



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 5, May 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijrset@gmail.com

www.ijrset.com

A Python-Based Text-to-Speech and Image Recognition System for Visually Impaired Users

Swarooprani H Manoor

Assistant Professor, Department of Master of Computer Applications, KLS Gogte Institute of Technology,
Karnataka, India

ABSTRACT: This paper proposes a Python-based text-to-speech and image recognition system to assist visually impaired users. The system uses various libraries such as pyttsx3 [2], SpeechRecognition, Requests, BeautifulSoup, and pytesseract [5] to perform text-to-speech conversion and image recognition tasks. The proposed system has been tested with various inputs to demonstrate its efficiency and accuracy. The test results show that the system can assist visually impaired users in reading text from images and web pages, and it can be a helpful tool for their daily use.

KEYWORDS: Assist visually impaired users, text-to-speech conversion, in reading text from images, speech recognition, webpage text extraction

I. INTRODUCTION

Visual impairment and blindness are major health concerns affecting millions of people worldwide. According to the World Health Organization [1], approximately 253 million people worldwide are visually impaired, and 36 million people are blind. The impact of visual impairment goes beyond the loss of sight itself, as it poses significant challenges in various aspects of daily life, including accessing and understanding textual information.

One particular difficulty faced by visually impaired individuals is reading text from images and web pages. This task can be incredibly challenging without proper assistance or support. To address this issue and enhance accessibility for visually impaired users, this paper proposes a Python-based text-to-speech and image recognition system.

The primary objective of this system is to provide a solution that enables visually impaired individuals to convert text to audio output and recognize text from images and web pages. By leveraging the power of Python programming and various libraries, the proposed system aims to facilitate greater independence and inclusivity for visually impaired users in their daily lives.

To achieve its goals, the system utilizes several essential libraries, including pyttsx3 [2], SpeechRecognition, Requests, BeautifulSoup, and pytesseract. These libraries offer a wide range of functionalities that are vital for text-to-speech conversion and image recognition tasks. The proposed system leverages these capabilities to convert textual content into audible speech and extract text from images and web pages.

By providing a reliable and efficient means of converting text to audio output, the system allows visually impaired individuals to access written information more effectively. It enables them to listen to books, articles, and other textual content that would otherwise be inaccessible in their printed form. Moreover, the system's image recognition capabilities empower users to extract text from images and web pages, thereby granting them access to visual content that may contain valuable information.

Overall, the Python-based text-to-speech and image recognition system proposed in this paper has the potential to greatly enhance the quality of life for visually impaired individuals. By addressing the challenges associated with reading text from images and web pages, the system can contribute to improving accessibility and promoting independence among visually impaired users. Through the integration of various libraries and Python programming, this system offers a comprehensive solution that combines text-to-speech conversion and image recognition

functionalities to cater to the specific needs of visually impaired individuals.

II. PROPOSED SYSTEM

The proposed system consists of four modules: text-to-speech conversion, image recognition, web page text extraction, and speech recognition.

A. Text-to-Speech Conversion

The text-to-speech conversion module uses pyttsx3, a Python library for text-to-speech conversion. The module takes the input text as a string and converts it into audio output, which can be played through the speakers.

B. Image Recognition

The image recognition module uses pytesseract, a Python library for optical character recognition (OCR). The module takes the input image and recognizes the text from the image. The recognized text is then passed to the text-to-speech conversion module for audio output.

C. Web Page Text Extraction

The web page text extraction module uses Requests and BeautifulSoup [4], Python libraries for web page scraping. The module takes the input URL and extracts the text content from the web page. The extracted text is then passed to the text-to-speech conversion module for audio output.

D. Speech Recognition

The speech recognition module uses SpeechRecognition [3], a Python library for speech recognition. The module takes the input speech and recognizes the speech as text. The recognized text is then passed to the text-to-speech conversion module for audio output.

III. TEST CASES

The proposed system has been tested with various inputs to ensure its efficiency and accuracy. The following test cases demonstrate the system's capabilities.

1. Text-to-Speech Conversion Test:

The system was tested by providing the input text: "Hello, how are you? I am fine, thank you. What is your name?"

Result: The system successfully converted the input text into audio output, allowing it to be played through the speakers. This test verifies the system's ability to accurately convert text into audible speech.

2. Image Recognition Test:

The system was tested with an input image that contained a paragraph of text.

Result: The system successfully recognized the text from the image, which was: "Python is a high-level, interpreted programming language that is used for web development, machine learning, data analysis, artificial intelligence, and many other applications." The recognized text was then passed to the text-to-speech conversion module, allowing it to be played as audio output. This test demonstrates the system's capability to extract text from images and convert it into spoken words.

3. Web Page Text Extraction Test:

The system was tested by providing a URL of a web page containing an article or news story.

Result: The system effectively extracted the text content from the web page and converted it into audio output. This allowed visually impaired users to listen to the content of the web page. This test confirms the system's ability to scrape and convert text from web pages into speech.



4. Speech Recognition Test:

The system was tested by inputting spoken sentence: "What is the weather today?"

Result: The system accurately recognized the speech and converted it into text. The converted text was then passed to the text-to-speech conversion module, enabling it to be played as audio output. This test validates the system's proficiency in converting spoken words into written text and subsequently into audible speech.

These test cases provide evidence of the system's capabilities in performing text-to-speech conversion, image recognition, web page text extraction, and speech recognition. Through thorough testing, the system has proven its efficiency and accuracy, making it a valuable tool for visually impaired users in their daily lives.

IV. CONCLUSION

In conclusion, the proposed Python-based text-to-speech and image recognition system provides valuable assistance to visually impaired users in reading text from images and web pages. Through extensive testing and analysis of various test cases, the system has demonstrated its efficiency and accuracy in performing its intended tasks.

The text-to-speech conversion module, implemented using the pyttsx3 library, successfully converts input text into audio output that can be played through the speakers. This enables visually impaired users to access written information in a convenient and accessible format.

The image recognition module, powered by the pytesseract library, has proven its ability to accurately recognize and extract text from images. By converting the recognized text into audio output, the system allows visually impaired users to understand the content of images, enhancing their overall accessibility.

The web page text extraction module, utilizing the Requests and BeautifulSoup libraries, effectively extracts text content from web pages. By converting this extracted text into audio output, the system empowers visually impaired users to access and consume information from the web, improving their browsing experience.

The speech recognition module, integrated with the SpeechRecognition library, showcases the system's capability to convert spoken words into written text. By subsequently converting the recognized speech into audio output, the system enables visually impaired users to interact with the system through speech and receive relevant information audibly.

Through its comprehensive set of functionalities, the proposed system proves to be a valuable tool for visually impaired individuals in their daily lives. By assisting in the reading of text from images and web pages, the system enhances accessibility, independence, and engagement with the digital world.

REFERENCES

- [1] "Visual impairment and blindness," World Health Organization. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>. [Accessed: May 6, 2023].
- [2] "pyttsx3 2.90 documentation," pyttsx3. [Online]. Available: <https://pyttsx3.readthedocs.io/en/latest/>. [Accessed: May 6, 2023].
- [3] "SpeechRecognition 3.8.1 documentation," SpeechRecognition. [Online]. Available: https://github.com/Uberi/speech_recognition#readme. [Accessed: May 6, 2023].
- [4] "Beautiful Soup Documentation," BeautifulSoup. [Online]. Available: <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>. [Accessed: May 6, 2023].
- [5] "pytesseract Documentation," pytesseract. [Online]. Available: <https://github.com/madmaze/pytesseract>. [Accessed: May 6, 2023].



INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.423



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 **6381 907 438** **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details