



HAL
open science

Monitoring of moving rail-road cars through infrared thermal vision: Study of hot box detection on reconstructed scenes with and without Deep Learning approaches

Boualem Merainani, Thibaud Toullier, Sarouthan Sriranjana, Jean Dumoulin

► To cite this version:

Boualem Merainani, Thibaud Toullier, Sarouthan Sriranjana, Jean Dumoulin. Monitoring of moving rail-road cars through infrared thermal vision: Study of hot box detection on reconstructed scenes with and without Deep Learning approaches. 17th Quantitative Infrared Thermography Conference - QIRT 2024, Jul 2024, Zagreb, Croatia. 10.2832/319371 . hal-04798560

HAL Id: hal-04798560

<https://inria.hal.science/hal-04798560v1>

Submitted on 22 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Monitoring of moving rail-road cars through infrared thermal vision: Study of hot box detection on reconstructed scenes with and without Deep Learning approaches

by B. Merainani*, T. Toullier*, S. Sriranjani**, J. Dumoulin*

* Univ. Gustave Eiffel, Inria, COSYS-SII, I4S Team, F-44344, Bouguenais, France, boualem.merainani@univ-eiffel.fr

** SNCF RESEAU, DGII, Département de la Signalisation Ferroviaire, 6 avenue François Mitterrand, F-93574 La Plaine Saint Denis, France

Abstract

Railway safety faces a critical challenge from overheated axle bearings, commonly known as hot boxes. Early detection is, therefore, of utmost importance. Departing from the complexity and resource-heavy nature of current hot box detectors, a new and affordable thermal vision-based wayside monitoring method is proposed. It combines panoramic thermal images with deep learning (DL)-based object detection algorithm. Panoramas were created, from thermal image sequences collected, by cooled and uncooled cameras, during field experiments on both freight and passenger trains. The results yielded promising perspectives, notably in maintaining high safety standards, at a more budget-friendly cost.

1. Introduction

Rail transport represents a crucial sector of the economy. In 2021, the freight transport activities in the European Union-27 are estimated to amount of 409.6 billion tonne-kilometre and the passenger transports to 265.2 billion passenger-kilometre [1]. As the demand for rail services continues to grow, safety remains a primary concern. As a critical vehicle component, axle bearing is subjected to intense mechanical and thermal stresses during operation. A hot box can arise, therefore, if the bearing experiences excessive heat due to friction caused by inadequate or contaminated lubricant. Material fatigue becomes a significant concern in this scenario, progressively compromising the structural integrity of the bearing components. The consequences of undetected hot box can be dramatic, posing imminent threats to the safety of the entire rail system, including the potential for derailment and fires.

Generally, hot box detectors incorporate two categories of sensing systems: inboard sensors and those integrated into the track also called wayside detectors. Inboard sensor systems, being directly mounted on individual vehicles on each axle box, provide distinct advantages such as real-time feedback on the condition of bearings, enabling prompt reactions to possible concerns. However, challenges include high installation costs, increased maintenance needs, and a significant coverage gap, as many existing freight trains lack these systems. On the other hand, wayside detectors are intended to measure the heat radiating remotely from passing trains axles, providing a broader perspective on the overall health of the rail network. Integrated into the track infrastructure, these detectors generate alerts when temperatures exceed the predefined thresholds set by infrastructure managers. While these systems are indeed effective, the high installation and maintenance costs can be a significant obstacle to their widespread deployment across the rail network. This problem cannot be understated given the fact that a bearing failure can occur within few minutes of passing a hot-box detector without triggering an alarm [2,3]. Additionally, the reliance on external facilities, like triggers, introduces a dependency that, if compromised, could impact the overall efficacy of the temperature monitoring system. To address these drawbacks, we propose an affordable thermal vision-based wayside monitoring system. It incorporates thermal camera and a cutting-edge DL-based object detection algorithm, which is YOLOV8, to detect the group axle/wheel components. Despite the very limited prior research in this specific application, this study stands out by processing panoramic thermal images of the entire train rather than individual image sequences, an area that has not been explored previously.

2. Procedures and thermal image database:

2.1. Data acquisition and creation of thermal panoramic images

In-situ experiments were conducted at various times of the day on an operational rail line. The image sequences were collected using two camera models, a cooled camera, operating at 250Hz (SWIR - MWIR), and an uncooled camera, operating at 100 Hz (LWIR). This choice was made to conduct a comparison.

Panoramic images are created through a multi-step process. Initially, the trains arrival and departure in the videos needs to be detected as the cameras were triggered manually. In the present study, it has been achieved by calculating the structural similarity index measure (SSIM) between two normalized frames, where the first frame corresponds to the initial frame of the video. A result approaching 1, means both frames display the background; otherwise, it indicates that the second frame captures the train body. Second, the train's displacement, in pixels, is estimated. Subsequently, a center strip of the size corresponding to the displacement is extracted from each frame. Finally, these center strips are stitched together to create panoramic images. It is noteworthy that precise pixel-level displacement is important for generating panorama devoid of visible stitching artifacts. To this end, the optical flow method, a technique that tracks pixel motion across consecutive frames, was employed [4].

2.2. Deep learning-based hot box detection

The hot box detection process relies on single-stage object detector algorithm. Different from two-stage detectors that typically involve a region proposal step followed by a refined object classification, like region-based convolutional neural networks (R-CNN) [5], fast and faster R-CNN [6,7], single-stage detectors streamline the process by directly



predicting bounding boxes and class probabilities in a single pass. This results in a more computationally efficient approach, making it well-suited for scenarios where a prompt decision is imperative.

Representative algorithms in this second category include Single Shot Detector [8] and the “You Only Look Once” (or YOLO) series [9], with YOLOV8 standing as the latest version. YOLOV8 improves the YOLOV5 framework, bringing advancements in both architecture and developer experience areas. It features an anchor-free detection head, an original backbone network, a loss function and other enhancements. In line with certain predecessors, YOLOV8 is available in five different variants ranging from nano (n) to extra-large (x), each determined by the number of parameters. In fact, the model size has a big influence on results. The bigger the size, the more coefficients the model will have to work with, which will allow it to solve more complex problems. Nevertheless, bigger the model, the longer it will take to process an image. Besides, the input image size affects latency and accuracy: if it's too large, latency increases significantly. But, on the same time, it generally boosts accuracy [10]. So, finding the right balance is key, in most application.

To tackle practical challenges, resource limitations and high-resolution nature of panoramas, YOLOV8n emerged as the optimal choice. Hence, it was trained on selected image sequences, ensuring the inclusion of diverse and representative samples, on a workstation powered by NVIDIA Quadro RTX 5000 (with 3072 CUDA cores, 384 Tensor Cores and 16 GB RAM) GPU. But the testing was performed on panoramic images with NVIDIA Jetson Orin AGX (2048-core NVIDIA Ampere with 64 Tensor Cores and 64 GB RAM). Figure 1 exhibits panoramas comprising three locomotives, reconstructed from image sequences taken by the two camera models. Within the depicted panoramas, the results of hot box detection, by the YOLOV8n algorithm, are showcased. Furthermore, we explore the feasibility of detecting the group axle/wheel components without relying on artificial intelligence-assisted tools. We also expose the limitations that impact the accuracy, particularly due to variations in experimental conditions. This investigation affirms the role of YOLOV8n, and highlights its adaptability in addressing these challenges. In the extended version of the paper, additional results are presented, and a more comprehensive discussion is undertaken.

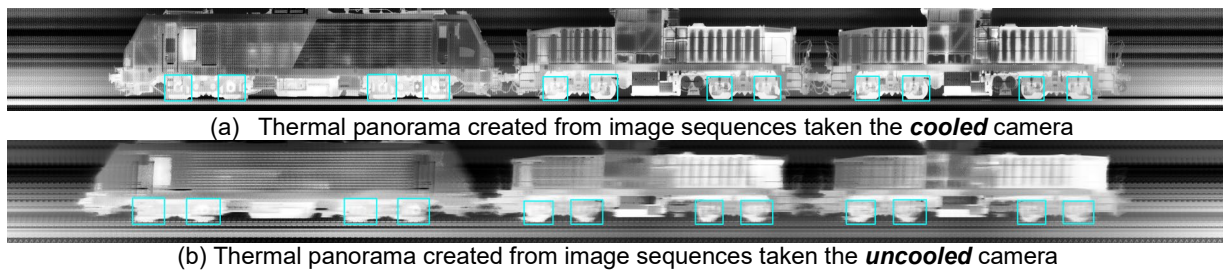


Figure 1: Hot box detection results using YOLOV8n.

3. Conclusion and perspectives

This research highlights the potential and prerequisites associated with employing panoramic thermal images for detecting and counting hot boxes. Seeking to maintain high standards of safety with cost-effective new hot box detector design, DL can effectively process thermal images from cooled but also uncooled camera, even under various sun influences in outdoor conditions. The natural next step involves dual objectives: deblurring images and evaluate performances in all weather conditions, validated through comparisons with existing axle counters.

Acknowledgments

SNCF Reseau and BRIGHTER project. BRIGHTER has received funding from the Chips Joint Undertaking (Chips JU) under grant agreement No 101096985. The JU receives support from the European Union's Horizon Europe research and innovation program and France, Belgium, Portugal, Spain, Turkey

REFERENCES

- [1] European Commission, Directorate-General for Mobility and Transport, EU transport in figures – Statistical pocketbook 2023 (pages 44, 51), Publications Office of the European Union, 2023, <https://data.europa.eu/doi/10.2832/319371>
- [2] Barke, D., & Chiu, W. K. (2005). Structural health monitoring in the railway industry: a review. *Structural Health Monitoring*, 4(1), 81-93.
- [3] Stewart, M., Flynn, E., & Marquis, B. P. (2019). An Implementation Guide for Wayside Detector Systems.
- [4] Horn, B. K., & Schunck, B. G. (1981). Determining optical flow. *Artificial intelligence*, 17(1-3), 185-203.
- [5] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2015). Region-based convolutional networks for accurate object detection and segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 38(1), 142-158.
- [6] Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440-1448).
- [7] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28.
- [8] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14* (pp. 21-37). Springer International Publishing.
- [9] Terven, J., Córdova-Esparza, D. M., & Romero-González, J. A. (2023). A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Machine Learning and Knowledge Extraction*, 5(4), 1680-1716.
- [10] Lema, D. G., Usamentiaga, R., & García, D. F. (2023). Quantitative comparison and performance evaluation of deep learning-based object detection models on edge computing devices. *Integration*, 102127.