HACETTEPE UNIVERSITY
Computer Engineering

Campus-Oriented Smart Assistant for the Visually Impaired

PA: Doç. Dr. Hacer Yalım Keleş
Ömer Faruk Horat // Doğukan Aytekin // Buğrahan Çıldır // Ecem Altınkeser

Introduction & Objective

Visually impaired individuals often struggle to navigate complex environments like university campuses. Existing tools lack real-time feedback and contextual awareness. This project aims to develop a mobile visual assistant that uses a smartphone camera and voice interaction to enhance mobility, safety, and communication by integrating object detection, scene understanding, and conversational AI.

You can try

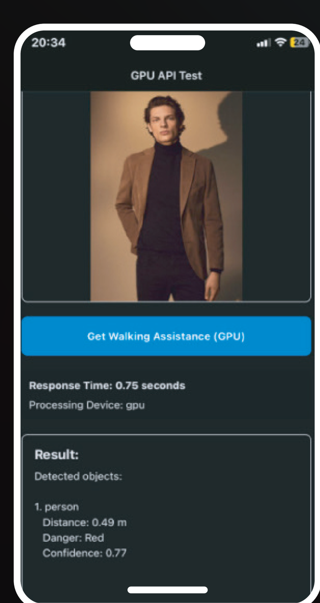
Website

Live Demo

1. Download CAMO Camera on your mobile app.
2. Scan QR code demonstrated in demo computer

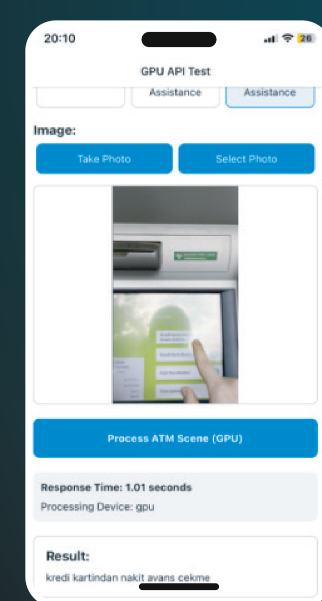
Attention

Detects and announces obstacles or objects (like stairs, doors, people) in real-time.



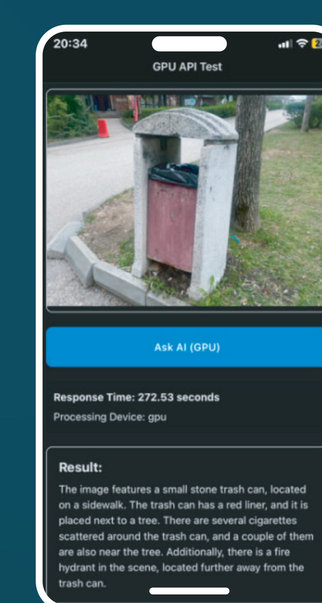
Apply

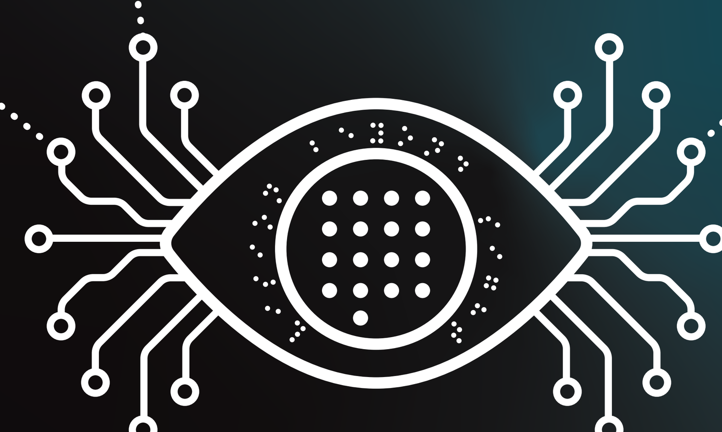
Understands structured scenes (like signs, devices) and allows the user to ask: "What's on this screen?"



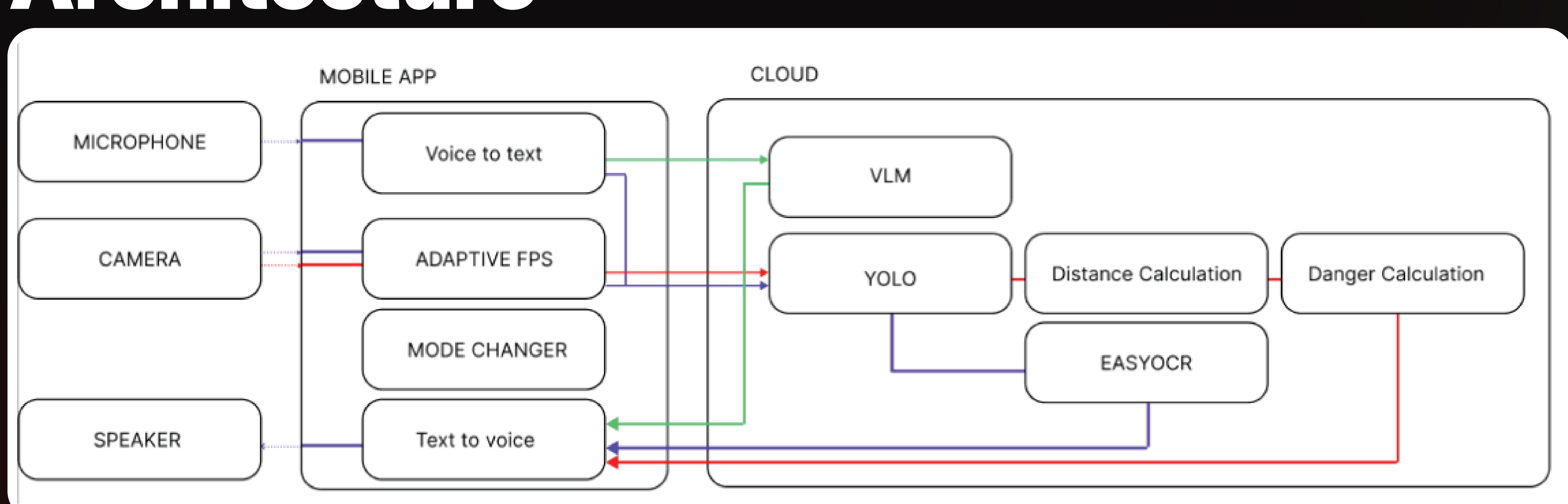
Communicate

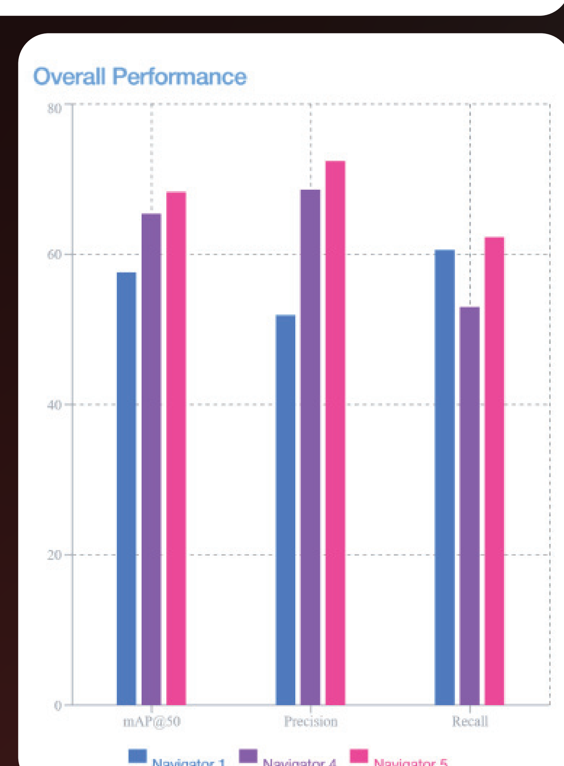
Engages in natural voice conversations for tasks, directions, or information.





Architecture





Dataset

Curated comprehensive dataset of urban environments, ATMs, obstacles, and navigation scenarios. Includes diverse lighting conditions, weather situations, and real-world scenarios to ensure robust model training and validation.

Yolo

Specialized YOLOv8 models trained for ATM detection and obstacle recognition. Features real-time object detection, segmentation, and tracking capabilities optimized for low-latency mobile deployment.

Visual LLM

Integration of LLaVA and GPT models for advanced scene understanding and natural language interaction. Enables detailed scene descriptions and contextual assistance for users.

Advanced Algorithms

Integration of EasyOCR for text recognition, sophisticated distance calculation, and danger assessment algorithms. Real-time processing for immediate hazard detection and safety alerts.

Mobile Application

User-friendly mobile interface with advanced voice interaction capabilities. Features intuitive navigation controls, real-time feedback, and customizable accessibility settings for enhanced user experience.

Cloud Infrastructure

Scalable cloud architecture for model deployment and real-time processing. Ensures high availability, low latency, and seamless updates while maintaining data security and privacy.

Results

The application was tested in simulated campus environments and real indoor spaces. It successfully:

- Detected and alerted users to objects with an average latency of < 300ms
- Provided accurate scene captions for over 85% of test images
- Handled voice interactions with GPT-based chat for custom queries like "Where is the nearest restroom?"

Analysis

- Performance: YOLOv8 achieved high FPS and reliable detection, though accuracy decreased in low-light conditions.
- Usability: The single-tap interaction model (tap-to-speak) was found intuitive, especially when paired with voice feedback.
- Limitations: The app depends on camera visibility and sufficient internet connectivity. Indoor localization was not implemented, limiting full navigation support.

Conclusion

This project demonstrates a feasible and effective approach to delivering real-time visual assistance to blind individuals using affordable, off-the-shelf mobile technologies. The integration of computer vision, language models, and voice interaction can significantly improve autonomy and confidence in unfamiliar environments such as university campuses. Future iterations will explore adding navigation capabilities, offline support, and wearable device integration.

- Redmon et al. (2016). YOLO: Real-Time Object Detection
- Li et al. (2022). BLIP: Bootstrapped Language-Image Pretraining
- OpenAI (2023). ChatGPT for Conversational Assistance
- Al-Razgan et al. (2020). Mobile Assistive Systems for the Blind
- Tapu et al. (2017). A Survey on Vision-Based Assistive Technologies