Article
A Road Behavior Pattern-Detection Model in Querétaro City Streets by the Use of Shape Descriptors

Antonio Trejo-Morales * and Hugo Jimenez-Hernandez *

Facultad de Informática, Universidad Autónoma de Querétaro, Av. de las Ciencias S/N, Juriquilla 76230, Querétaro, Mexico
* Correspondence: antonio.trj.mrls@gmail.com (A.T.-M.); hugojh@gmail.com (H.J.-H.)

Abstract: In this research, a proposed model aims to automatically identify patterns of spatial and temporal behavior of moving objects in video sequences. The moving objects are analyzed and characterized based on their shape and observable attributes in displacement. To quantify the moving objects over time and form a homogeneous database, a set of shape descriptors is introduced. Geometric measurements of shape, contrast, and connectedness are used to represent each moving object. The proposal uses Granger’s theory to find causal relationships from the history of each moving object stored in a database. The model is tested in two scenarios; the first is a public database, and the second scenario uses a proprietary database from a real scenario. The results show an average accuracy value of 78% in the detection of atypical behaviors in positive and negative dependence relationships.

Keywords: artificial intelligence; causality; extraction of information; inference of activities; intelligence system; temporal patterns

1. Introduction

The rapid growth of urban areas has led to the development of equipment that measures and analyzes information to make day-to-day decisions. This equipment and infrastructure are known as smart cities and require reliable communication and interoperability between technologies. The infrastructure measures urban variables and analyzes data to develop efficient, reliable, adaptable, and robust algorithms capable of identifying events and atypical behaviors of interest. This information is stored in a database using camera technologies, and historical trends can be analyzed over time. These data are useful for decision-making on events and generating video analytics on the activity of an area.

The increase in population density and the growth in vehicle use have led to a decrease in the standard of living for inhabitants due to the saturation of communication, transportation, and telecommunication routes. This has resulted in road congestion, making road monitoring systems more in demand. However, existing monitoring systems have general drawbacks that need to be addressed:

- Human operators are required to monitor day-to-day activities.
- Human intervention is required to locate the same object if multiple cameras are used.

The physical infrastructure of roads is sufficient for the collection of information on traffic flow and speed. However, since this method is invasive, it requires frequent maintenance to ensure proper functionality. Additionally, the presence of monitoring equipment may cause drivers to alter their driving behaviors [1]. To address this issue, a model based on Granger causal patterns can be used to efficiently monitor roadways and detect abnormal behaviors. A distributed monitoring system that utilizes conventional camera technology is chosen to identify the behavior of moving objects and situations of interest. This model can help identify periodic behaviors in traffic patterns and analyze the relationships between them. Currently, a monitoring system consists of cameras that
capture and store data in a database. These data include information such as the number of objects, their trajectories, size, color, and flow.

This article is divided into several sections. In Section 1.1, we provide information about the related work. In Section 2, we describe the materials and methods used in characterizing the behavior of the study scenario. In Section 3, we present the results. Section 4 concludes the article by discussing the model. In Section 5, we present the conclusion. Finally, we list the references that support this work.

1.1. Related Work

Smart cities are a new and growing concept that has attracted the interest of researchers, city authorities, and governments [2]. This offers a wide range of possibilities for research and industrial opportunities.

The issue of urban freight transport is not a new problem. Betanzo-Quezada and his team [3] carried out research on this topic in the Mexican city of Querétaro from 2003 to 2014. Their methodological approach included a multi-year research effort to create analytical tools, evaluation methods, and vehicle size categorization. They concluded that it is important to strengthen research on land use, economic trends, and freight activity characteristics. A similar study, focused on modeling and micro-simulation approaches to several loading/unloading bays, can be found in the references [4].

Trencher [5] conducted a study that provided evidence of how intelligence can be used to address social challenges [6]. The research examined the Aizuwakamatsu Smart City in Fukushima, Japan, to demonstrate how a smart city can be designed and implemented to meet the needs of its residents. The author suggests that social objectives should be formulated in response to social challenges and citizen needs, and that this is related to the development of apps and people-centered information communication technologies.

Wang et al. [7] presented data on the processing of traffic digital twins in smart cities using edge intelligent federation learning. This work utilized deep learning and digital twins (DTs) in the development of smart cities [8]. Edge computing technology was employed to build an intelligent traffic perception system based on edge computing combined with DTs. According to their experimental results, the SSD-ResNet50 and the improved DarkNet-53 algorithm showed fast training speeds, high recognition accuracy, and favorable training effects.

In their study, Amen, Afara, and Nia [9] examined the relationship between the centrality of street layout and walkability in promoting sustainable tourism in historic urban areas. The authors focused on the Turkish part of Nicosia’s old city, which is known for its narrow, winding streets, ancient stone houses, and lively markets. The study used digital maps for data collection and analysis. The results showed that betweenness centrality has a greater impact on tourist distribution than closeness centrality. This is because tourists tend to visit places with high betweenness centrality more frequently.

In a research paper published by Husain, Maity, and Yadav [10], a survey was conducted on vehicle detection in intelligent transport systems under hazy environmental conditions [11]. The study used cameras to capture data from an intelligent transport system (ITS) and reviewed various features such as traffic flow density, average velocity, and total vehicles passing through a point in a specific range of time. The authors also explored different technologies for data analysis, including machine learning, genetic algorithms, and recurrent neural networks. They concluded that weather conditions, illumination variability, and dynamic background scenes could affect the results of the detection algorithm [12].

In a research paper by Selvi and Amudha [13], they presented an automatic video surveillance system for pedestrian crossing using digital image processing [14]. The authors used a zebra crossing detection system to operate an intelligent vehicle and a self-similarity recognition method that yielded a 98.5% accuracy. The results suggest that this application can be used to track objects.
In the current era, computer vision research lines mention several approaches for vehicle classification and detection [15]. Thirumarai and Amudha [13] proposed a novel video surveillance technique with an image segmentation algorithm to track moving objects in a crosswalk of a road. They also applied several morphological filtering operations that improved the segmentation quality of moving objects in the video.

Shantaiya et al. [16] presented a multiclass image-based classification of vehicles using soft computing algorithms such as artificial neural network decision trees and support vector machines [17]. Gu et al. [18] proposed an online video object segmentation through boundary-constrained low-rank sparse representation. Their algorithm is based on a classical image partitioning algorithm (Graphcut, minimum of maximum current cut) [19].

Rawassizadeh et al. [20] developed a library of three algorithms for event detection in temporal spaces, clustering based on mobile contextual data, and identification of contrastive events within a cluster. Xiangyu et al. [21] developed a model that can produce automatic label assignments in images using label propagation. Their model is based on the theoretical foundation of the Bayesian conditional random field (BCRF) model.

Zambrano Martinez et al. [22] proposed an equation to model and characterize flow times in terms of vehicle load concerning travel time during peak traffic hours in urban environments.

Cities around the world are growing rapidly, but often in an uncontrolled way. This causes a negative impact on the quality of life, safety, and overall well-being of the population. Congestion and saturation of streets lead to car accidents, unplanned public investment costs, and general dissatisfaction among the people. Therefore, it is essential to measure urban variables accurately and develop efficient, reliable, adaptable, and robust algorithms to solve these problems. By analyzing typical road behaviors and unexpected events, we can propose a solution that can improve the situation. In this regard, researchers have applied causality techniques in various studies, such as Asuma-du-Sarkodie and Owusu [23], who implemented a research study in Kenya to examine the multivariate causality of carbon dioxide emissions [24,25]. They used a World Bank dataset spanning from the year 1961 to 2011 and the autoregressive distributed lag (ARDL) model for cointegration analysis [26]. On the other hand, Barnett and Seth [27] created a new method for Wiener–Granger causality inference, which avoids the explicit estimation of the standard Wiener–Granger causality model. This eliminates estimation errors and improves statistical power, while also facilitating fast and accurate estimation of the computationally cumbersome case of conditional Wiener–Granger causality in the frequency domain.

Currently, cities are growing in an undue way, thus affecting the quality of life and welfare of the population. The above is mainly due to the saturation and congestion of the streets, which results in car accidents, unplanned public investment costs, and generalized social dissatisfaction. In that sense, the measurement of urban variables, as well as the analysis and development of efficient, reliable, adaptable, and robust algorithms, becomes essential to developing a solution. The solution can be proposed in terms of data analyses of typical road behaviors and unexpected events.

1.2. Theoretical Background

Traffic image processing has attracted the attention of researchers in the recent past who have tried to use indirect methods for flow monitoring and behavioral analysis [28]. These approaches include non-invasive methods by computer vision [29], instrumentation using road pressure sensors, and the use of radar, to name the most commonly used [30], for pattern recognition, signal processing, communication, embedded computing, and image sensing, as well as object identification, multi-camera activity analyses and cooperative video surveillance with active and static cameras [10,31]. However, the amount of information generated by the geographical dispersion of sensors and the lack of computer and algorithmic infrastructures cause a lack of reliable and efficient criteria for decision-making.

The concept of causality, first formulated by Wiener [32] and Granger [33], has become a cornerstone theory in the analysis of dynamic relationships between variables.
This theory, which is used to determine the unidirectional, bidirectional, or independent causal relationship in terms of temporal precedence between two time series \{X, Y\}, has practical applications in traffic analysis. It is a statistical test based on the concept of cross-prediction [34], which can be used to forecast traffic patterns based on historical data, thereby aiding in traffic management and planning.

A binary digital image is a two-dimensional numerical matrix with two levels of gray hue at \{0, 1\}. In simpler terms, an image is a collection of connected dots, each representing an object in the picture [35]. From these dots, we can extract various features to describe the objects. These objects are sets of connected dots, or pixels, with a value of one [36–39].

On the other hand, the shape of objects is one of the most straightforward features to extract from a digital image [40], for which the basic contour-based and region-based methods are used [41]. The region-based method commonly implements moment descriptors, which include geometric moments, to name one. The contour-based method usually obtains the perimeter using methods based on the curvature of objects. The descriptors developed in different research range from basic simple shape descriptors such as perimeter, area, and circularity [42,43] to invariant descriptors such as Hu moments [44]. Fourier descriptors for contour recognition [45], Euler features [46] describing the structure of an image, and the Harris corner detector [47,48] used to extract and infer features in an image. These descriptive features are invariant to image transformations, such as translations, rotations, scale changes, and projections, and robust to changing motion and illumination conditions [35].

Granger causality [49] is a statistical concept of causality based on prediction given two stochastic signals or variables, \(X_t\) and \(Y_t\). In this concept, \(X_t\) (Granger) causes a variable \(Y_t\). The previous values of \(X_t\) must contain information that helps predict \(Y_t\) over and above the information contained in the previous values of \(Y_t\) alone and vice versa. Its mathematical formulation is based on linear regression modeling of stochastic processes [33].

2. Materials and Methods

The process described in Figure 1 involved the use of various materials (refer to Table 1). It began with a scenario composed of a sequence of images that were analyzed by a background model. This model helps to detect objects, enabling the application of geometric descriptors to extract characteristic features of moving objects. Feature extraction was carried out during offline periods of a week, specifically in the selected streets. The information obtained from this process was then used to generate a time series. Trajectories were grouped based on the number of detected objects, and causality analysis techniques were applied to characterize the behavior of the road scenario. The correlation matrix obtained in the causality analysis was then used to create an external dependence model using graphs.

Table 1. Description of materials used.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset TACC [50]</td>
<td>Scenery: Lamar and 38th street.</td>
<td>36,000 frames.</td>
</tr>
<tr>
<td>Dataset: study scenery</td>
<td>Scenery: Av. Paseo de la Constitución, Querétaro, Qro., México.</td>
<td>1,514,286 frames.</td>
</tr>
<tr>
<td>Lambda server</td>
<td>Workstation with NVIDIA GPU, Ubuntu S.O.</td>
<td>RTX4090, 16,384 CUDA.</td>
</tr>
<tr>
<td>Python</td>
<td>Programming language.</td>
<td>3.9.13 version.</td>
</tr>
<tr>
<td>Dome cameras</td>
<td>Three dome cameras PTZ VIVOTEK.</td>
<td>Model SD9364-EHL.</td>
</tr>
</tbody>
</table>
2.1. Scenery

The scenery is a planned monitoring scheme that employed three professional dome cameras from PTZ VIVOTEK, model SD9364-EHL. These cameras have real-time compression in formats H.265, H.264, and MJPEG (triple codec), a 30× zoom lens, a wide temperature range (−50–55 °C) for extreme weather conditions, and a vandal-resistant enclosure with NEMA 4X. In addition, an Alienware Aurora Dell-branded computer with a Linux Ubuntu 18.04.6 LTS operating system and Core i7-9700K with a 3.6 GHz to 4.9 GHz processor. The above is to capture and analyze data for this work. This is a technological resource used to identify the behavior of moving objects (vehicles). Thus, identifying situations of interest is employed to model and find an optimal representation of the object’s movement. Figure 2 shows the camera positioning along the Paseo de la Constitución Avenue, Querétaro city in México.

2.2. Features Extraction

The feature extraction ($T$, geometric measurements) was determined from the information related to the moving objects and the fixed objects. Thus, each particular position of the pixels in the image is $x = [x_1, y_1]$ and is indexed as $I_i(x)$ for the dimensions of the images $k \times l$. To characterize the behavior dataset, methods provided by the following references [50, 51] were employed. Then, the geometric measurement extraction algorithm worked from a sequence of images expressed as $\{I_0, \ldots, I_n\}$. The extraction pseudocode is shown in Algorithm 1, and its purpose is to generate a dataset.

Figure 1. Data acquisition methodology to be followed in the characterization of the behavior of the study scenery.

Figure 2. Geographic location of inspection points.
Algorithm 1. Pseudocode employed for geometric feature extraction.

<table>
<thead>
<tr>
<th>Input</th>
<th>Image sequence $I_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>Geometric measurements $T$</td>
</tr>
</tbody>
</table>

set intensity criterion;
set radius of the structuring object;
set connectivity pixel;

while has a sequence $I_t$ then
  get image($i$);
  set image($i$) to grayscale;
  set image($i$) to normalize;
  set background model;
  set pixel intensity criterion;
  set morphological process;
  [$L,n$] = get moving objects;
  for each object found $n$ then
    xy = get connected components;
    arx = calculate the area;
    per = calculate the perimeter;
    cxy = find the centroid in $x,y$;
    cir = calculate the circularity value;
    eul = find Euler’s number;
    std = calculate the standard deviation in $x,y$;
    [m1,..., m7] = calculating Hu’s inv. moments;
    [n,coord,puntos] = detect Harris corners;
    save and insert to $T$ dataset;
  end
  return $T$;
end

2.3. Time Series

To find time series, representative routes and trends on trajectories were identified. Thus, a histogram was built to represent and visually summarize the trajectory distribution of the objects. The data employed in the histogram construction were the measured frequency of the pixels (trajectories).

2.4. Causality Analysis

Clustering was used to identify representative routes and common trends of several trajectories. Then, graph clustering algorithms were employed to find subgraphs as representative routes or centroids. Then, the groups were formed according to the similarity of the trajectories to those centroids. After the clustering and correlation were performed, the $k$ value procedure was performed to obtain graphical information for its calculation, thus distinguishing between trajectories by some non-obvious characteristic, such as the type of movement, mode of transportation, or activity.

The correlation was classified as follows:
If $r = 1$: perfect positive correlation.
If $0 < r < 1$: it reflects that there is a positive correlation.
If $r = 0$: in this case, there is no linear relationship.
If $-1 < r < 0$: there is a negative correlation.

2.5. External Dependence Model

This model consists of finding the temporary relationships to build a graph where a relationship has a higher associated likelihood to be observed within a time interval. It means that, if a time series causes another, the knowledge of the first series could help to predict the future values of the other (after the influence of the other variables). Here, the main contribution is to find automatic relationships by using a simple camera. The above is obtained by following the model proposed in reference [33].
3. Results

A system is implemented where moving objects are analyzed and categorized (characterized) by the shape and attributes observable in the movement. A set of algorithms is introduced that help us quantify moving objects over time, and from the evidence (generated database) of the set of cameras, \( c_i \in C \), a state-based and probabilistic model is created that defines the main movement zones and finds the temporal patterns of behavior \( P(I_k, I_l) \) for \( I_k, I_l \in c_i \), which are related to the semantics of the scenario that is useful for analysis; the patterns are the result of associating through causal relationships. For example, determining the origin of the traffic at a certain time of day, the effects that an accident generates on a road artery and which ones will be affected, and changes or closures of roads due to events, parades, or demonstrations. Patterns are the result of associating through Granger causal relationships.

3.1. Scenery

The results of the data acquisition are described in this section. As mentioned in Section 2, such data belong to three cameras installed along the analyzed street. In that sense, the cameras are represented by \( c_i \in C \) to form \( c_i = \{I_1, \ldots, I_n\} \) and each \( I_j \) is an attribute generated by an intelligent process on the image. From the image dataset of the study scenery, we proceeded with the detection of object motion. Then, the background model was calculated from the information of moving objects and stationary objects. Then, the background model was calculated from the information of moving objects and stationary objects (see Figure 3), where each particular pixel position in the image \( x = [x_1, y_1] \) is indexed as \( I_k(x) \) for the \( k \times l \) image dimensions.

![Figure 3](image-url)

**Figure 3.** Dataset generation and ground truth of the study scenery. Although, a change in image quality can be noted, all the data were acquired correctly: (a) dataset of 504,762 images obtained from camera \( c_3 \); (b) dataset of 504,762 images obtained from camera \( c_2 \); (c) dataset of 504,762 images obtained from camera \( c_1 \). The above was recorded at an instant of time \( 0 \leq t \leq T \) for an image sequence \( \{I_1, \ldots, I_f\} \).

3.2. Features Extraction

As shown in Table 1 in Section 2, 1,514,286 images were generated for the study scenery, as well as 36,000 frames that were obtained from the Texas Advanced Computing Center [50]. Finally, 1700 images were generated as another dataset [51], to validate the functionality of the algorithm in the feature extraction task (see Algorithm 1). Shape descriptors are implemented in the algorithm (see Table 2) for normal dynamic detection. As a result, we have a dataset that has different objects and for each object, different measurements (see Table 3), so that, if viewed in time, we observe objects with measurements that can be related to the time instant \( t_1 \) with the time instant \( t_2 \). These characteristics allow objects to be represented by a set of numerical values; the characteristics obtained are invariant to scaling, rotation, and translation.
### Table 2. List of geometric characteristics.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Object number</td>
<td>Number of connected objects in the image.</td>
</tr>
<tr>
<td>Area</td>
<td>Area of objects, pixels of the area of objects in the image.</td>
</tr>
<tr>
<td>Perimeter</td>
<td>Pixels of the perimeter of the image objects.</td>
</tr>
<tr>
<td>Centroid</td>
<td>It is the geometric center of the body/point where the total area of a figure is considered to be concentrated.</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>A most common measure of dispersion that indicates how dispersed the data are concerning the mean.</td>
</tr>
<tr>
<td>Circularity [43]</td>
<td>Percentage of circularity of objects.</td>
</tr>
<tr>
<td>Euler number [46]</td>
<td>It is the total number of objects in the image minus the total number of holes in those connected objects.</td>
</tr>
<tr>
<td>Harris corners [47]</td>
<td>It is used to extract certain types of features and infer the content of an image. A corner is the intersection of two edges and/or a point for which there are two dominant edge directions.</td>
</tr>
<tr>
<td>Hu moments [48]</td>
<td>Set of seven invariant descriptors that quantify the shape of an object (centroid, area, and orientation)/[ordinary, centralized, and normalized].</td>
</tr>
</tbody>
</table>

### Table 3. Overview of the structure of the study scenery dataset.

<table>
<thead>
<tr>
<th>OBJ</th>
<th>AREA</th>
<th>PERIMETER</th>
<th>X</th>
<th>Y</th>
<th>STDX</th>
<th>STDY</th>
<th>CIR</th>
<th>EULER</th>
<th>HARRIS</th>
<th>MHU1</th>
<th>MHU2</th>
<th>MHU3</th>
<th>MHU4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>62</td>
<td>27.589</td>
<td>69.806</td>
<td>97.323</td>
<td>2.9578</td>
<td>1.8087</td>
<td>0.80645</td>
<td>1</td>
<td>2</td>
<td>0.20291</td>
<td>0.01238</td>
<td>0.00286</td>
<td>0.01120</td>
</tr>
<tr>
<td>2</td>
<td>37</td>
<td>19.815</td>
<td>132.57</td>
<td>57.162</td>
<td>2.2304</td>
<td>1.3645</td>
<td>0.86486</td>
<td>1</td>
<td>2</td>
<td>0.18919</td>
<td>0.00808</td>
<td>0.00068</td>
<td>0.00725</td>
</tr>
<tr>
<td>1</td>
<td>63</td>
<td>27.986</td>
<td>69.825</td>
<td>97.667</td>
<td>2.9267</td>
<td>1.9177</td>
<td>0.7619</td>
<td>1</td>
<td>2</td>
<td>0.20912</td>
<td>0.01300</td>
<td>0.00245</td>
<td>0.01360</td>
</tr>
<tr>
<td>2</td>
<td>24</td>
<td>15.316</td>
<td>132.5</td>
<td>57.5</td>
<td>1.7446</td>
<td>1.1421</td>
<td>0.83333</td>
<td>1</td>
<td>2</td>
<td>0.19444</td>
<td>0.00525</td>
<td>0.00091</td>
<td>0.01108</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>62</td>
<td>27.589</td>
<td>97.323</td>
<td>97.323</td>
<td>2.9578</td>
<td>1.8087</td>
<td>0.80645</td>
<td>1</td>
<td>2</td>
<td>0.20291</td>
<td>0.01238</td>
<td>0.00286</td>
<td>0.01120</td>
</tr>
<tr>
<td>1</td>
<td>62</td>
<td>27.589</td>
<td>97.323</td>
<td>97.323</td>
<td>2.9578</td>
<td>1.8087</td>
<td>0.80645</td>
<td>1</td>
<td>2</td>
<td>0.20291</td>
<td>0.01238</td>
<td>0.00286</td>
<td>0.01120</td>
</tr>
</tbody>
</table>

#### 3.3. Time Series

The dataset generated by the algorithm allows us to find the temporal relation $R(I_K, I_L)$ for $I_K, I_L \in c_i$ and, subsequently, with these data the associated graph for this relation can be constructed. Figure 4 represents the acquisition of frequency data versus the values of each trajectory of the objects captured by the cameras on the dataset mentioned in Table 3.
Figure 4. Histogram of trajectories. In subsections: (a) the histogram, the x-axis is a numerical line that visualizes the range of values of trajectories in the images that have been divided into ranges of numbers or bins (nBins = 100, changing the number of bins allows to see more or less detail in the structure of the data). A bar is drawn for each bin where the width of the bar represents the number density range of the bin, and the height of the bar represents the number of trajectories included in that range; (b–d) it can be seen that the most representative paths and the most common trends of the trajectories are concentrated in a range of values from approximately 30 to 110 sampling points.
3.4. Causality Analysis

The following figures (see Figures 5 and 6) show the result of applying the k-means clustering algorithm to the data set of the trajectories of the objects of each camera belonging to the experiment scenario. This grouping allows us to distinguish trajectories by some non-obvious characteristic, such as the type of movement, mode of transport, or activity carried out.

Figure 5. Result of k-means clustering of object trajectories with elbow point value $k = 5$, represented with different colors.

Figure 6. View of the grouping of trajectories in the experimentation scenery.

As the traffic dynamics are changing, the square correlation ($R$) “$n \times n$ matrix” is used to find the correlations of the flow intensity, which in this case is through geometric measurements/characteristics (see Figures 7 and 8). Based on Section 2.4, a correlation close to zero means no relationship, but high or maximum local correlations in their absolute values represent indications of a positive or negative relationship.
3.5. External Dependence Model

Based on the results of the correlation matrices, we proceed to illustrate a situation with a high correlation (see Figure 9), in this case using camera $c_3$ with the highest correlation, which represents this correlation at certain changes; thus, the dependency model is observed.

```
Figure 7. Correlation coefficient of geometric characteristics in cameras $c_1$, $c_2$, and $c_3$.

Figure 8. Dendrogram of geometric characteristics in the cameras $c_1$, $c_2$, and $c_3$.
```

3.5. External Dependence Model

Based on the results of the correlation matrices, we proceed to illustrate a situation with a high correlation (see Figure 9), in this case using camera $c_3$ with the highest correlation, which represents this correlation at certain changes; thus, the dependency model is observed.

```
Figure 9. Dependency model for a situation with high correlation in camera $c_3$.
```

The present results seek to find distantly related states; therefore, below, the high positive and negative correlations that are not within the same chamber are shown (see Figure 10).
This means that this relationship shows, when a flow is observed from the first state, which state will be associated with it and will be affected by that change in the flow (see Table 4). Low probabilities simply mean that this dependence is not always visible; only when there are peaks, which in certain situations are when there are many vehicles, are these dependencies observed, which is why they are close to zero.

- High positive correlations: if the correlation value is high and positive, it means that the direction of the flow of the vehicles is downward (they go down), that is, if there is flow in camera $c_1$, proportionally there will also be flow in camera $c_2$.
- High negative correlations: negative values represent a flow change; if high, it means that the number of vehicles decreases (opposite flow).

### Table 4. Interpretation of results of correlation coefficients between camera regions $c_1$, $c_2$, and $c_3$.

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>+0.93</td>
<td>The positive correlation of 0.93 in blue between the zero regions of camera $c_2$ and camera $c_3$ means that when the zero region of camera $c_2$ presents traffic or there are high increases, there will also be increases proportionally in the zero region of camera $c_3$. On the other hand, in the negative correlation of $-0.69$ in red between region one of camera $c_2$ and region zero of camera $c_3$, an inverse behavior is presented, which means that, if there is a lot of traffic in camera $c_2$ of region one, it will decrease in region zero of camera $c_3$. A similar case occurs in the negative correlation of $-0.67$ in red between region two of camera $c_2$ and region zero of camera $c_3$, where an inverse behavior occurs, that is, if a lot of traffic occurs in camera $c_2$ of region two, it will decrease in the zero region of camera $c_3$. Also, for the negative correlation of $-0.43$ in red between region three of camera $c_2$ and region zero of camera $c_3$, there is an inverse behavior, inferring that, if a lot of traffic occurs in camera $c_2$ of region three, it will decrease in the zero region of camera $c_3$. A similar case occurs for the negative correlation of $-0.46$ in red between region four of camera $c_2$ and region zero of camera $c_3$, where an inverse behavior occurs, deducing that, if there is a lot of traffic in camera $c_2$ of region four, it will decrease in the zero region of camera $c_3$. For the positive correlation of 0.85 (blue) between the zero regions of camera $c_1$ and camera $c_3$, it means that when the zero region of camera $c_1$ presents traffic (high increases), proportionally there will also be traffic in the zero region of camera $c_3$. Finally, in the negative correlation of $-0.43$ in red between region one of camera $c_1$ and region zero of camera $c_3$, an inverse behavior is presented, which means that, if there is a lot of traffic in camera $c_1$ of region one, it will decrease in region zero of camera $c_3$.</td>
</tr>
<tr>
<td>-0.46</td>
<td></td>
</tr>
<tr>
<td>+0.85</td>
<td></td>
</tr>
<tr>
<td>-0.43</td>
<td></td>
</tr>
</tbody>
</table>

Now, two distant cameras and two different states represent the cause–effect relationships that could occur within the avenue itself (all those that have a high positive correlation are identified). Therefore, by generalizing the concept of correlation because the avenue is connected and is being monitored by several cameras, we can find the temporal relationships of the flow dependence in terms of the measurements/characteristics used. The case (see Figure 11) that represents the maximum relationship between the state of...
camera $c_2$ and the other camera $c_3$ represents a cause–effect relationship in the behavior of the vehicles that are observed.

Finally, in the negative correlation of $-0.43$ in red between region one of camera $c_1$, an inverse behavior is presented, which means that, if there is a lot of traffic, this will decrease in region zero of camera $c_2$. An inverse behavior helps to infer that possible states are altered once the state of reference begins to change the dynamics of the vehicular flow. For example, the positive correlation of $+0.85$ in blue between the zero regions of camera $c_2$ and camera $c_3$ (see Figure 12) means that when the zero region of camera $c_2$ presents traffic or there are high increases in their geometric measurements of the objects (area and perimeter, see Figure 12d), proportionally there is also traffic or high increases in the zero region of camera $c_3$. These patterns are the result of associating through causal relationships to say what the origin of traffic is at a certain time of day, the effects that an accident generates on a road artery and which ones will be affected, and changes or closures of roads due to events, parades, or demonstrations.

It is shown that the high probability of $+0.85$ indicates the dependence between two states that are not adjacent but are at a distance and a behavior is being observed and that behavior helps to infer that possible states are altered once the state of reference begins to change the dynamics of the vehicular flow. For example, the positive correlation of $+0.85$ in blue between the zero regions of camera $c_2$ and camera $c_3$ (see Figure 12) means that when the zero region of camera $c_2$ presents traffic or there are high increases in their geometric measurements of the objects (area and perimeter, see Figure 12d), proportionally there is also traffic or high increases in the zero region of camera $c_3$. These patterns are the result of associating through causal relationships to say what the origin of traffic is at a certain time of day, the effects that an accident generates on a road artery and which ones will be affected, and changes or closures of roads due to events, parades, or demonstrations.
4. Discussion

In terms of application, this research showed that, by using feature extraction, it is possible to find the relationships between a set of urban variables (traffic, pedestrians, traffic light timing, and road obstructions, to name a few), measured by a fixed zone monitoring system, resulting in the generation of a stochastic causal model using graphs. The proposed model generates positive and negative dependence relationships with an average accuracy of 78%, which provides the information to interpret the cause–effect relationships and allows the detection of atypical behaviors, helping to better understand the dynamics of the monitoring scenario. Existing algorithms in the literature were implemented to detect the local behavior of objects in each camera. In conjunction, these results can be obtained by implementing the stochastic model on general-purpose electronic cards, which can be applied for inferences and modeling of periodic behaviors of scenarios in smart cities. The originality consists of finding the temporal and spatial dependencies between the behavior of a particular camera and the other adjacent cameras in the study scenario.

5. Conclusions

This research consisted of finding the time relationship \( R(I_K, I_L) \) for \( I_K, I_L \in c_i \), and constructing the associated graph for this relationship; the relationship \( R \) has associated with it a high probability \( P(R) \) of being observed for an interval \( t^* \subset 0 \leq t \leq T \). If one time series causes another, knowledge of the first process will help to predict the future values of the other after the influences of other variables have been taken into account. The fundamental contribution is to automatically find the \( R \) relationships using the evidence observed by a particular camera \( c_i \). Therefore, the method for finding \( R \) is to employ a graph-based causal model.


Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available at: https://github.com/trejoan1/patterndetectionmodel (accessed on 20 May 2024). The source code was typed in Python 3 using Jupyter Notebook.

Acknowledgments: Thanks to the UAQ Informatics faculty, CIDESI, and CONAHCYT for the facilities for the development of this work (materials used for experiments). The authors wish to thank the anonymous reviewers for their positive and valuable comments to improve and refine this work.

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

References
2. Camero, A.; Alba, E. Smart City and information technology: A review. *Cities* 2019, 93, 84–94. [CrossRef]
49. Lima, V.; Dellajustina, F.J.; Shimoura, R.O.; Girardi-Schappo, M.; Kamiji, N.L.; Pena, R.F.O.; Roque, A.C. Granger causality in the frequency domain: Derivation and applications. Rev. Bras. Ensino Física 2020, 42, e20200007-10. [CrossRef]
50. Dubrow, A. Artificial Intelligence and Supercomputers to Help Alleviate Urban Traffic Problems; Texas Advanced Computing Center: Austin, TX, USA, 2017.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.