Bearing Fault Diagnosis Using Convolution Neural Network and Support Vector Regression

Wei YOU\textsuperscript{1,a}, Chang-Qing SHEN\textsuperscript{1,b}, Xiao-Jie GUO\textsuperscript{2,c} and Zhong-Kui ZHU\textsuperscript{1,d,*}

\textsuperscript{1}School of Urban Rail Transportation, Soochow University, Suzhou, China
\textsuperscript{2}School of Mechanical and Electric Engineering, Soochow University, Suzhou, China
\textsuperscript{a}youwei678@hotmail.com, \textsuperscript{b}cqshen@suda.edu.cn, \textsuperscript{c}gxjcoco@outlook.com, \textsuperscript{d}zhuzhongkui@suda.edu.cn

*Corresponding author

Keywords: Deep Learning, Convolution Neural Network, Support Vector Regression, Fault Diagnosis.

Abstract. Rolling bearing is one of the most widely used parts in machinery. The vibration signals measured from a rolling bearing can reflect the conditions of the bearing, so the vibration signal is often used in the field of bearing fault diagnosis. A number of diagnostic techniques have been studied based on the vibration signals. Deep learning theory is usually used in image recognition or speech processing, and it has attracted more and more attentions in fault diagnosis. In this paper, a method based on deep learning and support vector regression is proposed. The convolutional neural network which is seldom used for one dimension signal is used for promoting feature extraction capability. Besides, the support vector regression (SVR) is an upgrading of the support vector machine which has great generalization ability and can be improved for classification by regression theory. The proposed method is achieved through a hybrid model which combines these two techniques. The presented method is used to classify the bearing fault patterns and obtains better results than that of the convolutional neural network and support vector regression machine separately.

Introduction

Rolling bearing is an essential element of rotating machinery, and their faults are one of the most recurrent reasons for machine breakdown [1]. The efficiency of bearing fault diagnosis plays a highly significant role in avoiding catastrophe and increasing the stabilization and the reliability of machinery. Bearing fault diagnosis is a challenging problem that has been continuously studied and various bearing condition monitoring techniques have been produced [2]. Numerous results have been achieved in the scientific research field of vibration signals either through traditional signal processing algorithms or newly up-to-date algorithms. Wang et al. [3] obtained a new sparse wavelet energy feature to recognized rolling bearings faults using discrete wavelet transform and sparse representation. In reference [4], the local mean decomposition-singular value decomposition and extreme learning machine were used to identify different rolling bearings’ fault types. An intelligent fault diagnosis scheme based on wavelet packet transform, distance evaluation technique and support vector regression were described and used in the fault pattern classification of bearings [5]. Deep learning is a class of techniques that can be used to identify objects [6]. Deep neural network is one kind of deep learning and is proposed to extract meaningful representations for bearing signal. Tao et al. [7] used deep belief networks which is a generative model for fault diagnosis of the rolling bearing. The convolutional neural network (CNN) is also a kind of deep learning model and the support vector regression (SVR) machine is an efficient classification method for fault recognition[8, 9]. In this paper, CNN, which has the powerful ability to extract features from the original data is combined with the SVR to exploit their advantages, and a novel hybrid method is thus proposed and used in bearing fault classification.
The New Hybrid Method

The Procedure of the Proposed Method

The new proposed hybrid method was composed of CNN feature extraction and SVR based classification. CNN is a typical model of deep learning which has a layer–layer structure. It is composed with linear convolutions and nonlinearities. SVR is an upgrade model of the famous support vector machine which is a promising classification approach in the fault diagnosis field. CNN and SVR are both supervised learning methods. First, the raw signal is fed to CNN due to its satisfactory ability to extract useful features from raw data, thus there is no need to manually extract and select the features for further processing. Then, the nonlinear model between the extracted feature vectors and their labels (fault patterns) is trained and tested by the proposed support vector regression classifier. That is, when building up the hybrid method, we use SVR as the top layer of CNN, which was usually a softmax classifier acquiescently. Because softmax classifier is a generalization of logistic model in multi classification problem but lacks of generalization ability whereas the proposed CNN-SVR classifier can solve this problem.

Mathematics Theory

The CNN-SVR is a hierarchical architecture which is built up layer by layer with different function layers. The input $S$ is one dimension. Each $s_i \in S, i = 1, 2, ..., n$ has a target value $y_i$ which demonstrated a classification type. $\{ (s_i, y_i) \}_{i=1}^{n}$ is the input matrix of the bottom layer, where $s_i$ is an input feature vector, $y_i$ is the target value represents the type label, and $n$ is the total number of training samples. CNN is utilized to extract features. A feature map is obtained as follows:

$$h^k_{ij} = \tanh((W^k * s)_ij + b_k)$$  \hspace{1cm} (1)

where $h^k_{ij}$ is the $k$-th feature map, $W^k$ is the weights between the two layers and $b_k$ is the bias. Tanh is the non-linear activate function.

The max-pooling trick is used after the convolution layer to down sampling, and here we use the $2\times2$ region. Then the features as output of last pooling layer are input to the SVR classifier for recognition. The main target of SVR is to find an optimal hyperplane $f(x) = 0$, and $f(x)$ which is defined as:

$$f(x) = w \cdot x + b$$  \hspace{1cm} (2)

where $x$ is the extracted feature point lying on the hyperplane, $w$ is the parameter for the orientation of the hyperplane, and $b$ is a scalar threshold which represents the bias from the SVR margins.

To obtain the optimal hyperplane, the positive slack variable $\xi_i$ is introduced to solve the following optimization problem:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} \xi_i$$  \hspace{1cm} (3)

subject to $y_i(w \cdot x_i + b) \geq 1 - \xi_i, i = 1, 2, ..., n$

where $C$ is a positive constant which penalizes the errors.

At last the linear function $f(x)$ is transferred into a regressive function by applying a kernel function. The regressive function is derived as follows:
\[ f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x, x_i) + b \]  

(4)

where \( \alpha_i \) and \( \alpha_i^* \) are the Lagrange multipliers. \( K(x, x_i) \) is the radial basis function (RBF) kernel function and its definition is in Eq. 5.

\[ K(x, \cdot) = \frac{-\|x - x_i\|^2}{e^{\frac{-\|x - x_i\|^2}{\sigma^2}}} \]  

(5)

where \( \sigma \) is a positive real number.

Based on the assumption that samples from the same pattern should have a similar output from the SVR, a SVR classifier can be constructed. The tested sample can be classified to class \( m \) when it satisfies the following function:

\[ \arg \min_{m=1,2,\ldots,M} \left| m - \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) e^{-\frac{-\|x - x_i\|^2}{\sigma^2}} + b \right| \]  

(6)

where \( M \) is the total number of categories \( x \) is the fault feature vector for the testing sample.

Validation of the Proposed CNN-SVR Method

The Rolling bearing data from the Case Western Reserve University [10] were used for the performance validation. The test bearings were 6205-2RSJEM SKF deep groove ball bearings and the single-point faults were arranged by using electrical discharge with fault diameters: 0.1778 mm, 0.3556 mm, 0.5334 mm. The vibration signals were collected from four conditions including normal condition (Norm), inner race fault (IF), rolling ball fault (BF), and outer race fault (OF). The data acquisition was conducted at the rotating speed of 1772 rpm under 1hp loading and the sampling frequency was 12 kHz.

We used two data sets for validation and the description is as follows: Dataset 1 consisted of Norm type and three faults with the same defect size of 0.1778 mm. Dataset 2 consisted of Norm type and three faults in different degrees. The outer race fault signal was acquired at 6:00 o’clock direction.

For both data sets, the label values were set as 0 for Norm, 1 for IF, 2 for OF, and 3 for OF. In dataset 1, there were 50 samples for each bearing condition. In dataset 2, there were 50 Norm signals and 150 samples for each fault type. It is equally split between the training and testing samples.

Results and Discussions

Fig. 1 and Fig. 2 show the training results and testing results for normal, inner race fault, rolling ball fault, and outer race fault for dataset 1. Fig. 3 and Fig. 4 present the results of training and testing results for dataset 2.

For dataset 1, the squared correlation coefficients which react to the classification accuracy is 0.9967 for training samples and 0.939 for testing samples. The training samples were properly recognized and kept in good accordance with their practical condition. Although the outputs fluctuate violently for some testing samples, most of them are correctly classified according to Eq. 6. Hence, the proposed method can identify different bearing health conditions properly. Thus the CNN-SVR model is proven to be effective in recognizing the bearing fault types.

For dataset 2, the squared correlation coefficients are 0.9925 for training samples and 0.9763 for testing samples. Compared with the results of dataset 1, the training accuracy of dataset 2 is similar while the testing accuracy is raised by almost 4 percent. Both dataset 1 and dataset 2 can be well
analyzed and recognized. Hence, a brief conclusion can be drawn that the proposed hybrid technique can classify diverse faults even with different fault sizes in a high accuracy.

In Fig. 3, one sample from ball fault was misclassified as inner fault in the training process while all the other training samples were correctly classified. In the testing process shown in Fig. 4, one sample of inner race fault was misclassified as ball fault and three samples of outer race fault were misclassified as ball race fault. The results were summarized in Table 1. It can be concluded that the proposed method achieved a good performance.

Because the proposed hybrid method was a combination of CNN and SVR, a comparison between the CNN-SVR and these two original methods was also carried out and their results were
shown in Fig. 5. It is obviously pointed that the classification accuracy of CNN-SVR is better than those of two original methods. The histogram showed the accuracy of all three methods can achieve 97% and above in the training process and reach 94% and above in the testing process. Moreover, the performance of CNN-SVR is quite good. The training accuracy of the CNN-SVR is similar to that of the SVR method, but the testing accuracy was 4.5% higher. When compared with CNN, both the training accuracy and testing accuracy had been improved.

Figure 3. The Diagnosed Results of Training Samples for Dataset 2.

Figure 4. The Diagnosed Results of Testing Samples for Dataset 2.

Figure 5. The Comparison of Three Methods.
Conclusions

In this paper, a new deep learning and SVR based method is proposed, and its application to bearing fault pattern recognition is conducted. This proposed model is a hybrid of a convolutional neural network (CNN) and a support vector regressive classifier. There are two convolutional layers and two pooling layers connected to each convolutional layers. At the top layer, a SVR classifier is employed instead of a fully connected softmax layer.

This new method named CNN-SVR can get a good accuracy in the bearing fault diagnosis and improves the testing accuracy significantly when compared with both original models. In details, the CNN part is a deep architecture that can extract more useful features that contain representational fault information automatically and the SVR part can exploit the advantages of the generalization ability from SVR.

Acknowledgment

This research was financially supported by the National Natural Science Foundation of China (Grant No. 51505311 and 51375322), the Natural Science Foundation of Jiangsu Province (No. BK20150339) and the Innovative Research Project for Postgraduates in Colleges of Jiangsu Province (Project No. KYZZ16_0085).

References