Crowding Sourcing and Pattern Recognition for Identification and Sharing of Thief's Information

Marcelo Felix Assis de Fonseca[1]; Marcus Vinicius Fonseca Trindade[1], Luiz Melk de Carvalho[1], Diva de Souza e Silva Rodrigues[1], Flávio Henrique Batista de Souza[1]

[1]Centro Universitário de Belo Horizonte UNIBH, Brazil

Abstract.

Security solutions based on facial recognition are widely employed. However, due to the lack of information and appropriate structures, this technology does not benefit small and medium-sized Brazilian businesses (the victim of most criminal occurrences). This research demonstrates a solution focused on this audience and based on: affordable hardware device; Multilayer Perceptrons (MLP); Cloud Computing (CC) and Crowdsourcing (CS). The main structure was developed with: Python programming language; MySQL as database management system; and OpenCV (Open Source Computer Vision Library) to perform the real-time detection of the faces of the people who are entering the establishment. The images are recorded and later analysed by a second algorithm responsible for returning an image vector. This vector is compared with the others vectors already registered in a database with the list of suspects, returning the percentage of similarity between them. A similarity higher than 70% enable an alert to the manager of the establishment, discreetly, for later consultations of the owners, or by judicial order. The structure works in CC and the filling of the database is done via CS by the establishments. Experiments with MLPs were performed to optimize the recognition process, considering 5 types of MLPs (Backpropagation Standard, Momentum, Weight Decay, Quick propagation; Resilient Propagation); 50 to 500 epochs, 7 to 10 neurons, learning rate of 0.01%, 25-35% validation and 75-65% training. It was performed 155 training processes (total: 19 hours and 31 minutes of test execution time). A maximum accuracy of 94% was reached. The solution can be implemented and integrated to the available Google cloud services.

Keywords: Multilayer Perceptron, Crowdsourcing, Cloud Computing, Security Structures, Facial Recognition

1. Introduction

The use of crowdsourcing and pattern recognition allows a structure to assist small and medium-sized businesses in obtaining information about thieves actions. According to Mendonça (2013) and Garcia & Viecili (2018), the criminal acts rationally. While engaged in the activities, the criminal observes places with a low level of vigilance and presents local administrators.

An analysis focused in the agent that intends to commit and/or participate in an illegal activity should be considered as a choice problem involving two variables: the risk of compromising the physical integrity of the agent, and the penalty of overlapping rules with the legal system. It is a clear trade-off between return and risk. However, a study of this situation, based only on a financial point of view would result in an incomplete understanding of the problem, failing to identify many points by which the criminal process operates on small and medium businesses, such as the legal assessment of the process.

The general objective of this research is to demonstrate a structure capable to collect information through Crowdsourcing (CS) and, with the aid of facial pattern recognition algorithms, to feed a centralized database in order to score and identify the facial pattern of thieves. Thus, as specific objectives, can be cited: analyse the process of an establishment theft; designing a structure for a process of gathering information, consolidating and disseminating potential criminals; present a sequence of experiments with pattern recognition algorithms for face recognition and, finally, evaluate the legal terms that apply to the process.

This study is justified by the increase in burglary and robbery rates in small and medium-sized businesses. They don't possess accessible tools for the identification of criminals, figuring a real problem of contemporary society. Fear of assault and violence is the most worrisome and frightening, plaguing 46.7% of traders (Garcia & Viecili, 2018). However, the methodology of crimes committed against small businesses may contain information relevant to the development of computational structures for the dissemination of warnings, since negligent methods on the part of the majority of the mourners at the time of the crime.

Contemporary technological advances must help combat this social problem. According to Du et al. (2019), the concept of pattern recognition can be considered as the development of theories and techniques aiming at the construction of machines or devices capable of presenting characteristics similar to those of beings recognize patterns.

Thus, it is presented a tool option to identify lawbreakers. However, algorithms that use pattern recognition are dependent on the data sets for machine learning. For the consolidation of data, the concept of Crowdsourcing was analysed. The term Crowdsourcing (CS) was originally created by Howe (2006) who defined the concept as an act of developing an activity by a designated agent or outsourcing through a call to a large group of people. CS became a business model using collective intelligence to solve problems or complete activities (Acar, 2019). According to Schwartz et al. (2015), the application of CS provided a simple way to access, supply and generate knowledge, offering benefits to perform services, production of ideas or content from the request of the contribution of a group (Benedek et al., 2015). This study is focused on managing Crowd Sourcing for the recognition of criminals in small businesses among the respective owners of the establishments, in addition to the availability of information by means of Cloud Computing.

According to Baun and Kunze (2019), Cloud Computing is a term to describe a computing environment based on a huge network of servers, whether virtual or physical. There is a demand for the understanding of the volume of data that will be worked (and later the information generated) and the risks implicit to the process. It tends to adapt the best solution, in order to avoid that the user must have the minimum contact with the situation. The focus of the proposed solution is not to prevent robbery or theft, but rather to evaluate the way of

acting and, above all, the availability of relevant and truthful information about this type of crime though the facial recognition.

Based on that, some questions guide this research: what are the possible risk situations that occur during the assault procedure? What would the procedure for a facial recognition solution look like for a criminal, who would not expose the victim and share relevant information about the occurrence? Which performance the chosen algorithms can have of good relevance to be used in this type of research? How should it be, and what should be considered by the police force in the information that will be released? Such questions will be addressed throughout the present paper.

2. Methods

This paper represents an experimental evaluation. The results are presented in stages:

- Stage 1: An overview of the business rules of the proposed system is demonstrated, where the activity sequence is defined. The algorithm language for face recognition and the storage in a local database is already addressed. The necessary structure for the system is presented, where the software for facial recognition and the types and set of equipment (in detail) are defined;
- Stage 2: A battery of tests with the algorithms is performed. A reference database was evaluated in order to identify the pattern recognition structure that most closely resembles the ability to understand facial recognition. The objective of this experiment is to measure the efficiency of each structure by measuring the accuracy of the AUC (Area Under the Curve) and the processing time taken;
- Stage 3: The behaviour of the proposed solution is evaluated, analysing the impacts to those involved, through the information entered in the database by the users themselves, to make a study of how the risk situation by region can be presented in the form of results. In this evaluation, the use of Cloud Solutions is demonstrated.

In addition, the legal consequences of the proposal are discussed.

3. Results and Discussion

3.1 Overview of the proposed work

The proposal is a development of an application to assist small and medium-sized businesses where it will have a mechanism to identify, through cameras, people entering the establishments, storing their identification as vectors in a local database.

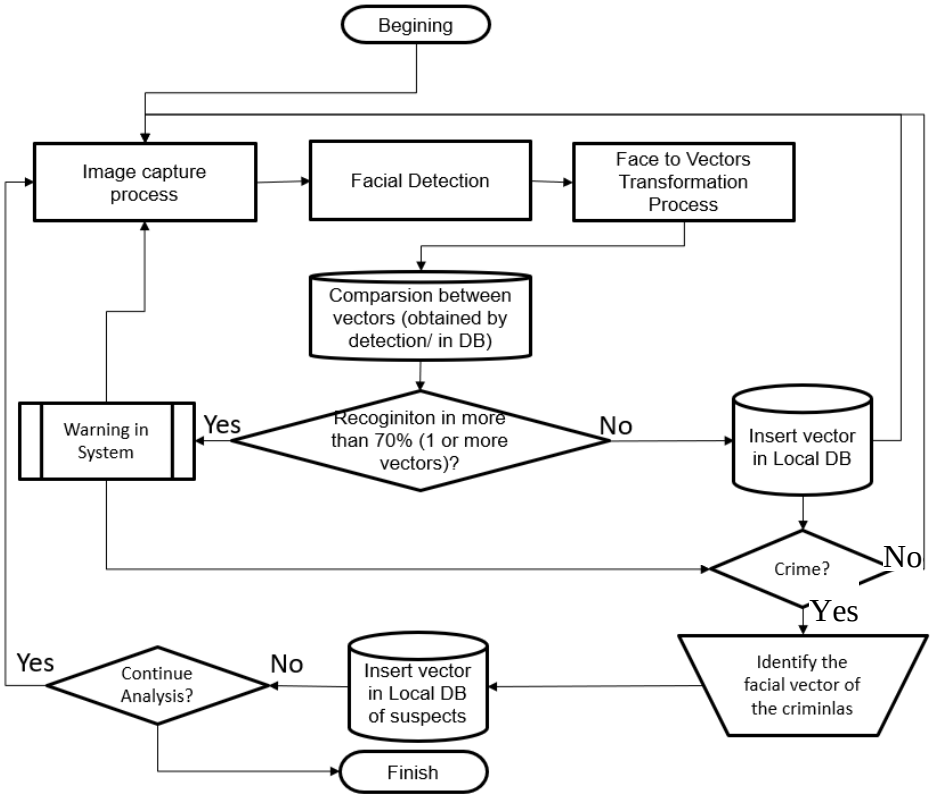
Figure 1 shows the process of capturing a user's image that occurs until it is stored in the database. Figure 2 shows a hypothetical arrangement of the equipment used in the solution. Thus, the characters of the process are defined. The user is the person that will have access to the images. It will be the responsibility of the owner of the business the identification of the "vectors" that store the information obtained in a secure database, since it can be used in cases of judicial verifications of the image disclosure.

Based on a classical closed-circuit television (CCTV) system, a collect process will be performed, with a database that has features similar to facial patterns, so that facial patterns

can be analysed. This database will be called "test", where it will test with cameras (of poor image quality) considering: it is a project aimed at small and medium-sized establishments, where they do not have cameras with good quality (HD - High Definition or 4K), and at the end a battery of tests with regard to the accuracy of the algorithm used for facial recognition, such as test addressed hypotheses of low, medium and high visibility.

The owners of the establishments would have a simple and safe interface (only they will have access).

Figure 1: Image Capture Process



Source: (authors)

Figure 2: Hypothetical Arrangement

Source: (authors)

This interface allows sending the videos of the security cameras in relation to the happened assault and will feed the test database. It is a responsibility of the owner of the establishment: send of the videos, which later could serve as an aid to the authorities, where there will be a term of responsibility to be defined, because with unauthorized disclosure of the citizen, the citizen may file a lawsuit against damages and injuries. The main structure of this project will be developed with Python programming language, the MySQL is used as database management system, where will be stored the facial vectors and incorporated into the proposed application and the OpenCV (Open Source Computer Vision Library – according Spizhevoi and Rybnikov (2018): developed by Intel in 2000 is a library that has data structure modules, linear algebra, GUI, image and video processing, among others).

Through the algorithm developed in OpenCV, it is possible to perform the real-time detection of the faces of the people who are entering the establishment. This image will be recorded and later analysed by a second algorithm responsible for returning an image vector. The image vector will be compared with the others vectors already registered in the database of suspects, returning the percentage of similarity between them. If the similarity is more than 70% an alert will be issued to the manager of the establishment discreetly for later consultations of the owners, or by judicial order.

3.2 Facial Recognition Experiments based on MLP

OpenCV is a library tool with algorithms for the recognition process. Considering the proposed work, it is relevant to make an evaluation of the algorithms performance. The proposal of these following results is to evaluate the performance of structures based on MLP.

The valuation performed is based on the AUC of each structure. The motivation of this experiment, as already mentioned, is to recommend the best structure (in a matter of time and quality of response measured by the AUC) to be used in the solution. With poor quality cameras or small storage and processing spaces, the solution to be implanted should maintain the capacity of evaluation, in face of a bearish structure.

During the process of selection of the algorithm that will integrate the core of the facial recognition solution, a definition of the experiments to be performed was previously made.

An analysis of the dataset to be used (its characteristics and singularities) was performed. Then a process defined the rules of the experiments. So, it was performed a sequence of experiments that correlate the Area Under Curve or AUC of each algorithm and the processing time spent, to have an analysis of the efficiency of each pattern recognition structure option.

3.2.1 Previews Dataset

Before starting the training of the MLP neural networks considering the different types (Backpropagation Momentum, Backpropagation with Weight Decay, Quick propagation and Resilient Propagation), it was obtained a facial data set on the website www.cs.umass.edu/lfw/. It contained 123 images, being 90 images of only person with different angles, and 33 images of different people. It was necessary to perform a process of evaluation of the facial dataset because it was not possible to use all images. There is a limitation on the algorithms that performed the experiments, and there is a need for all images to be in the same dimension.

The respective dataset consists of images of two dimensions (in pixels) 150X150px and 250X250px. In a process of transformation of the image into individual pixels, 67,500 pixels were obtained referring to the images of dimension 150X150px, and 187,500 pixels referring to the dimension of 250X250px. Having this limitation in view, the dataset was organized with images of dimension 150X 150px. Such criterion was chosen by the largest amount of images at this dimension. A total of 20 images of a single person were obtained, and 20 images of different people, bringing a balance between the images.

3.2.2 MLP Structure Calibration Experiments

A sequence of experiments was proposed. The first tests were done with the Backpropagation algorithm, working with 6 images of a certain person, but with different angles, and 2 images of different people from the first image, varying between: 50 to 500 epochs; 2 to 6 neurons; Resampling options: 25% validation and 75% training; 35% validation and 65% training.

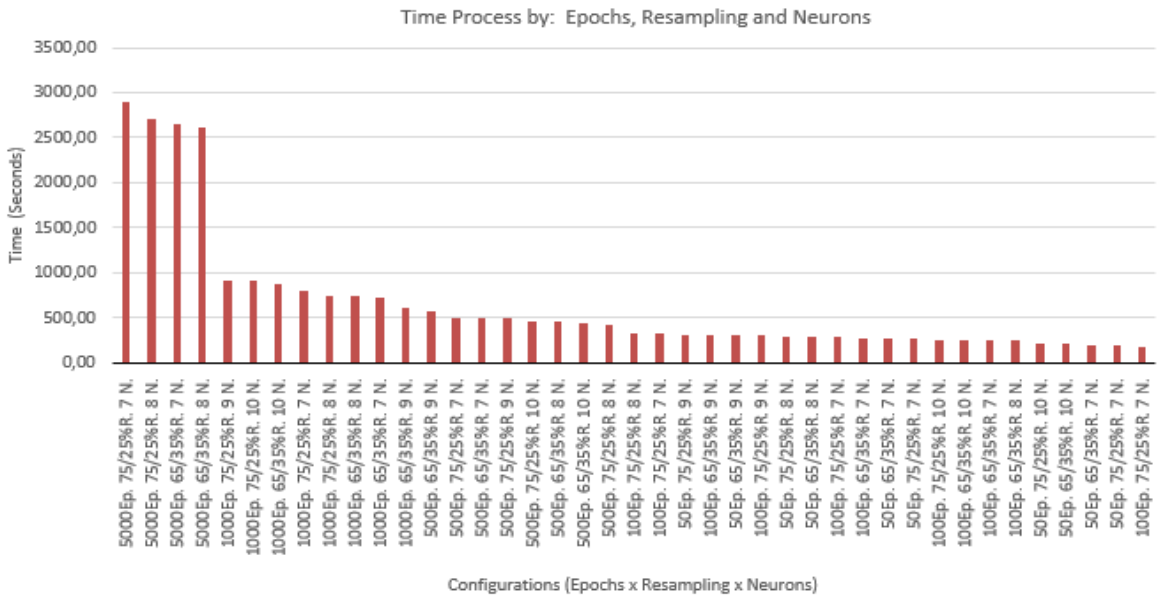
It was obtained an average run time of 4:24 minutes and these experiments reached AUCs lower than 0.3. For these experiments it was not possible to calculate the degree of similarity between the images. With an average of 40 tests in this first moment, some analyses were made regarding the parameters of the algorithm, concluding that:

- It was necessary to balance the images since the imbalance between them was directly influencing the AUC.
- The number of neurons was insufficient for the classification process via MLP.

In the second test, using the Backpropagation algorithm and a balanced dataset between 10 images of objective and 10 other images, with the configuration: 50 to 5000 epochs; 7 and 10 neurons; 25% validation and 75% training; 35% validation and 65% training;

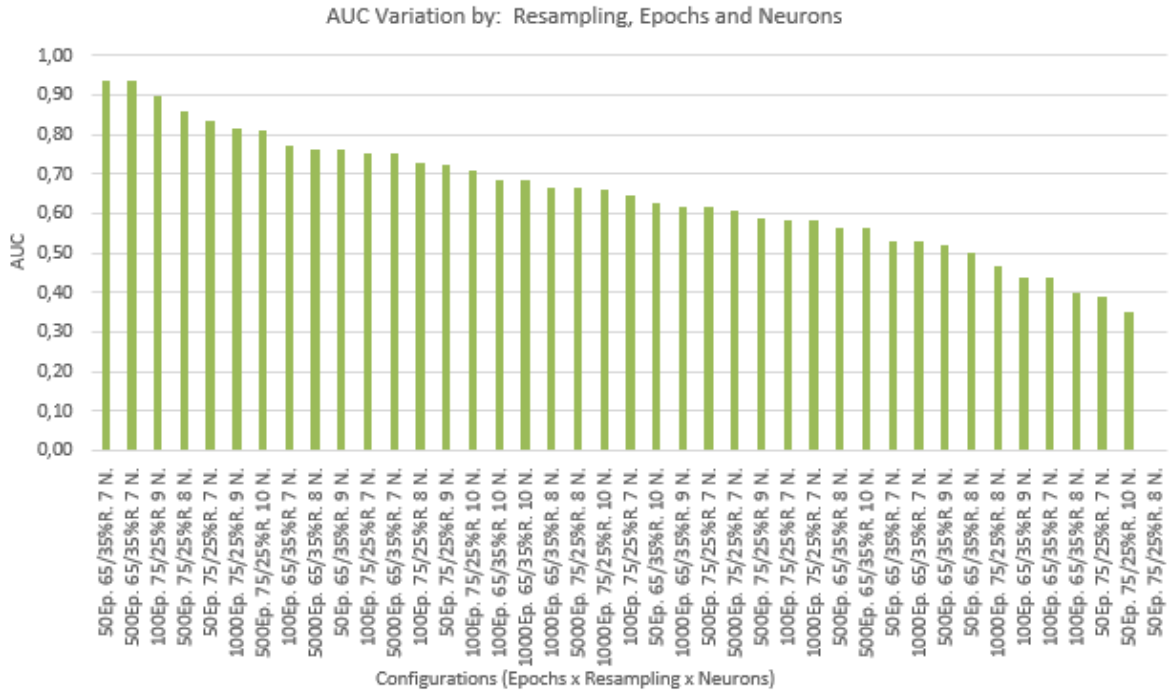
The AUC results ranged from 0.35 to 0.85. Taking those configurations as reference, it was proposed to use 20 images of the same person, and 20 images of different people for training. In a third test, it was possible to observe that the number of epochs was directly influencing the execution time of the algorithm, as shown in figure 3. However, it is possible to assimilate that the largest number of epochs is not precisely a guarantee of better AUC, as observed in figure 4, figuring a efficiency question.

Figure 3: Time variation by epochs, resampling and neuron



Source: (authors)

Figure 4: AUC variation by epochs, resampling and neuron.



Source: (authors)

Since the variations from 1,000 to 5,000 epochs have a training process achieving an execution time from 10 minutes to 50 minutes, the continuation of these respective epochs did not become viable, since the execution time was not influencing in a relevant way or positive in the AUC.

3.2.3 Efficiency Experiments

The previously process has guided the training with the other types of algorithms (Backpropagation Momentum, Backpropagation with Weight Decay, Quick propagation and Resilient Propagation). It was started in the following configuration: 50 to 500 epochs, 7 to 10 neurons, learning rate 0.01%, 25% validation and 75% training;35% validation and 65% training.

The graphs of Figures 5 and 6 show a box plot of the AUC and the time spent of those structures. In an analysis of the figure 5, it is observed that the algorithm Quick Propagation had a the best result when compared to the others (with maximum AUC of 0.9444).

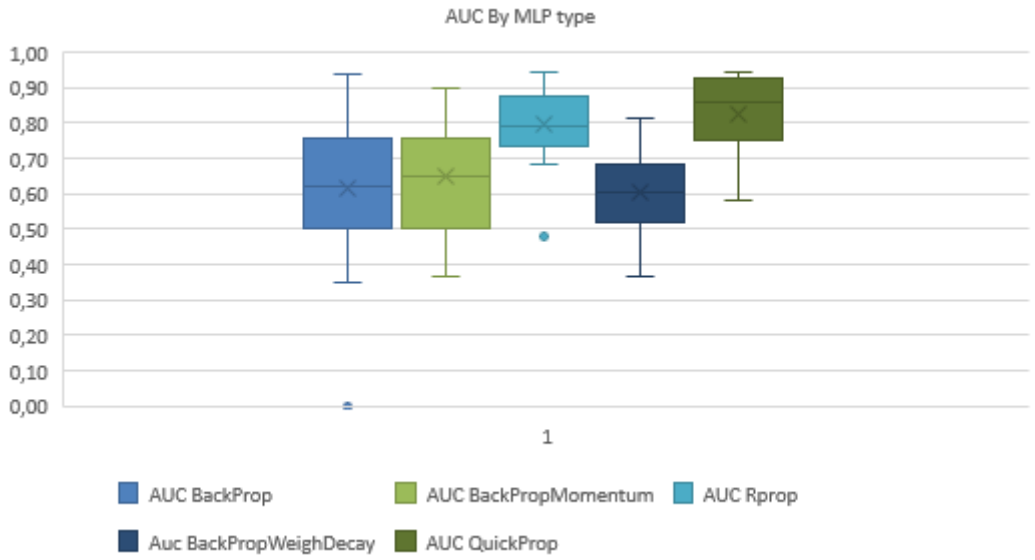
Its concentration varies between 0.7529 in its minimum value within the concentration, and 0.9255 in its maximum value within the concentration. It is also observed that this is the best concentration in relation to the other types.

Its AUC average reached 0.8571, figuring as the best average in relation to the other types of trained algorithms and has a minimum AUC of 0.5833 making the second algorithm with the best minimum AUC.

The type Backpropation with Weigh Decay did not obtain a result so relevant in relation to the other types. It is observed that its maximum AUC is 0.8125, with its concentration varying between 0.6883 as the maximum value and 0.5178 in its concentration. The mean AUC of

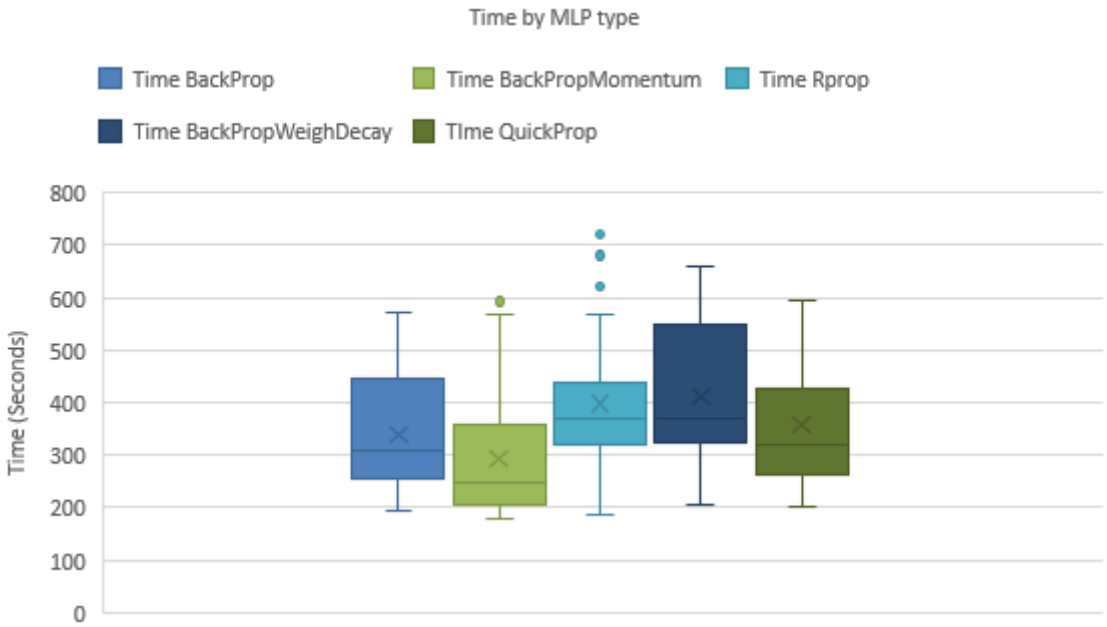
this type was 0.6041, and the second algorithm with one of the lowest AUC was the lowest, reaching 0.3650.

Figure 5: AUC variation by MLP type



Source: (authors)

Figure 6: Time variation by MLP type



Source: (authors)

The Resilient propagation algorithm has a maximum AUC of 0.9444 compared to the Quick Propagation and Backpropagation algorithms. It represents the highest AUC obtained during the training process of the algorithms. The AUC concentration ranged from 0.875 to 0.7366,

with the lowest variation being related to the other types. However, its mean is lower than the Quick propagation. The mean AUC is 0.7928 and the minimum AUC value is 0.6833. Note that there is a point outside the curve of the Resilient propagation algorithm, which means that it is not possible to have control over this point since it has moved away from the concentration, maximum, minimum and average value of the algorithm. This point corresponds toan AUC value of 0,4791.

The types Backpropagation Momentum and Backpropagation presented a slight difference in the concentration of AUCs.The Backpropagation Momentum algorithm varies its concentration between 0.7589 and 0.4992, with the average AUC of 0.6462, losing in the maximum AUC that reached the AUC of 0.9, but gaining in the minimum AUC with the value of 0.3666. The relationship between the Backpropagation algorithm and its AUCs concentration is between 0.76 and 0.50, but the result of the average of this algorithm was lower than the Backpropagation Momentum, with a value of 0.62, but there is a gain in relation to AUC maximum with 0.94, emphasizing that the AUC is equal for the types Quick Propagation and Resilient propagation, but loses in the minimum AUC relating to the Backpropagation Momentum. The Backpropagation had an outlier at 0.0.

3.3 Final Structure Definition

After Analysing the options of algorithms of pattern recognition, the other questions of the structure of the proposed solution are verified. A camera with integration to a Raspberry hardware for the identification of the faces was considered in the development of this work. The integrating of the Raspberry/camera can be obtained by the Raspberry PI camera. It is supported byRaspberry models PiB, B +, A +, 2.3B +, with a mean cost of R$ 240.00. This camera generates images up to 3,280X2,464px. The tests with the MLP algorithms were made with 150x150 pixels images and based on that it is necessary to configure the camera to capture images at a resolution of 150X150pixels.

As the focus of the camera is fixed, it is appropriate to position the camera strategically in order to capture images of any people entering the establishment, with enough face identification capability. A micro SD card with at least 64GB of storage space for the Ubuntu MATE OS install (OS) for being complete and compatible with the Raspberry Pi 2 and 3, together with a keyboard, mouse, monitor with HDMI input and cable HDMI, 5V source with micro USB output, Raspberry PI 2, desktop or notebook with card slot, requirements for the installation of Raspberry Pi according to the site www.robocore.net/tutoriais/instalando-ubuntu-na-raspberrypi.html, OS installation on Raspberry is done through the desktop or notebook.

The Raspberry it then connected to a computer/laptop for the installation of Python 3.6, because the MLP algorithm should be implemented in this language by reference to treat algorithms of facial recognition.

A network of at least 10Mbps of download and upload speed is necessary. It is worth mentioning that the operators in Brazil do not offer the same speed of download and upload. The upload is only 10% of the download and in this case a negotiation with the operator is necessary so that the 10Mb of uploading is guaranteed. Besides, if this amount of uploading is not guaranteed, there may be delays in sending the images to the cloud database service, since the images taken by the security camera can be generated up to 3280X2464px.Even with the

transformation of the images into matrices providing a smaller and redundant size, the upload is still affected and consequently the sending of the images is also affected, thus directly affecting the response time of the bank, and damaging the process as a whole.

Starting from the principle where the face of the citizen is recognized by the application developed and integrated in Raspberry, the image is sent to the Cloud database service to be stored and future consulted. In this work, for example, the Cloud service offered by Google was used.

This service is billed based on machine type parameters, virtual CPUs, RAM (GB), maximum storage capacity, maximum number of connections, price per use prolonging in (US $). In a simulation using machine type db-f1-micro, with shared virtual CPUs, 0.6 RAM (GB), maximum storage capacity 3,062 GB, maximum number of connections 250, price per use prolonging in (US $), the price of this simulation has the total cost of US$ 11.50 per month, with the possibility of paying this service per hour.

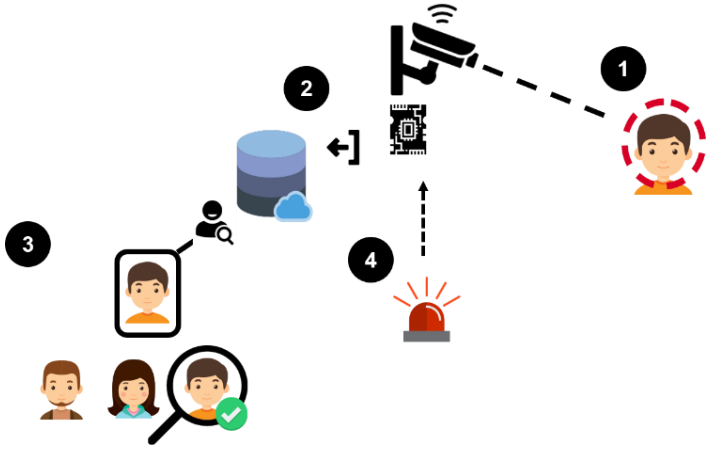
Also, Google offers US$ 300.00 for users to use in any of its services, being an opportunity to be dismembered. The application can still expand to an application using the result of the algorithm as an aid to the police.

Finally, the complete process is: The face of a person entering the establishment is detected with the facial recognition algorithm integrated with Raspberry along with the establishment's camera.

After the detection of the face the algorithm instantiates the database with the detected image at the moment, thus, not being able to have a resemblance to some criminal already registered in the base. After the first instance in the base, the algorithm is able to calculate a similarity between the person who has just entered the establishment with the database.

Returning results after people enter the establishment, it is possible to create an application that issues a warning to the owner of the establishment via SMS or WhatsApp, or even a partnership with the local police, letting the police themselves take action. The process is exemplified in the diagram of Figure 7.

Figure 7: Process of Evaluation



Source: (authors)

3.4 Legal Implications

Even with the structure defined, legal issues should be mentioned. The silent use of image or presumed consent is the most commonly practiced, although images of public persons are divulged every day in means of information dissemination, such as magazines, newspapers, among others, using as justifies the merely informative or educational use of the image, which is free in most countries. This use of the image for information purposes is lawful in Brazil and is linked to the collective right to freedom of information, which states that everyone has the right to inform and be informed.

On the one hand there is freedom of expression and the collective interest in access to information, and on the other hand it is necessary to consider possible situations of consumer embarrassment in the establishment, the right to privacy, as well as the objective honour of the citizen arrested in semi-open or open regime. Therefore, it is extremely important to consider the use that will be made of the image of third parties, according to Law 13.541/2003 applied in the state of São Paulo, determines the fixation of information boards at the entrance and exit of the monitored environments in these boards should contain the words "The establishment is being filmed" (Garcia & Viecili, 2018).

The images are confidential and protected in accordance with the law, the other laws that must be taken into account are officially recognized by the Brazilian civil code, in the Federal Constitution, art. 5, item X, the law provides that "privacy, private life, honour and image of persons shall be inviolable, assured the right to compensation for material or moral damage resulting from their violation". However, it is not taken into consideration only the public exposure of a person’s image to declare the direct violation of image according to the terms present in the Federal Constitution, art. 5, item X. In Brazil, a criminal action only takes place when the "victim" has evidence that the use of his image caused some real damage or damage in his life resulting from a certain action. In a legal and technological analysis is that the images will be stored in the cloud, already with the precept of people who have committed some crime. The security and reliability of the cloud service must be guaranteed, as it is a service aimed at society, which requires a greater evaluation of laws and terms.

4. Conclusion

The proposal developed in this work demonstrates one of the steps of the technology for the inhibition of robbery actions, configuring a solution option to protect society. The definition of the structure was based on solutions that are coherent with a situation present in small and medium-sized trades (low-power hardware, lack of information about technologies and access to contemporary innovations) Some MLPs structures do not have obtained a satisfactory results.

It was performed 155 training with a total of 19 hours and 31 minutes of execution time in the experiments. The solution can be implemented integrated to Google cloud services offered.

Finally, it is important to note the strong legal analysis demand of the involved processes. Such question It demands a thorough analysis and discussions with competent authorities, and if the information collected should be used by the police force.

As a future work of this project, a study with respect to the population of the database, when repairing the flow chart of the image capture process.

It is a fact that a conventional database will lose performance after a relative time, considering that on average 300 people enter a small establishment per day, and the images are captured from all these people, also considering an average image size of 2MG per image, remembering that the size of the image is relative to the quality of the camera, for 30 days, has an average disk space of the occupied database of 18GB per month, in view of this problem, it is an alternative to use a database with more performance, and consequently it has a way of storing the data, and that in the long run supports the storage of data.

References

[1] Garcia D. and Viecili J. (2018), "As consequências do assalto para o trabalhador do comércio vitimizado," *Revista Psicologia Organizações e Trabalho*, vol. 18, no. 2, pp. 396–402.

[2] Du, M., Liu, N., & Hu, X. (2019). Techniques for interpretable machine learning. *Communications of the ACM*, 63(1), 68-77.

[3] Howe J. (2006), "The rise of crowdsourcing," *Wired magazine*, vol. 14, no. 6, pp. 1–4.

[4] Acar, O. A. (2019), "Motivations and solution appropriateness in crowdsourcing challenges for innovation". *Research Policy*, 48(8), 103716.

[5] Schwartz, C., Borchert, K., Hirth, M. and Tran-Gia, P. (2015) "Modeling crowdsourcing platforms to enable workforce dimensioning," *in Telecommunication Networks and Applications Conference (ITNAC)*, 2015 International, pp. 30–37, IEEE.

[6] Benedek, A., Molnár,G., and Szuts, Z., (2015) "Practices of crowdsourcing in relation to big data analysis and education methods," in Intelligent Systems and Informatics (SISY), 2015 IEEE 13th International Symposium on, pp. 167–172,IEEE.

[7] Baun, C., and Kunze, M. (2019). Cloud computing Web-basierte dynamische IT-services. Springer-Verlag.

[8] Mendonça, M. J. C. d. (2013) "Criminalidade e violência no brasil: uma abordagem teórica e empírica," *Revista Brasileira de Economia de Empresas*, vol. 2, no. 1, 2013.

[9] Spizhevoi, A. and Rybnikov, A. (2018). *OpenCV 3 Computer Vision with Python Cookbook: Leverage the power of OpenCV 3 and Python to build computer vision applications*. Packt Publishing Ltd.