Generative AI for Braille Text Summarization

Smaran Bharadwaj
smaranrb.btech22@rvu.edu.in
School of Computer Science Engineering (SOCSE)
RV University, RV Vidyanikethan Post
8th Mile, Mysuru Road, Bengaluru – 560059

Srivijayalakshmi AT
srivijayalakshmiat.bsc22@rvu.edu.in
School of Computer Science Engineering (SOCSE)
RV University, RV Vidyanikethan Post
8th Mile, Mysuru Road, Bengaluru – 560059

***Abstract*—This study examines the urgent need for accessible reading aids for individuals with visual impairments by developing a novel method that translates scanned text documents into Braille. Utilizing Tesseract's Optical Character Recognition (OCR), the approach begins with image scanning and preprocessing to ensure high text recognition accuracy. The extracted text is then converted into Braille following the International Council on English Braille's (ICEB) guidelines, significantly increasing the availability of historical documents and educational materials for the blind. This research not only supports visually impaired readers and students, granting them autonomy and inclusion in accessing essential literature, but it also highlights the reliability and versatility of advanced OCR technologies, enhancing the efficiency of Braille translation. Furthermore, the study emphasizes the importance of integrating user feedback to continuously improve assistive technologies, ensuring they meet the evolving needs of the visually impaired community. Through these advancements, the paper lays the groundwork for further innovations in inclusive technology, ultimately promoting broader accessibility and equity in education and communication.**

***Index Terms*—GenAI - Generative AI, NLP - Natural Language Processing, BERT - Bidirectional Encoder Representations from Transformers**

I. Introduction

The Braille script plays a vital role for the visually impaired, especially in educational and professional settings where access to complex information is required. Modern Braille is divided into Grade 1 and Grade 2 systems; Braille Grade 1 is straight one-to-one character correspondence, and Grade 2, or ”contracted” Braille, does use shortcuts to save writing space; it compresses the volume of text relative to the printed version up to 20–30 percent. In spite of further growth of TTS use and other audio-based method, Braille remains especially important for those who become audibly or visually challenged. This versatility points to the continued need for generative AI in the development of Braille conversion technologies, which would make more contexts accessible and empower the visually impaired. Hersh et al. focus on multimodal conversion technologies as means of filling accessibility gaps for people who are blind or have low vision, placing their work at the core as electronic text. Speech to Text (STT), Text to Speech (TTS), Braille to Text (BTT), and Text to Braille (TTB) are a few technologies that are instrumental in the access of information digitally by visually impaired. However, challenges do still exist with such technologies. There is still a large scope for further development and optimization, especially as these technologies evolve to meet the diverse accessibility needs.
Bendel note that generative AI can be a game-changer in the improvement of the well-being of the visually impaired. Some of the technologies include the Be My Eyes app, which integrated OpenAI's GPT-4 model in 2023, and allowed users to capture images of their surroundings and obtain detailed descriptions of objects, environments, or texts. This capability gives the users much more independence and freedom to perform their daily routines without relying so much on their sighted volunteer or relatives for help. Generative AI models create a different kind of assistance technology based on personalized context-aware feedback, which aids users in easy navigation of their surroundings and access to information and enhanced participation in other activities not possible without seeing. [4]

II. Backround

In recent years, electronic text has become a key accessibility tool, especially for people who are blind or visually impaired. Advances in assistive technology now make it possible to convert electronic text into a range of formats, such as large print, Braille, and spoken audio, significantly improving access to information. This study explores how various conversion technologies enable electronic text to serve as a fundamental medium for information accessibility. These technologies include text-to-speech (TTS), Braille-to-text (BTT), text-to-Braille (TTB), and speech-to-text (STT). With these tools, electronic content can be read aloud, displayed in Braille, shared remotely, or shown in large print for greater accessibility. Human communication naturally combines different signals, like speech, gestures, and visual cues, to improve understanding. This reliance on visual information can pose challenges for blind individuals, making it essential to convert content into auditory and tactile formats. Although Braille remains a traditional and valuable tool, particularly for those who are deafblind, newer speech-based methods are also gaining popularity. For blind and visually impaired users, having access to multiple formats—text, audio, and tactile—offers flexibility and ensures a better, more comprehensive way to access information, especially for language-based content. This background provides the basis for examining the technological capabilities, practical uses, and scientific foundations of these conversion systems in the chapter. [8]
In recent years, electronic text has become an essential tool

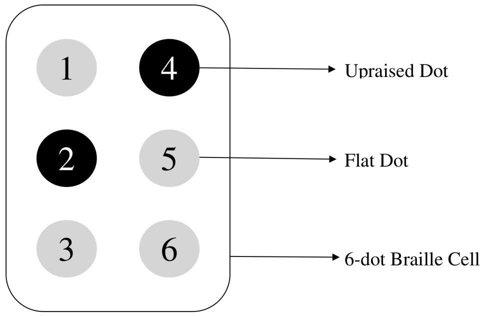


Figure 1.1 - Visualisation of a Braille Cell

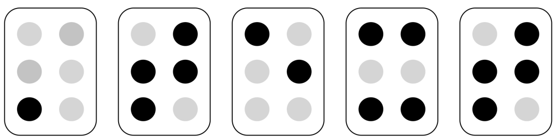


Figure 1.2 - The string “Text” written in Braille script using 5 Braille cells

for accessibility, particularly for people who are blind or visually impaired. Advances in assistive technology have made information more accessible by enabling electronic text to be converted into various formats, such as large print, Braille, and spoken audio. This study explores how these conversion technologies make electronic text a core medium for accessible information. Key technologies include text-to-speech (TTS), Braille-to-text (BTT), text-to-Braille (TTB), and speech-to-text (STT). Once transformed, electronic content can be read aloud, displayed in Braille, shared remotely, or presented in large type for easier reading.Human communication relies on multiple signals—speech, gestures, and visual cues—to enhance understanding. This dependence on visual information can create challenges for blind individuals, highlighting the importance of converting content into auditory and tactile formats. Braille remains a traditional tool, especially valuable for people who are deafblind, although newer speech-based methods are now emerging. For blind and visually impaired users, having access to multiple formats—text, audio, and tactile—provides flexibility in accessing information, especially language-based material. This context sets the stage for the chapter’s focus on the technological capabilities, practical applications, and scientific foundations of these conversion systems. [4]

III. Inputs to the Model

The Optical Arabic Braille Recognition (OBR) system begins by placing a Braille document on a standard flatbed scanner to capture its image. This method doesn’t require specialized equipment, making it accessible and affordable by using commonly available scanners. Once the document is scanned, the captured image undergoes several preprocessing steps to enhance clarity and accuracy in identifying the Braille dots.First, the image is converted to a binary format, which helps separate the Braille dots from the background. Then, edge detection is applied to define the boundaries of each Braille cell, making it easier for the system to identify the exact location of each dot. A hole-filling technique is used next to correct any flaws in the embossing, ensuring that partially scanned or incomplete dots are recognized as full dots. Noise reduction algorithms are then applied to remove any unnecessary details that might interfere with dot recognition.With preprocessing complete, the OBR system accurately locates the raised dots in each Braille cell and matches them against a database of Arabic Braille patterns. Each identified Braille cell is assigned a unique decimal code based on its dot pattern, which forms the foundation for constructing words and sentences. This structured approach to data entry ensures high recognition accuracy, enabling smooth and efficient transcription into text and audio formats. This research study introduces an Optical Arabic Braille Recognition (OBR) system designed to improve accessibility for people with visual impairments by converting Arabic Braille text into readable text and audio output. Originally developed by Louis Braille in 1821, the Braille system is widely used globally and is based on six-dot characters, where each pattern represents a specific letter or symbol. However, current OBR systems lack robust support for Arabic Braille, which this study aims to address by offering a new method for precise transcription and recognition.The proposed system uses a detailed scanning process that includes image capturing and preprocessing techniques like edge detection, binary conversion, and noise reduction to ensure clear and accurate dot recognition. A specialized algorithm extracts the precise dot positions, generating unique decimal codes for each character, which allows the system to build coherent words and sentences. The study also examines how lighting conditions affect dot recognition accuracy, addressing issues like shadows and inconsistent illumination with advanced image processing techniques.The system boasts a high recognition accuracy rate of over 99 percent and an average processing time of 32.6 seconds per page, demonstrating its efficiency across various environments. Beyond its immediate applications in education and communication, this system represents an important resource for individuals and organizations that rely on Braille. By bridging the gap in Arabic Braille recognition, this research contributes significantly to the usability, reliability, and practicality of OBR technology. [2]

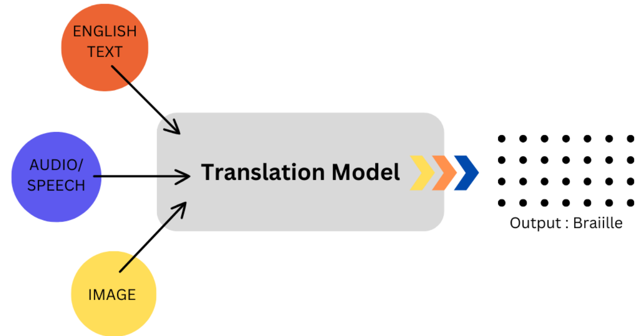
Our comparative evaluation study focused on two different braille input methods designed for blind individuals: the Swift Braille keyboard, which utilizes a traditional button layout, and BrailleEnter, a gesture-based input system. We conducted the study with six skilled blind volunteers to assess both methods across several performance criteria, including usability, error rates, and typing speed.The results showed that the Swift Braille keyboard was significantly faster, with an average typing speed of 4.37 words per minute (WPM) compared to BrailleEnter’s 2.45 WPM. However, BrailleEnter had a much lower error rate, averaging 7.8

percent compared to 28.96 percent for the Swift Braille keyboard. This suggests that while BrailleEnter is slower, it may provide higher accuracy in typing.Usability scores from the revised Lewis questionnaire indicated that users of BrailleEnter were more satisfied overall. Participants were able to quickly adapt to both methods thanks to their existing knowledge of Braille patterns. However, they noted that the multiple interactions required by BrailleEnter limited their speed.Based on these findings, we recommend design improvements for both input methods. For BrailleEnter, exploring alternative gesture movements could help increase typing speed. Meanwhile, optimizing the layout of the Swift Braille keyboard could reduce error rates. This study highlights the diverse preferences and needs of blind users when interacting with touchscreen devices, emphasizing the ongoing need for innovation in accessible technology. [5]

This paper addresses the critical need for accessible reading aids for people with visual impairments by presenting a new method that converts scanned text documents into Braille. Using Tesseract's Optical Character Recognition (OCR), the process begins with image scanning and preprocessing steps, such as grayscale conversion and adaptive thresholding, to ensure high text recognition accuracy. Once the text is extracted, it's translated into Braille based on the standards of the International Council on English Braille (ICEB). This approach facilitates the conversion of historical documents and educational materials into Braille, opening up valuable resources that were previously inaccessible to the blind and visually impaired. A key practical application of this research lies in its support for visually impaired readers and students, providing them with greater autonomy and inclusion in accessing essential books and materials. The paper also highlights the reliability and flexibility of modern OCR technologies, which enhance the effectiveness and efficiency of Braille translation. Beyond educational materials, the findings and methods outlined in this study lay a strong foundation for future advancements in assistive technology. Potential applications include improving general communication tools and adapting technologies to meet the needs of other disabilities, underscoring the paper's role in promoting inclusivity, education, and accessibility across diverse fields. [3].

IV. Current Status

Herman proposes an online Text to Braille translator application that is developed as an advancement over previous systems that usually have fewer translation capabilities and access limitations for the users. It supports up to 1,000 characters per session of translation. The proposed application is rules-based with a mapping of text onto Braille using pre-defined translation tables that support Grade 1 and Grade 2 Braille. Grade 1 Braille makes use of one-to-one letter mapping, whereas Grade 2 Braille uses contractions for common words or letter combinations. The



application uses PHP in the Windows environment, over MySQL, enhancing scalability and easy manageability of the database. The development process was organized on the lines of waterfall methodologies, which subdivide the entire project into phases such as requirement gathering, design, implementation, testing, and then to maintenance. However, the translator has an inherent limitation in its inability to process linguistically complex contractions or abbreviations across languages as well as cultural variations within Braille. Despite that, the system has provided an easy and inexpensive way of developing Braille texts, thus bringing forward a step in reducing the information accessibility gap for visually impaired people. [6]

Adnin and Das observed that GenAI tools are used by the visually impaired to generate content and search information. Participants reported that there were immense accessibility problems in mainstream GenAI interfaces such as ChatGPT which are difficult to use as there are many unlabeled buttons and navigation support by screen readers. Moreover, participants also found difficulties through hallucinations of GenAI responses, which mostly occurred with the tool like Be My AI, giving misleading visual descriptions. The authors further commented that blind users have different mental models about how GenAI works. Most of the participants imagined the GenAI tools as something more of a "knowledge database" or a "super search engine," though others formed more skeptical perspectives, regarding the GenAI as a "word-generating machine," creating answers according to some predicted words instead of proper comprehension. This reliance on GenAI introduces a risk of amplifying biases or misinformation, as visually impaired users often trust AI-generated information to a greater extent due to the extensive detail provided in responses. This research highlights both the potential benefits and the need for accessible design improvements in GenAI to mitigate risks and optimize usability for blind users. [1]

Yamuna et al. emphasize the significant gap in accessible and efficient summarization options for visually impaired people, especially for long documents in Braille. There are very few tools available to suit the specific needs of the visually impaired users. The authors propose a new method using a BERT-based extractive summarizer. This approach would help them condense large documents to more manageable Braille summaries. Using the BERT model, which has showed high performance in tasks in NLP, it is possible to extract important content from long texts. Concise and relevant information represented in Braille makes access and understanding easier for visually impaired readers. This would serve a dual purpose of condensing the content and making it available in Braille, thus greatly enhancing the access to information, save time, and empower readers with visual impairments to read a wider range of materials. The authors comment on the shortage of good quality translation options from English to Braille and the authors have, therefore, developed specific interconversion libraries to fulfill this shortage, thereby smoothing out the translation processes for researchers and developers. [10]

Blind individuals face significant challenges across mobility, information access, and social inclusion. Bendel et al. discuss how advancements in AI have provided targeted solutions, like screen readers, Braille displays, and now visual recognition tools that describe images and texts in real-time. Be My Eyes, with its GPT-4 integration, expands on this by helping users handle complex tasks like reading labels or navigating new areas independently. While generative AI solutions are not flawless and require careful ethical considerations, including privacy, data security, and potential dependencies, they represent a promising frontier in assistive technology, aligning closely with the societal need for inclusive solutions that empower visually impaired people in a world that is highly visual in nature. [4]

V. Future Scope

Joshi and Katyayan identified a scope of potential improvement in the case of India, with a specific need for a text-to-Bharti Braille system to benefit visually impaired readers of Indian languages. Using a hybrid approach from rule-based methods to LSTMs for converting written text to Bharti Braille ,this model can be expanded so that regional languages become accessible to the blind through reading and study materials in native languages, making education more inclusive. Bharti Braille for various languages will break massive information barriers because the visually challenged will be able to access content in different Indian scripts, from Devanagri to Gurumukhi. The translation model sets the groundwork for further innovations toward serving multiple regional contexts of India. Even introducing more languages in the system and perfecting its efficiency with cutting-edge machine learning techniques could bring full digital and physical accessibility of the space to the vision-impaired individual. This could be a mobile application and digital library that would support Bharti Braille, so the user could access literature, news, and educational content in his or her language and format of choice. Therefore, this technology marks a big step toward closing accessibility gaps for the visually impaired and provides a scalable solution to adapt India's multilingual landscape into Braille.

Sridhar et al. identified a large scope for the system in enabling digital access to visually challenged people by using it for processing video content into Braille-readable summaries. There are two phases to the system process: summarizing video content as text via automatic speech recognition (ASR) and computer vision followed by the translation of the summarized text to the Braille script. This innovation is likely to very much open multimedia content to users with any degree of blindness. Translating video into a tactile representation enables a wide variety of content types, such as videos of education, broadcasts, and others that will give such access to the information they had otherwise been barred from gaining. The system can be considered to have a future of becoming a real-time on-demand solution that assists users who are visually impaired to watch live or prerecorded videos independently. Further features, such as AI-driven contextual understanding and Braille refreshable displays, will make the system more versatile for various content styles. This technology would extend to regional languages to make video summarization more accessible, thus allowing the visually impaired of various linguistic settings around the country and more globally to access multimedia content in braille [9]

VI. Conclusion

This study sheds light on the potential and ongoing challenges faced in developing solutions for blind and visually impaired individuals within the realm of spoken language technology (SLT). While it recognizes the current limitations of the technology, it emphasizes that a comprehensive understanding of SLT principles is crucial for maximizing its capabilities. Notably, SLT programs have the ability to enhance accessibility for visually impaired users by addressing gaps in text and Braille conversion, which are vital for effective communication and information access. However, significant hurdles remain, particularly concerning system reliability and performance consistency in real-world environments.Output devices, such as embossers and Braille displays, play a critical role in the user experience, with usability being heavily influenced by the choice of device. Durable outputs, like Braille embossers, provide long-term access, while more temporary displays offer flexibility but may lack longevity. This dependence on device types indicates the need for tailored solutions to achieve optimal accessibility. For example, research suggests that the usability of speech-based kitchen devices—a key performance indicator in assistive technology—declines significantly when recognition accuracy falls below 85The report highlights a gap between the advancements in word recognition and synthesis quality and the systems available commercially. Comparative reviews of speech systems are rare, primarily due to proprietary restrictions that limit access to performance data, which could otherwise inform improvements across applications. While there is minimal implementation of these technologies for visually impaired individuals who could greatly benefit, many of the innovations remain confined to research laboratories.The findings underscore the importance of bridging the gap between research and industry to enhance accessibility, especially concerning the use of generative AI in translating Braille text. Generative AI models have the potential to improve translation quality significantly and could transform the creation of real-time Braille translations across various devices. Future SLT research should prioritize increasing user acceptance through enhanced accuracy, utilizing AI to address performance inconsistencies, and expanding device compatibility to ensure reliable and widespread accessibility. With ongoing research, generative AI-powered SLT solutions could overcome current challenges and better meet the functional needs of blind individuals. [7]

REFERENCES

[1] Rudaiba Adnin and Maitraye Das. "i look at it as the king of knowledge": How blind people use and understand generative ai tools. *people*, 16(54):92, 2024.

[2] Saad D Al-Shamma and Sami Fathi. Arabic braille recognition and transcription into text and voice. In *2010 5th Cairo International Biomedical Engineering Conference*, pages 227–231. IEEE, 2010.

[3] Mrim Alnfiai and Srini Sampali. An evaluation of the brailleenter keyboard: an input method based on braille patterns for touchscreen devices. In *2017 international conference on computer and applications (ICCA)*, pages 107–119. IEEE, 2017.

[4] Oliver Bendel. How can generative ai enhance the well-being of blind? In *Proceedings of the AAAI Symposium Series*, volume 3, pages 340–347, 2024.

[5] Joshua L Dela Cruz, Jonaida Angela D Ebreo, Reniel Allan John P Inovejas, Angelica Romaine C Medrano, and Argel A Bandala. Development of a text to braille interpreter for printed documents through optical image processing. In *2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, pages 1–6. IEEE, 2017.

[6] Nurafni Herman. *Text to Braille Translator*. PhD thesis, UMP, 2013.

[7] Rüdiger Hoffmann. Speech, text and braille conversion technology. In *Assistive Technology for Visually Impaired and Blind People*, pages 497–554. Springer, 2008.

[8] Sana Shokat, Rabia Riaz, Sanam Shahla Rizvi, Khalil Khan, Farina Riaz, and Se Jin Kwon. Analysis and evaluation of braille to text conversion methods. *Mobile Information Systems*, 2020(1):3461651, 2020.

[9] Bodanapu Sridhar, Gourishetti Saivishnu, Varla ManiShanker, D Dhana Lakshmi, and Shanmugasundaram Hariharan. Summarization of video into text and text to braille script. In *2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS)*, volume 1, pages 1–5. IEEE, 2024.

[10] K Yamuna, V Shriamrut, Drishti Singh, Vaishnavi Gopalasamy, and Vivek Menon. Bert-based braille summarization of long documents. In *2021 12th international conference on computing communication and networking technologies (ICCCNT)*, pages 1–6. IEEE, 2021.