© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1715263

NETRA-AI: An AI-Based Real-Time Vision Assistance System for Visually Impaired People Using Object Detection and Voice Guidance

SHUBHAM KUMAR[1], RAHUL KUMAR[2], SUNNY PATEL[3], ABHISHEK KUMAR GUPTA[4], DR. VANDANA KUMARI[5]

[1, 2, 3, 4, 5] *Department of Data Science, Lloyd Institute of Engineering and Technology, Greater Noida, India*

Abstract- ***The condition of visual impairment limits a person's ability to perceive what exists in their surroundings and to recognize objects which appear in their surroundings. People who experience visual impairment require artificial intelligence and computer vision solutions which enable them to perceive their surroundings while maintaining their independence. The researchers developed NETRA-AI which operates as an advanced vision assistant with real-time capabilities. The system uses speech synthesis to provide audio feedback after its deep learning system detects objects. The system enables users to perform object detection through voice commands while it provides automatic detection mode and emergency SOS alerts and reminder systems and memory assistance and multi-camera functionality. The system enables users to access multiple camera sources which include laptop cameras and mobile cameras and USB cameras and IP cameras for use in different settings. The experimental results demonstrate that the system effectively identifies common objects while producing instant auditory alerts which improve both safety and situational awareness for users who are visually impaired.***

Index Terms- ***Assistive Technology, Computer Vision, YOLOv8, Artificial Intelligence, Vision Assistance, Accessibility***

I. INTRODUCTION

About 285 million people worldwide suffer from vision loss which creates problems for them to move safely and use their surroundings effectively. The basic travel needs of white cane and guide dog users become difficult because their mobility tools do not reveal all environmental information. The development of intelligent assistive systems became possible through artificial intelligence and computer vision technologies which allow these systems to perform object recognition and generate instant environmental descriptions. The YOLO object detection algorithm operates with both speed and precision which makes it suitable for use in assistive technology systems. The NETRA-AI system operates as an AI-based visual assistant, which identifies objects in real time. The system offers voice-based guidance to all its users. The system unites six separate functions into one platform which combines object detection with voice command support and emergency alert systems and reminder services and memory storage capabilities. The system delivers total environmental understanding to users who cannot see which enables them to achieve independence through safe navigation.

II. PROBLEM STATEMENT

Visually impaired persons usually find difficulties in identifying objects and obstacles in their environments. The current assistive technologies available in the market have various shortcomings in the form of higher costs, dependency on hardware, and lack of integrated features such as emergency alerts, voice commands, reminder assistance, and support for multiple cameras.

Therefore, there is a need for the development of an intelligent assistive system that can assist the visually impaired in detecting objects in real-time and provide voice-based assistance while incorporating various safety and assistance features.

The proposed NETRA-AI system aims at resolving the issues associated with the current assistive technologies in the form of emergency alerts, voice

© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1715263

commands, reminder assistance, and support for multiple cameras within a single assistive system.

III. RELATED WORK

Research studies about computer vision and artificial intelligence studies have investigated their potential to help people who cannot see well. The camera sensors enable smart wearable systems to detect obstacles while they provide navigation support.

Developers have created mobile apps which use deep learning models to identify objects through smartphone cameras while they produce audio descriptions for users. The YOLO object detection framework by Redmon et al. brought revolutionary improvements to real-time object detection capabilities. The developers introduced YOLOv5 and YOLOv8 which brought better detection precision together with faster processing speeds.

The current systems operate with expensive hardware but they don't include essential features which combine emergency alerts with reminder systems and voice command interaction. The NETRA-AI system solves these problems by offering an adaptable AI-based assistive system which unites various functions into one unified platform.

IV. SYSTEM ARCHITECTURE

The NETRA-AI system contains different modules which perform image capturing and object detection and voice output and emergency alert generation.

Architecture Flow

Figure 1: System Architecture of NETRA-AI Vision Assistance

The system architecture has two new modules which developers integrated into the existing framework:

- Voice command recognition module
- Emergency SOS alert system
- Reminder system
- Memory assistance module
- Multi-camera switching module

These modules together provide a comprehensive assistive platform.

V. METHODOLOGY

5.1 Object Detection

The system operates through the YOLOv8 deep learning model which has received training from the COCO dataset that contains 80 different object categories.

YOLO processes images by splitting them into S × S grid sections. Each section of the grid generates bounding box predictions together with class probability estimates.

The detection confidence score is calculated as:
The formula for Confidence calculates object detection probability by multiplying P(Object) with IOU which represents the Intersection over Union between predicted and ground-truth bounding boxes.

Figure 2: Real-time object detection using YOLO8 in NETRA-AI

5.2 Voice Guidance System

The system transforms detected objects into audio through its text-to-speech conversion process. The system analyzes object positions and provides contextual voice feedback.

Example outputs:

- The chair stands before us
- There is someone present on the left side
- The right side contains an obstacle which blocks the path

The system helps people with vision loss by giving them audio directions which describe what exists around them.

5.3 Voice Command Interface

Users can operate the system through voice commands which enable them to control system functions.

Example commands:

- Start detection
- Stop detection
- Snapshot
- Auto detect
- Switch camera
- Reminder
- Emergency Commands

Voice interaction allows visually impaired users to access the system because they can use their voice to operate it

5.4 Multi-Camera Support

NETRA-AI supports multiple camera sources:

- Laptop camera
- Mobile camera
- USB camera
- IP camera

Users can switch cameras through voice commands.

5.5 Emergency SOS Alert System

The emergency alert module allows users to send alerts in dangerous situations.

The system activates to send emergency messages together with GPS location links which it delivers to a specific guardian contact.

The system activates a siren which produces sound to alert people who exist in the surrounding area.

5.6 Reminder System

NETRA-AI includes a voice-controlled reminder module.

Example commands:

- “10 second baad yaad dilana”
- “1 minute reminder set karo”

The system shows a countdown timer which produces an audio warning when the reminder period reaches its end.

5.7 Memory Assistance Feature

The system stores previously detected objects and allows users to query past observations.

Example queries:

- “Pehle kya dekha?”
- “Yaha kya tha?”

This system provides visually impaired users with enhanced ability to sense their environment.

VI. IMPLEMENTATION

Backend
Python
Flask
YOLOv8
OpenCV

Frontend
HTML
CSS
JavaScript

The system performs image processing through YOLOv8 at its backend while its frontend accepts voice commands and manages user interface operations.

Figure 3: User Interface of NETRA-AI SYSTEM

VII. EXPERIMENTAL RESULTS

© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1715263

The system underwent evaluation through its operation in an indoor setting which used live video feed from a camera. The system captured continuous video streams through the camera while it used the YOLOv8 model to perform object detection on the recorded footage.

The system underwent testing with various typical items to evaluate its ability to identify objects. The system achieved precise identification of persons and chairs and tables and laptops. The object detection ability of the system shows its results in Table 1 which displays the system's detection results.

Detection Performance

Object	Detection Accuracy
Person	92%
Chair	88%
Table	86%
Laptop	84%

The system produced real-time voice feedback which began when it identified objects. This feature benefits visually impaired users by helping them recognize their surroundings.

The system operational test of the emergency SOS function succeeded when it sent an emergency alert together with location data to the guardian device.

The experimental results show that the proposed system is able to detect objects and generate real-time audio feedback to the visually impaired person.

VIII. COMPARISON WITH EXISTING SYSTEMS

System	Object Detection	Voice Guidance	SOS Alert	Reminder
Smart Glass	Yes	Yes	No	No
Mobile App	Yes	Yes	No	No
NETRA-AI	Yes	Yes	Yes	Yes

IX. LIMITATIONS

The system works well in most situations but it has some specific restrictions which affect its performance.

The system shows decreased operational performance when users operate it during dark room conditions.

The system will struggle to detect targets because it needs to identify multiple objects which appear together in scenes with many people.

Internet connection may be required for certain features

X. CONCLUSION

The paper introduced NETRA-AI which functions as an intelligent vision assistant that works in real-time to help blind users identify objects through vocal commands and object recognition technology. The system offers various assistive functions which include voice command capabilities and SOS alert system and reminder service and support for multiple camera systems.

The experimental evaluation showed that the system detects objects while providing immediate auditory feedback which enables blind users to achieve better safety and environmental awareness.

FUTURE SCOPE

The system requires smart glasses to connect with its main interface.

The system needs GPS-based navigation assistance to guide users through their routes.

The development of deep learning models requires ongoing enhancement.

The AI system requires processing capabilities which operate without internet access.

REFERENCES

[1] J. Redmon and his team introduced their unified real-time object detection system which they named "You Only Look Once" at the 2016 IEEE CVPR conference.

[2] A. Bochkovskiy and his team published their work "YOLOv4: Optimal Speed and Accuracy of Object Detection" in 2020.

[3] Ultralytics, “YOLOv8 Documentation,” 2023.

[4] S. Gupta and his team published their research on assistive vision systems for blind people in IEEE Access during 2021.

[5] R. Girshick presented his Fast R-CNN research at the IEEE ICCV conference which took