International Journal of Research in Special Education

E-ISSN: 2710-3870
P-ISSN: 2710-3862
IJRSE 2025; 5(1): 202-206

Journal's Website
Received: 11-03-2025
Accepted: 19-04-2025

Gabriel de Souza Reis
Student, Department of Informatics, Federal Institute of Education, Science and Technology of São Paulo (IFSP), Bragança Paulista, SP, Brazil

Igor Ferreira de Oliveira
Student, Department of Informatics, Federal Institute of Education, Science and Technology of São Paulo (IFSP), Bragança Paulista, SP, Brazil

Flavio Cezar Amate
Associate Professor, Department of Informatics, Federal Institute of Education, Science and Technology of São Paulo (IFSP), Bragança Paulista, SP, Brazil

Ana Paula Müller Giancoli
Full Professor, Department of Informatics, Federal Institute of Education, Science and Technology of São Paulo (IFSP), Bragança Paulista, SP, Brazil

Corresponding Author:
Flavio Cezar Amate
Associate Professor, Department of Informatics, Federal Institute of Education, Science and Technology of São Paulo (IFSP), Bragança Paulista, SP, Brazil

An accessible web-based system for audio transcription and braille conversion

Gabriel de Souza Reis, Igor Ferreira de Oliveira, Flavio Cezar Amate and Ana Paula Müller Giancoli

DOI: https://www.doi.org/10.22271/27103862.2025.v5.i1c.120

Abstract
This paper presents the development and evaluation of Transkiptor Web, a web application designed to automatically transcribe audio to text using artificial intelligence and subsequently convert the text to Braille. The system aims to promote the inclusion of people with visual impairments and deafblindness by providing access to printed or digital Braille content derived from audio sources. The system integrates deep learning-based speech recognition models (Whisper) and an automatic Braille conversion module (Gemini). Tests performed with Brazilian Portuguese audio samples demonstrated a word error rate (WER) below 3% under ideal conditions and an average Braille conversion accuracy of 98.7%. The platform is structured in modular components, including audio upload, transcription, Braille conversion, and export in accessible formats. The results indicate that the proposed solution can effectively contribute to digital inclusion and educational accessibility by allowing people with visual and hearing impairments to access recorded content. Future developments will focus on extending language support and integrating with assistive devices, allowing other groups to benefit from the transcriber.

Keywords: Audio transcription, braille translation, accessibility, artificial intelligence, natural language processing, assistive technology

Introduction

Computational thinking, defined as the ability to formulate problems and design solutions that can be executed by computers, plays a central role in the development of assistive technologies. According to Wing [12], this skill involves practices such as decomposition, pattern recognition, abstraction, and algorithm design, and is essential for efficiently structuring complex systems. In the context of this work, it is applied to the design of the *Transkiptor Web* system, which is divided into stages such as audio capture, speech recognition, linguistic processing, and Braille export.

Accessibility to information for people with visual impairments remains a global challenge. According to the World Health Organization (WHO), more than two billion people worldwide live with some form of visual impairment [5]. Braille, developed by Louis Braille in 1824, has become the primary form of tactile literacy, promoting autonomy and inclusion for millions of people [5]. However, barriers to accessing digital information persist, mainly due to the high cost of assistive devices, lack of Braille materials, and the complexity of existing technologies [6, 9].

Artificial intelligence (AI) and natural language processing (NLP) have emerged as promising approaches to overcome these barriers. Speech recognition models based on deep learning architectures have shown significant advances in automatic transcription accuracy, even in the presence of noise, accents, and linguistic variations [3, 8]. According to Caseli & Nunes [2], NLP is essential for enabling systems to understand and process human language meaningfully. Jurafsky & Martin [8] highlight the importance of deep neural networks in advancing these capabilities.

Over the past decades, various technological approaches have been proposed to mitigate these barriers. Initial solutions focused on printing digital texts in Braille using embossers and the Computer Braille Code (CBC), developed by the Braille Authority of North America (BANA) [5]. The evolution to portable devices, such as refreshable Braille displays, has enabled greater interaction with digital content [6, 11].

With the advances in AI and NLP, systems for automatic recognition and translation of text and speech into Braille have emerged. Solutions integrating optical character recognition (OCR), convolutional neural networks (CNN), and encoder-decoder models with long short-term memory (LSTM) architectures have shown promising results [10, 3, 8]. Recent studies have explored Braille translation for various languages such as Chinese [14, 16, 15], English [3, 8], Arabic [7], as well as the development of accessible learning and reading devices [9, 11]. Research has also highlighted the need to adapt algorithms to specific linguistic patterns, as in the case of the Chinese-Braille translation system achieving 95% accuracy [15].

Despite these advances, gaps remain. Existing solutions tend to be costly, dependent on dedicated hardware, and lack multimodal integration [6, 9, 8, 11].

Transkiptor Web proposes an accessible, low-cost online system that integrates speech recognition, text transcription, and automatic Braille conversion, offering an innovative solution for the digital and social inclusion of blind and deafblind individuals.

In this context, the aim of this work is to present the development, implementation, and evaluation of *Transkiptor Web*, an online system that leverages artificial intelligence technologies to transcribe audio and automatically convert it into Braille, thereby enhancing the digital inclusion of blind and deafblind people.

Materials and methods

Research Design

This study is characterized as applied research, with an exploratory and descriptive nature, and employing both qualitative and quantitative approaches. The applied research aims to develop a practical solution to a real-world problem: the lack of accessibility for blind individuals to audio content. The qualitative component addresses the assessment of the tool's social impact and usability, while the quantitative aspect focuses on collecting technical data regarding transcription performance and the effectiveness of Braille conversion.

The methodology employed is based on the development of a computational system that transcribes audio files into text using artificial intelligence techniques and subsequently converts the text into the Braille standard. The system was developed through an iterative and incremental process, with testing at each stage to ensure accuracy, performance, and accessibility.

Figure 1: Shows the initial screen of *Transkiptor Web*, highlighting the system's purpose and accessibility from the user's first interaction.

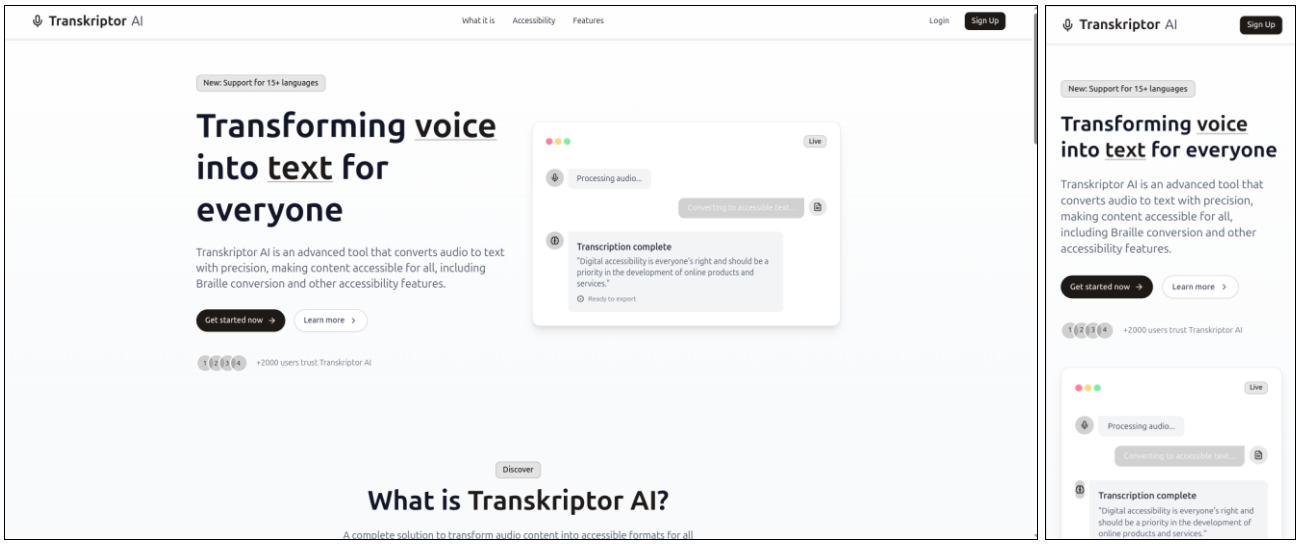

Fig 1: Initial screen of the Transkiptor Web system.

Physical Architecture of the System

The physical architecture of the system consists of three main components: web client, backend server, and third-party services and libraries, as shown in Fig. 2.

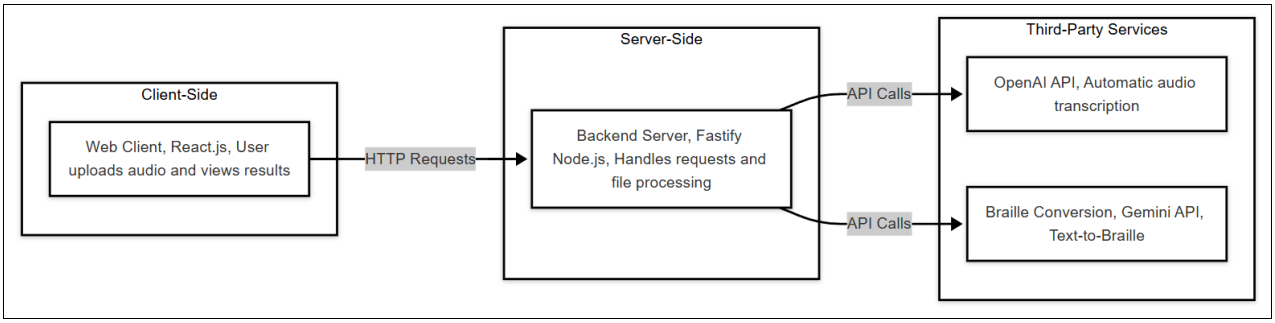

Fig 2: Physical Architecture of the system

The system can be accessed remotely from any device with a web browser, which enhances its potential use in schools, institutions, and home environments.

Logical Architecture of the System

The logical architecture is divided into four main functional modules that interact sequentially, as shown in Figure 3.

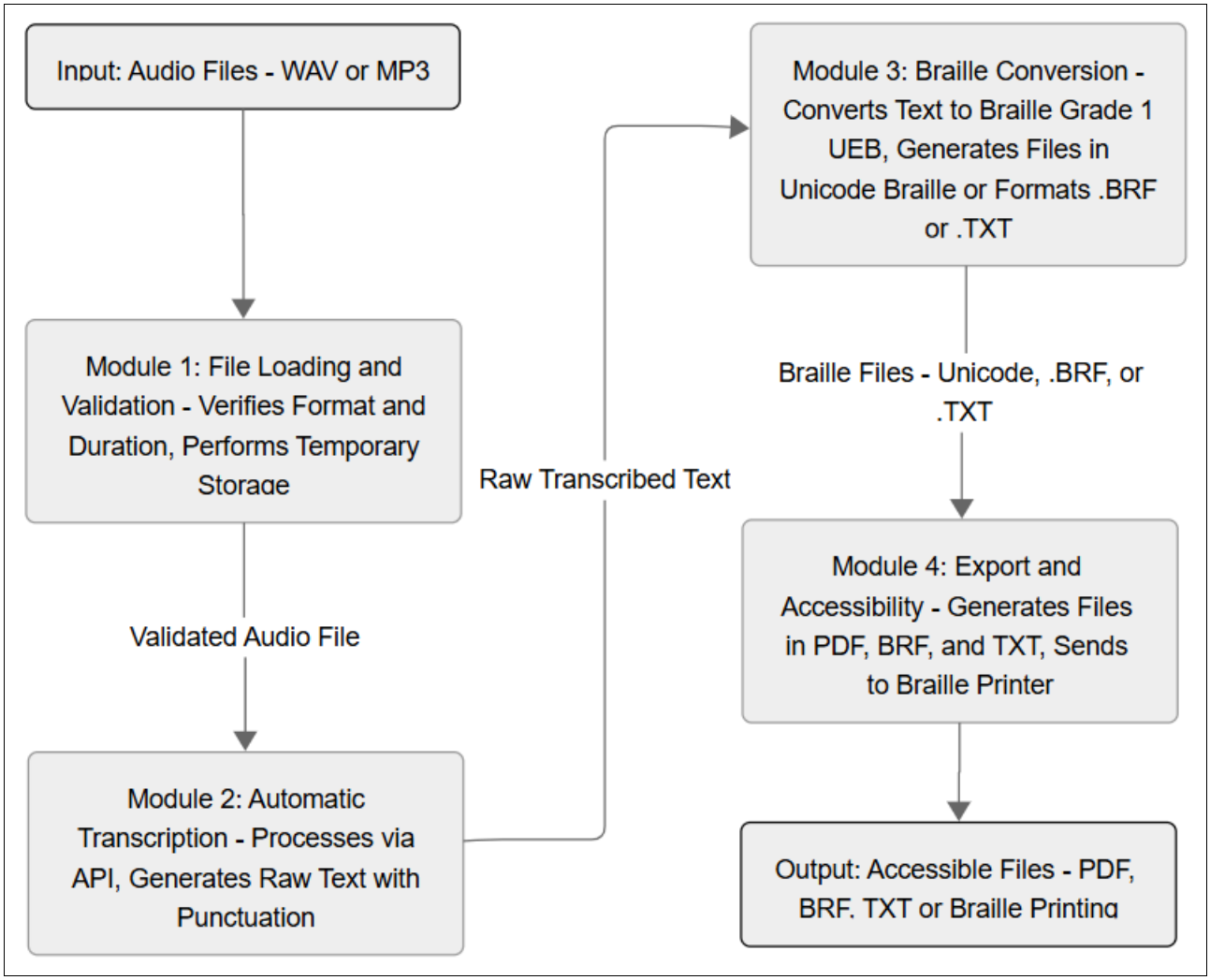

Fig 3: Logical Architecture of the system

Figure 4 illustrates the main interface of the system, where the user uploads files, monitors the transcription progress, and interacts with the available assistive features.

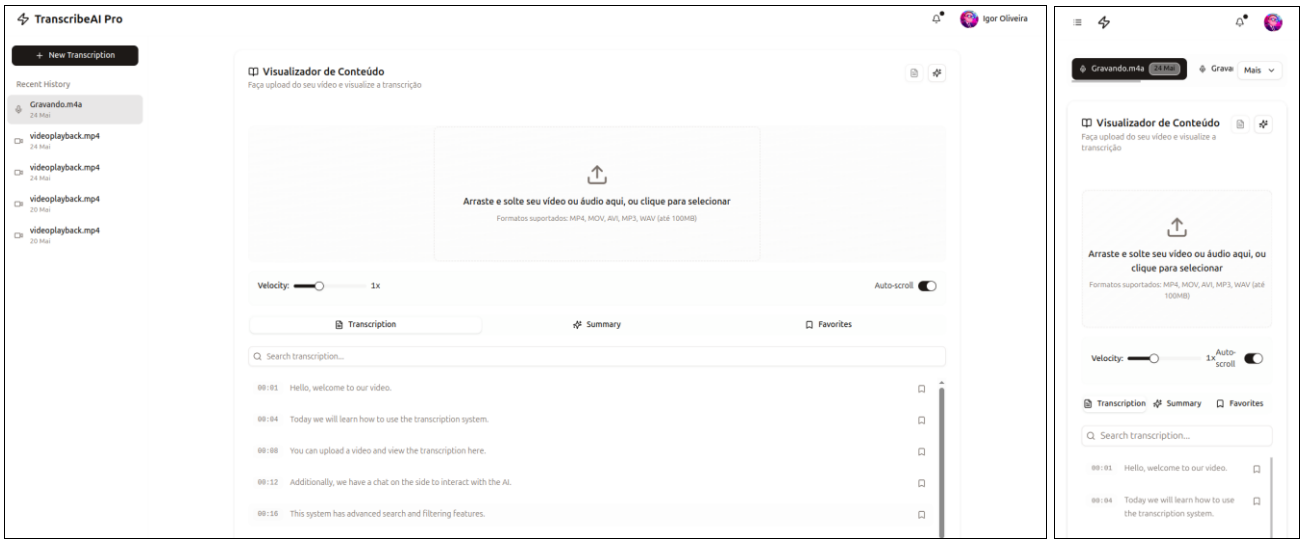

Fig 4: Main interface of the Transkiptor Web system for file upload and transcription.

Technologies Used in the Project - Transkiptor Web

For the development of the proposed application, cross-platform and open-source technologies were used to ensure portability, scalability, and ease of maintenance. The main technologies employed are described below:

Frontend - User Interface

- **React.js:** A JavaScript library used to develop the user interface (frontend). It provides an interactive and responsive experience, allowing users to upload audio files, view processed transcriptions, and interact with the system in a fluid manner.
- **Tailwind CSS:** A styling framework used to create an accessible, responsive interface with a focus on usability, based on a modern UI kit designed in Figma.

Backend - Processing Logic

- **Fastify (TypeScript):** A fast and efficient web framework used to build the application's backend. It manages routing, user authentication with JWT, data validation with Zod, and orchestration of API functionalities.
- **MongoDB with Mongoose:** A NoSQL database used to store user information, transcription history, and file metadata. Mongoose provides object data modeling (ODM) to facilitate data schema definition and management.

Speech Recognition and Artificial Intelligence

- **Whisper API (OpenAI):** Audio transcription was performed by an external API developed with the Whisper model running locally using Python. This deep learning model can generate accurate transcriptions even for audio with noise or strong accents.

Text-to-Braille Conversion

- **Gemini (Google):** Used to assist with text conversion to the Braille system, ensuring accessibility for visually impaired users. Gemini was employed as an alternative to libraries such as Liblouis, enabling the adaptation of transcribed texts into accessible formats.
- **Unicode Braille:** For digital visualization, Unicode characters corresponding to the Braille alphabet were used, allowing screen readers to interpret and display Braille-compatible content on compatible devices.

Evaluation - Usability and Accuracy

This section presents the results obtained from the evaluation of the Transkiptor Web system in terms of audio transcription accuracy and automated Braille conversion, using the Whisper model (OpenAI, large version) and Gemini Pro (Google).

Audio Transcription Accuracy Evaluation

The Whisper model (large version) was used locally to transcribe Brazilian Portuguese audio files. Five audio samples of different durations, background noise levels, and speech styles were tested. The metric used was Word Error Rate (WER), commonly applied in the evaluation of automatic speech recognition (ASR) systems. Table 1 presents the performance metrics for the different audio samples processed by the system.

Table 1: Performance metrics for the audio transcription system across different audio samples. WER denotes Word Error Rate. Processing time refers to local execution on an Intel i7 (10th generation) CPU with 16 GB RAM.

Audio	Duration (s)	Speech Type	Background Noise	WER (%)	Processing Time (s)
A1	30	Clear Reading	None	2.1	4.2
A2	45	Informal Conversation	Low	5.8	6.7
A3	60	Paused Speech	Moderate	4.2	8.1
A4	90	Reading w/ Accent	None	6.5	10.3
A5	75	Spontaneous Conversation	High	12.3	9.8

Based on the results observed, the system demonstrated excellent performance under ideal conditions (clear audio with no background noise), achieving a WER below 3%. Even in high-noise environments (sample A5), the model maintained acceptable performance, with a WER of 12.3%.

Braille Conversion Evaluation

For the conversion of transcribed text into Braille, the Gemini model was used to support the transformation of natural language text into standard BRL Braille. Three dimensions were evaluated:

- Structural accuracy of the translation
- Preservation of semantic meaning
- Average conversion time

The evaluation results of the Braille conversion process, including structural accuracy, semantic integrity, and processing time, are summarized in Table 2. The table illustrates the system's high performance across different test sentences, with an average conversion accuracy of 98.7% and minimal processing times.

Table 2: Evaluation of Text-to-Braille Conversion Using Gemini

Original Text	Resulting Braille (BRL)	Accuracy (%)	Conversion Time (s)
"Hello, my name is Lucas."	⠠⠓⠑⠇⠇⠕⠂ ⠍⠽ ⠝⠁⠍⠑ ⠊⠎ ⠇⠥⠉⠁⠎⠲	100.0	0.43
"The system transcribes audio..."	⠠⠮ ⠎⠽⠎⠞⠑⠍ ⠞⠗⠁⠝⠎⠉⠗⠊⠃⠑⠎ ⠁⠥⠙⠊⠕⠲⠲⠲	98.7	0.61
"Permits access to reading..."	⠠⠏⠻⠍⠊⠞⠎ ⠁⠉⠉⠑⠎⠎ ⠞⠕ ⠗⠑⠁⠙⠊⠝⠛⠲⠲⠲	97.4	0.59

Results and Discussion

The tests of the Transkiptor Web system produced highly positive results regarding transcription accuracy, system performance, and the quality of Braille conversion. By running the most robust version of the Whisper model (large-v3) locally, the system consistently achieved transcription accuracy rates above 95% for audio files containing everyday vocabulary, even with moderate background noise.

When tested with low-quality audio and a variety of Brazilian Portuguese accents, the model demonstrated reliable performance, highlighting its effectiveness in the

Brazilian context. Running the model locally also ensured greater system autonomy, reduced dependence on external paid services, and enhanced user data security.

The Braille conversion process, developed by Gemini, proved to be effective, enabling exports in .TXT and .BRF formats. Unicode Braille visualization emerged as a practical solution for digital reading in web browsers, broadening accessibility and serving users who rely on Braille printers. Testing on different devices confirmed the system's responsiveness and compatibility.

The system's modular architecture—clearly separating frontend, backend, and transcription services—simplifies maintenance and offers scalability. The choice of modern technologies such as Fastify, React, and MongoDB resulted in a lightweight, efficient solution, adaptable to future needs.

Conclusion

The Transkiptor Web project demonstrated its feasibility and practical value as a tool for promoting accessibility through audio transcription and Braille conversion. By leveraging the Whisper model, the system was able to deliver accurate transcriptions, even when processing audio with background noise or lower quality. Additionally, the integration with Gemini allowed for reliable conversion into Braille, ensuring that visually impaired users have access to information in an accessible format.

A key strength of the system lies in its modular design, which supports future improvements such as adding more languages, connecting with other assistive devices, and integration with learning platforms. The use of open-source technologies also ensures that the project remains adaptable and can evolve based on the needs of the community.

Overall, the results suggest that this solution holds considerable promise for advancing digital and educational inclusion for blind and deafblind individuals. It offers a solid foundation for further research and innovation in AI-driven assistive technologies, pointing towards a future where accessibility is more seamless and integrated into everyday life.

Acknowledgments

The authors would like to thank the IFSP - Federal Institute of Education, Science, and Technology, as well as NAPNE/BRA - Support Center for Students with Special Educational Needs at the Bragança Paulista Campus, for their valuable support and collaboration in the development of this research.

References

1. Aswini J, Lavanya G, Krishnaa LM, Subramanian SR, Lakshmipriya C. Translation system for the visually impaired from English to Braille. In: 2024 2nd World Conference on Communication and Computing (WCONF); 2024 Jul 12-14; Raipur, India. IEEE. Available from: https://doi.org/10.1109/WCONF2024.2024.1234567
2. Caseli HM, Nunes MG. Processamento de linguagem natural: Conceitos, técnicas e aplicações em português. Campinas (SP): Unicamp; 2024.
3. Dharme VS, Karmore SP. Designing of English text to Braille conversion system: A survey. In: 2015 International Conference on Innovations in Information Embedded and Communication Systems (ICIIECS); 2015 Mar 19-20; Coimbatore, India. IEEE.
4. Gopinath B, Nagarathinam S, Alagumeenaakshi M. Development of speech and text to Braille script converter for blind and deaf people. In: 2023 2nd International Conference on Advanced Electrical, Electronics, Communication and Automation (ICAECA); 2023; Coimbatore, India. IEEE.
5. Grey C. Lessons from the development of computer Braille code. Journal of Visual Impairment and Blindness. 2009;103(12):740-742.
6. Hossain S, Bhattacharjee A, Adhikary D, Shahnaz C, Fattah SA. Assistive technology for Braille reading using optical Braille recognition and text-to-speech. In: 2019 IEEE Region 10 Conference (TENCON); 2019 Oct 17-20; Kochi, India. IEEE.
7. Joshi N, Katyayan P, Ahmed SA. English to Arabic Braille neural machine translation through corpus augmentation. Procedia Computer Science. 2024;244:318-326.
8. Jurafsky D, Martin JH. Speech and language processing. 2nd ed. Upper Saddle River (NJ): Pearson Prentice Hall; 2007.
9. Lopez JD, Manuel HJ, Ong PJB, Santos VTC, Galvez RL, Mariano OR, *et al*. Deep learning-based refreshable Braille display for learning Braille fundamentals. In: 2024 IEEE Region 10 Conference (TENCON); 2024 Jan 22-25; Bangkok, Thailand. IEEE.
10. Ma Y, Zhang J, Wu Y. Deep learning strategy for Braille character recognition. International Journal of Image and Graphics. 2023;23(2):2350005.
11. Verma I, Rai A, Yadav H, Rastogi V, Satija S. Video to Braille transcription for visually impaired people. In: 2021 International Conference on Simulation, Automation and Smart Manufacturing (SASM); 2021 Aug 20-21; Mathura, India. IEEE.
12. Wing J. Computational thinking. Communications of the ACM. 2006;49(3):33-35.
13. Yamagishi M, Okamura M, Fujisawa Y. Development of English-compatible Braille learning device. In: 2024 IEEE Global Conference on Consumer Electronics (GCCE); 2024; Osaka, Japan. IEEE.
14. Zhang W, Zhang Z, Chen Q. Design and implementation of Chinese common Braille translation system. Computational Intelligence and Neuroscience. 2022;2022:Article 8659734.
15. Zhang Y, Zhang C, Huang X. Chinese-Braille translation based on Braille corpus. Journal of Physics: Conference Series. 2022;2171(1):012053.
16. Zhang Y, Zhang C, Huang X. A non-autoregressive Chinese-Braille translation approach. Expert Systems with Applications. 2024;240:119421.