Anomaly Detection for Vision-based Railway Inspection

Anomaly Detection for Vision-based Railway Inspection

Riccardo Gasparini[1], Stefano Pini[1], Guido Borghi[1], Giuseppe Scaglione[2], Simone Calderara[1], Eugenio Fedeli[2], and Rita Cucchiara[1]

[1] AIRI - Artificial Intelligence Research and Innovation Center
Università di Modena e Reggio Emilia, Italia
{riccardo.gasparini, s.pini, guido.borghi, rita.cucchiara}@unimore.it
[2] RFI - Rete Ferroviaria Italiana
Gruppo Ferrovie dello Stato, Firenze
{g.scaglione, e.fedeli}@rfi.it

Abstract. The automatic inspection of railways for the detection of obstacles is a fundamental activity in order to guarantee the safety of the train transport. Therefore, in this paper, we propose a vision-based framework that is able to detect obstacles during the night, when the train circulation is usually suspended, using RGB or thermal images. Acquisition cameras and external light sources are placed in the frontal part of a rail drone and a new dataset is collected. Experiments show the accuracy of the proposed approach and its suitability, in terms of computational load, to be implemented on a self-powered drone.

Keywords: Railway Inspection · Anomaly Detection · Computer Vision · Deep Learning · Self-powered drone

1 Introduction

A crucial element to guarantee the safety of rail transport is the visual inspection of railways, in order to ensure the absence of obstacles placed on the railroad track that could cause damages or even the derailment of trains. These inspection activities are generally conducted during nighttime, when the train circulation is usually suspended. In this context, due to the vastness of railroads, an automatic inspection system is strongly needed.
Therefore, in this paper, we propose a vision-based framework to tackle the obstacle detection task in videos acquired from a rail drone, *i.e.* a self-powered light-weight vehicle moving along railways and operated by remote control, which computes the analysis locally. To deal with the night time, the rail drone is equipped with thermal and RGB cameras, in addition to external light sources. The proposed framework is a combination of two sequential deep networks, an autoencoder and a binary classifier, as shown in Figure 1.

From the point of view of the computer vision research field, we interpret the detection of obstacles as the anomaly detection task, *i.e.* the ability to identify samples that exhibit significant differences with respect to a regularity.

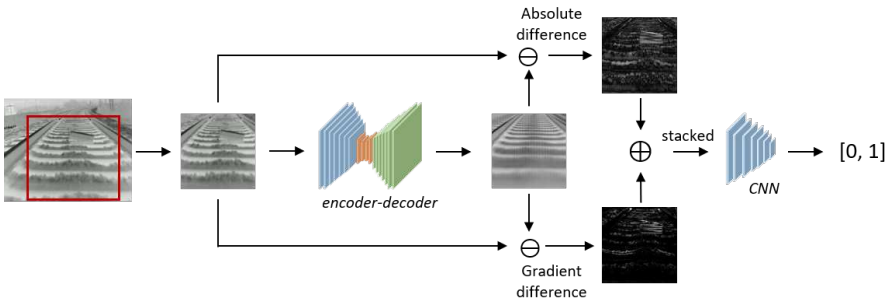


Fig. 1: Overall view of the proposed framework (using thermal data). From the left, the acquired frame is cropped, then fed into the autoencoder. The reconstructed frame is then used to compute the absolute and the gradient difference images that are stacked and fed into the classifier network. This classifier outputs the presence or absence of anomalies in the frame.

This task is a key element in many real-world applications, such as video surveillance [18], defect detection [20], reinforcement learning [27] and medical imaging [31]. In these applications, the acquisition sensor is often assumed to be in a raised and fixed position, resulting in images and videos with a static background [33]. In particular, this condition is present in industrial video-based systems [14] and video surveillance ones [30].
Furthermore, many approaches are based on supervised learning [16,3] that often require manual, time-consuming and expensive annotations along with the assumption that all anomalies are known during the training process.

Differently from these works, in this paper we investigate the anomaly detection task using images taken from a moving camera. Indeed, the acquisition devices are placed in the frontal part of the rail drone, close to the railroad.
Due to the lack of public railway datasets focused on the anomaly detection task, we collect more than 30k frames from a rail drone moving on the track during the night. The new dataset contains more than 50 recordings, with and without anomalies, acquired with multiple synchronized cameras, *i.e.* RGB and thermal cameras (used in this paper) in addition to stereo and depth sensors. As we focus on the railroad safety, we considered anomalies consisting in many categories of objects which are usually employed in rail yards, such as track lifting jacks, pickaxes, rail signals and so on. Samples of anomalous objects are depicted in Figure 2 in both RGB and thermal domains.

2 Related Work

2.1 Anomaly Detection on Railways

At the time of writing, there are no works that address the task of anomaly detection through visual data in the railway scenario during nighttime. Only

similar task have been addressed, such as track detection [12,17,36] and collision prediction [22,23]. Unfortunately, datasets are not often publicly available.

To detect obstacles on railways, many literature works exploit the use of infrared (IR) of ultra-sonic range sensors, usually placed in the frontal part of the train. For instance, [26] proposed a system based on a range sensor to perform obstacle detection. Specifically, an infrared emitter is exploited and a light turns on when an object is detected within the (limited) working distance. A framework using GSM and GPS modules is proposed in [29]: similar to the previous work, an infrared emitter, in combination with the other modules, is exploited to detect obstacles in front of the train. A LiDAR is exploited in [24]: the sensors is coupled with a camera to detect obstacles on railway tracks. In [15] pairs of infrared sensors are places on both the railway sides: a lack of connection between the two devices, specifically an emitter and a receiver, reveals the presence of obstacles.

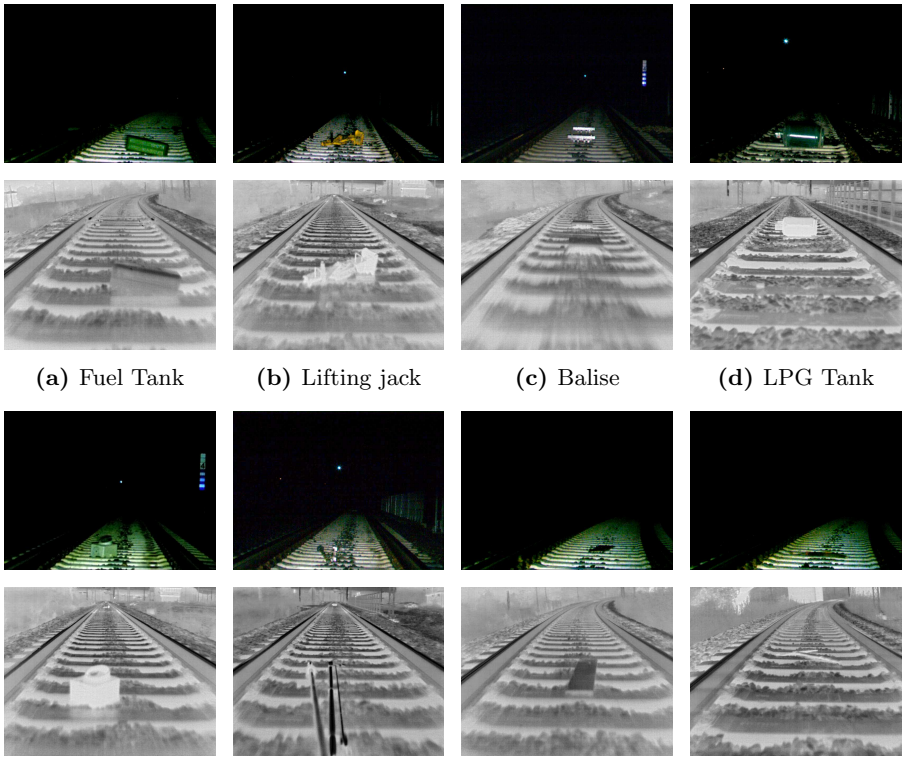
We note the scarcity of publicly-released dataset in this research field. Only recently, a public dataset for semantic scene understanding, acquired from the point of view of a train and a tram, namely *RailSem19* [34], has been introduced. RailSem19 contains specific annotations collected for a variety of tasks, including the classification of trains, switch plates, buffer stops and other elements related to the railway scenario, but not anomalies and obstacles.

In general, existing works addressing the anomaly detection on railways are often *ad hoc* systems, created for a specific scenario and employing specific infrared emitters. There is a lack of systems based only on vision-based systems.

2.2 Anomaly Detection in Computer Vision

From a general point of view, literature work are categorized in two different approaches: reconstruction-based models and probabilistic methods.
The former learn a parametric reconstruction of normal data through different methods, such as sparse-coding algorithms [35,11], deep encoder-decoder architectures [18] or GANs [30,13]. A similar approach is the future frame prediction, in which anomalies are detected comparing the differences between a predicted future frame and the current one [21].
The latter approximate a density function of motion features and normal appearance. In this case, optical flow and trajectory analysis, exploiting non-parametric [2] and parametric [5] estimators, are often used.

Highly-dynamic scenarios, such as images taken from a moving rail drone for the railway inspection, represent a tough challenge to these state-of-art methods based on fixed cameras. Only recently, an unsupervised approach has been proposed for the traffic accident detection [33] in which the acquisition device is a dashboard camera. In [10], a dataset of crowd-sourced dashcam images is presented and a supervised method that detect anomalies, in terms of motorbike and car collisions, is proposed. Abati *et al.* [1] introduce an anomaly detection method capable of working in the automotive scenario [25]. However, the visual content is purposely discarded and only eye fixations are employed.



(a) Fuel Tank **(b)** Lifting jack **(c)** Balise **(d)** LPG Tank

(e) Traffic light **(f)** Insulating stick **(g)** Rail signal **(h)** Pickaxe

Fig. 2: Some examples of anomalies included in the acquired data. The upper row contains frames acquired with the RGB camera (and external illuminators), while the lower row reports the same classes collected through a thermal device.

3 Data Acquisition

As mentioned above, we record a new dataset to overcome the lack of public railway datasets. Data has been collected placing multiple sensors in the frontal part of a rail drone, very close to the cobbled road.
The acquisition activity has been done during the night: to the best of our knowledge, this is the first dataset collected for the anomaly detection task in the nighttime railway scenario.
Therefore, the acquisition system needs to comply with three main requirements, derived from the automotive context [8,7]:

– **Fast Acquisition**: since cameras are placed on a rail drone, the frame rate and the shutter speed of the acquisition devices must be sufficiently high to avoid motion blur caused by the high speed of the drone (up to 100km/h).

- **Night Vision**: acquisition devices must deal with the night time. In this context, the adoption of external light sources and the use of thermal cameras is required. Since the acquisition system is places on a self-powered rail drone, it is important to limit the power consumption of the light sources.
- **High Resolution**: in order to detect even small-sized anomalies at long distances, the sensors must have a high spatial resolution.

To conform to these requirements, the following cameras and light sources are employed:

- **Basler acA800-510uc**[3]: this is an industrial camera with an extremely-high frame rate (more than 500 fps) that however is limited to a low spatial resolution (800 × 500 pixels). We equipped this camera with a 12.5 - 75mm zoom lens. With this camera, external light sources are needed.
- **Light sources**: we use two types of light source. The first one is the *LED Light Bar 470*[4]: this headlamp is a compact lightweight bar, having a low profile and a power consumption of only 35W. It is useful to illuminate wide areas close to the drone. The second light source is the *Comet 200 LED*[5]. Being a high-beam headlamp with a power consumption of 13W and only 495g of weight, it is useful to illuminate areas that are far from the drone.
- **Flir Boson 640**[6]: this is a high-resolution thermal camera, having a spatial resolution of 640 × 480 pixels, which is able to acquire up to 60 frames per second. Its small-size form factor (21 × 21 × 11mm), limited weight (7.5g) and low energy consumption (only 500mW) make it suitable to be installed on a rail drone. The camera is equipped with a 14mm lens.
- **Zed Stereo camera**[7]: this is a stereo camera carefully designed for the outdoor setting. The spatial resolution is 4416 × 1242 pixels, the acquisition range is up to 20 meters of distance and the acquisition rate ranges from 15 to 100 frames per seconds (depending on the resolution). To have real time performance at the maximum resolution, it requires a dedicated graphic processing unit (GPU).

In the acquired data, anomalies are objects placed on the railroad track. We select and employ the following objects, which are the common tools used in the construction sites along the railways:

- Electrical Insulator
- Fuel Tank
- Rail Signal
- Pickaxe
- Locking turnout
- Track lifting jack
- Traffic light
- Insulating stick
- LPG tank
- Balise
- Oiler

[3] https://www.baslerweb.com/en/products/cameras/area-scan-cameras/ace/aca800-510uc
[4] https://www.hella.com/truck/it/LED-LIGHT-BAR-470-Single-Twin-3950.html
[5] https://www.hella.com/offroad/it/Comet-200-LED-1626.html
[6] https://www.flir.it/products/boson
[7] https://www.stereolabs.com/zed

A sample of each of these classes is depicted in Figure 2 in RGB and thermal domains. Every frame is annotated with two labels: the presence of one or more obstacles (*i.e.* whether the frame contains an anomaly) and the location, expressed with bounding boxes, of each visible obstacle.

4 Proposed Framework

We propose a deep learning-based framework based on 2 sequential modules. The first one is an autoencoder network [6], *i.e.* an encoder-decoder architecture whose goal is to reconstruct the input frame, while the second one is a binary classifier network [4], predicting if the input frame contains or not an anomaly (*i.e.* an object placed on the railway).
The whole system is depicted in Figure 1 and described in the following.

4.1 Autoencoder

As mentioned above, the first module of the framework is an autoencoder that aims to reconstruct the input frame passing through an intermediate bottleneck. The input of the model is a single frame while the output is the reconstructed one. During the training, the network receives as input only regular frames, *i.e.* frames without any anomaly. In this way, the network should learn to reconstruct only normal frames thus the output should always result in a clean image devoid of any anomaly, even if the input frame contains anomalies.
Finally, the reconstructed frame is compared with the original input frame, through an absolute and a gradient difference, *i.e.* a difference computed on the gradients of the two images. The resulting 2 difference images are stacked and used as input for the second module, the classifier.

Model. This neural network accepts as input images with a spatial resolution of 192×192 pixels. The encoder architecture consists in 9 convolutional layers with kernel size 3×3. The first and last two layers have stride $s = 1$ while other layers have stride $s = 2$. The decoder architecture is symmetrical: it is composed of 9 transpose convolutional layers, to up-sample the input feature maps, with kernel size 3×3. The first two and the last layer have stride $s = 1$ while the remaining ones have $s = 2$.
Regarding the feature maps, their size is doubled (and then halved) at each layer, except for the first one, starting from 16, arriving to 1024 in the bottleneck, and then reducing down to 16 again at the end of the decoder architecture.
The final output is a 192×192 pixels image. We exploit the *Leaky ReLu* [32] activation function with slope $\alpha = 10^{-2}$. This deep architecture have $\simeq$ 22M parameters.

Training procedure. We train the autoencoder with an unsupervised approach since, as mentioned above, the network receives only frames without anomalies

during the training procedure. We adopt two different loss functions. The first one is the *Mean Squared Error* loss, here referred as L_{MSE}, defined as:

$$L_{MSE} = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \| I_I(m,n) - I_R(m,n) \|_2^2 \tag{1}$$

where I_I, I_R are the input and the reconstructed image, respectively, of size $M \times N$ pixels.
In addition, we propose to use a *Gradient Loss* (L_G) defined as:

$$L_G = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \| G_{I_I}(m,n) - G_{I_R}(m,n) \|_2^2 \tag{2}$$

where G_{I_I} and G_{I_R} are the gradients computed on the input (I_I) and the reconstructed (I_R) images with a spatial resolution of $M \times N$ pixels:

$$G = \sqrt{G_x^2 + G_y^2} \tag{3}$$

$$G_x = I * \begin{vmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{vmatrix}, \quad G_y = I * \begin{vmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{vmatrix} \tag{4}$$

in which the $*$ symbol is the convolution operator.
These equations, introduced in [28], perform the calculus of the gradients along both the horizontal and vertical dimension of an image. Minimizing this loss function is equivalent to improve the definition of lines and contours in reconstructed frames.
Finally, the general loss L is defined as a weighted sum, taking inspiration from [9], of L_{MSE} and L_G, as follows:

$$L = \alpha \cdot L_{MSE} + \beta \cdot L_G \tag{5}$$

In our experiments, we set $\alpha = \beta = 1$, the learning rate is set to 10^{-3} and the *Adam* [19] optimizer is used.

4.2 Classifier

This module is a deep binary classifier, predicting if a frame contains or not anomalies. The input is represented by the two (absolute and gradient) difference images stacked together. The output is a binary label representing the presence or the absence of any anomaly.
Using the two difference images as input, the network can use both the variations in terms of textures and the variations in terms of contours and lines.

Model. This neural network is a lightweight CNN that shares the architecture with the encoder module described previously, but the number of filters is halved and thus ranges from 8 to 256. Moreover, the last convolutional layer is removed and replaced with a flatten operation and 2 sequential linear layers with 48 and 2 units. We add a dropout regularization with drop probability $p = 0.3$ between the linear layers and, as in the autoencoder model, we exploit the *Leaky ReLu* [32] activation function with slope $\alpha = 10^{-2}$. This network contains $\simeq$ 700k parameters.
The output of the model is a binary classification which corresponds to 1 if the frame contains anomalies and to 0 if it does not.

Training procedure. We train the binary classifier with a supervised approach, using both frames with anomalies and frames without anomalies during the training procedure. The *Binary Cross Entropy* (L_{BCE}) loss is employed as objective function:

$$L_{BCE} = -(y \log(p) + (1 - y) \log(1 - p)) \tag{6}$$

in which log is the natural log, p is the predicted probability of a class and y is the binary label corresponding to anomalous and non-anomalous frames.

Table 1: Results of the proposed framework for both RGB and thermal data. The system achieves satisfactory results, confirming the applicability to real-world applications. The usage of thermal data results in higher scores.

Input type	Accuracy	Precision	Recall	F1-score
RGB	0.811	0.979	0.719	0.825
Thermal	0.966	0.989	0.957	0.973

5 Experimental evaluation

In this Section, we evaluate the proposed framework using as input the RGB images (converted in gray-scale), acquired with the *Basler* camera (supported by the external light sources), and the thermal images, collected by the *Flir Boson* thermal camera. Further details about the acquisition devices are reported in Section 3.
In order to train and test the proposed framework, we create appropriate splits: we group all the frames containing anomalies and randomly sample about 80% of the frames for the training and validation phases and the remaining 20% for the testing one. Then, we randomly sample an equivalent number of regular frames from the dataset and we add them to the training, validation and testing splits.

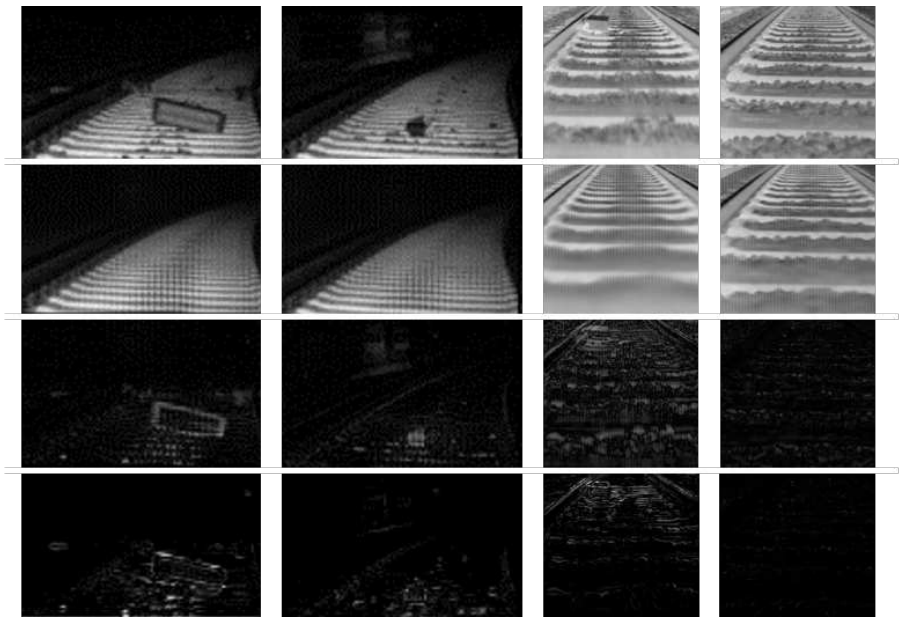


Fig. 3: Sample output of the autoencoder network of the proposed system, both for the intensity (left) and the thermal (right) domains. The first row represents the acquired frames used as input. The second row contains the reconstructed frames, while the third and fourth rows shows the absolute and the gradient difference images, respectively. The last columns reports the case in which no obstacles are placed on the railway.

For all experiments, we exploit the following common metrics: prediction accuracy, precision, recall and F1-score.

We report the obtained results in Table 1. We compare the use of RGB and thermal images as input data. We note that in general performance are good for both the data domains, revealing that the framework is able to deal with different data types and that the use of an autoencoder combined with the analysis of absolute and gradient difference images is a suitable approach in order to detect anomalies on the railways, as depicted in Figure 3.
Thermal data are probably a better choice than RGB data to achieve the best overall results: indeed, thermal cameras do not depend from external light sources (hence the energy consumption of the system is lower), but usually they are more expensive than the RGB ones and have a limited acquisition framerate and resolution. As shown in Figure 2, anomalies appear more evident, with a better contrast with respect to the railroad: this element could contribute to the better performance of the framework using thermal data. Experiments are conducted on sequences acquired during good weather conditions.

We also test the speed performance of the proposed framework, computing how many frames the architecture can process each second. In order to meet the requirement imposed by the use of a rail drone in terms of energy consumption and computation performance, we run the tests on a PC equipped with an *Intel i7-8700* CPU (3.60GHz, 60W) and a *Nvidia P4000* GPU (100W). The deep networks are implemented in *Pytorch*.
The framework runs in real-time, reaching about 190 frames per second. This result has been obtained carefully designing the two architectures, balancing between the number of layers and total parameters and the computational load of the overall system.

6 Conclusion

In this paper, we propose a deep vision-based framework capable of detecting anomalies (*i.e.* obstacles) on railways that could affect the safety of the train transport. The proposed system combines an autoencoder and a binary classifier in order to label input frames as normal or anomalous.
Experimental results are carried out on a dataset acquired on the railways during the night and confirm the feasibility and the accuracy of the proposed approach. In addition, the proposed system can operate in real-time.
Future work will regard the introduction of the stereo data in the framework and the usage of a GPU-based embedded board equipped with an ARM processor, such as the *Nvidia Jetson TX2*[8]. Moreover, future work will focus on the localization and classification of the detected anomalies as well as on adverse weather conditions that may influence the acquisition process.

Acknowledgements

We thank Ivan Mazzoni (RFI), Marco Plano (RFI) e Mattia Bevere (RFI) for the technical support and accurate annotations.

References

1. Abati, D., Porrello, A., Calderara, S., Cucchiara, R.: Latent space autoregression for novelty detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 481–490 (2019) 3
2. Adam, A., Rivlin, E., Shimshoni, I., Reinitz, D.: Robust real-time unusual event detection using multiple fixed-location monitors. IEEE transactions on pattern analysis and machine intelligence **30**(3), 555–560 (2008) 3
3. Akcay, S., Atapour-Abarghouei, A., Breckon, T.P.: Ganomaly: Semi-supervised anomaly detection via adversarial training. In: Asian Conference on Computer Vision. pp. 622–637. Springer (2018) 2

[8] https://developer.nvidia.com/embedded/jetson-tx2

4. Ballotta, D., Borghi, G., Vezzani, R., Cucchiara, R.: Head detection with depth images in the wild. In: 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. SCITEPRESS (2017) 6
5. Basharat, A., Gritai, A., Shah, M.: Learning object motion patterns for anomaly detection and improved object detection. In: 2008 IEEE Conference on Computer Vision and Pattern Recognition. pp. 1–8. IEEE (2008) 3
6. Borghi, G., Fabbri, M., Vezzani, R., Cucchiara, R., et al.: Face-from-depth for head pose estimation on depth images. IEEE transactions on pattern analysis and machine intelligence (2018) 6
7. Borghi, G., Frigieri, E., Vezzani, R., Cucchiara, R.: Hands on the wheel: a dataset for driver hand detection and tracking. In: 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). pp. 564–570. IEEE (2018) 4
8. Borghi, G., Gasparini, R., Vezzani, R., Cucchiara, R.: Embedded recurrent network for head pose estimation in car. In: 2017 IEEE Intelligent Vehicles Symposium (IV). pp. 1503–1508. IEEE (2017) 4
9. Borghi, G., Pini, S., Vezzani, R., Cucchiara, R.: Driver face verification with depth maps. Sensors **19**(15), 3361 (2019) 7
10. Chan, F.H., Chen, Y.T., Xiang, Y., Sun, M.: Anticipating accidents in dashcam videos. In: Asian Conference on Computer Vision. pp. 136–153. Springer (2016) 3
11. Cong, Y., Yuan, J., Liu, J.: Sparse reconstruction cost for abnormal event detection. In: CVPR 2011. pp. 3449–3456. IEEE (2011) 3
12. Espino, J.C., Stanciulescu, B.: Rail extraction technique using gradient information and a priori shape model. In: 2012 15th International IEEE Conference on Intelligent Transportation Systems. pp. 1132–1136. IEEE (2012) 3
13. Fabbri, M., Borghi, G., Lanzi, F., Vezzani, R., Calderara, S., Cucchiara, R.: Domain translation with conditional gans: from depth to rgb face-to-face. In: 2018 24th International Conference on Pattern Recognition (ICPR). pp. 1355–1360. IEEE (2018) 3
14. Filev, D.P., Chinnam, R.B., Tseng, F., Baruah, P.: An industrial strength novelty detection framework for autonomous equipment monitoring and diagnostics. IEEE Transactions on Industrial Informatics **6**(4), 767–779 (2010) 2
15. García, J.J., Losada, C., Espinosa, F., Ureña, J., Hernández, Á., Mazo, M., de Marziani, C., Jiménez, A., Bueno, E., Álvarez, F.: Dedicated smart ir barrier for obstacle detection in railways. In: 31st Annual Conference of IEEE Industrial Electronics Society, 2005. IECON 2005. pp. 6–pp. IEEE (2005) 3
16. Görnitz, N., Kloft, M., Rieck, K., Brefeld, U.: Toward supervised anomaly detection. Journal of Artificial Intelligence Research **46**, 235–262 (2013) 2
17. Gschwandtner, M., Pree, W., Uhl, A.: Track detection for autonomous trains. In: International Symposium on Visual Computing. pp. 19–28. Springer (2010) 3
18. Hasan, M., Choi, J., Neumann, J., Roy-Chowdhury, A.K., Davis, L.S.: Learning temporal regularity in video sequences. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 733–742 (2016) 2, 3
19. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014) 7
20. Kumar, A.: Computer-vision-based fabric defect detection: A survey. IEEE transactions on industrial electronics **55**(1), 348–363 (2008) 2
21. Liu, W., Luo, W., Lian, D., Gao, S.: Future frame prediction for anomaly detection–a new baseline. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 6536–6545 (2018) 3

22. Maire, F.: Vision based anti-collision system for rail track maintenance vehicles. In: 2007 IEEE Conference on Advanced Video and Signal Based Surveillance. pp. 170–175. IEEE (2007) 3
23. Maire, F., Bigdeli, A.: Obstacle-free range determination for rail track maintenance vehicles. In: 2010 11th International Conference on Control Automation Robotics & Vision. pp. 2172–2178. IEEE (2010) 3
24. Mockel, S., Scherer, F., Schuster, P.F.: Multi-sensor obstacle detection on railway tracks. In: IEEE IV2003 Intelligent Vehicles Symposium. Proceedings (Cat. No. 03TH8683). pp. 42–46. IEEE (2003) 3
25. Palazzi, A., Abati, D., Solera, F., Cucchiara, R., et al.: Predicting the driver's focus of attention: the dr (eye) ve project. IEEE transactions on pattern analysis and machine intelligence **41**(7), 1720–1733 (2018) 3
26. Passarella, R., Tutuko, B., Prasetyo, A.P.: Design concept of train obstacle detection system in indonesia. IJRRAS **9**(3), 453–460 (2011) 3
27. Pathak, D., Agrawal, P., Efros, A.A., Darrell, T.: Curiosity-driven exploration by self-supervised prediction. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 16–17 (2017) 2
28. Prewitt, J.M.: Object enhancement and extraction. Picture processing and Psychopictorics **10**(1), 15–19 (1970) 7
29. Punekar, N.S., Raut, A.A.: Improving railway safety with obstacle detection and tracking system using gps-gsm model. International Journal of Scientific & Engineering Research **4**(8), 282–288 (2013) 3
30. Sabokrou, M., Khalooei, M., Fathy, M., Adeli, E.: Adversarially learned one-class classifier for novelty detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3379–3388 (2018) 2, 3
31. Schlegl, T., Seeböck, P., Waldstein, S.M., Schmidt-Erfurth, U., Langs, G.: Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In: International conference on information processing in medical imaging. pp. 146–157. Springer (2017) 2
32. Xu, B., Wang, N., Chen, T., Li, M.: Empirical evaluation of rectified activations in convolutional network. arXiv preprint arXiv:1505.00853 (2015) 6, 8
33. Yao, Y., Xu, M., Wang, Y., Crandall, D.J., Atkins, E.M.: Unsupervised traffic accident detection in first-person videos. arXiv preprint arXiv:1903.00618 (2019) 2, 3
34. Zendel, O., Murschitz, M., Zeilinger, M., Steininger, D., Abbasi, S., Beleznai, C.: Railsem19: A dataset for semantic rail scene understanding. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 0–0 (2019) 3
35. Zhao, B., Fei-Fei, L., Xing, E.P.: Online detection of unusual events in videos via dynamic sparse coding. In: CVPR 2011. pp. 3313–3320. IEEE (2011) 3
36. Zwemer, M.H., van de Wouw, D.W., Jaspers, E.G., Zinger, S., et al.: A vision-based approach for tramway rail extraction. In: Video Surveillance and Transportation Imaging Applications 2015. vol. 9407, p. 94070R. International Society for Optics and Photonics (2015) 3