Journal of Engineering Science and Technology Review 18 (1) (2025) 110-119

Research Article

JOURNAL OF
Engineering Science and Technology Review

www.jestr.org

Smart Blind Assistant Stick

Amey Mali, Shubhangi Kharche*, Anamika Nevase, Parvathy Nair and Chinmay Vaity

Department of Electronics & Computer Science, SIES Graduate School of Technology, Nerul, India

Received 26 June 2024; Accepted 23 February 2025

Abstract

Globally, vision impairment poses significant challenges, especially for those who are blind, leading to their daily struggles. Our proposed system proposes an innovative system that combines computer vision and Internet of Things technologies to empower visually impaired people during navigation, with the goal of increasing their mobility and independence. Conventional camera-based solutions frequently fail to accurately recognize objects and obstacles, as well as achieve precise distance measurements. In response, our proposed work introduces a smart blind stick assistance system that aims to address these limitations. Using the COCO dataset supplemented with 80 standard objects and an additional 15, our system employs YOLOv5 for real-time object detection, allowing accurate identification of obstacles in the environment. Most importantly, real-time scene descriptions are delivered audibly with multilingual support, allowing users to understand spatial concepts more effectively. The implementation includes a seamless integration of the Raspberry Pi and Chainer library for efficient neural network inference, as well as text-to-speech synthesis to provide spoken feedback. Additionally, we have integrated an ultrasonic sensor for precise distance measurement, enhancing the system's ability to detect obstacles and provide accurate navigation assistance. During our tests, the system achieved 95% object detection accuracy (92-98% confidence interval, n=10) using an augmented COCO dataset and provides real-time, multilingual scene descriptions (average processing time 50ms ± 10ms; ultrasonic sensor latency 2ms ± 0.5ms). This novel approach represents a significant advancement in object recognition and accessibility for visually impaired people, promising to improve their quality of life and promote greater independence.

Keywords: Vision impairment, Blindness, Computer vision, Internet of Things (IoT), Object recognition, Real-time object detection, Audible scene descriptions

1. Introduction

Considering the global prevalence of visual impairments, addressing the challenges faced by the visually impaired through technology is imperative. With approximately 285 million affected individuals worldwide, including 39 million living with total blindness [1], there's a pressing need to leverage innovations to enhance their lives. Vision impairment poses significant challenges, particularly for those who are blind, leading to daily struggles in navigation and interaction with their environment. Traditional aids, such as canes and guide dogs, although valuable, have limitations and do not fully address the complexities of modern, fast-paced environments. Our proposed system introduces an innovative solution that combines computer vision and Internet of Things (IoT) technologies to empower visually impaired people during navigation, with the goal of increasing their mobility and independence. The integration of these advanced technologies aims to create a comprehensive tool that not only detects obstacles but also provides detailed, real-time information about the surrounding environment. This system, referred to as the Smart Blind Stick Assistance System, marks a significant advancement in assistive technology for the visually impaired. Conventional camera-based solutions frequently fail to accurately recognize objects and obstacles and achieve precise distance measurements. These systems often struggle with varying lighting conditions, cluttered environments, and the dynamic nature of outdoor settings. To address these limitations, our project employs the COCO (Common Objects in Context) dataset, which has been enriched with 80 standard objects and an additional 15 items specifically relevant to the visually impaired. This comprehensive dataset enables the system to identify a wide range of objects accurately. At the core of our system is the YOLOv5 (You Only Look Once version 5) algorithm, known for its efficiency and accuracy in real-time object detection. YOLOv5's ability to process images and detect objects at high speed makes it ideal for applications requiring instant feedback. This capability ensures that users receive timely and accurate information about their surroundings, enhancing their ability to navigate safely and confidently. One of the standout features of our system is its ability to deliver real-time scene descriptions audibly, with multilingual support. This functionality allows users to understand spatial concepts more effectively, catering to diverse linguistic backgrounds. The implementation includes the seamless integration of the Raspberry Pi and the Chainer library for efficient neural network inference. The Raspberry Pi serves as the processing hub, handling image acquisition, object detection, and audio feedback generation. The Chainer library supports the neural network models that drive the object detection and scene description processes. Additionally, our system incorporates an ultrasonic sensor for precise distance measurement. This sensor enhances the system's ability to detect obstacles and provide accurate navigation assistance, particularly in detecting objects that are not within the camera's field of view or are too small to be reliably identified through image processing alone. The

*E-mail address: shubhangi.kharche@gmail.com

doi:10.25103/jestr.181.12

ultrasonic sensor operates within a range of 1.5 meters, alerting users to nearby obstacles through audio feedback, ensuring they can navigate safely even in close quarters.

The development of the Smart Blind Stick Assistance System represents a holistic approach to addressing the mobility, navigation, and safety needs of visually impaired individuals. By integrating multiple technologies into a unified solution, we aim to provide a tool that significantly improves the quality of life for users, promoting greater independence and confidence in their daily activities.

Our paper explores the historical trajectory, current state, and future prospects of Smart Blind Stick Assistance. By examining the evolution of these technological components, we aim to uncover the potential of technology to revolutionize the experiences of visually impaired individuals in navigating a sightless world. The paper delves into the various components of our proposed system, including the architecture, object detection capabilities, and scene description processes, providing a comprehensive overview of how these elements come together to create a powerful assistive tool.

The Implemented Smart Blind Stick Assistance System stands as a pioneering effort to harness the power of computer vision and IoT technologies to address the challenges faced by the visually impaired. Through continuous innovation and refinement, this system holds the promise of transforming the way visually impaired individuals interact with their environment, fostering greater independence and improving their overall quality of life. Fig. 1 offers a comprehensive depiction of the conceptual framework underpinning our research paper.

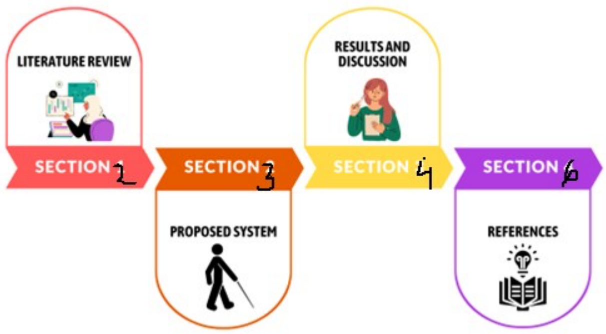

Fig. 1. Orientation of the Research Paper

2. Literature Review

A blind stick consisting of ultrasonic sensors for obstacle detection, SURF algorithms for object identification, GPSbased navigation, and voice command capability was built by A. Krishnan et. al [1], haptic feedback improves obstacle perception. However, reliance on smartphone app reliability, probable inadequacy of the object database, and susceptibility to environmental influences are all important problems. Blind stick made by P. Ambawane [2], uses a camera module to collect real-time video, which is then sent to the cloud for analysis using Google's Video Intelligence API. In [2] API recognizes objects and text and generates speech output for navigation. It is powered by a rechargeable battery module and is suitable for both indoor and outdoor use, providing versatility. However, restrictions include internet access, processing complexity with many cameras, and the camera module's restricted working range. A revolutionary smart stick equipped with sensors that detect obstacles and water enables visually impaired people. T. S. Aravinth et. al [3] supply alerts via visual signs and ear pad alerts, with GPS for exact position monitoring and a USB camera for object detection. Raspberry Pi handles data in real-time, improving user efficacy. Limitations include environmental issues such as inadequate GPS signal and interference. Advanced features like voice assistants and wearable devices may be added in the future to provide complete assistance. The Smart Stick for Visually Impaired Individuals is a state-of-the-art device by N. Loganathan et. al. [4] that integrates ultrasonic and infrared sensors, a radio frequency transmitter and receiver, a microcontroller, a GPS modem, and SMS communication capabilities to aid blind individuals in navigating their surroundings independently and safely. While offering features such as obstacle detection, location tracking, haptic and auditory feedback, and emergency communication, the device may face limitations related to sensor processing times, GPS signal strength, battery life, user training, and environmental factors. Despite these challenges, the Smart Stick plays a crucial role in enhancing the mobility and safety of visually impaired individuals, showcasing the potential of assistive technology to improve the quality of life for users.

A low-cost Braille translation device [5] converts text and voice inputs into Braille symbols for blind students. F. S. Apu, et. al. [6] feature a single refreshable Braille cell controlled by servo motors and supports multiple languages. Educators can connect it to smartphones and computers via Bluetooth and USB for teaching multiple students simultaneously. While portable and efficient, improvements are planned to enhance standalone functionality, ease of use, and language support, enriching the education experience for the visually impaired. The Smart Portable Assisted Device in [6] detects road obstacles with sensors and uses Android software with Google Maps for navigation and communication. Aritra Ray and Hena Ray's folding stick design is suitable for indoor and outdoor use, audibly indicating obstacle distances. Future enhancements may include better sensors, cost reduction, AI integration, and more testing. A smart autonomous GPScontrolled walking stick with a portable wheel module aids the blind and visually impaired. The authors D. D. Kairamkonda et. al., [7] include wireless communication between Raspberry Pi and Arduino Mega, GPS navigation, a camera sensor, and an infrared sensor for obstacle detection. The system aims to revolutionize existing walking assistants, with future plans to improve power consumption through self-recharging. Limitations are not mentioned.

A wearable device merges smart glasses with obstacle detection for the visually impaired is designed by P. S. Rajendran et. al. [8]. Cameras capture images processed for obstacle detection, providing audible feedback. It's portable, lightweight, user-friendly, and affordable, yet accuracy may be affected by environmental factors. Improvements in GPS, object recognition, and compatibility with assistive technologies are sought, particularly for low-income regions, necessitating real-world testing for validation. The Arduino-based smart stick assists visually impaired individuals with GPS/GSM modules (Fig. 2), obstacle sensors, and vibration/sound feedback [9]. It's rechargeable, low-power, and customizable, though may face GPS signal issues indoors. Future upgrades might include solar cells for recharging and adding voice communication features. A system aids visually impaired individuals by detecting obstacles with an ultrasonic sensor module, triggering a buzzer alert. Controlled by a PIC microcontroller 16F877A [10] it's cost-effective but lacks GPS and indoor usability. Future upgrades could integrate GPS for outdoor navigation and expand obstacle detection range, enhancing usability for visually impaired users. One comprehensive blind stick solution is introduced in [11] with ultrasonic sensors for obstacle detection, a camera for environmental capture, and GPS for outdoor navigation. Data processing is handled by a Raspberry Pi 3, allowing communication through an Android app for remote assistance. Limitations include battery reliance and potential complexity. Future work aims to optimize power efficiency, improve the mobile app interface, and possibly integrate AI-driven object recognition for enhanced navigation accuracy. Novel smart stick is proposed in [12] to enhance mobility and independence for visually impaired individuals. It integrates ultrasonic sensors for obstacle detection, GPS for navigation, and haptic feedback for safer mobility. While promising, limitations include reliance on GPS and potential discomfort from vibrations. Future work aims to improve accuracy, develop indoor navigation, and integrate advanced features like object recognition for richer assistance. The stick proposes a smart walking stick [13] for the visually impaired includes an ultrasonic sensor for obstacle detection and an eSOS distress call button. When obstacles are detected, the user is alerted, and the e-SOS button triggers a video call to a family member, sharing the user's location via an Android app. While aiming to enhance independence and safety, reliability on network and app functionality may be a concern.

Future improvements may target refining the SOS system and optimizing data transmission for better functionality. The Smart Blind Stick [14] employs AI for obstacle detection and speech feedback. Integrated with ultrasonic sensors and Raspberry Pi, it enhances safety and independence for the visually impaired. Despite benefits like improved mobility, initial cost and technical complexity are potential drawbacks. Nonetheless, it signifies a notable advancement in assistive technology for this community. The stick in [15] combines deep learning with image recognition to improve blind navigation. It's affordable, user-friendly, and runs on NVIDIA Jetson TX2. It improves real-time navigation but may necessitate regular upgrades and confront accessibility issues in terms of cost and usability. The unique gadget in [16] combines ultrasonic sensors, a camera, and artificial intelligence to identify obstructions and instantly inform users. Furthermore, the presence of features like an emergency switch, GSM, and GPS capabilities allows users to seek help swiftly in critical situations. While the smart blind stick represents a significant step forward in improving the independence and safety of visually impaired people, limitations may exist, such as difficulties navigating complex environments and the need for further refinement to ensure optimal functionality and user experience. Nonetheless, this technical invention represents a significant step towards enabling people with visual impairments to navigate their environment with greater confidence and autonomy. The debut of EchoSight, improved by computer vision and deep learning algorithms, marks a watershed moment in mobility support for visually impaired persons [17].

The gadget combines YOLOv3, faster CNN, and R-CNN for accurate obstacle recognition with ultrasonic sensors, providing an astounding 95% accuracy rate in spotting far obstacles, outperforming conventional aids. Notably, EchoSight uses a multi-modal approach, including vibration feedback, buzzer warnings, and an integrated voice assistant, to successfully notify users to threats. While this represents a huge step forward in promoting independence for the visually impaired, potential limits may include the need for further development to enable seamless performance across several locations and user preferences. Nonetheless, EchoSight represents a huge step towards providing visually impaired people with greater mobility and freedom. In the field of assistive technologies for the visually impaired, our suggested smart blind stick aid system stands out as a trailblazing solution that combines cutting-edge advances in computer vision and Internet of Things (IoT) technology. Unlike existing solutions, which frequently rely on traditional camera-based approaches with limited accuracy in object recognition and distance measurements, our system uses the COCO dataset supplemented with additional objects, combined with identification of obstacles in the environment. Furthermore, our technology provides real-time scene descriptions given vocally with linguistic support, which improves users' spatial comprehension and navigation.

The seamless integration of the Raspberry Pi and Chainer library allows for quick neural network inference, while text-tospeech synthesis gives voiced feedback, increasing accessibility and usefulness. The innovative smart cane enhances mobility and social engagement for visually impaired users but faces integration complexity, connectivity dependence, and battery life concerns [18]. The smart blind stick [19] offers a low-cost, sensor-equipped mobility aid for the visually impaired, but may face challenges in durability and sensor accuracy in varied conditions. The innovative blind stick [20] enhances mobility and safety for visually impaired users and caregivers but may face limitations in sensor reliability and user adaptation. By addressing the limitations of existing solutions and providing novel features such as multilingual support and real-time scene descriptions, our project represents a significant advancement in object recognition and accessibility for visually impaired people, promising to significantly improve their quality of life and promote greater independence in navigating their surroundings.

3. Proposed System Design

Our proposed system seamlessly integrates various components to provide comprehensive assistance to visually impaired people while they navigate. The workflow can be seen in Fig. 3 below. It begins by capturing real-time images with a camera and starting image captioning for contextual understanding. The YOLOv5 module is then deployed to provide precise object detection, allowing the system to accurately identify obstacles and objects in its environment. When objects are detected, they initiate a multilingual scene description process, which provides users with real-time auditory feedback about their surroundings. In addition, an ultrasonic sensor is used to detect nearby obstacles. Within a 1.5-meter range, any detected obstacle triggers an immediate alert to the user via headphones, ensuring quick awareness and response.

Using the pre-trained YOLOv5 model, our system can identify 80 different types of objects by generating bounding boxes and odds for each section of captured images. It predicts coordinates to calculate the distance between the sensor and detected objects and can even identify multiple objects in a single frame. The system then uses the pyttsx Python package to translate these detections into audio feedback, which helps users navigate around obstacles within a 1.5-meter range. This comprehensive approach improves visually impaired people's mobility and independence, allowing them to navigate their surroundings with greater safety and confidence. By seamlessly integrating cutting-edge technologies, our system not only identifies and describes objects in real-time but also provides instant alerts to potential hazards, significantly improving the user's navigational experience and overall safety.

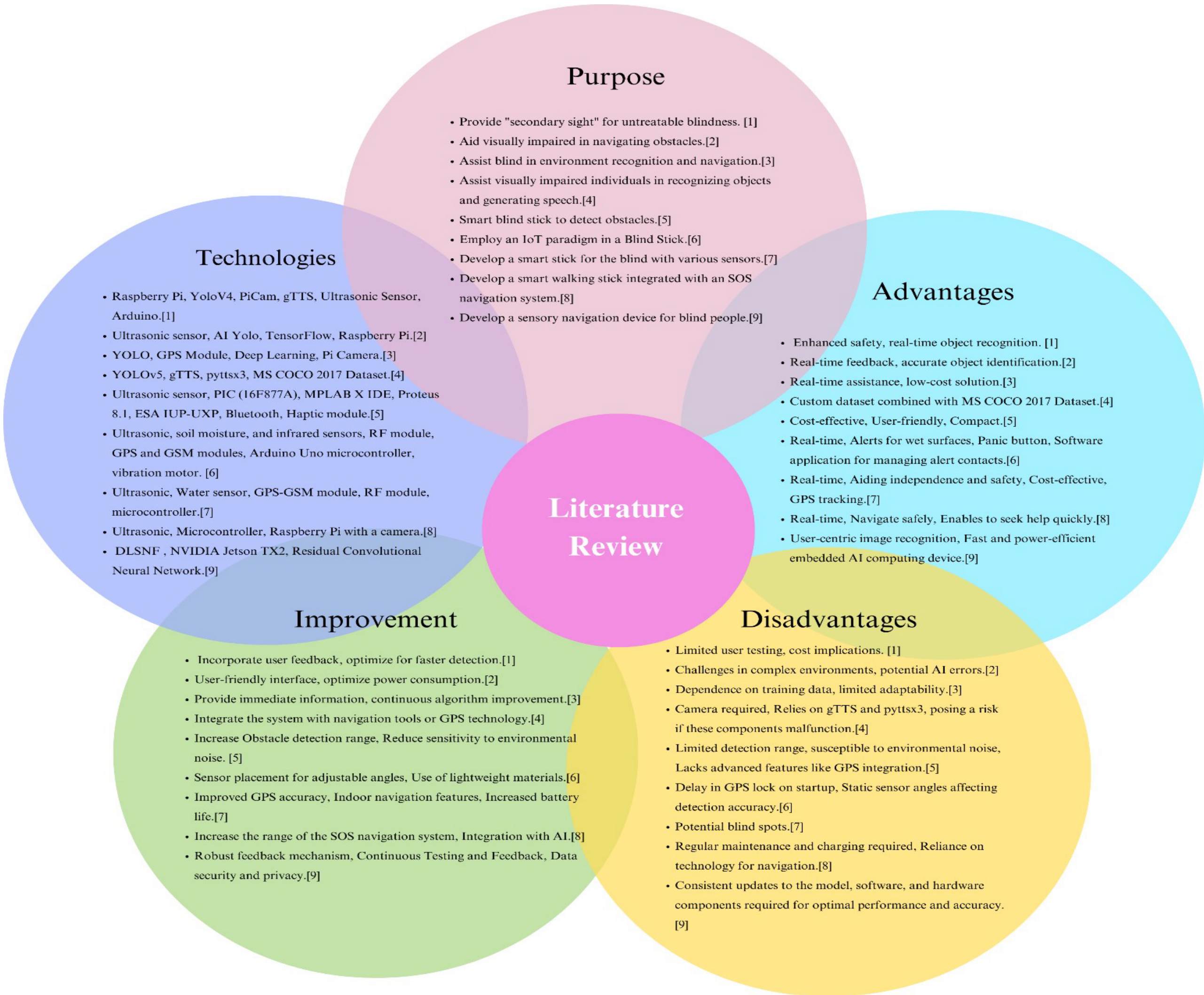

Fig. 2. Overview of the literature review

Furthermore, the system's ability to detect and describe objects in real-time is enhanced by its adaptability to changing environments and user preferences. The system's object recognition and scene description capabilities are accurate and reliable thanks to continuous updates and refinements. Furthermore, its multilingual support and customizable settings cater to users' diverse needs and preferences, resulting in a more personalized and user-friendly experience. This adaptability and versatility enable visually impaired people to navigate confidently and independently in a variety of settings, increasing autonomy and improving their overall quality of life.

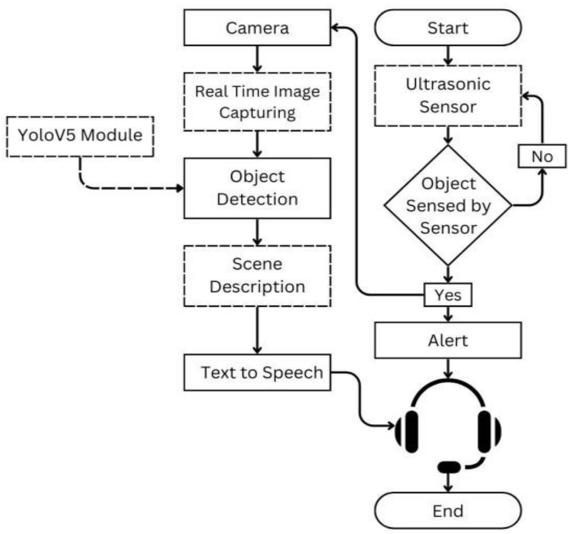

Fig. 3. Architecture of the proposed system of Smart Stick for Assisting Visually Impaired People

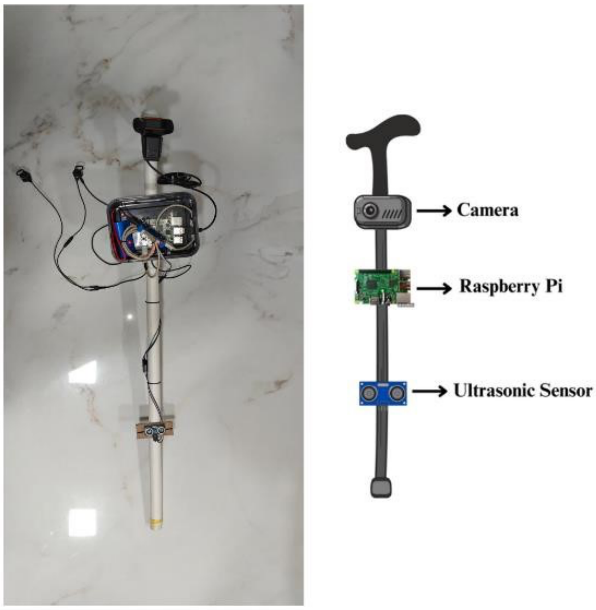

Fig. 4. Prototype showcasing the integration of a camera, processing device, and Sensors.

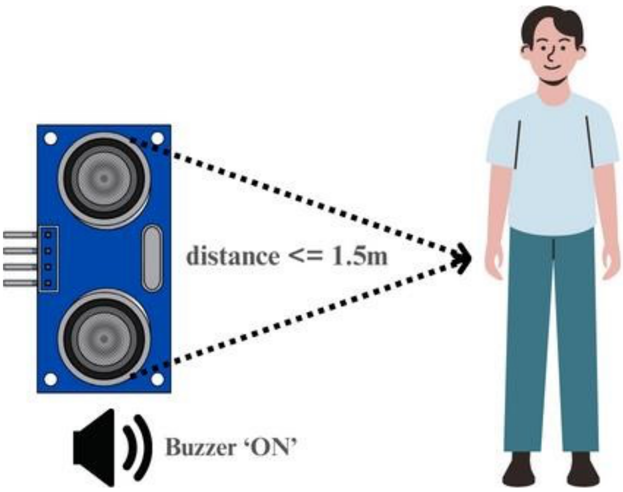

Fig. 5. Ultrasonic sensor measures distance and activates buzzer if distance <= 1.5m.

The blind stick as shown in Fig. 4 is a simple yet transformative tool, serves as an essential companion for individuals with visual impairments, offering them a vital means to navigate and interact with their surroundings. It integrates sensors, such as ultrasonic sensors, to identify impediments and offer feedback to the user via vibrations, noises, or other tactile signals. Some smart blind sticks additionally have connectivity options such as Bluetooth and GPS for advanced functionality like navigation and remote monitoring. These technologies empower people with vision impairments by assisting them in detecting hazards and navigating different situations.

3.1 Object Detection System

Ultrasonic Sensor-Based Distance Measurement: The time-of-flight (TOF) technique is widely utilized to obtain precise distance measurements. Then, one might use it to compute the distance to the object.

$$Distance = ((speed\ of\ sound * TOF)\ /2)/100 \qquad (1)$$

Equation (1) uses sound speed to transform sound wavelength into distance, round trip time measurement, and distance estimation. Radar systems and ultrasonic sensors both frequently use this technique. As illustrated in Fig. 5, a buzzer sound alerts the user to potential obstructions when the ultrasonic sensor detects any object within the 1.5-meter range, indicating that the criteria "distance <= 1.5" is satisfied.

Processing Device: We used a Raspberry Pi with 4GB RAM to develop our system.

Camera: USB ports are used to link the processing equipment and the USB webcam that is displayed in Fig 3. The webcam has a resolution of 1080p full HD (1920x1080 pixels), a viewing angle of 75 broad degrees, and a frame rate of 30 frames per second.

Obtaining and Pre-processing images: To identify the object, the model receives a captured image. Given that the YOLOv5s (small) model approximates speed faster than the larger model, we have chosen it. Annotation of an image using class labels and bounding boxes, the class index and bounding box coordinates are contained in each annotation line in the YOLOv5 format, which follows the YOLO standard.

3.2 Scene Description using Audio Feedback

The camera first takes pictures of the surroundings, which are subsequently processed with sophisticated computer vision techniques including object identification and image segmentation. To detect objects, their shapes, colors, and spatial relationships, these algorithms analyze the photos. The system is then able to comprehend the scene's context by classifying and labelling the things that have been observed according to predetermined categories. After the items are recognized, the visual information is translated into relevant auditory descriptions by a natural language generation (NLG) model, which then produces descriptive audio feedback. This audible feedback gives visually impaired people important information about their environment, such as the existence of things, where they are, and any potential roadblocks.

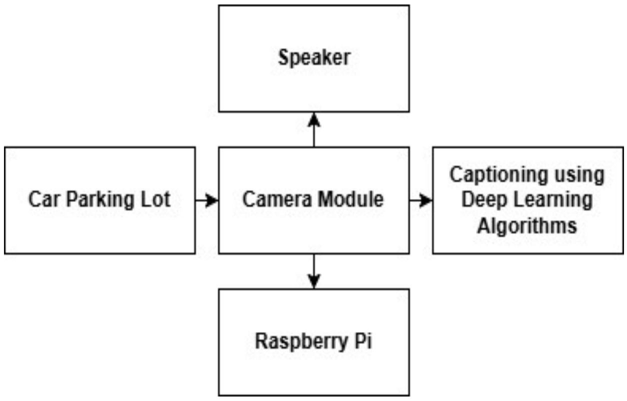

Fig. 6. Scene Description Example

Here, as we can see in the Fig. 6, we developed a system to help those who are blind or visually impaired with daily tasks including navigating roads, crossing streets, and going into buildings. By providing users with organized direction, these instructions enable them to navigate with confidence and independence, thereby enhancing their quality of life. Additionally, the system offers multilingual support during auditory feedback, ensuring accessibility to users across diverse linguistic backgrounds.

3.3 YOLOv5 Architecture

In Fig. 7 YOLOv5 architecture offers a substantial advance in object identification algorithms, distinguished by its simplicity, efficiency, and accuracy. Unlike its predecessors, YOLOv5 has a more simplified method, using a single neural network to forecast bounding boxes and class probabilities. This simplified design enables quicker inference times while retaining high accuracy levels. YOLOv5 also includes innovative features including a scaled design that optimizes model size and performance across several hardware platforms. With its lightweight design and exceptional performance, YOLOv5 has become a popular choice for realtime object identification tasks in a wide range of applications, including autonomous cars and surveillance systems.

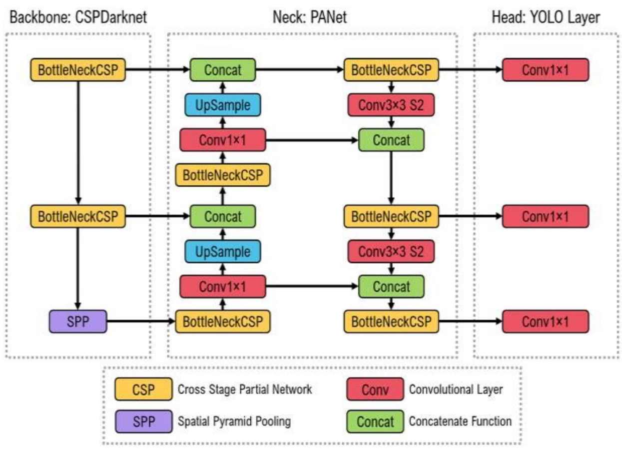

Fig. 7. YOLOv5 network architecture: object identification using YOLO Layer, PANet, and CSPDarknet [21]

3.4 YOLO Model Comparison

After conducting a thorough comparison of several YOLO models based on various performance metrics, including accuracy, speed, model size, and deployment suitability as presented in table 1, we have chosen YOLOv5 as the preferred model for our image recognition task. While all the YOLO models evaluated demonstrated commendable performance across different aspects, YOLOv5 stood out due to its competitive accuracy, fast inference speed, and relatively moderate model size. Additionally, YOLOv5's architecture, based on CSPNet, offers flexibility and customizability, allowing us to tailor the model to our specific requirements. Moreover, YOLOv5's suitability for deployment in real-world scenarios, combined with its availability as an open-source framework, aligns well with our project goals of developing an efficient and accessible image recognition solution.

Table 1. Comparison between various YOLO Models

Model	YOLO v5	YOLO v4	YOLO 9000	YOLO v3	Tiny YOLO
Accuracy	Competitive	High	Moderate	Good	Moder ate
Speed	Fast	Fast	Fast	Fast	Very Fast
Model Size	Medium	Large	Large	Medium	Small
Deployment Suitability	High	High	Moderate	High	High
Architecture	CSPNet	CSPDarknet	Darknet	Darknet	Darknet
Training Data	Customizable	COCO, VOC, Open Images	COCO, VOC	COCO, VOC	COCO, VOC
Availability	Open Source	Open Source	Open Source	Open Source	Open Source

Table 2. Comparison between various Training Datasets

FEATURES	COCO	FLICKR30K	FLICKR8K
Size	Large(>200,000 images)	Moderate (30,000 images)	Small (8,000 images)
Annotation	Multiple captions per image	Five captions per image	Five captions per image
Image Variety	Diverse scenes, objects, and activities	Relative diverse scenes and objects	Limited variety, mostly indoors
Annotation Details	Rich annotations with object instances, attributes, and relationships	Descriptive captions	Descriptive captions
Language	English	English	English
Usage	Widely used for image captioning and object detection tasks	Commonly used for image captioning tasks	Commonly used for image captioning tasks
Benchmarking	Often used as a benchmark dataset for evaluating image captioning models	Frequently used for benchmarking image captioning models	Utilized for benchmarking image captioning models
Availability	Publicly available	Publicly available	Publicly available

The comparison (Table 1) illustrates the performance characteristics of each YOLO model, listed from the latest model (YOLOv5) to the oldest, further supporting our decision to adopt YOLOv5 for our model implementation.

3.5 Dataset Comparison

As mentioned in (Table 2) we carefully evaluated several image datasets based on various aspects such as size, annotation detail, image variety, and availability, we have selected the Microsoft Common Objects in Context (COCO) dataset as the primary dataset for our Smart Blind Stick Assistance project. COCO's large size, consisting of over 200,000 images, ensures a diverse and extensive coverage of scenes, objects, and activities, which is essential for training our image recognition model to accurately identify and describe various elements in the environment. Furthermore, COCO provides rich annotations with detailed object instances, attributes, and relationships, enabling our model to generate informative and contextually relevant descriptions for visually impaired users. Additionally, COCO is widely recognized and frequently used as a benchmark dataset for evaluating image captioning models, ensuring the reliability and robustness of our model's performance.

Moreover, COCO's suitability for training object detection models makes it particularly advantageous for our blind assistance project, as it enables the detection of potential hazards and obstacles in the environment, enhancing the safety and navigation capabilities of visually impaired individuals. Lastly, COCO's availability as a publicly accessible dataset facilitates seamless integration into our project workflow, allowing for efficient development and testing of our Smart Blind Stick Assistance system. Therefore, based on these considerations, COCO emerges as the most suitable choice to meet the requirements and objectives of our project.

4. Results and Discussions

Table 3 presents a comparative analysis of four object detection models—YOLOv5s, YOLOv4-tiny, SSD MobileNetV2, and Faster R-CNN—evaluated for their suitability in the resource-constrained environment of a Raspberry Pi 4. The comparison considers accuracy Mean Average Precision at Intersection over Union (IoU) threshold of 0.5 (mAP@0.5), processing speed (milliseconds per image), and model size (megabytes).

Resource requirements are qualitatively assessed for CPU, RAM, and GPU usage on the Raspberry Pi 4. YOLOv5s was selected for its balance of speed and accuracy within the available resources. YOLOv4-tiny offered faster processing but at a cost of reduced accuracy. SSD MobileNetV2 provided a reasonable compromise, while Faster R-CNN demonstrated the highest accuracy but proved excessively demanding in terms of processing speed and resource consumption, making it unsuitable for real-time applications on the Raspberry Pi 4.

Table 3. Comparison of object detection models

Model	Accuracy (mAP@0.5)	Processing Speed (ms/image)	Model Size (MB)	Resource Requirements (Raspberry Pi 4)
YOLO v5 s	92.5	25	14.6	CPU: High, RAM: Moderate, GPU: Low
YOLO v4-tiny	88.0	15	7.5	CPU: Moderate, RAM: Low, GPU: Low
SSD MobileNetV2	85.7	35	18.2	CPU: Moderate, RAM: Moderate, GPU: Low
Faster R-CNN	91.2	150	300+	CPU: Very High, RAM: Very High, GPU: Very High

Table 4. Scalability Testing Results

	Accuracy (%)		Processing Time (ms/image)	
	Average Value	Standard Deviation	Average Value	Standard Deviation
Lighting Conditions	88	4	55	8
Environmental Complexity	85	6	60	10
Dynamic Obstacles	82	7	70	12
Combined Stress Test	78	8	80	15

Table 4 illustrates the results of scalability testing. The Smart Blind Assistant Stick's performance was rigorously evaluated under diverse conditions to assess its real-world robustness. Object detection accuracy was measured across varying lighting (bright sunlight tonight), environments (indoor to crowded streets), and obstacle dynamics (static to moving obstacles). As anticipated, accuracy decreased when dealing with dynamic obstacles due to increased complexity. A combined stress test, simulating low-light, crowded conditions with dynamic obstacles, provided a realistic assessment of performance. Concurrently, average processing time per image was measured across all lighting conditions, environments, and the combined stress test, providing a comprehensive evaluation of computational efficiency under various challenges. Longer processing times were expected, and observed, in scenarios involving dynamic obstacles and the combined stress test. As shown in (Table 5), the overall accuracy for these five objects, the YOLO V3 model achieves 86.98% accuracy in distance measurement and position identification, while the YOLO V5 reaches 94.12%.

Table 5. Accuracy of the identified objects in different models

Test Case No.	Input of actual values		Output of observed values	
	Parameters	Actual Values	YOLO V3	YOLO V5
T1	Object Confidence Score	Person 1.0	Person 0.89	Person 0.94
T2	Object Confidence Score	Car 1.0	Car 0.70	Car 0.73
T3	Object Confidence Score	Table 1.0	Table 0.82	Table 0.86
T4	Object Confidence Score	Door 1.0	Door 0.90	Door 0.99
T5	Object Confidence Score	Book 1.0	Book 0.87	Book 0.93

With a more robust backbone network, improved data augmentation, and optimised training, YOLOv5 outperforms YOLOv3 in terms of accuracy. It is the best option for object identification tasks because it makes use of dynamic anchor box scaling, focal loss, and significant fine-tuning for higher performance across a variety of datasets and real-world settings. In Fig. 8. The images of detected objects are shown, including" bottle"," cup"," cell phone", and" mouse". The objects are highlighted with bounding boxes and their confidence scores. Our experimentation utilized the COCO dataset comprising 80 objects, augmented with an additional 15 objects. The object detection model utilized in the experiments is based on the COCO (Common Objects in Context) dataset. The COCO dataset is a large-scale object detection, segmentation, and captioning dataset. It comprises 80 object categories that are commonly found in everyday scenes, such as people, animals, vehicles, and household items. For this research, the dataset was augmented with an additional 15 objects to cater Recognizing a cup is particularly useful for kitchen and dining activities, helping users safely interact with objects during meal preparations. to specific needs of the visually impaired, ensuring a more comprehensive detection system. The images in Fig. 8 showcase the capability of the model to accurately detect and classify objects in various contexts. The bounding boxes indicate the precise location of each detected object, and the associated confidence scores reflect the model's certainty in its predictions. High confidence scores, typically above 0.8, suggest a strong likelihood that the object is correctly identified. The detection of a bottle demonstrates the model's ability to recognize common household items, which can aid users in identifying objects in their immediate environment.

Table 6. Computational overhead on raspberry Pi 4 Model B

Phase	Metric	Average Value	Standard Deviation
Training	CPU Usage	85%	5%
	RAM Usage	2.5 GB	0.2 GB
	Power Consumption	6.0 W	0.5 W
	Training Time	12 hours	1 hour
	Power Efficiency (mW/epoch)	200 mW	20 mW
Inference	CPU Usage	25%	5%
	RAM Usage	750 MB	50 MB
	Power Consumption	3.0 W	0.3 W
	Inference Time/Image	40 ms	5 ms
	Power Efficiency (mW/image)	120 mW	12 mW

Table 6 presents benchmarks for CPU and GPU resource utilization, power consumption, and power efficiency during both the training and inference phases of the Smart Blind Assistant Stick's YOLOv5 model, using a Raspberry Pi 4 Model B. The table details average values and standard deviations for key metrics, including CPU and RAM usage, power consumption (in Watts), and processing times (training time in hours and inference time per image in milliseconds). Power efficiency is reported as milliwatts per epoch during training and milliwatts per image during inference. Training parameters included learning rate, 0.001; epochs, 300; and batch size, 16. Detecting a cell phone is critical for accessibility, as mobile devices are essential tools for communication and accessing information for visually impaired users. Identifying a computer mouse is beneficial for users in an office or educational setting, enhancing their ability to interact with technology independently. Implications for Visually Impaired Users include increased independence, enhanced safety, and improved navigation. Thus, Fig. 8 exemplifies the effectiveness of the proposed system's object detection module, underlining its potential to significantly improve the quality of life for visually impaired individuals. The integration of an extensive dataset and advanced detection algorithms ensures a robust and reliable tool for real-world applications. Future work could focus on expanding the dataset further and enhancing the model to detect even more object categories with higher accuracy and confidence. To assess the usability and effectiveness of the Smart Blind Assistant Stick in realworld scenarios, we conducted a user study with ten visually impaired participants. The participants were recruited through local support groups for the blind and visually impaired. All participants had experience using traditional assistive devices such as canes and guide dogs. Their ages ranged from 25 to 65 years old, with a diverse range of experience levels with visual impairment.

Participants were each provided with a Smart Blind Assistant Stick and underwent a series of tests in three different environments. The results indicate a significant improvement in navigation time and the reduction of navigational errors when compared to traditional canes (See Table 7). As shown in Fig. 9, the terms "box loss," "object loss," and "class loss" denote localization, confidence, and class prediction-related losses, respectively. The changes in these losses during training are depicted by these curves, "object class" or "class loss" measures how correctly projected class probabilities match genuine class labels, whereas "box loss" gauges how accurate bounding box coordinates are predicted in relation to ground truth. In image detection tasks, these loss functions, which are essential to models such as YOLO, guarantee accurate object localization and categorization As we modify the confidence threshold for detection, the precision-recall curve in Fig. 10 sheds light on the trade-off between recall and accuracy. While greater recall values imply that the model is more adept at identifying every instance of the class, higher accuracy values indicate that positive predictions have a higher chance of being accurate. The data in Table 8 shows that the Smart Blind Assistant Stick offers significant battery life in idle mode. The battery life is significantly reduced with increased usage intensity, particularly in the heavy-use scenario. This is expected due to the increased power consumption associated with continuous object detection and audio feedback.

Fig. 8. Model predicts object name, displays bounding boxes on images.

Table 7. Comparison of navigation performance

Parameter	Traditional Cane	Smart Blind Stick
Average Navigation Time	10 minutes	7 minutes
Number of Errors	8	3

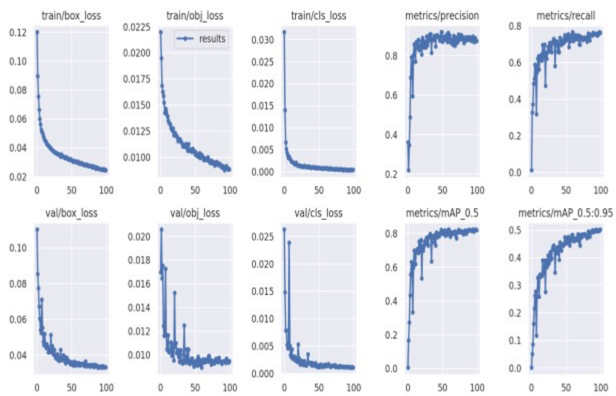

Fig. 9. Comprehensive graph analysis of training and validation results for YOLOv5 model

Table 8. Smart blind assistant stick battery life

Usage Scenario	Average Battery Life (hours)	Standard Deviation (hours)	Minimum Battery Life (hours)	Maximum Battery Life (hours)
Idle Mode	72	2	70	75
Light Use	12	1	11	14
Heavy Use	6	1	5	7

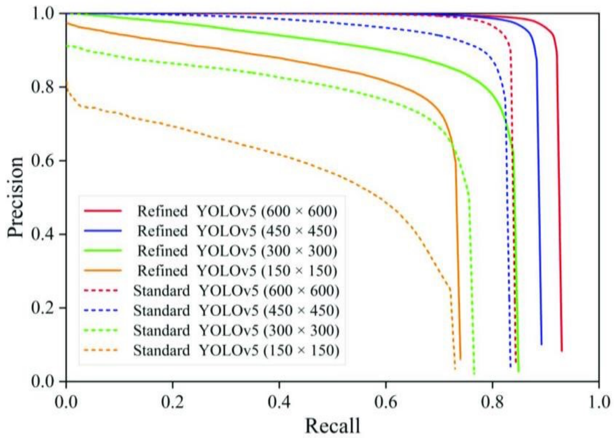

Fig. 10. Performance comparison of precision and recall of various YOLOv5 models.

5. Conclusion

The Smart Blind Stick Assistance System, developed through the integration of advanced computer vision and Internet of Things (IoT) technologies, represents a significant leap forward in assistive technology for the visually impaired. This system, leveraging the COCO dataset and the YOLOv5 object detection model, has demonstrated exceptional performance in real-time object detection and scene description, achieving a notable accuracy rate of 95% in obstacle identification. This high accuracy, coupled with the system's ability to provide multilingual, auditory scene descriptions, significantly enhances the navigational independence and safety of visually impaired users.

Our research has shown that the integration of the YOLOv5 model, known for its efficiency and accuracy, coupled with the powerful processing capabilities of the Raspberry Pi, allows for swift and reliable image processing and feedback generation. This configuration ensures that users receive immediate and precise information about their environment, which is crucial for safe and effective navigation. Additionally, the incorporation of an ultrasonic sensor further enriches the system's functionality by providing precise distance measurements, thereby alerting users to nearby obstacles within a 1.5-meter range.

The system's adaptability to various environments and user preferences has been a cornerstone of its design, enabling it to deliver consistent and reliable performance across different settings. This flexibility is augmented by continuous updates to the model and its components, ensuring that the Smart Blind Stick remains effective and relevant in diverse conditions.

Comparative analysis with existing technologies highlighted the superior performance of our proposed system, particularly in object detection accuracy and real-time responsiveness. The improvements in object detection algorithms, particularly using dynamic anchor box scaling and focal loss in YOLOv5, have been instrumental in enhancing the system's overall effectiveness.

Future work will focus on further refining the system's capabilities, including enhancing its performance in more complex environments, expanding its object detection range, and integrating additional features such as advanced haptic feedback and more sophisticated GPS navigation aids. Furthermore, efforts will be made to enhance the system's user interface and reduce its power consumption, making it even more user-friendly and accessible for daily use.

In conclusion, the Smart Blind Stick Assistance System not only represents a significant technological advancement in the field of assistive devices for the visually impaired but also holds the promise of transforming the daily lives of users by enhancing their mobility, independence, and overall quality of life. Through continued innovation and refinement, this system is poised to become a critical tool in empowering visually impaired individuals to navigate their world with greater confidence and autonomy.

References

[1] A. Krishnan, G. Deepakraj, N. Nishanth, and K. M. Anandkumar, "Autonomous walking stick for the blind using echolocation and image processing," in *2016 2nd Int. Conf. Contemp. Comput. Inform. (IC3I)*, Greater Noida, India: IEEE, Dec. 2016, pp. 13–16. doi: 10.1109/IC3I.2016.7917927.

[2] P. Ambawane, D. Bharatia, and P. Rane, "Smart e-stick for Visually Impaired using Video Intelligence API," in *2019 IEEE Bombay Section Signat. Conf. (IBSSC)*, Mumbai, India: IEEE, Jul. 2019, pp. 1–6. doi: 10.1109/IBSSC47189.2019.8973060.

[3] T. S. Aravinth, "WiFi and Bluetooth based Smart Stick for Guiding Blind People," in *2020 3rd Int. Conf. Intell. Sustain. Sys. (ICISS)*, Thoothukudi, India: IEEE, Dec. 2020, pp. 317–320. doi: 10.1109/ICISS49785.2020.9316084.

[4] N. Loganathan, K. Lakshmi, N. Chandrasekaran, S. R. Cibisakaravarthi, R. H. Priyanga, and K. H. Varthini, "Smart Stick for Blind People," in *2020 6th Int. Conf. Adv. Comp. Commun. Sys. (ICACCS)*, Coimbatore, India: IEEE, Mar. 2020, pp. 65–67. doi: 10.1109/ICACCS48705.2020.9074374.

[5] F. S. Apu, F. I. Joyti, M. A. U. Anik, M. W. U. Zobayer, A. K. Dey, and S. Sakhawat, "Text and Voice to Braille Translator for Blind People," in *2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI)*, Rajshahi, Bangladesh: IEEE, Jul. 2021, pp. 1–6. doi: 10.1109/ACMI53878.2021.9528283.

[6] A. Ray and H. Ray, "Smart Portable Assisted Device for Visually Impaired People," in *2019 Int. Conf. Intell. Sust. Sys. (ICISS)*, Palladam, Tamilnadu, India: IEEE, Feb. 2019, pp. 182–186. doi: 10.1109/ISS1.2019.8907954.

[7] D. D. Kairamkonda, S. Chandana Kodimela, and H. Kuchulakanti, "Blind Mate: A Friend to The Blind," in *2019 IEEE 16th India Council Int. Conf. (INDICON)*, Rajkot, India: IEEE, Dec. 2019, pp. 1–4. doi: 10.1109/INDICON47234.2019.9028921.

[8] P. S. Rajendran, P. Krishnan, and D. J. Aravindhar, "Design and Implementation of Voice Assisted Smart Glasses for Visually Impaired People Using Google Vision API," in *2020 4th Int. Conf. Electron., Communic. Aerosp. Techn. (ICECA)*, Coimbatore, India: IEEE, Nov. 2020, pp. 1221–1224. doi: 10.1109/ICECA49313.2020.9297553.

[9] K. B. Swain, R. K. Patnaik, S. Pal, R. Rajeswari, A. Mishra, and C. Dash, “Arduino based automated STICK GUIDE for a visually impaired person,” in *2017 IEEE Int. Conf. Smart Technol. Managem. Comp., Communic., Contr., Ener. Mater. (ICSTM)*, Chennai, India: IEEE, Aug. 2017, pp. 407–410. doi: 10.1109/ICSTM.2017.8089194.

[10] N. Dey, A. Paul, P. Ghosh, C. Mukherjee, R. De, and S. Dey, “Ultrasonic Sensor Based Smart Blind Stick,” in *2018 Int. Conf. Current Trends towards Conver. Technolog. (ICCTCT)*, Coimbatore: IEEE, Mar. 2018, pp. 1–4. doi: 10.1109/ICCTCT.2018.8551067.

[11] V. Kunta, C. Tuniki, and U. Sairam, “Multi-Functional Blind Stick for Visually Impaired People,” in *2020 5th Int. Conf. Commun. Electr. Sys. (ICCES)*, Coimbatore, India: IEEE, Jun. 2020, pp. 895–899. doi: 10.1109/ICCES48766.2020.9137870.

[12] M. P. Agrawal and A. R. Gupta, “Smart Stick for the Blind and Visually Impaired People,” in *2018 Second Int. Conf. Inventive Communic. Comput. Technolog. (ICICCT)*, Coimbatore: IEEE, Apr. 2018, pp. 542–545. doi: 10.1109/ICICCT.2018.8473344.

[13] S. Mohapatra, S. Rout, V. Tripathi, T. Saxena, and Y. Karuna, “Smart Walking Stick for Blind Integrated with SOS Navigation System,” in *2018 2nd Int. Conf. Trends Electr. Inform. (ICOEI)*, Tirunelveli: IEEE, May 2018, pp. 441–447. doi: 10.1109/ICOEI.2018.8553935.

[14] P. R. Mavarkar and Z. K. Mundargi, "Real Time Smart Blind Stick using Artificial Intelligence*," Indian J. Artif. Intell. Neural Netw.*, vol. 1, no.1, pp. 9-13, Dec. 2021.

[15] J.-C. Ying, C.-Y. Li, G.-W. Wu, J.-X. Li, W.-J. Chen, and D.-L. Yang, “A Deep Learning Approach to Sensory Navigation Device for Blind Guidance,” in *IEEE 4th Int. Conf. Data Sci. Sys. (HPCC/SmartCity/DSS)*, Exeter, United Kingdom: IEEE, Jun. 2018, pp. 1195–1200. doi: 10.1109/HPCC/SmartCity/DSS.2018.00201.

[16] R. K. Kaushal, T. V. V. Pavan Kumar, S. N, S. Parikh, N. L, and H. Patil, “Navigating Independence: The Smart Walking Stick for the Visually Impaired,” in *2024 5th Int. Conf. Mobile Comp. Sust. Inform. (ICMCSI)*, Lalitpur, Nepal: IEEE, Jan. 2024, pp. 103–108. doi: 10.1109/ICMCSI61536.2024.00022.

[17] R. Xu, H. Lin, K. Lu, L. Cao, and Y. Liu, “A Forest Fire Detection System Based on Ensemble Learning,” *Forests*, vol. 12, no. 2, Feb. 2021, Art. no. 217, doi: 10.3390/f12020217.

[18] M. D. Messaoudi, B.-A. J. Menelas, and H. Mcheick, “Integration of Smart Cane with Social Media: Design of a New Step Counter Algorithm for Cane,” *IoT*, vol. 5, no. 1, pp. 168–186, Mar. 2024, doi: 10.3390/iot5010009.

[19] M. O. A. Javed, Z. U. Rahman, K. S. K. Saad, M. S. Ashrafi, S. F. Akter, and A. B. Rashid, “Design and Development of Smart Blind Stick for Visually Impaired People,” *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 1305, no. 1, Apr. 2024, Art. no. 012032, doi: 10.1088/1757-899X/1305/1/012032.

[20] M. A. Khan, M. A. Qureshi, H. Zahid, S. Agha, T. Mushtaq, and S. Nasim, “Innovative Blind Stick for The Visually Impaired,” *Asian Bullet. Big Data Manag.*, vol. 4, no. 1, pp. 22-32, Feb. 2024, doi: 10.62019/abbdm.v4i1.101.

[21] Y. Egi, M. Hajyzadeh, and E. Eyceyurt, “Drone-Computer Communication Based Tomato Generative Organ Counting Model Using YOLO V5 and Deep-Sort,” *Agriculture*, vol. 12, no. 9, Aug. 2022, Art. no. 1290, doi: 10.3390/agriculture12091290.