






Article

ATRIP: Architecture for Traffic Classification Based on Image Processing

André Luis Cristiani ¹, Roger Immich ^{2,*} , Ademar T. Akabane ³ ,
Edmundo Roberto Mauro Madeira ⁴ , Leandro Aparecido Villas ⁴  and
Rodolfo I. Meneguette ^{5,*} 

¹ Computer Department, Federal University of São Carlos, São Carlos 13565-905, Brazil; aluis02@hotmail.com

² Digital Metropolis Institute, Federal University of Rio Grande do Norte, Natal 59078-970, Brazil

³ Center of Exact, Environmental and Technological Sciences, Pontifical Catholic University, Campinas 13087-571, Brazil; ademar.akabane@puc-campinas.edu.br

⁴ Institute of Computing, University of Campinas, Campinas 13083-852, Brazil; edmundo@ic.unicamp.br (E.R.M.M.); leandro@ic.unicamp.br (L.A.V.)

⁵ Instituto De Ciências Matemáticas e de Computação, University of São Paulo, São Carlos 13566-590, Brazil

* Correspondence: roger@imd.ufrn.br (R.I.); meneguette@icmc.usp.br (R.I.M.)

Received: 5 May 2020; Accepted: 28 May 2020; Published: 1 June 2020



Abstract: With the increase of vehicles in large urban centers, there is also an increase in the number of traffic jams and accidents on public roads. The development of a proper Intelligent Transport System (ITS) could help to alleviate these problems by assisting the drivers on route selections to avoid the most congested road sections. Therefore, to improve on this issue, this work proposes an architecture to aid an ITS to detect, analyze, and classify the traffic flow conditions in real time. This architecture also provides a control room dashboard to visualize the information and notify the users about the live traffic conditions. To this end, the proposed solution takes advantage of computer vision concepts to extract the maximum information about the roads to better assess and keep the drivers posted about the traffic conditions on selected highways. The main contribution of the proposed architecture is to perform the detection and classification of the flow of vehicles regardless of the luminosity conditions. In order to evaluate the efficiency of the proposed solution, a testbed was designed. The obtained results show that the accuracy of the traffic classification rate is up to 90% in daylight environments and up to 70% in low light environments when compared with the related literature.

Keywords: image processing; intelligent transport system; vehicle tracking; traffic classification; computer vision

1. Introduction

In the past few years, there has been a significant increase in the number of vehicles. A study by the US Department of Transportation in 2019 showed that, in the period from 2010 to 2016, the number of vehicles increased from 250 to 269 million and the distance traveled by these vehicles increased from approximately 2.97 to 3.17 billion of kilometers only in the United States [1]. This tendency is expected to increase even further in the foreseeing future.

More and more vehicles are being equipped with devices that can provide numerous services and applications for managing street traffic or even the entire traffic system in a large city [2]. Because of that, the Intelligent Transport Systems (ITSs) are envisioned to provide a number of new applications that spread from road and traffic alerts [3–5], distributed and cooperative route planning [6–10], several types of video distribution [11–15], driving strategy advisory systems [16–18], to autonomous

driving capabilities [19–21]. A report conducted by the Institute of Economic and Applied Research showed that more than 169,000 traffic accidents were registered on the federal highways in Brazil in 2014 [22]. Since then, this number has greatly decreased to 69,206 in 2018 mainly due to the strict legislation adopted and nationally enforced. However, the study still shows that 33.7% of these accidents are caused by head-on collisions, while 32.6% of them are due to a lack of attention [23]. This means that these kinds of accidents can be reduced by the use of devices and mechanisms designed to monitor the traffic, detect possible road incidents, and notify the drivers, increasing their alertness [24]. Thus, these alerts help the drivers take action for any event on the road, such as changing the route.

Several mechanisms have been proposed to alert drivers and provide traffic conditions notifications to improve on this issue. In the latter, the main objective is to detect which routes are coping with slow traffic, possible points of congestion, and sections with traffic interruption [25]. It is common for these mechanisms to use cables and sensors, such as piezoelectric, ultrasound, infrared, microwave, laser, and optics, to perform the counting of the vehicles [26]. Generally, inductive sensors are attached to the pavement to automatically detect the vehicles. This means that they only perform the counting of the vehicles in that specific section of the highway. Besides that, they have the capacity to classify them through the analysis of their magnetic profiles [27]. Another way to automatically monitor traffic is to use pre-installed cameras along the pathways. Recently, the use of video cameras became a promising and low-cost method to monitor traffic flow [28]. In addition to that, they can also be used to verify road conditions and different types of events, such as accidents and roadblocks. Because of that, the use of computer vision concepts, to monitor urban or countryside traffic, becomes an attractive ITS research field to develop new services and applications [29–31].

There are several methods to perform traffic analysis, for example, using mobile phone GPS [32] or cellular triangulation [33]. However, performing traffic analysis using cameras has many advantages than other monitoring methods. First of all, the cameras are easy to install; additionally, they do not damage the roads and also have relatively low installation and maintenance costs [34,35]. Furthermore, a large number of cameras are already installed in the roads for surveillance purposes. Besides that, a single camera is capable of monitoring more than one traffic lane over several hundred meters of road. Therefore, they provide a wide area of monitoring capabilities, which, in turn, leads to a better analysis of traffic flow and precise calculation of speed as well as counting and classification of vehicles [36]. Automatic traffic monitoring provides a rich source of information as the data collected through the cameras can be used in a great variety of situations, such as to discover the path traveled by the vehicle as well as the shape, dimensions, and color of the vehicles. These systems could potentially have the same powers of observation as a human agent, but without the damaging effects of tiredness and boredom caused by the repetitive task [37].

The use of video cameras brings many advantages; however, there are also some drawbacks. For example, the cameras are subject to vibration due to wind or heavy vehicles passing by. Additionally, the position of the camera also plays an important role. Generally, they are positioned high above the traffic and, because of that, depending on the position and angle that the camera captures the images, large vehicles may obstruct the view.

Another issue is that surveillance cameras that were not originally specified or installed to use computer vision algorithms may have to be tweaked to this end [25]. Moreover, the use of computer vision algorithms may be impacted by several other factors such as the weather, the luminosity, and foreign objects on the roads, just to give some examples. As a consequence of these adversities, it's been proven to be hard for the computer vision algorithm to take all these variations into account and to define parameters that work in all these distinct situations.

In the literature, several architectures use image processing for congestion detection [38–40]. Usually, they adopt image processing to identify and count vehicles. However, these systems do not take into account low-light scenarios such as vehicle detection at night. These environments present some challenges as they cannot accurately detect the vehicles' outline, therefore making its

classification difficult. Thus, in a darker environment such as at night, detecting and counting the vehicles to establish a level of congestion becomes a challenge.

In order to improve on these issues, such as how to better detect vehicles in low luminosity enrollments and to classify the traffic flow, this work proposes an architecture that allows the classification and users' notification about the flows of vehicles on the roads. The proposed architecture is divided into three main components: (i) monitoring and abstraction of information; (ii) information storage and processing; and (iii) control room dashboard.

These components are based on the concepts of computer vision, aiming to detect and classify the flows of vehicles that are traveling in highways settings. Although the proposed architecture is generic enough to work in a variety of situations, our approach was evaluated in a Brazilian multi-lane highways scenario. For this, real traffic video sequences were obtained from one of the Brazilian's highway concessionaires, namely AutoBAN [41]. The dataset used in this work can be found in [42].

The proposed architecture's main contribution is to perform the detection and classification of the flow of vehicles regardless of the luminosity conditions. The proposed approach performs the detection of vehicles through the analysis of the similarity of the objects [43]. Moreover, the proposed approach also performs the traffic flow classification through calculations with information on the highway and the number of identified vehicles. In addition to that, it can autonomously notify the drivers through the dashboard with the conditions of the roads. This allows drivers to choose an alternative path or delay the travel if they see fit.

The remainder of the work is structured as follows. Section 2 provides an overview of the related work. Section 3 describes the proposed mechanism for detecting and classifying the flows of vehicles on highways. The results of the proposed approach, as well as a comparison with other solutions found in the literature, are given in Section 4. The conclusions and future work are presented in Section 5.

2. Related Works

This section discusses several architectures found in the literature that use image and video processing to track and monitor the road's traffic.

Caballero et al. [44] proposed a mechanism based on images to detect vehicles as well as to verify and perform vehicle tracking over time. This solution is similar to the method proposed in [38], which describes a module for processing visual data images extracted from the scene. Both approaches use a temporal space analysis and high-level modules to perform the detection of vehicles and their attributes to acquire knowledge of traffic conditions. They can determine whether the vehicle actually left the scene as well as if it was overlapped or concealed by another vehicle. This work improves on several problems found in the identification of vehicles in images, such as the occlusion of vehicles; however, the method used in this work has a high cost of processing, which makes it not ideal for image processing in real time. The method used in this work does not perform the identification of vehicles and traffic flow classification in nocturnal or unfavorable climatic conditions.

Bhaskar et al. [45] use background subtraction techniques to identify and track vehicles. In each frame, Gaussian Mixture Model (GMM) techniques are applied to perform the differentiation of moving objects from the background of the image. To this end, the GMM technique uses a reference image and performs the comparison with a given sequence of images to identify the background. After identifying and modeling the background image, blob detection techniques are applied to identify groups of pixels representing moving vehicles. This type of technique may not be highly effective in certain environmental circumstances, such as dark environments, possible shadows of vehicles and trees, and scenarios where the camera is affected by vibration.

Xia et al. [40] presented a solution based on virtual ties to improve the quality of the vehicle counting method. The authors used the base station to improve vehicle motion detection by fusing the GMM with the Expectation Maximization (EM) algorithm for better tracking results. The combined use of these algorithms helped to improve the segmentation quality of the moving vehicles. Moreover, the authors developed a restoration method for noise suppression and filling

of cracks to obtain the best region of the object. Furthermore, the author also used a morphological feature and color histogram to solve obstruction problems. The combination of these improvements allowed some problems encountered in the background subtraction technique to be avoided, such as vehicle shadowing and unfavorable environmental conditions. However, the noise-filling technique ends up unduly joining vehicles when they are very close, either in the frontal, lateral or rear region, causing two or more vehicles to be identified as one.

Barcellos et al. [46] proposed a system to detect and perform the counting of vehicles in urban traffic videos with virtual links of defined users. The approach used motion linking and spatial adjacency to group particle samplings into urban video sequences. Furthermore, the author also used GMM and motion energy images to determine the ideal locations for particles to remove the samples and groups of convex particles are then analyzed for vehicle detection. The vehicles are tracked using the similarity of its colors in adjacent frames and counted in user virtual loops defined by intersection detection of vehicle tracking. The approach used by the author achieves good results in the identification of congestion on the highways since it can identify when the vehicle enters and leaves the virtual loop. However, it is inefficient in real-time analysis, since the areas of virtual ties need to be calculated, which makes the processing time of this method relatively high.

Sarikan et al. [47] proposed an image processing and machine learning vehicle classification technique. The authors adopted filtering techniques and image morphology to fill the regions of the vehicle, giving a better representation of the figure. A classification model was generated with images of cars and motorcycles to identify the vehicles, which, after being combined and normalized in the training set, after that, a heat map (HM) is generated and used to compare the similarity in other images. Decision trees were used to identify whether the object is a vehicle, a motorcycle, or is not a vehicle. This work presents an accurate technique for identifying different types of vehicles, which are, in this case, only motorcycles and sedans. This algorithm presents a vehicle identification technique, where it is possible to differentiate between motorcycles and sedans. The main disadvantage of this algorithm is that it only identifies motorcycles and sedan cars. For identification of smaller vehicles such as small cars and tricycles and large vehicles such as trucks and buses, other training data sets should be used. Another drawback of this algorithm is that it only performs the identification and classification of vehicles and does not classify the current traffic flow of the highway.

Table 1 gives a comparative summary of the above-mentioned related work and our proposed architecture.

Table 1. Related work comparison.

Article	Technique	Luminosity Identification	Traffic Flow Classification
Caballero et al. [44]	Background Subtraction	No	No
Bhaskar et al. [45]	Background Subtraction	No	No
Xia et al. [40]	Background Subtraction	No	No
Barcellos et al. [46]	Background Subtraction	No	No
Sarikan et al. [47]	Similarity (Decision Trees)	No	No
ATRIP	Similarity (Adaboost Classifier)	Yes	Yes

As seen, there are several proposed mechanisms; however, most of these solutions do not take into account the environmental settings, i.e., they lack the consideration of the luminosity of the ambient, such as nightfall, cloudy weather, or rain. Another point is that these solutions do not assess the traffic conditions as they only count the number of vehicles on the road. Therefore, the proposed architecture improves on these issues, by detecting, counting, and classifying the traffic conditions in scenarios with a wide range of luminosity.

3. Towards the Design of ATRIP

This section describes the proposed architecture for the detection and classification of traffic conditions, called ATRIP: Architecture for TRaffic classification based on Image Processing. The proposed architecture uses computer vision concepts to detect and classify vehicular traffic in a determined area. One of the advantages of the proposed mechanism is to use the already deployed cameras infrastructure along the pathways. It is a common practice nowadays for modern highways to have nearly full video surveillance. In any case, if there is no complete coverage, at least the important sections will probably be covered. ATRIP takes advantage of these traits as it uses Image Processing of real datasets, as well as real-time videos made available by companies that manage public roads such as AutoBAN.

Therefore, ATRIP provides information to the drivers through its web interface, so they can take action concerning how the traffic conditions are. In doing that, it will be possible to avoid aggravating even further traffic congestions.

The main advantages of ATRIP are the possibility to be easily integrated to any available ITS in place and also use the already installed surveillance cameras. This helps to cut down the system's deployment costs as well as keeping the maintenance costs down as the cameras are used for both surveillance and traffic control. Additionally, ATRIP is based on the client and server paradigm for both uploading and processing the information. This means that the system is reachable anywhere that is connected to the Internet network.

ATRIP is divided into three main components, aimed at visualizing image classification, preprocessing, and extracting information from images. The capture of videos can be done in two ways. The first one is to capture in real time the road footage sequences from monitoring platforms made available by the concessionaire who manages the highway. The second one is through videos obtained from the Internet where the systems have the traffic of the highways for a period of time.

3.1. Architecture Overview

As previously mentioned, the ATRIP architecture is divided into three main components: (i) information monitoring and abstraction; (ii) information storage and processing; and (iii) control room dashboard, as seen in Figure 1.

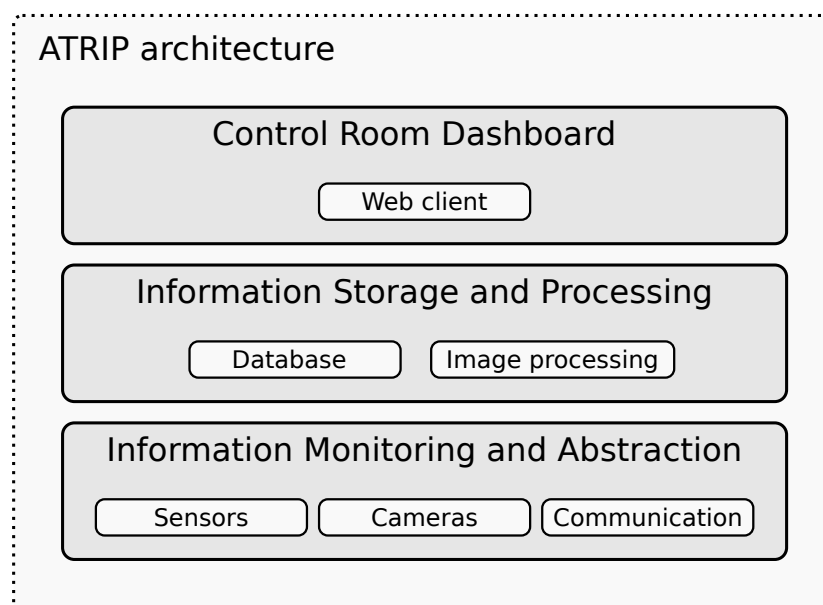


Figure 1. Overview of ATRIP architecture.

In a bottom-up order, the first component, (i) information monitoring and abstraction, communicates directly with the surveillance cameras and luminosity sensors. The adoption of this sensor is important as it makes the image processing faster. In general, it also provides more accurate results because it gives a precise measurement of the brightness of each spot where the cameras are located. This information is used as input for selecting the most appropriate traffic detection mechanism.

The second component, (ii) information storage and processing, is the proposed architecture main engine. It is responsible for receiving all the information such as the video frames from the surveillance cameras and the current light conditions. Another important task of this component is to process and classify all the received information.

The third component, (iii) control room dashboard, is the output interface of the system. This component allows for monitoring the routes, checking the output of the classification and traffic conditions. It was implemented as a web client, making it accessible anywhere and also providing on-screen alerts.

A detailed explanation of the above-mentioned components can be found below.

3.2. Information Monitoring and Abstraction

This layer is responsible for the management of public-road devices, such as cameras and sensors, besides communication between devices and a central server. The surveillance cameras placed in public locations, such as streets, avenues, and highways, are essential to capture real road footage sequences and transmit such footage to a central entity. Considering that this video footage can be accessed by the web, through a streaming mechanism.

Another important device in ATRIP architecture is the ambient light sensor. Such a sensor is applied to sense the brightness of its surroundings, that is, estimate the ambient light conditions of the environment where the sensor is currently operating.

Aiming to reduce the image processing time, and, consequently, the traffic classification time, an ambient light sensor placed close to the cameras was used to capture the local luminosity. Adding the luminosity captured in the training phase of the images helps to lower the computational cost, this will be explained in detail in Section 3.3.1. The communications between the sensor and the Information Storage and Processing module are made through the cellular network. It is worth mentioning that communication is only done when there is a change in the brightness of the surroundings.

3.3. Information Storage and Processing

This component is responsible for several procedures such as store and disseminates information in a scalable way, identify a vehicle through a video sequence, count the number of vehicles that are traveling within a monitored route, as well as classify and propagate this information to the drivers.

In other words, first, the vehicle identification training phase is described, detailing the adopted methods and how the training datasets are created. After that, the preprocessing of the real video frames as well as vehicle identification and counting are disclosed. Subsequently, the processing and classification of the live traffic condition are explained, as well as disseminating how this information is conveyed to users through the web interface.

3.3.1. Vehicle Identification Training Phase

ATRIP training phase is performed with real video sequences extracted from surveillance cameras. First of all, after the videos are gathered, they have to be converted to a series of static images. Once the static frames are obtained, the image processing engine will detect and count the vehicles in each one of them. The methods of getting the videos and transforming them into still images are trivial and will not be discussed in this work.

The identification and counting of the vehicles were performed based on the technique proposed by Viola-Jones [43], where vehicle identification was based on the main characteristics of vehicles

derived from a machine learning algorithm. These characteristics are responsible for realizing the differentiation between vehicles and other objects and shapes that may appear in the pictures. The use of Viola-Jones is due to the need for quick identification of the objects on stage. This technique is based on the characteristics of Haar [48], where rectangular features are used for the identification of objects.

According to Viola-Jones [43], the result found by the characteristic for a certain region is given by the sum of the pixels that are in the white rectangles that are subtracted by the sum of the pixels that they find in the gray rectangles. After the calculation of the rectangular region is obtained, the values of the Integral Image are obtained; with this, the classifier can traverse the image searching regions with the same patterns of characteristics of the desired object. AdaBoost is a machine learning method that combines several weak classifiers (weak learners–weak hypothesis) to obtain strong classifiers (i.e., a strong hypothesis). Boosting is used to select a feature set and train the classifier.

During the training phase, the rectangular features are located and analyzed, to determine whether or not they are useful to the classifier. When the features are applied to an image, the natural contrasts provided by the object features are examined, with regard to their spatial relations.

The objective of this mechanism is to carry out the detection and abstraction of the features of the vehicles. In order to be able to perform the detection of the vehicles, it is important to extract the characteristics of them. This process will be also responsible for differentiating a vehicle from other objects. To accomplish this, a set of characteristics found in the vehicles that are distinct concerning the sets of characteristics found in other objects are isolated.

For the machine learning algorithm to be executed, it is necessary to construct an image base that provides a training set that includes images of the object of interest, defined as positive images, and a set that does not contain the object, defined as negative images.

Two image datasets were created in which the first one has images with higher luminosity and the second one with less luminosity. Each dataset has the number of positive and negative images, the number of these images can influence the classification of the object [38,49]. Therefore, 5690 images were used for the training phase of high luminosity scenarios. Among these images, 2370 ones were positive and 3320 were negative images, as shown in Figure 2.

For the low luminosity environments, 962 images were used, including 193 positive and 769 negative images, as depicted in Figure 3. Therefore, the amount of images of both environments sum up to 6652. All the images of the dataset were handled manually by the authors of this work, selecting only the relevant objective in each image, i.e., tracing cars in the images to improve the accuracy of the classifier.

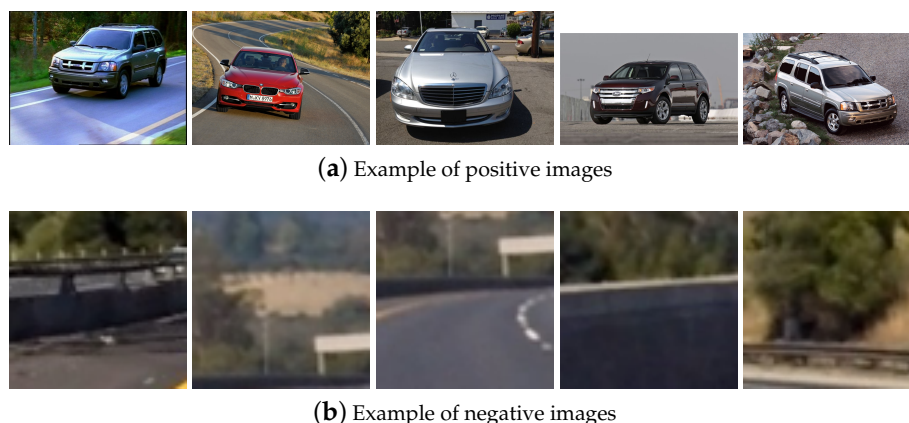


Figure 2. Training samples images (daytime scenario).

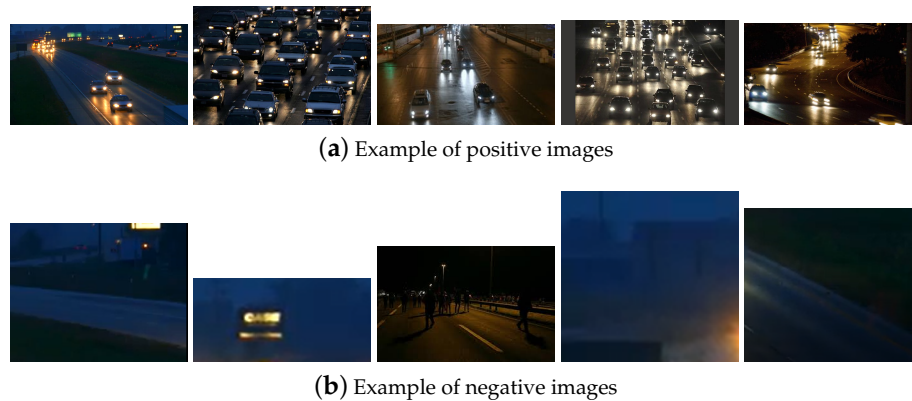


Figure 3. Training samples images (nighttime scenario).

3.3.2. Flow of Vehicles Classification

First, ATRIP captures the video segments from surveillance cameras or a real video dataset. These segments contain information about the last seconds of the highway's traffic status. Seeking to produce the desired output, the obtained data need to go through a preprocessing phase that comprises a set of steps. These steps are the same ones as in the aforementioned training phase. In other words, they are converted to a collection of static images; after that, the image processing engine detects and counts the vehicles.

Once the vehicles are detected, the mechanism counts the vehicles that are on the scene to classify the traffic flow on the road. For this, a predetermined threshold is used to characterize the flow of vehicles. It classifies the flow into three levels, namely free, moderate, and congested. This classification is based on the work [50]. However, the threshold value is very peculiar depending on each road configuration. Because of that, the values used here are based on the observation of a traffic engineering specialist.

Algorithm 1 describes the proposed classification, including the number of vehicles that are on the scene within a certain period on a given highway. Through the specialist assessment, the following thresholds were then found. It was defined that congestion is when the number of vehicles identified is above 20 (upper bound) in a single image, free when it is smaller than 12 (downbound), and moderate when the number is between 12 to 20. It is important to stress that these values are specific for the assessed situations; in the case of new scenarios, these thresholds would have to be revised.

Algorithm 1 Vehicles identification and traffic classification

```

1: Image ← getImage()
2: Light ← getLuminosity()
3: if (Light) then
4:   Base ← CarsDay
5: else
6:   Base ← CarsNight
7: end if
8: numberCars ← DetectVehicles(Image, Base)
9: if (numberCars > uperBound) then
10:  Return(Congestion)
11: else if (numberCars < downBound) then
12:  Return(Free)
13: else
14:  Return(Moderate)
15: end if

```

The result of both vehicle counting on the scene and the classification of traffic flows from the road is passed to the user interface so that the users can track traffic conditions in real time.

3.4. Control Room Dashboard

The last component of the ATRIP architecture is the control room panel. This is the main interface between the system and users as it can be used by the drivers to learn about the current traffic flow of the highways and the access points. With this information, it is possible to provide a better decision-making process, which can lead to avoiding highways with the greater traffic flow, and thus reduce congestion. The Web client is able to show the video sequences in real time. It can also display the results of the traffic counting and the current flow of the highway, as shown in Figure 4.

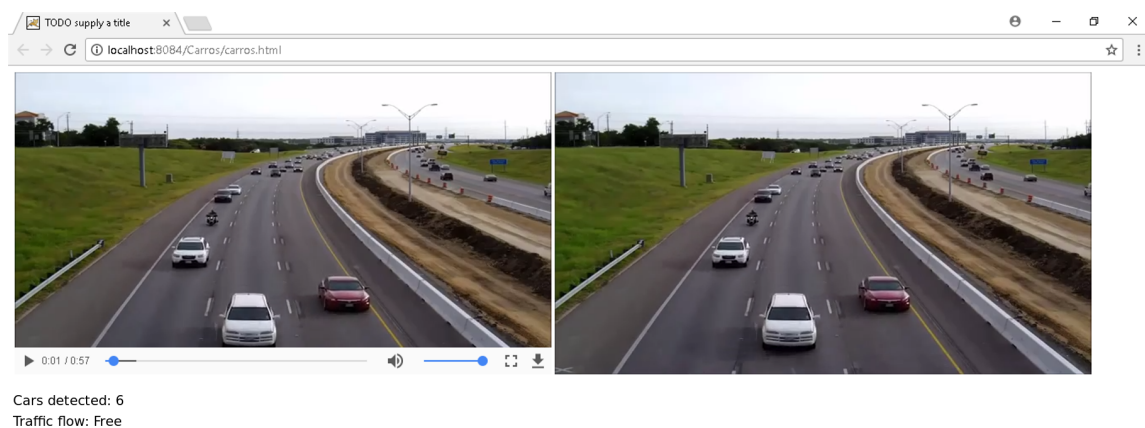


Figure 4. Web Client control room dashboard.

4. ATRIP Assessment

This section presents the performance evaluation of the proposed architecture through a set of experiments. The main component to be assessed is the Information Storage and Processing layer. It was implemented in Matlab version R2015A (The MathWorks, Inc., Natick, MA, USA) which has the functionalities needed to apply the computer vision techniques required to extract the information from videos. In the experiments, 70 videos were collected from the streaming provided by the AutoBAN company. Therefore, a real dataset with actual video footage of the highway traffic was used to train, classify, and validate the proposed architecture. A highway segment (Rodovia Anhanguera SP-330) was chosen for the performance evaluation. Such a highway is one of the most important in the state of São Paulo and connects Cordeirópolis, Campinas, Jundiaí, and São Paulo.

The main goal of the experiments is to evaluate the vehicle's detection and the traffic classification in both night and daytime environments. The performance evaluation for vehicle detection was compared against the Bhaskar et al. [45], called the Bhaskar method here. This method has been proven to have a high level of correct classification of objects with little false positive and false negative, as well as fast response time.

Furthermore, the output of the traffic flow classification was compared with the appraisal of a specialist in traffic engineering. This specialist observed the video frames used by the proposed architecture and performed his personal flow classification.

4.1. Detection of Vehicles

In this experiment, three factors that have a direct effect on the number of vehicles were evaluated, which are:

- The number of vehicles detected: represents the number of vehicles correctly detected;

- The number of false negatives: corresponds to the number of vehicles that were not detected although they were traveling on the highway;
- The number of false positives: gives the number of cases in which the objects were erroneously identified as vehicles.

The videos used in the experiments included several distinct traffic conditions. Moreover, a total of 35 videos were used for the nighttime environment and another 35 videos for the daytime one.

Figure 5a shows the vehicle detection results of the daytime environments. The figure presents the percentage of vehicles identified correctly. It is possible to notice that the proposed architecture has a better performance than our competitor, the Bhaskar method. This improvement in the detection rate can be attributed to the test group and training method used to classify the vehicles. As a result, ATRIP had an increase of approximately 10% in comparison to the other mechanism. This increase in detection is due to a reduction in the number of false positives and false negatives, as described below.

Figure 5c depicts the percentage of false negatives, i.e., the percentage of cases where the methods failed to correctly identify the vehicles. It can be observed in the graphs that the proposed method obtained a smaller number of false negatives in comparison to the Bhaskar method. This happens because the Bhaskar method was sometimes unable to identify the vehicles that were close to each other. This means that when a car is behind or very close to another one it is not correctly identified because of the image segmentation. Therefore, the ATRIP architecture reduced the number of false negatives compared to the Bhaskar method by approximately 11%.

Figure 5e gives the percentage of false-positive, i.e., the number of vehicles that were incorrectly identified. In other words, non-vehicles objects that were classified as vehicles. Here, again, ATRIP obtained a small improvement over its competitor. The percentage of false positives suffered a reduction of approximately 2%.

Taking everything into consideration, the proposed architecture was able to achieve an overall improvement of approximately 10% in high luminosity environments. This meant that a 92% accuracy rate was achieved. This was only possible because of the low number of both false negatives (approximately 8%) and false positives (approximately 1%).

Additionally, the ATRIP architecture was also assessed in low light conditions. In this condition, it is harder to correctly identify the vehicles. Figure 5b shows the percentage of vehicles correctly identified in videos with low luminosity. It is possible to notice that the proposed architecture has a better performance than the Bhaskar method. This can be explained by the difficulty experienced by the Bhaskar method when separating the background from the tracked object. In some cases, the image was too bright and in other ones was not bright enough. ATRIP improves on this issue and provided an increase of approximately 6% better than the competitor. This increase in the detection can be attributed to the reduction in the number of both false positives and false negatives, as can be seen in Figure 5d,f.

Figure 5d depicts the percentage of false negatives in a nighttime environment. In this scenario, the results reveal that the method obtained a smaller number of false negatives than the Bhaskar method. This happens because the competitor method is sometimes unable to identify the vehicles due to the dark background. As a result, ATRIP was able to reduce the number of false negatives by approximately 8%.

The percentage of false positives obtained by the video images captured at low light conditions is depicted in Figure 5f. The results follow a similar pattern as seen in the false-negative cases. The same as before, the ATRIP was able to reach an improvement of approximately 2% in comparison to the Bhaskar method. This can be attributed to the image database used by the proposed architecture in the training phase of the classification mechanism.

In the light of the aforementioned results, it is possible to conclude that ATRIP obtained an increase of 8% in its classification rate, in a low light environment, against the Bhaskar method. Moreover, it was able to achieve a 31% improvement on the accuracy rate as it reaches a low number of both false

negatives (approximately 34%) and false positives (approximately 35%), again, in comparison with the Bhaskar method.

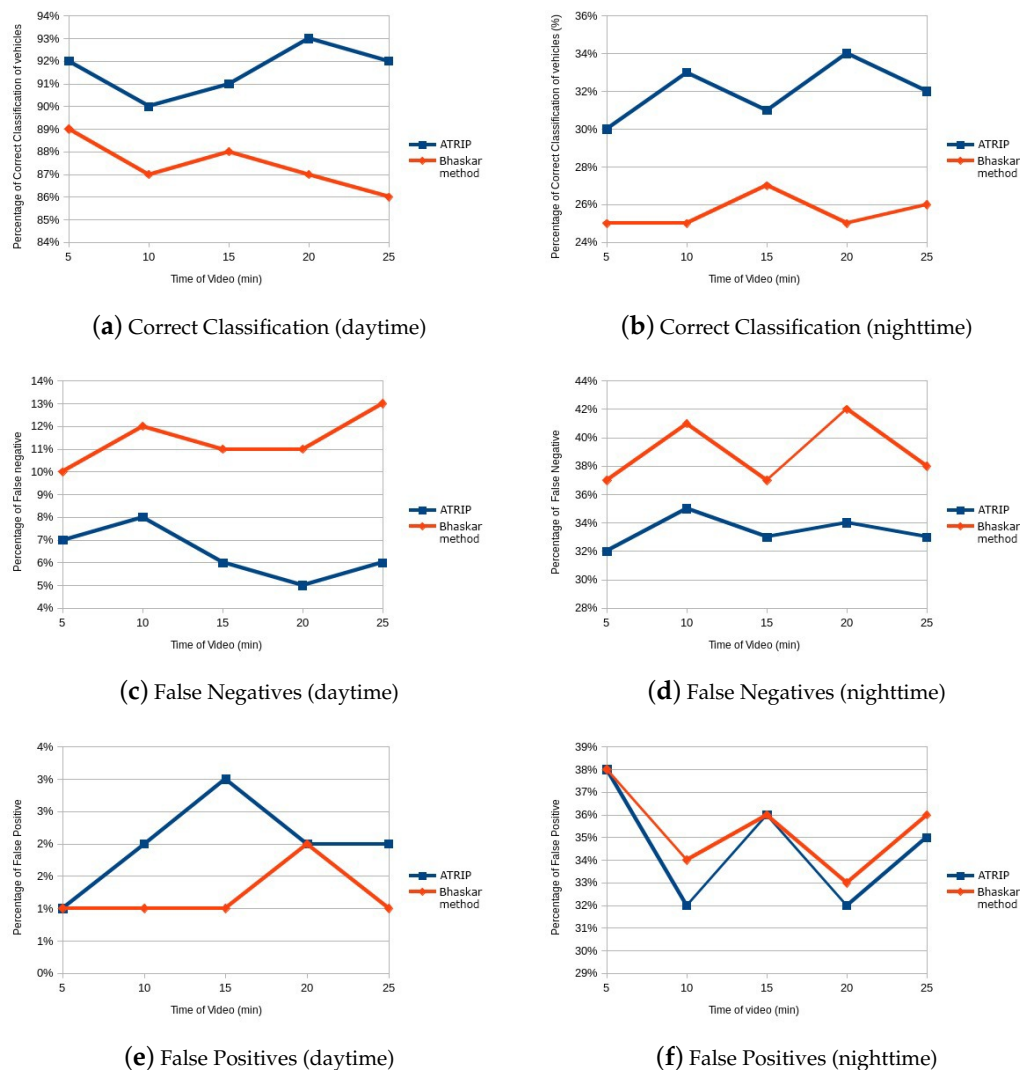


Figure 5. Experiment results of the proposed architecture in comparison to the Bhaskar method.

4.2. Traffic Classification

This section compiles the results obtained regarding the ATRIP traffic classification methods. These experiments aim to determine whether the proposed architecture can correctly classify the flows of vehicles on the highway. In order to do that, the degree of correctness of the classification performed by API needs to be assessed and compared with the results obtained from a human specialist who analyzed the same videos used. Several videos with distinct traffic conditions were assessed by both the proposed architecture and the human specialist. In the experiments, 35 videos were used for the nighttime experiments and another 35 videos for the daytime conditions, 70 videos in total.

Figure 6 shows the percentage of correct classifications of the traffic flow on the highway. The correct classifications mean that the proposed architecture output was the same as the human specialist opinion. The results demonstrated that ATRIP provides an accurate classification of the levels of traffic congestion. This is especially true in the case of pinpointing traffic congestion in the daylight environments, where it reached about a 90% accuracy level. On the other hand, the classification accuracy is lower in low light environments. However, this situation was expected as it is rather difficult to correctly detect the vehicles in these situations.

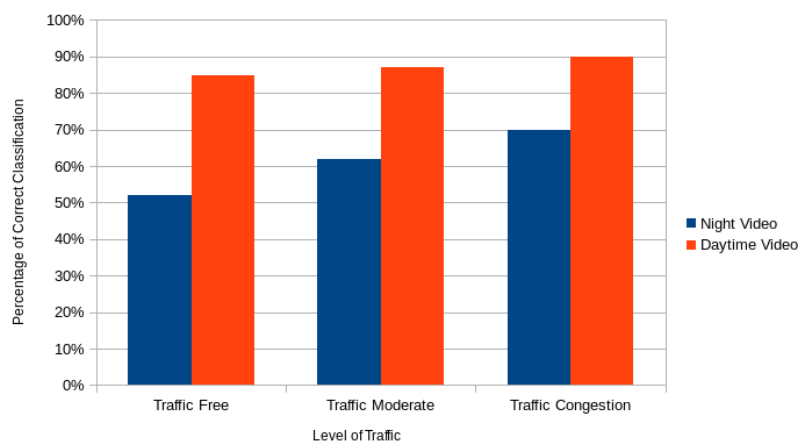


Figure 6. Percentage of correct classifications.

5. Conclusions

This information can be used reactively, making them available so that drivers can analyze the route, avoid congested roads, or be proactive in providing input information to the development of better traffic solutions by the city planners. Aiming to improve on this issue, this work proposed the ATRIP architecture which provides accurate vehicle detection and traffic classification, which can be used together with the ITS in place. The proposed architecture provides a holistic tool, being responsible for the video image retrieval and the preprocessing of the videos as well as the traffic classification up to a control room dashboard.

The results showed that, for an environment with daylight conditions, ATRIP obtained results with 10% better accuracy than the competitor. It was also able to achieve a 92% accuracy rate in vehicle detection. Additionally, it produced a very low number of both false negatives and false positives, approximately 8% and 6%, respectively. Furthermore, it also had a better performance in scenarios with low light conditions, improving the detection rate up to 8%. It also produced approximately 34% less false negatives and 35% less false positives. These results led to a traffic classification rate accuracy of around 90% in daylight environments and 70% in low light environments.

The advancement of the proposed solution in the identification of vehicles when compared to the competitor is clear. It is worth mentioning that ATRIP still has a low rate of accuracy in the traffic classification process during the night or in low-luminosity scenarios. One of the main reasons for this lack of success is the low amount of low-luminosity environment images used to train the classifier. It is known that the ATRIP uses only information about the highway and the number of vehicles to identify the current congestion status. Another piece of data that could be useful during the congestion identification process is vehicle speed because considering data regarding vehicle speed can increase the used classifier's performance.

As future work, we intend to conduct further research on other methods that can be employed to tackle other issues such as vehicle occlusion and vehicle speed to set the highway congestion status.

Author Contributions: Conceptualization, A.L.C., R.I., A.T.A., E.R.M.M., L.A.V. and R.I.M.; Investigation, A.L.C., R.I. and R.I.M.; Methodology, A.L.C., R.I., A.T.A., E.R.M.M., L.A.V. and R.I.M.; Supervision, R.I.M.; Validation, A.L.C., R.I. and R.I.M.; Writing—original draft, A.L.C., R.I. and R.I.M.; Writing—review & editing, R.I., A.T.A., E.R.M.M., L.A.V. and R.I.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by the grant 2018/02204-6 from São Paulo Research Foundation (FAPESP), and it is part of the INCT project called the Future Internet for Smart Cities (CNPq 465446/2014-0, CAPES 88887.136422/2017-00 and FAPESP 2014/50937-1). Rodolfo I. Meneguette would like to thank CNPq (process 407248/2018-8 and 309822/2018-1) for funding the research project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. US Department of Transportation (DOT). *National Transportation Statistics 2019*; Technical Report; 2019. Available online: <https://www.bts.gov/tsar> (accessed on 11 April 2020).
2. Meneguette, I.R.; De Grande, E.R.; Loureiro, A.F.A. Autonomous Vehicles. In *Intelligent Transport System in Smart Cities: Aspects and Challenges of Vehicular Networks and Cloud*; Springer International Publishing: Cham, Switzerland, 2018; pp. 37–52. [\[CrossRef\]](#)
3. Vourgidis, I.; Maglaras, L.; Alfakeeh, A.S.; Al-Bayatti, A.H.; Ferrag, M.A. Use of Smartphones for Ensuring Vulnerable Road User Safety through Path Prediction and Early Warning: An In-Depth Review of Capabilities, Limitations and Their Applications in Cooperative Intelligent Transport Systems. *Sensors* **2020**, *20*, 997. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Alam, M.; Ferreira, J.; Fonseca, J. Introduction to Intelligent Transportation Systems. In *Intelligent Transportation Systems: Dependable Vehicular Communications for Improved Road Safety*; Alam, M., Ferreira, J., Fonseca, J., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 1–17. [\[CrossRef\]](#)
5. Hernandez-Jayo, U.; De-la Iglesia, I.; Perez, J. V-Alert: Description and Validation of a Vulnerable Road User Alert System in the Framework of a Smart City. *Sensors* **2015**, *15*, 18480–18505. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Akabane, A.T.; Immich, R.; Bittencourt, L.F.; Madeira, E.R.; Villas, L.A. Towards a distributed and infrastructure-less vehicular traffic management system. *Comput. Commun.* **2020**, *151*, 306–319. [\[CrossRef\]](#)
7. Kim, M.; Kim, H.K.; Lee, S.H. A Distributed Cooperative Localization Strategy in Vehicular-to-Vehicular Networks. *Sensors* **2020**, *20*, 1413. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Akabane, A.T.; Immich, R.; Pazzi, R.W.; Madeira, E.R.M.; Villas, L.A. Distributed Egocentric Betweenness Measure as a Vehicle Selection Mechanism in VANETs: A Performance Evaluation Study. *Sensors* **2018**, *18*, 2731. [\[CrossRef\]](#) [\[PubMed\]](#)
9. An, H.; Jung, J.I. Design of a Cooperative Lane Change Protocol for a Connected and Automated Vehicle Based on an Estimation of the Communication Delay. *Sensors* **2018**, *18*, 3499. [\[CrossRef\]](#)
10. Akabane, A.T.; Immich, R.; Madeira, E.R.M.; Villas, L.A. iMOB: An Intelligent Urban Mobility Management System Based on Vehicular Social Networks. In Proceedings of the 2018 IEEE Vehicular Networking Conference (VNC), Taipei, Taiwan, 5–7 December 2018; pp. 1–8. [\[CrossRef\]](#)
11. Padilla, D.A.; Villaverde, J.F.; Magdaraog, J.J.T.; Oconer, A.J.L.; Ranjo, J.P. Vehicle and Weather Detection Using Real Time Image Processing Using Optical Flow and Color Histogram. In Proceedings of the 2019 5th International Conference on Control, Automation and Robotics (ICCAR), Beijing, China, 19–22 April 2019; pp. 880–883.
12. Immich, R.; Cerqueira, E.; Curado, M. Efficient high-resolution video delivery over VANETs. *Wirel. Netw.* **2018**. [\[CrossRef\]](#)
13. Gerát, J.; Sopiak, D.; Oravec, M.; Pavlovicová, J. Vehicle speed detection from camera stream using image processing methods. In Proceedings of the 2017 International Symposium ELMAR, Zadar, Croatia, 18–20 September 2017; pp. 201–204.
14. Immich, R.; Cerqueira, E.; Curado, M. Towards a QoE-driven mechanism for improved H.265 video delivery. In Proceedings of the Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net), Vilanova i la Geltru, Spain, 20–22 June 2016; pp. 1–8. [\[CrossRef\]](#)
15. Quadros, C.; Cerqueira, E.; Neto, A.; Riker, A.; Immich, R.; Curado, M. A mobile QoE Architecture for Heterogeneous Multimedia Wireless Networks. In Proceedings of the 2012 IEEE Globecom Workshops, Anaheim, CA, USA, 3–7 December 2012; pp. 1057–1061. [\[CrossRef\]](#)
16. Colombaroni, C.; Fusco, G.; Isaenko, N. A Simulation-Optimization Method for Signal Synchronization with Bus Priority and Driver Speed Advisory to Connected Vehicles. *Transp. Res. Procedia* **2020**, *45*, 890–897. [\[CrossRef\]](#)
17. Hussain, S.A.; Shahian Jahromi, B.; Cetin, S. Cooperative Highway Lane Merge of Connected Vehicles Using Nonlinear Model Predictive Optimal Controller. *Vehicles* **2020**, *2*, 249–266. [\[CrossRef\]](#)
18. Simchon, L.; Rabinovici, R. Real-Time Implementation of Green Light Optimal Speed Advisory for Electric Vehicles. *Vehicles* **2020**, *2*, 35–54. [\[CrossRef\]](#)
19. Wang, P.; Gao, S.; Li, L.; Sun, B.; Cheng, S. Obstacle Avoidance Path Planning Design for Autonomous Driving Vehicles Based on an Improved Artificial Potential Field Algorithm. *Energies* **2019**, *12*, 2342. [\[CrossRef\]](#)

20. Viana, B.; Aouf, N. Distributed Cooperative Path-Planning for Autonomous Vehicles Integrating Human Driver Trajectories. In Proceedings of the 2018 International Conference on Intelligent Systems (IS), Funchal, Portugal, 25–27 September 2018; pp. 655–661.
21. Chen, C.; Seff, A.; Kornhauser, A.; Xiao, J. DeepDriving: Learning Affordance for Direct Perception in Autonomous Driving. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015.
22. Andrade, F.R.d.; Antunes, J.L.F. Trends in the number of traffic accident victims on Brazil's federal highways before and after the start of the Decade of Action for Road Safety. *Cad. De Saude Publica* **2019**, *35*, e00250218. [[CrossRef](#)] [[PubMed](#)]
23. Confederaçao Nacional do Transporte. *Consultas Dinamicas de Acidentes Rodoviarios*; Confederaçao Nacional do Transporte: Brasília, Brazil, 2019.
24. de Souza, A.M.; Yokoyama, R.; Boukerche, A.; Maia, G.; Cerqueira, E.; Loureiro, A.A.; Villas, L.A. ICARUS: Improvement of traffic Condition through an Alerting and Re-routing System. *Comput. Netw.* **2016**, *110*, 118–132. [[CrossRef](#)]
25. Hankey, S.; Lu, T.; Mondschein, A.; Buehler, R. Spatial models of active travel in small communities: Merging the goals of traffic monitoring and direct-demand modeling. *J. Transp. Health* **2017**. [[CrossRef](#)]
26. Seenoung, N.; Watchareeruetai, U.; Nuthong, C.; Khongsomboon, K.; Ohnishi, N. A computer vision based vehicle detection and counting system. In Proceedings of the 2016 8th International Conference on Knowledge and Smart Technology (KST), Chiangmai, Thailand, 3–6 February 2016; pp. 224–227. [[CrossRef](#)]
27. Meshram, S.A.; Chintalacheruvu, N.; Muthukumar, V.; Kastrinaki, V.; Zervakis, M.; Kalaitzakis, K.; Hadi, R.A.; Sulong, G.; Meshram, S.A.; Malviya, A.V.; et al. Survey on Vehicle Detection and Tracking Techniques in Video Surveillance. *Int. J. Comput. Appl.* **2017**, *160*, 22–25.
28. Frank, A.; Khamis Al Aamri, Y.S.; Zayegh, A. IoT based Smart Traffic density Control using Image Processing. In Proceedings of the 2019 4th MEC International Conference on Big Data and Smart City (ICBDSC), Muscat, Oman, 15–16 January 2019; pp. 1–4.
29. Akabane, A.T.; Immich, R.; Pazzi, R.W.; Madeira, E.R.M.; Villas, L.A. TRUSTed: A Distributed System for Information Management and Knowledge Distribution in VANETs. In Proceedings of the 2018 IEEE Symposium on Computers and Communications (ISCC), Natal, Brazil, 25–28 June 2018; pp. 1–6.
30. Wang, X.; Liu, C.; Zhu, M. Instant Traveling Companion Discovery Based on Traffic-Monitoring Streaming Data. In Proceedings of the 2016 13th Web Information Systems and Applications Conference (WISA), Wuhan, China, 23–25 September 2016; pp. 89–94. [[CrossRef](#)]
31. Immich, R.; Cerqueira, E.; Curado, M. Adaptive QoE-driven video transmission over Vehicular Ad-hoc Networks. In Proceedings of the IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Hong Kong, China, 26 April–1 May 2015; pp. 227–232. [[CrossRef](#)]
32. D'Andrea, E.; Marcelloni, F. Detection of traffic congestion and incidents from GPS trace analysis. *Expert Syst. Appl.* **2017**, *73*, 43–56. [[CrossRef](#)]
33. Thiagarajan, A.; Ravindranath, L.; LaCurts, K.; Madden, S.; Balakrishnan, H.; Toledo, S.; Eriksson, J. VTrack: Accurate, energy-aware road traffic delay estimation using mobile phones. In Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems, Berkeley, CA, USA, 4 November 2009; pp. 85–98.
34. Ma, X.; Dai, Z.; He, Z.; Ma, J.; Wang, Y.; Wang, Y. Learning Traffic as Images: A Deep Convolutional Neural Network for Large-Scale Transportation Network Speed Prediction. *Sensors* **2017**, *17*, 818. [[CrossRef](#)]
35. Olaverri-Monreal, C.; Krizek, G.C.; Michaeler, F.; Lorenz, R.; Pichler, M. Collaborative approach for a safe driving distance using stereoscopic image processing. *Future Gener. Comput. Syst.* **2019**, *95*, 880–889. [[CrossRef](#)]
36. Khan, S.D.; Ullah, H. A survey of advances in vision-based vehicle re-identification. *Comput. Vis. Image Underst.* **2019**, *182*, 50–63. [[CrossRef](#)]
37. Tao, H.; Lu, X. Smoke vehicle detection based on multi-feature fusion and hidden Markov model. *J. Real Time Image Process.* **2019**, 1–14. doi:10.1007/s11554-019-00856-z [[CrossRef](#)]
38. Cucchiara, R.; Piccardi, M.; Mello, P. Image analysis and rule-based reasoning for a traffic monitoring system. *IEEE Trans. Intell. Transp. Syst.* **2000**, *1*, 119–130. [[CrossRef](#)]
39. Tai, J.C.; Tseng, S.T.; Lin, C.P.; Song, K.T. Real-time image tracking for automatic traffic monitoring and enforcement applications. *Image Vis. Comput.* **2004**, *22*, 485–501. [[CrossRef](#)]

40. Xia, Y.; Shi, X.; Song, G.; Geng, Q.; Liu, Y. Towards improving quality of video-based vehicle counting method for traffic flow estimation. *Signal Process.* **2016**, *120*, 672–681. [CrossRef]
41. AutoBAn, C. Live Stream. 2017. Available online: <http://www.autoban.com.br/ao-vivo> (accessed on 11 April 2020).
42. Cristiani, A. Dataset of Videos and Images. 2020. Available online: <http://rsd.ctd.ifsp.edu.br/wp-content/uploads/APIA/> (accessed on 11 April 2020).
43. Viola, P.; Jones, M.J. Robust Real-Time Face Detection. *Int. J. Comput. Vis.* **2004**, *57*, 137–154. [CrossRef]
44. Fernández-Caballero, A.; Gómez, F.J.; López-López, J. Road-traffic monitoring by knowledge-driven static and dynamic image analysis. *Expert Syst. Appl.* **2008**, *35*, 701–719. [CrossRef]
45. Bhaskar, P.K.; Yong, S.P. Image processing based vehicle detection and tracking method. In Proceedings of the 2014 International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, Malaysia, 3–5 June 2014; pp. 1–5.
46. Barcellos, P.; Bouvié, C.; Escouto, F.L.; Scharcanski, J. A novel video based system for detecting and counting vehicles at user-defined virtual loops. *Expert Syst. Appl.* **2015**, *42*, 1845–1856. [CrossRef]
47. Sarikan, S.S.; Ozbayoglu, A.M.; Zilci, O. Automated Vehicle Classification with Image Processing and Computational Intelligence. *Procedia Comput. Sci.* **2017**, *114*, 515–522, doi:10.1016/j.procs.2017.09.022. [CrossRef]
48. Viola, P.; Jones, M. Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), Kauai, HI, USA, 8–14 December 2001; pp. 511–518.
49. Tao, H.; Sawhney, H.S.; Kumar, R. Object tracking with Bayesian estimation of dynamic layer representations. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 75–89. [CrossRef]
50. Meneguette, R.I.; Fillho, G.P.R.; Bittencourt, L.F.; Ueyama, J.; Villas, L.A. A Solution for Detection and Control For Congested Roads Using Vehicular Networks. *IEEE Lat. Am. Trans.* **2016**, *14*, 1849–1855. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).