A Review on-Advancements in Autonomous Navigation Robots for Visually Impaired People

Purnima Hanumant Chabbi[1], Dr. Jaya Christiyan K.G[2], Dr. A.Ajina[3]

Department of Mechanical Engineering, Ramaiah Institute of Technology, Bengaluru-54

Corresponding Author: Purnima Hanumant Chabbi

Abstract:- Recent advancements in autonomous navigation technologies have markedly facilitated the development of intelligent assistive systems, aimed at augmenting the mobility and autonomy of individuals with visual impairments. This review paper critically examines contemporary innovations in autonomous navigation robots tailored for the visually impaired, with particular emphasis on the integration of computer vision, multimodal sensor fusion, machine learning, and real-time obstacle avoidance strategies. Notably, the deployment of state-of-the-art object detection algorithms, including You Only Look Once (YOLO) and Faster R-CNN, has substantially enhanced the precision and efficiency of environmental perception in dynamic settings. The study delineates a range of system architectures—spanning wearable technologies to robotic guide platforms—and evaluates their operational efficacy across diverse spatial contexts, encompassing both indoor and outdoor environments. Furthermore, the role of artificial intelligence in fostering situational awareness and autonomous decision-making is explored, with a view to optimising user interaction and navigational safety. The paper also engages with prevailing challenges, such as cost-effectiveness, adaptability to heterogeneous environments, and intuitive user interface design. In conclusion, this review provides a comprehensive synthesis of the current state-of-the-art, identifies salient research gaps, and proposes strategic directions for future inquiry in the domain of autonomous assistive navigation for visually impaired.

Keywords*:* Autonomous navigation robots; Visually impaired; Assistive technology; Object detection; YOLO; Faster R-CNN; Computer vision; Sensor fusion; Machine learning; Obstacle avoidance.

1. Introduction

Visual impairment continues to pose profound challenges to personal autonomy, mobility, and quality of life. As reported by the World Health Organization, over 2.2 billion individuals worldwide are affected by varying degrees of visual disability, necessitating the development of sophisticated assistive technologies. Conventional mobility aids such as the white cane or guide dogs, while valuable, are inherently constrained in their capacity to provide comprehensive spatial awareness, particularly in complex or unfamiliar environments. In light of these limitations, the advent of autonomous navigation robots heralds a paradigm shift in assistive mobility, leveraging advancements in robotics, sensor integration, and artificial intelligence to offer more intelligent and context-aware navigation solutions.

Recent innovations in this domain have centred on the amalgamation of computer vision, sensor fusion, and deep learning methodologies to augment situational awareness and navigational precision. Prominent among these are object detection algorithm[1] which facilitates the real-time identification and localisation of obstacles within dynamically changing environments. The integration of these technologies into wearable systems and robotic platforms has engendered significant improvements in path planning, obstacle avoidance, and user safety, both in indoor and outdoor contexts[2]. This review endeavours to provide a critical synthesis of current developments, interrogate prevailing challenges—such as cost, scalability, and user adaptability—and delineate prospective avenues for future research in the domain of autonomous assistive navigation for visually impaired individuals.

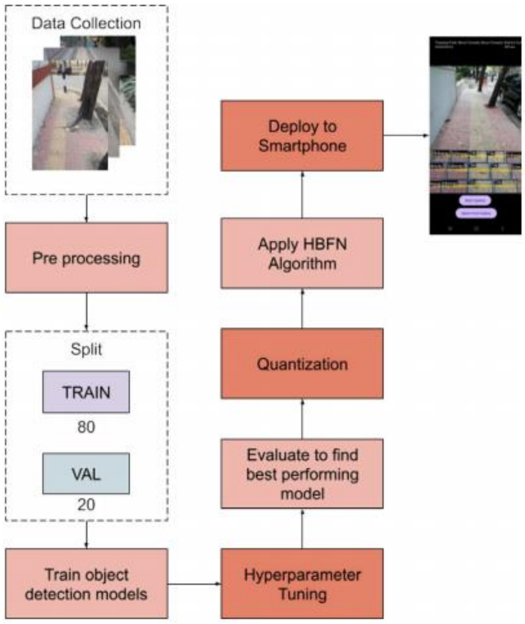

Fig. 1: Workflow for Smartphone-Based Object Detection Model Development and Deployment

Fig. 1 as illustrated in the present study, a structured series of procedures was undertaken to implement object detection mechanisms aimed at facilitating the navigation of visually impaired individuals along the footpaths of Bangladesh[3].

This review aspires to enrich the scholarly discourse on artificial intelligence–enabled assistive technologies by offering a comprehensive and critical examination of recent advancements in autonomous navigation robots tailored for individuals with visual impairments. By synthesising contemporary developments in robotics, computer vision, sensor integration, and human–machine interaction, this study seeks to illuminate the transformative capacity of such systems to enhance independent mobility, spatial cognition, and overall quality of life for the visually impaired[4]. Moreover, it aims to articulate prevailing limitations and propose informed directions for future inquiry within this rapidly evolving interdisciplinary domain.

Primarily attributable to the inherent limitations of conventional mobility aids—such as the white cane and guide dogs—the spatial autonomy of individuals with visual impairments remains profoundly constrained. The continual evolution of urban landscapes, coupled with the escalating intricacy of traffic systems and public infrastructure, exacerbates the navigational challenges encountered by this demographic[5]. In light of these complexities, contemporary innovations including autonomous navigation robots and wearable assistive technologies are gaining prominence as transformative, technology-driven interventions.

We propose the development of an intelligent navigational assistance system tailored for individuals who are blind or visually impaired (BVIP). The conceptual framework of the system is predicated upon a sophisticated decision-support mechanism underpinned by fuzzy logic[6]. Central to its architecture is the integration of a Raspberry Pi 4 microprocessor, facilitating real-time computational processing. The system further incorporates an array of high-precision sensors to ensure accurate environmental perception, alongside a multimodal haptic and voice-guided interface designed to provide intuitive and context-aware navigational feedback to the user[7].

Path planning has become a pivotal and complex area of research within autonomous navigation, especially for assistive technologies [8]. In unstructured or dynamic environments, systems must extract spatiotemporal data and adapt routes in real time. Existing solutions often rely on computationally intensive external services and numerous sensors, increasing energy consumption and reducing wearability. These constraints hinder real-time deployment in compact systems like smart glasses. To address this, we propose a novel two-level hierarchical architecture that combines global and local path planning for efficient, user-friendly navigation[9].

Together, these technologies enable adaptive, context-aware decision-making in dynamic and unstructured environments encountered by visually impaired individuals. While existing studies have explored various facets of autonomous navigation and assistive robotics, much of the prior research remains narrowly focused on isolated components such as obstacle detection or path planning. This review aspires to address the prevailing research gap by adopting a comprehensive and integrative perspective on the technological ecosystem.

The principal contributions of this study are outlined as follows:

1. The integration of advanced deep learning architectures has substantially augmented the precision of obstacle detection methodologies. These innovations significantly enrich environmental perception, thereby advancing autonomous navigational capabilities for visually impaired individuals in intricate outdoor settings

2. Introduces a cutting-edge edge intelligence paradigm that seamlessly integrates computer vision techniques to enable low-latency, decentralised processing. This advancement mitigates reliance on cloud infrastructure, thus rendering assistive systems more efficient, scalable, and suitable for real-time footpath navigation.

3. Presents an adaptive path planning algorithm capable of dynamically recalibrating navigation routes in response to environmental variability. This algorithmic innovation underpins the development of context-sensitive assistive navigation frameworks.

4. Offers a transformative perspective that reconceptualises disability through the lens of technologically mediated ability. By amalgamating insights from diverse disciplines, the proposed framework advances inclusive innovation and fosters autonomous spatial navigation as a right rather than a privilege.

5. Underscore the efficacy of real-time, sensor-integrated wearable systems in delivering autonomous navigation support. These systems exemplify the convergence of ergonomic design and intelligent control, enhancing mobility in both structured and unstructured environments.

The review of this paper is organized as follows: Section 2 provides an overview of the algorithms utilized in the proposed system. Section 3 discusses prevailing security technologies relevant to .Advancements in Autonomous Navigation Robots for Visually Impaired People Section 4 presents the system architecture in detail. Section 5 explores the practical applications of the proposed approach. Finally, Section 6 concludes the study and outlines potential future work.

2. Overview of Algorithms

2.1 YOLOv3

YOLOv3 epitomises a triadic single-stage detection paradigm undergirded by Darknet-53, an elaborate convolutional backbone that interlaces residual augmentations with alternating 3×3 and 1×1 kernels[10]. This architecture adroitly circumvents vanishing-gradient vicissitudes while engendering richly hierarchical feature abstractions, thereby striking a felicitous equilibrium between representational profundity and real-time throughput imperatives.

At its core lies a multiscale prognostication schema, deploying three discrete detection heads at progressively attenuated spatial resolutions (82×82, 41×41, 21×21). Each head employs k-means-derived anchor boxes to instigate objectness scoring via logistic regression and to regress bounding-box deltas vis-à-vis grid-cell centroids. This stratagem endows the model with perspicacious acuity across an expansive gamut of object dimensions, mitigating small-object occlusion proclivities endemic to monolithic detectors[11].

Complementing these structural refinements, YOLOv3's optimisation criterion amalgamates binary cross-entropy for objectness and classification with mean-squared error for spatial regression, forging a harmonised loss landscape[12]. Its introduction of class-agnostic objectness thresholds obviates class-imbalance perturbations, while its streamlined inference pipeline consummates a synthesis of alacrity and exactitude, cementing YOLOv3's pre-eminence in contemporary real-time detection praxis.

2.2 YOLOv5

YOLOv5 epitomizes a consummate single-stage object-detection paradigm, undergirded by a Cross-Stage Partial (CSP) backbone that mitigates representational redundancy whilst bolstering gradient propagation[13]. By bifurcating the feature map into dual conduits—one channel undergoing nonlinear transformation and the other circumventing intermediary layers—this architecture engenders computational parsimony without compromising discriminative richness. Such judicious stratification confers YOLOv5 with an exquisite equilibrium between parameter economy and inferential alacrity, rendering it eminently suitable for deployment on resource-constrained and edge-computing platforms[14].

Augmenting its foundational substrate, YOLOv5 incorporates a Path Aggregation Network (PANet) as its "neck," which orchestrates a synergistic fusion of multiscale feature hierarchies[15]. Through bidirectional lateral linkages, PANet seamlessly integrates coarse, context-laden embeddings with fine-grained spatial detail, thereby endowing the model with unrivalled locational acuity across heterogeneous object dimensions. The ensuing detection head is meticulously decoupled into discrete classification and bounding-box regression branches, a design choice that diminishes intertask interference and facilitates doctrinal specialisation, culminating in a felicitous harmony of precision and latency[16].

The training regimen of YOLOv5 is distinguished by avant-garde data-augmentation and automated optimisation schemes[17]. Mosaic augmentation artfully amalgamates quadripartite image fragments into a singular composite, diversifying scene contextualisation and attenuating overfitting propensities. Concurrently, MixUp interpolation engenders smoother decision boundaries by synthesising convex amalgams of inputs and labels. These techniques are complemented by an adaptive hyper parameter search—calibrating learning rates, momentum coefficients, and augmentation intensities—which expedites convergence and amplifies mean average precision. The resultant framework thus consummates a symbiosis of expediency and exactitude, cementing YOLOv5's preeminence within the contemporary object-detection vanguard[18][19].

2.3 YOLOv8

YOLOv8 inaugurates an anchor-free, single-stage object-detection paradigm that dispenses with heuristic anchor-box calibration, thereby obviating the need for manual anchor-tuning. Its backbone emanates from a Cross-Stage Partial (CSP) framework, augmented with efficient "C2f" modules that parsimoniously preserve gradient flow while curtailing representational redundancy[20]. A spatial pyramid pooling–fast (SPPF) layer atop the backbone aggregates multiresolution context, thereby enriching receptive fields without incurring onerous computational overhead[21].

The neck comprises a Path Aggregation Network (PANet) that orchestrates a bidirectional fusion of hierarchical embeddings, seamlessly melding coarse semantic cues with fine-grained spatial detail[22]. This integrative schema confers the detector with unrivalled locational acuity across a heterogeneous spectrum of object scales, ameliorating small-object omission and bolstering detection robustness in cluttered scenes[23].

At its terminus, the decoupled detection head bifurcates classification and regression pathways, attenuating intertask interference and fostering doctrinal specialization. YOLOv8's training regimen leverages dynamic label-assignment via a SimOTA-inspired optimiser, automated hyperparameter evolution, and advanced mosaic/copy-paste augmentations to expedite convergence and elevate mean average precision. Collectively, these innovations cement YOLOv8's preeminence in reconciling alacrity with exactitude for contemporary real-time object-detection tasks.

2.4 Faster RCNN

FasterRCNN epitomises an advanced paradigm in object detection, integrating a Region Proposal Network (RPN) with a deep convolutional architecture to generate and classify region proposals in a unified framework. This architecture significantly enhances computational efficiency and detection accuracy by obviating the need for exhaustive search strategies. Its capacity to localise and categorise multiple objects with high fidelity renders it well-suited for real-time scene interpretation in dynamic, unstructured environments[24].

In the context of assistive navigation for visually impaired individuals, Faster R-CNN facilitates nuanced environmental perception by reliably detecting obstacles, pedestrians, pathways, and critical landmarks. When embedded within wearable systems or autonomous robotic platforms, its outputs can be mapped to haptic or auditory feedback mechanisms, thereby enabling spatial cognition through sensory substitution. This synthesis of deep learning and assistive technology exemplifies the transformative potential of artificial intelligence in fostering inclusive, independent mobility.

2.5 Single Shot MultiBox Detector (SSD)

The Single Shot MultiBox Detector (SSD) represents a salient advancement in object detection, distinguished by its capacity to perform classification and localisation in a singular, end-to-end computational pass. In contrast to two-stage architectures such as Faster R-CNN, SSD obviates the need for a separate region proposal mechanism, thereby achieving markedly enhanced inference speed[25]. By leveraging multi-scale feature extraction and a diverse set of default bounding boxes, SSD exhibits robust detection capabilities across a spectrum of object sizes and aspect ratios, rendering it particularly efficacious for real-time embedded systems[26].

Within the purview of autonomous navigation for individuals with visual impairment, SSD offers a computationally efficient paradigm for the instantaneous recognition of obstacles, terrain features, and environmental affordances. Its aptitude for rapid processing with minimal hardware overhead makes it eminently suitable for deployment on portable assistive technologies. When integrated with haptic or auditory feedback interfaces, SSD facilitates timely spatial cognition, thereby augmenting the autonomy, safety, and situational awareness of visually impaired users navigating complex built environments[27].

2.6 Convolution Neural Networks (CNNs)

Convolutional Neural Networks represent a distinguished subset of deep neural architectures, meticulously engineered to process and interpret data exhibiting a spatial or temporal structure, such as visual imagery. Rooted in the principles of biological vision, CNNs emulate the human visual cortex by hierarchically learning features from input data through a series of structured layers[28]. At their core, CNNs are composed of convolutional layers, non-linear activation functions, pooling (subsampling) layers, and fully connected layers. The convolutional layer employs a set of learnable kernels or filters that systematically traverse the input tensor, executing discrete convolutional operations to extract localised features. This mechanism can be mathematically expressed as:

$$Y(i,j) = (X * K)(i.j) = \sum_{m} \sum_{n} X(i+m,j+n) . K(m, n) \text{---(6)}$$

where X signifies the input matrix, K denotes the convolutional kernel, and Y(i,j)) is the output feature map at spatial position (i,j).Following convolution, the incorporation of activation functions, such as the Rectified Linear Unit (ReLU), introduces non-linearity into the network, thereby augmenting its capacity to model intricate and non-trivial relationships[29]. Pooling layers, commonly implemented via max-pooling or average-pooling strategies, perform spatial downsampling, thereby reducing dimensionality, mitigating overfitting, and enhancing translational invariance. The concluding segment of a CNN involves fully connected layers, wherein the multidimensional feature maps are flattened and subjected to a dense neural structure to facilitate classification, detection, or regression tasks[30]. CNNs are celebrated for their exceptional proficiency in hierarchical feature extraction, progressively discerning elementary edges, textures, and shapes, culminating in the recognition of complex, abstract patterns. This renders them quintessential in a myriad of domains, including but not limited to computer vision, medical diagnostics, autonomous navigation, and remote sensing[31].

Table 1:Comparison of different Algorithms

Ref	Year	Title	Algorithm	Advantages	Drawbacks
[35]	2025	A hierarchical simulation-based push planner for autonomous recovery in navigation blocked scenarios of mobile robots	Hierarchical Push Planning Algorithm	This approach significantly enhances the robot's operational resilience and task continuity in dynamic environments.	Hierarchical simulation-based push planner incurs substantial computational demands due to its reliance on multi-level decision hierarchies and simulation-driven foresight
[36]	2025	Evolutionary optimization of spatially-distributed multi- sensors placement for indoor surveillance environments with security levels	Genetic Algorithm	Maximises resource utilisation while adapting to dynamic spatial configurations, enhancing system resilience and reliability.	The evolutionary optimisation of spatially-distributed multi-sensor placement may incur significant computational overhead, which could hinder real-time deployment in large-scale environments.
[37]	2023	Quasi-static balancing for a biped robot to perform extreme postures using a ducted-fan propulsion system	Quasi-Static Balancing Algorithm	The ducted-fan propulsion system enhances a biped robot's ability to maintain stability in extreme postures, providing rapid, precise balance adjustments.	The use of a ducted-fan propulsion system for quasi-static balancing in biped robots can lead to increased energy consumption, as continuous fan adjustments may require significant power.
[38]	2025	Robust localisation and tracking control of high-clearance robot system servicing high-throughput wheat phenotyping	Extended Kalman Filter	The robust localisation and tracking control system ensures precise navigation and data collection in dynamic agricultural environments. Enhancing wheat phenotyping efficiency.	The system may face high computational requirements, limiting its real-time performance in large-scale fields.
		Cyber-threat landscape of border control infrastructures	Recurrent Neural Network	Analysing the cyber-threat landscape of border control infrastructures	Analysing the cyber-threat landscape of border control infrastructures can be

[39]	2022			helps identify vulnerabilities proactively, enhancing system resilience.	resource-intensive, requiring advanced tools and skilled personnel.
[40]	2025	Evolutionary optimization of spatially-distributed multi- sensors placement for indoor surveillance environments with security levels	Genetic Algorithm (GA)	Evolutionary optimisation ensures efficient sensor placement, maximising coverage and minimising blind spots in indoor surveillance.	The evolutionary optimisation approach can be computationally expensive and time-consuming.
[41]	2024	Self-adaptive bifold-objective rate optimization algorithm for Wireless Sensor Networks	Self-Adaptive Genetic Algorithm	Enhances energy efficiency and prolongs the network lifespan by adaptively adjusting data transmission rates	High computational overhead, which is challenging for low- power sensor nodes.

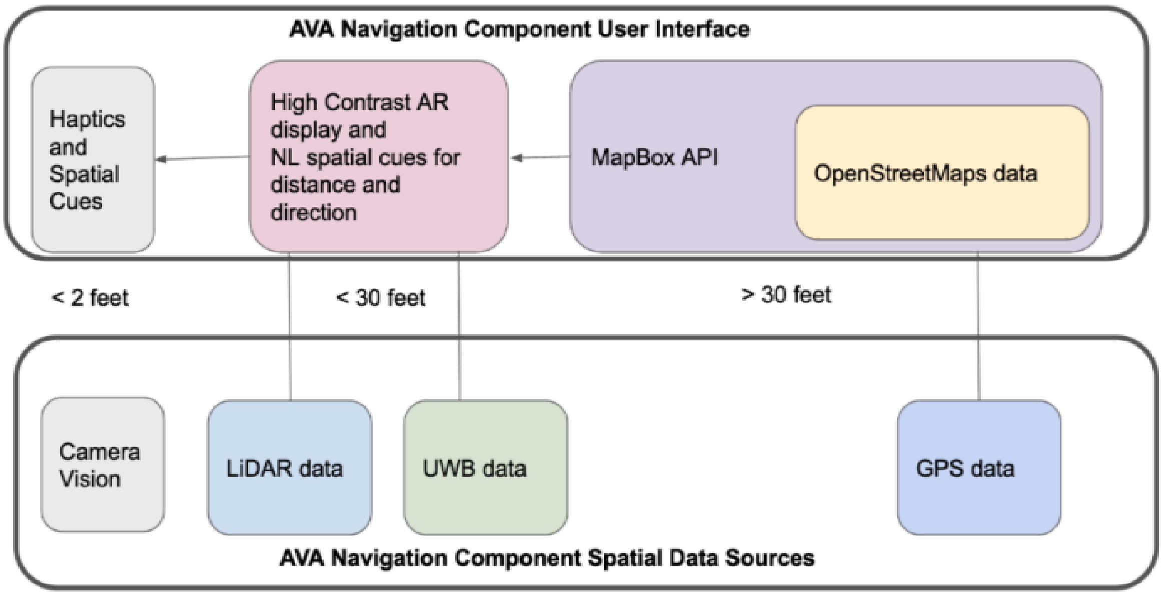

Fig.:2. High-level architecture of the AVA navigation component

Fig 2 delineates the participatory design paradigm underpinning the award-winning Autonomous Vehicle Assistant (AVA), conceived to furnish bespoke mobility solutions for individuals with visual impairment. Empirical insights gleaned from an initial survey (n = 90) and a sequence of semi-structured interviews (n = 12) informed AVA's innovative functional repertoire, which was subsequently assessed via a rigorous navigational protocol (n = 10) and collaborative evaluative workshops (n = 6). Confluent results intimate that AVA's multimodal sensor integration—amalgamating advanced computer vision, terminal-approach facilitation, and poly-sensory notifications—proffers pivotal enhancements for end-users optimally positioned to capitalise on this nascent transportation modality[38] [39].

This exegesis propounds a dilated tri-dimensional convolutional schema to accentuate perceptual-salience cues in navigation apparatuses for the visually impaired. The Spatio-Temporal Feature Fusion (STFF) construct amalgamates protracted spatio-temporal embeddings, thereby reinforcing temporal congruity across successive frames[40] [41]. In parallel, the Multi-Scale Feature Fusion (MSFF) apparatus orchestrates an integrative synthesis of disparate-scale representations and meticulously refines spatial encodings to preclude superfluous feature proliferation. Empirical validation on the DHF1K, Hollywood-2, and UCF Sports corpora—assessed via quintuplicate canonical metrics—corroborates its ascendancy over extant state-of-the-art paradigms[42] [43].

This treatise elucidates an avant-garde paradigm to ameliorate the latent vicissitudes engendered by pavement aberrations for visually impaired sojourners. Such subsurface cavities, by virtue of their inscrutable morphology, frequently precipitate corporeal trauma, imperilling ambulatory autonomy[44]. To obviate this hazard and promulgate unfettered perambulation, we advocate a schema predicated upon the You Only Look Once (YOLO) object-detection algorithm, which effectuates real-time demarcation of deformities. Upon identification, the system dispatches discretely calibrated auditory or haptic prompts, thereby endowing users with the perspicacity to preemptively eschew these menaces and sustain secure, self-determined locomotion irrespective of visual acuity[45].

3. Prevailing security technologies

Unmanned Surface Vehicles (USVs) encompass an extensive repertoire of sophisticated modalities to augment maritime situational awareness and safeguard operational integrity. High-resolution radar arrays and LiDAR systems confer meticulous obstacle delineation and precision collision avoidance, while acoustic sonar apparatuses extend subaqueous surveillance capabilities beyond visual confines[46]. Automatic Identification Systems (AIS) and satellite-facilitated Automatic Dependent Surveillance–Broadcast (ADS-B) mechanisms ensure persistent vessel monitoring and geospatial tracking across vast littoral expanses. Furthermore, Electro-Optical/Infrared (EO/IR) imaging suites, augmented with thermal and hyper spectral sensors, enable nuanced target identification under diverse environmental and temporal conditions. To fortify data security and preserve command-and-control sanctity, USVs deploy end-to-end encryption frameworks, frequency-hopping spread spectrum telemetry, and blockchain-anchored audit trails, collectively fostering a resilient and impervious maritime security architecture[47] [48].

4. System Architectures

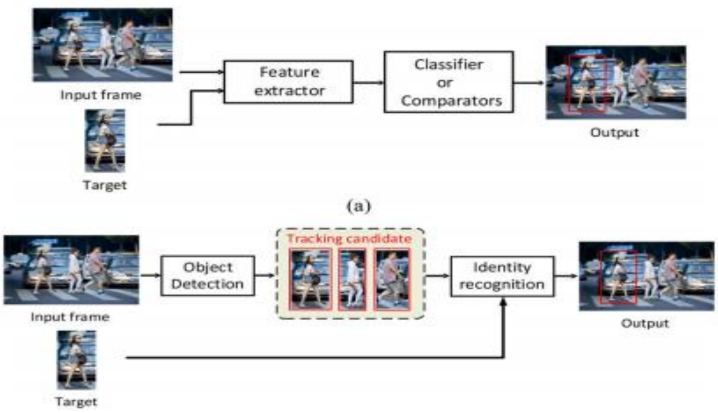

Fig: 3. Categories of human tracking algorithms (a) Template matching method, (b) Detection based method.

Numerous deep learning methodologies have been conceptualized and subsequently adapted for deployment on embedded development boards. Diverging from conventional paradigms, the advancement of deep learning algorithms necessitates meticulous consideration of hardware-specific constraints and capabilities. In Reference [49], C. Lee introduced a lightweight human detection model predicated on the YOLO object detection framework, specifically optimized for embedded system integration. proposed a comprehensive framework amalgamating deep learning techniques for environmental perception with variational Bayesian methodologies for robust trajectory prediction.

5. System Architectures

1. Enhanced Object Detection: The improved YOLOv8 model demonstrates superior performance in detecting various objects, such as vehicles and pedestrians, by effectively extracting and focusing on key features within images. This leads to more accurate identification and positioning of objects on roadways.

2. Real-Time Processing: The optimized model maintains high-speed processing capabilities, enabling real-time analysis of video streams or images[50]. This is crucial for autonomous vehicles to respond promptly to dynamic changes in their environment.

3. Robustness in Diverse Conditions: By incorporating RFAConv and Triplet Attention, the model exhibits increased robustness against various challenging conditions, such as occlusions, varying lighting, and complex backgrounds, ensuring reliable performance in real-world scenarios.

4. Improved Feature Extraction**:** The RFAConv module enhances the model's ability to extract relevant features by expanding the receptive field, while the Triplet Attention mechanism allows the model to focus more precisely on important areas within the image, leading to improved detection accuracy.

6. Conclusion

The evolution of autonomous navigation systems tailored for visually impaired individuals signifies a profound advancement in assistive technology and inclusive design. Cutting-edge progress in sensor integration, artificial intelligence, and machine perception has enabled these robots to navigate complex and dynamic environments with increased precision and reliability. Key developments in obstacle avoidance, spatial mapping, and user feedback mechanisms have substantially enhanced their operational efficacy and user acceptance. Nevertheless, persisting challenges—particularly those relating to cost-effectiveness, environmental adaptability, and user confidence—necessitate continued research. A multidisciplinary approach, grounded in ethical considerations and guided by the lived experiences of visually impaired users, remains imperative to realise the full potential of these technologies in fostering autonomy and social inclusion.

References

[1] Wasiq Khan 2023,Outdoor mobility aid for people with visual impairment: Obstacle detection and responsive framework for the scene perception during the outdoor mobility of people with visual impairment.

[2] Ahmed Ben Atitallah 2023 ,An effective obstacle detection system using deep learning advantages to aid blind and visually impaired navigation.

[3] Rashik Iram Chowdhury2024 ,A novel edge intelligence-based solution for safer footpath navigation of visually impaired using computer vision.

[4] Durgesh Nandini, K.R. Seeja2017.A novel path planning algorithm for visually impaired people.

[5] Ezra Pedzisai 2023, A novel framework to redefine societal disability as technologically-enabled ability: A case of multi-disciplinary innovations for safe autonomous spatial navigation for persons with visual impairment.

[6] Yassine Bouteraa 2022 ,Smart real time wearable navigation support system for BVIP.

[7] Charis Ntakolia 2023,Autonomous path planning with obstacle avoidance for smart assistive systems.

[8] Florian ScalviniuB-Visio Geoloc: An image sequences dataset of pedestrian navigation including geolocalised-inertial information and spatial sound rendering of the urban environment's obstacles.

[9] Rashik Iram Chowdhury 2024 ,A novel edge intelligence-based solution for safer footpath navigation of visually impaired using computer vision.

[10] Leandro C. Fernandes 2013, CaRINA Intelligent Robotic Car: Architectural design and applications.

[11] Jinhui Zhu 2020,A fog computing model for implementing motion guide to visually impaired.

[12] Ruixin Zhao 2024, Enhancing autonomous drivingsafety:Arobust traffic sign detection and recognition model TSD-YOLO.

[13] Fei WANG 2024,Deep-reinforcement-learning-based UAV autonomous navigation and collision avoidance in unknown environments.

[14] Yongchang Xie 2024,Reliable LiDAR-based ship detection and tracking for Autonomous Surface Vehicles in busy maritime environments.

[15] Homayun Kabir 2023, Internet of robotic things for mobile robots: Concepts, technologies, challenges, applications, and future directions.

[16] Dahang Wan 2024, YOLO-MIF: Improved YOLOv8 with Multi-Information fusion for object detection in Gray-Scale images.

[17] Yuhang Liu 2025, PV-YOLO: A lightweight pedestrian and vehicle detection model based on improved YOLOv8.

[18] Shizhou Xu 2024, YOLO-Hyper Vision: A vision transformer backbone-based enhancement of YOLOv5 for detection of dynamic traffic information.

[19] Rashmi 2024, SD-YOLO-AWDNet: A hybrid approach for smart object detection in challenging weather for self-driving cars.

[20] Yan-Lin Chen 2024, Transformer-CNN for small image object detection.

[21] Zhiguo Liu 2025,Rotating-YOLO: A novel YOLO model for remote sensing rotating object detection.

[22] Bowen Jiao2025, RS-YOLO: An efficient object detection algorithm for road scenes.

[23] Changdong Yu 2024, YOLO-MRS: An efficient deep learning-based maritime object detection method for unmanned surface vehicles.

[24] François Routhier 2024,Three pedestrian phasing with audible pedestrian signals configurations: Experience of blind or visually impaired persons in Quebec City (Canada).

[25] Jiaji Wang 2023,Artificial intelligence for visually impaired.

[26] Raihan Bin Islam 2023,Deep learning based object detection and surrounding environment description for visually impaired people.

[27] Aristides Mpitziopoulos 2011,A pervasive assistive environment for visually impaired people using wireless in manufacturing industry.

[28] Vladislav L 2025,Enhancing 3D object detection in autonomous vehicles based on synthetic virtual environment analysis.

[29] Cosimo Patruno 2024,Vision-based omnidirectional indoor robots for autonomous navigation and localization in manufacturing industry.

[30] Muhammad Faseeh 2024 ,Deep learning assisted real-time object recognition and depth estimation for enhancing emergency response in adaptive environment.

[31] Yu-Cheol Lee 2024 LSMCL: Long-term Static Mapping and Cloning Localization for autonomous robot navigation using 3D LiDAR in dynamic environments.

[32] Ruixin Zhao 2024, Enhancing autonomous drivingsafety:Arobust traffic sign detection and recognition model TSD-YOLO.

[33] Fei WANG 2024,Deep-reinforcement-learning-based UAV autonomous navigation and collision avoidance in unknown environments.

[34] Yongchang Xie 2024,Reliable LiDAR-based ship detection and tracking for Autonomous Surface Vehicles in busy maritime environments.

[35] Homayun Kabir 2023, Internet of robotic things for mobile robots: Concepts, technologies, challenges, applications, and future directions.

[36] Dahang Wan 2024, YOLO-MIF: Improved YOLOv8 with Multi-Information fusion for object detection in Gray-Scale images.

[37] Yuhang Liu 2025, PV-YOLO: A lightweight pedestrian and vehicle detection model based on improved YOLOv8.

[38] Paul D.S. Fink 2023 ,The Autonomous Vehicle Assistant (AVA): Emerging technology design supporting blind and visually impaired travelers in autonomous transportation.

[39] Fangfang Lu 2023,Visual saliency assistance mechanism based on visually impaired navigation systems.

[40] Arjun Paramarthalingam 2024, A deep learning model to assist visually impaired in pothole detection using computer vision.

[41] Anandh Nagarajan 2023,Hybrid Optimization-Enabled Deep Learning for Indoor Object Detection and Distance Estimation to Assist Visually Impaired Persons.

[42] Mahmoud M. Selim, 2025, Enhancing Arabic text-to-speech synthesis for emotional expression in visually impaired individuals using the artificial hummingbird and hybrid deep learning model.

[43] Huilin Gao 2024, Are real-time volunteer apps really helping visually impaired people? A social justice perspective.

[44] Jun Tang 2024, YOLO-Fusion and Internet of Things: Advancing object detection in smart transportation.

[45] K.P. Ajitha Gladis 2024, In-out YOLO glass: Indoor-outdoor object detection using adaptive spatial pooling squeeze and attention YOLO network.

[46] Zuxiang Situ 2023, A transfer learning-based YOLO network for sewer defect detection in comparison to classic object detection methods.

[47] Qamar Ul Islam 2025,advancing autonomous SLAM systems: Integrating YOLO object detection and enhanced loop closure techniques for robust environment mapping.

[48] Tsung-Han Tsai 2024, Design and implementation of deep learning-based object detection and tracking system.

[49] Gang Liu 2025, Object detection algorithm for autonomous driving: Design and real-time performance analysis of AttenRetina model.

[50] Zhipeng Ling 2024, Optimization of autonomous driving image detection based on RFAConv and triplet attention.