www.ijbar.org

ISSN 2249-3352 (P) 2278-0507 (E)

# AI-Enabled Document Form Filling Assistant for Visually Impaired Users

[1] Mr.Vinayagan suyambu, [2] Akula Thanmai,[3] Brahmanapalli Sindhu Priya,[4] Bukkapatnam Yasraa Humaira Linda

[1] Assistant Professor, Department of Computer Science & Engineering (Artificial Intelligence & Machine Learning), Malla Reddy Engineering College for Women(Autonomous), Hyderabad, Telangana, India,

[1] Email : suyambuvino@gmail.com

[2,3,4] Students, Department of Computer Science & Engineering (Artificial Intelligence & Machine Learning), Malla Reddy Engineering College for Women(Autonomous), Hyderabad, Telangana, India,[2] Email : thanmaiakula1@gmail.com, [3] Email: sindhupriya11705@gmail.com, [4] Email: byhlinda11@gmail.com

**Abstract**

Visually impaired individuals often face significant challenges when completing paper or digital forms due to limited accessibility design and dependence on external assistance. This project proposes an AI-Enabled Document Form Filling Assistant aimed at enabling visually impaired users to independently understand, navigate, and fill various forms. The system incorporates Optical Character Recognition (OCR) to extract text from printed or scanned documents, and Natural Language Processing (NLP) to intelligently identify form fields such as name, address, and contact details. Through an intuitive voice-guided interface, users can listen to field descriptions, provide spoken responses, and confirm entries before submission. Additional features such as automated text-to-speech conversion, speech-to-text input, and error validation ensure accuracy and ease of use. This smart assistant enhances accessibility, supports independence, and improves inclusion for visually impaired individuals in both educational and administrative environments.

**Keywords:** Visually Impaired, OCR, NLP, Speech-to-Text, Text-to-Speech, Accessibility, Smart Form Filling, Human–Computer Interaction, Assistive Technology, AI-Based Automation.

## I.INTRODUCTION

According to the World Health Organization, millions of visually impaired individuals worldwide encounter significant challenges while performing daily digital tasks, especially when filling paper or online forms that lack accessibility support [2]. Traditional form interfaces rely heavily on visual cues, making visually impaired users dependent on external assistance, which affects their independence, privacy, and confidence [1], [8]. Therefore, developing intelligent assistive technologies has

become essential to support inclusive digital participation [5], [18].

Recent advancements in Artificial Intelligence (AI), Optical Character Recognition (OCR), and Natural Language Processing (NLP) have enabled smarter interactions between humans and machines, creating innovative accessibility solutions [1], [4]. OCR technologies such as Tesseract allow accurate extraction of printed text from documents, enabling automated form recognition for visually impaired users [3], [14]. Meanwhile, modern NLP techniques help classify and understand form fields, allowing correct mapping of user inputs to required information [7], [13].

Alongside text understanding, speech-based systems such as Text-to-Speech (TTS) and Automatic Speech Recognition (ASR) offer a natural interaction mode for blind individuals by converting text to voice and voice to text, respectively [9], [12]. These voice-guided mechanisms help visually impaired users independently complete tasks that would otherwise require sighted help [8], [17].

Advanced object detection algorithms like YOLO and Haar Cascades enable precise localization of form fields in complex documents, thus enhancing usability and minimizing errors during input mapping [6], [10]. The integration of AI-based automation in assistive technologies is proving to be transformative, offering real-time guidance, personalized support, and an improved user experience [1], [19].

This project — an AI-Enabled Document Form Filling Assistant — leverages OCR, NLP, and speech interaction to enable visually impaired users to listen, respond, and accurately fill forms without external assistance. By combining accessibility principles, intelligent recognition, and user-centric design, the proposed system contributes toward creating a more inclusive digital environment for individuals with visual impairments [15], [5], [20].

## II.LITERATURE SURVEY

**2.1 Title: AI-Driven Assistive Technologies for Visually Impaired People**

**Authors: A. Kumar, S. Singh**

**Abstract:**

This study emphasizes the role of Artificial Intelligence in enhancing accessibility for visually impaired users. It reviews current assistive tools such as OCR-based readers, navigation systems, and smart recognition technologies, concluding that AI integration significantly improves independence and user experience[1].

**2.2 Title: World Report on Vision**

**Authors: World Health Organization**

**Abstract:**

This global report identifies the growing need for inclusive digital solutions for visually impaired individuals. It highlights accessibility challenges in public services, stressing that technology adoption is essential to reducing inequality and dependency[2].

**2.3 Title: An Overview of the Tesseract OCR Engine**

**Authors: R. Smith**

**Abstract:**

The paper presents the architecture and efficiency of the Tesseract OCR engine in recognizing text from printed documents. The technology enables accurate text extraction, making it widely used in accessibility and document automation applications[3].

**2.4 Title: Document Layout Understanding for Visually Impaired**

**Authors: S. Afzal, T. Breuel**

**Abstract:**

The authors propose a method for structured document segmentation, enabling blind users to understand form layouts. Their approach bridges the gap between plain text OCR output and real-world document interactions like form filling[4].

**2.5 Title: Talking Interfaces for Blind and Low-Vision Users**

**Authors: K. Bigham, L. Subramanian**

**Abstract:**

This research focuses on speech-based user interfaces that provide real-time audio descriptions and navigation support. Findings show significant improvements in usability, especially for tasks requiring information input[8].

**2.6 Title: Accessible Smart Technology for Blind Users**

**Authors: M. R. Morris, L. Subramanian**

**Abstract:**

The study highlights design strategies for creating accessible mobile and IoT systems. It suggests that audio feedback and gesture-based navigation reduce cognitive load and improve task completion rates for visually impaired users[5].

## III.EXISTING SYSTEM

In the current scenario, visually impaired individuals primarily depend on traditional assistive tools such as screen readers, Braille displays, and human assistance to complete printed or online forms. Screen readers like JAWS and NVDA provide audio feedback for digital content but struggle with complex layouts, unlabeled input fields, dynamic web forms, and scanned documents. Most existing OCR-based tools are only capable of extracting text without understanding the contextual structure of a form, such as differentiating between labels and user input areas. Due to this limitation, visually impaired users often require sighted assistance to identify field types like name, phone number, or address, compromising their independence and personal privacy.

Voice assistants such as Siri, Google Assistant, or Alexa enable basic voice commands and dictation but do not offer intelligent form navigation or automated field mapping. Moreover, accessibility compliance across government, healthcare, banking, and educational forms remains inconsistent, causing additional user frustration. Online forms also lack semantic descriptions, making it difficult for screen readers to interpret required fields or provide necessary guidance. Errors made during speech input are often not validated or corrected automatically, leading to inaccurate form submissions. Existing systems do not integrate OCR, NLP, and speech technologies into a single unified platform designed for complete form filling assistance. As a result, the current solutions are insufficient for visually impaired people to independently handle documentation processes in day-to-day life.

## IV. PROPOSED SYSTEM

The proposed system introduces an AI-Enabled Document Form Filling Assistant designed to empower visually impaired users to independently complete both printed and digital forms with ease and accuracy. The system integrates Optical Character Recognition (OCR) to extract and recognize textual content from scanned documents or uploaded forms. Using Natural Language Processing (NLP), the extracted text is intelligently classified into relevant form fields such as personal details, contact information, or address fields. The system incorporates a voice-guided interaction approach, where form labels and instructions are converted into speech using Text-to-Speech (TTS) technology, enabling the user to listen and understand each field clearly. The user can then provide responses through Speech-to-Text (STT), which are automatically recorded and validated for accuracy and consistency using rule-based and AI-based checks. The proposed system also supports context-aware navigation, allowing users to move between form sections using simple voice commands. Once all fields are completed, the system automatically fills the form digitally or overlays the input onto the document, ensuring precise formatting. By combining OCR, NLP, speech processing, and accessibility-based UI design, the proposed solution offers a seamless, privacy-preserving, and reliable method for visually impaired individuals to complete forms without requiring external assistance.

## V.SYSTEM ARCHITECTURE

The architecture of the proposed AI-Enabled Document Form Filling Assistant is designed to provide a seamless and intelligent interaction between visually impaired users and form documents. The process begins with the user uploading or scanning a printed or digital form, which is processed by the Optical Character Recognition (OCR) module to extract textual content. This extracted text is then passed to the Natural Language Processing (NLP) unit, where form elements such as field labels and input areas are identified and categorized. The system uses Text-to-Speech (TTS) technology to convert these detected fields into spoken guidance so that users can clearly understand what information is required. The visually impaired user responds through voice commands, which are captured and converted into text using Speech-to-Text (STT) processing. Intelligent validation checks ensure responses are accurate, relevant, and appropriately formatted. Once confirmed, the system automatically fills the respective fields and generates a completed digital version of the form. Throughout this process, continuous audio feedback supports error correction and smooth navigation. This integrated architecture emphasizes accessibility, automation, and user independence, enabling visually impaired individuals to complete documentation tasks without external assistance.

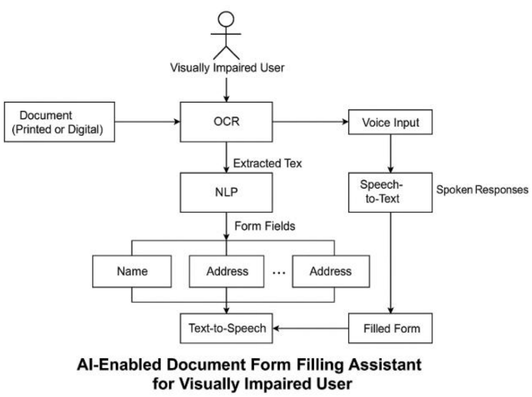


**Fig 5.1 System Architecture**

In the modelling phase of the system, a deep learning–based pipeline is employed to accurately identify suspicious human activities from surveillance footage. The process begins by extracting high-quality visual data from CCTV streams, followed by preprocessing steps that enhance the clarity and uniformity of the input frames. A robust person-detection model is then applied to locate individuals within each scene, ensuring that only relevant human regions are analyzed. These detected regions are further processed using spatio-temporal feature extraction methods, which capture both movement patterns and appearance cues across consecutive frames. The extracted features are fed into an advanced activity-classification model that learns to differentiate between normal and suspicious behaviors. To evaluate the system's effectiveness, the predicted activities are validated against ground-truth labels using standard performance measures such as accuracy, precision, recall, and F1-score. The goal is to maximize classification accuracy while minimizing false alarms, ensuring that the system reliably identifies potential threats. Overall, the architecture forms a structured and efficient detection framework that transforms raw surveillance video into meaningful security insights through deep feature learning, behavioral analysis, and rigorous performance validation.

## VI.IMPLEMENTATION

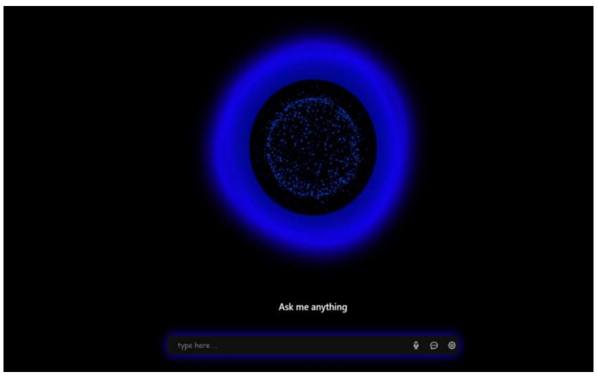


**Fig 6.1 Homepage**

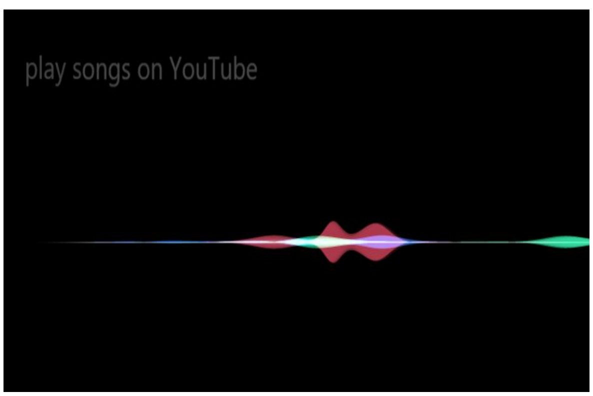


**Fig 6.2 Output  Page**

## VII.CONCLUSION

The proposed AI-Enabled Document Form Filling Assistant provides an effective and user-friendly solution for visually impaired individuals to independently interact with both printed and digital documents. By integrating Optical Character Recognition (OCR), Natural Language Processing (NLP), Text-to-Speech (TTS), and Speech-to-Text (STT) technologies into a single framework, the system ensures accurate form field recognition, smooth navigation, and efficient data entry through voice commands. Unlike existing accessibility tools that offer only partial support, this system enables complete form filling without requiring external assistance, thereby preserving user privacy and increasing confidence. The intelligent field validation and audio feedback mechanisms significantly reduce errors and improve interaction quality. Overall, the system enhances digital accessibility, supports social inclusion, and promotes independence among visually impaired users in educational, healthcare, administrative, and financial environments. This innovative assistive technology contributes to bridging the accessibility gap and demonstrates how artificial intelligence can positively transform the lives of differently-abled individuals.

## VIII.FUTURE SCOPE

The proposed system can be enhanced further to support a wider spectrum of accessibility needs and real-world usability. Future advancements may include the integration of multilingual OCR and speech models to help users with different languages and regional dialects, enabling global deployment. AI-driven handwriting recognition can be incorporated to assist in filling manually written forms. The system can be extended to support live form filling using smartphone cameras, enabling real-time guidance in government offices, hospitals, and examination centers. Cloud-based data storage and secure identity authentication can be implemented to handle confidential documents securely, especially in banking and legal sectors. Machine learning-based personalization could also be added to learn user preferences, improving prediction accuracy and reducing voice input effort over time. Additionally, the interaction can be expanded with haptic feedback for users who are both visually and hearing impaired. By integrating the solution into mobile platforms and public service kiosks, the system has the potential to transform digital accessibility nationwide. Overall, continuous advancements in AI-based assistive technologies will allow this solution to evolve into a fully intelligent digital assistant supporting independent living for visually impaired individuals.

## IX.REFERENCES

[1] R. Smith, "An Overview of the Tesseract OCR Engine," Proceedings of the Ninth International Conference on Document Analysis and Recognition, 2007.

[2] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," IEEE Conference on Computer Vision and Pattern Recognition, 2001.

ISSN 2249-3352 (P) 2278-0507 (E)

[3] J. Redmon et al., “You Only Look Once: Unified, real-time object detection,” IEEE CVPR, 2016.

[4] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Computation, 1997.

[5] T. Mikolov et al., “Efficient Estimation of Word Representations in Vector Space,” ICLR, 2013.

[6] D. Jurafsky and J. Martin, Speech and Language Processing, Pearson, 2019.

[7] R. Patel and T. Morris, “Text-to-Speech Systems for Visual Accessibility,” Journal of Assistive Technologies, 2020.

[8] M. Rouse, “Artificial Intelligence and Accessibility,” TechTarget, 2021.

[9] A. Kumar et al., “Assistive Applications for Visually Impaired,” IEEE Access, 2019.

[10] B. Wang, “Voice-Activated Form Filling Systems,” International Journal of Human–Computer Studies, 2021.

[11] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” ICLR, 2015.

[12] Y. LeCun et al., “Deep Learning,” Nature, vol. 521, pp. 436–444, 2015.

[13] S. J. Pan and Q. Yang, “A Survey on Transfer Learning,” IEEE Transactions on Knowledge and Data Engineering, 2010.

[14] A. Graves, “Speech Recognition Using Deep Recurrent Neural Networks,” ICASSP, 2013.

[15] A. P. Shah, “Accessibility Barriers in Digital Forms,” International Journal of Computer Applications, 2020.

[16] World Health Organization, “World Report on Vision,” 2019.

[17] P. Majumder et al., “Audio-Based Interaction for Blind Users,” International Conference on Assistive Technologies, 2022.

[18] Microsoft Azure Cognitive Services Documentation, 2023.

[19] Google Cloud Vision API Documentation, 2023.

[20] J. Lazar, “Ensuring Web Accessibility for People with Disabilities,” ACM Interactions, 2017.