International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Survey on Real-world Applications and Challenges of Deep Learning-Enhanced Techniques to Assist Visually Impaired

Rashmi B N[1], Dr,R Guru[2], Dr.Anasuya M A[3]

Submitted: 25/01/2024 **Revised**: 03/03/2024 **Accepted**: 11/03/2024

Abstract: This comprehensive survey report delves deeply into the real-world applications and complicated issues inherent in deep learning-enhanced wearable solutions for people with vision impairments. It stresses the global incidence of visual impairment, particularly in underserved areas, and follows the growth of assistive devices over time. The study examines deep learning's revolutionary function, demonstrating its impact through real-world case studies such as OrCam MyEye and Brain-Computer Interfaces. It does, however, rigorously identify the various technical challenges, such as data accessibility and real-time processing, as well as ethical concerns, such as privacy and fairness. In conclusion, while the paper highlights the potential of deep learning to empower people with disabilities, it also calls for the continual resolution of these obstacles to construct a more inclusive and accessible future. We need to focus on designing small size object detection and object recognition systems which consider varying size images, to address the problems faced by the visually impaired in their passive and active stages, according to the study.

Keywords: *Object Detection, Visually impaired, Accuracy Analysis, Assisting aid, Small object detection*

1. Introduction

According to data provided by the World Health Organization (WHO) in October 2017, visual impairment is one of the most important global health challenges, affecting a whopping 253 million people globally. [32] This population is divided into 36 million people who are completely blind and 217 million people with limited vision. The demographic distribution of this issue is particularly concerning; a disproportionate number of these people live in underdeveloped nations, where resources and assistive technologies are scarce. [30] The importance of vision in human life cannot be emphasized. Many daily activities rely on it, including navigation, obstacle avoidance, and social engagement [64]. The absence of this crucial sensory input complicates spatial orientation, navigation, and emotional well-being. The visually handicapped are frequently lost or intimidated, especially in strange surroundings, which has a negative influence on their personal, professional, and environmental interactions. For years, the visually impaired relied on conventional mobility and navigation aids such as white canes and guide dogs. [21] While these devices provide some assistance, their powers are severely limited. A white cane, for example, detects nearby obstructions but is ineffective for broader environmental awareness. Despite their navigation ability, guide dogs are costly to train and keep. Advancements in sensor technology, machine learning, and computational algorithms have paved the way for electronic traveling aids (ETAs) [15] and intelligent wearables. Such

[1] *JSSATEB,Bengaluru*
ORCID ID : /0000-0002-7416-4230
[2] *SJCE,JSS STU,Mysore*
ORCID ID : 0000-000-6108-1459
[3] *SJCE,JSS STU,Mysore*
ORCID ID : 0000-0002-5236-8400
** Corresponding Author Email:rashmibn@jssateb.ac.in*

devices have a variety of functions, including obstacle avoidance, route selection, and social engagement. Despite their technological complexity, however, these gadgets require assistance in user acceptance, especially due to pricing, usability, and data privacy concerns. Intelligent wearables are becoming increasingly common because of recent technical breakthroughs. According to IDC's 2019 estimate, the global shipment of smart wearable devices will reach 302.3 million units by 2024. These devices can potentially transform healthcare, lifestyle, and personal safety. Despite this promise, adoption rates may have been higher, partially due to usability concerns and partly due to the community's lack of understanding of the potential benefits and limitations of these devices. Integrating Deep Learning algorithms and Internet of Things (IoT) [55.] technologies represent an intriguing area for improving visually impaired wearable devices. These cutting-edge technologies provide an intelligent, adaptive interface to help with difficult activities like object recognition, spatial mapping, and real-time decision-making. However, real-world deployment and widespread acceptance are riddled with difficulties, including technical limits, ethical problems, and socioeconomic factors.

1.1. Deep learning techniques used in assistive devices for individuals with visual impairments:

Deep learning, an artificial intelligence subset, has emerged as a game changer in wearable assistance. Deep learning algorithms, applied in neural networks, enable wearable devices to absorb and interpret sensory information in previously unthinkable ways, like how the human brain learns and adapts. [7] Consider these algorithms the brain of the wearable device, allowing it to "see" and understand the world on the user's behalf. This technical breakthrough has ushered in a new era of assistive gadgets that have become more intelligent and reactive than ever. These advanced machine learning methods, particularly Convolution Neural Networks (CNNs) and Long Short-Term Memory Networks (LSTMs), serve as the computational backbone for systems that provide a richer, more engaging user experience. CNNs are most used for image recognition tasks; they excel at extracting complex features from visual data, making them perfect for recognizing objects, text, and even facial emotions in the visual field. These attributes are then used to deliver real-time input to the user, usually audio cues or haptic feedback. On the other hand, LSTMs, a form of Recurrent Neural Networks, are used for sequence prediction tasks such as natural language processing. They could turn identified text or features into coherent, contextual statements. This is very beneficial for summarizing the surroundings or converting visual data into understandable auditory descriptions. An LSTM, for example, can generate a descriptive caption for an image detected by a CNN, which is then translated to speech using Text-To-Speech APIs, allowing visually impaired people to "hear" the image. Integrating deep learning techniques into assistive devices is about more than just improving the functionality of classic equipment like canes and Braille readers. Nonetheless, technology is paving the way for greater autonomous navigation and better interaction with the environment for visually impaired people. [12] The real-world applications and problems of deep learning-enhanced wearable assistance for the visually handicapped are examined in this review study. Its goal is to thoroughly assess cutting-edge technologies, study their usefulness and limitations, examine the ethical implications, and make practical recommendations for future research and development. This review intends to add substantively to the ongoing discourse on this subject by bridging the gap between technological progress and its pragmatic use for improving the lives of visually impaired individuals. Deep learning, an artificial intelligence subset, has emerged as a game changer in wearable assistance. Deep learning algorithms, applied in neural networks, enable wearable devices to absorb and interpret sensory information in previously unthinkable ways, like how the human brain learns and adapts. [7] Consider these algorithms the brain of the wearable device, allowing it to "see" and understand the world on the user's behalf. This technical breakthrough has ushered in a new era of assistive gadgets that have become more intelligent and reactive than ever. These advanced machine learning methods, particularly Convolution Neural Networks (CNNs) and Long Short-Term Memory Networks (LSTMs), serve as the computational backbone for systems that provide a richer, more engaging user experience. CNNs are most used for image recognition tasks; they excel at extracting complex features from visual data, making them perfect for recognizing objects, text, and even facial emotions in the visual field. These attributes are then used to deliver real-time input to the user, usually audio cues or haptic feedback. On the other hand, LSTMs, a form of Recurrent Neural Networks, are used for sequence prediction tasks such as natural language processing. They could turn identified text or features into coherent, contextual statements. This is very beneficial for summarizing the surroundings or converting visual data into understandable auditory descriptions. An LSTM, for example, can generate a descriptive caption for an image detected by a CNN, which is then translated to speech using Text-To-Speech APIs, allowing visually impaired people to "hear" the image. Integrating deep learning techniques into assistive devices is about more than just improving the functionality of classic equipment like canes and Braille readers. Nonetheless, technology is paving the way for greater autonomous navigation and better interaction with the environment for visually impaired people. [12] The real-world applications and problems of deep learning-enhanced wearable assistance for the visually handicapped are examined in this review study. Its goal is to thoroughly assess cutting-edge technologies, study their usefulness and limitations, examine the ethical implications, and make practical recommendations for future research and development. This review intends to add substantively to the ongoing discourse on this subject by bridging the gap between technological progress and its pragmatic use for improving the lives of visually impaired individuals.

2. Evolution of assistive technologies for the visually impaired

It investigates the historical progression and technological advancement of assistive devices for visually impaired people. We go on a complete trip, charting the evolution of sensory substitution devices like ultrasonic and infrared systems to the modern landscape dominated by sophisticated computer vision-based solutions. This investigation emphasizes the significant achievements made in this domain while also emphasizing the persisting problems that have fueled innovation. The article demonstrates the enormous influence of technology on improving the autonomy and quality of life for persons with visual impairments while shedding light on the rising potential of assistive technology achievements in this area.

2.1 Historical Perspective: Wearable assistive devices for the blind have a fascinating history that illustrates how technology has advanced and how society's attitudes toward disability have changed. This section explores the tangled web of achievements in creating wearable aids and their enormous influence on the lives of people with visual impairments. i) The Early Days of Wearable Aids: Wearable assistive devices for the blind date back to the early 20th century, when innovators in education and assistive technology realized the importance of enhancing sensory perception. The creation of the white cane by James Biggs in the 1920s is one of the most well-known innovations from this time. This straightforward but clever device offered tactile input, enabling users to recognize impediments and move with confidence, and it signaled the beginning of wearable aids as we know them today.ii) The Emergence of Electronic Aids: With the development of electronic technology in the second part of the 20th century, the landscape of wearable aids experienced a fundamental change. Dr. Raymond Kurzweil developed the "Optacon," a ground-breaking gadget with a tiny camera that could record printed text, in the 1970s. The Optacon transformed visual information into tactile input, significantly affecting how people with visual impairments can access printed materials.

2.2 Milestones in Technology Adoption for the Visually Impaired: The visually impaired population has had significant developments in the adoption of technology over the years, including:

i) The Braille Display (1980s): Braille displays, small devices that gave real-time access to digital content in Braille, were developed in the 1980s. This breakthrough was critical in improving blind people's access to education and information, boosting independence and literacy.

ii) Voice Synthesis and Screen Readers (1990s): Screen reader software and voice synthesis technologies proliferated in the 1990s, making personal computers and the internet more accessible to blind people. These developments were a big step toward digital inclusion, allowing individuals to access and navigate the digital world.

iii) Smartphone’s and Mobile Apps (early 2000s): With the widespread use of smart phones with accessibility features in the early 2000s, a new age began. Concurrently, developers began developing many mobile apps customized specifically to the needs of the visually impaired. These programs varied from navigation tools to text recognition software, giving users unprecedented abilities.

iv) Deep Learning and Computer Vision (recent years): Deep learning and computer vision technology have ushered in a new era of wearable aids. Thanks to improved algorithms and cameras, modern devices can recognize objects, read text, and provide real-time audible feedback. Because of this technological advancement, individuals with visual impairments now have greater independence, autonomy, and quality of life. These landmarks highlight not only technology progress but also the increasing knowledge of disability and the critical role of accessibility and inclusivity. Wearable aids are no longer just tools; they are transforming tools that enable those with visual impairments to achieve independence, education, work, and social participation. We will explore deeper into the uses and problems of deep learning-enhanced wearable aids in the following sections of this essay, exposing their potential to transform the lives of the visually handicapped in our modern era.

V) Bionic Eye: Visual perception in humans and/or animals has been successfully attained by several kinds of bionic eyes. Nevertheless before a prosthetic eye can entirely and safely restore vision functions, a lot of problems need to be resolved and technical challenges need to be addressed. Future bionic eyes could produce greater medical outcomes if fresh techniques are used. Eleven major bionic eye preclinical and clinical experiments have been assessed as a result regarding their technological solutions and technical specifications used by bionic eye research worldwide.

2.3 Sensorial Networks ETAs: To aid visually impaired (VI) users in navigation and object recognition, the sensorial networks addressed in the offered text employ various technologies such as ultrasound, infrared, sonar-based systems, RFID, GPS, and others are shown in Table 1. These systems provide various functions, including estimating distances, detecting obstructions, providing geographic information, and assisting with interior or outdoor navigation. Each method, however, has disadvantages, such as the requirement for significant training, low resolution, and sensitivity to ambient variables, and accuracy concerns in 3D space. Furthermore, some systems necessitate prior knowledge of building layouts, whereas others are deemed invasive due to the requirement to install RFID tags. While these sensory networks are important in supporting VI users, none of them can properly detect, identify, and recognize specific items in an unknown area or estimate their relative hazard, and they are susceptible to interference and environmental influences.

Table 1. Sensorial Network Example

Sensorial Network	Technolog y Used	Purpose	Limitations
Mowat Sensors [13]	Ultrasound	Calculate the distance between the VI user and the obstructions.	Extensive training is required, it has a limited resolution, and it is less precise in 3D space.
Sonicguide and Trisensor [70]	Sonar-based	Give spatial information about potential obstructions.	Sensors with low resolution are less precise in 3D space.
Talking Braille [43]	Infrared	Aid VI users in confined spaces.	Restriction to identified buildings, vulnerability to sunshine
Binaural Sonar [8]	Sonar-based	Identify things on the left and right sides.	Input is delivered in the form of vibrotactile simulations.
GuideCane, SmartCane, UltraCane, Batcane, Necklace Cane [33]	Ultrasonic sensors	Enhance the possibilities of the white cane	Recognizing above-knee, hanging impediments , or sidewalk borders has some drawbacks.
EPFL Multi-Sonar Architecture [67]	Multi-sonar	Encourage indoor displacement	High rate of false positives under actual circumstance s
CyARM [63]	Ultrasonic transducers	Spatial localization of the VI users	While it is accurate for stationary objects, it is less so for moving ones.

Microcontroller Ultrasonic Sensor [29]	Ultrasonic sensor	Identify anything in the path of the walking stick.	Not tried in open-air, busy environments.
Smart-Robot [60]	RFID and GPS	Help VI users navigate both inside and outside	Accuracy issues and signal loss in urban settings
RFID Systems [68]	RFID technology	Enabling movement through indoor areas	Requires prior architecture of buildings knowledge and is seen as intrusive
BLE Beacon [42]	BLE beacon technology	Temporary environmental application	Performance is influenced by where BLE beacon devices are placed.
ShopTalk [62]	Helmets, barcode readers, keypad	Help VI users with their shopping	Inventory is hard to access and carry.
SUGAR System [54]	UWB technology	Provide VI users with indoor guidance	Limited to indoor conditions, no mention of power usage

Networks of Sensors users, especially those with visual impairments, can use ETAs to predict arrival times using several sensors and data sources. These systems are essential for improving independence and mobility.

Table 2. Defines the Types of Sensorial Networks ETAs.

Types of Sensorial Networks ETAs	**Applications**	**Benefits**
GPS-based ETAs	Turn-by-turn directions are provided via navigation apps.	Incredibly precise position tracking for on-time and safe travel.
IoT Sensors	Keeping an eye on environmental factors, such as temperature.	Real-time environmental information to aid in making wise travel choices.
Crowd sourced ETAs	Information about road conditions and current traffic conditions.	User-generated information for navigating barriers and traffic.

Table 3. Sensorial Network Example

System	Description	Strengths	Limitations
Monocular Camera-Based Systems[4]	Numerous systems use monocular cameras for object detection and indoor and outdoor navigation.	Portable and affordable	For arbitrary paths or situations, requiring prior environmental information is inappropriate.
stereo Camera Based Systems[75]	Systems that use stereo cameras to detect obstacles and guide users through both indoor and outdoor areas.	Give context to help viewers better understand the scene	It may not perform effectively in brightly illuminated outdoor conditions since it is sensitive to matching errors.
RGB-D Camera Based Systems[5]	For interior and outdoor navigation, object identification, and object recognition, assistive devices combining RGB-D sensors.	Can recognize depth and RGB data for greater comprehension	The disparity map's quality and sensitivity to shifting illumination conditions are limitations.

2.4 Computer Vision-based ETAs:

Based on their camera sensors, the three primary types of electronic travel assistance (ETA) systems are monocular, stereo, and RGB-D is shown in Table 3. Arianna and Mobile Vision are two examples of monocular camera-based ETAs that are affordable and portable but require prior environmental knowledge and are less effective for arbitrary courses. Smart Walker and Robot Vision are two examples of stereo camera-based systems that offer depth information for greater picture comprehension but are vulnerable to matching errors, particularly in bright outdoor settings. Depth sensing is integrated into RGB-D camera-based systems like SoV and NAVIG for indoor and outdoor navigation and object recognition, although these systems can have limitations due to the quality of the disparity map and shifting illumination conditions. These devices provide a range of features and trade-offs for helping visually impaired people move around and perceive the surrounding environment. Visually impaired individuals can receive context-aware guidance using computer vision-based ETAs, which rely on object tracking and visual recognition to estimate arrival timings. It becomes clear from looking back on the development of assistive technologies for the blind that innovation has been a key factor in fostering independence and inclusivity. These technologies have revolutionized how visually impaired people see the world and opened doors to new opportunities and experiences. They have progressed from the simple white cane to the sophisticated Sensorial Networks ETAs and Computer Vision-based ETAs of today. With improvements in sensor technology, machine learning, and artificial intelligence, the future is even more promising. These developments will keep removing obstacles and ensuring that people with visual impairments may independently and confidently navigate the world, achieving their full potential in a constantly changing digital environment. It's important to remember that technology is not just about gadgets as we grow it's also about empowering people, promoting inclusivity, and building a society where everyone can succeed, regardless of their talents.

3. Deep learning revolutionizing assistive technologies

Neural networks: Networks of neurons Numerous assistive technologies are built around neural networks, which were inspired by the human brain. Neural networks enable machines to comprehend and react to spoken or written words in applications like speech recognition and natural language processing (NLP), which benefits people with communication difficulties.[10].

Fig 1. Deep Learning Transforms Assistive Technologies

Convolution Neural Networks (CNNs): CNNs are excellent at processing images and videos. CNNs [22] are employed in assistive technologies to do tasks including facial recognition, scene description, and item detection. CNNs can use image analysis to describe the environment for people who are visually impaired, helping them better grasp it.

Recurrent neural networks (RNNs): RNNs [28] excel at handling sequential input and are, therefore, crucial for speech recognition, text prediction, and gesture recognition. RNNs are used in assistive technologies to translate sign language into text or speech, allowing deaf people to communicate.

Natural Language Processing (NLP) has revolutionized thanks to deep learning models, particularly Transformers. For tasks like text-to-speech and voice-to-text conversion, assistive technologies use NLP[29], enabling more accessible communication for those with speech or hearing problems.

Gesture Recognition: Deep learning algorithms are capable of recognizing and deciphering gestures, which is useful for assistive technology for those with mobility issues. With the aid of these technologies, users can interact with their surroundings, operate gadgets, and communicate with one another.[17]Personalization: Deep learning models can be made to fit the specific demands and preferences of each user. This personalization helps assistive technology by giving impaired users specialized assistance and accommodations.

Real-time Processing: Deep learning hardware and algorithm advancements have made it possible for assistive devices to process information in real time. By improving responsiveness, this feature makes assistive technology more capable of offering consumers quick assistance.[44]Enhanced Accessibility: Deep learning has made it possible to design user interfaces that are more accessible, making computers, smart phones, and other devices easier for people with impairments to use. This includes tools like screen readers, voice commands, and predictive text.

Autonomous Navigation: The development of autonomous navigation systems for the blind relies heavily on deep learning-based algorithms [18]. To assist users in securely navigating across varied settings, these systems combine sensors and deep learning.

Augmented Communication: Increased and Alternative Communication (AAC)[23] systems have been improved by deep learning, making it simpler for people with speech difficulties to communicate via text-to-speech and symbol-based interfaces. Finally, by offering solutions for speech recognition, picture analysis, natural language understanding, gesture identification, and personalization, deep learning principles have altered assistive technologies. For people with disabilities, these technologies have changed accessibility, autonomy, and communication, greatly enhancing their quality of life and independence. On the other hand, one-stage detectors contain a single feed-forward fully convolution network that directly provides the bounding boxes and the object classification.

3.1 our work: comparative study

Our Deep learning techniques focused study on Object detection in the real world highlights the limitations and challenges which affect good accuracy achievements in each technique and comparative analysis of the accuracy. Two-stage frameworks divide the detection process into the region proposal and the classification stage. These models first propose several object candidates, known as regions of interest (RoI), using reference boxes (anchors).

One-stage detectors: Object classification and bounding-box regression are done directly without using pre-generated region proposals (candidate object bounding-boxes).

Transform based object detectors: The astounding performance of transformers in natural language processing (NLP) has motivated researchers to explore their applications in computer vision tasks[78]. The transformer architecture has been shown to be effective in capturing long-range dependencies in sequential data, making it an attractive candidate for object detection tasks. In 2020, Carion et al. proposed a novel object detection framework called DEtection TRansformer (DETR) [7], which replaces the traditional region proposal-based methods with a fully end-to-end trainable architecture that uses a transformer encoder-decoder network. The DETR network shows promising results, outperforming conventional CNN-based object detectors [79]

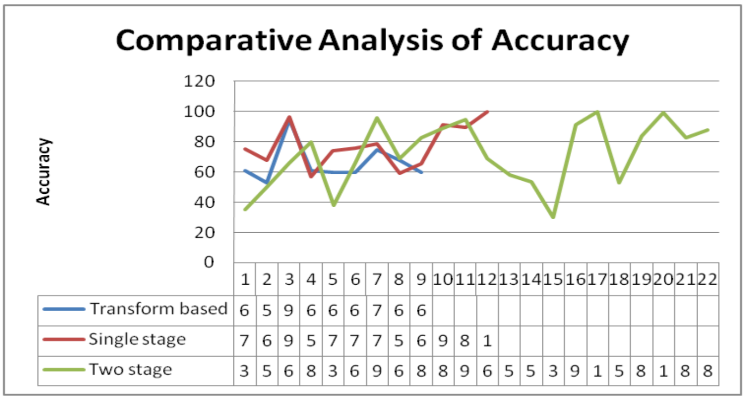

Fig 2. Comparative analysis of Accuracy stage detectors

From the above analysis transform based techniques achieves lesser accuracy than that of two stage and single stage detectors.

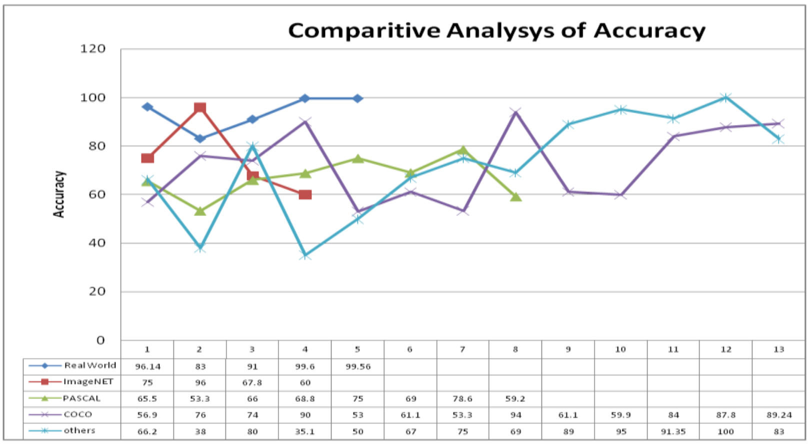

Fig 3. Comparative analysis of Accuracy Based on Data

From the above analysis its shows that the accuracy drawn on different datasets based on comparative study, where coco data sets shows linear to different techniques. This in turn we can conclude that coco data set includes the samples which test the different case study.

Table 4. Deep learning Techniques to detect the objects with accuracy

Deep learning Technique	Detector	Key Features	Data Base	Accuracy	Challenges
[14] YOLOv5	single	More parameters, included the uses of dynamic architecture	Image data set	75%	large in size
[48] SPP-NET	single	flexible solution for handling different scales, sizes, and aspect ratios, feature maps from the entire image only once, faster than R-CNN method,	Pascal VOC 2007.	59.2% mAP	deep-networks-based recognition
[19] YOLO	single	Detect real world objects and good accuracy	300 images	96.14%	Struggle to detect small objects
[72] YOLO	single	Detect real world objects and good accuracy	COCO	56.9%	Struggle to detect small objects
[46] YOLOv2	single	input size increased	VOC 2007	78.6%	increased number of layer
[11] SSDLite-MobileNet.net,	single	depthwise separable convolutions	live videos and COCO	76%	reduced parameters
[25] YOLOv1	single	spatial pyramid pooling layer is used, inception model structure is added	Pascal VOC 2007	65.5	unable to recognize small objects which are in group m
[40]Xception	single	excellent results with non-residual VGG-style models	FastEval14k	6.78mAP	depthwise separable convolution layers
YOLOv4	single	s single-frame-based and multi-frame-based recognition	MS COCO	74%	optimal hyperparameters are tuned to achieve the best performance
AlexNet 36 246 116-CNN	two	Models trained using two components	MSRA-B,HKU-IS	35.1% mAP	multiple levels of image segmentation
ZFNet	two	gives insight into the function of intermediate feature layers and the operation of the classifier	Caltech-256	50%mAP	less well to the PASCAL data
R-CNN O-Net BB	two	Faster, less parameters	PASCAL VOC 2012	66.0% mAP	training time more
[52]ConvNet VGGNet	two	emultinomial logistic regression	e ILSVRC-2012	8.0%	fixed-size 224 × 224 RGB image,More parameters
[11]ResNet	two	size is smaller compared to VGG	GoogLeNet	38%mAP	difficulty to detect the small object
[16]DetNet	two	convolutional decomposition of images	Caltech-101	66.2%mAP	unsupervised
[81]SqueezeNet	two	compress SqueezeNet to less than 0.5MB, Faster than Alexanet	PASCAL VOC 2012	68.8mAP	amenable to on-chip implementations
[50]SPP-net	two	Pyramid pooling,robust to object deformations. SPP is a flexible solution for handling different scales, sizes, and aspect ratios	PASCAL,Caltech101	68.8%	difficult to fine-tune the parameters of the network before the layer
[36]CNN	two	reduce number of the parameter	WORLDVIEW	89%	fixed-size input image
[35]CNN	two	reduce number of the parameter	PAVIA	95%	fixed-size input image
[49]CNN	two	reduce number of the parameter	VOC 2012	53.3%.mAP	fixed-size input image
[47]RCNN	two	-	PASCAL VOC	30% mAP	fine-tuning does not emphasize precise localization
[34]CNN VGG-16Fc6,	two	Deep Multi-Layer (DM-L) based Feature Extraction	Caltech-101	91.35%	multiple layers , selecting the best layer
[1]AlexNet	two	Computation complexity is low,less data and using a pre-trained network	Merch Data	100%	application specific, with a small amount of data can detect objects in the image
[77]Mask R-CNN	two	simple to train	COCO 2016	53%	instance-level recognition
R-CNN	two	Detect real world objects and good accuracy	VOC 2012 and ImageNet dataset	96.14%	Hybrid model
CNN	two	accurate in object classification,	1000 images	83%	struggled with locating objects within the image.
R-CNN	two	faster than sliding-window detector	VOC 2012	53.3%.mAP	domain-specific fine-tuning
MobileNet	Transform based	single topographical saliency map	ed soda(imges)108	2.5 ± 0.05 SFC	rapid selection
[38]MobileNetV2	Transform based	large scale geo-localization	2016 COCO	75%	depthwise separable convolutions
[39]ShuffleNet	Transform based	significantly outperforms on a larger image classification dataset	ImageNet Data set	6.78mAP	large number of towers
[45]ShuffleNetV2	Transform based	learning residual functions, residual networks are easier to optimize, and can gain accuracy from considerably increased depth	ImageNet, COCO	60%,95%	Restnet are faster than ShuffleNetV2
[37]DenseNet-BC	Transform based	good feature extractors for various computer vision tasks ,alleviate the vanishing-gradient problem, strengthen feature propagation	ImageNet	58.3%	hyperparameter not considered

Table 5: Transform based Techniques to detect the objects with accuracy

Transformed models	Data set	advantages	Limitations	accuracy
DETR	COCO 2017	DETR achieves competitive results compared to Faster R-CNN in quantitative evaluation on COCO	optimization and performances on small objects. to optimize the network as object queries concentrate on something other than specific regions	61.1mAP
SOF-DETR	MS COCO	better performance than DETR	well suited for small-sized objects	53.3%mAP
Anchor DETR	MSCOCO	proposing object queries as anchor points	predictions are near the anchor point	4.2 AP ,16 FPS
DESTR	MS-COCO	mini-detector is used to learn and initialize both content and positional embeddings of the decoder	mini-detector breaks the well-streamlined architecture of DETR	61.1% AP

Table 6. Assistive Technologies for the Visually Impaired with single and two stages object detectors

Technology	Deep learning Model with accuracy	Key features	Accuracy	Data Set	Functionality provided	Single Detector used	Text to speech converter
tensor-flow API -2022	CNN Mobile-Net	light weighted and provides good accuracy that makes it the best choice for computationally low devices/DSPs	91%	1000 real world images	outdoor and natural environment	single	SAPI
Azure Vision API-2021	SoftMax	better than as SSD 300, Faster-RCNN 300, and Faster-RCNN 600	84%mAP	MS COCO	OUTdoor assistance	Two	VAPI
DSP processor,processor, NVIDIA Tesla K80 GPU, having,OCR-2020	YOLOv3	faster than AlexNet VGG-16 [VGG-19 [40] YOLO-v3	99.69%	real world images with currency notes	for object detection and recognition,e	single	API
GPS,OCR,AT Based on Raspberry Pi-2022	CNN-LSTM	oncerns over predicting sequencesincluding spatial inputs such as photography or visual content	83%.	Flicker8k dataset	comprehend text and images,	two	Google Text-To-Speech API
Android app, Intel Xeon processor with 64 GB RAM and an NVIDIA GeForce GTX 1080 Ti GPU-2023	DeepNAVI	single-camerabased method	87.8%	COCO	obstacle detection, scene recognition, distance estimation, motion detection, and position estimation,	two	(TTS) library, Pyttsx5
ultrasonic sensors, a camera, breadboards, jumper wires, a buzzer-2022	CNN	utilized CNN architecture AlexNet yielded an impressive result	99.56%	real world images	detection of obstacles and potholes	single	API
quad-core Cortex A53 processor, bone conduction headphones, and a 1.2 GHz 64-bit quad-core Cortex A53-2022	YOLO	tional path which is augmented by a GPS-based smart Stick. Over 1	89.24%	COCO	Indoor and outdoor assist in detection and recognition	two	API

4. Real-world applications

These real-world case studies show how deep learning is used in wearable technology. Functionality of OrCam MyEye: OrCam MyEye[64] is a wearable gadget to help people with visual impairments. It comprises a camera that can recognize and understand visual data mounted on spectacles. Applying deep learning, OrCam MyEye can recognize and read a variety of text forms, including books, signs, labels, and handwritten notes, thanks to deep learning algorithms, in particular, object identification and text-to-speech conversion. Users can point to a word or item, and the gadget will read it out loud, giving them more independence and access to printed information. Users have noted a greater sense of independence and self-assurance when performing regular duties like reading, shopping, and navigating crowded places.

Functionality of SignAll: SignAll is a wearable device that makes it easier for those who use American Sign Language (ASL)[59] and those who do not communicate with one another. ASL is converted into spoken language. SignAll employs deep learning and computer vision to recognize ASL signals using a camera-based sensor. Deep learning algorithms decipher the signs and generate output corresponding to spoken language. ASL interpreters are not required for real-time communication between hearing and deaf people. This tool improves inclusivity and eliminates communication barriers, resulting in more inclusive interpersonal relationships and employment prospects. Functionality of Brain-Computer Interfaces (BCIs): BCIs, are wearable devices that let people with severe motor disabilities operate computers and other assistive technology by sending and receiving brain signals. Deep learning algorithms are used to decode and interpret the intricate brain signals captured by BCIs, giving users the ability to type, move a cursor, or operate a wheelchair with their thoughts. BCIs have allowed people with illnesses like ALS to reclaim their freedom and interact with the online community. Using brain impulses, these people can interact with others, access the internet, and manage their surroundings.[69]

EMG Sensors Used to Control Prosthetics: Electromyography (EMG) sensors integrated into wearable prosthetic limbs make it possible for users to operate their limbs intuitively. Deep learning models examine EMG[27] signals that are produced when muscles in the residual limb contract. The user's goals are translated into movements for the prosthesis by these models as they learn to understand them. It is easier and more accurate for users of these prosthetic limbs to execute tasks like gripping objects, walking, and even playing musical instruments because they have smoother, more natural control over their artificial limbs.

Hearing aids with brain enhancements: Hearing aids with deep learning capabilities can change their settings according to the wearer's surroundings and listening habits. These hearing aids utilize deep learning algorithms to assess background noise and instantly adjust their settings. They can offer a customized listening experience, lessen background noise, and improve speech clarity. Users of these hearing aids claim enhanced hearing in noisy restaurants and crowded settings, among other circumstances. Their total auditory experience is enhanced by significant learning-based adaptability.[24]

5. User-centric design

To create assistive technology for the visually impaired that truly suits their needs, User-Centered Design (UCD) is essential to produce wearable assistive devices for the blind and visually impaired that are practical and helpful in their daily lives. The value of UCD rests in its capacity to integrate end-user perspectives and experiences right from the start of the design process. By integrating people with visual impairments, designers can acquire important insights into their requirements, difficulties, and preferences. Human-Computer Interaction (HCI) and usability engineering research underpin UCD. It is consistent with accepted guidelines like ISO 9241[6], which highlights the significance of user-centeredness in system design [66]. Designers can make sure that the technology addresses real-world issues by giving user feedback priority and including them in usability testing. For instance, "Usability 101: Introduction to Usability" by Nielsen Norman Group emphasizes how UCD approaches result in products that are more effective, efficient, and pleasing to consumers [73].

Making User-Friendly and Intuitive Aids: Making user-friendly and intuitive wearable aids is crucial for guaranteeing that people with visual impairments can quickly accept and benefit from the technology. This strategy's primary goals are accessibility and simplicity.

Simple Navigation: Interfaces should be made to make navigation simple. Users ought to be able to access vital functions and data without needless complications. [41]

Tactile and Auditory Cues: People with visual impairments can benefit greatly from tactile and auditory cues. The usefulness of auditory feedback in information transmission is covered in studies like "Auditory Display: Sonification, Audification, and Auditory Interfaces" by Hermann, Hunt, and Neuhoff. Users can interact with technology more skillfully and perceive their environment better when such indicators are present [65].

Gesture recognition and voice commands: Wearable aids' usability can be greatly improved by integrating voice commands and gesture recognition technologies. [31].

In conclusion, user-centered design concepts are essential in developing wearable assistive devices for the blind that are not only practical but also simple to use. Designers may create solutions that enhance the quality of life for people with visual impairments by incorporating end-user views and building on well-established studies in usability and accessibility.

6. Future trends and innovations

Emerging technologies like augmented reality (AR) and cutting-edge sensors are on the verge of revolutionizing assistive technology for the blind. By bridging the gap between the visually handicapped and their surroundings, augmented reality (AR) glasses have the potential to provide real-time information about the user's surroundings. Additionally, as sensors grow, they can dramatically improve environmental awareness and obstacle recognition, leading to increased independence. The future holds increasingly intelligent and context-aware wearable aids thanks to deep learning's ongoing advancements, including reinforcement learning and generative models. Real-time object recognition, natural language understanding, and even more advanced navigation help may be made possible by these advancements. Furthermore, integrating such technology with smart city infrastructure and accessible navigation systems offers new opportunities for visually impaired people to navigate urban areas safely and confidently.

7. Challenges in implementation

It is crucial for the successful implementation of deep learning-enhanced wearable aids in assistive technology to address these technological issues and ethical issues, as shown in figure 2 and 3 respectively[3].

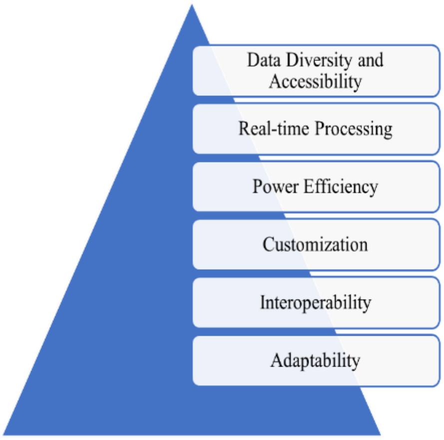

Fig 4. Technical issues

Technical and Real-World Difficulties:

1. Data Diversity and Accessibility: Deep learning models require a huge variety of datasets for training. It can be difficult to find complete datasets that include users with disabilities and varied settings. Model development is hampered by the lack of labeled data that appropriately reflects the needs of people with disabilities.

2. Real-time Processing: For prompt feedback, many assistive devices need to interpret sensory data in real-time. It is technically challenging to run deep learning algorithms at fast speeds and low latency on wearable devices. A primary focus is making sure users get prompt, accurate service.

3. Power Efficiency: Deep learning algorithms can be computationally demanding, and wearable devices are frequently battery-operated. Power efficiency and the necessity for precise projections must constantly be weighed in relation to one another. It is essential to optimize algorithms so that they can function well without depleting the device's battery.

4. Customization: Assistive technologies must consider the special demands and skills of each user. It is challenging to create deep learning models and algorithms that can be modified to satisfy unique needs and preferences. Personalization is crucial to ensuring that the tools offer useful support.

5. Interoperability: Deep learning-enhanced aids should smoothly interact with current assistive technologies and communicate with other devices to provide a seamless user experience. To build a cohesive ecosystem for assistive tools, compatibility and interoperability must be guaranteed.

6. Adaptability: Assistive technology must adjust to changing user settings and environmental conditions. It is difficult to create algorithms that dynamically modify settings to fit various circumstances. The equipment must be adaptable enough to offer dependable support in various circumstances.

Ethical and Privacy Issues:

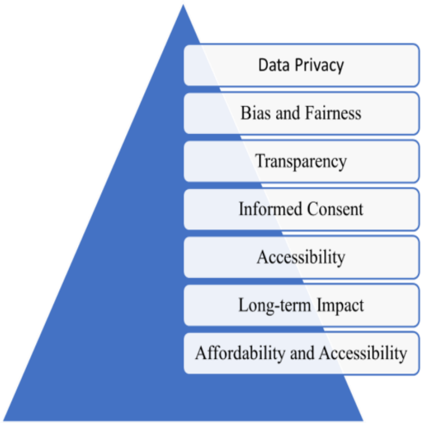

Fig 5. Ethical and privacy issues

1. Data Privacy: The collection and processing of user data, particularly sensitive biometric data, creates serious privacy issues. To avoid unwanted access or breaches, it is crucial to safeguard user data and establish strong data security procedures.
2. Bias and Fairness: Deep learning models may pick up biases from the training data, which could result in unjust or discriminating results. It is crucial to ensure fairness while creating and implementing these technologies, especially when it comes to assistive technology that affects people with impairments.
3. Transparency: Because of their intricacy, deep learning models are frequently referred to as "black boxes". Building trust and accountability requires that these models be transparent and easy to understand. It is important for users, developers, and regulators to comprehend how decisions are made.
4. Informed Consent: Users must be made fully aware of the collection, usage, and sharing of their data. Data control and obtaining informed consent are ethical requirements. The consequences of utilizing assistive technologies should be understood by users.
5. Accessibility: It's crucial to design deep learning-enhanced tools with accessibility in mind. Promoting inclusion and usability means making sure users with impairments can readily interact with and control the gadgets.
6. Long-term Impact: It's crucial to comprehend how deep learning-enhanced tools may affect users' long-term physical and mental health. Monitoring and mitigating potential negative effects to protect users' health and well-being are ethical considerations.
7. Affordability and Accessibility: It is ethical and a practical problem to make these technologies accessible and inexpensive for all people with impairments. For equal access, it is essential to ensure that price does not become a barrier to using assistive devices.

4. Results and conclusion

It is From the study focus on the Traditional aids have been superseded by cutting-edge wearable’s augmented by deep learning in the growth of assistive technologies for people with visual impairments, redefining independence, and inclusivity. Deep learning has driven OrCam MyEye, SignAll, BCIs, EMG-controlled prostheses, and adaptive hearing aids, delivering individualized, real-time help thanks to its neural networks, CNNs, RNNs, and NLP capabilities. To make sure that these technologies truly leave no one behind in creating a more accessible future, it is crucial to solve technological problems like data diversity, real-time processing, power efficiency, and cost, as well as ethical concerns like data protection, bias mitigation, and affordability.

We need to focus on designing small size object detection and object recognition systems which consider varying size images, to address the problems faced by the visually impaired in their passive and active stages, according to the study.

Acknowledgements

We thank our institute JSS Academy of Technical Education, Bengaluru and JSS STU Mysuru for providing support for carryout the research work

Author contributions

Rashmi B N: Conceptualization, Methodology, Software, Field study, Visualization, Investigation, Writing-Reviewing and Editing.

Dr R Guru: Reviewing and providing suggestions to improve the study

Dr Anasuya MA: Reviewing and providing suggestions to improve the study

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Sahand Shahalinejad, Mehdi Nooshyar ,Mir Majid Fakheri urmia, Nima ImaniObject Recognition and Detection Using Convolutional Neural Network in Deep Learning,in *bioRχiv ,(2023).*

[2] Satyaki Mukherjee; Tanya Sharma; Anshika Singh; Samiappan Dhanalakshmi,*in International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI)*, 2023.

[3] Mishra, S., Naqvi, T. K., & Dwivedi, P. K. Current Challenges and Perspectives of Flexible and Wearable Sensors. Flexible and Wearable Sensors*: Materials, Technologies, and Challenges, 351, (2023*).

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

[5] Wang, H., Wang, G., & Li, X. An RGB-D camera-based indoor occupancy positioning system for complex and densely populated scenarios. Indoor and Built Environment, 1420326X231155112. (2023).

[6] Mygal, G., Protasenko, O., Kobrina, N., & Mykhailova, E.). *Ergonomic thinking in the design of human-machine systems.]*, (2023).

[7] Seng, Kah Phooi, Li-Minn Ang, Eno Peter, and Anthony Mmonyi. "Machine Learning and AI Technologies for Smart Wearables" Electronics 12, no. 7: 1509. https://doi.org/10.3390/electronics12071509, 2023.

[8] Bergsten, P., & Kihan, M. Spatialisation of Binaural Audio With Head Tracking in First-Person Computer Games, (2023).

[9] Yanagawa, M., Ito, R., Nozaki, T., Fujioka, T., Yamada, A., Fujita, S., ... & Naganawa, S. New trend in artificial intelligence-based assistive technology for thoracic imaging. La radiologia medica, 1-14, (2023).

[10] Rahmayanti, A., Kahar, S., & Puspitasari, T. D. Utilizing Deep Learning Algorithm for Enhanced Visual Product Identification. EUROPEAN JOURNAL OF INNOVATION IN NONFORMAL EDUCATION, 3(7), 45-50, (2023).

[11] Renju Rachel Varghese; Pramod Mathew Jacob; Midhun Shaji; Abhijith R; Emil Saji John; Sebin Beebi Philip,An Intelligent Voice Assistance System for Visually Impaired using Deep Learning,in , 2022.

[12] Ganesan, J., Azar, A. T., Alsenan, S., Kamal, N. A., Qureshi, B., & Hassanien, A. E.Deep learning reader for visually impaired. Electronics, 11(20), 3335, (2022)..

[13] Adegboro, S., & Faniran, T. S. Development of Smart Intelligent Walking Aid 3rd Eye for the Blind Using Ultrasonic Sensor. University of Ibadan Journal of Science and Logics in ICT Research, 8(2), 27-36, (2022).

[14] .Harsh Shah; Rishi Shah; Shlok Shah; Paawan SharmaDangerous Object Detection for Visually Impaired People using Computer Vision,in International Conference on Artificial Intelligence and Machine Vision IEEE, (2022).

[15] Boldini, A., Ma, X., Rizzo, J. R., & Porfiri, M. A virtual reality interface to test wearable electronic travel aids for the visually impaired. In Nano-, Bio-, Info-Tech Sensors and Wearable Systems (Vol. 11590, pp. 50-56). SPIE, (2021, March).

[16] Mrozek, D., Kwaśnicki, S., Sunderam, V., Małysiak-Mrozek, B., Tokarz, K., & Kozielski, S. Comparison of Speech Recognition and Natural Language Understanding Frameworks for Detection of

Dangers with Smart Wearables. In International Conference on Computational Science (pp. 471-484). Cham: Springer International Publishing., (2021, June).

[17] Jiang, S., Kang, P., Song, X., Lo, B. P., & Shull, P. B. Emerging wearable interfaces and algorithms for hand gesture recognition: A survey. IEEE Reviews in Biomedical Engineering, 15, 85-102, (2021).

[18] Ntakolia, C., & Iakovidis, D. K. A route planning framework for smart wearable assistive navigation systems. SN Applied Sciences, 3(1), 104, (2021).

[19] .Nitin Kumar; Anuj Jain,Smart Navigation Detection using Deep-learning for Visually Impaired Person,in IEEE 2nd International Conference On Electrical Power and Energy Systems (ICEPES), 2021.

[20] Saravanan Alagarsamy; T. Dhiliphan Rajkumar; K. P. L. Syamala; Ch. Sandya Niharika; D. Usha Rani,K. Balaji,An Real Time Object Detection Method for Visually Impaired Using Machine Learning,in International Conference on Artificial Intelligence and Machine Vision,IEEE, 2021.

[21] Tapu, R., Mocanu, B., & Zaharia, T. Wearable assistive devices for visually impaired: A state of the art survey. Pattern Recognition Letters, 137, 37-52, (2020).

[22] Basu, A., Petropoulakis, L., Di Caterina, G., & Soraghan, J. Assistive technology evolving as intelligent system. New Trends in Computational Vision and Bio-inspired Computing: Selected works presented at the ICCVBIC 2018, Coimbatore, India, 289-303, (2020).

[23] Olwal, A., Balke, K., Votintcev, D., Starner, T., Conn, P., Chinh, B., & Corda, B. Wearable subtitles: Augmenting spoken communication with lightweight eyewear for all-day captioning. In Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (pp. 1108-1120), (2020, October).

[24] Alamdari, N., Lobarinas, E., & Kehtarnavaz, N. Personalization of hearing aid compression by human-in-the-loop deep reinforcement learning. IEEE Access, 8, 203503-203515, (2020).

[25] .Tanvir Ahmad,Yinglong Ma,Muhammad Yahya,Belal Ahmad,Shah Nazir,and Amin ul HaqObject Detection through Modified YOLO Neural Network,Volume 2020 | Article ID 8403262, 2020.

[26] Yanfen Li; Hanxiang Wang; L. Minh Dang; Tan N. Nguyen; Dongil Han; Ahyun Lee; Insung Jang Hyeonjoon Moon,A Deep Learning-Based Hybrid Framework for Object Detection and Recognition in Autonomous Driving,IEEE Access (Volume: 8) ,(2020)

[27] Jafarzadeh, M., Hussey, D. C., & Tadesse, Y. (2019, September). Deep learning approach to control of prosthetic hands with electromyography signals. In IEEE International Symposium on Measurement and Control in Robotics (ISMCR) (pp. A1-4). IEEE(2019).

[28] Tang, S. Y., Hoang, N. S., Chui, C. K., Lim, J. H., & Chua, M. C. Development of wearable gait assistive device using recurrent neural network. In IEEE/SICE International Symposium on System Integration (SII) (pp. 626-631). IEEE, (2019, January).

[29] Mohammed, S. L., Al-Naji, A., Farjo, M. M., & Chahl, J. Highly accurate water level measurement system using a microcontroller and an ultrasonic sensor. In IOP Conference Series: Materials Science and Engineering (Vol. 518, No. 4, p. 042025). IOP Publishing, (2019, May).

[30] Hakim, H., & Fadhil, A. Navigation system for visually impaired people based on RGB-D camera and ultrasonic sensor. In Proceedings of the International Conference on Information and Communication Technology (pp. 172-177), (2019, April).

[31] .Murad, C. Tools to support voice user interface design. In Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services (pp. 1-5), (2019, October).

[32] World Health Organization (WHO), Visual impairment and blindness, accessed on http://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment, (23.05.2018).

[33] Katzschmann, R. K., Araki, B., & Rus, D. Safe local navigation for visually impaired users with a time-of-flight and haptic feedback device. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26(3), 583-593, (2018).

[34] Hamayun A. Khan,DM-L Based Feature Extraction and Classifier Ensemble for Object Recognition,Journal of Signal and Information Processing,Vol.9 No.2, May 2018

[35] Wenzhi Zhao; Shihong Du; William J. Emery,Object-Based Convolutional Neural Network for High-Resolution Imagery Classification, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (Volume: 10, Issue: 7, July 2017)

[36] Wenzhou; Shihong Du; William J. Emery,Object-Based Convolutional Neural Network for High-Resolution Imagery Classification,in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (Volume: 10, Issue: 7, July 2017)

[37] Gao Huang; Zhuang Liu; Laurens Van Der Maaten; Kilian Q. Weinberger,Densely Connected Convolutional Networks, IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017

[38] .Andrew G. Howard Menglong Zhu Bo Chen Dmitry Kalenichenko Weijun Wang Tobias Weyand Marco Andreetto Hartwig

Adam,MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,arXiv:1704.04861v1 [cs.CV] 17 Apr 2017

[39] François Chollet,Xception: Deep Learning with Depthwise Separable Convolutions,IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017

[40] François Chollet,Xception: Deep Learning with Depthwise Separable Convolutions,IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017

[41] Lazar, J., Feng, J. H., & Hochheiser, H. Research methods in human-computer interaction. Morgan Kaufmann, (2017).

[42] Ishihara, T., Vongkulbhisal, J., Kitani, K. M., & Asakawa, C. Beacon-guided structure from motion for smartphone-based navigation. In 2017 IEEE Winter Conference on Applications of Computer Vision (WACV) (pp. 769-777). IEEE, (2017, March).

[43] .Hamzah, R., & Fadzil, M. I. M. Voice4Blind: The talking braille keyboard to assist the visual impaired users in text messaging. In 2016 4th International Conference on User Science and Engineering (i-USEr) (pp. 265-270). IEEE, (2016, August).

[44] Mohamedali, F., & Matoorian, N. Support dementia: using wearable assistive technology and analysing real-time data. In International Conference on Interactive Technologies and Games (ITAG) (pp. 50-54). IEEE, (2016, October).

[45] Kaiming He; Xiangyu Zhang; Shaoqing Ren; Jian Sun,Deep Residual Learning for Image Recognition,in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016

[46] Joseph Redmon, Ali FarhadiYOLO9000: Better, Faster, Stronger,in Computer Vision and Pattern Recognition ,arXiv:1612.08242,2016

[47] Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, in Computer Vision and Pattern Recognition,arXiv:1311.2524, (2015).

[48] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian SunSpatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognitio,Computer Vision and Pattern Recognition,arXiv:1406.4729,2015

[49] Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik, Rich feature hierarchies for accurate object detection and semantic segmentation,Computer Vision and Pattern Recognition,arXiv:1311.2524, 2015

[50] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun,Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition,aComputer Vision and Pattern Recognition,rXiv:1406.4729,2015

[51] Christian Szegedy; Wei Liu; Yangqing Jia; Pierre Sermanet; Scott Reed; Dragomir Anguelov; Dumit Erhan,,Vincent Vanhoucke,Andrew Rabinovich,Going deeper with convolutions in IEEE Conference on Computer Vision and Pattern Recognition (CVPR)IEEE 2015

[52] .Karen Simonyan, Andrew Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition,Computer Vision and Pattern Recognition,aXiv:1409.1556, (2015)

[53] .Guanbin Li, Yizhou Yu,Visual Saliency Based on Multiscale Deep Featuresin in Computer Vision and Pattern Recognition, arXiv:1503.08663, (2015)

[54] Martinez-Sala, A. S., Losilla, F., Sánchez-Aarnoutse, J. C., & García-Haro, J. Design, implementation and evaluation of an indoor navigation system for visually impaired people. Sensors, 15(12), 32168-32187, (2015).

[55] . Lane, N. D., Bhattacharya, S., Georgiev, P., Forlivesi, C., & Kawsar, F. An early resource characterization of deep learning on wearables, smartphones and internet-of-things devices. In Proceedings of the 2015 international workshop on internet of things towards applications (pp. 7-12), (2015, November).

[56] .Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik, Rich feature hierarchies for accurate object detection and semantic segmentation,Computer Vision and Pattern Recognition, arXiv:1311.2524v5, 2014

[57] Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik,Rich feature hierarchies for accurate object detection and semantic segmentation(in Computer Vision and Pattern Recognition,arXiv:1311.2524, 2014,

[58] .Matthew D Zeiler, Rob Fergus,Visualizing and Understanding Convolutional Networksin Computer Vision and Pattern Recognition arXiv:1311.2901,(2013),

[59] Oz, C., & Leu, M. C. American sign language word recognition with a sensory glove using artificial neural networks. Engineering Applications of Artificial Intelligence, 24(7), 1204-1213 (2011).

[60] Yelamarthi, K., Haas, D., Nielsen, D., & Mothersell, S. RFID and GPS integrated navigation system for the visually impaired. In 53rd IEEE International Midwest Symposium on Circuits and Systems (pp. 1149-1152). IEEE, (2010, August).

[61] Matthew D. Zeiler; Dilip Krishnan; Graham W. Taylor; Rob Fergus,Deconvolutional networks in : IEEE Computer Society Conference on Computer Vision and Pattern Recognition , 2010

[62] Nicholson, J., Kulyukin, V., & Coster, D. ShopTalk: independent blind shopping through verbal route directions and barcode scans. The Open Rehabilitation Journal, 2(1), (2009).

[63] Akita, J., Ito, K., Ono, T., & Okamoto, M.CyARM: haptic sensing device for spatial localization on basis of exploration by arms. Advances in Human-Computer Interaction, (2009).

[64] . Dakopoulos, D., & Bourbakis, N. G. Wearable obstacle avoidance electronic travel aids for blind: a survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40(1), 25-35, (2009).

[65] Peres, S. C., Best, V., Brock, D., Frauenberger, C., Hermann, T., Neuhoff, J. G., ... & Stockman, A. Auditory interfaces. HCI beyond the GUI: design for haptic, speech, olfactory, and other nontraditional interfaces, 147-195, (2008).

[66] . Rubin, J., & Chisnell, D. Handbook of usability testing: How to plan, design, and conduct effective tests. John Wiley & Sons., (2008).

[67] . Cardin, S., Thalmann, D., & Vexo, F. A wearable system for mobility improvement of visually impaired people. The Visual Computer, 23, 109-11, (2007).

[68] V. Kulyukin, C. Gharpure and J. Nicholson, "RoboCart: toward robot-assisted navigation of grocery stores by the visually impaired," IEEE/RSJ International Conference on Intelligent Robots and Systems, Edmonton, AB, Canada, pp. 2845-2850, doi: 10.1109/IROS.2005.1545107, (2005)

[69] .Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (Brain–computer interfaces for communication and control. Clinical neurophysiology, 113(6), 767-7912002)

[70] Kay, L. Auditory perception of objects by blind persons, using a bioacoustic high resolution air sonar. The Journal of the Acoustical Society of America, 107(6), 3266-3275, (2000).

[71] .L. Itti; C. Koch; E. Niebur,A model of saliency-based visual attention for rapid scene analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence (Volume: 20, Issue: 11, November 1998)

[72] .Object Detection and Recognition for Visually Impaired Using Cross-Stage-Partial Network Algorithms,

[73] . https://www.nngroup.com/articles/usability-101-introduction-to-usability/

[74] . https://www.noorcam.com/en-ae/noorcam-myeye

[75] .Obstacle and planar object detection using sparse 3D information for a smart walke

[76] . Kaiming He, Georgia Gkioxari, Piotr Dollár, Ross Girshick,arXiv:1703.06870,24 Jan 2018 .

[77] .Image et Reconnaissance de Forme-Systèmes Intelligents et Communicants (IRF-SIC) Laboratory, Faculty of Sciences, Ibn Zohr University, Agadir 80000, Morocco

[78] Tahira Shehzadi, Khurram Azeem Hashmi, Didier Stricker, Muhammad Zeshan Afzal,Object Detection with Transformers: A Review,Computer Vision and Pattern Recognition,10 Jul 2023

[79] .Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, Sergey Zagoruyko,Computer Vision and Pattern Recognition,arXiv:2005.12872 ,28 May 2020

[80] .Forrest N. Iandola1 , Song Han2 , Matthew W. Moskewicz1 , Khalid Ashraf1 , William J. Dally2 , Kurt Keutzer1 1DeepScale∗ & UC Berkeley 2Stanford Univers