QIRT'2022

10.21611/qirt.2022.3018
16th Quantitative InfraRed Thermography Conference

Toward the development of intelligent wayside hot bearings detector system : combining the thermal vision with the strength of YOLO-v4

by B. Merainani*, T. Toullier*, B. Zongo**, S. Sriranjan**, S. Zanaroli**, M. Guiraud**, J. Dumoulin*

* Univ. Gustave Eiffel, Inria, COSYS-SII, I4S Team, F-44344, Bouguenais, France
** SNCF RESEAU, DGII SF PRODUITS STEM, 6 avenue François Mitterrand, F-93574 La Plaine Saint Denis, France

Abstract

Overheated rail-road car wheels and bearings known as hot boxes, are a major threat for any railway operation. A hot box can lead to fractures of axle journals and might cause grave accidents if not detected in time. This paper proposes to automatically detect, track and count the hot boxes in infrared thermal image sequences (IRTIs), acquired with two infrared camera types, cooled camera (CC) and uncooled camera (UC), through a modified YOLO-v4 framework. In-service experiments have been carried out on freight and passenger trains. The training and testing sets are generated, including original and augmented IRTIs with a resolution of 448 $\times$ 239 pixels for CC and 640 $\times$ 240 pixels for the UC. The trained models, titled Y-CC and Y-UC, perform in a promising way in regards to detection precision, recall and F1-score.

1. Introduction

The volume of European rail traffic is growing rapidly. Estimates suggest increases in passenger and freight transport by 42% and 60% by 2050 respectively. This pushes, more than ever, infrastructure managers to develop cost effective and reliable solutions to improve safety and operational performance. Among monitored components, bogie is the most complex and important element in rail-road cars. It is responsible for the correct movement along the track. Wheels and bearings, are two critical components in bogie. However, no matter how excellent they are designed, they are prone to damages due to thermal or excessive loading ..etc. Therefore, to reduce maintenance cost and ensure the rail car running safety, monitoring technologies such as railway bearing acoustic monitoring (RailBAM), hot box detector (HBD) & hot wheel detector (HWD) and thermal imagery based infrared camera have been developed. However, previous thermal infrared imaging studies, [1–4], suffer from low detection accuracy which limit considerably their application. Recently, many researchers have investigated the possibility of detecting objects in infrared images based on machine learning. Some of them rely on region based convolutional neural networks (R-CNN), Faster R-CNN, single shot multi-box detector (SSD) and EfficientDet. To our knowledge, none of the methods mentioned above, in particular YOLO, have been investigated for train car overheated wheel and bearing, using visible or thermal cameras. In such context, this paper presents an approach to automatically detect, track and count the hot axle boxes in IRTIs using YOLO-v4 core fused with convolutional attention module. Results obtained with cooled infrared camera (SWIR - MWIR) and an uncooled one (LWIR) are presented and discussed.

2. Methodology and infrared image database

As a single-stage object detector, YOLO-v4 tackles object detection as a regression problem to spatially separated bounding boxes and associated class probabilities. The loss function used in the YOLO-v4 model consists of three main parts: bounding box location loss $\mathcal{L}_{CIoU}$, confidence loss $\mathcal{L}_{conf}$, and classification loss $\mathcal{L}_{cla}$. The formula of the loss function $\mathcal{L}$ is given as follows:

$$\mathcal{L} = \boldsymbol{\lambda_1}\mathcal{L}_{conf} + \boldsymbol{\lambda_2}\mathcal{L}_{cla} + \boldsymbol{\lambda_3}\mathcal{L}_{CIoU} \tag{1}$$

where $\boldsymbol{\lambda}$ is the balance coefficient.

For a faster convergence and more accurate regression during the training process, Complete-Intersection over Union (CIoU) loss was adopted. Different from IoU loss, the CIoU loss ($\mathcal{L}_{CIoU}$) considers geometrical factors such as aspect ratio and the normalized distance between the centers of the ground truth and the predicted bounding boxes. Furthermore, to evaluate the accuracy of object detection algorithm, performance evaluation indexes are needed such as precision (P), recall (R), F1-score. Precision gives the probability of network to identify only ground truth objects. Recall evaluates the model ability to correctly detect all ground truth objects while F1-score gives the balance between the first two metrics.

Image acquisition is the key step in training YOLO network. In this paper, an original dataset, partially used in [4], is exploited. The dataset was collected by two infrared cameras. One is cooled (operates in SWIR and MWIR regions) and the other one is uncooled (operates in LWIR region). Both are mounted on a tripods at distance of about $5m$ from the rail way. In-service experiments have been carried out, in France, on a rail line open to traffic which is taken by many freight and passenger trains. A set of infrared thermal data was, therefore, taken on the passage of two types of trains (freight and passenger trains), at different speed, and include a challenging data about the crossings between two trains. Besides, the IRTIs were recorded at different times of the day, under clear weather condition. Hence, scenarios where the sun was reflected on the trains body or pointed to the cameras, were also considered.

The training phase, was initiated with 2128 IRTIs (original and augmented data). They were randomly chosen and manually annotated with bounding boxes with a single class name called "*Axle*". 10% of the dataset was used for validation.

3. Results and analysis

In this section, the hot boxes detection results by the modified YOLO-v4 models are presented and analyzed statistically. For convenience, two YOLO-v4 models were proposed and trained independently, one for each camera type. They are titled Y-CC (for cooled camera) and Y-UC (for uncooled camera).

The models were trained and tested using the following hardware: Intel(R) Xeon(R) Gold 5218 CPU, 2 GPUs (NVIDIA Quaro RTX 5000) and $64\,Go$ RAM. The programming language was Python 3.8, the CUDA and cuDNN versions were 11.4 and 8.1, respectively. The network was optimized with Adam; The training parameters were tuned as follow: batch size of 10, learning rate of 10^{-4} and an iteration number of 2500.

6664 IRTIs were the data used to test the Y-CC and 6365 to test the Y-UC. Note that, the predicted bounding box that has the IoU value greater than or equal to 0.45 is considered as a correct prediction. The target confidence threshold (conf-thresh) was taken 0.25. The results are shown in Tab. 1 and 2.

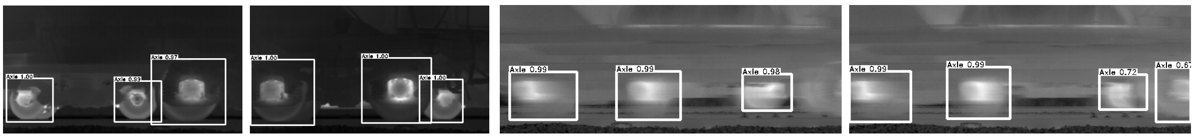
Table 1. *Performance of Y-CC: conf-thresh = 0.25 and IOU = 0.45*

Results (%) / Metrics	Dataset		
	Freight train	Crossing between passenger and freight trains	New unseen trains
Precision	99.64	99.71	99.25
Recall	99.69	99.64	99.49
F1-score	99.66	99.67	99.36

Table 2. *Performance of Y-UC: conf-thresh = 0.25 and IOU = 0.45*

Results (%) / Metrics	Dataset		
	Freight train	Crossing between passenger and freight trains	New unseen trains
Precision	98.93	99.08	99.11
Recall	98.43	100	99.9
F1-score	98.67	99.53	99.50

Visual representations of the Y-CC and Y-UC results are presented in Fig. 1. Looking separately to the performance of the two models. The Y-CC model achieves an average F1-score of 99.56% against 99.23% for the Y-UC. This difference may be explained by motion blur due to higher time integration for the UC.



(a) *IRTIs taken from CC*　　**(b)** *IRTIs taken from UC*

Fig. 1. *Example results of hot axle boxes detection using YOLO-v4: Crossing between freight and passenger trains*

Finally, discussion and perspectives on detection, tracking and counting will be proposed in the full version.

References

[1] Min-Soo Kim, Seh-Chan Oh, Geun-Yep Kim, and Seok-Jin Kwon. Underbody component monitoring system of railway vehicles using the infra-red thermal images. In *2014 International SoC Design Conference (ISOCC)*, pages 222–223. IEEE, 2014.

[2] Hanieh Deilamsalehy, Timothy C Havens, Pasi Lautala, Ezequiel Medici, and James Davis. An automatic method for detecting sliding railway wheels and hot bearings using thermal imagery. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 231(6):690–700, 2017.

[3] Hanieh Deilamsalehy, Timothy C Havens, and Pasi Lautala. Sensor fusion of wayside visible and thermal imagery for rail car wheel and bearing damage detection. In *ASME/IEEE Joint Rail Conference*, volume 50718, page V001T02A011. American Society of Mechanical Engineers, 2017.

[4] Thibaud Toullier, Jean Dumoulin, and Vincent Bourgeois. Comparative study of moving train hot boxes pre-detection and axles counting by in-situ implementation of two infrared cameras. In *QIRT Asia 2019-Quantitative InfraRed Thermography Conference*, 2019.