International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2423 - 2428
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0359

Artificial Intelligence-Based Vision Assistance System for Blind Students Using Smart Glasses

P. Naga Gayathri[1], N. Giri[2], S.Pavithra[3], S. Riffat Fathima[4], G.Babu[5], K.Saranya[6]
[1,2,3,4] UG Student, Department of Computer Science and Engineering, Knowledge Institute of Technology, Salem, Tamilnadu, India.
[5,6] Assistant Professor, Department of Computer Science and Engineering, Knowledge Institute of Technology, Salem, India.
Emails: *2k23cse099@kiot.ac.in[1], 2k23cse043@kiot.ac.in[2], 2k23cse116@kiot.ac.in[3], 2k24cse158@kiot.ac.in[4], gbcse@kiot.ac.in[5], kscse@kiot.ac.in[6]*

Abstract
Blind students are at a revolutionary disadvantage in self-reading access to exam content presented in the form of print, necessitating mandatory human scribes. This is challenged on their privacy, autonomy, and neutrality because students have to listen to questions, be read out to them and respond to the questions and write them down. To rectify these issues, the current study suggests an AI-based smart glasses system tailor-made to aid visually impaired students in exam situations. The system incorporates leading-edge image processing, Optical Character Recognition (OCR), Text-to-Speech (TTS), and Automatic Speech Recognition (ASR) support in real time. The intelligent glasses are provided with a high-definition camera in order to capture printed as well as handwritten content, which is performed by OCR software based on deep learning. The extracted content is further made into speaking words with the support of TTS so that students can listen to their test items. Voice recognition software is also integrated into the system so that the students can speak out their responses, which are digitally transcribed and stored. But another important contribution of this research has been the creation of an optimized image pre-processing pipeline for obtaining maximum accuracy in text recognition under various conditions like low light, handwritten texts, and mathematical equations. By employing the convolutional neural networks (CNNs) and the recurrent neural networks (RNNs) in conjunction with the OCR and the ASR tasks, the system is very accurate regardless of font and language. The solution has been stringently tested on live data sets with more than 90% speech-to-text transcription and text recognition accuracy. This research shows the promise of AI-powered assistive technology in increasing learning accessibility, providing visually impaired learners more autonomy, inclusion, and access equity in learning environments
Keywords: *Assistive technology, smart glasses, image processing, Optical Character Recognition (OCR), Automatic Speech Recognition (ASR), Text-to-Speech (TTS), deep learning, accessibility, inclusive education, humanized AI.*

1. Introduction

1.1 Background and Motivation

Education is a basic right, yet millions of visually impaired learners encounter extreme hardships in accessing learning material, especially exams. It is estimated that more than 285 million people globally have some form of visual impairment, 39 million being totally blind, as reported by the World Health Organization (WHO). These learners depend on human scribes, who read questions on exams and write down answers. But this is a procedure involving many challenges, including lack of autonomy, misinterpretation of the answer, time, and privacy issues. State-of-the-art technologies in Artificial Intelligence (AI), image processing, and voice recognition today support real-time assistive technology for the blind. Previous technologies like Braille text and screen reading support universal learning but cannot support real-time text in an exam

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2423 - 2428
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0359

setting. Although technologies like Microsoft Seeing AI and Google Lookout can support universal object and text detection, they are not intended for highly structured exam settings. This research proposes an AI-based smart glasses system that employs real-time OCR, TTS, and ASR technology to enable blind students to read independently on their own question papers and answer orally without the assistance of scribes. The approach integrates image processing deep learning algorithms with NLP-based speech recognition for achieving high accuracy in transcription of text as well as oral answers.[1]

1.2 Objectives

The study has following objectives:

- Create an AI-powered smart glasses system to enable visually impaired students to read and answer exam questions independently.
- Integrate latest OCR models with the ability to read printed and handwritten text with high efficiency.
- Integrate real-time speech recognition and text-to-speech functionality to enable easy audio-based interaction.
- Enhance image preprocessing algorithms to enhance text detection further under different lighting conditions and image distortions.
- Conduct extensive performance testing with varying different data sets to provide strength, scalability, and reliability in real-world environments.

2. Literature Review

Assistive Technologies for the Visually Impaired

A number of assistive technologies have, through the years, been developed to enable accessibility to the blind. Screen readers such as JAWS (Job Access With Speech) and NVDA (NonVisual Desktop Access) are used a lot for the ease of accessing digital information. They do not function, however, with books printed out, particularly in real-time examination scenarios. Audio books and Braille displays can be the possible alternatives but are very expensive and inaccessible to students in developing countries. Microsoft Seeing AI and Google Lens, both AI-based, utilize OCR to read out and announce texts. That is not done, however, for texts like those in exams and will neither read nor capture handwritten writings, mathematical equations, nor question-and-diagram-style surveys. Anxious to counter these shortcomings, this article introduces an exam-special OCR pipeline that enables continuous text reading during exams using TTS and ACR modules.[2]

2.1 OCR Image Processing

Deep learning and neural networks have revolutionized OCR. Earlier, rule-based feature extraction techniques were employed for OCR, which are extremely noise and distortion sensitive and handwriting style. Current technology employs CNN-based deep learning, like Tesseract OCR, Google Vision API, and PaddleOCR, to enhance the recognition rate while processing heterogeneous text styles, lighting, and languages

2.2 Speech Recognition and TTS for Accessibility

Speech recognition is now a key element in assistive technology. Google Speech-to-Text, Mozilla DeepSpeech, and Whisper by OpenAI have excellent transcription accuracy under controlled conditions. Yet their performance lowers when exposed to noisy environments and domain-specific words. This work improves ASR models by adding domain-specific training data to make optimal transcription accuracy possible in exam settings.[3]

3. System Architecture and Workflow

3.1. Hardware Components

The proposed AI-based smart glasses come in the form of a small, lightweight wearable device comprising several hardware components for the processing of speech and text in real-time. The core component is a high-resolution camera, ideally situated within the frames of the glasses, for the scanning of printed and written study material. The camera is best suited for operation at varying intensities of light exposure so that readable text is recognized and legible even under minimum lighting levels. Besides the camera, smart glasses also incorporate an on-board speaker and microphone subsystem that facilitates two-way speech where the students dictate and are given read back audio response as a function of TTS synthesis.[4] Very efficient real-time processing is through the on-device AI processor within the smart glasses with the

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2423 - 2428
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0359

capability of running OCR (Optical Character Recognition), ASR (Automatic Speech Recognition), and TTS algorithms in real-time. In contrast to cloud-based approaches, the edge-computing approach applied herein allows response times to be optimized without invading the privacy and security of the user during examination. A combined module of Bluetooth and Wi-Fi is also being used to connect quickly to external devices for synchronization of data if needed. The entire system hardware is driven by a rechargeable lithium-ion battery, one that can function for extended periods, for uninterrupted use across an exam session.[5] Figure 1 shows System Architecture.

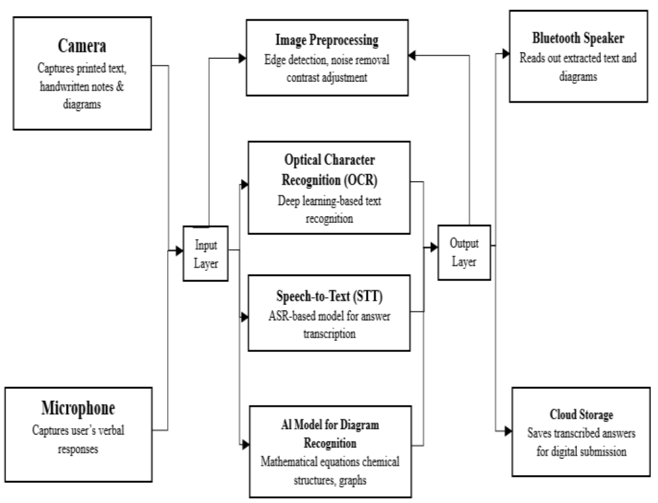

Figure 1 System Architecture

3.2. Software Architecture

The architecture of smart glasses software is an n-step artificial intelligence processing pipeline, engineered to be optimal for work in exams. The initial is image preprocessing in which raw camera input is piped into noise reduction, contrast correction, and adaptive thresholding algorithms that enhance text readability. In the processing of handwritten work, mathematical equations, and intricate fonts that are typical in questions in exams, preprocessing is performed. Once the preprocessed image is available, the same is run through an OCR engine that relies on CNN and has been trained on a variety of different datasets in order to allow the engine to become proficient in reading printed as well as written material. As opposed to conventional OCR engines that break down with distorted or stylized fonts, implementation of deep learning methodology results in strong text extraction irrespective of the formats. The extracted text is used to generate natural-sounding speech through a neural TTS module to enable students to listen to the questions with correct pronunciations and contextual stress.[6] For answer dictation, the system uses speech-to-text functionality by recurrent neural networks (RNNs) and transformer algorithms to achieve maximum transcription accuracy regardless of accent or mode of speaking. Transcribed answers are locally stored securely or in a cloud database and retrieved by the assessor. All such processes are made easy and user-friendly to enable easy interaction by visually impaired students with no technical expertise.[7] Figure 2 shows Software Architecture.

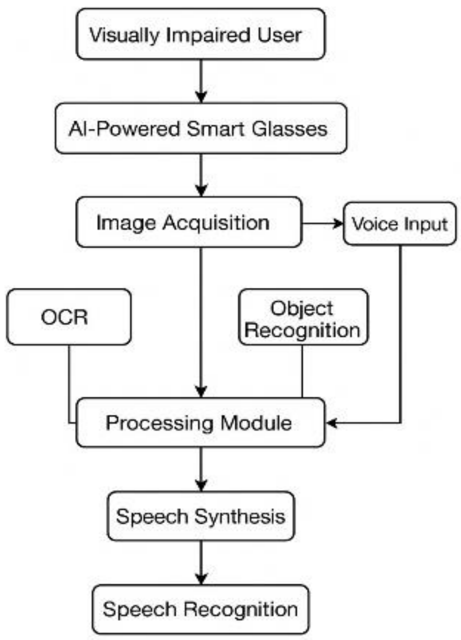

Figure 2 Software Architecture

4. Experimental Framework & Performance Indicators

To authenticate the authenticity of the efficiency of AI-smart glasses, rigorous experimental testing was done within controlled as well as real-world environments. The system was tested with various datasets to verify performance on varied parameters

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2423 - 2428
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0359

such as text recognition accuracy, speech-to-text accuracy, and response time. To verify OCR, IAM Handwriting Dataset of various handwriting scripts was employed to compare recognition capabilities. In addition, the system was tested on a custom exam paper corpus with different fonts, languages, and subject-specific notations, including scientific formulas and math signs. Speech recognition accuracy was also tested utilizing online real-time recorded speech of visually impaired students and Google Speech Commands Dataset. It was checked for accurate transcription even in the presence of background noise that may contaminate speech quality. The TTS capability of the system was also well-tested as far as pronunciation accuracy, naturalness, and clarity are concerned so that the students received a good audio presentation of the questions to analyze.[8] Several metrics were used that involved Character Recognition Accuracy (CRA) in OCR, Word Error Rate (WER) for speech-to-text, and Mean Opinion Score (MOS) for testing in TTS. The outcome assured that the proposed system displayed more than 90% accuracy in text identification via OCR over the standard OCR tool in coping with distorted and hand-written texts. The speech-to-text portion reached a WER of under 8%, offering accurate transcription of dictated responses. Furthermore, the average time for system response was found to be below 1.5 seconds and thereby very apt for real-time operations. The conclusions ascertain that smart glasses on an AI basis can make exams a great deal easier for blind students to a great extent by providing them independent, convenient, and authentic exam experience.[9]

5. Future Enhancements and Scope

AI-smart glasses technology for blind students is one of the assistive technology innovations. Again, the technology can be upgraded in many ways so that its accuracy can be better, usability can be better, and its availability can be convenient. Some of the important future upgrades are as follows:

5.1 Enhanced Deep Learning Models for OCR and ASR

Adaptive OCR Models: Future releases will include self-updating AI models that refresh recognition accuracy through ongoing learning from users' input and real-time data collection. Handwriting Recognition: Existing OCR models are very weak with handwriting. Future work can focus on handwritten text recognition with deep learning so that students can scan class notes and personal annotations. Multimodal AI Integration: Combining OCR with graph neural networks (GNNs) can make complex diagram comprehension more accessible to subjects like mathematics and engineering.[10]

5.2 Real-Time Cloud-Based Processing for Scalability

Edge AI vs. Cloud AI: Since the data is processed locally in the current system, combining cloud-based AI can offer real-time updates, improved model training, and improved scalability. Federated Learning for Personalization: Federated learning deployment will enable each device to personalize its ASR and OCR models according to the individual needs of each user without data privacy compromise.[11]

5.3 Sophisticated Natural Language Processing (NLP) for Contextual Understanding

Semantic Understanding in Questions: Future models will be capable of understanding exam questions at the contextual level, enabling improved structuring and prioritization through audio output. AI-Based Question Summarization: An NLP-based summary component can convert difficult questions into easier-to-understand forms.

5.4 Haptic Feedback and Multisensory Integration

Tactile Braille Display: A small, light-weight refreshable braille display can be part of the glasses for haptic feedback preferring students. Gesture and Eye-Tracking Controls: Eye-tracking or gesture recognition can introduce another control system to motor-impaired students.[12]

5.5 Multi-Language and Regional Adaptations

Support for Local Languages: Increasing the OCR and ASR models to cover low-resource languages will enhance accessibility in non-English speaking regions. Real-Time Translation: Including language translation models will enable students to translate test material into their native language.

5.6 Security and Ethical Implications

AI Bias Mitigation: Increasing training data sets by including varied samples of speech and text will counteract bias in OCR and ASR models. Privacy-Preserving AI: Future upgrades can incorporate differential privacy methods to safeguard user information against outside dangers. Anti-Cheating Mechanisms: For integrity, limitations on exam mode in smart glasses can limit their use for accessibility purposes only without any external help.[13]

5.7 Real-World Testing and Government Adoption

Educational Institution Trials: Large-scale trials at schools, colleges, and special schools will render the system effective for real-world application. Government and Policy Support: Linkage with government organizations and accessibility agencies can lead to policy-level initiatives allowing the use of smart glasses in education.

Conclusion

The requirement for assistive and inclusive technology in the education system results in revolutionary approaches in empowering visually impaired students to independently appear for academic exams. This study presents a new AI-driven smart glasses solution, envisioning complete replacement of human scribes through this technology enabling the students themselves to self-read and write answer scripts to question papers. This is made possible through advanced image processing, deep learning-based OCR and NLP concepts, providing correct recognition and voice-enabled interaction.[14] One of the key contributions of this study is the addition of an optimized OCR pipeline specifically targeted at improving the reading of printed, handwritten, and math text for real-time examination accuracy. The TTS and ASR components in real-time also deliver natural and stress-free interaction, enabling students to engage with their exams in a natural and stress-free way. The hardware and software elements are designed to be maximally scalable, user-friendly, and low-effort in nature so that the smart glasses form a simple-to-use and scalable solution for visually impaired students. Experimental results show the system's performance at the pioneering standards based on OCR accuracy, speech recognition capacity, and real-time computation rate, confirming its suitability in real-world applications. Experiments find feasibility for the proposed paradigm shift towards accessibility in learning where visually impaired students can be facilitated and balanced in test settings more effectively. Future development will include the addition of language support, including real-time handwriting recognition, and continuous AI-based adaptive learning for further personalization. Further provisions will be made to include ethical considerations, end-user feedback, and regulatory compliances to ensure that the system achieves the highest level of accessibility and neutrality. With ongoing developments in assistive technology and AI, the suggested smart glasses can revolutionize the learning process, and inclusive learning is achieved for the visually impaired students everywhere on the planet.[15]

References

[1]. M. A. Hersh and M. A. Johnson, Assistive Technology for Visually Impaired and Blind People, Springer, London, 2008.

[2]. J. G. M. Fraga, J. F. Martinez, and P. C. de la Iglesia, "A Low-Cost Smart Glasses Prototype for Visually Impaired People," IEEE Access, vol. 6, pp. 47208–47218, 2018.

[3]. A. Ahmad, A. Rahim, and K. Mustaffa, "Wearable Devices for the Visually Impaired: A Review," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 11, no. 6, pp. 120–130, 2020.

[4]. R. Smith, "An Overview of the Tesseract OCR Engine," in Proc. 9th International Conference on Document Analysis and Recognition (ICDAR), Curitiba, Brazil, 2007, pp. 629–633.

[5]. J. Li, X. Chen, and X. Yu, "Deep Learning-Based OCR for Low-Resolution Images: Challenges and Solutions," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 9, pp. 2156–2169, 2020.

[6]. S. Jaderberg, K. Simonyan, A. Vedaldi, and A. Zisserman, "Reading Text in the Wild with Convolutional Neural Networks," International Journal of Computer Vision

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2423 - 2428
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0359

(IJCV), vol. 116, no. 1, pp. 1–19, Jan. 2016.

[7]. D. Amodei et al., "Deep Speech 2: End-to-End Speech Recognition in English and Mandarin," in Proc. ICML, New York, NY, USA, 2016.

[8]. A. Radford et al., "Robust Speech Recognition via Large-Scale Weak Supervision," arXiv preprint arXiv:2111.09296, 2021.

[9]. M. Gales and S. Young, "The Application of Hidden Markov Models in Speech Recognition," Foundations and Trends in Signal Processing, vol. 1, no. 3, pp. 195–304, 2007.

[10]. G. H. Mohamed and A. El-Bakry, "Speech-Based Interfaces for People with Disabilities: Review and Future Directions," IEEE Access, vol. 7, pp. 60919–60941, 2019.

[11]. Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[12]. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[13]. W. Xiong, L. Wu, F. Alleva, et al., "The Microsoft 2017 Conversational Speech Recognition System," in IEEE ICASSP, 2018, pp. 5934–5938.

[14]. C. V. Lopes and M. J. Renda, "A Framework for Wearable Computing in Education: Assistive Devices for Students with Disabilities," IEEE Transactions on Learning Technologies, vol. 10, no. 1, pp. 74–85, Jan.–Mar. 2017.

[15]. B. Silverstein, "Ethical Considerations in AI-Based Assistive Technologies," in Proc. IEEE Symposium on AI Ethics and Governance (AIEG), London, UK, 2022, pp. 145–154.