



Structure-Based Pattern Recognition to Aid Identification and Criminal Proceedings of Disclosure in Collective Transport in Belo Horizonte

Welton Pereira Da Silva Neves¹, Gabriel Dias De Lima Novais¹, Luiz Melk de Carvalho¹, Diva de Souza e Silva Rodrigues¹, Flávio Henrique Batista de Souza¹

¹*Centro Universitário de Belo Horizonte UNIBH, Brazil*

Abstract.

During the criminal action of theft and robbery of public transportation, sudden movements and the lack of real-time information about the event's occurrence can cost the patrimony or even the lives of innocent people. This paper demonstrates a proposed structure, based on Multilayer Perceptron (MLP) and Support Vector Machines (SVM), Internet of Things (IoT), Microcomputers and Cloud Computing; to assist the monitoring of public transport in the city of Belo Horizonte (3 million inhabitants) and help on the recognition of the behavior of a criminal event. The proposed structure captures conversations on the bus and, when identifying terms from an assault, considers the bus as in a risk situation, without the need to expose the victims during the occurrence. A comparative analysis was carried out based on the Area Under the Curve (AUC), of the pattern recognition structures, where it was used: a database with a 57-word dictionary; six types of MLP, with variations of: neurons in the hidden layer, periods and the test and validation set (constituting 108 tests and an average execution time of 42.36 seconds); SVM with K-Fold Cross Validation, with three types of Kernels (averaging 8.51 seconds). An AUC of 0.948 was obtained using SVM. This structure has an implantation architecture based on Microphones, Raspberry Pi 3 Model B and Google Voice API for capturing and processing speech during the collective journey. The system administration consolidation uses a Cloud Computing platform and it was estimated that each unit could cost between 400 to 600 Brazilian reais.

Keywords: Pattern Recognition, Collective Transports, Criminal Occurrences, Multilayer Perceptrons, Support Vector Machines

1. Introduction

Currently, methodologies for identifying crimes in Brazilian public transport are flawed and often dangerous, limiting the possibility of action by both passengers and drivers who circulate in them daily. In combat and action, the field of action of the police in this type of occurrence is very limited, given that, in most cases, they are aware of accurate information only after the bus drivers have contacted the responsible company.



Several works contemplate the research aspect in the Brazilian scenario, such as Magalhães (2008), which presents a potential evaluation of GPS tracking and monitoring technologies already implemented in Brazil and abroad. Gularte et. al (2016) presents a prototype of a GPS bus monitoring solution for Android, providing positioning information and vehicle arrival forecast in relation to the user. Finally, Fedrechski et Al. (N.d.), uses mobile computing and Software Engineering techniques in its project to develop an Android prototype for real-time monitoring, by GPS, of public transport vehicles.

Thus, the objective of this research is to propose a methodology for capturing, identifying and retracting buses that have some type of assault in progress, with the purpose of alerting in real time in a discreet (and especially not alarming) manner and with greater possible assertiveness, not only to the competent security agencies about the collective situation, but also to other users, so that they do not take this vehicle.

In the discussion of the paper, a process of delimiting the occurrence of the robbery is made, in order to evaluate possibilities of capturing behavioral information, which can be integrated with a technology based on pattern recognition, to assimilate and publicize the occurrence (without exposing users to extreme risk situation).

For this purpose, pattern recognition structures were tested, based on Multilayer Perceptrons (MLP) and Support Vector Machines (SVM), with resampling processes based on test and validation sets and k-fold cross validation, respectively, to select the best configuration that adapts to the occurrence identification process (with the aid of literature techniques for classifier comparison, in this case the ROC curve and the AUC - Area Under the Curve). With this information in mind, a methodology and structure for collecting and disseminating information on theft of collectives was proposed, with evaluations of equipment and services based on microprocessors, Cloud Computing and IoT (Internet of Things).

The focus of this work is on the transport models that are used in Belo Horizonte, a Brazilian city of more than 3,000,000 inhabitants. Therefore, such research also includes analyzes and literary reviews on the issue of image dissemination with the evaluation of Brazilian laws on the subject.

2. Methods

2.1 Conceptual Evaluation

2.1.1 Selection of Pattern Recognition Algorithms and Evaluation Structures

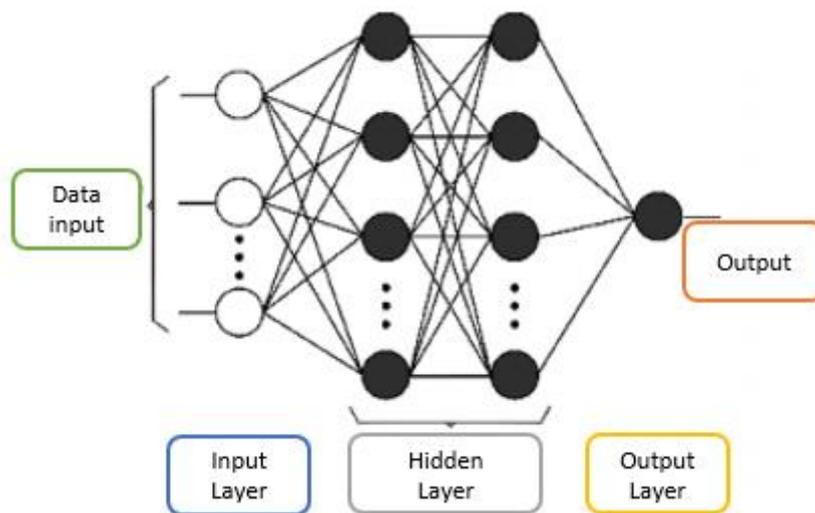
Two structures widely used in the literature for pattern recognition are MLP and SVM. An artificial neural network of the MLP type is made up of source nodes, forming the network's input layer, one or more hidden layers and an output layer. Among these layers, only the entry layer has no neurons in its constitution, unlike the others that have computational capacity. In order to be able to recognize patterns, a neural network needs to go through training steps, in order to modify the weights of neurons so that they perform the desired task.

Figure 1 shows the architecture of an MLP neural network with an input layer, two hidden layers and an output layer (Bisong, 2019). As soon as the input is received, the neurons send a signal to the posterior layer, through an activation function. Then, the signals propagate

through the neural network until they reach the output layer, calculating the output signal from the network.

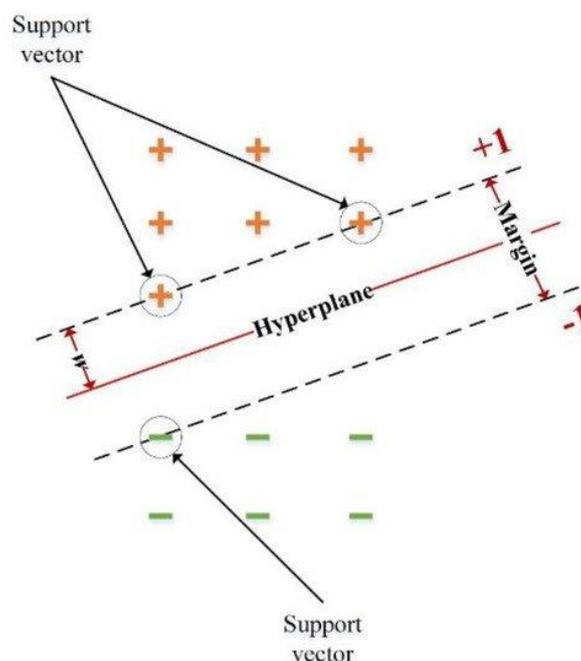
According to Armaghani et al (2020), SVMs are hyperplanes that separate the data to be trained in a limited space and with a maximum margin. In each of the hyperplanes, the vectors present in it receive a label where, on the one hand they receive -1 and on the other hand +1. The vectors on one side were labeled “+” and on the other side, they were labeled “-”, with a margin that separates them, as shown in Figure 2. For more complex cases, Kernel operators are used to aid in the separation process.

Figure 1: Graphical representation MLP with two intermediate layers.



Source: (Braga et Al. 2007 - Adapted)

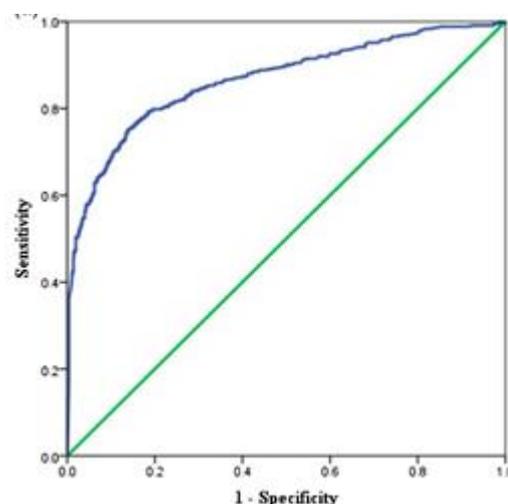
Figure 2: Linear SVM



Source: (Armaghani et al, 2020)

According to Janssens & Martens (2020), the ROC curve is a technique used to visualize, organize and select classifiers according to their performances. A binary classifier acts to obtain a mapping of the patterns for a predicted class. To distinguish between the current and predicted classes, the identification {Y, N} is used for the classes predicted as positive (p) and negative (n), respectively. The ROC curve is the graphical demonstration of the possible points of operation between classifiers by dimensioning points that refer to a presentation of the TP (True Positives) rate versus the FP (False Positives) rate for each decision threshold. The TP rate is usually referred to as "Sensitivity" and the FP as "Specialty". The curve starts at the point (0,0) and uses the concept generated by the confusion matrix; the decision threshold is set to 1 such that all standards are considered negative (figure 3).

Figure 3: Curve ROC and AUC.



Source: (Janssens & Martens, 2020)

One method to compare classifiers is to calculate the AUC, which can reduce the performance of the ROC curve to a single scalar value. Thus, the AUC is a part of the unit squared area, always staying between 0 and 1.

2.1.2 Selection of technologies for implementation

For the implementation of the resulting structure, some technologies were necessary:

Internet of Things – IoT: it is conceptualized as a global infrastructure that provides advanced resources through physical and virtual interconnection of objects based on the use of technologies that carry out communication and information processing. Identification, monitoring, communication network, computing, application, services and semantics are some of the subsystems available for the IoT architecture (Yu & Kim., 2019).

Raspberry PI: is a platform with a microprocessor system encapsulated in a single chip, composed with memories, clock and peripherals more limited than a computer, but with significant processing potential, providing flexibility and cost reduction. (Dow, 2018).

Speech API (Google): in order to translate a given voice, that is, to pass the audio content to a text content, it is first necessary to identify what that voice is transmitting. The Android



system, which belongs to Google, has its own speech recognition API, the Speech API, which has three main methods to perform speech recognition: Synchronous, Asynchronous and Transmission (Google Cloud, 2017). According to Google Cloud portal (2017), synchronous recognition is the simplest one to perform speech recognition. Synchronous request is a block. The Speech API normally processes audio faster than actual time. Thirty (30) seconds of audio are processed on average in fifteen (15) seconds. However, in cases of poor audio quality, the request may take longer.

2.1.3 Fundamentals of Image Law in Brazil

According to Rostelato (2016), image rights can be seen as any formal and sensitive expression of a man's personality. The idea of image is not restricted to visual representation only, but also by sound images, gestures and expressions. Therefore, as the system in question works with speech recognition in public transport, being careful with the information obtained is an essential part of the product, because to monitor possible robberies in real time, the system will be analyzing and recording everything that is said by citizens taking the public transport. Many personal matters can be recorded by the system.

Knowing how to separate what is significant for its correct functioning is of utmost importance so that this type of confidential information is not used in any way that could denigrate the image of the passenger and the professionals who are in the vehicle.

So, this is an experiment with capturing, recognizing and using voice, it is important to conceptualize the image right and its aspects in the legal aspect. Technological evolution and the development of the means of communication have made scholars of the legal environment begin to think about a way to preserve the right to use data disclosure, regardless of prior authorization from their owners (Siqueira & Guimarães. 2015). The Federal Constitution of 1988 mentions, in its Article 5, item X: "intimacy, private life, honor and image of people are inviolable, ensuring the right to compensation for material or moral damage resulting from their violation". However, even though it is an essential right, it is not considered an absolute right in the public interest, suffering some restrictions in this respect.

2.2 Experimental Methods

The type of research on which this work is based is exploratory research, where its methodology will take place in three stages. In step 1, the theoretical structure of collection and information and transformation process is defined (where the devices would be installed on the bus, how and why the data transmission would be). The second stage focuses on outlining and experimenting Neural Network and SVM configurations for a word dictionary, a methodology that is used in various sectors of the industry. Finally, in step 3, estimates of the implementation needs are made to configure the data transmission process as information consolidation.

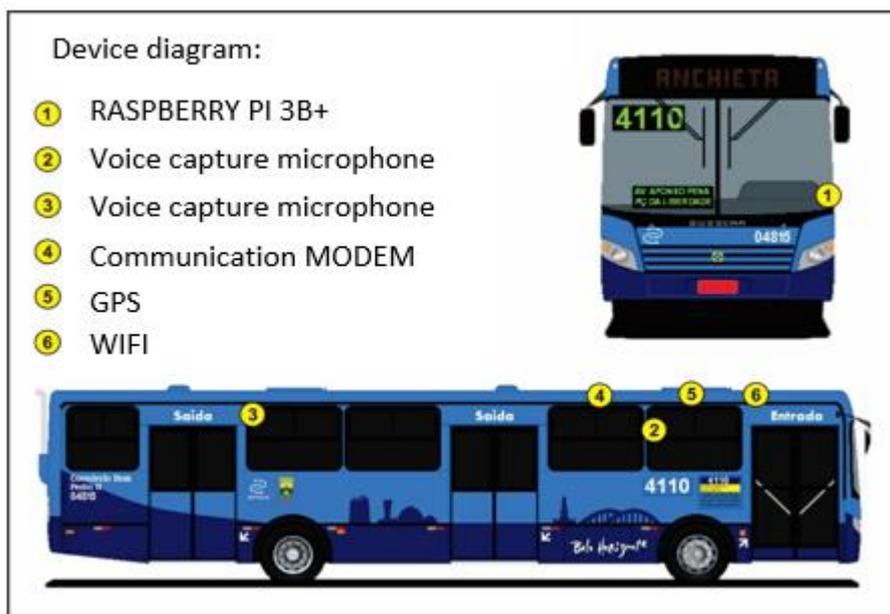
3. Results

3.1 Solution Structure

The availability of equipment in the public vehicle will consist of a Raspberry Pi, GPS / WI-FI modules, a communication modem and microphones for capturing ambient sound (Figure 4). Thus, based on the concept of ITS (Intelligent Transport System), this project contemplates a set of technological solutions and aims at security for users of public transport, also facilitating operational management by companies. It was found that most of the collectives in Belo Horizonte have the basic structure to receive the technological apparatus of this project (Figure 5).

Voice pickup would take place via microphones, transmitted to the Raspberry Pi board. After capture, the received data would be processed by Google's Speech Recognition API (Speech to Text), being compared to a trained database. Confirmed veracity, it would trigger information or an alert about the collective situation in real time (Figure 6).

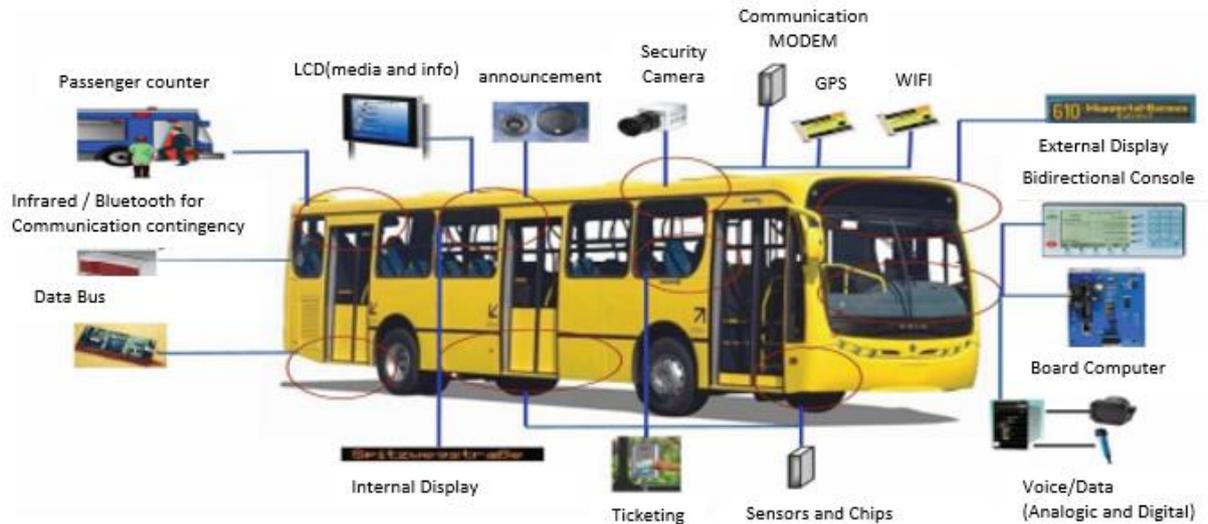
Figure 4: Diagram of Project Equipment



Source: (Authors)



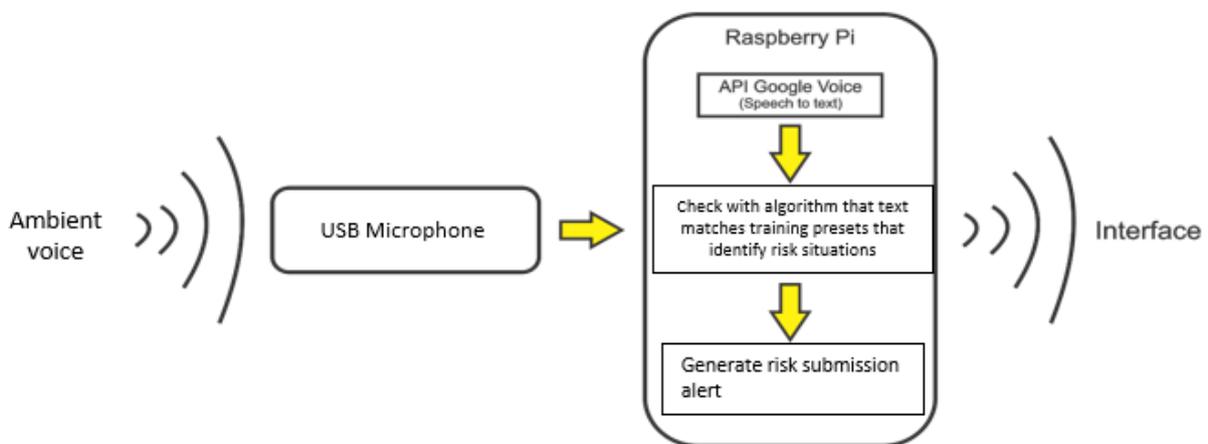
Figure 5: Technologies that can be used in Collectives



Source: (Portal da Prefeitura Municipal de Belo Horizonte, 2018 - Adapted)

The communication process between buses and interfaces would be based on continuous data transmission via GSM (Global System for Mobile Communications). A GSM module, coupled to the Raspberry card, would be responsible for maintaining an active data connection, allowing both the sending of information and remote access to devices embedded in the collectives.

Figure 6: Flowchart of the performance process.



Source: (Authors)



3.2 Evaluation of Pattern Recognition Algorithms

Experiments performed were intended to identify the classifier with the best accuracy for the proposed model.

3.2.1 Test Settings and Structure

The experiments were based on a desktop with the following configurations: Windows 10 Pro 64 Bit Operating System; Intel Core i5-6400 processor - 2.7 GHz; 8GB RAM Kingston Hyper X DDR4 - 2400 MHz; VGA Geforce GTX 1060 - 6 GB - 192-bit; Softwares - R Studio v_0.98.1103.0 / R X64 v_3.2.5. A database with a 57-word dictionary was used and it was the same database used in Souza et. Al. (2015), focused on classifying e-mails with SPAM. This basis became relevant due to two factors: it has an incidence of 57 words that can be in an e-mail (which can be linked to a dictionary of up to 57 words in the case of adapting the solution to a dictionary of terms used by burglars during the theft) and there is the dichotomy process, where there is only one classification for two types of situation (it is SPAM or it is not SPAM - which in a specific dictionary will be classified as in danger or not danger) .

3.2.2 Test Structure and Settings

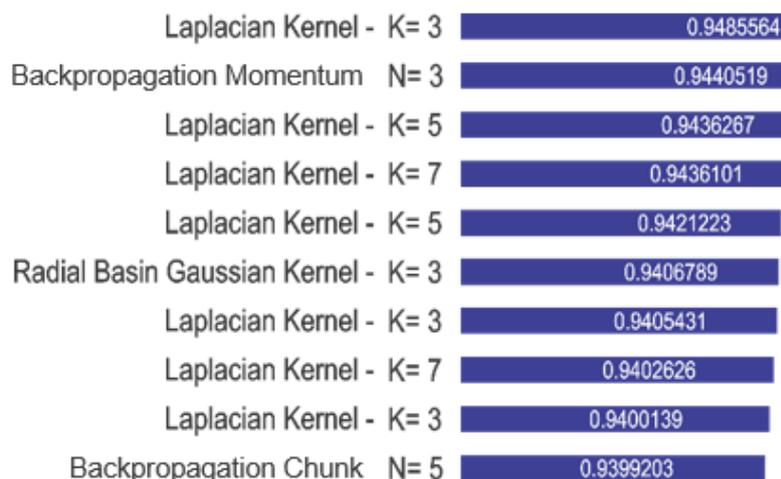
The algorithms used were the MLP algorithm and the SVM. The test for MLP was based on the number of neurons in the hidden layer, varying in 3, 5 and 7 neurons; the number of times varied from 1,000, 2,000 and 3,000 for each structure, the test and validation set between 35-65% and 25-75%. Six different types of MLP were also used: Backpropagation Batch, Backpropagation Momentum, Backpropagation with Weight Decay, Quickpropagation, Standart Propagation and Backpropagation Chuncck, resulting in 108 tests for MLP. When executing the algorithm for each of the listed functions, the Area Under Curve (or AUC) is obtained. The closer to 1 (one) the result, the better. The MLP tests, on the machine configuration mentioned above, took an average of 42.36 seconds.

As for the SVM, a smaller number of variables was used, consequently, resulting in a lesser number of tests. There is C the, which is the violation cost, varying in 3, 5 and 7. There is the K-Fold Cross Validation, varying in $k=\{3,5,7\}$. It has the kernel, varying in: Radial Basin Gaussian Kernel, Laplacian Kernel and Bassel Kernel (Totaling 27 tests). When executing the algorithm, as in MLP, the AUC is calculated. The closer to 1 (one) the result, the better. The tests performed on the SVM algorithm took, on average, 8.51 seconds in the machine configuration mentioned previously, making it considerably faster than MLP.

Classifying among all resulting tests, a ranking was obtained with the 10 best algorithms in execution times, listed in figure 7.



Figure 7: Ten best results between MLP and SVM.



Source: (Authors)

3.3 Deployment and Costs

3.3.1 Available Technologies and Their Combinations

The project is based on an IoT platform and the model evaluated for use was the Raspberry Pi 3 Model B, which has ANATEL (National Telecommunications Agency) certification and has several features previously described.

As it is a microcomputer, two variables may be adequate to optimize its use in the process of identifying risk situations by voice capture: combining the processing power of the Raspberry with the neural network algorithm now tested with the Google Cloud Speech to platform Text or attach to the hardware a Geetech speech recognition module, with built-in microphone.

In the Raspberry / Algorithm combination, the structure follows the flowchart of figure 6. The voice would be picked up by a microphone, which sends the data to the Raspberry; installed in it, the Google Voice API would encode the sound in text in real time, so that the Laplacian Kernel algorithm can compare it with the predefined data dictionary. If the comparison is positive, the algorithm generates an alert and sends an information bulletin to the competent bodies or owners of public transport companies through the network.

By varying the structure, the speech recognition module would be previously configured with words that refer to as-jumping behavior and, when identifying them, would send a command to the Raspberry, which would be in charge of communicating with the devices and of alert sending task. The advantage of using the algorithm over the module is given by its flexibility and learning power, as it is a neural network that can be trained. The module only has the option of storing up to 15 voice commands in its memory, which would greatly limit the scope of the problem. On the other hand, by being able to connect with the Raspberry from a distance, it would decrease the number of microphones to be installed in the collectives.



3.3.2 Implantation

As most of the buses in the capital of Minas Gerais have data collection technology, the implementation of the system would be linked to the devices responsible for transmitting data to a server, dealers, competent bodies or directly to the user.

Such a facility would decrease the cost of implementation, since a GSM / 3G module would be necessary to make the project with Raspberry totally independent from other technologies.

In an investigation with a collective in the capital, the possibility of housing the Raspberry next to the collector was envisaged, with a USB microphone attached to it. In order for the system not to be dependent only on events at the front of the collective, a second USB microphone would be installed next to the collective's rear door.

With regard to the voice recognition module, when coupled to the Raspberry, it would concentrate all the action only in the vicinity of the collector, since its module has a microphone in its structure and would do the entire procedure to identify the recurring situation. The sending of post-processing information would take place through the communication of the device with the network module belonging to the collective or with a GSM module destined for a device embedded with IoT.

3.3.3 Project Cost in its Variables

The feasibility of implementing the project studied so far is summed up in a structure based on the architecture of the IoTs, making the Raspberry an essential component in its assembly process. Glimpsing the project for 1 (one) and 100 (one hundred) collectives, budgetary surveys were carried out, in order to verify in real values the cost of the project in its variables.

As analyzed in topic 3.1, the proposed project is derived from three options, which are: Standard structure, with voice recognition module and independent. The standard structure of the project contemplates only the basics for the architecture itself, without counting on other components or specific modules for each purpose. Its cost serves as a basis to measure, roughly, the value of the basic hardware that would be installed on each bus (Table 1), emphasizing that they would communicate with the data transmission devices already loaded in the collectives. In the same system to connect to bus network devices, the structure with voice recognition module presents its cost values varying only when removing USB microphones and including the Geetech module that already has a built-in microphone (Tables 2 and 3).

Table 1: Cost Standard structure (Hardware).

Standard Structure			
Description	Units	Prize (R\$)	Total (R\$)
Raspberry Pi 3 Model B	1	229.90	229.90
Cartão Micro SD 32gb	1	89.90	89.90
Mini Microfone USB 2.0	2	38.50	77.00
Total			396.80

Source: (Authors)

Finally, evidenced as more complete and relatively more costly to implement, the independent structure presents values referring to all components that satisfy the project's ideology, without the need to communicate with devices already installed in the collectives (Table 3).



Table 2: Cost Structure with voice recognition module.

Structure – Voice Recognition Module (1 bus)			
Description	Units	Prize (R\$)	Total(R\$)
Raspberry Pi 3 Model B	1	229.90	229.90
Micro SD 32gb Card	1	89.90	89.90
Geetech Module with Microphone	1	199.90	199.90
Total			519,70

Source: (Authors)

Tabela 3. Cost of independent structure.

Independent Structure (1 bus)			
Description	Qtd.	Price. (R\$)	Total(R\$)
Raspberrry Pi 3 Model B	1	229.90	229.90
Cartão Micro SD 32gb	1	89,90	89,90
Mini Microfone USB 2.0	2	38.50	77.00
Módulo GPS	1	99.90	99.90
Módulo GSM/3G	1	129.90	129.90
Total:			626.60

Source: (Authors)

However, when measuring the value presented for each bus, the Google Speech API service should be added to this structure, which would capture and convert the audio in real time during the trips. This service has a differentiated feature with regard to billing, since its value is measured every fraction of seconds. In order to facilitate understanding, the amount was added and the calculation works with the value of the minute, already converted into real currency and based on the exchange rate in force at the time of making it (R\$ 3.80).

The spreadsheet shows the monthly service charge for each bus, with the assumption that the service would be active for 24 hours, which would result in 43,800 minutes (Table 4).

Table 4: Independent structure cost with addition of Google Speech API for 1 collective/month.

Independent Structure + Google Speech API			
Description	Units	Unid. (R\$)	Total(R\$)
Raspberrry Pi 3 Model B	1	229.90	229.90
Micro SD 32gb Class 10 Card	1	89.90	89.90
Mini Microphone USB 2.0	2	38.50	77.00
GPS Module	1	99.90	99.90
GSM/3G Module	1	129.90	129.90
Google Speech API (Minutes)	43800	0.024	3,994.56
Total:			4,621.16

Source: (Authors)

As the contracted service provides its customer with 1 million minutes, this amount could be divided between the lines of the concessionaire. For an exact value, if you divide the total value of the minutes that the plan has by the amount consumed by each bus monthly, there would be the possibility to serve with the same account 22.83 buses with only one contracted account. According to the company's website, if an IoT project has the need to use more than 1 million minutes, different values can be applied in the final cost pricing.



3.3.4 Limitations and Difficulties

Among the limitations and difficulties encountered in the process of adapting the project, three stand out. Firstly, the lack of a precise survey, with the competent security agencies, on a database with a dictionary of words that assist in the process of identifying patterns in theft of collectives. Secondly, the calibration of the sound pickup inside the collective, having seen the noise that some buses generate, which could impair the quality of word identification during the trip.

Finally, the complexity of training the pattern recognition algorithms to distinguish the tone of the voices captured during a trip, since the same word could be used both in a formal conversation and in the situation for which this study is intended.

3.3.5 Future Experiments

After obtaining experimental results and cost estimates of the project variables, some future projects became feasible. Firstly, the aim is to test the standard hardware structure in a real collective (with the appropriate licenses and permissions), in order to obtain performance data from the Laplacian algorithm operating on the Raspberry.

Another experiment will be based on the concept of Ontology, together with the competent security agencies, conducting a survey of requirements for a database with words that enrich a dictionary that makes it possible to identify patterns in cases of theft of collectives. The third perspective is to develop a kit of motion sensors within the collective, adding another technology in the process of identifying risk situations in public transportation. Thus, the development of the bus monitoring structure to make information available in real time about the status of each collective in circulation.

4. Conclusion

Such a solution features an automated and effective tool for combating crimes in public transport. This exploratory research returned a range of options relevant to its implementation.

Through an embedded technology device, based on IoT, there is the possibility of capturing, identifying and portraying (through sound) a possible situation that could be the result of an assault on an ongoing bus, without jeopardizing the physical integrity of its passengers. It was possible to identify structures that satisfy the proposed requirements, organized by flowcharts and execution capacity. Approaching 95% assertiveness with the proposed word dictionary for testing, the SVM Laplacian Kernel algorithm demonstrated a good training capacity and can very well be implemented, in order to assist in the recognition of patterns of this type of occurrence.

As a variant, a voice recognition module, with dictionary presets, is accredited to be a low-cost option for the project, respecting its configuration limitations. The best option found, in terms of assertiveness to satisfy the proposal, was with the Raspberry device together with the SVM Laplacian algorithm and the Google Speech to Text API.

Finally, a relevant consideration is the possibility for the tool to preserve passengers, collectors and drivers during the crime; however, the legal issues involved have considerable importance to be taken in future studies.



References

- [1] Armaghani, D. J., Asteris, P. G., Askarian, B., Hasanipanah, M., Tarinejad, R., and Huynh, V. V. (2020). Examining hybrid and single SVM models with different kernels to predict rock brittleness. *Sustainability*, 12(6), 2229.
- [2] Braga, A. P.; Ferreira, A. C. P; Ludermir, T. B (2007). *Redes neurais artificiais: teoria e aplicações*. Rio de Janeiro, Brazil:: LTC Editora.
- [3] Bisong, E. (2019). The Multilayer Perceptron (MLP). In *Building Machine Learning and Deep Learning Models on Google Cloud Platform* (pp. 401-405). Apress, Berkeley, CA.
- [4] Dow, C. (2018). *Internet of things programming projects: build modern IoT solutions with the Raspberry Pi 3 and Python*. Packt Publishing Ltd.
- [5] Fedrecheski, G.; Gonçalves, R. A. and Kuk, J. N. (N.D.) *Projeto de um Aplicativo Android de Monitoramento de Veículos do Transporte Público*. Departamento de Ciência da Computação, Universidade Estadual do Centro-Oeste. Guarapava-PR.
- [6] Google Cloud. *Speech-to-Text basics* (2017) Disponível em: <https://cloud.google.com/speech-to-text/docs/basics> .
- [7] Gularte, Á. C.; Ribeiro, V; G. and Silveira, S. R. (2016) *Um Protótipo para Monitoramento em Tempo Real do Transporte Público de Porto Alegre-RS por GPS (Global Positioning Sys-tem)*. RCT-Revista de Ciência e Tecnologia, v. 2, n. 2, 2016.
- [8] Janssens, A. C. J., & Martens, F. K. (2020). Reflection on modern methods: revisiting the area under the ROC curve. *International Journal of Epidemiology*.
- [9] Magalhaes, C. T. A. (2008) *Avaliação de Tecnologias de Rastreamento por GPS para o Monitoramento do Transporte Público por Ônibus*. Tese de Doutorado. Programas de Pós-Graduação em Engenharia de Transportes da Universidade Federal do Rio de Janeiro. Rio de Janeiro.
- [10] Prefeitura Municipal De Belo Horizonte. *Sitibus* (N.D.) Disponível em < <https://prefeitura.pbh.gov.br> .
- [11] Rostelato, T. A. (2016). Discursando sobre o direito à imagem: uma autêntica incidência de mutação constitucional. *Revista Direitos Sociais e Políticas Públicas (UNIFAFIBE)*, 4(1), 239-277.
- [12] Siqueira, M. S.; Guimarães, L. A. (2015) *Direito de Imagem*. *Revista Ciências Jurídicas e Sociais-UNG*, v. 5, n. 1, p. 04-10.
- [13] Yu, J. Y., & Kim, Y. G. (2019, January). *Analysis of IoT platform security: A survey*. In *2019 International Conference on Platform Technology and Service (PlatCon)* (pp. 1-5). IEEE.
- [14] Souza, F.H.B.; Maia, C.A.; Castro, C.L.; Braga, A. P.; Saldanha, R. R. (2015) *Proposta Orientada a Redes de Petri para Representação do Poder de Classificação*. In: *XII Simpósio Brasileiro De Automação Inteligente*, Natal.