Revista Científica de Pesquisa Aplicada à Engenharia | REPAE 2026
Brasília, Distrito Federal, Brasil

PROOF OF CONCEPT OF A LOW-COST TOOL FOR DATA COLLECTION AND TRAFFIC CONFLICT IDENTIFICATION AT CROSSWALKS

Samuel Seixas Martin*
Universidade de Brasília/Departamento de Engenharia Civil e Ambiental/Brasília/Brazil

Fabiana Serra de Arruda
Universidade de Brasília/Departamento de Engenharia Civil e Ambiental/Brasília/Brazil

ABSTRACT
Urban traffic conflicts pose significant safety risks and present real-time monitoring challenges in increasingly complex environments. Addressing these issues, especially at pedestrian crossings, this study proposes an Artificial Intelligence-based video analysis tool that can be deployed immediately and at low cost for integrated data collection, as literature suggests. Using Python, our program applies object detection, filtering, and temporal tracking with physical marker correlation to count vehicles and pedestrians, measure their speeds, and categorize interactions into risk levels. In a real-world pilot on the University Campus, the tool accurately identified speed violations and high-risk crossing events. Its performance surpasses traditional manual data collection methods in both technical precision and economic efficiency, provided the camera has an unobstructed view. There are vast possibilities for this solution, which enables traffic planners to detect and mitigate conflicts with data-driven tools for the safer traffic that all human beings inherently deserve.
KEYWORDS: *Computer Vision, Artificial Intelligence, Traffic Conflicts, Pedestrians, Vehicles, Road Safety.*

1. INTRODUCTION

Conflicts: why do they exist? Where do they come from? Would it be possible to live in a world without conflicts? In an increasingly connected society, it makes no sense to continue using outdated techniques to try to solve increasingly complex problems. Pedestrian-related crashes have been increasing for the last two decades in the U.S., which illustrates how the increase in the number of intersections leads to more traffic conflict points in cities. In the past, it would have been humanly impossible to know what is happening at each of them in real-time; however, with the Internet of Things (IoT) combined with Artificial Intelligence (AI), this becomes not only feasible but truly possible. (de Oliveira, Cunto, 2023; Zhang, Abdel-Aty, 2022)

The current work is exploratory, and its involved technology has a vast potential to disrupt many areas of our lives and even behave in unforeseen ways. That is why we must remember the Civil Engineering oath in Brazil, which urges everyone pursuing this field to "put all scientific knowledge at the service of the comfort and development of humanity" since we work "for the good of man, not of the machine." (UDESC, 2016) In popular culture, there is a saying that "with great powers come great responsibilities," and that is why the use of this technology must adhere to the strictest ethical standards to avoid any use that is detrimental to the human beings who appear in film footage, in compliance with the provisions of the Brazilian General Law on the Protection of Personal Data (LGPD).

Data from SES/DF (2024) indicate that Ceilândia, Taguatinga, and Plano Piloto are the Administrative Regions of the Federal District of Brazil with the highest number of pedestrian accidents and that since 2015, pedestrians who were run over account for approximately 35% of all deaths related to traffic accidents. Furthermore, using data from the Ministry of Health, the Brazilian Association of Traffic Medicine states that in 2023, 39,125 pedestrians were admitted to public hospitals in Brazil due to severe traffic accidents, second only to the number of motorcyclists hospitalized. This number represents an average of one pedestrian hospitalization in serious condition every 15 minutes. Although impossible, the goal is to eliminate this statistic; however, there is ample opportunity to drastically reduce the number of collisions resulting from traffic

*Autor correspondente: samuelseixasmartin@gmail.com

http://dx.doi.org
ISSN: 2595-0002

conflicts in urban centers, combining the best that technology has to offer in this regard with all the expertise of traffic authorities.

To prove this concept, we developed a computer program (software) using the Python programming language to achieve the proposed objectives. However, it is essential to say that the tool is not more important than the main objective; in the same way that even pliers can tighten a screw, the way in which someone solves a problem is not as important as the existence of a solution. From now on, it must be known that the identification, alert, and quantification of information concerning traffic conflicts is not only possible but that a scalable solution already exists for it.

This study proposes using a real-time video analysis tool to identify traffic conflicts, count traffic, and measure speed for both people and vehicles. Although the general use of AI is increasingly common, this work proposes a highly technical and low-cost proposition that can be immediately implemented for data collection, allowing effective measures to be taken.

2. LITERATURE REVIEW

2.1 Technical implementations of AI for Pedestrian Detection and Traffic Conflicts

Obviously, computers do not have eyes like humans and, therefore, cannot literally "see." However, making computers perceive the real physical world is still possible; this is why Computer Vision (CV) is such a relevant area within AI research and has received so much attention in recent times, even more so because it is broadly applicable in areas that go from cancer research to queue optimization. However, because the present paper results from applying a scientific computing tool in an Engineering context, one should use a technical approach from both areas.

On the computer science end, there has been a notable advance in computational power and modern algorithms. Mao et al. (2017) exemplify this in their report on how technical improvements in Convolutional Neural Networks (CNNs) aimed to improve pedestrian detection through finer detection of human features in an image. This reduced the number of false positives, such as saying that a tree is a human being or, conversely, finding the correct number of people in a group. At that time, there were already solutions being executed at 15 FPS (Frames Per Second) with an associated error rate of around 25%. Furthermore, Brunetti et al. (2018) highlighted the two most promising work fronts in the research of more efficient methods for pedestrian detection and classification, which explored the agility versus precision paradigm. These were, namely, high frame rate detection (100–135 FPS) on a low-resolution image (480x640 pixels) and Deep Neural Networks which are excellent at tasks that do not require instant classification but still do not produce results that can be applied in real-time.

On the other hand, more robust Engineering analyses were enabled through a large amount of real data, which would otherwise be impractical to make sense of. De Oliveira and Cunto (2023) analyzed more than 700 conflicts (called by the authors vehicle-pedestrian interactions) over 32 hours of recording at pedestrian crossings with and without traffic lights. They used a method that considers the distance between the vehicle and the pedestrian and their speed to calculate the probability of an accident occurring and concluded that 13% of these interactions were serious. As individual behavior changes according to their nationality, this Brazilian paper is important to establish a local reference.

Additionally, researchers such as Zhang and Abdel-Aty (2022) proposed, the previous year, a computational tool that predicts conflicts between pedestrians and vehicles one traffic light cycle before they occur, thus allowing preventive measures to be implemented. In the article, the authors talk about dynamically adjusting cycle times and alerting vehicle drivers to be more careful, but this is just a small glimpse of all this tool has potential. In a Traffic Engineering analysis, these researchers concluded that the four most important factors for increasing the risk of conflicts at intersections are: first, the volume of vehicles turning right; second, the volume of vehicles turning left; and, only in third and fourth place, the speed and volume of the main road.

However, there are still problems that not even AI can solve. Dollar et al. (2012) said there are two main difficulties for CV: the obstacles in the urban environment that partially hide pedestrians and the decrease in the resolution of the pedestrian's image as he moves away from the camera, both of which are physical issues related to the camera infrastructure.

2.2 Image Datasets for Training Computer Vision Models for Pedestrian Detection

This introduction to the computational approach to traffic conflicts leads us to the second paradigm within the world of Computer Vision: "how to teach the computer to recognize what I want"; for this, sets of data, or datasets, as they are known worldwide, are used. They are systematizations of cataloged data with related properties. Ronald Fisher created the oldest and most pioneering data set in this context in 1936, which related different properties of flowers to their colors. The analysis of this systematized data made it possible to infer the color of flowers based only on the physical measurements of the width and length of their petals and sepals, without ever even observing their color. This is the power of a well-organized data set.

In the context of person and vehicle recognition, some well-known datasets date back to 2013. One of the pioneers in this area was GRAM Road-Traffic Monitoring (GRAM-RTM), developed by Guerrero-Gomez-Olmedo et al. (2013), which, however, only had categorized videos of motor vehicles, such as cars, vans, and trucks. The following year, Urban Tracker was published by Jodoin et al. (2014), a Canadian initiative that, for the first time, combined pedestrians and vehicles in an environment that the researchers called "Urban Mixed Traffic". Also in 2014, the Ko-PER dataset was released, which, unlike the others, experimented with laser scanning of traffic circulating through an intersection, combined with a set of cameras that allowed data to be cross-referenced for a better perception of the spatial movements captured by the laser. The idea was excellent; however, this model was partially discarded due to the perception that, despite the process being more laborious, and financially and computationally more expensive, there were no significant improvements compared to the model with only cameras.

After a few years of stagnation in the work, 2018 saw a revolutionary launch: the JTA dataset from the University of Modena and Reggio Emilia, which used a realistic computer game to record scenes of pedestrians walking on sidewalks. Although it is a very large, high-resolution urban monitoring dataset used in several areas within CV, it is irrelevant to the present work because it only emphasizes pedestrians on the sidewalk, without conflicts with other means of transportation. In the same year, the WILDTRACK dataset was also published, with the identical focus on pedestrians specifically on sidewalks.

According to Dollar et al. (2012), one of the most significant academic challenges in this area at the time was measuring scientific progress in Computer Vision, since the datasets and algorithms used were difficult to correlate. Therefore, the publication of The Multiple Object Tracking Benchmark (MOTChallenge) was of crucial help to standardize the results. It comes from a partnership between major European universities and an Australian university, focusing on pedestrians in real urban environments including scenes with pedestrians crossing the street, near cars, and in the middle of the street during closed events, among others. It is the result of work that has been going on since 2014, and which, in 2020, during the period known as the pandemic, reached its peak with the launch of the MOT20 version.

3. WORK METHOD

To develop the program, it was necessary first to configure auxiliary tools and software to create a well-designed work environment for efficient, comfortable, and documented programming.

3.1 Nomenclature and Definitions

This program can process images individually (frames) or in sequence (videos). For a video, the program processes it frame by frame, individually, but not in isolation. Thus, for each video frame, the program first performs its isolated analyses on the image, detecting everything it can, and then compares it with the previous frames to establish a relationship of continuity for each detection, which allows the program to track them. The identification of any person or vehicle by the program will be generically referred to below as an "object" or "entity," even though it is known that people are not objects.

3.2 Desktop Setup

A Python interpreter, a text editor, and a console are required to program and execute the software. Therefore, the Python interpreter was installed on the local machine (computer), and all the necessary modules using the Anaconda distribution, which already manages compatibility between modules. Visual Studio Code was chosen for the text editor, which, in an integrated manner, allows writing, executing, and debugging the code.

3.3 Camera Positioning and Definition of Control Areas

This work deals with an interface between the physical world and the virtual world; it is, therefore, necessary to carry out a sequence of actions in order to be able to make the connection between points in the physical world and the virtual world. First, we must prepare the terrain in which the work will be conducted by choosing the positioning of the physical markers to be visible in the camera's field of view. Then, we must position the camera so that it is possible to see the entire area to be surveyed, with a small amount of spare space. The video can then be recorded once the camera's framing in the area in question has been confirmed. Once the video recording is finished, or when real-time monitoring begins, the images are pre-processed based on the definition of key points in the camera image; these are the polygons, which are called control regions from this point forward. It is indispensable that the camera remains still and at the same point and that the polygons are already defined before starting the next stage.

In the current implementation of the program, it is necessary to define three polygons that encompass the control regions, from smallest to largest: the crosswalk, the street, and the source region. The source region is the area in which the physical world is connected to the virtual world; it is the region in which speeds are measured, and detections are validated; detections outside the source region are discarded. For each of the polygons, it is necessary to indicate its position in the video capture in terms of pixels, and, only for the source region, it is also required to inform its measurements in the real world in rectangular coordinates (width and length), which correspond precisely to the coordinates of the physical markers. The program makes the correlation between the real and virtual polygons of the source region employing a "perspective transformation," which uses complex mathematical equations already implemented in one of the libraries used.

3.4 Program Execution

Once all the initial settings have been made, it is time to run the program. Its default setting is to save videos in real-time after they have been processed, creating a large video file that can be re-watched and calmly analyzed after capture. Another possibility is to use a camera disconnected from the program, which records a video that will only be processed later; this can be done by indicating a video file to the program.

3.5 Premises Used

As with all research work, arbitrary decisions must be made that invariably direct this work toward premises underlying the researcher's will. Knowing this, some of these choices are briefly reported so that the reader knows there are other possibilities.

3.5.1 Arbitrary Choices in the Algorithm

The choices made by the authors in each listed area are described below.

Polygons: The polygons chosen for the source region were right rectangles. For the other polygons, four vertices were used, not necessarily rectangular. However, the program is flexible enough to allow any number of vertices for any polygon to better adjust for the real geometry of the region. For the source region, it is only necessary to ensure the real-virtual pair has the same number of vertices.

Model: We used the Ultralytics module for Python, which, as a free and open-source repository, allows its use for the purposes of this research. The YOLOv8 model was chosen within this module because of its ease of use. This model is designed to be used in a wide variety of real-world situations and is therefore trained to recognize several classes of objects, which makes the model less efficient for use with known and restricted classes, as in the case of this research, which considers only pedestrians and some vehicles.

Reliability: The model used to process the video uses a parameter to decide the classification of the detected object. It was decided that only detections with more than 30% confidence would be considered valid. This value may seem low, but it is used to avoid losing the object when it is partially hidden in the image, as previously discussed.

Occlusions: The model's IOU ("intersection over union") parameter focuses on eliminating duplicate detections for the same object. A value of 0.7 was chosen for it. In short, this means that if the intersection area between two detections is equal to or greater than 70% of their combined area, they are treated as a single detection.

Speed calculation: After processing the frame, the program reports a calculated speed for each identified object. However, it is not an instantaneous speed; instead, it is calculated as a continuous average of the distance traveled in meters by the object in the last 1 second and converted to kilometers per hour (km/h). Thus, this speed is always "delayed," with the delay being directly proportional to the object's acceleration. This choice attempts to solve the problem of a constant but relatively small variability in the position the computer interprets the entity to be in. Thus, although it was possible to report a more precise speed, an interval of one second was chosen for a smoother average.

Legend: Below each identified entity is a legend composed of serial number and speed. The symbol # comes before the serial number, and after the speed comes its unit, in km/h.

3.5.2 Category Conventions

We decided to classify the detections into three categories: normal, warning, and danger. They are arbitrary but were chosen to facilitate the analysis of possible risks without calculating the actual degree of risk, since the literature in section 2 refers to degrees of risk (high, medium, and low) calculated from a correlation matrix between the identified factors. In the case of this research, the categories are purely spatial, depending only on the object's position around other objects, and are classified independently of the model classes (people, cars, buses, etc.), which are intrinsic to the objects.

In "danger," we placed three situations: all the jaywalkers, which are pedestrians who cross the street outside of a crosswalk, as well as all vehicles traveling outside the street, or all objects when there is at least one pedestrian and at least one vehicle in a crosswalk simultaneously. This last situation happens when, for example, a car invades the pedestrian crossing and there is still a person crossing, or when a person starts to cross while there is still a car passing by. It also considers the elderly and children, more vulnerable people, who may take longer in the crossing or impulsively cross the street outside the pedestrian crossing to catch a ball, for example.

In the "warning" class, we placed all people and vehicles on the pedestrian crossing but not simultaneously with objects from another category together (for example, only people in the crossing or only vehicles inside the crossing). This was done to differentiate them from detections of pedestrians on the sidewalk, which inherently have a lower risk.

All other situations not covered by the previous categories were placed in "normal." For example, pedestrians on the sidewalk or vehicles on the street.

3.5.3 Meaning of the Colors

The program uses only the following colors to overwrite the original image, each used in a different context. For polygons, green symbolizes the crosswalk, blue symbolizes the source region, and black symbolizes the street. As for the categories:

Red: symbolizes danger. It is used when there are conflicts between people and vehicles.

Yellow: symbolizes alert. It is used when an element is within the crosswalk, after it is verified that it is not in the danger category.

White: symbolizes normality; vehicles on the street, and pedestrians on the sidewalk. It is used on elements and their associated captions when it is verified that they are neither in the alert nor in the danger categories.

4. RESULTS

After developing the computer program in Python to identify normal, alert, and dangerous situations, it requires a live or recorded video to work.

4.1 Data Collection

To initially test the program's behavior, we searched the internet for videos that were freely available and allowed to be used with attribution; however, due to the enormous quantity of results retrieved, some parameters were introduced to search for the ideal video that would test all the functionalities of this work. These are: having a clear and unobstructed view of at least one pedestrian crossing, and of an intersection, which may or may not be traffic lighted; having vehicles and people being captured by the camera; and, perhaps the most difficult parameter to comply with, being a video that was recorded with the camera still. Two videos were found that fully met these criteria: Morina (2020), and Korb (2020), which were used as a basis for the initial tests of the program and will be further presented in their modified versions. Once the testing purposes were fully met, the author recorded a video on Monday, September 23, 2024, around 2:00 p.m., on the street in front of the University Restaurant at the Darcy Ribeiro Campus of the University of Brasília. This video was recorded with a tilted view, from top to bottom, but not completely vertical, as seen in Figure 2, which is the best framing for this CV situation, as it allows a less obstructed view of the entities passing by.

4.2 Data Processing

Initially, videos freely available by Morina (2020) and Korb (2020) were used to test the program's operation. Later, the author used footage taken in front of the University Restaurant as a real-life example to test its behavior in a realistic scenario.

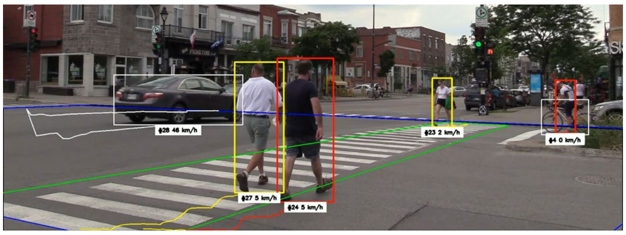

Figure 1. Several detection categories occurring simultaneously.
Source: Korb (2020) - adapted by the authors.

Figure 1 is very interesting because it simultaneously depicts a variety of situations incorporated into the classification algorithm. It is possible to identify at least five colors present in the image: red, yellow, white, green, and blue, and, perhaps the most hidden, black. Six detections can be observed from left to right, which will be treated according to their unique serial numbers. In the first, it is possible to observe a gray car framed with the white color and the caption "#26 46 km/h." This label represents that this object was identified with serial number 26, had its speed measured at 46 km/h, and, according to the criteria in section 3.4, explained previously, was classified as "normal," thus receiving the color white, since, being a car, it is on the street, and not on top of the crosswalk. In addition, behind the car is a trail of its position in the last 2 seconds as it moves across the screen, which is common for all detections.

Similarly, the following two cases received colors that represent their categories according to the program's analysis: detections 27 and 24 are of two men walking side by side; the man in yellow is in the "alert" category because he is on the crosswalk, and the man in red next to him is in the "danger" category because he is walking on the street, even though he is next to the crosswalk, but not in it. The subsequent detection, serial number 23, is of a woman who enters the crosswalk on the opposite side of the road and walks toward the two men: note that the detection model used in the program is well-trained enough to recognize people in different positions and angles; this also happens with vehicles, as can be seen in the following examples.

Finally, the last case is tricky because it involves a cyclist: typically, a human being thinks of a cyclist as a combination of person and bicycle; however, the program divides them into two separate entities. This is because this specific model only separately knows what a bicycle and a person are; that is why the person is framed in red ("person on the street"), and the bicycle is in white ("vehicle on the street").

Figures 2 and 3 show the vehicle and the pedestrians simultaneously occupying the crosswalk (marked in green), which should not occur. Therefore, as discussed in section 3.4.3, the algorithm determined that both the vehicles and pedestrians in the crosswalk would be colored red. This color was chosen to indicate visually that all involved are in a dangerous situation: the pedestrians because they are in real danger of being run over, and the vehicle because it is a real risk to the safety of the pedestrians.

Undoubtedly, there are situations in which a vehicle has safely entered the crosswalk area without people on the crosswalk, and then a pedestrian begins to cross the crosswalk, which can be one of the interpretations of Figure 2. Nevertheless, it is still classified as a dangerous situation. However, most cases are like the one in Figure 3, in which the pedestrians are already on the crosswalk, and the vehicle nevertheless begins to cross it.

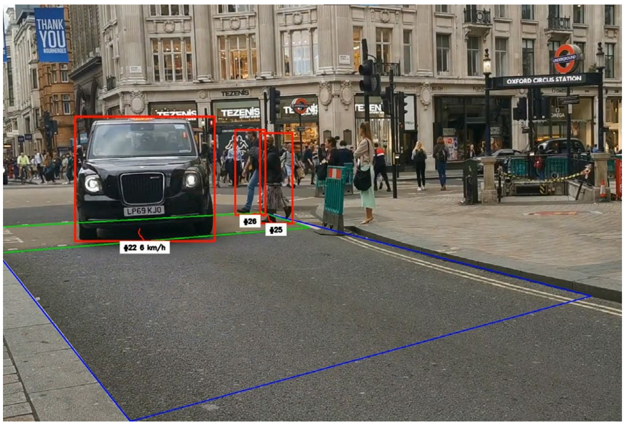

Figure 2. Identification of conflict between vehicles and pedestrians.
Source: Morina (2020) - adapted by the authors.

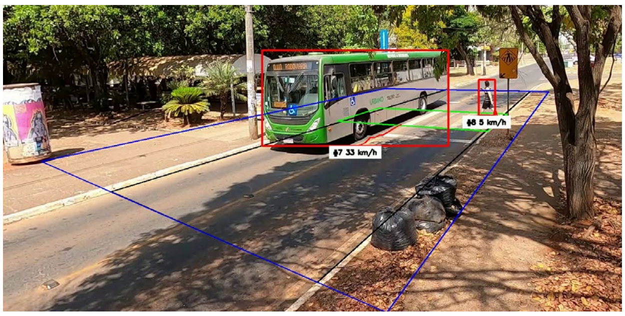

Figure 3. Identification of conflict between vehicles and pedestrians.

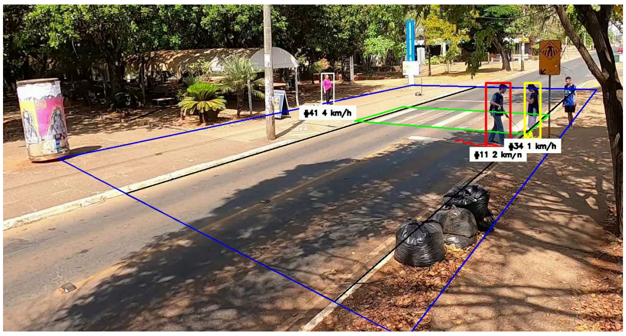

Figure 4. Identification of street and crosswalk occupancy with their respective colors.

In Figures 4 and 5, we can see the different categories in which pedestrians and vehicles can be classified, depending on their position relative to the street and the crosswalk, with the crosswalk marked in green, the street in black, and the source region in blue. As Figure 4 shows, pedestrian detections without the presence of cars are correctly marked as dangerous when outside the crosswalk but inside the street, alert when inside the crosswalk, and normal when outside the street (on the sidewalk). The fourth pedestrian in Figure 4, furthest to the right of the image, is not identified in this video frame because the midpoint of his detection is slightly outside the source region; however, he is considered after he enters the source region. As explained in section 3.2, this happens because detections outside the source region are discarded.

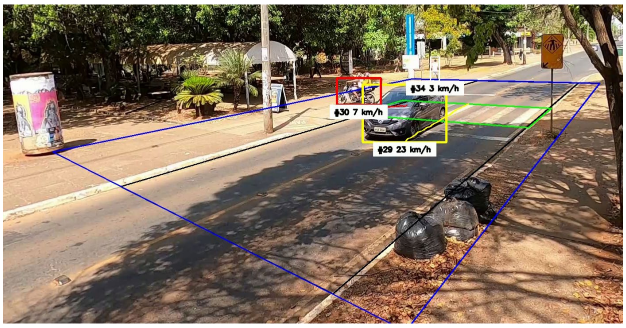

Figure 5. Identification of street and crosswalk occupancy with their respective colors.

Figure 5 shows that vehicles are also correctly assigned as dangerous when off the street, alert when inside the pedestrian crossing, and normal when inside the street; the latter situation is seen in Figure 1. The program also recognizes buses as vehicles, not just cars; trucks are also correctly identified, although not shown in this report.

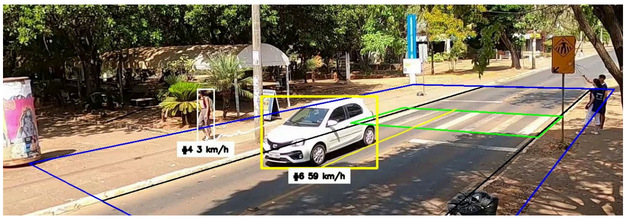

Figure 6. Measurement of speeds on the road.

In addition, Zhang and Abdel-Aty (2022) state that vehicle speed contributes significantly to the risk of conflict and increased accidents, even more so on a straight segment like the one in front of the University Restaurant, where the first two factors of the aforementioned study, which concern intersections at the turn, are not present, making speed the most important factor in conflict mechanisms.

The Brazilian metrology agency, INMETRO, is responsible for approving speed cameras used for issuing speeding tickets. Normally, speed cameras use copper coils embedded in the pavement that electrically detect moving iron, such as the motors of the vehicles. However, this proof-of-concept software enables the collection of realistic data without meddling with the transportation infrastructure, such as the pavement. Although INMETRO has not yet approved the speed measurement technique used in this work and, therefore, it cannot be used to issue speeding tickets, the software can be used for identifying critical points in cities with increased risk of accidents and conflicts. In the example in Figure 6, it is notable that the region where the video was filmed is within "Zone 30", where the regulatory speed limit of the road is 30 km/h, and the vehicle is traveling at an excess of almost 100% of this speed.

Finally, after analyzing the 15 minutes of video recorded in the field, the software counted 54 vehicles, 90 pedestrians, and 54 conflicts. This results in approximately 3.6 conflicts per minute and a rate of 37.5% in relation to the total number of vehicle and pedestrian detections in the video.

4.3 Identifying the Program's Positive and Negative Aspects

As the authors also participated in the manual counting of pedestrians and vehicles crossing an intersection, it was possible to have a critical view when comparing the two methods.

On the one hand, counting pedestrians and vehicles using computational methods is undoubtedly more efficient in using resources, whether time or money. Yes, the initial cost is much higher; however, it is diluted as the research period increases or with a high traffic volume. In addition, it is also possible to use parallel counting of several fronts with a single piece of equipment; there are videos on the internet in which a camera with an aerial view can count the vehicle traffic in all directions of a roundabout, something that would not be possible in the manual form of work without allocating several researchers. Furthermore, the error rate can be considerably reduced by using customized models and subsequent processing, since manual counting is often done without video recording, with no possibility of recounting, and, even with the analysis of recorded videos, there is still the effect of burnout on employees due to repetitive work.

On the other hand, the CV counting method requires special care and has limitations that do not exist in manual counting. As previously mentioned, the equipment for analyzing the data is quite expensive, with the quality of the video being almost directly proportional to the camera's price and its related add-ons: infrared functions for seeing at night, for example. Besides the camera, powerful computers are still needed to process the generated videos, which are also expensive, but now do not influence the quality of the research as much: a cheaper computer will take longer. In addition, it is impossible to use the cameras in bad weather conditions; not necessarily because it is not possible to use a camera in adverse situations (the GoPro works, for example), but rather because sufficient brightness and sharpness are required to discern the objects present in the video, something that does not happen at night, in rainy or in foggy periods. The weather can also physically damage the equipment and influence the lenses' cleanness, which can blur the images and make them unusable. Finally, the error rate of computational methods depending on the model used is still considerable.

The conclusion is that Computer Vision remains an imprecise method, but it provides an excellent estimate of the magnitude of the traffic volume. This is important for continuous monitoring of intersections so that timely measures can be taken based on instantaneous information. This is the main advantage of data collection with Artificial Intelligence processing: real-time information.

5. CONCLUSIONS

Therefore, the examples shown demonstrate the feasibility of implementing a low-cost, high-tech system for large-scale traffic counting, speed monitoring, and conflict detection, enabling accident mitigation. It no longer makes sense to treat intersections in increasingly connected cities as they did 50 or even 10 years ago: technology is an ally of human beings in preserving and improving quality of life, and it should be used as such.

While this proof-of-concept software works fully at any intersection where a camera can be attached, its optimal performance depends on an unobstructed, sharp view of the intersection. Despite this condition, the real-life application demonstrated the substantial technical and economic advantages of this proof-of-concept over manual data collection methods. The software efficiently performs multiple tasks simultaneously and collects diverse, integrated data that would typically require separate equipment and considerably more time and resources.

The integrated data, encompassing number of pedestrians, vehicle types and counts, speed, precise positioning relative to the road and the crosswalk, and conflict indicators, enables robust, continuous analysis of human-vehicle interactions, 24 hours a day. The 15-minute example showcased a high rate of conflicts, especially when considering vehicles invading the sidewalk, which represents both a threat to the drivers and especially to the pedestrians. This highlights the substantial number of uncatalogued conflicts that continuous monitoring could identify.

While urban surveillance raises serious privacy concerns, this research advocates strategically instrumenting critical intersections, which presents a technically sound and economically viable approach. Ultimately, by providing insights that enable the mitigation of traffic conflicts, this technology directly contributes to the fundamental human right to safety and the preservation of life in our shared urban spaces. In the dynamic environment of urban traffic, minimizing conflicts is not merely a technical challenge but a moral imperative, essential for preserving human life and enhancing the quality of urban living.

In the future, it is recommended to improve the detection of cyclists through a model specifically trained to recognize the person and their bicycle as a whole instead of separating them. Also, alternatives to deal with obstructed lenses must be sought.

ACKNOWLEDGMENTS

This study was funded by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq Work Plan PIBIC 9726A). The authors are also grateful to the Python Software Foundation and the developers of the other modules used in the program, as well as the University of Brasília and the Civil and Environmental Engineering Department for the world-class formation. To the Author and Sustainer of all life, our God and Father, be glory forever and ever. Amen.

REFERENCES

Associação Brasileira de Medicina do Tráfego (ABRAMET) (2024). *Quase 8 em cada 10 vítimas graves do trânsito são pedestres, ciclistas ou motociclistas*. Available at: https://abramet.com.br/noticias/quase-8-em-cada-10-vitimas-graves-do-transito-sao-pedestres-ciclistas-ou-motociclistas/. Accessed 19 September 2024.

BRUNETTI, A. et al. (2018). Computer vision and deep learning techniques for pedestrian detection and tracking: A survey. *Neurocomputing*, 300, ISSN 0925-2312, p. 17-33. Available at: https://doi.org/10.1016/j.neucom.2018.01.092.

DE OLIVEIRA, B. and CUNTO, F. (2023). Caracterização da interação pedestre-veículo em travessias semaforizadas. *TRANSPORTES*, v. 31, n. 3, p. e2768. Available at: https://doi.org/10.58922/transportes.v31i3.2768.

DENDORFER, P. et al. (2020). *MOTChallenge: A Benchmark for Single-Camera Multiple Target Tracking*. Available at: https://doi.org/10.48550/arXiv.2010.07548.

DOLLAR, P. et al. (2012). Pedestrian Detection: An Evaluation of the State of the Art. *IEEE Journals & Magazine*. Available at: https://doi.org/10.1109/TPAMI.2011.155. Accessed 19 September 2024.

GUERRERO-GOMEZ-OLMEDO, R. et al. (2013). Vehicle Tracking by Simultaneous Detection and Viewpoint Estimation. Natural and Artificial Computation in Engineering and Medical Applications. *IWINAC*; p. 306-316. Available at: https://gram.web.uah.es/data/publications/iwinac13-guerrero.pdf. Accessed 19 September 2024.

JODOIN, J. et al. (2014). Urban Tracker: Multiple Object Tracking in Urban Mixed Traffic. *IEEE Winter Conference on Applications of Computer Vision*, Steamboat Springs, CO, USA; p. 885-892. Available at: https://www.jpjodoin.com/urbantracker/download/JodoinetAlWACV2014.pdf. Accessed 19 September 2024.

KORB, G. (2020). Footage of a Busy Street. Quebec, Canada. Available at: https://www.pexels.com/video/ footage-of-a-busy-street-4779862/. Accessed 19 September 2024.

MAO, J., et al. (2017). *What Can Help Pedestrian Detection?* Available at: https://doi.org/10.48550/arXiv.1705.02757. Acessado em: 19/09/2024.

MORINA, G. (2020). View of Busy Street at Oxford Circus Station. London, United Kingdom. Available at: https://www.pexels.com/video/view-of-bust-street-at-oxford-circus-station-5330919/. Accessed 19 September 2024.

Secretaria de Saúde do Distrito Federal (SES/DF) (2024). Dia do Pedestre traz conscientização sobre segurança no trânsito. Available at: https://www.saude.df.gov.br/w/dia-do-pedestre-traz-conscientização-sobre-segurança-no-trânsito. Accessed 19 September 2024.

STRIGEL, E. et al. (2014). The Ko-PER intersection laserscanner and video dataset. *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Qingdao, China, p. 1900-1901. Available at: https://www.uni-ulm.de/fileadmin/website_uni_ulm/iui.inst.110/Bilder/Forschung/Datensaetze/20141010_DatasetDocumentation.pdf. Accessed 19 September 2024.

UDESC (2016). Juramentos e cores dos cursos. Available at: https://www.udesc.br/guiadeeventos/juramentos. Accessed 19 September 2024.

ZHANG, S. and ABDEL-ATY, M. (2022). Real-Time Pedestrian Conflict Prediction Model at the Signal Cycle Level Using Machine Learning Models. *IEEE Open Journal of Intelligent Transportation System*, 3, p. 176-186, Available at: https://doi.org/10.1109/OJITS.2022.3155126.