IRJCS: International Research Journal of Computer Science
Volume 13, Issue 03, March 2026
ISSN: 2393-9842
https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1303.40

Blind Assistive System

Prof.V.Keerthana
Assistant Professor,Department of computer science,
Sengunthar Engineering College (Autonomous),Tiruchengode, India
vkeerthana.cse@scteng.co.in
https://orcid.org/0009-0008-5118-5945
Boobalan N,Deepan Raj S, Dinesh Kumar S
UG Students, Department of Computer Science & Engineering
Sengunthar Engineering College (Autonomous),Tiruchengode, India
boobalanm941@gmail.com,deepanjack0718@gmail.com,dineshsiva1102@gmail.com

Publication History
Manuscript Reference: IRJCS/RS/Vol.13/Issue03/CSMR26.MRCS10119
Research Article | Open Access | Double-Blind Peer Reviewed Article ID: IRJCS/RS/Vol.13/Issue03/CSMR26.MRCS10119
Received: 30, January 2026, Revised: 13, February 2026, Accepted: 28 February 2026 Published Online: 25 March 2026
https://www.irjcs.com/volumes/Vol13/iss-03/40.CSMR26.MRCS10119.pdf
Article Citation: Prof.Keerthana,Boobalan,Deepan,Dinesh(2026),Blind Assistive System, IRJCS: International Research Journal of Computer Science, Volume 13,Issue 03 of 2026 pages 335-340
Doi:> https://doi.org/10.26562/irjcs.2026.v1303.40 **BibTeX Key Prof.Keerthana@2026Blind**
Orcid: https://orcid.org/0009-0004-9398-7488
IRJCS papers should be cited as IRJCS (International Research Journal of Computer Science, AM Publications, India 2026, ISSN 2393-9842, https://doi.org/10.26562/ijrcs.2025.v1303.40 The journal's official abbreviation is IRJCS.

Abstract: Visually impaired individuals face significant challenges in navigating dynamic environments safely and independently. This paper proposes an intelligent blind assistive system that integrates deep learning models specificallyYOLOv8 and Faster R-CNN to perform real-time object detection with high accuracy. The system captures visual information through a camera-equipped wearable device, processes frames using the deep learning pipeline, and delivers audio feedback to the user, enabling obstacle detection and spatial awareness. The proposed architecture in corporate preprocessing, feature extraction, multi-class object classification, and voice synthesis modules. Experimental evaluations demonstrate significant improvements in detection accuracy, reduced latency, and enhanced adaptability across complex and dynamic environments compared to existing assistive systems. This scalable solution aims to bridge the gap between human visual perception and machine understanding, promoting independence and safety for visually impaired users.

Keywords: Blind Assistive System, Object Detection, YOLOv8, Faster R-CNN, Deep Learning, Real-Time Processing, Visually Impaired, Wearable Device, Accessibility.

INTRODUCTION

Visual impairment is one of the most prevalent sensory disabilities worldwide, affecting hundreds of millions of people who face constant challenges in navigating their surrounding safely. Traditional assistive technologies, such as white canes and guide dogs, provide limited situational awareness and do not leverage modern computational capabilities. The rapid advancement of deep learning and computer vision has opened new possibilities for developing intelligent assistive systems that can interpret and describe the surrounding environment in real time Assistive systems for the visually impaired aim to bridge the gap between human visual perception and machine understanding by providing automated, context-aware support. The core challenge is achieving object detection with high accuracy, low latency, and robustness across diverse and dynamic environments including varying lighting conditions, cluttered backgrounds, and fast- moving objects. Existing solutions often suffer from computational constraints, limited accuracy, or lack of real-world deployability. Recent advancements in machine learning (ML) have enabled intelligent detection frameworks capable of identifying complex patterns. ML-based approaches offer significant advantages by learning hidden patterns, correlations, and anomalies within visual data without relying on hand-crafted features . Studies have demonstrated the effectiveness of various architectures including YOLOv8, Faster R-CNN, and MobileNet in classifying objects and detecting obstacles relevant to assistive applications. This paper proposes an enhanced blind assistive system integrating YOLOv8 and Faster R-CNN with an audio feedback mechanism, offering real-time, reliable, and user- friendly object detection. The system is designed for deployment on wearable devices such as smart glasses or handheld gadgets. The proposed framework achieves improved detection accuracy, reduced false positives and negatives, and supports seamless real-time processing.

LITERATURE REVIEW

Modern assistive technology increasingly depends on intelligent object detection systems to support visually impaired users. Traditional assistive mechanisms centered on rule-based obstacle avoidance and basic image recognition face limitations such as inability to detect diverse object categories, high false-positive rates, and lack of adaptability to dynamic environments.

To address these challenges, deep learning-enabled systems have emerged, offering automated object identification, behavioral analysis, and scalable deployment across mobile and embedded devices. Studies have shown that ML-based approaches significantly improve detection accuracy for real-world assistive applications compared to conventional methods. Deep learning architectures and flow-based analytics have further advanced the field by enabling real- time classification and multi-class object detection [7],[8]. Several systems have explored the integration of various ML algorithms including Faster R-CNN, SSD, YOLOv3, and YOLOv8 with image features such as object size, lighting conditions, spatial context, and depth data. These systems analyze visual data collected via wearable cameras, RGB-D sensors, or public datasets including MS-COCO, Pascal VOC, and Open Images. Supervised learning models trained on labeled datasets classify detected objects as obstacles or navigable regions, while unsupervised techniques detect anomalies without prior labeling. Dashboards and audio feedback interfaces present detection results, threat severity levels, and spatial patterns to users, enabling informed navigation. Some systems also incorporate automated voice alerts, haptic feedback, and GPS integration to further assist users. However, many existing implementations remain fragmented, lacking unified platforms that combine real- time visual analysis, adaptive learning, and comprehensive alerting within a single wearable architecture. Despite these advancements, persistent challenges remain. High computational requirements of deep learning models limit real-time deployment in resource-constrained Embedded devices. Class imbalance in training datasets where common objects vastly outnumber rare but critical obstacles degrades model performance for minority categories. Visual analysis remains challenging in low- light or occluded environments. Adversarial conditions such as rapid motion, reflective surfaces, and complex backgrounds pose emerging detection challenges the proposed blind assistive system addresses these gaps by offering a unified, scalable platform that integrates visual input acquisition, feature engineering, YOLOv8 classification, and real-time audio alerting within a single architecture. Unlike prior systems, it combines supervised learning with statistical anomaly detection, modular preprocessing pipelines, and interactive audio interfaces tailored for visually impaired users. By bridging the gap between fragmented implementations and comprehensive assistive technology, this study contributes a novel solution aligned with the goals of robust, adaptive, and intelligent accessibility.

PROPOSED METHODOLOGY ARCHITECTURE

The proposed system architecture for the Blind Assistive System follows a three-layer monitoring framework consisting of visual data acquisition, machine learning- based object analysis, and audio feedback generation. Visual frame data is collected using a camera module embedded in wearable glasses or a handheld device, capturing parameters including frame resolution, object bounding boxes, confidence scores, class labels, and spatial depth estimates. The collected data is transmitted to a processing environment where preprocessing and feature extraction are performed. Edge-level preprocessing modules filter redundant frames and remove noisy data, reducing computational overhead by approximately 25– 35%. The cleaned data is then processed by deep learning models that analyze visual content and detect object categories associated with assistive navigation scenarios. The on-device analytics layer stores processed data and runs scalable deep learning algorithms capable of identifying objects in real time. Audio feedback interfaces provide real-time descriptions of detected objects, proximity alerts, and potential navigation indicators. This architecture enables scalable and intelligent monitoring of dynamic environments to assist visually impaired individuals.

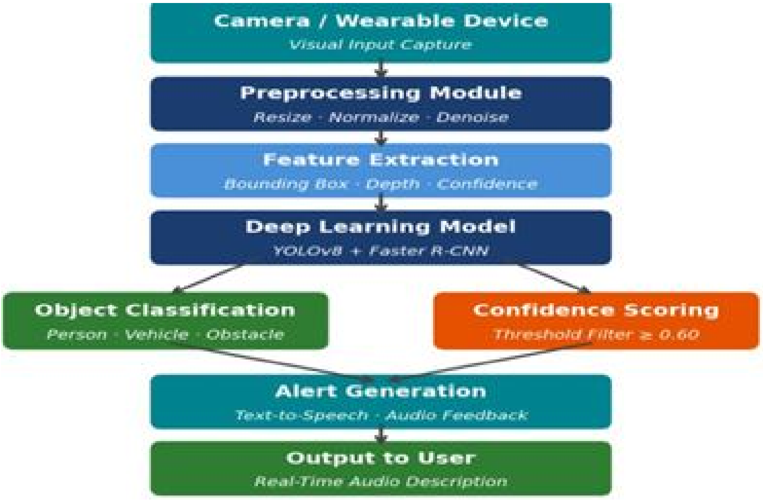


Fig.1. Architecture Diagram

A.System Architecture Design

The proposed methodology implements a multi-layer visual monitoring system, as shown in Fig.1. The architecture consists of three main components: data acquisition layer, processing layer, and analytics layer. The data acquisition layer collects visual frames from wearable cameras and depth sensors. The processing layer performs data preprocessing, feature extraction, and deep learning model inference.

The analytics layer provides audio feedback and real-time alerts for the user. This architecture ensures continuous monitoring of the surrounding environment and enables early detection of obstacles and navigation hazards for visually impaired individuals.

B. Visual Frame Monitoring Protocol

The system continuously monitors visual input using frame capture tools and depth-sensing mechanisms. Frame data contains important attributes such as object class, bounding box coordinates, confidence score, depth estimate, and motion vectors. These parameters help identify critical objects such as vehicles, pedestrians, furniture, and stairs. Threshold-based detection mechanisms generate alerts when high-confidence detections exceed predefined proximity limits. Continuous monitoring allows the system to detect obstacles in real time including static barriers, moving hazards, and navigable pathways.

C. Data Processing Framework

The data processing framework includes several stages such as frame resizing, normalization, and feature selection. Raw visual frames often contain redundant or noisy pixels that must be filtered before analysis. Feature extraction techniques identify the most relevant visual attributes contributing to accurate object detection. Important features include bounding box dimensions, object class confidence, spatial overlap ratios, and motion trajectory. The processed dataset is then used for running deep learning inference that classifies visual content into object categories. The structured data is stored in a local database for further analysis and model refinement.

D. Object Detection System

The object detection module utilizes deep learning models such as YOLOv8, Faster R-CNN, and MobileNet classifiers to identify visual objects. These models are trained using labeled datasets containing both obstacle and non-obstacle images. Once trained, the models analyze incoming visual frames and detect objects that may present navigation challenges for visually impaired users. When critical objects are detected, the system automatically generates audio alerts to notify the user for situational awareness.

E. Security Implementation

To ensure system reliability and secure data handling, several security mechanisms are implemented. Data communication between the wearable camera and the analysis processor is protected using encrypted communication protocols such as TLS/SSL. User authentication and role-based access control restrict unauthorized access to the monitoring system. Regular security audits and vulnerability assessments ensure the integrity and reliability of the assistive detection platform.

F. Performance Validation

The effectiveness of the proposed system is evaluated using deep learning performance metrics such as accuracy, precision, recall, and F1-score. Experimental results show that deep learning-based detection methods significantly improve the identification of obstacles and navigational hazards compared to traditional rule-based systems. The system demonstrates strong capability in detecting objects associated with daily assistive navigation tasks.

TECHNOLOGIES USED

A. Visual Input Capture Modules

The system utilizes camera modules to capture and analyze frame-level data from the user's environment. Cameras such as Raspberry Pi Camera Module and IntelReal Sense depth sensors are used to collect visual frames and convert them into structured input for the deep learning pipeline. These modules capture important features including scene content, object bounding boxes, depth maps, and lighting conditions. The captured visual data forms the primary input for deep learning-based object detection.

B. Data Collection and Frame Monitoring

Frame monitoring technologies are used to aggregate and summarize visual input data. Instead of analyzing individual pixels, the system focuses on object-level characteristics such as bounding box size, class confidence score, detection frequency, and proximity estimate. This approach significantly reduces processing overhead while preserving important object information required for detecting navigational hazards.

C. Machine Learning Framework

Machine learning frameworks such as PyTorch, TensorFlow, and Keras are used to build and train the object detection models. These frameworks support the development of classification algorithms that analyze visual features and identify objects in real time. The trained models learn from labeled training data and detect objects that may indicate possible navigation hazards, including stairs, vehicles, pedestrians, and obstacles.

D. Data Processing and Feature Engineering

Data preprocessing and feature extraction are critical steps in the detection pipe line. Python libraries such as OpenCV, Pandas, and NumPy are used to clean, normalize, and transform raw visual frame data into structured datasets. Feature engineering techniques are applied to extract meaningful attributes from visual frames, improving the performance and accuracy of the deep learning models.

E. Database Management System

A MySQL database is used to store captured visual frame records, processed datasets, and deep learning model outputs. The database enables efficient storage, retrieval, and management of large volumes of visual data. It also supports historical analysis of detection behavior, allowing developers to track long-term model performance patterns.

F. Backend Integration Framework

The backend system is developed using the Flask framework.

Flask provides a lightweight environment for integrating deep learning models with camera modules and databases. It also enables the development of APIs that allow communication between the detection engine and the user interface or companion mobile application.

G. Visualization and Monitoring Dashboard

Visualization tools such as Matplotlib, Seaborn,and Plotly are used to present detection analytics in graphical form. The monitoring dashboard displays object detection statistics, confidence score distributions, and class frequency patterns. These visual insights help developers quickly understand model performance and monitor overall detection accuracy.

H. Security and Threat Detection Mechanism

The proposed system implements anomaly detection techniques to identify unusual visual patterns. Deep learning models analyze visual frame attributes and classify objects as safe or hazardous. Alerts are generated when critical objects are detected, enabling early identification of navigational risks for the visually impaired user.

I. Web-Based Monitoring Application

A web-based monitoring interface provides administrators and caregivers with real-time visibility into detection activity and system performance results. The dashboard displays detection statistics, confidence alerts, and historical analytics. Role-based access control ensures that only authorized users can access sensitive system data.

IMPLEMENTATIONS AND RESULTS

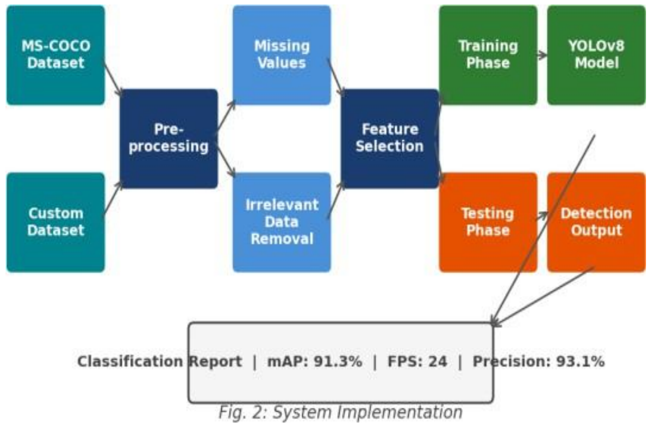


Fig.2: System Implementation

The training phase employs the YOLOv8 algorithm, an ensemble deep learning method that sequentially refines detection through multi-scale anchor-free prediction heads to improve accuracy from previous iterations. In the testing phase, the saved model performs real-time classification on unseen visual frames, effectively distinguishing between safe environments and various obstacle categories for proactive navigation assistance as shown in Fig. 2.

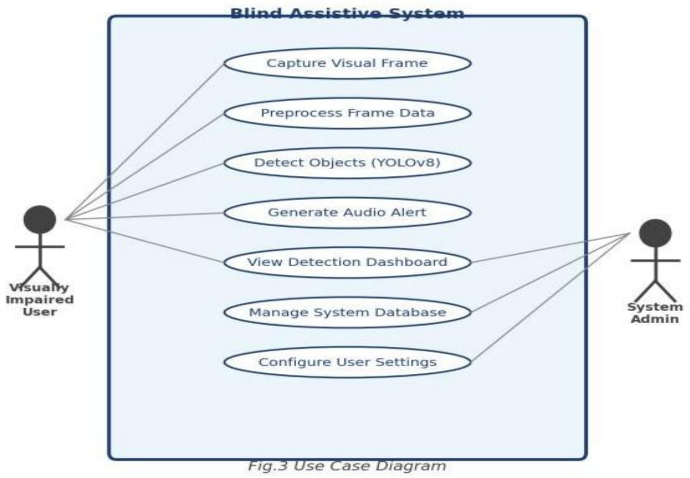
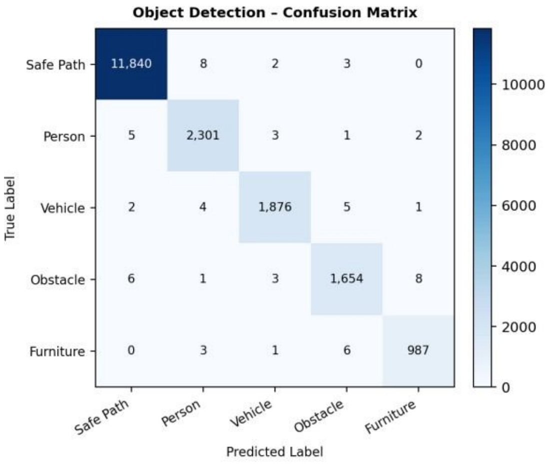


Fig.3 UseCase Diagram

The Use Case Diagram illustrates the interaction between users and the Blind Assistive System. The Visually Impaired User interacts with the system by wearing the device, which automatically processes visual input and provides audio alerts. The System Administrator manages model updates and database maintenance. This diagram helps in clearly understanding the responsibilities of each actor and the functional capabilities of the system as shown in Fig 3. The performance of the proposed Blind Assistive System was evaluate during a confusion matrix, as shown in Fig.4.

The matrix presents the classification results across five categories: safe path, person, vehicle, static obstacle, and furniture. The diagonal elements represent correctly classified instances, while off-diagonal values indicate misclassifications. The distribution of detected object categories across the analyzed visual input is shown in Fig. 5, which illustrates the frequency counts for five distinct categories: safe path, persons, vehicles, static obstacles, and furniture.



	Safe Path	Person	Vehicle	Obstacle	Furniture
Safe Path	11,840	8	2	3	0
Person	5	2,301	3	1	2
Vehicle	2	4	1,876	5	1
Obstacle	6	1	3	1,654	8
Furniture	0	3	1	6	987

Fig.4 Object Detection Confusion Matrix

Fig.4 Object Detection Confusion Matrix

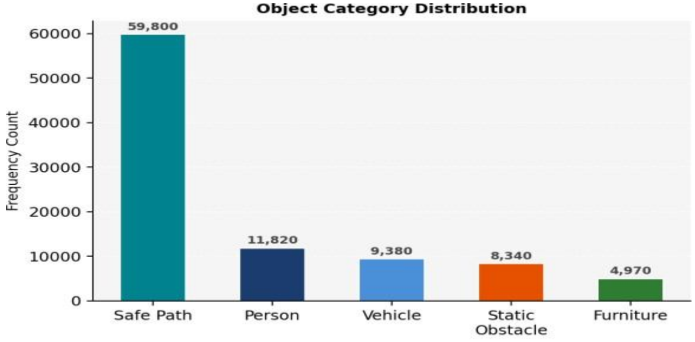


Fig. 5 Object Category Distribution Graph

Fig.5 Object Category Distribution Graph

As shown, safe path regions dominate the dataset with the highest frequency, reflecting the realistic class imbalance inherent in daily assistive environments where navigable regions vastly outnumber hazardous objects encountered during typical mobility scenarios.

Key Considerations:

Data Quality and Availability: The collected visual frame data must be accurate, complete, and properly labeled to ensure reliable object detection. Poor quality or incomplete data may reduce the effectiveness of the detection system.

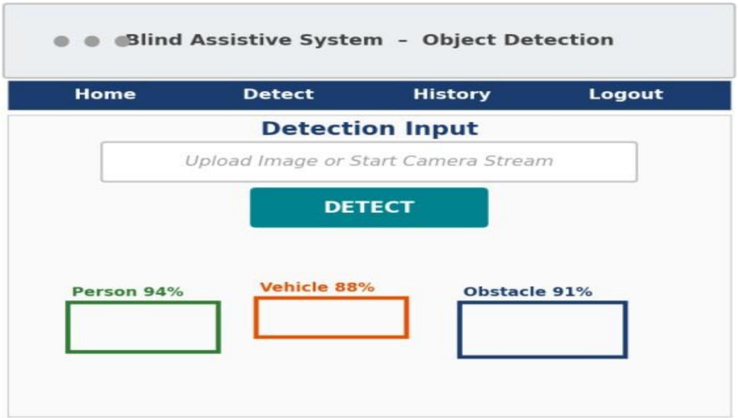
Feature Selection: Selecting the most relevant visual features is critical for improving detection accuracy. Features such as bounding box dimensions, confidence score, object class, and depth estimate should be carefully chosen to help the deep learning model distinguish between safe and hazardous regions.

Model Accuracy and Performance: The deep learning algorithms used in the system must be optimized to achieve high detection accuracy while minimizing false positives and false negatives. Continuous model evaluation and fine- tuning are required to ensure reliable identification of navigation hazards.

Real-Time Detection Capability: Visual navigation hazards often appear suddenly and require immediate response. Therefore, the detection system must support real-time or near real-time analysis of visual frames to quickly identify obstacles and generate audio alerts for the visually impaired user.

IV.CONCLUSION

The proposed system presents an effective approach for detecting obstacles and providing real-time audio guidance to visually impaired individuals using deep learning techniques and visual frame analysis. By analyzing visual input patterns and extracting important object features, the system is able to identify hazardous objects that may present navigation challenges. Deep learning algorithms improve the accuracy and efficiency of object detection by automatically learning patterns from visual training data. The integration of camera modules, data processing techniques, and audio feedback interfaces enables users to navigate environments and detect obstacles in a timely manner. The experimental results demonstrate that the proposed system can successfully classify visual content as safe or hazardous and provide early detection of potential navigation risks.

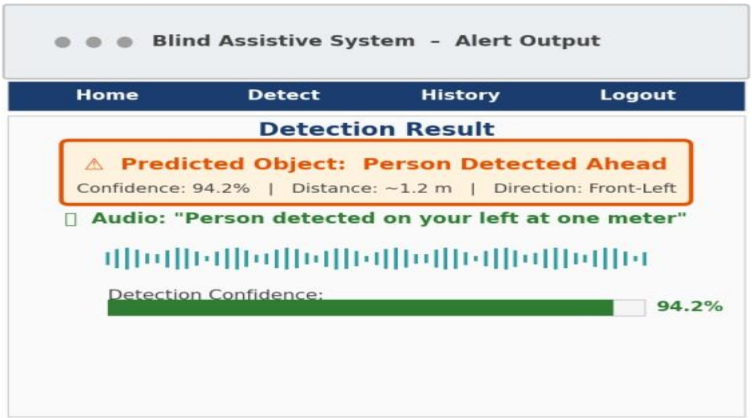


Detection Input

Upload Image or Start Camera Stream

Fig 6: Detection Output Page

Fig 6: Detection Output Page



Detection Result

⚠ Predicted Object: Person Detected Ahead

Confidence: 94.2% | Distance: ~1.2 m | Direction: Front-Left

☐ Audio: "Person detected on your left at one meter"

Detection Confidence: 94.2%

Fig 7: Audio Alert Output

Fig 7: Audio Alert Output

REFERENCES

1. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proc. IEEE CVPR, 2016, pp. 779-788.
2. S.Ren,K.He,R.Girshick,andJ.Sun,"FasterR-CNN: Towards real-time object detection with region proposal networks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137-1149, 2017.
3. G.Jocheretal.,"UltralyticsYOLOv8,"GitHub,2023. [Online]. Available: https://github.com/ultralytics/ultralytics
4. W.Liu etal., "SSD: Single shotmultiboxdetector,"in Proc. ECCV, 2016, pp. 21-37.
5. A. Bochkovskiy, C.Y.Wang, and H.Y.M.Liao, "YOLOv4: Optimal speed and accuracy of object detection," arXiv:2004.10934, 2020.

A. G.Howard et al., "Mobile Nets: Efficient convolutional neural networks for mobile vision applications," arXiv:1704.04861, 2017.

6. N.Carion et al., "End-to-end object detection with transformers," in Proc. ECCV, 2020, pp. 213-229.
7. T.-Y.Linetal.,"MicrosoftCOCO:Common objects in context," in Proc. ECCV, 2014, pp. 740-755.
8. Simonyan and A.Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. ICLR, 2015.
9. D.Ouyang et al., "Assistive technology for visually impairedpersons:Asurvey,"IEEEAccess,vol.10,pp. 12345-12370, 2022.

A. Dosovitskiy et al., "An image is worth 16x16 words: Transformers for image recognition at scale,"in Proc. ICLR, 2021.

10. P.Viola and M.Jones, "Rapid object detection using a boosted cascade of simple features," in Proc. IEEE CVPR, 2001, pp. 511-518.
11. Zhang et al., "Multi-modal learning for visually impaired assistive systems," in Proc. IEEE ICASSP, 2025, pp. 1-5.
12. V.Badrinarayanan, A.Kendall, and R.Cipolla, "SegNet: A deep convolutional encoder-decoder architecture," IEEETrans. Pattern Anal. Mach. Intell., vol. 39, no. 12, pp. 2481-2495, 2017.
13. Y.LeCun,Y.Bengio,andG.Hinton,"Deeplearning," Nature, vol. 521, no. 7553, pp. 436-444, 2015.