Smart obstacle & object detection glasses

Name: Omar Hussien Nady Mahmoud

Declaration of No Plagiarism

I hereby declare that this submitted report work is a result of my own efforts and I have not plagiarized any other person's work. I have provided all references of information that I have used and quoted in my work.

Name of Student: Omar Hussien Nady Mahmoud

Signature:

Student ID: 22510568

Abstract:

The Smart obstacle & object detection glasses project aims to help blind people in their daily lives by supporting both safe navigation and independent evaluation of food, in combination with cane or service pets. The system uses 3D-printed smart glasses with a Raspberry Pi camera, ultrasonic and IR sensors to capture the surrounding environment and estimate how dangerously close obstacles are to the user, while a gauntlet containing a Raspberry Pi 4B 8 GB performs all processing. Three YOLOv8n models are employed: one for pathways, one for obstacles, and one for fruit identification. A speaker module mounted on the glasses provides real-time audio feedback, giving simple guidance commands for navigation or clear messages about fruit safety, with the overall goal of offering a low-cost, energy-efficient and sustainable smart glasses solution for blind users.

Acknowledgement

I would like to express my deepest gratitude and sincere appreciation to **Dr. Eid Emary**, who has been a mentor, and a great teacher to me throughout the past four years, long before being an academic supervisor. Through his extensive knowledge, rich experience, and dedication to teaching, he has continuously guided me, shaped my way of thinking, and helped build a strong academic and practical foundation. What I have learned from him over these years will remain a cornerstone in my professional and personal journey.

I would also like to extend my sincere thanks to my mentor, **Eng. Ibrahim Adel**, for his continuous encouragement, motivation, and willingness to teach and guide me. His support, trust, and practical insights played a significant role in developing my skills and boosting my confidence, while helping me bridge the gap between academic knowledge and real-world applications.

I am deeply grateful to my family for their endless support, patience, and constant belief in me throughout my educational journey. Their encouragement has been my greatest source of strength.

Table of contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1. Background

In modern society, the number of blind people is increasing in a trend that is being observed worldwide. According to IAPB vision Atlas in Figure 1, 290,000,000 people were blind in 2020 estimating to reach about seven hundred million people in 2050 [6][7]. As the number of blind people increases the demand for assistive technologies rises. Kasowski et al. (2023) conducted a detailed review of assistive technologies for blind people, focusing on the role of AI technology [1]. Lin. Y et al. (2019) has shown that deep learning based assistive systems have shown promising results in real world applications [2]. While Manjari et al. highlighted the rapid progress in this field using a detailed survey of diverse groups of assistive technologies for blind people [3].

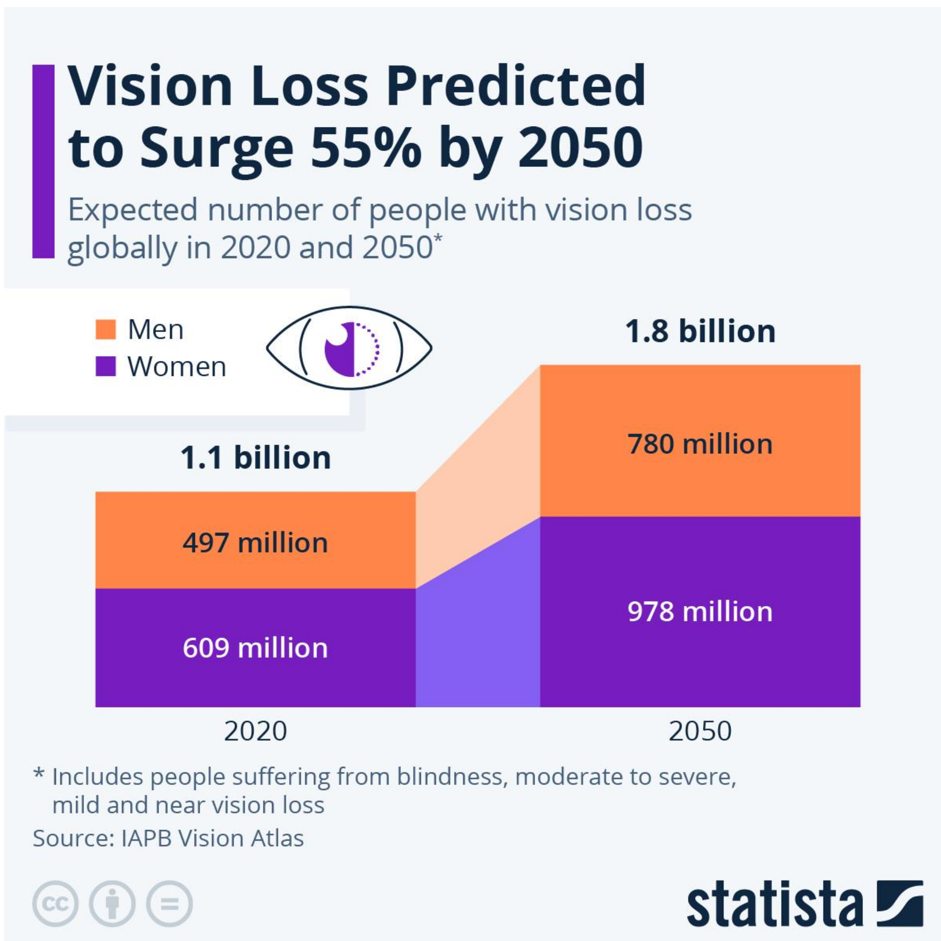

Figure 1 IAPB VISION ATLAS prediction on blind people

1.2. Problem with assistive technology

Even though there’s a great increase in assistive technology enhancements progress. Up till now, many individuals who are blind still prefer to use old assistive technologies like canes and service animals to guide them through the day in navigation. Usually, the main reason these individuals go to such old solutions is that they tend to have several advantages over newer technological assistive technologies like:

- Newer assistive technologies tend to be more expensive than old ones.
- Newer assistive technologies do not provide immediate feedback for dangerous situations, and if they do it can be overwhelming to the user.
- Newer assistive technologies are usually mildly heavy but in the long term it becomes extremely painful to carry for daily use.
- Newer assistive technologies usually do not last the whole day because of the intensive AI processing that happens to them, making the battery management system a critical factor on such solutions.
- Newer assistive technologies are impacted by environmental conditions like brightness, weather, and air quality.

These problems are currently being addressed in many newer research projects such as Jeong at al. (2025) article on using XR glasses to solve the weight problem and the slow or overwhelming feedback to the user [4].

1.3. Objectives of Smart glasses project

This project focuses on addressing some of these problems. Smart glasses address several problems, which are:

- Battery life efficiency for daily usage. Focusing on giving at most 6 hours of usage with a charging solution for extending usage to more than 6 hours.
- Weight distribution for daily usage. The system will be composed of two devices, one used for processing (gauntlet), and another used for video capturing and feedback (smart glasses).
- Cost cuts providing almost identical processing efficiency and precision as Jeong et al. (2025) model.

Such objectives might be constrained later due to budget limitations or knowledge limitations; however, this project aims to at least address two problems and solve them.

1.4. National legal obligations followed

This project complies and follows the following Egyptian laws governing assistive technologies for disabled people. Law No.10 of 2018 on the rights of persons with disabilities is the primary national framework governing the rights and protection of people with disabilities in Egypt. The law obligates the state and institutions to ensure accessibility, provide assistive technologies, and support equal participation of people with visual impairments. Hence, it regulates that assistive technologies should always be affordable and accessible to people with disabilities. Also, it encourages the enhancement of mobility, independence, and safety for people with disabilities [5]. Egyptian standard No.7866/2021 is a national standard that provides engineering requirements and testing methods for assistive products designed for individuals with disabilities. Even though, it's labeled "Not Obligatory", it's the official technical reference for safety, quality, and performance of assistive technology in Egypt. From ISO 21856 (on which Egyptian 7866/2021 is based) assistive products must satisfy a set of general requirements and pass related test methods [8].

Requirements	Definition
Mechanical Safety and structural integrity	The device must be mechanically robust: no sharp edges, no unstable parts. For wearable glasses: frame/shell strength, no breakage risk, secure fittings, safe under normal and reasonable misuse [].
Stability and Support (optional)	If the assistive product has components that could bear load, they must remain stable and not collapse under expected use [].
Ergonomics and comfort	Control of forces, pressure points, or stress on the user must be within safe limits to avoid injury [].
Materials safety and sustainability	Use of materials that are safe for prolonged contact with human body, avoidance of small loose parts that may be swallowed, insurance parts are secured
Electrical safety	If the device contains electronics, sensors, power sources, etc. then basic electrical safety must be ensured [].
Functional Performance and reliability	Device must perform its intended assistive function reliably: inputs, controls, outputs must work as planned, fail-safe behavior must be present for misuse, and instructions must be provided [].
Labeling/user manual	The manufacturer must supply sufficient instructions to the user manual like instructions for safe use, maintenance, warnings, limitations, storage conditions if needed, cleaning instructions, and any safety warnings [].
Packaging and transport safety	If the product is shipped, packaging should protect it during transport. Packaging and labeling should also be

	safe (no hazards). Instructions for safe transport, storage, and handling should be provided [].
Test/inspection documentation and traceability	For each device produced, there should be test reports or trial data showing compliance with safety and performance tests [].

Table 1 Egyptian 7866/2021 standard and regulations for product

By following the relevant Egyptian laws and recognized assistive technology standards, this project ensures that the proposed smart glasses are developed with full regard to user safety, ethical responsibility, and national regulatory expectations. Such guidelines form the foundation of a reliable, practical, and socially responsible solution intended to support individuals with visual impairments in their daily navigation and independence.

1.5. Project plan

Table 2 Project plan details

Start Date	End Date	Tasks
16-10	15-11	Research: Literature review on obstacle detection & navigation systems
15-11	30-11	Research: User needs analysis for blind navigation
30-11	16-12	Research: System requirements & architecture design
16-12	30-12	Sensor selection (LiDAR / Ultrasonic / IMU / audio output)
30-12	15-1	Microcontroller selection & hardware design planning
15-1	1-2	Hardware assembly & wiring setup
1-2	15-2	Sensor testing & calibration
15-2	1-3	Develop obstacle detection algorithm
1-3	15-3	Develop navigation logic
15-3	30-3	Develop audio/haptic feedback system
30-3	15-4	Full hardware–software integration
15-4	30-4	Smart glasses frame design + 3D printing
30-4	15-5	Indoor/outdoor testing & optimization
15-5	30-5	Final report, presentation, poster, demo

1.6. *Gantt chart*

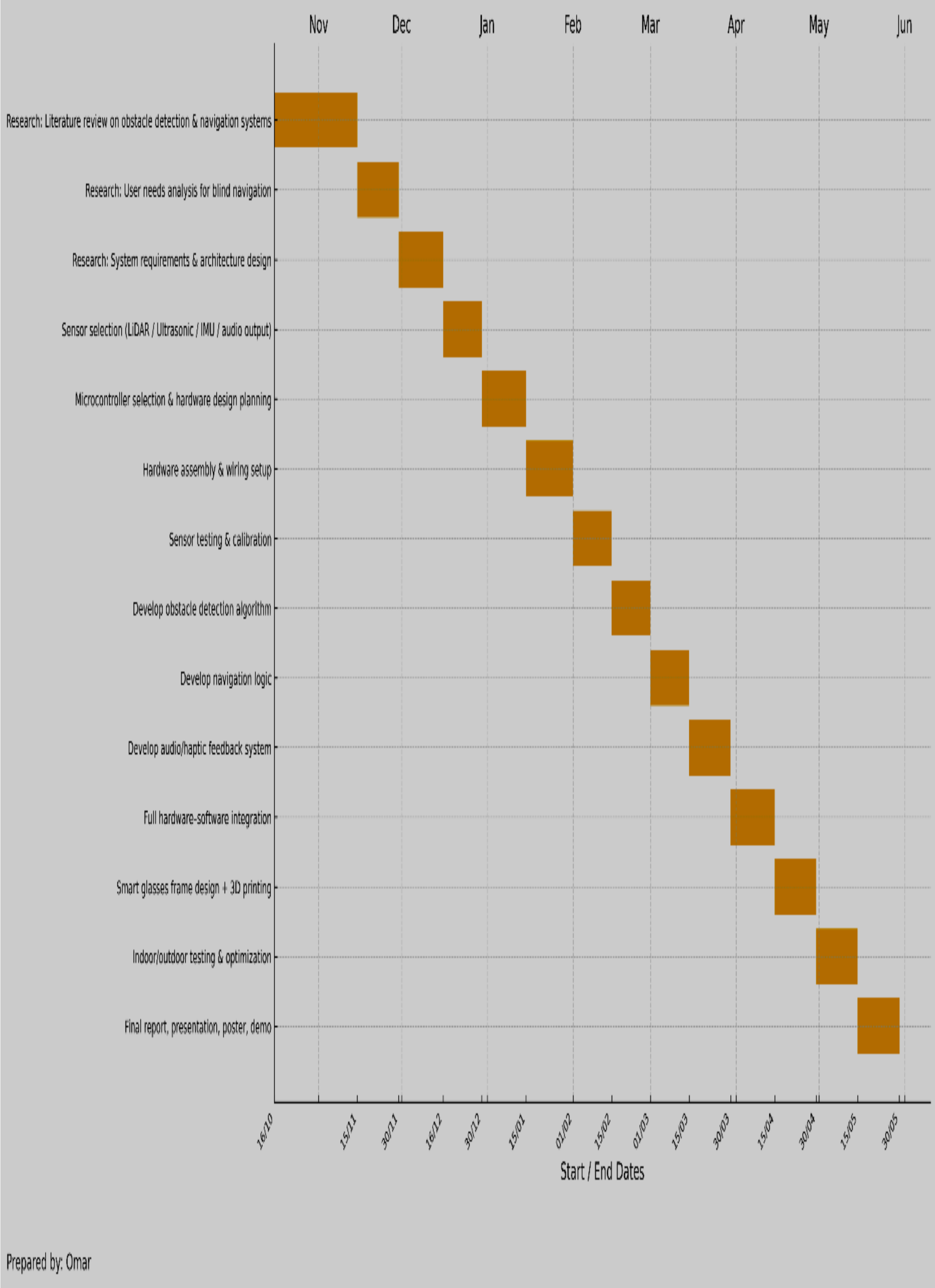

Figure 2 Gantt chart

Chapter 2: Literature Review

2.1. DEEP-SEE system

Tapu et al. (2013) [] have researched and created a device that consists of a wearable camera with computer vision technology that analyzes the user's surroundings in real time and provides voice guidance on the types and locations of obstacles.

Figure 3 DEEP-SEE system

In Figure 3 DEEP-SEE system Tapu et al (2013) is used for object detection, the system builds on an algorithm like YOLO. However DEEP-SEE combines a CNN algorithm with temporal tracking. For tracking, they used two CNNs trained offline, one based on motion pattern (to handle objects in movement) while the other one is based on visual appearances (to reidentify objects based on their looks). Hence, alternating between motion based tracking and visual similarity-based prediction.

2.2. *Multimodal walking assistance system*

Prattico et al. (2020) created a walking assistive system, shown in Figure 4, that combines cameras, LIDAR, and Ultrasonic Sensors [].

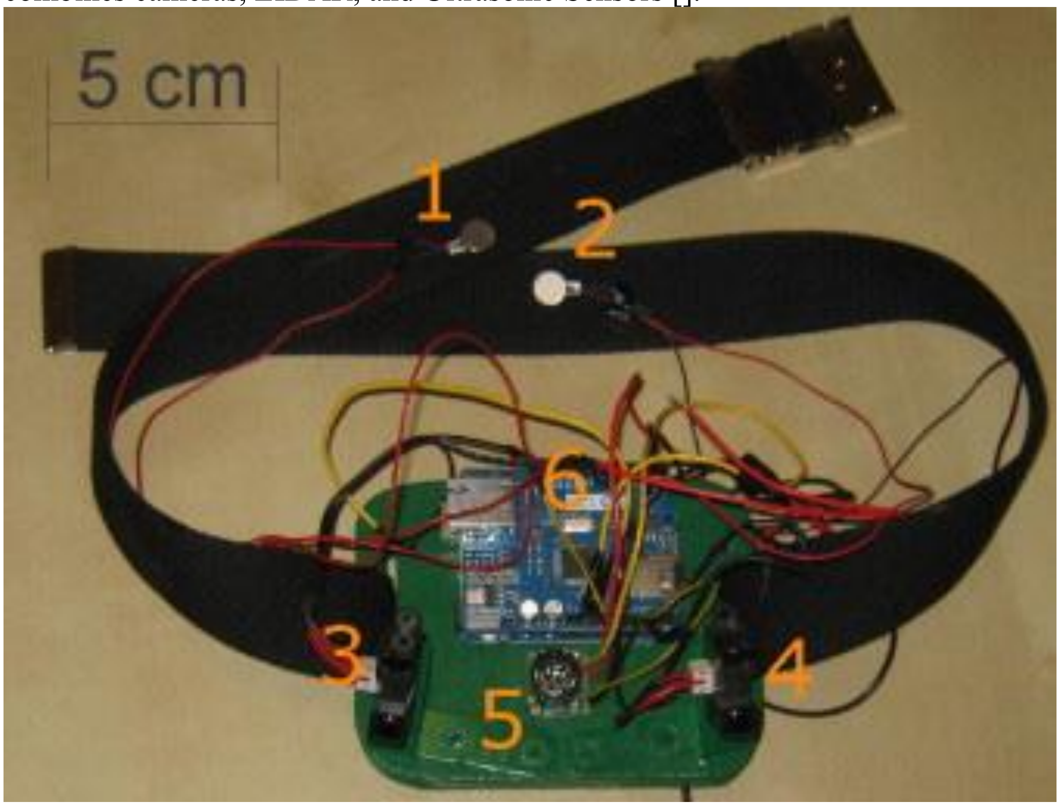

Figure 4 Walking assistive system created by Pattrico et al.

The system used two IR sensors, one ultrasonic sensor, and microcontroller that processes inputs and performs decision logic triggering vibration motors when an obstacle is within threshold. In tests the device was able to intercept obstacles and provide timely warnings to the user. However, the device only detects obstacles in the forward direction. Also, it doesn't detect ground level hazards like stairs, protrusions, or holes.

2.3. *AR glasses*

Hicks et al. (2013) created a depth sensing smart glasses, shown in Figure 5, that focus on basic depth sensing data []. Many developments had been adapted to this concept such as Huang et al. (2018) proposing a walking assistance system based on MR headsets that combine computer vision and spatial mapping techniques to give real time environments perception information []. Later, Maidenbaum et al. (2021) who used a HoloLens, shown in Figure 6, to combine 3d spatial information with voice guidance.

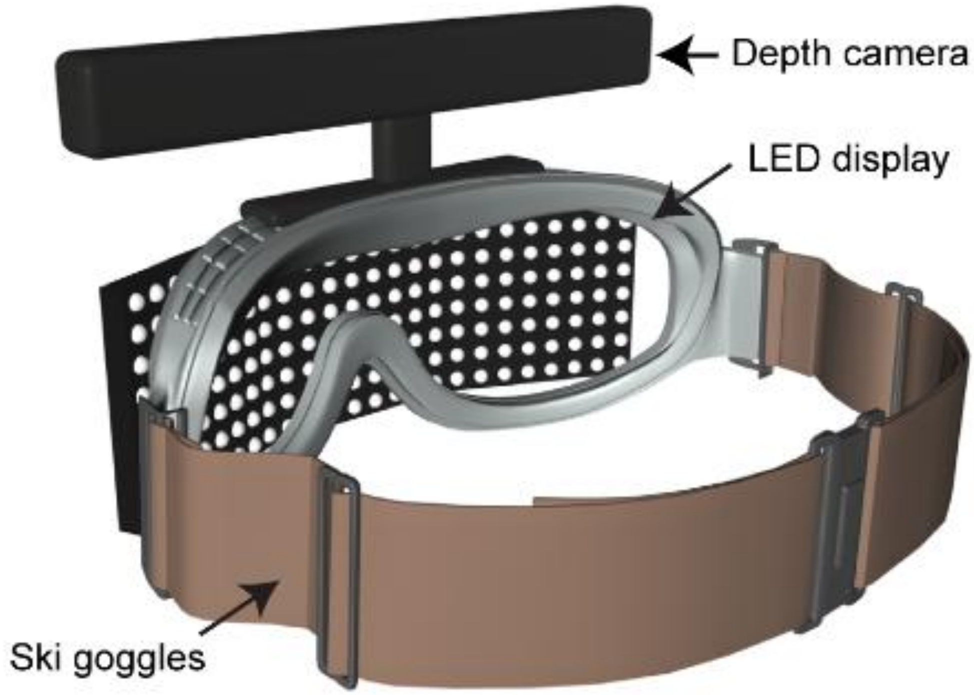

Figure 5 AR device developed by Hicks et al. (2013)

Figure 6 HOLOLENS

2.4. XR glasses

Jeong et al (2025) decided to use an affordable XR glasses that combine between augmented reality and virtual reality. Combining with it YOLOv8n Model for real time object detection. The glasses that were used are Xreal light smart glasses, released by Xreal light, shown in Figure 7. The glasses are connected to a smartphone (galaxy s22+), shown in Figure 8, that contains an unreal game engine application that uses a combination of three different AI models for processing walkways and pathways, bus stops and train stations, obstacles on the pathway to avoid during navigation [].

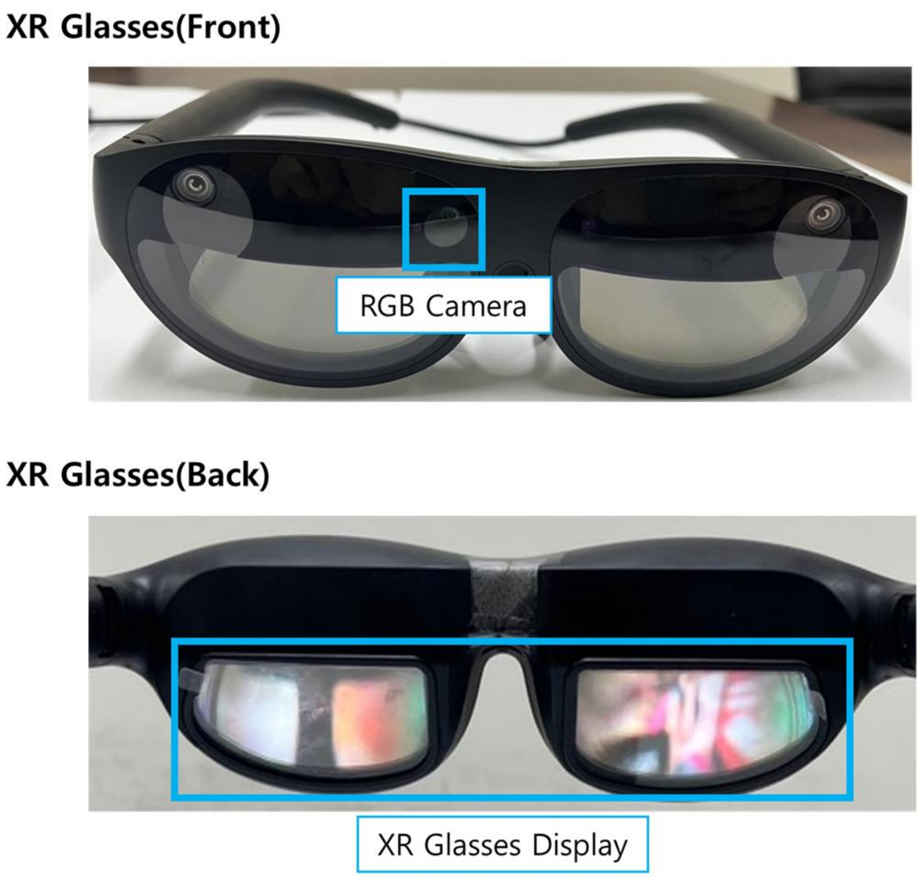

Figure 7 Xreal Light smart glasses

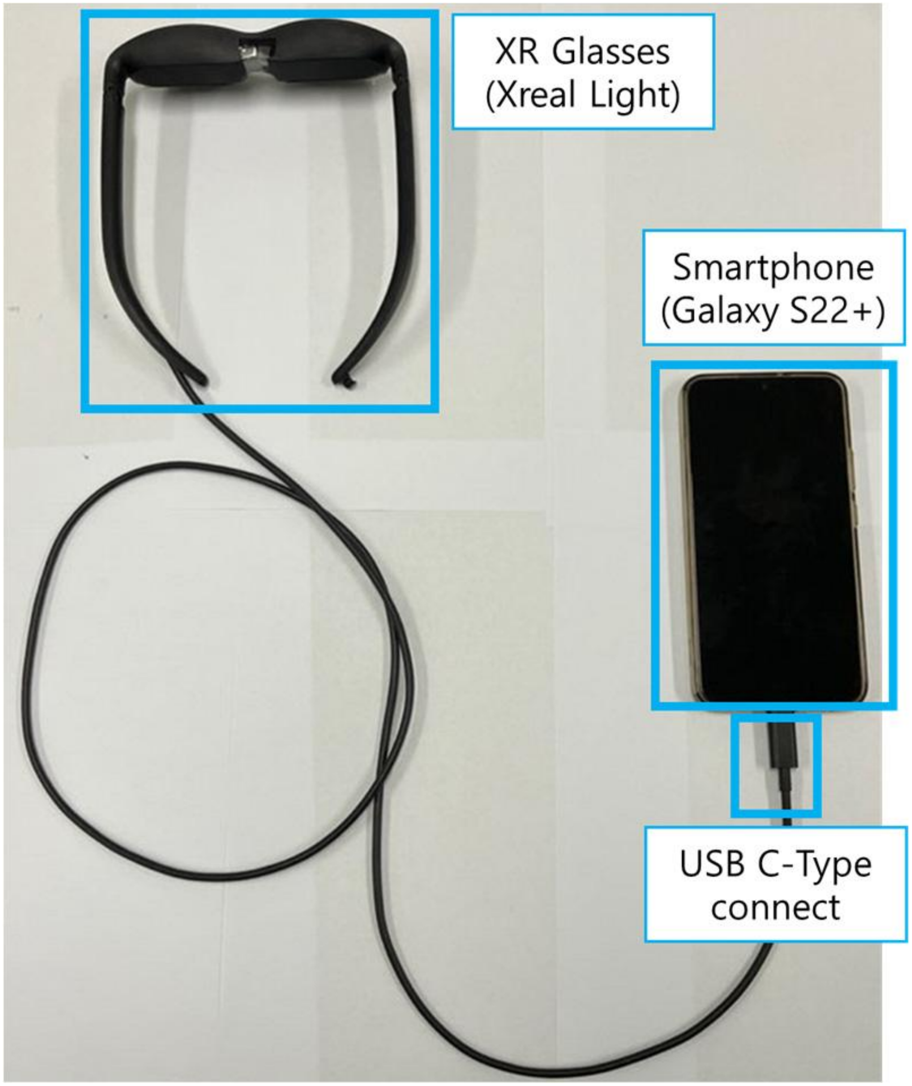

Figure 8 Xreal light connected to smartphone

However, the paper showed that further research needs to be done focusing on solving the problem with the poor performance of AI in obstacle detection in night time, also, the overwhelming alerts received that can paralyze a blind person in navigating safely, the battery only lasting for two hours which isn't acceptable for daily use, and the system being a bit heavy weighing about 75 grams that can cause annoyance to the user on prolonged uses.

2.5. YOLOv8n

Jeong et al. divided the AI model into three distinctive models (Walkway recognition model, transportation facility detection model, and obstacle detection model), shown in Figure 9, that works in parallel to each other to limit the processing power needed on the phone. The concept here is that camera captures 60 frames per second. In each 20 frames a model will process those frames to identify objects, walkways, or stations then integrate everything in the end to provide results for audio feedback all happening under a 500ms constrain from end-to-end latency requirements.

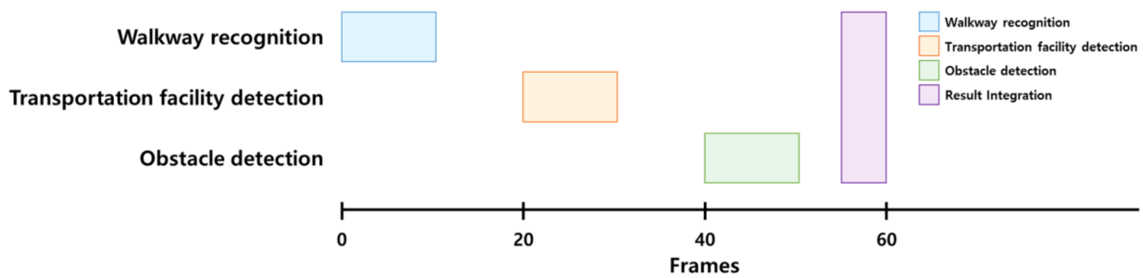

Figure 9 Model execution and integration timeline (60 frames) by Jeong et al.

Each model was trained separately. For walkway recognition model crosswalks pictures and road pictures were fed to it, shown in Table 3, while Transportation model was trained on pictures of bus stops and train stations, shown in Table 4, and for obstacle detection model, it was trained on pictures of cars, trucks, busses, people, movable signage, motorcycles, traffic lights, bicycles, bollards, which is shown in Table 5.

Table 3 Data distribution of the path identification model by Jeong et al.

Data	Image Count	Type	Label Count
Training data	11,071	Crosswalk Road	2377 17,642
Validation data	3163	Crosswalk Road	736 5446

Table 4 Data distribution of the public transportation recognition model by Jeong at al.

Data	Image Count	Type	Label Count
Training data	7144	Bus Stop Subway station exit	7407 354
Validation data	2032	Bus Stop Subway station exit	1890 257

Table 5 Data distribution of the obstacle detection model by Jeong et al.

Data	Image Count	Type	Label Count
Training data	296,949	Car Truck Bus Person Movable signage Motorcycle Traffic lights Bicycle Bollard	296,331 55,018 18,047 163,431 51,473 24,253 60,479 19,965 113,982
Validation data	82,826	Car Truck Bus Person Movable signage Motorcycle	98,701 19,040 6177 46,495 14,147 7358

		Traffic lights Bicycle Bollard	12,724 5433 32,063

The training parameters were united during the training of all models as it appears the dataset for each model was sufficient for YOLO AI model. The training parameters are provided in Table 6.

Table 6 Training parameters of the obstacle detection model by Jeong at al.

Item	Value
Resolution	640 x 640
Epoch	100
Batch size	16
Learning rate	0.01
Momentum	0.937
Weight decay	0.0005
FLOPS	4.2 GFLOPS

After training the three distinctive models to their sufficient datasets from "Walking on the Sidewalk" (https://aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=189) and Jeong et al.'s own dataset of bus stops pictures and train stations. Each model gave great precision, accuracy and recall ratings, shown in Tables 7 & 8 & 9. Also, each model was able to identify objects that they were trained on with an acceptable confidence level, as shown in Figure 10.

Table 7 Performance of the path identification model by Jeong et al.

Object Class	Precision	Recall	F1-Score	AP@0.5	AP@[0.5:0.95]
Crosswalk	0.749	0.655	0.698	0.689	0.55
Road	0.866	0.82	0.842	0.885	0.767

Table 8 Object detection performance of the traffic facility recognition model by Jeong et al.

Object Class	Precision	Recall	F1-Score	AP@0.5	AP@[0.5:0.95]
Bus Stop	0.749	0.655	0.698	0.689	0.55
Subway Station exit	0.866	0.82	0.842	0.885	0.767

Table 9 Object detection performance of the obstacle detection model by Jeong et al.

Object Class	Precision	Recall	F1-Score	AP@0.5	AP@[0.5:0.95]
Car	0.901	0.914	0.907	0.966	0.815
Truck	0.893	0.802	0.845	0.903	0.714

Bus	0.867	0.807	0.836	0.896	0.72
Person	0.899	0.846	0.871	0.927	0.672
Movable signage	0.835	0.738	0.783	0.842	0.647
Motorcycle	0.929	0.848	0.886	0.937	0.644
Traffic light	0.835	0.759	0.795	0.851	0.565
Bicycle	0.862	0.812	0.836	0.893	0.622
Bollard	0.855	0.75	0.799	0.843	0.56

(a) (b)

(c) (d)

(e) (f)

(g) (h)

(i) (j)

(k) (l)

Figure 10 Example of Jeong et al. three models identifying objects with confidence levels indicated (a) original crosswalk picture (b) identified crosswalk picture (c)original crosswalk with a passenger picture (d)identified crosswalk with a passenger picture (e)original train station (f)identified train station (g)original underground train station (h)identified underground train station (i)original car (j)identified car (k)original bicycle (l)identified bicycle

Jeong et al. research has proven YOLOv8n reliability and lightweight as each model during processing consumed about 30MB and a total of 80 MB for the integrated model, which proves how great YOLOv8n model in identification base small amounts of datasets and providing great real time object classification. However, the dataset lacked mostly other environmental conditions which affected the classification of objects during nighttime, foggy weather conditions, and rainy seasons. Smart glasses project aims to address such bottlenecks, making the identification and classification process more reliable by using bigger datasets with different conditions.

2.6. Literature review conclusion

To conclude it’s noticeable that significant progress in both hardware and software during the last ten years has been presented through assistive technology, in which, now devices have become more lightweight and more efficient, as well as the AI algorithms and models have become more efficient and more lightweight, helping in lowering the power consumption needed for assistive technologies. The smart glasses project aims to collectively and critically select the best solution for providing a lightweight system that can last for a longer period and not overwhelm the users on their daily use by using YOLOv8n model, combined with custom designed glasses that can replicate the effects of an AR or XR glasses in spatial awareness. Hence, providing enough data for helping blind users navigate safely and independently.

Chapter 3: Requirements and Analysis

3.1. *Overview of the system*

The suggested system is a smart obstacle and object detection system composed of 3d printed wearable glasses with a camera and audio output, and a gauntlet hosting raspberry pi 4B for real time processing. The system assists blind users in navigating through their daily lives by detecting obstacles, estimating distances, detecting safe pathways, and providing audible feedback to the safest pathway for them using YOLOv8n model, while aiming to be affordable, lightweight, and sustainable for daily use.

3.2. *Functional Requirements*

The system shall:

- Capture real time videos from the camera mounted on the smart glasses.
- Detect obstacles and relevant objects in the user's surroundings using a YOLOv8n model running on raspberry pi gauntlet.
- Estimate the distance to detected obstacles using ultrasonic and IR sensors.
- Determine safe walkable pathways and decide navigation directions for the user.
- Provide audio feedback through a speaker on the glasses to warn about obstacles and indicate safe directions.
- Generate immediate alerts for dangerous obstacles that are too close to the user.
- Support continuous operation for at least 6 hours of typical use per day.
- Classify detected objects into categories (pathway model, obstacle model, Fruit model) to support different audio messages.
- Use AI output (bounding boxes and classes) combined with distance sensor data to determine whether an object is dangerously close and requires immediate alerting to the user or not.
- Use AI detections to infer at least one safe direction (left, right, forward, stop).

3.3. *Non-functional requirements*

3.3.1. Performance requirements

The system shall:

- Process camera frames and update obstacle information in real time with a maximum end-to-end delay of 500 ms.

- The audio feedback shall be issued within 100 ms after detecting a dangerous obstacle.
- Follow the same logic of Jeong et al. (2025), in processing each AI model (pathway, obstacle, and fruit detection) logic by dividing the total time frames received in one second for each model. In other words, the system will process the pathway identification model from the first frame to 20, then the model for obstacle detection will be used from 20 frames to 40 frames, and the fruit model with be processed from 40 frames to 60 frames.

3.3.2. Reliability

The system shall:

- Operate for at least 6 hours without crashing during normal use.
- Operate in dimmed environments without misidentifying objects.
- Not overwhelm the user in alerting them from obstacles.

3.3.3. Usability

The system shall:

- Allow a new blind user to learn the basic audio feedback signals after no more than 1 hour of training.
- The audio messages shall be concise and non-overwhelming so the user can still make safe navigation decisions.
- Contain audible alerting outputs when the user needs to find either the gauntlet or the glasses.

3.3.4. Security/Safety

- All exposed surfaces of the glasses and gauntlet shall be free of sharp edges and unsafe protrusions, following the Egyptian standard 7866/2021 for assistive technologies.
- The electrical components shall be protected to avoid short circuits, overheating, uneven discharge, or uneven charging during normal and reasonable misuse.
- The material used shall not be toxic for long-term contact with human skin.

3.3.5. Maintainability

- The software shall be modular to allow replacing the object detection model without changing the whole system.
- The system shall run on Raspberry pi 4B 8GB with Linux OS, YOLO and OpenCV (if needed) libraries without including additional hardware upgrades.

3.4. Domain Requirements

Based on Egyptian disability law and assistive standard:

- The system shall support blind and visually impaired users in Egypt in accordance with Law No. 10 of 2018 on the rights of persons with disabilities, promoting independent and safe mobility.
- The system shall be designed to remain affordable for users, aligning with national obligations to provide accessible assistive technologies.
- The system shall follow the general safety and performance principles of Egyptian standard 7866/2021 for assistive products.

3.5. Interface Requirements

3.5.1. User-interface (UI)

The system shall:

- Use audio messages and simple sound patterns as the primary user interface for blind users.
- Provide different tones or phrases to distinguish between obstacle warning and safe direction messages.

3.5.2. Hardware interface

The system shall:

- Contain buttons with braille embedding to for changing modes and for controlling the volume of the alerting.
- Contain a microphone (optional) for inputs from users to provide ease of use in changing modes.

3.6. AI requirements

The AI model shall:

- Classify detected objects into three categories (pathways, obstacles, fruits) to safely guide the user in their daily lives and help them in avoiding rotten fruits.
- Combine camera detections with distance readings from IR sensors and Ultrasonic to estimate how dangerously near each obstacle from the user.
- Process video frames in a similar behavior to Jeong et al. by diving the camera feed rate by three giving each model about 20 frames to process.
- Model size shall be small enough with enough input resolution to save heavy loading on the CPU and memory of Raspberry PI 4. Allowing for a longer battery life duration of at least 6 hours.

- Provide a structured output (object class, confidence level, relative location) that the decision module uses it to generate audio feedback to guide the user.

3.7. Software and Hardware Requirements

3.7.1. Hardware Requirements

The system shall use:

- Raspberry pi 4b 8gb.
- PI cam.
- Lithium Battery.
- Battery charger.
- Speakers.
- Microphone.
- Protection circuits.
- Ultrasonic.
- IR.
- 3D printed casing for gauntlet.
- 3D printed glasses.

3.7.2. Software Requirements

The system shall use:

- Linux OS for saving processing power.
- YOLOv8n detection model.
- Walking on the sidewalk dataset.
- Fruit Computer Vision dataset from universe.
- Failsafe module for faulty sensor readings.
- Speech to text module for converting input commands by the user (optional).
- Text to speech module for converting ai feedback during navigation or fruit identification.

3.8. Diagrams

3.8.1. Use-case Diagram

The use case diagram illustrates the main interactions between the blind user and the smart obstacle and object detection glasses system. The primary actor, the blind user, starts and ends navigation sessions, while the system captures the environment, detects obstacles and pathways, and delivers audio guidance, as shown Figure 11 Use-case, also allows the blind user to change modes for detecting fruits.

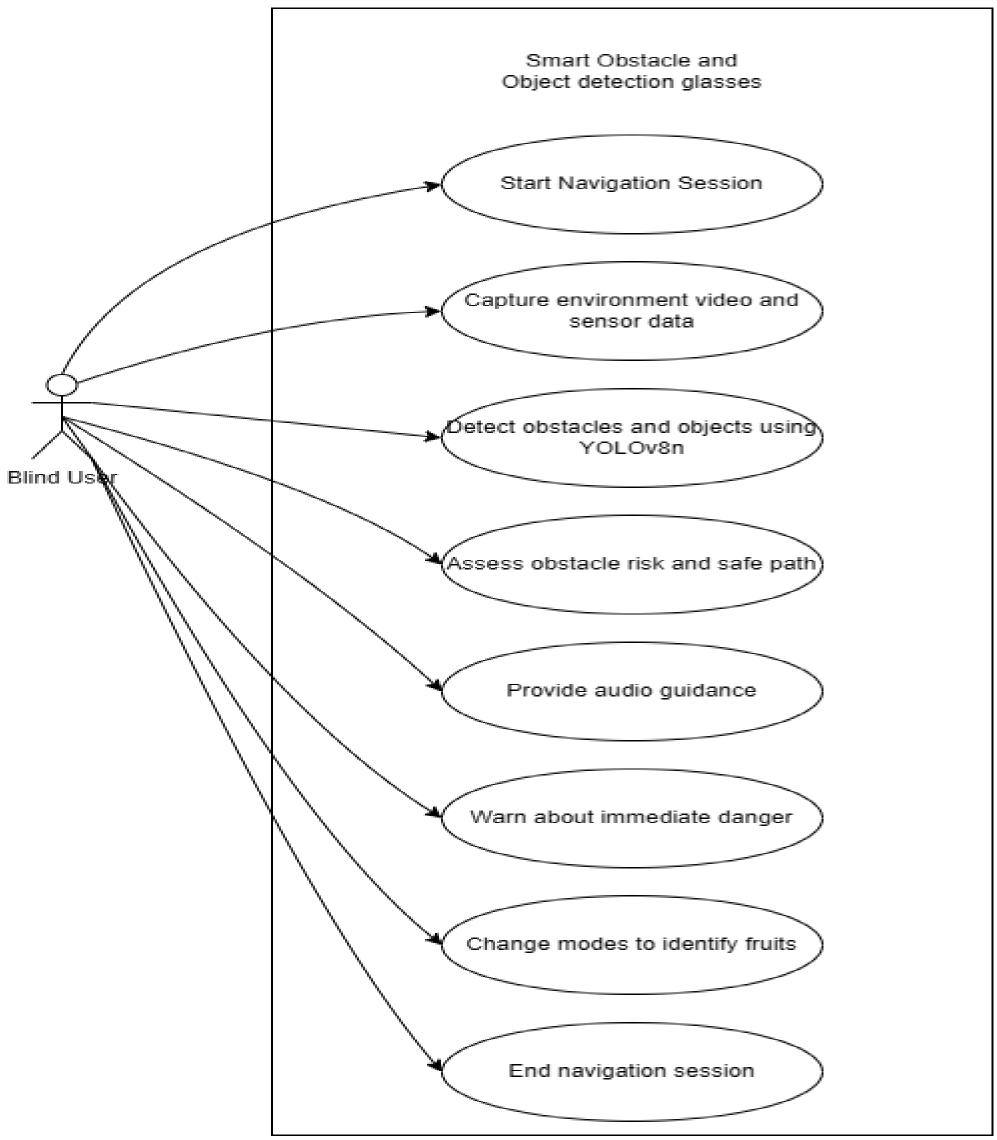

Figure 11 Use-case

3.8.2. Sequence Diagram

The sequence diagram models the real-time interaction flow when a blind user walks with the system and receives navigation assistance. The smart glasses continuously send camera frames to the Raspberry Pi gauntlet, while the sensor module provides ultrasonic and IR distance readings. The YOLOv8n detection module identifies relevant objects, the path and risk assessment module fuses detections with distance data to decide whether a dangerous obstacle exists, and the audio feedback module plays either a subtle confirmation or an urgent warning so the user can adjust their path.

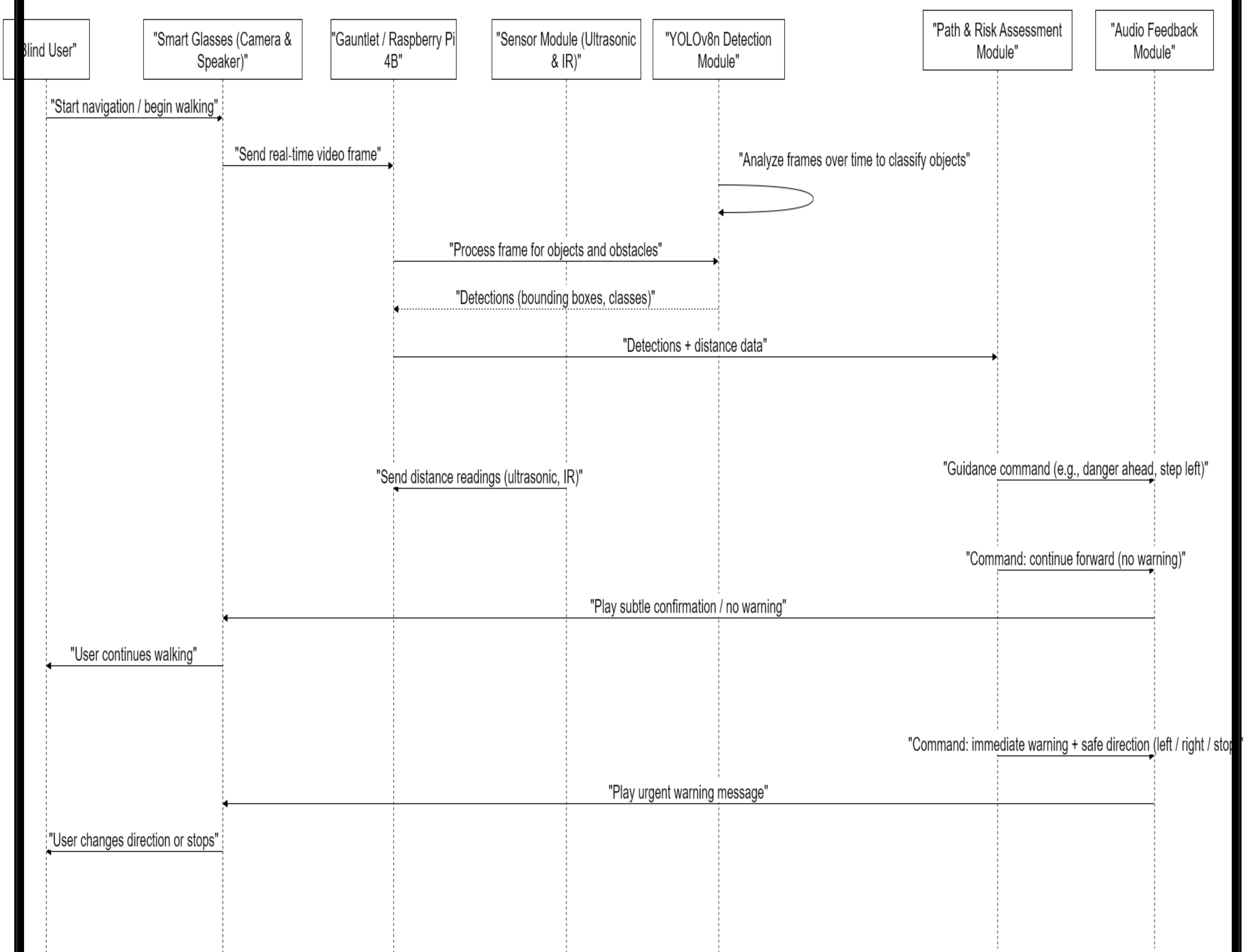

Figure 12 Sequence Diagram

3.8.3. Activity Diagram

The activity diagram describes how the smart glasses operate in two modes: navigation and fruit detection. After the user wears the glasses and gauntlet and powers on the system, the camera, ultrasonic and IR sensors, and YOLOv8n AI modules are initialized, then the user selects a mode. In navigation mode, the system repeatedly captures video frames and sensor data, runs the YOLOv8n navigation models to detect pathways, vehicles and obstacles, fuses detections with distance readings, and decides whether any obstacle is dangerous. If no danger is found it optionally confirms and loops, while if a dangerous obstacle is detected it determines a safe direction, generates and plays an audio warning, and the user adjusts the walking path before the loop continues until the user chooses to stop. In fruit-detection mode, the system captures a close-up image of a fruit, runs a YOLOv8n fruit model, decides whether the fruit appears safe or unsafe, and plays the appropriate message; the user can then choose to analyze another fruit, switch back to navigation mode, or power off the system, which leads to the final node.

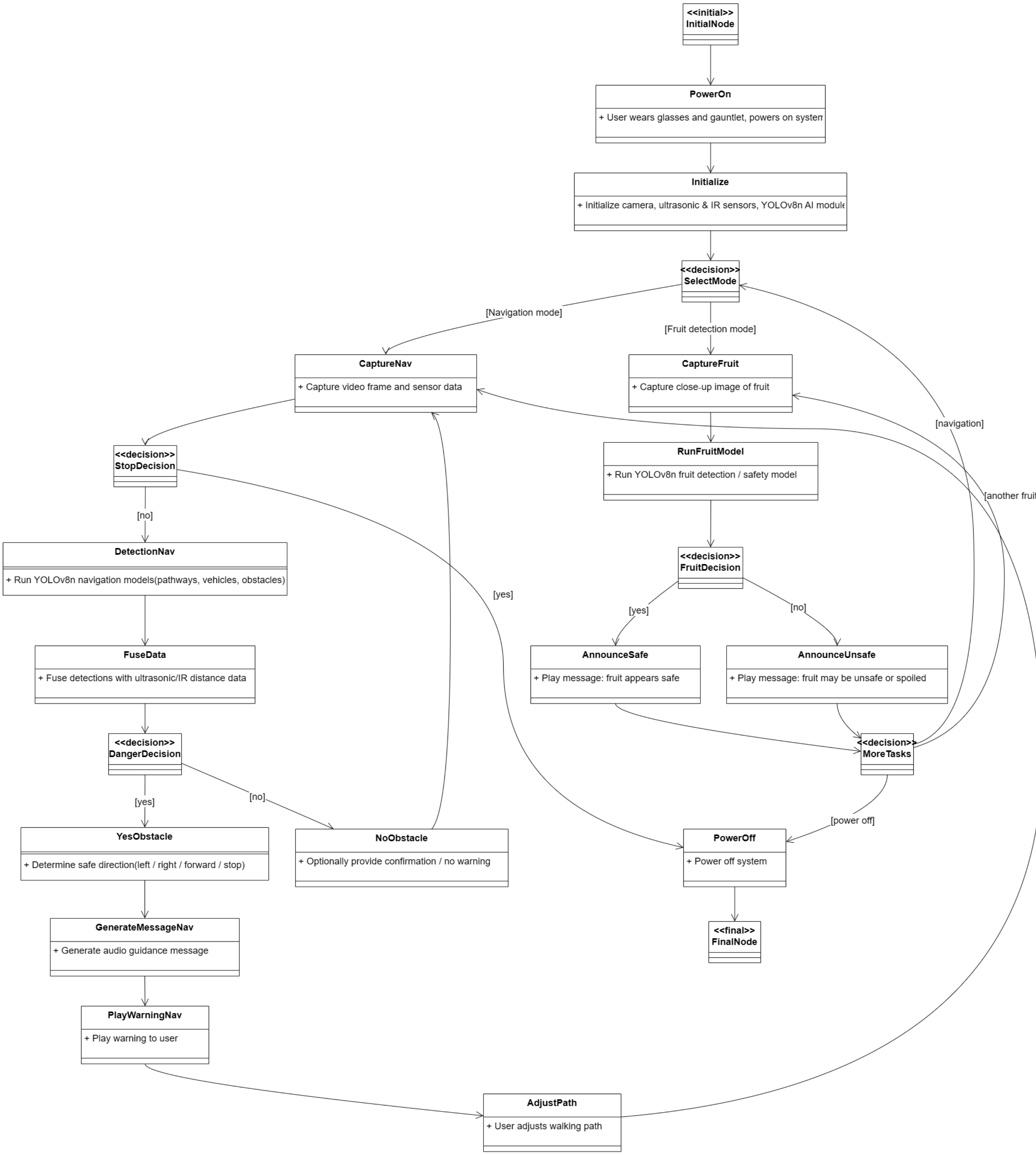

Figure 13 Activity Diagram

Chapter 4: Design and Implementation

4.1. Design approach

Smart glasses system follows a sensor – processing – decision – feedback pipeline, shown in Figure 14 Smart Glasses Pipeline, tailored for real time walking assistance. Environmental data is constantly collected through camera, ultrasonic, and IR sensors. Then this data is transferred to the raspberry pi gauntlet that processes this data. Captured images are resized for saving memory and processing power then are fed to YOLOv8n models. As mentioned before in chapter 3, the project contains three ai models that are focused on pathway identification, obstacle identification, and fruit identification. These models work in parallel to save processing power and memory, then delivered to the fuse module, where it combines the identification data with acquired distances of objects from other sensors. Such data is then delivered to decision module, where it determines how the user should move (forward, left, right, stop) or the type of fruit and if it's safe to eat through audio feedback given to the user using a speaker module attached to the glasses.

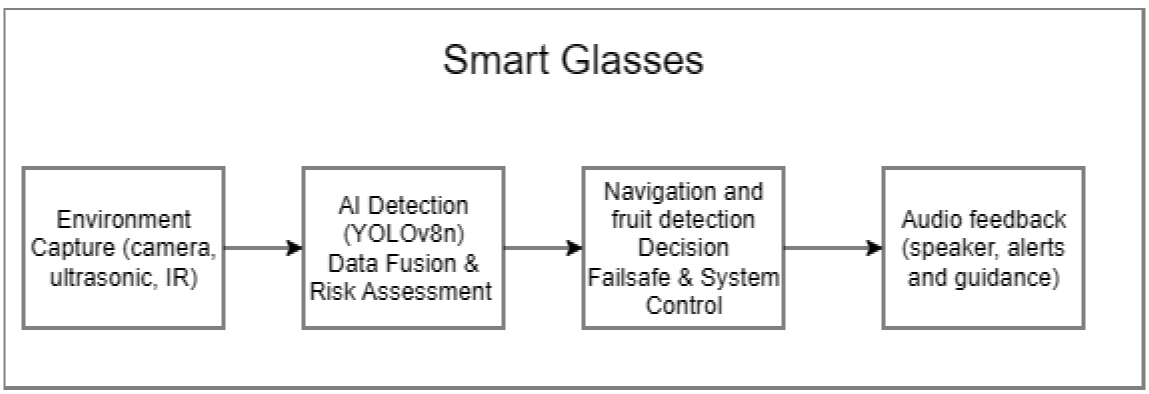

Figure 14 Smart Glasses Pipeline

At the heart of the design is an AI centric architecture built around YOLOv8n object detection model. YOLOv8n is chosen because it provides strong detection accuracy for multiple classes while staying lightweight enough to run in real time on limited hardware. As Jeong et al. (2025) shown that YOLOv8n can operate on mobile devices with an end-to-end latency of around 500 ms and a total memory usage of around 80 MBs, this proved that YOLOv8n is a great model for operating on micro computers with limited memory and processing power. However, unlike purely vision-based approaches for determining distances, the architecture combines camera-based detections with ultrasonic and IR distance readings to solve the nighttime and bad weather failure modes identified in Jeong et al.'s XR system.

The software is decomposed of a group of cooperating modules. The environment capture module acquires video frames and continuous readings from ultrasonic and IR sensors. Then the AI detection module runs the YOLOv8n model on resized frames to create bounding boxes, object classes, and confidence scores for pathways, obstacles, and fruits. Then the Data Fusions and risk assessment module combines these detections with distance values to estimate the level of danger for each object and to determine whether it's in the user's path or not. The Navigation decisions and fruit identification module maps the resulting risk levels into discrete commands such as "forward", "left", "right", and "stop" taking care not to overwhelm the user, and allow the user to determine the type of fruit and the status of the fruit before eating when he changes modes. Finally, the Audio feedback module converts chosen commands into short speech or audio tones and plays them through a speaker module embedded in the glasses, while a system control and failsafe module supervise the timing, monitors sensors' data and AI failures, and enforces a failsafe behaviour in case the system fails. This is shown in both Figure 14 Smart Glasses Pipeline & Figure 13 Activity Diagram.

Compared to Jeong et al. (2025) smartphone based multi-model XR system, the design follows the same core idea of YOLOv8n-based perception with multi-model structure working in parallel, but simplifies and adapts it for embedded, low-power hardware and augments it with sensor fusions to improve reliability of the system, as shown in Table 10 Comparison between XR system and smart glasses.

Table 10 Comparison between XR system and smart glasses

	XR smart glasses (Jeong et al.)	Smart glasses (This project)
Hardware	Xreal Light XR glasses + Galaxy S22 smartphone	3D-printed smart glasses + Raspberry PI 4B 8 GB gauntlet
Main Sensors	RGB camera on XR glasses	Pi Camera on glasses + Ultrasonic and IR distance sensors
AI model Structure	Three separate YOLOv8n models (walkway, transport, obstacles) running in parallel/ scheduled over 60 fps	Three separate YOLOv8n models (Pathways, Obstacles, Fruits) running in parallel/scheduled over 60 fps
Resource usage and latency	About 80 MB total model memory and 583 ms average processing time on smartphone	Designed to keep end-to-end latency less than or equal to 500 ms with a reduced model size and lower resolution
Power	Powered by smartphone battery, estimated 2 hours continuous use	Separate battery pack sized for at least 6 hours continuous usage
Analysis Techniques	Pure vision-based detection; performance	Vision + Ultrasonic + IR fusion to improve

	degrades in night-time, fog, and rain	detection in low light and adverse weather
User feedback	Visual overlays in XR display + voice guidance; reports of frequent alerts that can overwhelm users	Audio only guidance with explicit design goal to avoid overload and keep messages short

4.2 AI Module Design

4.2.1. Three model YOLOv8n pipeline

In Figure 15 Three model Pipeline, the AI detection module follows the same three model design concept introduced by Jeong et al., adapted to run on raspberry Pi instead of a smartphone. The camera on the smart glasses provides a continuous video stream, which is processed by three specialized YOLOv8n models: Walkway Model for detecting sidewalk and roads, Fruit Model for detecting safe fruits for eating, Obstacles model for detecting pedestrians, bollards, etc…

Each model is finetuned on a dataset tailored to its task, allowing it to focus on fewer object classes and achieve reliable detection with a compact YOLOv8n configuration suitable for embedded hardware.

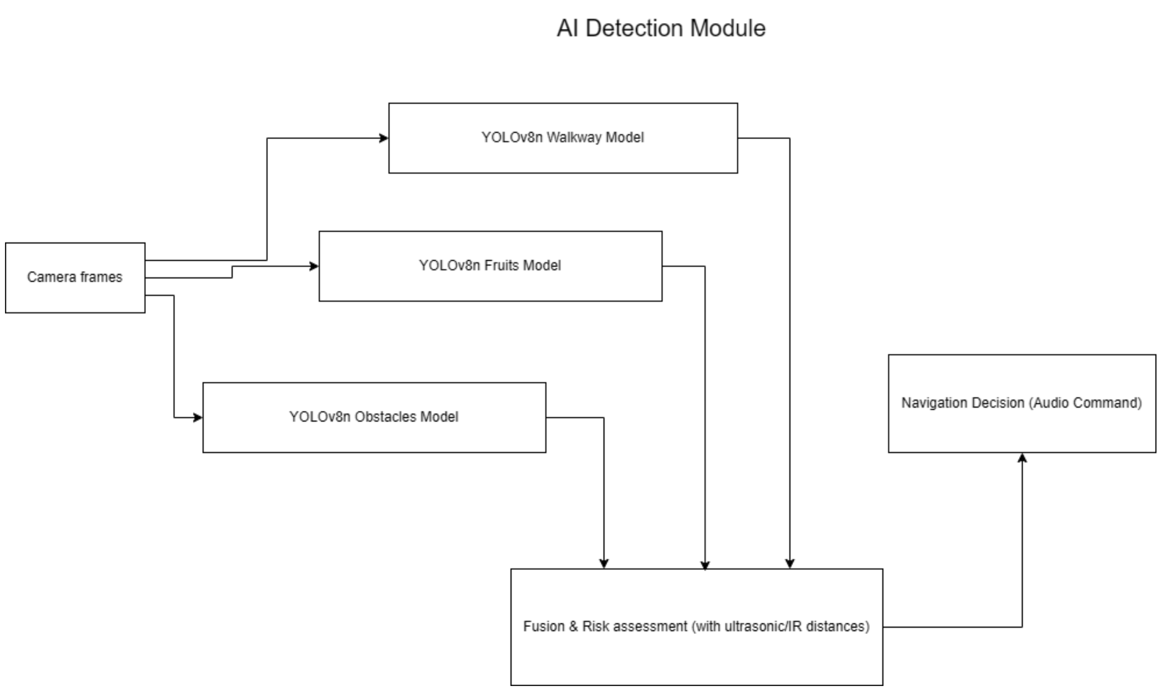

Figure 15 Three model Pipeline

4.2.2. Frame Scheduling

To meet the end-to-end delay requirement of approximately 500 ms while using limited Raspberry Pi resources, the system adopts the frame-scheduling strategy proposed by Jeong et al. The camera captures 60 frames per second, and within each one-second cycle the frames are divided into three consecutive groups: frames 1–20 are processed by the walkway recognition model, frames 21–40 by the Fruit detection model, and frames 41–60 by the obstacle detection model, shown in Figure 16 Planned frame allocation schedule. At the end of each cycle, the detections from all three models are integrated to produce a complete description of the environment, which is then used to generate navigation commands while the next cycle begins.

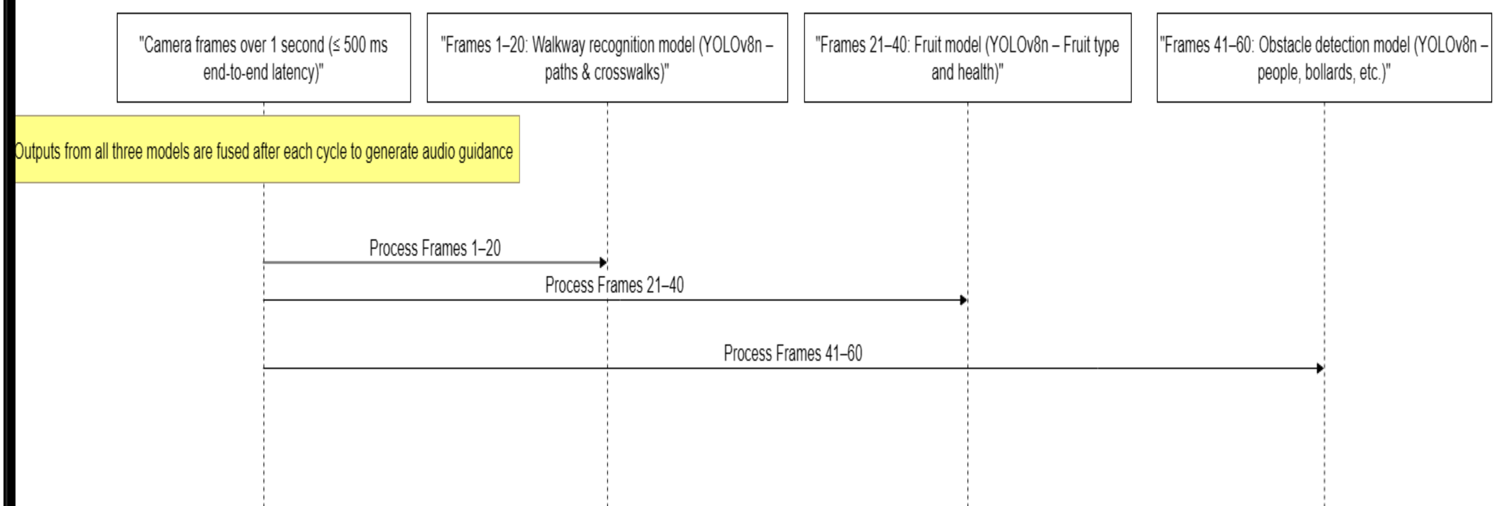

Figure 16 Planned frame allocation schedule

4.2.3. Fusion, risk scoring and command generation

After each cycle, all detections produced by the three models are merged with the latest ultrasonic and IR distance readings in a shared fusion layer. For every detected object, the fusion module stores the class, confidence score, bounding box location in the image, model source, and an estimated distance based on the depth sensors, then applies rule-based risk scoring that considers object type, distance thresholds and whether the object lies in front of the user or off to the side. Objects marked as dangerous are passed to the navigation decision logic, which prioritizes the highest-risk obstacle and maps it to a simple command such as “continue”, “turn left”, “turn right” or “stop”; however, if the user is using the fruit detection mode, he will be provided with audible feedback like: “orange – rotten”, “apple – not rotten” this command, together with its confidence level, is sent to the audio feedback module to be spoken to the user in real time.

Conclusion of Part A

The work in this project started from a clear real-world problem: many blind and visually impaired people still rely mainly on canes or guide dogs because many modern electronic aids are expensive, heavy, have short battery life, or give slow and overwhelming feedback. The introduction and literature review showed that, although systems such as DEEP-SEE, multimodal sensor devices and XR-based smart glasses have achieved promising results with AI and computer vision, they still suffer from issues like limited operating time, poor performance in difficult environments and complex user interfaces.

Building on these findings, the project proposed a smart obstacle and object detection system composed of 3D-printed glasses with a camera and audio output, and a gauntlet hosting a Raspberry Pi 4B with ultrasonic and IR sensors. The requirements chapter translated the main goals into detailed functional, non-functional, domain, interface, AI, and hardware/software requirements, including real-time operation with a target end-to-end delay of 500 ms, at least six hours of continuous use, and compliance with Egyptian assistive-technology standards and disability law. Use-case, sequence and activity diagrams were then developed to describe how the user, sensors, AI models and audio feedback interact during navigation, forming the basis for a structured system design.

The design chapter brought these elements together in a modular, AI-centric architecture organized around three YOLOv8n models for pathways, vehicles and other obstacles, combined with ultrasonic and IR distance readings. A frame-scheduling strategy adapted from Jeong et al. allocates camera frames to each model within a one-second cycle, and a fusion and risk-assessment layer converts detections into simple commands such as forward, left, right or stop that are delivered to the user through audio. Compared with Jeong et al.'s smartphone-based XR solution, the proposed design focuses on lower-cost hardware, longer battery life and sensor fusion to improve robustness in low-light and bad-weather conditions while keeping the interaction purely audio-based for blind users.

Overall, the report defines a complete path from identifying the needs of blind users, through reviewing existing assistive technologies, specifying detailed requirements, and proposing a realistic AI-driven design that can run on embedded hardware. Although the actual implementation and experimental testing are left for future work, the analysis and design presented here show that using lightweight YOLOv8n models and multi-sensor fusion on a Raspberry Pi platform is a promising direction for building smart glasses that are more affordable, safer and more comfortable for daily navigation by visually impaired people.

References:

1. Kasowski, J., Johnson, B.A., Neydavood, R., Akkaraju, A. & Beyeler, M. (2023) ‘A systematic review of extended reality (XR) for understanding and augmenting vision loss’, *Journal of Vision*, 23(5), p. 5. DOI: http://doi.org/10.1167/jov.23.5.5
2. Lin, Y., Wang, K., Yi, W. & Lian, S. (2019) ‘Deep learning-based wearable assistive system for visually impaired people’, in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (ICCVW 2019)*, Seoul, Republic of Korea, 27 October–2 November 2019, pp. 2549–2557. Available at: https://arxiv.org/abs/1908.03364
3. Manjari, K., Verma, M. & Singal, G. (2020) ‘A survey on assistive technology for visually impaired’, *Internet of Things*, 11, 100188. DOI: http://doi.org/10.1016/j.iot.2020.100188
4. Jeong, I., Kim, K., Jung, J. & Cho, J. (2025) ‘YOLOv8-based XR smart glasses mobility assistive system for aiding outdoor walking of visually impaired individuals in South Korea’, *Electronics*, 14, 425. (Full citation as per journal record; uses YOLOv8n-based XR mobility glasses system). DOI: https://doi.org/10.3390/electronics14030425
5. Law No. 10 of 2018 on the Rights of Persons with Disabilities (2018) *Official Gazette of Egypt*. Summary in: State Information Service (n.d.) ‘The most important features of Law No. 10 of 2018 on the rights of persons with disabilities’. Available at: https://hrightsstudies.sis.gov.eg/en/laws/the-most-important-features-of-law-no-10-of-2018-on-the-rights-of-persons-with-disabilities/
6. Statista (2024) ‘Vision loss predicted to surge 55% by 2050’. Available at: https://www.statista.com/chart/31502/expected-number-of-people-with-vision-loss-globally/
7. Vision Loss Expert Group & International Agency for the Prevention of Blindness (IAPB) (2020) ‘Magnitude & projections of vision loss’, *IAPB Vision Atlas evidence series*. Available at: https://www.iapb.org/learn/vision-atlas/magnitude-and-projections/projected-change/
8. Egyptian Organization for Standardization and Quality (EOS) (2021) *Egyptian Standard ES 7866/2021: Assistive products for persons with disability – General requirements and test methods* (based on EN 12182:2012): https://www.eos.org.eg/en/standard/14125.