IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Book Reader TTS Website

Jayant Umesh Shah
Student
Department of Computer Engineering
Pillai HOC College of Engineering and Technology
Mumbai, India

Aniket Shinde
Student
Department of Computer Engineering
Pillai HOC College of Engineering and Technology
Mumbai, India

Dipesh Parab
Student
Department of Computer Engineering
Pillai HOC College of Engineering and Technology
Mumbai, India

Abstract - The system aims to aid the visually impaired students to understand textual data. The objective of the system is to read data from the given input document and convert it into audio format using TTS technology. The motivation to design the system is to show that student's disability shouldn't be a major reason to hold them from pursuing their education and goals.

Index Terms— ***TTS-Text to Speech, OCR- Optical Character Recognition, Open CV- Open source Computer Vision, No SQL- Non Structured Query Language WHO- World Health Organization, Graphemes, Phoneme***

I. INTRODUCTION

Computers can perform the operations based on given instructions and the users can read the result of the same on the monitors or by printing the output. But there are situations where people need vocal communication with the computers. The TTS is a system for conversion of text documents to audio. One of the main applications of a TTS system is to help the people with visual problems / blind people and making them understand the text and number concept present on the system.

As stated by the statistics of the WHO-World Health Organization, 285 million people around the world are estimated to be visually impaired, of which 39 million are complete blind, out of which 90% live in growing countries. Thus, there is a need to develop a book reader device that is affordable to the low income sections of the society.

People who suffer from sight and visual impairment are not capable reading text in ordinary newsprint, books and magazines clearly. This can make the reading process difficult which can disturb the learning process and slow the person's intellectual development. Therefore, there is a need of a device that can scan and read any kind of text by changing it into audio form to help them read. This system will help the visual impaired people to listen the document, academic textual content, diagrams and any other text-based source which they are not able to read.

The TTS technology helps users to read the written material by listening to the audio output generated by the system. In TTS system the text is taken as input and then
converts letter to sound and finally converting the sound to speech. Letter are the basic unit of writing system of a language.Sound are the smallest meaningful sound element of the language. The TTS system converts text to a sound representation and then converts the sound representation to waveforms that can be obtained in output as converted speech of the text entered . There have been many different approaches and algorithms to implement TTS conversion.

Text and diagrams are detected using the input, which could be in the form of PDF or IMAGE (JPEG). The extracted text and diagrammatical data will be provided to the TTS conversion component and as a result, the audio file will be generated which reads out the text using human synthesized AI voice present in TTS. Reading and writing text is inherently difficult for visually blind people.
The main objective of our proposed system is help the visually impaired students to face and to develop an application that reads the textual content and educational data from the given input and converts it into the audio format using TTS technology. The basic idea behind the proposed application is helping blind students, also the normal users; understand text better via

audio. Our objective is to show that student's disability shouldn't be the major reason to hold them from pursuing their education and goals.

II. LITERATURE SURVEY

In this section a reference to the relevant past literature that use the various text-to-speech techniques for conversion of mathematical data is mentioned. Most of the researchers concentrate on developing an efficient text-to-speech system that will convert any mathematical data like equations, formulas and problems into speech for blind person.

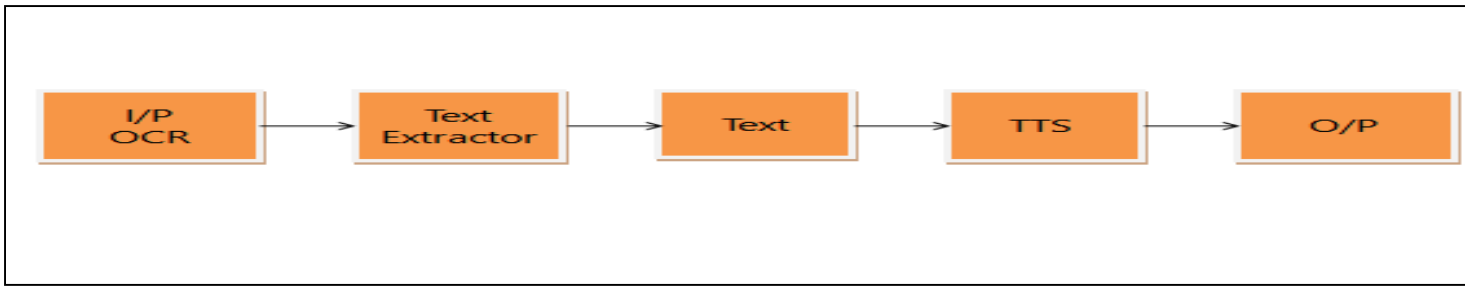
Sachin Kulkarni and Dr. Debajyoti Mukhopadhyay proposed a text-to-speech tool '*Math Says*', that will help legally blind students and also the normal users; understand mathematics better via audio form. The written mathematics material is examined as an image. Then that image is forwarded for the text detection. The detected text is then taken out from that image. The extracted text will be given to the TTS conversion component and as a result, the audio file will be generated which speaks the text in that image. The challenge in this work is how to tackle the formula which are not in proper English form and execute it properly . For this obstacle, the table will be maintained which has each formula and its corresponding text. Whenever the formula is encountered, the same will be searched in the table and the corresponding text will be extracted and spoken out loud. In this way, this system understand the mathematics concept .

K´evin VythelingumYannickEst`eve and Olivier suggested a routine to repeatedly detect letter-sound conversion errors by comparing equivalent phoneme hypothesis. A LBAC forced alignment system is implemented, allowing for signal dependent phoneme. He also implemented a sequence to sequence neural network model to obtain a context-dependent grapheme-to-phoneme conversion. On a French dataset, it was seen that it can detect upto 86.3% of the errors made by a commercial letter to sound system. Moreover, the amount of data interpreted as incorrect is kept under 10% of the total rectified data. The time spent for phoneme manual checking can thus been reduced without decreasing the phonemic file quality.

Manjare Anil, S.D. Shri. bahadurkar and Shaikh Shadab Shakil proposed a new method for a Devanagari (Marathi) Text To Speech system. As there is very less work has been done in Devanagari TTS systems especially for Marathi text. A new method is proposed with mapping and combining in order to compose the Marathi TTS system. The Marathi TTS system is build using existing English TTS Engine. Marathi text is to be compared with the text present in the database using simple Linear search algorithm then it is provided as input to the Existing English TTS. Contently Concatenative speech synthesis is the method mostly used in TTS systems. The proposed system introduces new concwhich is more feasible and easier than the earlier methods used for Marathi Text To-Speech. Also the proposed methodology supports high accuracy for text mapping.

Partha Mukherjee and Soumen Santra developed a useful TTS synthesizer in the form of a simple application that turns input text into merge speech and then reads out to user later on they can be stored as an mp3 file in the system . They developed a TTS synthesizer such that it converts text into spoken word, by analyzing and converting it using Natural Language Processing (NLP) and then using Digital Signal Processing (DSP) technology that helped in proper speech representation of the text.

III. PROPOSED FLOW SYSTEM



The primary goal is that the system is to convert the text into audio from. The system will take the PDF as an input and use the OCR system to convert the scanned book into text and then the same will be converted to audio form, later storing it in the database. The system is developed in such a way that it can be used by visually impaired students. The challenge was to develop a system making it useful to such users who are not capable to view the system due to sight disorder or poor vision power. The system's GUI is designed in such a way which makes the whole navigation process easy for such category of users. System consists of large tabs and voice guidance on mouse hover on every tab. A textbook surfing interface developed which will text aloud and having facility of moving back and forth using "previous page" functionality as well as "next page" functionality. Also reading desired page by giving page number as input is provided in the system.The system is constructed with fully voice guidance throughout all web pages, making it easier for the user to navigate. The system consists of a GUI so that the user can give different input using the mouse hover action to provide input actions such as play, pause, next page, previous page to interact with the system. The interface is designed by focusing mainly on visual impaired user making it very simple, responsive and mobile friendly so it can be used in

Algorithm:

- Start
- Input text
- OCR conversion and segregate input text
- Store it into the database and assign ID for relative text
- Parse the text with the suitable ID to convert into speech format
- Output is in human voice format for the given input text
- End

IV. RESULTS

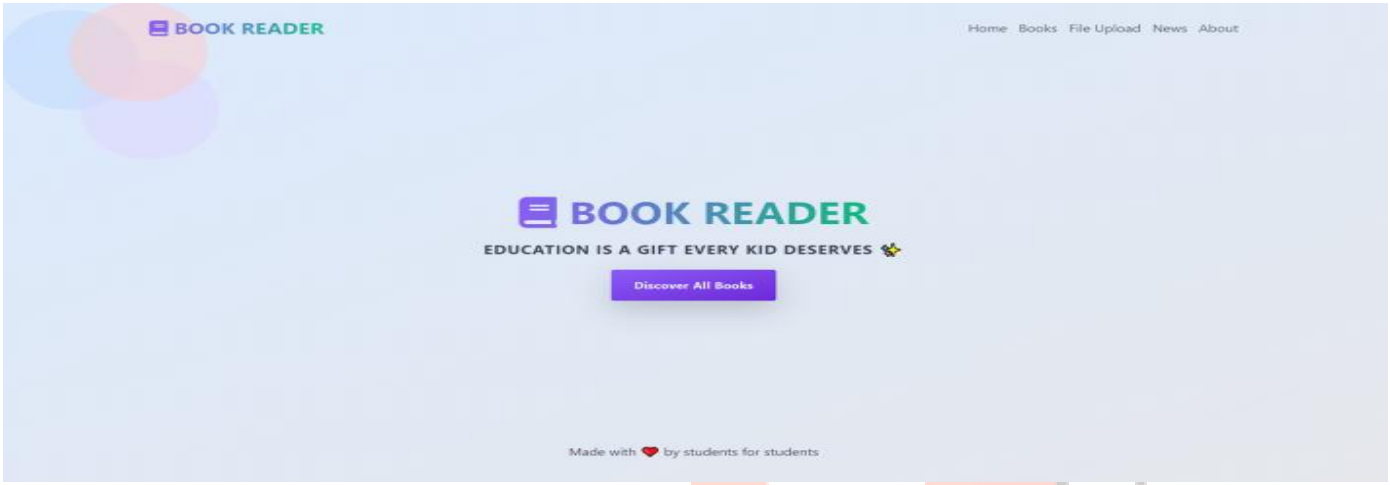


Fig 2. Home Menu Page

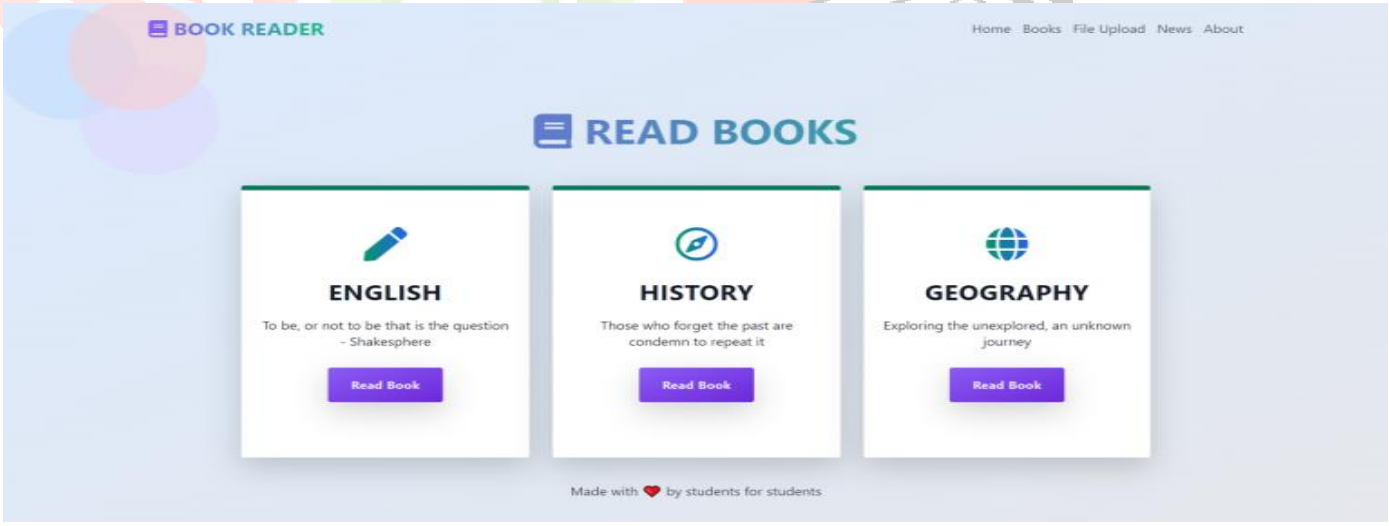


Fig 3. Read Books Page

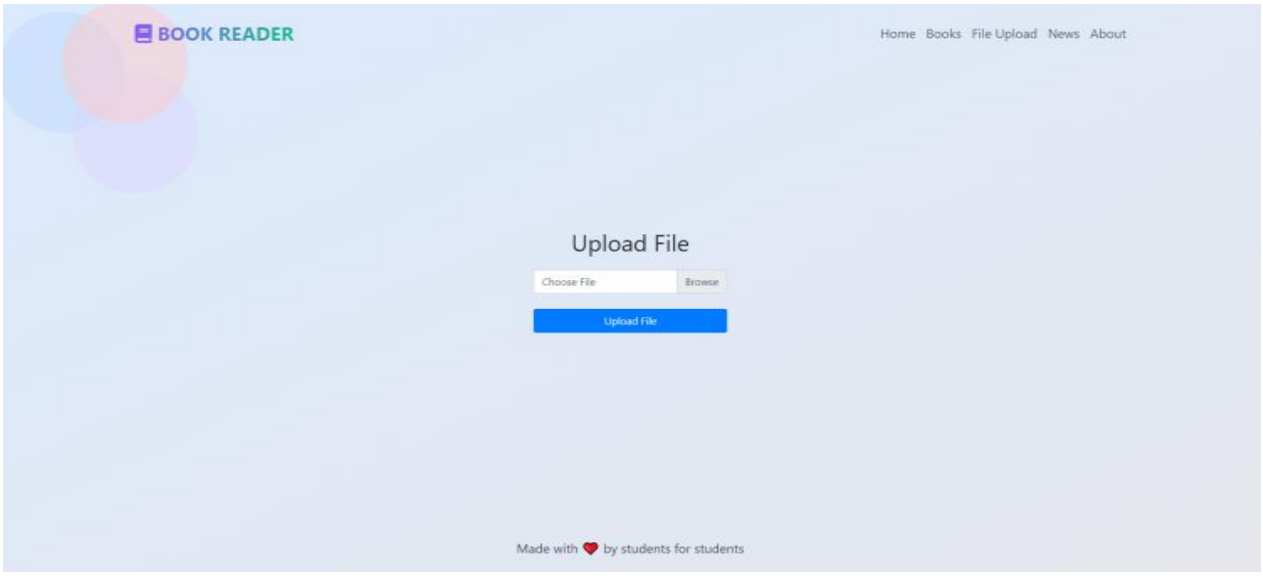


Fig 4. Upload Section

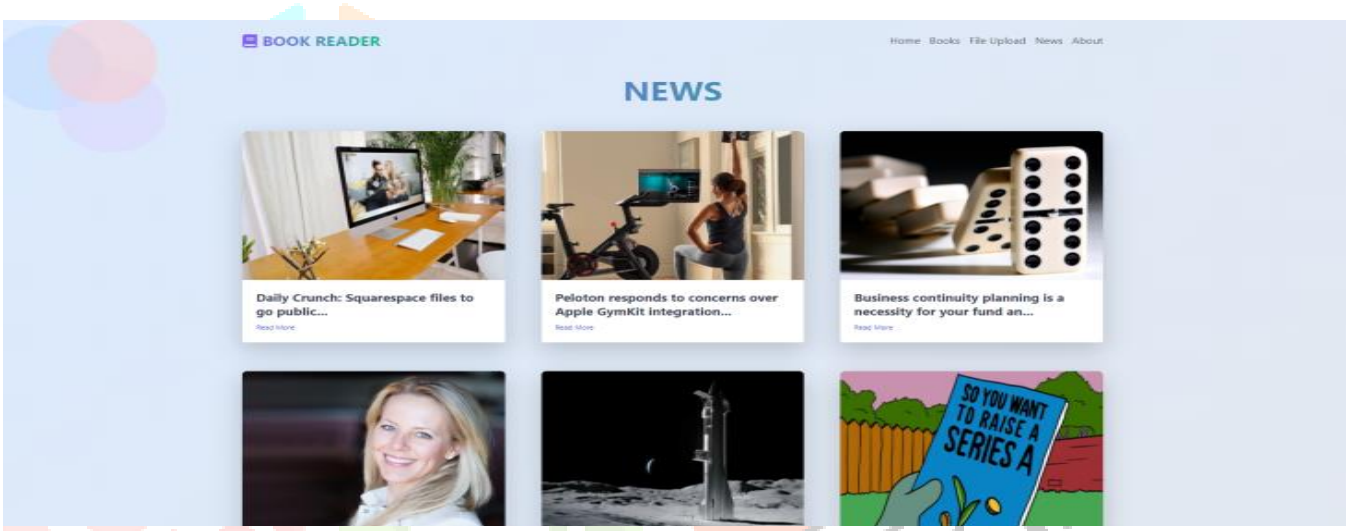


Fig 5. News Page

Fig 5. About Page

V. SCOPE AND FUTURE WORK

Currently we are supporting only English language, but the system can be scaled, and support can be provided for multiple languages, so that it can be used by various medium of instruction used by educational institutes. Therefore, students other than English medium can also use the system for their convenience; hence targeting large number of users. This system maintains a NoSQL database to store PDF's and audio to identify similar format of data, the system can be further extended to support runtime. The application interface is responsive and user friendly to ease the learning in best possible form. The system would be constantly upgraded for new features and regularly tested for errors and bugs, thus providing more accuracy and less error prone environment.

NLP has the capabilities to be included providing the ability of a system to understand, examine, manage, and generate human language. The TTS technology further can be developed for more accuracy.

VI. CONCLUSION

This system would help the visually impaired to learn and understand whatever they want at anywhere and at any point of time. They do not need any mentor or a guide always to help them study. Thus this system is developed with a motivation to increase interest in studies of visually impaired students by making knowledge extraction process simple.

VII. REFERENCES

[1] Sachin Kulkarni and Dr. Debajyoti Mukhopadhyay, "TEXT-TO-SPEECH CONVERSION FOR MATHEMATICAL DATA", April 2014

[2] KévinVythelingum, Yannick Estève, Olivier Rosee, "Error detection of grapheme-to-phoneme conversion in text-to-speech synthesis using speech signal and lexical context", 2017 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), December 2017

[3] ManjareChandraprabha Anil, S. D. Shirbahadurkar, Shaikh Shadab Shakil, "A mapper and combiner-based Marathi text to speech synthesis using English TTS Engine", 2015 Annual IEEE India Conference (INDICON), December 2015

[4] ParthaMukherjee, SoumenSantra, SubhajitBhowmick, Ananya Paul, PubaliChatterjee ArpanDeyasi, "Development of GUI for Text-to-Speech Recognition using Natural Language Processing", 2018 2nd International Conference on Electronics, Materials Engineering &Nanotechnology (IEMENTech), May 2018