Unsupervised anomaly detection in railway catenary condition monitoring using autoencoders

Wang, Hongrui

DOI
10.1109/IECON43393.2020.9254633

Publication date
2020

Document Version
Final published version

Published in
Proceedings - IECON 2020

Unsupervised anomaly detection in railway catenary condition monitoring using autoencoders

1st Hongrui Wang
Department of Engineering Structures
Delft University of Technology
Delft, The Netherlands
h.wang-8@tudelft.nl

***Abstract*—The condition monitoring of railway infrastructures is collecting big data for intelligent asset management. Making the most of the big data is a critical challenge facing the railway industry. This study focuses on one of the main railway infrastructures, namely the catenary (overhead line) system that transmits power to trains. To facilitate the effective usage of catenary condition monitoring data, this study proposes an unsupervised anomaly detection approach as a pre-processing measure. The approach trains autoencoders to reduce the dimensionality of multisensor data and generate discriminative features between healthy and anomalous data. By testing the reconstruction errors using the trained autoencoders, anomalous data that indicate potential defects of catenary can be identified without prior information and human intervention. A case study on a section of high-speed railway catenary in China shows that the approach can automatically distinguish between healthy and anomalous data. The output anomalous data can save a considerable amount of computation time and manpower in further interpretations aiming to pinpoint defects.**

***Index Terms*—anomaly detection, railway catenary, condition monitoring, unsupervised learning, autoencoders**

I. Introduction

Condition monitoring is widely deployed in worldwide railways to enable better maintenance decision making, which aims to prolong the life cycle of railway assets and reduce the probability of asset failures that disrupt train services. With the development of sensor technology, nowadays data are flowing in on a daily basis from a variety of sensors adopted to monitor railway assets in a comprehensive manner. It brings challenges as well as opportunities in using the railway big data to its full potential [1].

For railway infrastructures commonly distributed crossing lands and continents, there are mainly three ways to perform condition monitoring depending on the type and placement of sensors. First, sensors installed on a specialised inspection train or commercial train can measure the dynamic responses of on-board components and take photos or videos of surrounding infrastructures while the train is running [2], [3]. In this way, an entire railway network can be periodically inspected every year, half a year or lesser. Second, sensors installed on or near by infrastructures can monitor vibrations, deformations, appearances and temperatures of critical components at certain locations [4]. Data can be collected by track-side servers through cables or wireless transmission. Last, remote sensing techniques using satellites [5], drones [6], etc. are recently gaining attentions, which brings conveniences in monitoring railway infrastructures at a large scale from a distance. Because a railway infrastructure normally consist of tens or hundreds of component types, it may fail in providing its functionality in many ways. It is thus necessary and common that multiple sensors are adopted to monitor an infrastructure from different aspects in practice.

Until now, railway catenary (overhead line) systems, the most popular infrastructure for train power supply, are mainly monitored and inspected using sensors installed on board. To provide continuous and steady electric power to running trains, a pantograph mounted on the train roof should be able to collect electric current from the catenary through the pantograph-catenary sliding contact. This contact is essentially maintained by the contact force between the pantograph and the catenary, which is subject to fluctuation due to contact-induced vibrations. To monitor the quality of pantograph-catenary contact, sensors are often installed on the pantograph or the train roof. These sensors measure a variety of physical parameters or phenomenons that can reflect the contact quality, including the pantograph-catenary contact force [7], pantograph head acceleration [8], contact point or contact wire displacement [9], [10], electric arcing effect [11] and so on. In addition, components and foreign bodies of catenary that do not directly interact with the pantograph, including insulators, steady arms, cantilevers, fasteners, etc. [12], are inspected by cameras or scanners taking photos [13], videos [14] or point cloud data [15]. All data collected are commonly synchronized with auxiliary measurements such as train speed, GPS location, radio-frequency identification or similar, so that the measurement location and condition are known for further data analysis and on-site verification.

Transforming the collected multisensor data into qualitative or quantitative conditions of catenary is a challenging task. In practice and literature, condition assessment of catenary is mostly carried out using one type of data. Data processing techniques including Fourier transform, wavelet transform [16], empirical mode decomposition [17] and quadratic time-frequency distribution [18] have been adopted for feature extraction. Due to recent advancements in artificial intelligence, machine learning methods such as Bayesian network [19] and deep convolutional neural network [20] have also been applied to catenary condition assessment. However, existing studies

978-1-7281-5414-5/20/$31.00 ©2020 IEEE

and applications rely heavily on manual data processing to identify and label unhealthy conditions for feature extraction and training, while a majority of measurement data actually represents a healthy condition because inspection trains collect data by running through an entire railway line or network. When dealing with multisensor data, identifying anomalies becomes even more difficult because each data type reflects the catenary condition in a particular way. Combining the information contained in these data to make a correct judgment is still an ongoing research question.

To facilitate the intelligent data processing and decision making in catenary condition monitoring, this study proposes an unsupervised approach to detect anomalies reflected by the multisensor monitoring data. The approach adopts an unsupervised learning method, namely autoencoders [21] to distinguish between healthy and anomalous data samples without giving prior knowledge and interpretation of the original data. The rest of this paper is organized as follows. Section II introduces the theory of autoencoder neural networks. Section III presents the unsupervised anomaly detection approach using autoencoders. Section IV presents and discusses results applying the approach to a set of catenary condition monitoring data measured at a high-speed railway line in China.

II. Autoencoders

An autoencoder is a feed-forward neural network that can automatically learn features from unlabelled data. Figure 1 shows an autoencoder consisting of an input layer, a hidden layer and an output layer. The autoencoder first takes the input training set $\mathbf{x} = \{x_1, x_2, \cdots, x_n\}$ and transforms it into a hidden representation $\mathbf{h} = \{h_1, x_2, \cdots, h_m\}$ through a nonlinear activation function

$$\mathbf{h} = \varphi\left(\mathbf{W}\mathbf{x} + \mathbf{b}\right). \tag{1}$$

This phase is also called an encoder. Then, the hidden representation is mapped to an output representation $\mathbf{y} = \{y_1, y_2, \cdots, y_n\}$ similar to the input

$$\mathbf{y} = \varphi\left(\mathbf{W}'\mathbf{x} + \mathbf{b}'\right). \tag{2}$$

This phase is considered as a decoder. Through encoding and decoding, the autoencoder tries to learn model parameters $\theta = [\mathbf{W}, \mathbf{b}, \mathbf{W}', \mathbf{b}']$ that reconstruct the input so that the output $\mathbf{y} = f_\theta(\mathbf{x}) \approx \mathbf{x}$. Thus, the model parameters can be optimized by minimizing the reconstruction error between the output $\mathbf{y}$ and the input $\mathbf{x}$. This optimization problem can be formulated using the mean square error (MSE), also known as a cost function, as follows:

$$\min_\theta \frac{1}{N}\sum_i^N \left(\mathbf{x}_i - f_\theta\left(\mathbf{x}_i\right)\right)^2 \tag{3}$$

where $\mathbf{x}_i$ denotes the ith data sample and N is total number of samples.

If the training of autoencoders is performed by minimizing the MSE only, the learned transformation can be the identity one, which prevents the discovery of meaningful data

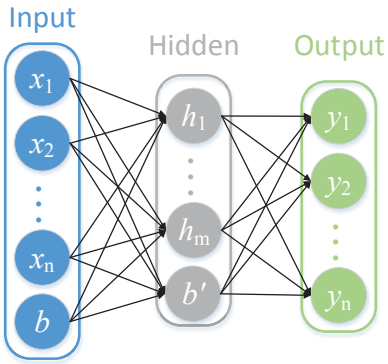

Fig. 1. An autoencoder neural network.

structures. This issue can be handled by imposing a sparsity constraint on the hidden units [21]. The constraint penalizes excessive activations of hidden units so that the hidden units are only sensitive to a certain type of training samples. It updates the cost function to

$$\min_\theta \frac{1}{N}\sum_i^N \left(\mathbf{x}_i - f_\theta\left(\mathbf{x}_i\right)\right)^2 + \beta\sum_j^m KL(p||p_j) \tag{4}$$

where the added term is the sum of the Kullback-Leibler (KL) divergence over the m hidden units and β is a coefficient for this sparsity penalty term. The KL-divergence measures the difference between two probability distributions. For the jth hidden unit,

$$KL(p||p_j) = p\log\left(\frac{p}{p_j}\right) + (1-p)\log\left(\frac{1-p}{1-p_j}\right) \tag{5}$$

where p_j is the average activation of the jth hidden unit over the training data, which is penalized if deviating from p, and p is a sparsity parameter typically with a value close to zero.

The hidden representation is an encoded and abstract representation of the input data $\mathbf{x}$. When the dimensionality of the hidden representation (the number of hidden units) m is smaller than that of the input n, the input data is compressed into a lower dimension. This compression is particularly meaningful when the input data vectors have certain correlations with each other. The autoencoder discovers and generates discriminative features in the hidden representation that can clearly separate potential clusters in the input data [22]. A set of condition monitoring data, which contains physically correlated data measured under the same system dynamics, is suitable for such compression.

III. An Unsupervised Anomaly Detection Approach

A. Data Correlation

Condition monitoring data of catenary measured by onboard sensors can be considered as the results of the pantograph-catenary interaction. Figure 2 shows a typical pantograph-catenary contact above a train roof. When a pantograph lifts up on the train roof, it gives an initial contact

force that presses the pantograph against the catenary. Electric current can thus flow from the catenary through the pantograph to the train locomotive. When the train starts running, the pantograph head slides through the contact wire of catenary with the contact force changing due to the variation of contact wire stiffness along the running direction. The change of contact force leads to vibrations of pantograph and catenary as the contact point is constantly moving. If this interaction cannot be maintained by a contact force regulated within a certain range, the pantograph will not be able collect electric current from the catenary during train operations, which should certainly be avoided as much as possible in practice. Therefore, physical parameters that reflect the status of pantograph-catenary interaction are monitored during train operations to evaluate the catenary performance. In this study, we consider five physical parameters that are measured and related to the pantograph-catenary interaction, as listed in Table I.

The five parameters are physically correlated with each other. These correlations have been elaborated in [19], which are summarized in short as follows:

1) The train speed is proportional to the average amplitude of all other parameters measured, because in general higher speeds cause stronger vibrations.
2) The inertia force F_i, as a part of the measured pantograph-catenary contact force, yields Newton's second law of motion that is

$$F_i = m_i \cdot a$$

where a is the pantograph head acceleration and m_i is the mass of pantograph head.
3) The pantograph head acceleration is theoretically the second derivative of contact wire dynamic height with respect to time, assuming the pantograph head is always in contact with the contact wire.
4) The contact wire dynamic stagger is proportional to the contact force, because a larger contact force results in a larger friction force in the lateral direction.

Therefore, it can be concluded that, by nature, the five types of parameter are strongly correlated. The fact that these parameters are measured by separate sensors makes it more likely and meaningful to discover a structure among the measured parameters using dimensionality reduction by autoencoders. This structure enables clear clustering of the original measurement data, which facilitates detecting anomalies in the measurement data.

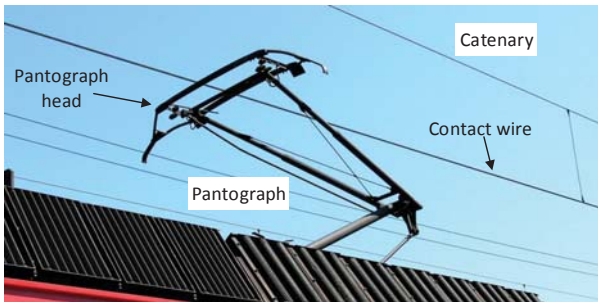

Fig. 2. Pantograph-catenary contact.

TABLE I
PHYSICAL PARAMETERS CONSIDERED IN THIS STUDY.

#	Parameter	Unit
1	Pantograph-catenary contact force (PCCF)	N
2	Pantograph head acceleration (PHA)	m/s^2
3	Contact wire dynamic height (vertical displacement) (CDH)	mm
4	Contact wire dynamic stagger (lateral displacement) (CDS)	mm
5	Train speed	km/h

B. Anomaly Detection

One main characteristic of catenary condition monitoring data is that a considerable amount of data samples represents a healthy catenary condition. From experiences and literature, most catenary defects, with a length from millimetres to meters [16], are locally and sporadically distributed along kilometres of catenary. In measurement data, these defects can only give rise to several anomalous data samples out of the entire data pool, depending on the sampling interval. When reducing the dimensionality of this type of data with strong correlations, the significant difference between the amount of healthy and anomalous data makes it even easier to generate discriminative features for identifying anomalies.

For condition monitoring data collected from a specific railway line, the proposed anomaly detection approach is described as follows:

Step 1: Train an autoencoder using historical data samples measured at the same railway line, with the data dimension n and the hidden representation dimension $m < n$ ($m \in \mathbf{N}^+$).
Step 2: Input a new set of measurement data samples $\mathbf{x}_{new}$, which should be measured at the same railway line as the training data, into the trained autoencoder and obtain a reconstructed version of the input data $\mathbf{x}'_{new}$.
Step 3: Calculate the MSE between the input samples and the reconstructed ones.
Step 4: Identify the data samples with reconstruction errors ranking above the pth percentile among all samples as anomalous data.

In this study, the input data for training and testing are the aforementioned five data types so that the number of input units $n = 5$. Thus, the number of hidden units m is selected to be 2 considering the aim is to classify data into only two clusters (healthy and anomalous ones) and the strong correlations among the five types of input data. Later, results when $m = 3$ will also be compared with those using $m = 2$. It is of course possible and advised to include more input data types reflecting the pantograph-catenary interaction into the autoencoder, by simply increasing the number of input units n. The selection of the percentile threshold p in Step 4 highly depends on the current condition of the targeted catenary. A combined factor concerning the age and the traffic density of the catenary should be taken into account. In general, an older

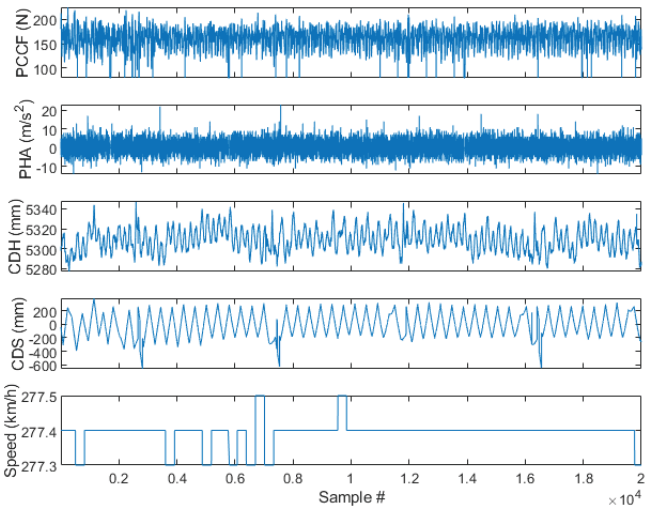

Fig. 3. A segment of training data.

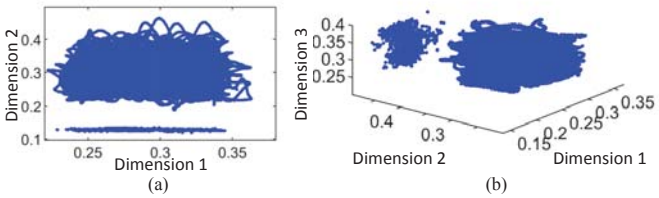

Fig. 4. The learned (a) 2D and (b) 3D representations of training data.

catenary structure with a higher traffic density has a smaller threshold p, because more anomalies are expected to be found. In addition, p can be slightly larger if the applied maintenance frequency or standard is higher than average. Empirically, p can range from 95 to 99.5.

IV. RESULTS AND DISCUSSIONS

A. Training

A set of historical data measured at a section of catenary in the Beijing-Guangzhou high-speed railway line is adopted for training and testing in this work. It contains data from more than 20 inspection runs on the same section between 2015 and 2018. Data from the latest inspection run is used as the test data, while the rest are used for autoencoder training. As a result, a total number of over 3 million data samples trains an autoencoder with the number of hidden units $m = 2$ and $m = 3$, respectively. Figure 3 shows a part of the training data set containing 20,000 samples. The five types of data are simultaneously measured and synchronized with a fixed sampling interval of 0.25m. All data types are rescaled through mean normalization to have the same range of values.

It takes 510 and 390 epochs for the training of autoencoders with two and three hidden units to converge, respectively. Both autoencoders find a minimum MSE of 0.0116 over all training samples. Figure 4(a) and 4(b) show the 2D and 3D representations learned by training the autoencoders, respectively. The dots are features generated by a piece of the training dataset containing more than 100 thousand samples. In both abstract representations, two clusters can be observed with one gathering a majority of samples and the other with much less samples. In the 2D representation, the big cluster is located above the small one. In the 3D representation, there is a small gap between the oval-shaped big cluster and the small one located closer to the vertical axis. Both representations demonstrate that a certain data structure has been learned by the autoencoders. Although it is still unclear what is the physical meaning of these clusters, the learned features are quite discriminative in its own right.

B. Result Discussions

The trained autoencoders are tested on a new set of inspection data measured from the same section of railway line, which consists of 12,500 data samples. Following the anomaly detection approach presented in the previous section, anomalous data are identified by ranking and comparing the reconstruction errors among data samples. The catenary evaluated in this case study is considered as well-maintained with a high maintenance standard applied. Thus, a percentile threshold $p = 99$ is chosen for reconstruction errors generated by the trained autoencoders. Some interesting clusters including imperfect ones are discussed in the following examples.

Figure 5 shows the anomaly detection results on a segment of the tested data. Figure 5(a) and Figure 5(c) show that there are differences between the anomalies identified by the autoencoder with two and three hidden units, which learns 2D and 3D features, respectively. From the 2D and 3D representations shown in Figure 5(b) and Figure 5(d), it can be seen that the 2D features are clearly clustered with anomalies separated perfectly from the healthy samples, whereas the 3D features do not represent a clear structure of the input data with anomalies mixed with healthy data. For the rest of the tested data, this discrepancy between 2D and 3D representations are consistent, indicating that an autoencoder with two hidden units are a more suitable choice in this case. In other words, the given data set is better learned by compressing its dimensionality to two. This choice only applies to the presented case and can be larger if the number of input data types increases.

Figure 6 shows the normalized input data corresponding to the detection results shown in Figure 5(a). It can be seen that this segment of data was measured under a constant train speed. The anomalies identified by the proposed approach are data samples with either very small PCCF or very small PHA. These samples indicate abnormal dynamic responses in the pantograph-catenary interaction, which potentially caused undesired electric arcing. It demonstrates that the autoencoder with $m = 2$ learns that the PCCF and PHA are the main physical parameters reflecting the catenary condition among the five types of input parameters. This is reasonable because the PCCF and PHA theoretically contain richer information about the dynamic responses of the interaction. It is important to validate these samples on site if given the opportunity, so

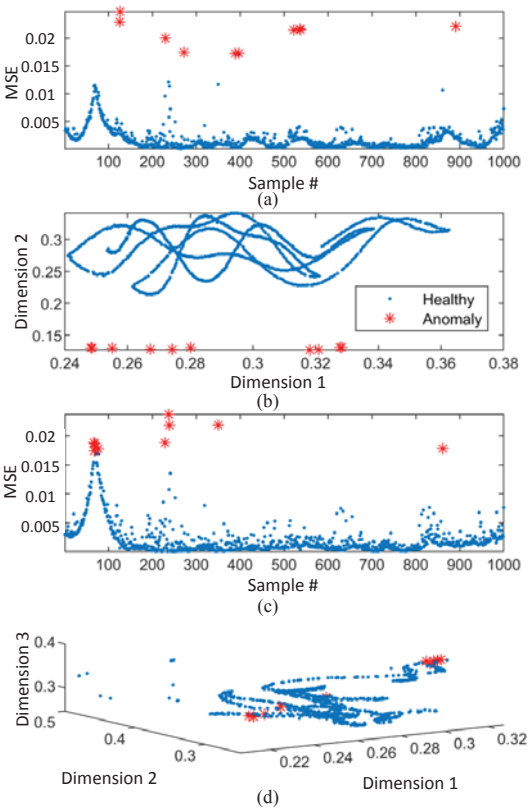

Fig. 5. An example of anomaly detection results and the corresponding clusters in 2D and 3D representations. (a) MSEs generated by the autoencoder with $m = 2$ and (b) the 2D features. (c) MSEs generated by the autoencoder with $m = 3$ and (d) the 3D features.

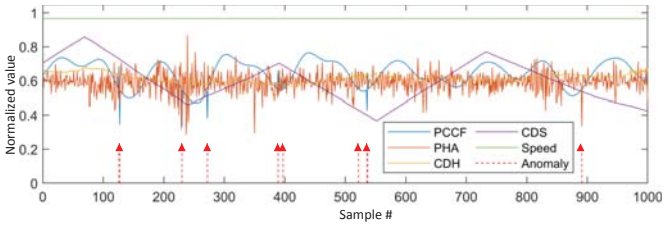

Fig. 6. Normalized input data with the detected anomalies indicated by dashed lines.

that the samples can be associated with a specific type of defect or condition, which is valuable for pinpointing defects in further research.

In comparison with one of the most popular traditional methods for dimensionality reduction, the principal component analysis (PCA), Figure 7 shows the MSEs of data reconstructed by PCA and the autoencoder. It can be seen that the level of MSEs is similar between the two methods. Although there are similar outliers with high MSEs around the sample 250, most of the anomalies identified by the autoencoder MSEs cannot be reflected by those of PCA. This indicates that the 2D feature space generated by the autoencoder can better represent the underlying structure among the data types

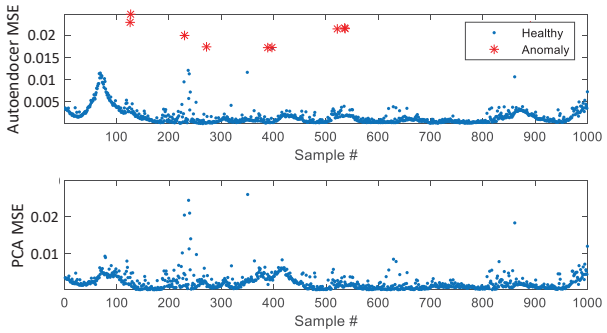

Fig. 7. Comparisons of MSEs resulting from the autoencoder (top) and PCA (bottom).

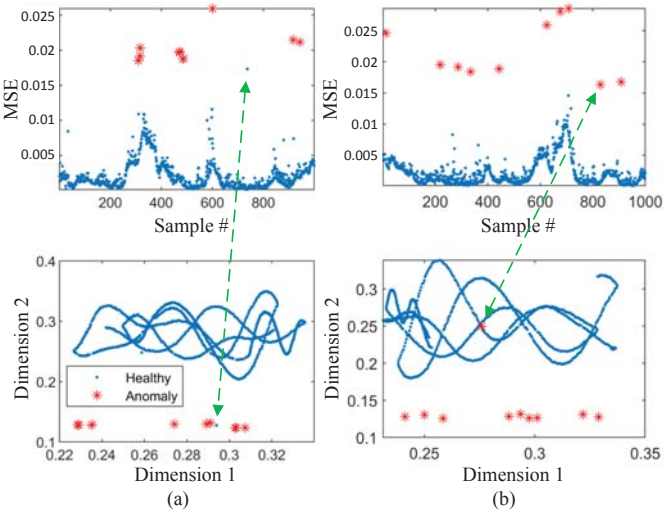

Fig. 8. Two examples of anomaly detection results (top) with mixed clusters in the 2D representations (bottom). (a) MSEs and 2D features with one sample falsely identified as healthy. (b) MSEs and 2D features with one sample falsely identified as anomaly. Both false samples are indicated by dashed lines with an arrow at both ends.

for catenary condition monitoring compared with PCA.

Figure 8 shows two examples with imperfect clusters in the 2D representation. In the first example shown in Figure 8(a), the dashed line indicates a data sample that should belong to anomalies in the 2D representation, but still being identified as healthy. This error is presumably caused by the fixed percentile threshold defined in the proposed approach. It can be improved by using an adaptive threshold or clustering directly in the 2D representation. On the contrary, Figure 8(b) highlights a data sample identified as an anomaly while appearing to be mixed with the 'healthy' cluster in the 2D representation. The causes of this error are likely related to the autoencoder learning. The autoencoder structure adopted in this study is a simple one with a single hidden layer. This structure can still be optimized to improve the performance of anomaly detection considering variants such as stacked autoencoders [23], denoising autoencoders [24] and variational autoencoders [25].

V. Conclusion

This study proposes an unsupervised anomaly detection approach for one of the main railway infrastructures, namely the catenary systems. The multisensor data from catenary condition monitoring are adopted as the input for the unsupervised approach. The approach adopts autoencoder neural networks with a single hidden layer to learn the structure hidden among the catenary data set with strong physical correlations. The anomalies are identified by percentile-based thresholding on the reconstruction errors generated by the trained autoencoders. Preliminary results on a historical data set measured at a section of catenary in the Beijing-Guangzhou high-speed line in China shows that, the multisensor data of catenary is suitable to be compressed into 2D representations by the autoencoder for the purpose of anomaly detection. The 2D representation generates discriminative features between healthy and anomalous data, which are learned by the autoencoder in an unsupervised manner. Real-life examples with detection results and comparisons with PCA are presented and discussed, which inspire the following potential improvements for future research:

1) Identifying anomalies by clustering features learned in the abstract representations instead of comparing the reconstruction errors.
2) Optimizing the autoencoder structure considering variants such as stacked autoencoders, de-noising autoencoders and variational autoencoders.
3) Including additional condition monitoring data sources of catenary such as images, 3D point cloud data or even the same data types measured by extra on-board sensors.
4) Extending from anomaly detection to defect identification by using more complex features learned by deeper autoencoders and potentially combining with supervised classifiers.

Acknowledgment

The author would like to thank Prof. Zhigang Liu with Southwest Jiaotong University and the Chinese Academy of Railway Sciences for sharing the data used in the case study of this research.

References

[1] F. Ghofrani, Q. He, R. M. P. Goverde, and X. Liu, Recent applications of big data analytics in railway transportation systems: A survey, Transportation Research Part C: Emerging Technologies, vol. 90, pp. 226-246, 2018.

[2] M. Molodova, Z. Li, A. Núñez, and R. Dollevoet, Automatic detection of squats in railway infrastructure, IEEE Transactions on Intelligent Transportation Systems, vol. 15, no. 5, pp. 1980-1990, 2014.

[3] J. Chen, Z. Liu, H. Wang, A. Núñez , and Z. Han, Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network, IEEE Transactions on Instrumentation and Measurement, vol. 67, no. 2, pp. 257-269, 2017.

[4] V. J. Hodge, S. O'Keefe, M. Weeks, and A. Moulds, Wireless sensor networks for condition monitoring in the railway industry: A survey, IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 3, pp. 1088-1106, 2014.

[5] L. Chang, R. P. B. J. Dollevoet, and R. F. Hanssen, Railway infrastructure monitoring using satellite radar data. Int. J. Railw. Technol, vol. 3, pp. 79-91, 2014.

[6] F. Flammini, C. Pragliola, and G. Smarra, Railway infrastructure monitoring by drones, in 2016 International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles & International Transportation Electrification Conference (ESARS-ITEC), pp. 1-6.

[7] S. Kusumi, T. Fukutani, and K. Nezu, Diagnosis of overhead contact line based on contact force, Quarterly Report of RTRI, vol. 47, no. 1, pp. 39-45, 2006.

[8] M. Carnevale and A. Collina, Processing of collector acceleration data for condition-based monitoring of overhead lines, Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, vol. 230, no. 2, pp. 472-485, 2016.

[9] C. J. Cho and H. Ko, Video-based dynamic stagger measurement of railway overhead power lines using rotation-invariant feature matching, IEEE Transactions on Intelligent Transportation Systems, vo. 16, no. 3, pp. 1294-1304, 2014.

[10] I. Aydin, M. Karakose, and E. Akin, Anomaly detection using a modified kernel-based tracking in the pantograph–catenary system, Expert Systems with Applications, vol. 42, no. 2, pp. 938-948, 2015.

[11] I. Aydin, M. Karaköse, and E. Akin, A robust anomaly detection in pantograph-catenary system based on mean-shift tracking and foreground detection, in 2013 IEEE international conference on systems, man, and cybernetics, pp. 4444-4449.

[12] W. Liu, Z. Liu, A. Núñez, and Z. Han, Unified deep learning architecture for the detection of all catenary support components. IEEE Access, vol. 8, pp. 17049-17059.

[13] J. Chen, Z. Liu, H. Wang, A. Núñez, and Z. Han, Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network, IEEE Transactions on Instrumentation and Measurement, vol. 67, no. 2, pp. 257-269, 2017.

[14] X. Wu, P. Yuan, Q. Peng, C. W. Ngo, and J. Y. He, Detection of bird nests in overhead catenary system images for high-speed rail, Pattern Recognition, vol. 51, pp. 242-254, 2016.

[15] Z. Han, C. Yang, and Z. Liu. Cantilever Structure Segmentation and Parameters Detection Based on Concavity and Convexity of 3D Point Clouds, IEEE Transactions on Instrumentation and Measurement, in press, 2019.

[16] H. Wang, Z. Liu , A. Núñez, and R. Dollevoet, Entropy-based local irregularity detection for high-speed railway catenaries with frequent inspections. IEEE Transactions on Instrumentation and Measurement, 68(10), 3536-3547, 2018.

[17] Z. Liu, H. Wang, R. Dollevoet, Y. Song, A. Núñez, and J. Zhang, Ensemble EMD-based automatic extraction of the catenary structure wavelength from the pantograph–catenary contact force. IEEE Transactions on Instrumentation and Measurement, vol. 65, no. 10, pp. 2272-2283, 2016.

[18] H. Wang, Z. Liu, Y. Song, X. Lu, Z. Han, J. Zhang, and Y. Wang, Detection of contact wire irregularities using a quadratic time–frequency representation of the pantograph–catenary contact force, IEEE Transactions on Instrumentation and Measurement, vol. 65, no. 6, pp. 1385-1397, 2016.

[19] H. Wang, A. Núñez, Z. Liu, D. Zhang, and R. Dollevoet, A Bayesian Network Approach for Condition Monitoring of High-Speed Railway Catenaries. IEEE Transactions on Intelligent Transportation Systems, in press, 2019.

[20] W. Liu, Z. Liu, A. Núñez, L. Wang, K. Liu, Y. Lyu, and H. Wang, Multi-objective performance evaluation of the detection of catenary support components using DCNNs, IFAC-PapersOnLine, vol. 51, no. 9, pp. 98-105, 2018.

[21] A. Ng, Sparse autoencoder, CS294A Lecture notes, vol. 72, no. 1-19, 2011.

[22] I. Goodfellow, Y. Bengio, and A. Courville, Deep learning, MIT press, 2016.

[23] H. I. Suk, S. W. Lee, D. Shen, and Alzheimer's Disease Neuroimaging Initiative, Latent feature representation with stacked auto-encoder for AD/MCI diagnosis, Brain Structure and Function, vol. 220 no. 2, pp. 841-859, 2015.

[24] C. Lu, Z. Y. Wang, W. L. Qin, and J. Ma, Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification, Signal Processing, vol. 130, pp. 377-388, 2017.

[25] Y. Pu, Z. Gan, R. Henao, X. Yuan, C. Li, A. Stevens, and L. Carin, Variational autoencoder for deep learning of images, labels and captions, Advances in neural information processing systems, pp. 2352-2360, 2016.