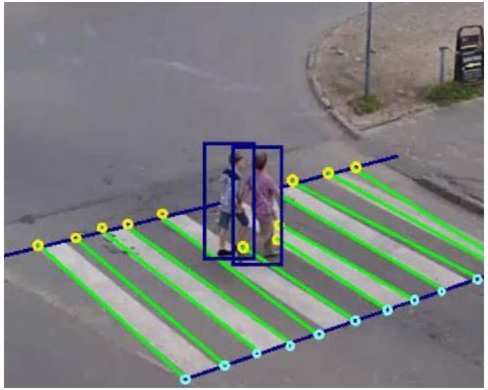
Use Case

CHALLENGE: Visually impaired pedestrians rely on external aids to cross the road.

- Not all crossroads have reliable or maintained aids
- Obstacles (like debris) on the road can further reduce safety

SOLUTION: A chest-worn device that aids the user in crossing the road by providing real time visual-to-audio guidance.

- “WALK” vs. “DON’T WALK” signals
- Helps user stay within the crosswalk
- Alerts to unexpected objects in the walking path



Use Case/Design Requirements - Hardware

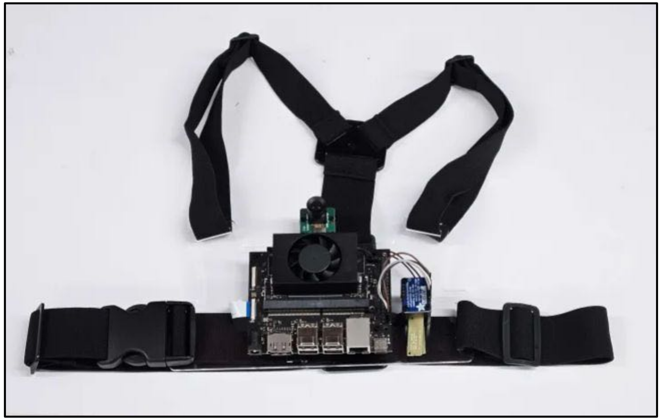
Use Case Requirement	Design Requirement	What we have
Detect walk signs and obstacles in real-time to guide user	IMX219 wide-angle camera: 1920x1080@30fps, 105° FOV	IMX219 1280x720@30fps, 105° FOV (Processing speed)
Provide real-time audio feedback within 0.5s of creating navigation event	USB sound card + on-ear headphones(<0.01s hardware latency)	Bluetooth + Wireless bone conduction headphones (higher latency, but acceptable)
Device must last ≥ 2 hrs on battery	10,000mAh power bank with DC barrel jack adaptor	24,000mAh Power Bank with 15V 5A USB-C to DC converter. (Lasts 11 hrs on battery)
System must not be uncomfortable	System must weigh < 5 pounds	System weighs < 3 pounds
Adjustability to fit user's dimensions	An adjustable strap system	An adjustable strap system, with stretchy straps.

Quantitative Design Requirements - Software

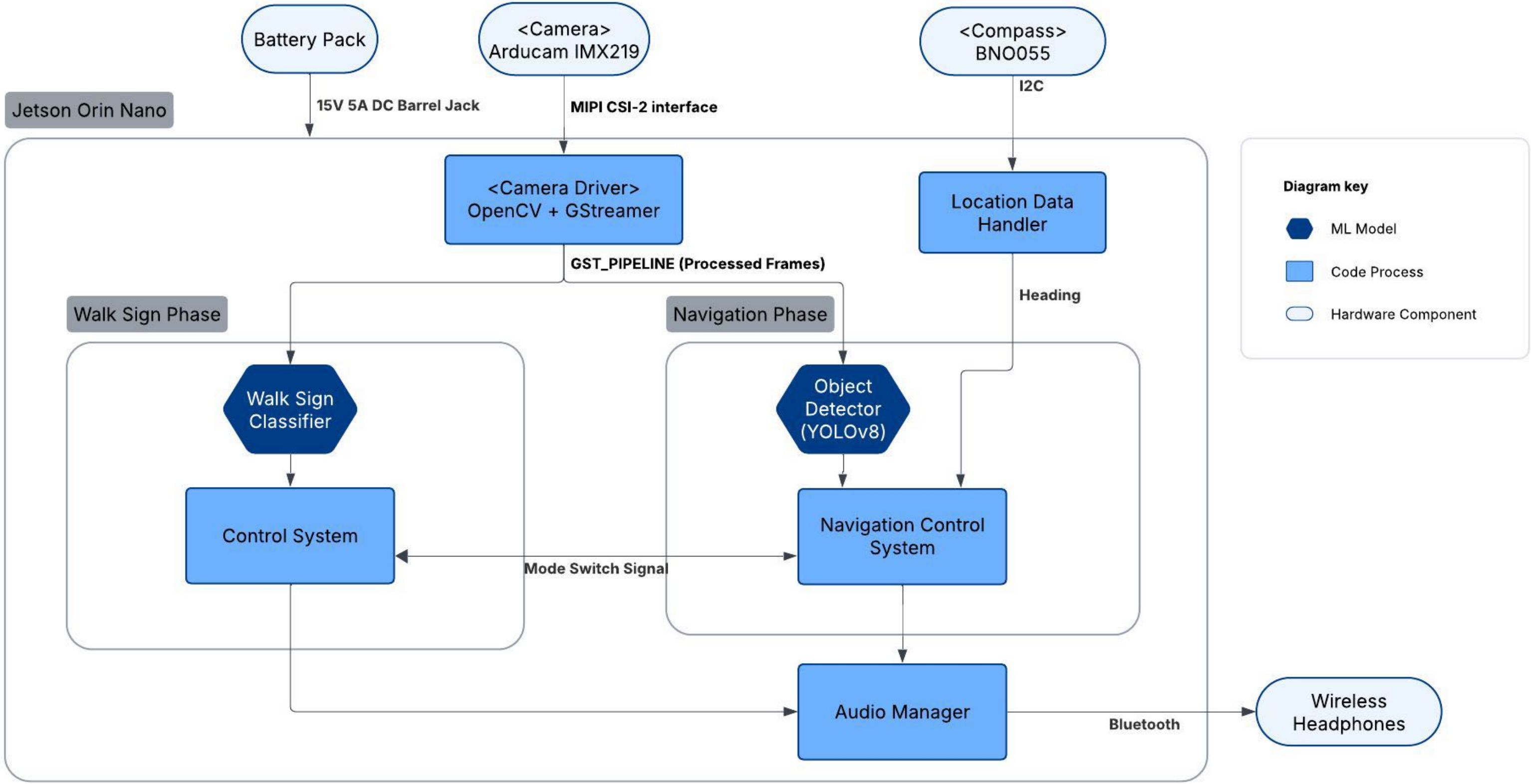
Use Case Requirement	Design Requirement
The system must correctly classify the walk signal with an accuracy of >95% to ensure user safety.	The ResNet image per-frame accuracy must be > 85% and AUROC > 90% to obtain a >95% majority vote accuracy over 5 frames.
The system must provide a decision within 2 seconds to allow users to react to a "GO" signal.	ResNet model inference latency must be ≤ 400 ms per frame.
The system must detect obstacles with at least 90% accuracy to ensure user safety.	The YOLOv12 computer vision model used for object detection must achieve ≥ 70% mAP (mean Average Precision).
The system must provide crosswalk navigation feedback within 0.5 seconds to allow users enough time to react.	YOLOv12 model inference latency must be ≤ 0.5 sec per frame.
The system must provide corrective audio feedback within 0.5 second of detecting veering, and clarity of feedback should be satisfactory as per the user.	Audio feedback system (e.g., using pyttsx3) must generate a response within ≤ 0.5 sec of deviation detection. Audio feedback must be tested for clarity and comprehension, with at least 90% user satisfaction in trials.

Solution Approach

- How We Are Solving the Problem:
  - Real-time ML on Jetson Orin Nano
  - Camera based identification
  - Audio Feedback
- Considerations:
  - Pro: Avoids reliance on inconsistent public infrastructure
  - Pro: Promotes independence and self sufficiency
  - Con: Cost of device may create economic inequality
- Design Evolution:
  - Head-worn device → chest harness for improved camera stability



Solution Approach: Block Diagram

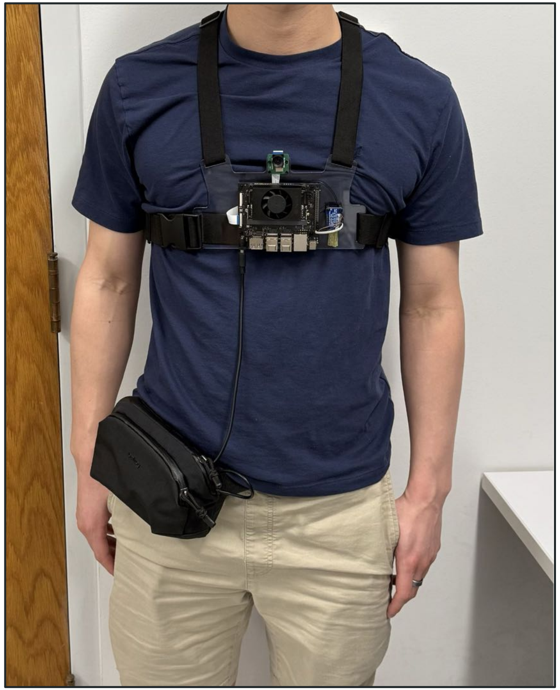


Complete Solution

Demonstration:

- User first waits at a crosswalk.
- Once the light displays “WALK”, the device alerts the user.
- The user crosses the street.
  - If there is an obstacle, the device alerts the user and provides directions to avoid it.
  - If the user strays off of the crosswalk, the device alerts the user and provides directions to correct their course.

Product Setup



Testing, Verification and Metrics - Results

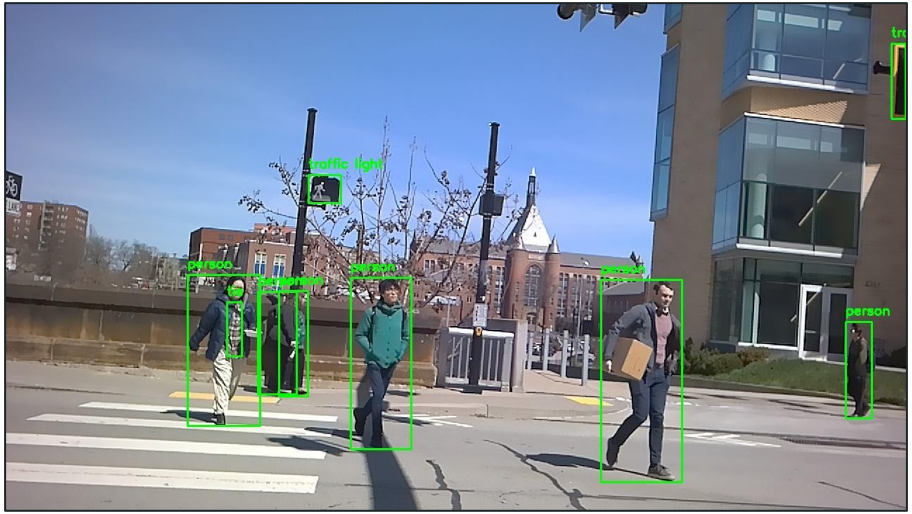
Test Name	Test Inputs	Test Outputs	Passing Criteria	**Results**
Per-frame Classification Performance	Real-world video frames from crosswalks under various lighting and weather conditions.	AUROC (Walk SIgn Classification) mAP (Object Detection)	AUROC > 0.90 accuracy > 85% mAP > 0.7	Walk Sign Model: - AUROC: 0.936 - accuracy: 87.2% Object Detection Model: - mAP: 0.3-0.6
Majority Vote Classification Accuracy	5-frame sequences from test dataset.	Accuracy: Correct majority vote predictions/Total number of sequences	≥ 95% accuracy across test sequences.	5-frame majority vote accuracy: 98.2%
Inference & Audio Response Time	5-frame inference batches	Inference time per frame and batch; audio delay.	Inference latency ≤ 100ms/frame; audio delay ≤ 500ms.	Walk Sign Latency: 33 ms Object Detection Latency: 90 ms Audio Latency: 100 ms
User Verification	10 users of different builds wear device and attempt to detect walk signals and obstacles 5 times each.	% of users successfully orienting camera to detect signals.	≥ 90% success rate within 5s.	To Be Determined

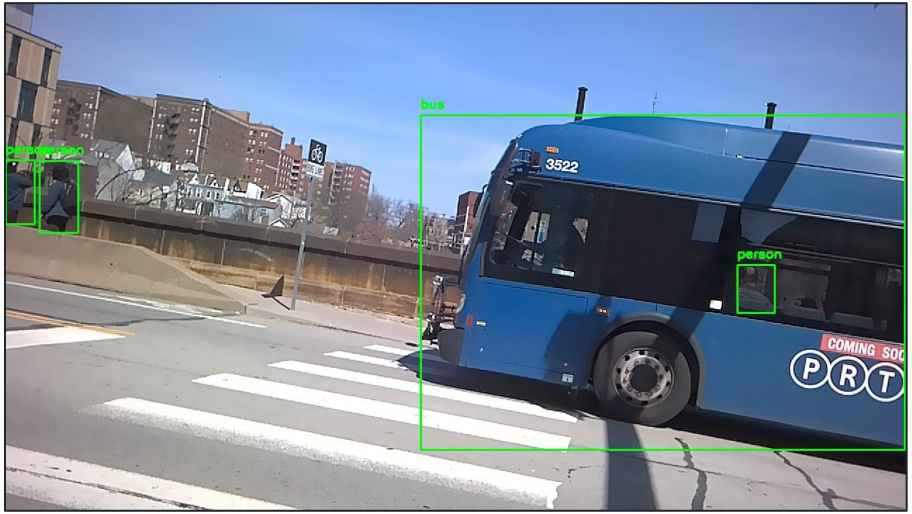
Testing, Verification and Metrics - Results

Test Name	Test Inputs	Test Outputs	Passing Criteria	**Results**
Veering Angle Detection	Controlled user tests walking along a straight path and at predefined angles (10°, 20°, 30°).	Correctly detecting 20°+ deviations/All 20°+ deviations	≥ 95% accuracy for 20°+ deviations.	To Be Determined (drift correction TBD)
Audio Response Time	Users simulate veering (>20°); measure delay from detection to audio.	Mean delay (ms).	≤ 300ms response time.	≤ 100ms response time.
Audio Feedback Clarity	Users intentionally veer off-course and provide subjective ratings.	Clarity score (1–5 scale).	Mean clarity rating ≥ 4.0	To Be Determined
Power Draw Test	Measure device power consumption under maximum load.	Battery life (hours) at peak wattage.	≥ 6 hours of operation.	Idle (4.1W) WalkSignClassifier (7.1W) ObjectDetection (17.0W) Max: 8 hrs Min: 3 hrs

Object Detection Evaluation

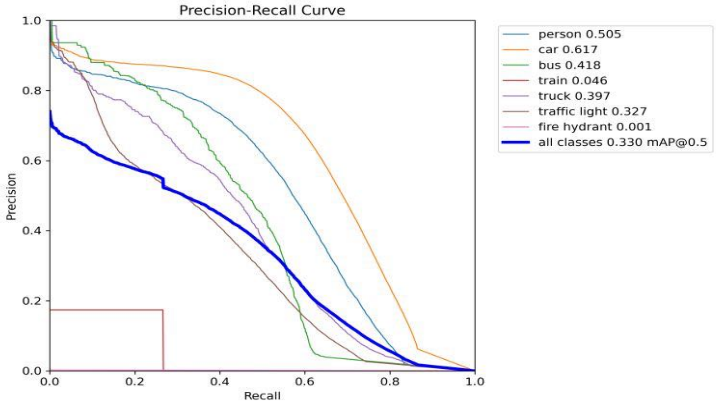
- mAP metric for object detection model did not meet 0.7 criteria, but model performed well in field tests
- Ex. correctly finding and labelling bounding boxes as shown below

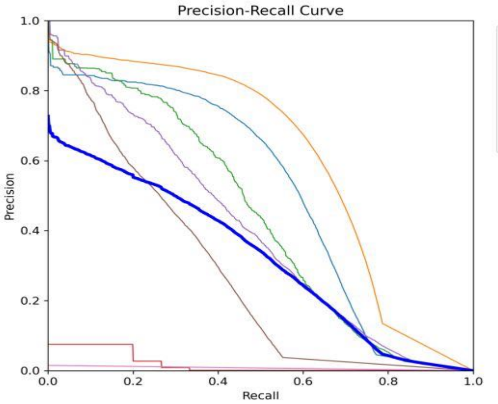




Pareto Tradeoff	Chosen Option	Considered Alternative Option
Model inference speed vs. precision	YOLOv12 (30ms faster inference)	RT-DETR (marginally higher precision)
Wireless vs. wired earbuds	Wireless (user comfort and convenience)	Wired (negligible latency)
Bone-conducting vs regular earbuds	Bone-conducting earbuds (higher spatial awareness)	Regular earbuds (higher audio clarity)
Battery weight vs lifespan	700g battery with 8 hr lifespan	400g battery with 4 hr lifespan

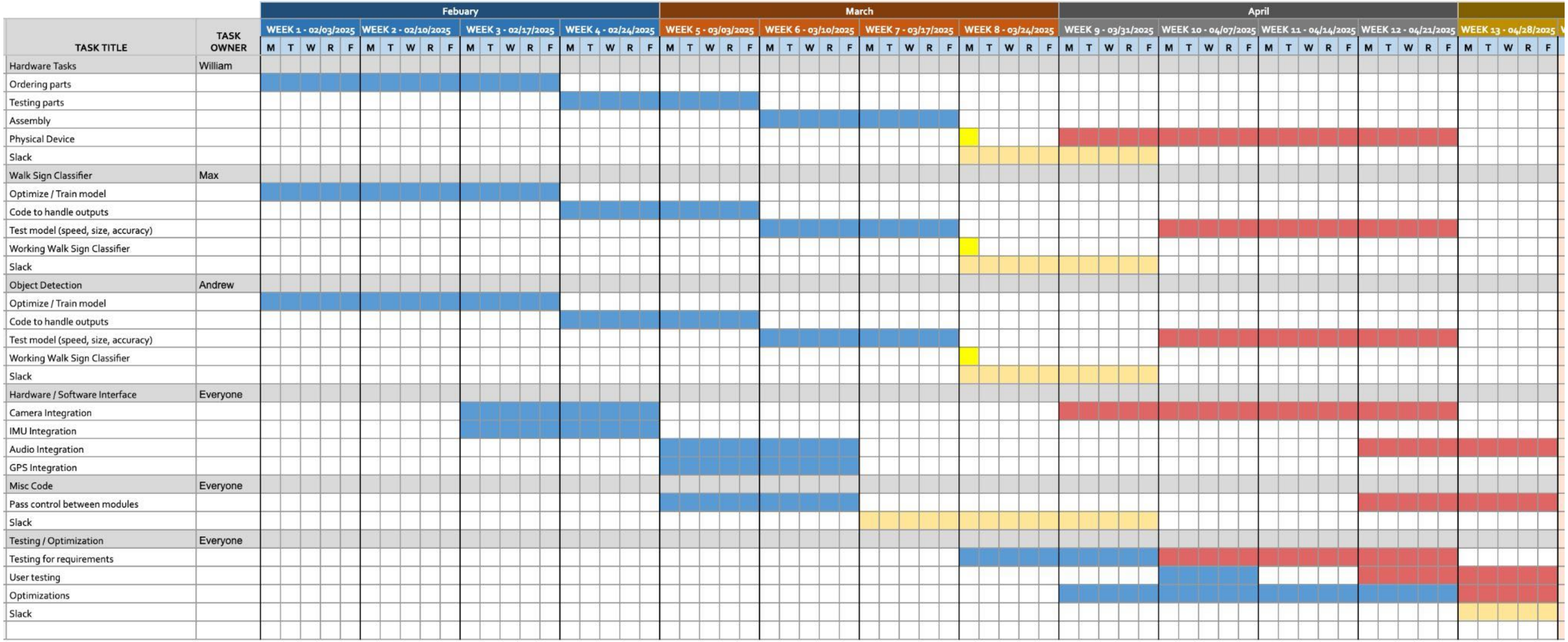
YOLOv12
(90 ms)





RT-DETR
(130 ms)

Gantt Chart


Febuary
March
April
WEEK 1 - 02/03/2025
WEEK 2 - 02/10/2025
WEEK 3 - 02/17/2025
WEEK 4 - 02/24/2025
WEEK 5 - 03/03/2025
WEEK 6 - 03/10/2025
WEEK 7 - 03/17/2025
WEEK 8 - 03/24/2025
WEEK 9 - 03/31/2025
WEEK 10 - 04/07/2025
WEEK 11 - 04/14/2025
WEEK 12 - 04/21/2025
WEEK 13 - 04/28/2025
TASK TITLE
TASK OWNER
M T W R F
Hardware Tasks
William
Ordering parts
Testing parts
Assembly
Physical Device
Slack
Walk Sign Classifier
Max
Optimize / Train model
Code to handle outputs
Test model (speed, size, accuracy)
Working Walk Sign Classifier
Slack
Object Detection
Andrew
Optimize / Train model
Code to handle outputs
Test model (speed, size, accuracy)
Working Walk Sign Classifier
Slack
Hardware / Software Interface
Everyone
Camera Integration
IMU Integration
Audio Integration
GPS Integration
Misc Code
Everyone
Pass control between modules
Slack
Testing / Optimization
Everyone
Testing for requirements
User testing
Optimizations
Slack