# Edge Solution with Machine Learning and Open Data to Interpret Signs for People with Visual Disability

Fabian Velosa, Hector Florez

Universidad Distrital Francisco Jose de Caldas, Bogota, Colombia

#### Abstract

Transport systems in growing cities present communication barriers for people with visual impairment that can prevent their caregivers' locomotion, making a city unequal, and limiting the full exercise of rights. Some current solutions and prototypes do not fully solve this problem. This paper proposes an application that uses machine learning and open data to identify, extend, and communicate the available information on the signs placed by the urban transport system of Bogota, Colombia using a mobile device without the need to use any internet connection. The solution presents a replicable model suitable for other transport systems. The result is a free mobile app available in the Android store covering the buses' routes in Bogota. Further works consider the extension of the strategy to the public transports platform of any city.

#### **Keywords**

Machine Learning, Object Detection, Optical Character Recognition, Accessibility, Transport System

## 1. Introduction

Transport systems determine different aspects of people's lives in cities that experience strong urban growth phenomena. However, these systems present different types of barriers for people with disabilities. To make a city accessible, it does not only require specialized infrastructure but also it is necessary to make it possible to reach the city using the transport system, which is itself a tool to guarantee the rights. The set of situations that prevent the possibility of fully exercising the rights that each citizen has known as barriers [1]. Those defined as communication barriers are the obstacles that hinder access to information and the flow of it. These barriers particularly affect partial or complete sensory disabilities, making it difficult to communicate and access information. There is then a communication barrier in the information systems related to urban transport and accessibility. Thus, it is evident the possibility of using appropriate technologies to increase the information available in the graphic signs to generate perceptual auditory accessibility.

In this project, we developed a model that has been tested through an application for mobile devices with the Android operating system, which implements Machine Learning functionalities and open data to consult the information of the routes in the zonal stops of the Integrated System of Urban Transport (SITP) of the city of Bogota Colombia without requiring an Internet

ICAIW 2020: Workshops at the Third International Conference on Applied Informatics 2020, October 29–31, 2020, Ota, Nigeria

velosaf@acm.org (F. Velosa); haflorezf@udistrital.edu.co (H. Florez)

**D** 0000-0002-5339-4459 (H. Florez)

<sup>© 2020</sup> Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

connection. This mobile application designed focused on people with the condition of partial or complete visual disability.

The paper is structured as follows. Section 2 contextualizes the city of Bogotá, the SITP transport system, and visual disability. Section 3 presents the most relevant technological solutions from different approaches, the available applications related to routes, academic prototypes, accessibility applications in general, and the current infrastructure. In section 4, we illustrate the solution developed based on the life cycle of a Machine Learning project. Section 5 reports the results obtained and finally, section 6 presents the most relevant conclusions of the project.

## 2. Disability and SITP

Disability is a demonstration of human diversity, which occurs either congenitally or at any time in life. It could be materialized when a person tries to interact in society or with the environment finding barriers in the context that is not prepared for this interaction generating exclusion [1].

According to the contract "Design and Evaluation of Accessibility Alternatives for SITP in Condition of Disability" between the town hall of Bogota and a private consultant, it was found that Bogota had approximately 195,821 people with disabilities between years 2011 and 2012. 20% of the characterized population have some form of visual disability, which represents a group of 21% of daily trips made by people with disabilities in the year 2011. Also, the contract identifies the absence of aids for sensory disabilities at stops and stations as a barrier that makes it difficult to carry out activities and travel. From all the people with disabilities who carry out some economic activity, 47% had some kind of visual disability, which reflects the importance of accessibility as a guarantee of rights.

Colombia's statutory law 1618 of 2013 defines accessibility as the conditions that must be met by services and facilities with the aim of providing equal access for people with disabilities to physical environments, transport, information, and communications as well as determines that technical aids shall be made with appropriate technology taking into account people's needs [2].

## 3. Related work

The related work is classified in the following aspects: accessibility, transport routes, prototypes, and infrastructure.

#### 3.1. Accessibility

Currently, it is feasible to find applications that make use of machine learning techniques with the aim of helping the visually impaired. In this field, different proposals have been developed such as:

Seeing-ai [3] is an application developed by the Microsoft team and available for iOS since mid-2017. It is a free application that uses the camera to facilitate the recognition and description of the environment of blind or low vision people identifying faces, objects, colors, and

even the value of paper money. One of the most attractive qualities of this application is that it is available in more than 5 languages, facilitating its use in different countries.

Google Lookout[4] is an initiative that uses artificial intelligence to recognize images and objects with the purpose of reducing accidents when visually impaired people travel. This application has a system of reading aloud that allows users to distinguish the objects in front of the device camera. It also has different modalities that condition its operation.

Tap Tap See[5] is an application available for most mobile devices. It is designed with internet APIs technology together with VoiceOver functions for the identification and recognition of objects and images for their subsequent description in voice reading.

Navilens [6] is an application for guiding people with disabilities through the city and with signaling functionalities for transport systems. It is a software capable of reading optimized QR codes arranged in a printed way, developed by Neosistec and the University of Alicante in 2017. This application is able to detect each code at a distance of 15 meters used in some transport systems and museums in Spain.

#### 3.2. Routes

Regarding technological applications that support the objective of providing information for route planning, it is possible to identify Moovit, Transmilenio and SITP, and Transmi App, which are available for free in the Google play store.

Moovit, is a world-class mobility application with over 720 million users and available in more than 3000 cities. It offers route information services through a combination of public information, authority information, and community information. In Bogota, it has been in operation since 2013 and its access is free for users. This app allows planning and suggesting a variety of modes of transport according to the availability and state of the traffic. Its main functions require online artificial intelligence services and its interface has TalkBack support for mobile devices as well as a system map in PDF format.

TransMilenio and SITP [7], is a proposal whose main function is to find and calculate the most optimal route for the user by reducing the number of stops within this transport system, even providing the possibility of requesting taxi services easily and safely. Users will be able to access information such as routes, stations, buses, schedules, recharging points as well as visualizing maps of the system pointing food service places, tourist sites of the city, and bike paths. It uses GPS and its access does not depend on an internet connection excepting for access to maps.

TransMi App [8] is the official application of the TransMilenio transport system. Its main function is planing the route given an origin and a destination. It includes the routes of the SITP and has an additional function that is monitoring in real-time of the routes and the time it takes to reach the stop. It makes use of an intelligent search to find the best paths and Google maps technology to suggest routes.

#### 3.3. Prototypes

The applications developed so far do not have a function to obtain information from the SITP stops by means of the camera of the mobile device. Likewise, they do not have a usability

interface designed for people with visual disabilities or low vision. However, some prototypes use spatial data to obtain information related to the routes or the state of the SITP. The next list of solutions proposes new approaches that contemplate the use of space and position of the user in relation to the bus stop.

NeuroSITP [9] is a prototype proposed in 2016. This application implements an artificial neural network to calculate the time that a bus takes for routes to reach different stops in one specific area providing real-time data to users. Its operation is relatively simple since it only requires indicating the destination neighborhood so that the system can provide the most appropriate options for a trip in the shortest possible time by means of client-server architecture through the Internet.

QRSITP [10] is a prototype proposed in 2016. It was planned as an information system related to the SITP in one specific area of Bogota. By means of software and hardware tools, it is possible to spatially consult the data related to destinations through the use of QR codes while its operation proposes a client-server architecture using an Internet connection.

ARComp [11] is a prototype created to serve as a technological tool for mobile devices showing the routes and access points belonging to the Integrated Transport System of Bogota. For this, it uses the space as a source of information using augmented reality techniques, also employing a client-server architecture through the Internet.

#### 3.4. Infrastructure

In Bogota, there are 4,881 bus stops with Braille plates as described in Figure 1 that offer information such as name, address, and telephone number for user service. Despite the fact that these initiatives contemplate a tangible accessibility option for the visually impaired sector of the population, In 2018, around 700 plates were replaced due to physical damage. The plates are limited by different situations such as lack of information, outdated information, and physical damage, which has generated a partial and out-of-context communication.

The information screens are part of a pilot test implemented at five points in the city. It provides real-time information on the status of services, routes, and estimated times for the arrival of buses thanks to a GPS system and the Internet that makes it possible to locate between routes and screens.

Therefore, it is possible to observe that no options are available as open-source software and they depend on specific technologies or hardware. The first three initiatives have been designed as general tools for the relationship of people in space, but they do not consider a specific approach to provide additional information. In most cases, these applications require Internet access. NaviLens [6] increases the perceived information but still has additional physical requirements as it needs previously programmed codes in the places of use.

Figure 2 summarizes the different software initiatives in this context grouping them in the following categories: Apps routes, prototypes, apps accessibility, and infrastructure.

## 4. Proposed solution

In this project, a machine learning model is proposed from all stages of its life cycle, which is described in Figure 3. This model is scalable and works in the deployment phase without the



Figure 1: Braille plates, by Bogota mobility department

need for an internet connection, which can be replicated with the signaling graphics of other transport systems and supported by the related open data sets. Based on this, it was necessary to create an application capable of using the camera of the Android device. The captured image is converted to a tensor and later evaluated by the model. Then, based on the results, the app is able to infer the position of the unique code of the signal. At this point, the implemented algorithm is used to recognize the optical characters of the image and generates as output an alphanumeric code that is compared to the available paths in the source of information obtained from the open data of Bogota. Later on, the information is presented to the user in an audible manner using the TalkBack accessibility assistant on the Android devices and optionally presented on the device screen.

The operating system Android was chosen due to its popularity around the world, low devices' prices, and the fact that most people who use the public system use Android devices city[1]. The software tools chosen to develop the project are open source, which advantage the support to run over Android the developed models.

#### 4.1. Open Data

Open data policies were born as an initiative of governments to provide citizens with the information available in public entities with the intention of using it and contributing to the development of the countries. They allow access to information as a tool for social control of

Apps Routes	Prototypes	Apps Accessibility	Infrastructure
Moovit ●	NeuroSITP	Seeing-ai* ● ●	Braille modules*
Transmilenio y Sitp ● ●	QRSITP*	Google Lookout* ● ●	Information screens
TransMi App ●	ARComp	Tap Tap See* ● ●	
		Neosistec*	

\* Use spatial user information

Depends on the Internet

- Depends on Infrastructure
- Limited information SITP
- No accessibility features
- Depends on the hardware

Figure 2: Related apps

the public administration. The most important value is the ability it gives citizens to understand their environment and find worth that improves their conditions.

The open datasets are used at two moments in this solution. The first moment is concerning the understanding of the data and the use of spatial geoinformation to locate the available signals and to be able to generate both the photographs and the images provided by Google Street View. The second moment is presented when the information at the end of the optical character recognition is obtained. For this, the application stores information about the routes and stops of the SITP. Information extracted from files come from various formats such as csv, json, geojson. For the normalization of the data, we used a script in python that is able to structure the information in the json standard and store them in two different files available for searching services in the application.

#### 4.2. Data sets

Machine Learning techniques depend on multiple data samples and variations for the correct training of the model and its learning of relationship by repetition [12, 13, 14]. The data necessary for this project was collected from different sources. A sample of each source is presented in Figure 4. The data sets were formed using two different sources and one was discarded.

• Pictures captured with Android devices from the mid-end range that usually have medium performance among the range of features offered by the market. The dimensions and resolutions of the photograph were limited to decrease the computational cost of processing

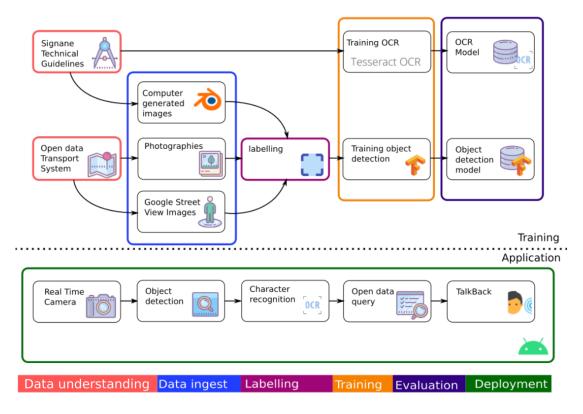


Figure 3: Proposed solution diagram

as tensors. This data source is accurate but ineffective due to the displacement required to obtain a large number of samples.

- Google Street View has a photo gallery of various cities around the world such as Bogota. This photo gallery is organized according to the streets and avenues. In some of these photos, it is possible to see the stops of the SITP. This information can be consulted using a web browser and using as parameters the latitude and longitude of the signal in the URL, however, the quality of the images obtained is low.
- Computer generated images. A workaround for the absence of data is to generate the images using a computerized method. There are multiple options for this purpose. In this solution, a design was modeled as a base in Blender open-source software for 3D creation using the technical specifications of the signage. Multiple samples were generated using a script in Python that varies randomly the position of the camera emulating the possible approaches of the users and varying the unique alphanumeric code available in the signal. Thus, 500 images were generated and are split into two data sets. The first one is used in the training process, while the second one is used for testing purposes.

The images obtained using the device's camera require more time and the process labeling of the code to be detected per image. The images obtained in Google Street View have a low



Figure 4: Dataset sources

quality that prevents the contours of the signs and their parts from being clearly identified; nonetheless, the use from data sets derived is not allowed and infringes the google permissions.

## 4.3. Object Detection Model

TensorFlow Object Detection API [15] is a framework supported by Google that works on top of TensorFlow<sup>1</sup> available in Python to reuse large CNN (Convolutional Neural Network) models by making use of Transfer Learning. Among the multiple options of previously trained models, we used the SSD MobilNet v1 model [16], which implements an architecture with depth-wise separable convolutions that improve some quality attributes over mobile architectures such as latency, accuracy, and performance. It was trained using 50.000 steps and a batch size of 12 with 300 x 300-pixel inputs over a data set of 350 training samples and 120 test samples achieving an approximate loss of 2.4. The outcome is a model that is a mathematical algorithm that uses the statistics to make detection or predictions.

## 4.4. Optical characters recognition trained data

Tesseract <sup>2</sup> is an OCR engine used to recognize the characters of the unique code of each signal. The outcome file to recognize the character optics was trained to limit the alphabet only to digits and the characters of the available modules from the technical stops guidelines, all codes match with the next pattern: three-digit followed by a letter and finally two digits.

## 4.5. Mobile application

The developed Android application called Supqua is available from Android API version 23, which corresponds to Android 6.0 Marshmallow as a minimum version with support for the TensorFlow Lite library. This version supports 84.9% of devices on the market according to

<sup>&</sup>lt;sup>1</sup>https://www.tensorflow.org/

<sup>&</sup>lt;sup>2</sup>https://tesseract-ocr.github.io/

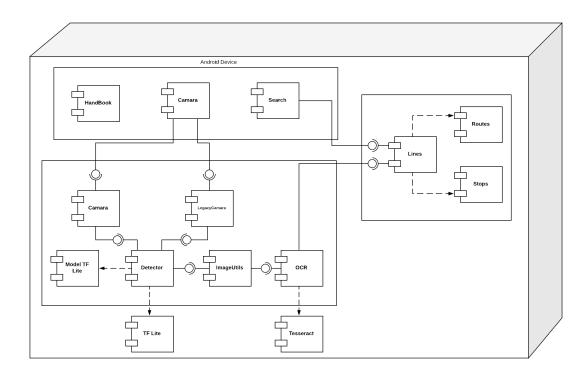


Figure 5: Application components diagram

Google statistics. The architecture of the app is described in the component application's diagram presented in Figure 5.

The application uses the device's camera passing its image to the object detector model, which locates the position of the signal border. This image is evaluated by optical character recognition using Tesseract and consults the resulting code in the open data files related to routes and stops. The workflow is illustrated in Figure 6

## 5. Results

As a result, we published a mobile application available in Google Play<sup>3</sup> available in Colombia. This application has a size of around 17 MB, which approximately 6 MB are for the object detection model, 142 KB is for alphanumeric character recognition, and 2.3 MB for open data files with routes information, with more than 100 downloads.

The results of the app on production are detailed in Table 1. These results were obtained using the debug mode of the application to log its time and performance. Thus, we can observe that the OCR recognition does not fulfill their job in two cases due to different environmental conditions. The second column shows the time used by the software to detect the signal. Finally, the confidence value is the probability that the object belongs to the class signal.

<sup>&</sup>lt;sup>3</sup>https://play.google.com/store/apps/details?id=com.catcode.supqua

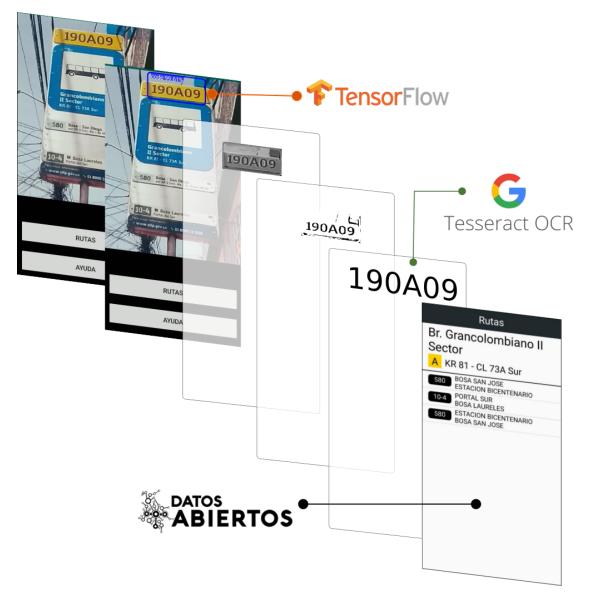


Figure 6: App result working

## 6. Conclusions

The proposed model can be extended to other transport systems in different cities by modifying the entries referring to technical guidelines and open data and creating new synthetic data enough to train a multi-class machine learning model.

Machine learning models are always exposed to continuous improvement by feedback with new hyperparameters, the better quality of training data, and different samples that is why this solution is ready for future updates.

#### Table 1 Results

Bus stop	<b>TF</b> Detection	OCR Recognition	Confidence
190A09	1 second	6 seconds	99.61%
659A09	1 second	did not recognize	73.05%
555A09	6 seconds	10 seconds	87.09%
055A04	26 seconds	did not recognize	80.05%
048A09	14 seconds	22 seconds	95.31%
069A06	9 seconds	15 seconds	73.05%

To get the expected results in this project, it was necessary to understand the problem and the available data. This volume of data gets better performance in the machine learning cycle that has as an outcome a model that could be used in mobile devices.

The use of machine learning in daily tasks could be contributing to the lack of planning of the infrastructure in cities in growth that do not include the different conditions of the citizens.

There is an opportunity to impact in a positive way the people of life with the new application of artificial intelligence.

## References

- A. F. C. Ordóñez, Acceso al transporte público para personas con discapacidad en bogotá: caso sitp, 2015. URL: http://bdigital.unal.edu.co/49986/, maestría en Derecho. Área de Profundización: Derecho Constitucional.
- [2] Congreso de la república, Ley estatutaria 1618 de 2013, 2013.
- [3] M. Corporation, Seeing-ai, 2019. URL: https://www.microsoft.com/en-us/ai/seeing-ai.
- [4] G. LLC, Lookout, 2020. URL: https://taptapseeapp.com/.
- [5] I. Cloudsight, Tap tap see, 2018. URL: https://taptapseeapp.com/.
- [6] Neosistec, Navilens, 2020. URL: https://www.navilens.com/.
- [7] MoviliXa SAS, Transmilenio y sitp, https://play.google.com/store/apps/details?id=com. rutasdeautobuses.transmileniositp&hl=es\_CO, 2020. URL: https://play.google.com/ store/apps/details?id=com.rutasdeautobuses.transmileniositp&hl=es\_CO.
- [8] TRANSMILENIO S.A., Transmi app | transmilenio, https://play.google.com/store/apps/ details?id=com.nexura.transmilenio&hl=es\_CO, 2020. URL: https://play.google.com/ store/apps/details?id=com.nexura.transmilenio&hl=es\_CO.
- [9] G. E. Palomino Contreras, P. L. Pineda Acero, et al., Diseño de un prototipo de software para el uso del sistema integrado de transporte público (sitp), en tiempo real en la localidad de chapinero, 2016. URL: http://hdl.handle.net/11349/3105.
- [10] B. Parra, D. Marcela, J. A. Parra Barrera, et al., Qrsitp (que ruede el sitp): Herramienta de software para consultar espacialmente las rutas del sitp, 2016. URL: http://hdl.handle.net/ 11349/3105.

- [11] J. Chawez, J. Orlando, J. L. Candamil Acevedo, et al., Diseño de aplicación de realidad aumentada en dispositivos móviles para usuarios del sistema integrado de transporte público de bogotá, 2015. URL: http://hdl.handle.net/11349/3050.
- [12] J. Hernandez, K. Daza, H. Florez, Alpha-beta vs scout algorithms for the othello game, in: CEUR Workshop Proceedings, 2019, pp. 65–79.
- [13] C. Vegega, P. Pytel, M. F. Pollo-Cattaneo, Evaluation of the bias in the management of patient's appointments in a pediatric office, ParadigmPlus 1 (2020) 1–21.
- [14] D. Sanchez, H. Florez, Improving game modeling for the quoridor game state using graph databases, in: International Conference on Information Technology and Systems, Springer, 2018, pp. 333–342.
- [15] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, S. Guadarrama, et al., Speed/accuracy trade-offs for modern convolutional object detectors, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 7310–7311.
- [16] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, Mobilenets: Efficient convolutional neural networks for mobile vision applications, arXiv preprint arXiv:1704.04861 (2017).