Simulation on neural networks for DUET-based delay estimation of abrasive debris signal separation

Tongyang Li, Shaoping Wang, Jian Shi, Zhonghai Ma, Enrico Zio

Simulation on neural networks for DUET-based delay estimation of abrasive debris signal separation

Tongyang Li, Shaoping Wang, Jian Shi, Zhonghai Ma
School of Automation Science and Electrical Engineering
Beihang University
Beijing, China
soredsun@buaa.edu.cn, shaopingwang@vip.sina.com,
shijian@buaa.edu.cn, zhonghaima@buaa.edu.cn

Enrico Zio
Energy Department
Politecnico di Milano
Via La Masa 34, 20156 Milano, Italy
enrico.zio@polimi.it

***Abstract*—Aliasing signals generate when two or more abrasive particles pass through an inductive debris detection sensor simultaneously, which will lead to an accumulative error for further diagnosis and prognosis of machinery equipment. The degenerate unmixing estimation technique (DUET) is an effective method for dividing aliasing signals into original sources and getting a more accurate number of the superimposed wear debris. By using the two-dimensional weighted histogram, two key parameters are estimated, which directly influences the following accuracy of source separation. To promote the precision of the parameter delay, neural networks methods including feedforward, cascade-forward, auto encoder (AE), sparse auto encoder (SAE), convolutional neural networks (CNN) are attempted and compared by using the simulative data. Different data structures are used for the testing and the result shows that the delays give the lowest mean square error (MSE) with the two-layer CNN.**

***Keywords*—aliasing signal separation, deep learning, convolutional neural networks, degenerate unmixing estimation technique**

I. INTRODUCTION

Wear and tear is one of the typical failures for mechanical components [1]. Usually, the wear status of a friction pair is difficult to be detected directly [2]. Debris detection methods, as indirect methods, provide online solutions to obtain the wear status of the mechanical components. These kind of methods have been applied on the diagnosis and prognosis of many mechanical components such as engines [3], rolling bearings [4] and many other machines [5, 6]. By using the debris detection methods, the characteristics of the direct wear productions, abrasive debris, can be obtained. The characteristics are more convincing than the vibration signals in describing the wear status and the failure process can also be presented quantitatively.

In a variety of debris detection methods, the inductive sensor is easier to be installed than those based on optics, x-ray and ultrasonic. Especially, it can distinguish between ferrous and non-ferrous debris compared with the methods based on capacitance and resistance. Many research works have been done on promoting the performance of the inductive debris detection methods, among which are a high throughput inductive pulse sensor by Du, et al [7], a radial inductive debris detection sensor by Hong, et al [8-10], a high sensitivity wear debris sensor by Zhu, et al [11] and other sensors in different structures [12]. These sensors have shown different performances on precision or sensitivity, while they share the problem of signal aliasing when two or more abrasive debris pass through an inductive debris detection sensor simultaneously and the problem is more severe when the methods are applied to high-flow-rate components like pumps.

The aliasing problem has not been paid much attentions by most researchers. On the one hand, most of the methods are aiming at the detection for lubricating oil under low flow rate which leads to a small probability of aliasing. On the other hand, some sensors are not practically applied and the performances are only proved by the simulations. In fact, when two signals are superimposed, the amplitude of the induced voltage is added and two wear debris may be recognized as one by the impulse counter method. These will cause an accumulative error and the wear status may be wrongly analyzed. Taking the aliasing into consideration, the method will have a wider application area and the precision of the method will also be promoted.

To solve the aliasing problem, a source separation method called degenerate unmixing estimation technique (DUET) [13, 14] for speech signal separation can be used [15]. The original assumption of the method is the anechoic environment which is invalid in the real acoustic environment. In other words, the DUET method is much easier to achieve for the aliasing signals separation than its original purpose: the speech signal separation. In real environment, the voice will be reflected by the wall and arrive at the sensors (normally the microphones) several times, which will cause an error. For debris sensors, one wear particle will only pass through the sensor once, which is the ideal situation of the method. In fact, there are also other methods for source separation like independent component analysis (ICA) [16], but the method cannot be applied when the sources are more than the sensors, which is commonly seen for the aliasing problem.

However, to achieve the application, the detection should be designed to meet the assumption. A serial structure has been proposed and the tested signals have been proved to agree with the assumption. By the DUET method, the separated signals are proved to be the solution of the aliasing signals, while they are not the only solution. The key parameter for the separation is obtained by a two-dimensional weighted histogram method, which cannot provide an accurate result of the parameters. The two key parameters, the attenuation and the delay, are influenced by at least three kinds of errors. The first one is caused by the

transformation: the method uses the short-time Fourier transform (STFT) to get the time-frequency domain characteristics of the aliasing signals. The two-dimensional weighted histogram clusters the key parameters by the weights, which is only to get the principal component of the transformed signals and the other components will be ignored. When conducting an inverse-transformation, there will be a transformation error. The second kind of error is caused by the clustering. If two delays or attenuations are very closed, they will be recognized as one by the current cluster method. There is another aliasing error which is caused by the noise when handling the aliasing problem. This error exists and shared by all the source separation methods. Several methods have been studied to promote the signal-to-noise ratio (SNR) so that the parameter can be obtained accurately [17].

Neural networks have been studied a lot as machine learning methods and widely used for the speech source separation [18-23]. In this paper, to obtain more accurate parameters for the source separation, several artificial neural networks are tried. Networks including the feedforward network, cascade forward network, auto encoder, sparse auto encoder and convolutional neural network are attempted. Networks with the same type but with different parameters are also tried, like the two-layer convolutional neural networks. In fact, in the aliasing signal separation, the attenuation is a constant and the most important is to get an accurate parameter delay. Both the network for the two parameters and only for the delay are tested. The results are listed in tables.

The rest of the paper is organized as follows. In section 2, the DUET-based aliasing separation is briefly described. In section 3, the networks used for delay estimation are proposed and structure are described. Section 4 shows the simulation results of different networks and the results are discussed. In Section 5, some conclusions and remarks are drawn

II. DUET-BASED ALIASING SEPARATION

The DUET method tries to separate several sources from two aliasing signals. The aliasing phenomenon can be described by Figure 1. When two or more wear debris go through an inductive sensor, the induced voltage of the debris will be superimposed. The actual wave shown to us is only the after-superimposed curve, which is the so-called aliasing signal.

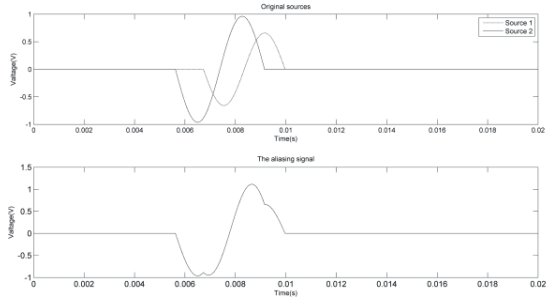

Figure 1 The aliasing phenomenon

As is shown in Figure 1, the aliasing is the superimposition of the two sources in time domain. Solving this problem is something like solving the linear equation in two unknowns like $x+y=a$, where a is a constant and x, y are unknowns. By the only one equation, we can get infinite solutions. If we want to get the only one solution, we need at least two equations and the equations cannot describe the same line in space. In other words, if the other equation is $n(x+y)=na$, where n is a constant, we cannot get the only solution. Hence, we also need two aliasing signals: $y_1(t)$ and $y_2(t)$ in DUET. The signals should meet

$$y_1(t)=\sum_{i=1}^{M} s_i(t) \tag{1}$$

$$y_2(t)=\sum_{i=1}^{M} a_i s_i(t-\delta_i) \tag{2}$$

where $s_i(t)$ is the i_{th} source and the delays δ_i should not be the same or the attenuations a_i should not be the same so that the aliasing can be separated.

Obviously, to separate the sources in time domain is very difficult. The core idea to separate the sources is to transform the signals into the time-frequency domain, which is also the key technique of the other source separation methods. This is the same like to watch the signal in another perspective, which is shown in Figure 2. The aliasing signals shows separated characteristics in the time-frequency domain. For a finite-length and discrete signal, the STFT is the proper transform function.

$$\hat{s}_i(\tau,\omega):=F^W[s_i](\tau,\omega):=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} W(t-\tau)s_i(t)e^{-j\omega t}\mathrm{d}t. \tag{3}$$

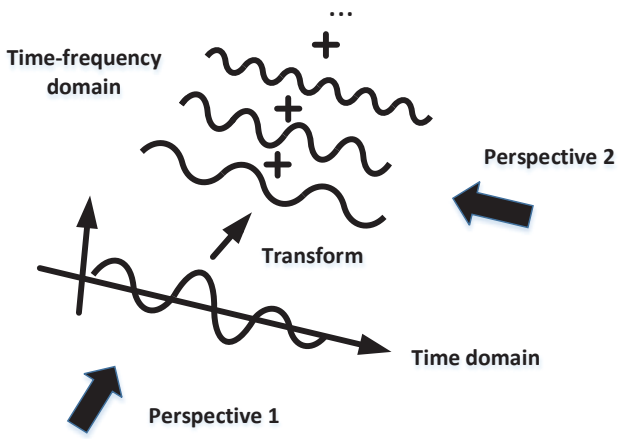

Figure 2 Transformed signals

In the STFT, if two of the aliasing sources satisfy

$$\hat{s}_i(\tau,\omega)\hat{s}_j(\tau,\omega)=0 \quad \forall\tau,\omega, \quad \forall i\neq j, \tag{4}$$

then, the two sources can be separated by the mask function

$$M_i(\tau,\omega):=\begin{cases}1 & \hat{s}_i(\tau,\omega)\neq 0\\ 0 & otherwise.\end{cases} \tag{5}$$

and the i_{th} source is

$$\hat{s}_i(\tau,\omega) = M_i(\tau,\omega)\hat{y}_1(\tau,\omega), \quad \forall \tau,\omega \tag{6}$$

The key parameters attenuation and delay can be calculated by

$$\tilde{a}(\tau,\omega) = \left| \frac{\hat{y}_2(\tau,\omega)}{\hat{y}_1(\tau,\omega)} \right| \tag{7}$$

$$\tilde{\delta}(\tau,\omega) = \left(\frac{1}{\omega}\right) \angle \left(\frac{\hat{y}_2(\tau,\omega)}{\hat{y}_1(\tau,\omega)} \right) \tag{8}$$

and the commonly used schematic diagram of a two-dimensional weighted histogram is shown in Figure 3. The peaks are the estimation of the attenuations and delays. In fact, this estimation does not always get a very accurate result because there are transformation errors, noise errors and clustering errors. To promote the accuracy, in the next section, different kinds of neural networks are compared and the results are shown in section 4.

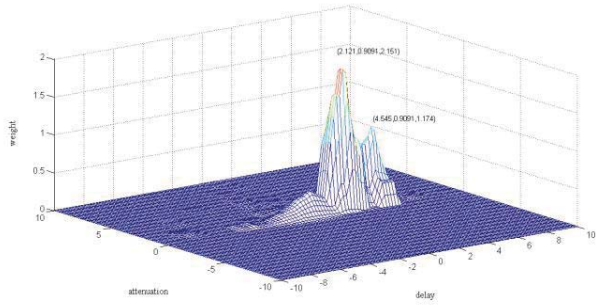

Figure 3 schematic diagram of a two-dimensional weighted histogram

III. Neural Networks Method for Delay Estimation

Neural network is a machine learning method based on large dataset. The first thing is to obtain the dataset for training. In fact, the data we got from the experiment are far away from the requirement and the experimental data cannot be separated perfectly for the training. Hence, the simulated data are used here for the training.

The simulated data are based on the sine wave. Each of the particles is described as a sine section with period from $\frac{\pi}{2}$ to $\frac{3\pi}{2}$. The sine sections with random amplitude and phase are superimposed to be the first aliasing signal. Modify the amplitude and phase of each sine sections and add them together to get the second aliasing signal. The two signals can be obtained and their delay characteristics can also be calculated during the process, which composed the dataset for training.

As is shown in Figure 4, the framework for the parameter estimation is displayed. The points of aliasing signals are the inputs and the outputs are the parameters.

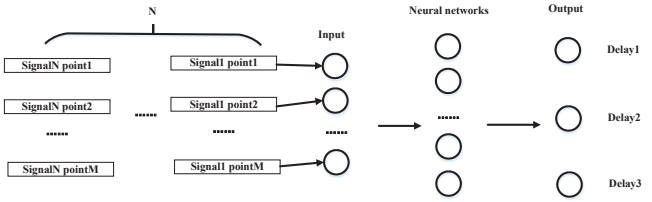

Figure 4 Framework for the parameter estimation

In fact, under the supervised learning, the number of the output should be confirmed firstly. In our framework, the output is relied on the number of sources which cannot be known. Fortunately, the DUET method provides a relative accurate number of the sources. So for each number of sources, a network is established and used only for the specified. For the debris sensors, because the attenuation is usually a constant, the estimation of the attenuation can be conducted by other method. The outputs of the proposed networks are only the delays.

The networks are actually the functions of the inputs and outputs. For the debris sensors, the delays are the characteristics of the two aliasing signals. The shape of inputs may also affect the performance of the networks. For different networks, different input shape is also tested in the next section.

IV. Simulation and Results

A. Feedforward networks

The simulation is conducted by using the Matlab neural network toolbox. One-layer and two-layer feedforward networks are tested and the structure is shown in Figure 5. Different number of hidden nodes are tested and the results are shown in TABLE I.

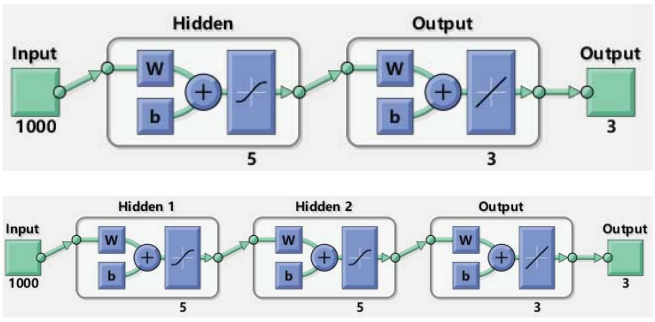

Figure 5 Feedforward networks structure

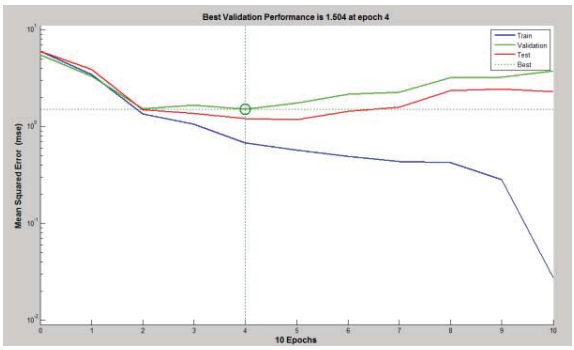

Figure 6 Mean square error

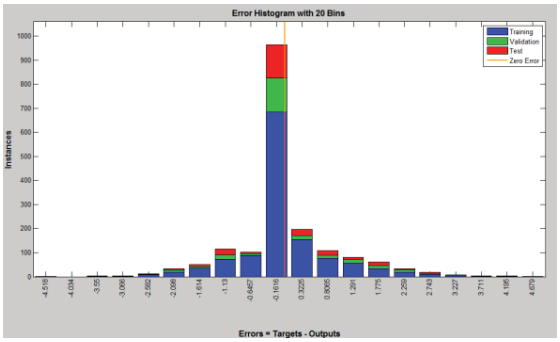

Figure 7 Error distribution

The inputs of the networks are the joint matrix of the two signals. Each signal has a length of 500 points data and the joint result is 1000 for the number of inputs. The outputs are three delays because we use three sources to generate the aliasing signals and the networks are only used to estimate the delays of three-source aliasing situation. Gradient descent is used for the training.

1000 groups of samples are used for training and 300 groups of samples are used for testing. Another 300 groups of samples with a noise of normal distribution whose mean value is zero and square error is 0.1 are also used for the test and the mean square error (MSE) are calculated for evaluation the performance of the networks. For each network structure, we can get the MSE of each training step which is shown in Figure 6 and the error distribution which is shown in Figure 7. The structure [5] means the feedforward network has one layer and the hidden nodes number is 5. The structure [5 5] means the feedforward network has two layers and the hidden nodes number of the first layer is 5, the number of the second layer is 5.

TABLE I SIMULATION RESULTS OF FEEDFORWARD NETWORKS

Network structure	MSE	MSE of data with noise
[5]	0.1819	0.2321
[8]	0.1931	0.2393
[10]	0.1816	0.2306
[12]	0.1827	0.2911
[15]	0.1564	0.2822
[5 5]	0.2016	0.2240
[8 5]	0.1782	0.2412
[10 5]	0.1925	0.2287
[5 10]	0.2038	0.2727
[8 8]	0.1929	0.2553

Because the initial weight of the networks are random, the results shown in TABLE I is the mean value of 10 times training. From the results we can easily see that the different structure of a feedforward network has little influence on the estimation. Generally, the MSE of the data with noise is higher than the MSE of the clear data.

B. *Cascade-forward networks*

The setting of the inputs and outputs are the same as the feedforward networks'. The structure of the network is shown in Figure 8. Both cascade-forward networks with one layer and two layers are tested. The results are listed in TABLE II.

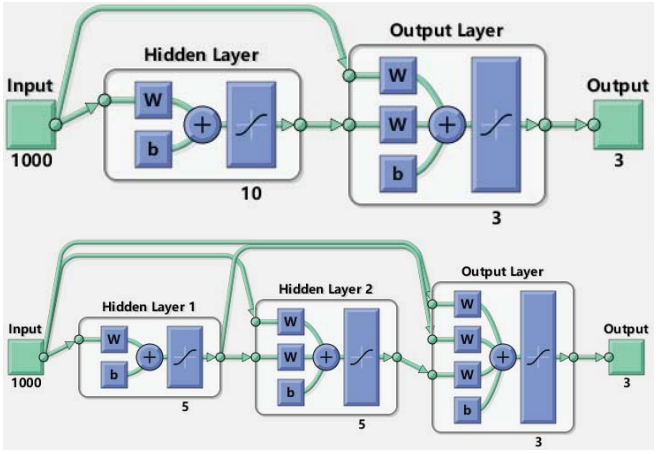

Figure 8 Cascadeforward networks structure

TABLE II SIMULATION RESULTS OF CASCADEFORWARD NETWORKS

Network structure	MSE	MSE of data with noise
[5]	0.1394	0.2803
[8]	0.2036	0.2393
[10]	0.1795	0.2101
[5 5]	0.2084	0.1976
[8 8]	0.1953	0.2480

The cascade-forward networks show a similar result with the feedforward networks. However, the time cost of the training is about 10 times than that of the feedforward networks. Different structures show an analogous MSE for both the clear data and the data with noise.

C. *Auto-encoder*

By using the traditional training method, the scale of the networks cannot be very large. The feedforward and cascade-forward networks proposed above consist of about one or two layers and totally under 20 nodes. However, the cost of the time for the training is very large. The auto-encoder changes the training method and reduces the cost largely. The scale of the network by auto-encoder can be thousands of nodes and the cost of time is approximately equals to train a 20 nodes two-layer feedforward network, which makes it an important branch in deep learning. The structure of the auto-encoder is shown in Figure 9.

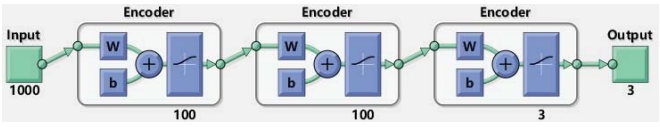

Figure 9 Auto-encoder structure

Because of the deep learning toolbox of the Matlab has stopped updated, the method is also tested by tensor-flow. The results of different node numbers are shown in TABLE III.

TABLE III SIMULATION RESULTS OF AUTO-ENCODER NETWORKS

Network structure	MSE	MSE of data with noise
[50 50]	0.1332	0.1295
[100 50]	0.1177	0.1221
[100 100]	0.1269	0.1227
[200 100]	0.1257	0.1265
[50 50 50]	0.1326	0.1220
[100 100 50]	0.1224	0.1260
[100 100 100]	0.1337	0.1365

The result of auto encoder is obviously better than the feedforward and cascade-forward networks. Less than 20 nodes may not fit the relation well. We can also see that the result is not better and better with the increasing of the number of the nodes. The best result is shown at [100 50]. The increasing of nodes does not promote the accuracy but only increases the cost of time. Deep network actually shows a better result.

D. Sparse auto-encoder

The sparse auto encoder is a modified auto encoder method with sparse constraint. This kind of method makes most nodes inactive compared with the auto-encoder. The structure of the sparse auto encoder is the same as the auto encoder shown in Figure 9.

The sparsity regularization of the tested networks is set as 4 and the sparsity proportion is set as 0.05. Purelin is chosen as the decoder transfer function. The testing result is shown in TABLE IV.

TABLE IV SIMULATION RESULTS OF SPARSE AUTO-ENCODER NETWORKS

Network structure	MSE	MSE of data with noise
[50 50]	0.0412	0.3673
[100 50]	0.0372	0.3691
[100 100]	0.0373	0.3585
[200 100]	0.0375	0.3538

Different from the auto encoder, result of the sparse auto encoder shows that it has a higher accuracy in the test of clear data but a lower accuracy in that of data with noise instead of a more accurate result.

E. Convolutional neural networks

The CNN simulation is under the tensor-flow environment. Generally, the structure of CNN is shown in Figure 10. The input signals are reshaped to a matrix of 2 rows and 500 columns. Convolutional core with different numbers and shapes are tested. Then, the max pooling layer is used. After the full connection layer, three outputs are obtained.

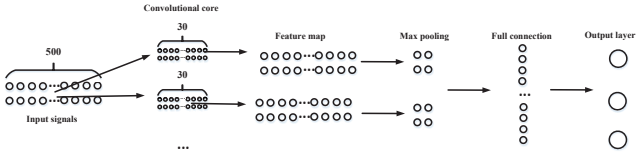

Figure 10 CNN structure

The two-layer CNN is after the first max pooling layer, there are another convolutional layer and max pooling layer. The number of nodes for the full connection layer is 256 in one-layer CNN and 1024 in two-layer CNN. Different structure of the CNNs are tested and the results are listed in TABLE V.

TABLE V SIMULATION RESULTS OF CONVOLUTIONAL NEURAL NETWORKS

Network structure	MSE	MSE of data with noise
10conv20*2 pool2*2	0.1297	0.1183
20conv20*2 pool2*2	0.1150	0.1169
30conv20*2 pool2*2	0.1251	0.1226
10conv30*2 pool2*2	0.1181	0.1136
20conv30*2 pool2*2	0.1177	0.1190
30conv30*2 pool2*2	0.1216	0.1159
10conv30*2 pool2*2 10conv15*1 pool2*1	0.0824	0.0841
20conv30*2 pool2*2 10conv15*1 pool2*1	0.0854	0.0852
30conv30*2 pool2*2 10conv15*1 pool2*1	0.0841	0.0859
10conv30*2 pool2*2 20conv10*1 pool2*1	0.0826	0.0881
20conv30*2 pool2*2 20conv10*1 pool2*1	0.0851	0.0827
30conv30*2 pool2*2 20conv10*1 pool2*1	0.0862	0.0865

The one-layer CNN shows a similar result with the auto encoder networks and the two-layer CNN shows a better result. Different parameters on the number of convolutional cores and shapes of convolutional cores are simulated. The result does not show a wide difference between networks with different parameters. The two-layer CNN shows the best performance in all the neural networks tested in this paper.

V. CONCLUSIONS

In this paper, neural networks including feedforward networks, cascade-forward networks, auto encoder, sparse auto encoder and convolutional neural networks are used for the estimation of delay for the aliasing source separation. Simulated data are used to do the estimation and the result shows that different networks show different performances. The difference of a special network with different structure like multi-layer or different number of nodes does not have an impact on the result. The two-layer CNN shows the best result that both the MSE for the testing group of clear data and the MSE for the testing group of data with noise are the lowest in all the networks used for testing. Especially, the sparse auto-encoder has shown the best result for the clear data but the worst result for the data with noise. In fact, in the real application, the data with noise are more often

to be seen that under this circumstance, the two-layer CNN shows a robust characteristic.

ACKNOWLEDGMENT

This study was co-supported by the National Natural Science Foundation of China (51575019, 51620105010), National Basic Research Program of China (2014CB046402) and Program 111 of China.

REFERENCES

[1] Anderson, D. P., Wear Particle Atlas. Revised. *FOXBORO ANALYTICAL BURLINGTON MA* **1982**.

[2] Li, T.; Wang, S.; Shi, J.; Ma, Z., An adaptive-order particle filter for remaining useful life prediction of aviation piston pumps. *Chinese Journal of Aeronautics* **2017**.

[3] Wu, T.; Peng, Y.; Wu, H.; Zhang, X.; Wang, J., Full-life dynamic identification of wear state based on on-line wear debris image features. *Mechanical Systems & Signal Processing* **2014,** 42, (1-2), 404-414.

[4] El-Thalji, I.; Jantunen, E., Dynamic modelling of wear evolution in rolling bearings. *Tribology International* **2015,** 84, (84), 90-99.

[5] Zhu, J.; Yoon, J. M.; He, D.; Bechhoefer, E., Online particle‐contaminated lubrication oil condition monitoring and remaining useful life prediction for wind turbines. *Wind Energy* **2015,** 18, (6), 1131-1149.

[6] Zhu, J.; He, D.; Bechhoefer, E., Survey of lubrication oil condition monitoring, diagnostics, prognostics techniques and systems. *Journal of Chemical Science & Technology* **2012,** 2, 100-115.

[7] Du, L.; Zhe, J., A high throughput inductive pulse sensor for online oil debris monitoring. *Tribology International* **2011,** 44, (2), 175-179.

[8] Hong, W.; Wang, S.; Tomovic, M.; Han, L.; Shi, J., Radial inductive debris detection sensor and performance analysis. *Measurement Science & Technology* **2013,** 24, (12), 5103.

[9] Hong, W.; Wang, S.; Tomovic, M. M.; Liu, H.; Wang, X., A new debris sensor based on dual excitation sources for online debris monitoring. *Measurement Science & Technology* **2015,** 26, (9).

[10] Hong, W.; Wang, S.; Liu, H.; Tomovic, M. M.; Chao, Z., A hybrid method based on Band Pass Filter and Correlation Algorithm to improve debris sensor capacity. *Mechanical Systems & Signal Processing* **2016,** 82, (14), 1-12.

[11] Zhu, X.; Zhong, C.; Zhe, J., A high sensitivity wear debris sensor using ferrite cores for online oil condition monitoring. *Measurement Science & Technology* **2017,** 28, (7), 075102.

[12] Zhan, H.; Song, Y.; Zhao, H.; Gu, J.; Yang, H.; Li, S., Study of the sensor for on-line lubricating oil debris monitoring. *Sensors & Transducers* **2014,** 175, (7), 214-219.

[13] Jourjine, A.; Rickard, S.; Yilmaz, O. In *Blind separation of disjoint orthogonal signals: demixing N sources from 2 mixtures*, Acoustics, Speech, and Signal Processing, 2000. on IEEE International Conference, 2000; 2000; pp 2985-2988.

[14] Yilmaz, O.; Rickard, S., Blind separation of speech mixtures via time-frequency masking. *Signal Processing IEEE Transactions on* **2004,** 52, (7), 1830-1847.

[15] Li, T.; Wang, S.; Zio, E.; Shi, J.; Hong, W., Aliasing Signal Separation of Superimposed Abrasive Debris Based on Degenerate Unmixing Estimation Technique. *Sensors* **2018,** 18, (3), 866.

[16] Singh, A.; Anand, R. S. In *Overview of the performance of fast ICA and DUET for blind source separation of multiple speakers*, Recent Advances in Electronics & Computer Engineering, 2016; 2016; pp 296-300.

[17] Zaidi, S. S. A.; Ali, M. M.; Aftab, F.; Shahid, Y.; Khurram, M. In *Name spotting over low signal-to-noise ratio (SNR) using Blind Source Separation and Connectionist Temporal Classification*, International Conference on Communication, Computing and Digital Systems, 2017; 2017; pp 320-325.

[18] Freisleben, B.; Hagen, C.; Borschbach, M. In *A neural network for the blind separation of non-Gaussian sources*, IEEE International Joint Conference on Neural Networks Proceedings, 1998. IEEE World Congress on Computational Intelligence, 1998; 1998; pp 837-842 vol.2.

[19] Huang, P.; Kim, M., Singing-Voice Separation From Monaural Recordings Using Deep Recurrent Neural Networks. **2014**.

[20] Huang, P. S.; Kim, M.; Hasegawa-Johnson, M.; Smaragdis, P. In *Deep learning for monaural speech separation*, IEEE International Conference on Acoustics, Speech and Signal Processing, 2014; 2014; pp 1562-1566.

[21] Weninger, F.; Eyben, F.; Schuller, B., *Single-channel speech separation with memory-enhanced recurrent neural networks*. 2014; p 3709-3713.

[22] Lee, H. Y.; Cho, J. W.; Kim, M.; Park, H. M., DNN-Based Feature Enhancement Using DOA-Constrained ICA for Robust Speech Recognition. *IEEE Signal Processing Letters* **2016,** 23, (8), 1091-1095.

[23] Nugraha, A. A.; Liutkus, A.; Vincent, E., Multichannel Audio Source Separation With Deep Neural Networks. *IEEE/ACM Transactions on Audio Speech & Language Processing* **2016,** 24, (9), 1652-1664.