

# Deep Learning Approach to Multiple Features Sequence Analysis in Predictive Maintenance

Jin Yuan, Kesheng Wang and Yi Wang

**Abstract** The representation learning of life cycle dataset has its particularity in the correlation of different features and the dependency of adjacent sampling time. This paper addresses the difficulty of segmentation to high-dimensional nonlinear life cycle long CBM data, and propose a new deep learning approach based on unsupervised representation learning named Autoencoder for rolling bearing diagnosis. Two kinds of Autoencoder with encoder and decoder model are developed respectively using fully connected and convolutional hidden layers to automatically extract the dataset's representative features. Compared to the fully connected one, the convolutional Autoencoder shows clearer in a lower dimensional feature space by preserving the local neighborhood structure, and more effective to discover subjectively the intrinsic structure of nonlinear high-dimensional data of deterioration process.

**Keywords** Predictive maintenance · Multiple features sequence  
Bearing life cycle · Deep learning · Autoencoder

---

J. Yuan  
College of Mechanical and Electronic Engineering,  
Shandong Agricultural University, Tai'an, China  
e-mail: jyuan@sdau.edu.cn; jin.yuan@ntnu.no

J. Yuan · K. Wang (✉)  
Department of Mechanical and Industrial Engineering, NTNU, Trondheim, Norway  
e-mail: kesheng.wang@ntnu.no

Y. Wang  
The School of Business, Plymouth University, Plymouth, UK  
e-mail: yi.wang@plymouth.ac.uk

## 1 Introduction

With the rapid development of automation, informationization and intelligent technology, the modern production equipment structure is becoming more and more complicated. The maintenance theory and method of the production system cannot be applied to the modern multi-device complex manufacturing system, mainly in the variable structures and working principles led to different evolution patterns of equipment health, which show the diversity of non-bathtub curve underwent strong randomness. Practice shows that frequent regular maintenance and static long-term planning cannot improve the reliability of equipment, thus convectional maintenance planning due to incapable dynamic prediction may easily fall into lack or excess of maintenance. That is, many expensive components are replaced prematurely or deteriorative parts still in service, even lead to accidents, which results in the lack of optimization of system maintenance.

The health prediction of the equipment refers to the use of accurate and effective model to carry out the state detection, symptom diagnosis and failure prediction of the equipment health recession process, and predict the future health deterioration trend by data mining and analysis of the current running state of the equipment. Predictive maintenance decision-making key technologies include the following aspects: data acquisition and monitoring technology, data preprocessing and features extraction method, health statuses classification and deterioration trend prediction algorithm, equipment maintenance planning model, system scheduling optimization strategy [1].

The life cycle of the equipment is different, the performance and status of the equipment is also different, the periodic maintenance measures should be different. At the same time, the degree of damage to the fault is different, the failure to take the maintenance measures should also be different. Therefore, it is necessary to introduce the concept of life cycle to qualitatively distinguish and describe the age and health status of the equipment. It is necessary to establish an evaluation model to evaluate the life cycle of the equipment.

On the other hand, Bearings are the most widely used mechanical parts. Due to natural degradation process and various reasons during operation, even in normal circumstances, the wear and tear of bearings that fatigues lead to premature damage. Because the tests and validation of fault diagnostic and prognostic methods are easy to introduce and to perform on the bearing test rig, furthermore, the long and slow degradation of bearing could to be acquired with monitoring signal sequences. Therefore, to study Condition-Based Maintenance (CBM) based bearings is an important research direction in predictive maintenance and health management field. Thus, we using this kind of dataset to test our algorithm in this paper.

Vibration signals are often contaminated by noise and thus unusable to distinguish direct from different machine statuses and fault types. Feature extraction can increase the signal to noise ratio to improve the correct identification by find some intrinsic ingredients. Several vibration of rotating machinery fault diagnosis techniques have been applied to. Some conventional time-domain, frequent-domain

and time-frequent domain's feature extraction by various signal processing algorithms to represent the manifold structure in vibration data space [2, 3].

The essence of deep learning [4] is to automatically learn more meaningful features of high dimensional complex data by constructing machine learning model and layer-by-layer feature transformation with multiple hidden layers, and discover the equipment inherent deterioration patterns, and ultimately improve the prediction accuracy. The deep Autoencoder network [5] is a mainstream representation learning model and high dimensional visualization tool, by studying the deep nonlinear network structure of the input, reveals the distributed characteristics of the input data, in addition to its ability to learn essential features from a number of samples, these excellent performance is far from the traditional shallow model is far from comparable. It can better model the rolling bearing vibration data and optimize the structure of continuous deep Autoencoder network model to further improve the prediction accuracy and effectively complete the fault prediction of rolling bearing.

This paper addresses the difficulty of handling high-dimensional nonlinear life cycle long CBM data, we propose a new deep learning approach based on unsupervised representation learning named Autoencoder for rolling bearing faults diagnosis. The rest of this paper is organized as follows: Section 2 illustrates the methodology of multiple features sequence analysis in predictive maintenance for the segmentation problem. Two kind of Autoencoder with encoder and decoder model is addressed respectively using fully connected and convolutional hidden layers to handle automatically extract the dataset's representative features. The experiments and discussion of the proposed solution is evaluated in Sect. 3. Finally, the conclusions are given in Sect. 4.

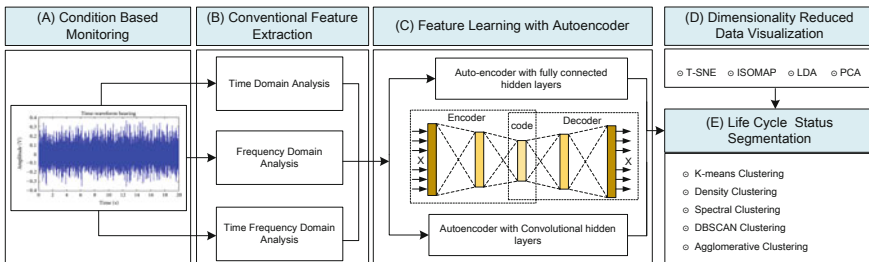
## 2 Methodology of Multiple Features Sequence Analysis in Predictive Maintenance

The evaluation of the life cycle of the equipment is the baseline of this paper approach, which is a basis for further determining the hazard level of a failure mode in the monitoring system of equipment, also as a foundation for determining maintenance implement. According to the fault occurrence, the life cycle of equipment conventionally is divided into three stages: run-in period (early failure period), maturity stage (occasional fault period) and recession period (loss failure period). Different maintenance strategies are used for these three stages in the traditional preventive maintenance. However, with the equipment failure rate and performance changes more complex and diverse, to address more reasonable and comprehensive qualitative description of equipment status and performance, it is very necessary and useful to make more detailed stage division in equipment predictive maintenance.

## 2.1 Architecture of Multiple Features Sequence Analysis and Segmentation

Data collected from the progressive deteriorated equipment often gather together, more complicated is that those data has time dependent, thus it is far difficult to segment naturally into clusters, where the characteristics of equipment status in the same cluster are similar and the characteristics of equipment status in different clusters are dissimilar. Therefore, the aim of this paper is to enhance the representative visualization of monitored multiple time series sequences leveraged by deep learning, and to give out a more reasonable segmentation.

Figure 1 shows the system overview of Autoencoder method for multiple features sequence analysis and segmentation in predictive maintenance. Our model consists of five main parts: condition based monitoring based data collection, conventional feature extraction, representation learning with Autoencoder, dimensionality reduced data visualization and life cycle status segmentation. Given some multiple sequences of multi-channels sensors or one main sensor signal (A), and then those multi-features could be extracted by time domain analysis, frequency domain analysis and time-frequency domain (B). The slices of multi-correlative time series could be generated, and the slices from different features are stacked together. Then, they are be feed to two kind of Autoencoder with encoder and decoder net components to train by representation learning process and multiple features reconstruct process (C). A dimensionality reduced latent feature representation are effectively predicted by the encoders, and then leveraged by T-SNE or PCA etc. visualization tools to show the intrinsic manifold structure mapping to 2D plane (D). From the distribution and the deterioration of points represented to every sampling statuses, using proper clustering algorithm, combined with subjective decided cluster number, to determine the corresponding cluster category representing current status in life cycle (E).



**Fig. 1** System overview of Autoencoder method for multiple features sequence analysis and segmentation in predictive maintenance

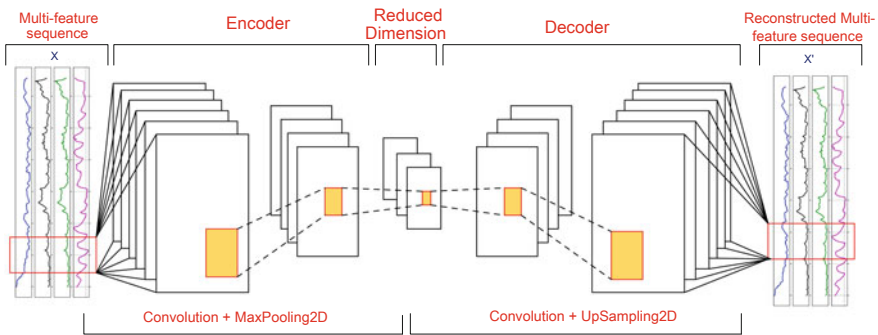
### 2.2 Two Autoencoder Network Structures

The purpose of dimensionality reduction is to obtain a more compact representation of the original high-dimensional data, a representation that nonetheless captures all the information necessary for higher-level decision-making. We propose two types deep encoder-decoder network for dimensionality reduce of multiple features sequence. Figure 2 shows the novel cascade convolution hidden layers and MaxPooling2D hidden layers formed the Encoder network structure, and cascade convolution hidden layers and UpSampling2D hidden layers formed the Decoder network structure, which together formed Autoencoder to be fed with multi-feature sequences  $X$  and output the reconstructed  $X'$ , and the Encoder could output the encoded and reduced low dimensional manifold, which is utilized to effectively discover the structure of this manifold by having the network to reconstruct the input despite some constraint and the Encoder model the spatial and sequential correlations of sequence slices with step size.

For the performance comparative, we also utilized the fully connected hidden layers to handle the same multi-feature sequences  $X$ , but without using slice of sequences. Both two types deep encoder network of Autoencoder could automatically extract the dataset's representative features.

### 2.3 Visualization and Segmentation

The deep Autoencoder successfully re-extracted the reduced representations by identified the patterns of vibration in every life cycle statuses. Visualization of high-dimensional data is a challenging problem of fault diagnosis and prognosis to deal with the high-dimensional and nonlinear data collected from the complete



**Fig. 2** The structure of Convolutional Autoencoder with encoder subnetwork and decoder subnetwork to learn the process of multi-sequence compression and reconstruction

information of operating machinery. The t-SNE algorithm [6] outperforms existing state-of-the-art techniques for visualizing a variety of real-world datasets.

Both Autoencoder and t-SNE try to find a lower dimensionality embedding of multi-features sequence. However, the Autoencoder tries to minimize the reconstruction error, while t-SNE tries to find a lower dimensional space and at the same time it tries to preserve the neighborhood distances. t-SNE is usually preferred for visualizations.

There is no uniform evaluation Index for clustering performance, unless there are labels of cluster category, because the objective functions of different clustering algorithms are very different. The segmentation of life cycle dataset has its particularity in the correlation of different features and the dependency of adjacent sampling time, therefore the clustering algorithm should be propagation based.

### 3 Experiments and Discussion

We conduct experiments on one bearing life cycle dataset to show the utilization of this approach. The convolution layers Autoencoders are compared with the traditional fully connected layers methods that stack the correlation of different features or different signal from channels. We evaluate our multiple features sequence learning scheme on typical bearing datasets to verify our method for modeling the temporal dependency.

#### 3.1 *Bearing Life Cycle Dataset*

The bearing life cycle dataset is captured from an experimentation platform named PRONOSTIA [7] shown in Fig. 3. The database provides real experimental data that characterize the degradation of ball bearings along the whole operational life (until their total failure). One of the dataset life cycle had the operating conditions: 1800 rpm and 4000 N, and the dataset had sampled every 10 s with 1/10 s sampling times with sampling frequency 25.6 kHz. That is, the life cycle dataset used in this experiments has 2803 bins data (whole raw signal shown in Fig. 4).

#### 3.2 *Feature Extraction and Representation Learning with Autoencoder*

In this Experiment, 12 dimensional classical features are extracted in time domain: Mean, Maximum Amplitude, Root Mean Square, Peak-to-Peak, Pulse Index, Crest

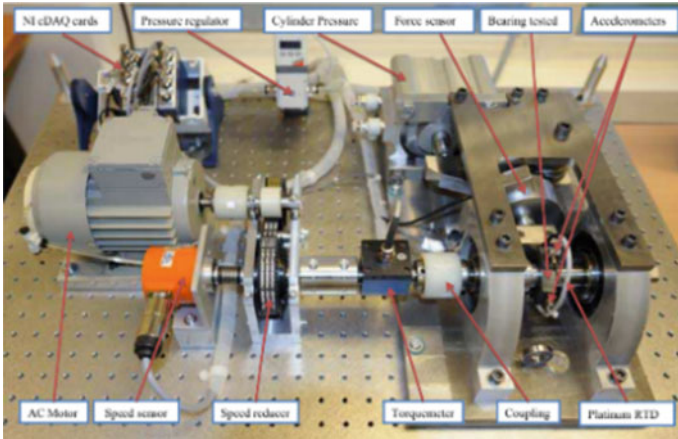
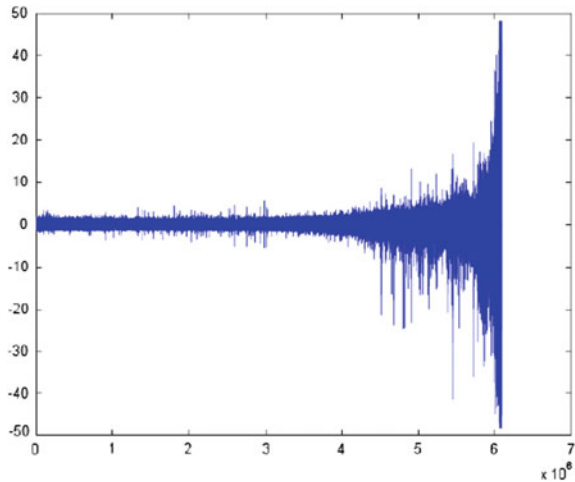


Fig. 3 Overview of PRONOSTIA

Fig. 4 A vibration raw signal



Factor, Waveform Factor, Standard Variance, Skewness, Kurtosis, Entropy, and Energy. Partly features is shown in Fig. 5.

### 3.3 Visualization and Segmentation

Figure 6a shows the deterioration process utilized by the color mapping its sampling points in bearing life cycle, which the representation of multi-feature sequences is learned by the fully connected Autoencoder and mapped it to 2D space by t-SNE. Figure 6b shows the same deterioration process, but the representation of

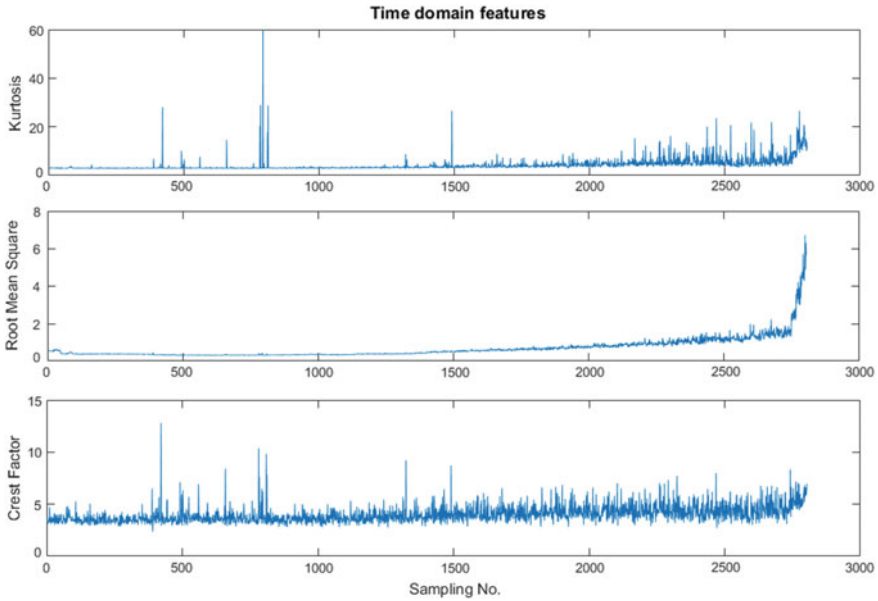


Fig. 5 Part of the features sequence of bearing life cycle extracted from time domain

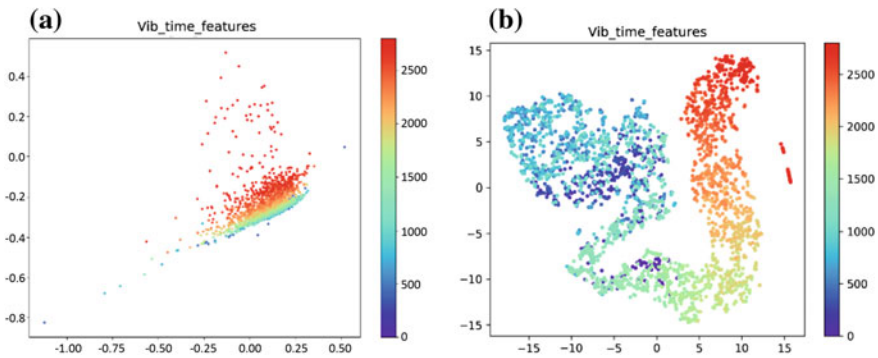
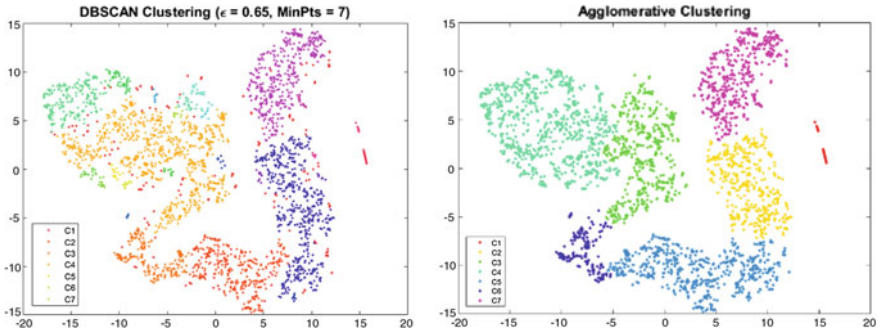


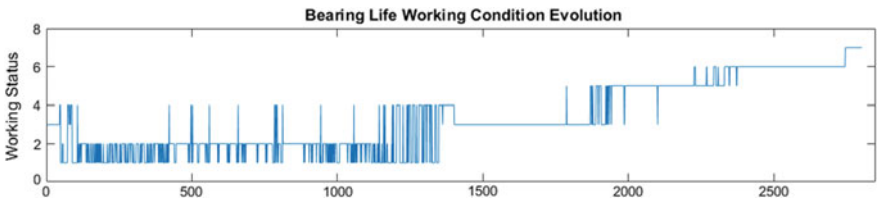
Fig. 6 Visualization of deterioration process of bearing life cycle dataset in 2D space. a Fully connected Autoencoder + t-SNE. b Convolutional Autoencoder + t-SNE

multi-feature sequences is learned by the convolutional Autoencoder and mapped it to 2D space by t-SNE with same parameters. Compared to the fully connected one, the convolutional Autoencoder shows clearer in a lower dimensional feature space by preserving the local neighborhood structure, and more effective to discover subjectively the intrinsic structure of nonlinear high-dimensional data of deterioration process.





**Fig. 7** Segmentation of deterioration process of bearing life cycle dataset with 7 clusters respectively by DBSCAN and agglomerative clustering



**Fig. 8** Status changing in the deterioration process of bearing life cycle dataset

Based on the result of visualization, we decide the life cycle dataset should be segmented to 7 clusters. We compared Hierarchical-based named agglomerative clustering, Density-based DBSCAN clustering and Distance-based K-means clustering to investigate grouping the life cycle bearing dataset [8]. The latter two clustering algorithms have big problem of partitioning observations according to locations and distances from each other, that cannot distinguish the variable of different stage in the deterioration process, for example, C7 the next to failure status at the end of sequence has some points in the early status belonged to same cluster in the DBSCAN clustering in Fig. 7, while this phenomenon is not shown in the right plot. The Agglomerative clustering has clearer centroid clusters of representative characterization. Finally, mapping the clustering category to the status changing (Fig. 8) in the deterioration process of bearing life cycle dataset, which is more interpretable for the decision making and more useful in predictive maintenance.

## 4 Conclusions

This paper addresses the difficulty of handling high-dimensional nonlinear life cycle CBM data, we propose a new deep learning approach based on unsupervised representation learning named Autoencoder for bearing fault diagnosis. Two kinds

of Autoencoder with encoder and decoder model are developed respectively using fully connected and convolutional hidden layers to handle automatically extract the dataset's representative features. Moreover, the achieved features also are visualized with state-of-the-art dimensionality reduction and clustering methods. The experiments show that convolutional Autoencoder has more feasibility and effectiveness at preserving the local neighborhood structure, and discovering the intrinsic structure of nonlinear high-dimensional data for the representation learning of multi-feature sequences.

**Acknowledgements** The work is part of project MonitorX. Which is supported by Norwegian Research Council (NFR).

## References

1. Wang KS (2014) Key techniques in intelligent predictive maintenance (IPdM)—a framework of intelligent faults diagnosis and prognosis system (IFDaPS). *Adv Mater Res* 1039:490–505
2. Mobley RK (2003) *An introduction to predictive maintenance*, 2nd edn. Butterworth Heinemann, Boston, USA
3. Grall A, Dieulle L, Berenguer C, Roussignol M (2002) Continuous-time predictive-maintenance scheduling for a deteriorating system. *IEEE Trans Reliab* 51(2):141–150
4. Goodfellow IJ, Le QV, Saxe AM et al (2009) Measuring invariances in deep networks *Adv Neural Inf Proc Syst* 22:646–654
5. Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. *Science* 313:504
6. Maaten LVD, Hinton G (2008) Visualizing data using t-SNE. *J Mach Learn Res* 9(Nov): 2579–2605
7. van der Maaten L, Hinton GE (2008) Visualizing high-dimensional data using t-SNE. *J Mach Learn Res* 9:2579–2605
8. Patrick N, Rafael G, Kamal M et al (2012) PRONOSTIA: an experimental platform for bearings accelerated life test. In: *IEEE international conference on prognostics and health management*