ESANN 2019 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2019, i6doc.com publ., ISBN 978-287-587-065-0. Available from http://www.i6doc.com/en/.

Deep Autoencoder Feature Extraction for Fault Detection of Elevator Systems

Krishna Mohan Mishra, Tomi R. Krogerus, and Kalevi J. Huhtala

Tampere University of Technology, Tampere, Finland

Abstract. In this research, we propose a generic deep autoencoder model for automated feature extraction from the elevator sensor data. Extracted deep features are classified with random forest algorithm for fault detection. Sensor data are labelled as healthy or faulty based on the maintenance actions recorded. In our research, we have included all fault types present for each elevator. The remaining healthy data is used for validation of the model to prove its efficacy in terms of avoiding false positives. We have achieved nearly 100% accuracy in fault detection along with avoiding false positives based on new extracted deep features, which outperform the results using existing features.

1 Introduction

In recent years, elevator systems have been used more and more extensively in apartments, commercial facilities and office buildings. Nowadays 54% of the world's population is living in urban areas [1]. Elevators transport 325 million passengers every day in the United States and Canada alone [2]. Therefore, elevator systems need proper maintenance and safety. The next step for improving the safety of elevator systems is the development of predictive and pre-emptive maintenance strategies, which will also reduce repair costs and increase the lifetime whilst maximizing the uptime. Modern elevator systems require intelligent fault monitoring and diagnosis.

Fault diagnosis methods based on deep neural networks [3] and convolutional neural networks [4] feature extraction methodology are presented as state of the art for rotatory machines similar to elevator systems. However, we have developed an intelligent deep autoencoder based feature extraction methodology for fault detection in elevator systems to improve the performance of traditional fault diagnosis methods.

In the last decade, neural networks [5] have extracted highly meaningful statistical patterns from large-scale and high-dimensional datasets. A deep learning network can self-learn the relevant features from multiple signals [6]. Autoencoding is a process for nonlinear dimension reduction with natural transformation architecture using feedforward neural network [7]. Autoencoders can increase the generalization ability of machine learning models by extracting features of high interest as well as making possible its application to sensor data [8]. Autoencoders were first introduced by LeCun [9], and have been studied for decades. Traditionally, feature learning and dimensionality reduction are the two main features of autoencoders. Recently, autoencoders have been considered as one of the most compelling subspace analysis techniques because of the existing theoretical relations between autoencoders and latent variable models [10]. Autoencoders have been used for feature extraction from the data in systems like localization [11] and wind turbines [12], different from elevator systems as in our research.

ESANN 2019 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2019, i6doc.com publ., ISBN 978-287-587-065-0. Available from http://www.i6doc.com/en/.

In our previous research, raw sensor data, mainly acceleration signals, were used to calculate elevator key performance and ride quality features, which we call here existing features. Random forest was used for fault detection based on these existing features. Existing domain specific features were calculated from raw sensor data, but that requires expert knowledge of the domain and results in a loss of information to some extent. To avoid these implications, we have developed a deep autoencoder random forest approach for automated feature extraction from elevator sensor data, and based on these deep features, faults are detected. In this paper, section 2 presents the methodology, section 3 includes the results and discussion, and section 4 provides the conclusions of our research.

2 Methodology

We have developed an automatic feature extraction technique in this research as an extension to the work of our previous research to compare the results using new extracted deep features. Figure 1(a) shows the fault detection approach used in this paper, which includes elevator sensor data extracted based on time periods provided by the maintenance data. We have analysed almost one year of the data from seven traction elevators in this research. Each elevator produces around 200 rides per day. Every movement of the elevator generates existing features from the vibration signal. Data collected from an elevator system is fed to a deep autoencoder model for new feature extraction and then random forest performs the fault detection task based on extracted deep features.

The deep autoencoder model is based on deep learning autoencoder feature extraction methodology. A basic autoencoder is a fully connected three-layer feedforward neural network with one hidden layer. Typically, the autoencoder has the same number of neurons in the input and the output layer and reproduces the input as its output. We use a five layer deep autoencoder (see Fig. 1(b)) including input, output, encoder, decoder and representation layers.

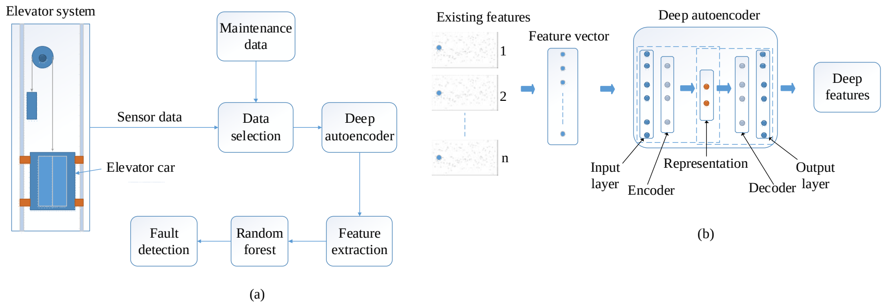


Fig. 1: (a) Fault detection approach (b) Deep autoencoder feature extraction approach.

In our approach, we first feed the elevator sensor data from each elevator movement in up and down directions separately in the deep autoencoder model to extract new deep features from the data. Then we apply random forest as a classifier for fault detection based on new deep features extracted from the data. The encoder

ESANN 2019 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2019, i6doc.com publ., ISBN 978-287-587-065-0. Available from http://www.i6doc.com/en/.

transforms the input x into corrupted input data x' using hidden representation h through nonlinear mapping

$$h=f(W_1x'+b) \tag{1}$$

where $f(\cdot)$ is a nonlinear activation function as the sigmoid function, $W_1 \in \mathbb{R}^{k*m}$ is the weight matrix and $b \in \mathbb{R}^k$ the bias vector to be optimized in encoding with k nodes in the hidden layer [13]. Then, with parameters $W_2 \in \mathbb{R}^{m*k}$ and $c \in \mathbb{R}^m$, the decoder uses nonlinear transformation to map hidden representation h to a reconstructed vector x'' at the output layer

$$x''=g(W_2h+c) \tag{2}$$

where $g(\cdot)$ is again nonlinear function (sigmoid function). In this study, the weight matrix is $W_2=W_1^T$, which is tied weights for better learning performance [14].

Random forest includes an additional layer of randomness to bagging. It uses different bootstrap samples of the data for constructing each tree [15]. The best subset of predictors is used to split each node in random forest. This counterintuitive strategy is the best feature of random forest, which makes it different from other classifiers as well as robust against overfitting. It is one of the most user-friendly classifiers because it consists of only two parameters: the number of variables and number of trees. However, it is not usually very sensitive to their values [16]. The final classification accuracy of random forest is calculated by averaging, i.e. arithmetic mean the probabilities of assigning classes related to all the produced trees. Testing data that is unknown to all the decision trees is used for evaluation by the voting method. Specifically, let sensor data value v_l^t have training sample l^{th} in the arrived leaf node of the decision tree $t \in T$, where $l \in \{1,...,L_t\}$ and the number of training samples is L_t in the current arrived leaf node of decision tree t . The final prediction result is given by [17]:

$$\mu=(\textstyle\sum_{t\in T}\sum_{l\in\{1,...,L_t\}} v_l^t)/(\sum_{t\in T} L_t) \tag{3}$$

All classification trees providing a final decision by voting method are given by:

$$H(a)=\arg\max_{y_j} \textstyle\sum_{i\in[1,2,...,Z]} I(h_i(a)=y_j) \qquad j=1,2,...,C \tag{4}$$

where the combination model is $H(a)$, the number of training subsets are Z depending on which decision tree model is $h_i(a)$, $i\in[1,2,...,Z]$ while output or labels of the P classes are y_j , $j=1,2,...,P$ and combined strategy is $I(\cdot)$ defined as

$$I(x)=1 \quad \text{If } h_i(a)=y_j \quad \text{else } I(x)=0 \tag{5}$$

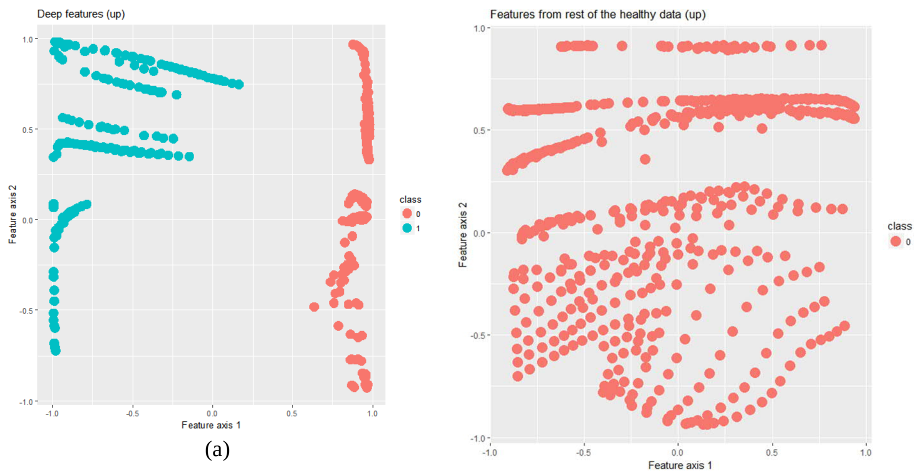
ESANN 2019 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2019, i6doc.com publ., ISBN 978-287-587-065-0. Available from http://www.i6doc.com/en/.

where output of the decision tree is $h_i(a)$ and i^{th} class label of the P classes is y_j, j=1,2,...,P .

3 Results and discussion

In this research, we first selected the faulty data based on time periods provided by the maintenance data. In the next step, an equal amount of healthy data was also selected and labelled as class 0 for healthy, with class 1 for faulty data. Finally, the deep autoencoder model is used for feature extraction from the data.

We have analysed up and down movements separately because the traction based elevator usually produces slightly different levels of vibration in each direction. Healthy and faulty data with class labels are fed to the deep autoencoder model and the generated deep features are shown in Fig. 2 (a). In Fig. 2 (a), we can see that both features with class labels are perfectly separated, which results in better fault detection. These are called deep features or latent features in deep autoencoder terminology, which shows hidden representations of the data. The extracted deep features are fed to the random forest algorithm for classification and the results provide 100% accuracy in fault detection, as shown in Table 1 (a). We have also calculated accuracy in terms of avoiding false positives from both features and found that the new deep features generated in this research outperform the existing features. We have used the rest of the healthy data to analyse the number of false positives. This healthy data is labelled as class 0 and fed to the deep autoencoder to extract new deep features from the data, as presented in Fig. 2 (b).

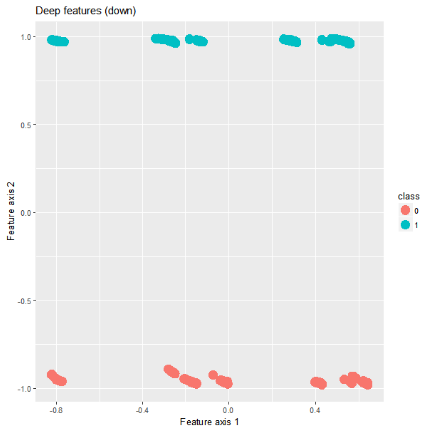


(a)

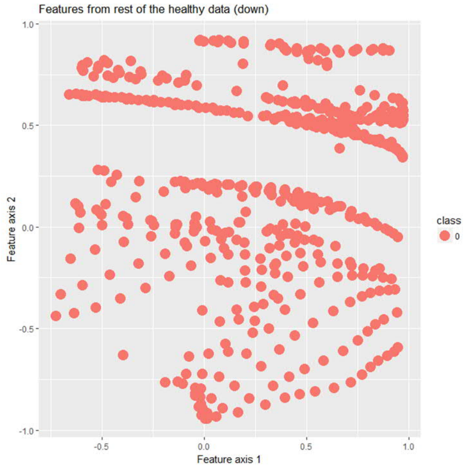
(b)

Fig. 2: Up (a) Extracted deep autoencoder features (visualization of the features w.r.t class variable) (b) Extracted deep features (only healthy data).

These new deep features are then classified with the pre-trained deep autoencoder random forest model to test the efficacy of the model in terms of false positives. For downward motion (see Fig. 3 (a, b)), just as in the case of up movement, we feed both healthy and faulty data with class labels to the deep autoencoder model for the extraction of new deep features. The new extracted deep features are classified with random forest model. After this, the rest of the healthy data with class label 0 is used to analyze the number of false positives and the results are shown in Table 1 (b).



(a)



(b)

Fig. 3: Down (a) Extracted deep features (b) Extracted deep features (only healthy data).

	Deep features	Existing features
Accuracy	1	0.65
False positives	1	0.61

(a)

	Deep features	Existing features
Accuracy	1	0.62
False positives	0.95	0.58

(b)

Table 1: (a) Fault detection analysis- up (false positives field related to analyzing the rest of the healthy data after the training and testing phase) (b) Fault detection analysis- down.

4 Conclusion

This research focuses on the health monitoring of elevator systems using a novel fault detection technique. The goal of this research was to develop a generic model for automated feature extraction and fault detection in the health state monitoring of elevator systems. Our approach in this research provided nearly 100% accuracy in fault detection and also in the case of analyzing false positives with new extracted deep features. The results support the goal of this research of developing a generic model which can be used in other machine systems for fault detection. We have used almost one year of data from seven traction elevators in this research, which proves the generalisation capability of our approach. The results are useful in terms of detecting false alarms in elevator predictive maintenance. The approach will also reduce unnecessary visits of maintenance personnel to installation sites if the analysis results are utilized to allocate maintenance resources. Our developed model can also be used for different predictive maintenance solutions to automatically generate highly informative deep features for solving diagnostics problems. The results outperform the existing features calculated from the raw sensor dataset of the same elevators. The automated feature extraction approach does not require any prior domain knowledge. It also provides dimensionality reduction and is robust against overfitting characteristics. The experimental results show the feasibility of our generic model, which will increase the safety of passengers as well as serve the public interest. We have tested the robustness of our model in the case of a large dataset, which proves the efficacy of our model.

References

[1] U. Desa, World urbanization prospects, the 2011 revision, Population Division, Department of Economic and Social Affairs, United Nations Secretariat, 2014.

[2] G. Niu, S. Lee, B. Yang and S. Lee, Decision fusion system for fault diagnosis of elevator traction machine, Journal of Mechanical Science and Technology, 22, pp. 85–95, 2008.

[3] R. Zhang, Z. Peng, L. Wu, B. Yao and Y. Guan, Fault diagnosis from raw sensor data using deep neural networks considering temporal coherence, Sensors, vol. 17, (3), pp. 549, 2017.

[4] M. Xia, T. Li, L. Xu, L. Liu and C.W. de Silva, Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks, IEEE/ASME Transactions on Mechatronics, vol. 23, (1), pp. 101-110, 2018.

[5] F. Calimeri, A. Marzullo, C. Stamile, and G. Terracina, Graph based neural networks for automatic classification of multiple sclerosis clinical courses, in Proceedings of the ESANN 2018.

[6] I. Fernández-Varela, D. Athanasakis, S. Parsons, E. Hernández-Pereira, and V. Moret-Bonillo, Sleep staging with deep learning: A convolutional model, in Proceedings of the ESANN 2018.

[7] J. Hänninen and T. Kärkkäinen, Comparison of four-and six-layered configurations for deep network pretraining, in Proceedings of the ESANN 2016.

[8] R. Huet, N. Courty and S. Lefèvre, A new penalisation term for image retrieval in clique neural networks, in Proceedings of the ESANN 2016.

[9] F. Fogelman-Soulie, Y. Robert and M. Tchuente, Automata Networks in Computer Science: Theory and Applications. Manchester University Press and Princeton University Press, 1987.

[10] P. Madani and N. Vlajic, Robustness of deep autoencoder in intrusion detection under adversarial contamination, in Proceedings of the 5th Annual Symposium and Bootcamp on Hot Topics in the Science of Security, 2018, pp. 1..

[11] S. Timotheatos, G. Tsagkatakis, P. Tsakalides, and P. Trahanias, Feature extraction and learning for RSSI based indoor device localization, in Proceedings of the ESANN 2017.

[12] G. Jiang, P. Xie, H. He, and J. Yan, Wind turbine fault detection using a denoising autoencoder with temporal information, IEEE/ASME Trans. Mechatronics, vol. 23, no. 1, pp. 89–100, 2018.

[13] N. Japkowicz, S. J. Hanson and M. A. Gluck, Nonlinear autoassociation is not equivalent to PCA, Neural Comput., vol. 12, (3), pp. 531-545, 2000.

[14] L. Breiman, Random forests, Machine Learning, vol. 45, no. 1, pages 5-32. 2001.

[15] A. Liaw and M. Wiener, Classification and regression by random forest, R News, vol. 2, (3), pp. 18-22, 2002.

[16] T. Huynh, Y. Gao, J. Kang, L. Wang, P. Zhang, J. Lian and D. Shen, Senior Member, IEEE, and Alzheimer's Disease Neuroimaging Initiative, Estimating CT image from MRI data using structured random forest and auto-context model, IEEE Transactions on Medical Imaging, vol. 35, (1), pp. 174-183, 2016.

[17] Z. Liu, Member, IEEE, B. Tang, X. He, Q. Qiu, and F. Liu, Class-specific random forest with cross-correlation constraints for spectral–spatial hyperspectral image classification, IEEE Geoscience and Remote Sensing Letters, vol. 14, (2), pp. 257-261, 2017.