

Smart Reader for Blind People Using AI and Computer Vision

Prof. Komal Rathod¹, Abhishek Bankar², Chetan Gangurde³, Onkar Ganjure⁴, Aniket Gite⁵

IT,
ZCOER, Pune
India

ABSTRACT

This paper presents a Smart Reader system designed to assist visually impaired individuals in reading printed text using Artificial Intelligence and Computer Vision. The system captures images using a camera, enhances image quality, and extracts text using a hybrid OCR approach combining AI-based OCR and Tesseract. The extracted text is processed using language detection and summarization techniques. A hybrid Text-to-Speech module converts text into speech, allowing users to hear the content. The system also supports offline voice commands using Vosk speech recognition. The proposed system provides accurate, real-time, and user-friendly reading assistance, improving accessibility and independence for visually impaired users.

Keywords — OCR, Computer Vision, Assistive Technology, Text-to-Speech, Voice Recognition.

I. INTRODUCTION

Visual impairment makes it difficult for individuals to access printed text such as books, documents, and labels. Traditional solutions like Braille and audio recordings are limited and not always available. With advancements in Artificial Intelligence and Computer Vision, smart assistive systems can now provide real-time reading assistance.

The proposed Smart Reader system helps visually impaired users by converting printed text into speech. It uses image processing, Optical Character Recognition (OCR), and Text-to-Speech (TTS) technologies. The system is easy to use and can be operated through voice commands, making it suitable for independent usage.

II. LITERATURE REVIEW

The development of assistive reading systems for visually impaired individuals has gained significant attention in recent years. Various researchers have proposed systems using Optical Character Recognition (OCR), Text-to-Speech (TTS), embedded platforms to improve accessibility

Impaired (IEEE 9243598)		
<i>Smart Reader for Blind People</i> (IJCRT, ISSN: 2320-2882)	To design a portable and low-cost reading device.	Lacks summarization and multilingual capability.
<i>OCR Assisted Translator</i> (IEEE 9202034)	To translate captured text into different languages.	Works only under good lighting; slower processing.
Smart Reader for Blind People. (IEEE 9155941)	To convert printed text to speech using camera.	Affected by image noise and blur.

Title of Paper	Objective	Limitations
Smart Reader for Visually	To assist visually impaired users	Limited to single-language support

Krishna Rao (2021) proposed a Smart Reader system for visually impaired users that integrates OCR and Text-to-Speech technologies for real-time

text reading. The system captures printed text using a camera and converts it into speech output. Although the system provides real-time assistance, it is limited by its support for a single language, which restricts its usability in multilingual environments.

Vishal More (2024) developed a portable Smart Reader system using Raspberry Pi. The system focuses on designing a low-cost and compact reading device that captures images and converts text into speech output. While the system is efficient in terms of cost and portability, it lacks advanced features such as text summarization and multilingual support, which are essential for handling large and diverse text content.

Nikhil Chigali (2020) introduced an OCR-assisted translator system that converts captured text into multiple languages using translation APIs. The system combines OCR with language translation to improve accessibility across different languages. However, the system is dependent on good lighting conditions for accurate text extraction and suffers from slower processing speed due to API-based operations.

From the analysis of existing systems, it is observed that most solutions rely on basic OCR and speech technologies but lack advanced features such as intelligent automation, hybrid OCR models, offline voice command processing, and AI-based summarization. Additionally, many systems are dependent on internet connectivity or are sensitive to environmental conditions such as lighting and image quality.

The proposed Smart Reader system addresses these limitations by incorporating a hybrid OCR approach, offline voice recognition, automatic image capture, and AI-based summarization. These enhancements improve accuracy, usability, and accessibility, making the system more robust and suitable for real-world applications.

III. SYSTEM ARCHITECTURE

The system consists of multiple modules working together. The system begins with capturing an image using a camera. The captured image is processed to enhance text visibility. OCR is applied to extract text from the image. The extracted text is then analyzed for language detection and summarization. Finally, the text is converted into speech and delivered to the user.

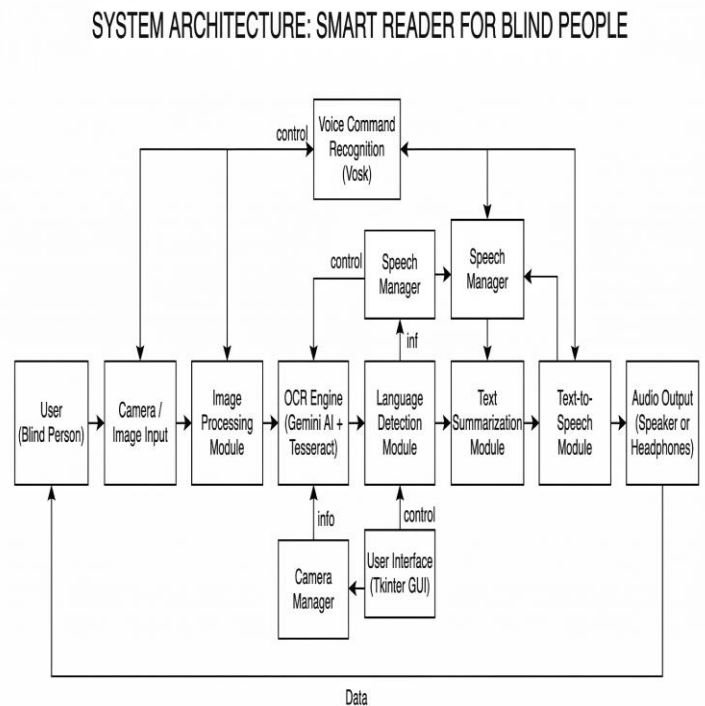


Figure 1: System Architecture

IV. METHODOLOGY

The methodology describes the step-by-step process of the system.

4.1 Image Acquisition

The system captures images using a webcam. Continuous frame capture is used to detect stability and automatically capture images when the page is steady.

4.2 Image Preprocessing

Preprocessing improves OCR accuracy by enhancing image quality.

Techniques used:

- Grayscale conversion
- Noise reduction
- Image sharpening
- Thresholding

4.3 Optical Character Recognition

The system uses a hybrid OCR approach:

- AI-based OCR (Gemini model)
- Tesseract OCR (fallback)

This ensures both accuracy and reliability.

4.4 Language Detection

The extracted text is analyzed using language detection algorithms to identify the language. This allows proper pronunciation during speech synthesis.

4.5 Text Summarization

For large text content, summarization reduces the amount of spoken output. AI models generate concise summaries.

4.6 Text-to-Speech Conversion

Text is converted into speech using:

- Google Text-to-Speech (online)
- pyttsx3 (offline)

4.7 Voice Command Processing

The system uses Vosk speech recognition to process commands such as:

- Capture
- Speak
- Summary
- Exit

V. IMPLEMENTATION

Software Tools	Hardware Requirements
<ul style="list-style-type: none"> • OpenCV • Tesseract OCR • Vosk Speech Recognition • gTTS • pyttsx3 	<ul style="list-style-type: none"> • Processor: Intel i5 or above • RAM: 8 GB • Camera: HD Webcam • Speaker: Audio output device

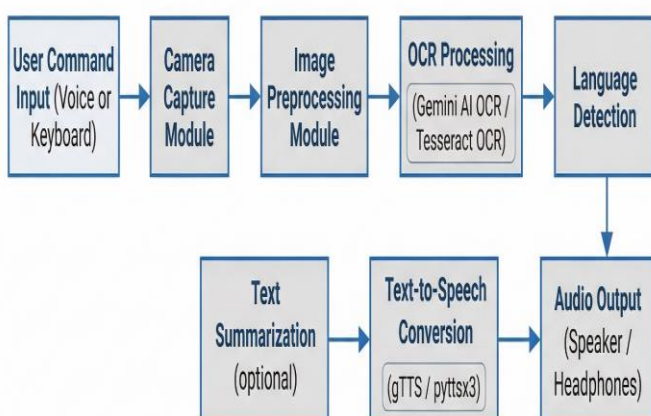
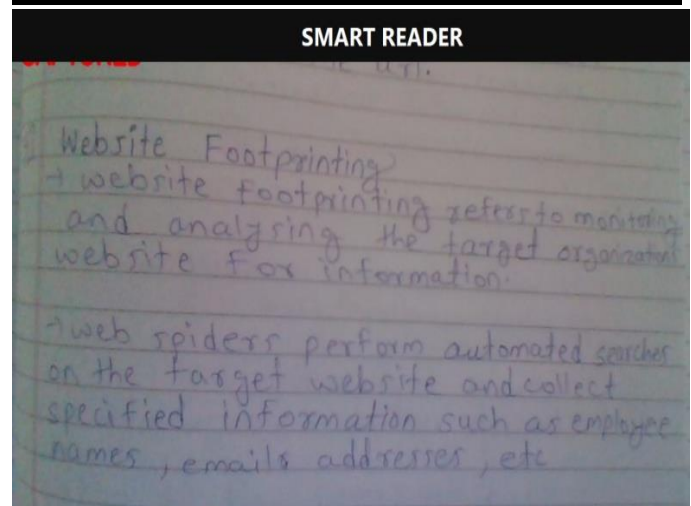
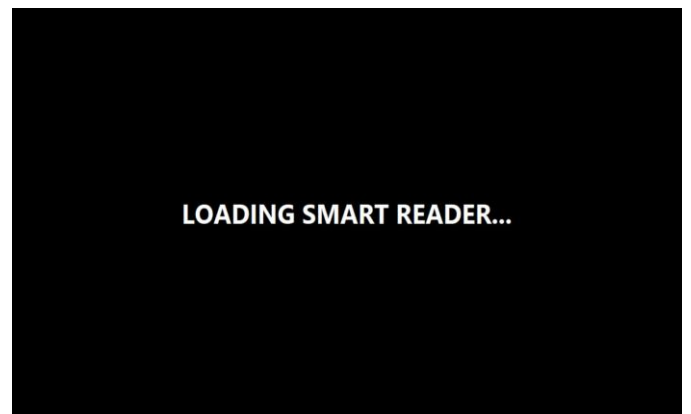


Figure 2: Workflow Diagram

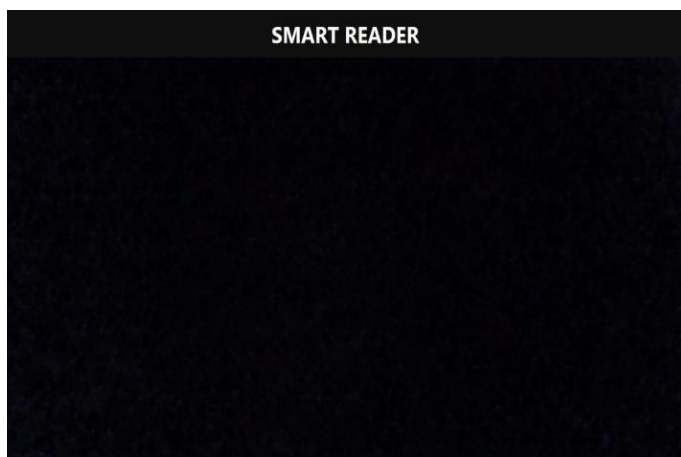


Figure 3: Application Interface

VI. RESULT AND ANALYSIS

Parameter	Value
OCR Accuracy	80-95%
Processing Time	10-20 sec
Speech Quality	High
Voice Accuracy	92%

VII. LIMITATIONS

- Low performance in poor lighting
- Handwritten text not accurate
- AI OCR needs internet

VIII. CONCLUSION

The Smart Reader system provides an efficient and intelligent solution for visually impaired individuals. The integration of AI, OCR, and speech processing enables real-time reading assistance. The system improves accessibility and independence for users.

IX. FUTURE SCOPE

- Mobile application
- Smart glasses integration
- Multi-language support
- Scene recognition

X. REFERENCES

- [1]. Krishna raj Rao, “Smart Reader for Visually Impaired,” IEEE International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1123–1127, 2021.
- [2]. Vishal More, “Smart Reader for Blind People,” International Journal of Creative Research Thoughts (IJCRT), Vol. 12, Issue 3, pp. 250–255, 2024.
- [3]. “OCR Assisted Translator,” IEEE Conference on Advances in Computing, Communication, and Control, 2020.
- [4]. Shantanu Chaudhari, “Optical Character Recognition: A Review,” International Journal of Computer Applications (IJCA), Vol. 175, No. 23, pp. 15–19, 2020.
- [5]. OpenCV Documentation – Open Source Computer Vision Library, [Online]. Available: <https://docs.opencv.org>
- [6]. Tesseract OCR Documentation – Tesseract OCR Engine, [Online]. Available: <https://tesseractocr.github.io/>
- [7]. Google Cloud Text-to-Speech API – Cloud-based Speech Synthesis Service Available: <https://cloud.google.com/text-to-speech>
- [8]. S. Patel, A. Sharma, and D. Verma, “AI-Based Assistive Technologies for the Visually Impaired,” International Journal of Advanced

- Research in Computer Science (IJARCS), Vol. 11, Issue 5, pp. 45–52, 2023.
- [9]. P. Banerjee and R. Das, “Deep Learning-Based Text Recognition Using Convolutional Neural Networks,” *IEEE Access*, Vol. 9, pp. 78365–78375, 2021.
- [10]. K. Jain and B. Ghosh, “Human-Centric Computing for Accessibility Enhancement,” *ACM Transactions on Accessible Computing (TACCESS)*, Vol. 14, No. 2, pp. 1–18, 2022.
- [11]. M. Gupta, “Real-Time Scene Text Detection Using EAST Model,” *International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)*, Vol. 10, Issue 6, pp. 8765–8772, 2022.
- [12]. R. Kulkarni, “Assistive Vision Systems Using Raspberry Pi and OCR,” *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, Vol. 11, Issue 8, pp. 1650–1656, 2023.
- [13]. A. Singh and N. Mehta, “Voice-Based Assistive Applications for Visually Challenged,” *Journal of Emerging Technologies and Innovative Research (JETIR)*, Vol. 10, Issue 4, pp. 980–985, 2023.
- [14]. S. Reddy, “Integration of OCR and Speech Synthesis for Accessibility Enhancement,” *International Journal of Computer Science Trends and Technology (IJCST)*, Vol. 9, Issue 2, pp. 55–60, 2022.