Bangla Book Reader For the Blind

Nazifa Hamid*, Shariar Azad*, Muhammad Zubair Hasan*, Shahir Abdullah Bin Shahnoor*, Takik Hasan*
Department of Computer Science and Engineering
Military Institute of Science and Technology*
Dhaka, Bangladesh
Email: nhrafa.96@gmail.com, ShariarRiday@gmail.com, mdzh1997@gmail.com, shahir2561@gmail.com, takikhasansa@gmail.com

***Abstract*—Recognizing text-image from mobile phones is a challenging task for their limited capacity and processing power and also the accuracy of the system is very important. In this system, we aimed to develop a Text-Image Recognition System for mobile environment using custom Bangla word dataset. Firstly, we trained a light neural network using the dataset and created a cloud API. The image captured from mobile phone's camera then sent to the cloud and it's converted to text and then to speech. This approach of detection process makes the process much faster. There are quite a few e-book readers available which are mostly in English and also expensive. The goal of this study is to develop an inexpensive Bangla e-book reader with high accuracy for the blind in order to give them easier access to Bangla printed books. .**

I. Introduction

Many among us are not blessed with the same gifts and abilities that we think as natural. Among them one is the power to see. To perceive the world that we live in, to take in the immense bundle of beauty and knowledge the world provides one has to be able to see. Without having this ability, many people are deprived of the knowledge that comes from books. In fact, just to live by one often has to be able to read papers or documents which is not possible for the blind. There are about 45 million people in the world who are completely blind and about 135 million visually impaired according to the World Health Organization (WHO) [1]. According to a survey done by non-profit organizations 800,000 of the Bangladeshi people are affected by blindness [2].

Although there is the braille system but it is not available for all documents we daily use such as receipts, bank statements, newspapers and numerous others. Braille is also not available for many books required for academics, especially- in the Bangla language. From restaurant menus to food cans, some parts of the world have braille included in many daily used products but that is not the case for Bangladesh. Also, according to studies the braille system is prone to error during reading and production because of low redundancy. Complexity of braille signs makes reading speed significantly slower and hard to learn for the newly blind adults. It also requires a lot of paper and space making documents in braille heavy to carry [3]. While in the outside world there are various mobile apps and devices that can indeed scan and read documents out loud, many of them are very expensive and none of them are available for the Bangla language. Therefore, blind students in Bangladesh greatly struggle and fall back compared to their sighted peers.

Our aim is to build a device that a blind student or anyone as a matter of fact can use to read a printed Bangla document or book by scanning it. Since blind people primarily have to depend on their friends and family for this, a system that can read these documents out loud and is easy to use will give them much more independence in their daily life. Considering these points, we will develop a system that is easy to use, inexpensive and designed to suit their needs.

II. Literature Review

There have been many works done regarding the application of technology in reading tools for the blind. Among them many are remarkable. One approach developed by A. Kulkarni [4] braille is used by building an e-book reading device which has two electronic refreshable braille display units and uses SD card for file storage.

Shilkrot R et al. [5] developed a device that user wears on their finger and it reads as user touches the printed text document on the go. It scans the text sequentially and the words are synthesized to audible speech, it also provides continuous feedback to user. Another finger-based approach developed by Lee Stearns et al. [6] called the HandSight enables readers to access documents by mounting a tiny camera on their finger and tracing along the text.

A Ravi and Sk Khasimbee [7] proposed a smart reader with the integration of Raspberry Pi in which user places book under a camera that takes picture of the page and converts text to audible speech. This system includes a page turning mechanism, dictionary and takes audio commands. A survey conducted by Geetha M.N. Et al. [8] on smart readers shows many drawbacks of the existing smart reader systems such as alignment, focus, accuracy, mobility, efficiency and proposes new smart reader which overcomes some of these problems.

Dadhich A. and Dutta K. [9] have provided a solution to convert text from printed document to digital format by detecting text with a text recognition API in real time and have also proposed an algorithm for the text selection and aggregation.

Another smart portable reader that makes use of braille is developed by R. Velázquez et al. [10] His device consists of a software that translates any eBook into Braille very fast which is then read by a blind person using a braille terminal but this device is only applicable for eBooks.

Notable work done in case of Bangla character recognition are a model developed by of A. Fardous and S. Afroge [11] to recognize handwritten Bangla compound characters using Convolutional Neural Network, S. Basu et al. [12] cused an MLP based classifier to recognize Bangla characters using a 76-element feature set. Keeping all related works in consideration we will try to build a system that is user friendly and inexpensive.

III. Design and Development

The hardware part of this project is comprised of the Arduino, the Bluetooth module that transmits button press to the android application and the book stand. The software part consists of both the android application and custom server implementation for deep learning inference. The server is used to process the image and extract the text from the text and then generate speech from this text.

Figure 1 shows the main flow diagram of the integrated hardware device and software based mobile application. User has to start by placing the book under a mobile stand and the phone on the mobile holder, then take picture of the book and use the mobile application to upload the image to the server. Image is then processed, text is extracted and converted to audio. If user wishes to listen to the audio, the audio in the server is converted to mp3 file, downloaded to the mobile application and played.

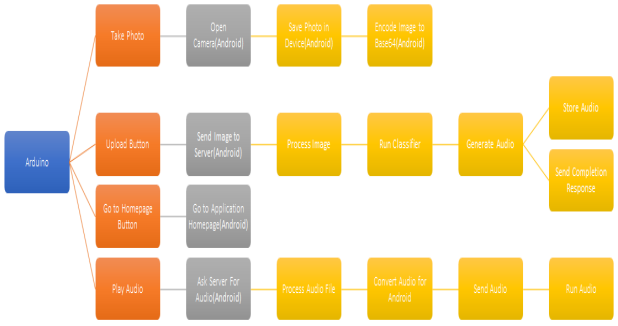

Fig. 1: A Conceptual Diagram for the mobile application

Figure 2 shows the system architecture of the hardware part where the inputs from the pressing of buttons that are attached to the stand are processed in the Arduino and sent to the the mobile application through the Bluetooth module. Along with the mobile application user can use these buttons to take photo, scan and listen too. This use of external buttons is kept to make the device easy to use.

The book stand is an important part of our project because from our feasibility study we found out that blind individuals have difficulties searching buttons in the UI of android applications. Figure 3 is an image of our Book Reader prototype. We made our system in such a way that a blind user will be

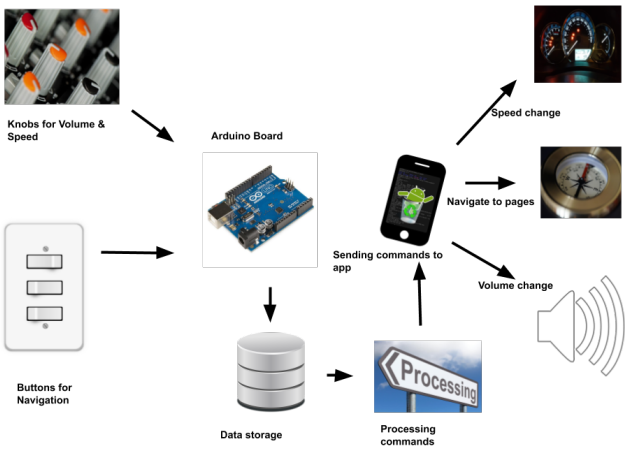

Fig. 2: System Architecture of the Hardware Part

able to place the books and their android device on the stand and through physical buttons and voice commands very easily take pictures of the books. This helps them to use the system very easily and efficiently.

Fig. 3: Prototype of the Book Reader for the Blind

The image is processed using OpenCV's EAST Text Detector. This crops all the words in the given image. Then all the individual words are sent to the efficientnet-b0 model which is a very lightweight but powerful model to classify all the words from this text. The accuracy of this model is 85%. We get the words for each image from this model. These words are then

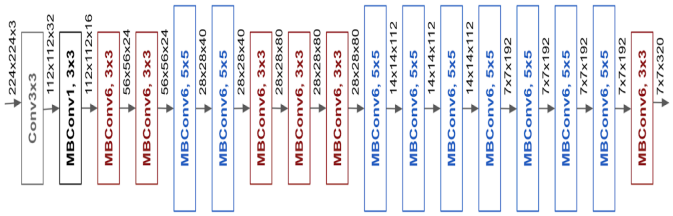

Fig. 4: RNN network architecture

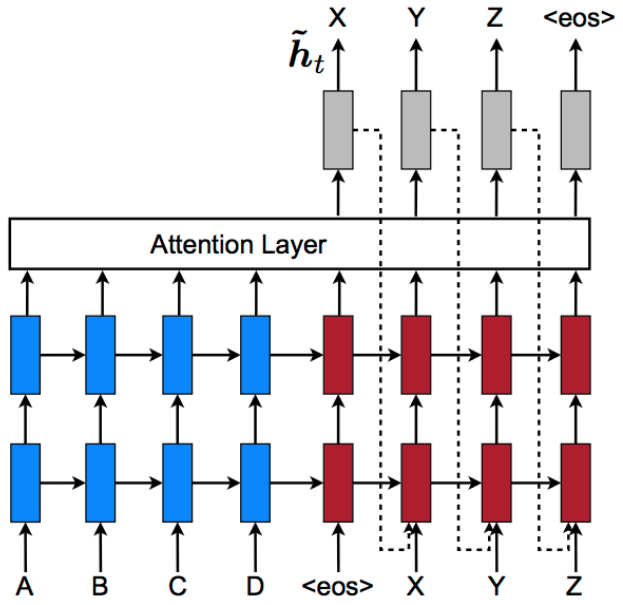

Fig. 5: Attention Layer

saved into a text file and after all the words are classified the text file is sent to the text to speech module which generates the audio. The audio is then generated using a RNN network with attention.

Figure 4 and 5 show the architectures of the RNN network along with attention. The Android application is used to take a picture and the picture is sent to the server encoded as a base64 string. The string is then decoded in the server and further work is done as mentioned above. Application also sends a GET request to server which starts buffering the audio file from server. Figure 6 and 7 show the graphical representations of epochs vs loss and accuracy for our text recognition model.

IV. Experiment Results

As mentioned earlier the text recognition accuracy of our model is 85%. The custom dataset contained about 3000 Bangla words. During testing, each page we tested on, contained about 250 to 305 pages.

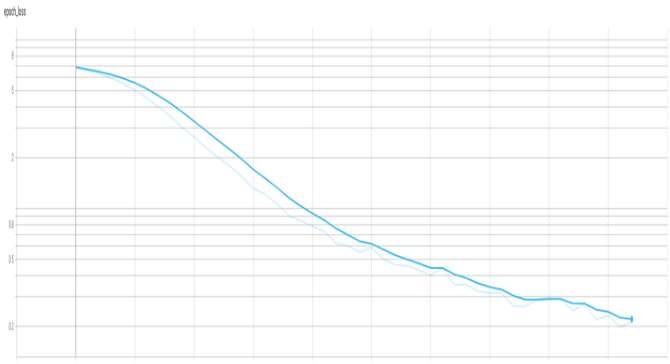

Fig. 6: Epochs vs Loss graph

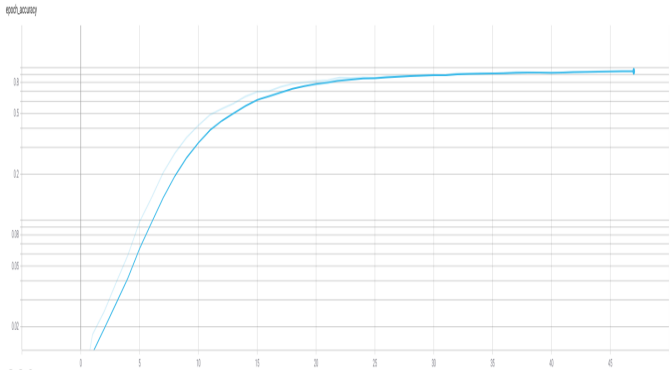

Fig. 7: Epochs vs Accuracy graph

Page No.	Total words	Correctly read	Failed to read
Page 1	261	221	40
Page 2	296	254	42
Page 3	275	236	39
Page 4	303	258	45
Page 5	278	238	40

1) page 1 contained 261 words. The device read 221 words correctly and failed to read 40 words.
2) Page 2 contained 296 words. The device read 254 of them correctly and failed to read 42 words.
3) Page 3 contained 275 words. The device read 236 of them correctly and failed to read 39 words.
4) Page 4 contained 303 words. The device read 258 correctly and failed to read 45 words.
5) Page 5 contained 278 words. The device read 238 correctly and failed to read 40 words.

V. Future Scope

We used our personal computers as servers. For this reason, the response time is slow. It takes about 3-4 minutes for text to speech conversion and download of the audio to app. A better server will give better response time. For the same reason it is not possible to deliver real time audio which can be incorporated in thhe future. Our dataset is not very robust and therefore, text detection accuracy is not state of the art.

We could make the model more accuratein the future. There is no way to verify if placement of book/document is correct the first time. Meaning, if the book placement is not correct image will not be captured correctly resulting in bad scan. A page detection system may be added to sort out this issue. Also, we have not added voice commands which also can be added in the future.

VI. CONCLUSION

Most blind students of Bangladesh still rely on the braille to read. Many books required in the studies of undergraduate or other do not have the braille version. So, students largely have to rely on their family members or friends when they need to read any printed text document in Bangla. Book reader for the blind is portable, inexpensive and an easy-to-use solution to this problem. It is an attempt to help them become more self-dependent which is designed to suit their needs.

REFERENCES

[1] W. H. O. P. for the Prevention of Blindness and Deafness, "Global initiative for the elimination of avoidable blindness," 1997.

[2] [Online]. Available: http://www.seva.org/pdf/Seva_Country_Fact_Sheets_Bangladesh.pdf

[3] M. Tobin and E. Hill, "Is literacy for blind people under threat? does braille have a future?" *British Journal of Visual Impairment*, vol. 33, pp. 239–250, 09 2015.

[4] A. Kulkarni and K. Bhurchandi, "Low cost e-book reading device for blind people," in *2015 International Conference on Computing Communication Control and Automation*, 2015, pp. 516–520.

[5] R. Shilkrot, J. Huber, M. E. Wong, P. Maes, and S. Nanayakkara, "Fingerreader: A wearable device to explore printed text on the go," 03 2015, pp. 2363–2372.

[6] L. Stearns, R. Du, U. Oh, C. Jou, L. Findlater, D. Ross, and J. Froehlich, "Evaluating haptic and auditory directional guidance to assist blind people in reading printed text using finger-mounted cameras," *ACM Transactions on Accessible Computing*, vol. 9, 10 2016.

[7] A. Ravi, S. Khasimbee, T. Asha, T. N. S. Joshna, and P. G. Jyothirmai, "Raspberry pi based smart reader for blind people," in *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 2020, pp. 445–450.

[8] G. M.N., S. H.V., S. S., A. Siddiqa, and C. H.C., "Survey on smart reader for blind and visually impaired (bvi)," *Indian Journal of Science and Technology*, vol. 12, pp. 1–4, 12 2019.

[9] A. Dadhich and K. Dutta, *Real-Time Printed Text Reader for Visually Impaired*, 01 2018, pp. 327–337.

[10] A. Kulkarni and K. Bhurchandi, "Low cost e-book reading device for blind people," in *2015 International Conference on Computing Communication Control and Automation*, 2015, pp. 516–520.

[11] A. Fardous and S. Afroge, "Handwritten isolated bangla compound character recognition," in *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 2019, pp. 1–5.

[12] S. Basu, N. Das, R. Sarkar, M. Kundu, M. Nasipuri, and D. Basu, "Handwritten bangla alphabet recognition using an mlp based classifier," 03 2012.