iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** XI **Month of publication:** November 2022

DOI: https://doi.org/10.22214/ijraset.2022.47547

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

Earbook

Prof. Dr S. M. Patil[1], Gayatri Gandhi[2], Arpita Moharir[3], Vaishnavi Pingle[4], Sai Pujari[5]
[1]*Professor,* [2,3,4,5] *Students, Department of Computer Engineering, SKN Sinhgad Institute of Technology and Science, Kusgaon (BK), Pune*

Abstract: ***it's not an era where you can take a book anywhere, and it's getting harder to carry and read a book every day. Here I came up with this solution. This yearbook application helps you convert hard copies of books to PDF format. In this application, the user simply clicks on the hard copy image with her mobile camera and it is converted to PDF format. Optical character recognition and text-to-speech are used to convert hardcopy pages to PDF format. The application allows you to update the book with a click of the hard copy and also saves this audio file in the database. This app helps you read books on your mobile phone just by installing this app and convert your notes and books to audio files (PDF) so that you can read your favorite books and notes with just a few clicks. This application also helps other visually impaired adults by allowing them to listen to audio files anytime, anywhere.***

Keywords: OCR, text-to-speech, audiobooks, machine learning, PDF.

I. INTRODUCTION

Audiobooks are becoming an admired and challenging resource for creating synthetic expression languages. Audiobooks are audio recordings of the text of a book that you listen to instead of reading. Audiobooks are exact word-to-word versions of books. These audiobooks can be accessed from smartphones, tablets and computers. The target users of the system are students who are visually impaired, who, among other adults, are more likely to encounter a visual impairment, for whom a visual impairment can be a life-changing experience, and for whom travel and learning make a difference. There are various audiobooks around the world, such as Audible, Storytel, Kuku FM. PDF to Audio System is a screen interpretation application designed for effective voice communication.

Books are an integral part of any student, book lover or knowledge seeker. Thanks to technology, soft his copies of books have been developed and made available anytime and anywhere. However, for students who commute to their respective universities every day and do not have enough time to study, in such cases it is difficult to read while traveling. Audiobooks can help in this scenario. In this project, we are developing an audiobook where the app receives input from the user. This input contains the soft copy of the book. A drawback of existing audiobooks is that you can only listen to the books provided by the app, not the user's choice. In this system, users log in and select files to convert to audio. Using ML models, text-to-speech conversion is done on the backend and displayed to the user. In today's busy life, people don't have time to pick up and read books. Instead, everyone needs alternate access to read the content. Reading a story, essay, or any other text can be tedious, but with audiobooks, reading the text makes the task easier. It is convenient because it does not Audiobooks also offer play and pause options based on usability.

PDF-to-Speech: This module allows users to take pictures of each paragraph of a book and read the user's words in the captured image.

II. LITERATURE SURVEY

In the digital age, almost everything is automated and information is stored and communicated in digital form**.** However, there are some cases where the data is not digitized and it is important to extract the text from the data for storage in digitized form. Modern technologies like text recognition software have completely revolutionized the process of text extraction using optical character recognition**.** Therefore, this paper introduces the concept, describes the extraction process and presents the latest techniques, techniques and current research in this field. Such reviews help other researchers in the field to get an overview of the technology. [1]

Handwritten character recognition is an ongoing research field that features machine learning, computer vision and pattern recognition. To do this, one scans a handwritten document and converts it into a simple text document. The basic Optical Character Recognition (OCR) process is to examine the text of a document and convert it into codes used for data processing. In this machine learning project, deep learning techniques were used to model a neural network that recognizes individual handwritten characters and handwritten numerals. To recognize them, a convolutional neural network (CNN) was built to train on alphabets and the digits datasets and further the predictions done by the trained model were visualized using OpenCV. [2]

©IJRASET: All Rights are Reserved

It discusses the historical and theoretical foundations of modern high-performance text-to-speech (TTS) systems and their current designs. The main elements of the TTS system are described with particular reference to the vocal tract model. The phases involved in the process of converting text to speech parameters are examined and include text normalization, word pronunciation, prosody, phonetic rules, language tables, and hardware implementation. The examples are mostly from Berkeley Speech Technologies' own text-to-speech system T-T-S, but other approaches are also briefly described. [3]

III. METHODOLOGY

In this system, Pdf file will be taken as input from user and processed using OCR model. This will include image preprocessing, the image recognition, handwriting recognition, feature extraction. And then converting the resulted text in OCR model to the audio format. This audio file will be made available to user at UI side.

A. Optical Character Recognition (OCR)

Optical Character Recognition (OCR) is the process of recognizing and reading text in images through computer vision. Optical Character Recognition is what the term "OCR" means. A commonly used method for recognizing text embedded in images such as scanned documents and photographs is Optical Character Recognition (OCR). If the image contains text, it uses Optical Character Recognition (OCR) technology to extract the text from the image. Perhaps best known is the use of optical character recognition to convert printed paper documents into machine-readable text files. Word processors such as Microsoft Word and Google Docs can be used to modify text in scanned paper documents that have been processed with optical character recognition.

A typical OCR system basically consists of image capture, preprocessing, segmentation, feature extraction and selection, and recognition components. Character images can be captured from a scanner or other electronic source. Any image format can be saved, and images can be color or grayscale, but the actual processing is done on binary images. The preprocessing module converts color/grayscale images to binary images. It also detects and corrects noise. Paper and print quality can introduce noise, and human intervention in image capture can introduce distortion.

1) *Image Preprocessing:* Data preprocessing, or data cleaning, is an important step for any machine learning engineer, and most ML engineers spend a lot of time preprocessing data before building models. Examples of data preprocessing include detecting outliers, handling missing values, and removing unnecessary or noisy data. Image preprocessing formats images before using them for model training and inference. This includes but is not limited to resizing, alignment, and color correction.

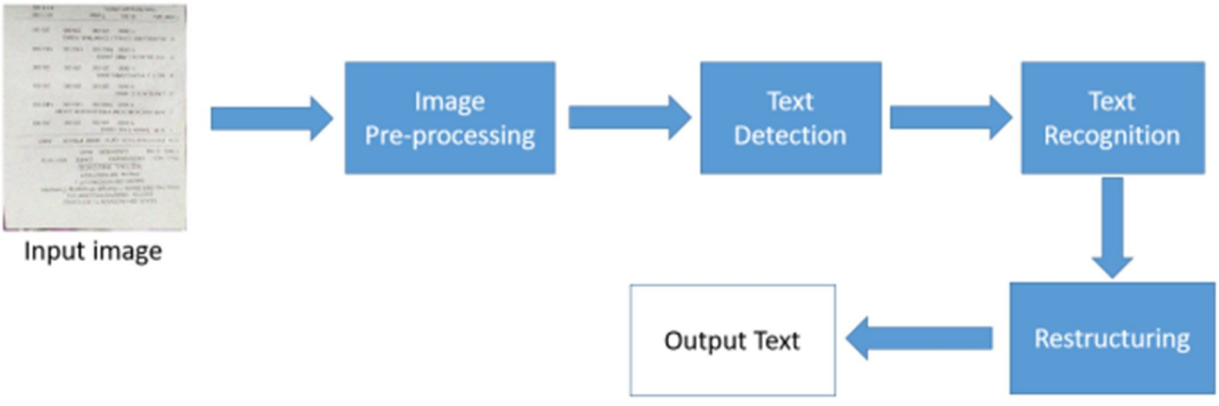


Fig. 1 Block diagram of a typical OCR System

2) *Text Detection- Text Recognition:* Text recognition is the technique of recognizing text in images and surrounding it with a rectangular bounding box. Text can be recognized using image-based or frequency-based algorithms. Split the image into multiple segments using an image-based approach. Each segment consists of connected pixels of the same quality. Classify and shape text using statistical properties of relevant components. Classify components as text or no text using machine learning techniques such as support vector machines and convolutional neural networks. The text recognition stage converts an image of text into a series of characters or phrases. Converting text images into words is important because words are the basic units that humans use for visual recognition.

3) *Feature Extraction:* The feature extraction phase is used to extract the most relevant information from the text image that helps in recognizing the characters in the text. The selection of a stable and representative feature set is central to the design of pattern recognition systems. In pattern recognition and image processing, feature extraction is a specialized form of dimensionality reduction. The main purpose of feature extraction is to obtain the most relevant information from the original data and represent that information in a low-dimensional space.

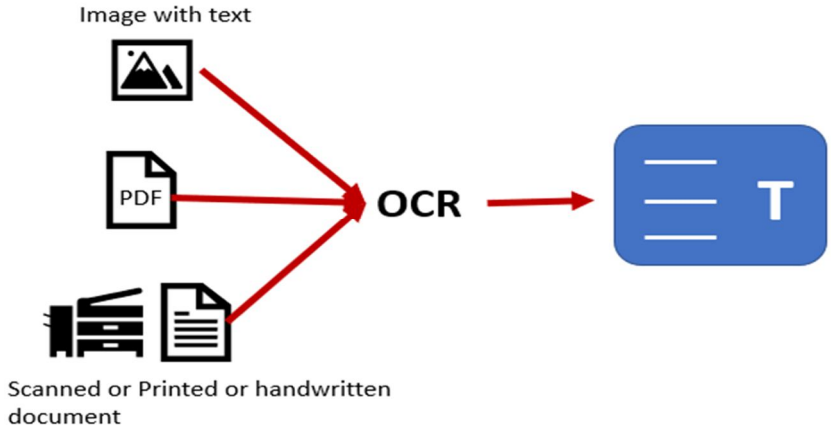


Fig. 2 OCR

B. PDF to Audio

PDF to Audio System is an interpretation application system designed for effective voice communication. PDF is the best method for electronic communication. It is also very easy to share and exchange through electronic information conversation systems. PDF documents are intended to contain links, buttons, forms, and sounds. PDF to Audiobook Converter helps users listen to audio his files while traveling or on a very busy schedule. This is also useful for visually impaired people who cannot read normally. This conversion also helps students listen to handwritten notes during exams. In this system

IV. RESULT AND CONCLUSION

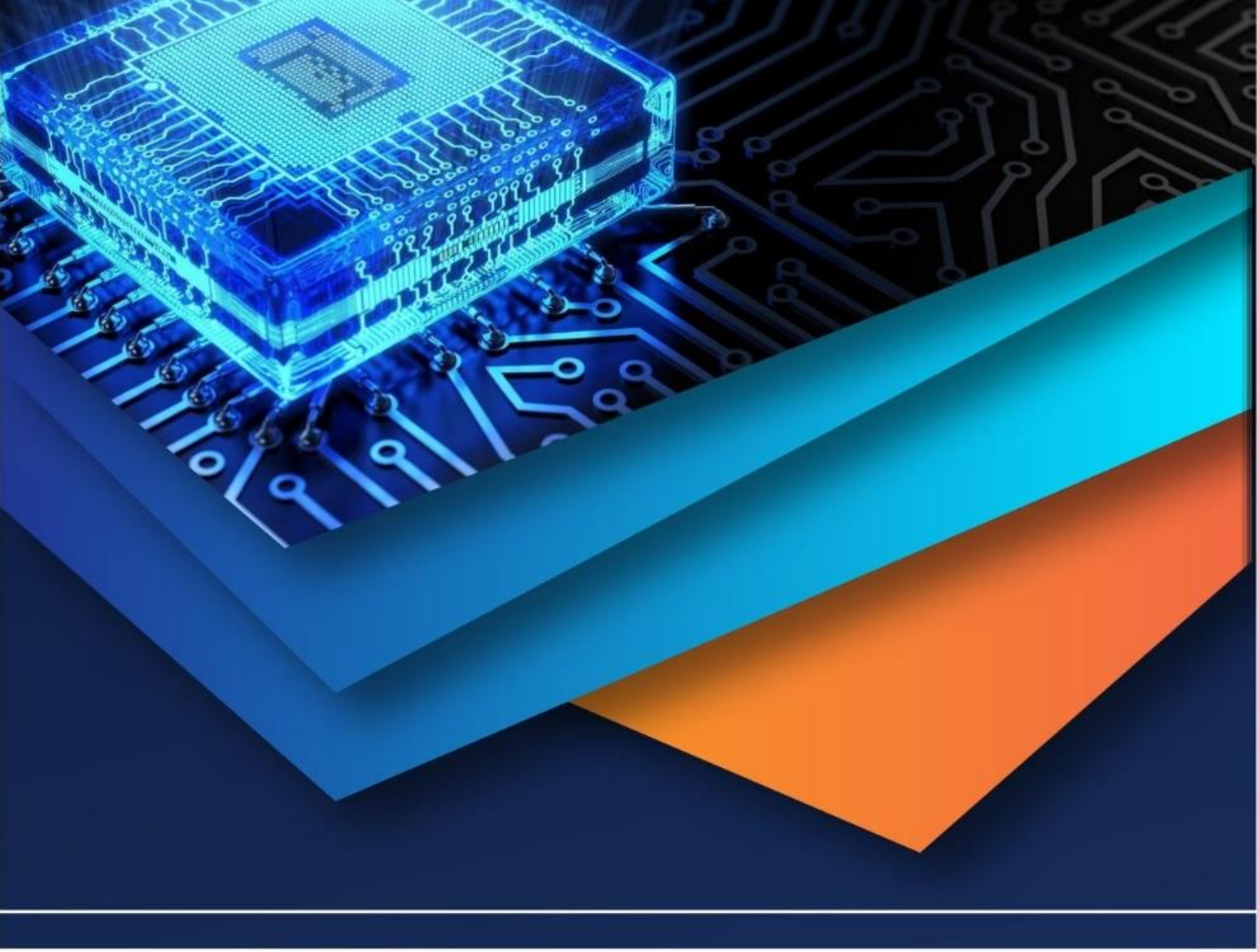
This paper described the methodology of the yearbook application. The application contains a user login page, a user registration page, an application dashboard with a files page, and an audio files page with a text file converted to an audio file. The project benefits traveling students, the blind, and book lovers. Here the hard copy of the book is replaced with a soft copy of the book with audio so you can read the book aloud. This application takes input files from selected users, converts the text to speech files using ML models, and displays these files on the user's page. This application helps students listen to handwritten notes, college books, or any book of their choice. This will help you learn more efficiently. This set of concepts is used to create a useful system for book readers and students.

V. ACKNOWLEDGEMENT

The authors would like to acknowledge the support and guidance provided by management and guides of SKN Sinhgad Institute of Technology and Science, Lonavala for providing the necessary support and guidance in carrying out this work.

REFERENCES

[1] IEEE Xplore Author-Rishabh Mittal, Achal Garg “Text Extraction with OCR: A Systematic Review” | IEEE Conference Publication, 1 September2020.
[2] Author – D. Saraswathi, Sanaa Mohamed Sherif “View of Handwritten Text Recognition System using Machine Learning (kristujayantijournal.com)” KJCS ,28 Jan 2021
[3] IEEE Xplore Author – M.H.O Malley “Text-to-Speech Technology” | IEEE Journals & Magazines | August 1990.
[4] Authors of IEEE Xplore – Rashmi S., Hanumanthappa M. “SupportVectorsMachine-Assisted Text-to-Speech Translation, an Approach to Finding Potential Paths for Human-Computer Speech Synthesizers” | IEEE Conference Publications |vol. 9 15 Jan 2016 .
[5] IEEE Xplore Author – Rahul kumar jaiswal, Rajesh kumar Dubey “Concatenative Text-to-Speech Synthesis System for Communication Recognition” | IEEE Conference Publications | 20 Jan 2022.
[6] IEEE Xplore Author – W. Ainsworth “Converting English Text to Languages” IEEE Journals & Magazines | IEEE Xplore Author – W. Ainsworth (June 1973).
[7] IEEE Xplore Author –F.Lee “Reading Engine: From Text to Speech” | IEEE Journals & Magazines |1969 December.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)