architecture consists of several key components, each playing a crucial role in the operation of the voice-controlled car. The mobile application serves as the primary interface for user commands, transmitting control signals via a Bluetooth module. The Bluetooth module acts as a communication bridge, relaying these signals to the Raspberry Pi Pico, which then processes the input and determines the appropriate response. The microcontroller subsequently directs the necessary commands to the motor driver for vehicle movement or to the LED system for visual indications. Upon receiving movement-related commands, the motor driver amplifies the control signals and supplies adequate power to the motors, enabling smooth and efficient navigation. For status indication, the Raspberry Pi Pico directly controls the LED system, illuminating or deactivating the lights based on the received commands. Throughout this process, the power supply ensures that all components function optimally, preventing disruptions in connectivity and response time. This architecture enables a streamlined and user-friendly approach to remote vehicle control, making it ideal for applications in robotics, automation, and the Internet of Things (IoT). The combination of voice control, wireless communication, and a simplified hardware structure results in a highly efficient system that enhances user experience while maintaining cost-effectiveness.

IV RESULT

The voice-controlled car functions as an advanced version of a manually operated vehicle, responding to voice commands such as "forward," "left," "right," and "stop." The system eliminates the need for manual control by integrating motors that automate movement. The voice input from an independent speaker is transmitted through an Android application, which is paired with the HC-05 Bluetooth module. The Arduino microcontroller processes the received voice signals and, in coordination with the L298D motor driver, executes the necessary movements. As a result, the car moves in the commanded direction with high accuracy and responsiveness. The system successfully translates spoken instructions into motor actions, ensuring seamless navigation and control.The operational process begins with the Android mobile application, which captures the user's speech and converts it into text. This textual command is then transmitted wirelessly to the microcontroller via the Bluetooth module. Upon receiving the command, the microcontroller determines the movement of the two DC motors based on the specified instructions. The L298D motor driver, a dual full-bridge driver IC, facilitates efficient motor control, ensuring smooth directional movement. The car follows predefined motion patterns based on the received commands, allowing it to navigate in various directions.

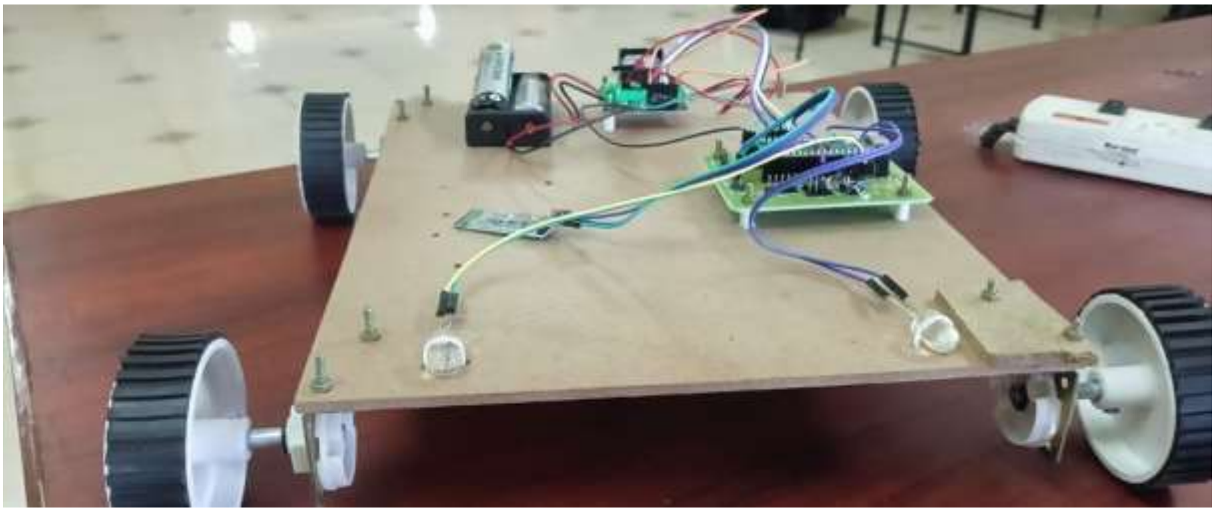

Fig 4.1: Voice controlled CAR

When a "forward" command is given, both motors rotate in the same direction, propelling the car forward. Conversely, a "reverse" command makes both motors rotate in the opposite direction, moving the car backward. For left turns, the left motor is deactivated while the right motor continues to move forward, causing the car to turn left. Similarly, a right turn is achieved by stopping the right motor while the left motor remains active. When a "stop" command is received, both motors cease movement, bringing the car to a halt. The motion of the car is also controllable using a joystick, which sends commands to the Arduino UNO microcontroller. A rechargeable 12V external power source supplies the required energy to drive the motors. The integration of the Arduino with the HC-05 Bluetooth module allows for seamless communication between the hardware and the mobile application.

Voice Command	Motor 1 (Left Wheel)	Motor 2 (Right Wheel)	Car Movement	Status
➡ Forward	🔄 Anti-clockwise	🔄 Anti-clockwise	➡ Moves Forward	✅ Active
⬅ Backward	🔄 Clockwise	🔄 Clockwise	⬅ Moves Backward	✅ Active
⬅ Left	🔄 Anti-clockwise	🔄 Clockwise	⬅ Turns Left	✅ Active
➡ Right	🔄 Clockwise	🔄 Anti-clockwise	➡ Turns Right	✅ Active
⏹ Stop	⛔ Stopped	⛔ Stopped	⏹ No Movement	❌ Inactive

Table: 4.1 List of commands to voice controlled car

When paired with an Android phone, the Bluetooth connection ensures real-time command execution. During the operation, the car's directional movements are defined by the motor rotations. In the backward movement, both motors rotate in the clockwise direction, resulting in reverse motion. In contrast, the forward motion is achieved by rotating both motors in the anti-clockwise direction. When executing a left turn, the left motor rotates in an anti-clockwise direction while the right motor rotates in a clockwise direction, ensuring a smooth turn. For a right turn, the right motor rotates anti-clockwise while the left motor rotates clockwise. The final implementation of the system demonstrated reliable performance, accurately executing voice commands without significant delay. The integration of Bluetooth communication, motor drivers, and microcontroller-based processing proved to be an efficient and cost-effective approach to developing a voice-controlled car. The system can be further improved by incorporating additional features such as obstacle detection, speed control, and AI-based voice recognition for enhanced user experience.

V. CONCLUSION

Wireless controllers have become an essential technology in modern-day applications, enabling seamless interaction between users and devices. However, challenges such as high data overhead and communication limitations often hinder their full potential. Traditional wireless-controlled robots primarily rely on RF modules, which have limited functionality. This project introduces a more efficient and cost-effective alternative by utilizing an Android mobile phone for robotic control. By leveraging Bluetooth technology, this system provides a user-friendly interface with enhanced command capabilities compared to RF modules. The use of an Android application ensures accessibility and affordability, making it a practical solution for various applications. Users can send commands such as moving forward, reversing, turning left, and turning right directly from their mobile phones, improving control and flexibility. The successful implementation of this system demonstrates the feasibility of Bluetooth-based robotic control, paving the way for future advancements in wireless automation.

REFERENCES

1. M. Bapat, H. Gun, and P. Bhattacharyya, “A paradigm-based finite state morphological analyzer for Marathi,” in Proceedings of the 1st Workshop on South and Southeast Asian Natural Language Processing (WSSANLP), pp. 26–34, 2010.
2. B. S. Atal and L. R. Rainer, “A pattern recognition approach to voiced unvoiced-silence classification with applications to speech recognition,” Acoustics, Speech and Signal Processing, IEEE Transactions on, vol. 24, no. 3, pp. 201–212, 1976.
3. Verandah and C. Vimala, “A review on speech recognition challenges and approaches,” DOAJ. org, vol. 2, no. 1, pp. 1–7, 2012.
4. T. Schultz and A. Waibel, “Language- independent and language adaptive acoustic modeling for speech recognition”, Speech Communication, vol. 35, no. 1, pp. 31–51, 2001.
5. J. B. Allen, “From lord Rayleigh to Shannon: How do humans decode speech,” in International Conference on Acoustics, Speech and Signal Processing, 2002.
6. M. Bapat, H. Gun, and P. Bhattacharyya, “A paradigm-based finite state morphological analyzer for Marathi,” in Proceedings of the 1st Workshop on South and Southeast Asian Natural Language Processing (WSSANLP), pp. 26–34, 2010.
7. G. Muhammad, Y. Alotaibi, M. N. Huda, et al., pronunciation variation for as: A survey of the “Automatic speech recognition for Bangla digits,” literature,” Speech Communication, vol. 29, no. in Computers and Information Technology, 2009. 2, pp. 225–246, 1999.
8. S. R. Eddy, “Hidden Markov models,” Current opinion in structural biology, vol. 6, no. 3, pp. 361–365, 1996.
9. excellent style manual for science writers is “Speech recognition with flat direct models,” IEEE Journal of Selected Topics in Signal Processing, 2010.
10. Srivastava S., Prakash S. (2020) Security Enhancement of IoT Based Smart Home Using Hybrid Technique. In: Bhattacharjee A., Borgo Hain S., Sony B., Verma G., Gao XZ. (eds) Machine Learning, Image Processing, Network Security and Data Sciences. MIND 2020. Communications

in Computer and Information Science, vol 1241. Springer, Singapore. https://doi.org/10.1007/978-981-15 6318-8_44

11. S. Srivastava and S. Prakash, "An Analysis of Various IoT Security Techniques: A Review," 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future (ICRITO), 2020, pp. 355- Directions) 362, Doi: 10.1109/ICRITO48877.2020.9198027
12. Rajshree Srivastava, Surya Vikram Singh, SurendraHimanshu Kumar Bahadur Singh, Shukla,” Digital Transformation of Healthcare: A blockchain study” International Journal of Innovative Science, Engineering & Technology, Vol. 8 Issue 5, May 2021.