



Smart Assistant System for Blind Users Based on Vision-Language Models and Contextual Location Assistance

Clay Aiken Mangeber Jr, Angeli Marselin Kolmate, Semmy Wellem Taju*, Regi Fernando Najooan

Informatics, Faculty of Computer Science, Universitas Klabat, Minahasa Utara, Indonesia

Jl. Arnold Mononutu, Airmadidi Bawah, Kec. Airmadidi, Kabupaten Minahasa Utara, Sulawesi Utara, Indonesia

Email: ¹s22210410@student.unklab.ac.id, ²s22210542@student.unklab.ac.id, ³*semmy@unklab.ac.id,

⁴reginajoan@unklab.ac.id

Correspondence Author Email: semmy@unklab.ac.id

Received: 10/04/2026, Revised: 03/08/2026, Accepted: 16/10/2026, Published: 18/10/2026

Abstract—In this study, we designed and developed an AI-based Smart Assistant for visually impaired users at Panti Sosial GMIM Baertemeus. The proposed smart assistant system integrates a Vision-Language Model (VLM) to identify and describe visual objects, and Contextual Location Assistance (CLA) to provide directional and positional information based on geographic data. This research followed the Prototyping method which consists of some user needs that had been identified, such as system design, prototype development and testing with visually impaired users. The implementation results show that the proposed system can recognize objects and their surroundings and deliver contextual voice feedback adapted to the user's location. The CLA feature also allows administrators to monitor user positions in real time via a web dashboard. This system improves the independence and safety of visually impaired individuals in daily activities and contributes to the development of inclusive AI-based assistive technologies through the integration of VLM and CLA for accessibility enhancement.

Keywords: Smart Assistant; Vision-Language Model; Contextual Location Assistance; Visually Impaired; Artificial Intelligence

1. INTRODUCTION

In this current digital era, Artificial Intelligence (AI) has fast-developing into a transformative technology that significantly affect different aspect of human life, such as healthcare, education and smart assistant technologies [1]. Integration of AI into daily based systems has enabled machines to perform complex tasks including image recognition, natural language processing and even decision making [2]. Enhancing accessibility for people with disabilities, especially those who are visually impaired, is one of the most significant uses of AI [3]. Despite these developments, many visually impaired people still face significant obstacles in their day to day lives, such as trouble identifying items in their surroundings, obtaining visual information, and independently navigating new locations [4]. For mobility and everyday tasks, visually impaired people frequently depend on outside help or conventional equipment like white canes. These technologies are useful, but they don't give contextual awareness or much information about the surrounding environment [5]. Therefore, users may still be at danger when navigating complicated or dynamic surroundings. Additionally, the majority of digital programs and systems have visual interfaces, which makes them less usable by those with visual impairments. This drawback emphasises how urgently intelligent assistive solutions that can close the gap between visual information and user comprehension are needed. Improving the independence of visually challenged residents is still a top priority at the GMIM Bartemeus Social Home. Based on observations and conversations with residents and carers, it was discovered that while people are able to carry out simple daily tasks, they still have trouble identifying items and comprehending their environment. Additionally, they frequently need help from others to move around, particularly in new or unfamiliar places. These challenges show that the current support networks are insufficient to completely meet the needs of people who are blind or visually impaired. As a result, there is a pressing need to create an intelligent solution that can offer environmental awareness and real-time help.

In recent years, a number of assistive devices have been created to help those who are blind or visually impaired. AI-based solutions have been widely used to increase accessibility and usability, according to a survey of assistive technology [6]. Additionally, systematic reviews highlight how important AI is for improving assistive devices for individuals with visual impairments [7]. A number of research have suggested computer vision-based object detection systems that can identify things in real time [8]. Nevertheless, these technologies frequently only identify objects and do not give the user complete contextual information [9]. Navigation devices have been created to help visually impaired people go from one place to another in addition to object detection. For location tracking and route direction, these devices usually make use of GPS technology or spatial mapping [10]. Although helpful, navigation systems are less successful in dynamic contexts because they frequently do not integrate with real-time visual perception [11]. Even after receiving instructions, users may still find it difficult to comprehend nearby items or barriers. This restriction suggests that in order to offer a more complete assistive solution, navigation and visual comprehension should be combined [12].

As we found the new possibilities to improve assistive technology, the idea have been presented by using recent developments in generative AI. Vision-Language Models (VLM) allows a system to comprehend and describe visual content in natural language, are the result of the combination of Large Language Models (LLMs) with Computer Vision [13]. This feature enables systems to not only identify items but also give consumers clear,

understandable explanations [14]. Additionally, new research like WalkVLM shows how VLM might help visually impaired people navigate in real-world situations [15]. These advancements point to a major change in the direction of assistive technologies that are more interactive and intelligent.

A number of earlier research have tried to accommodate visually impaired individuals by combining various technology. Although AI-based applications have been created to offer voice-based communication and object identification, they frequently lack conversational and contextual awareness [16]. Other research has concentrated on navigation systems that do not incorporate visual perception but offer route direction [17]. Furthermore, wearable assistive technology has been suggested to integrate ambient awareness and object identification; nevertheless, its usability and interface with mobile platforms are currently restricted [18]. These drawbacks show that a more thorough and coordinated strategy is required. A research gap in the creation of assistive systems that combine contextual navigation, natural language interaction, and visual recognition into a single platform can be found based on the examination of earlier works. The majority of current solutions concentrate on particular features rather than offering visually impaired people a comprehensive solution [19]. Because of this, users still have to rely on a variety of tools to complete various activities, which might lower usability and efficiency. To enhance user experience and independence, an integrated system that incorporates these features is necessary.

The goal of this project is to create an AI-based smart assistant that combines Contextual Location Assistance (CLA) and Vision-Language Model (VLM) into a mobile application in order to close this gap. Real-time item identification and description, contextual navigation guiding, and voice-activated information delivery are all possible with the suggested system. It is anticipated that by integrating these technologies, the system will offer visually impaired people a more complete and intuitive solution. The ultimate goal of this project is to improve the quality of life, safety, and independence of visually impaired users, especially in the GMIM Bartemeus Social Home setting.

2. RESEARCH METHODOLOGY

2.1 Research Stages

The research stages for the development of the Smart Assistant system follows a structured four-phase approach: Problem Identification, Design, Development, and Testing. Figure 1 illustrates the workflow of this research design.



Figure 1. Research Stages

The process begins with Problem Identification, where the researchers focused on understanding the specific needs and daily challenges faced by visually impaired individuals [20]. Data was collected through direct interviews with residents and caregivers at the GMIM Bartemeus Social Home, as well as through a literature review, to define the essential features required for the Smart Assistant, such as location-based navigation and audio-based visual information [21]. The second stage is the Design phase, where the system architecture and user interface were conceptually constructed. This involved designing an interface specifically tailored for accessibility to ensure ease of use for visually impaired individuals [1]. The design focuses on developing on accessible solutions that integrates object recognition and location-based assistance into a unified system [9]. Following the design, the Development phase involved translating these concepts into functional application code using the Prototyping method [11]. This approach allows iterative development, enabling continuous improvements based on feedback obtained during implementation. During this stage, the system components are integrated, including the mobile interface and backend services [22]. The result of this stage is a working prototype of the Smart Assistant application that is ready for testing.

The final stage was Testing, where the developed application underwent rigorous evaluation to ensure all functionalities met the research objectives. Testing focused on validating critical features such as navigation accuracy, voice command responsiveness, and the clarity of audio information delivery [23]. Trials were

conducted with visually impaired users using Android devices, and the results were utilized to evaluate system reliability and identify areas for further refinement in the prototype.

2.2 Conceptual Framework

Figures 2 illustrates the conceptual frameworks of the Smart Assistant system designed to support visually impaired users. The workflow begins with the User providing input through three primary channels: voice command, real-time camera visuals, and GPS location data. These inputs can be in the form of spoken questions or command, the states of the surrounding environment, or location information required to provide context-based assistant. These inputs are processed by three integrated components, including (1) Speech-to-Text (STT) can be used to convert the user's verbal commands into text format using the device's native capabilities. (2) Vision-Language-Model (VLM) is used by utilizing Gemini 2.5 flash model to analyze images captured using the camera, detecting objects or extracting text, and (3) Contextual Location Assistant (CLA) can be used to process GPS coordinates using OpenCage and Overpass APIs to identify the user's location and nearby places. The processed data from these components is aggregated and sent to Generative AI (Gemini). Gemini applies Natural Language Understanding (NLU) to comprehend the user's intent and context, then uses Natural Language Generation (NLG) to formulate an intelligent, contextual response. This text response is converted back into audio via Text-to-Speech (TTS) and delivered to the user.

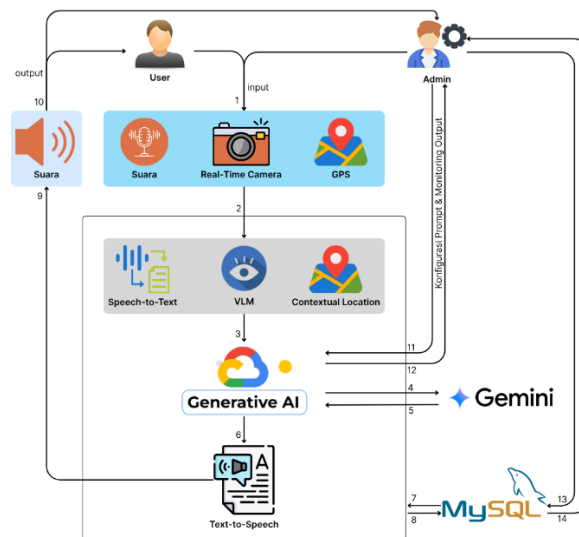


Figure 2. Conceptual Framework

2.3 Prototyping

As shown in Figure 3, the prototyping model has been adopted in our system development. This approach was specifically selected to facilitate active user participation throughout the entire development lifecycle [24]. By adopting this method, the researchers ensured that iterative feedback could be continuously integrated, guaranteeing that the final product effectively aligns with the specific needs and limitations of the visually impaired users [6]. The process began with the Communication phase, which focused on a comprehensive analysis of user requirements. Data collection was conducted through extensive literature reviews on Vision-Language Models (VLM) and Contextual Location Assistance (CLA). Additionally, direct interviews were held with visually impaired individuals and caregivers at the GMIM Bartemeus Social Home to identify specific daily challenges, resulting in a definitive list of user needs and an initial feature set for the system [6].

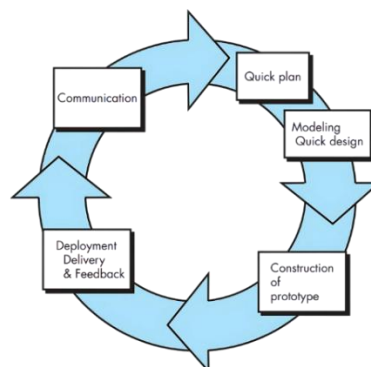


Figure 3. Prototyping Model

Subsequently, the Quick Plan phase was initiated to formulate a rapid development strategy. Based on the requirements gathered in the previous stage, the researchers defined the system's scope and prioritized key features for the initial development. This phase established a strategic roadmap for quick iteration, ensuring that the most critical functionalities were addressed first. This planning guided the Modeling Quick Design phase, where the initial system design was created based on the identified user needs. This involved designing an accessibility-friendly user interface (UI) using Figma to visualize the interaction flow. Furthermore, the system architecture was structured for the Android platform using React Native, producing a functional visual representation ready for immediate translation into a prototype [25]. The design concepts were then implemented into code during the Construction of Prototype phase. This stage involved building the core features, such as user input handling, Speech-to-Text integration, camera access for object recognition via VLM (Gemini), and audio output using Text-to-Speech. The development utilized React Native for the front-end, Python for back-end logic, and Node.js with MySQL for database integration to ensure robust performance [26].

The process concluded with the Deployment, Delivery and Feedback phase, where the developed prototype underwent iterative testing. Internal testing verified component functionality, specifically focusing on voice command responsiveness and location data consistency. Following this, user trials were conducted to gather feedback on usability and technical stability, which served as the primary basis for further system refinement and improvements in subsequent iterations [23].

3. RESULT AND DISCUSSION

3.1 System Overview

Figure 4 illustrates the use case diagram for the Smart Assistant System designed for visually impaired users. This diagram helps to visualize the interactions between different users (actors) and the system, based on the specific roles they have within the application. The primary actors involved are the visually impaired user (Pengguna) and Admin. The Pengguna refers to the visually impaired user who interacts with the system by giving voice commands such as asking about the location, requesting to read visible text, and receiving responses from the system. These tasks are broken down into different use cases such as "Menanyakan Lokasi Saat Ini" (asking the current location), "Membaca Teks yang Terlihat" (reading visible text), and "Mendeteksi Objek dan Menyampaikan Lokasi Objek" (detecting objects and communicating their location). The Admin is responsible for overseeing system management, including features like Registrasi (registration), Edit/Hapus (edit or delete users), and managing user data through features such as Lihat Data Pengguna (view user data). The Admin also manages the system configuration through options like Konfigurasi Prompt (configuring prompts) and can view system activity through History.

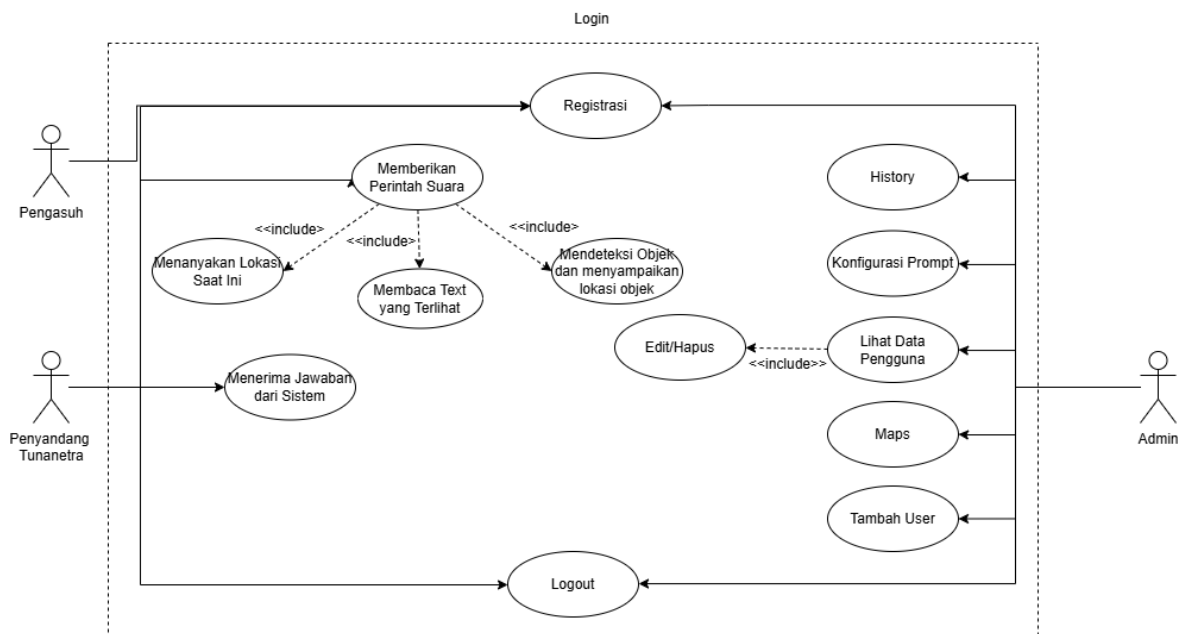


Figure 4. Use Case Diagram

The Pengasuh (Caregiver) plays a role in assisting the visually impaired user, specifically by interacting with the system to get answers on the user's behalf, such as location inquiries or object recognition. The diagram clearly outlines how the system operates by separating roles and responsibilities, making it easier to understand who interacts with what features and how each actor contributes to the overall functionality of the system. It

demonstrates the interaction flow from registration to voice commands, giving each actor distinct duties to ensure the system is functional and user-friendly.

3.2 Application Implementation

This section presents the implementation of the Smart Assistant system, focusing on the main functionalities developed to support visually impaired users. The admin system and the user application are the two primary parts of the system. Accessibility, real-time communication, and the incorporation of artificial intelligence technologies, such as Contextual Location Assistance (CLA) and Vision-Language Models (VLM), are prioritised in the implementation.

3.2.1 User

Figure 5 illustrates the implementation of the user signup interface within the Smart Assistant application. The registration process is structured into five sequential screens to ensure simplicity. The first screen serves as an introductory page where users can choose to proceed with registration by tapping the "Lanjut" (Next) button or navigate to the login page if they already possess an account. The subsequent screens guide the user step-by-step to input their required credentials, specifically First Name, Last Name, Username, and Password.



Figure 5. Users Signup

Figure 6 displays the implementation of the login interface for the Smart Assistant application. The login flow consists of three distinct screens. The first screen is the introductory page, where users can choose to proceed with registration by tapping "Lanjut" (Next) or select the option to login if they already have an account. Upon selecting login, the user is directed to the username input screen. After entering the username, the system proceeds to the password entry page. Finally, once both the username and password are validated, the user is authenticated and redirected to the main page.



Figure 6. Users Login

Figure 7 presents the implementation of the main interface for the Smart Assistant application tailored for visually impaired users. Upon entering this page, the application welcomes the user with a personalized greeting following the format "{greeting}, {name}. I am ready to help you today." (e.g., "Good Evening, User. I am ready to help you today."). This greeting is synchronized with the device's time to establish a natural interaction. The design prioritizes simplicity and high contrast to facilitate easy identification of key visual elements. The center

of the screen features a real-time camera viewfinder used to capture the user's surroundings for contextual analysis via text-to-speech. At the bottom, a microphone icon indicates the status of the voice command function. Additionally, a menu button is located in the top-right corner, serving as a toggle to reveal the logout option, allowing users to sign out of their accounts securely.



Figure 7. Main Page

3.2.2 Admin

Figure 8 presents the implementation of the Admin Dashboard interface, serving as the central control hub for administrators to monitor the system. This page displays a summary of essential information, including the total number of users and AI interaction history. It also features a navigation menu providing access to other management modules such as User, History, and Prompt. The interface design prioritizes simplicity and informativeness, ensuring that administrators can easily comprehend the overall system status and perform necessary actions efficiently.

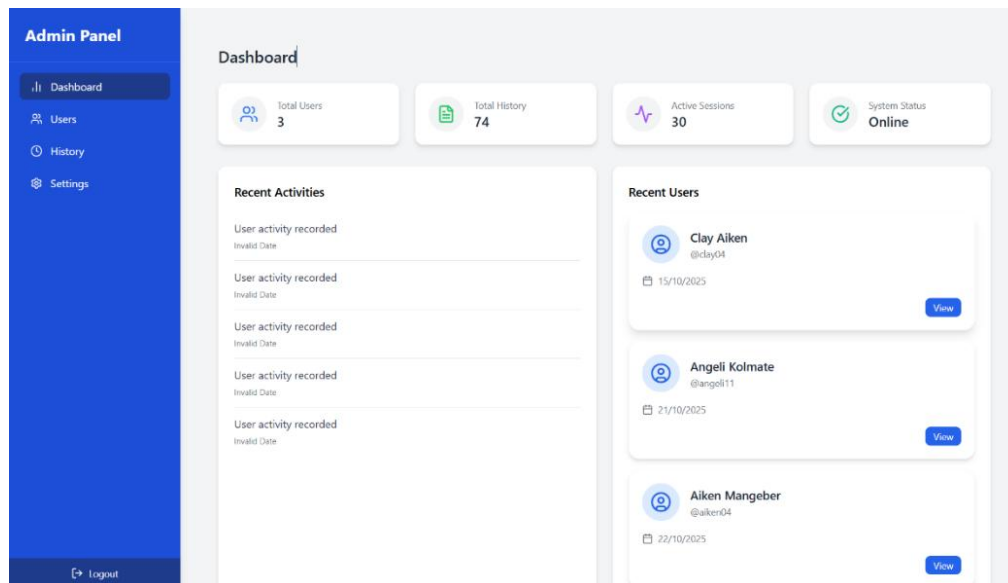


Figure 8. Admin Dashboard

Figure 9 presents the implementation of the History interface within the system. This page serves to display a comprehensive log of interactions between users and the AI system. Administrators can view specific details such as the commands or inputs provided by users, along with the corresponding timestamps of these interactions. The interface is designed with a structured list format, enabling administrators to browse and track the interaction history quickly and efficiently.

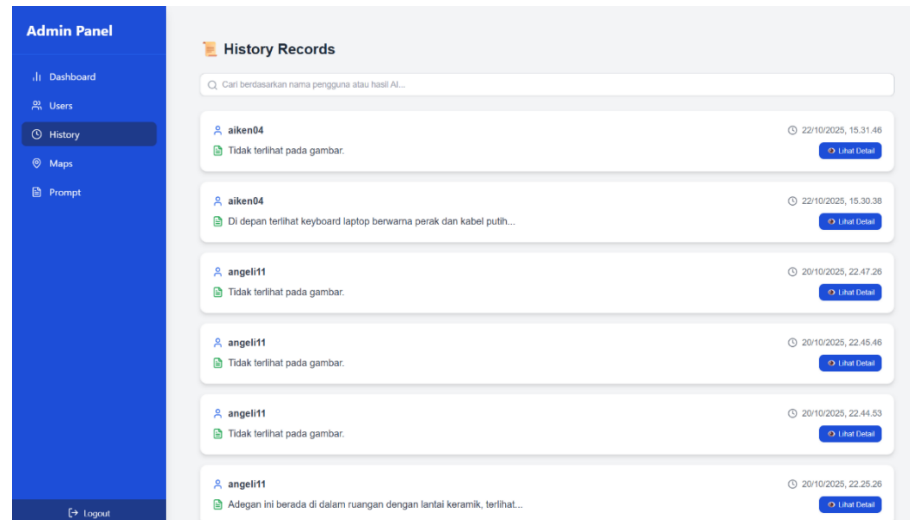


Figure 9. History Records of the Proposed Smart Assistant

Figure 10 displays the implementation of the History Details interface. This page provides a comprehensive breakdown of a specific interaction between a user and the AI system. Administrators can review both the user's inputs, including text queries and captured images and the corresponding response generated by the Gemini model. The History Details functionality supports the evaluation of system performance by enabling administrators to analyze the accuracy and relevance of AI-generated outputs. This contributes to improving the overall quality and reliability of the Smart Assistant.

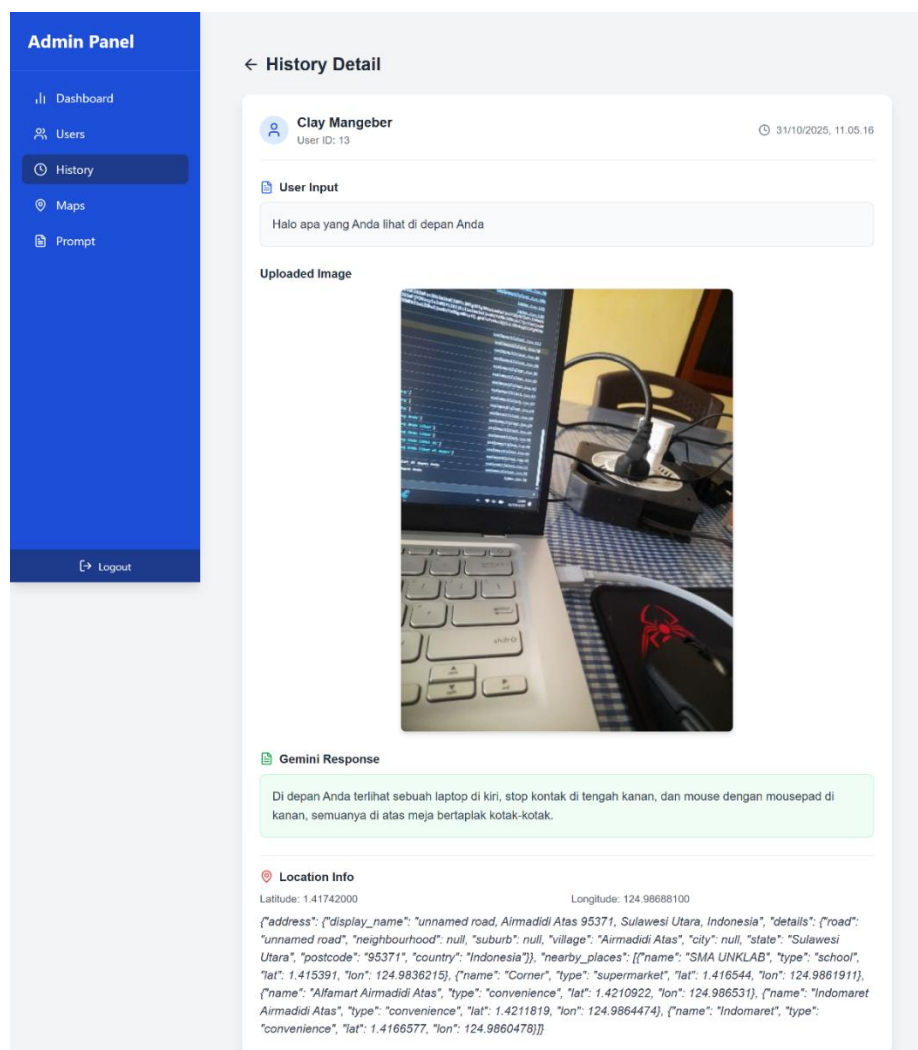


Figure 10. History Details

Figure 11 illustrates the implementation of the Maps interface within the admin panel of the Smart Assistant system. This page serves to display real-time user locations based on GPS coordinate data transmitted from their devices. The longitude and latitude information is visualized as markers on a digital map, allowing administrators to monitor user positions directly for safety and tracking purposes.

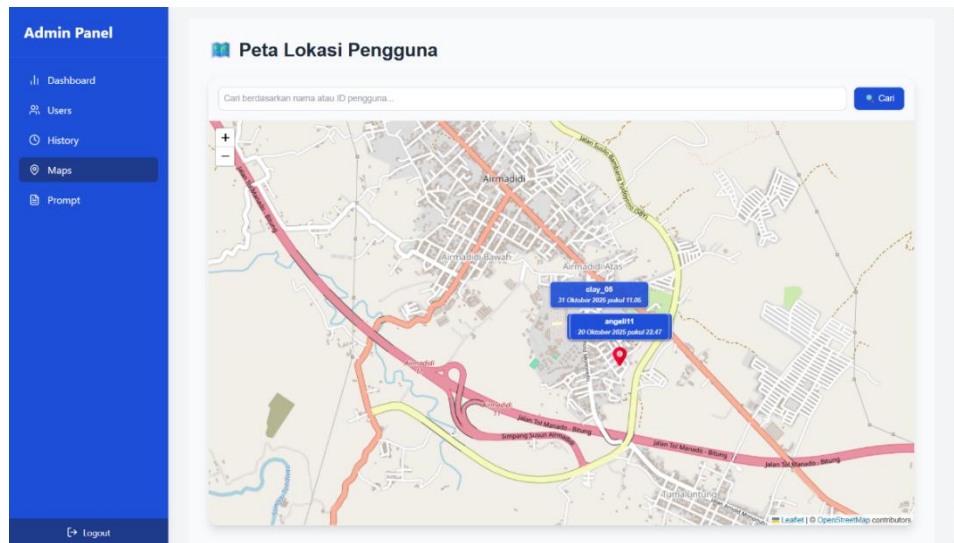


Figure 11. Maps

Figure 12 shows the Prompt configuration feature, which allows administrators to manage AI instructions to improve the relevance and accuracy of system responses.

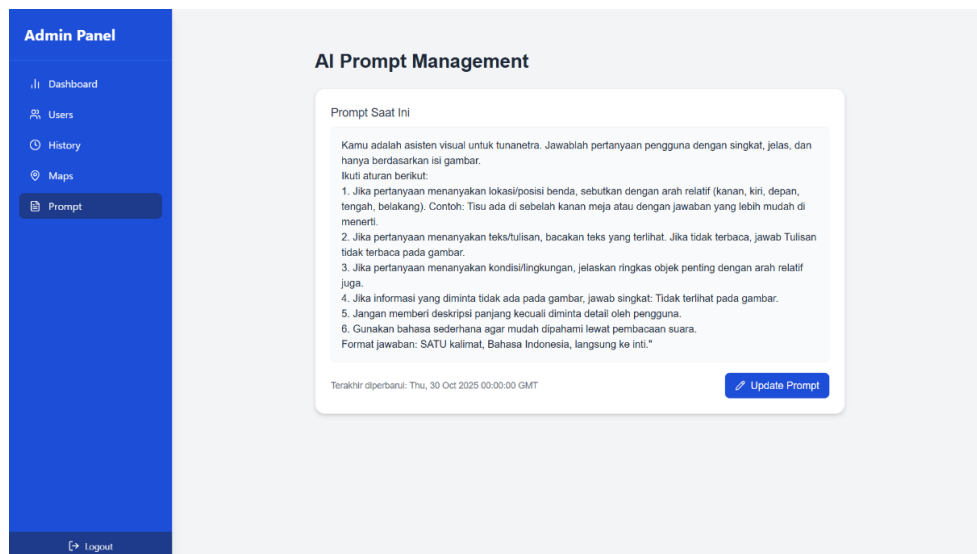


Figure 12. Prompt Engineering of the VLM

3.3 Discussion

To evaluate the performance of the Smart Assistant system, a series of functional tests were conducted involving users, administrators, and the AI component. The testing results are summarized in Table 1.

Table 1. System Testing Results

Actor	Input	Process	Expected Output	Status
Users	Registration	1. Enter the intro page, 2. Click the next button, 3. Input First Name, Last Name, Username, and Password, 4. Click the Finish button to complete registration.	Account creation. Data is saved in the database.	Successful

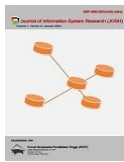
Actor	Input	Process	Expected Output	Status
	Login	1. Go to the login page by clicking the next button if an account already exists, 2. Input username and password, 3. Click the Login button.	Successfully logged into the app and the Smart Assistant main page is displayed.	Successful
	Logout	Click the Logout button.	Successfully logged out from the app, and the login page is displayed again.	Successful
	Voice Interaction (Asking for Object)	The user says the command, "What is in front of me?"	The system provides a voice output of the object analysis result.	Successful
	Voice Interaction (Asking for Location)	The user asks, "Where am I?"	The system provides a voice output with the user's location.	Successful
	Voice Interaction (Reading Text)	The user requests the system to read visible text.	The system recognizes the text from the camera and reads it aloud.	Successful
Admin	Login	Admin enters a valid username and password.	The system displays the admin dashboard page.	Successful
	View User Data	Admin selects the "View User Data" menu.	The system displays the list of all registered users.	Successful
	View History	Admin selects the "View History" menu.	The system displays the usage history of the Smart Assistant.	Successful
	Live Location Monitoring	Admin opens the dashboard and views the user's position on the map.	The user's position is displayed on the admin dashboard.	Successful
	Edit Prompt	Admin clicks the "Edit Prompt" button and saves a new prompt.	The new prompt is saved and activated.	Successful
	Logout	Click the Logout button.	The system redirects to the login page.	Successful
AI (Gemini)	Asking About Object in Front	The system receives voice input and an image from the user.	Gemini analyses and generates text and voice output.	Successful

Based on the testing results shown in Table 1, all functionalities of the proposed system were successfully executed according to the expected outcomes. This indicates that the proposed smart Assistant system is functionally stable and capable of supporting both user and administrator activities. From the user perspective, voice interaction features play a key role in improving accessibility. Users can interact with the system using natural voice commands, allowing them to obtain information about objects, text, and their surroundings without relying on visual input. This significantly enhances usability for visually impaired individuals.

The integration of the Vision-Language Model enables the system to analyze images and generate contextual responses. In addition, the location-based feature provides real-time information about the user's position, supporting spatial awareness and navigation. From the administrative side, features such as user data management, history tracking, and live location monitoring help ensure system reliability and performance evaluation. The prompt configuration feature also allows administrators to control and improve AI responses dynamically. However, several limitations were identified. The accuracy of object detection may be affected by environmental conditions such as lighting and object complexity. Additionally, the system depends on internet connectivity, which can impact performance in certain situations. Overall, the Smart Assistant system demonstrates strong potential as assistive technology by providing real-time, voice-based, and context-aware information for visually impaired users.

4. CONCLUSION

In this research, we successfully developed a Smart Assistant System designed to assist visually impaired users in recognizing objects and determining their position through voice interaction and the Android device's camera. The system integrates various technologies, such as Vision Language Model, Speech-to-Text, Text-to-Speech, and Contextual Location Assistance, developed using Flask API, Gemini AI, and MySQL. Based on the implementation and testing conducted, the system successfully recognizes objects around the user and provides



object descriptions in the form of speech, communicated in a simple manner for easy understanding by the visually impaired. In addition, the system can provide information about the relative direction of objects (right, left, or front) based on the image analysis from the camera. The live location tracking feature allows the admin to monitor the user's position accurately via coordinates (longitude and latitude) displayed on the web dashboard, which works effectively as long as there is a stable internet connection. However, the system still has limitations in analyzing more complex environments, such as recognizing distances between objects or small obstacles, due to the limitations of the camera and AI processing power on mobile devices. Overall, all the core functions of the system, such as login, object recognition, speech output, and location tracking, performed as designed, addressing the research problem and providing a functional solution for visually impaired users to recognize objects and positions around them, as well as facilitating admin monitoring of user locations through the web dashboard.

5. ACKNOWLEDGMENT

Although the system has been successfully implemented, there are still areas that can be further developed. One of these is the optimization of the location feature. Currently, the system only displays coordinates (longitude-latitude) and the direction of objects, so it is recommended to add a simple navigation feature that can provide step-by-step voice instructions to guide users toward specific locations or safe spots. This development is expected to make navigation easier for users, especially in more complex environments. Additionally, the system could be optimized to allow some AI processes to run directly on the device (on-device processing), which would reduce dependence on internet connectivity. This is crucial to ensure that the system continues to function well in areas with unstable networks. Lastly, it is recommended to improve the accessibility of the voice feature by adding settings for speech speed and voice tone, so users can adjust the speech settings according to their personal preferences and hearing capabilities. With these developments, the system will become more accessible and user-friendly for visually impaired users.

6. DECLARATION

6.1 Author Contributions

Conceptualization: Clay Aiken Mangeber Jr, Angeli Marselin Kolmate, Semmy Wellem Taju, and Regi Fernando Najoan;

Methodology: Clay Aiken Mangeber Jr and Angeli Marselin Kolmate;

Software: Clay Aiken Mangeber Jr;

Validation: Clay Aiken Mangeber Jr, Semmy Wellem Taju, and Regi Fernando Najoan;

Formal Analysis: Clay Aiken Mangeber Jr and Angeli Marselin Kolmate;

Investigation: Clay Aiken Mangeber Jr and Angeli Marselin Kolmate;

Resources: Semmy Wellem Taju and Regi Fernando Najoan;

Data Curation: Clay Aiken Mangeber Jr;

Writing – Original Draft Preparation: Clay Aiken Mangeber Jr and Angeli Marselin Kolmate;

Writing – Review and Editing: Semmy Wellem Taju and Regi Fernando Najoan;

Visualization: Clay Aiken Mangeber Jr.

All authors have read and agreed to the published version of the manuscript.

6.2 Data Availability

The data presented in this study are available on request from the corresponding author.

6.3 Funding

This research received no external funding. The publication and manuscript processing fees associated with the preparation and publication of this scientific manuscript were self-funded by the authors.

6.4 Institutional Review Board Statement

Not applicable.

6.5 Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

6.6 Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

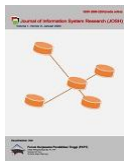


6.7 Generative AI and AI-assisted Technologies in The Writing Process

The authors used Generative AI to improve the writing clarity of this paper. They reviewed and edited the AI-assisted content and took full responsibility for the final publication.

REFERENCES

- [1] M. A. Shera et al., "Usability Evaluation of Blind and Visually Impaired Interface in Solving the Accessibility Problems," 4th International Conference on Innovative Computing, ICIC 2021, no. November, 2021, doi: 10.1109/ICIC53490.2021.9693084.
- [2] C. Zhu, "Artificial Intelligence Applications in Everyday Life," in Proceedings of the 5th International Conference on Signal Processing and Machine Learning, 2025, pp. 191–197. doi: 10.54254/2755-2721/121/2025.20074.
- [3] K. Chemnad and A. Othman, "Digital accessibility in the era of artificial intelligence—Bibliometric analysis and systematic review," *Frontiers in Artificial Intelligence*, vol. 7, p. 1349668, 2024, doi: 10.3389/frai.2024.1349668.
- [4] R. R. A. Bourne et al., "Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: The Right to Sight: An analysis for the Global Burden of Disease Study," *Lancet Glob. Health*, vol. 9, no. 2, pp. e144–e160, 2021, doi: 10.1016/S2214-109X(20)30489-7.
- [5] R. Susanto, W. Atmadja, R. Eko Prabowo, N. Julian, and U. Bina Nusantara, "Asisten Navigasi untuk Orang dengan Gangguan Penglihatan Menggunakan Ultra Sonic (Sonar Vision) Untuk Yayasan Mitra Netra," *SEEIJ (Social Economics and Ecology International Journal)*, vol. 4, no. 1, pp. 9–13, 2020.
- [6] A. Lavric, C. Beguni, E. Zadobrischi, A. M. Căilean, and S. A. Avătămăniței, "A Comprehensive Survey on Emerging Assistive Technologies for Visually Impaired Persons: Lighting the Path with Visible Light Communications and Artificial Intelligence Innovations," *Sensors*, vol. 24, no. 15, 2024, doi: 10.3390/s24154834.
- [7] L. Triyono, R. Gernowo, and S. R. Cholil, "A Systematic Review of Artificial Intelligence in Assistive Technology for People with Visual Impairment," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, vol. 8, no. 4, 2023, doi: 10.22219/kinetik.v8i4.1772.
- [8] J. Chai, H. Zeng, A. Li, and E. W. T. Ngai, "Deep learning in computer vision: A critical review of emerging techniques and application scenarios," *Machine Learning with Applications*, vol. 6, no. August, p. 100134, 2021, doi: 10.1016/j.mlwa.2021.100134.
- [9] S. Alhazmi, M. Kutbi, S. Alhelaly, U. Dawood, R. Felemban, and E. Alaslani, "Utilizing Artificial Intelligence Techniques for Assisting Visually Impaired People: A Personal AI-based Assistive Application," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 8, pp. 813–820, 2022, doi: 10.14569/IJACSA.2022.0130894.
- [10] A. McGibney, R. Pospisil, K. O'Mahony, J. F. Martinez, and S. Rea, "Spatial Mapping for Visually Impaired and Blind Using BLE Beacons," in Proceedings of the 6th Collaborative European Research Conference (CERC 2020), 2020, pp. 15–29. [Online]. Available: <https://www.ariadnegps.eu/>
- [11] N. Tyagi, D. Sharma, J. Singh, B. Sharma, and S. Narang, "Assistive Navigation System for Visually Impaired and Blind People: A Review," *Proceedings - 2021 1st IEEE International Conference on Artificial Intelligence and Machine Vision, AIMV 2021*, 2021, doi: 10.1109/AIMV53313.2021.9670951.
- [12] Z. Merchant, A. Anwar, E. Wang, S. Chattopadhyay, and J. Thomason, "Generating Contextually-Relevant Navigation Instructions for Blind and Low Vision People," *arXiv preprint arXiv:2407.08219*, Jul. 202, doi: 10.48550/arXiv.2407.08219.
- [13] B. Fu, A. Hadid, and N. Damer, "Generative AI in the context of assistive technologies: Trends, limitations and future directions," *Image Vis. Comput.*, vol. 154, Feb. 2025, doi: 10.1016/j.imavis.2024.105347.
- [14] R. Adnin and M. Das, "'I look at it as the king of knowledge': How Blind People Use and Understand Generative AI Tools," in The 26th International ACM SIGACCESS Conference on Computers and Accessibility, New York, NY, USA: ACM, Oct. 2024, pp. 1–14. doi: 10.1145/3663548.3675631.
- [15] Yuan, Zhiqiang, Ting Zhang, Yeshuang Zhu, Jiawei Zhang, Ying Deng, Zexi Jia, Peixiang Luo, Xiaoyue Duan, Jie Zhou, and Jinchao Zhang. "Walkvln: Aid visually impaired people walking by vision language model." In 2025 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 9845-9854. IEEE, 2025, doi: 10.1109/ICCV51701.2025.00918.
- [16] M. S. A. Baig, S. A. Gillani, S. M. Shah, M. Aljawarneh, A. A. Khan, and M. H. Siddiqui, "AI-based Wearable Vision Assistance System for the Visually Impaired: Integrating Real-Time Object Recognition and Contextual Understanding Using Large Vision-Language Models," *arXiv preprint arXiv:2412.20059*, 2024, doi: 10.48550/arXiv.2412.20059.
- [17] D. Plikynas, A. Indriulionis, A. Laukaitis, and L. Sakalauskas, "Indoor-Guided Navigation for People Who Are Blind: Crowdsourcing for Route Mapping and Assistance," *Applied Sciences (Switzerland)*, vol. 12, no. 1, Jan. 2022, doi: 10.3390/app12010523.
- [18] Z. Chen, Z. Liu, K. Wang, K. Wang, and S. Lian, "A Large Vision-Language Model based Environment Perception System for Visually Impaired People," in 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2024, doi: 10.1109/IROS58592.2024.10801813.
- [19] R. Nishant, M. Kennedy, and J. Corbett, "Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda," *Int. J. Inf. Manage.*, vol. 53, p. 102104, Aug. 2020, doi: 10.1016/J.IJINFOMGT.2020.102104.
- [20] D. Su and P. Fung, "Improving Spoken Question Answering Using Contextualized Word Representation," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2020-May, pp. 8004–8008, May 2020, doi: 10.1109/ICASSP40776.2020.9053979.
- [21] L. Manirajee, S. Q. Hanis Shariff, and S. M. Mohd Rashid, "Assistive Technology for Visually Impaired Individuals: A Systematic Literature Review (SLR)," *International Journal of Academic Research in Business and Social Sciences*, vol. 14, no. 2, 2024, doi: 10.6007/ijarbss/v14-i2/20827.



- [22] F. Atoillah, A. R. Gonzales, and J. Amirulloh, “Analisis Penggunaan Perangkat Android Dengan Perbandingan Parameter Benjamin Sparkman Fakultas Teknologi Informasi Yadika,” *Jurnal Sistem Informasi Aplikasi Teknologi Informasi*, vol. 1, no. 2, pp. 141–153, 2024.
- [23] Z. Huang and M. Benyoucef, “A systematic literature review of mobile application usability: addressing the design perspective,” *Univers. Access Inf. Soc.*, vol. 22, no. 3, pp. 715–735, Aug. 2023, doi: 10.1007/S10209-022-00903-W.
- [24] A. Obaid, “Using Prototypes in Agile Software Development,” *International Journal of Computers and Informatics*, vol. 3, no. 2, pp. 23–38, Feb. 2024, doi: 10.59992/ijci.2024.v3n2p2.
- [25] K. Kurniawan and Y. Yulhendri, “Pemanfaatan Framework React Native dalam Perancangan Aplikasi Penjualan Merchandise,” *Nucleus*, vol. 4, no. 2, pp. 84–97, 2023, doi: 10.37010/nuc.v4i2.1369.
- [26] S. Abdelhadi, K. Elbahnasy, and M. Abdelsalam, “A PROPOSED MODEL TO PREDICT AUTO INSURANCE CLAIMS USING MACHINE LEARNING TECHNIQUES,” *J. Theor. Appl. Inf. Technol.*, vol. 98, p. 22, 2020, [Online]. Available: www.jatit.org