International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

SpeakBraille: Revolutionizing Braille with Speech Synthesis

Poonam Shaylesh Lunawat[1], Ms. V.Anitha[2], Ms. D,Pavanikalyani[3], Ms. M.Nikitha[4]

[1,2,3,4]Computer Science and Engineering, BVRIT HYDERABAD College of Engineering for Women Hyderabad, India

[1]Poonam.lunawat20@gmail.com, [2]21wh1a05a0@bvrithyderabad.edu.in, [3]21wh1a05b8@bvrithyderabad.edu.in, [4]21wh1a05a7@bvrithyderabad.edu.in

Abstract

In an era of increasing technological advancement, accessibility remains a paramount concern. This project presents a robust and innovative solution for the visually impaired by harnessing the power of computer vision to seamlessly convert Braille into text and voice. The Braille to Text and Voice Conversion System (BTVCS) serves as a transformative tool that bridges the gapbetween tactile and auditory communication, empowering individuals with visual impairments to access written information independently. The core of the system lies in its ability to accurately recognize Braille characters through image processing and computer vision techniques. Utilizing state-of-the-art machine learning algorithms and deep neural networks, BTVCS interprets Braille patterns with precision, ensuring minimal error rates. The extracted Braille text is then seamlessly converted into natural language text, opening up a world of printed content to the visually impaired. Additionally, BTVCS features a robust text to- speech (TTS) engine that transforms the converted text into high quality, natural-sounding voice output. Users can select from a range of voices and settings to personalize their experience. This dynamic TTS capability extends beyond mere translation; it provides an immersive auditory experience that conveys tone, context, and emotion, making the content more engaging and informative

Keywords: Accessibility, Visually Impaired, Braille, Computer Vision, Text-to-Speech (TTS), Machine Learning, Deep Neural Networks, Image Processing, Natural Language Text, Voice Output, Personalization, Convolutional Neural Network

1. Introduction

Visually and vocally impaired individuals face numerous challenges in their daily lives, ranging from accessing written content to navigating social interactions. While the Braille system, developed in the 19th century by Louis Braille, has served as a cornerstone for enabling literacy among visually impaired individuals, it is not without limitations. The six-dot Braille system allows the representation of characters and words, yet reading and comprehending extensive material through touch remains a time-intensive process. This challenge is particularly acute in developing regions where access to Braille resources and training is scarce.

Globally, approximately 2.2 billion people suffer from some form of vision impairment, with 36 million being completely blind. Among them, a significant number also face vocal impairments, further complicating their ability to communicate effectively. Despite technological advancements, a large segment of this population relies on traditional Braille, which, while effective, does not fully address the modern needs of accessibility and speed.

Speak Braille aims to bridge this gap by introducing a novel integration of advanced image processing and speech synthesis technologies. This project proposes a system capable of converting Braille documents into audible speech, thereby enhancing accessibility for visually impaired individuals. The system leverages cutting-edge image recognition algorithms to detect Braille patterns from scanned documents or images and translates them into speech in real-time. By doing so, Speak Braille not only reduces the reliance on tactile reading but also empowers users to access information faster and more efficiently.

This innovative solution envisions a future where visually impaired individuals can independently engage with vast repositories of knowledge, participate in educational endeavors, and perform everyday tasks with greater autonomy. Through Speak Braille, the transformative potential of combining traditional Braille literacy with modern auditory technology is unlocked, fostering inclusivity and empowerment in an increasingly digital world.

2. LITERATURE REVIEW

Kishore et al. [1] proposed a cost-effective and efficient system for converting Braille script to voice. The system utilizes a Braille keyboard for data input and processes the input using Arduino. This method provides a robust platform for communication between visually and hearing-impaired individuals. However, it is constrained by its dependency on basic hardware components, which limits functionality and scalability. The system lacks real-time capabilities and multilingual support, which are crucial for broader applicability.

Deep learning techniques have significantly advanced Braille text recognition. For instance, methods employing Stacked Denoising Auto Encoders (SDAE) and image segmentation, as explored by Shokat et al., have achieved remarkable accuracy in Braille recognition, providing comprehensive guidelines for system development [2]. Their work highlights the potential of deep learning methods to provide comprehensive guidelines for developing robust Braille-to-text systems. Despite their merits, these systems often rely on screen-location-specific input or require memorization of gestures, imposing additional cognitive burdens on users. Additionally, the absence of multilingual features restricts their usability in diverse linguistic settings.

Recent studies, including those by Kumar et al., have explored Optical Character Recognition (OCR) technologies integrated with text-to-speech (TTS) systems to convert printed Braille into speech outputs. A notable implementation achieved 85% accuracy in real time processing, demonstrating the potential of such systems in improving accessibility [3]. However, the sensitivity to input quality, such as document wear and tear, remains a critical limitation. Furthermore, the system’s focus on English limits its global adoption. Another innovation involved the use of OpenCV for image preprocessing and Braille character recognition through contour detection and binary-to-decimal mapping, as demonstrated by Anuradha and images, provided 85% overall accuracy, improving to 95% with optimal input quality. The reliance on high-quality inputs underscores the need for robust preprocessing techniques. Thelijjagoda [4]. This approach, tested on high resolution brightness, and position shifts to enhance generalization. Achieving an accuracy of 96.30%, the system converts Braille characters into English text and subsequently generates audio using Google Text-to-Speech While highly effective for accessibility, the system's reliance on preprocessed data and a specific dataset may pose challenges in handling more diverse or noisy inputs.

A novel approach by AlSalman and Gumaei employed Deep Convolutional Neural Networks (DCNNs) for multilingual Braille image recognition, achieving over 98% accuracy in recognizing diverse Braille symbols [6]. This highlights the potential for developing multilingual solutions. Nevertheless, the computational real-world deployment. Hybrid systems integrating K-Means Clustering and heuristic algorithms for Kannada Braille script translation demonstrated innovative methodologies, as detailed by Sharma et al. [7]. Despite bridging communication gaps for specific languages, these systems struggle with processing hand-punched Braille data, underscoring the need for versatile algorithms.

A Novel Data Independent Approach for Conversion of Hand Punched Kannada Braille Script to Text and Speech proposes a system to convert hand-punched Kannada Braille to text and speech using image processing, k-means clustering, and Festival TTS, but faces accuracy challenges due to noise and inconsistencies in hand-punched Braille[8].

S.No	Models or Techniques used	Metrics and Result	Drawbacks
[1]	Arduino- Based System	Affordable; converts Braille input into voice and text	Affordable; converts Braille input into voice and text
[2]	OCR & Touch Screen- Based Input	Efficient; evaluates methods like scanned input and gesture- based touch input	Steep learning curve; location dependent inputs; reliance on gestures for validations.
[3]	OCR and TTS.	High accuracy converting Braille to ASCII text and vice versa; Python based	Limited to ASCII; no support for native languages.
[4]	OCR and OpenCV Image Processing.	85.3%accuracy, upto 95% for high quality inputs	Lower accuracy for noisy or poor quality inputs, Language specific limitations
[5]	Personal perspective on assistive tools	Highlights independence through Braille-enabled notetakers, talking gadgets, and screen readers	High costs; limited accessibility to digital platforms; dominance of sighted perspectives in design
[6]	OpenCV, Image Processing and Speech Synthesis	Effective for beginner level Braille learning; functional Android app	Limited to specific devices; lacks scalability.
[7]	KNN Algorithm, RaspberryPi and TTS	Cost-effective; converts Braille cells to speech.	Hardware dependence, suboptimal for large-scale applications
[8]	K-means Clustering, Gaussian Filtering and TTS.	85%OBR accuracy; MOS score of 4.54 for audio naturalness	Struggle with inconsistent Braille quality, limited to kannada braille
[9]	ANN, OCR and Image Processing	Converts Braille to text and sound; accurate pattern recognition	High computation al requirement s; dependence on specific hard ware/softwa re setups
[10]	BT-CNN(Braille Text Convolutional Network)	96.3% accuracy for braille-to-text conversion utilizes Google TTS for audio output	Limited dataset size(1560 images),dependency on specific datasets and pretrained models

Table 1 Comparison of existing works

complexity and resource demands of DCNN models pose challenges for Raspberry Pi-based solutions have demonstrated significant potential in converting text to audio for visually impaired individuals. For example, the "Vision Voice" [9] system utilizes a Raspberry Pi microprocessor, integrating a camera module to capture text images, Optical Character Recognition (OCR) software for text recognition, and a Text- to-Speech (TTS) engine to generate audible output. This cost-effective device provides an accessible and efficient assistive tool for text-to-audio conversion. However, the system's reliance on clear text input and absence of advanced contextual understanding may limit its adaptability in complex real world scenarios or multilingual environments. Deep learning-based models have shown great promise in Braille to- speech conversion. For instance, the "Braille to Speech Conversion Using BT-CNN" [10] system employs a Braille Text Convolutional Neural Network (BT-CNN) trained on a 28x28 grayscale Braille character dataset from Kaggle with augmentation techniques such as rotation.

3. RESEARCH GAP

Despite advancements in the field of Braille-to-audio conversion systems, several gaps remain unaddressed, hindering their scalability, accessibility, and practical implementation in real-world.

A. Hardware Dependence

Many existing systems, such as those proposed by Kishore et al. [1] and Ramiati et al. [7], rely on specialized hardware configurations, including platforms like Arduino or software environments such as MATLAB. These dependencies increase the overall cost of implementation and reduce the portability of such solutions, especially in resource-limited regions. Developing hardware-agnostic systems that can run seamlessly across various devices without requiring specialized setups is crucial for enabling scalable, costeffective adoption.

B. Dataset Limitations

A common drawback in most systems, including those by Shokat et al. [2] and AlSalman et al. [6], is their reliance on small and controlled datasets that often lack diversity. These datasets typically fail to account for variations in language, font styles, document quality, and real-world conditions, resulting in limited generalizability of the models. To address this, future work must focus on curating large-scale, diverse datasets that include multilingual text and real-world Braille documents, thereby improving the robustness and adaptability of these systems.

C. Real-Time Processing Challenges

Achieving real-time Braille-to-audio conversion remains a significant technical hurdle for many existing solutions, including those by Ramiati et al. [7]. Computational inefficiencies, coupled with hardware limitations, often lead to delays and high latency in processing. While localized systems, such as those leveraging Raspberry Pi setups, show promise, their scalability to larger networks and broader real-time applications is limited. For practical deployment, systems must achieve near-instantaneous processing while maintaining high accuracy across diverse operating environments.

D. Lack of Multilingual Support

Most existing systems are designed for single language Braille recognition, limiting their applicability for linguistically diverse user groups. For instance, DCNN-based systems have shown high accuracy in recognizing specific languages like Kannada Braille [8], but extending these models to support multiple languages simultaneously remains an unresolved challenge. Future research must focus on developing multilingual systems that can seamlessly interpret Braille in various languages while maintaining high accu racy and performance.

E. Ethical and Privacy Concerns:

The increasing use of advanced natural language processing (NLP) techniques, such as BERT and LSTM models, introduces ethical and privacy concerns due to their high computational demands and potential exposure of sensitive data. Tan et al. [13] demonstrated the efficacy of hybrid deep learning approaches but acknowledged that their computational intensity hinders scalability for large-scale or distributed deployments. To address these challenges, it is essential to optimize such models for efficiency while implementing secure mechanisms to protect user privacy during data processing.

F. Affordability and Accessibility

High implementation costs, driven by the need for specialized hardware and computational resources, create significant barriers to adoption, particularly in resource-constrained environments. While cost effective alternatives, such as Raspberry Pi-based systems, have been explored, these often sacrifice performance, functionality, or scalability. Future solutions must strike a balance between affordability and performance, ensuring that high quality Braille-to-audio conversion is accessible to a broader audience, including underserved communities.

4. Conclusion

The project on Braille to text and speech conversion using Optical Braille Recognition (OBR) enhances accessibility for visually impaired individuals. By converting Braille into text and high-quality voice output, it bridges tactile and auditory communication, allowing independent access to written information. The system uses advanced image processing and pattern recognition to ensure accurate Braille recognition with minimal errors. It also includes a dynamic text-to-speech engine that adds tone and context to the speech. This project aims to improve the quality of life for the visually impaired, promoting accessibility and innovation in this field.

5. Future work

Future work should focus on developing a portable, user- friendly wearable device with a camera for real world Braille text capture, ensuring durability in various environments. A mobile app compatible with Android and iOS could offer offline Braille- to-speech conversion, voice customization, and enhanced accessibility. Integration with screen readers, tactile feedback devices, and smart home systems would provide multi-sensory experiences. To ensure affordability, partnerships with non-profits could subsidize distribution. AI-driven contextual understanding and educational modules can enhance learning, while open-sourcing parts of the project could foster global collaboration. Long-term, the system can expand to tactile graphics and partnerships for digitizing Braille materials.

References

1. Kishore, K. K., Prudhvi, G., Naveen, M. (2017). “Braille script to voice conversion”. 2017 International Conference on Computing Methodologies and Communication (ICCMC). doi:10.1109/ICCMC.2017.8282637
2. 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, 1–14.

doi: 10.1155/2020/3461651. Downloaded on 2020-08-26

3. Kumar, R. Sathish, et al. ”EXTRACTING TEXT FROM IMAGE AND BRAILLE TO SPEECH CONVERSION.” https://www.researchgate.net/profile/Sathish-Rajamani3/publication/371476727EXT RACT INGTEXTFROMIMAGEANDBRAILLETOSP EECHCOT EXT−F ROM−IMAGE−AND−BRAILLE−T O−SP EECH−CONV ERSION.pdfhttps : K.S.AnuradhaandS.T helijjagoda, ”MachinetranslationsYstemtoconvertSinhalaandEnglishBr16, doi : https ://doi.org/10.1109/SCSE49731.2020.931302010.1109/SCSE49731.2020.9313020
4. K. S. Anuradha and S. Thelijjagoda, "Machine translatiosystem to convert Sinhala and English Braille documents into voice," 2020 International Research Conference on Smart Computing and Systems Engineering (SCSE), Colombo, Sri Lanka, 2020, pp. 7-16, doi:10.1109/SCSE49731.2020.9313020.
5. B. K. Rajan and V. Anjitha, ”Braille code conversion to voice in malayalam,” 2017 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2017, pp. 0710-0714, https://doi.org/10.1109/iccsp.2017.8286452
6. Abdulmalik AlSalman1, Abdu Gumaei ”A Deep Learning-Based Recognition Approach for the Conversion of Multilingual Braille Images”https://www.researchgate.net/profile/SuAlHadhrami/publication/350022456ADeepLearning−BasedRecognitionApproachfortheConversiDeep−Learning−Based−Recognition−Approach−for−the−Conversion−of−Multilingual−Braille−Images.pdf https://www.researchgate.net/prof ile/Suheer − Al − HaAbhishekSharma, SumanDevi, J itendraKumarV erma.HandBrailleBookReaderusingRaspbery
7. Ramiati, Siska Aulia, Lifwarda and S.Ningrum Nindya Satriani. Recognition of Image Pattern To Identification Of Braille Characters To Be Audio Signals For Blind Communication Tools .International Conference on Applied Sciences, Information and Technology 2019 1-3 November 2019, Padang, Indonesia, https://doi:10.1088/1757-899X/846/1/012008doi:10.1088/1757
8. K. Thopate, K. Amrutkar, T. Kasliwal, S. Karvir and O. Kumbhar, ”Vision Voice: A Raspberry Pi-Based Text-to-Audio Converter for the Visually Impaired,” 2024 5 International Conference on Emerging Innovations and Advanced Computing (INNOCOMP), Sonipat, India, 2024, pp. 231-234, https:// doi: 10.1109/INNOCOMP63224.2024.00045 .
9. K. Thopate, K. Amrutkar, T. Kasliwal, S. Karvir and O. Kumbhar, ”Vision Voice: A Raspberry Pi-Based Text-to-Audio Converter for the Visually Impaired,” 2024 5 International Conference on Emerging Innovations and Advanced Computing (INNOCOMP), Sonipat, India, 2024, pp. 231-234, https:// doi:10.1109/INNOCOMP63224.2024.00045
10. Shreekanth, T., Deeksha, M. R., Kaushik, K. R. (2018). ”Hand Punched Kannada Braille Script to Text and Speech”. International Journal of Image and Graphics, 18(2), 1850010 https://doi.org/10.1142/S0219467818500109doi.org/10.1109/ICACCT50842.202011
11. Aditya Upadhye, Tharunnesh R, Jeslin Shanthamalar J, C A Daphine Desona Clemency and N S Usha.”Braille to Speech conversion using BT-CNN”.International Journal of Microsystems and IoT".