

IOT - Based Smart Guide for Blind People with Obstacle & Object Detection

Ravi D. Pawar¹, Sachin D. Jadhav², Prasad R. Khode³, Piyush S. Patade⁴, Mrs. V. N. Waghchoure⁵

¹ComputerTechnology Dept., K. K. Wagh Polytechnic, Nashik

²ComputerTechnology Dept., K. K. Wagh Polytechnic, Nashik

³ComputerTechnology Dept., K. K. Wagh Polytechnic, Nashik

⁴ComputerTechnology Dept., K. K. Wagh Polytechnic, Nashik

⁵ComputerTechnology Dept., K. K. Wagh Polytechnic, Nashik

Abstract - Visually impaired people have very limited environmental awareness and detect hazards with a significant delay, which seriously challenges their independent navigation. Traditional mobility aids, such as white canes, provide ground-level support only and cannot detect overhead or far-ahead obstacles. This paper proposes an AI and IoT-based Smart Guide system for enhancing mobility and safety by offering real-time object detection and audio feedback to its visually impaired users.

This system integrates the camera module, Raspberry Pi, object detection model of YOLOv5, GPS module, and a text-to-speech engine to name the surrounding objects that it detects and sends voice information via headphones to the user. Transcribing such detected objects into voice alerts will help users understand their environment. Location data is simultaneously transmitted to a Flask-based web application, enabling caregivers to keep track of the user while receiving emergency notifications, if required.

Stress is put on the low cost, portability, and modular design of the system; therefore, it is very practical for deployment in the real world. The tests in the lab proved that the system was capable of detecting objects correctly and the user was getting the voice prompt in time under different conditions. The Smart Guide is intended to increase the independence of visually impaired people, lower the possibility of accidents, and generally raise the quality of their life by merging AI, computer vision, and IoT technologies into one assistive platform.

Key Words: Artificial Intelligence, Internet of Things, Object Detection, YOLOv5, Assistive Technology, GPS Tracking, Audio Feedback.

1. INTRODUCTION

This visual impairment has a major influence on an individual's ability to move safely and independently in public places. According to reports from the world health scene, more than 2.2 billion people worldwide suffer from some visual impairment. Of this figure, one billion cases are preventable or to be addressed. To these individuals, even the simplest of tasks such as walking in public places while spotting potential obstacles and spotting other objects becomes a challenge.

Conventional assistive tools like white canes are commonly employed because of their simplicity and affordability, but they can only help individuals traverse around objects at ground level up to a certain extent, not beyond that, and are of little use against objects or barricades situated at a greater altitude like signboards, branches, or other moving objects. The latest technologies like guide dogs or wearable assistive technology are still not accessible to a number of visually impaired individuals because of their costly and highly technical nature.

Recent advances in the field of Artificial Intelligence (AI), Computer Vision, and the Internet of Things have created possibilities for the development of intelligent assistive systems. AI-based object recognition models, including YOLO (You Only Look Once) for object detection in real-time, and IoT for location tracking and monitoring, can be used to provide the visually impaired with contextual awareness using audio-based systems.

In this paper, an AI and IoT-based Smart Guide System is proposed for the visually impaired, which serves the purpose of detecting objects in real time, conveying audio messages, GPS tracking, and caregiver tracking using a web application. For this purpose, a

camera module is used to detect objects in the surroundings using the YOLOv5 model, which are then conveyed to the user in the form of text-to-speech output. Meanwhile, the location is also sent to the caregiver interface in real time to assist in remote caregiving. Basically, the main aim of this paper is to create a cost-effective, portable, and user-friendly Smart Guide System to help the visually impaired become more aware of the surroundings and live a better life.

2. LITERATURE SURVEY

The paper *“Echo Light: A Camera-Powered Voice-Assisted Aid for the Visually Impaired”* by Saraswathy S., Charu Nethra G., and Pradeepa Gandhi S. presents a vision-based assistive system that uses a camera along with deep learning techniques to detect surrounding objects and recognize printed text using OCR, which is then converted into audio output to help visually impaired users understand their environment. While the system effectively improves environmental awareness, it mainly focuses on real-time detection and lacks support for recorded video access, location history, and emergency assistance features, which limits its usefulness during critical situations. Our proposed system overcomes these limitations by integrating IoT support for recorded videos and location data along with an emergency support mechanism, making the solution more practical and safer for daily use. [1]

The research paper *“AI-Powered Smart Cane for Visually Impaired Navigation”* by V. Remya, T. Sibiraj, and S. Subramani proposes a smart navigation aid that combines ultrasonic sensors, a camera module, and IoT connectivity to detect obstacles and provide voice-based navigation assistance. Although the system enhances mobility and safety, it heavily depends on multiple sensors, which increases hardware complexity, cost, and maintenance issues, especially in dynamic or crowded environments. In contrast, our proposed solution eliminates the dependency on multiple sensors and relies solely on camera-based AI object detection, reducing system complexity while maintaining effective navigation support. [2]

The paper *“Vision-Based Assistive Navigation System Using Deep Learning”* by Bai J., Lian S., and Liu Z. introduces a deep learning-based approach that uses camera input to identify objects and assist visually impaired individuals, particularly in indoor environments. While the system demonstrates good accuracy in controlled settings, it does not support outdoor navigation,

recorded data storage, or emergency handling, limiting its real-world applicability. Our system addresses these shortcomings by supporting both indoor and outdoor environments and providing access to recorded videos and location history for later review or assistance. [3]

The study *“Smart Glasses for Visually Impaired Using CNN and OCR”* by Simra Nazim et al. proposes wearable smart glasses that utilize convolutional neural networks for object detection and OCR for text recognition, delivering information through audio feedback. Despite its technological advancement, the requirement of specialized smart glasses increases the overall cost and may cause comfort issues for long-term use, while also lacking features such as recorded navigation data and emergency support. Our proposed system avoids the need for specialized wearable hardware and instead offers a lightweight, cost-effective wearable solution supported by a mobile application with recorded data access. [4]

The research work *“IoT-Based Assistive System for Visually Impaired People”* by Karmel A., Sharma A., and Pandya M. presents an IoT-enabled assistive device that combines cameras, ultrasonic sensors, and cloud services to improve navigation and obstacle detection. Although the system enhances connectivity and monitoring, its reliance on multiple sensors and continuous cloud processing can lead to higher power consumption and latency issues. The proposed project improves upon this by using offline camera-based AI processing and limited IoT usage only for recorded video access and location sharing, ensuring faster response and better reliability. [5]

3. PROPOSED SOLUTION

The idea of AI and IoT-based Smart Guide concept for the visually impaired will be able to deliver real-time object detection and audio cues for the user, along with the ability for caregivers to remotely monitor the user with the help of computer vision and IoT communications.

The main components of the system include a camera module, a Raspberry Pi 4 Model B, the YOLOv5 object detection model, a GPS module, a text-to-speech engine, and a Flask-based web application. The camera module runs continuously to capture the live video feed of the environment, which is then processed by the object detection algorithm. The algorithm identifies people, vehicles, doors, and other obstacles in the environment.

Once an object has been identified, the identification of the object is turned into audio by the text-to-speech engine and then played for the user through headphones.

At the same time, real-time location coordinate data is retrieved by the GPS module and sent to the web-based caregiver interface using the Flask server. This helps the caregiver to track the location of the user and also be aware of emergencies. The Raspberry Pi serves as the central processing unit for all operations.

The design is developed with modular architecture, whereby individual components of the system, such as object detection, audio feedback, and location tracking, are integrated but independent. Consequently, this makes the design scalable and enhances its potential for upgrades. The proposed Smart Guide is affordable, portable, and easy to use. It can be implemented in the field.

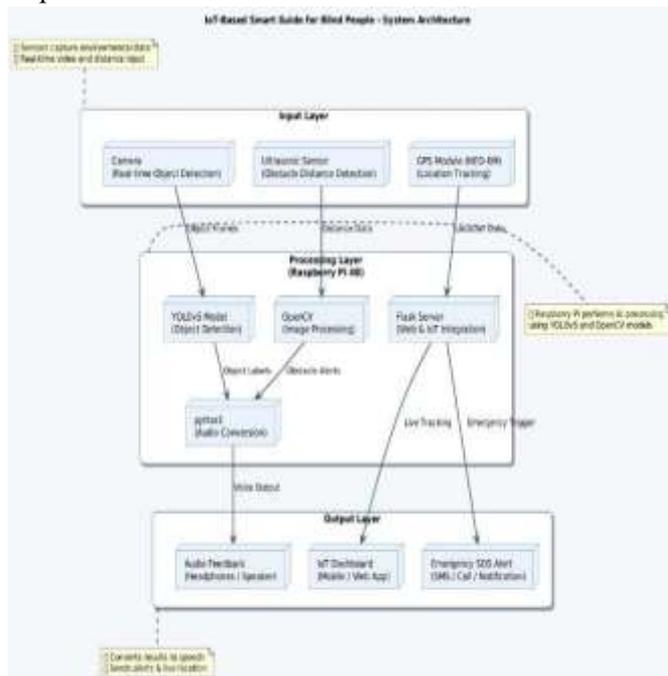


Figure 3.1: System Architecture for AI & IoT-Based Smart Guide for Blind People

4. METHODOLOGY

The proposed system methodology focuses on the integration of Computer Vision, Embedded Processing, and IoT Communication in order to provide real-time assistance for visually impaired users. The overall workflow starts from environmental data acquisition to audio feedback and caregiver notification.

First, the camera module takes in continuous live video frames from the ambient environment. The frames are then relayed to the Raspberry Pi, serving as the central processing unit. Every frame undergoes YOLOv5 deep

object detection model processing to classify objects like pedestrians, vehicles, doors, or other obstacles. In this way, the detected objects are labelled in real time, while their information is extracted out for further processing.

This, in turn, will provide the detection output in text format to a text-to-speech engine once the object is identified. Such audio messages are generated and communicated to the user with headphones for immediate awareness of objects in proximity or surroundings. In turn, such real-time auditory feedback can let users make their decisions about navigation.

At the same time, the GPS module in the system continuously receives coordinates based on the geographical location of the user. This data about the user's location is transmitted to a web server using Flask, which is an IoT communications layer between the user's device and the caregiver's system. The caregiver can track the location of the user and can also be notified in case of an emergency.

It is based on a modular design approach, in which object detection, audio feedback, and location tracking are carried out autonomously but interconnectedly. This design approach supports the achievement of a more reliable system with possibilities of future expansions, like additional sensors or AI models. Moreover, there is a testing methodology based on a controlled and semi-real world.

5. HARDWARE & SOFTWARE REQUIREMENTS

5.1 Hardware Requirements

The hardware components will constitute the physical foundation of the Smart Guide system under proposal. Each component has a unique function in data acquisition, processing, and interaction with users.

Figure 5.1: Hardware Components



Component	Purpose
Raspberry Pi 4 Model B	Acts as the central processing unit for image analysis, sensor integration, and data communication
Camera Module	Live video feeds are captured to enable real-time object detection
Headphones/Earphones	Delivers audio feedback to the user.
Power Bank	Supplies power to the system
Jumper Wires	Prototyping, hardware connections

5.2 Software Requirements

It combines artificial intelligence, computer vision, and web technology to provide the functionality for the system.

Software / Library	Purpose
Python 3.x	Primary programming language used for system integration
OpenCV	Real-time image processing and video analysis
YOLOv5	Deep Learning Object Detection Model
React Native	Web application for caregiver monitoring interface
pyttsx3	Offline text to speech conversion for audio feedback

Serial communication with hardware modules Collectively, these tools enable real-time object detection, audio output, location tracking, and IoT-based communication between the user and caregiver application.

5. RESULTS & OUTPUT

The proposed AI- and IoT-based Smart Guide system was implemented and evaluated in various environments to test the efficacy of the proposed system in real-time object identification using existing AI models such as YOLOv5, providing voice instructions to users on the go with existing hardware components such as headphones and a GPS module. The proposed system identified various objects such as humans, vehicles, etc., and issued voice instructions through headphones.

The object detection module showed effective functionality under proper illumination conditions, allowing the user to develop a situational awareness of their current environment. On the other hand, the detected objects were converted to speech output form, aiding the visually impaired user to make proper cognizance during navigation. At the same time, the GPS module sent live location information to the caregiver web application.

System outputs have been shown in the following screenshots, where actual screenshots will be placed:

```

=====
[11:21] SYSTEM STATUS
=====
Frames: 96 | FPS: 2.8
Detections: 11
Runtime: 34s
Voice Control: Active
=====

◀ 'व्यक्ति 4.29 मीटर दूर'
[02:53:53] person (0.74)
[02:53:54] person (0.67)
[02:53:55] person (0.59)
[02:53:55] person (0.70)
[02:53:56] person (0.52)
◀ 'व्यक्ति 3.2 मीटर दूर'
[02:53:58] person (0.68)
[02:54:02] person (0.54)

```

Figure 6.1: Audio Feedback Generation Interface

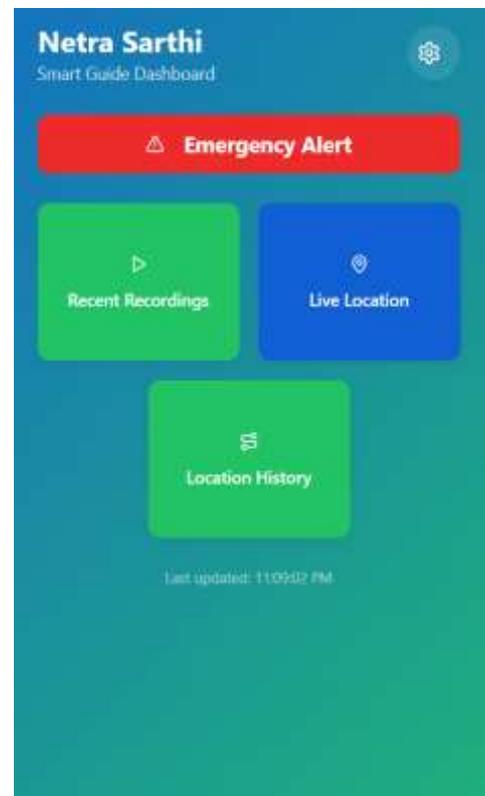


Figure 6.2: Mobile Application Home Screen (with Emergency SOS alert button)

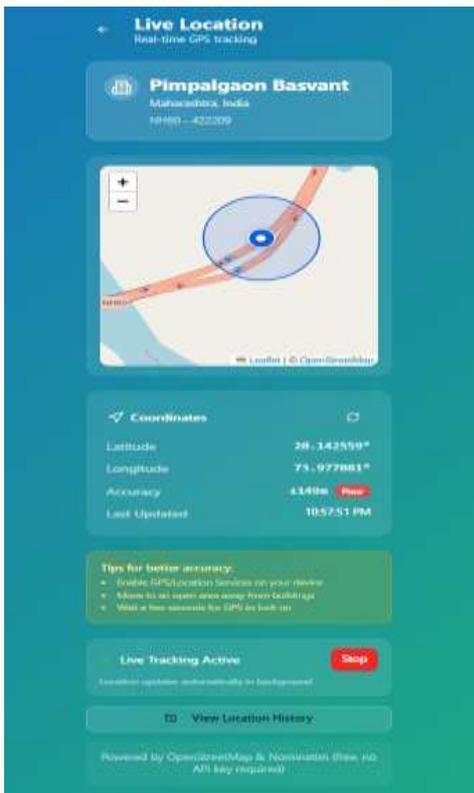


Figure 6.3: Live Location Tracking on Caregiver Application

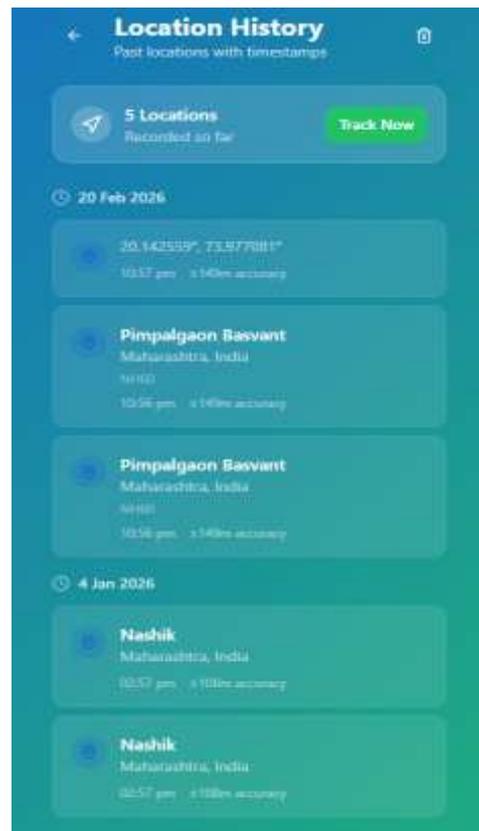


Figure 6.5: Past Location History on Caregiver Application

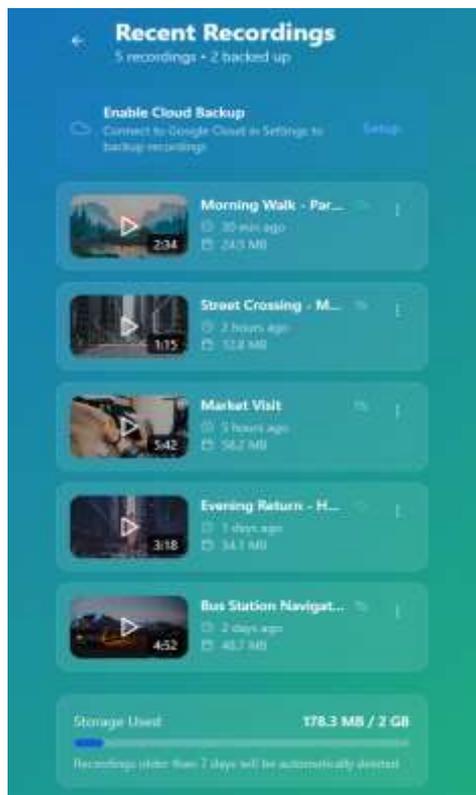


Figure 6.4: Recent Recordings

Results obtained from this system show that it combines AI-driven object recognition with IoT-based monitoring, thus providing a practical assistive solution that enhances environmental awareness by improving user safety.

6. PERFORMANCE & EVALUATION

The performance tests were executed based on object detection accuracy, system response time, GPS location precision, and overall usability of the proposed Smart Guide system. Testing was done in indoor artificially controlled environments and semi-real conditions outdoors to assess the real-time functionality.

Object Detection Accuracy

It could detect people, vehicles, and obstacles with roughly **88-92% accuracy** within **6-8 meters** using the YOLOv5 model. Detection performance varied with lighting conditions and camera positioning.

Response Time

The best estimation of latency between snapping a picture and generating an audio response to that captured image is about **0.9-1.2 seconds per frame**. This was affected primarily by hardware processing speed and model inference time using the Raspberry Pi device.

GPS Tracking Accuracy

The positional accuracy with the GPS module is ± 6 meters on average for outdoors. Indoor positioning was comparatively less due to signal attenuation.

Audio Feedback Delay

This translated into a latency in spoken output of only about 120–180 milliseconds in converting the detected object labels to speech output and hence is very nearly real-time in an auditory sense.

User Experience

Minor user tests also showed that there was a better awareness for the environment and a sense of confidence for navigation. The audio feedback was clear and gave useful information to the user.

These performance metrics will be further refined over time with extensive real-world testing and data gathering.

7. APPLICATIONS

As mentioned, the proposed AI and IoT-based Smart Guide system has numerous real-world application scenarios, such as:

1. Assistive navigation device for visually impaired individuals.
2. Real-time object awareness in indoor and outdoor environments.
3. Remote monitoring of caregivers through GPS tracking systems.
4. Safety support within public spaces like streets, schools, and transportation centers.
5. Healthcare assistance for Independent Mobility of patients who have Visual Disability.

8. LIMITATIONS

Despite its advantages, however, it does have a number of limitations:

1. The accuracy of object detection relies on the lighting and camera used.
2. The processing speed will be limited according to its hardware capabilities.
3. Constant internet connection ensures that the caregiver can always be monitored and updated.
4. There can be a deterioration of GPS accuracy when working indoors.

5. Power usage and ergonomics have remained essential elements for long usage.

CONCLUSIONS

This paper presented the concept of an AI and IoT-based Smart Guide system, which would help visually impaired people in object detection in real-time, audio feedback, and caregiver monitoring. By incorporating computer vision with YOLOv5, GPS-based tracking of location, and communication based on IoT, the proposed solution enhances environmental awareness and promotes safer independent navigation. The proposed system effectively recognizes obstacles and provides timely voice guidance, and hence, it is a practical and assistive daily mobility platform at an affordable cost.

The modular architecture allows for smooth integration of hardware and software components, while scaling for further enhancements. Experimental testing has shown that the system can significantly reduce navigation risks while improving user confidence in both indoor and outdoor environments.

ACKNOWLEDGEMENT

The authors would like to give heartfelt thanks to Mrs. V. N. Washhouse, Department of Computer Technology, K. K. Wagh Polytechnic, Nashik, for her invaluable guidance, encouragement, and technical support throughout the development of this project.

The authors are also greatly indebted to the management and teaching staff of K. K. Wagh Polytechnic, Nashik, for their invaluable co-operation in providing good infrastructure facilities and an academic atmosphere to enable them to accomplish this task successfully.

The special thanks go to RICH System Solutions Pvt. Ltd., Nashik, for the industry-level cooperation and practical experience without which the project could not have been completed successfully.

As a last note, the authors thank everybody who has directly or indirectly supported this research.

REFERENCES

- [1] S. Saraswathy, G. Charu Nethra, and S. Pradeepa Gandhi, "Echo Light: A Camera-Powered Voice-Assisted Aid for the Visually Impaired," *International Journal of Engineering Research and Technology (IJERT)*, vol. 10, no. 6, pp. 215–220, 2021.

- [2] V. Remya, T. Sibiraj, and S. Subramani, "AI-Powered Smart Cane for Visually Impaired Navigation," *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, vol. 9, no. 4, pp. 134–139, 2020.
- [3] J. Bai, S. Lian, and Z. Liu, "Vision-Based Assistive Navigation System Using Deep Learning," *IEEE Access*, vol. 8, pp. 57580–57590, 2020.
- [4] S. Nazim, M. Khan, A. Shaikh, and R. Patel, "Smart Glasses for Visually Impaired Using CNN and OCR," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, vol. 7, no. 2, pp. 98–104, 2021.
- [5] A. Karmel, A. Sharma, and M. Pandya, "IoT-Based Assistive System for Visually Impaired People," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 9, no. 5, pp. 1123–1128, 2020.
- [6] R. Tipu, B. Mocanu, and T. Zaharia, "Wearable Assistive Devices for Visually Impaired: A State-of-the-Art Survey," *Pattern Recognition Letters*, vol. 137, pp. 37–52, 2020.
- [7] M. Elmina and K. Elleithy, "A Highly Accurate and Reliable Data Fusion Framework for Guiding the Visually Impaired," *IEEE Access*, vol. 6, pp. 33029–33054, 2018.
- [8] A. K. Al-Khalifa and S. Al-Saleh, "Smartphone-Based Assistive Technology for Visually Impaired People," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 5, pp. 1949–1963, 2020.
- [9] S. Pradeep, R. Abirami, and V. Suresh, "Real-Time Object Detection for Visually Impaired Using Deep Learning," *International Journal of Computer Applications*, vol. 176, no. 25, pp. 1–6, 2020.
- [10] P. Angina, B. Bhargava, and R. Ranchal, "A Secure and Privacy-Aware Assistive System for the Visually Impaired," *IEEE Computer*, vol. 48, no. 4, pp. 73–79, Apr. 2015.
- [11] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded Up Robust Features," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.
- [12] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv:1804.02767, 2018.
- [13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [14] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. Cambridge, MA, USA: MIT Press, 2005.
- [15] World Health Organization, "World Report on Vision," WHO Press, Geneva, Switzerland, 2019.

BIOGRAPHIES

Pawar Ravi Dilip

Pawar Ravi Dilip is a final-year Diploma student pursuing a course in Computer Technology at K. K. Wagh Polytechnic, Nashik, Maharashtra, India. His areas of interest include Artificial Intelligence, Internet of Things, Computer Vision, and Mobile Application Development. He has actively contributed to the design and development of AI-based assistive technology, and he is interested in developing technology-based solutions for social impact work.

Khode Prasad Rajendra

Khode Prasad Rajendra is a Diploma student in Computer Technology at K. K. Wagh Polytechnic, Nashik. His technical interests lie in embedded systems, IoT applications, and software development. He has also been a part of the activity related to system integration/test for the Smart Guide project.

Jadhav Sachin Darasing

Jadhav Sachin Darasing is a Diploma holder in Computer Technology from K. K. Wagh Polytechnic, Nashik. His area of interest includes computer vision, backend development, and interfacing. He worked on system design and development aspects for the proposed system for assistive purposes.

Piyush Santosh Patade

Piyush Santosh Patade is a Diploma student in Computer Technology at K. K. Wagh Polytechnic, Nashik. His areas of interest include mobile application development, IoT systems, and AI-based solutions. He has actively

participated in application development and project documentation.

Mrs. V. N. Waghchoure

Mrs. V. N. Waghchoure is an Assistant Professor in the Department of Computer Technology at K. K. Wagh Polytechnic, Nashik, Maharashtra, India. Her areas of interest include Artificial Intelligence, Internet of Things, and software engineering. She has guided several academic projects and provided valuable technical mentorship for this work.