iJRASET

International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: XI Month of publication: November 2025

DOI: https://doi.org/10.22214/ijraset.2025.75763

Smart Assistive Device for Visually Impaired Individuals: A Survey

Aditya Pankaj Nikam[1], Purushottam Anil Gaikwad[2], Darshan Pandharinath Mekhe[3], Prof. Sonali Jadhav[4]
Dept. Artificial Intelligence and Data Science, MET Institute of Engineering Nashik, Maharashtra, India

Abstract: Recent advances in artificial intelligence and low- cost embedded computing have spurred the development of transformative technologies for the visually impaired. This paper provides a comprehensive survey of modern, vision-based wear-able systems designed to enhance navigation and environmental perception. We systematically review and analyze key research from 2015 to the present, categorizing systems based on their computational architecture (standalone vs. cloud-connected), form factor, and primary function. Our analysis covers the core technological pillars, including real-time navigation using GPS and sensor fusion, deep learning-based obstacle detection, and object recognition with Convolutional Neural Networks (CNNs). We specifically highlight the field's definitive shift from traditional computer vision algorithms to these more powerful deep learning models. By comparing the performance, cost, and usability of various approaches, we identify persistent challenges that hinder widespread adoption, such as limited battery life, performance in varied lighting conditions, and the high cost of commercial devices. This survey concludes by outlining the most significant research gaps and promising future directions, including the need for more efficient AI architectures and robust multi-modal sensor integration for creating the next generation of effective and accessible assistive devices.

Index Terms: Survey, Assistive Technology, Computer Vision, Deep Learning, Embedded Systems, Visually Impaired, Wearable Computing

I. INTRODUCTION

Enhancing the autonomy and safety of the world's 2.2 billion visually impaired individuals is a critical and pressing research area[1]. For this community, navigating unfamiliar environments and perceiving the details of the physical world pose significant daily challenges, often limiting mobility and independence. While traditional aids have been foundational, the rapid proliferation of artificial intelligence and low-cost embedded hardware has opened a new frontier for develop- ing powerful, wearable assistive technologies. These systems aim to translate complex visual information into accessible formats, but the sheer diversity of approaches has spurred the need for a systematic overview of the field.

Modern assistive systems have moved past simple sensors and older statistical methods to use deep learning (DL) tech- niques for better real-time performance. The goal is to balance speed and accuracy, which is crucial for navigation assistance. One-stage object detectors like YOLOv8 are current top choices, known for being fast and highly accurate (e.g., YOLOv8s achieved a high mAP of 0.974)[2],[3]. Compared to complex two-stage detectors (like Faster R-CNN), one-stage models are computationally quicker. For tasks requiring the highest precision and confidence scores (up to 99), models like the Detection Transformer (DETR) are utilized, although they may be slightly slower than YOLO models. Other core DL models employed include MobileNet architectures (often in versions like V2 or V3), which are intentionally designed to be lightweight for embedded vision applications. For complete environmental understanding, complex architectures and methods are now frequently used. Instead of just detecting obstacles, semantic segmentation models, like the Quantized and Pruned UNet-based Lightweight MobileNet Model (QPULM)[5], are trained to find the safe, walkable paths. Researchers also employ custom multi-branch architectures designed for efficient pathfinding. To ensure these complex models run quickly on small, affordable devices like Raspberry Pi or smartphones, techniques such as quantization are essential. Quantization reduces the model size significantly (e.g., by converting weights to 8-bit integers) without losing much accuracy, enabling fast inference speed on edge devices. Additionally, advanced systems integrate features like distance estimation (often using ultrasonic sensors or methods like SODD) and leverage Large Vision-Language Models (LVLMs) (like GPT-4o-mini) to give users rich, de- tailed, spoken descriptions and contextual awareness of their surroundings. This survey brings together the system architectures, soft- ware methodologies, and evaluation strategies from traditional sensor-based aids to cutting-edge, AI-powered wearable sys- tems. By gathering and synthesizing existing contributions, this paper provides a systematic overview of the assistive technology landscape for the visually impaired. It highlights persistent challenges such as real-time performance on embedded hardware, power consumption, and affordability, and proposes future research directions for the development of robust, user-centric, and truly accessible assistive devices.

II. LITERATURE SURVEY

The development of assistive technologies for visually im- paired individuals has been widely researched in recent years using a variety of methodologies, sensor integrations, and artificial intelligence approaches. This section provides an organized overview of previous studies focused on enhanc- ing mobility, safety, and independence for visually impaired users, with an emphasis on the technological contributions, AI techniques, and datasets utilized in these works.

A. Systematic Reviews and Surveys

Early comprehensive reviews classified assistive solutions based on hardware, such as camera-based and sensor-based systems, and identified critical challenges like low-light detection, prediction of dynamic obstacles, large prototype size, and high costs. A systematic review on smart glasses applications confirmed a steady increase in research since the release of Google Glass. Analysis of user needs highlights that a substantial portion of VI individuals report difficulties with independent movement (90 percent), object recognition (40 percent), and reading text (20 percent). Traditional assistive tools, such as white canes, are considered helpful but insufficient for providing comprehensive support. Consequently, user requirements prioritize devices that offer high-quality audio descriptions, accuracy of directions/information, alerts for unexpected events, and guaranteed real-time performance. Current systematic efforts also emphasize the necessity of robust solutions for navigating complex, changing urban areas, advocating for greater accuracy, improved planning, robust problem detection, and user-centered design to address the limitations of existing Electronic Travel Aids (ETAs)

B. Classical Machine Learning and Sensor-Based Methods

Initial assistive technologies relied heavily on non-vision- based sensors, including ultrasonic, infrared (IR), inertial, and magnetic sensing, for obstacle detection and distance estimation. Classic systems often utilized simple, low-cost components like Arduino and ultrasonic sensors, sometimes augmented with GPS and GSM systems for providing location services or connecting users to caregivers. Early software approaches utilized foundational computer vision techniques, such as the Viola Jones algorithm and Haar filters, and later incorporated conventional Convolutional Neural Networks (CNNs). For instance, widely recognized CNN architectures like VGG16, ResNet, and GoogLeNet (Inception v1) were influential in improving performance for image classification tasks in early prototypes. Specialized systems have also been developed using advanced non-visual sensors, such as compact mmWave radar (due to affordability and robustness against harsh environmental conditions) combined with Inertial Mea- surement Units (IMUs) for tracking natural body movement. However, traditional sensor-based systems often suffer from key limitations, including a short detection range (frequently less than 2 meters) and sensitivity to environmental factors

C. Advanced Contextual Understanding

The progression of assistive technology shows a movement toward providing rich contextual information beyond sim- ple object labels. The emergence of Large Vision-Language Models (LVLMs) is recognized for providing contextually rich environmental information, integrating real-time object recognition with contextual insights for tasks like shopping assistance. Early integration involved utilizing cloud services like the Google Vision API and Azure Cognitive Services (Custom Vision API) to generate scene captions, read text, or perform face recognition. Systems such as VisBuddy im- plemented advanced text reading capabilities, utilizing the Inception-V3 model as an encoder for scene recognition and developing a dynamic column determination algorithm to accurately read text organized in multiple columns (an improvement over previous works limited to two columns). Efforts also focus on enhancing social interaction. Research utilizes smart glasses and deep learning for real-time social assistance, frequently employing multimodal approaches that combine data from voice and facial expressions for emotion recognition. Future plans for some assistive devices include incorporating a multilingual feedback system to assist local users who may have difficulty speaking or understanding English

D. Obstacle and Early Hazard Detection Systems

In the context of VI assistance, ”early detection” refers to the timely and precise identification of immediate phys- ical hazards, obstacles, and path instability to prevent ac- cidents.Semantic segmentation models(like the Quantized and Pruned UNet-based Lightweight MobileNet Model, or QPULM) play a crucial role by classifying each pixel in an image to define walkable (safe) vs. non-walkable (unsafe) paths. Advanced semantic segmentation aims to achieve low computational overhead and real-time performance for foot- path navigation.Targeted detection efforts focus on hazards like staircases (upstairs/downstairs) using vision or 3D point cloud data, and low-lying obstructions such as potholes or pits

III. CHALLENGES

Despite the major advances in developing camera-based and sensor-based assistive devices, current methods still encounter a series of key challenges that hinder their efficacy, affordabil- ity, and generalization in real-world scenarios

A. Mobility and Environmental Challenges

Visually impaired individuals face significant challenges navigating urban environments, particularly in places like Dhaka or in India where footpaths are often unstructured and lack continuity, being frequently obstructed by construction materials, roadside hawkers, electrical poles, and trees. Beyond these dynamic obstacles, users struggle with identifying subtle ground hazards such as road curbs, changes in the surface of the road, and uneven floor surfaces. Traditional aids like white canes are acknowledged as having limitations, as they can only sense obstacles within their range and fail to detect dangers above ground level, such as overhanging objects. Furthermore, system efficacy is highly dependent on environmental factors, with performance metrics (like accuracy) notably decreasing in low-light conditions, and other factors like rain or snow also affecting basic sensor reliability

B. Technical Limitations and Performance Bottlenecks

The critical requirement for real-time object detection ne- cessitates careful balancing of model complexity against hard- ware limitations. Developers often utilize affordable embed- ded devices like Raspberry Pi, which suffer from constraints in processing power, memory, and battery life. This low computational capability makes computationally expensive two-stage detectors (like Faster R-CNN) less suitable than optimized one-stage models (like YOLO), which utilize an anchor-free detection approach. Despite using custom models, one prototype gadget still reported a rendering time delay of approximately 17.304 seconds per frame during dynamic testing. Model performance also degrades in poor visibility; for example, a system showed precision dropping from 0.88 (Good Lighting) to 0.76 (Low-Light), and overall accuracy falling from 90 percent to 80 percent. To improve deploy- ment efficiency on constrained devices, Quantization Analysis is performed on high-performing models like YOLOv8s to evaluate performance retention during size reduction.

C. Accessibility, Cost, and Usability Barriers

The widespread adoption of assistive technology is funda- mentally challenged by high cost and design barriers, partic- ularly because an estimated 2.2 billion people are visually impaired globally, with about 90 percent living in low-income settings. This financial reality makes many state-of-the-art detection systems too expensive for the target audience. Even when efforts are made to use cost-effective components for a lightweight and compact design, usability issues emerge, such as devices being too bulky or too heavy for comfortable daily wear. Furthermore, device operation can be problematic: some users found that overusing the application sometimes led to the device’s heating, while complexity in voice outputs—such as the requirement for voice accuracy and audibility—can reduce ease of use. Finally, systems that rely on offloading computa- tion or services, such as signboard detection or text recognition (OCR) using Cloud Vision API or database retrieval, face limitations due to limited internet connectivity in most rural areas

IV. FUTURE DIRECTIVES

A. Multilingual and Cross-Cultural Generalization

Future work needs to emphasize expanding the systems’ applicability across different geographical and linguistic con- texts. This involves extending existing datasets, such as those focused on only specific urban environments to capture scenes and conditions from other countries with tactile-paved foot- paths to enhance generalizability. For features like text recog- nition, future prototypes should focus on implementing a multilingual feedback system to benefit local users who have difficulty speaking or understanding English. Furthermore, studies recommend rigorous testing of assistive devices in dy- namic and diverse environmental settings to ensure consistent reliability globally.

B. Efficient and Scalable Architectures

To facilitate practical deployment at scale on portable, resource-constrained devices (like Raspberry Pi), future efforts must focus on optimizing deep learning models for com- putational efficiency. Research should investigate techniques like Quantization Analysis on best-performing models (such as YOLOv8s) to evaluate performance retention during size reduction, making them suitable for resource-constraint edge devices. Additionally, exploring other benchmark object de- tection architectures beyond conventional YOLO and Faster RCNN models, or incorporating custom hardware accelerators, can help streamline processing power, potentially reducing latency. Alternative approaches include shifting computation to a cloud-based assistive technology model by leveraging remote AI processing and high-speed networks to offload intensive tasks

C. Context-Aware and Explainable Systems

Future systems must move beyond simple object labeling to provide rich contextual understanding and enhance user trust through transparency. This involves integrating real-time contextual information using Large Language Models (LLMs), potentially leveraging high-performance GPUs and open- source models for seamless integration. Researchers should also explore multimodal data integration, such as developing a 360-degree video analysis system that combines audio and video data to provide comprehensive information about the user's surroundings. Furthermore, critical supplementary fea- tures must be added, including object distance estimation (po- tentially using LiDAR or stereo vision for enhanced precision) and GPS-based navigation for positioning and environmental awareness. Finally, adapting Explainable AI (XAI) techniques can be used to evaluate the interpretability of models and promote user trust and acceptance

V. CONCLUSION

The development of affordable and effective assistive tech- nologies for visually impaired people continues to be an important research area with great social value. This survey reviewed different modern vision-based wearable systems, ranging from early computer vision methods to the latest deep learning approaches. Traditional sensor-based and classical algorithms built a strong base, but they often fail to handle complex real-world situations effectively. Deep learning techniques, especially Convolutional Neural Networks (CNNs) used on small embedded devices like the Raspberry Pi, have shown much better accuracy. However, they still face challenges in achieving a balance between high performance, fast real-time response, and low power usage.

From the studies reviewed, one major gap identified is that most existing systems focus on only one task, such as navigation or object detection, instead of providing a complete, all-in-one solution. Many models that work well in controlled environments do not perform the same in outdoor or real- life conditions due to factors like poor lighting, glare, and unexpected obstacles.

Key issues such as real-time processing, battery life, afford- ability, and user comfort still need improvement. Future work should focus on creating lightweight AI models and using sensor fusion methods that combine cameras with technologies like LiDAR or ultrasonic sensors for better accuracy and reliability. It is also important to design easy-to-use interfaces and test these systems in real-life situations with visually im- paired users to ensure they are practical and socially accepted. Progress in this field will depend not only on technology but also on teamwork between engineers, designers, and the visually impaired community to truly improve independence and quality of life.

REFERENCES

[1] N. Yalcin and M. Alisawi, "Enhancing Social Interaction for the Visually Impaired: A Systematic Review of Real-Time Emotion Recognition Using Smart Glasses and Deep Learning," IEEE Access, Jun. 2025. doi: 10.1109/ACCESS.2025.3577106.

[2] S. Ikram, I. S. Bajwa, S. Gyawali, A. Ikram, and N. Alsubaie, "En- hancing Object Detection in Assistive Technology for the Visually Impaired: A DETR-Based Approach," IEEE Access, Apr. 2025. doi: 10.1109/ACCESS.2025.3558370.

[3] M. I. A. Hossain, J. Anjom, and R. I. Chowdhury, "Towards walkable footpath detection for the visually impaired on Bangladeshi roads with smartphones using deep edge intelligence,"Array, vol. 26, p. 100388, Apr. 2025. doi: 10.1016/j.array.2025.100388.

[4] R. Kra´l, P. Jacko, and T. Vince, "Low-Cost Multifunctional Assistive Device for Visually Impaired Individuals," IEEE Access, Mar. 2025. doi: 10.1109/ACCESS.2025.3554366.

[5] R. I. Chowdhury, J. Anjom, and M. I. A. Hossain, "A novel edge intelligence-based solution for safer footpath navigation of visually impaired using computer vision," Journal of King Saud University - Computer and Information Sciences, vol. 36, p. 102191, Sep. 2024. doi: 10.1016/j.jksuci.2024.102191.

[6] M. Alghamdi, H. A. Mengash, M. Aljebreen, M. Maray, A. A. Darem, and A. S. Salama, "Empowering Retail Through Advanced Consumer Product Recognition Using Aquila Optimization Algorithm With Deep Learning," IEEE Access, vol. 12, pp. 71055–71065, May 2024. doi: 10.1109/ACCESS.2024.3399480.

[7] M. S. A. Baig, S. A. Gillani, S. M. Shah, M. Aljawarneh, A. A. Khan, and M. H. Siddiqui, "AI-based Wearable Vision Assistance System for the Visually Impaired: Integrating Real-Time Object Recognition and Contextual Understanding Using Large Vision-Language Models," arXiv:2412.20059v1, 2024.

[8] A. M. Musharaff, G. S. Begum, P. R. Rao, and S. N. R. K., "iSight- Navigable Path Detection for Visually Impaired using Semantic Seg- mentation,"International Journal for Research in Applied Science and Engineering Technology, vol. 12, issue V, May 2024.

[9] F. Ashiq, M. Asif, M. B. Ahmad, S. Zafar, K. Masood, T. Mahmood, M. T. Mahmood, and I. H. Lee, "CNN-Based Object Recognition and Tracking System to Assist Visually Impaired People," IEEE Access, vol. 10, pp. 14819–14834, Jan. 2022. doi: 10.1109/ACCESS.2022.3148036.

[10] A. Dinodiya and R. R., "Smart Wearable Gadget to Detect Stairs, Potholes, and Other Objects for Visually Impaired Users Using Com- puter Vision,"Bachelor of Engineering Thesis, Dept. Mechatronics Eng., SATHYABAMA Institute of Science and Technology, Chennai, India, May 2022.

[11] H. Matsumura and C. Premachandra, "Deep-Learning-Based Stair De- tection Using 3D Point Cloud Data for Preventing Walking Accidents of the Visually Impaired," IEEE Access, vol. 10, May 2022. doi: 10.1109/ACCESS.2022.3178154.

[12] A. G. Song, X. Hu, and S. A. G. S, "StereoPilot: A Wearable Target Location System for Blind and Visually Impaired Using Spatial Audio Rendering,"IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2022. doi: 10.1109/TNSRE.2022.3182661.

[13] H. Ali A., S. U. Rao, S. Ranganath, T. S. Ashwin, and G. R. M. Reddy, "A Google Glass Based Real-Time Scene Analysis for the Visually Impaired,"IEEE Access, vol. 9, pp. 166351–166369, Dec. 2021. doi: 10.1109/ACCESS.2021.3135024.

[14] H. F. Á lvarez, G. Á lvarez-Narciandi, F. Las-Heras, and J. Laviada, ”Sys- tem Based on Compact mmWave Radar and Natural Body Movement for Assisting Visually Impaired People,” IEEE Access, vol. 9, pp. 129033- 129046, Sep. 2021. doi: 10.1109/ACCESS.2021.3110582.

[15] I. Sivakumar, N. Meenakshisundaram, I. Ramesh, S. E. D, and S. R. Raj C, ”VisBuddy - A Smart Wearable Assis- tant for the Visually Challenged,” 2020. [Online]. Available: https://link.springer.com/chapter/10.1007/978-981-15-5853-5-36

[16] P. S. Swami and P. Futane, ”Traffic Light Detection System for Visually Impaired Person with Voice System,” International Journal of Advance Engineering and Research Development, vol. 5, no. 04, Apr. 2018.

[17] A. Thakur, R. Singh, and A. Gehlot, ”Smart Blind Stick For Obstacle Detection and Navigation System,” International Journal of Emerging Technologies and Innovative Research, Oct. 2018.

[18] R. Kedia, K. K. Yoosuf, P. Dedeepya, M. Fazal, C. Arora, and M. Balakrishnan, ”MAVI: an embedded device to assist mobility of visually impaired,” in Proc. 30th International Conference on VLSI Design and 16th International Conference on Embedded Systems (VLSID), 2017, pp. 213-218. doi: 10.1109/VLSID.2017.38.

[19] R. Tapu, B. Mocanu, and T. Zaharia, ”DEEP-SEE: Joint object de- tection, tracking and recognition with application to visually impaired navigational assistance,”Sensors, vol. 17, no. 11, p. 2473, 2017. doi: 10.3390/s17112473.

[20] R. M. Y. B. Munger et al., “Apparatus and method for a dynamic “region of interest” in a display system,” US Patent 9,618,748, Apr. 11, 2017.