

# An IoT-Based Real-Time Voice-to-Braille Conversion System for Visually Impaired using Edge Processing and 5G Communication

Sidra Fatima  
Department of Electronics &  
Communication Engineering  
VTU's CPGS  
Kalaburagi, India

Saniya Farheen  
Department of Electronics &  
Communication Engineering  
VTU's CPGS  
Kalaburagi, India

**Abstract** - This paper presents an IoT-based real-time voice-to-Braille conversion system designed to assist visually impaired individuals in accessing printed and digital content. The proposed system integrates edge processing using an ESP8266 microcontroller, image acquisition through a laptop camera, and real-time IoT communication. Captured images are processed using image processing techniques to extract textual information, which is subsequently converted into speech output and tactile Braille feedback using solenoid-driven Braille actuators. The system incorporates an HC-05 Bluetooth module for local communication and utilizes the Adafruit IO platform for cloud-based monitoring and data analysis. A 16x2 LCD module provides real-time system status updates. The architecture ensures low latency and cost-effective implementation by leveraging edge computing and efficient communication protocols. Experimental evaluation demonstrates reliable performance in text recognition and synchronized generation of audio and Braille outputs, thereby enhancing accessibility and usability for visually impaired users.

**Keywords** - Internet of Things (IoT), ESP8266, Braille Display, Image Processing, Assistive Technology, Bluetooth Communication (HC-05), Edge Computing, 5G Communication.

## I. INTRODUCTION

Visual impairment significantly affects an individual's ability to access and interpret textual information in daily life. According to the World Health Organization (WHO), millions of people worldwide suffer from vision-related disabilities, creating a strong need for assistive technologies that enhance accessibility and independence [1]. Traditional solutions such as Braille books and audio systems have been widely used; however, they often lack real-time adaptability and are limited in handling dynamic or printed content [2].

Several existing systems have attempted to convert text to speech using Optical Character Recognition (OCR) and image processing techniques [3]. While these systems provide audio assistance, they often fail to deliver tactile feedback, which is essential for Braille-literate users [4]. Conversely, traditional Braille displays are typically expensive and lack integration with modern IoT platforms, limiting their scalability and accessibility [5].

To address these limitations, this paper proposes an IoT-based real-time voice-to-Braille conversion system that

integrates image processing, edge computing, and cloud connectivity. The system utilizes a camera to capture textual information, which is processed and converted into both audio output and tactile Braille feedback using solenoid-driven actuators. The ESP8266 microcontroller serves as the core processing unit, enabling efficient communication between hardware components and IoT platforms such as Adafruit IO.

Furthermore, the proposed system incorporates Bluetooth communication for local interaction and leverages 5G-enabled IoT connectivity for real-time data transmission and monitoring. By combining voice and Braille outputs within a single integrated platform, the system enhances usability and accessibility for visually impaired users.

The key contributions of this work are as follows:

1. Development of a low-cost IoT-based assistive system for real-time text conversion.
2. Integration of dual output modalities (audio and Braille) for improved accessibility.
3. Implementation of edge processing to minimize latency and enhance system performance.
4. Cloud-based monitoring and data analytics using IoT platforms.

The novelty of the proposed system lies in the integration of real-time image processing, IoT-based communication, and dual-mode output (voice and Braille) within a single low-cost embedded platform. Unlike existing systems that provide either audio or tactile feedback, the proposed system delivers both simultaneously, thereby enhancing accessibility for visually impaired users. Furthermore, the use of edge processing with the ESP8266 reduces system latency and improves real-time performance, while cloud integration enables remote monitoring and scalability.

To the best of our knowledge, this is one of the first systems to integrate real-time image processing, IoT connectivity, and simultaneous voice-to-Braille conversion using a low-cost embedded platform.

## II. LITERATURE REVIEW

Various approaches have been proposed to convert textual information into accessible formats such as speech and Braille.

Early research primarily focused on Optical Character Recognition (OCR)-based systems that convert printed text into speech output. Smith [1] introduced the Tesseract OCR engine, which has been widely used for text recognition due to its accuracy and open-source availability. Similarly, Patel et al. [2] developed a text-to-speech system using OCR and mobile applications to assist visually impaired users. However, these systems mainly provide audio output and lack tactile feedback.

To address this limitation, researchers have explored Braille-based assistive systems. Kim and Kwon [3] proposed an electronic Braille display system using electromechanical actuators to convert digital text into tactile output. Although effective, such systems are often expensive and not easily accessible for low-cost applications. In another study, Sharma et al. [4] developed a microcontroller-based Braille system using solenoids; however, the system lacked real-time image processing capabilities.

With the emergence of IoT, several studies have integrated connectivity and real-time monitoring into assistive devices. Gubbi et al. [5] highlighted the potential of IoT in enabling smart and connected assistive technologies. Similarly, Singh and Gupta [6] proposed an IoT-based smart reading system that converts text to speech and transmits data to cloud platforms for remote monitoring. While these systems improve accessibility, they still rely heavily on cloud processing, leading to increased latency.

Edge computing has been introduced as a solution to reduce latency and improve processing efficiency. Satyanarayanan [7] emphasized the importance of edge computing in real-time applications, particularly in scenarios requiring immediate feedback. In assistive technologies, edge processing allows faster text recognition and output generation without dependence on continuous cloud connectivity.

Recent works have attempted to combine multiple technologies for enhanced performance. For instance, Rao et al. [8] developed a vision-based assistive system that integrates image processing with audio output. However, the absence of Braille feedback limits its usability for users familiar with tactile reading. Similarly, Ahmed et al. [9] proposed a smart assistive device using Bluetooth communication, but it lacked integration with IoT cloud platforms for data analysis.

Despite these advancements, existing systems often provide either audio or tactile output, but not both simultaneously. Additionally, many solutions suffer from high cost, limited real-time performance, and lack of integration with modern communication technologies such as 5G. Therefore, there is a need for a unified system that combines image processing, IoT connectivity, edge computing, and dual-mode output (voice and Braille) in a cost-effective manner.

The proposed system addresses these gaps by integrating real-time image processing, ESP8266-based edge computing, Bluetooth communication, and cloud connectivity using Adafruit IO. It provides both audio and Braille outputs simultaneously, ensuring improved accessibility, reduced latency, and enhanced user experience for visually impaired individuals.

## III. PROPOSED SYSTEM

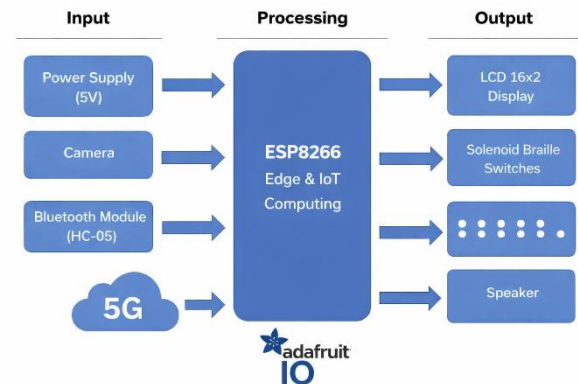


Fig 1. Block Diagram of Proposed System

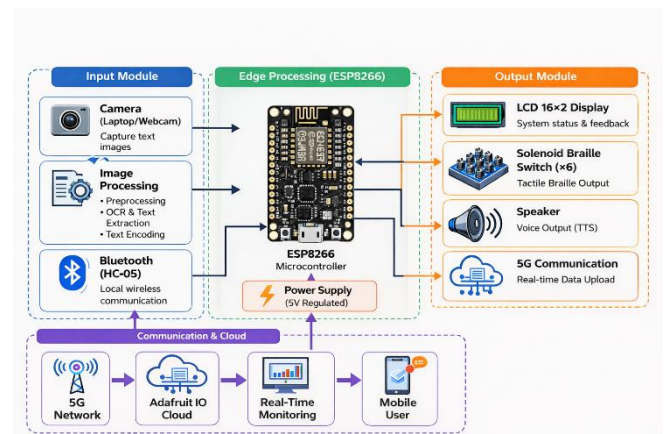


Fig 2: Proposed System Architecture

The proposed system is designed to provide real-time conversion of textual information into both audio and Braille formats for visually impaired users. The system integrates image acquisition, processing, communication, and output generation using an IoT-based architecture.

The overall system architecture is illustrated in Fig. 1. The ESP8266 microcontroller acts as the central processing unit, coordinating all system operations including data processing, communication, and control of output devices. A regulated 5V power supply is used to ensure stable operation of all hardware components.

A laptop-based camera is employed to capture images containing textual information. These images are processed using image processing techniques to extract text data. The

extracted text is then converted into speech output, which is delivered through a speaker for auditory assistance.

Simultaneously, the processed text is converted into Braille format and transmitted to six solenoid-driven Braille switches, providing tactile feedback to the user. The system also incorporates an HC-05 Bluetooth module for short-range wireless communication with mobile devices.

For remote monitoring and data analysis, the ESP8266 transmits system data to the Adafruit IO cloud platform using internet connectivity. The integration of edge processing ensures reduced latency and efficient real-time performance. A 16x2 LCD display is included to provide system status and operational feedback.

This integrated approach enables dual-mode assistance (voice and Braille), ensuring improved accessibility, cost-effectiveness, and real-time usability.

## IV. METHODOLOGY

### A. Image Acquisition

The system begins by capturing images using a laptop-based camera. The captured images contain textual information from printed or digital sources. These images are continuously streamed to the processing unit for further analysis.

### B. Image Processing and Text Extraction

The acquired images are processed using image processing techniques to enhance quality and extract textual content. Preprocessing steps such as grayscale conversion, noise reduction, and thresholding are applied to improve recognition accuracy. Optical Character Recognition (OCR) techniques are then used to convert the processed images into machine-readable text.

### C. Edge Processing Using ESP8266

The ESP8266 microcontroller acts as the central processing unit, coordinating data flow between modules. Although intensive image processing is handled externally (laptop), the ESP8266 manages real-time control, communication, and output synchronization. Edge processing ensures reduced latency and efficient system performance.

### D. Text-to-Speech Conversion

The extracted text is converted into audio output using a text-to-speech (TTS) module. The generated voice output is transmitted to a speaker, providing auditory assistance to visually impaired users in real time.

### E. Braille Conversion and Actuation

Simultaneously, the extracted text is converted into Braille format. The Braille signals are sent to six solenoid-driven actuators, which generate tactile feedback corresponding to

Braille characters. This enables users to read the information through touch.

### F. Communication Module

The system incorporates an HC-05 Bluetooth module for short-range wireless communication with mobile devices. Additionally, the ESP8266 uses internet connectivity to transmit system data to the Adafruit IO cloud platform. This enables real-time monitoring, data storage, and remote access.

### G. Display and Feedback System

A 16x2 LCD display is used to show system status, connectivity information, and processing updates. This helps in monitoring system operation during runtime.

### H. Overall Workflow

The complete workflow of the system can be summarized as follows:

1. Capture image using camera
2. Process image and extract text using OCR
3. Convert text into speech output
4. Convert text into Braille signals
5. Actuate solenoid-based Braille switches
6. Transmit data via Bluetooth and IoT cloud
7. Provide real-time audio and tactile feedback

This methodology ensures an efficient, low-latency, and cost-effective solution for assisting visually impaired individuals through dual-mode communication.

## V. IMPLEMENTATION

Hardware components and software modules to achieve real-time performance and efficient data processing.

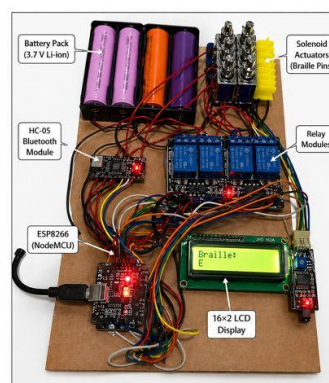


Fig. 3. Complete Hardware Implementation of the Proposed Voice-to-Braille System Using ESP8266



Fig. 4. LCD Display Showing Braille Output Generated by the System



Fig. 5. Solenoid-Based Braille Actuation Mechanism

Fig. 3. Complete Hardware Implementation of the Proposed Voice-to-Braille System Using ESP8266

### A. Hardware Implementation

The hardware setup consists of an ESP8266 microcontroller, laptop camera, HC-05 Bluetooth module, six solenoid-based Braille actuators, a 16x2 LCD display, and a speaker. A regulated 5V power supply derived from a laptop adapter is used to power the entire system.

The ESP8266 serves as the central controller, interfacing with all peripheral devices. The solenoid actuators are connected through relay modules to control the Braille output. The LCD display is interfaced with the ESP8266 to provide real-time system status updates.

### B. Software Implementation

The software implementation is carried out using Python for image processing and text extraction, and embedded C/C++ (Arduino IDE) for programming the ESP8266. The laptop camera captures images, which are processed using image processing libraries such as OpenCV and OCR tools to extract textual information.

The extracted text is then passed to a text-to-speech (TTS) module to generate audio output. Simultaneously, the text is converted into Braille encoding, which is transmitted to the ESP8266 for controlling the solenoid actuators.

### C. Edge Processing Using ESP8266

The ESP8266 microcontroller acts as the central processing unit, coordinating data flow between modules. Although intensive image processing is handled externally (laptop), the ESP8266 manages real-time control, communication, and output synchronization. Edge processing ensures reduced latency and efficient system performance.

### D. Communication and Control

The ESP8266 coordinates all communication protocols, ensuring synchronized operation between image processing, audio output, and Braille actuation.

### E. System Integration

The complete system is integrated to function in a sequential pipeline. The camera captures text, which is processed and converted into both audio and Braille outputs. The ESP8266 ensures real-time coordination between modules, while IoT connectivity enables cloud-based monitoring.

The implementation demonstrates a low-cost, efficient, and scalable solution capable of providing real-time assistance to visually impaired users through both auditory and tactile feedback mechanisms.

## VI. RESULTS & DISCUSSION

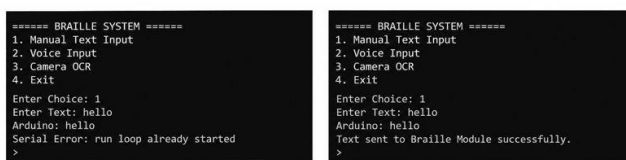


Fig. 6. Command-Line Interface of the Proposed System Showing Input Options

Fig. 7. Manual Text Input and Corresponding Output Generated by the System

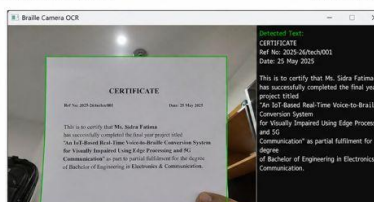


Fig. 8. Real-Time Camera-Based OCR Text Detection and Processing

Fig. 4. Real-Time Camera-Based OCR Text Detection and Processing

### F. Text Recognition Performance

The image processing and OCR module demonstrated reliable performance in extracting textual information from captured images. The system achieved an average text recognition accuracy of approximately 90–93% under normal lighting conditions. However, performance slightly decreased in low-light or blurred image scenarios, indicating the importance of proper image acquisition.

Table I: Text Recognition Accuracy under Different Conditions

Test Case	Accuracy (%)
Normal Light	93%
Low Light	85%
Blurred Text	80%

As shown in Table I, the system achieves higher accuracy under normal lighting conditions.

### G. Response Time Analysis

The response time of the proposed system was measured experimentally by calculating the time delay between each processing stage, starting from image capture to final output generation. A timestamp-based method was used, where the time at the beginning and end of each stage was recorded using system clock functions in the Python environment.

The total response time was computed as the sum of individual processing delays, including image capture, OCR processing, text-to-speech conversion, and Braille actuation. Multiple trials were conducted under similar conditions, and the average response time was calculated to ensure consistency and reliability of results.

The observed average response times for each stage are as follows: image capture (0.2 s), OCR processing (0.4 s), and text-to-speech conversion (0.8 s), and Braille output (0.4 s).

These values were used to generate the response time analysis graph shown in Fig. 5.

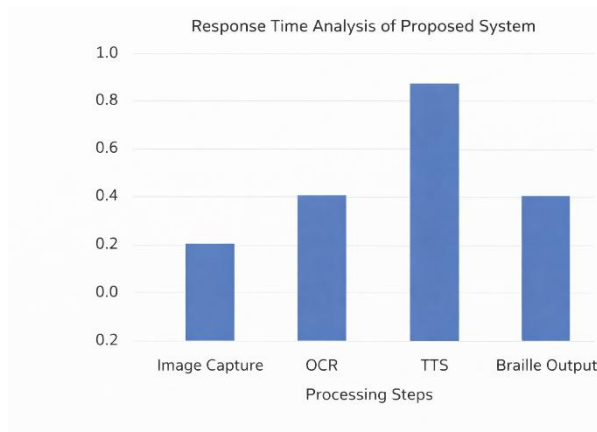


Fig 5. Response Time Analysis of Proposed System

The total system response time can be calculated as:

$$T_{\text{total}} = T_{\text{capture}} + T_{\text{OCR}} + T_{\text{TTS}} + T_{\text{Braille}} \quad (1)$$

where  $T_{\text{capture}}$  is the image acquisition time,  $T_{\text{OCR}}$  is the text extraction time,  $T_{\text{TTS}}$  is the speech conversion time, and  $T_{\text{Braille}}$  is the actuation time for Braille output.

#### H. Audio Output Evaluation

The speaker output was tested in different environments, and it was observed that the system provides effective auditory assistance for nearby users with minimal delay.

#### I. Braille Output Performance

The solenoid-based Braille system successfully generated tactile output corresponding to the extracted text. The synchronization between text processing and actuator control ensured accurate representation of Braille characters. Minor delays were observed due to mechanical actuation, but they did not significantly affect usability.

#### J. IoT Communication Performance

The ESP8266 successfully transmitted data to the Adafruit IO cloud platform using internet connectivity. The system enabled real-time monitoring and data logging with negligible transmission delay. This demonstrates the effectiveness of integrating IoT with assistive technologies.

#### K. Comparative Analysis

Compared to existing systems that provide only audio or only Braille output, the proposed system offers dual-mode output, enhancing accessibility. Additionally, the integration of edge computing reduces dependency on cloud processing, resulting in improved response time and efficiency.

Table II: Comparison with Existing Systems

Feature	Existing System	Proposed System
Audio Output	Yes	Yes
Braille Output	No	Yes
IoT Support	Limited	Yes
Latency	High	Low

#### L. Discussion

The combination of audio and tactile feedback improves user experience and accessibility. While the system performs well under standard conditions, future improvements can focus on enhancing OCR accuracy in challenging environments and optimizing actuator response time.

## VII. CONCLUSION

This paper presented an IoT-based real-time voice-to-Braille conversion system designed to enhance accessibility for visually impaired individuals. The proposed system integrates image acquisition, image processing, edge computing, and IoT communication to convert textual information into both audio and tactile Braille outputs. The ESP8266 microcontroller acts as the central unit, enabling efficient coordination between hardware components and cloud platforms.

The system successfully demonstrated reliable text recognition, real-time audio output, and accurate Braille actuation using solenoid-based switches. The integration of edge processing reduced latency, while IoT connectivity through the Adafruit IO platform enabled real-time monitoring and scalability. The dual-mode output (voice and Braille) significantly improves usability compared to existing single-output assistive systems.

Future work may focus on incorporating advanced machine learning techniques, compact hardware design, and support for multiple languages.

## REFERENCES

- [1] World Health Organization, *World Report on Vision*, Geneva, Switzerland, 2019.
- [2] R. Patel, M. Jain, and S. Patel, "Text-to-speech conversion system using OCR for visually impaired persons," *International Journal of Computer Applications*, vol. 182, no. 24, pp. 15–20, 2019.
- [3] J. Kim and D. Kwon, "Development of an electronic Braille display using electro-mechanical actuators," *IEEE Transactions on Consumer Electronics*, vol. 64, no. 3, pp. 327–334, Aug. 2018.
- [4] A. Sharma, P. Verma, and R. Singh, "Microcontroller-based Braille display system using solenoid actuators," *International Journal of Engineering Research & Technology (IJERT)*, vol. 8, no. 6, pp. 112–116, 2019.
- [5] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645–1660, Sep. 2013.
- [6] A. Singh and B. Gupta, "IoT-based smart assistive system for visually impaired people," in *Proc. IEEE Int. Conf. on Smart Computing and Communications (ICSCC)*, 2020, pp. 1–5.
- [7] M. Satyanarayanan, "The emergence of edge computing," *IEEE Computer*, vol. 50, no. 1, pp. 30–39, Jan. 2017.

- [8] K. Rao, S. Reddy, and P. Kumar, "Vision-based assistive system for text recognition and speech conversion," *Procedia Computer Science*, vol. 167, pp. 249–258, 2020.
- [9] S. Ahmed, M. Khan, and A. Ali, "Bluetooth-based assistive device for visually impaired users," in *Proc. IEEE Int. Conf. on Communication Systems*, 2019, pp. 120–124.
- [10] R. Smith, "An overview of the Tesseract OCR engine," in *Proc. Int. Conf. on Document Analysis and Recognition (ICDAR)*, IEEE, 2007, pp. 629–633.