A Hybrid CNN-RNN Based Framework for Transforming Braille to Multilingual Speech

G Naga Sreevidya [1], M Rajitha [2], K Mahesh Babu [3], K Venkata Tarun [4], KRMC Sekhar [5*]

[1-4] Student, Department of Information Technology, Kallam Haranadhareddy Institute of Technology(A), Chowdavarm, Guntur, India

[5] Assistant Professor, Department of Information Technology, Kallam Haranadhareddy Institute of Technology (A), Chowdavarm, Guntur, India

* Corresponding Author: **K.R.M.C Sekhar**

Article Info

P-ISSN: 3051-3383
E-ISSN: 3051-3391
Impact Factor (RSIF): 8.40
Volume: 07
Issue: 01
Received: 03-01-2026
Accepted: 05-02-2026
Published: 07-03-2026
Page No: 36-41

Abstract

Braille is a pivotal medium of communication for individuals with visual impairments, enabling independent access to written information. With the rapid advancement of digital technologies, there is a growing demand for systems that can convert traditional Braille documents into accessible digital formats. This work introduces a hybrid CNN-RNN based framework that leverages image processing and deep learning techniques to transform Braille images into readable text. The proposed system incorporates preprocessing steps such as noise reduction, image enhancement, and Braille dot segmentation to ensure accurate feature extraction. Convolutional Neural Networks (CNNs) are employed for image segmentation and pattern recognition, while Recurrent Neural Networks (RNNs) capture sequential relationships between Braille characters, enhancing recognition accuracy across variations in spacing, orientation, and image quality. Beyond text recognition, the system integrates multilingual translation and text-to-speech functionality, offering both visual and auditory access to information. The text-to-speech module delivers natural and customizable audio output, catering to diverse user preferences. In this work, a well-defined interface is implemented to provide ease of access with multilingual support for visually impaired individuals. Our experimental results justify that the proposed framework exhibits higher accuracy in converting Braille image into text which can further be converted into audio for better accessibility.

DOI: https://doi.org/10.54660/IJAIET.2026.7.1.36-41

Keywords: braille recognition system, CNN, RNN, Image Processing, Text-to-Speech, deep learning, pattern recognition

1. Introduction

Braille is a pivotal communication medium for individuals with visual impairments [[1]. Rapid advancements in digitization have enabled the development of technologies that convert Braille characters into digital text. Many researchers have worked in this direction by applying various CNN and RNN models. However, existing systems face challenges in accuracy, robustness, cost, multilingual support, and user personalization [2]. This work emphasizes on transforming braille characters into audible speech. Such a kind of system allows visually impaired users to access information on portable devices in a customized way without depending on Braille books. The proposed system integrates advanced image processing and deep learning techniques to detect and interpret Braille patterns from scanned or captured images. These patterns are then converted into digital text, which can subsequently be transformed into multilingual speech. CNNs are employed for segmentation and pattern recognition, while RNN capture sequential character relationships. This system enhances recognition accuracy under varying conditions such as noise, distortions, and irregular dot spacing. Beyond text recognition, Speak Braille incorporates multilingual translation and text-to-speech synthesis, enabling users to access information in their preferred language and voice settings.

This dual-mode accessibility eliminates the need of tactical reading and empowers visually impaired individuals to engage with digital content more efficiently. Unlike traditional assistive technologies that often depend on specialized hardware or manual intervention, the proposed system offers a well defined interface as a solution that can be deployed on widely available portable devices. This work provides new opportunities for inclusive education, supports seamless communication, and strengthens autonomy within visually impaired communities.

2. Problem Statement and Objectives

All existing systems have limited support for multilingual translation, lack of integrated text-to-speech functionality, and insufficient user-friendly design. These shortcomings highlight the need for a robust, Optical Braille Recognition (OBR) System that can automatically recognize Braille characters from images, improve accuracy through sequential learning models and provide multilingual text-to-speech conversion in a simple, portable, and accessible interface.

The primary objective of this system is to address the limitations of the existing works. The following objective are considered:

1. Automatic recognition of Braile Characters from images and convert them to text.
2. To improve recognition accuracy by employing RNN and CNN models.
3. To support the conversion of recognized text into speech with multilingual capabilities.
4. To provide well-defined and user-friendly interface

2.1. Related Work

This section summarizes the existing work, highlighting their strengths and limitations, and emphasizes the research gaps.

Kishore *et al.* [3] proposed a cost-effective Braille-to-voice conversion system in which input is provided through a Braille keyboard and processed using an Arduino-based embedded system. The system enables basic communication for visually and hearing-impaired individuals. However, it lacks support for real-time processing and multilingual translation, limiting its scalability and applicability in advanced systems.

D. R. Lim Roque et. al. [4] proposed a Brielle to text transformation system. That further produces speech from the recognized text. They used simple image processing techniques to reduce computational complexity. This system provides limited multilingual support and faces difficulty in detecting distorted or damaged Braille.

J. D. R. Cabangis *et al.* [5] proposed smart Braille translator using Raspberry Pi. This system supports both Braille-to-speech and speech-to-Braille, making it a bidirectional assistive communication system. It supports real time processing capability. The main limitation with this approach is hardware dependent and high implementation cost.

Premanathan G *et al.* [6] proposed multilingual Braille-to-Text using GTTS. The system provides a cost-effective multilingual Braille-to-text-to-speech solution using Raspberry Pi; however, it is limited by processing power, scalability, and dependency on input quality.

Vishnu Preetham Revelli *et al.* [7] proposed a deep learning-based system that converts Braille images into readable text and subsequently into speech. This model achieved an accuracy of approximately 96.15%, demonstrating robustness even under low-light conditions and varying image quality. The system supports only the English language, requires significant processing time and computational resources, and exhibits low performance when handling blurred or noisy images.

The following research gaps are identified:

- Inadequate Multilingual Capabilities
- Dependency on Specialized Hardware
- Reduced Accuracy in Noisy or Low-Quality Images
- High Computational Complexity
- Insufficient Integration of End-to-End Systems
- Limited User-Centric and Accessibility-Focused Design

3. Proposed Methodology

3.1. Proposed Framework

This section demonstrates a proposed framework that reads Braille characters and convert them into Text. Furthermore, the text is converted into speech, enabling visually impaired individuals to access and understand the content easily.

The proposed system is designed to convert Braille images into readable text and subsequently translate the text into multiple languages such as Hindi or Telugu. The application consists of both front-end and back-end components that work together to perform Braille-to-text and text-to-speech conversion.

On the front end, users interact with a simple interface that allows them to capture a Braille image using a camera or select an existing image from the media gallery. The selected image is then sent to the back end for processing.

In the back end, image processing techniques along with Google ML Kit are used to enhance the image and detect Braille dot patterns. The proposed framework uses MobileNet for efficient feature extraction from the Braille image. In our work, MobileNet is chosen due to its lightweight architecture and suitability for mobile and real-time applications. The extracted features are then passed to a Recurrent Neural Network (RNN), which learns the sequential structure of Braille characters and improves recognition accuracy. Based on this, the detected Braille patterns are converted into meaningful digital text through a text extraction process. The extracted text can further be translated into different languages using Google Translator. Finally, the translated or original text is converted into speech using a text-to-speech module, enabling visually impaired users to listen to the content.

Thus, the system ensures accurate recognition, efficient processing, and enhanced accessibility for visually impaired individuals.

The architecture of the proposed system is depicted in Fig.1.

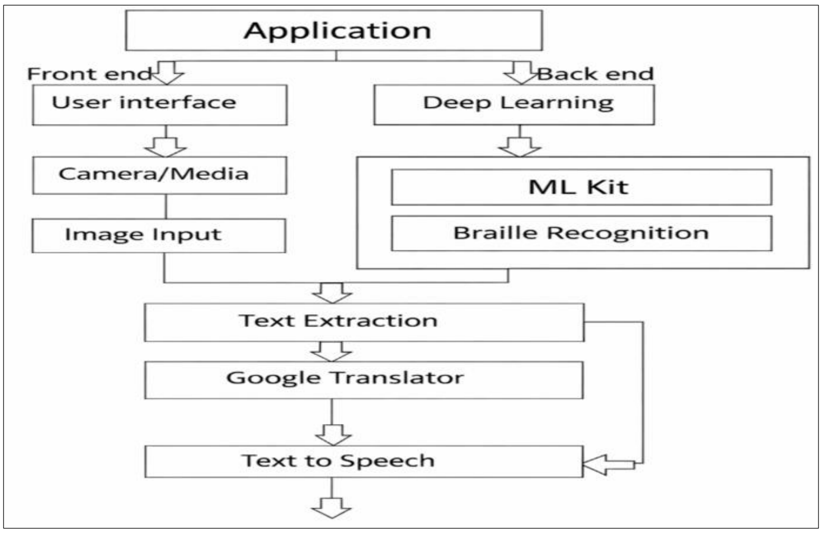

Fig.1: Architecture of Proposed System

The sequence of activities performed by proposed framework is illustrated in Fig.2.

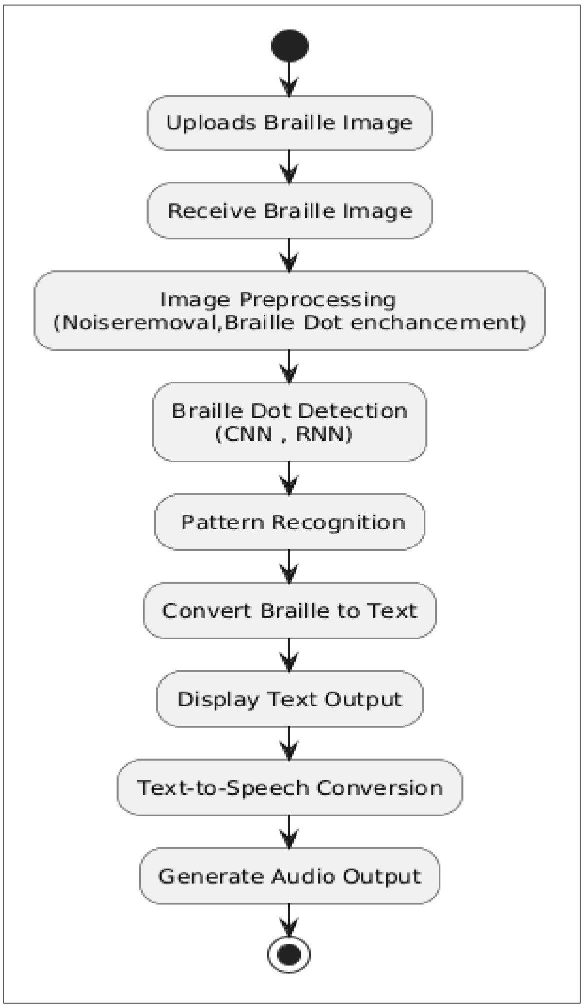

Fig.2: Activity Diagram of Proposed Framework

The Fig.2 demonstrates the workflow of our proposed framework. It starts with the user uploading a Braille image, followed by image preprocessing to enhance quality. The system then detects Braille dot patterns and uses a recognition model to convert them into readable text. The output text is displayed to the user and can also be converted into speech using a text-to-speech module. Overall, the system provides an efficient way to transform Braille images into understandable text and audio.

3.2. Implementation of Proposed Framework

The proposed system is implemented by using the following tools and packages listed in Table 1.

Table 1: Tools Used

Category	Tools
Operating System	Windows 10 or above
Programming Language	Python 3.10+
Framework	Flask
Image Processing	OpenCV
Numerical Computing	NumPy
Deep Learning	TensorFlow / Keras
Image Handling	Pillow (PIL)

First, the image is preprocessed using grayscale conversion and filtering. Then thresholding and morphological operations are applied to enhance Braille dots. Contour detection and segmentation identify individual cells. Features are extracted using MobileNet, and RNN converts them into text. Finally, the text is translated and converted into speech.

4. Results

Our proposed system is successfully implemented using Python, Flask, and deep learning techniques. The system was tested with multiple Braille captured images The experimental results demonstrate that the system is capable of accurately detecting Braille dot patterns and converting them into readable text. The output of proposed system is shown in figure 3, 4 & 5.

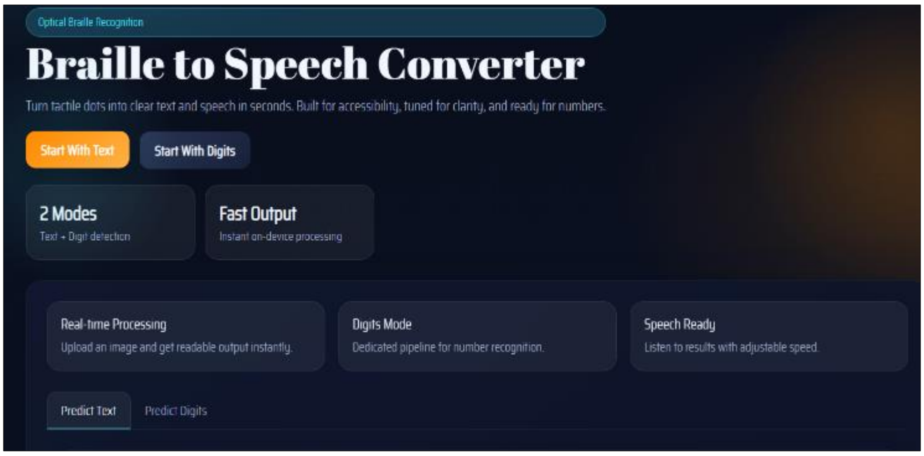

Fig 3: Home Page

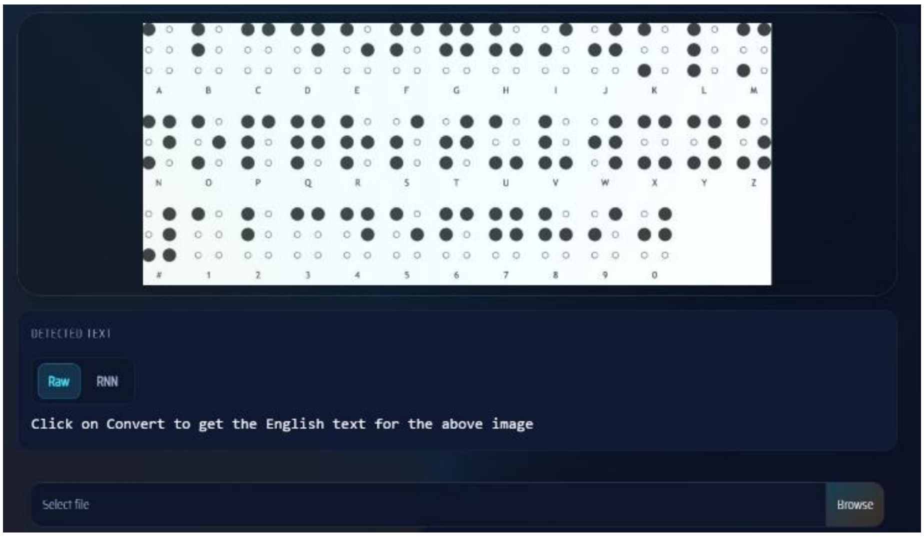

Fig 4: Braille Image Selection

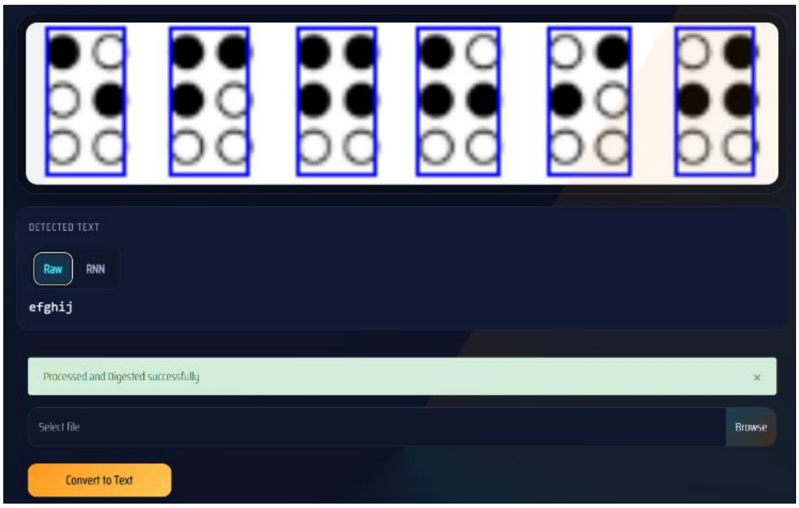

Fig 5: Braille Image to Text Generation

The above output screens show that the system correctly processes the input image, performs preprocessing, and generates corresponding text output. Furthermore, the text-to-speech module successfully converts the extracted text into audible speech, enabling users to listen to the content.

In addition, the system supports translation into multiple languages such as Hindi and Telugu, enhancing accessibility for diverse users. The results indicate that the system performs efficiently with minimal processing time and provides a user-friendly interface.

Overall, the proposed system achieves reliable Braille recognition and effective text and speech conversion, making it a useful assistive tool for visually impaired individuals.

5. Conclusion

The proposed hybrid CNN–RNN framework effectively converts Braille images into readable text and further transforms the text into speech. The proposed system provides a well-designed interface that ensures ease of interaction and usability. By integrating image processing techniques with MobileNet and RNN, the system achieves accurate recognition of Braille patterns. The inclusion of translation and text-to-speech modules enhances functionality by enabling multi-language support and audio output. Our experimental results indicate that the system is efficient, reliable, and user-friendly. Overall, the proposed approach serves as a robust assistive solution that helps visually impaired individuals to easily access information over portable devices. Furthermore, it supports independent learning and communication. In the future, this work can be extended by integrating advanced deep learning models along with federated learning to improve accuracy, performance and data privacy.

References

1. Cabangis JDR, *et al*. Smart Braille Translator: Integrating Braille-to-Speech and Speech-to-Braille Using Raspberry Pi. In: 2024 IEEE International Conference on Computing (ICOCO); 2024; Kuala Lumpur, Malaysia. Piscataway, NJ: IEEE; 2024. p. 202-7. doi: 10.1109/ICOCO62848.2024.10928214.
2. Herur S, Bhoomika GR, Arpitha KR, Shubhashini KH, Syeda Shaista Zainab P. Voice Assisted Text Reading System For Blind And Visually Impaired Persons. Int J Eng Res Technol. 2023;11(5).
3. Kishore K, *et al*. Braille script to voice conversion. In: 2017 International Conference on Computing Methodologies and Communication (ICCMC); 2017. p. 1080-2.
4. Lim Roque DR, Marcelo Mateo JT, Linsangan NB, Juanatas RA, Villamor I. Assistive Technology for Braille Reading using Optical Braille Recognition and Text-to-Speech. In: 2022 IEEE 14th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM); 2022; Boracay Island, Philippines. Piscataway, NJ: IEEE; 2022. p. 1-6. doi: 10.1109/HNICEM57413.2022.10109516.
5. Li T, Zeng X, Xu S. A Deep Learning Method for Braille Recognition. In: 2014 International Conference on Computational Intelligence and Communication Networks; 2014; Bhopal, India. p. 1092-5. doi: 10.1109/CICN.2014.229.
6. P G, K V S, R S, M M, V K V. Multilingual Braille-to-Text Conversion and Text-to-Speech System using Raspberry Pi. In: 2025 7th International Conference on Inventive Material Science and Applications (ICIMA); 2025; Namakkal, India. Piscataway, NJ: IEEE; 2025. p. 1238-43. doi: 10.1109/ICIMA64861.2025.11074178.
7. Revelli VP, Sharma G, Kiruthika Devi S. Automate extraction of braille text to speech from an image. Adv Eng Softw. 2022;172:103180. doi: 10.1016/j.advengsoft.2022.103180.
8. Latif G, Brahim GB, Abdelhamid SE, Alghazo R, Alhabib G, Alnujaidi K. Learning at Your Fingertips: An Innovative IoT-Based AI-Powered Braille Learning System. Appl Syst Innov. 2023;6(5):91. doi: 10.3390/asi6050091.
9. Kaur P, Ramu S, Panchakshari S, Krupa N. Conversion of Hindi Braille to Speech using Image and Speech Processing. In: 2020 IEEE 7th Uttar Pradesh Section International Conference on Electrical, Electronics and

10. Computer Engineering (UPCON); 2020; Prayagraj, India. Piscataway, NJ: IEEE; 2020. p. 1-6. doi: 10.1109/UPCON50219.2020.9376566.
11. Dholvan M, Mantha S, Mohanalaxmi Y, Keerthana R, Varunkumar K. Revolutionizing Braille with Speech Synthesis. In: 2025 11th International Conference on Smart Computing and Communications (ICSCC); 2025; Kochi, India. Piscataway, NJ: IEEE; 2025. p. 1-5. doi: 10.1109/ICSCC66177.2025.11233490

How to Cite This Article

Sreevidya GN, Rajitha M, Babu KM, Tarun KV, Sekhar KRMC. A hybrid CNN-RNN based framework for transforming Braille to multilingual speech. Int J Artif Intell Eng Transform. 2026;7(1):36–41. doi:10.54660/IJAIET.2026.7.1.36-41