International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 08 | Aug - 2021 ISSN: 2582-3930

Convert PDF to Audiobook using OCR (Optical Character Recognition) and Machine Learning

[1] Mohammad Adil Sheikh, [2] Prof. Deepa Bhattacharya, [3] Pritam Ghosh, [4] Khushal Sontakke, [5] Sharda Dharawat

Department of Computer Science and Engineering

[1,2,3,4,5] Ballarpur Institute of Technology Ballarpur, Gondwana University

Abstract:*Audiobooks are Favorable for everyone who is always zealous. It is just not possible to buy and store them in a bookshelf inside your house. Audiobooks can also be a great way to ease your eyes. A rest from the constant charge of digital screens. Yet others can use them to save time. For instance, you can keep up with books and multitask. It can not only solve the problems for millennial but can also be very beneficial tool for visually impaired person. The power to convert any documented file to audiobook is nothing but a pure gift to society. Our technology can be put on work to create such tools. LORO is one of the few contributions we make to our generation.*

Keywords:- Audiobook, OCR, PDF, Google Cloud Services.

I. INTRODUCTION

Text-to-speech and related read audio tools are being widely implemented in an attempt to assist students' reading comprehension skills.PDF to the audio system is a screen reader application designed and constructed for an effective audio communication system. The PDF to Audiobook Converter project provides an alternative way to access the PDF books for the blind, lazy, readers, and others. Using this PDF to Audiobook Converter the user will be able to listen to his\her favourite PDF and can do their daily routine.

LORO can convert any documented PDF to mp3 audio. It uses complex Deep learning algorithms to analyze the pattern of Title to body in a page and presents only the Important content. It has the ability to remove any miscellaneous content from the page, it can remove the iterative author name or Indexes of a file. It has the ability to select where the actual content is starting rather than reading every single bit of the PDF. It is annoying when the audiobook read out every punctuation in between of a very import part of the document, rather more disturbing is to listen to every bit of website mentioned. **LORO** filters out all those stuff for you so you can listen to what you actually intended for. **LORO** uses the very best API like Vision API, Auto ML API, TextToSpeech API from Google Cloud services. So, it delivers product which is best in class. **LORO** can not only converted the written documents into audio files but can also recognize text from a human handwriting, further extending its usability.

The main aim of this project is that Students, teachers, researchers and authors don't have enough time to read a book on an electronics device as that might problem with their eyes and facing a medical problem like headache, itching in the eye. So, to overcome those problem we have to create a simple audiobook that extracts the text from the pdf and read it loud to the user.

II. LITERATURE SURVEY

Overview of the Literature:-

There were six major concepts uncovered in the literature review:-

1) First major concepts reviews audiobook and its evolution and its utilization in education.

2) Second major concepts reviews the consideration of instructive strategies, assisted reading and how audiobooks have been applied to this long standing practice.

3) Third major concepts investigates the importance of fluency and prosody in reading, and what the relationship is between audiobooks and these two elements.

4) Fourth major concepts look at the research that involves reading comprehension and audiobooks.

5) Fifth major concepts is about the research covering the impact audiobooks have on reading enjoyment.

6) Finally, there is a reviews on audiobooks and distinguish instruction, and how audiobooks can scaffold content for students at varying skill levels.

Topic Specific Literature: -

1. Audiobooks:- Audiobooks should factor into any comprehensive reading program. Educators who do not consider the benefits of this technology are ignoring an effective strategy for struggling and proficient readers. Researchers have found that using audiobook technology in the classroom has long been a practicable instructional intervention for struggling readers . In regards to availability, audiobook popularity has grown in the last decade, and audiobook publishing has become a billion-dollar industry. Written text is derived from oral storytelling, so it follows that audiobooks are capturing the enthusiasm of an old tradition. With this recognition of the technology, educators find that there are now many more titles one could select.

2. Assisted reading with Audiobooks :- Audio based reading programs have a long history, starting first with an instructional strategy known as assisted reading. Assisted reading includes a fluent model in the form of a teacher and an audio recording. The process of assisted reading first starts with a student listening to an audio recording of a text selection while also reading a text based version. Following this, the teacher reads the selection, and the student follows along with a text version. Finally, the student reads the selection individually without the use of an audio version. This process found positive gains in vocabulary and comprehension.

3. Fluency and Prosody:- Literacy studies have long found that the single most important activity for developing reading skills is reading aloud. Audiobooks are especially adept at emulating the read aloud function and provides a fluent model of a competent reader. .Prosodic readers' segment text into meaningful units marked by appropriate prosodic cues such as pauses and varied duration of pauses. Students who exhibit prosodic reading display higher reading skill and comprehension.

4. Reading Enjoyment and Audiobooks:- An essential corollary to reading skill is reading enjoyment. If a student enjoys reading, they are generally found to have a high reading skill. It is plausible that the most insightful metric of reading skill is whether a person can describe the feeling of an internal movie playing in their mind. The greatest attribute that an audiobook, perhaps, allows is the ability for a struggling reader to complete a book. Successful completion of multiple books then fosters a reading habit where reading becomes a normal, consequential activity.

5. Reading Comprehension and Audiobooks:- The ultimate goal of reading a text is to gain meaning from that text. Reading comprehension skill and interest in the subject matter are then essential for learning. Students with a high level of skill in text comprehension are more apt to become expert learners. Reading skill develops through oral language experiences. So, audiobooks can benefit struggling readers by increasing comprehension of the text. The process of reading is a complex one.

6. Audiobooks and Scaffolding:- The use of audiobooks with struggling, reluctant, or second-language learners is powerful since they act as a scaffold that allows students to read above their actual reading level. This is critical for older students who may still read at a beginner level. While these students must have time to practice reading at their level, they must also have the opportunity to experience the plot structure, themes, and vocabulary of more difficult books.

III. RELATED WORK

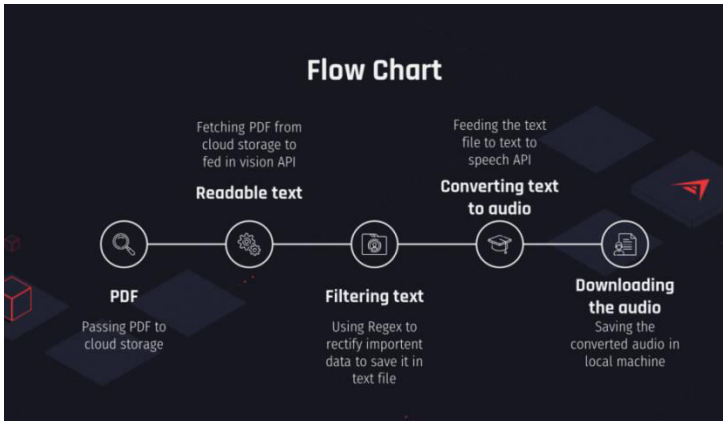
1) T. Rubesh Kumar etal [1] proposed reading is obviously essential in today's society. Printed text is everywhere in the form of reports, receipts, bank statements. There are already a few systems that have some promise for portable use, but they cannot handle product labeling. But a big limitation is that it is very hard for blind users to find the position of the bar code and to correctly point the bar code reader at the bar code [1]. T.Rubesh Kumar, C.Purnima have proposed a camera-based assistive text reading framework to help blind persons read text labels and product packaging from hand-held objects in their daily lives.

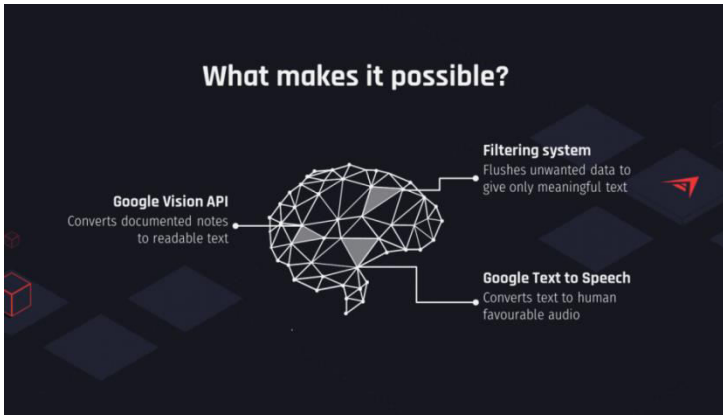
2) Pooja Sharmaetal [2] proposed Blindness is a state of lacking the visual perception due to physiological or neurological factors.In this proposed work by Pooja Sharma, Mrs.Shimi S. L. and Dr. S. Chatterji, a simple, cheap, friendly user, virtual eye will be designed and implementedto improve the mobility of both blind and visually impaired people in a specific area [2]. The proposed work includes a wearable equipment consists of head hat, mini hand stick and foot shoes to help the blind person to navigate alone safely and to avoid any obstacles that may be encountered, whether fixed or mobile, to prevent any possible accident.

3. Anusha Bhargavaetal [3] proposedMajority of the visually impaired use Braille for reading documents and books which

are difficult to make and less readily available. This gives rise to theneed for the development of devices that could bring relief to the agonizing tasks that the visually impaired has to go through says Anusha Bhargava, Karthik V. Nath, Pritish Sachdeva and Monil Samel. This project aims to study the image recognition technology with speech synthesis and to develop a cost effective, user friendly image to speech conversion system with help of Raspberry Pi [3]. The project has a small inbuilt camera that scans the text printed on a paper, converts it to audio format using a synthesized voice for reading out the scanned text quickly translating books, documents and other materials for daily living, especially away from home or office.

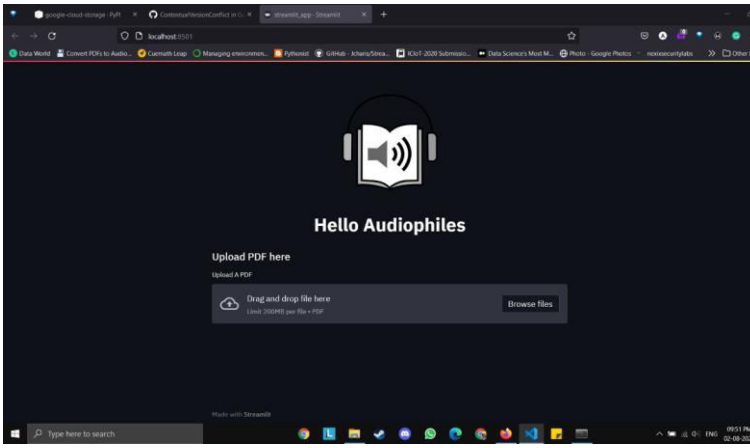
IV. SYSTEM ARCHITECTURE



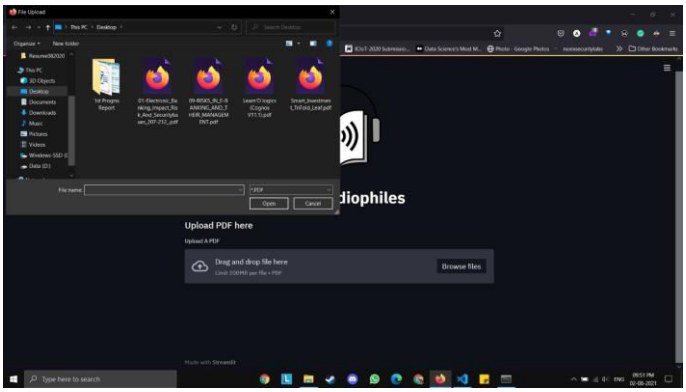


V. RESULTS

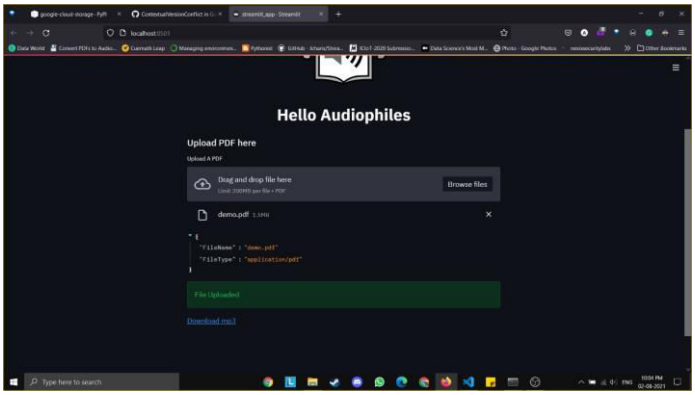
1. Starting Page:-



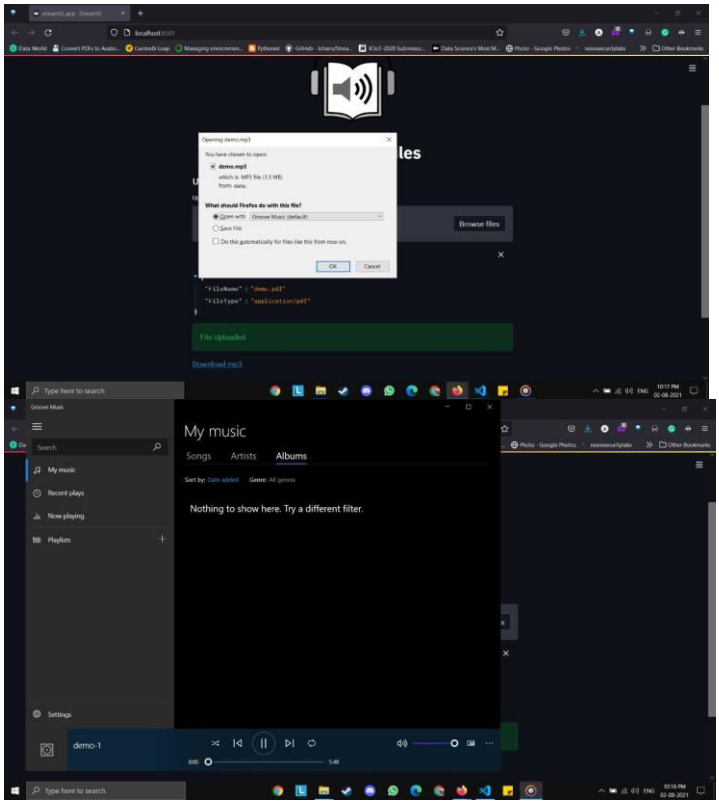
2. File Upload:-



3. File Uploaded:-



4. Audio File:-



VI. CONCLUSION

The application consists the OCR methodology that assists recognition of text (PDF file) into mp3 audio. OCR is very remarkable technology that hold a lot potential. We aim to develop our model further more to extend its processing for any documented file. We aim to give even picture its voice so it's easy for the children and visually impaired people to under the context easily. Our model will be more user friendly for the mass majority of users. It will be having good knowledge to make voice comfier for the users.

VII. FUTURE SCOPE

The enhancements of our project include providing the listener more overview information and applying the audio text concepts to electronic books, which display a single page at a time. Given the narrow view of the audio text the listener has, i.e., only the currently read words, it is easy for the listener to forget the context, i.e., what the current section is.

There is a possibility to provide the listener with context information more frequently.The possibility is for sections to be read in different voices. For example, a low pitch male voice reads introductions of papers, while a high pitch female voice reads the results. The difference in voice would be sufficient for the listener to immediately identify the current section, even without paying attention to the actual words. The reader's comprehension might improve because the text's details are more memorable this way and easily associable with the section.

VIII. REFERENCES

[1] T.Rubesh Kumar, C.Purnima "Assistive System for Product Label Detection with Voice Output For Blind Users" International Journal of Research in Engineering & Advanced Technology 2014.

[2] Pooja Sharma, Mrs. Shimi S. L and Dr. S.Chatterji, "Design of microcontroller based virtual eye for the blind" International Journal of Scientific Research Engineering & Technology,2014.

[3] Anusha Bhargava, Karthik V. Nath , Pritish Sachdeva and Monil Samel "Reading Assistant for the Visually Impaired" International Journal of Current Engineering and Technology,2015.

[4]https://www.lifewire.com/what-are-audiobooks-243853
[5]https://searchcloudcomputing.techtarget.com/definition/Google-Cloud-Platform [6]https://daleonai.com/pdf-to-audiobook
[4]https://techterms.com/definition/machine_learning
[7]https://www.investopedia.com/terms/d/deeplearning.asp#:~:text=DeeP%20learning%20 iS%20AN%20artificial,for%20use%20in%20decision%20making.&text=AlSO%20knowN%20aS%20DeeP%20neural%20learning%20or%20deeP%20neural%20network.
[8]https://docparser.com/blog/what-is-ocr/
[7]https://deepai.org/machine-learning-glossary-andterms/convolutional-neural-network
[9]https://www.researchgate.net/publication/285164623_An_Introduction_to_Convolutional_Neural_Networks
[9]https://machinelearningmastery.com/convolutional-layersfor-deep-learning-neural-networks/

[10]https://arxiv.org/ftp/arxiv/papers/1506/1506.01195.pdf
[11]https://www.tutorialspoint.com/applicationprogramming-interface-api

[12]https://heartbeat.fritz.ai/a-2019-guide-to-speech-synthesiswith-deep-learning-630afcafb9dd
[13]https://core.ac.uk/download/pdf/83592918.pdf