Mechanical Systems and Signal Processing 117 (2019) 293–318

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

A novel deep output kernel learning method for bearing fault structural diagnosis

Wentao Mao [a,*], Wushi Feng [a], Xihui Liang [b]

[a] School of Computer and Information Engineering, Henan Normal University, Xinxiang 453007, China
[b] Department of Mechanical Engineering, University of Manitoba, Winnipeg R3T 5V6, Canada

ARTICLE INFO

Article history:
Received 17 March 2018
Received in revised form 4 July 2018
Accepted 21 July 2018

Keywords:
Bearing fault diagnosis
Autoencoder
Deep learning
Output kernel
Extreme learning machine

ABSTRACT

In recent years, machine learning techniques have been proved a promising tool for bearing fault diagnosis. However, in the traditional machine learning-based diagnosis methods, the fault features tend to be relatively simple and couldn't work well for different fault type once a specific feature extraction method is determined. Meanwhile, although deep learning techniques can adaptively extract more representative features from bearing fault data, they are generally computationally expensive with slow convergence speed. Even if some deep learning algorithms like Multi-Layer Extreme Learning Machine (ML-ELM) can get fast training speed by means of non-tuned training strategy, they are inevitably of randomness to some extents. To solve this problem, a new deep learning method called *deep output kernel learning* is proposed in this paper to conduct collaborative diagnosis of multiple bearing fault types. The initial motivation is using the structural domain information among multiple bearing fault types to improve the diagnosis model's generalization ability and robustness. By adopting ML-ELM as baseline algorithm, this paper firstly utilizes autoencoder to adaptively extract deep features, and then uses them to construct an objective function with output kernel regularizer. Finally, after solving this optimization problem, an output kernel matrix is obtained, and with this matrix, the final diagnosis model is built by fusing the multiple outputs of fault classifier. Experimental results on CWRU and IMS bearing data sets show that, compared to one state-of-the-art signal analysis method and eight typical machine learning-based diagnosis methods including four shallow learning algorithms and four deep learning algorithms, the proposed method can effectively improve the accuracy of bearing fault diagnosis in an acceptable time. Moreover, the results from the Kruskal-Wallis Test also indicate the proposed method has good numerical stability.

1. Introduction

Rolling bearing is a common structural type of rotating machinery. Under complex working conditions like high load and strong impact, the rolling bearings are vulnerable to various faults, which will lead to deterioration of the whole machinery. Therefore, it's of great importance to conduct the efficient, robust and fast bearing fault diagnosis and prognosis for the rolling bearings. In recent years, along with the quick development of machine learning theory and techniques, the

* Corresponding author.
E-mail address: maowt@htu.edu.cn (W. Mao).

https://doi.org/10.1016/j.ymssp.2018.07.034
0888-3270/

data-driven intelligent fault diagnosis methods have received wide attention and been successfully used in real engineering applications. Improving the accuracy and robustness of intelligent bearing fault diagnosis is of clear academic significance and application demand.

Because vibration signal can reflect the working condition directly, most of the current fault diagnosis methods work with vibration signal. Generally speaking, machine learning-based bearing fault diagnosis includes two steps: 1) feature extraction and 2) diagnosis model construction. For 1), the widely-used fault features are the statistical characteristics of signals in time domain, frequency domain and time-frequency domain. Different types of feature have different levels of sensitivity for bearing fault. For example, signal's kurtosis [1] and bandwidth Empirical Mode Decomposition (EMD) [2] are fit for incipient fault diagnosis. Wavelet Packet Decomposition (WPD) is more suitable to make analysis for mid-high frequency fault, so it can be used to diagnose the bearing fault under high frequency impact vibration. EMD can adaptively determine the resolution ratio of signal at different frequency bands, therefore it suits to analyze nonlinear and non-stationary fault signal [2]. Moreover, GARCH series parameters can be used as temporal feature from the perspective of time series. For 2), the traditional diagnosis model mainly includes linear model, fuzzy discrimination model and similarity-based model, etc. For example, aiming at bearing fault diagnosis method under variable speed condition, Wang et al. [3] proposed a hybrid method for roller bearing's fault diagnosis based on computed order tracking and variational mode decomposition. Wang et al. [4] proposed a new analysis method using statistical locally linear embedding algorithm for bearing fault diagnosis. Lei et al. [5] introduced distance-based evaluation method and developed a clustering algorithm to analyze the bearing fault. In the past decade, machine learning techniques represented by Artificial Neural Network (ANN) and Support Vector Machine (SVM) have become a promising tool of bearing fault diagnosis. For example, due to the good generalization ability of SVM on small-scale data, Samanta et al. [6] utilized SVM and ANN to establish a diagnosis model based on bearing temporal fault feature. Caesarendra et al. [7] adopted relevance vector machine and logistic regression to assess the bearing's degradation level. Liu et al. [8] established an impact time-frequency dictionary for feature extraction, and used SVM for incipient fault diagnosis. Moreover, Barakat et al. [9] proposed a new kind of ANN, i.e., hard competitive growing neural network, to detect the small bearing fault. From the discussion above, we find that two key factors of these methods needs to be mainly considered: 1) the representative and discriminative ability of the extracted features; 2) the nonlinear mapping relationship of machine learning algorithm from feature (especially with noise) to different health conditions.

However, the methods mentioned above still have some drawbacks: 1) These methods heavily rely on the signal's understanding and domain knowledge, while the specific features are sensitive to fault type and working condition, and human intervention is usually involved; 2) Most of traditional machine learning algorithms adopt shallow model which is incapable of tackling classification or regression problems in complex situations well; 3) With the quick development of advanced measurement techniques, massive data are collected, but most of the traditional machine learning algorithms have high computational complexity to establish decision model on these data. For example, SVM's computational complexity is O(Nsv^3) where Nsv is the number of support vector [10]. Although SVM can provide sparse solution, Nsv^3 is still a very large number for large-scale data. Working on the base of ANN, deep learning, proposed by Hinton [11] in 2006, is good at extracting representative features directly from raw data. In contrast with most of traditional machine learning algorithms which usually have shallow model, e.g., the kernel mapping in SVM and the single hidden layer in ANN, deep learning utilizes deep network to exploit the implicit or complex functional relationship between raw data and final decision target, especially on large-scale data. Moreover, deep learning can also solve the problem of dimension disaster. Here deep learning is not an individual algorithm, but the combination of a series of algorithms. We notice that there already exists some works about deep learning-based rotating machinery fault diagnosis [12]. As a pioneer work, Lei et al. [13] introduced the classical deep neural network (DNN) to diagnose the bearing fault with high accuracy and robustness. Ahmed et al. [14] utilized a kind of deep learning, i.e., sparse Auto-encoder, to learn sparse and over-complete features, and realized bearing fault diagnosis from highly compressed measurement. Lei et al. [15] proposed a new locally connect deep network with normalized SAE for bearing fault diagnosis, and got satisfactory accuracy. Shao et al. [16] designed a new cost function of deep auto-encoder network based on maximum relevant entropy, and improved the feature representation ability of deep learning model for bearing fault diagnosis. Qu et al. [17] proposed a stacked multi-layer denoising autoencoder to improve the deep model's robustness and discriminative ability for fault diagnosis. Lei et al. [18] proposed a deep normalized Convolutional Neural Network (CNN) to improve the diagnosis performance on imbalanced health conditions of rolling machinery. Shao et al. [19] proposed an improved convolutional Deep Belief Network (DBN) with compressed sensing, and then got better representative ability of fault feature. However, these works mainly pay more focus on the automatic feature extraction of deep learning, lacking sufficient attention on the improvement of deep network's generalization ability. Moreover, to reach satisfactory training performance, some typical methods like CNN, DBN and Stacked Autoencoder (SAE) [20] generally need to set many network layers and iteration number, which will result in long training time in practical applications. Although fine tuning can effectively improve the performance of deep learning algorithms, the price is the increase of training time because of the utilized back-propagation strategy.

Based on the discussion above, one key issue of applying deep learning techniques into bearing fault diagnosis is developing a network architecture which can get satisfactory diagnosis performance in a relatively short time. To reach this target, in our previous work [21], we applied a kind of deep learning algorithm, i.e., Multi-Layer Extreme Learning Machine (ML-ELM) [22], for bearing fault diagnosis. Integrating ELM autoencoder in each layer, ML-ELM can learn deep features automatically. Without fine tuning process, ML-ELM utilizes feedforward strategy, and then is able to get high diagnosis accuracy