Application of Deep Belief Networks for Precision Mechanism Quality Inspection

Jianwen Sun, Alexander Steinecker, Philipp Glocker

Extended Abstract

Title:

Application of Deep Belief Networks for Precision Mechanism Quality Inspection

Suggested Topic Area: Metrology and Quality Control

Authors: Jianwen Sun[1,2,*], Alexander Steinecker[1], Philipp Glocker[1]

Affiliations:
1. Microassembly & Robotics, Centre Suisse d'Electronique et de Microtechnique S.A., Switzerland
2. Institute of Neuroinformatics, University / ETH Zurich, Switzerland

Address:
Untere Gründlistrasse 1, 6055 Alpnach Dorf, Switzerland

Email:
{jianwen.sun,alexander.steinecker,philipp.glocker}@csem.ch

Telephone: +41 41 6727551
Fax: +41 41 6727500

Abstract. Precision mechanism is widely used for various industry applications. Quality inspection for precision mechanism is essential for manufacturers to assure the product leaving factory with expected quality. In this paper, we propose a novel automated fault detection method, named Tilear, based on a Deep Belief Network (DBN) auto-encoder. DBN is a probabilistic generative model, composed by stacked Restricted Boltzmann Machines. With its RBM-layer-wise training methods, DBN can perform fast inference and extract high level feature of the inputs. By unfolding the stacked RBMs symmetrically, a DBN auto-encoder is constructed to reconstruct the inputs as closely as possible. Based on the DBN auto-encoder, Tilear is structured in two parts: training and decision-making. During training, Tilear is trained with the signals only from good samples, which enables the trained DBN auto-encoder only know how to reconstruct signals of good samples. In the decision-making part, comparing the recorded signal from test sample and the Tilear reconstructed signal, allows to measure how well a recording from a test sample matches the DBN auto-encoder model learned from good samples. A reliable decision could be made. We perform experiments on two different precision mechanisms: precision electromotors and greasing control units. The feasibility of Tilear was demonstrated first. Additionally, performance of Tilear on the acquired electromotor dataset was compared with the state-of-the-art machine learning based fault detection technique, support vector machine (SVM). First result indicates that Tilear excels the SVM in terms of the Area Under the Curve (AUC) obtained from the Receiver Operating Characteristics (ROC) curve plot: 0.960 achieved by Tilear, while 0.941 by SVM.

Application of Deep Belief Networks for Precision Mechanism Quality Inspection

1 Introduction

Precision mechanism is widely used for various industry applications, such as precision electromotor for industrial automation systems, greasing control units for microsystems, and so on. Quality inspection for precision mechanism is essential for manufacturers to assure the product leaving factory with expected quality.

Normally, quality inspection at the manufacturer side is performed by trained experts with different methods. Traditionally, it is accomplished by the experts listening to the sound emitted by the product under different conditions. This subjective assessment is expert individual dependent, and can be influenced by several factors. That brings in the variability in the quality inspection.

Certain algorithms and techniques have been developed to overcome the disadvantages introduced by traditional subjective assessment. These techniques roughly can fall into three categories: signal analysis based methods (SAMs), dynamic model based methods (DMMs), and knowledge based methods (KMs) [1]. With SAMs, experts directly analyze the characteristics of measured signal by performing certain time-frequency transforms, like Fast Fourier Transform (FFT). It is widely used by the industry, but it is always necessary to find a best signal feature before starting the threshold comparison. For some barely seen defect types, it takes a long time to select features. As for the DMMs, an accurate dynamic model for each specific mechanism model is required before performing the quality inspection. KMs have been widely studied recently with the development of machine learning algorithms. Most of these readily available techniques are on the basis of discriminative learning models. A certain amount of fault samples are required to perform the fault type classification [2,3]. However, in practical applications, it is extremely difficult to get fault samples in abundance. What makes matters worse is that even a single type of defect typically has many different sensory manifestations.

Alternatively, we treat the fault detection problem as an anomaly detection problem, to overcome the scarcity of defective samples in the production line. The core of anomaly detection is to recognize the inputs that differ from those under normal conditions. Different anomaly detection techniques have been proposed [4], such as classification based anomaly detection techniques or statistical anomaly detection techniques. J. McBain et al [5] applied the boundary based method, namely Support Vector Data Descriptor (SVDP), with the high dimension signal features extracted by autoregressive model for motor fault detection. They tried to maximize the distance between the average distance of normal class and the average distance of the defective class. B. Zheng et al [6] proposed to use different discriminative classification methods with extracted vibration signal features for bearing anomaly detection. These proposed methods are still discriminative based, which means the construction of the anomaly detector still needs the presence of defective samples, even if not a big number. Also, they usually train with carefully selected features, the choice of which may greatly influence the anomaly detector's performance. Rather than constructing the anomaly detector based on discriminative classifiers, Deep Belief Network (DBN) is selected since it is a generative model which has strong ability to perform fast inference and to learn features unsupervisedly [7-9]. Firstly proposed by Hinton in 2006, DBN has attracted great attention from both academia and industry, and shown promising future in many tasks, such as real time speech translation and image recognition. To our knowledge, our project is the first time that deep belief networks is applied for machinery quality inspection.

The objective of this work is to develop a new automated fault detection system for precision mechanism inspection either using acoustic signals or vibration signals. Treating the fault detection as an anomaly detection problem, this system is based on a Deep Belief Network (DBN) auto-encoder. It learns the sensory signals only from good samples, and makes decisions for test samples with the trained network.

2 Methods and Results

2.1 Theory basis: Deep Belief Networks

DBN is a probabilistic generative model, which employs a hierarchical structure constructed by stacking Restricted Boltzmann Machines (RBMs) [8,10]. The RBM is a two layer neural network modeling the joint distribution of its inputs and outputs. To construct a DBN, a number of RBMs are stacked on top of each other. The hidden layers of lower level RBMs are the visible layers of the adjacent higher level RBMs. A greedy layer-wise training algorithm is applied to train the DBN, which is actually training the RBMs individually under the contrastive divergence rule [7]. Trained in this way, the DBN can perform a fast inference and extract high level representations, or features, of the input data. Thorough descriptions of DBNs' mathematical and technical details are available elsewhere [8,10].

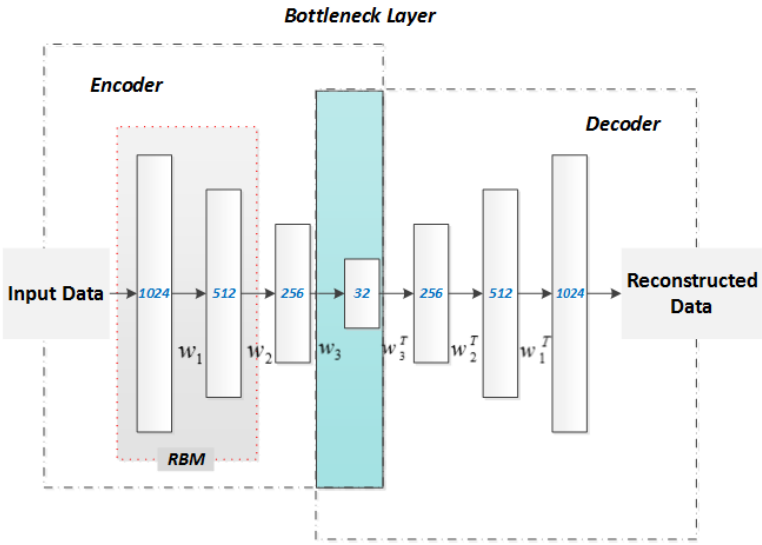

Fig. 1. Architecture of DBN based auto-encoder. Numbers in the blocks represent the number of nodes in each layer. Node number in the bottom layer represents the sampling points from the input data. Node number of the rest layers in the encoder represents the number of extracted high-order features for their respective input data. The number of nodes and layers are only examples. It is not required to have the same numbers in the experiments, or to be 2^n.

By unfolding the stacked n RBMs, an auto-encoder composed by $(2n-1)$ RBMs is constructed. This $(2n-1)$ directed auto-encoder can be fine-tuned with backpropagation [9]. As shown in Figure 1, the first n RBMs act as an encoder. High-level features of the input data are extracted by this encoder and stored at the hidden layer of the top RBM. The last n RBMs, including the top RBM of the encoder, form a decoder. This decoder reconstructs the input data with the extracted high-level features stored in the top RBM of the encoder. Generally speaking, a DBN based auto-encoder is to reconstruct training data as closely as possible.

2.2 Proposed Method: *Tilear*

Taking the advantage of DBN auto-encoder's capability to reconstruct the input data as closely as possible, we propose an anomaly detection model, named *Tilear* which learns the input data tile by tile for the purpose of performing fast inference.

Tilear has two functions: "Teacher" for the training phase, "Tester" for the decision making phase. The biggest difference between *Tilear* and other teacher/tester systems is that during training phase (teacher), only input data from good samples will be learned by the auto-encoder of *Tilear*, while other systems usually need the presence of defective samples for teacher training. In the training phase of *Tilear*, small anomalies in the "good" data are tolerable variances. The scarcity of anomalies prevents the DBN from learning and reconstructing those. This property results in an additional reconstruction error for the data containing anomalies. Therefore, the higher the reconstruction error, the more anomalies the data sample contains. An anomaly detector thus can be made by comparing the reconstruction error with a threshold. In *Tilear*, the reconstruction error S_i, also named score, is the Root Mean Squared Error (RMSE) between the input data I_i and corresponding reconstructed data R_i, averaged over n dimensions of the data, as expressed in Equation (1).

$$S_i = \sqrt{\frac{\sum_{j=1}^{n}(R_{ij} - I_{ij})^2}{n}} \tag{1}$$

The reconstruction error threshold S_{th} demarcating the anomaly boundary is another model parameter. This is determined by searching the reconstruction error space of a validation dataset containing labeled good samples and defective samples with anomalies. With the selected S_{th}, "Tester" can make a decision on the health status of test sample T_i by comparing its reconstruction score S_i to S_{th}.

2.3 Experiments and Results

Experiments on two different precision mechanisms were accomplished: precision electromotors with a 2 stage planetary gearbox, and greasing control units.

For the precision electromotor, vibration signals were acquired from 36 samples including 21 good samples and 15 defective samples with missing gears on different stages. Cepstrograms of these signals were used as the input for *Tilear*. The distribution of the reconstruction errors is shown in Figure 2.

Fig. 2. Cumulative distribution of the reconstruction errors of the precision electromotor dataset. The cumulative distribution of good samples is marked with green, while that of bad samples with red, which is flipped vertically to help examine the overlap between the two distributions. The less overlapped they are, the better performance *Tilear* has. No overlap means the detector can always make the right decision. The threshold selected by *Tilear* for decision making is shown as a red dashed line along with its actual value.

From the above figure, it is observed that most of the defective electromotors with missing gears can be detected. Although there were few samples misclassified with the self-selected threshold, it is possible to filter out all defective samples with the price of some false negative samples.

In order to evaluate *Tilear*'s performance, Area under Curve (AUC) obtained from the Receiver Operating Characteristics (ROC) curve plot is employed, due to the imbalanced class distribution in the electromotor dataset [11]. SVM as a state-of-the-art machine learning algorithm was used as a comparison benchmark. LIBSVM [12] was used to construct the SVM fault detector. Cepstrograms of the vibration signals were used as the input data for both *Tilear* and SVM. The comparison result is shown in table.1. It is observed that the AUC of *Tilear* is higher than that of SVM. This indicates that for *Tilear* had a better performance over SVM on this dataset, which has to be verified with further study. It is also worth pointing out that *Tilear* was faster for training compared to LIBSVM. The training time for *Tilear* was approximately 40 minutes, while LIBSVM always took at least several hours. In some sense, it is unfair to compare the training time here, since *Tilear* is developed to use Graphics Processing Unit (GPU) to achieve fast computation speed while LIBSVM not. The comparison of computational speed between *Tilear* and SVM on GPU platform is to be investigated in the future.

	TILEAR	**SVM**
AUC	0.960	0.941

Table 1. Comparison of AUC between *Tilear* and SVM

As for the greasing control units, acoustic signals were acquired from 47 samples consisting of 24 greased ones, which were considered as good samples, and 23 non-greased ones, which were considered as defective samples. Spectrograms were used as the input of *Tilear*. AUC was again used as the evaluation metric. AUC for these control units was 0.995. Only 1 out of 23 non-greased samples was misclassified, while all greased ones were correctly classified.

3 Conclusion

Tilear, a new automated fault detection method for precision mechanism inspection, which firstly uses the Deep Belief Networks (DBN) based auto-encoder, was proposed. *Tilear* is trained to reconstruct the data only from good samples as closely as possible. By comparing the reconstruction errors, a decision can be made. The feasibility of fault detection using *Tilear* is verified with two different kinds of precision mechanisms. It is shown that *Tilear* has comparable performance with the state-of-art technique, Support Vector Machine, using the Area under the Curve as the performance evaluation metric. It is believed that DBN not only can be used for fault detection, but also has the potential in the fault classification area, on condition that enough defective samples are collected for training.

References:

1. Liu, X., Zhang, H., Liu, J., and Yang, J.: Fault Detection and Diagnosis of Permanent-Magnet DC Motor Based on Parameter Estimation and Neural Network. In: IEEE Transactions on Industrial Electronics, vol. 47, no. 5, pp. 1021–1030 (2000).
2. Yang, J., Zhang, Y., and Zhu, Y.: Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimension. In: Mechanical Systems and Signal Processing, vol. 21, no. 5, pp. 2012–2024 (Jul. 2007)
3. Wu, S.D., Wu, P.H., Wu, C.W., Ding, J.J., and Wang, C.C.: Bearing Fault Diagnosis Based on Multiscale Permutation Entropy and Support Vector Machine. In: Entropy, vol. 14, no. 12, pp. 1343–1356 (Jul. 2012)
4. Chandola, V., Banerjee, A., and Kumar, V.: Anomaly detection. In: ACM Computing Surveys, vol. 41, no. 3, pp. 1–58 (Jul. 2009)
5. McBain, J., Timusk, M.: Feature extraction for novelty detection as applied to fault detection in machinery. In: Pattern Recognition Letters, 32(7), pp. 1054-1061 (2011)
6. Zhang, B., Georgoulas, G., Orchard, M., Saxena, A., Brown, D., Vachtsevanos, G., and Liang, S.: Rolling element bearing feature extraction and anomaly detection based on vibration monitoring. In: 16th Mediterranean Conference on Control and Automation, IEEE, pp. 1792-1797 (2008).
7. Hinton, G. E.: Training products of experts by minimizing contrastive divergence. In: Neural computation, vol. 14, no. 8, pp. 1771–800 (Aug. 2002)
8. Hinton, G., Osindero, S., and Teh, Y.: A fast learning algorithm for deep belief nets. In: Neural computation, vol. 1554, pp. 1527–1554 (2006)
9. Hinton, G. and Salakhutdinov, R.: Reducing the dimensionality of data with neural networks. In: Science, vol. 313, no. July, pp. 504–507 (2006)
10. Bengio, Y.: Learning Deep Architectures for AI. In: Foundations and Trends® in Machine Learning, vol. 2, no. 1, pp. 1–127 (2009)
11. Chawla, N. V.: Data mining for imbalanced datasets: An overview. In: Data mining and knowledge discovery handbook, Springer US, pp. 853-867 (2005)
12. Chang, C.C., and Lin, C.J.: Libsvm: A Library for Support Vector Machines. In: ACM Transactions on Intelligent Systems and Technology, vol. 2, no. 3, pp. 1–27 (Apr. 2011)