

CHITTAGONG UNIVERSITY OF ENGINEERING AND TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
CHITTAGONG-4349.

(Thesis/Project Proposal)

Application for the Approval of B. Sc. Engineering Thesis/Project
(Computer Science & Engineering)

Date: 25.02.2020

- | | |
|---|---|
| 1. Name of the Student | : Fahim Shahriar |
| Student ID | : 1504006 |
| Session | : 2018-2019 |
| 2. Present Address | : Room No. 320, Shaheed Tareq Huda Hall,
Chittagong University of Engineering & Technology |
| 3. Name of the Supervisor | : Tanusree Debi |
| Designation | : Lecturer,
Department of Computer Science & Engineering,
Chittagong University of Engineering & Technology |
| 4. Name of the Department | : Computer Science and Engineering |
| Program | : B.Sc. Engineering |
| 5. Date of First Enrollment
in the program | : February 25, 2016 |
| 6. Tentative Title | : Developing an OCR based system for visually
impaired people to read printed Bangla texts from
image |

7. Introduction

According to the official statistics from the World Health Organization, there are about 285 million visually impaired persons in the world up to the year of 2011; about 39 million are completely blind and 246 million have weak sight. In the context of Bangladesh, 7.5 million people who are aged more than 30 are blind and 1.5 million of children are suffering from low vision.

Blindness is a problem because it's increasing tremendously and technology isn't made for the blind, learning has to be altered, and is expensive.

Blind people are unable to perform visual tasks. For instance, a simple process 'learning from a book' requires the availability of the book in braille or a digital speech synthesizer (if the text is available in digital format). The majority of published printed works does not include braille or audio versions, and digital versions are still a minority.

The number of blind people will increase rapidly as the baby boomer generation ages. These visually impaired people have great difficulty in perceiving and interacting with the surroundings. Existing modes of education and visual training are expensive and almost never affordable for someone with a minimum wage. Thus, more often than not, visually impaired children from low income families don't even get the opportunity of receiving primary education, let alone explore further academically. It is often seen that visually impaired people become very proficient in both academic and job sector if they can overcome the obstacle of the primary education.

This project will smoothen the process of learning for the visually impaired people to a great extent. This project mainly comprises of two parts: an OCR (Optical Character Recognizer) and a speech synthesizer. The primary target is to make the text of an image audible to a visually impaired person. To get the audio output we must give a text file as an input. This text file is acquired from the image with the help of OCR.

Bangla is an Indo-Aryan language with approximately 228 million native speakers and another 37 million as second language speakers. Bengali is the fifth most spoken native language and the seventh most spoken language by total number of speakers in the world. Bangla characters are complex as the components are connected with a horizontal line as 'Matra' and there is a great number of modifiers as 'Kar' and 'Fola' and also compound letters as 'Juktoborno'. For many years a lot of attempts are made to make Bangla characters recognizable by an OCR. But for the aforesaid reasons the output was not so satisfactory.

In this project an OCR based approach for Bangla text is done to detect printed texts from an image. With a speech synthesizer, the extracted texts will be then read to the user; a visually impaired person. It will definitely ease the process of learning in Bangla for the visually impaired people.

8. Background and Present State of the System

Visual impairment has long been a great hindrance towards accessible education. If an aiding system for Bangla literature can be created it will ease the education for the visually impaired people.

Venkatesh et al [1] created a system to detect and extract text from image and then converted the text file to an audio. But it only extracted English texts. By applying OCR operation on an image file the text was extracted and a speech synthesizer gave output to an audio file.

Haque et al [2] created a system to aid the visually impaired people by extracting texts from road signs and converted them to audio. In this, Bangla text is extracted from road signs and converted to speech for the user. Here an outdoor image is taken and several pre-processing steps are performed on it to detect the region containing text. Then the text is extracted with the help of OCR technique. The extracted text is then converted to audio by a TTS system.

Another work was done by Rithika et al [3] where the texts in an image is read to a person in his desired language. Here the text is first extracted and later the extracted text is translated and read in the desired language. The translation and reading part is done using the google translator service. So, it needs an active internet connection to work. The purpose of accuracy of the meaning is depended on google translator, which is not fully dependable.

Visually impaired people generally read using Braille system. Now-a-days it's not being taught widely in schools, is not popular with parents of blind children, and social services departments often do not have the budget to help adults learn. Due to several technological developments like OCR, TTS, Image processing etc. it is becoming redundant day by day.

In our country visually impaired people have shown great proficiency in digital sectors after they were properly educated and trained. As there is almost no sustainable technological alternative to learn in Bangla, still most of them are deprived from proper education and training. With this project this obstacle in learning will be removed to a great extent.

9. Aims with Specific Objectives and Possible Outcomes

The proposal work will be carried out with view to achieve the following objectives:

- I. To design an efficient OCR to extract Bangla text from images.
- II. To develop a speech synthesizer to read aloud extracted text
- III. To implement the system as a dedicated device to ensure usability for a visually impaired person

10. Outline of Methodology

The working methodology for the proposed system is shown in Fig 1.

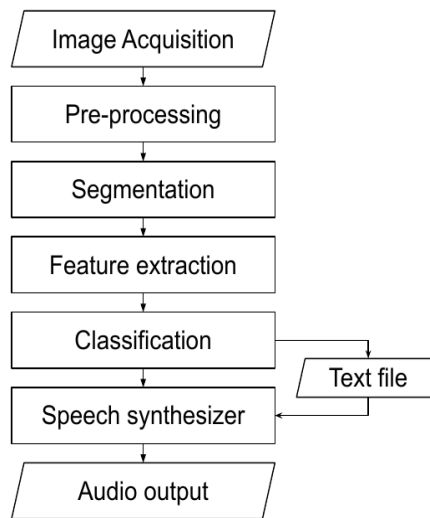


Fig 1: Methodology of the Proposed System

- **Image Acquisition:**

To start the full working methodology an image full of printed text will be needed and in this step the image will be acquired using the pi camera module. After the image is acquired the next step is to do some pre-processing on the image.

- **Pre-processing:**

The pre-processing step includes several operations to be done on the image. For an OCR system the accuracy of an OCR greatly depends on the pre-processing step. The pre-processing step includes:

- **Grayscale conversion:** As the first step of pre-processing, the image is converted to a grayscale image to reduce color complexities.
- **Noise removal:** Several noises might be present in the image. A bilateral filter is applied to reduce noises [4].
- **Binarization:** Binarization is done to make the image two-toned. Binarization resolves any image into foreground and background. As we are considering this for the printed texts it will easily distinguish text from the background. For Binarization several local methods [5] and Otsu method [6] can be used.
- **Skew detection and correction:** Necessary steps are taken here to detect the skew angle of the image acquired. Then the image is rotated accordingly to fix the skew angle.

- **Segmentation:**
This is the step where the raw text is identified and is differentiated into segments (eg. lines, words, characters etc). So basically segmentation has to be done in three levels:
 - **Line segmentation:** In printed Bangla scripts, text lines and sizes are almost of the same height. There is a small gap between two lines. These gaps can be easily detected by horizontal projection profile (HPP) [7]. Text lines are then easily separated by detecting the white spaces in HPP.
 - **Word segmentation:** Vertical projection profile (VPP) can be used in word segmentation from extracted lines [7]. In this process each line will be scanned vertically and number of black pixels in each column will be calculated. If no black pixel is found then that is considered as the gap between words.
 - **Character segmentation:** Character segmentation is the most difficult part of overall segmentation process. Two major properties of Bangla script which make it more difficult and erroneous in case of character level segmentation are – Matra line and Character overlap. Here, the matra line will be removed analyzing the frequency of black pixels in each rows of a line. Removing the matra lines will make all the characters disconnected which will result in a simpler scenario for character segmentation[7].
- **Feature extraction:** Feature extraction provides an important role in character recognition. It is a very difficult task to recognize a character. But, choice of good features significantly improves the recognition rate and minimizes the error in case of noise. In this stage a connected component based approach will be made to extract different features of a character.
- **Classification:** For classification based on the features selected a neural network is a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use. First knowledge is acquired by the network through a learning process and then storing knowledge is used for the recognition of character. After completion of this stage the output texts will be stored in a simple text file.
- **Speech synthesizer:** Speech synthesizer will take the text file and give an output in audio format.

11. Resources Required Completing the Work

- **Hardware**
 - 1) Personal Computer
 - 2) Raspberry pi
 - 3) Raspberry Pi Camera Module
 - 4) GPIO extension board
 - 5) Speaker
 - 6) Battery and a HAT board.

- **Operating System**
 1. Linux (Manjaro)
 2. Raspbian (Linux-based)
- **Platform**
 - Python

12. Cost Estimation

The costs that will occur to implement our proposed system are estimated below.

Material	Cost (BDT)
Raspberry pi 4	6200
Raspberry pi camera module	2600
Battery expansion board	1500
5.1v Battery	1000
GPIO extension board	500
A general speaker	500
Drafting	100
Binding	100
Printing	100
Total	= 12600

13. References

- [1] Venkatesh, R., Kumar, K. V. & Sudhakaran, M., “Optical character recognition based speech synthesis system using Raspberry pi”, International Research Journal in Advanced Engineering and Technology (IRJAET) pp. 3450 – 3455, 2018.
- [2] Haque, S. M. A., Arbi, S., Tamanna, T and Itu, S. M. (2007) “Automatic detection and translation of Bengali text on road sign for visually impaired”, Daffodil International University Journal of science and technology, Vol. 2, Issue 2.
- [3] Rithika, H. & Santhosi, B. N. (2016) “Image text to speech conversion in the desired language by translating with Raspberry Pi”, 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), pp. 1-4.
- [4] https://en.wikipedia.org/wiki/Bilateral_filter
- [5] John C. Russ, "The image processing handbook", CRC Press, Boca Raton, FL, USA, 1998.
- [6] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms", IEEE Transactions on Systems, Man, and Cybernetics, 1979.
- [7] Zahan, T., Selim, M. R., Rahman, M. S. & Iqbal, M. Z., “Connected component analysis based two zone approach for Bangla character segmentation”,

- International Conference on Bangla Speech and Language Processing (ICBSLP), pp. 1-4, 2018.
- [8] Hasnat, A.M., Habib, M. S. M., & Khan, M., "A high performance domain specific OCR for Bangla script", Int. Joint Conf. on Computer, Information, Systems, Sciences and Engineering (CISSE) , 2007.
 - [9] Ome, F. Y., Himel, S. S. & Bikas, M. A. N., "A complete workflow for development of Bangla OCR", International Journal of Computer Applications (0975 – 8887), Volume 21-(9), 2011.
 - [10] Chaudhuri, B.B. & Pal, U., "A complete printed Bangla OCR system", Pattern Recognition, Vol. 31-(5), pp. 531-549, 1998.
 - [11] Hasnat, M. A., Chowdhury, M. R. & Khan, M., "Integrating Bangla script recognition support in Tesseract OCR", Proc. of the Conference on Language and Technology 2009 (CLT09), Lahore, Pakistan, 2009.
 - [12] Smith, R., "An overview of the Tesseract OCR engine", Proc. of ICDAR, Vol. 2, pp. 629-633, 2007.
 - [13] Hasnat, M. A., Chowdhury, M. R. & Khan, M., "An open source tesseract based optical character recognizer for bangla script", Proc. 10th Int. Conf. Document Anal. Recogn., pp. 671-675, 2009.
 - [14] [http://elinux.org/RPi_Text_to_Speech_\(Speech_Synthesis\)](http://elinux.org/RPi_Text_to_Speech_(Speech_Synthesis))
 - [15] https://en.wikipedia.org/wiki/Visual_impairment
 - [16] <https://en.wikipedia.org/wiki/Braille>
 - [17] <https://www.classycyborgs.org/braille-literacy-statistics-india/>
 - [18] www.raspberrypi.org
 - [19] <http://www.zdnet.com/article/raspberry-pi-11-reasons-why-its-the-perfect-small-server/>
 - [20] <http://aishack.in/tutorials/opencv/>
 - [21] <http://www.zdnet.com/article/raspberry-pi-11-reasons-why-its-the-perfect-small-server/>
 - [22] <http://aishack.in/tutorials/opencv/>
 - [23] http://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_setup/py_intro/py_intro.html
 - [24] <http://hackaday.com/2016/02/28/introducing-the-raspberry-pi-3/>
 - [25] http://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_setup/py_intro/py_intro.html
 - [26] <http://hackaday.com/2016/02/28/introducing-the-raspberry-pi-3/>

14. CSE Undergraduate Studies (CUGS) Committee Reference:

Meeting No.:

Resolution No.:

Date:

15. Number of Under-Graduate Student(s) working with the Supervisor at Present:

Signature of the Student

Signature of the Supervisor

Signature of the Head of the Department