The 8th IEEE International Conference on E-Health and Bioengineering - EHB 2020
Grigore T. Popa University of Medicine and Pharmacy, Web Conference, Romania, October 29-30, 2020

Glasses Connected to Google Vision that Inform Blind People about what is in Front of Them

Michael Cabanillas-Carbonell[1], Alexander Aguilar Chávez [2], Jeshua Banda Barrientos[3]
Faculty of Engineering and Architecture, Universidad Autónoma del Perú, UA
Lima, Perú
mcabanillas@ieee.org[1], aaguilar23@autonoma.edu.pe[2], jbanda@autonoma.edu.pe[3]

***Abstract*— More than a billion people around the world have vision problems for different reasons, and these numbers are increasing every year. This leads us to make different innovations in the field of computer vision, with the aim of providing a better quality of life for these people. In this document we present as a resource, the development of an intelligent lens, which incorporates a Raspberry Pi ZW connected to the Google Cloud Vision API through the Wifi of the user's mobile phone, where at the touch of a button, the Raspberry camera, captures the image, processes it in a few seconds and retrieves its main features, obtaining important information for mobilization such as: pedestrian crossings, bus stop sign, vehicles, green light, etc. 150 people from the National Union of the Blind of Peru (NUBP) were evaluated with different degrees of blindness, obtaining a 40.5% increase of independence for their mobilization.**

Keywords—smart glasses, Google Cloud Vision, design thinking, blind people.

I. Introduction

According to the World Health Organization (WHO), in its October 2019 report, at least 2.2 billion people worldwide are visually impaired or blind, and according to reports from previous years, these numbers are increasing, affecting the most vulnerable people the most, such as the elderly, people with disabilities, ethnic minorities, indigenous populations and those on low incomes. [1]

In Perú about 160,000 people are blind for various reasons, the main cause of blindness is bilateral cataract with 47%, followed by uncorrected refractory errors with 15%, glaucoma with 14% and diabetic retinopathy with 5%. Thirty-seven percent of those who are blind due to cataracts are distributed in the Sierra and Selva regions, mostly in rural areas; the remaining 63% are located in urban-marginal areas of the Coast, including Lima and Callao. The prevalence of blindness in the country is about 0.6% if people with severe visual impairment (vision <20/200) are included [2], and most cases are preventable, treatable or recoverable.

This document aims to provide a better quality of life for people with visual impairment of the National Union of the Blind of Peru (NUBP) [3], through the development of smart glasses connected to Google Cloud Vision. Many of these people with blindness do not have the necessary means or economic availability to be able to acquire devices that help them improve their quality of life, even more so since most of them are over 50 years old.

There are different devices that help blind people to move around fluently, however, many of these have not been very successful due to the lack of support when developing this type of project. One of the first devices to appear was the white cane and guide dog [4]; with technological advances, new tools are being created such as the use of Google Cloud Vision in assistive technology scenarios [5], where the benefits of this Google tool, remote processing of images in the cloud and audio output on a speaker are presented; Today, devices have been created that further facilitate the daily activity of a person, as is the case of Google Assistant implementation in Raspberry Pi [6], processor that is used in this document, connecting it with the Api of Google Cloud, focusing on blind people.

Based on the assumption of such scenarios, the development of the first prototypes of smart glasses has started. Based on the new innovative technologies, the aim is to create a non-invasive and economic device that will provide a better way of life for the visually impaired, enabling these blind people to function outside their homes more independently and autonomously.

An embedded lens prototype will be developed with Rasberry Pi Zero W, this family of raspberries is considered one of the most ideal for IOT because of its low energy power, according to an evaluation of the performance of rasberry Pi Zero W [7], since our user will carry the lenses in his daily life, so their energy consumption should be very small; will be connected to a camera in the front, in which the user will press a button, the camera will take a picture, will process it remotely by means of the RES Google Vision API [8] and will inform by audio, connected by means of bluetooth, the result of the image processing, being able to detect between places, logos, letters in front, gestures in the faces, etc. The prototype is noninvasive, allowing the user to walk by the streets with more independence allowing him to make an image capture at will and to listen to the result of the processing in his hearing aids.

This report is structured in four parts. The next part presents the methodology used in the development of the research showing the initial evaluations, materials to be used, programming and tests carried out; the third part shows the results obtained and finally presents the conclusions of this report.

II. METHODOLOGY

For the development of the project, the Design Thinking methodology will be used. Design Thinking, consists of a series of steps forming a development method to generate innovative ideas, focusing on learning and solving the real needs of users. [9]

1. ***Understanding and Observing***: Discovering people (Empathy), the study was conducted with the people of the National Union of the Blind of Peru (NUBP) located at Plaza Bolognesi N° 477 - 479, Cercado Lima - Peru.

Fig. 1. Collection of information

We selected 150 people and used the International Classification of Diseases (2018) [10]- WHO, which classifies visual impairment into two groups according to the type of vision: far and near.

TABLE I. CLASSIFICATION OF VISUAL IMPAIRMENT

Impaired distance vision:	Volunteers
Mild: visual acuity less than 6/12	12
Moderate: visual acuity less than 6/18	31
Severe: visual acuity less than 6/60	42
Blindness: visual acuity less than 3/60	57
Impaired near vision:	
Visual acuity less than N6 or N8 at 40cm with existing correction	8
TOTAL	**150**

2. Find Patterns (Define the Problem): Interviews will be conducted on their autonomy to move on the streets without the use of the research lenses, their requirements will be consolidated.

TABLE II. REQUIREMENTS SURVEY

	Indicators	YES	NO
1	Needs assistance from a person to get around?	132	18
2	You must consult a person about the pedestrian walkway?	121	29
3	You need to check to see which vehicle you need to board?	148	2
4	Requires traffic light information?	139	11
5	You need to check to see if you are at an whereabouts?	133	17
6	You should check to see if a person is happy, sad or upset?	87	63
7	Need help getting on the streets?	144	6
8	You need additional instruments with audio features to get around?	148	2

3. ***Devise the possible solutions (design principles):*** The first research prototype is configured according to the suggestions given; the requirements of the prototype are established. For the development of the application and the prototype, the following materials were implemented:

- Glasses: Preferably medium size and lightweight plastic tool for comfort, this will support the wiring system.
- Raspberry Pi ZERO W: Used as our computer, it is very small, ideal for IOT projects, composed of an ARM processor, wifi connection and bluetooth. [11]
- Cámara Raspberry Pi: This high definition (HD) Raspberry Pi camera board connects to the Raspberry Pi W for photo capture. [12]
- Google Cloud Vision API: A powerful image recognition tool, this API recognizes and classifies images quickly, is able to detect individual objects and faces, and reads printed words that appear in images. [13].
- Internet: The user will provide internet access through his mobile phone by means of the wifi integrated in the Rasberry pi Zero W
- Battery. Integrated into the embedded system, considering that the Rasberry pi Zero W is a low power consumption hardware very useful for IOT projects.

4. ***Prototype the models (make them tangible):*** Tests are carried out with the prototype already created and configured.

A. Lens prototype

Diagram of the components incorporated into the lens:

Fig. 2. Lens prototype

B. Prototype architecture

Diagram of how the prototype is developed and interacts with the different components.

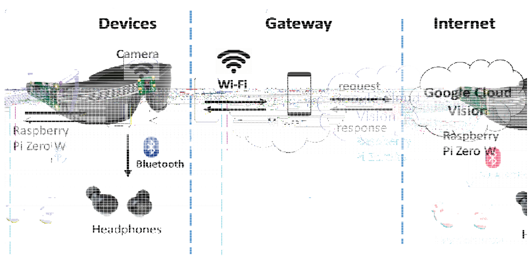

Fig. 3. Prototype architecture

The prototype is activated when the user presses the image capture button, the Raspberry pi W commands the camera to take a picture of whatever is in front of it at that moment, the information of the captured image is sent to the Google Cloud Vision RES API via the Wi-Fi of the user's cell phone. The image is processed and equivalents are searched in the database of Google Cloud Vision, the response is sent in a few seconds again to the Rasberry pi Zero W, which is programmed to send this result through the integrated bluetooth to the headphones previously configured in our processor.

C. System Configuration

1. *Configuration of the environment:*

We created a virtual machine hosted on the wifi ip to be able to enter remotely:

Sudo vncserver :2 –geometry 1920x1080 –depth 24

We updated Python3. We downloaded and installed the Google Vision API. Installed a Python library - installed a new version of PiCamera according to the model of the camera we are using:

python3 –m pip install --user pip
python3 –m pip install --user google-cloud-vision
python3 –m pip install --user Pillow
python3 –m pip install --user picamera

We must have the JSON file in the directory we have downloaded from the project created previously in google vision. We execute the Python script to recognize objects from the capture made by the camera.

2. *Development of the system:*

Next, we will show some of the phyton codes used, for example the geolocation:

```
import picamera
from google.cloud import vision

client = vision.ImageAnnotatorClient()
image = 'image.jpg'

def takephoto():
    camera = picamera.PiCamera()
    camera.capture(image)

def main():
    global image
    takephoto() # First take a picture
    """Run a label request on a single image"""

    with open(image, 'rb') as image_file:
        content = image_file.read()

    image = vision.types.Image(content=content)

    response = client.landmark_detection(image=image)

    landmarks = response.landmark_annotations
    print('Landmarks:')

    for landmark in landmarks:
        print(landmark.description)
        for location in landmark.locations:
            lat_lng = location.lat_lng
            print('Latitude {}'.format(lat_lng.latitude))
            print('Longitude {}'.format(lat_lng.longitude))


if __name__ == '__main__':
    main()
```

Fig. 4. Python code for Geolocation detection

D. Testing our prototype:

This information collected is very valuable for the autonomous mobilization of blind people.

i. As a support to know their location (geolocation)

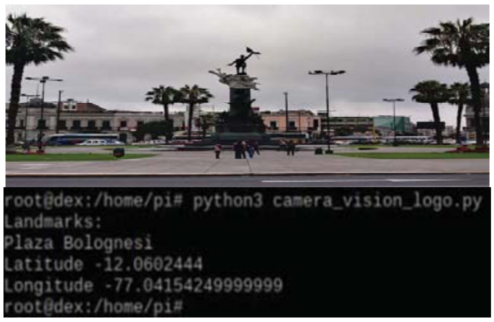

Fig. 5. (above) Photo taken from Plaza Bolognesi - Lima, (below) result of image processing.

The Json report correctly reports by audio "Plaza Bolognesi".

ii. To help you get to know your vehicle (object recognition)

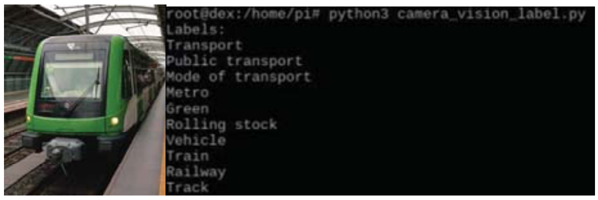

Fig. 6. (left) Photo taken at Lima's subway station, (right) result of image processing.

It correctly identifies the "Public trasnport", “Metro”, “Green”, and “Vehicle”, indicating these characteristics by audio.

iii. As a support in the recognition of vehicle whereabouts (object logotipe)

Fig. 7. (left) Photo taken image of a bus stop sign, (right) result of image processing.

The Json report reports audio "Parador".

iv. As a support in pedestrian crossing recognition

Fig. 8. (left) Photo taken of a pedestrian crossing, (right) result of image processing.

The Json report reports audio "Line", “Pedestrian crossing”.

v. As a support for traffic light recognition:

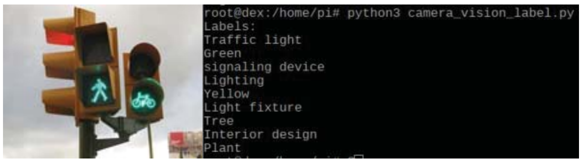

Fig. 9. (left) Photo taken from a traffic light at a green light, (right) result of processing the prototype.

The Json report reports audio "Traffic light", “Green”.

vi. Recognition of a person's state of mind:

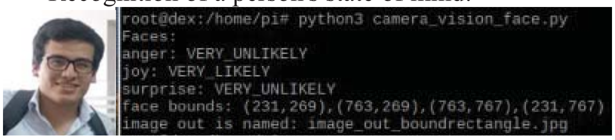

Fig. 10. (left) Photo of a person smiling, (right) result of processing the prototype.

The Json response evaluates the image coordinates by considering it to be 100 x 100 pixels, forming a boundary polygon around the entire image. The result of this image indicates very unlikely for anger and surprise, and very likely for joy. The very probable result is informed by audio to the user.

1. Evaluate and test the products: *We evaluate that the prototype achieves its objective, following up each user for approximately 5 hours during one week, obtaining our results and conclusions.*

III. RESULTS

The 150 volunteers were surveyed again according to Table II and the following results were obtained:

A comparison between Tables II (without the use of the lenses) and Table III (after the use of the lenses) shows an increase in "NO" responses, which are an indicator of autonomy. The comparison of the range responses is shown below:

TABLE III. REQUIREMENTS SURVEY

	Indicators	YES	NO
1	Needs assistance from a person to get around	109	41
2	You must consult a person about the pedestrian walkway	22	123
3	You need to check to see which vehicle you need to board	125	25
4	Requires traffic light information	72	78
5	You need to check to see if you are at an whereabouts	112	38
6	You should check to see if a person is happy, sad or upset	10	140
7	Need help getting on the streets	101	49
8	You need additional instruments with audio features to get around	9	141

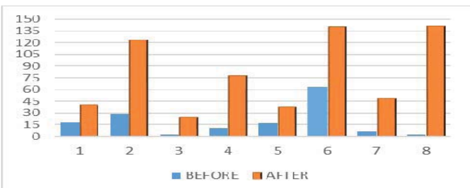

Fig. 11. Comparativo de increment de las respuestas de autonomía

IV. CONCLUSIONS

You get a non-invasive lens, aesthetic, portable and easy to use, achieving the implementation of the connection with Google Cloud Vision for effective capture, processing and audio output to the user, achieving the detection of the necessary images, generating greater autonomy in the mobilization of blind users.

Given the results obtained, the indicators of autonomy are improved by 40.5%; likewise, the persons evaluated indicated that the prototype of the intelligent lens helps to become independent from guide persons and that they would use the product.

REFERENCES

[1] WHO, «WHO launches first World report on vision,» WHO, 8 oct 2019. [En línea]. Available: https://www.who.int/es/news-room/detail/08-10-2019-who-launches-first-world-report-on-vision.

[2] “In Peru about 160,000 people are blind for various reasons”, Andina, 17 nov 2014. [Online]. Available: https://andina.pe/agencia/noticia-en-peru-cerca-160000-personasson-invidentes-diversas-causas-531943.aspx.

[3] “NATIONAL UNION OF THE BLIND OF PERU – NUBP”, [Online]. Available: https://uncp.pe/. [Último acceso: 14 Jan 2019].

[4] A. Garcia, “Products that make life easier for blind people”, Consumer, 05 May 2009. [Online]. Available: http://www.consumer.es/web/es/solidaridad/proyectos_y_campanas/2009/05/05/185103.php.

[5] D. Mulfari, A. Celesti, M. Fazio, M. Villar y A. Puliafito, Using Google Cloud Vision in Assistive Technology Scenarios, IEEE Workshop on ICT solutions for eHealth , 2016.

[6] S. Mischie, L. Mâțiu-Iovan y G. GăŠpăresc, «Implementation of Google Assistant on Rasberry Pi,» IEEE - International Symposium on Electronics and Telecommunications (ISETC), 2018.

[7] D. B. C. Lima, R. M. B. d. S. Lima, D. d. F. Medeiros, R. I. S. Pereira, C. P. d. Souza y O. Baiocchi, «A Performance Evaluation of Raspberry Pi Zero W Based Gateway Running MQTT Broker for IoT,» IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2019.

[8] Google Cloud Vision, «Get started with Vision AI, » [En línea]. Available: https://cloud.google.com/vision/overview/docs/get-started. [Último acceso: 18 jan 2020].

[9] “Design Thinking”, Dinngo, [Online]. Available: https://designthinking.es/inicio/index.php. [Último acceso: 12 Aug 2019].

[10] WHO, « Blindness and visual impairment » WHO, 20 mar 2020. [En línea]. Available: https://www.who.int/es/news-room/fact-sheets/detail/blindness-and-visual-impairment.

[11] Raspberrypi, «Raspberry Pi Zero W,» raspberrypi, [En línea]. Available: https://www.raspberrypi.org/products/raspberry-pi-zero-w/. [Último acceso: jan 2020]

[12] “Cámara Raspberry Pi v2 - 8 Megapixels”, BricoGeek, [En línea]. Available: https://tienda.bricogeek.com/accesorios-raspberry-pi/822-camara-raspberry-pi-v2-8-megapixels.html.

[13] Api, google cloud, “Google Cloud Vision API”, ackstorm, 11 Jan 2018. [Online]. Available: https://www.ackstorm.com/google-cloud-vision-api/