

Research Article

Intelligent Mobility System for Improving the Blind Pedestrian Independent Behavior in Unknown Outdoor Environments

Armida González-Lorence ¹, Ángel C. Navarrete-Fernández,² Rosana Ayala-Landeros,¹ Juan E. Soto-Osornio,³ and José G. Ayala-Landeros ²

¹*Sistemas y Computación, Tecnológico Nacional de México/Instituto Tecnológico de San Juan del Río, San Juan del Río, Querétaro 76800, Mexico*

²*Ingeniería Industrial, Tecnológico Nacional de México/Instituto Tecnológico de San Juan del Río, San Juan del Río, Querétaro 76800, Mexico*

³*Ingeniería Electrónica, Tecnológico Nacional de México/Instituto Tecnológico de San Juan del Río, San Juan del Río, Querétaro 76800, Mexico*

Correspondence should be addressed to José G. Ayala-Landeros; gabriel.al@sjuanrio.tecnm.mx

Received 14 July 2022; Revised 12 August 2022; Accepted 12 August 2022; Published 12 September 2022

Academic Editor: Zheng Yan

Copyright © 2022 Armida González-Lorence et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The second cause of disability in Mexico is visual impairment; 26% population is blind. Even though blind people maintain a lower quality's life than sighted people, there is not enough social interest to develop comprehensive solutions that improve it. Although many voice-activated emerging technologies use artificial intelligence to get human-machine communication through intelligent virtual assistants, such as Alexa, Siri, Google Home, and Cortana, among others, in reality, there have not been developed specific tools just for blind people that help them to improve their independent behavior because they always depend on other people in doing their daily tasks, more even when they need to move around unknown environments for them. This document reports the development of an autonomous mobility system to be operated by blind people. By audio reconstruction, it is possible to dictate, to the blind pedestrian, in real-time, the presence of traffic lights, crosswalks, and information about their current location. This system employs real-time computer vision tools, artificial intelligence, audio playback systems, and location systems, and it improves the independent behavior of blind people because they could move through unknown environments without the assistance of any sighted person, giving them greater independence and consequently increasing their life's quality.

1. Introduction

Currently, there exist 39 million completely blind people in the world and at least another 246 million people with some visual impairment.

In Mexico, according to figures from the INEGI 2020 Census [1], 16.5% of the total Mexican population has some disability; that is to say, there are 20,838,108 people with disabilities.

Of the 100% disability in Mexico, there are 12,727,653 people with visual disabilities; that is, 61.0% of people with disabilities have visual disabilities [2], which means that this is the most frequent disability in Mexico.

Many attempts have been made to help blind people through various tools, such as braille, the blind cane, or guide dogs. In reality, these tools have many limitations; that is why their mobility is too reduced; their independent behavior is diminished because they depend on somebody also.

The braille only helps them to read texts, and the blind cane has a very limited range just approximately 1.5 meters, and guide dogs are an inaccessible option for many since they require long training periods and their proper care represents a big difficulty for the blind people [3].

In recent years, the use of voice-based technologies on several hardware devices has become very useful; currently,

in many areas, it is usual to control processes and devices using these voice assistants [4–7], such as home applications, automotive industry [8], robotics, emergency warnings, and of course, including the health area [5, 9].

Therefore, this could be the principal way to help people with visual disabilities, because using these technological trends on mobile devices is possible to capture and immediately manage much information, through real-time image processing and IoT devices. Now, it is possible to inform accurately blind people about certain specific information and help them move independently without sight-human support and preserve their physical integrity.

This paper presents an innovative application for blind pedestrians' autonomous mobility in unknown outdoor changing environments; this improves their independent mobility and gives them a higher life's quality through a voice command operated system.

The developed system includes three interconnected components:

- (1) Voice vision application for blind people
- (2) GPS application for blind people
- (3) Smart lenses developed in previous research for indoor environments [10] but are now applied with special focus and features for outdoor spaces

The proposed system represents a low-cost technology option by which blind people can move autonomously and safely in unknown outdoor environments.

2. Materials and Methods

We carry out experimental research, on the streets of San Juan del Río, Querétaro, México. The hypotheses that guide this investigation process were as follows: H0: the X detection rate is less than 99%. H1: the X detection rate is equal or higher to 99%, where X is the traffic light, pedestrian crossing, and pedestrian location.

Here, we describe all the developed procedures. In order to understand the importance of this work, initially, the previous related works and its weakness are described here. After this, all other procedures are described.

3. Related Works

We carried a critical analysis of several closely related previous works.

Since years ago, there have been developed several technologies to help people with different types of disabilities. There are technologies that include advanced artificial intelligence mechanisms, such as Kerdivbulvech [11], who proposes robust gesture interaction methods with high recognition tools to increase interactions between people with diverse not visual impairments or inclusive an integrated application based on innovative technologies like Internet of Things and augmented reality [12] to assist people with diverse disabilities. However, in recent years, the

technology developed specifically for totally blind people is still too much limited, as described below.

The “WeWalk” smart cane was designed by the blind Kursat Ceylan [13] to help blind people to orient themselves and avoid obstacles when walking. It has two parts: (1) the cane and (2) the handle. At the handle, it has a speaker, a microphone, and sensors that allow the blind person to detect and avoid tripping over elements on his way. Its energy autonomy is five hours long, and its approximate price is 400 dollars [13].

Procer is a device to capture images similar to a digital tablet, processes images, and detects everything that is text to convert it to speech. It includes controls with functionalities such as rehearing paragraphs, generating summaries, and recording and exporting a PDF file. In addition, it includes the special functionality of recognizing the money bill denomination [14].

With the “AppStore” for blind people, mobile application developers have been able to take out the “Smartphone” advantage for the benefit of people with reduced vision [15]. It is a simple application that develops audiobook reading and works with voice commands.

“Reading ring” for blind people is a “ring” connected to a computer that interprets and reads text, developed by the MIT Media Laboratory researchers. It uses an algorithm specifically created to identify the words and then read them aloud; the users have to slide their fingers across the reading page while the ring emits sounds and vibrations to inform them when inadvertently change lines [16]. Currently, its creators are developing a new smartphone version, better than their computer version.

All reader is a device completely apart from the computer that clearly reads any printed document. It is a reading system for blind people that contains audio recordings of scanned documents [17] and can read in any language. It is composed of a scanner, two USB ports, a voice synthesizer, a CD drive, and a software system to be used as a digital media player; it allows you to read any type of document, digital or printed, without the help of another sighted person.

GPS Maptic is a device created by Emilios Farrington-Arnas of Brunel University in London [18]; it is a visual sensor that blind people can use as a collar; additionally, it has feedback sensor series placed on clothing and around the wrist like a bracelet, connected to a voice-controlled smartphone app which uses GPS to guide the user's movements through the body's left or right vibrations. However, this project is not for sale or working.

OrCam is another device developed for people with severe visual difficulties and not for blind people, which uses augmented reality tools, developed by OrCam technology, an Israeli company [6]. This device transforms written texts into sound and reads them to people's ears with visual difficulties. A small camera integrated into the glasses' frame recognizes the written sentences and reads them up to the user who cannot see them.

This device is a pair of glasses that allows reading everything that a finger pointed, facilitates the reading of texts, and can distinguish people's faces. However, it only works

in the English language, addressing a serious difficulty for all blind people who are not English speakers. The first version available on the market has an estimated price of about 2,500 dollars [19].

It is installed in the pocket computer [20], and its inventors say that even though we are living the 21st century, the available tools for severe visual impairment people are not adapted to the present, because of their size, complexity, or simply, because they are limited or obsolete methods [19].

Another device called “Smart Glasses” was created by Oxford University researchers, led by Dr. Stephen Hicks; they developed it to help people with severe vision loss. The glasses enhance images of people and objects next to their user through a program designed specifically for this function, giving a much clearer sense of the surrounding environment [16].

These smart glasses have a 3D camera, which captures images, and a computer processes and projects them in real time on small screens located in the commonplace of the normal glasses; in such a way, the surrounding objects become clearer. However, those researchers designed these glasses just for people with severe vision loss and not for totally blind people as we propose in this work.

Another partial solution identified is “the Intelligent Glasses,” created by the Mexican Daniel Martínez Macedo [21]. However, at this moment, it is just a prototype version; he explains that it needs a foundation for investment between USD 160,000 and USD 268,000 to build the real solution with better quality. Martínez estimates that the solution cost for the public would be between USD 160 and USD 321. It works with an integrated camera that captures images, then analyzes the information in them, and then gives details about animals, the environment, and people around. It also helps to read texts and translate them from English to Spanish.

Derived from the analysis of all the anterior works developed to support people with visual disabilities that we described before, we note their serious restrictions, such as these works are focused on people with severe vision loss and not on blind people. These will only partially solve the blind person’s contact with the outside real environment, and also, their development cost is too high.

All mentioned before definitively will not really solve the possibility that a blind person could move in exterior totally unknown environments without the help of a sighted person and preserving his physical integrity. However, it is conceivable to develop a new computer application to help the blind people to move independently because it is possible to generate a link between the user behavior and the new prototype developed in this work, because the blind people always generate mental models and the mental models can be inferred by observing user’s interactions with an application [5].

3.1. System Development. The activities developed to achieve the objectives established in this research are described below.

3.2. Selection of Important Information from the External Environment. Initially, we realize an analysis by making some interviews and coexisting with blind people to identify which

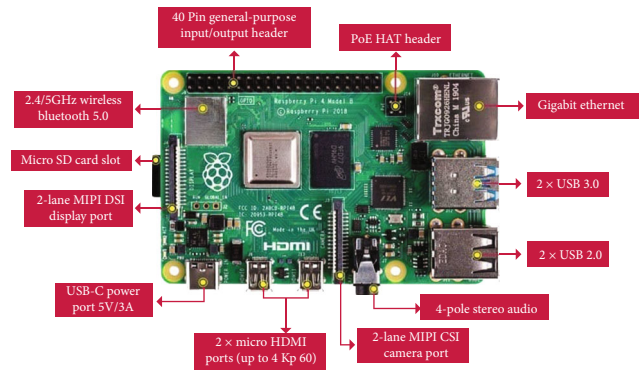


FIGURE 1: Raspberry Pi 4 connections.

is the most important information that the blind user must always know in real time, in order to preserve his physical integrity. From this analysis, we determined that the most important information corresponds to the following listed:

- (i) The presence of traffic lights and the current active light on these
- (ii) The presence of painted crosswalks on the streets
- (iii) Information about the user’s actual location

This information can assist blind pedestrians when they need to move autonomously through outdoor environments, such as streets or avenues.

Although this is the most important information to consider at the moment, there is the serious problem of poor road conditions in most countries; to solve this, deep learning models have been proposed, among which we can mention the called modified U-Net [22] and the downstream model proposed in [23] that purpose of detecting cracks in roads and/or structures. It would be valuable to use models like these to automatically detect problems in streets and avenues and thus be able to inform blind pedestrians, in real time, about the poor condition of the street or avenue that puts their physical integrity at risk.

For the above, we have identified that for later work, it would be very important to include additional information, such as the existence of damaged roads, warning signs, banquettes, steps, and potholes, among other present important warnings at the exterior environment.

3.3. Characteristics and Functionalities of Hardware Selected. A Raspberry Pi 4 board that is an economic minicomputer of the size of a credit card has been chosen as ideal, to work on this project. It provides the required portability and processing capacity and also brings the needed security for future scalability and to maintain the low cost.

Also, it is distinguished by its excellent use for the development of small prototypes [24].

The Raspberry Pi 4 is an ideal dispositive to connect all the devices that need the project to send and get information from the user because it has a big interconnection capability (see Figure 1), by its GPIO pins, USB ports, camera port, HDMI ports, Bluetooth, LAN, and Wi-Fi connections.

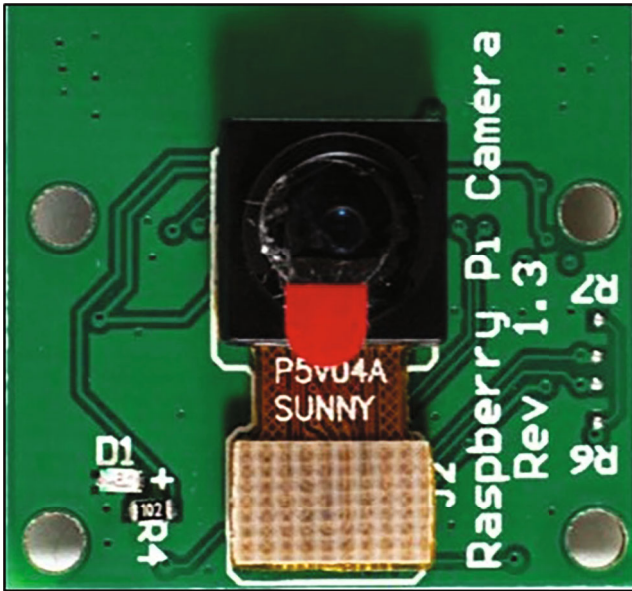


FIGURE 2: Raspberry Pi ECO 5-megapixel camera.

The Raspberry Pi 4's most important characteristics for the development of this work are as follows:

- (i) 1.5 GHz 64-bit quad-core ARM Cortex-A72 CPU (ARM v8, BCM2837)
- (ii) 8 GB RAM (LPDDR4)
- (iii) On-board wireless LAN (dual-band 802.11 b/g/n/ac)
- (iv) On-board Bluetooth 5.0, low-energy (BLE)
- (v) 2x USB 3.0 ports
- (vi) 2x USB 2.0 ports
- (vii) Gigabit Ethernet
- (viii) Power-over-Ethernet (this will require a PoE HAT)
- (ix) 40-pin GPIO header
- (x) 2x micro-HDMI ports (up to 4Kp60 supported)
- (xi) CSI camera port
- (xii) Combined 3.5 mm analog audio and composite video jack

The 64-bit quad-core processor runs multiple processes quickly and supports 4Kp60 hardware video codification in conjunction with its 8 GB RAM capacity. This feature gives the necessary processing speed for the development in real time of the functions that must be executed in this project. It has a dual-band 2.4/5.0 GHz Wi-Fi-LAN that addresses efficient sending and receiving information.

The CSI camera port is ideal to connect a Raspberry Pi ECO 5 megapixel camera (see Figure 2) [25]; we capture in real time the images from the outdoor scenes with this camera.

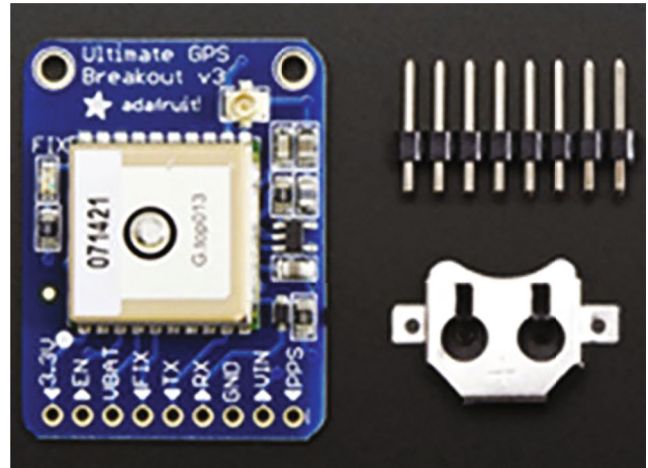


FIGURE 3: SIM900 GPS.

The selected camera can capture 2592×1944 pixels of static images and supports video recording of 1080 p to 30 fps, 720 p to 60 fps, and 640×480 to 60/90, with 5MP OV5647 1080 p camera sensor web; these characteristics are ideal to capture images in real time; it also maintains full compatibility with the Raspberry Pi 4.

The camera characteristics are as follows:

- (i) Sensor: OV5647
- (ii) Still resolution: 5 megapixels
- (iii) Sensor resolution: 2592×1944 pixels
- (iv) Sensor image area: 3.76×2.74 mm
- (v) Pixel size: $1.4 \mu\text{m} \times 1.4 \mu\text{m}$
- (vi) Lens: fixed-focus
- (vii) Field of view: 62°
- (viii) Aperture: F/2.4
- (ix) Video modes: 1080p30, 720p60, and 640×480 p60/90

We use the GPRS (General Packet Radio Service) card (see Figure 3) SIM900 GSM/GPRS for the location acquisition; it is an ultracompact card for wireless communication, and it is Raspberry Pi 4 compatible, configured and controlled via UART (Universal Asynchronous Receiver-Transmitter) with AT commands to send and receive SMS, make calls, and calculate the location in real time through GPS [26].

Technique characteristics are as follows:

- (i) Principal chip: SIM900
- (ii) External supply voltage: 5-12 V DC
- (iii) I/O voltage: 5 V TTL
- (iv) Operating voltage: 1.5 mA

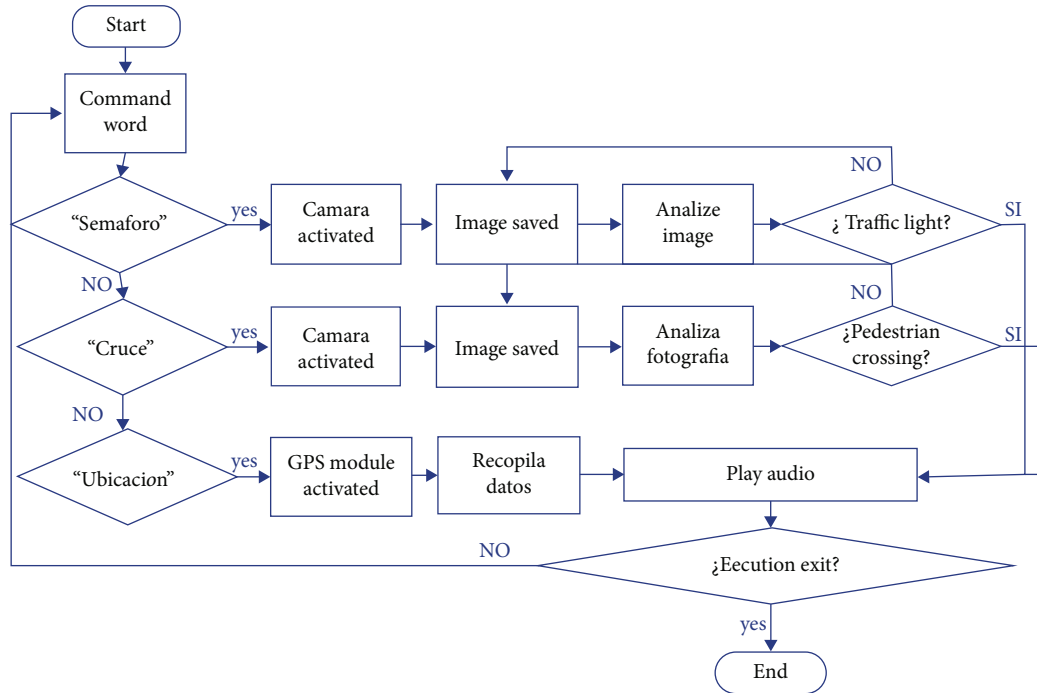


FIGURE 4: System design.

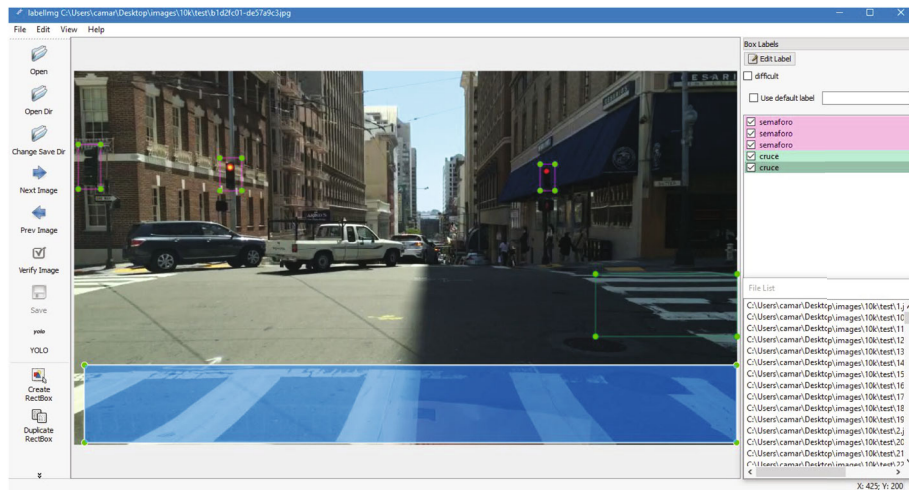


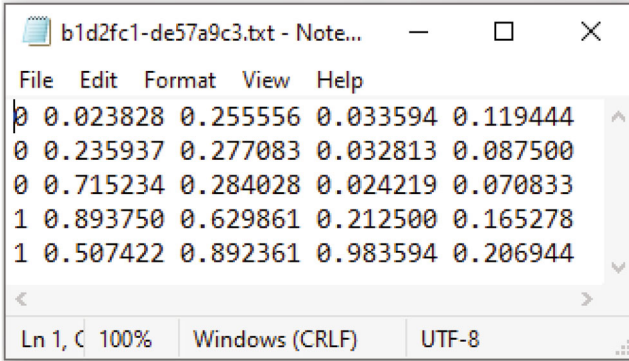
FIGURE 5: Object labelling example.

- (v) Four frequencies GSM/GPRS: 850, 900, 1800, and 1900 MHz
- (vi) Mobile station GPRS class B
- (vii) Supports data service: (850/900/1800/1900 MHz) GPRS
- (viii) Operating temperature: -40°C to +85°C

For sending images dictated by voice to the blind user, we use headphones without noise cancellation, in order to not to private the blind user of additional information from the surrounding, which gives him more important information about what is happening around him.

The headphones used at the project are wireless and connected via Bluetooth to the Raspberry. The headphones chosen are low cost and require only 2 to 3 hours of charge; their battery life lasts 22 hours of music, 24 hours of talk, or 60 days on standby approximately. So the charging time is not excessive. This feature is important because the user could charge the headphone overnight while he is sleeping and can use them all the next day without worries.

The “bee v5.0” 24-hour driving hand-free headphones with microphones (for iPhone, Android, Samsung, and laptop truck driver) meet the need for daily use and are comfortable designed (weight: 12g); these do not cause any load on ears, ideal situation for the project requirements.

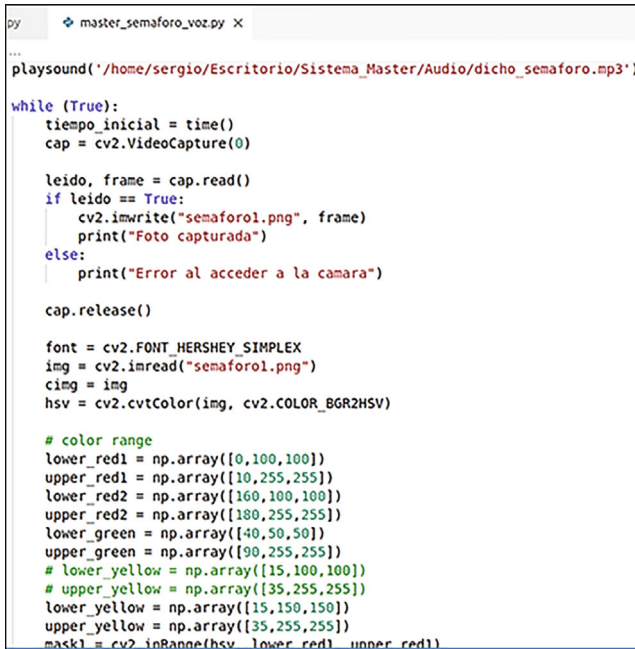


```

File Edit Format View Help
0 0.023828 0.255556 0.033594 0.119444
0 0.235937 0.277083 0.032813 0.087500
0 0.715234 0.284028 0.024219 0.070833
1 0.893750 0.629861 0.212500 0.165278
1 0.507422 0.892361 0.983594 0.206944
Ln 1, C 100% Windows (CRLF) UTF-8

```

FIGURE 6: Labelled object file.



```

...
playsound('/home/sergio/Escritorio/Sistema_Master/Audio/dicho_semaforo.mp3')

while (True):
    tiempo_inicial = time()
    cap = cv2.VideoCapture(0)

    leido, frame = cap.read()
    if leido == True:
        cv2.imwrite("Semaforol.png", frame)
        print("Foto capturada")
    else:
        print("Error al acceder a la camara")

    cap.release()

font = cv2.FONT_HERSHEY_SIMPLEX
img = cv2.imread("semaforol.png")
cimg = img
hsv = cv2.cvtColor(img, cv2.COLOR_BGR2HSV)

# color range
lower_red1 = np.array([0,100,100])
upper_red1 = np.array([10,255,255])
lower_red2 = np.array([160,100,100])
upper_red2 = np.array([180,255,255])
lower_green = np.array([40,50,50])
upper_green = np.array([90,255,255])
# lower_yellow = np.array([15,100,100])
# upper_yellow = np.array([35,255,255])
lower_yellow = np.array([15,150,150])
upper_yellow = np.array([35,255,255])
mask1 = cv2.inRange(hsv, lower_red1, upper_red1)

```

FIGURE 7: Traffic light detection.

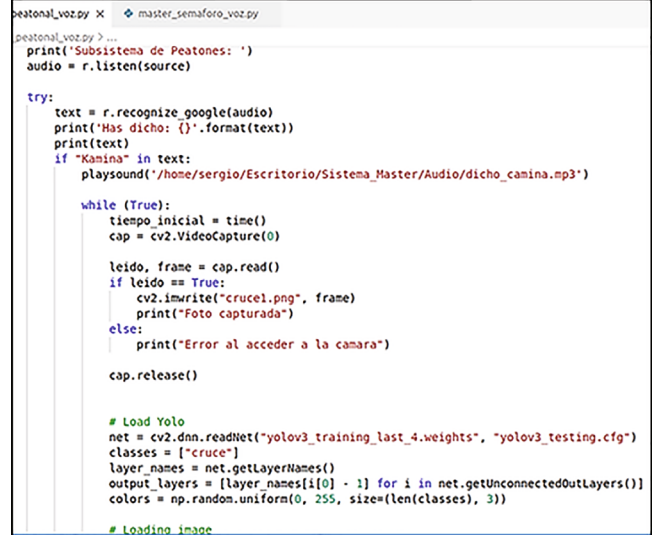
We use a 20,000 milliampere YICF Power Bank to power supply the prototype via USB. This battery contains the characteristics needed for the project, such as overload protection and high-temperature protection. Even, it also has two USB ports, which allow, plus charging the Raspberry Pi 4, the possibility of powering any other device that the blind pedestrian could need.

Finally, we store all the prototype programs developed for the Raspberry in a 60-gigabyte SD memory.

3.4. Characteristics and Functionalities of the Used Software.

The following are the software used:

Python (version 3) is a development software; we chose Python 3 as development software because it contains many libraries, data types, and incorporated functions in the same language. Its multiplatform compatibility is made possible by its interpreted, multiparadigm, dynamically typed, and multiplatform properties. Because it is free software, it helps us keep the cost of the system low. [8, 27]. Also, with Python,



```

peatonal_voz.py x master_semaforo_voz.py
peatonal_voz.py ...
print('Subsistema de Peatonos: ')
audio = r.listen(source)

try:
    text = r.recognize_google(audio)
    print('Has dicho: {}'.format(text))
    print(text)
    if "Kamina" in text:
        playsound('/home/sergio/Escritorio/Sistema_Master/Audio/dicho_camina.mp3')

    while (True):
        tiempo_inicial = time()
        cap = cv2.VideoCapture(0)

        leido, frame = cap.read()
        if leido == True:
            cv2.imwrite("cruce1.png", frame)
            print("Foto capturada")
        else:
            print("Error al acceder a la camara")

        cap.release()

# Load Yolo
net = cv2.dnn.readNet("yolov3_training_last_4.weights", "yolov3_testing.cfg")
classes = ["cruce"]
layer_names = net.getLayerNames()
output_layers = [layer_names[i[0] - 1] for i in net.getUnconnectedOutLayers()]
colors = np.random.uniform(0, 255, size=(len(classes), 3))

# Loading image

```

FIGURE 8: Pedestrian crossing detection.

it is possible to take advantage of all the connection capabilities of the Raspberry GPIO pins [28].

OpenCV (open source) is for managing exterior scene images; this real-time computer vision library [29] has libraries of static and dynamic data types (matrices, graphs, trees, etc.) and preserves high compatibility with various operating systems used around the world, such as Windows, Linux, and of course Raspbian (Raspberry Pi operating system) [30].

OpenCV provides flexible image processing and high-level tools [31] for the development needs of this project.

YOLO (English acronym of You Only Look Once) is an artificial intelligence code system for real-time object detection.

YOLO works with deep learning and CNN (convolutional neural networks); this algorithm can detect all trained objects in an image, and it is possible to train it in a personalized way in order to detect a new variety of objects, so in this project, we trained the model to detect traffic lights and their lights and identify pedestrian crossings.

GPS is an English acronym for Global Positioning System; we use Gisgraphy [32] to get the location. It is a reverse and direct geocoding tool that requires a unique global and consolidated address/POI database (+500 million entries).

3.5. *System Design.* We show the flow diagram of the complete system in Figure 3.

The developed system consists of three main parts: the vision subsystem, which is divided into two parts, the first one to execute the traffic lights identification, the second one to execute the pedestrian crossings identification, and the third one, the subsystem's localization to execute the identification of the user's current location (see Figure 4).

We train the artificial intelligence model to develop the vision subsystem; we use the free open-access tool of Python: "LabelImg" [33]; then, we capture the images to do the object recognition; it is addressed from a 682-image database with real photographs of several streets that contain different

```

def get_url(self):
    """Return the url of the query"""
    gisgraphy_url = GISGRAPHY_BASE_URL
    parameter_counter = 0
    self.parameters[GISGRAPHY_FORMAT_PARAMETER_NAME] = \
    PYTHON_PARAMETER_VALUE
    for parameter_name, parameter_value in self.parameters.items():
        if parameter_counter == 0:
            gisgraphy_url = gisgraphy_url+parameter_name+"="+ \
            str(urllib.quote((parameter_value)))
        else:
            gisgraphy_url = gisgraphy_url+"&" + parameter_name \
            +"="+str(urllib.quote((parameter_value)))
            parameter_counter = parameter_counter +1
    return gisgraphy_url

```

FIGURE 9: Real-time location identification.

```

indexes = cv2.dnn.NMSBoxes(bboxes, confidences, 0.5, 0.4)
print(indexes)
font = cv2.FONT_HERSHEY_PLAIN
for i in range(len(bboxes)):
    if i in indexes:
        x, y, w, h = bboxes[i]
        label = str(classes[class_ids[i]])
        color = colors[class_ids[i]]
        cv2.rectangle(img, (x, y), (x + w, y + h), color, 2)
        cv2.putText(img, label, (x, y + 30), font, 3, color, 2)
        playsound('/home/pi/Desktop/Sistema-Raspberry/cruce.mp3')

cv2.imshow("Image", img)
key = cv2.waitKey(0)

cv2.destroyAllWindows()

```

FIGURE 10: PlaySound "cruce peatonal."

```

# Leer los datos del puerto serial
def getPositionData(gps):
    data = gps.readline()
    message = data[0:6]
    if (message == "$GPRMC"):
        # GPRMC= datos
        parts = data.split(",")
        if parts[2] == 'V':
            # V = no hay señal gps
            print "advertencia GPS sin recepcion "
        else:
            # leer los datos d posicion
            longitude = formatDegreesMinutes(parts[5], 3)
            latitude = formatDegreesMinutes(parts[3], 2)
            print "Tu posición es: lon = " + str(longitude) + ", lat = " + str(latitude)

```

FIGURE 11: Localization acquiring. Source: self-made.

```

def get_url(self):
    """Return the url of the query"""
    gisgraphy_url = GISGRAPHY_BASE_URL
    parameter_counter = 0
    self.parameters[GISGRAPHY_FORMAT_PARAMETER_NAME] = \
    PYTHON_PARAMETER_VALUE
    for parameter_name, parameter_value in self.parameters.items():
        if parameter_counter == 0:
            gisgraphy_url = gisgraphy_url+parameter_name+"="+ \
            str(urllib.quote((parameter_value)))
        else:
            gisgraphy_url = gisgraphy_url+"&"+parameter_name \
            +"="+str(urllib.quote((parameter_value)))
            parameter_counter = parameter_counter +1
    return gisgraphy_url

```

FIGURE 12: Gisgraphy results.

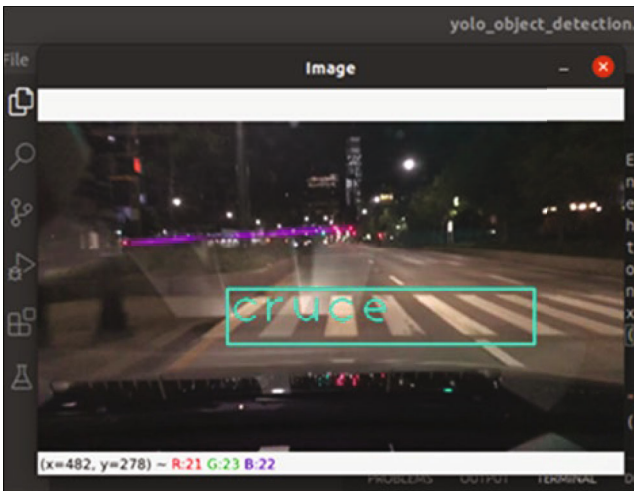


FIGURE 13: Pedestrian crossing recognition example.



FIGURE 15: Developed system.

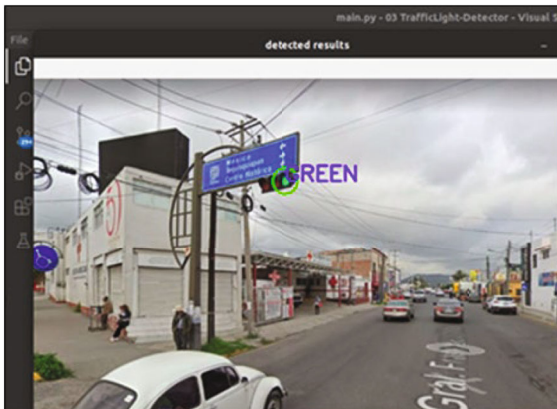


FIGURE 14: Traffic light recognition example.

kinds of traffic lights in operation, with different state's lights (red, yellow, or green) and pedestrian crossings on different perspectives.

To perform the 682-image labelling, we initially created a digital training folder with 682 images. Subsequently, the objects of interest's selection process are iterated for each traffic light and the pedestrian crossing on those images; see an example in Figure 5.

As the object labelling is done, we created a bank of text files, one file for each image where each file contains the coordinates of each identified object in the image.

Figure 6 shows an object labelling example, where index 0 represents the traffic lights identified and index 1 represents the pedestrian crossings identified.

3.6. *Training of the Objects of Interest Identification.* At this stage, we first created a file called "images.zip," which includes the labelled images and their corresponding label files, and then, we save this file in a new folder named "yolov3" in the Google Drive cloud to use it with Google Colab.

At Google Colab, we execute a Python script that makes the automatic learning; it has to be executed with machine learning libraries, fed with the "images.zip" file as an entry until we get the trained file: "yolov3_training_last.weights."

3.7. *Code Programming.* Next, we explain the source code made to perform the main functions of this project.

We show the main function to detect semaphores in Figure 7, with the trained model "yolov3_training_last.weights" and the object class "Semáforo."

Figure 8 shows the source code for pedestrian crossing detection, with the trained model "yolov3_training_last.weights" and the object class "Cruce."

TABLE 1: Tests executed; interest objects identified.

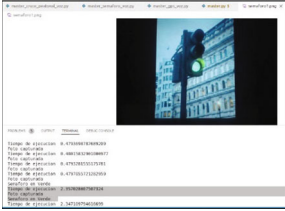



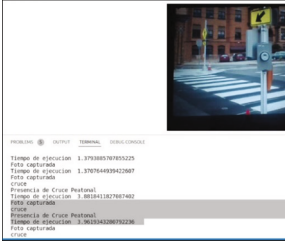
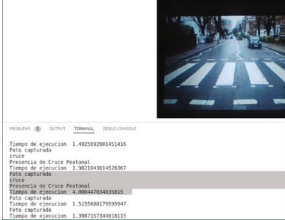
Command word	Image (identified object)	Sound	Help you?
“Semáforo”	<p>Traffic light identified</p> 	“Verde”	Si (yes)
Semáforo”	<p>Traffic light identified</p> 	“Rojo”	Si (yes)
“Cruce”	<p>Pedestrian crossing identified</p> 	“Cruce peatonal”	Si (yes)
“Semáforo”	<p>Traffic light identified</p> 	“Rojo”	Si (yes)
“Cruce”	<p>Pedestrian crossing identified</p> 	“Cruce peatonal”	Si (yes)
“Cruce”	<p>Pedestrian crossing identified</p> 	“Cruce peatonal”	Si (yes)

TABLE 1: Continued.

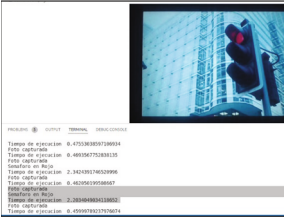
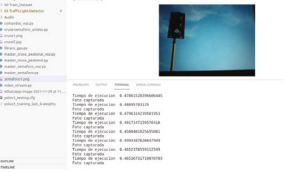

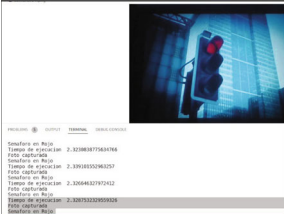
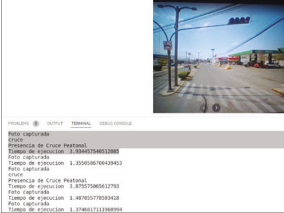

Command word	Image (identified object)	Sound	Help you?
"Semáforo"	<p>Traffic light identified</p> 	"Semáforo Rojo"	Si (yes)
	<p>Traffic light identified</p> 		
"Semáforo"	<p>Traffic light identified</p> 	"Verde"	Si (yes)
"Semáforo"	<p>Traffic light identified</p> 	"Rojo"	Si (yes)
"Cruce"	<p>Pedestrian crossing identified</p> 	"Cruce peatonal"	Si (yes)
	<p>Pedestrian crossing identified</p> 		

TABLE 1: Continued.







Command word	Image (identified object)	Sound	Help you?
	Traffic light identified		
“Semáforo”		“Rojo”	Si (yes)
	Traffic light identified		
“Semáforo”		“Rojo”	Si (yes)
	Pedestrian crossing identified		
“Cruce”		“Cruce peatonal”	Si (yes)
	Traffic light identified		
“Semáforo”		“Rojo”	Si (yes)
	Traffic light identified		
“Semáforo”		“Rojo”	Si (yes)
	Traffic light identified		
“Semáforo”		“Verde”	Si (yes)

TABLE 1: Continued.


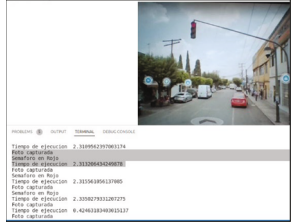
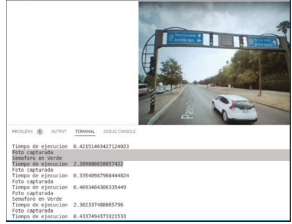
Command word	Image (identified object)	Sound	Help you?
"Semáforo"	<p>Traffic light identified</p> 	"Verde"	Si (yes)
	<p>Traffic light identified</p> 		
"Semáforo"	<p>Traffic light identified</p> 	"Verde"	Si (yes)

Figure 9 shows the source code for detecting the current location of the pedestrian.

The sounds are played with the Python audio library PlaySound [34]. This library contains the "PlaySound()" function that requires a unique argument that specifies the sound file path to play.

PlaySound works with WAV or MP3 files; to play ".wav" files, it uses the from_wav() method, and to play ".mp3" files, it uses the from_mp3() method, or it uses the play() method to play ".wav" and ".mp3" files.

The open-source Python toolkit PyAudio [34] is also used for voice recording and playback functions.

The speech recognition module [35] is a voice recognition library that gives support for various engines and APIs (online and offline). We use this module to recognize the voice dictated by the blind pedestrian.

Figure 10 shows the addition of the function "PlaySound" to the code shown before in Figure 8; the path of the file corresponding audio to pedestrian crossing is indicated: "cruce.mp3."

The same way it was auditioned the PlaySound() to the pedestrian crossing code (exemplified in Figure 9), we made the same preparation in the other function source code to reproduce with voice the current situation of the traffic light ("red," "green," and "yellow") in process.

For the blind pedestrian, current location identification encoding is used by the inverse and direct geocoding tool Gisgraphy [32], previously mentioned.

Figure 11 shows the main code programmed to get the current location; initially, the location is obtained with the GPS and later is made corresponding with the Gisgraphy database (Figure 12) to evaluate the pedestrian current location.

The source codes shown above allow the vision subsystems (traffic lights and crossroads) and GPS location subsystem to work together in real-time, getting information from the environment to sending it to the analyzing system and translating it to the blind pedestrian in clear and understandable voice.

3.8. Identification of Present Elements (in Exterior Scenes) Verification. Initially, the system's functioning is verified with the same images used in the labeling process training model, in order to validate the correct interest element identification (traffic light and pedestrian crossing).

We use Figure 13 to show an example of its correct operation; however, the complete test was driven. By this, we detect the correct identification of all the existing traffic lights and pedestrian crossings in each one of the training images.

Likewise, for each test, the system reproduces the corresponding audio to each identified element; specifically, it reproduces in loud the present light's color ("rojo," "verde," and "amarillo") for each identified traffic light and reproduces in loud the word "cruce" for each identified pedestrian crossing.

TABLE 2: Location identification. Graphic representation of some executed tests.




Test	Time	Real location	Identified location (said in loud)	Help you?	Image
1	3:20 pm	Av. Benito Juárez Ote. 54, Centro, San Juan del Río, Qro.	54, Av. Benito Juárez Ote., San Juan del Río, Qro.	Si (yes)	
2	3:28 pm	Calle Gral. Vicente Guerrero 1C, San Juan del Río, Qro.	5, Calle Vicente Guerrero, San Juan del Río, Qro.	Si (yes)	
3	3:32 pm	Av. Benito Juárez Ote. 9, Centro, San Juan del Río, Qro.	9, Avenida Benito Juárez Ote., Centro, San Juan del Río, Qro.	Si (yes)	
4	3:36 pm	Jardín de la Familia, San Juan del Río	Jardín de la Familia, San Juan del Río	Si (yes)	
5		56 Plaza Morelos Ote., San Juan del Río, Querétaro	4 Vicente Riva Palacio, San Juan del Río, Querétaro	Si (yes)	
6	3:41 pm	Av. Benito Juárez Ote. 29, Centro, San Juan del Río, Qro.	29, Av. Benito Juárez Ote., Centro, San Juan del Río, Qro.	Si (yes)	
7	3:50 pm	Calle Doctor Ruíz Olloqui,1, San Juan del Río	1, Calle Doctor Ruíz Olloqui, San Juan del Río	Si (yes)	
8	3:54 pm	Calle Miguel Hidalgo, 65 C, San Juan del Río	65, Avenida Miguel Hidalgo, San Juan del Río	Si (yes)	
9	4:02 pm	Plaza Morelos Ote., Centro, 76800 San Juan del Río, Qro.	10, Morelos Ote., Centro, San Juan del Río	Si (yes)	

TABLE 2: Continued.


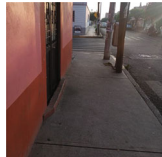



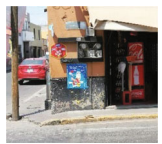




Test	Time	Real location	Identified location (said in loud)	Help you?	Image
10	4:05 pm	Calle José María Morelos 1C, centro, San Juan del Río	José María Morelos, San Juan del Río	Si (yes)	
11	4:08 pm	C. Miguel Hidalgo, Centro, 76800 San Juan del Río, Qro.	Avenida Miguel Hidalgo, San Juan del Río	Si (yes)	
12	4:17 pm	C. José María Morelos 12, Centro, 76800 San Juan del Río, Qro.	12, José María Morelos, San Juan del Río	Si (yes)	
13	4:21 pm	C. José María Morelos 14, Centro, San Juan del Río, Qro.	19C, José María Morelos, San Juan del Río	Si (yes)	
14	4:30 pm	C. José María Morelos 47, Centro, 76800 San Juan del Río, Qro.	47C. José María Morelos San Juan del Río	Si (yes)	
15	4:36 pm	C. José María Morelos 56C, Centro, 76800 San Juan del Río, Qro.	56 José María Morelos, San Juan del Río	Si (yes)	
16	4:43 pm	5 de Mayo 10, Centro, 76800 San Juan del Río, Qro.	10, 5 de Mayo, San Juan del Río	Si (yes)	
17	6:42 pm	Avenida Lomas del Pedregoso 230, INFONAVIT Pedregoso, San Juan del Río, Qro.	228 Avenida Lomas del Pedregoso San Juan del Río	Si (yes)	
18	3:28 pm	Avenida Lomas del Pedregoso 301, Loma de San Juan Sec. Campestre, 76806 San Juan del Río, Qro.	301 Avenida Lomas del Pedregoso San Juan del Río	Si (yes)	
19	7:22 pm	5 Calcita, INFONAVIT Pedregoso, San Juan del Río, Querétaro	5 Calcita, San Juan del Río	Si (yes)	

TABLE 2: Continued.










Test	Time	Real location	Identified location (said in loud)	Help you?	Image
20	5:14 pm	Calle Labradorita 15, INFONAVIT Pedregoso, San Juan del Río, Querétaro	15 Labradorita, San Juan del Río	Si (yes)	
21	5:22 pm	Calle Labradorita 8, INFONAVIT Pedregoso, San Juan del Río, Querétaro	8 Labradorita, San Juan del Río	Si (yes)	
22	5:27 pm	Labradorita 45, INFONAVIT Pedregoso, San Juan del Río, Qro.	56 Labradorita, San Juan del Río	Si (yes)	
23	5:32 pm	92 Labradorita, INFONAVIT Pedregoso, San Juan del Río, Querétaro	92 Labradorita, San Juan del Río	Si (yes)	
24	5:36 pm	Pedregoso 34, INFONAVIT Pedregoso, San Juan del Río, Querétaro	34 Pedregoso, San Juan del Río	Si (yes)	
25	5:37 pm	Pedregoso 4, INFONAVIT Pedregoso, San Juan del Río, Querétaro	4 Pedregoso, San Juan del Río	Si (yes)	
26	11:23 am	Avenida Lomas de San Juan 570, Lomas de San Juan, 76806 San Juan del Río, Qro.	570 Avenida Lomas de San Juan, San Juan del Río	Si (yes)	
27	11:26 am	Avenida Lomas de San Juan 151, Lomas de San Juan, 76806 San Juan del Río, Qro.	151 Avenida Lomas de San Juan, San Juan del Río	Si (yes)	
28		Avenida Lomas de San Juan 38, Lomas de San Juan, 76806 San Juan del Río, Qro.	38 Avenida Lomas de San Juan, San Juan del Río	Si (yes)	
29	8:03 am	162 Jurica-San Juan del Rio, San Juan del Río, Querétaro	3 Belisario Domínguez, San Juan del Río, Querétaro	Si (yes)	

TABLE 2: Continued.

Test	Time	Real location	Identified location (said in loud)	Help you?	Image
30	8:00 am	Avenida Benito Juárez Ote. 76807, Centro, 76807 San Juan del Río, Qro.	143, Avenida Benito Juárez Ote., San Juan del Río	Si (yes)	

Subsequently, we took the developed system to outdoor spaces in the city, in order to perform the experimental tests in real situations. Figure 14 shows an example of the gotten results. Here, we can appreciate the correct identification of the traffic lights. The system reproduces in loud the “verde” audio from the green light because this sound is the corresponding audio for the current state of the traffic light identified.

Similar to the example shown before, we developed another 100 tests. These tests help us to verify that the results are persistent and reliable at all times.

Subsequently, a series of positioning tests were carried out in different city places. With those tests, we observed that 99% of the tests coincide with the real location; the GPS works correctly and sends the precise location data to inform users of their localization.

4. Results and Discussion

The complete system developed in this research is shown in Figure 15. With this system, it is possible to guide the blind pedestrian in unknown outdoor environments getting the goal of this project: improving their independent behavior.

The blind pedestrian independence is achieved by informing him about if he is facing some traffic light (the current state of the light is said in loud), or if he is facing some “pedestrian crossing” (the word “cruce” is said in loud). Also, this system dictates in loud the information about the user’s current location. The system works each time the blind pedestrian requests it, through voice commands of the trigger words above-mentioned which gives him the confidence to always use this system, giving him consequently a behavior of higher independence through avoiding others help.

In short, with these results of this project tests, we can show that it is possible to guide the blind pedestrian with the aforementioned information, through audio reproductions aloud.

The system works in real time, and it provides the blind pedestrian real information about the exterior changing situations that are important for him, in order to move autonomously through the streets and/or avenues of the city.

Initially, when the execution of the developed system starts, it begins in a standby mode until any command word is listened to and activates, like “Semaforo,” “Cruce,” and “Ubicación”; the Bluetooth integrated microphone in the headset receives these command words; once the voice recognition is done, the system activates any of the 3 actions:

traffic light recognition, pedestrian crossing recognition, or current location identification.

Traffic lights, pedestrian cruises, and location were carried out between 9:00 am and 6:00 pm, with natural lighting. The number of tests performed for the traffic lights and pedestrian crossings was 957 and 700 tests for the location. In all cases, no faults were found in the system; the probability of success was 100% ($p = 1.0$). For these results, there is an error I type of zero; it was not necessary to estimate probabilities with the binomial distribution because there were no failures in the system. Therefore, the null hypothesis is rejected and the alternative accepted which establishes that the detection rate is greater than 99%.

In Table 1, we graphically present some of the 957 tests developed in different situations where the identification of the correct object is unvariable acquired (traffic light and pedestrian crossings), and in each proof, the blind user was asked (“Helped you?”) about his appreciation of the system; he answered that it was a useful system, even he said was reliable (see Tables 1 and 2).

Table 2 presents some tests of the 700 different situations that were correctly acquired as the current pedestrian location. It is necessary to explain that the column called “real location” is captured manually by a human tester, and the column “identified location (said in loud)” is automatically captured from the GPS system; however, although there are some differences on the captured text, in reality, it always corresponds to the same location.

5. Conclusions

Through the results of this work, it is possible to demonstrate that the independency of blind people could be higher if there were a bigger social interest in helping them, because, here, using IoT technology, artificial intelligence tools, and image processing, it was possible given them a better independent behavior with the system: “Intelligent Mobility System for improving de Blind Pedestrians independent behavior in unknown outdoor environments (integrated components: traffic lights, pedestrian crossings, and location).” Now it is possible to help the blind pedestrian to move independently in an unknown outdoor environment (streets/avenues). By providing the user with information about his surroundings in real-time via loud audio.

The vision subsystems (traffic lights and pedestrian crossings), the voice recognition and audio playback subsystem, and the current location identification subsystem complete the whole system. These subsystems complete the whole system. The tests carried out during the development,

likewise; the tests developed in the testing stage generated satisfactory results. With those results, it is possible to conclude that, using just low-cost tools and open-source software in this investigation, it was possible to generate a pertinent and alternative solution to solve a social problem that thousands of blind people experience daily.

This system represents a technological advance to improve the quality of life of blind pedestrians, since it allows them to behavior autonomous, to take their routes with greater safety and independence; getting real-time information from the external environment allows them to take care of their physical integrity without the help of any clairvoyant. And as Lee and Malcein establish that in the context of emerging technology and automation, people create mental models [5], so it is good to have a specific application for blind people, because then it is a good starting point for defining the new roles for related humans and for designing safe and acceptable an intelligent system just for blind people.

Currently, the work team is developing a glass model, printed in 3D, with the needed characteristics to embed safely the camera used and provide greater comfort to the user; this work is being realized by taking care of maintaining the low cost of the system; the actual system developed costs just \$4,900.00 Mexican pesos (\$245 USD).

However, it is a critical need to continue with this work, to improve the developed system in this research, in order to probe the system in other whether conditions as cloudy and rainy day and to add new functionalities to support the blind pedestrian, such as the identification and recognition of surrounding people, dangerous situation identification, identification of public transport routes or identification of interest places, and many other capabilities to improve their independent behavior.

Data Availability

The data used to support the findings of this study are available from the first author (armida.gl@sjuanrio.tecnm.mx) upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

A special thanks to the efforts of the undergraduate engineering students, María del Rosario Velasco Herrera, Sergio Cardoso Banda, and Martín Gudiño Sánchez, who collaborated in the development of this system. The authors express their gratitude about financing support of TecNM (Tecnológico Nacional de México) to develop this project, under the project registered as “Sistema de Movilidad Autónoma para Peatones Ciegos en Ambientes Desconocidos,” with code number 13237.21-P.

References

- [1] Censo 2020, “Instituto Nacional de Estadística, Geografía e Informática, (2000). XII Censo General de Población y Vivienda. México: INEGI. Instituto Nacional de Estadística, Geografía e Informática,” <http://www.inegi.gob.mx/>.
- [2] INEGI, “La discapacidad en México,2020. Obtenido de Instituto Nacional de Estadística y Geografía,” 2020, http://internet.contenidos.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/702825090203.pdf.
- [3] M. Quintana, *Áreas específicas para la atención de personas con discapacidad visual*, Secretaría de Educación y Cultura, Chihuahua, 2011.
- [4] A. Byalpi and Aush, “Alexa Based Real-Time Attendance System,” *3rd International Conference on Communication and Electronics Systems*, 2018, Coimbatore, India, 2018, 2018.
- [5] Y.-C. Lee and L. A. Malcein, “Users’ mental models for computer-mediated communication: theorizing emerging technology and behavior in eHealth applications,” *Hum Behav & Emerg Tech.*, vol. 2, no. 4, pp. 354–366, 2020.
- [6] B. Maharjan, J. Li, J. Kong, and C. Tao, “Alexa, what should I eat? A personalized virtual nutrition coach for Native American diabetes patients using Amazon’s smart speaker technology,” in *IEEE International Conference on E-health Networking, Application & Services*, Bogota, Colombia, 2019.
- [7] M. Schuager, “Aipoly, una app de reconocimiento de imágenes para personas con discapacidad visual,” 2016, <https://www.whatsnew.com/2016/03/28/aipoly-una-app-dereconocimiento-de-imagenes-para-personas-con-discapacidad-visual/>.
- [8] J. Solorio, J. Garcia-Bravo, and B. A. Newell, “Voice activated semi-autonomous vehicle using off the shelf home automation hardware,” *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 5046–5054, 2018.
- [9] J. Clemente, F. Li, M. Valero, and W. Song, “Smart seismic sensing for indoor fall detection, location, and notification,” *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 2, pp. 524–532, 2020.
- [10] A. González-Lorence, G. Sánchez-Orduña, G. Ayala-Landeros, and S. E. Román- Flores, “Smart glasses for blind people,” *Journal of Technological Prototypes ECORFAN*, vol. 32, pp. 32–38, 2020.
- [11] C. Kerdvibulvech, “A review of computer-based gesture interaction methods for supporting disabled people with special needs,” vol. 9759, pp. 503–506, 2016.
- [12] C. Kerdvibulvech, “A novel integrated system of visual communication and touch technology for people with disabilities,” *ICCSA*, vol. 2, no. 9787, pp. 509–518, 2016.
- [13] BluRadio, “Así es WeWalk, el bastón inteligente para ayudar a invidentes,” 2019, <https://www.bluradio.com/lanube/asi-es-wewalk-el-baston-inteligentepara-ayudar-invidentes-204043-ie5134696>.
- [14] G. Origlia, “Prócer, el dispositivo que transforma texto en voz está ya a la venta,” 2018, <https://www.lanacion.com.ar/sociedad/procer-el-dispositivo-quetransforma-texto-en-voz-esta-ya-a-la-venta-nid2099376>.
- [15] Y. Qu, S. Du, S. Li, Y. Meng, L. Zhang, and H. Zhu, “Automatic permission optimization framework for privacy enhancement of mobile applications,” *IEEE Internet of Things Journal*, vol. 8, no. 9, 2019.

- [16] J. Barbuzano, “5 ejemplos de tecnología para personas ciegas: más allá del Braille,” 2015, <https://www.bbvaopenmind.com/tecnologia/innovacion/tecnologia-parainvidentes-mas-alla-del-braille/>.
- [17] C. A. Hermida, “Tecnología asistiva para ciegos,” 2013, <http://www.youngmarketing.co/tecnologiaasistiva-para-ciegos/>.
- [18] E. Farrington-Arnas, *MAPTIC*, London, 2020 <https://emilios.co.uk/portfolio-maptic.html>.
- [19] ORCAM, “ORCAM,” 2018, <https://www.orcam.com/es/>.
- [20] Infobae, “Desarrollan en México lentes inteligentes para débiles visuales,” *Obtenido de Infobae*, vol. volume II, 2018 <https://www.infobae.com/america/mexico/2018/02/07/desarrollan-enmexico-lentes-inteligentes-para-debiles-visuales/>.
- [21] Infobae, “Volvieron las Google Glass en una versión más potente,” 2017, <https://www.infobae.com/tecno/2017/07/19/volvieron-las-google-glassen-una-version-mas-potente/>.
- [22] T. Siribornratanakul, “A deep learning based road distress visual inspection system using modified U-Net,” in *HCI International 2021- Late Breaking Papers: HCI Applications in Health, Transport, and Industry HCII 2021. Lecture Notes in Computer Science*, Springer, Cham, 2021.
- [23] T. Siribornratanakul, “Downstream semantic segmentation model for low-level surface crack detection,” *Advances in Multimedia*, vol. 2022, Article ID 3712289, 12 pages, 2022.
- [24] Teach (2020), “Learn, and make with Raspberry Pi,” <https://www.raspberrypi.org/>. Obtained on May 13th, 2020.
- [25] L. GNU, *Raspberry Pi*, 2019, https://es.wikipedia.org/wiki/Raspberry_Pi.
- [26] K. Townsend, “Adafruit Ultimate GPS with gpsd,” *Adafruit Industries*, 2021.
- [27] M. Á. Abellán, “¿Qué es Python?,” 2004, <https://www.programoergosum.com/cursos-online/raspberrypi/244-iniciacion-a-python-en-raspberry-pi/que-es-python>.
- [28] M. A. Alvarez, “Qué es Python,” 2003, <https://desarrolloweb.com/articulos/1325.php>.
- [29] P. García, *Reconocimiento de imágenes utilizando redes neuronales artificiales*, Universidad complutense de Madrid, Madrid, 2013.
- [30] Unipython, “Introducción Y procesamiento de imágenes en Python con OpenCV,” 2019, <https://unipython.com/introduccion-procesamiento-imagenespython-opencv/>.
- [31] G. Jonathan, “Reconocimiento de los objetos utilizando OpenCV y Python Thesis,” 2017.
- [32] L. Obuhov and E. Panidi, “Toward Correctness Control of Postal Addresses Geocoding,” *InterCarto InterGIS*, vol. 27, no. 2, pp. 114–127, 2021.
- [33] P. A. Letnes, “tzutalin/labelimg,” 2021, <https://github.com/tzutalin/labelimg>.
- [34] Z. Wang, D. Chen, and P. Xiao, “Design of a voice control 6DoF grasping robotic arm based on ultrasonic sensor, computer vision and Alexa voice assistance,” in *10th International Conference on Information Technology in Medicine and Education*, Qingdao, China, 2019.
- [35] J. Zhang, C. Muñoz, and L. H.-L. Como, “redes sociales, & 2020. (2020) Oportunidades y retos para la comunicación integrada y comercial en redes sociales de mensajería: los casos de WhatsApp y WeChat Academia Edu,” 2020, <http://www.academia.edu/download/62125595/Libro-01-v11-HIGH-PRINT-DEF20200217-34468-ei379i.pdf#page=439>.