UNIVERSIDADE DE SÃO PAULO
POLYTECHNIC SCHOOL

JOAQUIN BARREYRO

Vehicle classification based on binary images of optical barriers using convolutional neural networks

São Paulo

2022

JOAQUIN BARREYRO

Vehicle classification based on binary images of optical barriers using convolutional neural networks

Revised Version

Dissertation presented at the Polytechnic School, Universidade de São Paulo to obtain a degree in the Master of Science

Supervisor: Prof. Dr. Leopoldo Rideki Yoshioka.

São Paulo

2022

JOAQUIN BARREYRO

Vehicle classification based on binary images of optical barriers using convolutional neural networks

Revised Version

Dissertation presented at the Polytechnic School, Universidade de São Paulo to obtain a degree in the Master of Science

Area of concentration: Electronic Systems Engineering

Supervisor: Prof. Dr. Leopoldo Rideki Yoshioka.

São Paulo

2022

.

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 02 de MARÇO de 2022

Assinatura do autor: _____

Assinatura do orientador: _____

Catalogação-na-publicação

Barreyro, Joaquín
Vehicle classification based on binary images of optical barriers using convolutional neural networks / J. Barreyro -- versão corr. -- São Paulo, 2022.
p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Sistemas Eletrônicos.

1.Sistemas Inteligentes de Transporte 2.Classificação de veículos 3.Redes Neurais Convolucionais I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Sistemas Eletrônicos II.t.

Name: BARREYRO, Joaquin

Title Vehicle classification based on binary images of optical barriers using convolutional neural networks

Dissertation presented at the Polytechnic School, Universidade de São Paulo to obtain the degree in the Master of Science

Approved on: 01 / 20/ 2022

Examination committee

Prof. Dr. LEOPOLDO RIDEKI YOSHIOKA

Institution: USP – UNIVERSIDADE DE SÃO PAULO

Assessment: APPROVED

Prof. Dr. CLEDSON AKIO SAKURAI

Institution: UFSP-Santos

Assessment: APPROVED

Prof. Dr. ALESSANDRO SANTIAGO DOS SANTOS

Institution: IPT

Assessment: APPROVED

ACKNOWLEDGMENTS

To Prof. Dr. Leopoldo Rideki Yoshioka for dedication and guidance.

To Prof. Dr. Claudio Luiz Marte for his cooperation, guidance, and support.

To Tecsidel, for the support and trust.

RESUMO

Os Sistemas Inteligentes de Transporte (ITS) são um conjunto de ferramentas que facilitam a aquisição, processamento, integração e disponibilização da informação para os usuários e a administração do sistema de transporte. Uma das funcionalidades do ITS é voltada a identificar e classificar os veículos. No Brasil, as agências reguladoras e fiscalizadoras de transporte definem um conjunto de categorias de veículos baseadas na quantidade de eixos e rodas duplas. Dentre as diferentes abordagens existentes para classificar veículos, as soluções com melhor desempenho alcançam índices de acerto que variam entre 95% e 99%. A maoiria das estratégias utilizam sensores de pressão (piezo elétricos ou mecânicos) instalados no pavimento que detectam a passagem da roda do veículo. Esse tipo de solução apresenta alguns inconvenientes, como a impossibilidade de contar eixos suspensos, a degradação com o tempo e perda de confiabilidade. Diante disso, é necessário o desenvolvimento novas formas de classificação automática de veículos que trafegam pelas rodovias. Este trabalho propõe desenvolver um novo método de contagem de eixos veiculares em sistemas de análise de tráfego. Para isso foi desenvolvido um método de classificação baseado somente em imagens binárias dos perfis dos veículos, extraídas a partir de uma cortina constituída por barreiras ópticas. O método proposto utiliza uma rede neural convolucional (CNN). Utilizamos a rede AlexNet modificada, onde a última camada da rede foi adaptada ao padrão de classificação de veículos adotado pela Agência Reguladadora de Transporte do Estado de São Paulo (ARTESP). O treinamento da rede neural foi feito utilizando o método de *Transfer Learning* e a técnica de *Data Augmentation*. O método proposto foi avaliado utilizando um conjunto de 5329 imagens constituídas por 11 categorias de veículos. A acurácia do resultado alcançado foi de 98,39% de acertos, o que indica que a substituição de um conjunto de sensores pelo tratamento dos dados somente da barreira óptica poderá ser uma alternativa factível.

Descritores: Sistemas Inteligentes de Transporte; Praças de Pedágio; Contagem de eixos, Data Augmetation; Transfer learning.

ABSTRACT

Intelligent Transport Systems are tools that facilitate the acquisition, processing, integration, and availability of information to manage the road transport network. A part of these systems is dedicated to identifying and classifying vehicles. The regulatory and inspection agencies of transport in Brazil define a set of vehicle categories based on the number of axles and double wheels. There are different approaches to solving the problem of classifying vehicles. The best score among the solutions found varies between 95% and 99%. These strategies use discrete sensors installed on the pavement, activated by mechanical effort, in the vehicle wheel's passage. This approach has some drawbacks, such as the impossibility of counting suspended axles, time degradation, and reliability loss. Thus, it is necessary to continually develop and improve the tools and computational techniques that allow the automatic classification of vehicles traveling on highways. We propose developing methods and computational models to improve vehicle axle counting in traffic analysis systems - a classification method based on binary images of vehicle profiles extracted from a curtain composed of optical sensors. The proposed method uses a modified AlexNet convolutional neural network. The last layers of the network were adapted to the vehicle classification problem. To train the network, we use transfer learning and data augmentation techniques. The method was tested using 5329 images consisting of 11 vehicle categories. Results achieved 98.39% accuracy, indicating that replacing a set of sensors by processing data only from the optical barrier may be a feasible alternative.

Keywords: Intelligent Transport Systems; Transfer learning; Data Augmentation; Axle Count; Vehicle Classification.

LIST OF FIGURES

Fig 1. Vehicle detection and classification sensors 22
Fig 2. Simplified Layout of axle counting, double wheels, and direction travel 23
Fig 3. Imaging by using an optical barrier 23
Fig 4. Examples of vehicle profiles made with an optical barrier 24
Fig 5. CNN architecture diagram 25
Fig 6. Quantity of papers by year and by source. 30
Fig 7. Classification method diagram with binary images and CNN classification 42
Fig 8. Binary-image components illustration 43
Fig 9. Speed effect illustration on imaging with the Optical Barrier 44
Fig 10. Flowchart of the training method and classification 46
Fig 11. Image of the installation of the optical barrier 47
Fig 12. Different CNN architectures. Relative Prediction time vs. Accuracy 48
Fig 13. Examples of images with artificially generated noise 50
Fig 14. Accuracy comparison for different epochs and mini-batches 54
Fig 15. Confusion matrix for the CNN of 32 samples per mini-batch with 29 training epochs in experiment 1. 55
Fig 16. Confusion matrix for the CNN of 32 samples per mini-batch with 29 training epochs in experiment 2. 56
Fig 17. Accuracy by hour on P2. 57
Fig 18. Images were obtained in a rainy period in experiment 2 57
Fig 19. Example of color highlighting images with length information as a background. 59
Fig 20. Confusion matrix. Images with length information. Experiment 3. 59
Fig 21. Category 1 truck, detected as category 2 60
Fig 22. Visualization of the image dataset 71
Fig 23. Print Screen of the first page of the paper published in INDUSCON 72
Fig 24. Print Screen of the first page of the paper published in ITSC 72
Fig 25. Train and Classification algorithm main process flowchart 73
Fig 26. Toll system Architecture 74

LIST OF TABLES

Table 1.	*ARTESP vehicles classification categories*	*21*
Table 2.	*AlexNet Layers*	*26*
Table 3.	*Primary studies groups*	*31*
Table 4.	*Primary studies Group 1*	*31*
Table 5.	*Primary studies Group 2*	*32*
Table 6.	*Primary studies Group 3*	*32*
Table 7.	*Primary studies Group 4*	*32*
Table 8.	*Primary studies Group 5*	*33*
Table 9.	*Modified layers comparison*	*48*
Table 10.	*Number of images used in CNN training and validation*	*51*
Table 11.	*Hyper-parameters used to train the new CNN*	*51*
Table 12.	*The number of images and period of observation in experiment 2.*	*52*
Table 13.	*Number of images used in CNN training Experiment1 vs Experiment 3*	*53*
Table 14.	*Experiment 2. Accuracy by lane type.*	*58*
Table 15.	*Higher Accuracy studies*	*61*

LIST OF ABBREVIATIONS

ARTESP	Transport Agency of the State of São Paulo
CNN	Convolutional Neural Networks
DBN	dynamic Bayesian network
DFT	Discrete Fourier Transform
FBG	Fiber Bragg Grating
FCM	Fuzzy c-mean algorithm
FHT	Fast Hough Transform
FHWA	Federal Highway Administration Research and Technology
GMM	Gaussian Mixture Model
ILSVRC	Imagenet Large Scale Visual Recognition Challenge
ITS	Intelligent Transport Systems
MLP	Multi-Layer Perceptron
PCA	Principal Component Analysis
PVR	Per Vehicle Review
SGDM	Stochastic Gradient Descent with Momentum
SLR	Systematic Literature Review
SVM	Support Vector Machine

Table of contents

RESUMO **7**
ABSTRACT **8**
LIST OF FIGURES **9**
LIST OF TABLES **10**
LIST OF ABBREVIATIONS **11**
1 INTRODUCTION **14**
1.1 CONTEXT 14
1.2 AUTOMATIC VEHICLE CLASSIFICATION 14
1.3 PROBLEMS WITH THE TRADITIONAL APPROACH 15
1.4 ALTERNATIVE APPROACHES 15
1.5 OBJECTIVES 17
2 THEORETICAL BASIS **18**
2.1 CONCEPT OF VEHICLE CLASSIFICATION 18
2.2 MARKET TECHNIQUES 21
2.3 OPERATION OF THE OPTICAL BARRIER AND IMAGING 22
2.4 CONVOLUTIONAL NEURAL NETWORKS 24
2.4.1 AlexNet 25
2.4.2 Transfer Learning 26
2.4.3 Data Augmentation 27
2.4.4 Training, Epoch, and mini-batch 27
2.5 LITERATURE REVIEW, RELATED WORKS 28
2.5.1 Description of the SRL protocol 29
2.5.2 SRL Results 30
2.5.3 Primary studies based on images 33
2.5.4 Primary studies based on lasers and optical sensors 35
2.5.5 Primary studies based on magnetic and inductive sensors 36
2.5.6 Primary studies based on pressure, piezoelectric and pneumatic sensors 38
2.5.7 Other Technologies 39
2.5.8 Analysis 40
3 METHODOLOGY **41**
3.1 PREMISES, HYPOTHESES, AND CONSTRAINTS 41
3.2 METHODOLOGY 41

3.2.1 Imaging with optical barriers 41
3.2.2 Classification using CNN 45
4 EXPERIMENTS 47
4.1 EXPERIMENT 1 49
4.2 EXPERIMENT 2 51
4.3 EXPERIMENT 3 52
5 RESULTS 54
5.1 DETERMINING THE MINI-BATCH AND THE NUMBER OF EPOCHS 54
5.2 RESULTS OBTAINED IN EXPERIMENT1 54
5.3 RESULTS OBTAINED IN EXPERIMENT2 55
5.4 RESULTS OBTAINED IN EXPERIMENT3 58
5.5 A QUALITATIVE FAILURE ANALYSIS 60
5.6 ANALYSIS AND DISCUSSION OF RESULTS 61
6 CONCLUSION AND FUTURE WORK 63
REFERENCES 64
ANNEX 1 – DATASET DESCRIPTION 71
ANNEX 2 – PUBLICATIONS DERIVED FROM THIS RESEARCH 72
ANNEX 3 – CODE IMPLEMENTATION DETAILS 73
ANNEX 4 – TOLL COLLECTION SYSTEM DESCRIPTION 74

1 INTRODUCTION

1.1 CONTEXT

Brazil has a road network of approximately 1.8 million kilometers. Road transport has a greater weight in the transportation matrix with 61% of goods movement and 95% of passengers (CONFEDERAÇÃO NACIONAL DE TRANSPORTE, 2020). More than 1 million trucks circulate on Brazilian roads, and the total fleet of motor vehicles is approximately 50 million (DEPARTAMENTO NACIONAL DE TRÂNSITO, 2015).

To manage a road network with this extension, and with this high number of vehicles circulating, a wide range of tools for managing transport networks called Intelligent Transport Systems (ITS) are needed. Information acquisition, processing, integration, and delivery are at the heart of ITS (MARTE; MELLO; FERREIRA, 2012). The ITS systems can be divided into five functions: electronic toll payment, traffic management, transport inspection management, traveler information, emergency services, and user support (MARTE; MELLO; FERREIRA, 2012). At least three of these five groups are interested in identifying and classifying vehicles. Most federal and state transportation regulatory and inspection agencies in Brazil define a set of vehicle categories based on the number of axles and double wheels (AGÊNCIA NACIONAL DE TRNASPORTES TERRESTRES, 2015), (AGÊNCIA DE TRANSPORTE DO ESTADO DE SÃO PAULO, 2015). It is important to note that ITS involves transport systems that cover all modes of transport. This work is within the scope of the road ITS.

Vehicle classification systems that consider the number of axles is not exclusive to Brazil. One classification system that serves as a reference for many studies is defined by the Federal Highway Administration Research and Technology (FHWA) of the United States (FEDERAL HIGHWAY ADMINISTRATION, 2013).

1.2 AUTOMATIC VEHICLE CLASSIFICATION

The automatic classification of vehicles based on the number of axles is performed by processing information obtained from a set of sensors. Sensor technologies can be classified as intrusive and non-intrusive (BALID; TAFISH; REFAI, 2017). Intrusive sensors are embedded in the pavement, such as a pneumatic tube, piezoelectric sensor, and inductive loop, and require traffic interruption for installation and

maintenance. Non-intrusive sensors are based on remote observations. The most common technologies are infrared sensors, laser, microwave radar, Doppler radar, and image processing. They are usually installed close to or above the path of vehicles (SKSZEK, 2001).

1.3 PROBLEMS WITH THE TRADITIONAL APPROACH

The traditional axle counting approach is carried out with piezoelectric sensors or a variation. They are discrete contact or pressure sensors activated by mechanical effort. They are arranged perpendicular to the traffic direction. The drive is produced when the wheel passes over the sensor (LEDUC, 2008). This solution has been applied to Brazil's toll collection and control systems for many years and has become standard. This type of sensor has some drawbacks; for example, it cannot detect suspended axles. Many vehicles can temporarily suspend some axles, which prevents their detection by piezoelectric-based systems.

Another problem is related to the intrusive nature. The sensors need to be within the lane, making installing and maintaining them difficult. It is also relevant to point out that this type of sensor is degraded with increasing noise and, consequently, reliability. The lifespan is related to the number of vehicles that pass.

1.4 ALTERNATIVE APPROACHES

Due to the problems previously exposed, alternatives for axle counting are analyzed. Among the possibilities, we can mention the use of inductive loops. They are widely used in vehicle detection. Its primary function is to detect presence. It is possible to capture a vehicle's magnetic profile, obtain the signal's specific statistical characteristics, and infer the classification with a database of previously determined characteristics grouped by class. Inductive loops are found within the intrusive sensors group; therefore, they have the disadvantage of increased maintenance and installation difficulties (LEDUC, 2008).

Another kind of sensor that deserves attention is the individual optical sensor. It is used similarly to piezoelectric ones, with the advantage that it can be installed off-lanes (non-intrusive) and at different heights, allowing the detection of suspended axles.

Another solution found is based on Optical Barriers. They are formed by a set of optical sensors arranged vertically. They are placed next to the passageway (non-intrusive sensors). The vehicle's passage interrupts the light beams, producing a signal, which, viewed over time, constitutes a 2D profile of the vehicle's passage. It is possible to count axles and estimate the double wheel by processing this profile.

It is also possible to use lasers. Logical operation is similar to that of optical barriers. They have different installation characteristics, and as a result, they may be more suitable for specific locations. They can be placed on posts on the side of the road or porches over the road. Like optical barriers, they are within the group of non-intrusive sensors (MIMBELA; KLEIN, 2007).

The installation of video cameras at various points on the road has become common in traffic monitoring. Due to computers' low cost and increased processing capacity, there is a fast-growing collection of computer vision techniques and tools. However, the algorithms are not yet robust enough to solve the main difficulties found in vehicular traffic images, such as: operating in any weather condition, lighting variation, object occlusion, shadows, detection of objects that stop or start the movement in the image, among others (FEDERAL HIGHWAY ADMINISTRATION, 2013), (MIMBELA; KLEIN, 2007).

The approaches presented are not frequent among toll collection and control systems yet. Without a definitive solution, classification techniques with cameras, optical barriers, and lasers are still being developed. Some recent works in this sense can be found at (GOTHANKAR; KAMBHAMETTU; MOSER, 2019) (SASONGKO; FANANY, 2019) (NEZAFAT; SAHIN; CETIN, 2019) (LIU; HUYNH, *et al.*, 2020) (KHATRI; SOMANI, 2015).

Considering there is still room for improvement, we intend to find a way to classify vehicles using a single sensor that is not intrusive, with accuracy according to state of the art. The optical barrier complies with these requirements. Furthermore, classifying vehicles with a single non-intrusive sensor simplifies the lanes' operation, helps in operation and maintenance, guarantees longer lane time, and increases lifespan.

Regarding viability, access to data from several highway concessionaires is available. It is also possible to count on the support of a company that develops toll collection systems. Together, it is possible to obtain data from current systems, and to

test a proposed solution in a real environment, to be able to analyze and compare the results.

1.5 OBJECTIVES

This research aims to develop computational models to improve the correct vehicle axle counting rate in traffic analysis systems based on a binary image. The specific objectives are as follows:

- Assess the current techniques used (state of the art) to count axles and determine the appropriate approach concerning the sensor's choice to be used.
- Build a database containing vehicle information (photo; type; axle number and double wheels) and sensor signals from data collection to be carried out in the field and laboratory.
- Develop a controlled test environment that allows the generation of signals to study different situations, explore limit conditions, and confront other types of information.
- Identify behavior patterns from the database, infer an axle counting process, and develop a new algorithm.
- Validate the new algorithm using the database.

2 THEORETICAL BASIS

This chapter presents some concepts necessary to understand this research better - first, the main concepts related to vehicles' classification. The following is an analysis of market techniques for vehicle classification. Then we will present an introduction to the operation of an optical barrier and how images are generated. We will also show some basic concepts about machine learning techniques like Convolutional Neural Networks (CNN), AlexNet, Transfer Learning, and Data Augmentation. Finally, we present the State of the Art in vehicle classification based on the number of axles.

2.1 CONCEPT OF VEHICLE CLASSIFICATION

There are numerous ways to classify vehicles. Depending on the objective, it is possible to observe different characteristics and group them. In the Brazilian Traffic Code, Chapter IX, section I art. 96 vehicles are classified by:

Regarding traction:

a) self-propelled;
b) electric;
c) human propulsion;
d) animal traction;
e) trailer or semi-trailer.

Also as to the species:

a) passengers:
- bicycle;
- moped;
- scooter;
- motorcycle;
- tricycle;
- quadricycle;
- car;
- microbus;
- bus;
- tram;

- trailer or semi-trailer;
- cart.

b) cargo:

- scooter;
- motorcycle;
- tricycle;
- quadricycle;
- pickup truck;
- truck;
- trailer or semi-trailer;
- wagon;
- handcart.

c) mixed:

- bus;
- utility;
- others;

d) competition;

e) traction:

- tractor truck;
- wheeled tractor;
- crawler tractor;
- mixed tractor;

f) special;

g) collection.

Regarding the category:

a) official;

b) diplomatic representation, consular career offices, or organizations international instruments accredited to the Brazilian Government;

c) private;

d) rent;

e) learning.

In toll collection systems, government agencies in different countries have defined different metrics to calculate vehicle categories.

FHWA defines a set of thirteen categories based on the number of axles (FEDERAL HIGHWAY ADMINISTRATION, 2013). This set of categories is one of the most used among researchers in the area, but it is not the only one.

Nordic countries have developed the Nordic System for Intelligent Vehicle Classification project (NorSIKT) (VAA; MELÉN, *et al.*, 2012). This system defines four classification levels, from general to the most detailed. The most detailed level defines fourteen different categories based on the vehicle's function: ride, passenger and cargo transport, and the vehicle's weight.

In Portugal, the criteria are set in base XIV, Decree-Law 294/97, 24 October. According to the Decree-Law, there are two criteria for defining this classification:

- Vehicle height measured at first axle vertical position
- A total number of vehicle axles.

In Brazil, the different federal and state authorities responsible for highways define the categories used in their respective jurisdictions. They all have in common the following criteria:

- Number of axles on the floor
- Number of suspended axles
- Total number of vehicle axles (rolling and suspended)
- Axle with double wheels.

We use the vehicle classification established by the Transport Agency of the State of São Paulo (ARTESP). The classification is made up of 12 different categories. We can see the categories in Table 1.

Each category is defined by the number of total axles, including suspended axles, and the presence or absence of axles with double wheels. Each of the different Brazilian states agencies adopts its classification, but all are based on the number of axles and double wheels. Please, note that even the table considers category 9, motorcycle, in this research, we are not to consider category 9. This is because motorcycles do not pass throw the toll lanes we will use. Also, we remark that other Brazilian transport agencies have a similar table with minimal changes. Another critical point is that some agencies do not consider the suspended axles when defining the categories. This work considers all axles, ground, and suspended.

Table 1. ARTESP vehicles classification categories

CAT	Number of Axles	Double Wheel	Example
1	2	No	Car- Pickup
2	2	yes	Truck - Bus
3	3	yes	Truck - Bus
4	4	yes	Truck
5	5	Yes	Truck
60	6	Yes	Truck
7	3	No	Car with single axle trailer
8	4	No	Car with dual axle trailer
9	2	Yes	motorcycle
61	7	Yes	Truck
62	8	Yes	Truck
63	9	Yes	Truck

Source: ARTESP (2015)

2.2 MARKET TECHNIQUES

In the example shown in Fig 1, it is possible to see the region of detection and classification of vehicles in a toll lane located in São Paulo State. With some variants, which depend on the solution adopted and the year of installation, the detection system's operation principle is composed of the same sensors. In Fig 2, we show a simplified diagram of the operation of sensors.

The most intuitive way of counting axles with piezoelectric sensors is to place a single sensor transversely to the vehicle roll direction and count the sensor's number of drives. The main drawback is that, on toll roads, vehicles often go backward, making it necessary to identify this situation. For this reason, pairs of sensors are used. Observing the sensors in pairs separated by a distance small enough to be activated simultaneously by a tire, it is possible to identify the axles and direction, as shown in sequences A and B of Fig 2.

The double-wheel sensors are placed on a 45-degree slope and with a distance between them so that a double-wheel axle can drive both sensors simultaneously, but a single-wheel axle cannot, as shown in sequences C and D of Fig 2, respectively.

Some trucks can lift axles when they are lightly loaded. These axles are suspended or lifted axles, and the piezoelectric sensors cannot count them. Therefore, toll systems have adopted optical sensors placed in the same arrangement as the axle sensor but with a specific height, as shown in Fig 1. The operating logic is the same as the piezoelectric sensor.

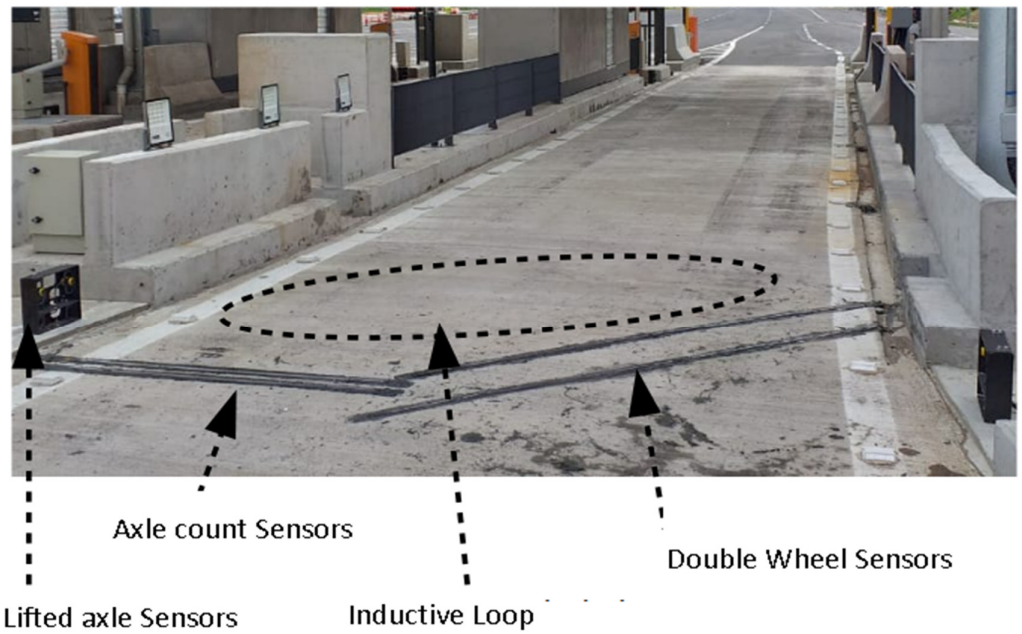


Fig 1.Vehicle detection and classification sensors

2.3 OPERATION OF THE OPTICAL BARRIER AND IMAGING

The optical barrier consists of two columns, approximately 1.80 m high, installed transversely to the vehicles' lane. LEDs arranged at regular intervals emit parallel light beams from one column to another. If there is no obstruction, the light beams are captured by optical detectors installed in the other column on the opposite side. When a vehicle passes between the two columns that make up the optical barrier, the beams of light included in the region of the body and wheels will be blocked. However, the beams that are outside these regions reach the detectors. Thus, a binary image of the

vehicle side view profile can be constructed by sampling the light beams with enough frequency, as shown in Fig 3.

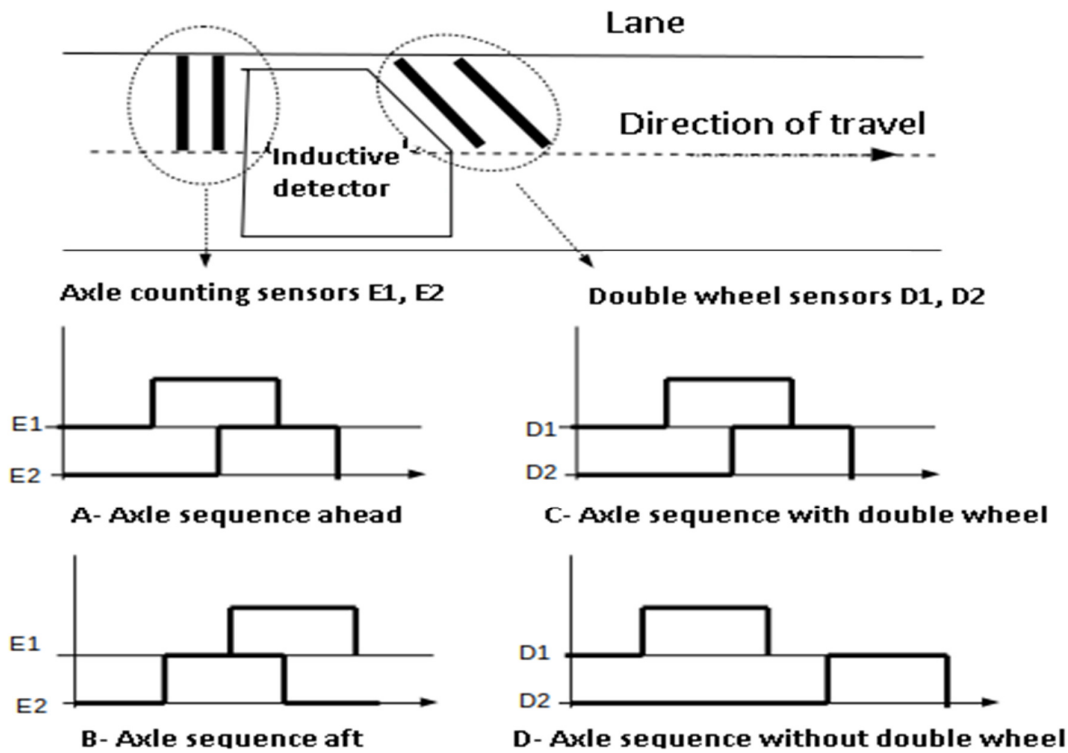


Fig 2.Simplified Layout of axle counting, double wheels, and direction travel

.

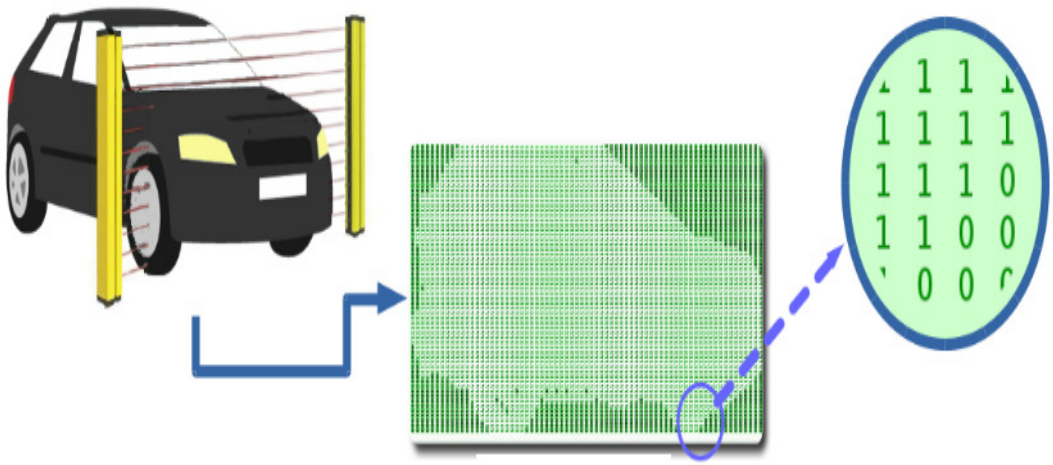


Fig 3.Imaging by using an optical barrier

Fig 4 shows some side view profiles obtained with the proposed method.

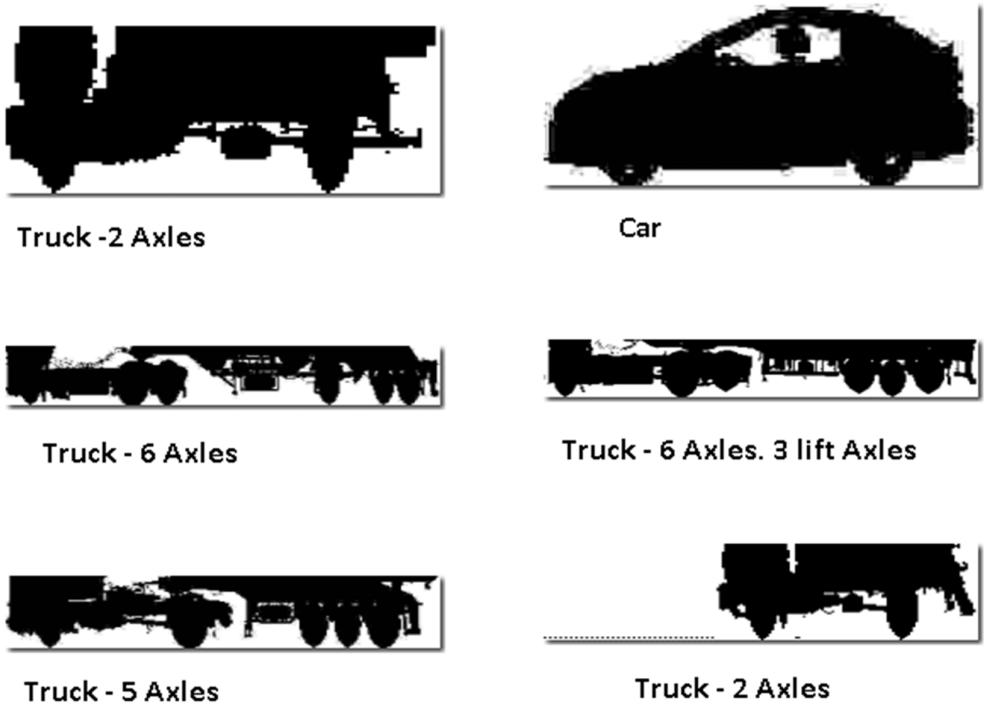


Fig 4.Examples of vehicle profiles made with an optical barrier

2.4 CONVOLUTIONAL NEURAL NETWORKS

A CNN is a feed-forward artificial neural network type. It has its origins in the 70s but gained visibility with LeCun's publication in 1998 (LECUN, 1998). In Fig 5, we present the architecture of a CNN. The network is organized into layers. Each type of layer has a specific function. The layer types are Input, Convolutional Layer, Pooling, and Fully connected.

- **Input Layer:** It is the layer that receives the input image data. The input consists of a certain number of features extracted from the image. In the case of CNN, the entry is the value of each pixel. That is a three-dimensional matrix with height and width, and depth. It increases the number of parameters but simplifies processing. The structure of the network is fixed. Therefore, if the input images are of variable size, they must be adapted.
- **Convolutional Layer:** This layer has the function of extracting different image characteristics. It is composed of simple matrices that act as a filter. The convolution operation used to apply these filters gives the layer its name.

- **Pooling layer:** This layer receives each output from the convolutional layer characteristics map and condenses the information. A typical procedure for pooling is known as Max-Pooling, which consists of taking the maximum value of a given region.
- **Fully connected Layer:** This layer connects all neurons in the last pooling layer to each N category that the network can classify. Depending on the network, it is possible to have several layers until reaching the final layer.

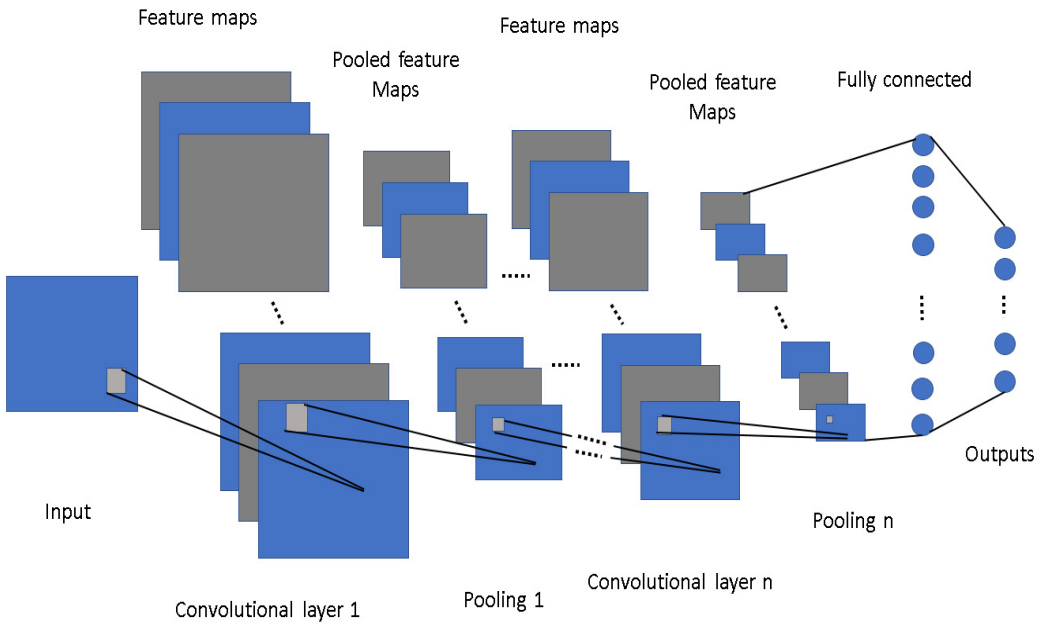


Fig 5. CNN architecture diagram

Some papers divide the network architecture into feature extraction and classification. The characteristics extraction ranges from the network entrance to the last layer of pooling. Classification begins at the first Fully Connected layer and extends to the exit.

2.4.1 AlexNet

AlexNet is a CNN proposed by Krizhevsky et al. (KRIZHEVSKY, 2012) . The authors won the Imagenet Large Scale Visual Recognition Challenge (ILSVRC) competition held in 2012. The results obtained aroused researchers' and companies' interest that gave new impetus to convolutional neural networks. Although newer

networks have improved, it is still widely used due to their simplicity and a good relationship between speed and performance. We can see the network architecture in Table 2. It is eight layers deep, with five convolution layers, and the last three are fully connected.

Table 2.AlexNet Layers

Layer	Description
1	Image Input 227x227x3
2	Convolution 55x55x96
3	ReLU 55x55x96
4	Cross Channel Normalization 55x55x96
5	Max Pooling 27x27x96
6	Grouped Convolution 27x27x256
7	ReLU 27x27x256
8	Cross Channel Normalization 27x27x256
9	Max Pooling 13x13x256
10	Convolution 13x13x384
11	ReLU 13x13x384
12	Grouped Convolution 13c13x384
13	ReLU 13x13x256
14	Grouped Convolution 13x13x256
15	ReLU 13x13x256
16	Max Pooling 6x6x256
17	Fully Connected 1x1x4096
18	ReLU
19	50% dropout
20	Fully Connected 4096
21	ReLU
22	50% dropout
23	Fully Connected 1000
24	Softmax
25	Cross entropy with 1000 classes

Source: Online Matlab help, 2020

2.4.2 Transfer Learning

The AlexNet network was designed to classify 1000 categories of images. It has been trained with millions of examples. In new classification problems, it is possible to use the Transfer Learning technique with different categories to take advantage of this

effort. It replaces the last layers of the network responsible for the classification with new layers adapted to the new set of categories to be classified. The new network needs to be retrained, but a few examples will be enough to obtain good results. The part of the network responsible for extracting characteristics from the images has been conserved. Shin et al. (SHIN, 2016) discuss the technique and assess whether it is better to train from scratch or use Transfer Learning.

2.4.3 Data Augmentation

When training a network, the number of training images of a specific category is often unavailable. Therefore, it is possible to increase the training set by making small changes to the available images. These possible transformations are rotation, translation, mirroring, random cuts, and different noise types. This technique of artificially increasing the data set is known as Data Augmentation. Perez and Wang discuss the technique's efficiency (PEREZ, 2017).

2.4.4 Training, Epoch, and mini-batch

The process of training a neural network consists of adjusting the weights of the neurons that compose it so that, given an input, the network calculates the desired output. If the input image is a two-axle passenger vehicle in our specific vehicle classification case, the network output must indicate category 1. One of the most used training algorithms is the stochastic gradient descent with momentum (SGDM), but there is a wide variety. Dogo et al. (DOGO, 2018) discuss the most used algorithms.

In the training process, sets of previously classified images are used. The network processes a known image. The output estimated by the network is compared to the actual output. This comparison is quantified in an indicator used to correct neurons' weights. The complete circuit of processing a known image, comparing the estimated output with the real one, and updating neurons' weights is known as an iteration. For each image of the training set, we will have an iteration. The iteration in the complete training data set defines an Epoch. Usually, more than one Epoch is needed. The number of epochs we use is a parameter to be defined. If we use only a few epochs, we may not achieve the best result. On the other hand, if we use many epochs, we risk

learning the training data very well but failing to classify new images accurately. This phenomenon is known as over-fitting.

When we work with many images, a resource problem is created. We will need a lot of memory. To solve this problem, the mini-batch concept arose. Instead of using a single image for each iteration, we use images to build the indicator and return by adjusting the weights. However, we have to be careful, and we cannot use very large mini-batches. We take the risk of not adjusting the weights as often as necessary. There is no consensus in the literature regarding the best mini-batch size. Master and Luschi (MASTERS; LUSCHI, 2018) defend that we will have better results using smaller mini-batches than other studies.

2.5 LITERATURE REVIEW, RELATED WORKS

To identify and analyze the current state of the art in vehicle classification, we carried out a Systematic Literature Review (SLR). Different metrics and characteristics are adopted to classify the vehicles depending on the purpose. Some studies estimate volume, width, height, or other characteristics, such as brand and color. We are interested in the classification system used in tolls in Brazil and other countries, which considers the number of axles, the number of suspended axles, and the presence of double wheels. We seek to identify the solutions in the scientific and academic community for similar problems. The results and conclusions are intended to help identify existing research gaps to suggest future research areas considering the mentioned context.

A typical solution involves some hardware or sensor that produces input to a software system. Depending on the nature of the data collected by the sensors, a different software solution is derived. Previous works summarize many of the approaches commonly used, but their focus is on the hardware point of view. As an example, we can refer to Klein (KLEIN; MILLS; GIBSON, 2006), Leduc (LEDUC, 2008), and Mimbela (MIMBELA; KLEIN, 2007). This review also considers the software component a fundamental part of the solution. Another difference is that those works cover a wide area of Traffic Monitoring Systems. We focus on a more specific sub-area of vehicle classification.

The following section describes the SRL's protocols, and then we present the results and a discussion.

2.5.1 Description of the SRL protocol

We carried out a systematic review of the literature, following the guidelines of Kitchenham (KITCHENHAM, 2004), to analyze the primary studies that address vehicle classification, primarily, but not exclusively, those studies based on the classification by axle count. The research questions addressed by this study are:

RQ. What is currently known about vehicle classification? More specifically, this study focuses on the following issues:

RQ1. What strategies or practices are commonly used to determine a vehicle category based on axles?

RQ2.What strategies or practices are commonly used to determine the number of axles on a vehicle?

RQ3.What strategies or practices are used to determine the number of axles suspended from a vehicle?

RQ4.What strategies or practices are used to determine if a vehicle has a double wheel?

RQ5. What strategies or practices are used to evaluate the results?

RQ5.1 Which data sets are used for evaluation?

RQ5.2 Are the data sets publicly available?

RQ5.3 What are the sizes of the data sets?

We searched for articles available on the web in the following databases:

Web of Science (https://webofknowledge.com/)

Google Scholar, ((http://scholar.google.com)

CAPES Journal Portal (www.periodicos.capes.gov.br)

IEEE Explore (ieeexplore.ieee.org)

Science Direct. (www.sciencedirect.com)

The following keywords were used: axle, count; vehicle; classify.

Articles published in the last ten years were searched. Articles that were cited at least once were selected. Unfortunately, many primary studies cited articles that did not fit the established keywords and filters.

2.5.2 SRL Results

Before doing a fine filter, we found 89 potential papers. In Fig 6, we show the frequency histogram by year and where we found them. Of course, this information is irrelevant, but it can indicate the topic's relevance.

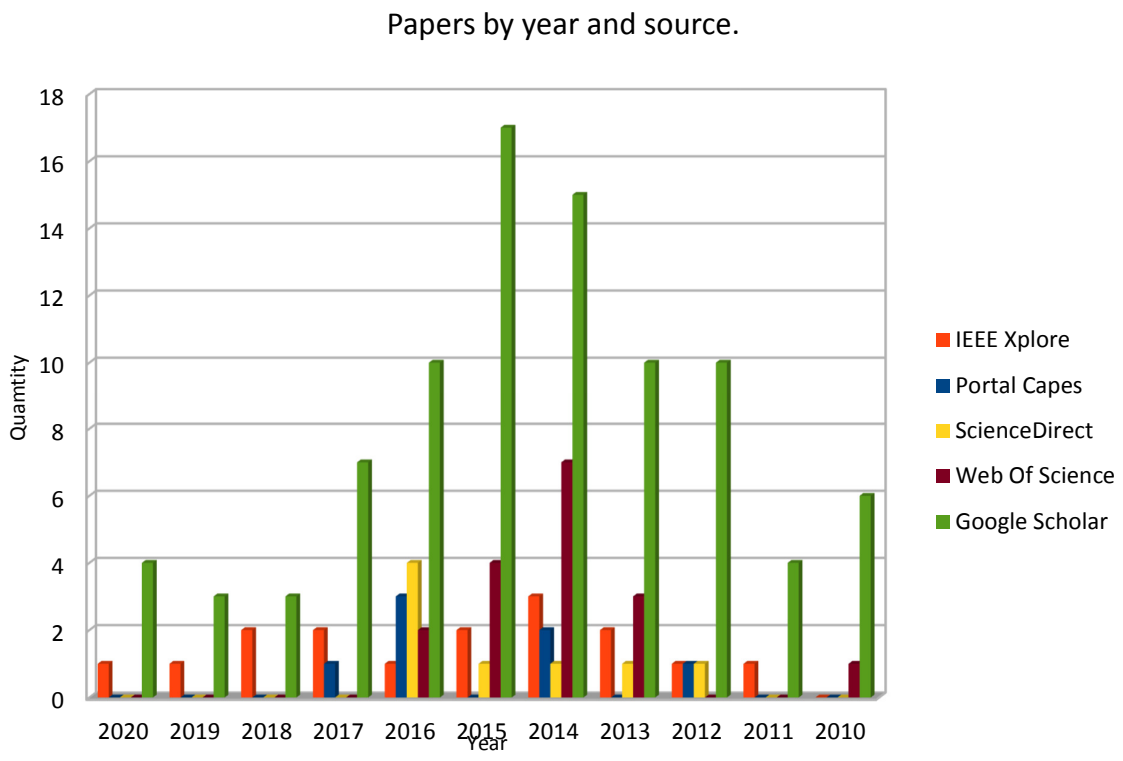


Fig 6.Quantity of papers by year and by source.

After filtering according to the defined criteria remains 44 papers. To better analyze and compare the primary studies, they were grouped according to the hardware technology used. The different groups are listed in Table 3

.

Table 3. Primary studies groups

Group	Description
1	Primary studies based on images/video
2	Preliminary studies based on lasers and optical sensors.
3	Primary studies based on magnetic and inductive sensors.
4	Primary studies based on pressure sensors/ Piezo /Pneumatic
5	Other Technologies

The summary of primary studies found is listed in the following five tables. Each table corresponds to one primary study group. The data summarized here attempt to give answers to the protocol questions.

Table 4. Primary studies Group 1

Primary study	Floor Axle	Lift Axle	Double wheel	Motor-cycle	Results
S1. (YAO, 2016)	✓			✓	65% over 1500 vehicles, in FHWA categories
S3 (Khanipov, 2015)	✓				99% in 24hs test, in 3 categories
S4 (Grigoryev,2015)	✓	P			98% over 1400 vehicles, in 4 categories.
S5 (Grigoryev, 2011)	✓	P			Four categories. No results discussed
S6 (KUMAR; GOYAL; KUMAR, 2014)	✓				Four categories. No results discussed
S7 (PENG; YAN, *et al.*, 2014)					94% over 1000 vehicles, in 4 categories
S8 (PENG; YAN, *et al.*, 2014)					98% in 4 categories.
S9 (WEN; SHAO, *et al.*, 2015)					94% of detections (No classification)
S10 (LIU; WANG, 2014)					80% over 80 vehicles in 4 categories
S26 (LI; YU, *et al.*, 2014)					92% over 2900 vehicles in 2 categories
S37 (NEZAFAT; SALAHSHOUR; CETIN, 2018)	✓				96.5% over 240 vehicles, in 2 categories
S40 (GOTHANKAR; KAMBHAMETTU; MOSER, 2019)	✓				89.4% over 2500 vehicles in 5 categories
S41 (SASONGKO; FANANY, 2019)	✓			✓	99% over 2300 vehicles in 5 categories
S44 (LIU; HUYNH, *et al.*, 2020)					98% over 21400 vehicles in 4 categories
S45 (ZHANG; ZHANG, 2020)	✓	✓			97% over 100 vehicles in 3 categories

Table 5. Primary studies Group 2

Primary study	Floor Axle	Lift Axle	Double wheel	Motor-cycle	Results
S27 (UEDA; TAKEMURA, *et al.*, 1997)	✓				100% over 100 vehicles. Axle count. Do not Classify
S28 (MOGELMOSE; MOESLUND, 2016)	✓		✓		84% counting axles. Do not Classify
S29 (LEE; COIFMAN, 2015)				✓	98% over 21000 vehicles, in 4 and 6 sub-categories of FHWA
S30 (LEE; COIFMAN, 2012)				✓	99.5% of non-occluded vehicles in 6 categories.
S31 (KHATRI; SOMANI, 2015)	✓				Three categories Result not quantitative.
S32 (DA COSTA FILHO; DE BRITO FILHO, *et al.*, 2009)					Shape-based classification. Not quantitative results.
S42 (NEZAFAT; SAHIN; CETIN, 2019)					97% over 4500 vehicles, in 4 categories.
S43 (WU; XU, *et al.*, 2019)	✓				91.98% over 1095 vehicles, in 10 categories.

Source: Joaquin Barreyro, 2020

Table 6. Primary studies Group 3

Primary study	Floor Axle	Lift Axle	Double wheel	Motor-cycle	Results
S17 (TAFISH; BALID; REFAI, 2016)					87% over 2000 vehicles, in 3 categories.
S18 (KLEYKO; HOSTETTLER, *et al.*, 2016)					76% over 3000 vehicles in 4 categories
S20 (TAGHVAEEYAN; RAJAMANI, 2013)					Four categories. No quantitative results.
S21 (CICARELLI; RIBEIRO; LEANDRO, 2011)				✓	90% over 2000 vehicles, in 8 categories.
S22 (YANG; LEI, 2014)					96% over 2000 vehicles, in 5 categories
S24 (DE ANGELIS; DE ANGELIS, *et al.*, 2016)					98% over 50 vehicles, in 10 categories.
S25 (HE; DU, *et al.*, 2015)					92.5% over 125 vehicles, in 3 categories.
S38 (DONG; WANG, *et al.*, 2018)	✓				80.5% over 4507 vehicles in 4 categories.
S47 (GONZÁLEZ; JIMÉNEZ; DE FRUTOS, 2020)	✓				90% over 457 vehicles, in 10 categories

Source: Joaquin Barreyro, 2020

Table 7.Primary studies Group 4

Primary study	Floor Axle	Lift Axle	Double wheel	Motor-cycle	Results
S12 (KIM; COIFMAN, 2013)	✓		P	✓	97% over 18000 vehicles, in FHWA categories.
S19 (RAJAB; REFAI, 2014)	✓		P	✓	98.9% over 1000 vehicles in FHWA categories
S33 (MCGOWEN; SANDERSON, 2011)	✓			✓	93% in intervals off 15 minutes in FHWA categories

Source: Joaquin Barreyro, 2020

Table 8.Primary studies Group 5

Primary study	Floor Axle	Lift Axle	Double wheel	Motor-cycle	Results
S2 (ALEXANDRE; CUADRA, *et al.*, 2015)				✓	93.4% over 475 vehicles in 3 categories
S11 (YU; PREVEDOUROS; SULIJOADIKUSUMO, 2010)	✓			✓	There are not quantitative results. FHWA categories
S13 (HASSELBRING, 1998)					There are not quantitative results. Four categories
S15 (MA; XING, *et al.*, 2013)	✓			✓	99% over 600 vehicles in FHWA categories
S16 (BAJWA; RAJAGOPAL, *et al.*, 2011)	✓			✓	100% over 53 vehicles, in FHWA categories
S23 (IWASAKI; MISUMI; NAKAMIYA, 2013)					Only detection, Not class. 92.8% over 1500 vehicles.
S35 (HUANG; LU; TOLLIVER, 2018)	✓				94% over 358 vehicles, in 3 categories
S36 (ZHAO; WU, *et al.*, 2018)	✓				80.36% over 415 vehicles in 10 categories.
S39 (AKULA; GHOSH, *et al.*, 2018)	✓				93%% in 4 categories

Source: Joaquin Barreyro, 2020

We present a qualitative and comparative analysis of the studies found in the following. It is common to refer to each study using the identification in the first column of Primary Studies tables listed above. Therefore, if the reader wants to know the complete reference, it must go back to the table.

2.5.3 Primary studies based on images

The number of studies based on images in recent years is significantly higher than other approaches. It is possible to group these primary studies into three main groups defined by the camera's position:

a) side camera close to the lane (the vehicle never appears entirely in a single frame);

b) the camera is located above the road;

c) the camera is located on the side, with the vehicle appearing in a single frame.

The first group uses a camera looking sideways at the vehicle's passage and uses object detection and tracking to identify and track each wheel. For each frame, the detector output is the coordinated position of each wheel in the image. We can see this approach in S3, S4, and S5. They use the Viola-Jones method (VIOLA; JONES, 2004) (MINKINA; NIKOLAEV, *et al.*, 2014) as an object detector. We found a different solution shown in S6. The authors use a Canny edge detector followed by Fast Hough Transform

(FHT). This wheel-per-frame detection serves as input to another algorithm. Considering the information of all the frames together, it must determine the total number of axles. To do this, some authors generate a new image representing the X coordinate on one axis and the time on the other. Counting the axles becomes a problem of counting how many continuous lines are in the image. S3, S4, and S5 are examples of the described approach. In S5, the authors use an FHT with pre-smoothing to count the number of output lines. The method is robust and precise, but it has a limitation: it applies only to almost constant movement. To solve the non-uniform movement in S3 and S4, the authors propose an algorithm based on the analysis of connected components.

A different approach is used in S26, S7, S8, S9 S10. In these studies, the camera is located directly above the target road. It is impossible to count axles directly with this type of image. It is necessary to extract some key characteristics that help infer the different classes of vehicles. In S26, the authors classify vehicles only in two categories. Small and large. They developed an adaptive background subtraction based on a color image that allows tracking vehicles to determine speed and size. In S7 and S8, the authors determine four categories: sedan, taxi, van, truck. In S7, the spatial pyramid correspondence method is used with sparse coding to extract characteristics from the images and classify them. In S8, the images are partitioned into several fragments, and a hash function is calculated for each one. The categories are found using the hash calculated to input a Support Vector Machine (SVM).

In S10, the authors present a vehicle classification system based on a dynamic Bayesian network (DBN). Three main features are employed in this system: the vehicle's geometric feature, the location and shape of the license plate, and the vehicle's position. First, a vehicle detection method and a tracking method are used to locate the vehicle. Then the characteristics of the video sequences are extracted. a Gaussian Mixture Model (GMM) is used to construct the probability distribution of the characteristics. In S9, Haar-type features are extracted to represent vehicles.

In S1, a different camera position was adopted. A side camera was used, with an open view. The main difference with the first approach is that it is possible to observe the complete vehicle in a frame. For background separation, an average filter was used. It calculates the average image of a finite number of frames and then subtracts it from the actual image. It is also necessary to apply morphological operations such as

opening, closing, erosion, and dilation. Detection and segmentation of the wheels are made based on pixels' color using a fuzzy c-mean algorithm. (FCM).

In S37, the camera still captures the side view, but the focus is not on the axle count. They use a ResNet CNN network to identify two different types of trucks.

In S40, with side view pictures where the vehicle appears complete in the image, the authors use a CNN network to identify the vehicle. They count the vehicle's wheels with two different approaches: one based on CNN and the other using Hough's circular transform. Finally, they propose a decision algorithm that unifies the two criteria.

In S41, the authors use an image with the vehicle side view already segmented and use these images as input to a ResNet CNN to classify vehicles into five categories based on the number of axles.

In S44, the proposed method employs existing road monitoring cameras. The authors use CNN to locate the vehicles in the image. They then classify them using a subgroup of FHWA categories (4 categories) based on the vehicle's size.

In S45, the authors continue with the same type of images as before. To identify the vehicles, they use Mask R-CNN. To identify the axles, they compare three methods: a Canny-based ellipse detector, another detector based on Template Matching, and a method based on Mask R-CNN. The best results were obtained with the last one.

2.5.4 Primary studies based on lasers and optical sensors

It is possible to identify two groups of devices, those based on laser and based on led. The concept is the same: it consists of a curtain defined by several infrared light beams. As these beams are interrupted, a vehicle profile is formed. Depending on the position, spatial resolution, and temporal resolution, it is possible to count, determine the speed, and classify them.

In S32, the authors use a modulated infrared sensor at the top of the lane. The light reflected by the cars is captured. Using this data, a time series signal is constructed for each vehicle. A comparison of these signals is made using an algorithm called Dynamic Time Warping, and the category of each vehicle is determined. Axles are not counted. In the experiment done, only two different categories are used.

In S31, the authors use a vertical optical barrier to extract profiles from passing vehicles. A set of horizontal optical sensors determines the number of axles and a

magnetic loop to determine vehicle arrivals. To identify a vehicle class, they use the number of axles and then compare them with a set of profiles stored in a database. Few categories are used.

In S29, the authors use an adaptation of the S30 to develop a portable non-intrusive vehicle classification system to validate the performance of Vehicle Classification Stations. In S42, the approach is similar. The authors use profile images made with a LIDAR. They explore the possibility of classifying this type of image with CNNs. They use three different types of CNNs (AlexNet, VggNet, and ResNet.) and compare results.

In S43, the authors also use LIDAR to capture a profile. They prefer to obtain six signal characteristics: maximum length, maximum and minimum heights. They compare four classification methods: A Bayesian network, K-NN, Random Forest, and SVM. The best result is obtained with Random Forest, 91.98%.

In S28, a system based on LIDAR is proposed to determine the width of the wheels. This information can be used to determine tires mounted (double wheel). The system can also determine the number of axles. The authors conclude that the system works as a proof of concept but that the results do not reach the expected level. This article was the only one we found which concentrates efforts on detecting the double wheel. In Brazilian classification systems, double-wheel drive determines whether a vehicle is commercial or not.

2.5.5 Primary studies based on magnetic and inductive sensors

Electromagnetic sensors, especially those based on inductive loops, are the most common sensors used in traffic management applications. There are many configurations and variations, but the most popular sensors are inductive loop detectors and magnetic sensors. The operating principles are described in (MIMBELA; KLEIN, 2007).

In S20, there is a system of portable magnetic sensors installed beside the road. This system can count vehicles, classify, and measure speed. The speed is estimated using two sensors and the correlation between the data provided in each one. The classification is made by the magnetic signal's length and the magnetic signal's height. They define four categories but do not count axles directly.

In S22, the authors use a sensor like that used in S20, but they choose different features. Five signal features are extracted, including the signal duration, the signal energy, the average signal energy, the positive and negative energy ratio of the X-axis signal, and the Y-axis signal's positive and negative energy relationship. They perform a simple classification based on a hierarchical tree methodology in five categories.

In S25, the authors use similar hardware to those used in S22 and S20. The main difference is that they estimate 17 parameters. They developed an intelligent selection of these parameters based on their relevance in each category for use in an SVM. They fall into three different categories.

In S17, the proposed classification system uses the Discrete Fourier Transform (DFT) and Principal Component Analysis (PCA) to automatically extract the magnetic signatures' characteristics and classify the vehicles into three categories, determined by an SVM.

In S18, the main contribution is that the authors study the applicability of Data Smashing in vehicle classification. In contrast to most traditional machine learning algorithms, this approach's main advantage is that it does not require the extraction of raw signal parameters.

In S24, the authors use a pair of magnetometers to determine the speed and classify vehicles by analyzing the magnetic signature using a DTW algorithm. This algorithm estimates a degree of similarity between two-time series. Then, they determine the category by comparing the time series recorded from sensors with a database and using a simple decision tree algorithm. They tested ten different vehicles and were 95% correct.

In S21, the authors use an inductive sensor to obtain a vehicle's magnetic profile and obtain parameters such as the number of peaks, magnetic compliance, average peak value, variance, and symmetry. The classification system is based on a Multi-Layer Perceptron (MLP) neural network, and the output parameters serve as input to the network. The authors defined eight categories: Motorbike Car, Bus Truck Trailer, Van Truck Pickup Truck. These categories are not based on the number of axles; it is possible to make an analogy.

In S38, the authors obtain 42 characteristics of the magnetic signal from the sensors. The Gradient Tree Boosting algorithm is used to classify.

2.5.6 Primary studies based on pressure, piezoelectric and pneumatic sensors

Contact sensors such as piezoelectric, pneumatic tubes, and the like are the most used devices to acquire traffic volume data. The moving vehicle wheel's pressure activates it, like a switch, allowing a counter or software to register and process the event. It is common to find these sensors installed in pairs, placed in traffic lanes (SKSZEK, 2001) (MIMBELA; KLEIN, 2007).

Pneumatic tube sensors emit air pressure pulse along a rubber tube when the vehicle's tires pass over the tube. The air pressure pulse closes an air switch, producing a signal transmitted to a counter or analysis software (MIMBELA; KLEIN, 2007).

A piezoelectric is a specially processed material that converts kinetic energy into electrical energy. Piezoelectric materials generate tension when subjected to impacts or mechanical vibrations. Piezoelectric sensors can be used similarly to pneumatic tubes, registering a signal emitted every time a tire passes over it (MIMBELA; KLEIN, 2007).

Suppliers claim accuracy rates of about 99%. However, in S33, the authors assess the accuracy of pneumatic tubes in counting and classifying vehicles. They conclude that errors are close to 4% for daily counts. Compared to checking the manual count for shorter periods, the error is more remarkable, 12 % to 14 %.

In S12, the authors evaluate three classification stations composed of a double electromagnetic loop and a piezoelectric sensor. They do a per vehicle review (PVR). More than 18000 vehicles were analyzed. The three stations' performance was good (97 %), but the performance is 60 % correct on trucks.

In S19, the authors introduce a vehicle classification system that employs a single-element piezoelectric sensor placed diagonally across a traffic lane. Each tire of a passing vehicle produces an output pulse. An average vehicle width (distance between the tires on the same axle) is assumed, and, with that, it is possible to determine the speed and then the distance between the axles.

In S47, the authors use piezoelectric transducers that do not need to be installed inside the floor. They can be placed on top. First, they classify counting axles traditionally, identifying the signal's peak energy corresponding to the axles. Then, they use a pair of sensors separated by a known distance to determine the wheelbase's speed.

2.5.7 Other Technologies

Other sensors can be used in vehicle classification. In S2, the authors propose using a system based on acoustic data to solve automatic vehicle classification. They select some sound parameters to correctly classify the vehicles using a hybrid genetic algorithm and an Extreme type neural network (Due to the excellent generalization performance and fast learning speed). It classifies only in a few categories: motorcycle, car, and truck. It does not seem to be able to handle many categories.

In S11, there is an interesting comparison between three different systems. The first is based on infrared sensors and classification based on axle count; the second is based on video (and classification based on length), and the third is based on radar (classification based on length). The authors observed many video and radar systems problems in congested periods and occluded vehicles. They conclude that only the infrared-based system behaves satisfactorily (more than 91% accuracy).

In S23, the authors use a thermal camera. The proposed algorithm uses the thermal energy reflection of the tires. As the purpose of the article is the detection and not classification, they used frontal images. It is possible to use lateral images to count tires (axles) and determine categories based on the number of axles. In S39, the same type of sensor is used to obtain images of the vehicle's side. Some characteristics of these images are extracted to be classified in categories based on axles.

In S16, the authors use an array of accelerometers and magnetometers to count the axles and determine the spacing. The magnetometer determines the detection of a vehicle. The accelerometer signal is processed with an algorithm in the time domain (filter, peak detection, and peak selection). Each peak is an axle, and, with speed (calculated using magnetometers, but not discussed in the article), it is possible to determine the distance between the axles. In S15, the authors use an approach like the one used on S16, but with sensors located in different locations.

In S35, the authors use sensors formed by a Fiber Bragg Grating (FBG) grid wrapped in polymer and reinforced with glass fiber. The passing of the vehicles' axles produces a voltage signal that is recorded. They use an SVM algorithm to classify vehicles.

In S36, the authors use optical sensors on the pavement to record the passing axles' vibrations. They register the signal and extract some characteristics using an empirical method.

2.5.8 Analysis

It can be deduced from the bibliography that methods using images have increased considerably. It is also possible to observe the increasing number of methods that use some type of CNN. The success rate achieved by the different authors has an extensive range. It is difficult to compare studies by the accuracy index without considering the number and type of categories that the authors propose. It would not be logical to say that the 99% accuracy rate in a study employing only three categories is better than an 80% accuracy rate in another study that classified vehicles in eleven different categories. In general, all studies are concerned with counting the axles, either directly or indirectly. In our research, we are interested in suspended axles in addition to the axles. Papers' quantity focused on suspended axles is smaller, and some treat the suspended axles as false positives. Even fewer studies are concerned with the double wheel. Only the primary study S28 dealt with the problem in detail, using a LIDAR. A mapping could be made between the categories of the FWHA and the categories based on double wheels and axles. In this way, those works that use all the categories of FWHA could be considered apt to classify using double wheels systems.

Another essential point to be analyzed is the location chosen for the sensor. Most studies use a highway with a broad and open view. Few studies used a toll plaza environment. Within the toll lane, the view is restricted. It is not possible, for example, to have a single image of the complete vehicle because there is not enough space between the road and the camera to have a critical distance.

Another question that arises and that is not addressed by the vast majority is: when dealing with open environments, what happens with the different methods in the face of different types of lighting, day, night, sun from the front, sun from the side, shadows, and variations in weather conditions such as rain, snow or fog.

3 METHODOLOGY

This chapter will formulate the hypotheses that support the method and then describe the proposed method.

3.1 PREMISES, HYPOTHESES, AND CONSTRAINTS

We believe it is good to address vehicle classification in toll plazas using vehicle images and a CNN to classify them. The optical barrier is the best sensor to generate these images. Moreover, it solves two fundamental drawbacks: lighting and little space on toll lanes. For this, we need to consider some essential points.

The height of the optical barrier must allow capturing of enough information from the vehicle to classify it unambiguously.

The speed of vehicles traveling on the Toll lane is variable. The maximum speed allowed is 40km / h. However, frequently some drivers exceed this speed. Without the aid of another sensor and being perpendicular to the vehicle, the optical barrier cannot determine vehicle speed. Therefore, we must find the influence of speed variations over the classification.

An optical barrier placed perpendicular to the vehicle cannot directly pick up information regarding the double wheel. One of the hypotheses of this research is a relationship between the vehicle profile and the double wheel. An efficient method must use this relationship to infer the vehicle category.

3.2 METHODOLOGY

In Fig 7, we can see a graphic summary of the proposed method. This figure can be divided into two blocks: The binary image creation from the optical barrier data, illustrated in the upper part of Fig 7, and the classification itself, illustrated in the lower line.

3.2.1 Imaging with optical barriers

In Section 2.3, we present the operation of the optical barrier. It periodically sends the status of each sensor that comprises it. This information can be represented with a

one-dimensional binary vector of N elements with binary values, where N is the number of sensors in that barrier. The initial state has all the optical barrier sensors in logical state 0. When a vehicle arrives, we will have at least one sensor with logic status 1. During the passage of the vehicle, we store M vectors. The number of vectors depends on the relationship between the sample rate and the vehicle speed. Then we proceed to the construction of binary images. Each vector stored in the first stage constitutes a column with N pixels' height. If we combine all the stored M vectors, we will have an NxM image representing the vehicle. We can see the details in Fig 8.

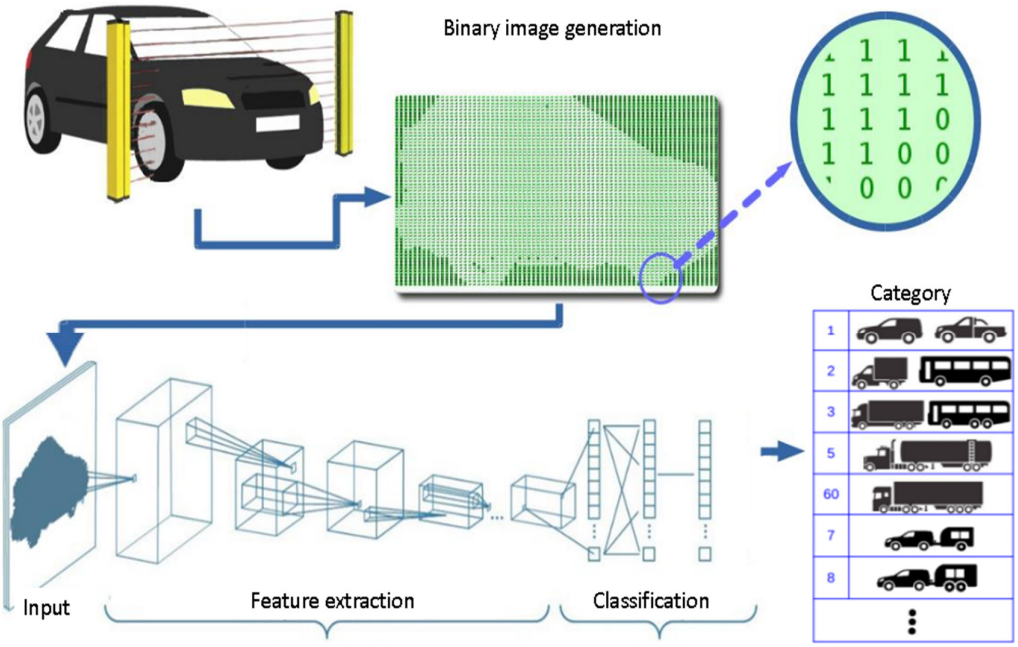


Fig 7.Classification method diagram with binary images and CNN classification

The optical barrier has a height of h. This height is generally less than that of many vehicles traveling along the road, limiting the amount of information. On the other hand, as already mentioned, the barrier is composed of N equidistant sensors. Therefore, the barrier's resolution will be h / N. The barrier periodically sends the N sensors' status. The speed with which the hardware can send this information determines the temporal resolution. This temporal resolution associated with the vehicle's passing speed determines the resulting image's number of columns (width). The vehicle's speed is unknown, and it cannot be assumed that it is constant.

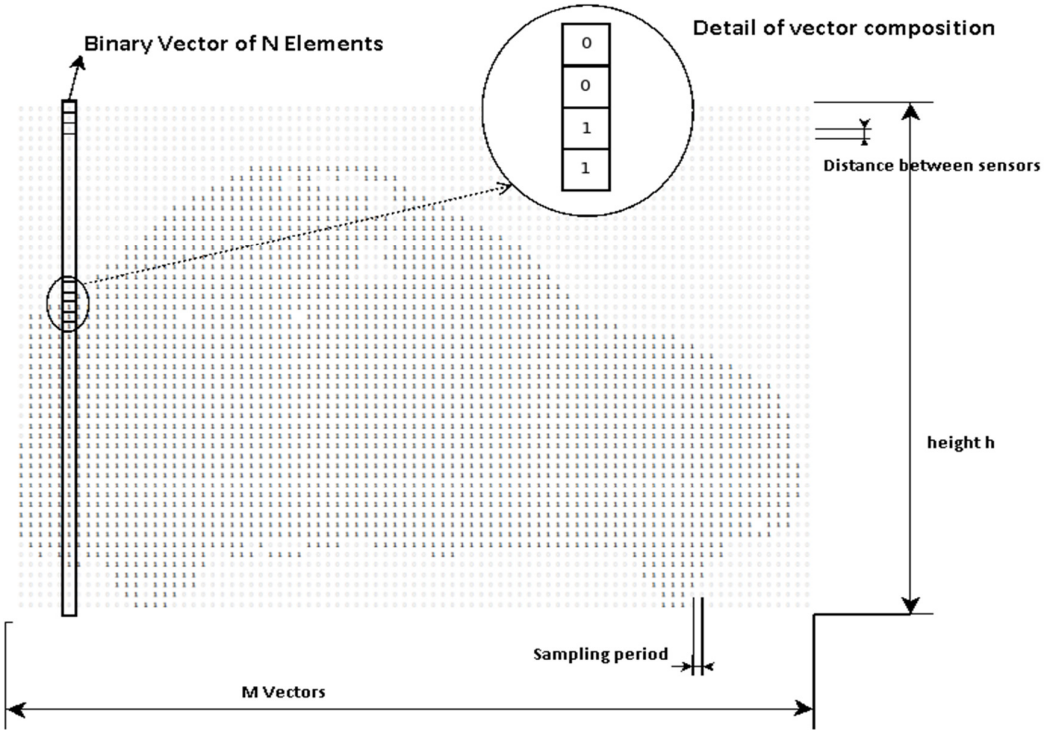


Fig 8.Binary-image components illustration

Considering that the barrier's sampling period is constant and that the vehicle's speed is variable, it is expected that the faster vehicles will generate compressed images in the horizontal direction. Conversely, slower vehicles tend to generate more elongated images. Although in Brazil, passage speed through the automatic lanes is regulated at 40km / h. However, the range of minimum and maximum speeds is typically broader, so working with an estimated average speed does not seem to be the solution. The strategy adopted was to preserve only the consecutive vectors with at least one different value. Thus, if a vehicle's passage generates M vectors, but only L vectors are not repeated consecutively, the resulting image will have NxL dimensions. This way, in uniform regions of the vehicle, image compression occurs. In the wheel region, which is the region of particular interest in this research, the number of vectors with at least one different point will be more significant due to its circular shape. The image will have a more significant amount of information in that region. The representations become realistic when the vehicle's passing speed is uniform. In Fig 9, we can observe the commented situation. In the upper part of the figure, we show the exact vehicle's front at three different speeds. The same three images appear at the bottom of the image with the speed filter applied.

To use the image in the classifier, we need to adapt their size. As previously mentioned, the images have variable width, but the classifier has a fixed number of input parameters. Therefore, we use a simple interpolation by the nearest neighbor to make the image size compatible with CNN to circumvent this difficulty.

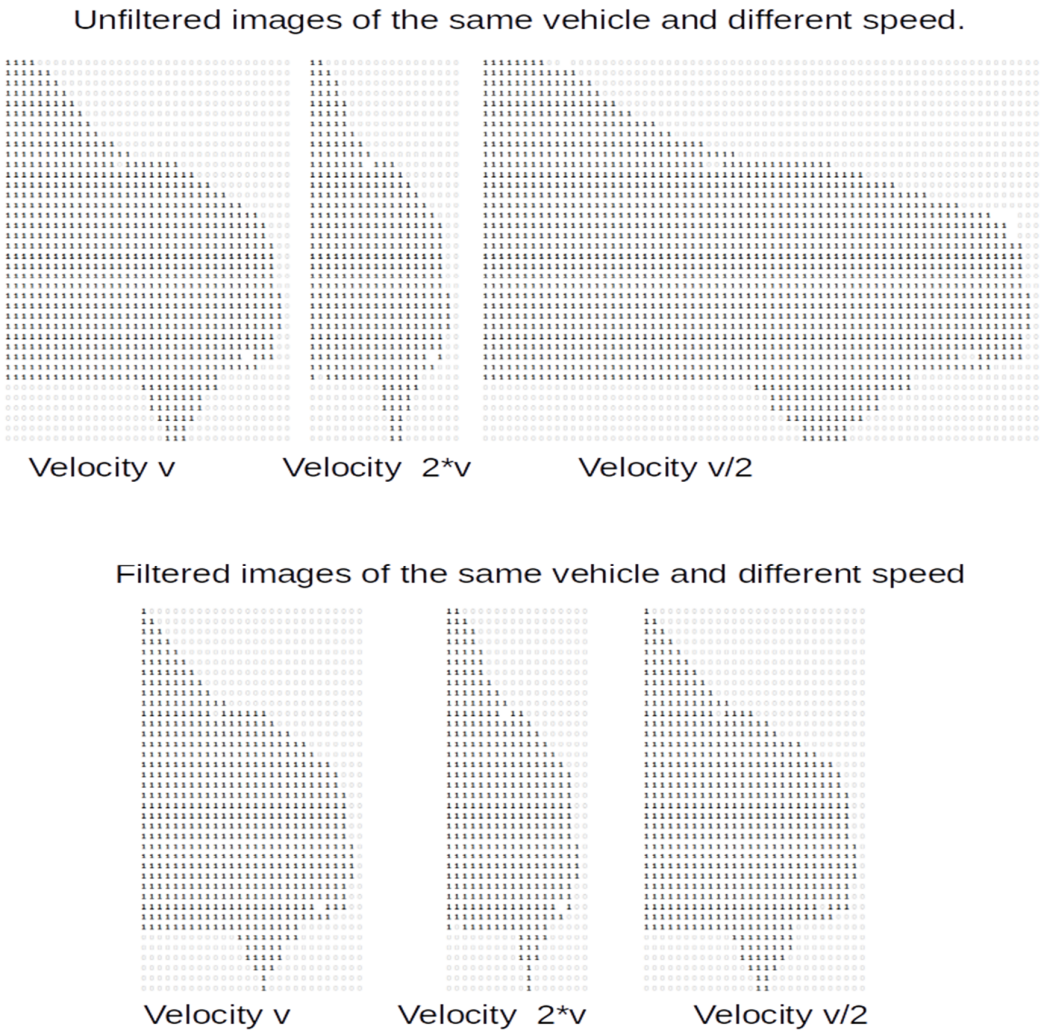


Fig 9.Speed effect illustration on imaging with the Optical Barrier

3.2.2 Classification using CNN

The CNN network that we will use to classify needs to be adapted to the interest categories. For this, it is necessary: choose an architecture; modify the last layers; build a validated data set to train; complement the data set with artificial images in those categories where there are not enough examples; prepare a set of real images for validation; determine the set of hyper-parameters to train; train the network; validate results. If results are not satisfactory, we must change hyperparameters; retrain, and compare the results. The process is illustrated in Fig 10 flowchart.

1. **Choose an architecture**. We start from an existing and trained architecture that saves effort and labor. There are several networks. Most of them trained with millions of images from the ImageNet database (RUSSAKOVSKY; DENG, *et al.*, 2015). They were made to classify objects into 1000 different categories. Choosing an architecture depends on the specific problem. The essential characteristics are accuracy, speed, and size when choosing a network.
2. **Modify the last layers**. To use this classifier in our problem, we need to incorporate the defined categories. It is done by modifying the last layers of the network. We are replacing the generic categories that the original network has with our problem categories.
3. **Build a validated data set for training**. Since the network was modified in the previous step, we must retrain it. Luckily as the network is already pre-trained, we do not need millions of images to train from scratch. However, thousands of images must be enough. These images need to represent the different types of vehicles classified in the best possible way.
4. **Complement the data set**. For some categories, there are not enough examples. Using the Data Augmentation technique, it is possible to complement the data set.
5. **Determine the set of hyper-parameters**. Regardless of the type of architecture chosen, it will be necessary to determine a series of hyper-parameters when training the network. It is necessary to make several iterations by fine-tuning these values to find an optimal result.

6. **Train the neural network**. It is the most time-consuming and computer-intensive process. Bearing in mind that this process will be repeated each time the hyper-parameters are changed, it is good to have a tool, for example, scripts, that automate this training.
7. **Validate results**. Once the neural network is trained, we will validate the network using images reserved exclusively for validation. As the data set was previously classified and verified, the validation process compares the results. For this purpose, we will calculate global and category accuracy rates and correctness.
8. **Compare results**. Evaluate the results against other systems and methods.

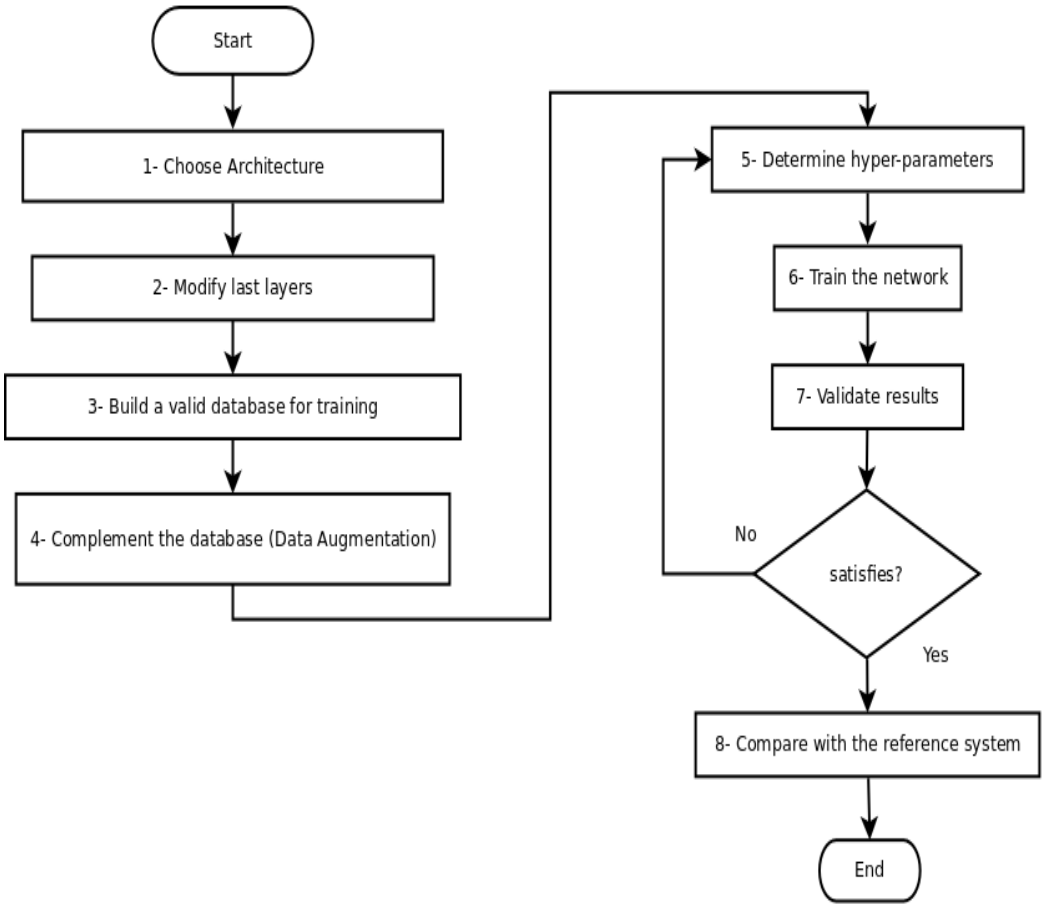


Fig 10.Flowchart of the training method and classification

4 EXPERIMENTS

In order to evaluate the proposed methodology in section 3.2, we performed three experiments. Experiment 1 aimed to evaluate classification performance individually and identify specific problems related exclusively to each category. Experiment 2 evaluates the method's performance in an actual toll collection system environment. In experiment 3, we went back to the first experiment and tried to improve the results by doing minor modifications

. We installed an optical barrier on an automatic toll in the state of São Paulo. The barrier is 180cm high and has 59 equidistant sensors. In Fig 11, we show the installation. The sample rate that the barrier offers is 240Hz. The images generated by the optical barrier have a fixed height of 59 pixels. However, the length is variable because, as stated in Section 3.2.1, it depends on the vehicle's speed and length.

Fig 11.Image of the installation of the optical barrier

The CNN architecture chosen to classify is the AlexNet architecture introduced in section 2.4.1. We chose it because it is one of the fastest pre-trained networks available and uses fewer resources. In the MatLab documentation (MATLAB, 2020) , we can find Fig 12, in which the performance of various types of CNNs is compared. The comparison is only an estimate to serve as a reference, but the actual results depend on the application context. Nevertheless, the graph clearly shows a direct relationship between speed and accuracy. As the results depend a lot on the nature of the application, it is intuitive to start working with faster networks and if the results are not satisfactory, try with slower and more complex networks.

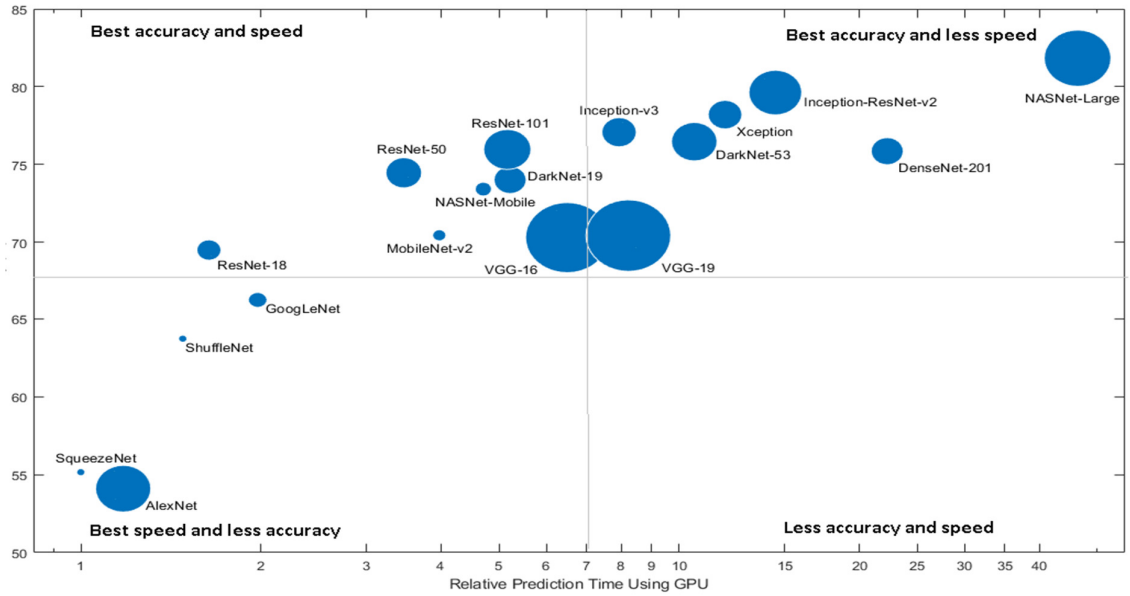


Fig 12.Different CNN architectures. Relative Prediction time vs. Accuracy

Source: Adapted from (MATLAB, 2020)

The AlexNet network was initially trained to classify images in 1000 different categories. In the context of vehicle classification, we want the network to classify into 11 categories defined in Table 1. For this, the last layers were modified. In Table 9, we can see the detail of the replaced layers.

Table 9.Modified layers comparison

Layer Number	Original Layer (Table 2)	Modified Layer
23	Fully Connected 1000	Fully Connected 11
24	Sofmax	Sofmax
25	Crossentropyex with 1000 classes	Crossentropyex with 11 classes(Table 1)

Source: Joaquin Barreyro, 2020

In training the neural network, a set of previously classified data is required. The installed optical barrier was linked to the toll production collection system. Each transit detected by the collection system generates a unique vehicle identifier. We use this same identifier to name the images obtained by the optical barrier. In this way, it is possible to link the transits within the collection system and the category detected by it with the images generated by the barrier. The collection system's automatic classification is based on piezoelectric, loops, and optical sensors described in section 2. In addition to the automatic classification system, the concessionaire has a team dedicated to manually validating those transits that present some anomaly. This validation is done with the help of the photographic record of each transit. All this information is stored in a database. More information on how the concessionary system works can be found in Annex 4.

We made available all the images used in this study in Mendeley's Open Repository (BARREYRO; YOSHIOKA; MARTE, 2021a).

4.1 EXPERIMENT 1

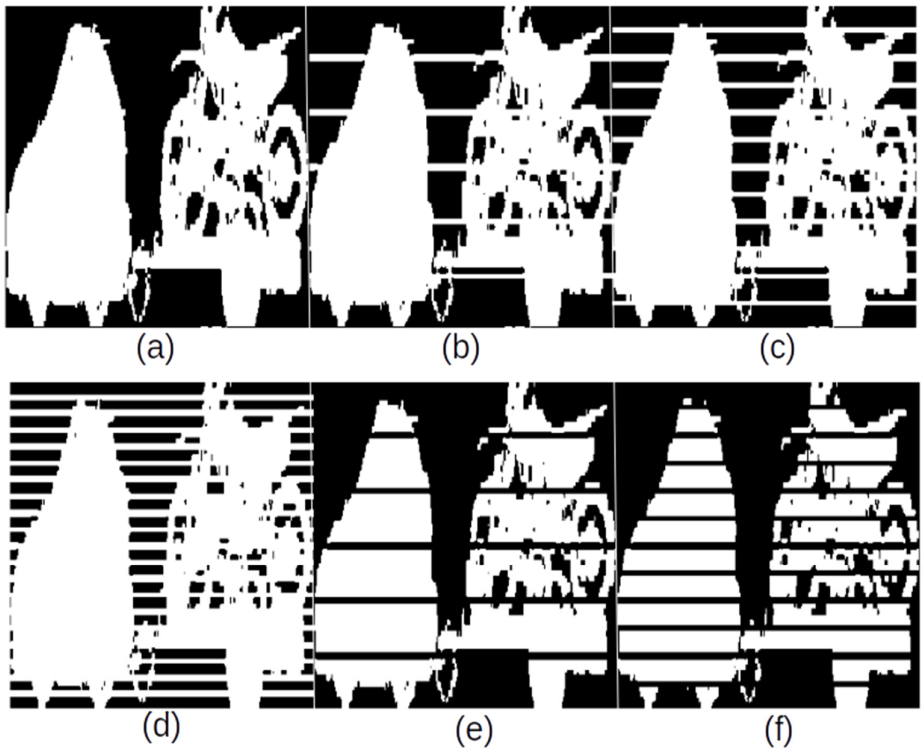
The barrier was in operation from 04/13/2018 to 11/11/2018. We collected 270541 images. When working with neural networks and CNNs, a frequent question is how many images to train the network. The ImageNet database (RUSSAKOVSKY; DENG, *et al.*, 2015) used to train the AlexNet network provides between 732 and 1300 images for each category; therefore, we selected approximately 1000 images from each category of the total images collected following this criterion. Another 1000 different images per category, when available, were reserved for validation.

Some categories do not have a significant amount of vehicle passes. To complement the dataset, we use the Data Augmentation technique introduced in section 2.4.3. The transformations made consist of:

a) Increase the number of columns by 10%, 20%, and 30%;
b) decrease the number of columns by 10% and 20%;
c) vertically shift the image by a certain number of lines;
d) Incorporate noise into the images: The type of simulated noise is characteristic of this sensor. For example, it is common to obstruct one or more sensors that make up the barrier due to some obstacle dirt, resulting in a logical state 1 of the sensor and a corresponding horizontal line throughout the image.

Therefore, we incorporated the noise of 10% and 20% of clogged sensors. It is also frequent when a sensor is defective and always remains in a logical state 0. Therefore, we also incorporate 10% and 20% of defective sensors' noise.

Some examples of the transformations are shown in Fig 13. The criterion for choosing the transformations was to maintain the vehicles' fundamental characteristics concerning the classification, that is, to maintain the number of axles and respect the morphology.



Original image(a) , Obstructed sensor 10%(b) , 20% (c), 30%(d), Sensor off 10%(e), 20%(f)

Fig 13.Examples of images with artificially generated noise

Table 10 shows the set of images used grouped by vehicle category. We differentiate between original images and artificially modified images. To train, we used 11233 images distributed among the 11 categories. The training algorithm chosen was SGDM. In Table 11, we show a list of the hyper-parameters used. Hyper-parameters were chosen based on frequent use in the literature. Once it is verified that the solution converges, we concentrate our efforts on determining two parameters: The mini-batch

size and the number of training epochs. We trained the network with three different mini-batch size values 32, 64, and 128, and we varied the number of epochs from 1 to 50.

Table 10.Number of images used in CNN training and validation

Class	Training			Validation
	Original	Modified	Total	
1	1034	0	1034	1000
2	1000	0	1000	973
3	1000	0	1000	1006
4	1029	0	1029	510
5	613	399	1012	324
60	1001	0	1001	497
7	90	999	1089	45
8	29	998	1027	24
61	208	799	1007	112
62	2	1000	1002	10
63	682	350	1032	828

Source: Joaquin Barreyro, 2020

Table 11.Hyper-parameters used to train the new CNN

Parameter	Value	Parameter	Value
Momentum	0.9000	**L2Regularization**	1.0000e-04
InitialLearnRate	1.0000e-03	**GradientThresholdMethod**	l2norm
LearnRateSchedule	’none’	**GradientThreshold**	Inf
LearnRateDropFactor	0.1000	**MaxEpochs**	150
LearnRateDropPeriod	10	**MiniBatchSize**	32-64-128

Source: Joaquin Barreyro, 2020

4.2 EXPERIMENT 2

We used the same network trained in Experiment 1, but we collected data from three toll stations. We call the stations P1, P2, and P3. We observe the three stations in a continuous space of time. The total number of vehicles classified was 194,361. The global accuracy will depend on the type of vehicles and the environmental conditions.

We collected 92,841 samples from P1 and 75,869 from P2 between 10 February 2021 and 14 February 2021. In addition, we collected 25,651 vehicles from P3 between 11 February 2021 and 14 February 2021. We show summarized data in Table 12.

Table 12.The number of images and period of observation in experiment 2.

Station	Period		Number of vehicles
	Start	End	
P1	10/02/2021 00:00hs	14/02/2021 23:59hs	92841
P2	10/02/2021 00:00hs	14/02/2021 23:59hs	75869
P3	11/02/2021 00:00hs	14/02/2021 23:59hs	25651

Source: Joaquin Barreyro, 2020

This experiment wanted to determine two factors: the rain and the lane type impact classification accuracy. To take a closer look, we focus on a specific toll station, P2, and observe performance on a specific day every hour. We chose two different periods of 12 hours, one with rain and one without rain. We believe that the worst performance will be at the time of the most significant rain. Also, we compare the accuracy between manual lanes and automatic lanes.

4.3 EXPERIMENT 3

In this experiment, we want to know how the quality of the training images influences the result. We will use the same dataset used in experiment 1 and optimize the training dataset. The first factor we are going to observe is the quality of images. We suspect that we could have some images in the wrong categories. The method used to pre-classify the images is imperfect and could have some mistakes. Besides, we are going to observe the artificial images we use carefully. The worst results were between categories with artificial data, and we suspect there is an opportunity to improve. Table 13 shows differences in data from Experiment 1 to experiment 3. We eliminated 731 images, near 6.5% — most of the eliminated images, 672, resulted from the data augmentation. We eliminate them because, in the transformation, the represented vehicle lost an axle and became into another category. Another modification we will introduce in this experiment is to include the length of the vehicle in the CNN. We believe that there is a relationship between the length of the vehicle and the category. Small vehicles tend to be category 1, and larger vehicles tend to be trucks with many axles. However, because of the fixed size of the input of the CNN, we normalized images, and

we lost the length information. A 9 axles truck has the same image size as a category one car. Here, we work with binary images. They could be represented only by two values, 0 and 1, and consequently only one-color channel. However, the CNN we are using is a general porpoise CNN, and it is prepared to receive the colored image as an input. The strategy we propose is to use one color channel to include the length of the vehicles in the image. To do this, we need to map the maximum and minimum possible length in the range of the channel.

Table 13.Number of images used in CNN training Experiment1 vs Experiment 3

Class	**Training Experiment 1**			**Training Experiment 3**		
	Original	Modified	Total	Original	Modified	Total
1	1034	0	1034	1000	0	1000
2	1000	0	1000	1000	0	1000
3	1000	0	1000	999	0	999
4	1029	0	1029	1021	0	1021
5	613	399	1012	613	334	947
60	1001	0	1001	1001	0	1001
7	90	999	1089	75	894	969
8	29	998	1027	28	816	844
61	208	799	1007	208	677	885
62	2	1000	1002	2	802	804
63	682	350	1032	682	350	1032

Source: Joaquin Barreyro, 2020

5 RESULTS

This chapter presents the results achieved by applying the defined methodology. First, as a newly adapted AlexNet network refinement, we describe how we arrived at the best value for the mini-batch size and the number of epochs. Next, we present the results achieved in each experiment. Finally, we do a qualitative analysis of failures followed by a discussion of the results.

5.1 DETERMINING THE MINI-BATCH AND THE NUMBER OF EPOCHS

We trained the network by varying the mini-batches size between 3 possible values, 32, 64, and 128 samples. For each mini-batch size, we train up to a maximum of 50 epochs. Every new Epoch, we validate the result. In Fig 14, we show the result of the exercise. We can see that the best result was obtained with values for 32 mini-batch, with 29 training epochs.

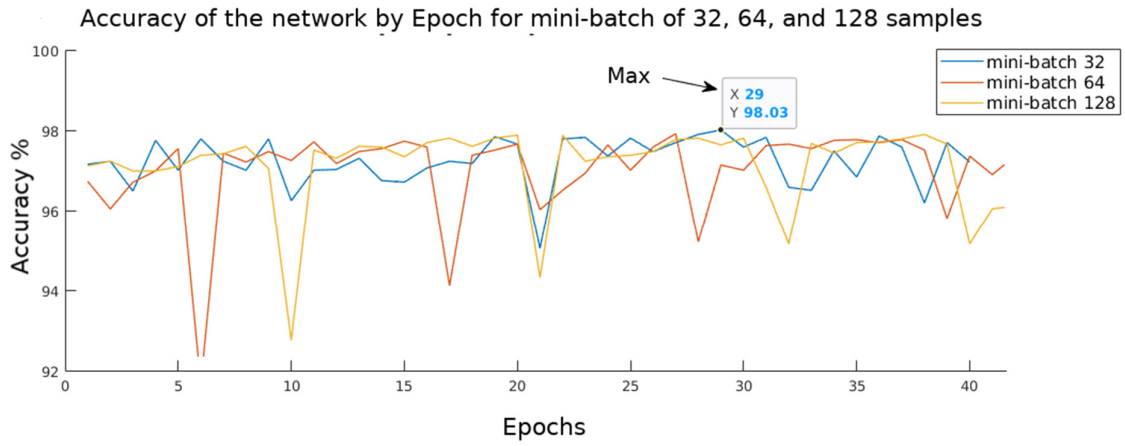


Fig 14.Accuracy comparison for different epochs and mini-batches

5.2 RESULTS OBTAINED IN EXPERIMENT1

In Fig 15 -the category confusion matrix-. We present the results obtained in experiment 1. Each row of the matrix corresponds to a natural category. Each column of the matrix corresponds to a predicted category. The number in each cell corresponds to several observations. Diagonal and off-diagonal cells correspond to correctly and incorrectly classified observations, respectively.

At the right side of Fig 15, a row summary displays the classification accuracy and the misclassification rate for each real category presented as a percentage. The classification accuracy is the ratio of correct predictions to total predictions made (all the values in the same row). Misclassification rate is the complement of the accuracy.

The global accuracy, 98,02%, is calculated as the ratio between the correct answers and the total number of tested images.

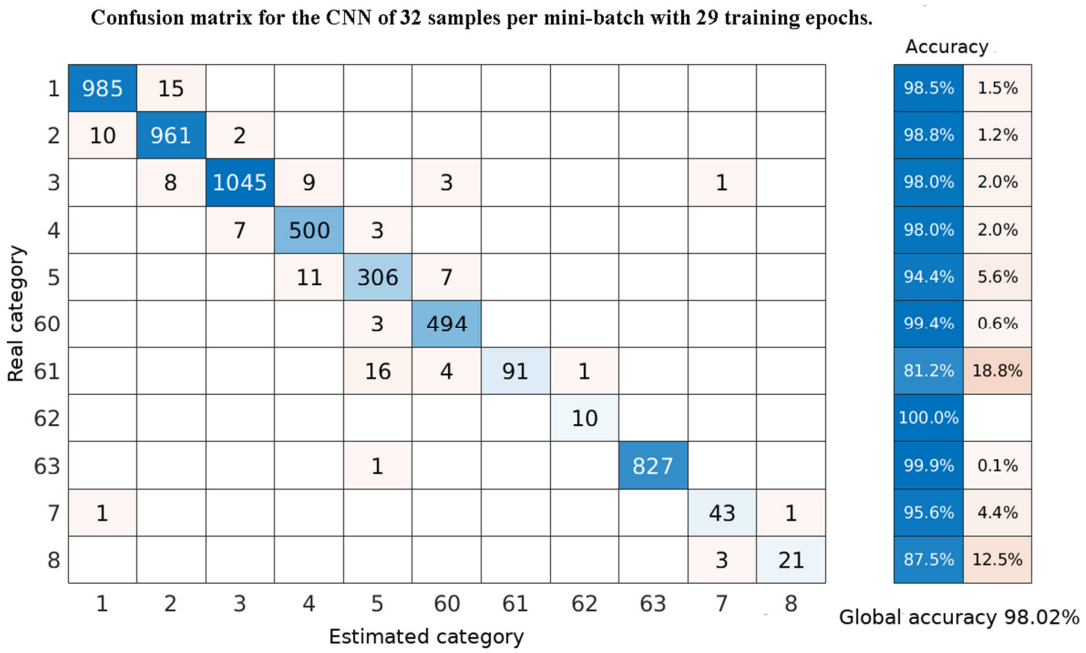


Fig 15.Confusion matrix for the CNN of 32 samples per mini-batch with 29 training epochs in experiment 1.

5.3 RESULTS OBTAINED IN EXPERIMENT2

The global accuracy obtained over all the datasets was 96,48%. In Fig 16, we show the confusion matrix. We can see that results are similar to experiment 1. However, performance in vehicle category 1 was better than the results in experiment 1. Moreover, we can see that over 60% of vehicles are category 1.

On the other hand, we have a loss of accuracy in the other categories, but the general accuracy is mitigated due to vehicles' lower occurrence in those categories. As we collect data in a rain period, it could be one reason for the field experiment's accuracy loss. To examine this possibility deeply, we focus on a specific toll station, P2, and look at the performance at each hour. We expect that the worst performance will be at the

moment of the heaviest rain. We compare a specific period in four different days. The number of vehicles observed each day in the same period, from 15:00hs to 16:00hs, is similar and representative: 903, 1084, 1372, and 701 for days 10/02, 11,12, and 13 of February, respectively. In Fig 17, we show the method's accuracy, per hour, on four different days. The referred period is highlighted with a circle. Days 10/02 and 11/02 are the days with rain. The accuracy at 15:00hs was 92.40% and 92.99%, respectively.

Days 12/02 and 13/02 are the days without rain. The accuracy at 15:00hs was 96.42% and 98,00%. Thus, we lose accuracy of at least 3% in rainy periods.

In Fig 18, we can see some examples of vehicles. Unfortunately, the water accumulated on the pavement is splashed in the vehicle's passage and generates noise in the region of the wheels' image. This is difficult for the correct classification.

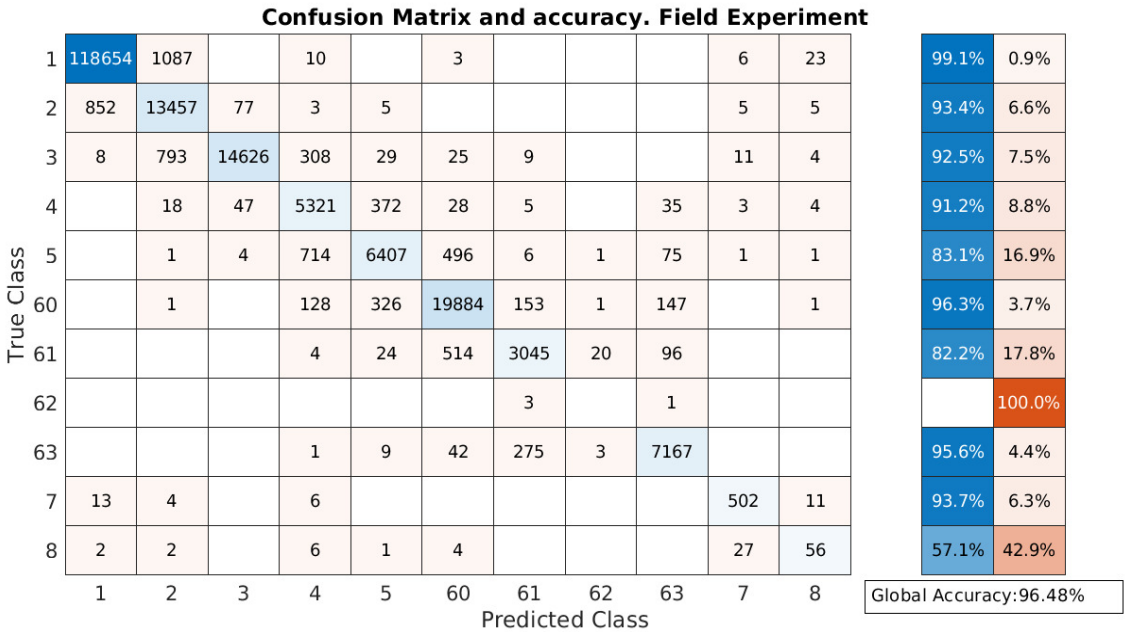


Fig 16.Confusion matrix for the CNN of 32 samples per mini-batch with 29 training epochs in experiment 2.

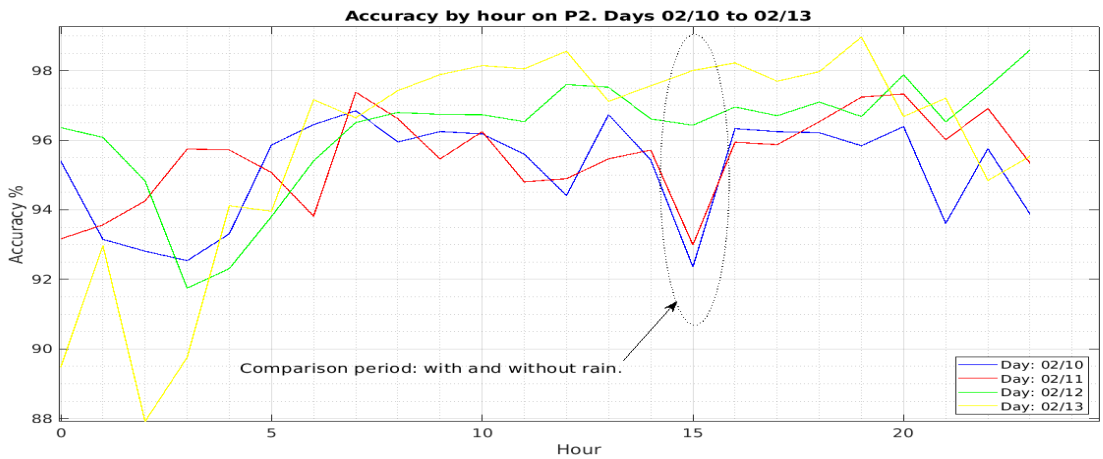


Fig 17.Accuracy by hour on P2.

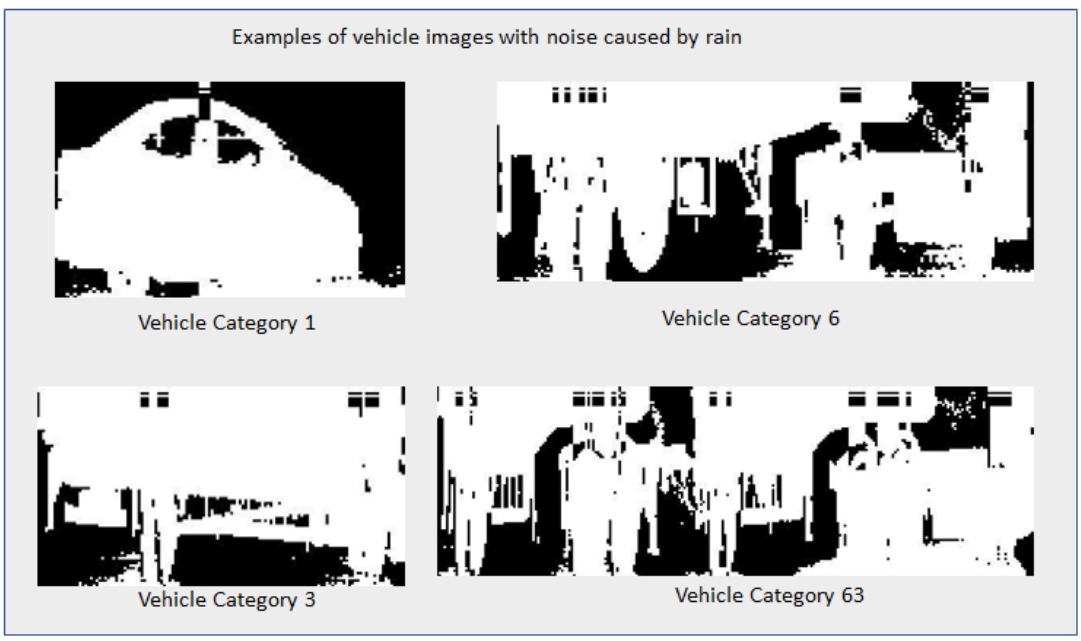


Fig 18.Images were obtained in a rainy period in experiment 2

Another interesting point to compare is manual and automatic lanes' performance. Automatic lanes are faster than manual lanes because the vehicle never stops when go throw the lane. The max allowed speed is 40km/h. In Manual lanes, the vehicle must stop to make the payment. Therefore, going out is relatively slower than an automatic lane.

We expect to have better images in manual lanes and better performance too.

On the other side, we can assume that the vehicle speed is relatively constant when passing in an automatic lane. In manual lanes, the vehicle is accelerating, trying to return to its average trip speed. We implemented a strategy to avoid this type of variation illustrated in Fig 9. That is why we expect that this effect does not affect the classification. To help in Table 14, we show the accuracy obtained in the different scenarios. The Highest and Lowest accuracy was calculated by hour.

We confirm that manual lanes have better accuracy in all scenarios than automatic lanes. Also, we can see that rain has more influence in automatic lanes than in manual lanes. One explanation for this may be that the automatic lanes are more exposed to the weather, unlike the manual lanes protected by the marquee.

Table 14.Experiment 2. Accuracy by lane type.

		Accuracy		
		Global (%)	Highest(%)	Lowest(%)
Global Results P1+P2+P3	Total	96.48	98.24	90.13
	Automatic	95.77	97.81	87.21
	Manual	96.95	98.83	90.01
P2 Global Results	Total	96.40	98.71	86.81
	Automatic	95.08	100	71.42
	Manual	96.74	99.39	86.81
P2 12hs Rain period	Total	95.69	96.66	92.18
	Automatic	94.07	97.97	88.59
	Manual	96.11	96.98	93.09
P2 12hs Without rain	Total	97.90	98.71	95.97
	Automatic	96.88	98.92	93.55
	Manual	98.20	98.97	96.68

Source: Joaquin Barreyro, 2021

5.4 RESULTS OBTAINED IN EXPERIMENT3

The first part of this experiment consists of training again, the same CNN we trained in experiment 1, but with a clean dataset. We did the same process as in experiment 1. The best result was with a mini-batch of 32 samples and 25 epochs. We had a global accuracy of 98.18%. It is not an expressive result, but considering that we did not modify the CNN or the images, it shows the relation between a good dataset and good results.

In the second part of this experiment, we included the length of vehicles as a background color variation. In Fig 19, we show some examples. Longer vehicles have a white background, and shorter ones have a yellow background. The global accuracy obtained in experiment 3 was 98,39%. In Fig 20, we show the confusion matrix. We can see that results are better than in experiment 1. We had a slight gain in category one

vehicles. This is important when applying the solution in the field because category 1 is the highest frequency in toll collection systems. Besides this, we had a slight improvement in almost all categories. We need to consider that we already had a good performance in experiment 1, and at this point, we need considerable efforts for minor gains.

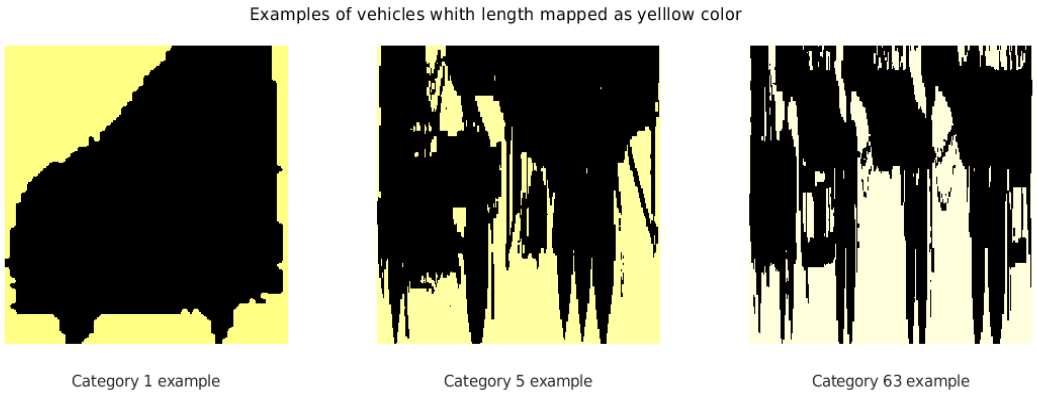


Fig 19.Example of color highlighting images with length information as a background.

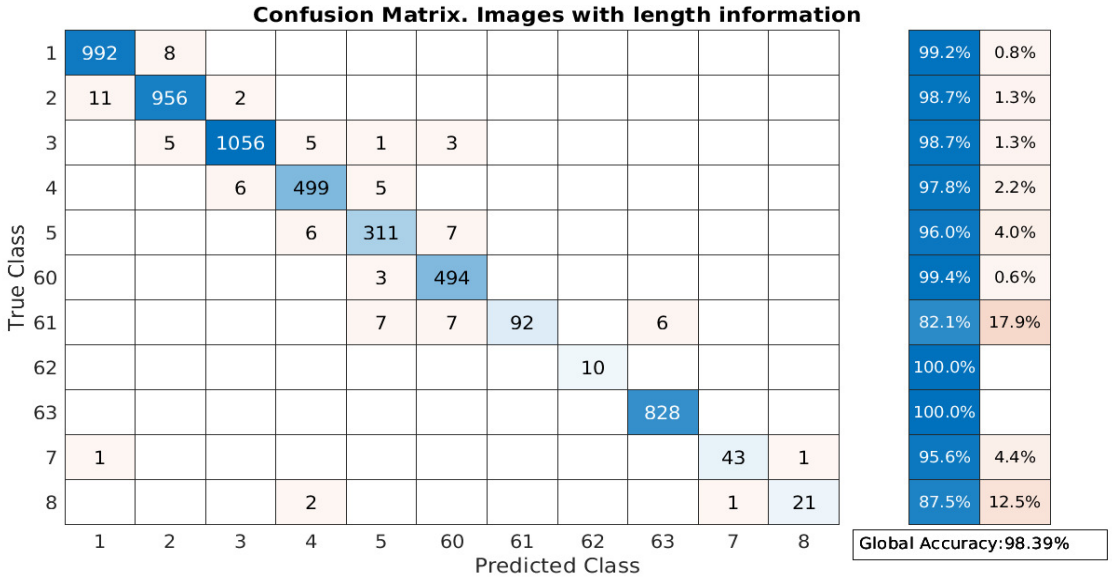


Fig 20.Confusion matrix. Images with length information. Experiment 3.

5.5 A QUALITATIVE FAILURE ANALYSIS

This section addresses a particular case's failures between categories 1 and 2. They are two categories with the highest frequency in tolls. For example, category 1 had 1.5% failures, 15 vehicles omitted from that category, but commissioned in category 2. The two categories have two axles, and the difference between them is the double wheel.

Fig 21 shows the image generated with the optical barrier and an actual photo of a typical case. This type of truck, which does not have double wheels, has similar characteristics to many category 2 trucks. A possible alternative to treat cases like this is to analyze the penultimate layer of the network. The output of this layer is a number that represents the confidence index from 0.0 to 1.0. In the example analyzed, the confidence index calculated for category 2 was 0.66, with the same case having 0.30 confidence for category 1. In 9 of the 15 failures, the confidence index was less than 0.8. However, of the 985 hits in category 1, 975 had a confidence index greater than 0.8, indicating some possibilities of treatments to be analyzed in future work to improve the index.

Fig 21.Category 1 truck, detected as category 2

5.6 ANALYSIS AND DISCUSSION OF RESULTS

In this section, we analyze the results obtained. Looking at the global result, the accuracy of 98,39% in experiment 3, the proposed method is near state of the art. It is difficult to compare studies because of the different Classification criteria, fields tests conditions, and the number of samples evaluated. We made a list of the top ten higher global accuracy studies analyzed in Section 2.5. We choose only those studies that classify vehicles in 8 categories or more. The result is shown in Table 15. The studies are ordered by accuracy. We can see that this research is in fourth place. The three studies with better performance are based on intrusive technologies. The first and second ones are based on accelerometers, and the third is based on piezoelectric sensors. If we consider only non-intrusive technologies, this research has higher accuracy. We analyze many studies based on non-intrusive methods, but only three are classified into eight or more categories. We emphasize the fact that this comparison is not accurate. The scenarios are different between studies, and the intention is only an estimate.

Table 15.Higher Accuracy studies

Primary study	Ref	Intrusive?	Results
S16 (BAJWA; RAJAGOPAL, *et al.*, 2011)		Yes	100% over 53 vehicles, in FHWA categories
S15 (MA; XING, *et al.*, 2013)		Yes	99% over 600 vehicles in FHWA categories
S19 (RAJAB; REFAI, 2014)		Yes	98.9% over 1000 vehicles in FHWA categories
THIS		No	98.39% over in 11 categories
S12 (KIM; COIFMAN, 2013)		Yes	97% over 18000 vehicles, in FHWA categories.
S43 (WU; XU, *et al.*, 2019)		No	91.98% over 1095 vehicles, in 10 categories.
S47 (GONZÁLEZ; JIMÉNEZ; DE FRUTOS, 2020)		Yes	90% over 457 vehicles, in 10 categories
S21 (CICARELLI; RIBEIRO; LEANDRO, 2011)		Yes	90% over 2000 vehicles, in 8 categories.
S36 (ZHAO; WU, *et al.*, 2018)		Yes	80.36% over 415 vehicles in 10 categories.
S1. (YAO, 2016)		No	65% over 1500 vehicles, in FHWA categories

Source: Joaquin Barreyro, 2020

If we look at the performance of the proposed method in each category in a particular way, we can see that it has high accuracy in all categories. Moreover, even those categories with more axles have good results. The exception is category 61, with 17.9% misclassifications in experiment 3.

We use artificial data in many categories in the training process, and we were concerned about how this affects results. Categories 7, 8, 61, and 62 used more artificial data than other categories. More than 80%. Three of them have the worst accuracy. Category 63 has an accuracy of 99,9% in experiment 1 and 100% in experiment 3, and it uses artificial images. However, the percentage of artificial data is only 33%. As expected, the number of artificial data employed affects the result. We can see an evolution from experiment 1 to experiment 3. We improved the train dataset's quality, and we had a better performance in almost all categories.

Another point of attention was whether the method would differentiate vehicles with different categories due to the double wheels and the same number of axles. It is impossible to view the double wheel directly in the vehicle side profile. It means that there must be some other feature in the profile strongly linked to that double wheel feature. Examples of double wheel influence are omission errors and commission between categories 1 and 2, with two axles, and between categories 3 and 7. In Section 5.5, we presented an example of this issue. Although we see errors of this type, 26 misclassifications in experiment 1, this number is smaller than the number of misclassifications because of the axles, 80 errors. In experiment 3, we had 21 misclassifications of the double wheel and 65 because of the axles. This shows that in experiment 3, we improve the detection of both types of errors proportionally.

Finally, some observations about experiment 2. First, it shows the complexity of deploying a controlled solution in the field. Second, the performance is affected by external factors that, in many cases, are unpredictable. We isolated only one of them, the rain, and show that the accuracy decreased by 3%. Finally, in experiment 2, we show that the distribution of categories is not homogeneous. In our experiment, 60% of vehicles were category 1, followed by category 60 with 10%. On the other hand, category 62 had only four vehicles.

6 CONCLUSION AND FUTURE WORK

This research proposed a set of techniques to classify vehicles using binary images obtained from optical barriers, with the data processed in an adapted convolutional neural network - AlexNet.

The classification obtained was satisfactory since the global accuracy index was 98,39%. It was shown that the proposed method has good performance indexes in almost all categories. The performance is comparable with the best results within the literature.

We did the first experiment, analyzed results, identified some opportunities, and proposed a new experiment, including the length of vehicles inside the image and refining the training dataset. As a result, we improved the accuracy by 0,37%.

We also did an in-field experiment. We confirm the results obtained in the controlled experiment. We identified a loss of accuracy of 3% in rainy periods. We also identified a slight difference in performance between manual and automatic lanes. This difference could be because of the speed of vehicles.

There is the possibility of improvement since 25% of errors occur in categories with the same number of axles. In these cases, analyzing the network's penultimate layer proved to be a path to be explored.

In future work, we can study the possibility of correcting the result according to the image's other characteristics. In cases such as the example analyzed, where the difference in the confidence index is not more significant than a specific value, it would be possible to do another type of treatment, studying other characteristics of the image, for example, the height of the vehicle in the first axle, or size of the wheels. A statistical approach could also be adopted. We could correct the confidence index by multiplying by a factor related to the probability of this category's occurrence (The probability of occurrence of each category can be calculated based on historical data from the concessionaire).

REFERENCES

AGÊNCIA DE TRANSPORTE DO ESTADO DE SÃO PAULO. ARTESP - Tarifas de pedágios., 2015. Disponível em: <http://www.artesp.sp.gov.br/rodovias-tarifas-de-pedagios.html>.

AGÊNCIA NACIONAL DE TRANSPORTES TERRESTRES. ANTT - Concessões rodoviárias, 2015. Disponível em: <https://appweb.antt.gov.br/faq/concessoes_rodoviarias.asp>.

AKULA, A. et al. Towards an optimal bag-of-features representation for vehicle type classification in thermal infrared imagery. **In Optics and Photonics for Information Processing XII**, p. 107511C, 2018.

ALEXANDRE, E. et al. Hybridizing extreme learning machines and genetic algorithms to select acoustic features in-vehicle classification applications. **Neurocomputing**, 2015. 58-68.

BAJWA, R. et al. **In-pavement wireless sensor network for vehicle classification**. 10th ACM/IEEE International Conference on Information Processing in Sensor Networks. [S.l.]: IEEE. 2011. p. 85-96.

BALID, W; TAFISH, H; REFAI, H H. Intelligent vehicle counting and classification sensor for real-time traffic surveillance. **IEEE Transactions on Intelligent Transportation Systems, vol. 19, no. 6**, 2017. 1784–1794.

BARREYRO, Joaquín; YOSHIOKA, Leopoldo R.; MARTE, Claudio L. **Side Profile Vehicle Binary Images**. [S.l.]. 2021a.

BARREYRO, Joaquín; YOSHIOKA, Leopoldo R.; MARTE, Claudio L. **A vehicle classification system based on a non-intrusive sensor's binary image and convolutional neural networks**. 2021 14th IEEE International Conference on Industry Applications (INDUSCON). [S.l.]: IEEE. 2021b. p. 813-819.

BARREYRO, Joaquín; YOSHIOKA, Leopoldo R.; MARTE, Claudio L. **A non-intrusive category identification method based on the binary image of profile vehicles and CNN classification algorithm**. 2021 IEEE Intelligent Transportation Systems Conference (ITSC). Indianapolis, USA.: IEEE. 2021c. p. 1803-1808.

CICARELLI, L. C.; RIBEIRO, E. P.; LEANDRO, G. V. **Classificação de veículos automotivos através de sensores indutivos e redes neurais**. 10th Brazilian Congress

on Computational Intelligence (CBIC'2011). Fortaleza, Ceará Brazil: Brazilian Society on Computational Intelligence. 2011.

CONFEDERAÇÃO NACIONAL DE TRANSPORTE. **Anuário CNT do transporte - Estatísticas consolidadas 2019**. [S.l.]. 2020.

DA COSTA FILHO, A. C. B. et al. **Infrared-based system for vehicle classification**. International Microwave and Optoelectronics Conference (IMOC). [S.l.]: IEEE. 2009. p. 537-540.

DE ANGELIS, G. et al. **A simple magnetic signature vehicles detection and classification system for smart cities**. 2016 IEEE International Symposium on Systems Engineering (ISSE). [S.l.]: IEEE. 2016. p. 1-6.

DEPARTAMENTO NACIONAL DE TRÂNSITO. http://www.denatran.gov.br/download/. **Frota de veículos**, 2015. Disponível em: <http://www.denatran.gov.br/download/>.

DOGO, E. M. . A. O. J. . N. N. I. . T. B. . &. A. C. O. **December). A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks.** International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS). [S.l.]: IEEE. 2018. p. 92-99.

DONG, H. et al. Improved robust vehicle detection and identification based on single magnetic sensor. **Ieee Access**, 2018. 5247-5255.

FEDERAL HIGHWAY ADMINISTRATION. T. F. H. Administration. Traffic monitoring guide., 2013. Disponível em: <https://www.fhwa.dot.gov/ policyinformation/tmguide/tmg_fhwa_pl_13_015.pdf>.

GONZÁLEZ, B.; JIMÉNEZ, F. J.; DE FRUTOS, J. A virtual instrument for road vehicle classification based on piezoelectric transducers. **Sensors**, p. 4597, 2020.

GOTHANKAR, N.; KAMBHAMETTU, C.; MOSER, P. **Circular Hough Transform Assisted CNN Based Vehicle Axle Detection and Classification**. International Conference on Intelligent Transportation Engineering (ICITE). [S.l.]: IEEE. 2019. p. 217-221.

GOTHANKAR, N; KAMBHAMETTU, C; MOSER, P. **Circular hough transform assisted CNN based vehicle axle detection and classification**. 4th International Conference on Intelligent Transportation Engineering (ICITE). IEEE. [S.l.]: [s.n.]. 2019. p. 217-221.

GRIGORYEV, A., B. D. . T. A. P. . &. N. D. P. Vision-Based Vehicle Wheel Detector And Axle Counter. **ECMS**, p. 521-526, 2015.

GRIGORYEV, A; KHANIPOV, T; NIKOLAEV, D. **Determination of axle count for vehicle recognition and classification**. 8th open german-russian workshop «pattern recognition and image understanding. Nizhny Novgorod: [s.n.]. 2011. p. 89–91.

HASSELBRING, R. E. **Doppler-radar based automatic vehicle-classification system**. 5,717,390, 1998.

HE, Y. et al. Improved waveform-feature-based vehicle classification using a single-point magnetic sensor. **Journal of Advanced Transportation,** 2015. 663-682.

HUANG, Y.; LU, P.; TOLLIVER, D. vehicle classification system using in-pavement fiber Bragg grating sensors. **IEEE Sensors Journal**, p. 2807-2815, 2018.

IWASAKI, Y.; MISUMI, M.; NAKAMIYA, T. Robust vehicle detection under various environmental conditions using a thermal infrared camera and its application to road traffic flow monitoring. **Sensors**, p. 7756-7773, 2013.

KHANIPOV, T et al. Vision-based industrial automatic vehicle classifier. **Seventh International Conference on Machine Vision (ICMV 2014)**, p. 944511, 2015.

KHATRI, H, S.; SOMANI, S, B. **Infrared-based system for vehicle counting and classification," in**. Pervasive Computing (ICPC) 2015 International Conference on. IEEE, 2015. [S.l.]: [s.n.]. 2015. p. 1–5.

KHATRI, H. S.; SOMANI, S. B. **Infrared-based system for vehicle counting and classification**. International Conference on Pervasive Computing (ICPC). [S.l.]: IEEE. 2015. p. 1-5.

KIM, S.; COIFMAN, B. Evaluation of axle-based and length-based vehicle classification stations. **Transportation research record**, p. 1-12, 2013.

KITCHENHAM, B. **Procedures for performing systematic reviews**. Keele, UK. 2004.

KLEIN, L. A.; MILLS, M. K.; GIBSON, D. R. **Traffic detector handbook:** Volume I.

KLEYKO, D. et al. **Vehicle classification using roadside sensors and feature-free data smashing approach**. 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). [S.l.]: IEEE. 2016. p. 1988-1993.

KRIZHEVSKY, A., S. I. . &. H. G. E. Imagenet classification with deep convolutional neural networks. **Advances in neural information processing systems**, 2012. 1097-1105.

KUMAR, R.; GOYAL, D.; KUMAR, R. Real-Time Toll Rate Determination Using Image Processing & Network Database. **International Journal of Digital Application & Contemporary research**, 2014. 6.

LECUN, Y., B. L. . B. Y. . &. H. P. Gradient-based learning applied to document recognition. **Proceedings of the IEEE,** 1998. 2278-2324.

LEDUC, G. Road traffic data: Collection methods and applications. **Working Papers on Energy, Transport and Climate Change, vol. 1, no. 55**, 2008.

LEE, H.; COIFMAN, B. Using LIDAR to validate the performance of vehicle classification stations. **Journal of Intelligent Transportation Systems,** 2015. 355-369.

LEE, H; COIFMAN, B. Side-fire lidar-based vehicle classification. **Transportation Research Record**, p. 173-183, 2012.

LI, S et al. Video-based traffic data collection system for multiple vehicle types. **IET Intelligent Transport Systems**, p. 164-174, 2014.

LIU, C et al. A vision-based pipeline for vehicle counting, speed estimation, and classification. **IEEE Transactions on Intelligent Transportation Systems,** 2020.

LIU, C. et al. A vision-based pipeline for vehicle counting, speed estimation, and classification. **IEEE Transactions on Intelligent Transportation Systems.**, 2020.

LIU, Y; WANG, K. **Vehicle classification system based on dynamic Bayesian network**. Proceedings of 2014 IEEE International Conference on Service Operations and Logistics, and Informatics. [S.l.]: IEEE. 2014. p. 22-26.

MA, W. et al. A wireless accelerometer-based automatic vehicle classification prototype system. **IEEE Transactions on Intelligent Transportation Systems**, 2013. 104-111.

MARTE, C. L.; MELLO, F.; FERREIRA, M R. **Sistemas inteligentes de transportes, serie cadernos técnicos**.

MASTERS, D.; LUSCHI, C. Revisiting small batch training for deep neural networks. **arXiv preprint arXiv:1804.07612.**, 2018.

MATLAB. Pretrained deep neural networks. **MathWorks Help Center**, 01 Jan. 2020. Disponível em: <https://www.mathworks.com/help/>.

MCGOWEN, P.; SANDERSON, M. **Accuracy of pneumatic road tube counters**. Institute of Transportation Engineers (ITE) Western District Annual Meeting. [S.l.]: Institute of Transportation Engineers (ITE). 2011.

MIMBELA, J L.; KLEIN, L A. **Summary of vehicle detection and surveillance technologies used in intelligent transportation systems**.

MINKINA, A. et al. **Generalization of the Viola-Jones method as a decision tree of robust classifiers for real-time object recognition in the video stream**. Seventh International Conference on Machine Vision (ICMV 2014). [S.l.]: International Society for Optics and Photonics. 2014. p. 944517.

MOGELMOSE, A; MOESLUND, T. B. Analyzing wheels of vehicles in motion using laser scanning. **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops.**, p. 1-8, 2016.

NEZAFAT, R, V.; SAHIN, O; CETIN, M. "Transfer learning using deep neural networks for classification of truck body types based on side-fire lidar data". **Journal of Big Data Analytics in Transportation, vol. 1, no. 1**, 2019. 71–82.

NEZAFAT, R. V.; SAHIN, O.; CETIN, M. Transfer learning using deep neural networks for classification of truck body types based on side-fire lidar data. **Journal of Big Data Analytics in Transportation**, 2019. 71-82.

NEZAFAT, R. V.; SALAHSHOUR, B; CETIN, M. **Classification of truck body types using a deep transfer learning approach**. International Conference on Intelligent Transportation Systems (ITSC). [S.l.]: IEEE. 2018. p. 3144-3149.

PENG, Y et al. **Binary coding-based vehicle image classification.** 12th International Conference on Signal Processing (ICSP). [S.l.]: [s.n.]. 2014. p. 918-921.

PENG, Y. et al. **Vehicle classification using sparse coding and spatial pyramid matching**. 17th International IEEE Conference on Intelligent Transportation Systems (ITSC). [S.l.]: [s.n.]. 2014. p. 259-263.

PEREZ, L., &. W. J. **The effectiveness of data augmentation in image classification using deep learning**. [S.l.]. 2017.

RAJAB, S. A.; REFAI, H. H. **A single element piezoelectric sensor for vehicle classification using the ratio of track width to length**. 2014 IEEE Intelligent Vehicles Symposium Proceedings. [S.l.]: IEEE. 2014. p. 1463-1468.

RUSSAKOVSKY, O. et al. Imagenet large scale visual recognition challenge. **International journal of computer vision**, p. 211-252, 2015.

SASONGKO, A, T.; FANANY, M, I. **Indonesia toll road vehicle classification using transfer learning with pre-trained resnet models**. 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI). IEEE. [S.l.]: [s.n.]. 2019. p. 373.

SASONGKO, A. T.; FANANY, M. I. **Indonesia toll road vehicle classification using transfer learning with pre-trained resnet models.** International Seminar on Research of Information Technology and Intelligent Systems (ISRITI). [S.l.]: IEEE. 2019. p. 373-378.

SHIN, H. C. . R. H. R. . G. M. . L. L. . X. Z. . N. I. .. &. S. R. M. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. **IEEE transactions on medical imaging**, 2016. 1285-1298.

SKSZEK, S L. **state-of-the-art report on non-traditional traffic counting methods**. [S.l.]. 2001.

TAFISH, H.; BALID, W.; REFAI, H. H. **Cost effective vehicle classification using a single wireless magnetometer**. 2016 International Wireless Communications and Mobile Computing Conference (IWCMC). [S.l.]: IEEE. 2016. p. 194-199.

TAGHVAEEYAN, S.; RAJAMANI, R. **Portable roadside sensors for vehicle counting, classification, and speed measurement**. IEEE Transactions on Intelligent Transportation Systems. [S.l.]: IEEE. 2013. p. 73-83.

UEDA, T. et al. Optical axle-counting equipment. **Proceedings of Conference on Intelligent Transportation Systems**, p. 326-331, 1997.

VAA, T et al. Norsikt–nordic system for intelligent classification of traffic. In: _____ **Procedia-Social and Behavioral Sciences**. [S.l.]: [s.n.], 2012. p. 1702–1712.

VIOLA, P.; JONES, M. J. Robust real-time face detection. **International journal of computer vision**, p. 137-154, 2004.

WEN, X. et al. A rapid learning algorithm for vehicle classification. **Information sciences,** p. 295-406, 2015.

WU, J. et al. Automatic vehicle classification using roadside LiDAR data. **Transportation Research Record**, p. 153-164, 2019.

YANG, B.; LEI, Y. Vehicle detection and classification for low-speed congested traffic with anisotropic magnetoresistive sensor.. **IEEE Sensors Journal**, 2014. 1132-1138.

YAO, Z., W. H. . L. Z. . &. C. J. Fuzzy C-Means Image Segmentation Approach for Axle-Based Vehicle Classification. **Transportation Research Record,** p. 68-77, 2016.

YU, X.; PREVEDOUROS, P. D.; SULIJOADIKUSUMO, G. Evaluation of autoscope, SmartSensor HD, and infra-red traffic logger for vehicle classification. **Transportation research record**, p. 77-86, 2010.

ZHANG, B.; ZHANG, J. Zhang, B., & Zhang, J. (2020). A traffic surveillance system for obtaining comprehensive information of the passing vehicles based on instance segmentation. **IEEE Transactions on Intelligent Transportation Systems.**, 2020.

ZHAO, H. et al. A vibration-based vehicle classification system using distributed optical sensing technology. **Transportation Research Record**, p. 12-23, 2018.

ANNEX 1 – DATASET DESCRIPTION

This is a dataset of vehicle's binary images obtained with an optical barrier. It is possible to obtain a copy of the complete dataset online (BARREYRO; YOSHIOKA; MARTE, 2021a). The main goal of this dataset is to train and test a vehicle classification method. Categories are defined according to the number of axles and the presence of a double wheel or not.

This dataset is composed of 3 folders:

Train: Images used to train a CNN. (Over 1000 images/cat) Total of 11232 images.

Experiment1: Images used to validate: 5389 images.

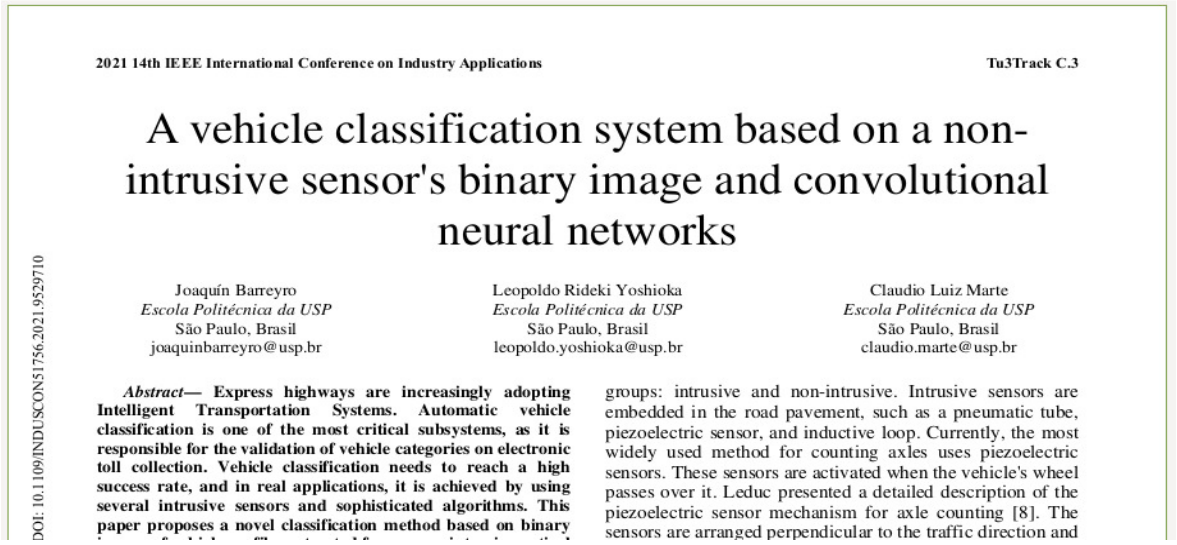
Experiment2: A field test. Four days of images: 196018 images.

Train and Experiment1 are selected images from the same lane in different periods over three months. Experiment 2 has images from several lanes and 3 different toll stations. We do not delete any image in this set. In each folder, there is a compressed file. The images will be organized in folders by category when the file is downloaded and extracted, as we show in Fig 22. First, the name of the image file is composed by date, followed by the identification of the station, lane, and transit id that generated it. All the images are in png format.

Fig 22.Visualization of the image dataset

ANNEX 2 – PUBLICATIONS DERIVED FROM THIS RESEARCH

As a partial result of this research, we elaborated two papers: the first one (BARREYRO; YOSHIOKA; MARTE, 2021b) was published in the 14th IEEE International Conference on Industry Applications (IEEE/IAS), and the second one (BARREYRO; YOSHIOKA; MARTE, 2021c) was published in the 24th IEEE International Conference on Intelligent Transportation Systems (IEEE\ ITSC).



2021 14th IEEE International Conference on Industry Applications Tu3Track C.3

A vehicle classification system based on a non-intrusive sensor's binary image and convolutional neural networks

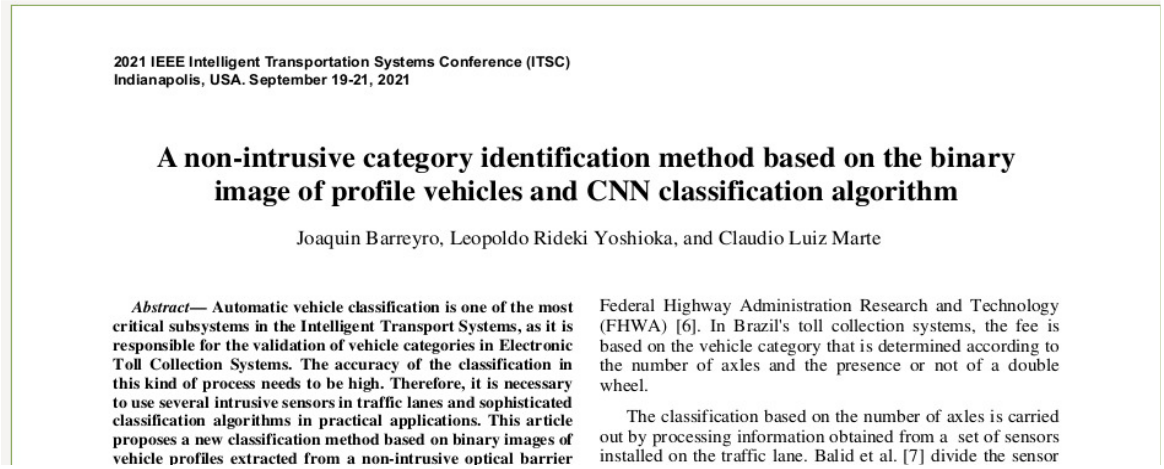
DOI: 10.1109/INDUSCON51756.2021.9529710

Joaquín Barreyro, *Escola Politécnica da USP*, São Paulo, Brasil, joaquinbarreyro@usp.br
Leopoldo Rideki Yoshioka, *Escola Politécnica da USP*, São Paulo, Brasil, leopoldo.yoshioka@usp.br
Claudio Luiz Marte, *Escola Politécnica da USP*, São Paulo, Brasil, claudio.marte@usp.br

Abstract— **Express highways are increasingly adopting Intelligent Transportation Systems. Automatic vehicle classification is one of the most critical subsystems, as it is responsible for the validation of vehicle categories on electronic toll collection. Vehicle classification needs to reach a high success rate, and in real applications, it is achieved by using several intrusive sensors and sophisticated algorithms. This paper proposes a novel classification method based on binary**

groups: intrusive and non-intrusive. Intrusive sensors are embedded in the road pavement, such as a pneumatic tube, piezoelectric sensor, and inductive loop. Currently, the most widely used method for counting axles uses piezoelectric sensors. These sensors are activated when the vehicle's wheel passes over it. Leduc presented a detailed description of the piezoelectric sensor mechanism for axle counting [8]. The sensors are arranged perpendicular to the traffic direction and

Fig 23.Print Screen of the first page of the paper published in INDUSCON



2021 IEEE Intelligent Transportation Systems Conference (ITSC)
Indianapolis, USA. September 19-21, 2021

A non-intrusive category identification method based on the binary image of profile vehicles and CNN classification algorithm

Joaquin Barreyro, Leopoldo Rideki Yoshioka, and Claudio Luiz Marte

Abstract— **Automatic vehicle classification is one of the most critical subsystems in the Intelligent Transport Systems, as it is responsible for the validation of vehicle categories in Electronic Toll Collection Systems. The accuracy of the classification in this kind of process needs to be high. Therefore, it is necessary to use several intrusive sensors in traffic lanes and sophisticated classification algorithms in practical applications. This article proposes a new classification method based on binary images of vehicle profiles extracted from a non-intrusive optical barrier**

Federal Highway Administration Research and Technology (FHWA) [6]. In Brazil's toll collection systems, the fee is based on the vehicle category that is determined according to the number of axles and the presence or not of a double wheel.

The classification based on the number of axles is carried out by processing information obtained from a set of sensors installed on the traffic lane. Balid et al. [7] divide the sensor

Fig 24.Print Screen of the first page of the paper published in ITSC

ANNEX 3 – CODE IMPLEMENTATION DETAILS

In this research, we used MATLAB release 2020a. In addition, we used the Deep Learning toolbox. The operating system used was Ubuntu 20.04. In Fig 25, we show the flowchart of the main process of the developed program.

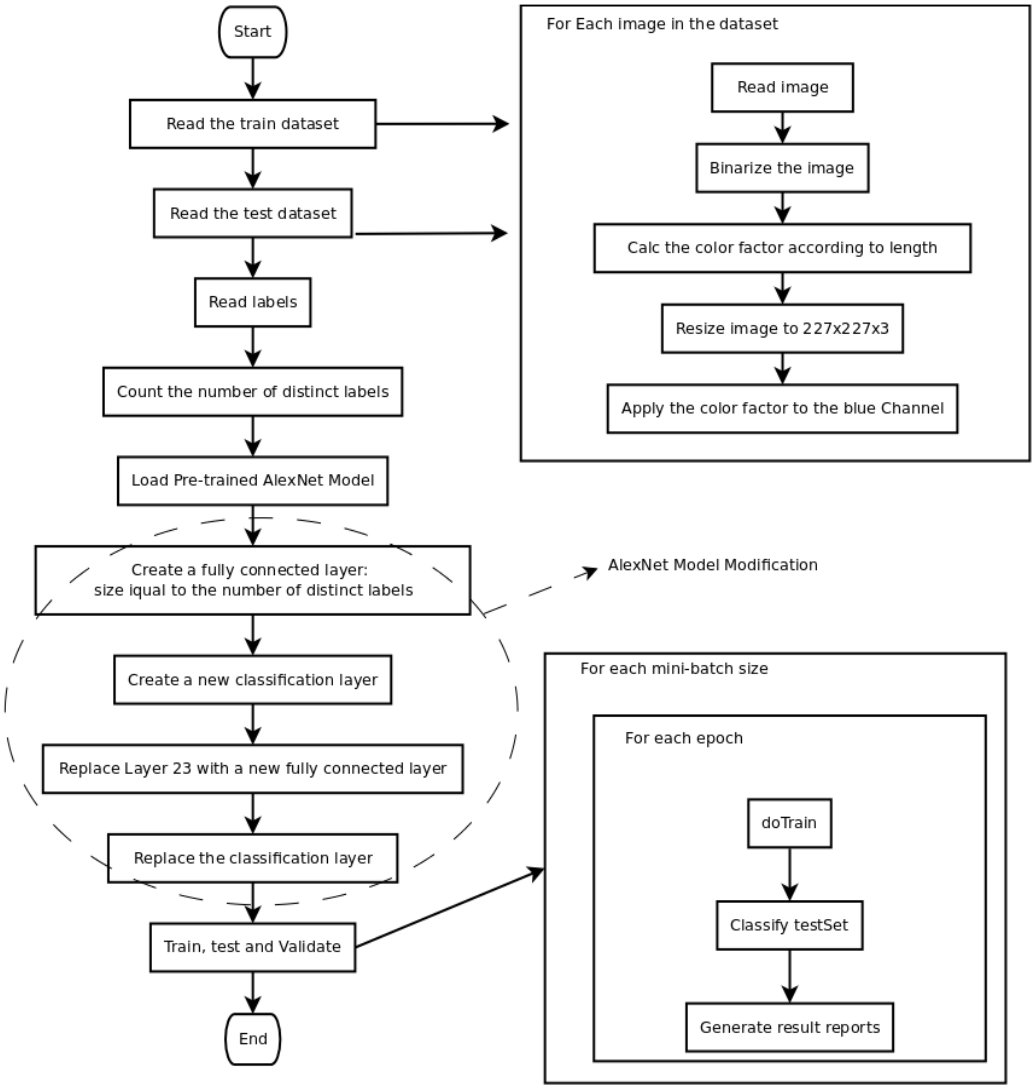


Fig 25.Train and Classification algorithm main process flowchart

ANNEX 4 – TOLL COLLECTION SYSTEM DESCRIPTION

Toll collection systems are tools, hardware, and software that help operate and control the entire toll collection process. There are several ways to do it, and it depends on the company that has the concession. In this annex, we describe one possibility, but not the only one. The toll control and collection system are modular in a pyramid, where each part has a specific function. We can see the architecture in Fig 26.

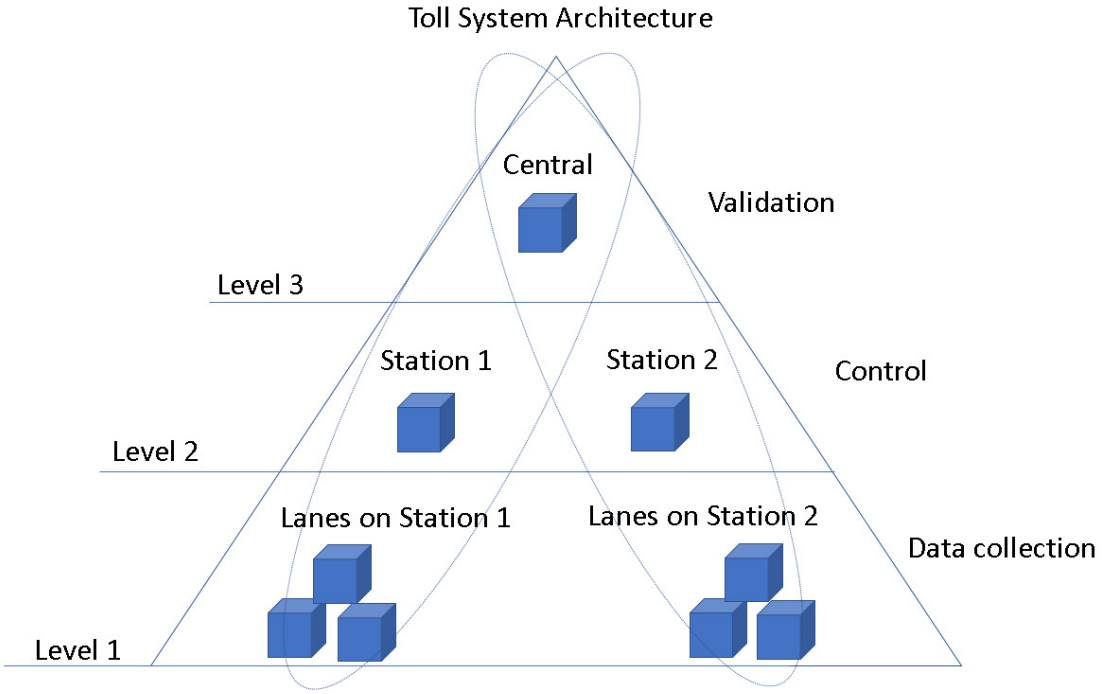


Fig 26.Toll system Architecture

Level 1 comprises the software and hardware installed in each lane. It integrates control and registration of vehicle passage, toll collection, road equipment control, and generated anomalies registration. Its central component is a PC per lane that commands and centralizes all the necessary peripherals, including cameras, printers, card readers, and all kinds of sensors.

Level 2 comprises the software and hardware installed in each toll station: Its function is to acquire and record the information generated by the toll lanes. Also, control and monitor the operation of lanes and peripherals, and make information available to plaza operators in real-time for decision making.

Level 3 - It is the core system. It integrates acquisition, registration, consolidation, and data processing of toll stations. Here, the preparation of operation reports on traffic and money collected, fraud control, generation and exchange of information with external systems, and treatment and distribution of different types of data necessary for collection, such as categories and rates, is produced. One of the activities that level 3 performs is validating traffic data. A team of people is dedicated to verifying if the category associated with each vehicle is correct. The system provides the information collected at Level 1. Depending on the lane type, the level 1 system makes available a category indicated by a collector, a category made available by a set of sensors, and sometimes an additional category made available by a redundant sensor system. With all this information, the visualization of images, and the help of algorithms that automate the process, level 3 operators validate the category of all vehicles.