

Article

Image Recognition-Based Architecture to Enhance Inclusive Mobility of Visually Impaired People in Smart and Urban Environments

Sara Paiva ^{1,*}, António Amaral ^{2,3}, Joana Gonçalves ¹, Rui Lima ¹ and Luis Barreto ¹¹ ADiT-LAB, Instituto Politécnico de Viana do Castelo, 4900-367 Viana do Castelo, Portugal² Departamento de Engenharia Mecânica, ISEP Politécnico do Porto, 4249-015 Porto, Portugal³ INESC TEC—Institute for Systems and Computer Engineering, Technology and Science, 4200-465 Porto, Portugal

* Correspondence: sara.paiva@estg.ipv.pt

Abstract: The demographic growth that we have witnessed in recent years, which is expected to increase in the years to come, raises emerging challenges worldwide regarding urban mobility, both in transport and pedestrian movement. The sustainable development of cities is also intrinsically linked to urban planning and mobility strategies. The tasks of navigation and orientation in cities are something that we resort to today with great frequency, especially in unknown cities and places. Current navigation solutions refer to the precision aspect as a big challenge, especially between buildings in city centers. In this paper, we focus on the segment of visually impaired people and how they can obtain information about where they are when, for some reason, they have lost their orientation. Of course, the challenges are different and much more challenging in this situation and with this population segment. GPS, a technique widely used for navigation in outdoor environments, does not have the precision we need or the most beneficial type of content because the information that a visually impaired person needs when lost is not the name of the street or the coordinates but a reference point. Therefore, this paper includes the proposal of a conceptual architecture for outdoor positioning of visually impaired people using the Landmark Positioning approach.

Keywords: visually impaired people; urban mobility; image processing; urban positioning; landmark techniques



Citation: Paiva, S.; Amaral, A.; Gonçalves, J.; Lima, R.; Barreto, L. Image Recognition-Based Architecture to Enhance Inclusive Mobility of Visually Impaired People in Smart and Urban Environments. *Sustainability* **2022**, *14*, 11567. <https://doi.org/10.3390/su141811567>

Academic Editors: Dhananjay Singh, Paulo J. Sequeira Gonçalves, Pradeep Kumar Singh, Pradip Sharma and Pao-Ann Hsiung

Received: 29 July 2022

Accepted: 10 September 2022

Published: 15 September 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

In recent decades, we have witnessed an exponential demographic growth of cities, which substantially increases the difficulties of urban planning in terms of mobility, which presents itself as an emerging challenge in all countries [1]. Whether in terms of transport or pedestrian navigation, mobility has taken on a new dimension and represents a problem to which many researchers respond, focusing on the sustainable development of cities. Infrastructures and city design are constantly changing to respond to the daily displacement needs of millions of people from the outskirts to city centers [2]. Sustainable development will be difficult to achieve with a large number of private vehicles and the current public transport network [3]. The mobility as a service (MaaS) paradigm emerges as an innovative concept that intends to be a strong ally towards sustainability, which is user-centric and combines different means of transport to minimize CO₂ emissions and, at the same time, maximize the use of sustainable transportation and/or means of active mobility.

Navigation within cities is a daily activity inherent to all citizens, performed several times throughout the day for commuting to work, schools, shopping, and leisure purposes, among others. For all those in a well-known city, the navigation turns out to be intuitive, something simple that does not pose any significant difficulties. However, when we are in an unknown city, the aspects of orientation and mobility become more relevant.

The orientation task in a city involves several elements, such as positioning, navigation, and direction. Urban spatial orientation is, in turn, intrinsically linked to the aspect of mobility, which can be described as the ability of someone to move in an environment, known or not, and which may contain several obstacles that are related to how citizens are informed about their position or how to navigate to the desired location [4]. Suppose several challenges arise for urban mobility in general terms, such as precision. In that case, additional questions are brought up when we talk about visually impaired people (VIP) that are immediately noticed by the different ways VIP use their senses to orient themselves [5]. The World Health Organization (WHO), in its report on “Visual Impairment 2010”, refers to approximately 320 thousand people per million with some visual impairment worldwide, and about 47 thousand people per million are considered blind [6]. In Portugal, the 2001 reports had values of 160 thousand visually impaired individuals [7], while the 2011 reports point to 900 thousand visually impaired individuals, 28 thousand of whom are blind [8]. Although the mobility challenges for this segment are varied, we will focus, throughout this paper, on the problem of positioning when a VIP loses track of where they are, which represents one of the main challenges they feel and experience in their daily routines. Receiving information that they are close to a particular store, near a given known intersection, etc., can be decisive and make all the difference for a VIP’s orientation [9,10].

One of the techniques that could be used would be the GPS combined with the information on Google Maps. Still, the information usually available does not have the necessary refinement and detail because the points of interests (POIs) inserted are typically thought of more globally and not with all the detail VIPs need. This paper proposes a conceptual framework for an outdoor positioning system for VIP that uses Landmark Positioning, an approach that fits inside the broad concept of Visual Positioning. The framework assumes a VIP with a smartphone camera that captures the image of surroundings as the VIP is walking and a backend server that can process those images and get information about its positioning, which is informed to the VIP. The approach mainly uses images with landmarks that represent locations in the cities and, for that reason, constitute helpful information for a VIP [11].

The paper is organized as follows. Firstly, Section 1 describes the problem context and its main implications for defining a proper solution. Section 2 presents a literature review on smart mobility and its main challenges, trends, and inclusiveness. Section 3 offers the related work regarding the VIP. Section 4 presents an architecture proposed for Outdoor Positioning for Visually Impaired People using Landmarks (OPIL) framework and model. Finally, in Section 5, the conclusions and future work are addressed.

2. Urban Mobility Challenges, Trends, and Inclusiveness

The advent of technologies, systems integration, and the era of Big Data and Artificial Intelligence, among others, gave birth to the concept of intelligent “things”. Smart cities are born from integrating all these concepts and the possibility of their application to different areas of intervention and domains. According to [12], there are six main axes/domains of interventions within a smart city: Government, Mobility, Environment, Economy, People, and Living. The world population has been growing significantly in recent years. In 1974, the world population numbered approximately 4 billion people, 6 billion in 1999, 7.8 billion in 2020, and an increasing trend according to United Nations projections, presented in Figure 1.

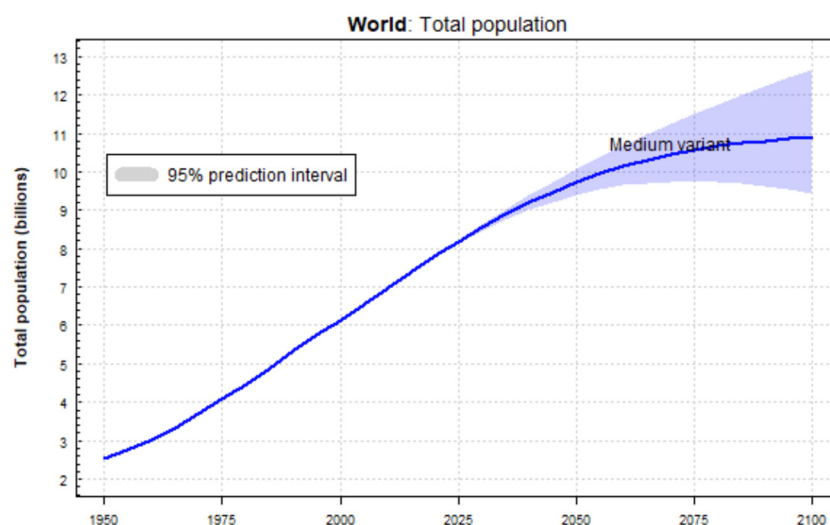


Figure 1. World population evolution [12].

This growth undoubtedly challenges cities at various levels in terms of urban mobility. The congestion problems arising from the high population density can be partly solved with efficient transport networks. Still, the current proposals do not stop there since the concept of smart mobility goes further and calls for more innovative solutions [13]. These include recommendations for sustainable mobility solutions, such as active transport that use environmentally friendly fuels and citizens' accountability [14,15]. Mobility must be seen as a path that allows for increasing sustainable development. The impact of smart mobility may, at first, seem to focus on solving what appears to be the most obvious problem: traffic congestion. But the impacted dimensions are much higher and encompass sustainability, economy, and living and, therefore, with a direct impact not only on citizens but also on governmental entities [13]. The "smart" character of mobility can be seen in various ways: public access to information and making it available in real-time to make travel planning more efficient, save time and money, and reduce CO₂ emissions [16,17].

Several challenges and barriers to providing inclusive and equitable mobility services amongst vulnerable user groups, such as people with disabilities and impairments, arise. These commonly include poverty [18] budget constraints, lack of accessibility, inadequate service integration, lack of correspondence between user needs and service provision, preliminary safety measures or unsafe infrastructure, and lack of access to or understanding of technology [19,20]. Another common problem is improper parking in places reserved for people with disabilities or even in public spaces in such a way that interfere with the mobility of people with reduced mobility. An example comes from the growing use of e-scooters in the urban context, which, in the absence of the proper legal and safety framework, translates into parking on public roads and sidewalks in ways that impair the pedestrian movement of people with disabilities, with considerable impact on blind and partially sighted people who encounter unexpected obstacles [21].

It is acknowledged that challenges to providing mobility services while affecting different users are also often related to an area's geography and economic characteristics and have additional prominence in other prioritized regions. Therefore, better communication with communities of need is necessary, particularly during planning processes. Taking a more inclusive, communicative approach to mobility planning will allow for better, more efficient mobility plans that more robustly serve the needs of the most vulnerable users [22].

Spatial and temporal coverage of services needs to be considered for all users. While often centered on 'peak hour' needs, shifting employment patterns, changing family demographics, and more dispersed activity-based lifestyles require that better mobility services coverage is needed in terms of both space and time to ensure that all users have equal access to opportunities that may not fit into traditional working/leisure hours.

Better service integration is needed [23] with more heavily dispersed populations, it and is unlikely that a one-seat trip provided by one operator will adequately serve the needs of all users. Integration, however, should be a multidimensional consideration, encompassing integration of services (including nontransport services), payment, and information, thus promoting the transition to a fully mobility as a service (MaaS) system [24–26].

A more effective information-sharing process should be developed to fully serve all users' needs. Information should be accurate and reliable, accessible from various platforms (including digital and paper) and in different formats (including, for example, images, text, and multiple languages), and secure to guarantee its consistency and confidence to all users [27]. New service provisions should be explored, including considerations beyond traditional mobility services or systems and incorporating walking, cycling, car-share, carpool, taxis, and other services [28]. Some resistance is expected as innovation pushes towards a shared transport model. A study presented in [29] shows that car-sharing service providers have shown the greatest resistance to this adaptation, contrary to e-bike-sharing services that are more open to this change.

Adequate financial and policy support is needed from the local and national governments. Namely, financial concerns are the primary concern regarding the cost of transport for users and the cost of providing service to vulnerable populations and areas. The cost was identified as a significant barrier to mobility services and transport use, indicating that more considerations should be made for ways to offset its impact.

The transportation system must always be planned and designed to ensure maximum inclusiveness for all members of society, especially those with disabilities, seniors, and children/youth, without limiting or restricting their mobility options [30], with the final purpose to make sure all means of transport are prepared to receive every citizen without any limitation or exclusion, thus promoting a more inclusive society.

Therefore, developing solutions for all these cases should be highly prioritized, namely for ensuring the embedding of new intelligent assistive technology conceived for easing the navigation and routing experience of blind and visually impaired people throughout their daily commuting safely and independently [31–33].

3. Related Work on VIP

Visual assistive technology can be divided into three main categories: vision enhancement, vision substitution, and vision replacement [34,35]. Due to the fast presence of sensors and their functions in almost every activity of society, these systems are becoming increasingly available in terms of applications and devices, providing different services, such as localization, detection, and avoidance. Sensors can help navigate and orient VIP and give them a sense of their external environment. Sensors can identify any surrounding object property and facilitate VIP's mobility tasks [36]. Vision replacement is the most complex category, as it is related to medical and technological issues. In such systems, sensor data will be sent to the brain or a specific nerve. Vision replacement and vision enhancement are comparable yet different. In the vision enhancement category, the process data that is sensed by a sensor will be displayed in a device. In the vision substitution category, the result of any data sensed by any sensor will not be displayed, and the output result will be acoustic, tactile, vibrating, or a combination. This makes the vision substitution systems very similar to intelligent transportation systems and mobile robotics.

This research work is focused on the visual substitution category, which can be further subdivided into three other categories: electronic travel aids (ETAs), electronic orientation aids (EOAs), and position locator devices (PLDs). Table 1 briefly presents each visual substitution category subcategory and its services.

Table 1. Visual substitution subcategories.

Subcategory Name	Description	Services
Electronic Travel Aid (ETAs)	Collected and sensed data about surrounding areas are sent to the user or a remote server via sensors, laser, or sonar.	Surrounding obstacles are identified. Texture and gaps on the surface can be provided. The distance between the user and an obstacle is determined. Great locations are identified. Self-orientation throughout an area is improved with obstacle avoidance information.
Electronic Orientation Aid (EOAs)	Guidelines and instructions about a path are given to the user through a device.	The best path for a particular user is determined. Clear direction and path signs are given by calculating the user's position and tracking the path.
Position Locator Devices (PLDs)	The user's location is identified, for example, using the global positioning system (GPS)	Guidance from one point to another point is given.

3.1. Navigation of Visually Impaired People

A study involving VIP and control groups aims to assess the ability to travel and learn routes and the consequent capacity for mobility in an urban environment [37]. The results show that VIP can learn complex routes efficiently when assisted and guided. They experience greater difficulties when they are expected to do so autonomously. In [38], a system for navigation in outdoor environments for VIP is proposed using a geographic information system (GIS) as a fundamental supporting system. Given the current position, the GIS allows the system to find a route to the desired destination, thus supporting the VIP in its movement along the route. Information about the user's positioning is obtained through a GPS receiver, a magnetic compass, and a gyrocompass. System communication with the VIP uses a keyboard and voice messages. In [39], the authors present an app developed to help VIP move autonomously, identifying obstacles that may exist along the route through the smartphone's camera. The Walking Stick app fulfills three objectives: (1) to identify obstacles, (2) to allow the app to run without having to open it explicitly (it is accessible through shaking movement), and (3) safety requirements by detecting falls from the smartphone and emitting a sound so it can be found. The application considers development requirements to avoid rapid battery consumption. In [40], the X-EYE app is presented to promote VIP independent mobility, using a wearable camera which reduces the cost of the solution. The application allows the detection of obstacles, people recognition, tracking and sharing of location, SMS reader, and a translation system. The authors indicate some limitations of the solution, such as the partial appearance of images and rapid changes in the background or lighting conditions. In [41], the authors introduce a navigation system for VIP with the main objective of addressing obstacle detection, which is useful for navigation in outdoor and indoor environments. The system uses ultrasonic sensors that, together with a buzzer and a vibrator, can be placed on a jacket or other garment to be easily included in the daily life of the VIP and, therefore, more easily be accepted as an accessory that is easily transported. In [42], a system for indoor navigation using the computer vision approach is presented. The user must use a camera in their hand while navigating the environment prepared with fiducial markers augmented with audio information. The system has two modes of operation. In the first mode—free mode—the user freely navigates through the pre-prepared environment and is informed of the environment around them through audio and the detected markers. In the second mode—guided mode—the user navigates from the current location to the desired destination using the shortest path (using the Dijkstra algorithm). In [43], a support system for indoor navigation is presented, consisting of a navigation system and a map information system installed on a cane. The system works on the assumption that there is a colored navigation line on the pavement. The cane has a color sensor that, when leaning against

the line on the floor, manages to inform the user through a vibration system that they are moving over the line. A one-chip microprocessor controls the color recognition system. The map information system uses RFID tags along the line on the floor and RFID receivers placed on the cane, allowing a navigation system from a given source to a given destination to be made available. In [44], the authors present a navigation system that detects obstacles and guides VIPs through the most appropriate paths. Obstacles are detected using an infrared system, and feedback is sent to the user through vibration or sound to inform them of their position. Aware of the limitation of canes regarding protection at the head level, the system has incorporated a sensor in a cap (and, therefore, at the level of the user's head area) that allows information about obstacles and navigation. In [45], the authors present a navigation system that aims to fill the gaps in the GPS (only valid in outdoor environments). Its main focus is to provide the VIP with accurate information about its spatial positioning. The proposed system can be used in an indoor or outdoor environment and uses the camera carried by the user and computer vision techniques. As the user follows a previously memorized path, the camera's image is recognized to calculate its direction and positioning precisely. In [46], the authors propose a system that intends to complement the use of the cane when traveling in indoor environments. The system consists of a geographic information system related to a building that contains information about visual markers that allow the system to create a mental map of the environment and later allow navigation in space. In [47], a system is proposed to help VIPs move around outdoors with the possibility of detecting obstacles. The system, which aims to combat the limitations of other solutions in terms of cost, dependency, and usability, makes use of a mobile-based camera vision system to allow great usability in the use of the system in places with which the VIP is not familiar, such as parks, streets, among other places. Deep learning algorithms are used to recognize objects and interact with a mobile application which is also part of the system. In [48], a system is proposed based on a mobile application that stores important information for the VIP during the route, making their journey easier and safer simultaneously. This concept goes against generalized navigation, favoring a personalized recommendation based on the information each person prefers to receive, based on the most favorable input for them to feel located and safe. The system uses voice feedback consisting of multi-sensory clues combined with microlocation technology. In [49], the authors propose an android app to assist in indoor navigation based on transmitting information to the user via audio messages and the existence of QR codes. The system takes predesigned paths for VIPs along which QR codes are placed to infer the position and recommend the quickest and shortest route to an intended destination. The system allows detecting deviations in the route from what is expected and informs the user of the deviation that has occurred. In [50], the authors propose a navigation system based on simultaneous location and mapping (SLAM) to help VIPs navigate indoor environments. This system integrates sensors and feedback devices, including an RGB-D sensor and an inertial measurement unit (IMU) on the waist, a camera on the user's head, and a microphone together with a headset. The authors developed a visual odometry algorithm based on RGB-D data to estimate the user's position and orientation, refining the orientation using the IMU. The head-mounted camera recognizes port numbers, and the RGB-D sensor detects important landmarks, such as corners and corridors. In [51], the authors describe an integrated navigation system with RFID and GPS, entitled Smart-Robot. The system uses location based on RFID and GPS operating, therefore, both in indoor and outdoor environments, respectively. The system uses a portable device composed of an RFID reader, GPS, and an analog compass to obtain location and orientation information. During navigation, the system avoids obstacles using ultrasonic and infrared sensors. User feedback is made through vibration in a glove and loudspeakers. In [52], the authors propose a VIP system that can operate indoors and outdoors. The system uses an ultrasonic sensor for navigation and voice outputs to prevent obstacles. Table 2 summarizes the challenges and approaches.

Table 2. Summary of challenges and approaches of navigation systems for visually impaired people.

Reference	Challenge	Approach
[38]	Outdoor navigation	SIG, GPS receiver, magnetic compass, and gyrocompass
[39]		Smartphone camera, accelerometer
[47]		Deep learning, camera vision-based system
[48]		Voice feedback, multi-sensory clues, microlocation technology
[44]	Obstacle detection and navigation	The infrared sensor in a cap for head protection
[40]	Obstacle detection	Wearable camera connected to the smartphone
[41]		Ultrasonic sensors, vibrato rand buzzer
[42]	Indoor navigation	Webcam, Dijkstra Algorithm, Computer Vision
[50]		SLAM, inertial sensors, feedback devices
[43]		Walking stick, color sensor, colored floor, RFID tags
[49]		QR codes
[46]		GIS, visual markers
[45]	Indoor and outdoor navigation	Camera, computer vision
[51]		RFID, GPS, ultrasonic and infrared sensors
[52]		Ultrasonic sensor

3.2. Outdoor Positioning Using Landmarks

In [53], the authors present an indoor and outdoor navigation system based on landmarks and artificial intelligence. Template matching techniques are used to find patterns to estimate the distance to the landmark. Several techniques are used, such as high boost filtering and histogram equalization, to improve the images for the processing process. The dynamic exclusion heuristic converts the estimated distance into the actual space [54] and presents a system that uses QR codes displayed in registered landmarks. The system thus allows the user, based on QR codes, to identify their position and, on the other hand, to have access to the navigation functionality based on landmarks. The system includes images of each landmark, allowing users to navigate freely using these images without needing guided navigation. In [55], the authors present a system for localization in outdoor environments using landmarks based on deep learning. The proposed location method is based on the faster regional convolutional neural network (Faster R-CNN), which is implemented for the detection of landmarks in images, and the feed-forward neural network (FFNN), which is used to retrieve the location coordinates and compass directions for the device implemented in the real world based on landmarks detected from R-CNN. In [56], the authors report on how to drive a mobile robot using natural landmarks such as trees and plants on Japan's University's outdoor campus. An essential component of the system is the acquisition of natural landmarks being proposed for the landmark agent (LmA). This component has three states: SLEEP, WAKE UP, and TRACK. The SLEEP status means that there is no candidate for a landmark. WAKE UP means that a candidate for a reference point has been detected, and the agent begins to locate it, and finally, the TRACK state means that the object is repeatedly detected two or more times. When you make the transition order SLEEP-WAKE UP-TRACK, the reference point is detected and recorded on the LmA map. Another component of the system is autonomous navigation using the acquired natural landmarks. The navigation strategy uses the perceived route map (PRM) generated by automatically acquiring natural landmarks through the teaching of routes. In [57], the authors propose an approach to improve the information provided by most navigation systems that are based on information based on street names and step-by-step instructions. The authors suggest improving this information based on landmarks collected

and inserted in the recommendation system to provide information that the user can more easily identify.

3.3. *The Role of Machine Learning in Image Recognition*

In this section, we highlight some of the research being carried out regarding the use of machine learning algorithms and techniques in the field of image recognition, which has been a field of research for the past two decades. These algorithms can recognize landmarks and help VIP gain positional awareness within an urban environment. Those landmark images can be acquired in real-time from a smartphone camera or other mobile devices. The identification can be read out loud with the help of accessibility tools such as Google TalkBack.

To solve this problem, Yeh et al. propose an image search approach in web search engines [58]. In this approach, location recognition is performed not only through images but also through text. The process includes three phases: image search on Google to identify similar images, collection of keywords from pages where similar images are found, and a new search on Google with the keywords obtained to obtain similar images. The technique used is a hybrid search for images and keywords. Li and Kosecká present a method of probabilistic location recognition for visual navigation systems [59]. The proposed system has two phases: in the first phase, the environment is divided into several locations categorized by key points; in the second phase, the key points are used to estimate the most likely location. This process demonstrates a high localization rate using only 10% of the characteristics of an image. To perform location recognition in large datasets, Schindler et al. present a method that uses a generalized version of the vocabulary tree search algorithm [60]. This method allows using datasets with ten times the number of images. The algorithm recognized 20 km of the urban landscape, storing 100 million characteristics in a vocabulary tree. Gallagher et al. use a K-nearest-neighbor algorithm in a dataset with more than 1 million georeferenced photographs. This method proves to be more effective in determining the location of an image than using only the visual information of an image [61]. One of the biggest challenges for visual localization is the search for images similar to the original in a database with georeferenced images. For the visual location to be carried out, the objects must be identified in different poses, sizes, and background environments. Faced with this challenge, Schroth et al. use a “bag of words” quantization and indexing method based on a tree search [62]. Cao and Snavely propose a method of representing images from a database in a graph [63]. The method selects a group of subgraphs and creates a function that calculates the distance between each group. Each image is classified according to this distance function. A probabilistic method to increase the diversity of these classified images is also presented. The paper states that this method improves the performance of location recognition in image datasets compared to “bag of words” models. Wan et al. present research about using neural networks to find similar images based on the content. This work investigates a convolutional neural network with several changes in its definitions. Convolutional neural networks are very effective in recognizing images [64]. Liu et al. use a semi-supervised machine learning model using a convolutional neural network. The image dataset was created with images from surveillance cameras and smartphones and took into account the changes that variations in the weather can cause in the panorama of an image [65]. Zhou et al. present research on the methods used to obtain images based on the content. In this study [66], searches for images are grouped according to the method used: by keyword, by example image, by sketch, by color layout, and by concept layout. Zhou et al. highlight two pioneering studies carried out around the 2000s. The first is the introduction of local visual invariant characteristics capable of capturing the invariance to the rotation and resizing of images. The second work highlighted was the introduction of the “bag of words” model capable of creating a compact representation of images based on the number of local characteristics present. It has been revealed that recently the convolutional neural networks have shown an above-average performance compared to the other methods. In this work, the authors explain that

convolutional neural networks are a variant of deep neural networks and are usually used in recognition and image acquisition. In [67], Tzelepi and Tefas propose a convolutional neural network model to obtain image characteristics representation. Subsequently, the neuronal network is again trained to create more effective image descriptors. In this method, three different approaches are proposed: in the case that no information is provided, in the case the classification of the images is provided, and in the case, user feedback is provided. The proposed method outperforms all other unsupervised methods used in the comparison. To reduce the time spent obtaining images based on the content, Saritha et al. present a deep learning method [68] capable of extracting the properties of an image and classifying them according to those properties. In this work, properties such as color histogram, image edges, and texture are taken into account. The extracted properties are saved in a small file for each image, where files with similar properties must belong to similar images. The images' properties are extracted using a convolutional neural network.

3.4. Convolutional Neural Networks for Image Recognition

With recent advances in machine learning, the most used solution for image recognition is neural networks, specifically convolutional neural networks. We next explain how artificial neural networks work and their more specific subgroup of convolutional neural networks.

One of the types of machine learning algorithms, neural networks, demonstrates a remarkable ability to learn complex relationships between objects belonging to a large amount of data. Neural networks are also valuable for the analysis of unstructured data. Examples of unstructured data are text, audio files, and images [69]. However, neural networks can be difficult to program, train, and evaluate. There are several frameworks capable of facilitating this work for a programmer. Examples of these tools in Python are TensorFlow to deal with neural networks in a more particular way or Keras, which works on top of TensorFlow, offering a higher level of abstraction. Neural networks are composed of an input layer that receives the input data, an output layer that classifies the data, and one intermediate layer that processes the information. A simple neural network has only one middle layer, whereas a deep neural network has two or more layers. The unity of neural networks is called a neuron, and the connection between each neuron is called a synapse (Figure 2) [70].

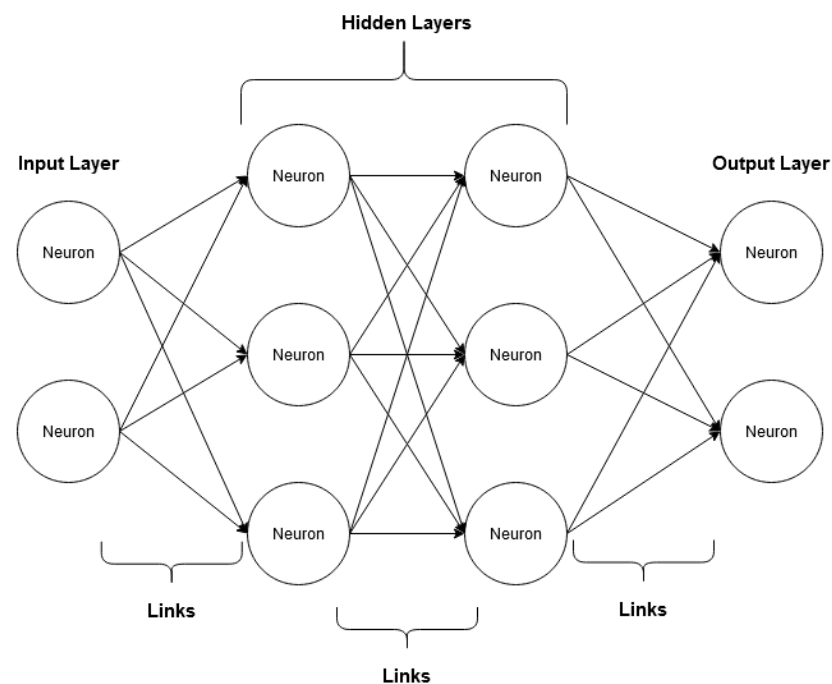


Figure 2. Example of the operation of a deep neural network (adapted from [67]).

Convolutional neural networks are the most used neural networks in image processing. Like neural networks, they are also composed of “neurons” that receive input and process them using activation functions. However, unlike the layers of neural networks, convolutional neural networks use images as input, not data vectors [69]. Convolutional neural networks have several applications: they can identify pictures, group them by their similarity or recognize objects and places portrayed in the image. The components of a convolutional neural network are as follows [69]:

- Data input
- Convolutional layer;
- Pooling layer;
- Flattening.

Image input is the first component of a convolutional neural network. In this component, the image is converted into a matrix (Figure 3), where each value represents the color of each pixel in the image [69].

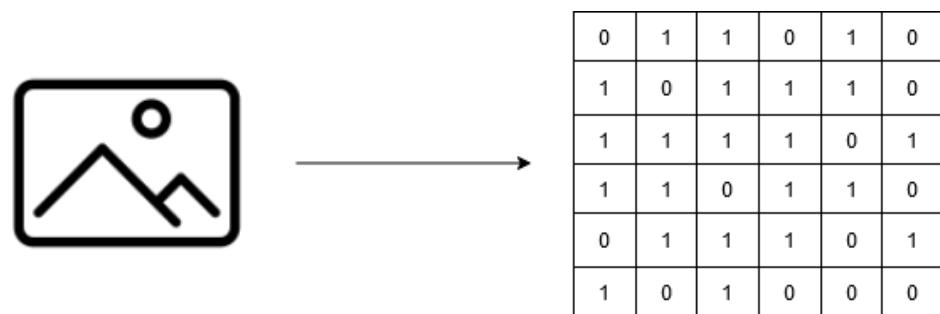


Figure 3. Example of inputting images into a convolutional neural network (adapted from [66]).

The convolutional layer is where image processing begins and consists of two parts:

- Filter;
- Characteristics map.

The filter is a matrix or pattern combined with the original image’s matrix to obtain the feature map. The feature map is the reduced image produced by the seizure between the original image and the filter (Figure 4). In a real convolutional neural network, several filters are used to create multiple feature maps for each image [69].

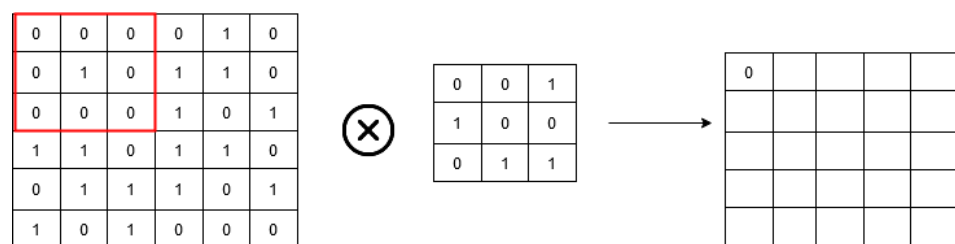


Figure 4. Example of the convolutional layer in a convolutional neural network (adapted from [66]).

The pooling layer helps the network to “ignore” the least important data in an image, further reducing the image and preserving the most critical features (Figure 5). This component consists of functions such as “max pooling”, “min pooling”, and “average pooling”.

Flattening is the component where the image is arranged to serve as an input to a shared neural network. In this part of the process, the matrix is “flattened” and converted into a single column (Figure 6).

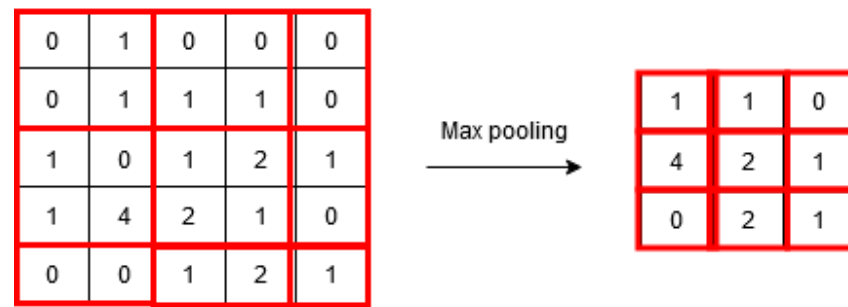


Figure 5. Example of the pooling layer in a convolutional neural network (adapted from [66]).

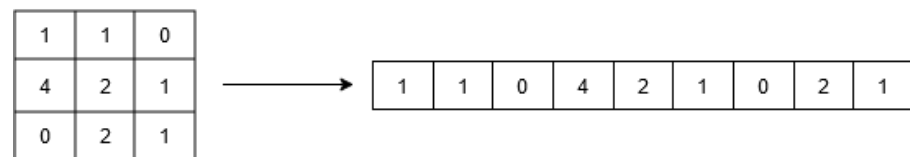


Figure 6. Example of flattening in a convolutional neural network (adapted from [66]).

4. Proposed OPIL Framework

4.1. Overview and Architecture

The architecture proposed for Outdoor Positioning for Visually Impaired People using Landmarks (OPIL), presented in Figure 7, aims to allow a VIP to obtain information on the location where they are positioned, one of the scenarios that represent a challenge for this segment of people when they move autonomously in a city. The architecture comprises two main components: a mobile application that the VIP must use and a backend server that does the image recognition processing and obtains the information about a helpful description that can be transmitted to the VIP so they can orient themselves.

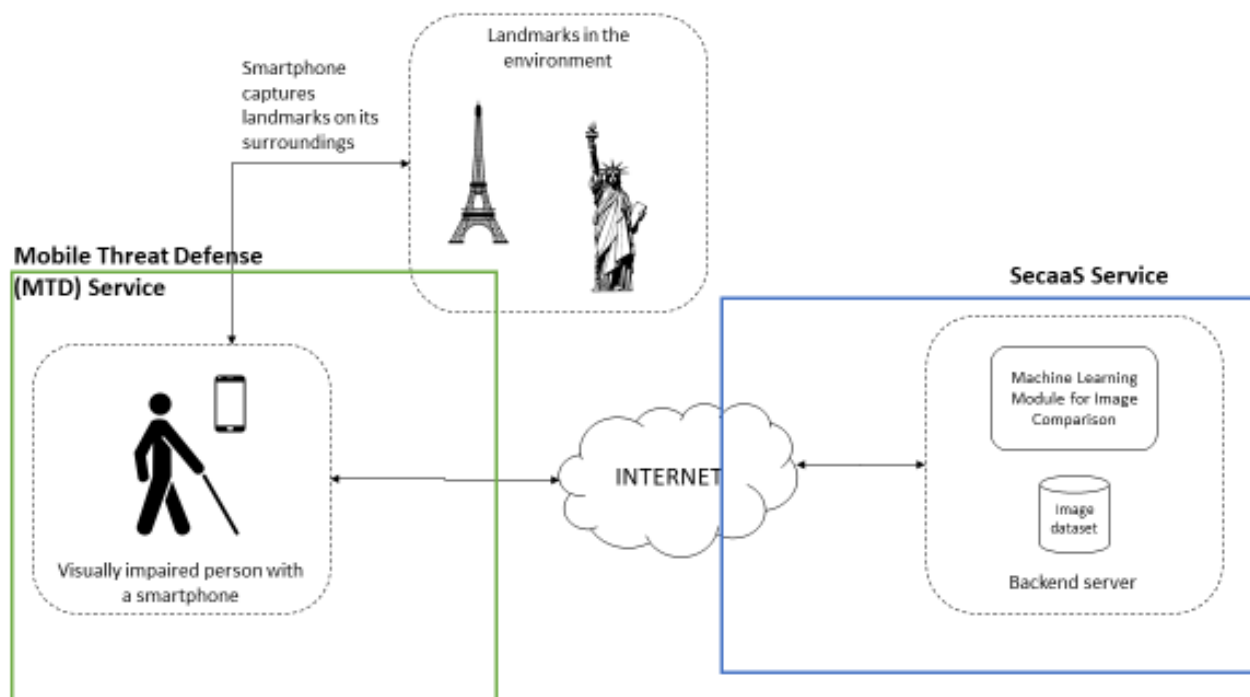


Figure 7. Core components of the proposed framework.

In the architecture proposed for OPIL, it is assumed that each VIP has a mobile phone and uses an app that is one of the components of the system. This app allows the capture of the image through the mobile phone's camera, which helps collect information about

where the user is. This app must consider the battery consumption and the processing power it takes from the mobile phone. The user must keep the mobile phone fixed for some time, and only when they get the result sent by the backend as a result of processing its positioning (in some cases, the location may not be identified) can they move the mobile phone to another location. Whenever the backend receives an image, a comparison is made with all the images in a prepopulated and georeferenced image database. If the image is successfully identified, a description is returned by the algorithm in such a way that it is helpful for the VIP.

4.2. Components

4.2.1. Mobile Application

The mobile application represents the process's starting and ending point, as seen in Figure 8. The VIP triggers the process when the user opens the app and points the camera to the place it is facing, allowing the image collection. The image capture module (ICM) is responsible for presenting the UI that captures the image, collecting it, and sending it to the backend server communication module (BSCM), which is responsible for communicating over the Internet with the backend server that will process the image (this process will be described in more detail in the next section). As soon as there is a response from the backend server, the callback method module (CMM) is triggered automatically to be able to process the response and send it to the notification module (NM), which is responsible for the interaction with the VIP to inform the user of the result about their positioning.

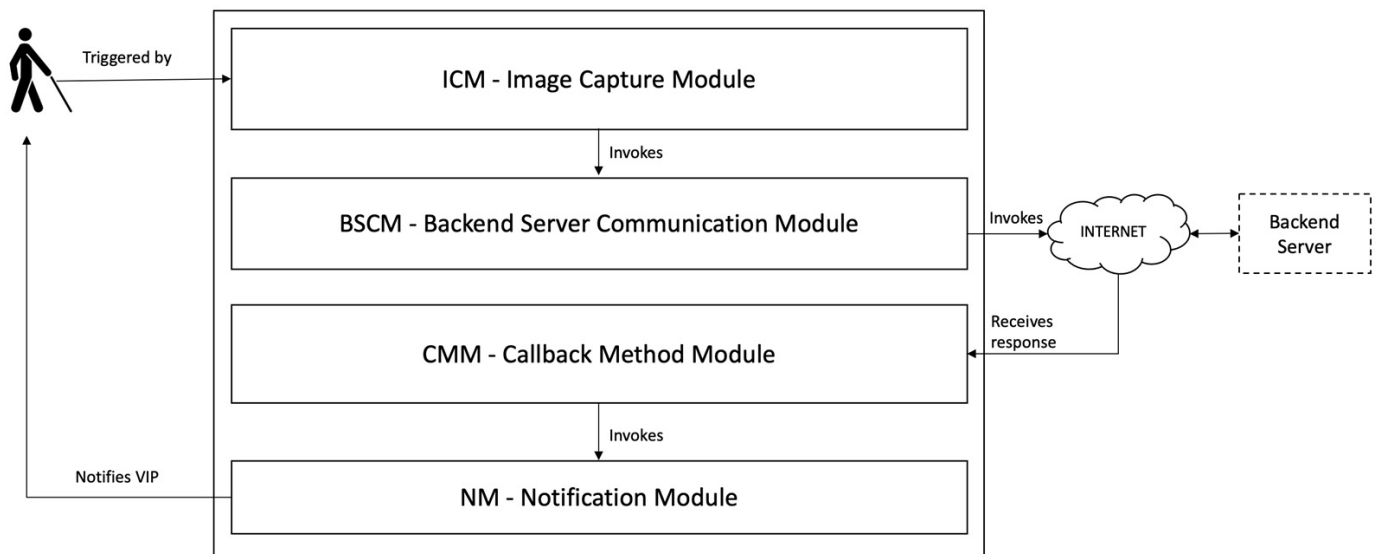


Figure 8. The architecture of the mobile app component.

4.2.2. Backend Server and Proposed Algorithm

The proposed algorithm for image recognition is divided into three phases (Figure 9). In the initial phase, datasets are obtained where all points of interest necessary for the project (from different angles, light levels, weather conditions, etc.) are photographed and identified. The second phase trains the model using a convolutional neural network. In this phase, the identification attributes are the images of the photographed objects, and the target attribute is the identification of these objects. The final phase corresponds to the trained model's use in identifying new images of the objects.

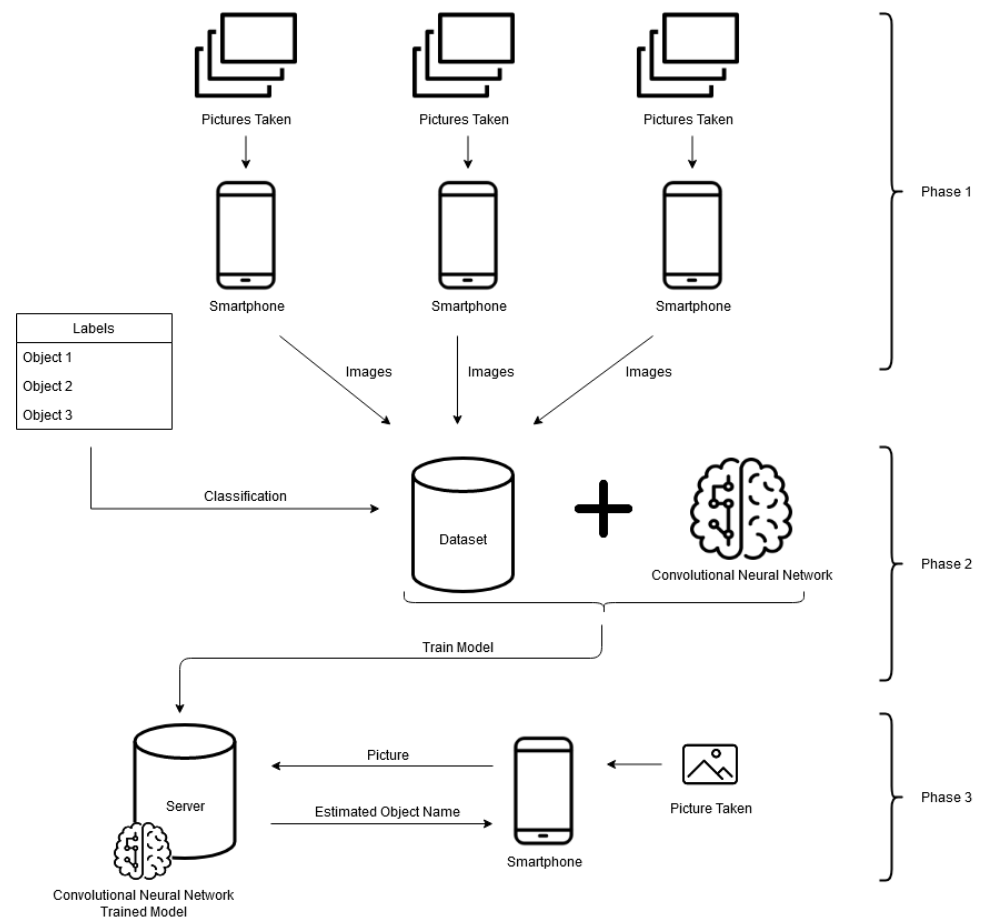


Figure 9. The 3 phases of the algorithm for the recognition of points of interest.

The first phase corresponds to data collection, where images are gathered. Later, the identification model is trained using those same pictures. The collection of images must have an equal number of images with points of interest and those without points of interest so that the model's training is not biased. Images without points of interest should be similar to images with points of interest. All images must be changed to have the exact dimensions. The image distribution should be about 80% training and 20% test images. Changing image properties such as opacity, converting them to black and white, or reducing the definition may be necessary to reduce model training time at the expense of accuracy. An image marking tool may be required to identify and select points of interest in the photos manually. Some examples of these tools are LabelImg (<https://github.com/tzutalin/labelImg> (accessed on 1 July 2022)) and OpenLabeling (<https://github.com/Cartucho/OpenLabeling> (accessed on 1 July 2022)).

The second phase corresponds to the training of the convolutional neural network. In this phase, the model is trained to differentiate the landmarks based on their geometric-functional typologies, such as punctual elements (e.g., monuments, stores, schools, city hall), linear elements (e.g., rivers, roads, sidewalks), or areal elements (e.g., squares, green parks). When creating the model, there are some general guidelines to follow. The use of convolution and pooling should be alternated, with convolution initially being used and pooling at the end before connecting to the neural network (e.g., convolution-pooling-convolution-pooling-neural network-output). During convolution, there is little need to use a filter more significant than a three-by-three matrix. It is advisable to use max pooling along the neural network and average pooling at the end before connecting to the neural network layer. Generally, the more layers a neural network has, the greater its accuracy and execution time. Preprocessing of the images should be done only if necessary to increase the accuracy or speed of the model. Creating new pictures with existing images (data

augmentation) almost always helps improve a model's accuracy. The number of nodes in the first intermediate layer of the neural network should be half of the nodes in the input layer, and in the second layer, half of those in the first layer. The number of nodes in the neural network's output layer must equal the number of classes identified by the network. The number of nodes in the middle layers of a neural network must follow a geometric progression (2, 4, 8, 16, 32, ...).

The third phase corresponds to the use of the model. A smartphone application must be created to obtain images from the camera and send the data to an external API. If access to the Internet via mobile data is impossible, then the pretrained model must be integrated into the application. To be able to use the android model, the pretrained model must be in the ".pb" format (TensorFlow) and must be converted to this format if it is in the ".h5" format (keras). Suppose the solution is used in more than one geographical area. In that case, it is advisable to add the GPS coordinates to the characteristics matrix of the images that serve as input data to the convolutional neural network. Suppose the solution is used in more than one geographical area. In that case, the smartphone application must also obtain the GPS position when taking images, which must be used to forecast the point of interest.

The model's usefulness will depend on the imageability of the urban environment. This is because, and according to Lynch, an imageable city has, in principle, a better degree of legibility [71]. Lynch also states, in his 1960 article [71], that imageability is the "quality in a physical object which gives it a high probability of evoking a strong image in any given observer. It is that shape, color, or arrangement which facilitates the making of vividly identified, powerfully structured, highly useful mental images of the environment".

4.3. Framework Security Aspects

The preferred framework is supported by mobile devices and other digital technologies for its operation, i.e., cloud computing. The technological development and the sophistication of such devices usually reveal unexpected and underestimated security vulnerabilities. As there is an increased reliance on technology, it is also necessary to increase efforts to protect the technological resources and infrastructure, ensuring that they operate safely and uninterruptedly.

As the amount of data collected increases, new systems and applications are integrated, the number of users increases, and many that lack computer skills, the potential occurrence of information security incidents, breaches, and threats also increases. Thus, it is almost mighty to have adequate security services and systems, as well as good practices, to manage all the interactions within the framework; being fundamental to have mechanisms that ensure the confidentiality, integrity, and availability (CIA) properties of the framework operation and information [72].

As a smartphone supports the proposed framework for capturing landmarks and surroundings, on the Internet for bidirectional communication, and in cloud computing for the backend server operation and image comparison, the CIA properties must also be guaranteed. To guarantee the CIA properties in the proposed framework, a cybersecurity layer in the form of a mobile threat defense (MTD) system must be considered. MTD can be viewed as sophisticated, dynamic protection against cyber threats targeted against mobile devices, and the safety is applied to devices, networks, and applications [73]. This will maintain the security information of the data generated and received by the smartphone (or another mobile device), protecting against malicious applications, network attacks, and device vulnerabilities [74].

For the backend server, a cloud service system can be used. As the information and data, the server manages, and stores are compassionate. If the CIA properties are not met, this can cause a high-level impact on the users and organizations that use the proposed framework, possibly causing a severe or catastrophic adverse effect on organizational operations, organizational assets, and individual users. It is suggested to use the security as a service (SecaaS) provided by the cloud provider. This is a package of security services offered by a service provider that offloads much of the security responsibility from an en-

terprise to the security service provider. The services typically provided are authentication, antivirus, antimalware/spyware, intrusion detection, and security event management [75]. SecaaS is considered part of any cloud provider's software as a service (SaaS). A segment of the SaaS offering of a CP. In a significantly more straightforward manner, the Cloud Security Alliance defines SecaaS as the provision of security applications and services via the cloud either to cloud-based infrastructure and software or from the cloud to the customers' on-premises systems [76]. The Cloud Security Alliance has identified the following SecaaS categories of service: Identity and access management; and data loss prevention. The combination of the previously referred systems and services will improve the level of usage confidence of the framework and can be viewed as an adequate strategy to secure both the end user, that relies on the utilization of mobile devices, and server-side data and information (at rest and being used).

Cybersecurity is critical nowadays because safety, security, and trust are inseparable. The relative insecurity of almost every connected thing is top of mind due to the heightened awareness of the impact of cyberattacks and breaches, highlighted in the media and popular culture depictions of nefarious hacking exploits. This makes achieving security from the outset extremely critical to avoid becoming any computing service and application highlighted by this media attention [77].

5. Conclusions and Future Work

Urban mobility is currently in a massive evolving period due to the digitalization of society, the effects of the COVID-19 pandemic, the global economic dynamics, and the ongoing pattern of people concentrated in the urban area. The mobility challenges are one of the most visible and notable elements that profoundly affect urban metabolism and its decurrent effects. The solutions being developed towards ensuring more inclusive and universal access to the mobility system are essential to guarantee that visually impaired people could benefit from this reality and substantially improve their daily commuting activities and overall life in the urban ecosystem.

The presented solution addresses a specific problem related to the difficulty of orientation by a blind or visually impaired person when they lose their direction in an urban environment. Making use of the massive use of mobile phones, the solution assumes the existence of one of these devices to capture an image of the environment that surrounds the user to be able, through image recognition techniques, to obtain the user's location, by comparison to a georeferenced image data and a trained model.

Some factors need to be addressed such as reducing the time to acquire data and the performance of the increased number of users that could use such a solution simultaneously in a city. These factors are decisive for the success of the solution.

This solution aims to contribute to more inclusive mobility in urban environments. The main limitations of the solution are mainly in terms of the scalability for large areas, although the potential for historic centers or smaller cities is promising. In future terms, we intend to create a functional prototype that allows the execution of tests in a controlled environment. The evolution to an approach that contemplates the cooperation of citizens, through crowdsourcing, is a possible way to overcome the scalability problems that the solution may have now.

Author Contributions: Formal analysis, A.A.; Investigation, S.P., A.A., J.G., R.L. and L.B.; Methodology, A.A.; Supervision, S.P.; Writing—original draft, S.P., Rui Lima and Luis Barreto; Writing—review & editing, S.P. and A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work is funded by the European Regional Development Fund (ERDF) through the Regional Operational Program North 2020, within the scope of Project TECH—Technology, Environment, Creativity and Health, Norte-01-0145-FEDER-000043.

Data Availability Statement: Not applicable.

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

References

- Giduthuri, V.K. Sustainable Urban Mobility: Challenges, Initiatives and Planning. *Curr. Urban Stud.* **2015**, *3*, 261–265. [CrossRef]
- Bezbradica, M.; Ruskin, H. Understanding Urban Mobility and Pedestrian Movement. In *Smart Urban Development*; IntechOpen: London, UK, 2019. [CrossRef]
- Esztergár-Kiss, D.; Mátrai, T.; Aba, A. MaaS framework realization as a pilot demonstration in Budapest. In Proceedings of the 2021 7th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), Heraklion, Greece, 16–17 June 2021; pp. 1–5. [CrossRef]
- Riazi, A.; Riazi, F.; Yoosfi, R.; Bahmeei, F. Outdoor difficulties experienced by a group of visually impaired Iranian people. *J. Curr. Ophthalmol.* **2016**, *28*, 85–90. [CrossRef] [PubMed]
- Lakde, C.K.; Prasad, D.P.S. Review Paper on Navigation System for Visually Impaired People. *Int. J. Adv. Res. Comput. Commun. Eng.* **2015**, *4*, 166–168. [CrossRef]
- WHO. Blindness and Visual Impairment. Available online: <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment> (accessed on 29 July 2022).
- ANACOM. ACAPO—Associação dos Cegos e Amblíopes de Portugal. Available online: <https://www.anacom.pt/render.jsp?categoryId=36666> (accessed on 29 July 2022).
- ACAPO. Deficiência Visual. Available online: <http://www.acapo.pt/deficiencia-visual/perguntas-e-respostas/deficiencia-visual#quantas-pessoas-com-deficiencia-visual-existem-em-portugal-202> (accessed on 29 July 2022).
- Brito, D.; Viana, T.; Sousa, D.; Lourenço, A.; Paiva, S. A mobile solution to help visually impaired people in public transports and in pedestrian walks. *Int. J. Sustain. Dev. Plan.* **2018**, *13*, 281–293. [CrossRef]
- Paiva, S.; Gupta, N. Technologies and Systems to Improve Mobility of Visually Impaired People: A State of the Art. In *EAI/Springer Innovations in Communication and Computing*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 105–123. [CrossRef]
- Heiniz, P.; Krempels, K.H.; Terwelp, C.; Wuller, S. Landmark-based navigation in complex buildings. In Proceedings of the 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Sydney, NSW, Australia, 13–15 November 2012. [CrossRef]
- Giffinger, R.; Fertner, C.; Kramar, H.; Pichler-Milanovic, N.Y.; Meijers, E. *Smart Cities Ranking of European Medium-Sized Cities*; Centre of Regional Science, Universidad Tecnológica de Viena: Vienna, Austria, 2007.
- Arce-Ruiz, R.; Baucells, N.; Moreno Alonso, C. Smart Mobility in Smart Cities. In Proceedings of the XII Congreso de Ingeniería del Transporte (CIT 2016), Valencia, Spain, 7–9 June 2016. [CrossRef]
- Van Audenhove, F.; Dauby, L.; Korniiuchuk, O.; Poubaix, J. *Future of Urban Mobility 2.0*. Arthur D. Little; Future Lab. International Association for Public Transport (UITP): Brussels, Belgium, 2014.
- Neirotti, P. Current trends in Smart City initiatives: Some stylised facts. *Cities* **2012**, *38*, 25–36. [CrossRef]
- Manville, C.; Cochrane, G.; Cave, J.; Millard, J.; Pederson, J.; Thaarup, R.; Liebe, A.; Wissner, W.M.; Massink, W.R.; Kotterink, B. *Mapping Smart Cities in the EU*; Department of Economic and Scientific Policy: Luxembourg, 2014.
- Carneiro, D.; Amaral, A.; Carvalho, M.; Barreto, L. An Anthropocentric and Enhanced Predictive Approach to Smart City Management. *Smart Cities* **2021**, *4*, 1366–1390. [CrossRef]
- Groth, S. Multimodal divide: Reproduction of transport poverty in smart mobility trends. *Transp. Res. Part A Policy Pract.* **2019**, *125*, 56–71. [CrossRef]
- Paiva, S.; Ahad, M.A.; Tripathi, G.; Feroz, N.; Casalino, G. Enabling Technologies for Urban Smart Mobility: Recent Trends, Opportunities and Challenges. *Sensors* **2021**, *21*, 2143. [CrossRef]
- Barreto, L.; Amaral, A.; Baltazar, S. Mobility in the Era of Digitalization: Thinking Mobility as a Service (MaaS). In *Intelligent Systems: Theory, Research and Innovation in Applications*; Springer: Cham, Switzerland, 2021; pp. 275–293.
- Turoń, K.; Czech, P. The Concept of Rules and Recommendations for Riding Shared and Private E-Scooters in the Road Network in the Light of Global Problems. In *Modern Traffic Engineering in the System Approach to the Development of Traffic Networks*; Macioszek, E., Sierpiński, G., Eds.; Advances in Intelligent Systems and Computing; Springer: Cham, Switzerland, 2020; Volume 1083. [CrossRef]
- Talebkhah, M.; Sali, A.; Marjani, M.; Gordan, M.; Hashim, S.J.; Rokhani, F.Z. IoT and Big Data Applications in Smart Cities: Recent Advances, Challenges, and Critical Issues. *IEEE Access* **2021**, *9*, 55465–55484. [CrossRef]
- Oliveira, T.A.; Oliver, M.; Ramalhinho, H. Challenges for Connecting Citizens and Smart Cities: ICT, E-Governance and Blockchain. *Sustainability* **2020**, *12*, 2926. [CrossRef]
- Muller, M.; Park, S.; Lee, R.; Fusco, B.; Correia, G.H.d.A. Review of Whole System Simulation Methodologies for Assessing Mobility as a Service (MaaS) as an Enabler for Sustainable Urban Mobility. *Sustainability* **2021**, *13*, 5591. [CrossRef]
- Gonçalves, L.; Silva, J.P.; Baltazar, S.; Barreto, L.; Amaral, A. Challenges and implications of Mobility as a Service (MaaS). In *Implications of Mobility as a Service (MaaS) in Urban and Rural Environments: Emerging Research and Opportunities*; IGI Global: Hershey, PA, USA, 2020; pp. 1–20.
- Amaral, A.; Barreto, L.; Baltazar, S.; Pereira, T. Mobility as a Service (MaaS): Past and Present Challenges and Future Opportunities. In *Conference on Sustainable Urban Mobility*; Springer: Cham, Switzerland, 2020; pp. 220–229.
- Barreto, L.; Amaral, A.; Baltazar, S. Urban mobility digitalization: Towards mobility as a service (MaaS). In Proceedings of the 2018 International Conference on Intelligent Systems (IS), Funchal, Portugal, 25–27 September 2018; pp. 850–855.

28. Al-Rahamneh, A.; Javier Astrain, J.; Villadangos, J.; Klaina, H.; Picallo Gueembe, I.; Lopez-Iturri, P.; Falcone, F. Enabling Customizable Services for Multimodal Smart Mobility With City-Platforms. *IEEE Access* **2021**, *9*, 41628–41646. [[CrossRef](#)]
29. Turoñ, K. Open Innovation Business Model as an Opportunity to Enhance the Development of Sustainable Shared Mobility Industry. *J. Open Innov. Technol. Mark. Complex.* **2022**, *8*, 37. [[CrossRef](#)]
30. Abbasi, S.; Ko, J.; Min, J. Measuring destination-based segregation through mobility patterns: Application of transport card data. *J. Transp. Geogr.* **2021**, *92*, 103025. [[CrossRef](#)]
31. 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](#)]
32. Chang, W.-J.; Chen, L.-B.; Chen, M.C.; Su, J.P.; Sie, C.Y.; Yang, C.H. Design and Implementation of an Intelligent Assistive System for Visually Impaired People for Aerial Obstacle Avoidance and Fall Detection. *IEEE Sens. J.* **2020**, *20*, 10199–10210. [[CrossRef](#)]
33. El-taher, F.E.; Taha, A.; Courtney, J.; Mckeever, S. A Systematic Review of Urban Navigation Systems for Visually Impaired People. *Sensors* **2021**, *21*, 3103. [[CrossRef](#)]
34. Dakopoulos, D.; Bourbakis, N.G. Wearable obstacle avoidance electronic travel aids for blind: A survey. *IEEE Trans. Syst. Man Cybern. Part C* **2010**, *40*, 25–35. [[CrossRef](#)]
35. Renier, L.; De Volder, A.G. Vision substitution and depth perception: Early blind subjects experience visual perspective through their ears. *Disabil. Rehabil. Assist. Technol.* **2010**, *5*, 175–183. [[CrossRef](#)]
36. Tapu, R.; Mocanu, B.; Tapu, E. A survey on wearable devices used to assist the visual impaired user navigation in outdoor environments. In Proceedings of the 2014 11th International Symposium on Electronics and Telecommunications (ISETC), Timisoara, Romania, 14–15 November 2014; pp. 1–4.
37. Jacobson, D.; Kitchin, R.; Golledge, R.; Blades, M. Learning a complex urban route without sight: Comparing naturalistic versus laboratory measures. In Proceedings of the Mind III Annual Conference of the Cognitive Science Society, Madison, WI, USA, 1–4 August 1998; pp. 1–20.
38. Kaminski, L.; Kowalik, R.; Lubniewski, Z.; Stepnowski, A. ‘VOICE MAPS’—Portable, dedicated GIS for supporting the street navigation and self-dependent movement of the blind. In Proceedings of the 2010 2nd International Conference on Information Technology (ICIT 2010), Gdansk, Poland, 28–30 June 2010; pp. 153–156.
39. Ueda, T.A.; De Araújo, L.V. *Virtual Walking Stick: Mobile Application to Assist Visually Impaired People to Walking Safely*; Lecture Notes in Computer Science (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics); Springer: Berlin/Heidelberg, Germany, 2014; Volume 8515 LNCS, pp. 803–813. [[CrossRef](#)]
40. Minhas, R.A.; Javed, A. X-EYE: A Bio-smart Secure Navigation Framework for Visually Impaired People. In Proceedings of the 2018 International Conference on Signal Processing and Information Security (ICSPIS), Dubai, United Arab Emirates, 7–8 November 2018; pp. 2018–2021. [[CrossRef](#)]
41. Kaiser, E.B.; Lawo, M. Wearable navigation system for the visually impaired and blind people. In Proceedings of the 2012 IEEE/ACIS 11th International Conference on Computer and Information Science, Shanghai, China, 30 May–1 June 2012; pp. 230–233. [[CrossRef](#)]
42. Zeb, A.; Ullah, S.; Rabbi, I. Indoor vision-based auditory assistance for blind people in semi controlled environments. In Proceedings of the 2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA), Paris, France, 14–17 October 2014; pp. 1–6. [[CrossRef](#)]
43. Fukasawa, A.J.; Magatani, K. A navigation system for the visually impaired an intelligent white cane. In Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 28 August–1 September 2012; pp. 4760–4763. [[CrossRef](#)]
44. Bhardwaj, P.; Singh, J. Design and Development of Secure Navigation System for Visually Impaired People. *Int. J. Comput. Sci. Inf. Technol.* **2013**, *5*, 159–164. [[CrossRef](#)]
45. Treuillet, S.; Royer, E. Outdoor/indoor vision-based localization for blind pedestrian navigation assistance. *Int. J. Image Graph.* **2010**, *10*, 481–496. [[CrossRef](#)]
46. Serrão, M.; Rodrigues, J.M.F.; Rodrigues, J.I.; Du Buf, J.M.H. Indoor localization and navigation for blind persons using visual landmarks and a GIS. *Procedia Comput. Sci.* **2012**, *14*, 65–73. [[CrossRef](#)]
47. Shadi, S.; Hadi, S.; Nazari, M.A.; Hardt, W. Outdoor navigation for visually impaired based on deep learning. *CEUR Workshop Proc.* **2019**, *2514*, 397–406.
48. Chen, H.E.; Lin, Y.Y.; Chen, C.H.; Wang, I.F. BlindNavi: A Navigation App for the Visually Impaired Smartphone User. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, Seoul, Korea, 18–23 April 2015; pp. 19–24. [[CrossRef](#)]
49. Idrees, A.; Iqbal, Z.; Ishfaq, M. An Efficient Indoor Navigation Technique to Find Optimal Route for Blinds Using QR Codes. In Proceedings of the 2015 IEEE 10th Conference on Industrial Electronics and Applications (ICIEA), Auckland, New Zealand, 15–17 June 2015; pp. 690–695. [[CrossRef](#)]
50. Zhang, X.; Li, B.; Joseph, S.L.; Muñoz, J.P.; Yi, C. A SLAM based Semantic Indoor Navigation System for Visually Impaired Users. In Proceedings of the 2015 IEEE International Conference on Systems, Man, and Cybernetics, Hong Kong, China, 9–12 October 2015; pp. 1458–1463. [[CrossRef](#)]

51. Yelamarthi, K.; Haas, D.; Nielsen, D.; Mothersell, S. RFID and GPS integrated navigation system for the visually impaired. In Proceedings of the 2010 53rd IEEE International Midwest Symposium on Circuits and Systems, Seattle, WA, USA, 1–4 August 2010; pp. 1149–1152. [[CrossRef](#)]
52. Tanpure, R.A. Advanced Voice Based Blind Stick with Voice Announcement of Obstacle Distance. *Int. J. Innov. Res. Sci. Technol.* **2018**, *4*, 85–87.
53. Salahuddin, M.A.; Al-Fuqaha, A.; Gavirangaswamy, V.B.; Ljucovic, M.; Anan, M. An efficient artificial landmark-based system for indoor and outdoor identification and localization. In Proceedings of the 2011 7th International Wireless Communications and Mobile Computing Conference, Istanbul, Turkey, 4–8 July 2011; pp. 583–588. [[CrossRef](#)]
54. Basiri, A.; Amirian, P.; Winstanley, A. The use of quick response (QR) codes in landmark-based pedestrian navigation. *Int. J. Navig. Obs.* **2014**, *2014*, 897103. [[CrossRef](#)]
55. Nilwong, S.; Hossain, D.; Shin-Ichiro, K.; Capi, G. Outdoor Landmark Detection for Real-World Localization using Faster R-CNN. In Proceedings of the 6th International Conference on Control, Mechatronics and Automation, Tokyo, Japan, 12–14 October 2018; pp. 165–169. [[CrossRef](#)]
56. Maeyama, S.; Ohya, A.; Sciences, I.; Tsukuba, T. Outdoor Navigation of a Mobile Robot Using Natural Landmarks. In Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robot and Systems. Innovative Robotics for Real-World Applications. IROS '97, Grenoble, France, 11 September 1997; Volume 3, pp. V17–V18. [[CrossRef](#)]
57. Raubal, M.; Winter, S. *Enriching Wayfinding Instructions with Local Landmarks*; Lecture Notes in Computer Science (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics); Springer: Berlin/Heidelberg, Germany, 2002; Volume 2478, pp. 243–259. [[CrossRef](#)]
58. Yeh, T.; Tollmar, K.; Darrell, T. Searching the web with mobile images for location recognition. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2004, Washington, DC, USA, 27 June–2 July 2004; IEEE: Piscataway, NJ, USA, 2004; Volume 2, p. II.
59. Li, F.; Kosecka, J. Probabilistic location recognition using reduced feature set. In Proceedings of the 2006 IEEE International Conference on Robotics and Automation ICRA 2006, Orlando, FL, USA, 15–19 May 2006; IEEE: Piscataway, NJ, USA, 2006; pp. 3405–3410.
60. Schindler, G.; Brown, M.; Szeliski, R. City-scale location recognition. In Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–22 June 2007; IEEE: Piscataway, NJ, USA, 2007; pp. 1–7.
61. Gallagher, A.; Joshi, D.; Yu, J.; Luo, J. Geo-location inference from image content and user tags. In Proceedings of the 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Miami, FL, USA, 20–25 June 2009; IEEE: Piscataway, NJ, USA, 2009; pp. 55–62.
62. Schroth, G.; Huitl, R.; Chen, D.; Abu-Alqumsan, M.; Al-Nuaimi, A.; Steinbach, E. Mobile visual location recognition. *IEEE Signal Process. Mag.* **2011**, *28*, 77–89. [[CrossRef](#)]
63. Cao, S.; Snavely, N. Graph-based discriminative learning for location recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 23–28 June 2013; pp. 700–707.
64. Wan, J.; Wang, D.; Hoi, S.C.H.; Wu, P.; Zhu, J.; Zhang, Y.; Li, J. Deep learning for content-based image retrieval: A comprehensive study. In Proceedings of the 22nd ACM International Conference on Multimedia, Orlando, FL, USA, 3–7 November 2014; pp. 157–166.
65. Liu, P.; Yang, P.; Wang, C.; Huang, K.; Tan, T. A semi-supervised method for surveillance-based visual location recognition. *IEEE Trans. Cybern.* **2016**, *47*, 3719–3732. [[CrossRef](#)]
66. Zhou, W.; Li, H.; Tian, Q. Recent advance in content-based image retrieval: A literature survey. *arXiv* **2017**, arXiv:1706.06064.
67. Tzelepi, M.; Tefas, A. Deep convolutional learning for content based image retrieval. *Neurocomputing* **2018**, *275*, 2467–2478. [[CrossRef](#)]
68. Saritha, R.R.; Paul, V.; Kumar, P.G. Content based image retrieval using deep learning process. *Clust. Comput.* **2019**, *22*, 4187–4200. [[CrossRef](#)]
69. Bhagwat, R.; Abdolahnejad, M.; Moocarme, M. *Applied Deep Learning with Keras: Solve Complex Real-Life Problems with the Simplicity of Keras*; Packt Publishing Ltd.: Birmingham, UK, 2019.
70. Lima, R.; da Cruz, A.M.R.; Ribeiro, J. Artificial Intelligence Applied to Software Testing: A literature review. In Proceedings of the 2020 15th Iberian Conference on Information Systems and Technologies (CISTI), Seville, Spain, 24–27 June 2020.
71. Lynch, K. The image of the environment. *Image City* **1960**, *11*, 1–13.
72. Pfleeger, C.; Shari, L. *Security in Computing*, 4th ed.; Prentice Hall PTR: Hoboken, NJ, USA, 2007.
73. Khan, J.; Abbas, H.; Al-Muhtadi, J. Survey on Mobile User's Data Privacy Threats and Defense Mechanisms. *Procedia Comput. Sci.* **2015**, *56*, 376–383. [[CrossRef](#)]
74. Becher, M.; Freiling, F.C.; Hoffmann, J.; Holz, T.; Uellenbeck, S.; Wolf, C. Mobile Security Catching Up? Revealing the Nuts and Bolts of the Security of Mobile Devices. In Proceedings of the 2011 IEEE Symposium on Security and Privacy, Oakland, CA, USA, 22–25 May 2011; pp. 96–111. [[CrossRef](#)]

-
75. Stallings, W. *Cryptography and Network Security: Principles and Practice*, 7th ed.; Pearson: London, UK, 2020.
 76. Varadharajan, V.; Tupakula, U. Security as a Service Model for Cloud Environment. *IEEE Trans. Netw. Serv. Manag.* **2014**, *11*, 60–75. [[CrossRef](#)]
 77. Paiva, S.; Ahad, M.A.; Zafar, S.; Tripathi, G.; Khaliq, A.; Hussain, I. Privacy and security challenges in smart and sustainable mobility. *SN Appl. Sci.* **2020**, *2*, 1175. [[CrossRef](#)]