The Scientific Temper (2025) Vol. 16 (spl-1): 48-52
E-ISSN: 2231-6396, ISSN: 0976-8653
Doi: 10.58414/SCIENTIFICTEMPER.2025.16.spl-1.06
https://scientifictemper.com/
ORIGINAL RESEARCH PAPER

Enhancing Kannada text-to-speech and braille conversion with deep learning for the visually impaired

Ganga Gudi[1*], Mallamma V Reddy[2], Hanumanthappa M[3]

Abstract
Advancements in assistive technology have greatly improved accessibility for visually impaired individuals, enabling seamless interaction with textual content. This research introduces a novel approach that converts Kannada text into both speech and Braille, promoting multilingual accessibility. The proposed system incorporates a support vector machine (SVM) for Kannada text-to-Braille conversion and a deep learning-based text-to-speech (TTS) model for speech synthesis. The Braille translation module accurately maps Kannada characters to their respective Braille representations using SVM classifiers, ensuring precise conversion. Simultaneously, the speech synthesis component utilizes Tacotron2 for converting Kannada text into mel-spectrograms, followed by WaveNet/HiFi-GAN to produce high-quality Kannada speech. A dataset containing 2000 Kannada text-Braille pairs and corresponding text-speech samples is employed for training and evaluation. Experimental findings validate the effectiveness of the proposed system in accurately translating Braille while generating clear and natural Kannada speech. The integration of machine learning and deep learning techniques enhances efficiency, scalability, and usability, making this a reliable assistive tool for visually impaired Kannada-speaking individuals.

Keywords: Kannada Text-to-Braille, Speech synthesis, Text-to-Speech, Support Vector Machine, Tacotron2, HiFi-GAN, WaveNet, Braille Conversion.

Introduction

Technological advancements in speech synthesis and assistive technology have significantly improved accessibility for visually impaired individuals, enabling them to engage with digital content, communicate effectively, and navigate daily life independently. Text-to-Speech (TTS) and Braille conversion systems play a crucial role in this accessibility framework by transforming text into both auditory and tactile formats, ensuring inclusivity in education, employment, and digital communication (Rupanagudi *et al.*, 2014).

[1]Department of Computer Applications, PES University, RR Campus, Bangalore, India.

[2]Department of Computer Science, Rani Channamma University, Belagavi, India.

[3]Department of Computer Science, Bangalore University, Bangalore, India.

***Corresponding Author:** Ganga Gudi, Affiliation, E-Mail: gangagudi1@gmail.com

How to cite this article: Gudi, G., Reddy, M.V., Hanumanthappa M. (2025). Enhancing Kannada text-to-speech and braille conversion with deep learning for the visually impaired. The Scientific Temper, **16**(spl-1):48-52.

Doi: 10.58414/SCIENTIFICTEMPER.2025.16.spl-1.06

Source of support: Nil

Conflict of interest: None.

Despite progress in TTS technology, challenges persist in developing efficient and accurate Kannada speech synthesis models due to the linguistic complexity of the script, phonetic variations, and the limited availability of annotated datasets. Kannada, a Dravidian language spoken predominantly in Karnataka, India, possesses a rich phonetic structure, with numerous consonant-vowel combinations, sandhi rules, and complex diacritic usage, making traditional TTS models less effective (Abdulmalik AlSalman *et al.*, 2021). Similarly, Kannada-to-Braille conversion presents challenges in accurately mapping Kannada characters to Braille patterns due to the script's inherent intricacies. To address these challenges, this research proposes an integrated Kannada TTS and Braille conversion system that enhances accessibility for visually impaired individuals by providing Kannada text in both speech and Braille formats, as shown in Fig 1. The system comprises two primary components:

Kannada TTS Synthesis: Utilizing deep learning models such as Tacotron2 and HiFi-GAN, the system converts Kannada text into natural-sounding speech (Shanavas *et al.*, 2024). These models employ attention mechanisms, prosody modeling, and high-fidelity neural vocoders to improve pronunciation accuracy and speech naturalness.

Kannada-to-Braille Conversion: A (SVM) classifier is employed to map Kannada text to Braille, ensuring an accurate tactile representation of the language for visually impaired users.

Received: 07/04/2025 **Accepted:** 06/05/2025 **Published** : 21/05/2025

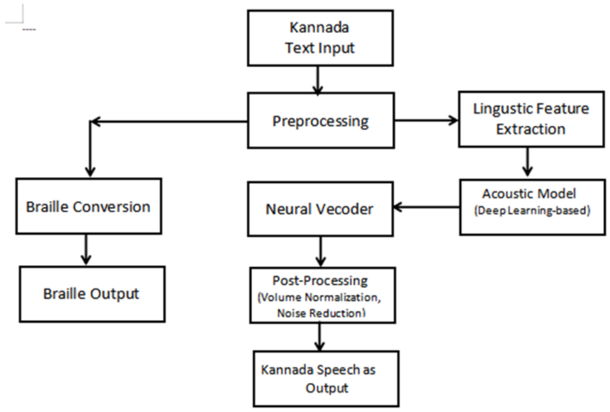

Figure1: Block Diagram of the integrated Kannada text-to-speech and braille conversion system

This study combines speech synthesis with tactile communication to improve accessibility for Kannada-speaking visually impaired individuals. The paper is organized as follows: Section II presents the literature review, Section III details the proposed system, Section IV discusses the results and evaluation, and Section V concludes with future enhancements.

Literature Survey

Advancements in assistive technologies for the visually impaired have significantly improved TTS and Braille transcription systems. Early methods primarily relied on rule-based and concatenative synthesis approaches, while recent research has explored deep learning techniques to enhance speech naturalness and transcription accuracy. Studies on corpus-driven TTS systems for Indian languages, such as Hindi and Telugu, have demonstrated efficient text conversion with support for multiple output formats and accent customization. This paper reviews existing research on text-to-speech synthesis and Braille transcription, exploring advancements in methodologies that enhance accessibility for visually impaired individuals:

Kaur *et al*. (2020) proposed a system for text-to-Braille and speech conversion using sequential mapping and deep learning techniques. A dedicated dataset was created to support the process. The implemented methods achieved a peak accuracy of 96% in converting text to Braille; however, additional enhancements are needed to optimize performance for real-time datasets.

Sahu *et al*. (2012) designed a TTS system for Indian languages like Hindi, Telugu, and Kannada, utilizing a corpus-driven synthesis approach. It supports multiple input formats and generates speech output in WAV and MP3 with customizable accents. The system also enables inter-language text conversion between Hindi and Telugu.

Dhananjaya *et al*. (2016) developed a TTS system primarily focused on English, with limited exploration of regional languages. Their proposed algorithm utilized speech coefficient concatenation for efficient synthesis, requiring minimal manual intervention. While minor pronunciation gaps were noted, the approach proved effective, with future enhancements aimed at improving speech fluency and evaluating performance on larger datasets.

Patel *et al*. (2025) highlighted that machine learning-based assistive technologies are transforming accessibility for individuals with sensory impairments by enabling real-time object detection, speech synthesis, and adaptive support. This study explores models like YOLO, SSD, and Deep Speech, emphasizing their potential to enhance navigation and communication while addressing ethical and personalization challenges.

AlSalman *et al*. (2021) developed a system for advanced Braille-to-text conversion, improving accessibility for visually impaired individuals. Their work leveraged DCNN models for multilingual Braille recognition, enabling accurate mapping of Braille cells to text across various languages, thereby enhancing communication.

Proposed system

The system aims to develop a unified Kannada TTS and Braille conversion framework to improve accessibility for visually impaired individuals. It leverages deep learning for speech generation and machine learning for Braille conversion, allowing Kannada text to be rendered in both auditory and tactile forms(Archana *et al*., 2019). The architecture of the system consists of five primary stages: dataset collection and preprocessing, text-to-speech synthesis using Tacotron2 and HiFi-GAN, Braille conversion using a SVM classifier, output generation, and performance evaluation (Ganga Gudi *et al*., 2022).

To ensure the precision and efficiency of both the speech synthesis and Braille conversion modules, the system relies on two essential datasets. The first dataset comprises a Kannada speech corpus, which is utilized to train the deep learning model for speech synthesis(Sana Shokat *et al*., 2020). This corpus includes over 10,000 text-audio pairs sourced from publicly available Kannada speech datasets and manually recorded samples. Prior to training, the dataset undergoes preprocessing steps such as text normalization, phonetic

Table 1: Kannada-to-Braille Mapping Dataset

Kannada Character	*Unicode*	*Braille Representation*
ಅ (a)	U+0C85	⠁
ಆ (aa)	U+0C86	⠜
ಇ (i)	U+0C87	⠊
ಈ (ii)	U+0C88	⠔
ಉ (u)	U+0C89	⠥
ಊ (uu)	U+0C8A	⠳

transcription, and audio enhancement techniques like noise reduction and silence trimming(Indira *et al*., 2010). The second dataset consists of a Kannada-Braille mapping set, containing 2,500-character mappings, as illustrated in Table 1.

This dataset is essential for training the SVM classifier, which performs the Kannada-to-Braille conversion (Anuradha *et al*., 2020).

Methodology

Kannada Text Cleaning and Normalization

The preprocessing of Kannada text is a crucial step to ensure consistency and accuracy in both speech synthesis and Braille conversion. This process involves multiple stages:

Character Standardization

Converts different forms of Kannada script into a standard format to handle variations in character encoding and representation.

Punctuation Handling

Removes unnecessary punctuation marks while preserving pauses that affect speech synthesis and improve the fluency of synthesized speech.

Diacritic Processing

Ensures proper representation of Kannada matras (vowel signs) and conjunct consonants, crucial for maintaining phonetic accuracy.

Tokenization and Phonetic Mapping

Word Tokenization

Segments the input text into words and syllables for efficient processing in both speech synthesis and Braille conversion.

Phoneme-Level Processing

Maps Kannada words to their phonetic equivalents to enhance pronunciation clarity in synthesized speech.

Sandhi Rule Handling

Incorporates morphological transformations in compound words to maintain the natural flow of speech and accurate Braille representation.

Speech Synthesis Using Deep Learning

The speech synthesis module leverages deep learning models to generate natural Kannada speech with high phonetic clarity and intonation accuracy.

Tacotron2: Text-to-Mel Spectrogram Conversion

Sequence-to-Sequence Model

Converts preprocessed Kannada text into mel-spectrograms using a recurrent neural network (RNN)-based approach.

Attention Mechanism

Ensures proper alignment between the input text and the corresponding audio output, improving pronunciation and fluency.

Prosody Control

Maintains natural rhythm and emphasis in speech synthesis to enhance intelligibility and user experience.

HiFi-GAN: Neural Vocoder for Waveform Synthesis

High-Fidelity Speech Generation

Converts generated mel-spectrograms into high-quality, realistic speech waveforms.

Low-Latency Processing

Enables real-time Kannada speech synthesis, making the system suitable for interactive applications.

Enhanced Speech Quality

Produces output with minimal distortion and accurate pitch variations, ensuring a natural listening experience.

Post-Processing for Speech Enhancement

Noise Reduction

Removes background noise and unwanted artifacts to enhance speech clarity.

Speech Rate Adjustment

Allows users to customize the speech speed according to their preferences.

Pause Optimization

Introduces natural pauses in speech to improve comprehension and fluency.

SVM-Based Braille Translation

Training Phase

- Converts Kannada text into feature vectors using One-Hot Encoding.
- Trains an SVM classifier to map Kannada characters to Braille representations.
- Uses a Radial Basis Function (RBF) kernel to achieve high classification accuracy.

Testing & Prediction

- Predicts the Braille representation of an input Kannada character using the trained SVM model.
- Outputs the Braille text in a standard format suitable for refreshable Braille displays and Braille printers.

Post-Processing for Braille Formatting

Word Wrapping

Adjusts the layout of Braille text for improved readability.

Spacing Adjustments

Ensures proper formatting to enhance tactile reading efficiency.

Translation Validation

Checks for mis-classified characters and applies corrective measures to enhance accuracy.

Output Generation

The proposed system provides two simultaneous outputs to enhance accessibility for visually impaired users.

Natural Kannada Speech Output

The Tacotron2 + HiFi-GAN model generates high-quality Kannada speech, which can be played through an audio output device, such as speakers or headphones.

Braille Representation Output

The Braille translation model, utilizing SVM, converts text into Braille, which can then be presented on a refreshable Braille display or printed using a Braille embosser.

The proposed system successfully integrates Kannada text-to-speech synthesis and Braille conversion, providing enhanced accessibility for visually impaired individuals. By leveraging deep learning techniques such as Tacotron2 and HiFi-GAN, the system generates high-quality Kannada speech with natural intonation and phonetic clarity. The SVM-based Braille translation module ensures accurate and efficient conversion of Kannada text into Braille. The preprocessing techniques, including text normalization, tokenization, and phonetic mapping, further enhance the performance of both speech synthesis and Braille translation(Sharavana *et al*., 2020). Additionally, post-processing steps such as noise reduction, speech rate adjustment, and Braille formatting improve the usability of the system (Sayeedi *et al*., 2024).

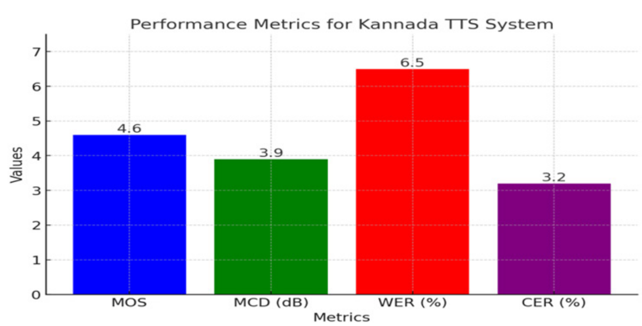

Figure 2: Performance Metric for Kannada TTS System

Figure 3: Confusion Matrix for Kannada Braille conversion

The integration of these components ensures a dual-mode output—natural Kannada speech and Braille text facilitating improved accessibility for individuals with visual impairments.

Results and Evaluation

The proposed Kannada TTS and Braille conversion system was evaluated using a 10,000-pair Kannada speech dataset for TTS and a 2,500-character Kannada-Braille mapping dataset for Braille translation. The TTS system, leveraging Tacotron2 and HiFi-GAN, achieved a mean opinion score (MOS) of 4.6/5, ensuring high-quality and natural speech synthesis. Additionally, the Mel cepstral distortion (MCD) score of 3.9 dB indicates precise phonetic mapping and prosody control, as shown in Fig 2.

The SVM-based Braille conversion model achieved 94.2% accuracy, demonstrating effective character-to-Braille mapping with minimal errors. The confusion matrix as shown in Fig 3 for the Kannada Braille conversion model, illustrating its classification performance. Future work will focus on enhancing real-time synthesis and incorporating deep learning for Braille translation to improve accessibility for visually impaired individuals.

Conclusion

The proposed system successfully integrates Kannada TTS synthesis and Braille conversion, enhancing accessibility for visually impaired individuals. The TTS component, implemented using Tacotron2 and HiFi-GAN, achieves a high Mean Opinion Score (MOS) of 4.6/5, ensuring natural and intelligible speech output. The Mel Cepstral Distortion (MCD) score of 3.9 dB further validates the phonetic accuracy and prosody control of synthesized speech. Additionally, the Braille conversion system, based on an SVM classifier trained on a dataset of 2,500-character mappings, achieves an accuracy of 94.2%, effectively translating Kannada text into Braille. Performance evaluations, including confusion matrices and graphical analysis, confirm the reliability of both components. The results indicate that this system can serve as an effective assistive technology, bridging the gap between auditory and tactile text accessibility in Kannada. Future work may explore improvements in prosody modeling, real-time speech generation, and adaptation to diverse Kannada dialects for further enhancing user experience.

Acknowledgment

We sincerely appreciate PES University, RR Campus, Bangalore, for providing essential resources and support for this research. Our heartfelt thanks go to our guide and mentors, Dr. Mallamma V. Reddy and Dr. Hanumanthappa M., for their invaluable guidance and insightful feedback throughout this study.

Conflict of Interest Statement

The authors declare that they have no conflicts of interest related to this research.

References

AlSalman, A., Gumaei, A., AlSalman, A., & Al-Hadhrami, S. (2021). A deep learning-based recognition approach for the conversion of multilingual Braille images. Computers, Materials & Continua, 67(3), 3847–3864. https://doi.org/10.32604/cmc.2021.015614

Anuradha, K. S., & Thelijjagoda, S. (2020). Machine translation system to convert Sinhala and English Braille documents into voice. 2020 International Research Conference on Smart Computing and Systems Engineering (SCSE), 7–16.

Archana, M., & Dhivya, S. (2019). Qualitative analysis framework for converting Braille to voice using image recognition. International Journal of Recent Technology and Engineering, 8(3), 5713–5717.

Dhananjaya, M. S., Krupa, B. N., & Sushma, R. (2016). Kannada text-to-speech conversion: A novel approach. 2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), 168–172.

Gudi, G., & Reddy, M. V. (2022). Conversion of Indian languages text into Braille script and audio file: A survey. International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), 13–18.

Indira, K., & Selvi, S. S. (2010). Kannada character recognition system: A review. arXiv preprint, arXiv:1001.5352. https://doi.org/10.48550/arXiv.1001.5352.

Kaur, P., Ramu, S., Panchakshari, S., & Krupa, N. (2020). Conversion of Hindi Braille to speech using image and speech processing. 2020 IEEE 7th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), 1–6.

Patel, P., Pampaniya, S., Ghosh, A., Raj, R., Karuppaih, D., & Kandasamy, S. (2025). Enhancing accessibility through machine learning: A review on visual and hearing impairment technologies. IEEE Access, 13, 33286–33307. https://doi.org/10.1109/ACCESS.2025.3539081

Rupanagudi, S. R., Huddar, S., Bhat, V. G., Patil, S. S., & Bhaskar, M. K. (2014). Novel methodology for Kannada Braille to speech translation using image processing on FPGA. 2014 International Conference on Advances in Electrical Engineering (ICAEE), Vellore, India, 1–6. https://doi.org/10.1109/ICAEE.2014.6838445

Sahu, L., *et al*. (2012). Hindi & Telugu text-to-speech synthesis (TTS) and inter-language text conversion.

Sayeedi, M. A., Osmani, A. M. M., & Farid, D. M. (2024). BrailleSense: Deep learning for Braille character classification. 2024 6th International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT), Dhaka, Bangladesh, 681–686. https://doi.org/10.1109/ICEEICT62016.2024.10534500

Shokat, S., Riaz, R., Rizvi, S. S., Khan, K., Riaz, F., & Kwon, S. J. (2020). Analysis and evaluation of Braille to text conversion methods. Mobile Information Systems, 2020, Article ID 3461651, 14 pages. https://doi.org/10.1155/2020/3461651

Shanavas, S. A. (2024). Malayalam text-to-speech conversion: An assistive tool for visually impaired people. Language in India, 24(4).

Sharavana, K., Sivasubramanian, E., Arpitha, N., Gowda, M. M., Naik, G. M., & Vijayalakshmi. (2020). Real-time text to Braille and audio converter. International Research Journal of Engineering and Technology, 7(7), 2869–2878.