A Novel Educational Approach for Volumetric Vehicle Counting Using Artificial Intelligence in Smart Cities

Bruno Sousa Dias
Institute of Urban Attention
Federal University of Tocantins
Palmas, Brazil
email: bsdias42@gmail.com

Nailson Martins Dantas Landim
Institute of Urban Attention
Federal University of Tocantins
Palmas, Brazil
email: nailsonlinux@gmail.com

Humberto Xavier de Araujo
Institute of Urban Attention
Federal University of Tocantins
Palmas, Brazil
email: hxaraujo@mail.uft.edu.br

Yasmin Barbosa de Carvalho
Institute of Urban Attention
Federal University of Tocantins
Palmas, Brazil
email: yasminbc@gmail.com

Starley do Nascimento Lobo
Institute of Urban Attention
Federal University of Tocantins
Palmas, Brazil
email: starley.lobo@mail.uft.edu.br

Thiago Campos Silva
Institute of Urban Attention
Federal University of Tocantins
Palmas, Brazil
email: campos.thiago@mail.uft.edu.br

Abstract — **This article presents a solution for volumetric vehicle counting using artificial intelligence based on images of specific areas of interest in smart cities. The focus is on collecting data to support urban mobility planning. It includes the development of equipment for collecting images, easy to install with its own power generation, development of software for data processing, covering the identification, tracking and counting of vehicles, to identify the detailed traffic flow with the separation of vehicles by their respective categories: cars, motorcycles, trucks and buses. In this way, the solution developed in this work can directly contribute to the planning of mobility in urban areas through the automated collection of traffic flow data. This solution was developed within a university environment and can also be used in practical educational approaches as an applied framework for learning in artificial intelligence and computer vision, with direct application to smart city contexts.**

Keywords: Volumetric vehicle counting; Urban mobility; Smart city.

I. INTRODUCTION

The term "smart cities" is commonly associated with the application of technology in the urban environment; however, its concept is broad and extends beyond this notion. Cities globally have been exploring methods to address contemporary challenges through solutions that yield long-term economic benefits [1].

The upward trend in human population growth has rendered sustainable urban mobility a significant challenge for municipal authorities [2]. With social and economic development in urban areas, issues such as congestion and increased pollution emerge. To implement effective improvements in this context, it is crucial to understand traffic demand through indicators such as traffic flow [3].

This study presents an Artificial Intelligence (AI)-based solution to acquire traffic flow data on urban roads through volumetric vehicle counting, utilizing images captured in designated zones of interest. To facilitate image acquisition, a custom, easy-to-install station powered by solar energy was constructed.

The image processing component predominantly focuses on the detection of objects of interest, specifically motor vehicles categorized by type. Following detection, each object must be tracked to establish its movement, enabling the software to perform counting. Subsequently, the collected data is recorded to create a historical database.

The following sections present the context and development of the system addressed in this work. In Section II, the need for traffic studies in urban mobility is discussed. In Section III, the techniques used for object detection and vehicle tracking in this approach are described. In Section IV, the system architecture and the equipment used are presented. The results obtained are discussed in Section V, and finally, the conclusions are presented in Section VI.

II. URBAN MOBILITY

In various cities around the world, the association between mobility and transportation has encouraged an increasing reliance on motor vehicles and a tendency towards the expansion of urban road networks. Consequently, road structures such as overpasses, bridges, tunnels, and pedestrian walkways have become common and characteristic elements of the modern city [4].

Urban mobility is a critical subject for city management as it pertains to the movement of goods and people. Thus, the analysis of traffic flow is necessary for public transportation planning, parking management, and other related aspects.

According to the Traffic Study Manual by National Department of Transport Infrastructure (DNIT), the quantity, direction, and composition of vehicle traffic flow in a road system over a specified period are determined through Volumetric Counting. This information is invaluable for capacity analysis, identifying the causes of congestion and accidents, as well as for pavement sizing and traffic channeling projects [5].

III. COMPUTER VISION

The demand for computer vision has grown over the years, focusing on solving real-world problems [6]. However, there are several challenges, as many models are deficient in handling adverse situations such as partial occlusion, low contrast, and changing environmental conditions [6]. The challenges and deficiencies highlighted in Sebe's 2005 work remain relevant even after nearly two decades.

You Only Look Once (YOLO) is an algorithm designed for object detection [7]. Its name is self-descriptive since detection occurs in a single stage; in its output, identified objects are displayed with rectangles around each one according to their class [8]. These rectangles are called bounding boxes. The entire image is utilized to predict each bounding box for all classes for which the model is trained; the input is divided into an SxS grid, meaning that if the center of an object lies within one of the cells, that cell will detect it. In each case, a confidence value is assigned to statistically determine which of the generated boxes best fits the detected class [7].

To train the model, although it is possible to create a custom dataset, gathering thousands of images and categorizing objects individually is a slow and labor-intensive process. On the other hand, ready-to-use datasets exist to meet various demands. Launched in 2014, the Common Objects in Context (COCO) dataset consists of 328,000 images, of which over 200,000 are labeled, covering 80 categories [9].

Object tracking becomes useful in certain applications, and for vehicle counting, it is indispensable since it is necessary to obtain information about the movement performed. Once detected in a frame, the object of interest is tracked in subsequent frames; this process is usually lighter than detection [10]. The choice of tracker must take into account the demand and available resources.

Simple Online and Realtime Tracking (SORT) is a method for tracking multiple objects. Proposed in 2016, it was created with an emphasis on real-time applications, considering only the current frame and the previous one for tracking [11]. The object's trajectory is predicted using a Kalman filter, and frame-by-frame data is associated using the Hungarian algorithm.

Although SORT is capable of tracking multiple objects, the fact that it considers only the current frame and its predecessor makes it susceptible to identification changes whenever occlusion occurs. To solve this problem, a derivative called DeepSORT was created, which improves data association [12]. DeepSORT adds information related to the object's appearance, so even if occlusion occurs, when the object reappears, tracking remains, thus maintaining the same identification [13].

IV. SOLUTION DEVELOPMENT

This section details the architecture, as can be seen in Figure 1, of the developed system, highlighting the primary hardware and software components. A prototype for image acquisition was constructed, with images later processed in a cluster equipped with NVIDIA GPUs, providing automated volumetric vehicle counting. Vehicles are categorized into cars, motorcycles, trucks, and buses.

A key component in this equipment set is the camera, specifically the Hikvision DS-2CD1123G0E-IC model used in this research. This device features a 2MP resolution, infrared LEDs with a range of up to 30 meters suitable for capturing nighttime images, an energy demand of up to 5W, and a memory card interface supporting up to 128GB [14]. Additionally, it has IP67 protection, making it resistant to dust and water, ideal for outdoor installations, and IK10 vandalism protection [15].

Based on these specifications, the image collection equipment was installed in a metallic hermetic box. This design protects the water-sensitive battery and charge controller. Besides protection, the metallic box was chosen for its durability, allowing it to be used as part of the structure. Thus, a metal support for ground installation was designed to be fixed directly to the bottom wall of the box. Similarly, a support was created for the photovoltaic module on top, adjustable in angle to fit its size.

To ensure the system is versatile and secure, all energy consumed is generated locally via a solar generator. Thus, no extra installations are required at the collection points. After collecting the images, the data is taken to the cluster, where it will be processed. Seeking to reduce costs and make logistics viable, edge computing was not used; in addition, post-processing allows recounting and subsequent adjustments.

To meet energy demands, a 20W photovoltaic module and a charge controller were included. The module-generated energy powers a Uninterruptible Power Supply (UPS) battery with a 7Ah capacity; the charge controller manages battery charging, preventing overcharging and overvoltage.

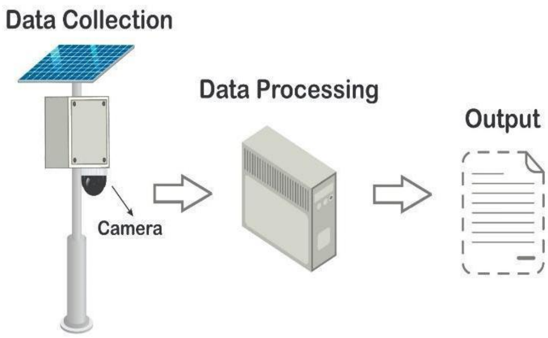


Figure 1. Solution architecture.

The camera was securely mounted at the bottom of the box using screws and nuts. To facilitate field installation and support, the previously mentioned metal bracket was employed. This setup enables the image capture system to be easily transported, adjusted, and deployed at various points of interest. Upon activation, the camera captures images and stores them on a memory card.

For image processing, the captured images are transferred to a cluster outfitted with six NVIDIA RTX 3070 GPUs, as depicted in Figure 2.

The algorithm for volumetric vehicle counting operates through a series of steps including detection, tracking, counting, and logging of objects of interest, such as cars, buses, motorcycles, and trucks.

For object detection, the YOLOv5 algorithm was employed. This version of YOLO offers a range of model variations, from YOLOv5n (nano) to YOLOv5x (extra-large) [16].

Upon detection, each vehicle is assigned a unique ID and tracked using the DeepSORT algorithm. The DeepSORT algorithm's robustness allows it to maintain tracking continuity even in the presence of short occlusions, preventing ID switches and enhancing accuracy in the volumetric vehicle counting process.

Additionally, a virtual grid system was implemented to precisely define the movement patterns of each vehicle, as illustrated in Figure 3. With this approach, it is possible to identify the traffic flow in different locations, based on the quadrants present in the image. To achieve this, it was necessary to separate the grid system into zones, so that the movement carried out by the vehicle is established by the zone in which it is first identified and the zone in which it disappears. This counting strategy is flexible, can identify complex movements in busy locations and eliminate unwanted movements. As an example, Figure 4 shows four zones with their respective movement in relation to the origin and destination.

Figure 2. GPU cluster for processing artificial intelligence software.

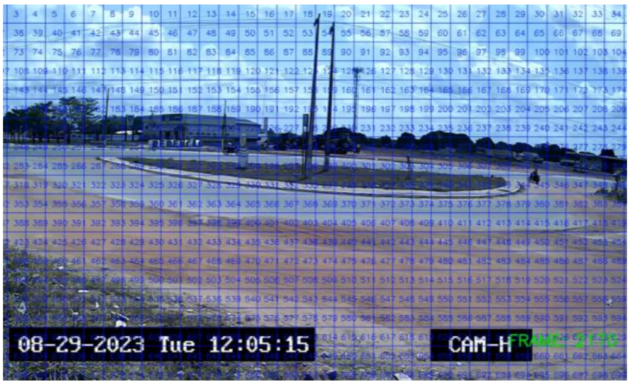
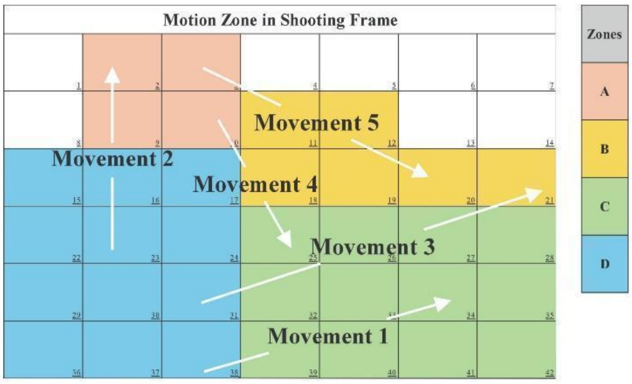


Figure 3. Virtual grids for identifying conversions.



	Movement 1	Movement 2	Movement 3	Movement 4	Movement 5
Origin Zone	D	D	D	A	A
Destination Zone	C	A	B	C	B

Figure 4. Zones delimitation.

Defining zones and including movements are separate processes from generating the analysis spreadsheet, so that traffic engineers have freedom in delimiting the scope of what will be evaluated. Furthermore, by considering each vehicle category individually, it is possible to identify peak hours in a segmented version. The data is generated in spreadsheet form, as this makes access to the data more palatable and opens space to calculate other metrics, create graphs, etc.

V. Results

A prototype for image capture was successfully developed at the Federal University of Tocantins. This setup enabled the validation of its continuous image recording capabilities, encompassing both camera functionality and the reliability of the self-sustaining power generation system. It was verified that image capture remained uninterrupted, with the generated energy effectively maintaining battery charge and system operation, including nocturnal periods.

Following operational validation, an additional eight image capture units were manufactured. These units were meticulously installed to proactively identify potential operational issues during a week-long continuous runtime. Six units performed as anticipated, seamlessly operating without interruptions as per the initial testing phase, while two necessitated battery replacements.

With all nine image capture units functioning at full capacity, they were transported to Redenção-PA, a city in the northern region of Brazil boasting a population exceeding 100,000 inhabitants. Strategically placed in areas characterized by high vehicular activity, one of the deployment units is shown in Figure 5.

Following the installation, all units underwent testing. By connecting a notebook directly to the camera using an Ethernet cable, it was ensured that the camera was effectively recording images.

Figure 5. Image acquisition system installed and operating.

Utilizing the implemented grid system, specific movement identification zones were established for each counting area based on the monitoring roadways' requirements, as depicted in Figure 5. Therefore, after processing, it was possible to obtain the traffic flow in each zone established by day and time, as shown in Figure 6. Complementarily, Figure 7 shows how the results are exported after the counting process.

Figure 6. Vehicles and their conversions identified through the AI algorithm.

date_time	Truck	Car	Motorcycle	Bus	Total
22/08/2023 06:00:00	0	4	1	0	5
22/08/2023 06:15:00	1	2	4	0	7
22/08/2023 06:30:00	0	5	4	0	9
22/08/2023 06:45:00	2	16	23	0	41
22/08/2023 07:00:00	1	13	12	0	27
22/08/2023 07:15:00	0	21	22	0	44
22/08/2023 07:30:00	0	19	18	0	37
22/08/2023 07:45:00	3	20	20	0	43
22/08/2023 08:00:00	1	21	15	0	37
22/08/2023 08:15:00	3	21	12	0	36
22/08/2023 08:30:00	4	20	14	0	38
22/08/2023 08:45:00	2	19	14	0	35
22/08/2023 09:00:00	2	18	12	0	32
22/08/2023 09:15:00	1	18	3	0	24
22/08/2023 09:30:00	3	22	10	0	35

Figure 7. Traffic flow in one of the observed zones.

Considering that the image capture stage is completely independent of processing, in challenging scenarios such as the city of Redenção – PA, the analyses can be repeated, if necessary, until all the desired criteria are observed, which would not be possible in a situation with local processing, without video storage.

VI. Conclusion And future work

This paper sets out to develop an artificial intelligence-based solution for volumetric vehicle counting in urban environments, aiming to support smart city mobility planning. To achieve this, an integrated system was designed combining a self-powered image acquisition unit with a computer vision pipeline for vehicle detection, tracking, and counting.

The proposed system for volumetric vehicle counting aligns with the concept of smart cities by integrating advanced technologies such as artificial intelligence, Internet of Things (IoT), and renewable energy sources. By implementing a smart image acquisition system powered by a photovoltaic generator, the solution promotes autonomous and sustainable operations, two key pillars in the development of smart cities. The software component, comprising vehicle detection, tracking, and data analysis, contributes to enhancing urban mobility planning and traffic management efficiency, vital aspects for the development of smarter and more interconnected urban environments.

The successful field testing in Redenção demonstrates the practical application of these smart city principles to improve traffic monitoring and city infrastructure, showcasing how innovative solutions can contribute to the advancement of urban living standards.

In addition to its technical contributions, the proposed system also contributes to the educational context, as higher education institutions play a fundamental role in training professionals capable of addressing smart city challenges and overcoming contemporary urban issues.

As future work, it is intended to improve the system to operate under adverse conditions, such as low illumination, occlusions, and weather variations. In addition, the inclusion of edge computing techniques will be investigated to enable real-time data processing at the point of collection, as well as to expand traffic analysis capabilities, including vehicle speed estimation and the detection of traffic irregularities.

Acknowledgment

The authors would like to thank the Secretaria da Educação do Estado do Tocantins (SEDUC-TO) for supporting the publication of this work. The authors also acknowledge the Universidade Federal do Tocantins (UFT) and the Fundação de Apoio Científico e Tecnológico do Tocantins (FAPTO) for the institutional support necessary for the development of this project.

References

[1] V. Albino, U. Berardi, and R. M. Dangelico, “Smart cities: Definitions, dimensions, performance, and initiatives,” *Journal of Urban Technology*, vol. 22, no. 1, pp. 3–21, 2015, doi: 10.1080/10630732.2014.942092.

[2] G. Lyons, “Getting smart about urban mobility – Aligning the paradigms of smart and sustainable,” *Transp. Res. Part A Policy Pract.*, vol. 115, pp. 4–14, Sep. 2018, doi: 10.1016/j.tra.2016.12.001.

[3] C. He, D. Wang, Y. Yu, and Z. Cai, “A Hybrid Deep Learning Model for Link Dynamic Vehicle Count Forecasting with Bayesian Optimization,” *J. Adv. Transp.*, vol. 2023, 2023, doi: 10.1155/2023/5070504.

[4] UN-Habitat, “Planning and Design for Sustainable Urban Mobility: Global Report on Human Settlements 2013,” 2013. [Online]. Available: https://unhabitat.org/planning-and-design-for-sustainable-urban-mobility-global-report-on-human-settlements-2013. [retrieved: Apr. 11, 2026].

[5] National Department of Transport Infrastructure (Departamento Nacional de Infraestrutura de Transportes – DNIT). Traffic Studies Manual. 2006. [Online]. Available: https://www.gov.br/dnit/pt-br/assuntos/planejamento-e-pesquisa/ipr/coletanea-de-manuais/vigentes/723_manual_estudos_trafego.pdf [retrieved: Apr. 11, 2026].

[6] N. Sebe, I. Cohen, T. S. Huang, and A. Garg, Machine Learning in Computer Vision. New York, NY, USA: Springer, 2005.

[7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” arXiv:1506.02640, 2015. [Online]. Available: http://arxiv.org/abs/1506.02640 [retrieved: Mar. 2026].

[8] M. Durve et al., “DropTrack - Automatic droplet tracking with YOLOv5 and DeepSORT for microfluidic applications,” *Physics of Fluids*, vol. 34, no. 8, Aug. 2022, doi: 10.1063/5.0097597.

[9] T.-Y. Lin et al., “Microsoft COCO: Common Objects in Context,” arXiv:1405.0312, 2014. [Online]. Available: http://arxiv.org/abs/1405.0312 [retrieved: Mar. 2026].

[10] D. S. Bolme, J. R. Beveridge, B. A. Draper, and Y. M. Lui, “Visual Object Tracking using Adaptive Correlation Filters,” San Francisco, CA, USA: IEEE, Aug. 2010.

[11] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, “Simple Online and Realtime Tracking,” Feb. 2016, doi: 10.1109/ICIP.2016.7533003.

[12] T. Bai, “Multiple Object Tracking Based on YOLOv5 and Optimized DeepSORT Algorithm,” in *Journal of Physics: Conference Series*, Institute of Physics, 2023. doi: 10.1088/1742-6596/2547/1/012022.

[13] Y. Chen, B. Wu, G. Luo, X. Chen, and J. Liu, “Multi-target tracking algorithm based on YOLO+DeepSORT,” in *Journal of Physics: Conference Series*, Institute of Physics, 2022. doi: 10.1088/1742-6596/2414/1/012018.N

[14] Hikvision, Pro & Value Series Quick Guide – 2020 H1. Hangzhou, China: Hikvision Digital Technology Co., Ltd., 2020. [Online]. Available: hikvision.com/content/dam/hikvision/eu/support/brochures/product-catalog/PRO%20&%20VALUE%20SERIES%20QUICK%20GUIDE%20-%20 2020%20H1%20-%20EU%20.pdf. [retrieved: Apr. 11, 2026].

[15] Associação Brasileira de Normas Técnicas, “NBR IEC 60529: Degrees of protection provided by enclosures (IP Code),” Brazilian Standard, Apr. 2017.

[16] Ultralytics, “YOLOv5.” [Online]. Available: https://github.com/ultralytics/yolov5 [retrieved: Mar. 2026].