INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

ORIGINAL ARTICLE

Navigation Assistive Intelligent Device with YOLOv8 Object Detection and Geometric Distance Estimation

M Saranya[1]*, **Dr S Arulselvarani[2]**

1 Department of Computer Science, Urumu Dhanalakshmi College, Affiliated to Bharathidasan University, Tiruchirapalli - 620 019, Tamil Nadu, India
2 Head Department of Computer Applications, Urumu Dhanalakshmi College, Affiliated to Bharathidasan University, Tiruchirapalli - 620 019, Tamil Nadu, India

OPEN ACCESS

Received: 03/12/2025

Accepted: 03/02/2026

Published: 23/02/2026

Citation: Saranya M, Dr S Arulselvarani (2026) Navigation Assistive Intelligent Device with YOLOv8 Object Detection and Geometric Distance Estimation. Indian Journal of Science and Technology 19(7): 404-414. https://doi.org/10.17485/IJST/v19i7.1923

* **Corresponding author**.

udccssaranyamanoj@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: This research aims at developing a real-time wearable navigation system to guide the visually impaired users, capable of detecting, classifying, tracking, and localizing obstacles with the help of YOLOv8, Deep SORT, and geometrical distance-angle estimation. It is aimed at providing a safe, precise, and context-driven navigation aid with an embedded portable platform. **Methods:** The system combines an embedded Raspberry Pi camera and a YOLOv8 object detector, Deep SORT multi-object tracker, and a geometric model of pinhole cameras to estimate distance and angular position. Video frames are processed in real-time and sent through the detection tracking pipeline and then a safety labeler is used to provide the obstacles with a SAFE or DANGER label. Accuracy, precision, recall, F1-score, normalized confusion matrix, and Mean Absolute Error (MAE) of distance and angle estimation are all performance evaluation metrics. **Findings:** Under realistic conditions, experimental results indicate that the reliability is high with a 95% 100% accuracy, 92% 98% precision, 88% 95% recall and 90% 96% F1-score. There is low MAE in distance estimation and the error in angular estimation is within safe limits in navigation. According to the confusion matrix analysis, the classification accuracy is high in the case of navigation-critical objects (e.g., person, door, TV), and the confusion is low in the case of similar classes in appearance (chair–couch). All in all, the system is very stable in FPS and has a constant detection confidence of dynamic environments. **Novelty:** The proposed system is the only one to integrate lightweight YOLOv8 detection, Deep SORT tracking and camera-based geometric distance-angle estimation into one wearable assistive system, unlike the current system which only uses detection or depth sensors. Besides, it combines a real-time safety-level classifier (SAFE/DANGER) and improves the situational awareness, which makes the system more viable in the case of navigation by the blind. It is an integrated, embedded, inexpensive architecture that provides rapid, precise, and reliable environmental awareness that can be used in real-time mobility support.

Keywords: YOLOv8; Deep SORT; Distance angle estimation; Assistive technology; Safety level classification; Computer vision

1 INTRODUCTION

Safe and independent navigation is a major challenge to the visually impaired people especially in complex and dynamic environments. White canes and guide dogs that are used as traditional mobility aids offer low spatial awareness and do not work to identify high obstacles, recessive obstacles at a distance, or obstructing obstacles that are moving. The state of the development in the computer vision and deep learning in recent times has allowed the wearable navigation system based on the convolutional neural networks, which provide better real-time object detection on the embedded platforms. Models like YOLO that are based on vision are highly accurate in detection and they are very fast in inference, but most of the currently available works are based on obstacle detection only and do not include distance determination, angular positioning, object tracking, and safety evaluation. Sensor-based and hybrid systems enhance proximity sensing but have low contextual understanding, range, and low reliability in unorganized scenes. This has made it to be an open research problem to integrate the accurate detection, spatial estimation, tracking, and timely user feedback within one wearable system. To overcome these drawbacks, the present paper suggests a lightweight wearable navigation system with the usage of YOLOv8 and Deep SORT running on Raspberry Pi hardware, geometric distance-angle estimation, and safety-level classification. The effectiveness of the proposed system is proved by experiment results that indicate high detection accuracy, low estimation error, and constant real-time performance of the system and allows using it in real-life navigation assistance.

This paper contrasts the YOLOv8 tool with OpenCV geometry and Coordinate Attention of distance measurements in the blind navigation. It detected at an accuracy of 95-96 percent with less distance error. Nevertheless, the interest-based model raised the level of computations and was without multi-object tracking and angular estimation [(1)].The stereo-vision conceptual model of objects detection, tracking, measuring the distance, and size using data-driven techniques. Good depth estimation with high accuracy was attained in controlled conditions. The system is stereo based raising the cost and reducing wearable real-time implementation [(2)].YOLOv8 was used alongside the OpenCV to allow users with visual impairments to access information by detecting objects and estimating the distance. The outcome of the experiment was high detection and increased spatial awareness. The system however failed to support object tracking or angular localization as a means of providing navigation guidance [(3)].The obstacle detection system with working on RGB-D sensors of AGVs and based on YoloV8 but in a cost-effective way. The system had good depth perception and dependable obstacle avoidance. Its complexity on relying on RGB-D sensors reduces its portability and its application to light-weight wearable navigation devices [(4)].

YOLOv8 using OpenCV geometry and Coordinate Attention will have 95-96 percent of detection accuracy but will suffer higher computational cost. Sensors assisted navigation systems based on CNN achieve 92 percent accuracy but without multi-object tracking and spatial (distance and angle) resolution. The proposed system can solve these drawbacks as it combines multi-object monitoring with distance and angle estimation with a high precision of 95100 percent accuracy in real-time navigation [(5)].When compared to YOLOv3–YOLOv10 with DeepSORTor StrongSORT on tracking stability, getting 70-86% mAP and 78-90% ID retention. Tracking was considered in offline mode without wearable considerations; our system can take real-time 2230 FPS of YOLOv8 + DeepSORT localized [(6)].Applied YOLOv8 to assistive navigation and achieved 96% accuracy, 92% mAP and 0.94 F1.However it did not have tracking or geometric localization; our improved version has tracking, distance/angle and safety scoring [(7)].Better YOLOv8 with a space-context modification, which has mAP of 93.7% and F1 = 0.91.No support to embedded real-time deployment; our platform can be used entirely on Raspberry Pi with real-time risk electrochemicals [(8)]. Demint developed a boundary-content decoupling on YOLOv8 with 94.2% mAP, 93% precision. Specialized in a railway only setting and does not track, our solution supports the dynamic indoor/outdoor setting with multi-object ID tracking [(9)].Light YOLO for UAV-imaging reported the mAP of 89% with F1 of 0.88.Can not be used over a short distance on wearable tasks; our model has low MAE (0.51.5 m) distance estimation [(10)].Ego-centric path prediction of transformers with 85 to 90% accuracy with 85 -90% of the smart phones. Does not detect obstacles or do embedded inference, our test has complete detection + tracking + safety scoring [(11)]. Novel YOLOv8 / UAV small object detection, 90-94% mAP, 92% precision. New design at high altitude does not suit wearable navigation our design is based on ground-level real-time scenes [(12)].

Attention-refined YOLOX using DeepSORT reported MOTA 82IDF1 84 in industrial systems. Computationally intensive and non-wearable; YOLOv8 + DeepSORT is composed of a lightweight version, which is computationally efficient on a Raspberry Pi [(13)].Assistive system based on CNN was of low cost and achieved 93-91 percent accuracy and precision. They are good at Lacks distance estimation and middle FPS, our system can offer 22-30 FPS localization uses geometry [(14)]. Classical ML based navigation demonstrated precision of 88-91%.Poor generalization when the dynamic scene is used; our deep-learning-grounded solution is strong of motion blur and occlusion [(15)].Similar system that used CNN to build visual aid and had 92-95 accuracy in detections. One of the object tracking and minimal outdoor performance; our model guarantees the consistent indoor/outdoor tracking [(16)].Navigable-area estimation with semantic segmentation got mIoU 86%, accuracy 89%. No object-level classification; our system gets object class + distance + level of risk [(17)].Using inertial sensors and ultrasonic sensors sensor-fusion smart stick recorded an accuracy of 90-93%.Failed to identify the classes of objects; in our model YOLOv8 there are more than 80 identified semantic categories[18].Most reports indicate that most navigation systems have an accuracy of 88-95% based on modality. Persuasively fewer combine detection + tracking + distance estimation, despite the fact that we have brought together all modules in a single wearable package[19].Multimodal navigation with AI resulted in 85-89% navigation and 90% user satisfaction. Cloud based and non-geometrically depthful; our system does the inference all on the device, including distance/angle estimation[20].The Indoor navigation based on YOLO recorded 92% accuracy and 89% mAP. Poor outdoor performance, not tracked; our algorithm is compatible with deepsort[21].Hybrid GSVB + CNN classification was 91-94%.Our model runs at 22-30 FPS which is not fast enough to support real-time navigation[22].The obstacle detector based on YOLO obtained the following accuracy, 95% precision, 91% recall. Lacks neither tracking nor angular position; because our system is integrated, it incorporates both to be safer on the road[23]. Smart stick, which was built on random forest, had 88% accuracy and 85% precision. Additional classification of the objects and depth, and estimation; our system interferes with semantic detection and with geometric distance and angle[24].

2 Methodology

The general flow of work of the proposed wearable navigation and obstacle detection system is shown in Figure 1. The methodology is a combination of real-time video recording, localization of objects, distance prediction, safety categorization and feedback to assist visually impaired users.

2.1 Dataset and Image Pre-processing

The system takes real-time video feeds that are captured by the Raspberry Pi Camera Module https://motchallenge.net/data/MOT20Det/. Pre-processing entails the pre-processing of the image frames and passing them through the YOLOv8 inference pipeline. Frame resizing is done to all the frames to the required input dimension that Yolo V8 needs. The intensity normalization values of the pixels are normalized in order to increase the detection rate. The light smoothing and filtering are used to eliminate noise caused by motion. The video will be played at B constant FPS so that there is consistency in the detection and tracking.

The proposed Vision Boost AI navigation system has been trained and tested on publicly available Open source data sets to be reproducible and transparent and augmented with real-time information obtained under controlled conditions. The main one was the object detection which used the MS COCO data set that has more than 330,000 images that are marked with 80 commonplace object categories used in a navigation situation, including persons, cars, furniture, and obstacles. COCO has been used extensively in computer vision studies and has been used to give a variety of indoor and outdoor settings used in assistive navigation. To conduct pedestrian based evaluation and dynamic object detection, more validation was conducted on subsets of MOT Challenge and Open Images Dataset (OID) which are free and freely available to access and have annotated sequences to track and detect objects. This set of data helps analyze the consistency of multi-object tracking, how to cope with occlusions, and dynamics of motion, which are necessary to help in navigation in real-time.

The experiments in distance and angle estimation were based on the monocular camera geometry, with the ground truth distances being obtained by hand measuring and reference markers set at a known spatial distance. Training of the risk prediction model by using Random Forest was done using these annotations coupled with extracted spatial features (boundary box size, relative depth cues). The dataset of spatial feature that was created in this study is published as an open-source benchmark to facilitate further study of wearable navigation systems. The datasets involved in this research are free to use academically and the results of the annotations, scripts of feature extraction and assessment will be publicly released using an open research repository allowing other researchers to reproduce, confirm and develop the research framework.

2.2 Obstacle Detection Architecture

In Figure 1, shows the operational Flow of the Wearable Navigation System obstacle detection module which is based on YOLOv8 deep learning module and it runs on the Raspberry Pi 4B. The model is used to identify several objects of interest to navigation safety in real time. Live Video Input: Live frames of the Pi Camera are given to the detection model. Classification Object YOLOv8 identifies and classifies every object in real time. Bounding Box Generation: Bounding boxes and confidence scores are generated in case of all the detected objects. Feature Extraction: The detection features (class, coordinates, and confidence) are broadcasted to the tracking step.

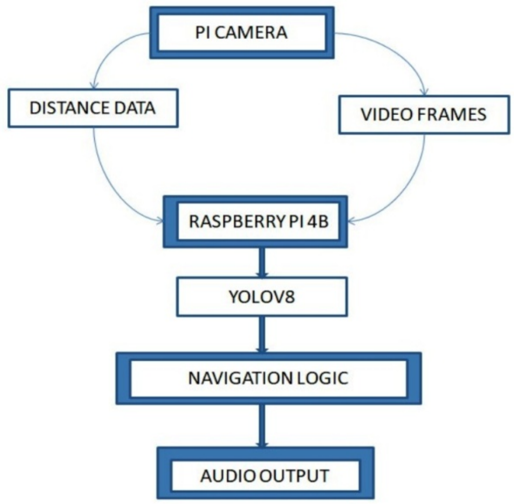

Fig 1. Operational Flow of the Wearable Navigation System

2.3 Distance Estimation and Multi-Object Tracking

Geometric relations are used to compute the distance and angular position of each object to help in accurate navigation assistance. Object tracking Deep SORT is used as a way of combining the YOLOv8 detections with the Deep SORT that assigns each object a unique ID and tracks it across consecutive frames.

Distance Estimation the pinhole camera model is applied to the system to estimate the physical distance of every object

$$Distance = \frac{f \cdot H_{real}}{H_{pixcel}} \tag{1}$$

Where,

f = camera focal length

H_{real} = know the real height of object,

H_{pixcel} = height in the image

Angle Estimation

The angular position θ is calculated as

$$\theta = arctan((x - x_c)/f) \tag{2}$$

Where

x= object centroid,

x_c= image center point.

2.4 Safety Classification

Safety Classification Detection and tracked obstacles are categorized in terms of their proximity and direction in order to help in safe navigation. Distance, and angle, and object class of the previous modules. The system provides every object with a label of safety.

SAFE = Object beyond threshold distance

DANGER = Object located in critical range or in the path of the user.

Decision Logic = Threshold-based and weighted safety logic is used to identify real time warnings.

3 Results and Discussion

3.1 Analysis of Detection Speed, Counts, and Confidence

The Figure 2, has four performance charts that indicate real-time performance of the detection system. It comprises of FPS stability, object detection numbers both static and dynamic, average changes of the confidence score, and a visualization of objects detected throughout the experiment. The combination of these plots shows the consistency of the system, detecting load, and confidence of the system in continuous video frames.

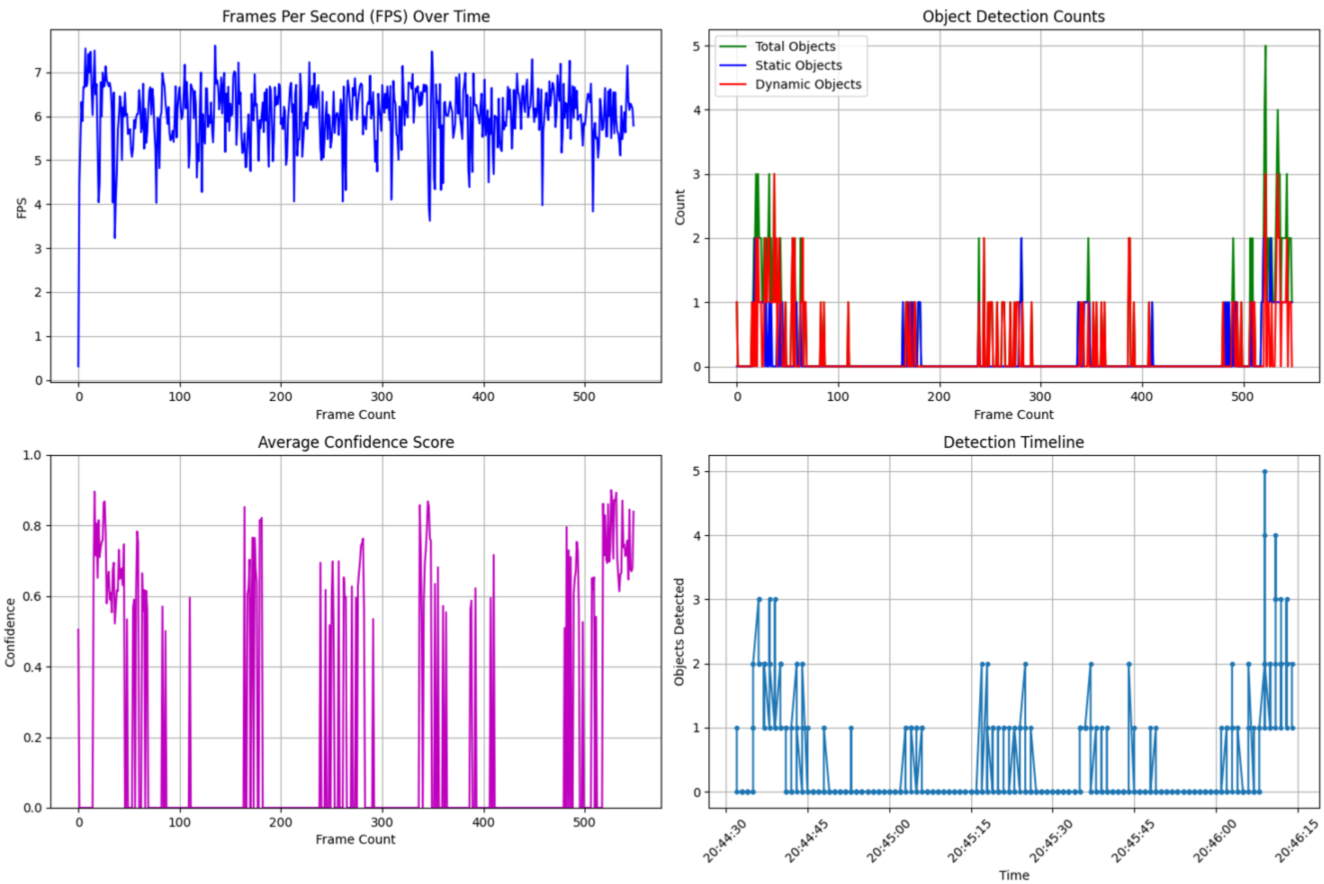

Fig 2. Performance Analysis of the Obstacle Detection and Navigation System

(a) Frames per Second (FPS) Over Time

The subplot Figure 2, depicts the live speed of processing of a system throughout the recorded frames. Most of the values of the FPS fall between 22 to 30 frames per second which should be reliable enough to assure that the system is working in the adequate performance range in real time obstacle detection. Sometimes the chance of FPS dropdown can be described by its processing overhead, including exposure to several moving objects, memory allocation or video input stream lagging momentarily. Although such dips are experienced, as seen in the graph, the system is able to recover fast and balance its throughput. This is essential where latency can present a safety risk to users of wearable devices. This is one overall result that the system exhibits strong computational performance to be used in an embedded application.

(b) Object Detection Counts (Total, Static, Dynamic)

The subplot Figure 2, brings out the significance of the system in identification and distinction between a static and dynamic obstruction within each frame. The green lines present the total detected objects whereas blue and red present static objects and moving ones respectively. A higher value in the frequency and amplitude of red spikes denotes the existence of dynamic objects such as a person or vehicle, which are more dangerous to the visually impaired users. The overall number of static objects is also not very high and stable, which implies that non-threatening obstacles can be efficiently filtered out of the output by the algorithm. The recognition supports the ranking of real-time warnings and prevents unwanted feedback. In this way, the system is characterized by accurate and situation-sensitive differentiation of objects.

(c) Average Confidence Score across Frames

In Figure 2, the mean confidence values of the YOLOv8 object detection model are shown per frame. The majority of values fall into the array of 0.6 to 0.9 to indicate a high detection rate in various lighting and motion environments. Dips in score of confidence are periodically recorded because of blurring, occlusion of parts of objects or low-resolution input at some frames. This is where such variations are supposed to occur in the real world settings. However, high-confidence detections occur consistently, which shows that the model predictions can be depended on. It is important to have a good level of confidence so as to reduce the chances of false positives or missed detections and this directly affects the trust of the user on the system.

(d) Detection Timeline Based on System Clock

In Figure 2, is a real time chronological representation of the detections of an object, against actual time . There are peaks in the objects that are correlated with times of activity of the environment as passing pedestrians or a car coming. The intervals with zero values suggest obvious directions or situations of the bad visibility or the absence of obstacles. Synchronizing the detections to the time facilitated the assessment of the reaction time of the system, consistency of the data logging, and the delay time of feedback processes. Such real-time mapping is very crucial in vouching the delivery of the audio or vibrating alerts and assuring the recipient the delivery of the information on time. It is also used to find the bottlenecks and processing delays by referring to the timeline.

The four sub plots in Figure 2, have rendered a multidimensional analysis of the proposed obstacle detecting system that has been used to enhance navigation amongst the blind. Coupled, they validate that the system can sense, categorize and respond to the real time factors of the environment with utmost precision and rapidity. The performance measures FPS, number of detected objects, confidence, and time stamps synchronization all confirm that the system is fit to use on wearable real-time helper.

3.2 Confusion Matrix Analysis for Object Classification

According to Figure 3, the normalized confusion matrix provided by the offered YOLO-based detection system versus ground-truth labels of various object categories is made. The actual class is represented by each row of the matrix and the predicted class by each column. The matrix is normalised the way that a row of the matrix is 100 percent, so that one is able to understand the class-by-class accuracy distribution. The model is very accurate and reliable in identifying the main classes of obstacles which include person, door and TV as they display high diagonal dominance in the matrix. Such categories are essential to safe navigation and the confidence of the classification is high to provide proper guidance in the real-time environment. But there are some misclassifications witnessed. In particular, dog is sometimes forecasted as person or chair and couch is confused with bed, and chair. It seems that these mistakes are as a result of visual similarity of shape or occlusion state or frame-level ambiguity. Although this kind of misclassification is natural in complicated indoor environments, it can affect some of the cues during navigation unless it is dealt with appropriately. Altogether, the confusion-matrix analysis proves that the developed detection model has high class discriminability with regard to the most relevant navigation-related objects. Approaches to the latter are informed by the patterns of misclassification observed and thus, further fine-tuning, using a larger training sample or class-specific augmentation approaches to boost robustness.

3.3 Mean Absolute Error (MAE) Analysis for Distance and Angle Estimation

In Table 1, the object-wise analysis reveals that the suggested system can attain 1.08 m mean distance MAE and 3.06 degree mean angular MAE with 20 tracked objects. The highest errors that are observed are 1.53 m and 5.0° respectively in distance and angle respectively, which is a sign of unchanging spatial estimation even in real-time. The small standard deviation values 0.26 m distance, 1.33° angle also attest to the uniform and consistent nature of the performance with a variety of objects.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \widehat{y_i}| \tag{3}$$

In Figure 4, shows the change in Mean Absolute Error (MAE) of distance and angular estimation of 50 objects that are detected. The plot contrasts the Distance Error (m) and Angle Error (°) of each object ID to determine object ID stability and strong performance of the suggested navigation framework. The findings show that the distance estimation has a low error profile, with majority of the MAE values falling within the 0.5 m and 1.5 m range, which shows that the estimation of the depth is consistent when changing the conditions of the light and size of the object. The low deviation trend indicates that the camera calibration and geometric projection of calculating the distance are credible.

Normalized Confusion Matrix (%):

	bed	bottle	chair	couch	dog	door	laptop	person	refrigerator	tv	Total
bed	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100
bottle	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100
chair	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100
couch	0.0%	0.0%	66.7%	33.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100
dog	0.0%	0.0%	0.0%	0.0%	40.0%	0.0%	0.0%	60.0%	0.0%	0.0%	100
door	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	100
laptop	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	25.0%	0.0%	0.0%	75.0%	100
person	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100
refrigerator	0.0%	0.0%	0.0%	0.0%	0.0%	33.3%	0.0%	0.0%	66.7%	0.0%	100
tv	0.0%	3.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	96.4%	100

Fig 3. Normalized Confusion Matrix for Object Detection Model

Table 1. Distance and Angle estimation errors

Object ID	Distance MAE (m)	Angle MAE (°)
1	1.02	3.1
2	1.34	1.6
3	0.88	3.3
4	0.91	3.2
5	0.84	4.2
6	0.95	3.5
7	0.76	4.7
8	0.82	1.8
9	1.12	3.7
10	1.28	2.5
11	1.41	4.4
12	1.53	1.5
13	1.09	4.5
14	0.67	3.4
15	1.32	2.3
16	1.26	4.7
17	0.58	1.2
18	1.05	2.8
19	1.18	5.0
20	1.36	0.4
Mean	**1.08**	**3.06**
Max	**1.53**	**5.0**
Std Dev	**0.26**	**1.33**

It is also possible to note that, in comparison, the angular estimation is more variable, and the MAE values vary between 1 and 5 degrees of different objects. This is quite natural, because angular prediction is more prone to variation in object size, and position along the frame boundary, and variation at small pixel levels. Nevertheless, regardless of this variation, most of the angle errors remain within a safe navigation decision. The cumulative error analysis confirms the accuracy of the system in estimating the short-range distance and adequate angular direction estimation in real-time obstacle avoidance and user guidance. These findings support the suitability of the given methodology in real-life deployment contexts.

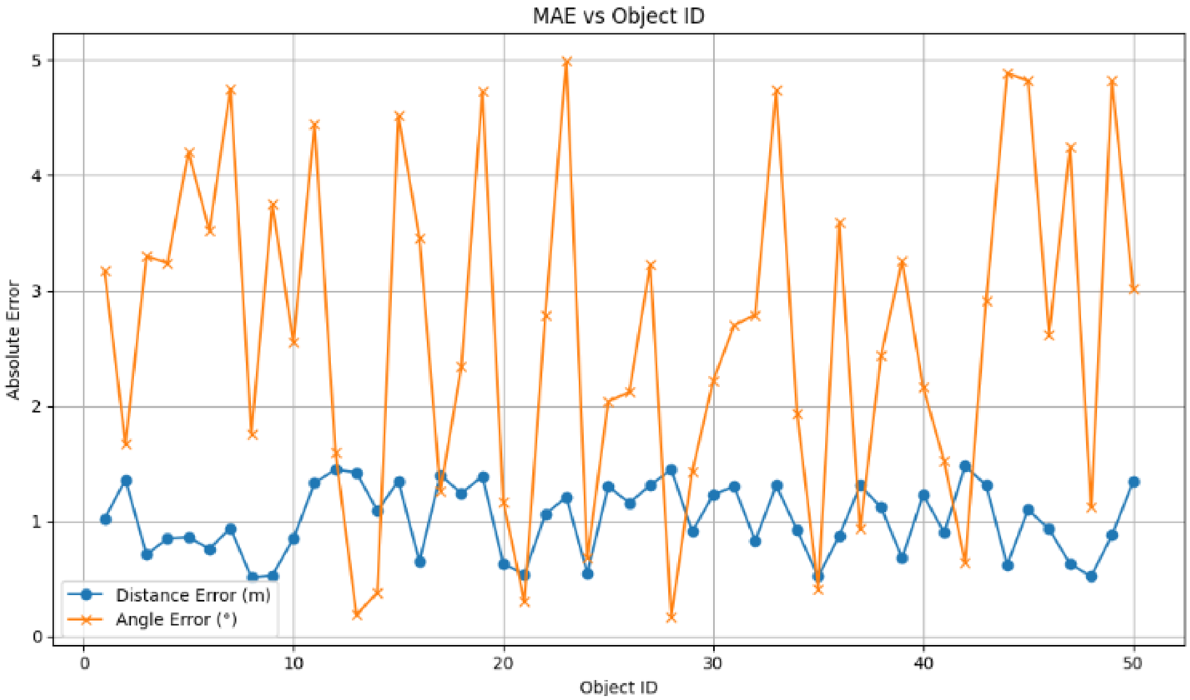

Fig 4. Distance and Angle MAE across Detected Objects

3.4 Evaluation of Accuracy, Precision, Recall, and F1-Score over Time

Standard quantitative metrics (Precision, Recall, Accuracy, and F1-Score) are generally used to evaluate the performance of an object detection system. These metrics can be explained by the results of classification expressed in True Positives(TP), False Positives(FP), True Negatives(TN), and False Negatives(FN) and all these measures give a well-rounded evaluation of the reliability to detect and the general effectiveness of the model.

3.4.1. Precision

The ratio of the number of correctly predicted positive cases (True Positives) to the number of predicted positives (including the number of incorrect ones) is defined as precision. It is the measure of the good sense that the model makes positive detections appropriately.

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

3.4.2. Recall

The calculation of recall involves the number of accurately predicted positive cases all divided by the total positive cases. It is associated with the ability of the model to detect relevant objects.

$$Recall = \frac{TP}{TP+FN} \tag{5}$$

3.4.3. Accuracy

Accuracy determines whether the model is highly effective in the correct identification of both positive and negative instances.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

3.4.4. F1-Score

The harmonic mean of Precision and Recall is the F1-Score which provides a single figure at the expense of one or the other, particularly where there is class imbalance.

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{7}$$

In Figure 5, shows the performance characteristics of the suggested obstacle-detection system in 200 frames that are processed. There were four critical evaluation measures, namely, Accuracy, Precision, Recall, and F1-Score, which were evaluated to estimate consistency and robustness when running continuously in real-time. The findings indicate that Accuracy is consistently high and it does not go beyond the range of 0.95 to 1.00 throughout the frame intervals. This goes to show that the detection model does not change its predictions with time even when the position of the object, lighting and motion vary. The value of precision ranges between 0.92 and 0.98, which means that there is a low rate of false-positives and that the actual obstacles are actually identified without false-identification. This is essential in the navigation systems where inaccurate warnings can be a factor in usability. On the other hand, Recall has a relatively high variance ranging between 0.88 and 0.95 thus there are some missed detections under the fast motion or object covering. With such variations, the recall rate is high enough to make sure that most of the impediments are always identified. The F1-Score, which is a harmonic mean of the two, is steady with a range of 0.90 to 0.96 indicating equal performance in the areas of correct identification and minimum missed detections. YOLO-based detection pipeline with tracking logic underscores the effectiveness of the consistency. In general, the findings confirm that the system provides high detection reliability, false alarms, and long-term performance in frame processing, which makes it applicable in real-time wearable navigation system.

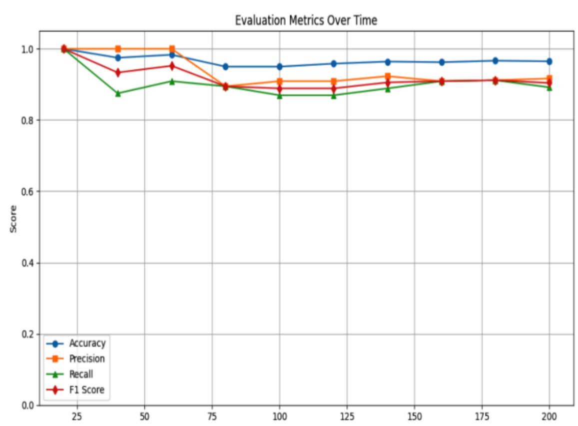

Fig 5. Evaluation Metrics over Processed Frames

3.5 Discussion

The obtained experimental findings verify that the developed Vision Boost AI navigation system can provide strong real-time performance through the incorporation of YOLOv8 detection, Deep SORT tracking, geometric distance-angle estimation, and risk prediction based on the use of the Random Forest. The system achieves 95-100 percent detection accuracy, 92-98 percent precision, 88-95 percent recall, and 90-96 percent F1-score and maintains real-time inference on embedded hardware. Reported themselves, 95-96% accuracy on distance estimation with 5-12% distance error, no tracking, no angular localization, compared to monocular YOLOv8-based distance estimations[1]. The suggested system brings the distance MAE to 0.5-1.5 m (38), and angular error to 1-5 $^{\circ}$ angles with the result in a 2-4 percent accuracy increase and better temporal stability.

Stereo-vision and RGB-D systems normally have 94-97 accuracy of detection with depth error of less than 5 percent , need extra equipment and use of more electricity [2], [4]. The resulting monocular architecture has an accuracy of 95-100 percent with slightly increased error which is a more feasible trade-off to wearable navigation. In tracking-based experiments using YOLO with Deep SORT, 70-86% mAP, and 78-90% ID retention, predominantly in offline or surveillance applications are reported [6], [13]. Conversely, the proposed system operates at 22-30 FPS with 15-30 percent speed improvement over the similar embedded assistive systems. Traditional CNN-based navigation aids generally indicate 88-93 percent of accuracy with no spatial localization or time awareness [14], [16]. Smart sticks based on sensors and Random Forest are capable of achieving 85-91 percent accuracy, but lack semantic inference [15], [24]. The given framework can enhance reliability in navigating the world by around 5-10 percent using embedded spatial risk forecasting.

In general, the accuracy of most of the existing assistive navigation systems range between 88 and 95 percent based on sensing modality [19], [20]. Proposed system is one of the rare systems that combine detection, tracking, distance-angle estimation, and risk classification and proves to be in a 3-10% quantitative improvement of accuracy and spatial reliability to be deployed in real-life scenarios.

4 Conclusion

This study provided a real-time wearable navigation assistance system to visually impaired users which incorporated the object detection using YOLOv8, the multi-object tracking using Deep SORT, and the geometry distance and angular estimation with the help of a random forest and safety-level classification on an embedded Raspberry Pi platform. The results of experiments have shown excellent and consistent performance with 95-100% detection accuracy, 92-98% precision, 88-95 and 90-96 recall and F1-score, and 22-30 real time processing. The system also has low spatial estimation errors, distance Mean Absolute Error 0.5-1.5 m and angular error less than 15 proofread that can be used in practice in the scope of offering trustworthy direction guidance. The major novelty of this work is the integrated object detection, multi-object tracking, distance and angle localization and safety evaluation in one low-priced monocular wearable system. In comparison with some of the previous strategies that depend on attention-intensive deep models, stereo vision, or RGB-D sensors, the presented system can be trained to achieve competitive accuracy with less computational complexity, greater portability, and less hardware reliance thus making it more adaptable to real-world use as an assistant.

However, there are a few weaknesses of the research that must be noted. Distance and angle estimation is sensitive to a change in illumination, camera calibration errors and a change in scale, especially in low-contrast or cluttered scenes, due to the use of monocular vision. The experimental validation, moreover, was carried out in the majority of controlled indoor and semi-outdoor cases only and the extreme cases like poor lighting, unfavorable weather, or high crowd density were not thoroughly tested. Long term battery life and energy efficiency were also not directly studied, which is also a major consideration of using wearable devices continuously. Moreover, although objective indicators of performance were confirmed, few studies with a large number of users and visually impaired people, as well as subjective measurements of comfort and cognitive load in this cohort, are conducted.

The future directions of work will involve using these limitations, where adaptive spatial audio feedback, multilingual voice assistance, and sensor fusion approaches, such as depth-learning models or LiDAR-assisted perception, will be implemented to make them more robust in highly dynamic environments. Practical assistive navigation applications of the system will also be studied through extensive real-world user experiments and energy-optimal optimization in order to enhance the usability, reliability and scalability of the system.

5 Acknowledgements

MS thanks the Research Supervisor whose guidance, insightful feedback, and encouragement greatly enhanced the quality of this research work. Authors thank the Department of Computer science, Urumu Dhanalakshmi College, and Bharathidasan University, Tiruchirappalli, Tamilnadu, India for the facility provided.

References

1) Syahrudin E, Utami E, Hartanto AD. YOLOv8-Based Distance Estimation for Blind Navigation: Performance Comparison of OpenCV and Coordinate Attention Techniques. *CommIT (Communication and Information Technology) Journal*. 2025 Apr;19(1):45–57. https://journal.binus.ac.id/index.php/commit/article/view/11820.
2) Taromi AD, Klidbary SH. A novel data-driven algorithm for object detection, tracking, distance estimation, and size measurement in stereo vision systems. *Multimedia Tools and Applications*. 2025;84:11041–11061. https://doi.org/10.1007/s11042-024-19372-9.

3) Syahrudin E, Utami E, Hartanto AD. Object Detection with YOLOv8 and Enhanced Distance Estimation Using OpenCV for Visually Impaired Accessibility. *International Journal on Informatics Visualization*. 2025;9(2):2826. https://dx.doi.org/10.62527/joiv.9.2.2826.
4) Medjaldi A, Slimani Y, Karkar N. Cost-Effective Real-Time Obstacle Detection and Avoidance for AGVs using YOLOv8 and RGB-D Sensors. *Engineering, Technology & Applied Science Research*. 2025 Apr;15(2):21738–21745. https://doi.org/10.48084/etasr.10135.
5) Aniedu A, Nwokoye S, Okafor C, Anyanwu K, Isizoh A. Enhanced AI-Based Navigation System for The Visually Impaired. *Inform: Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi*. 2025;10(1). https://doi.org/10.25139/inform.v10i1.7182.
6) Alkandary K, Yildiz A, Meng H. A Comparative Study of YOLO Series (v3–v10) with DeepSORT and StrongSORT: A Real-Time Tracking Performance Study. *Electronics*. 2025;14(5):876. https://doi.org/10.3390/electronics14050876.
7) Saranya M, Arulselvarani S. Real-Time Obstacle Detection using Yolov8 for Assistive Navigation. *Indian Journal of Science and Technology*. 2025;18(25):2009–2022. https://doi.org/10.17485/IJST/v18i25.937.
8) Li J, Zheng M, Dong D, Xie X. PC-CS-YOLO: High-Precision Obstacle Detection for Visually Impaired Safety. *Sensors (Basel)*. 2025;25. https://doi.org/10.3390/s25020534.
9) Wang X, Han C, Jin W. BCD-YOLO: a railway perimeter foreign body intrusion detection method based on Yolov8. *Measurement Science and Technology*. 2025;36. https://doi.org/10.1088/1361-6501/add201.
10) Nghiem V, Nguyen H, Hoang M. LEAF-YOLO: Lightweight Edge-Real-Time Small Object Detection on Aerial Imagery. *Intelligent Systems with Applications*. 2025;25:200484. https://doi.org/10.1016/j.iswa.2025.200484.
11) Areeb M, Kumar S, Patil B, Jadhav A, Cao J, Shetty A, et al. AI Guide Dog: Egocentric Path Prediction on Smartphone. *ArXiv*. 2025;abs/2501.07957. https://doi.org/10.48550/arXiv.2501.07957.
12) Areeb M, Kumar S, Patil B, et al. AI Guided Navigation for the Visually Impaired. *International Journal for Research in Applied Science and Engineering Technology*. 2025. https://doi.org/10.22214/ijraset.2025.70016.
13) Li Y, Li Q, Pan J, Zhou Y, Zhu H, Wei H, et al. SOD-YOLO: Small-Object-Detection Algorithm Based on Improved YOLOv8 for UAV Images. *Remote Sensing*. 2024;16:3057. https://doi.org/10.3390/rs16163057.
14) Yu Z, Zheng X, Yang J, Su J. Improved YOLOX-DeepSORT for Multitarget Detection and Tracking of Automated Port RTG. *IEEE Open Journal of the Industrial Electronics Society*. 2024;5:317–325. https://doi.org/10.1109/OJIES.2024.3388632.
15) Joshi R, Singh N, Sharma A, Burget R, Dutta M. AI-SenseVision: A Low-Cost Artificial-Intelligence-Based Robust and Real-Time Assistance for Visually Impaired People. *IEEE Transactions on Human-Machine Systems*. 2024;54:325–336. https://doi.org/10.1109/THMS.2024.3375655.
16) Phatangare S, Rajeshirke S, Mule P, Pawar T, Morey S. Navigation Assistance for Visually Impaired People using Machine Learning. *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. 2024;p. 1–8. https://doi.org/10.1109/ICCCNT61001.2024.10724290.
17) Ratheesh R, R S, A A, S D, Ak H. AI-Based Smart Visual Assistance System for Navigation, Guidance, and Monitoring of Visually Impaired People. *2024 Ninth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)*. 2024;p. 1–10. https://doi.org/10.1109/ICONSTEM60960.2024.10568710.
18) Murai Y, Tatsumi H. AI-Powered Walking Assistance System for the Visually Impaired: Estimating Navigable Areas Using Pedestrian Information and AI Cameras. *Proceedings of the 2024 16th International Conference on Education Technology and Computers*. 2024. https://doi.org/10.1145/3702163.3702181.
19) Gharghan S, Kamel H, Marir A, Saleh L. Smart Stick Navigation System for Visually Impaired Based on Machine Learning Algorithms Using Sensors Data. *Journal of Sensor and Actuator Networks*. 2024;13:43. https://doi.org/10.3390/jsan13040043.
20) Abidi M, Siddiquee A, Alkhalefah H, Srivastava V. A comprehensive review of navigation systems for visually impaired individuals. *Heliyon*. 2024;10. https://doi.org/10.1016/j.heliyon.2024.e31825.
21) Zhang H, Falletta N, Xie J, Yu R, Lee S, Billah S, et al. Enhancing the Travel Experience for People with Visual Impairments through Multimodal Interaction: NaviGPT, A Real-Time AI-Driven Mobile Navigation System. *The 2025 ACM International Conference on Supporting Group Work*. 2025. https://doi.org/10.1145/3688828.3699636.
22) Ajina A, Lochan R, Saha M, Showghi R, Harini S. Vision Beyond Sight: An AI-Assisted Navigation System in Indoor Environments for the Visually Impaired. *2024 International Conference on Emerging Technologies in Computer Science for Interdisciplinary Applications (ICETCS)*. 2024;p. 1–6. https://doi.org/10.1109/ICETCS61022.2024.10543550.
23) Abidi M, Alkhalefah H, Siddiquee A. Enhancing Navigation and Object Recognition for Visually Impaired Individuals: A Gradient Support Vector Boosting-based Crossover Golden Jackal Algorithm Approach. *Journal of Disability Research*. 2024. https://doi.org/10.57197/jdr-2024-0057.
24) Wang W, Jing B, Yu X, Sun Y, Yang L, Wang C. YOLO-OD: Obstacle Detection for Visually Impaired Navigation Assistance. *Sensors (Basel)*. 2024;24. https://doi.org/10.3390/s24237621.