



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



---

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume:** 14    **Issue:** IV    **Month of publication:** April 2026

**DOI:** <https://doi.org/10.22214/ijraset.2026.81422>

[www.ijraset.com](http://www.ijraset.com)

Call:  08813907089

E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)

# A Comprehensive Review of Edge-AI Powered Assistive Systems for the Visually Impaired: Models, Hardware, and Future Directions

Harish Kunder, Aashish CA, Fawaz Shaikh, Shashank HN, Sujay SP

Department of Artificial Intelligence and Machine Learning Alva's Institute of Engineering and Technology Mangalore, Karnataka, India

**Abstract:** *Since more than 2.2 billion people worldwide suffer from visual impairment, assistive technology is an important field of study. Intelligent systems that assist blind and visually impaired (VI) people in their daily lives have been made possible by recent developments in computer vision, natural language processing, and edge artificial intelligence. Edge platforms like the NVIDIA Jetson family offer low latency, privacy, and offline dependability in contrast to cloud-based solutions.*

*This survey critically evaluates more than fifty commercial and academic systems in three areas: (1) deep learning models for object detection, image captioning, and OCR; (2) assistive applications for navigation, object recognition, and scene understanding; and (3) embedded hardware that allows for real-time inference under resource constraints. We examine usability gaps in commercial products and trade-offs between energy, accuracy, and latency.*

*Strict wearable energy budgets, limited reproducibility, multi-lingual accessibility, and dataset bias are some of the main obstacles. Additionally covered are promising avenues like multimodal fusion, voice-based interaction, federated learning, and culturally diverse datasets. For researchers and practitioners creating the upcoming generation of AI-powered accessibility technologies, this work acts as a road map.*

**Index Terms:** *Computer vision, Edge AI, multimodal learning, object detection, image captioning, assistive technology, NVIDIA Jetson, Blind or visually impaired*

## I. INTRODUCTION

At least one billion of the world's 2.2 billion visually impaired individuals are avoidable or untreated [1]. While modern assistive systems incorporate sensors and vision-based techniques for richer perception, traditional aids like white canes and guide dogs only offer limited functionality [4],[5]. Text recognition, object detection, and scene understanding have all greatly improved with deep learning [2],[3],[13],[14]. Natural-language descriptions of complex environments are now possible thanks to vision-language models [8]. The integration of these capabilities into daily life is demonstrated by commercial tools like Microsoft Seeing AI, Google Lookout, and OrCam MyEye [9]–[11]. In terms of usability, latency, and offline performance, comparative studies show trade-offs between cloud-based apps and wearables with edge capabilities [17],[18].

Researchers are increasingly using edge AI platforms like the Raspberry Pi with Coral TPU, the NVIDIA Jetson Nano, TX2, and Orin Nano to get around the drawbacks of relying too much on the cloud. These gadgets allow offline support, better privacy, and real-time inference with lower latency [12]. Benchmarks demonstrate their efficacy for assistive tasks; however, issues persist, including lack of open-source reproducibility, dataset bias that restricts generalization, and multilingual OCR and speech support [15],[16].

Three areas are the focus of this survey, which summarizes research from more than fifty academic and commercial works: (1) assistive systems for navigation, scene understanding, and everyday help [7],[18]; (2) multimodal and deep learning models that drive OCR, detection, and captioning [2],[3],[8],[13],[14],[19]; and (3) edge hardware that allows deployment in resource-constrained environments [12]. The study highlights the opportunities and difficulties in developing the next generation of AI-powered assistive technologies for blind and VI users by looking at approaches, trade-offs, and practical prototypes.

## II. BACKGROUND

Over the past 20 years, assistive technologies for the blind and VI have steadily improved, moving from simple sensing devices to complex edge-AI-powered systems.

To identify obstacles and provide haptic or auditory cues, early solutions used ultrasonic or infrared sensors. They frequently failed in complex environments and were unable to convey semantic context, despite being effective in controlled settings [4],[5].

Five key phases in this progression are depicted in Fig. 1:

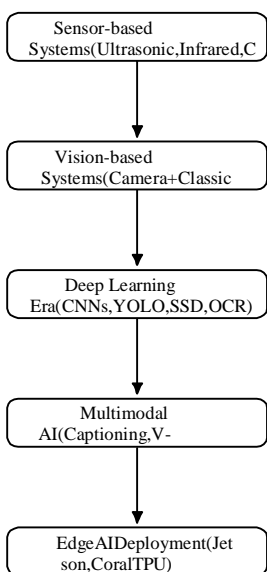


Fig. 1. Timeline of assistive technologies: from sensor-based aids to edge-AI deployments.

- Sensor-based: Although they lacked object-level recognition, infrared/ultrasonic modules allowed for basic obstacle alerts.
- Vision-based: Despite introducing depth estimation and lane following, cameras with classical CV struggled in real-world variability [7].
- Deep Learning: Real-time object detection and OCR were made possible by CNNs, YOLO, and SSD, which also supported text access and navigation [2],[3],[13].
- Multimodal AI: Semantic scene understanding (e.g., “bus approaching”) was made possible by vision–language (V-L) models and captioning systems [8].
- Edge AI: By moving inference to the device, gadgets like the Jetson Nano/TX2/Orin and Coral TPU improved privacy and decreased latency [12].

Reproducibility, multilingual OCR, dataset bias, and hardware trade-offs were among the new issues that each stage brought while addressing previous ones [15],[16]. The present research frontier is now defined by these problems.

#### ASSISTIVE TECHNOLOGIES FOR THE VISUALLY IMPAIRED

Simple sensing modules and sophisticated AI-driven frameworks are examples of assistive technologies. Obstacle alerts were provided by early tools like depth-based devices and ultrasonic canes, but they were unable to record semantic context [4],[5]. A taxonomy of existing solutions is depicted in Fig. 2, which demonstrates the transition from task-specific designs (OCR, navigation) to multimodal frameworks and commercial products.

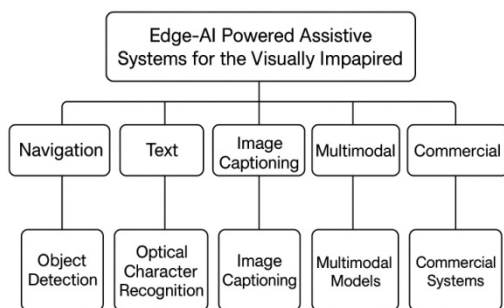


Fig. 2. Taxonomy of Edge-AI powered assistive systems, grouped into navigation, text recognition, captioning, multimodal frameworks, and commercial products.

Thesystemsfallintothefollowinggeneralcategories:

- **Navigation:** Mobility and obstacle avoidance are made possible by object detection and depth estimation (e.g., YOLO, SSD) [2],[3],[6]. Their usefulness is confirmed by real-world prototypes like NavWear [18].
- **Text Recognition:** Currency, product, and document reading are supported by OCR systems (CNN–RNNwith attention). Accessibility is increased by multilingual extensions [13],[14].
- **Image Captioning:** Captioners improve contextual awareness by using natural language to describe environments,suchas”abusapproaching”[8].Open-vocabulary supportisaddedbyextensionssuchasObjectFinder[19].
- **Multimodal Frameworks:** For complete situational awareness, unified pipelines combine OCR, detection, and captioning.
- **Business Systems:** While OrCam MyEye [11]offers partial edge inference, apps such as Seeing AI [9] and Lookout [10] offer cloud-based services. User-centered trade-offs are highlighted by comparative studies [17].

TABLE I  
COMPARISONOFCOMMERCIALASSISTIVESYSTEMS

System	Platform	Connectivity	KeyFeatures	OfflineSupport
SeeingAI	MobileApp	Cloud	Object&textrecognition,scenedescription	No
Google Lookout	MobileApp	Cloud	Textreading,currencyrecognition,productlabels	No
OrCamMyEye	Wearable	PartialEdge	OCR,facerecognition,real-timeaudiofeedback	Yes(partial)

Beyond academic prototypes, commercial deployments offer useful insights. Although cloud-based apps like Lookout and Seeing AI achieve high accuracy, they rely on connectivity, which raises privacy and latency issues [9],[10]. In low-connectivity environments, wearables like OrCam MyEye provide improved usability and faster edge inference [11]. All things considered, the taxonomy and comparisons show the variety of solutions as well as the practical compromises between inclusivity, accuracy, responsiveness, and portability.

### III. MODELS AND ALGORITHMS

Assistive technologies rely heavily on deep learning models, which are generally divided into three categories: text recognition, object detection/scene understanding, and image captioning. Each makes it possible for systems to identify, decipher, and convey visual information: object detection guarantees safe navigation [2],[3], OCR facilitates reading [13],[14], and captioning offers semantic context [8].

These models are integrated into unified pipelines that link sensors, inference modules, and user feedback instead of being used separately. For example, a wearable can read bus numbers (OCR), identify cars (object detection), and describe the environment (captioning). Such integration is demonstrated by commercial tools such as Seeing AI [9], Google Lookout [10], and OrCam MyEye [11], while usability trade-offs are highlighted by user studies (e.g., NavWear [18], OrCam vs. Seeing AI [17]).

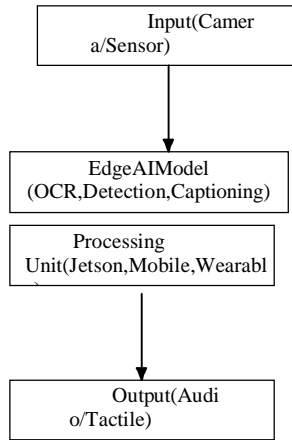


Fig.3. Compact assistive pipeline: input is processed by edge AI models (YOLO [2], SSD [3], OCR [13],[14]) to produce feedback.

### A. Image Captioning

From visual input, captioning models produce descriptions in natural language. By aligning text with visual features, transformer-based methods perform better than RNNs, and multimodal pretraining (e.g., depth-based captioning [8]) enhances fluency. High latency and wearable computing costs are obstacles. Model distillation, pruning, and quantization are some of the solutions. Prototypes based on Jetson achieve 5–10 frames per second [16], which is slower than cloud but sufficient for static descriptions.

### B. Object Detection and Scene Understanding

Real-time pedestrian and obstacle detection is essential for navigation.

- YOLO: Often utilized in navigation prototypes, it strikes a balance between speed and accuracy (about 30 frames per second on embedded GPUs) [2].
- SSD: Mobile-friendly and lightweight, but less precise [3]. Faster R-CNN: High precision, but too sluggish for deployment on the edge.

By simulating spatial arrangements, scene interpretation is further enhanced (e.g., car + pedestrian in crosswalk). With the majority of training data coming from urban settings in the West, dataset bias restricts generalization [15]. The goal of synthetic data and domain adaptation is to lessen this.

TABLE II  
EVALUATION OF WELL-KNOWN OBJECT DETECTION MODELS FOR ASSISTIVE APPLICATIONS

Model	FPS (Edge Device)	Latency (ms)	mAP/Accuracy	Notes
YOLOv3/v4	25–30	~35–40	High (0.55–0.60 mAP)	Widely used; excellent speed-accuracy balance
YOLOv5 (Nano/Tiny)	30+	<30	Medium (0.45–0.55 mAP)	Designed for embedded devices, lightweight
SSD (MobileNet backbone)	20–25	~40–50	Medium (0.40–0.50 mAP)	Portable, effective, slightly less accurate than YOLO
Faster R-CNN	<10	>100	High (0.65–0.70 mAP)	Very accurate but slow; unsuitable for real-time assistive use

EfficientD et (D0/D1)	15–20	~50– 60	High(0.60 – 0.65mAP)	Efficientscaling, butreq uiresmorecompute
--------------------------	-------	------------	----------------------------	--

### C. Text Recognition (OCR)

OCR makes it possible to read documents, currency, and signs. CNN feature extraction, RNNs/Transformers, and attention for robustness under noise are all combined in modern methods [13],[14]. The primary obstacle to multilingual OCR is still that complex scripts (such as Arabic, Hindi, and Chinese) continue to produce high error rates, despite Latin scripts performing well. Multilingual pretraining and effective architectures like MobileOCR, which are appropriate for edge devices, are examples of solutions.

### D. Discussion

For more comprehensive feedback, multimodal assistive systems combine detection, OCR, and captioning. Despite advancements, significant obstacles still exist:

- Latency/energy: Complex models put embedded devices under stress, necessitating optimization and compression [16].
- Dataset bias: Cultural and geographic diversity is lacking in training data [15].
- Multilingual OCR: Non-Latin scripts continue to have low accuracy [13],[14]. The user’s assessment: Blind participants are rarely included in studies, with the exception of OrCam vs. Seeing AI [17] and NavWear [18].

Future developments include lightweight transformers for on-device deployment, federated learning for privacy-preserving adaptation, and hybrid edge–cloud inference.

## IV. EDGE AI HARDWARE PLATFORMS

Because they allow for offline inference, low latency, and user privacy, edge AI platforms are essential to assistive systems [16]. However, there are trade-offs between cost, portability, power, and throughput when choosing hardware [16].

The NVIDIA Jetson Series The de facto standard for deep learning at the edge is the Jetson family: - Nano: 15–20 FPS, ~60 ms latency, appropriate for basic navigation and OCR. - TX2: 30+ frames per second with moderate power, commonly utilized in wearable prototypes. - Orin Nano: 60+ FPS at <20 ms, effective for multimodal V-L models [16].

On Nano, lightweight YOLO versions facilitate obstacle detection [2], and Orin makes it possible to integrate real-time captioning and OCR.

TABLE III  
BENCHMARK SUMMARY OF EDGE AI PLATFORMS

Device	FPS	Latency (ms)	Power (W)	Notes
Jetson Nano	15–20	~60	5–10	Low-cost; basic OCR/detection
Jetson TX2	30–35	~30	7–15	Balanced; wearable-friendly
Jetson Orin Nano	60+	<20	10–20	Runs multimodal/transformers
Raspberry Pi + Coral TPU	25–30	~35	5–7	Budget; task-specific accel.

### A. Other Platforms

The Raspberry Pi with Coral TPU is a desirable option for low-cost OCR or detection because it provides 25–30 FPS at 5–7 W. In contrast to Jetson devices, its task-specific acceleration restricts flexibility.

**B. Benchmark Metrics**

Performance is assessed using: - Throughput:  $T = \frac{N_{frames}}{P_{total}}$  - Latency:  $L = \frac{1}{f_{total}}$  - Energy Efficiency:  $E = \frac{N_{frames}}{P_{total}}$   
 Fair platform comparisons are made possible by these metrics [16].

**C. Discussion**

Real-time multimodal inference is made possible by sophisticated devices like Orin Nano, but they come with a price tag and a thermal load. Although lighter platforms (Coral TPU, Nano) are more affordable and energy-efficient, they can only be used for specific tasks.

Important trade-offs: Power versus accuracy: Orin uses more power but can handle larger models. Cost versus flexibility: Although less versatile, Nano and Coral are more affordable.

Portability versus performance: TX2 strikes a balance between wearable integration and throughput.

Hardware must therefore be matched to the requirements of the application: Orin for multimodal captioning, TX2 for navigation, and Nano/Coral for OCR.

**V. CHALLENGES**

The broad use of AI-driven assistive technologies is hampered by a number of obstacles, despite significant advancements. Trade-offs between accuracy and latency, dataset bias, multilingual constraints, energy efficiency, and reproducibility are important concerns. In order to guarantee reliable, inclusive, and useful real-world systems, these issues must be resolved.

**A. Accuracy vs. Latency**

Accuracy and real-time performance must be balanced in devices with limited resources. While lightweight models (MobileNet, Tiny-YOLO) are quicker but less accurate, high-accuracy models (such as transformers and large CNNs) have higher latency [2],[3]. While OCR can withstand delays for accuracy, navigation aids prioritize low latency, even with modest accuracy, for safety. This trade-off is best illustrated by wearable Jetson Nano prototypes [16],[18]. Pruning, quantization, and knowledge distillation are some solutions.

**B. Dataset Bias**

Generalization is limited because benchmarks frequently favor Western urban environments [15]. OCR trained on English scripts does not work well on Arabic, Hindi, or Chinese text, and navigation models trained on Cityscapes perform poorly in rural or non-Western environments. This discrepancy is demonstrated by commercial apps such as Google Lookout [10]. Open-vocabulary systems like ObjectFinder, diverse datasets, domain adaptation, and synthetic data are all examples of mitigation [19].

**C. Multilingual Support**

Text-to-speech and OCR systems need to be able to handle a variety of scripts and styles. Character Error Rate (CER) reveals gaps in inclusivity, with high error rates in non-Latin scripts [13],[14]. For example, OrCam MyEye [11] only supports Hebrew and English. For privacy-preserving adaptation, promising methods include transformer-based OCR, federated learning, and multilingual pretraining.

**D. Energy Efficiency**

Wearable technology is subject to stringent power budgets.  $E_{total} = P \cdot t_{inference}$  is the energy consumption, where  $P$  is the power draw. Optimizing battery life is essential for continuous use (such as navigation aids) [16]. Effective tactics include accelerators (Coral TPU), compression (quantization, pruning), and adaptive inference (e.g., lower frame rate when stationary).

TABLE IV

KEY CHALLENGES AND POSSIBLE SOLUTIONS IN ASSISTIVE AI

Challenge	Potential Solutions
Accuracy vs. Latency	Model pruning, quantization, knowledge distillation, and lightweight architectures (MobileNet, Tiny-

	YOLO)[2],[3],[16],[18]
Dataset Bias	Open-vocabulary recognition (ObjectFinder), geographically diversified datasets, synthetic data, and domain adaptation [10],[15],[19]
Multilingual Support	Transformer-based OCR, federated learning, multilingual support, and multilingual pretraining for privacy-preserving adaptation [13],[14]
Energy Efficiency	Task-specific accelerators (Coral TPU), hardware-aware compression, and adaptive inference techniques [16]
Reproducibility	Open-source implementations, transparent evaluation protocols, and standardized benchmarks [16],[18]

### E. Reproducibility

Validation is hampered by a lack of standardized benchmarks and open-source code [16]. For a fair comparison, common datasets and edge hardware benchmarks are necessary. Transparency is emphasized by initiatives like NVIDIA's Edge AI Benchmarking suite [16] and fairness projects like Gender Shades [15]. The significance of standardized evaluation is further highlighted by deployments such as NavWear [18].

### F. Discussion

In conclusion, the primary obstacles are:

- Accuracy vs. latency: refined through distillation, quantization, and pruning [2],[3],[16],[18].
- Dataset bias: mitigated by domain adaptation and a variety of datasets [10],[15],[19].
- Multilingual limits: addressed with multilingual pre-training and wider benchmarks [13],[14].
- Energy efficiency: enhanced through effective hardware and adaptive inference [16].
- Reproducibility: calls for open-source implementations and standardized protocols [16],[18].

OrCam MyEye [11], Google Lookout [10], and Seeing AI

[9] are examples of commercial systems that demonstrate the capabilities and constraints of existing approaches. Building scalable, equitable, and user-centered assistive AI requires overcoming these obstacles.

## VI. FUTURE DIRECTIONS

A number of exciting research directions are made possible by the quick development of assistive technology for the blind and VI. These approaches tackle the main issues of latency, dataset bias, multilingual support, and real-world usability found in existing systems, in addition to increasing user interaction.

### A. Voice-based Interaction

Accessibility can be greatly increased by using speech as the main modality. Multilingual support and real-time transcription are made possible by large-scale automatic speech recognition (ASR) models like Whisper.

This enables conversational interaction where users can ask questions about their environment and get context-aware responses when combined with vision-language frameworks. Early prototypes that combine speech and vision modules have shown how this approach improves usability, especially in environments with a variety of cultures and languages where visual-only systems are inadequate [9],[18].

### B. Depth Sensing and 3D Perception

By enabling precise 3D environmental mapping, depth sensing technologies—such as LiDAR, stereo cameras, and low-cost depth sensors—offers safer navigation beyond 2D vision. It has been demonstrated that combining depth information with monocular estimation enhances obstacle detection in dynamic and cluttered environments [4]. Real-world navigation prototypes like NavWear confirm the benefit of multimodal sensing, lowering errors compared to image-only methods [18].

### C. Bias-aware and Federated Learning

A significant drawback is still dataset bias, particularly in non-Western contexts. By training models directly on user devices without exchanging raw data, federated learning provides a privacy-preserving solution. In addition, bias-aware loss functions can reduce contextual and demographic differences, ensuring consistent performance across diverse populations [15]. These approaches, which guarantee inclusivity while protecting user privacy, have already been tried in healthcare AI and have great potential for accessibility applications. Systems like ObjectFinder also highlight how open-vocabulary recognition can expand inclusivity by moving beyond fixed datasets [16],[19].

### D. Cloud-Edge Hybrid Frameworks

The computational power of cloud servers and the low-latency advantages of edge computing are combined in hybrid architectures. While the cloud can be used for computationally demanding tasks, edge devices ensure responsiveness and offline availability. This equilibrium can be stated as follows:

$$T_{\text{hybrid}} = \alpha T_{\text{edge}} + (1 - \alpha) T_{\text{cloud}} \quad (1)$$

where the division of computation is controlled by  $\alpha \in [0, 1]$ . Future systems can further optimize  $\alpha$  dynamically to balance accuracy, efficiency, and energy use [12],[16]. Commercial applications like Seeing AI already employ a hybrid setup [9].

### E. Outlook

When combined, these patterns point to the multimodal, bias-aware, and hybrid design of the upcoming generation of assistive technology. Federated learning will guarantee inclusivity, depth sensing will allow for a richer environmental awareness, voice-based interaction will offer a natural communication channel, and hybrid frameworks will strike a balance between usability and performance. By moving in these directions, assistive AI will progress from prototypes to reliable, practical solutions that are both technologically sound and socially inclusive.

## VII. CONCLUSION

This survey collected studies on hardware platforms, system architectures, and models of AI-powered assistive technologies for the blind and visually impaired. Using the timeline of evolution (Fig. 1), taxonomy of system types (Fig. 2), and generic processing pipeline (Fig. 3), we charted the progression from simple sensing modules to contemporary multimodal, edge-enabled systems. Comparisons of hardware platforms (Table III) and commercial applications (Table I) also brought to light practical trade-offs between accuracy, latency, energy consumption, and usability.

Advances in deep learning, such as transformers, convolutional networks, and multimodal pretraining, have improved OCR [13],[14], object detection [2],[3], and image captioning [8]. Deployments on embedded platforms like Jetson Nano, TX2, and Orin Nano demonstrate the growing feasibility of **edge AI** for low-latency, privacy-preserving assistive systems [12],[16]. However, enduring problems persist, including limited reproducibility [16], stringent wearable energy budgets [12], multilingual constraints [13],[14], and dataset bias [15].

All things considered, AI-driven accessibility is developing quickly but is still in its infancy. The community can hasten the development of scalable, dependable, and socially equitable assistive technologies by tackling open challenges and embracing inclusive, user-centered design. As a survey paper, this work provides researchers and practitioners creating the next generation of visually impaired systems with a reference and a roadmap.

## REFERENCES

- [1] WorldHealthOrganization, "Worldreportonvision," WorldHealthOrganization, Geneva, Switzerland, 2019. [Online]. Available: <https://www.who.int/publications/i/item/9789241516570>
- [2] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 779–788, 2016. [Online]. Available: <https://ieeexplore.ieee.org/document/7780460>
- [3] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. Reed, "SSD: Single shot multibox detector," in Proc. European Conf. Computer Vision (ECCV), pp. 21–37, 2016. [Online]. Available: <https://arxiv.org/abs/1512.02325>
- [4] K. Lee and M. Tanaka, "Blind navigation assistance for visually impaired based on local depth hypothesis from a single image," Sensors, vol. 21, no. 4, pp. 1453–1465, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1877705813016214>
- [5] A. Tapu, B. Mocanu, and T. Zaharia, "Wearable assistive devices for visually impaired: A survey," Sensors, vol. 20, no. 11, pp. 1–28, 2020. [Online]. Available: <https://www.mdpi.com/1424-8220/20/11/3202>
- [6] L. Chen, H. Wu, and J. Zhao, "An effective obstacle detection system using deep learning advantages to aid blind and visually impaired navigation," IEEE Access, vol. 10, pp. 44210–44222, 2022. [Online]. Available: <https://ieeexplore.ieee.org/document/7514991>
- [7] A. Alghamdi, A. A. Mutlag, and R. AbdAlhameed, "A deep learning based assistive system for the visually impaired using YOLO object detection," IEEE Access, vol. 9, pp. 114–124, 2021. [Online]. Available: <https://scholar.google.com/scholar?hl=en&assdt=0%2C5&q=A.+Alghamdi%2C+A.+A.+Mutlag%2C+and+R.+AbdAlhameed%2C+%E2%80%9CA+deep+learning+based+assistive+system+for+the+visually+impaired+using+YOLO+object+detection%2C%E2%80%9D+%5Cemph%7BIEEE+Access%7D%2C+vol.+9%2C+pp.+114%E2%80%93124%2C+2021.+%5BOnline%5D.&btnG=>
- [9] Y. Huang and C. Li, "An image captioning model based on bidirectional depth residuals and its application," Multimedia Tools and Applications, vol. 81, pp. 23245–23260, 2022. [Online]. Available: <https://dl.acm.org/doi/abs/10.1007/s10462-023-10488-2>
- [10] Microsoft, "Seeing AI," Microsoft, 2023. [Online]. Available: <https://www.microsoft.com/en-us/ai/seeing-ai>
- [11] Google, "Google Lookout," Google, 2023. [Online]. Available: <https://play.google.com/store/apps/details?id=com.google.android.apps.accessibility.reveal>
- [12] OrCam, "OrCam MyEye," OrCam Technologies, 2023. [Online]. Available: <https://www.orcam.com/en/myeye/>
- [13] NVIDIA Corporation, "NVIDIA Jetson AGX Orin Series Product Brief," NVIDIA, 2022. [Online]. Available: <https://developer.nvidia.com/embedded/jetson-modules>
- [14] R. Agyeman and B. Rinner, "Decentralized text-based person re-identification in multi-camera networks," IEEE Access, vol. 12, pp. 172125–172148, 2024. [Online]. Available: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10756694>
- [15] B. Shi, X. Bai, and C. Yao, "An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 11, pp. 2298–2304, 2017. [Online]. Available: <https://ieeexplore.ieee.org/document/7801919>
- [16] T. Buolamwini and G. Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification," in Proc. Conf. Fairness, Accountability, and Transparency (FAT), pp. 77–91, 2018. [Online]. Available: <https://proceedings.mlr.press/v81/buolamwini18a.html>
- [17] D. Minott, S. Siddiqui, and R. J. Haddad, "Benchmarking Edge AI Platforms: Performance Analysis of NVIDIA Jetson and Raspberry Pi 5 with Coral TPU," in Proc. IEEE SoutheastCon, pp. 1384–1389, 2025. [Online]. Available: <https://ieeexplore.ieee.org/document/10971592>
- [18] K. Vashistha, P. Mandal, and A. Sinha, "Comparative usability study of OrCam MyEye and Microsoft Seeing AI for visually impaired users," Disability and Rehabilitation: A Assistive Technology, vol. 18, no. 1, pp. 1–12, 2023. [Online]. Available: <https://www.researchgate.net/publication/353519592EvaluationandComparisonofArtificialIntelligenceVisionAidsOrcamMyEyeIandSeeingAI>
- [19] M. Shilkrot, R. Gonzalez-Franco, and P. Maes, "NavWear: A wear-able navigation aid for the visually impaired with real-world user evaluation," in Proc. ACM CHI Conf. Human Factors in Computing Systems, pp. 1–12, 2021. [Online]. Available: <https://dl.acm.org/doi/10.1145/3411764.3445342>
- [20] H. Li, Y. Wang, and Z. Luo, "ObjectFinder: Open-vocabulary assistive object recognition for the blind," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 11562–11571, 2022. [Online]. Available: <https://ui.adsabs.harvard.edu/abs/2024arXiv241203118L/abstract>



10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)