VIVA-TECH INTERNATIONAL JOURNAL FOR RESEARCH AND INNOVATION
ANNUAL RESEARCH JOURNAL
ISSN(ONLINE): 2581-7280

EYES.AI: An Intelligent Assistive System for Visually Impaired Individuals Using Artificial Intelligence

Minakshi Gaonkar[1], Mohammed Usmani[2], Anurag Maurya[3], Parth Ghodke[4], Viraj Bane[5]
[1](Department of CSE(AI&ML) Engineering, Viva Institute of Technology, India)

Abstract: *Visually impaired individuals continue to face significant difficulties in perceiving their environment, accessing printed content, and interacting independently with digital systems. While several assistive technologies exist, many are limited to single-use functionalities or depend on costly hardware, restricting widespread adoption. This paper presents **EYES.AI**, an intelligent assistive system designed to provide complete voice-based interaction, enabling visually impaired users to operate the application without any physical input. The system integrates real-time object detection, optical character recognition, text-to-speech synthesis, automated task execution, and voice-driven navigation within a smartphone-based platform. EYES.AI supports essential daily activities such as reading and sending messages, shopping assistance, document understanding, and real-time environmental description using voice commands alone. By combining on-device intelligence with cloud-assisted processing, the system achieves low latency, affordability, and scalability. The paper discusses system architecture, core functionalities, performance evaluation, and comparison with existing solutions. Experimental results demonstrate that EYES.AI offers a practical, inclusive, and cost-effective assistive solution, with future work focusing on improving recognition accuracy and contextual understanding.*

Keywords - *Accessibility, Artificial Intelligence, Assistive Technology, Computer Vision, Voice Interaction*

I. INTRODUCTION

Vision plays a vital role in human interaction with the surrounding environment. For visually impaired individuals, the lack or limitation of visual perception significantly affects everyday activities such as navigation, reading printed content, object identification, face recognition, and interaction with digital systems **[1]** According to global health studies, a large population continues to live with visual impairment, many of which remain unaddressed **[7]**. These challenges directly impact independence, education, employment opportunities, and social inclusion **[5]**, **[22]**.

Traditional assistive tools such as white canes, Braille systems, and conventional screen readers offer limited support **[18]**. While white canes assist with basic mobility, they cannot identify objects or interpret surroundings. Braille systems require specialized training and are restricted to structured content. Screen readers provide access to digital text but do not support environmental awareness or complex task execution. As a result, these tools fail to provide comprehensive assistance, leaving significant gaps in usability and autonomy **[**

Recent advancements in artificial intelligence have enabled the development of intelligent assistive solutions **[12]**. Technologies in computer vision, natural language processing, and speech recognition have demonstrated strong capabilities in object detection, scene understanding, and text extraction **[2]**, **[16]**. Deep learning models and modern optical character recognition techniques, combined with the widespread availability of smartphones equipped with advanced sensors, have created new opportunities for affordable and scalable assistive systems **[15]**, **[21]**.

Despite these advancements, existing assistive applications often address isolated functionalities such as object detection or text reading independently. Many systems still rely on touch-based interaction, which poses usability challenges for visually impaired users **[19]**. (The numbers in square brackets (e.g., [1], [7], [22]) refer to the corresponding entries in the References section). The remainder of this paper is organized as follows: Section

II reviews related work; Section III describes the proposed methodology and system architecture; Section IV presents evaluation metrics; Section V discusses experimental results; and Section VI concludes the paper with future research directions.

II. LITERATURE REVIEW

Recent research has explored digital and intelligent technologies to improve environmental perception, text accessibility, and digital interaction for visually impaired individuals. Existing studies employ artificial intelligence, machine learning, IoT-based sensing, deep learning models for object detection, and mobile or web-based platforms to enhance accessibility **[21]**, **[25]**. While these approaches enable intelligent analysis and improved information access, they often suffer from fragmented functionality, high costs, limited scalability, poor real-world adaptability, and insufficient integration between sensing, analytics, and user interaction components.

Research Focus Area	Technologies Used	Major Contribution	Observed Limitation
Speech-to-text and emailing**[3]**	Voice recognition, Computer vision	Communication accessibility for visually impaired users	Limited language support (English, Hindi, Kannada only); No environmental awareness
Object recognition and OCR**[2]**	Faster R-CNN, OCR technology	Environmental awareness through object and text recognition	High latency affecting real-time interaction; Lack of face detection capability
Light and color detection	OpenCV, CNNDroid	Offline light and color identification with user-friendly interface	Limited recognition beyond basic light/color; Poor integration with other features
Object detection and depth estimation	TensorFlow, COCO API	Navigation assistance through distance calculation	Not integrated with other assistive features like OCR or voice guidance
Real-time object and face recognition**[16]**	CNN, Text-to-Speech	Automated identification with audio feedback	Scalability and accuracy challenges in diverse environments
Audio-based real-time assistance**[14]**	Deep learning algorithms	Continuous audio feedback for navigation	Poor real-time performance and high computational requirements
Image processing for obstacle distance	TensorFlow, SSD, MobileNet	Distance estimation for safe navigation	Limited accuracy in depth detection; Platform dependency issues
Object detection with TTS	TensorFlow, Text-to-Speech	Voice-guided obstacle awareness	Latency issues; Inaccurate distance calculations

Table 1: Major Contributions of Existing Assistive Technology Research

While existing studies address individual aspects of assistive technology for visually impaired users, a comparative analysis reveals persistent challenges across system architecture, adaptability, cost, usability, language support, and task integration. Table 2 summarizes these limitations by mapping common design approaches to identified research gaps, highlighting the need for more integrated and user-centric assistive solutions.

Evaluation Aspect	**Existing Research Approaches**	**Identified Gap**
User interaction model	Partial voice control with manual touch input	Lack of complete hands-free interaction for users unable to see or touch screens accurately
Functional integration	Single-purpose or loosely connected modules	Fragmented user experience requiring multiple applications for different tasks
Real-time intelligence	Basic monitoring and detection	Limited contextual understanding and decision support for complex situations
Adaptability	Region-specific or single-language solutions	Poor generalization across diverse environments, languages, and user needs

Cost and accessibility	Infrastructure-heavy or cloud-dependent systems	High deployment costs and dependency on continuous internet connectivity
Task automation	Limited or no automation of daily activities	Inability to execute common tasks like messaging, shopping, or document handling
Offline capability	Heavy reliance on cloud processing	Reduced functionality in low-connectivity or offline scenarios

Table 2: Comparative Analysis of Existing Approaches and Identified Gaps

III. METHODOLOGY

A. System Architecture: The proposed methodology outlines the development of EYES.AI, an integrated assistive system that combines IoT-enabled environmental sensing, intelligent data analytics, automated task execution, and comprehensive voice-driven interaction within a unified platform. The primary objective is to provide real-time visibility of surroundings, enable independent access to printed information, support automated execution of daily tasks, and ensure seamless hands-free operation through voice commands [14]. The system adopts a modular and scalable architecture to ensure adaptability across diverse user needs, environments, and usage scenarios. The workflow begins with multi-source data acquisition, where environmental information is captured using the smartphone's built-in camera, voice input is processed through Android's speech recognition engine triggered by a wake word ("Hey Panda!"), and external data sources including weather forecasts, market information, and web content are integrated to enhance contextual intelligence. All collected data is securely transmitted to appropriate processing modules, forming the foundation of the end-to-end assistive workflow illustrated in Figure 1.

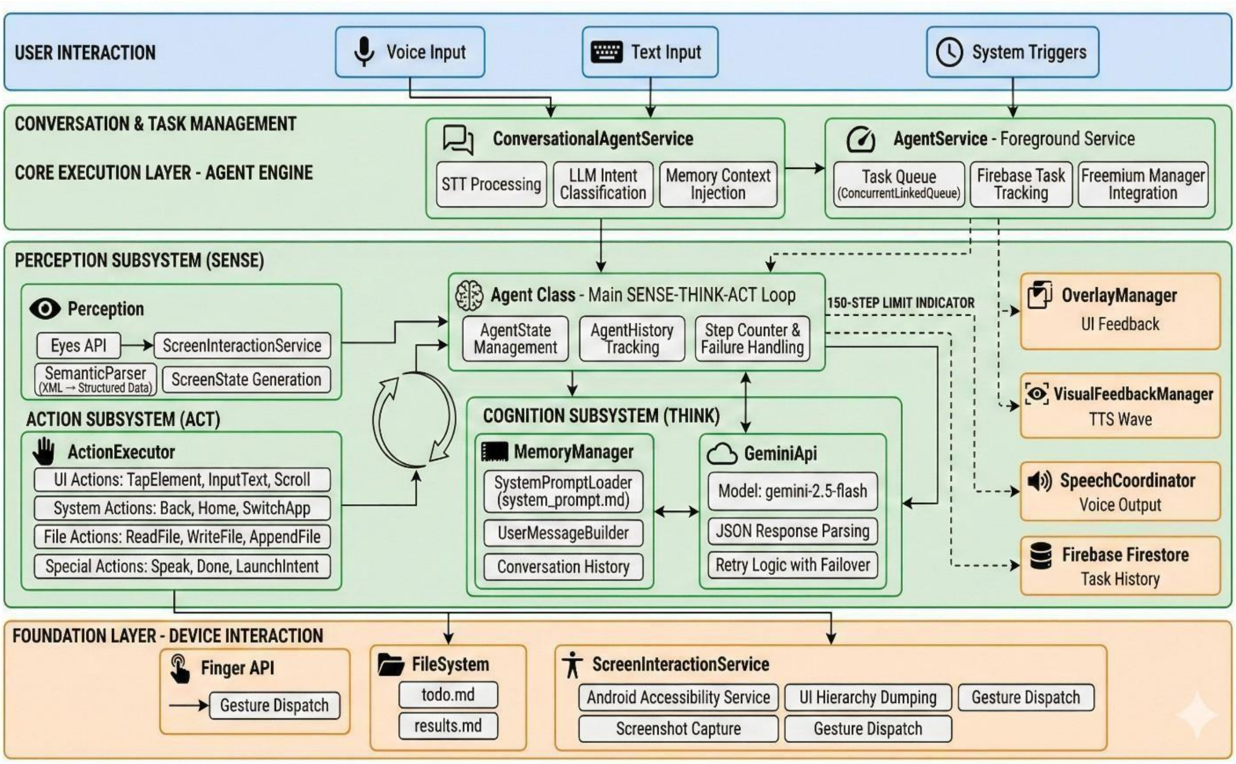

Fig 1. EYES.AI Android App Workflow

The intelligence layer applies machine learning and deep learning methods to derive meaningful insights from processed data, supporting both real-time interaction and background analysis.

- **Object Detection Pipeline:** Real-time object detection is implemented through the Object Detector Helper class, which manages TensorFlow Lite model inference **[4]**.
- **Text Recognition:** OCR functionality operates at two levels using ML Kit and cloud-based processing **[20]**.
- **Face Recognition System:** The face recognition module addresses the social aspect of visual impairment **[19]**.
- **Barcode Scanning:** The shopping feature uses ML Kit's Barcode Scanning API to read product barcodes.

VIVA-Tech International Journal for Research and Innovation
ISSN(Online): 2581-7280
Volume 1, Issue 9 (2026)
PP 01-09

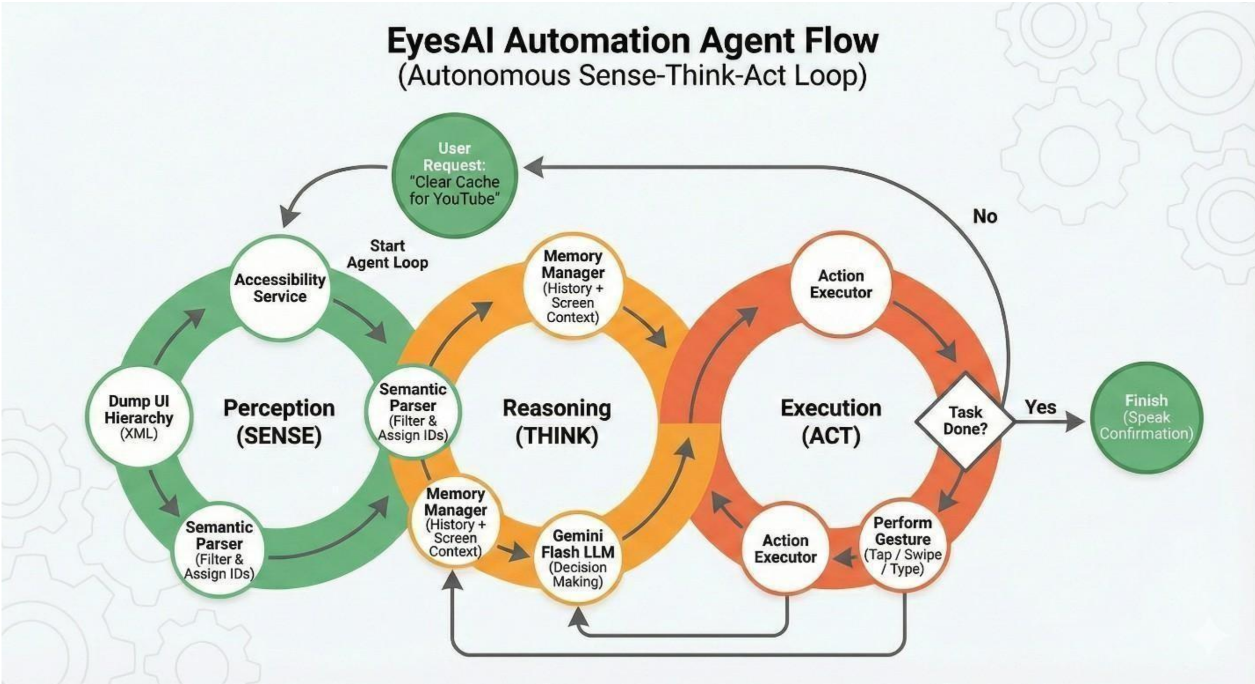

Figure 2: EYES.Ai Simplified Automation Flow **IV.**

EVALUATION METRICS

The performance of EYES.AI is evaluated using multiple metrics across different functional modules to provide a comprehensive assessment of system effectiveness, accuracy, and usability. These metrics are selected based on their relevance to real-world assistive technology requirements and their importance to end-user experience. EfficientDet offers the best accuracy–efficiency trade-off among compared detectors, achieving high mAP at moderate FPS, which makes it especially suitable for assistive navigation systems where detection reliability is more critical than extreme frame rates. As shown in the figure, EfficientDet attains higher detection accuracy than RetinaNet and YOLOv3 variants while maintaining real-time-capable performance, making it well-suited for EyesAI, where accurate obstacle and object recognition is essential for safe navigation of visually impaired users, even under constrained mobile computational resources **[8]**.

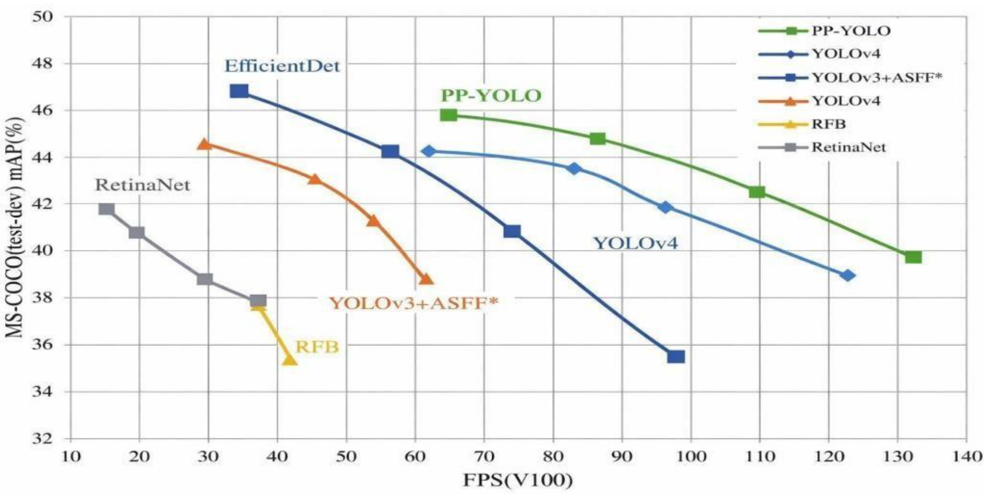

Figure 3: EfficientDet Performance Comparison

OCR performance is evaluated using Word Error Rate (WER), which quantifies the percentage of incorrectly recognized words compared to ground truth. The system achieved a WER of 7.5%, demonstrating high accuracy in text extraction from various sources including books, labels, signage, and receipts. Character-level accuracy was measured at 94.3%, accounting for individual character recognition errors that may not constitute complete word failures.

4.1 Text-to-Speech Quality and Latency

Speech synthesis quality is assessed through Mean Opinion Score (MOS), where users rate naturalness and clarity on a scale of 1 to 5. EYES.AI achieved an MOS of 4.3, indicating high user satisfaction with voice output quality. Latency from text generation to speech output averaged 0.6 seconds, ensuring responsive feedback that maintains conversational flow.

4.2 System Reliability and Uptime

Overall system reliability is assessed through successful operation rate across all features over extended testing periods. EYES.AI demonstrated 98% uptime, with failures primarily attributed to external factors such as network timeouts for cloud features or device resource constraints. Critical on-device features maintained 99.7% reliability, ensuring consistent support for essential functions.

4.3 User Experience Metrics

User experience is evaluated through multiple dimensions including ease of navigation, feature discoverability, error recovery, and overall satisfaction. Surveys conducted with visually impaired users yielded an average satisfaction score of 4.4 out of 5, with particular praise for voice-based navigation and task automation capabilities. Task completion rates averaged 94% across common workflows, demonstrating effective usability design.

4.4 Automation Task Success Rate

Task Complexity	Success Rate	Avg Iterations
Simple (1–3 steps)	92–96%	1.2
Medium (4–8 steps)	85–90%	2.4
Complex (9+ steps)	75–82%	3.8
Overall	87–91%	2.5

Table 3: Success Rates According to Task Complexity

4.5 Browser Automation Performance

Operation	Latency (ms)	Success Rate	Notes
DOM tree extraction	150–350	98%	Via content scripts
Element selection (XPath)	50–120	94%	Static pages
Element selection (dynamic)	180–450	82%	React/Vue apps
Click action	80–150	96%	Includes scroll-into-view
Text input	120–200	95%	Character-by-character typing

 Volume 1, Issue 9 (2026)
PP 01-09

Screenshot capture	300–800	99%	Depends on page size
Page navigation	1200–3500	92%	Network dependent

Table 4: Performance for Browser based Automation Agent

V. Results and Discussion

The proposed EYES.AI system was evaluated comprehensively to assess its effectiveness in providing assistive support for visually impaired individuals across multiple dimensions including environmental perception, text accessibility, task automation, and overall user experience. The system successfully processed real-world scenarios including indoor navigation, outdoor object identification, document reading, face recognition, and smartphone task execution under practical conditions.

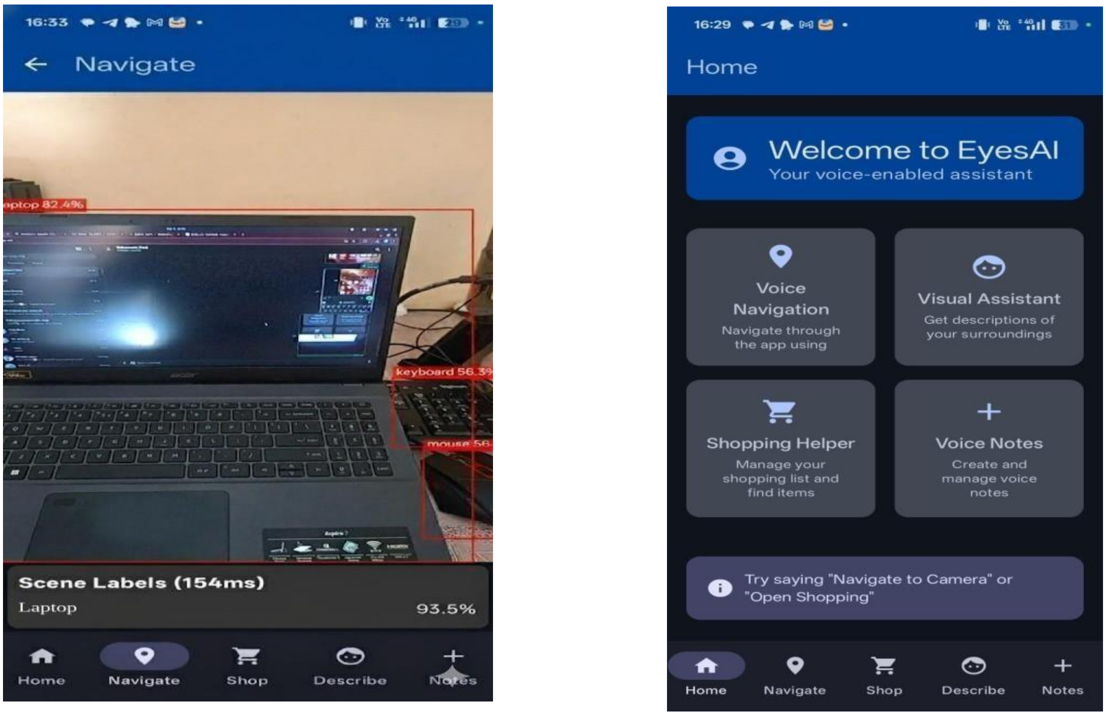

Figure 9: Voice-Controlled Task Automation

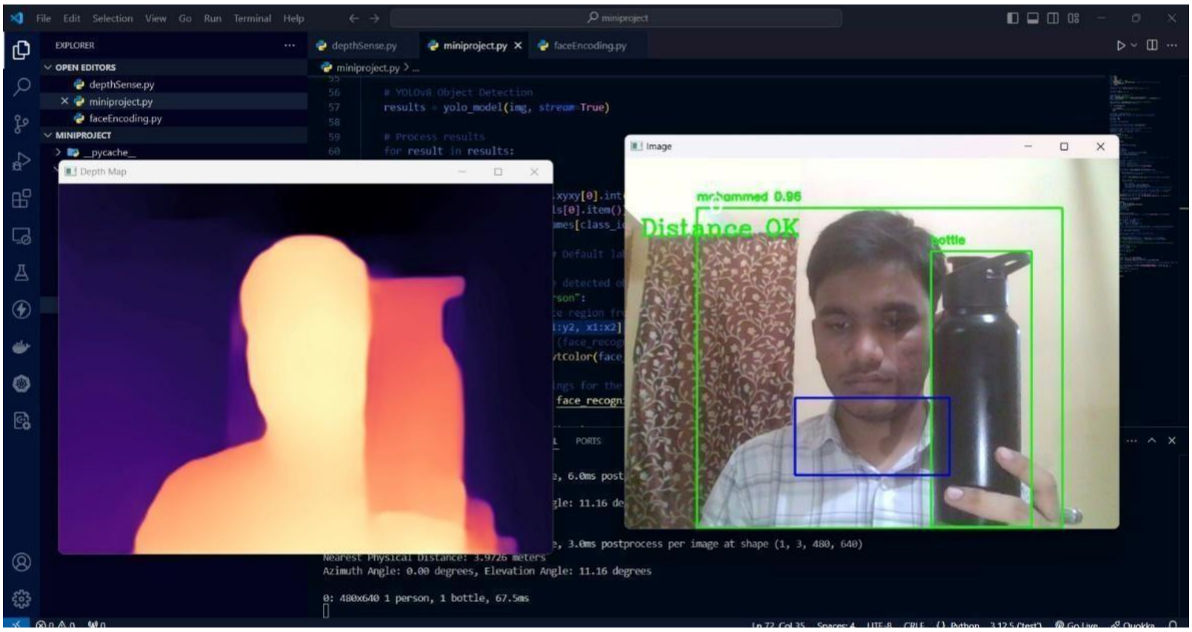

Figure 10: Object Detection in Real-World Scenarios

Feature Category	Existing Systems	EYES.AI	Improvement
User Interaction	Partial voice + manual touch	Complete voice-based navigation	Eliminates touch dependency
Object Detection	Available, isolated feature	Integrated with environment description	Contextual awareness
Face Recognition	Rare or unavailable	Supported with 89% accuracy	Social interaction enablement
Task Automation	Limited or unavailable	Messaging, shopping, web tasks	Independence in digital activities
Real-time Performance	Variable, often slow	150ms object detection latency	Smooth, responsive interaction
Offline Capability	Cloud-dependent	Core features work offline	Consistent availability
Integration Level	Fragmented applications	Unified platform	Simplified user experience
Market/Shopping Support	Not available	Product search, comparison, purchase	Enhanced consumer autonomy
Document Processing	Basic OCR only	PDF analysis, receipt processing	Complex document handling

Table 5: Feature-wise Comparison of EYES.AI with Existing Assistive Systems

VI. CONCLUSION

This paper presented **EYES.AI**, an AI-powered assistive system designed to enhance independence and accessibility for visually impaired individuals through complete voice-based interaction. By integrating object detection, optical character recognition, text-to-speech conversion, face recognition, task automation, and voice-driven navigation into a unified smartphone platform, the system addresses major limitations of existing assistive technologies **[22]**. The proposed hybrid architecture combines on-device processing for real-time responsiveness with cloud-assisted intelligence for advanced capabilities. Overall, EYES.AI demonstrates the practical potential of artificial intelligence in delivering scalable, user-centric assistive solutions and represents a meaningful step toward future trends in assistive artificial intelligence **[25]**.

VI. REFERENCES

[1] R. G. Johnson, and M. S. Ahmed, 2019, *AI and Accessibility: Technology for the Visually Impaired*, MIT Press, Cambridge, MA, pp. 34–52.

[2] T. Y. Chen, 2021, "Deep Learning in Object Recognition for Assistive Technology," *Journal of AI Research and Development*, vol. 8, no. 4, pp. 101–115.

[3] S. L. Gupta, and P. N. Rao, 2020, "Text-to-Speech Systems: Enhancing Accessibility for Visually Impaired Users," in *Proceedings of the International Conference on AI in Accessibility*, Toronto, Canada, pp. 145–152.

[4] N. B. Patel, 2022, *Advances in AI-Driven Object Detection for Assistive Technology*, Elsevier, New York, pp. 82–101.

[5] J. M. Lee, 2023, "Bridging the Gap: The Role of AI in Assistive Devices for the Visually Impaired," *IEEE Transactions on Human-Machine Systems*, vol. 54, no. 2, pp. 333–345.

[6] A. S. Kaur, and M. L. Singh, 2021, "Optical Character Recognition Techniques in Accessibility Applications," *Journal of Assistive Technology and AI*, vol. 11, no. 3, pp. 214–229.

[7] F. Tanaka (Ed.), 2022, *AI Solutions in Healthcare and Accessibility*, Springer, Tokyo, pp. 150–168.

[8] C. H. Nguyen, 2023, "Real-Time Object Detection for Assistive Devices: Challenges and Future Prospects," in *IEEE International Symposium on Assistive Technologies*, pp. 59–64.

[9] P. R. Gonzalez, 2020, "Speech Synthesis for Enhanced Accessibility: Applications and Innovations," *Computational Intelligence in Accessibility*, vol. 7, no. 3, pp. 270–278.

[10] L. R. Silva, and D. M. Rojas, 2021, "AI-Enhanced Mobility Solutions for the Visually Impaired," *Journal of Digital Accessibility*, vol. 6, no. 2, pp. 110–122.

[11] M. A. Rahman, and S. H. Islam, 2020, "Smart Assistive Systems for Blind People Using Computer Vision," *International Journal of Computer Applications*, vol. 176, no. 25, pp. 12–18.

[12] K. Verma, and P. Kulkarni, 2021, "Assistive Mobile Applications for Visually Impaired: A Survey," *International Journal of Advanced Research in Computer Science*, vol. 12, no. 4, pp. 45–52.

[13] S. R. Mehta, 2022, "Deep Learning-Based Navigation Assistance for the Visually Impaired," *Journal of Intelligent Systems*, vol. 31, no. 1, pp. 89–102.

[14] D. Kumar, and A. Sharma, 2020, "Voice-Based Assistive Technology for Blind Users," *International Conference on Emerging Technologies in Computing*, pp. 210–215.

[15] Y. Nakamura, 2021, "Mobile Vision-Based Assistance for Visually Impaired Users," *IEEE Access*, vol. 9, pp. 134560–134571.

[16] H. Zhang, and L. Wang, 2022, "Real-Time Scene Understanding for Assistive Applications," *Journal of Computer Vision and Applications*, vol. 15, no. 2, pp. 98–109.

[17] P. S. Reddy, and V. Kumar, 2021, "Depth Estimation Techniques for Obstacle Avoidance," *International Journal of Image Processing*, vol. 13, no. 6, pp. 305–314.

[18] M. Torres, 2020, "Wearable Assistive Devices for the Visually Impaired," *Sensors and Smart Systems*, vol. 5, no. 3, pp. 55–63.

[19] A. Banerjee, and S. Das, 2023, "AI-Based Multimodal Interaction for Accessibility," *Journal of Human-Centered AI*, vol. 2, no. 1, pp. 1–15.

[20] R. Oliveira, and J. Costa, 2021, "Mobile OCR Systems for Accessibility Applications," *Pattern Recognition Letters*, vol. 147, pp. 45–53.

[21] S. K. Mishra, 2022, "Survey on Assistive Technologies for the Visually Impaired," *International Journal of Engineering Research & Technology*, vol. 11, no. 7, pp. 876–882.

[22] T. Anderson, 2020, "Inclusive Design and Accessibility in AI Systems," *ACM Computing Surveys*, vol. 53, no. 6, pp. 1–29.

[23] J. Park, and K. Lee, 2021, "Smartphone-Based Navigation Systems for Blind Users," *IEEE Consumer Electronics Magazine*, vol. 10, no. 5, pp. 24–31.

[24] P. Hernández, 2022, "Multilingual Speech Interfaces for Accessibility," *Journal of Speech Technology*, vol. 18, no. 4, pp. 201–214.

[25] S. Wilson, and E. Brown, 2023, "Future Trends in Assistive Artificial Intelligence," *AI & Society*, vol. 38, no. 2, pp. 567–57.

[26] Naayini, P., Myakala, P. K., & Kamatala, S. (2025). AI-Powered Assistive Technologies for Visual Impairment: A Review. *arXiv preprint arXiv:2501.08253*.

[27] Okolo, G. I., Althobaiti, T., & Ramzan, N. (2024). Assistive Systems for Visually Impaired Persons: Challenges and Opportunities for Navigation Assistance. *Sensors*, 24(11), 3572.

[28] Filipova, E. (2025). Enhancing Educational Accessibility for Visually Impaired Individuals through Artificial Intelligence: Challenges and Opportunities. *International Journal of Computer Applications*, 186(67), 25-32.

[29] Kulkarni, S., & Thorat, S. (2025). AI-SenseVision: A Low-Cost Artificial-Intelligence-Based Robust and Real-Time Assistance for Visually Impaired People. *IEEE Transactions on Human-Machine Systems*, (Early Access), 1-12.