Adaptive Vision Assistant (AVA): A Smartglass-Integrated Assistive System for People With Visual Impairment

Yiran Zhou[1,2], Saad Alrazzi[3], Ava Chang[4], Hannah Boike[1], Ava Fischer[1], James Weiland[1,5]

Departments of [1]Biomedical Engineering, [2]Robotics, [3]Electrical & Computer Engineering, [4]Computer Science & Engineering, and [5]Ophthalmology & Visual Sciences, University of Michigan, Ann Arbor, MI

BIOMEDICAL ENGINEERING
UNIVERSITY OF MICHIGAN

Poster # 4478652

PHOTOGRAPHY ALLOWED

Introduction

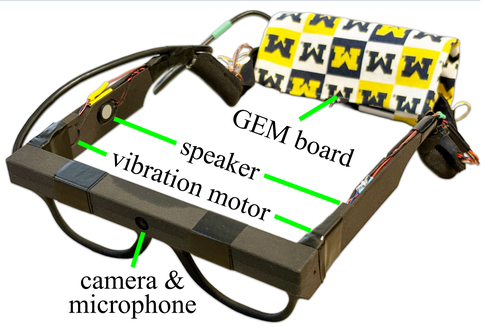
People with visual impairments often lack hands-free, real-time support for daily tasks. AVA is an integrated system designed to solve this by combining custom smartglasses with a powerful Android app.

AVA.glasses capture real-time visual input and deliver hands-free assistance through audio cues and directional vibration feedback, enabling independent mobility.

AVA.app processes all vision and voice models locally on-device. Users can simply issue voice commands to:

- Launch third-party apps (e.g., Google Maps, Be My Eyes)
- Detect doors, handles, and obstacles
- Answer visual questions about their surroundings

Hardware Design



AVA.glasses feature a front-facing camera for real-time visual input, two open-ear speakers for audio cues, and four motors for vibration feedback. The custom GEM board routes data and power between AVA.glasses and the Android phone running AVA.app.

Door Handle Detection

- **Goal**: Independent door access in various environments
- **Model**: A YOLOv8n[1] model trained on an open source dataset[2]
- **Classes**: Standard, cabinet, refrigerator doors, and their handles

Door with handle detected on right: "Right" voice instruction + right-arm vibration for precise navigation.

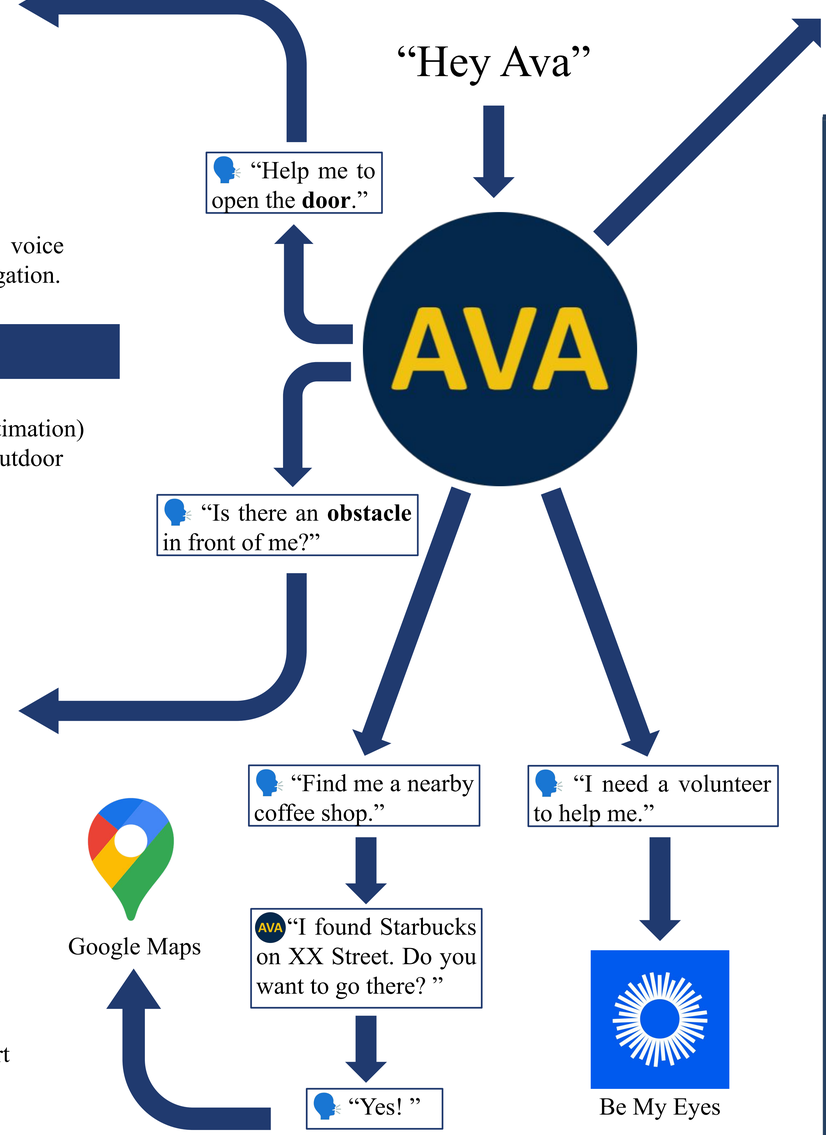
Obstacle Detection

- **Goal**: Obstacle avoidance in various environments
- **Model**: Depth Anything V2[3] (monocular depth estimation)
- **Adaptation**: Uses GPS to switch between indoor/outdoor

Obstacle detected on right: "Right" voice alert + right-arm vibration for hazard avoidance.

Voice Assistant

- **Goal**: Enables hands-free control of the app
- **Multimodal Activation Methods**:
  - Voice: "Hey Ava" detection via Picovoice Porcupine.
  - Manual triggers: in-app button | side button
  - Haptic gesture: device back-tap gesture
- **Speech Recognition**: Google Speech-to-Text API
- **Command Handling**:
  - Direct: Instant execution via keyword detection
  - Fallback: Gemma 4 E2B[4] (Small Language Model) for on-device intent classification and natural language interactions



Visual Question Answering (VQA)

- **Goal**: Answer specific, user-initiated questions about objects, text, or dynamic events within their field of view
- **Model**: Gemma 4 E2B (Vision-Language Model) for zero-shot object recognition and interactive spatial awareness

"Take a look and tell me what bus is coming."

AVA "This is a M-bus going to Bursley-Baits."

"Look at this and tell me what room is in front of me."

AVA "This is room 014-190."

Results

Internal testing demonstrated highly reliable performance:

- **Door Detection**: **90%** accuracy across diverse styles and backgrounds, operating at **~8 FPS** on-device.
- **Obstacle Detection**: 100% success rate in standard indoor environments, operating at **~1.5 FPS** on-device.
- **Voice Assistant**: **98%** wake-word detection accuracy and **96%** correct command interpretation.

Conclusion and Future Work

AVA integrates smartglass hardware with on-device computer vision and voice control to provide better assistance for individuals with visual impairments.

Future work:

- End-to-end navigation (Google Maps API)
- Hardware improvement (Meta AI Glasses)
- Expanded computer vision features
- Context-aware adaptation

Funding

MICHIGAN ENGINEERING UNIVERSITY OF MICHIGAN
MEDICAL SCHOOL UNIVERSITY OF MICHIGAN

NIH Supported by the NIH under Award #1R61EY037514

Coulter Translational Research Partnership Program
Rosalind and Brian Chamberlain Research Fund

References:

[1] Redmon et al., 2016 CVPR.
[3] Arduengo et al., 2021 Intelligent Service Robotics.
[2] Yang et al., 2024 NeurIPS.
[4] Google DeepMind, 2026 Gemma 4 Technical Report.