DOI: 10.3217/978-3-99161-057-1-090

AI-Powered vehicle classification for scalable infrastructure monitoring

Leonardo Iacussi[1, 0009-0007-5321-3826], Nicola Giulietti[2, 0000-0001-9922-3201], Alessandro Lucci[1]
Giuseppe Lucenti[1], Emanuele Zappa[1, 0000-0003-3320-9030], Paolo Chiariotti[1, 0000-0001-6287-5521], Alfredo Cigada[1, 0000-0001-6861-8374]

[1]Politecnico di Milano, Department of Mechanical Engineering, Via Privata Giuseppe la Masa, 20156 Milano, Italy
[2]University of Pavia, Dipartimento di Ingegneria Industriale e dell'Informazione, 27100 Pavia, Italy

email: leonardo.iacussi@polimi.it, nicola.giulietti@unipv.it, alessandro.lucci@mail.polimi.it, giuseppe.lucenti@mail.polimi.it, emanuele.zappa@polimi.it, paolo.chiariotti@polimi.it, alfredo.cigada@polimi.it

ABSTRACT: Real-time monitoring of road infrastructure is crucial in addressing the challenges posed by the increasing volume of vehicles and the need for timely maintenance to manage structural aging. Traditional Weight-in-Motion (WIM) systems provide accurate measurements of vehicle load, axle configurations, and speed but are costly to install and require road closures, hindering widespread deployment. This study introduces an innovative method for estimating traffic load by repurposing acceleration-based Structural Health Monitoring (SHM) systems integrated with an AI powered vision system which enables to classify vehicles, estimate their weight, speed and finally assess traffic load over time with a scalable and cheaper solution. Vehicles have been classified into three macro classes: cars, lightweight trucks and heavy trucks. A comparative analysis has been performed between load estimation using only the AI-powered vision system, based on YOLO object detection, and an enhanced approach that integrates acceleration data. The combined method demonstrated significantly improved accuracy in weight estimation. The methodology was tested on an highway viaduct and the results validated by using a reference WIM system. The findings underscore the potential of this integrated approach to provide cost-effective and scalable solutions for traffic load estimation and structural health assessment.

KEY WORDS: SHM; Weight in motion; Road traffic monitoring; Sensor fusion

1 INTRODUCTION

In recent decades, highway infrastructure has been subjected to escalating stress due to the combined impacts of increasing traffic volumes and the aging of structural components. Transportation networks, which serve as vital conduits for economic activity and goods distribution, are facing growing demands. A major driver of this strain is the rapid expansion of road-based freight transport, fueled by the rising demand for efficient and adaptable logistics. In Europe, for instance, road freight accounted for more than 75% of all inland freight transport in 2023, with heavy-duty vehicles (HDVs) playing a central role in long-haul logistics and supply chains [1]. Bridges and viaducts are particularly susceptible to the combined effects of increased loading and structural aging. These assets often endure long service durations, facing environmental wear, and intensified mechanical stresses. Among the various contributors to their deterioration (e.g. material fatigue, environmental factors, seismic forces and inadequate maintenance) overloading by HDVs stands out as one of the most widespread and damaging. Numerous studies have identified overloading as a critical factor in both the progressive degradation and sudden failure of bridge structures across different contexts and typologies [2, 3].

Given this context, the need for effective and continuous traffic monitoring systems has become increasingly urgent. In particular, identifying and quantifying the load contribution from heavy vehicles is critical for estimating cumulative damage, supporting load rating decisions, and optimizing maintenance schedules. Traditional static weighing stations, while accurate, are inefficient for large-scale deployment due to their reliance on vehicle stops, high operational costs. Consequently, Weigh-In-Motion (WIM) systems have emerged as a valuable alternative, capable of measuring axle loads, gross vehicle weight (GVW), and vehicle classification in real time without interrupting traffic flow [4]

WIM systems are generally classified into two main categories: Pavement-based WIM (P-WIM) and Bridge-based WIM (B-WIM). P-WIM systems involve the installation of strain or piezoelectric sensors within the roadway surface to directly record the forces exerted by passing axles. While effective in certain applications, these systems require an high installation cost and frequent maintenance and recalibration due to their direct exposure to traffic and weather. In contrast, B-WIM systems leverage the dynamic or static responses of bridge structures to estimate vehicle weights typically using strain, displacement, or acceleration sensors. This approach is advantageous as it utilizes existing infrastructure, minimizes road surface interventions, and offers spatial scalability over the entire highway network [5].

The concept of B-WIM was first introduced by Moses in 1979, where an inverse problem formulation was used to estimate axle weights based on strain measurements recorded during vehicle crossings [6]. This methodology laid the foundation for modern B-WIM systems and has since been refined through the integration of improved sensing hardware, robust signal processing algorithms, and advanced calibration procedures [7]. Strain-based B-WIM systems are currently the most widely adopted, offering high accuracy for vehicle weight estimation under controlled conditions. However, they still face limitations related to temperature sensitivity, sensor drift, and

DOI: 10.3217/978-3-99161-057-1-090

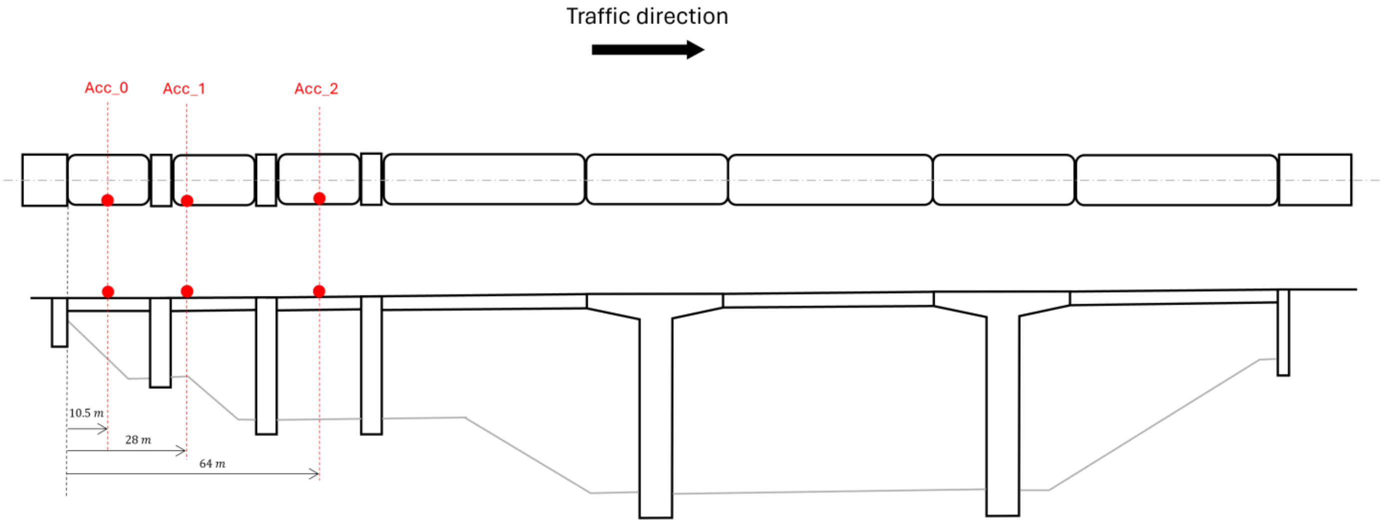

Figure 1: Viaduct schema and accelerometer position.

the need for detailed knowledge of bridge geometry and boundary conditions.

More recently, research has explored the use of acceleration-based B-WIM systems as a low-cost and scalable alternative. These systems utilize MEMS (Micro Electro-Mechanical System) accelerometers to measure bridge vibrations induced by vehicle crossings. Their non-intrusive nature and ease of deployment make them attractive for widespread monitoring applications. Nevertheless, acceleration-based systems introduce new challenges, particularly related to signal variability caused by vehicle dynamics, road roughness and environmental noise [8]. Despite these challenges, several promising studies have demonstrated the feasibility of such approaches. For instance, Sekiya et al. [9] deployed MEMS accelerometers on a steel bridge and demonstrated the potential to estimate GVW and axle positions from a single vehicle crossing. A follow-up study [10] extended this investigation to a year-long deployment, highlighting the need for temperature compensation and advanced data filtering techniques to ensure accuracy.

Further developments in the field have explored hybrid and data-driven methods. O'Brien et al. [11] used statistical analysis of acceleration signals to jointly estimate vehicle weight and assess bridge integrity. While effective for GVW estimation, the method struggled to resolve axle weights and spacings. Wang et al. [12] addressed some of these limitations by integrating vision-based systems with acceleration data to enhance vehicle detection and classification capabilities. The incorporation of transfer learning also demonstrated improved generalization across different bridge types.

These advancements underscore the growing interest in low-cost, scalable, and intelligent structural health monitoring systems. The evolution of WIM technologies, particularly B-WIM systems enhanced by MEMS sensors and machine learning algorithms, opens new possibilities for real-time infrastructure assessment. However, significant challenges remain in terms of robustness, environmental adaptability, and the reliable estimation of axle-level loads. As urban infrastructure continues to age and traffic volumes increase, the development of resilient and accurate WIM systems becomes not only desirable but essential for the future of bridge maintenance and safety.

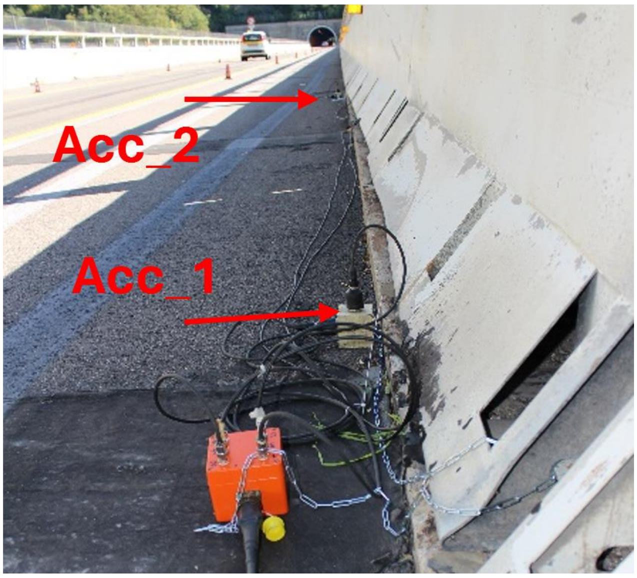

Figure 2: Piezoelectric accelerometers (Acc_1 and Acc_2) positioned on the viaduct spans.

This study aims to demonstrate the feasibility of estimating highway traffic loads by utilizing existing structural monitoring systems installed on viaducts, offering a cost-effective alternative to conventional, high-cost systems such as P-WIM technologies. The proposed methodology integrates vision-based systems for vehicle detection and classification with acceleration data acquired using accelerometers positioned on the bridge spans of the viaduct. The effectiveness of this approach is illustrated through a case study involving an operational highway viaduct located in Italy (schematized in Figure 1) where a reference P-WIM system is present as a reference for the vehicle weight.

In this work, sensitive data regarding highway traffic and viaduct accelerations were used; therefore, the actual acceleration values will be masked in the figures, and no

DOI: 10.3217/978-3-99161-057-1-090

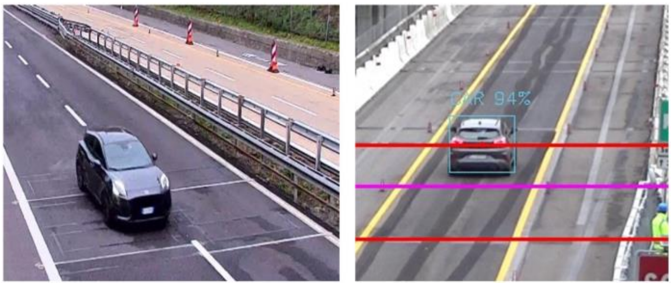

Figure 3: Vehicle detection when it crosses the P-WIM system (**left**) and while it crosses the first viaduct span (**right**).

absolute values will appear in the text. This does not affect the validity of the results.

2 MATERIALS AND METHODS

The viaduct features a simply supported beam configuration and consists of six spans. This study focuses on the first three spans, each measuring 21 meters, as they correspond to the direction of traffic flow. The primary structural components are made of conventional reinforced concrete, while the decks and pier caps are constructed using pre-stressed reinforced concrete to enhance load-bearing capacity and durability. Overhangs are located at piers 5 and 6, supporting the adjacent suspended spans through Gerber saddles, which facilitate simple support articulation and allow for structural continuity. The entire structure is founded on direct foundation systems.

The monitoring setup implemented on the viaduct includes three PCB 393A03 piezoelectric IEPE accelerometers positioned on the first three spans of the viaduct (see Figure 1 and Figure 2). The use of multiple accelerometers enables vehicle speed estimation by correlating the time delays between the signals captured by the sensors. Furthermore, it allows for a comparative analysis of acceleration responses at mid-span locations (Acc_0 and Acc_2), where the dynamic amplification of structural response is more significant and near the joint (Acc_1), where vehicle impacts contribution on the acceleration data. Data acquisition was performed synchronously using a National Instruments CompactDAQ system (model cDAQ-9172 equipped with NI 9230 boards) with a sampling frequency of 1000 Samples/s.

For vehicle detection and classification, a vision system comprising a Sony Handycam HDR-CX405 camera was installed above the tunnel exit, providing a clear field of view of the traffic flow immediately following the tunnel (as shown in Figure 3). To establish a ground-truth reference, a P-WIM system located upstream of the viaduct was used. Given the absence of highway exits along between the P-WIM system and the viaduct, it was ensured that all vehicles recorded by the P-WIM system subsequently traversed the viaduct under investigation. To account for potential overtaking between the two locations, an auxiliary camera installed at the P-WIM site was used to reorder vehicles and ensure accurate matching with the viaduct observations.

The reference data provided by the P-WIM system included detailed vehicle information, such as speed, length, lane position, axle count and spacing, individual axle loads, gross vehicle weight (GVW), and vehicle classification based on the ASTM E1318-09 standard [13]. The vehicle flow was sparse enough to ensure that only one vehicle crossed at a time, simplifying the analysis and reducing the complexity associated with multi-vehicle events. The dataset analyzed comprises a total of 96 vehicles that traversed the viaduct during the time period in which all monitoring systems were simultaneously acquiring data. Approximately 90% of the recorded vehicles were two-axle vehicles, with an average GVW of around 2 tonnes. This category includes not only passenger cars, but also vans, light trucks, and motorcycles. The remaining 10% consisted of heavier vehicles with three, four, or five axles, corresponding primarily to trucks, with an average GVW of approximately 20 tonnes. Regarding vehicle speed estimation, it was derived by correlating the data from the accelerometers with the output from the vision system. The velocity measurements provided by the P-WIM system could not be considered a reliable reference, as the system is located upstream of the viaduct and vehicle speeds may vary along the intervening highway segments.

2.1 Vision-based vehicle detection and speed estimation

Vehicle detection and classification within the region of interest were performed using a Python-based application that integrates OpenCV with a pre-trained object detection algorithm. Each video frame was processed to identify and localize vehicles using the YOLOv3 (You Only Look Once [14]) model, trained on the COCO dataset [15], which enables vehicle classification in only four classes: car, motorbike, bus, and truck. To achieve more stable classification and refine detections, post-processing was performed by applying a confidence threshold and non-maximum suppression. A tracking algorithm was then employed to assign persistent IDs to detected vehicles across consecutive frames by comparing the centroids of bounding boxes. This enabled consistent object identification throughout the video. Vehicle counting was carried out by monitoring object trajectories across a user-defined detection zone delimited by three virtual lines. When an object centroid crossed the designated thresholds, it was counted and classified accordingly (see Figure 3 (**right**)).

Given the relevance of vehicle speed as a parameter, a method for its estimation was also implemented. During video processing, the frame numbers at which each vehicle entered and exited the detection area is recorded. By combining this information with the known frame rate and the estimated physical distance (Δs) between the entry and exit lines, the vehicle speed (v_ID) was approximated using the following relation:

$$v\_ID = \frac{\Delta s \times fps}{frame_1 - frame_0} \tag{1}$$

where $frame_0$ and $frame_1$ denote the frame indices corresponding to the vehicle entry and exit points, respectively, and fps (frame per second) corresponds to 50 for the camera used. The accuracy of this estimate depends on factors such as camera placement, resolution, and perspective distortion. Several limitations in the detection and classification process were addressed. First, the object class assigned by YOLOv3 may vary across frames, leading to misclassification. To

DOI: 10.3217/978-3-99161-057-1-090

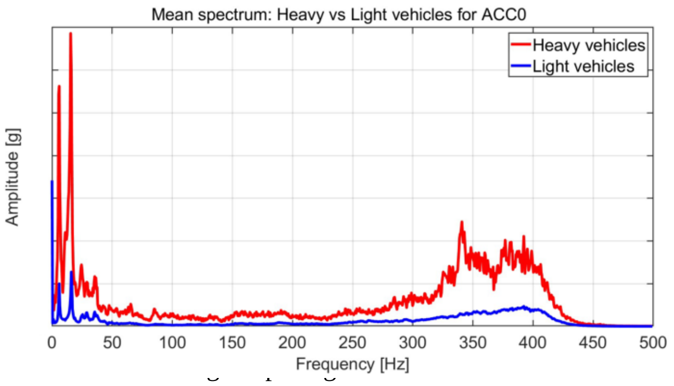

Figure 5: Average spectrum for heavy vehicles and light vehicles during their passage over the accelerometer Acc_0.

mitigate this, all class labels assigned to an object during its passage were stored, and the most frequently assigned class (mode) was used as the final classification. Second, incomplete or intermittent detections, often due to low confidence scores, could cause the same vehicle to be assigned to different IDs, affecting both vehicle counting and speed estimation. To address this, the ID matching process was extended to compare across both temporary detection lists, improving continuity in object tracking. However, if a vehicle was detected only in the latter part of the detection zone, the estimated speed would be significantly overestimated. To prevent this, the system recorded the initial detection point for each object and suppressed speed calculations for vehicles detected too late in the zone, where time-distance correlation becomes unreliable.

2.2 Acceleration-based vehicle detection and sped estimation

This section focuses on vibration analysis using accelerometer data, aiming to correlate sensor-derived information with data obtained from video recordings. Specifically, the analysis addresses two objectives: (i) the identification of light and heavy vehicles using time-domain features extracted from acceleration signals, and (ii) the estimation of vehicle speed through the synchronization of signals from three accelerometers, followed by comparison with speed estimates derived from the vision-based system.

By analyzing the time histories from raw accelerometer data (see Figure 4), vehicle pass-by events were initially identified by observing distinct signal peaks. A representative segment of the vibration data, collected during one of the experimental campaigns, illustrates this concept. To improve detection clarity, a moving Root Mean Square (RMS) function was applied to the signal. The RMS provides a measure of signal energy over a defined time window and is particularly useful for distinguishing between light and heavy vehicles, as heavier vehicles are expected to induce greater energy in the structure. However, fixed thresholds for vehicle classification cannot be determined a priori, as they depend on the specific structure and sensor placement.

Frequency domain analysis (see Figure 5) reveals that the primary distinction between light and heavy vehicles lies in the low-frequency range, specifically below 50 Hz. Based on this observation, the signal energy within this band, quantified using the RMS of the band-pass filtered signal (5–50 Hz), was selected as the key feature for vehicle classification.

Vehicle speed can be estimated also correlating signals from different accelerometers, based on their spatial position and the time taken by a vehicle to traverse the corresponding distance. Several methods can be used to estimate the time lag between two signals; in this case, the cross-correlation method was used, which computes the time lag between two signals considering the well-known correlation function:

$$Corr(\tau) = \int_0^T a_{en}(t)\, a_{ex}(t+\tau)dt \qquad (2)$$

where $a_{en}(t)$ and $a_{ex}(t)$ represent the signals from entry and exit sensors, respectively, and τ denotes the time shift. The correlation was applied either directly to the moving RMS of the signal or to the envelope of the signal. These two acceleration-based methods will be compared to the vision-based estimates in the results section to evaluate their performance.

3 RESULTS

This section presents the results of the study, beginning with the comparison of speed estimations comparison between the two systems and subsequently discussing the outcomes of vehicle classification between light and heavy vehicles.

Although vehicle speed estimation is not the primary focus of this study, it remains a valuable parameter, as the velocity at which a load travels over a viaduct can significantly influence the dynamic response of the structure. Moreover, since speed is also estimated by P-WIM systems, utilizing existing infrastructure such as cameras or accelerometers for speed estimation can enhance the spatial coverage of traffic monitoring. This approach has the potential to provide a more comprehensive mapping of vehicle speeds across the highway network, contributing to a more accurate assessment of moving loads on the structure.

The comparison between the acceleration-based and vision-based vehicle speed estimation methods is summarized in Table 1.

Table 1. Vehicle speed estimation comparison between acceleration-based and vision-based methods.

Method	$\Delta < 20\,\frac{\text{km}}{\text{h}}$	$\Delta \geq 20\,\frac{\text{km}}{\text{h}}$
Moving RMS	91.7%	8.3%
Envelope	82.5%	17.5%

Both the Moving RMS and the Envelope methods have been used for speed estimation from the acceleration based system. The parameter Δ represents the absolute difference in speed estimates between the two systems. Results are categorized based on whether this difference is less than or greater than 20 km/h. The Moving RMS method exhibited a high level of agreement, with 91.7 % of the estimates falling within a $\Delta < 20$

DOI: 10.3217/978-3-99161-057-1-090

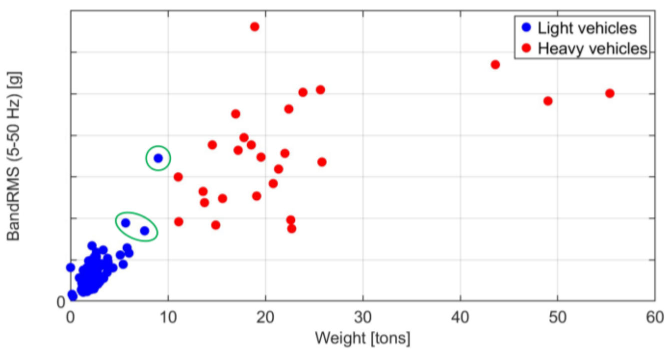

Figure 6: RMS value of the acceleration in the time window when the vehicle passes by the accelerometer Acc_0.

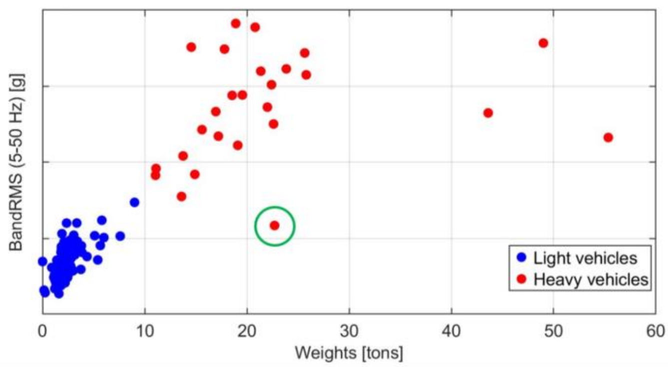

Figure 7: RMS value of the acceleration in the time window when the vehicle passes by the accelerometer Acc_1.

km/h and only 8.3 % exceeding this threshold. Similarly, the Envelope method showed that 82.5 % of estimates were within the acceptable margin, while 17.5 % exceeded it. These findings suggest that both acceleration-based techniques, particularly the Moving RMS approach, demonstrate promising accuracy and consistency when compared with the vision-based system.

4 CONCLUSION AND FURTHER DEVELOPMENT

This study presents a preliminary investigation into the feasibility of leveraging existing monitoring systems - , specifically, accelerometers installed on viaducts and highway surveillance cameras - to estimate traffic loads acting on highway infrastructure.

Although not the primary objective, traffic load estimation is highly relevant for informing maintenance strategies, forecasting potential structural damage and design new infrastructure.

Currently, P-WIM systems serve this purpose but are often associated with high installation and maintenance costs. As an alternative, the proposed methodology offers a low-cost, albeit less precise, approach by estimating vehicle loading by identifying the number of heavy vehicles (>10 tons) traversing the viaduct. A reference P-WIM system was used as ground truth for weight calibration, enabling a daily estimate of total traffic-induced load, potentially useful in design or assessment contexts where knowledge of acting loads is required.

Results demonstrate that employing an intelligent vision-based system (e.g. using pre-trained deep learning object detection models), can distinguish between light and heavy vehicles. However, intermediate vehicle categories (e.g., vans, minibuses, RVs) are occasionally misclassified. One inherent limitation is the high variability in heavy vehicle weights depending on loading conditions, making it challenging to detect overloading or estimate precise weight using classification alone. In this regard, the inclusion of structural acceleration data improves classification accuracy, as the energy transferred by the vehicle, expressed through the root mean square (RMS) of the acceleration signal, provides an additional informative metric for estimating vehicle mass.

Furthermore, the study compares vehicle speed estimates derived from both systems. By correlating signals from three accelerometers and cross-referencing them with visual detections, the comparison shows that in 90% of cases, the speed difference between the two methods is below 20 km/h, suggesting consistency and potential for dual-system validation.

Despite its promising results, the study faces several limitations that define avenues for future research. Most notably, the dataset includes only 239 vehicles, limiting the use of more sophisticated data-driven algorithms that could enhance load estimation using richer input features beyond RMS. Expanding the dataset would enable the exploration of machine learning models trained directly on raw acceleration signals.
Concerning vision-based algorithms, having a bigger dataset would allow model retraining with more vehicle classes specifically tailored for this application, enabling a more precise weight estimation. Additionally, while the present study analyzes the two systems separately, future work should consider integrating their outputs into a unified model, enabling sensor fusion to enhance overall system robustness and accuracy.

ACKNOWLEDGMENTS

The data in this article refer to the project "Argo Innovation Lab – Comparison of the metrological performances of measurement systems to get the cumulative damage by heavy traffic", funded by the private companies Movyon and Elis Innovation Hub. Both companies are respectfully acknowledged.

REFERENCES

[1] Who is driving what, and where? EU road freight trends iru.org. https://www.iru.org/news- resources/newsroom/who-driving-what-and-where-eu-road-freight-trends, 2024. [Accessed 07-04-2025].

[2] Lu Deng, Wei Wang, and Yang Yu. State-of-the- art review on the causes and mechanisms of bridge collapse. Journal of Performance of Constructed Facilities, 30(2):04015005, 2016.

[3] Wanshui Han, Jun Wu, C. S. Cai, and Suren Chen. Characteristics and dynamic impact of overloaded extra heavy trucks on typical highway bridges. Journal of Bridge Engineering, 20(2), February 2015.

[4] Rahim F. Benekohal, Yoassry M. El-Zohairy, and Stanley Wang. Truck travel time around weigh stations: Effects of weigh in motion and automatic vehicle identification systems. Transporta- tion Research Record: Journal of the Transportation Research Board, 1716(1):135–143, January 2000.

[5] Lu Cheng, Hongjian Zhang, and Qing Li. Design of a capacitive flexible weighing sensor for vehicle wim system. Sensors, 7(8):1530–1544, August 2007.

 DOI: 10.3217/978-3-99161-057-1-090

[6] Fred Moses. Weigh-in-motion system using in- strumented bridges. Transportation Engineering Journal of ASCE, 105(3):233–249, May 1979.

[7] Myra Lydon, S. E. Taylor, D. Robinson, A. Mufti, and E. J. O. Brien. Recent developments in bridge weigh in motion (b-wim). Journal of Civil Struc- tural Health Monitoring, 6(1):69–81, May 2015.

[8] Enrico Oliveira Rocheti and Rodrigo Mor- eira Bacurau. Weigh-in-motion systems review: Methods for axle and gross vehicle weight estima- tion. IEEE Access, 12:134822–134836, 2024

[9] Hidehiko Sekiya, Kosaku Kubota, and Chitoshi Miki. Simplified portable bridge weigh-in-motion system using accelerometers. Journal of Bridge Engineering, 23(1), January 2018.

[10] Hidehiko Sekiya. Field verification over one year of a portable bridge weigh-in-motion system for steel bridges. Journal of Bridge Engineering, 24(7), July 2019.

[11] Eugene OBrien, Muhammad Arslan Khan, Daniel Patrick McCrum, and Aleˇs ˇZnidariˇc. Using statistical analysis of an acceleration-based bridge weigh-in-motion system for damage detection. Applied Sciences, 10(2):663, January 2020.

[12] Haoqi Wang, Tomonori Nagayama, Takaya Kawakatsu, and Atsuhiro Takasu. A data-driven approach for bridge weigh-in-motion from impact acceleration responses at bridge joints. Structural Control and Health Monitoring, 2023:1–14, May 2023.

[13] Standard specification for highway weigh-in- motion (wim) systems with user requirements and test methods, 2009.

[14] Ultralytics. YOLOv3 — docs.ultralytics.com. https://docs.ultralytics.com/it/models/yolov3/. [Accessed 07-04-2025].

[15] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zit- nick, and Piotr Doll´ar. Microsoft coco: Common objects in context, 2014.

