ISBN: 978-628-95207-4-3. ISSN: 2414-6390. Digital Object Identifier: 10.18687/LACCEI2023.1.1.1447

Intelligent system to detect violations in pedestrian areas committed by vehicles in the City of Cartagena de Indias

Nicolás Hernández-Díaz; Eng[1], Yersica C. Pañaloza, Eng[2], Y. Yuliana Rios, PhD[3], Juan Carlos Martinez-Santos, PhD[4], Edwin Puertas, PhD[5]
[1,4,5] Universidad Tecnólogica de Bolívar, Colombia, nihernandez@utb.edu.co, epuerta@utb.edu.co , jcmartinezs@utb.edu.co
[2] Universidad de Pamplona, Pamplona, Colombia, yersica.carrillo@unipamplona.edu.co
[3] Universidad Industrial de Santander, Bucaramanga, Colombia, yeriosdi@uis.edu.co

***Abstract*– This article presents applied research which proposes an approach to a low-cost electronic system that allows the detection of violations in pedestrian areas. For this purpose, we proposed an electronic system that implements artificial intelligence algorithms and Internet of Things architectures to detect in real-time these events. The proposal focuses on experimental laboratory work through which we developed and validated a system by analyzing a real pedestrian zone scenario and invasion of the pedestrian zone. As a result of this work, we deliver a labeled dataset to the scientific community that can serve as input to new projects and the development of better offender detection algorithms.**

***Keywords*—intelligent system, object detection, computer vision.**

I. Introduction

Within the Sustainable Development Goals (SDGs), the third objective, Good Health, and Well-being hope to guarantee a healthy life and promote well-being at all ages, thus ensuring that non-compliance generates uncertainty in the search for a sustainable system [1]. In addition, following the SDGs, the eleventh Sustainable Cities and Communities goal alleges that rapid urbanization is resulting in an increasing number of inhabitants in slums, inadequate and overburdened infrastructure, and services (such as waste collection and water and sanitation systems, roads, and transportation), which is worsening air pollution and urban sprawl [2]

Accidents between two or more subjects on the road due to inefficient traffic control can vary in nature and severity. Additionally, certain types of accidents occur more frequently than others. The WHO (World Health Organization) points out that around 3,500 people die daily on the roads. In addition, millions of people become injured or disabled each year. Children, pedestrians, cyclists, and the elderly are the most vulnerable users of public roads [3]. In addition, a publication by the same entity states that approximately 1.25 million people die each year on the world's roads due to road crashes. Furthermore, they are the leading cause of death in the young population between 15 and 29 years of age [4].

These data are related to the situation in Colombia. According to the National Road Safety Observatory (ANSV - Agencia Nacional de Seguridad Vial), they reveal that during the years 2017, 2018, and 2019, road accidents represented the second cause of death in this population. Furthermore, 1,616 children and adolescents died in the last three years in road accidents. Therefore, and even, an agreement that is part of the strategy of the Colombian National Government to generate a public policy aligned with the 2030 agenda for compliance with the Sustainable Development Goals (SDG 3.6 and SDG 11.2) and that is added to the transversal public policy on road safety. It addresses the regions in four dimensions: victims, education, optimal vehicles and infrastructure, and suitable users. Furthermore, the Ministers of Transport, Health, and Education signed with representatives of the Global Alliance of nongovernmental organizations for Road Safety, DITRA, and the Business Safety Committee [5].

On the other hand, in Cartagena de Indias's most recent development plan, there is a need to migrate toward a smart city. For this, two strategies become relevant: (i) have a modern technological infrastructure; (ii) have data for decision-making. However, this development plan [6] also highlights that "Cartagena is at level 1 according to the maturity model of smart cities defined by MinTIC", indicating a significant gap in achieving the goal of migrating to a smart city. For this reason, any initiative that contributes to the technological development of intelligent systems that contribute to solving the city's problems (such as the detection of violations by motorcyclists in pedestrian areas) would contribute to the vision of migrating towards a smart city.

Currently, traffic control systems do not have a methodology based on local feedback, which means that on-site supervision is not the most optimal, so an adequate traffic flow is not achieved, much less any degree is guaranteed. of traffic education, in this particular case, about the appropriate use of the pedestrian level crossing, which facilitates the occurrence of accidents, or the violation of the traffic laws established at the national level, this situation directly affects the public policy aligned to the 2030 agenda for the fulfillment of the Sustainable Development Goals. Therefore, the development of an intelligent system that allows detection, within the boundary arranged for a pedestrian level crossing, Within the Sustainable Development Goals (SDGs), the third objective, Good Health, and Well-being, hopes to guarantee a healthy life and promote well-being at all ages, thus ensuring that non-compliance generates uncertainty in the search for a sustainable system [1]. In addition, following the SDGs, the eleventh Sustainable Cities and Communities goal alleges that rapid urbanization is resulting in an increasing number of inhabitants in slums, inadequate and overburdened infrastructure, and services (such as waste collection and water and sanitation systems, roads, and transportation), which is worsening air pollution and urban sprawl [2]

Accidents between two or more subjects on the road due to inefficient traffic control can vary in nature and severity. Additionally, certain types of accidents occur more frequently than others. The WHO (World Health Organization) points out that around 3,500 people die daily on the roads. In addition, millions of people become injured or disabled each year. Children, pedestrians, cyclists, and the elderly are the most vulnerable users of public roads [3]. In addition, a publication by the same entity states that approximately 1.25 million people die each year on the world's roads due to road crashes. Furthermore, they are the leading cause of death in the young population between 15 and 29 years of age [4].

These data are related to the situation in Colombia. According to the National Road Safety Observatory (ANSV - Agencia Nacional de Seguridad Vial), they reveal that during the years 2017, 2018, and 2019, road accidents represented the second cause of death in this population. Furthermore, 1,616 children and adolescents died in the last three years in road accidents. Therefore, and even, an agreement that is part of the strategy of the Colombian National Government to generate a public policy aligned with the 2030 agenda for compliance with the Sustainable Development Goals (SDG 3.6 and SDG 11.2) and that is added to the transversal public policy on road safety. It addresses the regions in four dimensions: victims, education, optimal vehicles and infrastructure, and suitable users. Furthermore, the Ministers of Transport, Health, and Education signed with representatives of the Global Alliance of nongovernmental organizations for Road Safety, DITRA, and the Business Safety Committee [5].

On the other hand, in Cartagena de Indias's most recent development plan, there is a need to migrate toward a smart city. For this, two strategies become relevant: (i) have a modern technological infrastructure; (ii) have data for decision-making. However, this development plan [6] also highlights that "Cartagena is at level 1 according to the maturity model of smart cities defined by MinTIC", indicating a significant gap in achieving the goal of migrating to a smart city. For this reason, any initiative that contributes to the technological development of intelligent systems that contribute to solving the city's problems (such as the detection of violations by motorcyclists in pedestrian areas) would contribute to the vision of migrating towards a smart city.

Currently, traffic control systems do not have a methodology based on local feedback, which means that on-site supervision is not the most optimal, so an adequate traffic flow is not achieved, much less any degree is guaranteed. of traffic education, in this particular case, about the appropriate use of the pedestrian level crossing, which facilitates the occurrence of accidents, or the violation of the traffic laws established at the national level, this situation directly affects the public policy aligned to the 2030 agenda for the fulfillment of the Sustainable Development Goals. Therefore, the development of an intelligent system that allows detecting, within the boundary arranged for a pedestrian level crossing, the interaction between vehicles and pedestrians using artificial vision techniques combined with machine learning techniques for real-time data processing to guarantee and strengthen the government's road safety strategy.

II. RELATED WORK

The improper use of pedestrian areas by vehicles is one of the principal cases of vehicular recklessness that represents a great danger to society and pedestrians in general. A clear example of this recklessness occurs in Cartagena de Indias, a city in the north of Colombia, which experiences motorcycle recklessness daily. These motorized invade platforms and pedestrian crossings (also known as "zebras"), disrespecting the space and violating the lives of pedestrians due to their irrationality in driving and lack of road education [7, 8]. Despite a law that penalizes this imprudence (law 1801 of 2016, National Police and Coexistence Code), in the case of Cartagena de Indias, there have been multiple strategies that address to counteract said imprudence committed in the pedestrian areas by motorcycles. For example, in the Cartagena 2020/2023 development plan, "Save Cartagena Together". Different actions are evidenced in the 2021 action plan [9] of the Administrative Department of Traffic and Transportation DATT to reduce road accidents. Taking it as a strategy, the implementation of training programs in driver education. However, recklessness by motorcyclists toward pedestrians is still in force.

The accident situation is recurrent in the city of Cartagena. According to the DATT, although they carry out controls and operations, various road safety education campaigns, and different road safety strategies, the reduction of accidents from one year to the next varies by only 2.3%, as observed in Table 1.

TABLE I
ROAD ACCIDENTS IN CARTAGENA [10]

Year	**2019**	**2018**	**Variation**
January	416	430	-3.4%
February	418	444	-6,2%
March	423	496	-17.3
April	475	452	4.8
May	474	467	1.5%
June	445	422	5,2%
Total	2652	2711	-2.3%

It is evident that this decrease, although positive, does not represent a high impact in terms of traffic control in the city, and this is due, among other reasons, to the fact that the DATT cannot guarantee for each square meter of the city a traffic agent. DATT Cartagena pointed out that traffic is an issue that concerns all of us and, therefore, it is essential to do as a people that we continuously and politely participate in the foundation of a road culture that makes it possible to travel through the city safely [11]. This statement is susceptible in this regard, and the reality that the town lives due to its citizens' lack of road culture is complicated unless its citizens act as traffic agents.

Therefore, several authors have worked on artificial vision to mitigate the automatic detection of pedestrians and vehicles. For example, the authors [11] present a comprehensive method that combines the Gaussian mixture model based on the foreground detection method with Hough line detection, color recognition, and centroid detection methods to detect whether the pedestrian crosses the specified boundary in the video surveillance system. Along the same lines, the authors in [12] describe the detection of a crosswalk using a single camera to detect pedestrians, the classification is binary, and the processing is performed using an improved method of the art by name scattered optical flow. They evaluated the proposed algorithms using public datasets such as (Caltech and ETH) and the authors' data. The performance result shows that the correct pedestrian detection rate is 99.50% at 0.09 false positives per image.

On the other hand, the authors in [12] propose a new method of reidentification of pedestrians that adapts to the change of light, for which artificial intelligence technology is used, such as extraction of local characteristics of pedestrian images based on the Bag of Words model and the KISSME algorithm. Finally, they established a comprehensive decision scheme based on the recognition results of multiple functions to obtain the outcome of pedestrian recognition. In the same way, the authors [13] propose to build a driving assistance device that can detect sudden pedestrians crossing the road using a Raspberry Pi microcontroller. To do this, they use libraries such as Open CV in Python and a buzzer to call the driver's attention to reduce the vehicle's speed or stop it.

On the other hand, the semantic segmentation of images is one of the critical applications in image processing and the domain of computer vision, since it allows a linguistic classification of the elements present in a study image [15]. It will enable us to use this technique in multiple domains, such as the medical area and intelligent transportation. The extraordinary rise of this technique has caused great interest in the great community of researchers around the world, to such an extent that databases organized and structured in classes or types of objects are published and freely accessible through the cloud, which facilitates the verification of any algorithm designed, an example of these datasets can be seen at [15] Several authors have studied semantic segmentation for many years. Since the appearance of the Deep Neural Network (DNN), which generally are neural networks with at least two layers capable of modeling complex systems [16], segmentation has made enormous progress.

To date, developments in this field are in high demand. It is evident by the different approaches used and recent technologies applied. The results are very positive, and indeed, the research concerning the applicability of artificial intelligent techniques in machine vision for process control and its potential is still in the exploration process.

III. INTELLIGENT SYSTEM

One of the great difficulties the scientific community faces when developing artificial intelligence models is obtaining good quality, well-structured, and publicly accessible data, especially if the semantic context involves classes such as pedestrians, motorists, and their respective interactions with the road as themselves. Moreover, building a dataset is time-consuming and tedious, especially if there are no tools or methodologies to facilitate its generation and development.

The methodology developed for constructing datasets consists of two main lines of execution. However, only the image acquisition line from the CCTV server uses a Comparative Chronological Analysis to determine whether it is necessary to download the image. The nth camera observed them at an instant t of time. We use ROIs of the rectangular type and hotkeys for the labeling process.

Comparative Chronological Analysis is the inference process of analyzing the current image observed by the nth camera and the magnitude of the observable difference on a previous image from the same camera. The following equation represents the operation done on it to the data observed by each camera.

$$ACC(A, B) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}(A_{(i,j)} - B_{(i,j)})_{(i,j)} \qquad (1)$$

Where:

- *ACC (A, B)* Chronological Comparative Analysis (Scalar Value).
- *n* is fxc the number of rows times the number of matrix columns.
- *(i, j)* the state of indexing over the array.
- ***A*** the current image observed by the nth camera.
- *B* the previously stored image of the nth camera.

The criteria for downloading an image for the nth camera. analyzed depends on the value ACC (A, B) previously obtained, therefore,

$$BD = \begin{cases} True & \text{if } ACC(A,B) > threshold \\ False & \text{if } ACC(A,B) \leq threshold \end{cases} \quad (2)$$

Where:

- *BD* Flag to download.
- ACC(A, B) Chronological Comparative analysis (Scalar Value).
- *threshold* that sets the download criteria.

The Comparative Chronological Analysis for the Image Download method can ensure that the system does not store identical images based on the download criteria due to possible communication errors or the asynchronous update behavior the connected IP cameras typically connect to the CCTV. It achieves the construction of a raw dataset that does not contain repeated or too similar images.

Fig. 1 Example of a selected ROI for an object of a specific class.

After obtaining the raw dataset, labeling based on the classes defined in the project is necessary. In an image, we represent the classes with two-dimensional data. Therefore, the user can extract the information of interest. A technique widely used by different manual labelers is using rectangular ROIs.

Rectangular ROIs is a tool that makes it possible to manually select a range of interest in an image by selecting the area in the image and returning a vector containing the indexing of the top starting vertex of the rectangle and the bottom opening vertex. Figure 1. presents a sample of the dataset built in this document, along with the ROI selection.

The fast access keys are a procedure based on the continuous sensing of the state of a key during the observation of the nth image or general sample so that each key is assigned a class or task. For this work, the following equation summarizes the cases.

$$Assign = \begin{cases} classOne & \text{if } keyState == A \\ classTwo & \text{if } keyState == B \\ classThree & \text{if } keyState == C \\ classFour & \text{if } keyState == D \\ classFive & \text{if } keyState == E \\ continue & \text{if } keyState == Q \end{cases} \quad (3)$$

Where:

- *Assign* corresponds to the path within a dictionary. the cutout, the ROI coordinates, and the class label should be stored.
- *classOne*, classTwo... are the distant classes recognized during the project.
- *continue* to tell the system that the general sample information does not contain classes or that there are no more classes to classify.

Combining both concepts makes it possible to build refined. and structured dataset, as shown in Figure 2.

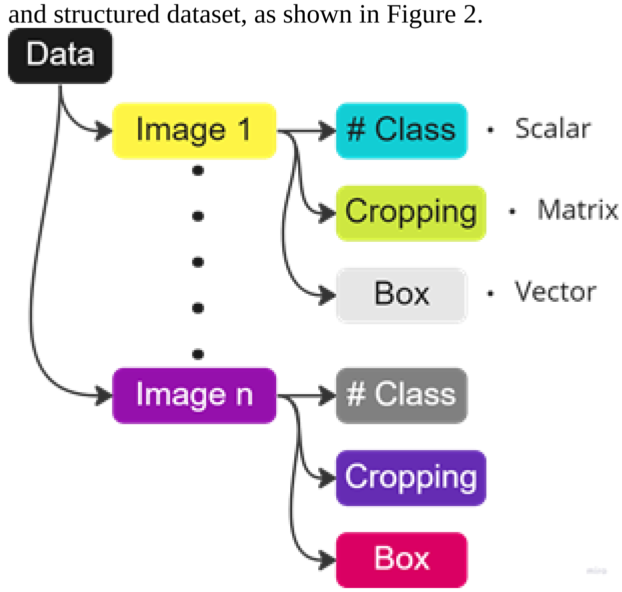

Fig. 2 Structure of the obtained clean dataset.

The data come from a free CCTV comprising 80 cameras arranged publicly in the city of Medellín [16]. They are part of the Intelligent Mobility System of Medellín (SIMM), a pioneering program in Colombia of the Medellíin Mobility Secretary, which emerged in 2020 as a response to the need to optimize the city's road network.

The nature of the data represents a challenge to its excellent acquisition because although they are entirely free, each camera updates the information on the server after approximately five minutes and asynchronously to the other cameras.

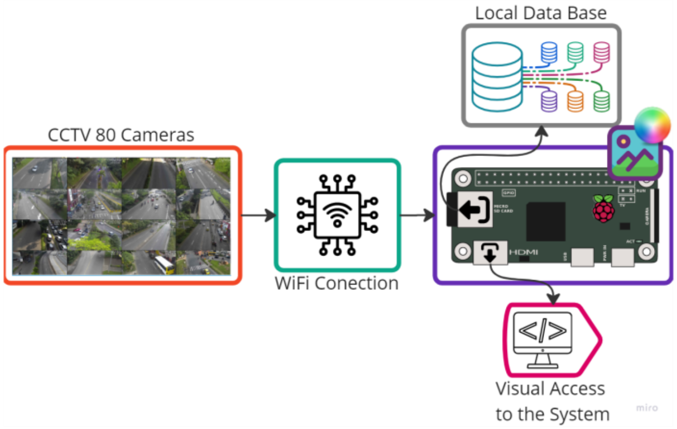

Fig. 2 Methodology implemented.

Figure 2. shows a diagram of the proposed methodology. As we can discern, the system depends on the availability and state of the data present in the CCTV. The request for information (RGB Image) is made to each camera independently through a Raspberry Pi Zero W with an internet connection.

This embedded device, employing a Comparative Chronological Analysis, downloads images in JPG format to construct the raw dataset during the time indicated by the user or until the execution is interrupted. Finally, the user manually initializes an algorithm based on the selection of Regions of Interest (ROI) and hotkeys to decompose the information in the nth image of the raw dataset into the required classes, thus getting the dataset clean and ready to train an object classifier.

Finally, three algorithms were developed for the following purposes.

Algorithm 1
The code block below presents the algorithm in charge of the data acquisition process. It follows the proposed methodology and corresponds to the implementation of the Comparative Chronological Analysis for Image Download; the flag type variables are used to verify the connection to the internet and the CCTV server, the variables old and new check correspond to the variables (A, B) respectively.

Algorithm 2
The code below shows how extracting Regions of Interest (ROI) was put into practice and using hotkeys to direct the clipping, class, and bounding box object to the Data variable and the hierarchy in its data structure. Similarly, at the end of the code, it can see the redefinition of the data structure by going from general samples in .jpg format to a single binary file in Pickle format that contains all the information necessary to train and validate a multiclass object classifier model. Figure 4. shows the general use of this section.

Fig. 4 General operation of the labeling system.

Algorithm 3
After completing the labeling process, using the code illustrated below, it is possible to review the labels made following the data hierarchy, presenting both the color assigned to the label and the object bounding box and the text of the class type. Figure 5., shows for the umpteenth image observed by one of the cameras, the labeling was carried out according to the different established classes.

Fig. 5 General operation of the label verification system.

RESULT

The dataset contains a total of 11361 images in JPG format for a total of 0.99 GB, the resolution of the images is 1280x720 (HD), and the weight per image is 56. 7 KB. We obtained the images from a closed-circuit television CCTV free public access located in Medellín, Colombia, consisting of 80 cameras on the road. The system captured these images during both night and day hours. The time range in which the system acquired them is from 04/10/2022 to 28/10/2022, a total of 24 days. The format of the name of each image is "image(camera number)_(dateFormat)_(timeFormat). The date format is AAAAMMDD, where AAAA is the year, MM is the month, and DD is the day. Finally, the time format is hhmmss, where hh is the 24-hour hour, mm is the minute, and ss is the second.

The file (annotations. pickle) corresponds to a file in PICKLE format that contains the labels made to each image and corresponds to a single primary variable of dictionary type, whose main key is the name of the image (image2_20221004_054300.jpg). The subkeys are box, which contains a list of lists of bounding boxes, clippings, which presents a list of arrays containing clippings made to the image, class, which includes a list of integer values representing the class to which corresponds the nth bounding box or clipping.

The dataset for Detecting Motorcyclists in Pedestrian Areas is hosted at https://doi.org/10.5281/zenodo.7935298 [17].

CONCLUSION

We built a dataset [17] following the desired semantic object using the tools implemented in Python and presented throughout this work. The dataset has 11,513 general samples and around 5,000 samples for each class. The Dataset is in a dictionary and stored in a binary file in pickle format. In addition, a tool was developed to download the images under the criterion of Comparative Chronological Analysis with an error of less than 0.5%, which guarantees its use in any project of a similar nature to the one presented in this work.

Another significant contribution is that the developed labeling system can process the raw dataset and, together with the user, perform the respective labeling of the information in a general example. It is independent of whether in the same example, there is more than one class or for each class more than one instance is, since, due to the interaction methodology through hotkeys, this is indifferent.

The environment for reviewing or checking the status of the refined dataset is very intuitive and minimalistic as it only requires the user to press a key to review the data, which for the nth sample, comes with text and bounding boxes.

The model for detecting vehicles in pedestrian zones is under training. In this phase, we have made several iterations where several object detection models have been used to generate an intelligent model with an accuracy higher than 85%.

REFERENCES

[1] Naciones Unidas (UN), “Health – United Nations Sustainable Development.”

[2] Naciones Unidas (UN), “Ciudades – Desarrollo Sostenible.”

[3] OMS, “OMS | Accidentes de tránsito.”

[4] C. O. WHO, “World health organization,” Diabetes, 2020.

[5] Mineduación, “Colombia ratifica ante la ONU su compromiso para reducir siniestros viales desde la infancia - Ministerio de Educación Nacional de Colombia,” 2019.

[6] A. de Cartagena, “Plan de desarrollo cartagena 2020-2023 – sed cartagena.”

[7] T. O. B. Fram, “Puntos críticos de las principales vías terrestres de cartagena y su incidencia en la movilidad para la búsqueda de una ciudad inteligente,” Revista Cultural Unilibre, no. 1, pp. 65–84, 2018.

[8] E. U. Cartagena and R. Goméz-Stand, “Así se puede contrarrestar el uso indebido y apropiación de andenes,” 2021.

[9] J. prensa DATT, “DATT, Disminuyó accidentalidad en un 2,3% durante primer semestre de 2019,” 2019.

[10] A. Fascioli, R. I. Fedriga, and S. Ghidoni, “Vision-based monitoring of pedestrian crossings,” in 14th International Conference on Image Analysis and Processing (ICIAP 2007), pp. 566–574, 2007.

[11] J. Hariyono and K.-H. Jo, “Detection of pedestrian crossing road,” in 2015 IEEE International Conference on Image Processing (ICIP), pp. 4585–4588, 2015.

[12] C. Zhao and X. Chen, “The study of pedestrian re-identification with the illumination change,” in 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), pp. 133–137, 2018.

[13] [13] M. A. Alias, S. N. Sulaiman, I. S. Isa, R. Boudville, and Z. H. C. Soh, “Detection of sudden pedestrian crossing for driving assistance systems,” in 2020 1st International Conference on Information Technology, Advanced Mechanical and Electrical Engineering (ICITAMEE), pp. 220–225, 2020.

[14] MATLAB & Simulink, “Segmentación semántica .” Humansintheloop.org, “Humans in the Loop - Continuously better models with a human in the loop.”

[15] ITELLIGENT, “Deep learning & Convolutional Neuronal Network: qué es y en qué consiste.”

[16] A. de Medellín, “Cámaras de CCTV.” https://www.medellin.gov.co/simm/camaras-de-circuito-cerrado, 2013. [Online; accessed 11-December-2020].

[17] Hernández-Díaz, Nicolás, Puertas, Edwin, Martinez-Santos, Juan Carlos, Archbold, George, Rios, Yuliana, & Peñaloza, Yersica. (2023). Dataset for Detecting Motorcyclists in Pedestrian Areas (1.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7935299.