ISSN: 2584-0495 Vol. 2, Issue 4, pp. 785-791

International Journal of Microsystems and IoT

ISSN: (Online) Journal homepage: https://www.ijmit.org

Braille to Speech conversion using BT-CNN

Aditya Upadhye, Tharunnesh R, Jeslin Shanthamalar J, C A Daphine Desona Clemency and N S Usha

Cite as: Upadhye, A., Tharunnesh, R., Jeslin, S. J., Clemency, C. A. D. D., & Usha, N. S. (2024). Braille to Speech conversion using BT-CNN. International Journal of Microsystems and IoT, 2(4), 785-791. https://doi.org/10.5281/zenodo.12205060

Published online: 22 April 2024

DOI: https://doi.org/10.5281/zenodo.12205060

International Journal of Microsystems and IoT
Vol. 2, Issue 4, pp.785-791; DOI: https://doi.org/10.5281/zenodo.12205060

Braille to Speech conversion using BT-CNN

Aditya Upadhye, Tharunnesh R, Jeslin Shanthamalar J, C A Daphine Desona Clemency and N S Usha

Department of Computer Science and Engineering, Sathyabama Institute of Technology, Chennai, Tamil Nadu, India

ABSTRACT

Braille is a universally accepted system of writing and understanding scripts for blind people. It is a touch-based system for reading and writing. The alphabet is composed of specific dot configurations. An image containing a word or statement in Braille can be given to the model. Although there are various ways to implement it, the proposed model uses Convolutional Neural Network (CNN) to analyze the dots in the Braille image and convert them into letters, eventually forming words and statements. Keras library is to perform several operations like generating image data, splitting the dataset and evaluating the model's accuracy. This text is further processed by the model and the final output is an audio file which is a speech of the text obtained from the Braille Image. This text to speech conversion is done by using Google Text to Speech. By training the model with a large number of epochs, an accuracy of about 96.30% is achieved.

KEYWORDS

Convolutional Neural Network; Deep Learning; Machine Learning; Speech Synthesis

1. INTRODUCTION

Traditionally, Braille is used by blind, partially sighted, and deafblind people. Louis Braille created braille in France during the 1820s. The spatial units used to form braille symbols are called braille cells. Imagine the number 6 on a dice [1]. A complete braille cell is made up of six raised dots organized in two parallel vertical rows of three dots each. The digits one through six represent the locations of the [2] dots. When one or more of these six dots are used, there are 64 possible combinations. Both filled and empty cells are also used. Punctuation, numbers, alphabet letters, and even entire phrases can be represented by a group of one or more dots positioned inside the cell [3].

There are several ways to produce Braille, including:

- Braille Typewriter: This is a special typewriter with just six keys, one for each of the dots in a Braille cell. When a key is pressed, a metal stylus embosses the dots for that character onto the paper [4].
- Braille Printer: This is a kind of printer that embosses (puts raised dots on-to) paper to create Braille characters.
- Refreshable Braille Display: This is a device containing row upon row of mechanical or electronic Braille cells. Each cell has multiple pins that can be raised or lowered to form any Braille characters [5].

A Braille to Speech conversion tool whispers the secrets of the written word through raised dots, providing a unique means to encode text. But, remarkable as Braille is, it comes with its own set of challenges [6]-[8]. The specialized knowledge required for reading and writing often leaves it an enigma. Not just about algorithms and data, braille to speech conversion brings a means to empower individuals with the most important of gifts -- the written word.

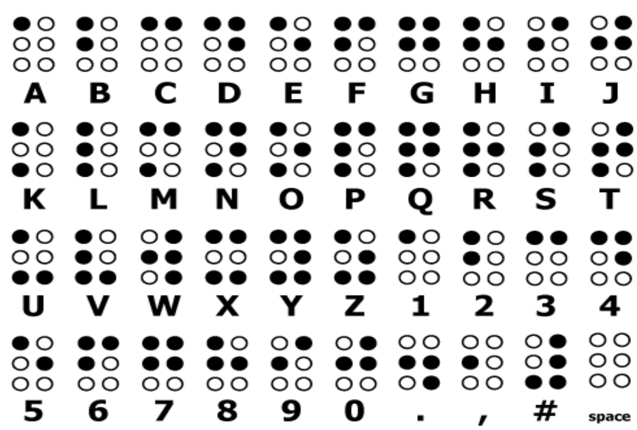


Fig. 1 Braille Characters

For people with visual impairment, the first and foremost advantage of Braille to Speech conversion is accessibility. It allows users to hear printed material, signages and other written information, which promotes independence and inclusion with the help of audio output. Another key advantage of a Braille to Speech translation model is its ability to translate Braille text into spoken language in real-time [9]-[11]. Using a well-designed Braille to Speech model, a blind person would be able to immediately hear spoken output of any Braille text they place their fingers on.

In an educational context, Braille to Speech conversion provides an important form of auditory reinforcement for Braille students [12]. This reinforcement can help students to develop a firm association between new Braille characters and their corresponding spoken language in a faster manner than trial and error alone would allow.

2. RELATED WORK

Using preprocessing, image segmentation, and feature extraction approaches, M. Hanumanth Appa et al. demonstrated that several researchers' efforts were still unsuccessful in achieving 100% accuracy with a high translation rate [13]-[14]. In his research, Chollet demonstrated how Inception modules acted as a midway point, situated at opposite ends of a discrete spectrum, between convolutions and depthwise separable convolutions. Kirill Smelyakov et al. used Optical Braille Character Recognition system using Artificial Neural Networks. K M Naimul Hassan et al. developed a dual-purpose system, in 2019, which can help to detect objects around them and a refreshable braille display to display text in the form of Braille Characters. A. A. Choudhury et al. introduced a novel technique that uses the YOLO (You Only Look Once) algorithm to enhance real-time recognition of Braille characters. Known for its speed and recognition accuracy, Yolo was modified to recognize Braille characters, with the aim of improving accessibility and independence for the visually impaired [15]-[17].

A useful method for pronouncing and converting scenes to text was proposed by S. M. Qaisar et al. in 2019. This method can aid visually impaired people in understanding written text found in everyday objects such as caution notices, medicine cartons, and basic food goods. V Jha and K Parvathi proposed a transliteration system, in 2019, which converts Hindi characters into their corresponding Braille Characters using Optical Braille Recognition (OBR) Braille dot pattern. C. Park at al. proposed a 3-stage method using R-CNN as the first stage for banknote images captured in disturbed environments using a smartphone camera [18]. A device was demonstrated by G. C. Bettelani et al. that uses an electromagnetic field to actuate six independently controlled dots. The alphabet letters can be encoded with this device in accordance with the Braille system. Additionally, the blind participants said the gadget was well-accepted, practical, economical, and intuitive. B.-M. HSU used the Radio Character Segmentation Algorithm (RCSA) to recognize complex Braille characters and translate them into plain English [19]. In a pivotal look, P. Kaur et al. harnessed both conventional and contemporary strategies to broaden a laptop program that converts Hindi Braille into spoken language. By combining those methods, they created a gadget that complements accessibility for the visually impaired.

Shokat et al. did groundbreaking research in assistive technology for visually impaired individuals in 2020. Their use of the GoogleNet Inception model achieved an impressive 95.8% accuracy in predicting Braille characters. Mao developed a model that converts Chinese Braille into spoken language, increasing communication accessibility and inclusion [20]. This pioneering work extends beyond text, providing a natural understanding of the sounds of Braille letters, and contributing to universal inclusion and accessibility. In 2021, T. Kausar et al.'s innovative approach included replacing modules in traditional convolutional neural networks (CNNs) with techniques designed to significantly reduce computational costs. This system addresses a fundamental challenge in Braille character recognition by streamlining computational requirements, especially in critical or real-time. A method to develop a low-cost, portable communication device that can translate any Braille input into alphanumeric text format was presented by Ramachandran et al.

A mechanism that prints the Braille alphabet on a sheet after receiving an alphabet voice command was proposed by M. A. Hussain. In the same year, Bhatia et al. Automatic Arabic language has contributed to the translation system in Braille and contributed to the simple technology, that "Convolutional Neural Network Real Time Arabic Speech Recognition to Arabic Braille" is known as the requirement of the classification, that in real time in Arabic for translation capabilities [21]. Leveraging the power of convolutional neural networks (CNNs), this revolutionary technology empowers individuals with sensory impairment, providing effective communication and access to information, exemplifying that they are committed to being inclusive and improving their lives. Ramani RG and Shanthamalar JJ have conducted extensive prior research on feature extraction methods utilising full-reference and no-reference image quality measures for picture quality classification. This work presents an automatic feature extraction system for fundus images utilising image processing techniques, followed by an automatic image quality classification system employing different classification algorithms. A camera embedded hand glove with OCR for conversion to Braille and speech for blinds have been created by Seung-Bin Park and Bonghyun Kim to improve the quality of life [22].

3. METHODOLOGY

3.1 Dataset:

In the proposed approach, the idea is to use the resources available inside the "Braille Character Dataset" from Kaggle to train our model. In BW scale, each image is a letter and is 28x28 pixels. The image number, the alphabet of characters, and the kind of data augmentation applied to each image make up their names. The dataset consists of 26 characters, 3 augmentations, and 20 distinct images with varying augmentation values (i.e., distinct shift, rotational, and brightness values). 1560 images make up this dataset.

3.2 Convolutional Neural Network:

Convolutional Neural Network (CNN) is like an intelligent computer system that is really good at understanding images and videos. It can see and recognize different things in images, similar to how humans see and understand things. CNNs are used for many applications such as recognizing objects in images, reading medical images, and even understanding words in text. They're like the super detectives of the computer world for visuals.

The first and essential step in the digital representation of visual data is image acquisition (Figure 2). It involves special devices such as cameras, scanners and sensors to capture visible information from the physical world. These devices convert analog information such as light and electromagnetic radiation into digital images. The quality of these acquired images is extremely important as it directly affects the success of image production and subsequent analysis. Factors such as lighting, camera settings, and sensor sensitivity play an important role in ensuring the accuracy and realism of digital images. Image acquisition is widely used in a variety of industries, including photography, medical imaging, surveillance, and computer vision. It serves as the basis for such vital processes as medical diagnosis, discovery, and independent guidance. By converting real-world visual information into a digital format, image acquisition enables computers and digital systems to make sense of and interact with the physical world, enabling product delivery of many practical and new uses.

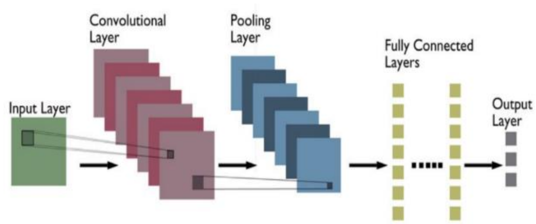


Fig. 2 Convolutional Neural Network

In Image Preprocessing, a comprehensive image records preprocessing strategy is implemented with the usage of two key strains. TensorFlow's Keras library imports the image module, which contains functions for loading and preprocessing images. This ensures finest formatting for superior version performance. Concurrently, the Image Data Generator module line employs the Image Data Generator module for dynamic dataset augmentation all through schooling, enhancing the model's adaptability and generalization throughout numerous scenarios.

To make the proposed version even more reliable, the plan is putting preprocessing strategies into play, refining the process of entering pictures before they undergo analysis. The operational method of proposed version entails a meticulous dissection of the complete photograph into smaller fragments, accomplished by way of figuring out the feature six dots that compose a Braille character.

The pooling process is implemented using the L. MaxPooling2D layers. Specifically, the lines x = L. MaxPooling2D (2,2)) (x) after the separable convolution layers down sample the spatial dimensions of the input feature map through max-pooling. A down sampled (pooled) feature map is produced by the pooling process known as "max pooling," which determines the maximum value for each patch in a feature map.

In the proposed work, the feature extraction process occurs in the layers of the neural network model, specifically within the SeparableConv2D and GlobalMaxPooling2D layers. These layers are tasked with capturing essential features from the input images. The SeparableConv2D layers perform convolution operations to capture spatial hierarchies of features, while the GlobalMaxPooling2D layer aggregates the most important features by taking the maximum value across spatial dimensions. Therefore, the lines involving SeparableConv2D and GlobalMaxPooling2D contribute to the feature extraction process in this model. The extraction method happens when the use of the model. Predict(x) line. This line sends the enter records x through the skilled version, and the output represents the extracted capabilities or predictions primarily based at the discovered patterns from the input. The np.expand_dims (x, axis=zero) line is used to adjust the shape of x by adding a further dimension, which is a common preprocessing step, but it does not perform the characteristic extraction itself.

3.3 BT-CNN Model:

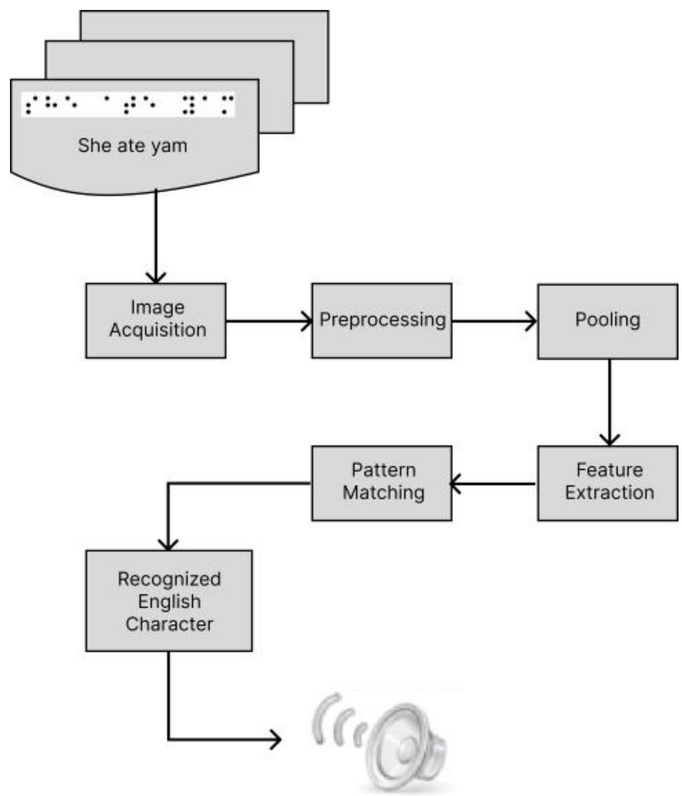


Fig. 3 Architecture Diagram

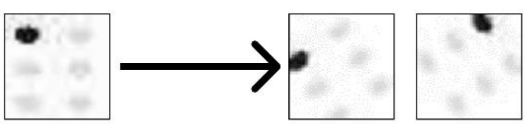
Figure 3 presents a visible depiction of the proposed version, illustrating its structure and functionality within the system of changing Braille characters into an audio file. This graphical illustration serves as a schematic overview highlights the important components and connections implicated in the operation of the model. The figure captivates a clear visual representation of how the proposed structure is employed to facilitate the transformation of Braille characters, illuminating the sequential steps and interactions that yield the final output—an audio clip. This visible representation aids in comprehending the design of the model and its role in the transformation process of Braille character conversion, by providing a visible reference for readers to better hold close the method hired inside the study.

A CNN architecture was employed, comprising several layers: separable convolutional layers followed by max pooling operations. The network further utilized global max pooling to reduce spatial dimensions. Dense layers with ReLU and Leaky ReLU activation functions, along with a softmax output layer, were implemented for feature extraction and classification. The final output layer consisted of 26 units, indicating a multi-class classification task.

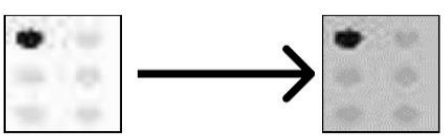
Table. 1 Layers in BT-CNN

Layer (type)	Output Shape	Parameter
Input	28, 28, 3	0
Separable Convolution	26, 26, 64	283
Max Pooling	13, 13, 64	0
Separable Convolution	11, 11, 128	8896
Max Pooling	5, 5, 128	0
Separable Convolution	4, 4, 256	33536
Global Max Pooling	256	0
Dense	256	65792
LeakyReLU	256	0
Dense	64	16448
LeakyReLU	64	0
Dense	26	1690
Output	26	0

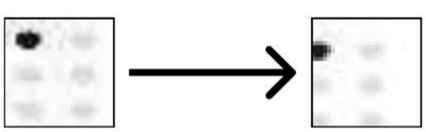
The images of the training dataset are augmented by applying image augmentation techniques (such as rotating, brightness, width-height shifting etc), thus increasing the dataset further and significantly increasing the diversity of the images available for training the model thus generalizing the model. In Figure 4 Image Augmentation performed on an image is shown.



(a) Rotation



(b) Brightness



(c) Width-Height Shifting

Fig. 4 Image Augmentation performed on an image

After Image Augmentation, the training dataset is splitted into the training (80%) and the testing (20%) data. The model architecture is defined and compiled using the Adam optimizer and categorical cross-entropy loss function for multi-class classification tasks.

The system specifications used for developing the proposed model were:
Processor: Intel Core i7
RAM: 16 GB
GPU: Intel Iris Xe
OS: Microsoft Windows 11
An alternative platform would be Google Colab that provides CPU and T4 GPU hardware runtime for free.

Training and validation data generators were used to train the model across 100 epochs. The training process is monitored and optimized using callbacks, which include model checkpointing, learning rate decrease, and early termination. Epoch, in machine learning, is referred to one-time passing of the entire training data through the model or algorithm. After certain number of epochs, if there isn't significant increase in the model's accuracy, the early-stopping callback method executes to stop the training process. The best version Then, the accuracy of model is evaluated using a validation data generator. The achieved model accuracy, computed using the evaluation generator, was printed with four decimal places for precision. The evaluate generator function is used to return a list of metrics values and extract the accuracy value.

The model. Predict method is used to generate predictions for the input data provided by Val generator. This typically occurs after the training phase when the model has learned to extract relevant features from the training data. In the proposed work, plt.imshow (img) uses Matplotlib to plot the image represented by the img variable.

The image.img_to_array (img) function is used to convert the image into a NumPy array. Then, the dimensions of an array are found using the shape function and this is the shape of the array - (28, 28, 3). Preprocessing of input character is done by expanding its dimensions and then use it to make predictions using the proposed model. The predicted output (desired English Alphabet) was obtained by finding the index of the highest probability prediction.

As the final step in this system, the model is designed to generate an audio file that aligns with the Braille word or sentence first of all supplied as an image. In order to streamline this process, Google Text to Speech is used.

Once trained, the network can synthesize speech for new input text by:

- Text pre-processing: The input text is analyzed and segmented into phonemes (basic units of sound).
- Feature prediction: The network predicts the acoustic features corresponding to each phoneme based on the surrounding context.

- Waveform generation: The predicted acoustic features are used to generate a speech waveform that resembles the target speaker's voice

4. RESULTS

This paper deals with the image recognition of Braille characters and convert Braille Text into Audio Speech, a CNN approach is proposed which uses various different types of layers in the convolutional neural network to recognize the Braille character images and classify each into one of the 26 characters of the English alphabets. The highest accuracy obtained by the proposed model is 96.30%. The advantage of the proposed system is that it is able to predict the braille characters and convert them into equivalent English Characters in a very short span of time. One disadvantage of the system could be that in certain scenarios the segmentation engine is not able to correctly segment the braille characters, leading to incorrect prediction for a few characters. An attempt to overcome this incorrect prediction is the use of Google Text to Speech so even if few letters of an English word are not predicted properly, listening to the audio speech file could help one assume what was the message being conveyed. In Figure 5, Figure 6, Figure 7 and Figure 8 shows Graph of accuracy at each epoch, Graph of loss function at each epoch, Application Homepage, Application showing converted Text respectively.

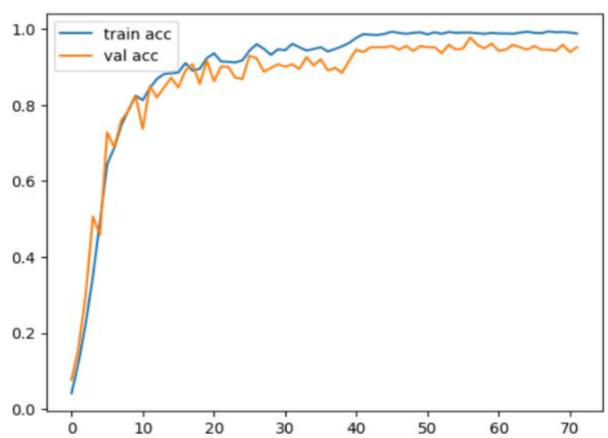


Fig. 5 Graph of accuracy at each epoch

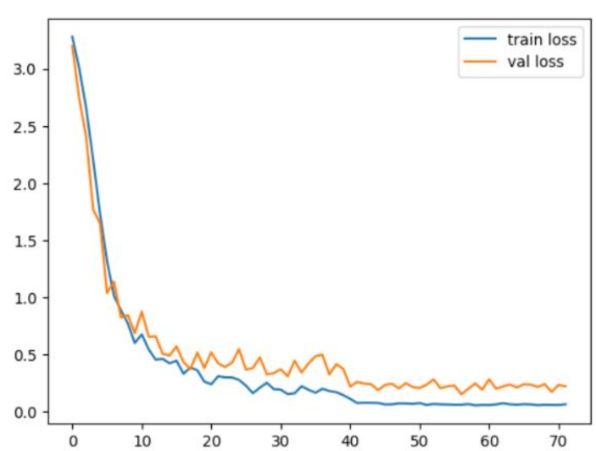


Fig. 6 Graph of loss function at each epoch

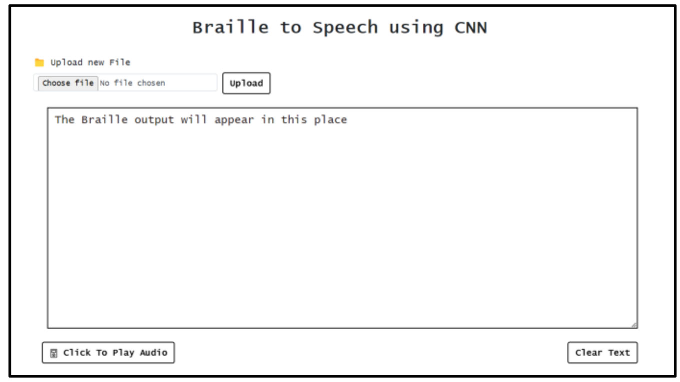


Fig. 7 Application Homepage

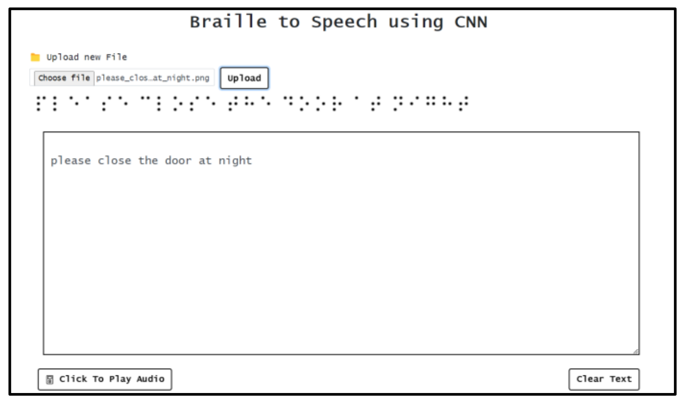


Fig. 8 Application showing converted Text

5. COMPARATIVE ANALYSIS

Table. 2 Comparative Analysis of Existing Methods vs proposed Method

No.	Researcher	Algorithm	Accuracy
1	K. Smelyakov (2018)	ANN	95.00%
2	S. Shokat (2020)	GoogleNet	95.80%
3	J Mao (2021)	Prosody Prediction	94.42%
4	T Kausar (2021)	DenseNet201	95.20%
5	M. A. Hussain (2022)	VGG16	92.42%
6	**Proposed Work**	**BT-CNN**	**96.30%**

A comparison between the proposed system and recent methods is shown in Table 2. Our system evaluated the overall performance with existing approaches; however, it is challenging to compare them because they were trained on different datasets. J Mao and M. A. Hussain's approaches convert Braille text to only Chinese and Bangla languages, respectively, which caters a relatively small set of people as compared to the Internationally accepted English language. S. Shokat's study does not mention incorporating feedback from visually impaired individuals during the development and evaluation of the system. Involvement of end-users in the design process could provide valuable insights into their needs, preferences, and usability challenges, ultimately leading to a more user-centric and effective solution. While T. Kausar's preprocessing stage involves the application of Principal Component Analysis (PCA) to enhance feature representation, they also add complexity to the system and may require additional computational resources.

6. CONCLUSION AND FUTURE WORK

In this study, efforts were made to use Deep learning for braille to speech conversion and got accuracy of 96.30%. A potential drawback of the method could be that under specific conditions, the segmentation engine could not be able to accurately segment the braille characters, which could result in some characters being predicted incorrectly. Using Google Text to Speech is an attempt to address this inaccurate prediction; even if a few letters in an English word are mispronounced, listening to the audio speech file can help determine the meaning being expressed. This system can be proved as a major breakthrough in the field of automation and innovative technologies and can be used in various websites and places with diverse set of people and can further be incorporated with systems which can help all the visually challenged people as well as normal people around the world to understand and visualize their respective surroundings and communicate with people in a simpler and efficient manner. This study can further be enhanced in the future using different and more accurate, high-resolution datasets or using different, more complex methods like Extreme Machine Learning (EML), or Classification and Regression Trees (CART) for better performance of the model which may require higher computation power and much advanced software for efficient working.

REFERENCES

1. M. Hanumanthappa and V. V. Murthy, "Optical Braille recognition and its correspond-ence in the conversion of Braille script to text — A literature review," 2016 International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Bengaluru, India, 2016, pp. 297-301, doi: https://doi.org/10.1109/CSITSS.2016.7779374.
2. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1800-1807, Jul. 2017. https://doi.org/10.48550/arXiv.1610.02357
3. K. Smelyakov, A. Chupryna, D. Yeremenko, A. Sakhon and V. Polezhai, "Braille Character Recognition Based on Neural Networks," 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP), Lviv, Ukraine, 2018, pp. 509-513, doi: https://doi.org/10.1109/DSMP.2018.8478615.
4. K. M. Naimul Hassan, S. K. Biswas, M. S. Anwar, M. S. Iman Siam and C. Shahnaz, "A Dual-Purpose Refreshable Braille Display Based on Real Time Object Detection and Optical Character Recognition," 2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON), Dhaka, Bangladesh, 2019, pp. 78-81, doi: https://doi.org/10.1109/SPICSCON48833.2019.9065110.
5. A. A. Choudhury, R. Saha, S. Z. Hasan Shoumo, S. Rafsun Tulon, J. Uddin and M. K. Rahman, "An Efficient Way to Represent Braille using YOLO Algorithm," 2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Kitakyushu, Japan, 2018, pp. 379-383, doi: https://doi.org/10.1109/ICIEV.2018.8641038.
6. S. M. Qaisar, R. Khan and N. Hammad, "Scene to Text Conversion and Pronunciation for Visually Impaired People," 2019 Advances in Science and Engineering Technology International Conferences (ASET), Dubai, United Arab Emirates, 2019, pp. 1-4, doi: https://doi.org/10.1109/ICASET.2019.8714269.
7. V Jha and K Parvathi, "Braille Transliteration of hindi handwritten texts using machine learning for character recognition", Int J Sci Technol Res, vol. 8.10, pp. 1188-1193, 2019. https://www.semanticscholar.org/paper/Braille-Transliteration-Of-Hindi-Handwritten-Texts-Jha-Parvathi/195bf27cb108bb7c1a67a355430f8c55c2225a5f
8. C. Park, S. W. Cho, N. R. Baek, J. Choi and K. R. Park, "Deep Feature-Based Three-Stage Detection of Banknotes and Coins for Assisting Visually Impaired People," in IEEE Access, vol. 8, pp. 184598-184613, 2020, doi: https://doi.org/10.1109/ACCESS.2020.3029526.
9. G. C. Bettelani, G. Averta, M. G. Catalano, B. Leporini and M. Bianchi, "Design and Validation of the Readable Device: A Single-Cell Electromagnetic Refreshable Braille Display," in IEEE Transactions on Haptics, vol. 13, no. 1, pp. 239-245, 1 Jan.-March 2020, doi: https://doi.org/10.1109/TOH.2020.2970929.
10. M. A. Khan, P. Paul, M. Rashid, M. Hossain and M. A. R. Ahad, "An AI-Based Visual Aid With Integrated Reading Assistant for the Completely Blind," in IEEE Transactions on Human-Machine Systems, vol. 50, no. 6, pp. 507-517, Dec. 2020, doi: https://doi.org/10.1109/THMS.2020.3027534
11. B.-M. Hsu, ''Braille recognition for reducing asymmetric communication between the blind and non-blind,'' Symmetry, vol. 12, no. 7, p. 1069, Jun. 2020, doi: https://doi.org/10.3390/SYM12071069.
12. P. Kaur, S. Ramu, S. Panchakshari, and N. Krupa, ''Conversion of Hindi Braille to speech using image and speech processing,'' in Proc. IEEE 7th Uttar Pradesh Section Int. Conf. Elect., Electron. Comput. Eng. (UPCON), Nov. 2020, pp. 1–6, doi: https://doi.org/10.1109/UPCON50219.2020.9376566.
13. S. Shokat, R. Riaz, S. S. Rizvi, A. M. Abbasi, A. A. Abbasi, and S. J. Kwon, ''Deep learning scheme for character prediction with position-free touch screen-based Braille input method,'' Hum.-centric Coput. Inf. Sci., vol. 10, no. 1, pp. 1–24, Dec. 2020, doi: https://doi.org/10.1186/s13673-020-00246-6.
14. J. Mao, J. Zhu, X. Wang, H. Liu, and Y. Qian, ''Speech synthesis of Chinese Braille with limited training data,'' in Proc. IEEE Int. Conf. Multimedia Expo (ICME), Jul. 2021, pp. 1–6, doi: https://doi.org/10.1109/icme51207.2021.9428160
15. T. Kausar, S. Manzoor, A. Kausar, Y. Lu, M. Wasif and M. A. Ashraf, "Deep Learning Strategy for Braille Character Recognition," in IEEE Access, vol. 9, pp. 169357-169371, 2021, doi: https://doi.org/10.1109/ACCESS.2021.3138240.
16. S. Ramachandran, N. Rajan, K. N. Pallavi, J. Subashree, S. Suchithra and B. Sonal, "Communication Device for the Visual and Hearing Impaired Persons to Convert Braille Characters to English Text," 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2021, pp. 587-592, doi: https://doi.org/10.1109/ESCI50559.2021.9396859.
17. M. A. Hussain, R. I. Rifat, S. B. Iqbal, S. Biswas, M. G. Rabiul Alam and M. T. Reza, "Deep Learning based Bangla Voice to Braille Character Conversion System," 2022 IEEE 13th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 2022, pp. 0262-0267, doi: https://doi.org/10.1109/IEMCON56893.2022.9946619.
18. Bhatia, Surbhi & Devi, Ajantha & Alsuwailem, Razan & Mashat, Arwa. (2022). Convolutional Neural Network Based Real Time Arabic Speech Recognition to Arabic Braille for Hearing and Visually Impaired. Frontiers in Public Health. https://doi.org/10.3389/fpubh.2022.898355.
19. Ramani RG, Shanthamalar JJ. Automated image quality appraisal through partial least squares discriminant analysis. Int J Comput Assist Radiol Surg. 2022 Jul;17(7):1367-1377. https://doi.org/10.1007/s11548-022-02668-2
20. Seung-Bin Park and Bonghyun Kim. "Design of CNN-based Braille Conversion and Voice Output Device for the Blind" Journal of Internet of Things and Convergence 9, no.3 (2023) : 87-92.doi: https://doi.org/10.20465/KIOTS.2023.9.3.087.
21. Kumar, A. (2023). Study and analysis of different segmentation methods for brain tumor MRI application. Multimedia Tools and Applications, 82(5), 7117-7139. https://doi.org/10.1007%2Fs11042-022-13636-y

22. Dhyani, S., Kumar, A., & Choudhury, S. (2023). Arrhythmia disease classification utilizing ResRNN. Biomedical Signal Processing and Control, 79, 104160. https://doi.org/10.1016/j.bspc.2022.104160

AUTHORS

Aditya Upadhye is currently pursuing his B.E. degree in Computer Science and Engineering from Sathyabama Institute of Science and Technology, Chennai. His areas of interests are Distributed Systems, Cloud Computing, System Security and Machine Learning.

Corresponding Author Email: adityaupadhye47@gmail.com

Tharunnesh R is currently pursuing his Bachelor of Engineering program in computer science at Sathyabama Institute of Science and Technology in Chennai, India in 2024. His academic pursuits are focused on the fields of Data Science, Artificial Intelligence, and Machine Learning. Email: imtharunr@gmail.com

J Jeslin Shanthamalar is currently working as Assistant Professor in the Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai. She completed her PhD program in the Department of Information Science and Technology, Anna University, Chennai, India, as a full-time scholar. She has more than 10 years of experience in Teaching, Research and Industry. Her areas of specialization include Data Mining, Image Processing, Machine Learning, Deep Learning and Medical Image Analysis. She has published more than 10 papers in reputed journals, patents and conferences.

Email: jeslin.cse@sathyabama.ac.in

C A Daphine Desona Clemency is currently working as Assistant Professor in the Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai. She is pursuing her PhD in the field of Machine Learning and Deep Learning. She has more than two years of experience in Teaching and Research. She has published a few patents and papers in conferences and journals.

Email: daphine.cse@sathyabama.ac.in

N S Usha is currently designated as an Associate Professor, Department of Computer Science and Engineering at Sathyabama Institute of Science and Technology, Chennai. She completed her B.E. degree in Computer Science and Engineering at Crescent Engineering College, Madras University in 2002 and her ME degree in Computer Science and Engineering at S.A. Engineering College, Anna University in 2008. She was awarded with PhD degree in Computer Science and Engineering from St. Peter's Institute of Higher Education and Research, in the year 2022. She has got nearly 20 years of teaching experience in Engineering. She has published 5 Design patterns, 9 reputed journals and many papers in both National and International conferences. She also organized and coordinated many Guest Lectures and served as an eminent Guest Speaker for many events. Her current area of research includes Network Security, Cyber Security, IoT, Artificial Intelligence, Machine Learning, Deep Learning, Computer Vision and Data Science.

Email: ushans.cse@sathyabama.ac.in