Inclusive Mobility Solution for Visually Impaired People using Google Cloud Vision

1st Joana Gonçalves
Instituto Politécnico de Viana do Castelo
Viana do Castelo, Portugal
joanagoncalves@ipvc.pt

2nd Sara Paiva
ADiT-LAB, Instituto Politécnico de Viana do Castelo
Rua Escola Industrial e Comercial Nun'Álvares
4900-347 Viana do Castelo, Portugal
sara.paiva@estg.ipvc.pt

***Abstract*—Mobility in cities is of particular and growing importance nowadays due to the demographic increase and the existence of people with reduced mobility, as is the case of visually impaired people. Of the various situations where mobility represents a challenge, obtaining the notion of positioning, at times when the person loses track of where he is and becomes disoriented, can be extremely useful and a way to contribute to greater autonomy for this segment of people. This paper proposes a visual positioning system using the Google Cloud Vision API. The architecture includes a mobile application that captures an image via the mobile phone and sends it to a backend server that makes use of Google Cloud Vision to recognize the image, which may consist of text, logos or landmarks. In a first phase, the solution was evaluated individually and, in a second phase, on a route chosen in the city of Braga, in Portugal. Logo recognition achieved an accuracy of 98% and proved to be sensitive to image resolution. The frontal text recognition obtained an accuracy of 100% while the lateral recognition and at a 3 meters distance obtained lower values, with worse results in images with more text and of reduced dimensions. Landmark recognition always returned the correct result, although the average accuracy is 82%. The processing time was around 3 seconds in tests done with Wi-Fi network and about 2 seconds in field tests made with mobile network. The obtained results prove the adequacy of using this solution to be adapted in a real scenario.**

***Index Terms*—Visually Impaired People, Mobility, Inclusive Mobility, Outdoor Positioning, Image Recognition.**

I. Introduction

The World Health Organization (WHO), in its report on "Visual Impairment 2010", refers approximately 320 thousand people per million with some visual impairment, worldwide, and about 47 thousand people per million considered blind [1]. In Portugal, the 2001 reports had values in the order of 160 thousand visually impaired individuals [2] while the 2011 reports point to 900 thousand visually impaired individuals, 28 thousand of whom are blind [3]. On another hand, the United Nations agenda for 2030 [4] defines 17 Sustainable Development Goals (SGDs) where the aspect of inclusive mobility is highlighted so that people with reduced mobility, where visually impaired people (VIP) are included, can have access to solutions that help them to get to relevant places they need. One of the main challenges for VIP in navigation in a city includes the difficulty of moving from a source to a given destination in an autonomous way, being able to position themselves when they lose their orientation, or even detecting obstacles in their path, such as cars and holes that hinder their circulation [5]. The navigation systems, widely used for several years, are intended to help a person to locate themselves in unknown places and is mainly based on the use of GPS (Global Positioning System). The GPS, although useful in some scenarios, has some limitations: 1) low signal accuracy (may present errors in the order of 1 to 10 meters of accuracy [6]), 2) fluctuations due to atmospheric conditions and 3) delay in the response by the satellites, resulting in the slowness of the response in real time [7]. One of the alternative techniques for outdoor positioning and navigation is the visual positioning, which includes positioning that makes use of landmarks, which are known places in a city [8], as being squares or monuments. The visual positioning technique for navigation purposes is effective when the landmarks are clearly visible and the directions are properly incorporated in the photograph, and also when it is integrated with a map-based system [9].

This paper presents a solution based on visual positioning to assist in the positioning of VIP in a city, using the Google Cloud Vision API. The solution includes a mobile application, which is used to capture the image of the surrounding environment and inform the VIP of nearby places through which he is passing and which are useful him to position himself and regain his orientation. Recognition tests of images containing text (such as store names), logos or known landmarks were carried out. The solution was evaluated in terms of its accuracy and processing time of images recognition. The rest of the paper is structured as follows. The next section presents the state of the art, highlighting applications that address the issue of navigation for visually impaired people. Section III presents the system architecture as well as the developed prototype. Section IV presents the evaluation of the use of the API in terms of recognition accuracy and processing times. Finally, the discussion and conclusion are presented in section V.

II. Related Works

In [10], an outdoor system is proposed in which a geographic database of pedestrian paths in a city is used and also a GPS to provide the user's position, in which the objective is to determine the optimum route from the origin to destination. In [11], authors present a mobile navigation system to assist in the navigation of people with visual impairments when navigating in outdoor environments. The suggested solution is

2021 IEEE International Smart Cities Conference (ISC2) | 978-1-6654-4919-9/21/$31.00 ©2021 European Union | DOI: 10.1109/ISC253183.2021.9562892

based on the use of a mobile device camera and *Deep Learning* algorithms to recognize and detect different objects, as well as providing additional information to help VIP. In [12], authors present an application entitled ”BlindNavi” that intends to make life easier for people with visual impairments. The main objective is to provide a new mobility aid solution, in the form of a navigation application that stores significant information during the trip to make it safer. The application uses voice feedback that consists of multisensory cues combined with micro-location technology to help VIP to explore the outdoor environment. In [13], a case study is presented to promote and enhance transport sustainability by providing a mobile app that allows visually impaired people to independently use public transportation. The work, a follow-up of previous work on this topic [14] [15], presents the main functionalities available in a mobile app, including a careful and detailed layout definition to improve usability. Tests made in the city with people from institutions are presented allowing to draw some important conclusions regarding the usefulness of such an application as well as future directions.

In [16], authors introduce a navigation system for people with visual impairments for both outdoor and indoor environments. The components of this system include a laser that provides the distance and angle obtained from different positions, a laptop computer, headphones and an inertial measurement unit that contains accelerometers, gyroscopes and magnetometers that can be located anywhere on the body of the person. The algorithms used are Simultaneous Localization and Mapping (SLAM), Pedestrian Dead Reckoning (PDR) and GPS. In [17], authors describe a navigation system through a Smart Robot integrated with RFID and GPS for the VIP, both in indoor and outdoor environments. The system guides the user to a predefined destination or to create a real-time route for later use. In navigation mode, the Smart Robot arrives at the destination avoiding obstacles using ultrasonic and infrared sensors. In [18], authors propose a system for VIP both in indoor and outdoor environments that focuses mainly on providing a voice-based outlet for preventing obstacles and also for navigation using an ultrasonic sensor. The solution also uses GPS and voice that alert the visually impaired.

Many of the contributions by the scientific community to help VIP in their navigation are focused on indoor environments. Next some of those proposals are presented, with the main objective of understanding the methodologies, technologies and approaches that, in some cases, can be replicated for outdoor environments as well. In [19], authors present an indoor navigation system for VIP, using an approach based on computer vision. Fiducial markers with audio information are placed in the interior environment. The navigation system provides two types of guidance: free mode guidance and guided mode navigation. In the first mode, the user navigates freely and the system only informs him of his current position through audio based on the fiducial marker. In guided navigation mode, the system uses the Dijkstra algorithm, in which the user is guided from the beginning to the destination by the shortest path. In [20], authors describe an indoor navigation system using a modified cane that includes color sensors and an RFID reader, along with a micro controller, speaker and vibration equipment. The systems installed on the cane are a navigation system and a cartographic information system. The navigation system follows a colored navigation line that is placed on the floor. A color sensor is installed on the tip of the cane and detects the color of the line that the user is walking and the user is informed that he is walking the line through the sensation of vibration. In [21], authors present a navigation system that detects obstacles and helps in the navigation of visually impaired people on the best way forward. The obstacle is detected using an infrared-based detection system and it sends the feedback to the receiver through vibration-tactile or audible feedback to inform the user of its position. A sensor is attached to the user’s head and allows the user to obtain information about obstacles. In [22], authors present a system that integrates data from a building’s geographic information system with detection of landmarks to locate the user in the building and to plot and validate a route for the user’s navigation. The developed system complements the cane to improve the user’s autonomy when navigating inside the building. In [23], authors propose an Android application to be used for indoor navigation with audio instructions where QR codes are used. The application provides assistance in navigating predefined paths for the blind. These QR codes are placed in different floor sections that, after a specific distance, act as an input for the detection and navigation of the current location. Whenever this QR code is scanned, it provides the user with the current location information and asks the user to select the destination, after which the shortest and most optimized path is made available. In [24], solutions and technologies that are being used to assist in the mobility of visually impaired people, both in indoor and outdoor environments, as well as in the detection of obstacles are reviewed. Recent research in these three main categories is presented as well as a discussion and summary of main technologies, approaches, and equipment used.

III. System Overview

This section presents the architecture of the system and its components, as well as the prototype developed for the mobile application to be tested in a real environment.

A. Architecture

Figure 1 shows the system architecture where two main components can be identified: the mobile application and the image recognition API, which in this case is Google Cloud Vision. The system works as follows: the user, using the mobile application, should point the camera at a given location that surrounds him and, automatically, a photo is taken and sent through the application to the image recognition API of the Google Cloud Vision. After processing, the result is returned and the VIP is informed of the result that is transmitted via audio through Talk Back. The result will allow the person to identify the location to which he is facing, which represents very relevant information for a VIP to be able to locate and return guidance if he has lost it. The image recognition algorithm is currently prepared to process well-known logos and landmarks or text (in which an example is the name of stores).

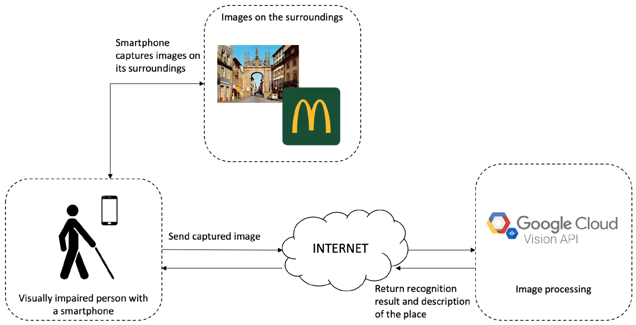

Fig. 1. Architecture of the system with two main components: the mobile application and the Google Cloud Vision API for image recognition

B. Mobile App Prototype

Figure 2 shows the prototype of the mobile application that was developed taking into account the maximum ease of use. Thus, when the application is opened, the camera view is automatically presented, and the capture of images is done every 7 seconds, which allows the image to be captured and the output of the recognition process is received. The user can, in this way, walk down a street with his mobile phone, and receive information about the places he is passing through, provided that they consist of text, logo or landmarks. The images shown in Figure 2 represent the application being executed and the result of the recognition process being presented to the user. With TalkBack or VoiceOver enabled, this information is read to the user automatically. From left to right, it is possible to see the application recognizing a landmark (Garden Central Avenue), a logo (McDonald´s) and text (Santander bank).

Fig. 2. Prototype images. The application captures an image and sends the result to the user. From left to right: landmark, logo and text.

IV. Evaluation

This section starts by presenting the methodology used to carry out the tests and then presents the evaluation of the solution from two different perspectives. First, the results of the individual assessment of text recognition, logo and landmarks are presented and, in a second phase, field tests results are presented.

A. Methodology

The individual tests performed on the recognition of text, logo and landmarks aimed to measure the time of recognition and the obtained accuracy. The field test additionally measures the smartphone's battery and CPU consumption. All tests were performed with Redmi Note 8 which is running on the Android version *10 QKQ1.200114.002* and the MIUI version *-Miui Global 12.0.3*. For monitoring individual tests, Android Debug Bridge (ADB) is used, which is a versatile command line tool that allows communication with an emulator instance or with an Android device connected via Wi-Fi. The individual tests, and as a result of the confinement situation due to COVID-19, were carried out using printed photographs with a resolution of 3840x2160. The tests followed the following methodology:

- the app points to the image and awaits the recognition result;
- the results (processing time and accuracy) are stored;
- the app points to a neutral location, where a result of ”no elements were recognized” is returned
- the process is repeated 20 times.

B. Evaluation of Logo Recognition

To assess the recognition of logos, three were chosen: McDonald's, Starbucks and Burger King, as shown in Figure 3. In order to assess the impact of the image resolution, one of the logos was printed with a resolution of 1036x1024 (all other has a resolution of 3840x2160).

Fig. 3. Logos used for the recognition tests

The results obtained in terms of processing time and accuracy are illustrated in Figure 4. Its analysis allows to verify a significant difference in the recognition accuracy of the McDonald's logo of low resolution (presented as Mc (low-res)) in relation to all the other logos. Another conclusion is that there are three measurements in all four logos in which there is an processing time of approximately 4,500 milliseconds. This value corresponds to the first request made to the API, and after continuous requests this value tends to decrease. Table I indicates the average values of processing times and recognition accuracy for each of the logos. The obtained values are very similar for the higher resolution logos; the only value to be highlighted is the lesser accuracy of the recognition of the logo with a lower resolution.

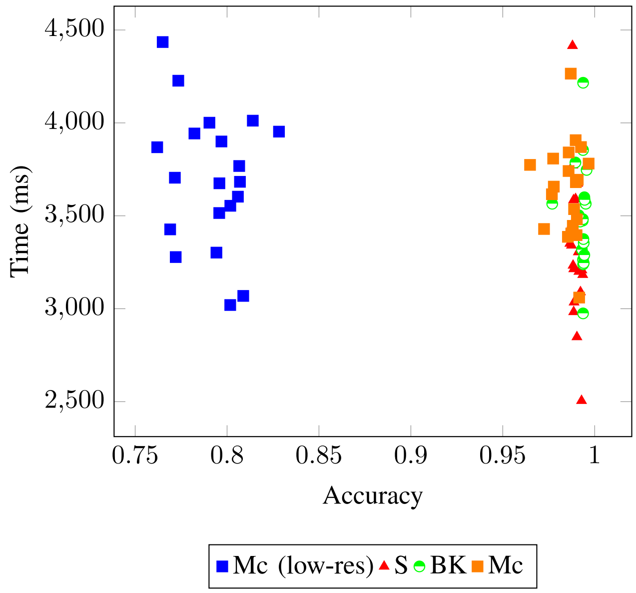

Fig. 4. Evaluation results of logo recognition

TABLE I
AVERAGE PROCESSING TIME AND AVERAGE ACCURACY OF THE FOUR LOGOS

	Burger King	McDonald's	McDonald's low-res	Starbucks
Avg accuracy	0,99	0,98	0,79	0,98
Avg proc. time (ms)	3457,47	3605,78	3658,10	3605,78

C. Evaluation of Text Recognition

For the evaluation of text recognition, three images of places with text in it were chosen and are presented in Figure 5. The first one from the left represents a pizzaria and the text in it is "MAMMA MIA Ristorante Pizzeria; the second represents a street name and in smaller letter the postal code and the entire text is "4710-079 Rua José Antunes Guimarães Gualtar"; the third image is the name of a clothes store and the text is "ARRANJOS DE ROUPA D.AMÉLIA".

Fig. 5. Images used for the evaluation of text recognition. From the left to the right: (a) "MAMMA MIA Ristorante Pizzeria"; (b) "4710-079 Rua José Antunes Guimarães Gualtar" and (c) "ARRANJOS DE ROUPA D.AMÉLIA"

The test carried out follows the same procedure explained in the Methodology section. In this specific test, for each image, frontal, side and distance tests were carried out to understand the impact of each of these factors on the accuracy of the recognition algorithm. To measure the accuracy, we used the accuracy formula presented in Figure 6 [25].

1) Frontal angle: The frontal angle tests all returned an accuracy of 100%. In Table II it is possible to verify a slightly longer processing time in Figure 5 (b), the reason for which is

$$accuracy = \frac{\#(character/word \times correctly\ recognized)}{\#(character/word \times recognized)}$$

Fig. 6. Accuracy formula used for text recognition evaluation

understood to be because there are more letters in this specific sample and some of them with small dimensions.

TABLE II
AVERAGE PROCESSING TIME AND AVERAGE ACCURACY OF FRONTAL ANGLE TEXT RECOGNITION

	Fig 5 (a)	Fig 5 (b)	Fig 5 (c)
Avg. accuracy	1	1	1
Avg. proc. time (ms)	3927,4	4572,7	3803,25

2) Side angle: In Figure 7, the processing time of recognition of the text of each image, as well as the accuracy, according to the formula indicated in Figure 6 are presented. The side text recognition performed worse in Figure 5 (b), due to the small size of the letters related to the postal code. In the case of Figure 5 (a), in the 20 tests made, only in 3 of them, the letter "N" of the word "RISTORANTE" was wrongly recognized by the letter "M". In the case of Figure 5 (b), in the 20 tests made, in 6 times it did not recognize correctly the postal code (4710-079). In the case of Figure 5 (c), in 2 of the tests, the letter "D." was not recognized correctly.

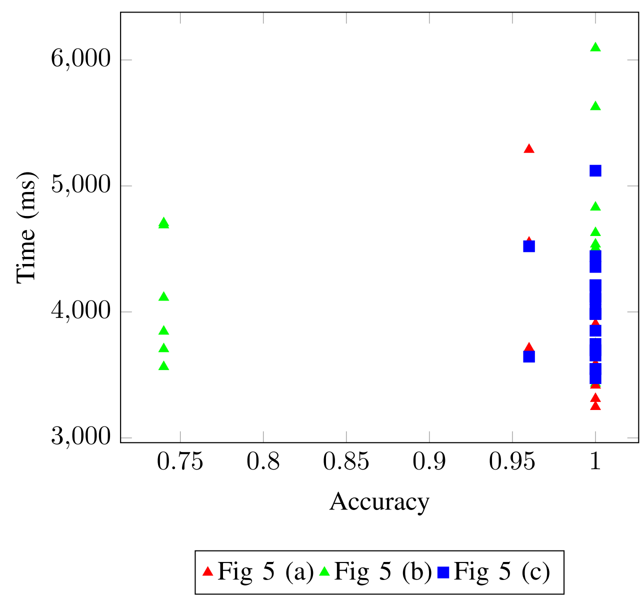

Fig. 7. Recognition results in side angle

In Table III, and following the results presented in Figure 7, it is possible to verify that the average accuracy recognition value of Figure 5 (b) was the lowest, and the processing time was also the highest.

3) Distance tests: For these tests, the recognition of the images was done from 3 meters away. In Figure 8, there is some heterogeneity in the results obtained. In the case

TABLE III
AVERAGE PROCESSING TIME AND AVERAGE ACCURACY OF SIDE ANGLE TEXT RECOGNITION

	Fig 5 (a)	Fig 5 (b)	Fig 5 (c)
Avg. accuracy	0,99	0,93	0,99
Avg. proc. time (ms)	3849,8	4235,25	4011,05

of Figure 5 (a), in 4 of the tests the letter ”R” of the word ”Ristorante” was not correctly recognized (the algorithm recognized it as ”Q”). In the case of Figure 5 (b), the postal code alone was not recognised in 13 of the 20 tests (due to the reduced size of the letters) and in 4 times the words were also not correctly identified. In the case of Figure 5 (c), the words ”D.Amélia” were never correctly recognized due to the existence of a strong luminosity over this part of the image.

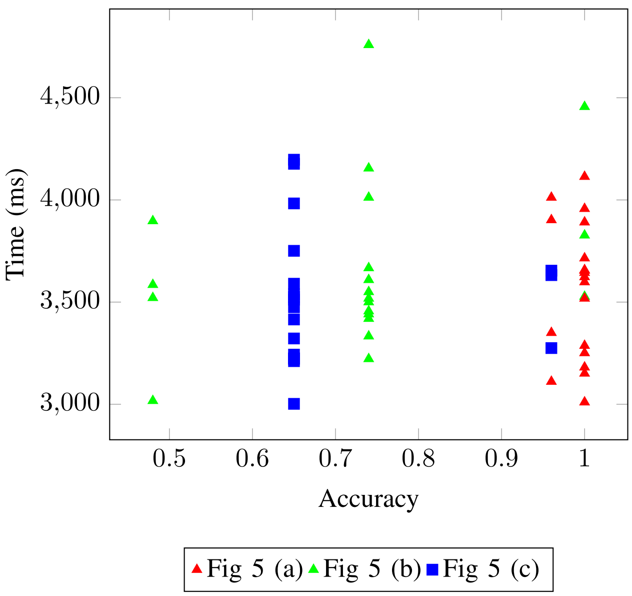

Fig. 8. Recognition results from a 3 meters distance

In Table IV it is possible to see the average processing time and average accuracy for the recognition of the text in the three images, that are coherent with the results presented in Figure 8.

TABLE IV
AVERAGE PROCESSING TIME AND AVERAGE ACCURACY BY DISTANCE

	Fig 5 (a)	Fig 5 (b)	Fig 5 (c)
Avg. accuracy	0,99	0,72	0,69
Avg. proc. time (ms)	3564,15	3673,55	3508,65

D. Landmark recognition

For the evaluation of the recognition of landmarks, three points of reference illustrated in Figure 9 were chosen: (a) Bom Jesus, in the city of Braga, (b) Santa Luzia, in the city of Viana do Castelo and (c) Avenida Central, in the city of Braga. The Google Cloud Vision API returns a confidence score associated with the recognition of landmarks, which was used in Figure 10, where the score for the recognition of the three images is presented. There was never a score above 0.9 and there are values that are close to 0.65, but in all of them the returned text was the correct one. The score is therefore not directly related to the correct identification of the landmark. To eliminate variables such as resolution, brightness and other factors that could be associated with these lower values in the recognition score, a test was made with an image of ”Bom Jesus” directly on the Google API web console and the score was 0.8 which indicates that these are the average values returned by the API for the identification of landmarks. It should be noted that in all tests performed for the 3 landmarks, and although the scores are those shown in Figure 10, the location was always correctly identified in terms of the returned text.

Fig. 9. Images used for landmark recognition: (a) Bom Jesus, in the city of Braga, (b) Santa Luzia, in the city of Viana do Castelo and (c) Avenida Central, in the city of Braga.

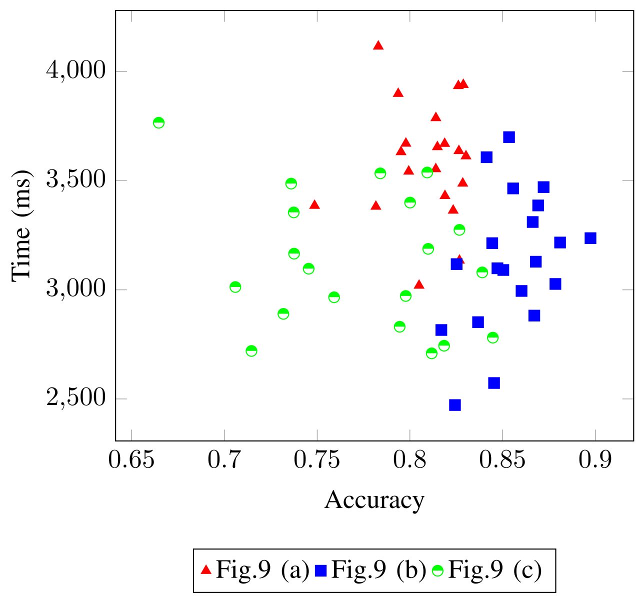

Fig. 10. Landmark Recognition Results

Table V presents the processing time and the accuracy of the landmark recognition. As previously mentioned, the average values for the accuracy are normal for a landmark recognition using the Google Cloud Vision API and the place was always correctly identified.

E. Fields Tests

In addition to the individual tests for logos, texts and landmarks presented in the previous sections, a route was

TABLE V
AVERAGE PROCESSING TIME AND AVERAGE ACCURACY OF LANDMARK RECOGNITION

	Fig.9 (a)	Fig.9 (b)	Fig.9 (c)
Avg. accuracy	0,81	0,86	0,78
Avg. proc. time (ms)	3592,95	3133,2	3125,6

designed in a central location in the city of Braga, in Portugal, so a real test scenario could be carried out in the field, representing a route a person could do in a real situation, passing by places of the three types considered. Figure 11 represents the location where the tests were performed, with two types of points being identified:

1) places to be recognized (represented with red color). Four places were chosen: (1) - McDonald's (Logo), (2) - Santander Bank (Text), (3) - Garden Central Avenue (Landmark) and (4) - Jupial (Text)
2) places from where the recognition was made (represented in black color)

Fig. 11. Map scenario of the field tests showing the places to be recognized (1 to 4) and the places where the recognition was made (5 to 11)

The following route was performed five times: (5) - (9) - (6) - (7) - (10) - (11) - (8). Table VI shows the places where the recognition of each of the four places was made. Text places (1 and 4) were only recognized from one location and at a short distance. The logo (1) and the landmark (3) were recognised in more places since they are larger and more noticeable at a distance.

TABLE VI
MATCHING BETWEEN THE PLACES TO RECOGNIZE AND THE PLACES WHERE THE RECOGNITION WILL BE MADE

Places where the recognition is made	Place to recognize
(5) (9)	(1)
(6)	(2)
(7) (10) (11)	(3)
(8)	(4)

Table VII represents the accuracy and processing time of the text recognition of the locations (1) and (4). In the 5 courses performed as a test, the texts of the two locations were correctly identified in all situations, with an accuracy of 100%, according to the formula defined in Figure 6.

TABLE VII
AVERAGE PROCESSING TIME AND AVERAGE ACCURACY OF TEXT RECOGNITION IN FIELD TESTS

	Santander	Jupial
Avg. accuracy	1	1
Avg. proc. time (ms)	1694,4	1852,2

Table VIII represents the accuracy and processing time of the logo recognition (1) from two different locations, one closer (5) and one with some distance (9). In the 5 courses performed as a test, the logo was recognized with an accuracy of 97% from the most distant location and with an accuracy of 98% from the closest location.

TABLE VIII
AVERAGE PROCESSING TIME AND AVERAGE ACCURACY OF LOGO RECOGNITION IN FIELD TESTS

	From location (5)	From location (9)
Avg. accuracy	0,98	0,97
Avg. proc. time (ms)	1687,4	1605

Table IX represents the accuracy and processing time of the landmark recognition (3) from three different locations. From the closest location (7), the landmark was recognized with an accuracy of 49% which is due to the fact that the landmark is not fully captured by the camera due to its proximity to the place. From location (11) there was a fountain in front of the landmark, which also reduced the accuracy of recognition. The recognition from location (10) was superior to the others, given that the captured image does not have any significant obstacles and the landmark can be captured in its entirety. Despite the accuracy of the recognition, in two situations, present values around 50%, the name of the landmark returned to the user was always correct.

TABLE IX
AVERAGE PROCESSING TIME AND AVERAGE ACCURACY OF LANDMARK RECOGNITION IN FIELD TESTS

	From location (7)	From location (10)	From location (11)
Avg. accuracy	0,49	0,64	0,48
Avg. proc. time (ms)	2196,8	1976,6	2278,6

V. DISCUSSION AND CONCLUSIONS

In this paper, a mobile app was proposed, implemented and tested with the end purpose to help visually impaired people to move around cities and obtain information about their current location and thus solving situations where they lose orientation. The solution uses the Visual Positioning approach and the Google Cloud Vision API is used to recognize images, namely text, logos and landmarks. The solution was individually assessed on the three aspects using photographs, and a field test was also carried out in the city of Braga, in Portugal.

With regard to logo recognition, an accuracy of 98% was achieved with recognition using photographs with a resolution of 3840x2160. The decrease in image resolution to 1036x1024 had a considerable impact on the accuracy, which in this case was 79%. The average processing time was 3.5 seconds, using a Wi-Fi network. With regard to text recognition, three different images were used, and three positions were considered: frontal, lateral and with 3 meters distance. In the frontal angle, the recognition accuracy was 100%. In the lateral angle, an accuracy of 99% was achieved in two of the images and an accuracy of 93% in another, which is justified by the fact that it is an image with smaller letters and which will not be as visible from a side angle. Finally, at a distance of 3 meters, the accuracy was noticeably lower, with many letters being incorrectly identified. The average accuracy value in this case was 80%. With regard to processing times, the average was 4.4 seconds, with the highest values being recorded in the image recognition with smaller letters and in greater numbers. Landmark recognition obtained an average accuracy of 82%, which is considered a good value, as this is the average value obtained when a sharp and good resolution photo is submitted for recognition directly on the API's web console. Even taking into account the accuracy value, the result returned on the landmark identification was always correct. The average time was 3.2 seconds.

Field tests were carried out in the city of Braga. For this purpose, a route was designed that included the recognition of two texts, one landmark and one logo. The accuracy of the text recognition was 100% with an average processing time of 1.8 seconds. The recognition of logos had an accuracy of 98%, and the greater distance to the logo was not significant. The processing average in this case was 1.6 seconds. Finally, recognition of the landmark always gave the correct location as a return, although the accuracy varied when it was too close and when there were obstacles ahead. The average processing time, in this case, was 2.2 seconds. The decrease in average processing times in tests carried out in a real scenario is highlighted, which may eventually be justified by the use of mobile data.

The use of this API proved to be adequate for the defined problem research, and it can be a viable solution for incorporation in a mobile application with the objective of helping visually impaired people to have a greater orientation when moving in cities.

REFERENCES

[1] WHO, "Global data on visual impairments 2010," 2010, last accessed on April 14th 2021. [Online]. Available: https://www.who.int/blindness/GLOBALDATAFINALforweb.pdf?ua=1

[2] ANACOM, "Acapo - associação dos cegos e amblíopes de portugal," 2002, last accessed on April 14th 2021. [Online]. Available: https://www.anacom.pt/render.jsp?categoryId=36666

[3] ACAPO, "Deficiência visual," 2021, last accessed on April 14th 2021. [Online]. Available: http://www.acapo.pt/deficiencia-visual/perguntas-e-respostas/deficiencia-visualquantas-pessoas-com-deficiencia-visual-existem-em-portugal-202

[4] U. Nations, "Transforming our world: the 2030 agenda for sustainable development," 2020, last accessed on March 31st 2021. [Online]. Available: https://sustainabledevelopment.un.org/post2015/transformingourworld

[5] A. Riazi, F. Riazi, R. Yoosfi, and F. Bahmeei, "Outdoor difficulties experienced by a group of visually impaired iranian people," *Journal of Current Ophthalmology*, vol. 28, pp. 85–90, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.joco.2016.04.002

[6] N. A. Giudice and G. E. Legge, "Blind navigation and the role of technology," *The Engineering Handbook of Smart Technology for Aging, Disability, and Independence*, pp. 479–500, 2008.

[7] C. K. Lakde and D. P. S. Prasad, "Review paper on navigation system for visually impaired people," *Ijarcce*, vol. 4, pp. 166–168, 2015.

[8] P. Heiniz, K. H. Krempels, C. Terwelp, and S. Wuller, "Landmark-based navigation in complex buildings," *2012 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2012 - Conference Proceedings*, pp. 13–15, 2012.

[9] H. Hile, R. Vedantham, G. Cuellar, A. Liu, N. Gelfand, R. Grzeszczuk, and G. Borriello, "Landmark-based pedestrian navigation from collections of geotagged photos," *MUM'08 - Proceedings of the 7th International Conference on Mobile and Ubiquitous Multimedia*, pp. 145–152, 2008.

[10] L. Kaminski, R. Kowalik, Z. Lubniewski, and A. Stepnowski, ""voice maps" - portable, dedicated gis for supporting the street navigation and self-dependent movement of the blind," *Proceedings of the 2010 2nd International Conference on Information Technology, ICIT 2010*, pp. 153–156, 2010.

[11] S. Shadi, S. Hadi, M. A. Nazari, and W. Hardt, "Outdoor navigation for visually impaired based on deep learning," *CEUR Workshop Proceedings*, vol. 2514, pp. 397–406, 2019.

[12] H.-E. Chen, Y.-Y. Lin, C.-H. Chen, and I.-F. Wang, "Blindnavi: A navigation app for the visually impaired smartphone user," p. 19–24, 2015. [Online]. Available: https://doi.org/10.1145/2702613.2726953

[13] P. Faria, A. F. Curralo, and S. Paiva, *A Case Study for the Promotion of Urban Mobility for Visually Impaired People*. Cham: Springer International Publishing, 2019, pp. 65–82.

[14] A. Lima, D. Mendes, and S. Paiva, "Mobile solutions for visually impaired people: Case study in viana do castelo historical center," in *2017 12th Iberian Conference on Information Systems and Technologies (CISTI)*, 2017, pp. 1–6.

[15] D. Brito, T. Viana, D. Lourenço, and S. Paiva, "A mobile solution to help visually impaired people in public transports and in pedestrian walks," in *Journal of Sustainable Development and Planning*, vol. 13, no. 2, 2018, pp. 281–293.

[16] E. B. Kaiser and M. Lawo, "Wearable navigation system for the visually impaired and blind people," *Proceedings - 2012 IEEE/ACIS 11th International Conference on Computer and Information Science, ICIS 2012*, pp. 230–233, 2012.

[17] K. Yelamarthi, D. Haas, D. Nielsen, and S. Mothersell, "Rfid and gps integrated navigation system for the visually impaired," *Midwest Symposium on Circuits and Systems*, pp. 1149–1152, 2010.

[18] R. A. Tanpure, "Advanced voice based blind stick with voice announcement of obstacle distance," vol. 4, pp. 85–87, 2018.

[19] A. Zeb, S. Ullah, and I. Rabbi, "Indoor vision-based auditory assistance for blind people in semi controlled environments," *2014 4th International Conference on Image Processing Theory, Tools and Applications, IPTA 2014*, pp. 0–5, 2015.

[20] A. J. Fukasawa and K. Magatani, "A navigation system for the visually impaired an intelligent white cane," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 4760–4763, 2012.

[21] P. Bhardwaj and J. Singh, "Design and development of secure navigation system for visually impaired people," *International Journal of Computer Science and Information Technology*, vol. 5, pp. 159–164, 2013.

[22] M. Serrão, J. M. Rodrigues, J. I. Rodrigues, and J. M. D. Buf, "Indoor localization and navigation for blind persons using visual landmarks and a gis," *Procedia Computer Science*, vol. 14, pp. 65–73, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.procs.2012.10.008

[23] A. Idrees, Z. Iqbal, and M. Ishfaq, "An efficient indoor navigation technique to find optimal route for blinds using qr codes."

[24] S. Paiva and N. Gupta, *Technologies and Systems to Improve Mobility of Visually Impaired People: A State of the Art*. Cham: Springer International Publishing, 2020, pp. 105–123.

[25] M. Junker, R. Hoch, and A. Dengel, "On the evaluation of document analysis components by recall, precision, and accuracy," 04 2000.