Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Smart Mobility Assistant for the Visually Impaired: Integrating Machine Learning and Sensor Technology

Shruti Mallikarjun [1], Dr. Sheetalrani R Kawale [2]
[1] *Research Scholar, Department of Computer Science, Karnataka State Akkamahadevi Women's University, Vijayapura, Karnataka, India. Email: shrutipasargi290@gmail.com, ORCID ID: 0009-0009-6326-4859*
[2] *Assistant Professor, Department of Computer Science, Karnataka State Akkamahadevi Women's University, Vijayapura, Karnataka, India. Email: sheetalrani@kswu.ac.in*

ARTICLE INFO

Received: 22 Dec 2024
Revised: 14 Feb 2025
Accepted: 26 Feb 2025

ABSTRACT

Facilitating the safety and autonomy of visually impaired persons in traversing their environments is an essential goal. This work presents the development of a comprehensive assistive kit that leverages Raspberry Pi technology to empower visually impaired users. The kit integrates ultrasonic sensors for obstacle detection, machine-learning models for object recognition, voice alerts, facial recognition for person identification, and indoor localization using RSSI (Received Signal Strength Indicator) modules. The core functionality of the system lies in real-time environmental monitoring through ultrasonic sensors, which detect obstacles in the user's path. Upon detection, a pre-trained deep learning model identifies the nature of the obstacle—whether it be a chair, door, or other objects - and delivers auditory alerts to the user via voice synthesis. The system also employs facial recognition algorithms to identify familiar individuals, enabling users to confidently recognize and interact with acquaintances.

Keywords: object detection, deep learning, obstacles, Visually Impaired, GPS module, ResNet.

INTRODUCTION

Computer vision is an extensively studied domain, focused on equipping computers to understand and analyze intricate visual information. The identification and localisation of objects of interest in photos or videos is tough in this subject [1-3]. The popularity of artificial neural networks, multilayer perceptrons, and support vector machines propelled deep learning in the early 2000s. The scalability issues and processing resource constraints of deep learning initially limited its adoption [4-6]. Since 2006, large datasets and powerful processing power have boosted deep learning. Computer vision object detection recognises and localises objects in pictures and videos. The main goal is to reliably recognise, position, and size objects in an image or video and classify them [7-10]. This technique is used to predict stock values, recognize speech, detect objects, recognize characters, detect intrusions, detect landslides, solve time series problems, classify text, express gene expression, create microblogs, handle data, classify irregular data with fault-classification, caption text from images, and perform aspect-based sentiment analysis [11-16]. Real-world object identification models employ multiple approaches and deep learning architectures to classify objects. Object detection works on images, movies, and sounds. Computer and software systems can identify things in images. Object detection in video closely resembles its function in photos. This technology would allow the computer to identify, recognize, and classify objects present in the supplied video footage. Object detectors may recognize things based on several auditory cues as well. Machine learning algorithms can automatically detect and locate objects in images and movies. These models recognise visual objects using feature extraction, selection, and classification. These models are trained using annotated pictures containing class-specific items of interest. The program then learns category-specific traits from these annotated photographs. SVM, decision trees, and random forests are object recognition machine learning models [17]. These models identify features differently and perform various tasks based on the work and data. Some models need human feature engineering, whereas others may automatically gather features from input data. Obstacles affect visually impaired people's quality of life and everyday activities. The conventional approach involves that of a supporting guide. An individual is required to adeptly navigate them from

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

one location to another [18]. The white cane is an effective tool for negotiating obstacles in close proximity to the individual. It may only assist in identifying obstructions that are in close proximity to the individual. This raises the concern over the safety of those who are blind or visually challenged. Independent mobility is likewise a source of worry. An object detecting system is essential to address safety concerns and facilitate autonomous movement [19]. The object detecting system should be straightforward and, more significantly, user-friendly. It should also be inexpensive. Modern technologies such as computer vision, item detection, and recognition are available. Smartphone cameras enable object detection. It then recognises the stuff in front of the camera [20]. Real-time object identification in video streams using YOLO V3. The user receives an audio message about the item. It may be done using headphones or speakers. This software is for Android smartphones. Mobile phones with cameras enable the blind and visually challenged make calls. The software aims to recognise the camera's subject. Blind persons may benefit from computer vision-based object detection systems. This simplified method is user-friendly. It is easy and affordable. This addresses the challenges of a costly, supplementary advanced system. Consequently, the contributions of this study are as follows:

- To adopt innovative assistive devices tailored specifically for visually impaired individuals. By integrating multiple NodeMCUs with sensors, GPS, and communication features (voice command in indoor and outdoor), it provides a practical and scalable solution to enhance mobility and safety.
- The NodeMCU module on a blind stick, which calculates the Received Signal Strength Indicator (RSSI) values to determine the user's location within a home.
- By integrating the YOLOv8n framework with ResNet, allowing for the training of deeper networks without suffering from vanishing gradient issues. YOLOv8n's lightweight design renders it appropriate for implementation on edge devices with constrained processing capabilities, such smartphones and embedded systems, therefore broadening the accessibility of object detection technologies.

This is the arrangement of the article. Section 2 provides specifics on the literature review. Section 3 describes the proposed method in detail. The Comparative analysis is used to investigate performance and results in Section 4. Section 5 provides conclusion with recommendations.

RELATED WORKS

Numerous devices have been created to establish a system that offers visual assistance, aiding those with vision impairments in their everyday activities by alerting them to impediments in their direction. This literature study delineates the results and constraints of several available assistive solutions for those who are blind or visually impaired.

In [21], the CNN approach is used for optimum object detection. The gathered items will undergo machine learning training to provide a model for integration into the system's primary device, namely the Raspberry Pi 4B. It attains an object detection accuracy of 84.3%. In [22], the Yolov5 (You Only Look Once, Version 5) DL model utilizes a camera to recognize and monitor individuals with disabilities in real-time. An effective and inclusive use of technology might be a ubiquitous system that uses deep learning-based YOLO models to identify individuals with disabilities. GAN and CNN are used to recognise objects in [23]. HAAVO optimises DCNN classifier and DRN distance estimation training. HAAVO combines HBA, Adam Optimiser, and African Vultures Optimisation. Its accuracy is 0.940. In [24], DeepNAVI uses deep learning for smartphone navigation. Our system can display obstacles' kind, position, proximity to the user, movement status, and contextual scene information. Image and video processing methods for real-time camera inputs were developed in [25]. Deep Neural Networks in combination with Google's Text-To-Speech (GTTS) API module predicted items and places, identifying their class and position. A deep CNN model in [26] achieves 83.3% accuracy on a dataset of over 1000 categories. The gadgets' capabilities are used to perform a quantitative score comparison research. A CNN-based image identification model employing an SSD and a pretrained model from Mobile V2 is built at Google dataset in [27] using the TensorFlow object API. In [28], a smartphone navigation tool for blind or visually impaired persons used EfficientNet-Lite-based scene identification. In [29], visually impaired persons were able to navigate using a smart stick and a vision-based, three-dimensional

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

wearable model. The GPS-enabled smart stick and neural network-based YOLO algorithm in the wearable model determined the navigational path. The limitations across these studies primarily revolve around the lack of detailed performance metrics and broader contextual insights. While the models described, such as CNNs, YOLO, GANs, and hybrid optimization techniques, show promising applications for object detection, real-time tracking, and navigation assistance, there is a noticeable absence of comprehensive quantitative data. Many of the works do not specify the exact accuracy, precision, or recall rates in a way that allows for clear comparison across methods. Additionally, there is often a lack of information regarding the datasets used, such as their size, diversity, or real-world applicability, which makes it difficult to assess how these systems would perform in varied or complex environments.

SYSTEM MODEL

The figure-1 illustrates a process flow for a video surveillance system that uses a machine learning model, specifically YOLO (You Only Look Once), to detect and track objects. It begins with an image/video dataset that undergoes preprocessing to prepare the data for the next steps. Preprocessing might involve operations like noise reduction, normalization, or resizing of images to ensure compatibility with the neural network model. Subsequent to preprocessing, the data is input into a training network where the YOLO model undergoes training using machine learning methodologies. The model recognises and tracks objects in real time in video surveillance applications after training. We then assess the system's performance using True Positive, True Negative, FP, and FN. Precision, recall, and the confusion matrix evaluate the model's accuracy and usefulness in actual surveillance circumstances.

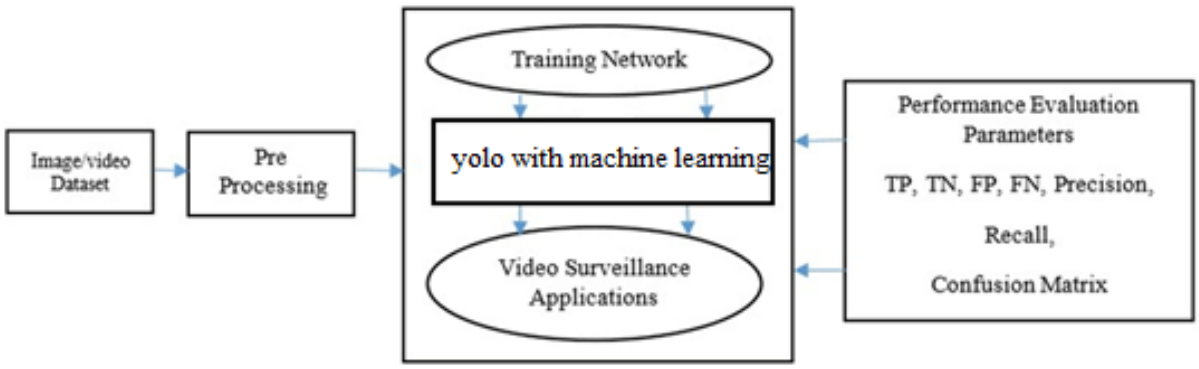

Figure 1. Process flow for a video surveillance system

3.1 Dataset Description

Collecting images for the ten target objects [30] (Suitcase-Luggage, Switch-Box, Bottle, Dustbin, Mirror, Person, Staircase, Stove, Toilet, and Toothbrush) using the Roboflow platform. Each image must contain at least one of the ten target objects. The objects should be clearly visible and identifiable in the images, without significant obstructions or occlusions. Under Indoor and Outdoor Settings, we can collect images from both indoor and outdoor environments to cover different use cases (e.g., suitcases in outdoor, staircases in buildings, toilets in homes).

3.2. Hardware Configurations

The figure-2 a system architecture centred around a Raspberry Pi, which functions as the main processing unit. This setup includes several essential components, such as a power supply to ensure reliable operation of the Raspberry Pi and its peripherals. An ultrasonic sensor is integrated to detect distance and proximity, which is crucial for applications like robotics and environmental monitoring.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

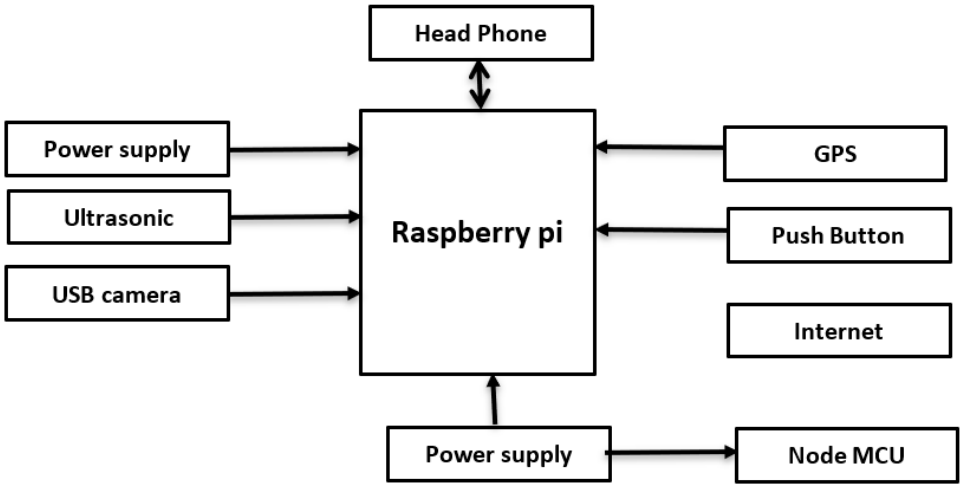

Figure-2 Hardware representation of Smart Movement and Control Assistant

A USB camera is also included, enabling visual data capture for tasks such as surveillance or object detection. The GPS module provides geographical location data, essential for tracking and location-based services. Additionally, a push button serves as a simple user interface, allowing for manual interaction to trigger actions or alerts. The inclusion of a Node MCU adds Wi-Fi capabilities, facilitating internet connectivity for remote monitoring or control. Finally, headphones are incorporated for audio feedback or communication, allowing users to receive alerts or interact with the system audibly. Overall, this combination of components indicates a versatile system designed for various applications, including smart home automation, robotics, and environmental monitoring, enhancing interactivity and data collection in modern contexts.

- **NodeMCU Modules**: Three NodeMCU modules are strategically placed in key locations within the home—namely the bedroom, hall, and kitchen. These modules continuously broadcast signals that are detected by a fourth NodeMCU module attached to a blind stick, which the user carries.
- **RSSI-Based Localization**: The NodeMCU module on the blind stick receives signals from the fixed modules in different rooms. By calculating the RSSI values, the system determines the distance between the blind stick and each of the fixed modules. The module with the strongest signal (i.e., the highest RSSI value) is inferred to be the nearest to the user, thus identifying the user's current location within the home. The system uses RSSI measurements to compute the distance between the user and each room's NodeMCU. The relationship between RSSI and distance is governed by the following equation: Where:

$$\text{Distance} = 10^{\left(\frac{A-\text{RSSI}}{10 \cdot n}\right)}$$

- RSSI is the measured signal strength from the Wi-Fi network,
- A represents the reference RSSI value at a distance of 1 meter (typically around -50 dBm),
- n is the path loss exponent, which varies depending on the environment (commonly between 2 and 4).

3.3. Pre-Processing of Dataset

The dataset employs an annotation approach for each item denoted by an oriented bounding box (OBB). The orientated bounding box (OBB) format is (x1, y1, x2, y2, x3, y3, x4, y4, categories, complex) and the i-th vertex coordinates are (xi, yi). The object's bounding box is defined by clockwise vertices. This research presents an innovative pre-processing technique that may be used for the purpose of training the YOLOv8 object detection system. The noise reduction, data presentation, and standardization processes are all included in the preprocessing. These actions improve the effectiveness of the YOLOv8 dataset.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

- Noise handling: Dataset noise is addressed at the first stage of the pre-processing process. The collection is comprised of both images and labels. Label files include item addresses, class names, and unnecessary text that may noisy the dataset. Regular expressions reduce unnecessary dataset strings.

3.4. YOLOv8n with Machine Learning Model

The input image is divided into grid cells using the YOLOv8 model, which then forecasts bounding boxes and class probabilities for each grid cell respectively. The model divides input into N grid cells. It predicts the bounding box location using four coordinates (x, y, w, h) for each cell i and matching bounding box j and a confidence score c. The vector P encodes the class probabilities. The anticipated coordinates of the bounding box, $((x'^{j}_{i}, y'^{j}_{i}, w'^{j}_{i}, h'^{j}_{i})$, are computed based on the subsequent equations.

$$x'^{j}_{i} = \sigma\left(b^{j}_{x,i}\right) + i \quad (1)$$

$$y'^{j}_{i} = \sigma\left(b^{j}_{y,i}\right) + i \quad (2)$$

$$w'^{j}_{i} = p^{j}_{w,i} . e^{b^{j}_{w,i}} \quad (3)$$

$$h'^{j}_{i} = p^{j}_{h,i} . e^{b^{j}_{h,i}} \quad (4)$$

The confidence score c_i^h for each bounding box is determined by the sigmoid function σ, predicted parameters $b^{j}_{x,i}, b^{j}_{y,i}, b^{j}_{w,i}, b^{j}_{h,i}$, and anchor box dimensions $p^{j}_{w,i}, p^{j}_{h,i}$

$$c^{h}_{i} = \sigma(b^{j}_{c,i}) \quad (5)$$

The expected confidence parameter is b(c,i)^j. It calculates class probabilities pi using softmax activation function. Deep transfer learning feature extraction network ResNet-50.

The proposed feature extraction approach uses ResNet-50 as a deep transfer model. ResNets are residual neural networks for deep transfer learning. 16 bottleneck blocks remain in ResNet-50. As shown in Fig. 8 (ResNet section), each block has convolutional sizes of 1x1, 3x3, and 1x1, employing feature maps of 64, 128, 256, 512, 1024. The detection network (YOLO v2) is a multi-layer convolutional neural network with a transformation and output layer. The transform layer stabilises deep neural networks by pulling activations from the convolutional layer. The transform layer converts bounding box predictions into target box contours. Unmodified goal bounding box coordinates are produced by the output layer. The YOLO v2 loss function determines the mean squared error loss between predicted bounding boxes and the target,

$$\text{YOLOv8n}_{\text{loss}} = \text{localization loss} + \text{confidence loss} + \text{classification loss} \quad (6)$$

Localisation loss measures the target-expected bounding box difference. The localisation loss coefficients include the bounding box's grid cell width (w) and height (h). The localisation loss coefficients are calculated using equation (7).

$$\text{localization loss} = \text{q1} \sum_{\text{a=0}}^{\text{g}^2} \sum_{\text{b=0}}^{\text{v}} 1_{\text{ij}}^{\text{obj}} [(\text{x}_\text{i} - \text{x'}_\text{i})^2 + (\text{y}_\text{i} - \text{y'}_\text{i})^2] + \text{q1} \sum_{\text{a=0}}^{\text{g}^2} \sum_{\text{b=0}}^{\text{v}} 1_{\text{ij}}^{\text{obj}} [(\sqrt{\text{w}_\text{i}} - \sqrt{\text{w'}_\text{i}})^2 + (\sqrt{\text{h}_\text{i}} - \sqrt{\text{h'}_\text{i}})^2 \quad (7)$$

Where q1 is a weight, g is the grid cell count, and v is the bounding box count within each g. The centre of v in $g(x_i, y_i)$. The width and height of v in g are denoted by (w_i, h_i), whereas the centre of the target is represented by (x'_i, y'_i) and the centre of the target is represented by (w'_i, h'_i). The object is present in v in g if $1_{ij}^{obj}$is 1; otherwise, it is When an item is recognised within g's bounding box, confidence loss determines the error's confidence score. The confidence loss coefficients calculates by Equation (8),

$$\text{confidence loss} = \text{q2} \sum_{\text{a=0}}^{\text{g}^2} \sum_{\text{b=0}}^{\text{v}} 1_{\text{ij}}^{\text{obj}} (\text{cs}_\text{i} - \text{cs'}_\text{i})^2 + \text{q3} \sum_{\text{a=0}}^{\text{g}^2} \sum_{\text{b=0}}^{\text{v}} 1_{\text{ij}}^{\text{noobj}} (\text{cs}_\text{i} - \text{cs'}_\text{i})^2 \quad (8)$$

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Where, q2 and q3 represent confidence error weights, cs_i represents 's confidence score in g, cs'_i represents the target's confidence score in g, 1_{ij}^{obj} = 1 when an object is present in v in each g, otherwise = 0. Conversely, 1_{ij}^{obj} equals 1 when no item is present in v within each g; otherwise, it equals 0. The Classification loss is the grid cell i class conditional probability difference. The classification loss parameters by equation 9,

$$clasification\ loss = q4 \sum_{a=0}^{g^2} 1_{ij}^{obj} \sum_{zeclasses} (p_i(z) - p'_i(z))^2 \quad (9)$$

Where $q4$ represents the weight of the classification mistake $p_i(z)$ and $p'_i(z)$ denotes the estimated and real conditional probabilities of an item inside grid cell a.

PERFORMANCE ANALYSIS

Experimental setup - The training process spanned 100 epochs, allowing the model ample opportunity to learn and refine its understanding of the features and patterns associated with each object class. This extended training duration helped in improving the model's accuracy and robustness. The batch size of 16 was utilised to find a compromise between computational effort and the capacity to generalise successfully from the training data. It ensures that each training step is efficient, leveraging the parallel processing capabilities of the Tesla T4 GPU, while also providing enough variability in each batch to enhance learning.

Performance metrics- It is crucial to evaluate the model in order to ascertain its compatibility with the research objectives and its performance. There are numerous evaluation metrics that can be employed in the detection of bus objects, including those that pertain to the model's efficiency, speed, and accuracy. The primary focus of bus object detection is typically on evaluation metrics that pertain to accuracy, as the model must be capable of accurately identifying buses. The performance are analysed by precision (P) and recall (R). Table 1 shows the evaluation parameters. These parameters are compared to Convolutional Neural Network [21], GAN [23].

Table-1 Performance Metrics

Performance matrix	Formula	description
accuracy	$ac = \frac{TP + TN}{TP + TN + FP + FN}$	Where, $pt = true\ positive\ rate$ $pf = false\ positive\ rate$ $nf = false\ negative\ rate$ $nt = true\ negative\ rate$
recall	$re = \frac{pt}{pt + nf}$	
precision	$pr = \frac{nt}{nt + pf}$	
F1-score	$ppv = \frac{pt}{pt + pf}$	

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

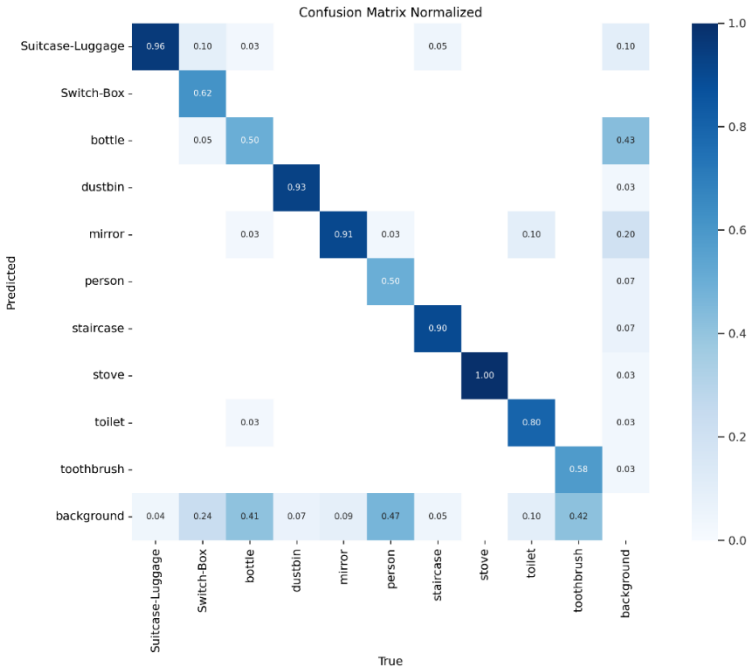

Figure-3 Confusion Matrix

The figure-3 is a normalized confusion matrix for a machine learning model tasked with classifying images into ten categories: Suitcase-Luggage, Switch-Box, Bottle, Dustbin, Mirror, Person, Staircase, Stove, Toilet, Toothbrush, and a Background class. The confusion matrix is normalized, meaning that the values are proportions rather than raw counts, with each row summing up to 1.0.

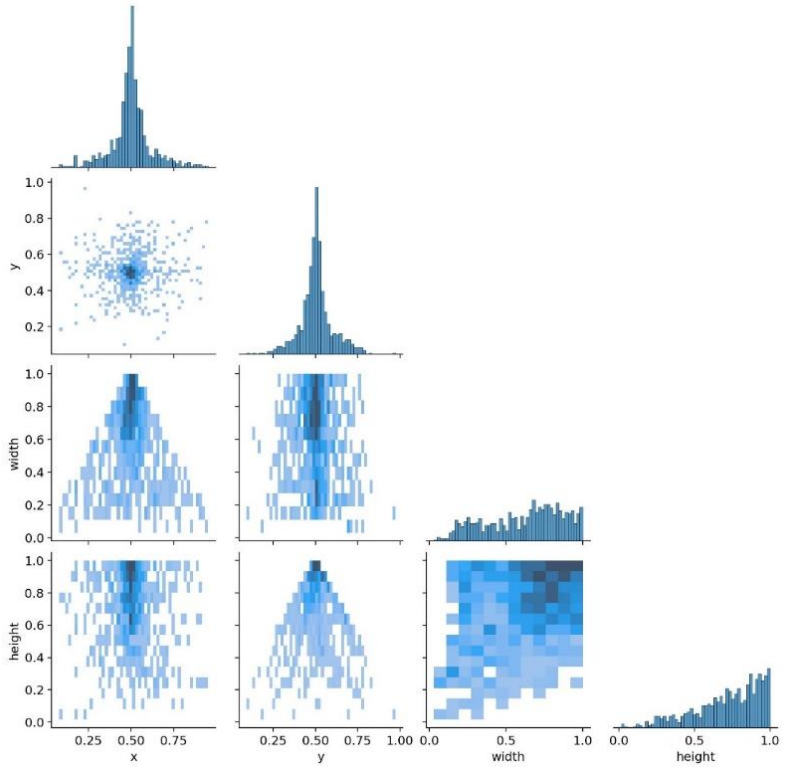

Figure-4 Label Correlogram

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The figure-4 is a correlogram, specifically a pair plot, displaying the relationships between four variables related to object detection bounding boxes: x, y, width, and height. These variables usually reflect the bounding box centre coordinates and dimensions in normalised image space. It reveals that bounding box centers are typically located near the center of the image and that there is considerable variability in the sizes of the bounding boxes. This information is crucial for understanding the dataset's characteristics and for guiding further model improvements, such as data augmentation strategies to handle the variability in object sizes and positions.

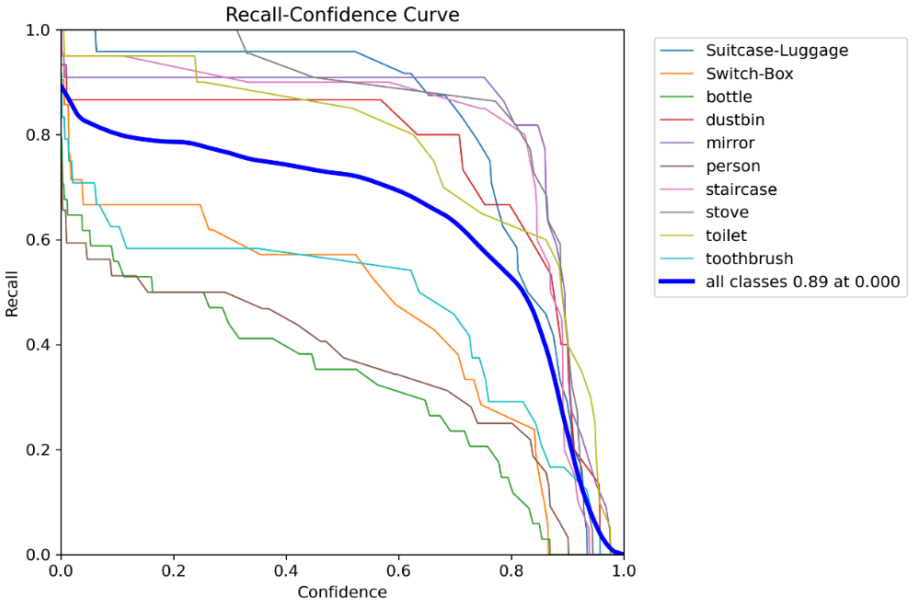

Figure-5 Recall-Confidence Curve

The figure-5 is a Recall-Confidence curve for a YOLOv8n object detection model trained to detect ten different object categories. A recall of 1.0 means that the model correctly detected all instances of the object. The x-axis represents the confidence threshold, which is the model's predicted probability that an object belongs to a particular class. A higher confidence threshold means the model is more certain about its predictions, but it might miss some true positives,

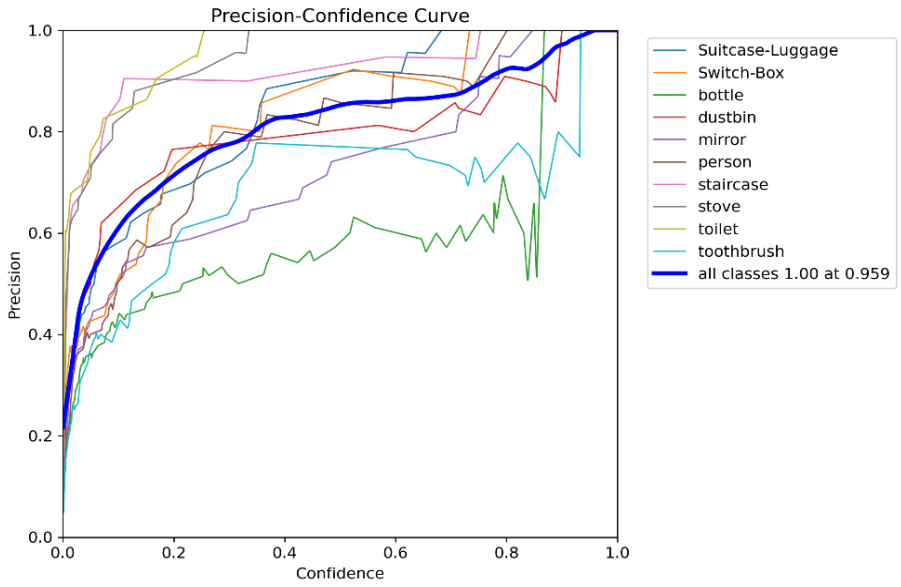

Figure- 6 Precision-Confidence Curve

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

This figure-6 is Precision-Confidence Curve, often used in evaluating the classification models, particularly in object detection tasks. The overall performance can be assessed by looking at the thick blue line. A higher and more consistent curve suggests that the model is performing well across different classes

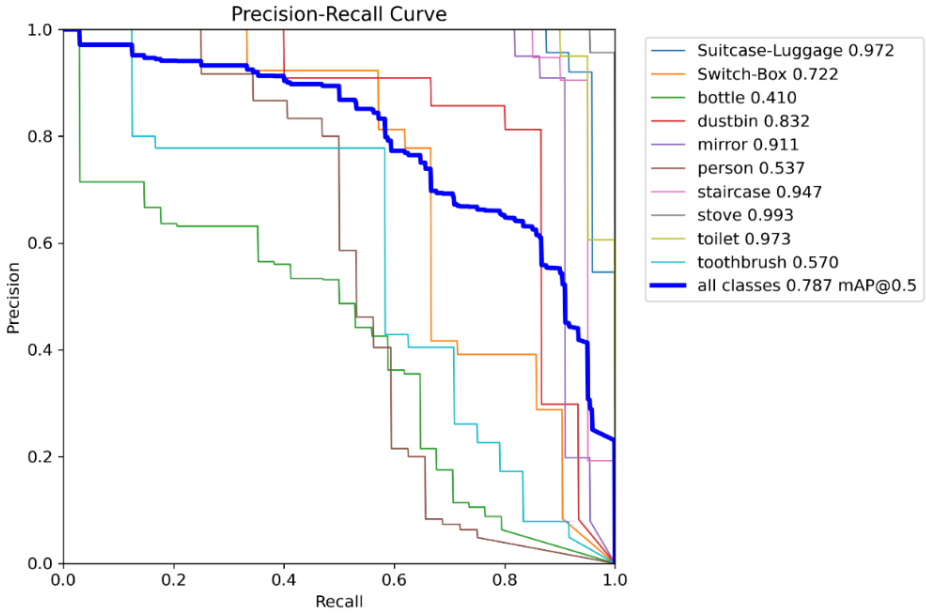

Figure-7 Calculation of PR curve

Figure-7 shows a Precision-Recall (PR) curve, which is used to test classification models, notably for unbalanced datasets. At a threshold of 0.5, the mean Average Precision (mAP) score is 0.787, and the thick blue line shows the total performance across all classes. The better the model performs for that class, the closer the curve is to the plot's upper-right corner. The area under the curve shows the model's ability to balance recall and accuracy.

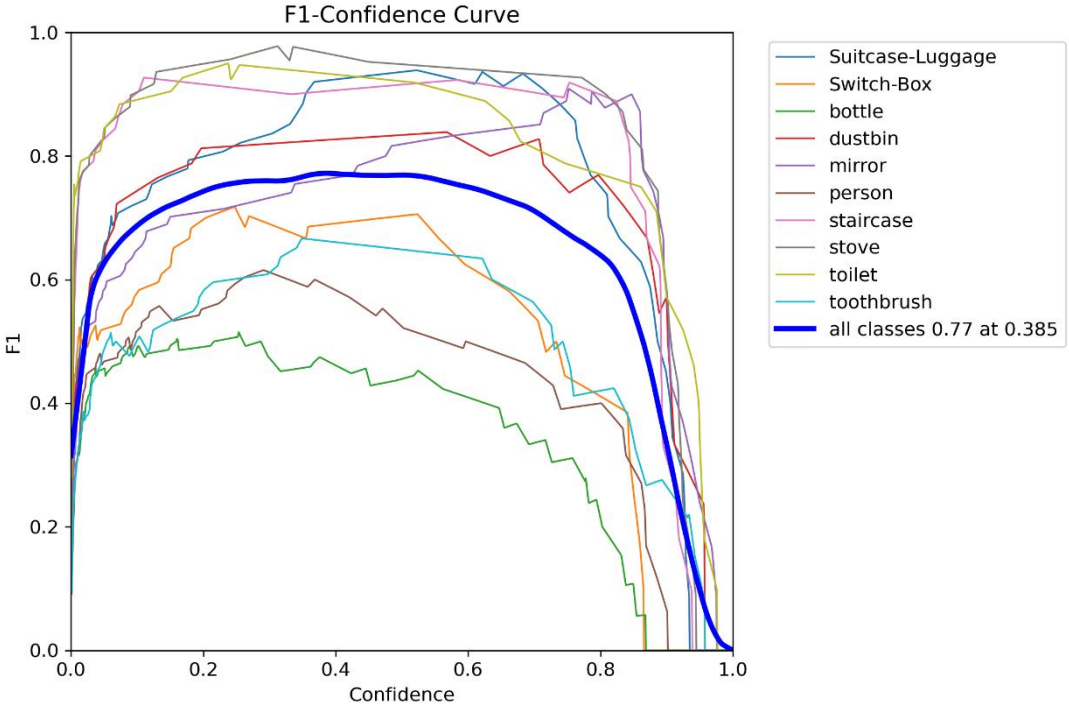

Figure-8 F1-Confidence Curve

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The model's estimated probability that an item belongs to a certain class is shown by the x-axis, which is similar to the Recall-Confidence curve in figure 8. The F1 score is the harmonic mean of accuracy and memory, therefore it may balance both variables.

Table-2 Comparison of existing and proposed method

Model	Accuracy	Precision	Recall
[21]	-	80%	68.2%
[22]	73.26 %,	72.84%	73.47%
Proposed model	78.3	82.9	74.2

CONCLUSION

This substantial Raspberry Pi-powered assistive package improves visually impaired people's safety and autonomy. By integrating ultrasonic sensors for obstacle detection, deep learning models for object recognition, voice alerts, facial recognition for person identification, and RSSI-based indoor localization, the system addresses visually challenged users' special needs. Real-time environmental monitoring and adaptive navigation help users move more confidently, while the inclusion of GPS-based navigation ensures that they remain informed about their surroundings, especially in emergencies. All things considered, this integrated strategy helps visually impaired people live better lives and be more accessible while also increasing their mobility and autonomy. Future research is on forming alliances with healthcare providers to include the assistive gear with health monitoring systems, enabling users to communicate their location and health condition to carers or medical professionals immediately in an emergency.

Declaration of Interest Statement:

- The authors declare that there are no conflicts of interest regarding the publication of this article. This research was conducted without any financial support from external funding sources.
- The authors have no financial relationships or affiliations that could be perceived as influencing the work reported in this manuscript.
- All authors have contributed equally to the design, implementation, and analysis of the results.

Funding Information:

Not Available OR Not Applicable**.**

Author Contribution:

1. Dr. Sheetalrani R Kawale[1]: Review

2. Shruti Mallikarjun[2]: Research, Data Collection, Analysis.

Research Involving Human and /or Animals:

This research is entirely technology-based, and no experiments were conducted involving humans or animals.

Informed Consent:

This study adhered to ethical guidelines and obtained informed consent from all participants involved. Participants were fully informed about the purpose, procedures, potential risks, and benefits of the study, as well as their right to withdraw at any time without consequences. Written consent was obtained and documented for all participants prior to their involvement in the research.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Data Availability Statement:

No new Data Set: This study did not generate any new datasets; all data analyzed are available in the referenced publications.

Ethical Statement:

The research presented in this manuscript adheres to the ethical principles and guidelines for scientific research and publication. The study was conducted with the primary objective of advancing accessibility and safety for visually impaired individuals through the development of a smart mobility assistant.

No human participants or animals were involved in the experimental procedures, ensuring compliance with ethical standards. All data utilized in this study were either publicly available or generated using simulated environments. Strict measures were taken to ensure the integrity, security, and confidentiality of any data processed during the research.

The authors confirm that this work does not involve any conflicts of interest, plagiarism, or fraudulent data. The manuscript is original and has not been submitted to other journals for publication. Additionally, the outcomes of this research are intended solely for academic and technological advancements without any misuse or harm to individuals or communities.

The authors are committed to responsible innovation, ensuring the developed technology promotes inclusivity and benefits society.

REFERENCE

[1] Rather, A.M., Agarwal, A., Sastry, V.N.: Recurrent neural network and a hybrid model for prediction of stock returns. Expert Syst. Appl. 42(6), 3234–3241 (2015)

[2] Sak, H., Senior, A., Rao, K., Beaufays, F.: Fast and accurate recurrent neural network acoustic models for speech recognition. arXiv preprint arXiv:1507.06947 (2015)

[3] Liang, M., Hu, X.: Recurrent convolutional neural network for object recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, pp. 3367–3375. (2015) https://doi.org/10.1109/CVPR.2015. 7298958

[4] Zhang, X.Y., Yin, F., Zhang, Y.M., Liu, C.L., Bengio, Y.: Drawing and recognizing Chinese characters with recurrent neural network. IEEE Trans. Pattern Anal. Mach. Intell. 40(4), 849–862 (2017)

[5] Kim, J., Kim, J., Thu, H.L.T., Kim, H.: Long short term memory recurrent neural network classifier for intrusion detection. In: 2016 International Conference on Platform Technology and Service (PlatCon), Jeju, Korea, pp. 1–5. (2016). https://doi.org/ 10.1109/PlatCon.2016.7456805

[6] Mezaal, M.R., Pradhan, B., Sameen, M.I., Shafri, M., Zulhaidi, H., Yusoff, Z.M.: Optimized neural architecture for automatic landslide detection from high resolution airborne laser scanning data. Appl Sci 7(7), 730 (2017). https://doi.org/10.3390/ app7070730

[7] Swamy, S.R., Praveen, S.P., Ahmed, S., Srinivasu, P.N., Alhumam, A.: Multi-features disease analysis based smart diagnosis for COVID-19. Comput. Syst. Sci. Eng. 45, 869–886 (2023)

[8] Lai, S., Xu, L., Liu, K., Zhao, J.: Recurrent convolutional neural networks for text classification. In: Twenty-ninth AAAI Conference on Artificial Intelligence (2015), Austin, Texas, USA.

[9] Quang, D., Xie, X.: DanQ: a hybrid convolutional and recurrent deep neural network for quantifying the function of DNA sequences. Nucleic Acids Res 44(11), e107–e107 (2016). https:// doi.org/10.1093/nar/ gkw226

[10] Arava, K., Chaitanya, R.S.K., Sikindar, S., Praveen, S.P., Swapna, D.: Sentiment analysis using deep learning for use in recommendation systems of various public media applications. In: 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), pp. 739–744. IEEE (2022)

[11] Liao, S., Wang, J., Yu, R., Sato, K., Cheng, Z.: CNN for situations understanding based on sentiment analysis of twitter data. Procedia Comput. Sci. 111, 376–381 (2017). https://doi.org/10. 1016/j.procs.2017.06.037

[12] Wei, D., Wang, B., Lin, G., Liu, D., Dong, Z., Liu, H., Liu, Y.: Research on unstructured text data mining and fault classification based on RNN-LSTM with malfunction inspection report. Energies 10(3), 406 (2017). https://doi.org/10.3390/en10030406

[13] Sirisha, U., Bolem, S.C.: Semantic interdisciplinary evaluation of image captioning models. Cogent Eng. 9(1), 2104333 (2022)

[14] Sirisha, U., Bolem, S.C.: GITAAR-GIT based Abnormal Activity Recognition on UCF Crime Dataset. 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT). IEEE (2023)

[15] Sirisha, U., Bolem, S.C.: Aspect based sentiment & emotion analysis with ROBERTa, LSTM. Int. J. Adv. Comput. Sci. Appl. (2022). https://doi.org/10.14569/IJACSA.2022.0131189

[16] Xu, N., Liu, A.A., Wong, Y., Zhang, Y., Nie, W., Su, Y., Kankanhalli, M.: Dual-stream recurrent neural network for video captioning. IEEE Trans. Circuits Syst. Vid Technol. 29(8), 2482– 2493 (2018). https://doi.org/10.1109/TCSVT.2018.2867286

[17] Thai, L.H., Hai, T.S., Thuy, N.T.: Image classification using support vector machine and artificial neural network. Int. J. Inform. Technol. Comput. Sci. 4(5), 32–38 (2012)

[18] Chen LB, Su JP, Chen MC, et al (2019) An Implementation of an intelligent assistance system for visually impaired/blind people. In: 2019 IEEE International conference on consumer electronics ICCE 2019, pp 4–5. https://doi.org/10.1109/ICCE.2019.8661943

[19] Lin Y, Wang K, Yi W, Lian S (2019) Deep learning based wearable assistive system for visually impaired people. In: Proceedings of 2019 International conference on computer vision workshop ICCVW 2019, pp 2549–2557. https://doi.org/10.1109/ICCVW.2019.00312

[20] Matusiak K, Skulimowski P, Strurnillo P (2013) Object recognition in a mobile phone application for visually impaired users. In: 2013 6th international conference on human system interaction (HSI 2013), vol 17, pp 479–484. https://doi.org/10.1109/HSI.2013.6577868

[21] Parenreng, J. M., Kaswar, A. B., & Syahputra, I. F. (2024). Visual Impaired Assistance for Object and Distance Detection Using Convolutional Neural Networks. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, *8*(1), 26-32.

[22] Alruwaili, M., Siddiqi, M. H., Atta, M. N., & Arif, M. (2024). Deep learning and ubiquitous systems for disabled people detection using YOLO models. *Computers in Human Behavior*, *154*, 108150.

[23] Nagarajan, A., & Gopinath, M. P. (2023). Hybrid optimization-enabled deep learning for indoor object detection and distance estimation to assist visually impaired persons. *Advances in Engineering Software*, *176*, 103362.

[24] Kuriakose, B., Shrestha, R., & Sandnes, F. E. (2023). DeepNAVI: A deep learning based smartphone navigation assistant for people with visual impairments. *Expert Systems with Applications*, *212*, 118720.

[25] Hussan, M. I., Saidulu, D., Anitha, P. T., Manikandan, A., & Naresh, P. (2022). Object detection and recognition in real time using deep learning for visually impaired people. *International Journal of Electrical and Electronics Research*, *10*(2), 80-86.

[26] Ashiq, F., Asif, M., Ahmad, M. B., Zafar, S., Masood, K., Mahmood, T., ... & Lee, I. H. (2022). CNN-based object recognition and tracking system to assist visually impaired people. *IEEE access*, *10*, 14819-14834.

[27] Nasir, H. M., Brahin, N. M. A., Aminuddin, M. M. M., Mispan, M. S., & Zulkifli, M. F. (2021). Android based application for visually impaired using deep learning approach. *IAES International Journal of Artificial Intelligence*, *10*(4), 879.

[28] Kuriakose, B., Shrestha, R., & Sandnes, F. E. (2021, October). SceneRecog: a deep learning scene recognition model for assisting blind and visually impaired navigate using smartphones. In *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 2464-2470). IEEE.

[29] Kumar, N., & Jain, A. (2022). A Deep Learning Based Model to Assist Blind People in Their Navigation. *J. Inf. Technol. Educ. Innov. Pract.*, *21*, 95-114.

[30] https://universe.roboflow.com/blind-people-object-detection-klpbg/blind-people-object-detection-iuhho