Article

A Survey on Recent Advances in AI and Vision-Based Methods for Helping and Guiding Visually Impaired People

Hélène Walle ^{1,*}, Cyril De Runz ^{1,*} , Barthélemy Serres ² and Gilles Venturini ¹

¹ LIFAT, University of Tours, FR-37000 Tours, France; helene.walle@univ-tours.fr (H.W.); gilles.venturini@univ-tours.fr (G.V.)

² CETU ILIAD3, University of Tours, FR-37000 Tours, France; barthelemy.serres@univ-tours.fr

* Correspondence: cyril.derunz@univ-tours.fr

Abstract: We present in this paper the state of the art and an analysis of recent research work and achievements performed in the domain of AI-based and vision-based systems for helping blind and visually impaired people (BVIP). We start by highlighting the recent and tremendous importance that AI has acquired following the use of convolutional neural networks (CNN) and their ability to solve image classification tasks efficiently. After that, we also note that VIP have high expectations about AI-based systems as a possible way to ease the perception of their environment and to improve their everyday life. Then, we set the scope of our survey: we concentrate our investigations on the use of CNN or related methods in a vision-based system for helping BVIP. We analyze the existing surveys, and we study the current work (a selection of 30 case studies) using several dimensions such as acquired data, learned models, and human–computer interfaces. We compare the different approaches, and conclude by analyzing future trends in this domain.

Keywords: blind and visually impaired people; assistive technologies; artificial intelligence

Citation: Walle, H.; De Runz, C.; Serres, B.; Venturini, G. A Survey on Recent Advances in AI and Vision-Based Methods for Helping and Guiding Visually Impaired People. *Appl. Sci.* **2022**, *12*, 2308. <https://doi.org/10.3390/app12052308>

Academic Editor: Giuseppe Andreoni

Received: 10 January 2022

Accepted: 15 February 2022

Published: 23 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

According to the World Health Organization, 285 million people suffer from important sight loss (39 million blind and 246 million with impaired vision), and the figures will keep rising as populations grow older. Assisting blind and visually impaired people in their everyday life is a long-standing research topic, with traveling a particular concern. Traditionally, white canes and guide dogs have acted as walking assistants, but recent advances in deep learning and computer-vision technologies have broadened the spectrum of possibilities.

Despite being a classical topic that has been investigated for decades, research teams are continually innovating and offering hope for a future where vision disability is not a constant struggle. From radars in the mid-20th century to the latest AI emerging presently, assistive technologies have used an exceptionally diverse set of technologies to design tools for blind and visually impaired people. The bloom of new, better-performing algorithms are paving the way for future developments, becoming essential aspects on which to focus our survey.

Beyond leading to important innovations, numerous studies conducted over recent years have also brought clearer classifications of the developed assistive tools that are now widely used to define them. Assistive technologies aiming at easing travel are often divided into three areas: electronic travel aids, electronic orientation aids, and position locator devices, described by Elmannai et al. [1] as “devices that gather information about the surrounding environment and transfer it to the user”, as “devices that provide pedestrians with directions in unfamiliar places”, and as “devices that determine the precise position of its holder”, respectively. Tapu et al. [2] proposed another classification, depending on the kinds of skills involved, between perceptual tools that replace vision (images and

distances) with other sensory stimuli such as acoustic or haptic signals, and conceptual ones that develop orientation strategies (spatial modeling or surface mapping) to represent the environment and prepare for unpredictable situations during navigation.

Assistive systems need three modules to help blind and visually impaired people (BVIP). The first is the navigation module, or wayfinding, defined by Kandalan and Namuduri [3] as “the set of efficient movements required to reach the desired destination, which benefits from the knowledge of the user’s initial location and constant update of the user’s orientation”. This module should ideally provide path and surface description, a selection of the optimal path taking several criteria into account (user preference, avoidance of hazardous zones, etc.), accident-free navigation, and completion within a reasonable amount of time. Ideally, it must also work indoors and outdoors, in various light conditions (night/day, sun/rain, etc.), regardless the location has already been visited or not, and perform real-time analysis without sacrificing robustness and accuracy. The second module performs object detection with two purposes. First, it allows the avoidance of obstacles that might cause harm to the user, and warn the user in a helpful way. It must be able to detect static and dynamic obstacles and their locations (preferably ahead) and nature, and estimate their distance to provide timely feedback. Object detection should also provide scene description whenever the user asks, to give the BVIP a good understanding of their environment and build cognitive maps of a place. The last module is the human–machine interface that encompasses the tools that will be manipulated by the user. This is comprised of several devices meant to acquire data, process it, and finally return the information to the user. Many combinations are possible, and they must be chosen according to several criteria, such as algorithmic methods, user preference or wearability of devices.

Several surveys have been conducted in recent years on assistive technologies for BVIP, and the range of technologies available that could be used to develop them. Some of those articles were read when preparing this paper, to understand the different ways this subject has been addressed in the past. In their articles, Bhowmick and Hazarika [4] and Khan et al. [5] provided detailed descriptions of their research methodologies, highlighting the connections between the different fields pertaining to assistive technologies for BVIP. They spotted the most frequent keywords found in the papers they analyzed and their common sources, and stressed the trends of research. Some papers concentrated on the technological aspects of navigation and object detection. Zhao et al. [6] and Jiao et al. [7] proposed a survey on object detection with deep learning, its history, its possible techniques, and its current trends. Ignatov et al. [8] and Leo et al. [9] also reviewed recent deep-learning techniques for processing images, with Ignatov et al. concentrating on their use in assistive technologies, and Leo et al. concentrating on hardware and frameworks that allow them to run on Android smartphones. El-Zahraa El-Taher et al. [10] and Kandalan et Namuduri [3] described several techniques for constructing navigation systems, the tasks that need to be performed, the hardware available, and the possible interfaces. They chose two different scopes: urban travel for El-Zahraa El-Taher et al. and indoor navigation for Kandalan and Namuduri. Several papers [1,2,11–14] have provided an analysis of different systems aimed at assisting BVIP in their everyday lives. They did so through several analysis dimensions, such as hardware components and techniques used (sensor or computer-vision-based systems) or types of interface, proposing more or less detailed descriptions. Some papers have also described current or future trends and challenges, or even given recommendations for future development. It should be noted that Plikynas et al. [13] and Tapu et al. [2] associated with blind experts to conduct their surveys, and benefited from an end-user point of view on the proposed assistive devices.

This paper has tried to adopt an original approach. First, only the most recent papers were analyzed to build this survey, aiming to highlight the most recent advances in a constantly evolving research field. The whole development process was taken into account, aiming to understand all the aspects involved when helping BVIP, and not concentrating on a single AI technology or assistive function. Those aspects are highlighted separately to

provide a detailed overview of possibilities in this field, with statistics on the most common and relevant choices among researchers.

The main contributions of this paper are:

- to provide a state-of-the-art survey based on the latest publications and updates;
- to provide a detailed comparison of several possible configurations of assistive tools for BVIP;
- to emphasize AI techniques, especially CNN methods.

The rest of this paper is organized as follows: Section 2 presents the research methodology. Sections 3–5 provides paper analysis on human–machine interfaces, AI technologies, and testing methods, respectively. Finally, Section 6 ends with a conclusion and some perspectives, highlighting achievements and difficulties to solve.

2. Method

The survey was conducted with the following search methodology (see Figure 1): the first step was determining the scope of the survey. The selected articles should be:

1. on assistive devices;
2. preferably using deep-learning technologies;
3. aiming at fulfilling navigation, obstacle detection and/or object recognition tasks;
4. designed for BVIP.

Figure 1. PRISMA selection diagram that explains the main steps of our survey methodology (see text for explanation).

A decision was also made to favor recent articles, ideally ranging from 2017 until the present, to provide only the latest technological advances in AI methods. Article quality was particularly important. All publications were assessed using the following tools:

- Qualis from the Computing Institute of the Federal University of Mato Grosso in Brazil (<https://qualis.ic.ufmt.br/>, visited 30 September 2021), with venues ranked A1 or A2;
- Core Conference Portal (<http://portal.core.edu.au/conf-ranks/>, visited 30 September 2021), with conferences ranked A or B;
- Scimago Journal and Country Ranks (<https://www.scimagojr.com/journalrank.php>, visited 30 September 2021), with journals or conferences ranked Q1 or Q2.

The articles were found using Google Scholar and by scanning through scientific databases such as IEEE Xplore, Elsevier ScienceDirect, ACM Digital Library, and PubMed. The search was made using a combination of the following keywords:

- “assistive technologies” or “assistive devices”;
- “navigation”;
- “obstacle detection” and/or “object recognition”;
- “artificial intelligence”, “deep learning” or “computer vision”;
- “blind” and/or “visually impaired people”.

A total of 78 identified articles were first filtered by reading their abstract to estimate the correspondence with the chosen topics, thus excluding unsuitable ones (33). A few more articles (12) removed from the list after quick reading, mostly due to being technologically inadequate regarding the subject of this survey. In this paper, it was decided to concentrate on external assistive devices and to set aside research based on vision replacement using medical bionic prosthetic systems such as the project described by Ge et al. in [15]. A total of 30 case studies, from 33 research papers, were finally incorporated in the analysis: papers [16–20] were considered to be sole entities, as they were parts of the same research study. Some more generalist articles were included to highlight important steps and challenges when building assistive tools for BVIP.

Analysis was done following three main themes, divided in the following sections:

- First, human–machine interface (data acquisition and processing, feedback transmissions);
- Second, artificial intelligence techniques (scope, algorithms, datasets, training techniques);
- Third, testing methods (context, end-user participation).

For each part, the results summarized in tables with their percentage of occurrence.

3. Human–Machine Interfaces for Data Acquisition and User Feedback

In this section, we consider the interface between BVIP and AI-based assistive systems in a broad sense: it encompasses data-acquisition techniques, data-processing approaches, and feedback given to the user. Several aspects must be taken into account when designing such human–machine interfaces. First, some general aspects need to be considered:

- The overall interface must be robust and reliable, given that failure of the system could potentially be very harmful (even physically harmful) to users, and capable of functioning over important time durations (effective management of energy consumption);
- The interface must be comfortable to wear, unobtrusive and discrete to avoid carrier stigmatization;
- The hardware components must be easily accessible, and the total cost must be affordable to most of the public;
- The system must also be user-friendly and require minimum training from users; a too-complex interface often results in misuse or abandonment of the device.

3.1. Data Acquisition and Processing

The acquisition and processing hardware must be adapted to deep-learning techniques chosen to perform the tasks (choice of sensors and camera types; remote or wearable processors). Furthermore, the acquisition step, as with many computer-vision studies or applications, is crucial. In the case of embedded assistive technologies, this data acquisition can be even more difficult, because the user is often a part of the system (the user holds or

warns the acquisition devices, while moving) and the conditions of acquisition cannot be easily controlled (day, night, rain, etc.).

3.1.1. Type of Acquisition Interface

Several solutions are possible when designing the acquisition part of interfaces (see Table 1), each with its own pros and cons. Smart glasses are frequently chosen as the interface type because of their many advantages, including being able to carry several acquisition tools, or adopting the eyes' point of view. They are also easy to wear and can be found quite easily (several models with built-in sensors, cameras and headsets are currently available). Their main drawback is high cost, making them unaffordable to many. Acquisition devices can also be mounted on a white cane, with the main advantage being the use of a tool that is already used by most BVIP. The problem is that electronic tools on the canes tend to make them heavy and uncomfortable to carry. The most common type of interface in the analyzed articles is the smartphone, a tool that is both easy to find and already used by most people. Smartphones are becoming increasingly technologically advanced, and some models are now well equipped with sensors and high-resolution cameras. Despite these advances, many older or cheaper models do not hold sufficient data-acquisition tools, and replacing them is sometimes a problem for some users. Some research teams have decided to create their own wearable interfaces to deal with the issues mentioned above. For example, Chen et al. [21] designed an acquisition interface worn on a headband, while Wang et al. [22] and Malek et al. [23] opted for small bags worn across the chest.

Table 1. The type of acquisition devices used in each paper.

Acquisition Device	% of Papers	Ref. to Papers
smart glasses	30	[19,24–31]
smart cane	6.7	[32,33]
smartphone	36.7	[18,34–42]
other wearable device	26.7	[21–23,43–47]

3.1.2. Data-Acquisition Tools

The types of sensors that acquire data must be chosen carefully, taking several factors into account (see Table 2). A varied set of sensors will increase system accuracy, which is crucial in obstacles and object-detection tasks. On the other hand, a huge amount of diverse data needs to be processed and fused, thus increasing the computational cost and achievement time for a task. When designing their systems, researchers had to find the right balance between those points, depending on the purpose at which they were aiming. One of the most common solutions is to use smartphones equipped with RGB-D or monocular cameras and position sensors (accelerometers, gyroscopes, and magnetometers), as shown in examples [27–29,33,42,44]. Another solution is the introduction of infra-red or laser sensors in the acquisition system, a choice made in articles [23,32,45]. For outdoor navigation, GPS is almost always chosen as the navigational support, being easily accessible, cheap, and having very wide coverage.

Table 2. Data-acquisition tools. By “position sensors”, we consider techniques and sensors such as inertial measurement units, odometry, or accelerometers/gyroscopes/magnetometers from a smartphone.

Acquisition Sensors	% of Papers	Ref. to Papers
monocular camera(s)	33.3	[18,28,36–40,42,45,46]
stereo vision (two cameras)	10	[26,31,47]
RGB-D camera	43.3	[19,21–25,27,29,30,33,41,43,44]
wide angle camera	6.7	[19,33]
GPS	10	[21,29,47]
position sensors	33.3	[21,27–29,32–34,42,44,47]
IR/laser	20	[19,23,31,32,44,45]

3.1.3. Types of Processors

Another choice to be made when designing the systems is the ways the data are going to be processed and analyzed (see Table 3). Several papers chose a smartphone as the sole processor, thus having limited treatment and energy capacities, but a cheap and easily accessible solution. Another option is to use a remote server to analyze the data, a solution with high computation power but with important risks of failure due to connection issues, especially when navigating indoors. Among the papers analyzed in this survey, the most common method was the addition of another wearable device, such as a tablet or a laptop, to act as the processor. This solution offers better computational power than a smartphone, and no connection issues; however, this option may cause a potential loss of comfort for users when carrying the system.

Table 3. Devices used for data processing. In some papers (mentioned with a “*”), a hybrid client–server architecture was used (smartphone/laptop + remote server).

Type of Device	% of Papers	Ref. to Papers
smartphone	23.3	[29,34,37,40–42] [46] *
tablet, laptop, etc.	53.3	[19,21–27,30,33,36,38,39,43,44] [32] *
remote server	23.3	[16,28,31,32,45–47]

When acquiring data, smartphones can be carried in the hand or worn on the body with a specific outfit (see Table 4). Despite being natural, occupying one hand with a phone may cause annoyance in everyday life activities, not to mention the risk of accidental dropping, or the risk of theft. Wearing the device on the body may be more comfortable, but localization must be chosen carefully to avoid social stigma: Tapu et al. [38] and Sato et al. [16] opted for a system worn on a belt, while Neugebauer et al. [37] chose a specific headset to place the smartphone on the head.

Table 4. How the smartphone is held by the users. Not all systems have been implemented on a smartphone.

Smartphone Position	% of Papers	Ref. to Papers
in hand	20	[34–36,40–42]
worn	13.3	[18,37–39]

3.2. Feedback

The design of the feedback module must be carefully chosen to avoid difficult user experiences. Feedback is defined by El-Zahraa El-Taher et al. [10] as “the means used by the system to convey information to the blind and visually impaired people.” To be efficient, the navigation instructions must be delivered quickly and clearly. They must be adapted to the difficulty of the task being performed (for example, scene description delivers more complex

information than instructions to turn left or right), but also to the type of environment encountered (level of noise). This module must respect user knowledge (choice of symbolic representations that will be easy to understand) and avoid sensory overload (information prioritization). Developers must also ensure that the interface delivering the feedback will not have a negative impact on other senses and communication capabilities (for example, by favoring bone-conducting headphones rather than traditional ones).

3.2.1. Type of Feedback

Feedback is essential in navigation tasks, especially when helping BVIP (see Table 5). A wrong choice may result in hazardous and potentially accident-prone situations for the user. The most popular choice is audio-feedback through speech instructions or sonification guidance (the use of sound or music to depict different elements present in a given scene). A tactile interface, such as a Braille display [22], or haptic systems, such as a vibrating smartphone [16,42], are among other possibilities. The most effective solution seems to be a combination of several feedback interfaces to adapt to the different situations encountered by users. For example, audio-feedback might not be adequate in noisy environments, but the best solution when much information needs to be delivered, as in scene description tasks. A combination of feedback types were illustrated by Bauer et al. [46], where navigational instructions (turn left or right) are delivered by vibrating smartwatches that can also perform an audio scene description, or by Li et al. [33] with two vibrating motors on a cane for turning left and right, and an Android smartphone with text-to-audio software to provide more precise instructions.

Table 5. General methods and techniques for providing feedback to users. Some tools have only been tested online, and their types of feedback remain theoretical.

Type of Feedback	% of Papers	Ref. to Papers
speech	70%	[16,20,21,23,24,26,28–30,32,33,35–39,41,43,44,46,47]
<i>in combination with other types</i>	20%	[16,30,33,35,41,46]
vibrations	13.3%	[16,22,33,46]
sonification	13.3%	[25,27,35,41]
tactile	10%	[22,30,42]

3.2.2. Feedback Conveyors

Several tools are available to deliver audio-feedback (see Table 6). In some cases, the instructions can be transmitted directly by smartphone, a method quite incompatible with user privacy. The use of headsets tends to be privileged by researchers: traditional ones are cheap and easily accessible, but often disrupt the sense of hearing on which BVIP rely heavily to understand their environment. Another solution is bone-conducting earphones that allow the delivery of instructions without covering the ears, therefore not impairing sound perception and offering the possibility to communicate with other people.

Table 6. Devices for audio feedback.

Audio-Feedback Devices	% of Papers	Ref. to Papers
earphones/headsets	46.7	[20,21,23,26,28–32,37,38,43,44,47]
bone-conducting earphones	16.7	[16,24,25,27,39]
phone/tablet	13.3	[33,36,40,41]

The devices for the other types of feedback are as follows: motors on a cane [33], vibrating belt and Braille interface [22], two smartwatches [46], and smartphone vibrations [16,42].

The users can also interact with the interface, either to choose a mode (wayfinding or scene description), or just to switch the system on or off. In most papers, this is done

via the smartphone, thanks to built-in voice controllers (iOS' Voice Control or Voice Over, Android's Voice Access or TalkBack) to set up the system. In [22], instructions are given through a Braille tablet, or with smartwatches in [46].

4. Artificial Intelligence Techniques

As research on AI and computer vision keep expanding, more and more solutions are becoming available to develop assistive tools for BVIP. Technical choices are motivated by several factors, and it is crucial for all the necessary features to be previously known for the engineering process to be carried out smoothly. One of the aspects that will influence technical choice is whether the system will be working indoors, outdoors, or both, with some tools having a limited use range (such as GPS, which is only accessible outside). Development teams must also take into account the kind of tasks that will be performed by the system—sole wayfinding, additional scene description—and the available data sources and computational capabilities, both of which depend on interface design. Some techniques, such as RFID tags and BLE beacons, require the installation of materials before the deployment of a system in a specific site, thus increasing the costs of use and maintenance. After deciding which algorithm models will be employed, researchers must set up the training procedure with the appropriate datasets and data treatment techniques.

4.1. Scope of System

Navigation systems can be designed to cover either indoor or outdoor navigation, thus only partially meeting the everyday needs of BVIP (see Table 7). Almost half the systems analyzed in this survey potentially propose solutions for both situations (although this double coverage was often only tested in one situation).

Table 7. Scope of assistive systems. Some studies do not clearly specify their targeted scope.

Scope of Systems	% of Papers	Ref. to Papers
indoor	30	[16,22,23,26,32–35,45]
outdoor	10	[24,31,36]
both	43.3	[20,21,25,27–30,39,42,43,46–48]

4.2. Machine- or Deep-Learning Algorithms

A lot of methods are currently available for developing tools based on navigation or object-detection tasks (see Table 8). In the field of navigation and wayfinding, the most popular methods in this survey are:

- SLAM (Simultaneous Localization And Mapping) algorithms, which are used to construct maps of the encountered environment and localize the user within it. Several methods were investigated in the analyzed papers: semantic visual SLAM (ORB SLAM) in [21], 2-STEP Graph SLAM in [32], VSLAM in [29], and ORB-SLAM2 in [19].
- RANSAC (RANDOM Sample Consensus) is a method for detecting (and eliminating) outliers, and was developed to solve the Location Determination Problem (extracting feature points and localizing them on a projection).
- A* is a search algorithm for wayfinding that uses heuristics to determine the path with the smallest cost.
- The Kalman filter algorithm (Linear Quadratic Estimation) is a method to estimate unknown variables from the observation of a series of measurements that can be employed for many tasks, such as pose estimation [26], obstacle motion estimation [33], or error reduction [42].

Other architecture is available for dealing with object recognition tasks, mostly used in the papers for the detection of obstacles and to describe scene content if needed:

- YOLO (You Only Look Once) is a CNN designed for real-time object detection (it recognizes what objects are present in a scene, and where) created in 2015. Several updated versions are currently available (YOLOv1-v3).

- VGG is a deep CNN architecture deriving from AlexNet, developed and already trained (on the ImageNet dataset) by Oxford University's Visual Geometry Group. It is mainly designed for image classification and object recognition tasks, and is adapted to the transfer-learning method. Two versions are available: VGG16 (16 convolutional layers) and VGG19 (19 layers).
- Inception is a CNN classifier developed by Google (and named after the movie) that serves the analysis of images and object detection. Among the several versions that have been released, only Inceptionv2 [42] and v3 [40,41] are used in the analyzed papers. Others are: Inceptionv1 (aka GoogLeNet), Inceptionv4 and InceptionResNetv1 and v2 (hybrid Inception and ResNet architecture).

To perform the complex tasks needed to build navigation and object recognition assistive tools, encoder/decoder architecture is also a popular choice, as shown in [21,23,25,31,48]. Numerous algorithms and architectures are possible when developing navigation and object-detection systems, and only the most frequent (among the analyzed samples) have been described in this survey. Despite the wide range of possibilities currently available, some papers opted for the development of specific algorithms tailored to their research objectives. They are used to perform specific tasks of the overall development such as list construction and object detection in [33], or object extraction and obstacle avoidance in [47].

Table 8. most frequent methods.

Method	% to Papers	Ref. to Papers
SLAM	16.7	[19,21,29,32,41]
Encoder/Decoder	16.7	[21,23,25,31,48]
RANSAC	16.7	[26,27,29,32,47]
A*	16.7	[19,28,29,32,33]
Kalman's filter	16.7	[26,28,29,33,42]
YOLO	16.7	[35,36,38,43,46]
VGG	13.3	[28,31,39,45]
Inception	10	[40–42]
specific algorithm	20	[22,31,33,43,47,48]

4.3. Choices of Datasets

A wide range of datasets is currently available. The most popular in the described papers were ImageNet, followed by COCO and PASCAL ensembles (see Table 9). Those datasets are quite general, and do not always fit the requirements of the systems developed. They are often used for pre-training models, before using specific datasets created specifically for the project. This last option was chosen in [31,36,38–40,42,45].

Table 9. The most frequent image datasets.

Datasets	% of Papers	Ref. to Papers
specific	36.7	[23,28,30,31,36,38–42,45]
ImageNet	30	[24,25,31,36,38–40,45,46]
PASCAL-VOC	10	[25,36,46]
COCO	13.3	[25,29,35,42]

4.4. Data-Processing Methods

Often, when designing systems for a specific task, finding or creating the right dataset is an important obstacle (see Table 10). To train model efficiently, researchers need to use special methods. The most frequent within the papers analyzed is transfer learning: the system is entirely pre-trained with an already available general dataset (such as ImageNet), then the last layers are trained with a smaller, more specific one. The other method used here is data augmentation. New synthetic images are created from the original dataset with

several techniques such as random flipping, cropping, scaling, rotation, and color jittering for [25], and rotation, skewing, mirroring, flipping, brightness and noise in [43].

Table 10. Techniques used for model training.

ML Techniques	% of Papers	Ref. to Papers
data augmentation	6.7	[25,43]
transfer learning	16.7	[28,39–42]

4.5. Type of Model Training

Two methods are possible when training a model (see Table 11). Most of the papers opted for offline training: the model is trained once with pre-defined datasets and then deployed. However, five papers [38–41,48] chose incremental learning, where the model keeps being trained with data acquired by the users themselves, to acquire increasing accuracy and deal more easily with complex situations uncovered by the datasets.

Table 11. Incrementality of learning procedures.

Type of Training	% of Papers	Ref. to Papers
incremental	16.7	[38–41,48]
offline	46.7	[24,25,28–31,34–36,39,40,42,43,46]

4.6. Solving Challenges

Despite promising results, these systems still must be improved to be deployed on smartphones and small devices. As Bertheliet et al. [49] state in their survey, “deep-learning-based methods have achieved state-of-the-art performance in many applications such as face recognition, semantic segmentation, object detection, etc. [but] to run these applications on embedded devices, the deep models need to be less-parametrized in size and time-efficient.” To address these issues, several compression techniques for CNNs are currently being studied:

- Pruning, which consists of removing unused parameters of a network while still achieving state-of-the-art results;
- Quantization, which approximates a neural network by reducing floating-point number precision, with higher risk of error and lower accuracy;
- Hash methods, which convert original features into low-dimensional hash codes, regrouping data according to similarity and avoiding redundancy. Some hashing systems are already available, such as HashedNets by NVIDIA;
- Knowledge distillation, which is based on the process of transferring knowledge from a deep neural network to a shallow one, while keeping the same efficiency and learning capacities.

Compression methods have been widely investigated in recent years, and are already proposed as part of several frameworks, such as TensorFlow Lite or Apple’s Core ML.

Another way of reducing computational needs is by optimizing the system architecture itself. Several methods to reduce the cost of convolutional operations have been successfully tested, such as the conception of modules (with specifically organized and sized layers), or the use of separable convolutional layers. However, the most promising results have emerged in the field of Neural Architecture Search (NAS), which aims to develop self-organized structures that automatically shape their design to fit targeted tasks (neural gas methods, neuroevolution, network morphism, supergraphs).

Incremental learning is another crucial challenge in research on assistive technologies with CNNs. Its implementation on embedded devices represents an important step for BVIP, as it will allow them to customize their system to detect personal belongings or recognize friends, colleagues, and family members. Many research teams are currently developing solutions, and proposals are being made, such as the ones described by Luo et al. [50]

in their recent survey. Incrementality is even more difficult to implement on embedded devices because of the limited computational power available.

5. Testing Methods

There are several ways of assessing an assistive system. First, the testing phase can be theoretical (virtual scenarios performed on computer) and/or practical (pre-determined paths or real-life environments). Systems can be assessed with objective performance metrics evaluating for instance accuracy, time of response, and best choice of path. Subjective evaluations can also be conducted to measure the level of acceptance and usefulness of the system for BVIP, with, for instance, wearability, appropriateness of feedback, or integrability in everyday life.

5.1. Types of Tests

Several phases of testing can be observed in the papers analyzed (see Table 12). Usually, the system is first investigated theoretically on a computer to check if its accuracy meets the standards, before progressing to practical tests in a given environment (either in pre-set paths or buildings, or unknown real-life conditions). Some studies, probably less advanced in development, only tested their proposed system online. Others skipped the theoretical phase, opting to deploy their system directly in real conditions.

Table 12. Method of testing.

Method	% of Papers	Ref. to Papers
only simulation	26.7	[23,31,34,35,38,42,45,48]
only on field	20	[16,20,22,27,37,44]
both	53.5	[21,24–26,28–30,32,33,36,39–41,43,46,47]

5.2. End-User Testing

The majority of analyzed systems were tested directly by blind or visually impaired users, which is important for checking whether the needs of the end user have been met (see Table 13). A third of papers were only tested by sighted people, blindfolded or not, leaving uncertainty about the adequacy of the proposed solution and the highlighted issues.

Table 13. Tests with BVIP or not (in general blindfolded people).

Tests with BVIP	% of Papers	Ref. to Papers
yes	56 %	[16,20,22,24–30,33,36,37,39–41,43]
no	34 %	[21,23,31,32,34,42,45–48]

5.3. Testing Panel

The number of visually impaired testers varied greatly between papers, depending on the possibilities of recruitment or time available to conduct testing phases (see Table 14). Gathering enough end users to test the system is important when developing an assistive tool to check if all the requirements are met. Furthermore, unique feedback will allow the correction of several aspects of the system, such as interface or wearability, to make it as easy and comfortable to use as possible.

Table 14. Number of VIP testers.

Nb of Testers	% of Papers	Ref. to Papers
<5	10	[28,33,36]
[5–15]	26.7	[22,24,25,27,37,39–41]
>15	16.7	[16,26,29,30,43]

6. Conclusions and Perspectives

6.1. Achievements

In this paper, we studied recent advances in the field of AI techniques for developing assistive technologies for BVIP. This survey was made by assessing recent case studies in the field with a well-defined research methodology. Several aspects of the proposed systems were analyzed, including their human–machine interface (acquisition and processing hardware, feedback types and conveyors), the AI techniques chosen, and the ways they have been assessed.

Helping disabled people, such as BVIP, has been a major research topic for a long time, and the burst of technologies in everyday life has dramatically accelerated the development of the field. Although lacking functionality, several smartphone apps are already available for BVIP and are being adopted by many to ease their everyday activities. Some examples are:

- *Lookout* (Android) and *SeeingAI* (iOS) to identify objects or people;
- *TapTapSee* and *VizWiz* (iOS/Android) to identify elements in pictures taken by the user;
- GPS apps *Ariadne GPS*, *Microsoft Soundscape* and *Blindsquare* (iOS) to provide customized environment descriptions;
- *Evelity* (iOS/Android) to navigate in equipped buildings using a specific GPS system,
- *MyMoveo* (iOS/Android) to activate connected crosswalks and spot important elements (such as entrance doors).

Benefiting from the recent innovations of AI, assistive technologies have developed possibilities while increasing robustness and efficiency. In the very near future, it will be possible to propose safe and reliable wayfinding systems for BVIP, as well as detailed scene description and customized object recognition. Devices are also becoming cheaper and more comfortable to wear, especially smartphones, which are now equipped with many sensors, such as cameras, gyroscopes, and accelerometers, and endowed with enough computing capacity to run deep- or machine-learning algorithms.

6.2. Limitations and Challenges

Despite varied and interesting proposals of robust and accurate systems, several challenges and perspectives must be studied and addressed. They are mainly related to the following points (as further discussed in this section):

- most of the reviewed systems fail to become fully operational, and reasons for this must be analyzed;
- human factors are very important for BVIP, and they must be carefully understood and taken into account, raising challenges for user-centered AI approaches;
- the use of embedded devices is mandatory, but it raises important constraints on AI-based techniques, due to limited storage and computational resources, and to higher power consumption. In addition, incrementality is becoming important for BVIP, as a way to adapt object recognition to specific objects, including personal ones;
- in comparison with other general fields, the available image datasets or pre-trained CNNs specific to BVIP are not widely accessible (or even defined). The adaptation of such models to specific conditions (cultural, dynamic, etc.) is difficult.

In most of the studied papers, the presented prototypes are not yet ready to be deployed. This difficulty in achieving the last step of development is caused by several factors that will have to be solved, with the lack of complete and available solutions representing an additional difficulty for BVIP. First, despite the continuing innovations in the field of AI and computer vision, researchers still face serious issues when developing navigation tools for BVIP. Many situations can be challenging for navigation systems:

- obstacles situated at ground or head levels are still difficult to detect, and those are very important for BVIP;
- many elements present in urban environments are geographically or culturally dependent (for example, New York yellow cabs are quite different from London black ones);

- appropriate scene description is quite subjective (what information is really important depends on the user and the context).

These issues are due to the huge complexity of the problems, but also to the fact that some research fields have been more developed than others in recent years, and not for the specific case of BVIP. For example, object detection and outdoor localization are much more advanced today than safe street crossing or environment mapping. Salient object detection was once a major issue when developing assistive technologies for BVIP, but is now about to be overcome thanks to recent innovation, such as that described by Ji et al. in [51]. In addition, it might seem a minor point, but one should note that most existing CNNs available for use are trained with images of objects or scenes that are not specific to BVIP. Many relevant objects or situations (for BVIP) are missing from training sets. Building such training sets and specific CNNs, and making them available to the scientific (and developer) community, would certainly help expand AI for BVIP.

Additionally, human factors are very important in such AI applications. There is often a gap between end-user needs and developer decisions. As shown in this paper, BVIP are not always engaged in the final testing phase, and even more rarely in the engineering process. Some important requirements often remain unfulfilled, thus leading users to abandoning the developed systems. Including BVIP as well as health professionals in the development process of AI-based systems is certainly necessary to increase the chance of those systems becoming operational. More generally, the human factors involved in such specific AI applications play a crucial role, and are part of user-centered AI approaches. Taking into account such factors raises challenging issues. For instance, the design of human-machine interfaces does not always take into account the user relationships to their environment. For example, many BVIP do not wish to walk around with smartphones in their hand, fearing dropping them or being robbed. Furthermore, scene description methods should receive additional attention for BVIP, who mostly rely on audio information.

Running deep-learning systems on small devices has become a central question for researchers worldwide, as the demand for ever-more efficient, sophisticated, and easily accessible AI applications increases. Beyond the search for neural network compression and optimization, as described in Section 4.6, several developments have been made for more adequate hardware units to host these systems. The subject of customizing deep-learning methods to fit targeted tasks and types of hardware has been widely studied on recent years, in particular the use of graphics processing units (GPUs), field programmable gate arrays (FPGAs) and application-specific integrated circuits (ASICs), thus highlighting their pros and cons. The developments made in recent years have been analyzed and summarized by several authors: Seng et al. [52] for FPGAs, Moolchandani et al. [53] for ASICs and Ang et al. [54] for GPUs.

Other aspects of deep-learning models increase the difficulty of such applications: for instance, an incremental learning model is important, as it can allow the user to teach an AI new objects to recognize. These can be personal objects that belong to the user, objects that are culturally dependent, or even people (such as family members) or pets. Incrementality raises the problem of dynamic learning, which becomes even more complex on embedded devices with limited resources.

Finally, in the future, intelligent assistive technologies for BVIP will have to interact with connected areas, such as smart cities. Integrating disabled people and their assistants in the development of future environments is a question that needs to be taken into account by researchers. A few proposals are already available, such as those proposed by Chang et al. [55] in their recent article.

Author Contributions: Conceptualization, H.W., C.D.R., B.S. and G.V.; methodology, H.W., C.D.R. and G.V.; validation, H.W., C.D.R., B.S. and G.V.; formal analysis, H.W.; investigation, H.W.; data curation, H.W.; writing—original draft preparation, H.W. and G.V.; writing—review and editing, C.D.R., B.S. and G.V.; supervision, C.D.R. and G.V.; project administration, C.D.R.; funding acquisition, C.D.R. and G.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Elmannai, W.; Elleithy, K. Sensor-based assistive devices for visually-impaired people: Current status, challenges, and future directions. *Sensors* **2017**, *17*, 565. [[CrossRef](#)] [[PubMed](#)]
2. Tapu, R.; Mocanu, B.; Zaharia, T. Wearable assistive devices for visually impaired: A state of the art survey. *Pattern Recognit. Lett.* **2020**, *137*, 37–52. [[CrossRef](#)]
3. Kandalan, R.N.; Namuduri, K. Techniques for Constructing Indoor Navigation Systems for the Visually Impaired: A Review. *IEEE Trans. Hum.-Mach. Syst.* **2020**, *50*, 492–506. [[CrossRef](#)]
4. Bhowmick, A.; Hazarika, S.M. An insight into assistive technology for the visually impaired and blind people: State-of-the-art and future trends. *J. Multimodal User Interfaces* **2017**, *11*, 149–172. [[CrossRef](#)]
5. Khan, S.; Nazir, S.; Khan, H.U. Analysis of Navigation Assistants for Blind and Visually Impaired People: A Systematic Review. *IEEE Access* **2021**, *9*, 26712–26734. [[CrossRef](#)]
6. Zhao, Z.Q.; Zheng, P.; Xu, S.t.; Wu, X. Object detection with deep learning: A review. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 3212–3232. [[CrossRef](#)]
7. Jiao, L.; Zhang, F.; Liu, F.; Yang, S.; Li, L.; Feng, Z.; Qu, R. A survey of deep learning-based object detection. *IEEE Access* **2019**, *7*, 128837–128868. [[CrossRef](#)]
8. Ignatov, A.; Timofte, R.; Chou, W.; Wang, K.; Wu, M.; Hartley, T.; Van Gool, L. AI benchmark: Running deep neural networks on android smartphones. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 288–314.
9. Leo, M.; Furnari, A.; Medioni, G.G.; Trivedi, M.; Farinella, G.M. Deep learning for assistive computer vision. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 3–14.
10. El-Zahraa El-Taher, F.; Taha, A.; Courtney, J.; McKeever, S. A systematic review of urban navigation systems for visually impaired people. *Sensors* **2021**, *21*, 3103. [[CrossRef](#)]
11. Islam, M.M.; Sadi, M.S.; Zamli, K.Z.; Ahmed, M.M. Developing walking assistants for visually impaired people: A review. *IEEE Sens. J.* **2019**, *19*, 2814–2828. [[CrossRef](#)]
12. Real, S.; Araujo, A. Navigation systems for the blind and visually impaired: Past work, challenges, and open problems. *Sensors* **2019**, *19*, 3404. [[CrossRef](#)]
13. Plikynas, D.; Žvironas, A.; Budrionis, A.; Gudauskis, M. Indoor navigation systems for visually impaired persons: Mapping the features of existing technologies to user needs. *Sensors* **2020**, *20*, 636. [[CrossRef](#)]
14. Kuriakose, B.; Shrestha, R.; Sandnes, F.E. Tools and Technologies for Blind and Visually Impaired Navigation Support: A Review. *IETE Tech. Rev.* **2020**, 1–16. [[CrossRef](#)]
15. Ge, C.; Kasabov, N.; Liu, Z.; Yang, J. A spiking neural network model for obstacle avoidance in simulated prosthetic vision. *Inf. Sci.* **2017**, *399*, 30–42. [[CrossRef](#)]
16. Sato, D.; Oh, U.; Guerreiro, J.; Ahmetovic, D.; Naito, K.; Takagi, H.; Kitani, K.M.; Asakawa, C. NavCog3 in the wild: Large-scale blind indoor navigation assistant with semantic features. *ACM Trans. Access. Comput. (TACCESS)* **2019**, *12*, 1–30. [[CrossRef](#)]
17. Murata, M.; Ahmetovic, D.; Sato, D.; Takagi, H.; Kitani, K.M.; Asakawa, C. Smartphone-based indoor localization for blind navigation across building complexes. In Proceedings of the 2018 IEEE International Conference on Pervasive Computing and Communications (PerCom), Pisa, Italy, 21–25 March 2018; pp. 1–10.
18. Sato, D.; Oh, U.; Naito, K.; Takagi, H.; Kitani, K.; Asakawa, C. Navcog3: An evaluation of a smartphone-based blind indoor navigation assistant with semantic features in a large-scale environment. In Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility, Baltimore, MD, USA, 29 October–1 November 2017; pp. 270–279.
19. Bai, J.; Lian, S.; Liu, Z.; Wang, K.; Liu, D. Virtual-blind-road following-based wearable navigation device for blind people. *IEEE Trans. Consum. Electron.* **2018**, *64*, 136–143. [[CrossRef](#)]
20. Bai, J.; Lian, S.; Liu, Z.; Wang, K.; Liu, D. Smart guiding glasses for visually impaired people in indoor environment. *IEEE Trans. Consum. Electron.* **2017**, *63*, 258–266. [[CrossRef](#)]
21. Chen, Z.; Liu, X.; Kojima, M.; Huang, Q.; Arai, T. A Wearable Navigation Device for Visually Impaired People Based on the Real-Time Semantic Visual SLAM System. *Sensors* **2021**, *21*, 1536. [[CrossRef](#)]

22. Wang, H.C.; Katschmann, R.K.; Teng, S.; Araki, B.; Giarré, L.; Rus, D. Enabling independent navigation for visually impaired people through a wearable vision-based feedback system. In Proceedings of the 2017 IEEE International Conference on robotics and Automation (ICRA), Singapore, 29 May–30 June 2017; pp. 6533–6540.
23. Malek, S.; Melgani, F.; Mekhalfi, M.L.; Bazi, Y. Real-time indoor scene description for the visually impaired using autoencoder fusion strategies with visible cameras. *Sensors* **2017**, *17*, 2641. [\[CrossRef\]](#)
24. Lin, S.; Cheng, R.; Wang, K.; Yang, K. Visual localizer: Outdoor localization based on convnet descriptor and global optimization for visually impaired pedestrians. *Sensors* **2018**, *18*, 2476. [\[CrossRef\]](#)
25. Yang, K.; Wang, K.; Bergasa, L.M.; Romera, E.; Hu, W.; Sun, D.; Sun, J.; Cheng, R.; Chen, T.; López, E. Unifying terrain awareness for the visually impaired through real-time semantic segmentation. *Sensors* **2018**, *18*, 1506. [\[CrossRef\]](#)
26. Simões, W.C.; Silva, Y.M.; Pio, J.L.d.S.; Jazdi, N.; F de Lucena, V. Audio Guide for Visually Impaired People Based on Combination of Stereo Vision and Musical Tones. *Sensors* **2020**, *20*, 151. [\[CrossRef\]](#)
27. Hu, W.; Wang, K.; Yang, K.; Cheng, R.; Ye, Y.; Sun, L.; Xu, Z. A comparative study in real-time scene sonification for visually impaired people. *Sensors* **2020**, *20*, 3222. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Son, H.; Krishnagiri, D.; Jeganathan, V.S.; Weiland, J. Crosswalk guidance system for the blind. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020; pp. 3327–3330.
29. Bai, J.; Liu, Z.; Lin, Y.; Li, Y.; Lian, S.; Liu, D. Wearable travel aid for environment perception and navigation of visually impaired people. *Electronics* **2019**, *8*, 697. [\[CrossRef\]](#)
30. Lin, Y.; Wang, K.; Yi, W.; Lian, S. Deep learning based wearable assistive system for visually impaired people. In Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, Seoul, Korea, 27–28 October 2019; pp. 2549–2557.
31. Dimas, G.; Diamantis, D.E.; Kalozoumis, P.; Iakovidis, D.K. Uncertainty-Aware Visual Perception System for Outdoor Navigation of the Visually Challenged. *Sensors* **2020**, *20*, 2385. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Zhang, H.; Ye, C. An indoor wayfinding system based on geometric features aided graph SLAM for the visually impaired. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2017**, *25*, 1592–1604. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Li, B.; Muñoz, J.P.; Rong, X.; Chen, Q.; Xiao, J.; Tian, Y.; Arditi, A.; Yousuf, M. Vision-based mobile indoor assistive navigation aid for blind people. *IEEE Trans. Mob. Comput.* **2018**, *18*, 702–714. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Mahida, P.; Shahrestani, S.; Cheung, H. Deep Learning-Based Positioning of Visually Impaired People in Indoor Environments. *Sensors* **2020**, *20*, 6238. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Yang, G.; Saniie, J. Sight-to-Sound Human-Machine Interface for Guiding and Navigating Visually Impaired People. *IEEE Access* **2020**, *8*, 185416–185428. [\[CrossRef\]](#)
36. Lin, B.S.; Lee, C.C.; Chiang, P.Y. Simple smartphone-based guiding system for visually impaired people. *Sensors* **2017**, *17*, 1371. [\[CrossRef\]](#)
37. Neugebauer, A.; Rifai, K.; Getzlaff, M.; Wahl, S. Navigation aid for blind persons by visual-to-auditory sensory substitution: A pilot study. *PLoS ONE* **2020**, *15*, e0237344. [\[CrossRef\]](#)
38. Tapu, R.; Mocanu, B.; Zaharia, T. DEEP-SEE: Joint object detection, tracking and recognition with application to visually impaired navigational assistance. *Sensors* **2017**, *17*, 2473. [\[CrossRef\]](#)
39. Mocanu, B.; Tapu, R.; Zaharia, T. Deep-see face: A mobile face recognition system dedicated to visually impaired people. *IEEE Access* **2018**, *6*, 51975–51985. [\[CrossRef\]](#)
40. Kacorri, H.; Kitani, K.M.; Bigham, J.P.; Asakawa, C. People with visual impairment training personal object recognizers: Feasibility and challenges. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 5839–5849.
41. Ahmetovic, D.; Sato, D.; Oh, U.; Ishihara, T.; Kitani, K.; Asakawa, C. Recog: Supporting blind people in recognizing personal objects. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–12.
42. Lo Valvo, A.; Croce, D.; Garlisi, D.; Giuliano, F.; Giarré, L.; Tinnirello, I. A Navigation and Augmented Reality System for Visually Impaired People. *Sensors* **2021**, *21*, 3061. [\[CrossRef\]](#)
43. Joshi, R.C.; Yadav, S.; Dutta, M.K.; Travieso-Gonzalez, C.M. Efficient Multi-Object Detection and Smart Navigation Using Artificial Intelligence for Visually Impaired People. *Entropy* **2020**, *22*, 941. [\[CrossRef\]](#)
44. Grayson, M.; Thieme, A.; Marques, R.; Massiceti, D.; Cutrell, E.; Morrison, C. A dynamic AI system for extending the capabilities of blind people. In Proceedings of the Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–4.
45. Cornacchia, M.; Kakilioglu, B.; Zheng, Y.; Velipasalar, S. Deep learning-based obstacle detection and classification with portable uncalibrated patterned light. *IEEE Sens. J.* **2018**, *18*, 8416–8425. [\[CrossRef\]](#)
46. Bauer, Z.; Dominguez, A.; Cruz, E.; Gomez-Donoso, F.; Orts-Escolano, S.; Cazorla, M. Enhancing perception for the visually impaired with deep learning techniques and low-cost wearable sensors. *Pattern Recognit. Lett.* **2020**, *137*, 27–36. [\[CrossRef\]](#)
47. Elmannai, W.M.; Elleithy, K.M. A highly accurate and reliable data fusion framework for guiding the visually impaired. *IEEE Access* **2018**, *6*, 33029–33054. [\[CrossRef\]](#)
48. Wang, L.; Famouri, M.; Wong, A. DepthNet Nano: A Highly Compact Self-Normalizing Neural Network for Monocular Depth Estimation. *arXiv* **2020**, arXiv:2004.08008.

-
49. Bertheliet, A.; Chateau, T.; Duffner, S.; Garcia, C.; Blanc, C. Deep Model Compression and Architecture Optimization for Embedded Systems: A Survey. *J. Signal Process. Syst.* **2021**, *93*, 863–878. [[CrossRef](#)]
 50. Luo, Y.; Yin, L.; Bai, W.; Mao, K. An Appraisal of Incremental Learning Methods. *Entropy* **2020**, *22*, 1190. [[CrossRef](#)]
 51. Ji, Y.; Zhang, H.; Zhang, Z.; Liu, M. CNN-based encoder-decoder networks for salient object detection: A comprehensive review and recent advances. *Inf. Sci.* **2021**, *546*, 835–857. [[CrossRef](#)]
 52. Seng, K.P.; Lee, P.J.; Ang, L.M. Embedded Intelligence on FPGA: Survey, Applications and Challenges. *Electronics* **2021**, *10*, 895. [[CrossRef](#)]
 53. Moolchandani, D.; Kumar, A.; Sarangi, S.R. Accelerating CNN Inference on ASICs: A Survey. *J. Syst. Archit.* **2021**, *113*, 101887. [[CrossRef](#)]
 54. Ang, L.M.; Seng, K.P. GPU-Based Embedded Intelligence Architectures and Applications. *Electronics* **2021**, *10*, 952. [[CrossRef](#)]
 55. Chang, I.; Castillo, J.; Montes, H. Technology-Based Social Innovation: Smart City Inclusive System for Hearing Impairment and Visual Disability Citizens. *Sensors* **2022**, *22*, 848. [[CrossRef](#)]