

VisionAI

GIVING SIGHT THROUGH SOUND

Bhoomika Babu

235890506

Nandini Patil

235890096

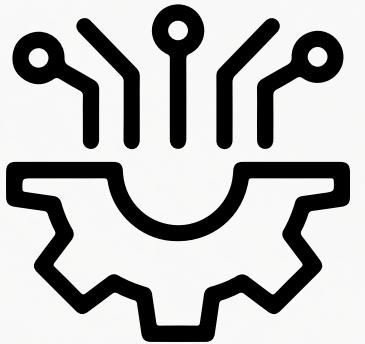
Table of Contents

1	Problems	2	Solutions
3	Methodology	4	Expected Output
5	Issues Faced	6	Future Aspects

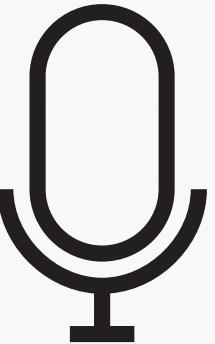
Problems

- ① Visually impaired individuals face difficulty identifying objects in their surroundings.
- ② Existing assistive tools (canes, GPS) lack real-time object awareness.
- ③ Many AI-based solutions are expensive or hardware-dependent.
- ④ Limited accessibility to affordable, software-based assistive systems.
- ⑤ Lack of systems that provide natural, context-aware audio feedback.

Solutions



Develop a real-time AI system that detects objects through a camera



It converts detection into natural voice feedback.



The software offers an affordable assistive foundation.

Methodology

Data Collection:

- Use COCO dataset for diverse object samples

Model Selection:

- Implement YOLOv5 or MobileNet-SSD for real-time detection.

Programming Environment:

- Python with libraries – OpenCV, PyTorch/TensorFlow.

Integration:

- Capture input through webcam.
- Detect and classify objects using the trained model.
- Convert text labels to speech using gTTS or pyttsx3.

Evaluation:

- Use metrics such as accuracy, precision, recall, and F1-score.

— Expected Output

A working software prototype that:

- Detects and classifies multiple objects in real time.
- Provides audio output describing detected items.
- Operates efficiently with minimal lag.
- Enhanced accessibility tool for visually impaired users.
- Basis for future wearable AI vision systems (e.g., smart glasses).



Issues Faced

Difficulty achieving real-time performance on limited hardware.

Challenges in integrating object detection and TTS smoothly.

Environmental factors (lighting, distance) affected accuracy.

High processing time for certain models.

Limited dataset scope for specific objects.

More

Future Aspects

1

Integration with
wearable devices
(e.g., AI-powered
glasses).

3

Deploy on mobile or
edge devices for
complete portability.

2

Expand dataset for
better recognition in
varied environments.

4

Include voice
command interaction
for user control.

The End

THANK YOU