

Vision Mate: An AI Based Mobile Assistive System for Visually Impaired Users

SARTHAK MARUTI MORE¹, VIJAYRAJ CHAVAN², ABHISHEK KUNDKAR³, SOHEL NADAF⁴,
SMT. M. M. CHAITANYA⁵

^{1, 2, 3, 4, 5}Information Technology, Government Polytechnic Solapur

Abstract- *VisionMate is a mobile accessibility application that utilizes artificial intelligence (AI) and helps visually impaired and blind people to carry out daily tasks independently. The device proposed integrates computer vision, speech recognition, and GPS to utilize the camera, microphone, and location services on a user's smartphone to assist users with a variety of tasks. VisionMate also provides a multitude of functions for people with visual impairments, including Optical Character Recognition (OCR) for reading text, Object Detection for finding people, currency detection for identifying denomination, Barcode Scanning to read product/food labels, and Voice Navigation for providing direction to users via voice command. Uses of cloud services (Google Vision API & Speech-to-Text API) to store incoming images and voice commands allow for the analysis of received images and voice commands. The image results are returned to the user via real-time sound using Text-to-Speech (TTS) technology. VisionMate will be developed using the React Native framework, and MongoDB will serve as the backend database for user management and authentication. Ultimately the goal of the proposed system is to improve access to and independence for visually impaired users by allowing users to better understand the environment using intelligent audio guidance.*

Keywords- *Artificial Intelligence, Assistive Technology, Computer Vision, Mobile Accessibility, Optical Character Recognition (OCR)*

I. INTRODUCTION

The visual impairment epidemic presents a serious problem all over the world because it has a direct impact on everyday living and independent functioning for millions of people. Global health statistics indicate that many individuals are partially or completely blind and, therefore, cannot perform basic daily tasks such as finding their way around, reading printed materials, recognizing items, or identifying dollar bills. As a result, assistive devices have become an important area of research focused

on creating devices and systems that enhance the accessibility and independence of individuals with visual impairments. New advancements in AI, Computer Vision, and Mobile Computing technologies have created new opportunities to develop intelligent systems that assist visually impaired individuals in navigating through everyday conditions [1], [2].

Traditional assistive tools (white cane/glass dog) have very limited capabilities in terms of helping users comprehend their immediate environment in detail. To overcome the limitations of traditional assistive tools have led to the development of smart assistive systems that use cameras, sensors, and computer vision technologies to help the visually impaired recognize and interpret their surroundings. Computer vision technology such as object recognition, facial recognition, and optical character recognition (OCR) are currently being used by different researchers to determine the effectiveness of AI-based vision systems for the visually impaired, such as detecting obstacles, identifying objects, and converting text from printed materials into sound for visually impaired users.

The growing capabilities of mobile devices, including their state-of-the-art high-resolution cameras, microphones, GPS modules, and processing power, have made these devices a strong platform for developing assistive technologies. As a result, mobile applications can capture visual information from the environment, analyze it through machine learning algorithms, and provide users with real-time feedback. Many researchers have investigated the potential of using a smartphone as a tool to assist people who are blind or visually impaired through solutions such as text recognition, scene understanding, and navigation. In many instances,

these projects have combined computer vision algorithms with speech recognition and text-to-speech technologies to create hands-free, voice-based interfaces for interacting with smartphones [6], [7], [8].

In a short period, the cloud-based AI services that are converting images into text (e.g., Google Vision API) and those that convert speech into text have increased the efficiency and accuracy of mobile assistive applications. These new technologies enable mobile applications to perform other complex and difficult tasks, such as determining the location of an object, analyzing the emotion on a person's face, and identifying a product by scanning its barcode, reliably. Many GPS-based navigation applications provide voice instruction to assist their users in navigating through unfamiliar environments. The combination of computer vision and speech technologies clearly enhances the ability of visually impaired individuals to gain access to various modes of transportation and navigate their communities more effectively [9], [10], [11].

Current technologies for assisting disabled persons are ridiculously expensive, hard to use, nonfunctioning, and usually designed for one specific function (i.e., detecting obstacles or reading text) instead of providing a comprehensive assistive solution. This leaves a significant gap in the availability of a fully integrated and usable mobility assistive device that integrates all types of functions into one mobile device [12], [13], [14].

This paper presents VisionMate, an artificial intelligence-powered mobility assistance application to help blind users through an enhanced level of assistance. VisionMate integrates three new technologies together (computer vision, speech processing, and navigation) to create a comprehensive assistive system. The application contains multiple functionalities by utilizing the combination of available hardware on smart devices used by the user and utilizing cloud-based artificial intelligence services on a server to perform multiple tasks including reading printed text via OCR, identifying objects and faces, identifying currency, scanning barcodes, and providing audio navigational

directions based on GPS. By transforming visual information into audible information, VisionMate provides a user-friendly means for blind users to access the world around them in a safe and efficient manner.

This proposed system's primary objective is to give people with visual impairments more independence and mobility by providing them with real-time environmental information in an accessible voice interface. VisionMate clearly shows how current mobile technology combined with artificial intelligence creates useful assistive technologies that improve the daily lives of users who have vision problems.

II. IDENTIFY, RESEARCH AND COLLECT IDEA

The first step involved reading already published research work in the same field to understand the current state of assistive technologies. Numerous studies have explored the use of computer vision algorithms and deep learning techniques to detect objects, recognize faces, and read text through Optical Character Recognition (OCR). These studies demonstrated that integrating camera-based vision systems with speech output can significantly enhance accessibility for visually impaired users. Research on mobile assistive applications also showed that smartphones provide a powerful platform due to their built-in sensors such as cameras, microphones, and GPS modules [6], [7], [8], [9], [10].

The second step involved searching and collecting information through online resources and academic databases to explore different approaches and technologies used in assistive systems. By reviewing research articles, journals, and conference papers available online, various technological solutions were analyzed, including object detection models, cloud-based vision APIs, speech recognition systems, and GPS-based navigation tools. These studies highlighted how machine learning and artificial intelligence can be integrated into mobile applications to provide real-time assistance and environmental awareness for visually impaired individuals [11], [12], [13], [14], [15].

Another important aspect of the research process involved participating in academic conferences, workshops, and symposiums related to artificial intelligence, computer vision, and assistive technologies. Such events provide valuable insights into the latest advancements in the field and allow researchers to learn from experts working on similar problems. Discussions and presentations in these forums often reveal practical challenges and innovative solutions for developing assistive systems. Research presented at international conferences on accessibility technologies and intelligent systems has contributed significantly to the development of integrated assistive applications for visually impaired users.

The final step involved understanding the scientific concepts, terminology, and technical frameworks associated with the research area. Key concepts such as computer vision, deep learning, object detection, optical character recognition, speech recognition, and GPS-based navigation were studied in detail. Familiarity with these technologies helped in designing a comprehensive system that integrates multiple assistive features into a single platform. Previous research has shown that combining these technologies can significantly improve the efficiency and usability of assistive applications for visually impaired individuals.

Based on this extensive review and analysis of existing research, the concept of the VisionMate system was formulated. The proposed system integrates computer vision, speech processing, and navigation technologies within a smartphone application to assist visually impaired users in real time. The application includes functionalities such as text reading through OCR, object and face detection, currency identification, barcode scanning, and voice-guided navigation. By converting visual information into audio feedback, the system aims to provide a practical and accessible solution that improves mobility, independence, and safety for visually impaired individuals. The feasibility of this approach is supported by several previous studies that emphasize the effectiveness of AI-based assistive

systems in enhancing accessibility and quality of life for visually impaired users.

III. WRITE DOWN YOUR STUDIES AND FINDINGS

After identifying the research problem and reviewing previously published work, the next step is to organize and present the collected ideas and findings into a structured research paper. For the proposed VisionMate: An AI-Based Mobile Assistive System for Visually Impaired Users, the research findings were compiled by analyzing various studies related to assistive technologies, computer vision systems, and mobile accessibility applications. These studies provided valuable insights into the design of systems capable of converting visual information into voice feedback for visually impaired individuals. Research on object detection, optical character recognition (OCR), and speech processing helped in understanding how artificial intelligence techniques can be integrated with smartphone hardware to assist users in daily activities such as reading text, recognizing objects, and navigating their surroundings [1], [2], [3], [4], [5].

A. Bits and Pieces together

In this approach, the information gathered from different research papers, journals, and technical resources was carefully combined to form the foundation of the proposed system. Existing research works on AI-based assistive technologies were studied to understand their methodologies, strengths, and limitations. Many studies have demonstrated the effectiveness of computer vision algorithms in detecting objects and recognizing text from images. Similarly, research on speech recognition and text-to-speech systems showed how voice-based interaction can improve accessibility for visually impaired users. By reviewing these studies, the key concepts and techniques required for building an integrated assistive system were identified and adapted for the VisionMate application. The integration of multiple functionalities such as OCR-based text reading, object detection, face recognition, barcode scanning, and currency identification was inspired by previously accomplished research in the field of intelligent assistive technologies [6], [7], [8], [9], [10], [11], [12].

The “Jump Start” approach was also followed during the development of this research work. In this approach, the ideas and methodologies were refined with guidance from fellow researchers, project mentors, and academic supervisors. Continuous discussions and feedback helped in improving the design and implementation strategy of the system. Peer suggestions and expert comments contributed to refining the research objectives, selecting appropriate technologies, and improving the clarity of the research methodology. Such collaborative interaction plays an important role in strengthening the quality and reliability of the research work, as it allows researchers to validate their ideas and incorporate improvements suggested by experts in the field [13], [14], [15].

B. Use of Simulation software

Another approach used during the research process involved the use of simulation tools and development platforms to evaluate the proposed system. Simulation and software-based testing help researchers analyze the behavior and performance of a system before full deployment. Although many research studies use specialized simulation tools such as MATLAB to model algorithms and evaluate system performance, the VisionMate system was primarily developed and tested using mobile development frameworks and cloud-based artificial intelligence services. The system utilizes React Native for mobile application development, while Google Vision API and Speech-to-Text services are used to perform image analysis and voice recognition tasks.

Testing and experimentation were conducted by capturing images through the smartphone camera and processing them using cloud-based computer vision services. The system successfully detected objects, extracted text using OCR, recognized faces, and identified currency denominations. Additionally, the navigation module was tested using GPS location data and routing services to provide voice-guided directions. These experimental results demonstrate that integrating artificial intelligence services with smartphone sensors can effectively assist visually impaired users in understanding their surroundings

and performing everyday tasks more independently. The use of development platforms and testing tools allowed the researchers to validate the feasibility and performance of the proposed system and compile the results into a structured research paper suitable for peer review and publication [1], [2].

By adopting the above approaches—combining information from existing research, collaborating with experts, and testing the system using appropriate software platforms—the major components of the research paper were successfully developed. These practices enabled the researchers to transform theoretical concepts into a practical assistive application and present the findings in a systematic manner for academic evaluation and publication.

IV. GET PEER REVIEWED

Peer review is one of the most important stages in the research publication process. After completing the initial draft of the research paper titled “VisionMate: An AI-Based Mobile Assistive System for Visually Impaired Users,” the manuscript should be carefully evaluated by peers, mentors, and subject matter experts before submission to a journal. Peer review helps in identifying possible errors, improving the clarity of the research methodology, and strengthening the overall quality of the paper. Experts in the fields of artificial intelligence, computer vision, mobile application development, and assistive technologies can provide valuable feedback regarding the technical accuracy, experimental validation, and presentation of the research work.

During the peer review stage, the reviewers analyze different aspects of the paper such as the clarity of the research problem, relevance of the literature survey, effectiveness of the proposed system, and accuracy of the results and conclusions. In the case of the VisionMate system, reviewers may evaluate whether the integration of computer vision, speech recognition, and GPS navigation effectively addresses the challenges faced by visually impaired individuals. They may also review the system architecture, implementation approach, and performance analysis to ensure that the proposed

solution is technically sound and practically applicable.

Receiving constructive feedback from peers and experts helps the authors refine their work and improve the presentation of their research. Suggestions may include improving the explanation of the system methodology, adding additional experimental results, enhancing diagrams or system architecture illustrations, and correcting formatting issues according to the journal guidelines. Even if the researchers are confident about their work, obtaining multiple review comments from different experts is beneficial because it provides diverse perspectives and helps in identifying areas that require improvement.

Therefore, before submitting the VisionMate research paper to a journal, it is essential to ensure that the manuscript undergoes thorough peer evaluation. Incorporating the suggestions and recommendations provided by reviewers will enhance the scientific credibility, readability, and technical quality of the paper. This step significantly increases the chances of successful acceptance and publication in a reputed research journal.

V. IMPROVEMENT AS PER REVIEWER COMMENTS

Peer review is one of the most important stages in the research publication process. After completing the initial draft of the research paper titled “VisionMate: An AI-Based Mobile Assistive System for Visually Impaired Users,” the manuscript should be carefully evaluated by peers, mentors, and subject matter experts before submission to a journal. Peer review helps in identifying possible errors, improving the clarity of the research methodology, and strengthening the overall quality of the paper. Experts in the fields of artificial intelligence, computer vision, mobile application development, and assistive technologies can provide valuable feedback regarding the technical accuracy, experimental validation, and presentation of the research work.

During the peer review stage, the reviewers analyze different aspects of the paper such as the clarity of the research problem, relevance of the literature

survey, effectiveness of the proposed system, and accuracy of the results and conclusions. In the case of the VisionMate system, reviewers may evaluate whether the integration of computer vision, speech recognition, and GPS navigation effectively addresses the challenges faced by visually impaired individuals. They may also review the system architecture, implementation approach, and performance analysis to ensure that the proposed solution is technically sound and practically applicable.

Receiving constructive feedback from peers and experts helps the authors refine their work and improve the presentation of their research. Suggestions may include improving the explanation of the system methodology, adding additional experimental results, enhancing diagrams or system architecture illustrations, and correcting formatting issues according to the journal guidelines. Even if the researchers are confident about their work, obtaining multiple review comments from different experts is beneficial because it provides diverse perspectives and helps in identifying areas that require improvement.

Therefore, before submitting the VisionMate research paper to a journal, it is essential to ensure that the manuscript undergoes thorough peer evaluation. Incorporating the suggestions and recommendations provided by reviewers will enhance the scientific credibility, readability, and technical quality of the paper. This step significantly increases the chances of successful acceptance and publication in a reputed research journal.

VI. RESULT

The proposed VisionMate: An AI-Based Mobile Assistive System for Visually Impaired Users was implemented and tested using a smartphone-based environment to evaluate its effectiveness in assisting visually impaired individuals. The system integrates computer vision, speech recognition, and navigation technologies to provide real-time assistance through voice feedback. Various functional modules such as text recognition, object detection, face detection, currency identification, barcode scanning, and

navigation assistance were tested under different real-world conditions to analyze the performance and reliability of the system.

The text reading module using Optical Character Recognition (OCR) successfully extracted text from printed documents, product labels, and signboards captured using the smartphone camera. The extracted text was converted into speech using the text-to-speech engine, allowing users to listen to the content without visual interaction. The object detection module was able to identify common objects such as bottles, chairs, doors, and electronic devices in the surrounding environment. The system announced detected objects through voice output, helping visually impaired users understand their surroundings more effectively. The face detection module successfully detected the presence of people in the camera frame and provided audio feedback indicating the number of faces detected.

The currency recognition module demonstrated the ability to identify different denominations of currency notes based on text and visual patterns detected through the OCR system. Similarly, the barcode and QR code scanning module accurately detected product codes and provided the corresponding information through voice feedback. The navigation module used GPS and routing services to provide turn-by-turn walking directions. The system converted navigation instructions into spoken guidance, enabling users to move safely in outdoor environments. Experimental testing showed that the integration of artificial intelligence services with smartphone sensors allows the system to perform assistive tasks efficiently in real-time scenarios.

Overall, the results indicate that the VisionMate system provides a practical and effective assistive solution for visually impaired individuals. The system successfully combines multiple intelligent features within a single mobile application and provides real-time voice guidance to improve accessibility and independence for users. The testing results demonstrate that the proposed approach can enhance environmental awareness and mobility for visually impaired people, making it a promising

solution in the field of AI-based assistive technologies.

VII. CONCLUSION

This paper presented VisionMate, an AI-based mobile assistive system designed to help visually impaired individuals perform daily activities independently. The system integrates computer vision, speech recognition, and GPS navigation to provide features such as text reading using OCR, object detection, face detection, currency recognition, barcode scanning, and voice-guided navigation. By converting visual information into audio feedback, the application enables users to understand their surroundings and access information easily. The results show that the VisionMate system can effectively assist visually impaired users by providing real-time voice guidance and environmental awareness. The proposed solution demonstrates how smartphone technologies and artificial intelligence can be combined to create a practical and accessible assistive tool that improves mobility, safety, and independence for visually impaired individuals.

REFERENCES

- [1] D. Dakopoulos and N. G. Bourbakis, "Wearable obstacle avoidance electronic travel aids for blind: A survey," *IEEE Transactions on Systems, Man, and Cybernetics Part C*, vol. 40, no. 1, pp. 25–35, 2010.
- [2] M. Manduchi and J. Coughlan, "(Computer) Vision without sight," *Communications of the ACM*, vol. 55, no. 1, pp. 96–104, 2012.
- [3] J. M. Coughlan and H. Shen, "A smartphone-based system for assisting visually impaired users to navigate complex indoor environments," *IEEE International Conference on Systems, Man and Cybernetics*, 2013.
- [4] A. Ahmetovic, C. Gleason, K. Kitani, H. Takagi, and C. Asakawa, "NavCog: A navigation system for visually impaired users," *ACM SIGACCESS Conference on Computers and Accessibility*, 2016.

- [5] S. Mascetti, A. Ahmetovic, C. Bernareggi, and A. Gerino, "TypeInBraille: A braille-based typing interface for touchscreen devices," *ACM Transactions on Accessible Computing*, 2016.
- [6] J. Bigham et al., "VizWiz: Nearly real-time answers to visual questions," *Proceedings of the ACM Symposium on User Interface Software and Technology*, 2010.
- [7] A. Kane, J. Wobbrock, and R. Ladner, "Usable gestures for blind people: Understanding preference and performance," *ACM SIGCHI Conference on Human Factors in Computing Systems*, 2011.
- [8] S. Shoval, J. Borenstein, and Y. Koren, "Mobile robot obstacle avoidance in a computerized travel aid for the blind," *IEEE Transactions on Robotics and Automation*, vol. 14, no. 3, pp. 485–492, 1998.
- [9] J. M. Coughlan and A. Yuille, "The Manhattan world assumption: Regularities in scene statistics," *Neural Computation*, vol. 15, no. 4, pp. 833–854, 2003.
- [10] M. A. Goodrich and A. C. Schultz, "Human–robot interaction: A survey," *Foundations and Trends in Human–Computer Interaction*, vol. 1, no. 3, pp. 203–275, 2007.
- [11] R. Smith, "An overview of the Tesseract OCR engine," *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR)*, 2007.
- [12] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, 2012.
- [13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified real-time object detection," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [14] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.
- [15] A. Howard et al., "MobileNets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.