A Context-Aware Artificial Intelligence-based System to Support Street Crossings For Pedestrians with Visual Impairments

Aleksandro Montanha, Andreea M. Oprescu & MCarmen Romero-Ternero

Taylor & Francis
Taylor & Francis Group

A Context-Aware Artificial Intelligence-based System to Support Street Crossings For Pedestrians with Visual Impairments

Aleksandro Montanha [a], Andreea M. Oprescu [b], and MCarmen Romero-Ternero [b]

aPrograma Doctorado Ingeniería Informática, Escuela Técnica Superior de Ingeniería Informática, Universidad de Sevilla, Seville, Spain; bDepartamento de Tecnología Electrónica, Universidad de Sevilla, Seville, Spain

ABSTRACT

Artificial intelligence has the potential to support and improve the quality of life of people with disabilities. Mobility is a potentially dangerous activity for people with impaired ability. This article presents an assistive technology solution to assist visually impaired pedestrians in safely crossing the street. We use a signal trilateration technique and deep learning (DL) for image processing to segment visually impaired pedestrians from the rest of pedestrians. The system receives information about the presence of a potential user through WiFi signals from a mobile application installed on the user's phone. The software runs on an intelligent semaphore originally designed and installed to improve urban mobility in a smart city context. This solution can communicate with users, interpret the traffic situation, and make the necessary adjustments (with the semaphore's capabilities) to ensure a safe street crossing. The proposed system has been implemented in Maringá, Brazil, for a one-year period. Trial tests carried out with visually impaired pedestrians confirm its feasibility and practicality in a real-life environment.

Introduction

Information and communication technology (ICT) interventions have an undeniable potential to help people with disabilities live their lives more independently and improve their quality of life (Lancioni et al. 2020). Artificial intelligence (AI) and mobile technologies have been shown to support people with disabilities by increasing their participation in society, helping them communicate, learn, shop, travel, move around town, among many others (Baumgatner, Rohrbach, and Schönhagen 2021), (Molina-Cantero et al. 2019), (Balasuriya et al. 2017), (Mahmud et al. 2020).

Specifically, assistive technologies using AI can improve the quality of life of people with visual disabilities. According to the World Health Organization, 2.2 billion people live with near or far-reaching vision impairment throughout

CONTACT Aleksandro Montanha mcromerot@us.es Programa Doctorado Ingeniería Informática, Escuela Técnica Superior de Ingeniería Informática, Universidad de Sevilla, Avda. Reina Mercedes S/n, 41012 Seville, Spain

the world (World Health Organization 2021). A challenging everyday activity for people with blindness or visual disabilities is mobility and, more specifically, street crossing (Hakobyan et al. 2013). Recently, more cities have already implemented accessibility features for street crossings, mainly in the form of sound alerts emitted by traffic lights. This alert can guide pedestrians by emitting a sound when the traffic light is green.

However, in cities where noise pollution is an issue, this type of solution may not be sufficient to provide the accessibility that people with visual disabilities need. In addition, cases where intersections are close to each other can also create confusion for visually impaired pedestrians. If there are multiple traffic lights that emit sound alerts at a single intersection, it can be confusing and time consuming for visually impaired pedestrians to recognize when it is safe to cross. Therefore, the sound alert system presents many deficiencies in terms of its ability to transcribe the environment. With solutions that can transcribe the environment and act on the traffic situation if necessary, besides potentially providing a safer option, could be empowering and improve the interaction of visually impaired users with the urban environment.

Currently, with the advent of the Internet of Things (IoT), which combines ubiquitous computing, sensing technologies, Internet communication protocols, and embedded devices (Suzuki 2017), a scenario with numerous solutions for smart city environments has emerged. Smartphones, as an important part of this paradigm, are widely available today to the general public. More importantly, in this context, visually impaired people use smartphones extensively as assistive technology (Khan & Khusro, 2020), (Retorta and Cristovão 2017).

We introduce a context-sensitive system that guides visually impaired people at street crossings by automatically detecting pedestrians and vehicles in traffic. The hardware installation at an intersection consists of smart traffic lights (Seebot Agent)[1] paired with cameras and WiFi antennas. Communication between the system and the visually impaired pedestrian is achieved through a mobile application installed on the user's smartphone. This method uses a signal trilateration technique and a computer vision pre-trained model to compute the user's location and the traffic situation.

Throughout the design and development process of this solution, potential users with visual impairment participated. Furthermore, field tests have been carried out with visually impaired users to test the application, which yielded positive results. The innovation of this work lies in 1) the solution has been built on existing hardware, an intelligent semaphore that was initially installed in Brazilian cities for smart city solutions related to traffic management; 2) the ability to transcribe the traffic situation to the user and to act on the situation to ensure user safety, since the semaphore can alter its behavior when a visually impaired pedestrian is crossing, if needed.

Related Work

Mobility solutions for visually impaired users have been widely addressed in the literature.

Advances in pedestrian detection and tracking using computer vision and machine learning techniques are discussed in (Brunetti et al. 2018). The authors call for more architectures to be implemented and tested. Additionally, one solution they point toward includes the use of RGB cameras paired with a DL strategy. (Chinchole and Patel 2017) developed a stick solution using an Arduino Nano microcontroller, two ultrasonic sensors, and an accelerometer. The authors also used a smartphone to obtain images of the user's surroundings. After processing with artificial intelligence techniques, these images provided information about the user's environment. A path hole detection system has been developed by (Islam and Sadi 2018) to assist visually impaired pedestrians using convolution neural networks (CNN). The authors achieved positive accuracy results in the experiments performed. A stick prototype for staircase and ground detection using a camera and an ultrasonic sensor has been proposed by (Habib et al. 2019). They used a pre-trained CNN model for training the system.

In addition, there is no doubt that crossing the streets is a major challenge for visually impaired pedestrians.

In this sense, an important number of solutions have been proposed in the literature. Some efforts have been devoted to developing an intelligent solution that includes traffic lights (through a connected device) in the cloud (Amazon Web Services-based system) and then alerts visually impaired users when approaching an intersection (Ihejimba and Wenkstern 2020). This work presents some limitations given the rough precision provided by GPS at ground level, especially at an intersection with several crossing directions. Other authors evaluated the feasibility of machine learning algorithms to identify traffic lights and track their status so that the visually impaired user can cross the road safely (Ghilardi et al. 2018). Furthermore, custom hardware implementations that apply image processing and machine learning to identify and track traffic lights have been reported in (Li et al. 2019). Information and feedback are provided to the visually impaired user through a voice synthesizer. Another similar version is based on a mobile phone device that must be held horizontally when crossing the street (Yu, Lee, and Kim 2019). The user is provided with guidance from the vibrations and sounds of the device. (Cheng et al. 2017) presented an assistive solution for crosswalk detection, consisting of an adaptive extraction and consistency analysis algorithm and a wearable navigation system (Intoer) for people with visual impairments. To achieve a more robust proposal, the authors considered several crosswalk situations that users can encounter: crosswalks at far distances from the user, crosswalks at close distances from the user, and crosswalks in front of the user. In addition, field tests were performed considering different variables,

such as time of day and weather. (Cheng et al. 2018) proposed a real-time Pedestrian Crossing Lights (PCL) detection algorithm using machine learning. The novelty of this work is based on its high precision and recall, robustness, and low complexity. Additionally, the authors applied a temporal-spatial analysis to track the detected PCL in order to improve the system's results in challenging scenarios. Communication between the user and the device, consisting of an Intel RealSense R200 camera, a pair of bone-conduction earphones and a portable PC, is achieved through voice instructions. (Tian et al. 2021) described an assistive solution to understand crosswalk scenes and provide users with key information and directions to cross the street. The device consists of a head-mounted mobile equipped with an Intel RealSense camera and a mobile phone. In their solution, the authors carried out field tests on crosswalk detection, pedestrian traffic light recognition, and distance measurement, and proved their practical usability. One of their main contributions is the SensingAI dataset with more than 11,700 images for the recognition of pedestrian traffic lights, key objects on the crossroads, and the crosswalk. (Yang et al. 2018) described a real-world framework using a wearable device to help visually impaired pedestrians with navigation tasks. Specifically, they study the context situation at an intersection by detecting the crosswalk, traffic light, pedestrian and vehicle. The authors carried out comprehensive experiments as part of their framework, comparing their own approach with other algorithms against several datasets.

The difficulties of visually impaired people experience using mobile applications have been studied (Qureshi and Wong 2019). The application of the user-centered design (UCD) philosophy is especially important for "not mainstream" user groups, and the resulting products have high acceptance rates (Krajnc, Feiner, and Schmidt 2010). Accessibility makes user interfaces perceivable, operable and understandable by people with a wide range of abilities and people in a wide range of circumstances, environments, and conditions. Therefore, accessibility also benefits people with disabilities and organizations that develop accessible products (Henry 2007).

In the literature, an increasing number of crosswalk assistive solutions for visually impaired pedestrians have been reported. Most of these advances are possible due to the increasingly higher processing power of devices, computer vision techniques, and different wearable devices and smartphones. Our solution also benefits from this aspect; however, in our case, the user only needs a smartphone. To our knowledge, this work is the first to report an ecosystem in which traffic alterations can be performed in favor of visually impaired pedestrians.

This work adds to our previous work in which the signal trilateration technique is proposed (Montanha et al. 2016). Compared to our previous work, the current proposed system incorporates the ability to segment visually impaired pedestrians automatically and in real time. In the previous version, they could not be distinguished from the rest of the pedestrians, so personalized

guidance could not be provided to cross the street. The previous version of the system could detect the presence of users and inform users carrying the smartphone application about the color of the traffic light. This work incorporates the visually impaired pedestrian segmentation technique, which allows the system to guide and assist them when crossing the street for the entire duration of the process. The app was designed and developed following the UCD approach.

Pedestrian Tracking Proposal

In this section, a proposal to detect and track pedestrians will be explored. The method implements a heuristic to estimate the pedestrian's location using a signal trilateration process and a deep learning model.

Signal Location Method to Support Pedestrian Tracking

The system establishes a communication canal between the smart traffic lights and the application installed on the smartphone of the visually impaired user. This is a wireless communication achieved via WiFi signals between the two parts. At least three WiFi antennas are placed at an intersection, as can be seen in Figure 1.

Figure 1. Proposal of camera and smart traffic lights installation in an intersection.

The smart traffic lights continuously monitor intersections using IP cameras. When a visually impaired pedestrian (with the mobile application opened) is detected to be within the WiFi range of at least one of the antennas, the localization process is requested. The system starts to compute a signal trilateration technique to identify in which quadrant the user is located (Montanha, Polidorio, and Romero-Ternero 2021). We explain this technique in detail in one of our previous works (Montanha, Polidorio, and Romero-Ternero 2021). The process continues when the user is located (the system detects in which quadrant they are). We will refer to this quadrant as the monitoring area. A processing thread is started for each pedestrian (visually impaired or not) that is located within the range of the monitoring areas, as they are being monitored by the cameras at the traffic light. The images captured from the monitoring area are further submitted to a pre-trained model whose function is to search for traces that identify pedestrians. This technique will be explained in more detail in the next section.

Neural Network Pre-trained Model to Support Pedestrian Tracking

The proposed system uses a pre-trained deep neural network model from the open-source toolkit OpenVINO[2] (Open Visual Inference and Neural Network Optimization), offered by Intel. The OpenVINO toolkit allows one to quickly develop and deploy a wide range of computer vision applications. The pre-trained model used, *person-detection-retail-0013*, is part of the Object Detection Models group of Intel's Pre-Trained Models. This model is based on the MobileNetV2-like backbone that includes depth-wise convolutions to reduce the amount of computation for the 3×3 convolution block. The single SSD head from 1/16 scale feature map has 12 clustered prior boxes. The model has an area under the precision/recall curve of 88.62%. The system implementing the neural network is capable of performing inference at a rate of one frame per second using the pre-trained model. The dimensions of the image are 640×480 pixels, and the color is RGB using codec H.265. The image format is MJPEG.

In the proposed solution, the information detailed below is collected by the pre-trained algorithms at a once-per-second acquisition rate: image data, the position relative to the image, date and time of the event, dimensions of the segmented pedestrian, and predominant colors. The image data is stored using the Base64 encoding algorithm, which is commonly used to embed the image data within other formats, JSON in our case.

Data Fusion for Signal Location and Image Processing Techniques

Data obtained by signal trilateration and image processing techniques are combined to improve pedestrian segmentation precision.

In each processing thread, a statistical data correlation process is performed between the position of the detected pedestrian in the spatial plane calculated by the signal trilateration technique (x_{tri}, y_{tri}) and the relative position of the pedestrian in the image, obtained from the image processing algorithm (x_{img}, y_{img}). Both pairs of coordinates refer to the pedestrian's location. If a correlation is established, the system then starts to look for the reidentification of the individual in the next frame to confirm the segmentation.

The Pearson correlation coefficient is used because it measures the strength of the linear relationship between two variables. Correlation analysis is calculated in the R programming language and the *pacman* package. As input, the script receives a .csv file generated by the algorithm. The following variables are stored in the .csv: date and time, trilateration ID, trilateration coordinates (x_{tri}, y_{tri}), pedestrian ID, pedestrian pixel coordinates in the image (x_{img}, y_{img}), predominant color of the pedestrian image, height and width of the pedestrian.

The system relies on the computation of the Pearson coefficient to guide the pedestrian through the street crossing process. Under certain circumstances (excess noise in the data, low visibility), it may occur that during the computation process a correlation is not found or is below the threshold of 90%. This implies that identification and tracking of the blind pedestrian are not possible, since we cannot merge information from cameras and signals. Since this is a possible event, an alternative procedure must be defined to allow the pedestrian to cross the street safely. The system prepares a contingency assistive procedure. This assistive procedure informs the visually impaired pedestrian about the status (color) of the traffic light.

System Design and Architecture

This section addresses the main components of the design and architecture of the system.

There are three main hardware components in the proposed system: smartphone, smart traffic lights, and WiFi antennas. The smart traffic lights monitor intersections using IP cameras and computing the signal trilateration and image processing algorithms. Meanwhile, the WiFi antennas installed at the intersections are responsible for the exchange of the WiFi signal with the smartphone. An overview of the system is shown in Figure 2.

There are two main software components: the computations running on the smart traffic light (signal trilateration technique, neural network-based image processing algorithm, data correlation process), and the mobile application.

The transmission of information between the mobile application and the smart traffic light is carried out using high-speed cellular network technology such as 3 G or 4 G. WiFi signals are used solely for localization purposes.
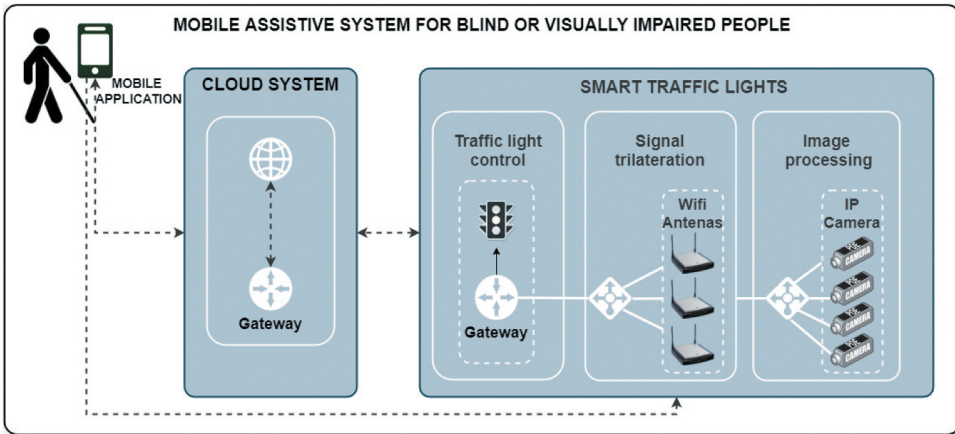
Figure 2. Overview of the system design for the proposed mobile assistive system.

As explained previously, the target user localization is first estimated in one of the quadrants (monitoring area) using our signal trilateration technique. Then, the image processing technique is added to the localization system, to segment the target user from the rest of the pedestrians. The images captured by the IP cameras from the monitoring areas are submitted to the pre-trained model. For each segmented pedestrian, the system stores information in a buffer. Each buffer has a space for analysis of 100 frames. Once the buffer is filled, the system starts the data fusion method. Once a correlation is established, the system then starts to search for the reidentification of the individual in the next frames to confirm the segmentation of the visually impaired pedestrian. OpenCV version 4.5.2 and Python 3 have been used to capture and process the images on a Debian 10 Linux distribution (Buster version).

The data resulting from the correlation process between the signal trilateration data and the image processing data need to exceed the selected threshold of 90% to start the assistance process. In addition, information from the traffic situation is checked, adding the following conditions to be met before initiating safe crossing: 1) the traffic light is closed for vehicles on the lane; 2) there are no vehicles on the lane, or, if there are, the first line of vehicles are stopped, waiting for the traffic light to turn green. If the traffic light is red for vehicles but there are vehicles moving in the lane, at high or low speeds, the system inspects the area closest to the crosswalk. If there is no line of stopped vehicles in front of the crosswalk, the system does not send instructions for crossing (we need to assume that a car could not stop at a red light).

Information is sent to the mobile application, alerting about the status of the traffic lights and the presence of vehicles in motion. The interaction between the mobile application interface and the user is explained in more detail in Section 4.2. In addition, smart traffic lights can increase the duration of the
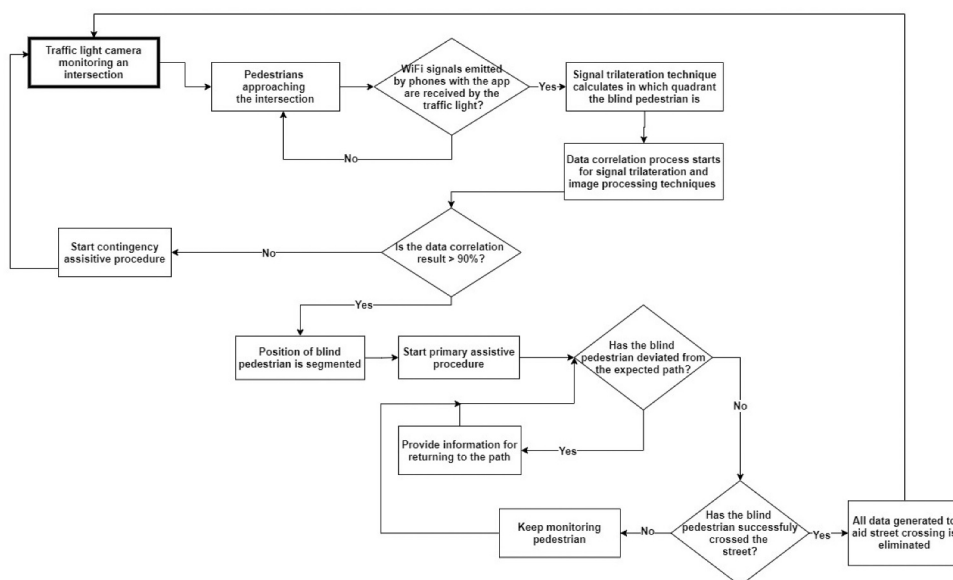
Figure 3. Process flow diagram of the mobile assistive system for blind or visually impaired people.

green light cycle when a visually impaired pedestrian has been detected. This change in the duration of the traffic light cycle must be made according to local traffic regulation laws.

If the user takes a deviation from the street crossing (gets out of the monitoring area), the system dynamically prepares a set of instructions to amend the trajectory. Once the system detects that the user has finished crossing the street and the light cycle ends, the algorithm terminates the instance and removes the data collected for the crossing, according to and complying with the local data protection regulation (Brazilian General Data Protection Law, in this case).

As mentioned above, if the system does not consolidate the data, the contingency system is activated, informing pedestrians about the status of the traffic light and about the inability to guide them during the crossing process. A flow diagram of the system operation process is presented in Figure 3. The system can assist more than one user at a time, since mobile devices are uniquely identified via the MAC address.

The Smart Traffic Light

The traffic light used in the proposed system is an Agent Seebot Smart Traffic Light.[3] This device incorporates an Intel I5 7th generation processor with a high-resolution LED display interface. This device has an Axis HD IP camera with image capture and a transmission rate greater than 30 fps. The connection protocols are Ethernet, Bluetooth, and WiFi. The device also incorporates an inclinometer, an accelerometer, and a motorized camera angle adjustment

Figure 4. Image of a pedestrian crossing retrieved from a Seebot agent installed at an intersection.

system, capable of tilting 80 on the Y-axis (vertical). The images captured by the smart traffic lights are at least 5.5 meters high, and the distance in relation to the area destined for the crossing is dynamic and can change according to the characteristics of the place. An image recovered from an installed semaphore can be seen in Figure 4. The crossing area is detected and marked in light green, the pedestrian is marked in fluorescent green, and the car lane is marked in red.

Generally, the area intended for pedestrian crossing is marked with a horizontal painting; however, in many places, they do not have constant maintenance and therefore may have a discrete appearance or even not exist. Therefore, it is necessary to inform the system of the location in the space plan where there is an area for pedestrians to cross. This feature allows drawing a polygon that serves as a pedestrian segmentation parameter. The system uses a camera with infrared lighting and, in some environments, has additional infrared lighting for places with low lighting during the night period. Each traffic light has the computational capacity to evaluate at least two crossing points in real time in both directions. Due to electronic control of the vertical angle of the camera installed on the device, the camera can be directed both at the viewing angle ahead, below, and even after crossing the lanes. This allows the traffic light to be installed in different intersection scenarios. Image capture is continuous, 24 hours a day.

The Mobile Application Interface

An Android-based mobile app to help urban mobility for visually impaired people has been developed. An Android-based smartphone device was chosen because of its popularity. The hardware and software needed for the solution are already available on the device (GPS, WiFi, Bluetooth, voice system, etc.). Additionally, many visually impaired people are familiar with these devices, so we decided to introduce this technology proposal to a group of target users. They

provided valuable insight and knowledge that we used to develop this mobile app interface. The group had experience with smartphones and was able to move autonomously in the urban environment. We were interested in understanding how they interact with signals, other individuals, obstacles, and more importantly technology. The group oriented us on how to establish an appropriate communication channel between a visually impaired user and the application, using speech recognition features, vibration, sounds, and voice messages. This application follows the set of accessibility compliance testing (ACT) rules recommended by the World Wide Web Consortium (W3C) (World Wide Web Consortium" 2008), and the international standards for the Web.

Several screenshots of the application are shown in Figure 5 (main screen for the initialization of the location of the pedestrian, an aerial view of the smart intelligent traffic lights installed on a street and their status, and an example of an alert created for a user to wait for the traffic light to open).

When a user approaches a street crossing, the application starts providing information by voice to the user on what to do next. Users are instructed to point the mobile device toward the direction where they wish to cross parallel to the ground, and once that has been done, the application instructs the user to confirm their direction by tapping twice on the screen. The application informs the user about the selected path (street name) and asks the user to confirm by sliding the finger over the screen. The application notifies the user of the confirmed path and informs the user of the state (color) of the traffic light. On the other hand, the traffic light is notified of the presence and desired direction of a visually impaired user. If the color is red, the system starts a countdown for the color to change to green. Finally, when the traffic light is
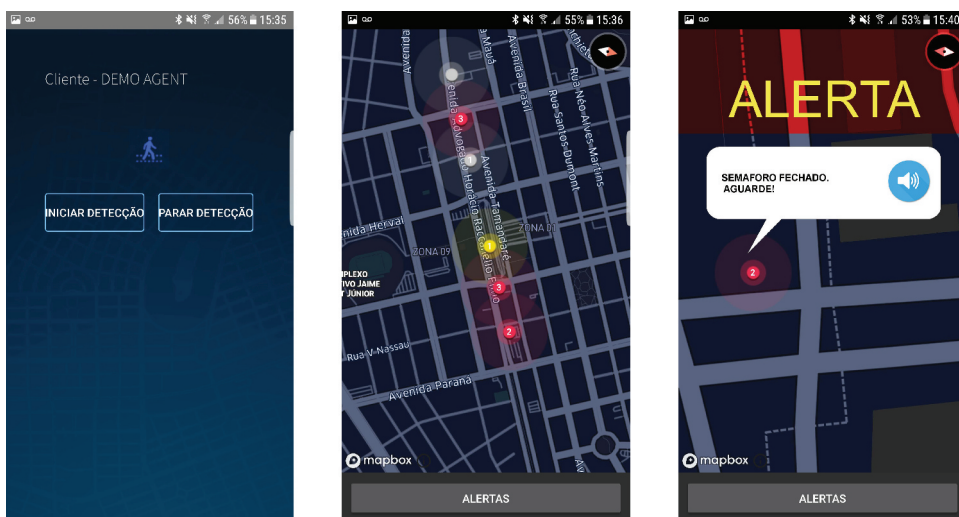
(a) App's main screen to initiate or stop localization process

(b) Aerial view of the intelligent semaphores on a street

(c) Alert to hold (closed semaphore, hold on)

Figure 5. Mobile application for street crossing assistance screenshots.

green, the user is informed that there are no vehicles on the street and that the light is green. Therefore, traffic light communication with the mobile application includes information about the conditions of the local environment (if there are vehicles, if vehicles are moving, or if the traffic light is opened or closed). Even if the visually impaired user has already started crossing the street without interacting with the application, guidance and information about the state of the road and traffic light are provided. If the user leaves the correct and safe route, a route error warning message is sent with instructions on how to correct the error.

Installation Process

To physically install the solution, we faced several challenges, as smart traffic lights were already installed. This represented our biggest operational challenge, since we wanted to interfere as little as possible with vehicle flow. A logistical operation was established in the city to divert traffic. We installed the WiFi antennas and adjusted the IP cameras so that they have enough aperture to capture both vehicles on the roads and waiting pedestrians (not only in the crossing, but also in the crosswalk area). Additionally, an uninterruptible supply system with a capacity of 4 hours of operation was installed at the intersection to guarantee autonomy in the event of a momentary power failure, common on stormy days.

Once the system was running, our next challenge was to integrate the proposed solution with existing semaphore software. As part of this adaptation, the main task was to calibrate the WiFi signal at each point. During the operation, a drone was used to measure the signal of each WiFi antenna in the waiting areas of the crosswalks.

The assistive system was in operation for tests for 30 days. After 30 days, the system was made available for testing so visually impaired users could validate the application.

Results

In this work, we will present some preliminary results collected from the installation in Maringá, Brazil.

The proposal to develop this feature raised the need to understand how blind people interact with mobile devices. We invited a group of visually impaired users to help us develop the app (applying UCD methodology). We observe the essential features that a mobile application interface must present to work with locomotion guidance in an urban environment. In this way, we carefully elaborated the messages and information the system offers as audio indications: indicating to the users their position in relation to the intersection, indicating the name of the streets at the intersection, and asking

which street the user would like to cross. An appropriate way to capture the user's response is by sliding on the screen in the horizontal direction for one choice or vertically for another choice. We found a practical way to notify users when approaching the intersection using smartphone vibrations and asking if they would like to access the assistant to cross the intersection. When the user responds affirmatively, they receive audible information (literal transcription) on the status of the traffic light, the presence of vehicles, and the assistance process for crossing.

In Table 1 the dynamic of the operation of the system over a period of 3 months can be observed (number of connections to the app, number of crossing assistance triggered, and number of segmentations). During the first month (07/2019), the app was made available to a user who made 7 connections to the application. Of the 7 connections, the system activated safe crossing assistance 6 times (when the data fusion result was above the 90% threshold). In the second month, the number of users who use the app increased due to the advertisement of the application in a social networking group. The number of connections also increased, and, in all cases, safe crossing assistance was successfully activated. During the third month, we experienced a high volume of app usage, as the group was motivated to try the solution. A significant number of cross-requests achieved good segmentation (96.5%). Safe crossing assistance was completed successfully in 80.8% of the total of app connections. The remaining 19.2% did not cross or crossed with human assistance and did not inform the platform.

The hardware on which the system is deployed is very potent, based on an Intel i5 7° Generation processor. It is an already existing and installed software (a Seebot agent). As it has been explained before, there are two main processes running in the system: the signal trilateration process and the image processing process. From our observations, as for the computational power needed for the signal trilateration technique, the usage is very low in comparison to the device's capabilities, it consumes 1–3% of one of the five cores and 10–20MB of memory with a time processing of 0.1 seconds. On the other hand, the image processing process (which includes frame capture, frame processing, pedestrians detection and reidentification) consumes 10–30% of two cores and 50–80MB of memory with 32 frames per second. In this scenario, there is a user detected in the signal trilateration system and there is a constant segmentation process of a visually impaired individual using the system. The detection of more than one user in the environment increases the processing consumption around 3–5% and the

Table 1. Number of app connections, assistance for safe crossings, and successful data fusion of the system operating over a 3 month period.

Period	App connections	Crossing assistance completed	Segmentation (trilateration and image processing)
07/2019	7	6	6
08/2019	14	14	14
09/2019	115	93	111

memory consumption around 10 MB. As we indicate before, considering this configuration and the availability of resources for the application in the traffic light, the system is able to handle up to 5 simultaneous individuals.

For a qualitative analysis, the mobile application was made available to 11 visually impaired user volunteers. We have interviewed users using a qualitative interview about their experience, usability, and user experience with the app. 27% of this population had reduced mobility. 83% of volunteers are familiar with the use of the app (50% always uses mobile apps to get information and navigate the city and 36.3% uses it routinely), while 8.3% rarely and 8.3% never. 100% of the users thought that the app was not complex to use and that the app does not need technical support for its use. Most people with visual impairments would learn to use the app quickly, and they felt safe and confident using the crossing system provided by the app. 58.3% did not need to learn anything else before using the app. 100% did not find it too complicated to use the app, thought that everything was well integrated and easy to use. All of them felt safe to cross while using the app. 66,6% were not instructed by other people to follow the crossing, even using the app.

Regarding the usability of the app, 33.3% ranked it with 10 out of 10, 33.3% with 9, 16.6% with 8 and 16.6% with 7.

Volunteers were asked what they would suggest to improve the app, and those who responded said: "That a campaign be made so that in more traffic lights the same was installed.," "Extending your range to other traffic lights and natural enhancements.," "It is great," "That there was a policy to have more of these traffic lights in the city.," "The app is just perfect and regarding some-thing like that, it is hard to give any kind of suggestions.," "The app is excellent, but it should have other ways to use it without having to use the cell phone." and "No suggestions."

Discussion

One of the key aspects of the proposed system is its ability to translate and adapt the environment to user capabilities, which is a crucial feature when developing solutions for disabled individuals. When creating this solution, we have followed the principles of user-centered design by interacting with users throughout the design and development process. It was especially important to understand how these specific users interact with the urban environment and how they use mobile application interfaces to design and implement a solution that would be useful and easy to use.

This system is aware of the presence and location of visually impaired people. It not only describes the traffic situation, but can also inform the user about the crossing situation in real time, offering a personalized set of instructions. The system can adapt the environment to user needs by prolonging the duration of the interval to cross the street (the traffic light will remain green longer for

pedestrians). The real-world pilot study has shown great results; however, we are aware that our research presents several limitations. There is room to improve the accuracy of the system in more extreme circumstances, such as low visibility due to weather conditions or nighttime. Second, the performance and usage data could have been collected over a longer period to develop statistical information. Another drawback of the proposal could be that the system is highly dependent on the hardware installation. However, we consider this to be beneficial for the user who only needs to carry a smartphone. A usability problem was detected in situations where there is noise pollution due to heavy traffic or construction near the site. It can be difficult for the user to hear the instructions in this situation, so we are continuing to investigate alternative methods. Lastly, an analysis of different variants of the approach could have been beneficial to the study.

Although there are more solutions to include people with visual impairments in the urban environment, there are still obstacles to overcome. Specifically, for urban environments that aim to become inclusive smart cities. We believe that there is a need to design and implement technological solutions that are safe and useful to users, widely available, and economically viable.

This system was built into the existing computational platform of the intelligent traffic lights in use. These traffic lights already perform the task of segmenting vehicles and reading license plates. At each intersection, these tasks typically consume 70% of the available processing power. To avoid saturating the processor and therefore offering some level of slowness, there is an additional processing limit that can be used by additional resources to the primary tasks of the smart semaphore. For this reason, it is not possible to carry out more than three simultaneous processes of segmentation of people with visual impairments at the same intersection, and a fourth request would enter the process queue. If this happens, the system activates the emergency system and puts plan B into practice, notifying the user who cannot perform the assisted crossing.

Conclusion

In the context of AI-based technology, this work contributes to the development of assistive technology for the mobility of visually impaired pedestrians, specifically for street crossings. To our knowledge, this is the first study designed to aid visually impaired pedestrian street crossings that combines the use of smart traffic lights as an element to detect traffic and pedestrian situation through sensors, image processing, AI techniques, and mobile devices.

This solution is based on the paradigms of Smart Cities and Internet of Things, seeking the fusion of data and the use of intelligent infrastructure. The functionality of the existing smart traffic light has been extended by adding WiFi antennas and developing segmentation software. This system as a whole can carry out a great number of tasks: acquire and transmit real-time images of their coverage locations to a traffic control center, communicate with pedestrians in

possession of mobile devices or devices installed in vehicles via WiFi and/or Bluetooth signals, assistive crossing for visually impaired pedestrians, among many other traffic regulation tasks. Furthermore, the data collection performed by this device contributes to the creation of a training database.

We can conclude that the proposed solution adds as an innovating factor the use of existing infrastructure to deploy assistive technology for visually impaired pedestrians. Often, solutions to assist visually impaired people on a street crossing require the use of a portable device (in the form of glasses or a stick). We can see that there are benefits to a solution that limits the use of a smartphone on the user's end and relies more heavily on a common infrastructure facilitated by a smart city context. Furthermore, communication between assistive software and traffic regulation software can be beneficial for visually impaired users, as the traffic light can adapt to the traffic situation to ensure safe crossing.

Future Work

Further research is required to improve the functionality of the system at night. Without adequate light, most of the time the system is unable to segment the visually impaired pedestrian, and the emergency assistive system has to be applied.

Our results are encouraging, so further development of techniques and algorithms to improve accuracy can be carried out within the scientific community. As a continuation of this work, open-source code can be released, making it available on GitHub, so researchers around the world can contribute to the code and implement the system in their cities.

Notes

1. https://seebot.com.br/es/agent.html
2. https://docs.openvinotoolkit.org
3. https://seebot.com.br/es/agent.html

Disclosure Statement

ORCID

Aleksandro Montanha ⓘD http://orcid.org/0000-0002-6288-2753
Andreea M. Oprescu ⓘD http://orcid.org/0000-0003-3308-4688
MCarmen Romero-Ternero ⓘD http://orcid.org/0000-0001-6965-9485

References

Balasuriya, B. K., N. P. Lokuhettiarachchi, A. R. M. D. N. Ranasinghe, K. D. C. Shiwantha, and C. Jayawardena. 2017. Learning platform for visually impaired children through artificial intelligence and computer vision 2017 11th International Conference on Software, Knowledge, Information Management and Applications (SKIMA) (Manhattan, New York, U.S.: IEEE) Colombo, Sri Lanka 1–7 doi:10.1109/SKIMA.2017.8294106.

Baumgatner, A., T. Rohrbach, and P. Schönhagen. 2021. 'if the phone were broken, i'd be screwed': Media use of people with disabilities in the digital era. *Disability and Society* 1–25. doi:10.1080/09687599.2021.1916884.

Brunetti, A., D. Buongiorno, G. F. Trotta, and V. Bevilacqua. 2018. Computer vision and deep learning techniques for pedestrian detection and tracking: A survey. *Neurocomputing* 300:17–33. Retrieved from https://www.sciencedirect.com/science/article/pii/S092523121830290X

Cheng, R., K. Wang, K. Yang, N. Long, J. Bai, and D. Liu. August 2018. Real-time pedes- trian crossing lights detection algorithm for the visually impaired. *Multimedia Tools and Applications* 77(16):20651–71. doi: 10.1007/s11042-017-5472-5.

Cheng, R., K. Wang, K. Yang, N. Long, and H. Weijian. October 2017. Crosswalk navigation for people with visual impairments on a wearable device. *Journal of Electronic Imaging* 26:1.

Chinchole, S., and S. Patel. 2017. Artificial intelligence and sensors based assistive system for the visually impaired people. 2017 International Conference on Intelligent Sustainable Systems (ICISS) SCAD Institute of Technology, Palladam, India . 16–19 doi:10.1109/ISS1.2017.8389401.

Ghilardi, M. C., G. Simões, J. Wehrmann, I. H. Manssour, and R. C. Barros. 2018 Real-Time Detection of Pedestrian Traffic Lights for Visually-Impaired People 2018 International Joint Conference on Neural Networks (IJCNN) Rio de Janeiro, Brazil. 1–8 doi:10.1109/IJCNN.2018.8489516.

Habib, A., M. Islam, M. Kabir, M. Mredul, and M. Hasan. 2019, November. Staircase detec- tion to guide visually impaired people: A hybrid approach. (Retrieved from) *Revued'Intelligence Artificielle* 33 (5):327–34. doi: 10.18280/ria.330501.

Hakobyan, L., J. Lumsden, D. O'Sullivan, and H. Bartlett. September 2013. Mobile assistive technologies for the visually impaired. *Survey of Ophthalmology* 58(6):513–28. doi: 10.1016/j.survophthal.2012.10.004.

Henry, S. (2007). *Just ask: Integrating accessibility throughout design*. Lulu.com. Retrieved from https://books.google.es/books?id=hRnpXbFB06cC

Ihejimba, C., and R. Z. Wenkstern. 2020. Detectsignal: Acloud-based traffic signal notification system for the blind and visually impaired, 2020 IEEE International Smart Cities Conference (ISC2) , 1–6 doi:10.1109/ISC251055.2020.9239004.

Islam, M. M., and M. S. Sadi. 2018. Path Hole Detection to Assist the Visually Impaired People in Navigation 2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT) Bangladesh (Manhattan, New York, U.S.: IEEE) 268–73 doi:10.1109/CEEICT.2018.8628134.

Khan, A., S. Khusro, and J. D. Camba. 2020. An insight into smartphone-based assistive solutions for visually impaired and blind people: Issues, challenges and opportunities. *Universal Access in the Information Society* 1–34. doi:10.1007/s10209-020-00776-x.

Krajnc, E., J. Feiner, and S. Schmidt. 2010. User centered interaction design for mobile applications focused on visually impaired and blind people. In Leitner, G., Hitz, M., Holzinger, A. (eds) HCI in Work and Learning, Life and Leisure. USAB 2010. Lecture Notes in Computer Science. 195–202. Berlin, Heidelberg: Springer-Verlag. doi:10.1007/978-3-642-16607-5_12

Lancioni, G. E., N. N. Singh, M. F. O'Reilly, J. Sigafoos, G. Alberti, V. Chiariello, and L. Carrella. 2020. Everyday technology to support leisure and daily activities in people with intellectual and other disabilities. *Developmental Neurorehabilitation* PMID: 32118503. 23 (7):431–38. doi:10.1080/17518423.2020.1737590.

Li, X., H. Cui, J.-R. Rizzo, E. Wong, and Y. Fang. 2019. Cross-safe: Acomputer vision-based approach to make all intersection-related pedestrian signals accessible for the visually impaired , April. *Arai, K., Kapoor, S. (eds) Advances in Computer Vision. CVC 2019. Advances in Intelligent Systems and Computing* 944 . In (pp. 132146). 132146). Switzerland: Springer International Publishing 978-3-030-17798-0. doi:10.1007/978-3-030-17798-0_13

Mahmud, S., R. Haque Sourave, M. Islam, X. Lin, and J.-H. Kim. 2020. A vision based voice controlled indoor assistant robot for visually impaired people. 2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS) (Manhattan, New York, U.S.) . 1–6 doi:10.1109/IEMTRONICS51293.2020.9216359.

Molina-Cantero, A. J., C. Lebrato-Vázquez, M. Merino-Monge, R. Quesada-Tabares, J. A. Castro- García, and I. M. Gómez-González. 2019. Communication technologies based on voluntary blinks: Assessment and design. *IEEE Access* 7:70770–98. doi:10.1109/ACCESS.2019.2919324.

Montanha, A., M. J. Escalona, F. J. Dominguez-Mayo, and A. M. Polidorio. 2016. A technological innovation to safely aid in the spatial orientation of blind people in a complex urban environment 2016 International Conference on Image, Vision and Computing (ICIVC) . August. Manhattan, New York, U.S.: IEEE 102–107. doi:10.1109/ICIVC.2016.7571281.

Montanha, A., A. M. Polidorio, and M. Romero-Ternero April 2021. New signal location method based on signal-range data for proximity tracing tools. *Journal of Network and Computer Applications* 180:103006. doi: 10.1016/j.jnca.2021.103006.

Qureshi, H. H., and D. H.-T. Wong. 2019. A systematic literature review on user-centered design (UCD) interface of mobile application for visually impaired people, HCI International 2019 – Late Breaking Posters . (pp. 168175). Springer International Publishing, 168–175. doi:10.1007/978-3-030-30712-7_23.

Retorta, M., and V. Cristovão. July 2017. Visually-impaired Brazilian students learning English with smartphones: Overcoming limitations. *Languages* 2(3):12. doi: 10.3390/languages2030012.

Suzuki, L. R. 2017. Smart cities IoT: Enablers and technology road map. In *Smart City Networks: Through the Internet of Things*, 167–90. ISBN: 978-3-319-61313-0. doi:10.1007/978-3-319-61313-0_10. Berlin: Springer International Publishing.

Tian, S., M. Zheng, W. Zou, X. Li, and L. Zhang. 2021. Dynamic crosswalk scene understanding for the visually impaired. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 29 (29):1478–86. doi:10.1109/TNSRE.2021.3096379.

World Health Organization. (2021, October). *Blindness and vision impairment*. Retrieved from https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment

World Wide Web Consortium". (2008, December). *Web content accessibility guidelines 2.0*. Retrieved from https://www.w3.org/

Yang, K., R. Cheng, L. M. Bergasa, E. Romera, K. Wang, and N. Long. 2018. Intersection perception through real-time semantic segmentation to assist navigation of visually impaired pedestrians. In *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)* (Manhattan, New York, U.S.: IEEE), 1034–39 doi:10.1109/ROBIO.2018.8665211.

Yu, S., H. Lee, and J. Kim. 2019. Street Crossing Aid Using Light-Weight CNNs for the Visually Impaired IEEE/CVF International Conference on Computer Vision Workshop (ICCVW) 2019 , October. Manhattan, New York, U.S.: IEEE doi:10.1109/ICCVW.2019.00317.