Tampere University

KRISHNA MOHAN MISHRA

Deep Neural Networks for Elevator Fault Detection

Tampere University Dissertations 460

Tampere University Dissertations 460

KRISHNA MOHAN MISHRA

Deep Neural Networks for Elevator Fault Detection

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty of Engineering and Natural Sciences
of Tampere University,
for public discussion in the Pieni Sali 1,
of the Festia Building, Korkeakoulunkatu 8, Tampere,
on 17th September 2021, at 12 o'clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Engineering and Natural Sciences,
Finland

Responsible supervisor and Custos	Prof. Kalevi Huhtala Tampere University Finland	
Pre-examiners	Prof. Leif Kari KTH Royal Institute of Technology Sweden	Associate Prof. Matti Rantatalo Luleå University of Technology Sweden
Opponents	Prof. Leif Kari KTH Royal Institute of Technology Sweden	Dr. Jari Vepsäläinen, Aalto University Finland

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

Cover design: Roihu Inc.

ISBN 978-952-03-2074-4 (print)
ISBN 978-952-03-2075-1 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/URN:ISBN:978-952-03-2075-1

PunaMusta Oy – Yliopistopaino
Joensuu 2021

PREFACE

This study was carried out at Tampere University, Finland between 2017 and 2020.

First and foremost, I would like to express my sincere gratitude to my supervisor Prof. Kalevi Huhtala, for accepting me into his research group as a doctoral student in 2017, and supporting, guiding and encouraging me in my academic career since then.

I would like to thank all of my co-authors, Dr. Tomi Krogerus, John-Eric Saxen and Dr. Jerker Björkqvist. This thesis would not be possible without your deeply valuable contributions.

I would like to thank the pre-examiners of this thesis, Prof. Matti Rantatalo and Prof. Leif Kari, and the opponents of the public defense of this thesis, Prof. Leif Kari and Dr. Jari Vepsäläinen.

I would like to thank all the members of our research group for creating such a friendly and motivating research atmosphere.

Finally, I would like to thank my parents, my wife, and my family for supporting me in this academic journey.

ABSTRACT

The objective of this thesis is to develop novel data extraction, feature extraction and fault detection techniques for the task of elevator fault detection in real-world environments. Aim of the research is to develop systems that can automatically detect the elevator faults commonly present in the systems. In addition, this research will help various predictive maintenance systems to detect false alarms, which will in turn reduce unnecessary visits of service technicians to installation sites. The proposed solutions answer two research questions: how can we detect elevator faults efficiently and how can we detect false alarms in elevator predictive maintenance systems?

Five publications have been developed to address these issues from various perspectives. In this thesis, modern machine learning method called deep learning is applied for elevator fault detection. The relationship between the commonly used time series representations for elevator movement and the target fault event labels are highly complex. Deep learning methods such as deep autoencoder and multilayer perceptron utilize a layered structure of units to extract features from the given vibration input with increased abstraction at each layer. This increases the network's capacity to efficiently learn the highly complex relationship between the elevator movement and the target fault event labels. This research shows that the proposed deep autoencoder and multilayer perceptron approaches perform significantly better than the established classifying techniques for elevator fault detection such as Random forest algorithm.

An off-line profile extraction algorithm is also developed based on low-pass filtering and peak detection to extract elevator start and stop events from sensor data. These profiles are used to calculate motion and vibration related features, which is called here existing features. Profile extraction algorithm and deep autoencoder model are

combined to calculate new deep features from the data to improve the results in terms of elevator fault detection. The approaches 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 generic models which can be used in other machine systems for fault detection.

CONTENTS

List of Figures

List of Tables

ABBREVIATIONS

A95	95% of acceleration or vibration
ANN	Artificial neural network
AR	Autoregressive
BDI	Belief-desire-intention
DNN	Deep neural network
DTW	Dynamic time warping
EMD	Empirical mode decomposition
FN	False negative
FNN	Feed-forward neural network
FP	False positive
IMF	Interplanetary magnetic field
ISO	International organization for standardization
MAS	Multi-agent system
MKL	Multiple kernel learning
MLP	Multilayer perceptron
NMEA	National marine electronics association
PP	Peak-to-peak
RF	Random forest
RMS	Root mean square
RP	Research problem
SNR	Signal to noise ratio

TN	True negative
TP	True positive
UAV	Unmanned aerial vehicle
WHO	World health organization

ORIGINAL PUBLICATIONS

Publication I	K. M. Mishra, T. Krogerus and K. Huhtala. Deep autoencoder feature extraction for fault detection of elevator systems. *in Proceedings of the 27th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)* 27 (2019), 191–196.
Publication II	K. M. Mishra, T. Krogerus and K. Huhtala. Fault detection of elevator systems using deep autoencoder feature extraction. *in Proceedings of the IEEE 13th International Conference on Research Challenges in Information Science (RCIS)* 13 (2019), 43–48.
Publication III	K. M. Mishra, J. Saxen, J. Bjorkqvist and K. Huhtala. Fault detection of elevator system using profile extraction and deep autoencoder feature extraction. *in Proceedings of the 33rd annual European Simulation and Modelling Conference (ESM)* 33 (2019).
Publication IV	K. M. Mishra and K. Huhtala. Elevator Fault detection using profile extraction and deep autoencoder feature extraction for acceleration and magnetic signals. *Applied Sciences* 9 (2019), 15p.
Publication V	K. M. Mishra and K. Huhtala. Fault detection of elevator systems using multilayer perceptron neural network. *in Proceedings of the 24th IEEE Conference on Emerging Technologies and Factory Automation (ETFA)* 24 (2019), 904–909.

1 INTRODUCTION

Elevators (see Figure 1.1) are considered as an efficient way of transportation that allows a get on and get off facility to passengers at the floor specified. There are 12 million elevators in the world [13], which makes it the biggest global transportation system. High-speed elevator systems required by high-rise buildings to provide rapid access inside buildings have been built in modern cities. Remote monitoring introduced in the 1980s only provides information about breakdowns but cannot reduce the amount of them. In the 1990s, usage-based maintenance was introduced and adapted by the automobile industry later on in which after a certain distance travelled the motor oil must be changed.

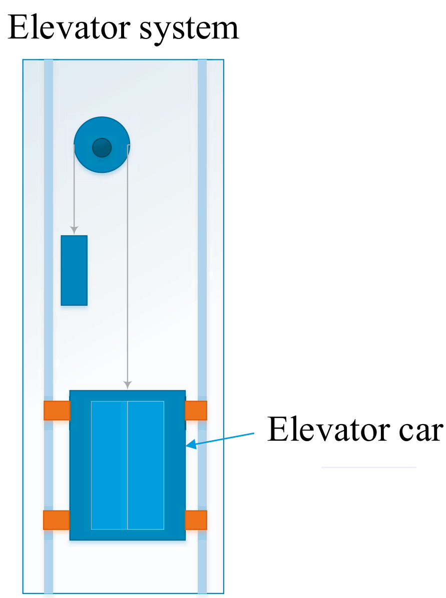


Figure 1.1 Elevator system

Elevator companies are currently focusing on these issues and finding optimal solutions. Traditional elevator maintenance service is not suitable for the current scenario and needs to be improved in order to support better elevator service business. Both horizontal and vertical transportation due to growing urbanization face transport logistics problems that require intelligent transportation systems with increased reliability, capacity and efficiency. In this thesis, the focus is on the automatic detection of elevator faults in real-world environments. Methods are proposed and developed that can be used for automatic elevator fault detection.

1.1 Research Problems

In recent years, elevator systems have been used more and more extensively in apartments, commercial facilities and office buildings [V]. Nowadays 54% of the world's population lives in urban areas [13]. 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 of the system [16], [15]. Elevator production and service companies are now opting for a predictive maintenance policy to provide better service to customers. They are remotely monitoring faults in elevators and estimating the remaining lifetime of the components responsible for faults. Elevator systems require fault detection and diagnosis for healthy operation [74]. Condition monitoring is a multidisciplinary problem, which includes many engineering branches. It also required implementation of variety of complex technologies. Fault detection and diagnosis is performed using pre-existing tools. Objective of the research was to answer two research questions with multidisciplinary approaches.

RP.I Fault Detection of Elevator System
How can elevator faults be detected efficiently?

RP.II False Alarms in Elevator Predictive Maintenance Systems
How can false alarms be reduced in elevator predictive maintenance systems?

Existing hardware and software modules have been utilized for fulfilling the requirement of commercial application of outputs.

1.1.1 RP.I: Fault Detection of Elevator System

Condition monitoring is an essential part of every machine maintenance system. Elevators are frequently used by people nowadays, which requires proper maintenance and safety of elevators. Fault detection and diagnosis is very important in smooth functioning of elevator systems. Traditional methods are not very efficient in detecting faults; thus this research focused on developing a deep learning model for efficient fault detection and diagnosis. The research also addressed the challenges of dimensionality reduction and robustness against overfitting characteristics. This research includes calculation of highly informative deep features from raw sensor data along with existing features. State of the art includes fault diagnosis methods having feature extraction methodologies based on deep neural networks [73], [37], [6] and convolutional neural networks [77], [31] for rotatory machines similar to elevator systems. Fault detection methods for rotatory machines are also using support vector machines [44] and extreme learning machines [79]. However, to improve the performance of traditional fault diagnosis methods, an intelligent deep autoencoder model is developed for feature extraction from the data and random forest performs the fault detection in elevator systems based on extracted features.

1.1.2 RP.II: False Alarms in Elevator Predictive Maintenance Systems

False alarms is a challenging task to handle in condition monitoring of every machine maintenance system. Fault diagnosis methods based on deep neural networks [85], [26] provide above 90% accuracy in fault detection, but false alarms are not considered into analysis. In this research, false alarm analysis is also considered along with fault detection analysis. False positives are considered as evaluation parameter for this research in addition to the accuracy, sensitivity and specificity. False alarms are directly related to false positives in this research, which is calculated by validating the pre-trained model with the remaining healthy data along with training and testing phase. Therefore, both RPs should be considered to accomplish elevator fault detection using deep learning approaches.

Studies focusing on these RPs are included in this thesis. For example, RP.I and RP.II were considered by Publication-I to Publication-V from various perspectives.

1.2 Objectives of the Thesis

The main objective of this thesis is to study and utilize the recently proposed advanced machine learning techniques in the context of elevator fault detection. These techniques include deep autoencoders, multilayer perceptron (MLP) neural network and off-line profile extraction algorithm. While utilizing these techniques for various elevator fault detection tasks, we aim to develop an understanding of the working principles of the neural networks for a specific elevator fault detection problem. The wide range of elevator fault detection tasks that are tackled in this thesis provide a clear idea on the scalability and the robustness of these machine learning techniques. Lastly, the aim is to investigate the end-to-end elevator fault detection and propose a novel method which is obtained through the hidden layer outputs of a neural network.

In this thesis, the main research questions that can be asked listed as follows. Research investigates whether the modern machine learning techniques such as deep learning methods can be used to develop elevator fault detection systems with robust performance in real-life conditions. Thesis tests the multi-label learning capabilities of the deep learning methods for elevator fault detection in various levels. Research investigates on which deep learning methods perform the best for a given elevator fault detection task, and how should the model architecture be adjusted for the optimal performance. Thesis analyzes the effectiveness of the established elevator fault detection techniques, and searches for ways to make deep learning methods model the elevator fault events directly from raw sensor signals.

1.3 The Author's Contribution to the Publications

This section briefly explains the role of the author in each of the listed publications.

Publication I: The author developed the initial idea and developed the methods presented and the theoretical framework and the coding of the algorithms and wrote the paper. Dr. Tomi Krogerus has assisted for the initial idea and the data extraction code. As academic supervisor, Professor Kalevi Huhtala reviewed the paper and made corrections and suggestions.

Publication II: The author developed the initial idea and developed the methods presented and the theoretical framework and the coding of the algorithms and wrote the paper. Dr. Tomi Krogerus has assisted for the initial idea and the data extraction code. As academic supervisor, Professor Kalevi Huhtala reviewed the paper and made corrections and suggestions.

Publication III: The author developed the initial idea and developed the methods presented and the theoretical framework and the coding of the algorithms and wrote the paper. M.Sc. John-Eric Saxen has assisted for the profile extraction algorithm code. Dr. Jerker Björkqvist reviewed the paper. As academic supervisor, Professor Kalevi Huhtala reviewed the paper and made corrections and suggestions.

Publication IV: The author developed the initial idea and developed the methods presented and the theoretical framework and the coding of the algorithms and wrote the paper. As academic supervisor, Professor Kalevi Huhtala reviewed the paper and made corrections and suggestions.

Publication V: The author developed the initial idea and developed the methods presented and the theoretical framework and the coding of the algorithms and wrote the paper. As academic supervisor, Professor Kalevi Huhtala reviewed the paper and made corrections and suggestions.

1.4 Outline and Structure of the Thesis

The organization of the remainder of this thesis is as follows.

The background information about elevator fault detection, and the topics of problem formulation , applications, challenges, data representations, and different datasets used in the context of elevator fault detection are presented in Chapter 2.

Chapter 3 presents methodology including feature learning, machine learning, artificial neural networks, profile extraction algorithm, tree based algorithms and evaluation methods in the context of elevator fault detection.

Chapter 4, summary of publications includes summaries for each of the five publications. This chapter explains the connection between the thesis RPs and the publications.

Conclusions of this thesis and discussions for the current and future research on elevator fault detection are provided in Chapter 5. This chapter is followed by the publications, which are relevant published papers.

2 BACKGROUND

2.1 Problem Formulation

The goal of elevator fault detection is to automatically estimate the start and end times of the fault events present for a given collection of sensor data, and then to associate a textual label to each of these fault events. These textual labels are often called classes. Elevator fault detection can be formulated in two stages: data pre-processing and classification. In the data pre-processing stage, deep features are extracted for each short time frame t in the sensor data to obtain a feature vector $x_t \in \mathbb{R}^M$, where M is the number of features per frame. In the classification stage, the task is to learn a deep model that would estimate the event presence probabilities $p(y_t \mid x_t, \theta) \in [0, 1]^t$ for each pre-defined fault event class, where θ represents the deep model parameters of the classifier. In the usage case, the event presence probabilities are binarized by e.g. constant thresholding to obtain the event presence predictions $y_t \in [0, 1]^t$. By combining the presence predictions for consecutive time frames, one can determine the start and end times of the fault event classes.

In the scope of this thesis, deep model parameters θ are optimized using supervised learning (Section 2.5). There are also other learning methods to optimize the model parameters, such as unsupervised learning (often used when target outputs are unavailable/unused) and semi-supervised learning (used when the target outputs are available for only a portion of the data). In supervised learning for elevator fault detection, the binary target outputs y_t for each frame t are obtained from the reference annotations, which include the start and end time of each fault event in the sensor data. If the signals are obtained from the real-life environment, then the reference annotations are often collected manually, i.e. by a human through the signals and labeling the start and end times of the fault events that they could notice. Supervised learning for elevator fault detection is formulated in such a way that the fault event classes of interest are defined beforehand. This makes the elevator fault detection task concrete, and helps the human annotator in omitting the irrelevant fault events and grouping different fault events under a certain class (e.g. different kinds of faults under a single fault class as visualized in Figure 2.1). Healthy data is assigned binary 0 and faulty data 1 for machine learning classification. Data inside the time period starting from the complaint being reported and ending when maintenance is done are considered faulty, while all other data are considered healthy. To remove suspicious data we have left out a time of two weeks before the complaints were registered and data before that is considered healthy.

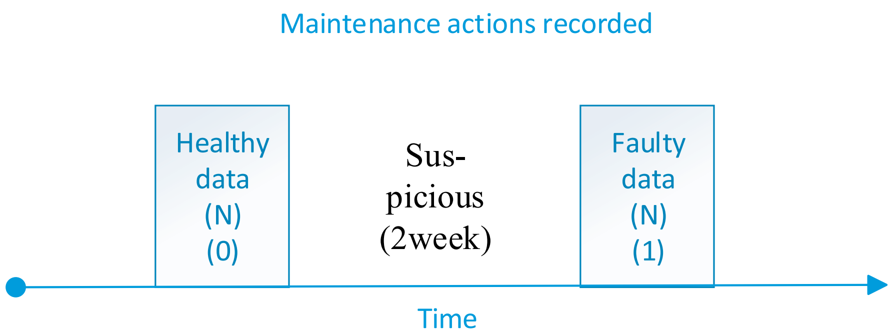


Figure 2.1 Data labelling

2.2 Applications

In recent years, there has been rapid growth in the use of elevator systems, mostly in apartments, commercial facilities and office buildings. Elevators are considered an efficient way of transportation that allows a get on and get off facility to passengers at the floor specified. There are 12 million elevators in the world, which makes it the biggest global transportation system. High-speed elevator systems required by high-rise buildings to provide rapid access inside buildings have been built in modern cities. Urban areas contain 54% of the world population, and that is continuously increasing [13]. According to the World Health Organization (WHO), 70% of the world's population will live in large cities by 2050. Therefore, automatic elevator fault detection can be utilized for aircraft subsystem solutions [9], robot transportation system [1], smart cities including smart buildings [82], satellite/climber system

[34] and indoor navigation systems [51].

Recently, elevator fault detection systems have been commercialized into automatic alarm systems [39], propulsion controlled aircraft [71], large civil aircraft [76], small unmanned aerial vehicle (UAV) [65], traffic analysis [67], quality evaluation system [87] and energy recovery systems [59].

2.3 Challenges

It can be claimed that the research progress on elevator fault detection has been stagnant until the recent years. One of the reasons is that there are several challenges for a robust elevator fault detection system that can operate in real-life conditions. The challenges for elevator fault detection systems can be listed as, but not limited to, unstructured data, missing values, feature selection, data balancing, human behaviour and false alarms.

Unstructured Data

Unstructured data [19] is a type of data, which do not have class information. Data labelling is an expensive process, therefore most of the real-life data do not have label information. In this research, some maintenance records have been used to provide the class information for machine learning algorithms.

Missing Values

Missing values [4] are also common problem in analyzing real-life data. It is often found in data due to some data collection errors. In this study, the developed methods are capable of handling these missing values during the analysis. In addition, it has been removed during the data pre-processing stage.

Feature Selection

Feature selection [40] is one of the most important part of data analysis. Usually, the real-life data contain many features but it is not necessary that all of them are important for certain analysis. Therefore, in this research feature selection plays a significant role in selecting suitable features for elevator fault detection.

Data Balancing

Real-life data usually do not have large amount of faulty data because in general the faults had been repaired quickly as they reported. Therefore, to avoid the problem of overfitting [36] data balancing is required. In this study, balanced dataset is used for machine learning algorithms to detect faults in the elevator systems.

Human Behaviour

Human behaviour plays a significant role in fault detection analysis because it will be a case of uncertainty while selecting healthy and faulty data based on maintenance actions recorded. It can be a possibility that the time of actual fault is different than the time reported because of the human behaviour, which is considered in this research.

False Alarms

False alarms [80] need to be reduced in predictive maintenance systems because it costs huge amount of money in unnecessary visits of service technicians to installation sites. In this study, false alarms are analyzed in terms of false positives as an evaluation parameter for developed machine learning methods.

2.4 Data Representation

Sensor data for elevator fault detection [39] are obtained by digitally recording the elevator movement in a real-life environment. The time series representation [57] of a elevator movement is considered as the lowest level representation, since the signal is not much processed before using it as the representation of a elevator movement. On the other hand, this representation is quite redundant for a classifier to learn which elevator movement it belongs to. For this reason, sensor data for elevator fault detection are often represented by extracting certain features. Different types of sensor datasets are used in this research e.g. existing features sensor data, most frequent floor patterns raw sensor data, profiles from acceleration signals of raw sensor data and profiles from both acceleration and magnetic signals of raw sensor data.

Existing Features Sensor Data

Raw sensor data, mainly acceleration signals, were used to calculate elevator key performance and ride quality features, which we call here existing features. In publications [I] and [V], 12 different existing features derived from raw sensor data describing the motion and vibration of an elevator for fault detection and diagnostics [54] of multiple faults are utilized. These existing features (see Figure 2.2) are e.g. peak-to-peak (PP), A95 and root-mean-square (RMS) defined by ISO standards [24].

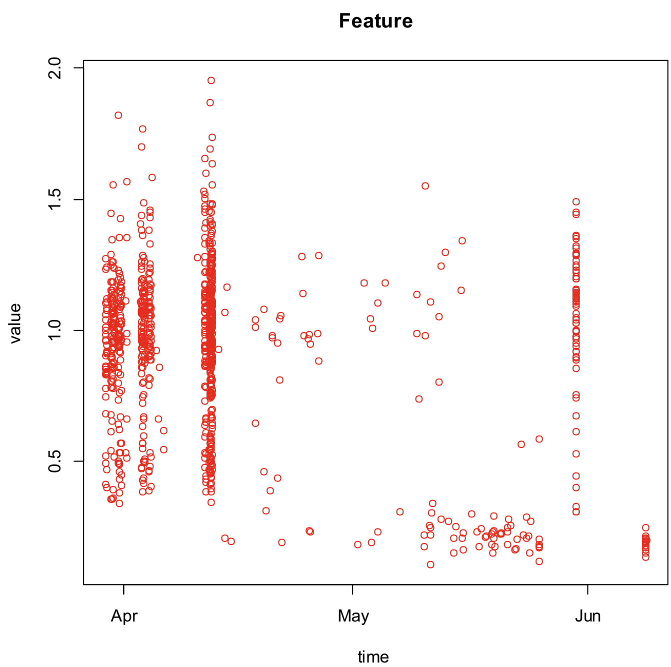


Figure 2.2 Existing features

Most Frequent Floor Patterns Raw Sensor Data

In publication [II], data is selected from the most frequent floor patterns of the data, i.e. floor patterns which consist of the maximum number of rides between specific floor combinations. Only the vertical component of acceleration data is selected in this research because it is the most informative aspect consisting of significant changes in vibration levels [29] as compared to other components. Data is selected based on elevator movement as shown in Figure 2.3.

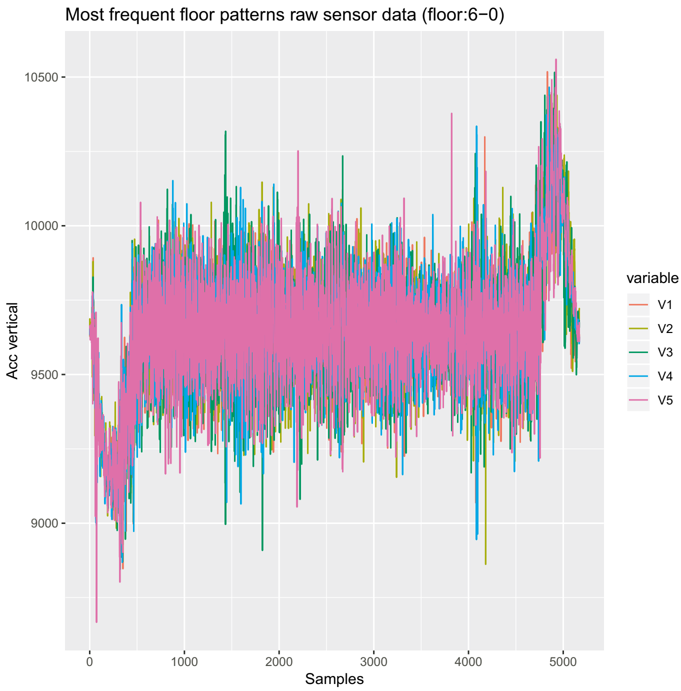


Figure 2.3 Most frequent floor patterns raw sensor data (Acc vertical represents acceleration signals in z-direction)

Profiles from Acceleration Signals of Raw Sensor Data

In publication [III], start and stop profiles [17] are extracted from the rides because of the different lengths of rides for each floor combination due to the constant speed phase, which is longer when there is longer travel. Up and down movements have analyzed separately because the traction based elevator [41] usually produces slightly different levels of vibration in each direction. Profiles from acceleration signals of raw sensor data have been visualized in Figure 2.4.

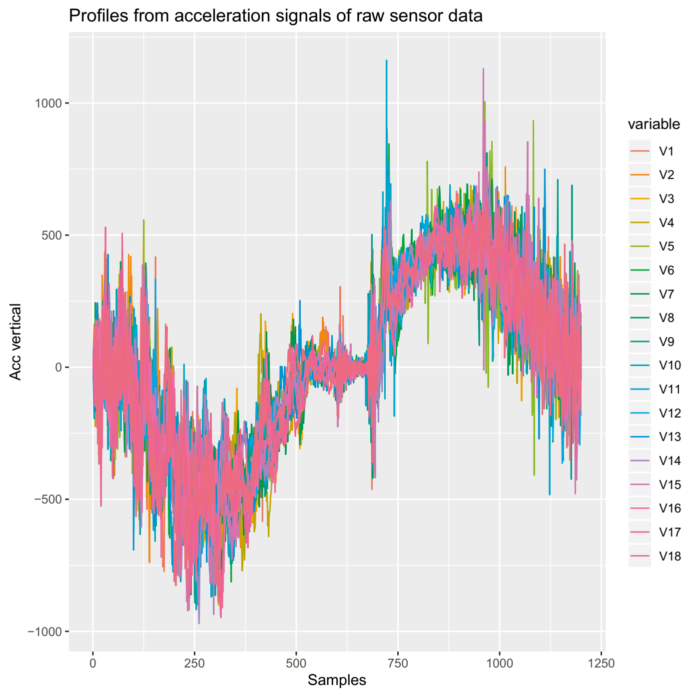


Figure 2.4 Profiles from acceleration signals of raw sensor data (Acc vertical represents acceleration signals in z-direction)

Profiles from Magnetic Signals of Raw Sensor Data

In publication [IV], profiles from magnetic signals [42] of raw sensor data are extracted along with acceleration signals as an extension to the publication [III] . This approach has provided more data to validate the machine learning model and robustness against overfitting characteristics [83]. Only the vertical component of the signal is used and up and down movements have been analyzed separately similar to the previous approaches. Profiles from magnetic signals of raw sensor data are presented in Figure 2.5.

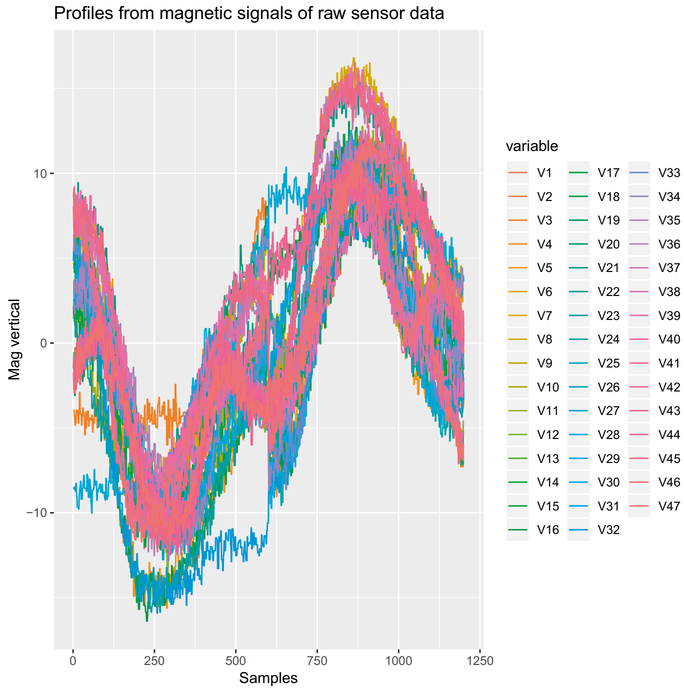


Figure 2.5 Profiles from magnetic signals of raw sensor data (Mag vertical represents magnetic signals in z-direction)

2.5 Datasets

The datasets used in the experiments for this thesis are explained below in detail.

Raw Sensor Data

Raw sensor datasets are collected according to the movement of elevators. Each elevator has different raw sensor dataset, however each raw sensor dataset includes three axis acceleration and magnetic field signals. In addition, raw sensor dataset has information about floor movement of elevator system. Currently, the dataset is not publicly available due to copyright issues. Most frequent floor patterns from the data, i.e. floor patterns which consist of the maximum number of rides between specific floor combinations are used in publication [II]. In publication [III], an off-line profile extraction algorithm is developed based on low-pass filtering and peak detection to extract elevator start and stop events from raw sensor data. The off-line profile extraction algorithm extracts elevator start and stop events from both acceleration and magnetic signals separately in publication [IV].

Existing Features Data

Existing features are calculated from the raw sensor data based on elevator movement. Every movement of the elevator generates one set of existing features from the vibration signal. In addition, each elevator usually produces around 200 rides per day. Each ride used in analysis contains around 5000 rows of the raw sensor data. Currently, the dataset is not publicly available due to copyright issues. In publication [I], a generic deep autoencoder model has been proposed for automated feature extraction from the existing features data. In publication [V], a generic multilayer perceptron (MLP) neural network model is proposed for automated feature extraction from the existing features data for elevator fault detection.

3 METHODOLOGY

3.1 Feature Learning for Elevator Fault Detection

Recently, most of the proposed elevator fault detection methods simply pick one of the well-known deep feature extraction methods for sensor data, and focus on developing a novel and better performing deep model through machine learning. Apart from the feature extraction methods, machine learning can also be used to learn discriminative features from low-level representations for elevator fault detection. Low-level representations may refer to time series signal, or the existing features. The method of using machine learning for the deep feature extraction is often called feature learning (or representation learning).

There are several feature learning methods that have been proposed for elevator fault detection. Song et al. [68] proposed multiple kernel learning (MKL) for analysing elevator dataset. It is using deep neural network architecture to learn the parameters through sophisticated optimization procedures for combining the kernels. Accuracy of 90.2% through MKL method is better than other methods used in the analysis. Li [38] proposed three reinforcement feature learning algorithms e.g. Q-learning, Q-value and multi-step Q-learning for elevator control system. Average waiting time used as evaluation parameter and multi-step Q-learning algorithm performs best out of three algorithms applied. Results prove that reinforcement algorithms are better than classical methods for elevator control system. Kim et al. [33] designed safety elevator monitoring system for marine elevators based on NMEA 2000 network. It uses big data feature learning from the server to provide prediction for maintenance. Accuracy of slope prediction model with load and platform tilt is 99% while roll and pitch based model provides 94%. Skog et al. [66] have proposed a signal processing scheme with a smart sensor for the Internet-of-Elevators. The sensor node behaves like a self-contained black box unit, which supports condition monitoring capabilities in a cost efficient way by modernizing the conventional elevator system. Features related to condition monitoring of elevator systems similar to ride quality parameters are calculated by vibration versus frequency spectrum, and the vibration versus position spectrum. Low cost sensors such as smartphones were used in experiments that reduced the error in estimating the position of the elevator up to less than 1 meter with 99.9% probability in a 43 s long travel. Wang et al. [74] presented a fault diagnosis approach for elevator braking systems using a wavelet packet algorithm and fuzzy neural network. The wavelet packet method was used to decompose fault signals and to extract eight frequency components of signals, from low to high frequency. A fault diagnosis model was designed using a fuzzy neural network along with B-spline, and the model input included eight obtained eigenvalues. According to the authors the method provides 100% accuracy in brake fault recognition and 87% for fault type identification. Zhang et al. [84] applied intelligent feature learning agent technology to the traditional fault diagnosis method in fault diagnosis systems. A belief-desire-intention (BDI) agent based fault diagnosis system was constructed using an elevator door multi-agent system (MAS). The authors argue that the advantages of the system are autonomy of the diagnosis process, the reusability of diagnostic resources (knowledge diagnosis, diagnostic agent, etc.) and the scalability of the system. Zhao et al. [88] studied fault detection in elevator systems using feature learning. Their research explains the types of faults occurring in elevator systems as well as the reasons for their occurrence, and also proposed a method for fault detection.

3.2 Machine Learning for Elevator Fault Detection

Machine learning [7] is a field of computer science that aims to design machines (or software) that are able to learn directly from the given data. Machine learning systems are especially useful for certain tasks for whom the solutions are very challenging to implement as an algorithm by a human engineer. Elevator fault detection [78] can be given as an example for such tasks. Due to the challenges explained in Section 2.3, it is difficult to come up with an engineered algorithm that could map sensor signals with their corresponding fault events with satisfactory accuracy to be utilized in real-world applications. By studying how machine learning systems for elevator fault detection process sensor data, it may even be possible in the future to gain more insight about how humans perceive fault events.

Some of the earlier work on machine learning for elevator fault detection are studied here. In [86] the structural health monitoring of elevators is performed by sensor network data. The authors argue that wireless sensors will play an important role in future studies, though currently wired systems still dominate. In this research, data transmission is performed with a wired system, which is still a valid choice for a life-long monitoring system. The authors also state that structured data, which can be further fed to machine learning algorithms for classifying elevator health states, is difficult to produce and interpret. Li [30] studied the measurement systems and fault diagnosis of elevators. The author states that elevator faults are electromechanical, and during the running phase mechanical faults have a frequency attached that can rise randomly. Vibrations of the elevator are the main cause of these mechanical faults attached to the running phase. The following analysis were conducted: frequency response analysis, modal analysis, cepstrum analysis, spectrum analysis, autocorrelation analysis, time-domain parameter analysis and waveform analysis. Jinjin [32] and Jian-can [27] studied the fault diagnosis of elevator systems using vibration signals. The evaluation of vibration signals included the beginning of the acceleration process to the end of the travel, and contained starting, smooth operation and braking. Horizontal and vertical vibration signals were studied. Horizontal vibration occurs due to the guiding system that includes wire ropes, the degree of tightness of the guide shoes and the verticality of the guide slides in elevator cars. The driving lift systems, including damage to the gearbox and wheel wear, are responsible for vertical vibration. Niu et al. [53] developed a decision fusion system for intelligent fault diagnosis to satisfy the requirement of advanced maintenance and to realize real-time and convenient diagnosis of an elevator system. A faulty elevator motor system was analysed using vibration and current signals. Four-fusion algorithms: modified borda count, multi-agent, bayesian belief and majority voting were compared. Support vector machine provides the best accuracy of 75% for current signal among all the classification algorithms applied. Chen and Liu [11] proposed a mechanical fault diagnosis method using autoregressive (AR) dual spectrum. Nonlinear coupling described by AR bispectrum was used for gaussian noise elimination and retaining of phase information. Vibration signals produced by a running elevator were used to establish an AR time series model for analysis and diagnosis of elevator faults. Based on the study, the elevator has different dual spectral in normal and fault situations that can be easily distinguished. Qifeng et al. [60] studied a wavelet multi-threshold de-noising method for fault detection and diagnosis of elevators. Fault characteristic information was extracted from the vibration signal by de-noising it first. The method uses a correlation coefficient, signal to noise ratio (SNR) and root mean square (RMS) error of the original signal, and includes three methods: wavelet packet multi-threshold analysis, wavelet packet analysis and wavelet analysis. Chen and Liu [10] presented a fault diagnosis method for elevator systems using empirical mode decomposition (EMD) and box dimension. Interplanetary magnetic field (IMF) components are produced due to the decomposition of acceleration signals in normal and fault conditions by EMD in elevator systems. The calculation of IMF components and signals were performed by box dimension after the removal of noise and background signals from the recombined signal. Effective discriminations and ranges are produced by box dimension between normal and failure conditions, as shown in this research. Yaman et al. [78] developed an image-processing system for elevators to detect wear on guide-rail surfaces. Cameras were used in this research to monitor real-time conditions using a built-in system. Images were captured via four digital cameras of elevator guide-rail surfaces fixed onto the elevator cab. Detection of wear on the surface of the guide-rails was done by analysing the images captured by the cameras using image-processing methods.

3.3 Artificial Neural Networks

An artificial neural network (ANN) is a type of machine learning method, which is similar to human brain in terms of information processing. Neurons are the main constituting elements of human brain, which can be up to 15-20 billion in a single human brain. These are inter-connected nodes stimulated by various electrochemical signals [58]. Each type of signals e.g. visual, audio or sensory etc. has their own path of neurons, which require a set of neurons to process their information inside the brain. Every human brain has its own speciality for processing certain type of signals, which always improves with time to create a mapping between certain input signals to its output representation. Deep learning, or deep neural networks (DNN) is also a type of artificial neural network with more than one hidden layer. This research investigates the application of deep neural network with feed-forward neural network having multiple hidden layers on elevator fault detection. Deep learning [85], [26], [I] has also been found in many state-of-the-art methods for rotatory machine fault detection similar to elevator systems.

3.3.1 Feed-forward Neural Networks

Feed-forward neural network (FNN) (see Figure 3.1) is a type of neural network, which has fully connected neurons as a set of sequential layers.

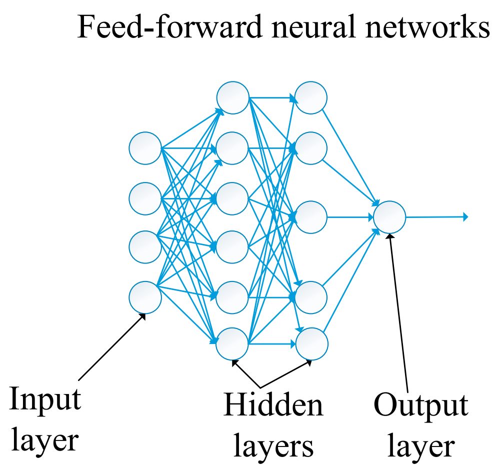


Figure 3.1 Feed-forward neural network

Input for all the neurons in one layer come as output of each neuron from previous layer. Feed-forward neural network [35] has no feedback connection for the neuron outputs i.e. layer outputs are calculated through a forward pass. The feed-forward neural networks [20] with multiple hidden layers, called as deep neural networks are often utilized for rotatory machine fault detection similar to elevator systems. Deep neural networks can learn high level representations in several layers of abstraction for modeling the complex input - target output relationships. Research [6], [56] and [II] provide examples of deep neural networks proposed for rotatory machine fault detection similar to elevator systems.

3.3.2 Deep Autoencoder Neural Networks

Deep autoencoder has been successfully used in various applications, which makes it a popular deep learning model [70]. In this research, deep autoencoder feature learning method is proposed for elevator fault diagnosis. A three-layer network including an encoder and a decoder is called an autoencoder. The encoder maps the high-dimensional input data into low-dimensional codes, while reconstruction of input data from these codes are performed by decoder [81]. A five layer deep autoencoder (see Figure 3.2) used in this study, which is a different approach than in [28], [72].

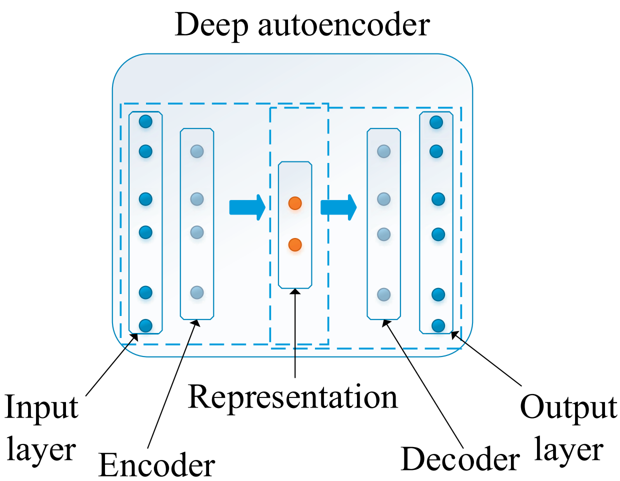


Figure 3.2 Deep autoencoder neural network

The encoder transforms the input x into corrupted input data $x^{'}$ using hidden representation H through nonlinear mapping

$$H = f(W_1 x^{'} + b) \tag{3.1}$$

where $f(.)$ 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 [72]. 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

$$x^{"} = g(W_2 H + c) \tag{3.2}$$

where *g(.)* 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 [25].

3.3.3 Multilayer Perceptron Neural Networks

Multilayer perceptron learns a non-linear function approximator, which is a supervised learning algorithm [21]. Non-linear layers called as hidden layers are existing between the input and the output layer. Multilayer perceptron is different from other algorithm, which can have one or more hidden layers (see Figure 3.3). Multilayer perceptron is a type of feedforward neural network, which can distinguish nonlinearly separable patterns. Multilayer perceptron includes several nodes called as neurons, those are arranged as a directed graph in multiple layers. Multilayer perceptron is a type of fully connected neural network, which is also known as universal approximators. Multilayer perceptron with one hidden layer having enough neurons can approximate any given continuous function [50].

At first, preparing the training dataset $D = \{(x_i, y_i)\}^n{}_{i=1}$, $x_i \in \mathbb{R}^{m*l}, y_i \in \mathbb{R}$. Where, n is the number of samples. $x_i (i = 1, 2, ..., n)$ is m-dimensional phased feature vector $I_i (i = 1, 2, ..., m)$ as the input of multilayer perceptron. y_i is the label of fault and the weighted input of j node in the hidden layer can be expressed as [18]:

$$h_{\mathrm{j}} = \sum_{i=1}^{m} W_{\mathrm{ij}} * I_{\mathrm{i}} + b_{\mathrm{j}} \tag{3.3}$$

Where $\mathrm{W_{ij}}$ is the connection weight which from the input layer i node to the hidden

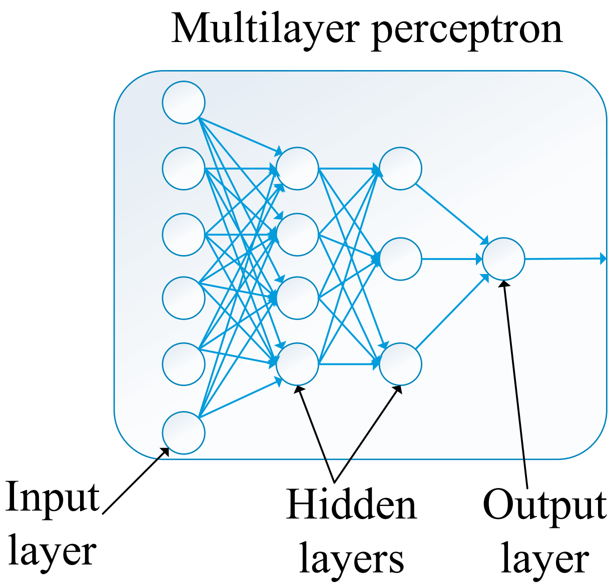


Figure 3.3 Multilayer perceptron neural network

layer j node, b_j is bias for the corresponding node, the output of the j node in the hidden layer is H_j .

$$H_{\mathrm{j}} = tanh(h_{\mathrm{j}}) \tag{3.4}$$

After several iterations, the input o_k of output layer k note from hidden layers is

$$o_{\mathrm{k}} = \sum_{j=1}^{J} W_{\mathrm{jk}} * H_{\mathrm{j}} + b_{\mathrm{k}} \tag{3.5}$$

Where output layer contains K notes $(k = 1,2,...,K)$. The output O_k of the k node in the output layer corresponding to different activation functions.

3.4 Profile Extraction Algorithm

The algorithm consists of two stages. First stage includes signal pre-processing and normalization. After this low-pass filter is applied to reduce noise spikes. Peak detection uses the low-pass filtered signal, which corresponding to acceleration and deceleration (start and stop) events, detects a local minimum and maximum for each elevator travel (see Figure 3.4).

Second stage includes alignment and collection of equal length profiles. It is based on dividing the acceleration signal in many windows near the peak events. Raw acceleration signal is used in this stage instead of the filtered signal. State of the art includes various time domain alignment methods. Commonly used method is dynamic time warping (DTW), e.g. in speech recognition [14]. Research [62] studied various alignment techniques for sensor data.

The off-line profile extraction algorithm is described as following.

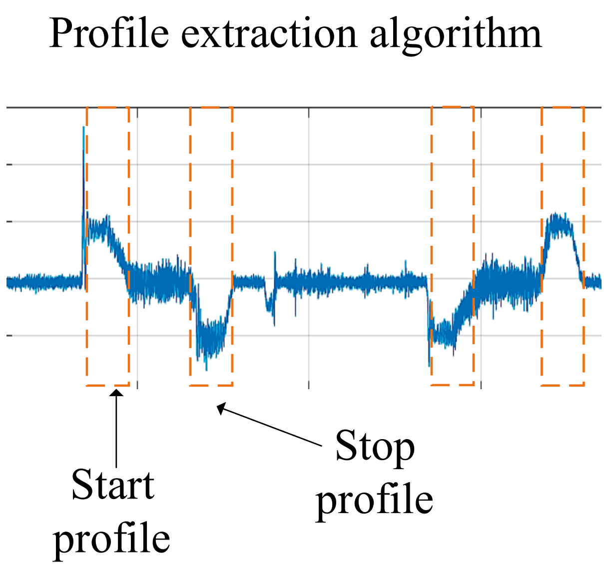


Figure 3.4 Profile extraction algorithm

Off-line profile extraction algorithm

Pre-procession

1. Read a vector of raw acceleration data containing k elevator travels. Define the zero mean transformed dataset as X.

2. Perform low-pass filtering on X and obtain denoised dataset Y.

Initialization

3. Define parameters for reference profile. Set the approximated maximum window length to m samples and height h to the 99th percentile of the low-pass filtered dataset.

4. Define alignment window size a and set k=1.

Iteration

5. From *Y(k)*, detect peak acceleration points y_{min} and y_{max}

6. Align reference profile P against raw dataset X in the vicinity of detected peaks by minimizing the L_2 norm according to

$$min \sum_{i=-a/2}^{a/2} \sum_{j=1}^{m} [-p_\mathrm{j} - x_{\mathrm{min+i+j}}]^2 \tag{3.6}$$

$$min \sum_{i=-a/2}^{a/2} \sum_{j=1}^{m} [p_\mathrm{j} - x_{\mathrm{max+i+j}}]^2 \tag{3.7}$$

7. Add aligned data points from *X(k)* as rows into an $n \times m$ profile matrix, alternatively separate matrices according to direction of travel (min/max).

8. Set travel window $k = k + 1$ and repeat steps 5–7 until end of dataset.

9. Update reference profile P with the signal-averaged profile obtained from the column-wise mean of the new profile matrix.

10. Reduce window length m by s samples, where s is the number of elements in P that satisfy

$$p \leq \epsilon, p \in P \tag{3.8}$$

where ϵ is a close to zero number indicating no acceleration.

11. Set $k = 1$ and continue with new batch iterations by repeating steps 5–8.

3.5 Tree Based Algorithms

Different tree based algorithms have been tested in this research and found random forest as the best classifier among all. Random forest is type of ensemble classifier selecting a subset of training samples and variables randomly to produce multiple decision trees [5]. High data dimensionality and multicollinearity can be handled by a RF classifier while imbalanced data affect the results of the RF classifier. It can also be used for sample proximity analysis, i.e. outlier detection and removal in train set [2]. The final classification accuracy of RF is calculated by averaging the probabilities of assigning classes related to all produced trees *(e)*. Testing data *(d)* that is unknown to all the decision trees is used for evaluation by voting method. Selection of the class is based on the maximum number of votes (see Figure 3.5). Random forest classifier provides variable importance measurement that helps in reducing the dimensions of hyperspectral data in order to identify the most relevant features of data, and helps in selecting the most suitable reason for classification of a certain target class.

Specifically, let sensor data value v_l^e have training sample l^{th} in the arrived leaf node of the decision tree $e \in E$, where $l \in [1, ..., L_e]$ and the number of training samples is L_e in the current arrived leaf node of decision tree e. The final prediction result is given by [23]:

$$\mu = \frac{\sum_{e \in E} \sum_{l \in [1,...,L_\mathrm{e}]} v_\mathrm{l}^\mathrm{e}}{\sum_{e \in E} L_\mathrm{e}} \tag{3.9}$$

All classification trees providing a final decision by voting method are given by [43]:

$$H(a) = \arg\max_{\mathrm{y_j}} \sum_{i \in [1,2,...,Z]} I(h_\mathrm{i}(a) = y_\mathrm{j}) \tag{3.10}$$

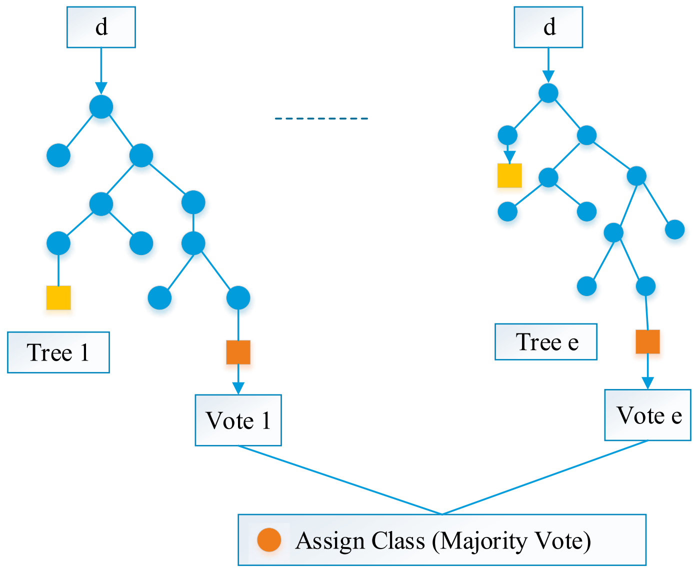


Figure 3.5 Classification phase of RF classifier

where $j = 1,2,...,C$ and 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 $I(.)$ defined as:

$$I(x) = \begin{cases} 1, & h_i(a) = y_j \\ 0, & \text{otherwise} \end{cases} \tag{3.11}$$

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

3.6 Evaluation of Elevator Fault Detection Methods

In order to effectively measure the improvements offered by a proposed scientific method, systematic evaluation is crucial. The evaluation parameters in this research are accuracy, sensitivity and specificity for the machine learning algorithm. The number of true negatives (TN), false negatives (FN), true positives (TP) and false positives (FP) can compute the efficiency of the classifier [64]. Statistical measurements of the tests are sensitivity and specificity. There are four possible outcomes for binary choice prediction (see Table 3.1):

- True Positive - Positive instance correctly classified as positive
- False Positive - Negative instance incorrectly classified as positive
- True Negative - Negative instance correctly classified as negative
- False Negative - Positive instance incorrectly classified as negative

Table 3.1 Confusion matrix.

	Predicted (P)	**(N)**
Actual (P)	True positive (TP)	False negative (FN)
(N)	False positive (FP)	True negative (TN)

The rate of positive test result is sensitivity,

$$Sensitivity = \frac{TP}{TP+FN} * 100\% \tag{3.12}$$

The ratio of a negative test result is specificity,

$$Specificity = \frac{TN}{TN+FP} * 100\% \tag{3.13}$$

The overall measure is accuracy,

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} * 100\% \tag{3.14}$$

4 SUMMARY OF PUBLICATIONS

This chapter summarizes each publication of this thesis, and answering the research problems were the primary aim of these publications. As previously addressed, there were two main variations of test datasets. The publication [I] and publication [V] addressed the existing features data for elevator fault detection, and publication [II], publication [III] and publication [IV] addressed the raw sensor data for elevator fault detection.

4.1 Publication 1: Deep Autoencoder Feature Extraction for Fault Detection of Elevator Systems

In publication [I], a generic deep autoencoder model for automated feature extraction from the elevator sensor data has been proposed for elevator fault detection. Almost one year of the data from seven traction elevators are used in this research. Deep autoencoder offers an exceptional 100% accuracy in fault detection based on new extracted deep features, which outperform the results using existing features with the conventional random forest classifier for the task of elevator fault detection. In addition, nearly 100% accuracy is achieved for avoiding false positives i.e. reducing false alarms in elevator predictive maintenance solutions.

4.2 Publication 2: Fault Detection of Elevator Systems Using Deep Autoencoder Feature Extraction

In publication [II], raw sensor data is used for calculation of new deep features with a generic deep autoencoder model. Most frequent floor patterns from the data, i.e. floor patterns which consist of the maximum number of rides between specific floor combinations are used for data extraction. In addition, almost one week of the data from one traction elevator is used in this research. Existing features were calculated from the same raw sensor dataset. Aim of the research was to compare the results from both features in terms of elevator fault detection. Fault detection accuracy is same for both features, while deep features outperform existing features in terms of avoiding false positives i.e. reducing false alarms in elevator predictive maintenance solutions.

4.3 Publication 3: Fault Detection of Elevator System Using Profile Extraction and Deep Autoencoder Feature Extraction

In publication [III], an off-line profile extraction algorithm is developed based on low-pass filtering and peak detection to extract elevator start and stop events from raw sensor data. Data is extracted from all the floor combinations available in the raw sensor data of one traction elevator. Furthermore, almost two weeks of the data is analyzed from one traction elevator in this research. The capability of profile extraction algorithm to extract elevator start and stop events from raw sensor data is combined with the capability of deep autoencoder model to calculate new deep features from extracted profiles. This combined approach outperforms state-of-the-art method by a considerable margin.

4.4 Publication 4: Elevator Fault Detection Using Profile Extraction and Deep Autoencoder Feature Extraction for Acceleration and Magnetic Signals

Research approach used in publication [III] is extended in publication [IV] for validation of the proposed method over bigger dataset. Furthermore, almost two months of the data from five traction elevators are used in this research as an extension to one elevator in publication [III]. The off-line profile extraction algorithm extracts elevator start and stop events from both acceleration and magnetic signals separately.

Profiles are combined as a vector from all five traction elevators before feature extraction. This combined approach outperforms state-of-the-art method by a considerable margin.

4.5 Publication 5: Fault Detection of Elevator Systems Using Multilayer Perceptron Neural Network

In publication [V], a generic multilayer perceptron (MLP) neural network model is proposed for automated feature extraction from the elevator sensor data for elevator fault detection. Almost one year of the data from seven traction elevators are used in this research. Multilayer perceptron offers 99% accuracy in fault detection based on new extracted deep features, which outperform the results using existing features with the conventional random forest classifier for the task of elevator fault detection. In addition, an exceptional 100% accuracy is achieved for avoiding false positives i.e. reducing false alarms in elevator predictive maintenance solutions.

5 CONCLUSIONS AND DISCUSSIONS

5.1 Conclusions

This thesis investigates the effectiveness of neural networks methods for various elevator fault detection tasks. These tasks can be identified from different types of data e.g. raw sensor data and existing feature data.

To the author's knowledge, publication [I] is the first work that proposes utilizing deep autoencoder random forest for elevator fault detection. Publication [I] 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. The 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. Almost one year of the data from seven traction elevators were used in this research, which proves the generalisation capability of this approach. The results outperform the existing features calculated from the raw sensor dataset of the same elevators.

In publication [II], aim was to investigate whether proposed deep autoencoder random forest approach in publication[I] is also suitable for elevator fault detection based on raw sensor data. The goal of the research was to develop a generic model for automated feature extraction and fault detection in the health state monitoring of elevator systems. Most frequent floor patterns from the data, i.e. floor patterns which consist of the maximum number of rides between specific floor combinations are used for data extraction. The approach provided 100% accuracy in the fault detection and in analyzing false positives for new extracted deep features. The model outperforms because of new deep features extracted from the dataset as compared to existing features calculated from the raw sensor dataset of the same elevator.

The profile extraction method proposed in publication [III] provided state-of-the-art results for elevator fault detection, beating the previous state-of-the-art by a considerable margin. The goal of this research was to develop generic models for profile extraction and automated feature extraction for fault detection in the health state monitoring of elevator systems. The approach in this research provided nearly 100% accuracy in fault detection and also in the case of analyzing false positives for all floor combinations with new extracted deep features. The models outperform because of new deep features extracted from the dataset as compared to existing features calculated from the same raw sensor dataset.

Publication [IV] focuses on the health monitoring of elevator systems using a novel fault detection technique. The goal of this research was to develop generic models for profile extraction and automated feature extraction for fault detection in the health state monitoring of elevator systems. The approach in this research provided above 90% accuracy in fault detection and in the case of analyzing false positives for all floor combinations with new extracted deep features from sensor data including both acceleration and magnetic signals. The results support the goal of this research of developing generic models which can be used in other machine systems for fault detection. The results are useful in terms of detecting false alarms in elevator predictive maintenance. Visualization of the extracted profiles and features support the goal of developing generic models for profile and feature extraction for fault detection.

In publication [V], a generic multilayer perceptron (MLP) neural network model has been proposed based on deep learning algorithm for automatic calculation of highly informative deep features from the elevator time series data and based on extracted deep features faults are detected. The approach in this research provided nearly 100% accuracy in the fault detection and also in the case of analyzing false positives for new extracted deep features. The results support the goal in this research of developing a generic model which can be used to other machine systems for automated feature extraction and fault detection. Almost one year of data from seven traction elevators have been used in this research, which proves the generalization capability of the approach. The results are useful in terms of detecting false alarms in elevator predictive maintenance.

In summary this thesis successfully answered the two research questions RP.I and RP.II with the help of publications [I] to [V] mentioned in this thesis. RP.I mainly focused on elevator fault detection problem which have been answered through the "Accuracy" evaluation parameter used in the results section of publications [I] to [V]. Deep neural networks developed in this thesis provided nearly 100% accuracy in the elevator fault detection which also answered RP.I. RP.II mainly focused on reducing number of false alarms problem in elevator predictive maintenance systems which have been answered through the "False positives" evaluation parameter used in the results section of publications [I] to [V]. Deep neural networks developed in this thesis provided nearly 100% accuracy in reducing number of false alarms which also answered RP.II.

5.2 Discussions

In publication [I] and [II], a generic deep autoencoder model has been proposed for elevator fault detection on realistic machine operations. Fault events often occur simultaneously in real-world environments. Therefore, fault events are essential to get a robust elevator fault detection system that would provide high performance in complex machine operations such as real-world environments.

Fault diagnosis methods based on deep neural networks [85], [26], [6] and convolutional neural networks [77], [31] feature extraction methodology are presented as state of the art for rotatory machines similar to elevator systems. Support vector machines [44] and extreme learning machines [79] are also used as fault detection methods for rotatory machines. However, in publication [I] an intelligent deep autoencoder random forest based feature extraction methodology has been developed for fault detection in elevator systems to improve the performance of traditional fault diagnosis methods.

The state-of-the-art results obtained with deep autoencoder random forest for elevator fault detection in publication [I] has shown that it is worth investigating deep learning methods in various ways for fault detection in elevator systems. Following this idea, in publication [II], raw sensor data is used for calculation of new deep features with a generic deep autoencoder model. Most frequent floor patterns from the data, i.e. floor patterns which consist of the maximum number of rides between specific floor combinations are used for data extraction. Existing features were calculated from the same raw sensor dataset. Aim of the research was to compare the results from both features in terms of elevator fault detection.

The methods that are proposed in publication [III] and [IV] are profile extraction algorithm for various elevator fault detection tasks based on raw acceleration signal, and both acceleration and magnetic signals profile extraction respectively. While the framework for the methods in these works is the almost similar, there are several variations between the methods mainly due to the difference between datasets used for publication [III] and [IV].

Acceleration profile extraction for health monitoring is a major issue in automated industrial applications like elevator system, computer numerical control, machinery and robotics [55]. Although rotating machine have been running for decades, but acceleration profiles extraction and processing methods are not widely available [8]. Acceleration profile extraction methods have applied in electric vehicles [3], computer numerical control systems [52] and horizontal planes [69]. Kalman filter [75] is one of the methods being used for acceleration profile extraction. However, in publication [III] an off-line profile extraction algorithm (see Section 3.4) has been developed based on low-pass filtering and peak detection to extract elevator start and stop events from raw sensor data.

The main motivation for publication [IV] is to apply the successful profile extraction method (see Section 3.4) proposed in publication [III] on the both acceleration and magnetic signals profile extraction for elevator fault detection. In publication [III], only acceleration signal have been used, which represents vibration related features. In this research, the approach have been extended to include magnetic signals, which represents position related features. This will validate the goal of this research to develop generic models for profile extraction and automated feature extraction for fault detection in the health state monitoring of elevator systems. In addition, almost two months of the data from five traction elevators have been analyzed in this research as an extension to one elevator in publication [III]. Each elevator usually produces around 200 rides per day. Each ride used in analysis contains around 5000 rows of the data, which proves robustness of the algorithms over large dataset.

Linear discriminant analysis [22], [12], artificial neural networks [61] and kalman filter [63] are used as fault detection methods for rotatory machines similar to elevator systems. However, in publication [V] an intelligent multilayer perceptron neural network model has been developed based on deep learning feature extraction methodology for fault detection in elevator systems to improve the performance of traditional fault diagnosis methods.

REFERENCES

[1] A. A. Abdulla, H. Liu, N. Stoll and K. Thurow. A robust method for elevator operation in semi-outdoor environment for mobile robot transportation system in life science laboratories. *2016 IEEE 20th Jubilee International Conference on Intelligent Engineering Systems (INES)*. IEEE. 2016, 45–50.

[2] M. Belgiu and L. Drăguţ. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing* 114 (2016), 24–31.

[3] C. Bingham, C. Walsh and S. Carroll. Impact of driving characteristics on electric vehicle energy consumption and range. *IET Intelligent Transport Systems* 6.1 (2012), 29–35.

[4] C. Bonander and U. Strömberg. Methods to handle missing values and missing individuals. *European journal of epidemiology* 34.1 (2019), 5–7.

[5] L. Breiman. Random forests. *Machine learning* 45.1 (2001), 5–32.

[6] J. Bulla, A. D. Orjuela-Cañón and O. D. Flórez. Feature extraction analysis using filter banks for faults classification in induction motors. *2018 International Joint Conference on Neural Networks (IJCNN)*. IEEE. 2018, 1–6.

[7] E. Cakir. Deep Neural Networks for Sound Event Detection. (2019).

[8] C. E. Carr, N. C. Bryan, K. N. Saboda, S. A. Bhattaru, G. Ruvkun and M. T. Zuber. Acceleration profiles and processing methods for parabolic flight. *NPJ microgravity* 4.1 (2018), 14.

[9] I. Chakraborty, D. N. Mavris, M. Emeneth and A. Schneegans. A methodology for vehicle and mission level comparison of more electric aircraft subsystem solutions: application to the flight control actuation system. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering* 229.6 (2015), 1088–1102.

[10] J.-c. CHEN and X.-m. LIU. Application of EMD and box dimension in fault diagnosis of elevator machinery [J]. *Journal of Huaqiao University (Natural Science)* 5 (2013), 494–499.

[11] J. Chen and X. Liu. AR bispectrum in fault diagnosis of elevator machinery. *2012 IEEE Fifth International Conference on Advanced Computational Intelligence (ICACI)*. IEEE. 2012, 1149–1152.

[12] L. Ciabattoni, G. Cimini, F. Ferracuti, A. Freddi, G. Ippoliti and A. Monteriu. A novel LDA-based approach for motor bearing fault detection. *2015 IEEE 13th International Conference on Industrial Informatics (INDIN)*. IEEE. 2015, 771–776.

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

[14] J. Di Martino. Dynamic time warping algorithms for isolated and connected word recognition. *New systems and architectures for automatic speech recognition and synthesis*. Springer, 1985, 405–418.

[15] T. Ebeling and M. Haul. Results of a field trial aiming at demonstrating the permanent detection of elevator wear using intelligent sensors. *Proc. ELEVCON*. 2016, 101–109.

[16] T. Ebeling. Condition monitoring for elevators–An overview. *Lift Report* 6 (2011), 25–26.

[17] E. Esteban, O. Salgado, A. Iturrospe and I. Isasa. Design methodology of a reduced-scale test bench for fault detection and diagnosis. *Mechatronics* 47 (2017), 14–23.

[18] C. Fang, X. Zhang, Y. Cheng, S. Wang, L. Zhang and Y. Yang. Fault Diagnosis for Brake System in High-Speed Trains Using the Phased Features and Multi-layer Perceptron. *IOP Conference Series: Materials Science and Engineering*. Vol. 470. 1. IOP Publishing. 2019, 012007.

[19] A. Gandomi and M. Haider. Beyond the hype: Big data concepts, methods, and analytics. *International journal of information management* 35.2 (2015), 137–144.

[20] P. L. Gentili, H. Gotoda, M. Dolnik and I. R. Epstein. Analysis and prediction of aperiodic hydrodynamic oscillatory time series by feed-forward neural networks, fuzzy logic, and a local nonlinear predictor. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 25.1 (2015), 013104.

[21] X. Glorot and Y. Bengio. Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. 2010, 249–256.

[22] R. Z. Haddad, C. A. Lopez, J. Pons-Llinares, J. Antonino-Daviu and E. G. Strangas. Outer race bearing fault detection in induction machines using stator current signals. *2015 IEEE 13th International Conference on Industrial Informatics (INDIN)*. IEEE. 2015, 801–808.

[23] T. Huynh, Y. Gao, J. Kang, L. Wang, P. Zhang, J. Lian and D. Shen. Estimating CT image from MRI data using structured random forest and auto-context model. *IEEE transactions on medical imaging* 35.1 (2016), 174.

[24] ISO. Measurement of ride quality – Part 1: Lifts (elevators). 18738.1 (2012).

[25] N. Japkowicz, S. J. Hanson and M. A. Gluck. Nonlinear autoassociation is not equivalent to PCA. *Neural computation* 12.3 (2000), 531–545.

[26] F. Jia, Y. Lei, J. Lin, X. Zhou and N. Lu. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mechanical Systems and Signal Processing* 72 (2016), 303–315.

[27] C. Jian-can. The fault diagnosis of the elevator mechanical based on higher-order spectral analysis[D]. *Journal of Huaqiao University* (2013).

[28] G. Jiang, P. Xie, H. He and J. Yan. Wind turbine fault detection using a denoising autoencoder with temporal information. *IEEE/ASME Transactions on Mechatronics* 23.1 (2018), 89–100.

[29] H. Jiang, Y. Shi and L. Qi. Design of Elevator Monitoring and Alarm System Based on WiMAX. *2015 2nd International Conference on Electrical, Computer Engineering and Electronics*. Atlantis Press. 2015.

[30] L. Li-jing. A Study on the Measurement System and Technology of Fault Diagnosis for Elevators [D]. *Tianjin: Tianjin University* (2013).

[31] L. Jing, T. Wang, M. Zhao and P. Wang. An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox. *Sensors* 17.2 (2017), 414.

[32] Z. Jinjin. The fault analysis and treatment measures for elevator car to vibration[J]. *Technology Wind* (2013).

[33] J. Kim, B. Jang, Y. Cho and Y. Park. Design and Implementation of Marine Elevator Safety Monitoring System based on Machine Learning. *International Conference on Convergence Technology*. Vol. 6. 1. 2016, 143–144.

[34] H. Kojima, K. Fukatsu and P. M. Trivailo. Mission-function control of tethered satellite/climber system. *Acta Astronautica* 106 (2015), 24–32.

[35] A. Laudani, G. M. Lozito, F. R. Fulginei and A. Salvini. On training efficiency and computational costs of a feed forward neural network: a review. *Computational intelligence and neuroscience* 2015 (2015), 83.

[36] J. Lever, M. Krzywinski and N. Altman. *Points of significance: model selection and overfitting*. 2016.

[37] F. Li, X. Pang and Z. Yang. Motor Current Signal Analysis Using Deep Neural Networks for Planetary Gear Fault Diagnosis. *Measurement* (2019).

[38] H. Li. The implementation of reinforcement learning algorithms on the elevator control system. *2015 IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA)*. IEEE. 2015, 1–4.

[39] J. Li. Research on Internet of Things Elevator Fault Detection and Automatic Alarm System Based on Android. *2018 3rd International Workshop on Materials Engineering and Computer Sciences (IWMECS 2018)*. Atlantis Press. 2018.

[40] J. Li, K. Cheng, S. Wang, F. Morstatter, R. P. Trevino, J. Tang and H. Liu. Feature selection: A data perspective. *ACM Computing Surveys (CSUR)* 50.6 (2018), 94.

[41] J. LIU and S.-b. WEI. Elevator Traction Machine Fault Detection Based on Contourlet Image Denoising. *Control Engineering of China* (2015), S1.

[42] Z. Liu, W. Cao, P.-H. Huang, G.-Y. Tian and J. L. Kirtley. Non-invasive winding fault detection for induction machines based on stray flux magnetic sensors. *2016 IEEE Power and Energy Society General Meeting (PESGM)*. IEEE. 2016, 1–6.

[43] Z. Liu, 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* 14.2 (2017), 257–261.

[44] D. Martınez-Rego, O. Fontenla-Romero and A. Alonso-Betanzos. Power wind mill fault detection via one-class ν-SVM vibration signal analysis. *Neural Networks (IJCNN), The 2011 International Joint Conference on*. IEEE. 2011, 511–518.

[45] K. M. Mishra and K. Huhtala. Elevator Fault detection using profile extraction and deep autoencoder feature extraction for acceleration and magnetic signals. *Applied Sciences* 9 (2019), 15p.

[46] K. M. Mishra and K. Huhtala. Fault detection of elevator systems using multilayer perceptron neural network. *in Proceedings of the 24th IEEE Conference on Emerging Technologies and Factory Automation (ETFA)* 24 (2019), 904–909.

[47] K. M. Mishra, T. Krogerus and K. Huhtala. Deep autoencoder feature extraction for fault detection of elevator systems. *in Proceedings of the 27th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)* 27 (2019), 191–196.

[48] K. M. Mishra, T. Krogerus and K. Huhtala. Fault detection of elevator systems using deep autoencoder feature extraction. *in Proceedings of the IEEE 13th International Conference on Research Challenges in Information Science (RCIS)* 13 (2019), 43–48.

[49] K. M. Mishra, J. Saxen, J. Bjorkqvist and K. Huhtala. Fault detection of elevator system using profile extraction and deep autoencoder feature extraction. *in Proceedings of the 33rd annual European Simulation and Modelling Conference (ESM)* 33 (2019).

[50] R. P. Monteiro, M. Cerrada, D. R. Cabrera, R. V. Sánchez and C. J. Bastos-Filho. Using a Support Vector Machine Based Decision Stage to Improve the Fault Diagnosis on Gearboxes. *Computational intelligence and neuroscience* 2019 (2019).

[51] J. P. Muñoz, B. Li, X. Rong, J. Xiao, Y. Tian and A. Arditi. An assistive indoor navigation system for the visually impaired in multi-floor environments. *2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*. IEEE. 2017, 7–12.

[52] S.-H. Nam and M.-Y. Yang. A study on a generalized parametric interpolator with real-time jerk-limited acceleration. *Computer-Aided Design* 36.1 (2004), 27–36.

[53] G. Niu, S.-S. Lee, B.-S. Yang and S.-J. Lee. Decision fusion system for fault diagnosis of elevator traction machine. *Journal of Mechanical Science and Technology* 22.1 (2008), 85–95.

[54] I. O. Olalere and M. Dewa. Early fault detection of elevators using remote condition monitoring through IoT technology. *South African Journal of Industrial Engineering* 29.4 (2018), 17–32.

[55] R. A. Osornio-Rios, R. de Jesús Romero-Troncoso, G. Herrera-Ruiz and R. Castañeda-Miranda. FPGA implementation of higher degree polynomial acceleration profiles for peak jerk reduction in servomotors. *Robotics and Computer-Integrated Manufacturing* 25.2 (2009), 379–392.

[56] S. E. Pandarakone, M. Masuko, Y. Mizuno and H. Nakamura. Deep Neural Network Based Bearing Fault Diagnosis of Induction Motor Using Fast Fourier Transform Analysis. *2018 IEEE Energy Conversion Congress and Exposition (ECCE)*. IEEE. 2018, 3214–3221.

[57] J. Pang, D. Liu, Y. Peng and X. Peng. Intelligent pattern analysis and anomaly detection of satellite telemetry series with improved time series representation. *Journal of Intelligent & Fuzzy Systems* 34.6 (2018), 3785–3798.

[58] D. Pelvig, H. Pakkenberg, A. Stark and B. Pakkenberg. Neocortical glial cell numbers in human brains. *Neurobiology of aging* 29.11 (2008), 1754–1762.

[59] O. R. Peña and M. J. Leamy. An efficient architecture for energy recovery in hydraulic elevators. *International Journal of Fluid Power* 16.2 (2015), 83–98.

[60] F. Qifeng, C. Guoqing and S. Zibo. Application of wavelet de-noising method in vibration signal analysis of elevator car. *2016 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*. IEEE. 2016, 610–614.

[61] R. Rahimilarki and Z. Gao. Grey-box model identification and fault detection of wind turbines using artificial neural networks. *2018 IEEE 16th International Conference on Industrial Informatics (INDIN)*. IEEE. 2018, 647–652.

[62] M. Rhudy. Time alignment techniques for experimental sensor data. *Int. J. Comput. Sci. Eng. Survey* 5.2 (2014), 1–14.

[63] G. Rigatos, N. Zervos, D. Serpanos, V. Siadimas, P. Siano and M. Abbaszadeh. Condition monitoring of wind-power units using the derivative-free nonlinear kalman filter. *2018 IEEE 16th International Conference on Industrial Informatics (INDIN).* IEEE. 2018, 472–477.

[64] H. Sahin and A. Subasi. Classification of the cardiotocogram data for anticipation of fetal risks using machine learning techniques. *Applied Soft Computing* 33 (2015), 231–238.

[65] Z. Sahwee, A. S. Mahmood, N. A. Rahman and K. S. M. Sahari. Flight testing of control allocation for elevator actuator fault in small UAV. *2016 International Conference on Control, Automation and Information Sciences (ICCAIS).* IEEE. 2016, 232–237.

[66] I. Skog, I. Karagiannis, A. B. Bergsten, J. Härdén, L. Gustafsson and P. Händel. A smart sensor node for the internet-of-elevators—non-invasive condition and fault monitoring. *IEEE Sensors Journal* 17.16 (2017), 5198–5208.

[67] A. So, L. Al-Sharif and A. Hammoudeh. Traffic analysis of a simplified two-dimensional elevator system. *Building Services Engineering Research and Technology* 36.5 (2015), 567–579.

[68] H. Song, J. J. Thiagarajan, P. Sattigeri, K. N. Ramamurthy and A. Spanias. A deep learning approach to multiple kernel fusion. *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* IEEE. 2017, 2292–2296.

[69] F. Soyka, P. R. Giordano, K. Beykirch and H. H. Bülthoff. Predicting direction detection thresholds for arbitrary translational acceleration profiles in the horizontal plane. *Experimental brain research* 209.1 (2011), 95–107.

[70] S.-Z. Su, Z.-H. Liu, S.-P. Xu, S.-Z. Li and R. Ji. Sparse auto-encoder based feature learning for human body detection in depth image. *Signal Processing* 112 (2015), 43–52.

[71] R. Takase, N. Yoshikawa and S. Suzuki. Combined Fault Detection, Isolation, and Control: Propulsion Controlled Aircraft in Case of Elevator Failure. *2018 IEEE Conference on Control Technology and Applications (CCTA).* IEEE. 2018, 754–759.

[72] P. Vincent, H. Larochelle, Y. Bengio and P.-A. Manzagol. Extracting and composing robust features with denoising autoencoders. *Proceedings of the 25th international conference on Machine learning*. ACM. 2008, 1096–1103.

[73] J. Wang, S. Li, Z. An, X. Jiang, W. Qian and S. Ji. Batch-normalized deep neural networks for achieving fast intelligent fault diagnosis of machines. *Neurocomputing* 329 (2019), 53–65.

[74] P. Wang, W. He and W. Yan. Fault diagnosis of elevator braking system based on wavelet packet algorithm and fuzzy neural network. *Electronic Measurement & Instruments, 2009. ICEMI'09. 9th International Conference on*. IEEE. 2009, 4–1028.

[75] X. Wang, S. Wu and K. Wang. A run-to-run profile control algorithm for improving the flatness of nano-scale products. *IEEE Transactions on Automation Science and Engineering* 12.1 (2015), 192–203.

[76] X. Wang, S. Wang, Z. Yang and C. Zhang. Active fault-tolerant control strategy of large civil aircraft under elevator failures. *Chinese Journal of Aeronautics* 28.6 (2015), 1658–1666.

[77] 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* 23.1 (2018), 101–110.

[78] O. Yaman, M. Baygin and M. Karakose. A new approach based on image processing for detection of wear of guide-rail surface in elevator systems. *International Journal of Applied Mathematics, Electronics and Computers* 4 (2016), 296–300.

[79] Z.-X. Yang and P.-B. Zhang. ELM Meets RAE-ELM: A hybrid intelligent model for multiple fault diagnosis and remaining useful life predication of rotating machinery. *Neural Networks (IJCNN), 2016 International Joint Conference on*. IEEE. 2016, 2321–2328.

[80] J. T. Yoon, B. D. Youn, M. Yoo, Y. Kim and S. Kim. Life-cycle maintenance cost analysis framework considering time-dependent false and missed alarms for fault diagnosis. *Reliability Engineering & System Safety* 184 (2019), 181–192.

[81] J. Zabalza, J. Ren, J. Zheng, H. Zhao, C. Qing, Z. Yang, P. Du and S. Marshall. Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging. *Neurocomputing* 185 (2016), 1–10.

[82] V. Zarikas and N. Tursynbek. Intelligent Elevators in a Smart Building. *FTC 2017-Future Technologies Conference 2017*. 2017.

[83] D. Zhang, L. Qian, B. Mao, C. Huang, B. Huang and Y. Si. A data-driven design for fault detection of wind turbines using random forests and XGboost. *IEEE Access* 6 (2018), 21020–21031.

[84] G. Zhang, S. Huang and Y. Yuan. The study of elevator fault diagnosis based on multi-agent system. *2009 International Conference on Computational Intelligence and Software Engineering*. IEEE. 2009, 1–5.

[85] 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* 17.3 (2017), 549.

[86] T. Zhang, D. Wang, J. Cao, Y. Q. Ni, L.-J. Chen and D. Chen. Elevator-assisted sensor data collection for structural health monitoring. *IEEE Transactions on Mobile Computing* 11.10 (2011), 1555–1568.

[87] X. Zhang and D. Liang. Construction of Elevator Inspection Quality Evaluation System Based on Big Data. *2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC)*. IEEE. 2019, 238–242.

[88] C.-G. Zhao, H.-Y. Xu and L. Jia. Research of elevator fault diagnosis based on decision tree and rough set. *2012 International Conference on Computer Science and Information Processing (CSIP)*. IEEE. 2012, 1318–1322.

PUBLICATIONS

PUBLICATION

I

Deep autoencoder feature extraction for fault detection of elevator systems

K. M. Mishra, T. Krogerus and K. Huhtala

in Proceedings of the 27th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN) 27.(2019), 191–196

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.

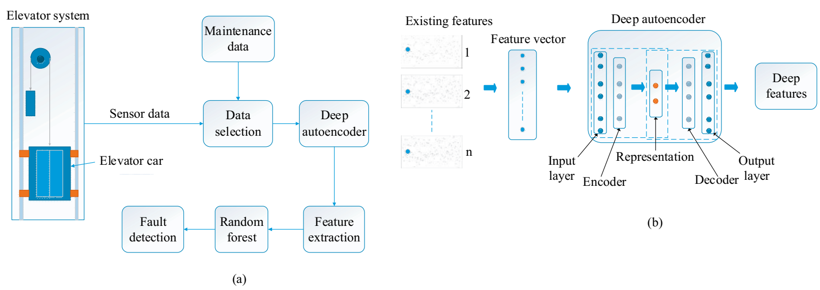


Fig. 1: (a) Fault detection approach (b) Deep autoencoder feature extraction approach.

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 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]:

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

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

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

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

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

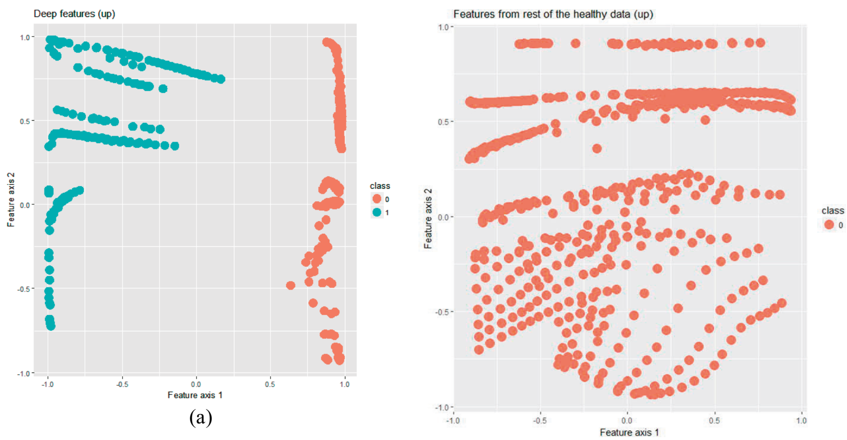
ESANN 2019 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2019, i6doc.com publ., ISBN 978-287-587-065-0. Available from http://www.i6doc.com/en/.

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)

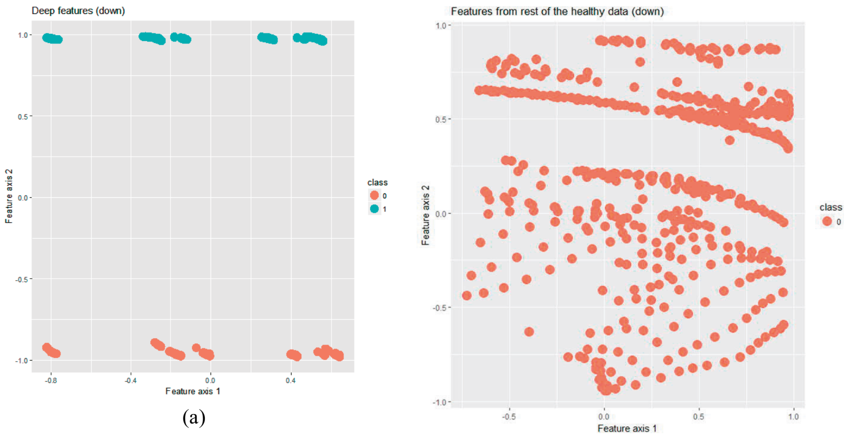
(b)

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

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/.

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.

PUBLICATION

II

Fault detection of elevator systems using deep autoencoder feature extraction

K. M. Mishra, T. Krogerus and K. Huhtala

in Proceedings of the IEEE 13th International Conference on Research Challenges in Information Science (RCIS) 13.(2019), 43–48

Fault Detection of Elevator Systems Using Deep Autoencoder Feature Extraction

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

Unit of Automation Technology and Mechanical Engineering
Tampere University
Tampere, Finland
E-mail: (krishna.mishra, tomi.krogerus, kalevi.huhtala)@tuni.fi

***Abstract*—In this research, we propose a generic deep autoencoder model for automated feature extraction from the raw sensor data. Extracted deep features are classified with random forest algorithm for fault detection. Sensor data are labelled as healthy and faulty based on the maintenance actions recorded. The remaining healthy data is used for validation of the model to prove its efficacy in terms of avoiding false positives. We have achieved 100% accuracy in fault detection along with avoiding false positives based on new extracted deep features, which outperform the results using existing features. Existing features are also classified with random forest to compare results. Deep autoencoder random forest provides better results due to the new deep features extracted from the dataset when compared to existing features. Our model provides good classification and is robust against overfitting characteristics. This research will help various predictive maintenance systems to detect false alarms, which will reduce unnecessary visits of service technicians to installation sites.**

***Index Terms*—Elevator System, Deep Autoencoder, Fault Detection, Feature Extraction, Random Forest**

I. Introduction

In recent years, elevator systems have been used more and more extensively in apartments, commercial facilities and office buildings. Nowadays 54% of the worlds population lives 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 of the system [3], [4]. Elevator production and service companies are now opting for a predictive maintenance policy to provide better service to customers. They are remotely monitoring faults in elevators and estimating the remaining lifetime of the components responsible for faults. Elevator systems require fault detection and diagnosis for healthy operation [5].

Fault diagnosis methods based on deep neural networks [6], [7], [8] and convolutional neural networks [9], [10] feature extraction methodology are presented as state of the art for rotatory machines similar to elevator systems. Linear discriminant analysis [11], [12], artificial neural networks [13] and kalman filter [14] are also used as fault detection methods for rotatory machines. However, we have developed an intelligent deep autoencoder random forest 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 [15] 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 [16]. Deep learning algorithms [17] are frequently used in areas such as natural language processing, signal processing, speech recognition, computer vision and image classication. Deep learning algorithms have also been used in various applications e.g. biotechnology [18], sentiment analysis [19], survival analysis [20] and information systems [21]. Autoencoding is a process for nonlinear dimension reduction with natural transformation architecture using feedforward neural network [22]. Autoencoders have proven powerful as nonlinear feature extractors [23]. 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 [24]. Autoencoders were first introduced by LeCun [25], and have been studied for decades. Traditionally, feature learning and dimensionality reduction are the two main features of autoencoders. Recently, autoencoders have been considered one of the most compelling subspace analysis techniques because of the existing theoretical relations between autoencoders and latent variable models [26]. Autoencoders have been used for feature extraction from the data in systems like industrial processes [27], induction motor [28] and wind turbines [29], different from elevator systems as in our research.

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 raw sensor data, and based on these deep features, faults are detected. The rest of this paper is organized as follows. Section II presents the methodology of the paper including deep autoencoder and random forest algorithms. Then, section III includes the details of experiments performed, results and discussion. Finally, section IV concludes the paper and presents the future work.

II. Methodology

We have developed an automatic feature extraction technique in this research as an extension to the work of our previous research [30] to compare the results using new extracted deep features. Fig. 1 shows the fault detection approach used in this paper, which includes raw sensor data extracted based on time periods provided by the maintenance data. Data collected from an elevator system is fed to the deep autoencoder model for feature extraction and then random forest performs the fault detection task based on extracted deep features.

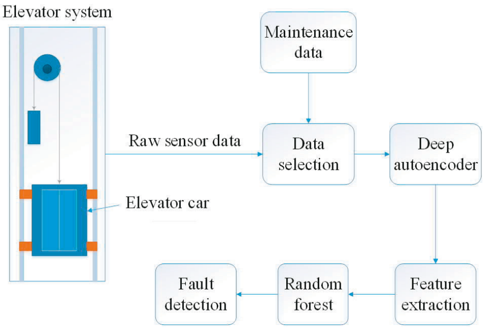


Fig. 1. Fault detection approach

A. Deep Autoencoder

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 output layer and reproduces its inputs as its output. We are using a five layer deep autoencoder (see Fig. 2) including input, output, encoder, decoder and representation layers, which is a different approach than in [29], [31]. In our approach, we first analyze the data to find the most frequent floor pattern and then feed the segmented raw sensor data windows in up and down directions separately to the deep autoencoder model to extract new deep features from the raw data. Lastly, we apply random forest as a classifier for fault detection based on new deep features extracted from the data.

The encoder transforms the input x into corrupted input data $x^{'}$ using hidden representation h through nonlinear mapping

$$h = f(W_1 x^{'} + b) \tag{1}$$

where *f(.)* 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 [31]. 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

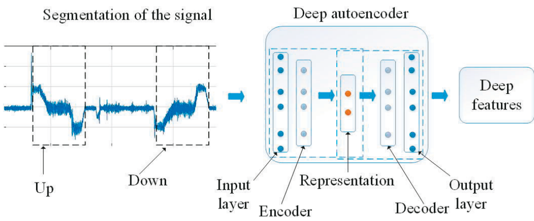


Fig. 2. Deep autoencoder feature extraction approach

$$x^{''} = g(W_2 h + c) \tag{2}$$

where *g(.)* 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 [32].

B. Random Forest

Random forest includes an additional layer of randomness to bagging. It uses different bootstrap samples of the data for constructing each tree [33]. 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 [34]. The final classification accuracy of random forest is calculated by averaging, i.e. arithmetic mean of the probabilities of assigning classes related to all the produced trees *(t)*. Testing data *(d)* that is unknown to all the decision trees is used for evaluation by the voting method (see Fig. 3).

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 [35]:

$$\mu = \frac{\sum_{t \in T} \sum_{l \in [1,...,L_t]} v_l{}^t}{\sum_{t \in T} L_t} \tag{3}$$

All classification trees providing a final decision by voting method are given by [36]:

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

where $j = 1,2,...,C$ and 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 $I(.)$ defined as:

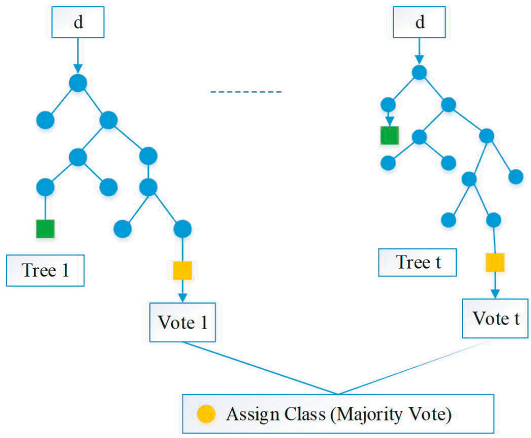


Fig. 3. Classification phase of random forest classifier

$$I(x) = \begin{cases} 1, & h_i(a) = y_j \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

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

C. Evaluation Parameters

Evaluation parameters used in this research are defined with the confusion matrix in Table I.

TABLE I
CONFUSION MATRIX

	Predicted (P)	**(N)**
Actual (P)	True positive (TP)	False negative (FN)
(N)	False positive (FP)	True negative (TN)

The rate of positive test result is sensitivity,

$$Sensitivity = \frac{TP}{TP + FN} * 100\% \quad (6)$$

The ratio of a negative test result is specificity,

$$Specificity = \frac{TN}{TN + FP} * 100\% \quad (7)$$

The overall measure is accuracy,

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} * 100\% \quad (8)$$

III. RESULTS AND DISCUSSION

In this research, first, we selected the most frequent floor patterns from the data, i.e. floor patterns which consist of the maximum number of rides between specific floor combinations. The next step includes the selection of faulty data from the most frequent floor patterns based on time periods provided by the maintenance data. An equal amount of healthy data is also selected and labelled as class 0 for healthy, with class 1 for faulty data. Only the vertical component of acceleration data is selected in this research because it is the most informative aspect consisting of significant changes in vibration levels as compared to other components. A combined version of healthy and faulty data is used for feature extraction with the deep autoencoder model.

A. Up Movement

We have analyzed up and down movements separately because the traction based elevator usually produces slightly different levels of vibration in each direction. First, we have selected the floor patterns 0 to 6 and faulty data based on time periods provided by the maintenance data as shown in Fig. 4. Each ride used in analysis contains around 5000 rows of the data, which proves robustness of the algorithm over large dataset.

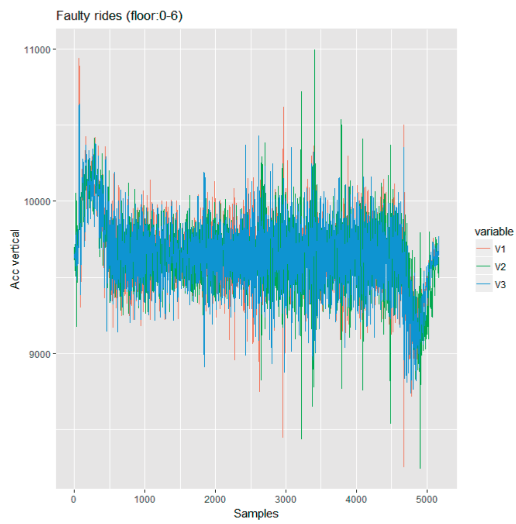


Fig. 4. Rides from faulty data

Then, we have selected an equal number of rides for healthy data, as shown in Fig. 5. The next step is to label both the healthy and faulty data with class labels 0 and 1 respectively. Healthy and faulty data with class labels are fed to the deep autoencoder model and the generated deep features are shown in Fig. 6. These are called as deep features or latent features in deep autoencoder terminology, which shows hidden representations of the data.

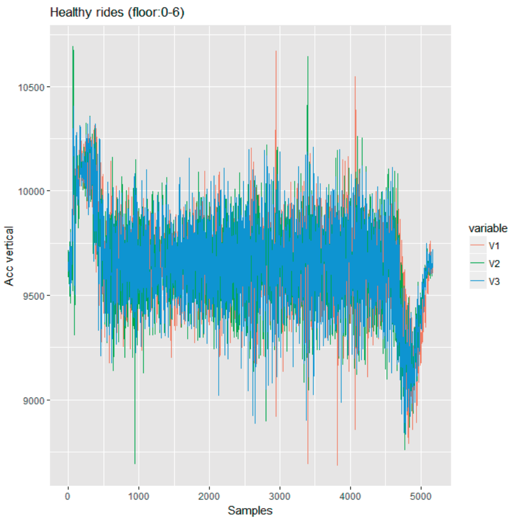


Fig. 5. Rides from healthy data

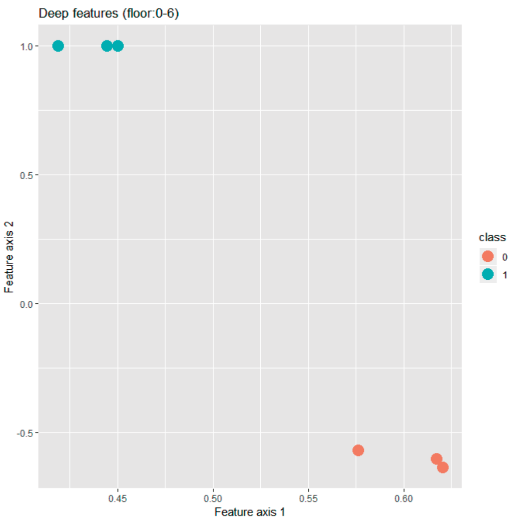


Fig. 6. Extracted deep autoencoder features. (Visualization of the features w.r.t class variable)

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 II. 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 similar as Fig. 5 to analyze 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. 7. 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.

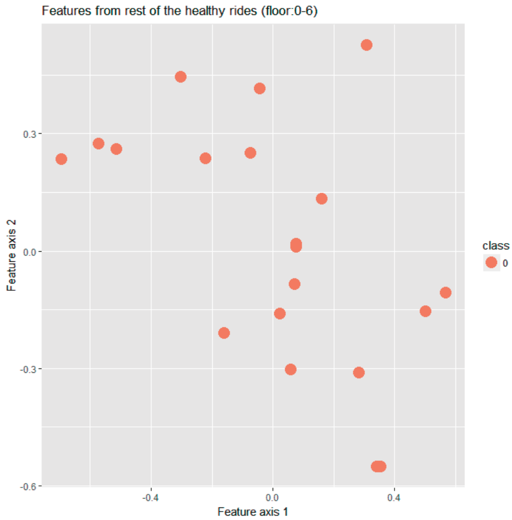


Fig. 7. Extracted deep features (only healthy rides)

Table II presents the results for upward movement of the elevator in terms of accuracy, sensitivity and specificity. We have also included the accuracy of avoiding false positives as evaluation parameters for this research. The results show that the new deep features provide better accuracy in terms of avoiding false positives from the data, which is helpful in detecting false alarms for elevator predictive maintenance strategies. It is extremely helpful in reducing the unnecessary visits of maintenance personnel to installation sites.

TABLE II
FAULT DETECTION ANALYSIS (FALSE POSITIVES FIELD RELATED TO ANALYZING REST OF THE HEALTHY DATA AFTER THE TRAINING AND TESTING PHASE)

	Deep features	**Existing features**
Accuracy	1	1
Sensitivity	1	1
Specificity	1	1
False positives	1	0.90

B. Down Movement

For downward motion, we have repeated the same analysis procedure as in the case of upward motion. We feed both healthy and faulty data with class labels to the deep autoencoder model for the extraction of new deep features, as shown in Fig. 8.

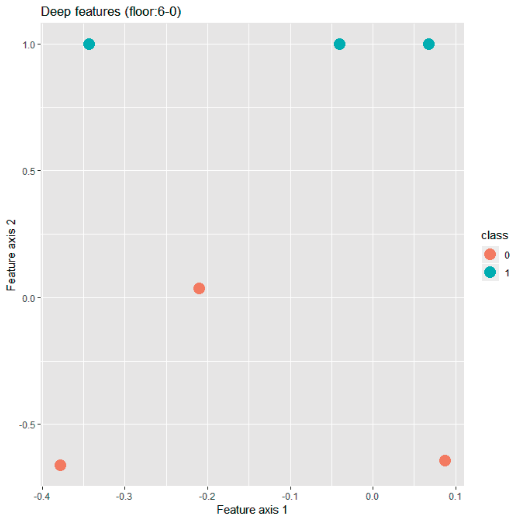


Fig. 8. Extracted deep features. (Visualization of the features w.r.t class variable).

Finally, the new extracted deep features are classified with random forest model, and the results are shown in Table III. After this, the rest of the healthy data with class label 0 is used to analyze the number of false positives. The extracted deep features are presented in Fig. 9. Table III presents the results for fault detection with deep autoencoder random forest model in the downward direction. The results are similar to the upward direction, but we can see significant change in terms of accuracy when analyzing the number of false positives with new deep features.

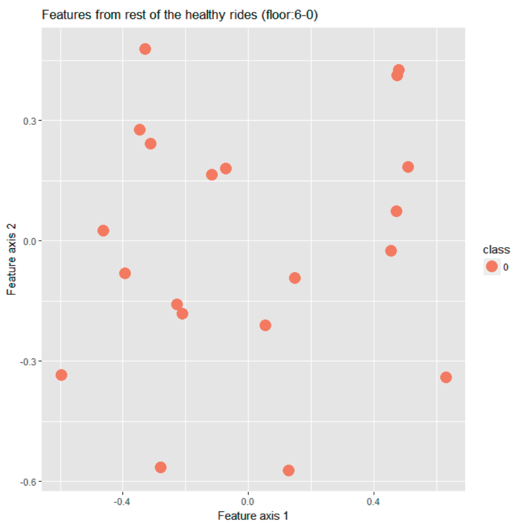


Fig. 9. Extracted deep features (only healthy rides)

TABLE III
FAULT DETECTION ANALYSIS

	Deep features	Existing features
Accuracy	1	1
Sensitivity	1	1
Specificity	1	1
False positives	1	0.55

IV. CONCLUSIONS AND FUTURE WORK

This research focuses on the health monitoring of elevator systems using a novel fault detection technique. The goal of the research was to develop a generic model for automated feature extraction and fault detection in the health state monitoring of elevator systems. Our approach provided 100% accuracy in the fault detection and in analyzing false positives for new extracted deep features. The results support our goal in this research of developing a generic model which can be used to other machine systems for automated feature extraction and fault detection. 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. Our model outperforms because of new deep features extracted from the dataset as compared to existing features calculated from the raw sensor dataset of the same elevator. 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.

In future work, we will extend our approach on more elevators including multiple floor patterns and real-world big data cases to validate its potential for other applications and improve its efficacy.

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, 8595, 2008. doi:10.1007/s12206-007-1010-0.
[3] T. Ebeling, "Condition monitoring for elevators - An overview," Lift-Report, vol. 6, pp. 25, 2011.

[4] T. Ebeling and M. Haul, ”Results of a field trial aiming at demonstrating the permanent detection of elevator wear using intelligent sensors,” in Proc. ELEVCON, Madrid, Spain, pp. 101109, May 2016.
[5] P. Wang, W. He and W. Yan, ”Fault diagnosis of elevator braking system based on wavelet packet algorithm and fuzzy neural network,” 9th International Conference on Electronic Measurement Instruments, ICEMI ’09, 10281031, 2009.
[6] 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.
[7] F. Jia, Y. Lei, J. Lin, X. Zhou, and N. Lu, ”Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data,” Mechanical Systems and Signal Processing, vol. 72, pp. 303-315, 2016.
[8] J. Bulla, A. D. Orjuela-Can and O. D. Flrez, ”Feature extraction analysis using filter banks for faults classification in induction motors,” in 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1-6, 2018.
[9] 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.
[10] L. Jing, T. Wang, M. Zhao and P. Wang, ”An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox,” Sensors, vol. 17, (2), pp. 414, 2017.
[11] R. Z. Haddad, C. A. Lopez, J. Pons-Llinares, J. Antonino-Daviu, and E. G. Strangas, ”Outer race bearing fault detection in induction machines using stator current signals,” in Industrial Informatics (INDIN), 2015 IEEE 13th International Conference on, 2015, pp. 801-808.
[12] L. Ciabattoni, G. Cimini, F. Ferracuti, A. Freddi, G. Ippoliti, and A. Monteriu, ”A novel LDA-based approach for motor bearing fault detection,” in Industrial Informatics (INDIN), 2015 IEEE 13th International Conference on, 2015, pp. 771-776.
[13] R. Rahimilarki and Z. Gao, ”Grey-box model identification and fault detection of wind turbines using artificial neural networks,” in 2018 IEEE 16th International Conference on Industrial Informatics (INDIN), 2018, pp. 647-652.
[14] G. Rigatos, N. Zervos, D. Serpanos, V. Siadimas, P. Siano, and M. Abbaszadeh, ”Condition monitoring of wind-power units using the derivative-free nonlinear kalman filter,” in 2018 IEEE 16th International Conference on Industrial Informatics (INDIN), 2018, pp. 472-477.
[15] 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.
[16] I. Fernndez-Varela, D. Athanasakis, S. Parsons, E. Hernndez-Pereira, and V. Moret-Bonillo, ”Sleep staging with deep learning: A convolutional model,” in Proceedings of the ESANN 2018.
[17] D. Xie, L. Zhang and L. Bai, ”Deep learning in visual computing and signal processing,” Applied Computational Intelligence and Soft Computing, vol. 2017, 2017.
[18] A. L. Palacio and . P. Lpez, ”From big data to smart data: A genomic information systems perspective,” in 2018 12th International Conference on Research Challenges in Information Science (RCIS), 2018, pp. 1-11.
[19] M. ElAbdi, B. Smine and S. B. Yahia, ”DFBICA: A new distributed approach for sentiment analysis of bibliographic citations,” in 2018 12th International Conference on Research Challenges in Information Science (RCIS), 2018, pp. 1-6.
[20] J. O’Donoghue, M. Roantree and A. McCarren, ”Variable interactions in risk factors for dementia,” in Research Challenges in Information Science (RCIS), 2016 IEEE Tenth International Conference on, 2016, pp. 1-10.
[21] A. Eberhard, K. Tilvawala, G. Peko, and D. Sundaram, ”Engaging a class of 2200 digital natives a blended approach,” in Research Challenges in Information Science (RCIS), 2014 IEEE Eighth International Conference on, 2014, pp. 1-7.
[22] J. Hnninen and T. Krkkinen, ”Comparison of four-and six-layered configurations for deep network pretraining,” in Proceedings of the ESANN 2016.
[23] A. Albuquerque, T. Amador, R. Ferreira, A. Veloso, and N. Ziviani, ”Learning to rank with deep autoencoder features,” in 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1-8, 2018.
[24] R. Huet, N. Courty and S. Lefvre, ”A new penalisation term for image retrieval in clique neural networks,” in Proceedings of the ESANN 2016.
[25] F. Fogelman-Soulie, Y. Robert and M. Tchuente, ”Automata networks in computer science: theory and applications.” Manchester University Press and Princeton University Press, 1987.
[26] 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, pp. 1, 2018.
[27] J. Fan, W. Wang and H. Zhang, ”AutoEncoder based high-dimensional data fault detection system,” in Industrial Informatics (INDIN), 2017 IEEE 15th International Conference on, 2017, pp. 1001-1006.
[28] W. Sun, S. Shao, R. Zhao, R. Yan, X. Zhang, and X. Chen, ”A sparse auto-encoder-based deep neural network approach for induction motor faults classication,” Measurement, vol. 89, pp. 171-178, 2016.
[29] 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. 89100, 2018.
[30] K. M. Mishra, T. Krogerus and K. Huhtala, ”Deep autoencoder feature extraction for fault detection of elevator systems,” in Proceedings of the 27th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN), pp. 191-196, 2019.
[31] P. Vincent, H. Larochelle, Y. Bengio and P.A Manzagol, ”Extracting and composing robust features with denoising autoencoders,” in Proceedings of the 25th International Conference on Machine Learning, pp. 1096-1103, 2008.
[32] 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.
[33] L. Breiman, ”Random forests”, Machine Learning, vol. 45, no. 1, pages 5-32. 2001.
[34] A. Liaw and M. Wiener, ”Classification and regression by random forest,” R News, vol. 2, (3), pp. 18-22, 2002.
[35] T. Huynh, Y. Gao, J. Kang, L. Wang, P. Zhang, J. Lian and D. Shen, Senior Member, IEEE, and Alzheimers 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.
[36] Z. Liu, Member, IEEE, B. Tang, X. He, Q. Qiu, and F. Liu, ”Class-specific random forest with cross-correlation constraints for spectralspatial hyperspectral image classification,” IEEE Geoscience and Remote Sensing Letters, vol. 14, (2), pp. 257-261, 2017.

PUBLICATION

III

Fault detection of elevator system using profile extraction and deep autoencoder feature extraction

K. M. Mishra, J. Saxen, J. Bjorkqvist and K. Huhtala

in Proceedings of the 33rd annual European Simulation and Modelling Conference (ESM)
33.(2019)

FAULT DETECTION OF ELEVATOR SYSTEM USING PROFILE EXTRACTION AND DEEP AUTOENCODER FEATURE EXTRACTION

Krishna Mohan Mishra[1], John-Eric Saxen[2], Jerker Bjorkqvist[2], and Kalevi J. Huhtala[1]
[1]Unit of Automation Technology and Mechanical Engineering, Tampere University, Tampere, Finland
[2]Department of Information Sciences, Abo Akademi University, Turku, Finland
email: {krishna.mishra, kalevi.huhtala}@tuni.fi
email: {John-Eric.Saxen, jerker.bjorkqvist}@abo.fi

KEYWORDS

Elevator system, deep autoencoder, fault detection, feature extraction, random forest, profile extraction

ABSTRACT

In this paper, we propose a new algorithm for data extraction from time series data, and furthermore automatic calculation of highly informative deep features to be used in fault detection. In data extraction elevator start and stop events are extracted from sensor data, and a generic deep autoencoder model is also developed for automated feature extraction from the extracted profiles. After this, extracted deep features are classified with random forest algorithm for fault detection. Sensor data are labelled as healthy and faulty based on the maintenance actions recorded. The rest of the healthy data are 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 outperforms results using existing features. Existing features are also classified with random forest to compare results. Our developed algorithm provides better results due to the new deep features extracted from the dataset when compared to existing features. This research will help various predictive maintenance systems to detect false alarms, which will in turn reduce unnecessary visits of service technicians to installation sites.

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 lives in urban areas (Desa 2014). Therefore, elevator systems need proper maintenance and safety. Fault diagnosis methods based on deep neural networks (Jia et al. 2016) and convolutional neural networks (Xia et al. 2018) feature extraction methodology are presented as state of the art for rotatory machines similar to elevator systems. Support vector machines (Martínez-Rego et al. 2011) and extreme learning machines (Yang and Zhang 2016) are also used as fault detection methods for rotatory machines. However, we have developed an intelligent deep autoencoder random forest based feature extraction methodology for fault detection in elevator systems to improve the performance of traditional fault diagnosis methods.

Acceleration profile extraction methods have applied in electric vehicles (Bingham et al. 2012) and horizontal planes (Soyka et al. 2011). Kalman filter (Wang et al. 2015) is one of the methods being used for acceleration profile extraction. However, we have developed an offline profile extraction algorithm based on low-pass filtering and peak detection to extract elevator start and stop events from sensor data.

Autoencoders were first introduced by LeCun (Fogelman-Soulie et al. 1987), and have been studied for decades. Traditionally, feature learning and dimensionality reduction are the two main features of autoencoders. Recently, autoencoders have been considered one of the most compelling subspace analysis techniques because of the existing theoretical relations between autoencoders and latent variable models. Autoencoders have been used for feature extraction from the data in systems like induction motor (Sun et al. 2016) and wind turbines (Jiang et al. 2018) for fault detection, different from elevator systems as in our research.

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 are 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 an algorithm for profile extraction from the raw sensor data rides and a generic algorithm with deep autoencoder random forest approach for automated feature extraction from raw sensor data profiles for fault detection in elevator systems. The rest of this paper is organized as follows. The next section presents the methodology of the paper including profile extraction, deep autoencoder and random forest algorithms. This is followed by section that includes the details of experiments performed, results and discussion. Finally, the last section concludes the paper and presents the future work.

METHODOLOGY

In this study, we have utilised 12 different existing features derived from raw sensor data describing the motion and vibration of an elevator for fault detection and diagnostics of multiple faults. We have developed an automated feature extraction technique for raw sensor data in this research as an extension to the work of our previous research (Mishra et al. 2019) to compare the results using new extracted deep features. We only extract start and stop profiles from the rides because of the different lengths of rides for each floor combination due to the constant speed phase, which is longer when there is longer travel.

Profile extraction algorithm

The algorithm works in two stages. In the first stage, the signal is pre-processed and normalized, followed by low-pass filtering in order to reduce noise spikes. The low-pass filtered signal is used for peak detection, which for each elevator travel detects a local minimum and maximum, corresponding to acceleration and deceleration (start and stop) events.
In the second stage, alignment and collection of equal length profiles is performed based on windowing of the acceleration signal near the peak events. In this stage, the raw acceleration signal is used instead of the filtered signal. A number of time domain alignment methods have been proposed in the literature. Dynamic time warping (DTW) has been commonly applied, e.g. in speech recognition (Di Martino 1985), whereas various alignment techniques for sensor data have been presented in (Rhudy 2014). Here, alignment is performed against a reference profile, which is initialized to the known approximate length of the acceleration and deceleration windows. The reference profile is aligned against the raw data in the window of the detected peaks. The criterion for optimal alignment was defined as the alignment that minimizes the sum of the Euclidean or L_2 norm. The output from this operation is an $n \times m$ matrix of aligned profiles describing n acceleration and deceleration events of length m.
In order to improve the alignment accuracy, the reference profile is updated iteratively following each run. Each sequence in the profile matrix is of the same sample size and closely synchronized in time and can hence be considered a repetition of the same signal. Using signal averaging, the new reference profile is calculated as the mean of the n extracted profiles. This both maintains the main characteristics of the signal and reduces the noise. Assuming white noise and perfect synchronization, signal averaging improves the signal-to-noise ratio *(SNR)* by a factor of $\sqrt{n}$. The reference profile is updated on-line during the alignment stage or in batch mode by multiple iterations through the same dataset. The off-line profile extraction algorithm is described as follows.

Off-line profile extraction algorithm
Pre-procession
1. Read a vector of raw acceleration data containing k elevator travels. Define the zero mean transformed dataset as X.
2. Perform low-pass filtering on X and obtain denoised dataset Y.
Initialization
3. Define parameters for reference profile. Set window length to m samples and height h to the 99th percentile of the low-pass filtered dataset.
4. Set threshold limit t for triggering peak detection as a fraction of h.
5. Define alignment window size a and set k=1.
Iteration
6. From $Y(k)$, detect peak acceleration points y_{min} and y_{max} satisfying $abs(y_{min,max}) \geq t$
7. Align reference profile R against raw dataset X in the vicinity of detected peaks by minimizing the L_2 norm according to

$$min \sum_{i=-a/2}^{a/2} \sum_{j=1}^{m} [-r_{\mathrm{j}} - x_{\mathrm{min+i+j}}]^2 \tag{1}$$

$$min \sum_{i=-a/2}^{a/2} \sum_{j=1}^{m} [r_{\mathrm{j}} - x_{\mathrm{max+i+j}}]^2 \tag{2}$$

8. Add aligned data points from $X(k)$ as rows into an $n \times m$ profile matrix, alternatively separate matrices according to direction of travel (min/max).
9. Set travel window k=k+1 and repeat steps 6-8 until end of dataset.
10. Update reference profile with the signal-averaged profile obtained from the column-wise mean of the new profile matrix. Set k=1 and continue with new batch iterations by repeating steps 6-9.

Deep autoencoder

We are using a five layer deep autoencoder (see Figure 1) including input, output, encoder, decoder and representation layers, which is a different approach than in (Jiang et al. 2018), (Vincent et al. 2008). In our approach, we first analyze the data to find all floor patterns and then feed the segmented raw sensor data windows in up and down directions separately to the algorithm for profile extraction. Extracted profiles are fed to the deep autoencoder model for extracting new deep features. Lastly, we apply random forest as a classifier for

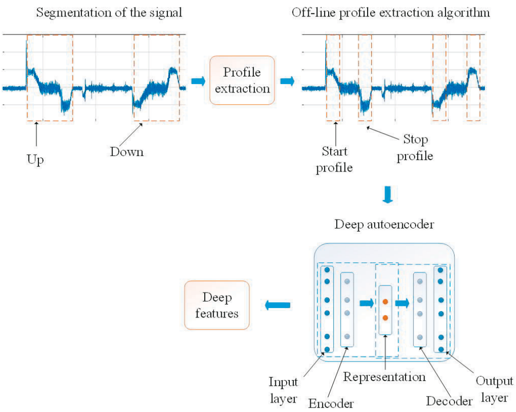


Figure 1: Off-line profile extraction and deep autoencoder feature extraction approach.

fault detection based on new deep features extracted from the profiles.
The encoder transforms the input x into corrupted input data $x^{'}$ using hidden representation H through nonlinear mapping

$$H = f(W_1 x^{'} + b) \tag{3}$$

where $f(.)$ is a nonlinear activation function as the sigmoid function, $W_1 \in R^{k*m}$ is the weight matrix and $b \in R^k$ the bias vector to be optimized in encoding with k nodes in the hidden layer (Vincent et al. 2008). Then, with parameters $W_2 \in R^{m*k}$ and $c \in R^m$, the decoder uses nonlinear transformation to map hidden representation H to a reconstructed vector $x^{"}$ at the output layer

$$x^{"} = g(W_2 H + c) \tag{4}$$

where $g(.)$ 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.

Random forest

The final classification accuracy of random forest is calculated by averaging, i.e. arithmetic mean of the probabilities of assigning classes related to all the produced trees (e). Testing data (d) that is unknown to all the decision trees is used for evaluation by the voting method. Specifically, let sensor data value $v_l{}^e$ have training sample l^{th} in the arrived leaf node of the decision tree $e \in E$, where $l \in [1, ..., L_e]$ and the number of training samples is L_e in the current arrived leaf node of decision tree e. The final prediction result is given by (Huynh et al. 2016):

$$\mu = \frac{\sum_{e \in E} \sum_{l \in [1,...,L_e]} v_l{}^{\mathrm{e}}}{\sum_{e \in E} L_{\mathrm{e}}} \tag{5}$$

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

$$H(a) = \arg\max_{\mathrm{y_j}} \sum_{i \in [1,2,...,Z]} I(h_{\mathrm{i}}(a) = y_{\mathrm{j}}) \tag{6}$$

where $j = 1,2,...,C$ and 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 $I(.)$ defined as:

$$I(x) = \begin{cases} 1, & h_{\mathrm{i}}(a) = y_{\mathrm{j}} \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

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

RESULTS AND DISCUSSION

In this research, we first selected all floor patterns like floor 2-5, 3-8 and so on from the data, some of which are shown in Table 1.

Table 1: Floor patterns.

Start floor	Stop floor
0	1
2	5
3	8
4	6

The next step includes the selection of faulty rides from all floor patterns based on time periods provided by the maintenance data. An equal amount of healthy rides are also selected. Only the vertical component of acceleration data is selected in this research because it is the most informative aspect, consisting of significant changes in vibration levels as compared to other components. Healthy and faulty rides are fed to the algorithm for profile extraction separately. Start and stop profiles are of equal length, irrespective of floor combination.

Up movement

We have analyzed up and down movements separately because the traction based elevator usually produces slightly different levels of vibration in each direction. First, we have selected faulty rides based on time periods provided by the maintenance data, including all floor patterns, which is fed to the algorithm for profile extraction, as shown in Figure 2.
Then we have selected an equal number of rides for healthy data, similar to Figure 2. The next step is to

Figure 2: Profiles from faulty rides (Acc represents acceleration signal).

label both the healthy and faulty profiles with class labels 0 and 1 respectively. Healthy and faulty profiles with class labels are fed to the deep autoencoder model and the generated deep features are shown in Figure 3. These are called deep features or latent features in deep autoencoder terminology, which shows hidden representations of the data.

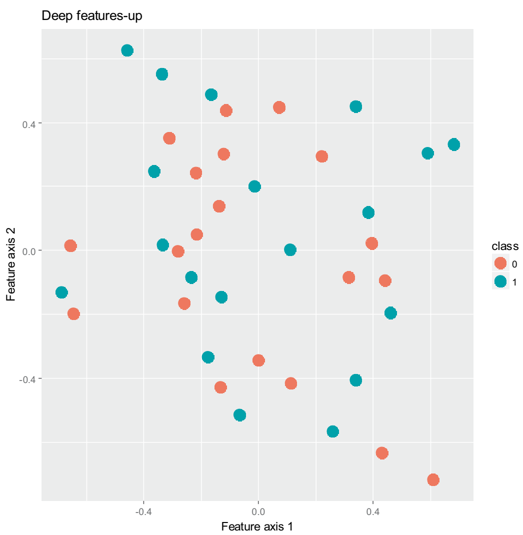


Figure 3: Extracted deep autoencoder features (visualization of the features w.r.t class variable).

Extracted deep features are fed to the random forest algorithm for classification, and the results provide 100% accuracy in fault detection in Table 2. We have compared accuracy in terms of avoiding false positives from both features and found that new deep features generated in this research outperform the existing features. We have used the rest of the healthy rides for extracting profiles to analyze the number of false positives. These healthy profiles are labelled as class 0 and fed to the deep autoencoder to extract new deep features from the profiles.

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. Table 2 presents the results for upward movement of the elevator in terms of accuracy, sensitivity and specificity. We have also included the accuracy of avoiding false positives as an evaluation parameter for this research. The results show that the new deep features provide better accuracy in terms of fault detection and avoiding false positives from the data, which is helpful in detecting false alarms for elevator predictive maintenance strategies. It is extremely helpful in reducing unnecessary visits by maintenance personnel to installation sites.

Table 2: Fault detection analysis (False positives field related to analyzing rest of the healthy profiles after the training and testing phase).

	Deep features	**Existing features**
Accuracy	1	0.55
Sensitivity	1	0.33
Specificity	1	0.80
False positives	1	0.48

Down movement

For downward motion, we have repeated the same analysis procedure as in the case of upward motion. Table 3 presents the results for fault detection with deep autoencoder random forest model in the downward direction. The results are similar to the upward direction but we can see significant change in terms of accuracy of fault detection and when analyzing the number of false positives with new deep features.

CONCLUSIONS AND FUTURE WORK

This research focuses on the health monitoring of elevator systems using a novel fault detection technique. The goal of this research was to develop generic models for profile extraction and automated feature extrac-

Table 3: Fault detection analysis.

	Deep features	Existing features
Accuracy	1	0.78
Sensitivity	1	0.60
Specificity	1	1
False positives	0.98	0.66

tion for 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 for all floor combinations with new extracted deep features. The results support the goal of this research of developing generic models which can be used in other machine systems for fault detection. Our models outperform others because of new deep features extracted from the dataset as compared to existing features calculated from the same raw sensor dataset. The automated feature extraction approach does not require any prior domain knowledge. It also provides dimensionality reduction and is robust against overfitting characteristics.
In future work, we will extend our approach on more elevators and real-world big data cases to validate its potential for other applications and improve its efficacy.

REFERENCES

Bingham C.; Walsh C.; and Carroll S., 2012. *Impact of driving characteristics on electric vehicle energy consumption and range. IET Intelligent Transport Systems*, 6, no. 1, 29–35.

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

Di Martino J., 1985. *Dynamic time warping algorithms for isolated and connected word recognition.* In *New systems and architectures for automatic speech recognition and synthesis*, Springer. 405–418.

Fogelman-Soulie F.; Robert Y.; and Tchuente M., 1987. *Automata networks in computer science: theory and applications.* Manchester University Press and Princeton University Press.

Huynh T.; Gao Y.; Kang J.; Wang L.; Zhang P.; Lian J.; and Shen D., 2016. *Estimating CT image from MRI data using structured random forest and auto-context model. IEEE transactions on medical imaging*, 35, no. 1, 174.

Jia F.; Lei Y.; Lin J.; Zhou X.; and Lu N., 2016. *Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. Mechanical Systems and Signal Processing*, 72, 303–315.

Jiang G.; Xie P.; He H.; and Yan J., 2018. *Wind turbine fault detection using a denoising autoencoder with temporal information. IEEE/ASME Transactions on Mechatronics*, 23, no. 1, 89–100.

Martínez-Rego D.; Fontenla-Romero O.; and Alonso-Betanzos A., 2011. *Power wind mill fault detection via one-class ν-SVM vibration signal analysis.* In *Neural Networks (IJCNN), The 2011 International Joint Conference on.* IEEE, 511–518.

Mishra K.M.; Krogerus T.; and Huhtala K., 2019. *Fault detection of elevator systems using deep autoencoder feature extraction.* In *Research Challenges in Information Science (RCIS), 2019 International Conference on.* IEEE, 43–48.

Rhudy M., 2014. *Time alignment techniques for experimental sensor data. Int J Comput Sci Eng Survey*, 5, no. 2, 1–14.

Soyka F.; Giordano P.R.; Beykirch K.; and Bülthoff H.H., 2011. *Predicting direction detection thresholds for arbitrary translational acceleration profiles in the horizontal plane. Experimental brain research*, 209, no. 1, 95–107.

Sun W.; Shao S.; Zhao R.; Yan R.; Zhang X.; and Chen X., 2016. *A sparse auto-encoder-based deep neural network approach for induction motor faults classification. Measurement*, 89, 171–178.

Vincent P.; Larochelle H.; Bengio Y.; and Manzagol P.A., 2008. *Extracting and composing robust features with denoising autoencoders.* In *Proceedings of the 25th international conference on Machine learning.* ACM, 1096–1103.

Wang X.; Wu S.; and Wang K., 2015. *A run-to-run profile control algorithm for improving the flatness of nano-scale products. IEEE Transactions on Automation Science and Engineering*, 12, no. 1, 192–203.

Xia M.; Li T.; Xu L.; Liu L.; and de Silva C.W., 2018. *Fault Diagnosis for Rotating Machinery Using Multiple Sensors and Convolutional Neural Networks. IEEE/ASME Transactions on Mechatronics*, 23, no. 1, 101–110.

Yang Z.X. and Zhang P.B., 2016. *ELM Meets RAE-ELM: A hybrid intelligent model for multiple fault diagnosis and remaining useful life predication of rotating machinery.* In *Neural Networks (IJCNN), 2016 International Joint Conference on.* IEEE, 2321–2328.

PUBLICATION
IV

Elevator Fault detection using profile extraction and deep autoencoder feature extraction for acceleration and magnetic signals

K. M. Mishra and K. Huhtala

Applied Sciences 9.(2019), 15p

applied sciences

MDPI

Article

Elevator Fault Detection Using Profile Extraction and Deep Autoencoder Feature Extraction for Acceleration and Magnetic Signals

Krishna Mohan Mishra * and Kalevi Huhtala

Unit of Automation Technology and Mechanical Engineering, Tampere University, 33720 Tampere, Finland
* Correspondence: krishna.mishra@tuni.fi

Received: 26 June 2019; Accepted: 22 July 2019; Published: 25 July 2019

check for updates

Abstract: In this paper, we propose a new algorithm for data extraction from time-series data, and furthermore automatic calculation of highly informative deep features to be used in fault detection. In data extraction, elevator start and stop events are extracted from sensor data including both acceleration and magnetic signals. In addition, a generic deep autoencoder model is also developed for automated feature extraction from the extracted profiles. After this, extracted deep features are classified with random forest algorithm for fault detection. Sensor data are labelled as healthy and faulty based on the maintenance actions recorded. The remaining healthy data are used for validation of the model to prove its efficacy in terms of avoiding false positives. We have achieved above 90% accuracy in fault detection along with avoiding false positives based on new extracted deep features, which outperforms results using existing features. Existing features are also classified with random forest to compare results. Our developed algorithm provides better results due to the new deep features extracted from the dataset when compared to existing features. This research will help various predictive maintenance systems to detect false alarms, which will in turn reduce unnecessary visits of service technicians to installation sites.

Keywords: elevator system; deep autoencoder; fault detection; feature extraction; random forest; profile extraction

1. Introduction

In recent years, elevator systems have been used increasingly extensively in apartments, commercial facilities, and office buildings. Presently 54% of the world's population lives in urban areas [1]. 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 while maximizing the uptime of the system [2,3]. Elevator production and service companies are now opting for a predictive maintenance policy to provide better service to customers. They are remotely monitoring faults in elevators and estimating the remaining lifetime of the components responsible for faults. Elevator systems require fault detection and diagnosis for healthy operation [4].

Fault diagnosis methods based on deep neural networks [5–7] and convolutional neural networks [8,9] feature extraction methodology are presented as state of the art for rotatory machines similar to elevator systems. Support vector machines [10] and extreme learning machines [11] are also used as fault detection methods for rotatory machines. However, we have developed an intelligent deep autoencoder random forest-based feature extraction methodology for fault detection in elevator systems to improve the performance of traditional fault diagnosis methods.

Profile extraction for health monitoring is a major issue in automated industrial applications such as elevator systems, computer numerical control, machinery, and robotics [12]. Although rotating machine have been running for decades, but profile extraction and processing methods are not widely available [13]. Profile extraction methods have applied in electric vehicles [14], computer numerical control systems [15] and horizontal planes [16]. Kalman filter [17] is one of the methods being used for profile extraction. However, we have developed an off-line profile extraction algorithm based on low-pass filtering and peak detection to extract elevator start and stop events from sensor data including both acceleration and magnetic signals.

In the last decade, neural networks [18] have extracted highly meaningful statistical patterns from large-scale and high-dimensional datasets. Neural networks [19] has also been used to improve elevator ride comfort via speed profile design. Neural networks [20] has been applied successfully to nonlinear time-series modeling. A deep learning network can self-learn the relevant features from multiple signals [21]. Deep learning algorithms are frequently used in areas such as bearing fault diagnosis [22], machine defect detection [23], vibration signal analysis [24], computer vision [25] and image classification [26]. Autoencoding is a process for nonlinear dimension reduction with natural transformation architecture using feedforward neural network [27]. Autoencoders have proven powerful as nonlinear feature extractors [28]. 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 [29]. Autoencoders were first introduced by LeCun [30], and have been studied for decades. Traditionally, feature learning and dimensionality reduction are the two main features of autoencoders. Recently, autoencoders have been considered one of the most compelling subspace analysis techniques because of the existing theoretical relations between autoencoders and latent variable models [31]. Autoencoders have been used for feature extraction from the data in systems such as induction motors [32] and wind turbines [33] for fault detection, different from elevator systems as in our research.

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 are 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 an algorithm for profile extraction from the raw sensor data rides including both acceleration and magnetic signals. In addition, a generic algorithm with deep autoencoder random forest approach for automated feature extraction from raw sensor data profiles for fault detection in elevator systems.

Our off-line profile extraction algorithm is signal-based and deep autoencoder random forest method is model-based. First, it extracts profiles from time-series signal and then, calculates highly informative deep features from extracted profiles. It is better than other algorithms because it provides better results, dimensionality reduction and is robust against overfitting characteristics.

We have proposed a reliable fault detection model with above 90% accuracy in fault detection, which will increase the safety of passengers. In addition, we have validated the efficacy of the pre-trained model in terms of false positives with the remaining healthy rides, which is helpful in detecting false alarms for elevator predictive maintenance strategies. It is extremely helpful in reducing unnecessary visits by maintenance personnel to installation sites.

Figure 1 shows the fault detection approach used in this paper, which includes raw sensor data rides extracted based on time periods provided by the maintenance data from all floor patterns. Acceleration and magnetic signal rides collected from an elevator system are fed to the algorithm for profile extraction separately. These extracted profiles from all five traction elevators including both acceleration and magnetic signals are then fed to the deep autoencoder model for feature extraction, and then random forest performs the fault detection task based on extracted deep features. We only extract start and stop profiles from the both acceleration and magnetic signal rides because of the different lengths of rides for each floor combination due to the constant speed phase, which is longer when there is longer travel.

This paper provides the following novelties. (1) We propose a new off-line profile extraction algorithm for extracting elevator start and stop events from time-series data. (2) In addition, we propose a new deep autoencoder model to automatically generate highly informative deep features from sensor data for fault detection. The rest of this paper is organized as follows. Section 2 presents the methodology of the paper including profile extraction, deep autoencoder, and random forest algorithms. Then, Section 3 includes the details of experiments performed, results, and discussion. Finally, Section 4 concludes the paper and presents the future work.

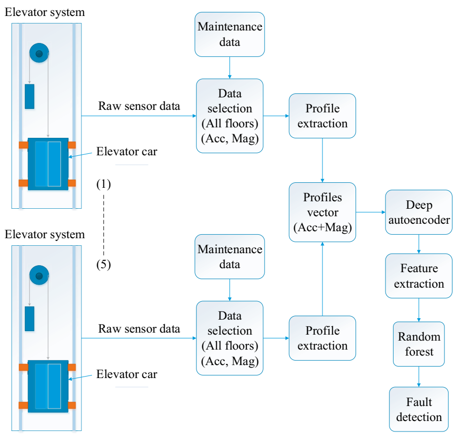


Figure 1. Fault detection approach (Acc and Mag represent acceleration and magnetic signals respectively).

2. Methodology

In this study, we have used 12 different existing features derived from raw sensor data describing the motion and vibration of an elevator for fault detection and diagnostics of multiple faults. We have developed an automated feature extraction technique for raw sensor data in this research as an extension to the work of our previous research to compare the results using new extracted deep features. In our previous research [34], we have used only acceleration signal, which represents vibration related features. In this research, we have extended our approach to include magnetic signals, which represents position related features. This will validate our goal of this research to develop generic models for profile extraction and automated feature extraction for fault detection in the health state monitoring of elevator systems. In addition, we have analyzed almost two months of the data from five traction elevators in this research as an extension to one elevator in our previous research. Each elevator usually produces around 200 rides per day. Each ride used in analysis contains around 5000 rows of the data, which proves robustness of the algorithms over large dataset. We have excluded around 20 rides before and after the time period of faulty rides in selecting healthy rides, which will help us to remove suspicious data from the analysis with our algorithm. We have used 70% of the data for training and rest 30% for testing.

2.1. Profile Extraction Algorithm

Raw sensor data collected from elevator systems typically encompass a large collection of data points sampled at high frequency. In order to feed large sensor data to cloud-based applications, it is often desirable to pre-process the data and perform compression before transmission, for example in the form of edge computing performed in the device end. Here we assume that raw data is in the form of a one-dimensional time-series vector with equidistant sampling times. The goal of the proposed method is to compress the raw time series obtained from machinery while maintaining the information about key events, and secondly, to make the data more applicable for machine learning.

The algorithm works in two stages. In the first stage, the signal is pre-processed and normalized, followed by low-pass filtering. The low-pass filtered signal is used for peak detection, which for each elevator travel detects a local minimum and maximum corresponding to acceleration and deceleration (start and stop) events. The algorithm uses crude low-pass filtering with a low cut-off frequency for peak detection, which ensures that a sustained period of acceleration is required for a peak to be registered. This prevents short bursts of noise from being detected as a movement window. Low-pass filtering is applied to ensure that only a sustained acceleration or deceleration event is registered as a peak as opposed to noise.

Low-pass filter is used to avoid noise spikes being detected as peaks. From low-pass filter signal, the algorithm cannot detect the precise magnitude or timing of peaks, but it will detect the approximate region in which to align the event profiles. Unfiltered data is then used for profile alignment.

In the second stage, alignment and collection of equal length profiles is performed based on windowing of the acceleration signal near the peak events. In this stage, the raw acceleration signal is used instead of the filtered signal. A number of time domain alignment methods have been proposed in the literature. Dynamic time warping (DTW) has been commonly applied, e.g., in speech recognition [35], whereas various alignment techniques for sensor data have been presented in [36]. Here alignment is performed against a reference profile. The reference profile is aligned against the raw data in the window of the detected peaks. The length of the initial profile window m is selected empirically based on the sample frequency and the maximum estimated length of the elevator acceleration events. The criterion for optimal alignment was defined as the alignment that minimizes the sum of the Euclidean or L_2 norm. The output from this operation is an $n \times m$ matrix of aligned profiles describing n acceleration and deceleration events of length m.

To improve the alignment accuracy, the reference profile is updated iteratively following each batch run. Each sequence in the profile matrix is closely synchronized in time and can hence be considered a repetition of the same signal. Using signal averaging, the new reference profile is calculated as the mean of the n extracted profiles. This both maintains the main characteristics of the signal and reduces the noise. Assuming white noise and perfect synchronization, signal averaging improves the signal-to-noise ratio *(SNR)* by a factor of $\sqrt{n}$.

Information in the obtained reference profile can be used to update the window size m. Assuming an overestimated size of the event window, the averaged reference profile will contain superfluous close to zero values corresponding to no acceleration. The number of elements s below this threshold in the reference profile can be used to estimate the optimal window length by reducing the window length m by s for the following iteration.

The off-line profile extraction algorithm is described as following.

Off-line profile extraction algorithm

Pre-procession

1. Read a vector of raw acceleration data containing k elevator travels. Define the zero mean transformed dataset as X.

2. Perform low-pass filtering on X and obtain denoised dataset Y.

Initialization

3. Define parameters for reference profile. Set the approximated maximum window length to m samples and height h to the 99th percentile of the low-pass filtered dataset.

4. Define alignment window size a and set $k = 1$.

Iteration

5. From $Y(k)$, detect peak acceleration points y_{min} and y_{max}

6. Align reference profile P against raw dataset X in the vicinity of detected peaks by minimizing the L_2 norm according to

$$min \sum_{i=-a/2}^{a/2} \sum_{j=1}^{m} [-p_j - x_{min+i+j}]^2 \tag{1}$$

$$min \sum_{i=-a/2}^{a/2} \sum_{j=1}^{m} [p_j - x_{max+i+j}]^2 \tag{2}$$

7. Add aligned data points from *X(k)* as rows into an $n \times m$ profile matrix, alternatively separate matrices according to direction of travel (min/max).

8. Set travel window $k = k + 1$ and repeat steps 5–7 until end of dataset.

9. Update reference profile P with the signal-averaged profile obtained from the column-wise mean of the new profile matrix.

10. Reduce window length m by s samples, where s is the number of elements in P that satisfy

$$p \leq \epsilon, p \in P \tag{3}$$

where ϵ is a close to zero number indicating no acceleration.

11. Set $k = 1$ and continue with new batch iterations by repeating steps 5–8.

2.2. Deep Autoencoder

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 output layer and reproduces its inputs as its output. We are using a five-layer deep autoencoder (see Figure 2) including input, output, encoder, decoder, and representation layers, which is a different approach than in [33,37]. In our approach, we first analyze the data to find all floor patterns and then feed the segmented raw sensor data windows in up and down directions separately to the algorithm for profile extraction. Extracted profiles from both acceleration and magnetic signals are fed to the deep autoencoder model for extracting new deep features. Lastly, we apply random forest as a classifier for fault detection based on new deep features extracted from the profiles. We have combined healthy and faulty profiles as a vector from all five traction elevators including both acceleration and magnetic signals before feature extraction.

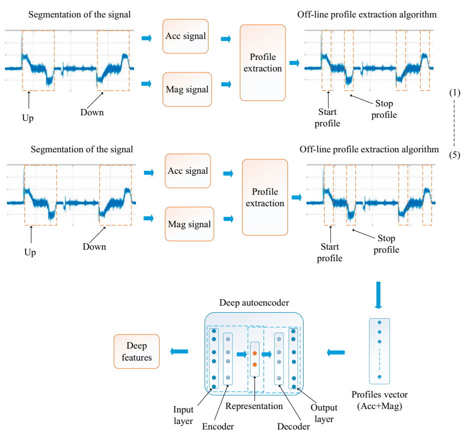


Figure 2. Off-line profile extraction and deep autoencoder feature extraction approach.

The encoder transforms the input x into corrupted input data x' using hidden representation H through nonlinear mapping

$$H = f(W_1 x' + b) \tag{4}$$

where $f(.)$ 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 [37]. 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.

$$x^{''} = g(W_2 H + c) \tag{5}$$

where $g(.)$ 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 [38]. Among multiple input variables the use of nonlinear activation functions provides us better opportunity to capture nonlinear relationships. Effective fault detection is a challenge due to nonlinearity of elevator systems and as a result, time-series data will have temporal dependencies. Our proposed approach can capture nonlinear relationships among multiple sensor variables, which has improved the performance in terms of fault detection.

2.3. Random Forest

Random forest includes an additional layer of randomness to bagging. It uses different bootstrap samples of the data for constructing each tree [39]. 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 [40]. The final classification accuracy of random forest is calculated by averaging, i.e., arithmetic mean of the probabilities of assigning classes related to all the produced trees (*e*). Testing data (*d*) that is unknown to all the decision trees is used for evaluation by the voting method (see Figure 3).

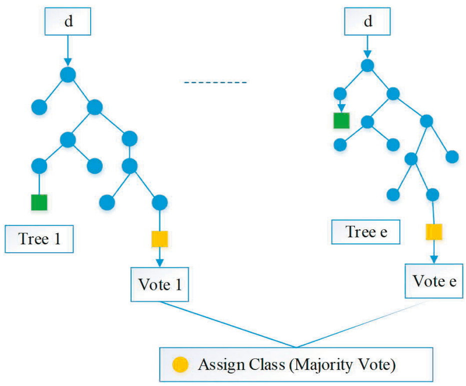


Figure 3. Classification phase of random forest classifier.

Specifically, let sensor data value ${v_l}^e$ have training sample lth in the arrived leaf node of the decision tree $e \in E$, where $l \in [1, \ldots, L_e]$ and the number of training samples is L_e in the current arrived leaf node of decision tree e. The final prediction result is given by [41]:

$$\mu = \frac{\sum_{e \in E} \sum_{l \in [1,\ldots,L_e]} {v_l}^e}{\sum_{e \in E} L_e} \tag{6}$$

All classification trees providing a final decision by voting method are given by [42]:

$$H(a) = \arg\max_{y_j} \sum_{i \in [1,2,\ldots,Z]} I(h_i(a) = y_j) \tag{7}$$

where $j = 1, 2, ..., C$ and 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 $I(.)$ defined as:

$$I(x) = \begin{cases} 1, & h_\mathrm{i}(a) = y_\mathrm{j} \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

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

2.4. Evaluation Parameters

Evaluation parameters used in this research are defined with the confusion matrix in Table 1.

Table 1. Confusion matrix.

	Predicted (P)	**(N)**
Actual (P)	True positive (TP)	False negative (FN)
(N)	False positive (FP)	True negative (TN)

The rate of positive test result is sensitivity,

$$Sensitivity = \frac{TP}{TP + FN} * 100\% \tag{9}$$

The ratio of a negative test result is specificity,

$$Specificity = \frac{TN}{TN + FP} * 100\% \tag{10}$$

The overall measure is accuracy,

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} * 100\% \tag{11}$$

3. Results and Discussion

In this research, we first selected all floor patterns like floor 2–5, 3–8, and so on from the data, some of which are shown in Table 2.

Table 2. Floor patterns.

Start Floor	**Stop Floor**
0	1
2	5
3	8
4	6

The next step includes the selection of faulty rides from all floor patterns based on time periods provided by the maintenance data. An equal number of healthy rides are also selected. Only the vertical component of both acceleration and magnetic signal data is selected in this research because it is the most informative aspect, consisting of significant changes in vibration levels as compared to other components. Healthy and faulty rides are fed to the algorithm for profile extraction separately. Start and stop profiles are of equal length, irrespective of floor combination.

First, we have selected all floor patterns from the data and then divided the data into up and down directions. Next, we selected the rides from healthy and faulty parts of the data and extracted profiles from them. These profiles are fed to deep autoencoder model for feature extraction and based on these feature faults are detected.

3.1. Up Movement

We have analyzed up and down movements separately because the traction-based elevator usually produces slightly different levels of vibration in each direction. First, we have selected faulty rides based on time periods provided by the maintenance data, including all floor patterns, which is fed to the algorithm for profile extraction, as shown in Figure 4. Then, we have selected an equal number of rides for healthy data, and the extracted profiles are shown in Figure 5. Visualization of the profiles proved that our proposed algorithm extracted elevator start and stop events have equal length irrespective of the floor combination as shown in Figures 4 and 5.

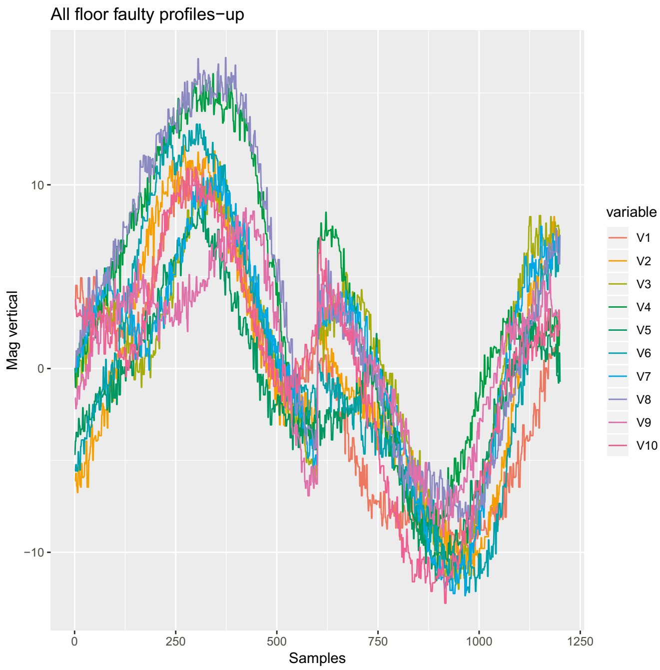


Figure 4. Profiles from faulty rides (The figure shows first 10 rides from the combined version of elevator start and stop events from magnetic signals in z-direction i.e., Mag vertical. Profiles are extracted from all floor combinations during the time period of fault occurred and fixed by maintenance personal in upward movement of elevator system. Color labelling is based on maximum number of rides available in the data during this time period i.e., variable.).

The next step is to label both the healthy and faulty profiles with class labels 0 and 1, respectively. Healthy and faulty profiles with class labels are fed to the deep autoencoder model and the generated deep features are shown in Figure 6. These are called deep features or latent features in deep autoencoder terminology, which shows hidden representations of the data. In Figure 6, we can see from visualization that both features with class labels are perfectly separated, which results in better fault detection.

Extracted deep features are fed to the random forest algorithm for classification, and the results provide 90% accuracy in fault detection as shown in Table 3. We have compared accuracy in terms of avoiding false positives from both features and found that new deep features generated in this research outperform the existing features. We have used the remaining healthy rides for extracting profiles to analyze the number of false positives. These healthy profiles are labelled as class 0 and fed to the deep autoencoder to extract new deep features from the profiles, as shown in Figure 7.

Figure 5. Profiles from healthy rides (The figure shows first 10 rides from the combined version of elevator start and stop events from magnetic signals in z-direction i.e., Mag vertical. Equal number of profiles as in Figure 4 are extracted from all floor combinations before the time period of fault occurred in upward movement of elevator system. Color labelling is based on number of rides available in the data during this time period i.e., variable.).

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. Table 3 presents the results for upward movement of the elevator in terms of accuracy, sensitivity, and specificity. We have also included the accuracy of avoiding false positives as an evaluation parameter for this research. The results show that the new deep features provide better accuracy in terms of fault detection and avoiding false positives from the data, which is helpful in detecting false alarms for elevator predictive maintenance strategies. It is extremely helpful in reducing unnecessary visits by maintenance personnel to installation sites.

Table 3. Fault detection analysis (False positives field related to analyzing remaining healthy profiles after the training and testing phase).

	Deep Features	Existing Features
Accuracy	0.90	0.54
Sensitivity	0.92	0.50
Specificity	0.88	0.58
False positives	0.86	0.31

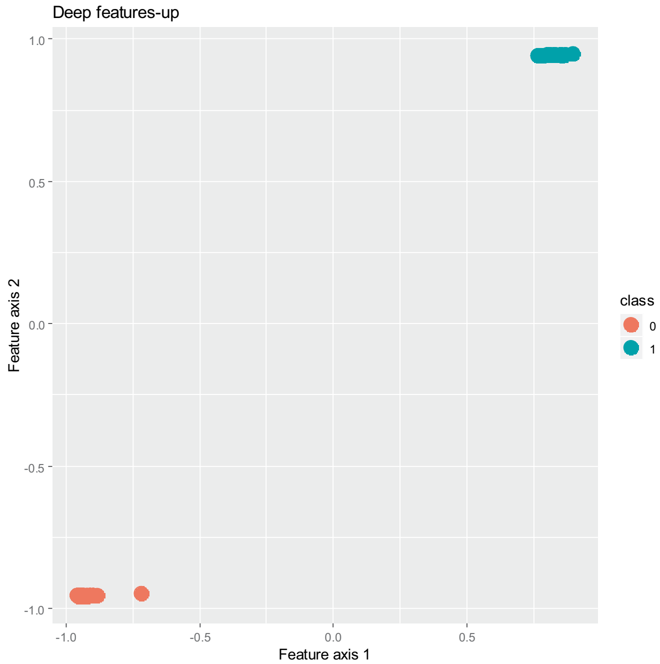


Figure 6. Extracted deep autoencoder features (visualization of the features w.r.t class variable).

Figure 7. Extracted deep features (only healthy profiles).

3.2. Down Movement

For downward motion, we have repeated the same analysis procedure as in the case of upward motion. First we have selected the faulty rides and an equal amount of healthy data rides for profile extraction, as shown in Figures 8 and 9.

Figure 8. Profiles from faulty rides (similar to Figure 4 but in downward movement of elevator system).

Again, we fed both healthy and faulty profiles with class labels to the deep autoencoder for the extraction of new deep features, as shown in Figure 10.

Finally, the new extracted deep features are classified with random forest model and the results are shown in Table 4. After this, the remaining healthy rides are used to analyze the number of false positives. The extracted deep features are shown in Figure 11.

Table 4 presents the results for fault detection with deep autoencoder random forest model in the downward direction. The results are similar to the upward direction, but we can see significant change in terms of accuracy of fault detection and when analyzing the number of false positives with new deep features.

Table 4. Fault detection analysis.

	Deep Features	Existing Features
Accuracy	0.95	0.59
Sensitivity	0.92	0.67
Specificity	0.97	0.50
False positives	1	0.61

Figure 9. Profiles from healthy rides (similar to Figure 5 but in downward movement of elevator system).

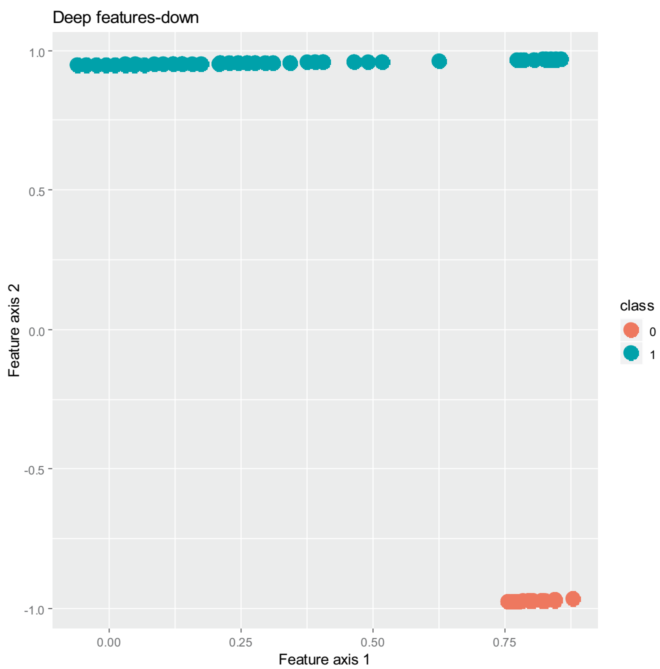


Figure 10. Extracted deep features.

Figure 11. Extracted deep features (only healthy profiles).

4. Conclusions and Future Work

This research focuses on the health monitoring of elevator systems using a novel fault detection technique. The goal of this research was to develop generic models for profile extraction and automated feature extraction for fault detection in the health state monitoring of elevator systems. Our approach in this research provided above 90% accuracy in fault detection and in the case of analyzing false positives for all floor combinations with new extracted deep features from sensor data including both acceleration and magnetic signals. The results support the goal of this research of developing generic models which can be used in other machine systems for fault detection. 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 used to allocate maintenance resources. Our developed models can also be used for different predictive maintenance solutions to automatically generate highly informative deep features for solving diagnostics problems. Our models outperform others because of new deep features extracted from the dataset as compared to existing features calculated from the same raw sensor dataset. 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 models, which will increase the safety of passengers as well as serve the public interest. Visualization of the extracted profiles and features support our goal of developing generic models for profile and feature extraction for fault detection.

In future work, we will extend our approach on other real-world big data cases to validate its potential for other applications and improve its efficacy.

Author Contributions: Conceptualization, K.M.M. and K.H.; Data curation, K.M.M.; Formal analysis, K.M.M.; Investigation, K.M.M.; Methodology, K.M.M.; Resources, K.M.M.; Software, K.M.M.; Supervision, K.H.; Validation, K.M.M. and K.H.; Visualization, K.M.M.; Writing-original draft, K.M.M.; Writing-review and editing, K.M.M. and K.H.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Desa, U. *World Urbanization Prospects*; the 2011 Revision; Population Division, Department of Economic and Social Affairs, United Nations Secretariat: New York, NY, USA, 2014.
2. Ebeling, T. Condition monitoring for elevators—An overview. *Lift Rep.* **2011**, *6*, 25–26.
3. Ebeling, T.; Haul, M. Results of a field trial aiming at demonstrating the permanent detection of elevator wear using intelligent sensors. *Proc. ELEVCON* **2016**, *20*, 101–109.
4. Wang, P.; He, W.; Yan, W. Fault diagnosis of elevator braking system based on wavelet packet algorithm and fuzzy neural network. In Proceedings of the 2009 9th International Conference on Electronic Measurement & Instruments, Beijing, China, 16–19 August 2009; pp. 4–1028.
5. Zhang, R.; Peng, Z.; Wu, L.; Yao, B.; Guan, Y. Fault diagnosis from raw sensor data using deep neural networks considering temporal coherence. *Sensors* **2017**, *17*, 549. [CrossRef] [PubMed]
6. Jia, F.; Lei, Y.; Lin, J.; Zhou, X.; Lu, N. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mech. Syst. Signal Process.* **2016**, *72*, 303–315. [CrossRef]
7. Bulla, J.; Orjuela-Cañón, A.D.; Flórez, O.D. Feature extraction analysis using filter banks for faults classification in induction motors. In Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazi, 8–13 July 2018; pp. 1–6.
8. Xia, M.; Li, T.; Xu, L.; Liu, L.; de Silva, C.W. Fault Diagnosis for Rotating Machinery Using Multiple Sensors and Convolutional Neural Networks. In Proceedings of the 2018 IEEE International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazi, 8–13 July 2018; pp. 101–110.
9. Jing, L.; Wang, T.; Zhao, M.; Wang, P. An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox. *Sensors* **2017**, *17*, 414. [CrossRef] [PubMed]
10. Martínez-Rego, D.; Fontenla-Romero, O.; Alonso-Betanzos, A. Power wind mill fault detection via one-class v-SVM vibration signal analysis. In Proceedings of the 2011 IEEE International Joint Conference on Neural Networks, San Jose, CA, USA, 31 July–5 August 2011; pp. 511–518.
11. Yang, Z.; Zhang, P. ELM Meets RAE-ELM: A hybrid intelligent model for multiple fault diagnosis and remaining useful life predication of rotating machinery. In Proceedings of the 2016 IEEE International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; pp. 2321–2328.
12. Osornio-Rios, R.A.; de Jesús Romero-Troncoso, R.; Herrera-Ruiz, G.; Castañeda-Miranda, R. FPGA implementation of higher degree polynomial acceleration profiles for peak jerk reduction in servomotors. *Robot. Comput. Integr. Manuf.* **2009**, *25*, 379–392. [CrossRef]
13. Carr, C.E.; Bryan, N.C.; Saboda, K.N.; Bhattaru, S.A.; Ruvkun, G.; Zuber, M.T. Acceleration profiles and processing methods for parabolic flight. *NPJ Microgravity* **2018**, *4*, 14. [CrossRef] [PubMed]
14. Bingham, C.; Walsh, C.; Carroll, S. Impact of driving characteristics on electric vehicle energy consumption and range. *IET Intell. Transp. Syst.* **2012**, *6*, 29–35. [CrossRef]
15. Nam, S.; Yang, M. A study on a generalized parametric interpolator with real-time jerk-limited acceleration. *Comput. -Aided Des.* **2004**, *36*, 27–36. [CrossRef]
16. Soyka, F.; Giordano, P.R.; Beykirch, K.; Bülthoff, H.H. Predicting direction detection thresholds for arbitrary translational acceleration profiles in the horizontal plane. *Exp. Brain Res.* **2011**, *209*, 95–107. [CrossRef] [PubMed]
17. Wang, X.; Wu, S.; Wang, K. A run-to-run profile control algorithm for improving the flatness of nano-scale products. *IEEE Trans. Autom. Sci. Eng.* **2013**, *12*, 192–203. [CrossRef]
18. Calimeri, F.; Marzullo, A.; Stamile, C.; Terracina, G. Graph Based Neural Networks for Automatic Classification of Multiple Sclerosis Clinical Courses. In Proceedings of the ESANN, Bruges, Belgium, 25–27 April 2018.
19. Seppala, J.; Koivisto, H.; Koivo, H. Modeling elevator dynamics using neural networks. In Proceedings of the 1998 IEEE International Joint Conference on Neural Networks Proceedings, IEEE World Congress on Computational Intelligence (Cat. No. 98CH36227), Anchorage, AK, USA, 4–9 May 1998; pp. 2419–2424.
20. Lee, C. Gender classification using M-estimator based radial basis function neural network. In Proceedings of the 2014 International Conference on Signal Processing and Multimedia Applications (SIGMAP), Vienna, Austria, 28–30 August 2014; pp. 302–306.

21. Fernández-Varela, I.; Athanasakis, D.; Parsons, S.; Hernández-Pereira, E.; Moret-Bonillo, V. Sleep Staging with Deep Learning: A Convolutional Model. In Proceedings of the ESANN, Bruges, Belgium, 25–27 April 2018.
22. Zhuang, Z.; Lv, H.; Xu, J.; Huang, Z.; Qin, W. A Deep Learning Method for Bearing Fault Diagnosis through Stacked Residual Dilated Convolutions. *Appl. Sci.* **2019**, *9*, 1823. [CrossRef]
23. Wang, J.; Sanchez, J.; Iturrioz, J.; Ayesta, I. Geometrical Defect Detection in the Wire Electrical Discharge Machining of Fir-Tree Slots Using Deep Learning Techniques. *Appl. Sci.* **2019**, *9*, 90. [CrossRef]
24. Lin, W.; Lo, S.; Young, H.; Hung, C. Evaluation of Deep Learning Neural Networks for Surface Roughness Prediction Using Vibration Signal Analysis. *Appl. Sci.* **2019**, *9*, 1462. [CrossRef]
25. Donati, L.; Iotti, E.; Mordonini, G.; Prati, A. Fashion Product Classification through Deep Learning and Computer Vision. *Appl. Sci.* **2019**, *9*, 1385. [CrossRef]
26. Gu, Y.; Wang, Y.; Li, Y. A Survey on Deep Learning-Driven Remote Sensing Image Scene Understanding: Scene Classification, Scene Retrieval and Scene-Guided Object Detection. *Appl. Sci.* **2019**, *9*, 2110. [CrossRef]
27. Hänninen, J.; Kärkkäinen, T. Comparison of Four-and Six-Layered Configurations for Deep Network Pretraining. In Proceedings of the ESANN, Bruges, Belgium, 27–29 April 2016.
28. Albuquerque, A.; Amador, T.; Ferreira, R.; Veloso, A.; Ziviani, N. Learning to rank with deep autoencoder features. In Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazi, 8–13 July 2018; pp. 1–8.
29. Huet, R.; Courty, N.; Lefèvre, S. A New Penalisation Term for Image Retrieval in Clique Neural Networks. In Proceedings of the ESANN, Bruges, Belgium, 27–29 April 2016.
30. Fogelman-Soulie, F.; Robert, Y.; Tchuente, M. *Automata Networks in Computer Science: Theory and Applications*; Manchester University Press: Manchester, UK; Princeton University Press: Princeton, NJ, USA, 1987.
31. Madani, P.; Vlajic, N. 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, Raleigh, NC, USA, 10–11 April 2018; p. 1.
32. Sun, W.; Shao, S.; Zhao, R.; Yan, R.; Zhang, X.; Chen, X. A sparse auto-encoder-based deep neural network approach for induction motor faults classification. *Measurement* **2016**, *89*, 171–178. [CrossRef]
33. Jiang, G.; Xie, P.; He, H.; Yan, J. Wind turbine fault detection using a denoising autoencoder with temporal information. *IEEE/ASME Trans. Mech.* **2017**, *23*, 89–100. [CrossRef]
34. Mishra, K.M.; Krogerus, T.; Huhtala, K. Profile extraction and deep autoencoder feature extraction for elevator fault detection. In Proceedings of the 16th International Conference on Signal Processing and Multimedia Applications (SIGMAP), Prague, Czech Republic, 26–28 July 2019; pp. 319–326.
35. Di Martino, J. Dynamic time warping algorithms for isolated and connected word recognition. In *New Systems and Architectures for Automatic Speech Recognition and Synthesis*; Springer: Berlin/Heidelberg, Germany, 1985; pp. 405–418.
36. Rhudy, M. Time alignment techniques for experimental sensor data. *Int. J. Comput. Sci. Eng. Surv.* **2014**, *5*, 1–14. [CrossRef]
37. Vincent, P.; Larochelle, H.; Bengio, Y.; Manzagol, P. Extracting and composing robust features with denoising autoencoders. In Proceedings of the 25th International Conference on Machine Learning, Helsinki, Finland, 5–9 July 2008; pp. 1096–1103.
38. Japkowicz, N.; Hanson, S.J.; Gluck, M.A. Nonlinear autoassociation is not equivalent to PCA. *Neural Comput.* **2000**, *12*, 531–545. [CrossRef]
39. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
40. Huynh, T.; Gao, Y.; Kang, J.; Wang, L.; Zhang, P.; Lian, J.; Shen, D. Estimating CT image from MRI data using structured random forest and auto-context model. *IEEE Trans. Med. Imaging* **2015**, *35*, 174–183. [CrossRef] [PubMed]
41. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
42. Liu, Z.; Tang, B.; He, X.; Qiu, Q.; Liu, F. Class-specific random forest with cross-correlation constraints for spectral–spatial hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 257–261. [CrossRef]

PUBLICATION

V

Fault detection of elevator systems using multilayer perceptron neural network

K. M. Mishra and K. Huhtala

in Proceedings of the 24th IEEE Conference on Emerging Technologies and Factory Automation (ETFA) 24.(2019), 904–909

Fault Detection of Elevator Systems Using Multilayer Perceptron Neural Network

Krishna Mohan Mishra and Kalevi J. Huhtala

Unit of Automation Technology and Mechanical Engineering
Tampere University
Tampere, Finland
E-mail: {krishna.mishra, kalevi.huhtala}@tuni.fi

***Abstract*—In this research, we propose a generic multilayer perceptron (MLP) neural network model based on deep learning algorithm for automatic calculation of highly informative deep features from the elevator time series data and based on extracted deep features faults are detected. Sensor data are labelled as healthy or faulty based on the maintenance actions recorded. The remaining healthy data are 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. Existing features are also classified with random forest (RF) to compare results. Multilayer perceptron neural network model based on deep learning approach provides better results due to the new deep features extracted from the dataset compared to existing features. Cross-validation method used with multilayer perceptron plays a significant role in improving accuracy of fault detection. Our model provides good classification and is robust against overfitting characteristics. This research will help various predictive maintenance systems to detect false alarms, which will reduce unnecessary visits of service technicians to installation sites.**

***Index Terms*—Elevator System, Multilayer Perceptron, Fault Detection, Feature Extraction, Random Forest**

I. INTRODUCTION

In recent years, elevator systems have been used more and more extensively in apartments, commercial facilities and office buildings. Nowadays 54% of the worlds population lives 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 of the system [3], [4]. Elevator production and service companies are now opting for a predictive maintenance policy to provide better service to customers. They are remotely monitoring faults in elevators and estimating the remaining lifetime of the components responsible for faults. Elevator systems require fault detection and diagnosis for healthy operation [5].

Fault diagnosis methods based on deep neural networks [6], [7], [8] and convolutional neural networks [9], [10] feature extraction methodology are presented as state of the art for rotatory machines similar to elevator systems. Linear discriminant analysis [11], [12], artificial neural networks [13] and kalman filter [14] are also used as fault detection methods for rotatory machines. However, we have developed an intelligent multilayer perceptron neural network model based on deep learning feature extraction methodology for fault detection in elevator systems to improve the performance of traditional fault diagnosis methods.

In the last decade, neural networks [15] have extracted highly meaningful statistical patterns from large-scale and high-dimensional datasets. A neural network can self-learn the relevant features from multiple signals [16]. Neural network algorithms are frequently used in areas such as signal processing [17], condition monitoring [18], fault detection [19], image processing [20] and industrial systems [21]. Multilayer perceptron neural network is one of the most commonly employed forms of artificial neural networks for fault diagnosis. It is the trained with the back-propagation algorithm and is also referred as the backpropagation neural network. The back propagation neural network was proposed by McClelland and Rumelhart [22]. A back-propagation neural network consists of an input layer, one or more hidden layers and an output layer. Synapses are the connecting links between neurons in the different layers associated with a synaptic weight [23]. Dimensionality of the input feature vectors decide number of neurons in the input layer while the number of classes into which the dataset is to be classified are responsible for the number of output neurons. Multilayer perceptron is typically a simple artificial neural network having a linear mapping between the input and output without a hidden layer. Multilayer perceptron neural networks have been used for fault detection from the data in systems like centrifugal pump [24], brake system [25] and bearing system [26], different from elevator systems as in our research.

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 are 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 generic multilayer perceptron neural network model based on deep learning approach for automated feature extraction from elevator sensor data, and based on these deep features, faults are detected. The rest of this paper is organized as follows. Section II presents the methodology of the paper including the data descriptions, multilayer perceptron neural network model, random forest algorithm and evaluation parameters used in this research. Then, section III includes the details of experiments performed, results and discussion. Finally, section IV concludes the paper and presents the future work.

II. Methodology

In this study, we have utilised 12 different existing features derived from raw sensor data describing the motion and vibration of an elevator for fault detection and diagnostics of multiple faults. We have developed an automated feature extraction technique in this research as an extension to the work of our previous research [27] to compare the results using new extracted deep features. We have analyzed almost one year of the data from seven traction elevators in this research. Each elevator usually produces around 200 rides per day, which proves robustness of the algorithms over large dataset. Every movement of the elevator generates existing features from the vibration signal. We have used 70% of the data for training and rest 30% for testing. Fig. 1 shows the fault detection approach used in this paper, which includes elevator sensor data extracted based on time periods provided by the maintenance data. Data collected from the elevator systems are fed to the multilayer perceptron neural network model based on deep learning approach for new feature extraction. Then, fault detection task is performed based on extracted deep features. In addition, random forest is used for fault detection based on the existing features to compare results.

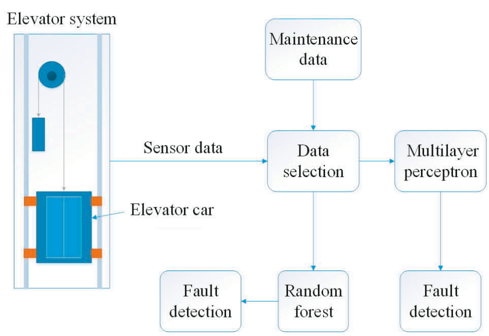


Fig. 1. Fault detection approach

A. Multilayer Perceptron

Multilayer perceptron is a supervised learning algorithm, which learns a non-linear function approximator [28]. Hidden layers are non-linear layers situated between the input and the output layer. Multilayer perceptron can have one or more hidden layers, which makes it different from other algorithm (see Fig. 2). Multilayer perceptron is a feedforward neural network. It can distinguish nonlinearly separable patterns. Multilayer perceptron consists of several nodes called as neurons. Neurons are arranged as a directed graph in multiple layers. Each layer is fully connected to the next layer. Multilayer perceptrons are also called as universal approximators. Any given continuous function can be approximated by multilayer perceptron having one hidden layer with enough neurons [29]. In our approach, we first feed the elevator sensor data from each elevator movement in up and down directions separately to the multilayer perceptron neural network model based on deep learning approach for extracting new deep features from the data and based on new deep features faults are detected. We have combined healthy and faulty existing features as a vector from all seven traction elevators before feature extraction. In this paper, back propagation neural networks are trained in supervised manner and have been applied successfully to elevator fault diagnosis. We have used multilayer perceptron neural network model with two hidden layers, each contains 20 neurons.

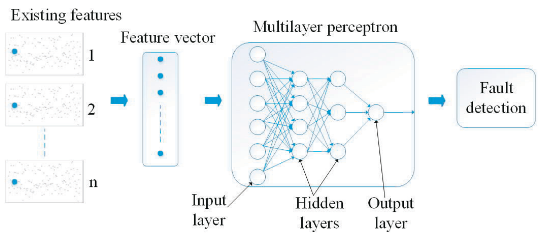


Fig. 2. Multilayer perceptron neural network approach

At first, preparing the training dataset $D = \{(x_i,y_i)\}^n{}_{i=1}$, $x_i \in \mathbb{R}^{m*l}, y_i \in \mathbb{R}$. Where, n is the number of samples. x_i $(i = 1,2,...,n)$ is m-dimensional phased feature vector I_i $(i = 1,2,...,m)$ as the input of multilayer perceptron. y_i is the label of fault and the weighted input of j node in the hidden layer can be expressed as [25]:

$$h_\mathrm{j} = \sum_{i=1}^{m} W_\mathrm{ij} * I_\mathrm{i} + b_\mathrm{j} \tag{1}$$

Where $\mathrm{W_{ij}}$ is the connection weight which from the input layer i node to the hidden layer j node, b_j is bias for the corresponding node, the output of the j node in the hidden layer is H_j .

$$H_\mathrm{j} = tanh(h_\mathrm{j}) \tag{2}$$

After several iterations, the input o_k of output layer k note from hidden layers is

$$o_\mathrm{k} = \sum_{j=1}^{J} W_\mathrm{jk} * H_\mathrm{j} + b_\mathrm{k} \tag{3}$$

Where output layer contains K notes ($k = 1,2,...,K$). The output O_k of the k node in the output layer corresponding to different activation functions.

B. Random Forest

Random forest includes an additional layer of randomness to bagging. It uses different bootstrap samples of the data for constructing each tree [30]. 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 [31]. The final classification accuracy of random forest is calculated by averaging, i.e. arithmetic mean of the probabilities of assigning classes related to all the produced trees *(t)*. Testing data *(d)* that is unknown to all the decision trees is used for evaluation by the voting method (see Fig. 3).

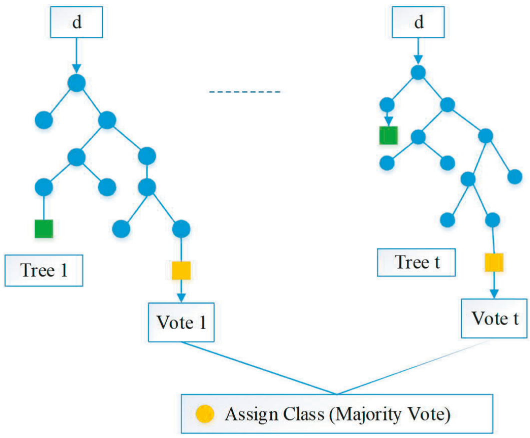


Fig. 3. Classification phase of random forest classifier

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 [32]:

$$\mu = \frac{\sum_{t \in T} \sum_{l \in [1,\ldots,L_t]} v_l{}^t}{\sum_{t \in T} L_t} \tag{4}$$

All classification trees providing a final decision by voting method are given by [33]:

$$H(a) = \arg\max_{y_j} \sum_{i \in [1,2,\ldots,Z]} I(h_i(a) = y_j) \tag{5}$$

where $j= 1,2,...,C$ and 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 $I(.)$ defined as:

$$I(x) = \begin{cases} 1, & h_i(a) = y_j \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

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

C. Evaluation Parameters

Evaluation parameters used in this research are defined with the confusion matrix in Table I.

TABLE I
CONFUSION MATRIX

	Predicted (P)	**(N)**
Actual (P)	True positive (TP)	False negative (FN)
(N)	False positive (FP)	True negative (TN)

The overall measure is accuracy,

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} * 100\% \tag{7}$$

III. 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 multilayer perceptron neural network model based on deep learning approach is used for feature extraction from the data. Some of the existing features are shown in Fig. 4, 5 and 6

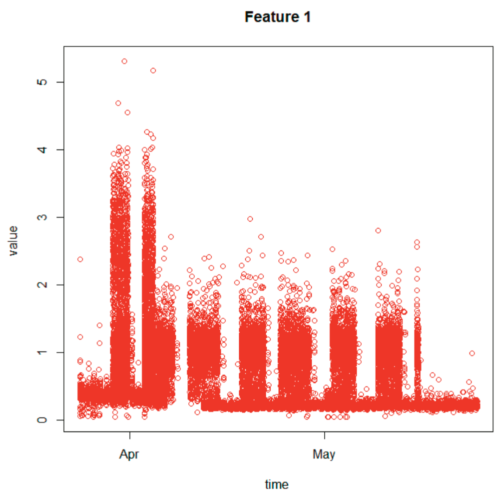


Fig. 4. Existing feature 1

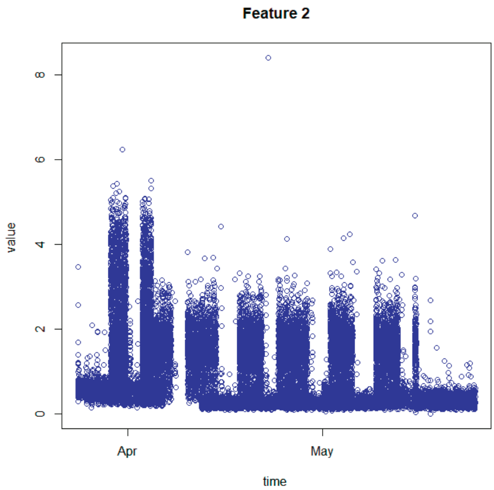


Fig. 5. Existing feature 2

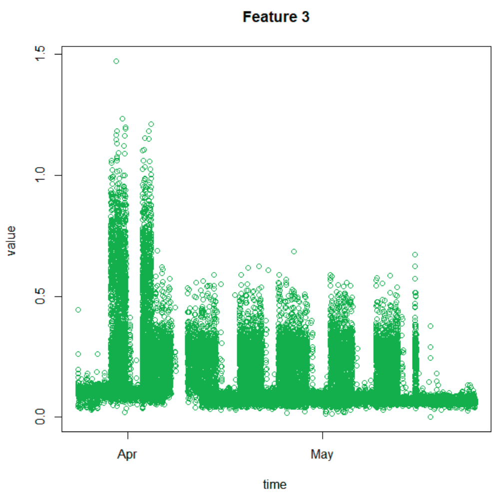


Fig. 6. Existing feature 3

A. Up Movement

We have analyzed 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 multilayer perceptron model and the generated deep features are shown in Fig. 7. In Fig. 7, we can see from visualization 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 learning terminology, which shows hidden representations of the data.

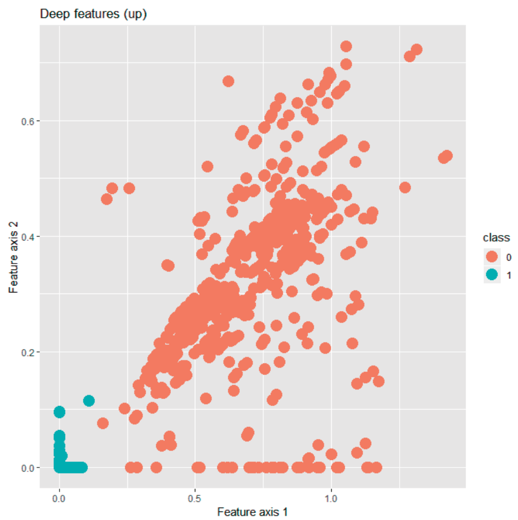


Fig. 7. Extracted multilayer perceptron features (visualization of the features w.r.t class variable)

The extracted deep features are used for classification and the results provide nearly 100% accuracy in fault detection, as shown in Table II. 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 remaining healthy data to analyze the number of false positives. The remaining healthy data are labelled as class 0 and classified with the pre-trained multilayer perceptron model to test the efficacy of the model in terms of false positives. Table II presents the results for upward movement of the elevator in terms of accuracy of fault detection. We have also included the accuracy of avoiding false positives as evaluation parameters for this research. The results show that the new deep features provide better accuracy in terms of fault detection and avoiding false positives from the data, which is helpful in detecting false alarms for elevator predictive maintenance strategies. It is extremely helpful in reducing the unnecessary visits of maintenance personnel to installation sites.

TABLE II
FAULT DETECTION ANALYSIS (FALSE POSITIVES FIELD RELATED TO ANALYZING THE REMAINING HEALTHY DATA AFTER THE TRAINING AND TESTING PHASE)

	MLP (Deep features)	RF (Existing features)
Accuracy	0.99	0.65
False positives	1	0.61

B. Down Movement

For downward motion, just as in the case of up movement, we feed both healthy and faulty data with class labels to the multilayer perceptron model for the extraction of new deep features, as shown in Fig. 8.

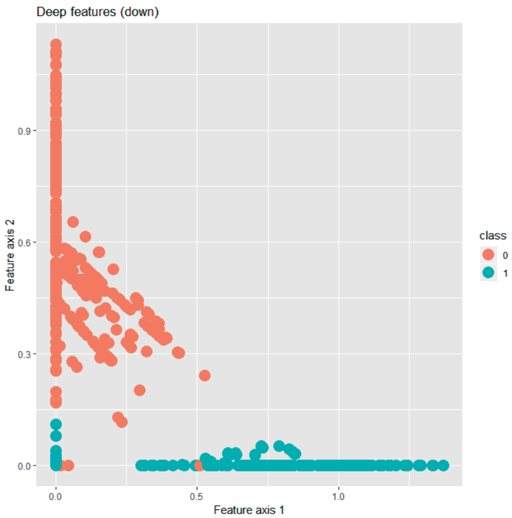


Fig. 8. Extracted deep features

Finally, the new extracted deep features are classified for fault detection, and the results are shown in Table III. After this, the remaining healthy data with class label 0 is used to analyze the number of false positives. Table III presents the results for fault detection with multilayer perceptron model in the downward direction. The results are similar to the upward direction, but we can see significant change in terms of accuracy of fault detection and when analyzing the number of false positives with new deep features.

TABLE III
FAULT DETECTION ANALYSIS

	MLP (Deep features)	RF (Existing features)
Accuracy	0.99	0.62
False positives	1	0.58

IV. CONCLUSIONS AND FUTURE WORK

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 the fault detection and also in the case of analyzing false positives for new extracted deep features. The results support our goal in this research of developing a generic model which can be used to other machine systems for automated feature extraction and fault detection. We have used almost one year of data from seven traction elevators in this research, which proves the generalization 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. Our model outperforms others because of new deep features extracted from the dataset as compared to 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.

In future work, we will extend our approach on other real-world big data cases to validate its potential for other applications and improve its efficacy.

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, 8595, 2008. doi:10.1007/s12206-007-1010-0.

[3] T. Ebeling, "Condition monitoring for elevators - An overview," Lift-Report, vol. 6, pp. 25, 2011.

[4] T. Ebeling and M. Haul, "Results of a field trial aiming at demonstrating the permanent detection of elevator wear using intelligent sensors," in Proc. ELEVCON, Madrid, Spain, pp. 101109, May 2016.

[5] P. Wang, W. He and W. Yan, "Fault diagnosis of elevator braking system based on wavelet packet algorithm and fuzzy neural network," 9th International Conference on Electronic Measurement Instruments, ICEMI '09, 10281031, 2009.

[6] 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.

[7] F. Jia, Y. Lei, J. Lin, X. Zhou, and N. Lu, "Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data," Mechanical Systems and Signal Processing, vol. 72, pp. 303-315, 2016.

[8] J. Bulla, A. D. Orjuela-Can and O. D. Flrez, "Feature extraction analysis using filter banks for faults classification in induction motors," in 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1-6, 2018.

[9] 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.

[10] L. Jing, T. Wang, M. Zhao and P. Wang, "An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox," Sensors, vol. 17, (2), pp. 414, 2017.

[11] R. Z. Haddad, C. A. Lopez, J. Pons-Llinares, J. Antonino-Daviu, and E. G. Strangas, "Outer race bearing fault detection in induction machines using stator current signals," in Industrial Informatics (INDIN), 2015 IEEE 13th International Conference on, 2015, pp. 801-808.

[12] L. Ciabattoni, G. Cimini, F. Ferracuti, A. Freddi, G. Ippoliti, and A. Monteriu, ”A novel LDA-based approach for motor bearing fault detection,” in Industrial Informatics (INDIN), 2015 IEEE 13th International Conference on, 2015, pp. 771-776.
[13] R. Rahimilarki and Z. Gao, ”Grey-box model identification and fault detection of wind turbines using artificial neural networks,” in 2018 IEEE 16th International Conference on Industrial Informatics (INDIN), 2018, pp. 647-652.
[14] G. Rigatos, N. Zervos, D. Serpanos, V. Siadimas, P. Siano, and M. Abbaszadeh, ”Condition monitoring of wind-power units using the derivative-free nonlinear kalman filter,” in 2018 IEEE 16th International Conference on Industrial Informatics (INDIN), 2018, pp. 472-477.
[15] 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.
[16] I. Fernndez-Varela, D. Athanasakis, S. Parsons, E. Hernndez-Pereira, and V. Moret-Bonillo, ”Sleep staging with deep learning: A convolutional model,” in Proceedings of the ESANN 2018.
[17] D. Zurita, E. Sala, J.A. Carino, M. Delgado, and J.A. Ortega, ”Industrial process monitoring by means of recurrent neural networks and self organizing maps,” in 2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA), 2016, pp. 1-8.
[18] R. Silva and R. Reis, ”Adaptive self-organizing map applied to lathe tool condition monitoring,” in 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), 2017, pp. 1-6.
[19] G. S. Chadha and A. Schwung, ”Comparison of deep neural network architectures for fault detection in tennessee eastman process,” in 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), 2017, pp. 1-8.
[20] W. Ren, G. Wen, S. Liu, Z. Yang, B. Xu, and Z. Zhang, ”Seam penetration recognition for GTAW using convolutional neural network based on time-frequency image of arc sound,” in 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), 2018, pp. 853-860.
[21] S. Potluri, C. Diedrich and G. K. R. Sangala, ”Identifying false data injection attacks in industrial control systems using artificial neural networks,” in 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), 2017, pp. 1-8.
[22] J. L. McClelland, D. E. Rumelhart and PDP Research Group, Parallel Distributed Processing. MIT press Cambridge, MA:, 2, 1987.
[23] A. Singh and A. Parey, ”Gearbox fault diagnosis under non-stationary conditions with independent angular re-sampling technique applied to vibration and sound emission signals,” Appl. Acoust., vol. 144, pp. 11-22, 2019.
[24] M.A.S. ALTobi, G. Bevan, P. Wallace, D. Harrison, and K.P. Ramachandran, ”Fault diagnosis of a centrifugal pump using MLP-GABP and SVM with CWT,” Engineering Science and Technology, an International Journal, 2019.
[25] C. Fang, X. Zhang, Y. Cheng, S. Wang, L. Zhang, and Y. Yang, ”Fault diagnosis for brake system in high-speed trains using the phased features and multi-layer perceptron,” in IOP Conference Series: Materials Science and Engineering, 2019, pp. 012007.
[26] R. S. Gunerkar, A. K. Jalan and S. U. Belgamwar, ”Fault diagnosis of rolling element bearing based on artificial neural network,” Journal of Mechanical Science and Technology, vol. 33, (2), pp. 505-511, 2019.
[27] K. M. Mishra, T. Krogerus and K. Huhtala, ”Deep autoencoder feature extraction for fault detection of elevator systems,” in Proceedings of the 27th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN), pp. 191-196, 2019.
[28] X. Glorot and Y. Bengio, ”Understanding the difficulty of training deep feedforward neural networks,” in Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, 2010, pp. 249-256.
[29] R. P. Monteiro, M. Cerrada, D. R. Cabrera, R. V. Snchez, and C. J. Bastos-Filho, ”Using a Support Vector Machine Based Decision Stage to Improve the Fault Diagnosis on Gearboxes,” Computational Intelligence and Neuroscience, vol. 2019, 2019.
[30] L. Breiman, ”Random forests”, Machine Learning, vol. 45, no. 1, pages 5-32. 2001.
[31] A. Liaw and M. Wiener, ”Classification and regression by random forest,” R News, vol. 2, (3), pp. 18-22, 2002.
[32] T. Huynh, Y. Gao, J. Kang, L. Wang, P. Zhang, J. Lian and D. Shen, Senior Member, IEEE, and Alzheimers 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.
[33] Z. Liu, Member, IEEE, B. Tang, X. He, Q. Qiu, and F. Liu, ”Class-specific random forest with cross-correlation constraints for spectralspatial hyperspectral image classification,” IEEE Geoscience and Remote Sensing Letters, vol. 14, (2), pp. 257-261, 2017.