Rethinking transport
27–30 April 2020

TRA 2020 HELSINKI

Proceedings of 8th Transport Research Arena TRA 2020, April 27-30, 2020, Helsinki, Finland

Cargo Wagon Structural Health Estimation Using Computer Vision

Andrés Felipe Posada-Moreno [a]*, Christopher Klein[b], Marc Haßler [a], Damir Pehar [b], Alexia Fenollar Solvay [a], Christian Kohlschein [a]

[a]*Institute of Information Management in Mechanical Engineering at the RWTH Aachen University, Dennewartstr. 25-27, 52068 Aachen,Germany*
[b]*DB Cargo AG, Rheinstr. 2, 55116 Mainz, Germany*

Abstract

European Railway companies have to guarantee the safety of their operations as well as compliance with the European standards. This includes the diagnosis and maintenance of each wagon during their operations. This process is reactive and timely based, which impedes an optimal working schedule of the wagons and hinders the lifecycle management of their components. This publication describes the vision of a support system for wagon health assessment, a detailed description of the underlying methodology as well as results of the image acquisition and object detection steps. Said system aims to assist the overall maintenance process by delivering pertinent information regarding the state of the wagons as well as diagnostic results of automated components analysis. The proposed system consists of four stages, image acquisition, wagon identification, health estimation, and a decision support system.

Keywords: cargo wagons; computer vision; deep learning; safety; anomaly detection

* Corresponding author. Tel.: +49 241 80-911-73;
E-mail address: andres.posada@ima-ifu.rwth-aachen.de

1. Introduction

With the European Union growing ever closer together and the interlocking of the rail transport network as a result, it is even more essential than ever to guarantee the functioning of the rail network. Increasing the reliability of the European rail transport system is still a focus of railway companies and transport operators. The ongoing digitalization of railway assets, combined with the development of novel computer vision algorithms, opens major possibilities for supporting and automating the current railway quality and security processes.

In order to guarantee the functionality and safety of the rail network, rail freight traffic is regulated at European level. The General Contract for the Use of Freight Wagons (GCU) (*General Contract for the Use of Freight Wagons*, 2019) regulates all parameters, from wheels and undercarriages to outer ladders and tarpaulins of the cargo rail vehicles.

If a freight wagon no longer complies with the European directives, it is separated to be returned to a rail worthy condition. At this point, the wagon is inspected by expert personnel and necessary maintenance requirements are determined. Depending on the weather and lighting conditions, the process becomes more difficult and is overall very time-consuming.

To counteract the above shortcomings and working in collaboration with the German railway company Deutsche Bahn Cargo, this publication presents a structured approach to build a full pipeline tackling the automatic recognition or identification of cargo vehicles as well as the detection of anomalies (such as deformed metal parts of the wagons) through image processing. The general idea is to spot anomalies within the structure of the cargo vehicles and generate a record of said detection by taking a picture of the wagon at the shunting yard, where the vehicles are separated into different tracks.

As a basis of such a system, a four stages process is presented. As a first stage, a structure for automated image taking has to be set up, followed by a second stage in which the complete identification of a vehicle within a picture must be automated. With the first two stages of the system as a basis, the last two stages consist of the health estimation of the wagons as well as the visualization of the historical results through a decision support system.

In this publication, a novel system design, as well as an in-depth view of the automated computer vision vehicle identification is presented. In more detail, the identification stage is split into three steps. First, the position of the vehicle in the picture must be identified. Second, the International Union of Railways (UIC) wagon identification number must be recognized and read. Finally, the orientation of the wagon and therefore the picture of the left and/or right side must be established.

After a recap of the state of the art of maintenance procedures in railway operations and the development of anomaly and damage detection through computer vision (c.f. chapter 2), an overview of the different components of the decision support system is given (c.f. chapter 3) as well as an explanation of the corresponding methodology (c.f. chapter 4). In chapter 5, the results are presented and discussed. Finally, the work is concluded and an outlook into future research is given (c.f. chapter 6).

2. Related works

In recent years, multiple authors have proposed and implemented computer vision based approaches for assisting different processes in the railway transportation sector. Said research has focused on three main subjects, identification, detection of regions of interest (ROI), and health estimation.

The first subject frequently treated in the related works is the ID detection of the wagons. Authors have tackled this problem using a range of algorithms from template matching (Ningning *et al.*, 2016), region extraction and character recognition (Xiang *et al.*, 2016), to a mixture between edge detection and geometric analysis (Lisanti *et al.*, 2018). These papers have shown multiple viable approaches regarding wagon identification systems. Nevertheless, none of the papers tackles the recurrent cases of occluded characters or dirt/degradation of the characters when deploying said systems at an industrial scale.

In the Area of ROI detection for railroad vehicles, different approaches have already been considered. Multiple articles tackle the detection of specific parts, for example, a speed sensor at the axis of a locomotive (Li *et al.*, 2019) or dust collectors, cock handles and fastening bolts at freight trains (Zhan *et al.*, 2018). None of these approaches included industrial scenarios, like changes in illumination, heterogeneous wagon types or scalability.

The third subject concerns health estimation or safety assurance. Authors have tackled this problem with different techniques and usability concepts. A first group such as (Hamey, Watkins and Yen, 2007) and (Fernandes *et al.*, 2018) tackled the single component inspection problems, developing algorithms for the inspection of a certain component. Another more descriptive approach by (Lisanti *et al.*, 2018) was to install thermal cameras as well as normal cameras enhanced with pattern recognition algorithms to provide the location of multiple interest components for the inspector to analyze. A comparison of different health estimation methods, such as texture analysis, SVM and CNN, was done by (Rocha *et al.*, 2018). The mentioned approaches neither tackle the high diversity of wagons present at industrial scale nor the different image conditions due to changes in the acquisition such as light and wagon speed.

Further research regarding computer vision approaches in the railway sector has also been done. In the following, some up-to-date examples are mentioned. First (Liu *et al.*, 2017) verified if the posture of the vehicle is coherent with the rails by using laser projection. Multiple approaches to detect “fasteners” in rails have been used by (He *et al.*, 2019) and (Peng *et al.*, 2019). At last, (Karakose *et al.*, 2018) and (Gavai, Alto and Wu, 2019) went a step further and detected defects on the rails using different methods.

All these publications show a general interest of researchers and the industry for using computer vision techniques to facilitate identification, inspection, and monitoring of the state of the vehicles or the rails without perturbing the normal operations. Our focus is on the development of methods capable of learning from highly heterogeneous wagons data, as well as solutions to the practical challenges of their implementation.

3. System and concept

In response to the GCU (*General Contract for the Use of Freight Wagons*, 2019) requirements and to enable a more automated maintenance process, the proposed system aims to automatically deliver useful and timely insights to the persons in charge of the maintenance process. The proposed overall system consists of four main stages (see figure 1, I - IV).

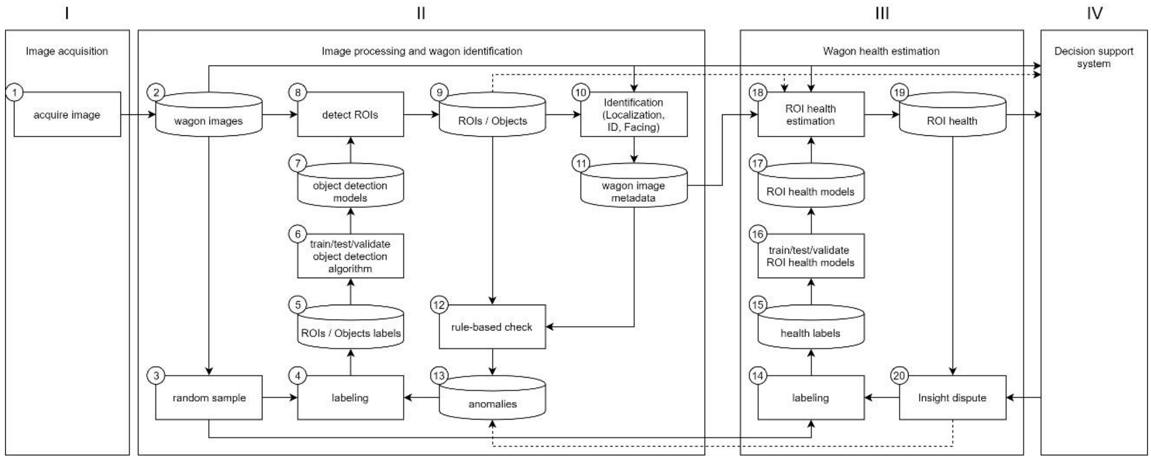

Fig.1 Full concept of the proposed system. Roman numbers represent stages, Arabic numbers the components and the arrows the data flow

In order to acquire usable images of cargo wagons for component inspection and further health estimation, the process must fulfill three initial requirements to properly work in an industrial environment. First, the acquisition must be done without disrupting normal operations. Second, the device must be able to illustrate the wagon from different perspectives. Third, the system must be able to continuously acquire images from multiple wagons linked together. Thus, a multi-camera bridge was implemented to fulfill the mentioned requirements. Said camera bridge use multiple line scan cameras from multiple perspectives to scan the wagons passing through, and composes the images by taking into account the speed of the wagons. This multi-camera bridge handles the image acquisition (figure 1, component 1) in the overall system and is explained in detail in Section 4.

The second stage of the system deals with the initial image processing and wagon identification. Its main goals are the localization of the wagon within the image, the extraction of the wagon UIC ID and the recognition of the facing direction of the wagon. To achieve the goals in this stage, the algorithm must reliably locate multiple objects on high resolution images with heterogeneous conditions (illumination, type of wagons, image corruption, and ratio), and therefore needs meaningful labeled data. The required steps are shown in figure 1 by components 2 to 13, which can be divided into three parts:

- The first step is a state of the art object detection algorithm (components 2 to 9). It starts by receiving and storing the acquired wagon images in a cloud based object storage (component 2). Then, the incoming dataset is sampled randomly (component 3) to obtain a subset of images for labeling (component 4). As a result, a set of ground truth annotations in VOC format is obtained (component 5). With the obtained labels, and relying on data augmentation, multiple state of the art object detection algorithms where trained and compared (component 6), before saving the resulting models as well as their validation metrics (component 7). Consequently, the object detection models are used to detect the necessary regions of interest (ROI) of the wagon images (component 8) and the resulting bounding boxes are saved for further analysis (component 9).
- The second step is the identification of the wagons. Component 10 starts by localizing the wagon in the image. This is directly obtained with the location of the bumpers at each side of the wagon, and allows the cropping of the excessive background, as well as to obtain reference points for further analysis. Subsequently, the facing direction is determined by localizing a maintenance plate (present in the left face of the wagon) or the lack of it. Later, the ROI containing the whole UIC ID number is extracted and then the cropped image containing the characters is parsed by an OCR algorithm. Finally, the results are saved as part of the metadata of the image (component 11).
- Finally with the third step, components 12 and 13 focus on identifying anomalies in previous processes. By using the obtained metadata and knowledge about the characteristics of a wagon, a set of rules is used to check the images regarding coherence. The identified corruptions or anomalies are then saved for reprocessing, enabling a second loop of labelling and training similar to other active learning approaches. It also facilitates the transfer of the trained model between different types of wagons.

The third stage of the system is the estimation of the health of the wagons. Its goal is to use the previously detected areas of interest to systematically estimate the deterioration of different parts of the wagons, as well as to compare them historically and fleet-wise. Some of the challenges to handle in this stage are the artifacts generated by the image acquisition, the lack of a large label database, as well as the unbalanced occurrence of faulty parts in any random sample of images. The solution proposed on the diagram includes a training and inference loops of classification models (components 14 to 19), together with a dispute loop (enabling active learning) to include any feedback given by the experts that will use the generated insights (component 20).

Lastly, the fourth stage is called the ‘decision support system’, it aims to communicate the obtained insights to the maintenance and inspection staff, as well as to collect the feedback of the users for the continuous improvement of the detection and health estimation algorithms. Thus, this component is planned as a set of web interfaces and rest services that allow (a) the selection and visualization of wagons and the associated insights, (b) visualization of historic evolution of ROIs per wagon, (c) fleet comparison of single ROIs, (d) annotation of comments, feedback and disputes, and (e) annotations of new ROI elements to be analyzed.

4. Methodology

The current section briefly describes the methods and ideas on how to acquire images of cargo wagons without disrupting normal operations as well as how to identify wagons (location, wagon UIC ID and facing direction) by using deep learning algorithms. The methods, as well as the results of the current paper, will focus on these two steps, which are the basis to continue the development of the health estimation and the decision support system.

4.1. Image acquisition

The multi-camera bridge was set up with six line scan cameras, three area scan cameras and two laser scanners that are connected to a customized server via Ethernet through a switch. In addition, six LEDs were installed to illuminate at the arrival of a train. The arrival detection and velocity measurement are managed by a laser scanner (Sick LMS111). Two line scan cameras (Basler racer raL2048-48gm) with a 20 mm optic are responsible for capturing images from the lower wagon part. The upper wagon part is recorded through four line scan cameras

(Basler racer ruL2098-10gc and Basler racer raL4096-24gm). Two-dimensional images are built out of multiple line scans depending on the velocity measurement which is done once per wagon at arrival. The resulting images are 2048, 2098 and 4096 pixels high respectively. The length of the images varies depending on the amount of lines used to create the images (between 10K and 70K).

The selected result and the compressed images are stored in a database. If a train consists of multiple wagons, a second laser scanner (Sick LMS111) is responsible for the division of wagons. After a train has passed the LEDs are switched off and the system turns into the idle state waiting for the next event.

Fig. 2 Upper image of cargo wagon. Green rectangles are labels of bumpers and UIC ID

4.2. Image processing and wagon identification

The starting step in this stage was the selection of two datasets. The first dataset is a random sample of the acquired images, containing multiple wagon types and camera angles. It was used to label objects of interest (see figure 3). The second dataset is composed of images of the lower section of wagons type 305. The wagon type was selected, as it is the most prevalent wagon type in the analyzed fleet and it will be used to verify the performance of the algorithms in combination with rule-based checks.

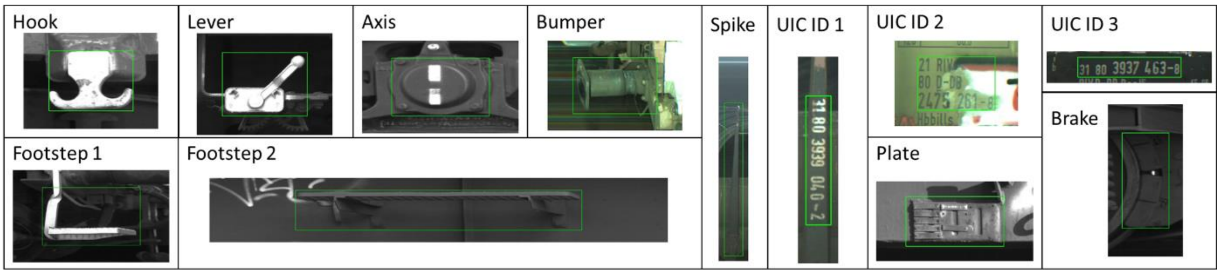

Fig. 3 List of labeled objects

Following the annotation of the first dataset, multiple deep learning algorithms were trained to detect the labeled objects, from which two main variants will be presented in the results, Faster RCNN (Ren *et al.*, 2017) and Cascade RCNN (Cai and Vasconcelos, 2018). Both models use a ResNet-101 backbone and a balanced feature pyramid network. These architectures not only tackle the traditional object unbalance problems in training (Pang *et al.*, 2019), but also allow the detection of bigger, medium and contained objects in the cargo wagon images. Data augmentation techniques were used during the training to increase the generalization capabilities of the networks, as well as to ensure the robustness of the detectors. One type of augmentation where based on photometric distortions, by modifying the brightness, contrast saturation and hue of the images. These transformations are meant to help the algorithms generalize different acquisition conditions with a small set of labeled data. Other transforms were also implemented to increase the robustness of the algorithms, such as random crops in areas of interest (to account for occlusion or deterioration), flipping and the addition of Gaussian noise.

Subsequently, the wagon identification step used the acquired images and the trained models to obtain the needed information regarding the location of the wagon, the wagons UIC ID and the facing direction of the wagon. The wagons are not always centred within the high-resolution images resulting from the image acquisition stage (see figure 2). This means that large parts of the images may not contain wagons at all and that the relative position of the wagon in the images is unknown. Hence, before continuing the analysis, the starting and ending points of the wagons must be detected. This detection can be done in multiple ways, by using the change in axis-wise metrics, using ORB features (Rublee *et al.*, 2011) and homography, patch matching, or by using the previously trained object detection algorithms to locate the bumpers of the wagons. All the mentioned methods were tested, with the object detection approach showing superior results as well as synergy with the rest of the analysis.

Even without considering the travelling direction, each wagon has two facing possibilities. Because most wagons are symmetrical, the facing direction cannot be determined by the general structure of the wagon. As part of the wagon standards, each wagon must have a maintenance plate (see figure 3) on its left side. Hence, the previously trained models are used to identifying this plate or the lack of it.

The final part is the identification of the Wagon UIC ID. The GCU state that each side of a wagon must contain at least one UIC ID. Said UIC ID (see figures 3) will be used to map a specific wagon and their type to the corresponding pictures. False positives are a serious concern in the identification, as the wagon contains other text unrelated to the ID as well as other “artistic” artefacts (see figure 2). Ultimately, the presence of occlusions, as well as the size of the image, can pose problems for OCR algorithms, thus, making the ROI extraction an instrumental goal towards more robust OCR training. By using the trained models, the region containing the UIC ID can be extracted and passed to an OCR algorithm to obtain the wagon ID in a usable format, or detection of an irregular set of characters.

With the wagon correctly identified, a list of rules related to the wagon type was used to verify the detected ROIs. The wagon type 305 for example always has two double wheel axles, thus four brakes, in the lower image. The images that do not pass the rule-based check are redirected to be annotated again, either to locate non-detected objects or to annotate severe anomalies in the parts of interest.

5. Results and discussion

The overall vision of this system is the transformation of the current maintenance process, from reactive and time based to an intelligent predictive process. The most elementary requirement for analyzing components lifecycle cost and identifying relevant factors for the wear development is the continuous knowledge of their condition.

5.1. Image acquisition

On average 1200 wagons are passing the multi-camera bridge every day with a data volume of 30 megabytes per wagon. About 57% of the captured wagons are owned by DB Cargo. To improve the identification rate, all wagons of DB Cargo are equipped with radio-frequency identification (RFID) tags and can be interpreted by an RFID reader installed in front of the multi-camera bridge. The resulting information obtained by the RFID reader can later be used to validate the image-based identification process.

The location of the multi-camera bridge generates a sporadic non-constant stretching or compression of the images. The multi-camera bridge is installed at bottom of the hump at a marshalling yard. Consequently, wagon acceleration occurs and can disrupt the composition of the multiple line scans since the train velocity is measured only once per wagon (see figure 4).

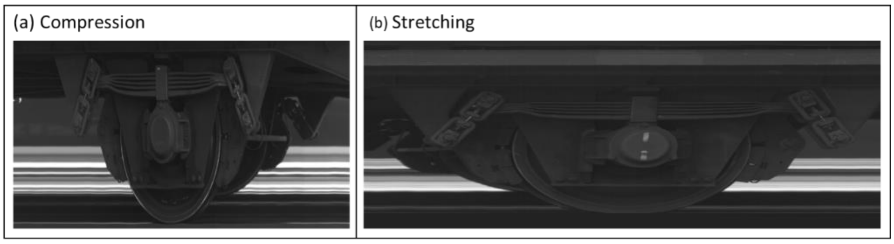

Fig. 4 (a) Image compression; (b) Image stretching

A preventive remedy for this restriction could be a continuous velocity measurement of the wagons to avoid defective compositions of line scans or a post-processing of the images. With a larger set of the detected ROIs or key points, their relative positions can be used to compute the stretching factor in multiple parts of the image.

5.2. Image processing and wagon identification

Table 1 presents the results of training the selected object detection models (Cascade RCNN and Faster RCNN) for a variety of objects. In a first experiment both models were trained to detect all the objects presented in figure 3. Three more experiments were run, training said models to detect Bumpers, UIC ID and Plates individually. The algorithms were implemented using the Pytorch framework, in an Ubuntu 18.04 server with an Intel Xeon Silver 4116 CPU of 2.10 GHz, 128 GB of RAM and a Tesla V100 GPU.

Table 1. Object detection comparison.

Model	Objects	Accuracy*	mAP_{50}	mAP_{75}
Cascade RCNN-Resnet-101	All	98.628	0.652	0.528
Cascade RCNN-Resnet-101	Bumper	99.530	0.956	0.854
Cascade RCNN-Resnet-101	UIC ID	99.772	0.936	0.694
Cascade RCNN-Resnet-101	Plate	99.716	0.366	0.002
Faster RCNN-Resnet-101	All	98.549	0.632	0.525
Faster RCNN-Resnet-101	Bumper	99.447	0.955	0.876
Faster RCNN-Resnet-101	UIC ID	99.826	0.839	0.593
Faster RCNN-Resnet-101	Plate	99.843	0.020	0.000

The initial results of the trained models show the feasibility of this approach. With a Cascade RCNN model to detect all the objects shown in figure 3, a mean Average Precision (mAP_{50}) of 0.652 was obtained, and the head of the detector was able to differentiate the correct objects with an accuracy of 98.628%. When training to detect single objects, the mAP_{50} of Bumpers and UIC IDs increases up to 0.956. The maintenance plate only obtained a mAP_{50} of 0.366, which indicates difficulties in its detection, possibly because of its lack of notable features. Given the acquisition challenges explained previously, most of the models are robust enough to work under different industrial conditions.

The bumpers of the wagons have proven easy to detect, with a mAP_{50} of 0.956 and an accuracy of 99.53%. With said results, the models can be used to properly locate the wagons in the images. As seen in figure 5, the bumpers can be detected successfully even in harsh contrast conditions.

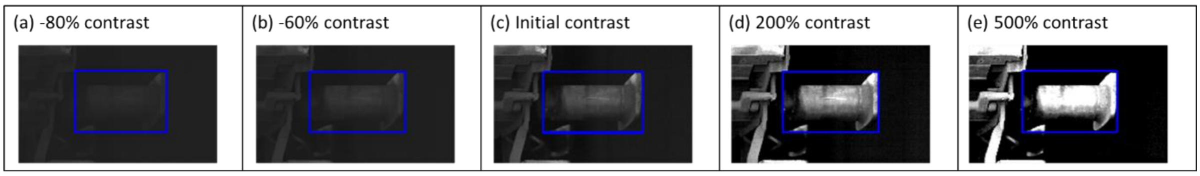

Fig. 5 bumper detection at different contrast levels

The mean Average Precision (mAP) of the plate detection is unsatisfactory. Although the algorithms can successfully differentiate an extracted ROI as a plate with a 99.71% accuracy, the precise location of the plate is not easily obtained (see figure 6). These results can be attributed to the lack of constant salient features in plates as well as their small size compared to the scale of the images. It must also be noted that the plate is detected to assess the facing direction of the wagon and thus, a mAP_{50} of 0.366 is insufficient and will require further work to increase the localization of the maintenance plates to an acceptable level.

Fig. 6 Plate detection with bounding box offset

In the case of UIC ID numbers, the results are satisfactory, with a mAP_{50} of 0.936 and an accuracy of 99.77%. This also shows that the object detection approach is a valid alternative to methods such as MSER proposed in (Xiang *et al.*, 2016). This method has the advantage of also solving the issues of positioning and facing direction estimation. In some cases, non-standard paintings can also trigger the detections as seen in figure 7. The occurring false positives can be minimized with the increasing number of labelled images.

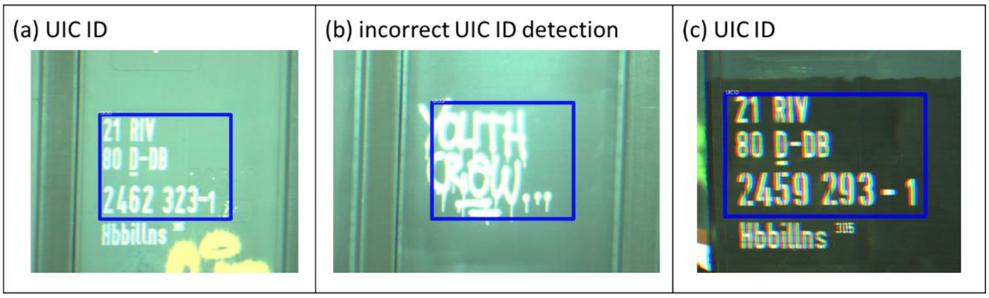

Fig. 7 (a) UIC ID detection; (b) incorrect UIC ID detection; (d) UIC ID detection

After the overall analysis, a closer look at the wagon type 305 was done. For this type of wagon, the following rules apply. In the lower section of the wagon, two hooks, two axis, four brakes, two bumpers, one footstep of each type and three levers are expected. On the upper images, exactly one UIC ID is expected. In order to comply with the changes of scale, two groups of bumpers separated by at least a third of the image have to be found. This set of rules allowed the identification of edge cases such as figure 7-b. Other acquisition artifacts were also identified with this technique; Examples of these are incomplete (cut) wagons and images with sudden changes in illumination. In this test, 1729 images of upper and lower wagon (type 305) sections where used. A total of 96.64% of the UIC ID were correctly identified. From the 58 non-compliant images, 9 consisted in occluded UIC IDs, 4 where non-standard formats and the rests where double detection including the UIC ID and an unrelated text (see figure 7-b). The bumpers where detected 89.64% of the times. In the remaining 179 images, one side of the wagon was limiting the border of the image, leaving the bumpers outside. In the case of the maintenance plate, half of the images should trigger a detection, but only in 16.18% of the images, the plate was detected. This can be attributed to the lack of consistency regarding the plate characteristics (see figure 6), as well as the light and contrast changes in the images. Thus, further labelling and training iterations are needed for a reliable result.

5.3. Wagon health estimation

While the first two sections of the proposed system are being tested, the concept for the health estimation of the cargo wagons has also been designed. The general pipeline has been described in section 3 and resembles significantly stage 2 (image preprocessing and wagon identification). The most notable differences are the types of models that will be trained to assert the health of the previously detected components, and the insight dispute mechanism.

Each detected region of interest contains objects of different natures and said objects must comply with a different set of requirements from the GCU. This means that for regions such as the UIC ID, occlusion or paint deterioration are significant health issues that must be corrected, which is not the case for brakes and bumpers. Based on related works, a classification approach using deep learning techniques can be used to assert the health status of the components. However, this approach raises the challenge of correctly labelling and modeling highly unbalanced datasets. In the case of structural components such as the brakes, key points can be detected (in a similar fashion as the ROI detection) in multiple standard locations of the pads, asserting their relative width and using this metric for the health estimation. Lastly, autoencoders can be used to detect anomalies and to generate an explanation of the affected/deviant sections of the analyzed objects. It must be clarified that in all cases, dirt and acquisition artifacts are challenges that will require further work and consideration.

Once the health of the wagon components has been estimated, it will be communicated through the decision support system. These insights will be shown as part of the wagon diagnostics, pointing to the regions that must be verified by the inspection staff. Said staff will have the option to dispute the automatic health assessment and create new labels that after verification will be propagated to stages 2 and 3. This insight dispute mechanism and a double verification of the odd cases will ensure continuous improvement of the overall system.

5.4. Decision support system

The current concept for the decision support system will be implemented as a set of web services and interfaces as described in section 3. As opposed to stand-alone applications, the web interfaces will allow a faster integration with the rest of the data infrastructure. Users will be able to navigate from one type of web interface to the next while keeping their focus on the current task at hand. Said tasks are mainly divided into single wagon inspection, ROI comparison, insight explanations, and annotations.

The single wagon inspection interfaces will focus on conveying the insights generated from the previous analysis, by highlighting over the wagon images, the regions that contain anomalies or low health indications. This interface will serve as a fast overview of the health of the wagons and is meant to be used on the normal operations of the inspection staff. Once an anomaly has been highlighted and the users desire to know more about the insight, the interface will provide links to either a ROI comparison or an insight explanation.

In the case an explanation is needed, an interface will present the selected ROI, the result of the health estimation algorithms as well as any provided explanations. Given the current progress in the field of explainable AI (XAI), multiple types of explanations can be used to generate a better understanding of the detected health issues. Such explanations can be on the forms of saliency maps (Simonyan, Vedaldi and Zisserman, 2013), healthy and unhealthy sample images (Ribeiro, Singh and Guestrin, 2018) or confidence of the detection. This interface will also allow the user to dispute the insights, giving valuable edge case feedback for the knowledge base.

When a ROI is selected for comparison, the user may decide to compare it on a fleet level or through time. The fleet comparison will display ROI through a fleet of wagons, grouping them by health levels and similarity. For example, a user may compare the state of the brakes through a wagon type, to gain a better understanding of the most common types of issues. An ROI of a specific wagon may also be compared through time, by showing the evolution of the region, the previous health estimations and possible health evolutions. In both cases, the most important challenge lies in the user interaction and creating an interface that displays high volumes of information without bloating the users.

Finally, an annotation tool will also be provided, to enable expert staff to resolve any dispute and to label new images. In this case, a set of suggested labels will be shown along with the wagons images and the user will decide to correct them or confirm the annotations.

6. Conclusion and Outlook

In the preceding publication, a novel system to advance the semi-automated inspection of cargo vehicles was presented. The overall vision of this system is the transformation of the current maintenance process, from reactive and time based to an intelligent predictive process. Such a complete pipeline from acquisition of the image up to the delivery of the insight to the corresponding user has not been explored in the state of the art. Especially, the feedback loops for active learning by using rule-based checks and human in the loop disputes are a novelty that focuses on the scalability of this approach.

The design of the system can be summarized in four parts, namely the image acquisition phase, the identification phase, the health estimation phase and finally the subsequent decision support system. The non-invasive way of acquiring high-resolution pictures of moving wagons, the automated identification of the vehicle as well as the detection of regions of interest are the focus of this paper.

The technical structure of the build multi-camera bridge was presented and the techniques for taking the images described. Subsequently, different algorithms for identifying several objects within the captured images were applied and evaluated. The recognized objects can be divided into two groups. The first group was used to identify the wagons UIC ID, its facing direction as well as its exact location within the picture. The second group contains different regions of interest, such as brakes, axis, footsteps, for the health estimation stage. At the same time, all objects serve to identify erroneous images as well as to increase the training base.

The best achieved mAP_{50} for multiple ROIs such as bumpers and UID ID are 0.956 and 0.936 respectively. Therefore, this paper serves as proof of concept for the application of AI methods for semi-automated condition analysis of freight wagons in real-life application scenarios.

Future work will focus on the health estimation part and the decision support systems based on it. Novel insights of the wear and tear can be generated using the described system by comparing historic data of the vehicles as well as using fleet analysis techniques. The most elementary requirement for analyzing components lifecycle cost and identifying relevant factors for the wear development is the continuous knowledge of their condition.

This requires the integration of artificial intelligence (AI) as support for the employees. Pre- and supporting analysis using AI can automate routine tasks and focus employees' attention on anomalies and deviations. The idea of implementing various (use case specific) assistance systems will effectively speed up and specify the diagnosis of maintenance tasks for freight wagons. The approaches described in this paper form the basis for these optimizations.

Acknowledgement

The authors acknowledge the financial support by the Federal Ministry of Transport and Digital Infrastructure of Germany in the project QUISS (project number 19F2060).

References

Cai, Z. and Vasconcelos, N. (2018) 'Cascade R-CNN: Delving into High Quality Object Detection', *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 6154–6162. doi: 10.1109/CVPR.2018.00644.

Fernandes, E. *et al.* (2018) 'An Ensemble of Convolutional Neural Networks for Unbalanced Datasets : A case Study with Wagon Component Inspection', *2018 International Joint Conference on Neural Networks (IJCNN)*. IEEE, pp. 1–6. doi: 10.1109/IJCNN.2018.8489423.

Gavai, G., Alto, P. and Wu, W. (2019) 'Hybrid Image-based Defect Detection for Railroad Maintenance', pp. 1–7.

General Contract for the Use of Freight Wagons (2019).

Hamey, L. G. C., Watkins, T. and Yen, S. W. T. (2007) 'Pancam: In-service inspection of locomotive pantographs', *Proceedings - Digital Image Computing Techniques and Applications: 9th Biennial Conference of the Australian Pattern Recognition Society, DICTA 2007*, pp. 493–499. doi: 10.1109/DICTA.2007.4426837.

He, B. *et al.* (2019) 'Railway Fastener Defects Detection Using Gaussian Mixture Deformable Part Model Railway Fastener Defects Detection Using Gaussian Mixture Deformable Part Model'. doi: 10.1088/1742-6596/1302/2/022102.

Karakose, M. *et al.* (2018) 'A new approach for condition monitoring and detection of rail components and rail track in railway', *International Journal of Computational Intelligence Systems*, 11(1), pp. 830–845. doi: 10.2991/ijcis.11.1.63.

Li, B. *et al.* (2019) 'A Cascade Learning Approach for Automated Detection of Locomotive Speed Sensor Using Imbalanced Data in ITS', *IEEE Access*, 7, pp. 90851–90862. doi: 10.1109/access.2019.2928224.

Lisanti, G. *et al.* (2018) 'A multi-camera image processing and visualization system for train safety assessment', *Multimedia Tools and Applications*, 77(2), pp. 1583–1604. doi: 10.1007/s11042-017-4351-4.

Liu, D. *et al.* (2017) 'A real-time posture monitoring method for rail vehicle bodies based on machine vision', *Vehicle System Dynamics*, 55(6), pp. 853–874. doi: 10.1080/00423114.2017.1284339.

Ningning, L. *et al.* (2016) 'An Algorithm of Freight Train Number Locating Based on Template Matching and Morphological Operations', (2), pp. 102–107.

Pang, J. *et al.* (2019) 'Libra R-CNN: Towards Balanced Learning for Object Detection', (2). Available at: http://arxiv.org/abs/1904.02701.

Peng, Z. *et al.* (2019) 'A Multifeature Hierarchical Locating Algorithm for Hexagon Nut of Railway Fasteners', *IEEE Transactions on Instrumentation and Measurement*. IEEE, PP, pp. 1–7. doi: 10.1109/tim.2019.2905309.

Ren, S. *et al.* (2017) 'Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), pp. 1137–1149. doi: 10.1109/TPAMI.2016.2577031.

Ribeiro, M. T., Singh, S. and Guestrin, C. (2018) 'Anchors: High-precision model-agnostic explanations', *32nd AAAI Conference on Artificial Intelligence, AAAI 2018*, pp. 1527–1535.

Rocha, R. L. *et al.* (2018) 'A deep-learning-based approach for automated wagon component inspection', *Proceedings of the ACM Symposium on Applied Computing*, (September), pp. 276–283. doi: 10.1145/3167132.3167157.

Rublee, E. *et al.* (2011) 'ORB: An efficient alternative to SIFT or SURF', *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2564–2571. doi: 10.1109/ICCV.2011.6126544.

Simonyan, K., Vedaldi, A. and Zisserman, A. (2013) 'Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps', pp. 1–8. Available at: http://arxiv.org/abs/1312.6034.

Xiang, X. *et al.* (2016) 'ID localization and recognition for railway oil tank wagon in the industrial scene', *Proceedings of the World Congress on Intelligent Control and Automation (WCICA)*. IEEE, 2016-Septe, pp. 826–829. doi: 10.1109/WCICA.2016.7578476.

Zhan, Y. *et al.* (2018) 'A Unified Framework for Fault Detection of Freight Train Images under Complex Environment', *Proceedings - International Conference on Image Processing, ICIP*, pp. 1348–1352. doi: 10.1109/ICIP.2018.8451188.